Volume 7
Issue 3 *Computer/Law Journal - Summer 1987*

Article 4

Summer 1987

Artificial Intelligence: Can Computers Understand Why Two Legal Cases Are Similar?, 7 Computer L.J. 409 (1987)

Celeste Tito

ARTIFICIAL INTELLIGENCE: CAN COMPUTERS UNDERSTAND WHY TWO LEGAL CASES ARE SIMILAR?

I. INTRODUCTION

Can computers be programmed to understand analogies between fact situations as lawyers do when they decide that two fact situations are alike? This question does not ask whether computers can arrive at the same *result* as lawyers; it asks whether they understand legal analogical reasoning. Computers can only do what humans program them to do. Unless legal theorists can describe how lawyers find and understand analogies between legal cases and hypotheticals, programmers cannot instruct computers to do the same. Legal theorists differ on whether there are rules which determine when two factual situations are similar. If no rules exist to describe why two cases are similar, or if theorists are unable to discover those rules, then computers cannot be programmed to emulate the legal reasoning process. If there is a plausible descriptive theory of how lawyers reason about similarities, then, theoretically, programmers may devise a set of rules that instruct computers to reason as lawyers do.

Part II of this Note will discuss competing theories of how lawyers reason about similarities. The discussion concludes that rules governing this reasoning process must exist. This Note does not propose an actual rule. Rather, it illustrates why any plausible rule describing why two legal cases are similar must answer these two questions:

(1) What properties of the entities being compared are legally relevant?
(2) What degree of sharing of those properties is sufficient to find similarity in the legal context?

Part III of this Note discusses two reasoning processes that lawyers must engage in to answer these questions. One process is understanding the meanings of words. Deciding which properties are legally relevant is at least a function of whether particular facts count for instances of general legal categories (whether they are "like" the other members of those categories) that judges recognize. Because no principle of deductive logic dictates when a particular entity is an instance of a general category, an inferential reasoning process must be used. This Note proposes that this inferential process *at least* depends upon understanding word meanings. Immanual Kant's theory of knowledge and Andrew Ortony's theory of metaphorical language are discussed to support the proposed theory of meaning. This theory, if accurate, gives reason to doubt whether digital computers can understand word meanings in the manner humans do.

The second reasoning process that lawyers must engage in to answer the questions about similarity is the process of moral decisionmaking. Choosing the appropriate level of generality at which to compare two entities requires a substantive moral theory about which level of detail is important in various legal contexts. This Note does not propose a substantive moral theory. Rather, it illustrates why a theory about moral decisionmaking is necessary for reasoning about legal similarities.

Part III of this Note concludes that understanding the similarities between two legal cases requires an ability to:

(1) *understand* natural language word meanings; and
(2) make moral decisions about which level of generality is appropriate for comparing facts in the legal context.

Part IV of this Note is a theoretical discussion of the way computers process information. This section illustrates why the limits of digital computing pose problems for machine understanding of the meanings of words. Comparing two hypotheticals illustrates how the difference between digital computing and the proposed theory of meaning causes problems for machine understanding of word meanings. Part IV also illustrates that, absent a substantive moral theory for choosing the appropriate level of generality to compare factual situations within the legal context, the computer's ability to reason as humans do about similarity is limited.

Part V of this Note distinguishes true artificial intelligence (AI) from expert systems and illustrates why natural language processing is the most troublesome obstacle to programming computers to reason as humans do. Part V describes several advanced natural language processing systems. This section discusses the strengths and weaknesses of various systems in understanding word meanings and in choosing relevant levels of generality at which to compare the factual situations that words describe. Because the problem of computer understanding of word meanings is a threshold to computer understanding of analogical thinking, this Note does not examine current research regarding computer modeling of moral decisionmaking. Natural language processing presents the greatest challenge for AI researchers. If researchers can determine how to program computers to understand *the meanings of words*, there is no apparent theoretical obstacle to programming any type of reasoning, including moral reasoning.

II. COMPETING THEORIES OF HOW LAWYERS DECIDE THAT TWO CASES ARE SIMILAR

What are some theories about how lawyers decide whether a fact situation is *like* a precedent case? Do lawyers apply *rules* to decide when two situations are similar within a legal context? Or, do lawyers just get a "hunch" or have a "flash of insight" that alerts them to similarities?[1] These questions are critical because even if "hunches" or other ruleless theories are plausible psychological descriptions of legal reasoning,[2] general rules for determining similarity are necessary to instruct computers to perform such reasoning.

Proponents of ruleless theories suggest that lawyers just "know similarities when they see them" and that no general principles determine when similarities exist. These theorists describe a reasoning process that is neither deductive nor inductive.[3] They argue that rules are unable to describe a property or set of properties that explain why two situations are similar.

Michael Moore criticizes this view as a plausible psychological description of how lawyers reason about similarities, but an implausible model of how lawyers know a finding of similarity is "right," or "similar enough" within the legal context:

> [W]ithout recourse to rules that include both the precedent case and the case to be decided, how do we get at the question of whether one case is similar, or similar enough, to another?[4] A ruleless theory supposes a world that I cannot imagine, a world in which particulars can be similar to one another but not similar in any particular respect, i.e., there need be no properties (describable by rules) that make one particular similar to another.[5]

The ruleless view is also objectionable because it treats "similarity" as a distinct property that can be possessed independently of any other properties.[6] "There is no property of 'similarity'; similarity is only relational."[7] It only makes sense to say that one negligence case is like another if there are shared properties. Similarity must have a basis; it cannot exist in a vacuum.

A more sensible position suggests that similarity need not be based on one specific describable property, but rather, similarity exists when

---

1. Moore, *A Natural Law Theory of Precedent*, manuscript, at 34, to be published in PRECEDENT IN LAW (L. Goldstein ed. 1987).

2. *Id.*

3. *E.g. id.* at 37 (citing J. WISDOM, PHILOSOPHY AND PSYCHOANALYSIS 248 (1969)).

4. Moore, *supra* note 1, at 34.

5. *Id.* at 38-39.

6. *See id.* at 39-40.

7. *Id.* at 40.

two entities share a significant portion of their properties.[8] This view is objectionable because it presupposes the question of *how many* properties the two entities must share to be similar.[9] This question is wrong. Similarity is not merely an issue of how many properties are shared; it is a question of the degree to which *relevant* properties are shared.[10] Knowing what common properties are relevant depends on *why* two entities are compared.[11] Moore illustrates this point:

> There is no context-independent way to specify the degree of likeness required for similarity. Are two emerald cut diamonds of different size and different imperfections similar? Who knows, until we know why we are classifying the stones. They may be alike enough to be called similar by a mineralogist interested in a general taxonomy of gems, not like enough to be called similar by a jeweler replacing a lost diamond.[12]

A plausible theory of how lawyers determine similarities between cases requires a general rule that specifies:

(1) what properties of the entities being compared are legally relevant; and
(2) what degree of sharing of those properties is needed for similarity *in the legal context*.[13]

III. NECESSARY REASONING PROCESSES FOR DETERMINING SIMILARITY

A. Understanding the Meanings of Words

In legal analysis, what properties are relevant? Particular facts are relevant if they constitute an example of general legal categories recognized by judges.[14] How do lawyers decide whether a particular fact counts for an example of a general legal category?[15] Fitting particular facts into general legal categories like "offer" or "breach" depends in part upon understanding the meanings of words. Lawyers ask themselves whether particular fact situations are *like* the concepts they have in mind when they use legal words such as "offer" or "breach." Because lawyers use words to describe both the particular facts and the general legal categories, the questions of word meanings arises. In this

8. *Id.*
9. *See id.* at 41.
10. *Id.* at 42.
11. *Cf. id.* (Moore indicates that one needs to know the purpose of the comparison in order to determine the level of generality appropriate for the comparison).
12. *Id.* at 41-42.
13. *Cf. id.* at 42.
14. This analysis is one step removed from the more fundamental question of what values inform judicial and legislative decisions about what *categories* are legally relevant.
15. No deductive law warrants the move from particular to general.

context, "words" refers to the actual physical symbols on a page. It is necessary to distinguish the discrete physical symbols from the concepts. The concepts include the particular facts and general legal categories that the words represent.

