Gerhard Ernst / Jakob Steinbrenner / Oliver R. Scholz (Eds.)

From Logic to Art

Themes from Nelson Goodman

P h i l o s o p h i s c h e F o r s c h u n g
P h i l o s o p h i c a l R e s e a r c h

Herausgegeben von / Edited by

Johannes Brandl • Andreas Kemmerling
Wolfgang Künne • Mark Textor

Band 7 / Volume 7

Gerhard Ernst / Jakob Steinbrenner
Oliver R. Scholz (Eds.)

From Logic to Art

Themes from Nelson Goodman

ontos
verlag

Frankfurt I Paris I Lancaster I New Brunswick

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliographie;
detailed bibliographic data is available in the Internet at http://dnb.ddb.de

North and South America by
Transaction Books
Rutgers University
Piscataway, NJ 08854-8042
trans@transactionpub.com

United Kingdom, Ireland, Iceland, Turkey, Malta, Portugal by
Gazelle Books Services Limited
White Cross Mills
Hightown
LANCASTER, LA1 4XS
sales@gazellebooks.co.uk

Livraison pour la France et la Belgique:
Librairie Philosophique J.Vrin
6, place de la Sorbonne ; F-75005 PARIS
Tel. +33 (0)1 43 54 03 47 ; Fax +33 (0)1 43 54 48 18
www.vrin.fr

P.O. Box 15 41, D-63133 Heusenstamm
www.ontosverlag.com

ISBN: 978-3-86838-019-4

2009

Printed on acid-free paper
ISO-Norm 970-6
FSC-certified (Forest Stewardship Council)

This hardcover binding meets the International Library standard

Printed in Germany
by buch bücher **dd ag**

Preface

Nelson Goodman was born on August 7th 1906 in Somerville, Massachusetts. When he died on November 25th 1998 at the age of 92 in Needham, Massachusetts, the world lost one of its great philosophers.

In August 2006 a couple of Goodman aficionados met in Munich to celebrate the Centennial. The proceedings of the ensuing international conference *Nelson Goodman: From Logic to Art – Looking Back on the Occasion of his 100th Birthday* are documented in this volume. The conference was organized by the editors and carried out in the Seminar for Philosophy, Logic and Philosophy of Science at Ludwig Maximilians University Munich. We want to thank Godehard Link for his organizational support and the University for its hospitality. Alexander Oldemeier and Philipp Richter provided invaluable help in preparing the manuscripts for publication. We are very much indebted to them. Last, but not least we want to thank the contributors for their lectures at the conference and, of course, for their papers.

The contributions in this volume attest the fact that Goodman's thinking holds many treasures waiting for being digged up. May they inspire further research.

Gerhard Ernst Jakob Steinbrenner Oliver R. Scholz

SA: *The Structure of Appearance*, third edition, Dordrecht 1977 (first edition 1951)
FFF: *Fact, Fiction, Forecast*, fourth edition, Cambridge, Mass. 1983 (first edition 1954)
LA: *Languages of Art: An Approach to a Theory of Symbols*, second edition, Indianapolis, Ind. 1976 (first edition 1968)
PP: *Problems and Projects*, Indianapolis 1972
WW: *Ways of Worldmaking*, second printing 1981, Indianapolis 1978
MM: *Of Mind and Other Matters*, Cambridge, Mass. 1984
R: N. Goodman u. C.Z. Elgin: *Reconceptions in Philosophy and Other Arts and Sciences*, Indianapolis 1988

Contents

Oliver R. Scholz

The Life and Opinions of Nelson Goodman – A Very Short Introduction[1]

Abstract
Nelson Goodman was an eminent philosopher and a passionate art collector. The first section of this introductory essay provides some biographical information with special emphasis on Goodman's education and academic career. The second part gives a survey of his major works and highlights their most distinctive contributions. The final section throws a glance at the still imperfect reception of his rich legacy.

1. Life

Nelson Goodman was one of the outstanding thinkers of the 20th century. In a memorial note in the Harvard University Gazette (1998), Hilary Putnam considers him to be "one of the two or three greatest analytic philosophers of the post-World War II period".[2]

Goodman has left his mark in many fields of philosophical investigation. Whether Epistemology, Philosophy of Science, Logic, Metaphysics, the General Theory of Symbols, Philosophy of Language or Philosophy of Art, all have been challenged and enriched by the problems he has shown up, the projects he developed from them and the solutions he has suggested. Non-philosophical disciplines, too, – from measurement theory (cf. e.g. Grunstra 1972) to linguistics (cf. e.g. Heydrich 1982, 1993), literary criticism (cf. e.g. Scholz 1984; Ihwe 1985; Thürnau 1994) and art education – build on his findings. In further fields we may yet expect his ideas to be taken up.

---

[1] This is a revised and extended version of (Scholz 2005). I am grateful to Morton G. White, Catherine Z. Elgin, Henry S. Leonard, Jr., Karlheinz Lüdeking, Daniel Cohnitz and Marcus Rossberg for personal communication and advice. For his translation of my German text I want to thank Rudolf Owen Müllan (Hughes). A much more comprehensive and thorough introduction in Goodman's philosophy is now available in (Cohnitz; Rossberg 2006).

[2] Quoted in (Cohnitz; Rossberg 2006, 11).

Goodman's contributions are characterized by their fundamental and far-reaching importance, great originality as well as a rare combination of rigor and elegance. *Reconceptions in Philosophy and Other Arts and Sciences,* the title of a book he wrote together with Catherine Z. Elgin, could stand as the heading for his entire œuvre. Goodman has analyzed, criticized and rejected a host of traditional answers to fundamental philosophical questions in order to consequently reframe the issues and suggest innovative and often also provocative new solutions.

On August 7th 1906 Nelson Goodman was born the son of Sarah Elizabeth (Woodbury) Goodman (1874-1964) and Henry L. Goodman (1874-1941) in Somerville, Massachusetts. Entering Harvard in the 1920s, he was to witness the beginning of what has been termed "a second golden age in philosophy" (Lowe 1990, 149)[3]. If a first golden age that had been shaped by William James (1842-1910), Josiah Royce (1855-1916), George Herbert Palmer (1842-1933) and Hugo Münsterberg (1863-1916) came to an end with the death of the figureheads James and Royce, the inception of the second golden age was marked, in 1924, by the coming of Alfred North Whitehead[4] (1861-1947) and tenure being granted to Clarence Irving Lewis, who had been a member of the Philosophy Department since 1920.[5] Furthermore William Ernest Hocking (1873-1966),[6] Ralph Barton Perry (1876-1957),[7] James Haughton Woods (1864-1935)[8] as well as brilliant

[3] On this period of American philosophy cf. (Kuklick 2001, Chapter 9: Pragmatism at Harvard, 1878-1913).

[4] Alfred North Whitehead (1861-1947) taught from 1924 through 1937 (official retirement) – and finally until 1941 (farewell lecture) – at Harvard University.

[5] Clarence Irving Lewis (1883-1964) taught from 1920 until 1953 at Harvard-University, since 1924 as a Full Professor. His main fields of scholarly interest were logic, epistemology and ethics. His textbooks on logic, *A Survey of Symbolic Logic* (Berkeley 1918) and (together with C. H. Langford) *Symbolic Logic* (New York 1932), which went through several editions, were widely disseminated. On his works cf. the contributions in (Schilpp 1968). Cf. also (Kuklick 2001, 214-220).

[6] William Ernest Hocking (1873-1966) taught at Harvard-University from 1914 until 1943; he was influenced by Kuno Fischer, Wilhelm Dilthey and Wilhelm Windelband. He has given a vivid image of Harvard during these years in Hocking 1961 (cf. also Kuklick 1977; Lowe 1990, chapter VII).

[7] Ralph Barton Perry (1876-1957), as a student and successor of William James, taught at Harvard-University from 1913 until 1946. Among his works is book with the

young scholars like Henry Maurice Sheffer (1883-1964),[9] Ralph Monroe Eaton (1892-1932)[10] und Harry Austryn Wolfson (1887-1974)[11] were members of the department. In Gordon Allport (1897-1967), Edwin G. Boring (1886-1968) and L.T. Troland (1889-1932) three outstanding psychologists were added to the mix.

Despite Goodman's interests in literature[12] and mathematics as a student, he already decided on philosophy in his first semester – in a seminar "on the early history of philosophy, on Thales, Empedocles and the other Presocratics. I immediately knew that philosophy was the right thing for me."[13] C.I. Lewis became his most influential philosophical teacher and

---

programmatic title *The New Realism: Cooperative Studies in Philosophy* (1912). Cf. (Kuklick 2001, 202ff).

[8] James Haughton Woods (1864-1935), after studying in Harvard, Cambridge (England), Berlin and Straßburg, taught at Harvard-University since 1913. In addition to Plato, Asian philosophy and religion were among his fields of research. In grateful memory Goodman mentions him in his dissertation – "To the late James H. Woods I owe a great debt for the kindness and encouragement which first led me to undertake this work." (Goodman 1940, i) – and in the preface to the first edition of *The Structure of Appearance*: "[...] I owe lasting gratitude to the late Professor James Haughton Woods for the indispensable initial spark of encouragement." (*SA* (1951), VII; *SA* (1977), XVII.)

[9] Henry M. Sheffer (1883-1964), a distinguished logician, taught at Harvard-University from 1908 until 1952. Among his works are: "A set of five independent postulates for Boolean algebras, with application to logical constants" (Sheffer 1913).

[10] Ralph Monroe Eaton (1892-1932) was Harvard instructor from 1919 until 1926, Assistant Professor from 1926 until 1932, Department Chairman from 1926 until 1930. His book *Symbolism and Truth: An Introduction to the Theory of Knowledge* published in 1925 was compared to Wittgenstein's *Tractatus*. On his life, which came to an early and tragic end, cf. the obituary in *The Philosophical Review* 6 (1933), 212-213, also (Kuklick 1977, 458f).

[11] Harry Austryn Wolfson (1887-1974) taught at Harvard University until 1958 (cf. Kuklick 1977, 457). He made a name for himself as an expert on Neoplatonist and Medieval Philosophy and as a Spinoza scholar. A professorship at the Brandeis University was named after him, which Goodman held from 1964 until 1967. (See below.)

[12] Goodman initially wanted to become a writer.

[13] Nelson Goodman in conversation with Karlheinz Lüdeking (Goodman 2005, 264). It seems reasonable to conjecture that he was talking of a seminar by the above-mentioned James Haughton Woods.

gave him several ideas for his early works. In a recollection of his fellow student and year-long comrade-in-arms Henry S. Leonard (1905-1967)[14] Goodman has given a witty account of their first years as students:

> Our introduction to philosophy included the historic running debate over idealism versus realism between W.E. Hocking and Ralph Barton Perry, and over monism and pluralism in logic between C.I. Lewis and Harry Sheffer. We were absorbed by Lewis' courses in the theory of knowledge, based on the just published *Mind and the World Order*;[15] and we were first exasperated and then enthralled by the nearly incoherent but inspired and profound lectures of James Haughton Woods on Plato. We sharpened our philosophical teeth in almost daily discussions of such matters as Berkeley's idealism, Plato's theory of ideas, Whitehead's extensive abstraction,[16] and problems in logic (Goodman 1969, ix).

Goodman received his B.Sc. in 1928 (Phi Beta Kappa, magna cum laude) and his Ph.D. – also from Harvard – in 1941 with *A Study of Qualities* (1940), which he later elaborated in his first major work, *The Structure of Appearance* (1951). The relatively long period of time to his dissertation is explained by a number of circumstances.

---

[14] Henry S. Leonard (1905-1967) studied under Whitehead and Lewis. Together with Leonard Goodman developed a new version of mereological nominalism, the "calculus of individuals". Leonard "also first proposed trying direct application of symbolic logic to some philosophical problems we had been working on. This resulted in our joint article "The Calculus of Individuals", and in some other material incorporated in his doctoral thesis, 'Singular Terms', and in my own later work." (Goodman 1969, x). Cf. Leonard's Dissertation (Leonard 1930), (Leonard; Goodman 1949) as well as (Goodman, 1940) and *SA*.

[15] *Mind and the World Order* published in 1929 is Lewis's first major work in epistemology; In it Lewis develops his influential strict version of epistemological fundamentalism. (The distinction between qualia and object properties was adopted from this work by Goodman. Cf. (Goodman 1940, 5n) and (*SA* (3. Aufl. 1977), 95n) Cf. also the further developments in: (Lewis 1946) and (Lewis 1952). In many works Goodman dealt with the writings of his teacher, cf. especially "Sense and Certainty" (Goodman 1952) and "Snowflakes and Wastebaskets" (in *PP*, 416-419).

[16] Whitehead developed his method of extensive abstraction, which Goodman mentions in his report, in his book *An Enquiry concerning the Principles of Natural Knowledge* (Whitehead 1919). Of Whitehead's works the contributions to mereological geometry must have been most interesting to Goodman. With reference to Whitehead's definition of geometrical points Goodman liked to explicate methodological questions of constructional definition (cf., for instance, *SA*, 5-7).

First and foremost, in this context, one has to mention Goodman's great passion for the arts. An intense commitment to all practical and theoretical questions of the artworld formed a major part of his life since the 1920s. In the seminars and colloquia given by the famous Museum director, collector and connoisseur Paul Joseph Sachs (1878-1965) on "Museum Work and Museum Problems", Goodman had not merely learned the basics of art history and art appreciation. He had become aware of how profound the difference is between cursorily glancing at pieces of art in museums or catalogues and an intense immersion in the original. Sachs invited his best students into his home and showed them originals in his possession or from the inventory of the William Hayes Fogg Art Museum, which belongs to Harvard. With his lively and practically oriented teaching style he managed to fill generations of students with enthusiasm for curatorial work and collecting art.[17] Another student of Sachs', John Walker put it thus: "He was someone who really made you want to be a collector. None of the rest could. They could teach you the history of art, but Paul was the one that made us all want to be collectors" (Ciniglio 1976; also quoted in Alexander 1997, 213). Goodman's passion for collecting art, too, was awakened. With a wink he commented that, ever since, he had been constantly bankrupt.[18]

From 1929 to his army draft in 1941, Goodman was the director of the Walker-Goodman Art Gallery at Boston's Copley Square. There he was also to meet his later wife, Katherine Sturgis (1904-1996), a Cambridge artist, who visited the gallery in order to exhibit her watercolour paintings and Indian ink drawings.[19] Henceforth, Goodman led a double life between philosophy and the arts. Museums, art fairs and auction houses were home to him just as much as conventions on logic, epistemology or aesthetics. His friend, Curtis L. Carter of the Haggerty Museum of Art at Marquette University retrospectively comments:

---

[17] On the pre-eminent importance of Sachs for the American museum world cf.. (Alexander 1997, Chapter 13: Paul Joseph Sachs Teaches a Pioneering Course in Museum Studies).

[18] Nelson Goodman in conversation with Karlheinz Lüdeking (Goodman 2005, 264).

[19] Goodman's readers know one of her works from the essay "The Status of Style" (reprinted in *WW*; the picture of the drawing is found on p. 30).

> Goodman's professional role as a gallery director and his private art collecting were sources of great satisfaction. His life-long pursuit of collecting art began in his student days. He was well known in the artworld for his discriminating aesthetic perception and equally for his astuteness in negotiating the price of an object. A visit to his home in Weston, Massachusetts would reveal a collector with enthusiasm an in-depth knowledge over a wide range of art (Carter 2000, 252).

The second reason for his delayed academic career is to be sought in his contemporary circumstances, which are easily forgotten today. "Because of his Jewish heritage, Goodman was ineligible for graduate fellowships" (Elgin 2000, 2).[20] W.V. Quine (1908-2000), who was aware of Goodman's circumstances of living, recommended him to Rudolf Carnap (1891-1970) for a job as an assistant. It is worthwile quoting from this letter of recommendation:

> My next recommendation would be H. Nelson Goodman, 607 Boylson St., Boston. Candidate for Ph.D. Age about 30. Very competent and industrious. An authority on the *Logischer Aufbau*, and conversant with your later work. Well grounded in philosophy and logic. As you know, he is in business – which accounts for the lateness of his Ph.D. I am not sure that he would not be interested in the assistantship. Teaching experience and publications none (Creath (ed.) 1990, 233).[21]

As the institutional framework of Carnap's professorship eventually did not turn out to be what had initially been promised to him, nothing came of this assistantship, from which both philosophers would undoubtedly have profited.

A third reason, which is not to be forgotten, is the high standard of the topics dealt with by Goodman. Among experts, *A Study of Qualities* has the reputation of being the most demanding and best doctoral dissertation ever handed in at Harvard. To his later co-author, Catherine Z. Elgin, he described his policy concerning scholarly projects as follows: "When you finish one project, ask yourself what is the most difficult outstanding problem in philosophy. Then work on that" (Elgin 2000, 2).

---

[20] Cf. (Schwartz 1999, 8): "This detour in career reflected both his passion for art and the difficulties with which someone of Jewish parentage then faced in academia".

[21] The correspondence between Carnap and Quine contains further material on Goodman's temporary co-operation with Quine, the first meetings with Carnap and other, related events. Cf. (Creath (ed) 1990, 191, 194, 201-204, 212, 235, 295, 397, 420, 422f., 434, 454, 460, 465).

After World War II, throughout which Goodman had been entrusted with carrying out psychological experiments for the U.S. Army, he worked as an Instructor in Philosophy at Tufts University for a brief interval. From there he went to the University of Pennsylvania, where he taught from 1946 to 1964, first as an associate professor and after 1951 as a full professor. From 1964 to 1967 he was Harry Austryn Wolfson Professor for Philosophy at Brandeis University. In 1968 he finally returned to Harvard University, where he taught till 1977.

It is there also that he founded "Project Zero",[22] in the context of which psychologists and philosophers study the cognitive skills that are developed while creating, understanding and judging works of art. This project, one of the foci of which is investigating how theoretical and practical competence in the arts can be refined, has had a significant impact on Art Education in the United States. At Harvard, Goodman was also involved in the founding of the Dance Center and of the Institute for Arts Administration, he was head of the Arts Orientation Series (1969-1977) and for many years was an advisor of Arts for Summer School (1971-1977).

A less well known fact is that Goodman designed three multimedia performance events, in the realisation of which he actively participated:

1) *Hockey Seen: A Nightmare in Three Periods and Sudden Death*, together with the choreographer Martha Armstrong Gray, the composer John Adams and Goodman's wife, the artist Katherine Sturgis, performed 1972 at Harvard, 1973 in Philadelphia and 1980 at Knokke-le-Zoute (Belgium), made into a Belgian Television production in 1980 and into a film at Harvard in 1984;[23]

---

[22] "I founded Project Zero in 1967 at the Harvard Graduate School of Education and directed it for four years. Since then it has been directed by David Perkins and Howard Gardner. The research is conducted by a varying group of psychologists, philosophers, and others, paid and unpaid, and has been reported in a number of books and papers." (*MM*, 147) The name "Project Zero" alludes to the awareness of having to start at zero. First results have been published in: (Leondar; Perkins 1977) as well as in *WW* and *MM*, especially Part V.

[23] "One production of *Hockey Seen* was sponsored by a professional hockey team and attended by many of its members. They gave it a good review." (Elgin 2000, 3.) – From September 28, 2006 through January 14, 2007 *Hockey Seen* was presented again

2) *Rabbit Run*, after a novel by John Updike, together with the choreographer Martha Armstrong Gray and the composer Joel Kabakov, as well as

3) *Variations, An Illustrated Lecture Concert*. At the heart of the latter stood twenty-two variations by Pablo Picasso on Diego Velazquez' painting *Las Meninas*, which Goodman had selected and arranged in a certain order. To this picture and its variations, which were shown as slides, the composer David Alpher composed a theme, also with twenty-two variations, which provided the score to the Lecture Concert. The programme was staged in different shapes at the University of Helsinki, at the Wayne State University, at the Rockport (Massachusetts) Chamber Music Festival, at Harvard University and at Trinity University in Texas.[24]

Over the years Goodman received several honours, among them four honorary doctorates, one of them from the Technische Universität Berlin. He delivered the Sherman Lectures at the University of London, the John Locke Lectures in Oxford, the Alfred North Whitehead Lecture at Harvard, the Immanuel Kant Lectures at Stanford and the Howison Lecture at Berkeley. In 1991, Goodman was honoured with an Author's Colloquium at the Centre for Interdisciplinary Research (Bielefeld, Germany).[25]

In his academic environment Goodmann had a reputation for being very strict, demanding and unapproachable. His impatience with slow-witted colleagues and students was feared. A description recurrent in all obituaries runs as follows: "he did not suffer fools gladly" (cf., for instance, Elgin; Scheffler; Schwartz 1999, 207; Elgin 2000, 1; Mitchell 1999). He placed high demands on himself and felt justified in also expecting a lot from others.

---

in the context of a multi-media exhibition at the Haggerty Museum of Art (Marquette University, Milwaukee).

[24] Cf. (*R*, Chapter IV), especially the references on p. 81 as well as (Carter 2000, 252) and (Elgin 2000, 3). – According to David Alpher's homepage, *Variations* has by now seen over 25 performances worldwide.

[25] The records of this colloquium organized by Peter Bieri and myself were published in a special edition of the journal *Synthese*, Vol. 95, No. 1, April 1993, cf. my introduction (Scholz 1993a).

The best scholars took their hats off to Goodman's unrelenting sharp wit. In such a vein Sir Alfred Jules Ayer (1910-1989) in his autobiographical writings freely reports on an encounter unforgettable to him:

> My inaugural lecture "Thinking and Meaning" was published as a pamphlet in 1947. I still believe that it posed some of the right questions but doubt if it gave many of the right answers to them. When I repeated it before a small audience of philosophers in New York in 1948 Nelson Goodman raised objections to it which I could not meet (Ayer, in: Hahn 1992, 23). [26]

From the published replies to objections one gets an idea of the mental presence characteristic of Goodman in conversation.

After having formally been made Professor Emeritus in 1977 Goodman can hardly be said to have retired. He published three further books full of new and provocative ideas: 1978 *Ways of Worldmaking* (*WW*), 1984 *Of Mind and Other Matters* (*MM*) and together with Catherine Z. Elgin 1988 *Reconceptions in Philosophy and Other Arts and Sciences* (*R*). He wrote replies and retractations for the volume *Starmaking* (McCormick 1996), which contains old and new contributions to the issue of pluralism about worlds. With the imaginativeness and optimism characteristic of him he further investigated the perspectives laid out in his early works.

That his works incurred increasing interest in Europe, too, filled Goodman with obvious joy. On his countless travels he would not be robbed of visiting the great art fairs and exhibitions of the whole world.

A fact hardly known is that Goodman and his wife were passionate animal lovers and committed animal conservationists. They always had dogs around them; their names are listed in the dedication of *Of Mind and Other Matters*: "To Snubby, Tweedledee, Randy, Angie, Debby, Susan, Trushka – for help and hindrance". As members of the World Society for the Protection of Animals and similar organisations, the Goodmans generously supported campaigns to save animals which were endangered by wars such as, for instance, the first Gulf War or the War in Bosnia, or by natural disasters.

---

[26] Cf. also the detailed report in the second part of Ayer's autobiography (Ayer 1985, 47). Ayer has respectfully acknowledged Goodman's œuvre in (Ayer 1982, 253-262).

Until shortly before his death, Goodman was working on a lecture which he wanted to deliver at a conference in Heidelberg.[27] The fact that until the very end he was absolutely determined to travel to Germany in order to deliver his lecture and reconvene with his friends is very characteristic of his constantly future-oriented thinking and striving. However, it never came to this journey. On November 25th 1998 Nelson Goodman died at the age of 92 in Needham, Massachusetts.

His steady gaze ahead and the corrolary reluctance to comfortably look back on his achievements are to be seen as the reason why we know so little about his life first-hand, whether spoken or written.[28] Thus, he rejected the not insignificant honour of a volume in the series *Library of Living Philosophers* (cf. Elgin 2000, 2). An Intellectual Autobiography as it characteristically opens the volumes of this series held no attraction for him whatsoever. He preferred to look ahead and continue his work.

2. Opinions

Even though Goodman placed highest demands on his publications[29], over the course of his long life, a considerable œuvre developed, which comprises eight books[30] and several essays and other smaller works.

---

[27] A volume dedicated to the commemoration of Nelson Goodman with the title *Wirklichkeit und Welterzeugung. In memoriam Nelson Goodman* (Fischer; Schmidt 2000) came out of this conference. It is, however, doubtful whether everything that is debated in this volume under the ambiguous heading "constructivism" really is in the spirit of Goodman's constructional philosophy.

[28] Goodman only rarely gave interviews; the most insightful one – the conversation with Karlheinz Lüdeking – is printed as a translation into German in (Steinbrenner; Scholz; Ernst 2005, 261-269). (The only further authorized interview, which was recorded for Belgian television in August 1980, has been published as an appendix to *MM,* pp. 189-200.)

[29] The high standard is epitomized by the fact that his first works were accepted and published by the *Journal of Symbolic Logic*.

[30] Counting *A Study of Qualities* (1940; printed 1990), which was elaborated in *The Structure of Appearance* (1951; 3rd edition 1977). This is justified because it contains important passages that were not taken over in *SA*. The volume *Starmaking* (McCormick 1996), which also contains, aside from Goodman's replies, contributions by other philosophers.

That Goodman was quite the opposite of a hasty and profuse writer is most impressively witnessed by the genesis of his dissertation and its history of publication.

> First thoughts toward the thesis began to occur during my undergraduate years, 1924-1928, at Harvard.[31] A 'finished version' was ready by 1933, but the project had by then expanded enough to take another seven years. Still, hundreds of changes were made from thesis to book (Goodman 1940, preface).

The constantly reworked book was finally published in 1951 under the title *The Structure of Appearance*.[32]

Goodman's first major work is doubtless his most demanding and difficult book; unfortunately it has also remained his least well-known work. This is regrettable for several reasons. First, of course, because it contains so many ideas and develops projects, which were taken up and developed further only by a few specialists. Second, this book already contains the seeds of later works, which are hardly comprehensible without this background. This is true particularly of *Ways of Worldmaking,* the "radical relativism under rigourous restraints" (*WW*, x) which is often received in a watered-down form, because the results of the earlier work are not taken into account.

In continuation and critical analysis of *Our Knowledge of the External World* (1914) and other works by Bertrand Russell (1872-1970), *Mind and the World Order* (1929) by C.I. Lewis (1883-1964) and most importantly *Der logische Aufbau der Welt* (1928) by Rudolf Carnap (1891-1970), Goodman develops his own constructional philosophy and provides a host of constructional systems for characterising the formal structure of

---

[31] Goodman stressed in conversation that he had already been working on the questions he deals with in his dissertation before the publication of *Der logische Aufbau der Welt* (1928). This work having been pointed out to him (probably by C.I. Lewis, maybe also by Henry S. Leonard, who was in Munich in 1929), it did, of course, become the work, against which he most set himself of throughout his further work. (As Carnap's work would only be translated into English in 1967 (cf. Carnap 1967), Goodman had to translate it himself). In any case, the widespread opinion that Goodman's early works are simply a reaction to Carnap's *Aufbau* has to be revised.

[32] A second, revised edition was published in 1966, a third one, once again revised, in 1977. The typed version of *A Study of Qualities* that Goodman had handed in was only to be published in 1990.

experience, in particular a detailed phenomenalist system, as it were: "a structure of appearances". Among the many further riches contained within this work, one may mention: "the calculus of individuals", a version of formal mereology which Goodman had developed together with Henry S. Leonard as a graduate student, a theory of simplicity (of predicates), systematic investigations of ordering, measurement, time and a theory of indexical expressions, especially temporal deixis.[33]

At that stage already, an anti-absolutist pluralism had become the hallmark of Goodman's philosophy. In the preface to his dissertation (1940!) he writes: "Perhaps the conviction I should acknowledge most willingly is that absolutism must be rejected" (Goodman 1940, v). He continues his confession in a way particularly characteristic of his constructional philosophy:

> But nothing is much emptier than a relativism that, without seeking to find any solution to a given problem, merely expresses a willingness to admit alternative solutions. Freedom has little value if we use it solely to declare that we have it. The recognition that theoretically there are many equally satisfactory systems is but the first step in the laborious task of realizing at least one of them (Goodman 1940, v).

Goodman has taken up this task in the main parts of *A Study of Qualities* and *The Structure of Appearance*; the result is a detailed phenomenalist system with qualia as atoms. The cognitive aim of such system building is explained by Goodman as follows:

> [...] the purpose of constructing a system is to interrelate its predicates. The same purpose is served by reducing to a minimum the basis required. Every definition at once both increases the coherence of the system and diminishes the number of predicates that need be taken as primitive. Thus the motive for seeking economy is not mere concern for superficial neatness. To economize and to systematize are the same (*SA*, 47f).

An appropriate appreciation of this first major work, pioneering methodically as well as with regard to content, is yet to be awaited.[34]

---

[33] The best introduction to Goodman's first major work, its position in the œuvre as a whole and in contemporary philosophy still is Geoffrey Hellman's "Introduction", which is prefaced to the third edition (Dordrecht, Boston 1977), pp. XIX-XLVII.

[34] Cf. the abovementioned "Introduction" to the third edition by Geoffrey Hellman.

In the summer of 1947 a very lively correspondence[35] ensued between Goodman, W.V. Quine and Morton G. White, from which influential works emerged that attacked the analytic-synthetic distinction and connected terms like intensional synonymy, apriority, metaphysical necessity etc. At the end of this correspondence, Morton White was chosen to give a critical overview of the issues, a task which he carried out in an exemplary way in "The Analytic and the Synthetic: An Untenable Dualism" (White 1950). Quine's essay "Two Dogmas of Empiricism", in which he drew far-reaching philosophical and metaphilosophical conclusions from the failure of attempts to give satisfactory definitions of the incriminated terms, became much more famous.

Goodman's reaction was different in a telling way. While he gave an extensional explication of "likeness of meaning" in "On Likeness of Meaning", which is capable of replacing an absolute concept of synonymy by a workable one[36], he did not participate in the hustle and bustle caused by the discussion of the analytic-synthetic distinction and, in particular, the essay "Two Dogmas". Instead of painting a gloomy picture of the possible consequences of the absence of an absolute distinction, Goodman preferred to continue his work – without relying on terms as questionable as analyticity, necessity, essential property or natural kind. Goodman was always more interested in philosophy than in meta-philosophy.

In accordance with his above-quoted policy to always seek out the greatest challenge, in the works culminating in the book *Fact, Fiction, Forecast* (*FFF*) (1954), he turned his attention to the most intricate complex of problems in epistemology and the philosophy of science. The issues can be outlined by the following titles: "dispositions", "potentiality", "counterfactual conditionals", "inductive confirmation" and "natural laws". *Fact, Fiction, Forecast,* in a particularly impressive way, exemplifies Goodman's method of "reconception" or, as one might say, fruitful transference of problems. A number of significant achievements may be

[35] On this episode of importance to the philosophy of the following decades cf. (Quine 1985, 226; 1960, 67, n. 7- 68; Wang 1986, 130-133; White 1999 as well as Creath 1990, 35). This volume also contains Carnap's first reaction to "Two Dogmas", cf. (ibid., 427-432). The correspondence between White, Quine und Goodman is now published in the appendix to (White 1999, 337-357).

[36] Cf. (Heydrich 1993) with an important amendment and further references.

distinguished within this small book: As a start, Goodman has put forward an analysis of the specific character of the above-mentioned philosophical problems. Goodman not only shows that all of these problems are so far unsolved, but also that they are thoroughly intertwined. Their solution presupposes the resolution of a much more general problem so far hardly recognised, let alone overcome: How do lawlike statements differ from non-lawlike, accidental generalisations? In so doing, in particular, the traditional problem of induction is reconceived. For one part it can be solved, for another part it has to be dissolved to uncover a real and general problem: the problem of law-likeness, that is, more generally the problem of projection (of statements and the predicates occurring in them) from one set of cases to another one. Finally, Goodman has undertaken the first steps towards developing a general solution to this ubiquitous problem (Chapter 4).

The so-called problem of justifying induction is dissolved if it is understood as the search after guarantees for hypotheses. Understood in this way the problem turns out to be simply a curious attempt to justify knowledge we do not possess. It is, however, possible to solve another problem: One can give a general answer to the question of how rules of inductive inference are justified. Rules of inductive reasoning are codifications of an antecedent practice; they are justified if they adequately represent this practice:

> A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend. The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either (*FFF*, 64).

For this method of mutual adjustment the term "reflective equilibrium" has become commonplace since John Rawls' *A Theory of Justice* (1971), in which it has been applied to moral principles (Rawls 1971, 20; cf. also Daniels 1996; Hahn 2000).

The new riddle of induction shows that syntactic (formal) and semantic conditions are not sufficient to draw the dividing line between valid and invalid inductive inferences. In retrospect, Goodman notes: "[...] we may by now confidently conclude that no general distinction between projectible and non-projectible predicates can be drawn on syntactic or

even on semantic grounds." (*PP*, 357). Inductive Logic, if at all one can speak of such a thing, cannot be formal in the same way as deductive logic is.

In the conclusion of his book Goodman has suggested a solution, which makes use of the pragmatic term "entrenchment", i.e. the historical anchoring of a predicate in linguistic practice. The relative projectibility of generalisations is primarily determined by the relative anchoring of the predicates occurring in it, i.e. by the extent to which these predicates have been used in previously projected generalisations.

In this way the concepts of confirmation and lawlikeness are relativised to linguistic practices. In the pragmatic-historical theory past projections, i.e. earlier projections from known to unknown cases, have to be taken into account in judging the lawlike character. It is insightful to see how Goodman locates his suggestion between Hume and Kant:

> Like Hume, we are appealing here to past recurrences, but to recurrences in the explicit use of terms as well as to recurrent features of what is observed. Somewhat like Kant, we are saying that inductive validity depends not only upon what is presented but also how it is organized; but the organization we point to is effected by the use of language and is not attributed to anything inevitable or immutable in the nature of human cognition. To speak very loosely, I might say that in answer to the question what distinguishes those recurrent features of experience that underlie valid projections from those that do not, I am suggesting that the former are those features for which we have adopted predicates that we have habitually projected (*FFF*, 97).

Had Goodman written only this book, he would already have been granted a place of honour in the history of philosophy.[37] The renowned philosopher of science Ian Hacking in retrospect commented on the problems and riddles brought up by Goodman: "They combine precision of statement, generality of application, and difficulty of solution to a degree greater than any other philosophic problem broached in this century" (Hacking 1965, 41).

---

[37] The discussion is documented in: (Stalker 1994). This volume contains an extensive commented bibliography. – A selection of the most important contributions with an instructive introduction is reprinted in: (Elgin 1997). Applications and Conclusions in the fields of social sciences and medicine are examined in: (Douglas; Hull 1992).

In the 60ies, for the first time, Goodman brought together his two great passions – theoretical philosophy and the arts – in his research and his publications, too. From a thorough epistemological and semiotic study of the arts in 1968 the third book emerged, *Languages of Art*, which was also to achieve epoch-making status.

In this approach to and outline of a general theory of symbols the structures of various symbol systems as they function in everyday life, in the arts and in the sciences are illuminated by astute analyses and constructive comparisons. The manifold forms of symbolisation in the arts are, and this is one of Goodman's major points, of an importance in our cognitive endeavours equal to that of theories, formulae and diagrams in the sciences. Works of art are, just like scientific systematizations, complex symbols, which we create, apply and interpret in our striving for knowledge and understanding.

In the book semiotic, aesthetic and epistemological investigations run together. As subheading and introduction emphasise the primary aim is the development of a general theory of symbols which is intended to illuminate the various symbol systems in everyday life, science and technology as well as in the arts. The term "symbol" is here understood in a very broad manner, covering "letters, words, texts, pictures, diagrams, maps, models, and more" (*LA*, xi). Aesthetic questions serve as starting points and are given inventive solutions.

The first two chapters investigate two traditional aesthetic issues: the nature of pictorial representation (I) and the nature of aesthetic expression (II). Both relations are characterised semiotically as forms of reference. In the course of a sweeping critique of resemblance theories of pictures, representation is provisionally characterised as a subcase of denotation. In analysing the concept of expression, attention is drawn to the neglected semiotic relation of exemplification, in which an object, as a sample, refers to predicates it instantiates. That a work expresses a feeling or the likes comes down to its metaphorically exemplifying suitable predicates.

In chapter III a new strand begins that is only in the end brought together with the others. Goodman brings up the question of how it is possible in a given art that there are forgeries and thus a distinction between original and fake. There are fundamental differences between so-called autographical arts (like painting), in the case of which authenticity can only be confirmed

historically, and allographical arts (like music), in the case of which the constitutive features of a work can be fixed by a notation. In this way, clarifying the differing criteria of identity for works of art leads to the task of investigating the nature and function of notational systems.

Chapter IV, the systematic centre-piece of the book, in the theory of notation develops a conceptual instrument, which makes possible the description and classification of symbol systems in general. Musical and other notational systems represent an illuminating borderline case of symbol systems: They are characterized by five requirements which are only partially or not at all satisfied in other symbol systems. Notational systems are syntactically disjoint and differentiated systems, they are free of ambiguities and are also semantically disjoint and differentiated. To simplify a bit, in a notational system a maximally high degree of syntactic and semantic definiteness is realised.[38] Chapter V demonstrates up the fertility of the theory by applying it to music, the visual arts, literature, dance and architecture.

In the final chapter a few loose threads are then tied together. It is now possible to draw a clearer distinction between pictorial and verbal symbol systems (VI.1-2). Finally (VI.3-7), a cognitivist conception of the aesthetic and the arts is sketched (cf. Steinbrenner 1996; Ernst 2000; Scholz 2001). The arts, which, like the sciences, are cognitive practices, actively contribute to structuring, understanding and constructing the worlds in which we live. The contrast between the sciences and the arts is based on differences as to the dominance of features of the symbols used. Tentatively four "symptoms of the aesthetic" are given, to which Goodman added a fifth in *Ways of Worldmaking* (1978) (*WW*, 68). *Languages of Art* today already counts as a classic work in aesthetics; increasingly it is also receiving attention in discussions on semiotics and epistemology.

The collection of essays *Problems and Projects* (1972) contains important retractions from and further developments of ideas in the first three books as well as replys to criticisms. In part I Goodman speaks out on his understanding of philosophy, its aims and methods (cf. Cohnitz; Rossberg 2006, Chapter 3).

---

[38] For more elaborate accounts cf. (Scholz 1991; 2004, Chapter 4) and (Cohnitz; Rossberg 2006, Chapters 6 and 7).

In subsequent works Goodman follows up on the far-reaching epistemological and metaphysical consequences of this general theory of signs. The symbols of the manifold systems do not provide passive representations of a world to be discovered, but go into constituting that which is referred to. In other words: They are constitutive of worlds. We are thus confronted with a plurality of scientific and artistic world versions, none of which can exclusively lay claim to be the true one. This pluralism is, however, not to be confused with an irresponsible relativism á la "anything goes"; for the difference between right and wrong world versions is not at all abandoned. The investigation of the respective standards of correctness moves centre-stage in Goodman's work, and it is plainly seen that the fulfillment of these standards is anything other than trivial. As mentioned above, in *Ways of Worldmaking* (1978) Goodman speaks of a "radical relativism *under rigorous restraints*" (*WW*, x; my emphasis). He adds:

> Nevertheless, I think of this book as belonging in that mainstream of modern philosophy that began when Kant exchanged the structure of the world for the structure of the mind, continued when C.I. Lewis exchanged the structure of the mind for the structure of concepts, and that now proceeds to exchange the structure of concepts for the structure of the several symbol systems of the sciences, philosophy, the arts, perception, and everyday discourse. The movement is from unique truth and a world fixed and found to a diversity of right and even conflicting versions or worlds in the making (*WW*, x).

The responses by top-class colleagues of Goodman as to this pluralism concerning worlds have been compiled in the volume *Starmaking* (1996) edited by Peter J. McCormick; in it Goodman responded, for the last time, to what he called "worldly worries".[39]

In 1984, a further collection of essays was published, *Of Mind and Other Matters* (*MM*). Like *Problems and Projects* (*PP*) it also contains replies to criticisms and further developments. Most noteworthy are the developments of the general theory of symbols in Part III, particularly in the important paper "Routes of Reference", as well as few smaller papers on philosophical psychology in Part I.

---

[39] It may be noted that Goodman's essay "On Some Worldly Worries" reprinted in this volume by McCormick had initially been published in the special edition of *Synthese* (Bieri; Scholz 1993).

Goodman's provocative contributions have suggested again and again that prevalent conceptions of philosophical ideals like certainty, knowledge and truth have to be reconceived and revised. How this could be done is the subject of Goodman's last greater work, *Reconceptions in Philosophy and Other Arts and Sciences* (1988) (*R*), which he co-authored with Catherine Z. Elgin.

If all fields of cognition, all kinds of symbols and reference are considered, new guideline concepts need to be accorded central status, for instance: current adoption, understanding and correctness.[40] Against the backdrop of the general theory of symbols and an epistemological and metaphysical pluralism Goodman and Elgin, in this book, propagate a revision and reorientation of philosophy.

The introductory chapter recapitulates the symbol-theoretic research program and its epistemological and metaphysical consequences: The symbols of everyday life, of the sciences and the arts do not simply describe an independently given world, they go into the constitution of what is referred to.

Nonetheless, rigorous restraints (nominalism) and standards of correctness (consistency, deductive and inductive correctness, rightness of categorisation etc.) hold for the construction of adequate world versions. In this sense constructionalist pluralism in no way comes down to an "irresponsible relativism".

Having given a diagnosis of the plight of prevalent epistemologies, the conclusion makes suggestions for due reforms. To account for all fields of cognition, all kinds of symbols and reference guideline concepts like "truth", "certainty" and "knowledge" need to be expanded or even replaced; as successors a concept of rightness of symbolic functioning, a concept of tentative adoption and a concept of understanding in a broad sense are recommended.

In the detailed studies in the middle part the fertility of this reorientation is shown for a semiotics of pictorial representation and architecture, for literary theory and for the cognitive sciences. In the final chapter "A Reconception of Philosophy" Goodman and Elgin also give a brief

[40] Cf. my review in the *Frankfurter Allgemeine Zeitung* (November 14th, 1989) as well as the relevant article in (Volpi 1999, 582-583).

characterisation of the overall project as it emerged since *Languages of Art* (1968) and Elgin's *With Reference to Reference* (1983):[41]

> We work *from* a perspective that takes in the arts, the sciences, philosophy, perception, and our everyday worlds, and *toward* better understanding of each through significant comparison with the others. Speaking schematically, the first phase of this effort begins by observing that the use – that is, the fabrication, application, and interpretation – of symbols is centrally involved in all these fields. Accordingly, a general theory of symbols and their functions is outlined (*LA*; *RR*). The *second* phase confronts the consequences of recognizing that symbols are not merely devices for describing objects, events, a world waiting to be discovered, but enter into the very constitution of what is referred to (*WW*). The present third phase starts from the realization that the prevailing conception of philosophy is hopelessly deficient when all fields of cognition, symbols of all kind, and all ways of referring are taken into account, and so goes on to search for more comprehensive and responsive concepts (*R*, 164).

However different the objects and applications of Goodman's philosophy have been, continuous characteristics are a strict form of nominalism, persistent criticism of the idea of "the given" and concomitantly an emphasis on the active, constructive and creative elements of perception, cognition and understanding. His epistemology is anti-fundamentalist; instead, Goodman recommends a coherentist conception of epistemic justification. Metaphysical points of contention should make room for methodological questions. In an allusion to a hope, which he had raised in the preface to the first edition of *The Structure of Appearance,* Goodman comments in the preface to the third edition:

> Unfortunately, the hoped-for day when philosophy will be 'discussed in terms of investigation rather than controversy, and philosophers, like scientists, be known by the topics they study rather than the views they hold' has not yet come (*SA*, XIII).

3. Reception

During his lifetime Nelson Goodman already counted as one of the most important philosophers of the 20th century. Anyhow, a comprehensive reception of his complex œuvre is still in its beginnings.

---

[41] *With Reference to Reference* (*RR*) is Elgin's attempt to systematize the theory of symbols developed by Goodman. Cf. my review (Scholz 1988, 336-340).

If one looks to the United States, to begin with, it is to be noted that, while Goodman is being quoted and discussed all over, he has not, in the narrow sense, become the figurehead of a particular philosophical school. There are, however, a number of excellent students of his.

Two of his most famous students, Noam Chomsky (born 1928) and David K. Lewis (1941-2002) have, at an early stage, developed in directions quite different from their teacher. Chomsky and his students, most prominently Jerry Fodor, have revived the Cartesian doctrine of innate ideas, which Goodman finds mistaken, even incomprehensible.[42] And David Lewis tried to rehabilitate concepts, which had been discredited by Goodman and Quine.

Among those who sought to further develop and communicate his ideas to a wider audience one might note Israel Scheffler, Catherine Z. Elgin, Joseph Ullian, Richard Rudner, Robert Schwartz, Margaret Atherton, Geoffrey Hellman and Marsha Hanen. [43]

With other outstanding philosophers there was temporary cooperation. The intermittent collaboration with Henry S. Leonard and W.V. Quine has already been mentioned. More recently Hilary Putnam, Goodman's colleague at Harvard, has pointed out a convergence of their positions.[44] But despite this illustrious series of followers and allies Goodman has not, unlike, e.g., W.V. Quine, Donald Davidson or Michael Dummett, become the figurehead of an influential school in the Anglo-Saxon world.

Several reasons for this may be assumed. The most important of these may be taken to lie in the versatility of Goodman's œuvre and an unfortunately fragmented reception. Only a few have studied all of his major works, even less are capable of seeing the connections in terms of content and method between his contributions.

---

[42] Goodman dedicated two witty essays to the revival of the doctrine of innate ideas by Chomsky, which are reprinted in *Problems and Projects* (*PP*, 69-75, 76-79). Goodman's student's Margaret Atherton and Robert Schwartz, too, have repeatedly and critically dealt with innateness hypotheses à la Chomsky and Fodor.

[43] Cf. also the works mentioned in the bibliography, which are, of course, only a selection.

[44] Cf. the preface to his book *Truth, Reason and History* (1981). Putnam wrote a bright and empathic preface to the fourth edition of *FFF* (1983).

In Great Britain Alfred Jules Ayer, whom we have quoted above, was one of the few admirers of Goodman's work. It seems that through Michael Dummett's harsh criticism of *The Structure of Appearance* the reception of Goodman's philosophy in England was all in all ill-omened (Dummett 1955, 1956, 1957). Peter Hacker sums it up thus: "It is true that philosophers in Oxford placed no faith in the kind of system-building Goodman had embarked upon in *The Structure of Appearance*" (Hacker 1996, 230). Retrospectively, however, Dummett commented: "I regret that I did not make more explicit the admiration that I felt for *The Structure of Appearance*" (Dummett 1978, xli).[45]

In contrast to the British Isles, there is a growing reception of Goodman's work in many other European countries, especially in Scandinavia, Belgium and, perhaps somewhat surprisingly, in France (cf. e.g. Genette 1992, 1994; Morizot 1996; Pouivet 1996).

In the German-speaking world the reception began very early. As early as in the 40ies Heinrich Scholz (1884-1956), then Professor for Philosophy at Münster University, asked Goodman to send him his first works in logic. (Goodman kept the postcards and later showed them to me.) In the *Deutsche Literaturzeitung* he tried to introduce *The Structure of Appearance* and *Fact, Fiction, Forecast* to the German audience (cf. Scholz 1951, 1958). Since the 50ies Wolfgang Stegmüller (Munich) has regularly pointed out Goodman's seminal contributions (cf. Stegmüller 1957, 1958/59, 1969, 1983). The journal *Erkenntnis* edited, at the time, by Carl G. Hempel, Wolfgang Stegmüller and Wilhelm K. Essler dedicated two special issues to Goodman's works in 1978 (*Erkenntnis* 12/1, 1978). Besides Stegmüller Franz von Kutschera (Regensburg),[46] Guido Küng (Fribourg, Switzerland),[47] Gottfried Gabriel (Jena),[48] Jens Kulenkampff (Erlangen)[49] and Günter Abel (Berlin, Technical University)[50] deserve to be mentioned. In 1990, Goodman received an honorary doctorate from the

[45] A thorough investigation and rejection of Dummett's criticism is given by Farrell (1974, 223-249).

[46] Cf. (von Kutschera 1972, 1975, 2005).

[47] Cf. (Küng 1967, 1993; Hottinger 1988).

[48] Cf. (Gabriel 1986, 1991, 1997, 2000; Thürnau 1994).

[49] Cf. (Kulenkampff 1978, 1981, 1997, 2005a, 2005b).

[50] Cf. (Abel 1991, 1993, 1999).

Technical University Berlin. In 1991, he was honoured by an Author's Colloquium at the Centre for Interdisciplinary Research in Bielefeld, Germany (cf. Bieri; Scholz 1993).

In more recent years, a younger generation of Goodman aficionados emerged. In 1996, some of us met near Munich to celebrate his 90th birthday (cf. Steinbrenner; Scholz; Ernst 2005). Finally, in 2006, a conference *Nelson Goodman: From Logic to Art – Looking Back on the Occasion of his 100th Birthday* was organized by Gerhard Ernst, Jakob Steinbrenner and me in Munich. This conference is documented in this volume.

References

Abel, G. (1991): "Logic, Art, and Understanding in the Philosophy of Nelson Goodman", in: *Inquiry* 34 (1991), 311-321.

Abel, G. (1993): *Interpretationswelten: Gegenwartsphilosophie jenseits von Essentialismus und Relativismus*, Frankfurt am Main: Suhrkamp, 1993.

Abel, G. (1999): *Sprache, Zeichen, Interpretation*, Frankfurt am Main: Suhrkamp, 1999.

Alexander, E.P. (1997): *The Museum in America: Innovators and Pioneers*, Walnut Creek: AltaMira Press, 1997.

Atherton, M. (1970): *Nativism*, Doctoral Dissertation, Brandeis University, 1970.

Atherton, M. (1990): *Berkeley's Revolution in Vision*, Ithaca; N.Y.: Cornell University Press, 1990.

Ayer, A.J. (1982): *Philosophy in the Twentieth Century*, London: Weidenfeld, 1982.

Ayer, A.J. (1985): *More of My Life*, Oxford: Collins, 1985.

Bieri, P.; Scholz, O.R. (eds) (1993): *Probing into 'Reconceptions'*, in: *Synthese* 95, No. 1 (1993).

Carnap, R. (1928): *Der logische Aufbau der Welt*, 2nd Edition, Hamburg: Meiner, 1961.

Carnap, R. (1967): *The Logical Structure of the World and Pseudoproblems in Philosophy*, translated by Rolf A. George, Berkeley; Los Angeles: University of California Press, 1967.

Carter, C.L. (2000): "A Tribute to Nelson Goodman", in: *Journal of Aesthetics and Art Criticism* 58 (2000), 251-253.

Ciniglio, A.V. (1976): "Pioneers in American Museums: Paul J. Sachs", in: *Museum News* 55 (1976), 69-70.

Cohnitz, D.; Rossberg, M. (2006): *Nelson Goodman*, Chesham: Acumen, 2006.

Creath, R. (ed.) (1990): *Dear Carnap, Dear Van. The Quine-Carnap Correspondence and Related Work, edited with an Introduction by Richard Creath*, Berkeley; Los Angeles; London: University of California Press, 1990.

Daniels, N. (1996): *Justice and Justification. Reflective Equilibrium in Theory and Practice*, Cambridge: Cambridge University Press, 1996.

Douglas, M. and Hull, D. (eds) (1992): *How Classification Works: Nelson Goodman among the Social Sciences*, Edinburgh: Edinburgh University Press, 1992.

Dummett, M. (1955): "Critical Notice: The Structure of Appearance", in: *Mind* 64 (1955), 101-109.

Dummett, M. (1956): "Nominalism", in: *The Philosophical Review* 65 (1956), 491-505.

Dummett, M. (1957): "Constructionalism", in: *The Philosophical Review* 66 (1957), 47-65.

Dummett, M. (1978): *Truth and Other Enigmas*, London; Cambridge (Mass.): Harvard University Press, 1978.

Eberle, R. (1970): *Nominalistic Systems*, Dordrecht: D. Reidel, 1970.

Elgin, C.Z. (1983): *With Reference to Reference*, Indianapolis: Hacket, 1983.

Elgin, C.Z. (1992): "Nelson Goodman", in: Jonathan Dancy und Ernest Sosa (eds): *A Companion to Epistemology*, Oxford: Blackwell, 1992, 162-164.

Elgin, C.Z. (1996): *Considered Judgment*, Princeton: Princeton University Press, 1996.

Elgin, C.Z. (1997): *Between the Absolute and the Arbitrary*, Ithaca, N.Y.: Cornell University Press, 1997.

Elgin, C.Z. (1997b): “Nelson Goodman’s New Riddle of Induction”, in: Elgin, C.Z. (ed.): *The Philosophy of Nelon Goodman: Selected Essays*, New York: Garland, 1997.

Elgin, C.Z. (1998): “Goodman, Nelson”, in: E. Craig (ed.), *Routledge Encyclopedia of Philosophy*, Vol. 4, London; New York: Routledge, 1998, 135-140.

Elgin, C.Z. (2000): “Worldmaker: Nelson Goodman (1906-1998)”, in: *Journal for General Philosophy of Science* 31 (2000), 1-18.

Elgin, C.Z. (2001): “The Legacy of Nelson Goodman”, in: *Philosophy and Phenomenological Research* 62 (2001), 679-690.

Elgin, C.Z. (2004): “Denying a Dualism: Goodman’s Repudiation of the Analytic/Synthetic Distinction”, in: *Midwest Studies in Philosophy* 28 (2004), 226-238.

Elgin, C.Z. (ed.) (1997): *The Philosophy of Nelson Goodman. Selected Essays*, Vol. 1-4, New York: Garland, 1997.

Elgin, C.Z.; Scheffler, I.; Schwartz, R. (1999): “Nelson Goodman 1906-1998”, in: *Proceedings and Addresses of the American Philosophical Association* 72:5 (1999), 206-208.

Ernst, G. (2000): “Ästhetik als Teil der Erkenntnistheorie bei Nelson Goodman”, in: *Philosophisches Jahrbuch* 107 (2000), 315-340.

Farrell, R. (1974): “Michael Dummett on the Structure of Appearance”, in: *Synthese* 28 (1974), 223-249.

Fischer, H.R.; Schmidt, S.J. (eds) (2000): *Wirklichkeit und Welterzeugung. In memoriam Nelson Goodman*, Heidelberg: Carl-Auer-Systeme, 2000.

Gabriel. G. (1986): “Ein Mann von Welten. Besprechung von Nelson Goodmans ‚Weisen der Welterzeugung’”, in: *Philosophische Rundschau* 33 (1986), 48-55.

Gabriel, G. (1991): *Zwischen Logik und Literatur. Erkenntnisformen von Dichtung, Philosophie und Wissenschaft*, Stuttgart: Metzler, 1991.

Gabriel, G. (1997): *Logik und Rhetorik der Erkenntnis*, Paderborn: Ferdinand Schöningh, 1997.

Gabriel, G. (2000): “Kontinentales Erbe und Analytische Methode. Nelson Goodman und die Tradition”, in: *Erkenntnis* 52 (2000), 185-198.

Genette. G. (1992): *Fiction et Diction*, Paris: Seuil, 1992.

Genette, G. (1994): *L’Œuvre de l’Art*, Paris: Seuil, 1994.

Gochet, P. (1972): *Esquisse d'une Théorie Nominaliste de la Proposition*, Paris: Armand Colin, 1972.

Gombrich, E.H.J. (1960): *Art and Illusion. A Study in the Psychology of Pictorial Representation*, Princeton: Princeton University Press, 1960.

Gombrich, E.H.J. (1982): *The Image and the Eye. Further Studies in the Psychology of Pictorial Representation*, Oxford: Phaidon, 1982.

Goodman, N. (1940): *A Study of Qualities* [1940], Ph.D. Thesis, New York: Garland, 1990.

Goodman, N. (1952): "Sense and Certainty", in: *Philosophical Review* 61 (1952), 160-167. Reprinted in (*PP*, 60-68).

Goodman, N. (1969): "Memorial Note", in: K. Lambert (ed.): *The Logical Way of Doing Things*, New Haven; London: Yale University Press, 1969, ix-x.

Goodman, N. (1983): "Semiotische Ästhetik und pragmatischer Irrealismus", in: *Zeitschrift für Semiotik* 5 (1983), 256-257.

Goodman, N. (1997): "Some Reflections on my Philosophies", in: *Philosophia Scientiae* 2 (1) (1997), 15-20.

Goodman, N. (2005): "Gewissheit ist etwas ganz und gar Absurdes" (Nelson Goodman im Gespräch mit Karlheinz Lüdeking), in: Steinbrenner, J.; Scholz, O.R.; Ernst, G. (eds): *Symbole, Systeme, Welten. Studien zur Philosophie Nelson Goodmans*, Heidelberg: Synchron, 2005, 261-269.

Gosselin, M. (1990): *Nominalism and Contemporary Nominalism: Ontological and Epistemological Implications of the Work of W.V. Quine and N. Goodman*, Dordrecht: Kluwer, 1990.

Grunstra, B.R. (1972): "Empirical and Conventional Elements in Certain Numerical Laws", in: Rudner, R. and Scheffler, I. (eds): *Logic and Art: Essays in Honor of Nelson Goodman*, Indianapolis: Bobb-Merrill, 1972, 21-42.

Hacker, P. (1996): *Wittgenstein's Place in Twentieth-Century Analytic Philosophy*, Oxford: Blackwell, 1996.

Hacking, I. (1965): *Logic of Statistical Inference*, Cambridge: Cambridge University Press, 1965.

Hahn, L.E. (ed.) (1992): *The Philosophy of A.J. Ayer* (= *The Library of Living Philosophers*, Volume XXI), La Salle, Ill.: Open Court Press, 1992.

Hahn, S. (2000): *Überlegungsgleichgewicht(e)*, Freiburg; München: Alber, 2000.

Hanen, M. (1970): *An Examination of Adequacy Conditions for Confirmation*, Doctoral dissertation, Brandeis University, 1970.

Hellman, G. (1977): "Introduction", in: *SA*, XIX-XLVII.

Heydrich, W. (1982): *Gegenstand und Sachverhalt: Bausteine zu einer nominalistisch orientierten Semantik für Texte*, Hamburg: Buske, 1982.

Heydrich, W. (1993): "A Reconception of Meaning", in: *Synthese* 95, No. 1 (1993), 77-94.

Hocking, W.E. (1961): "Whitehead as I Knew Him", in: *The Journal of Philosophy* 58 (1961), 505-516.

Hottinger, S. (1988): *Nelson Goodmans Nominalismus und Methodologie*, Bern; Stuttgart: Haupt, 1988.

Ihwe, J.F. (1985): *Konversationen über Literatur: Literatur und Wissenschaft aus nominalistischer Sicht*, Braunschweig: Vieweg, 1985.

Küng, G. (1963): *Ontologie und logistische Analyse der Sprache*, Wien: Springer, 1963.

Küng, G. (1967): *Ontology and the Logistic Analysis of Language*, Revised Edition, translated by E.C.M. Mays, Dordrecht: D. Reidel, 1967.

Küng, G. (1977): "Nominalistische Logik heute", in: *Allgemeine Zeitschrift für Philosophie* 2 (1977), 29-52.

Küng, G. (1993): "Ontology and the Construction of Systems", in: *Synthese* 95 (1993), 29-53.

Kuklick, B. (1977): *The Rise of American Philosophy: Cambridge, Massachusetts 1860-1930*, New Haven; London: Yale University Press, 1977.

Kuklick, B. (2001): *A History of Philosophy in America 1720-2000*, Oxford: Oxford University Press, 2001.

Kulenkampff, J. (1978): "Nelson Goodman, Languages of Art/dt.: Sprachen der Kunst", in: *Philosophische Rundschau* 25 (1978), 161-176.

Kulenkampff, J. (1981): "Music Considered as a Way of Worldmaking", in: *Journal of Aesthetics and art Criticism* 39 (1981), 254-258.

Kulenkampff, J. (1997): "Von Einhörnern und Kentauren. Schwierigkeiten mit Goodmans Theorie der Darstellung", in: J. Nida-Rümelin (ed.): *Rationalität, Realismus, Revision*, Berlin; New York: de Gruyter, 1997, 761-768.

Kulenkampff, J. (2005a): “Sind Bilder Zeichen?”, in: Steinbrenner, J.; Scholz, O.R.; Ernst, G. (eds): *Symbole, Systeme, Welten. Studien zur Philosophie Nelson Goodmans*, Heidelberg: Synchron, 2005, 185-201.

Kulenkampff, J. (2005b): “Kunst und Erkenntnis. Überlegungen im Anschluß an Nelson Goodman”, in: Jäger, C.; Meggle, G. (eds): *Kunst und Erkenntnis* , Paderborn: Mentis, 2005, 43-61.

von Kutschera, F. (1972): *Wissenschaftstheorie*, 2 Volumes, München: Fink, 1972.

von Kutschera, F. (1975): “Nelson Goodman: Das neue Rätsel der Induktion”, in: J. Speck (ed.): *Grundprobleme der großen Philosophen: Philosophie der Gegenwart III*, Göttingen: Vandenhoeck & Ruprecht, 1975, 51-86.

von Kutschera, F. (2005): “Kunst und Erkenntnis”, in: Ch. Jäger and G. Meggle (Hrsg.): *Kunst und Erkenntnis* , Paderborn: Mentis, 2005, 73-92.

Lammenranta, M. (1991): “Do We Make Worlds with Symbols?”, in: *Semiotica* 86 (1991), 277-287.

Lammenranta, M. (1992): “Goodman’s Semiotic Theory of Art”, in: *Canadian Journal of Philosophy* 22 (1992), 339-352.

Leonard, H.S. (1930): *Singular Terms*, Doctoral Dissertation, Widener Library, Harvard University, 1930.

Leonard, H.S., Goodman, N. (1940): “The Calculus of Individuals and Its Uses”, in: *Journal of Symbolic Logic* 5 (1940), 44-55.

Leondar, B.; Perkins, D. (eds) (1977): *The Arts and Cognition*, Baltimore: Johns Hopkins University Press, 1977.

Lewis, C.I. (1946): *An Analysis of Knowledge and Valuation*, La Salle, Il: Open Court, 1946.

Lewis, C.I. (1952): “The Given Element in Empirical Knowledge”, in: *Philosophical Review* 61 (1952), 168-175.

Lowe, V. (1990): *Alfred North Whitehead: The Man and His Work, Vol. II: 1910-1947*, Baltimore; London: Johns Hopkins University Press, 1990.

Mahrenholz, S. (1998): *Musik und Erkenntnis. Eine Studie im Ausgang von Nelson Goodmans Symboltheorie*, Stuttgart; Weimar: Metzler, 1998, ²2000.

McCormick, P.J. (ed.) (1996): *Starmaking. Realism, Anti-Realism, and Irrealism*, Cambridge, Mass.; London: MIT Press, 1996.

McDonell, N. (1979): *Sight and Symbol*, Doctoral Dissertation, Harvard University, 1979.

Mitchell, W.J.T. (1986): *Iconology: Image, Text, Ideology*, Chicago: University of Chicago Press, 1986.

Mitchell, W.J.T. (1994): *Picture Theory. Essays on Verbal and Visual Representation*, Chicago; London: University of Chicago Press, 1994.

Mitchell, W.J.T. (1999): "Vim and Rigor", in: *Art Forum*, May 1999 (URL: http://www.artforum.com/goodman.html).

Morizot, J (1996): *La Philosophie de l'Art de Nelson Goodman*, Nimes : Editions Jacqueline Chambon, 1996.

Pouivet, R. (1996): *Esthétique et logique*, Liège: Mardaga, 1996.

Putnam, H (1981): *Reason, Truth and History*, Cambridge: Cambridge University Press, 1981.

Putnam, H. (1990): *Realism with a Human Face*, Cambridge, Mass.; London: Harvard University Press, 1990.

Putnam 1992: H. Putnam: *Renewing Philosophy*, Cambridge, Mass.; London: Harvard University Press, 1992.

Quine, W.V.O. (1985): *The Time of my Life*, Cambridge, Mass.; London: MIT Press, 1985.

Ridder, L. (2002): *Mereologie*, Frankfurt am Main: Klostermann, 2002.

Rudner, R. (1967): "Nelson Goodman", in: P. Edwards (ed.), *The Encyclopedia of Philosophy*, Vol. 3, London; New York: MacMillan, 1967, 370-374.

Rudner, R.; Scheffler, I. (eds) (1972): *Logic and Art: Essays in Honor of Nelson Goodman*, Indianapolis: Bobbs-Merrill, 1972.

Sachs-Hombach, K.; Rehkämper, K. (eds) (1999): *Bildgrammatik. Interdisziplinäre Forschungen zur Syntax bildlicher Darstellungsformen*, Magdeburg: Scriptum, 1999.

Scheffler, I. (1963): *The Anatomy of Inquiry*, New York: Knopf, 1963.

Scheffler, I. (1979): *Beyond the Letter: A Philosophical Inquiry into Ambiguity, Vagueness and Metaphor in Language*, London: Routledge & Kegan Paul, 1979.

Scheffler, I. (1986): *Inquiries: Philosophical Studies of Language, Science and Learning*, Indianapolis: Hacket, 1986.

Scheffler, I. (1991): *In Praise of the Cognitive Emotions and Other Essays in the Philosophy of Education*, New York; London: Routledge, 1991.

Scheffler, I. (1997): *Symbolic Worlds: Art, Science, Language, Ritual*, Cambridge: Cambridge University Press, 1997.

Scheffler, I. (2001): "Nelson Goodman (1906-1998)", in: Martinich, A.P.; Sosa, D. (eds): *A Companion to Analytic Philosophy*, Oxford: Blackwell, 2001, 160-168.

Schilpp, P.A. (ed.) (1968): *The Philosophy of C. I. Lewis* (= *The Library of Living Philosophers*, Volume XIII), La Salle, Ill.: Open Court Press, 1968.

Scholz, H. (1951): "Nelson Goodman, The Structure of Appearance", in: *Deutsche Literaturzeitung* 72 (1951), 433-434.

Scholz, H. (1958): "Nelson Goodman, Fact, Fiction and Forecast", in: *Deutsche Literaturzeitung* 79 (1958), 297-298.

Scholz, O.R. (1984): "Fiktionale Welten, mögliche Welten und Wege der Referenz", in: Finke, P.; Schmidt, S.J. (eds): *Analytische Literaturwissenschaft*, Braunschweig: Vieweg, 1984, 70-89.

Scholz, O.R. (1988): "Catherine Z. Elgin: With Reference to Reference (Book Review)", in: *Zeitschrift für philosophische Forschung* 42 (1988), 336-340.

Scholz, O.R. (1993a): "Introduction: *Reconceptions* in Context", in: *Synthese* 95 (1993), 1-7.

Scholz, O.R. (1993b): "When is a Picture?", in: *Synthese* 95 (1993), 95-106.

Scholz, O.R. (1995): "Künste, Symbolsysteme und die Verbesserung des Verstehens", in: *Kunstforum International* 131 (1995), 348-351.

Scholz, O.R. (2000): "A Solid Sense of Syntax", in: *Erkenntnis* 52 (2000), 199-212.

Scholz, O.R. (2001): "Kunst, Erkenntnis und Verstehen", in: Kleimann, B.; Schmücker, R. (eds): *Wozu Kunst? Die Frage nach ihrer Funktion*, Darmstadt: Wiss. Buchgesellschaft, 2001, 34-48.

Scholz, O.R. (2004): *Bild, Darstellung, Zeichen. Philosophische Theorien bildlicher Darstellung*, Second, revised edition, Frankfurt am Main: Klostermann, 2004.

Scholz, O.R. (2005): "In memoriam Nelson Goodman", in: Steinbrenner, J.; Scholz, O.R.; Ernst, G. (eds): *Symbole, Systeme, Welten. Studien zur Philosophie Nelson Goodmans*, Heidelberg: Synchron, 2005, 9-32.

Schwartz, R. (1986): “I’m Going to Make You a Star”, in: *Midwest Studies in Philosophy* 11 (1986), 427-439.

Schwartz, R. (1994): *Vision: Variations on Some Berkeleian Themes*, Oxford: Blackwell, 1994.

Schwartz, R. (1997): “Goodman, Nelson”, in: Garrett, D.; Barbanell, E. (eds): *Encyclopedia of Empiricism*, Westport, Conn.: Greenwood Press, 1997, 126-128.

Schwartz, R. (1999): “In memoriam Nelson Goodman (August 7, 1906 – November 25, 1998)”, in: *Erkenntnis* 50 (1999), 7-10.

Sheffer, H. (1913): “A set of five independent postulates for Boolean algebras, with application to logical constants”, in: *Transactions of the American Mathematical Society* 14 (1913), 481-488.

Sheffer, H. (1921): *The general theory of notational relativity*, Manuscript, copies in: Sheffer Papers, Cambridge (Mass.): Harvard University, Houghton library.

Stalker, D. (ed.) (1994): *Grue! The New Riddle of Induction*, Chicago: Open Court, 1994.

Stegmüller, W. (1957): “Nelson Goodman: The Structure of Appearance & Willard Van Orman Quine: From a Logical Point of View”, in: *Philosophische Rundschau* 5 (1957), 280-292.

Stegmüller, W. (1958/59): “Conditio Irrealis, Dispositionen, Naturgesetze und Induktion (Zu N. Goodman: Fact, Fiction, Forecast)”, in: *Kant-Studien* 50 (1958/59), 363-390.

Stegmüller, W. (1969): *Probleme und Resultate der Wissenschaftstheorie und Analytische Philosophie, Band I: Wissenschaftliche Erklärung und Begründung*, Berlin; Heidelberg; New York: Springer, 1969.

Stegmüller, W. (1983): *Probleme und Resultate der Wissenschaftstheorie und Analytische Philosophie, Band I: Erklärung, Begründung, Kausalität*, Second, enlarged edition, Berlin; Heidelberg; New York: Springer, 1983.

Steinbrenner, J. (1996): *Kognitivismus in der Ästhetik*, Würzburg: Königshausen & Neumann, 1996.

Steinbrenner, J.; Winko, U. (eds) (1997): *Bilder in der Philosophie & in anderen Wissenschaften und Künsten*, Paderborn: Schöningh, 1997.

Steinbrenner, J.; Scholz, O.R.; Ernst, G. (eds) (2005): *Symbole, Systeme, Welten. Studien zur Philosophie Nelson Goodmans*, Heidelberg: Synchron, 2005.

Thürnau, D. (1994): *Gedichtete Versionen der Welt. Nelson Goodmans Semantik fiktionaler Literatur*, Paderborn: Schöningh, 1994.

Volpi, F. (ed.) (1999): *Großes Werkelexikon der Philosophie*, Vol. I., Stuttgart: Kröner, 1999.

Vuillemin, J. (1971): *La Logique et le Monde Sensible*, Paris: Flammarion, 1971.

Wang, H. (1986): *Beyond Analytic Philosophy. Doing Justice to What We Know*, Cambridge, Mass.; London: MIT Press, 1986.

White, M. (1950): "The Analytic and the Synthetic: An Untenable Dualism", in: Hook, S. (ed.): *John Dewey: Philosopher of Science and Freedom*, New York: The Dial Press, 1950, 316-330.

White, M. (1999): *A Philosopher's Story*, University Park, PA: Pennsylvania State University Press, 1999.

Whitehead, A.N. (1919): *An Enquiry concerning the Principles of Natural Knowledge*, Cambridge: Cambridge University Press, 1919.

Daniel Cohnitz

The Unity of Goodman's Thought[1]

Abstract
I argue that Goodman's philosophy should not be characterised in *opposition* to the philosophy of the logical empiricists, but is more fruitfully interpreted as a *continuation* of their philosophical programme. In particular, understanding Goodman's philosophy as a continuation of the ideal language tradition makes explicable how a radical ontological relativist could be such a staunch nominalist at the same time.

1. Introduction

The 100th anniversary of Henry Nelson Goodman's birth is certainly not the only reason why it might seem timely to have a new look at his philosophical achievements. Goodman's systematic contributions in the philosophy of logic and mathematics, as well as his work on constructional systems, are at present again at the centre of interest for contemporary systematic endeavours and enjoy a certain revival.

His anti-foundationalist attitude towards logic reappears in the work of modern logical pluralists. His metamathematical theory of part/whole-relations as codified in his "Calculus of Individuals" (more familiar under the name "mereology") is nowadays becoming a standard tool to supplement set theory (formerly known as the "calculus of classes") in ontological representation systems or even to replace set theory completely

[1] This paper or parts of it were presented at several occasions, including the Congress of the European Society for Analytic Philosophy (ECAP) in 2005, the Congress of the German Association for Analytic Philosophy (GAP) in 2006, the conference on occasion of Goodman's centenary in Munich that is documented in this volume and my *venia legendi* lecture at the University of Tartu in 2006. I'd like to thank the audiences of these presentations for valuable comments. Special thanks go to Marcus Rossberg with whom I carried out the research on this topic, and to Alan Baker for very valuable comments on an earlier draft. This paper was written with the support of grant ETF 7163 of the Estonian Science Foundation.

in nominalist programs in the philosophy of mathematics. His work on similarity relations, quality classes, and phenomenalism is inspiring and intriguing for anyone working in the cognitive sciences and interested in the ways our mind structures appearances. On the other hand, his work on the languages of art and his so-called cognitive turn in aesthetics, as well as his infamous "New Riddle of Induction" have never ceased to be of central interest in their areas, the philosophy of art and the philosophy of science, respectively.

However, the timeliness of Goodman's work will for now not be the focus of this paper. What I am going to concentrate on is a topic usually completely neglected when studying Goodman – it is the unity of Goodman's thought. How are his contributions connected with each other?[2] Is it just an idiosyncratic coincidence of interests or is there something more behind it? A philosopher specialized in logic is typically also interested in philosophy of language or the philosophy of science but very seldom is a logician in addition interested in the philosophy of art and almost never in all these areas at the same time. Even if it was not for historical curiosity on our side to find out whether or not there is a common basis in Goodman's interests and proposed solutions, being ourselves working in certain systematic sub-disciplines should make us curious. Maybe Goodman hit on a connection between areas we, so far, have ignored. Maybe it would for our own systematic endeavours be helpful to see beyond our own noses, from time to time, the way Goodman constantly did.

This enterprise, to find unity in Goodman's thought, might be doomed from the start. In what I believe to be Goodman's last publication he writes:

> There is no such thing as the philosophy of Nelson Goodman any more than there is such a thing as the finger of Nelson Goodman. There are many philosophies, but on the other hand there is no nice neat order of different complete philosophies: there are lots of ideas, conjectures about various fields. A few months ago, at the Technische Universität in Berlin, […] I gave an impromptu talk called "Untangling Nelson Goodman" and I said, 'Well, here's all this mess and can I do anything about untangling these things?'. The answer was that I couldn't do very much.

[2] This question we answer in far more detail in (Cohnitz; Rossberg 2006).

> I mean, for instance, I had dealt with certain topics many different times and in many different contexts; but it is not always clear how these relate to one another. All I could do is suggest some of the different attacks that I made on some of the problems at different times and at least note that these were not all part of a well organized scheme. They were all different attempts to deal with different aspects of the problem. And then it occurred to me that untangling this mess might entail a good deal of loss, the kind of loss you get if you try to untangle a plate of spaghetti: you would end up with some rather uninspiring strings of dough which would not have anything of the central quality of the whole meal (Goodman 1997, 16-17).

So, is there really a unity in Goodman's work? Despite Goodman's pessimism, I hope to be able to point out the coherence and unity that Goodman denies in the quote just given.

Perhaps it is possible to "untangle" Goodman a bit, or at least provide a starting point for further studies of his philosophy. Of course I cannot deal with all and everything – especially not in view of the space constraints. I will limit my project here to highlighting two central topics that could be followed through all his work, although I will not have the space then to follow them. The two main topics I will suggest should – I hope – be self-explanatory. I have chosen these two, since these are the points most often misrepresented by historians of post-war analytic philosophy and even by prominent Goodman specialists.

2. On Foundations

If one is looking for commonalities it is often advisable to start historically and look for common roots. If we start this way with Goodman we should start with his major methodological work, his PhD thesis *A Study of Qualities*, defended in 1941, which was published in revised form as *The Structure of Appearance* ten years later, in 1951. Indeed, or so I will argue, this work is the key to the unity of Goodman's thought. However *A Study of Qualities* is also the most complicated book Goodman ever wrote and few people have studied it in any detail. If it contains the key to the unity of his philosophy it is certainly well hidden. I will thus spend some time explaining what this work is about and what its major insights were and where it came from. This is necessary also because even Goodman scholars who advocate that *The Structure of Appearance is* the key to Goodman's thought have, to my mind, failed to understand what that key

really is. In particular there are two major misunderstandings concerning the motivation and origin of Goodman's *Structure of Appearance.*

1. Goodman's project in *The Structure of Appearance* is intended as an anti-foundationalist rival to Carnap's *Aufbau*, and thus his philosophy can best be characterized in opposition to logical positivism.

2. Goodman's nominalism led to the construction of the *Calculus of Individuals*, Goodman's mereological system, and its implementation in *The Structure of Appearance*, and thus his philosophy is throughout driven by nominalism.

Both these claims are common, but nevertheless wrong. The second part of each claim which generalizes to Goodman's philosophy as a whole, could of course still be true for all the rest of Goodman's philosophy, even if the first part, which is specific for *The Structure of Appearance*, should turn out to be false. In that case, *The Structure of Appearance* would simply not contain *the* key to Goodman's philosophy. I will in the next sections argue that *both* parts of *both* claims are false. Since I nevertheless think that *The Structure of Appearance* is the key to the unity of Goodman's thought I will have to locate it elsewhere – neither in his opposition to logical positivism nor in his nominalism.

So let us turn first to the first half of the first claim.

The received view:
Goodman's *The Structure of Appearance* was intended as an anti-foundationalist reconception of Carnap's *Der logische Aufbau der Welt.*

Catherine Elgin, for example, writes:

> Traditionally, phenomenalism maintains that all *a posteriori* knowledge derives from what is given in experience. If so, the goal of a phenomenalist construction is to provide the derivation. That is what Carnap attempts in the *Aufbau*. Goodman changes the subject. He believes the myth of the given cannot survive the repudiation of the scheme/content distinction. He denies that, independent of any

> prior systematization, some things are and other things are not really primitive (Elgin 2001, 681).

And so does Geoffrey Hellman:

> Turning to epistemology it must be stressed that, despite Goodman's indebtedness to Carnap and the positivists on constructionalism, *Structure* represents a sharp break away from the foundationalism that characterized the *Aufbau* [...] (Hellman, *SA* (3rd ed.), XXIII).

As is well known, Carnap's *Der logische Aufbau der Welt* (Carnap 1928) and Goodman's *The Structure of Appearance* are both studies in constitution theory. Carnap investigated the example of a world built up from primitive temporal parts of the totality of experiences of a subject (the so-called "elementary experiences" or just "erlebs") and thus faced the *problem of abstraction*: how can qualities, properties and their objects in the world be abstracted from the phenomenal experiences of ours. Goodman on the other hand investigated in his thesis the example of a system built on phenomenal qualities, so-called qualia (phenomenal colors, phenomenal sounds, etc.) and faced the *problem of concretion*: how can concrete experiences be built up from abstract particulars. Since Carnap's system is – as we have said – built on phenomenal experience, it might be conjectured that it is thus an exercise in epistemological foundationalism. Was that really Carnap's intention?

Thanks to the efforts of Friedman (1987), Richardson (1998), Pincock (2002) and others, many are nowadays convinced that phenomenalistic foundationalism was not a major (if any) concern for Rudolf Carnap and that therefore the project of *Structure of Appearance* does not differ with respect to foundationalism from the *Aufbau*. Thus we should not interpret the *Aufbau* the way Elgin did in the quote, which is considered an instance of the "received view".

Surprisingly, this "received" interpretation of the *Aufbau* is usually seen as having emerged from Quine's and Goodman's interpretations. Pincock, for example, writes:

> In "Two Dogmas of Empiricism" Quine claims that Carnap's project in *Der logische Aufbau der Welt* is empiricist and reductionist [...] In a similar way, Nelson Goodman has claimed that the *Aufbau* is phenomenalist in nature [...] Quine and Goodman see in Carnap's many references to Russell a sign that he is continuing the traditional project of British empiricism (Pincock 2002, 1-2).

Did Goodman misunderstand Carnap? If that were true, Goodman's interpretation of the *Aufbau* would be wrong, and the characterization of his philosophy as a program directed against Carnap's logical positivism almost tragic. A closer look at Goodman's reading of Carnap reveals, however, that Goodman considered Carnap's project to be in the same anti-foundationalist spirit as his own. Reading Goodman's doctoral thesis *A Study of Qualities*, from which, albeit with considerable rewriting, Goodman developed *The Structure of Appearance*, it becomes clear that Goodman was absolutely aware from the very beginning of his dealings with the *Aufbau*, that phenomenalist foundationalism was not the main theme, if indeed a theme at all, of Carnap's. Neither in *A Study of Qualities*, nor in *The Structure of Appearance* is Carnap accused of foundationalism. In fact, Goodman correctly says already in *A Study of Qualities*:

> [...] Carnap has made it clear that what we take as ground elements is a matter of choice. They are not dignified as the atomic units from which others must be built; they simply constitute one possible starting point. [...] In choosing erlebs, Carnap is plainly seeking to approximate as closely as possible what he regards the original epistemological state [...] Yet whether it does so or not is no test of the system. [...] Hence [...] argument concerning whether the elements selected are really primitive in knowledge is extraneous to the major purpose of the system (Goodman 1940, 96-98).

Here it is obvious that Goodman himself did not consider his constructionalism as an epistemological alternative to Carnap's. Insofar as epistemology does play a role in *The Structure of Appearance* or *A Study of Qualities*, this criticism was directed at the philosophy of C.I. Lewis, who was Goodman's teacher at Harvard. Lewis indeed held the view that empiricism must presuppose the incorrigibility and indubitability of what is given in experience. According to Lewis, I might need to revise, for example, that I saw a plane crossing the sky when I learn that what I mistook for a plane was Superman. However, nothing can make me revise that there was a blue and a red spot in the centre of my visual field that then led to the (false) belief that there was a plane.

