

## Knowledge-Based Natural-Language Processing

László Kálmán<sup>1</sup>, László Balázs<sup>2</sup>, and Miklós Erdélyi Szabó<sup>3</sup>

<sup>1</sup> Applied Logic Lab, Budapest

Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest  
Theoretical Linguistics Programme, Loránd Eötvös University, Budapest  
kalman@nytud.hu

<sup>2</sup> Applied Logic Lab, Budapest  
bazsi@all.hu

<sup>3</sup> Applied Logic Lab, Budapest  
Alfréd Rényi Institute for Mathematics, Hungarian Academy of Sciences, Budapest  
mszabo@renyi.hu

**Keywords:** computational linguistics, syntax, semantics, construction grammar, knowledge-based approaches

Computational linguistics' enthusiasm for the formal models of syntax and semantics introduced by Chomsky and Montague in the 70's declined during the 80's and 90's. It turned out that an autonomous syntax, without access to semantics, or an autonomous semantics, without access to pragmatics and world knowledge, are hopelessly untractable computationally and not suitable for applications.

The alternative that we propose is that modules are needed, but they must correspond to different types of linguistic knowledge and different modes of operation (such as statistical, associative and logical modes) rather than different linguistic 'levels'. The statistical modules perform jobs like speech recognition and synthesis, associative modules take care of invoking forms, structures and meanings from memory, and logical modules perform reasoning in understanding and planning tasks.

The main ingredient of our system is abduction: both understanding and producing utterances requires us to complete an array of pieces of information about the formal and semantic aspects of an utterance. All sources of information, including linguistic, pragmatic and cognitive sources can be used in order to achieve this goal. This is the key procedure that our system performs.

Linguistic knowledge itself is stored in a format that does not draw a sharp line between linguistic, pragmatic and cognitive aspects of signs. The framework that makes this possible is called construction grammar. Construction grammar is about systematic (conventional) associations of form types and meaning types, and information about the probabilities of their co-occurrence. The starting-point of the abduction process in our system is a loosely connected network of 'activised' constructions.