SYNTHESE LIBRARY / VOLUME 299

ERIC CHARLES STEINHART

THE LOGIC OF METAPHOR

Analogous Parts of Possible Worlds

SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

ISBN 978-90-481-5712-9 ISBN 978-94-015-9654-1 (eBook)
DOI 10.1007/978-94-015-9654-1

Originally published by Kluwer Academic Publishers in 2001.

Softcover reprint of the hardcover 1st edition 2001

Table of Contents

Acknowledgements

I'm standing on the shoulders of giants. The theory I present evolved from three main sources: Kittay's semantic field theory of metaphor, Thagard & Holyoak's work on analogy, and Hintikka & Sandu's work on metaphor and possible worlds semantics. I appreciate the support and encouragement of Eva Kittay, Paul Thagard, Jaakko Hintikka, James Moor, Daniel Kolak, Donn Welton, David Warren, Marcello Dascal, Earl MacCormac, Bipin Indurkhya, Robert Fogelin, Walter Sinnott-Armstrong, Kathleen Wallace, Jhan Hochman, Patrick Grim, and Celian Schonbach. Long ago, Carl Hausman at Penn State and Richard Cobb-Stevens at Boston College got me interested in metaphor. Particular credit goes to Eva Kittay for her constant attention to this project over many years. Many loud squawks of thanks to Kathleen Wallace. I appreciate the support of the philosophy departments of SUNY at Stony Brook, William Paterson University, and Dartmouth College.

1
Introduction

Eric Charles Steinhart[1]

(1) William Paterson University of New Jersey, Wayne, New Jersey, USA

1. Metaphors and Logic

Metaphors are among the most vigorous offspring of the creative mind; but their vitality springs from the fact that they are logical organisms in the ecology of language. I aim to use logical techniques to analyze the meanings of metaphors. My goal here is to show how contemporary formal semantics can be extended to handle metaphorical utterances. What distinguishes this work is that it focuses intensely on the logical aspects of metaphors. I stress the role of logic in the generation and interpretation of metaphors. While I don't presuppose any formal training in logic, some familiarity with philosophical logic (the propositional calculus and the predicate calculus) is helpful. Since my theory makes great use of the notion of structure, I refer to it as the *structural theory of metaphor* (STM). STM is a *semantic theory of metaphor:* if STM is correct, then metaphors are cognitively meaningful and are nontrivially logically linked with truth.

I aim to extend possible worlds semantics to handle metaphors. I'll argue that some sentences in natural languages like English have multiple meanings: "Juliet is the sun" has (at least) two meanings: the literal meaning "$(\text{Juliet is the sun})_{\text{LIT}}$" and the metaphorical meaning "$(\text{Juliet is the sun})_{\text{MET}}$". Each meaning is a function from (possible) worlds to truth-values. I deny that these functions are identical; I deny that the metaphorical function is necessarily false or necessarily true. I'll argue that most (but not quite all) metaphors are based on analogies. Analogy is the relative structural indiscernibility of parts of worlds. So: a metaphor is true at a world if and only if certain parts of that world are relatively structurally indiscernible (they are analogous). I'll also argue that metaphors that aren't based

on analogies are still based on relative indiscernibilities of parts of worlds. I'll talk about the meanings of metaphors in terms of analogical accessibility and analogical counterparts.

I'll develop all this both informally and formally. Formally: I'll develop an intensional version of the predicate calculus (an extended predicate calculus) that has mechanisms for interpreting English sentences in terms of thematic roles and eventlike entities (occurrences). My formal truth conditions for analogies and metaphors are based on the extended predicate calculus and its models. Since truth conditions are sometimes uninformative (even if correct), I'll also talk about confirmation conditions for metaphors. I'll indicate how metaphor is related to abductive inference and explanatory coherence. I'll show how to extend natural deduction systems to handle inferences that (dis)confirm metaphors. Science makes extensive use of theory-constitutive metaphors (e.g. "Light is a wave", "The brain is a spinglass", and "The immune system is the nervous system"). I aim to explain why such ampliative metaphors are scientifically legitimate.

While striving for formal precision, I also strive for empirical generality. One of the most important aspects of my project is its insistence on developing rules for generating and interpreting a wide variety of grammatical classes of metaphor.[1] Most metaphor theories consider only noun-identifications ("Juliet is the sun") or noun-predications ("Socrates is a midwife"); but STM is grammatically general, and is able to handle metaphors involving verbs ("Theaetetus gives birth to an idea"), adjectives ("Sharp minds are intelligent"), and so on. STM is therefore superior to theories able only to handle a few grammatical classes of metaphor. Moreover, by using standard linguistic and logical notions, such as re-write rules and truth-conditions, STM is syntactically and semantically extendible.

To validate the consistency and empirical adequacy of STM, I built a computational model of it. The result is a working computer program, called NETMET, that realizes the rules I posit for competence regarding metaphors. Using NETMET, you can examine STM yourself. The success of NETMET has methodological consequences: it shows how the computer can be applied to a philosophical problem. NETMET is a very general

analogy and metaphor engine serviceable for a variety of philosophical tasks.[2] In what follows, I attend both to building my theory of metaphor and to my realization of it in NETMET.

---

2. Metaphor and Possible Worlds Semantics

2.1 Logical Truth-Conditions for Metaphors

According to Davidson (1979), metaphors are supposed to mean just what their words literally mean. So: what does the sentence "Juliet is the sun" literally mean? One answer is that "Juliet is the sun" literally means that (is literally true if and only if) Juliet is the sun. However: there are at least 5 logically distinct senses of "is".[3] Each sense involves its own truth-conditions. There is the "is" of numerical-identity: "Cicero is Tully" the "is" of sortal-predication: "John is human"; the "is" of property-predication "John is white" the "is" of intertheoretic-reduction: "Temperature is the average kinetic energy of molecules"; the "is" of role-occupancy: "Mel Gibson is Hamlet" (Shapiro, 1997: 83). Syntax alone does not always decide the sense: "Pauline Reage is Dominique Aury" and "Jim Carrey is Andy Kaufmann" are syntactically equivalent but semantically distinct uses of "is".[4] Which is the literal sense of "is"? It would be better to speak of the *logical senses* of "is".

