Linguistics 575:
Semantic Representations

April 22, 2013

Emily M. Bender

Overview

- f-structures to KR, via resource-sensitive rewrite rules (Crouch 2005, Crouch & King 2006)
- GMB (Basile 2012, Venhuizen et al 2013)
- Evaluation plans

Reading questions: Crouch 2005

- What is the relationship between packing and underspecification?

    - “Free choice packing is a generalization of the use of charts in context free parsing. Context-freeness guarantees that alternative analyses of disjoint word spans are independent and do not interact. Hence one can freely combine any analyses of disjoint spans; charts exploit this to compute all complete analyses in cubic time and quadratic space. For non-context free grammars (e.g. LFG), analyses of disjoint word spans are not always independent. But they are mostly independent, so that interactions are minimal.”

- Do LFG parses *not* have a context-free backbone? Does that mean that neither do those of the ERG? I never thought to ask because I just assumed that the ARGS < > branching was formally CFG.

- "Maxwell-Kaplan wager"?

Reading questions: Crouch 2005

- What do they mean by ‘chart approach’ to KR? What’s a ‘span’?
- What MRS concepts do the ‘contexts’ correspond to?
- What is skolemization, and why do it?
- Why is the counterfactual world part of the KR for *The technician prevented an accident*. (2)?
- Won’t leaving the ambiguity in the KR inhibit further processing?

Crouch 2005: Example

(5) Two terrorists intended to attack a factory.

```
A1: exists2(t. terrorist(t) &
     intend(t, exists(f. factory(f) & exists(a,attack(a,t,f)))))
A2: exists2(t. terrorist(t) & exists(f. factory(f) &
     intend(t,exists(a,attack(a,t,f)))))
A3: exists(f. factory(f) & exists2(t. terrorist(t) &
     intend(t,exists(a,attack(a,t,f)))))
```

(6)

A2 v A3:	ist(context3,factory(sk_factory2))
	ist(context0, terrorist(sk_terrorist1))
	ist(context0, cardinality(sk_terrorist1,2))
A1:	ist(context0,factory(sk_factory2))
A2 v A3:	ist(context3,factory(sk_factory2))
	ist(context0, intend(sk_terrorist1, context3))
	ist(context3, attack(sk_attack3, sk_terrorist1, sk_factory2))
	skfn(sk_terrorist1, context0)
A2:	sk_function(sk_factory2, sk_terrorist1)
A3:	sk_function(sk_factory2, context0)
A2 v A3:	sk_function(sk_factory2, context3)
	sk_function(context3, sk_terrorist1)
	sk_function(sk_attack3, context3)

Reading questions: Crouch 2005

- What is the ‘Curry Howard’ correspondence?

- “Rule ordering can be exploited in a number of useful ways (section 3), but it can sometimes be a nuisance. A way of marking blocks of rules that apply in parallel is due to be added to the system.”

- Why would it sometimes be a nuisance?

- Within (putative) parallel blocks, would there still be rule ordering or could this be a way to introduce declarativity into the rewriter in a controlled, incremental way?

Reading questions: Crouch & King 2006

- What's the difference to the system described in Crouch 2005?
- How is context_head() determined and what causes multiple context_head() expressions to be generated?
- What are the main differences between these f-structures from an LFG and MRSs from the ERG, and what would this mean for the different rewriting rules? I was wondering specifically about how deverbal nouns might be handled differently in MRS
- The rules are generated by people right? They're not somehow aggregated from a corpus, right?
- In fact, is there any automated way of generating the semantics?

Reading questions: Crouch & King 2006

- “Although the system described here could by no means be described as a theory of semantic construction or the syntax-semantic interface, from a practical stand point it can efficiently and robustly produce theoretically defensible semantic structures from broad-coverage syntactic ones.”

- What does 'robustly' mean? How is 'robustness' demonstrated in the paper? There are no evaluation metrics to compare the system against anything else or a hold-out data set that is tested to determine how well the system performs on unseen data, or even a test suite of sentences that one could claim coverage on.

- I may be just being too picky, but it seems that the XOR they use in (6) is too strong. It may seem unlikely, but it's not impossible that John is using a telescope to look at a man who is looking into a telescope.

Reading questions: Crouch & King 2006

- “This is an alternative to using type hierarchy for producing compact rule sets.”: Does the HPSG -> MSR translation make use of the type hierarchy?

- Regarding deverbal nouns: “The goal of the rules is to take the nominal and map it to its verbal counterpart.”: Are there some properties of nominal that are lost in this process? I would imagine. For example, I don't see how *the first burning of the whitehouse* can be translated into a verbal form: Where in the verbal frame would *first* fit?

