Streamlining Formal Construction Grammar[1]

Vanja Sophie Cangalovic, Robert Porzel & John A. Bateman
University of Bremen

Keywords: Parsing, Formalisation, Language Understanding

Streaming Construction Grammar (SCG) is both a Construction Grammar formalism (Fillmore 1988; Goldberg 1995) as well as a parser implementation for simulation-based natural language understanding. It was developed in the CRC EASE, where it transforms underspecified natural language instructions from the kitchen domain into interpretations of an ontological model (Beßler et al. 2020). SCG draws inspiration from Embodied Construction Grammar and Fluid Construction Grammar, combining constructions as mental simulation-parameterising operations tightly integrating external knowledge bases, with the computational model of unification and merging (Feldman et al. 2009; Steels et al. 2006).

However, SCG is uniquely simplified and generalised. Normalising the classical attribute-value-matrices into triples, also known as *entity-attribute-value tuples*, and using them for input, output, and internal data allows for a fundamental reengineering of the match-merge operation. Most computational work is pushed to an input-independent pre-computation step, so that parsing itself is reduced to unification with constraint propagation and the weighted set cover problem, opening new avenues for performance improvements through indices and heuristics. This also yields reactive incremental parsing capabilities and the ability to simultaneously explore and combine different interpretation hypotheses, similar to chart-based approaches.

A novel query operator between the match and merge operations allows constructions to access a shared monotonic knowledge base, i.e. a tuplespace that is based on a common data representation, defined by an ontology and its constraints. Using a common knowledge framework for the grammar, the attached knowledge base, and the robotic systems allows for seamless communication and enables realtime bi-directional data streams between them. As data, such as input words or world information, becomes available in the shared knowledge base, it is picked up by reactive queries and, in turn, causes constructions to fire. Moreover, queries using knowledge from an ontology, a simulation, or robotic sensors can:

- act as a success criterion, e.g. by expressing role filler type constraints, such as those in Embodied Construction Grammar
- enable context-aware constructions
- facilitate the processing of syntactically ambiguous instructions, e.g. distinguishing CAUSEDMOTION from PROPERTYRESULTATIVE
- resolve coreferences via both syntactic heuristics as well as semantic constraints (Raghuram et al. 2017)

Conversely, schemas evoked by the instruction can parameterise subsequent simulations which are able to determine unknown parameter values and the overall plausibility of the instruction interpretations. The implementation also provides a notebook-based sandbox with tools and visualisations to aid grammar engineering and inspect the parsing process.

[1] The research reported in this paper has been supported by the FET-Open Project 951846 "MUHAI -- Meaning and Understanding for Human-centric AI'' funded by the EU Program Horizon 2020 as well as the German Research Foundation DFG, as part of Collaborative Research Center 1320 "EASE -- Everyday Activity Science and Engineering'', University of Bremen. The research was conducted in subproject H02.

References

Beßler, Daniel, Robert Porzel, Mihai Pomarlan, Abhijit Vyas, Sebastian Höffner, Michael Beetz, Rainer Malaka & John A. Bateman. 2020. Foundations of the Socio-physical Model of Activities (SOMA) for Autonomous Robotic Agents.

Feldman, J., Ellen Dodge & John Bryant. 2009. A Neural Theory of Language and Embodied Construction Grammar. In B. Heine & H. Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*, 111-138.

Fillmore, C. 1988. The Mechanisms of "Construction Grammar". *Fourteenth Annual Meeting of the Berkeley Linguistics Society*. 35-55.

Goldberg, A. E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure.* Chicago: University of Chicago press.

Raghuram, V., Trott, S., Shen, K., Goldberg, E.M., & Oderberg, S. 2017. Semantically-Driven Coreference Resolution with Embodied Construction Grammar. *AAAI Spring Symposia*.

Steels, L. & De Beule, J. 2006. Unify and Merge in Fluid Construction Grammar. *Third international conference on Emergence and Evolution of Linguistic Communication: Symbol Grounding and Beyond (EELC'06)*. 197–223.