Lecture Notes in Computer Science 2428
Edited by G. Goos, J. Hartmanis, and J. van Leeuwen

Springer
Berlin
Heidelberg
New York
Barcelona
Hong Kong
London
Milan
Paris
Tokyo

Holger Hermanns

Interactive Markov Chains

And the Quest for Quantified Quality

Springer

Series Editors

Gerhard Goos, Karlsruhe University, Germany
Juris Hartmanis, Cornell University, NY, USA
Jan van Leeuwen, Utrecht University, The Netherlands

Author

Holger Hermanns
University of Twente, Faculty of Computer Science
Formal Methods and Tools Group
P.O. Box 217, 7500 AE Enschede, The Netherlands
E-mail: hermanns@cs.utwente.nl

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Hermanns, Holger:
Interactive Markov chains : and the quest for quantified quality / Holger Hermanns. - Berlin ; Heidelberg ; New York ; Hong Kong ; London ; Milan ; Paris ; Tokyo : Springer, 2002
(Lecture notes in computer science ; 2428)
ISBN 3-540-44261-8

CR Subject Classification (1998): D.2.4, F.3.2, F.1.2, C.4, D.2.2, G.3

ISSN 0302-9743
ISBN 3-540-44261-8 Springer-Verlag Berlin Heidelberg New York

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH

http://www.springer.de

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Christian Grosche, Hamburg
Printed on acid-free paper SPIN 10873887 06/3142 5 4 3 2 1 0

Preface

Markov chains are widely used as stochastic models to study and estimate a broad spectrum of *performance and dependability characteristics*. In this monograph we address the issue of *compositional specification* and analysis of Markov chains. Based on principles known from *process algebra*, we develop an algebra of *Interactive Markov Chains* (IMC) arising as an *orthogonal* extension of both continuous-time Markov chains and process algebra. In this algebra the interrelation of delays and actions is governed by the notion of *maximal progress*: Internal actions are executed without letting time pass, while external actions are potentially delayed. IMC is more than 'yet another' formalism to describe Markov chains. This claim is substantiated by a number of distinguishing results of both theoretical and practical nature. Among others, we develop an *algebraic theory* of IMC, devise *algorithms* to mechanise *compositional aggregation* of IMC, and successfully apply these ingredients to analyse state spaces of several million states, resulting from a study of an ordinary telephone system.

The contents of this monograph is a revised version of my PhD thesis manuscript [96] which I completed in spring 1998 at the University of Erlangen, Germany. I am deeply indebted to Ulrich Herzog and Ed Brinksma for their enthusiastic support when preparing its contents, and when finalising this revision at the University of Twente, The Netherlands.

Many researchers had inspiring influence on this piece, or on myself in a broader context, and I take the opportunity to express my gratitude to all of them. I am particularly happy to acknowledge enjoyable joint research efforts with Christel Baier, Salem Derisavi, Joost-Pieter Katoen, Markus Lohrey, Michael Rettelbach, Marina Ribaudo, William H. Sanders, and Markus Siegle which have led to various cornerstones of this book. Henrik Bohnenkamp, Salem Derisavi, and Marielle Stoelinga read the manuscript carefully enough to spot many flaws, and gave me the chance to iron them out in this monograph. Cordial thanks go to Alfred Hofmann at Springer-Verlag for his support in the process of making the manuscript a part of the LNCS series. And finally, there is Sabine and the tiny crowd. Those who know her are able to assess how perfectly happy I account myself.

June 2002 Holger Hermanns

Foreword

To devise methods for the construction of high quality information processing systems is a major challenge of computer science. In most contexts, however, the definition of what constitutes (high) quality in a more concrete sense is problematic, as invariably any definition seems to fall short of its essence. Computer science proves no exception to the rule, and its quest for quality in relation to the analysis of system designs has given birth to two main interpretations: quality as correctness, and quality as performance.

The first interpretation assesses quality by showing formally that (a model of) a system satisfies the functional requirements of its formal specification. Its methods are rooted in logic and discrete mathematics, and are based on the all-or-nothing game imposed by the Boolean lattice: unless satisfaction has been demonstrated completely, nothing can be said. This is both the strength and the weakness of the approach: results have the utmost precision, but are hard to obtain.