This Note proposes that humans understand word meanings by bringing to mind the concepts that they associate with the "things in the world" referred to by those words. The concept humans have in mind when they think of a word is not a finite entity. Rather, the concept of a word is a fluid, everchanging continuum of all the worldly experiences, both real and vicarious, which can "count for an instance of" or "be the referent of" a particular digital, noncontinuous symbol. This Note refers to these continuums as "continuums of referential space." To understand the meaning of a word is to understand the continuum of referential space that the discrete, finite symbol denotes. The conceptualization of the continuum of referential space that a word denotes changes over time as a person encounters more examples and counterexamples of the things in the world that the word describes.

Each discrete symbol attempts to capture a continuous idea; words are a shorthand reference to analogue thoughts. This notion is sensible because words cannot be defined by any set of necessary and sufficient conditions that apply in all cases.[16] Any discete list of attributes chosen to define a word will be adequate in some cases and inadequate in others.[17] For example, Webster defines "bird" as "warmblooded, two-legged vertebrates with feathers and wings."[18] Is a featherless "bird" still a bird? Are feathered bodies necessary for a creature to be a bird? Future data, such as examples and counterexamples, may show that certain attributes are or are not necessary and/or sufficient conditions to apply the term "bird".

Definitions that list attributes have value because they allow one to pick out the referent. Definitions, however, do not give the *meaning* of a word because they fail to offer every conceivable extension to which the word could refer.[19] This is true even of natural kind words because one can imagine future data showing that gold is not yellow or that cats are not animals.[20] "That examples of the kind in question have the

---

16. *See generally* Moore, *The Semantics of Judging*, 54 S. CAL. L. REV. 151 (1981) [hereinafter *The Semantics of Judging*].

17. *Id.*

18. WEBSTER'S NEW WORLD DICTIONARY OF THE AMERICAN LANGUAGE 143 (2d ed., 1979).

19. Donnellan, *Reference and Definite Descriptions*, in NAMING, NECESSITY, AND NATURAL KINDS 42 (S. Schwartz ed. 1977).

20. Kripke, *Identity and Necessity*, in NAMING, NECESSITY, AND NATURAL KINDS 66 (S. Schwartz ed. 1977).

properties they do is a matter of nature, not language."[21] Our certainty that water is H_20 is the certainty of a well established empirical theory and not the certainty that issues from knowledge of a definition."[22]

This concept applies especially to legal words, such as "contract" or "intent." Legal words are summary terms that can be explained only with the aid of exceptions that depict situations where the term does not apply or where it applies weakly. Lawyers build and modify (based on precedent) their idea of the continuum of referential space that is "what they mean" when they use a legal or any other word. This parallels Kant's notion of building and modifying a synthetic *a priori* idea of an object.

Immanuel Kant's theory of knowledge supports the contention that human thought and perception is continuous rather than digital. Kant maintained that human perceptions presuppose the existence of space and time. The sensory data that humans receive does not tell them anything about the world unless it is interpreted within a pre-existing mental structure that sees time and space as continuous.

To illustrate Kant's theory, imagine you are looking at a car. Looking at the car from the front, you see two headlights, the convex curve of two rubber tires, two windshield wipers, and numerous other items. From the side, you see two round tires, a door, perhaps only one windshield wiper, and no headlights. Absent a logical basis for knowing that objects exist,[23] no logical basis exists for concluding that what gave rise to your front view and side view is the same object. Do humans view two different cars at two discrete points in time, or do they have two views of the same car?

Kant would argue that the *a priori* concept of space and time continuity allows humans to constantly build and modify a synthetic *a priori* construct of "car," which is the source of the sense data. This *a priori* "car" has been built in the human mind over many years from continuous sense data, both examples and counterexamples, and changes constantly as new data is received. Humans have an *a priori* "car" in their mind that has continuity in space and time and changes appearances when viewed from different angles. A person's discrete sensory perceptions received when viewing the car from different angles do not contradict the belief that both views are of the same car.

21. NAMING, NECESSITY, AND NATURAL KINDS 27 (S. Schwartz ed. 1977).

22. *Id.* at 30.

23. Hume believes that *all* knowledge comes solely from sensory impressions. He believes that because there is no sensory impression of *actual* objects there can be no knowledge of the actual substance of objects in the world. Hume believes that we only *imagine* that the objects of our perceptions exist. Kant agrees with Hume that knowledge *begins with* sense data. Kant does not believe that this concession dictates that all knowledge arises out of the sense data.

This is because humans think about the world and interpret their perceptions of it within the context of a belief that time and space are continuous. The human belief that time and referential space are continuous enables humans to build these conceptualizations about the referents of words over time.[24] These conceptualizations are the *meanings of words*.

Humans use words, the discrete symbols of natural language, to communicate about similarities. The natural language techniques of simile, analogy, and metaphor all involve a reasoning process that allows for communication about likeliness between concepts. Andrew Ortony's theory of metaphorical language illustrates why understanding likeliness between concepts or factual situations depends upon understanding the meanings of words as continuums of referential space, as analogues for the discrete symbols.[25]

Metaphor provides a means to implicitly compare the concepts denoted by two terms.[26] Ortony's theory expands upon a model developed by I. A. Richard that describes the relationship between two terms. The "tenor", sometimes referred to as the "topic," is the term one is trying to say something about. The "vehicle" is the term being used metaphorically to form the basis of the comparison. The "ground" is that which the two terms have in common. The "tension" is the dissimilarity between the two terms.[27]

In understanding language, humans reconstruct a described event by including a great deal of what they already know about the world.[28] Humans form a continuous mental image that fills in the details between the digital linguistic sign posts, words. For example, Ortony describes his thoughts when reading that "a man swam the English channel in mid-winter:"

> I build a representation which invokes what I know about men and their capacity to swim, about what I know or believe (or even imagine) to be some characteristics of the English Channel and so on. What I invoke is largely experiential, perceptual and cognitive, and to this extent generally similar, but probably almost never identical, to what others invoke. I infer that the man was probably covered with oil, that he was strong and muscular, that the sea was likely cold and rough,

24. If the concept of continuity in space and time is innate, there is a serious question of how to build such a construct into a digital computer. If it is learned, then maybe a computer can be instructed to interpret words within the context of continuous space and time by building greater technological capacity for memory of past experiences.

25. Ortony, *Why Metaphors are Necessary and Not Just Nice*, 25 EDUC. THEORY 45, 46-48 (1975).

26. *Id.* at 45, n.3 (citing ARISTOTLE, RHETORIC, III., iv., 1-3).

27. *Id.* at 45 (citing I.A. RICHARDS, THE PHILOSOPHY OF RHETORIC (1936)).