A Study of Qualities, on the other hand, begins with the argument that even the simplest judgements of this sort – as the one about a blue and a red spot in the centre of my visual field – might be revised in the light of

new evidence. My judgement that I had a blue spot in the middle of my visual field a few seconds ago when I looked at a ripe apple under normal conditions might be revised when I now judge that I have a red spot in my visual field, looking at the same object under the same conditions and know that it could not have changed its colour. However, if such revisions can be made in retrospect, nothing of the "given" is indubitable or incorrigible. Judgements about qualia, in this sense, are *decrees*. Which are accepted is a matter of the overall coherence of my system of beliefs and my other qualia judgements.[3]

> The literal unverifiability of such quale-recognition is, nevertheless, in the last analysis beyond question. If I say the green presented by that grass now is the same as the green presented by it at a certain past moment, I cannot truly verify that statement because I cannot revive that past moment. The statement therefore constitutes an arbitrary and supreme decree. But a decree, simply because it is arbitrary, is not therefore necessarily haphazard. My quale-identifications are influenced; I do not feel equally inclined to identify the color presented by the grass now with the color presented by a cherry a moment ago, though such a decree if made would be equally supreme and unchallengeable on strict grounds. We are all much in the same position of absolute but sane monarchs; our pronouncements are law, but we use our heads in making them (Goodman 1940, 17; cf. *SA* (2nd ed.), 134).[4]

Also in this respect Goodman was *following* Carnap and the logical empiricists. Unlike the picture that most people (and even most historians of philosophy) have of logical positivism, the given in experience as a foundation of knowledge played no special role in it, especially not in the philosophy of Carnap.

C.I. Lewis emphasises this in his 'Logical Positivism and Pragmatism' (Lewis 1941). There he explains that the main difference between the empiricism of the pragmatists and the empiricism of the logical positivists

---

[3] See also (Goodman 1952) and (Lewis 1952).

[4] "When an identification causes us too much trouble, it becomes simpler to abandon it; and the sort of trouble we generally find ourselves in because of our decrees is the most serious possible trouble of thought: logical inconsistency." (Goodman 1940, 18) – "Normally, we have not a conflict of two decrees, but a conflict between a new decree and a whole background of accepted decrees. We could uphold the discordant newcomer, but only at the exorbitant price of reconstructing our whole picture of the past" (*SA* (2nd ed.), 135).

(especially the Carnap of *Philosophy and Logical Syntax* (Carnap 1935)) is that the latter were ready to analyse empirical knowledge fully in the so called "formal mode", as more or less coherent systems of accepted sentences, some of which are "protocols", some are sentences of mathematics and logic, some are generalisations, etc. The content of such knowledge would be explicated only in terms of deductive consequences and logical relations, but not in terms of the experiences connected with certain terms and statements.

In particular, the formal mode would not distinguish between statements such as 'This object looks red.' and 'This object is red.' Instead, logical positivism would (according to Lewis' interpretation of it) recognise both statements as one and the same "observation-sentence".

For Lewis, this sort of empiricism was not worthy of the name. After all, the experiential element did not seem to show up at all in this kind of formal analysis. Lewis claims instead that a proper empiricism must treat sentences of the form 'This looks red.' as special, indubitable statements. We might err when classifying things as *being* red, but we cannot err when it comes to recognising things as *looking* red. This is "the given" in experience, the phenomenal states we find ourselves in when making experiences. Without such an indubitable element, Lewis feared that our epistemology would necessarily collapse into a coherence theory of truth:

> [E]ither there must be some ground in experience, some factuality it directly affords, which plays an indispensable part in the validation of empirical beliefs, or what determines empirical truth is merely some logical relationship of a candidate-belief with other beliefs which have been accepted. And in the latter case any reason, apart from factualities afforded by experience, why these antecedent beliefs have been accepted remains obscure. Even passing that difficulty, this second alternative would seem to be merely a revival of the coherence theory of truth, whose defects have long been patent (Lewis 1952, 112-113).

Thus, besides the chicken and egg problem of why we should ever actually come to accept any system, we would not know how to choose between the many equally coherent systems that are all logically possible.

Goodman was ready to bite that bullet when throwing away the indubitable given. Lewis, the major advocate of pragmatism, commented on this move by Goodman that his "proposal is, I fear, a little more pragmatic than I dare to be" (Lewis 1952, 118).

Indeed, Goodman's early and later philosophy was anti-foundationalist. This is truly a characteristic of his work on induction, metaphysics, logic and even the languages of art. It should, however, not be interpreted as a counter program to logical positivism. What Goodman did – in all these areas – was to continue Carnap's program and enlarge it to cover new areas. This is obvious if we consider Goodman's relativism and irrealism. It is also apparent, when we think about his pluralism in logic and his insistence that there are more cognitively valuable representation systems than just the sciences, namely the languages of art.

His anti-foundationalism therefore was more than just a restatement that there is no bed rock for knowledge – as was also argued by Karl Popper and Otto Neurath, but also that there are no fundamental ontological objects, that there are no fundamental logical principles, and that there are no privileged representation systems. All of these echo Rudolf Carnap's famous principles of tolerance. Tolerance with regards to ontology, to logical principles and to representation systems in general.

But wait, you might object, if anti-foundationalism in this broad sense, in epistemology, logic and ontology does best characterize the unity in his philosophy and places him in the tradition of the logical positivists, as you claim, how come that Goodman is one of the most famous nominalists? After all, nominalism is usually seen as an ontological claim and – moreover – a not very permissive one. How can anybody claim of himself that he is an ontological pluralist and at the same time be a nominalist? Here we must turn to a second misunderstanding of Goodman's philosophy.

3. On Constructions

The second misunderstanding concerns Goodman's nominalism. First, it has to be made clear that Goodman's nominalism is not concerned with the rejection of abstracta or universals, but is precisely *the rejection of the use of classes in constructional systems*. Second, it is a common misunderstanding that the technique to actually develop constructional systems without any classes – namely Goodman's (or rather Goodman's adaptation of Leonard's) Calculus of Individuals – was designed with that

purpose in mind.[5] Goodman simply was no nominalist when he began his work on the Calculus. He was not a nominalist in any sense of the word.

3.1 The Calculus of Classes and the Calculus of Individuals

Carnap's constructional systems, as well as Russell and Whitehead's *Principia Mathematica*, included in their logical apparatus the formation of classes, classes of classes, and so forth, of whatever entities function as individuals, and it is in terms of such classes that the various non-primitive terms are defined.

Indeed, in *The Structure of Appearance* Goodman eschews the use of the class concept as "platonistic" and adheres instead to the version of nominalism that he had at that time developed with Quine. However, the road towards this nominalism is not as straight as it might seem.

Goodman collaborated already as an undergraduate student with Henry Leonard, who was at that time a PhD student at Harvard. Henry Leonard's PhD thesis discusses issues related to Whitehead and Russell's *Principia Mathematica*. In the 'Preface' of *Singular Terms* (defended in 1930) Leonard thanks his supervisor Alfred North Whitehead and "Mr. H. N. Goodman", with whom he "discussed together nearly every point developed in this thesis" (Leonard 1930, v). Leonard summarizes the topic of this thesis as follows:

> The Calculus of Classes achieved success through its disregard of the traditional abstract terms. The object of this thesis is to show that this disregard was justified, because abstract terms are singular and not general, that on the other hand, symbolic logic should and can provide a treatment of these terms in a calculus of singular terms and that with this development, it is possible to see that symbolic logic has not thrown over intension, but only its traditional association with abstract terms (Leonard 1930, p. 1, abstract to the PhD thesis).

From these remarks alone it is difficult to see what significance *Singular Terms* might have for Goodman's work on *The Structure of Appearance*, however, I believe that Leonard's thesis was probably the initial

[5] That nominalism is a mark of all of Goodman's philosophy is, for example, claimed in (Scholz 2005).

motivation for Goodman to start the project carried out in *A Study of Qualities*, presumably long before he had even heard of Carnap's project.[6]

This becomes clear when we consider what at that time was understood to be an abstract term. "Abstract universal terms" in traditional logic were names of attributes (properties, such as 'redness') and relations. They were taken to be different from concrete general terms (such as 'red thing') on metaphysical grounds. Leonard's conviction that logic should not settle metaphysical disputes one way or the other led him to favour a treatment of names of attributes as names of additional individuals. *Principia Mathematica*, being primarily interested in the reconstruction of mathematics, could not provide the means for doing this:

> Ideally our symbolic logic should offer us a calculus, which analyses immediately the structure of the propositions that employ abstract terms. And until it does that, our symbolic logic is incomplete. [...] The object of the present paper is to seek a solution to this problem. In general, our solution consists in holding that the Calculus of Classes is the calculus of general terms, that abstract terms are rightfully excluded from the type called "general", but that they belong to the type called "singular." Thus, instead of omitting them entirely, we add a new chapter to symbolic logic, and it is this addition which characterizes our whole position (Leonard 1930, 7-8).

In other words, Leonard did not intend to *replace* any part of the Calculus of Classes, but to add new resources to it to deal with problems the original calculus was not designed for. Leonard introduces the calculus of singular terms, which turns out to be a mereological system. In this system, variables range over individuals that can be summed together and form new units, just as the later Calculus of Individuals would.

---

[6] Quine (as he says in his autobiography, 1985, p. 86) first heard of Carnap in 1932 from John Cooley, one of his fellow graduate students. In the same year Feigl was at Harvard and talked to Quine about the Vienna Circle. Presumably Goodman's knowledge of Carnap was not before this year either (cf. also Quine 1985, 122, where he describes how he and Goodman discovered that Leonard and Goodman's project had resemblance to the *Aufbau* and how Quine noted that the Calculus of Individuals was essentially Lesniewski's mereology). Scholz (2003, 17, footnote 32) conjectures that Goodman might have learned about the *Aufbau* from Leonard who was in Munich in 1929. This seems unlikely though, since Leonard does not refer to Carnap's work in (Leonard 1930) although it is very closely related to his chapter V.

Applications considered in Leonard's examples are volumes and qualities. The calculus deals with part-of and crossing relations, etc.[7]

In part V of his dissertation, Leonard discusses different possible interpretations of his calculus. His interesting result is that the calculus whose variables should be interpreted as ranging over parts of the world can serve just as well for a *realistic* conception of the world as for a *nominalistic* conception. In a realistic conception qualities would be treated as basic units and all concrete objects or other units defined as complexes built up out of these. Leonard also considers realistic systems built on phenomenal qualia rather than qualities.

> All units are equally real, but in order to introduce system in our view of the world, we must take certain ones as basic and describe others in terms of these. In the suggestions which we have just outlined, we have taken quality units as basic, where under "quality units" we include units of space and time. In terms of these, we describe units of other types. On this view quality units are real parts of our world, the basic units in a world view (Leonard 1930, 238).

Nominalistic conceptions, on the other hand, would take concrete particulars as basic and construct qualia or qualities out of these. For Leonard, just as for Carnap and Goodman, what you start with seems completely arbitrary from a logical point of view. Eventually, Leonard comes to his own principle of tolerance, claiming that for different purposes ("every day activity", "science", "art", Leonard 1930, 242-244) different systems might be adequate and that none of them can claim to be more real than any other. Logic is considered to be the study of what is common to all "unitations" [ways to build up a representation of the world from a given basis], but is neither concerned with developing a new unitation nor with singling out one such unitation as privileged.

3.2 The Platonism and Realism of A Study of Qualities

Goodman's dissertation project might in its first conception be stated as a further elaboration of Leonard's dissertation. Just as *Singular Terms*, *A Study of Qualities* does not *replace* the Calculus of Classes in the construction, but *adds* the mereological system to it. This is why in *A Study of Qualities* nominalism is not an issue at all. In *A Study of Qualities*,

[7] Cf. Marcus Rossberg's contribution in this volume.

Carnap's system is called "nominalistic", for the reason that Carnap started with erlebs (*Elementarerlebnissen*), concrete particulars. The system of *A Study of Qualities*, on the other hand, is called "realistic" because its basis is comprised of phenomenal qualia. The method of construction in both systems, however, is *platonistic*, for Goodman and Carnap both use the Calculus of Classes (although Goodman makes less use of it).

This latter point changes in *The Structure of Appearance*. Set theory, the Calculus of Classes, is completely replaced by the mereological system, the Calculus of Individuals. Now Goodman considers his system to be "nominalistic", this time with respect to the method of construction, and Carnap's system to be "platonistic". The old division, based on the different choice of bases, is now expressed as the difference between "particularism" (Carnap) and "realism" (Goodman). Both systems are, however, phenomenalistic rather than physicalistic (a basis that Carnap later was flirting with, cf. Carnap 1931).

The main impact of the Calculus of Individuals since its first implementation in *A Study of Qualities* seems to be the fact that it made it possible to circumvent a certain constructional difficulty of Carnap's *Aufbau* known as the *difficulty of imperfect community*. The problem was that Carnap's method of abstracting qualities from elementary experiences – the so-called "quasi-analysis" – results under unfavourable circumstances in certain quality classes although not all members in these classes properly share a common quality.

As Goodman and Leonard already showed in their 1940 paper 'The Calculus of Individuals and Its Uses', this problem does not arise if the resources of the Calculus of Individuals are used to construct qualities as quality wholes. It was only much later, when Goodman reworked *A Study of Qualities* into *The Structure of Appearance*, that Goodman made nominalism the prime motivation for his use of the calculus of individuals.

To summarize: Goodman's nominalism is thus no aspect of his philosophy that would unify all his work, although, of course, you will find Goodman mentioning his nominalism also in his other writings, as for example in *Languages of Art*. Nominalism is rather a symptom of another really unifying aspect of his work, which also explains how a pluralist – like Goodman – can at the same time defend such a prima facie non-pluralistic idea as nominalism.

4. On Explication

I emphasized already that Goodman continued the tradition of the logical empiricists. He did so not only with respect to their anti-foundationalism, but also in the way that he conceived of the aim and purpose of philosophy as such. Whereas the British branch of analytic philosophy was rooted in the ordinary language philosophy of the later Wittgenstein, the American branch, founded by the Viennese emigrants, and by Hans Reichenbach, Charles Morris, Willard Van Orman Quine, Morton White, Goodman and Alfred Tarski (to name but a few), was rooted in the so called *ideal language philosophy* that dates back to Gottlob Frege (and in a broader perspective to the work of Gottfried Wilhem Leibniz).

In contrast to the ordinary language philosophers, the logical empiricists did not think that ordinary language is basically in good order and that just a better understanding of it would reveal the dissolutions to our philosophical (pseudo-)problems. To the contrary, the positivists believed ordinary language to be inexact and misleading and were engaged in constructing a better substitute. This made necessary a systematic re-construction of discourse in an artificial ideal language, the language of formal logic.

Goodman stood in this tradition and hence in opposition to British ordinary language philosophy. One of the main differences between the two types of analytic philosophy was their attitude towards systematicity in philosophy. As is well-known, the later Wittgenstein abjured systematicity in philosophy completely. As Goodman interpreted him, Wittgenstein regarded philosophical problems as diseases spread via natural language. The philosopher is accordingly a therapist who in single cases of confusion comes to help with a cure that is specific to the case at hand. The nature of this ambulance-model of philosophy allows philosophers to stop doing philosophy whenever they please. Since they are not interested in constructing a systematic theory, to stop philosophising will not prevent them from reaching a final goal. Their aims are all temporary, to help the poor souls who find themselves trapped in a puzzle of natural language.

Goodman rejected this view. First of all, philosophical puzzles do not arise for the man in the street who just tries to make a living. Philosophical puzzles arise for philosophers and they arise only because philosophers

have set up standards of understanding, which might or might not be met by a literal understanding of natural language.

> [T]he philosopher's puzzlement about language is always a puzzlement about interpreting ordinary statements in a philosophical way. The puzzlement or confusion is a function not only of the language but of our standards or sense of philosophical acceptability.
> Wittgenstein triumphantly exclaims that his conception of philosophy allows him to stop doing philosophy whenever he pleases. But he can stop doing philosophy, or at least stop needing to do philosophy, only when all philosophical puzzlement and confusion are resolved (*PP*, 43-44).

Thus philosophers cannot stop philosophising as they please; they can stop philosophising only if either all philosophical confusions are resolved by a reinterpretation of ordinary language that conforms to the standards of philosophical acceptability or if they relax their requirements enough to take language as it is. Philosophers are not therapists of ordinary people. They are the ones having the problem of understanding and will be able to stop doing philosophy no sooner than such understanding is achieved.

There is also a second major difference between Goodman's and Wittgenstein's understanding of the aim of philosophy. For the later Wittgenstein, there are no real philosophical problems, merely "puzzles about language". What seems to be a philosophical problem inevitably must turn out to be a misunderstanding of what our words mean, or resulting from a misuse of those words. Philosophy, therefore, cannot be revisionary. The correct use of the words in natural language is the standard that needs to be respected, it cannot be that a philosopher uncovers that this use is not in good order. Goodman, on the other hand, allows for surprises. In the course of doing philosophy, so-called "common sense" (what Goodman calls "the repository of ancient error") and the pre-systematic use of words can – and in interesting cases will – be declared defunct. Philosophy leads to revisions, and discovers errors in, for example, ordinary language categorisations. Goodman's philosophy is critical in that sense in a way that Wittgenstein or ordinary language philosophy is not. For Goodman there are real philosophical problems that can and need to be solved, not just verbal confusion and language puzzles.

Goodman's nominalism, as well as much of the work in *Languages of Art*, his solution of the problem of induction and projection and especially

his work of *Reconceptions in Philosophy* is unified under the umbrella of the ideal language-conception of rational reconstruction and explication.

In this conception, problematic notions that led to philosophical problems were to be replaced by new constructions in an ideal language that should not again be polluted by incomprehensible notions, as far as possible. Such replacement was called "explication"; these replacements did not conserve the meaning of the old notions.

There is a long list of notions that Goodman attempted to explicate in this way. The connection to his nominalism is that in explicating the problematic notions in a clarified language, Goodman could not accept explications in the language of set theory. A language that contained the notion of sets wasn't comprehensible to him, like a language that contained any other intensional notions. Nominalism is thus a symptom of this special approach to philosophy. It is upheld because Goodman's very own standards of clarity did not allow the notion of sets. It is, however, not upheld because the corresponding objects were supposed to be less real than others. Nominalism is not a metaphysical thesis in Goodman's work, but a constraint on acceptable explications.

We can thus find two unifying features of Goodman's work. One concerns the philosophical content of Goodman's philosophy. A thorough anti-foundationalism, developed as a radicalization of views that are rooted in logical positivism. The second feature is methodological in nature. Since philosophy is seen as an activity that aims at understanding and elucidation, it gets done by explicating problematic notions in a clarified language.

Goodman's anti-foundationalism keeps together the different fields he was working in. His view on explication, on the other hand, as the method by which philosophy gets done, keeps together his interest in philosophical tools such as logic and mereology and explains his nominalism.

References

Carnap, Rudolf (1928): *Der logische Aufbau der Welt. Scheinprobleme in der Philosophie*, Hamburg: Meiner, 1961.

Carnap, Rudolf (1931), "Die physikalische Sprache als Universalsprache der Wissenschaft", *Erkenntnis* II (1931), 432-465.

Carnap, Rudolf (1935), *Philosophy and Logical Syntax*, London: Kegan Paul 1935.

Cohnitz, Daniel; Rossberg, Marcus (2006): *Nelson Goodman*, Chesham: Acumen 2006.

Elgin, Catherine (2001): "The Legacy of Nelson Goodman", *Philosophy and Phenomenological Research* 62 (2001), 679-690.

Friedman, Michael (1987): "Carnap's *Aufbau* Reconsidered", *Noûs* 21 (1987), 521-545.

Goodman, Nelson (1940): *A Study of Qualities* [1940], Ph.D. Thesis, New York: Garland, 1990.

Goodman, Nelson; Leonard, Henry (1940): "The Calculus of Individuals and Its Uses", *Journal of Symbolic Logic* 5 (1940), 45-55.

Goodman, Nelson (1952): "Sense and Certainty", *Philosophical Review* 61, 160-167.

Goodman, Nelson (1997): "Some Reflections on my Philosophy", *Philosophia Scientiae* 2, Actes du Colloque Nelson Goodman 1997, "Manières de faire les mondes", 15-20.

Leonard, Henry (1930): *Singular Terms*, PhD dissertation thesis, Harvard University, 1930.

Lewis, C.I. (1941): "Logical Positivism and Pragmatism", reprinted in Goheen, J.D.; Mothershead, J. L. (eds): *Collected Papers of Clarence Irving Lewis*, Stanford: Stanford University Press 1970, 92-112.

Lewis, C.I. (1952): "The Given Element in Empirical Knowledge", reprinted in Elgin, C.: *The Philosophy of Nelson Goodman Vol. 1: Nominalism, Constructivism, and Relativism*, New York: Garland 1997, 112-119.

Pincock, Christopher (2002): "Russell's Influence on Carnap's Aufbau", *Synthese* 131 (2002), 1-37.

Quine, Willard Van Orman (1985): *The Time of My Life: An Autobiography*, Cambridge, MA: MIT Press, 1985.

Richardson, Alan (1998): *Carnap's Construction of the World: The Aufbau and the Emergence of Logical Empiricism*, Cambridge: Cambridge University Press, 1998.

Scholz, Oliver (2005): “In Memoriam Nelson Goodman 1906 - 1998”, in Steinbrenner, J.; Scholz, O.R.; Ernst, G. (eds): *Symbole, Systeme, Welten: Studien zur Philosophie Nelson Goodmans*, Heidelberg: Synchron 2005, 9-32.

Marcus Rossberg

Leonard, Goodman, and the Development of the *Calculus of Individuals*

Abstract
This paper investigates the relation of the Calculus of Individuals presented by Henry S. Leonard and Nelson Goodman in their joint paper, and an earlier version of it, the so-called Calculus of Singular Terms, introduced by Leonard in his Ph.D. dissertation thesis *Singular Terms*. The latter calculus is shown to be a proper subsystem of the former. Further, Leonard's projected extension of his system is described, and the definition of a non-extensional part-relation in his system is proposed. The final section discusses to what extent Goodman might have contributed to the formulation of the Calculus of Individuals.

1. The Calculus of Individuals

In 1936, Henry S. Leonard and Nelson Goodman presented a joint paper at the meeting of the Association for Symbolic Logic which was held at the meeting of the Eastern Division of the American Philosophical Association in Cambridge, Massachusetts. Eleven years later, they published an elaborated version of this paper under the title "The Calculus of Individuals and its Uses" (Leonard; Goodman 1940). The calculus they introduce in this paper is today usually taken as a basis for the study and use of formal part-whole relations (often called "mereology") in analytic metaphysics, sometimes mediated by Goodman's presentation of the calculus in his *The Structure of Appearance* (*SA*).

Goodman used the Calculus of Individuals in his Ph.D. dissertation thesis *A Study of Qualities* of 1940 (Goodman 1940), which eventually became *The Structure of Appearance*. As in his joint paper with Leonard, he used the calculus as an addition to set theory to solve a problem known as the *difficulty of imperfect community* in Rudolf Carnap's *Aufbau* (Carnap 1928). Only in *Structure* Goodman abandoned set theory and

presented a nominalistic construction that used only the Calculus of Individuals.[1]

The focus of this paper, however, will be an investigation of the system that Leonard presents in his Ph.D. dissertation thesis *Singular Terms* (Leonard 1930), which is the first occurrence of a system akin to the Calculus of Individuals in this line of development. As is well known, Stanisław Leśniewski developed his mereology, and also coined the term, well before this (1916, 1927, 1929), but the work of Leonard and Goodman is independent of Leśniewski's, until W.V. Quine recognised the similarities between the systems.[2] In their joint paper, Leonard and Goodman acknowledge

> The calculus of individuals we shall employ is formally indistinguishable from the general theory of manifolds developed by Leśniewski (Leonard; Goodman 1940, 46).

They go on to justify their project by pointing out that Leśniewski's system is "rather inaccessible, lacks many useful definitions, and is set forth in the language of an unfamiliar logical doctrine and in words rather than symbols" (*ibid.*).

Indeed, Leśniewski's system, Tarski's axiomatisations of it (1929, 1937), the Calculus of Individuals of Leonard and Goodman's paper, and Goodman's version of it that he presents in *Structure*, all turn out to be equivalent.[3] We will now turn to the question, how Leonard's Calculus of Singular Terms fits in.

2. Singular Terms

In his Ph.D. dissertation thesis *Singular Terms* (Leonard 1930), Leonard introduced the Calculus of Singular Terms which is an early version of the Calculus of Individuals. Alfred Whitehead supervised Leonard on his Ph.D. project which was conceived of as an extension to Whitehead and Russell's *Principia Mathematica* (Russell; Whitehead 1910-1913). The

[1] See (Cohnitz; Rossberg 2006, 121-139), and (Cohnitz 2008, §3.2) for a discussion.

[2] (Quine 1985, 122); see the quotation in §4 below.

[3] Given some straightforward assumptions; for details see Lothar Ridder's *Mereologie* (Ridder 2002). See also Rolf Eberle's (1967) and (1970).

formal sections of his thesis are slotted into those of *Principia* and bear the numbers *16 to *18, and thus sit in the space between the introduction of definite descriptions, *14, and the theory of classes, *20 and following.

Leonard's aim was to bring clarity into the confusion that surrounded the distinction of universal and particular and the conflation of this distinction with that of concreta and abstracta:[4]

> The distinction of particular and universal has suffered from an ambiguity. It has covered (1) the distinction between entities with and entities without spatio-temporality, and (2) the distinction between entities which are complex in the sense of a realistic interpretation and those which are simple. On the last basis, the distinction is relative; on the first, arbitrary, and what is more, the two are not coextensive (Leonard 1930, abstract, 4).

He aims to give a precise and formalised treatment of the distinction between universal and particular, the beginning of which is the Calculus of Singular terms, since he deems the apparatus of the theory of classes developed in *Principia Mathematica* unfit for this purpose.

> The calculus of classes has, we said, offered us no treatment of the traditional abstract and universal terms, "redness," "coloredness," "weight," and the like, but only of what have been called in traditional logic concrete general terms, "red things," "colored things," "stones," and the like. And yet we do use such terms, and not only as predicates but as subjects of predication as well. We say, for example, "Scarlet is red," "Scarlet is a color." It is not enough that we can always present a materially equivalent proposition, such as "Scarlet things are red things," or "The class of scarlet things is a member of the class of classes, 'color.'" Ideally our symbolic logic should offer us a calculus which analyses immediately the structure of these propositions that employ abstract terms. And until it does that, our symbolic logic is incomplete. For if it cannot do that, it cannot symbolise the equivalence of these two propositions, and the claim that they are equivalent is a claim made outside the symbolic system, made by the interpretor rather than by the system builder.
>
> The object of the present paper is to seek a solution to this problem. In general, our solution consists in holding that the calculus of classes is the calculus of general terms, that abstract terms are rightfully excluded from the type called "general," but that they belong to the type called "singular." Thus, instead of omitting them

[4] Since Leonard's thesis is unfortunately unpublished, I will quote extensively from it, in order to make the relevant passages available.

> entirely, we add a new chapter to symbolic logic, and it is this addition which characterizes our whole position (Leonard 1930, 6–8).

Leonard thus proceeds to develop the Calculus of Singular Terms as an addition to the system of *Principia Mathematica* and its calculus of classes.

> The calculus of classes achieved success through its disregard of the traditional abstract terms. The object of this thesis is to show that this disregard was justified, because abstract terms are singular and not general, that on the other hand, symbolic logic should and can provide a treatment of these terms in a calculus of singular terms and that with this development, it is possible to see that symbolic logic has not thrown over intension, but only its traditional association with abstract terms (Leonard 1930, abstract, 1).

2.1 The Calculus of Singular Terms

The formal system of the Calculus of Singular Terms is based on a term-forming operation in which we nowadays might recognise the binary mereological sum. Leonard prefaced his axiom system with the remark:

> For the development of this chapter, we require one primitive idea, which we introduce as a descriptive function and denote by "$x+y$." By this we denote a symbol of the form "$(\iota x)(\phi x)$," where the specific determinant function is not indicated. [...] By "$x+y$" we mean to describe that individual which arises from the most general togetherness of any two other individuals (Leonard 1930, 187).

In other words, '+' is taken as primitive and treated as a definite description. Next, the defined notions are introduced (Leonard 1930, 190) which are the relations '<' for 'is part of', '○' for 'overlap', and two further definite descriptions, namely, concatenation of terms for 'product', and '-' for 'difference'.[5]

Definitions

*16.01 $x<y =_{df} x+y=y$

*16.02 $x\circ y =_{df} \exists z(z<x \wedge z<y)$

*16.03 $xy =_{df} (\iota z)(z<x \wedge z<y \wedge \forall t[(t<x \wedge t<y) \supset t<z])$

*16.04 $x\text{-}y =_{df} (\iota z)(\neg z\circ y \wedge (x\circ y \supset z+xy=x) \wedge (\neg x\circ y \supset z=x))$

[5] Leonard gives the formalism in the notation of *Principia Mathematica* (Russell; Whitehead 1910-1913); for ease of legibility, the notation is modernised in the exposition below. Note that '(ιx)' is the definite description operator.

Next the postulates, or axioms, that govern '+' are listed (Leonard 1930, 191):

Postulates

*16.1 $x+y = y+x$ (PM *22.57)
*16.12 $E!x+x \supset x+x = x$ (PM *22.56)
*16.14 $[E!x+y \wedge E!y+z \wedge E!(x+y)+z \wedge E!x+(y+z)] \supset (x+y)+z = x+(y+z)$ (PM *22.7)
*16.16 $x \circ y \supset E!xy$ (PM *22.36)
*16.17 $x \circ (y+z) \supset (x \circ y \vee x \circ z)$
*16.18 $x < (y+z) \supset (x<y \vee x<z \vee x<xy+xz)$

The numbers in the right hand column (PM *22.57, etc.) denote the corresponding propositions of the calculus of classes of *Principia Mathematica*. Leonard gives these references throughout his presentation. Note that 'E!' is the existence predicate of *Principia* (Russell; Whitehead 1910-1913), which is defined as

PM *14.02 $E!(\iota x)(\phi x) =_{df} \exists b \forall x(\phi x \equiv x=b)$

Sum, product and difference are definite descriptions, as noted above. As such they are guaranteed to refer to at most one object, but it is not thereby guaranteed that they pick out an object at all. If we want to read '+', etc., as functions, we thus have to take them to be *partial* functions. Indeed, the product must not be a total function; the existence of the product of some objects x and y is only guaranteed, if x and y overlap (*16.16).

The first theorem that Leonard proves is that, in fact, any two objects have a sum:

*16.2 $E!x+y$

As Leonard explains (Leonard 1930, 192), *16.2 follows immediately from *16.1 by the PM proposition

PM *14.21 $\Psi(\iota x)(\phi x) \supset E!(\iota x)(\phi x)$

The first postulate, *16.1, thus packs two into one: the commutativity of the sum and the existence of arbitrary sums. Leonard notes that *16.1

could have been split into two postulates: one in the form of theorem *16.2, and the other in a conditionalised form analogous to *16.12 and *16.14. With the help of *16.2, the antecedents of *16.12 and *16.14 can now be discharged.

Leonard proceeds to prove a large number of theorems which are familiar to us today as theorems of classical mereology. Here is a (comparatively) small selection:

Some theorems

*16.21 $\exists z(x+y) = z$
*16.213 $\mathrm{E!}(x+y)+z$
*16.214 $(x+y)+z = x+(y+z)$
*16.221 $(x+y) = z \supset x<z$
*16.25 $x<x$
*16.26 $x<(x+y) \wedge y<(x+y)$
*16.271 $x = y \supset z<x \equiv z<y$
*16.274 $x = y \equiv (x<y \wedge y<x)$
*16.28 $(x<z \wedge y<z) \equiv \mathrm{x+y<z}$
*16.281 $(x<y \wedge y<z) \supset x<z$
*16.3 $x \circ y \equiv \exists z(z<x \wedge z<y)$
*16.31 $(x<y \wedge x<z) \supset y \circ z$
*16.32 $x \circ y \equiv y \circ x$
*16.321 $x<y \supset (x \circ y \wedge y \circ x)$
*16.323 $x \circ x$

2.2 Universals and the crossing relation

Next, Leonard sets out to incorporate universals into the system. He introduces the crossing relation to this end.

> By "$x\mathrm{k}y$" (read "x crosses y"), we denote a specific propositional form, corresponding in meaning to what "y is x" would often mean. As we shall point out in the next two chapters, this symbol is amenable to a variety of interpretations and in this variety shows that logic itself does not dictate a philosophic world-view. [...] [On a realist] interpretation, "$x\mathrm{k}y$" means that "x" stands to "y" in the relation of an ingressing quality to the object it constitutes, or partially constitutes, that "x" qualifies, or is a quality of, "y." Thus, for example, we should say that whiteness crosses this sheet of paper (Leonard 1930, 214–215).

Using ‘crossing’, other notions become definable, for example, ‘uniform crossing’ (‘k_u’), ‘portion’ (‘′’), and ‘occuring together’ of qualities (‘w’):

*17 .01 $x\mathrm{k}_u y =_{df} \forall z(z<y \supset x\mathrm{k}z)$

*17.201 $x'y =_{df} (\iota x)(x\mathrm{k}_u z \wedge z<y \wedge \forall t[(x\mathrm{k}_u t \wedge t<y) \supset t<z])$

*17.301 $x\mathrm{w}y =_{df} \neg x\circ y \wedge \exists z(x\mathrm{k}z \wedge y\mathrm{k}z \wedge x'z\circ y'z)$

Leonard does not present an axiom system of the crossing relation and the related notions. He merely lists some desiderata for such an axiomatisation and remarks:

> It is our ambition to work out something more definite for this science in the next few years. We offer the following as a suggestion of the direction in which we should seek a solution (Leonard 1930, 214).

Alas, Leonard never returned to this project in print. Amongst the desiderata he lists are (Leonard 1930, 217):

*17.1 $(x\mathrm{k}_u y \wedge y\mathrm{k}_u z) \supset x\mathrm{k}_u z$

*17.11 $(x\mathrm{k}_u y \wedge y\mathrm{k}z) \supset x\mathrm{k}z$

*17.12 $(x\mathrm{k}_u y \wedge z\mathrm{k}_u t) \supset (x+z)\ \mathrm{k}_u\ (y+t)$

*17.13 $(x\mathrm{k}y \wedge y<z) \supset x\mathrm{k}z$

*17.14 $(x\mathrm{k}y \wedge x<z) \supset z\mathrm{k}y$

*17.142 $x\mathrm{k}_u y \supset x\mathrm{k}y$

Leonard remarks that “which of these must be taken as primitive propositions [i.e., postulates, or axioms], we do not as yet know” (Leonard 1930, 217). He also observes, however, that using *17.13 and *17.14, together with *16.2, the existence of arbitrary sums, it is possible to derive the undesirable consequence

*17.141 $x\mathrm{k}y \supset (x+y)\ \mathrm{k}\ (x+y)$

which says that any entity is a quality of itself. Leonard discusses two responses, both of which he finds unsatisfactory. The first is to bite the bullet on the result; the second, to restrict summation to individuals that are in the same category.

> The difficulty is that we do not know what a category is. It would seem to us that logic cannot know more about categories than what can be formulated in such abstract rules as these which we are giving. As a result, I do not at present see just

> how the limitations for the existence of sums can be formulated. (Leonard 1930, 218)

A further desideratum has even more severe problems:

*17.15 $x\mathrm{k}y \supset \neg x \circ y$

"is inconsistent with those which we have already given, for '$x+y \circ x+y$' is true by *16.33, and is false by *17.141 [and *17].15. But by denying *16.2, we could, by this theorem, deny *17.141" (Leonard 1930, 218).

Being a portion of something is not to be identified with being a part of this object. Rather, x is a part of y if and only if any quality of x is a quality of y:

*17.16 $x<y \equiv \forall z(z\mathrm{k}x \supset z\mathrm{k}y)$

which "says that the whole has all the qualities of the part and that if one thing had all the qualities of another, then it must be the whole of which that other is a part" (Leonard 1930, 219). Desiderata for 'portion' ('´') and 'occurring together' ('w') follow, as does a "more or less speculative" definition of "that unit which two units occurring together constitute in virtue of their occurrence together", '$\underline{\mathrm{w}}$', (think of product and overlap) (Leonard 1930, 221):

*17.401 $x\underline{\mathrm{w}}y =_{df} (\iota z)(\forall t[t\mathrm{k}z \supset (t<x \vee x<t \vee t\mathrm{k}x \vee x\mathrm{k}t \vee t<y \vee t<y \vee t\mathrm{k}y \vee y\mathrm{k}t])$

*18 is developed in an even more rudimentary way. It gestures at defining 'complex' ('ex', as in 'complex') and makes use of the calculus of classes (and should thus rather appear after *24 in the system of *Principia Mathematica*). To define 'ex', the auxiliary definition which extends the definition of '$\underline{\mathrm{w}}$' given above to the case of classes, is introduced:

*18.01 $x\underline{\mathrm{w}}\alpha =_{df} (\iota z)(\forall y[y \in \alpha \supset \forall t((x\mathrm{w}y \wedge (x\underline{\mathrm{w}}y)\mathrm{k}z \wedge (x\underline{\mathrm{w}}y)\mathrm{k}t) \supset z<t)])$

*18.02 $x \mathrm{ex} =_{df} x\underline{\mathrm{w}}\hat{}y(y\mathrm{w}x)$

Note that 'α' is a class variable and that '$\hat{}x$' is the *Principia* notation for class abstraction.

The complexity of the planned extension of Leonard's project should be sufficiently clear at this point.

2.3 Pluralism in Singular Terms

As we could already see in the first quote given in the section, Leonard's philosophy is pluralistic in a way that will be characteristic for Goodman's philosophy too (cf. Cohnitz 2009 and chapter 8 of Cohnitz; Rossberg 2006). In the final chapter of *Singular Terms*, this feature of his views becomes even more pronounced:

> All units are equally real, but in order to introduce system in our view of the world, we must take certain ones as basic and describe others in terms of these. In the suggestions which we have just outlined, we have taken quality units as basic, where under "quality units" we include units of space and time. In terms of these, we describe units of other types. On this view quality units are real parts of our world, the basic units in a world view (Leonard 1930, 238).

Specifically, this pluralism pertains to the nominalism-realism debate:

> Realism says that the quality crosses the object, nominalism that the object crosses the quality.
>
> Let us consider for a moment some uniform object, say a white billiard ball. It is uniformly white, uniform in density and in hardness. There are, of course, a multitude of other qualities but let us restrict ourselves to these three. We may say that the billiard ball is these qualities occurring with each other and with a certain space-time unit. This would be the realistic view […]. On the nominalistic view we should reverse the account. We should say that whiteness is the complex of all white objects occurring together. On the realistic view objects do not occur together, but form a system of units which have relations of part and whole, discreteness, and the like. One the nominalistic view, objects do occur together, and in their occurrence together constitute quality units, which are thereby derivative and relatively complex (Leonard 1930, 238-239).

Realism and nominalism, thus, appear merely as different ways to interpret the formal system that characterises the relation between universals and particulars (Leonard 1930, 241 & 247). But Leonard is prepared to adopt an even more far-reaching position, clearly foreshadowing Goodman's pluralistic irrealism of *Ways of Worldmaking* (*WW*). In section D of the penultimate chapter of *Singular Terms* (242–247) Leonard describes how choosing different units as basis of the system he presents can serve different purposes. He mentions different special sciences that might require different such unitations, and also discusses an application to the arts—a further analogy to Goodman. In sum, he asks,

> why should every function of life be suited by the same mode of unitation and why should even every branch of knowledge and culture find the same mode most satisfactory? Why, finally, should philosophy be concerned to develop a new and distinctive unitation? Is it not possible that the end of philosophy would be the careful correlation of those systems of unitation which other disciplines and other cultures spin out or exemplify? (Leonard 1930, 245)

This is, however, not the place to investigate Leonard's pluralism any further. Instead, let us venture into the formal comparison of the Calculus of Singular Terms with the Calculus of Individuals.

3. The Calculus of Individuals *vs* the Calculus of Singular Terms

For the formal comparison of the two calculi, the Calculus of Individuals too will briefly be described. Leonard and Goodman here take discreteness, '$\llcorner$', as primitive relation (Leonard; Goodman 1940, 46) and define parthood ('<'), overlap ('$\circ$'), fusion ('Fu') and sum ('+'), and nucleus ('Nu') and product (by concatenation of terms). Note that fusion and nucleus are generalised notions of binary sum and product, respectively.

Definitions

I.01 $x<y =_{df} \forall z(z \llcorner y \supset z \llcorner x)$
I.02 $x \circ y =_{df} \exists z(z<x \wedge z<y)$
I.03 $x \text{ Fu } \alpha =_{df} \forall z(z \llcorner x \equiv \forall y(y \in \alpha \supset z \llcorner y))$
I.04 $x \text{ Nu } \alpha =_{df} \forall z(z<x \equiv \forall y(y \in \alpha \supset z<y))$
I.06 $x+y =_{df} \text{Fu}`(\{x\} \cup \{y\})$
I.07 $xy =_{df} \text{Nu}`(\{x\} \cup \{y\})$

'Fu`' and 'Nu`' are so-called *descriptive functions*, introduced in *Principia Mathematica*:

PM *30.01 $\text{R}`y =_{df} (\iota x)(x\text{R}y)$

Thus, since 'x Fu α' can be read as 'x fuses the members of α', 'Fu`α' should be understood as 'the fusion of the members of α'.

I.03–I.07 mention classes: a class variable, ‘α’, occurs in I.03 and I.04, on which I.06 and I.07 build. Leonard and Goodman suggest that instead of the definitions mentioning classes the following could be used (Leonard; Goodman 1940, 48):

I.06’ $x+y =_{df} (\iota z)(\forall w[w \sqsubset z \equiv (w \sqsubset x \wedge w \sqsubset y)])$
I.07’ $xy =_{df} (\iota z)(\forall w[w<z \equiv (w<x \wedge w<y)])$

The three axioms they propose are (Leonard; Goodman 1940, 48–49):

Postulates

I.1 $\exists x\, x\in\alpha \supset \exists y\, y \text{ Fu } \alpha$
I.12 $(x<y \wedge y<x) \supset x=y$
I.13 $x\circ y \equiv \neg x \sqsubset y$

Some Theorems

I.3 $(x<y \wedge y<z) \supset x<z$
I.31 $x<x$
I.331 $x\circ y = y\circ x$
I.332 $x<y \supset x\circ y$
I.333 $x\circ x$
I.53 $\exists x\, x\in\alpha \equiv \text{E! Fu‘}\alpha$
I.6 $\text{E!}x+y$
I.62 $x+y = y+x$
I.66 $(x+y)+z = x+(y+z)$

It is relatively straightforward to see that the Calculus of Individuals satisfies all definitions of the Calculus of Singular Terms and also proves all axioms of the latter. The Calculus of Singular Terms is thus a subsystem of the Calculus of Individuals. The converse is much harder to establish: in fact, it is impossible. Neither I.01 (or rather, the right-to-left direction of the definitional equivalence) nor I.06, can be derived in Leonard’s system. The Calculus of Singular Terms is hence a *proper* subsystem of the Calculus of Individuals. (See Appendix A for a proof sketch.)

The definitions of ‘overlap’ and ‘product’ fall short of the characterisation that these notions are given in the Calculus of Individuals,

which we nowadays associate with classical extensional mereology (which, essentially, is a Boolean algebra with the zero-element removed[6]).

3.1 Non-extensional parthood

The "shortcoming" in the definition of 'overlap' in the Calculus of Singular Terms might turn out to be a blessing in disguise, however. Using 'overlap', it is possible to define in this system a "non-extensional"[7] part-relation in the following way (compare I.01):

$$D_{\prec} \quad x \prec y =_{df} \forall z(z \circ x \supset z \circ y)$$

Call the resulting system $ST^{\prec}$. This non-extensional part-relation, '$\prec$', might be of interest in contemporary discussions in analytic metaphysics: think of the problem of the statue and the clay. A statue and the lump of clay it consists of arguably share — at least for some given time, *t*, say — all their parts. Nevertheless, one might want to resist identifying the statue with the clay. The thought behind this is that the statue will survive the loss of a bit of matter, while the lump of clay will not. At the given time *t* the statue and the clay will coincide, however, in the sense that they completely overlap each other. But in the Calculus of Individuals and other classical extensional mereologies complete overlap means identity:

$$\forall z(x \circ z \wedge y \circ z) \supset x = y$$

Conveniently, in the Calculus of Singular Terms this does not hold, and $\prec$ inherits this non-extensionality of $\circ$. Thus, while mutual extensional parthood means identity:

$$(x < y \wedge y < x) \supset x = y$$

this is not true of the non-extensional part-relation:

$$\nvdash (x \prec y \wedge y \prec x) \supset x = y$$

So, while the statue and the lump of clay can share all $\prec$-parts and be $\prec$-part of each other, this does not entail that they are identical.

---

[6] See (Eberle 1970, 36), and (Simons 1987, 25); see (Ridder 2002, chapter 3), for a detailed investigation.

[7] Following the terminology suggested by (Simons 1987).

This is not the place to investigate the philosophical potential of the here introduced $\prec$-relation. Note, however, before we move on, that $\prec$ appears to have all right to be called a part-relation, since it has many of the characteristic properties. In particular, the following theorems hold (see Appendix B for hints on the proofs):

Some theorems of $\mathrm{ST}^{\prec}$

$\mathrm{ST}^{\prec}.1$	$x \prec x$	(compare *16.25)
$\mathrm{ST}^{\prec}.2$	$(x \prec y \wedge y \prec z) \supset x \prec z$	(compare *16.281)
$\mathrm{ST}^{\prec}.3$	$x<y \supset x \prec y$	
$\mathrm{ST}^{\prec}.4$	$(x \prec y \wedge x \circ z) \supset y \circ z$	
$\mathrm{ST}^{\prec}.5$	$(x \prec y \wedge z \prec x) \supset y \circ z$	
$\mathrm{ST}^{\prec}.6$	$(x \prec y \wedge x \prec z) \supset y \circ z$	(compare *16.31)
$\mathrm{ST}^{\prec}.7$	$x \prec y \supset (x \circ y \wedge y \circ x)$	(compare *16.321)
$\mathrm{ST}^{\prec}.8$	$x \prec (x+y)$	(compare *16.26)
$\mathrm{ST}^{\prec}.9$	$(x \prec z \wedge y \prec z) \supset (x+y) \prec z$	(compare *16.28, r.t.l.)

Extensionality fails for $\prec$ however. Accordingly, $\prec$ is a proper subrelation of $<$.

$\nvdash (x \prec y \wedge y \prec x) \supset x=y$	($\prec$ is non-extensional)
$\nvdash x \prec y \supset x<y$	($\prec$ and $<$ do not coindcide)

Note further:

$\nvdash x \prec y \supset y \prec x$	($\prec$ is not symmetric)
$\nvdash x \prec y \supset \neg y \prec x$	($\prec$ is not antisymmetric)

4. Goodman's input

Despite not being discussed in print, it has been subject of speculation for a long time, what exactly Goodman's contribution to the Calculus of Individuals was. It seems clear that the third section of Leonard and Goodman's paper (1940), which deals with the application of the calculus to Carnap's *Aufbau* and the solution of its difficulty of imperfect community, was largely Goodman's work. It is an open question, however, to what extent Goodman contributed to the construction of the calculus

itself and how much of it was developed by Leonard. Some light has been shed on the question here by outlining how much of the system was already contained in *Singular Terms*.

The problem is complicated by the fact that Leonard co-operated with Goodman already before he finished his Ph.D. dissertation thesis. Leonard thanks Goodman in the preface of *Singular Terms*, and states that the two "discussed together nearly every point developed in this thesis" (Leonard 1930, v). Indeed, in *The Structure of Appearance*, Goodman remarks in a footnote:

> The calculus to be outlined here was developed by Henry S. Leonard and Nelson Goodman. It was first presented in Leonard's doctoral thesis *Singular Terms* […] (*SA*, 33, fn. 8).

The reference to Leonard's *Singular Terms* is missing from Goodman's *A Study of Qualities* (Goodman 1940). In *Problems and Projects* Goodman vaguely concedes:

> The first thought of developing some of the basic ideas by the use of the (then rather new) techniques of symbolic logic came from my fellow student Henry Leonard, and our collaboration led to "The Calculus of Individuals and its Uses" (*PP*, 149).

Recently a document surfaced that might be of help. In 1966, Leonard started preparing a collection of his papers for publication; a project that was never finished since Leonard sadly died of a heart attack on July 11, 1967, aged 61, on a vacation in Frankfurt, Germany. The collection was meant not only to contain his joint paper with Goodman, but also introductory comments on each paper. The comment on "The Calculus of Individuals and its Uses" was written, and is now due to be published (see Leonard 1967). In this comment, Leonard writes:

> Although Goodman and I published "The Calculus of Individuals" only in 1940, such a calculus had for a long time been occupying our attention, both independently and collaboratively. Concern with a part-whole relation between individuals was a major one in Goodman's Honors Thesis, submitted to the Harvard Department of Philosophy when he was a senior in 1928. A formal development of the calculus constituted Chapter IV of my doctoral dissertation, submitted to the Harvard department in December, 1930. In the fall of that year, as I was writing the thesis, Goodman and I met together many times for exchanges of ideas. […]
>
> The earliest version of the calculus, in my dissertation in 1930, differed from the later versions in certain significant respects. It was presented as an interpolation in

> Whitehead and Russell's *Principia Mathematica* between *14 and *20. Hence, it did not include such general notions as those of the fusion and the nucleus of a class (I.03 and I.04, below). Instead of taking the relational expression '$x \perp y$' (i.e., 'x is discrete from y') as primitive, it took the operation '$x+y$' as primitive. It rested on more postulates than did the later versions. (In fact, the 1936 version still had five postulates.) […]
>
> If responsibilities can be divided in a collaborative enterprise, I believe that it may be fairly stated that the major responsibility for the formal calculus (Part II, below) was mine, while the major responsibility for discussions of applications (Part III) lay with Goodman.

Goodman's Honors Thesis unfortunately appears to be lost; at least, it is not contained in any of the known collections of his *Nachlass*.[8]

Despite Leonard's comments, it is still not entirely clear what impact Goodman had on the Calculus of Individuals. It is also unknown what exactly Quine's role was. Quine writes in his autobiography that his intervention took place in 1935 when Leonard, Goodman and Quine had a stop-over for the night on their way back to Cambridge from the famous trip to Baltimore they made with Carnap:

> Goodman, Leonard, and I drove back to Cambridge, stopping in New York at the Lafayette, a French hotel off Washington Square. On the way they told of a project of theirs. They broached it diffidently, for I had seemed unsympathetic when Henry spoke of it on an earlier occasion. I became interested as I heard more, and I was able to help them on a technical problem. We talked in our hotel room until four in the morning. They were concerned with constructing a systematic theory of sense qualities, and their effort had much in common with Carnap's *Logischer Aufbau der Welt*. As an auxiliary they had developed a logic of part–whole relation, which I recognized as Leśniewski's so called mereology (Quine 1985, 122).

What this technical problem was appears to be lost to history.

The fact that the Calculus of Singular Terms is a proper subsystem of the Calculus of Individuals, and is also formulated without the use of classes in contrast to the latter, opens up the possibility of a significant input of

[8] Part of Goodman's *Nachlass* is held by the Harvard University Archives, part by his literary executor, Catherine Elgin.

Goodman's. But given the lack of further evidence, this must remain speculation.[9]

Appendix A

To prove that the Calculus of Singular Terms (ST) is a proper subsystem of the Calculus of Individuals (CI), it suffices to show that the definitions and axioms of ST can be derived in CI, but that the converse does not hold.

We leave out of consideration the notions defined in either system that do not interact with the axioms (such as the difference in ST, or the universe, proper part, and complement in CI). ST does not contain the discreteness relation, '$\llcorner$', which, for the purpose at hand, is taken to be defined thus (compare I.13):

$x \llcorner y =_{df} \neg x \circ y$

Further, only the class-free fragment of CI is considered. I.06' and I.07' thus replace I.03–I.07; an axiom that asserts the existence of arbitrary binary sums is used in lieu of the fusion axiom, I.1, which contains a variable ranging over classes.

I.1' $\exists z\, x+y = z$

Obviously, I.1 implies I.1'. The tables below show what postulates and definitions are used in CI to derive the axioms and definitions of ST, and *vice versa*.

ST definitions	derived in CI using
*16.01	I.01, I.06', I.1'
*16.02	=I.02
*16.03	I.01, I.02, I.06', I.07', I.1', I.12, I.13

ST postulates	derived in CI using
*16.1	(=I.62) I.01, I.02, I.06', I.1', I.12, I.13
*16.12	I.06', I.1'

[9] I would like to thank Ralf Bader, Daniel Cohnitz, Philip Ebert, Karl-Georg Niebergall and an anonymous referee for helpful comments on earlier versions of this paper.

*16.14	I.01, I.02, I.06', I.07', I.12, I.13
*16.16	I.01, I.02, I.06', I.07', I.12, I.13
*16.17	I.06', I.1', I.13
*16.18	I.01, I.02, I.06', I.07', I.12, I.13

Thus, ST is a subsystem of CI. The converse, however, does not hold:

CI definitions	derived in ST using
I.01 l.t.r.	*16.01, *16.02, *16.1, *16.14
I.01 r.t.l.	*fails*
I.02	= *16.02
I.06'	*fails*
I.07'	*16.01, *16.02, *16.03, *16.1, *16.12, *16.14, *16.16

CI postulates	derived in ST using
I.1'	(= *16.21) *16.1
I.12	(cf. *16.274) *16.01, *16.1
I.13	n/a

Appendix B

The system $ST^{\prec}$ introduced in this paper extends the language of ST by the binary relation symbol, '$\prec$', for non-extensional parthood, and the axiom system by the definition:

$D_{\prec}$ $\quad x \prec y =_{df} \forall z(z \circ x \supset z \circ y)$

The following table lists some theorems of $ST^{\prec}$ and what is required for their proofs:

$ST^{\prec}$ theorems	derived using $D_{\prec}$ and:
$ST^{\prec}$.1	($D_{\prec}$ suffices)
$ST^{\prec}$.2	($D_{\prec}$ suffices)
$ST^{\prec}$.3	*16.01, *16.02, *16.1, *16.14
$ST^{\prec}$.4	*16.02
$ST^{\prec}$.5	*16.01, *16.02, *16.1, *16.12
$ST^{\prec}$.6	*16.01, *16.02, *16.1, *16.12

$ST^{\prec}.7$	*16.01, *16.02, *16.1, *16.12
$ST^{\prec}.8$	*16.01, *16.02, *16.1, *16.14
$ST^{\prec}.9$	*16.1, *16.17

References

Carnap, R. (1928): *Der logische Aufbau der Welt*. Berlin: Weltkreis, 1928.

Cohnitz, D.; Rossberg, M. (2006): *Nelson Goodman*, Cheshalm: Acumen, 2006.

Cohnitz, D. (2009): "The Unity of Goodman's Thought", in this volume.

Eberle, R.A. (1967): "Some Complete Calculi of Individuals", in: *Notre Dame Journale of Formal Logic* 8 (1967), 267-278.

Eberle, R.A. (1970): *Nominalistic Systems*, Dordrecht: Reidel, 1970.

Goodman, N. (1940): *A Study of Qualities* [1940], Ph.D. dissertation Thesis, New York: Garland, 1990.

Leonard, H.S. (1930): *Singular Terms*, Ph.D. dissertation Thesis, Widener Library, Harvard University, 1930.

Leonard, H.S. (1967): "Comments on "The Calculus of Individuals and its Uses" (1967)", edited by Henry S. Leonard Jr., in: Burkhardt, H.; Imaguire, G.; Seibt, J.: *Handbook of Mereology*, Munich: Philosophia Verlag, (forthcoming).

Leonard, H.S.; Goodman, N. (1940): "The Calculus of Individuals and its Uses", in: Journal of Symbolic Logic, 5 (1940), 44-55.

Leśniewsky, S. (1916): *Podstawy ogólnej teorii mnogości* I [Foundations of the General Theory of Manifolds, I], Moscow: Poplawski, 1916.

Leśniewsky, S. (1927): "O podstawach matematyki" [On the Foundations of Mathematics], in: *Przegląd Filozoficzny* 30 (1927): 164–206; 31 (1928): 261–291; 32 (1929): 60–101; 33 (1930): 77–105; 34 (1931): 142–170.

Leśniewsky, S. (1929): "Grundzüge eines neuen Systems der Grundlagen der Mathematik", in: *Fundamenta Mathematicae* 14 (1929), 1-81.

Quine, W.V. (1985): *The Time of My Life: An Autobiography*, Cambridge, MA: MIT Press, 1985.

Ridder, L. (2002): *Mereologie*, Frankfurt: Klostermann, 2002.

Ridder, L. (2003): “Gegenstände der Zeit”, in: Metaphysica 4 (2003): 29-58.

Simons, P. (1987): *Parts: A Study In Ontology*, Oxford: Oxford Clarendon Press, 1987.

Tarski, A. (1929): “Les fondements de la géométrie des corps”, in: *Annales de la Société Polonaise de Mathématique*, suppl. vol., 1929, 29-33; revised version, translated by J.H. Woodger: “Foundation of the Geometry of Solids”, in: Tarski, A.: *Logic, Semantics, Metamathematics*, edited by John Corcoran, Indianapolis: Hackett, 1983, 24-29.

Tarski, A. (1937): “Appendix E”, in: Woodger, J.H.: *The Axiomatic Method in Biology*, Cambridge: Cambridge University Press, 1937, 161-172.

Whitehead, A.N.; Russell, B. (1910-1913): *Principia Mathematica*, 3 vols., Cambridge: Cambridge University Press, 1910-1913.

Wolfgang Heydrich

Counterfactuals beyond Paradox[1]

Abstract

Goodman has proposed a basically enthymematic analysis of counterfactuals:

$$A \langle C := \exists S.S\&F[S]\&(S\&A\rightarrow C)\& \neg\exists S.S\&F[S]\&\neg(S\&A\rightarrow\neg C),$$

which is to support contraposition ($A \langle C \mid\!\!- \neg C \langle \neg A$) and Boethius' thesis ($A \langle C \mid\!\!- \neg(A \langle \neg C)$). There are two main problems with this account: how to specify the condition F[_] and how to interpret '→'. As is well known, the second problem leads Goodman *via* the concept of *law* to the theory of *entrenchment* of predicates and *projectibility* of hypotheses. The former leads him *via* the notion of *cotenability* into circularity or infinite regress, "for cotenability is defined in terms of counterfactuals, yet the meaning of counterfactuals is defined in terms of cotenability" (*FFF*, 16). Hence, Goodman's attempt to specify truth conditions for counterfactual conditionals fails.

As a remedy, I am proposing a para-consistent interpretation of '→' which rejects the *paradox of law-like implication* ($\mid\!\!- A\&\neg A \rightarrow C$) and does not consider *logical* consequence ($A \mid\!\!- C$) as a sub-relation of *law-like* consequence ($\mid\!\!- A\rightarrow C$). As a result, contraposition may be retained, but Boethius' thesis has to go. The para-consistent framework in which I am working is a decidable fragment of what I call the "*logic of entrenchment*". Goodman's puzzles about *Jones in Carolina* and *the wet match* are reconsidered in light of the modified account.

---

[1] This paper is an (hopefully) improved version of (Heydrich 2007). The material has also been presented at the conference "Nelson Goodman: Von der Logik zur Kunst – Ein Rückblick anläßlich des 100. Geburtstags", Munich, August, 5-8, 2006. Thanks to Greg Rastall and Catherine Elgin for advice and encouragement.

1. Introduction

"*Great minds against themselves conspire …*"
Henry Purcell: Dido and Aeneas

Nelson Goodman's lecture, "The problem of counterfactual conditionals", held in New York in 1946, published in 1947, in the *Journal of Philosophy*, and adopted later as the introductory chapter of *FFF* (first edition: 1954), has been recognised as occupying a central position, not only in Goodman's *oeuvre* itself, but in the 20th century's (analytical) philosophy in general[2].

In this seminal essay, Goodman identifies and discusses several difficulties in explicating truth conditions for conditional sentences like

(1) *If the match had been scratched, it would have lighted.*

His observations and analyses are rich and insightful. However, they do not (according to him) lead to a satisfactory result. After several pages of semantic exertion, Goodman sees himself caught in circularity: "[…] we can never explain a counterfactual except in terms of others, so that the problem of counterfactuals must remain unsolved" (*FFF*, 16f). He continues somewhat resignedly: "Though unwilling to accept this conclusion, I do not at present see any way of meeting the difficulty."

The purpose of this paper is to propose a simple way out.

2. Goodman's basic idea

Goodman's basic intuition about (1) is that it is kind of enthymematic. (1) is true, he suggests, if and only if there are sound inferential steps leading from its antecedent part (*the match has been scratched*) to its consequent part (*the match has lighted*). The inference involved is enthymematic, depending on support premises (*the match has been dry*, *it is well made*, *enough oxygen has been present*, etc.), which fix relevant obtaining circumstances that are left implicit in (1).

[2] See Putnam's "Foreword to the fourth edition" (*FFF*, vii ff).

Using '⟨' for the interesting conditional link between antecedent A and consequent C, variable *S* for a support proposition and '→' for the required notion of implication, this idea might be captured by the following definition[3]:

(2) $A \langle C := \exists S.\ S\&(S\&A \rightarrow C)$.

(2) is meant as a first approximation. Certainly, refinement is needed (see below). As it stands, (2) does not restrict support propositions beyond truth. How to formulate further constraints is Goodman's first problem with the basic idea (cf. *FFF*, I.2, 9ff, "*The Problem of Relevant Conditions*"). Apart from the question of necessary refinement, however, there is need for a general comment.

The definiens of (2) is an object-language expression making use of quantification over values of propositional variables and an (yet unexplained) intensional implicational connective.[4] This, of course, is at odds with Goodman's general nominalistic convictions, which shun

---

[3] For a treatment of enthymematic inferences along these lines, see Michael Dunn in (Anderson & Belnap 1992, §35, 45-55), where '→' is used as entailment in the sense of system E of relevance logic. Dunn does not note the parallel between his treatment of enthymemes and Goodman's treatment of counterfactuals.

I am inclined to call this the *bound* kind of enthymematicity. Given some context (represented by a set of propositions Σ), ⌜A ⟨ C⌝ is *true* in context Σ, if the propositions in Σ cannot be true, unless there is a proposition *S* such that (*inter alia*) ⌜*S*&(*S*&A → C)⌝ is true. This should be kept apart from the *indexical* kind, according to which ⌜A ⟨ C⌝ is *enthymematically true* in context Σ, if there is a proposition *S* (true or not) such that the propositions in Σ cannot be true, unless (*inter alia*) ⌜*S*&A → C⌝ is true. Goodman tries to make clear (*contra* Chisholm (1946, 1955), perhaps, cf. (Bennett 2003, 304ff)) that in his explication, he is concerned with the bound kind: "[T]he connection we affirm may be regarded as joining the consequent with the conjunction of the antecedent and other statements that truly describe relevant conditions. [… O]ur assertion of the counterfactual is *not* conditioned upon these circumstances obtaining. We do not assert that the counterfactual is true *if* the circumstances obtain; rather, in asserting the counterfactual we commit ourselves to the actual truth of the statements describing the requisite relevant conditions" (*FFF*, 8). In Section 6, I will have occasion to consider the indexical kind as well.

[4] Obviously, '→' cannot be interpreted as material conditional. Otherwise the definiendum of (2) would come out as valid, if ⌜A⌝ is false. (See *FFF*, 4).

propositions as entities in the domain of quantifiers. Hence, Goodman (without making much ado about this point) prefers to speak of (a conjunction of) supporting *sentences* (instead of *propositions*). Accordingly, he treats the connection between the augmented antecedent and the consequent as a relation between sentences and not in terms of an operation on values of propositional variables. Maybe we should render Goodman's basic idea as:

(2') $\ulcorner A \langle C \urcorner$ is true iff for some true sentence S: $\ulcorner S\&A \urcorner \Rightarrow \ulcorner C \urcorner$,

where '$\Rightarrow$' has to be understood as a meta-linguistic counterpart of '$\rightarrow$'. The meta-linguistic character of Goodman's analysis – made explicit in (2') – has been interpreted as its central feature by some commentators (e.g., Lewis 1973, 65ff). In contrast, I consider the difference between (2) and (2') as rather superficial.

Of course, (2) requires the specification of a domain of values for propositional variables which may be dispensed with if we stick to (2'). Inversely, however, (2') – but not (2) – involves an hierarchy of semantic meta-languages as soon as we consider logical relations among counterfactuals (e.g., between $\ulcorner A \langle C \urcorner$ and $\ulcorner \neg C \langle \neg A \urcorner$) or – worse still – nested constructions (like $\ulcorner (A \langle C) \langle (\neg C \langle \neg A) \urcorner$). Notwithstanding the difficulties of working within a stratified semantic meta-theory[5], it seems to me that, in principle, we may translate back and forth between both accounts. (At least, I assume, we are not gaining or losing much.)

Anyway, (2) and (2') confront us with parallel (if not identical) conceptual queries: in using (2), we are in need of an explication of an implicational *operation* $\rightarrow$ on *propositions*. In using (2'), we have to gain an understanding of a corresponding *relation* between *sentences*.

In order to constrain different accounts – intensionalistic ones along the line of (2) and meta-linguistic ones using (2') – , one might express this parallelism by postulating

(3) $\ulcorner A \rightarrow C \urcorner$ is true iff $\ulcorner A \urcorner \Rightarrow \ulcorner C \urcorner$.

[5] Note that if (2') is meant as a definition, sentence S may have counterfactual constituents itself, only if there is some ranking of counterfactual degrees.

One possibility to interpret '⇒' which suggests itself is to take it as logical consequence in the official sense[6]:

(2'') ⌜A ⟨ C⌝ is true iff for some true sentence *S*: *S*&A |– C.

However, Goodman observes that in cases like (1), this seems to be inadequate. Generally, the relation between ⌜A⌝ (*plus* a conjunction of sentences specifying *special* obtaining circumstances) and ⌜C⌝ will not be *logical* consequence.

Perhaps, then, we should understand ⌜X ⇒ Y⌝ as ⌜X leads to Y by *law*⌝ or: ⌜the inference from X to Y is *law-like*⌝. This is effectively how Goodman talks (cf. *FFF*, 11ff).

One might think (but Goodman does not – cf. *FFF*, 18f) that we can evade the problem of law-likeness by sticking to (2'') if we allow the inclusion of universal generalisations in the supporting premise. Indeed, if we add the true generalisation

(4) *All well-made matches, which are scratched under the obtaining relevant circumstances, will light,*

to a specification of the pertinent *special* circumstances of an utterance of (1), we are led from the augmented antecedent to the consequent by logic alone.

This move, however, only conceals the need for a concept of *law-likeness*. Note that, given the truth of a generalisation like

(5) *All coins in Nelson's pocket are silver*,

we will derive along these lines,

(6) *This very penny made of copper would be a silver coin, if it were in Nelson's pocket.*

Instead, one would be inclined to conclude

---

[6] I.e., A_1, …, A_k |– B iff ⌜A_1& … & A_k & ¬B⌝ is logically inconsistent (false in every model) and |– B iff ⌜B⌝ is logically true (true in every model). As usual, '|–' is taken to absorb (quasi-) quotes around its argument(s). In the sequel, I will use '|–' only in its official sense. This needs mention because I will explicate '→' in the framework of relevance logics, where one also deals with (non-official) concepts of logical consequence which are stronger. Cf. (Heydrich 1993a).

(7) *Some coin in Nelson's pocket would not be silver, if this penny were in it*.

What is required for universal generalisations to be added to antecedents (plus specification of special relevant circumstances) is for them to not only be true but be *laws*. (4) is, if true, acceptable as a law, while (5) is a universal generalisation, but definitely not law-like.

The best we can do in using (2'') (this is the strategy Goodman advocates) is to specify relevant conditions that identify obtaining circumstances for support, *including* law-like generalisations that connect special conditions (*plus* the antecedent) to the consequent. Following this strategy, we need an analysis of the concept of *general law*.

But obviously, this strategy simply amounts to a restriction of sentence *S* such that $\ulcorner S\&A\urcorner \Rightarrow \ulcorner C\urcorner$ and we have gained nothing. So, we should stick to (2'). Whether additional restrictions for admissible supporting sentences are required remains to be answered and will depend on how exactly we explicate $\Rightarrow$ (see Section 3).

Alternatively, we can use (2) (this is the strategy I am advocating in this paper). We then need an interpretation of '$\rightarrow$' as an operator for *law-like* implication. Again, we will have to discuss whether (2) needs refinement or not (see Section 5).

Either way, the concept of *law-likeness* is central. This is Goodman's second problem with the basic idea (cf. *FFF*, I.3, 17ff, "*The Problem of Law*"). His analysis of general laws makes use of the notion of *entrenchment*. Perhaps the theory of entrenchment (*FFF*, 84ff, cf. also *PP*, 357ff) is the most constructive part of his overall sceptical essay. The central idea is that a universal generalisation's being law-like crucially depends on the degree to which the predicates involved have turned out to be useful in our past practice of induction and pertinent habits of description and explanation. Without discussing this theory here, I will take up Goodman's term and basic idea. In Section 4, I shall try to explicate the implicational operation $\rightarrow$ (along with truth-functional connectives and propositional quantification) in terms of truth value *and* degree of entrenchment of the propositions involved.[7]

[7] An atomic proposition $\ulcorner F(t_1, \ldots, t_k)\urcorner$ may be considered to inherit the degree of entrenchment of its predicate F.

In taking (2) (instead of (2')) as my starting point, I am certainly not following the letter of Goodman's account; but in accepting his intuitions concerning the enthymematic character of counterfactual conditionals and the central role of entrenchment for law-like implication between propositions (or sentences), I intend to pay tribute to the basic intuition and gist of his analysis.

3. Restrictions and criteria of adequacy

Let us start with some remarks concerning certain restrictions Goodman imposes on his analysis.

It should be stressed that Goodman's essay is not intended as a linguistic study in the compositional semantics of grammatical constructions of the form ⌜If *p* [+PAST], then *q* [+would, +INF]⌝ or the like. Such an account should certainly take seriously syntactic structure and matters like tense, aspect, and modality.[8] Rather, Goodman aims at a specific logico-philosophical link in the background, a conditional connection between component sentences that happens to be instantiated by grammatical constructions of the considered type. E.g., Goodman is not engaged in highlighting subtle (but semantically puzzling) differences between so-called *indicative* and *subjunctive* conditionals, which have been discussed in the literature since.[9] "The problem," as conceived by him, "is independent of the form in which a given statement happens to be expressed" (*FFF*, 4).

Goodman notes that his own label "*counterfactual* conditional" is not felicitous. Do we always *claim* or *presuppose* the falsity of *p* and *q* in statements of the form "If *p* were the case, than *q* would be the case"? It seems that at times, we only indicate improbability, insecurity, or merely highlight suspended commitment to truth or falsity (cf. Karttunen & Peters (1978)). In his essay, Goodman mainly confines discussion to "counterfactuals as such" (*FFF*, 5), i.e., sentences "in which antecedent

[8] The same is true for conjunctions like *since* and adverbs or particles like *still* and *even* (see below).

[9] Cf. e.g., *If Shakespeare hadn't written Hamlet, then someone else would have* (false, we presume) and *If Shakespeare didn't write Hamlet, then someone else did* (true). Cf. (Adams 1970) and the discussion in (Bennett 2003).

and consequent are inalterably false", but "a general solution," he maintains, "would explain the kind of connection involved irrespective of any assumption as to the truth or falsity of the components". Accordingly, he considers contrapositions of counterfactuals (so-called "*factual* conditionals") like

(8) *Since the match did not light, it has not been scratched,*

which come with the presumption of the truth of both antecedent and consequent as mere transpositions of counterfactuals. "The possibility of such transformations," Goodman writes, "is of no great importance except to clarify the nature of our problem." Hence, he seems to suggest that contra-posing counterfactuals preserves truth conditions. In other words, I take Goodman to assume the principle

(9) $A \langle C \mid\!- \neg C \langle \neg A$

as a criterion of adequacy for his analysis. This is remarkable, if only for the fact that distinguished students of counterfactuals, like David Lewis (1973) and Robert Stalnaker (1968), think otherwise.

Goodman even wants his analysis (the "general solution") to cover what he calls "*semifactuals*" (*FFF*, 5) like

(10) *Even if the match had been scratched, it still would not have lighted*,

which strongly indicate the falsity of their antecedent and the truth of their consequent. The fact that we "uncompromisingly reject" contraposition for semifactuals – as witnessed, according to Goodman, by

(11) *Even if the match lighted, it still wasn't scratched –*

does not bother him. Goodman attributes this fact to pragmatic factors connected with the "auxiliary terms" (*FFF*, 6) *even* and *still,* and takes it to show a difference between "practical" and "literal import", which is irrelevant on the level of conceptual analysis in which he is engaged.

However, even granting the relevance of such an abstract logico-philosophical level, one may still think that Goodman is exaggerating by rejecting *each* "assumption as to the truth or falsity of the components" of counterfactuals as irrelevant. At least in the case of a true antecedent and a

false consequent, ⌜A ⟨ B⌝ should be considered false. This means, I think, that Goodman should accept

(12) $A \langle C \vdash A \supset C$.[10]

Another constraint one would like to impose is

(13) $A \rightarrow C \vdash A \langle C$[11],

because it seems to be inconsistent to deny that *this* (pointing to a dice) would be sugar, if it were a sugar cube and admit that something's being sugar is implied by its being a sugar cube. Goodman, however, must reject (13) in its general form.

The reason is this: Goodman wants to take measures against selecting true ⌜¬A⌝ as a value of the support sentence *S* in the (standard) case of counterfactuals with false ⌜A⌝. He believes that a characteristic feature (sometimes called "paradoxical") of classical logic – the feature that contradictions have arbitrary consequences – is shared by laws in general. Hence, he assumes

*(14) $\vdash A\&\neg A \rightarrow C$[12].

Given this, allowing the choice of ⌜¬A⌝ for *S* clearly leads to

*(15) $\neg A \vdash A \langle C$.

This, indeed, would amount to the Armageddon of Goodman's analysis: every counterfactual (with a false antecedent) would turn out to be true.

The requirement Goodman thinks necessary is that the range of admissible support *S* should be restricted such that ⌜*S*&A⌝ is "self-compatible" (*FFF*, 11ff), i.e., does not lead by law to its own negation. Note that *if* (14) is adopted, lack of self-compatibility is equivalent with implying any proposition.

---

[10] Hence, we have counterfactual *modus ponens* ($A \langle C, A \vdash C$) and *modus tollens* ($A \langle C, \neg C \vdash \neg A$).

[11] Taking (12) and (13) together, we get (by transitivity of '|–') the constraint: $A \rightarrow C \vdash A \supset C$. Law-like implication is a sub-case of material implication.

[12] Principles I will reject in the constructive part of this paper are marked with an asterisk. Principles without asterisks will be adopted.

Obviously, Goodman's requirement has far-reaching consequences. It implies (and is intended to imply (*FFF*, 12, footnote 5)) that ⌜A⌝ alone must be self-compatible as well. Hence, Goodman does not have in full generality either

(16) $\vdash A \langle A$,

(17) $\vdash A\&B \langle A$,

or – in view of (9), (17), De Morgan and Double Negation –

(18) $\vdash A \langle A \vee B$.

(16)–(18), however, seem to be plausible general principles.

As a remedy, Goodman exempts so-called *counterlegals*, like (19), or *counteridenticals*, like (20), from his analysis[13]:

(19) *Even if this sugar cube were* (*also*) *spherical, it still would consist of sugar.*

(20) *If Goodman were Cicero, Goodman would not have lived in the 20th century.*

By this move, Goodman may retain (16)–(18). Note that, given contraposition – i.e., (9) –, this means that (21) and (22) must also be excluded from consideration.

(21) *Even if this sugar cube did not consist of sugar, it still would not be* (*also*) *spherical.*

(22) *Since Goodman lived in the 20th century, he wasn't Cicero.*

Again, I consider it to be rather unfortunate to exclude such sentences from analysis even in preliminary stages of the theory (cf. *FFF*, 12, footnote 5).

By excluding counterlegals and counteridenticals, Goodman may be seen as favouring a connexive view[14] on counterfactuals[15]. Effectively, he

[13] Alternatively he treats all of them as false (cf. *FFF*, 12, footnote 5): "This is convenient for our present purpose […] If it later appears desirable to regard all or some counterlegals as true, special provisions may be introduced."

[14] Cf. Storrs McCall in (Anderson & Belnap 1975, §29.8).

[15] Cf. *FFF*, 15, footnote 9: "If two conditionals having the same counterfactual antecedent are such that the consequent of one is the negate of the consequent of the other, the conditionals are contraries and both may be false."

uses Boethius' thesis as a criterion of adequacy for counterfactual conditionals:

*(23) $A \langle C \vdash \neg(A \langle \neg C)$.

In view of (16), he also has Aristotle's thesis:

*(24) $\vdash \neg(A \langle \neg A)$.

In favour of (23), Goodman writes: "*S* by itself must not decide between C and ¬C, but *S* together with A must lead to C but not to ¬C" (*FFF*, 13). Accordingly, he restricts possible choices of support sentences by requiring that not only must ⌜*S*&A⌝ be "self-compatible" (see above), but also that ⌜*S*⌝ be compatible with ⌜C⌝ and ⌜¬C⌝[16]. He explicitly adds the requirement that there be *no* true[17] support sentence *S*' compatible with ⌜C⌝ and ⌜¬C⌝ such that ⌜*S*'&A⌝ is self-compatible and leads to ⌜¬C⌝. By urging that no such *S*' exists, Goodman clearly insists on (23).

In reaction to criticism of William T. Parry (1957), this already quite involved explication has been further complicated. Parry argued that the weakest truth-functional links between ⌜A⌝ and ⌜C⌝ (i.e., ⌜A ⊃ C⌝) and ⌜A⌝ and ⌜¬C⌝ (i.e., ⌜A ⊃ ¬C⌝) satisfy the restrictions imposed on *S* and *S*', respectively, if ⌜A⌝ is false. If so, the positive requirement of Goodman's explication turns out to be trivially true and its negative part is trivially false. Again, this means Armageddon: no counterfactual turns out to be true according to Goodman's original explication.

---

[16] This is the requirement that *S* should lead by law neither to ¬C nor to C. Strictly, the requirement of *S*'s compatibility with C (i.e., ¬(*S*→¬C)) is redundant, since *S*&A is required to lead by law to C, but not to any contradiction. The requirement of *S*'s compatibility with ¬C (i.e., ¬(S→C)) was introduced by Goodman (*FFF*, 12) to disqualify ¬C as support whenever C is false and compatible with A (which would lead to *¬C, ¬(A→¬C), $A \langle C \vdash A \langle \neg C$). It might also be used to disqualify C (in factual conditionals or semifactuals) as support proposition (which would lead to *C |– $A \langle C$).

[17] Goodman suppressed "true." This strengthens the requirement and shortens its wording. Whether this was an error (as supposed by Parry (1957, 86) and assumed silently by (von Kutschera 1975, 64)) or intended, I do not dare to decide. Goodman did not amend (or comment on) the original formulation of this point either in his answer to Parry (Goodman 1957) or in the third (or fourth) edition of his book (*FFF*, 13), where he took up Parry's criticism.

But note that Parry's argument relies on a *law-like* connection between a material conditional *plus* antecedent and its consequent:

*(25) $\vdash (A \supset C)\&A \rightarrow C$.

If (as I assume) this implies both

(25') $\vdash A\&C \rightarrow C$ and *(14) $\vdash \neg A\&A \rightarrow C$,

Parry's argument clearly depends on the assumption (shared by Goodman) that a law-like link between contradictions and arbitrary propositions exists[18].

Parry proposed skipping the negative part of Goodman's explication and adding the clause that $\ulcorner\neg S\urcorner$ is compatible with $\ulcorner\neg A\urcorner$ (i.e., $\ulcorner\neg(\neg A \rightarrow S)\urcorner$). Goodman accepted Parry's criticism but intended retaining the negative part (thus, insisting on Boethius' thesis). Hence, he added the requirement that neither *S* (in the positive part of his explication) nor *S*' (in the negative part) should follow by law from $\ulcorner\neg A\urcorner$.[19]

All of the difficulties of Goodman's basic idea discussed so far originate from (14). This constraint might be labelled "the paradox of law-like implication." Without (14), the choice of $\ulcorner\neg A\urcorner$ for *S* does not lead to (15). (16)–(18) may hold unrestrictedly, there is no need to exclude counterlegals or counteridenticals and Parry's criticism loses its point. Accordingly, (23) and (24) are not to be accepted. Instead, we may adopt the (more conservative) principle of counterfactual *reductio ad absurdum*:

(26) $A \langle \neg A \vdash \neg A$.

This should be motivation enough to rethink Goodman's intuitions in a paradox-free framework – a framework that rejects (14) and countenances instances of $A \vdash C$ without $\vdash A \rightarrow C$, i.e., a logic where logical

---

[18] Of course, this is a meta-logical claim: Any logic accepting $\vdash B\&A \rightarrow (B \vee C)\&A$ *and* the meta-logical principle: if $\vdash F \rightarrow G$ and $\vdash G \rightarrow C$, then $\vdash F \rightarrow C$ (as Parry's logic doubtlessly does), will have (14) along with (25). Just define $\ulcorner A \supset C\urcorner$ as $\ulcorner\neg A \vee C\urcorner$ and substitute $\ulcorner\neg A\urcorner$ for $\ulcorner B\urcorner$, $\ulcorner\neg A\&A\urcorner$ for $\ulcorner F\urcorner$ and $\ulcorner(A \supset C)\ \&\ A\urcorner$ for G.

[19] This amendment concerns "counterfactuals as such" (with a false antecedent and consequent). To account for factual conditionals and semifactuals as well, Goodman could have added the further requirement that neither *S* (in the positive part) or *S*' (in the negative part of his explication) should follow by law from C.

consequence (in the official sense[20]) is not presumed to be a sub-relation of law-like consequence[21]. Such an account will be presented in the next paragraph.

4. Logic of entrenchment

In this section, I will give a rough sketch of a propositional logic within which the definiens of (2) can be interpreted.[22]

Let $L_{\rightarrow,\forall}$ be the usual language of truth-functions ('¬' and '&' taken as primitive) augmented with the connective '→' and a device of propositional quantification '∀': If p is a propositional variable and ⌜A⌝ $\in L_{\rightarrow,\forall}$, then ⌜$\forall p$A⌝$\in L_{\rightarrow,\forall}$.[23] Adopting definition (2), ⌜A ⟨ C⌝ turns out to be well-formed in $L_{\rightarrow,\forall}$, iff ⌜A⌝ and ⌜C⌝ are well-formed.

Semantically, we interpret formulae by means of (implicational) *values*. A value is a pair of a *polarity* (+ or –, *vulgo*: a truth value) and an integer d which is to represent a *degree* (of entrenchment). For each i, j$\in$**Z** (i<0≤j), we call the closed interval k = {d$\in$**Z**: i≤d≤j} = [i,j] an *entrenchment structure*.

Examples of entrenchment structures include: [-1,0] = {-1,0}, [-1,1] = {-1,0,1}, [-1,2] = {-1,0,1,2} and [-3,4] = {-3,-2,-1,0,1,2,3,4}.

$V_k = \{+,-\} \times k$ is the set of values corresponding to k.

