Invariance Hierarchies in Metaphor Interpretation

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

Abstract

Interpreting metaphors is an integral and inescapable process in human understanding of natural language. The current investigation analyzes analogical mappings underlying metaphors and their implications for inference and memory organization. Regularities have been observed indicating that certain types of conceptual relations are much more apt to remain invariant in analogical mappings than other relations, resulting in an induced invariance hierarchy. The central thesis is that human inference processes are governed by the same analogical mappings manifest as metaphors in language.[1]

1. Introduction

Metaphor is a pervasive phenomenon in almost all human written and spoken language [7, 9, 6]. Recently, I proposed a model of metaphor comprehension based on the identification of general metaphor mappings and subsequent recognition of metaphors as instances of previously-encountered generalized mappings [3]. This method was meant to be a computationally-effective means of interpreting "common" or "mundane" metaphors. As such, it did not address issues of how the underlying analogical mappings structure inference processes; nor did it consider the more difficult task of understanding truly novel metaphors. Here, I investigate the inference processes underlying novel-metaphor comprehension and their implications for memory organization.

Recognition and initial interpretation of metaphors in written or spoken language represents only the tip of the proverbial iceberg. Metaphors in narratives, dialogs and informative text are the observable linguistic manifestations of a central, underlying cognitive process. My thesis is that *cognition is dominated by analogical reasoning* [4], in contrast to more rigid reasoning models based on "sounder" principles of formal logic (e.g., deduction, resolution, etc.). Metaphor is the reflection, on the language medium, of analogical thought processes; as such it provides essential clues of the inner functioning of human inference processes. This paper discusses the utility of metaphor as a tool to investigate various cognitive processes.

2. Two Metaphors are Better Than One

Consider the metaphor ***inflation is war***, as discussed by Lakoff. Newscasters talk of "*fighting* inflation", "workers *taking a beating* from inflation", "Carter *loosing another round to inflation*", "inflation *overrunning* our economy", "savings being *attacked* by inflation", etc. Of what possible use is this metaphor to the reader (or the writer)? Inflation is an economic phenomenon whose causes, implications, and methods of control are not understood by the public at large[2]. (Indeed, some would say that inflation is poorly understood by politicians, economists and business men alike.) Therefore, the metaphor helps to enrich the knowledge brought to bear in the comprehension process by transfering corresponding appropriate information from the more familiar adversary-conflict situation. The necessity to enrich and elaborate upon concepts in the understanding process has been amply demonstrated by Bransford and Johnson [2], Anderson [1] and others. The only feasible way to bring knowledge to bear on an ill-understood domain is to construct a metaphor suggesting a useful transfer mapping of factual information and inference rules from a better-understood domain.

Given the general notion that metaphors transfer knowledge from well-understood to more ill-structured domains, three questions arise:

1. How can the transfered knowledge be used? (i.e., what does the metaphor provide that a literal description may lack?)
2. How does the transfer of knowledge actually take place? (i.e., what would constitute a computationally-viable mechanism?)
3. What implications does the utility and pervasiveness of metaphor have on cognitive processes such as memory organization, inference, and learning?

We consider each question in turn.

3. Why Metaphor?

The knowledge transfered from a richer domain to a more impoverished one via an analogical mapping, triggered by the use of linguistic metaphor, plays a key role in inference processes. Let us return to the ***Inflation is War*** metaphor. If a newspaper article opens with this metaphor, the entire text can be organized around it. Equating inflation with a personified "enemy" enables one to draw upon the knowledge organized under "adversary conflict" to suggest courses of action (in terms of the organizing metaphor)[3]. For instance, we can now understand that in order to "*vanquish* inflation", we clearly must: "formulate a *battle-plan*", "*marshal our forces*", "*take the initiative*", "*go on the offensive*", and "make a *determined attack* on inflation" in order to "stamp it out of our society" and remain on the alert for "future *bouts* with inflation. In short, we must "*whip inflation now* (WIN)" as President Ford said when he "*launched his campaign against inflation*" When the metaphor has been drawn, it is reasonably easy to formulate subgoals based on the better-understood source domain. This is the first step in planning purposive behavior.

