

Treebanks

Building and Using Parsed Corpora

Linguists and engineers in Natural Language Processing tend to use electronic corpora more and more. Most research has long been limited to raw (unannotated) texts or to tagged texts (annotated with parts of speech only), but these approaches suffer from a word by word perspective. A new line of research involves corpora with richer annotations such as clauses and major constituents, grammatical functions and dependency links. The first parsed corpora were the English Lancaster treebank and Penn Treebank. New ones have recently been developed for other languages.

This book:

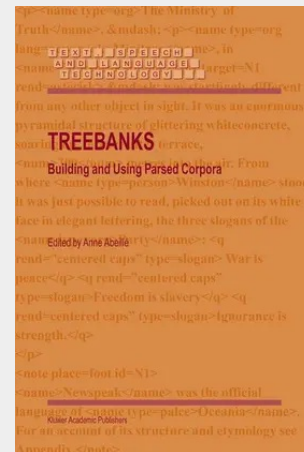
provides a state of the art on work being done with parsed corpora;

gathers 21 papers on building and using parsed corpora raising many relevant questions;

deals with a variety of languages and a variety of corpora;

is for those working in linguistics, computational linguistics, natural language, syntax, and grammar.

Linguists and engineers in Natural Language Processing tend to use electronic corpora more and more. Most research has long been limited to raw (unannotated) texts or to tagged texts (annotated with parts of speech only), but these approaches suffer from a word by word perspective. A new line of research involves corpora with richer annotations such as clauses and major constituents, grammatical functions and dependency links. The first parsed corpora were the English Lancaster treebank and Penn Treebank. New ones have recently been developed for other languages. This book: provides a state of the art on work being done with parsed corpora; gathers 21 papers on building and using parsed corpora raising many relevant questions; deals with a variety of languages and a variety of corpora; is for those working in linguistics, computational linguistics, natural language, syntax, and grammar.



160,49 €

149,99 € (zzgl. MwSt.)

Lieferfrist: bis zu 10 Tage

Artikelnummer: 9781402013348

Medium: Buch

ISBN: 978-1-4020-1334-8

Verlag: Springer

Erscheinungstermin: 30.09.2003

Sprache(n): Englisch

Auflage: 2003

Serie: Text, Speech and Language Technology

Produktform: Gebunden

Gewicht: 1750 g

Seiten: 407

Format (B x H): 160 x 241 mm

