



The Holy Grail of Sense Definition

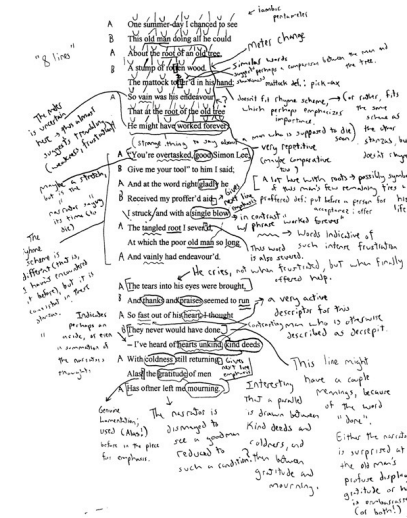
Anna Rumshisky, Marc Verhagen and Jessica L. Moszkowicz

Part Deux

Application to a
Semantic Annotation Task

Using the Sense Clusters

- This approach may help other semantic annotation efforts
- The case of GLML annotation
- Very very preliminary (just a plan really)



GLML Annotation

- Procedure for annotating argument selection and coercion
 - Pustejovsky et al., 2009 (IWCS-8).
GLML: Annotating argument selection and coercion.
 - Pustejovsky and Rumshisky, 2009 (NAACL HLT 2009).
SemEval- 2010 Task 7: Argument Selection and Coercion.
 - The aim was to identify the nature of the compositional relation rather than merely annotating surface types of entities involved in argument selection.
 - (a) *Mary* called yesterday
(b) *The Boston office* called yesterday

Two Modes of Composition

- Pure selection
 - the type a function requires is directly satisfied by the argument
 - *Mary* called yesterday
- Type coercion
 - the type a function requires is imposed on the argument by exploitation or introduction
 - *The Boston office* called yesterday

Annotation

- Three tasks
 - *verb-based annotation*
 - qualia in modification structures
 - type selection in modification of dot objects
- All three tasks involve two stages
 - a data preparation phase with selection of annotation domain and *construction of sense inventories*
 - *the annotation phase.*

Verb-based Annotation

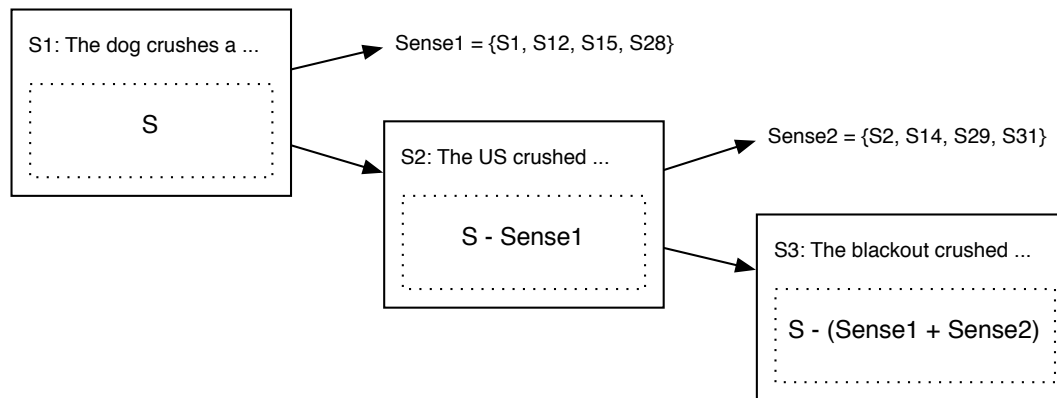
- Data preparation
 - Select target set of verbs
 - Create sense inventory for each verb
 - Assign type templates for each sense
 - The child *threw* a stone
[PERSON throw PHYSICAL-OBJECT]
 - Automatically extract sentences with the target verb with its arguments
- Annotation
 - Select verb sense
 - Does the argument match the type from the template?
 - If not, what type is it?

Problems?

- Creation of sense inventory
 - Avoid theoretical bias
 - P&R used Corpus Pattern Analysis
- Shallow type system
 - Informs template creation and selection of actual types of arguments
 - Unclear how to pick the types
 - Choice of types determines whether there is coercion

Adapting GLML Annotation

- What do we have after the sense definition procedure outlined above?



- A cluster of senses
- For each sense, a set of sentences that illustrate the usage of the sense

Using the Senses

- The sense clusters define the sense inventory called for in the data preparation phase of the argument selection task.
- The clusters also trivially provide for the sense selection, the first step of the annotation phase

Using the Examples

- The exemplars can guide a trained linguist in creating the templates
 - For each sense, the linguist/lexicographer needs to distinguish between instances of pure selection and instances of type coercion
 - For each set, a succinct way to describe the type of the argument needs to be defined
- The exemplars can also inform the annotator

Conclusion

- We took the specialist out of the construction of the GLML sense inventory
- However, it is still up to the expert to analyze and process the clusters defined by Mturkers
- “We are currently experimenting with this adapted GLML annotation procedure”