The inferences one can draw on how to deal with inflation are all structured by the initial metaphor. Different metaphors will yield markedly different sets of inferences. In order to illustrate that there is nothing inherently special about the ***inflation is war*** metaphor, consider another metaphor used to discuss inflation, encountered frequently in the Spanish press, but easily understood once stated: ***Inflation is a disease***. Here, the economy is the patient, inflation is the infecting organism that must be driven out of the patient with the help of the physician (the economist who sets national economic policy).[4] Hence, one can "*take the pulse* of the economy", "*diagnose* the cause (always placing the blame on external forces - just as disease organisms are an external cause of illness)", "*prescribe treatment*", "put the economy on a *lean diet*", "make the *medicine palatable* to the poor", "wait for private enterprise to *recuperate*", "perform *radical surgery*, cutting *swollen* budgets", "*treat the symptoms* while the inflation continues to ravage", and "*relieve the pain* by subsidizing the price of necessities"

---

[1] This research was sponsored in part by the Office of Naval Research (ONR) under grant number N00014-79-C-0661, and in part by the Defense Advanced Research Projects Agency (DOD), Order No. 3597, monitored by the Air Force Avionics Laboratory under Contract F33615-78-C-1551. The views and conclusions contained in this document are those of the author, and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the U.S. Government.

[2] The popular adage that inflation is "Too much money *chasing* too few goods" is itself a metaphor, one that suggests different corrective action.

[3] This realization is due to Lakoff.

[4] One can imagine this author's confusion upon reading in a Spanish newspaper about the "national malady", the proposal to "inoculate workers" by cost-of-living adjustments, and a "prescription for the national health". What sort of epidemic was on the loose? However, once the metaphor was understood, the text made perfect sense. Since this metaphor ***is the way in which inflation is always discussed***, there appeared to be no need to introduce it explicitly. Moreover, no one would admit that inflation was not a disease, as the metaphor so deeply permeated discussions of inflation that metaphorical terms were not recognized as such. In a conversation with local person, the following statement was made in reaction to my statement suggesting inflation was being discussed in terms of a medical metaphor: "*Of course*, our economy is sick and must be cured, literally! I mean just what I said." This episode should help us step outside our own metaphor and realize that no one can literally *battle* inflation, but that the metaphor is so ingrained in our thinking that we can draw inferences and make statements easily *only if* we rely on the accepted metaphor to structure our reasoning processes.

The moral that can be drawn from these two examples is twofold:

1. **Inference and planning are directly structured by the analogical mapping underlying a dominating metaphor.** The first inflation metaphor suggested *tactics* against inflation, whereas the second suggested *cures* for inflation. Therefore, the inference mechanism consists of mapping corresponding solutions from the source to the target domains. Hence, the metaphor must equate two problems, one of which is better understood and therefore suggests inferences and plans presumed applicable in solving the second problem. (See [4] for a detailed discussion of analogical problem solving.)
2. **Solutions to problems generated by metaphors are ONLY useful as heuristic problem-solving advice.** No **detailed** solutions in ***war*** or ***curing disease*** can be applied directly to inflation. How would one "shoot bullets at inflation" or "get inflation to sign the Geneva convention"? Similarly, one cannot "intern the economy in a hospital" or "give it an intravenous penicillin injection" Clearly, the underlying analogy does not extend to the object level. However, the *planning* level provides useful information. Therefore the **intentions of the actors** are preserved in the mapping, as is the **causal structure** of the events, but the **instantiations of** the events themselves are lost in the analogy. This observation accords with Winston's analogy mappings based on preservation of causal structure [12] and Gentner's discussion of analogy in scientific theory [5].

4. How Metaphors Structure Inference Processes

As we discussed in the preceding section, metaphors can establish an expectation setting for comprehending large portions of text. This expectation setting is generated by transfering inferences from the source to the target domain, including the goals and plans that actors in the target domain are expected to pursue. (It is important to realize that the goals of "defeating" or "curing" inflation come from the respective metaphors -- not from the concept of inflation itself. Therefore a language understanding system must tap the metaphor to comprehend exactly what problems are caused by inflation, and what their respective solution strategies ought to be.)

Let us define *creative metaphors* to be the linguistic realizations of large scale analogical mappings that can structure entire planning episodes. Creative metaphors include the two inflation examples above, Gentner's scientific-theory metaphors [5], and each of the roughly 50 generalized metaphors discussed in [3]. Non-creative metaphors are frozen metaphor instances with fixed meanings, or figures of speech (such as "kick the bucket") whose metaphorical roots can only be traced through their etymology. Non-creative metaphors do not map inferences, as their source domain has been lost in their history, and therefore is not available to the understander. In short, if a metaphor enables one to bring knowledge to bear from an existing domain to a new, less understood domain, we define it as a creative metaphor. The discussion below centers on the process that brings knowledge to bear in understanding new domains.

In order extract information from an existing domain to a new domain via a metaphor, it is crucial to know what aspects of the existing domain should remain invariant in the mapping, which should be transformed, and which should be ignored. As we saw in the previous section, objects are seldom, if ever, preserved in a metaphorical mapping, whereas planning structures are mapped invariant to the new domain in fact it is precisely because planning structure and inferences can be preserved by analogical mappings that metaphors are powerful means of helping an understander formulate reasonable behavior in uncharted domains.

