

Informal Notes

No. IN-CSLI-85-4

Analogy

Todd Davies



CENTER FOR THE STUDY  
OF LANGUAGE  
AND INFORMATION

Ventura Hall • Stanford University • Stanford, CA 94305



# Informal Notes

No. IN-CSLI-85-4

## Analogy

**Todd Davies**

CSLI was founded early in 1983 by researchers from Stanford University, SRI International, and Xerox PARC to further research and development of integrated theories of language, information, and computation. CSLI headquarters and the publication offices are located at the Stanford site.

CSLI/SRI International  
333 Ravenswood Avenue  
Menlo Park, CA 94025

CSLI/Stanford  
Ventura Hall  
Stanford, CA 94305

CSLI/Xerox PARC  
3333 Coyote Hill Road  
Palo Alto, CA 94304

The preparation and publication of this paper have been made possible in part through a award from the System Development Foundation.

Copyright ©1985  
Todd Davies

## Preface

This essay is a revised version of my senior thesis, which was submitted to Stanford University for the bachelor of science degree in the spring of 1985. In early 1984, I was exploring possible topics for an essay in the humanities honors program, which is open to undergraduates in any department at Stanford who wish to do a thesis incorporating disciplines outside their majors. I narrowed down the project to one that would incorporate ideas from logic and the study of language in an empirical analysis of arguments, and I applied through the Stanford Overseas Studies program to spend the autumn quarter of that year in Oxford, on the idea that debates at the Oxford Union Society and the philosophical traditions of Oxford would provide good fodder for my research. I cajoled Professor John Perry in the Philosophy Department, whom I had known from a course in 1982, to be my advisor for the project. His laboratory, the Center for the Study of Language and Information, was an attractively interdisciplinary environment for a statistics major with interests in data analysis and artificial intelligence trying to tackle issues in linguistics, philosophy and logic.

The spring and summer months of 1984 provided my first introduction to some of the diverse fields I was attempting to cross. Mary Pratt's class on "Discourse and Ideology" introduced me to some issues in language, and a directed reading under Marsh McCall that same spring provided me with some background in classical rhetoric and traditional logic. A summer internship at NCR Corporation in Colorado gave me the chance to learn quite broadly about issues and techniques in artificial intelligence, including the problems involved in getting a computer to reason by analogy. Yet I set off to Europe late in the summer without a clearly defined topic for my thesis. Professor Perry had set me up to meet with Professor J. O. Urmson in Oxford, who welcomed me into his home during several autumn afternoons for some tutoring in linguistic philosophy and discussion of my essay. The staffs of the Stanford Program in Oxford, the Oxford Union, Balliol College, and the Bodleian and British Libraries also helped to make my three months in Britain a productive time. After attending some debates, philosophy society meetings and lectures during the Michaelmas term and spending long hours in Oxford's many bookstores, I decided on the problem of analogy in late October.

I returned to Stanford in January, 1985, and began working through ideas about analogical reasoning on a weekly basis with John Perry. A small amount of funding from the System Development Foundation grant to CSLI helped to tide me over during the remaining two quarters over which this project ranged, supplementing the generous support from my parents, Berwyn and Gail Davies, and the moral support of my sister, Lyn. After my graduation I presented the CSLI Talk, "A Situational Theory of Analogy", based on my thesis, and received helpful feedback over the summer as a result of the seminar and follow-up talks given at NCR in Colorado and MCC in Texas. The revised version reflects that feedback and the help I received from many who were willing to discuss the thesis with me at CSLI and later during the fall when I began work at SRI.

Todd Davies  
November 6, 1985  
Menlo Park, California



I. Analogy and 'Analogy' .....	1
II. <i>Analogia Rationis</i> and the Problem of the Cogent Quality .....	5
III. A Situational Theory of Analogy .....	26
IV. Tests of the Theory and the Analysis of Analogies .....	48
V. Conclusion .....	61
Acknowledgements .....	63
Works Consulted .....	64



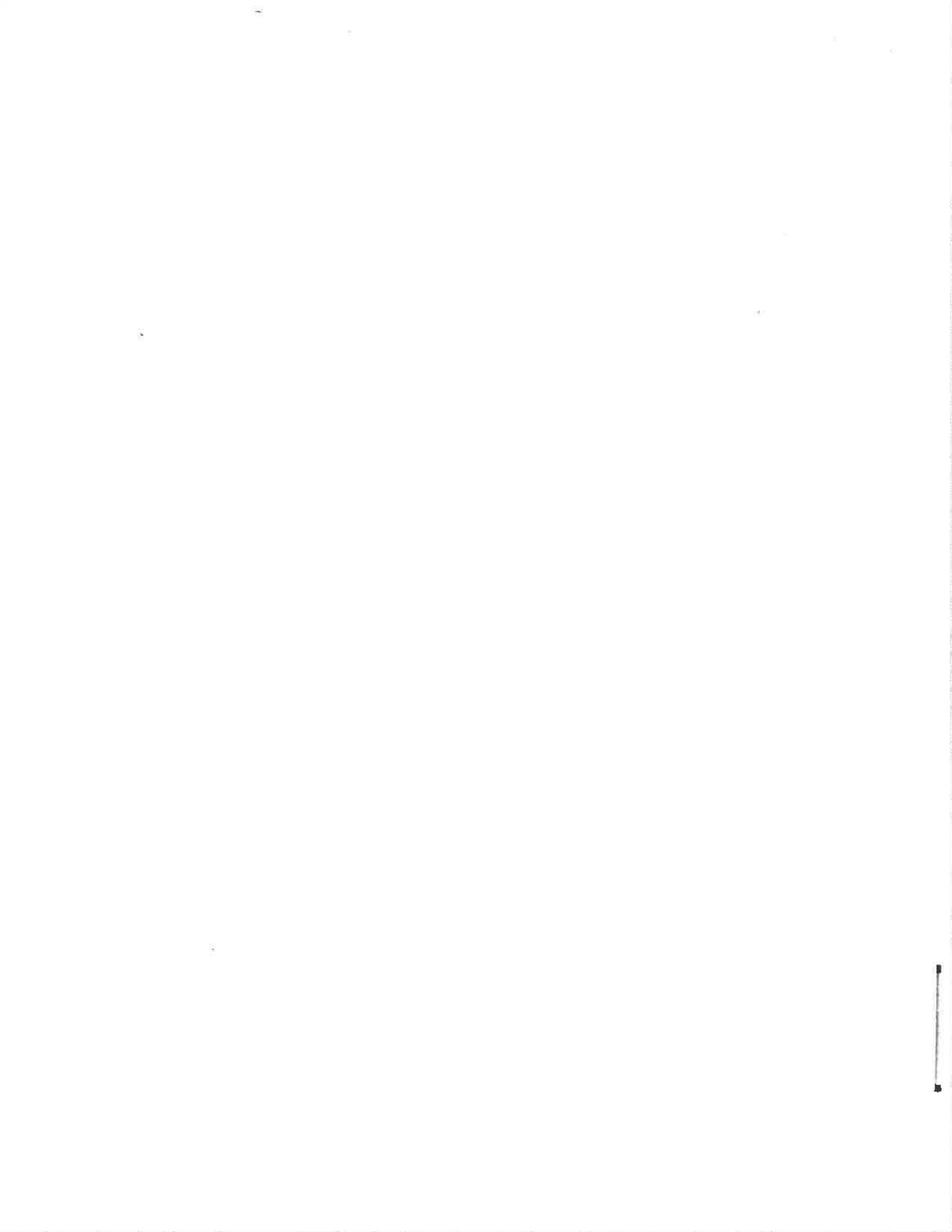


Arguments that make their point by means of similarities are imposters, and unless you are on your guard, will quite readily deceive you.

Simmias in Plato's  
*Phaedo*

All perception of truth is the detection of an analogy. We reason from our hands to our head.

Henry David Thoreau  
*Journal*  
September 5, 1851



## I. Analogy and 'Analogy'

In an analogy, we suppose, there is something of value. It may compel a new conclusion, or merely hint at one; in the absence of direct evidence it seems to allow us to infer what we are unable to know with certainty. I cannot really know, for instance, that other people have minds which give them the kinds of experiences my mind gives me, but an inference from analogy, perhaps exclusively, gives me the feeling that they do. Many philosophical problems are in fact questions about the soundness of some analogical argument. The existence of God has not been proven to the satisfaction of science, but arguments by analogy allow one to conclude that God exists, or perhaps that no God exists.

So these "conclusions" are strangely inconclusive. Perhaps conclusions are not what we get from analogy at all; indeed perhaps analogy is a fraud and our intuition that it is useful rests on a mistake. In any case, the status of what we infer by analogy has proven surprisingly difficult to characterize. J. L. Austin pointed out that we do not ordinarily claim to *know* by analogy, only to argue by it.<sup>1</sup> If we are convinced by an analogy, then, our inference is a kind of belief, the grounding for which is the subject of this essay.

In beginning a study of analogy one might well ask the simple question, "What is it?" Following the tradition of linguistic philosophy, I would like firstly to explore the uses of the word 'analogy', being mindful that the concept I wish to analyze may be only one among a collection of its possible meanings. In his writings, John Stuart Mill gave the term a dubious distinction: "There is no word...", he wrote, "which is used more loosely or in a greater variety of senses, than Analogy."<sup>2</sup>

Still the word does seem to have a general meaning, being roughly "likeness, not including identity". What proves more difficult to state is just what "an analogy" is, or rather what *sort* of thing it is. We speak of an analogy being "invoked" or being "good or bad", but what other things might we speak this way about? "Analogy" may variously be the basis for a mode of reasoning, or a relation between uses of words, or a type of relation that can exist between objects, but "the analogy" seems to refer to something other than just single instances of these. For example, we may say with equal confidence that "similarity" is a basis for reasoning or is a relation between senses or things; but we would not, I think, say "the similarity between A and B is good" or "he invoked a likeness" or "let me try this similarity out on you". 'Analogy' basically means "similarity"<sup>3</sup>, but 'an analogy' often means something close to "an idea that there is a similarity" rather than "a similarity".

All of this is just a way to say that the word 'analogy' functions *analogously*, not identically, in different types of sentences. Furthermore, the modern uses of 'analogy' have been derived by analogy from older uses. Since the word's evolution and its different modern uses are a barrier to deciphering the literature on analogy, sorting these out is an important first

---

<sup>1</sup>Austin, "Other Minds", *Philosophical Papers*, p. 77

<sup>2</sup>Mill, "Of Analogy" (Ch. 20), *A System of Logic*, p. 393

<sup>3</sup>Of course, the language is ever changing and recently analogy has often been referred to as a mode of reasoning itself rather than the basis for one

step in researching how others have approached the particular problems of analogy with which I am concerned.

The concept of analogy is as old as philosophy itself. Ironically, the general modern sense of 'analogy' ("likeness in non-identity") derives not from a single Greek word but, apparently, from two words which had certain similarities of meaning.<sup>4</sup> The word *αναλογια*, or "proportion", was imported from mathematics by Plato to characterize a resemblance of relation between two pairs of terms in the general form "a is to b as c is to d" or, symbolically, a:b::c:d. Meanwhile, Plato was also developing the notion of *παραδειγμα*, or example cases for use by comparison with a lesser-known case "in order to discover the true nature of both".<sup>5</sup> The meaning of *παραδειγμα* was further defined by Aristotle to be a method of "rhetorical induction" or argument by example.<sup>6</sup> This resemblance of particular to particular<sup>7</sup> was kept distinct from four-term relational resemblance in Greek texts, but *αναλογια* and the Aristotelian sense of *παραδειγμα* seemed in Latin works to collapse into the term *analogia* while the Latin *paradigma* came to mean something close to "paradigm".<sup>8</sup> Thus could Seneca the Younger say that concepts are "comprehended by analogy",<sup>9</sup> although, despite Quintilian's objections,<sup>10</sup> there remained those who kept the word *analogia* exclusively to mean *αναλογια*.

In ordinary language we now use 'analogy' to refer generally to likeness between different situations. We do not say "argument by paradigm", and "argument by example" usually refers to reasoning from single examples, used without a target situation, to yield a generalization. What Aristotle analyzed in the *Rhetoric* and in the *Organon* (see section II of this essay) has become the modern concept of analogy. How this happened has been the subject of research by fine scholars in theology, rhetoric, and the

---

<sup>4</sup>Measell, "Classical Bases of the Concept of Analogy", *Journal of the American Forensic Association*, 10, p. 1

<sup>5</sup>Ibid., p. 2

<sup>6</sup>Ibid., pp. 3-5

<sup>7</sup>Lloyd, *Polarity and Analogy: Two Types of Argumentation in Early Greek Thought*, p. 408

<sup>8</sup>Op. Cit., Measell, p. 8; Liddell and Scott *A Greek-English Lexicon*; Lewis and Short, *A Latin Dictionary*; Hesse, "Aristotle's Logic of Analogy", *Models and Analogies in Science*, pp. 130-156 argues that the origin of the modern sense of 'analogy' comes during Aristotle's discussion of likeness in *Topics*

<sup>9</sup>Seneca the Younger, *Ad Lucilium Epistulae Morales*, LXX, tr. R. M. Grummere, Loeb Classical Library (3 vols: Cambridge, Massachusetts: Harvard University Press, 1925) cited in Op. Cit., Measell, p. 9

<sup>10</sup>Quintilian, *Institutio Oratoria*, I, vi, 3, tr. H. E. Butler, Loeb Classical Lib., 1920, cited in Ibid., p. 9; I disagree with Measell that *αναλογια* maps onto the modern notion of "figurative analogy"—Aristotle, *The "Art" of Rhetoric*, II, xx, pp. 273-279, speaks of paradigms which are historical parallels, close comparisons, and fables, so it seems to me that figurative analogy, as well as literal, derives from *παραδειγμα* since fables are metaphorical analogies.

classics, whom I have merely been summarizing.<sup>11</sup> Of course, the story is not so simple because in addition to the general modern meaning of 'analogy' there is a host of more specialized definitions which might apply depending on who is writing about it. The Miller Analogies Test, for example, involves solving for the fourth term in classical analogies. Research on "analogy" in psychology and education has often focused on this type.<sup>12</sup> Thomists and philosophers of theology use 'analogy' to refer to the analogy of names as delineated by Aquinas and Cajetan.<sup>13</sup> Natural language semanticists and philosophers of language concern themselves, for the most part, with analogous word usage.<sup>14</sup> Biologists writing about 'analogy' are usually referring to functionally similar physical features of different species (as opposed to 'homology').<sup>15</sup> Language acquisition theorists study analogy as a process for learning morphology and syntax.<sup>16</sup> 'Analogy' in mathematics refers to similarities between different operations,<sup>17</sup> while in literature it may refer to a "figure of speech" as do 'hyperbole', 'metaphor', and 'simile'.<sup>18</sup> Finally, in cognitive science there is a growing literature on analogy as a form of mental or knowledge representation, in which the representation resembles in some way the object of one's thought (e.g., a diagram) as opposed to thoughts being about things symbolically or applicatively.<sup>19</sup>

But I am not going to examine any of the issues on the foregoing list, at least not specifically. What I propose to discuss is 'analogy' in its general, ordinary language sense, which I construe as analogy, or similarity, between situations. Philosophical inquiries about and attempts to account for analogy have roughly fallen into three categories since classical times, according to James Ross.<sup>20</sup> There is *analogia nominum* (the analogy of meaning or of names), about which Ross writes, *analogia rationis* (analogy of reason), and *analogia entis* (analogy of being, or among things).

---

<sup>11</sup>In addition to Measell and Lloyd, see also the history of the Thomistic concept of analogy in Lyttkens, *The Analogy Between God and the World*; The speculations of Joseph in *Introduction to Logic*, p. 537, seem to be supplanted by Measell's findings.

<sup>12</sup>See, for example, Sternberg, *Intelligence, Information Processing, and Analogical Reasoning*

<sup>13</sup>See Op. Cit., Lyttkens; Burrell, *Analogy and Philosophical Language*; McInerny, *Studies in Analogy*; and Bochenski, "On Analogy", *The Thomist*, 11, pp. 425-447

<sup>14</sup>See Ross, *Portraying Analogy* (which also discusses analogy in religion and in law); MacDonald, "The Philosopher's Use of Analogy", *Logic and Language*; Hofstadter, "Metamagical Themas" (on Analogy), *Scientific American*, Sept. 1981, pp. 18-30; Anttila, *Analogy* and Anttila and Brewer, *Analogy: A Basic Bibliography*

<sup>15</sup>Curtis, *Biology*, pp. 377-378

<sup>16</sup>See Esper, *Analogy and Association in Linguistics and Psychology*; and Op. Cit., Anttila

<sup>17</sup>See Polya, *Induction and Analogy in Mathematics*

<sup>18</sup>Perrin, *Writer's Guide and Index to English*, pp. 235-238

<sup>19</sup>Sloman, "Intuition and Analogical Reasoning", *The Computer Revolution in Philosophy*, pp. 144-176

<sup>20</sup>Op. Cit., Ross, pp. 17-19

*Analogia nominum* is a linguistic problem concerning the extent to which a word, used in different sentences and which cannot really be said to be used equivocally, is used *by analogy* in the different sentences. Ross's hypothesis is that the meaning of a word always depends somewhat on the words it is used with, so that univocity in practice never occurs and word usage is always seen as either univocal or analogous, the latter of which includes metaphor.<sup>21</sup> This theory, Ross feels, does not bode well for truth-conditional semantics,<sup>22</sup> or, as the saying goes, "a cat can only be called a cat by metaphor." His view generalizes Wittgenstein's ideas about "family resemblance" words like 'game', which, Wittgenstein contended, have no unique biconditional to determine their appropriateness in a given context.<sup>23</sup>

*Analogia rationis* is the problem I am primarily considering. It concerns the question of how an analogy (and here again I am referring to its general meaning) achieves its appropriateness. This problem seems more general than that of *analogia nominum* because if one could determine theoretically what the use of analogy requires then the answer could presumably be specialized to cover analogous word usage as well as perceptual recognition and decisionmaking about new situations. It is for this reason that so much recent attention has focused on analogical reasoning in artificial intelligence research.

*Analogia entis* gives analogy the status of a phenomenon that exists between real objects or situations. The question of univocal predication, as a problem in the analogy of being,<sup>24</sup> is not one about whether our words can ever be assigned the same meaning in two sentences but is rather a question about whether relations and properties are language-independent conceptual identities. A central question, and one that is directly relevant to the implementation of analogical reasoning in computers, is whether things which we call similar always have a list of common properties which account for the similarity or whether their similarity, beyond some point, can no longer be analyzed and must itself be expressed as a primitive relation between the items. In my theory of analogy I will assume that similarity can always be explicated without having just to say, "A is similar to B." It is not obvious that this is justified, however, and this problem in the analogy of being, which is perhaps the deepest philosophical issue of the three traditional problems of analogy, is discussed further in the conclusion of this essay.

By understanding analogy as just similarity between non-identical situations, I think it should be clear that we have a general notion, of which the various restricted definitions of 'analogy' and of some other words are special cases. This, I suggest, can eliminate much confusion about the relation of the ordinary language 'analogy' to the *αναλογία* sense of the word, as well as to 'metaphor', 'simile', and 'model'. All are example types for the analogy of situation. Classical, four-term analogies like Sun:Earth::Saturn:Titan are analogies in which there are no more than four different objects in two pairs, but in which there is an equivalent relation between each pair (as opposed to non-classical analogies which may have many objects, more than one resemblance of relation, or property equiv-

---

<sup>21</sup>Ibid.,p. 8

<sup>22</sup>Ibid.,p.2

<sup>23</sup>Wittgenstein,*Philosophical Investigations*,pp. 31-35

<sup>24</sup>See Anderson,*Reflections on the Analogy of Being*

alences on the parts of some objects). Metaphors are special cases in the analogy of meaning, since there is some analogous word usage we would not refer to as metaphorical (like 'plow' as a tool versus 'plow' as an activity<sup>25</sup>), but all metaphorical usage of words is analogous. Likewise, 'simile' and 'model' can be defined using 'analogy', the former being a case in which A is explicitly likened to B in a sentence and the latter being a representation which bears a close and useful analogy to the situation that it models.

The following words also appear to be definable as sub-classes of 'analogy': 'equivalence', 'equality', 'isomorphism', 'homomorphism', and 'association'. The following words might be used as synonyms for 'analogy': 'likeness', 'comparison', 'similarity', 'resemblance', 'correspondence', 'closeness', 'parallel', 'relatedness', 'nearness', and 'proximity'.

The objective of all this discussion about 'analogy' (the word) has been to set the right tone for the analysis of analogy (the concept). The conclusions that I hope have emerged are:

1. Analogy is a very basic concept; its ordinary language meaning is fairly easy to define and that definition reveals that analogy is what is going on in many processes with more specialized names (including metaphor and the paronymous 'analog' and 'analogue'). And,
2. The original meaning of analogy, A:B::C:D, is outdated except for its use on standardized tests. Someone is always trying to revive it<sup>26</sup> as the "proper" meaning of 'analogy' and even suggest that other uses of 'analogy' are reducible to it. But this ignores the fact that 'analogy' derives from *analogia* which developed by analogy between *αναλογια* and *παραδειγμα*. The result of that ancient analogy between two concepts was the formation of a very slippery but powerful idea.

## II. *Analogia Rationis* and the Problem of the Cogent Quality

There are three modes of inference which might be considered basic. Of these, one (deduction) is sound in that its conclusions are always true given the truth of its premises. Another (induction) produces plausible, but not sound, conclusions. And finally, a third (abduction) produces conclusions which are hypothetical, suggesting new possibilities.<sup>27</sup> This last (abduction) is not ordinarily admitted as a part of "logic", but interest in its role in learning and discovery has increased during this century because of the work of philosophers of method like Charles Peirce.<sup>28</sup> Abduction, in fact, appears to be the main rule of inference in hypothetico-deductive reasoning, in which the hypothesis does not necessarily follow from known information but would help to explain that information.<sup>29</sup>

<sup>25</sup>Example from Op. Cit., Ross, p. 7

<sup>26</sup>See, for example, discussion of the argument by analogy in Perelman and Olbrechts-Tyteca, *The New Rhetoric: A Treatise on Argumentation*, pp. 371-398

<sup>27</sup>Op. Cit., Anttila, pp. 13-16

<sup>28</sup>See Peirce, Charles, *Philosophical Writings*, ed. by Justus Buchler (New York: Dover, 1955), pp. 129-134 & 150-156, cited in Op. Cit., Anttila, p. 13

<sup>29</sup>See Churchland, *Matter and Consciousness*, p. 71

The exhaustiveness of deduction, induction, and abduction as basic rules of inference can be established in the traditional domain of the syllogism. In that domain, a *result* ("Socrates is a mortal") is inferred from that which is the known *case* ("Socrates is a man") by means of a *rule* ("All men are mortals") which relates the case and result. Symbolically, a property  $Q$  holds for an individual object  $A$  if (a) there is a property  $P$  such that  $A$  has  $P$  and (b)  $Q$  holds for all individuals  $x$  which have  $P$ . Formally, we have the following:

Deductive Rule of Inference (The Syllogism)

Rule  $P(x) \Rightarrow Q(x)$

Case  $\underline{P(A)}$

Result  $\therefore Q(A)$

The statements above the line in the syllogism are the *premises* and that below is the *conclusion*. The syllogism can be generalized to modus ponens and ultimately to resolution in order to characterize deduction completely.<sup>30</sup> Induction and abduction can be given this symbolic form when one notices that, given the above set of domain objects, there are exactly two other possibilities for assigning statements to the premises and the conclusion, namely:

Induction

Case  $P(A)$

Result  $\underline{Q(A)}$

Rule  $\therefore P(x) \Rightarrow Q(x)$

and

ABDUCTION

Result  $Q(A)$

Rule  $\underline{P(x) \Rightarrow Q(x)}$

Case  $\therefore P(A)$

Now the question remains, given that conclusions based on deduction are always logically valid and those of the other two modes never are, under what circumstances, if any, should we believe conclusions based on induction or abduction? The conventional answer for the case of induction is that the inference of the rule is more powerful the more sets of individuals there are, without any counterexamples, for which the case and result hold true. Thus if  $n$  individuals  $A_1, \dots, A_n$  are confirming instances of  $P(x) \Rightarrow Q(x)$  where  $P(A_i)$  and  $Q(A_i)$  are the case and result, respectively, then we have the enumerative induction,

	$P(A_1) \wedge Q(A_1)$
	$P(A_2) \wedge Q(A_2)$
	⋮
premises	⋮
	$\underline{P(A_n) \wedge Q(A_n)}$
conclusion	$\therefore P(x) \Rightarrow Q(x)$

<sup>30</sup> Genesereth and Nilsson, *Fundamentals of Artificial Intelligence*, Ch. 6, pp. 3-5



The conclusion thus becomes more probable as  $n$  approaches the number of values which  $x$  could take on.

