Joint Satisfaction of Syntactic and Pragmatic Constraints Improves Incremental Spoken Language Understanding

Andreas Peldszus, Okko Buß, Timo Baumann & David Schlangen

Abstract

We present a model of semantic processing of spoken language that

- is **robust** against ill-formed input, such as can be expected from automatic speech recognisers,
- respects both **syntactic and pragmatic** constraints in the computation of most likely interpretations,
- uses a **principled, expressive semantic representation formalism** (RMRS) with a well-defined model theory, and
- works **continuously** (producing meaning representations on a word-by-word basis, rather than only for full utterances) and **incrementally** (computing only the additional contribution by the new word, rather than re-computing for the whole utterance-so-far).

We show that the joint satisfaction of syntactic and pragmatic constraints **improves** the performance of the NLU component (around **10% absolute**, over a syntax-only baseline).

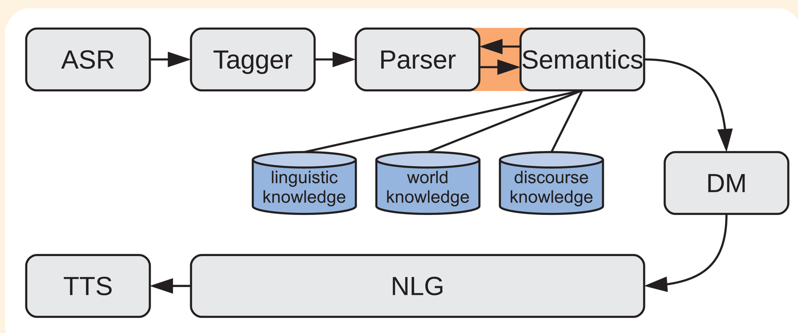
1. Introduction

Motivation for Incremental Processing in SDSs

- psycholinguistic plausibility
- descriptive adequacy (backchannel feedback, intervening corrections, clarifications, completions etc.)
- computational benefits (by making lower-level processing results earlier available to higher-level modules and by providing higher-level feedback to lower-level modules)

Aim

Investigate the potential advantage of **pragmatic feedback to the parser**, specifically of reranking syntactic readings according to their success in real-world reference.



2. The Model

Parser

- A variant of Roark's (2001) parser: incremental, probabilistic, beam-searching top-down parser.
- Monotonic growth of fully connected syntactic structures. Left factorisation delays syntactic decisions.
- Robust lexical operations (insertions, deletions and repairs) for ASR input.

Semantic Construction

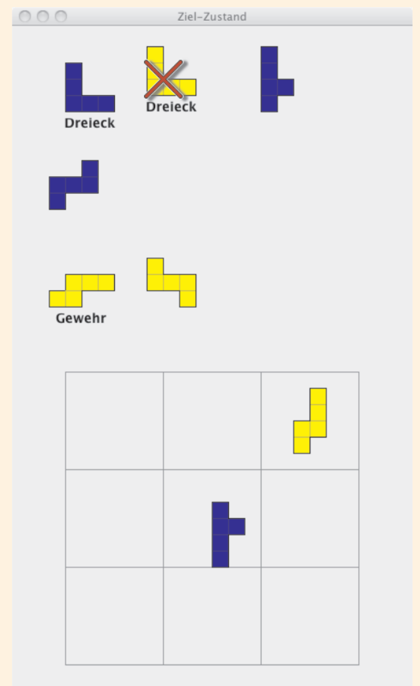
- Combination is linearised in a **top-down left-to-right** fashion (instead of bottom-up).
- No need to define semantics of open projected nodes, no re-interpretation of existing parts of the tree.
- **Robust Minimal Recursion Semantics** (Copestake, 2007) as semantic formalism, adapted to the chosen combination order.
- Semantic construction **synchronised** with syntactic expansion: monotonically, continuously and incrementally.

Reference Feedback

- Test each representation for it's semantic plausibility, depending on whether definite NPs refer uniquely, ambiguously or fail to refer to pieces in the corresponding world state.
- Degrade unplausible readings (lower probability in the next parsing step).

The model is implemented in the InproTK (Schlangen et al., 2010).

3. Experiment Data



- Corpus of Pentmino puzzle games (see Fernández & Schlangen, 2007), collected in a WOz-study with 284 games of 20 participants
- Users instruct the wizard to manipulate different puzzle pieces in order to reach a specified goal state.
- Subset of 1026 utterances with reliable semantics (extracted from the wizard's next action) & without pronouns referring to pieces (no discourse model) in gold transcript and in ASR output version
- **Natural spontaneous speech** representative for interactions of such a domain

- "das blaue ähm das teil links oben neben dem winkel"
  the blue-one ehm the piece left top next-to the angle — regular complex ambiguities
- "genau und jetzt nochmal drehen"
  right and now again rotate — elliptical constructions
- "löschen unten"
  delete bottom — non-standard constructions

Grammar

- Small handcrafted core grammar (30 rules), easy to engineer
- Weights set according to intuition

4. Evaluation

Baselines & Settings

Evaluation of semantic accuracy by comparing the extracted "gold" semantic alignment with the resolved reference of ...

