

The MIT Press

Journals

Sign In / Register

Books

Journals

Digital

Resources

About

Contact

Home | Computational Linguistics | List Article navigation of Issues | Volume 27 , No. 1 | Bootstrapping Morphological Analyzers by Combining Human Elicitation and Machine Learning



Bootstrapping Morphological Analyzers by Combining Human Elicitation and Machine Learning

Kemal Ofazer, Sergei Nirenburg and Marjorie McShane

Posted Online March 13, 2006
<https://doi.org/10.1162/089120101300346804>

© 2001 Association for Computational Linguistics

Computational Linguistics
Volume 27 | Issue 1 | March 2001
p.59-85

Download Options

Abstract **Authors**

Journal Resources

- Editorial Info
- Abstracting and Indexing
- Release Schedule
- Advertising Info

Author Resources

- Submission Guidelines
- Publication Agreement
- Author Reprints

Reader Resources

- Rights and Permissions
- Most Read

Most Cited

More About Computational Linguistics ▼

Metrics ▼



14 Total citations

1 Recent citation

2.81 Field Citation Ratio

n/a Relative Citation Ratio

Open Access ▼



Computational Linguistics Computational Linguistics is Open Access. All content is freely available in electronic format (Full text HTML, PDF, and PDF Plus) to readers across the globe. All articles are published under a [CC BY-NC-ND 4.0 license](#). For more information on allowed

This paper presents a semiautomatic technique for developing broad-coverage finite-state morphological analyzers for use in natural language processing applications. It consists of three components – elicitation of linguistic information from humans, a machine learning bootstrapping scheme, and a testing environment. The three components are applied iteratively until a threshold of output quality is attained. The initial application of this technique is for the morphology of low-density languages in the context of the Expedition project at NMSU Computing Research Laboratory. This elicit-build-test technique compiles lexical and inflectional information elicited from a human into a finite-state transducer lexicon and combines this with a sequence of morphographemic rewrite rules that is induced using transformation-based learning from the elicited examples. The resulting morphological analyzer is then tested against a test set, and any corrections are fed back into the learning procedure, which then builds an improved analyzer.

Forthcoming

Most Read

[See More](#)

Lexicon-Based Methods for Sentiment Analysis (14087 times)
Maite Taboada et al.
Computational Linguistics
Volume: 37, Issue: 2, pp. 267-307

Computational Linguistics and Deep Learning (10542 times)
Christopher D. Manning
Computational Linguistics
Volume: 41, Issue: 4, pp. 701-707


Near-Synonymy and Lexical Choice (3675 times)
Philip Edmonds et al.
Computational Linguistics
Volume: 28, Issue: 2, pp. 105-144


(Note that the Most Read numbers are based on the number of full text downloads over the last 12 months.)


Most Cited

[See More](#)

uses, please view the CC license. [Support OA at MITP](#)

 **Lexicon-Based Methods for Sentiment Analysis** (436 times)
Maite Taboada et al.
Computational Linguistics
Volume: 37, Issue: 2, pp. 267-307



 **A Systematic Comparison of Various Statistical Alignment Models** (174 times)
Franz Josef Och et al.
Computational Linguistics
Volume: 29, Issue: 1, pp. 19-51

 **Opinion Word Expansion and Target Extraction through Double Propagation** (147 times)
Guang Qiu et al.
Computational Linguistics
Volume: 37, Issue: 1, pp. 9-27

(Note that the Most Cited numbers are based on Crossref's [Cited-by service](#) and reflect citation information for the past 24 months.)

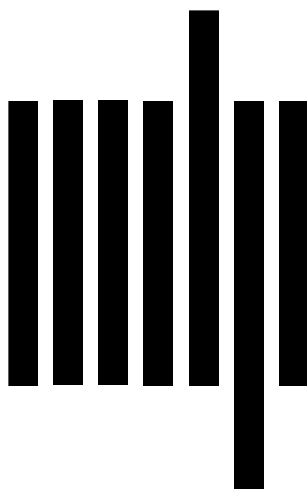
 **Download Options** >

Favorite  Sign up for Alerts 

Download Citation  RSS TOC 

RSS Citation  Submit your article

[Support OA at MITP](#) 



Journals

Terms & Conditions

Privacy Statement

Contact Us

Books

US

UK

Connect

One Rogers Street
Cambridge MA
02142-1209

Suite 2, 1 Duchess
Street London,
W1W 6AN, UK



© 2018 The MIT Press
Technology Partner:
[Atypon Systems, Inc.](#)
[CrossRef Member](#)
[COUNTER Member](#)
The MIT Press colophon is registered in the

U.S. Patent and Trademark Office