There are problems with this criterion as a test for any given induction. Our judgment about whether the conclusion holds depends on background knowledge about the *projectibility* of the property  $Q$  for which we are trying to make a rule. Nelson Goodman's famous example of a non-projectible property is 'grue', which is defined as applying to all emeralds examined thus far which are green, as well as to all thus far unexamined emeralds which are blue. Since, in reality, emeralds are green, all examined emeralds (a large number) are "grue" also. So if, by induction, we conclude that all emeralds are grue then we expect emeralds examined tomorrow to be blue.<sup>31</sup> Clearly this makes no sense. Grue is not a projectible property. I introduce this example now because I will refer to it in section IV in considering the difference between analogy and enumerative induction.

Abduction, one might say, is made more believable the more properties of the individual  $A$  are explained by the inferred property. So in the more general form for abduction,

$$\begin{array}{l} \text{premises} \quad Q_1(A) \wedge Q_2(A) \wedge \dots \wedge Q_m(A) \\ \quad \quad \quad \underline{P(x) \Rightarrow Q_1(x) \wedge \dots \wedge Q_m(x)} \\ \quad \quad \quad \therefore P(A), \end{array}$$

attributing property  $P$  to  $A$  might be more acceptable as the number  $m$  of properties explained by the conclusion approaches the total number of  $A$ 's properties. This seems to me highly suspect as a general rule, since one can always invent an infinite number of properties for any object which are entailed by  $P(x)$ , but the notion has some intuitive value, perhaps. The question of its justification is related to the "strength of similarity" justification for analogical inference discussed later in this section.

Having discussed these three modes of inference, we might well ask how analogical reasoning fits into this framework. As a first stab, one might assert that analogy is represented by the following, where  $A$  and  $B$  are separate individuals:

$$\begin{array}{l} \text{premises} \quad P(A) \wedge Q(A) \\ \quad \quad \quad \underline{P(B)} \\ \text{conclusion} \quad \therefore Q(B) \end{array}$$

It should be noticed that this reduces to a two-step process in our framework, namely

$$\begin{array}{l} \text{(1) induction} \quad \begin{array}{l} P(A) \\ \underline{Q(A)} \\ P(x) \Rightarrow Q(x) \end{array} \\ \text{(2) deduction} \quad \begin{array}{l} \underline{P(B)} \\ Q(B) \end{array} \end{array}$$

---

<sup>31</sup>Goodman, *Fact, Fiction, and Forecast*, pp. 72-83

But this simple conception of analogy as the inference of a property for an individual by syllogism from a single-instance generalization seems inadequate when we apply the induction criterion. Suppose that  $Q$  is projectible and that we use the number of confirming instances, given that no contradicting ones are known, as a guide to the plausibility of the analogy. Well, there is only one confirming instance out of perhaps many possible individuals over which  $x$  could range, so the criteria for inductive plausibility seem to give us a rather low probability for our generalization-syllogism.

The problem with this verdict on the plausibility of  $Q(B)$  is that it does not jibe with the high degree of plausibility that is attached by humans to *some* arguments from analogy. Such arguments have a quality of cogency which goes beyond the plausibility of a single instance induction. An example should help to illustrate this. Suppose you are an American student in Berlin. Yesterday, you went into East Berlin for some touring, and when you entered at Checkpoint Charlie, the border guard stamped your passport with a day visa. Now here you are the next day, at Checkpoint Charlie, getting ready to enter East Germany again and the border guard asks for your passport. You reason by analogy that she too will stamp your passport, just as the border guard did yesterday. This seems like a cogent analogy--your prediction is likely to be accurate. But now suppose you are that same student traveler boarding a bus in Patras bound for Athens. The first people you notice on the bus are two girls speaking to each other in Italian. You spot two more people as you meander down the aisle, so perhaps you should infer by analogy that they too are Italian. This seems not to be a cogent inference by analogy. You are, after all, in Greece, not Italy, so you would expect most of those on the bus to be Greeks, or perhaps, from your experiences elsewhere in Europe, Americans. But according to the induction criteria, it should be at least as plausible that the next people you see on the bus will be Italians as that today's East German border guard will stamp your passport like yesterday's did. In the bus example you have not just one but *two* confirming instances, and the properties of being Italian and of stamping passports are both projectible properties. If, for instance, in Greece you go on to find that the next five or ten people you see on the bus are Italian then induction starts doing its work, and when two teenagers on the bus begin pointing at you and laughing as more Italians board the bus, you are not surprised when an Italian woman comes over to you and says that you have mistakenly boarded a private, Italian tour bus. So an enumerative induction leads one to conclude, as the number of Italians gets larger, that all others on the bus are Italian just as an increased number of experiences at Checkpoint Charlie would engrain the belief that all East German border guards stamp passports.

Clearly there is a factor in our cogency evaluation for analogies with a small number of confirming instances which goes beyond both that number and whether or not the properties being inferred are non-projectible like "grue". Furthermore, it is this factor which must account for the difference between good analogies and bad ones when both cite only a single instance to support the conclusion. One candidate that has been proposed for this extra factor is the degree-of-similarity between the confirming individual and the one for which we are trying to infer a property by analogy. If two individuals  $A$  and  $B$  share properties  $P_1, \dots, P_p$  and  $A$  also has property  $Q$  then by analogy we infer, in this formulation, that  $B$  also has  $Q$ . The rule of inference now looks like:

	$P_1(A) \wedge P_2(A) \wedge \dots \wedge P_p(A) \wedge Q(A)$
premises	<u><math>P_1(B) \wedge P_2(B) \wedge \dots \wedge P_p(B)</math></u>
conclusion	$\therefore Q(B)$

But if our method for evaluating cogency involves asking how similar A and B are then we have moved away from enumerative induction, for whether  $Q(B)$  can be concluded does not just depend on how probable the generalization " $P_1(x) \wedge \dots \wedge P_p(x) \Rightarrow Q(x)$ " is from a single example but really depends on how many properties are shared by A and B and on whether any dissimilarities between A and B make the analogy a poor one. In this formulation, analogy is possibly a separate type of inference from enumerative induction. Jennifer Trusted posits that induction involves the citing of a large number of *confirming instances* to support the inference that new cases will resemble these confirming instances. But analogy, for Trusted, cites a large number of *similarities* of a case with one or a few other instances in order to infer that the case is also similar to the analogous instance(s) in other respects.<sup>32</sup>

There are problems with this view of analogy also. A given individual may have virtually an infinite number of properties, loosely defined, in common with another. We want to say that some similarities are more important than others and that cogent analogy requires a large number of essential similarities, and no essential dissimilarities, between individuals.<sup>33</sup> To go a step further, we might say that it doesn't really matter *how many* similarities two individuals have, only that the similarities they do have should be enough to imply the existence of the inferred property. So perhaps cogent analogies are really deductive in character rather than inductive: the analogy helps us to verify that a set  $\underline{P}$  of B's properties do in fact imply that B has property Q. Since we find  $\underline{P}$  to be held by A also and that A has property Q we hypothesize that B has Q, and the conclusion is valid if Q follows from  $\underline{P}$  combined with one's beliefs.

This deductive explanation seems to reduce the analogous case to the role of generating hypotheses. One may agree with this but hold that what people really do in accepting conclusions by analogy is to take a leap of faith; that analogies can be cogent without their conclusions logically following from one's beliefs. Such a person might hold that analogies are primarily abductive: their conclusions are accepted (a) if the conclusions are consistent with the properties of the individual or situation which they are about, and (b) if the conclusions help to explain those properties. But if analogy is abductive then we have no method for weighing its hypotheses against other hypotheses which would also explain an individual's properties. We can take the argument by analogy for other minds as an example. I have privileged access to only one mind -- my own. I know to what I am referring when I say that I see green or that I have a pain in my leg. For any given person I can infer by analogy when that person says "I see green" or "I have a pain in my leg" that it is for him or her the same sensation that I have when I say those things. But these statements on the part of another are also explained by all kinds of alternative hypotheses ranging from solipsism to inverted spectrums and wicked conspirators. These alternative

<sup>32</sup>Trusted, *The Logic of Scientific Inference*, pp. 8-14

<sup>33</sup>Fearnside, *About Thinking*, p. 266, takes this view

hypotheses all have one thing in common: they each posit some fundamental difference between what it is like to have privileged access to another mind and what my own mental experience is. The argument that other minds are like mine is based on a real case -- my mind, and I thereby put more probability on the conclusion based on analogy than I do for any other randomly chosen, though equally explanatory, hypothesis. But the abduction principles just do not account for this difference between the plausibility of the analogy versus the plausibility of, say, an inverted spectrum hypothesis.

One way to account for the greater plausibility of the argument based on analogy would be to say that there is no logical explanation for why people choose to believe the conclusion by analogy -- logically, it is no better than any of the other hypotheses. Instead, such a person would argue, the cogency of the argument by analogy can only be accounted for by a psychological explanation. What causes people to accept analogies, by such a theory, is an irreducible habit of associating ideas in a way which, through instinct or motivation, helps people to interact with their environment and to fulfill their human drives. On the other hand, analogies seem to function well even at the highest level of rational debate. Extremely intelligent people use analogies and are convinced by them, and the position that analogies which are considered by a person to be cogent do not have a logical form that incorporates that person's beliefs, is probably question-begging. So I propose that psychological explanation should be considered only as a last resort if no logical form adequate to describe analogy can be found.

The argument that inspired me to write this essay was Judith Thomson's famous analogy for abortion in cases of rape. She assumes, for this argument, that a fetus is in fact a person and therefore has a "right to life" equivalent to that of an adult. The argument concerns whether it is morally permissible for a woman to abort the fetus given that she has become pregnant despite having no intention to risk pregnancy, i.e. when she has been raped. I want to call this situation the *central case* since it is the one about which we are trying to decide. Thomson wants to argue that the woman should be allowed to abort in this case. Her argument cites the following *analogous case*:

"You wake up in the morning and find yourself back to back in bed with an unconscious violinist. A famous unconscious violinist. He has been found to have a kidney ailment, and the Society of Music Lovers has canvassed all the available medical records and found that you alone have the right blood type to help. They have therefore kidnapped you, and last night the violinist's circulatory system was plugged into yours, so that your kidneys can be used to extract poisons from his blood as well as your own. The director of the hospital now tells you, 'Look, we're sorry the Society of Music Lovers did this to you -- we would never have permitted it if we had known. But still, they did it, and the violinist now is plugged into you. To unplug you would be to kill him. But never mind, it's only for nine months. By then he will have recovered from his ailment, and can safely be unplugged from you.'"<sup>34</sup>

Thomson asserts that most people would find it outrageous to require anyone to remain hooked up to the violinist, particularly for a long period.

---

<sup>34</sup>Thomson, "A Defense of Abortion", *Philosophy and Public Affairs*, I, pp. 48-49; This was discussed in Jean Roberts' fall 1983 course at Stanford on "Ethical Theories", in which I was enrolled.

She uses this intuition to argue that, even granting that the fetus has right to life, that right does not supersede a woman's right to decide what happens in and to her body. Prior to reading the argument I was sure that whether abortion is permissible hinges on whether the fetus has a right to life. After reading it I brooded about how the analogous case could possibly be relevant in helping one to decide whether abortion is permissible in the central case, given that the fetus has a right to life. Although I felt that the right to abortion could be justified for other reasons, my intuition before the analogy told me that abortion just cannot be permitted if the right to life is granted to the fetus. Yet I agreed with Thomson that a person should be allowed to be disconnected from the violinist, and my intuition after the analogy was confused.

I wanted to know, before buying the argument, whether there could be any logical basis for changing my position based on the analogy, or whether in succumbing I would merely be falling prey to a clever trick of intuition. Would agreeing that the analogy is apt mean that my view of the central case had been blurred or poorly thought through? Or would it mean that I had bought a straw man which jettisoned exactly those parts of the central case that would make direct argument for abortion hard to support given the right to life? If I analyzed those aspects of the analogous case which were unparallel to the central case, would I find therein the crucial difference which would account for why the analogous conclusion seemed easy to support while the central one did not? If I could not find the relevant disanalogy would I have enough reason to buy the argument? And finally, would analogies have any use at all if people were completely logical, like computers, or would analogies then be unnecessary because all the right conclusions could be reached by implication?

I cannot claim to have found complete answers to all of these questions in the past year and a half. They have been reformulated and I have learned some fairly interesting facts about logic. But I have concentrated on answering one question in particular which the other questions all seem to be circling around, namely, why would any rational person ever be convinced by an argument from analogy? Perhaps no one should be convinced by them, but then how do I account for the fact that people have this intuitive ability<sup>35</sup> to tell the difference between what they regard as good and bad analogies?<sup>36</sup>

Relative to other philosophical problems like induction and the analogy of meaning there have been few theories proposed to answer this question. Despite finding a fair number of journal articles on the subject I am aware of no full-length books in English devoted to this problem as a general phenomenon, and only a few which focus on analogy in limited domains like theology or science. The recent upsurge in artificial intelligence research into analogical reasoning<sup>37</sup> makes examination of the basic issues important, especially if a reappraisal of the assumptions underlying current formalisms suggests new approaches to the unsolved problems.

---

<sup>35</sup>I don't mean by "intuitive ability" that the ability cannot involve reason, only that how this reasoning or judgment works does not have to be explainable by those who use it

<sup>36</sup>This formulation of the question is due to my thesis advisor at Stanford, John Perry

<sup>37</sup>Russ Greiner and Stuart Russell at Stanford were both writing doctoral theses on analogical reasoning as I worked on this essay; and Keith Bellairs at Minnesota, Smadar Kedar-Cabelli at Rutgers, and Kevin Ashley at Massachusetts were all graduate students working on analogical reasoning formalisms for legal expert systems

In searching for a theory of analogy, it makes sense to ask what we are looking for. More specifically, what should one *require* of a theory of analogy? I suggest that a complete theory would give solid answers to questions about analogy's form, its place in logic, its justification, soundness, limitations, purpose, methods for its use, and its nature as a mental activity. What follows is an outline of these questions:

1. Form. A theory of analogy should first tell us what analogy *is* and whether the sense of 'analogy' as a type of reasoning really denotes a clear set of phenomena. So one might ask: What is analogy? Is it one type of inference or more than one distinct type? How can its structure be expressed symbolically? Is there a real difference between traditional rhetoric's literal (direct resemblance) analogies and figurative (more metaphorical) analogies, or is it just a matter of degree?
2. Logic. The theory should indicate the place of analogy among types of inference. Some questions would be: What relationships does analogy have to deduction, induction, and abduction? Is it reducible to any or a combination of them? What is analogy's inferential power relative to that of other modes of reasoning?
3. Justification. A theory of analogy should answer whether one ought to believe arguments by analogy, and if so under what circumstances. Questions connected with this are: What, if anything, makes a cogent analogy? What are the roles of considerations like the degree of similarity between cases, the degree of dissimilarity, number of analogous cases called on for support, strength and amount of detail in the conclusion, explanatory power of the conclusion, and background information?
4. Soundness. The theory should take a position on whether conclusions reached by analogy can ever be logically valid, or, if not, what their logical status is. Questions include: Do analogies have any place in sound reasoning? Are they always dispensable and redundant in making deductive conclusions? Can the analogous case ever be said to prove anything? Or is analogy never more than plausible? Never more than suggestive? A complete sham?
5. Limitations. The theory should give us a way to critically evaluate analogical inferences and to detect bad ones. Some questions would be: What makes a poor analogy? A misleading one? In bad analogies is the disanalogy always what makes the analogy fail? Or can it fail for other reasons?<sup>38</sup>
6. Purpose. Related to answering why and whether people should be convinced by analogies, one would want to know why an arguer would want to use them. One might ask: Why do people invoke analogies in arguments instead of applying reasoning to the central case? Would any valid conclusions not always emerge if they just did the latter? Does analogical argument always contain an implicit generalization? Is it always either a straw man argument or unnecessary?
7. Methods. A theory of analogy, to be useful, should indicate how

---

<sup>38</sup>Professor Urmson suggested to me that if one is unconvinced by an analogy, it may not be that one's reasoning is clouded or that the disanalogy is to blame, but may instead sometimes be that the similarities are just unimpressive

one should go about finding good analogies, if such exist. Questions include: Is it best, in drawing an analogy, to make the analogous case as similar as possible to the central case? What types of situation, given a central case, make good analogous cases? Which don't? And, finally,

8. Nature. The theory should take a position on the status of analogy in the Kantian debate over whether certain modes of reasoning are intuitive and non-logical or whether any useful inference must have a logical structure.<sup>39</sup> Analogy would seem to be a good candidate for an intuitive mode, so if it can be shown to be logical then that should have some impact on the debate. Some questions are: Can analogies that are cogent to humans all be evaluated computationally? Symbolically (digitally)? Is analogy an intuitive mode of inference that cannot be fully explained in logic? How do we account for genuine differences of opinion between people concerning the cogency of a given analogy? Is there an element of randomness in the evaluation?

In reviewing the literature on analogy as a logical problem I have discovered that the first four of the above considerations seem to be of prime importance for distinguishing different theories, while the latter four are very roughly dictated by the positions that a theory takes on the first four. The way in which theories of analogy work appears to be the following: a characteristic logical form for analogy is postulated and is seen to correspond either with one of the three basic modes of inference or with none of them. The standards for what make that form or mode a cogent form of argument are then assumed to hold for analogy, and the theory states under what conditions analogies are sound, or plausible, or suggestive. The status of analogy as intuition or reason usually follows from the above considerations, and the theorist may make comments about when, how, and why analogy should be invoked.

Most theories of analogy are in the spirit of one out of five possible accounts for analogical inference which I want to label: logical necessity, inductive generalization, abductive hypothesis, strength of similarity, and psychological explanation. In reviewing and testing ideas about analogy that have cropped up at various points in the history of thought, it is good to bear in mind that those who have written about analogy have sometimes envisaged fields of application more narrow, and uses of analogy more specific, than I am considering in this essay. What follows is an analysis of the sides which philosophers and others seem to have taken in this debate, but there is no reference list of "-isms" onto which writers about the analogy of reason have projected their ideas. People have explored various aspects of analogy and have often considered only arguments used in science or ethics, or religion or oratory; full articulations of general theories about what makes a good analogy are scarce relative to theories about what types of analogies there are. Nonetheless, each of those who have written about analogy seems to have chosen a set of assumptions which characterize one of the theories presented here. I will therefore be careful not to misrepresent ideas espoused by a given writer on, say, science analogies, as constituting that writer's view about analogy in general. I am aware that people often dislike having someone else put their ideas into categories, but hope that the following classification reveals, with some benefit, the traditions within which philosophers and others who study cognitive activity have been working.

---

<sup>39</sup>Op. Cit., Sloman, p. 144

*Theories Based on Logical Necessity*

The first inclination one might have in trying to explain why we judge conclusions based on analogy to be good or bad might be to conjecture that cogent analogy must entail some form of deduction. Under strong logical necessity theories, an analogical argument or inference is judged cogent if and only if the conclusion is provable from one's knowledge and beliefs. In weaker forms of this position, the conclusion may be judged as cogent to the extent that it is either given a specific, non-zero probability or implied by default rules. So a deductivist might hold, for instance, that what has happened when we are convinced by an analogy is that a set of inferences rapidly calculated in our heads has confirmed the truth of the conclusion, and the function of the analogous case has just been to let us identify a possible theorem or default inference before we prove it to ourselves.

This descriptive view would hold that logical deduction adequately describes how people really, actually are convinced. But I know of no author who has sincerely advocated it. Instead, one might hold the deductivist viewpoint only normatively, proposing that a conclusion  $Q$  from analogy *should* be accepted only when it is logically implied from (a) the data base  $\Delta$  of believed propositions and (b) the set of facts  $\underline{P}$  which are shared by the central and analogous cases,  $B$  and  $A$ . Actually, no philosopher I know of has quite embraced this normative view either, but a few have suggested a validity demarcation for analogies in science. Ernest Nagel gives examples from science in which certain physical problems of one type, like heat conduction, can be solved in lieu of analogous problems in another domain, like gravitation, to yield perfect predictions.<sup>40</sup> Although Nagel feels it is impossible to know in advance whether a model (and hence an analogy) will work for a given domain,<sup>41</sup> he suggests that if identity of logical structure is shown to hold then its predictions are valid, and this is the criterion for a useful analogy as opposed to an analogy which is an "obstacle to fruitful development."<sup>42</sup> Nagel's more general view in his writing with Morris R. Cohen is that the validity of a conclusion from analogy depends on a generalization being true, and as such requires fair sampling and an absence of counter-examples.<sup>43</sup> This is the view that I will call analogy as "inductive generalization" (see "Theories Based on Inductive Generalization").

R. O. Anderson argues that one should distinguish between "conclusive" analogical arguments and those which are merely plausible.<sup>44</sup> In particular, conclusive results are obtained when there is a logical isomorphism between two "different representations of a field". By this is meant, for instance, using predicate calculus as an "analogy" to electronic circuits. So on this view, the conclusion  $Q$  is either logically implied by the shared formal structure between  $B$  and  $A$  or it is not "conclusive". But if this is

<sup>40</sup>Nagel, *The Structure of Science*, p. 109

<sup>41</sup>Ibid., pp. 115-116

<sup>42</sup>Ibid., p. 116; Nagel is cited and interpreted in Achinstein, "Models, Analogies and Theories", *Philosophy of Science*, 31, pp. 329-330.

<sup>43</sup>Cohen and Nagel, *An Introduction to Logic and Scientific Method*, pp. 286-288

<sup>44</sup>Anderson, "Conclusive Analogical Argument", *Revue Internationale de Philosophie*, 23, pp. 44-57



the logical foundation for reasoning by analogy then sound analogical reasoning seems to require that the analogous case A be *redundant* in determining whether Q should be inferred for the central case B. For if one's set of beliefs  $\Delta$  and the set of shared properties  $\underline{P}$  logically imply Q then  $\Delta$  and all of the propositions true about B also logically imply Q, given consistency, and one does not need to examine A or construct the set  $\underline{P}$ .

One who holds to the view that convincing (as opposed to merely plausible) analogies are verified by deductions independent of knowledge about A could point out that this would not make analogy useless. Descriptively, such a person might argue that logical implication is just a check for validity and that it may be more efficient or convenient for people to look at the analogous case A to see whether Q is true of B. But if a validity check is the only way in which one can conclude Q(B) with confidence then one is limited to what can be derived from present knowledge and can never infer Q(B) if it is not decidable from  $\Delta$ . This seems to undercut the whole purpose for using analogy which should be to give one a way to infer Q(B) when  $\Delta$  and the known facts about B are logically incomplete.

For analogy to have this useful place in logic, it needs to be shown that analogical inferences can be valid when the conclusion is not decidable from  $\Delta$  and  $\underline{P}$ . This requires (a) that Q(B) be a sound conclusion given  $\Delta$ , the facts  $\underline{P}$  shared by A and B, and the fact that Q is true of A, and (b) that A should not be logically unnecessary to the proof. Such a set of circumstances can in fact exist. This is proven in part III of this essay for two types of premises, one of which was suggested last year by Julian S. Weitzenfeld<sup>45</sup> and the other of which supports the need for clauses having truth value variables. In general, conclusions by analogy follow soundly from the belief that two or more cases share properties which are adequate to determine whether or not Q obtains. So if A has Q being true, and  $\underline{D}$  is such a "determining structure"<sup>46</sup> and is a subset of the conjunction of shared predicates  $\underline{P}$ , then by analogy one can infer Q(B). However, there are substantial reasons to doubt that this is the end of the story about analogy.

Weitzenfeld asserts that, since "the schema for reasoning by analogy is valid", the uncertainty in analogy is accounted for by uncertainties in the "tacit premise" that Q belongs to a determining structure.<sup>47</sup> But this does not seem to account for how people often use analogies. In particular, the argument for other minds is not explained under such a theory because the whole problem is that a person's actions are *not* enough to tell whether she or he has a mind. The analogy seems very convincing nonetheless.<sup>48</sup> It is not just that determination is uncertain in this case -- I *know* that the similarities between others and myself do *not* determine the presence of a mind in others. So it looks like we will have to continue searching for justifications for analogy, turning now to a very old form of account.

---

<sup>45</sup>Weitzenfeld, "Valid Reasoning by Analogy", *Philosophy of Science*, 51, pp. 137-149

<sup>46</sup>This is Weitzenfeld's term, *Ibid.*, p. 141

<sup>47</sup>*Ibid.*, p. 139

<sup>48</sup>In *Human Knowledge and Its Limits*, p. 482, Bertrand Russell asserts that skepticism on the part of philosophers "is professional rather than sincere."

