Automatic Summarization

Organizing committee
Udo Hahn, University of Freiburg
Chin-Yew Lin, USC Information Sciences Institute
Inderjeet Mani, MITRE
Dragomir Radev, University of Michigan

April 30, 2000
Seattle, Washington
Automatic summarization aims at producing a concise, condensed representation of the key information content in an information source for a particular user and task. Interest in automatic summarization continues to grow, motivated by the explosion of on-line information sources and advances in natural language processing and information retrieval. In fact, some form of automatic summarization may be indispensable given the massive information universes that lie ahead in the 21st century.

The problem of automatic summarization poses a variety of tough challenges in both NL understanding and generation. A spate of recent papers and tutorials on this subject at conferences such as ACL/EACL, AAAI, ECAI, IJCAI, and SIGIR point to a growing interest in research in this field. Several commercial summarization products have also appeared. There have been several workshops in the past on this subject: in Dagstuhl (1994), Madrid (1997), and Stanford (1998). It is our great pleasure to bring the fourth such event to you.

While the field continues to progress, there are also many problems that need to be addressed before the promises of automatic text summarization can be fully realized. The papers included in this volume present a snapshot of recent progress in solving some of the problems: concept identification, identification of discourse markers, multi-document summarization, content visualization, evaluation, and use. Out of 29 submissions to the workshop, 10 are included in these proceedings.

We would like to thank all authors for their hard work. This workshop would not be possible without them. We would also like to thank the program committee and all the reviewers for their valuable feedback. In addition, we are grateful to the ACL for sponsoring the workshop and to the MITRE Corporation for supporting the publication of the proceedings.
ORGANIZING COMMITTEE

Udo Hahn
Chin-Yew Lin
Inderjeet Mani
Dragomir Radev

University of Freiburg
USC/Information Sciences Institute
MITRE
University of Michigan

PROGRAM COMMITTEE

Elisabeth André
Branimir Boguraev
Chris Buckley
Michael Elhadad
Takahiro Fukushima
Udo Hahn
Eduard Hovy
Hongyan Jing
Elizabeth Liddy
Chin-Yew Lin
Inderjeet Mani
Daniel Marcu
Shigeru Masuyama
Mark Maybury
Vibhu Mittal
Sung Hyon Myaeng
Akitoshi Okumura
Dragomir Radev
Chris Paice
Karen Sparck-Jones
Tomek Strzalkowski
Simone Teufel
Benjamin Tsou

DFKI GmbH
IBM Research
SabiR Research
Ben Gurion University
Telecommunications Advancement Organization of Japan
University of Freiburg
USC/Information Sciences Institute
Columbia University
Syracuse University
USC/Information Sciences Institute
MITRE
USC/Information Sciences Institute
MITRE
Just Research
Chungnam National University
NEC
University of Michigan
Lancaster University
University of Cambridge
GE CRD
University of Edinburgh
City University of Hong Kong

ADDITIONAL REVIEWERS

Ulf Hermjakob
Hsin-Hsi Chen

USC/Information Sciences Institute
National Taiwan University, Taiwan
# TABLE OF CONTENTS

**Concept Identification and Presentation in the Context of Technical Text Summarization**  
Horacio Saggion and Guy Lapalme ................................................................. 1

**Mining Discourse Markers for Chinese Textual Summarization**  
Samuel W. K. Chan, Tom B. Y. Lai, W. J. Gao, and Benjamin K. Tsou .................. 11

**Centroid-based summarization of multiple documents: sentence extraction, utility-based evaluation, and user studies**  
Dragomir R. Radev, Hongyan Jing, Malgorzata Budzikowska ..................................... 21

**Extracting Key Paragraph based on Topic and Event Detection --- Towards Multi-Document Summarization**  
Fumiyo Fukumoto and Yoshimi Suzuki ...................................................................... 31

**Multi-Document Summarization By Sentence Extraction**  
Jade Goldstein, Vibhu Mittal, Jaime Carbonell, and Mark Kantrowitz ......................... 40

**Text Summarizer in Use: Lessons Learned from Real World Deployment and Evaluation**  
Mary Ellen Okurowski, Harold Wilson, Joaquin Urbina, Tony Taylor, Ruth Colvin Clark, and Frank Krapcho .............................................................. 49

**Evaluation of Phrase-representation Summarization based on Information Retrieval Task**  
Mamiko Oka and Yoshihiro Ueda .............................................................................. 59

**A Comparison of Rankings Produced by Summarization Evaluation Measures**  
Robert L. Donaway, Kevin W. Drummey, and Laura A. Mather ............................... 69

**Multi-document Summarization by Visualizing Topical Content**  
Rie Kubota Ando, Branimir K. Boguraev, Roy J. Byrd, and Mary S. Neff ................... 79

**Using Summarization for Automatic Briefing Generation**  
Inderjeet Mani, Kristian Concepcion, and Linda van Guilder .................................. 89
AUTHOR INDEX

Branimir Boguraev ..................................... 79
Roy Byrd ..................................................... .79
Malgorzata Budzikowska ........................... 21
Jaime Carbonell .......................................... 40
Samuel W. K. Chan ..................................... 11
Ruth Colvin Clark ....................................... 49
Kristian Concepcion ................................... 89
Robert L. Donaway ..................................... 69
Kevin W. Drummey .................................... 69
Fumiyo Fukumoto ...................................... 31
W. J. Gao .................................................. 11
Jade Goldstein .......................................... 40
Hongyan Jing ............................................. 21
Mark Kantrowitz ........................................ 40
Frank Krapcho .......................................... 49
Rie Kubota Ando ........................................ 79
Tom B. Y. Lai ............................................. 11

Guy Lapalme............................................. 1
Inderjeet Mani ........................................... 89
Laura A. Mather ........................................ 69
Vibhu Mittal ............................................. 40
Mary Neff ............................................... 79
Mamiko Oka ............................................. 59
Mary Ellen Okurowski ................................ 49
Dragomir R. Radev .................................... 21
Horacio Saggion ....................................... 1
Yoshimi Suzuki ......................................... 31
Tony Taylor ............................................. 49
Benjamin K. T'sou .................................... 11
Yoshihiro Ueda ......................................... 59
Joaquin Urbina .......................................... 49
Linda Van Guilder ..................................... 89
Harold Wilson ......................................... 49