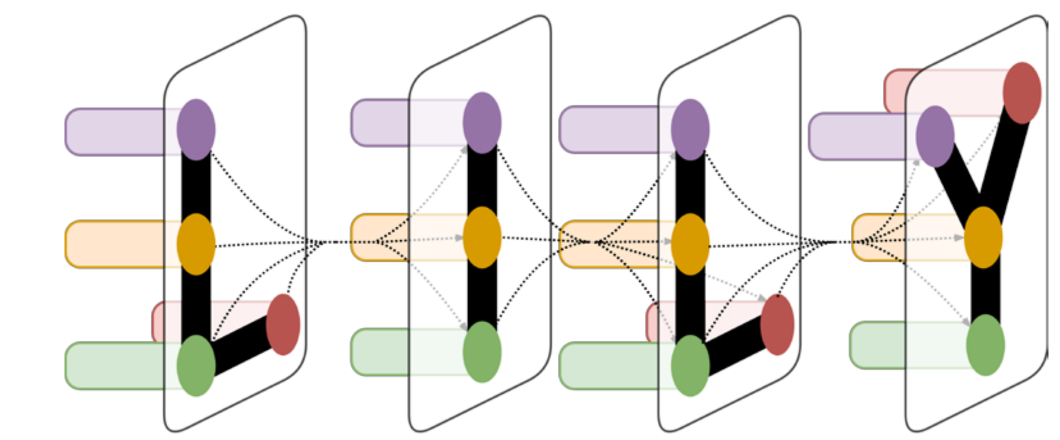
Interpreting Metaphors using Lessons drawn from Psychology

Daniel Keller

[1]LILY Lab, Yale University



Introduction

Ever since the field of computational metaphor has started, many researchers have approached the issue from a linguistic viewpoint. Shutova's survey on the computational metaphor field starts with a discussion of linguistic considerations, representing the importance linguistics has had on the field (Shutova, 2015).

A fair amount of work has revolved around conceptual metaphors, but this work has primarily focused on using common linguistic metaphorical mappings such as FEELINGS ARE LIQUIDS ("my anger was stirring within me"), LIFE IS A JOURNEY ("His sudden death was an unexpected arrival at the end of his life."), and others. These mappings are determined through a study of corpuses, and projects that have tried to build corpuses with these annotations have often given little thought to the psychologically processes that lead annotators to allocate phrases as metaphorically or literal (Shutova and Teufel, 2010).

A large body of work has also clearly established the importance of metaphors in NLP.

There is a large amount of psychology and cognitive science research on metaphor that often isn't used to motivate computational models of metaphor. There is consensus among cognitive linguists that metaphorical inference is the basis for the use of metaphor(Hobbs 1981; Carbonell 1982; Rohrer 1997; Turner and Fauconnier 2003; Feldman 2006); however, the study of this inference from the side of psychology has rarely been used. Furthermore, there is a large amount of psychology research that has analyzed how humans understand metaphor. For example, Tversky (1977) went to significant length to model how words metaphorical meanings can often be determined through a feature matching between two words' bags of features. This work was motivated by Tversky and Kahneman's (1973) study of the availability heuristic, a mental heuristic that human minds use that causes us to give unnecessary amounts of importance to the first things that come to our mind. This work, as it addresses metaphor, could easily be leveraged to create an automatic model.

Figure 1. Example Feature Set

'writings': ('other', 0.07482993197278912), ('own', 0.034013605442176874), ('recent', 0.027210884353741496), ('religious', 0.013605442176870748), ('extreme', 0.013605442176870748)

Results for Metaphor: love is a gold.
K = 5 Results:

Word Net:

Feature: more Strength: 0.194444444444.
Feature: individual Strength: 0.0298637588572.
Feature: more Strength: 0.0204081632653.
Feature: relay Strength: 0.0180786319324.
Feature: much Strength: 0.00294678530762.
Feature: other Strength: 0.0.

Figure 2. Sample output from the interpreter

Topic-Vehicle	Properties as Metaphor				Properties as Simile			
Alcohol-Crutch	Response	Freq	Sal	Conn	Response	Freq	Sal	Conn
	Helpful	5	1.80	3.50	Dependent	5	1.20	3.81
	Dependable	4	1.25		Supportive	4	1.25	
	Addictive	3	1.67		Bad	2	1.50	
	Support	3	3.00		Addictive	2	2.00	
	Disability	2	2.00		Hard	2	2.00	
	Aid	2	2.00		Necessity	2	2.00	
	Problem	2	2.00		Limiting	2	2.00	
					Help	2	2.00	

Figure 3. Example Test Data

Materials and Methods

In this experiment, we used Tversky's bag of fea- tures theory along with Tversky and Kahneman's availability heuristic to design a model that in- terprets metaphors. The whole process of the metaphor interpretator can be broken into 2 steps: 1. creation of the bags of features from the cor- pus and 2. interpreting metaphors using feature matching.

3.1 Creating Feature Sets

In order to create feature sets, we will use a cor- pus based approach. Using the GigaWord corpus, we will count how many times a word is "related" to another word. We will define two words as "related" when one word is the governor and the other word is the dependent in a dependency re- lation. Dependency parses of sentences are taken from the GigaWord corpus.

Next, we take the 100,000 most frequently oc- curing words that were tagged as nouns or verbs by the GigaWord Corpus. For each of these words, we find the 100 most "related" words, and add these as a words features. Along with every fea- ture, we give a numerical value determined by the number of times the relation between the two words occured divided by the number of times the non-feature word occured. This list of 100 related words will be our best guess at the feature set of the frequently occuring words in question.

This is how we build our feature sets.

3.2 Interpreting Metaphors

Once we have built our feature sets, we define the functions that will be used to interpret a metaphor.

In order to interpret a metaphor, we take the top k features in both of the considered words' feature sets. With each of these top features, we scan through the opposing word's feature set and find the opposing word's feature with the highest similarity. We then offer up this top feature as a possible interpretation of the metaphor, with a score determined by the value of the top feature multiplied by the similarity of the opposing word's feature. We then return the n highest or "best" interpretations.

This model is motivated by the availability heuristic. We only check the top k metaphors for any of the words associated with the metaphor since the human mind will only latch onto the first interpretation that presents itself in a given context.

Results

We ran three experiments, all using WordNet to calcu- late similarity and the NYTimes section of the Gi- gaword corpus as our corpus for feature extrac- tion. Test metaphors were taken from a small cor- pus of 84 copular metaphors and their interpreta- tions made by Carlos Roncero and Roberto G. de Almeida (2014). Three trials were run per copu- lar metaphor: one with the top 3 features of each word extracted, one with the top 5 features of each word extracted, and one with the top 12 features of each word extracted. Results were very disappointing. Precision and recall are shown below.

Trial	Pre.	Rec.
Top 3	0.03	0.02
Top 5	0.03	0.04
Top 12	0.01	0.08

Table 1: Results

Conclusion

The results of this experiment were not promising, but there are many reasons that the results could have been decreased through error. Further study into this project is required. While the results of this project may not be a large jumping off point for further study into new areas, it was a wealth of experience for me, and the experience gained here can certainly help me approach other areas in NLP research with faster execution time, more thorough planning, improved background research, and greater independence of design. While I clearly would've wanted better results, this project is very much a stepping stone that can be used to inform and direct new research, both through the practice of techniques that should be pursued and techniques that should not be.

Acknowledgement

This is a body of text that will repeat over and over until it fills up the page. It requires little thought but says a great deal about layout design.