*Theories Based on Inductive Generalization*

It is often asserted that analogical arguments in practice are not deductively valid but are at most a form of plausible inference. One way to make this point is to note that the logical form of analogy collapses into single-instance, enumerative induction. Aristotle, in the first analysis of arguments by analogy, saw "rhetorical induction" (or "argument by paradigm") as the two-step inductive/deductive process discussed earlier in this section.<sup>49</sup> Its truth, then, depends on an "incomplete" induction<sup>50</sup> -- a universal statement made, in Aristotle's terms, without examining all the known cases in which it could apply. The assumption, then, is that analogy involves a generalization based on one or a few instances, and it should be cogent if and only if the generalization is true, i.e. if the examination of other known cases (perhaps every other known case) sharing the antecedent does not produce a counter-example, and the consequent is a projectible property.

This is a popular position for those who are very skeptical about the value of analogy since, in this account, analogy is a weakly-supported induction and the conclusion is therefore assumed to be less probable than one based on many instances. For example, M. C. Beardsley contends that there is always a hidden generalization in an analogy and that the analogy can be thrown away when that generalization is made explicit.<sup>51</sup> His position that analogy can never be a cogent form of argument but is merely a plausible one<sup>52</sup> takes the hard-nosed Aristotelian view<sup>53</sup> and is like other "analogy is bogus" theories of analogy which dismiss this form of reasoning from any place in logic. The briefly sketched theory of Henry Kyburg<sup>54</sup> claims that analogy is not a primitive but is instead always reducible, oddly enough, to *statistical* induction with a single sample, a view similar to that of Cohen and Nagel who hold that inferring the generalization requires fair sampling and no counter-examples and is probable, if it passes these tests, in proportion to the number of analogues examined.<sup>55</sup> Joseph Agassi argues that analogies which turn out to be right in scientific prediction are eventually turned into generalizations. On his view, *ad hoc* analogies, like "Nerves probably conduct electricity because, like telegraphs, they transmit information", only suggest possible inferences and would not be judged convincing by a scientist until the underlying principle finds other support.<sup>56</sup>

---

<sup>49</sup>Aristotle, "Prior Analytics", *Organon I, II, xxiv*, pp. 515-517

<sup>50</sup>Op. Cit., Lloyd, p. 408

<sup>51</sup>Beardsley, *Thinking Straight*, p. 113

<sup>52</sup>Ibid.

<sup>53</sup>As opposed to Aristotle's more sympathetic view of analogy in his ruminations about likeness in the *Topics*--see "Theories Based on Abductive Hypothesis" for accounts in that tradition

<sup>54</sup>Kyburg, *Probability and the Logic of Rational Belief*, pp. 278-279

<sup>55</sup>Op. Cit., Cohen and Nagel, pp. 286-288

<sup>56</sup>Agassi, "Discussion: Analogies as Generalizations", *Philosophy of Science*, 31, pp. 351-356

So the accounts which follow what, as we saw earlier, is a very obvious two-step breakdown of analogical inference, have generally asserted that since analogy is just a weak form of induction, the logical justification for it must also be weak, and that when analogies are cogent there must be some additional investigation or introspection which bears out the conclusion. As was illustrated earlier with the Italian tourist bus and Berlin border guard examples, the view that analogy can be explained as enumerative induction falls into problems in trying to account for how two separate projectible conclusions, confirmed by an equally small number of instances, can differ so widely in their apparent levels of cogency. An inductivist might try to answer this within the theory of inductive justification by suggesting that in evaluating cogency one really is reminded of other generalizations based perhaps on stronger inductions, and the consistency of these generalizations with the analogical conclusion is really what tells one whether to believe it. This works as a way to explain why, say, a close physical resemblance in size and features between two youngsters does not warrant that, if one is known to be named "Skippy", the other one should be also.<sup>57</sup> Past induction has lead us to generalize that similar looking kids probably do not have the same first name, so this analogy is rejected. Experience does allow us the conclusion that if one is in grammar school then the other probably is also, since generalizations like "Kids of similar sizes are around the same age" support the induction. But the theory has a more difficult time explaining the cogency of some other arguments. Suppose you are the world's first inter-planetary traveler, and while in some distant solar system you spot that staple of modern philosophy, Putnam's twin earth. Orbiting the planet, you are struck by the uncanny resemblance to home -- that same climate, the oceans, the mountains, deserts; but you know that you are light years from the earth. You reason that there must be living things on this planet, even though you cannot see them or any direct evidence that they exist. I suggest that there is no past experience to appeal to here. The conclusion seems very plausible, but it just is not founded on any generalizations from induction based on more than just the earth.<sup>58</sup> Actually, I do not think that making the planet an exact twin earth adds to the intuitiveness of this example. The example works as well if one imagines spotting a planet with oceans and continents and clouds like the earth but with a different "geography" -- this is imaginable. Most of us would still conclude some probability that the planet has life on it, and the probability would be higher than for a planet very dissimilar to the earth. The point about this example is that it is a pure analogy, like the argument for other minds and the existence of God; in each of these only one instance gives us any positive reason for our conclusion (although the other planets in our own solar system give us a number of reasons for a negative conclusion in the similar-planet example).

In short, while some analogies do seem to be inductive in character as opposed to deductive or abductive, if the principle of induction is roughly defined to be that confirming instances increase one's confidence in a generalization then this does not seem to be the principle at work in explaining why a single known example can give great confidence in its prediction. In induction, one's goal is to seek a large number of confirming instances,

---

<sup>57</sup>This is based on an example from Professor Urmson

<sup>58</sup>Except, perhaps, for the similarity principle, which I inveigh against in "Theories Based on Strength of Similarity"

and one sacrifices specificity in the reference class common to the instances because one is looking for a law with a simple antecedent. In analogy, the goal is to minimize the scope of the reference class, thereby gaining a power which enumerative induction gives up because one is not seeking, in analogy, a concise generalization -- the whole point of using analogy may be that a generalization would be too hard to state. So analogy seems to have its own set of problems and advantages which are, respectively, not solved and not justified by single-instance induction accounts.

#### *Theories Based on Abductive Hypothesis*

If an analogy is not primarily based on deduction or induction, perhaps it is based on abduction. One might hold that the real function of most analogies is not to prove things, or even to establish their probabilities, but instead to suggest possible explanations and propositions which are consistent with the known facts about the central case. In like manner does Raimo Anttila take the opinion, in his book on linguistic analogies, that analogy "is weaker than induction" and rather "feeds on abduction".<sup>59</sup>

A similar position underlies both Mary Hesse's treatment of "the logical problem of analogy"<sup>60</sup> and Bertrand Russell's attempt to justify analogical arguments.<sup>61</sup> Hesse treats the problem of justifying analogies as models for suggesting possibilities in science and rejects the idea that validity or even plausibility are required of conclusions based on analogy. Accounting for the cogent quality of an analogy thus requires one to show that "given the choice between a hypothetical explanation based on a model and one which is not, it is *more reasonable* to select the former."<sup>62</sup> This is the task facing an abductivist. For if one's view is that analogies generate scientific hypotheses which merely have to be consistent with and to explain the observed facts about a situation then it is not obvious on what grounds such a hypothesis should be preferred over others.

I discussed this problem for the other minds argument earlier in this section. Russell's theory for how analogy works, which, I think, is clearly an abductive theory, tries to justify the analogical argument for other minds by citing the fact that since particular thoughts of mine cause particular behaviors and that other *conceivable* causal relationships are not in my experience, the type I have seem likely to be the only kind which actually exist.<sup>63</sup> This is equivalent to Hesse's first justification for conclusions based on analogy, namely "inductive support". According to such attempts at justification, not only is the rule that my type of thoughts cause behavior valid, but also there is at least one instance to support the converse, that behavior implies that thoughts like the ones that would explain it in me are its cause in another. *Ergo*, this latter rule is preferable to any other hypothesis which may be consistent with a different causal rule but which is not instantiated

---

<sup>59</sup>Op. Cit.,Anttila,p. 18

<sup>60</sup>Op. Cit.,Hesse,pp. 101-129

<sup>61</sup>Op. Cit.,Russell,pp. 482-486

<sup>62</sup>Op. Cit.,Hesse,p. 101

<sup>63</sup>Op. Cit.,Russell,pp. 485-486

even once in one's experience.<sup>64</sup>

This seems like a dubious argument to me. If I can concoct a model in my imagination which supports the hypothesis that it is some other thing (like a wicked conspirator) which causes the behavior of others then I see no reason why this could not count as an instance in support of the rule that behavior in others implies that a conspirator caused it. Likewise, I see no reason to think that falsifiability is a positive reason, as Hesse claims it is,<sup>65</sup> to prefer the hypothesis based on analogy with my own mind, for the problem brought on by skeptical arguments is that there seems to be no way to check the different hypotheses -- they are all equally unfalsifiable. The criterion of simplicity, which is the last one suggested by Hesse,<sup>66</sup> might give one reason to prefer the model of one's own case to one dreamed up because it asks us to believe nothing of others more than what we believe of ourselves. However, why one should prefer simple theories is not a question as yet answered in logic, and also that criterion may be a double-edged sword for supporters of the argument by analogy for other minds since the analogy asks one to believe that others have minds, whereas many forms of skepticism do not ask that one believe anything about others and suggest that what we see is an illusion, a phenomenon for which there are other examples in our experience.

So if one holds that analogies are abductive in character then their cogency appears hard to justify with the tools we have. One can say that conclusions are more believable the more aspects of a situation they explain, or the fewer there are of other hypotheses which are alternatives to the conclusion, but preferring a conclusion based on analogy to another based on an alternative view seems unjustified if both conclusions do equally well at explaining the known facts. The problem is that alternative explanations are almost always imaginable, so perhaps abduction is too weak to account for the cogent quality of some analogies.

#### *Theories Based on Strength of Similarity*

If analogy is not fundamentally based on principles of deduction, induction, or abduction then perhaps its justification is simply a principle of analogy itself. The most popular traditional account of analogy treats the analogical inference as a primitive, axiomatizing the notion that the greater the similarity one sees between the analogous and central cases, the more one should expect unexamined aspects of the central case to resemble corresponding aspects of the analogous case. Under such theories, analogies are judged cogent if and only if they pass some threshold test for similarity between two cases. This test is usually seen as a separate criterion from that of enumerative induction, because the latter requires a large number of instances of a similarity (as opposed to a large number of similarities) relative to the number possible.

The idea that similarity is a logical justification for analogy seems to have originated in the philosophy of John Stuart Mill. It was he who launched a tradition by declaring what I want to call the "Similarity Principle":

---

<sup>64</sup>Op. Cit., Hesse, pp. 104-106

<sup>65</sup>Ibid., p. 127

<sup>66</sup>Ibid.

"...we conclude (and that is all which the argument of analogy amounts to) that a fact *m*, known to be true of A, is more likely to be true of B if B agrees with A in some of its properties (even though no connection is known to exist between *m* and those properties), than if no resemblance at all could be traced between B and any other thing known to possess the attribute *m*. ... There can be no doubt that every such resemblance which can be pointed out between B and A, affords some degree of probability, beyond what would otherwise exist, in favor of the conclusion drawn from it. If B resembled A in all its ultimate properties, its possessing the attribute *m* would be a certainty, not a probability; and every resemblance which can be shown to exist between them, places it by so much nearer to that point."<sup>67</sup>

The Similarity Principle sounds plausible on first reading (perhaps I should call this the *seductive* theory of analogy) but I think that one can almost completely reject it after close examination. To begin with, the basic idea -- what John Maynard Keynes calls "our instinctive principle that likeness breeds the expectation of likeness"<sup>68</sup> -- is extremely difficult to justify logically. Weitzenfeld writes that "nothing, not even an increase in likelihood, follows from mere similarity,"<sup>69</sup> and I am inclined to agree. One could attempt to give a justification based on statistical induction, but it is not clear at all that even a majority of cases one encounters would support the principle. The main problem for such attempts is an abundance of example analogies which violate the Similarity Principle. In some analogies, the analogues appear to be extremely similar and yet we have no confidence in the analogy. Furthermore, such situations may exhibit more *a priori* similarity than another pair of situations whose analogy provides a very convincing conclusion.

Here is an example of such situations. Suppose you are in Palo Alto on the day of the annual Big Game between Stanford and Cal. As a new resident in town and not quite sure where the game is being played, you emerge from buying some kitchenware at the Town and Country Shopping Center. Standing in the parking lot, you notice a gold Honda Civic on El Camino Real with a blond-haired woman in it. She turns toward the Bayshore at Embarcadero. Seeing another blond-haired woman in a gold Honda Civic, heading toward you on El Camino, you reason that she too will turn toward the Bayshore. Now this is not a cogent conclusion because what type of car one is driving and one's hair and sex are just irrelevant to the direction in which one is going. But there is more similarity between the two car-approach situations than there is between two very dissimilar objects, a walking man with a UC Berkeley sweatshirt, say, and a tacky looking car with a "Go Bears" bumper sticker on it, which also approach the intersection at different times. If the first object turns onto Galvez toward Stanford, it is a good bet the second one will. Clearly background knowledge is coming into the evaluation despite Mill's assertion that the connection between the similar properties and the inferred property is not a necessary part of the justification. A defender of similarity-based accounts could argue that the man and the tacky looking car are in fact more similar

---

<sup>67</sup>Op. Cit., Mill, p. 394

<sup>68</sup>Keynes, *A Treatise on Probability*, p. 222

<sup>69</sup>Op. Cit., Weitzenfeld, pp. 137-138

than the two Honda Civics with blond women in them because the similarity between a UC Berkeley sweatshirt and a "Go Bears" bumper sticker has a very high *weight* in measuring similarity. But the difference between the two analogies appears not to be one merely of degree of similarity but of qualitative significance -- however many properties the man and the car share we can only conclude that they will turn in the same direction because we know that Berkeley people would be traveling to the game on this day, and people traveling to the same destination tend to turn at the same places. Most people would judge the two approaching Civics to be more similar than the man and the tacky looking car, even on Big Game day, given no information about what purpose the similarity judgment is to serve. The evaluation must not only be made relative to the conclusion but must also involve some questioning of what real reason one has, beyond their *a priori* similarity, for deciding that two objects will turn at the same place.

There are also examples to count against the Similarity Principle in which two arguments by analogy share the same central and analogous cases. For instance, suppose you happen to be in a parking lot full of brand new cars -- all the same model, body and color, outside the Ford plant in Dearborn, Michigan.<sup>70</sup> If you look under the hood of one and find a 2.3 liter engine, you might conclude by analogy that the car next to it also has a 2.3 liter engine. But if you see a wadded up piece of paper in the back seat, you probably would not conclude that the next car also has a wadded up piece of paper in its back seat. The two analogies rest on the *same* degree of similarity because they are based on the same situations; yet one analogy is cogent and the other is not. Strength of similarity just cannot account for the difference in cogency because there is *no* difference in similarity.

A remarkably large number of writers, though, have followed Mill in assuming that strength of similarity and an absence of strong dissimilarity must be the keys in determining whether an analogy is any good. Theories which express the probability increase for the conclusion based on a specific *number* of known similarities have been developed by Keynes, who argues for such a theory based on the uniformity of nature,<sup>71</sup> and also by P. R. Wilson,<sup>72</sup> Hugues Leblanc,<sup>73</sup> G. H. von Wright,<sup>74</sup> H. W. B. Joseph,<sup>75</sup> and a host of others working in the tradition of Rudolf Carnap.<sup>76</sup> All of this work is an attempt to give formulas for how probable an inference is based on the number of similarities between the cases and the total number of features each one has. The problem is that one can make up an infinite number of shared properties for any pair of cases. A property can be anything -- distance from a given point in space, having a different number of letters than the word 'green', not having another property, and so forth; and the

---

<sup>70</sup>This example is due to Professor Perry

<sup>71</sup>Op. Cit., Keynes, pp. 222-232

<sup>72</sup>Wilson, "On the Argument by Analogy", *Philosophy of Science*, 31, pp. 34-39

<sup>73</sup>Leblanc, "A Rationale for Analogical Inference", *Philosophical Studies*, 20, pp. 29-31

<sup>74</sup>von Wright, *A Treatise on Induction and Probability*, pp. 264-272

<sup>75</sup>Op. Cit., Joseph, pp. 532-542

<sup>76</sup>See discussion under "Other Theories and Hybrid Theories" on confirmation theory approaches to analogy

hope that this can all be reduced to a finite, let alone a manageable, number of primitive, "real" properties seems to be wistful.

It should follow then that attempts to define a threshold test for whether two situations are "similar enough" to project any properties one may hold onto the other is just not a productive approach to automating the kind of analogical reasoning done by humans. However, early work on analogy in artificial intelligence has for the most part rested on some form of the Similarity Principle. Efforts to measure how similar two things are by counting equivalent corresponding attributes in a frame (e.g., by Patrick Winston<sup>77</sup>) and at requiring situations to match at a preset threshold (e.g. by Jaime Carbonell<sup>78</sup>) were useful investigations that have spawned much work, but my feeling is that the way forward is to reject the Similarity Principle as a guide to good inference and to use similarity only as a first pass indicator of whether two situations might be analogous for the purposes of drawing a particular conclusion.

### *Theories Based on Psychological Explanation*

Against theories which see the cogency of analogy as reflecting itself in logic, many would argue that convincing analogies can only be explained in terms of some extra-logical persuasive factors. For instance, a cognitive dissonance theory of analogy might state that cogent analogies are ones which minimize, against the rejection of the analogical inference, the unhappy contradictions in one's beliefs. A behaviorist, on the other hand, might say that analogy is a conditioned response made more likely by greater neural association between concepts.<sup>79</sup>

Psychological explanation does not seem to be needed in order to explain every analogy, since valid analogies are possible. Nor would one want to say that one's being convinced by an analogy can never be the result of motivational or non-logical factors, particularly in arguments about ethics or God. So the type of account, based on psychology, that would be useful is probably one which would demonstrate not only that analogies can be psychologically compelling beyond their logical justification (as J. Hospers, for one, has claimed<sup>80</sup>) but that the habit of analogizing is a necessary component of intelligence despite the failure of logic to justify such analogies.

If a psychological account is primarily what is needed in order to explain why we judge some analogies as good and others as bad then there are important consequences for artificial intelligence and cognitive science. The debate over whether logic programming and symbol processing are adequate approaches both in describing and in creating intelligence is surely a fas-

---

<sup>77</sup>Winston, "Learning and Reasoning by Analogy", *Communications of the Association for Computing Machinery*, 23, pp. 689-703

<sup>78</sup>Carbonell, "Derivational Analogy and Its Role in Problem Solving", *Proceedings of the National Conference on Artificial Intelligence (AAAI-83)*, pp. 64-69

<sup>79</sup>For a discussion of psychological theories of persuasion, see Ross, *Persuasion: Communication and Interpersonal Relations*

<sup>80</sup>Hospers, J., *An Introduction to Philosophical Analysis* (Routledge and Kegan Paul), 1960, p. 355, cited in Op. Cit., Wilson, p. 34



inating one,<sup>81</sup> and I would like to discuss the analogy between computers and brains, with which it is connected, in part IV of this essay. But for now, we can say that it is not known at what level (if any) cognitive processes are isomorphic with those governed by the rules of presently-defined logic, and so it may be a mistake to assume that analogies are more than just neural phenomena which, like seeing and hearing, have proven beneficial in the evolution of humans.

One who held that the human inclination to accept analogies is just a habit of the brain was David Hume. In the *Treatise of Human Nature*, Hume argued that people have a "propension" to be led to opinions like those they have formed about other situations when their *perceptions* of the two situations resemble each other.<sup>82</sup> Hume seems to be advocating that the strength of such resemblance is what accounts for our judgment about analogy,<sup>83</sup> but of course his associationist account does not attempt to justify this principle. Instead, what we regard as similar objects of thought are just plain, unconnected objects for Hume. That people believe they have anything to do with each other is therefore just an "inexplicable fact"<sup>84</sup> about how our minds work; it has no justification.

Hume's idea that people apply analogies to new instances out of the force of habit is shared and extended by W. V. Quine and J. S. Ullian, who argue that in reasoning by analogy we "slur over" any generalization and that plausibility is achieved by the *virtues* that a belief can have: conservatism, generality, simplicity, refutability, and modesty. They see our ability to make good analogical inferences as determined in part by natural selection.<sup>85</sup> Quine and Ullian do not assert that analogy is a phenomenon outside of logical justification and in fact their view has also been likened to theories based on inductive generalization.<sup>86</sup>

There are some criticisms to be made of those who hold that an extralogical explanation for the cogency of analogy is necessary. Firstly, it can be pointed out that most such "theories" are not really accounts of analogy at all -- many just point in the direction of an account and malign analogical arguments for their supposed lack of validity. While one can condemn the use of analogy as an "intuition pump",<sup>87</sup> the leap to assuming that analogy cannot have a justification seems itself unjustified. For instance, one may admit that association occurs in some way prior to logic in the mind, but this does not mean that a logical account cannot be given for cogent analogies.

In short, if a theory can be constructed for how analogy can be jus-

---

<sup>81</sup>See Israel, "The Role of Logic in Knowledge Representation", *IEEE Computer*, October 1983, pp. 37-41

<sup>82</sup>Hume, *Treatise of Human Nature*, p. 142 & 208-209

<sup>83</sup>Laing, *David Hume*, p. 137, shares this interpretation

<sup>84</sup>Laird, *Hume's Philosophy of Human Nature*, pp. 94-95

<sup>85</sup>Quine and Ullian, *The Web of Belief*, pp. 42-62

<sup>86</sup>See Shaw and Ashley, "Analogy and Inference", *Dialogue: Canadian Philosophical Review*, XXII, p. 423

<sup>87</sup>Dennett and Hofstadter, *The Mind's I*, p. 375

tified in logic then it will be useful independent of whether it represents exactly how the mind works. If there are cogent conclusions from analogy that cannot be explained by such a theory then demonstrating this with examples is something the advocates of psychological explanation can shoot for.

#### *Other Theories and Hybrid Theories*

Some recent theories of analogy incorporate aspects of more than one of the above types of account and may offer alternative explanations for what makes a good analogy.

In particular, a major philosophical effort in this century, which is apparently still going on, has been the attempt to incorporate analogical inference into confirmation theory, following on the "inductive logic" of Rudolf Carnap. In *Logical Foundations of Probability*, Carnap defined a confirmation function which could give the increase in the probability that a conclusion holds for the central case B, given that it holds in the analogous case A.<sup>88</sup> This seems to require one to count the number of properties which characterize A, B, and their intersection as well as the number of properties which logically entail these sets. I have already remarked that the idea of using measures of similarity based on numbers of properties is a highly questionable one.<sup>89</sup> Nonetheless, there have been many who have worked on this approach, including Peter Achinstein (1963),<sup>90</sup> Carnap (1963),<sup>91</sup> Hesse (1964),<sup>92</sup> Achinstein again (1964),<sup>93</sup> Ilkka Niiniluoto,<sup>94</sup> Wolfgang Spohn,<sup>95</sup> Domenico Constantini,<sup>96</sup> and T. A. F. Kuipers.<sup>97</sup>

The basic idea in the confirmation theory approach is that individuals or situations A and B are characterized by conjunctions of predicates applied to them and that cogency corresponds roughly to confirmation. Confirmation in turn corresponds to the probability of the conclusion increasing given the analogous case, so that if  $F_1$  is the conjunction of shared predi-

---

<sup>88</sup>Carnap, *Logical Foundations of Probability*, pp. 569-570

<sup>89</sup>Op. Cit., Leblanc, p. 30, says that the issue of how much likelihood increases due to shared properties "is academic, since rules for reckoning how many features a thing has, and how many of those features another thing doesn't have, are (to say the least) still wanting."

