

CAMBRIDGE

JOURNALS

Journal of Functional Programming

Editors-in-Chief

Matthias Felleisen, *Northeastern University, USA*
Jeremy Gibbons, *University of Oxford, UK*

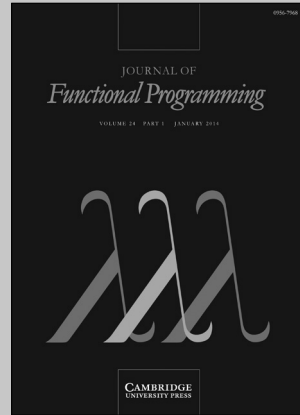
The *Journal of Functional Programming* is the only journal devoted solely to the design, implementation, and application of functional programming languages, spanning the range from mathematical theory to industrial practice. Topics covered include functional languages and extensions, implementation techniques, reasoning and proof, program transformation and synthesis, type systems, type theory, language-based security, memory management, parallelism and applications. The journal is of interest to computer scientists, software engineers, programming language researchers and mathematicians interested in the logical foundations of programming.

Price information

is available at: <http://journals.cambridge.org/jfp>

Free email alerts

Keep up-to-date with new material – sign up at
<http://journals.cambridge.org/jfp-alerts>



Journal of Functional Programming

is available online at:
<http://journals.cambridge.org/jfp>

To subscribe contact Customer Services

in Cambridge:

Phone +44 (0)1223 326070
Fax +44 (0)1223 325150
Email journals@cambridge.org

in New York:

Phone +1 (845) 353 7500
Fax +1 (845) 353 4141
Email
subscriptions_newyork@cambridge.org

For free online content visit:
<http://journals.cambridge.org/jfp>



CAMBRIDGE
UNIVERSITY PRESS

The Knowledge Engineering Review

Editors-in-Chief

Peter McBurney, *King's College, London, UK*

Simon Parsons, *Brooklyn College, City University of New York, USA*

The Knowledge Engineering Review is committed to the development of the field of artificial intelligence and the clarification and dissemination of its methods and concepts. *KER* publishes analyses – high quality surveys providing balanced but critical presentations of the primary concepts in an area; technical tutorials – detailed introductions to an area; application and country surveys commentaries and debates; book reviews; and a popular ‘from the journals’ section, providing the contents of current journals in theoretical and applied artificial intelligence.

Price information

is available at: <http://journals.cambridge.org/ker>

Free email alerts

Keep up-to-date with new material – sign up at
<http://journals.cambridge.org/ker-alerts>



The Knowledge Engineering Review

is available online at:
<http://journals.cambridge.org/ker>

To subscribe contact Customer Services

in Cambridge:

Phone +44 (0)1223 326070
Fax +44 (0)1223 325150
Email journals@cambridge.org

in New York:

Phone +1 (845) 353 7500
Fax +1 (845) 353 4141
Email
subscriptions_newyork@cambridge.org

For free online content visit:
<http://journals.cambridge.org/ker>



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE

JOURNALS

Robotica

An Official Journal of the International Federation of Robotics

Editor-in-Chief

G. S. Chirikjian, *Johns Hopkins University, USA*

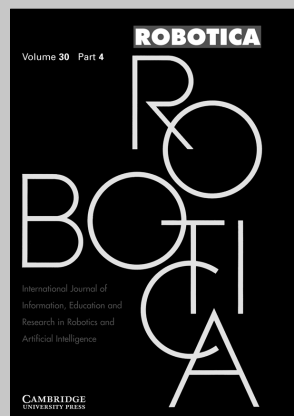
Robotica provides an international forum for the multidisciplinary subject of robotics and encourages developments in this important field of automation with regard to industry, education and research. It covers the many aspects of robotics, including sensory perception, software, kinematics and dynamics involved in robot design, robot task planning and description, intelligibility of skilled motion, applications of robots in the service industries, world model representation, artificial intelligence, development of relevant educational courses, training methods, economic and cost problems and other items of theoretical and practical interest.

Price information

is available at: <http://journals.cambridge.org/rob>

Free email alerts

Keep up-to-date with new material – sign up at
<http://journals.cambridge.org/rob-alerts>



Robotica

is available online at:
<http://journals.cambridge.org/rob>

To subscribe contact Customer Services

in Cambridge:

Phone +44 (0)1223 326070
Fax +44 (0)1223 325150
Email journals@cambridge.org

in New York:

Phone +1 (845) 353 7500
Fax +1 (845) 353 4141
Email
subscriptions_newyork@cambridge.org

For free online content visit:
<http://journals.cambridge.org/rob>



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE

JOURNALS

Theory and Practice of Logic Programming

Published for the Association for Logic programming

Editor-in-Chief

I Niemelä, *Aalto University School of Science, Finland*

Theory and Practice of Logic Programming emphasises both the theory and practice of logic programming. Logic programming applies to all areas of artificial intelligence and computer science and is fundamental to them.

