COMPUTATIONAL CONSTRUCTION GRAMMAR and NATURAL LANGUAGE UNDERSTANDING

AAAI Spring Symposium, Stanford, March 27-29 2017.
http://www.aaai.org/Symposia/Spring/sss17.php
https://ai.vub.ac.be/aaai-construction-grammar-2017/

DESCRIPTION

TOPIC: Computational construction grammar has been emerging as a new line of research into language modeling that promises to give us advances in natural language processing/understanding with applications to web querying, intelligent tutoring, translation, summarization, language grounding in robots, etc. Natural Language Understanding (NLU) entails systems that take action without human intervention and remains intractable with current template or statistical approaches.

The characteristic features of the construction grammar approach and their relevance to NLP/NLU include:
1. Tight integration of morpho-syntax and semantics/pragmatics - supporting better semantic parsing and language production as well as cleaner handling of non-standard language such as reduced constructions and idioms.
2. Smooth integration of many different linguistic perspectives (phrase structure, functional and dependency structure, frame semantics, information structure, etc.) - so that constraints from all these perspectives can be applied as early as possible in language processing and learning.
3. Bi-directionality, in the sense of using the same architecture and the same language model for comprehension and formulation - supporting faster learning, a more compact representation of language knowledge, self-monitoring in language production, and better prediction in language comprehension.
4. Closer fit with human language processing and learning - allowing enhanced contact with the psycholinguistic community and leading to more intuitive language interfaces to autonomous systems.

Although there is an active international community working on construction grammars, the impact on AI and NLP/NLU has so far been limited due to three main factors: (i) any computational implementation of construction grammar is a significant problem, involving sophisticated programming and advanced AI techniques - much more complex than syntactic parsing (ii) broad coverage constructicons and corpora annotated from the viewpoint of construction grammar are still rare, and (iii) no well-organized research community pursuing computational construction grammar exists. Nevertheless, the first usable computational formalisms and implementations developed in research groups have started to emerge and some of the efforts in building constructicons and constructional corpora are beginning to bear fruit. There are now moderate scale

construction-based systems that demonstrate full NLU for autonomous systems like robots.

GOALS: The primary goals of this AAAI symposium are (i) to draw the attention of the AI community to the challenging technical issues and opportunities that the constructional perspective allows, (ii) to nurture an emergent community of computational construction grammar developers that is well integrated within the broader AI community, (iii) to compare existing implementations, understand open challenges, exchange technical solutions, and build up user communities, (iv) to standardize emerging corpora and establish challenges and evaluation criteria, similar to those successfully driving other areas of AI, (v) to survey the application potential of construction grammar in NLP/NLU applications.

AREAS

The symposium will address different aspects of computational construction grammar research. There will be sessions for each of these, raising the following questions:

1. Theory: The symposium should clarify what the constructional perspective entails, in a way that is informative to AI researchers. How is this perspective different from other approaches to language? Why is this perspective enthusiastically embraced by researchers engaged in empirical grammar writing, child language acquisition, diachronic linguistics, and language education? What psycholinguistic evidence is available to indicate that human language knowledge is structured in terms of constructions?
=> This session is steered by linguists active in construction grammar. They can present short tutorial introductions, state-of-the art surveys, and open problems in the foundations of construction grammar.

2. Formalisms for construction grammar: There have been a number of long-term efforts towards formalization of construction grammar and implementation of constructional processing (e.g. Sign-based Construction Grammar, Template Construction Grammar, Fluid Construction Grammar, Embodied Construction Grammar). How are these formalisms expressed? What are the computational mechanisms on which they rest? What kind of semantics do they interface with? How are problems such as managing combinatorial search, robustness, or flexible parsing being tackled? How is the complexity of writing large-scale grammars being supported? What interfaces exist for analyzing language processing? What is the current state of implementation? What are open technical problems? What grammar subsets have already been developed? Can we agree on common evaluation criteria to compare the scope, efficiency, and robustness of construction grammar implementations? Can we agree on a shared set of challenging cases to demonstrate the versatility of construction grammars and compare them to other formalisms such as semantic systems based on dependency grammars?