<sup>90</sup>Achinstein, "Variety and Analogy in Confirmation Theory", *Philosophy of Science*, 30, pp. 207-221

<sup>91</sup>Carnap, "Discussion: Variety, Analogy, and Periodicity in Inductive Logic", *Philosophy of Science*, 30, pp. 222-227

<sup>92</sup>Hesse, "Analogy and Confirmation Theory", *Philosophy of Science*, 31, pp. 319-327

<sup>93</sup>Achinstein, "Models, Analogies, and Theories", *Philosophy of Science*, 31, pp. 328-350

<sup>94</sup>Niiniluoto, "Analogy, Transitivity, and the Confirmation of Theories", *Applications of Inductive Logic*, pp. 218-234; and "Analogy and Inductive Logic", *Erkenntnis*, 16, pp. 1-34

<sup>95</sup>Spohn, "Analogy and Inductive Logic: A Note on Niiniluoto", *Erkenntnis*, 16, pp. 35-52

<sup>96</sup>Constantini, "Analogy by Similarity", *Erkenntnis*, 20, pp. 103-114

<sup>97</sup>Kuipers, "Two Types of Inductive Analogy by Similarity", *Erkenntnis*, 21, pp. 63-88

cates and  $F_2$  is that conjunction conjoined with  $m$  other predicates also true of  $A$ , then

$$\Pr[F_2(B) \mid F_1(B) \wedge F_2(A)] > \Pr[F_2(B) \mid F_1(B)],^{98}$$

where the increase in probability comes from some variant of the Similarity Principle. In fact, the confirmation theories of analogy are very much like traditional strength of similarity theories except that they are part of an elaborate inductive calculus. In addition to the objections to similarity criteria which I have already stated, I would argue that the above formula is violated too often to be a good assumption. One could certainly imagine situations in which  $F_2(A)$  would give no useful information (recall the two similar looking children, one of whom is named "Skippy") so that the two probabilities would really be equal, but I also think the inequality could go the other way -- that the probability of the conclusion given the analogous case could be *less* than its *a priori* probability. For instance, suppose  $m=1$  and the property conjoined to  $F_1$  to form  $F_2$  is the property of being 93 million miles from the sun. Now if  $A \neq B$  and  $A$  and  $B$  are both planets then we *know* that the probability is now zero that  $B$  is 93 million miles from the sun because  $A$  has that property and there is only one planet for which it holds true. Similarly, suppose two of the people in the crowd at a Stanford Debate Club public forum are named "Jones" and the other is named "Smith", and all three of those in the crowd have brown hair. Now given that  $A$  in the crowd is a Jones and has brown hair and  $B(\neq A)$  in the crowd has brown hair,  $\Pr[\text{Jones}(B)]$  is *less* than it was before we knew that  $A$ 's name was Jones (down to one-half from two-thirds).

Spohn recently sounded the toll of skepticism about the usefulness of inductive logic in exploring analogy.<sup>99</sup> The problem is that it requires too many assumptions (including the Similarity Principle) which are just blatantly violated in practice. Some of the work on analogy which has appeared in the last few years has been an attempt to do away with such assumptions and to concentrate on how the projection of properties from the analogous case (sometimes called the "base" or "source") to the central case (or the "target") are constrained. Cognitive psychology and artificial intelligence have produced theories based on "structure mapping" in which the projection of properties is often constrained by rules.<sup>100</sup>

Theorists of argumentation logic have sometimes adverted to *relevance* as the crucial criterion for inferring a conclusion from similarities with an analogous case. That such similarities should be relevant to the conclusion has often been stated in logic texts,<sup>101</sup> and recently W. H. Shaw and L. R. Ashley have suggested that analogical arguments do not rely so much on resemblances as on background information about the conditions of

<sup>98</sup>Op. Cit., Niiniluoto, "Analogy and Inductive Logic", pp. 6-7; The vertical bar in the probability expression means "given that..."

<sup>99</sup>Op. Cit., Spohn

<sup>100</sup>See Burstein, "A Model of Learning by Incremental Analogical Reasoning and Debugging", *Proceedings of the National Conference on Artificial Intelligence (AAAI-83)*, pp. 45-48; Gentner, "Structure Mapping: A Theoretical Framework for Analogy and Similarity", *Proceedings of the Fourth Annual Conference of the Cognitive Science Society*, pp. 13-15, cited in Burstein above; Gick and Holyoak, "Schema Induction and Analogical Transfer", *Cognitive Psychology*, 15, pp. 1-38

<sup>101</sup>See, for example, Copi, *Introduction to Logic*, p. 360

relevance.<sup>102</sup> I think this is the right direction in which to travel if one is seeking a workable theory of analogy, but there are problems with the picture of analogy painted by Shaw and Ashley. The disanalogy can be relevant to the conclusion, too, and so can facts about the central case which are not known. Furthermore, Shaw and Ashley do not define relevance. Perhaps it really just means that the conclusion follows from the similarities, in which case we are back where we started -- at logical necessity -- and the analogy looks redundant again. I shall attempt to solve these and other problems in the account which follows.

### III. A Situational Theory of Analogy

Language tells us what there is: *Individuals*<sup>103</sup> are the subjects and objects of our discourse. *Relations* and *properties* are what predicate upon them. Sentences refer to *locations* in space and time. They describe situations and *events* which may or may not be real.

The theory of situations developed by Jon Barwise and John Perry<sup>104</sup> suggests that these facts about language are not incidentally related to the universe that language is about -- they are a natural ontology for it. And if we really do see the world the way our language indicates we do, then situation theory seems like a good framework within which to build a theory of analogy because analogies, as nearly as I can tell, are just similarities between different situations.

#### A Summary of the Theory of Situations

Situation theory provides a convenient, precisely defined terminology for describing analogy and similarity.<sup>105</sup> My own experience indicates that, once the terms of situation theory are understood, the concepts essential to analogical inference are more intuitively understandable in the vocabulary of situation theory than are the corresponding concepts in the first order logic associated with it. This should become clearer as the presentation unfolds, since I will use both situation theory and predicate calculus definitions. But because situation theory is relatively new, I need to summarize its basic terminology before proceeding with the theory of analogy.

The primitives of situation theory are *individuals*, *relations*, *locations*, and *polarities*. Individuals are real things, like people and bicycles, and are represented symbolically by  $a, b, c, \dots$ . Relations,  $r^n$ , are  $n$ -place predicates, with 0-ary (situational states), 1-ary (properties), and 2-ary (binary relations,  $n=2$ ) being the most important kinds. Locations,  $l$ , are

---

<sup>102</sup>Op. Cit., Shaw and Ashley, p. 430

<sup>103</sup>The term 'individuals' as used in philosophy means "individual objects" as opposed to "individual persons"

<sup>104</sup>Barwise and Perry, "A Theory of Situations" (Part B), *Situations and Attitudes*, pp. 47-116

<sup>105</sup>I will for the most part stick to the vocabulary of Ibid. (Part B of *Situations and Attitudes*). Perry, Barwise, and a host of others in a working group at the Center for the Study of Language and Information at Stanford have revised the theory quite a bit since the book came out in 1983, but the older, more familiar theory seems adequate for my purposes.

regions of space-time, with the universal location  $l_u$  containing every other location. Polarities,  $i$ , are Boolean values, 0 or 1.

A *situation-type*,  $s$ , is an assignment of polarities to a conjunction of *constituent sequences*,  $y$ , where each  $y$  consists of a relation predicating on a different number  $n$  of individuals, viz:  $r^n, a_1, \dots, a_n$ . So an atomic proposition, for instance, is a constituent sequence assigned polarity 1 (= "yes" or "is the case"). For example, a clumsy student having lunch with a professor and his wife would be represented as the situation-type

$s :=$  professor,  $a$ ; yes  
 student-of,  $b, a$ ; yes  
 husband-of,  $a, c$ ; yes  
 eating-with,  $b, a, c$ ; yes  
 graceful,  $b$ ; no.

A *state of affairs* or *event* is a situation-type occurring at a location, and a *course of events* or *coe*,  $e$ , is a set of events occurring in a spacio-temporal sequence having certain relations between the locations. The coe is *factual* if it accurately describes an *actual* coe as far as it goes, and is non-factual if it gets something wrong.<sup>106</sup> For example, a factual, though probably not a complete, description of an actual coe might be

in  $e$ : at  $l$ :  $s$   
 at  $l'$ :  $s'$   
 temporally-precedes,  $l, l'$ ; yes

where  $s$  and  $s'$  are situation-types, and "at  $l$ :  $s$ " is an event. An *event-type*,  $E$ , is an abstraction of a coe in which any of the primitives for any event in the coe may be *indeterminate*. Indeterminates are represented by  $r$ ,  $a$ ,  $l$ , and  $i$  where these refer, respectively, to any relation, individual, location, or polarity instead of denoting a particular one. So, for example, if

$E :=$  at  $l$ : professor,  $a$ ; yes  
 student-of,  $b, a$ ;  $i$

then  $e$  would be of type  $E$  and  $a$ ,  $b$ ,  $l$ , and 1 (= "yes") would be *anchors* for  $a$ ,  $b$ ,  $l$ , and  $i$ . Furthermore,  $E$  could be viewed as *abstracting* a coe in which  $a$ ,  $b$ ,  $l$ , and  $i$  are replaced (anchored) by  $a$ ,  $b$ ,  $l$ , and 1. Event-types can be and-ed together to form more descriptive event-types (e.g., the event-type  $E$  and  $E'$ ) as long as the combined event-type is consistent. The component event-types are then each *part of* the combined one. A *schema*,  $S$ , is an inclusive disjunction of event-types  $\{E_1, \dots, E_n\}$ .

In situation theory, facts about particular objects are represented by coes, while events of different types are related by *constraints*,  $C$ . A constraint is a rule which tells what one may logically conclude about two schemata, e.g. having lunch involves eating it, spilling it, or both:

$E_1 :=$  at  $l$ : person,  $a$ ; yes  
 having,  $a, b$ ; yes  
 lunch,  $b$ ; yes

<sup>106</sup>In *Ibid.*, pp. 49-51, Barwise and Perry distinguish between *abstract* coes  $e$  and *real* coes  $e$ , where the former are descriptions within the theory and the latter are the parts of the world which correspond to actual coes; For my purposes, all discussion of situations is entirely descriptive.

$$E_p := \text{at } l : \text{spilling}, b; \text{yes}$$

$$E_g := \text{at } l : \text{eating}, b; \text{yes}$$

$$S := \{E_p, E_g\}$$

$$C := \text{at } l_u : \text{involves}, E_p, S; \text{yes}$$

where  $C$  is the constraint.<sup>107</sup>

### The Definition of Analogy

Having reviewed the parts of situation theory of which I will make use, the first task in describing my theory of analogy should probably be to say just what analogy corresponds to in situation theory.

Two coes, which I will henceforth just call "situations" or "cases", are similar if they have a shared event-type (i.e., if there is an event-type of which they are both type) with determined (all anchored) polarities. I will assume that the event-type under discussion contains no relation indeterminates, so the requirement that it also have no polarity indeterminates means that it will contain at least one literal.<sup>108</sup> The representations of our knowledge about the two situations will always fall short of describing everything there is to know about the situation, if it is real, so when I write of a similarity existing between two cases, I mean a known, or encoded, similarity, which is only part of the total similarity between the two real situations.

Analogy also requires that the situations be non-identical, so they must either differ in location or there must be an abstracted situation-type which is known to be part of one situation and known not to be part of the other. The known disanalogy will therefore be a single event-type with indeterminate polarities assigned to each constituent sequence which is a "yes" (1) in one situation and a "no" (0) in the other. If we assume that the real situations are correctly described, then these are the abstracted facts which are anchored to true in one situation only. Furthermore, if the analogy is to be useful to us then there should be another event-type which is anchored in the analogous case and undetermined in the central case, onto which we want to project at least part of the anchored event-type (the conclusion).

In the definitions below, the analogous and central cases should be taken to mean what is known about the two situations both from their initial descriptions and from any constraints that have been "fired" (invoked) to yield new facts about them.

**Definition 1:** An analogy exists between a central case  $e_C$  and an analogous case  $e_A$  iff (i) there is an event-type  $E_S$  (the *similarity*), with anchored polarities for all of its constituent sequences, such that  $e_A$  and  $e_C$  are both of type  $E_S$ , and (ii) there is an event type  $E_D$  (the *disanalogy*), with indeterminate polarities, such that  $e_A$  and  $e_C$  anchor  $E_D$  to opposite polarities for each constituent sequence in  $E_D$ .

<sup>107</sup>Ibid.

<sup>108</sup>A literal is a term from predicate calculus which corresponds to a constituent sequence with a specific truth polarity (0 or 1) assigned

The non-identity requirement (ii) may be fulfilled by differences in location or individuals, or both, in addition to explicit polarity differences, when an opposite polarity is deducible for some anchoring. Thus if Mimi is Syd's mother and Bibi is Gus's mother and we know that the two are not equal and that each person has only one mother, we can conclude a dissimilarity for analogue individuals Syd and Gus which says that Mimi is the mother in one situation and Bibi is the mother in the other.

Merely having an analogy does not make an analogical inference possible. For that, we need to be able to project at least some abstracted facts about  $e_A$  onto  $e_C$ . Therefore, we have,

**Definition 2:** An *argument* may be made *by analogy* between  $e_A$  and  $e_C$  iff (i) there is an analogy between  $e_A$  and  $e_C$ , and (ii) there is an event-type  $E_U$  (the *unknown*) such that  $E_U$  is completely anchored in  $e_A$  (no indeterminates) and is undetermined for  $e_C$  (not completely anchored). The unknown,  $E_U$ , must be of type  $E'$  (the *conclusion*), which is the event-type whose anchoring in  $e_A$  is projected onto its undetermined portion in  $e_C$ , i.e. the indeterminates in  $E'$  which are unanchored in  $e_C$  prior to the argument become anchored in  $e_C$  as they are in  $e_A$ .

### Determination Constraints

I should now be in a position to tackle the fundamental problem of analogy with which this essay is concerned. It may be stated thusly: Given an analogy  $\{e_C, e_A, E_S, E_D\}$ , a data base  $\Delta$  of constraints, and an unknown  $E_U$ , under what conditions might any part  $E'$  of  $E_U$  be anchored in  $e_C$  for its undetermined parts as it is anchored in  $e_A$ ?

The problems which need to be solved in answering this question are those which arise from the theories considered in part II. Firstly, it should not be the case that the justification condition is that the anchored conclusion for  $e_C$  is provable from  $\Delta$  and  $E_S$  or from  $\Delta$  and  $e_C$ , or even that  $E_S$  is provable from the anchored  $E'$  and  $\Delta$ , for then one would not need  $e_A$  except perhaps to identify a possible  $E'$ . Yet it is not that the conclusion  $E'$  should be based *just* on the fact that it occurs in  $e_A$ , for then any conclusion by analogy would be equally good. There should be some *test condition*, other than logical implication, for deciding whether or not the conclusion is cogent, and the degree of similarity between  $e_A$  and  $e_C$  does not appear, from earlier investigations, to be a good candidate.

Human judgments about the cogency of an analogy seem to center around how "relevant" the similarity  $E_S$  is in deciding whether the conclusion  $E'$  holds. Moreover, we may require that  $E_S$  contain a set of conditions which are more or less exhaustively relevant, or sufficient to decide the question. But how does one test for relevance or whether  $E_S$  is sufficient to determine  $E'$ ? I propose that we should simply add the relation 'determines' as a constraint verb like 'involves', and define it as follows:

**Definition 3:** An event-type  $E$  *determines* an event-type  $E'$  iff for every complete anchoring of the set  $\underline{u}$  of unshared indeterminates in  $E$  (those indeterminates which do not appear in  $E'$ ) there is a unique complete anchoring of the set  $\underline{z}$  of unshared indeterminates in  $E'$  which holds for all complete anchorings of the set  $\underline{x}$  of indeterminates which are shared by the two

event-types.<sup>109</sup>

I think that this definition is correct, but it certainly requires some explaining. Perhaps the easiest way is with an example. Let us assume that the make and design of a car and whether or not it is American-made are jointly sufficient to determine that car's engine-type and whether or not it is fuel efficient. Specifically, we have

$C := \text{at } l_u : \text{determines}, E, E'; 1$

where  $E := \text{at } l : \text{car}, a; 1$   
            $\text{make}, a, b; 1$   
            $\text{design}, a, c; 1$   
            $\text{American}, a; i_1$

and  $E' := \text{at } l : \text{car}, a; 1$   
            $\text{engine-type}, a, d; 1$   
            $\text{fuel-efficient}, a; i_2$ .

In this example, the set  $\underline{x}$  of shared indeterminates is  $\{l, a\}$ . The set  $\underline{y}$  of other indeterminates in  $E$  is  $\{b, c, i_1\}$ , and the set  $\underline{z}$  of other indeterminates in  $E'$  is  $\{d, i_2\}$ . A complete anchoring for  $\{\underline{x}, \underline{y}\}$  is a set of values for these indeterminates, so the definition above simply says that if the determination constraint holds and one knows the anchoring for  $\{\underline{x}, \underline{y}\}$  in a situation, then there is only one possible anchoring for  $\underline{z}$  even though we might not know what that anchoring is. A corollary to the definition would be that if  $E$  involves  $E'$  then  $E$  determines  $E'$ , so the set of ordered pairs of event-types in the extension of 'involves' is a subset of those which stand in the relation of determination.

#### Detail of an Argument by Analogy

Before defining the theory more rigorously, it is probably worthwhile to demonstrate exactly how a particular (simple) argument by analogy can be constructed using determination. The argument concerns two Frenchmen, Pierre and Francois. Last year, Pierre came to the United States to become a research associate at Stanford. He was required to get a visa from the U.S. Government before he came. This year, his friend Francois is coming from France to attend Berkeley as a graduate student. Francois concludes, by analogy, that he too will need a visa.

Francois's situation is the *central case*:

in  $e_C$ :  $\text{at France-1985: nationality}, \text{Francois}, \text{French}; 1$   
            $\text{visiting}, \text{Francois}, \text{U.S.}; 1$   
            $\text{student}, \text{Francois}; 1$   
            $\text{school}, \text{Francois}, \text{Berkeley}; 1$

The *unknown* is whether Francois needs a visa:

$E_U := \text{at France-1985: needvisa}, \text{Francois}; i_1$

The *analogous case* is his friend Pierre's, last year:

<sup>109</sup>The 'determines' constraint requires that no constituent sequence in  $E'$  both contain an indeterminate individual and be assigned an indeterminate polarity. This is a technical decision which is too complicated to discuss here but which should not hamper us very much if at all. I have yet to decide whether this assumption is necessary.



in  $e_A$ : at France-1984: nationality,Pierre,French;1  
 visiting,Pierre,U.S.;1  
 student,Pierre;0  
 school,Pierre,Stanford;1  
 needvisa,Pierre;1

So, between them, the strongest *similarity* we can state (without firing constraints on their respective situations) is:

$E_S := \text{at } t$ : nationality,a,French;1  
 visiting,a,U.S.;1  
 school,a,b;1

The *disanalogy* is the set of facts which are true in one situation but not in the other:

$E_D := \text{at } t$ : student,a; $i_2$

Of course, the actual disanalogy will be virtually infinite, but in the absence of further knowledge or information provided by constraints,  $E_D$  is the full provable dissimilarity. The disanalogy only plays a role if the argument is not sound, in which case it needs to be enumerated more completely than in this example.

Now suppose we know that whether you will need a visa is determined, in general, by which country you are from and which one you are going to. We know the *determination constraint*:

$C := \text{at } t_u$ : determines,E,E';1

where the *determinant* is:

$E := \text{at } t$ : nationality,a,c;1  
 visiting,a,d;1

and the *conclusion* is:

$E' := \text{at } t$ : needvisa,a; $i_1$ .

To make the argument, we verify that  $E_U$  is of type  $E'$  and that  $E_S$  is of type  $E$ . The sets of shared indeterminates ( $\underline{x}$ ), of unshared indeterminates in  $E$  ( $\underline{y}$ ), and of unshared indeterminates in  $E'$  ( $\underline{z}$ ), are  $\underline{x} = \{t,a\}$ ,  $\underline{y} = \{c,d\}$ , and  $\underline{z} = \{i_1\}$ . Anchoring the determinant's set  $\underline{y}$  to {French,U.S.} as it is anchored in  $E_S$  gives the *antecedent*:

$E'' := \text{at } t$ : nationality,a,French;1  
 visiting,a,U.S.;1

and anchoring the conclusion event-type's set  $\underline{z}$  to {1}, as dictated by  $e_A$ , gives the *consequent* (or anchored conclusion):

$E''' := \text{at } t$ : needvisa,a;1

which sets up a new "involves" constraint (the *generalization*):

$C' := \text{at } t_u$ : involves,E'',E''';1

Since  $e_C$  is of type  $E_S$ , and therefore of type  $E''$ , it must by  $C'$  be of type  $E'''$ , so  $\underline{x}$ , the indeterminates in  $E'''$ , can be anchored for  $e_C$  to {France-1985,Francois} and we can say

in  $e_C$ : at France-1985: needvisa,Francis;yes.

This example illustrates the general mechanism of analogical inference when determination constraints are known for given similarity and conclusion event-types. One possible criticism of the approach is that we may not know any such constraint for a particular analogy. I do not assume that all reasoning by analogy entails a deductive application of a determination rule. Other forms are discussed later in this section. But it seems to me that determination is a very useful form of constraint to be aware of, because it is more general than involvement and is often the kind of information we have, or ask whether we have. I do not know, for instance, whether Turkish people need visas to go to Greece, but I have a feeling that they either all need visas or none of them do. Let me talk to a Turk who has been to Greece, and I will tell you which one it is. So determinative knowledge facilitates learning. From a general rule I can collect new, more specific rules as I examine, in the theory's terms, new anchorings for the unshared indeterminates in  $E$  and  $E'$ . Reasoning by analogy in such cases is both valid and indispensable.

### Semantics in First Order Logic

Although situation theory is intended as a framework for a semantics of natural and artificial languages, in this case it seems appropriate to reverse roles and give an interpretation of situation theory and determination constraints in terms of the first order predicate calculus. This will facilitate the proof of soundness and should help to clarify the exact meanings of some terms.

The translation initially requires a few more definitions. To set up the interpretation of event-types and coes in first order logic, I first define a weak notion of truth:

**Definition 4:** A set  $A$  of anchors ("values") for each of the indeterminates in an event-type  $E$  is a *factual complete anchoring* of  $E$  iff  $E$  is anchored to  $A$  ("instantiated") in some actual coe.

Event-types and coes can be defined in terms of first order predicate schemata:

**Definition 5:** A predicate schema<sup>110</sup>  $P(x_1, \dots, x_n)$  is an *associated first order expression*, or *afoe*, for an event-type  $E$  having  $n$  indeterminates iff the extension of  $P$  is equal to the set of all factual complete anchorings of  $E$ .

With the afoe, a predicate calculus interpretation for definition 4 may be given:

**Axiom 6:** An event-type  $E$  with  $n$  indeterminates is anchored by some actual course of events  $e$  (there is an  $e$  of type  $E$ ) iff  $\exists x_1, \dots, x_n P(x_1, \dots, x_n)$ , where  $P(x_1, \dots, x_n)$  is an afoe for  $E$ . Similarly, an event type  $E'$  is of type  $E$  iff an afoe of  $E'$  is coextensive with an afoe of  $E$  anchored for its unshared indeterminates as in  $E'$ .

Intuitively,  $P$  can be thought of as the scaffolding of the event-type: all the determined relations, locations, individuals, and polarities which appear in it, together with their structure. The free variables  $x_1, \dots, x_n$  are the

<sup>110</sup>A predicate schema (not to be confused with a situation theory schema, as defined earlier) is, for my purposes, a predicate on free variables

spaces between scaffolding, and anchors for these are the boards and poles which are placed in the spaces and thereby become part of the structure.

The relation of involvement between event-types corresponds to a kind of implication:

**Axiom 7:** For event-types  $E$  and  $E'$  with  $n$  and  $m(\leq n)$  indeterminates, respectively,  $E$  involves  $E'$  iff  $\forall x_1, \dots, x_m \forall y_1, \dots, y_{n-m} P(x_1, \dots, x_m, y_1, \dots, y_{n-m}) \Rightarrow Q(x_1, \dots, x_m)$ , where  $P(x_1, \dots, y_{n-m})$  and  $Q(x_1, \dots, x_m)$  are afoes for  $E$  and  $E'$ , respectively, and the set  $\{x_1, \dots, x_m\}$  corresponds to the shared indeterminates.

The relation of determination may be seen to correspond to a somewhat more complicated expression in logic:

**Axiom 8:** For event-types  $E$  and  $E'$  with  $n_P$  and  $n_Q$  indeterminates, respectively,  $E$  determines  $E'$  iff  $\forall x_1, \dots, x_m \forall y_1, \dots, y_{n_P-m} \forall z_1, \dots, z_{n_Q-m} P(x_1, x_2, \dots, x_m, y_1, y_2, \dots, y_{n_P-m}) \wedge Q(x_1, x_2, \dots, x_m, z_1, z_2, \dots, z_{n_Q-m}) \Rightarrow [\forall w_1, \dots, w_m P(w_1, \dots, w_m, y_1, \dots, y_{n_P-m}) \Rightarrow Q(w_1, \dots, w_m, z_1, \dots, z_{n_Q-m})]$  for afoes  $P(x_1, \dots, y_{n_P-m})$  and  $Q(x_1, \dots, z_{n_Q-m})$  of  $E$  and  $E'$ .