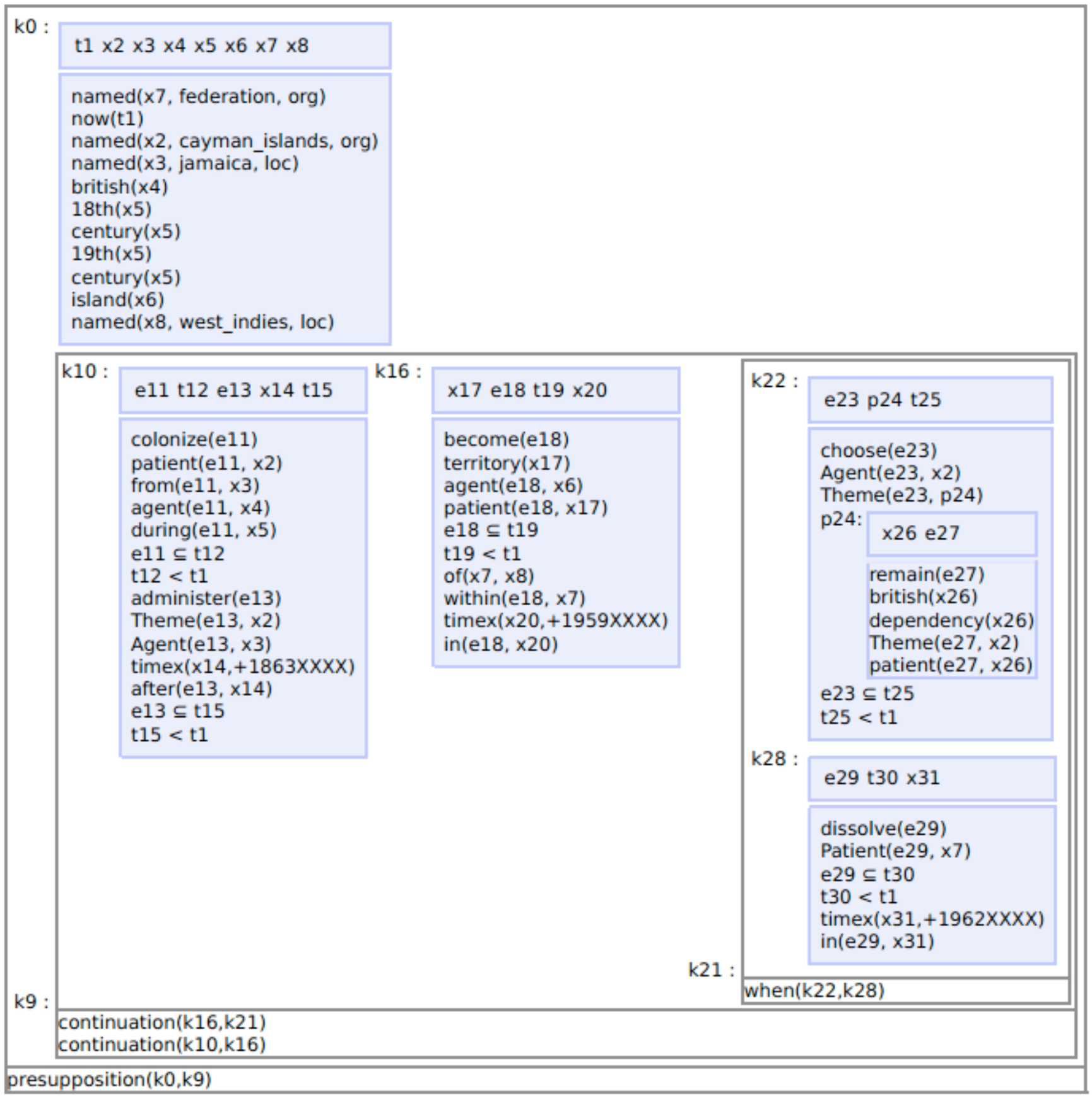
- What is Glue Semantics?

- How does recursion work in this system?

- What does “oblique” mean?

