Towards a Computational Model of Metaphor in Common Sense Reasoning

Jaime G. Carbonell
Carnegie-Mellon University
Pittsburgh, PA 15213

1. Introduction

The theory that metaphor dominates large aspects of human thinking, as well playing a significant role in linguistic communication, has been argued with considerable force [10, 8, 3, 1]. However, the validity of such a theory is a matter of continuing debate that appears neither to dissuade its proponents nor convince its detractors. Being among the proponents, I propose to develop a computationally effective, common sense reasoning system based on underlying metaphors. I claim that if such a system exhibits cognitively plausible common sense reasoning capabilities, it will demonstrate the utility of metaphorical reasoning. Moreover, if the model can account for observed instances of naive human reasoning better than existing inference systems, it will provide convincing evidence in favor of the metaphorical reasoning theory. This brief paper investigates aspects of the metaphorical reasoning phenomenon and describes the initial steps towards developing a computational model.

2. Experiential Reasoning vs Formal Systems

Humans learn from experience to a degree that no formal system, AI model, or philosophical theory can match. The statement that the human mind is (or contains) the sum total of its experiences is in itself rather vacuous. A more precise formulation of experience-based reasoning may be structured in terms of coordinated answers to the following questions: *How* are experiences brought to bear in understanding new situations? *How* is long term memory modified and indexed? *How* are inference patterns acquired in a particular domain and adapted to apply in novel situations? *How* does a person "see the light" when a previously incomprehensible problem is viewed from a new perspective? *How* are the vast majority of irrelevant or inappropriate experiences and inference patterns filtered out in the understanding process? Answering all these "how" questions requires a ***process model*** capable of organizing large amount of knowledge and mapping relevant aspects of past experience to new situations. Some meaningful starts have been made towards large-scale episodic-based memory organization [14, 15, 12, 9] and towards episodic-based analogical reasoning [5, 4, 2]. Bearing these questions in mind, I turn towards the issue of common sense reasoning in knowledge-rich mundane domains.

My central claim is that reasoning in mundane, recurrent situations is qualitatively different from reasoning in more abstract and experientially unique situations (such as some mathematical or puzzle-solving domains). The former consists of recalling appropriate past experiences and inference patterns, whereas the latter requires knowledge-poor search processes more typical of past and present AI problem solving systems. Since computer programs perform much better in simple, elegant, abstract domains than in "scruffy" experience-rich human domains, it is evident that a fundamental reasoning mechanism is lacking from the AI repertoire. The issue is not merely that AI systems lack experience in mundane human scenarios -- they would be unable to benefit from such experience if it were encoded in their knowledge base. I postulate that the missing reasoning method is one of metaphor-based transfer of proven inference patterns and experiential knowledge across domains. This is not to say that humans are largely incapable of more formal reasoning, but rather that such reasoning is seldom necessary and when applied requires a more concerted cognitive effort than mundane metaphorical inference.

3. Towards Metaphorical Reasoning: The Balance Metaphor

Consider a prevalent metaphor: reasoning about imponderable or abstract entities as though they were objects with a measurable weight. One of several reasoning patterns based on this simple metaphor is the *balance principle*. The physical analog of this reasoning pattern is a prototypical scale with two balanced plates. Large numbers of metaphors appeal to this simple device coupled with the processes of bringing the system into (and out of) equilibrium. First, consider some examples of the basic metaphor, in which *the relevant aspect of an abstract concept maps onto the weight*[1] *of an unspecified physical object.*

> Arms control is a ***weighty*** issue.
>
> The worries of a nation ***weigh heavily*** upon his shoulders.
>
> The Argentine air force launched a ***massive*** attack on the British fleet. One frigate was ***heavily*** damaged, but only ***light*** casualties were suffered by British sailors. The Argentines payed a ***heavy*** toll in downed aircraft.
>
> Not being in the mood for ***heavy*** drama, John went to a ***light*** comedy, which turned out to be a piece of meaningless ***fluff.***
>
> Pendergast was a real ***heavyweight*** in the 1920s Saint Louis political scene.
>
> The crime ***weighed heavily*** upon his conscience.
>
> The ***weight*** of the evidence was overwhelming.

Weight clearly represents different things in the various metaphors: the severity of a nation's problems, the number of attacking aircraft, the extent of physical damage, the emotional affect on audiences of theatrical productions, the amount of political muscle (to use another metaphor), the reaction to violated moral principles, and the degree to which evidence is found to be convincing. In general, more is heavier; less is lighter. One may argue that since language is heavily endowed with words that describe weight, mass and other physical attributes (such as hight and orientation [10]), one borrows such words when discussing more abstract entities [13] -- for lack of alternate vocabulary. Whereas this argument is widely accepted, it falls far short of the conjecture I wish to make.