28. *Id.* at 47. This is analogous to Schank's view that words serve as cues for retrieving expectations from memory, *see* R. SCHANK, *infra* note 68.

that the sky was perhaps gray and gloomy. I might also invoke my knowledge of likely public reaction, a reaction of admiration, incredulity, indifference or even alleged insanity.[29]

The mental process of filling in the details allows for language comprehension without requiring the message to explicitly spell out all the details.[30] Ortony urges that the mental process of filling in the details between digital words "is the language comprehender's digital to-analog converter; it takes him nearer to the continuous mode of perceived experience by taking him further away from the discrete mode of linguistic symbols."[31] Humans *possibly* move only nearer to the continuous mode of perception, and not completely to the full analogue meaning. Humans might add just enough detail, (like filling in the dots between the numbers on a clock so that it looks more like a circle), to enable them to make an adequate theoretical reconstruction of the analogue thought. This is consistent with Wilk's "laziness hypothesis," which proposes that humans seek out only enough information to solve a problem and no more.[32]

When a person states, "He dived into the icy water like a fearless warrior," a certain continuum of characteristics are brought to mind. Different people will think of different characteristics based on their experience and memories of warriors.[33] The simile directs attention first, to the salient characteristics of warriors, and second, to the subset of those salient characteristics that are transferable to the "topic" of the simile, "men who dive in icy water."[34] "Sets" of characteristics are continuums of cognitive and perceptual characteristics rather than lists of discrete attributes. For this Note's purpose it is easier to speak about characteristics as "sets," but the reader should keep in mind that the concepts are analogue rather than digital.

This second subset is determined by eliminating those characteristics that create a "tension," or somehow contradict or are conceptually incompatible with what one knows of the topic.[35] Ortony says, "what the simile is doing is, effectively, saying 'take all those aspects you know peculiar to fearless warriors which could *reasonably* be applied to a diving swimmer and predicate the entire set of them to the swimmer.' "[36]

Metaphor then, efficiently transfers a continuum of unspecified characteristics from the vehicle (fearless warriors) to the topic (men

29. Ortony, *supra* note 25, at 47.
30. *Id.*
31. *Id.*
32. *See* Wilks, *infra* note 109, at 67.
33. Ortony, *supra* note 25, at 48.
34. *Id.*
35. *Id.*
36. *Id.* (emphasis added).

who dive into icy water). Ortony argues that metaphor, simile, and analogy are not only efficient means for communicating about similarity, they are necessary for effecting such communication:

> Metaphors are necessary as a communicative device because they allow the transfer of coherent chunks of characteristics — perceptual, cognitive, emotional and experiential — from a vehicle which is known to a topic which is less so. In so doing they circumvent the problem of specifying one by one each of the often unnameable and innumerable characteristics; they avoid discretizing the perceived continuity of experience and are thus closer to experience.[37]

That is, human understanding of word meanings as continuums of referential space (developed over time by human perception, cognition, emotion, and experience) is necessary for communicating about likeness between concepts. Discrete words alone are insufficient for this task. Only by understanding the analogues of those symbols can one understand similarity. Thus, deciding whether or not a word describing a particular fact in a case is an instance of a general legal category requires an understanding of the word's meanings, the analogue of the digital symbol.

B. Making Moral Decisions About the Appropriate Level of Generality for Comparing Facts in the Legal Context

Choosing the appropriate level of generality for comparing two entities in a given context is important because even seemingly disparate things are alike at some level of abstraction. Likewise, apparently similar things can be distinguished if compared in detail, or at a low level of generality. What degree of sharing of relevant properties (those facts that count for instances of legal categories) is "similar enough" within the legal context? An understanding of *why* lawyers compare undecided cases to precedent cases answers this question.

A general rule about the *purpose* of using precedent cases is a nec-

---

37. *Id.* at 53. Ortony claims that in addition to serving as a short hand for predicating a "chunk" of characteristics, the metaphor enables predication or transfer of unnameable characteristics. *Id.* at 48. Ortony claims that for any given language, certain things are inexpressible. *Id.* at 49. He argues that "the continuous nature of experiences precludes the possibility of having distinctions in word meanings capable of capturing every conceivable detail that one might wish to convey; and this is in spite of the flexibility of individual word meanings." *Id.* at 49. This is really just a conclusion. Perhaps a better argument, which Ortony mentions, is that humans do not have literal language to describe what thoughts do. *Id.* For example, if one tried to unpack the statement, "the thought slipped my mind," any characteristic of "things slipping" that one tried to predicate of "thoughts" would be limited to metaphorical predicates of "thoughts," such as, "thoughts evading," or "thoughts escaping." This part of Ortony's theory is objectionable. The same thought could be expressed directly in the statement, "the thought was inaccessible to my consciousness."

essary prerequisite for a general rule describing what makes two cases "similar enough." Lawyers use precedent for two purposes. First, they may use it to predict the outcome of a case[38] to assess their chances of winning (for example, in deciding whether to litigate or settle).[39] Second, lawyers also use precedent to persuade judges that a certain rule is a "good" way to decide a particular case.[40] Focusing on the second purpose, if what counts as "similar enough" within a particular legal context depends on a certain outcome being "good," then the value choice of what is "good" would seem to determine the appropriate level of abstraction for comparing two cases. This requires a substantive moral theory about what makes an outcome "good." A substantive moral rule about "goodness" is a necessary prerequisite for a rule of similarity.[41]

For computers to understand "similarity,"

(1) they must understand both the analogue meanings of words and
(2) they must understand moral decisionmaking.

IV. HOW COMPUTERS PROCESS INFORMATION AND THEIR CAPACITY TO REASON ABOUT SIMILARITY THE WAY HUMANS DO

A. How Digital Computers Process Information

Computers are digital information processors. Described at the lowest conceptual level, computers process information as a series of on/off electrical impulses. Communicable information is limited to concepts that can be defined or understood as the sum of discrete entities. For example, if a "glug" is anything that has properties *A*, *B*, and *C*, a computer understands a "glug" by recognizing the discrete symbols *A*, *B*, and *C*. The presence of properties *A*, *B*, and *C* are necessary and sufficient for a thing to be recognized by a computer as a "glug."

Computers only understand digital information. All computer understanding of concepts must be solely the sum of digital pieces of information in its database. Theoretically, a database as rich as all human

38. Moore, *supra* note 1.

39. This use of precedent is beyond the scope of this Note and will not be discussed.

40. Moore, *supra* note 1.

41. *The Semantics of Judging*, *supra* note 16. What values do judges employ to determine whether a particular outcome is "good" and are these values reducible to general rules? Judges use both conservative and reform values such as fairness, notice, efficiency, and separation of powers to decide cases. Judges use values to determine whether to overrule cases. Judges use values to decide whether to interpret the words in statutes broadly or narrowly. Judges also use values to determine which line of reasoning to follow when there are concurring opinions. All these moral decisions figure into judges decisions of whether a particular outcome is "good." Thus, the values are used to decide whether two cases are "similar enough." Further inquiry is necessary to determine whether these values are describable by rules.

perception, cognition, emotion, and experience, despite being comprised of digital information, may function as a continuum. The operative notion here is similar to the way that adding dots between the marks on a clock eventually produces the functional equivalent of a continuous line — a circle. Part V of this Note will illustrate that state of the art programming technology falls short of approaching such a rich database.

B. How the Digital-to-Analogue Problem Limits Computer Understanding of Similarities Between Legal Cases

The way digital computers must process language limits their ability to understand word meanings. Humans, unlike computers, use words to communicate thoughts which can not be reduced to sets of discrete properties. Word meanings are not sets of discrete necessary and sufficient attributes.[42] Humans use digital definitions (necessary and sufficient conditions for application of a term) as a short-hand device for "pointing"[43] or directing the information recipient's attention to the object or concept they wish to communicate about. Digital definitions only supply the "linguistic signposts." The digital definition does not capture the continuum of referential space (or the speaker's conceptualization of what is being referred to). Humans connect the gaps between linguistic signposts by drawing on their memories, perceptions, emotions, and experiences. As a result, a person conceptualizes the continuum of things in the world a word could refer to. Digital computers can not fill the gaps between linguistic signposts and thus cannot understand analogue thought — word meanings.