I will write 'm_k' short for <+, min(k)>, 'M_k' for <+, max(k)>, '$\mathbf{m}_k$' for <-, min(k)> and '$\mathbf{M}_k$' for <-, max(k)> as well as ⌜d⌝ for <+, d> and ⌜$\mathbf{d}$⌝ for <-, d>. (Hence, *italic*s indicate positive polarity, **bold**s negative polarity. Degrees are rendered neutral. They are unmarked with respect to polarity). With respect to entrenchment structure k: $\mathbf{M}_k$, M_k, $\mathbf{m}_k$ and m_k are also called *the absurd*$_k$, *the trivial*$_k$, *mere falsity*$_k$ and *mere truth*$_k$, respectively. Where no confusion is likely to arise, indices may be dropped.

We naturally define the *complement* of a value as the value with identical degree but different polarity:

---

[20] Cf. footnote 7 above.

[21] I.e., the relation obtaining between ⌜A⌝ and ⌜C⌝ iff |– A→C.

[22] Likewise, its refinement to be presented in Section 5.

[23] Let '∨', '⊃', '≡', '↔' and the existential quantifier '∃' be defined as expected. I trust that my conventions of saving brackets are sufficiently transparent.

(27) For all d$\in k$: $\neg d = \mathbf{d}$ and $\neg\mathbf{d} = d$.[24]

I assume that values are related to one another by *involvement*. E.g., the absurd$_k$ involves every value in V_k and the trivial$_k$ is involved by every value. Mere truth$_k$ involves all values in V_k with positive polarity, mere falsity$_k$ is involved by all values with negative polarity.

Let us call values with positive (negative) polarity *positive* (*negative*) and values with a degree ≥0 (<0) *strong* (*weak*). In each V_k, there is at least one weak and one strong value and there are as many negative as positive values. The trivial and the absurd are strong, mere truth and falsity are weak. The idea is that weak values neither involve nor are involved by their complements, whereas strong values involve their complements iff they have negative and are involved by them iff they have positive polarity.

Models for $L_{\rightarrow,\forall}$ will be introduced by means of model structures and interpretations (assignments of values to formulae).

A *model structure* S_k is an entrenchment structure with the relation of involvement:

(28) $S_k = \langle k, \Rightarrow\rangle$, where
- (a) k is an entrenchment structure,
- (b) $\Rightarrow$[25] is the smallest transitive and reflexive relation on V_k, such that for all d, e $\in k$:
  - (i) $m_k \Rightarrow d$ and $d \Rightarrow M_k$,
  - (ii) $d \Rightarrow d+2$, if d+2 $\in k$,
  - (iii) $\mathbf{d} \Rightarrow \mathbf{e}$, if $e \Rightarrow d$,
  - (iv) $\mathbf{d} \Rightarrow d$, if d≥0,
  - (v) $\mathbf{m}_k \Rightarrow M_k$ and $\mathbf{M}_k \Rightarrow m_k$.

It is not hard to convince oneself that, according to this explication, involvement is antisymmetric (hence, given transitivity and reflexivity, a

[24] I am using '¬' systematically ambiguous: as an operation on values and a connective for sentences. The same is true for '∨', '&' and '→' which will be introduced immediately.

[25] Although my choice of one and the same symbol is deliberate, strictly, the relation between implicational values should be distinguished from the relation among sentences discussed above.

partial order) and that, with respect to $\Rightarrow$, each pair of values of V_k has a greatest lower and a least upper bound in V_k.

We define greatest lower bound (*inf*imum) and least upper bound (*sup*remum) in the usual way:

(29) For all $X \subseteq V_k$, $y \in V_k$:

inf $X = y$ iff (a) for all $x \in X$: $y \Rightarrow x$ and (b) if for all $z \in V_k$, $x \in X$: $z \Rightarrow x$, $z \Rightarrow y$.

sup $X = y$ iff (a) for all $x \in X$: $x \Rightarrow y$ and (b) if for all $z \in V_k$, $x \in X$: $x \Rightarrow z$, $y \Rightarrow z$.

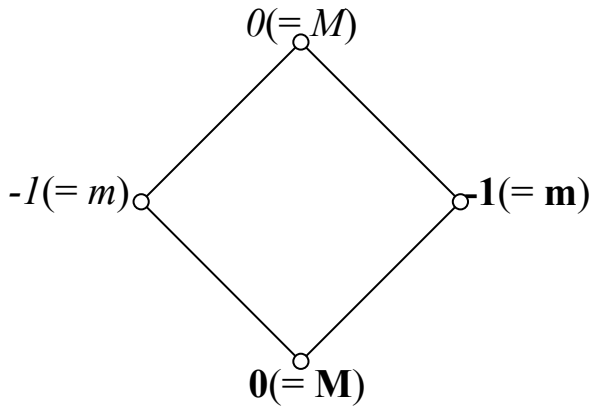
For all $x, y \in V_k$: $x\&y = inf\{x,y\}$ and $x \vee y = sup\{x,y\}$.

Hence, model structures are lattices. Because of (27) and (28)(b)(iii), we have for all $x, y \in V_k$: (a) $x = \neg\neg x$ and (b) $x \Rightarrow y$ iff $\neg y \Rightarrow \neg x$, hence $x\&y = \neg(\neg x \vee \neg y)$ and $x \vee y = \neg(\neg x \& \neg y)$. Therefore, $\neg$ is a so-called *De Morgan* complement. However, it is not *Boolean*: in general we do not have $x\&\neg x = \mathbf{M}_k$ or $x \vee \neg x = M_k$ for each $x \in V_k$.

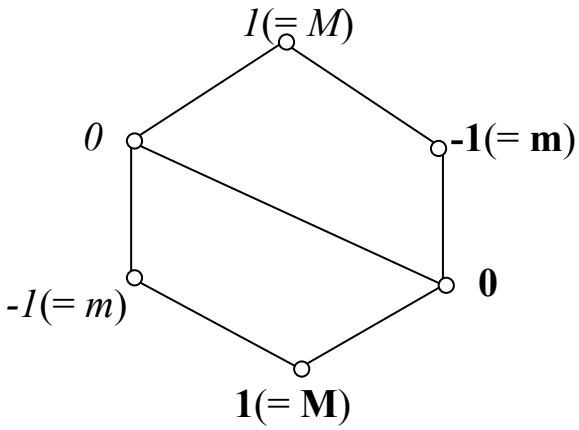
Let us have a look at some examples of model structures which I represent by their Hasse diagrams. As usual, $x \Rightarrow y$ is displayed by putting x in a connected path below y.

Among these examples, $S_{[-1,0]}$, $S_{[-1,1]}$ and $S_{[-1,2]}$ are distributive, but $S_{[-3,4]}$ is not:

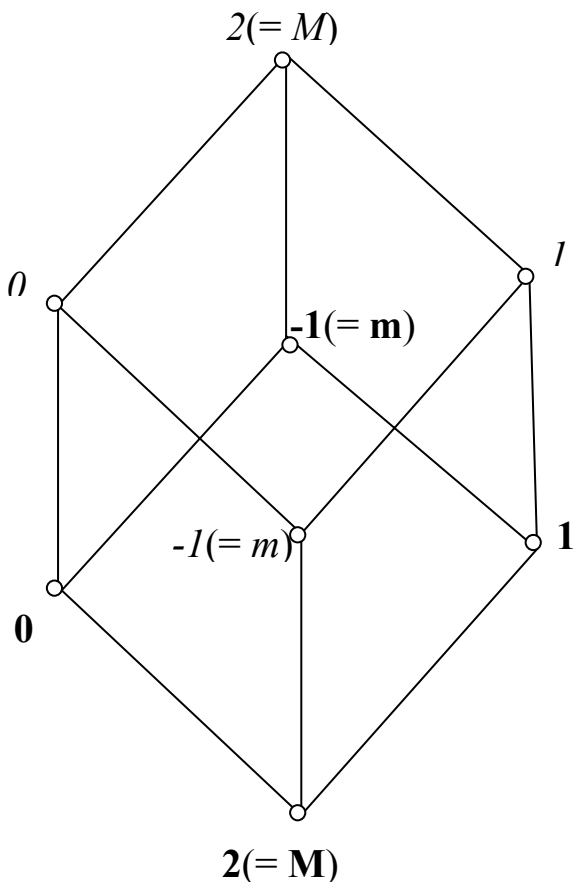
$S_{[-1,0]}$



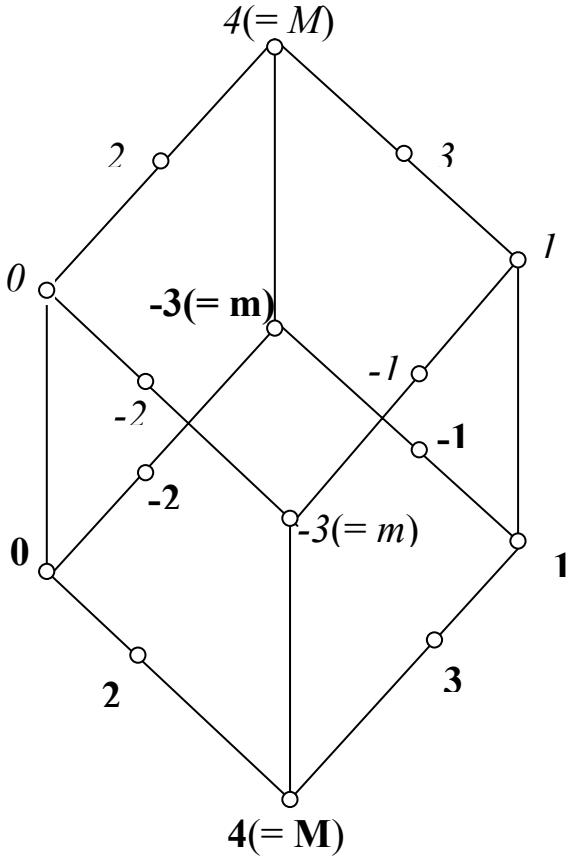
$S_{[-1,1]}$



$S_{[-1,2]}$



$S_{[-3,4]}$



(30) For all x, y, z$\in S_{[i,j]}$, (i≥-1,j≤2): x&(y∨z) ⇒ (x&y)∨z.

There are no other distributive model structures according to definition (28)[26]. The only distributive model structure with a Boolean complement is $S_{[-1,0]}$. Hence, it is the only Boolean lattice among our model structures.

Interpreting the connectives '¬', '&' and '∨' (as well as propositional quantification) by means of complement, infimum and supremum seems straightforward. In order to interpret '→', however, we need an additional operation (note that ⇒ is a *relation* between values, not an *operation* assigning values to pairs of values).

When interpreting '→', we would certainly like to respect the constraint (3), reformulated here as

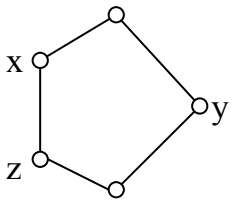
(3') $\ulcorner A \rightarrow B \urcorner$ is true$_k$ iff $I_k(A) \Rightarrow I_k(B)$,

where a formula is true$_k$ iff its value in V_k has a positive polarity and 'I_k' is an interpretation function assigning elements of V_k to formulae[27]. Still, this constraint leaves several possibilities for operation →. Some simple solutions would require that implications always get maximal or minimal degree of entrenchment, i.e., $I_k(A \rightarrow B) \in \{M_k, m_k\}$ if $I_k(A) \Rightarrow I_k(B)$ and $I_k(A \rightarrow B) \in \{\mathbf{M}_k, \mathbf{m}_k\}$, else[28]. - Here I opt for another proposal:

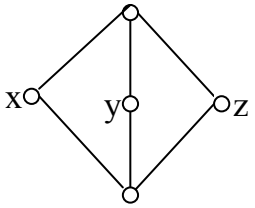
(31) For all p, q ∈ {+,-}, d, e ∈ *k*, x = <p, d>, y = <q, e>:

---

[26] If i<-1 or j>2, then $\{\mathbf{M}_k, \mathbf{m}_k, -1, m_k, M_k\}$ or $\{m_k, 0, 1, 2, M_k\}$ will be a sub-lattice of $S_{[i,j]}$. Due to a theorem of Dedekind, lattices with sub-lattices of this type (called N_5) are known to be non-modular; they contain elements x, y and z with z ⇒ x but without x&(y∨z) ⇒ (x&y)∨z. (In case i<-1, let x = *-1*, y = $\mathbf{m}_k$ and z = m_k; in case j>2, let x = *2*, y = *1* and z = *0*.) It is well-known that non-modularity implies non-distributivity. Interestingly, the only "source" of non-distributivity of model structures according to definition (27) is non-modularity. The other possible "source" which (due to a theorem of Birkhoff) is having a sub-lattices of another type (called M_5) does not occur. We have x = y = z, if x&y = y&z = x&z and x∨y = y∨z = x∨z (for distributive *vs*. non-distributive lattices, cf. Burris & Sankappanavar 1981, 12ff).



N_5



M_5

[27] Like '|–', 'I' (with or without subscript) is taken to absorb (quasi-) quotes around its argument.

[28] This would give us the paradox (14), whenever $\ulcorner A \urcorner$ is an implication.

(a)(i) $\mathbf{d}\rightarrow\mathbf{d} = d$,
(a)(ii) $\mathbf{d}\rightarrow d = \mathbf{M}_k$, if $d<0$,
else: (b)(i) $x\rightarrow y = y$, if $d\leq e$ and either $x\Rightarrow y$ or $x\Rightarrow\neg y$,
(b)(ii) $x\rightarrow y = \neg x$, if $e<d$ and either $x\Rightarrow y$ or $\neg x\Rightarrow y$,
else: (c) $x\rightarrow y = \mathbf{M}_k$.

Now, I define a *model for* $L_{\rightarrow,\forall}$ *based on k* as a model structure, together with an interpretation of $L_{\rightarrow,\forall}$ in *k* (a function from $L_{\rightarrow,\forall}$ to V_k):

(32) $M_k = \langle S_k, I\rangle$, where S_k is a model structure and for all formulae A, B and propositional variables $p \in L_{\rightarrow,\forall}$:
(a) $I(p) \in V_k$,
(b) $I(\neg A) = \neg I(A)$,
(c) $I(A\&B) = I(A)\&I(B)$,
(d) $I(A\rightarrow B) = I(A)\rightarrow I(B)$ and
(e) $I(\forall pA) = inf\{I'(A)$: I' = I with the possible exception of $I(p)\}$.

Finally, we define $\ulcorner A\urcorner \in L_{\rightarrow,\forall}$ as *true in model* $M_k = \langle S_k, I\rangle$ iff I(A) has positive polarity. $\ulcorner A\urcorner$ is *valid with respect to k* ($\mid\!\!-_k$ A) iff it is true in each model M_k. It is a *logical consequence of* $\ulcorner B_1\urcorner, \ldots, \ulcorner B_n\urcorner$ *with respect to k* ($B_1, \ldots, B_n \mid\!\!-_k A$) iff $\mid\!\!-_k \neg(B_1\&\ldots\&B_n\&\neg A)$. $\ulcorner A\urcorner$ is *valid* (*simpliciter*) ($\mid\!\!- A$) iff it is valid with respect to each entrenchment structure (hence, true in all models) and a *logical consequence* (*simpliciter*) *of* $B_1, \ldots, B_n$ ($B_1, \ldots, B_n \mid\!\!- A$) iff $\mid\!\!- \neg(B_1\&\ldots\&B_n\&\neg A)$.

For each entrenchment structure *k*, validity with respect to *k* is decidable.[29] Each *k* is finite, and so is lattice S_k. For each model structure,

[29] Whether validity (*simpliciter*) is also decidable, I do not know. Likewise, I am not aware of any precise syntactic (proof-theoretical) characterisation of the logic presented here. Obviously, lattice-based postulates ($\mid\!\!-$ A&B $\rightarrow$A, $\mid\!\!-$ A&B $\rightarrow$B, $\mid\!\!-$ (A$\rightarrow$ B)&(A$\rightarrow$C) $\rightarrow$ (A$\rightarrow$ B&C)) and postulates based on De Morgan complements ($\mid\!\!-$ (A$\rightarrow\neg$C)$\rightarrow$(C$\rightarrow\neg$A) and $\mid\!\!-$ $\neg\neg$A$\rightarrow$A), as well as *reductio ad absurdum* ($\mid\!\!-$ (A$\rightarrow\neg$A)$\rightarrow$ $\neg$A) hold. Likewise, the rules of adjunction (if $\mid\!\!-$ A and $\mid\!\!-$ B, then $\mid\!\!-$ A&B) and implicational detachment (if $\mid\!\!-$ A and $\mid\!\!-$ A$\rightarrow$B, then $\mid\!\!-$ B). As for pure implicational nestings, I conjecture that permutation ($\mid\!\!-$ A$\rightarrow$((A$\rightarrow$B)$\rightarrow$B)), contraction ($\mid\!\!-$ (A$\rightarrow$(A$\rightarrow$B))$\rightarrow$(A$\rightarrow$B)) and transitivity ($\mid\!\!-$ (A$\rightarrow$B)$\rightarrow$((B$\rightarrow$C)$\rightarrow$(A$\rightarrow$C))) hold. If I am right, entrenchment logic for $L_{\rightarrow}$ (without propositional quantification) comprises

the operations defined can be rendered as finite matrices. Propositional quantification is reducible to finite conjunction and complement. Hence, we can check, in a finite number of steps, whether for all assignments of values to the free variables in a formula, the entire formula receives a positive value of V_k[30]. If so, the formula is valid with respect to k, otherwise it is not.

Restricting ourselves to $L_{\rightarrow}$ (i.e., $L_{\forall,\rightarrow}$ without quantification), we may note that, given $k_1 \subseteq k_2$, any formula is valid with respect to k_1 if it is valid with respect to k_2.[31] The converse does not hold. E.g., $\ulcorner\neg A\&A \rightarrow C\vee\neg C\urcorner$ is valid with respect to [-1,1], but not with respect to [-1,2].

In the sequel – especially when discussing two outstanding puzzles in Goodman's analysis of counterfactuals in Section 6 – I will mainly be concerned with $S_{[-3,4]}$. By relying (strategically) on the decidability result just mentioned, I will provisionally restrict the notion of validity to validity with respect to entrenchment structure [-3,4].[32] Henceforth, '|–' strictly

---

the system of relevance logic R *minus* distribution (*|–A&(B∨C) → (A&B)∨C) but includes material detachment (also called 'rule γ': if |– A and |– A ⊃ B, then |– B. In view of the fact that we do not have (14) or the mingle axiom *|–A→(A→A), the logic of entrenchment looks very R-ish indeed. Note that rule γ is known to be admissible in R (including distribution) (Anderson & Belnap 1975, 300ff) and that R *minus* distribution is known to be decidable (Anderson & Belnap 1992, 335) but does not seem to license rule γ (Anderson & Belnap 1975, 313). Note that lack of distribution has consequences for propositional quantification: of the five postulates PQ1-PQ5 listed by Dunn (Anderson & Belnap 1992, 19), we lose the last one: *|–$\forall p$(A∨B)→ A∨$\forall p$B (p not free in A).

[30] Of course, this becomes rather tedious as the numbers of free propositional variables in a formula, (possibly nested) quantifiers or elements of k increase. You better write a programme and let your computer do the job.

[31] Assume that $k_1 \subseteq k_2$ and some arbitrary formula $\ulcorner A\urcorner$ ($\in L_{\rightarrow}$) is not valid with respect to k_1. Then, there is a model $M_{k1} = \langle S_{k1}, I_1\rangle$ such that $\ulcorner A\urcorner$ is not true in M_{k1}. Now, consider $M_{k2} = \langle S_{k2}, I_2\rangle$ such that for each propositional variable p, $I_2(p)$ and $I_1(p)$ have identical polarity and $I_2(p)$ has maximal (minimal) degree in k_2, if $I_1(p)$ has maximal (minimal) degree in k_1. Otherwise, let $I_2 = I_1$. Obviously, $\ulcorner A\urcorner$ is not true in M_{k2}, and hence, is not valid with respect to k_2. By contraposition, $\ulcorner A\urcorner$ is valid with respect to k_1 if it is valid with respect to k_2.

[32] This is also due to my ignorance confessed in footnote 29.

means '$|\!-_{[-3,4]}$'. (I hope that this simplification will not seriously affect my arguments.[33])

Given I(A) and I(B), the values I(¬A), I(A&B) and I(A∨B) can easily be read from the Hasse diagrams. Calculation of I(A→B) is not that easy. For ease, I give the matrix induced by (30) for $S_{[-3,4]}$[34], where empty cells are occupied by **4** (= **M**):

→	**4**	**3**	**2**	**1**	**0**	**-1**	**-2**	**-3**	*-3*	*-2*	*-1*	*0*	*1*	*2*	*3*	*4*
4	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*
3		*3*		*3*		*3*		*3*					*3*		*3*	*4*
2			*2*		*2*		*2*	*2*				*2*		*2*		*4*
1		**3**		*1*		*1*		*1*					*1*		*3*	*4*
0			**2**		*0*		*0*	*0*				*0*		*2*		*4*
-1						*-1*		*-1*								*4*
-2							*-2*	*-2*								*4*
-3								*-3*								*4*
-3		**3**	**2**	**1**	**0**	**-1**	**-2**	**-3**	*-3*	*-2*	*-1*	*0*	*1*	*2*	*3*	*4*
-2			**2**		**0**		**-2**	**-2**		*-2*		*0*		*2*		*4*
-1		**3**		**1**		**-1**		**-1**			*-1*		*1*		*3*	*4*
0			**2**		**0**		**0**	**0**				*0*		*2*		*4*
1		**3**		**1**		**1**		**1**					*1*		*3*	*4*
2			**2**		**2**		**2**	**2**				**2**		*2*		*4*
3		**3**		**3**		**3**		**3**					**3**		*3*	*4*
4																*4*

[33] In the sequel, proofs as to the *invalidity* of a principle with respect to [-3,4] will be given by indication (mainly in footnotes) of refuting valuations of the occurring propositional variables. Obviously, such a refutation also proves that the principle in question is invalid *simpliciter*. The proof method I am using for the *validity* of a principle with respect to [-3,4] is *brute force*: calculation of each valuation. (Since in model structure $S_{[-3,4]}$ we are working with a 16-valued logic, I should vary the common dictum "The proof is left to the reader": *the proof is left to the computer*). This, of course, only supports, but does not prove, that the principle is valid *simpliciter*.

[34] Note that the matrices for $S_{[-1,2]}$ induced by the operations defined are just Anderson & Belnap's for M_0 (Anderson & Belnap 1975, 252f).

The /-diagonal in this matrix (from <*4*, **4**> to <**4**, *4*>) is a symmetry axis for values (i.e., for both polarities and degrees). Hence, we have (33) as well as its implicational version (33'):

(33) $A\rightarrow C \mid\!- \neg C\rightarrow \neg A$ (33') $\mid\!- (A\rightarrow C)\rightarrow(\neg C\rightarrow \neg A)$

By similar considerations, we may, in many cases, show a parallelism between logical consequence and validity of implication, e.g.:

(34)(a) $A \mid\!- A$ (34')(a) $\mid\!- A \rightarrow A$

(b) $A\&(B\vee A) \mid\!- (A\&B)\vee A$ (b) $\mid\!- A\&(B\vee A) \rightarrow (A\&B)\vee A$

(c) $(A\rightarrow C)\&A \mid\!- C$ (c) $\mid\!- (A\rightarrow C)\&A \rightarrow C$

(d) $A\rightarrow B \mid\!- (B\rightarrow C)\rightarrow(A\rightarrow C)$ (d) $\mid\!- (A\rightarrow B)\rightarrow((B\rightarrow C)\rightarrow(A\rightarrow C))$

However, we cannot generally (as in these cases) make a single premise the antecedent of its conclusion in order to state a valid implication. Since $\ulcorner A\&\neg A\urcorner$ is false in each model, we have (35). However, the corresponding implication (14) is not valid:

(35) $A\&\neg A \mid\!- C$ *(14) $\mid\!- A\&\neg A \rightarrow C$

As stressed above, logical consequence is not a sub-relation of law-like (implicational) consequence in the logic presented here.

More generally, although we have rule (36), we don't have rule (36'):

(36) If $A_1, \ldots, A_n, A \mid\!- C$, then $A_1, \ldots, A_n \mid\!- A \supset C$

*(36') If $A_1, \ldots, A_n, A \mid\!- C$, then $A_1, \ldots, A_n \mid\!- A\rightarrow C$

Cases in point are *disjunctive syllogism*, (*truth-functional*) *modus ponens* and so-called (*implicational*) *antilogism*.[35]

(37)(a) $A\vee C, \neg A \mid\!- C$ *(37')(a) $A\vee C \mid\!- \neg A\rightarrow C$

(b) $A \supset C, A \mid\!- C$ (b) $A \supset C \mid\!- A\rightarrow C$

(c) $S\&A \rightarrow C, S\&\neg C \mid\!- \neg A$ (c) $S\&A \rightarrow C \mid\!- S\&\neg C \rightarrow \neg A$

In view of (37)(b) and (37')(b), we note (in comparison with (34)(c) and (34')(c)) that, whereas Parry is quite right in maintaining that $\ulcorner A \supset B\urcorner$ is the weakest *truth-functional* link between some premise $\ulcorner A\urcorner$ and an arbitrary conclusion $\ulcorner C\urcorner$, it is not a *law-like* link.

---

[35] To refute (37')(a)-(c) choose, e.g., the values **3**, *-3* and **2** for 'A', 'C' and 'S', respectively.

5. Re(de)fining '〈'

Having identified (14), the paradox of law-like implication, as the principle source of Goodman's troubles with (2) and having recommended a paradox-free notion of law-like implication as a remedy, let us reconsider (2).

For <*d*, **e**>, (d, e ∈ [-3,4]) we arrive a the following sub-matrix for $S_{[-3,4]}$:

	4	**3**	**2**	**1**	**0**	**-1**	**-2**	**-3**
-3	**-3**	**-3**	**-3**	**-3**	**-3**	**-3**	**-3**	**-3**
-2	**-3**	**-3**	**-3**	**-3**	**-3**	**-3**	**-3**	**-3**
-1	**-3**	**-3**	**-3**	**-3**	**-3**	**-3**	**-3**	**-3**
0	**-1**	**-1**	**-3**	**-1**	**-3**	**-1**	**-3**	**-3**
1	**-2**	**-3**	**-2**	**-3**	**-2**	**-3**	**-2**	**-3**
2	**-1**	**-1**	**-3**	**-1**	**-3**	**-1**	**-3**	**-3**
3	**-2**	**-3**	**-2**	**-3**	**-2**	**-3**	**-2**	**-3**
4	**4**	**3**	**2**	**1**	**0**	**1**	**0**	**-3**

<*4*, *e*> (e ∈ [-3,3]) is assigned *e*. In the remaining cases, we obtain value *4*.

This is unsatisfactory for at least three reasons.

First and foremost, counterfactuals with a false antecedent or true consequent (e.g., semifactuals) turn out as true. We not only have (15) above, but also (38) and (39), the converse of (12):

*(38) $C \vdash A \langle C$

*(39) $A \supset C \vdash A \langle C$ (12) $A \langle C \vdash A \supset C$

Of course, the result that the counterfactual and material conditional of ⌜A⌝ and ⌜C⌝ are logical consequences of each other is truly the Armageddon of any account reconstructing Goodman's intuition about the enthymematic character of counterfactuals. Things could hardly be worse. Refinement of definition (2) is urgently needed.

Second, the unrefined definition is at odds with Goodman's intuition that counterfactual and factual conditionals are *mere transpositions* of each other.

Note, however, that in fact, the definiens of (2) licenses contraposition (9). Since the /-diagonal in the matrix just described is a symmetry-axis *with respect to polarities*, we have

(40) $\exists S.\ S\ \&\ (S\&A \rightarrow C) \mid\!- \exists S.\ S\ \&\ (S\&\neg C \rightarrow \neg A)$.

However, the implicational strengthening of (40) does not hold.

If we then take Goodman's intuition to imply that $\ulcorner A \langle C \urcorner$ and $\ulcorner \neg C \langle \neg A \urcorner$ should have identical values, i.e., not only the same *truth* value but the same *degree of entrenchment* as well, we have a second (independent) reason to be dissatisfied with (2).

A definition of '$\langle$' should be required to exhibit a matrix with a /-diagonal that is a symmetry-axis for values. In other words, we would like to license the implicational version of (9), i.e.,

(9') $\quad \mid\!- (A \langle C) \rightarrow (\neg C \langle \neg A)$.

At this point, we observe that – given the presumption (rejected here) that logical consequence is a sub-relation of law-like consequence – Goodman's intuition that contrapositives of counterfactuals preserve implicational value (are *mere transpositions*) could be justified by implicational antilogism ((37')(c) above). If $\ulcorner S\&A \rightarrow C \urcorner$ and $\ulcorner S\&\neg C \rightarrow \neg A \urcorner$ *were* indeed logical consequences of each other, this would license (9') under the definition (2). But, as we have seen, they are not (according to our explication of '$\rightarrow$').

Of course, (9') could be gained cheaply by conjoining $\ulcorner \exists S.\ S\&\neg C \rightarrow \neg A \urcorner$ to the definiens of (2). But this move would not help against (15) or (38). Therefore, a more adequate adjustment, one which takes Goodman's idea seriously, seems to be making the assumption of implicational antilogism (implicit in Goodman's reasoning) explicit in the revised definition. Accordingly, I take it that the antecedent and consequent of a counterfactual have to be such that there is no true *S* such that $\ulcorner S\&A \rightarrow C \urcorner$ and $\ulcorner S\&\neg C \rightarrow \neg A \urcorner$ have different polarities[36]. This leads to the addition of $\ulcorner \forall S.S \supset ((S\&A \rightarrow C) \equiv (S\&\neg C \rightarrow \neg A)) \urcorner$ to the definiens in (2).

Third, we observe that <**m**, *m*> is assigned *M* according to the matrix described above. Hence, it assigns triviality to the pair of mere falsity (the

[36] Note that in order to guarantee identity of implicational value under contraposition, I am somewhat imitating Goodman's strategy of forcing connexivity (i.e., Boethius' thesis). Like him, I am adding a negative clause $\ulcorner \neg \exists S.S\&\neg G[A,C] \urcorner$ to the core-explanation $\ulcorner \exists S.S\&F[A,C] \urcorner$ which accounts for the enthymematic character of counterfactuals.

supremum of the set of negative values) and mere truth (the infimum of the set of positive values). This amounts to the claim, that everything true would be the case, if anything false were the case – and even trivially so. I urge that the claim is absurd and that <**m**, *m*> should be assigned **M**.

I fix this point by defining a constant '*t*' (for 'mere truth') as '$\forall p.p \rightarrow p$' and add '$(C \rightarrow t) \vee (\neg t \rightarrow A) \rightarrow (A \rightarrow C)$' to the definiens of (2).

We arrive at the following redefinition of '$\langle$':

(41) $A \langle C := \exists S.S\&(S\&A \rightarrow C)\ \&$
$\forall S.S \supset ((S\&A \rightarrow C) \equiv (S\&\neg C \rightarrow \neg A))\ \&$
$(C \rightarrow t) \vee (\neg t \rightarrow A) \rightarrow (A \rightarrow C)$

The emerging $\langle$-matrix for [-3,4][37] looks like this (empty cells are occupied by **4**):

[37] For the distributive model-structures $S_{[-1,0]}$, $S_{[-1,1]}$ and $S_{[-1,2]}$ (cf. footnote 26), the $\rightarrow$-matrix and the $\langle$-matrix coincide. Hence, we have: If $\vdash_k A\&(B \vee C) \rightarrow (A\&B) \vee C$, then $\vdash_k (A \langle C) \rightarrow (A \rightarrow C)$. I conjecture that the converse holds as well.

⟨	**4**	**3**	**2**	**1**	**0**	**-1**	**-2**	**-3**	*-3*	*-2*	*-1*	*0*	*1*	*2*	*3*	*4*
4	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*	*4*
3		*3*		*3*		*3*		*3*			**3**		*3*		*3*	*4*
2			*2*		*2*		*2*	*2*		**2**		*2*		*2*		*4*
1		**1**		*1*		*1*		*1*			**1**		*1*		*3*	*4*
0			**0**		*0*		*0*	*0*		**0**		*0*		*2*		*4*
-1		*3*		**1**		*-1*		*-1*			*4*		**1**		**3**	*4*
-2			*2*		**0**		*-2*	*-2*		*4*		**0**		**2**		*4*
-3								*-3*								*4*
-3		**1**	**0**	**1**	**0**	**-3**	**-3**	**-3**	*-3*	*-2*	*-1*	*0*	*1*	*2*	*3*	*4*
-2			**0**		**0**		**-3**	**-3**		*-2*		*0*		*2*		*4*
-1		**1**		**1**		**-3**		**-3**			*-1*		*1*		*3*	*4*
0			**0**		**0**		**0**	**0**		**0**		*0*		*2*		*4*
1		**1**		**1**		**1**		**1**			**1**		*1*		*3*	*4*
2			**0**		**0**		**0**	**0**		*2*		**0**		*2*		*4*
3		**1**		**1**		**1**		**1**			*3*		**1**		*3*	*4*
4																*4*

Here, the /-diagonal is a symmetry axis with respect to values (both polarities and degrees). Hence, (9') is licensed.

As can be seen by comparison, the ⟨-matrix retains all values with positive polarity from the →-matrix. Hence, we have (13), repeated here. As can easily be verified by inspection, its implicational strengthening (13') holds as well:

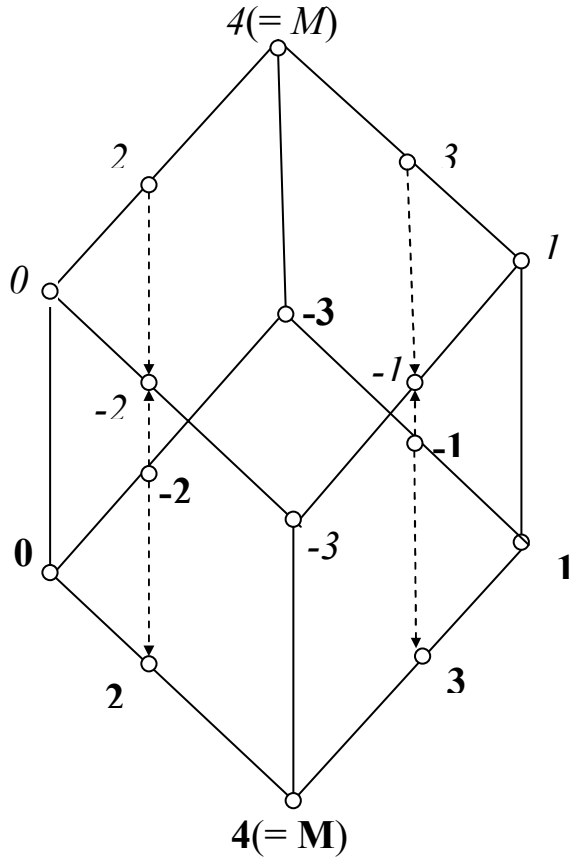
(13) $A \rightarrow C \vdash A \langle C$ (13') $\vdash (A \rightarrow C) \rightarrow (A \langle C)$

We also have (16)-(18) above[38].

[38] Other postulates are immediately derivable by means of (12) and (13) from the postulates (without stars) mentioned in footnote 29 (all of which are demonstrably valid with respect to [-3,4]). E.g., $\vdash (A \rightarrow B) \& (A \rightarrow C) \langle (A \rightarrow B \& C)$, $(A \rightarrow B) \& (A \rightarrow C) \vdash A \langle (B \& C)$, $\vdash (A \rightarrow \neg C) \langle (C \rightarrow \neg A)$, $A \rightarrow \neg C \vdash C \langle \neg A$, $\vdash \neg\neg A \langle A$, $\vdash (A \rightarrow \neg A) \langle \neg A$, $\vdash A \langle ((A \rightarrow B) \rightarrow B)$, $A \vdash (A \rightarrow B) \langle B$, $\vdash (A \rightarrow (A \rightarrow B)) \langle (A \rightarrow B)$, $A \rightarrow (A \rightarrow B) \vdash A \langle B$, $\vdash (A \rightarrow B) \langle ((B \rightarrow C) \rightarrow (A \rightarrow C))$, $A \rightarrow B \vdash (B \rightarrow C) \langle (A \rightarrow C)$ and $A \rightarrow B, B \rightarrow C \vdash A \langle$

However, the converse of (13) (let alone (13')) does not hold, since **-1**⟨**3**, **-1**⟨*-1*, **-2**⟨**2**, **-2**⟨*-2*, *2*⟨*-2* and *3*⟨*-1* are positive values and the corresponding implications have negative polarity.

Let us indicate the additional links between values (*c-involvements*, for convenience) by arrows in the Hasse diagram for $S_{[-3,4]}$.



All involvements are c-involvements. However, whereas the former "chain up", the latter do not. Although, e.g., **0** c-involves **-2** and **-2** c-involves *–2*, **0** neither involves or c-involves *-2* (nor *vice versa*). Note also that *-2* and *2* are different although they c-involve each other. Hence, c-involvement is neither transitive nor antisymmetric. It is not monotonic either; it does not allow for 'conjunctive strengthening' of the first argument. E.g., **-1** c-involves *-1*, but **-1**&*1* (= **1**) does not.

Because of constraint (3'), reformulated here as

C. We also have the rule of counterfactual detachment: if |– A and |– A ⟨ B, then |– B, which is derivable by means of A ⟨ B |– A ⊃ B and material detachment: if |– A and |– A ⊃ B, then |– B.

(3'') $\ulcorner A \rightarrow B \urcorner$ is true in model $M_k = \langle S_k, I \rangle$ iff $I(A) \Rightarrow I(B)$,

law-like implication clearly inherits transitivity and monotonicity from involvement, but counterfactuals do not. We have (42) and (43) but neither (42') or (43'):

(42) $A \rightarrow B, B \rightarrow C \mid\!- A \rightarrow C$ *(42') $A \langle B, B \langle C \mid\!- A \langle C$.

(43) $A \rightarrow C \mid\!- A \& B \rightarrow C$ *(43') $A \langle C \mid\!- A \& B \langle C$.

Not yet explicitly discussed by Goodman in *FFF*, the features of transitivity and monotonicity have gained prominence in recent investigations of counterfactuals. It has widely been accepted that counterfactuals diverge from other consequential concepts like (material, strict or law-like) implication and logical consequence for being non-transitive and non-monotonic.

The following examples witnessing the non-validity of the principles of transitivity and monotonicity are due to Stalnaker (1968) and Lewis (1973), respectively.

(44) *If J. Edgar Hoover had been born a Russian, then he would have been a Communist.*
If he had been a Communist, he would have been a traitor.

∴ If he had been born a Russian, then he would have been a traitor.

(45) *If Otto had come, it would have been a lively party; but if both Otto and Anna had come, it would have been a dreary party.*

It seems to me that examples against transitivity are somewhat strained compared to those against monotonicity. To explain why is hard.[39] For note that abandoning monotonicity while retaining transitivity is no easy job. Remember that (17) ($\mid\!- A \& B \langle A$) is valid. Hence, since

*(46) $A \& B \langle A, A \langle C \mid\!- A \& B \langle C$

is an instance of (42'), cutting the first premise of (46) immediately leads to monotonicity (43').

[39] I suppose, it would involve us in a deeper analysis of the situated character of counterfactual reasoning and, perhaps, the explication of non-official notions of logical consequence (cf. footnote 6 above). I leave that for another occasion.

As noted above, non-exclusion of counterlegals and counteridenticals (with self-incompatible antecedents) is at odds with Goodman's use of Boethius' thesis as a criterion of adequacy. Actually, definition (41) does not license (23), which is repeated here[40]:

*(23) $A \langle C \vdash \neg(A \langle \neg C)$.

Even weakening (23) by adding the requirement that $\ulcorner A \urcorner$ be self-compatible does not help[41]:

*(47) $\neg(A \rightarrow \neg A), A \langle C \vdash \neg(A \langle \neg C)$.

This is what Stalnaker (1968) and Lewis (1973) assume for their notions of counterfactual conditionals (they interpret '$\rightarrow$' as strict implication, i.e., the necessitation of material implication[42]). For them, all counterfactuals with impossible antecedent are true.

If, however, we add the stronger requirement that $\ulcorner A \urcorner$ be self-cotenable, we obtain a valid principle[43]:

(48) $\neg(A \langle \neg A), A \langle C \vdash \neg(A \langle \neg C)$.

In Section 3, I proposed counterfactual *reduction ad absurdum* (26) as a substitute for Aristotle's thesis (24), both repeated here[44]:

*(24) $\vdash \neg(A \langle \neg A)$ (26) $A \langle \neg A \vdash \neg A$

---

[40] This can readily be seen in the $\langle$-matrix above: except <**m**, *m*>, the upper halves of the \-diagonal and the /-diagonal are exclusively occupied by positive values. Hence, $\ulcorner (A \langle A) \,\&\, (A \langle \neg A) \urcorner$ is satisfiable and Boethius' thesis is not valid.

[41] To refute (47), choose a weak negative value (except **m**) for $\ulcorner A \urcorner$ and a value with an identical degree for $\ulcorner C \urcorner$.

[42] Given this interpretation, $\ulcorner \neg(A \rightarrow \neg A) \urcorner$ is equivalent with $\ulcorner \Diamond A \urcorner$ ($\ulcorner$it is possible, that $A \urcorner$).

[43] "Cotenable" and "self-cotenable" are Goodman's coinage. He defines (*FFF*, 15) $\ulcorner A \urcorner$ to be *cotenable* with $\ulcorner C \urcorner$, iff it is not the case that $\ulcorner C \urcorner$ would not be true if $\ulcorner A \urcorner$ were, i.e., $\ulcorner \neg(A \langle \neg C) \urcorner$. Note that on our account, we have $A \vdash \neg(A \langle \neg A)$. Hence, we also have $A, A \langle C \vdash \neg(A \langle \neg C)$. Another valid, perhaps more interesting principle is $\neg(A \rightarrow \neg A), A \langle C \& \neg C, A \langle C \vdash \neg(A \langle \neg C)$.

[44] To refute (24), choose any negative value (except **m**) for $\ulcorner A \urcorner$. As for (26), observe that positive values on the /-diagonal of the $\langle$-matrix are exclusively situated on its upper half.

Note that we neither have the converse of (26) or its implicational strengthening[45]:

*(49) $\neg A \vdash A \langle \neg A$ *(26') $\vdash (A \langle \neg A) \rightarrow A$

As for the converse of Boethius' thesis, we note that it is equivalent with so-called Conditional Excluded Middle (CEM) (cf. Bennett 2003, 183).

*(50) $\vdash (A \langle C) \vee (A \langle \neg C)$.

The validity of (50) has been a matter of controversy between Stalnaker (1968) and Lewis (1973). Consider the following:

(51) *If Bizet and Verdi were compatriots, Bizet would be Italian.*
(51') *If Bizet and Verdi were compatriots, Bizet would not be Italian.*

Lewis thinks that both (51) and (51') are false. *Neither* can we say that possible circumstances with the two composers being Italian are more similar to the facts than circumstances with the two composers being French[46] *nor vice versa*. Hence, for Lewis, (50) is not valid. Stalnaker thinks that both (51) and (51') lack truth value, because there are no possible circumstances with both composers of identical nationality qualifying as *the* circumstances which are most similar to the facts. However, Stalnaker maintains that each *choice* of such circumstances *as most similar* resolves the question of whether Bizet is Italian or not. By this move, using techniques of supervaluation (van Fraassen 1966), Stalnaker may retain (50) as valid.

Without much argument, Goodman considers counterfactuals with identical antecedent and opposing consequent as "contraries" such that both may be false (cf. footnote 15). Hence, Goodman rejects (50).

As indicated by the star, the present account takes sides with Goodman and Lewis against Stalnaker.[47]

---

[45] (49) is refuted by **m** for $\ulcorner A \urcorner$, (26') by each weak positive value and each negative value (except **m**) for $\ulcorner A \urcorner$.

[46] I assume that among all the possible circumstances with Bizet and Verdi being compatriots and Bizet or Verdi not being Italian those with both composers being French are most similar to the facts.

[47] In order to refute (50), choose, e.g., **-1** or *3* as value for $\ulcorner A \urcorner$ and **1** as value for $\ulcorner C \urcorner$.

One last remark concerning the Stalnaker-Lewis account of counterfactuals should be made. As mentioned above, Lewis and Stalnaker not only reject (42') and (43') – as we do – but also contraposition (9), which – following Goodman – we accept. As a consequence, we (but not they) have to reject (52) along with (43'):

*(52) $A \langle C \mid\!\!- A \langle C \vee B$

Perhaps, one might cite examples like (53):

(53) *If Bernd was born in Berlin, he would have been born in East Berlin.*

∴ If he was born in Berlin, he would have been born in East Berlin or in East-Africa.

I am hesitant to claim that fallacies like (53) are as forceful a case against (52) as examples of type (45) are against (43'). But if they were, this would be a problem for the Stalnaker-Lewis-account and an argument for Goodman's basic intuition.

6. Goodman's two tough problems

6.1 Bernd born in Berlin

Assume that Bernd was born in Chemnitz, GDR, in 1962. At that time, Berlin was divided by the wall in a western part that belonged to the FRG[48] and an eastern part, which was the capital of the GDR. The division was complete and without overlap. Obviously, the counterfactual assumption that Bernd was born in Berlin is compatible both with him not being born in West Berlin and with his not being born in East Berlin, but it is incompatible with him being born *neither* in West *or* in East Berlin. Additionally, the counterfactual assumption that Bernd was born in West Berlin or in East Berlin is incompatible with him not being born in

[48] At least, I take it this way here. As a matter of fact, the status of West Berlin was controversial under international law in 1962. According to the doctrine of the GDR-government, West Berlin was no part of the FDG but a 'special political unit' 'on the territory' of the GDR.

Berlin.[49] In other words, by using appropriate abbreviations[50], we assume, among other things, the following:

(54) $\neg((B\rightarrow W)\vee(B\rightarrow E))\&(B\leftrightarrow W\vee E)$.

Now, think of (55) as a counterfactual about Bernd.

(55) $B \langle E$

Consider '¬W', which is certainly true, since Bernd was born in Chemnitz. Furthermore, '¬W&B' does not imply each proposition (as intended by Goodman when requiring self-compatibility[51]) and '¬W' alone neither implies 'E' or '¬E'. Hence, '¬W' seems to be – according to Goodman's original criteria in the positive part of his explanation[52] – a reasonable candidate for support. Indeed, 'B' and one half of the last conjunct of (54) (the direction from left to right) lead, *via* implicational *modus ponens* (34)(c), to 'W∨E', from where we reach (using '¬W' and *disjunctive syllogism* (37)(b)) 'E' as a logical consequence:

(56) $B\rightarrow W\vee E, \neg W\&B \vdash E$.

For Goodman, this seemed sufficient to license (55). Given ¬W, there is a true proposition, which together with the antecedent of (55) renders its consequent.

At this point, however, Goodman's first tough problem arises, if we realise that we could equally choose '¬E' for support. Using

(56') $B\rightarrow W\vee E, \neg E\&B \vdash W$,

we arrive by a strictly parallel argument at

(55') $B \langle W$.

---

[49] This example mimics Goodman's case of *Jones in Carolina* (*FFF*, 11).

[50] 'B', 'W', 'E' are to represent that Bernd was born in 1962 in Berlin, in West-Berlin and in East-Berlin, respectively.

[51] As noted above in section 2, Goodman did not (and because of his acceptance of (14) he could not) differentiate between self-incompatibility of a sentence (implying its own negation) and absurdity (implying each proposition).

[52] I.e., without the amendment answering to Parry's criticism (see Section 2).

With respect to a slightly different example of the same brand, Goodman suggests that his analogues to (55) and (55') "cannot both be true" (*FFF*, 11). At least, it is strange enough to suppose that the innocent fact of a man being born in Chemnitz in 1962 should licence the eccentric counterfactual that if he was born in Berlin, his birth would have happened in the intersection of the western and the eastern sector – an intersection that (as far as I know) never existed.[53]

Goodman used this example to motivate the connexive twist of his analysis, which led him – as we have seen in Section 3 – to the addition of the negative part in his explication. Parry (1957) argued that the difficulty with the example could be solved at a lower cost by strengthening the criteria for support, adding the requirement that the negations of antecedent and support should be compatible. This disqualifies '¬W' and '¬E' as support propositions, since their negations imply 'B' (W∨E →B). Goodman, although accepting Parry's amendment (if, however, for somewhat different reasons – see Section 3), argued that substitutes for '¬W' and '¬E' without violation of Parry's condition are "readily at hand" such that, as he writes, "the gate-crashers [(55) and (55')] are back at our party" (Goodman 1957, 444).

To give Goodman-style substitutes for (56) and (56'), we think of two true propositions[54]: (a) the assumption that Bernd was born both in Berlin *and* in the GDR implies that he was born in East Berlin, and (b) the assumption that he was born in Berlin *but* not in a capital implies that he was born in West Berlin. Obviously, since Bernd was born in a province town of the GDR, the second conjuncts in the antecedent of (a) and (b) are

[53] Taking (55) and (55') as contraries, Goodman silently assumes 'E→¬W' as true. No overlap of sectors where Bernd could have been born! Does Goodman accept the principle *B ⟨ W, B ⟨ E |– B ⟨ (W&E)? I don't know. Anyhow, the present account rejects it. (Use **-1**, **-1** and *-1* as values for ⌜B⌝, ⌜W⌝ and ⌜E⌝, respectively, to refute it). Also, we do not have *B ⟨ E, E→¬W |– B ⟨ ¬W. (Use **-1**, **3**, and **3** for ⌜B⌝, ⌜W⌝ and ⌜E⌝, respectively.) But note that given the special circumstances of the example, B ⟨ W, B ⟨ E, (54) |– B ⟨ (W&E) is licensed, if only for the fact that the premises are not jointly satisfiable.

[54] This mimics the examples in (Goodman 1957, 444).

true in the circumstances described above. Using appropriate abbreviations[55], we may state:

(57) B&O $\rightarrow$E, O&B |– E (57') B&C $\rightarrow$W, C&B |– W

Here, Parry's monition has been respected: neither 'O' nor 'C' implies 'B'. But again, we arrive at conflicting consequences.

In view of the present account, (56) and (56') as well as (57) and (57') are irreproachable. If 'W' and 'E' are incompatible, then this convincingly shows that 'B' is false (which we knew in advance). What they do *not* demonstrate – *contra* Goodman –, is that there is need for a condition in the explication of counterfactuals that forces connexivity. According to the present account, (56) or (57) simply do not license (55), nor do (56') or (57') license (55').

Look at (56) again. In order to license 'B $\langle$ E', we need for some proposition *S*, $\ulcorner S\&(S\&B \rightarrow E)\urcorner$. What we get is '$\neg$W&($\neg$W&B $\supset$ E)' (cf. (36)). This is not strong enough. The case (56') is analogous.

As for (57), observe that we *do* get 'O&(O&B $\rightarrow$E)', hence, as required, for some proposition *S*, $\ulcorner S\&(S\&B \rightarrow E)\urcorner$. But we need more. According to the second conjunct in the definies of (41), we should take care that for no true proposition *S*, 'B' and 'E' are such that $\ulcorner S\&B \rightarrow E\urcorner$ and $\ulcorner S\&\neg E \rightarrow \neg B\urcorner$ diverge in truth value. This is a requirement that Goodman, subsuming logical consequence under law-like consequence, could take for granted. In our reconstruction, however, we lack *implicational antilogism* (37')(c). Hence, we had to explicitly integrate the requirement in the definition.

Now, observe that the premises in (57) are not sufficient to guarantee the required transposition. Although we do in fact have (58), we do not obtain (58')[56]:

(58) B&O $\rightarrow$E, O&B |– O&$\neg$E $\supset$ $\neg$B
*(58') B&O $\rightarrow$E, O&B |– O&$\neg$E $\rightarrow\neg$B

[55] Let 'O' represent that Bernd is an *Ossi* (i.e., born in the GDR) and 'C'– within the bounds of this example – that Bernd was not born in the capital of any state.

[56] To refute (58'), take, e.g., *-2*, *3* and *3* as values for 'B', 'O' and 'E', respectively.

In other words, given the premises in (57), ‘O’ might very well be true but still such that ‘O&¬E’ does not imply ‘¬B’. Hence, (57) is not sufficient to license ‘B ⟨ E’. The case of (57’) is analogous.

Thus far, I have shown that Goodman’s troubles with his first tough problem evaporate, if we abandon the paradox of law-like implication.

But what about Bernd and the question of whether he would have been born in East Berlin, if he was born in Berlin?

We notice that the assumptions about Bernd in (54), far from licensing both (55) and (55’), even warrant their negations:

(59) (54) |– ¬((B ⟨ E) ∨ (B ⟨ W)).

It seems that we should content ourselves with the conclusion that *neither* would Bernd from Chemnitz have been born in East Berlin if he was born in Berlin *or* would he have been born in West Berlin, if he was born there.

Incidentally, note that by manipulating the premises, we can alter this result. First, let us skip one half of the last conjunct in (54) (the direction from right to left: ⌜E∨W →B⌝) to get (54’):

(54’) ¬((B→W)∨(B→E))&(B→ W∨E)

Second, add ⌜¬(E∨W)⌝. Obviously, the result is still a correct (though partial) description of Bernd’s situation (we may keep ‘B’ and ‘E’ as indicated above, but are free now to interpret ‘W’ as representing, say, that Bernd was born in the FRG[57]). Now, we may state:

(59’) ¬(E∨W), (54’) |– ¬((B ⟨ E) & (B ⟨ W))

(59’) allows the consist addition of one of the two counterfactuals, (55) or (55’), to its premises, from whence we get the negation of the other as a consequence. Thus, we have specified conditions which allow one of the counterfactuals to be true (and condemn the other one to be false).

Still, this is not satisfactory. Note that ⌜W∨E →B⌝ *is* true in the situation of our concern[58] and adding it to (54’) brings us back to the conclusion that neither of the two counterfactuals is true.

---

[57] I.e., ‘W’ represents the counterfactual assumption that Bernd is a *Wessi*.

[58] Under the original interpretation of ‘W’, of course.

In a way, our problem is dual to Goodman's. Where he was horrified by the prospect of having to warrant both of the conflicting counterfactuals, we have trouble accepting even one of them.

Let us briefly reflect about what might make us favour one counterfactual *vis à vis* the other. For me, the author of the example, I am thinking of Bernd from Chemnitz stereotypically as an *Ossi*, sharing this property with people being born in the former capital of the GDR in 1962. I am not, or at least not predominantly, concerned with him being born in a *non*-capital, a property he is actually sharing with people born in (at that time) the *non*-capital but otherwise rather metropolitan city of West Berlin. I am using the *obvious* (at least for me) proposition that Bernd is an Ossi to arrive at the conclusion that he would have been born in East Berlin if he was born in Berlin. I am doing so in the inferential mood of *indexical enthymematicity* (cf. footnote 3). I.e., what I non-enthymematically infer is that Bernd would have been born in East Berlin if it were the case that he was born in Berlin *and* was (as he actually is) an Ossi. In expressing my belief about Bernd, I am suppressing the true (and obvious) conjunct in the false antecedent. Of course, I might equally infer in the *indexical* mood that Bernd would have been born in West Berlin if he was born in Berlin. I might do so using the premise that Bernd was born in a non-capital. What I would infer non-enthymematically then is that Bernd would have been born in West Berlin if he was born in Berlin and (as he actually was) in a non-capital. In expressing this belief, I might suppress the true (and obvious) conjunct in the false antecedent. What I cannot infer validly in a non-enthymematic way is that Bernd would have been born in East Berlin (West Berlin) if he was born in Berlin. This counterfactual is inconsistent with the facts given. Likewise, I cannot infer *both* that Bernd would have been born in East Berlin (West Berlin) if he was born in Berlin as an Ossi *and* that he would have been born there if he was born in Berlin and in a non-capital. The conjunction of these counterfactuals is inconsistent with the facts, because the first and the negation of the second (the second and the negation of the first) are logical consequences of them.

What condition may hold for 'O' such that – given the facts about Bernd – we must conclude 'B&O $\langle$ E' and '$\neg$(B&O $\langle$ W)'? (By analogy, such a condition should license 'B&C $\langle$ W' and '$\neg$(B&C $\langle$ E)' for 'C').

I propose as a sufficient condition that 'O' is true[59] and such that its conjunction with 'B' both implies and is compatible with 'E' (By analogy, 'C&B' has to be compatible with and imply 'W'). Obviously, these conditions obtain in Bernd's situation. As required, we get:

(60) $O, B\&O \rightarrow E, \neg(B\&O \rightarrow \neg E), (54') \mid\!- (B\&O \langle E) \& \neg(B\&O \langle W)$

(60') $C, B\&C \rightarrow W, \neg(B\&C \rightarrow \neg W), (54') \mid\!- (B\&C \langle W) \& \neg(B\&C \langle E)$

In a way, we manage to have the cake and eat it, too. Strictly, we do not have either (55) or (55'). Enthymematically, we may have both.

It is tempting to transfer this kind of analysis to the case of the compatriot composers (cf. Section 5).

Using 'C_{bv}' for *Bizet and Verdi are compatriots*, 'I_b' for *Bizet is Italian*, 'I_v' for *Verdi is Italian* and 'F_b' for *Bizet is French*, we get

(61) $\neg(C_{bv} \vee I_b \vee (C_{bv} \rightarrow I_b) \vee (C_{bv} \rightarrow \neg I_b) \vee (C_{bv} \& I_v \rightarrow \neg I_b)), I_v, C_{bv} \& I_v \rightarrow I_b \mid\!- (C_{bv} \& I_v \langle I_b) \& \neg(C_{bv} \& I_v \langle \neg I_b)$

(61') $\neg(C_{bv} \vee I_b \vee (C_{bv} \rightarrow I_b) \vee (C_{bv} \rightarrow \neg I_b) \vee (C_{bv} \& F_b \rightarrow I_b)), F_b, C_{bv} \& F_b \rightarrow \neg I_b \mid\!- (C_{bv} \& F_b \langle \neg I_b) \& \neg(C_{bv} \& F_b \langle I_b)$

Strictly, it does not follow that Bizet would be Italian if he and Verdi were compatriots *or* that he would not be Italian if the composers were of identical nationality. (It even follows that both counterfactuals are false.) But, if we rely on the obvious fact that Verdi is Italian, we may use (61) to infer enthymematically that Bizet would be Italian if he and Verdi were compatriots. Using (61') instead and assuming as obvious that Bizet is French, we may enthymematically infer that Bizet would not be Italian if he and Verdi were compatriots.

I hasten to add that this analysis is not without problems.

First, we note that we may strengthen the premises of (61') by substituting '$\neg(C_{bv} \rightarrow I_b \vee \neg I_b)$' for '$\neg((C_{bv} \rightarrow I_b) \vee (C_{bv} \rightarrow \neg I_b))$' without rendering them jointly unsatisfiable. This strengthening seems legitimate enough, since it is hard to see how the (false) assumption of Bizet and Verdi being compatriots could possibly lead *by law* to the (logically true but independent) consequence of Bizet's being Italian or not. However, the

[59] As it turns out, the requirement of truth is not necessary here.

same strengthening leads to inconsistency in (61)[60]. This means that, if we used the obvious 'I_v' along with the strengthened premises of (61) to conclude that Bizet would be Italian if he and Verdi were compatriots *and* Verdi was Italian (as he surely is) we would – given the facts – have to conclude everything. This is unintuitive.

Second, without any strengthening of the premises in (61) or (61') we get the additional conclusion '$I_b \rightarrow C_{bv}$'. This is too strong[61]. What we want is that Verdi and Bizet *would be* compatriots if Bizet *was* Italian, not that Bizet being Italian *implies* that he and Verdi are compatriots.

Third, there are still problems with Bernd from Chemnitz. Suppose that we are concerned not with the counterfactual 'B $\langle$ E' but with the counterfactual 'N $\langle$ E' meaning that Bernd would have been born in East Berlin if he was born in *North* Berlin. (We think of North Berlin in 1962 as including districts both in the eastern and in the western sector (e.g., Wedding and Pankow) as well as excluding quarters both in the eastern and in the western sector (e.g., Dahlem and Treptow)). According to my intuition, we should *not* be able to conclude from the facts about Bernd that he would have been born in West Berlin if he was born in North Berlin *and* was (as he actually is) an Ossi. But given the present account, we have to, because (62) turns out to be unsatisfiable. [62]

(62) $\neg((N\rightarrow E)\vee(N\rightarrow W)\vee(E\rightarrow N)\vee(W\rightarrow N)\vee N\vee W\vee E)\&(N\&O\rightarrow E)\&$
$\neg(N\&O\rightarrow\neg E)$[63]

---

[60] This can be confirmed by calculation with respect to [-3,4]. I conjecture, however, it is true with respect to each entrenchment structure as defined in Section 4.

[61] At least, if it turns out (as I conjecture) that it is a conclusion *simpliciter*.

[62] Again, I conjecture that (62) is unsatisfiable *simpliciter* (not satisfiable in any entrenchment structure).

[63] I have to confess that I am insecure, how to handle these problems. We may have to alter the logic developed in Section 4. There are some principles which might be responsible for our present difficulties and seem dubious anyway (E.g., we have A, B, (A&B $\rightarrow$C), $\neg$(C$\rightarrow$A), $\neg$(C$\rightarrow$B) |– (B$\rightarrow$C)$\vee$(A$\rightarrow$C), which looks strange, even given the reminder that '|–' is used in the official sense (cf. footnote 6). Is |– (A&B&(A&B $\rightarrow$C))&$\neg$((C$\rightarrow$A)$\vee$(C$\rightarrow$B)) $\supset$ (B$\rightarrow$C)$\vee$(A$\rightarrow$C) derivable in R? I guess: no). It is also possible that we have to revise our account of indexical enthymemes. Perhaps they should not be based on logical consequence (in the official sense) but on law-like

6.2 The wet match

Let us go back to where we started:

(1) *If the match had been scratched, it would have lighted.*

According to Goodman's analysis, (1) being true requires there to be a support proposition *S* which, in conjunction with the antecedent 'S_m' (*The match has been scratched*), leads by law to the consequent 'L_m' (*The match has lighted*). Most plausibly, *S* will imply the condition 'D_m' (*The match has been dry*) along with a background condition ⌜B⌝. Taking ⌜$B\&D_m$⌝ as support proposition, we represent the requirement as

(63) $B\&D_m\&S_m \rightarrow L_m$.

Goodman argues (as I will explain momentarily) that it is hard to see how we can, in following his account, claim (1) but deny, as we certainly want to,

(1') *If the match had been scratched, it would not have been dry.*

This is the second tough problem, which, as we noted above, led Goodman to give up his project of specifying truth-conditions for counterfactuals.

Assume that ⌜B⌝ is chosen such that ⌜$B\&D_m$⌝ is true and related to 'S_m' and 'L_m' – both presumed to be false – in such a way that whatever criteria assumed to be relevant to license '$S_m \langle L_m$' are satisfied. Goodman argues that the same must then be said of ⌜$B\&\neg L_m$⌝ in relation to 'S_m' and '$\neg D_m$'. "By means of a legitimate general law," Goodman writes (*FFF*, 14), "we can infer"

(64) $B\&\neg L_m\&S_m \rightarrow \neg D_m$.

arriving at identical truth conditions for (1) (i.e., '$S_m \langle L_m$') and (1') (i.e., '$S_m \langle \neg D_m$').

Goodman urges that here, there must be a difference between 'D_m' and '$\neg L_m$'. In a sense, to him, it seems that 'D_m' is more legitimate or better qualified as a component of the support proposition than '$\neg L_m$'. The

---

consequence. Perhaps, both revisions are required. In a sense, I have to admit that Goodman's first tough problem remains a puzzle.

difference, Goodman thinks, is that $\ulcorner B\&D_m \urcorner$ is *cotenable* with 'S_m', whereas $\ulcorner B\&\neg L_m \urcorner$ is not. That $\ulcorner C \urcorner$ is *cotenable* with $\ulcorner A \urcorner$ means, according to him (*FFF*, 15), that it is not the case that $\ulcorner C \urcorner$ would be false if $\ulcorner A \urcorner$ were true (i.e., $\ulcorner \neg(A \langle \neg C) \urcorner$) (cf. footnote 43).

Hence, when applied to the present case, Goodman assumes $\ulcorner \neg(S_m \langle \neg B \vee \neg D_m) \urcorner$ but $\ulcorner S_m \langle \neg B \vee L_m \urcorner$. Being unable, however, to explicate the notion of cotenability without recourse to the notion of counterfactual conditionals itself, Goodman considers his account to be trapped by circularity and throws in the towel.

Given the perspective of *our* account, Goodman's scepticism is not unavoidable at this point. To recapitulate his argument in the present framework, we should show that (64) is a logical consequence of (63). Note that we actually *do* have

(65) $B\&D_m\&S_m \rightarrow L_m, B\&\neg L_m\&S_m \mid\!- \neg D_m,$

since the first premise and '$\neg L_m$' (from the second premise) lead by implicational *modus tollens* to $\ulcorner \neg(B\&S_m) \vee \neg D_m \urcorner$. Hence, the remaining two conjuncts of the second premise and *disjunctive syllogism* (37)(a) render the conclusion. However, this is not sufficient. What is required instead to arrive at (64) as a logical consequence of (63) is implicational *antilogism*, i.e., (37')(c). However, its instance (65') is not available for us:

*(65') $B\&D_m\&S_m \rightarrow L_m \mid\!- B\&\neg L_m\&S_m \rightarrow \neg D_m.$

In the logic presented in Section 4, antilogism is simply not a "legitimate general law."[64]

Perhaps, however, the fatal inference from (63) to (64) is actually justified, given special assumptions concerning the example at hand and constraints to be imposed on the background condition $\ulcorner B \urcorner$ for the required counterfactual link between 'S_m' and 'L_m'. Let us see.

To begin with, we assume that the values of $\ulcorner B \urcorner$, 'D_m', '$\neg S_m$' and '$\neg L_m$' are positive, since the support condition should be true and the example is a "counterfactual as such" (with false antecedent and false consequent).

[64] To refute (65'), take, e.g., *4*, *0*, **-1** and **3** as values for $\ulcorner B \urcorner$, 'D_m', 'S_m' and 'L_m' respectively.

Second, we assume that 'S_m' is compatible with '$\neg L_m$' (i.e., $\neg(S_m \rightarrow L_m)$). This assumption motivated recourse to a support proposition in Goodman's enthymematic account from the start. "Otherwise," he writes, "nothing would be required along with ['S_m'] to reach the opposite [of '$\neg L_m$'] as the consequent of the true counterfactual statement [(1)]" (*FFF*, 14).

Third, as has been mentioned above, Goodman requires as a general condition that the support proposition (here: $\ulcorner B\&D_m \urcorner$) "by itself must not decide between" (*FFF*, 13) the consequent (here: 'L_m') and its negation (here: '$\neg L_m$'). This obviously constrains the choice of the background assumption: $\ulcorner B \urcorner$ has to be such that $\ulcorner B\&D_m \urcorner$ is compatible both with 'L_m' and its negation. As for the negation, this goes without saying, since we know that $\ulcorner \neg(A \rightarrow C) \urcorner$ is true if $\ulcorner A \urcorner$ is true and $\ulcorner C \urcorner$ is false.[65] However, let us adopt (if only for the sake of the argument[66]) the requirement $\ulcorner \neg(B\&D_m \rightarrow \neg L_m) \urcorner$.

These three points give us additional premises for (65'):

(66) $B, D_m, \neg(S_m \vee L_m \vee (S_m \rightarrow L_m) \vee (B\&D_m \rightarrow \neg L_m)), B\&D_m\&S_m \rightarrow L_m$

A fourth point can be made due to Goodman's acceptance of Parry's criticism (cf. Section 3) and adding to the positive part of his explication the general requirement that the negation of antecedent and support should be compatible, here: $\ulcorner \neg(\neg S_m \rightarrow B\&D_m) \urcorner$. By the same line of argument one will require that the consequent and negation of support must be

---

[65] $A, \neg C \mid\!\sim \neg(A \rightarrow C)$. As reported in footnote 16, Goodman retained $\ulcorner \neg(S \rightarrow C) \urcorner$ as a general requirement for support $\ulcorner S \urcorner$ even in the case of a false consequent $\ulcorner C \urcorner$, because he wanted to take measures against the principle $*\neg C, \neg(A \rightarrow \neg C), A \langle C \mid\!\sim A \langle \neg C$. We, however, do not need to do so. First, we do not share Goodman's motivation to avoid $\ulcorner (A \langle C) \,\&\, (A \langle \neg C) \urcorner$ for all antecedents $\ulcorner A \urcorner$ under consideration and, second, the principle is not valid in our account anyhow. (To refute it, choose, e.g., **-1** and **3** as values for 'A' and 'C', respectively.)

[66] As noted in footnote 16, Goodman could have dispensed with the requirement $\ulcorner \neg(S \rightarrow \neg C) \urcorner$ for the support proposition $\ulcorner S \urcorner$, since it is a logical consequence of the condition $\ulcorner S\&A \rightarrow C) \urcorner$, together with the charge that its antecedent be self-compatible. We have not adopted the requirement as a condition in our definition (41), but in view of our *not* insisting on the self-compatibility of support *plus* antecedent, it is a substantial restriction in cases like the present one.

compatible (cf. footnote 19), here: $\ulcorner\neg(L_m\rightarrow B\&D_m)\urcorner$. In the context of the present account, we remain unimpressed by Parry's argument, since it depends on the acceptance of (14), the paradox of law-like implication, which we emphatically reject. However (if only for the sake of argument), we take the two requirements into account. This gives us an extended list of conditions:

(67) $\neg(\neg S_m\rightarrow B\&D_m)$, $\neg(L_m\rightarrow B\&D_m)$, (66)

The first thing to note is that the assumptions in (67) (hence in (66)) are jointly satisfiable.[67]

Next, notice that in gathering special assumptions about (1) and constraints concerning the background condition $\ulcorner B\urcorner$ we have been ignoring the negative part of Goodman's definition, which he insisted on in order to force connexivity (i.e., Boethius' thesis, here repeated once more):

*(23) $A \langle C \mid\!\!- \neg(A \langle \neg C)$.

As has been stressed in Section 2, we do not (*contra* Goodman) exclude self-incompatible antecedents (and along with them counteridenticals and counterlegals) from consideration. Hence, we are quite content in refuting (23) in our account[68]. Of course, however, in a case like (1), we do not want to have '$S_m \langle \neg L_m$' together with '$S_m \langle L_m$'. Hence, perhaps, something akin to the negative part of Goodman's definition is constraining the support proposition for (1).

But this is not the case. For if it was, it would be possible to consistently add both '$S_m \langle L_m$' and $\ulcorner\neg(B\&D_m\&S_m\rightarrow\neg L_m)\urcorner$ to (67). Instead, we have:

(68) $S_m \langle L_m$, (67) $\mid\!\!- B\&D_m\&S_m\rightarrow\neg L_m$[69].

Whenever '$S_m \langle L_m$' is true, there is a true proposition (*viz.* $\ulcorner B\&D_m\urcorner$) satisfying all constraints imposed upon the background condition $\ulcorner B\urcorner$ which – together with the antecedent of (1) (i.e., 'S_m') – implies the negation of

[67] Use again the valuation from footnote 64 (*4*, **-1**, **3** and *0* for $\ulcorner B\urcorner$, 'S_m', 'L_m' and 'D_m', respectively) to confirm this.

[68] Cf. footnote 40 above.

[69] We even have $S_m \langle L_m$, (66) $\mid\!\!- B\&D_m\&S_m\rightarrow \neg L_m$.

its consequent (i.e., '$\neg L_m$'). The non-existence of such a proposition is just what the negative part in Goodman's explication postulates.[70]

Finally, notice that we do not impose or even consider as an additional constraint on $\ulcorner B \urcorner$ the requirement that $\ulcorner B\&D_m\&S_m \urcorner$ be self-compatible. After all, the main point of the present paper is that the requirement of self-compatibility for antecedent *plus* support may be dropped as a necessary condition in the definition of $\ulcorner A \langle C \urcorner$ if the paradox of law-like implication is rejected.

Actually, in the present case we *must* avoid the additional constraint on $\ulcorner B \urcorner$. Adding it, would render (67) unsatisfiable:

(69) (67) $\vdash B\&D_m\&S_m \rightarrow \neg(B\&D_m\&S_m)$

It might be objected here that *self-compatibility* of $\ulcorner B\&Dm\&Sm \urcorner$ (in the sense of not implying its own negation) is too strong a constraint on $\ulcorner B \urcorner$ anyway. After all, what Goodman *intended* in imposing self-compatibility upon antecedent *plus* support was to rule out their joint *absurdity* (in the sense of implying every proposition). It was only because his adherence to the paradox that he could not keep the notions apart and took the stronger one (self-compatibility) in exchange for the intended weaker one (non-absurdity).

However, it turns out that this objection is not efficient. Even adding the weaker constraint that $\ulcorner B\&D_m\&S_m \urcorner$ is non-absurd renders (67) unsatisfiable.

Let us define a constant '$\boldsymbol{F}$' for $L_{\rightarrow\forall}$ as '$\forall p.p$'[71]. We may then represent the non-absurdity of $\ulcorner B\&D_m\&S_m \urcorner$ as $\ulcorner \neg(B\&D_m\&S_m \rightarrow \boldsymbol{F}) \urcorner$. We obtain:

(70) (67) $\vdash B\&D_m\&S_m \rightarrow \boldsymbol{F}$

Interestingly however, if we add the requirement of $\ulcorner B\&D_m\&S_m \urcorner$'s non-absurdity, not to (67), but to (66), we do in fact retain satisfiablity[72]. But if we conceive of the background condition $\ulcorner B \urcorner$ in this way, we lose what the

[70] Note that the validity of (68) is compatible with the invalidity of $* S_m \langle L_m$, (67) $\vdash S_m \langle \neg L_m$. We actually do have $S_m \langle L_m$, (67) $\vdash \neg(S_m \langle \neg L_m)$.

[71] Among several possibilities, we might equivalently define '$\boldsymbol{F}$' as '$t\&\neg t$', where 't' is defined as '$\forall p.p \rightarrow p$' (cf. Section 5).

[72] Witnessed, e.g., by the valuation **-1**, **1**, *3*, *4* for 'S_m', 'L_m', 'D_m' and 'B', respectively.

support proposition, the conjunction of ⌜B⌝ and 'D_m', was actually supposed to provide: a counterfactual link between S_m and L_m:

(72) $\neg(B\&D_m\&S_m\rightarrow \boldsymbol{F})$, (66) |– $\neg(S_m \langle L_m)$.

Coming back to the question of whether requisite assumptions concerning (1) would, when added as premises to (65'), restore the fatal validity of the inference from (63) to (64), our result is negative. We still have[73]

*(73) (67) |– $B\&\neg L_m\&S_m \rightarrow \neg D_m$

I think this shows that Goodman's devastating argument against his own project is not sound (at least as far as our reconstruction of the project is concerned). We cannot conclude that (1) and (1') have identical truth conditions: [74]

*(74) (67) |– $(S_m \langle L_m) \equiv (S_m \langle \neg D_m)$.

Actually, their truth conditions are demonstrably different. Even for the reduced list of conditions (66), we have

(75) (66) |– $(S_m \langle L_m) \supset \neg(S_m \langle \neg D_m)$.

(66) (and *a fortiori* (67)) is consistent with (1) and with (1') but not with both.

As for Goodman's intuition that the difference between (1) and (1') consists in the cotenability of 'S_m' with ⌜$B\&D_m$⌝ (but not with ⌜$B\&\neg L_m$⌝), we obtain as a confirmation in our account:

(76) $S_m \langle L_m$, (66) |– $\neg(S_m \langle \neg(B\&D_m)) \,\&\, S_m \langle (\neg B \vee L_m)$.

But of course, we have defined '⟨' in a non-circular way.

I conclude my discussion of Goodman's second tough problem with the remark that the analysis of (1) proposed here has to be taken with a grain of salt. We not only have *(74) and (75) but also (77) and (78):

(77) $S_m \langle L_m$, (66) |– $L_m \rightarrow S_m$

(78) $S_m \langle L_m$, (66) |– $(B\rightarrow D_m)\vee(D_m\rightarrow B)$

[73] Use, e.g., **-1**, **3**, *0*, *4* for 'S_m', 'L_m', 'D_m' and 'B', respectively, to refute (73).

[74] Use, e.g., **3**, **2**, *-1*, *1* for 'S_m', 'L_m', 'D_m' and 'B', respectively, to refute (74).

Perhaps, one could accept that – given the conditions specified in (66) *plus* (1) – the match would have been scratched if it lighted. That, however, even under these circumstances, its lighting *implies* its having been scratched seems too strong. Likewise it seems hard to digest that the conditions in (66) *plus* (1) constrain background ⌜B⌝ in such a way, that it either implies or is implied by the match being dry. Intuitively, ⌜B⌝ might be specified quite independently of 'D_m'. I think Goodman had an independent specification in mind when he wrote, "As our total [⌜$B\&S_m$⌝] we may have: 'Match *m* is scratched. It does not light. It is well made. Oxygen enough is present … etc'" (*FFF*, 14)[75]. As far as our account goes, such an independent specification does not exist.

7. Conclusion

> "*… and shun the cure they most desire.*"
> Henry Purcell, *loc.cit.*

In this paper I have proposed the reconsideration of Goodman's basic ideas about counterfactuals within a paradox-free framework. The basic features of these ideas, I take it, are:

- The enthymematic character of counterfactuals,
- Their dependence on law-like implication,
- Conditionality (*modus ponens* and *modus tollens*) and
- Contraposition.

In addition, the account presented respects features not discussed explicitly by Goodman but in accordance, I presume, with the general direction of his endeavour:

- Non-monotonicity and
- Non-transitivity.

A paradox-free conception, like the one used here, essentially consists in the abandonment of the paradox of law-like implication (i.e., (14)).

I have been showing that getting rid of Goodman's pertinent presumption at this point allows for technical simplification and conceptual broadening. The revised account is technically more smooth insofar as

[75] I add these scruples to the list of problems mentioned in footnote 63.

certain general and intuitively acceptable principles of counterfactual inference (like *identity* (16), *conjunctive reduction* (17), *disjunctive extension* (18), etc.) can be retained without restriction. More importantly, perhaps, the revision gains conceptual generality insofar as Goodman's exclusion of counterlegals and counteridenticals (which seem hardly independently motivated) is no longer required. Boethius' thesis can no longer be vindicated for couterfactual conditionals.

The price we have to pay for getting rid of the paradoxes is quite moderate. We do not have to buy non-official notions of logical consequence, advocated by champions of relevance logic (Anderson & Belnap 1975, 1992, Routley et.al. 1982, Dunn 1986, cf. also Read 1988, Heydrich 1993a and Mares 2004). The only thing one has to swallow is that the relation between (law-like) implication ($\rightarrow$) and logical consequence is not quite the same as the relation between material implication ($\supset$) and logical consequence. This, however, can hardly come as a surprise. After all, law-like implication is not material implication (or any other truth-function) and no one expects it to be.

The main point stressed repeatedly above is that logical consequence should not be construed as a sub-relation of law-like consequence. Mainly, the rule of *conditionalisation*, valid for the horse-shoe in (36), is not unrestrictedly operative for the arrow in (36').

(36) if $A_1, \ldots, A_k, A \mid\!- C$, then $A_1, \ldots, A_k \mid\!- A \supset C$
*(36') if $A_1, \ldots, A_k, A \mid\!- C$, then $A_1, \ldots, A_k \mid\!- A \rightarrow C$.[76]

We especially do not adopt implicational conditionalisation of *disjunctive syllogism* (37')(a), *antilogism* (37')(c), or truth-functional *modus ponens* (37')(b).

To implement a specific paradox-free framework, I have proposed, in terms of polarity (truth values) and degree of entrenchment, a *logic of entrenchment*[77]. About the precise syntactic (proof-theoretical) characterisation of this logic and its meta-logical properties I am uncertain. The same is true of its *final* explicational value for the treatment of

[76] The converse of (36') holds unrestrictedly.

[77] To philologically-minded readers who feel that I do not use Goodman's term "entrenchment" properly or have even hijacked it, I apologise.

counterfactuals[78]. Frankly, the logic mainly has an instrumental character in the present paper. I have adopted it to back up claims and dispose of some reliable ground in order to prove or disprove principles one would like to accept or reject. In doing so, I have restricted myself to a decidable fragment of that logic. Perhaps a modified (or quite different) paradox-free framework is necessary.

Be that as it may, whatever the ultimate (or an improved) logical framework for the enthymematic analysis of counterfactuals will look like, readers of Goodman may feel that any project of this brand is severely at odds with Goodman's deep philosophical convictions. Goodman has described his position repeatedly as "radical relativism under rigorous restraints" (*WW*, x). This *dictum* nicely combines open-mindedness and austerity. However, it is not quite clear what "rigorous restraints" exactly means, what is comprised and what is excluded. Reading Goodman, one gets the impression that inscriptions and other individuals[79] are innocent and accepted, whereas intensions (thoughts, concepts, properties and propositions) are disreputable and have to be expelled – as long as, of course, they cannot be understood in terms of talk about individuals (at least in principle). The logical bounds of this talk are the tools of classical symbolic logic, restricted to first order quantificational theory. "Rigorous restraints," it seems, means austere extensionalism (one without classes).

It is not easy to see how this position could be reconciled with a *para*-consistent logical framework with propositional quantification and an intensional core-connective. After all, it is Goodman's scepticism concerning intensionality (cf. Heydrich 1993b) that led him to a meta-linguistic treatment of law-like implication (cf. Section 2).

However, it is not too hard to uncover below the guise of Goodman's use of the meta-linguistic relation of ⌜X *leading by law to* Y⌝, a logical conception strongly influenced by the notion of *strict implication* in modal logic. Supposedly, this is the logical source from where Goodman inherited his presumption of the paradox (14).[80]

---

[78] Cf. footnote 63 and 75.

[79] "Individuals" in the technically elaborate sense of mereology (cf. Goodman & Leonard, 1940 and *SA*, 33ff).

[80] Cf. footnote 1 in *FFF*, 3, where Goodman testifies "indebtedness" to the work of the founder of modal logic, C.I. Lewis, as "too obvious to call for detailed mention".

If this speculation is correct, the upshot of the present paper is to advocate a reorientation in the *sources of inspiration* for Goodman's analysis, not the abandonment of his program of subduing – at least in the long run – everything one wants to maintain under the rigorous restraints of austere extensionalism. This program, I do not discuss. What I want to recommend is: do not take Clarence I. Lewis' *strict* implication as your starting point in understanding law-likeness involved in the enthymematic analysis of counterfactuals, but instead take Wilhelm Ackermann's *strong* implication (cf. Lewis & Langford (1932), Ackermann (1958)).

In short, I propose to reconstruct Goodman as a crypto-relevantist.

