

METAPHOR: THE GOOD, THE BAD AND THE UGLY¹

Dedre Gentner Brian Falkenhainer* Janice Skorstad**
Psychology Department Computer Science Computer Science
University of Illinois at Urbana-Champaign

Metaphor is a pervasive and important phenomenon, both in literature and in ordinary language. It is also an immensely variable phenomenon. The term 'metaphor' is often used to refer to nonliteral comparisons that are novel and vivid and that convey ideas that might otherwise be difficult to express (Ortony, 1975). But the term has also been used to refer to systems of extended meanings that are so familiar as to be almost invisible, such as the spatial metaphors 'soaring spirits' or 'falling GNP' (Lakoff & Johnson, 1979; Nagy, 1974). Even if we restrict ourselves to literary metaphors, there is still an enormous range of metaphor types, as shown in the following list:

1. She allowed life to waste like a tap left running (Virginia Wolfe).
2. I have ventured, /Like little wanton boys that swim on bladders, /This many summers in a sea of glory; /But far beyond my depth: my high-blown pride /At length broke under me; and now has left me, /Weary and old with service, to the mercy /Of a rude stream, that must forever hide me. (Shakespeare)
3. For the black bat, night, has flown (Tennyson)
4. The glorious lamp of heaven, the sun (Robert Herrick)
5. On a star of faith pure as the drifting bread, /As the food and flames of the snow (Dylan Thomas)
6. the voice of your eyes is deeper than all roses (Cummings)

Perhaps because of this staggering variety, there is little consensus on how metaphor should be defined and analyzed. Most would agree that metaphors are nonliteral similarity comparisons (though not everyone would agree on how literality should be defined), and that they are typically used for expressive-affective as opposed to explanatory-predictive purposes. But beyond this, metaphor has remained elusive of analysis. In this paper we offer a partial solution. We use Gentner's (1980, 1983, 1986) structure-mapping framework to distinguish three classes of metaphors — two that are computationally tractable within the framework and one that is not. Then we demonstrate how the analysis works, using the Structure-mapping Engine, a simulation written by Brian Falkenhainer and Ken Forbus (Falkenhainer, Forbus, & Gentner, 1986).

This research was supported in part by the Office of Naval Research under Contract No. N00014-85-K-0559, NR667-551.

* The author is currently supported by an IBM Graduate Fellowship.

** The author is currently supported by a University of Illinois Cognitive Science/AI Fellowship.

1. We mean 'ugly' here in the sense of 'computationally intractable.' We use 'metaphor' here to refer to both metaphor and simile.

The basic intuition of structure-mapping theory is that an analogy is a mapping of knowledge from one domain (the base) into another (the target) which conveys that a system of relations that holds among the base objects also holds among the target objects. Thus an analogy is a way of noticing relational commonalities independently of the objects in which those relations are embedded. In interpreting an analogy, people seek to put the objects of the base in 1-to-1 correspondence with the objects of the target so as to obtain maximum structural match. The corresponding objects in the base and target don't have to resemble each other at all; object correspondences are determined by roles in the matching relational structures. Central to the mapping process is the principle of **systematicity**: people prefer to map systems of predicates that contain higher-order relations with inferential import, rather than to map isolated predicates. The systematicity principle is a structural expression of our tacit preference for coherence and deductive power in interpreting analogy.

Besides analogy, other kinds of similarity matches can be distinguished in this framework, according to whether the match is one of relational structure, object descriptions, or both. Recall that *analogies* discard object descriptions and map relational structure. *Mere-appearance* matches are the opposite: they map aspects of object descriptions and discard relational structure. *Literal similarity* matches map both relational structure and object-descriptions.

Kinds of Metaphors: Now let us apply this framework to metaphor. We can distinguish three rough categories of metaphors: relational metaphors, attributional metaphors, and complex metaphors that cannot be simply analyzed. Relational metaphors — e.g., metaphors (1) and (2) — are mappings of relational structure. They can be analyzed like analogies. Attributional metaphors — e.g., metaphors (3) and (4) — are *mere-appearance* matches: their focus is on common object attributes. Among these two classes, adults (but not children) seem to prefer relational metaphors (Gentner, 1980; 1986). So far both these classes can readily be described in structure-mapping terms: both utilize 1-to-1 object mappings and are characterizable by their distribution of relational and attributional predicates. The third class, which we will not attempt to analyze, is exemplified by metaphors (5) and (6). These metaphors lack clear 1-to-1 mappings; they are characterized many cross-weaving connections with no clear way of deciding exactly how the base predicates should attach in the target (See Gentner, 1982).