That involvement is just a special case of determination may now be proven:

**Theorem 9:** If  $E$  involves  $E'$  then  $E$  determines  $E'$ .

**Proof:** Let the number of indeterminates in  $E$  and  $E'$  equal  $n$  and  $m$ , respectively. Map the set of values  $x_1, \dots, x_m$  one-to-one onto a range for a new variable  $x$  and do a similar mapping for  $y_1, \dots, y_{n-m}$  onto  $y$ . By axiom 7 and the hypothesis,  $\forall x, y P(x, y) \Rightarrow Q(x)$ , where  $P(x, y)$  and  $Q(x)$  are afoes for  $E$  and  $E'$  and therefore there are no unshared indeterminates in  $E'$ . A substitution gives  $\forall w \forall y P(w, y) \Rightarrow Q(w)$ , and by quantifier reversal and implication introduction,  $\forall y \forall x P(x, y) \wedge Q(x) \Rightarrow [\forall w P(w, y) \Rightarrow Q(w)]$ , which, by axiom 8, is the condition for determination of event-types for  $n_Q = 0$ .

### Proof of Soundness

To prove the soundness of analogical inference under determination, I will first prove the situation theory analogue of the rule of the syllogism and then prove that an "involves" constraint can be inferred from a determination constraint combined with a single coe. The validity of the argument by analogy as defined in definition 2 follows from the two lemmas.

**Lemma 10:** If  $E$  involves  $E'$  then any actual coe  $e$  of type  $E$  must also be of type  $E'$ .

**Proof:** If  $e$  is of type  $E$  then by axiom 6,  $\exists x, y P(x, y)$  where  $P(x, y)$  is an afoe of  $E$  with  $n$  indeterminates and  $x$  and  $y$  are collapsed variables with ranges equivalent to those of  $\{x_1, \dots, x_m\}$  and  $\{y_1, \dots, y_{n-1}\}$ , respectively. By axiom 7,  $\forall x, y P(x, y) \Rightarrow Q(x)$ , where  $Q(x)$  is an afoe for  $E'$ , so by universal instantiation and modus ponens,  $Q(A)$  is an afoe for  $e$  where  $A$  is the complex object corresponding to the factual complete anchoring of  $\underline{x}$  in  $e$  and where  $e$  is treated as an event-type. Therefore  $e$  is of type  $E'$ .

The byproduct of determined arguments by analogy is a new rule inferred as a generalization from a single instance:

**Lemma 11:** If  $E$  determines  $E'$  and  $e$  is of type  $E$  and  $E'$ , then there is a pair of event-types  $E''$  of type  $E$ , and  $E'''$  of type  $E'$ , with anchors for each of the  $n_P$ -m and  $n_Q$ -m indeterminates unshared in  $E$  and  $E'$  and with the same set of shared indeterminates as  $E$  and  $E'$ , respectively, such that  $e$  is of type  $E''$  and  $E'''$ , and  $E''$  involves  $E'''$ .

**Proof:** By axiom 8 and the hypothesis, for afoes  $P(x,y)$  and  $Q(x,z)$  of  $E$  and  $E'$ , respectively, where  $x$ ,  $y$ , and  $z$  are collapsed as before,  $\forall x,y,z P(x,y) \wedge Q(x,z) \Rightarrow [\forall w P(w,y) \Rightarrow Q(w,z)]$ . Let  $A$  be the anchoring for  $x$ ,  $B$  be the anchoring for  $y$ , and  $C$  be the anchoring for  $z$  in  $e$ . Then by universal instantiation and modus ponens,  $\forall w P(w,B) \Rightarrow Q(w,C)$ , which may be rewritten as  $\forall w P'(w) \Rightarrow Q'(w)$ , where  $P$  and  $Q$  have been replaced by primed predicates which have extensions equal to the range of  $w$ . The implication is a sufficient condition for  $E''$  to involve  $E'''$ , where these two event-types have afoes  $P'(w)$  and  $Q'(w)$ . By axiom 6, since these are coextensive with  $P(w,B)$  and  $Q(w,C)$ , respectively,  $E''$  and  $E'''$  are of types  $E$  and  $E'$ , respectively, and since for  $e$ ,  $\exists w P'(w) \wedge Q'(w)$ ,  $e$  is of type  $E''$  and  $E'''$ .

Finally, the validity of analogical inference can be proven formally under the constraint of determination:

**Theorem 12:** Given an analogy  $\{e_C, e_A, E_S, E_D\}$  with an unknown,  $E_U$ , which is completely anchored in  $e_A$  and is of type  $E'$  (the conclusion event-type): if there is an event-type  $E$  such that  $E_S$  is of type  $E$  and  $E$  determines  $E'$ , then the argument by analogy is sound.

**Proof:** From lemma 11, since  $E$  determines  $E'$  and  $e_A$  is of type  $E_S$  and  $E_U$  and therefore of type  $E$  and  $E'$ , there are event-types  $E''$  (the antecedent) and  $E'''$  (the consequent) which anchor the unshared indeterminates in  $E$  and  $E'$ , respectively, as they are anchored in  $e_A$  such that  $E''$  involves  $E'''$ .  $E''$  anchors the unshared indeterminates in  $E$  to values in  $E_S$ , which is of type  $E$ , so  $E_S$  must also be of type  $E''$ . Since  $e_C$  is of type  $E_S$ ,  $e_C$  is also of type  $E''$ , so by lemma 10,  $e_C$  must be of type  $E'''$ . The anchoring of indeterminates as prescribed in definition 2 has been accomplished by a deductive procedure.

Having given a general account of analogy in situation theory, we might want to ask how determination rules could be expressed in logic for arbitrarily complex expressions, viz when the predicates, constants, and variables are not collapsed representations of event-types but direct representations of properties holding or not holding for objects. It is this task to which I now turn.

### Solution and Decision

The determination relation as I have presented it allows assignment of location, individual, and polarity indeterminates based on a single constraint verb, 'determines'. In traditional predicate logic, however, the fact that clauses cannot have variable truth values (there is no argument place for polarity) means that the effect of having polarity variables must be achieved without actually having them. The allowance for indeterminate polarities seems to be an advantage for situation theory in addition to its perspicuous

ontology for the various combinations of facts and schemata and for the relations between the different forms. But situation theory is a more complicated framework for representing propositions in a data base, and it has not yet been fully axiomatized in the manner of first order logic. So it is probably worthwhile to try to experiment with determination in the more familiar formal language of the predicate calculus, in which it is possible to define a pair of logical connectives to correspond to individual determination, in one case, and to polarity determination in the other.

Individual determination corresponds to what I want to call *functional determination* or *solution* in predicate calculus. It is this type of determination which was anticipated by Weitzenfeld<sup>111</sup> although he did not define it in terms of logical connectives. A new logical operator for functional determination may be defined as follows:

**Definition 13:** A function  $F$  solves another function  $G$  for all  $x$  iff  $\forall x [F(x)=y \wedge G(x)=z] \Rightarrow [\forall w F(w)=y \Rightarrow G(w)=z]$ , which can also be written as  $\forall x \forall y \forall z [P(x,y) \wedge Q(x,z)] \Rightarrow [\forall w P(w,y) \Rightarrow Q(w,z)]$ .<sup>112</sup>

This, of course, looks very much like the definition of determination using afoes, but in this case the variables to be determined are restricted to object variables. Suppose, for instance, that we wanted to split up the determination constraint defined in the example after definition 3 into two parts, namely (a) the make and design of a car determine its engine-type, and (b) whether or not a car is American-made determines whether it is fuel-efficient. The first constraint is a case of functional determination: make and design solve for engine-type. The second is an example of what I call *truth value determination* or *decision*. Its realization in predicate logic is

**Definition 14:** A clause with an unknown truth value  $P?(x)$  decides the truth value of another clause  $Q?(x)$  iff  $\{[\forall x P(x) \Rightarrow Q(x)] \vee [\forall x P(x) \Rightarrow \neg Q(x)]\} \wedge \{[\forall x \neg P(x) \Rightarrow Q(x)] \vee [\forall x \neg P(x) \Rightarrow \neg Q(x)]\}$ . I write this as  $P?(x) \succ Q?(x)$ .

In traditional first order logic, the truth value of a clause (which I will henceforth assume to be synonymous with its polarity) cannot itself be a variable. Given the set of logical operators in predicate calculus, this makes sense because any variable attachment to a predicate's polarity would simply make it a tautology ( $P$  or not  $P$ ). Having polarity variables becomes useful, however, when we define connectives like 'decides' which allow statements about what one needs to know in order to know other things.

A number of little theorems can now be proven which make clear how these connectives operate. Some of these have been pointed out to me by others, as noted.

**Theorem 15:** If  $P(x,y)$  solves  $Q(x,y)$  and there are objects  $A$ ,  $B$ , and  $C$  such that  $P(A,B)$  and  $Q(A,C)$ , then we can infer the functional generalization:  $\forall w P(w,B) \Rightarrow Q(w,C)$ .

**Proof:** Follows by syllogism (universal instantiation and modus ponens) from definition 13.

<sup>111</sup>Op. Cit., Weitzenfeld, p. 141

<sup>112</sup>This formula may look more complicated than one would expect. Unfortunately, simpler formulas like " $\forall x \forall y \exists z P(x,y) \Rightarrow Q(x,z)$ " break down in proofs of the validity of functionally determined conclusions. The conjunctive normal form clause for the solution sentence is " $\neg P(x,y) \vee \neg Q(x,z) \vee \neg P(w,y) \vee Q(w,z)$ ".

The above theorem allows inferences like "All Spaniards speak Spanish" from, for example, the premise that nationality solves for native language, together with an observed case of a Spaniard whose native language is Spanish.<sup>113</sup>

**Theorem 16:** If  $P(x)$  decides whether  $Q(x)$  or  $\neg Q(x)$ , and there is a case  $A$  such that  $P(A)$  and  $Q(A)$ , then we can infer the truth value generalization:  $\forall x P(x) \Rightarrow Q(x)$ .

**Proof:** Assume  $\exists x P(x) \wedge \neg Q(x)$ . Definition 14 can be And-simplified to  $[\forall x P(x) \Rightarrow Q(x)] \vee [\forall x P(x) \Rightarrow \neg Q(x)]$ . The assumption eliminates the first disjunct, and the hypothesis that  $P$  and  $Q$  both hold for  $A$  eliminates the second disjunct, so by *reductio ad absurdum* the theorem is proven.

This justifies the first step (the generalization) in the two-step Aristotelian model of analogy given earlier. The analogous case  $A$  serves to eliminate one of two possibilities: that all  $P$ 's are  $Q$ 's or that they are all not  $Q$ 's. We have this kind of disjunctive knowledge often. For example, I do not know whether history majors at the University of Texas are required to take a foreign language, but it seems likely that being a history major at that University determines which it is.

**Theorem 17:**  $P(x,y)$  solves  $Q(x,z)$  iff  $\forall w,x,y,z P(x,y) \wedge Q(x,z) \wedge P(w,y) \Rightarrow Q(w,z)$ . (Paul Rosenbloom, Stuart Russell)

**Proof:** From definition 13,  $\forall x,y,z P(x,y) \wedge Q(x,z) \Rightarrow [\forall w P(w,y) \Rightarrow Q(w,z)]$ , so taking the " $\forall w$ " outside,  $\forall w,x,y,z P(x,y) \wedge Q(x,z) \Rightarrow P(w,y) \Rightarrow Q(w,z)$ , which can be rewritten as above.

The above theorem suggests an alternative way to prove the validity of analogy, by instantiating  $w$ ,  $x$ ,  $y$ , and  $z$  to the values in the analogues and performing a syllogism.

**Theorem 18:**  $P?(x) \succ Q?(x) \Rightarrow P(x) \succ Q?(x)$ .

**Proof:** Follows from definition 14 by And-simplification.

In practice, the test condition for analogy will be the consequent in this (the above) theorem rather than being the antecedent, since it is known prior to the analogy whether  $P$  holds for the two analogues.

**Theorem 19:**  $P?(x) \succ Q?(x) \Leftrightarrow P?(x) \prec Q?(x)$ . (Jerry Hobbs)

**Proof:** By definition 14,  $P?(x) \succ Q?(x)$  iff  $\{[\forall x P(x) \Rightarrow Q(x)] \vee [\forall x P(x) \Rightarrow \neg Q(x)]\} \wedge \{[\forall x \neg P(x) \Rightarrow Q(x)] \vee [\forall x \neg P(x) \Rightarrow \neg Q(x)]\}$ , where the implications can be written truth functionally as  $\{[\forall x \neg P(x) \vee Q(x)] \vee [\forall x \neg P(x) \vee \neg Q(x)]\} \wedge \{[\forall x P(x) \vee Q(x)] \vee [\forall x P(x) \vee \neg Q(x)]\}$ , which is the same as the truth functional form for the definition of  $Q?(x) \succ P?(x)$ .

This points out that definition 14 for  $P?(x) \succ Q?(x)$  would serve equally well as a definition for  $Q?(x) \succ P?(x)$ . I wrote the definition as a conjunction of disjunctions to show the form that of the general case logical equivalence

<sup>113</sup>This example was suggested to me by Stuart Russell

when the determinant is a conjunction of predicates. If some, but not all, of the predicates have determined polarities then the biconditional (mutual) determination will not hold -- it is only valid in the simple case in which all of the predicates are of variable truth value.

**Theorem 20:**  $F(x)$  solves  $G(x)$  iff  $\forall w, x F(x)=F(w) \Rightarrow G(x)=G(w)$ . (Stan Rosenschein)

**Proof:** By theorem 17,  $F(x)$  solves  $G(x)$  iff  $F(x)=y \wedge G(x)=z \wedge F(w)=y \Rightarrow G(w)=z$ , which is equivalent to  $F(x)=F(w) \wedge G(x)=z \Rightarrow G(w)=z$ . Rewriting,  $F(x)=F(w) \Rightarrow [G(x)=z \Rightarrow G(w)=z]$ , which is true if and only if  $F(x)=F(w) \Rightarrow G(x)=G(w)$ .

In the special case in which the object variables ( $x$  and  $y$ ) range over instants in time, the function  $F$  is the state of a system (e.g., a collection of particles in space) for one possible history, and the function  $G$  is the state for another possible history, the above theorem expresses Montague's condition for theories to allow one to deduce, given the instantaneous state of a system, its state at any time thereafter.<sup>114</sup>

**Theorem 21:**  $P(x) \succ Q?(x)$  iff  $\forall x, y P(x) \wedge P(y) \Rightarrow [Q(x) \Leftrightarrow Q(y)]$ . (Stan Rosenschein)

**Proof:** By definition 14 and And-simplification,  $P(x) \succ Q?(x)$  iff  $[\forall x P(x) \Rightarrow Q(x)] \vee [\forall x P(x) \Rightarrow \neg Q(x)]$ , which has the same conjunctive normal form as the above expansion:  $\neg P(x) \vee Q(x) \vee \neg P(y) \vee \neg Q(y)$ .

Theorem 21 provides another definition for decision which is particularly well suited to analogical inference. Finally,

**Theorem 22:** If  $P(x) \succ Q?(x)$  and  $P(A) \wedge Q(A) \wedge P(B)$ , then  $Q(B)$ .

**Proof:** The premises fulfill the requirement for theorem 16, from the consequent of which  $Q(B)$  follows by syllogism.

Theorem 22 completes the proof of soundness for Aristotle's two-step process of analogy when  $P(x)$  decides  $Q?(x)$ .

The connectives 'solves' and 'decides' are useful for describing analogy in the simplified domain of the syllogism. It is also possible to reconstruct the notion of 'determines' in predicate calculus based on solution and decision. An event-type may be represented as a conjunction of  $m$  clauses ( $m$ =the number of clauses) where each clause is either a literal (anchored polarity) or a two-proposition disjunction (indeterminate polarity) and in which the  $n$ th argument of each  $n$ -place predicate is the location referred to by the clause. Thus the event-type  $E$  defined above would be written thusly:

<sup>114</sup>Montague, "Deterministic Theories" (Ch. 11), *Formal Philosophy*, pp. 303-359; See especially Montague's discussion of why this sentence "cannot be taken quite literally", pp. 303-304; There is a large philosophical literature on determinism which I have not attempted to review, but which is not directly relevant to this essay.

$$\begin{aligned} & \text{Car}(a,l) \\ & \wedge \text{Make}(a,b,l) \\ & \wedge \text{Design}(a,c,l) \\ & \wedge [\text{American}(a,l) \vee \neg \text{American}(a,l)]. \end{aligned}$$

In general, an event-type with  $m$  constituent sequences, of which  $t$  have positive polarities and  $v$  have indeterminate polarities, can be represented as

$$\begin{aligned} & \bigwedge_{i=1}^t P_i(\underline{a}_i, l_{1i}) \\ & \bigwedge_{j=1}^{m-v-t} \neg Q_j(\underline{b}_j, l_{2j}) \\ & \bigwedge_{k=1}^v \{R_k(\underline{c}_k, l_{3k}) \vee \neg R_k(\underline{c}_k, l_{3k})\} \end{aligned}$$

where  $\underline{a}_i$ ,  $\underline{b}_j$ , and  $\underline{c}_k$  are the parametric vectors of individuals predicated upon, respectively, by the  $i$ th true relation  $P_i$ , the  $j$ th false relation  $Q_j$ , and the  $k$ th relation with an undetermined truth value,  $R_k$ , and where  $l_{1i}$ ,  $l_{2j}$ , and  $l_{3k}$  are the location parameters for each of these predicates.<sup>115</sup>

The connectives 'solves' and 'decides' can now be combined in predicate logic to form a definition for 'determines':

**Definition 23:** A conjunction  $E$ <sup>116</sup> with  $m$  clauses,  $t$  of which have polarity one and  $v$  of which have indeterminate polarities, *determines* a conjunction  $E'$  with corresponding parameters  $m'$ ,  $t'$ , and  $v'$  iff

(i) For every possible vector of truth value assignments to the undetermined clauses  $R_k^?$ :

$$\{ \{ \bigwedge_{i=1}^t P_i(\underline{a}_i, l_{1i}) \} \wedge \{ \bigwedge_{j=1}^{m-v-t} \neg Q_j(\underline{b}_j, l_{2j}) \} \wedge \{ \bigwedge_{k=1}^v R_k^?(\underline{c}_k, l_{3k}) \} \}$$

solves each conjunct in

$$\{ \{ \bigwedge_{i=1}^{t'} P'_i(\underline{a}'_i, l'_{1i}) \} \wedge \{ \bigwedge_{j=1}^{m'-t'-v'} \neg Q'_j(\underline{b}'_j, l'_{2j}) \} \}$$

and (ii) For every clause  $R_k^?$  of indeterminate polarity in  $E'$ ,

$$\{ \{ \bigwedge_{i=1}^t P_i(\underline{a}_i, l_{1i}) \} \wedge \{ \bigwedge_{j=1}^{m-v-t} \neg Q_j(\underline{b}_j, l_{2j}) \} \wedge \{ \bigwedge_{k=1}^v R_k^?(\underline{c}_k, l_{3k}) \} \}$$

decides  $R_k^?(\underline{c}'_k, l'_{3k})$ .

Essentially, then, the general case of determination is somewhat awkward to express using our new connectives. It might be noted that location determination becomes just a special case of functional determination in this definition, as long as the location parameter is simply added to each predicate in the translation. Furthermore, the predicate calculus representation of "E determines E'" is exponential in  $v$  and  $v'$ , and depends on the exact form of the definition chosen for 'decides' (i.e., definition 14 or theorem 21)

<sup>115</sup>The vectors or location parameter may contain constant or variable objects

<sup>116</sup>The  $E$  without italics means that I am referring to the conjunction which represents the event-type  $E$



The types of determination that have been defined here (location, individual, and polarity) do not exhaust the possibilities. A determinative operator can be defined for anything which can be variable. While I do not see any benefit to relation indeterminates, other logical objects normally taken to be absolute could be made into variables. What type of quantifier a variable is bound to could itself be a variable, as could numerical functions and entire situations, event-types, or schemata, e.g. in a constraint. Even whether-or-not-a-thing-is-a-variable can be a variable. I have especially thought about punctuation indeterminates, which might be useful for describing the chalkboard writing of some professors.

### Types of Situation Similarity

The similarity event-type  $E_S$  can range from highly determined and specific to very abstract and general. I wish to discuss briefly the different types of similarity abstraction and the different levels of event-type involvement.

Any situation description of the form invented by Barwise and Perry can be represented as a conjunction of constituent sequences each assigned to a polarity and a location. I want to call these conjuncts *situation conjuncts*. They have the general form,

$$r^n(a_1, \dots, a_n, l, i).$$

In the strictest terms, more than one location might be specified (for instance in relations between locations), and  $n$  may be zero (when  $r$  is a locational relation or situational state). The abstracted situation conjuncts (which I call *event-type conjuncts*) of the similarity event-type  $E_S$  may thus have different types of indeterminates among the parameters of the predicate  $r^n$ , as long as all the  $i$ 's are anchored. For instance, a conjunct in  $E_S$  may be a property anchored to an individual and to a polarity but with an indeterminate in the location parameter. This type of abstraction I call "individual equivalence".

In general, two situation conjuncts can have the following types of similarity, in roughly descending order of strength: individual equivalence, locational equivalence, state equivalence, property equivalence, and  $n$ -ary relational equivalence.

*Individual Equivalence:* Two situation conjuncts have individual equivalence if the event-type conjunct they share contains no individual indeterminates. The event-type conjunct in question may or may not have locational equivalence, but in either case individual equivalence is a strong type of similarity: a conjunct in  $E_S$  with this type of similarity indicates that a relation is true of the same individual(s) in two different situations.

*Locational Equivalence:* Two situation conjuncts have this type of similarity if their shared event-type conjunct is anchored to a spacio-temporal location. At a given place and time, a relation may be true of more than one set of individuals. For instance, here and now, five is greater than three and four is greater than two. These two situation conjuncts are equivalent relations true at the same location but of different individuals.

The remaining types of similarity classify abstraction according to the arity of the relation predicate in the situation conjuncts. Any similar situation conjuncts will fulfill exactly one type from among these:

*State Equivalence:* Two situation conjuncts have state equivalence if they share an event-type conjunct with a 0-ary relation predicate. This occurs, for example, when it is raining in both situations, regardless of whether the rain is falling at the same location in each situation.

*Property Equivalence:* In this case, the shared event-type conjunct has a 1-ary relation predicate. The property may hold for different individuals at different locations. For example, a red block here has a property equivalence with a red sock there.

*N-ary Relational Equivalence:* Similarly, two situation conjuncts may share a predicate of arity  $n$ , greater than one. Classical analogies are cases of binary relational equivalence, for example beak:bird::mouth:human. A bird's beak and a human's mouth are both food intake features, so a situation in which a bird has a beak is similar to one in which a human has a mouth because the two situations share the binary predicate 'is the food intake feature of'. It should be noticed that the situation conjuncts retain their relational (state, property, or  $n$ -ary) equivalence even if they have locational and/or individual equivalence as well.

These, then, are the types of equivalence which two situation conjuncts can have. The event-type conjuncts in  $E_S$ , which are shared by situation conjuncts in  $e_C$  and  $e_A$ , are each anchored to a specific relation and polarity. Some of them may also be anchored to individuals and locations. We may define three grades of similarity for each type of equivalence: "total", "partial", and "no".

*Total (type) Equivalence:* The similarity  $E_S$  has the total grade of a type of equivalence (locational, individual, etc.) if every event-type conjunct in  $E_S$  has that type of equivalence.

*Partial (type) Equivalence:* If only some of the conjuncts in  $E_S$  have a particular type of equivalence, then the situations have that type only partially.

*No (type) Equivalence:* Two situations (or the shared event-type) have a type of equivalence not at all if none of their shared event-type conjuncts have that type.

Some pairs of situations, then, will be similar only by relational equivalence and will have no similarities for individuals. Other analogies may relate situations at the same locations, and so forth. There are many possible combinations that can arise in a similarity event-type  $E_S$ . For instance, two situation conjuncts might share one individual out of two in a binary relation -- they would not have individual equivalence. Two conjuncts with different predicates or polarities also cannot have individual or locational equivalence because they do not share an event-type conjunct with an anchored relation and polarity.

One might ask, "For what is all this useful?" I doubt that knowing whether two situations have total locational equivalence, for example, can assist in doing inference or in telling how similar two situations are relative to two others. But knowing the many forms that the similarity event-type can take does seem, nonetheless, to be helpful in creating the general case algorithms for reasoning by analogy. It also reminds us that analogy can mean any type of similarity in the theory as constructed. A space probe on Mars and a candy bar are analogous in their not being Bob Hope. This may seem a ludicrous case to allow, but we want to take advantage of any explicit knowledge one way or the other, so if a data base of constraints specifically implies that the probe and the candy bar are both not Bob Hope, then this should be honored as a similarity. The probe and candy bar have a property equivalence, however weak. Analogy extends also to situations like my playing the eighteenth hole at Fort Collins today with the same golf ball and clubs that I used yesterday. Descriptions of these two situations might have total individual equivalence, although the time difference between yesterday and today means that they have no locational equivalence.