I suggest there is a sixth logical sense of "is": the "is" of *counterpart correspondence*. In this sense, x is y if and only if x in situation T is the counterpart of y in situation S. Here's how Dretske (1991) uses the "is" of counterpart correspondence:

> Let this dime on the table be Oscar Robertson, let this nickle (heads up) be Kareem Abdul-Jabbar, and let this nickle (tails uppermost) be the opposing center. These pieces of popcorn are the other players, and this glass is the basket.... I can now by moving coins and popcorn around on the table, represent the positions and movements of these players.... The coins and the popcorn have been assigned a temporary function, the function of indicating (by their positions and movement) the relative positions and movements of certain players during a particular game. (pp. 52–53)

Dretske uses the "is" of counterpart correspondence when he says "These pieces of popcorn *are* the other players, and this glass *is* the basket." He also uses it in the explicit statements of correspondence: "Let this dime *be* Oscar Robertson, let this nickle (heads up) *be* Kareem Abdul-Jabbar, and let this nickle (tails uppermost) *be* the opposing center". The counterparts are based on a positional analogy. The "is" of (analogical) counterpart correspondence is the metaphorical sense of "is". The metaphorical sense of "is" is one sense of "is" (along with the numerical-identity, sortal and property predication, role-occupancy, and intertheoretic-reduction senses). Each particular sense has its own logical truth-conditions. I'll argue that the metaphorical sense is logical whether or not it is "literal".

I'll agree that any sentence of the form $(x \text{ is } y)_{LIT}$ is true at world W if and only if x at W is numerically identical with y at W. So: "$(\text{Juliet is the sun})_{LIT}$" is true at W if and only if Juliet at W is numerically identical with the sun at W. But that's not the whole story: "Juliet is the sun" is ambiguous. Its other meaning is metaphorical. I will argue that any sentence of the form $(x \text{ is } y)_{MET}$ is true at world W if and only if there are situations S and T in W such that x in its situation T is the counterpart of y in its situation S. The term "situation" is often linked with the semantic theories of Barwise & Perry (1999). I won't follow Barwise & Perry, but I will use the term situation — it's the best English word for the kinds of structures I want to talk about. Situations are parts of logical space. Any situation contains some individuals with some properties that stand in relations to one another.

For example: "$(\text{Juliet is the sun})_{MET}$" is true at W if and only if there are situations S and T in W such that Juliet in T is the counterpart of the sun in S. The situations S and T are plain from Shakespeare's text: "(Juliet appears above at a window) ROMEO: But, soft! what light through yonder window breaks? It is the east and Juliet is the sun! Arise, fair sun, and kill the envious moon" (Shakespeare, *Romeo and Juliet*, 1974, Act II Scene II, p. 751). Just as the sun appears in the east, so too Juliet appears at her window.

I assume that literal and metaphorical truth-conditions are both expressed in some logical language — an intensional predicate calculus. I'll use an extended predicate calculus (the XPC) that

involves thematic roles and generalized events (occurrences). I do not aim to give literal paraphrases of metaphors; I aim to give *logical paraphrases* of metaphors.[5] I don't doubt that logical paraphrases miss some of the most exciting aspects of metaphor: the tension, the aesthetic brilliance, the rhetorical force. But I'm interested in truth. I'm interested in the *cognitive meanings* of metaphors.

2.2 Analogical Access and Counterparts

Possible worlds semantics argues that reality in its least restricted form fills a logical space that is divided into parts known as *worlds*. We inhabit one of these worlds: the actual world. Other worlds are possible relative to our world. The other worlds contain the ways things might have been: Bob Dole might have won the election in 1996. So: in some world, Bob Dole does win the election in 1996. But there's a problem: it is hard to see how the Bob Dole who wins the election in that world is numerically identical with the Bob Dole who loses the election in our world. Identicals are supposed to be indiscernible. One way to solve the problem of *trans-world identity* is to deny that one individual inhabits many worlds: Bob Dole in our world has a *counterpart* at some other world who is exactly like our Bob Dole up to the election in 1996. Our Bob Dole loses; his counterpart wins. This is David Lewis's idea.[6] It is controversial; it is not the only way to deal with the problem of trans-world identity (which may not even really be a problem). Lewis's notion of counterparts won't work for metaphors. Analogical counterpart theory has to allow individuals to have many counterparts in distinct situations in the same world.

Hintikka has often argued that worlds can be small; he calls them scenarios (1983). I follow his lead: situations are small worlds; they are fine-grained parts of logical space. Possible worlds in the traditional sense are just special situations (the spatio-temporally-causally closed ones, or the maximally consistent ones, or whatever). I refine these ideas later. For now I just want to sketch some links between accessibility, counterparts, and metaphors. If reality in some grand sense consists of a plurality of worlds, then:

> We often quantify restrictedly over worlds, limiting our attention to those that somehow resemble ours, and we

> call this a restriction to "accessible" worlds. And we often quantify restrictedly over possible individuals, limiting our attention to those that somehow resemble some given this-worldly individual, and I call this a restriction to "counterparts" of that individual. (Lewis, 1986: 234).

