Making Syntactic Relations Recognizable

WordSeer Currently

Grammatical Relations Can be Searchable & Visualized

But:
- Menu labels for the dependency parses are unclear.
- Difficult for a non-expert to find the desired relation.

Principle:
Recognition over Recall

Question: how to show the context of syntactic relations in an understandable manner?

A crowdsourced recognizability experiment

We focused on the specific problem of recognizing English syntactic dependency relationships between words. Our question was: how can dependency relationships be represented in order to make them more recognizable?

Method

Amazon's Mechanical Turk

Non-expert users

Moby Dick by Melville, Stanford Dependency parser

Participants given a series of identification tasks:
- In each task, see a list of sentences containing a syntactic relationship between highlighted words.
- Asked to identify the relationship type from a list of 4 options.
- Options shown in 3 different ways; each participant saw the same presentation for all 12 tasks.

Payment: 50c (U.S.) + 50c bonus if they correctly identified 10 or more of the 12 relationships.

Used the 12 most common grammatical relationships in English syntactic dependency relationships between words. Our question was: how can dependency relationships be represented in order to make them more recognizable?

Hypothesis

Grammatical relations are identified more accurately when shown with examples of contextualizing words or phrases than without.

Builds on the success of auto-suggest in search interfaces.

Experiment Materials

Figure 1: A sample recognition task. Target word in yellow ("life"). 8 example sentences containing the relation were shown, with the words that entered into the relationship highlighted. Participants had to identify the relationship and correctly select it from a list of four options on the left.

The presentation styles:

Baseline.

Baseline + 4 example words.

Baseline + 4 example phrases.

Experiment Materials

Figure 2: Results – recognition success rate for different types of relations under the three presentations.

Implications

- A list of phrases is the most recognizable presentation (34% better than the baseline). However, there is room for improvement. Even the best strategy had a success rate of only 55%.
- Auto-suggest interfaces for syntactic search should show candidate relationships augmented with a list of phrases in which they occur.

Results: examples improve recognizability

400 participants completed the study distributed randomly over the 3 presentations.

The results confirm our hypothesis.

Participants in conditions that showed examples (phrases and words) were significantly more accurate at identifying the relations than participants in the baseline condition.

The average success rate in the baseline condition was 41%, which is significantly less accurate than words: 52%, (p=0.00019, W=6136), and phrases: 55%, (p=0.00014, W=5546).1

For the non-clausal relations, there was no significant difference between phrases and words, although they were both overall significantly better than the baseline.

Implications

- A list of phrases is the most recognizable presentation (34% better than the baseline). However, there is room for improvement. Even the best strategy had a success rate of only 55%.
- Auto-suggest interfaces for syntactic search should show candidate relationships augmented with a list of phrases in which they occur.

Aditi Muralidharan
aditi@cs.berkeley.edu
Marti A. Hearst
hearst@berkeley.edu