***

In an interview in August 1980, Goodman reported about "the tremendous event it was when symbolic logic burst upon the scene. It so quickly developed that […] we may forget how […] gratefully it was welcomed by many of us who were looking for a new aid to clarity in philosophy." (*MM*, 190). Undoubtedly, Goodman remained loyal to classical symbolic logic in the Frege-Russell-Tarski tradition (and shared with Quine a preference for ontologically motivated first-order). Actually, I know of no place in his work where he discusses (or only flirts with) heterodox (non-classical) conceptions of logic as a potentially fruitful "aid to clarity in philosophy". Irrespective of how rigorously alternatives to classical logic might have been restrained and worked out formally, this seems to constitute a confinement of Goodman's open-minded relativism.

However, I think even at this point (of faithfulness to classical first order logic), Goodman is not dogmatic. This can be appreciated from passages where he writes about how we can justify deductive logic:

> The validity of a deduction depends […] upon conformity to valid rules. […] Principles of deductive inference [in turn] are justified by their conformity with accepted deductive practice. […] This looks flagrantly circular. […] But this circle is a virtuous one. The point is that rules and particular inferences alike are justified by being brought into agreement with each other. […] The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either (*FFF*, 63f).

Even here, in questions concerning the most fundamental conceptual tools including deductive rules and inference of which we try to dispose,

there is no self-evidence and there are no pre-established principles, dogmatically fixed once and for all. Nothing is given in advance. Even logic is – far from being either arbitrary or unalterable – a matter of "rightness of rendering." (*WW*, 109).

References

Ackermann, W. (1958): "Über die Beziehung zwischen strikter und strenger Implikation", in: *Dialectica* 12 (1958), 213-22.

Adams, E.W. (1970): "Subjunctive and Indicative Conditionals", in: *Foundations of Language* 6 (1970), 89-94.

Anderson, A.R. and N.D. Belnap (1975): *Entailment. The Logic of Relevance and Necessity.* Vol. I, Princeton, London: Princeton University Press, 1975.

Anderson, A.R. and N.D. Belnap (1992): *Entailment. The Logic of Relevance and Necessity.* Vol. II, Princeton, Oxford: Princeton University Press, 1992.

Bennett, J. (2003): *A Philosophical Guide to Conditionals.* Oxford: Clarendon Press, 2003.

Burris, S. and H.P. Sankappanavar (1981): *A Course in Universal Algebra.* New York, Heidelberg, Berlin: Springer (The Millennium Edition 2001)

Chisholm, R.M. (1946): "The Contrary-to-Fact Conditional", in: *Mind* 55 (1946).

Chisholm, R.M. (1955): "Law Statements and Counterfactual Inference", in: *Analysis* 15, 97-105.

Dunn, M. (1986): "Relevance Logic and Entailment", in: Gabbay, D. and F. Guenthner (eds): *Handbook of Philosophical Logic*, Vol. III, Dordrecht: Reidel, 1986, 117-224.

Goodman, N. (1957): "Parry on Counterfactuals", *Journal of Philosophy* 54 (1957), 442-445.

Goodman, N. and H.S. Leonard (1940): "The Calculus of Individuals and its Uses", in: *The Journal of Symbolic Logic* 5 (1940), 45-55.

Heydrich, W. (1993a): *Relevanzlogik und Situationssemantik*. Berlin, New York: de Gruyter, 1993.

Heydrich, W. (1993b): “A Reconception of Meaning”, in: *Synthese* 95 (1993), 77-94.

Heydrich, W. (2007): "Defending Goodman against himself", in: E. Franzini and M. La Martina (eds): *Nelson Goodman, la filosofia e i linguaggi,* Macerata: Quodlibet, 2007, 75-107.

Karttunen, L. and S. Peters (1979): “Conventional Implicatures“, in: Ch.-K. Oh and D.A. Dinneen (eds): *Presupposition*. Syntax and Semantics, Vol.11, New York, San Francisco, London: Academic Press, 1979, 1-56.

Lewis, C.I. and C.H. Langford (1932): *Symbolic Logic*. New York, London: Dover Publications, 1932.

Lewis, D. (1973): *Counterfactuals*. Cambridge (Mass.), London: Harvard University Press, 1973.

Mares, E.D. (2004): *Relevant Logic. A Philosophical Interpretation*, Cambridge : Cambridge University Press, 2004.

Parry, W.T. (1957): “Reexamination of the Problem of Counterfactual Conditionals”, in: *Journal of Philosophy* 54 (1957), 85-94.

Read, S. (1988): *Relevant Logic*, Oxford: Bleckwell, 1988.

Routley, R, V. Plumwood, R.K., Meyer and R.T. Brady (1982): *Relevant Logics and their Rivals*. Atascadero: Ridgeview Publishing Company, 1982.

Stalnaker, R.C. (1968): “A Theory of Conditionals”, in: N. Rescher (ed.): *Studies in Logical Theory*, Oxford: Basil Blackwell, 1968, 98-112.

van Fraassen, B.C. (1966): “Singular Terms, Truth-value Gaps, and Free Logic”, in: *Journal of Philosophy* 63 (1966), 481-495.

von Kutschera, F. (1975): “Nelson Goodman: Das neue Rätsel der Induktion”, in: Speck, J. (1975): *Grundprobleme der großen Philosophen. Philosophie der Gegenwart III*, Göttingen: Vandenhoeck und Ruprecht, 1975, 51-86.

Ansgar Seide

Contextualist References in Nelson Goodman's Solution to the "New Riddle of Induction"[1]

Abstract
This paper pursues two aims. The first aim is to appraise Goodman's solution to the "new riddle of induction" by relating it to different theories of epistemic justification. The second aim is to look at the conclusions that can be drawn from Michael Williams' claim that Goodman's account is best interpreted as a kind of contextualism. Especially, this claim can be seen as a hint to Williams' understanding of his own contextualist theory. It sheds light on the ranking of those elements in his theory that play an important role in Goodman's account.

Introduction

In his book *Fact, Fiction, and Forecast*, Nelson Goodman addresses David Hume's well known Problem of Induction. Unlike many other contributors to the debate, Goodman takes it that the problem has been solved (or dissolved). By introducing his famous predicate "grue", he poses the "new riddle of induction" and proposes a surprising solution to it.

This solution will be surveyed in the following passages. For this the connection between Hume's Proplem of Induction and Goodman's "new riddle of induction" plays an important role. Following an idea of Michael Williams', this connection will be investigated in the light of an analogy between approaches to the justification of inductive inferences and theories of epistemic justification. A special focus will be laid on Williams' thesis that Goodman's solution to the "new riddle of induction" can be interpreted as an application of a (namely Williams' own) contextualist theory of justification.

This paper pursues two aims. The first aim is to appraise Goodman's solution by relating it to different theories of epistemic justification. The

[1] For helpful comments on the translation of this paper I am indebted to Rudolf Owen Müllan.

second aim is to look at the conclusions that can be drawn from Williams' claim that Goodman's account is best interpreted as a kind of contextualism. Especially, this claim can be seen as a hint to Williams' understanding of his own contextualist theory. It sheds light on the ranking of those elements in his theory that play an important role in Goodman's account.

1. The old problem of Induction

The *Problem of Induction* is often called *Hume's Problem* because David Hume raises and investigates it in his books *A Treatise of Human Nature* and *An Enquiry Concerning Human Understanding*. It is a problem concerning the justification of inductive inferences.

Inductive inferences are generalizing inferences that take us from a finite set of observations to a general statement.[2] They play an important role in everyday life as well as in science: When I sit down on a chair, I do not expect the chair to dissolve into thin air. This is because I have sat down a lot of times in my life and nothing as strange as this ever happened to me. Scientists infer general laws from individual observations: A general law of falling bodies can only be inferred from a finite set of observations of bodies that are released near the surface of the Earth (Williams 2001, 201).

Hume raises the question of the justificatory status of such inferences. At first he notices that we cannot know *a priori* that future or unobserved cases resemble the observed cases.[3] It is in principle not contradictory to conceive a change in the course of Nature to the effect that future cases

[2] At least, this is the way induction is understood today. Hume seems to think of inductive inferences as inferences from past experience to future experience. As Michael Williams puts it, Hume is more interested in prediction than in generalization.

The Problem of Induction as raised by Hume can be transferred to today's understanding of induction easily, as is shown in n. 3. Therefore I will disregard for the differences of the two understandings of induction in the following (cf. Williams 2001, 204ff).

[3] This characterization of the problem, stated in temporal terms by Hume, has to be generalized in the face of today's characterization of inductive inferences: The fundamental problem is not that the course of Nature might change to the effect that especially *future* cases will not resemble the observed cases. Rather, the question is how we know that the observed cases are a representative part of the world *at all* (and not only with regard to different *periods of time*). "[T]he problem of prediction from past to future cases is but a narrower version of the problem of projecting from any set of cases to others" (*FFF*, 83; cf. Williams 2001, 205).

will not turn out to be the way we expect them to be on the basis of an inductive inference: The thought that a heavy stone released near the surface of the Earth will not fall down does not amount to a logical contradiction.

So we have to look for another kind of justification for inductive inferences. A natural suggestion is that we rely on a principle to the effect that the future resembles the past, or rather future unobserved cases resemble the past observed ones. This principle would allow us to infer from a finite set of observations to a general statement.[4] But how can we justify this principle of uniformity in Nature itself? As we have already seen, the thought that the future does *not* resemble the past does not amount to a contradiction. So there seems to be no possibility of justifying the principle *a priori*. But an *a posteriori* justification seems to be ruled out as well: The principle is itself a general statement that could be justified empirically only by an inductive inference. But such an inductive inference, as we have seen above, would have to rely on the principle of uniformity in Nature. So there is no non-circular way of justifying the Uniformity Principle empirically (Hume 1975, 37f).

Hume's negative result is that there is no rational defence of the validity of inductive inferences.[5] Still it cannot be denied that we rely on inductive inferences frequently and that we are convinced of their validity. But what brings us to this unshakable conviction? Hume's answer is that it is from sheer force of habit that we assume future cases to resemble past ones (Hume 1975, 43). So the reason why we rely on inductive inferences is not that we have a rational insight into their validity but that it is a principle of human nature that we expect future cases to resemble past ones.

At first sight this answer does not seem to be a solution to the Problem of Induction. It rather sounds like a sceptical conclusion: While Hume sheds light on the *source* of inductive reasoning, his answer does not seem to address the question of the *legitimacy* of inductive inferences.

---

[4] Strictly speaking this principle has to be broadened out to cover an inference to a general statement. See n. 3.

[5] "[I]t is not reasoning which engages us to suppose the past resembling the future [...]. This is the proposition which I intended to enforce in the present section" (Hume 1975, 39).

2. Goodman's (dis)solution of Hume's Problem

In his book *Fact, Fiction, and Forecast*, Nelson Goodman tells us that he takes the Problem of Induction to have been solved or rather dissolved. On Goodman's view, the basic idea for the answer to the problem can already be found in the writings of Hume.[6]

Goodman draws attention to the question of what we can expect from a justification of inductive reasoning. In approaching this issue, he first takes a look at how we justify *de*ductive reasoning (*FFF*, 63).

So how do we justify a deductive inference? According to Goodman, we do this by showing that it conforms to the general rules of deductive inference. But how are the rules of deductive inference themselves to be justified? Goodman explicitly rejects the idea that the rules of deductive inference follow from self-evident axioms or are grounded in the nature of the human mind.[7] His, at first sight, somewhat surprising answer is that the principles of deductive inference are justified by their conformity with accepted deductive practice.

As Goodman readily concedes, this line of argument seems to be hopelessly circular: Deductive inferences are justified by their conformity to valid rules of deductive inference, while these rules of inference, in turn, are justified by their conformity to valid deductive inferences. According to Goodman, however, this circle is not vicious: The rules of inference on the one hand and the inferences on the other hand are justified by the agreement that is achieved by making mutual adjustments between them.[8]

On Goodman's view, this model of justification can be transferred to the case of inductive reasoning: Accepted inductive inferences and general

---

[6] "What is commonly thought of as the Problem of Induction has been solved, or dissolved [...]" (*FFF*, 59). "I think Hume grasped the central question and considered his answer to be passably effective. And I think his answer is reasonable and relevant, even if it is not entirely satisfactory" (*FFF*, 61).

[7] "Here again we encounter philosophers who insist that these rules follow from some self-evident axiom, and others who try to show that the rules are grounded in the very nature of the human mind. I think the answer lies much nearer the surface" (*FFF*, 63).

[8] "The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences; and in the agreement achieved lies the only justification needed for either" (*FFF*, 64).
For an illuminating illustration of this process of making mutual adjustments, see (Haack 1978, 15f).

rules of inductive inference are justified by the agreement that is achieved through making mutual adjustments between them. Inferences and rules support one another. The Problem of Induction, Goodman tells us, is not a problem of *demonstrating* that inductive inferences are valid but of definig the difference between valid and invalid inductive inferences.

Does Goodman really offer a satisfactory solution to the Problem of Induction? To answer this question, we will have to take a further look at the kind of justification Goodman proposes for deductive and inductive inference.

3. Excursus: Theories of epistemic justification

Obviously, Goodman's account of justification for inductive and deductive inference resembles coherence theories of justification. It is reasonable to assume that a comparison with theories of this kind will shed light on Goodman's account.

Coherence theories of justification can be seen in relation to the problem of scepticism. The most famous kind of scepticism is scepticism about knowledge of the external world, which is also called *Cartesian Scepticism,* because its first clear formulation can be found in the wiritings of René Descartes. Descartes points out that all our sense impressions are compatible with the hypothesis that we are only dreaming and that the external world is completely different from what we normally expect it to be on the basis of our sense impressions. To put it generally, the sceptical line of argument goes like this: A (sometimes bizarre) possibility is pointed out that is compatible with all our evidence but that is completely different from what we expect to be the actual state of our world. In the face of such a *sceptical scenario* we are obliged to hold that all our beliefs about the external world are unjustified.

The classical reaction to this problem is the attempt to argue for a positive theory of knowledge that is able to explain how knowledge is possible. Historically, *foundationalism* is the standard choice for such a constructive project.

The common idea behind all versions of foundationalist theories of justification is to ground our empirical beliefs on a foundation. The foundation is supposed to consist of so-called basic beliefs that have a certain degree of noninferential justification, i.e. justification that does not

derive from any further beliefs. In early foundationalist theories it was even demanded that these basic beliefs have only self-evident truths as their contents or be in some way beyond rational doubt. The justification of all other empirical beliefs is supposed to derive directly or indirectly from this set of epistemically privileged basic beliefs.

There are a lot of problems for foundationalist accounts especially in connection with Cartesian Scepticism. The main problem seems to be choosing the scope of the basis: On the one hand, it has to be rich enough to include basic beliefs that can serve as an inferential basis for *all* our empirical beliefs, on the other hand, all the basic beliefs have to be good candidates for beliefs that have a certain degree of intrinsic justification. One of the classical accounts is to suggest that beliefs about our immediate sense impressions are beyond reasonable doubt and so can be classified as basic beliefs. But the Cartesian dream hypotheses just shows that our beliefs about the external world are underdetermined by our beliefs about our immediate sense impressions.

Coherence theories of justification are an alternative to foundationalism and are often advocated by philosophers exactly because they hold that the problems faced by foundationalism are unsolvable. The basic idea is that a system of beliefs is holistically justified by inferential relations between the beliefs. Intuitively, coherence is a measure of how good a system of beliefs "hangs together" (BonJour 1985, 93). This "hanging together" of the beliefs is ensured by various inferential and explanatory relations.[9] The crucial point is that the involved relations are not one-way relations that correspond to an objective epistemic priority of certain beliefs over others, as is advocated by foundationalists. Rather, coherentists argue, these relations are reciprocal relations of mutual support.

Coherence theories too are faced with serious problems.[10] I will focus here on one objection against coherence theories that has a special significance for the subject of this paper. This is the so-called *alternative coherent systems objection*: According to coherence theories as described above a system of beliefs is justified solely by its internal coherence. But no matter how the concept of coherence is analysed in detail, it will always be possible to construct competitive systems of beliefs that are equally

[9] Cf. (BonJour 1985, 95ff), and in greater detail (Barthelborth 1996, 148ff, 172ff).

[10] A clear structured summary of various objections against coherence theories of justification is offered by (BonJour 1999, 127ff).

coherent. So on a coherence theory we have no more reason to think that our actual beliefs are true than to think that the beliefs of another, equally coherent system are true (BonJour 1985, 107).

4. The relation between the Problem of Induction and Cartesian Scepticism

As Michael Williams remarks, the Problem of Induction is analogous to Cartesian Scepticism in an interesting way: Both are underdetermination problems (Williams 2001, 211). As we have seen above, the Cartesian dream hypotheses shows that our attempt to justify our beliefs about the external world by showing that they are inferable from immediate knowledge about our own sense impressions is doomed to fail: In view of the fact that it is possible that we are only dreaming and that the external world is completely different from what we expect it to be on the basis of our sense impressions, we have to realize that beliefs about the external world are underdetermined by our beliefs about our immediate sense impressions. Analogously, Hume's considerations show that beliefs about future events are underdetermined by past observations: It is not contradictory to think that future cases will be completely different from what we expect them to be on the basis of an inductive inference.

The underdetermination problem raised by the Cartesian dream hypotheses is strongly connected with a foundationalist structure of justification. It arises from the attempt to infer our beliefs about the external world from our beliefs about our sense impressions. Analogously, it can be said that the attempts to justifiy inductive inference considered and rejected by Hume have a foundationalist structure. The basis can be said to consist of beliefs about past events; beliefs about future events are to be jusified by being inferred from them.

As we have seen above, the problems of foundationalism have led to the consideraton of giving up the idea of a foundationalist structure of justification and of developing a coherence theory instead. Goodman seems to have something similar in mind in connection with the Problem of Induction. Agreeing with Hume, he holds that it is impossible to justify inductive inferences in a foundationalist way. That is why he proposes justifying the rules of inductive inference and the accepted inductive

inferences by making mutual adjustments between them to the effect that they support each other. Obviously, this thought is coherentist in spirit.

It seems that, by making this move, Goodman incurs the typical problems faced by coherence theories. In particular, the alternative coherent systems objection can be raised against Goodman's account: It seems to be very easy to construct alternative systems consisting of well adjusted inference rules and inferences that are just as coherent as our accepted deductive and inductive inferences and inference rules.

To see how this objection can be dealt with within the framework of Goodman's account, we turn now to the famous "new riddle of induction".

5. The "new riddle of induction"

As indicated above, Goodman's way of dealing with the Problem of Induction is connected with a reassessment of the problem:

> The problem of induction is not a problem of demonstration but a problem of definig the difference between valid and invalid inferences (*FFF*, 65).

On Goodman's view, the debate concerning the Problem of Induction in the wake of Hume has been misguided. While most of the contributors to the debate tried to *demonstrate* that inductive inferences are valid, Goodman thinks that this is not neccesary. Instead of trying to demonstrate that inductive inferences are valid, Goodman analyses the difference between valid and invalid inferences.[11]

As we have seen in the last section, this means that Goodman rejects attempts at justifying inductive inferences in a foundationalist way. Instead, he puts forward the coherentist idea of making mutual adjustments

[11] With this in mind it can be said that Goodman *dis*solves the problem of the justification of inductive inference: "We no longer demand an explanation for guarantees that we do not have, or seek keys to knowledge that we cannot obtain" (*FFF*, 64).

On Goodman's view, Hume has already recognized that the more relevant question is the question of the distincton between valid and invalid inferences: "And we owe belated apologies to Hume. For in dealing with the question how normally accepted inductive judgments are made, he was in fact dealing with the question of inductive validity. The validity of a prediction consisted for him in its arising from habit, and thus in its exemplifying some past regularity" (*FFF*, 64f).

It has to be said, though, that this reading of Hume probably deviates from what can be said to be the standard interpretation.

between accepted inductive inferences and the rules of inductive inference. A successful model is the case of *de*ductive inference: There are familiar and highly developed laws of logic that are in agreement with our accepted deductive inferences (*FFF*, 65).

But inductive inference is a much more problematic case. As Goodman's famous "new riddle of induction" illustrates, there can be no purely syntactic rule that defines the difference between valid and invalid inductive inferences. To show us this, Goodman defines a predicate "grue":

> [T]he predicate 'grue' [...] applies to all things examined before t just in case they are green but to other things just in case they are blue (*FFF*, 74).

Let us now have a look at an example of an inductive inference that we normally take to be valid: Suppose that all emeralds examined before a certain time t are green. We assume, at least if we have tested enough cases, that this supports the hypotheses $\mathbf{H_1}$ that all emeralds are green.[12]

According to the definition of the predicate "grue", all emeralds that have been examined before t and are green are grue as well. So it seems that the observations made up to the time t also support the hypotheses $\mathbf{H_2}$ that all emeralds are grue. But the predictions that can be inferred from H_1 and H_2 contradict each other: According to H_1 all emeralds found after t will be green. According to H_2 they will be grue and therefore blue.

This example illustrates that there are hypotheses we do not assume to be confirmable: While we think that the hypothesis that all emeralds are green is confirmed by positive instances, we evidently do not think that the same holds for the hypothesis that all emeralds are grue. The "new riddle of induction" is the question what distinguishes hypotheses that are confirmable from those that are not (*FFF*, 80f).

To answer this question, Goodman tells us, we have to take into account the *entrenchment* of the predicates contained in the hypotheses. The entrenchment of a predicate results from the repeated actual projection of it. A predicate is more entrenched than another predicate if it was projected, i.e. used in inductive inferences, more often (*FFF*, 94). The predicate "green", having been projected so many times, is obviously much

[12] As indicated by Hume, the observations do not *guarantee* the hypotheses to be true. Nevertheless we normally assume that certain hypotheses are *confirmed* by positive instances. Goodman's task is to define the difference between inductive inferences of this kind and inductive inferences that we take to be obviously invalid.

more entrenched than the predicate "grue" that we only know since Goodman invented it. To put it in a nutshell, Goodman's solution of the "new riddle of induction" is that the hypothesis H_1 *overrides* the conflicting hypothesis H_2 because the predicate "green" contained in it is more deeply entrenched than the predicate "grue" that is contained in H_2 (*FFF*, 101). So on this account H_2 is not confirmed by positive instances because there is a hypothesis H_1 disagreeing with H_2 about future cases and containing a predicate that is more deeply entrenched.

It is a remarkable feature of this solution that it reverses our common sense picture of the connection between the actual projection of predicates and their projectibility (cf. Ernst 2005, 102). We in general assume that we got into the habit of projecting predicates like "green" just because we judge them to be projectible. On Goodman's view, it is just the other way round:

> [...] I submit that the judgment of projectibility has derived from the habitual projection, rather than the habitual projection from the judgment of projectibility. The reason why only the right predicates happen so luckily to have become well entrenched is just that the well entrenched predicates have thereby become the right ones (*FFF*, 98).

6. Interim conclusion

Let us take a look at the results up to now: Following Hume, Goodman holds that it is impossible to demonstrate the validity of inductive inferences. Unlike many other contributors to the debate, however, Goodman does not think that this is a problematic result. Instead of attempting to justify inductive inference in a foundationalist way, he puts forward the idea of making mutual adjustments between accepted inductive inferences and the rules of inductive inference. As indicated in section 4, this coherentist idea gives rise to the alternative coherent systems objection: If accepted inferences and rules of inductive inference were justified solely by relations of mutual support, we could construct indefinitely many systems consisting of well adjusted inferences and rules of inference that were justified to the same degree.

In the light of Goodman's solution to the "new riddle of induction", however, we can now see how this objection can be dealt with: Valid inferences are not only justified by being part of a coherent system of

inferences and inference rules. In addition to this, they have to meet the requirement of being accepted by us in practice. This kind of entrenchment makes the difference between a *constructed* and an *accepted* system of mutually adjusted inferences and inference rules.

In connection with the attempt to compare Goodman's account with theories of epistemic justification, we can now raise the question wether Goodman's solution to the "new riddle of induction" is really genuinely coherentist. As indicated in the beginning, Michael Williams holds that Goodman's solution is rather to be interpreted as an application of a contextualist theory of justification (Williams 2001, 216). To be able to investigate this thesis, we first have to take a look at Williams' contextualist theory of justification.

7. Excursus: Contextualism

In the face of the problems connected with foundationalist and coherentist attempts to respond to the challenge of scepticism, some epistemologists look for other ways of facing the sceptic. Michael Williams, for instance, puts forward an *indirect* or *diagnostic* approach to scepticism, the aim being to uncover the epistemological presuppositions of the sceptical arguments. His main thesis is that the sceptical arguments presuppose theoretical ideas that are not as intuitive or natural as they appear.[13] So the aim of his theoretical diagnosis is not to refute scepticism once and for all, but to shift the burden of argument. If it can be shown that the sceptic needs to base his arguments on highly theoretical presuppositions that are not as simple and intuitive as they appear, the sceptic has to come up with arguments supporting them (Williams 1996, xvii).

According to Williams, the decisive theoretical presupposition of the sceptical arguments is the thesis that our beliefs have an *intrinsic epistemological status*, i.e. that they "arrange themselves into broad, theoretically coherent classes according to certain *natural* relations of epistemological priority" (Williams 1996, 116).[14]

---

[13] By *intuitive* or *natural ideas* Williams means ideas that can be understood "[...] without prior indoctrination in contentious theoretical ideas [...]" (Williams 1996, xv) and that "[...] almost any reflective person can understand and be moved by" (Williams 2001, 58).

[14] To show that the sceptic is commited to this thesis is one of the main goals in (Williams 1996).

In order to shift the burden of argument to the sceptic, Williams develops a contextualist theory of justification based on the thesis that there are *no* objective relations of epistemic priority between beliefs or classes of beliefs.[15] This complete negation of objective relations of epistemic priority leads to the conclusion that the epistemological status of a proposition is not only dependent on contextual factors, but that independently of all contextual influences a proposition has no epistemic status at all (Williams 1996, 119). A context is generated by a set of legitimate presuppositions that are not currently under scrutiny. As Williams emphasizes with reference to Wittgenstein, the question "What has to be tested by what?" can only be answered against the setting of a context (Williams 1996, 118). Wittgenstein writes in *On Certainty*:

> If a blind man were to ask me "Have you got two hands?" I should not make sure by looking. If I were to have any doubt of it, then I don't know why I should trust my eyes. For why shouldn't I test my *eyes* by looking to find out wether I see my two hands? *What* is to be tested by *what*? (Who decides *what* stands fast?) (Wittgenstein 1969, 125)

All in all, these considerations result in a contextualist theory of justification according to which the standards of knowledge attribution are not objective, but are dependent on contextually variable factors. Besides other factors listed by Williams, especially the so-called *intelligibility or semantic constraints* are a source of entitlements to contextual presuppositions: Unless we assume that most of our beliefs are true, Williams points out, we cannot make sense of our beliefs and are not in a position to raise questions at all:

> At some point mistakes shade of into unintelligibility. Someone who cannot do the simplest calculations, or perform the simplest counting operations, is not making arithmetical mistakes: he does not understand numbers (Williams 2001, 160).

In foundationalist theories we often encounter this thesis as the claim that we have a priviledged access to our immediate sense impressions and that beliefs concerning our sense impressions are therefore epistemically prior to beliefs about the external world.

Coherentists, although emphasizing that there are no objective relations of epistemic priority between the beliefs of our belief system, hold that the justification of individual beliefs depends on the coherence of the whole system of beliefs which is genuinely prior. So even coherentists do not dismiss the idea of epistemic priority completely (cf. BonJour 1985, 241, n. 20).

[15] While his arguments against the sceptic are spelled out in detail in (1996), Williams develops his own account in particular in (2001).

According to Williams, this example illustrates that although individual beliefs can be legitimately challenged and then have to be defended or to be given up, we always have to assume that *most* of them are true.

So Williams' contextualist account gives rise to a picture of justification according to which we can presuppose a lot of beliefs without having gone through a prior process of justification: Making such presuppositions is not epistemically irresponsible, it is also not just a practical limitation but it is a condition on which questions concerning epistemic justification make sense at all.

A highly interesting feature of this contextualist account is that both foundationalist and coherentist ideas can be connected with it. On the one hand beliefs can be justified by being inferred from beliefs that are contextually basic. Those contextually basic beliefs are not basic in a substantially foundationalist sense, i.e. they do not have a positive epistemic status just because of their content. Rather, in every context there are beliefs that are exempted from doubt and can be used as premisses in justifying arguments for other beliefs while they are in need of justification themselves in other contexts. On the other hand, contexts, since they involve a lot of contextually basic beliefs, can be assessed by coherentist standards.[16] Roughly speaking, this can be taken as the idea that a context is a better basis for justification the more the beliefs contained by it are connected by inferential relations.

This combination of foundationalist and coherentist features takes us back to Goodman's solution to the "new riddle of induction".

8. A (Goodman's?) contextualist solution to the "new riddle of induction"

As we have seen, contexts are characterized by sets of presuppostions that are not currently under scrutiny because they enjoy a positive default status. Williams emphasizes that contextually basic beliefs concern matters of general as well as particular fact. In particular, a lot of presuppositions concerning the projectibility of predicates and the confirmability of hypotheses are implicit in our inductive practice (Williams 2001, 211).

[16] At least sometimes Williams seems to point to such a possibility: "A contextualist account of justification may be, so to say, *locally holistic*, in that a context of justification always involves a significant array of beliefs" (Williams 1996, 287).

There seem to be good reasons to count these presuppositions as contextually basic beliefs, many of which have to be (and are) presupposed to make sense of questions concerning justification at all.

According to Williams, Goodman's solution to the "new riddle of induction" suggests a contextualist reading along these lines.[17] At first sight Goodman's idea of making mutual adjustments between rules and accepted inferences seems to hint at a genuinely coherentist account. His "new riddle of induction" shows that inductive logic can never be purely formal. Finally, Goodman's solution to the "new riddle" is based on the idea that the process of making mutual adjustments between rules and accepted inferences is, as Williams puts it, "constrained by prior views about how the world can usefully be described: in effect, by general views about the way the world works" (Williams 2001, 216).

As we now can see, foundationalist as well as coherentist ideas play an important role in this account: On the one hand the assumptions concerning the projectibility of predicates that are implicit in our record of past projections are part of the set of *contextual basic beliefs* we have to presuppose to be able to make sense of questions concerning the validity of inductive inferences. On the other hand there is the demand for a *reflective equilibrium* between rules and accepted inferences.

9. Concluding remarks

As we have seen, Goodman's solution to the "new riddle of induction" can indeed be interpreted as an application of a contextualist theory of justification.

Following Hume, Goodman holds that it is impossible to justify inductive inferences in a foundationalist way. Instead, he puts forward the coherentist idea of making mutual adjustments between rules and accepted inferences. As Goodman himself shows by introducing his famous predicate "grue", this idea of a reflective equilibrium is not sufficient for a definition of the difference between valid and invalid inferences. To approach such a definition, Goodman adds a pragmatic component to his account: The distinction between valid and invalid inferences has to be made against the setting of our previous inductive practice. This

[17] "It seems to me that this contextualist outlook is reinforced by Goodman's arguments" (Williams 2001, 216).

combination of a coherentist and a pragmatic component can be interpreted as an application of a contextualist theory of justification to the Problem of Induction.

Because Williams himself holds that Goodman's account is contextualist, we can also draw conclusions concerning Williams' understanding of his own contextualist theory. Firstly, we can now see that the semantic constraints introduced as context factors by Williams give his account a strong pragmatic touch: If Williams wants to solve the "new riddle of induction" in Goodman's way, the semantic constraints of contexts must lead to a higher initial credibility of those beliefs that are entrenched in our practice. Secondly, by taking up Goodman's idea of a reflective equilibrium between rules and accepted inferences Williams indicates the inclusion of coherentist ideas in his account. While Williams otherwise mentions coherentist elements of his account only in passing and concentrates on foundationalist features of a contextualist structure of justification, this application of contextualism to the Problem of Induction shows the importance of coherentist ideas to a contextualist theory of justification.

References

BonJour, L. (1985): *The Structure of Empirical Knowledge*, Cambridge, Mass.: Harvard University Press, 1985.

BonJour, L. (1999): "The Dialectic of Foundationalism and Coherentism", in: Greco, J.; Sosa, E. (eds): *The Blackwell Guide to Epistemology*, Oxford: Blackwell, 1999, 117-142.

Ernst, G. (2005): "Induktion, Exemplifikation und Welterzeugung", in: Steinbrenner, J.; Scholz, O.R.; Ernst, G. (eds): *Symbole, Systeme, Welten – Studien zur Philosophie Nelson Goodmans*, Heidelberg: Synchron Wissenschaftsverlag der Autoren, 2005, 99-109.

Haack, S. (1978): *Philosophy of Logics*, Cambridge: Cambridge University Press, 1978.

Hume, D. (1975): *Enquiries Concerning Human Understanding and the Principles of Morals*, edited by L.A. Selby-Bigge, third edition, revised by P.H. Nidditch, Oxford: Oxford University Press, 1975.

Hume, D. (1978): *A Treatise of Human Nature*, edited by L.A. Selby-Bigge, second edition, revised by P.H. Nidditch, Oxford: Oxford University Press, 1978.

Williams, M. (1996): *Unnatural Doubts. Epistemological Realism and the Basis of Scepticism*, corrected paperback edn., Princeton, N.J.: Princeton University Press, 1996.

Williams, M. (2001): *Problems of Knowledge*, Oxford: Oxford University Press, 2001.

Wittgenstein, L. (1969): *On Certainty*, Oxford: Blackwell, 1969.

Karl-Georg Niebergall

On "About": Definitions and Principles

Abstract
This paper contains: (i) a list of adequacy conditions for the use of the phrase "φ is about *a*"; (ii) arguments that Goodman's definition of "φ is absolutely about *a*" is unacceptable both from his viewpoint and from the perspective of that list; (iii) general reflections on related topics: on principles vs. definitions, on semantic vocabulary and on free logic.[1]

1. Intuitive basis and simple definitions

1.1 Introduction

Beginning with the paper *About* (1961, reprinted in *PP*), Nelson Goodman has dealt with the question of what it means for a sentence φ to be *about* an object *a* (see also Goodman 1965, the Foreword to the chapter *Relevance* in *PP* and Ullian/Goodman 1977). Somewhat surprisingly, his contributions to this theme have not received much attention.[2] The topic itself ("What is a *sentence* about") should not be viewed as philosophically unimportant or neglected, however. Indeed, Willard Van Orman Quine's well known and highly influential work on *ontological commitment* may well be regarded as a contribution to it. Eventually, Goodman gave two types of explications for the phrase "φ is about *a*": the first he called "φ is absolutely about *a*", the second "φ and ψ are about *a* relative to each

[1] This paper is an elaboration of a talk I gave under the title "About" at the conference "Nelson Goodman: Von der Logik zur Kunst – Ein Rückblick anlässlich des 100. Geburtstags", Munich, August, 5-8, 2006. I would like to thank the participants for comments on the presentation and Sebastian Paasch. Jakob Steinbrenner and, in particular, Wolfgang Heydrich for many valuable suggestions for the paper.

[2] As for texts which are specifically about them, I am aware of (Ullian 1962), (Rescher 1963), (Putnam/Ullian 1965), (Patton 1965a, 1965b), (Hart 1981) (a discussion of which could unfortunately not be included in this paper) and the detailed exposition in (Elgin 1983).

other".[3] This paper is about "about", in general, and about Goodman's "φ is absolutely about *a*", in particular.

I first address the question whether the *definiens* Goodman has put forward for "φ is absolutely about *a*" is adequate to *his own* predefinitional understanding of "φ is about *a*". I think *probably not*. This claim is suggested by a theorem (in section 2) which implies that, if φ is absolutely about *a*, then φ contains a term designating *a* (for first-order φ). And this was probably not intended by Goodman.[4] I am, of course, also interested if Goodman's definition is adequate to *my* understanding of "φ is about *a*". Here, the answer will be *certainly no*. This leads to the second theme of this paper: formulating adequacy conditions for possible explications of "φ is about *a*". The principles I come up with (in section 3) turn out to be incompatible with what Goodman assumes about "φ is absolutely about *a*"; they are, however, also at odds with a definition for "φ is about *a*" of my own making (see section 1.3). Being left without any definition whatsoever may be unsatisfactory; but then, I will show that the adequacy conditions themselves can be turned into a *definiens* of "φ is about *a*" (see section 4). In the concluding section, I address two themes which have not received a proper treatment in this paper (nor in Goodman's work): the explication of "designates" and the replacement of classical first-order logic by free logic.

1.2 Examples and a fallacious definition

To obtain some clarity as to Goodman's intuitions about "about", let's start with some of his own examples. In *About*, he endorses

(1) "Maine has many lakes" is about Maine,

[3] The variants of aboutness presented in *About* (called "φ is *a*-about" there and "φ is rhetorically about *a*" in (Elgin 1983)) and in (Ullian/Goodman 1977) (called "φ is true about *a*"; see also Elgin 1983) are not treated here.

[4] Actually, it is not that easy to determine what Goodman *did* intend. For in *About*, he is somewhat undecided as to whether he should treat "φ is absolutely about *a*" as an *explicans* of "φ is about *a*" or rather as its replacement (see *PP*, 246; when referring to *About*, I follow the page numbering from *PP*). In the latter case, Goodman should have told the reader what his criteria were for a successful definition of "φ is absolutely about *a*". But it is hard to find explicit statements on this in *About*. Let me note that in (Goodman 1956), Goodman adopts a similar attitude towards the use of "individual".

(2) "Paris is a city" is about Paris,

(3) "Maine is smaller than Texas" is about Maine and
"Maine is smaller than Texas" is about Texas.

Moreover, he also is disposed to accept

(4) "Maine is smaller than Texas" is about the smaller-than-relation;

and this plays an considerable role in *About* (see also Elgin 1983). Given Goodman's adherence to nominalism – which means for him, in particular, the rejection of the assumption of classes – I am somewhat puzzled by (4).[5] Be this as it may: I share Goodman's intuitions expressed by (1) to (3), but will not presuppose claims of the sort (4) in what follows.

Important as such examples may be, it is general definitions that are desirable, too. Goodman, for one, was certainly highly interested in them. Indeed, in *About*, he supplies a rather precise *definiens* for "φ is absolutely about *a*" for which no constraints are intended as to the languages L to which φ may belong (see (D3) below).[6] When it comes to the application of his definition, however, Goodman usually makes two – if common – assumptions. *First*, his examples are sentences from English, i.e., ordinary language sentences (or semi-formal translations); actually, many of them are, when viewed as belonging to a quantificational language, atomic sentences.[7] *Second*, general results are stated and substantiated only for formal (in particular, second-order) languages L.

Both with respect to the examples and the definitions, my approach is

---

[5] I do not believe that Goodman's treatment of that problem in the concluding section of *About* is of any help.

[6] Goodman takes the first *relatum* of the aboutness-relation to be a sentence; I will follow him in this and, moreover, use φ, ψ, ... as variables ranging over sentences.

[7] Goodman also discusses several sample sentences "containing" quantifiers in *About*. Thus, he claims: '"Crows are black" is about crows and about black things ...' (*PP*, 253); and Hilary Putnam regards that sentence as being about crows, too (see Putnam 1958). I am reluctant to accept even Putnam's assessment.

similar. That is, I will discuss the phrase "φ is about *a*" for φ belonging to a classical first-order language L which is enriched by English sentences that play the role of atomic sentences of L. For definiteness, L should be relational, possibly with constants. In particular, it should not contain non-classical sentential operators or quantifiers;[8] but intentional predicates may belong to L.

With this background, let's consider as a first attempt the following definition:

(D1) φ is about *a* :⇔ "*a*" occurs in φ.

It is granted that (D1) is simple; but as a general explication of "φ is about *a*" it is intolerable. I present three reasons for that claim. The first two have been addressed already by Goodman (see *PP*, 254 and 265). The third criticism points to a modification (D2) of (D1) which I will investigate afterwards.

The *first* problem with (D1) concerns non-denoting terms. Thus, consider

(I) Paris is a city,

(II) Pegasus has wings.

Whereas I take the truth of (1) to (3) for granted, I think it is far from clear whether (II) is about Pegasus – although the term "Pegasus" occurs in it: for Pegasus does not exist. As a matter of fact, there *are* philosophers who, while agreeing that Pegasus does not exist, still tend to make the claim

(5) (II) is about Pegasus.[9]

[8] In particular, already from an intuitive point of view it may be hard to tell about which objects sentences containing modal operators are (and with this remark, I do not want to address the difficulties surrounding *quantifying-in*). Thus, consider "Possibly, Maine has many lakes". Is it about Maine?

[9] Others might claim that (II) is *meaningless* or *does not express a proposition*. I disagree, and regard the sample sentences considered in this paper as meaningful.

For them, "φ is about *a*" carries no ontological commitment to *a*.[10] My understanding of English (or, for that, of German) is quite different, however. I am inclined to assume that a sentence φ is about *a* only if *a* exists. That is, I adopt the following principle

(P1) If φ is about *a*, then $\exists x(x = a)$.

Actually, the above discussion suggests that the *schema* 'If "*a*" occurs in φ, then φ is about *a*' cannot be sustained. Now, Goodman did not only reject this schema (and accepted (P1); see section 7 of *About*); he held the same attitude with respect to its *converse*, too.

Accordingly, the *second* objection to (D1) is this: (I) is about Paris; therefore, it is also about the capital of France, for Paris *is* the capital of France (thus Goodman). But the expression "the capital of France" does not occur in (I); therefore, given (D1), (I) fails to be about the capital of France. Contradiction.

Put more generally, Goodman took "is about" to be extensional at its second argument place, i.e., he assumed the following principle of extensionality

(Ext1) If φ is about *a* and $a = b$, then φ is about *b*

– and I fully agree. In sum, for a definition or, more generally, a theory of "φ is about *a*" to be acceptable in Goodman's eyes (and also in mine), it must not have (D1) as a consequence, but should include (P1) and (Ext1).

As for the *third* criticism of (D1), let me merely note that (D1) contains a use-mention flaw (which gives sort of an explanation for the failure of (P1) and (Ext1)). In the *definiendum*, "*a*" is used, in the *definiens*, it is mentioned.

[10] It may rather be another way to express '"*a*" occurs in φ' or '"*a*" occurs meaningfully in φ' or 'Some term *k* occurs meaningfully in φ which is synonymous with "*a*"', say.

1.3 A serious definition of "φ is about a"

For (D1) as it stands, the third criticism taken by itself is demolishing.[11] Now, what could be an enhancement of (D1) which *first*, avoids quantification into quotation complexes, and *second*, is nonetheless merely sort of a minimal mutilation of (D1)? Here is an idea: treat all the names of an object as on a par, and replace quantification into quotation complexes by quantification *over these names*. This leads to an *explicans* for "φ is about *a*" which certainly improves on (D1):[12]

(D2) φ is about$_N$ a :⇔ $\exists k$ (k is a constant ∧ k occurs in φ ∧ k designates a)

It may be noted that, in dealing with the semantic predicate "mention", Goodman writes: "A sentence may be said to mention whatever any expression in it designates" (*PP*, 249). Replace "mention" by "be about" in that quotation, and you almost have (D2).

What about (P1) and (Ext1) in the case of (D2)? It is high time to address a certain peculiarity of (P1) and (Ext1): they are logical truths of first-order predicate logic with identity and as such do not exclude anything. First, for each term "*a*", "$a = a$" and, thus, "$\exists x(x = a)$" is logically provable; and (P1) is merely a weakening of this. And, second,

---

[11] Let me nonetheless mention two possible methods for "saving" (D1) from it:

– On the one hand, when (D1) is conceived of as a definition, regard it as a definition-in-use. That is, take the whole context "φ is about *a*" as being introduced by (D1) as (sort of) an abbreviation of '"*a*" occurs in φ'. Under this construal of (D1) it is, for example, easily seen that it is not correct to infer "φ is about *b*" from "φ is about *a*" and "$a = b$". The reason is that '"*b*" occurs in φ' does not follow from '"*a*" occurs in φ' and "$a = b$"; for, as it is well known, the quotation marks introduce a strong form of nonextensionality.

– On the other hand, that way out seems to be closed if (D1) is reformulated as an axiom (let's call it "(PD1)"). For if, e.g., the "*a*" occurring on the left side of (PD1) is a variable, then both (PD1) and its "universal closure" (i.e., '$\forall x$ (φ is about x ⇔ "x" occurs in φ)') are hardly well-formed or understandable. It is possible, however, to view (PD1) as a *schema*: and this, in turn, can be taken to be a set of suitably formed sentences (i.e., the set of *instances of the schema*).

[12] I often abbreviate "k is a constant ∧ k occurs in φ" by "φ contains the constant k". In using "designates" rather than "refers to" or "denotes", I follow Goodman's usage.

(Ext1) is simply an instance of Leibniz's law. Now, since the *definiens* of (D2) is formulated in an extensional first-order language supplied with classical logic, it is therefore trivially guaranteed that (D2) implies (P1) and (Ext1).

Being logical truths, (P1) and (Ext1) should be worthless as criteria for a successful explication of "about". Yet in particular when it comes to (P1), this is not what has been intended. For there is a specific motive for its acceptance: (P1) is – perhaps in conjunction with other assumptions – supposed to yield "¬ ((II) is about Pegasus)", i.e., the negation of (5), as a result. Now, ¬(5) *is* a consequence of (P1) amended with the *negation* of

(III) $\exists x(x = \text{Pegasus})$.

But, using classical first-order logic, it is (III) that can be proved.

In a nutshell: given classical logic, (P1) does not serve its purpose. I regard this as a serious difficulty, on which I will have more to say in section 5.2.

2. Goodman: a theorem and a problem

2.1 Goodman's definition and a theorem

Given a language L and a sentence φ from L, the definition of "φ is absolutely about *a*" Goodman finally adopts is (see *PP*, 253):[13]

(D3) φ is absolutely about $a :\Leftrightarrow \exists\psi k$ [ψ is a sentence of L $\wedge$ ψ contains the constant $k \wedge k$ designates $a \wedge \varphi \models \psi \wedge \forall u$ (u is a part of $k \Rightarrow \varphi \not\models \forall \text{v}\ \psi(\text{v}/u)$)].

Especially when compared with (D2), the *definiens* of (D3) is rather complicated and opaque. I must admit that I find it hard to see what its intuitive motivation and what its consequences are.[14] The same applies to

[13] $\psi(x/u)$ is supposed to be the simple substitution of x for at least one occurrence of u, no matter which sort of term u is.

[14] Thomas Patton (1965a) seemed to have the same difficulties.

the only slightly simpler *definiens* considered earlier by Goodman (see *PP*, 252), for which I write "φ is weakly about *a*":

$\exists\psi k$ [ψ is a sentence in L $\wedge$ ψ contains the constant $k \wedge k$ designates a $\wedge$ φ $\models$ ψ $\wedge$ φ $\not\models \forall$v ψ(v/k)].[15]

A systematic reason for the choice of such complex definitions could be their superiority over simpler competitors. But it is so far anything but clear whether (D2) is inferior. It is, at any rate, not obvious that (D2) should be unacceptable; with respect to the adequacy conditions (P1) and (Ext1), the *definientia* of (D2) and (D3) are at par; and both contain the up to now unexplained predicate "designates" (for more on this, see section 5.1).

In this section, I carry on with the comparative evaluation of (D2) and (D3) by proving a theorem which establishes a close connection between them and thereby threatens Goodman's approach to aboutness. To be specific, let L be equipped with the usual definition of "(first-order) consequence" (cf. Enderton 1972, Prestel 1986), and let φ be a sentence of L. I claim that in this case

(*) *If φ is absolutely about a, then φ is about$_N$ a.*

Actually, I will show: (**) *If φ is weakly about a, then φ is about$_N$ a*. Since, as usually understood, each expression is a (nonproper) subexpression of itself, (**) trivially implies the aimed-at (*).

In order to obtain (**), a certain M*eta*-L*ogical* L*emma* (MLL) turns out to be of crucial importance. For our purposes, it can be stated as follows:

(MLL)[16] If φ $\models$ ψ and $\neg$ (*k* occurs in φ), then there is a *new* variable v

[15] Goodman begins *About* with a discussion of the principle (*): *If φ is about a, and if a is a part of b or b is a part of a, then φ is about b.* He concedes (*) some initial plausibility (an assessment I do not share), but quickly dismisses it for good reasons. Nonetheless, he views (*) as a motivating force for the definitions considered by him later on.

such that $\varphi \models \psi(v/k)$.

Theorem 1:[17] If φ is weakly about a, then φ is $about_N$ a.

Proof: Assume: ψ is a sentence of $L \wedge k$ is a closed term in $L \wedge k$ occurs in $\psi \wedge k$ designates $a \wedge \varphi \models \psi \wedge \varphi \not\models \forall v\, \psi(v/k)$ (i)

Claim: k is a closed term in $L \wedge k$ designates $a \wedge k$ occurs in φ (!)

Since the first and second conjunct of (!) hold by (i), let's assume as the premise of an attempted *reductio*: $\neg$ (k occurs in φ) (+)

Then: $\neg$ (k occurs in φ) $\wedge \varphi \models \psi \wedge k$ occurs in ψ (ii)

By (MLL) (using a new variable w), we have

$\neg$ (k occurs in φ) $\wedge \varphi \models \psi \wedge k$ occurs in $\psi \Rightarrow \varphi \models \psi(w/k)$ (iii)

Now (ii) and (iii) imply that $\varphi \models \psi(w/k)$; and this yields

$\varphi \models \forall w\, \psi(w/k)$ (iv)

by way of $\forall$-introduction. But (iv) is incompatible with (i).

Therefore, given (i), (+) has to be false – whence (!) has to be true. In sum, what we have shown is this: if (i), then (!). The theorem follows.

Corollary 1: If φ is absolutely about a, then φ is $about_N$ a.

---

[16] A proof can be found, e.g., in (Enderton 1972, 116) and, with many details, in (Prestel 1986, 37 ff).

[17] Similar, but weaker results are presented in (Putnam/Ullian 1965) and (Patton 1965a).

2.2 A problem for Goodman

It has already been pointed out by Patton (1965a) that Corollary 1 presents a problem for Goodman's conception of aboutness. For, given that the *definiens* of (D2) implies "φ mentions *a*", Corollary 1 yields: if φ is absolutely about *a*, φ mentions *a*. And Patton interprets Goodman as rejecting just that conditional: "A statement need not mention an object that it is about" (Patton 1965a, 314). Although I have not found such an explicit claim as this one in *About*, I nonetheless think it is fair to ascribe that position to Goodman. For Goodman does state: "... mention of *k* by *S* is neither a necessary nor a sufficient condition for *S* to be absolutely about *k*" (*PP*, 255); and he certainly does not regard that "result" as suggesting the inadequacy of (D3) as an explication of "about".

Now the theorem and the corollary have been established only for first-order languages. Yet, these are the language which Goodman, being a nominalist, preferred anyway.[18] Thus, it really seems that (D3) should be unacceptable given Goodman's own biases. But perhaps *we* should not share them. Let's try a new and different type of approach.

3. Adequacy conditions for "φ is about *a*"

3.1 Preparatory remarks

Both of (D2) and (D3) make (P1) and (Ext1) true; yet, are they admissible in the light of further (general) principles concerning the use or meaning of "φ is about *a*"? In this section, I formulate and investigate such principles, and employ them to appraise (D2) and (D3).

In distinction to section 1, I will here not take Goodman's intuitions about the meaning of "about" as a starting point and criterion for adequacy – but *my own*. In general, I think that when trying to lay down an explication, one should first be as close and faithful as possible to one's own *linguistic* intuitions. In a second step, one could try to do justice to those of the other competent speakers by anticipating them. This is

[18] Let me stress that Theorem 1 and Corollary 1 have been stated merely for first-order languages. It seems that they are not applicable in the case of, e.g., second-order languages; see *PP*, 255, for an example and (Ullian 1962), (Putnam/Ullian 1965), (Patton 1965a) and Goodman's Foreword to Relevance in *PP* for more on this topic.

obviously very hard: for there is a wide range of unpredictable possibilities in which linguistic intuitions may differ from one another. I prefer a more economical strategy: the idea is to put forward principles which are so weak that "everybody" can accept them. Of course, I may have found the "wrong" principles; and agreement cannot be forced, anyway.[19]

The principles formulated in this section will be used as *adequacy conditions* for attempts to explicate "φ is about *a*", where φ is again a sentence from a first-order language L.[20] When laying down adequacy conditions, many approaches are conceivable. I found it convenient to move along the built-up of the formulas from L. It is well known that the set of formulas of L is defined inductively: first come the atomic formulas, and more complex expressions are declared to be formulas if certain components of them have already been taken to be formulas. In addition, when doing model theoretic semantics for L, the procedure of constructing formulas is copied on the semantic level: the evaluation of atomic formulas (in a model, under an assignment) is defined directly; and if a formula φ is built up from formulas φ_1, ..., φ_n, the evaluation of φ is determined by the structure of φ and the evaluations of φ_1, ..., φ_n (relative to that model and that assignment).

When it comes to principles for "φ is about *a*", let me in the spirit of the approach just sketched first state them for atomic sentences φ, and then extend them to sentences containing classical propositional operators and quantifiers (or definite descriptions). But let's begin with *logic*.

3.2 Logical truth and equivalence

First, I think that at least some logical truths (of first-order logic) are not about anything. As examples, I would like to suggest sentences of the form $\ulcorner\varphi\vee\neg\varphi\urcorner$ and $\ulcorner\neg(\varphi\wedge\neg\varphi)\urcorner$.[21] And second, at least some contradictions are not

[19] Moreover, there is the general threat that weak assumptions may yield merely weak conclusions. From a logical point of view, the nontrivial task is to avoid just that.

[20] Let me add that they can also be used in two other roles: they may be taken to be axioms of a *theory of aboutness* and they may be transformed into clauses of a *definiens* of "φ is about *a*". More on this can be found in section 4.

[21] "$\ulcorner$ $\urcorner$" will be used as quasiquotes in Quine's sense (see Quine 1940).

about anything. Third, I think that logically equivalent sentences are about the same objects. In sum, I suggest as principles:

(PTheo) There is a logical truth φ such that $\neg\exists x$ (φ is about x),
(PCon) There is a contradiction φ such that $\neg\exists x$ (φ is about x),
(P↔) If φ is about a and φ is logically equivalent to ψ, ψ is about a.

Let me remark that in all cases, I have Goodman on my side (cf. Elgin 1983): indeed, he takes (P↔) to be particularly evident; see footnote 6 in *PP*, 255, and *PP*, 258. He also endorses that no logical truths and falsities should be about anything; see *PP*, 256. I didn't postulate that; but it is an easy consequence of (PTheo) and (PCon) plus (P↔), of course.

3.3 Atomic sentences

For atomic sentences, one might take part of (D1) as a starting point: thus,

(PAt) If φ is an atomic sentence and φ contains the constant "*a*", then φ is about *a*.

Of course, this is as unacceptable as (D1) itself; (II) is already a counterexample (and we have again the use-mention flaw). Is there a special type of atomic sentences φ such that (PAt) holds for them? The most promising examples that come to my mind are sentences of the sort "$a = a$" (where "*a*" is a constant); they should be about *a*. Yet, take "Pegasus = Pegasus": it would be about Pegasus, then; but Pegasus does not exist.

Now, since one may declare (II) to be false (if not meaningless), a reasonable variant of (PAt) could be

If φ is a true atomic sentence containing the constant "*a*", φ is about *a.*

Intentional predicates provide counterexamples to this principle: take, e.g., "seeks"; now consider

(IV) Reinhold Messner is seeking the Yeti

This is atomic, contains the constant "the Yeti" and is true; but it is not about the Yeti (granted that no Yeti exists). Let's than assume in addition that the closed terms occurring in φ designate:

> If φ is an atomic sentence and φ contains the constant *k* and *k* designates *a* and φ is true, then φ is about *a*.

This might be acceptable: at any rate, I have not found a counterexample. But φ being true does not seem to be a relevant assumption; for example,

(V) Paris is a river

is about Paris – even if it is false. Thus, one could as well adopt

(PAtD) If φ is an atomic sentence and φ contains the constant *k* and *k* designates *a*, then φ is about *a*.

3.4 Propositional operators: "∧", "∨" and "¬"

When it comes to general principles concerning propositionally complex sentences, it is the case of the conditional where I have the least firm intuitions. I therefore will assume that the language L considered here contains only "∧", "∨" and "¬" as sentential operators. Let's treat "∧" first.

(P∧i) φ is about *a* and ψ is about *b* ⇒ ⌜φ∧ψ⌝ is about *a* and ⌜φ∧ψ⌝ is about *b*,
(P∧ii) φ is about *a* and ψ is about *a* ⇒ ⌜φ∧ψ⌝ is about *a*,
(P∧iii) φ is about *a* ⇒ ⌜φ∧ψ⌝ is about *a*.

Clearly, (P∧i) implies (P∧ii) and (P∧iii) implies (P∧ii). But (P∧iii) should be rejected: take ⌜¬φ⌝ for ψ. In what follows, I will accept (P∧i). This is sort of a provisional decision: eventually, weaker principles could be preferable; in this case, some of the arguments presented below may

cease to be correct.[22]

Let's come to "∨" now. To start with, I state all principles resulting from (P∧i) to (P∧iii) by replacing "and" by "and" or "or":

(P∨i) φ is about *a* and ψ is about *b* ⇒ ⌜φ∨ψ⌝ is about *a* and ⌜φ∨ψ⌝ is about *b*,

(P∨ii) φ is about *a* or ψ is about *b* ⇒ ⌜φ∨ψ⌝ is about *a* or ⌜φ∨ψ⌝ is about *b*,

(P∨iii) φ is about *a* and ψ is about *b* ⇒ ⌜φ∨ψ⌝ is about *a* or ⌜φ∨ψ⌝ is about *b*,

(P∨iv) φ is about *a* or ψ is about *b* ⇒ ⌜φ∨ψ⌝ is about *a* and ⌜φ∨ψ⌝ is about *b*,

(P∨v) φ is about *a* and ψ is about *a* ⇒ ⌜φ∨ψ⌝ is about *a*,

(P∨vi) φ is about *a* or ψ is about *a* ⇒ ⌜φ∨ψ⌝ is about *a*,

(P∨vii) φ is about *a* ⇒ ⌜φ∨ψ⌝ is about *a*.

The following can be shown: (P∨ii) is equivalent to (P∨vi); (P∨iv) implies (P∨ii), which implies (P∨i), which implies (P∨iii), which implies (P∨v); (P∨vii) implies both (P∨i) and (P∨ii), provided that sentences which are logically equivalent are about the same.

It should be obvious that whatever implies (P∨vi) is unacceptable (take ⌜¬φ⌝ for ψ in (P∨vi)). But (P∨v) may have its problems, too: for even if φ and ψ are no logical truths, ⌜φ∨ψ⌝ might be one. To be on the safe side, I will only postulate (see section 3.6 for "extended tautology")

(P∨) φ is about *a* and ψ is about *a* and ⌜φ∨ψ⌝ is no extended tautology ⇒ ⌜φ∨ψ⌝ is about *a*.

Finally, let's deal with negation. Here is a typical simple example:

---

[22] If merely (P∧ii) holds, the derivation of (P¬) at the end of section 3.6 can probably not be sustained. And if (P∧ii) is replaced by

φ is about *a* and ψ is about *a* and ⌜¬(φ∧ψ)⌝ is no extended tautology ⇒ ⌜φ∧ψ⌝ is about *a*,

the third argument against (D3) from section 3.7 cannot be upheld.

(VI) ¬ (Paris is a city).

I certainly do not claim

(6) (VI) is about Paris.

I think that an adequate definition or theory of aboutness should not yield (6). For (VI) is the negation of the entire sentence (I). That is, (VI) can also be true if Paris does not exist. In fact, by (P1), (VI) *is* true if $\neg\exists x(x = \text{Paris})$. Let me add that (VI) should not be conflated with, say

(VII) Paris is a non-city.[23]

These considerations may suggest a principle to the effect that negated sentences are always about nothing. But this is no viable idea: for consider

(VIII) ¬¬ (Paris is a city);

it is not implausible that (VIII) is about Paris.[24]

In answer to that, one may contemplate that a sentence which is in the scope of an odd number of negation signs is about nothing. Yet, I fear that the actual development of this idea may get rather complicated. I am therefore content with the principle:

(P¬) If $\exists x$ (φ is about x), then $\neg\ \exists x$ ($\ulcorner\neg\ \varphi\urcorner$ is about x).

To see how the principles considered so far interact, take sentence (VI). (I) is an atomic sentence which contains the constant "Paris"; and let's

[23] Here, "is a non-city" is an artificial predicate which should be understood to be true of an *a* precisely if "is a city" fails to be true of *a*, given that *a* exists. I have introduced this because in ordinary language, one would say neither (VI) nor (VII), but rather "Paris is not a city". This, however, is ambiguous between (VI) and (VII); and the use of "non-city" provides an easy way to make the disambiguation of "Paris is not a city" apparent.

[24] Indeed, given (2), the principle (Pp↔) accepted below does yield (VIII).

assume that “Paris” designates Paris. Then, by (PAtD), (I) is about Paris. Now (P¬) implies that (VI) fails to be about anything; and that is exactly what I wanted.

3.5 Quantifiers

In principle, I would like to stay faithful to Quine’s intuitions and assume that the archetypical examples of sentences which are about something are existentially quantified ones. But which of them? And about which objects are they? With respect to the following sentences, I couldn't tell: “$\exists$v(v = v)” and “$\exists$v(v is a city)”. Actually, I tend to claim that they are about nothing. “$\exists$v(v is a city)”, for example, is not about Paris, not about Rome, not about any other city that has a name in L, and not about anything which is no city. Thus, nothing that can be named (in L, at least) remains about which “$\exists$v(v is a city)” could be.[25]

What is missing here are closed terms. This is not at all in the spirit of Quine; but it is them which introduce the characteristic needed for a sentence to be about a particular thing. Thus, I regard

(IX) $\exists$v(v = Paris)

as being about Paris. More generally, I find this principle agreeable:

$\forall k$ (k is a constant $\wedge$ k designates $a \wedge \ulcorner\exists$v(v = k)$\urcorner$ is true
$\Rightarrow \ulcorner\exists$v(v = k)$\urcorner$ is about a).

What about simplifications? “$\forall k$ (k is a constant $\Rightarrow \ulcorner\exists$v(v = k)$\urcorner$ is about some a)” is unacceptable. Again, “Pegasus” provides the counterexample: “$\exists$v(v = Pegasus)” is not about Pegasus (or else Pegasus would exist). There is, however, a close connection between designation and the truth of existence assertions: for each constant k,

[25] Universally quantified sentences of these types present intuitive problems, too. What about “$\forall$v(v is a city)”, for example? Given one intuition, it is about everything. But then, again, it is not about anything specific. I therefore don’t think that ‘“$\forall$v(v is a city)” is about a’ should be accepted as a principle for aboutness (whatever a may be).

$\ulcorner \exists v(v = k) \urcorner$ is true $\Leftrightarrow \exists y$ (k designates y).[26]

Now this yields: "k designates $a \Rightarrow \ulcorner \exists v(v = k) \urcorner$ is true." Thus, the above principle may be simplified to:

(P∃) $\forall k$ (k is a constant $\wedge$ k designates $a \Rightarrow \ulcorner \exists v(v = k) \urcorner$ is about a).[27]

It is this principle for the quantifiers which I accept.

Of course, (P∃) is not enough. Additional principles dealing with the quantifiers (and with definite descriptions) would be highly welcome. I have some ideas for them; but even a sketchy elaboration would result in a new paper. For my purposes, (P∃) suffices anyway: if principles are used as *adequacy conditions*, it is not so much their completeness (whatever that means) that is required, but their plausibility. And we will see that the principles put forward here *are* strong enough to shed light on the acceptability of (D2) and (D3).

3.6 The accepted principles

By what has been suggested so far, a first try should be (PTheo), (PCont), (P↔), (PAtD), (P∧i), (P∨), (P¬) and (P∃). In tests with sample sentences, they have been rather successful. But here, a problem arises: they are not consistent with each other (given the existence of certain sentences). For:

(PTheo) and (P↔) imply that, being a logical truth, "∃v(v = Paris)" is not about anything. But from (P∃), we obtain "∃v(v = Paris)" is about Paris (assuming that "Paris" designates Paris). Contradiction.

What to do? One has to choose. Here I give an answer which introduces merely minimal changes; section 5.2 contains a different – and, as I think,

---

[26] Note that this need not be taken as a schema: by using "⌢" for concatenation, it can be written as: $\forall x$ (x is a constant $\Rightarrow$ ("∃" ⌢ "v" ⌢ "(" ⌢ "v" ⌢ "=" ⌢ x ⌢ ")" is true $\Leftrightarrow \exists y$ (x designates y))).

[27] Let me be more explicit here, too. (P∃) is: $\forall x$ (x is a constant $\wedge$ x designates $a \Rightarrow$ "∃" ⌢ "v" ⌢ "(" ⌢ "v" ⌢ "=" ⌢ x ⌢ ")" is about a).

superior – treatment.

Let me note first that, although the language L from which our sentences φ are taken is a first-order language, it is nonetheless determined whether such a φ is *of the form* of a logical truth (or a logical falsity) of *propositional logic*. This is the case exactly when φ (or ⌜¬φ⌝) is derivable, employing only *modus ponens*, from logical axioms (stated in L) which have the shape of the common propositional axioms. I will call such a φ an "extended tautology". Now, since it is specifically first-order logical truths such as "∃v(v = Paris)" that create the above mentioned difficulty, the new idea would be to replace mention of *logical truth* (and related notions) in (PTheo), (PCon) and (P↔) by mention of *extended tautologies*. By this procedure we obtain:

(PTaut) If φ is an extended tautology, then ¬∃*x* (φ is about *x*),
(PpCon) If ⌜¬φ⌝ is an extended tautology, then ¬∃*x* (φ is about *x*),
(Pp↔) If φ is about *a* and if ⌜(φ∧ψ) ∨ (¬φ∧¬ψ)⌝ is an extended tautology, then ψ is about *a*.[28]

Note, moreover, that (P¬) follows from (P∧i) and (PpCon). For assume that φ is about *x* and ⌜¬φ⌝ is about *y*; then, by (P∧i), ⌜φ ∧ ¬φ⌝ is about *x*. Thus, if ∃*x* (φ is about *x*) and ∃*x* (⌜¬φ⌝ is about *x*), then ∃*x* (⌜φ ∧ ¬φ⌝ is about *x*) – which, by (PpCon), cannot be the case.

Eventually, this section's reasoning results in a "final" list of principles for "about", which I adopt: (PTaut), (PpCon), (Pp↔), (PAtD), (P∧i), (P∨), (P∃).

3.7 The adequacy conditions applied to (D2) and (D3)

In what follows, I am going to evaluate (D2) and (D3) from the point of view of the principles accepted in section 3.6.[29] To start with, let me cite some relevant results concerning (D3):

[28] I also write "φ ≡ ψ" instead of "⌜(φ∧ψ) ∨ (¬φ∧¬ψ)⌝ is an extended tautology".

[29] Thus, the word "about" occurring in the principles has to be replaced by "absolutely about" or by "about$_N$" where appropriate.

(Gi) If φ and ψ are logically equivalent and if φ is absolutely about *a*, then ψ is absolutely about *a*.

(Gii) Logical truths and contradictions are not absolutely about anything (see *PP*, 256).

(Giii) If φ is absolutely about *a*, then ⌜¬φ⌝ is absolutely about *a* (see Putnam/Ullian 1965).

In the case of (D2), we also have: (*) If φ is about$_N$ *a*, then ⌜¬φ⌝ is about$_N$ *a*.

Now, this implies that (D2) is unacceptable as an explication of "φ is about *a*". The reasons are:

First: (*) is inconsistent with (P¬).

Second: (D2) is inconsistent with (PTaut). – Under the assumption that "Paris" designates Paris, (D2) yields that "Paris is a city ∨ ¬Paris is a city" is about Paris. But, being an extended tautology, it is about nothing.

Third: (D2) is inconsistent with (PpCon). – Under the assumption that "Paris" designates Paris, (D2) yields that "Paris is a city ∧ ¬Paris is a city" is about Paris. But, being the negation of an extended tautology, it is about nothing.

Fourth: (D2) is inconsistent with (Pp↔). – Since "(Paris is a city ∨ ¬Paris is a city)" ≡ "(Munich is a city ∨ ¬Munich is a city)", "Paris is a city ∨ ¬Paris is a city" and "Munich is a city ∨ ¬Munich is a city" sentences have to be about the same. But under the assumption that "Paris" designates Paris and "Munich" designates Munich, (D2) yields that this is not the case.

Furthermore, (D3) is unacceptable as an explication.[30] The reasons are:

First: (Giii) is inconsistent with (P¬).

Second: (Gii) is inconsistent with (P∃).

Third: (Gii) plus (Giii) is inconsistent with (P∧i). – For take a sentence φ which is absolutely about some *a*. Then by (Giii), ⌜¬φ⌝ is absolutely

[30] There are additional problems for (D3). Here is one: consider

(*) If φ and φ' are about the same objects, and if ψ and ψ' are about the same objects, then ⌜φ∧ψ⌝ is about the same objects as ⌜φ'∧ψ'⌝.

Given (D3), (*) is not true (take φ for ψ and φ', and ⌜¬φ⌝ for ψ').

about *a*, and therefore, by (P∧i), ⌜φ ∧ ¬φ⌝ is absolutely about *a*. This contradicts (Gii).[31]

With (D2) and (D3) I have considered two definitions for "φ is about *a*" which, in the light of this sections' adequacy conditions, cannot be sustained. I conclude from this that it is *far from evident* that (D2) and (D3) are adequate as explications of "φ is about *a*". In reaction, one could try to find some amendments to these definitions. When it comes to (D3), however, I personally wouldn't even attempt to do that: for I suspect that my understanding of "φ is about a" is so different from Goodman's that those modifications would turn out to be far-reaching, if feasible at all. At the same time, it has to be granted that (so called) adequacy conditions themselves are not beyond any possible censure. In general, definitions, adequacy conditions and examples have to be well-balanced (see also *FF*, III, 2, for more on topic).[32]

As an extreme, there is the threat of inconsistency: adequacy conditions may be convincing, if not evident – but may nonetheless turn out to be inconsistent. Now, *is* the list of principles chosen here consistent? See the next section for an answer.

4. Principles and inductive definitions

In this part, I will take the correctness of the principles from section 3.7 for granted. What I am interested here is: apart from these principles, what else do we need for the phrase "φ is about *a*" to be philosophically respectable?

For many logicians and philosophers working in the analytic tradition it is of considerable importance to fix the use or meaning of expressions in an explicit and precise way. In addition, some of them think that the only method to attain *that* goal is by providing definitions; while others may believe that this is at any rate the best method. Thus, one answer to the above question is: a definition. Now, the adequacy conditions are certainly

[31] In these refutations of (D2) and (D3), I have at several places assumed without mention that for some φ and some *a*, φ is about$_N$ / absolutely about *a*.

[32] In fact, I think that in principle Goodman was aware of the option of the kind of criticism put forward here: see especially section 4 from *About*. He simply wouldn't have agreed to all of the adequacy conditions put forward here.

no *definiens* for "φ is about *a*". But they also discredit (D2) and (D3) as its *explications*. From the point of view taken here, we are therefore left in the intolerable state of lacking an adequate definition of "φ is about *a*" altogether.

A brief response to this is: in general, axioms are quite as good as definitions; and the adequacy conditions may easily be regarded as axioms. Moreover, both in principle and in practice, axioms are unavoidable: for in giving definitions, one always draws on other expressions – which is a procedure that eventually leads to a circle or to an infinite regress or to a halt point; and that means, at its best, formulating and perhaps accepting axioms.

But let me not press *this* here, but rather emphasize another reply: occasionally, principles can be converted into clauses of *positive inductive definitions*. And whatever an inductive definition really is – it is known at least since Gottlob Frege's investigations how to rewrite it as an *explicit* one (see Frege 1879, Aczel 1977). In such a case, the desired explicit definition is for free.

Let me explain this assertion with the example of our phrase "φ is about *a*".[33] To start with, consider the principles (PAtD), (P∧i), (P∨) and (P∃).[34] They share a certain noteworthy feature: they ensure that ordered pairs <φ, *a*> (φ being a sentence of L) are put *into* the aboutness-relation. Some ((PAtD) and (P∃)) do it outright; while others express that <φ, *a*> should be an element of the aboutness-relation under the assumption that one or several other pairs <φ', *a*'> already belongs to it ((P∧i) and (P∨)). Now, let's replace "about" by some 2-place predicate (alternatively: set-) variable *X*, rewrite the above principles accordingly and conjoin the resulting formulas. By this procedure, we obtain a certain formula $\alpha(X)$:

$\alpha(X) :\Leftrightarrow$
$\forall\varphi\forall x\forall a$ (*x* is a constant $\wedge$ *x* designates $a \Rightarrow <\ulcorner\exists v(v = x)\urcorner, a>\in X$)
and

[33] Actually, I will present a construction which is a bit more complex than the one that would be expected. The reason is that I want to get Corollary 2.

[34] (PTaut) and (PpCon) have not been included here because they are not of the right form: they rather bar the aboutness-relation from being enlarged.

$\forall\varphi\forall x\forall a$ (φ is an atomic sentence $\wedge$ φ contains the constant x
$\wedge$ x designates $a \Rightarrow \langle\varphi, a\rangle \in X$)
and
$\forall\varphi\psi\forall ab$ ($\langle\varphi, a\rangle \in X \wedge \langle\psi, b\rangle \in X \Rightarrow \langle\ulcorner\varphi\wedge\psi\urcorner, a\rangle \in X \wedge \langle\ulcorner\varphi\wedge\psi\urcorner, b\rangle \in X$)
and
$\forall\varphi\psi\forall a$ ($\langle\varphi, a\rangle \in X \wedge \langle\psi, a\rangle \in X \wedge \ulcorner\varphi\vee\psi\urcorner$ is no extended tautology
$\Rightarrow \langle\ulcorner\varphi\vee\psi\urcorner, a\rangle \in X$).

Now, let's define *explicitly*:

(D4-) φ is about$_I$ $a :\Leftrightarrow \forall X\,(\alpha(X) \Rightarrow \langle\varphi, a\rangle \in X)$.

Theorem 2: For "φ is about$_I$ a", the principles (PAtD), (P$\wedge$i), (P$\vee$) and (P$\exists$) hold.

This is a well-known consequence of the fact that all of the clauses of the *definiens* of $\alpha(X)$ are *positive in X* (see Aczel 1977). Furthermore, by the minimality of the aboutness$_I$-relation as explained in (D4-), we also obtain the missing principles (PTaut) and (PpCon).

Theorem 3: For "φ is about$_I$ a", the principles (PTaut) and (PpCon) hold.

Proof: In order to obtain (PTaut), set

$A := \{\langle\varphi, a\rangle \mid \varphi \text{ is about}_I\ a \wedge \varphi \text{ is no extended tautology}\}$.

It can be checked that $\alpha(A)$ is the case (see below for $\alpha(C)$, which is similar but more involved). Now assume "φ is about$_I$ a". By (D4-), $\langle\varphi, a\rangle \in A$ follows; but then, φ is no extended tautology.

In the case of (PpCon), let me first indicate a definition: (for φ from L) φ is a *$\wedge$-$\vee$-combination* iff it is an atomic or an existentially quantified sentence of L or built from such sentences by means of conjunctions and adjunctions only.

Lemma 1: If ζ is a $\wedge$-$\vee$-combination, then $\ulcorner\neg\zeta\urcorner$ is no extended tautology.

Proof: Let ζ be a ∧-∨-combination and fix an assignment *f* mapping each atomic or existentially quantified subsentence of ζ to *the true*. *Claim*: under *f*, ζ is *true* (and therefore, ⌜¬ζ⌝ is no extended tautology). This is shown by an induction on the built-up of ζ:

– If ζ itself is atomic or existentially quantified, the claim holds trivially.

– If ζ is a conjunction or adjunction of φ and ψ, then φ and ψ are ∧-∨-combinations, and each atomic or existentially quantified subsentence of φ as well as of ψ gets assigned the value *true* under *f*. But then, by induction hypothesis, φ as well as ψ are true under (the same assignment) *f*. Thus, in each case, ζ is *true* under *f*.

Proof of Theorem 3, continued: In order to obtain (PpCon), set

$C := \{\langle \varphi, a\rangle \mid \varphi \text{ is about}_I\ a \wedge\ \varphi \text{ is a } \wedge\text{-}\vee\text{-combination}\}$.

It can be checked that $\alpha(C)$ is the case:

– If φ is atomic and contains the constant *x* which designates *a*, then by Theorem 2, φ is about$_I$ *a*. Moreover, φ is a ∧-∨-combination.

– If *x* is a constant which designates *a* and $\varphi = \ulcorner \exists v(v = x) \urcorner$, then by Theorem 2, φ is about$_I$ *a*. Moreover, φ is a ∧-∨-combination.

– If $\langle \varphi, a\rangle \in C$ and $\langle \psi, b\rangle \in C$, then φ is about$_I$ *a* and ψ is about$_I$ *b*; whence by Theorem 2, $\ulcorner \varphi \wedge \psi \urcorner$ is about$_I$ *a* and $\ulcorner \varphi \wedge \psi \urcorner$ is about$_I$ *b*.

Since because of $\langle \varphi, a\rangle \in C$ and $\langle \psi, b\rangle \in C$, φ and ψ are ∧-∨-combinations, also $\ulcorner \varphi \wedge \psi \urcorner$ is a ∧-∨-combination. Thus, $\langle \ulcorner \varphi \wedge \psi \urcorner, a\rangle \in C$ and $\langle \ulcorner \varphi \wedge \psi \urcorner, b\rangle \in C$.

– If $\langle \varphi, a\rangle \in C$, $\langle \psi, a\rangle \in C$ and $\ulcorner \varphi \vee \psi \urcorner$ is no extended tautology, then φ is about$_I$ *a*, ψ is about$_I$ *a* and $\ulcorner \varphi \vee \psi \urcorner$ is no extended tautology. Therefore, by Theorem 2, $\ulcorner \varphi \vee \psi \urcorner$ is about$_I$ *a*.

Since because of $\langle \varphi, a\rangle \in C$ and $\langle \psi, a\rangle \in C$, φ and ψ are ∧-∨-combinations, also $\ulcorner \varphi \vee \psi \urcorner$ is a ∧-∨-combination. Thus, $\langle \ulcorner \varphi \vee \psi \urcorner, a\rangle \in C$.

Now assume "φ is about$_I$ *a*". By (D4-), $\langle \varphi, a\rangle \in C$ follows; but then, φ is a ∧-∨-combination and, by Lemma 1, $\ulcorner \neg \varphi \urcorner$ is no extended tautology.

Theorems 2 and 3 establish that, apart from (Pp↔), all principles

accepted for "φ is about *a*" are true for about$_I$, too. Now, let's define:

(D4) φ is about *a* :⇔ ∃ψ (φ ≡ ψ ∧ ψ is about$_I$ *a*).

It is easy though somewhat lengthy to check that (PTaut), (PpCon), (PAtD), (P∧i), (P∨) and (P∃) also hold for "φ is about *a*" as defined by (D4); in addition, we obviously have (Pp↔) in this case. And since definitions cannot introduce contradictions, we finally obtain

Corollary 2: The list (PTaut), (PpCon), (Pp↔), (PAtD), (P∧i), (P∨), (P∃) is consistent.[35]

Let me sum up. If you already have adequacy conditions at hand, you never need to invest additional work into finding axioms: the adequacy conditions are your axioms. And if the adequacy conditions are of the right form – that is, of the form of clauses of a positive inductive definition – you even do not need to struggle for definitions: the adequacy conditions can be rephrased as components of a suitable *definiens*.[36] It still remains to find good adequacy conditions, of course: and it is here where the specific challenge lies.

5. Two open problems

Let me finish this paper with a brief discussion of two of its features that may be regarded as unsatisfactory: one is that the phrase "*s* designates *a*" is treated as if it were understood; the other is that the languages considered are supplied with classical first-order logic.

---

[35] That is: this list plus set theory is consistent. We could also add (DES) (cf. section 5.1) without introducing an inconsistency.

[36] It may be argued that if principles for "φ is about *a*" are employed as clauses of a *definiens* of "φ is about *a*", they cannot also play the role of adequacy *criteria* (for this definition). That is, we have lost adequacy criteria in employing them as *definientia*. This may be correct; but I do not think it is a problem. If the original adequacy conditions have delivered a useful procedure to determine whether φ *is* about *a*, it's hard to see that they will not do the same when they are components of an inductive definition.

5.1 Designation

The *definiens* of Goodman's (D3) contains the expression "designates". And I, too, have not eschewed employing that predicate in (D2) and (D4). Moreover, "designates" plays a noticeable role in all of these definitions. Thus, the question arises: how is it understood and how shall it be explained?

To start with, there is no definition of "designates" in *About*. Moreover, this text supplies no references to other works in which the missing definitions could be found. Goodman simply *used* "designates". Now, I think he was quite clear about what he was doing. Thus, in the one passage where he explicitly addresses the topic of designation, we find: "Our sole problem (and it will prove troublesome enough) is to determine what a sentence is about, *given* what its terms designate" (*PP*, 248). Yet, I do not know whether that means that he considered the task of explaining "designates" as easy, whence not necessary to carry out, or as difficult, whence intractable in his paper.

In some sense, one could perhaps agree with Goodman's reaction. Since in practice, the procedure of giving definitions has to stop somewhere (be it only provisionally, though), why not with "designates" in the context considered here? In my view, the main (general) reason is that one should not regard the predicate "designates" as being understood well enough to stop the analysis with it: it is in need of analysis itself. Actually, I think there is really some work to do in providing an explication of "designates" or a theory of designation. In particular, as far as I know, there simply does not exist a commonly accepted definition of "*k* designates *a*", be it in philosophy or in logic.

Let me point out here that neither Quine's conception of ontological commitment nor model theoretic semantics seem to be of any help for this task. It is granted that the several *definientia* of "*T* is committed to *A's*" formulated by Quine do not contain the expression "designates"; and in the model theoretic framework, "*k* designates *a* in *M*" or "*k* designates *a* relative to *I*" (with *M* being a model and *I* being an interpretation function) are employed instead of "*k* designates *a*". So there might be a chance of explaining the troublesome "designates" via expressions that are better understood. But in the model theoretic approach there is no convincing

way to define the unrelativized "*k* designates *a*".[37] That is, even if, with "*k* designates *a* in *M*" at hand, we obtain a definition of "φ is about *a* in *M*", model theoretic semantics simply does not provide any contribution to the task I started with. And in the case of Quine's definitions, I think that we have a dilemma: either they are taken to be stated in the framework of model theoretic semantics – and don't work properly; or they are viewed as belonging to a non-model theoretic formal semantics – but this has not been developed by Quine (cf. Niebergall 2005 and the literature mentioned there).[38]

As a possible reply, it may be claimed that a precise explanation of "designates", though desirable in principle, is not necessary for the above *definientia* to work properly. Here I strongly disagree; let me explain why.

(D2) to (D4) should, for example, yield statement (2). Now, it is rather obvious that this is not so in the case of (D2) and (D4); and because of Theorem 1, (2) also does not follow from (D3) alone. The main missing assumption is:

(7) "Paris" designates Paris.