The objection to using discrete or digital systems to describe continuous processes or analogues applies to the use of natural language techniques such as simile, analogy, and metaphor to communicate about similarities between cases. Because communication deteriorates when factual situations are compared by a digital process, differences in referential space create an important barrier to reasoning about similarities. Two different people may conceptualize the meaning of a word in two different ways. For example, a precedent case may hold that an "offer" existed. Lawyers compare the facts of the precedent case to the facts of a new case to determine if they are similar enough to warrant a decision that an "offer" also exists in the new case. The digital problem arises because the facts that constituted an "offer" in the precedent case may be neither necessary nor sufficient to constitute an "offer" in all future cases. The analogue thought or meaning behind "offer" in the precedent case contains something more than the discrete list of facts in that case. The precedent case describes facts with discrete symbols — the

42. *Id.*
43. *Id.*

words. The same digital words may appy to a different underlying story, a different underlying continuous thought in another case that did not find an "offer." For example, the facts "P, a salesman, asked D if he wanted to buy a car" may have constituted an "offer" in one case. Yet, the existence of the words "P, a salesman, asked D if he wanted to buy a car" in a second case does not necessarily mean that an "offer" occurred. The underlying story that gave rise to the choice of words in the second case may be different.

A digital computer is limited to processing discrete lists of attributes that constitute an "offer." A computer can only distinguish between the facts in the precedent case and the facts in the new case by pre-programming probabilities of the existence of an "offer" given specific discrete attributes. For example, the precedent case may say, in digital form, that conditions one, two, three, and four constitute an "offer." The programmer can add instructions that the existence of the first three conditions produce an eighty-five percent chance that an "offer" exists. Programmers can also instruct the computer to follow additional chains of inquiry. For example, computers can be instructed to ask questions "a," "b," and "c" if condition four is absent. Then, perhaps, three positive answers may increase the probability to ninety percent. A computer with a very large memory, and a very complete database of all the meanings of every word, might provide the same result as a lawyer. If hardware and programming technology progress to this point, such a system may be useful to lawyers despite a lack of complete accuracy. This Note, however, does not explore the practical question of whether computers can *conclude* as lawyers do; its purpose is to explore whether computers can *reason* as lawyers do.

1. *Ortony's Theory Applied to Legal Reasoning*

Ortony's theory of metaphorical language showed that understanding metaphors, similes, and analogies requires an understanding of word meanings as continuums of referential space.[44] A computer's capacity to understand why two legal cases are similar is limited because its understanding of word meanings is limited to identifying discrete, noncontinuous attributes.

Ortony's theory of how humans understand metaphorical language provides a useful framework for discussing how lawyers reason about similarities in legal cases. That is, how lawyers decide first,

(1) what properties are legally relevant; and
(2) what degree of sharing of those properties is needed for similarity in the legal context.

---

44. *See supra*, text accompanying notes 31-39.

Applying this theoretical discussion by comparing two hypothetical legal cases helps to illustrate that if computers are to understand legal similarity, they must be able to

(1) understand word meanings as analogue thoughts rather than as lists of discrete necessary and sufficient conditions; and
(2) understand moral decisionmaking.

Attending to the "vehicle's" "salient" characteristics is analogous to attending to the "precedent case's" holding that certain "facts count for an instance of an accepted legal category." Lawyers do not know which analogue-facts (real world events or circumstances) the court relied on; they only know which fact-words (digital symbols) the court *says* it relied on. Lawyers do not know what extensions of the words the judge had in mind when writing the opinion. That is, they do not know what aspects of the continuous meanings of digital words were truly the "facts" which counted for instances of legal words. Lawyers also do not know what values persuaded the judge to choose the words ultimately used to describe the facts. Even if lawyers accept what the court *says* constitutes an instance of a legal category, the digital-to-analogue problem still hinders computer understanding of similarity.

The digital-to-analogue problem arises in both requirements of a theory of similarity. Taking the subset of salient characteristics that are transferable from the "vehicle" to the "topic" is analogous to asking what facts that constituted an instance of accepted legal categories in the precedent case are transferable or "relevant" to the new case. An understanding of the analogue meanings that each legal word and each fact-word refers to in the precedent case answers this question.

Ortony suggests that "transferability" of characteristics is a question of which ones can be "reasonably applied" to the "topic."[45] He urges that the transferable or relevant subset of characteristics is determined by eliminating those characteristics that create a "tension," somehow contradict, or are conceptually incompatible with what we know of the "topic."[46] This is similar to eliminating the range of objects and concepts that the fact-words refer to in the precedent case which are not "reasonable" extensions of the fact-words in the new case.

Relevancy then becomes a question of reasonable applicability. This analysis presupposes an understanding of the analogue meanings of digital words. A plausible rule of relevancy may be that the accepted legal categories in the precedent case are relevant to the facts of the new case if the analogue meanings (referential space) denoted by fact-words in the precedent case are reasonable extensions of the fact-words in the new case. Computers must understand the analogue meanings of

45. Ortony, *supra* note 25, at 49.
46. *Id.* at 48.

words if they are to answer the question, "What properties are relevant?" This question is the first component of a theory of similarity.

The second component of a theory of similarity requires that the legal context determines the relevant degree of shared attributes necessary for similarity. Computers must understand the *legal context* (why lawyers compare cases) before they can determine if the referential space covered in the precedent case is "similar enough" to the referential space covered in the new case. To account for all the various legal contexts, a general rule would have to take the form of a string of disjunctives. For example, such a rule might suggest that two things are similar if in context *A*, *x* is shared; in context *B*, *y* is shared; in context *C*, *z* is shared, and so forth.

2. *An Illustrative Comparison of Two Legal Hypotheticals*

Comparing two legal hypotheticals helps illustrate how these requirements limit a computer's capacity to reason as lawyers do about similarity:

Hypothetical 1	Hypothetical 2
On a rainy night	On a sunny day
in January	in June
Roger Able	Carrie C.
drove	drove
a truck	a motorcycle
on a highway	on a residential street
at an excessive speed	at an excessive speed
and hit the rear	and hit the rear
of a brown Buick	of a green Dodge
driven by	driven by
Kathy B.	John D.
and injury occurred	and injury occurred
to Kathy B.'s	to John D.'s
neck.	leg.

Is the first hypothetical similar to the second?

Assume the defendant in the first hypothetical was negligent. What are the salient or relevant facts in the case that led the judge or jury to find that Roger was negligent? What facts constituted a "duty," a "breach," "cause-in-fact," "proximate cause," and "damage?" This involves two questions. First, what extensions of the words of the opinion did the judges have in mind; in other words, what aspects of the continuous meanings of digital words are truly the "facts" that constituted a "breach?"[47] Second, what *values* or *undescribed facts* influenced the judge's choice of words to describe the facts in the opinion? For example, assume the speed of Roger's car was an issue. What persuaded the judge to describe Roger's speed as "excessive?" Of all the possible interpretations of the event, what tipped the scale from "normal speed" to "excessive speed?" What factors influenced the judge's or jury's *a priori* conceptualizations of the continuum of things that could constitute "excessive speed"? Maybe the jury perceived Roger as a "long-haired punk with no sense of responsibility" and this influenced their decision that his speed was excessive. Maybe Roger was a driver for a large trucking company and the jury was unsympathetic to a "deep pocket." This example illustrates how undescribed facts, such as "long hair" and "deep pockets" can invoke value judgments. These value judgments influence the concepts that decisionmakers have in mind when they find "excessive speed." This undescribed element is lost in the digital description of the facts.