An analysis of some two hundred creative metaphors yields the following empirical observation. There is a well-defined invariance hierarchy among the aspects of a situation that are mapped by a metaphor. This perceived regularity is remarkably consistent across metaphors in different domains. In fact, metaphors that are rated as "bad metaphors" often violate the invariance hierarchy presented below.[5] Hence, a plausible hypothesis is in that people *expect* certain aspects of the source domain to remain invariant and other aspects to be coerced into corresponding entities in the new domain. These expectations can focus the search for metaphor interpretations. The regularities observed over a large number of metaphors are summarized by the normative invariance hierarchy presented below. The conceptual relations in the hierarchy are listed in decreasing order of expected invariance:

- *A goal-expectation setting for the animate actors involved (if any).* Goals, if present in the source domain, are almost always mapped invariant into corresponding entities in the target domain. If the source domain contains animate actors and the target does not, then the goals of the actors will be attributed to the corresponding personified entities in the target domain. E.g., inflation becomes an anthropomorphized malevolent agent in ***inflation is war***, therefore the *the goals* or a nation at war are mapped invariant in that inflation must be *fought and defeated*.
- *Planning and counterplanning strategies among competing or cooperating actors.* -- These strategies, almost always preserved intact by an analogy, provide a priority ordering among the goals and suggests possible means for pursuing each goal. Often, the most useful aspect of a metaphor is to enable purposive planning in what previously was too ill-structured a domain.
- *Causal Structures.* -- When the causal structure of the source domain is explicit, it will typically be preserved by the mapping. E.g., medicine cures disease; therefore economic measures will "cure" inflation. In Reddy's conduit metaphor for how people talk of language [9], causal structure abounds. E.g., a blocked conduit prevents physical transfer; therefore press censorship will also block dissemination of ideas.
- *Functional Attributes.* -- The function to which an object in the source domain is typically applied will often be coerced onto an analogous function for the corresponding object in the target domain.
- *Temporal Orderings.* Normative planning sequences in the source domain map into potentially applicable planning sequences (instantiated differently) in the target domain, often preserving temporal relations.
- *Natural tendencies.* -- In the celebrated analogy between electric circuits and and hydraulic systems (used to explain Ohm's Law), water "tends" to go down hill, therefore electricity "tends" to go towards the voltage "*drop*" Moreover, thin pipes resist the flow of water, therefore thin wires "resist" the "*flow*" of electricity.[6]
- *Social roles.* -- Social relations are sometimes preserved and sometimes not. In a battle there are generals and foot-soldiers; therefore, the war against inflation must be fought by many wage-earners (soldiers) under the direction of economic planners (generals). Since doctors cure the disease directly, the economic planner must shoulder the entire burden, and wage-earners (patients) are essentially powerless with respect to taking an active role in the cure. Both mappings preserve the inferences associated with the social roles in the source domain. However, the more specific roles of "spy" and "submarine commander", are not preserved by the ***inflation is war*** metaphor.
- *Structural relations.* -- Occasionally structural relations remain invariant in an analogy, but often they are transformed or suppressed. For instance, in the Rutherford solar-system model of the atom, physical relations between the electrons (planets) and nucleus (sun) are remain invariant. (In both case there is an orbit relation as a function of an inverse-square centripetal force). However, Saying "John is at the head of his class" does not preserve the physical structure normally found between a body and a head (the latter being connected to and nourished by the former).
- *Descriptive Properties.* -- These are the least likely properties to be preserved in a metaphor. Wires and pipes are both long and narrow (in the hydraulics metaphor) However, Generals

---

[5] This observation is based on data collected by Lakoff, Gentner, Ortony and others

[6] Electricity is actually not a flow of electrons, but we always think of it that way because the hydraulics metaphor pervades our discussions of electrical phenomena

are military men, economic planners are usually academics. The sun is yellow-orange, very large and has sunspots, but as Gentner points out, none of these monadic descriptors apply to the nucleus of the Rutherford atom.

- *Object identity.* Objects in the source domain are almost never mapped onto objects of identical type in the target domain. Therefore, there are no tanks, bullets, M16's, attack submarines, uniforms, or field hospitals in the battle against inflation.