For metaphors, accessibility is (almost always)[7] analogy: situation S is accessible from situation T if and only if S is analogous to T. If S is analogous to T, then there is some function *f* that correlates the individuals (and events and maybe properties) of S with those of T; that function *f* is usually known as an *analogical mapping function*. More philosophically, it is a counterpart function. If x is in situation S, then *f*(x) in T is the counterpart of x in S. The counterpart relation was intended to resolve problems of trans-world identity. It looks to me like metaphorical identifications like "Juliet is the sun" are *trans-situation identities* — identifications across small analogous parts of logical space. So I think that some appropriately modified version of counterpart theory that is able to handle situations and analogies is able to provide truth-conditions for metaphors generally.

Possible worlds semantics provides some good resources to deal with metaphors: logical space, situations, worlds, accessibility, counterparts. It is useful for other reasons. For example: some metaphors (as well as similes) involve comparisons with things that don't actually exist: "Tornadoes are vacuum cleaners from the sky" compares tornadoes with things that don't actually exist; so, possible worlds are needed to avoid vacuous reference. Indeed: in Shakespeare's metaphor "Juliet is the sun", Juliet is a fictional character who does not inhabit our world. If that metaphor is true, it is true only in worlds of which Shakespeare's *Romeo and Juliet* is true. Another reason is that Hintikka and Sandu (1994) have sketched a theory of meaning for metaphors using PWS, and their theory in many ways complements Kittay's semantic field theory of metaphor (SFTM). I don't doubt that there are problems with both SFTM and PWS. Both those theories require work. Doubtless their combination is even more troublesome than either theory by itself. Still, an extension of PWS to metaphors brings metaphor into the best semantic theory available today. It also brings metaphor closer to metaphysical issues.

I aim to show that much of the semantic machinery used in PWS for literal statements (e.g. intensions) can be successfully applied to metaphorical statements as well. To do this, it will be necessary to dispel two big myths about the nature of metaphor. First: proponents of PWS, fearful that PWS could not easily accommodate metaphor, have tended to treat it as a matter of pragmatics, a matter of language use. For example: in their recent textbook of PWS, Chierchia & McConnell-Ginet (1991: 161) assign only one sentence to metaphor: “Fresh metaphors and many other figurative uses of language are also to be understood in terms of multiple levels of speaker’s meaning”. Second: proponents of metaphor have tended to treat metaphor as a counter-example to PWS. Lakoff’s (1987) & Johnson’s (1987) declamations against PWS (which they refer to as “objectivism”) are by now well known. I think that Kittay, and Hintikka & Sandu, have shown that the conflict between PWS and metaphor is misguided. I aim to show that metaphors have truth-conditions that fit in just fine with logical projects like Chierchia & McConnell-Ginet’s.

3. Analogical Counterparts

3.1 Hobbes’s Analogy A STATE IS AN ORGANISM

Metaphors involve the comparison of target and source situations. Aspects of the source situation are *systematically correlated* with aspects of the target situation. The correlation is a function from the set of source objects to the set of target objects. Specifically: the correlation is an analogy. To see how source objects are correlated with target objects, it helps to look at a text in which the correlations are explicit. Hobbes (1962) thinks of the state as an artificial human organism:

> Art goes yet further, imitating that rational and most excellent work of nature, man. For by art is created that greate LEVIATHAN called a COMMONWEALTH, or STATE, in Latin CIVITAS, which is but an artificial man... and in which the *sovereignty* is an artificial *soul*, as giving life and motion to the whole body; the *magistrates*, and other *officers* of judicature and execution, artificial *joints; reward* and *punishment*, by which fastened to the seat of

the sovereignty every joint and member is moved to perform his duty, are the *nerves*, that do the same in the body natural; the *wealth* and *riches* of all the particular members, are the *strength; salus populi*, the *people's safety*, its *business; counsellors*, by whom all things needful for it to know are suggested unto it, are the *memory; equity*, and *laws*, an artificial *reason* and *will; concord, health; sedition, sickness;* and *civil war, death*. (Introduction, p. 5)

In the Hobbesian text, the target situation is a (generic) state; the source situation is a (generic) man. Table 1 lists counterpart correspondences between the state and man. Hobbes is using the "is" of counterpart correspondence in "the sovereignty *is* an artificial soul" (my italics). He is not using the "is" of sortal-predication. So: "the sovereignty is an artificial soul" doesn't have the same logical form as "the sovereignty is a political institution"; so, it does not have the same kind of truth-conditions. When he writes "reward and punishment... *are* the nerves," (my italics) he is using the "is" of counterpart-correspondence. Thus "reward and punishment... are the nerves" doesn't have the same logical form as "reward and punishment are moral operations"; so, it does not have the same kind of truth-conditions.

Table 1. Correlations from state to body.

state	→	man	equity	→	reason
sovereignty	→	soul	laws	→	will
magistrates, officers	→	joints	concord	→	health
reward, punishment	→	nerves	sedition	→	sickness
wealth, riches	→	strength	civil war	→	death
counsellors	→	memory			

One interpretation of "the sovereignty *is* an artificial soul," provides it with truth-conditions like this: "the sovereignty *is* an artificial soul," is true if and only if (means that) the functional role of the sovereignty in a commonwealth is the same as the functional role of the soul in an organism. More precisely: "the sovereignty is an artificial soul" is true if and only if there is some role R such that the sovereign plays R in the commonwealth and the soul plays role

R in an organism. There is such a role: x plays role R in system y if and only if x rationally orders the activities of y. Likewise: “reward and punishment... *are* the nerves” is true if and only if the role of reward and punishment in the commonwealth is the same as the role of the nerves in an organism. If we say: x plays role Q in system y if and only if x enables the controlling center of y to regulate the parts of y by arousing or inhibiting their activity, then we have found a role played both by reward and punishment in the commonwealth and by the nerves in an organism; if that is right, then we have found a role that makes “reward and punishment... are the nerves” true.

