Embodied Semantics
Towards a cognitive grounding of formal semantics

Tillmann Pross

Institute for Computational Linguistics/Institute of Philosophy, University of Stuttgart

TUM,17.04.2009
Outline

1. Computational Linguistics and Robotics?

2. Example: the proper processing of events
A starting point

- Acceptance of robotic assistance crucially depends on the possibility of natural interaction between man and machine.

  'Natural':
  
  - Multimodal, Realtime, Goal-directed, . . .
  - In general: Aspects of spatiotemporal embodiedness of both cognitive and bodily abilities.

  ⇒ Grounding of higher mental functions connects Computational Linguistics (CL) and Robotics.
State of the art, roughly

CL and Robotics figuratively approach each other:

- CL: Top-Down to the model theory of formal NL-semantics (e.g. CandC Tools: CCG+DRT).
- Robotics: Bottom-Up to behaviour-based pragmatics (e.g. BDI-based planning).
Computational Linguistics and Robotics?
Example: the proper processing of events

State of the art, figuratively

What exactly is the point of contact between CL and Robotics?

Robotics (Embodied)

Computational Linguistics (Disembodied)
Objective: Connect CL and Robotics

→ Extend the sensomotoric capabilities of a robot with the ability to construct, maintain and manipulate complex symbolic representations and corresponding models
← Ground complex symbolic representations and corresponding models in the sensomotoric embodiedness of a robot
⇒ How to achieve this?
Events

- The proper processing of events constitutes the core mechanism of natural human-machine interaction.
- E.g., events carry the pragmatic force of speech acts by means of their semantic structure.
- A proper account to events must consider both semantics and pragmatics of events.

⇒ Does the traditional approach to events in CL fulfill these requirements?
Davidson [1967]: Introduce a new ontological class of entities besides individuals: events.

E.g. 'x build a house':
\[ \exists e. \exists x. \exists y. \text{agent}(x) \land \text{house}(y) \land \text{build}(e, x, y) \]
Vendler Classes

- Vendler [1957]: different verbs can have very different ‘temporal profiles’ in that they are used to describe very different event complexes:
- E.g. ’build a house’ refers to a process of construction that brings about a house
- E.g. ’reach the top’ refers to a punctual event
Moens and Steedman [1988] capture Vendler’s observation by the introduction of a subatomic structure of events:

- Event Nucleus := preparatory state, culmination, consequent state

E.g. ‘build a house’ := process of construction, topping-out ceremony, existence of the house
Simple-minded DRT

In Discourse Representation Theory (Kamp et al. [2007])

Meaning Postulate 1:

\[
\begin{align*}
  & x, y, e \\
  & e : \text{build}(x, y) \\
  & \text{house}(y) \\
  \Rightarrow & \quad s^{res} \\
  & e : \text{build}(x, y) \\
  & (s^{res}) \\
  & s^{res} : \text{house}(y)
\end{align*}
\]

Meaning Postulate 2:

\[
\begin{align*}
  & x, y, e \\
  & e : \text{build}(x, y) \\
  & \text{house}(y) \\
  \Rightarrow & \quad s^{prep} \\
  & e : \text{build}(x, y) \\
  & s^{prep} \subseteq e \\
  & s^{prep} : \neg \text{house}(y)
\end{align*}
\]
Given a set of events and states \( EV \) structured by \(<\), a universe of individuals \( U \) and an interpretation function \( I \),

\[
\text{if} \quad g \models_M e : R(x_1, \ldots, x_n) \quad \text{iff} \quad \langle g(e), g(x_1), \ldots, g(x_n) \rangle \in I(R)
\]

where \( g \) is an assignment that maps \( e \) onto an element of \( EV \) and \( x_1, \ldots, x_n \) onto elements of \( U \).

For ’x build a house’: Events described by occurrences of ’build a house’ are events that stand in some ’build’-relation to the one who is doing the building (or the ones who are doing the building) and the thing that is built.
Some consequent Problems

- This semantics does not identify (Searle [1969]) the building of a house as an action but as a relation → No pragmatic dimension of meaning.
- Where do $U, I, EV, g$ come from? → Requires cognitive grounding, vs. the purely structural nature of Tarski-Models
- How can a robot draw any information from such a semantics about an appropriate understanding of what it means to build a house?
  → 'Blind alley' with respect to the desiderata of natural human-robot interaction.
Improve on the model theory

⇒ Provide the robot with mechanisms to construct and maintain $U, I, EV, g$
How do humans do that?

Given a perception of unsegmented temporal variation:

- Psychology: Humans structure a perceived temporal variation along the lines of plan-goal relationships and causal structures. (Zacks et al. [2001], Zacks and Swallow [2007])

- Philosophy: Causal, behavioural and intentional explanation of temporal variation. (E.g. Dretske [1988], Dennett [1989], Hartmann and Janich [1991])

  - E.g. ‘x is building a house’: ascription of an underlying intention to x to predict x’s behaviour.

⇒ Conservative transfer of these insights to a DRT-like setup.
Grounding

- Sensomotoric grounding: Output of a perpetual flow of snapshots from the object recognition.
- Behavioural grounding: Output from a BDI-based multi-agent planner.
- Combined grounding results in a branching time model, where the past and present is anchored by sensomotoric grounding and the future is anchored in behavioural grounding.
Grounded branching-time Models

Example: the proper processing of events
Dynamic online models

- Combined grounding can be formalized with a modal Kripke semantics/CTL.
- But: the model must incorporate a notion of dynamics resulting from the permanent temporal move of the present and consequent revisions of the future (vs. the timelessness and offline construction of Tarski-Models).

⇒ Dynamic 'online' model of an agent x at time t: \( M(x)(t) \)

- How must the semantic representations of events look like to match such dynamic models?

⇒ Dovetail semantic representations and model structures.
Event anchoring

  - Involved agents
  - The explanation type that identifies/constructs the event
  - Causes/Goals/Intended states of affairs
  - Tense

- Pragmatic specification of the event $\Rightarrow$ Course of action corresponding to the event
  - Causal Chains, Plans, Intentions, other action-related information
  - Specifies a partial structure of the model (!)
E.g. ‘x is building a house’:

\[
\begin{array}{c}
\text{SEM} \\
x, e, n \\
\langle e, x\text{INT} \langle y, \text{house}(y) \rangle \langle y, ? \rangle \rangle \\
\text{PRG} \quad t_0 - \text{get-materials} - t_1 - \text{lay-bricks} - t_2 - \text{make-roof} - t_3
\end{array}
\]
Embodied event semantics comprises both the semantic and pragmatic dimension of events by means of anchoring.

- **Semantic**: truth-conditional structure embedding (word-to-world fit).
- **Pragmatic**: model manipulations via the adoption of new goals, beliefs or the execution of actions (world-to-word fit) specified by the semantic structure of the representation.

⇒ Reciprocal influence of semantics and pragmatics.
⇒ Computation of event meaning via incremental anchor resolution.
⇒ Leads to a notion of dynamic realtime interpretation.
The connection between CL and Robotics

⇒ Grounding of higher mental functions connects Computational Linguistics (CL) and Robotics.
Thank you for your attention.

- Discussion
- Questions


References II


References III