To illustrate the way in which relational metaphors can be analyzed, we now describe the operation of SME on metaphor (1):

She allowed life to waste like a tap left running.

The representations for base and target are shown in Figure 1. We assume the reader starts off with some notion of water flowing through a tap into a drain, and with the idea that waste occurs if an agent allows such a flow to occur with no purpose. In the target domain of life it is less clear exactly what to assume as initial knowledge. In this example we have chosen a rather sparse description. We assume that the reader has the

idea that life flows from present to past. Since the information that the protagonist's life is being wasted is given directly, we also include that knowledge in the initial life representation.

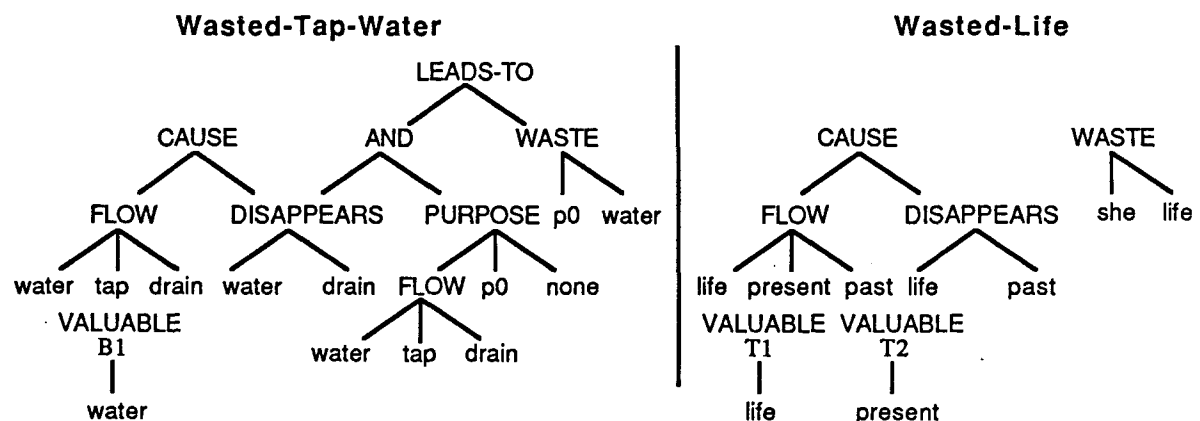


Figure 1. Wasted-Tap-Water and Wasted-Life Descriptions

SME starts by finding *local matches* — potential matches between single items in the base and target. For each entity and predicate in the base, it finds the set of entities or predicates in the target that could plausibly match that item. These potential correspondences (*match hypotheses*) are determined by a set of simple rules:²

- (1) If two relations have the same name, create a match hypothesis;
- (2) For every match hypothesis between relations, check their corresponding arguments: if both are entities, or if both are functions, then create a match hypothesis between them.

Here, rule (1) creates match hypotheses between the FLOW relations which occur in base and target. Then rule (2) creates match hypotheses between their arguments: water-life, tap-present, drain-past. At this stage the program may have a large number of local matches, possibly mutually inconsistent. Another set of rules assigns *evidence scores* to these local matches:

- (1) Increase the evidence for a match if the base and target predicate have the same name.
- (2) Increase the evidence for a given match if there is evidence for a match among the parent relations — i.e., the immediately governing higher-order relations.

Rule (1) reflects a preference for relational identity and rule (2) reflects a preference for systematicity. Here, match between the FLOW predicates discussed above gains evidence from the identity of the FLOW predicates themselves (by evidence rule (1)) and also from the identity of the parent CAUSE relations (by evidence rule (2)).

2. This description is for analogy. SME can also be run with different match rules to simulate mere-appearance matches and literal similarity matches.

The next stage is to collect these local matches into *global matches* — systems of matches that use consistent entity-pairings. SME propagates entity-correspondences upward and finds the largest possible systems of matched predicates with consistent object-mappings. These global matches, called Gmaps, are the possible interpretations of the analogy. Figure 2a shows the Gmap for the life/water example.³ Associated with each Gmap is a (possibly empty) set of *candidate inferences* — predicates that are part of the base system but were not initially present in the corresponding target system. These will be hypothesized to be true in the target system. In this case, the system brings across the inference that the protagonist is letting her life pass with no purpose, and that this purposeless flow is causing her life to be wasted. Finally, each Gmap is given a structural evaluation, which depends on its local match evidence.⁴

SME can also operate in mere-appearance mode to process attributional metaphors. Figure 2b shows the interpretation that metaphor (1) receives under these matching rules. Clearly the relational interpretation is preferable in this case.

