

# The MIT Press

## Journals

[Sign In / Register](#)
[Books](#)
[Journals](#)
[Digital](#)
[Resources](#)
[About](#)
[Contact](#)


Home | Computational Linguistics | List Article navigation  
of Issues | Volume 34 , No. 1 | Feature  
Forest Models for Probabilistic HPSG  
Parsing



# Feature Forest Models for Probabilistic HPSG Parsing

Yusuke Miyao and Jun'ichi  
Tsuji

Quarterly (March,  
June, September,  
December)

160pp. per issue

6 3/4 x 10

Founded: 1974

2018 Impact

Factor: 1.319

2018 Google

Scholar h5-index:

32

ISSN: 0891-2017

E-ISSN: 1530-9312

Posted Online March 07, 2008

<https://doi.org/10.1162/coli.2008.34.1.35>

© 2008 Massachusetts Institute of Technology

Computational Linguistics

Volume 34 | Issue 1 | March 2008

p.35-80

 **Download Options** >

**Abstract Authors**

**Journal  
Resources**

Editorial Info

Abstracting and  
Indexing

Release Schedule

Advertising Info

Probabilistic modeling of lexicalized grammars is difficult because these grammars exploit complicated data structures, such as typed feature structures. This prevents us from applying common methods of probabilistic modeling in which a complete structure is divided into sub-structures under the assumption of statistical independence among sub-structures. For example, part-of-speech tagging of a sentence is decomposed into tagging of

## Author Resources




Submission Guidelines  
Publication Agreement  
Author Reprints

## Reader Resources

Rights and Permissions  
Most Read  
Most Cited

More About Computational Linguistics ∨

Metrics ∨

 Patents (1)  
 Wikipedia (1)  
 Mendeley (72)



**69** Total citations

**11** Recent citations

**11** Field Citation Ratio

**n/a** Relative Citation Ratio

Open Access ∨

each word, and CFG parsing is split into applications of CFG rules. These methods have relied on the structure of the target problem, namely lattices or trees, and cannot be applied to graph structures including typed feature structures.


This article proposes the feature forest model as a solution to the problem of probabilistic modeling of complex data structures including typed feature structures. The feature forest model provides a method for probabilistic modeling without the independence assumption when probabilistic events are represented with feature forests. Feature forests are generic data structures that represent ambiguous trees in a packed forest structure. Feature forest models are maximum entropy models defined over feature forests. A dynamic programming algorithm is proposed for maximum entropy estimation without unpacking feature forests. Thus probabilistic modeling of any data structures is possible when they are represented by feature forests.


This article also describes methods for representing HPSG syntactic structures and predicate-argument structures with feature forests. Hence, we describe a complete strategy for developing probabilistic models for HPSG parsing. The effectiveness of the proposed methods is empirically evaluated through parsing experiments on the Penn Treebank, and the promise of applicability to parsing of real-world sentences is discussed.


## Forthcoming

### Most Read

[See More](#)

 **Lexicon-Based Methods for Sentiment Analysis** (14057 times)  
Maite Taboada et al.  
Computational Linguistics  
Volume: 37, Issue: 2, pp. 267-307

 **Computational Linguistics and Deep Learning** (10535 times)  
Christopher D. Manning  
Computational Linguistics  
Volume: 41, Issue: 4, pp. 701-707

 **Near-Synonymy and Lexical Choice** (3670 times)  
Philip Edmonds et al.  
Computational Linguistics  
Volume: 28, Issue: 2, pp. 105-144



**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 uses, please view the [CC license](#).

[Support OA at MITP](#)

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

### Most Cited

[See More](#)

**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. )

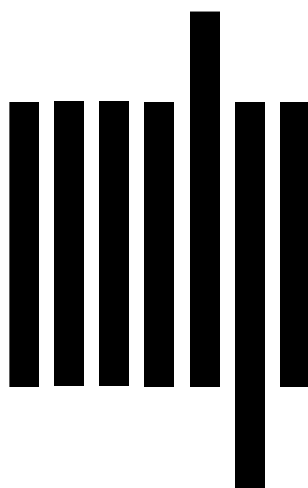
### Download > Options

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](#)

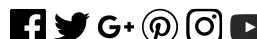
[US](#)

[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.  
[Site Help](#)