5. Analogical Mappings in Problem Solving

Our discussion suggests that metaphors are a useful means of indexing mappings between goals, planning structures, causal connections, tendencies, relations, and descriptions (in decreasing order of invariance and significance). Not all components are present in every metaphor. The preferred-invariance ordering helps us understand how metaphors may be used to facilitate reasoning processes in new domains, namely:[7]

1. Establish the invariant components of the mapping
2. Establish initial correspondences among the entities in the source and target domains. (This is a very partial correspondence -- only entities that are referenced by an invariant component in the explicit mapping can be directly related.)
3. Goal-correspondence identifies the problems that must be solved in the target domain. [What should one do about inflation? The ***disease*** metaphor metaphor states that it should be eradicated. The ***war*** metaphor suggests subjugating it. A comparison of inflation with an overindulging *gourmand* would yield the goal of trimming it down and controling its scope, but not eliminating it.] Therefore, metaphor actually determines the goals that one ought to pursue in the target domain. Without knowledge of goals little purposive action can take place (i.e., problem solving becomes meaningless, as there are no goal states in the problem space.)
4. Planning strategies invariant under a metaphorical mapping transfer operators from the source to the target domain, hence establishing a problem space [8] and suggesting potentially troublesome interactions among operator preconditions. The ***inflation is a disease*** metaphor suggests that since administering medicine is a useful operator in the medical domain, a correlate operator ought to be useful in the economic domain. Moreover, medicine is usually an unpleasant experience, therefore the inference is made that its economic correlate would be unpleasant as well. Hence, we speak of giving the economy a strong *dose* of anti-inflationary monetary restraints, and making the policy *palatable* to workers.
5. Causal connections classify operators in the target domain by the differences they reduce (analogized from the source domain) "The pressure of the water is determined by the product of the rate of flow and the cross-section of the pipe" suggests that in order to know the voltage, one can measure the current and resistance. Therefore a way of reducing the KNOW-V goal is to apply the multiply(I, R) operator, reducing the KNOW-V goal to the subgoals KNOW-R and KNOW-I.
6. Natural tendencies, social roles and structural relations provide information about the applicability conditions of operators [E.g., who can administer medicine (decide economic policy)?], and provide heuristic guidance to planning processes in the new domain. [E.g., Wars are costly and people must make personal sacrifices; therefore in battling inflation the cost should be taken into account, and the planner should be ware of potential problems caused by those who are unwillingly called upon to make the sacrifices.]
7. Temporal-progressions suggest macro-operators (typically useful sequences of operators) In treating a disease we first must identify the cause, then prescribe medicine then wait patiently for it to take effect. In war we marshal our forces (no searching for a suitable the enemy is needed, as the enemy is known at the start of hostilities), then attack (no waiting for the attack to take effect is necessary). Therefore we see two very different general plans suggested by the two metaphors. However, recall that the metaphors shared the same general goal. It is typically the case that most metaphors used to explain a particular ill-defined situation will share common goals and diverge increasingly as one traverses down the invariance hierarchy.

Reiterating the central theme of this section: metaphors provide a problem space, including a goal state, operators indexed by differences they may reduce, and normative plans that may prove useful. In essence, they make problem solving possible in what may previously have been too ill-structured a situation to make any progress. Metaphors do ***not***, however, provide any ***canned*** solutions applicable directly to new problems such would be an unreasonable expectation.

6. Exploiting the Invariance Hierarchy

The invariance hierarchy provides a first-pass solution to an apparently simple phenomenon that had perplexed some investigators, including this writer. When we hear that "John is a fox" we interpret it to mean "John is sly", not "John has pointed ears and a bushy tail." Similarly, we interpret "John is a pig" as a remark on his personal habits or his obesity, rather than a statement that John lives in a farm and has a curly tail. A partial answer to this problem lies in knowing the most salient feature of the animal to whom a person is compared. However, a more complete answer is provided by exploiting the invariance hierarchy in the following manner: Consider the animal (source domain) and scan down the hierarchy stopping at the first entry for which we have a commonly known fact. For foxes we stop at planning/counterplanning -- folk wisdom tells us that foxes are very adept at devious counterplanning behavior. Hence, we never reach the physical descriptors of a fox. For pigs we may stop at either "natural tendency" (if we believe that pigs tend to get fat) or at "social role" -- folk wisdom asserts that pigs play a distinct role in the animal kingdom as the least hygienical of all animals. If we heard "John is a Giraffe", we find no common knowledge anywhere in the hierarchy until we reach physical attributes. Here we pick the most salient ones (e.g., height and/or length of neck), to understand the metaphor. The key to the process is that comparisons along the higher-invariance entries in the hierarchy are preferred. Once a high invariant property is found, no lower ones are considered. This is crucial to understand "John is an elephant" as a remark on the length of his memory (or his capacity for work), not the length of his nose (trunk), although the latter is perhaps the single most salient feature of elephants. Physical descriptors, however, are ranked low in the hierarchy.