The statements “the sovereignty is an artificial soul” and “reward and punishment are the nerves” are metaphors. They are rather prosaic metaphors; but they are metaphors.[8]

3.2 Swanson’s Analogy THE CELL IS A FACTORY

Metaphors are often used to introduce theories. Such pedagogical uses of metaphor are among the most effective techniques for teaching unfamiliar concepts. Metaphors are quite frequently used to teach scientific theories. If metaphors are creatively used falsehoods, any inferences from the metaphor to the content of the theory are accidental; the speaker can hardly be said to have any intentions, and it is truly surprising that such pedagogical uses of metaphor are effective. It is even more surprising (if metaphors are creatively used falsehoods) to find metaphors used to teach scientific concepts. So far from *facilitating* teaching, metaphor ought to *hinder* teaching. Yet it does not.

On my view, such uses are successful (that is, reliably effective) because the inferences from the metaphor to the content of the theory are rule-governed. There are rules for the interpretation of metaphors. Naturally, these rules are like all other linguistic rules: they have exceptions — a fact about natural language that seems to drive philosophers mad. In mathematics, a single counterexample refutes a theorem; in semantics of natural language, a single counterexample is merely an exception. Indeed, there are even classes of counterexamples (such as English verbs that form their past tenses by internal vowel changes). But back to metaphor. I will argue that metaphors are reliable and effective tools for communication because the same idea of truth is at work in both

metaphorical and literal language. Here, for example, is a metaphor used to illustrate the operation of the cell (Swanson, 1960: 26 – 41). The cell is thought of in terms of a factory:

> The cell... can be considered as a chemical factory. It may, of course, be a general-purpose factory, capable of performing all the services and of manufacturing all the products necessary to continue life; this must obviously be true in unicellular organisms. Or it may be a specialty shop, doing only a single job, such as serving as nerve cells for communication or as muscle cells for movement. Regardless of its nature, however, a cell, like a factory, must possess a certain organization in order to be efficient; it must contain a controlling or directing center, a source of supplies, a source of energy, and the machinery for making its product or performing its service. (p. 26)... The cell membrane, therefore, not only provides mechanical support and exterior form for our cellular factory, it also is very much a part of the living machinery of the cell. (p. 28)... the nucleus... is the controlling center, the board of directors of our cellular factory, for in it are found the chromosomes and the genes which somehow guide and determine the character, activities, and destiny of each individual cell. (p. 31)... The cytoplasm... is the main assembly line of the cell and its output is either a product... or a service... or a combination of the two. To do these things, the cytoplasm requires a source of raw materials, a source of power, the machinery necessary to do the job, and mechanisms for distributing its product or service. (p. 35)... The “powerhouse” of the cell is... another particle in the cytoplasm, the mitochondrion. (p. 37)... the highly ordered arrangement of membranes that make up this cellular background, which is called the... endoplasmic reticulum... we therefore look upon the endoplasmic reticulum as the principle manufacturing portion of our factory. (p. 41)... An efficiently operated factory is a planned, not a haphazard affair; its continuous operation requires, as we have pointed out, direction, power,

> machinery, and raw materials, and the parts must be related to the functions they perform. Nature has constructed cells along the same lines as we have constructed factories. (p. 41)

Table 2 lists the correlations of the cell and its parts with the factory and its parts. These correlations facilitate comprehension of something initially unfamiliar to students — the cell — in terms of something that is likely to be more familiar — the factory.

Table 2. Correlations from cell to factory.

cell	→	chemical factory
unicellular organism	→	general-purpose factory
nerve cell	→	specialty shop
membrane	→	mechanical support
nucleus	→	board of directors
cytoplasm	→	main assembly line
mitochondrion	→	powerhouse
endoplasmic reticulum	→	manufacturing area

In this metaphor, as in the Hobbesian metaphors, the correlations preserve the functional roles of the components in their respective systems. For example: “The nucleus is the board of directors of the cellular factory” is true if and only if (means that) the functional role of the nucleus in the cell is the same as the functional role of the board of directors in a factory (or manufacturing company). So: the nucleus in the cell is the counterpart of the board of directors in a manufacturing company. In what follows, I will bring out some heavy logical machinery to analyze counterpart correspondence for metaphors. I’ll analyze it in terms of analogical mapping functions across situations.

---

4. Theory-Constitutive Metaphors

4.1 Some Different Kinds of Metaphors

There are many different kinds of linguistic metaphor.[9] I distinguish between *poetic* or literary metaphors and what Boyd (1979) has called *theory-constitutive metaphors*. Theory-constitutive metaphors

are used productively in the sciences: “Electricity is a fluid”, “Light is a wave”, and “The mind is a computer program”, are all instances of theory-constitutive metaphors. But theory-constitutive metaphors also occur outside of the sciences: they occur in philosophy (“Memory is a wax tablet”), in theology (“God is light”), in politics (“A nation is a body”), in logic (“The null individual is the null set”), and in other disciplines. Theory-constitutive metaphors occur in engineering as well as in the more abstract disciplines: “A paintbrush is a pump” (Schon, 1979: 257 – 260) is an ampliative metaphor in which the theory of paintbrush function is reorganized according to the theory of pumps in order to solve an engineering problem. Theory-constitutive metaphors are *ampliative* — they use creative reasoning by analogy to generate novel and informative hypotheses that are significantly true or false, and that are tested like other theoretical hypotheses. My thoughts on creative reasoning by analogy are guided by Thagard’s *Mental Leaps: Analogy in Creative Thought* (1995).