> **Conjecture:** *Physical metaphors directly mirror the underlying inference processes. Patterns of inference valid for physical attributes are mapped invariant and reinstantiated in the target domain of the metaphor.*

In order to illustrate the validity of this conjecture consider a common inference pattern based on the weight of physical

---

[1] Mass is virtually synonymous with weight in naive reasoning.

objects: The inference pattern is the *balance principle* mentioned earlier as applied to a scale with two plates. The scale can be in balance or tipped towards either side, as a function of the relative weights of objects placed in the respective plates. Inference consists of placing objects in the scale and predicting the resultant situation -- no claim is made as to whether this process occurs in a propositional framework or as visual imagery, although I favor the former. How could such a simple inference pattern be useful? How could it apply to complex, non-physical domains? Consider the following examples of metaphorical communication based on this inference pattern:

> The jury found the ***weight*** of the evidence favoring the defendant. His impeccable record ***weighed heavily*** in his favor, whereas the prosecution witness, being a confessed con-man, carried ***little weight*** with the jury. ***On balance*** the state failed to ***amass*** sufficient evidence for a ***solid*** case.
>
> The SS-20 missile ***tips the balance*** of power in favor of the Soviets.
>
> Both conservative and liberal arguments appeared to ***carry equal weight*** with the president, and his decision ***hung on the balance***. However, his long-standing opposition to abortion ***tipped the scale*** in favor of the conservatives.
>
> The Steelers were the ***heavy*** pre-game favorites, but the Browns started ***piling up*** points and accumulated a ***massive*** half-time lead. In spite of a late rally, the steelers did not score ***heavily*** enough to pull the game out.
>
> The job applicant's shyness ***weighed*** against her, but her excellent recommendations ***tipped the scales*** in her favor.

In each example above the same basic underlying inference pattern recurs, whether representing the outcome of a trial, statements of relative military power, decision-making processes, or the outcome of a sporting event. The inference pattern itself is quite simple: it takes as input signed quantities -- whose magnitudes are analogous to their stated "weight" and whose signs depend on which side of a binary issue those weights correspond -- and selects the side with the maximal weight, computing some qualitative estimate of how far out of balance the system is. Moreover, the inference pattern also serves to infer the rough weight of one side if the weight of the other side and the resultant balance state are known. (E.g., If Georgia won the football game scoring only 17 points, Alabama's scoring must have been ***really light***)

The central issue in my discussion is that this very simple inference pattern based on a physical metaphor accounts for very large numbers of inferences in mundane human situations. Given the existence of such a simple and widely applicable pattern, why should one suppose that more complicated inference methods explain human reasoning more accurately? It is my belief that there exist a moderate number of general inference patterns such as the present one, which together span most mundane human situations. Moreover, the few other patterns I have found thus far are also rooted on simple physical principles or other directly experienced phenomena. However, since the current study is only in its initial stages, the hypothesis that metaphorical inference predominates human cognition retains the status of a conjecture, pending additional investigation. I would say that the weight of the evidence is as yet insufficient to tip the academic scales.

4. Requirements on a computational model

Metaphorically-based general patterns of inference do not appear confined to naive reasoning in mundane situations. Gentner [7] and Johnson [8] have argued the significant role that metaphor plays in formulating scientific theories. In our preliminary investigations, Larkin and I [11] have isolated general inference patterns in scientific reasoning that transcend the traditional boundaries of a science. For instance, the notion of equilibrium (of forces on a rigid object, or of ion transfer in aqueous solutions, etc.) is, in essence, a more precise and general formulation of the balance metaphor. Reasoning based on recurring general inference patterns seems to pervade every aspect of human cognition. These patterns encapsulate sets of rules to be used in unison, and thereby bypass the combinatorial problems in traditional rule-based deductive inference. The inference patterns are frozen from experience and generalized to apply in many relevant domains.

I have started working on a computational model that acquires and generalizes recurring inference patterns from prior experience [6], but let us focus on the equally basic issue of how such patterns may be used in the reasoning process. Conceptually, the process may be divided into three stages:

1. Index the relevant inference patterns appropriate to the situation at hand. The establishment of the appropriate metaphor is the really difficult part. This is why it is much easier to understand someone's description of observed or experienced events (the metaphor is explicitly referenced by the choice of words), than to generate appropriate action -- the typical distinction between planning and plan comprehension.
2. Instantiate the inference patterns in the specific situation. Computationally, the process of instantiation and the process of searching for appropriate inference patterns are two aspects of the same mechanism.
3. Carry out the inferences stipulated in the retrieved patterns, and check whether additional inference patterns are invoked as a result of the expanded knowledge state.