Of course, I do not want to claim that for each constant "*k*" it is the case that "*k*" designates *k*: "Pegasus" does not designate Pegasus. But at least for some constants "*k*", it should be assumed that "*k*" designates *k*: i.e., for

---

[37] Note that, whether reasonable or not, a definition of the sort: *k* designates *a* :⇔ *k* designates *a* in $M \wedge M = \langle$the real world W, the correct interpretation (of L) in W$\rangle$ is in any case beyond what can be done in the model theoretic framework proper.

[38] This may be an appropriate place for two remarks on a comparison of Quine's "*T* is committed to *A's*" with Goodman's "φ is absolutely about *a*". First, it is noteworthy that, although the *definientia* of these predicates are obviously quite different from each other, they seem to have in common that they are not stated in a model theoretic framework. Second, I think that those predicates are supposed to play quite different roles. For whereas Goodman's "φ is absolutely about *a*" does and should yield "$\exists x(x = a)$", that sentence is not intended to be a consequence of the Quinean "*T* is committed to *a*". In this respect, "*T* is committed to *A's*" is closer to Goodman's "φ is *a*-about" (see *About* for its definition) then to his "φ is absolutely about *a*".

"Paris", "Maine", "Texas". Thus, we really have a *schema* here:[39] a set (DES) of sentences of the specific form

"*k*" designates *k*,

among them (7), but not '"Pegasus" designates Pegasus'.

That there are terms "*k*" and objects *k* such that "*k*" designates *k* is not part of or implied by the *definientia* of (D2) to (D4) – it is rather presupposed when applying them. As a matter of fact, I have assumed the truth of (7) at several places: at the end of section 3.4, in section 3.6 and in the critique of (D2) in section 3.7. Thus, we employ (DES) in order to get definitions (D2) to (D4) off the ground. Moreover, it is easy to obtain results for the latter the moment we have (DES). The reason is that, apart from "designates", all expressions occurring in the *definientia* of (D2) to (D4) may be taken to be sharply defined. In fact, widely accepted *explicantia* for them are known and in use; and we have a rather clear informal understanding of them.[40]

In sum, of the expressions occurring in the *definientia* of (D2) to (D4), "designates" seems to carry most of the weight. For this reason, its explication is too important to be skipped from the program of understanding and justifying (D2) to (D4). A failure to analyze "designates" is a weakness that my approach to aboutness shares with Goodman's.

5.2 Logical truths reconsidered: free logic

Consider sentences (IX), i.e., "∃v(v = Paris)", and

(X) ∃v(v = Pegasus)

---

[39] '∀*x* ("*x*" designates *x*)' is not adequate, given the common understanding of quotation marks

[40] It is granted that the *definientia* of (D2) to (D4) may be regarded as problematic because they contain set-theoretical vocabulary. But since, apart from "designates", that vocabulary may be taken to consist entirely of syntactically defined expressions (replace "|=" by "|–"), a reformulation of these definitions along nominalistic lines seems feasible.

which belong to our extended first-order language L. It is plausible that the expressions “Paris” and “Pegasus” occurring in them are construed as constants; in particular, I will not suppose them to be eliminated in favour of definite descriptions. But then, (IX) and (X) are logical truths of classical first-order predicate logic with identity. Indeed, for each constant *k* from L, the sentence ⌜∃v(v = *k*)⌝ is a logical truth of classical first-order logic.

Now, it is the conventional wisdom to regard (IX) (and (X)) as formal paraphrases of (IX)’ and (IX)’’ (and (X)’ and (X)’’):[41]

(IX)’ Paris exists.
(IX)’’ There is something which is Paris.
(X)’ Pegasus exists.
(X)’’ There is something which is Pegasus.

Intuitively, however, these sentences are not at all logical truths. On the one hand, this should be especially obvious for (X)’ and (X)’’, which, in my understanding, are even false. On the other hand, I take it that (IX)’ and (IX)’’ are about Paris; yet, for a sentence φ to be a logical truth it should not be allowed that φ be about something, let alone about something specific.

Note that also for (IX), the list of axioms presented in section 3.6 implies that it is about Paris. For (7) is accepted (and “Paris” is a constant) and (P∃) is assumed to hold. With (IX), we therefore have in this formal setting, too, a logical truth which is about something specific. But that I regard as untenable.

To put it more generally: the assumption that, for each constant *k*, a sentence of the form ⌜∃v(v = *k*)⌝ is a logical truth should be relinquished anyway. Actually, it is well known that classical first-order logic cannot adequately deal with sentences containing non-denoting closed terms (apart from treating them as being eliminated). Yet, to be fair, it was not invented to accomplish that: for this, one should turn to so-called *free logics* (see Bencivenga 1983 for a survey). Furthermore, even when in

[41] In what follows, a referential reading of the quantifiers is presupposed. I regard it as the most natural one, anyway; and I will not attempt to defend it.

⌜∃v(v = *k*)⌝ the constant *k* does designate, I view it as counterintuitive to claim (a) ⌜∃v(v = *k*)⌝ *is a logical truth*; and I certainly reject the conjoined assumption of (a) with (b): ⌜∃v(v = *k*)⌝ *is about a specific object*.

Now, since I have no problems with (7) and (P∃), and since these plus "(IX) is a logical truth" yield: "(IX) is about Paris" (giving us (a) and (b) in this case), it is obvious (for me, at any rate) what to abandon: "(IX) is a logical truth".[42] This, then, is my first suggestion:

(α) Adopt free logic for L. – More precisely: replace the set of classical logical truths formulated in L by the set of logical truths of free logic (in L).

This way of putting things has a, albeit minor, weakness: it presupposes that there is, given L, a uniquely determined set of the free logical truths; but this is not the case. The reason is that there are many different calculi that come under the heading "free logic". This is not the place for a discussion and comparison of them (see, again, Bencivenga 1983). Let me merely note that all of them seem to share the replacement of "∀vψ → ψ(*t*/v)" by "∀vψ ∧ ∃v(v = *t*) → ψ(*t*/v)" as a logical truth. And ⌜∃v(v = *k*)⌝ is not always, i.e., for each constant *k*, declared to be a logical truth.

Personally, I think that there is no constant *k* at all for which ⌜∃v(v = *k*)⌝ should be considered a *logical truth*: not even "God" or "1". Accordingly, I opt for an understanding of "free logic" in this minimal sense: the set of logical truths of free logic (in L) should for no constant *k* (of L) contain a sentence of the form ⌜∃v(v = *k*)⌝.[43]

---

[42] Note that the replacement of (PTheo) plus (P↔) by (PTaut) plus (Pp↔) in section 3.6 helps us in avoiding a contradiction, but still leaves us with "(IX) is about Paris".

[43] Whether it is defined or determined by principles – the phrase "φ is about *a*" may fall victim to the notorious *slingshot* argument (see Barwise/Perry 1981). Here is an example:

Start with (2), which is supposed to be true. Now (I) is logically equivalent to

(*) {v | v = v} = {v | v = v ∧ Paris is a city}.

Thus, by (P↔), (*) is about Paris. Now, apply the principle (!):

if φ(*k*) is about *a* and *k* is coreferential with *k'*, than φ(*k'*/*k*) is about *a*

This much for free logic when dealing with the object language L. It is reasonable, however, that it should also enter the picture with respect to *Meta-L*, i.e., that metalanguage of L in which the *definientia* of and the principles for "φ is about *a*" have been stated. For already in section 1.3, I had pointed out that (III) should not be a accepted as a logical truth if (P1) were to fulfil his task.[44] Now, what this suggests is that the logic of Meta-L should also be a free instead of a classical one. I therefore propose:

(β) Adopt free logic for Meta-L. – And interpret (β) in the same way as (α).[45]

When dealing with Meta-L, one pressing question remains, however: what happens to the principles for "φ is about *a*" adopted above? I think they all can be retained. (PTaut), (PpCon), (Pp↔), (P∧i) and (P∨) should

---

to the formula "(*) is about Paris" and the coreferential terms (*k* =) "{v | v = v ∧ Paris is a city}" and (*k'* =) "{v | v = v ∧ Munich is a city}" (forget about set-theoretic niceties, like: there is no universal set, given set theories such as ZF). Replacing *k* by *k'* results in

(**) {v | v = v} = {v | v = v ∧ Munich is a city};

and (!) therefore yields: (**) is about Paris. Now, by (P↔) again, we obtain:

"Munich is a city" is about Paris.

Let me answer this with some brief remarks: (a) (P↔) has already been dropped in favour of (Pp↔). Moreover, (I), for example, is not logically, but merely set theoretically equivalent to (*). (b) (!) holds in the case of (D2) and, at least for atomic sentences, in the case of (D4); I have no problems with this. (c) Apart from any slingshot, (*) *is* about Paris when given (D2): for (*) contains the constant "Paris", which designates Paris. This is indeed a further problem for (D2). (d) A slingshot can also be stated with definite descriptions instead of class terms. Now, it is granted that, e.g., the claim "φ is logically equivalent to ⌜ιv(v = *k*) = ιv(v = *k* ∧ φ)⌝" is more plausible than "φ is logically equivalent to ⌜{v | v = *k*} = {v | v = *k* ∧ φ}⌝" – *if* the constant *k* denotes. But the latter simply need not be the case when free logic is adopted.

[44] Let me emphasize here that (III) is not (X): (X) is a sentence of L. By contrast, (III) is a sentence which is supposed to be the consequent of (P1), and thus belongs Meta-L.

[45] Of course, both (α) and (β) allow that sentences of the form ⌜∃v(v = *k*)⌝ (or ⌜∃*x*(*x* = *k*)⌝) may be true (though not logically true).

present no problem, since there is no essential occurrence of quantifiers in them. And in (PAtD) and (P∃), the assumptions that certain relevant expressions do designate are even made explicitly.

It may moreover be necessary, now that we have a free logic as the logic of Meta-L, to add further principles. One that contains "about" is (P1). Empty as it is in the framework of classical first-order logic, does it become substantial and even of decisive importance when merely a free logic is adopted. In addition, one should also consider to add principles for "designates". Thus, speaking of section 1.3: with only free logic at hand, the claim that (D2) yields (P1) and (Ext1) is not trivial any longer. Here is a way to argue for (P1) and (Ext1), given (D2), in our new framework.

Ad (P1): If φ is about$_N$ a, then some constant k occurring in φ designates a; therefore $\exists x(x = a)$.
Ad (Ext1): If φ is about$_N$ $a \wedge a = b$, then some constant k occurring in φ designates a; and since $a = b$, k designates b, too. Therefore, some constant occurring in φ designates b, whence φ is about$_N$ b.

In this reasoning I have employed two principles which are the analogues of (P1) and (Ext1) on the level of terms. They are: for constants k from L,

(P2) If k designates a, then $\exists x(x = a)$,
(Ext2) If k designates a and $a = b$, then k designates b.

Like (P1) and (Ext1), I take them to be evident also in the context of free logic. But they are no longer logical truths: they have to be postulated separately.[46]

[46] It could be worthwhile to envisage an axiomatization of "k designates a". A possible choice is: (P2), (Ext2) and (DES). I do not regard this as being particularly convincing, however. First and foremost, an account of the conditions stating for which constants "k", a sentence of the form '"k" designates k' should belong to (DES), is missing; and it seems hard to formulate one. In addition, (P2), (Ext2) and (DES) seem to be rather weak in their consequences. Finally, I think that for a convincing axiomatization of "designates", its interaction with other (e.g., semantic) predicates should be made explicit.

Of course, replacing classical first-order logic by free logic may have considerable consequences as to the viability of the reasoning presented in other parts of this paper. Take sections 3.6 and 3.7, in particular: as a means to discredit certain principles and (D3), for examples, I used both that (IX) is about Paris[47] and that (IX) is a logical truth, whence not about anything. Now, in the framework of free logic, the second horn of the dilemma thereby created simply disappears. And what happens to the inductive definition of section 4; must free set theory be applied? It remains to be investigated if such modifications would lead to a decisive reevaluation of the principles and *definientia* considered here.

References

Aczel, P. (1977): "An introduction to inductive definitions", in: Barwise, J. (ed.): *Handbook of Mathematical Logic*, Amsterdam: North-Holland, 1977, 739-815.

Barwise, J. / Perry, J. (1981): "Semantic innocence and uncompromising situations", in French, P.; Uehling, T.; Wettstein, H. (eds): *Midwest Studies in the Philosophy of Language*, VI, Minneapolis: University of Minnesota Press, 1981, 369-381.

Bencivenga, E. (1983): "Free logics", in Gabbay, D.; Guenther, F. (eds): *Handbook of Philosophical Logic. Vol. 3: Alternatives to Classical Logic*, Dordrecht, Holland: Reidel, 1983, 373-426.

Elgin, C. (1983): *With Reference to Reference*, Indianapolis, Cambridge: Hackett Publishing, 1983.

Enderton, H.B. (1972): *A Mathematical Introduction to Logic*, San Diego: Academic Press, 1972.

Frege, G. (1879): *Begriffsschrift: eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*, Halle, 1879.

Goodman, N. (1956): "A world of individuals", in: *The Problem of Universals: A Symposion*, Notre Dame, Ind.: U of Notre Dame P, 1956. Reprinted in *PP*, 155-171.

[47] It would have been enough to assume that "$\exists v(v = \text{Paris})$" is about something.

Goodman, N. (1961): “About”, in: *Mind* 70 (1961), 1-24; reprinted in PP, 246-272.

Goodman, N. (1965): ‘“About” mistaken’, in: *Mind* 74 (1965), 248; reprinted in PP, 273.

Hart, A.M. (1981): *Aboutness: a Logico-Philosophical Analysis*, Dissertation, Michigan State University, 1981:

Niebergall, K.G. (2005): “Zur nominalistischen Behandlung der Mathematik”, in: Steinbrenner, J.; Scholz, O; Ernst, G. (eds): *Symbole, Systeme, Welten: Studien zur Philosophie Nelson Goodmans*, Heidelberg: Synchron Wissenschaftsverlag der Autoren, 2005, 235-260.

Patton, T.E. (1965a): ‘Some comments on “About”’, in: *Journal of Philosophy* 62 (1965), 311-325.

Patton, T.E. (1965b): ‘A reply to Rescher on Goodman’s “About”’, in: *Mind* 74 (1965), 592-593.

Prestel, A. (1986): *Einführung in die Mathematische Logik und Modelltheorie*, Braunschweig/Wiesbaden: Vieweg & Sohn, 1986

Putnam, H. (1958): ‘Formalization of the concept “About”’, in: *Philosophy of Science* 25 (1958), 125-130.

Putnam, H. / Ullian, J.S. (1965): ‘More about “About”’, in: *Journal of Philosophy* 62 (1965), 305-310.

Quine, W.V.O. (1940): *Mathematical Logic*, New York: Norton, 1940.

Rescher, N. (1963): ‘A note on “About”’, in: *Mind* 72 (1963), 268-270.

Ullian, J.S. (1962): ‘Corollary to Goodman's explication of “About”’, in: *Mind* 71 (1962), 545.

Ullian, J.S. / Goodman, N. (1977): ‘Truth about Jones’, in: *Journal of Philosophy* 74 (1977), 317-338.

Richard Schantz

Goodman on Truth

Abstract
Nelson Goodman belongs to the camp of those philosophers who attempt to understand the concept of truth in epistemic terms – in such terms as verifiability, warranted assertibility, permanent acceptability, superassertibility or justifiability in ideal epistemic situations. Goodman's original proposal was to define truth as permanent credibility. Later, however, he weakened this claim and merely suggested that permanent credibility is a sufficient condition for truth. The author argues that an epistemic status is neither a necessary nor a sufficient condition for truth and thereby demonstrates that neither Goodman's nor any other epistemic conception of truth can work.

I

In the contemporary philosophical debate on the concept of truth, we can, with a certain simplification, distinguish two factions (cf. Schantz 1996, 2002). On the one hand, there are those philosophers who maintain that the concept of truth is an important, indispensable, a profound or substantial concept – in any case, a concept worth defending. Within this faction, we can further differentiate between those philosophers who have committed themselves to a realist view on truth from those who support an *epistemic* analysis of the concept. On the other hand, there are philosophers who assert that the advocates of the first faction suffer from a major delusion, that, in principle, they are chasing after a chimera when they assume that truth has some underlying nature that could be analyzed epistemically, ontologically, or semantically. The advocates of the second faction make the radical assertion that truth is not a substantial or explanatory concept, not a concept that expresses some interesting property or some interesting relation. On the contrary, in their view truth is a purely formal or logical concept, a concept whose correct explanation requires far less extravagant conceptual resources than adherents of substantive theories believe. Partisans of the second faction advocate a deflationary or minimalist analysis of truth (cf. Horwich 1990; Quine 1990).

I advocate a realist conception that sees truth as a relationship between a truth bearer – a statement or belief – and a truth maker, a reality that has an objective status with respect to the truth bearer. Truth makers are familiar entities: objects having certain properties and standing in certain relations to each other at various spatiotemporal locations. Further, I think that truth can be explained by objective referential relationships between language and the thought expressed on the one hand and the external world on the other (cf. Schantz 1996, 2001a). Such a referential explication, then, allows us to develop a plausible form of the correspondence theory of truth that can take seriously the intuitive idea that a statement is true if and only if the fact or state of affairs actually obtains whose obtaining is asserted by the statement. Correspondence does not require something as pretentious as a relation of structural isomorphism between statements and facts, as early Wittgenstein und Russell once thought (Wittgenstein 1922; Russell 1912). Ordinary reference is enough.

The fact that tomatoes are red is both a necessary and a sufficient condition for the truth of the statement that tomatoes are red. It is true that *p* if and only if *p*. Facts, as I see them, are not completely artificial, "sentence-shaped" objects, mere shadows cast by our linguistic practice of making statements, as Peter Strawson and Donald Davidson suggest (cf. Strawson 1950; Davidson 1990). And facts are not true thoughts either, as Frege (1919) contended. Rather, facts are components of the objective, mind-independent spatiotemporal world. Neither the existence nor the nature of facts depends, in any philosophically significant sense, on what we believe, think or perceive. Moreover, an essential element of the realist view is that truth transcends verifiability, that a statement can be true even though *we*, beings with our sensory and cognitive nature, are not at all able to verify it. Truth and objective reality cannot be reduced to what *we* are able to ascertain. Truth, according to alethic realism, is evidence-transcendent; it exceeds rational acceptability (cf. Schantz 1996).

II

It is well-known that Nelson Goodman belongs to the sharpest critics both of the correspondence theory of truth and of alethic realism in general. In fact, he is opposed to the prevailing tendency to overestimate the scope and

the philosophical significance of the concept of truth. According to him, truth is just one species of a more general and more important virtue he calls "rightness" (*WW*, 1-22). Whereas truth applies only to statements, rightness comprises all standards of acceptability of linguistic and nonlinguistic symbol systems. So works of art, pictures and symphonies, e.g., also constitute versions of the world, and are the bearers of semantic properties. Both scientific and artistic symbol systems have a cognitive function. They do not just mirror aspects of the world, but are, rather, productive forms of representation.

It is one of Goodman's central ideas that there is no such thing as a unique right version of the world. There is, rather, a plurality of incompatible and yet right versions. He likes to illustrate his point with geometrical examples. Points, for example, can be identified with sets of concentric spheres or, alternatively, with intersections of three planes. One can also invoke examples from physics. In Newtonian physics, physical events can be reduced either to particles acting at a distance, across empty space, or to particles interacting locally with electro-magnetic fields.

Here we seem to have two incompatible theories about the world. For either there are or there aren't fields which mediate the action of separated particles on each other. Relative to field theory, the statement "There are fields" is true, relative to its competitor, the particle theory, however, this statement is false. That is, apparently, what the theory-relativity of this statement consists in. The question whether there really are electro-magnetic fields does not seem to have a theory-independent answer. Therefore, it is tempting to reject the question which of the two theories is really true as a metaphysical pseudo-problem. Both theories, so Goodman would claim, can be simultaneously true. After all, they are empirically and mathematically equivalent; they are mathematically intertranslatable.

Surprisingly, Goodman's conviction that there is no unique right version of the world is accompanied by the idea that we do not live in one world but in many different worlds at the same time. And these worlds, he contends, are real worlds, not the merely possible worlds that fans of the semantics of possible worlds go into raptures about (*MM*, 31). Moreover, he even maintains, in a manner reminiscent of Fichte's and Hegel's idealism, that we have made these worlds – not, of course, ex nihilo, but out of previous worlds. According to Goodman, there is no ready-made

world, no world dictating its only true description. So he defends pluralism against monism, antirealism against realism, and a form of relativism against absolutism – a form of relativism, however, that retains the difference between rightness and falsity and hence discards the principle that "everything goes".

Israel Scheffler (1980) has drawn attention to an important ambiguity in Goodman's use of the key expression "worlds". Sometimes Goodman uses this expression simply to mean versions; this is a "versional" use with "versional reference". The pluralism resulting from this use resumes the commendable and fruitful doctrine, developed already in the *Structure of Appearance*, that all prephilosophical subject matters can be systematized in ways that conflict with one another. Elsewhere, however, he means by "worlds" objects that the versions refer to or describe. This "objectual" use with "objectual reference" includes: "the many stuffs – matter, energy, waves, phenomena – that worlds are made of." (*WW*, 6) Goodman himself gives emphasis to the weighty difference between versions with reference and versions without reference (*WW*, 96). "Don Quixote" and "unicorn" are empty terms, terms without any referential function. "Stephen Hawking" and "planet" possess referential functions; by using them we refer to existent things. It is important to distinguish between these two senses of "world". For surely, as Scheffler grants, by using symbols we make worlds in the sense of versions, but it does not follow therefrom that we also make the worlds to which these versions refer. It is indeed we who make world versions, but it is not we who make these versions true or right, or who make the terms they contain into non-empty, referential terms.

Goodman concedes the ambiguity in the use of "world" diagnosed by Scheffler (*MM*, 119). His answer is that, although the versional and the objectual interpretation conflict with one another, they are both right and often exchangeable. He does admit that versions are not worlds; after all, some versions consist of words, but their worlds do not. Nonetheless he insists that worlds depend on versions because, as he expounds, the true versions make worlds. But he still owes us a really satisfying explanation of how in making versions we also make worlds. In the end, for Goodman, in contrast to Scheffler and the point of view of reflected common sense, there is no significant difference between regarding true versions as

descriptions of worlds, or supposing instead that there are no worlds at all but only versions. He does not attach great importance to ontological or metaphysical statements of a problem. So he writes:

> The realist will resist the conclusion that there is no world; the idealist will resist the conclusion that all conflicting versions describe different worlds. As for me, I find these views equally delightful and equally deplorable – for, after all, the difference between them is purely conventional (*WW*, 119).

In keeping with his anti-metaphysical stance, he designates his position not as "antirealism" but rather as "irrealism". By employing this label he does not want to give expression to the view that everything or even anything is unreal, but that the whole dispute between realism and antirealism is undermined by the insight that the boundary between convention and content is arbitrary and variable.

In his counter-move Goodman challenges Scheffler to specify an aspect of the world, e.g. of stars, that is independent of all versions or discourses. I share with Scheffler the unwavering persuasion that we did not ourselves make the stars and their physical properties. We have, by creating and shaping our language, also created the concept *star* with its, as must be admitted, partly conventional boundaries. Our language cannot be neatly divided into two parts, one part describing the world as it is in itself, and the other part describing our conceptual contribution. But from the fact that the concept *star* comprises conventional elements it does not follow that we are responsible for the concept's being true of certain things, or that we have made those things. The discourse-dependence of the concept *star* by no means implies that the stars themselves are discourse-dependent. Obviously, we cannot describe the world independently of our versions. But we should take care not to draw from this truism the absurd conclusion that there are only versions or discourses in the end. His incontestable philosophical brilliancy and the acumen of many of his argumentations cannot protect Goodman from manoeuvring himself, with several of his formulations, into dangerous closeness to Jacques Derrida's project of deconstruction. I wished we could overlook his idealist jargon as mere rhetoric. But Goodman explicitly denies this.

III

This is, roughly sketched, the background against which Goodman rides his attacks against the correspondence theory and against the apparently closely allied doctrine of metaphysical realism. In *Ways of Worldmaking* he says:

> We cannot test a version by comparing it with a world undescribed, undepicted, unperceived, but only by other means … all we learn about the world is contained in right versions of it; and while the underlying world, bereft of these, need not be denied to those who love it, it is perhaps on the whole a world well lost (*WW*, 4).

According to Goodman, we can compare a version only with other versions, not with the world itself. When we compare our theories with sensory data, we compare them with that version of the world that experience is presenting us. In his eyes, experience is not a simple, unvarnished givenness but itself the product of a process of interpretation regulated by theory and by culture (*WW*, 92-93).

I think, however, that his assumption that the correspondence theory requires a comparison with naked, unconceptualized reality is pretty implausible. Why should an advocate of the correspondence theory not be able to describe, to conceptually articulate, the facts that make our statements true? The statement that tomatoes are red corresponds to the fact that tomatoes are red which I have just described by using the words "tomatoes" and "red".

It is clear, however, that Goodman's critique of the correspondence theory is epistemologically motivated. What he seems to miss in this theory is a useful test or an effective criterion of truth. After all, we also want to know whether the statements we make and the beliefs we hold are true or false. And he is deeply convinced that the correspondence theory cannot succeed in establishing our cognitive access to reality. For there is, according to his central objection, that in similar forms has also been raised by the Logical Empiricists Neurath and Hempel, and, more recently, by Davidson (1986, 307), Putnam (1983, VII) and Michael Williams (1977, 112), no pure, unmediated consciousness of external objects or facts as they are in themselves, independently of our ways of conceptualizing them. Therefore, we cannot compare our statements and beliefs – our linguistic and nonlinguistic representations – with the world itself in order to see

whether they agree or correspond with it. The alleged comparison of a belief with raw reality finally turns out to be merely a comparison of a belief with other beliefs. We cannot break out of the circle of our beliefs to get hold of the objective reality itself. Hence beliefs, so runs the conclusion, can only be justified by the internal coherence in a system of beliefs.

This popular objection overlooks the epistemological neutrality of the correspondence account. This account tries to tell us only what the nature or essence of truth is. It does not pretend to provide a criterion of truth telling us how to find out what is true and what is false. Nevertheless, I think it is worthwhile to throw a short, critical glance at this influential objection which attempts to show that the correspondence theory leads to epistemological skepticism in the end because it implies that objective facts are beyond our cognitive ken. I have scrutinized and rejected this argument elsewhere and defended the intuitive standpoint that we are indeed capable of comparing our beliefs with the facts (cf. Schantz 1999, 2001b). Why is such a comparison supposed to be impossible? I have found no compelling reason for this supposition. Indeed I think that no mysterious feat is required for comparing our beliefs with the world. All that is required are commonly accepted methods for acquiring knowledge. So we have, under standard conditions, through perception a direct cognitive access to external facts, an access that enables us to ascertain whether or not these facts render our beliefs true. But the directness or immediacy is not really essential. To determine whether the statement that *p* corresponds to a fact we just have to determine whether *p* – no matter how, whether directly or by more or less indirect routes, such as inference to the best explanation. Thus the correspondence theory by no means leads to the absurd conclusion that beliefs and judgments are the only things to which we have a cognitive access.

Goodman, however, thinks otherwise. His credo is coherence instead of correspondence or confrontation. And surely, the internal coherence of our beliefs is of crucial importance to the theory of epistemic justification. In this context, Goodman rightly emphasizes that coherence alone is not sufficient to determine truth, but has to be supplemented by judgments that express an "initial credibility" (*WW*, 125). But, in my opinion, the theory of epistemic justification is one thing, and the definition or analysis or

theory of truth is another. I myself defend a correspondence theory of truth and a theory of justification according to which both the coherence of our versions and sensory experience are relevant to the warrant of our beliefs (cf. Schantz 1996, 1999, 2001b).

Goodman is of a different opinion. He thinks that truth and test of truth cannot be separated from each other at all. He even goes so far as to contend: “Truth, like intelligence, is perhaps just what the tests test; and the best account of truth may be an 'operational' one in terms of tests and procedures used in judging it” (*WW*, 122). Hence he seems to have some kind of an epistemic analysis of truth in mind. In his reply to Hilary Putnam’s fervent commentary on *Ways of Worldmaking* he speaks of: “my suggestion that truth might be equated with permanent credibility” (Goodman 1980, 211), and elsewhere he recommends „construing truth as ultimate acceptability: as acceptability that is not subsequently lost“ (Goodman 1982, 359).

Goodman seems to join the camp of those philosophers who support an epistemic analysis of truth, an analysis which tries to elucidate the concept of truth by an epistemic conceptual repertoire, that is, by such concepts as permanent credibility, verifiability, warranted assertibility, superassertibility or justifiability under ideal conditions. According to the point of view of alethic antirealism, truth consists, roughly speaking, not in a relationship to an external fact, but rather in the positive epistemic status of the truth bearer within our language, our thought or our experience.

IV

Putnam, and already much earlier Carnap, have shown that truth cannot be simply identified with verification or justification (cf. Putnam 1981, 55; Carnap 1936). Against this simple-minded identification the convincing reason can be brought to bear that truth is a stable property of statements, a property that cannot be lost, whereas justification is tensed and therefore can be lost. The statement “The earth is flat” was, most probably, justified 3,000 years ago, but it is no longer justified today. It would, however, be wrong to say that the statement “The earth is flat” was true 3,000 years ago, for that would mean that the earth has fundamentally changed its shape in the meantime. The conclusion we have to draw from this is that

truth cannot be simply equated with justification. Of course, Goodman is well aware of this objection and tries to dispel it in the following way:

> Consider for a moment the notion of permanence - taken here to mean lasting forever after some given time. Although we can never establish permanence of an object or material, we can establish durability in varying degrees short of permanence. Likewise, although we can never establish total and permanent credibility, we can establish strength and durability of credibility in varying degrees short of that. Shall we then identify unattainable total and permanent credibility with unattainable truth? To the ready protest that we might have total and permanent belief in falsehood – that what is totally and permanently credible might not be true – perhaps the answer is that so long as the belief or credibility is indeed total and permanent, any divergence from the truth could not matter to us at all. The proposal seems to be that if there is any such divergence, so much the worse for truth: scrap it in favour of total and permanent credibility, which though equally unattainable is explicable in terms of what is attainable, just as permanence is explicable in terms of durability (*WW*, 123-124).

Hence we should, so runs Goodman's radical proposal, throw the concept of truth overboard, if it diverges from permanent credibility. A quite similar conception of truth has been later developed by Crispin Wright - without, however, mentioning Goodman. Wright suggests that the best way to capture the spirit of the pragmatist conception of truth is to conceive it globally as "superassertibility". According to him, a statement is superassertible just when it is enduringly justified or, a bit more precisely: a statement is superassertible if and only if it is justified by some accessible state of information and will remain justified no matter how that state of information is enlarged upon or otherwise improved (cf. Wright 1992, 44-57).

Goodman's and Wright's proposals can, however, be refuted by considering cases of the following kind, which figure prominently in current discussions about indefeasibility theories of knowledge (cf. Skorupski 1988). Suppose that, under standard conditions of perception, I see Mary, whom I know pretty well, and that, on this stable evidential basis, I am completely warranted in asserting that it is Mary whom I see. Imagine now, however, that Mary's mother tells me that Mary is not in town today, but that her qualitatively identical twin sister Anna, whom she herself often confuses with Mary, is in town. This new information defeats my warrant. Suppose further that, quite unknown to me, what Mary's

mother said was quite wrong – perhaps because she was in a state of bewilderment. By virtue of the increment to my original state of information, the statement that I see Mary is no longer superassertible. But it is true nonetheless. So a true statement need not be superassertible or permanently credible. Hence we are well advised not to scrap the intuitive concept of truth.

In his book *Of Mind and Other Matters*, which appeared in 1984, Goodman himself has abandoned the view that permanent credibility or ultimate acceptability is a necessary condition for truth. Those passages in the texts reprinted here which seemed to point in this direction were appropriately revised. Unfortunately, he does not even mention his change of mind, and hence he also remains silent about the reasons for it. Instead, he now proclaims: "Ultimate acceptability – acceptability that is not subsequently lost – serves as a sufficient condition for truth …" (*MM*, 38; cf. also 40), and he accentuates that he does not propose to define truth as ultimate acceptability (ibid.).

In the meantime Goodman himself seems to have come upon some of the serious problems that an epistemic analysis of truth is confronted with. Nonetheless, he still adheres to the assertion that permanent credibility is at least a sufficient condition for truth. We will come back to this point soon.

V

First, however, I want to ask the question of whether there are not other, perhaps more promising theoretical approaches to understanding the concept of truth in epistemic terms.

In my opinion, the conception of truth which Putnam advocated during his internal-realist phase, the conception of truth as justifiability under epistemically ideal conditions, i.e., under conditions in which all relevant facts are available, has been the most interesting, fruitful, and informative attempt thus far to definitively bridge the logical gap between truth and justification. Putnam characterizes his internalist concept of truth in the following way:

> "Truth", in an internalist view, is some sort of (idealized) rational acceptability – some sort of ideal coherence of our beliefs with each other and with our experiences as those experiences are themselves represented in our belief system – and not correspondence

with mind-independent or discourse-independent "states of affairs" (cf. Putnam 1981, 49-50).

Putnam's theory, which equates truth with an idealization of rational acceptability or, using John Dewey's term, of "warranted assertibility", incorporates essential elements of both coherence theories and pragmatist theories of truth. Moreover, Putnam likes to lay emphasis on the affinities to Kant's equally epistemically constrained conception of truth (cf. Putnam 1981, 60-64). Yet, he attributes the main idea behind his new view to the antirealist philosophy of Michael Dummett. For here he felt that he had found a coherent alternative to both the correspondence theory and the redundancy theory of truth (cf. Putnam 1978, 127-130; 1983, XVI, 81-86). Expressed somewhat simplistically, Dummett's view is that the meaning of a sentence is not provided by its realist, recognition-transcendent truth conditions, but, instead, by its verification conditions (cf. Dummett 1975 and 1976).

Even though he ascribed his rejection of the correspondence theory to Dummett's increasing influence on his position, Putnam immediately began to emphasize that there were essential differences between his own view and that of Dummett. For the most part, these differences of opinion are related to Putnam's much greater concern with the differences between mathematical discourse – an important source of inspiration for Dummett's semantic approach – and empirical discourse. Consequently, Putnam was much more careful than Dummett in applying verificationist semantics to empirical discourse.

Dummett equates truth with justification tout court. Whereas he originally understood "justification" largely in the restrictive sense of conclusive justification, he later began to use the term in a weaker sense, meaning a justification of the strongest available kind (cf. Schantz 1996, 240-46). Putnam is very skeptical about the idea that statements about external material objects can be conclusively justified (cf. Putnam 1983, XVI-XVIII, 22). A mathematical statement is either proven or not. But in areas beyond mathematics, in the realm of contingency, we are hardly in a position to fully exclude the possibility of error. Putnam reminds us that in the empirical sciences, verification is always a matter of degree. In stark contrast to Dummett, Putnam is a staunch epistemological holist. There is such a thing as a formal proof of an isolated mathematical theorem. But – and Putnam

insists on this point no less than Willard Van Quine – the unity of verification in the empirical sciences is the entire theory and not an isolated statement. The sentences of a theory are not separately vulnerable to adverse observations, because it is only jointly as a theory that such sentences imply their observable consequences.

Putnam's main objection, however, aims at Dummett's suggestion of simply equating truth with justification (cf. Putnam 1981, 55; Putnam 1983, 84-6). As pointed out above, truth is stable, whereas justification is tensed.

VI

I think that Putnam actually succeeded in eliminating various flaws of earlier, much more radical approaches. It is also important to note that Putnam explicitly dissociates himself from Peirce's suggestion that there is such a thing as an epistemic situation which is simultaneously ideal for the justification of every true statement. Rather, Putnam insists that the question of what are better or worse epistemic situations varies from statement to statement (cf. Putnam 1990, VIII). So the notion of ideal epistemic conditions must be specialized to the topic at issue. Nevertheless, I will now demonstrate that his own epistemic definition is also not able to stand up to a critical examination.

First of all, strikingly enough, Putnam makes no serious attempt to define the notion of epistemically ideal conditions more precisely. Certainly, this is no easy task, for specifying the details of such conditions would require stating one's position on fundamental epistemological problems. Thus, it can be expected that the distinct views of internalists and externalists on the one hand and of fundamentalists and coherentists on the other result in varying criteria of what an epistemically ideal situation actually is. In any case, Putnam clearly interprets an epistemically ideal situation in an anthropocentric way – as a situation that is ideal for *us*, for finite entities with our perceptual and cognitive faculties. Other possible creatures that may have cognitive abilities far beyond our own rightly remain overlooked. Of course, the reason why epistemic definitions of truth are motivated by an anthropocentric position is obvious: they are to demonstrate that truth is accessible to *us,* so that, in principle, truth cannot transcend the evidence that *we* can attain for it. Within this context, it is apparent that any omniscient

creature is of no avail. Without any qualification whatsoever, realism can readily acknowledge that for a statement to be true it is a necessary condition that an omniscient creature would accept the statement.

But it is precisely this restriction to cognitive human subjects that makes epistemic analyses vulnerable to the objection that there could be true statements that are not justifiable for us in an epistemically ideal situation. Even in such a situation, they might be unjustifiable for us because the reasons that tell for their truth might remain, in principle, inaccessible to us due to our sensory and cognitive limitations. And, in an analogous way, false statements might be justified for us in an epistemically ideal situation because the reasons that tell against their truth might be, in principle, inaccessible to us. It becomes evident that justifiability under epistemically ideal conditions is neither a necessary nor a sufficient condition for truth. Hence we also see that, and why, Goodman's new, weakened claim that permanent credibility is at least a sufficient condition for truth cannot be sustained.

In light of the entire body of scientific knowledge pertaining to our sensory systems and cognitive mechanisms and to how they acquire and process information about the external world, it is a highly plausible assumption that there are aspects of reality that will forever remain unapproachable to us, even if perhaps not to all possible cognitive subjects. If we are actually aware of our sensory and cognitive limitations, if we have good reason to believe that man is not the measure of all things, then why shouldn't we acknowledge that there could be facts that, in principle, exceed human cognitive faculties? Certainly, this is a sound argument against equating truth and justification under conditions that are epistemically ideal for us.

VII

My second objection to the epistemic definition of truth proposed by Putnam, the objection I wish to focus on here, is that the definition is circular since its main concept of an epistemically ideal situation can ultimately only be adequately defined with reference to the concept of truth. The truth norm underlies the norms of justification in all their ramifications and refinements. This fact becomes apparent if we center our attention on the concept of

epistemic justification, and especially if we inquire into the relationship between this concept and the concept of truth. It is indeed possible that some belief is epistemically justified and at the same time actually false. For this reason, truth cannot simply be identified with justification. Thus, the relationship between justification and truth cannot be the relationship of logical implication.

Nonetheless, I do think that there is a conceptual connection between epistemic justification and truth. In agreement with numerous other epistemologists, such as William Alston (1989), Alvin Goldman (1986), Laurence BonJour (1985), and Paul Moser (1989), I insist that, in a certain sense, it is constitutive of justification that it leads to truth, that it counts toward truth. In the final analysis, the essential goal of cognition is to attain true beliefs and avoid false ones, to maximize truth and minimize falsehood. We want our beliefs to represent the world as it really is. The fundamental role of epistemic justification is that of a subjective means to this objective, cognitive goal. Justification helps us to approximate to truth and hence is essentially directed toward it. Accordingly, any belief is epistemically justified for us if and only if we have good reason to believe that it is true. Accepting some belief without good reason would mean disregarding the cognitive goal of truth.

There are varying approaches to explaining more precisely this conceptual connection between epistemic justification and truth. Yet, they all share the view that it is an essential element of the concept of epistemic justification that it is "truth-conducive". This means that the conditions of justification must be such that satisfying them can guarantee that it is extremely probable that a particular belief is true. If attaining justified beliefs did not considerably increase the probability of attaining true beliefs, justification would be insignificant for attaining our primary cognitive goal. Why should we then be concerned about whether our beliefs are justified or not? On the other hand, a conception of justification that includes a connection to truth makes it obvious why justification is considered a desideratum in the search for truth.

Of course, the relevant literature also contains manifold attempts – especially from the internalist faction – to explain the concept of justification without using the concept of truth. In such views, the justification of a belief does not consist in the objective probability of its being true, but, on the

contrary, in how things present themselves to the specific perspective of the human subject. The internalists insist that a belief is justified for some person only if she is consciously aware of the reasons that support the belief or if the reasons are at least cognitively accessible to her. Roderick Chisholm, for example, states the following:

> The internalist assumes that, merely by reflecting upon his own conscious state, he can formulate a set of epistemic principles that will enable him to find out, with respect to any possible belief he has, whether he is justified in having that belief. […] In a word, one need consider only one's own state of mind (Chisholm 1989, 76).

Those who wish to make justification a purely internal matter often advocate a deontological conception of justification according to which a belief is justified if a person having this belief satisfies her intellectual duties or does not violate these. Then, a justified belief is that which is permissible according to the relevant intellectual norms. Yet, a closer look reveals that most advocates of a deontological version of internalism do not go so far as to completely disrupt the connection between justification and truth. For they usually assign a prominent position among our intellectual duties to the obligation of believing what is true and disbelieving what is false. Thus, Chisholm writes:

> We may assume that every person is subject to a purely intellectual requirement – that of trying his best to bring about that, for every proposition h that he considers, he accepts h if and only if h is true (Chisholm 1977, 14).

Although Chisholm does not advocate the idea of truth conducivity, just like other internalists he also takes recourse to truth in order to make the concept of justification comprehensible. He uses the concept of truth to describe the content of our most important intellectual obligations. According to Chisholm, our beliefs are justified if and only if we have exerted the greatest effort to accept them only if they are true. So the theory of epistemic justification does not seem to be able to work without some recourse to the concept of truth.

Of course, there are other scholars, such as Carl Ginet (1975) or John Pollock (1974), who have attempted to characterize epistemic rationality without referring to the concept of truth in any way at all. I do not think that such attempts could have much chance of success. Without taking recourse to the concept of truth they can neither succeed in differentiating epistemic

justification from other forms of justification, be these moral, legal or pragmatic, nor in determining adequate criteria that would allow for a rational choice between competing rules of justification. But this is not a topic that need be explored in detail here. The crucial point is that those who advocate equating truth and justification under epistemically ideal conditions cannot find support in such attempts to sever the conceptual connection between justification and truth. They cannot find any support here because their entire argument that truth is nothing other than epistemically ideal justification fails if the concept of justification does not involve truth conducivity. For if the concept of justification does not even include a guarantee of the probability that a statement is true, it is obvious that a belief, regardless of how ideally justified it may be, can nonetheless be false. To put it briefly: if the conceptual connection between justification and truth is denied, identifying truth and justification in an epistemically ideal situation becomes totally implausible. But if this close connection is indeed acknowledged then the identification proposed becomes circular.

VIII

During the last stage of his internal realism, Putnam himself began to admit that an explication of the concept of an epistemically ideal situation is dependent on the concept of truth (cf. Putnam 1988, 115). Nonetheless, he did not believe that this would lead him to a vicious circle. At this point, he asserted that what he meant by stating that truth was idealized justification was not to be interpreted as if the concept of idealized justification were more fundamental than the concept of truth. Putnam now assured us that he had never aimed at a reduction of truth to epistemic concepts. What he really meant is that the concepts of truth and of idealized justification are interdependent concepts. He readily admitted that his error lay in having hitherto emphasized only one side of this interdependence, truth's dependence on justification, and having neglected the other side, justification's dependence on truth. He now wanted to make up for this deficit by stating that determining whether an epistemic situation is ideal depends on whether many diversified statements are true in the situation.

What can we make of Putnam's more recent assertion that truth and idealized epistemic justification are interdependent concepts? I do not wish

to deny that the acquisition of the concepts of justification and of truth involves mutual dependence. Yet, I do think that there is a significant asymmetry between these two concepts, an asymmetry that cannot adequately be expressed by the handy phrase of "conceptual interdependence". For although an adequate explication of justification is dependent on the concept of truth, an adequate explication of truth does not at all require epistemic concepts. The fact that *p* is both a necessary and a sufficient condition for *p* to be true. There is simply no place for an epistemically necessary or sufficient condition of truth. That *p* is true requires no more than the fact that *p*, and nothing less than this fact will be sufficient.

References

Alston, William (1989): *Epistemic Justification: Essays in Theory of Knowledge*, Ithaca: Cornell University Press, 1989.

BonJour, Laurence (1985): *The Structure of Empirical Knowledge*, Cambridge/MA: Harvard University Press, 1985.

Carnap, Rudolf (1936): "Wahrheit und Bewährung", in: Skirbekk, Gunnar (ed.): *Wahrheitstheorien*, Frankfurt: Suhrkamp,1977, 89-95.

Chisholm, Roderick (1977): *Theory of Knowledge*, 2d ed., Englewood Cliffs, N.J.: Prentice-Hall, 1977.

Chisholm, Roderick (1989): *Theory of Knowledge*, 3d ed., Englewood Cliffs, N.J.: Prentice-Hall, 1989.

Davidson, Donald (1986): *"A Coherence Theory of Truth and Knowledge"*, in: LePore, Ernest (ed.): *Truth and Interpretation. Perspectives on the Philosophy of Donald Davidson*, Oxford: Basil Blackwell, 1986, 307-319.

Davidson, Donald (1990): "The Structure and Content of Truth", in: *Journal of Philosophy* 87 (1990), 279-328.

Dummett, Michael (1975): "What is a Theory of Meaning? (I)", in: Guttenplan, Samuel (ed.): *Mind and Language*, London: Oxford University Press, 1975, 97-138.

Dummett, Michael (1976): "What is a Theory of Meaning? (II)", in: Evans, Gareth; McDowell, John (eds): *Truth and Meaning*, Oxford: Oxford University Press, 67-137.

Frege, Gottlob (1919): Der Gedanke. Eine logische Untersuchung, *Beiträge zur Philosophie des deutschen Idealismus* 1, 58-77, transl. A. M. and A. Quinton as "The Thought: A Logical Inquiry", in: *Mind* 65 (1956), 289-311.

Ginet, Carl (1975): *Knowledge, Perception, and Memory*, Dordrecht: Reidel, 1975.

Goldman, Alvin (1986): *Epistemology and Cognition*, Cambridge/MA: Harvard University Press, 1986.

Goodman, Nelson (1951), The Structure of Appearance, Dordrecht: Reidel, 1951.

Goodman, Nelson (1978), Ways of Worldmaking, Indianapolis: Hackett, 1978.

Goodman, Nelson (1980): "On Starmaking", in: *Synthese* 45 (1980), 211-215.

Goodman, Nelson (1982): "Notes on the Well-Made World", in: Leinfellner et al. (eds): *Language and Ontology. Proceedings of the 6th International Wittgenstein Symposion*, Wien: Hölder-Pichler-Tempsky, 1982, 31-38.

Goodman, Nelson (1984), Of Mind and Other Matters, Cambridge/MA: Harvard Universtity Press, 1984.

Horwich, Paul (1990): *Truth*, Oxford: Basil Blackwell, 1990.

Künne, Wolfgang (1993): "Truth, Rightness, and Permanent Acceptability", in: *Synthese* 95 (1993), 107-117.

Moser, Paul (1989): *Knowledge and Evidence*, Cambridge: Cambridge University Press, 1989.

Peirce, Charles Sanders (1958): *Collected Papers*, Cambridge/MA: Harvard University Press, 1958.

Pollock, John (1974): *Knowledge and Justification*, Princeton: Princeton University Press, 1974.

Putnam, Hilary (1978): *Meaning and the Moral Sciences*, London: Routledge & Kegan Paul, 1978.

Putnam, Hilary (1981): *Reason, Truth and History*, Cambridge: Cambridge University Press, 1981.

Putnam, Hilary (1983): *Realism and Reason: Philosophical Papers*, Vol. 3, Cambridge: Cambridge University Press, 1983.

Putnam, Hilary (1988): *Representation and Reality*, Cambridge/MA: MIT Press, 1988.

Putnam, Hilary (1990): *Realism with a Human Face*, Cambridge/MA: Harvard University Press, 1990.

Quine, Willard Van (1990): *Pursuit of Truth*, Cambridge/MA: Harvard University Press, 1990.

Russell, Bertrand (1912): *The Problems of Philosophy*, Oxford: Oxford University Press, 1912.

Schantz, Richard (1996): *Wahrheit, Referenz und Realismus*, Berlin & New York: de Gruyter, 1996.

Schantz, Richard (1999): "The Role of Sensory Experience in Epistemic Justification: A Problem for Coherentism", in: *Erkennnis* 50 (1999), 177-191.

Schantz, Richard (2001a): "Truth and Reference", in: *Synthese* 126 (2001), 261-281.

Schantz, Richard (2001b): "The Given Regained, Reflections on the Sensuous Content of Experience", in: *Philosophy and Phenomenological Research* 61 (2001), 167-180.

Schantz, Richard (2002) (ed.): *What is Truth?* Berlin & New York: de Gruyter, 2002.

Scheffler, Israel (1980): "The Wonderful Worlds of Goodman", in: *Synthese* 45 (1980), 201-209.

Skorubski, John (1988): "Critical Study [of Crispin Wright *Realism, Meaning, and Truth*, 1st edn.]", in: *Philosophical Quarterly* 38 (1988), 500-525.

Strawson, Peter (1950): "Truth", in: *Aristotelian Society Supplementary Volume* 24 (1950), 129-156.

Williams, Michael (1977): *Groundless Belief,* New Haven: Yale University Press, 1977.

Wittgenstein, Ludwig (1922), *Tractatus Logico-philosophicus*, London: Routledge, 1922.

Wright, Crispin (1992): *Truth and Objectivity*, Cambridge/MA: Harvard University Press, 1992.

Thomas Splett

How Much of a Relativist Is Goodman?

Abstract
In this paper I search for an understanding of what relativism consists in, especially Goodman's version of it. In three attempts I try to make sense of the concept of truth relativism. Each attempt leads to a distinctive kind of relativism: meaning pluralism, indexicalism, and slippery relativism. All three kinds are evaluated with regard to their respective potential to account for widespread relativistic intuitions. Based on these considerations, the adequacy of labelling Goodman a relativist is pondered.

1. Introduction

When thinking about Goodman's position, certain labels such as constructivism, pluralism, nominalism and relativism come to mind almost inevitably. What is Goodman's attitude towards such labels? On the one hand, he has a relaxed attitude towards them. For example, he advocates not making a big deal about the difference between monism and pluralism of worlds (*WW*, 2). Then again, he insists that his claim of the existence of competing truths and a plurality of real worlds should not be regarded as merely rhetorical (*WW*, 110). My following discussion is based on the assumption that exploring a label might help to improve the understanding of the position it is used to label. In this paper I want to clarify the meaning of at least one label - that of relativism. In philosophy and cultural sciences, we find a varying usage for this label: expressing a rather practical attitude, referring to vague intuitions, or even naming elaborate theorems, for instance concerning the interpretation of the truth predicate.

Goodman's usage of this label, even connected with constructivist ideas, mainly concerns the theory of truth. Sometimes “radical relativism” means to him “only that truth must be otherwise conceived than as correspondence with a ready-made word” (*WW*, 94). Sometimes his understanding of relativism is more bold, namely when revising the dictum “‘Snow is white’ is true if and only if snow is white” to something like “‘Snow is white’ is true in a given world if and only if snow is white in

that world" (*WW*, 120).

Relativism generally concerns validity. The core idea is that certain norms, ideals, values, or the truth of statements are considered valid only with reference to certain individuals, groups, cultures, theories, languages, discourses, systems etc. Following Goodman I will speak mainly of truth relativism. This is not just an arbitrary restriction. As I will point out later, those aspects of relativism which are actually provoking in a philosophical sense concern truth relativism. Truth relativism is not simply one relativism among others, as, for instance, the relativism of beauty or taste, but rather the culmination of relativism. Consider the relativism of beauty. It states: "Beauty is relative (to individuals, cultures ...)." One way of putting this is to say "The truth of statements about beauty is relative (to individuals, cultures ...)." This way, a relativism regarding beauty is transformed into a relativism regarding the truth of statements about beauty.[1] To put it another way: since truth is the truth of statements about things other than truth, there is no truth relativism without relativism regarding certain substantial objects which a given discourse is about. In this paper, I shall discuss two examples: first, ideals of beauty, and second, the following:

The ancient Greeks said: *"There are gods dwelling atop of Mount Olympus."*
Modern people say: *"There are no gods dwelling atop of Mount Olympus."*

I shall examine in what way a relativist would deal with these statements. I am not asking whether relativism is the adequate way to deal with these statements. An easy, non-relativist way to handle these statements would be to say: they contradict one another and the statement of the ancient Greeks is false, whereas the modern one is true. The relativist cannot follow this path. He must grant the ancients more than merely having the belief that gods are dwelling atop of Mount Olympus.

[1] One condition for this transformation is cognitivism regarding statements about beauty. But even on that condition, one could speak of beauty as being relative without claiming that the truth of statements about beauty is relative. I will comment on this idea later.

He must concede that they are in some sense correct in believing so. The relativist I have in mind is not someone totally indifferent, simply disinterested in the question of gods and Mount Olympus. My relativist is not an opportunist. Rather, he is someone who claims there are no gods dwelling atop of Mount Olympus. But at the same time he responds to the ancient statement in a different way than an absolutist or objectivist, who would simply regard the statement as false.[2] I am not asking whether this is a good way to respond as I am not arguing against or in favour of relativism in the first place. What I am interested in is what the point of relativism is. I will distinguish three different ways a relativist could interpret the above statements, from which three kinds of relativism derive. Subsequently, we will have enough at hand to determine in how far it is adequate to label Goodman's position relativistic.

2. Pluralism of meaning

The simplest way to interpret the seemingly conflicting statements without judging one to be true and the other to be false is to take them to express different meanings. Our modern relativist could take the ancient speaker to use the expressions “gods” or “are dwelling atop of Mount Olympus” with improper meaning (displayed by the quotation marks).

Ancient: *“There are gods dwelling atop of Mount Olympus.”*
Modern: *“There are no gods dwelling atop of Mount Olympus.”*

Interpreted this way, the modern relativist has no problem to judge both statements as true. Having spoken of gods dwelling atop of Mount Olympus, the ancient Greeks were not making a statement about observable characters atop of a physical mountain. Rather, their statement must be understood in a “metaphorical” way: for them, the gods have a symbolic presence. They dwell atop of Mount Olympus in a way comparable to what a modern Christian has in mind when saying that God exists in heaven, but not in the blue sky above us, where he could be hit by

[2] At this point, I am not concerned with the subject of truth over time. One could imagine a community living near Mount Olympus nowadays which believes there are gods atop of the mountain.

aeroplanes or spaceships.

So there is no disagreement. What is claimed by one statement and what is denied by the other are different and independent issues. This is an easy way to render both statements compatible. But is it really the relativist's way? Let's have a look at the following statement:

> *There are no unicorns, but there are unicorns for the director of the movie "The Last Unicorn" - at least one unicorn, which is the movie's protagonist.*

The above statement combines the following two statements:

Director: *"There is a unicorn (in my movie)."*
We: *"There are no unicorns."*

Reconciliation works here in the same way it did in the case of gods and Mount Olympus. But it does not seem plausible to call this an example of relativism. There are two different meanings of "there being a unicorn", but no relativity. Specification is not relativization.

Of course, if somebody still wants to call this a way of treating the statements in a relativistic manner, it will be of little use to argue about the mere term. But then it is important to distinguish more significant relativistic positions from this pluralism of meaning.

Before I move on to do just that, I want to indicate a variant of meaning pluralism, namely when it concerns the term "true". Up to now, we proceeded without using the truth predicate. There were simply statements as "There are gods dwelling atop of Mount Olympus." But the Greeks could have also formulated "It is true that gods are dwelling atop of Mount Olympus." And the modern speaker: "It is not true that there are gods dwelling atop of Mount Olympus." Both statements are compatible either when the phrase "gods are dwelling atop of Mount Olympus" is used with different meanings (as considered above) or when the term "true" is used with different meanings, which is characteristic for the variant. A modern relativist could accept the ancient statement, if it was to express only a metaphorical truth. Whereas it is true (or literally true or maybe physically true) that there are no gods dwelling atop of Mount Olympus, it is

metaphorically true or symbolically true that there are gods dwelling atop of Mount Olympus.

Again, I am not debating whether this would be a good way to deal with the ancient statement. I only want to clarify whether it is an interpretation which is typical for a relativistic position. Again, I would deny this. The fact that the term "true" is being used in different ways here does not mean there is significant relativism. So why do I mention the variant? Because it indicates a way to implant relativity: by relating truth to speakers, cultures, or - using Goodman's phrasing - to worlds or versions of worlds. So let us now have a closer look at:

3. Indexicalism

Two steps lead from the meaning pluralism described above to a position that is best understood as indexicalism or indexical relativism. One step is to replace specification with relativization. The other is to proceed from looking at bare propositions like gods-dwelling-atop-of-Mount-Olympus to statements about the truth of these propositions. What the ancient Greeks and we today actually express, sounds to the indexicalist relativist as follows:

Ancient Greek: *"It is true for us, that there are gods dwelling atop of Mount Olympus."*
Modern people: *"It is not true for us, that there are gods dwelling atop of Mount Olympus."*

Even if both occurrences of "that there are gods dwelling atop of Mount Olympus" have identical meanings, both statements are compatible because of the relativization of truth. The ancient Greeks spoke of what was true for them, modern people speak of what is true for them. Both express different matters which are independent of one another. There is no point of direct contradiction.

To better understand what happens due to relativization, I want to point out that both steps, the one of relativizing and the other of proceeding to the truth predicate, are not necessarily connected. One can consider the is-

true statements without relativization. (Remember the variant of meaning pluralism.) And one can relativize the proposition instead of the truth talk. In that case we get something like:

Ancient Greeks: *"For us, there are gods dwelling atop of Mount Olympus."*
Modern people: *"For us, there are no gods dwelling atop of Mount Olympus."*

One can use the „for us"-phrase to refer to different parts of the proposition, for instance to "the gods" or "dwelling". What we considered as difference in meanings of "gods" before, may alternatively be conceived as a case of relativity under certain circumstances. There may be only one meaning of the term "gods", but different classes of objects that it refers to, depending on the speaker's location in cultural history. The structure is similar to saying "you" standing in front of different persons, once in front of Paul and another time in front of Paula. The difference is that the ancient way of speaking about gods does not differ from the modern way with regard to the gods it is used to refer to, say Zeus or Hera, but something like "gods of art and myth" and "real gods" respectively.

Considering another example, namely the claim that beauty is relative, may clarify this point. Let us distinguish (at least) two kinds of beauty of the human body, heroin chic beauty and Rubens beauty. Now imagine a model agency sorting its models according to their kind of beauty. There is one folder for heroin chic models and another for Rubens type models. We could then simply state: There are different kinds of models, just as there are different kinds of beauties. Kate Moss is heroin-chic-beautiful, but not Rubens-like-beautiful. Construed that way, we are again dealing with a matter of specificity, not of relativity. To let the idea of relativity come into play, one first has to procede from talking about two kinds of beauty to talking about two standards of beauty, namely "beautiful with regard to the standard of heroin chic" and "beautiful with regard to the Rubens-standard". Understood this way, they can be transformed to an indexical formula like "beautiful relative to the particular standard". Unlike talking of "gods" or "dwelling" in a relativistic manner, this talk of relative beauty sounds more familiar. Because of the apparent implausibility of

relativizing the phrase “gods dwelling atop of Mount Olympus” the relativist chooses the truth-predicate for relativization. Given that a relativization has to take place at all, it should take place there.[3] Involving the truth predicate is the second step in proceeding from meaning pluralism to indexicalism.

When is it appropriate to relativize truth and when is it inappropriate? Why could one consider being a relativist concerning beauty (in an interesting sense), but not to be a relativist concerning left and right? Compare the following statements:

(a) *Kate is beautiful for us, but nor for Rubens.*
(a1) *It is true for us, that Kate is beautiful, but not for Rubens.*
(b) *The chess clock is on the right from Karpow's point of view, but not from Kasparow's point of view.*
(b1) *It is true for Karpow that the chess clock is on the right, but it is wrong for Kasparow.*

Whether one accepts (a1) as a proper statement or not, one has to understand why one might be more inclined to accept (a1) than to accept (b1). Why is it more plausible to express the meaning of (a) by asserting (a1) than to express the meaning of (b) by asserting (b1)?[4]

If we want to understand the idea underlying relativism we will need criteria to distinguish cases of interesting relativism from cases of simple relativity or indexicality of the left/right-kind. Consider the following suggestions for such criteria:

---

[3] What about the case of beauty? Here it seems that the transition is not that necessary. For the sake of the argument, let us assume beauty to be relative without assuming statements about beauty to be truth relative. This will not be a relativism of validity and therefore a philosophically less interesting kind of relativism. Relativism of beauty understood that way may be more interesting for a theory of beauty than for the theory of relativism. Speaking of “left” or “right” is also relative (left to what?), but it would be beside the point to speak of a left relativism.

[4] We all understand (b) very well. If someone utters (b1) we quickly interpret his utterance as a strange way to express (b). That means we do not accept (b1) as a correct transformation of (b), let alone as an explication of (b).

A. The compatibility of the statements considered must not be too obvious. There must be at least a strong semblance of conflict. There will only be relativism if relativity is not too self-evident, when not all people agree to the compatibility of the statements in question immediately.

Compare this to labels such as constructivism. There is not much sense in speaking of social constructivism with regard to entities like paper money. Of course, the value of paper money is a result of social valuation. There is no constructivism in a significant sense unless it concerns entities like molecules, sex/gender or anything else that is subject to disagreement about the extent of social categorization and construction. Similarly, speaking of relativism only makes sense regarding entities which do not already seem to imply relativity.

Recently there was a discussion under the title of "(faultless) disagreement" (see Kölbel 2003; MacFarlane 2007), whether one even has to admit some kind of real conflict between the relevant statements in order to be a relativist. Goodman often speaks of "conflicting" statements or versions (*WW*, 110ff). For him they actually conflict and not only seem to do so.

B1. The compatibility is only abstract in the sense that there is no straight picture, no single description, and no view available on the subject the statements are about, that integrates both statements without missing important aspects. Consider both statements as expressing different perspectives on the same subject. This talk of perspectives is heavyweight in case it means there are only specific perspectives available, not a view on the subject as it is. The right/left case is excluded, since there is a way to capture the spatial situation in a single description like: "Karpow and Kasparow are sitting vis-à-vis and the chess clock is near the right hand of Karpow." It is also possible to make a bird's eye view drawing of the scene. There might be a slight loss when integrating both statements in this manner.[5] But in comparison to the case of gods dwelling or not dwelling atop of Mount Olympus, the loss is way less extensive.

[5] As Goodman considers with regard to descriptions of relative movements (cf. *WW*, 113f).

B2. Understanding of and empathy with the conflicting position is difficult. There have to be hurdles which make it difficult to switch to the other perspective. In the right/left-case there is no such difficulty. Kasparow can easily imagine himself in Karpow's position in space. In the case of dwelling or not dwelling gods it takes much more effort for the modern man to put himself in the position of the ancient one (not to mention vice versa). We could guess what it means for the ancient Greeks to be of the opinion that there are gods dwelling atop of Mount Olympus only if we knew a lot about their culture. Maybe so much that it is practically impossible for us to understand their view perfectly without sharing their form of life; that is without becoming ancient Greeks ourselves.

Regarding different ideals of beauty we may now distinguish more clearly between a less and a more relativistic way of understanding this difference. Presenting the standards more as two kinds of beauty does not imply any sort of significant relativism. For the model agency it is only a task of attributing a certain kind of beauty to a person. It comes to relativism when different people apply different standards of beauty without being able to really comprehend someone else's differing standard. Maybe an admirer of today’s skinny models cannot understand what Rubens liked about plump women. Then it would be more appropriate to speak of relativism. (Of course only under the condition that he concedes statements made by admirers of Rubens beauty as somehow valid.)

These (connected) points present aspects of relativism and make the indexicalist’s idea more sound. If the conditions are met (to a certain degree), then there is not only relativity but relativism (to a certain degree).

But are we really at the end of defining relativism? I don’t think so. Within the framework of indexicalism, the suggested conditions will not be met to the highest degree. There are features of indexicalism which are also pointed out in the conditions that match intuitions often connected with relativism only to a comparatively small extent. Therefore one has to think about stronger requirements for relativism.

Let's have a look at condition A, the one regarding the role of conflict. This is a crucial point for the relativist. When conceding too much conflict between the statements, there is no relativism, but absolutism or objectivism. If a modern speaker who is convinced that there are no gods dwelling atop of Mount Olympus judges the ancient Greek's statement to

be simply false, then he is not a relativist, of course. But on the other hand: if there is no conflict, no disagreement at all, then there is no need to speak of relativism. The indexicalist grants apparent conflict. So the first condition is met, at least to a certain degree. But an apparent conflict is no real conflict, no real disagreement, since disagreement only exists when there are different opinions with regard to the same proposition. It could be wise to demand more than only the appearance of conflict. At least it seems possible to find another kind of relativism that will allow for stronger conflict.

Not only the matter of conflict but also the questions of non-integrability (B1) and non-comprehensibility (B2) may be understood in a stronger sense, as I will point out later. Within indexicalism they are only addressed in a weak sense: as merely practical matters. One lacks information about the background beliefs of another which would help him better understand his counterpart's position. But a relativist may reach out for a kind of non-integrability and non-comprehensibility which goes deeper and is closer to incommensurability. The indexicalist may be better described as a perspectivist who doesn't speak directly of a certain thing he envisages, but of his own view of the thing. He does not speak of what is true, but what is true for him. Furthermore, does such perspectivism not imply objectivity? In talking about different perspectives, one seems to accept that there is (not one topic opposed statements are about but) one and the same thing or situation which is viewed from these perspectives. Even if there is not a single straight description that integrates the different relative statements, the implication of objectivity remains - which seems to contradict strong relativist intuitions. Is the same true for indexicalism? At least the contrary has not been proven so far. The phrasing of a true-for-statement does not rule out the possibility to accept an objective entity, it does not even rule out the possibility that there is one true description in a not relative sense of truth. There could be one true description, a very long and complex one. That is not ruled out by condition B1 above. But maybe a strong relativist would want to rule it out. (With respect to condition B2 the argument would be roughly the same.)

Furthermore, indexicalism has one essential feature that could be judged to be of (too) high cost, namely that it talks of “being true for someone” instead of talking about “being true”. This is not merely a new way of

speaking. For instance, it affects the disquotational scheme of truth, which in consequence cannot be upheld in its hitherto existing form. Goodman claims the dictum "'snow is white' is true if and only if snow is white" should be revised to something like "'snow is white' is true in a given world if and only if snow is white in this world" (*WW*, 120). If this were true, we would have to change not only our usage of the "is true"-expression, but all moves in the game of assertion. Saying "snow is white" would be elliptical, short for „snow is white for us/in our world". Equally, "there are no gods dwelling atop of Mount Olympus" would be short for "there are no gods dwelling atop of Mount Olympus for us/in our world". As we have already seen, such relativization makes it difficult to maintain the idea of disagreement between competing opinions with regard to the same thing or affair. But by abandoning the standard disquotational scheme of truth delicate effects on our philosophical as well as on our ordinary treatment of truth are to be expected.

4. Slippery relativism

When searching for an alternative conception of relativism[6], three conditions should be met:

1. Conflict and disagreement must be considered in a stronger sense than the idea of merely apparent conflict implies.
2. Non-integrability in a single description or picture and non-comprehensibility of one's position with another's position should be obtained more in principle.
3. There should not be too much revision concerning the way we make statements and use the truth predicate (if possible).

Is there a relativistic interpretation of the ancient and modern talk of gods dwelling or not dwelling atop of Mount Olympus? The proposal would be to let them express the following:

[6] Other approaches to develop stronger forms of relativism that match the one I present in important aspects are to be found in (Kölbel 2004) and (MacFarlane 2005). Both authors are of great help for an understanding of the recent debate and work out their ideas in many more argumentative details than I can do here.

Ancient Greeks: *"There are gods dwelling atop of Mount Olympus. And it is true that there are gods dwelling atop of Mount Olympus."*

Modern people: *"There are no gods dwelling atop of Mount Olympus. And it is true that there are no gods dwelling atop of Mount Olympus."*

According to the idea of relativism, they should (if they really understand what they are saying, which means: if they understand their statements relativistically) both be thinking of something like: … *whereas truth is relative to contexts of evaluation and conflicting statements from different contexts are not without warrant.*

This kind of relativism does not take the meaning of the judged proposition (gods-dwelling-atop-of-Mount-Olympus) to be relative to speakers or worlds. It is the judgement of the proposition with regard to its truth value itself which is considered relative to speakers or worlds. The statement with the "true"-expression has the same meaning as the statement without the expression. The relativist does not challenge this assumption. Of course, there is no truth relativism without relativity of truth. But taking "is true" to be elliptical and introducing the term "true for us/in our world", as standard, is not the only possible approach to establish relative truth. Slippery relativism, as I want to call this position, distinguishes the level of statements and the underlying pragmatics. Relativism is an explicitly theoretical way of understanding our game of making statements. The slippery relativist leaves our way of making statements largely untouched, but proposes to change our understanding of what we do when we make statements.

Assuming that ancient and modern people speak of largely different things is not relativism. Relativism means taking their statements to be on (roughly) one and the same topic but expressing different, conflicting opinions about that topic, and nevertheless accepting these different opinions in a stronger way than simply attributing belief. What do I mean by that? A relativist who is of the opinion that there are no gods dwelling atop of Mount Olympus himself needs to attribute more to the ancient Greeks than to have had the belief that there are gods dwelling atop of Mount Olympus. Then again, he must not attribute truth to their statement either, not in the same sense he thinks his own view of there not being gods atop of Mount Olympus is true. The modern relativist could say something

like: “It is true that there are no gods dwelling atop of Mount Olympus, and it is not true that there are gods, as the Greek said, but nevertheless their statement is not simply false. For them their view is as legitimate as mine is for me and for them their view is as correct as mine is for me.”

It is not easy to understand what exactly is meant by that. The relativist's statement does not seem very clear. In any case it sounds a bit slippery, and that is why I call this position slippery relativism.

Let us consider to what extent slippery relativism fulfills our conditions.

Condition 1 is fulfilled, since there is more conflict than the indexicalist implies. People are not only in apparent, but in real conflict: they express different opinions with regard to the truth value of (roughly) the same proposition. (That is what disagreement consists in.) Slippery relativism is a strong kind of relativism. Relativizing truth is to consider the proposition to be the same, not only the expression as in meaning pluralism, or an indexical formula as in indexicalism. (The more important the common basis for divergence, the stronger the kind of relativism.)

This is why condition 2 is fulfilled. Non-integrability and non-comprehensibility go deeper than they do within indexicalism. As it is the same proposition that is held true by one and false by another, the solution of speaking of different perspectives on the same subject is out of reach. I cannot imagine a picture that reconciles both the trueness and the falseness of one and the same proposition. There is a greater gap between the disagreeing parties. Someone cannot fully explain the statement of his counterpart which contradicts his own view. There always remains some lack of understanding. A purporter of slippery relativism accepts that as a fact.[7]

Does that lead to such dramatic changes in our practice of making statements that condition 3 is not fulfilled? I do not think so. Whereas the basics of our way of speaking are left untouched, there is some change on the level of the meta-theory on making statements and having opinions. Is it only an abstract theoretical affair or does it affect our practices? Relativists are to be expected to treat statements of others in a way which partly differs from the way non-relativists do. It is an interesting question

[7] For some people that would be a big disadvantage attached to slippery relativism. Maybe it is. But then it would be a point against relativism. It is not my aim at this point to prove that relativism is a better position than objectivism.

whether there would actually be important changes if we all became relativists. Imagine us talking that way: “Yes, you are right in your world as I am in mine, but for me it is my world, so I have to think the way I do and enforce my claims and not yours.” (Even if this were an idealized meta-view from which the others were as right in their respective worlds as I am in mine, it would not be possible for anyone to take this point of view.) Maybe there would be no big changes in the end: no changes in how we use the truth-predicate, and no drastic changes in our practices, but only new façons de parler used to comment on situations of disagreement?

But maybe there are heavy costs which come along with accepting slippery relativism. One has to adopt a complex, a double, in the end a slippery attitude. For example, stating that it is true that there are no gods dwelling atop of Mount Olympus, and at the same time accepting that truth is in some kind relative to contexts of evaluation. This may look self-contradictory, like some pragmatic inconsistency. But is there really a contradiction? To be honest: I cannot recognize one. Granted that there are high costs, aren't we paying the bill already? To me it seems as though we are accustomed to these complex attitudes in everyday life, but tend to simplify them when it comes to philosophical discussions about the status of statements.

5. Goodman

Having distinguished three types of relativism, how much of a relativist is Goodman? Like many other thinkers calling themselves “relativists”, Goodman mixes aspects of all three types.

A considerable amount of slippery-relativistic ideas can be found in Goodman's position, since he stresses issues which are considered best in this kind of relativism, like the significance of conflict and fundamental non-integrability. He points out that it is impossible to integrate different versions without loss and gives an ontological interpretation: If it is impossible to integrate versions into a single view, then there is no objective thing.

That sounds much like a strong and slippery relativism. But think of Goodman’s pluralism of worlds and versions. One reason Goodman has

for talking of different worlds is just to avoid conflict (cf. *WW*, 110). Opposed statements would conflict if settled in one and the same world. To avoid a real contradiction, they are assigned to different worlds. Emphasizing this idea of plurality, Goodman turns out to be more a meaning pluralist than a slippery relativist.

How serious do we want to take his talk of different worlds? When taking it too seriously, it becomes incomprehensible how different worlds, versions or even statements could be in conflict. This may lead one to take Goodman's talk of conflicting worlds (*WW*, 109ff) not that literally.

Last but not least, there are moments of indexicalism to be found in Goodman's account, first of all in his mentioned proposal to replace "'snow is white' is true if and only if snow is white" by "'snow is white' is true in a given world if and only if snow is white in that world".

For indexicalism, the propositions of the uttered statements depend on contexts of utterance, whereas for slippery relativism the attribution of the truth value to a proposition is dependent on contexts of evaluation. Goodman's talk of worlds or versions does not seem very interested in distinguishing these two types of context. If Goodman conceives of different worlds, i.e. of the world of the ancient Greeks and the world of modern people, as rather closed and locates the standard for judging the statements made by a group in that group's respective world - ancient matters are to be judged within the ancient framework - then he is a relativist more in the meaning pluralistic and indexicalistic sense. On the other hand, when he considers that there are bridges between the worlds which make it possible to reach out from one world into another, allowing us to take sides, to propose the modern view (in another way than the objectivist does) even on an ancient subject, then he is a relativist in the stronger and slippery sense.

Goodman's remarks on relativity, worlds and versions leave room for interpretation. The mentioned aspects of different kinds of relativism partly contradict each other. If we want to draw a consistent and unitary picture of Goodman's position, we have to weigh and choose which aspects to stress.

There is one last point: I discussed truth relativism. After all, for Goodman truth relativism does not have priority. To him, there is something wider, more fundamental and more interesting than truth:

rightness (cf. *WW*, ch. 7). With regard to rightness there seems to be no space for relativistic ideas in Goodman's position. Whole versions are right or wrong, absolutely right or wrong (where rightness is construed as ultimate acceptability; cf. *WW*, 139), not again relative to some version. So if we do label Goodman a relativist, we should keep in mind that this is not the whole story.

References

Kölbel, M. (2003): "Faultless Disagreement", in: *Proceedings of the Aristotelian Society* 104 (2003), 53-73.

Kölbel, M. (2004): "Indexical vs Genuine Relativism", in: *International Journal of Philosophical Studies* 12 (2004), 297-313.

MacFarlane, J. (2005): "Making Sense of Relative Truth", in: *Proceedings of the Aristotelian Society* 105 (2005), 321-339.

MacFarlane, J. (2007): "Relativism and Disagreement", in: *Philosophical Studies* 132 (2007), 17-31.

Siegel, H. (1987): *Relativism Refuted. A Critique of Contemporary Epistemological Relativism*, Dordrecht: Reidel, 1987.

Mark Textor

Exemplification and Idealisation[∀]

Abstract
I will discuss the role of the exemplarisation of properties in scientific reasoning, especially in scientific idealisation. After criticising the view that idealisations are fictional objects, I will argue that in scientific idealisation a real object (for example, a gas) is jointly pretended to fit a law.

1. Exemplification

An example from the history of philosophy can serve to introduce our topic. In Kant's philosophy of mathematics the symbolic construction of concepts plays an important role. He takes us to be able to construct symbols that stand for numbers and that reveal the properties of the numbers. But I can't come to know properties of cows by constructing a cow representation! ("The word "cow" does not give milk") According to Kant, this is different in arithmetic and geometry. In these sciences "signs are sensible means of knowledge" ("sinnliche Erkenntnismittel", cf. Kant 1764, Erste Betrachtung §2). Here is Parsons's illustration of the Kantian idea:

> It is natural to think of the natural numbers as represented to the senses (and of course in space and time) by numerals. This does not mean mainly that numerals function as names of numbers, although of course they do, but that they provide instances of the structure of natural numbers (Parsons 1983, 135-136).

Take the number three to be represented by three strokes: / / /. This numeral has exactly one successor ("/ / / /") and one predecessor ("/ /"). In this, the numeral "/ / /" is like the number for which it stands. The perceivable sequence of numerals is isomorphic to the sequence of numbers the numerals stand for. One might even confuse the numerals for the numbers (cf. Frege 1979, 266).

[∀] Many thanks go to Jakob Steinbrenner, Oliver Scholz and Georgina Tsang for helpful written comments.

Let us focus on the important feature of the example: the numerals are symbols that refer to numbers *and* they share higher-order properties like *having exactly one successor* with their referents. The series of the numerals is an intuitive model of the number series: it is a sample of what it stands for or symbolises.

If one looks for further examples of the same phenomenon, it will soon turn out to be ubiquitous. Nelson Goodman supplies another example:

> Consider a tailor's booklet of small swatches of cloth. These function as samples, as symbols exemplifying certain properties (*LA*, 53).

It is indeed independently plausible to say that the swatches in the tailor's booklet are symbols standing for properties they have. The tailor can communicate with his customer by using swatches. For example, he can answer questions by using (exhibiting) a swatch in the right way.

The relation between the sample and the property it stands for is called "exemplification" (Goodman, Elgin) or "exemplarisation" (Lehrer). Not all samples stand for one or more properties they possess. A counterexample to the hypothesis that all swans are white need not stand for the property of being a black swan. Simply possessing the property is sufficient to refute the hypothesis.

Goodman seems to explain exemplification in *Languages of Art*:

> Exemplification is possession plus reference. To have without symbolizing is merely to possess, while to symbolise without having is to refer in some other way than by exemplifying (*LA*, 53).

A play exemplifies properties, because it possesses them and refers to them in a broad sense according to which reference is singling out properties for our attention. A work of art will single out at least some of its properties not by describing, but by *displaying* them.

A notorious problem for Goodman is that an object displays many properties under optimal conditions that are not exemplified. Which of the properties the object displays are exemplified? Goodman's final answer to this problem is to take exemplification as not further explainable in independently graspable concepts. Exemplification cannot be explained, only be exemplified. In this paper I will not take issue with Goodman and

his followers about this point.[1] Whether exemplification can be explained or only exemplified, it is uncontroversial that there are cases of exemplification.

Goodman distinguishes between exemplification and expression. The series of numerals exemplifies the order properties of the series of natural numbers. These are properties, which can be instantiated by abstract and non-abstract objects alike. It is therefore straightforward to say:

> *x* exemplifies F-ness iff *x* refers to F-ness and *x* is F.

In contrast, a picture can exemplify sadness, although it cannot be sad. It is a category mistake to call it "sad". Goodman calls this variety of exemplification 'expression':

> *x* expresses F-ness if *x* refers to F-ness and *x* is metaphorically F.

Exemplification is philosophically interesting because it is an underexplored but fundamental form of representation. Lehrer and Goodman see it as a key to understand how art represents the world and Lehrer uses exemplification in his theory of self-consciousness. Recently, Elgin has argued that exemplification can help us to understand the role of, as she puts it, *felicitous falsehoods*. The felicitous falsehoods that I will focus on are scientific idealisations. I will discuss Elgin's proposal that some felicitous falsehood are fictions (sect. 2) and replace it with another one in section 3. My general conclusion is that exemplification helps to understand scientific idealisation, but not in the way proposed by Elgin.

2. Idealisation as Exemplification by Fictions?

Elgin has applied the notion of exemplification to explain how obviously false representations can be cognitively valuable. Let us consider what she says about scientific idealisations:

> Some laws never obtain. They characterize ideal cases that do not, perhaps cannot occur in nature. The ideal gas law represents gas molecules as perfectly elastic

[1] See (Textor 2008).

> spheres that occupy negligible space and exhibit no mutual attraction. There are no such molecules (Elgin 2004, 118).

There cannot be frictionless surfaces, but physicists give explanations in terms of them; there cannot be ideal gases, but scientists are happy to explain the behaviour of actual gases by seeing them approximate the ideal gases.

How can we understand the role of idealisations in scientific theories? Let us start by thinking about the cognitive role of exemplification in general. Why do scientists build a model of a ship to study its behaviour? They do so, because the model exemplifies now some of the properties the future ship will possess. The practical advantage is that we don't need to build the real ship to test it. In effect, anything will serve the engineering purpose if it exemplifies the properties that the ship will have. Elgin generalises this case:

> If the cognitive contribution of an exemplar consists in the exemplification of select features, then anything that exemplifies exactly those features can, in a suitable context, make the same contribution (Elgin 2004, 126).

The model allows us *now* to study the behaviour of the ship that is likely to exist in the future.

Elgin connects the idea that anything with the right properties may serve as a model to felicitous falsehood and fiction. According to her,

> [Models etc.] qualify as fictions because they diverge from the phenomena in unexemplified properties (Elgin 2004, 126).

This seems misguided to me. The model is a true representation of DNA because it literally shares with DNA properties like *having the relation R between its constituents*. If we would require for truth or correctness of a sign that refers by displaying a property that the sign shares with what it models *every* property, only DNA could be a true (correct) representation of DNA. But imposing this requirement would make models and representation via exemplification, in general, pointless. Hence, the requirement for truth (correctness) is too strong. If we drop it and only require that the representation and the represented object share some properties, there is no longer any reason to say that the DNA model is a fiction. The DNA model may be false, but this does not make it a fiction.

Things are different when it comes to idealisations. What is an ideal gas in particular or an idealisation in general?

> I suggest that idealizations in science function similarly [to samples exemplifying select features]. *The ideal gas is a fiction that exemplifies features that exist, but are hard to discern in actual gases.* The idealisation affords epistemic access to those features, and enables us to explore them and their consequences by prescinding from complications that overshadow the features to ideal gases. (Elgin 2007, 40, my emphasis)

There are several problems with this idea.

