About the Linguistic Phrasing project

1. Some mistakes have been found in json file .For example, the term '8:00 pm' occurs in the right where should be the person class .Another example comes that the template should be 'person verb person' but the truth is 'person verb noun'.
2. There is some confusion about the meaning of the phrase or word such as 'participate in show' ,'perform in show ' and 'star in show' .Eventually ,I set them as the meaning for the fundamental version of the relation sheet when considering the integrity of the relation.
3. Our extraction work or entity similarity construct could be better under the situation that some entity is still not accurate .For example ,the company name in our files is not complete .It only includes the term '.INC' or '.CO' .It just work perfect when the company name like 'Wal-Mart' which is a proper noun.

LSTM (Christopher Olah)

It first talks about the RNN .Recurrent neural networks allow information to persist with loops in them .Long-Term Dependencies can be describes as following: we need to look at recent information to perform the present task. But the gap between the relevant information and the point where it is needed to become very large .But the LSTM can figure it out.

Long Short Term Memory networks – usually just called “LSTMs” – are a special kind of RNN, capable of learning long-term dependencies. It includes:

1. Decide what information we’re going to throw away from the cell state.
2. Decide what new information we’re going to store in the cell state.
3. Update the old cell state into the new cell state.
4. Decide what we’re going to output.

Skip-Gram (CS224D)

It aims to predict surrounding context words when given a center work .This slide has also mentioned the CBOW compared the Skip-Gram model. The author invoke as Naïve Bayes assumption to break out the probabilities.

Ensemble Semantics for Large-scale Unsupervised Relation Extraction (2012)

Because our project is based on this paper, I reviewed it carefully and found the following 2 papers.

This paper present a fully unsupervised algorithm WEBER for large-scale open-domain relation extraction .It can be listed as these parts: problem analysis, entity similarity graph, hypernym graph, relation phrase similarity and 2 phases for relation extraction.

Phase 1: Discovering Type-A Relations

This phase aims to generate Type-A relations which have exactly one relation phrase and two argument entity semantic classes.

Phase 2: Discovering Type-B Relations

Merge similar Type-A relations which have a set of synonymous relation phrases and more complete argument entity classes.

Some details have not been mentioned in this paper, such as the 'ctx' preprocessing in the 1st phase. I have discussed this problem with Xi Chen, we need more source to test our algorithm quality and improve it.

DIRT-Discovery of Inference Rules from Text (Lin and Pantel 2001)

The former paper has mentioned this algorithm and it points out the difference between the WEBRE and DIRT. So I read this paper and try to figure it out.

It apply the Harris' Distributional Hypothesis to paths in the dependency trees of a parsed corpus. WEBRE use stemmed lexical sequences instead of dependency paths as phrase candidates because of the very large scale of the corpus.

It is the first attempt to dis cover such knowledge automatically from a large corpus of text. It has introduced the Extended Distributional Hypothesis, which states that paths in dependency trees have similar meaning s if they tend to connect similar sets of words.

Two questions remain to be addressed. One is to recognize the polarity in inference relationships. Another is to extend paths with constraints on the inference rule's variables.

Leveraging Linguistic Structure for Open Domain Information Extraction (Stanford 2013)

This paper have presented a system for extracting open domain relation triples by breaking a long sentence into short, coherent clauses, and then finding the maximally simple relation triples which are warranted given each of these clauses.

It first pre-process the sentence in linguistically motivated ways to produce coherent clauses.

This approach consists of two stages:

1. Learn a classifier for splitting a sentence into shorter utterances
2. Appeal to natural logic to maximally shorten these utterances while maintaining necessary context.