------------------------------------------------------------- CS 224U Lit Review ------

**Sources**

1. [Metaphor Comprehension: A Computational Theory](http://download-v2.springer.com/static/pdf/59/art%3A10.3758%2FBF03212981.pdf?token2=exp=1430431545~acl=%2Fstatic%2Fpdf%2F59%2Fart%3A10.3758%2FBF03212981.pdf*~hmac=5f8f0d45d652f9a7b1ef3359044a56a97f3f38ae4693f53d4edf0c552e811eec)
2. [Metaphor – A Key to Extensible Semantic Analysis](http://aclweb.org/anthology/P80-1004)
3. [On Understanding Metaphor: The Role of Context](http://ac.els-cdn.com/S0022537183903559/1-s2.0-S0022537183903559-main.pdf?_tid=8ea2c9ac-ef82-11e4-9e7e-00000aacb362&acdnat=1430430655_86f8d393a491dd38eec619619958f872)

**Motivation**

Why study metaphors? In reading, and especially critical reading, the ability to interpret metaphors is integral to high-level understanding of texts. In fact, on the SAT, the most challenging questions test obscure metaphor comprehension:

*“On line 28, the author’s allusion to \_\_\_\_\_\_\_ most likely means…”*

Clearly, a robust Natural Language Understanding model would include the ability to recognize and understand the deep meaning and intentions behind metaphors. A successful Natural Language Understanding model, especially with respect to the SAT, must therefore be robust enough to comprehend metaphors at a sufficient level. The following three papers detail the little research performed on computational metaphor comprehension in the natural language processing domain; they detail both the current importance, successes and problems of understanding metaphors.

**1 Metaphor Comprehension: A Computational Theory**

Metaphor comprehension involves an interaction between the meaning of the topic and the vehicle of the metaphor. Meaning is represented by vectors in high-dimensional semantic space. The topic is modified by the predication by merging the topic with selected features of the vehicle vector. It defines the Argument as the topic of the metaphor and the Predicate as the vehicle of the metaphor. As an example, in the metaphor “That girl is a Barbie,” the argument/topic is “the girl” and the predicate/vehicle is “the Barbie”. This method is based two assumptions. The first, to understand the metaphor “Billboards are warts on a landscape”, the paper makes the assumption that understanding the metaphor does not require abstract conceptions of the argument, “billboards”. The second assumption assumes that metaphors are not reversible: “my surgeon is a butcher” does not have the same meaning as “my butcher is a surgeon”.

This paper comments that common methods of vectorizing words like LSA capture semantic meaning, but not perception and real world manifestation. It also comments that averaging the vectors of both the topic and the vehicle makes little sense, as the average of two unrelated vectors results in a vector pointing to a semantic no-man’s land. Instead, the paper proposes an algorithm for metaphoric composition:

1. Compute the semantic neighborhood of the predicate (P) of size *m*, by taking the m nearest neighbors of P’s vector. For metaphors, *m* has to be fairly large (500 < *m* < 1,500) because the predicate and argument in a metaphor often are quite unrelated. This step ensures that all terms that enter into the predication are in fact related to P.
2. The next step picks those terms from the neighborhood of P that are also related to the argument (A). The cosines between the *m* neighbors of P and A are computed, and the *k* terms with the highest cosine are selected. This step obviates the need for setting up a huge self-inhibitory network and yields much the same results, because there are usually only a few items related to both P and A and these would be selected in either case.
3. In order to avoid introducing noise by selecting the strongest terms even when their absolute strength is low, the terms selected must have a cosine with P and A above some minimum threshold. Additionally, any terms that have a cosine distance lower than a certain threshold (two standard deviations above the mean is recommended) from P and A will be eliminated.
4. The vector representing the meaning of the metaphor can then be computed as the centroid of A and the terms selected above (P and the *k* terms from the neighborhood of P, subject to the restriction that their cosine with A is above threshold).

The results of this paper are promising: this method achieved 83.3% accuracy. The example it missed was “my marriage is the icebox”. The paper attributed this error to the length of the “icebox” word-vector – its length was too small, meaning this model was not fed enough prior knowledge of “iceboxes” to interpret the metaphor.

While these results are promising, they must be taken with a hefty grain of salt. This paper only analyzes seven metaphors – quite blatantly insufficient to make any substantial claim to metaphoric comprehension. Additionally, this paper makes a lofty assumption of metaphoric non-reversibility. This is not necessarily true: “Food is bae” holds the same meaning and implications as “Bae is food”.