What facts that constituted "breach" in the first hypothetical are transferable or relevant to the second hypothetical? Legal words such as "breach" are open-ended and defeasible because there are no sets of necessary and/or sufficient conditions that define their referents in all cases. Humans can take what they know about the vehicle (the first hypothetical) and see what is conceptually incompatible with what they know about the topic (the second hypothetical). All that can be inferred from the text is that "excessive speed" counted for a "breach." To determine what facts are conceptually incompatible with the second hypothetical, one needs to determine the conceptualizations that the decisionmakers had in mind when they used the words "excessive speed." That is, one must identify the continuum of referential space covered by the digital or discrete symbol "excessive speed." One then needs to determine if this continuum can be a reasonable extension of the words "excessive speed" in the second hypothetical.

How does one determine whether the hypotheticals are "similar enough" within the legal context? What values tell a judge that the ruling against Roger is a "good" reason to decide against Carrie? Rules about the values that inform judicial decisions are necessary.

Computers must understand natural language, the analogue meanings of words, and they must understand moral decisionmaking to understand "similarity."

47. This view of precedent differs from both the classical and the realist views. While classical theorists look to what judges *say*, and realists look to what judges *do*, this view looks to the analogue meanings judges have in mind when they use digital symbols (words) to communicate their decisions.

V. THE EXTENT TO WHICH PROGRAMMERS HAVE BEEN ABLE TO ACHIEVE MACHINE UNDERSTANDING OF NATURAL LANGUAGE

The purpose of Part V is to determine to what extent digital computer programming techniques could effect machine *understanding* of natural language, the analogue meanings of the digital symbols. This addresses the issues of current computer technology and the future direction of Artificial Intelligence (AI). AI is defined as " 'the attempt to (1) understand the nature of intelligence and (2) produce new classes of intelligence machines through programming computers to perform tasks which require reasoning and perception.' "[48] Programming moral decisionmaking will not be discussed because computer understanding of natural language processing is a prerequisite to computer understanding of any reasoning process.

The difference between a true AI system and an expert system is important. A true AI system involves natural language understanding, causal and inferential reasoning, and pattern recognition. An AI system should be able to process any problem, given an appropriate data base, because its understanding of language delimits its causal and inferential reasoning capabilities.

In contrast, an expert system incorporates the opinions and methods of experts in a substantive field to solve particular problems. Although expert systems provide causal and inferential reasoning to some extent, as part of the experts' methods, these systems do not attempt to understand natural language. The success of an expert system depends upon accurate encoding of human experts' rules for problem solving within well defined and limited domains.

Thus, it appears that an expert system would not be sufficient for understanding legal cases the way humans do. Legal cases, arising from real world events, encompass a universe of problems and discourse far too broad to be encoded. This is, however, a practical, not a theoretical limit. The problem descriptions and their judicial solutions are not well defined in the same way that other fields are. This is not to say that decisionmaking rules do not exist; legal theorists have just not yet described them completely. The infinite variety of problems and semantic descriptions makes natural language understanding a prerequisite to building a program to understand and predict legal decisions. That is the domain of true AI.

One of the first attempts at simulating human thinking with com-

48. Nycum, *Artificial Intelligence and Certain Resulting Legal Issues*, USC SIXTH ANNUAL COMPUTER LAW INSTITUTE (1985), at 1 (quoting Waltz, *Artificial Intelligence: An Assessment of the State-of-the-Art and Recommendation for Future Directions*, THE AI MAG., Fall 1983, at 55).

puters was the "General Problem Solver" (GPS).[49] GPS applied laws of reasoning or "heuristics" that instructed the computer to solve a restricted domain of puzzles.[50] GPS was unable to handle complex "real-world" problems, but the program's limited success made heuristic programming devices the focus of AI research.[51]

Two expert systems using heuristic devices were developed in the 1970's: DENDRAL, a system for identifying organic molecules, and MYCIN, a system for diagnosing bacterial infections.[52] Programmers develop expert systems by consulting with human experts in a substantive field to extract the rules of reason or heuristics from which human experts *deduce* their decisions.[53] Expert systems work by logically applying these principles to problems in a particular substantive field.[54] "Researchers found that by removing the medical information, or the knowledge base, from Mycin, what was left was the generalized logic of the system. This system, dubbed Emycin for Essential Mycin, could then be connected to a database containing expert heuristics from other fields."[55]

Developing an AI legal system capable of understanding similarities between cases requires rules of reasoning about similarities. Part II of this Note concluded that a rule about similarity is dependent upon rules describing how humans understand the analogue meanings of words and rules describing how humans make moral decisions. Successful development of AI systems for determining similarity between cases is limited to the state-of-the-art and future direction of natural language processing.

AI researchers have studied natural language processing for over twenty years.[56] "[P]rogramming computers to understand ordinary natural language is one of the most difficult challenges now facing the discipline of artificial intelligence."[57] A computer that will understand natural language must have a database relevant to the universe of discourse.[58] In law, the universe of discourse covers all human and machine actions, all possible human states of mind, all conceivable at-

---

49. Nycum, *supra* note 48, at 2.

50. *Id.* at 3.

51. *Id.*

52. *Id.* at 4 (citing E. Feigenbaum, *On Generality and Problem Solving*, in MACHINE INTELLIGENCE 6 (B. Meltzer & D. Michie eds. 1971)); *see also* E. SHORTLIFFE, COMPUTER BASED MEDICAL CONSULTATIONS: MYCIN (1976).

53. Nycum, *supra* note 48, at 5-6.

54. *Id.* at 6.

55. *Id.* at 5.

56. *Id.* at 8 (citing H. TENNANT, NATURAL LANGUAGE PROCESSING, AN INTRODUCTION TO AN EMERGING TECHNOLOGY (1981)).

57. Waltz, *Artificial Intelligence*, 247 SCI. AM. 118, 130 (1982).

58. M. BODEN, ARTIFICIAL INTELLIGENCE AND NATURAL MAN 114 (1977).

tendant circumstances, and all the forces of nature. Understanding natural language requires implicit and explicit inferences that draw on knowledge of the world.[59] It follows that a computer program that could understand discourse about the law must understand the referential space, analogue meanings, that natural language denotes at least as well as humans do.

Examination of current natural language processing technology demonstrates that there are technical and possibly theoretical limits to achieve complete computer understanding of natural language. Theoretical limits *may* exist if there are no rules about values (for example, "goodness" of outcomes) that inform decisions about what level of detail is relevant in particular legal contexts.

A difference exists between a computer program that *appears* to understand natural language and a program that *actually* understands the analogue meanings of digital words. Weizenbaum's ELIZA, written in 1966, "bypasses any real linguistic processing and instead relies on a clever system of stored, fixed patterns of response that give an imitation of language understanding many people find convincing."[60] ELIZA simulates a psychotherapist interacting with patients. ELIZA's strategy is to prompt the patient to discuss the emotions associated with the patient's remarks.[61] ELIZA scans a patient's sentence for key words, such as "father," "family," and "I," and uses pre-programmed rules associated with those words to formulate its response.[62] These techniques enable ELIZA to persuade patients to think that the computer really understands what they are saying. ELIZA, however, does not understand anything. ELIZA simply *responds* to natural language.[63]

ELIZA's success results from the well defined nature of the programmed task. The nature of nondirective psychotherapy allows for the deception. Patients expect the therapist to make noncommittal statements such as "tell me more about boats"; they do not expect the therapist to launch into a substantive discussion about sails and rud-

---

59. *Id.*

60. Waltz, *supra* note 57, at 130.

61. M. BODEN, *supra* note 58, at 114. ELIZA "avoids making statements or expressing attitudes on her own behalf; she concentrates instead of asking questions about topics previously introduced by the patient or guides the discussion onto subjects like "father" and "family" that will very likely prove to be emotionally significant." *Id.*

62. Often, more than one rule applies to each key word. To choose a rule, ELIZA matches the patterns of the input with the patterns of pre-programmed "context" rules. For example, ELIZA's response to an input sentence containing the word "you" will depend on which pattern or "context" it appeared in: "you are . . .," "you . . . me," or "you are like . . . ." ELIZA also has a pre-programmed ranking of key words in case more than one key word appears in an input sentence.