7. Implications for Memory Organization

We have outlined how reasoning based on metaphors may proceed. Now, consider another aspect of metaphorical reasoning: *How are metaphors formed in the first place?* Given the ubiquity of metaphor, it becomes strikingly apparent that humans generate metaphors as readily as they understand them, occasionally unconscious of the fact that they are creating (or more often instantiating) metaphors. The question that must be posed is more specific: *What memory organization could enable, facilitate and encourage the continuous creation of metaphors?*

If we assume that the invariance hierarchy is roughly correct, it provides a best-first criterion for searching a content-addressed episodic memory, organized along the general lines of Schank's MOPS [11, 10]. In investigating reminding and inference phenomena, Schank asserts that detecting similarities at every level of abstraction is the key to human memory organization. Accepting this notion requires one to have a means of computing similarities among large numbers of potentially-relevant episodic traces, both for memory access and update. The hierarchy above suggests that goal similarities are crucial, planning-level similarities are almost as important, and similarities across other dimensions are of progressively lower importance. Hence, if memory were organized according to the computational criteria required for metaphor comprehension, it follows that a

[7] Here I adopt Newell and Simon's Means-Ends Analysis framework for problem solving [8]. The reader is referred to [4] for a more detailed discussion of analogical problem solving.

hierarchical structure would result, where the categories formed are largely determined by groupings along the entries in the invariance hierarchy, the more invariant entries corresponding to more global organizing categories. The actual content of the hierarchical memory is determined primarily by the idiosyncratic experience of the individual. Therefore, memory searchers for "good metaphors" (those preserving high-invariance properties) require less work (either to generate or comprehend) and may prove more rewarding for the understander as they index relevant memory more readily.

Metaphor is a linguistic realization of an inference phenomenon. As such, it should reflect underlying memory structure, as well as suggest the types of inferences people can perform most readily. If we ask *why* creative metaphors are used, the most logical answer appears to be that the writer is trying to induce the reader to perform the necessary inferences required to comprehend the new material. Metaphor serves as a vehicle to suggest a fruitful domain from which the relevant inferences can be mapped onto the new domain. Hence, when Senator Joe McCarthy refered to Communism as a "dreaded plague", he was inducing, in the minds of his listeners, the inference that communism must be actively "eradicated" or it will spread. The metaphor is effective only because the appropriate inference structure was already in existence in the source domain, and McCarthy knew this at the time he created the metaphor.

An interesting avenue of future research is automating metaphor generation. If the model discussed here is essentially correct, metaphor generation requires that the writer have a model of the knowledge state (including goals, strategies, beliefs, etc.) of his reader, as well as an integrated episodic memory where the computed similarity metric incorporates the invariance hierarchy. (I.e., two domains are considered similar if the same types of problems and inference processes are present in planning effective behavior in both domains.)

In order to clear possible misconceptions, I emphasize that no distinct, localized, "conscious" existence for the invariance hierarchy is postulated as part of a human memory model. My hypothesis is that *the regularities manifested in the hierarchy are epiphenomenal reflections of human memory organization and inference mechanisms.* As such, the invariance hierarchy summarizes a phenomenon that must be explained by comprehensive memory-organization models, and hence it ought to be taken into account in the model-formulation process.

References

1. Anderson, J. R., *Language, Memory and Thought,* New Jersey: Erlbaum, 1976.

2. Bransford, J. D. and Johnson, M. K., "Considerations of Some Problems of Comprehension," in *Visual Information Processing,* W. G. Chase, ed., New York: Academic Press, 1973.

3. Carbonell, J. G., "Metaphor - A Key to Extensible Semantic Analysis," *Proceedings of the 18th Meeting of the Association for Computational Linguistics,* 1980 .

4. Carbonell, J. G., "Learning by Analogy: Skill Acquisition in Reactive Environments," in *Machine Learning,* R. S. Michalski, J. G. Carbonell and T. M. Mitchell, eds., Palo Alto, CA: Tioga Pub. Co., 1981.

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

6. Hobbs, J. R., "Metaphor, Metaphor Schemata, and Selective Inference," Tech. report 204, SRI International, 1979.

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

8. Newell, A. and Simon, H. A., *Human Problem Solving,* New Jersey: Prentice-Hall, 1972.

9. Reddy, M., "The Conduit Metaphor," in *Metaphor and Thought,* A. Ortony, ed., Cambridge University Press, 1979.

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

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

12. Winston, P. H., "Learning and Reasoning by Analogy," *CACM,* Vol. 23, No. 12, 1979 , pp. 689-703.