I am mainly interested in theory-constitutive metaphors. They are usually more extensively elaborated than literary metaphors, and the discourses in which they occur already privilege truth over non-cognitive values such as aesthetic merit. My focus on theory-constitutive metaphors is not meant to exclude literary metaphors. I think the methods I develop to analyze theory-constitutive metaphors apply just as well to poetic metaphors. For instance, Kittay’s (1987: 287–8) analysis of the “bees of England” metaphor in Shelley’s “Song to the Men of England” uses the same techniques I use for more prosaic metaphors. Poetic metaphors also depend on counter-part-correlations in analogical situations.

Boyd (1979: 359–60) distinguishes theory-constitutive metaphors from pedagogical or exegetical metaphors. Pedagogical metaphors play a role in teaching theories already well-understood. For instance, a physics teacher might say “The atom is a miniature solar system” to introduce the notion of the atom to students already familiar with the solar system. The description of the solar-system in terms of the atom is dispensable once the theory of the atom is learned. But theory-constitutive metaphors are “those in which metaphorical expressions constitute, at least for a time, an irreplaceable part of the linguistic machinery of a scientific theory” (p. 360). Boyd offers the computer metaphor for the mind as an

example of a theory-constitutive metaphor.

Boyd lists further distinctive characteristics of theory-constitutive metaphors:

1. if successful, a theory-constitutive metaphor becomes "the property of the entire scientific community, and variations on [it] are explored by hundreds of scientific authors without [its] interactive quality being lost" (p. 361);
2. there is no reason to doubt that complete explication of theory-constitutive metaphors is possible, nor to doubt that "complete explications are often the eventual result of the attempts at explication which are central to scientific theory" (p. 362);
3. theory-constitutive metaphors display a kind of "inductive open-endedness" (p. 363); they suggest "strategies for future research" (p. 363) and have a power to programmatically orient research. Indeed, one can even view theory-constitutive metaphors as *discipline-constitutive*. For instance, the computer metaphor for the mind orients and drives the discipline of cognitive science.

4.2 Explication of Theory-Constitutive Metaphors

A theory-constitutive metaphor for some target consists of a metaphorical redescription of the target along with an analogy on the basis of which the target was metaphorically redescribed. Such a metaphorical redescription directs research concerning the target by assigning to it the task of interpreting all the metaphorical concepts and propositions in the metaphorical redescription of the target. Research into the target consequently involves the *production of a series of ever better logical approximations to the meaning of the metaphor*. Each of these logical approximations is a theory of the target that better approximates its metaphorical redescription. Each successive member of this series more precisely fixes the references or extensions of metaphorical concepts and provides more accurate truth-conditions for metaphorical propositions. The limit of this series is a theory in which all metaphorical concepts have fully fixed references or extensions, and all metaphorical propositions have perfectly accurate logical truth-conditions. In principle, it is possible for this series to converge towards this limit without ever reaching it. In practice, there is usually some point at

which the productivity of the metaphor is exhausted.

For instance, the LIGHT IS A WAVE metaphor constitutes a theory of light by directing optical research at the production of a series of ever better logical approximations to the metaphorical concepts and propositions analogically transferred from the domain of waves to that of light. Particularly, a logical approximation is better if it more precisely fixes the reference of the concept "luminiferous ether", and so provides more accurate logical truth-conditions for the proposition "Light travels through the luminiferous ether".

The construction of better logical approximations is often difficult, indirect, and fortuitous. Very accurate logical approximations to the meaning of the MEMORY IS A WAX TABLET metaphor are provided by thermodynamic theories of memory (Smolensky, 1986). Surprisingly, such theories, based on scientific understandings of information in terms of entropy, provide logical definitions for concepts like *computational temperature*, and so provide highly accurate logical truth-conditions for metaphors like "Heat melts memory" and "Memory freezes as it gets colder."

4.3 Examples of Theory — Constitutive Metaphors

Theory-constitutive metaphors are alive and well in recent scientific research (Thagard, 1995: ch. 8). Three theory-constitutive metaphors deserve further philosophical study: (1) the metaphorical interpretation of neural networks in terms of logical switching circuits; (2) the metaphorical interpretation of neural networks in terms of spin glasses; and (3) the metaphorical interpretation of the immune system in terms of the nervous system. All these metaphors are controversial; the fact that they are controversial is good — it means that they are (or have been) actively debated; but cognitively meaningless statements (i.e. nonsense) isn't actively debated; so, these metaphors aren't cognitively meaningless. There are many other recent theory-constitutive metaphors besides these. Here they are:

1. NEURAL NETWORKS ARE SWITCHING CIRCUITS. The discovery that neurons appear to act like switches lead to the idea that neural networks are networks of logic gates. The analogy was developed by McCulloch & Pitts (1943); it became the basis for the computational model of the nervous system used

extensively in cognitive science. The metaphor has seen much discussion (Boden, 1981: ch. 1; MacCormac, 1985: ch. 1).

2. NEURAL NETWORKS ARE SPIN GLASSES. Spin glasses are semi-crystalline materials with the ability to store patterns (Stein, 1989). Cragg & Temperley (1954) recognized some similarities between neural networks and semi-crystalline lattices but did not develop those parallels. Hopfield (1982, 1984) explicitly developed the formal analogies between spin glasses and neural networks. Hopfield's work became part of the connectionist approach to cognition (Cowan & Sharp, 1988: 97 – 101).

3. THE IMMUNE SYSTEM IS A NEURAL NETWORK. The human immune system is enormously complex and very poorly understood (Benjamini et al., 1996). To account for the complex adaptive behaviors of the immune system, Jerne (1974) proposed that the IS is a kind of network. Since then an enormous amount of work has been done on immune networks. I list only a few items here: Segel & Perelson (1988); Varela et al.(1988); Perelson (1989); Varela & Coutinho (1991); Vertosick & Kelly (1991); Rowe (1994); Takumi & De Boer (1996); Leon (1998). Immune networks are sometimes explicitly developed in terms of neural networks (Vertosick & Kelly, 1991; Roshi, 1996; Dasgupta, 1997). The analogy has also been turned the other way: principles of the immune system have been used to develop theories of neural nets (Hoffman, 1986). In general, immune networks are thought of as connectionist or parallel distributed processing models just like neural networks, even when the comparisons are not made explicit. All network models of immune functions are controversial; it is far from clear where these models will go. More philosophical work should be done in this area (see Levy, 1988; Tauber, 1994).