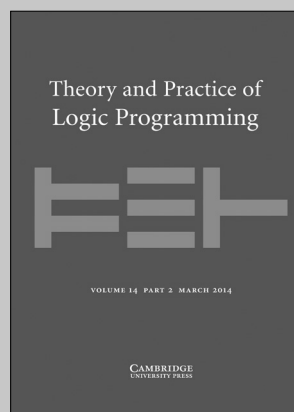
Among the topics covered are AI applications that use logic programming, logic programming methodologies, specification, analysis and verification of systems, inductive logic programming, multi-relational data mining, natural language processing, knowledge representation, non-monotonic reasoning, semantic web reasoning, databases, implementations and architectures and constraint logic programming.

Price information

is available at: <http://journals.cambridge.org/tlp>

Free email alerts

Keep up-to-date with new material – sign up at <http://journals.cambridge.org/tlp-alerts>



Theory and Practice of Logic Programming is available online at: <http://journals.cambridge.org/tlp>

To subscribe contact Customer Services

Americas:

Phone +1 (845) 353 7500

Fax +1 (845) 353 4141

Email

subscriptions_newyork@cambridge.org

Rest of world:

Phone +44 (0)1223 326070

Fax +44 (0)1223 325150

Email journals@cambridge.org

For free online content visit:
<http://journals.cambridge.org/tlp>



CAMBRIDGE
UNIVERSITY PRESS

Instructions for Contributors

Aims and Scope

Natural Language Engineering encourages original papers reporting research with a clear potential for practical application. The journal also invites surveys presenting the state of the art in important areas of Natural Language Engineering and Natural Language Processing as well as squibs discussing specific problems. Book reviews and reports on industrial applications are welcomed, as are conference reports, comparative discussions of Natural Language Engineering products and policy-oriented papers examining, for example, funding programmes or market opportunities. All contributions are peer reviewed.

The journal is part of the Thomson Reuters Citation Index and is indexed in many of their products, but most importantly in both the Science and the Social Sciences editions of the Journal Citation Reports (JCR).

Editorial statement

Natural Language Engineering encourages papers reporting research with a clear potential for practical application. Theoretical papers that consider techniques in sufficient detail to provide for practical implementation are also welcomed, as are shorter reports of on-going research, conference reports, comparative discussions of NLE products, and policy-oriented papers examining e.g. funding programmes or market opportunities. All contributions are peer reviewed and the review process is specifically designed to be fast, contributing to the rapid publication of accepted papers.

Submissions

Submission of a paper to *Natural Language Engineering* is held to imply that it represents an original contribution not previously published and that it is not being considered elsewhere. Authors of articles published in the journal assign copyright to Cambridge University Press (with certain rights reserved) and you will receive a copyright assignment form for signature on acceptance of your paper.

Contributors are responsible for obtaining permission to reproduce any material in which they do not own copyright, to be used in both print and electronic media, and for ensuring that the appropriate acknowledgements are included in their manuscript.

All manuscripts must be submitted online via the website: <http://mc.manuscriptcentral.com/nle>

Papers should be preceded by an abstract of approximately 300 words.

Manuscript requirements

Manuscripts should be double spaced throughout, with wide margins. Pages should be numbered consecutively.

The first page of the manuscript should give the title, the name(s) and full mailing address(es) of the author(s), together with e-mail addresses(es) when possible.

The preferred formatting system is LaTeX, which can be used for direct typesetting, and a style file is available through https://mc.manuscriptcentral.com/societyimages/nle/NLE_LaTeX_Style_File.zip

In case of difficulty, please contact cup-texsupport@aptaracorp.com

Conventions

Spelling, British or American English spelling may be used provided it is used consistently throughout the paper.

Footnotes should be kept to a minimum. Any acknowledgements, or explanation of the genesis of an article, should appear as the first note keyed to the article title by an asterisk (*). Note indicators in the text should follow punctuation.

Bibliographical references should be given in parentheses in standard author-date form in the body of the text: (Lee and Devore 1968: 236). When a second or subsequent work by a particular author in the same year is cited, references should be distinguished by letters (a, b, c, etc.) placed after the date. When a work is written by three or more authors, all names should be given in the first citation: (Fraser, Brown and Smith 1989). In subsequent citations, the first name only should be given with et al. added. A group of references within the text should be date ordered, the earliest first.

A complete list of references cited, arranged alphabetically by author's surname, should be typed double-spaced at the end of the article. The style adopted for particular types of publication should be as follows:

Akmajian, A. and Lehrer, A. 1976. NP-like quantifiers and the problem of determining the head of an NP.

Linguistic Analysis 2: 295-313.

Huddleston, Rodney. 1984. *Introduction to the Grammar of English*. Cambridge: Cambridge University Press.

McCord, Michael C. 1990 Slot grammar: a system for simpler construction of practical natural language grammars. In R. Studer (ed.), *Natural Language and Logic: International Scientific Symposium*, pp. 118-45. Lecture Notes in Computer Science. Berlin: Springer Verlag.

Salton, Gerald, Zhao, Zhongnan, and Buckley, Chris. 1990. A simple syntactic approach for the generation of indexing phrases. Technical Report 90-1137, Department of Computer Science, Cornell University.

