

Sentence Generation: Input, Algorithms and Applications

Claire Gardent

CNRS/LORIA, Nancy (France)

Abstract

(Joint work with Paul Bedaride, Eric Kow, Shashi Narayan and Laura Perez-Beltrachini)

Sentence Generation maps abstract linguistic representations into sentences. A necessary part of any natural language generation system, sentence generation has also recently received increasing attention in applications such as transfer based machine translation (cf. the LOGON project) and natural language interfaces to knowledge bases (e.g., to verbalise, to author and/or to query ontologies).

One outstanding issue in Sentence Generation is what it starts from. What is the abstract linguistic representation it generates from? In my talk, I will explore sentence generation from two main input formats (flat semantic formulae and dependency structures) and discuss their impact on efficiency, algorithms and applications.

I will start by describing an algorithm that generates from flat semantic formulae, explain why it is computationally intractable and presenting ways of optimising it to make it usable in practice. I will then show how this algorithm can be used to generate paraphrases; to support error mining and to generate teaching material for language learners from an ontology.

In the second part of the talk, I will focus on generation from dependency structures. Based on the input data recently made available by the Generation Challenges Surface Realisation Shared Task, I will show how the algorithm previously used to generate from flat semantic formulae can be adapted to generate from dependency structures. I will moreover discuss various issues raised by the GenChal data such as, missing lexical entries and mismatches between dependency and grammar structures.

Bio of Claire Gardent

Claire Gardent is a senior researcher at the French National Center for Scientific Research (CNRS). She graduated in linguistics at the University of Toulouse in 1986, received an MSc in Artificial Intelligence from the University of Essex in 1987 and defended a PhD in Cognitive Science at the University of Edinburgh in 1991. From 1991 to 2000, she worked as a researcher at the Universities of Utrecht and Amsterdam (The Netherlands), Clermont-Ferrand and Sarrebruecken (Germany). Since 2001 she has been working for the CNRS at the Lorraine Laboratory for Research in Computer Science (LORIA) in Nancy, France.

Claire Gardent's research focuses on the computational treatment of natural language meaning. She has worked on the automatic acquisition of lexical resources for French, on syntactic parsing and semantic role labelling and on text generation. Recently, she has become interested in exploring the interaction between virtual worlds and natural language processing.

Claire Gardent has published a textbook on analysis and generation (with Karine Baschung) and about 100 articles in journals and conference proceedings. She has been nominated Chair of the European Chapter for the Association of Computational Linguistics (EACL), editor in chief of the french journal "Traitement Automatique des Langues" and member of the editorial board of the journals "Computational Linguistics", "Journal of Semantics". Each year she is on the programme committee of half a dozen international conferences or workshops, she also acted as scientific chair for various international conferences (EACL), workshops (TAG+, ENLG, DIALOR, SIGDIAL) and summer schools (ESSLI).