At the present stage in the investigation, I am searching for general inference patterns and the metaphors that give rise to them, both in mundane and in scientific scenarios. As these patterns are discovered, they are cataloged according to the situational features that indicate their presence. The basic metaphor underlying each inference pattern is recorded along with exemplary linguistic manifestations. The internal structure of the inference patterns themselves are simple to encode in an AI system. The difficulty arises in connecting them to the external world (i.e., establishing appropriate mappings) and in determining the conditions of applicability for each inference pattern (which are more accurately represented by continuous functions than simple binary tests). For instance, it is difficult to formulate a general process capable of drawing the mapping between the "weight" of a hypothetical object and the corresponding aspect of the non-physical entity under consideration, so that the balance inference pattern my apply. It is equally difficult to determine the degree to which this or any other inference pattern can make a useful contribution to novel situations that bear sufficient similarity to past experience [4].

5. Future Directions

If one lends credence to the metaphorical reasoning hypothesis, several avenues of continued research suggest themselves.

- Continue the development of a computational model to test the theory of metaphorical inference and thereby force a finer-grain analysis of the phenomenon.
- Examine the *extent* to which linguistic metaphors reflect underlying inference patterns. The existence of a number generally useful inference patterns based on underlying metaphors is not incompatible with the possibility that the vast majority of metaphors remain mere linguistic devices, as previously thought. In essence, the existence of a phenomenon does not necessarily imply its universal

presence. This is a matter to be resolved by more comprehensive future investigation.

- Investigate the close connection between models of experiential learning and metaphorical inference. In fact, my earlier investigation of patterns of analogical reasoning in learning problem solving strategies first suggested that the inference patterns that could be acquired from experience coincide with those underlying many common metaphors [4, 3].
- Exploit the human ability for experientially-based metaphorical reasoning in order to enhance the educational process. In fact, Sleeman and others have independently used the *balance metaphor* to help teach algebra to young or learning disabled children. Briefly, a scale is viewed as an equation, where the quantities on the right and left hand sides must balance. Algebraic manipulations correspond to adding or deleting equal amounts of weight from both sides of the scale, hence preserving balance. First, the child is taught to use the scale with color-coded boxes or different (integral) weights. Then, the transfer to numbers in simple algebraic equations is performed. Preliminary results indicate that children learn faster and better when they are able to use explicitly this general inference pattern. I foresee other applications of this and other metaphorical inference patterns in facilitating instruction of more abstract concepts. The teacher must make the mapping explicit to the student in domains alien to his or her past experience. As discussed earlier, establishing and instantiating the appropriate mapping is also the most problematical phase from a computational standpoint, and therefore should correspond to the most difficult step in the learning process.

6. References

1. Burstein, M. H., "Concept Formation Through the Interaction of Multiple Models," *Proceedings of the Third Annual Conference of the Cognitive Science Society*, 1981

2. Carbonell, J. G., "A Computational Model of Problem Solving by Analogy," *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, August 1981 , pp. 147-152.

3. Carbonell, J. G., "Metaphor: An Inescapable Phenomenon in Natural Language Comprehension," in *Knowledge Representation for Language Processing Systems*, W. Lehnert and M. Ringle, eds.. New Jersey: Erlbaum, 1982.

4. Carbonell, J. G., "Learning by Analogy: Formulating and Generalizing Plans from Past Experience," in *Machine Learning*, R. S. Michalski, J. G. Carbonell and T. M. Mitchell, eds., Palo Alto, CA: Tioga Pub. Co., 1982.

5. Carbonell, J. G., "Invariance Hierarchies in Metaphor Interpretation," *Proceedings of the Third Meeting of the Cognitive Science Society*, August 1981 , pp. 292-295.

6. Carbonell, J. G., "Acquiring Problem Solving Skills by Analogy," *Proceedings of the Second Meeting of the American Association for Artificial Intelligence*, 1982 , Pittsburgh.

7. Gentner, D., "The Structure of Analogical Models in Science," Tech. report 4451, Bolt Beranek and Newman, 1980.

8. Johnson, M., "Metaphorical Reasoning," 1982. Unpublished manuscript.

9. Kolodner, J. L., *Retrieval and Organizational Strategies in Conceptual Memory: A Computer Model*, PhD dissertation, Yale University, Nov. 1980.

10. Lakoff, G. and Johnson, M., *Metaphors We Live By*, Chicago University Press, 1980.

11. Larkin, J. H. and Carbonell, J. G., "General Patterns of Scientific Inference: A Basis for Robust and Extensible Instructional Systems," 1982. Proposal to the Office of Naval Research.

12. Lebowitz, M., *Generalization and Memory in an Integrated Understanding System*, PhD dissertation, Yale University, Oct. 1980.

13. Ortony, A. (Ed.), *Metaphor and Thought*, Cambridge University Press, 1979.

14. Schank, R. C., "Reminding and Memory Organization: An Introduction to MOPS," Tech. report 170, Yale University Comp. Sci. Dept., 1979.

15. Schank, R. C., "Language and Memory," *Cognitive Science*, Vol. 4, No. 3, 1980 , pp. 243-284.