=> This section will form the main core of the symposium. It is organized by system developers and grammar engineers. They present short tutorials and live demonstrations of their systems, and participants are given opportunities for hands-on usage.

3. Natural Language Understanding (NLU): Natural Language Understanding has been a long-standing goal of AI, and many related fields, but is often dismissed as intractable. Natural Language Understanding (NLU) entails systems that take action without human intervention. This inherently involves strong semantic (meaning) capabilities. A system failing to understand a query causes no serious damage, but an error by a robot or automated vehicle could be disastrous. Because constructions integrate form and meaning, they constitute a potentially ground-breaking approach to NLU. For communication with autonomous systems such as robots, an action-oriented embodied semantics seems especially promising. Some moderate scale NLU systems of this kind have been demonstrated. One question is whether there are better semantic formalisms for this purpose. Another is whether the embodied semantics could work equally well with more conventional dependency grammar. Also, the construction grammar approach should be compared with other research on NLU.
=> This session is organized by people working on NLU from both constructions and other approaches. They can present short tutorials and give live demonstrations of their systems as well as general questions like the habitability problem.

4. Constructicons and corpora annotated for construction grammar. There are several efforts on the way for the development of constructional corpora and constructicons (e.g. the Berkeley Framenet Constructicon, Japanese FrameNet, Swedish Constructicon(SweCxn), German Constructicon, Brazilian Portuguese constructicon, etc.). What is the status and availability of these resources? What problems have been encountered in building them? In how far can constructional corpora be standardized? Can constructicons be bootstrapped by using or enhancing already existing broad coverage grammars? What tools have been developed to browse very large constructicons? How can we foster more effort in this area?
=> This session is steered by developers of constructicons and corpora. They present short tutorials and give live demonstrations of their systems, showing how access can be gained. General issues such as standardization are discussed in a panel.

5. Construction grammar learning and adaptation. The number of constructions known by a normal language speaker is estimated to be between half a million and one million. Can compositional structure significantly reduce the number of constructions required? If not then realistic constructicons can only be developed through machine learning. In any event, all NLP/NLU systems need to be adaptable to differing goals and contexts. Efforts in this direction are still in a very early stage. We find two approaches: statistical machine learning approaches using symbolic or neural techniques, and insight grammar learning approaches that acquire constructions incrementally and embedded in situated embodied interactions. What is the state of the art in the machine learning of construction grammar learning and adaptation? Which machine learning methods have

been tried and what are the results? Do new machine learning techniques need to be developed? Can we develop common benchmarks against which progress can be measured? Are there any particular advantages or disadvantages in learning language and adaptation from a constructional point of view?
=> This session is expected to be steered by machine learning researchers focusing on constructional learning. It will primarily be structured around posters of concrete experiments and panel discussion by poster presenters.

5. Applications. Construction grammars are currently being applied already for building intelligent tutoring systems, language grounding in robots, and automated translation, although these applications efforts are so far still in a very early stage of development. What are the motivations for using the constructional perspective for these applications? To what extent have they been developed? Why do developers believe that they may eventually go beyond the capacities of current more mature NLP systems?
=> This session is to be steered by application developers. It will primarily be structured around posters and demonstrations of applications, and a panel discussion by poster presenters.

6. Community organization. Although we begin to see special theme sessions in some conferences, such as at the International Construction Grammar Conference (http://www.ufjf.br/iccg9/home/theme-sessions/computational-semantics-with-frames-and-constructions/), there are at the moment no regular workshops nor conferences on computational construction grammar. Is there now enough critical mass to create a community organization dedicated to computational construction grammar? How can we optimally interface with the rest of the AI community? How can we stimulate a better exchange of data, software, documentation, and papers?
=> This session will take the form of group discussion, animated by a panel of researchers implicated in computational construction grammar.

AREAS OF INTEREST

This symposium is primarily targeting researchers in natural language processing/understanding and computational linguistics, but will have broader appeal to the whole of the AI community, specifically researchers in knowledge representation, man-machine interaction, and machine learning. The interest for a wide range of AI applications (from intelligent robots to web querying) is obvious.