First, are ideal gases fictions? There is no ideal gas and there can be none in our world. But that doesn't make the ideal gas a fiction. There is no planet Vulcan and there can be none in our universe. However, it seems clear that Vulcan is no fiction. If one likes non-existing objects, Vulcan is one of them; but not all non-existing objects are fictions.

Second, whether the ideal gas is a fiction or not is not important for Elgin's purposes. Elgin presents an account of the scientific value of idealisations that makes them a variety of a sample or a model. Elgin's account requires that a fiction can share features with an actual object. Take again the ideal gas. It is a non-existent object that consists of molecules with perfectly elastic spheres that occupy negligible space and exhibit no mutual attraction. Since the molecules are perfectly elastic etc., their behaviour is not influenced by factors like friction.

The ideal gas is supposed to possess properties of real gases, select a group of them and, as Elgin says, highlight them. But it is nomologically impossible that there are molecules which perfectly elastic spheres that occupy negligible space and exhibit no mutual attraction. If there can be no such molecules, how can the ideal gas share properties with real gases and make them accessible? They can't and they don't need to share such properties to be epistemically valuable. To see this compare ideal gases with averages.[2] The average British family has 2.4 children. However, there is no British family with 2.4 children. Nonetheless working and calculating with the average British family is useful for planning purposes. This consideration suggests the explanatory value of ideal gases cannot be based on the fact that ideal gases exemplify properties of actual gases.

[2] Thanks to Georgina Tsang for this example.

Third, how can the ideal gas possess properties that actual gases approximately possess? I think it is clear that the ideal gas can't possess any property; it can only *fictionally* exemplify properties that actual gases approximately have.

There is no problem with saying that things fictionally exemplify properties. For example, Sherlock Holmes fictionally exemplifies being an eccentric genius. The Sherlock Holmes' stories either explicitly state or one can deduce from them and plausible background assumptions that Sherlock Holmes is an eccentric genius. As a first stab approximation we can say:

> *x* can exemplify F-ness if *x* refers to F-ness and there is a fiction that has it that *x* is F.

There is no problem of applying this idea to the ideal gas:

> It is fictionally the case that the ideal gas exemplifies F-ness iff according to some fiction, there is an ideal gas that has F-ness.

But if we have to go this way, Elgin's account of idealisations in terms of exemplification loses its explanatory power. Compare Elgin's idealisations with the ship model. We can actually study the ship model and draw conclusions from it. However, we cannot study the properties of an ideal gas, for things that cannot exist have no properties and cannot be telling examples of some properties of gases. The explanatory work must be done by the fiction that has it that there is gas whose molecules are perfectly elastic etc. Instead of making the detour through exemplification, we could directly ask why this fiction is helpful. If we have an answer to this question, appealing to exemplification in addition is superfluous and does not add anything. I propose therefore that we try to connect idealisation and exemplification in a different way.

3. Exemplification, Correction and Pretence

Let us make a fresh start by considering a further example. I may teach you a song by singing it, although I am very bad singer. My singing may even

be barely recognisable as a singing of the song. Nonetheless I manage successfully to teach you the song by the sounds I produce. Why? You recognise that my bad example needs correction and you supply it.[3]

Do I exemplify 'Wind of Change' by my singing, although I don't sing it correctly? If we say YES, something can have a property without having it either metaphorically or literally.

Now one might argue that my incorrectly singing "Wind of Change" is singing it correctly according to some low standard of correct performance. For here as everywhere else we need to be aware that exemplification is relative to contextually determined standards. You ask me what hexagonal is. I draw a figure which is, given my limited talent, not quite hexagonal, point to it and say "This is hexagonal". In this situation it would be unnecessarily pedantic to say that the figure does not exemplify being hexagonal, although in comparison to a very good drawing of a hexagon it hardly qualifies as hexagonal. Hence, we must say:

> *x* can exemplify F-ness if *x* refers to F-ness and *x* is F according to the standard in play in the conversation.

I think the example cannot be explained away in this way. For I might sing incorrectly even according to the standards in play in the context, and yet you might learn the song.

Another way to avoid the result that properties that the sample does not possess are exemplified is to argue that I did sing correctly a song similar to "Wind of Change". But this is highly counterintuitive. A bad performance does not make a new song. We must have room for mistakes in performing a work of art. What makes my singing a singing of "Wind of Change" is my intention in singing, what makes a bad performance are the limits of my musical ability.

Why should one say that my singing exemplifies how the song "Wind of Change" goes? Because the distinctive cognitive effect of exemplification is achieved in the way distinctive of exemplification: properties are selected and highlighted without describing or denoting them. The cognitive effect is achieved although the "sample" does not possess the

[3] Lehrer made this point in a talk in (Kirchberg 2005).

properties because correction mechanisms are in play and there is sufficient background knowledge to draw on.

This modified account of exemplification faces the problem of correction-only cases. Even an unrecognisable singing of a song might be corrected to a proper singing of it. In this case we hardly want to say that my singing exemplified the song. Compare two cases.

First, I sing something which has no similarity to "Wind of Change" by *The Scorpions*, but (i) you know the song and (ii) my intention to sing the song is clearly recognisable for you. In this situation, you will be able to sing the song, but you haven't learned it on the basis of my singing. You can sing the song because you recognised my intention in singing and you knew the song before.

Second, I sing "Winds of Change" slightly of cue. You don't know the song, but you can bring your general understanding of music and musical ability to bear on my performance. In this case, you will be able to sing the song correctly because I sang it incorrectly.

The general recipe for a solution of the worry is that in exemplification the sample and its properties must contribute to a causal explanation of the success of my action. In the correction-only case all the work is done by your recognition of my intention and my singing drops out as idle. Even if I had sung radically different, you would have learned the song. In the exemplification case I got you to know the song because my singing had certain properties and you recognised my intentions and brought your musicality to bear. If you shall be able learn the song because of my 'singing', it must be similar enough to a correct performance of the song that someone who does not yet know it can 'get' it from hearing my singing and applying his general expertise. What "similar enough" is will vary from case to case. This needs refinement, but let's press on.

Generalising from the example we can say that something can exemplify a property even if it does not possess the property, but only approximates the property given that the audience can make the necessary corrections. In general, we must reckon mechanisms similar to correction to be at work in 'understanding' samples. In order to understand a sample I must, for instance, be able to generalise and abstract features. Lehrer makes this point with respect to abstract painting:

> It is not enough to look at it [an abstract painting] in some cases. The person may not understand it because they don't know how to generalise. The functional state of generalising fails and the person has no clear conception of what they are seeing (Lehrer 2004, 10).

Understanding a sample standing for a property requires that one is able to make the step from the particular to the general. Lehrer has described this feature of exemplification as follows:

> Exemplarisation requires an exemplar and the generalisation of it to stand for a class of states or objects.[4]

A critical comment is in order here. The class of Fs is the class of Gs iff these classes have the same members. For instance, the class of things with a kidney is the same as the class of things with a heart. But something can exemplify *having a heart* without *eo ipso* exemplifying *having kidneys*. Hence, classes are to coarsely individuated to be of use in a theory of exemplification. Kinds and properties are better suited, for they have more demanding identity-conditions.

The observation that something that is not F can exemplify F-ness can now be our starting point for an account of idealisation that dispenses with non-existents as samples. How? Think for a start to take the idea of correction to a limit. A non-F may exemplify F-ness if it is to some degree like an F and correcting mechanisms are in play. If you overcorrect, you can arrive at something that is not only an F, but a paradigm case of an F.

Now even if we acknowledge that non-Fs can exemplify F-ness, this move does not straightforwardly allow us to understand idealisations. The ideal gas contains only perfectly elastic molecules. If we know that there are no perfectly elastic molecules we can hardly say that we correct the imperfect elasticity of real gas molecules and thereby exemplify a property they don't possess. The mechanism in play cannot be correction. I will now argue that the mechanism is play is pretence or make-believe.

Consider an independent example. Shakespeare asks his audience for help:

---

[4] See (Lehrer 2003, 428) and (Lehrer 2000, 308). Lehrer himself explains exemplification, roughly, as the use of something *as an exemplar* (Lehrer 2000, 305).

> Piece out our imperfections with your thoughts;
> into a thousand parts divide one man,
> And make imaginary puissance, … (Henry V, Prologue)

The cooperative audience must imagine that the props on stage have other properties than the ones they actually have. Only actual objects are involved in the imaginative project of the audience, namely the props on stage and their properties. These properties are neither possibilia nor fictional objects. According to common sense ontology, properties are actual non-spatiotemporal objects. The audience pretends that the props has properties they actually don't have.

Let us pursue this thought further. As a theatre audience can pretend that a rocking horse is a horse, scientists can pretend that molecules have properties they actually don't have, for instance, perfect elasticity. It may be nomologically impossible that real gas molecules are perfectly elastic. But that does not limit our imagination. Little children can jointly imagine a stone to be a piece of cake. Yet, if the essentiality of origin is true, it is impossible for something that actually is a stone to be a cake. We can imagine counter-possibilities and nomological impossibilities.

If scientists knowingly engage in a joint game of make-believe, real gases can serve as samples of ideal gases. The real gas can exemplify the properties, although they can't have these properties, but only approximate them. For understanding the samples involves not only generalisation, but pretending that they have properties they can only approximately have.

Why is it cognitively valuable to engage in such a joint pretence? Consider again an independent example. You are working for Scotland Yard. You assume that the crime you are investigating has been committed by members of one family. By going through the pictures of the suspects, you can spot similarities, but you are not yet sure. Now you start to ignore distracting features (difference in hair cut) and stress the main features of the people on the picture. In other words: you pretend that they have certain properties to a higher degree than the actually have and other properties to a lesser degree. This pretence is epistemically valuable because it enables you to focus on what the suspects have in common and to ignore features that distort the likeness. This is exactly how Elgin describes the role of the ideal gas:

> The fictional ideal […] serves as a sort of least common denominator that facilitates reasoning about and comparison to real gases (Elgin 2004, 127).

But the valuable cognitive effect has now been achieved without assuming that there is an ideal gas. The real gas does not contain perfectly elastic molecules, nonetheless it may serve to exemplify perfect elasticity in the scientific community.

The difference between Elgin's view of idealisations and the one proposed here can illuminatingly be brought out if we take a leaf from Hume. Following Berkeley, Hume argues that there are no abstract ideas. There are only individuals or particular ideas. Generality in representation is achieved not by entertaining – per impossible – a general idea, but by entertaining a particular idea – what else – and apply it as a general idea is supposed to be applied:

> So abstract ideas are in themselves individual even when they become general in their representation. The image in the mind is only that of a particular object, though the application of it in our reasoning may be same as if it were universal (Hume, Treatise, Book I, I, 7).

We may now state the proposal above in Hume's terminology: The image in the mind can only be of real object yet the application in our reasoning may be the same as if it were ideal. Real things become ideal when we pretend to them to have properties they only approximate and bring them thus under ideal laws.

References

Elgin, C. (1981): *Considered Judgment*, Princeton: Princeton University Press, 1981.

Elgin, C. (2004): "True Enough", in: *Philosophical Issues: Epistemology* 14 (2004), 113-141.

Elgin, C. (2007): "Understanding and the facts", in: *Philosophical Studies* 132 (2007), 33-42.

Frege, G. (1979): *Posthumous Writings*, Oxford: Basil Blackwell, 1979.

Hume, D. (1739): *A Treatise on Human Nature*, edited by Ernest C. Mossner, London: Penguin Books, 1969.

Kant, I. (1764): “Untersuchungen über die Deutlichkeit der Grundsätze der natürlichen Theologie und Moral”, in: *Kant Akademie Ausgabe* Band II, 273-302.

Lehrer, K. (2000): “Meaning, Exemplarisation and Metarepresentation”, in: Sperber, D. (ed.): *Metarepresentation: A Multidisciplinary Perspective*, Oxford: Oxford University Press, 2000, 299-311.

Lehrer, K. (2002): “Self-Presentation, Representation and the Self”, in: *Philosophy and Phenomenological Research* 64 (2002), 412-430.

Lehrer, K. (2004): “Representation in Painting and Consciousness”, in: *Philosophical Studies* 117 (2004), 1-14.

Parsons, C.: “Kant’s Philosophy of Arithmetic”, in: Parsons, C.: *Mathematics in Philosophy*, Ithaca; New York: Cornell University Press, 1983, 110-149.

Textor, M. (2008): “Samples as Symbols”, *Ratio* 21 (2008), 344-359.

Inga Vermeulen, Georg Brun, Christoph Baumberger

Five Ways of (not) Defining Exemplification

Abstract
The notion of exemplification is essential for Goodman's theory of symbols. But Goodman's account of exemplification has been criticized as unclear and inadequate. He points out two conditions for an object x exemplifying a label y: (C1) y denotes x and (C2) x refers to y. While (C1) is uncontroversial, (C2) raises the question of how "refers to" should be interpreted. This problem is intertwined with three further questions that consequently should be discussed together with it. Are the two necessary conditions (C1) and (C2) conjointly sufficient? Do they amount to a definition of "exemplification"? Which notions of Goodman's theory are basic, and hence undefined? In this paper, we address these questions and defend a reconstruction of the notion of exemplification that interprets "refers to" in (C2) as exemplificational reference and hence treats "exemplification" as a basic notion of Goodman's theory. Firstly, we argue that even though the notion of exemplification is not defined, it is still sufficiently clear. This ensures its contribution to Goodman's theory of symbols. Secondly, we show that our account is plausible as an interpretation of Goodman's and Elgin's writings, although it implies that some of Goodman's theorems about self-reference have to be weakened. Thirdly, we argue that it is the only materially adequate reconstruction of Goodman's notion of exemplification, whereas the alternative definitional accounts fail.

In his theory of symbols, Nelson Goodman uses "reference" as a very broad notion covering all sorts of symbolization, all cases of "standing for". He neither attempts to define "reference" nor to explain how referential relationships are established. Rather, he distinguishes and compares different forms of reference. There are at least two fundamental ones, denotation and exemplification. Denotation is reference from a symbol to one or many objects it applies to. Thus, a name denotes its bearer and a predicate each object in its extension. Denoting symbols are called *labels*. Labels do not have to be verbal. A portrait denotes its subject

We would like to thank Catherine Z. Elgin for helpful comments on an earlier draft.

and a music-score the performances complying with it. Exemplification runs in the opposite direction and is reference from an object back to a label that applies to it. Symbols that exemplify are called *samples*. Thus, a chip of paint on a manufacturer's sample card exemplifies, in normal use, a colour-label that applies to it and a tone produced by the concert-master before the performance a label which denotes its pitch. Importantly, no object is in itself a label or a sample, but only if used as part of a *symbol system*. In particular, this implies that a symbol is always a member of a *scheme*; that is, a set of alternative symbols, that is correlated with a *realm*, which is a set of objects to which at least one of the symbols refers.

The notion of exemplification is characteristic of and essential for Goodman's theory of symbols. Goodman introduces it in *Languages of Art* (*LA*, II) and uses it to explain artistic expression as well as the symbolic functions of fictional labels, metaphors, abstract paintings and other nondenotational symbols. It furthermore serves as a basis for analysing important parallels and contrasts between these symbolic functions and the role samples play in the sciences. But Goodman's account of exemplification is not without problems. His writings raise questions of interpretation and his account has been criticized as unclear and inadequate. However, giving up the notion of exemplification would be devastating for Goodman's theory, for it would severely reduce its range of application and its explanatory power. In this paper, we suggest a reconstruction of the notion of exemplification that is materially adequate as well as sufficiently clear to secure the contribution of this notion to Goodman's theory of symbols. Since this theory was systematized and further developed by Catherine Elgin (especially in 1983), we rely both on Goodman's and Elgin's writings.

Section 1 of this paper gives an overview of possible interpretations of "exemplification" that remain in a broadly Goodmanian framework and presents our reconstruction of the notion of exemplification from a purely systematic point of view. We then discuss in section 2 the relevant passages in the writings of Goodman and Elgin and argue that our reconstruction gives a plausible interpretation. Finally, in section 3 we analyse the debate on the notion of exemplification and show that our

reconstruction is the only materially adequate one of the proposals considered.

1. Reconstructing the notion of exemplification

Goodman points out two conditions for something being a sample of a given label. It must be denoted by the label and it must refer to the label (*LA,* 52; 53; *MM,* 59):

(1) There are two conditions for x exemplifying y: (C1) y denotes x and (C2) x refers to y.

1.1 Two necessary conditions

According to condition (C1), exemplification presupposes denotation in the opposite direction. The paint chip on the manufacturer's sample card can be used as a sample of "off-white" only if this label applies to, that is, denotes the chip. But an object does not exemplify all labels that apply to it. Maybe "rectangular" and "manufactured by a well known company" apply to the chip, but those labels are not exemplified if the chip is used in the usual way. Consequently, exemplification cannot be identified with the converse of denotation, but only with a certain subrelation thereof. Condition (C1) is thus necessary but not sufficient for exemplification. Only if y denotes x, may x exemplify y. However, although exemplification presupposes denotation in opposite direction, exemplification is reference in one direction only, namely from sample to label.

Condition (C2) additionally requires that an object refers to the label it exemplifies. The paint chips do not exemplify all the labels they instantiate, but only those they refer to. In normal use, these are colour-labels. However, it is not clear how exactly this second condition should be interpreted. The primary question is how one should interpret "refers to" in (C2). This problem is intertwined with three further questions. Firstly, are the two necessary conditions (C1) and (C2) sufficient, if taken together? That is, can (1) be read as (1.1)?

(1.1) x exemplifies y $\leftrightarrow$ y denotes x $\wedge$ x refers to y.

Secondly, do the two conditions (C1) and (C2) taken together amount to a definition of "exemplification", such that the biconditional in (1.1) could be replaced by "$=_{df}$"? Even though Goodman's semantics is extensional, it does not eliminate the difference between a true biconditional and a definition, since coextensive expressions can differ in secondary extensions.[1] Thirdly, there is the question of which notions are basic and therefore undefined in Goodman's theory of symbols. These four questions must be discussed together since the interpretation of "refers to" in (C2) may have implications for the question of whether (1) states necessary and sufficient conditions, as well as for the question of whether (1) should be interpreted as a definition, which in turn depends on the question of which notions should be taken as the basic ones. Before we introduce and defend our position, we present an overview of possible answers, beginning with the question of how "exemplification" may be defined.

1.2 Five accounts of exemplification

Since denotation and exemplification are the two most important elementary forms of reference, one may opt for interpreting "refers to" in (C2) as either denotational or exemplificational reference. The first option corresponds to the following definition:

(D) x exemplifies y $=_{df}$ y denotes x $\wedge$ x denotes y.

The second option threatens to yield a circular definition: x exemplifies y $=_{df}$ y denotes x $\wedge$ x exemplifies y. This can only be avoided if this option is understood as claiming that (1) should not be interpreted as defining "exemplification", but as introducing a basic notion. In short, it claims:

(E) "Exemplification" is a basic notion (and therefore undefined).

According to proposal (D), "reference" and "denotation" are basic notions of the theory of symbols; according to (E), "exemplification" is a third basic notion. Variants of both (D) and (E) could be generated by

[1] The secondary extensions of an expression P are the extensions of compounds featuring this expression, such as "P-description". For the explication of likeness of meaning in terms of secondary extensions see (*PP,* 221–238) and (Elgin 1983, 54–58); this approach was further developed by Heydrich (1993).

reducing the basic notions to "denotation", or to "denotation" and "exemplification" respectively. For this purpose, "reference" would have to be defined disjunctively: x refers to y $=_{df}$ x denotes y $\vee$ x exemplifies y. The consequences of such a reduction will be discussed towards the end of this section.

A third proposal results if "refers to" in (C2) is interpreted as a basic form of reference, different from both denotation and exemplification. This implies that an additional term, say, "X-reference" has to be taken as basic:

(X) x exemplifies y $=_{df}$ y denotes x $\wedge$ x X-refers to y.

If, however, one wants to stick as closely as possible to the original (1), it will be most natural to leave open the question which form of reference is meant by "refers to" in (C2). Thus, we get a fourth proposal:

(RD) x exemplifies y $=_{df}$ y denotes x $\wedge$ x refers to y.

Like (D), (RD) assumes that "reference" and "denotation" are basic notions. Nevertheless, the two proposals are not equivalent because "refers to" in (RD) may be denotational, but need not be so.[2]

A fifth proposal results if "reference" is assumed to be the only basic notion. This means that both "denotation" and "exemplification" have to be defined as unidirectional and bidirectional reference respectively. So, (R) comprises the following two definitions:

(R1) x denotes y $=_{df}$ x refers to y.

(R2) x exemplifies y $=_{df}$ y refers to x $\wedge$ x refers to y.

Of all these proposals, (X) may be dismissed straightaway. Firstly, it is implausible as an interpretation of Goodman since he never writes about such a third form of reference. This is confirmed by Elgin's explicit claim that there are exactly two basic forms of reference, denotation and exemplification (Elgin 1983, 5; 1993, 172; 1995, 66; 1997a, xvii; 1997b,

[2] Again, a variant of (RD) could be introduced by defining "reference" disjunctively: x refers to y $=_{df}$ x denotes y $\vee$ x exemplifies y $\vee$ x X-refers to y. This would turn (RD) into: x exemplifies y $=_{df}$ y denotes x $\wedge$ (x denotes y $\vee$ x exemplifies y $\vee$ x X-refers to y). But this proposal is rather pointless. Instead of introducing "exemplification" as a disjunct in the second condition, one would better accept (E).

101). Secondly, (X) is not a sensible proposal. Instead of introducing a mysterious basic notion in order to define "exemplification", it would be better to introduce "exemplification" itself as basic. This leaves us with the four proposals (D), (E), (RD) and (R).

The differences between these proposals can be analysed with the help of some statements about reference, denotation and exemplification. To begin with, the following three should hold in any defensible account:

(2) x denotes y → x refers to y.

(3) x exemplifies y → x refers to y.

(4) x exemplifies y → y denotes x.

In the context of Goodman's theory, these are uncontroversial statements. (2) and (3) express that denotation and exemplification are forms of reference; and (4) is just condition (C1) of (1). In all four proposals the three statements can either be proved as theorems or postulated without any problems. (All three are implied by (R1) and (R2), (RD) implies (3) and (4), and (D) implies only (4).) Consequently, there are no important differences between the four proposals with respect to the claims (2), (3) and (4).

But there are remarkable differences between the four proposals with respect to the following two statements, both of which must be rejected:

(5) x exemplifies y → x denotes y.

(6) x denotes y ∧ y denotes x → x exemplifies y.

As we will show, (5) and (6) have utterly implausible consequences and (5) cannot be defended as an interpretation of Goodman's writings. But there is no way to defend (D), (R) or (RD) without accepting (5) or (6), since (5) follows from (D) and from (R), and (6) is implied by all three proposals.[3] (E), however, implies neither (5) nor (6). Therefore, if one is not willing to accept (5) or (6), one has to defend (E), in other words, one has to accept "exemplification" as a basic notion. As far as we know, this has not been proposed unambiguously in the literature, although perhaps many writers have assumed it tacitly. We will defend it in this paper.

[3] In case of (RD), the uncontroversial postulate (2) is needed as well.

Although (E) makes it possible to get rid of the unwanted theorems (5) and (6), this comes at a price. (1.1) cannot be accepted in its original form since (6) is implied by (1.1) and (2). Instead one may accept only a weaker reading of (1) such as:

(1.2) x exemplifies y $\leftrightarrow$ y denotes x $\wedge$ x exemplifies y.

(1.3) x exemplifies y $\rightarrow$ y denotes x $\wedge$ x refers to y.

Of course, these claims look suspicious. For (1.2) is equivalent to (4) and (1.3) is equivalent to the conjunction of (4) and (3). But if we adopt (E), we may interpret (1.2) as emphasizing once more that denotation in the opposite direction is only a necessary condition for exemplification. (1.3) provides an alternative, conditional instead of biconditional, reading of (1).

Such moves make sense only if one is going to defend (E). For the stronger claim (1.1) is implied by (R) and (RD); and in case of (D), nothing substantial can be gained by avoiding (1.1) since (6) follows from (D) directly (the same goes for (R) as well). In short, there is no way to avoid (6) by replacing (1.1) with a weaker claim in any of the alternatives to (E).

1.3 Introducing basic notions

According to (E), "exemplification" is a basic notion of Goodman's theory of symbols. This raises the question of how exemplification may be introduced into his theory. For this purpose, a range of strategies can be used, preferably in combination. They can all be illustrated with passages from Elgin's *With Reference to Reference*. Firstly, examples may be used to point out the intended use of the predicate "x exemplifies y":

> A tailor's swatch, for example, simultaneously *exemplifies* labels denoting fabric, pattern, colour, and weave (Elgin 1983, 79; our emphasis).

> Thus, e.g., if disconnecting the respirator that keeps a patient alive is described as 'killing him', the action is brought to *exemplify* morally blameworthy labels. If it is described as 'ending his suffering', it is brought to *exemplify* morally praiseworthy labels (Elgin 1983, 90; our emphasis).

Secondly, one may refer to similar notions in ordinary language and show how they differ in meaning or usage:

> The *samples* we encounter are various, and the uses to which they are put diverse. The *model* home on a development site, the *prototype* of a jet plane, and the free bottle of shampoo which arrives in the mail are integral parts of sales campaigns. A *sample* problem worked out in a text *is an illustration of* characteristic problems and acceptable modes of solution in a given discipline. And *an example of* the way you can expect to be treated or the sort of person you are likely to become, may serve as a promise or threat (Elgin 1983, 71; our emphasis; cf. *WW*, 65).

But Elgin is not concerned with cataloguing the different roles that samples play. Rather, she wants to determine just what it is for something to function as a sample. This is completed by "Something serves as a sample when it functions as a symbol for a label it instantiates." (Elgin 1983, 72), which in this context indicates that "is a sample of" may be used in a much more general sense than in ordinary discourse. Thirdly, possible misunderstandings can be addressed:

> Let us begin by considering a familiar sample – a chip of paint on a manufacturer's sample card. This particular chip is blue. […] Under its standard interpretation, it is a sample of 'blue', […] Its being a sample of 'blue' *does not depend* on its instantiation of the label being particularly conspicuous or striking as compared with its instantiation of other labels (Elgin 1983, 71; our emphasis).

Finally, one may discuss the relation of the newly introduced term to other technical terms that are already in use:

> Exemplification then is like denotation in being a mode of reference, but it differs from denotation in direction. […] Exemplification is not, however, the converse of denotation. To denote an object, a term need only refer to it. But to exemplify a term, an object must both refer to and instantiate that term (Elgin 1983, 73).

All in all, this procedure of introducing a basic term corresponds to the classic method of explication. According to Carnap (1971, §§2–3), explicating a term means replacing a pretheoretical term ("explicandum") by another term ("explicatum") that is more exact and embedded in a theory. As a preliminary step, the relevant use of the explicandum has to be identified and distinguished from other uses of the same term. The explicatum is required to meet four conditions. (i) In important cases, it can be used instead of the explicandum ("similarity"). (ii) Its use in a system of terms is laid down as precisely as possible. This can but need not be done

by introducing an explicit definition. Finally, the explicatum needs to be (iii) as fruitful and (iv) simple as possible. All these points can be identified in the examples presented above. The first two groups of quotations show how "exemplifies", "is a sample of" and some further expressions are identified as explicanda and how the relevant uses of these terms are explained and illustrated by examples. The predicate "x exemplifies y" is then introduced as explicatum that can be used in place of the explicanda. Its use is regulated by statements such as (1) or the passage quoted last, which put forward the two conditions (C1) ("x exemplifies y → y denotes x") and (C2) ("x exemplifies y → x refers to y"). These conditions ensure that the use of the explicatum is more exactly regulated than the use of the explicanda, even though the explicatum is not defined explicitly, but introduced as a basic notion. A more comprehensive look at Goodman's theory of symbols also shows that the explicatum is extraordinarily fruitful. The notion of exemplification plays a crucial role in Goodman's systematic accounts of artistic expression (*LA*, 85–95) and of the symbolic functions of, for example, fictional labels (*LA*, 21–26; 66), metaphors (*LA* 68–85) and abstract paintings (*WW*, 63–66); furthermore, as Elgin (1983, 81–93; 1996, 171–186) has shown, this explication also serves as a basis for analysing important parallels and contrasts between those symbolic functions and the role samples play in the sciences. Finally, the two conditions (C1) and (C2) give a simple account of the core features of "exemplification".

Nevertheless, one may feel inclined to object that introducing "exemplification" as a basic term leaves this notion unexplained and consequently turns "exemplification" into an unintelligible and useless technical term. But if such objections were justified, they would be similarly valid against denotation – or any other basic notion for that matter. After all, no theory can be developed without relying on one or more undefined notions. Still, one may want to insist on admitting as basic no technical terms but only notions from ordinary language (cf. Textor 2008, 349). But this strategy would result in giving up the core advantages of Goodman's theory. For the ordinary notions of exemplification and sample are far less fruitful. They are, for example, quite useless for explaining how metaphors and fictional labels function.

There is one marked difference between Goodman's introduction of exemplification in *Languages of Art* and the one given by Elgin. Goodman starts with the example of a tailor's swatch and then proceeds to introduce the following conditions (*LA,* 53):

(7) There are two conditions for x exemplifying the property y: (C3) x possesses y and (C2) x refers to y.

A carmine swatch exemplifies carmine because it possesses this colour and refers to it. If one starts with such an example, it is more natural to explain exemplification in terms of properties and possession rather than labels and denotation. But right after the passage just reported, Goodman argues that exemplification should be explained by appeal to labels and denotation as in (1). There is an obvious reason for this move. Goodman's nominalism does not allow for properties but explains possession of properties in terms of being denoted by labels (*LA,* 54–57; cf. *MM,* 59, note 4). Yet, there is another reason not mentioned by Goodman. (1) introduces "exemplification" using the semiotic terms "label", "denotation" and "reference", whereas (7) relies on "properties" and "possession (of properties)", which are not semiotic terms. Therefore, choosing the latter alternative would place non-semiotic terms at the heart of the theory of symbols, as well as introduce an ontological commitment to properties. All this is avoided by (1). Nonetheless, Goodman's strategy of replacing properties by labels is often challenged by pointing out that there are apparently not enough labels (for some properties do not have a label) as well as too many labels (for some different labels refer to the same properties). Goodman (*LA*, 54–57; 67) and Elgin (1983, 76–78) have addressed both problems but the details of this issue are quite irrelevant to the topic of this paper. We therefore restrict the following discussion to Goodman's official theory, which explains exemplification in terms of labels, denotation and reference.

As we have introduced (E), all three notions, "denotation", "exemplification" and "reference", are taken to be basic, and hence undefined. But as mentioned before, "reference" could easily be defined as the disjunction of its basic forms. However, such a definition would presuppose a complete list of the basic forms of reference and it is not clear

that such a list can be given once and for all. In one passage, for example, Goodman points out that reference cannot always be interpreted as denotation or exemplification, since an object may be used as a symbol for something on the basis of almost any relation (*LA,* 65). While this rules out defining "reference" by enumerating all its basic forms, it leaves open the possibility of introducing "reference" using an open list of such forms. On the other hand, treating "reference" as a third basic notion also has its drawbacks. The resulting symbol theory will have more basic notions and consequently be less simple (in the sense explained in *The Structure of Appearance*), systematic and coherent (cf. *SA*, III; *PP*, VII). Nevertheless, we think it is better on the whole to accept these disadvantages than to rely on the rather dubious assumption that denotation and exemplification can be identified as the only basic forms of reference once and for all. For this reason, we stick to the original version of (E) that includes "reference" as a basic notion.

2. Goodman and Elgin on exemplification

In this section we investigate whether Goodman's and Elgin's writings support our interpretation (E) of (1) that treats "exemplification" as a basic notion of Goodman's symbol theory. Since at least Goodman never explicitly formulates (E) we have to proceed indirectly and ask first whether there is an official definition of "exemplification" in Goodman's writings.

2.1 Is there an official definition of "exemplification"?

In all of Goodman's and Elgin's writings there is but one place with an explicit hint that "exemplification" might be defined. The index of *Languages of Art* contains an entry "exemplification, definition of" referring to page 52. Presumably the statement in question is:

> An object that is literally or metaphorically denoted by a predicate, and refers to that predicate or the corresponding property, may be said to exemplify that predicate or property (*LA,* 52).

However, it is quite questionable whether this hint can be taken seriously. Should this passage be interpreted as providing a definition in a strict

sense? Even Goodman's official definition of "exemplification"? A positive answer is implausible for several reasons. Firstly, Goodman did not write the Index to *Languages of Art* (cf. *LA*, vi). Secondly, the structure of Goodman's text speaks against such an interpretation. The quoted passage is nowhere referred to as a definition in the text itself, it occurs *before* the chapter heading *Exemplification* and there are many more or less similar passages which could be interpreted as definitions of "exemplification" with equal (im)plausibility. Thirdly, an entry "…, definition of" in the index does not imply that the page it refers to contains a definition of "…", for such a definition may only be talked about. There is, for example, an entry "expression, definition of" referring to a passage (*LA,* 95) where Goodman explicitly states that he is not defining "expression". Finally, it is implausible that Goodman should include the term "property" in an official definition, since he uses it only to conform to customary ways of speaking.

2.2 Can "exemplification" be defined?

All relevant alternatives to our proposal (E) are suggestions for defining "exemplification". In order to judge their plausibility from an exegetic point of view, we discuss passages that throw some light on the question of whether and how "exemplification" *could* be defined. At least for (RD) there are a number of formulations in Goodman's writings that could, with some plausibility, be read as definitions of "exemplification", namely those stating that exemplification *is* being denoted (or possession) plus reference (e.g. *LA,* 53). If "is" here is read in a definitional sense, (RD) results directly. But let us first turn to proposal (D).

According to (D), the relation called "reference" in (C2) is denotation. But this view is challenged by passages in *Of Mind and Other Matters* where Goodman explicitly claims that exemplification is not a denotational form of reference. For example:

> An extremely important but often overlooked form of *nondenotational* reference is exemplification […] Exemplification, then, far from being a variety of denotation, runs in the opposite direction, […] Exemplification indeed involves denotation, by inversion, yet cannot be equated with the converse of denotation (*MM,* 59; Goodman's emphasis)

> Although exemplification is not a denotational relation at all, […] (*MM,* 60)

A notably clear statement can be found in Goodman's reply to Richard Martin's complaint that aesthetic relations, such as exemplification, should not be subsumed under reference or denotation (Martin 1981, 261):

> I do not, as Martin implies in one place, "subsume [these] relations under denotation" but rather under the broader relation of reference. Exemplification and expression are *non*denotational but referential relations (*MM,* 92, note 27; Goodman's emphasis).

This is Goodman's interpretation of his position in *Ways of Worldmaking*. His formulations in *Languages of Art* are somewhat less unambiguous. However, Goodman writes in a letter to Monroe Beardsley:

> […] *presenting for apprehension* is something more than merely *having* – it is emphasizing, calling attention to, showing forth, exemplifying (and so referring to though not denoting) the properties in question. I use "reference" for the general relation of symbolization; exemplification and denotation are different kinds of reference (see, e.g., *Languages of Art,* pp. 51–52, 57ff, 92) (Goodman 1975, 25; his emphasis).

This leaves no doubt that Goodman considers the claim that exemplification is a nondenotational relation to be part of the position he presented in *Languages of Art*. Hence, Goodman explicitly rejects (D) as an interpretation of *Languages of Art*. This, however, leaves open the question of whether denotation might be a necessary or a sufficient condition for the relation of reference mentioned in (C2). We will address this issue a bit further on. For now, we may note that the passages discussed so far imply that Goodman accepts a second form of reference that cannot be reduced to denotation. Interpretation (E) assumes that this nondenotational form of reference simply *is* exemplification, which consequently needs to be accepted as a basic notion. We will therefore briefly look at what Goodman and Elgin write about the basic notions of their theory of symbols, extending the discussion to include proposals (R) and (RD).

In *Routes of Reference* and *Reconceptions* Goodman comments on the conceptional structure of his theory of symbols:

> "Reference" as I use it is a very general and primitive term, covering all sorts of symbolization, all cases of *standing for.* As a primitive relation, reference will not be defined but rather explicated by distinguishing and comparing its several forms (*MM,* 55; Goodman's emphasis).

> The basic notion is *reference* or symbolization, the relation between a symbol and whatever it stands for in any way. As primitive, this relation is not defined but explicated in terms of its several varieties, among them (1) *denotation*, reference by a word or other label to something it applies to, as in naming or predication, and (2) *exemplification,* reference by an instance, as a sample, to a label denoting it (*R,* 124; Goodman's emphasis).

These passages leave no doubt that reference is a basic notion of Goodman's theory of symbols, and hence undefined. With respect to the various forms of reference, Goodman also makes clear that he does not intend to define but only to distinguish and compare these notions. We should therefore not expect definitions of "denotation" or "exemplification". Neither should we assume that Goodman is committed to the claim that these notions could be defined with reference to "reference" (and logical notions) alone – or in any other way.

In other respects, however, these passages pose some problems of interpretation. How should one understand Goodman's announcement that he will *explicate* the notion of reference? Why is only reference called "basic" and "primitive"? And what is meant by these attributes?[4] We start with Goodman's use of "explication". Firstly, explications are often thought of as explicit definitions satisfying certain criteria. But this cannot be what Goodman wants to express since he does not intend to define "reference". Secondly, it is implausible that Goodman intends to give an explication in Carnap's classical sense. Although this would not necessarily call for an explicit definition, it would imply that Goodman intends to give a characterization of the notion of reference meeting Carnap's criteria of adequacy (cf. section 1.3). But Goodman adopts only

[4] In *Routes of Reference,* exemplification is also called an "elementary" form of reference. But, as the headings of the main sections in *Routes of Reference* suggest, this probably means only that exemplification is a non-complex form of reference; that is, no chain of reference (cf. *MM*, 61f; 70).

the less ambitious goal of distinguishing and comparing forms of reference. It may therefore be best to interpret him in a third way as using "explication" not terminologically but only in the sense of "clarifying a notion", corresponding to the first step of Carnap's method of explication.[5] This interpretation could be supported by interpreting "basic" and "primitive" in the way they are used in *The Structure of Appearance* (45–46). This would imply that "reference" is introduced without being defined. On the other hand, the basic and primitive notions in *The Structure of Appearance* are introduced as a basis for defining further notions. But, as we have already seen, this is not what Goodman declares as his intention in the quoted passages. He does not claim to define "denotation" or "exemplification", much less to develop a system of semiotic notions using the constructional techniques of *The Structure of Appearance.* Consequently, it is not clear whether the quoted passages can sensibly be interpreted with reference to Goodman's theory of constructional systems in *The Structure of Appearance*. All in all, it seems fairly questionable whether the fact that Goodman calls only reference "basic" and "primitive" should be taken to imply that "denotation" and "exemplification" are notions that are not basic but defined. The best interpretation may be that Goodman leaves open the question of whether and how "denotation" and "exemplification" can be defined in a strict sense. This respects at least the general line of *The Structure of Appearance,* which calls for selecting the primitive notions according to criteria that can be assessed only on the basis of a constructional system (e.g. simplicity). To sum up, Goodman does not claim to give a definition of "exemplification", but only to distinguish and compare exemplification and other forms of reference. This neither excludes nor implies that "exemplification" could be defined. It therefore seems sensible to treat

[5] In Goodman's writings, "explication" and "explicative" appear only casually and with various meanings. In the context of *The Structure of Appearance* explications with explicit definitions are called "definitions" (*SA,* I) or "(explicative) definitions" (*PP,* 4); informal explanations of notions are called "informal explications" (*SA,* 19, note 8). In *Fact, Fiction, Forecast* (65, note 2), his use of "explication" can be interpreted in all three ways we have distinguished.

exemplification as an undefined notion (at least at the present stage of development of Goodman's theory). This interpretation can be supported by Goodman's passages that introduce the notion of exemplification (e.g. *LA*, 51ff). There we find all the strategies for introducing an undefined notion that we have discussed in section 1.3, although Goodman is somewhat less strict, especially in his liberal use of property-talk.

Elgin's writings, however, are unambiguous:

> Goodman […] recognizes two primitive semantic relations – denotation and exemplification – and constructs a variety of complex modes of reference out of them (Elgin 1997a, xvii; 1997b, 101).
>
> Like Goodman, I recognize two basic modes of reference – denotation and exemplification (Elgin 1995, 66).

According to Elgin, denotation and exemplification are the two basic and primitive forms of reference, and this is also her interpretation of Goodman. We read these quotes as claiming that "denotation" as well as "exemplification" are basic notions of the theory of symbols. Even without explicitly using the word "undefined", treating "denotation" and "exemplification" as notions which are on the same conceptual level not only implies that "exemplification" may not be defined as a form of denotation but also rules out that "exemplification" is defined any more than "denotation". But all interpreters we know of agree that Goodman and Elgin do not define "denotation", and there is not a single passage that could plausibly be read as such a definition (cf. *MM*, 55–59; *LA*, 3–6; Elgin 1983, 19).[6]

This point can also be made more generally. The whole discussion about basic, primitive and undefined notions could be repeated for "denotation" instead of "exemplification". The only difference would be that there is not even a suspicious entry in the index of *Languages of Art*. Consequently, it is hard to see why one should believe that Goodman intended to treat the

[6] One could venture the claim that Goodman defines "x denotes y" as "x applies to y" (in analogy to Quine's alleged definition as "x is true of y", which would be too narrow for Goodman). But this is not a definition in any interesting sense but only the introduction of an alternative label for the same relation.

two notions "denotation" and "exemplification" markedly different with respect to their introduction and definition. To put it more polemically, whoever thinks that Goodman has or needs a definition of "exemplification" should also tell us where to find a definition of "denotation" or at least why we should not need it. It is much more plausible to assume that both terms are not defined but treated as basic in Goodman's theory.

All in all, the points discussed so far undermine interpretations (D), (R) and also (RD) since these proposals all assume that Goodman defines "exemplification";[7] and they speak in favour of (E) since (D), (R) and (RD) are all the relevant alternatives to (E). To confirm this result, we turn to the question of whether denotation in the same direction might be a necessary or sufficient condition for exemplification.

2.3 The relationship between denotation and exemplification

In his letter to Beardsley, Goodman explicitly rejects the idea that denotation in the same direction is necessary for exemplification: "referring to though not denoting" (Goodman, 1975, 25). Further support for this interpretation of *Languages of Art* is provided by the text itself. Firstly, there is not a single passage suggesting that the labels a sample can exemplify should be restricted to those it denotes. Rather, Goodman introduces exemplification with examples that make it utterly implausible to think that the samples denote the labels they exemplify. Even though they exemplify, a tailor's swatches in general do not denote at all, and Goodman does not claim that they do, but only that they refer to the labels they exemplify. Since denotation has not only been introduced before exemplification, but is also used in the explanation of exemplification, Goodman surely would have explicitly pointed out that denotation in the same direction were necessary for exemplification if this indeed were true.

---

[7] An additional problem challenging (R) is Goodman's remark "denotation implies reference between two elements in one direction while exemplification implies reference between the two in both directions" (*LA,* 59). According to (R), we should find "is" in place of "implies". If (R) were correct, Goodman had no reason to prefer the weaker claim to the stronger.

Secondly, Goodman draws attention to the fact that usually the practical distinction between instances of denotation and exemplification is easily made in the case of ordinary language:

> In ordinary language, the reference of "man" to Churchill, and of "word" to "man", is unequivocally denotation; while if Churchill symbolizes "man", and "man" symbolizes "word", the reference is unequivocally exemplification (*LA,* 57–58).

Writing that the reference from Churchill to "man" is *unequivocally* exemplification would be highly misleading if this reference were also an instance of denotation. If denotation in the same direction were necessary for the reference mentioned in (C2), there would be an ambiguity in every instance of exemplification; for a sample then could not exemplify a label without denoting it as well.

Finally, Goodman's extensive discussion of a difference in direction between denotation and exemplification (cf. *LA,* 52, 57ff, 65; *MM,* 59, 82; *R*, 36) would completely miss the point if denotation in the same direction were necessary for exemplification. If this were true, the crucial difference between denotation and exemplification would have nothing to do with a difference in direction, but would amount to exemplification being bidirectional denotation while "normal" denotation would be just unidirectional denotation, as claimed in proposal (R).

In short, denotation in the same direction is not a necessary condition for exemplification and, hence, statement (5) – "x exemplifies y → x denotes y" – is false. As a consequence, neither (D) nor (R) can be defended as an interpretation of Goodman's position, since they both imply (5).

However, whether denotation is sufficient for the relation of reference in (C2) is a more difficult question. Goodman addresses it when he deals with the special case of self-reference (*LA,* 59). After mentioning some examples he notes four "theorems":

(a) If x exemplifies y, then y denotes x. [= (4)]
(b) x and y denote each other if and only if they exemplify each other.
(c) x exemplifies x if and only if x denotes x.
(d) If x exemplifies and is coextensive with y, then x denotes and exemplifies x (*LA,* 59, note 9).

The crucial claim is (b) since (a) is uncontroversial, (c) follows from (b), and (d) from (a) and (b).[8] That Goodman calls (b) a "theorem" suggests that he thinks it can be derived from some more fundamental statements about denotation and exemplification. But the uncontroversial (a) only implies "half" of (b): "If x and y exemplify each other, then they denote each other". What is missing for deriving (b) is the already mentioned claim:

(6) x denotes y $\wedge$ y denotes x $\rightarrow$ x exemplifies y.

In (D), (R) and (RD), (6) can be derived from the respective definitions of "exemplification" (together with, in case of (RD), the uncontroversial (2)). In interpretation (E), however, (6) cannot be derived as a theorem unless one postulates (1.1) or some other suitable statement. Thus, from an exegetic point of view, the fact that Goodman's theorems (b)–(d) presuppose (6) supports (D), (R) and (RD), since (6) is a *theorem* in those interpretations, but an independent claim in (E). From a systematic point of view, however, the opposite is true. Unlike (a), (6) is highly problematic, as will be shown in section 3. This will amount to a powerful argument for (E): in (E), (6) can be given up, but not in (D), (R) and (RD).

Before picking up this argument, we may note that, again, the writings of Elgin fit our proposal (E) even better. Elgin formulates neither the theorems mentioned nor (1.1) which, together with the uncontroversial postulate (2), implies the problematic claim (6). Instead of (1.1) she uses the unproblematic (1.3), which is compatible with (E), as the "official" characterization of exemplification: "When an object exemplifies a label, it both refers to and instantiates that label" (Elgin 1983, 73).

3. The debate on the notion of exemplification

We now turn to systematic arguments in favour of (E), analysing the debate on the notion of exemplification. Since (E), (D), (R) and (RD) are

[8] The plausibility of (d) depends on how exemplification of coextensive labels is dealt with. Goodman considers some alternative approaches in *LA* (54–56). We will not discuss this issue.

all proposals for interpreting the two conditions for exemplification that Goodman puts forward in (1), we will focus on the question of how "refers to" has to be understood in (C2). We will not discuss positions that cannot claim to be at least broadly Goodmanian. This includes not only the suggestion that the notion of exemplification should be abandoned, but also all non-semiotic accounts of exemplification; that is, all proposals that replace "refers to" in condition (C2) with a non-semiotic relation.[9] Some authors advocate such a move since they think that Goodman's account of exemplification faces insurmountable problems. We attempt to show that these problems vanish if interpretation (E) is accepted.

3.1 Dempster's dilemma

Douglas Dempster published a widely read critique of Goodman's notion of exemplification. He considers three possible readings of "refers to" in condition (C2) of (1) (Dempster 1989, 407–410):

(8) x exemplifies y.

(9) x denotes y.

(10) x refers to y in some other way.

Option (10) leads to definition (X). It has already been criticized in section 1 and Dempster rejects it as well. Since Dempster argues that the two remaining proposals lead to unsolvable problems, his attack may be reconstructed as a dilemma.

The first horn of the dilemma is generated by (8), if "exemplification" is defined as

(11) x exemplifies y $=_{df}$ y denotes x $\wedge$ x exemplifies y.

Since this is obviously a circular definition, no author has seriously considered adopting it (cf. Dempster 1989, 407; Steinbrenner 2005, 230).

---

[9] For non-semiotic accounts of exemplification see, e.g., Beardsley 1975, 13–19; 1978, 102–106; Dempster 1989, 410–412; Textor 2008, 351–359. Beardsley sometimes also shows a tendency to give up the notion of exemplification altogether (cf. 1978, 106).

The second horn is the result of choosing option (9); that is, of adopting definition (D) of "exemplification":

(D) x exemplifies y $=_{df}$ y denotes x $\wedge$ x denotes y.

Dempster (1989, 407) argues that (D) provides the only acceptable interpretation of Goodman's position since (11) has to be rejected for the obvious reason already mentioned. Probably, this interpretation has also been inspired by Goodman's theorems of self-reference. But we may note that although the theorems follow from (D), one can accept them without being committed to (D). As we have discussed in section 2, denotation could be sufficient but not necessary for the relation of reference that is called for in (C2).[10] Nevertheless, proposal (D) is materially inadequate for three reasons (Dempster 1989, 407f).

Firstly, (D) blurs the distinction between exemplification and denotation, and consequently between samples and labels. The reason is that according to (D) there are no longer two fundamental forms of reference but only one, denotation, which can be unidirectional or bidirectional. Exemplification then simply is bidirectional denotation. But this contradicts not only Goodman's and Elgin's frequent assurance that there are two fundamental forms of reference, but also their insisting on a difference in direction. For if exemplification is denotation in both directions, there is no "opposite" direction left that could distinguish exemplification from denotation.[11] Furthermore, labels would no longer be distinguished from samples as being denotational symbols because samples would denote as well.

---

[10] This may be what Dempster wants to express when he writes: "This [(D)] may be too strong. Goodman seems to hold only that compound denotation is sufficient for exemplification, indeed sufficient for compound exemplification" (1989, 407, note 24). Dempster's use of "compound" is ill-advised since it suggests a chain of reference, but exemplification is not a complex form of reference. We will use therefore "bidirectional" instead of "compound".

[11] The difference in direction could be saved if (D) were interpreted as requiring not bidirectional denotation, but only denotation that presupposes denotation in the opposite direction. But then it would be impossible to say which of the two relations is exemplificational. This is the objection of symmetry, which is discussed next.

Secondly, (D) renders exemplification symmetric. If the conditions for x exemplifying y are met, y exemplifies x as well. But surely, a sample is not exemplified by everything it exemplifies. That a colour sample exemplifies "red" does not imply that "red" exemplifies that colour sample. Even restricted to cases of bidirectional denotation, exemplification is not symmetric. If "word I have not thought of yesterday at 3:07 a.m." is used as a sample for a predicate and hence exemplifies "predicate", then this does not imply that "predicate" exemplifies "word I have not thought of yesterday at 3:07 a.m.", although the two expressions denote each other.

Thirdly, according to (D), bidirectional denotation is sufficient for exemplification since (D) implies (6). But there are countless counter-examples featuring bidirectional denotation without exemplification. On the one hand, there are cases of bidirectional denotation as in the preceding example. On the other hand, there are cases of self-denotation, which perhaps are even more convincing counter-examples. Here are two:[12] When asked to illustrate what counts as a German word, I may start with "kurz". This word denotes itself, although in the context mentioned it does not exemplify itself, but rather "German word". Or suppose I am writing in blue ink on a postcard "The blue sea is here and I wish you were beautiful". The word-token "blue" denotes itself without exemplifying itself, for it does not exemplify at all.[13]

Constructing a dilemma with the two horns (11) and (D), as Dempster does, presupposes that "exemplification" should be defined. In a nutshell, the dilemma is this: in Goodman's definition of "exemplification" (x exemplifies y $=_{df}$ y denotes x $\wedge$ x refers to y) "reference" must either be

---

[12] Dempster's counter-examples are of limited use (1989, 408) since he does not specify any context of use. Therefore, one may imagine circumstances in which the label in question would indeed be used to exemplify as well. Similar counter-examples can be found in Textor (2008, 349f).

[13] The problem is not that we sometimes do not know what a symbol exemplifies. This indeed happens if we face difficulties in finding out about the symbol system in use or in making it explicit; this difficulty affects denotational systems as well, but it is more common with exemplificational systems because they are less standardized. The relevant problem is that according to (D) a symbol would sometimes have to exemplify something we know it does not exemplify in the given context.

interpreted as exemplificational or as denotational reference; but in the first case the definition is circular, in the second it is inadequate; hence Goodman's definition of "exemplification" has to be rejected.

The dilemma vanishes if one gives up the presupposition that "exemplification" should be defined. This opens the possibility of interpreting "refers to" in (C2) as exemplificational without running into an inacceptable circularity. Goodman's statements expressing (1) may be read not as defining a term but as introducing a basic notion. Since "refers to" in (C2) now is exemplification, it can no longer be denotation. This does not imply that an exemplificational relation cannot be accompanied by a denotational relation in the same direction. But even if there is such a denotational relation it is different from the exemplificational one. Consequently, the problems of the second horn disappear. Firstly, it is guaranteed that the difference between exemplification and denotation as well as between samples and labels remains clear. There are two basic forms of reference, which differ in direction. Of course, it can happen that we run into problems when we try to decide whether a particular symbol denotes or exemplifies what it refers to. But the epistemological question of what criteria may be used for finding out about whether something denotes or exemplifies must be distinguished from the question of what the conditions are for something to denote or exemplify something. In this paper we exclusively deal with such conditions, but not with epistemological criteria. Secondly, exemplification is not forced to be symmetrical if bidirectional denotation is not sufficient for exemplification. Even in cases of bidirectional denotation, there are still four possibilities: no exemplification, exemplification in one or in the other direction only, or exemplification in both directions. Thirdly, the counter-examples can easily be dealt with since bidirectional denotation and self-denotation are no longer sufficient for bidirectional exemplification and self-exemplification respectively. Even though the expressions "word I have not thought of yesterday at 3:07 a.m." and "predicate" denote each other, there can now be situations where the second does not exemplify the first one; and although "kurz" and "blue" written in blue ink denote themselves, they need not exemplify themselves. For example, they do not exemplify themselves in the contexts described above.

Incidentally, we may note that the problems discussed are independent of the definitional nature of (D). They also challenge the corresponding claim that bidirectional denotation is necessary and sufficient for exemplification. If we give up the idea of defining "exemplification", we can replace (11) by (1.2) and thereby avoid the charge of circularity. If, on the other hand, we replace (D) by

(12) x exemplifies y $\leftrightarrow$ y denotes x $\wedge$ x denotes y,

we still face all three objections.

Dempster's dilemma could also be attacked in another way, by claiming that the alternatives he considers are not exhaustive. After all, there are more than three possibilities how "exemplification" could be defined. Specifically, there are (R) and (RD), which are quite obvious proposals, although Dempster manifestly has overlooked them completely. However, this strategy for dealing with Dempster's dilemma ultimately fails: (R) and (RD) are materially inadequate as well.

3.2 Extending the argument

No new objections are needed for attacking (R) and (RD). According to (R), "reference" is the only basic notion and "denotation" as well as "exemplification" are defined in terms of reference:

(R1) x denotes y $=_{df}$ x refers to y.

(R2) x exemplifies y $=_{df}$ y refers to x $\wedge$ x refers to y.

This proposal has the obvious advantage that it manages to do with one basic notion only. But it has to be rejected for the same reasons as (D). Firstly, it obscures the distinction between exemplification and denotation and hence between samples and labels. Suggesting that there is only one basic form of reference and that exemplification *is* bidirectional reference is neither compatible with claiming that there are two fundamentally distinct forms of reference, nor with a difference in direction between denotation and exemplification. Secondly, (R) turns exemplification into a symmetric relationship, which is implausible as we have seen already. Thirdly, the counter-examples challenging (D) cause the same troubles for advocates of (R). They show that bidirectional denotation is not sufficient

for exemplification, but (R1) and (R2) imply exactly the contrary, namely (6).

Proposal (RD) accepts two notions as basic and defines

(RD) x exemplifies y $=_{df}$ y denotes x $\wedge$ x refers to y.

Although this has some plausibility as an interpretation of Goodman's text (cf. section 2.2), it nonetheless has to be rejected since it cannot deal with the counter-examples any more than (D) and (R). The reason is that (6) follows from (RD) and the uncontroversial (2) ("x exemplifies y $\rightarrow$ x refers to y"). Bidirectional denotation therefore is sufficient for exemplification according to (RD) and this is all that is needed for setting up the counter-examples. Again, the problems challenging (R) and (RD) cannot be evaded by giving up the definitions (R) and (RD) and resorting to the corresponding biconditional claims.

To sum up, (D), (R) and (RD), the alternatives to our proposal (E), are inadequate since (6), which is responsible for the counter-examples, cannot be given up in those accounts. Furthermore, although (1.1) could be held together with (E), this makes little sense since (1.1) and (2) together imply (6). Consequently, if (E) is adopted, (1.1) must be replaced by the weaker (1.2) or (1.3). This move is possible only because (E) does not imply (1.1) as (R) and (RD) do. Albeit (1.1) could also be given up in (D), this would be of little help, since (D) directly implies (6); consequently the unwanted implications cannot be avoided by replacing (1.1) with (1.2) or (1.3).

3.3 Self-reference

Importantly, the counter-examples discussed provide good reasons for rejecting not only proposals (D), (R) and (RD) but Goodman's theorems of self-reference as well. The counter-examples directly attack (b) and (c), which consequently must be given up. Once more, this can be done only in (E), since (b) and (c) are not theorems in a strict sense in (E). (D), (R) and (RD), on the other hand, imply (6),[14] which together with the uncontroversial (a) implies the theorems (b)–(d).[15]

---

[14] In case of (RD), (2) is needed as well.

[15] (d) or its negation can be added to the postulates of (E), which are (2), (3), (4),

In (E) the problems are solved by accepting only (a), whereas (b) and (c) have to be replaced by weaker claims, which follow from (1.2) as well as from (1.3):

(13) If x and y exemplify each other then they denote each other.

(14) If x exemplifies x then x denotes x.

This completes our argument in favour of (E) and against (D), (R) and (RD). In the rest of this section we briefly look at two interpretations of (1) we have not discussed so far.

3.4 Peltz's account

Richard Peltz has presented a further interpretation of "refers to" in (1). He complains that Goodman characterized the reference mentioned in (C2) only negatively as non-denotational and consequently did not make clear that it is indeed a form of reference. As a remedy, he suggests replacing reference with denotation in (C2), but in a different way than (D) or (12) do. His proposal can be reconstructed as follows:[16]

(15) x exemplifies y $\leftrightarrow$ y denotes x $\wedge$ x denotes everything denoted by y.[17]

This proposal is meant not as an interpretation but as a modification of Goodman's account; Peltz does not claim to explain what Goodman wrote or meant, but what he should have written. According to (15) the red colour-sample exemplifies "red" if and only if "red" denotes the sample and the sample denotes everything red. Peltz's move is to replace the condition that x must refer to y by the condition that x must denote everything denoted by y, which includes x itself. Now, Goodman himself

and (1.2) or (1.3); (6) still does not follow.

[16] The reason why Peltz favours (15) may be related to his view that denotation is the paradigmatic form of reference. He mistakenly attributes this stance to Goodman: "For Goodman the paradigm of symbolism is denotation. To be a sign or symbol is to denote" (Peltz 1972, 73).

[17] Peltz writes (1972, 81): "'A exemplifies B' can mean ''B' denotes A (A possesses B) and 'A' denotes what 'B' denotes'". This formulation raises some further problems, which are discussed by Dempster (1989, 401–403) and Coldron (1982, 90–93). Additionally, Peltz would need to understand (15) as a definition given his motivation for introducing such a claim.

notes that a sample which lacks an established denotation sometimes can be understood as denoting everything that is denoted by the label it exemplifies; it then will function as a label that exemplifies itself (*LA,* 63). But this is typically not possible for samples that already have an established denotation. Even if "long" is used as a sample for short words and consequently exemplifies "short", it cannot take over the denotation of "short" since it denotes long objects, not short ones. Peltz tries to avoid this problem by introducing a notion of semi-metaphorical exemplification that would allow for claiming that "long" metaphorically denotes the objects literally denoted by the exemplified label "short" (Peltz 1972, 84).

Apart from introducing the rather dubious notion of semi-metaphorical exemplification, Peltz's proposal fails for two reasons. Firstly, every symbol that denotes itself meets the conditions mentioned in (15) and consequently exemplifies itself. Hence we have the same counter-examples as with (D). Secondly, while Goodman requires that a sample refers to the label it exemplifies, Peltz requires that it denotes what the exemplified label denotes. This raises the problem of coextensive labels. If Socrates exemplifies "rational" and "rational" is coextensive to "risible", then Socrates exemplifies "risible" according to (15); but surely, Socrates can exemplify "rational" without exemplifying "risible" (cf. *LA,* 55; Elgin 1983, 76).[18]

3.5 Steinbrenner's account

Steinbrenner has presented an interpretation of (1) which is designed to counter the charge of circularity. His proposal can be put as follows (Steinbrenner 1996, 89; 2005, 231):

(16) An object x exemplifies a label y $\leftrightarrow$ x belongs to the extension of y $\wedge$ x refers to y $\wedge$ x indirectly refers to the extension of y.

[18] Dempster already raises these objections (1989, 403f). But he claims that Peltz faces the problem of coextensive labels because he allegedly holds that the objects referred to by exemplification are extensions. This is a misinterpretation of Peltz, who shares Goodman's view that samples exemplify labels. Rather, the problem of coextensive labels is generated by the second condition in (15).

Presumably, the third condition is to be understood as equivalent to "x indirectly refers to the objects belonging to the extension of y" since he gives the following reference for this condition: "Exemplification relates the symbol to a label that denotes it, and hence to the things (including the symbol itself) in the range of the label" (*LA,* 92). Proposal (16) reformulates condition (C1) of (1) by replacing "is denoted by" with the equivalent "belongs to the extension of" and it adds a third condition that is trivially met and hence superfluous. But it gives no answer to the question of how "refers to" is to be understood in (C2) since (16) retains the original wording of condition (C2). Consequently, Steinbrenner's proposal provides no solution to the problem of circularity that results if "refers to" in (16) is interpreted as exemplification. Steinbrenner avoids deciding on an interpretation of "refers to" and thereby he avoids the problem of circularity without getting rid of it. It would be eliminated if (16) was meant as introducing an undefined notion, similar to (E). At the beginning of his paper (Steinbrenner 2005, 228), Steinbrenner gives a hint in this direction when he points out that Goodman does not intend to define "denotation" and "exemplification" (with reference to *MM,* 55, where Goodman makes such a claim for "reference" only). But later on he explicitly treats (16) as a definition (Steinbrenner 2005, 230, note 5).

Perhaps Steinbrenner is worried by another threat of circularity. Indeed he writes that one may suspect a circularity "because Goodman seemingly wants to define the term 'reference' with the help of the terms 'denotation' and 'exemplification', but at the same time cannot explain the term 'exemplification' without the help of the terms 'denotation' and 'reference'" (Steinbrenner 2005, 228). This suggests that the circularity is not generated by the definition of "exemplification" alone but by its interplay with the definition of "reference"; it would result because "reference" is defined in terms of exemplification, which is defined in terms of reference. But Steinbrenner's proposal cannot avoid this circularity because in (16) "exemplification" is still defined with the help of "refers to". It is no surprise that Steinbrenner does not eliminate this circularity since, as he rightly insists, there is no such circularity because Goodman does not intend to define "reference" (Steinbrenner 2005, 228).

Sometimes, however, Steinbrenner seems to think that circularity might arise since Goodman characterizes "reference" with the help of two terms while explaining one of them ("exemplification") in terms of the other ("denotation") (cf. Steinbrenner 1996, 85f). His proposal is intended to solve this problem by keeping denotation out of (16). However, it is not easy to see what the circularity should be, and even less how it could be eliminated by resorting to (16) since, not surprisingly, Steinbrenner explains "extension" (in connection with his characterization of systems) with the help of "denotation" (cf. Steinbrenner 1996, 89; 2005, 231).

All in all, it seems that, despite his own announcements, Steinbrenner is not so much trying to avoid circularity but to make sure that exemplification cannot be confused with denotation at least in theory (cf. Steinbrenner 2005, 231). For this purpose, he proposes to rely on a presupposed denotational system and the difference it introduces between labels and denoted objects, which in turn may serve as samples. But this would amount to characterize denotation and exemplification in terms of labels and samples instead of the other way around.[19]

4. Conclusion

Besides the last two accounts that not only fail but are also unnecessary as ways out, we have discussed four proposals to interpret the two conditions for exemplification Goodman formulates in (1). According to our proposal (E), "exemplification" is a basic notion, and hence undefined; the other alternatives suggest various definitions. (D) defines "exemplification" as denotation in both directions, (RD) as unqualified reference which presupposes denotation in the opposite direction and (R) takes "reference" as the only basic notion and defines "denotation" as unidirectional and "exemplification" as bidirectional reference.

---

[19] There is a similar tendency in (*LA,* 58). But at this point Goodman is not dealing with the semantical conditions for denotation and exemplification, but with the epistemological criteria that might be used for deciding in specific cases whether a given reference is an instance of denotation or of exemplification.

We argued that (D) as well as (R) fail from both an exegetic and a systematic point of view. Goodman unambiguously makes clear, firstly, that exemplification is a nondenotational form of reference and consequently cannot be defined in terms of denotation as in (D); and, secondly, that denotation in the same direction is not a necessary condition for exemplification as (D) and (R) imply. Both suggestions are also materially inadequate since they face all three of Dempster's objections. They obscure the distinction between exemplification and denotation and consequently between samples and labels, they turn exemplification into a symmetric relationship (which is implausible) and they are subject to countless counter-examples. For (RD) the situation seems, at least exegetically, to be less clear. There are frequent formulations in Goodman's writings stating that exemplification *is* being denoted (or possession) plus reference; if those statements could be read as definitions, (RD) would result directly. But Goodman never characterizes any of his formulations as a definition of "exemplification" and there is not even a passage which makes us expect such a definition. From an exegetic point of view, (RD) is therefore at least not inevitable. And it fails from a systematic point of view since it is subject to the same counter-examples as (D) and (R). Hence, only proposal (E) remains. The price for accepting (E) is that Goodman's theorems (b) and (c) have to be weakened and (1.1) has to be replaced by (1.2) or (1.3). Consequently, in statements that claim that exemplification is being denoted (or possession) plus reference either "reference" has to be interpreted as exemplification or the "is" as a conditional. Elgin does the second in her "official" introduction of the term and nowhere formulates Goodman's problematic theorems. Thus, while our suggestion (E) is at least defensible as an interpretation of Goodman, it is highly plausible as an interpretation of Elgin. The price for accepting (E) is well worth paying since (E) is the only proposal not challenged by the counter-examples. It is thus the only materially adequate reconstruction of Goodman's notion of exemplification.

References

Beardsley, M. (1975): “Semiotic Aesthetics and Aesthetic Education”, in: *Journal of Aesthetic Education* 9 (1975), 5–26.

Beardsley, M. (1978): “‘Languages of Art’ and Art Criticism”, in: *Erkenntnis* 12 (1978), 95–118.

Carnap, R. (1971) [1950/62]: *Logical Foundations of Probability,* London: University of Chicago Press; Routledge and Kegan Paul, 1971.

Coldron, J. (1982): “Peltz on Goodman on Exemplification”, in: *Journal of Aesthetic Education* 16 (1982), 87–93.

Dempster, D. (1989): “Exemplification and the Cognitive Value of Art”, in: *Philosophy and Phenomenological Research* 49 (1989), 393–412.

Elgin, C. Z. (1983): *With Reference to Reference*, Indianapolis; Cambridge: Hackett, 1983.

Elgin, C. Z.(1993): “Relocating Aesthetics: Goodman’s Epistemic Turn”, in: *Revue Internationale de Philosophie* 46 (1993), 171–186. Reprinted in Elgin, C. Z.: *Between the Absolute and the Arbitrary*, Ithaca; London: Cornell University Press, 1997, 63–80; and in Elgin, C. Z. (ed.): *Nelson Goodman’s Philosophy of Art*, New York: Garland, 1997, (= *The Philosophy of Nelson Goodman. Selected Essays*. Vol. 3), 1–17.

Elgin, C. Z.(1995): “Metaphor and Reference”, in: Radman, Z. (ed.): *From a Metaphorical Point of View. A Multidisciplinary Approach to the Cognitive Content of Metaphor,* Berlin; New York: de Gruyter, 1995, 53–72.

Elgin, C. Z.(1996): *Considered Judgment*, Princeton; New Jersey: Princeton University Press, 1996.

Elgin, C. Z.(ed.) (1997a): *Nelson Goodman's Theory of Symbols and Its Applications,* New York: Garland, 1997 (= *The Philosophy of Nelson Goodman. Selected Essays.* Vol. 4).

Elgin, C. Z.(1997b): “The Power of Parsimony”, in: *Philosophia Scientiae* 2 (1997), 89–104.

Goodman, N. (1975): [Letter to Monroe Beardsley], in: *Journal of Aesthetic Education* 9 (1975), 25–26.

Heydrich, W. (1993): “A Reconception of Meaning”, in: *Synthese* 95 (1993), 77–94 (Reprinted in: Elgin 1997a, 155–172).

Martin, R. (1981): “On Some Aesthetic Relations”, in: *Journal of Aesthetics and Art Criticism* 39 (1981), 258–264.

Peltz, R. (1972): “Nelson Goodman on Picturing, Describing, and Exemplifying”, in: *Journal of Aesthetic Education* 6 (1972), 71–86.

Steinbrenner, J. (1996): *Kognitivismus in der Ästhetik*, Würzburg: Königshausen & Neumann, 1996.

Steinbrenner, J. (2005): “Exemplifikation und Bezugnahmefeld”, in: Steinbrenner, J.; Scholz, O., R.; Ernst, G. (eds): *Symbole, Systeme, Welten. Studien zur Philosophie Nelson Goodmans*, Heidelberg: Synchron, 2005, 227–234.

Textor, M. (2008): "Samples as Symbols”, in: *Ratio* 21 (2008), 344–359.

Jakob Steinbrenner

Art-Samples.
On the Connection between Art and Science

Abstract
I want to illustrate the closeness of art and science by drawing on the example of the scientific practice of sampling. I want to show that Goodman's idea of seeing works of art as samples may lead to a deeper understanding of art, and maybe also of science. By way of doing so I will first explain the function of samples in science and in a second step show how typical features of samples in science may be applied to pictures, here especially those of abstract art.

> Painting is a science and should be pursued as an inquiry … of which pictures are but the experiments (Constable).

Here I want to follow up on the idea of understanding works of art, especially pictures, analogously to samples in science (e.g. spot checks and also experiments). To my knowledge, this idea was first explicitly stated by Nelson Goodman. Of course, I do not wish to deny that the comparison and, by extension, an approximation of artistic and scientific endeavours is as old as art, respectively science itself (cf. Gombrich 1977, I.1). Furthermore, if one wishes to do so, one can see the whole enterprise of "philosophical disenfranchisement" (cf. Danto 1986, chap. I) as an attempt to deny the equality of art and science respectively philosophy. Also one should take note of the fact that, for most artists too, art is not to be understood as science.

Despite these internal attempts at distinction as well as external attempts at disenfranchisement, I want to followingly illustrate the closeness of art and science by drawing on the example of the scientific practice of sampling. I want to show that Goodman's idea of seeing works of art as samples may lead to a deeper understanding of art, and maybe also of science. By way of doing so I will first explain the function of samples in

science and in a second step show how typical features of samples in science may be applied to pictures, here especially those of abstract art.

1. Samples in Science

Certainly no one doubts that samples play an important role in science. Consider, for example, samples in medicine, biology, chemistry, criminology, meteorology etc. Samples in these sciences are typically characterized by the act of taking a part of a larger whole and analyzing it. We have a different kind of sample, for instance, if a scientist examines something new and artificially developed in an experiment. In this case the sample is not part of the whole in the way a water sample is part of a lake, but is, at best, part of a future whole. This is the case, for instance, if a pharmacist attempts to develop a new drug. A further distinction in the practice of sampling is whether the whole, from which the sample is taken, is exhaustible or inexhaustible. A water sample taken from a bottle of mineral water is, for example, exhaustible, while a sample from the sea is practically inexhaustible.

Despite the differences among kinds of samples, one common feature is easily identified. Samples depend on a frame, i.e. a theory or a system, within which the relevant features of a sample are defined. To clarify these so far cursory remarks two examples can be adduced: Let's say I decide to lay in a stock of potatoes for the winter and, for this reason, go to a farmer in order to get a sample. The potatoes are tasty and have all features of good potatoes. I buy two hundredweights, but unfortunately, in the course of the winter, find out that many are worm-eaten and bitter. The sample taken was bad. It was bad because the essential features of the taken sample did not correspond to those of the potatoes eaten in the winter. We should note the following: The sample is bad, but not wrong. It is bad because it has not got the essential features that would be representative of the respective potato harvest. [1]

[1] A sample can, of course, also be simply false, for instance, if a red fabric pattern is to exemplify green fabric. It is, in this case, just as false as if "red" were applied to a green object (cf. *WW*, 133). Therefore a distinction has to be made whether something is not a sample of a feature or just not a "fair" one (ibid., 134).

An interesting question in this context is which conclusions can be drawn from this failure for the practice of taking samples in the future. It is obvious that, for practical reasons, I can impossibly sample all potatoes that I intend to eat in the coming winter. On the other hand, what I can do is take a number of samples in order to increase the likelihood that the sample is representative. I can furthermore try to improve my practice of sampling by paying attention to different features. I can, for example, develop a "theory" of the way the colour, the consistency etc. of potatoes give indications as to their shelf life and taste. The consequence for future samples of modifying my "theory" like this is that features will be examined that were so far considered irrelevant. That is, I examine whether the sample has or has not got certain features that I have so far not examined. Modifications of this sort are likely to be found in the sampling practices of all sciences and are often the basis of scientific progress. While, in the case of my potato example, the practice of sampling is exhausted when all potatoes bought are used up and I thus know whether the sample was representative or not, in the case of most scientific samples, the samples are inexhaustible. Consider the following example. On behalf of the waterworks a hydrologist takes a sample from Lake Constance. She wants to find out whether the water in Lake Constance is potable. The quality of the water sample indicates that the water in Lake Constance is potable. A number of swimmers, who have accidentally drunk water from Lake Constance, lose their trust in science, when they fall ill shortly afterwards. Was the sample bad?

Whereas in the first example (the potato sample) the question can clearly be answered in the positive, the case is slightly more complicated in the second. The water sample had properties other than those of the water, in which the swimmers have swum. The fundamental difference between the two samples described above is that, in the first case, there is ultimately a fact of the matter whether the sample is good or bad when the hypothesis is exhausted, i.e. when all potatoes are eaten and they correspond to the sample in all their essential features. This is in fact impossible in the second case. It will never be possible to finally decide whether a sample is good, since it will always be possible that further samples will be inconsistent with the samples on which induction has been based so far.

But how can the quality of the samples taken so far be assessed in view of a new sample?

By way of example, let us imagine that the hydrologists of the waterworks in Constance have so far taken 1000 samples according to a certain practice of taking samples. This practice is based on certain theoretical background assumptions, for instance, probability theory, fundamentals of water properties etc. and a more or less explicit way of handling the practice of sampling, for instance, demanding a standard of cleanliness for the containers used for sampling or defining areas of the lake, from which samples can and cannot be taken. Against this background a hydrologist takes a further sample, which is not potable. It is obvious that one or two samples hardly justify the assumption that Lake Constance no longer has the quality of drinking water nor that the practice of sampling should be called into question for this reason. The hydrologist has to take further samples and only if they are not potable with a certain frequency will she come to the conclusion that the water in Lake Constance is no longer potable. We can assume this to typically be the case if we think of the role of samples in science. The closest case, on the other hand, probably is one, in which the samples still have the quality of drinking water, but in which people who drink water from Lake Constance exhibit certain disease patterns with increased frequency. This will lead to the assumption of a causal link between drinking water from Lake Constance and these disease patterns. This can lead to changing the features, on the basis of which the quality of drinking water will be assessed. So effects of this kind do not only affect the practice of sampling but also the background theory. If we take a look at the historical development of the identification of which features water must possess in order to be potable, we find that in the beginning there were relatively few criteria, for instance, smell and clarity, and only gradually over time further requirements were made on the quality of drinking water and others were dropped. The considerations so far as regards the water sample may be summed up as follows.

In the example there are four possibilities: (a) the sample taken by the waterworks was merely accidentally bad. (b) The practice of sampling is revised. It is no longer sufficient to take one sample from 20 meters below the surface and a kilometre from the shore. Instead samples have to be

taken from many places and at varying depths. (c) The methods of analysis can also be changed. In contrast to the potato sample it is also the case that (d) hypotheses, which are to be supported by samples similar to the water sample, are firstly never, in practice, exhaustible. Secondly it will always be possible that the hypotheses are replaced by better-founded hypotheses or entire theories. Thirdly it cannot be ruled out that a negative case is on hand.

Against the backdrop of points (a) to (d) we can now give a first answer as to which criteria may be used as a standard according to which the quality of water samples, respectively samples in general, may be assessed. Samples have to be projectible[2], i.e. they have to have the features given in (d).

What this amounts to is that a sample is good only if further samples can be taken from the respective material (i.e. the material is not exhausted and we can therefore take further samples), if we can use the sample to support alternative hypotheses, the predicates of which are more firmly entrenched. According to Goodman we can take better entrenchment to be the following: A predicate P1 is more firmly entrenched than a predicate P2 if P1 or another predicate with the same extension has actually been projected more often than P2. Further constraints are that the predicates in question not change their extension and that all inscriptions of a predicate have the same entrenchment. An example of the second case can be seen in the following hypothesis:

H: Overweight persons have a disadvantage when cycling.

This hypothesis may be true today, but 1947 in post-war Germany it was false. The reason for this is, for one thing, that in 1947 persons would have been considered overweight that today would be considered normal weight, for another, that many thin persons suffered from physical exhaustion and could not have competeted with overweight persons when cycling.

A crucial point in Goodman's considerations to be observed here is that the choice of hypotheses (and accordingly theories or systems) depends not

[2] At this point it becomes clear that the concept of projection in Goodman's sense precedes that of rightness (cf. Chokr 1993, 64).

only on truth, but on the historicity of linguistic practice (cf. Ernst 2005, especially p.102). I.e. when we examine the entrenchment of predicates, we actually examine languages and our past linguistic practices. Not alone truth guides us, but our use of language.

> [B]ut the organisation we point is effected by the use of language and is not attributed to anything inevitable or immutable in the nature of the human cognition (*FFF***,** 96).

The concept of entrenchment, one the one hand, illustrates Goodman's early interest in and appraisal of a pragmatic theory of linguistic practice and, on the other hand, provides a first criterion for, or at least a symptom of, the choice of (competing) categorial systems. It should furthermore be noted that the concept of entrenchment therefore is a gradual one.

A further expression of central importance for the following considerations, next to entrenchment, is that of projection. According to Goodman a hypothesis is "projectible when and only when it is supported, unviolated, and unexhausted, and all such hypotheses that conflicted with it are overridden" (*FFF***,** 108). Competing hypotheses are displaced, because they are not as well supported (i.e. there are few actual cases of application) or simply false. [3]

As has already been suggested, projectibility plays a crucial role in the choice of hypotheses that are supported by samples. Goodman himself does not want to restrict it to hypotheses, but sees its application wherever theories, categorial systems or symbol systems are used.[4] This, for example, is not done for the sole purpose of predicting events, as in the case of the hypothesis, but also for the purpose of organisation or, in the case of metaphor, the re-organisation of the most diverse of domains. How are we to understand this?

Not only in science is it common practice to demarcate domains by the use of predicates (e.g. fish, fowl, mammals). To what extent such a

---

[3] Following this definition Goodman defines the term „*non-projectible* [as follows:] when and only when it and a conflicting hypothesis are supported, unviolated, unexhausted, and not overridden; and *unprobjectible* when and only when it is unsupported, violated, exhausted, or overridden" (*FFF***,** 108).

[4] What is important is that projection is not restricted to language codified in words, but comprises all other symbol systems (cf. *LA*, 201; *WW* 128, 135).

categorisation displaces others (e.g. no longer counting whales as a kind of fish) is also a matter of projectibility. We will prefer one categorial system to another, because it is more easily projectible, i.e. because it is better supported, does not lead to contradictions, is open to new applications and has displaced competing systems.

In applying to works of art the whole breadth of the considerations made so far it is necessary to point out a further idea of Goodman's, namely that works of art mostly refer metaphorically. In what follows we will therefore give a brief sketch of his theory of metaphor and will examine whether it is also subject to his theory of projectibility.

1.1 Metaphors in the Sciences

Goodman's theory of metaphor is a further development of the Interaction Theory of Metaphor as put forward by I.A. Richard (1936) and M. Black (1954). The central characteristics of Goodman's theory are the following:
First, metaphors cannot be less true or false than literal statements. [5]
Second, metaphors are pervasive throughout language and especially so in art.

Third, metaphors arise when a schema of labels (for Goodman, linguistic as well as non-linguistic symbols) for categorising a certain domain (sphere) is transferred to the categorisation of another domain.

For such a transfer to be successful three conditions must be met:

(i) A symbol **M** has to already denote or exemplify literally, [6] i.e. there has to be at least one schema **S** of labels and a sphere $\mathbf{G}_1$, to which the categorisation, i.e. the signs are applied in a denotative sense.

(ii) There has to be a further sphere $\mathbf{G}_2$ and

(iii) a transfer **T** of **S** to $\mathbf{G}_2$.

---

[5] It has to be stressed that Goodman's constraints are contrary to all those theories of metaphor, according to which metaphors are essentially evocative and therefore truth predicates cannot be applied to them.

[6] How Goodman understands these two expressions can be illustrated by way of the following example: The expression "Lake Constance" denotes Lake Constance and, in turn, it exemplifies the label „inland water". Material pictures and other symbols, too can denote or exemplify something.

If we apply conditions (i-iii) to the metaphorical expression "the lake is a sapphire", we get the following structure or chain of reference:[7]

1. "Sapphire" denotes a sapphire.
2. A sapphire exemplifies "smooth" and "blue"
3. "Smooth" and "blue" denote the lake.

This chain of reference thus clarifies the three constraints Goodman sets for metaphors:

First, the word "sapphire" literally denotes sapphires. This is accomplished by the fact that there exists a schema **S**, call it the schema for describing gems, that categorises a given sphere $\mathbf{G}_1$, namely gems, by the use of labels.

Second, there is a further sphere $\mathbf{G}_2$, the sphere of water bodies.

Third, the schema $\mathbf{S}_1$, the schema for describing gems ("sapphire", "smooth", "blue" etc.), is transferred to the sphere $\mathbf{G}_2$, the sphere of water bodies.