The similarity between two situations may also be classified according to the levels of involvement at which they have some abstracted type of equivalence. This refers to how many implication constraints need to be invoked on the descriptions of two situations before the shared relation in question is found. It may be known, for instance, that individual  $a$  in situation  $e_C$  is a firetruck and that individual  $b$  in  $e_A$  is also a firetruck. And further, it may be that there is no event-type which is part of either  $e_C$  or  $e_A$  (and which does not include a conjunct with the predicate 'firetruck') such that the event-type implies that  $a$  or  $b$  is a firetruck. If there is in fact no such event-type then the property of being a firetruck is information as specific as any we know about the individuals  $a$  and  $b$ . This is a case of *0-level equivalence*. Suppose instead that  $a$  and  $b$  are equivalent only in that they are both vehicles, and that while  $a$  is a firetruck as before,  $b$  is a family car. In this case, firing constraints on the two situations,

in  $e_C$ : at  $l$ : firetruck, $a$ ;1

and

in  $e_A$ : at  $l'$ : car, $b$ ;1,

reveals that  $a$  and  $b$  are both vehicles, and that the resulting event-type conjunct in  $E_S$ :

in  $E_S$ : at  $l$ : vehicle, $a$ ;1,

represents *n-level equivalence*, where  $n$ =the maxi-min number of constraints required to be invoked to tell that  $a$  and  $b$  are both vehicles. If (a) Firetruck  $\Rightarrow$  Truck  $\Rightarrow$  Vehicle, and (b) Car  $\Rightarrow$  Vehicle, then the maximum of the two minimum constraint firings is two (invoked in the case of a), so the situations have 2-level equivalence. If there were, in the description of  $e_C$ , no more specific information about  $a$  than that it is a vehicle, then if  $b$  is a car and Car  $\Rightarrow$  Vehicle, the event-type conjunct in  $E_S$  would have 1-level equivalence because one constraint is fired in  $e_A$ , zero are fired in  $e_C$ , and one is greater than zero.

Given a data base of constraints which have several levels of entailment possible in them, this might be useful as a way of describing why, say, two Mars bars are equivalent at a more basic level than are a candy-bar and a space probe. In the latter pair, constraint firings might eventually tell us that the bar and the probe are both "things", but this level of similarity is clearly less impressive than a 0-level equivalence. Otherwise, I really have worked out no theories about levels of equivalence and I doubt that they can be useful in formulating rules about inference.

If two objects are similar, we might ask whether this implies that they are  $n$ -level equivalent for some finite  $n$ . The assumption in my theory is that the answer is yes, but we might imagine cases in which  $\infty$ -level equivalence would destroy this assumption. That is, it may be that two things can seem to be similar even though they share no explicable properties which account for our intuitive grasp of their similarity. This possibility has been called "dyadic similarity"<sup>117</sup> and its importance for analogy will be discussed later in this essay.

---

<sup>117</sup>Myers, "Inexplicable Analogies", *Philosophy and Phenomenological Research*, 22, p. 327

### Classifying the Forms of Analogical Justification

In part II of this essay, the form of analogy was given as follows:

$$\frac{P(A) \wedge Q(A)}{P(B)} \\ Q(B)$$

where  $P(A)$ ,  $Q(A)$ , and  $P(B)$  are the premises and  $Q(B)$  is the conclusion. It was shown that this reduces to an induction followed by a deduction:

$$\frac{P(A)}{Q(A)} \\ P(x) \Rightarrow Q(x) \\ \frac{P(B)}{Q(B)}$$

This two-step formulation was discovered by Aristotle, as I mentioned in part II of this essay, but there is another way of arriving at  $Q(B)$ . Specifically, analogy may also consist of an induction followed by an *abduction*, as follows:

$$\frac{P(A)}{Q(A)} \\ Q(x) \Rightarrow P(x) \\ \frac{P(B)}{Q(B)}$$

The problem of analogy is traditionally that of justifying the first step -- what Aristotle called the "incomplete induction". Theorem 16 established that a sufficient condition for concluding  $P(x) \Rightarrow Q(x)$  from  $P(A) \wedge Q(A)$  is that  $P(x)$  decides whether  $Q(x)$ , which I write " $P(x) > Q(x)$ ". If the second step is a syllogism, then the inference from analogy is logically necessary given the premises and that  $P(x) > Q(x)$ . But this is not the only type of analogical reasoning that is possible. It may be, for instance, that the conclusion of the first step is merely plausible instead of being valid, and it may be that the second step is abductive instead of deductive.

I contend that there are four types of analogical inference that have any kind of logical justification. The four kinds are presented initially in the table below, in which "the generalization" refers to the conclusion of the first step in the analogy and the "rule application" refers to the mode of inference whereby the second step's conclusion is reached:

Form of Justification	Status of the Generalization	Rule Application
Pure Deductive	Valid	Deduction
Pure Abductive	Valid	Abduction
Strong Inductive	Plausible	Deduction
Weak Inductive	Plausible	Abduction

In the table above, a valid *generalization* means that the rule (e.g.  $P(x) \Rightarrow Q(x)$ ) is inferred from both  $P(A) \wedge Q(A)$  and a determinative *test condition* (e.g.  $P(x) \supset Q?(x)$ ). This is the first inference step in the analogy. The second step is the *application* of a rule of inference. If the generalization is  $P(x) \Rightarrow Q(x)$ , and  $P(B)$  is observed, then  $Q(B)$  is concluded by application of the syllogism, a deductive rule of inference. If the generalization is  $Q(x) \Rightarrow P(x)$ , then  $Q(B)$  is concluded by abduction. The question that remains, then, is what it means for the generalization to be merely *plausible*.

When the generalization from  $P(A) \wedge Q(A)$  is valid, it follows soundly, but we may sometimes wish to conclude the generalization even when it is not strictly valid, in which case the first step in the analogy is inductive instead of deductive. We can get into trouble, as was demonstrated in part II, if we allow the generalization to be concluded automatically from just one instance (A). For one thing, merely knowing  $P(A) \wedge Q(A)$  does not indicate that we should prefer  $P(x) \Rightarrow Q(x)$  over  $Q(x) \Rightarrow P(x)$ . In addition, Q may not be a projectible property in Goodman's terms, or there may be abundant counter-examples to either generalization. Finally, there may simply be no reason to think that the occurrence of P has anything to say about whether Q occurs with it, as in the example of the car with a wadded up piece of paper in the back seat. We would not want to conclude that all the other cars in Ford's lot have wadded up pieces of paper in their back seats.

A test condition which is less demanding than determination but which should make the induction step have some plausibility is the condition that  $P(x)$  be *relevant* to whether  $Q?(x)$ , which I write as " $P(x) \# Q?(x)$ ". In the example above, then, the generalization that all the Ford cars in the lot have wadded up pieces of paper in them would not be concluded because the fact that a car is a Ford in a given lot is irrelevant to whether or not it has a wadded up piece of paper in it. In order to test for relevance, one might just include facts like " $P(x)$  is relevant to  $Q?(x)$ " in the data base of constraints, but I think that relevance can be defined in terms of determination. The type of relevance I have been discussing so far is relevance in deciding the truth value of  $Q?(x)$ , so let's define that first.

**Definition 24:**  $P?(x)$  is *relevant to deciding*  $Q?(x)$  iff there is a minimally sufficient<sup>118</sup> conjunction  $\bigwedge_{i=1}^m D_i?(x)$  such that (i) the

<sup>118</sup>By minimally sufficient, I mean that it should not be possible to remove any conjunct or conjuncts and have the remaining conjunction still be a determinant

conjunction decides  $Q?(x)$ , (ii) there is no  $i$  in  $\{1, \dots, m\}$  for which  $Q=D_i$ , and (iii)  $P=D_i$  for some  $i$ .

This definition says that  $P?(x) \dagger Q?(x)$  if and only if  $P?(x)$  is part of a determinant of  $Q?(x)$ . It can be generalized in situation theory to cover functional as well as truth value relevance as follows:

**Definition 25:** An event-type  $E$  is *relevant to the determination* of a conclusion event-type  $E'$  iff there is an event-type  $E^*$  such that (i)  $E^*$  determines  $E'$  and is minimal (viz, there is no event-type  $E^\circ$  such that [a]  $E^\circ$  determines  $E'$ , [b]  $E^\circ \neq E^*$ , and [c]  $E^*$  is of type  $E^\circ$ ), (ii)  $E'$  is not part of  $E^{*119}$ , and (iii)  $E$  is part of  $E^*$ .<sup>120</sup>

The four varieties of analogical justification may now be expounded more completely:

#### *Pure Deductive Analogy*

This is the only type of sound inference by analogy. Given an analogy  $\{e_C, e_A, E_S, E_D\}$ , an unknown  $E_U$  and a conclusion  $E'$  of which  $E_U$  is type, the conclusion follows by pure deduction if there is an  $E$  (with  $E_S$  of type  $E$ ) such that  $E$  determines  $E'$  (the test condition). The test condition is fulfilled if  $E$  involves  $E'$ , in which case the analogous case is *redundant*, although it might be useful in identifying the determinant and/or the conclusion. If, on the other hand, the anchoring of unshared indeterminates in  $E'$  by the consequent event-type  $E'''$  cannot be completely discovered without examining  $e_A$ , then the analogous case is *informative*. It may also be true that a conclusion about  $e_A$  can be reached by examining  $e_C$  as an analogous case, so that the situations are analogically *interactive*. And finally, the mere fact that an analogy exists between  $e_C$  and  $e_A$  may give us some more information about the central case, so that the two situations are actually synergistic. For instance, I dropped my backpack on an Oxford don's feet while leaving the Old Library one day. He reacted kindly, but if I had done it again the next day his reaction might have been different just because I had inflicted this unpleasantness on him a second time. Inertia effects may also be examples of synergy, e.g. Congress will pass more readily what it has passed before.

#### *Pure Abductive Analogy*

If the conclusion  $E'$  determines  $E$  (with  $E_S$  of type  $E$ ), then the inference that the conclusion event-type is anchored in  $e_C$  by analogy with  $e_A$  is not sound but is instead hypothetical. Such analogies are abductive because they provide us with an explanation, though not necessarily the only

<sup>119</sup>This requirement was pointed out to me by Stuart Russell

<sup>120</sup>It may be challenged that being part of a determinant is not a necessary, while it may be a sufficient, condition for relevance in determining. Such a challenge might run as follows: A person's sex is relevant to determining his or her name, yet it is difficult to think of a minimally sufficient determinant for name which would include sex. My present answer to this is that the relevance which I have defined may not match exactly with an ordinary language definition of 'relevance' but may be closer to the condition for plausible generalization. I cannot see how knowing that a person is male helps one to decide *plausibly* whether or not his name is 'Ray'. It only tells us that it is *possible*.

possible one, for at least part of the central case. It is difficult to say what constitutes a cogent abductive conclusion by analogy. One might regard the conclusion as a hypothesis to be tested. If one provisionally concludes  $E'''$ , then other constraints may allow one to conclude additional facts about  $e_C$ . These may be tested against conclusions about  $e_C$  which are drawn soundly without assuming  $E'$ , such as conclusions implied by the data base and sound conclusions by analogy with other situations.

#### *Strong Inductive Analogy*

If  $E$  is relevant to  $E'$  but is not sufficient to determine  $E'$  then the conclusion by analogy is merely plausible, or inductive. For instance, in an earlier example, the make and design of a car were said to determine the car's engine-type. If two cars are only known to have the same *design*, then it is not certain that they have the same type of engine, but there is at least some grounding for this conclusion. By an induction it is assumed that all cars with the design of the car in  $e_A$  have its type of engine, and then the rule is applied to the car in  $e_C$ . As one would expect, this is a very unstable form of inference. Its cogency clearly depends on how close  $E$  is to being a determinant of  $E'$ , i.e. *how relevant*  $E$  is. I have not come up with a good way to measure this. The conclusion may also be tested using certain *adjunct conditions* (such as whether the generalization has any known counter-examples) which are discussed in the next section.

#### *Weak Inductive Analogy*

The final form of analogical justification occurs when the conclusion  $E'$  is relevant to  $E$ . In this case, the inductive generalization that the  $E'''$  anchoring of  $E'$  involves the abstracted similarity portion  $E''$  is not a sound inference but has some grounding, and the projection of  $E'$  onto  $e_C$  is an abduction. The net result is a conclusion that is *possibly* an explanation for the abstracted portion  $E$  of  $E_C$ . This form of inference is weaker than strong inductive analogy because it not only involves an unstable generalization but also generates a mere hypothesis even given that the generalization holds. Thus, it must pass all the adjunct tests for both inductive and abductive analogy, to be discussed shortly. Obviously, a particular weak inductive analogy may be more cogent than a particular strong inductive analogy, the difference in terminology just reflects differences in the number of steps which are not sound, which is one in the strong case and is two in the weak case.

The four forms of legitimate analogy may be summarized using simple predicate symbols as follows:

<b>Pure Deductive Analogy</b>  $P(x) \supset Q(x)$ $P(A) \& Q(A)$ $\frac{P(B)}{Q(B)}$	<b>Pure Abductive Analogy</b>  $P(x) \prec Q(x)$ $P(A) \& Q(A)$ $\frac{P(B)}{Q(B)}$
<b>Strong Inductive Analogy</b>  $P(x) \uparrow Q(x)$ $P(A) \& Q(A)$ $\frac{P(B)}{Q(B)}$	<b>Weak Inductive Analogy</b>  $P(x) \uparrow Q(x)$ $P(A) \& Q(A)$ $\frac{P(B)}{Q(B)}$

### Inference Using Situations and Adjunct Conditions

The theory of analogy as I have outlined it suggests a means by which a computer could infer constraints based on examination of one or more situation descriptions. For purely deductive analogies, the new rule is inferred with certainty, but if the form of justification is one of the other three varieties, then there is a lot of ways in which the conclusion might be contradicted by other knowledge and beliefs. I wish to discuss some computational methods for both sound and non-sound reasoning by analogy.

Consider first the case of pure deductive analogy. I imagine that the knowledge base of a situation-based system might contain a large set of separate situation descriptions (coes), all entirely anchored. I further suppose that the knowledge base would contain a set of constraints relating event-types in varying degrees of abstraction, and that some constraints would be "involves" constraints and some would be "determines" constraints, with perhaps some "is relevant to" constraints and any other desired kind thrown in. Let us say that the system is working on a situation in which a 22-year-old American student is entering East Berlin at Checkpoint Charlie. A similar situation at a different time (yesterday) is retrieved from the knowledge base, and in that situation, the student had his passport stamped. Should the computer conclude that the student will have his passport stamped in the present situation? The test condition to be applied is whether or not there is a subset of the similarities between the two situations which determines whether or not the passport will be stamped. By firing constraints on the similarity, it may be learned that Checkpoint Charlie is a border entrance for East Germany, that East Germany is a country, and that passport stamping is a border control procedure. Another constraint may tell the computer that which country a border entrance is in determines its control procedures, and hence the conclusion that the passport will be stamped again is valid by application of a new "involves" constraint formed from the determinative one. The key maneuver in finding out whether the similarity contains a determinant for the conclusion is to fire constraints on the event-types in the proposed test condition.

In this example, the similarity contains much extraneous information -- that the person is a student, American, 22, etc. This information is a hindrance to deciding which test condition should be tried because there are so many possible ones. In this observation I think we have a clue to what



makes a good analogous case, namely, that its similarity with the central case should be close to the determinant in the test condition. A second thing to notice about this example is that the determination constraint is a default rule. Clearly, it would be violated if the student forgot his passport. The rule is probably useful nonetheless because it applies most of the time. An extension of this theory would be to allow "multi-valued" truth assignments (or probabilities) for constraints, instead of just 0 or 1. Finally, it should be pointed out that a by-product of the inference by analogy is the addition of a new "involves" constraint to the knowledge base, namely that East German border guards always stamp passports.

If the form of the justification is purely abductive then the same considerations as above apply but for discovering that the conclusion implies some portion of the observed similarity instead of the other way around. However, since the conclusion is no longer guaranteed to hold even by default, the system will need to do some checking of *adjunct conditions*, or tests for the consistency and reasonableness of the conclusion. For abductive analogy, these criteria include the following:

1. The conclusion should predict no results which are contradicted by those facts about the central case which are not part of the explained event-type.
2. There should either be no contradicting alternative conclusion which implies the same subset of the facts about  $e_C$ , or else any such alternatives should be able to be judged less likely than the conclusion, by some means.
3. There should be no counter-analogy with some other situation which predicts a conclusion at odds with the present conclusion.
4. There should either be no counter-example situations in which  $E''$  occurs without  $E'''$ , or else any such examples should be judged disanalogous in some significant way.
5. If heuristic risk values can be assigned to rejecting and accepting the conclusion, then the risk involved in accepting should be less than the risk involved in rejecting. And,
6. In general, the *a priori* probability of the conclusion should compare favorably with those of other hypotheses which are not inconsistent with the known situation and are therefore competing hypotheses.

This is not a complete, rigorous list, but merely an indication of the types of tests one should do. The second disjuncts in criteria 2 and 4 are obviously very fuzzy.

For inductive analogy, the adjunct conditions are safeguards against concluding an "involves" constraint which is contradicted in the knowledge base or which would otherwise be ill-founded:

1. The conclusion  $E'$  should be projectible, unlike "grue".
2. The inferred constraint that  $E''$  involves  $E'''$  (or that  $E'''$  involves  $E''$ , in the case of weak inductive analogy) should have no counter-examples in the knowledge base of situations.
3. There should be no counter-analogy which would result in a sound or very plausible conclusion contradicting the original conclusion  $E'$ . This would be any situation sharing some set of conjuncts other than  $E$  (though possibly with overlap) which would determine or be relevant to part of  $E'$  and result in a different polarity for at least one conjunct in  $E'$ . And,

4. There should be no part of the determinant  $E^*$  of type  $E$  (or  $E'$ , in the case of weak inductive analogy) which is unknown or contained in the disanalogy and is "crucial" (whatever that means) to determining  $E'$  (or  $E$ ). (Of course even a fully determinative disanalogy does not disprove the conclusion since both an antecedent and its complement can imply a consequent, but finding such a disanalogy mandates some counter-checking to see if a disproof of the conclusion is possible.)

This is probably not a complete list either, although I think it captures most of what can go wrong. The adjunct conditions must be fulfilled in inductive analogies in addition to the test condition, which is either that  $E$  be relevant to  $E'$  (strong) or that  $E'$  be relevant to  $E$  (weak). Strong inductive analogies need to satisfy the inductive adjunct criteria (for plausible generalization), while weak inductive analogies must fulfill both the inductive and the abductive adjunct conditions.

#### IV. Tests of the Theory and the Analysis of Analogies

Theories are usually evaluated according to their explanatory power, so it seems proper to try applying this theory of analogy to a mix of problems. In particular, the theory should be able to answer some, if not all, of the questions about analogy which motivated its study. I think, as it happens, that the theory can also explain some of what we already knew about analogy, but in a more precise way, and it may now be clearer just where other accounts of analogy either work or fall short in this regard. Finally, I will try to put the theory and my perspective on analogy to work in analyzing some example analogies.

##### Answering the Requirements for an Analogy Theory

I would like firstly to consider the eight categories of questions for theories of analogy which were proposed in part II of this essay. I do not wish to go through the questions one by one, since some of them have obvious answers already set forth in the theory and others are presently unanswerable from it. Rather, I would like to point out implications of the situational theory of analogy which might not be obvious and which appear to answer some of the questions.

One question about the form of analogy concerned how and whether one should distinguish between figurative and literal analogies. An example of a figurative analogy is one of the Reagan/Bush Committee's political commercials from Ronald Reagan's 1984 re-election campaign. The commercial depicted a situation in which a human is trapped in a woods with a hungry bear on the prowl. By analogy, the argument appears to be that, just as the woods dweller should be armed with a gun to defend against the bear, the United States should have a strong defense against the Soviet Union. In a figurative analogy, property equivalences between analogues do not seem to form the determinant; thus it is not that a woods dweller is much like the U.S. or that the bear is much like the U.S.S.R., or even that the woods are like the world. Instead, the relations *between* these individuals seem to be the salient similarity, namely the relation of one individual posing a threat to another. Literal analogies, in general, seem to be those in which there is a great deal of 0-level property equivalence in the determinant event-type. As properties get less specifically similar and as n-level relational equivalences predominate in the determinant, the analogy

gets more figurative, but the theory seems to indicate that there is no way to draw a line between the two since any combination of types of similarity is possible.

I think the theory can also shed some light on questions about analogy's place among modes of inference. The second step in analogy as I construe it is a straightforward application of either deduction or abduction. The first step -- concluding the rule -- is either deductive or what I have termed, by default, inductive. This latter covers cases in which the generalization is merely plausible, not valid, but the requirement for plausibility (namely, relevance) is more stringent than inductive criteria which guard against contradictions in the knowledge base and against non-projectibility. The problem of justifying induction as the number  $n$  of instances approaches infinity is different from the problem of justifying analogies. The difference is that questions about relevance and determinability go away as  $n$  gets very large, whereas whether or not the conclusion is projectible continues to be a problem even when, say, a thousand green emeralds have been observed before midnight and thus been called "grue", for one would not want to say that emeralds observed tomorrow will be blue. In the case of the wadded up piece of paper in the back seat of a car in Ford's lot, the conclusion that all these cars have this property is a projectible conclusion because as one observes more and more cars with pieces of paper in them one eventually believes that all cars in the lot have the property. One would expect that for inferences based on some number of instances neither large nor small, a mix of principles of good analogy (determinability or relevance) and for believable induction (number of confirming instances, projectibility, consistency) would be applied.

The theory demonstrates, in response to questions about the justification and soundness of reasoning by analogy, that inferences from parallel cases can, under certain circumstances, be both valid and necessary for concluding new constraints. Whether a substantial number of explicit analogies in real life are of this type is unclear, but it does seem to be a mechanism for *everyday* conclusions by analogy which we make without labelling them as such. What is really lacking in the theory is a way to evaluate the cogency of non-sound arguments by analogy. It does, however, indicate how an analogy might go wrong. If no part of the similarity is relevant to determining the conclusion and *vice versa*, then the argument has no legitimacy, as I see it. Furthermore, inferences by analogy that are not sound need to pass the adjunct conditions, precluding conclusions which are inconsistent with sound predictions or other examples, or which are of ludicrous probability versus the alternatives, are non-projectible, or are realistically undecidable given the unknown or disanalogous status of crucially relevant factors.

Some hints are given in the theory regarding the purposes and methods of analogy. Even redundant analogues can be useful in identifying possible conclusions which can then be verified or in pruning the set of facts from which one attempts to prove the conclusion, and when a relevant or determining similarity can be demonstrated to exist between two cases, a conclusion may be reached which would not otherwise be possible to infer. I think the theory also argues strongly that similarity should not be the criterion which is optimized in seeking good analogues. Salient similarities seem to be just those which either determine or are determined by a conclusion from the unknown.

### Explaining Some Introspective Data

Questions about the cognitive nature of analogical inference, which formed the last category of requirements set forth in part II, are very difficult to answer. I am not sure to what extent human analogizing can be mapped onto the logical structure which I have proposed, but I would like to cite three applications of parts of the theory which have occurred to me since I began working on this subject. The three "case studies" involve observations about how analogy functions in arguments, in thinking, and in learning.

#### *Counter-examples and Counter-analogies*

It struck me, as I thought about analogies I had collected from arguments and debates, that the function of analogy in arguments is often two-tiered. In arguing about what is the right conclusion for the central case at issue, a way of opposing someone else's conclusion is to argue that the central case has an analogous case, and that the analogous case (a) disproves the opponent's conclusion and (b) establishes either a hypothetical or a plausible or even a valid counter-conclusion. It accomplishes (a) in one of two ways: either by counter-example or by counter-analogy. In the former method the analogous case is an attempt to capture the antecedent conditions in the generalization which one's opponent is assumed to be making, either because the opponent has explicitly stated the generalization or because it is marginally inferable from his or her position on the central issue. If the antecedent (or determinant or relevant event-type) describes the analogous case and this case contradicts what the other is concluding, then the effect of the analogy is to disprove the generalization by counter-example. A counter-analogy disproves the generalization in a slightly different way, by finding a determinant or relevant similarity (for the conclusion) which is not the antecedent of the generalization but which is nonetheless shared by the central and analogous cases such that the analogous case contradicts the conclusion.

Clearly, if two determinants lead to different conclusions or if the same determinant has different consequences in two analogies then one's data base is inconsistent, but if the generalizations being made are merely plausible then it is easy to see how counter-examples and counter-analogies could be found in a consistent system of beliefs. This, I propose, is a lot of what is going on in arguments. Thomson's violinist analogy, for instance, is a counter-example to the generalization that any person with a right to life cannot be disconnected from a solely-capable life-sustainer at the sustainer's request. I am now satisfied that I had just been taking the generalization for granted before I read Thomson's article and that she provides good reason to reject it. Usually, one's opponent does not state a precise generalization, so one picks a generalization whose scope appears to be wide enough to be covered by the analogy and hopes that this will force a retreat. The situational theory seems to be a help in analyzing these arguments, about which I will have more to say in the example analyses.

