The Working Group Formerly Known as Dynamic Distributional Semantics at MIC

September 11, 2015

1 Motivation

Consider:

(1)   a.   The doctors are old.
      b.   The cucumbers are old.

We expect a good Distributional Semantic (DS) model to build sensible representations of these two sentences. For example, the model should be able to capture how the subject noun disambiguates the predicative adjective, so that (1-b) is very similar to *"The cucumbers are rotten"*, but (1-a) is most definitely not a paraphrase of *"The doctors are rotten."* Consider next:

(2)   a.   (i)   I saw the doctors...
             (ii)   I saw the cucumbers...
      b.   They are old.

The interpretation of (2-b) depends on whether it is a continuation of (2-a-i) or (2-a-ii). It is not clear that DS would be able to construct an appropriate representation of (2-b), e.g., one that is closer to *"They are rotten"* only if *they* refers to the cucumbers.

Making sense of (2-b) crucially depends on a notion of *reference*. Correctly interpreting *they* requires referring to some entities that are currently in the universe of discourse of the unfolding conversation, e.g., a specific set of cucumbers. Reference to discourse entities is actually also needed to make full sense of the sentences in (1) (*which* doctors or cucumbers are we referring to?), but, there, at least predicate disambiguation (assigning the right sense to *old*) can be performed by considering *generic* properties of doctors and

cucumbers (the latter are more likely to go bad than the former), that do not require looking at the specific subsets of entities we are talking about.

DS models, by extracting word meanings from large amounts of text, can be very effective at capturing generic properties, but, for the same reason (word representations are averages across many contexts), they are not well-equipped to handle phenomena depending on specific reference (see Section 2 for other phenomena of this sort).

Still, the generic knowledge encoded in DS might provide evidence that can help to solve specific reference ambiguities. Consider:

(3) a. The professors finally found the cucumbers. . .
   b. They were rotten.

It is easy to think of approaches that could harness DS to decide that *they* must refer to cucumbers. For this specific example, something as simple as measuring distributional similarity of *professor/rotten* vs. *cucumber/rotten* might suffice. Of course, most real-life cases of coreference resolution require a much more sophisticated exploitation of DS-based cues (and in many cases DS will not suffice), but at least the issues at hand are clear: What kind of DS models are best for the task? How should DS-produced scores be integrated in a larger coreference system? Etc.

We would like to use the opportunity offered by the symposium to instead focus most directly on the relation between DS and reference from the point of view of a general theory of DS. **That is, we are not going to discuss how to use DS evidence as a component in other systems that handle reference, but rather how to handle reference-related phenomena *within* DS.**

2 Examples

To make the working group discussion concrete, we propose a number of examples we would like you to think of, and we invite you to add more to the list.

Plain pronouns, where disambiguation really depends on context and not on generic conceptual properties.

(4) a. The teacher entered the classroom. . .
   b. He looked scary.

How do we process (4-b) in a DS framework such that, e.g., the fact is captured that it entails: *"The teacher looked scary"*?

The issue of how **current discourse** contributes to the representation of sentences with pronouns is also of interest, consider the following example from Bonnie:

(5)   a.   (i)   The professors were unable to find the cucumbers.
           (ii) They were too old. (i.e., because the professors were too old.)
      b.   (i)   The professors were able to find the cucumbers.
           (ii) They were too old. (i.e., but the cukes were too old.)

Definiteness.   In a DS framework, how do we deal with the difference between:

(6)   a.   The teacher entered the classroom.
      b.   A teacher entered the classroom.

How do we account for / model the fact that the meaning of these sentences changes if, instead of being uttered out of context, they are produced while speaking of the classroom where Mr. Smith teaches?

Proper nouns.   Assuming Mr. Smith (the teacher) is called John, how do we capture the relation between:

(7)   a.   A teacher entered the classroom.
      b.   The teacher entered the classroom.
      c.   John entered the classroom.

How do we capture the fact that, given that we know which John we are talking about, *John* in (7-c) carries presuppositions that are more like those of *"the teacher"* in (7-b) than those of *"a teacher"* in (7-a)?

Referentially identical: same or different vectors? *Sinn* or *Bedeutung?*   The following example is from Gemma Boleda's ongoing Marie Curie project on DS and reference:

(8)   a.   The Beatles broke up in 1970.
      b.   The best-selling band of the sixties broke up in 1970.

3   Questions

We would like you to think about the previous examples (and of course more that you come up with) while pondering the following questions:

1. How do we **represent the broader discourse context** and the entities it defines in DS?

2. How do we **build such discourse representations compositionally and dynamically**?

3. How do we **learn a reference-aware DS model** from the data? From **which** data?

4. How do we **represent sentences** such as those in these examples in DS?

5. Actually, is **representing sentences** the right approach or do alternatives exist in which our DS model represents the subpart of **discourse context that is necessary to interpret subsequent discourse**, but does not represent sentences as vectors?

Of course, we realize that a reasonable answer to these questions is that DS was never meant to handle reference, so that they should not even be asked. This is fair game for what we should discuss in the working group.