But to what extent are metaphors projectible? It is obvious that a new metaphor cannot be supported in the same way as a hypothesis. In the event of the first use of a metaphor there are no literally true cases of its previous application. They are unsupported and refuted in a literal sense, because they are literally false.[8] Nonetheless, metaphors can be assessed as regards their projectibility. A metaphor is more useful than its alternatives if transferring it to a new sphere makes possible more correct applications and, by extension, more new insights than would have been made possible by alternative metaphors. For example, up to this day one can say that the metaphor "man is a wolf" conveys more insights than the metaphor "man is a rowing boat". For both metaphors it is the case that they have not been exhausted. However, the first metaphor is certainly better supported. Furthermore it is the case that a metaphor can be assessed according to

[7] The expression "Chain of reference" was coined by Goodman (*MM*, chap. 3.1). It is important here to stress that Goodman does not hold that such chains of reference can be paraphrased, but that they serve to illuminate the structure of metaphorical reference (but also that, for instance, of allusions.).

[8] Possibly metaphors of the kind "This is a stony path" form an exception. Both can be true, that the path actually is stony and that it is hard to walk on, wearisome etc.

how many different domains it may reasonably be transferred to. In this sense the first metaphor may certainly better be applied to the sphere of politics than to the domain of flower-lovers.

1.2 True and other Samples

For the following considerations it is of importance to once again point out that the concept of truth is contained within the concept of projectibility as it has been characterized so far. Only hypotheses, for which there are *positive cases*, are projectible. "Positive Cases" here means that there were cases or events, to which the hypothesis (a simple or complex sentence) applied, i.e. was, respectively is true. In cases of schemata not coded in the form of words the notion of truth is absent. For example, systems of musical notation or pictorial systems (e.g. cubist painting) are neither true nor false. For cases of this kind, the question arises whether there is something that takes the place of truth and therefore makes it possible to still talk of projection in these cases. For the cognitivist this question is of key importance, as she views art and science as cognitively equivalent (cf. Steinbrenner 1996, chap. I). In answering this question Goodman employs the following strategy. He replaces the central concept of truth by that of rightness. While rightness plays an important role in all symbol systems, truth is only of relevance for certain linguistic systems coded in words. No less holds true for the languages of art than for the languages of science: not truth, but rightness is their actual goal.[9] For the practice of sampling this entails that only projectible samples fit and therefore are right.

> Such fairness or projectibility, rather than requiring or guaranteeing agreement [being true J.St.] between the projection made and an actual feature of the whole or of further samples, depends upon conformity to good practice in interpreting samples (*WW*, 135).

Samples of the kind of the potato sample have as their criterion agreement. A sample from the stock of potatoes for the winter is good and

[9] Here the starting point of Goodman's line of thought is the problem of induction, which can be seen as the paradigm of his claim that we cannot obtain a solution to the problem of induction by the concept of truth alone. (cf. *FFF*). Projectibility has to also be taken into consideration and it has to do primarily with rightness (cf. Steinbrenner 1996, 2.72).

right if it by and large agrees, as regards taste and consistency, with the entirety of potatoes eaten throughout the winter. Such agreement can in practice never be achieved for the water sample from Lake Constance.

A key feature in which samples differ from hypotheses is that samples are mostly[10] not linguistically coded in the form of words. What samples exemplify is always contingent upon a practice. As has been shown, it is the case that if samples do not support the hypotheses, for example, because they are false in most cases, the hypothesis can be replaced or the practice of sampling can be changed (cf. *WW*, 136)[11]. The formerly good practice becomes a bad one and is replaced by another practice if it repeatedly leads to unwanted results. The quality of a sampling practice can be assessed through the projectibility of its results. Better entrenchment, in such cases, is the hallmark of a handed-down and successful sampling practice. If two samples support two conflicting hypotheses, we will reject the sample that is less well entrenched (cf. the previous Goodman quote).

Before I followingly turn to art, let me briefly address a term that I have already used repeatedly and that is of central importance for the following considerations, namely exemplification. The reason for this is, in my opinion, that samples not only support hypotheses, but also exemplify them if they are right. In this sense they are similar to works of art that, as samples, can support as well as exemplify hypotheses and interpretations. For example, a picture by Monet can exemplify “impressionist painting” and, at the same time, be used as a sample to support certain hypotheses concerning impressionism. What matters in both cases is that successful exemplification as well as right samples are subject to the criterion of projectibility. This entails that the properties, predicates or interpretations that are exemplified by a work of art are ones, which, in contrast to their alternatives, are more easily projectible. I will address this issue in more detail in what follows.

---

[10] Exceptions are linguistic and similar investigations.

[11] The difference in later writings as against Goodman 1954 is that Goodman no longer is a follower of Popper. This means that a falsification ("negative cases") of a hypothesis does not necessarily lead to abandoning the hypothesis, but possibly to changing the sampling practice.

1.3 Exemplification

First, we should note that samples, respectively exemplifying objects do not exemplify all theories, predicates etc., to which they are subject. Goodman's example here is a swatch that a tailor sends to his customer. This sample normally exemplifies colour and quality of the fabric, but not its shape. On the other hand, the paper-mâché cake in the confectioner's shop window exemplifies the shape and colour, but not the taste. To sum up, what is exemplified by a sample depends on the context, in which the sample is used. In this sense "[e]xemplification is possession plus reference" (*LA*, 53). I.e. depending on the context certain properties that the sample has are exemplified by the sample.

Within Goodman's nominalist framework this means that the sample exemplifies certain predicates, labels, hypotheses etc. Goodman writes:

> Exemplification relates the symbol to a label that denotes it, and hence indirectly to the things (including the symbol itself) in the range of the label (*LA*, 92).

As we have seen, however, a sample cannot only literally exemplify, i.e. properties that the sample actually has, but also metaphorically. Concerning this matter, Goodman writes:

> An object is literally or metaphorically denoted by a predicate, and refers to that predicate or the corresponding property, may be said to exemplify that predicate or property (*LA*, 52).

Exemplification, according to Goodman, is the most important symptom of the aesthetic and therefore plays a pivotal role in his theory of art: The reasons for this are as follows:

First, exemplification is to explain reference-in-fiction.

Second, it is to explain how abstract art functions.

Third, Goodman's theory of expression is based on exemplification.

Fourth, it is to explain the similarities of the modes of operation in works of art and samples in the sciences.

I will give a few short examples for the first three points before I go on to address the fourth point in detail. The latter point is the one of crucial importance for my topic here.

Reference-in-fiction

Turning to fictional signs, namely pictures or texts that are depictions or descriptions of nonexistent objects, we may ask what it means to understand such symbols. As objects such as, for example, unicorns do not exist, understanding them cannot be knowing their denotation. Goodman's solution to this challenge is that understanding, in these cases, is recognizing to which type the symbol belongs and under which descriptions it falls, respectively in which descriptions it can appear. For the picture of a unicorn this means that the picture is of the type *picture of a unicorn* and falls under the description "fabulous animal with a horse's body and single straight horn". The expression "unicorn", on the other hand, is of the type "description of a unicorn" and falls under the description "expression for fabulous animal with a horse's body and single straight horn". It can, for example, occur in the following descriptions: "Unicorns are fabulous animals with highly sexual connotations" or "the unicorn is creeping up on us." Understanding fictional expressions, then, is knowing which linguistic types and descriptions are exemplified by the fictional symbol. In this sense a picture of a unicorn neither exemplifies the type "picture of a mermaid" nor the description "picture of a fabulous animal that lives under water." Knowing what fictional signs mean is knowing what they exemplify. If a poet writes a new fable about a unicorn, we can see the fable as a sample, which is based on our knowledge of unicorns and the quality of which can be assessed by the degree to which it can be metaphorically projected onto certain aspects of life.

Abstract painting

Abstract paintings are similar to fictional paintings in a certain sense. They, too, do not denote. Like fictional paintings, they do not represent existing objects. Understanding abstract paintings does, however, depend on whether the observer knows how to properly categorize them i.e. whether she knows under which descriptions they fall. If, for instance, an observer believes to be seeing a depiction of a chess board in an abstract painting by Ellsworth Kelly, she is mistaken. If she recognizes that the picture exemplifies the label "checkerboard pattern", this can, however, be of advantage for a better understanding of the picture. She will, in doing

so, for example, recognize how Kelly moved from abstraction to non-figurative painting and how his pictures become sheer lessons in perception. That is, the observer might recognize that Kelly's pictures have features similar to perception experiments (samples!), but go beyond this comparison, as they exemplify certain aspects of the question "What is a picture?" (e.g. "Can a mere pattern be a picture?") Knowing that a picture exemplifies such questions, of course, presupposes knowledge of a sampling practice. (Namely, that of the visual arts and more specifically that of abstract painting.) [12]

Artistic Expression

If we assume that phrases like "the picture expresses joy" are meaningful, we can, of course, also ask how it is possible that an inanimate object like, for example, a picture can express feelings or anything else. Usually we say only of people that they can express feelings.

Goodman's solution here is that while persons can literally express feelings, symbols can only exemplify them metaphorically. In this sense expression, for Goodman, is a special kind of metaphorical exemplification. As far as I know, Goodman does not provide a criterion how expression is to be distinguished from metaphorical exemplification. My suggestion therefore is to restrict the use of the expression "expression" to cases, in which feeling predicates are metaphorically exemplified (cf. Steinbrenner 1996, II. 1.33.2). In this sense a colourful painting, for example by Delaunay, can, at the same time, express joy and metaphorically exemplify its owner's wealth. In a nutshell: A symbol can only express feelings, respectively, expressions that express feelings, whereas it can metaphorically exemplify just about any predicate imaginable.

---

[12] Cf. (Hölscher 2005). Christoph Baumberger pointed out to me that it is questionable whether a picture can exemplify a question. Two solutions seem possible: Either one firstly accepts that pictures can exemplify questions (not, for example, hypotheses, sentences or predicates) or secondly one tries to paraphrase the questions as ordinary declamatory sentences. A work by Robert Rymans, for example, might then exemplify the hypothesis "Alberti's thesis that pictures have properties P1 - Pn is questionable today" instead of "Is Alberti's theory of pictures still valid today?".

Without having here addressed all subtleties concerning Goodman's understanding of exemplification (cf. Steinbrenner 2005), my hopes are high that it has transpired how important this concept is in the context of his theory of art. What needs to be further examined, however, is the fourth point, namely that artistic exemplification is parallel to the scientific practice of sampling in its crucial features.

2. Works of Art as Samples

If we follow Goodman in assuming that it is not the truth that is the guideline in formulating scientific hypotheses, but rightness and, by extension, projectibility, a former criterion for categorially distinguishing art and science has become obsolete (cf. Steinbrenner 1996, II.2). According to Goodman it is the common goal of scientific samples as well as works of art to provide new insights, that is, essentially to construct new categorial systems, which make us aware of useful or new connections among and similarities between different systems. As we have seen in the case of the water sample, it is precisely the goal to deduce certain features (e.g. potability) from other features. The case is similar for works of art.

An art historian, for example, is interested in finding out the common features of Rembrandt's works and from these deducing the general stylistic features of his oeuvre. To achieve this it is necessary to compare the works with one another, to form groups and, if necessary, to develop new predicates, with which it is possible to form new hypotheses concerning Rembrandt's oeuvre. In the case of Rembrandt this has recently meant that his oeuvre became smaller and several of the works so far attributed to him were attributed to his disciples. This, in turn, has resulted in his disciples being credited with more artistic significance (cf. Wiegand 2005). Aside from such stylistic issues that can concern individual artists, epochs or even more general considerations, it is, in the case of abstract art, crucial to determine what works exemplify. In order to do so it is often necessary to pay attention to features that to that point might have only been marginally acknowledged.

An example of such a feature is the frame, respectively the outer delimiting of a picture. Robert Ryman, for example in his works examined the boundary between the picture and the wall and, more specifically, the

question of how a picture is affixed to the wall. Questions of this kind, of course, could only be asked after the picture as a medium had been fundamentally called into question. That I today can so naturally speak of such matters presupposes a "sampling practice", part of which it is that several pictures exemplify such features concerning affixation and the frame. An observer unaware of such questions will mostly be at a loss when looking at abstract paintings, since she does not know the relevant features. She might, for example, look for objects in the painting and discard Ryman's works as boring snowscapes. A sampling practice of the kind illustrated above does not only have an influence on the observation of abstract painting, but furthermore has a retrospective effect in the sense that we re-examine how painters have traditionally used frames. Ideally it follows from such a practice that we develop new stylistic features concerning the frame, which give us a deeper insight into art or even life, for that matter. We achieve this feat of cognition through works of art that may well be expressions of exactly those kinds of considerations on the part of their creators. This does not, however, always have to be the case.

In many cases features are only recognized in retrospect. Take, for example, the hairstyles in films from the 60ies. We can, for example, in retrospect detect typical colours in the hair and the film as a whole, which the contemporary stylists could hardly have been aware of, since they did not know the alternatives that are available today (cf. Steinbrenner 1998). The same holds for editing, colour quality etc., that is, features to which we would be more inclined to attribute artistic relevance.

But let us return to the example of Ryman. Not least do his works have their significance, because they allow us to "project" well founded predicates into them, that is, let them exemplify these predicates. As we have seen, this is possible, because they are predicates that denote features characteristic of Modernity. This can, for example, be seen in the fact that these features, at least subconsciously, already played an important role with the predecessors of Modernity (e.g. Manet, Degas and Monet).

Put another way, just as in the case of samples, we look to the hypotheses that we want to confirm through them, in the case of works of art, we often look for stylistic features of an epoch, an oeuvre etc. that we want to see exemplified by the work (cf. Ernst 2005, 105). For this reason it also becomes clear why we are more interested in a single work by a

great artist than in a work similar in quality by an artist, who only achieved artistic greatness in this one work. We appreciate the work of the great artist more, because it allows us to locate it in a significant context. In Goodman's terminology we can explain it thus: The work of the great artist exemplifies more better entrenched predicates than the work of his colleague. The reason for this is that the oeuvre of the great artist, to which the work belongs, provides more possibilities for exemplification than the oeuvre of his colleague.[13]

So far I have mainly focussed on connections internal to art and similarities that can be exemplified by a work of art (cf. Hölscher 2005). Nonetheless it should have become clear that such a restriction is hardly necessary. Works of art can, of course, also bear witness to general human themes and their times. Just think of great novels, classical tragedies, that is, any masterpiece imaginable. It is precisely these that are our "samples" or examples of what man is, what he should do and where he wants to go. These works exemplify the leading properties (values) of a culture. At the same time, however, the way of historically dealing with these masterpieces bears witness to the way in which our sampling practice changes. Raffael's Madonna, for instance, does not essentially exemplify "naive piety" for us, but depending on context, quite a number of different things. And, of course, our canon of masterpieces is also subject to continuous revision. This runs in parallel with a change of the features that are important in our everyday lives. One need only think of the re-evaluation that women writers have undergone in the wake of the feminist movement.

Goodman's significant contribution, in my opinion, lies in the fact that he has developed a theory of scientific sampling practice that we can apply to our way of dealing with art. In this sense works of art are comparable to water samples from an inexhaustible body of water. That is, works of art

[13] Interestingly and not accidentally, Goodman is here close to Hume's inductive theory of art. For Hume too, those works are great that again and again are used in cultivating taste. But how do we know that these works are really great? If our answer to this question is that, in the course of history, we continually use them as evaluative standards, the spectre of circularity looms. (cf. Levinson 2002).

can, depending on context, exemplify very different things and are, like the water sample, taken from an inexhaustible whole.[14]

Translated by R. Owen Hughes (Müllan)

References

Black, M. (1954): "Metaphor", in: *Proceedings of the Aristotelian Society* 55 (1954), 273-294.

Chokr, N. (1993): "Nelson Goodman on Truth, Relativism, and Criteria of Rightness", in: *Dialectica* 47 (1993), 55-73.

Danto, A. (1986): *The Philosophical Disenfranchisement of Art*, New York: Columbia University Press, 1986.

Ernst, G. (2005): "Induktion, Exemplifikation und Welterzeugung", in: Steinbrenner, J.; Scholz, O.R.; Ernst, G (eds): *Symbole, System, Welten. Studien zur Philosophie Nelson Goodmans*. Heidelberg: Synchron, 2005, 99-110.

Gombrich, E. (1977): *Art and Illusion*, Oxford: Phaidon, 1977.

Hölscher, T. (2005): "Goodman und die Kunstgeschichte", in: Steinbrenner, J.; Scholz, O.R.; Ernst, G (eds): *Symbole, System, Welten. Studien zur Philosophie Nelson Goodmans*. Heidelberg: Synchron, 2005, 89-98.

Levinson, J. (2002): "Hume's Standard of Taste: The Real Problem", in: *Journal of Aesthetics and Art Crtitcism* 60 (2002). 227-238.

Richard, I. A. (1936): *The Philosophy of Rhetoric*. New York; London: Oxford University Press, 1936.

Steinbrenner, J. (1996): *Kognitivismus in der Ästhetik*. Würzburg: Königshausen & Neumann, 1996.

Steinbrenner, J. (1998): "The Unimaginable", in: *Theory & History: Danto and his Critics: Art History, Historiography and After the End of Art*: Theme Issue 37 (1998), 115-126.

Steinbrenner, J. (2005): "Exemplifikation und Bezugnahmefeld", in: Steinbrenner, J.; Scholz, O.R.; Ernst, G (eds): *Symbole, System, Welten.*

[14] For thorough reading of and critical comments on this essay I wish to express my gratitude to Christoph Baumberger and Gerhard Ernst.

Studien zur Philosophie Nelson Goodmans. Heidelberg: Synchron, 2005, 227-234.

Wiegand, W. (2005): “Als würde mitten im Barock ein Fenster aufgestoßen”, in: *Frankfurter Allgemeine Zeitung* (16.05.05).

Remei Capdevila Werning

Nelson Goodman's Autographic-Allographic Distinction in Architecture: The Case of Mies van der Rohe's Barcelona Pavilion

Abstract
In *Languages of Art*, Nelson Goodman proposes the categories of the autographic and the allographic to philosophically distinguish between an original work and its copy, reproduction, or forgery. However clear these distinctions might be in cases such as painting or music, difficulties arise when dealing with architecture, for our judgments regarding the identity of an architectural work cannot be indisputably classified according to Goodman's categories. Thus, first, I will discuss the scope of Goodman's autographic and allographic criteria to identify an architectural work. Second, I will examine the specific case of Mies van der Rohe's Barcelona Pavilion, first built in 1929 and reconstructed in 1986, to show how the allographic and the autographic are inextricably linked when establishing the identity of architectural works. Furthermore, I will argue that Goodman's statements on the identity status of architecture need to be reconsidered.

1. The Identity of Architectural Works: the Autographic and the Allographic

In the chapter entitled "Art and Authenticity" (*LA*, 99-123), Goodman develops the distinction between autographic and allographic arts based on the observation that some works of art cannot be faked. While it is possible to have a perfect copy of a painting, no forgery of a known musical piece can be made.[1] In Goodman's words:

> A work of art [is] *autographic* if and only if the distinction between original and forgery of it is significant; or better, if and only if even the most exact duplication of it does not thereby count as genuine. If a work of art is autographic, we may also call that art autographic. Thus painting is autographic, music nonautographic, or *allographic* (*LA*, 113).

[1] It is theoretically possible to forge an unknown musical work by using allographic processes, as it is argued in the following articles, based on Goodman's philosophy: (Kivy 2000, 2002) and (Pillow 2002).

Hence, the autographic identifies a category of works that cannot be replicated, i.e., every difference between a work and even its closest copy makes a difference to the work's identity. The history of production is the criterion for establishing the identity of an autographic work and what allows us to discriminate between the original and something else. The autographic does not mean singularity, for there are autographic and multiple works, such as engravings and cast sculptures. That is to say, an artistic discipline is autographic if the identification of the works in this discipline necessarily depends on the history of production; this is not related to singularity or multiplicity.[2]

By contrast, the allographic identifies a category of works that can be replicated, i.e., the difference between an original and its duplication is irrelevant. The criteria for establishing the identity of allographic works are a notational scheme or a notational system, which allow us to affirm that pieces sharing certain determinate and identifiable characteristics are the same work. This implies that allographic works are potentially multiple and reproducible; there can be several instances of the same allographic work. Although these works of art also have histories of production, this is not the criterion for defining the work as it is; music, literature, and performing arts are allographic, for, despite having a history of production, a score, a text, or a script are needed to identify a work as such (*LA,* 122; *MM*, 140).

According to Goodman, the primary logical function of a notational scheme and a notational system is to identify a work, even though it has other functions such as being a "practical aid to production" (*LA*, 128). In order to fulfill its allographic identificatory function, every notational system has to accomplish certain syntactic and semantic requirements, whereas a notational scheme must only fulfill the syntactic requirements (*LA*, 130-141; 148-154). In general, music, literature, and dance are allographic arts, although there are some exceptions: music with no score, such as jazz improvisations, is autographic; however, since it is possible to actually create a notational system that fixes their identities, they can be taken as allographic too. In opposition, painting, etchings, and sculpture

[2] As Goodman affirms: "Concerning the distinction between autographic and allographic arts or works, the first point is that this distinction does not coincide with that between singular and multiple arts; for some multiple arts, such as etching, are autographic" (*MM*, 140).

are usually autographic.[3] And architecture, according to Goodman, can be classified under both categories:[4]

> We are not as comfortable about identifying an architectural work with a design rather than a building as we are about identifying a musical work with a composition rather than a performance. In that architecture has a reasonably appropriate notational system and that some of its works are unmistakably allographic, the art is allographic. But insofar as its notational language has not yet acquired full authority to divorce identity of work in all cases from particular production, architecture is a mixed and transitional case (*LA*, 221).

In this passage, Goodman comments on the hybrid status of architecture as an artistic discipline that is both autographic and allographic, since both criteria for establishing the identity of a work are valid. Some buildings are identified by the history of their productions, while others are identified by means of a notation; moreover, a single building can be judged according to both criteria at the same time. Note that the autographic and the allographic help us clarify and classify our judgments regarding the identity of a work of art, i.e., these criteria are a reflection of our practice of distinguishing the constitutive elements of a work. If we consider a building to be only autographic, it is because important constitutive elements of the building's identity would be lost if considered from the allographic perspective.

This in-between position of architecture is not only synchronic but also diachronic: defining architecture as a "mixed case" means that the autographic and the allographic criteria coexist simultaneously; defining it as a "transitional case" means that the allographic criterion to establish the identity of an architectural work succeeds the autographic, and thus

---

[3] Some exceptions can also be found in the autographic arts. Kirk Pillow discusses the case of Sol LeWitt, and argues that that Goodman's autographic-allographic distinction does not apply for some of LeWitt's paintings (Pillow, 2003), and Michel Weemans also comments on LeWitt as well as on Claude Rutault's and Lawrence Weiner's works to argue that the allographic criterion is valid in those cases and, hence, painting can be also allographic (Weemans, 2001).

[4] I will take for granted that architecture is or can be considered an art, although not every building is a work of art taking into account Goodman's statement according to which "[a] building is a work of art insofar as it signifies, means, refers, symbolizes in some way" (*R*, 33). Nevertheless, since also according to Goodman it is possible to make a "work work," i.e., that almost any object can be considered from an aesthetic point of view, then it seems that it is, at least theoretically, always possible to consider a building aesthetically (see: *MM*, 174-187.)

presupposes an inherent historical process. This is precisely what Goodman assumes when he states that architecture will become completely allographic once its notational system has "acquired full authority to divorce identity of work in all cases from particular production" (*LA*, 221). In other words, this affirmation acknowledges that the decisive factor in determining whether architecture is autographic or allographic is the perfection of the correspondent notation, in this case, plans, elevations, and sections.

When discussing architectural plans, Goodman compares them with musical scores, which are the paradigmatic example of notational systems in art. In Languages of Art, Goodman provides two different characterizations of plans as notations: in the first one he considers plans as scores, i.e., as notational systems that fulfill both syntactic and semantic requirements (*LA*, 219); in the second one, he maintains that plans are as scripts, i.e., notational schemes that only fulfill syntactic requirements (*LA*, 219). The reason for this shift from system to scheme is that whereas it is possible to construct a building from a set of given plans, it is not possible to derive a unique plan from a given building. That is to say, whereas the semantic requirements are fulfilled in the former case, they are not in the latter. Architecture, then, would function in a way more similar to drama rather than to music, because from the work multiple notational schemes can be created.[5] In other aspects, however, architecture functions in a way closer to music. For instance, just as a book does constitute a literary work, plans and scores do not constitute architectural and musical works: they are only their notations. Moreover, plans and scores include integral parts that constitute the notation itself, and others that cannot be considered notational. The notations of architecture (e.g., plans, elevations, and sections) result from the combination of accurate drawings and measurements represented to scale, which is parallel to the musical scale system; the non-notational parts of plans are the written specifications or other non-codified indications regarding certain aspects of the building (construction materials, for instance), which are comparable to the verbal specifications of tempo in a score. In a certain sense, the non-notational aspects of plans over-define, as it were, a building insofar as they establish more properties apart from the ones that are strictly constitutive elements,

[5] To determine whether plans are notational systems or notational schemes requires a more thorough discussion that is far beyond of the purposes of this paper. In what follows I will talk about notations in general, without specifying whether they are systems or schemes.

i.e., that are fixed by a notation that fulfills the syntactic and semantic requirements (*LA*, 218-221).

Despite these similarities between architectural and musical notations, scores are much more codified than plans, and for this reason Goodman affirms that architecture does not share the entire characteristic features of a mature notation. While musical scores can resort to very precise codifications to establish how every single note has to be played, architectural plans lack a similar system to fix how each and every one of their parts has to be executed. Therefore, extranotational information is (still) required to explicitly state constructive elements such as the wall texture, or the exact color; these are usually exemplified by means of a sample: a piece of a wall with the correspondent definitive surface and color serves as a model for the rest of the building (*WW*, 63-70). Since for Goodman architecture will not become absolutely allographic until its notation is able to completely fix the identity of architectural works, it is fundamental to explore whether it is feasible for plans to develop into a perfect notation to include the features now expressed in an extranotational way (acknowledging, of course, that these characteristics are constitutive for the work's identity). This is exactly what Saul Fisher undertakes when he argues, based on Goodman's theory, that Computer Aided Design (CAD) theoretically allows the complete specification of the elements that are part of the architect's extranotational language by means of codification (Fisher, 2000).[6] In this way, identification of architectural works could be independent of any particular building or, in other words, with the progressive transformation of plans from analogical to digital, the

[6] In this article, Fisher not only argues that through CAD it is possible to develop a notational system that fixes the constitutive elements of a building, but also the "expressive" or aesthetic ones. I disagree with this statement, because the role of notational systems is only to identify a work as such and not to uniquely fix its aesthetic properties. First, Fisher should clearly define what an aesthetic property is. Second, the inclusion of aesthetic properties in a notation would have as a consequence that two instances of a same work with different aesthetic properties could not be identified as the same work or with the same notation. As Goodman affirms: "Thus, even where the constitutive properties of a work are clearly distinguished by means of a notation, they cannot be identified with the aesthetic properties" (*LA*, 120). Apart from confusing constitutive and aesthetic properties, Fisher also mixes up the artist's intentionality and constitutive features of a work and this brings him to limit the range of Goodman's allographic criterion and, to a certain extent, to blur the distintion between the autographic and the allographic.

notational system can potentially become the only criterion to determine the identity of a building.

However, even though a perfect notation could be available, some architectural works resist being considered only from an allographic perspective. As Goodman affirms:

> [A]ll houses complying with the plans for Smith-Jones Split-Level #17 are equally instances of that architectural work. But in the case of an earlier architectural tribute to womanhood, the Taj Mahal, we may bridle at considering another building from the same plans and even on the same site to be an instance of the same work rather than a copy (*LA*, 220-221).

Although it would be possible to build a copy of the Taj Mahal, the criterion to identify this building is still the history of its construction and not a notation. It is important to note that the fact of not knowing exactly every detail of its history of production does not invalidate the possibility of establishing the work's identity autographically. Moreover, even if we did not possess any information about the history of production of a work, this would not invalidate the autographic criterion. This is the case of anonymous buildings, whose history of production is impossible to trace back and yet these buildings can only be autographically identified.

Apart from stressing that an exact reproduction of the Taj Mahal would only be a copy, the passage by Goodman just cited also implies that the difference between multiple allographic buildings and singular autographic architectural works is related to the distinction between what could be called high and low architecture.[7] While tract housing sprawls throughout suburban areas and this process has no effect on its identity, no masterpiece of architecture is reproduced without emphasizing that the subsequent instances are not the original but simply replicas (the Parthenon in Nashville's Centennial Park is simply a copy of the one in Athens).[8] It

[7] Note that it is also possible to have two or more identical, original, and autographic buildings. In that case, it could be considered that the architectural work is composed of two or more buildings, for a copy or a reproduction of this work would mean to construct all of them. This is the case, for instance, of Mies van der Rohe's two Lake Shore Drive Apartments in Chicago, or M.I. Pei three Silver Towers in New York.

[8] I use the term "replica" in the sense of reconstruction, copy, duplicate, or reproduction of a work of art, and not in Goodman's technical sense according to which a work is a replica of another if and only if they are instances of the same work and share the same syntactical properties. This implies that there are only replicas of allographic works that have the same notation (see: *LA*, 131-132).

thus seems that, in general, for buildings aesthetically less valuable, the identity criterion is the allographic, and, for the others, it is the autographic. In other words, there exists a dichotomy between potential reproducibility and actual limitation in the construction of architectural works.[9]

Architecture, thus, has a double nature; it is a "mixed and transitional case" not only because some buildings are considered autographic while at the same time others are taken to be allographic, but also because some of their parts are autographic and others allographic, as would be the case of autographic buildings containing an allographic inscription. Yet architecture can be considered partly autographic and partly allographic from another point of view. According to Goodman, arts can be classified into one-stage and two-stage arts (*LA*, 114-115). Architecture is a two-stage art, because first it is necessary to design a building and then to actually build it, which implies two phases in the construction process. Precisely for this reason, a notation was developed: to assure the correct transition from design to construction in a process in which several people intervene (*LA*, 220). Architecture can be considered allographic in its first phase and autographic in the second phase. The plan is allographic because its relevant aspect is the accurate reproduction of a series of specifications that fix the identity of a building, and, for this sake, the history of its production is irrelevant; in other words, it does not matter if the plan is the first and original one (if existing) or not. The actual building, which corresponds to the second phase, is autographic, for the history of production is the criterion necessary to distinguish between two buildings that follow the same plan, i.e., the first phase of an architectural work. This is the only way to distinguish between two or more identical works, such

[9] Gérard Genette expresses this idea as follows: "The case of architecture may seem paradoxical, because this art, which today possesses systems of (de)notation powerful enough to make possible the indefinite multiplication of its realizations, never exploits this possibility, except in aesthetically least prestigious productions […]. As to the unique character of works of high fashion (actually, I am told, one realization of each is authorized *per continent*, whatever that rather hazy geographical term may mean), it stems in part from a deliberate restriction whose motives are obvious, and in part from the autographic nature of this practice – for not only does nothing prevent an art from functioning in one regime in the case of certain of its works, and in the other in the case of certain others; it is also quite possible for one and the same work to be autographic in one of its parts and allographic in another" (Genette 1997, 97).

as the Cinderella Castle built in 1971 in the Walt Disney Resort in Florida and the one built in 1983 for the Tokyo Disney Resort.

Obviously, the distinction among identical works is possible only if site-specificity is taken as an indispensable requisite of their histories of production. Although at first glance it might seem that architecture is always site-specific, this statement cannot be generalized. Otherwise, some buildings that have been moved would not be the same building anymore: the Pergamon Altar now in the Pergamon Museum in Berlin would not be the Pergamon Altar, for it is not in Pergamon, but in Berlin. This is one difficulty that leads Maurice Lagueux to affirm that architecture is neither autographic nor allographic (Lagueux, 1998). According to him, an architectural work cannot be considered as allographic because "the strictest fidelity to its plans will not be enough to turn a 'reconstruction' of a building into another instance of it" (Lageux 1998, 24). However, it must be considered that an accurate "reconstruction" is possible only if plans are available, i.e., if an allographic system exists; thus, although autographically they are another work, reconstructions can be allographically considered as an instance of a work and this does not invalidate the allographic criterion for identifying architectural works. On the other side, for Lagueux, buildings cannot be autographic

> since their identity is not really defined by the historical conditions of their production, unless we consider that the site is in some way included in what Goodman means by the 'historical conditions' of the production of a work of art (Lageux 1998, 24).

In this passage, Lagueux rejects the possibility of including site in the history of production because other autographic works such as paintings are "typically context independent and can be exhibited anywhere in the world without loss of authenticity" (Lagueux 1998, 24). Yet the fact that some autographic works are site-specific, such as the Taj Mahal, does not imply that all other autographic works, such as Rembrandt's Lucretia or Zaha Hadid's Chanel Pavilion, have to be site-specific. Site is a constitutive element of the history of production of an autographic work only if it is considered relevant for establishing the work's identity; and this is not the case of most paintings (and of some buildings). Moreover, even though site is not an indispensable requisite in the history of production of a building, the autographic criterion is still valid for distinguishing between two identical works located on different sites: it is enough to know, for instance, that there is one Cinderella Castle built in 1971 and another in 1983 to discriminate one from the other. And if both

had been built at exactly the same time, different people would have participated in the construction, for no human being has the gift of ubiquity, and thus their histories of production would continue being different and valid as an identification criterion.

Indeed, as Goodman himself acknowledges, "not every art can be classed as autographic or as allographic. This classification applies only where we have some means of sorting objects or events into works – that is, where there is some criterion for identity of a work" (*MM*, 139). Yet architecture can indeed be classified according to these two categories, as the several examples discussed here illustrate. Moreover, the affirmation that architecture is neither autographic nor allographic would leave us with no criteria to establish the identity of architectural works. In other words, there would be no manner to conceptually differentiate between an original and a copy, a reproduction, or even a forgery of a building,[10] and it would be difficult to define the status of another kind of constructions, namely, restorations.

Restorations are another kind of architectural constructions that illustrate the complexity inherent in trying to establish what constitutes the identity of an architectural work. If architecture is considered only as allographic, then a building is restored by following the notational system and by repairing or reconstructing the missing or damaged elements of that building according to the features established in plans. Since, if allographic, its history of production is irrelevant, the restored building has to be taken as another instance of the work. However, in some cases the restorations undertaken in a damaged building are visible. Technically speaking, there are two kinds of restorations: an integral restoration, which intends to repair a work to make the whole look original, and a purist or archeological restoration, which contends that any substitution or addition has to be visible to avoid any pretense of authenticity and, thus, the missing parts are not replaced with pieces indistinguishable from the

[10] According to Goodman, "a forgery of a work of art is an object falsely purporting to have the history of production requisite for the (or an) original of the work" (*LA*, 122). Although technically complicated, it is not impossible to forge a building, at least an unknown one: by creating a false history of production a temple can be "discovered" in the middle of the jungle, or the plans of an unidentified palace can be "found."

already existing ones, but clearly discernible from the old ones.[11] That is to say, since in this latter kind of restoration the distinction between original and newer parts is relevant, some restorations cannot simply be considered as allographic works. However, if architecture is considered as only autographic, then a restoration can be regarded as an alteration of the building, since it implies replacing, adding, and also eliminating some traces of the history of production of the building. Nevertheless, these traces could also disappear through the effect of aging, and the autographic criterion would nonetheless remain valid for determining the identity of the work. Although a restored building would not wholly be like the initial building, it is still the same building as long as the original can be identified as such.[12]

Yet it is not clear if we could still consider a restored building to be the same if, for preservation's sake, each and every one of the original pieces were replaced with new ones. This is similar to the logical paradox raised by the ship of Theseus, whose planks were substituted for new ones when they became damaged until there was no trace left of the original ship and thus posing the question whether it was the same ship or not. As a corollary, one can question what happens if the replaced planks were used to build a second ship. If the stones that constituted a specific building – a church, for instance – were used to build the exact building, we could ask which of the churches, if either, is the original work. If allographic, then both buildings would be instances of the same work. If autographic, then it

[11] For a detailed discussion of these two kinds of restoration and its consequences on the aesthetic experience of a work of art see Mark Sagoff's work (Sagoff 1978**,** 457-458).

[12] Catherine Z. Elgin argues that, even though the history of production allows the distinction between an original and an exact copy, this does not involve restorations. When referring to paintings she affirms: "The identity of a painting is fixed, Goodman and I contend, by its history of production. A painting is a syntactically dense symbol, and such symbols admit of no replicas. No difference is so small as to be insignificant. Even the closest copy is bound to diverge from the original somewhere. So nothing with a different history can be the same painting. This explains why even the best reproductions are not instances of the paintings they purport to reproduce. It does not, however, tell against restorations. Although every difference *between instances* constitutes a difference between the symbols they instantiate, it does not follow that every change *in an individual instance* constitutes a difference in the symbol it instantiates" (Elgin 1997, 106). For the question of identity and restoration see the entire chapter "Restoration and Work Identity" (Elgin 1997, 97-109).

would be first necessary to determine which elements of the history of production of both buildings are relevant to the work's identity: to consider that the materials make the work the original has as a consequence the fact that only the second church would be the original; to consider that place is relevant leads us to affirm that the complete reconstructed church is the original.

To sum up, Goodman's categories of the autographic and the allographic allow us to establish what constitutes the identity of an architectural work. Although there are some architectural works that can be identified as clearly autographic (such as the Taj Mahal) or clearly allographic (such as the plentiful examples of tract housing), the majority of buildings can be considered from both the autographic and the allographic perspective (as has been shown in the cases of copies and restorations). Thus, not only is architecture as an artistic discipline a "mixed and transitional case" because it simultaneously comprises autographic as well as allographic works, but these two criteria are also intertwined within one and the same building. Both criteria to establish the identity of a work are interlaced insofar as the first stage of the construction process of a building is allographic and the second stage autographic; both criteria are inextricably linked insofar as the two are necessary to identify one and the same building, i.e., the autographic criterion serves to distinguish two identical buildings (such as the two Cinderella Castles), and, at the same time, the allographic criterion allows us to actually build two identical buildings (such as the Cinderella Castle in Orlando and the one in Tokyo). This array of possibilities does not have to be interpreted as a failure in the process of identifying works, but rather as evidence of the inherent complexity and richness of architecture. In other words, the affirmation that architecture is a "mixed and transitional" case is a reflection of the way in which we judge buildings' identities. On the one hand, some buildings are judged according to their history of production; on the other, some buildings are judged taking a notation as the criterion; and sometimes both criteria are used for one and the same building. The cause of this mixture lies not only in the lack of a perfect notation to identify works, but in our reluctance to judge some buildings exclusively from an allographic or an autographic perspective. As discussed in the next section, this hybridism in defining what constitutes the identity of architectural works is paradigmatically illustrated by the case of Mies van der Rohe's Barcelona Pavilion.

2. The Barcelona Pavilion: Autographic versus Allographic

Mies van der Rohe's German Pavilion (now known as Barcelona Pavilion), originally designed as a temporary structure for the 1929 International Exposition in Barcelona and rebuilt in 1986 as a permanent building, is a perfect case for testing Goodman's categories of the autographic and the allographic in architecture. Not by accident, Mies's building is discussed in existing publications on Goodman's criteria for establishing the identity of a work. Gérard Genette takes the Pavilion as an example of a work rebuilt after the demise of its author thanks to the availability of plans, thus acknowledging that the 1929 and the 1986 Pavilion can be (allographically) considered as two instances of the same work.[13] This opinion is shared by Maurice Lagueux, who affirms that the two Pavilions "constitute two instances of the same work and can be identified allographically."[14] Conversely, Sylvain Malfroy concludes that, using Goodman's categories, the logical status of the 1986 Pavilion is neither that of a copy nor that of an instance of the same work. Rather, according to her, "[w]hile transposing Mies's project, the architects of the 1986 Pavilion have indisputably claimed authorship. There is thus a real interest in distinguishing, in this respect, two Pavilions of Barcelona […]. The most recent is not a copy, but an original work à propos of a preceding

[13] "In contrast, buildings are "always" (as we say) being built or rebuilt from the plans of an allographic architect after his death: take a look, without leaving Barcelona, at Mies van der Rohe's German Pavilion" (Genette 1997, 37). And also: "But the fact that architectural (de)notations exist at least makes it possible to finish a work after the death of its author, like Spreckelsen's Great Arch de La Défense, or, as I have already indicated, to rebuild a demolished building exactly as it was, like the German Pavilion in Barcelona" (Genette 1997, 97).

[14] "In this manner, the famous Barcelona Pavilion built by Mies van der Rohe for the International Exposition of 1929, held in the city that gave it its name, was disassembled shortly afterwards – under conditions such that the materials from which it was constructed disappeared mysteriously – but was recently rebuilt on the site in accordance with its architect's plans using similar materials. Here we can state that the 1929 and 1989 [sic] pavilions, which are undoubtedly much more alike than are some interpretations of the same symphony, constitute two instances of the same work and can be identified allographically, even if, when faced with purists upset at not being able to admire Mies's *own* work, we might hesitate to assure them that these are equivalent instances of the same building" (Lagueux 1998, 21).

work" (Malfroy 2002, 87). Although Malfroy insists that her position is undeniable, both Genette and Lagueux defend the opposite position, also basing their arguments on Goodman's thought.

In this section I will argue that to consider the 1986 Pavilion simply as a copy does not completely define its identity status. Moreover, to conclude either that the two Pavilions are instances of the same work or that they are two different buildings entails a restriction of the elements that constitute the identity of Mies's work. By analyzing the reconstruction process of the Barcelona Pavilion, I will show how there is a continuous shift between the autographic and the allographic and that neither of them can be eliminated without losing some of the work's constitutive elements. This examination will help to answer whether or not the 1929 and the 1986 Pavilions are the same work.

One of the main arguments in favor of rebuilding the Pavilion was that it was considered a unique work of modern architecture. If we consider that the building that was unique was the 1929 Pavilion, or, using Goodman's terms, that one of the features that make the Pavilion autographic is its uniqueness, i.e., that in this case being autographic follows from uniqueness, then the 1986 Pavilion is a replica of the original one. Architecture is here considered as an autographic art, i.e., the history of production of the building is significant. That is to say, it is precisely the history of production that allows us to distinguish between the original and the replica: even though both Pavilions look the same, we know that they are different because one has a history of production (or construction) that took place under certain conditions in the spring of 1929, and the other has a history of construction under other conditions that took place between 1983 and 1986. That we are able to make this distinction between both buildings, and that we consider the distinction between an original and its replica relevant, brings us to affirm that the Barcelona Pavilion is an autographic work of architecture. This seems to be the general view, particularly because the 1986 building is defined as a reconstruction, and not as the original. Yet the distinction between reconstruction and original is not conclusive enough for determining the identities of the 1929 and the 1986 Pavilions and for incontestably assuming that they are two different buildings, as the following example illustrates.

When the reconstructed Pavilion was dedicated, Georgia van der Rohe, Mies's daughter affirmed that "for a second time, the German Pavilion of Barcelona ha[d] been given to the world" (Amela 1986, 52). Only a work that is not unique can be given "for a second time" and still continue to be

the same; only allographic works can be reproduced and maintain the same identity. This compels us to affirm that the building recovered in 1986 is the same as the one erected in 1929, i.e., that the world has been given back the "German Pavilion of Barcelona," assuming that what has been recovered is the same as what had been lost in 1930 due to its dismantling. In this context, the autographic distinction between original and replica is meaningless, because a unique and original work – a work of which there is only one – cannot be given twice. In other words, the Pavilion "can be lived in today for a second time" (Amela 1986, 52) only if it is interpreted and judged as an allographic work. Here we find a first oscillation between autographic and allographic criteria; a closer look at the process of reconstruction will show how this shifting is continuous, in other words, how autographic and allographic features are inextricably interwoven when trying to determine the identity of the Barcelona Pavilion.

The main obstacle faced by Solà-Morales, Cirici, and Ramos, the architects responsible for the reconstruction of Mies's Pavilion, was the absence of reliable and definitive plans, for the originals were lost.[15] In their words: "One thing that has been a constant problem in recreating the scheme for the German Pavilion in Barcelona is precisely its nonexistence" (Solà-Morales 1993, 9). Thus the first step was to fix its identity by means of a notational system, i.e., new plans, cross sections, and elevations. Although the autographic character of the Pavilion is what allows us to distinguish between the two buildings, to build the 1986 Pavilion the criterion followed was not the history of production of the building, but the recreation of a "scheme." Moreover, it is literally impossible to reproduce the history of production of the 1929 Pavilion to build the 1986 one and, therefore, the only possible way to rebuild the Pavilion is by developing a series of plans. In other words, the identity of the Barcelona Pavilion is established by allographic processes, i.e., by the creation of a notation. Goodman expresses this idea as follows:

> [W]hile availability of a notation is usually what establishes an art as allographic, mere availability of a notation is neither a necessary nor a sufficient condition. What *is necessary* is that identification of the or an instance of a work be

[15] Since its dismantling in January 1930, there have been several attempts to reconstruct the Pavilion. In 1957, Mies offered to redraw the plans, for "[t]he original construction drawings of the Pavilion were lost or misplaced in Germany" (Cirici 1983, 6), as Mies himself wrote to the then promoter of the Pavilion's reconstruction. For a summary of the reconstruction attempts see (Solà-Morales 1993, 26-27) and (Capdevila 2007, 40-42).

> independent of the history of production; a notation as much codifies as creates such an independent criterion (*MM*, 139).

This is precisely the case of the Barcelona Pavilion. Although a notational system that uniquely determined the work was not available, it was possible to identify the work without relying on the history of its construction, and the work of Solà-Morales and colleagues was to create an allographic criterion. Thus the Pavilion is here rather allographic than autographic, because it is actually possible to create a notation and because the history of its construction is not decisive to fix its identity. Nevertheless, the process of allographically establishing the identity of the Pavilion was open-ended. The architects found themselves "faced with the singular circumstance of a task of reconstruction for which there was an abundance of information relating to what we were to rebuild, yet not enough to establish, beyond all doubt, the characteristics belonging to Mies's mythic building" (Solà-Morales 1993, 5). The impossibility to completely fix the Pavilion's identity lies in the fact that not all the available sources provided the same information and thus could not be relied on.

It could be argued that this degree of uncertainty, even if minimal, invalidates the possibility of establishing the identity of Mies's work in an allographic way. However, the fact that it is not possible to exactly determine each and every one of the features that constituted the 1929 building does not mean that we cannot establish its identity by means of a notation, for some of the properties are irrelevant when establishing identity. A notation – a score, a script, an architectural plan – is abstract and "defines a work solely in terms of its necessary features, even if its contingent features are legion" (Genette 1997, 19). In other words, idiosyncratic attributes are excluded from notation and only unrenounceable properties are allographically fixed. Parallel cases in other arts, in which similar processes are undertaken to establish the identity of the work and whose validity is not questioned, reveal that uncertainty is not an impediment to define an architectural work by means of a plan: critical editions of literary works compare several versions of the same work and, in case of conflict, one of them (or a specific passage) is established as the canonical one. Much the same occurs when there are several versions of a musical score and the identity of the piece is established by fixing a unique score.

New plans, elevations for each façade, and various sections and details of the Pavilion were developed, specifying several features that had not

been defined in previous plans, and thus fixing what would constitute its identity in an unprecedented way. From that moment, following Goodman's theory, only the building that fulfills the conditions established by these notational systems can be considered the Barcelona Pavilion. The architects responsible for the 1986 Pavilion described their main reconstruction criterion as follows:

> [A]n undisputed premise here was the concept of a reconstruction that would interpret as faithfully as possible the idea and the material form of the 1929 Pavilion. If we have made a distinction between idea and material form, it is because the study of the materials used in the project, alongside other contemporary schemes by its architect, indicates that the physical execution of the building, for reasons of economy, haste or simple technological limitations, did not always come up to the level of its ideal character before, during and after construction (Solà-Morales 1993, 29).

Here Goodman's terms prompt us to consider that the Pavilion is allographic; the "ideal character" of the work would correspond to a work whose features are completely established by a notation, and the "material form" would be an actual built instance of this work. Yet, according to this premise, the 1929 Pavilion would not be a perfect version of the Barcelona Pavilion: it is known, for instance, that due to the scarcity of green marble and travertine, some walls were made out of brick and painted green and yellow (Solà-Morales 1993, 14). That is to say, the built 1929 Pavilion was an approximation of what it actually should have been; a building that apparently was the Barcelona Pavilion, but did not possess some of its properties. Strictly speaking, hence, the 1929 Pavilion cannot be considered as an instance of the Barcelona Pavilion as planned by Mies, for the "physical execution of the building" does not "come up to the level of its ideal character," as noted by Solà-Morales, Cirici, and Ramos. Obviously, one is very reluctant to affirm that the Pavilion built in 1929 is not the actual Barcelona Pavilion, and that the 1986 building is indeed an instance of the Pavilion, or, more precisely, the first and only instance of the Pavilion, because it accurately fulfills the requirements established by the notational system to be considered an instance of the work. As a matter of fact, this is literally the aim the reconstruction had – to complete the uncompleted 1929 work.[16]

[16] "It should be noted that it was our firm decision to totally clad all three faces of this wall, although we were well aware that on the 1929 building two of the exterior wall surfaces were without the green marble, which was replaced by a simple rendering in the same tone as the stone. In this respect, too, we were perfectly clear

Even if, according to the architects responsible for the reconstruction, the 1986 building is more the Barcelona Pavilion (and the 1929 Pavilion less), our intuitive reluctance to conclude that the 1929 Pavilion is not the Barcelona Pavilion – that the original Pavilion is actually not the Pavilion designed by Mies and built under his direction in 1929 because it was "not completed" – points to the need to shift the chosen identity criterion again, from allographic to autographic. In other words, this leads us to question the way in which the identity of the Barcelona Pavilion had been established by Solà-Morales, Cirici, and Ramos. Their premise when reconstructing Mies's work was to fulfill the Pavilion's "ideal character" fixed by them in a notation even though this would exclude the 1929 Pavilion from being an instance of the work. This means that their plan fixed as constitutive elements of the work some characteristics that the 1929 Pavilion did not fulfill. Nevertheless, a notation, or plan in this case, does not necessarily have to fix the "ideal," but only determine the constitutive elements of a work. It is perfectly possible to design a plan that establishes the identity of both the 1929 and the 1986 Pavilions. That is to say (and returning to Goodman's categories), the fact that the 1929 Pavilion does not fulfill the requirements established by a certain plan does not mean that the Barcelona Pavilion cannot be considered as an allographic work.

Another shift from allographic to autographic identity processes is necessary at this point. To elude the counterintuitive statement that the 1929 Pavilion is not the Barcelona Pavilion, it should be taken into account that the Barcelona Pavilion is not completely allographic, i.e., that its identity is not totally defined by a plan. While scores and scripts seem to be sufficient in fixing the identity of musical pieces and plays, and we consider that the differences among instances do not matter for establishing their identity, this seems not to be the case of architectural plans, for extranotational information (referring to the construction materials, for instance) is required to absolutely determine a building. It can be argued that the extranotational information is required because the notation in architecture is not accurate enough to completely establish the identity of a certain building, as Goodman also maintains (*LA*, 221). Nevertheless, it should be considered whether the extranotational information is necessary

---

that our criterion was not to reproduce the building as it had actually been when it was completed – or not completed – in 1929, but to carry through to its conclusion an idea with regard to which we had an abundance of information and the support of an architectural logic that was beyond all doubt (Solà-Morales 1993, 32).

to determine the identity of a work because some features that constitute its identity are not able to be allographically fixed. That is to say, even if we had a perfect notation, a specific work could also be determined by autographic features, and then its identity would still be a mixed case of allographic and autographic characteristics. Actually, this is the case of the Barcelona Pavilion: on the one hand, we distinguish between original and replica through the different histories of constructions – autographically; and, on the other, the reconstruction was possible by means of the creation of a notational system – allographically. Thus Goodman's statement about architecture being a transitional case is incorrect: architecture is not a "mixed and transitional case" (*LA*, 221). It is a mixed but *not* transitional case.

Let us suppose that we would consider the Barcelona Pavilion only as an allographic work, i.e., that we would establish its identity based only on a notation. Then, not only would the 1929 and the 1986 Pavilion be considered instances of the work, but so would every building that fulfilled the requirements established by the plan. Consequently, if the specific location of the Pavilion is not determined by its notation, it would be possible to have an indeterminate number of Barcelona Pavilions spread around the world. Yet it can also be argued that the location of the Pavilion can be fixed by a notation by including its latitude, longitude, and orientation in a plan. In this way, only one building at a time would be able to fulfill the requirements allographically established. Thus, only the 1929 and the 1986 Pavilion – and not any other Pavilion built in any other place – would be instances of the Barcelona Pavilion. This would mean that the Pavilion is an allographic work with two instances (1929 and 1986) – provided, of course, that the 1929 Pavilion is an actual instance of the Pavilion, which, as has been discussed above, is questionable depending on the notation or notations (plans, elevations, and sections) used to establish its identity. Even if it is possible to consider that the Barcelona Pavilion is only allographic, this does not seem to be completely adequate: even if it is technically possible to endlessly reproduce the Barcelona Pavilion following a plan, its site-specificity limits this allographic potential. Considering this, then Goodman's argument, according to which the Pavilion is not completely allographic because "its notational language has not yet acquired full authority to divorce identity of work in all cases from particular production" (*LA*, 221), is not quite right. The Pavilion is autographic because intrinsic autographic features preclude its being a completely allographic work; even if a perfect notation were available, the

Pavilion would still be partly autographic. This resistance to becoming completely allographic is a reflection of the values that prevail when we classify works as autographic or allographic. That is to say, if we are reluctant to consider the Pavilion completely allographic, it is because an important element of its identity would be lost; the Pavilion has indelible autographic features that limit its allographic characteristics.

As has been shown, site is one of the elements that can tip the balance in favor of considering that the Pavilion is autographic. Indeed, that site was one of the Pavilion's constitutive elements was a given for the architects responsible for the reconstruction: as Solà-Morales affirmed during the dedication ceremony, the Pavilion "should not and could not be in any other place" (Amela 1986, 52). However, not everybody agrees that the Pavilion is site-specific; instead, some consider it a "prototype; a perfect, autonomous spatial experiment capable of being considered as an object" (Solà-Morales 1993, 28). Thus, the Barcelona Pavilion can be regarded as both site-specific and not site-specific. Considering the Pavilion as not site-specific makes of it an allographic building reproducible everywhere (because its identity is determined by a notation that does not define its location). However, the fact of being site-specific does not make it immediately autographic (for site-specificity is not a necessary condition for a work to be autographic). If the Barcelona Pavilion is taken to be autographic, it is because its location has been considered a constitutive element of its identity, i.e., a relevant feature in the history of its construction. Consequently, the 1986 Pavilion is only a replica, or, more concretely, the only possible replica.

However, this is not the Pavilion's definite status, for another shift in the identity criteria is possible. It can be argued that the Barcelona Pavilion is autographic and that its ephemeral character is a constitutive element of the history of its production, because Mies designed it taking into account that it would stand only nine months. The 1986 Pavilion would then be, if it were only autographic, a partial reproduction of the 1929 Pavilion because it ignores the ephemeral character of the original having lasted (as of this writing) twenty-one years longer. Since this ephemeral character is not respected in the reproduction, either the durability of the Pavilion is not a constitutive element of the history of its construction – which seems difficult to maintain given the circumstances for which it was built – or the Pavilion is rather allographic than autographic – which is actually one of the underlying and implicit conditions for the Pavilion to be rebuilt,

because the possibility of creating a notation is what allowed the reconstruction.

After having discussed the main elements of the Pavilion's reconstruction process, it seems that the initial question, whether the 1929 Pavilion and the 1986 Pavilion are the same or not, does not have a univocal answer. In Goodman's terms, the Pavilion is both autographic and allographic: it is autographic, considering the history of its production (and then the work we can visit now is a replica), and allographic, considering its notation (and the present work is the 1986 instance of it). As has been shown, the two criteria to establish what constitutes the identity of a work (history of production and notation) are both pertinent for determining the identity of the Barcelona Pavilion. More than that, they are inseparable, for arguing that the Pavilion is only autographic or only allographic does not lead to convincing conclusions. Maintaining that the Barcelona Pavilion is only autographic ignores the fact that the identity of Mies's work can be established by a notation. Arguing that the Pavilion is only allographic could lead to the conclusion that the 1929 Pavilion itself is not the Barcelona Pavilion at all (if its notation fixes as relevant features the 1929 Pavilion did not possess). The Pavilion clearly has both autographic and allographic characteristics. Since we are reluctant to admit that the 1929 Pavilion is not an instance of the work and, rather, we consider it as the original, the Barcelona Pavilion is an autographic work and, consequently, the 1986 Pavilion is a reproduction. Nevertheless, since there exists the possibility to create a notation to fix its identity (and it has actually been created), the Barcelona Pavilion is allographic and the 1986 Pavilion is an instance of it. Yet not all the features of the Pavilion, such as site and materials, can indisputably be established by means of a notation: this makes the Pavilion more autographic than allographic – of course, only if these features are considered to be constitutive for the Pavilion. Furthermore, even though these features would be established by a plan, it would be possible to change them and the work continue to be considered the same, which is what actually happened: Mies decided to use brick instead of marble on some of the exterior walls and thus violated the initial project, but the Barcelona Pavilion continued being the Barcelona Pavilion. Taking this fact into account, the Pavilion is autographic rather than allographic.

According to Goodman, "[t]he allographic art has won its emancipation not by proclamation but by notation" (*LA*, 122). After challenging the notions of autographic and allographic with the case of the Barcelona

Pavilion, we should specify and affirm that notation is a necessary condition for architecture to achieve "full authority to divorce identity of work in all cases from particular production," but not a sufficient one (*LA*, 221). Apart from a precise and accurate notation, it is necessary that the difference between two architectural works based on the same plan be considered irrelevant, and this is (still) not the case in the Barcelona Pavilion: the distinction between original and reproduction matters. Although the Barcelona Pavilion could be a completely allographic work, it is a hybrid: it is a mixed work of architecture whose identity is established by autographic and allographic features – neither reducible to the other.

3. Conclusion

Determining the identity of the Barcelona Pavilion or any other architectural structure is an initial and indispensable step in construing the meaning of architecture, since both our understanding and experience of an architectural work change depending on whether or not we know that the work that we are contemplating is authentic. Furthermore, if, like Goodman, we grant aesthetics and artistic understanding a preeminent role in our comprehension and acknowledge that our construals actually construct worlds, then it is crucial to examine the conceptual elements that actively participate in these constructions, especially those referring to identity. Establishing what constitutes the identity of buildings, determining whether an architectural structure is original or not, and providing criteria to undertake these tasks are also constitutive in the creation of a work and its world.

As has been shown, the case of architecture is particularly complex because of the intrinsic features of this discipline. Goodman's statement that "insofar as its notational language has not yet acquired full authority to divorce identity of work in all cases from particular production, architecture is a mixed and transitional case" (*LA*, 221) has to be understood as a reflection of the current judgments concerning the identity of architecture. While it is certainly true that architecture is a mixed case, i.e., that the autographic and the allographic coexist as criteria to identify a work, the affirmation that architecture is transitional needs to be reconsidered. Goodman's prediction, according to which architecture will become completely allographic once a notational system that completely identifies a work is achieved, does not have as a necessary consequence

that the work's identity will be independent from its history of production. In other words, the autographic criterion for identifying works still concerns contemporary buildings with an accurate notation that fulfills its function of fixing the work's identity. Thus, the distinction between autographic and allographic is not related to an intrinsic historical development of architecture, a process which would bring us to affirm (with Goodman) that architecture is currently a transitional case. Nevertheless, it must be noted that architecture could only become allographic with the establishment of plans, which is indeed a historical process. Furthermore, since architecture is a discipline continuously evolving, it cannot be definitively excluded that, in the future, the allographic criterion will be enough for identifying architectural works. That is to say, it is uncertain whether or not the autographic features will cease being considered as constitutive elements of a work's identity and, in that way, when establishing the identity of a work, no loss of any fundamental property would be experienced. Since works of architecture incessantly age, change, and evolve, we are immersed in a never-ending process of construing and constructing architecture.[17]

References

Amela, V. (1986): "Inaugurado en Barcelona el pabellón alemán de Mies van der Rohe con la presencia de su hija" in: *La Vanguardia*: Barcelona, 3 June 1986, 52.

Capdevila, R. (2007): *Construing Reconstruction: the Barcelona Pavilion and Nelson Goodman's Aesthetic Philosophy* (SMArchS Thesis), Cambridge, MA: Massachusetts Institute of Technology, 2007.

Cirici, C.; Ramos, F.; Solà-Morales, I.; (1983): "Proyecto de reconstrucción del Pabellón Alemán de la Exposición Internacional de Barcelona de 1929" in: *Arquitecturas bis* 44 (1983), 6-17.

Elgin, C.Z. (1997): "Restoration and Work Identity", in: Elgin, C.Z.: *Between the Absolute and the Arbitrary,* Ithaca: Cornell University Press, 1997, 97-109.

Fisher, S. (2000): "Architectural Notation and Computer Aided Design" in: *The Journal of Aesthetics and Art Criticism* 58 (2000), 273-289.

[17] This article is an abridged version of my master thesis (Capdevila 2007). I am thankful to Catherine Z. Elgin for her comments and suggestions.

Genette, G. (1997): *The Work of Art. Immanence and Transcendence*, Ithaca: Cornell University Press, 1997.

Kivy, P. (2000): “How to Forge a Musical Work”, in: *The Journal of Aesthetics and Art Criticism* 58 (2000), 233-235.

Kivy, P. (2002): “Versions and "Versions," Forgeries and "Forgeries": A Response to Kirk Pillow”, in: *The Journal of Aesthetics and Art Criticism* 60 (2002), 180-182.

Lagueux, M. (1998): “Nelson Goodman and Architecture”, in: *Assemblage* 35 (1998), 18-35.

Malfroy, S. (2002): “Combien y a-t-il de *Pavillons de Barcelone*? Réflexion sur le statut logique de la “réplique” *in situ* d’un édifice détruit”, in: *Matières* 5 (2002), 76-89.

Pillow, K. (2002): “Versions and Forgeries: A Response to Kivy”, in: *The Journal of Aesthetics and Art Criticism* 60 (2002), 177-179.

Pillow, K. (2003): “Did Goodman’s distinction survive LeWitt?”, in: *The Journal of Aesthetics and art Criticism* 61 (2003), 365-380.

Sagoff, M. (1978): “On Restoring and Reproducing Art” in: *The Journal of Philosophy* 75 (1978), 453-470.

Solà-Morales, I.; Cirici, C.; Ramos, F. (1993): *Mies van der Rohe: Barcelona Pavilion*, Barcelona: Gustavo Gili, 1993.

Weemans, M. (2001): “Pratiques allographiques et reproduction: Sol LeWitt, Claude Rutault, Lawrence Weiner”, in: Goudinoux, V.; Weemans, M. (eds): *Reproductibilité et irreproductibilité de l'oeuvre d'art*, Bruxelles: Ante, 2001, 136-165.

Christoph Baumberger

Ambiguity in Architecture

Abstract
Buildings are frequently described as ambiguous and, indeed, they often involve the ambivalence associated with ambiguous symbols. In this paper, I develop a theory of architectural ambiguity within the framework of a Goodmanian symbol theory. Based upon Israel Scheffler's study of verbal and pictorial ambiguity, I present a theory of denotational ambiguity of buildings which distinguishes four types of ambiguity: elementary ambiguity, interpretation-ambiguity, multiple meaning and metaphor, which proves to be a special case of multiple meaning. Denotationally ambiguous buildings are exceptions, because buildings usually exemplify rather than denote. I therefore add a theory of exemplificational ambiguity. The crucial distinction between mere multiple exemplification and genuine exemplificational ambiguity leads to two versions of each of the first three types of ambiguity. The resulting extension of Goodman's symbol theory is of interest beyond architecture.

In architectural critique and theory, buildings and their parts are frequently described as ambiguous. Charles Jencks speaks about the ambiguity of Le Corbusier's chapel at Ronchamp which "sets the mind off on a wild goose chase where it actually catches the goose, among other animals" (Jencks 1984, 48). Roger Scruton points to "the ambiguous arrangement of columns employed by Peruzzi in the entrance loggia to the Palazzo Massimo" (Scruton 1979, 87). According to Rudolf Arnheim "ambiguity can make one and the same building look tall when it is perceived in one context and small when it is perceived in another" (Arnheim 1977, 179). And in his discussion of Palladio's Villa Foscari, called the Malcontenta, Colin Rowe mentions "the ambiguity, profound in both idea and form, in the equivocal conjunction of the temple front and domestic bloc" (Rowe 1976, 14).

But talk of ambiguity in architectural critique and theory is too elastic to be of theoretical interest. First of all, ambiguity should not be confused with the discrepancy between what something is and what it seems to be: The mere fact that a bank looks like a church does not make it ambiguous.

Neither should ambiguity be confounded with mere multifunctionality: Pillars of a church are not ambiguous simply because they simultaneously support the roof and shape the rooms. Thirdly, ambiguity does not just amount to generality or vagueness: It may be the case that some hotdog stand in form of a hotdog generally refers to hotdogs, and also that it fails to prepare us to decide whether a given snack is a hotdog or not, but this doesn't make the stand ambiguous. Finally, ambiguity should not be confused with a mere difference in reference by different parts of a building: the fact that the roof of a church refers to sailboats, while the frame of the nave is an inverted image of the skeletons of fishing boats neither makes the parts of the church nor the whole church ambiguous.

To be ambiguous a building must admit for multiple interpretations according to which it symbolizes or refers to different things. An interpretation is, in this paper, an assignment of a referent to a symbol. In the first two cases above, there need not be any reference at all. In the last two cases, the reference need not be ambiguous. Neither generality nor vagueness imply ambiguity; and that different parts of a building differ in their reference implies neither that the parts are, nor that the whole building is, ambiguous.

Nevertheless, buildings often do involve the sort of ambivalence associated with ambiguous symbols. In this paper, I propose a theory of architectural ambiguity in the framework of a Goodmanian symbol theory. I will confine myself to the two basic forms of reference acknowledged in this theory, denotation and exemplification. Usually ambiguity in architecture concerns exemplification, but there are a number of cases in which it concerns denotation. Starting with these exceptions, I present a theory of denotational ambiguity which is based upon Israel Scheffler's studies of verbal ambiguity (1979, 11–36; 1997, 25–49) and particularly of pictorial ambiguity (1997, 50–63). On that background I develop in section 2 a theory of exemplificational ambiguity.

1. Denotational ambiguity

Denotation is the semantic relation between a symbol and the objects to which it applies. Thus, a name denotes its bearer, a variable its values, a predicate the objects which it is true of, a passage the event it describes, a

portrait its subject and a plan the building whose form it determines. Symbols which denote or pretend to denote are *labels*. Labels are seldom taken in isolation; they typically function as members of families of alternatives which sort the objects of a given domain. Such a family is a *scheme* and the objects it sorts constitute its *realm*. The scheme consisting of "odd" and "even" sorts the realm of integers; the scheme consisting of "blue", "red", "yellow" etc. the realm of the colours. A *system* is a scheme correlated with a realm (*LA*, 71f; Elgin 1983, 37). This notion of a system is applicable to systems with only two alternatives ("odd"/"even") as well as to whole discourses (talk about art) or complete languages (English).

Even if parts of buildings – sculptures, frescos, mosaics, inscriptions, and so on – frequently function as labels, that does not hold for whole buildings, which usually neither denote nor pretend to denote. But there are a number of interesting exceptions. The notorious examples are the already mentioned hotdog stand in form of a hotdog and comparable structures of dubious taste. Architecturally more important cases are Claude-Nicolas Ledoux' design for a house of pleasure (Fig. 1), a brothel with a phallic ground plan, and the famous TWA-terminal built by Eero Saarinen in 1962 for the Kennedy Airport (Fig. 2), which, according to general opinion,

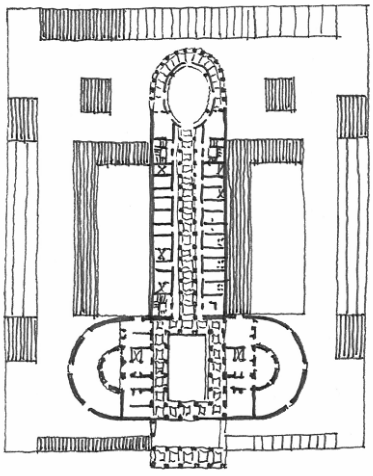


Figure 1. Claude-Nicolas Ledoux, House of Pleasure

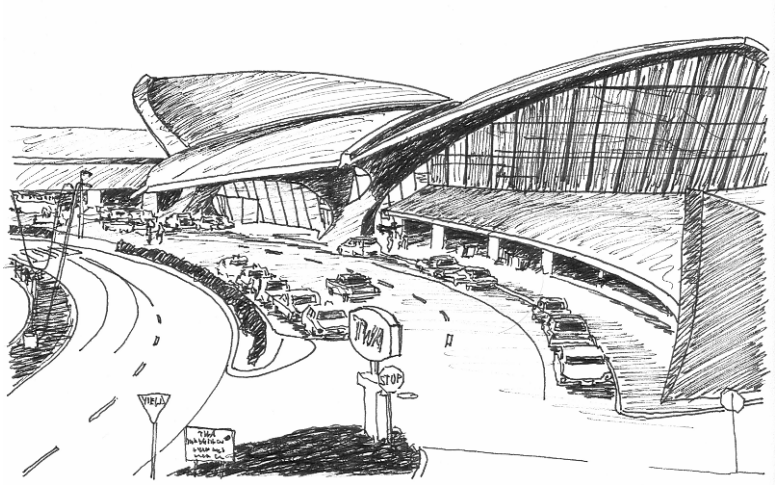


Figure 2. Eero Saarinen, TWA-Terminal, Kennedy Airport, New York

denotes an eagle flying away. In these cases denotation is an architectural counterpart to pictorial representation. But that need not be so. A building may (like a flag) denote a state even if it does not pictorially represent it. While in all these examples, the denotation is relatively clear and definite,

denoting buildings are often ambiguous. In what follows, it will be shown in what ways they are ambiguous.

1.1 Elementary ambiguity

Let us take language as a model. Here, denotational ambiguity may either concern the abstract sign type or the concrete sign tokens. Type ambiguity is basic, since a type may be ambiguous even if none of its tokens are ambiguous, but the ambiguity of a single token depends on the ambiguity of its type.

Israel Scheffler has presented an inscriptional explication of the basic case which he calls "elementary ambiguity". The explication makes use of neither intensions nor abstract types. Instead of asking whether different tokens are tokens of the same type, in such an approach one asks directly whether they are replicas of each other. That is the case if and only if they are spelled the same (*LA*, 115; Elgin 1983, 25). The explication then reads as follows: two tokens *x* and *y* are in a discourse *D* elementarily denotationally ambiguous with respect to one another if and only if *x* and *y* are in *D* replicas of one another (i.e. spelled the same) and diverge extensionally (i.e. either one denotes something not denoted by the other) (Scheffler 1979, 13, cf. 37). Besides lexical ambiguity, elementary ambiguity also covers the phenomena of metaphor and indexicality. The relativity to a discourse should prevent that a token counts already as ambiguous if there is an extensional divergent replica in another language or a remote context.

The explicated notion of elementary ambiguity is not directly applicable to architecture. It presupposes that several tokens can stay in the replica relation to one another. But buildings cannot stay in that relation because they are not composed of characters of an alphabet and thus cannot be spelled the same. Two ways out suggest themselves.

The first consists in replacing the replicas by the instantiations of an architectural work (which does not, as I use the expression, have to be an artwork). A particular building, then, is an instantiation of an architectural work if it complies with the plans of the architect; and an architectural work is a compliance-class of the plans. Such a semantic criterion for the identity of an architectural work is reasonable for buildings of series architecture. The notion of elementary ambiguity could thus be extended to

cover such buildings. But, firstly, there will hardly be any buildings which comply with the same plans but differ in their denotation. And, secondly, the buildings described as ambiguous typically belong to a context-specific architecture, the works of which require a historical criterion for their identity. Accordingly, a building counts as a specific architectural work if it is built at the relevant time in the relevant place; and an architectural work is typically an individual building.

However, such buildings have different temporal parts or time slices which can differ in their denotation. A governmental building may at a certain time denote a monarchy and at another time a democracy. It is recommended, therefore, to choose the second way out and take time slices instead of instantiations of buildings as substitutes for the replicas. The time slices *x* and *y* are, then, (in a discourse *D*) elementary ambiguous with respect to one another if and only if they are time slices of the same building and differ in their denotation. This explication calls, of course, for a criterion which determines whether different time slices are time slices of the same building. But such a criterion can be formulated as an expanded historical criterion which also requires a continuous history of the object. It is therefore clearly applicable to the cases of context-specific architecture.

Elementary ambiguity usually dissolves when one chooses a context and thus a time slice. If a time slice is ambiguous in a specific context, then we either have a case of interpretation-ambiguity or a case of multiple meaning. These two forms of ambiguity are the subject of the rest of section 1. Instead of talking of "time slices", I simply talk of "buildings"; and I speak of "buildings" even if the ambiguity frequently concerns only a part of a building.

1.2 Interpretation-ambiguity

A building is in a certain context denotationally interpretation-ambiguous if there is an indecision between different conflicting denotational interpretations, each of which makes maximally good sense of the building in the given context. For Le Corbusier's chapel Nôtre-Dame-du-Haut at Ronchamp (1950–55) (Fig. 3), Charles Jencks (1984, 48–49) proposed the following interpretations which may be understood in the sense of denotation: praying hands, a ship, a duck, a head-covering of a cleric and two brothers embracing each other. In order to count as interpretation-

ambiguous, a building has to fulfil three further conditions. Firstly, it does not only have to admit several rival denotational interpretations, but also to rule out certain interpretations as wrong. Rorschach inkblots preclude none of the proffered interpretations and do not, hence, count as ambiguous. Secondly, the conflicting interpretations have to concern the building as a whole. This condition is fulfilled if the interpretations concern a view of the building from one side; they do not have to concern the views of the building from every side. Thirdly, there have to be maximally satisfying, though conflicting, resolutions of the indecision concerning the denotation of the building which assign different denotations to the building. If there are no such resolutions, the building is not ambiguous, but only vague (cf. Scheffler 1997, 53). The chapel fulfils the three conditions, because it does not admit of every interpretation, each of the admitted interpretations concerns a view of the chapel from one side, and several (if not all) of them are satisfying in the given context. It does not matter that they gain

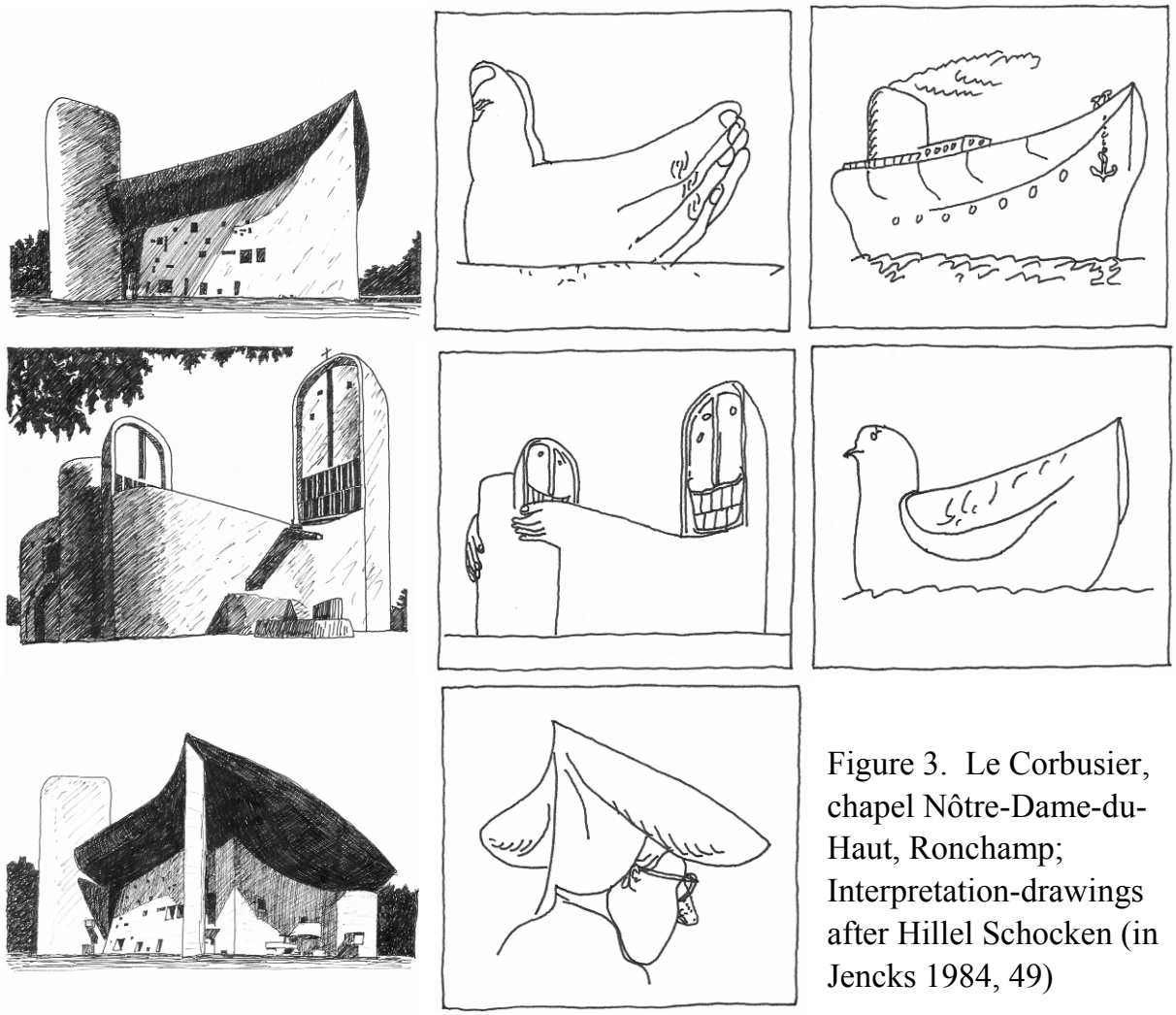


Figure 3. Le Corbusier, chapel Nôtre-Dame-du-Haut, Ronchamp; Interpretation-drawings after Hillel Schocken (in Jencks 1984, 49)

their pervasiveness only through the drawings by Hillel Schocken which accompany Jencks Interpretations. It does not have to be obvious what a building symbolizes, and drawings as well as descriptions may help us to see what it symbolizes.

So, the general *notion* of interpretation-ambiguity is applicable to architecture. But it is more difficult to extend to buildings Scheffler's *explication* of that notion as applied to language. The reason is simple. In order to prevent interpretation-ambiguity from collapsing into vagueness, the indecision concerning a token has to be related to the ambiguity of its type, i.e. to elementary ambiguity. In the case of language, that can be done as follows. If the tokens *x*, *y* and *z* are replicas of each other and *y* and *z* have decidably divergent extensions and are thus elementarily ambiguous with respect to each other, then the following holds: *x* is in a context *C* interpretation-ambiguous if and only if it will make in *C* both maximally good sense to interpret *x* as coextensive with *y* and to interpret *x* as coextensive with *z* (cf. Scheffler 1979, 16; 1997, 30–31, 52–53). But for the same reasons as in the case of elementary ambiguity such an explication is impossible for buildings. And to fall back on time slices does not provide a way out in the case of interpretation-ambiguity, since a building like the chapel at Ronchamp seems to denote different things simultaneously.

The proposed explication could perhaps be extended to buildings by replacing the replica relation by a looser connection between the ambiguous building and other extensionally divergent representations which do not have to be buildings. That is exactly what the drawings by Schocken achieve.

But buildings can be ambiguous even if the conflicting interpretations are coextensive. According to two of the countless interpretations of Frank Gehry's Guggenheim Museum in Bilbao (1993–97), the building represents, respectively, a sea monster with gleaming scales made of titanium zinc metal sheets and a dragon with an immense tail. Since there are neither sea-monsters nor dragons to be denoted, so that both interpretations assign the museum the same null-extension, the indecision between the two interpretations cannot be understood in terms of the relation to extensionally divergent representations. But how can we understand it then?

Now, according to one interpretation the building has null-denotation and is a sea-monster-representation, according to the other interpretation it has null-denotation and is a dragon-representation. The difference, then, does not concern the denotation of the building, but its characterization. Each of the two characterizations links the building to other representations which, in general, differ. The following explication results: a building *B* is in a context *C* denotationally interpretation-ambiguous if and only if it makes in *C* both maximally good sense to characterize *B* as an *F*-representation and to characterize *B* as a *G*-representation, and *C* does not provide a sufficient reason to decide between the two conflicting interpretations.[1] I add two remarks.

Firstly, the explication makes it clear that in architecture too, interpretation-ambiguity doesn't collapse into mere vagueness, even if the ambiguity of the token cannot be related to the ambiguity of its type here, since there is no such type. While the indecision in the case of vagueness is an indecision whether one should connect a specific characterization with a building or not, it is, in the case of interpretation-ambiguity, an indecision whether one should connect one rather than another characterization with a building (cf. Scheffler 1997, 57). Secondly, the explication is also applicable to buildings whose conflicting interpretations are not coextensive: the chapel at Ronchamp can be characterized both as a "ship-representation" and as a "duck-representation".

---

[1] A more elaborated treatment of the subject would reformulate this explication in terms of mention-selection. This term was introduced by Scheffler (1979, 31–36; cf. 1997, 11–21) to designate a kind of symbolization in which a symbol is applied not to what it denotes, but to mentions thereof. The term "dragon", for example, mention-selects dragon-pictures and dragon-descriptions. As explication we get: a building *B* is in a context *C* denotationally interpretation-ambiguous if and only if it makes in *C* both maximally good sense to interpret *B* so that it mention-selects *"F"* and to interpret *B* so that it mention-selects *"G"*, and *C* does not provide a sufficient reason to decide between the two conflicting interpretations. Such an explication in terms of mention-selection has two advantages. It is closer to our way of talking: instead of characterizing the Guggenheim Museum in Bilbao as a "dragon-representation", we just describe it as a "dragon". And it reflects the fact that ambiguity is a property of symbols which concerns their reference, since if a symbol mention-selects another one, the converse holds as well (cf. Scheffler 1997, 56).

1.3 Multiple meaning

In the case of interpretation-ambiguity it is assumed that only one of the rival interpretations is correct. If that is so, the rivalry can usually be resolved in favour of one or the other by broadening the original context and importing additional information. In the cases of the chapel at Ronchamp and the Guggenheim Museum in Bilbao, it certainly makes sense to weigh up the different interpretations against each other in order to decide which are more plausible than the others. But it is doubtful whether a complete resolution is possible.

If it turns out that several interpretations of the same building are simultaneously correct, we have an example of multiple meaning. Here, the ambiguity does not consist in the indecision between one meaning and another, but in a multiplicity of simultaneous meanings. Consider a fictitious example. In Italo Calvino's marvellous book *Invisible Cities*, Marco Polo reports to Kublai Kahn of a town called Despina which appears different depending on the side from which it is approached. Coming from the land with the camel, the town is seen as a ship which takes you away from the coast; coming from the sea by ship, the town is seen as a camel which takes you away from the desert of the sea to a fresh water oasis (Calvino 1997). In the case of Despina, it does not make sense to try to decide the rivalry between the two interpretations; it should rather be regarded as part of the functioning of the symbol. Someone who catches only one meaning does not catch the whole meaning, just like someone who only sees the rabbit head in the case of the famous duck-rabbit picture misses out on the whole meaning of the drawing.