#### *Analogy and Memory in Thinking*

Because the topic of analogy has pre-occupied me for the past seven or eight months, I have often stopped myself in mid-thought realizing that I am thinking analogically. What my theory is not equipped to explain is the rapidity with which the human mind is able to call up an appropriate

analogous case given a new situation. A student thinks about which of his week's worth of ideas he might tell to his advisor. Immediately the student is reminded of an old saying he heard back in his photography days: "A good photographer is one who never shows his bad pictures." A good student, then, should never reveal his bad ideas. The thought is momentary; the analogy clicks, or it does not.

How the saying about photographers just popped into my head, I do not know. This is the problem of searching for and matching an analogue, and it may well be that Hume was right in holding that the recognition of resembling ideas is just a case of association. My theory is aimed at explaining one's evaluation of the analogue after it is retrieved from memory, which seems to be a separate process, taking place after matching. If, say, I have an overdue book at the mathematical sciences library, I tend automatically to think that I can avoid a fine up to a week after it is due because, although I have not tested this at that particular library, I know that at Green and Meyer Libraries the overdue notices do not come out right away. Once this thought occurs to me, I evaluate the analogy and its conclusion. Is it reasonable to assume in this case that what holds for Meyer and Green will hold for the math library? I ponder the likelihood that the math folks send out notices by the same procedure as the other libraries do. My inquiry has turned to the test condition: Does the library's being a Stanford University Library determine its billing procedures? Cautiously, plausibly, I think yes. I returned the book and was not fined.

#### *Dreyfus' Theory of Skill Acquisition*

In a well attended lecture at the Stanford Center for the Study of Language and Information this year, Hubert Dreyfus argued that lists of if-then rules are not sufficient for constructing systems that are truly "expert".<sup>121</sup> Instead, Dreyfus says, experts are experts in virtue of their vast accumulated experience with specific examples, and skill acquisition moves away from applying abstract rules and toward recalling particular cases.

Dreyfus' example domain is chess-playing. A novice chess player, he says, knows a few general rules like the rules of the game and heuristics like "Get control of the center." At this stage, the player has to think about what moves are possible under the rules and calculates which move to make. At the advanced beginner and intermediate levels, according to Dreyfus, the player can draw on his own rules which have been formed from experience and, more importantly, recognize playing positions similar to those that have been encountered in the past and try to apply them. By the time a player is an expert, Dreyfus says that he or she just "sees" the right move or a very good one almost instantly, based on his or her "holistic memory" of common configurations and on some intuitive mechanism for grasping the relevant position and applying it. This theory, based partly on empirical studies of chess experts, is Dreyfus' alternative to the old idea that a true expert has evolved a set of rules which only need to be articulated and programmed by computer scientists in order for the expertise to be duplicated.

I suggest that one does not have to throw away the concept of rules in order to represent such an expert's knowledge, for the rules appear just to be getting more situation-specific. Whether the specific rules are actually represented and calculated upon as such or just acted in accordance with in

---

<sup>121</sup>Dreyfus, "From Socrates to Expert Systems: The Limits of Calculative Rationality" (CSLI Colloquium), March 14, 1985

some more basic mental process is a question I cannot answer. What seems to me interesting is how the situational theory can explain the process of forming this large and specific set of rules. Perhaps a novice knows or believes a set of constraints, some determinative and some implicative. To one who is inexperienced in chess, each new situation will probably have to be calculated upon by applying constraints. But as she or he gains a wider memory of chess situations, the advanced beginner and intermediate player can make moves by analogy with past positions, using determination and relevance rules as justification. Finally, at an expert level, the player can apply a constraint to an impending move immediately because so many have been formed from earlier analogies. Determination considerations are no longer needed at this point; the player zooms in parallel on the "involves" constraint with the appropriate antecedent event-type and applies the consequent to the situation at hand.

This might work for a chess program, but Dreyfus would probably argue that it is hopeless as a model of how human chess players become experts. People usually cannot recall very many specific positions from past games or articulate the antecedent conditions in the highly specific constraints I am construing. Still, I think this exercise at least demonstrates that one needn't posit some holistic mechanism to model expertise -- it may just be a matter of explaining how the mind could form new, more specific rules, and the fact that past situations or rules often cannot be articulated may just be because they are known at some subconscious level. Dreyfus cites as support for his theory the fact that chess grandmasters who can usually make a good move instantly in real games must nonetheless calculate like everyone else when solving chess "puzzles", presumably because these are problems which would rarely have occurred in the expert's past experience. This phenomenon would be expected under the situational theory because there would presumably be no constraints in one's knowledge base to cover such "puzzle" situations as there would be for more common positions.

### Comparisons and Contrasts with Other Theories

In part II of this essay I reviewed the types of account of analogy that I have identified in the literature. These included deductive, inductive, abductive, similarity-based, and associationist theories plus some hybrids. I see my theory as incorporating the first three, for the most part rejecting the fourth and leaving the psychological association theories as possible explanations for how analogies are found.

Since my theory allows for analogies which are either valid, hypothetical, plausible, or plausibly hypothetical, this helps to explain why cogent analogy has been looked upon variously as logically necessary, inductively plausible, or just abductively suggestive. Weitzenfeld<sup>122</sup> demonstrated validity for cases in which the conclusion is what I call "functionally determined" (solved) by the similarity, but I think that this does not account for all analogies. Firstly, truth value determination (decision) also seems to be necessary, especially for test conditions. Secondly, sometimes we *know* that the conclusion is not fully determined by the similarity rather than just being uncertain about it and yet we still believe the conclusion. And thirdly, sometimes the conclusion is merely hypothetical and determines or is relevant to the similarity as opposed to the reverse. My strongest dis-

---

<sup>122</sup>Op. Cit., Weitzenfeld, pp. 137-149

agreement with Weitzenfeld is that I think there is no reason to assume, as he appears to, that the determining structure for any given function is unique.<sup>123</sup> A football player's uniform or the coach he or she plays for can either of them determine the player's team membership; neither is unique in this respect.

Although my theory allows for analogies whose generalizations are merely plausible, there are some differences between my inductive forms of analogy and inductive forms proposed by others. I hold that the relevance test condition exists in addition to projectibility in cogent analogies, so they are somewhat more justified than arbitrary single-instance inductions. Also, other theories do not admit an abductive application, as an alternative to syllogism, of the plausible generalization. Perhaps the most important thing to notice about the condition for plausible rule formation in my theory is that it merely stipulates that *some* relevant event-type must exist in the similarity or conclusion. For inductive analogy, one has the option not to infer a specific constraint but to allow the anchored conclusion to be projected to the central case based on the belief that there is a relevant event-type *somewhere* in the similarity (or conclusion, in the weak case). I therefore do not feel that the generalization step needs to be explicit, only that it be known that one is possible.

I disagree with Hesse and Russell that explanatory hypotheses from analogies are intrinsically more believable than other hypotheses.<sup>124</sup> For it seems to me that there is no logical reason to prefer a conclusion based on analogy to another hypothesis which is equally explanatory if the analogized conclusion is not determined either totally or plausibly by the similarity. In making it clear how such analogies entail abduction, I think my theory can explain why the analogical argument for other minds, for instance, is so logically problematic however intuitively appealing it may be.

My theory does not use any similarity measure as a condition for analogical inference. A determinant or relevant similarity event-type may be very weak in the degree of similarity it represents, which is consistent with my position that strength of similarity has no direct bearing on the legitimate cogency of an analogy. Heuristically, the amount of similarity may be a good way to match initial analogues, but it does not work as a justification for conclusions. Theories based on psychological explanation may yet be necessary to account for how analogies occur to us and for the cogency of some analogies like the other minds argument, but I think the situational theory at least shows that there is a broad range of analogies whose logical justifications can be identified.

The relevance-based theory of Shaw and Ashley<sup>125</sup> is a relatively new type of theory, odd as that fact seems, in the game of justifying analogy. Shaw and Ashley do not give a definition of relevance in their paper or make it clear that relevance is not enough if the disanalogy is relevant also and can lead to a different conclusion by counter-analogy. But I think they had the right idea in emphasizing relevance, and now it seems that the question is, "How much relevance is enough?"

---

<sup>123</sup>Ibid.,pp. 139-141

<sup>124</sup>See discussion under "Theories Based on Abductive Hypothesis" in part II of this essay

<sup>125</sup>Op. Cit.,Shaw and Ashley,pp. 415-432

### Analyzing Analogies

Philosophers of argument have often condemned the use of analogies. William Sacksteder, who does not condemn analogy, asserts that many philosophers and rhetoricians "suppose analogies to be notoriously 'illogical'," belonging "to realms in which validity is not in question: to inventive inquiry, to refinement of exposition and teaching, or to literary grace."<sup>126</sup> Irving Copi says that "no argument by analogy is ever valid" though some are very cogent,<sup>127</sup> and M. C. Beardsley, as I noted earlier, maintains that analogy is not cogent at all.<sup>128</sup> I have proven that Copi was mistaken -- analogies can be valid, but there is also this feeling in many, including me, that analogies can be extremely misleading despite their power to convince people. I can explain the existence of analogies that are illegitimate; namely, they are ones in which the similarity event-type neither determines nor is relevant to the conclusion and *vice versa*. The question is, why would people be convinced by illegitimate analogical arguments or attach greater cogency to them than is theoretically justified? Put another way, under what circumstances can an argument by analogy be misleading?

Here I believe that strength of similarity and psychological explanation can find their rightful places in my theory. My experience indicates that, however fallacious I think it is, human beings are in fact inclined to act in accordance with the Similarity Principle, expecting further likeness when some likeness is observed. It is this tendency in human behavior which, if I would stop short of calling it the source of all evil in the world, nonetheless seems to be responsible for the bad analogizing that goes on. An analogy is misleading, in short, if the Similarity Principle predicts a conclusion which would fail the test condition or adjunct criteria for that type of analogy.

Humans are endowed with two gifts. One is the ability to perceive analogies, which is intuitive, psychological, associative at a deep level and hard to explain. The other is the ability to test analogies, which is analytical, logical, and argumentative and the methods for which I have discussed. The problem is that it is too easy, without an opposition debating us, to exercise our first ability without exercising the second: mental laziness allows one to draw analogical conclusions without testing them. The Rt. Hon. Lord Wilson of Rievaulx, former Prime Minister of the United Kingdom, said last year that one of the reasons for his resignation in 1976 was that he was "afraid of getting stale". What he meant, he said, was that after you have been in charge for a while, 'you see new problems which look like something you've encountered before, so you do what you did then, only the new situation may be different in some way that calls for a different approach.' Furthermore, he said that a fresh prime minister would handle such situations with more sensitivity to their nuances.<sup>129</sup> However important it was to Wilson's decision, this is an interesting comment because we often think that the more experiences one has to draw upon, the better will be the decisions that emerge. Maybe the key to avoiding staleness is to remain wary of the temptation to project unjustified conclusions by analogy.

---

<sup>126</sup>Sacksteder, "The Logic of Analogy", *Philosophy and Rhetoric*, 7, p. 234

<sup>127</sup>Op. Cit., Copi, p. 358

<sup>128</sup>Op. Cit., Beardsley, p. 113

<sup>129</sup>Wilson, "The Kennedy Memorial Lecture" (Oxford Union Society), November 30, 1984; The passage in single quotes is paraphrased.



Sometimes analogical thinking gives us the right answer even when we do not examine the test criteria. The alternative to concluding based on something we know may be to conclude nothing at all, that we just DON'T KNOW. This non-conclusion seems to be repugnant to humans when we have a question; it runs against our grain. So we draw an analogy, assume the conclusion it generates is right, and sometimes, perhaps often, we get lucky. The habit of leaping to analogized conclusions becomes a bit more dangerous, however, when we are judging on matters of great importance and thus are prey to arguments contrived by those who are biased. That people have the capacity to be misled by analogies, combined with the fact that arguments in favor of one or another position may cite seductive but illegitimate analogies, implies that philosophy should be of some benefit in separating good analogies from bad ones if the logical criteria are well defined. A comfortable seat in a debating chamber or in front of one's television set makes one vulnerable to arguments based on the Similarity Principle. Since I have sought to replace this common sense, but invalid, standard for drawing analogical inferences, I should try to apply my theory as well as the ideas about analogy that come out of it to some specific analogical arguments and to say what philosophy can tell us about these arguments.

In argument, most analogies play a dual role. Prior to their role in suggesting new conclusions, they are attempts to refute some other conclusion which the arguer is opposing. Such an analogy was invoked during the famous "King and Country" debate at the Oxford Union Society in 1933 by Quinton Hogg. Speaking against the motion "That this House will in no circumstances fight for its King and Country",<sup>130</sup> Hogg asked his opponents, "What would you do if you saw a man raping your wife? Would you interfere, or would you confine yourself to a polite invitation to desist?"<sup>131</sup> Hogg had generalized the motion to something like "Pacifism is always the best policy -- one should never defend with force," and he had given what he saw as the most intuitively clear counter-example possible in attempting to refute it. But the analogy really suggested its own conclusion, which was that Britain's young men should be willing to fight for Britain when it is threatened. Notice that the conclusion about the central case (Britain's defense) does not follow even if one agrees that the broader generalization chosen by Hogg is refuted by the rape example. For this, Hogg would have had to show that there is a national defense situation which shares those aspects of having one's wife raped which merit force. One can imagine that the bombing of Britain by Nazi Germany would have been such an example, though perhaps an invasion of Poland would not have.

In general, I think it is very important to distinguish (a) whether the generalization refuted by a given analogy is crucial to the scope of one's opposition from (b) whether the analogy establishes the conclusion being argued for in the central case. An analogy is always a counter-example to some generalization. For instance, Stephen Hellman of Hertford argued in a 1984 Oxford Union debate against having nuclear weapons in Britain by asserting that the empirical argument that nuclear weapons have kept the peace for forty years and are therefore likely to continue doing so is "like the young man who, having fallen off the edge of the cliff, says halfway

---

<sup>130</sup>The motion passed, 275 to 153, which sent a shock wave throughout Britain in February of that year -- Hollis, *The Oxford Union*, pp. 186-193

<sup>131</sup>Walter, *The Oxford Union: Playground of Power*, p. 82

down, 'Well, I'm all right so far.'<sup>132</sup> This analogy really only shows that positive experience when one initially had assumed danger does not necessarily prove that the experience will not wind up dangerous after all. That is the counter-example aspect of the argument. A stronger conclusion by analogy would be that Britain under the nuclear umbrella is deceived by its apparent success while disaster remains inevitable. As a counter-example, the argument serves quite nicely, but as an analogy to nuclear defense it appears to beg the question, for the whole problem is that we just don't know whether committing to nuclear weapons is like jumping off a cliff or not. Furthermore, the opposition could argue that the real generalization they advocate is not that results thus far are always a good predictor of results to come but that prolonged peace under a weapons system is a good indication that it prevents war. The scope thus reduced, a new analogy would be needed in order to refute the generalization deemed crucial by the opposition.

An example of a counter-*analogy* was provided by a speaker in the first Oxford Union debate I attended. Nicholas Prettejohn, arguing against "This House believes that sexual freedom leads to ruin", responded to claims that such freedom is responsible for violence and disease by saying that one seldom hears, in Britain, arguments that other freedoms, like political freedom and the freedom of movement, which are also demonstrably linked to violence and the spread of disease, but are not themselves the causes of these, should therefore be abridged.<sup>133</sup> That freedom of movement, say, shares with sexual freedom the properties of leading to violence and disease, and is not considered abridgeable, provides a case analogous to the central case in which the conclusion that freedom should be held responsible is refuted. This is a counter-example to the generalization that freedoms which allow violence and disease are therefore responsible for these, but if the generalization is just that *sexual* freedom which allows violence and disease is responsible for them then the argument is a counter-*analogy* -- it assumes that a freedom which does not specifically protect against the spread of violence and disease but also does not cause these is a determinant for whether or not the freedom is responsible, and it concludes that the cases of freedom of movement and political freedom anchor the polarity of "is responsible for" to "no" for the central case.

In the above example, the conclusion by analogy and the conclusion it attempts to refute are polar opposites. This should be the case whenever the issue for debate is a yes or no question. Thus Thomson's violinist analogy argues, for instance, not only that disconnecting another person is not always impermissible but also that abortion is *permissible* in the case of rape. The stronger conclusion succeeds if the antecedent of the generalization being refuted is a determinant -- in that case the analogous case is both a counter-example (counter-*analogy*) and a way to establish a definite result, that not only does the conclusion not necessarily hold for all such cases but that it specifically does not hold in the central case. In other counter-examples, the analogous case may establish that a generalization is false without establishing the specific conclusion being argued, as in the analogy from the King and Country debate cited earlier.

---

<sup>132</sup>Hellman proposing the motion "This house calls for a non-nuclear defence policy for Britain" (Oxford Union Society), November 27, 1984

<sup>133</sup>Prettejohn speaking sixth in opposition to "This House believes that sexual freedom leads to ruin" (Oxford Union Society), October 12, 1984

Analogies in political and ethical debates tend to have this refutational character. But in some types of analogical argument, the generalization being argued for is much more difficult to identify. These arguments rely on a perceived isomorphism between the analogues which is very complex. For instance, if one were to ask what it is about the similarity between rats and people that allows substances which cause cancer in rats to be assumed carcinogenic for humans, one would probably have a difficult time enumerating the determinant. Thomas Reid argued that because of the similarities between the earth and other planets like Mars and Venus (each gets its light from the sun, rotates on its axis, etc.), one might assume that the other planets could have "living creatures" like the earth does.<sup>134</sup> We know now that they probably do not (largely from having visited the moon and having probed Mars and Venus) and that the similarity between this world and the others does not seem to include a determinant for life as we know it.

For other analogies, the verdict is not in yet. One of the most interesting posited isomorphisms in recent years is the analogy between computers and brains, which may allow us to conclude that digital silicon machines can have intentionality. I do not plan to resolve this debate here, but only to formulate it somewhat within my theory of analogy. Human beings can understand discourse at a very deep level -- we know what the words mean. I would like to bypass distinctions like that between external and internal meaning and to assume that for a computer to have intentionality means that it would have essentially the same understanding capabilities that I have. The test condition for concluding that this is possible in computers is whether the similarity between brains and computers could ever contain a sufficient condition for human-type understanding. It seems to me that there are three possibilities for the determinant. At its most portable level, the requirement for thought would be a software requirement: a thing understands if it demonstrates enough flexibility in manipulating symbols, attaching them to objects, and so forth. At this level, a computer made of silicon could do human-type thinking if it were programmed correctly -- the conditions for thought at this level are medium-independent. At the opposite extreme, one might hold that nothing can think unless it has a soul; minds are not just brains but are something magical and perhaps can only be made by God or produced naturally. In between these two views is the position of philosophers such as John Searle, who argues that mental processes like understanding are fundamentally biological.<sup>135</sup> One of Searle's comments sparks me to draw the following analogy. Suppose we wanted to make a machine that could do digestion. This process goes on in humans as an elaborate exchange of molecules through the membranes of cells and involves chemical breakdown and synthesis. To do digestion in silicon or gallium arsenide, one would have to implement an analogous low-level chemical process allowing food material like chocolate and lettuce, but not shoe leather or book bindings, to be partially incorporated into the material of the machine and partially excreted. This is what digestion is. Maybe mental activity has this low-level character too, so that it would be impossible to implement the necessary conditions for thought in silicon. On the other hand, perhaps thinking is exactly coincident with the activity (digital computation) that goes on in present-

---

<sup>134</sup>Reid, Thomas, *Essays on the Intellectual Powers of Man* (Essay I, Chapter 4), cited in *Op. Cit.*, p. 352

<sup>135</sup>Searle, *Minds, Brains, and Science*, pp. 28-41

technology computers. Perhaps analog computation is required but beyond that there is no physical requirement.<sup>136</sup> Or maybe thinking is beyond material specification altogether. These are the main possibilities as I see them. What is at issue is the scope of the determinant (material-independent, material-dependent, or extra-material) which will make the conclusion by analogy valid if and only if human-type intentionality does not require something lying outside the similarity assumed to exist between computers and brains. I would class those whom Daniel Dennett calls "computationalists" as believing in the material-independent determinant, "holists" as believers in extra-material requirements, and "connectionists" as believing in material dependence.<sup>137</sup>

The most commonly debated analogical arguments in philosophy are abductive: the conclusion, if true, would help to explain observed effects and is therefore accepted as their most probable cause or explanation. The problem arises when more than one hypothesis is consistent with what is observed. In particular, two questions to which arguments by analogy have been applied are the problem of other minds and the question of God's existence. I have discussed the analogical argument for other minds at various points already. C. W. K. Mundle has criticized philosophers for not paying enough attention to this argument, and he concludes his *Critique of Linguistic Philosophy* with a formulation of the basic argument:

"I shall end with the challenge to identify the fallacy, if there is one, in arguing as follows:

In my own case, S1 [some specific kind of stimulus] is always or usually followed by M1 [some specific kind of mental act or process], M1 is always or usually followed by B1 [some specific kind of bodily behaviour] and/or B1 is always or usually preceded by M1;

John Doe reacted to S1 with behaviour B1, so in view of the many similarities between us, anatomical, cultural, etc., there is some degree of probability that in John Doe S1 and B1 were mediated by M1."<sup>138</sup>

I take it as given that this argument has abductive legitimacy, which only gives it the status of generating a hypothesis. Mundle seems to be going further, saying that the similarity between another person and me contains a determinant for whether or not M1 follows S1. Actually, he asserts only that the existence of the determinant is probable "in some degree", but I do not know what this means so I will just treat the conclusion as definite. The problem in trying to make the conclusion by analogy sound is just that the known similarity does *not* include a determinant for mental activity. We cannot know what it is like to experience someone else's thoughts without privileged access to those thoughts, and so without a determinant in the similarity, my answer to Mundle's challenge would be that the fallacy in the argument is that it relies on the Similarity Principle. Nothing follows from the fact that others are like me in observed respects -- that could all be part of the wicked conspirator's (or my mind's) plan. In fact, we don't even

---

<sup>136</sup>Smith, "Is Computation Formal?" (Stanford Philosophy Department Colloquium), May 24, 1985

<sup>137</sup>Dennett, "Computer Models and the Mind -- a view from the East Pole", *Times Literary Supplement*, Dec. 14, 1984

<sup>138</sup>Mundle, *A Critique of Linguistic Philosophy*, p. 277

know whether any of these observed anatomical similarities are even *relevant* to mental activity, so I am afraid that the argument by analogy for other minds is destined to remain non-sound and implausible. There must be some reason why we believe in other minds; I agree with A. M. Turing that the idea that everyone thinks is a "polite convention".<sup>139</sup> The convention is appealing and convenient, and these appear to be the motivation for accepting it.

Of equal philosophical importance is the argument by analogy for the existence of God, expressed as the "argument from design" by Hume in his *Dialogues*,<sup>140</sup> and earlier "in the general scholium to Sir Isaac Newton's *Principia Mathematica*".<sup>141</sup> Hume's theist debator Cleanthes expounded it in the *Dialogues*:

"Look round the world: Contemplate the whole and every part of it: You will find it to be nothing but one great machine, subdivided into an infinite number of lesser machines, which again admit of subdivisions to a degree beyond what human senses and faculties can trace and explain. All these various machines, and even their most minute parts, are adjusted to each other with an accuracy which ravishes into admiration all men who ever contemplated them. The curious adapting of means to ends, the productions of human contrivance; of human design, thought, wisdom, and intelligence. Since therefore the effects resemble each other, we are led to infer, by all the rules of analogy, that the causes also resemble, and that the Author of Nature is somewhat similar to the mind of man, though possessed of much larger faculties, proportioned to the grandeur of the work which he has executed. By this argument, *a posteriori*, and by this argument alone, do we prove at once the existence of a Deity and his similarity to human mind and intelligence."<sup>142</sup>

The odd thing about the *Dialogues*, for many, is that Cleanthes' opponent Philo, a skeptic like Hume, in the end concedes "That the cause or causes of order in the universe probably bear some remote analogy to human intelligence."<sup>143</sup> To me, the argument from design differs from the argument by analogy for other minds in that no one has come up with an alternative to Philo's conclusion which has its explanatory power. One can say that what causes universes just isn't known, but Cleanthes offers a possibility we can understand. One can say that the universe is more like a tree than a machine, but concluding that no one designed the machine by analogy with the tree is still begging the question because we don't know whether or not the tree has a designer. The claim that a universe could create itself cannot be disproved, but the idea lacks an intuitive analogous case as far as we know. So the argument from design is an abductive analogy whose hypothesis is not strictly valid but to which an equally comprehensible alter-

---

<sup>139</sup>Turing, "Computing Machinery and Intelligence", in Op. Cit., Hofstadter and Dennett, p. 60

<sup>140</sup>Hume, *Dialogues Concerning Natural Religion and the Posthumous Essays*

<sup>141</sup>Popkin, "Editor's Introduction", *Ibid.*, p. x

<sup>142</sup>Hume, *Ibid.*, p. 15

<sup>143</sup>Hume, quoted by Popkin, *Ibid.*, p. xiii

native has not been found. It seems to me that skepticism about God's existence is based not so much on the non-determination of the similarity in the argument as in observations which its hypothesis has a hard time explaining, e.g. Why would a Deity create evil, allow suffering, etc.? In the sciences, analogies are often abductive also and, like Cleanthes' argument, have only one currently defined explanatory hypothesis, but the hypothesis is usually accepted unless there are observations it does not explain. Perhaps this is an adjunct requirement which is needed for good abductive analogies to be convincing. When an abductive analogy without competing alternatives is unable to meet this additional requirement, new theories are sought, but consistent new theories in science appear much easier to come by than ones for explaining how the order of the universe came to be.