**2 Metaphor – A Key to Extensible Semantic Analysis**

This paper suggests a generalized method to analyze metaphors based on the existence of a small number of metaphor mappings. It claims that each generalized metaphor contains (1) a recognition network, (2) a basic mapping, (3) additional transfer mappings and (4) an implicit intention component. By breaking down a metaphor into these four categories, this paper attempts to reduce metaphor interpretation from a reconstruction task to a recognition task. The paper attempts this approach because it is much more tractable to recognize a class and framework of a metaphor than to construct the conceptual meaning of it each time a new instance is encountered.

The paper deviates from the traditional interpretation of the metaphor (vehicle/topic) and argues that metaphors are mappings of the from “X is used to mean Y in context Z”. Specifically, it argues that each metaphor can be interpreted in the following manner:

1. A metaphor contains a recognition network, which contains enough information to identify the vehicles and topics of the metaphor.
2. A metaphor is a basic mapping that maps an interpretation of two related arguments of the metaphor to another in some directly quantifiable way.
3. Metaphors contain an implicit-intention component: each metaphor has attached “invisible” semantic connotations which adds to the semantic meaning of the metaphor.
4. Each metaphor may contain both literal and conceptual semantics after basic mapping, which then may have to be additionally mapped by Transfer Mapping.

This paper then outlines a high-level methodology/algorithm to understand these metaphors:

1. Attempt to analyze the input utterance in a literal, conventional fashion. If this fails, and the failure is caused by a semantic case-constraint violation, go to the next step. (Otherwise, the failure is probably not due to the presence of a metaphor).
2. Apply the recognition networks of the generalized metaphors. If one succeeds, then retrieve all the information stored with that metaphorical mapping and go on to the next step. (Otherwise, we have an unknown metaphor or a different failure in the original semantic interpretation. Store this case for future evaluation by the system builder.)
3. Use the basic mapping to establish the semantic framework of the input utterance.
4. Use the transfer mapping to fill the slots of the meaning framework with the entities in the input, transforming them as specified in the transfer map. If any inconsistencies arise in the meaning framework, either the wrong metaphor was chosen, or there is a second metaphor in the input (or the input is meaningless).
5. Integrate into the semantic framework any additional information found in the implicit-intention component that does not contradict existing information.
6. Remember this instantiation of the general metaphor within the scope of the present dialog. It is likely the same metaphor will be used again with the same transfer mappings present, but with additional information conveyed.

While this is a comprehensive algorithm and interesting paper, there are several problems that I see with it. The first, the algorithm is so high-level that it loses all use; it is unreasonable to be able to assume that you can already detect a metaphor (step 1 of the algorithm). Second, while this algorithm has been implemented (in a system called MIDAS), it requires a large initial base of metaphors and a code complexity that might not be achievable in our short timespan. The largest and most important problem with this paper is that it doesn’t attempt to *comprehend* or *understand* metaphors, it only attempts to recognize them and interpret them in a more traditional Natural Language Processing fashion. While this algorithm may be effective in translating metaphors to more literal semantics, it circumvents truly understanding the metaphor. However, this paper offers an interesting and alternative methodology on how to construct and interpret the metaphor, which gives valuable insight in how we could algorithmize metaphoric comprehension.

**3 On Understanding Metaphor: The Role of Context**

This paper documents non-computational research on how metaphors are verbally comprehended with requisite context. A note about its generalizability: it focuses on only metaphors with the structure “*All (Topic)s are (Predicate)s”*. It hypothesizes that any context which activates a property of the predicate that is informative about the topic is sufficient to trigger immediate metaphoric comprehension. For example, consider the metaphor “All marriages are iceboxes”. According to this theory, the minimal context that includes additional semantic information on *iceboxes* (the predicate) that may semantically apply to *marriages* (the topic) is a sufficient minimal context.

In another paper, they cite that humans comprehend metaphors non-optional; we comprehend some metaphors automatically without trying to understand them. On the other hand, when some metaphors such as “*all marriages are iceboxes*” are unintuitive, humans have to very intentionally comprehend them. This is analogous to the computer case: we start with the assumption that all metaphors are unintuitive to computers. The fact that humans can consciously and methodologically attempt to understand metaphors suggests that computers may be able to do so, too.

This paper details an experiment where they primed metaphors with both literal and figurative, semantically-similar information to aid in metaphor comprehension. They found that both these primes helped metaphor comprehension. They further found that by providing a priming dimension for the metaphor to be evaluated on, metaphoric comprehension became much easier. For example, the metaphor “John is basically an elephant” is naturally ambiguous – there are many dimensions of an elephant that may be attributed to John. By priming one attribute, for example memory, this metaphor becomes more clear: “Elephants have great memory. John is basically an elephant.” In this way, metaphoric comprehension becomes somewhat of a disambiguation task.

In application to our project, what defines a minimal context? According to this paper, it would be defined as the closest preceding sentence that helps define the salient dimension of the predicate (elephant) that should be applied to the topic (John).