4. TECHNOLOGY IS AN ORGANISM. The exosomatic organ theory of technology (first articulated by Ernst Kapp) is based on an analogy between technology and organisms (Feibleman, 1979). Artifacts are said to be exosomatic organs (imitations or extensions of natural organs that are projected or externalized outside of human bodies). So, clothes are exosomatic skin; the Internet is an exosomatic nervous system. Since it ties

the erotic-tactile pleasure he experiences in kissing her lips is the same kind of pleasure as the gustatory-tactile pleasure he experiences in eating cherries.

(3) "(Her lips are cherries)$_{MET\text{-}2}$" is true if and only if (means that) the functional role of her lips in the cannibal feast is the same as the functional role of cherries in our cooking: her lips play the role of garnishes in the cannibal culinary arts just as cherries play the role of garnishes in our culinary arts.

5.2 Meanings and Contexts

I'm interested only in utterances (words, phrases, sentences) that are generated (said, written) and interpreted (heard, read) by competent language-users. I'm interested only in utterances that have contexts. Each meaning of an utterance is true or false relative to some context. To see that context matters, consider the following three ways contexts fix truth-values for "Her lips are cherries":

(1) "(Her lips are cherries)$_{LIT}$" is *false* when said of the context of him kissing her insofar as it is sensuous; it is *false* when said of the context of cannibals who use human lips in their gruesome feasts.

(2) "(Her lips are cherries)$_{MET\text{-}1}$" is *true* when said of the context of him kissing her insofar as it is sensuous; it is *false* when said of the context of cannibals who use human lips in their gruesome feasts.

(3) "(Her lips are cherries)$_{MET\text{-}2}$" is *false* when said of the context of him kissing her insofar as it is sensuous; it is *true* when said of the context of cannibals who use human lips in their gruesome feasts.

The three meanings of "Her lips are cherries" are distinct because they distinctly correlate circumstances (contexts or situations) with truth-values. The literal meaning of "Her lips are cherries" involves only her lips and cherries. But the metaphorical meanings of "Her lips are cherries" involve many additional items in the contexts of which they are true. For instance: MET-1 involves some man, an act of kissing, an act of eating, and two different pleasures of the same kind; MET-2 involves cannibals, feats,

are oral sensations.	they are oral sensations.

5.4 Descriptions as Contexts for Metaphors

We can render the two stories from Table 3 more precisely using some logical notation. The logical notation clarifies the complexities and ambiguities of English. The stories are treated as lists of expressions written in a notation that resembles the logician's predicate calculus. Each expression is a *proposition*. For example: "P1: x1 is a person" and "P7: x1 kisses x4" are propositions. Propositions have parts. Each proposition contains an *index* ("P1", "P7"), a *predicate* ("is a person", "kisses"), and some *names* ("x1", "x4"). Propositions are true or false. I assume the propositions in the target and source stories are true. Each part of any true proposition *refers to* some part of reality. The index usually refers to an *event;* the predicate refers to a *property* or *relation*; the names refer to *individuals*. Table 4 shows the propositions in the source and target stories S and T. The indexes of the relevant propositions are shown in bold in Table 4.

Table 4. Target and source descriptions.

Description T: Kissing	**Description S: Eating**
P1: x1 is a person;	**Q1:** y1 is a person;
P2: x2 is a woman;	Q2: y2 is some cherry tree;
P3: x3 are the lips of x1;	**Q3:** y3 are the lips of y1;
P4: x4 are the lips of x2;	Q4: y4 are some cherries on y2;
P5: x4 are red;	Q5: y4 are red;
P6: x4 are plump;	Q6: y4 are plump;
P7: x1 kisses x4;	Q7: y1 eats y4;
P8: x3 makes oral contact with x4;	**Q8:** y3 makes oral contact with y4;
P9: if P7 then P8;	**Q9:** if Q7 then Q8;
P10: x5 is a pleasurable sensation;	**Q10:** y5 is a pleasurable sensation;
P11: x5 comes from the mouth of x1;	**Q11:** y5 comes from the mouth of y1;
P12: x1 experiences x5;	**Q12:** y1 experiences y5;
P13: if P8 then P12;	**Q13:** if Q8 then Q12;

Every true proposition corresponds to some *fact* or state-of-affairs. States-of-affairs are ways things are, either by themselves ("Phantom is a cat") or in relation to other things ("Phantom sits on

the mat"). Lists of propositions are *descriptions*. A description is true if and only if each of its member propositions is true. True descriptions correspond to systems of facts. Systems of facts are *situations*. Facts are parts of situations. Situations are parts of worlds. True propositions correspond to facts; sets of true propositions are descriptions that correspond to (are true of) objective situations. What is true in the description is real in the situation. If the stories are true, the situations they describe are real: the describing story and the described situation have the same logical structure. If the source and target descriptions are analogous, and if they are both true, then the source and target situations are analogous. The analogy is real.