63. M. BODEN, *supra* note 58, at 111.

ders.[64] ELIZA focuses on recognizing and commenting on a very limited number of key words in a limited number of contexts.[65] This restriction on the possible domain and context of word meanings is necessary for all current natural language processors.[66] Current systems are inadequate to perform the task of unrestricted natural language processing; they have only been successful while operating in well structured and defined universes.[67]

One of the most comprehensive attempts at natural language processing is Conceptual Dependency Theory (CD) developed by Roger Schank of Yale University.[68] Schank intended CD to be a sufficiently explicit description of natural language processing so that it could be programmed on a computer.[69] Schank attempts to describe natural language in terms of the concepts underlying the sentences.[70]

The programs are based on what Schank calls the "primitive" concepts describing human action.[71] Schank views language understanding as a "top down" process. In this process, words serve as cues for retrieving expectations from memory and as evidence for or against estab-

---

64. *Id.* at 108.

65. *See id.* at 109.

66. Lenat, *Computer Software for Intelligent Systems*, 251 SCI. AM. 204 (1984).

67. *Id.*

68. *See generally* R. SCHANK, CONCEPTUAL INFORMATION PROCESSING (1st ed. 1975, 2nd printing 1984).

69. *Id.* at 3.

70. *Id.* at 7.

71. Schank requires that these concepts be represented without using the words of the language they are representing, because he believes thoughts are independent of language. *Id.* at 8.

These concepts take the form of basic primitives. There are six roles that concepts could play. Something can be an "ACTOR," "ACT," "OBJECT," "RECIPIENT," "DIRECTION," (the location that an ACT is directed toward)," or "STATE." *Id.* at 37. There are rules for what could be within any role, and for how these conceptual categories can combine. For example, a conceptual ACTION is defined as something that can be done to an OBJECT by an ACTOR. There are two categories of ACTIONS: PHYSICAL ACTIONS and MENTAL ACTIONS. *Id.* at 37-38.

There are five ACTS that describe the PHYSICAL ACTIONS that people can perform: "PROPEL," "MOVE," "INGEST," "GRASP," and "EXPEL." *Id.* at 41-42. The real meaning of each ACT is a set of pre-programmed inferences that are possibly true when the ACT is present; they do not have their ordinary English meanings. *Id.* For example "INGEST" means:

> 'Take something to the inside of an animate object.' The object of INGEST must be smaller than the particular body opening of the actor that it is entering. If the object is bigger then it can be inferred that it was divided up somehow into smaller bits previous to the INGESTing. The Directive case for INGEST is always to a body opening and from the original position of the object. *Id.*

Each ACT has similar deep meaning that accompanies it. These concepts focus on the search for semantic similarities between acts; these may be important for semantic groupings, but also ignore subtle differences in word meanings.

lished expectations.[72] Schank maintains that the representations of underlying concepts must represent all the information that is implicit in a language.[73] Theoretically, Schank's system should be able to represent continuous concepts as well as unnameable attributes. The system should be able to unpack metaphors.[74]

To understand analogue meaning effectively, CD must account for the inferences or expectations triggered by cues in a sentence. Schank claims that a natural language processor or parsor can be programmed to predict the most likely interpretation of a given sentence. The parsor will proceed to other interpretations only if the original inference is proven wrong by additional data.[75] Schank believes that the parsor should not look for all the ambiguities in a sentence because doing so would inhibit understanding.[76] Many legal issues involve the ambiguity of word meanings. Legal rules, both common law and statutory, are conditional imperatives. Legal questions center on the application of ambiguous words in the rule to the facts of a case. It seems then, that CD may not be an adequate language representation system for use in the legal context.

Schank also requires the parsor to have world knowledge, which enables it to draw inferences and make predictions.[77] Schank concedes that a complete parsor is impossible because of the infinite number of concepts in the real world.[78] Thus, CD, like other natural language processing systems, may only be successful in a very restricted domain.[79] This also poses problems for building a legal AI system.

The complexity of the real world, an unlimited domain, makes each legal decision unique. The unique events which constitute legal cases inhibit concept formation of "sufficient" and "necessary" facts. The discrete or digital limitations of words forces some part of the known events to be classified by words that either incompletely or inaccurately describe the events. Since the knowledge required to draw legal infer-

---

72. This is similar to the view that our perceptions are continuous and our language forces us to describe experiences in discrete, digital symbols. The crucial question is whether CD's representations of actions and concepts captures the whole continuum of thoughts and perceptions encompassed by natural language.

73. R. SCHANK, *supra* note 68, at 9.

74. Schank's system of conceptual representation would have abstract units that unambiguously represent the speaker's meaning. *Id.* at 15. A sentence that had two different meanings in two different contexts would have two different representations. *Id.* at 17. Two different sentences having the same meaning in the same context would have the same representation because the underlying concepts are the same. *Id.*

75. R. SCHANK, *supra* note 68, at 19.

76. *Id.*

77. *Id.* at 20.

78. *Id.* at 21.

79. *Id.*

ences encompasses nearly the entire universe of discourse, such a programming task would be impractical, if not impossible.

Schank's admission that CD can only construct an adequate parsor for a limited universe may not be as damaging to the AI cause as it seems. Lawyers and judges often make decisions about the meanings of words based on incomplete data. The human brain is finite and does not have a "database" that encompasses the infinity of world experiences. Humans, however, do have the capacity to select the level of generality appropriate for a given task. Programs successful at finding the relevant level of generality at which to draw inferences or ask for more information have a precisely structured syntax, which limits their universe.[80] Humans make these decisions in a much larger universe of discourse. MARGIE, a program that uses CD, is unable to determine the level of generality necessary to interpret communication.[81] This is fatal to the CD system's understanding of metaphor. A language representation system that can reason about similarities must have a context sensitivity, which enables it to choose the appropriate level of generality at which to make the comparison.

"Frames" are another representation technique used in natural language processing. In 1974, Marvin Minsky suggested that language, and all thinking, might depend heavily on processes driven by expected structures of knowledge called "frames."[82] Frames are categories of knowledge. They are essentially stereotypes, and are similar to a primitive concept in CD.[83] Each frame is assigned a set of probable characteristics for each variable such as "actor," "object," and "action." Stereotyping provides an intuitively plausible model of how people fill in implicit information about a situation. For example, a frame representing "going to restaurants" would instruct the computer about the

---

80. *See* M. BODEN, *supra* note 58, at 159.

81. "The computational work that has been done in connection with Schank's theory has concentrated on assigning conceptual dependencies to sentences and drawing limited inferences therefrom — so limited that the problem of selection in accessing the data in memory has been postponed rather than solved." *Id.* at 160.

For example, MARGIE, a program that uses CD, cannot decide what level and area of detail to access in interpreting a given statement. *Id.* at 159. To determine if a "man diving into icy water" is like a "fearless warrior" MARGIE must determine the salient features of being a warrior such as bravery and strength, as opposed to non-distinctive characteristics, such as "has a nose" and "breathes". Then MARGIE must decide which of the distinctive characteristics of "fearless warriors" could be predicated of "men diving in icy water." If the program cannot determine the level of generality at which to interpret "he dived into the icy water" and "fearless warrior," then it cannot understand the metaphorical meaning of "dived into the icy water like a fearless warrior."