This raises the problem of showing how a single time slice can possibly have several correct denotational interpretations without leading to contradiction. Neither is Despina a ship-and-camel-representation; because the town denotes what is denoted by a view of the town from one side, it does not contain a ship-representation and a camel-representation as parts. Nor is Despina a ship-or-camel-representation, since it does not simply represent everything which is either a ship or a camel. But it is also not the case that Despina is either a ship-representation or a camel-representation, with us being unable to decide which of the two. We are rather tempted to say that it is both: a ship-representation (as a whole) which denotes a ship

but not a camel, and also a camel-representation (as a whole) which denotes a camel but not a ship. But that is contradictory (cf. Scheffler 1997, 58).

The seeming contradiction can be resolved by ascending from the denotation of a label to its characterization, since the characterization of a label does not determine its denotation. A picture may be a dragon-picture without denoting dragons (because there are no dragons to be denoted). Once the link with denotation is broken, there is not longer any inconsistency in supposing Despina to be both a ship-representation and a camel-representation (cf. Scheffler 1997, 61). As for denotation, we can either say that Despina has no denotation or that it denotes both a ship and a camel. As explication we get: a building *B* has in a context *C* multiple meaning if and only if *B* is in *C* both an *F*-representation and a *G*-representation.[2]

1.4 Metaphor

Since a symbol with a metaphorical and a literal interpretation has several denotational interpretations, metaphor is a form of denotational ambiguity. However, while in cases of simple ambiguity as discussed until now the interpretations are independent of each other; one interpretation of a metaphor depends on the other: the metaphorical interpretation of a label is guided by its antecedent literal interpretation (*LA*, 71; *MM*, 129).

In the case of verbal metaphor, the different interpretations concern different tokens of the same type. Thus, verbal metaphor is a form of elementary ambiguity (Scheffler 1979, 15; 79). In the case of architectural metaphor, the different interpretations typically concern the same time slice of the same building. Since both interpretations are correct, architectural metaphor is a form of multiple meaning.

Although there are many metaphorical descriptions of buildings, only few buildings function as metaphors. The reason is that a building can only count as a metaphor if it admits of both a metaphorical and a literal interpretation. An example is the temple of the sun, built in the thirteenth

[2] This explication, too, could be reformulated in terms of mention-selection: a building *B* in context *C* has multiple meaning if and only if *B* both mention-selects *"F"* and mention-selects *"G"*.

century at Konarak in the Bay of Bengal, which has the form of a huge carriage. The temple, let us assume, denotes literally a carriage and metaphorically the sun.[3]

The sun-temple is at the same time a carriage-representation and a sun-representation. Thus it has multiple meaning. The first characterization is a literal one and the second a metaphorical one, since "carriage" itself is a metaphor for the sun (cf. Scheffler 1997, 63). Accordingly, a building *B* is in a context *C* a metaphor if it is literally an *F*-representation and metaphorically a *G*-representation.[4] The obvious circularity in this explanation can be eliminated by combining it with a Goodmanian explication of metaphor (cf. *LA*, 78–80; Elgin 1983, 59–65).

2. Exemplificational ambiguity

Examples of denotational ambiguity are exceptions, for the sole reason that buildings only rarely denote. Usually, when buildings are described as ambiguous, the ambiguity in question can be interpreted as concerning exemplification. So, let us now turn to exemplificational ambiguity and start with some introductory remarks on the notion of exemplification.

2.1 Exemplification

Consider the notorious tailor's swatch. It is, let us suppose, 10 x 10 cm, blue, of fine texture, made of linen, with zigzag edges. Although these labels, and many others as well, apply to the swatch, it functions as a sample only for those labels to which it refers as well; and these labels are exemplified by the swatch. In customary use, the exemplified labels may be "blue", "of fine texture" and "made of linen", but not "10 x 10 cm" and "zigzag edged". So there are two conditions for a symbol to exemplify a

[3] According to an alternative interpretation, the temple represents metaphorically the Indian sun-god Surya whose carriage the temple denotes literally (Behera 1996, xviii). This second and more common interpretation is less suitable for present purposes since it raises the question whether we face a case of fictional denotation.

[4] In terms of mention-selection we get: a building *B* is in a context *C* a metaphor if it literally mention-selects *"F"* and metaphorically mention-selects *"G"*.

label: it must be denoted by the label and it must refer to it (*LA*, 52; 53; *MM*, 59; Elgin 1983, 73).[5]

What is often overlooked is that the tailor's swatch functions in customary use as a sample only of a conjunction of labels. Someone who orders materials according to the sample will not be satisfied just by getting something blue. Rather, the delivered goods ought to instantiate also the other labels exemplified by the swatch. Thus, a symbol is a *sample* of the conjunction of the labels it exemplifies; and I suggest calling it an *exemplar* of each of these labels. The swatch is, thus, at the same time a "blue"-exemplar, a "of-fine-texture"-exemplar and a "made-of-linen"-exemplar.

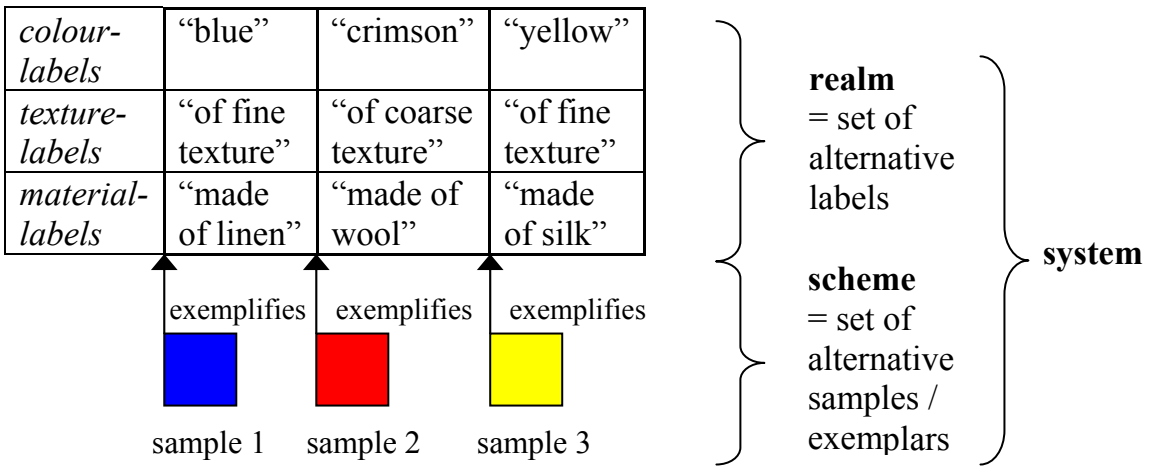
Like the interpretation of labels, the interpretation of samples is relative to a system (*LA*, 53; Elgin 1983, 73). An exemplificational system consists of a scheme which is correlated with a realm. Consider again the simplified example of the tailor's swatch (Fig. 4). The scheme is constituted by a set of alternative samples, the realm by a set of alternative labels (Elgin 1983, 121). Now, the realm is structured by the scheme. It is, firstly, structured into different alternatives which may serve as the referents of the sample. Thus, "crimson, of coarse texture and made of wool" and "yellow, of fine texture and made of silk" may function as alternatives to "blue, of fine texture and made of linen". The alternatives of a realm consist, thus, themselves of several components which can vary independently. Each component is a conjunct of an alternative of the realm. These conjuncts constitute the referents of the exemplars. The realm is thus, secondly, structured into different components which may serve as the referents of the exemplars. Thus, "crimson" und "yellow" may function as alternatives to "blue". The alternative components constitute the different parts of the realm. For the parts we often have categories like "colour-labels", "texture-labels" and "material-labels" at our disposal; but this needn't be so in every case.

Now, the same sample can simultaneously function in different exemplificational systems; and the system in use determines which labels the sample exemplifies. Which system is appropriate depends on the

[5] For the question how these two conditions should be interpreted, see the paper by Vermeulen, Brun and Baumberger in this volume.

context. In the normal context of selling fabric, the swatch functions as a fabric sample and exemplifies colour-, texture- and material-labels, but not size- and form-labels; in a seminar on symbol-theory, the swatch may function as a sample of a fabric sample and exemplify size- and form-labels, but not colour-, texture- and material-labels (*LA*, 53; *WW,* 64; Elgin 1983, 72–73).

Figure 4. Structure of an exemplificational system



categories: „colour-labels", „texture-labels", „material-labels"
alternatives of the realm: „blue, of fine texture and made of linen", „crimson, of coarse texture and made of wool", „yellow, of fine texture and made of silk"
components of the first alternative: "blue", "of fine texture", "made of linen"
parts of the realm: "blue/crimson /yellow", "of fine texture/of coarse texture", "made of linen/made of wool/made of silk"

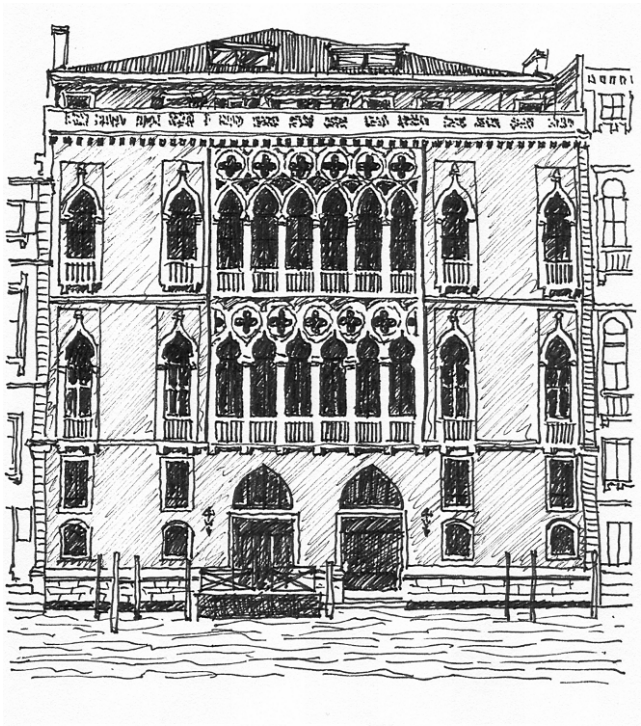
To understand how and what a sample symbolizes, one has to determine the system in which it functions. In the case of architecture, this is far more difficult than in the case of the tailor's samples, since the architectural systems are far less standardized. They are usually less discovered than developed by elaborating an interpretation. In such an elaboration, new labels exemplified by a building can always be figured out. Hence, in the case of architecture, the focus shifts from the level of the samples to the level of the exemplars; from the issue of the alternatives of a realm to the issue of the parts of the realm and the alternatives of which these parts are composed.

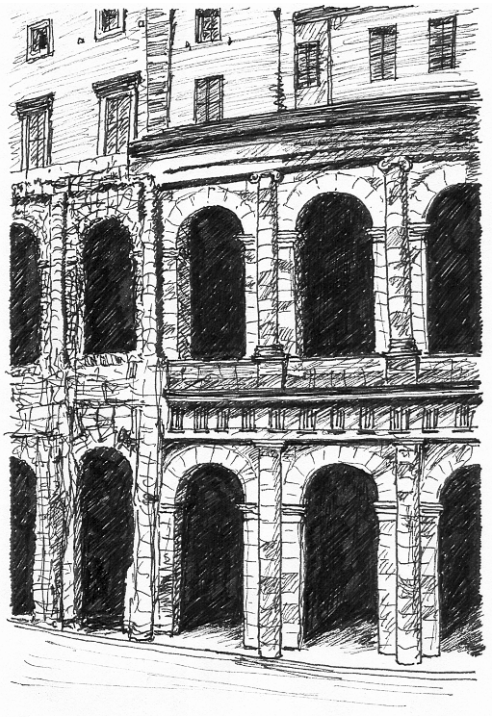
2.2 Conditions for exemplificational ambiguity

After these preliminary remarks, we can turn to exemplificational ambiguity. I start with a decisive difference between exemplificational and denotational ambiguity. In the case of exemplification, not every symbol exemplifying several labels is genuinely ambiguous. The fabric sample, for instance, already exemplifies several labels in the standard use without thereby being ambiguous. And a building like Charles Garnier's Opera in Paris (1861–74) may simultaneously exemplify labels like "massive", "overarticulated", "splendiferous", "exuberant" and "pretentious" (Jencks 1984, 71) without thereby being ambiguous. Thus, one has to distinguish between mere multiple exemplification and genuine ambiguity. Catherine Elgin (1983, 79–80) made the following proposal. A symbol has multiple exemplification if it refers to more than one label in a single system; and a symbol is exemplificationally ambiguous if it has different exemplificational interpretations in different systems. In the standard use, the fabric sample has only multiple exemplification; but it is ambiguous since it functions in other contexts as a sample for a fabric sample.

In a given case, there are various ways for deciding between mere multiple exemplification and genuine ambiguity. Such a decision depends on how the system in operation is constructed, and systems can be constructed in different ways. The Palazzo Strozzi in Florence (1489-1507) may simultaneously exemplify "well proportioned" and "unfriendly". If both labels are classified as belonging to different realms, the Palazzo is ambiguous; if they are classified as different components of the same alternative, the Palazzo refers unambiguously to both labels.

Elgin's suggestion, however, cannot be the whole story, because a building can also be ambiguous if it admits of different exemplificational interpretations within the same system. That is shown by the following three examples (cf. Scruton 1979, 85–94). The tracery from the central loggia on the upper floor of the Palazzo Pisani-Moretta in Venice (around 1460) (Fig. 5) can be seen both in such a way that neighbouring columns form aedicules, and in such a way that columns joined by the semi-circle of





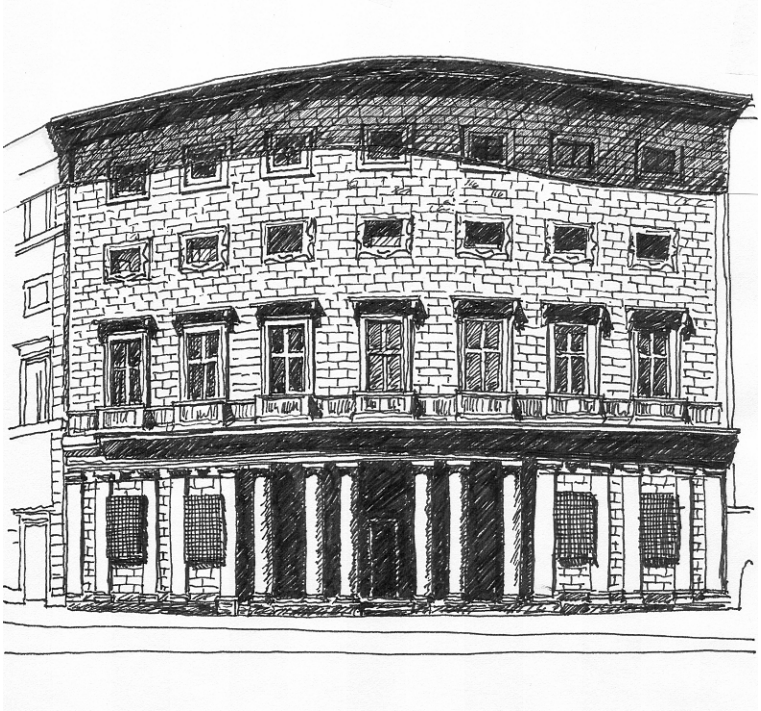


Figure 5, above on the left. Palazzo Pisani-Moretta, Venice.
Figure 6, on the left. Palazzo Massimo alle Colonne, Rome
Figure 7, above on the right. Theatre of Marcellus, Rome.

tracery form aedicules. The six columns of the entrance loggia of Baldassare Peruzzi's Palazzo Massimo alle Colonne in Rome (began 1535) (Fig. 6) can be seen both as two pairs and two single columns which form two further pairs together with the two adjacent pilasters, or as three pairs whereby the middle one frames the door and the others agree with the frames of the windows on the floor above. And in the case of a Roman building like the theatre of Marcellus in Rome (completed 13 BC) (Fig. 7), the half-columns and pilasters can both be seen as vertical decorations to the wall and as columns and pillars between which wall elements are filled in. All three examples constitute genuine cases of ambiguity, since in each case there is the same oscillation and tension between the competing interpretations as in the case of the duck-rabbit picture, which counts, according to general opinion, as ambiguous. And in each example the conflicting interpretations belong to the same system, since the labels which are exemplified according to these interpretations are part of the same realm.

As a result, the question arises how we can acknowledge the three examples as cases of genuine ambiguity without thereby being committed to count the fabric sample in its customary use as ambiguous. Accordingly, if a building admits of different exemplificational interpretations within the same system, we have to decide whether it is an instance of mere multiple exemplification or of genuine ambiguity. To do so, it is helpful to ask how in each of the two cases the exemplified labels are related to each other. In the case of mere multiple exemplification, the exemplified labels are different components of an alternative and belong to different parts of the realm. In the customary use of the fabric sample the labels "blue", "of fine texture" and "made of linen" belong to the three parts of the realm we designate by the categories "colour-labels", "texture-labels" and "material-labels". In the case of genuine ambiguity the exemplified labels are different alternatives of a part of the realm and accordingly belong to the same part of the realm. In the example of the Palazzo Massimo alle Colonne the labels "three pairs of columns" and "two pairs of columns plus two single columns" belong to the same part of the realm we can designate by the category "column-assigning-labels". The other two examples allow for analogous descriptions.

However, the fact that the exemplified labels belong to the same part of a realm is not sufficient for a building to be genuinely ambiguous, because the parts of a realm can be constructed in different ways. We can, for example, put together the colour-labels and the material-labels to form a part of the realm of the customary system for fabric samples; but that does not make the tailor's swatch in its customary use ambiguous. In order to be genuinely ambiguous, the relevant parts of a realm have to be minimal parts. But how should we determine minimal parts?

My suggestion is the following: labels belong to a minimal part of a realm if and only if they are either contrary to each other or coextensive but of a different kind.[6] Two labels are contrary to each other if there is no object to which they both apply; and two labels are coextensive but of a different kind if they denote the same object but each denotes it as something different from what the other label denotes it as. A label *x* denotes an object *y* as *z* in the relevant sense if and only if *x* denotes *y* and is a *z*-label (*LA*, 28).[7] In the case of the theatre of Marcellus, the two labels are contrary to each other: nothing can be a wall with a vertical decoration in front of it and at the same time a row of columns with wall elements filled in between. In the case of the Palazzo Massimo alle Colonne, the two labels are coextensive but of a different kind: both denote the six columns of the entrance loggia, but one describes them as three pairs of columns, the other as two pairs with two single columns; the same holds for the tracery of the Palazzo Pisani-Moretta.

We may call an exemplificational scheme which is correlated with a part of a (minimal) realm a (minimal) "subsystem". Then, the distinction can be

---

[6] At least verbal labels could also stand to each other in the relation of superordinate to subordinate labels. But this third possibility can be ignored, since a symbol can only in different systems simultaneously exemplify a label and one of its subordinate labels. If a fabric sample exemplifies both "blue" and "marine blue", it will do that in different systems, since according to one interpretation things will match the sample which will not match it according to the other interpretation.

[7] A more elaborated explication has to acknowledge both the fact that *y* may be denoted by *x* as a whole or by a part of *x* and the fact that *x* may be a *z*-label as a whole or merely through containing a *z*-label. As a result we get: a label *x* denotes an object *y* as *z* if and only if *x* is or contains a label that as a whole both denotes *y* and is a *z*-label (cf. *LA*, 27–31).

made as follows. A symbol which admits of different exemplificational interpretations within the same system is genuinely ambiguous if the conflicting interpretations take place within the same minimal subsystem; otherwise, it merely has multiple exemplification. Taking this for granted, we can formulate two conditions for exemplificational ambiguity which are disjunctively necessary and individually sufficient. A building, or a part of a building, is exemplificationally ambiguous

(E1) if it admits in different exemplificational systems (but within the same discourse) of different exemplificational interpretations; or

(E2) if it admits in the same minimal exemplificational subsystem of different exemplificational interpretations.

These two conditions generate for each of the three *types* of elementary ambiguity, interpretation-ambiguity and multiple meaning two *versions* of exemplificational ambiguity.

2.3 Elementary ambiguity

To extend the notion of elementary denotational ambiguity to architecture, I suggested replacing the replicas as the bearers of the divergent meanings by instantiations or time slices of the same building. Both suggestions can be transferred to exemplification. But here, too, elementary ambiguity of different instantiations of an architectural work plays virtually no role, since there are, at least in the relevant cases of context-specific architecture, no different instantiations. But elementary ambiguity of different time slices of a building plays a crucial role. I will therefore focus on this second case.

If we transfer the explication given for denotation to exemplification and add the implications of the two conditions (E1) and (E2), we get: the time slices *x* and *y* are (in a discourse *D*) elementarily ambiguous with respect to one another if and only if they are (in *D*) time slices of the same building and exemplify, interpreted in different contexts according to different systems or according to the same minimal subsystem, different labels. According to this explication, changes in meaning over time are construed as a form of elementary ambiguity. Let us look at an example for each of the two cases.

The Colosseum in Rome (72–80 AC) may serve as an example of the first case. A time slice of it from the period of Roman emperors may, amongst other things, have exemplified “eternal greatness of Rome”, a time slice from the Middle Ages instead “transitoriness of all earthly matters”, one from the Renaissance “unequalled model” and one from our times may exemplify “tourist attraction” or simply “Rome”. These different interpretations obviously belong to different systems.

But different interpretations may also belong to the same minimal subsystem. A time slice of the façade of Marcel Breuer’s department store De Bijenkorf (beehive) in Rotterdam exemplified after its opening, amongst other things, “regular” and “even”. Due to the impact of weather, a time slice of the same façade from today exemplifies “irregular” and “uneven”. In this case, the conflicting labels are contrary to each other, which seems typical for elementary ambiguity. Since different time slices exemplify contrary labels, no contradiction results.

Furthermore, elementary ambiguity usually dissolves in the case of exemplification if one chooses a context and thus a time slice. If a single time slice of a building is ambiguous in a specific context, then we have a case of interpretation-ambiguity or a case of multiple meaning.

2.4 Interpretation-ambiguity

A building is, in a specific context, exemplificationally interpretation-ambiguous if there is an indecision between conflicting exemplificational interpretations, which meet the following conditions: they make, in the given context, maximally good sense, only one is correct and they belong either to different systems or to the same minimal subsystem.

Very little is known about the meaning of the buildings of the Incas. Because there is a lot of speculation, they may serve as examples for conflicting interpretations in different systems. The pilaster-façade of the already mentioned theatre of Marcellus is an example for conflicting interpretations within the same minimal subsystem. In this second case, the conflicting labels are contrary to each other. This seems to be the typical case for interpretation-ambiguity. But again, no contradiction arises, since only one of the two interpretations is correct.

For a building to be not just vague but interpretation-ambiguous, the indecision must have maximally satisfying, though conflicting, resolutions which assign different exemplificational references to the building. This can be achieved by relating the building to other exemplars which unambiguously exemplify in their context but differ in their

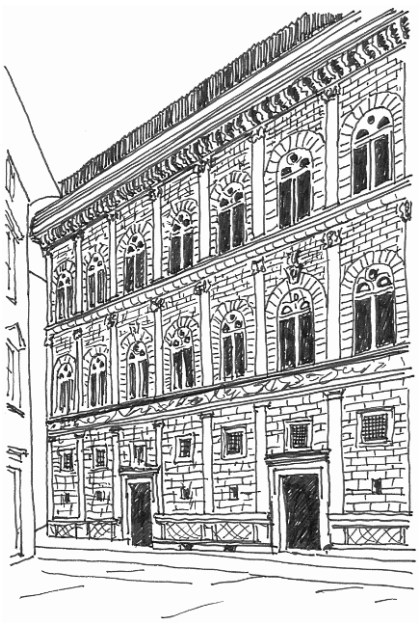


Figure 8. Leon Battista Alberti, Palazzo Ruccelai, Florence

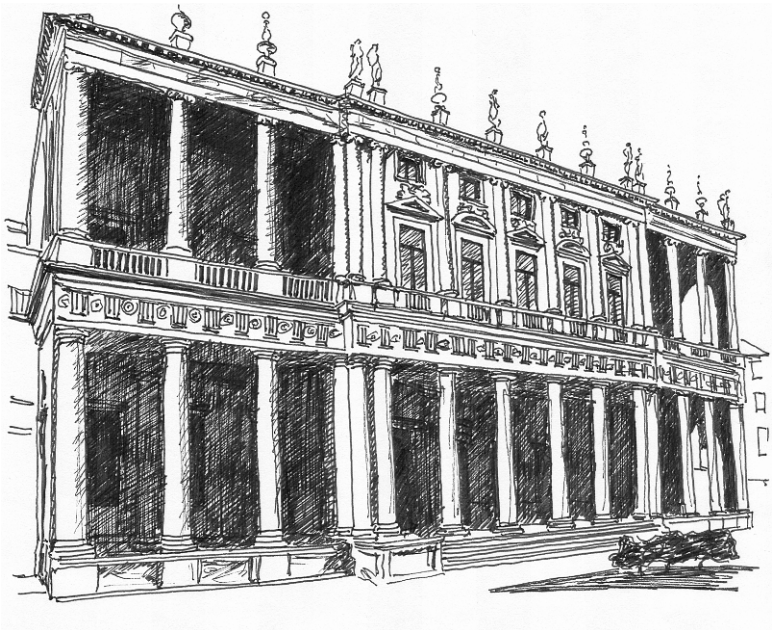


Figure 9. Andrea Palladio, Palazzo Chiericati, Vicenca

exemplification. Again, this relation can neither be a relation between different replicas nor between different time slices. But maybe, analogous to the case of denotation, a looser relation will do the job, since we often specify the conflicting interpretations of ambiguous buildings by relating them to other buildings which exemplify unambiguously. If we bring the theatre of Marcellus in connection with Leon Battista Alberti's Palazzo Rucellai (1446–51) (Fig. 8) whose pilasters are clearly to be interpreted as mere decorations, this supports the interpretation that reads the pilasters and half-columns of the theatre as mere decorations in front of the wall. Relating the theatre to Andrea Palladio's Palazzo Chiericati (began in 1550) (Fig. 9), whose half-columns are clearly to be interpreted as columns with wall elements filled in between, supports the interpretation that the pilasters and half-columns of the theatre are also to be read as pillars and columns between which wall elements are introduced.

Unfortunately this suggestion fails for two reasons. Firstly, there are not always corresponding unambiguous exemplars with which the building in

question can be connected. The reason is that exemplification is often specific, so that the exemplified labels are found in the exemplifying symbol itself (*LA*, 65; Elgin 1983, 77–87). Secondly, even if there are such exemplars, not every correct interpretation can be specified by saying that the building in question is coreferential with such an exemplar. Because the pilaster-façade can only be denoted by one of the two contrary labels: The pilasters can either be mere decorations or pillars but not both. Now, many theoreticians and historians argue that exactly that interpretation is correct according to which the façade *seems* to exemplify those labels which do not denote it (cf. e.g. Scruton 1979, 93). But that interpretation cannot be specified by saying that the façade is to be interpreted as coreferential with the façade of the Palazzo Chiericati.

The example presupposes that there is something like fictive exemplification. But Goodman stresses that there cannot be such a thing because a symbol can only exemplify labels which denote the symbol (*MM*, 60).[8] That is, of course, right. But in this sense, denotation cannot be fictive either, because a symbol can only denote objects that exist. Nonetheless, there can be, and there are, fictive labels which admittedly do not denote but pretend to denote. Accordingly, the relevant question in these cases is not what they denote, but rather how they are to be characterized: of what kind they are (*LA*, 21–26). A label only pretends to denote *x* if it is an *x*-label but does not denote *x* (since there is no *x* to be denoted). I suggest that exactly in this sense there can also be fictive exemplars which admittedly do not exemplify but pretend to exemplify. And in these cases, too, the relevant question is not what they exemplify

[8] Goodman acknowledges two cases in which one could speak of fictive exemplification. In the first case the label which exemplifies is fictive, in the sense that it is an actual label with null-denotation. “Pickwick” can in that sense exemplify “clown-label”; and that is, according to Goodman, meant when we say that Pickwick exemplifies clownishness. In the second case the label which is exemplified is fictive, in the sense that it is an actual label with literal null-denotation. That case is possible if, and only if, the exemplified label denotes the exemplifying symbol metaphorically. A man may in that sense exemplify “Don Juan” (*LA*, 66–67). Both cases differ from the pilaster example I am dealing with, where contrary to the first case the exemplar is not a fictive label and, contrary to the second case, not even metaphorically denoted by the label which it seems to exemplify. But this is exactly what, according to Goodman, makes a further sense of fictive exemplification, or exemplars, impossible (*MM*, 60).

but rather how they are to be characterized: of what kind they are. An exemplar only pretends to exemplify *F* if it is an *F*-exemplar but does not exemplify *F* (since *F* does not denote the exemplar).[9] What makes this possible is that the kind of an exemplar determines the exemplification no more than the kind of a label the denotation.

These considerations lead to an explication of interpretation-ambiguity which is parallel to the case of denotation. A building *B* is in a context *C* exemplificationally interpretation-ambiguous if and only if it makes in *C* either according to different symbol systems or according to the same minimal subsystem both maximally good sense to characterize *B* as an *F*-exemplar, and to characterize *B* as a *G*-exemplar.[10]

This explication avoids the two problems of the previous suggestion. Firstly, it does not demand alternative exemplars; it may well be that the ambiguous building is the only *F*-exemplar or the only *G*-exemplar. And secondly, the interpretations are no longer specified in terms of coreferential exemplars, but through the characterization of the ambiguous exemplars. Since the kind of the exemplar does not determine its exemplification, this is even then possible for fictive exemplars. Even if the pilaster façade does not consist of pillars with wall elements filled in between them, but rather of a wall with added decoration, it may still be a pillar-with-wall-elements-filled-in-exemplar.

---

[9] Elgin suggested (in personal communication) an explication which can do without fictive exemplars: An exemplar only pretends to exemplify *F* if it is not denoted by *F* and exemplifies a label like *F-ish* which has in its extension instances of *F* and other instances that are in suitable respects sufficiently similar to the instances of *F*. If this explication allows speaking of *F-ish*-exemplars, it will suit my purpose as well, but I prefer my own explication since the resulting formulations are simpler.

[10] The corresponding explication in terms of mention-selection is not transferable to the case of exemplification. Because in the case of exemplification, the labels mentioned would simply coincide with the exemplified labels, so that nothing would have been achieved. Consequently, the two advantages of the explication in terms of mention-selection (cf. note 1) are lost in the case of exemplification. That may be deplorable but it certainly does not constitute a knock-down argument against the proposed explication.

2.5 Multiple meaning

Interpretation-ambiguity can usually be resolved in favour of one interpretation by broadening the original context and taking further information into consideration. Whether that is possible in the case of the pilaster example may be questionable. If it turns out that both interpretations are correct, then we have an example of multiple meaning. Also, in the case of multiple meaning the conflicting interpretations belong either to different systems or to the same minimal subsystem.

In typical examples of the first version, different social groups use different symbol systems. In that way the flat roof buildings of Le Corbusier's working-class housing at Pessac near Bordeaux (1925) may for architects and critics of the modern movement have exemplified labels like "progressive", "user-friendly" and "rational", while they exemplified for the inhabitants labels like "insecure", "unfinished" and "decapitated", as is shown by the modifications the inhabitants made – ground floors were walled up, pitched roofs added, the ribbon windows divided up – so that the houses came to exemplify for them "security", "home" and "ownership" (cf. Jencks 1984, 54–55).

The loggias of the Palazzo Pisani-Moretta (Fig. 5) and the Palazzo Massimo alle Colonne (Fig. 6), mentioned in section 2.2, are examples of the second version. A less obvious and more complex example is provided by the façade which Michelangelo designed for the chapel San Lorenzo in Florence (1516–34). By drawing no less than sixteen figures with different structures emphasizing different axes of symmetry, Colin Rowe and Robert Slutzky pointed out that the façade oscillates between the proposed structures, which the façade exemplifies simultaneously (Rowe/Slutzky 1956, 212–218). Such multiple meaning is, of course, not deplorable. It is rather the result of hard work, as is shown by Michelangelo's design sketches. The overweight of the middle symmetry axis dissolves gradually and makes room for a balanced composition which admits multiple readings. The conflicting labels in the San Lorenzo example, too, are coextensive but of different kinds rather than contrary to each other. The different figures drawn by Rowe and Slutzky all denote the façade of San Lorenzo, but each of them as differently structured.

Do we need an explication for multiple meaning in the case of exemplification, or is it sufficient to describe it as a form of multiple exemplification in which the exemplified labels belong to the same minimal subsystem? Such a description is sufficient as long as the conflicting labels are coextensive and differ only in their kind, as it is the case in the foregoing examples. But there are also examples in which the exemplified labels are contrary to each other. In the case of certain pilaster façades, for example, it turns out that both interpretations are correct. Such cases cannot be understood as forms of multiple exemplification, since the façade can exemplify only one of the contrary labels – the one which applies to the façade. We thus need an explication that shows how such cases of multiple meaning are possible. The explication can be formulated in analogy to the one of interpretation-ambiguity. A building *B* has in a context *C* multiple meaning if and only if *B* is in *C* either according to different correct symbol systems or according to the same correct minimal subsystem both an *F*-exemplar and a *G*-exemplar.

This explication is also applicable in cases where the conflicting labels are coextensive but of a different kind: Even though the fact that something exemplifies *F* implies that it is an *F*-exemplar, it is not the case that something being an *F*-exemplar implies that it exemplifies *F*.

2.6 Metaphorical exemplification

Besides and precisely because of the existence of metaphorical denotation, there is also metaphorical exemplification. When an object exemplifies a label metaphorically, then it is metaphorically denoted by the label and refers to that label. Goodman explicates, as is well known, artistic expression as a form of metaphorical exemplification (*LA*, II, especially 85–95).

In the case of denotation, metaphor is a form of ambiguity which is distinguished from the simple forms of ambiguity by the fact that the different interpretations depend on each other. Metaphorical exemplification, on the other hand, is no extra form of ambiguity, since the dependence between the interpretations does not concern the exemplification, but the denotation. If a church literally exemplifies its cruciform and metaphorically “moved”, then the metaphorical interpretation does not depend on the literal interpretation. Rather, the

metaphorical application of “moved” to the church is guided by the literal application of the label to literally moved things like blowing clothes or waves.

But, of course, the types and versions of exemplificational ambiguity I have discussed up to now cover metaphorical exemplifications as well as literal exemplifications. For the exemplificational ambiguity of a building it is irrelevant whether the different exemplificational interpretations are literal or metaphorical. But the cases where a building admits of metaphorical as well as literal interpretations are typically cases of mere multiple exemplification. The church that exemplifies both its cruciform and “moved” is not ambiguous, because while the different interpretations belong to the same system, they do not belong to the same minimal subsystem since the simultaneously exemplified labels are neither contrary to each other nor coextensive but of a different kind.

3. Conclusion

Many buildings simultaneously symbolize in different ways: they denote, exemplify and allude at the same time. But that does not make them ambiguous. That a hotdog stand in form of a hotdog simultaneously denotes hotdogs, exemplifies its practical function and alludes to the American way of life does not turn the stand into an ambiguous symbol. Multiple symbolizations have, thus, to be distinguished from genuine ambiguity.

A building which is genuinely ambiguous has different interpretations within the same way of symbolization. Following Israel Scheffler’s explication of pictorial ambiguity, I have distinguished four types of architectural ambiguity by three conditions, each of which is based on the foregoing one. Firstly, if the interpretations concern different time slices of the same building in different contexts, we have a case of elementary ambiguity; if they concern the same time slice in the same context, we have one of the other types of ambiguity. Secondly, if only one of the interpretations is correct, we have a case of interpretation-ambiguity; if several interpretations are simultaneously correct, we have a case of multiple meaning. Thirdly, if one of the interpretations depends on another one, we have a case of metaphor.

The transfer of this scheme to the case of exemplification demands two modifications. Firstly, since not every case in which a time slice admits of several exemplificational interpretations is a case of ambiguity, one has to distinguish between mere multiple exemplification and genuine ambiguity. I suggested that we have a case of genuine ambiguity if the conflicting interpretations belong to different systems or to the same minimal subsystem; otherwise we have a case of mere multiple exemplification. A minimal subsystem is a scheme which is correlated with a minimal part of a realm. Such a minimal part is constituted by labels which are either contrary to each other or coextensive but of a different kind. This suggestion makes it necessary to distinguish for each of the first three types two versions of ambiguity. Secondly, metaphorical exemplification does not constitute a further, non-simple form of ambiguity. In the case both of denotational and exemplificational ambiguity, appeal to characterizations of symbols proved a suitable instrument for their explications.

Many results of the theory presented in this paper are of interest beyond the philosophy of architecture. The proposed structure of exemplificational systems is not specific to architecture, but holds for all such systems. Furthermore, the distinction between samples and exemplars as well as the idea of characterization, which admits of talk of fictive exemplars, will prove useful in the philosophy of the other arts as well. Since we can never be sure that we have figured out all the labels exemplified by an artwork, the focus in the arts shifts from the level of samples to the level of exemplars; and other artworks than buildings can also function as fictive exemplars. Finally, the theory of exemplificational ambiguity presupposes all devices and distinctions mentioned and it can easily be generalized beyond architecture. To get a general formulation, one can simply substitute “symbol” for “building” in the two individually sufficient and disjunctively necessary conditions for exemplificational ambiguity, (E1) and (E2). The same holds for the explications of interpretation-ambiguity and multiple meaning. Elementary ambiguity is somewhat more complicated. In systems with an alphabet, as in our verbal languages, elementary ambiguity typically concerns different symbols that are replicas of each other. Where there are several instantiations of a work, as in etching and design, it typically concerns different instantiations of a work. And where there are neither replicas nor multiple instantiations of a work,

as in pictures and carved sculptures, elementary ambiguity concerns different time slices of a given object. To get a general explication, we may introduce a relation *R* which is either the relation between different replicas, the relation between different instantiations of a given work or the relation between different time slices of the same object.[11]

References

Arnheim, R. (1977): *The Dynamics of Architectural Form*, Berkley; Los Angeles; London: University of California Press, 1977.

Behera, K. S. (1996): *Konarak. The Heritage of Mankind*, Vol. 1, New Delhi: Aryan Books International, 1996.

Calvino, I. (1997): *Invisible Cities*, London: Vintage, 1997.

Elgin, C. Z. (1983): *With Reference to Reference*, Indianapolis; Cambridge: Hackett Publishing Company, 1983.

Jencks, Ch. (1984): *The Language of Post-Modern Architecture.* Fourth revised and enlarged edition, New York: Rizzoli, 1984.

Rowe, C. (1976): "The Mathematics of the Ideal Villa", in: Rowe, C.: *The Mathematics of the Ideal Villa and Other Essays*, Cambridge, Mass.; London: MIT Press, 1976.

Rowe, C.; Slutzky, R. (1956): "Transparency: Literal and Phenomenal (part 2)", in: Ockman, J. (ed.): *Architecture Culture 1943-1968*, New York: Rizzoli, 1993, 206-225.

Scheffler, I. (1979): *Beyond the Letter. A Philosophical Inquiry into Ambiguity, Vagueness and Metaphor in Language*, London; Boston; Henley: Routledge & Kegan Paul, 1979.

Scheffler, I. (1997): *Symbolic Worlds. Art, Science, Language, Ritual*, Cambridge; New York: Cambridge University Press, 1997.

Scruton, R. (1979): *The Aesthetics of Architecture*, Princeton; New Jersey: Princeton University Press, 1979.

[11] I would like to thank Georg Brun for the fruitful discussions on the subject, Catherine Z. Elgin and Katia Saporiti for their helpful comments, and Ellen Milnes and Dominique Kuenzle for their help with the English version.

Sabine Ammon

Language of Architecture. Some Reflections on Nelson Goodman's Theory of Symbols[1]

Abstract
This article discusses the status of buildings as works of art. Judging architecture cannot be restricted to aesthetics, since architecture is an extremely interdisciplinary art. As a ‚joint venture', the example of architecture shows strengths as well as shortcomings of Goodman's theory of symbols.

When we try to describe architecture, a serious problem arises. As far as the eye can see, there are buildings. People in modern society live in a built environment that encompasses everything from the simple wooden shed to the skyscraper, including the town house, the tenement block and the office block, roads and bridges and landscapes shaped by human design – the list goes on and on. But not everything that is built is architecture. So what distinguishes the mere built from architectural art? In brief and succinct terms, architectural art exists when a building is a work of art. But this answer immediately raises another question: when are buildings art? When it comes to architecture, most popular definitions of art do not hold.

Art is what is created by an artist.

But being an artist's work does not mean that the object is art. Even celebrated architects count among their designs buildings that were simply carried out to keep the architectural firm going. And to explain art as what an artist creates *intentionally* is no help either, as artists can have a wide array of intentions. Having intentions is no guarantee of keeping them!

---

[1] Presented in an earlier version at the Congress CATH 2004: Philosophy of Architecture. Architecture of Philosophy, 11.7.2004, Bradford, UK.

Art is what is created using certain techniques and certain materials.

Again, architecture demonstrates the absurdity of the attempted definition. The building trade today is international, and the same techniques and materials are used the world over. Common standards of construction apply regardless whether the building is art or not. The way something is built does not decide on its status as art.

Art is what can be judged by aesthetic criteria.

But an architectural judgement cannot stop at aesthetic questions of shape, proportion, or style. Many additional factors have to be considered: 1. constructional aspects, from structural and technical aspects to the specific realization of details, choice of materials, and physics, 2. economic aspects like the relation of cost and output, 3. ethical aspects of sustainability, the consumption of resources and energy, and the impact on future generations, 4. psychological aspects investigating the relationship of the building to those who live or work in and around it, and 5. sociological aspects considering the consequences of the building for a community. This list is by no means exhaustive; in fact depending on the specific building, other factors differing in priority and relevance may also apply. It is crucial to point out that architecture is more than just aesthetics; it is a composition of a broad variety of needs and demands.

If the art of architecture cannot be defined with reference to the artist, materials and techniques, or aesthetic criteria, what is left? What is it that characterizes architectural art? What distinguishes a work of architectural art from the broad mass of buildings? The aim of my contribution is to tackle these questions with the help of the philosophy of Nelson Goodman.

1. The Language of Architecture

Nelson Goodman made an important contribution to the philosophy of architecture with his books *Languages of Art* (1968) and *Reconceptions in Philosophy and Other Arts and Sciences* (1988, with Catherine Z. Elgin). Embedded in a theory of symbols, he is able to thoroughly investigate the arts and sciences as symbolic systems – or, like the book's title suggests – as languages, if 'language' is used in a very general sense. Goodman

structures his investigation on logical considerations, using terms usually restricted to the philosophy of language. This enables him to systematically compare such diverse fields as astronomy, biology, music, and architecture.

His theory is based on a broad notion of the symbol. Symbols can be anything from demands and queries, words, categories, pictures, diagrams, musical passages and performances to buildings (*R*, 155f). According to Goodman, all symbolic functioning goes back to 'reference' in its basic sense of 'standing for'. It is the relationship of "a term or other sign or symbol and what it refers to" (*MM*, 55). In its simplest articulation, a symbol refers to an object, like most of our words do: for example, the term 'big' refers to objects of an extraordinary size.

But a symbol is never isolated; it is always related to other symbols. Like colours, symbols differentiate, sort, and systematize. 'Red', 'green', 'blue', 'orange', and many other labels divide our visual field into different areas. But none of these categories can act independently; they need their partners to define themselves. In this way, clusters of related symbols are established. These so-called symbol systems can differ significantly in their syntactic and semantic properties. With the help of these criteria, Goodman is able to explain the differences among the various symbol systems – for example, the differences between physics and literature, the everyday world and music, or painting and architecture.

Within this theoretical framework, Goodman pursues two lines of architectural investigation. First, he studies notational systems and their relation to the work of architecture in its ultimate built form. By focussing on the extent to which the plan determines what is actually built, statements can be made about the identification of the building.[2] Further dimensions of this problem become apparent when the whole planning process is examined, which can include everything from the initial sketch to the final plan, tender descriptions, CAD, and scale models. Questions arise as to whether a virtual architectural model can already be considered architecture, and what elements of the planning process determine the form in which the building is realized.

---

[2] S. Fisher has investigated this topic in 'Architectural Notation and Computer Aided Design' (Fisher 2000).

Secondly, Goodman looks at the diverse ways of symbolisation in buildings. His investigation is centred on the peculiarities of buildings that can be traced back to distinctive features of the symbol system of architecture. As my objective is to be able to differentiate among buildings in greater detail, I will concentrate on this aspect in the following.

2. Symbolic Functioning of Buildings

Goodman describes various ways of symbolic functioning, which all originate from two main forms of symbolisation: denotation and exemplification. *Denotation* is the simplest form of reference. When we study our verbal language, the situation is well known: the label ‘building’ refers to buildings. But when do real buildings denote? Denotation can be often found in two areas of architecture. On the one hand, there is the architecture of follies, which were designed to furnish country estates. The entrance of the greenhouse in the grounds of Dunmore Castle in Stirlingshire, shaped as a huge stone pineapple, refers to pineapples; the many sham castles, built from the outset as decorative ruins, denote ancient castles (Jones 1974). On the other hand, we find similar examples when looking at what may be termed ‘advertisement architecture’: the juice shop in the form of a huge orange symbolizes this fruit, the roof of a fish shop in the shape of a fish's body denotes seafood. However, apart from these fields, denotation in architecture is rare. A recent example is the planetarium of Santiago Calatrava in Valencia, which, together with its reflection, denotes an eye composed of eyelid, iris, and lashes.

Another way of basic symbolic functioning is *exemplification:* the ”reference by a sample to a feature of the sample” (*MM,* 59). Goodman speaks of features when a label denotes an object. But not each of an object’s features is also exemplified. Only if the sample refers back to the label is exemplification taking place. The paradigmatic example of this is a tailor’s swatch. The swatch stands for a certain colour or pattern, usually not for its size or shape. Exemplification in architecture takes place when certain features of a building are exposed. Sometimes the construction of buildings is designed to reveal the structure, as architects like Ludwig Mies van der Rohe or Norman Foster did. Here, the construction exemplifies, in its shape, the building’s structure and the interplay of forces at work within

it. In more general terms, exemplification occurs when a certain feature is brought out, be it the building's structure, shell, or roof.

Based on these rudimentary forms of symbolization are more complex ways that buildings make reference. *Expression* is a special case of exemplification, a process of symbolization that is often found in architecture. It happens when the denotation of the label is not literal but metaphorical. Deconstructivist buildings metaphorically refer to instability or dissolution, a palace to grandeur, a church to calm. The New National Gallery in Berlin of Mies van der Rohe expresses clarity, plainness, and simplicity through its reduced construction and its straight lines. *Mediated reference* takes place if a chain of related symbol processes has been established. In the above example of the greenhouse entrance, the chain of reference does not stop at the pineapple. During the epoch in which the greenhouse was built, the pineapple exemplified the enormous costs involved in cultivating it, and thus, in a metaphorical way, luxury and the wealth of its proprietor. The shape of the New National Gallery denotes a Greek temple, which in turn relates to a temple of art, an allusion to the Gallery's function as a museum of modern painting and sculpture. With this, I end the short overview of the various symbol processes at work in architecture. The crucial question now is how these forms of symbolization help to distinguish architectural art from mere buildings.

3. Symptoms of Architectural Art

When can we speak of architecture as art? Goodman has a straight answer: "A building is a work of art only insofar as it signifies, means, refers, symbolizes in some way" (*R*, 33). The premise is symbolisation: without symbolisation, no art. Although symbolisation is the basis of architectural art, there are restrictions. Not any reference makes a building a work of architectural art. There are many ways buildings mean that are not connected to architecture. Political or historical events in particular can make buildings signify specific meanings. A building can stand for the birth of a nation, a prison can stand for torture, an estate for wealth. In order to exclude these forms of signification, symbolisation must be understood to somehow depend on the architecture. According to

Goodman, architecture only is art when the building functions as an *aesthetic* symbol.

This specification immediately raises the question of what it is that makes a symbol aesthetic. In Goodman's writings we find a general answer that is not restricted to architecture. Instead of a definition, he indicates several *symptoms of the aesthetic* that "distinguish or are indicative of the symbolizing that constitutes functioning as a work of art" (*WW*, 67). On the level of the symbol, three features occur.[3] First, *exemplification* very often indicates art, through both literally and metaphorically possessed features. Second, *multiple and complex reference* is often related to art. Here, chains of symbols in various functions refer to an object. A third criterion is *repleteness* when not only one, but rather many aspects of a symbol turn out to be relevant. Given a work of art, these symptoms usually are fulfilled. However, they need not necessarily be so, they can also appear independently of one another.

To sum it up: architectural art presupposes aesthetic symbolization. When we return to the example of the New National Gallery, its status as a work of art can be demonstrated with the above criteria. Exemplification is part of the symbolisation and there are long chains of reference. Because not only the shape, but also the structure and many other elements such as columns and facade symbolize, the symbolic functioning of the building can be called 'replete'. When the building refers to other symbols and objects as in this case, when it interacts, then it is a symbol in a system of architecture – or metaphorically: it is part of the language of architecture. Not all buildings are symbols, however: there are also non-symbols and non-referential aspects of symbols (*R*, 156), the paradigm being an ordinary wooden shed. When a structure is inactive, when it does not relate, when it has no language, then it is just a built edifice, not art. It is a mute building.

[3] On the level of the system, two further symptoms must be mentioned: syntactic and semantic density.

4. Judging architecture

So far, the investigation has discussed criteria for aesthetic symbolisation. But considering only aesthetics, as Goodman does in his writings, does not do justice to architecture. As outlined above, architectural art cannot be described according to aesthetic criteria alone. Rather, social, psychological, economic, structural, and ethical aspects must also be included. These aspects are not opposed to aesthetics or architecture, but on the contrary are part of what constitutes architectural art. Only if a design is an intelligent solution to the economic and structural problem at hand are we willing to call the building a successful piece of work. The quality of a building shows not only in its shape and the relation of its parts, but also, for example, in its ethical implications. It is impossible to separate these aspects into different categories, looking for the best solution to each and presenting their final sum as architectural art. All these components are interdependent, and trade-offs must be made to get good results. Architectural art is a highly interdisciplinary 'joint venture'. There are many different questions and tasks a building is confronted with, but they must all be solved in order to produce a well-made building, a work of art. Therefore, not only aesthetic symbolic functioning has to be considered when we try to identify architectural art. Many systems are at work within it at the same time, many different ways of symbolic functioning are brought together simultaneously.

How can this be achieved? Putting all these different components together is a complex process of interpretation and construction. Any time we discuss architectural art – when we design, when we teach, when we assess architecture – judgements are involved. Judging means testing the symbol, checking its various functions, and trying out its position within the system. A judgement in this sense is an active undertaking; it includes making and creating. But it is not a situation where anything goes. Rather, as Goodman puts it: "Judgements of rightness of a building as a work of architecture ... are often in terms of some sort of good fit – fit of the parts together and of the whole to context and background" (ibid., 46). The fit is constituted by the tension between existing order and innovation, by the breaking of rules and the assessment of new ones. In this way, the different aspects and components are brought together, their limits are tested, and

novel creations emerge. Considering the many systems involved in architecture, the fit has to be achieved not only for one system but for many systems at the same time. Additionally, dependencies and interactions between the systems and resulting effects must be taken into account. Precisely these demands represent the specific situation of architecture which distinguishes it from many other arts: a situation which is not yet adequately grasped by Goodman's theory of symbols. This is left for future research.

5. Conclusion

Architecture is multi-disciplinary. Its difficulty, but also its challenge consists in bringing together dependencies, differing demands, and antagonistic methods from such diverse fields as construction, economics, ethics, psychology, and sociology. Nelson Goodman's theory of symbols offers only rudimentary assistance and must be developed further as it remains restricted to aesthetics. However, Goodman has provided a highly interesting approach regarding the problem of differentiation of the building as a work of art, an approach that can be extended beyond purely aesthetic considerations. His answer to the question of what distinguishes a piece of architectural art introduces a new perspective to the discussion of architecture. The way a building refers, the way it builds up multiple contextual relations, singles it out as a work of art. "A work of art typically means in varied and contrasting and shifting ways and is open to many equally good and enlightening interpretations" (ibid., 44). Or, to put it metaphorically, a work of art is able to break the silence, to start a dialogue in its own language.

References

Fisher, S. (2000): "Architectural Notation and Computer Aided Design", in: *Journal of Aesthetics and Art Criticism* 58, no. 3 (2000), 273-289.

Jones, B. (1974): *Follies and Grottoes*, London: Constable.

Axel Spree

Fiction, Truth, and Knowledge

Abstract
Starting with a reflection on different types of fictionality in literary texts, this article states (1) that there are different degrees of fictionality which can be seperated by means of some of Goodman's epistemological concepts, (2) that there is truth in fiction which nevertheless should *not* be regarded as relational or gradual, and (3) that we can gain knowledge from fictional texts which, in contrast to truth, may be "non-propositional", but which in its "gradation" is independent from the different degrees of fictionality.—This article understands itself both as a recapitulation of the effects that Goodman's theses had in the fields of aesthetics and epistemology and as a suggestion for a possible continuation of Goodman's thinking especially in the field of the theory of fiction.

On comparing the two best-known books by George Orwell, *1984* and *Animal Farm*, regarding their status as fictional literature on the one hand and their statements about the "real", nonfictional world on the other, one finds a number of similarities. In both books there is a description of fictive events, and in both works fictive characters act in more or less fictive places. Nevertheless, both books tell much about nonfictive events (past or expected) and real, so-called historical people. For example, without difficulty the two boars Napoleon and Snowball can be traced back to the initial "comrades-in-arms" and later antagonists Stalin and Trotzki, and the totalitarian state of "Big Brother" contains both elements of Nazism and Stalinism. However, from a philological perspective, one can see a number of significant differences between the two texts, differences that make it difficult to categorize them as simply literary, fictional texts. How can we gain similar assertions about the real world from both books while one of them describes human beings and the other describes animals behaving like human beings? Is the fictive person Winston Smith, who behaves like a real person (he works, eats, sleeps, falls in love, gets tortured), fictive in the same way as the pig Napoleon that does not exactly behave like a pig (at least not in the literal sense of the word)? Is it enough to say that the setting of both cases is "fictive", although in the one case it is just any

imaginary farm, in the other case it is London—even though it is a fictive London, namely one that did not exist in Orwell's time? And is the farm in *Animal Farm*, which is fictive but in its description very close to a nonfictive farm (aside, of course, from the strange behavior of the animals) in the same degree or in the same manner fictive as, for example, the magical cockaigne? Using the general characterization of "fictional" or "fictive" for literary texts and objects or persons in literary texts, respectively, seems all too general.[1] Thus, as the above comparison suggests, there are different degrees regarding the fictionality of literary texts.[2] Rephrasing Orwell, one could say that all fictions are equal, but some fictions are more equal than others. To be more precise, all fictions are false, but some fictions are more false than others. Conversely, this could also be true for the "truth" in fictions: Some fictions are true, but some fictions are truer than others. Starting with a comparatively innocent observation in the field of literary theory thereby leads to a truly profound epistemological problem, namely, whether truth is a relational concept, and consequently whether there are sentences which are *truer* than others.

The theses just hinted at—that all fiction is literal falsehood and that some fictions can be metaphorically true—come from Nelson Goodman.[3] These theses will pose the starting point for an inquiry into the question concerning the different degrees of fictionality and fictional truth. First of all, I want to elaborate on the concept of fiction as used by Goodman. In a second step, going beyond Goodman, this concept will be explored as a gradual and relational one. The problem of metaphorical truth then leads to the question of whether truth itself is gradual. Finally, the epistemological consequences stemming from the relationality of fiction will be analyzed in detail.

This article, on the occasion of Goodman's hundredth birthday, thus understands itself both as a recapitulation of the effects that Goodman's theses had in the fields of aesthetics and epistemology and as a suggestion

---

[1] I am using the terms "fictional" and "fictive" in the way explicated by Gottfried Gabriel. Accordingly, *fictional* is a predicate of texts, stories, discourses etc., *fictive*, in contrast, a predicate of objects; see (Gabriel 1979; 1982; 1997, 596).

[2] For "degrees of fictionality" ("Grade von Fiktionalität") see also (Gabriel 1997, 595).

[3] See his "Fiction for Five Fingers" in (*MM*, 123–126).

for a possible continuation of Goodman's thinking, especially in the field of aesthetics and, more particularly, the theory of fiction.

Fiction

In the first two chapters of *Languages of Art,* Goodman analyzes forms of reference, giving an example from the arts. Among other things, he differentiates between representation and description, exemplification and expression. Representation, in pictorial arts as well as in verbal description, Goodman conceptualizes as forms of denotation, which itself must be distinguished as unique denotation, multiple denotation and null denotation. In the context of the problem of fictionality, the third form is of the greatest interest. While, Goodman's example, "the duke of Wellington" is a unique denotation, a picture next to a definition in a dictionary or encyclopedia is an illustration or representation with multiple denotation. Such pictures do not denote, for example, a certain eagle singularly or the class of eagles collectively, but rather eagles in general (*LA*, 21). In contrast, a picture of a unicorn or of Pickwick or the expression "the man with three heads" neither has single nor multiple denotation; these are in fact representations or descriptions with null denotation.

But if the denotation of a picture of a unicorn and a picture of Pickwick is the same in both cases (namely, "null") then those things which both pictures represent should be the same as well. Nevertheless, obviously we are capable of differentiating between pictures of unicorns and pictures of Pickwick:

> much as most pieces of furniture are readily sorted out as desks, chairs, tables, etc., so most pictures are readily sorted out as pictures of Pickwick, of Pegasus, of a unicorn, etc., without reference to anything represented (*LA,* 21).

According to Goodman, the confusion emerges from the fact that a phrase like "image of a unicorn" appears as a mannerly two-place predicate, but in fact it is rather an unbreakable one-place predicate: "From the fact that *P* is a picture of or represents a unicorn we cannot infer that there is something that *P* is a picture of or represents" (*LA*, 22). Goodman therefore recommends a technical phrasing such as "unicorn-picture",

"Pickwick-picture" or "Pegasus-picture". Hereby nothing is said about what these images denote, yet it makes clear what kind of pictures they are:

> Obviously a picture cannot, barring equivocation, both represent Pickwick and represent nothing. But a picture may be of a certain kind—be a Pickwick-picture or a man-picture—without representing anything (*LA,* 22).

In fiction — be it fictional texts or fictitious pictorial representation — we typically are dealing with descriptions and depictions, respectively, with null denotation. Consequently, such descriptions are fictive or literally false.

Goodman's thesis regarding literary fiction thus is: *All fiction is literal, literary falsehood*. In this thesis, three aspects should be noted:

(1) "Fictional" or "fictive" is not the same as "literal". The literally false description of an object ("This man over there has three heads") may be fictional, yet it is by no means necessarily literature—and certainly not "literarily valuable". In reference to texts, it is possible to think up texts that are fictional but not literary. Among those are, for example, "scholarly fictions," that is, scholarly texts in which facts, observations, etc. are made up or at least are sugarcoated in order to back up a certain thesis that might indeed be right. On the other hand, there are texts that by no means are fictional but that still are literary (this case can probably be found more often; a classic example seems to be Friedrich Nietzsche's *Thus Quoth Zarathustra*).

(2) Not all literal falseness is fictive at the same time. A statement as 2 + 2 = 5 is literally false but it is "only false", "simply false" or also a lie, but obviously not fiction: "'Plain lies, damn lies, and statistics' are not fiction; neither are mistakes, whether computer or human, whether misprints, miscalculations, or misconceptions" (*MM,* 124).

This is the case because of (3): Only the quality of "literariness" turns literal falseness into literary fiction. In contrast to scholarly fiction, literary texts usually feature a number of specific symptoms which are characteristic for aesthetic modes of reference; among those are syntactic and semantic density, syntactic repleteness and exemplification.[4]

[4] For details on the "Symptoms of the Aesthetic" see (*LA*, 252–255; *WW*, 67–69).

Degrees of Fictionality

According to the general view, the concept "true" is like the concept "pregnant": *A little bit* pregnant (or true) does not exist, just as little as a woman (or proposition) can be *more* or *less* pregnant (or true). Usually, we also use the concept "fictive" as an "absolute": Either an object *is* fictive or it is *not*. It would appear very strange if one tried to put fictiveness into relation with other fictions or nonfiction—who would claim that the man with *two* heads is less fictive than the man with *three* heads? Yet, Winston Smith and the obstreperous boar Napoleon obviously are both fictive characters. Even if the reader is capable of identifying with both characters in a similar manner or to feel repelled by both of them in a similar manner, it remains questionable whether both can be called "fictive" in the same degree. Technically speaking, both "Winston Smith" and "Napoleon" (the pig from *Animal Farm*) are names with null denotation, and we are, according to Goodman, quite capable of differentiating between "Winston-Smith-descriptions" and "Napoleon-descriptions". But is the Winston-Smith-description not somehow more probable and more realistic—and therefore less fictive—than the Napoleon-description? After all, one can indeed imagine human beings who look like the fictive character Winston Smith and who behave like him under such circumstances. Yet, can one also imagine that a pig becomes a revolutionary? That it speaks? That it has its rivals killed in a Stalin-like fashion?

One answer suggests itself: On the one hand, there are fictions which (like all fictions) are literally false, but whose "realization" at least seems possible; and on the other hand, there are fictions which are completely impossible. The difference between those two obviously has something to do with probability and realism: The Winston-Smith-descriptions in *1984* are comparably realistic, as the attributes or characteristics from which they are put together could be attributes of nonfictional descriptions; for instance, the description of Winston-Smith is not a man-with-three-heads-description. In contrast, the descriptions of Napoleon in *Animal Farm* are comparably unrealistic, as there are elements which would make a nonfictional description of a pig a literally false description. This does not mean that the descriptions of Winston Smith are not literally false (otherwise, *per definitionem* they would not be considered as fiction), but

rather that a nonfictional description of a man *could* contain the same elements as a fictional description of Winston-Smith without *necessarily* turning into being false.

However, the literally false descriptions of the pig Napoleon contain something realistic, and this is the case in two aspects. Firstly, Napoleon is not able to do just anything that a pig otherwise cannot do (for example, he cannot use tools with his forelegs); secondly, even the unrealistic capabilities are realistic in the sense that they can realistically be ascribed to other beings, that is, people (Napoleon is not given any capabilities that no being possesses, such as moving through a vacuum at the speed of light). If one is able to classify the described objects of a text in what Goodman calls a "realm",[5] one is also able to state an even higher degree of fictionality. A description is *fictional in the highest degree* when the ascribed attributes of the described objects (here, the pigs) cannot realistically be ascribed to any of the objects of the realm (here, all beings).

With this distinction one can list *ad hoc* four degrees of fictionality: (1) There is *no fictionality* when descriptions are literally true (or have a unique denotation).[6] (2) There is a *low degree of fictionality* when the description of an object is literally false (or has null denotation), yet the nonfictional description of an object of the same realm does not necessarily turn fictional by adopting the attributes of the fictional description. (3) There is a *higher degree of fictionality* when adopting the attributes of a fictional description turning a similar nonfictional description *necessarily* into a fictional one. (4) There is the *highest degree of fictionality* when the ascribed attributes of the described objects cannot be ascribed realistically to any of the objects of the realm.

This is only a rough and preliminary classification whose purpose is less to give an exact gradation of fictionality but rather to emphasize the observation that there are such diverse degrees of fictionality. A problem could arise from the criteria of the realistic that may have been posed too naïvely here. However, this does not change anything when one defines realism, according to Goodman, as a "matter of habit" (*LA,* 38) and accepts

---

[5] For the difference between scheme, realm, and system see (*LA,* 71f; *R*, 7f).

[6] I leave out the aforementioned case that there can be literally false descriptions that are not fictional.

that description or picture as realistic that coincide the most with the norms of the current standard system.[7]

The Truth of Fiction

The first thesis *(All fiction is literal, literary falsehood)* is followed by the second: *Yet some fiction is true*. To clarify how "falseness" can be "true" I again go back to a very famous example given by Goodman himself:

> Before me is a picture of trees and cliffs by the sea, painted in dull grays, and expressing great sadness. This description gives information of three kinds, saying something about (1) what things the picture represents, (2) what properties it possesses, and (3) what feelings it expresses (*LA*, 50).

The first two propositions certainly inhere some truth value: to argue that the discussed picture is painted in brilliant yellow obviously would be false, and it would be just as false to argue that it depicts Churchill or a unicorn or Pickwick. Even if one might argue for a long time if single shades of the picture are still gray or already black or white and also if someone might think to recognize other figures in the trees or cliffs it appears that, given a clear and unambiguous case, the sentences "The picture is painted in dark gray" and "The picture depicts trees and cliffs at sea" are true propositions about the picture. But what about the proposition that the picture is sad or expresses deep sadness? A picture can be gray or square or expensive but, strictly speaking, it cannot be sad, as only sentient beings can feel sadness. "The picture is sad" thus is a literally false proposition. However, the proposition "The picture is cheerful" obviously is as false as the proposition "The novel *1984* is optimistic".

One should resist a temptation which in this context possibly suggests itself, namely to argue that the proposition "The picture is sad" is *less false* or *truer* than the proposition "The picture is cheerful". Goodman's solution is more elegant: the proposition "The picture is sad" is true, yet metaphorically true because the picture contains this sadness only figuratively or, as Goodman puts it, metaphorically. However, regardless

[7] Goodman's third thesis on fictions also belongs in this context: *Truth of fiction has nothing to do with realism*. See also (*LA,* 34–39) and "Three Types of Realism" in (*MM*, 126–130).

whether metaphorical or literal: the picture *possesses* the feature of being sad just as much as it possesses the feature of being grey. While the proposition "The picture is cheerful" is both literally and metaphorically false the proposition "The picture is sad" is literally false but metaphorically true.[8]

When transferred to fictions, or, more precisely, to fictional texts, one can state the following: *Animal Farm* gives a literally false description of the October Revolution and its consequences. Strictly speaking, the novel does not give any literal description of this revolution at all. It does give, however, much information about this and other revolutions. One can now try to convey this information in a literally true proposition—for example, of the kind of this rather poor description: "With time the former revolutionaries become worse than their former oppressors", or: "For the truly suffering, the animals or the people respectively, nothing changes, if at all it changes for the worse". Yet in such reformulations one will never grasp all the information given by *Animal Farm*. This is due to the fact that a literary text cannot be reduced to its implicit or explicit literally true statements. In any case, *Animal Farm* gives a literally false, yet a metaphorically true, description of the October Revolution. Or at least, for those disagreeing with Orwell, the descriptions in *Animal Farm* can be checked; they potentially have (metaphorical) truth value.

"Truer than …"?

According to the hitherto presented thoughts, it can be stated that there are different degrees of fiction and that "fictive" therefore cannot be seen as an

[8] Goodman's position in this context certainly is disputable. On the one hand he concedes, "The general question remains: what does a metaphor say and what makes it true?" (*LA*, 77) But, on the other hand, his answer might not convince everyone. The strategy does not rest on rigorously differentiating literal and metaphorical truth but in aligning them: "The question why predicates apply as they do metaphorically is much the same as the question why they apply as they do literally. […] Standards of truth are much the same whether the schema used is transferred or not." (*LA*, 78–79) Goodman's position, I argue, is indeed supportable; however, this important discussion cannot be given in this context and it also is not necessary for the scope of this article.

absolute, but rather as a relative concept. Furthermore, fictions can be true, namely in the sense that we can speak of metaphorically true propositions. Now the question arises if the relationality of fiction can be transported to truth as well. Put differently, can we call certain propositions truer than others?

First of all, the question of the relationality of truth has to be clearly differentiated from the question if propositions standing side by side can be similarly true at the same time. The second question certainly is to be affirmed. Even tolerating conflicting truths appears comparatively unproblematic. For example, literal and metaphorical truths conflict since metaphorical truths are literally false *per definitionem*. But also the praxis in the humanities, such as literary interpretations, knows many cases of similarly 'true' yet conflicting statements. The truth of such statements—this is almost a commonplace—depends on the theories or contexts in which they are stated.[9] In this sense truth is indeed a relative concept. But is truth also a gradual or relational concept?

Does one have to put truth (or at least some truths) in relation to other truths, and can one then decide on the "degree of truth"? This is not envisaged in Goodman's theory. Just as in the case of fiction, truth, for him, is absolute: a proposition is either true or it is not. While Goodman differentiates between literal and metaphorical truth he does not differentiate between more or less true. To call the aforementioned picture sad is literally false, yet metaphorically true. To argue it is not sad is literally true but metaphorically false. The two contradicting propositions are thus similarly true or false but one cannot say that one is more or less true than the other. The reason for this is the fact that Goodman exclusively accepts propositional truth, which is truth that limits itself to the truth of sentences. If he accepted as truth—as, for example, Martin Heidegger did—such a thing as the "Entbergung" of truth (as we find especially in the arts; see Heidegger 1986), it might also be possible to differentiate between different degrees of truth. One should follow Goodman's example and reserve truth for the truth of propositions, thus for propositional truth. A

[9] As far as one should speak of the truth in interpretations at all; and as far as one can say of conflicting statements formulated in the context of different theories that they relate to one and the same object.

non-propositional truth therefore does not exist and, consequently, there are no degrees of truth. Either a statement is true or it is not; a statement that is truer than another thus cannot be thought.

Literature—or better, fiction—thus can indeed contain truths insofar as it can contain metaphorically true sentences. Furthermore, by all means not all sentences in the text classified as fictional refer to invented or fictive objects. If, for example, the Joyce scholarship time and again claims that one could rebuild the destroyed Dublin by means of *Ulysses,* then it is a misjudgement based on too much respect for the devotion to details in this novel. All the same, it is correct to say that Joyce paid a great deal of attention to his placing the invented story in a historical and elaborately coherent setting, namely, the Dublin of 1904. Thus, such descriptions as Leopold Bloom's house in Eccles Street No. 7 indeed have an extension; they are descriptions with a single denotation and also, according to all we know, in respect to the described place at the described time they are true. Therefore, one can neither claim that all fictional sentences are literally false nor that *all* sentences in fiction cannot have any truth value at all.[10] However, from the fact that fiction, for the most part, is a mixture of sentences with null denotation and sentences with single or multiple denotation,[11] one cannot deduce that some sentences of fiction are truer

[10] As is well known, Gottlob Frege held the view that sentences in poetry must not be either true or false (see the introduction of the editor to (Frege 1990), esp. XVIII). The proceeding assertion that a text which only has a single sentence of that kind must altogether be regarded as poetry, in my opinion, is not only wrong—who would argue that Nietzsche's *Zarathrustra* is *exclusively* poetry?—but also little helpful in the context of the inquiry of epistemological functions of art in general and poetry in particular.

[11] "Mixture" does not only mean that sentences with an overall singular or multiple denotation stand next to sentences with a null denotation but (and this might even be the rule) that the denotation varies within single sentences. Let us keep the example of *Ulysses*. The first sentence is: "Stately, plump Buck Mulligan came from the stairhead, bearing a bowl of lather on which a mirror and a razor lay crossed." In this sentence, "Buck Mulligan" certainly is an expression with null denotation (the fact that a real person was a model for this character may be disregarded at this point). The scene, however, occurs in a place described in detail, namely the so-called Martello Tower in the bay of Dublin, where Joyce lived for a while. The mentioned "stairhead" thus might be understood as an expression with a singular denotation. Furthermore, it could

than others. If reducing the concept of truth to that of propositional truth (which I think would make sense) then this concept simply cannot be thought of as gradual.

But it remains a fact that some works of fiction tell us more about our "real" world than others; and this is independent from the probability or the degree of realism within the texts. *Animal Farm* certainly gives more information about the October Revolution than some historical descriptions from the field of a doctrinaire or only idealist communist. And this is the case although the historical description probably matches the established standards of realism much more than the descriptions of the "improbable" animals. According to what has been said, *Animal Farm* could not be any "truer"; but the book can transport a different and perhaps also a somewhat "superior" kind of knowledge in comparison to the historical account. This is also the case if this knowledge cannot be put in propositional sentences. For, in contrast to truth, this knowledge can also be non-propositional.

Art and Knowledge

A prerequisite for the study of the relationship between art and knowledge is the clarification of the relationship between truth and knowledge. The common belief that truth always also imparts knowledge is as false as the belief that knowledge always has to be true. In the first case truth is overrated, in the second case knowledge (or better, the concept of knowledge) is too limited. Firstly, the criteria making a proposition true are different from those criteria for a proposition imparting knowledge. According to Goodman, scholars could rattle off the most banal or trivial truths *ad infinitum* (*LA,* 262–263; *WW*, 18), but from this by no means does it follow that these truths impart any knowledge. Knowledge apparently has, to put it provisionally, something to do with the progress of understanding, and for this it is necessary that the meant truths contain a certain "quality"; "nothing but the truth" is decidedly too little. Secondly, if one requires that knowledge always has to possess truth value, then it can also always be put in propositional sentences. Yet in this way the

---

be reflected if expressions such as "bowl of lather," "mirror," and "razor" can be understood as expressions with multiple denotation.

concept of knowledge would be too limited. Following Goodman, Gottfried Gabriel and others have shown that it is especially the experience of art that supplies examples for non-propositional knowledge, thus examples that cannot be formulated as propositions with truth value.[12]

This aesthetic cognitivism, according to which the arts are not (as in the so-called emotivism) reducible to an expression and arousal of emotions, was not completely new when *Languages of Art* appeared in 1968, yet the book was spectacular since it was written by a representative of analytical philosophy, and, more particularly, on the basis of a sophisticated symbol theory which makes possible a rather detailed nomenclature of various functions of knowledge of the arts. In the meantime, also concerning aesthetics, one has to speak of a "cognitivist turn"—numerous publications in the last years attest to this fact, not least of which by the editors of this book.[13] I therefore can confine myself to naming the most important characteristics of this aesthetic cognitivism.

The point of origin of this thesis is that the mediation of knowledge is not limited to the sciences or philosophy: "Art and science are *in equal measure* part of our knowledge process" (Steinbrenner 1996, 12; my translation, A.S.). Yet, if one limits knowledge to propositional knowledge, that is, knowledge based on propositional knowledge formulated in propositional sentences, one cannot meaningfully speak of knowledge regarding the arts, especially those not using words (such as painting and music). Consequently, this calls for the broadening of the concept of knowledge on the one hand and the relativization of the concept of truth on the other. Firstly, the concept of propositional knowledge is too narrow to do justice to the diverse cognitive competencies and goals. Gabriel's exploration of non-propositional knowledge is just *one* proposal on broadening the concept of knowledge; Goodman goes further, arguing for the replacement of the concept of knowledge completely with the more copious concept of understanding. Oliver R. Scholz voices a certain skepticism about such a nobilitation of the concept of understanding—a skepticism that certainly suggests itself more clearly when considering the

[12] For the comparison of propositional truth and non-propositional knowledge see (Gabriel 1991, 214–217).

[13] See for example (Steinbrenner 1996), (Ernst 2000), and (Scholz 2001). For an overview and further literature see (Gaut 2003).

German tradition of hermeneutic philosophy than when considering American analytical philosophy. Therefore, given the aforementioned diversity of cognitive capabilities and goals, Scholz proposes using the language already existing in the rich epistemological vocabulary:

> It appears […] questionable if in all contexts this term [understanding] can fulfill the function of the concept of knowledge in an improved form. It might be better to redefine the meaning of the concept of knowledge insofar that it no longer is the key term for epistemology. The concept of knowledge would then be embedded in an ample net of concepts for the cognitive capabilities and achievements" (Scholz 2001, 38–39; my translation, A.S.).

Secondly, the object of this non-propositional knowledge obviously is not the (propositionally thought and thus "one-dimensional") truth but a multiple, "multi-dimensional" "rightness," to which truth as a special case has to be subordinated. It is obvious that through their diverse kinds of representation and mediation, the arts are an excellent object for demonstrating this broadened theory of knowledge—and this, in the end, was an impetus for Goodman's study of the arts in *Languages of Art.*

In conclusion, also in the case of non-propositional knowledge, one has to pose the question of whether or not there can be a more-or-less. In fact, this question certainly has to be answered in the affirmative. For example, the change in perspective—to name only one example of aesthetic experience that is also examined by Goodman[14]—can be more or less far-reaching although it might be difficult to specify a criterion according to which one can measure the change in knowledge. Furthermore, such knowledge can be more or less intensive; there are "degrees of intensity of knowledge" (as the Swiss writer Ludwig Hohl put it), which in the course of time can vary for one and the same person. However, the feeling for language resists the discourse of a more-or-less-known. At least in this context, it therefore seems best to replace the rather demure term of a non-propositional knowledge with a term used in everyday language, namely, with "understanding". It is rather uncommon to say: "Now I know it better

[14] Of course, the accomplishments of art cannot be reduced to the production and change of perspectives on the non-artistic world (see, for example, Scholz 2001); furthermore, the far-reaching impact of this accomplishment should not be underestimated. For reasons of textual economy in this context I, however, must focus on this seemingly modest aspect.

but I don't know it completely." According to the common use of language, to know something means to know it completely—otherwise you don't know it at all; knowledge is absolute. In contrast to this one can indeed say: "Now I understand it better but I don't understand it completely". This also equals Goodman's suggestion:

> Whereas knowledge typically requires truth, belief, and substantiation, understanding requires none of these. Propositions can be understood regardless of their truth and regardless of belief in them; and we can understand requests and queries and verbs and dances, though these are neither true nor false, neither believed nor disbelieved, and subject neither to demonstration nor refutation (*R*, 161).

Hereby the concept of the artistic production and change of perspectives achieves an intentional dimension: One does not see things just "somehow" differently but in a particular way; by understanding *Animal Farm* one also understands the "real" world in a particular, more coherent way. The understanding on the part of the recipients thus corresponds an aesthetic rightness on the part of the piece of art.

So what is won by including the concept of understanding when considering our problem of the relationality of fiction and of fictional truth? Goodman does not discuss understanding as a relational concept; one either understands a sentence, a question, a picture, a piece of music, or one does not. But this much can be said: the broadening of understanding is independent of the degree of fictionality; the reading of texts such as *Animal Farm* which are considered to be "fictional in a higher degree" can lead to a higher understanding of the October Revolution and its consequences than a nonfictional historical description. A possible "gradation of understanding" therefore cannot be designed parallel to a "gradation of fictionality". If it exists at all, then there are degrees of understanding both in nonfictional and, to a lesser or greater extent, in fictional texts. Questions of the quality and the quantity of understanding are also not dependent on the fictional or nonfictional status of a text: in the example constructed here, *Animal Farm* might be superior to the historical description; but the other way around is just as possible.

Coming back to the initially discussed epistemological problem, one can say that there are no degrees of truth (despite the fact that there can indeed be fictional truths). Yet the examination of diverse degrees of intensity of

understanding confirms once again the claim that in epistemological concerns, truth by no means plays the main role (or that it should). Truth, even fictional truth, is only one aspect of an all-encompassing rightness, which has more to do with success or failure than with true or false. Even though *Animal Farm* gives a metaphorically true description of the October Revolution, this fictional truth is still just *one* aspect of the aesthetic success of this text. With this acknowledgement the problem, of course, is neither gone nor resolved. All the same, inquiring into the problem of the relationality of fiction shows that it is not a problem solely of an epistemological thinking reduced to truth value or propositional sentences.

References

Ernst, G. (2000): "Ästhetik als Teil der Erkenntnistheorie bei Nelson Goodman", in: *Philosophisches Jahrbuch* 107 (2000), 316–340.

Frege, G. (1990): *Schriften zur Logik und Sprachphilosophie. Aus dem Nachlaß.* Ed. Gottfried Gabriel, Hamburg: Meiner, 3. ed. 1990.

Gabriel, G. (1979): "Fiction—A Semantic Approach", in: *Poetics* 8 (1979), 245–255.

Gabriel, G. (1982): "Fiction and Truth, Reconsidered", in: *Poetics* 11 (1982), 541–551.

Gabriel, G. (1991): *Zwischen Logik und Literatur*. *Erkenntnisformen von Dichtung*, *Philosophie und Wissenschaft*, Stuttgart/Weimar: Metzler, 1991.

Gabriel, G. (1997): "Fiktion", in: Weimar, K. (ed.): *Reallexikon der deutschen Literaturwissenschaft,* Vol. I, Berlin/New York: de Gruyter, 1997, 594–598.

Gaut, B. (2003): "Art and Knowledge", in: Levinson, J. (ed.): *The Oxford Handbook of Aesthetics,* Oxford: Oxford UP 2003, 436–450.

Heidegger, M. (1986): *Vom Wesen der Wahrheit*, Frankfurt a.M.: Vittorio Klostermann, 7. ed. 1986.

Steinbrenner, J. (1996): *Kognitivismus in der Ästhetik,* Würzburg: Königshausen & Neumann, 1996.

Scholz, O.R. (2001): “Kunst, Erkenntnis und Verstehen. Eine Verteidigung einer kognitivistischen Ästhetik”, in: Kleimann, B.; Schmücker, R. (eds): *Wozu Kunst? Die Frage nach ihrer Funktion,* Darmstadt: Wissenschaftliche Buchgesellschaft, 2001, 34–48.

Contributors

Sabine Ammon is Lecturer in Philosophy at the Technical University of Berlin.

Christoph Baumberger is Senior Research Fellow, Environmental Philosophy Group, Institute for Environmental Decisions, ETH Zürich.

Georg Brun is Senior Research Fellow, Environmental Philosophy Group, Institute for Environmental Decisions, ETH Zürich.

Remei Capdevila Werning gained her PhD at the Universitat Autònoma de Barcelona and is a Visiting Scholar at Harvard University.

Daniel Cohnitz is Professor of Philosophy at the University of Tartu, Estonia.

Gerhard Ernst is Professor of Philosophy at the University of Stuttgart.

Wolfgang Heydrich is Lecturer in German Linguistics at the Ruhr Universität Bochum.

Karl-Georg Niebergall is Professor of Philosophy at the Humboldt-Universität zu Berlin.

Marcus Rossberg is Assistant Professor of Philosophy at the University of Connecticut.

Richard Schantz is Professor of Philosophy at the University of Siegen.

Oliver R. Scholz is Professor of Philosophy at the University of Münster.

Ansgar Seide is Lecturer in Philosophy at the University of Münster.

Thomas Splett is Lecturer in Philosophy at the University of Munich.

Axel Spree currently holds a stand-in Professorship at the University of Münster.

Jakob Steinbrenner is Professor of Philosophy at the University of Munich.

Mark Textor is Reader in Philosophy at King's College London.

Inga Vermeulen studied philosophy at the University of Zürich and is now doing a PhD.