5.5 Shared Syntactical Patterns of Descriptions

Writing the stories as lists of propositions makes their logical structure apparent. It's easy to see from Table 4 that the form of the description S resembles the form of the description T. The descriptions are arrangements of symbols that share a common syntactical pattern. Much of chapters 4 and 5 is devoted to developing symbol-manipulation techniques (computational tools) to discover and manipulate they shared syntactical patterns that occur in analogous descriptions. One way to specify that common pattern is to use *abstract propositions*. If we use Greek letters as meta-variables, then we can say that "P3: x3 are the lips of x1;" and "Q3: y3 are the lips of y1;" both instantiate the abstract proposition "Φ3: α3 are the lips of α1". Another way to specify the common pattern is to build a *translation dictionary*. Start with "Q3: y3 are the lips of y1"; replace "Q3" with "P3", "y3" with "x3", and "y1" with "x1"; the result is "P3: x3 are the lips of x1". Since the translation transforms a proposition in the source into one in the target, the replacements capture shared structure. Table 5 shows the shared structure of S and T. If descriptions S and T share any propositions that involve relations, then S is analogous to T; the translation dictionary from S to T is a correspondence that specifies their analogy.

Table 5. Propositions common to target and source descriptions.

Shared Abstract Propositions	**Translation Dictionary**	
Φ1: α1 is a person;	Q1 → P1	y1 → x1
Φ3: α3 are the lips of α1;	Q3 → P3	y3 → x3

any metaphor specifies the structure shared by the two situations involved in the metaphor. The ground is the analogy. Metaphors are true if and only if the analogies on which they are based are true. Analogies take many forms. Note that I use angle brackets "<" and ">" to enclose schemas (they're like Quine's corner-quotes). Here is one form: <A is to B as C is to D> is true if and only if there exists some relation R such that R(A, B) and R(C, D). Such analogies are proportional. Proportional analogies define counterpart relations: A is the counterpart of C and B is the counterpart of D.

Analogy is logically deep: there is much more to be said about it than I say in this section. This initial account is merely introductory. In general, analogy is a partial relational indiscernibility. It is a kind of relative indiscernibility. Analogy in its purest and most powerful form is *isomorphism* — perfect relational correspondence. Analogies are real if and only if parts of worlds (situations) have similar relational structure. Here I briefly analyze three metaphors in terms of analogy: (1) "Juliet is the sun" (2) "Her lips are cherries"; and (3) "Sally is a block of ice". These analyses are not intended to be complete. To serve my introductory purposes. They focus only on truth. I do not aim to deal with the aesthetic or rhetorical aspects of these metaphors here.

5.7.1 Juliet is the sun

Consider this short text from Shakespeare's *Romeo and Juliet*, Act II Scene II (1974: 751): "(Juliet appears above at a window) ROMEO: But, soft! what light through yonder window breaks? It is the east and Juliet is the sun! Arise, fair sun, and kill the envious moon". The text lays out two correspondences: the window is the east, Juliet is the sun; it suggests another: the moon is something too. The text specifies an analogy: Juliet is to the window as the sun is to the east; just as the sun appears in the east, so also Juliet appears in the window.

The metaphor "Juliet is the sun" means that Juliet at her window is the counterpart of the sun in the east. Since "Juliet" denotes an individual in the target, I write it a "$(\text{Juliet})_T$"; since "the sun" denotes an individual in the source situation S, we can write it "$(\text{the sun})_S$". If "$(x)_T$" and "$(y)_S$" denote individuals, then "$((x)_T\ (\text{is})\ (y)_S)_{MET}$" is true if and only if $(\exists R)(\exists C, D)(R(x, C) \& R(y, D))$. More precisely:

pleasurable sensation y5, and if the event of kissing her lips is the counterpart of the event of eating the cherries, the truth-conditions in (1) can be refined and edited to become more specific truth-conditions for "((Her lips)$_T$ (are) (cherries)$_S$)$_{MET\text{-}1}$". For example:

(2) "((Her lips)$_T$ (are) (cherries)$_S$)$_{MET\text{-}1}$" is true if and only if (means that) the erotic-tactile pleasure he experiences in kissing her lips is the same kind of pleasure as the gustatory-tactile pleasure he experiences in eating cherries.

5.7.3 Sally is a block of ice

Consider "Sally is a block of ice" (Searle, 1979). One analysis says "(Sally is a block of ice)$_{MET}$" means that Sally metaphorically has the property of being cold while a block of ice literally has the property of being cold. I reply that "(Sally is a block of ice)$_{MET}$" means that the feelings aroused in the emotional sensory system by social contact with Sally are analogous to the feelings aroused in the thermal sensory system by tactile contact with a block of ice. Sally and the block of ice are counterparts in a pair of fairly complex situations that share relational structure. The analogy is illustrated in Table 7.

Table 7. Sally is a block of ice.

Source Situation S	**Target Situation T**	**Analogy *f***
S1: person(A1):	T1: person(B1):	$S_i \rightarrow T_i$
S2: block-of-ice(A2);	T2: person(B2);	$A_i \rightarrow B_i$
S3: touches(A1, A2);	T3: touches(B1, B2);	
S4: if S3 then S5;	T4: if T3 then T5;	
S5: contacts(A1, A2);	T5: contacts(B1, B2);	
S6: physical(S5):	T6: physical(T5);	
S7: cold(A3);	T7: emotion(B3);	
S8: arouses(S5, A3);	T8: arouses(T5, B3);	

6. Philosophical Application of the Computer

Although the central goals of this work concern metaphor, its purpose is peripherally methodological. This work aims to contribute to the advance of philosophical method through its use of the

	∴ (My car guzzles gas)$_{MET}$ is partially and approximately equivalent to (My car rapidly burns or leaks gas)$_{LIT}$.

2.3 Inference to the Best Literal Paraphrase

The use of forward and backward inference seems to have led to a significant literal paraphrase of the original metaphor (My car guzzles gas)$_{MET}$. Specifically: (My car guzzles gas)$_{MET}$ is true if and only if (My car rapidly burns or leaks gas)$_{LIT}$. Which is to say: "My car guzzles gas" *is metaphorically true* if and only if "My car rapidly burns or leaks gas" *is literally true*. Although my goal is to produce logical paraphrases of metaphors (to give them intensional truth-conditions in the XPC), I have nothing against literal paraphrases. Indeed: *metaphors produced by perfect analogies have exact literal paraphrases*.