82. Waltz, *supra* note 57, at 132.

83. *Id.* at 132.

background expectations most people have about going to restaurants.[84] A language processing program using such a frame should be able to infer from the following passage that John also ate the spaghetti: "John took the bus from New Haven to New York. On the way his pocket was picked. He went to a restaurant and ordered spaghetti. John could not pay the bill so he washed dishes."[85]

Frames may facilitate machine understanding of metaphors by enabling the computer to detect similarities and differences between two concepts. Frames facilitate understanding the novel frame or the "topic" by relating what is known in the familiar frame or "vehicle".[86] Richer or more detailed representations allow the computer to extend and explore the analogy more completely.

The fundamental analogue meaning/digital description problem still remains. The computer must understand the analogue meaning behind the words that describe the variables. In addition, although stereotypes may facilitate drawing the proper inferences for novel frames, like CD representations, they face the difficulty of selecting the proper level of generality at which to make the comparison. The system must be able to understand *why* two variables are alike. Current frame representation techniques lack a substantive moral theory to inform decisions about what is important for the purposes of the comparison.

Another proposal to understand metaphor and analogy is Waltz's "Event Shape Diagrams." Waltz claims that his system "can support judgments of the degree of plausibility of various interpretations of a sentence's meaning, and may be useful in figuring out the meaning of certain kinds of metaphors."[87] Waltz attempts to represent the nuances in meaning rather than capture the similarity among objects and verb actions.[88]

The basic diagrams represent actions or verb-based metaphors on two dimensions: over time, and through values reflecting degrees of

---

84. *Id.*

85. *Id.*

86. M. BODEN, *supra* note 58, at 326.

87. Waltz, *Event Shape Diagrams*, in AMERICAN ASSOCIATION FOR ARTIFICIAL INTELLIGENCE: PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE, 84. (Carnegie-Mellon, University of Pittsburg, Aug. 18-20, 1982) [hereinafter *Event Shape Diagrams*].

88. This is what Waltz sees as the problem and misguided focus of Conceptual Dependency theory. For example, Schank's primitive act "INGEST" represents several verbs, including "eat, eat up, overeat, nibble, gulp, wolf, drink, swig, swallow, inject, and smoke." *Id.* at 84-85. These meanings are lumped together because they give rise to similar inferences. However, Waltz claims that substituting INGEST for all the verbs above makes the system incapable of making some predictions or inferences that humans would normally make. *Id.* Particular verbs describe certain nuances in meaning. To "nibble" some food or to "wolf" some food invokes different analogue meanings. These inferences are

each action.[89] Time lines and values scales purport to capture many of the nuances lost in lumping similar verbs into a single primitive concept. For example, on the time continuum the action "INGEST" stops where "desire to eat" goes to zero.[90] The value scale for "desire to eat" accounts for coercion, habit, and other factors as well as hunger.[91]

Event Shape Diagrams focus on verbs, and reflect two general principles. First, both verbs and nouns have inherent selection restrictions; and second, nouns are less likely to be metaphorical than verbs.[92] Although Waltz explains the first principle by reference to semantic hierarchies,[93] he merely announces the second principle without justification and states its corollary: "If a verb and object do not match each others' selection restrictions, the object should be taken as referring literally, and the verb as referring metaphorically."[94]

To understand verb-based metaphors this system uses pre-programmed lists of predicates that each noun normally prefers.[95] When a verb is used metaphorically, the system compares the diagram of that verb with the diagram of the preferred predicate.[96] Next, the system identi-

---

lost when the subtle differences in the meaning of verbs are ignored in the representation of INGEST. *Id.*

Waltz claims that Schank concentrated on developing larger memory structures that contain many CD structures, rather than rectifying the shortcomings in representing subtle meanings. *Id.* Schank may be focusing on a larger number of concepts and larger parsors because he believes, like others, that inferences arise from stored experiences. The greater the memory base, the greater the ability to draw inferences from known to novel situations. *See* Lenat, *supra* note 66, at 211. But though larger memory structures help with inference drawing, a finer-grained representation system is necessary for capturing nuances in meaning.

89. *Event Shape Diagrams*, *supra* note 87, at 85-86.

90. *Id.* at 85.

91. *Id.*

92. DeJong and Waltz, *Understanding Novel Language*, 9 INT'L J. COMPUTERS AND MATHEM. 131, 141 (1983).

93. Semantic hierarchies are concepts associated and to some degree linked with a trigger word. For example, "to eat" requires the concept of food, which also triggers limited acts to be done on food such as growing, cooking, and preparing. Semantic hierarchies are good models for suggesting how humans categorize and understand cues, but do not function as real limits on the inferences to be made. The very nature of analogy and metaphor is to make comparisons between actions and objects that are *not* normally associated with those topics.

94. DeJong, *supra* note 92, at 141-42.

95. Waltz calls this theory verb-based because he gives the verb priority in assigning meaning. In analyzing a metaphor, the verb is given literal meaning, and the object takes the metaphorical meaning. Thus, Waltz believes that objects have certain "preferred" actions; the metaphorical comparison is between the literal meaning of the chosen verb and the unnamed "preferred" verb that usually accompanies the object in a similar context. Of course, Waltz pre-programs the object preferences into the system.

96. DeJong, *supra* note 92, at 142-44.

fies the parts of the verb diagram that differ from the preferred predicate diagram.[97] It then transfers the identified parts to the preferred predicate diagram.[98] This procedure aims to illuminate the meaning of the metaphorical verb.

Event Shape Diagrams do offer the potential of reflecting more dimensions of information than other systems (like CD's primitives or frames). However, Waltz seems to have the metaphorical comparison backwards. In understanding the metaphor, the known characteristics of the vehicle are transferred to the topic, thereby fleshing out the metaphorical meaning of the topic. Waltz suggests curiously that one should compare the topic to the vehicle and transfer the dissimilar parts to the known vehicle in order to increase understanding of the unknown topic![99]

Finally, the system does not answer the analogue meaning/discrete description problem raised in Part II of this Note. Asserting that the diagram values scale accounts for all the inferences that accompany a concept only presents the issue; it does not answer it. Using a single value to label a continuous referent is no different than using a digital word to label a continuous thought.

Other systems are based upon theories of human learning and concept formation. Patrick Winston developed a program that learns by example. The program "learns" the concept of an "arch" by building a structural description of the arch in terms of objects and their relations.[100] This representation is gradually refined as the computer experiences examples and counterexamples of arches.[101] Winston's program decides the importance or conceptual salience of the features

97. *Id.*

98. *Id.*

99. Waltz illustrates the two principles by showing how the sentence "John ate up the compliments" would be interpreted:

> Using principle (1) above, one notes first that "ate up" prefers *food* of some kind as a semantic object, that "compliments" is not a food, and itself prefers an MTRANS-type verb, in particular either "tell" or "hear". Next, using principle (2), one can judge that "compliments" refers literally, and so either "tell" or "hear" is probably the true basic verb. *Id.* at 142.

The next step is to "match up" the diagrams for "tell" and "hear" and with the diagram for "eat up." *Id.* Waltz seems to be using the likely predicates of compliments as the common vehicle for illuminating the metaphorical meaning of the "topic" verb "eat up." To do this, he transfers (by superimposing or modifying) the parts of the diagrams for "eat up" that do not have matching diagrams in "hear" or "tell" to the diagrams for "hear" and "tell". If Waltz uses the preferred predicate for "compliments" as the "vehicle" for interpreting "eat up," then the diagrams in "hear" and "tell" that are not present in "eat up" should be transferred to the "eat up" diagram. After all, "eat up" is the metaphorical language.