Example

(Basile et al 2012)



Reading Questions: Basile et al 2012

- What do ⊆, ⊇, ⊂, and ⊃ represent?
- The paper doesn't describe the online NLP pipeline in detail. Is the pipeline extensible? Could one use the output from an alternate syntactic parser? Could one simply upload pre-processed BOWs into the database for each judgement stage?
- Do they maintain provenances of the applied BOWs in their current-best system, so they can re-apply the exact decisions that were made if they introduce new or improved modules or techniques?
- The Bits of Wisdom seem confined to the immediate context for which they were determined. It seems for this to be scalable there'd need to be some sort of feedback mechanism into the actual grammar and/or POS tagger to incorporate the common/expert wisdom insofar as they diverge.

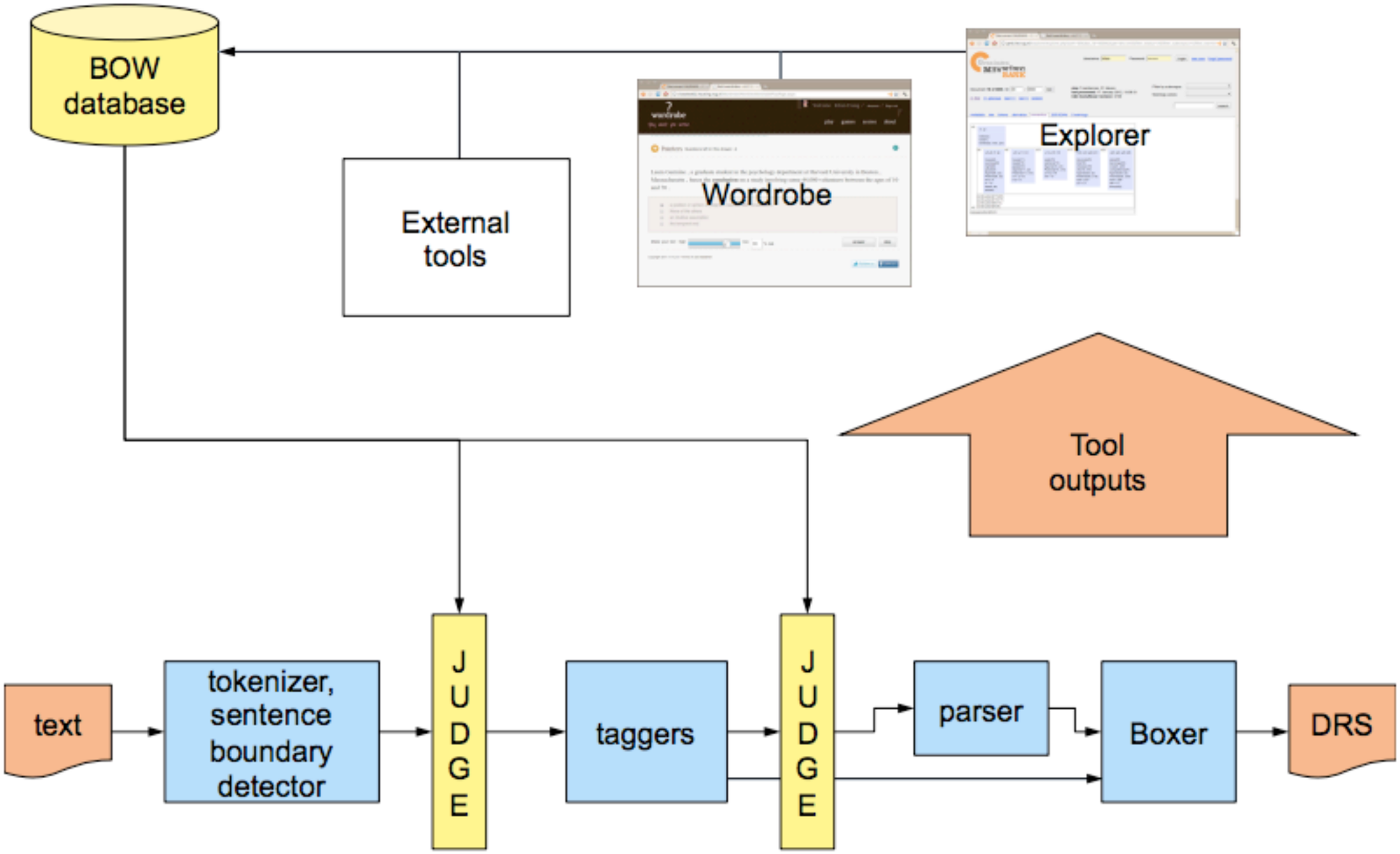


Figure 3: Graphical representation of the workflow for constructing the GMB.

(Basile et al 2012)

Reading Questions: Basile et al 2012

- Based on the workflow diagram, it looks like the expert and crowdsourced annotations can only affect the output of the parts of the pipeline *before* the syntactic parser and Boxer. Is there a way for experts and/or the crowdsourced game to offer insight on the syntax or semantics directly?