- **Just Syntax (JS):** the single-best derivation of syntax only
- **External Filtering (EF):** of the 5 best syntax-only derivations, the best referring one
- **Syntax/Pragmatic Interaction (SPI):** the single-best reference-feedback derivation
- **Combined Interaction & Filtering (CIF):** of the 5 best reference-feedback derivations, the best referring one

Incremental Score Metric

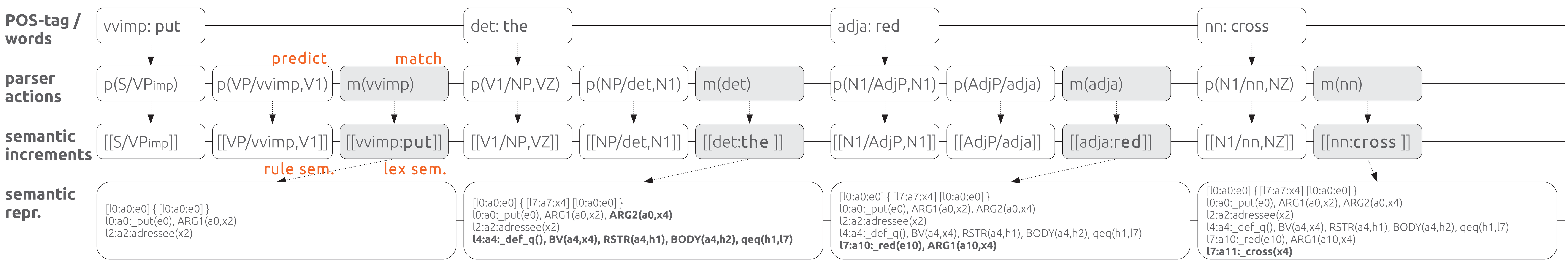
- A measure of how the resolved reference matches over time, with increasingly stronger influence of later changes

5. Results

With reference feedback:

- Less mismatches (-1)
- More partial matches (0)
- More unique matches (1)
- 12.5% abs. improvement for strict accuracy and 19.3% for relaxed accuracy (allowing partial matches) for transcript SPI over JS
- Effect not only at the final word, but within the utterance (incremental score)
- Similar but slightly smaller improvements for transcript EF over SPI
- Similar but slightly smaller improvements for ASR input
- No further improvement by additional filtering (SPI vs CIF)
- Low baseline due to the complexity of authentic data (see examples)

		JS	EF	SPI	CIF
transcript	−1	563	518	364	363
	0	197	198	267	268
	1	264	308	392	392
	str.acc.	25.7 %	30.0 %	38.2 %	38.2 %
	rel.acc.	44.9 %	49.3 %	64.2 %	64.3 %
	incr.scr	−1568	−1248	−536	−504
	avg.incr.scr	−1.52	−1.22	−0.52	−0.49
recognition	−1	362	348	254	255
	0	122	121	173	173
	1	143	158	196	195
	str.acc.	22.6 %	25.0 %	31.0 %	30.8 %
	rel.acc.	41.2 %	44.1 %	58.3 %	58.1 %
	incr.scr	−1906	−1730	−1105	−1076
	avg.incr.scr	−1.86	−1.69	−1.01	−1.05



Copestake. 2007. Semantic composition with (robust) minimal recursion semantics. In Proceedings of the Workshop on Deep Linguistic Processing, DeepLP '07, pages 73–80, Stroudsburg, PA, USA. Association for Computational Linguistics. — **Fernández & Schlangen. 2007.** Referring under restricted interactivity conditions. In Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue, pages 136–139, Antwerp, Belgium. — **Roark. 2001.** Robust Probabilistic Predictive Syntactic Processing: Motivations, Models, and Applications. Ph.D. thesis, Department of Cognitive and Linguistics, Brown University — **Schlangen & Skantze. 2009.** A general, abstract model of incremental dialogue processing. In EACL '09: Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics, pages 710–718. Association for Computational Linguistics. — **Schlangen, Baumann, Buschmeier, Buß, Kopp, Skantze & Yaghoubzadeh. 2010.** Middleware for Incremental Processing in Conversational Agents. In Proceedings of SigDial 2010, Tokyo, Japan. — **Stoness, Allen, Aist & Swift. 2005.** Using real-world reference to improve spoken language understanding. In AAAI Workshop on Spoken Language Understanding, pages 38–45.