Lecture Notes in Computer Science 1293
Edited by G. Goos, J. Hartmanis and J. van Leeuwen

Advisory Board: W. Brauer D. Gries J. Stoer

Springer
Berlin
Heidelberg
New York
Barcelona
Budapest
Hong Kong
London
Milan
Paris
Santa Clara
Singapore
Tokyo

Charles Nicholas Derick Wood (Eds.)

Principles of Document Processing

Third International Workshop, PODP'96
Palo Alto, California, USA, September 23, 1996
Proceedings

Springer

Series Editors

Gerhard Goos, Karlsruhe University, Germany

Juris Hartmanis, Cornell University, NY, USA

Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editors

Charles Nicholas
University of Maryland Baltimore County
Department of Computer Science and Electrical Engineering
1000 Hilltop Circle, Baltimore, MD 21250, USA
E-mail: nicholas@cs.umbc.edu

Derick Wood
The Hong Kong University of Science and Technology
Department of Computer Science
Clear Water Bay, Kowloon, Hong Kong
E-mail: dwood@cs.ust.hk

Cataloging-in-Publication data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Principles of document processing : third international workshop ; proceedings / PODP '96, Palo Alto, California, USA, September 23, 1996 / Charles Nicholas ; Derick Wood (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Budapest ; Hong Kong ; London ; Milan ; Paris ; Santa Clara ; Singapore ; Tokyo : Springer, 1997
(Lecture notes in computer science ; Vol. 1293)
ISBN 3-540-63620-X

CR Subject Classification (1991): I.7, H.5, I.3.7, I.4

ISSN 0302-9743
ISBN 3-540-63620-X Springer-Verlag Berlin Heidelberg New York

Typesetting: Camera-ready by author
SPIN 10546383 06/3142 – 5 4 3 2 1 0 Printed on acid-free paper

Preface

The Third International Workshop on Principles of Document Processing took place in Palo Alto, California, on September 23, 1996. PODP'96 was the third in a series of international workshops that provide forums to discuss the modeling of document processing systems using theories and techniques from, for example, computer science, mathematics, and psychology. PODP'96 took place in conjunction with EP'96 at Xerox Corporation's conference center in Palo Alto, California.

The charter of the PODP workshops is deliberately ambitious and its scope broad. The current state of electronic document processing can be characterized as a plethora of tools without a clear articulation of unifying principles and concepts underlying them. The practical and commercial impact of these tools, which include formatters, composition systems, word processing systems, structured editors, and document management systems, among others, is too pervasive and obvious to require further elaboration and emphasis. However, with the rapid development in hardware technology (processors, memory, and especially high bandwidth networks) the notion of a document and of document processing itself is undergoing a profound change. It is imperative that this change be fueled, not only by enabling technologies and tools, but also by precise, computational, and conceptual models of documents and document processing. To this end, we hope to bring to bear theories and techniques developed by researchers in other areas of science, mathematics, engineering, and the humanities (such as databases, formal specification languages and methodologies, optimization, work-flow analysis, and user interface design.)

The PODP workshops are intended to promote a happy marriage between documents and document processing, and theories and techniques. PODP provides an ideal opportunity for discussion and information exchange between researchers who are grappling with problems in *any* area of document processing.

We invited researchers to submit papers with a good balance between theory and practice in document processing. Papers that address both on a somewhat equal basis were preferred. Each paper was subjected to rigorous peer review. We extend our thanks to the other members of this Workshop's Program Committee, for their hard work under tight time pressure: Howard Blair (USA), Heather Brown (UK), Anne Brueggemann-Klein (Germany), Richard Furuta (USA), Heikki Mannila (Finland), Ethan Munson (USA), Makoto Murata (Japan), and James Sasaki (USA).

There was considerable discussion about the workshop name during and after PODP'96. As a result of this discussion, and to more accurately reflect the workshop's focus on the processing of documents in digital form, it was decided to change the workshop's name to Principles of Digital Docu-

ment Processing. The next workshop in this series, PODDP'98, is planned for April 1998, in Saint Malo, France, in conjunction with EP'98, the seventh International Conference on Electronic Publishing, Document Manipulation and Typography.

July 1997

Charles Nicholas

Derick Wood

Table of Contents