Style

Subheadings should be typed with prefatory numbers indicating the level of importance, 1, 1.1, 1.1.1. No more than three levels of subheading should normally be used.

Quotations. Single inverted commas should be used except for quotations within quotations, which should have double inverted commas. Longer quotations of more than 60 words, or quotations which are of particular importance or the focus of your discussion, should be set off from the text with an extra line of space above and below, and typed without inverted commas.

Hyphenation should be kept to a minimum: lifetime, cooperation, subheading.

Numbers should be written out up to 100, except where they refer to precise measurements. Above 100, use a comma rather than a space where four or more digits are involved (2,000 not 2 000). The words 'per cent' should be written out rather than abbreviated to %. Centuries too should be written out in full ('the nineteenth century' rather than 'the 19th century'). Make all journal numbers arabic.

Abbreviations. Do not use pp. before page numbers if the volume number is also given. Use Zip code form (MA, IL) for US state names. Elide page numbers to the shortest pronounceable form: 56-7, 281-3, but 215-16. Contractions and acronyms should have no full points (Dr, NATO), but abbreviations and their plurals should retain them (vol., vols., ed., eds.).

Note also:

- Make the titles of published works italic (not bold) by underlining or using an appropriate word processor font.
- Do not use inverted commas around chapter titles in edited books, journal articles, and the titles of unpublished dissertations.
- Use minimum capitalisation, that is proper names and the first word of the title only capitalised (French, Spanish, Italian citations follow this rule, but German nouns retain capitals in the normal way)
- Anglicise places of publication: Brunswick, The Hague, Florence, rather than Braunschweig, 's-Gravenhage, Firenze

Illustrations

Tables should be numbered consecutively and designed to fit a printed page of 247 x 174mm. Vertical lines should not be used and horizontal lines should be used only at the top and bottom of the table and below column headings. Totals and percentages should be labelled and units identified.

Illustrations should be provided on separate pages, numbered consecutively in a single sequence whether they are line figures or photographs. Captions should be typed on a separate sheet, double spaced. Indicate in the margin of the typescript approximately where in the text tables and figures should fall.

Figures On acceptance, authors will be asked to provide artwork of a professional standard suitable for direct reproduction. Line drawings, good photo prints and sharp copy are acceptable. If you request colour figures in the printed version, you will be contacted by CCC-Rightslink who are acting on our behalf to collect Author Charges. Please follow their instructions in order to avoid any delay in the publication of your article.

Photographs should be provided with a resolution of 300 dpi, numbered sequentially with any other illustrations (Fig. 4, Fig. 5, not Plate I, Plate II).

English Language Editing Services

Authors, particularly those whose first language is not English, may wish to have their English-language manuscripts checked by a native speaker before submission. This is optional, but may help to ensure that the academic content of the paper is fully understood by the editor and any reviewers. We list a number of third-party services specialising in language editing and / or translation, and suggest that authors contact as appropriate.

Please note that the use of any of these services is voluntary, and at the author's own expense. Use of these services does not guarantee that the manuscript will be accepted for publication, nor does it restrict the author to submitting to a Cambridge published journal.

Proofreading

First proofs may be read and corrected by contributors provided that they can guarantee to return the corrected proofs within two working days. Contributors should correct printers' errors but not introduce new or different material at this stage. The publisher reserves the right to charge authors for correction of non-typographical errors.

Offprints

Contributors will not be sent paper offprints but will instead receive a PDF version upon publication online.

Open Access Policies

Please visit journals.cambridge.org/openaccess for information on our open access policies, compliance with major funding bodies, and guidelines on depositing your manuscript in an institutional repository.

Printed in the UK by Bell & Bain Ltd, Glasgow

Cambridge University Press

Journals Fulfillment Department, UPH, Shaftesbury Road, Cambridge CB2 8BS, UK.

32 Avenue of the Americas, New York, NY 10013-2473, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

Printed in the Netherlands by Cambridge University Press

Dock House, The Waterfront, Cape Town 8001, South Africa

NATURAL LANGUAGE ENGINEERING

Volume 21 Part 4 August 2015

CONTENTS

ARTICLES

- Authorship analysis of aliases: Does topic influence accuracy? 497
Robert Layton, Paul A. Watters and Richard Dazeley
- Exploring the effectiveness of linguistic knowledge for biographical relation extraction 519
Marcos Garcia and Pablo Gamallo
- (Un/Semi-)supervised SMS text message SPAM detection 553
Chris R. Giannella, Ransom Winder and Brandon Wilson
- A tree-based learning approach for document structure analysis and its application to web search 569
F. Canan Pembe and Tunga Güngör
- Construction and evaluation of event graphs 607
Goran Glavaš and Jan Snajder
- Industry Watch
NLP meets the cloud 653
Robert Dale

The cover illustration is computer generated from an engraving of the Tower of Babel in an eighteenth-century bible

Cambridge Journals Online

For further information about this journal
please go to the journal website at:
journals.cambridge.org/nle



MIX
Paper from
responsible sources
FSC® C007785

CAMBRIDGE
UNIVERSITY PRESS