SYMPOSIUM FORMAT

The symposium will be highly technical in nature, emphasizing demonstrations of working systems and opportunities to obtain hands-on experience and sufficient information to allow users/developers to assess systems and possibly start to use them. Each session is

structured around a series of short (10 minutes) presentations, posters further developing the presentations, live demonstrations, and software installation and hands-on testing.

COMMITTEE

Conveners

1. Luc Steels steels@arti.vub.ac.be
Institut de Biologia Evolutiva, Universitat Pompeu Fabra - CSIC
Doctor Aiguader, 88. 08003 Barcelona - Spain

2. Jerome Feldman feldman@icsi.berkeley.edu
International Computer Science Institute
1947 Center St. Suite 600
Berkeley, CA 94704-1198 - USA

Members

3. Adele Goldberg adele@Princeton.EDU
Psychology of Language Lab
Princeton University - USA

4. Katrien Beuls katrien@arti.vub.ac.be
Artificial Intelligence Lab. Vrije Universiteit Brussel
Pleinlaan 2 1050 Brussels, Belgium.

5. Nancy Chang nchang@gmail.com
Machine Intelligence Group. Google Research.
Mountain View Ca, USA

BIOs of committee:

Luc Steels studied linguistics at the University of Antwerp and computer science, specifically Artificial Intelligence, at MIT. In 1983 he became founding director of the Artificial Intelligence Laboratory at the University of Brussels (VUB) where he worked on different areas of AI, ranging from knowledge-level expert systems and knowledge representation systems to artificial life and behavior-based robotics. In 1996 he became the founding director of the Sony Computer Science Laboratory in Paris where he focused on language, specifically how language can originate and evolve in populations of artificial agents, including robotic agents. As part of this work he designed Fluid Construction Grammar in 2003. Currently Steels leads the Language Evolution Lab at the Institute for Evolutionary Biology in Barcelona. Relevant book publications include: '(ed.) Design patterns in Fluid Construction Grammar, John Benjamins Pub, 2011' and `(ed.) Computational issues in Fluid Construction Grammar, Springer Verlag, 2012'.

Jerome Feldman is Professor of Electrical Engineering and Computer Science at UC Berkeley and a member of the Institute for Cognitive and Brain Sciences. From 1988 to 1998 he was the director of the International Computer Science Institute, where he remains the head of the AI group. Feldman has a long standing interest in computational construction grammar and the neural basis of language. He initiated research on Embodied Construction Grammar at ICSI and is still one of its leading architects. Relevant publications include: `From Molecule to Metaphor: A Neural Theory of Language, MIT Press'.

Adele Goldberg is a linguist well known for her pioneering work in Construction Grammar. Since 2004, she has been a full professor at Princeton university in linguistics and psychology, after receiving her PhD at the University of California (Berkeley) and working as associate professor of linguistics at the University of California (San Diego). Goldberg's research focus is on the psychology of language, including theoretical and experimental aspects of grammar and its representation, acquisition of form-function correspondences, and syntactic priming. Her works aim to illuminate parallels between language and other cognitive processes. Relevant publications include: 'Constructions at Work: the nature of generalization in language, Oxford University Press, 2006' and 'Constructions: A construction grammar approach to argument structure. University of Chicago Press, 1995'.

Katrien Beuls is currently assistant professor at the Computer Science Department of the University of Brussels (VUB) where she leads a group on computational construction grammar. She is one of the leading developers of Fluid Construction Grammar, extending the formalism to deal with complex morphology and phonology. Her work has focused on the application of construction grammar in intelligent tutoring systems, the bootstrapping of a Spanish constructicon using machine learning techniques, and multi-agent models of ongoing language change. Relevant publications include: 'Towards an agent-based tutoring system for Spanish verb conjugation, VUB Press, 2013.'

Nancy Chang is currently a Google engineer at the Machine Intelligence group. She has a long-term interest in constructional learning, particularly the modeling of child language acquisition. She was one of the co-developers of Embodied Construction Grammar and Relevant publications include: `Constructing grammar: A computational model of the emergence of early constructions. UC Berkeley, 2008.