Project sketch 1: Incrementality in compositional distributional semantics
Distributional semantics

• Traditional focus: Word meaning
  – Words = Vectors
  – Meaning of word: Statistics about its context

• Recent strand of work (2008--): Extension to phrase (and sentence) level
  – Method does not generalize straightforwardly: Sparsity!

• Fregean compositionality: “Meaning of phrase is a function of the meaning of its constituents and their manner of combination”
Compositional distributional semantics

• Fregean Compositionality
  – “Meaning of phrase is a function of the meaning of its constituents and their manner of combination”

• Wide range of models
  – Most of them fairly *ad hoc*
    • Linguistic limitations
    • Concentration on verb-argument or noun-modifier combinations, no “bigger picture”
Baroni et al. (under review)

• A proper syntax-semantics interface for compositional distributional semantics
  – Functional application-based approach
  – Strong parallel to model-theoretic semantics:

<table>
<thead>
<tr>
<th>Model-theoretic semantics</th>
<th>Compositional distributional semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>lambda calculus</td>
<td>linear algebra</td>
</tr>
<tr>
<td>functional application</td>
<td>matrix multiplication</td>
</tr>
<tr>
<td>lower- and higher-order types: nouns are sets, modifiers are</td>
<td>vectors and tensors: nouns are vectors, modifiers are</td>
</tr>
<tr>
<td>functions from sets to sets</td>
<td>matrices (linear transformations)</td>
</tr>
<tr>
<td>types determined by CG</td>
<td>types determined by CCG</td>
</tr>
</tbody>
</table>
Example
Example: Types
Project goals

1. Investigate an *incremental* version of this framework
   – Incrementality: central property of language processing – SFB topic!
   – Needs ability to deal with incomplete input
2. Develop a notion of *expectation* within it
   – What semantic concept are we likely to see next?
3. Use to predict reading times
   – Either on whole-sentence eye tracking data or on a specific phenomenon (to be determined)

• “Spiritual successor” of D6: Combination of modeling + cognitive component
Step 1: Incremental distributional compositional semantics

• Syntax side: Incremental CCG [Hefny et al. 2011, Demberg 2012, ...]
  – Typically involves some form of type raising
    • E.g. NP → (S/(S\NP)) – Syntactic expectations
    • Must be mirrored by a corresponding “lifting” of semantic objects

• Semantics side: Consequences of type raising?
  – How to compute higher-order tensors?

• Parameter: “Granularity” of incrementality
  – Word level: higher-order types, high-dim. tensors
  – Chunk level: lower-order types, low-dim. tensors
Step 2: Expectations

• How to compute: Open question...
• Previous work: ICE (Brew & McDonald 2002), Selectional Preference models (e.g. Erk et al. 2010, Lenci 2011), D6 work (Utt et al. 2012)
• Necesssary: Translation into new framework
Step 3: Prediction of reading times

- Existing work falls into two categories:
  1. Predicate-argument plausibility (e.g. Pado et al. 2006)
  2. Syntactic/n-gram-based “surprisal” (e.g. Levy 2008)
- Compositional distributional semantics fills the gap:
  - Compared to (1): Covers contributions of wider range of constructions (modification, adverbs, ...)
  - Compared to (2): Uses a “proper” semantic model
- Question: Choice of materials
  - Whole-sentence eye tracking: attractive, but large syntactic component?
  - Individual phenomena (modification?)
  - Also related to size of project
Language

• German or English
• Tendency towards German
  – Whole-sentence reading times: only English datasets available..
    • Schilling corpus, Dundee corpus
  – ..on the other hand, German-speaking participants are much easier to get in Stuttgart :)
• Modeling part possible in both languages
  – SFB collaboration considerations...?