If there were no good literal paraphrases of metaphors, then they would not become conventionalized (dead). Since dictionaries contain many examples of dead metaphors, there are good literal paraphrases of metaphors. If some literal paraphrase is sufficiently useful,[1] it becomes conventional, and the metaphor dies. For instance: "guzzles" gains a literal meaning in the automotive field. Forward-backward inference has produced a novel sense for the concept [guzzles] when applied to an AGENT that is an automobile and PATIENT gas. It produces the novel meaning postulate: [if x is a car and y is gas, then x guzzles y if and only if x rapidly burns or leaks y]. Metaphor interpretation is often (forward-backward) inference to the best literal paraphrase . Three worries: (1) dead metaphors can become reanimated ("the mouth of the river is laughing at us"); (2) sensory metaphors ("cold person") appear to be semantically primitive; (3) ampliative metaphors seem to lead to the creation of novel semantically primitive conventional meanings.

Literal paraphrase is really only useful for mere substitution metaphors: in the MINDS ARE KNIVES analogy, "sharp" is substituted for "smart" and "dull" for "stupid"; just so, "guzzles" is substituted for "rapidly consumes". Such metaphors are based wholly on positive analogies. They are not ampliative. Confirmed neutral analogies often generate novel primitive target meanings for terms with literal meanings in the source. The meanings of those terms are explained by their logical relations with other terms in the theories

that confirm them. For me the literal is essentially the conventional.[2] If neutral analogies generate ampliative metaphors, it may be necessary to reach more deeply into the logical foundations of language to construct their conventional meanings (e.g. metaphorical meaning postulates). I prefer to say that metaphors have logical paraphrases. Their logical paraphrases serve as the basis for any literal paraphrases they might have.

2.4 Inference in Metaphor Justification

Metaphors are often offered as *explanations*. Ampliative metaphors, also known as theory-constitutive metaphors, are explanatory metaphors. So: THE BRAIN IS A COMPUTER or SOUND IS A WAVE are explanatory metaphors.

For instance, Socrates offers the metaphors made by the SOCRATES IS A MIDWIFE analogy to explain at least four empirical facts: (1) the painful perplexity Theaetetus feels trying to articulate his thoughts; (2) the fact that Socrates only asks questions; (3) the fact that Socrates never states any positive doctrine; (4) the anger some of his students bear towards him.

Explanatory metaphors function logically as *hypotheses*. If the logic of scientific (dis)confirmation does not apply as well to metaphorical hypotheses as it does to literal hypotheses, then metaphorical hypotheses should not occur in scientific theorizing; but metaphorical hypotheses do occur in scientific theorizing; in fact, they occur regularly and as a matter of normal practice; so, assuming the logic of scientific (dis)confirmation applies to literal statements, I conclude that it also applies to metaphorical hypotheses. The logic of scientific (dis)confirmation involves both deductive and abductive inference.

Metaphorical hypotheses require *justification*. There are no reasons to restrict the search for justification to metaphorical hypotheses. We can wonder about the justification for any metaphor. Justification involves inference. It involves the same kind of inference as metaphor interpretation. Interpretation starts with a metaphor whose meaning is unknown; it searches for the meaning. Justification ends with a metaphor whose evidence is unknown; it searches for the evidence. Insofar as meaning is given by its truth-conditions, the literal meaning of any metaphor serves equally well as its literal evidence. Justification is just the inverse of interpretation.

Interpretation reasons from the metaphor to its literal evidence; justification reasons from the literal evidence to the metaphor. The inferential steps used in interpretation are reversed in justification. Table 2 illustrates this for the "Theaetetus gives birth to an idea" metaphor. Metaphor justification is important because it links metaphors with evidence in a way that makes them non-trivially cognitively meaningful.

Table 2. Forward-backward inference in metaphor justification.

Forward	**Forward**
L L ⇒ E ∴ E	(Theaetetus anxiously expresses an idea)$_{LIT}$. IF (Theaetetus anxiously expresses an idea)$_{LIT}$, THEN (Theaetetus painfully produces an idea)$_{LIT}$, (Theaetetus painfully produces an idea)$_{LIT}$.
Backward	**Backward**
E M ⇒ E ∴ M	(Theaetetus painfully produces an idea)LIT· IF (Theaetetus gives birth to an idea)$_{MET}$, THEN (Theaetetus painfully produces an idea)$_{LIT}$ ∴ (Theaetetus gives birth to an idea)$_{MET}$.
Forward-Backward	**Forward-Backward**

L
L ⇒ E
M ⇒ E
∴ M
	(Theaetetus anxiously expresses an idea)$_{LIT}$.
	IF (Theaetetus anxiously expresses an idea)$_{LIT}$
	THEN (Theaetetus painfully produces an idea)$_{LIT}$.
	IF (Theaetetus gives-birth to an idea)$_{MET}$,
	THEN (Theaetetus painfully produces an idea)$_{LIT}$.
	∴ (Theaetetus gives birth to an idea)$_{MET}$.

3. Metaphor Justification

3.1 Metaphors as Hypotheses

An argument that justifies a metaphor M is one that (1) has plausible premises; (2) uses plausible inference methods; and (3) has the metaphor M as its conclusion. I discuss such arguments here. This section is very technical. It's not as important as section 4 on the Black-Tourangeau method; if you're comfortable with the informal sketch of abductive justification of metaphors in section 3, and you don't want to get bogged down by logical detail, just read this introductory subsection (section 3.1) and skip to section 4. Two kinds of arguments exist that justify metaphors made by analogical