

CyMON

Natural Language Processing

Natural Language Processing

The **CyMON Natural Language Processing Engine** is based on a shallow parsing approach using both, statistical and deterministic mechanisms in order to analyze textual input. To each sentence, it will return an appropriate *contextualized meaning representation*.

Application Areas

CyMON-NLU can be customized to fit different approaches to Customer Relationship Management: chatting, informing, data mining and queries to product databases. They all can be employed in vortals and e-business web sites.

Knowledge Base

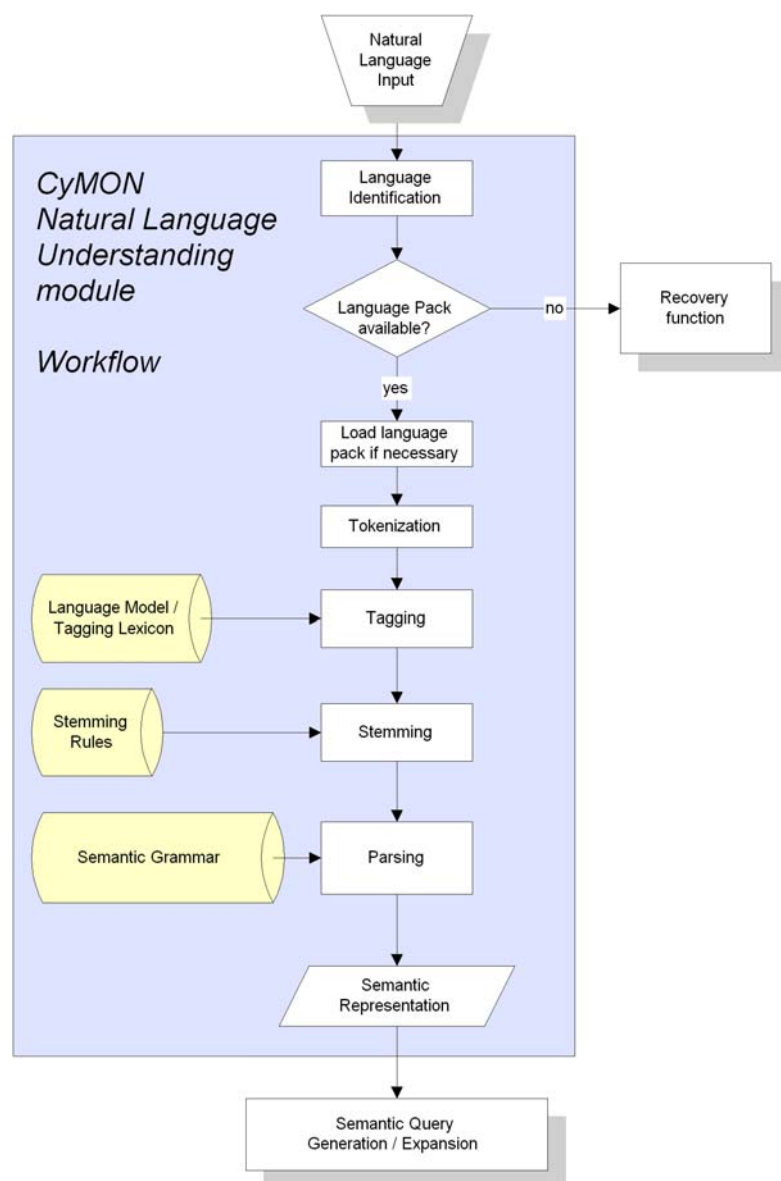
A linguistic knowledge base module providing extensive chat functionality is included in the CyMON distribution. Its structure allows for customization of the CyMON-NLU to specific support-related needs. Additional language packages are also available. Agentscape offers full support for all customization needs.



Shallow parsing algorithms and statistical methods recognize the input. Even if a phrase is unknown, such as "sofas", the system will answer appropriately.

Semantic Grammars

Input is parsed using a two-level semantic phrase-structure grammar. The first level of rules is an ordered set, assigning priority to the rules. The second level of rules deals with semantically simple terms without priority information. Using the two-level approach, the risk of syntactic and semantic ambiguities is eliminated. The Grammar rules also open up ways to track discourse topics. They can be related to topics in the CyMON semantic knowledge network module.



Various steps ensure that the textual input is understood correctly. The components provide internal recovery methods in order to process unexpected input and simulate human behavior.

Statistical Disambiguation Technology

In addition to disambiguation by priority rules, CyMON-NLU contains a morpho-syntactic disambiguation scheme using statistical tagging methods. The tagging results are fully integrated in the semantic grammar. The tagger can be trained for all tag sets and for a variety of languages that allow for normalization to an eight bit char set. Language Models and dictionaries for German and English are currently available from Agentscape.

Stemming tailored to every Language

For highly inflective languages like German, a stemming algorithm can be used from the grammar. This normalizes inflected forms in order to cover semantically similar derived forms with just one entry in the semantic grammar.

Semantic Representations for Low-Cost Localization

Semantic representations can be built compositionally and non-compositionally at both levels of grammar rules. They can reference unknown, textual input in order to implement a pattern-matching based dialog system. Semantic representations are not language-specific. The localization of interaction rules, which reference dialogs as reactions to a natural language input, is a matter of a translator's work. Databases may be kept centralized.



X^{treme}
Personalisation



How can you contact us?

Address

Agentscape AG
Bülowstraße 66
10783 Berlin, Germany

Fon +49-30-59 00 478-0
Fax +49-30-59 00 478-99

Contact

<http://www.agentscape.de>
info@agentscape.de