100. M. BODEN, *supra* note 58, at 253.

101. This is analogous to building synthetic *a priori* conceptualizations over time.

in the examples being compared. This decisionmaking process makes Winston's program interesting as a representation of intelligence, as opposed to a mere "arch recognizer."[102]

How does the computer recognize the salient features? All observed characteristics are recorded in the initial description built by the program. But how does the computer know whether a particular feature is essential or merely incidental to the concept? "The program (or a person in like case) has no way of knowing, until an example of an arch is encountered that does not have this feature."[103] Similarly, if all the examples a computer encounters share a certain feature, should the computer assume that it is necessary to the concept, or should it assume the feature is incidental but universally present? A counterexample answers this question.[104] Further, if the machine does not observe any example that possesses a certain feature, should the computer conclude that the feature is forbidden or should it conclude that it is a coincidence that the feature is not present? The computer can encode it as a contingent feature[105] pending more data. The concept is refined and evolves when the computer compares each successive example or counterexample of the concept with the current version of the concept.[106] The order in which new examples and counterexamples are introduced changes the nature of a concept at any point in time. For example, one of the above problems may arise at time one; a clarifying counterexample may not be introduced until time four. Thus, at time two, the computer has a different concept than if the examples were received in reverse order. In addition, introducing more than one difference at a time poses grave difficulties for the program, as it does for people.

Being limited to serial concept building and understanding is not problematic because humans share similar limitations. It is unreasonable to expect a computer to do something humans cannot do (even though humans do expect computers to peform more quickly the same tasks that humans do). The primary advantage of Winston's program is that it specifies the nature and closeness of an analogy;[107] the program does not find similarity in an all-or-nothing fashion. Notwithstanding these concessions, Winston's program has three problems.

First, the program cannot determine the level of generality of particular features upon which to focus. The program is unable to distin-

102. M. BODEN, *supra* note 58, at 253.

103. *Id.*

104. *Id.* at 258.

105. *Id.* at 259.

106. *Id.* at 254-59.

107. *Id.* at 267.

guish between salient features for the purposes of the comparison because the program does not understand context.[108]

Second, a general concept learner must be sensitive to both abstract semantic relations, like contrast, and to provisional structuring, where particular features become more salient with the presence or absence of certain other features. Because Winston's concept learner has no way of knowing which features are salient in various contexts, it lacks this sensitivity. These two problems may be only technical. The ability to program context sensitivity in a legal computer system depends on whether legal theorists can formulate rules about the values that determine the appropriate level of generality in various legal contexts.

Third, similar to other programs, the program is only successful when operating in a very restricted universe such as arches. The uniqueness of real world cases in a legal system makes this limitation a nearly intractable barrier to concept formation. Even though the human brain is finite, it forms concepts within the context of a much larger universe of discourse.

Wilks developed "Preference Semantics," a program that aims to draw the necessary inferences about world events in order to resolve pronoun references and ambiguities in word senses.[109] "Word senses" are the *intended* word meanings, those aspects of the continuous meaning the author intended to denote. For this Note, it suffices to describe only the part of the program that attempts to determine word senses.

Two operative characteristics distinguish this program from others. The first is the inferential use of nonanalytic or nondeductive information.[110] The second is the ability to *prefer* one representation of meaning over another.[111] Words always have alternative competing representations. The program chooses one inferential chain over another to settle ambiguity. The program prefers representations or inferential chains that increase the semantic density or richness of a previously derived structure.[112] This concept is identical to having the most "matches" in a frame,[113] or using the most similar Event Shape Diagram to compare the vehicle and topic. An example will illustrate the preference rules:

> [I]n understanding "My ideas followed hers closely", we want to accept

108. Winston's program does not have any expectations about what particular features are salient. The program, however, prefers the least detailed differences, and places the detailed differences on an "alternate list" for comparison if necessary.

109. Wilks, *A Preferential, Pattern-Seeking, Semantics for Natural Language Inference* 6 ARTIFICIAL INTELLIGENCE 53 (1975).

110. *Id.* at 54.

111. *Id.*

112. *Id.* at 56.

113. *Id.* at 69.

> the ideas as the apparent agent, even though our information about the concept of following is that it normally prefers an animate agent if one can be found. Only in that way can the animate sense of fly be chosen correctly as the agent in "The fly followed the ladybird into the web". The point is to prefer the normal, but to *accept* the unusual.[114]

Preference rules allow the program to derive richer representations. Common sense inference rules operate upon these representations to provide deeper understanding and representation.[115] Preferential Semantics uses context based rules about the most likely meaning to eliminate ambiguities. The program prefers the shortest inference chain, so as to never introduce more information than is necessary.[116] The general principle is one of "always being prepared to complexify, or deepen, a representation, but never doing so unless the problem cannot be solved at a more superficial level."[117] This program is distinguished from the programs of Schank and others, which rely entirely on deductive chains. The deductive chains allow "preferences," but only those preferences which fit the logical or deductive pre-defined structure. Wilks' program has a logical rule, the shortest inference chain, but will allow for illogical results, because the shortest chain is not necessarily the most logical metaphorical meaning. This approach seems necessary, since people are often just plain illogical. Wilks attempts to allow illogical results by incorporating a psychological assumption called the "laziness hypothesis." Wilks hypothesizes that humans think through the chain no further than is necessary to understand the analogy.[118]

This program is promising because it determines which ambiguities are important in the particular context. Future theories about "good" outcomes (the reason we look to precedent) may make this program a good candidate for reasoning about legal similarity. Although this program uses rules that arrive at a "solution" at the appropriate level of meaning, the program still has no way to flag words that may be interpreted in significantly different ways at a later time. This would be fatal to a system purporting to understand legal analogy. Without a method to determine which word ambiguities are important in one case but not another, the program would merely reach an explainable answer — not necessarily an appropriate or correct one. Wilks points out this defect in his own system:

> [T]here is the question of how we get outside a pragmatics of local inference, like this one, so as to take account of important facts in a dis-

114. *Id.* at 56.
115. *Id.* at 67.
116. *Id.*
117. *Id.*
118. *Id.*

course that change all standard interpretations, in the way that a single fact in a detective story can do. No advance will be made there, I am sure, until we have some idea what it is to select out certain salient facts as potential sources of future reinterpretation.[119]

VI. CONCLUSION

Computer *understanding* of natural language requires the digital machine's ability to understand the meanings behind the physical symbols, words. This requires an understanding of analogue thoughts because humans understand word meanings as continuums of referential space, rather than as discrete lists of necessary and sufficient conditions. That is, human conceptualizations of word meanings develop continuously as humans experience examples and counterexamples of things in the world that "count for" instances of particular words. Because these conceptualizations are not reducible to discrete lists of conditions that define word applications, word meanings cannot be fully understood as the sum of the series of yes and no questions.

Understanding similarity requires a human understanding of word meanings, because humans understand metaphor by reconstructing the underlying analogue meanings behind the words describing the situations compared. It is possible that humans understand metaphor by filling in only enough detail to adequately reconstruct the analogue meanings. If this is true, then computers may not need databases as large as the universe of discourse to understand word meanings adequately. The goal of AI is to reproduce *human* reasoning with a computer; it is not to create an omniscient being. Nevertheless, state-of-the-art programming technology falls far short of the level of understanding necessary to reconstruct human reasoning adequately. Programs have been successful only in very limited, well defined domains. Technological advances may remedy this shortcoming.

Computers also need to understand moral decisionmaking if they are to understand similarity. Determining the relevant level of generality at which to compare two entities depends on the context of the comparison. It is necessary to know *why* one is comparing two entities before determining whether they are "similar enough" in the particular context. This requires a substantive moral theory because *values* inform choices about the relevant level of generality.

Before machines can understand moral decisionmaking in the legal context, legal theorists will have to reduce this reasoning process to a series of rules that describe what values are important in the various legal contexts. Whether this is a technical or a theoretical limit for

119. *Id.* at 72.

computer understanding of similarity depends upon the future research of legal theorists.

Celeste Tito