- In playing the wardrobe games it seems that some answers are very obvious and some are less obvious. It seems like obvious answers would not need so many non-expert opinions, but less obvious ones would require more opinions.

Reading Questions: Basile et al 2012

- How would the feedback loop work where the DRS output is used for intelligently creating new Wordrobe tasks?

- I am wondering about the differences between the CCG/DRT representations and an HPSG/MRS representation.

- I noticed that they said DRT is translatable into FOL, but is it currently available in FOL form, or are they expecting us to do the translation algorithm on our own?

Reading Questions: Venhuizen et al 2013

- I wasn't familiar with the terms "presupposition" (at least in this context) and "conventional implicature". After reading the paper I get the impression that a presupposition is a bit of information that is already assumed to be known in the discourse context but whose repetition is linguistically necessary to connect the meanings of various bits of the dialog together (most prominantly in the paper, definite descriptions and proper names). Hence, one can try to link their a semantic representation to the location where the information was originally presented.

- Conventional implicature, on the other hand, seems to refer to situations where *new* information is asserted in mid-discourse, which somehow *behaves* as though it were asserted somewhere else.

- Have I got this right? What is the real deal with these terms?

Reading Questions: Venhuizen et al 2013

- In the example "Someone did not see the cyclist", I do not understand why the "the cyclist" is a presupposition and "someone" is not a presupposition.

- Section 3.4.2 regarding Conventional Implicatures says that CIs and presuppositions accommodate to the global discourse context, but presupposition pointer is available for binding while the CI pointer refers to the most global context. I didn't understand the distinction.

- The author says projective DRT is a system where "all linguistic material is associated with a pointer to indicate its accommodation site", but I think they are using a specific interpretation of "linguistic material". Their examples leave off many lexemes present in the sentence, so perhaps they mean verbs and their core arguments? How do they decide what is "linguistic material"?

Reading Questions: Venhuizen et al 2013

- Do the pointers for all presuppositions point to a free variable? Or would something like a pronoun point to the PDRS that asserts its antecedent?

- In example 4c, does the pointer pointing to the PDRS labeled "1" refer to the entire PDRS labeled "2", or just the negation portion?

- What does it mean to bind a presupposition?

- Do free pointers always point at the outermost PDRS, regardless of how deeply the pointer is nested (if it may be nested)?

- What is the difference between global accommodation and local accommodation?

Reading Questions: Venhuizen et al 2013

- In section 3.1 is talking about free variables and how it is a way to reference the top most global PDRS and that in figure 4 (a) and (b) are equivalent. (c) seems to suggest that the top most global layer would always be 1 (or whatever label you chose), so why bother with f (free variable) at all? What would be a situation where the top most global layer might be ambiguous? Is it simply a shorthand code so you don't have to type out the top most label (1 or otherwise) all the time?

- I'm a little fuzzy on how they merge the pointers. For example, "I did not see the cyclist. Mary didn't either." How does that get parsed/represented semantically? (Because "Marry didn't [see the cyclist] either." And the representation should cover the fact that both cyclists are the same presupposed cyclist, but the "see" actions are different.)

Reading Questions: Venhuizen et al 2013

- From 3.4.1 lexical semantics:

- Where is the line drawn between lexical items which do and don't add presuppositional content?

- Or, in a sort of high-level, philosophical way: doesn't every word carry with it a set of association, relations, and other specific content which can connect with the context in varied and unique ways?

- This article seems to be dealing with only a very limited kind of presupposition, the case of *the*. I wonder about other lexical items which are (sometimes) analyzed as providing presuppositions, such as *even* or non-intersective adjectives like *fake* and *former*.

Reading Questions: Venhuizen et al 2013

- What linguistic phenomena are judged by the crowdsourcing component of the GMB annotation lifecycle? Are PDRT judgements mapped to questions in one of the games?

- Finally, what can be done with a DRT that incorporates this projection information?

- I guess my whole discussion here is to pose the question about the NLP community and what is the current common thought in within it, as I am unaware. Is it to lean toward having the correct representation even if it means increased complexity or is it make the representations simpler even if it means odd shortcuts so that people can achieve it / do something with it now? Perhaps somewhere in the middle?

Overview

- f-structures to KR, via resource-sensitive rewrite rules (Crouch 2005, Crouch & King 2006)
- GMB (Basile 2012, Venhuizen et al 2013)
- Evaluation plans