4. NATURAL LANGUAGE PROCESSING

GÁBOR PRÓSZÉKY, Professor; BORBÁLA SIKLÓSI, Assistant Professor; LÁSZLÓ LAKI, PhD; ATTILA NOVÁK, PhD

PhD Students: ISTVÁN ENDRÉDY, BALÁZS INDIG, GYŐZŐ ZIJIAN YANG

SHORT DESCRIPTION OF THE ACTIVITIES

The Natural Language Processing Group consists of faculty researchers, post-graduate researchers, PhD students, undergraduate students and programmers who work together developing algorithms that enable computers to process and understand human languages. Our research interest covers:

- corpus linguistic applications
- statistical machine translation
- syntactic parsing
- medical text mining
- morphologies
- spelling correction
- part-of-speech tagging

One of the most ambitious aims of the research group is to develop new methods and algorithms for syntactic parsing of the Hungarian language. Such a method must handle grammatically possible, but not correct analyses. That is why to deal with problems efficiently, parallelism is necessary. In practice for a human understanding is cooperatively done by several parts of the brain. To incorporate this knowledge the consideration of current state of the neurolinguistics and psycholinguistics is indispensable. The model that is to be researched is characterized by performance, while the state-of-the-art research results are considered from various field of applied linguistics. Since the current state of research does not provide any deeper understanding of how the ambiguous phrases are understood, in our project we incorporate parallel corpora to handle these—not necessarily multilingual—problems. With this, new aspects of corpus linguistic research are being revealed. The developed new methods are adapted to many aspects of Hungarian language. For using it for medical text processing tasks, we expect growing performance and precision. The developed algorithms are also planned to be adapted to other agglutinating languages and are expected to behave similarly well.
SELECTED PUBLICATIONS