Hume was skeptical not only about proofs but also about disproofs.<sup>144</sup> So my interpretation would be that Philo's grudging acceptance of the weak, "ambiguous" hypothesis is a response to the lack of apparent alternatives, but that there is, in his earlier denial that the causes of the universe are knowable, the lurking suspicion that the alternative is out there somewhere, inconceivable to us because it has, in our world, no analogue.

---

<sup>144</sup>Popkin, *Ibid.*, p. xiv

## V. Conclusion

What I do not have is a theory that will yield the same analysis of a given analogy from every analyst. I am aware, for instance, that Prettejohn's freedom analogy could be analyzed as a counter-example instead of as a counter-analogy, and that the argument for other minds might be classed as weakly inductive instead of purely abductive. But I think that my theory of analogy does have some fairly clear consequences, which might be enumerated as follows:

1. Treating analogy as a single form of argument and making global statements about its value should be seen to be untenable. Analogies run the gamut from logically valid to completely illegitimate.
2. There is good reason to seek novel analogues which may have less *a priori* similarity with the central case than others have. Creative thinking involves the discovery of such novel analogies rather than just looking for the most similar analogous case and projecting a conclusion. Since determination and relevance are what are important, an analogous case can be very dissimilar to the central case and still provide useful conclusions.
3. There is good reason to develop the critical thinking abilities needed to recognize when our natural acceptance of the Similarity Principle may be misleading us. A recognition that this principle has no validity is a good first step. And,
4. Analogy has a unique, non-redundant function in logical problem-solving. When one knows a determination rule, one can examine a case analogous to that being worked on and conclude an implication rule at some level of legitimacy.

Relative to traditional implication rules, determination constraints lessen the informational detail needed in order to infer specific-case results. The best way to test how useful this is would probably be to implement the situational theory in a computer program. The computer could then be experimented upon, i.e. given an analogy and asked whether the analogy has any cogency. Such results could be compared to those produced by human beings who are asked the same questions.

Experiments might also help in discovering whether or not "dyadic similarity" is crucial to the understanding of analogies. This, it might be recalled, is a level of equivalence that is really beyond analysis. I have so far assumed that a similarity can always be summarized as an identity of certain relations, but William James argued against this, saying,

"...any theory which would base likeness on identity, and not rather identity on likeness, must fail. It is supposed perhaps, by most people, that two resembling things owe their resemblance to their absolute identity in respect of some attribute or attributes combined with the absolute non-identity of the rest of their being. This, which may be true of compound things, breaks down when we come to simple impressions."<sup>145</sup>

One may imagine examples in which two terms might seem similar to us without being explicably similar, e.g. James's example of "*blotting paper*

---

<sup>145</sup>James, William, *Principles of Psychology* (London: Macmillan and Co., 1891), pp. 532-533, cited in Op. Cit., Myers, p. 327

voices", two musical tones,<sup>146</sup> and word pairs like 'up' and 'good' which are similar in connotation.<sup>147</sup> The real question is whether this primitive level of similarity can be coded in a symbolic processing machine.

While inferences cannot be justified solely on the basis of similarity, discovering similarity is crucial to the generation of testable analogies. For this reason, the study of similarity and of the analogy of being is a necessary step in defining a complete theory of analogy which I have not taken. The limitations of univocal predication may require a more elaborate form of semantic inference than symbolic implication, and the use of analog computing and analogue representations would surely require a clearer notion of the role of association in similarity cognition.<sup>148</sup>

There are many additional aspects of analogical inference which I have not had the time to explore. These would include tidying up the adjunct criteria for inductive and abductive analogy. Implementation of the theory would require algorithms for managing the complexity problems in similarity searches and matching, constraint chaining, contradiction checking and the like. The cogency of an inductive analogy cannot be determined as the theory stands because there is no measure of how relevant each part of a determinant is to the conclusion. Finally, I have not investigated how people discover, or how machines should infer, determination constraints -- an ability which is clearly needed if a person or machine is to learn on his, her, or its own. But that is a problem for another day and an untired mind.

---

<sup>146</sup>Myers, *Ibid.*, p. 326

<sup>147</sup>This last example is from Professor David Wellbery, who mentioned it at my senior thesis colloquy

<sup>148</sup>See Evans, *The Varieties of Reference*, pp. 292-294, for a discussion of the philosophy of similarity in reference; and Funt, "Analogical Modes of Reasoning and Process Modeling", *IEEE Computer*, October 1983, pp. 99-104 on analogue representation, perception, and parallel processing



### Acknowledgements

A great many people helped me, on both sides of the Atlantic, as I was researching and putting together this thesis. I would especially like to thank Professor and Mrs. James Urmson for welcoming me into their home in Cumnor on numerous occasions during the autumn. I think Edwin Good put it best when he remarked that getting Professor Urmson to tutor me in Oxford amounted to "pure luck".

I would also like to thank Professors Good and David Wellbery for their comments at my colloquy; Marsh McCall, Jr. for tutoring me in classical rhetoric despite being swamped as the chairman of classics at Stanford; and Doug Edwards, Russ Greiner, Jeff Lang, Paul Filseth, Bryn Ekroot, Anne Gardner, Jerry Hobbs, David Israel, Stan Rosenschein, and Paul Rosenbloom for helpful discussions.

Stuart Russell and David Helman each read through earlier drafts, and I am very grateful for their comments and corrections as well as for the hours of their time which they gave to me in discussing this work.

Finally, I would like to thank my advisor, John Perry, who was never too busy to squeeze me in between the hosts of others wanting to see him and spent many hours with me during the year over which this project ranged. If there is one thing I have learned in five years at Stanford it is how to pick a tutor.

### Works Consulted

- Achinstein, Peter. "Models, Analogies and Theories", *Philosophy of Science*, 1964 (Volume 31), pp. 328-350
- Achinstein, Peter. "Variety and Analogy in Confirmation Theory", *Philosophy of Science*, 1963 (Volume 30), pp. 207-221
- Agassi, Joseph. "Discussion: Analogies as Generalizations", *Philosophy of Science*, 1964 (Volume 31), pp. 351-356
- Anderson, James F. *Reflections on the Analogy of Being*. The Hague: Martinus Nijhoff, 1967
- Anderson, R. O. "Conclusive Analogical Argument", *Revue Internationale de Philosophie*, 1969 (Volume 23), pp. 44-57
- Anttila, Raimo. *Analogy*. New York: Mouton Publishers, 1977
- Anttila, Raimo and Warren A. Brewer. *Analogy: A Basic Bibliography*. Amsterdam: John Benjamins B.V., 1977
- Aristotle; John Henry Freese, translator. *The "Art" of Rhetoric*. Cambridge, MA: Harvard University Press, 1925
- Aristotle; Hugh Tredennick, translator. "Prior Analytics", *The Organon I*. Cambridge, MA: Harvard University Press, 1938
- Austin, J. L.; J. O. Urmson and G. J. Warnock, editors. *Philosophical Papers*. Oxford: Oxford University Press, 1979
- Avramides, Anita. "Understanding the Gricean Account of Meaning". Wolfson Philosophy Society Meeting, Oxford, November 20, 1984
- Ayer, A. J. *Philosophy in the Twentieth Century*. New York: Vintage Books, 1982
- Barwise, Jon and John Perry. "Shifting Situations and Shaken Attitudes". Stanford Center for the Study of Language and Information Report No. CSLI-84-13, August 1984
- Barwise, Jon and John Perry. *Situations and Attitudes*. London: MIT Press, 1983
- Beardsley, M. C. *Thinking Straight*. Englewood Cliffs, NJ: Prentice-Hall Inc., 1975
- Black, Max. *Models and Metaphors*. Ithaca: Cornell University Press, 1962
- Bochenski, I. M. "On Analogy", *The Thomist*, 1948 (Volume 11), pp. 424-447
- Braithwaite, Richard Bevan. *Scientific Explanation*. Cambridge, UK: Cambridge University Press, 1959
- Burrell, David Bakewell. *Analogy and Philosophical Language*. New Haven: Yale University Press, 1973
- Burstein, Mark H. "A Model of Incremental Analogical Reasoning and Debugging", in *Proceedings of the National Conference on Artificial Intelligence (AAAI-83)*. Menlo Park, CA: American Association for Artificial Intelligence, 1983, pp. 45-48

- Carbonell, Jaime G. "A Computational Model of Analogical Problem Solving", in *Proceedings of the Seventh International Joint Conference on Artificial Intelligence (IJCAI-81)*. Menlo Park, CA: American Association for Artificial Intelligence, 1981, Volume 1, pp. 147-152
- Carbonell, Jaime G. "Derivational Analogy and its Role in Problem Solving", in *Proceedings of the National Conference on Artificial Intelligence (AAAI-83)*. Menlo Park, CA: American Association for Artificial Intelligence, 1983, pp. 64-69
- Carnap, Rudolf. "Discussion: Variety, Analogy, and Periodicity in Inductive Logic", *Philosophy of Science*, 1963 (Volume 30), pp. 222-227
- Carnap, Rudolf. *Logical Foundations of Probability*. Chicago: University of Chicago Press, 1963
- Carney, James Donald and Richard K. Scheer. *Fundamentals of Logic*. New York: Macmillan Publishing Co., Inc., 1974
- Chappell, U. C., editor. *Hume*. Garden City, NY: Anchor Books, 1966
- Churchland, Paul M. *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind*. Cambridge, MA: MIT Press, 1984
- Cohen, L. Jonathan. *The Implications of Induction*. London: Methuen & Co Ltd, 1970
- Cohen, Morris R. and Ernest Nagel. *An Introduction to Logic and Scientific Method*. New York: Harcourt, Brace and Company, 1934
- Constantini, Domenico. "Analogy by Similarity", *Erkenntnis*, 1983 (Volume 20), pp. 103-114
- Copi, Irving M. *Introduction to Logic*. New York: The Macmillan Company, 1972
- Cousins, Adrian. "Recognition and Similarity". Wolfson Philosophy Society Meeting, Oxford, December 11, 1984
- Curtis, Helena. *Biology*. New York: Worth Publishers Inc., 1983
- Davidson, Donald. "How Do I Know What I Think?". First Gareth Evans Memorial Lecture, Oxford, October 30, 1984
- Davidson, Donald. "What Metaphors Mean", *Inquiries into Truth and Interpretation*. New York: Oxford University Press, 1984, pp. 245-264
- Dennett, Daniel C. "Computer Models and the Mind -- a view from the East Pole", *The Times Literary Supplement*, December 14, 1984
- Dennett, Daniel C. and Douglas R. Hofstadter, arrangers and composers. *The Mind's I: Fantasies and Reflections on Self and Soul*. New York: Bantam Books, Inc., 1981

- Dorrough, Douglas C. "A Logical Calculus of Analogy Involving Functions of Order 2", *Notre Dame Journal of Formal Logic*, July 1970 (Volume 11, Number 3), pp. 321-336
- Dreyfus, Hubert. "From Socrates to Expert Systems: The Limits of Calculative Rationality". CSLI Colloquium, Stanford, March 14, 1985
- Esper, Erwin A. *Analogy and Association in Linguistics and Psychology*. Athens, GA: University of Georgia Press, 1973
- Evans, Gareth. *The Varieties of Reference*. New York: Oxford University Press, 1982
- Fearnside, W. Ward. *About Thinking*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1980
- Fearnside, W. Ward. and William B. Holther. *Fallacy: The Counterfeit of Argument*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1959
- Funt, Brian V. "Analogical Modes of Reasoning and Process Modeling", *IEEE Computer*, October 1983, pp. 99-104
- Genesereth, Michael And Nils Nilsson. *Fundamentals of Artificial Intelligence*. Unpublished Manuscript, Jan. 13, 1985
- Genesereth, Michael R. "Metaphors and Models", in *Proceedings of the National Conference on Artificial Intelligence (AAAI-80)*. Menlo Park, CA: American Association for Artificial Intelligence, 1980, pp. 208-211
- Gick, Mary L. and Keith J. Holyoak. "Schema Induction and Analogical Transfer", *Cognitive Psychology*, 1983 (Volume 15), pp. 1-38
- Goodman, Nelson. *Fact, Fiction and Forecast*. Cambridge, MA: Harvard University Press, 1983
- Hendel, Charles W. *Studies in the Philosophy of David Hume*. Indianapolis: The Bobbs-Merrill Company, Inc., 1963
- Hesse, Mary. "Analogy and Confirmation Theory", *Philosophy of Science*, 1964 (Volume 31), pp. 319-327
- Hesse, Mary B. *Models and Analogies in Science*. Notre Dame: University of Notre Dame Press, 1966
- Hodges, Wilfred. *Logic*. Harmondsworth, UK: Penguin Books, 1977
- Hofstadter, Douglas R. *Godel, Escher, Bach: An Eternal Golden Braid*. New York: Vintage Books, 1980
- Hofstadter, Douglas R. "Metamagical Themas: How might analogy, the core of human thinking, be understood by computers?", *Scientific American*, September 1981, pp. 18-30
- Hollis, Christopher. *The Oxford Union*. London: Evans Brothers Limited, 1965
- Hume, David; Richard H. Popkin, editor. *Dialogues Concerning Natural Religion and the Posthumous Essays*. Indianapolis: Hackett Publishing Company, 1980
- Hume, David. *A Treatise of Human Nature*. Oxford: The Clarendon Press, 1960

- Israel, David J. "The Role of Logic in Knowledge Representation", *IEEE Computer*, October 1983, pp. 37-41
- Joseph, H. W. B. *An Introduction to Logic*. Oxford: The Clarendon Press, 1957
- Kahane, Howard. *Logic and Contemporary Rhetoric*. Belmont, CA: Wadsworth, Inc., 1980
- Keynes, John Maynard. *A Treatise on Probability*. London: Macmillan & Co Ltd, 1957
- Kling, Robert E. "A Paradigm for Reasoning by Analogy", *Artificial Intelligence*, 1971 (Volume 2), pp. 147-178
- Kneale, William. *Probability and Induction*. Oxford: The Clarendon Press, 1963
- Kuipers, Theo A. F. "Two Types of Inductive Analogy by Similarity", *Erkenntnis*, 1984 (Volume 21), pp. 63-88
- Kyburg, Henry E., Jr. *Probability and the Logic of Rational Belief*. Middletown, CT: Wesleyan University Press, 1961
- Laird, John. *Hume's Philosophy of Human Nature*. Oxford: Archon Books, 1967
- Laing, B. M. *David Hume*. New York: Russell & Russell, 1968
- Leblanc, Hugues. "A Rationale for Analogical Inference", *Philosophical Studies*, 1969 (Volume 20) pp. 29-31
- Leech, Geoffrey. *Semantics: The Study of Meaning*. Harmondsworth, UK: Penguin Books, 1981
- Lewis, Charlton T. and Charles Short. *A Latin Dictionary: Founded on Andrews' Edition of Freund's Latin Dictionary*. Oxford: The Clarendon Press, 1975
- Liddell, Henry George and Robert Scott, et. al. *A Greek-English Lexicon*. Oxford: The Clarendon Press, 1966
- Lloyd, G. E. R. *Polarity and Analogy: Two Types of Argumentation in Early Greek Thought*. New York: Cambridge University Press, 1966
- Lyttkens, Hampus. *The Analogy Between God and the World: An Investigation of its Background and Interpretation of its Use by Thomas of Aquino*. Uppsala: Almqvist & Wiksells Boktryckeri AB, 1952
- MacDonald, Margaret. "The Philosopher's Use of Analogy", in Antony Flew, editor. *Logic and Language*. Garden City, NY: Anchor Books, 1965, pp. 85-106
- MacNabb, D. G. C. *David Hume: His Theory of Knowledge and Morality*. Oxford: Basil Blackwell, 1966
- Magee, Bryan. *Men of Ideas: Some Contemporary Creators of Philosophy*. Oxford: Oxford University Press, 1978
- McInerny, Ralph. *Studies in Analogy*. The Hague: Martinus Nijhoff, 1968

- McKerrow, Ray E., editor. *Explorations in Rhetoric: Studies in Memory of Douglas Ehninger*. Glenview, IL: Scott, Foresman and Company, 1982
- Measell, James S. "Classical Bases of the Concept of Analogy", *Journal of the American Forensic Association*, 1973 (Volume 10), pp. 1-10
- Mill, John Stuart. *A System of Logic*. New York: Harper & Brothers Publishers, 1900
- Miller, Jonathan, editor. *States of Mind: Conversations with Psychological Investigators*. London: British Broadcasting Corporation, 1983
- Montague, Richard; Richard H. Thomason, editor. *Formal Philosophy*. New Haven: Yale University Press, 1974
- Mundle, C. W. K. *A Critique of Linguistic Philosophy, with Second Thoughts -- an Epilogue after Ten Years*. London: Glover and Blair Limited, 1979
- Myers, Charles M. "Inexplicable Analogies", *Philosophy and Phenomenological Research*, March 1962 (Volume 22, Number 3), pp. 326-333
- Nagel, Ernest. *The Structure of Science*. New York: Harcourt, Brace & World, 1961
- Niiniluoto, Ilkka. "Analogy and Inductive Logic", *Erkenntnis*, 1981 (Volume 16), pp. 1-34
- Niiniluoto, Ilkka. "Analogy, Transitivity, and the Confirmation of Theories", in L. Jonathan Cohen and Mary Hesse, editors. *Applications of Inductive Logic: Proceedings of a Conference at The Queens College, Oxford, 21-24 August 1978*. Oxford: The Clarendon Press, 1980
- Norton, David Fate. *David Hume: Common Sense Moralist, Sceptical Metaphysician*. Princeton: Princeton University Press, 1982
- Ortony, Andrew, editor. *Metaphor and Thought*. New York: Cambridge University Press, 1979
- Papineau, David. "Probabilities and Causes", *Journal of Philosophy*, February 1985 (Volume 82, Number 2), pp. 57-74
- Perelman, Ch. and L. Olbrechts-Tyteca. *The New Rhetoric: A Treatise on Argumentation*. London: University of Notre Dame Press, 1969
- Perrin, Porter G. *Writer's Guide and Index to English*. Dallas: Scott, Foresman and Company, 1942
- Pietarinen, Juhani. "Lawlikeness, Analogy and Inductive Logic", *Acta Philosophica Fennica*, 1972 (Volume 26), whole issue
- Pollock, John L. "A Solution to the Problem of Induction", *NOUS*, 1984 (Volume 18), pp. 423-461
- Polya, George. *Induction and Analogy in Mathematics*, Volume 1 of *Mathematics and Plausible Reasoning*. Princeton: Princeton University Press, 1954

- Price, H. H. *Thinking and Experience*. Cambridge, MA: Harvard University Press, 1962
- Pylyshyn, Zenon W. *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, MA: MIT Press, 1984
- Quine, Willard van Orman. *The Roots of Reference*. La Salle, IL: Open Court Publishing Co., 1974
- Quine, W. V. and J. S. Ullian. *The Web of Belief*. New York: Random House, 1970
- Reitman, Walter R. *Cognition and Thought: An Information Processing Approach*. New York: John Wiley & Sons, Inc., 1965
- Rescher, Nicholas. *Introduction to Logic*. New York: St Martin's Press, 1964
- Rich, Elaine. *Artificial Intelligence*. San Francisco: McGraw-Hill Book Company, 1983
- Rissland, Edwina. Hypothetical Examples in Legal Reasoning. Computers and Law Seminar, Stanford Law School, March 13, 1985
- Ross, J. F. *Portraying Analogy*. New York: Cambridge University Press, 1981
- Ross, Raymond S. *Persuasion: Communication and Interpersonal Relations*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1974
- Russell, Bertrand. *Human Knowledge: Its Scope and Limits*. New York: Simon and Schuster, 1948
- Sacksteder, William. "Analogy: Justification for Logic", *Philosophy and Rhetoric*, 1979 (Volume 12, Number 1), pp. 21-40
- Sacksteder, William. "The Logic of Analogy", *Philosophy and Rhetoric*, 1974 (Volume 7, Number 4), pp. 234-252
- Scriven, Michael. *Reasoning*. New York: McGraw-Hill, Inc., 1976
- Searle, John. *Minds, Brains and Science*. Cambridge, MA: Harvard University Press, 1984
- Shaw, William H. and L. R. Ashley. "Analogy and Inference", *Dialogue: Canadian Journal of Philosophy*, 1983 (Volume 22), pp. 415-432
- Sloman, Aaron. *The Computer Revolution in Philosophy: Philosophy, Science, and Models of Mind*. Hassocks, UK: The Harvester Press Limited, 1978
- Smith, Brian. "Is Computation Formal?". Philosophy Department Colloquium, Stanford, May 24, 1985
- Smith, Norman Kemp. *The Philosophy of David Hume: A Critical Study of Its Origins and Central Doctrines*. London: Macmillan and Co., Limited, 1941
- Solomon, Robert C. *Introducing Philosophy: Problems and Perspectives*. San Diego: Harcourt, Brace, Jovanovich, Inc., 1981
- Spohn, Wolfgang. "Analogy and Inductive Logic: A Note on Niiniluoto", *Erkenntnis*, 1981 (Volume 16), pp. 35-52

- Sternberg, R. J. *Intelligence, Information Processing and Analogical Reasoning: The Componential Analysis of Human Abilities*. Hillsdale, NJ: Laurence Erlbaum Associates, 1977
- Swinburne, Richard. *The Justification of Induction*. New York: Oxford University Press, 1974
- Thibadeau, Robert. "Reaching for an Understanding about Analogy". Rutgers Department of Computer Science Report No. DCS-TM-8, May 1977
- "This House believes that sexual freedom leads to ruin". Oxford Union Society Debate, Oxford, October 12, 1984
- "This House calls for a non-nuclear defence policy for Britain". Oxford Union Society Debate, Oxford, November 27, 1984
- Thomson, J. J. "A Defense of Abortion", *Philosophy and Public Affairs*, Fall 1971 (Volume 1), pp. 47-66
- Trusted, Jennifer. *The Logic of Scientific Inference: An Introduction*. Basingstoke, UK: Macmillan Press, 1979
- Turing, A. M. "Computing Machinery and Intelligence", in Daniel C. Dennett and Douglas R. Hofstadter, arrangers and composers. *The Mind's I: Fantasies and Reflections on Self and Soul*. New York: Bantam Books, Inc., 1981
- Tversky, Amos. "Features of Similarity", *Psychological Review*, July 1977 (Volume 84, Number 4), pp. 327-352
- Urmson, J. O. "The History of Philosophical Analysis" in Richard Rorty, editor. *The Linguistic Turn*. Chicago: University of Chicago Press, 1975, pp. 294-301
- von Wright, George Henrik. *A Treatise on Induction and Probability*. Paterson, NJ: Littlefield, Adams & Co., 1960
- Walter, David. *The Oxford Union: Playground of Power*. London: Macdonald & Co (Publishers) Ltd, 1984
- Weitzenfeld, Julian S. "Valid Reasoning by Analogy", *Philosophy of Science*, 1984 (Volume 51), pp. 137-149
- Wilcox, James R. and H. L. Ewbank. "Analogy for Rhetors", *Philosophy and Rhetoric*, 1979 (Volume 12, Number 1), pp. 1-20
- Wilson, Harold. The Kennedy Memorial Lecture. Oxford Union Society, Oxford, November 30, 1984
- Wilson, P. R. "On the Argument by Analogy", *Philosophy of Science*, 1964 (Volume 31), pp. 34-39
- Winston, Patrick H. "Learning and Reasoning by Analogy", *Communications of the Association for Computing Machinery*, December 1980 (Volume 23, Number 12), pp. 689-703
- Wittgenstein, Ludwig; G. E. M. Anscombe, translator. *Philosophical Investigations*. New York: The Macmillan Company, 1953