Gmap #1: { (WASTE ↔ WASTE) (FLOW ↔ FLOW) (DISAPPEARS ↔ DISAPPEARS)
 (CAUSE ↔ CAUSE) (p0 ↔ she) (tap ↔ present) (water ↔ life) (drain ↔ past) }
 Weight: 6.7018
 Candidate Inferences: { (LEADS-TO (AND (DISAPPEARS life past)
 (PURPOSE (FLOW life present past) she none))
 (WASTE she life)) }

(a)

Gmap #1: { (VALUABLE_{B1} ↔ VALUABLE_{T2}) (water ↔ present) }
 Weight: 0.9500
 Candidate Inferences: { }
 Gmap #2: { (VALUABLE_{B1} ↔ VALUABLE_{T1}) (water ↔ life) }
 Weight: 0.9500
 Candidate Inferences: { }

(b)

Figure 2. (a) Analogy Match Rules, (b) Mere Appearance Match Rules

Comments: A few points about the simulation model should be noted. First, SME's interpretations are extremely sensitive to the knowledge representations of base and target. We think this roughly reflects the state of affairs in human processing of analogy and metaphor. Second, SME's matching process is entirely

-
3. Because of the sparseness of the representations, only one Gmap is discovered. When we run this example with richer representations, adding such potentially confusing information as "Life consumes water." in the life domain, we find more Gmaps, although the highest evaluation still goes to the Gmap shown here.
 4. The system also has the capability to consider the number of candidate inferences and the graph-theoretic structure in determining the evaluation, but their ramifications need to be explored. It is interesting that the simple version of systematicity embodied in the local evidence rules seems to lead to very reasonable interpretations.

structural. SME arrives at its interpretation by finding the most systematic mappable structure consistent with the 1-to-1 mapping rule. The reason that relatively interesting interpretations are found is that the systematicity principle operates to promote predicates that participate in causal chains and in other constraining relations. Unlike some current models of analogy (e.g., Holyoak, 1985), structure-mapping does not need to use a prior goal-structure to select its interpretation.⁵ This makes it particularly apt for the interpretation of novel metaphors, in which we may have no advance knowledge of the content of the interpretation.

In conclusion, structure-mapping can handle the good and the bad — ie., either relational or attributional mappings that are 1-to-1. Whether it can handle the ugly — the complex n-to-1 mappings — remains to be seen.

Acknowledgements: The authors wish to thank Ken Forbus for his invaluable assistance.

References:

- Burstein, M. H. (1983). Concept Formation by Incremental Analogical Reasoning and Debugging. *Proceedings of the 1983 International Machine Learning Workshop*, University of Illinois, Monticello, IL.
- Carbonell, J. G. (1983). Learning by Analogy: Formulating and Generalizing Plans from Past Experience. In R. S. Michalski, J. Carbonell, and T. Mitchell (Eds.), *Machine learning*. Palo Alto, CA: Tioga Publishing Company.
- Falkenhainer, B., Forbus, K. D. & Gentner, D. (1986). The Structure-Mapping Engine. *Proceedings of the American Association for Artificial Intelligence*. Philadelphia, PA.
- Gentner, D. (1980). *The Structure of Analogical Models in Science* (BBN Rpt. No. 4451). Cambridge, MA: Bolt Beranek and Newman Inc.
- Gentner, D. (1982). Are Scientific Analogies Metaphors? In D. Miall (Ed.), *Metaphor: Problems and Perspectives*. Brighton, England: Harvester Press Ltd.
- Gentner, D. (1983). Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, 7(2), 155-170.
- Gentner, D. (1986). Mechanisms of Analogy. To appear in S. Vosniadou and A. Ortony (Eds.) *Similarity and Analogical Reasoning*.
- Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. Chicago, IL: University of Chicago Press.
- Nagy, W. (1974). Figurative Patterns and Redundancy in the Lexicon. PhD. dissertation, University of California at San Diego.
- Ortony, A. (1975). Why Metaphors are Necessary and Not Just Nice. *Educational Theory*, 25, 45-53.

5. Of course, if there were a specified contextual goal, then the output of the Structure-Mapping engine would have to be evaluated with respect to that goal by a further processor. (See Burstein, 1983; Carbonell, 1983)