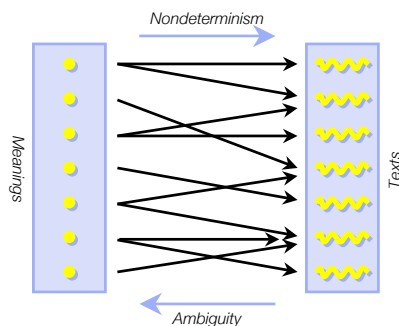


## Wide-coverage generation

For a given meaning, a wide-coverage realiser can produce multiple texts; for a given text, there may be a multiple meanings from which it could be generated. Viewed from the perspective of generation, this is nondeterminism, viewed from the perspective of analysis, it is ambiguity. The fundamental question which this project addresses is, assuming this is a 'closed' system (that is, no additional context is available, not represented in the input representations), to what extent can or should nondeterminism be constrained during the generation process by consideration of the potential for ambiguity in the output.



## Reflective generation

We use the term *reflective generation* to refer to a generator which reflects on its own behaviour, evaluating a goal state by comparing it both with other states achieving the same goal and with other possible goals which result in the same state. At the most global level, a reflective generator may evaluate a given text by comparing it with other texts from the same meaning, and with other meanings which generate the same text. The first aim of the project is to build a generation platform which supports this kind of reflection, through the use of (a) a bidirectional wide-coverage grammar, the *LinGO English Resource Grammar*, so that the relationships between meanings and texts can be computed in both directions and (b) a generate-and-test architecture using reflective filters to select among competing realisations of a given meaning. Subsequently we intend to explore more fine-grained reflective control, intercepting the LinGO generator's interactions with its chart, so that internal generation states can be evaluated against competing alternatives in the same way.



## Controlled Generation of Text

The COGENT project is developing a platform for the systematic study of **wide-coverage realisation** and using it to address three key issues for large-scale generation:

**Nondeterminism** – what is the set of outputs that can be generated from a given input, what form of input expands to the 'right' set of outputs, and how can the expansion be constrained;

**Ambiguity** – what is the set of inputs that can generate a given output, how can that set be predicted or constrained during generation, and how does ambiguity of an output affect its suitability;

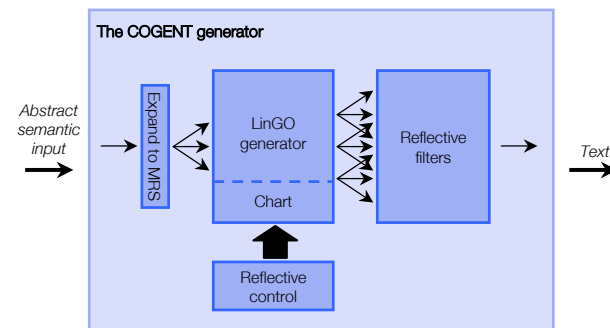
**Efficiency** – how can the search among potentially millions of alternative solutions be effectively controlled.

ITRI  University of Brighton



NATURAL LANGUAGE & COMPUTATIONAL LINGUISTICS  
DEPARTMENT OF INFORMATICS, UNIVERSITY OF SUSSEX

Research supported by 



## Current work

In the first year of this three-year project, we are developing the initial software framework to support systematic study of wide-coverage generation. This framework has three main components:

**An abstract semantic input language.** We are developing a formalism to support fine-grained underspecification of semantic inputs, which will allow researchers to explore the consequences of different aspects of underspecification in a uniform well-defined formal framework.

**A platform for reflective generation.** Built around the LinGO generator, the platform provides support for (a) nondeterministically mapping from the abstract input language to concrete MRS (Minimal Recursion Semantics) (b) incorporating reflective modules, which take the set of output texts generated from an input and make decisions about their relative suitability as candidate realisations.

**A semantic input corpus for the realiser.** We aim to develop a corpus of around 100,000 abstract semantic inputs, all realisable using the framework. We will achieve this by using the LinGO parser to parse existing corpora, and then automatically abstract away from the resulting MRS structures to create underspecified semantic inputs.

### The COGENT team

Investigators: John Carroll, Roger Evans, Kees van Deemter, David Weir  
Researchers: Anja Belz, Daniel Paiva  
PhD students: Eva Esteve Ferrer, Jason Teeple

[cogent@itri.brighton.ac.uk](mailto:cogent@itri.brighton.ac.uk)  
[www.itri.brighton.ac.uk/cogent](http://www.itri.brighton.ac.uk/cogent)