The second interpretation aims to assess quality on a continuous scale that allows for quantification: using stochastic system models one tries to calculate system properties in terms of mathematical expectation, variation, probability, etc. The strong point of this approach is that it allows for quality in other than absolute terms, e.g. a message loss of less than 0.01%, service availability of more than 99.99%, etc. Its weaker side is that it cannot handle very well system properties that are not directly related to repeatable events, including many functional system properties, such as e.g. absence of deadlock, reachability of desirable system states, etc.

It is clear that the analysis of the quality of system designs must ultimately encompass both of the approaches above. A first step in this direction was the development of stochastic Petri net models, which combine a classical functional model for (concurrent) systems with stochastic features. The latter allow the derivation of performance models in the form of continuous-time Markov chains directly from a system description using such nets. Thus the formalism in principle allows functional and performance analysis of systems in terms of an integrated model and perspective.

This potentially great leap forward from the existing practice of studying correctness and performance through unrelated models (and by different scientific communities) proved harder to materialise than was initially hoped for. One of the main causes was the infamous state space explosion: the fact that the number of global states of a system grows exponentially with the number of components of the system. Because of this, non-trivial system designs give rise to large Petri net models, which in turn yield huge Markov models that can no longer be effectively manipulated, even with the aid of computers.

In the early 1990s this observation motivated the study of what is now referred to as *stochastic process algebras*. In the preceding decade process algebras had proven an effective means for the modelling and analysis of the functionality of concurrent systems. They address the problem of state space explosion by a powerful formalisation of system composition by process algebraic operators, combined with the study of *observational congruence* of behaviours. The latter allows for a *compositional* control of state space complexity: replacing components with observationally congruent but simpler components the state space can be reduced without explicitly generating it first.

The study of stochastic process algebra has for a considerable part been driven by the non-trivial question of how best to add stochastic features to process algebra, combining sufficient stochastic expressivity with compatibility with existing process algebraic theory. The present LNCS volume by Holger Hermanns contains his answer to this question for Markovian process algebra, i.e., where the stochastic model of interest is that of continuous-time Markov chains. Written in a clear and refreshing style it demonstrates that it is not only Hermanns' answer, but really 'the' answer.

Where others before him treated stochastic delay as attributes of system actions, Hermanns saw the enormous advantages of a completely different approach: treating delays as actions in their own right that silently consume exponentially distributed amounts of time, and treating system actions as instantaneous actions. This separation of concerns bears all the signs of a great idea: it is (retrospectively) simple and leads to very elegant results. A complication that mars the other approaches, viz. the synchronisation of delays as a by-product of synchronising actions, is completely avoided. Only system actions are subject to synchronisation, and delays in different components of a system are interleaved. Due to the memoryless nature of exponential distributions this yields a perfectly natural interpretation of the passage of (stochastic) time. It is the Platonic discovery that interleaving process algebra and Markov chains are a perfect couple. Another advantage of the decoupling of system actions and delays is that there can be more than one delay preceding an action. This extends the class of (implicit) action delays far beyond that of the exponential distribution, viz. to the (dense) class of phase-type distributions.

The author must be commended for the technical skills with which he has reaped the full benefits of this idea. In addition to defining and applying his formalism, he has also firmly embedded it in standard process algebraic theory by providing full axiomatisations for the stochastic varieties of observation congruence which are conservative extensions of the non-stochastic cases. Also, the link between the concepts of lumpability in Markov chains and bisimilarity in process algebra that was first observed by Hillston, comes to full fruition in the hands of Hermanns. Based on standard algorithms for bisimulation a low complexity algorithm is devised for lumping (Markov chain) states that can be applied compositionally. The latter is a fine example of the advantages of interdisciplinary research, as such an algorithm was not available in the standard theory of Markov chains.

We believe that this monograph by Holger Hermanns represents an important step in the quest for the integrated modelling and analysis of functional and performance properties of information processing systems. It is also written in a very accessible and, where appropriate, tutorial style, with great effort to explain the intuition behind the ideas that are presented. With a growing number of researchers in the performance analysis and formal methods communities that are interested in combining their methods, we think that this monograph may serve both as a source of inspiration and a work of reference that captures some vital ingredients of quality.

May 2002 Ed Brinskma, Ulrich Herzog

Contents