Cognition and Intractability

Intractability is a growing concern across the cognitive sciences: while many models of cognition can describe and predict human behavior in the lab, it remains unclear how these models can scale to situations of real-world complexity. *Cognition and Intractability* is the first book to provide an accessible introduction to computational complexity analysis and its application to questions of intractability in cognitive science. Covering both classical and parameterized complexity analysis, it introduces the mathematical concepts and proof techniques that can be used to test one's intuition of (in)tractability. It also describes how these tools can be applied to cognitive modeling to deal with intractability – and its ramifications – in a systematic way. Aimed at students and researchers in philosophy, cognitive neuroscience, psychology, artificial intelligence, and linguistics who want to build a firm understanding of intractability and its implications in their modeling work, it is an ideal resource for teaching or self-study.

Iris van Rooij is a psychologist and cognitive scientist at the Donders Institute for Brain, Cognition and Behaviour and the School for Psychology and Artificial Intelligence at Radboud University, the Netherlands.

Mark Blokpoel is a computational cognitive scientist at the Donders Institute for Brain, Cognition and Behaviour at Radboud University, the Netherlands.

Johan Kwisthout is a computer scientist at the Donders Institute for Brain, Cognition and Behaviour and the School for Psychology and Artificial Intelligence at Radboud University, the Netherlands.

Todd Wareham is a computer scientist at the Department of Computer Science at Memorial University of Newfoundland, Canada.

Cambridge University Press
978-1-107-04399-2 — Cognition and Intractability
Iris van Rooij , Mark Blokpoel , Johan Kwisthout , Todd Wareham
Frontmatter
More Information

Cognition and Intractability

A Guide to Classical and Parameterized Complexity Analysis

IRIS VAN ROOIJ
Radboud University Nijmegen

MARK BLOKPOEL
Radboud University Nijmegen

JOHAN KWISTHOUT
Radboud University Nijmegen

TODD WAREHAM
Memorial University of Newfoundland

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,
New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107043992
DOI: 10.1017/9781107358331

First published 2019

Printed and bound in Great Britain by Clays Ltd, Elcograf S.p.A.

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data

Names: Rooij, Iris van, 1973- author. | Blokpoel, Mark, 1984- author. |
Kwisthout, Johan, 1976- author. | Wareham, Todd, 1963- author.
Title: Cognition and intractability : a guide to classical and parameterized complexity analysis / Iris van Rooij (Radboud Universiteit Nijmegen), Mark Blokpoel (Radboud University Nijmegen), Johan Kwisthout (Universiteit Leiden), Todd Wareham (Memorial University of Newfoundland).
Description: Cambridge ; New York, NY : Cambridge University Press, 2019. |
Includes bibliographical references and index.
Identifiers: LCCN 2018057971| ISBN 9781107043992 (hardback : alk. paper) |
ISBN 9781108728973 (pbk. : alk. paper)
Subjects: LCSH: Computational complexity. | Decision making. | Probabilities.
Classification: LCC QA267.7 .C64 2019 | DDC 511.3/52–dc23 LC record available at
https://lccn.loc.gov/2018057971

ISBN 978-1-107-04399-2 Hardback
ISBN 978-1-108-72897-3 Paperback

Cambridge University Press
978-1-107-04399-2 — Cognition and Intractability
Iris van Rooij , Mark Blokpoel , Johan Kwisthout , Todd Wareham
Frontmatter
More Information

Contents

Part III Reflections and Elaborations

Part IV Applications

CAMBRIDGE
Cambridge University Press
978-1-107-04399-2 — Cognition and Intractability
Iris van Rooij , Mark Blokpoel , Johan Kwisthout , Todd Wareham
Frontmatter
More Information

Figures

Cambridge University Press
978-1-107-04399-2 — Cognition and Intractability
Iris van Rooij , Mark Blokpoel , Johan Kwisthout , Todd Wareham
Frontmatter
More Information

Cambridge University Press
978-1-107-04399-2 — Cognition and Intractability
Iris van Rooij , Mark Blokpoel , Johan Kwisthout , Todd Wareham
Frontmatter
More Information

Tables

Preface

P.1 Why This Book?

Intractability has been a growing concern in cognitive science: many theories of cognition can describe and predict human behavior in the lab, but it remains unclear how these theories can scale to situations of real-world complexity given that they postulate intractable (e.g., NP-hard) computations. Unbeknownst to many cognitive scientists, there exist complexity theoretic tools for dealing with the problem of intractability. These methods have been developed over the last decades in theoretical computer science. This book aims to make those methods visible and accessible for a broad cognitive science audience, in effect growing the community of researchers that will have these complexity theoretic tools as part of their standard theoretical toolkit. With this, we believe cognitive science will be in a significantly better position to tackle one of its key theoretical problems – namely, how can cognition in its full richness and complexity be computationally tractable?

In addition, we hope to contribute to the unification of otherwise disconnected work in the cognitive science literature. At the end of this book, you will find a compendium of computational-level theories, together with their known complexity results. These theories come from different content domains (decision-making, analogy, reasoning, vision, language, action, etc.) and fall under different theoretical frameworks (e.g., symbolic, neural network, probabilistic, dynamical, logic, robotic, and heuristic models of cognition). It is our experience that researchers from different areas often are not aware that from a formal perspective these (superficially distinct looking) models and their computational complexity are in fact closely related. By bringing these models together in one place and by highlighting their common and distinct structure, we hope to generate new and more general ways of thinking about computational-level theorizing in cognitive science and the associated problem of intractability. We hope that our compendium, not unlike the compendium of NP-complete problems in Garey and Johnson's classical book of 1979, will become invaluable as both a reference for existing work and an aid to future work.

P.2 For Whom Is This Book?

The book is written for advanced undergraduate students, graduate students, and researchers in cognitive science and its affiliated disciplines. Among these disciplines we count computational- and psycholinguistics, cognitive psychology, artificial intelligence, cognitive neuroscience, computer science, logic, and philosophy of mind, science, and computation.

The book assumes an inherently interdisciplinary perspective. It is written primarily for cognitive scientists interested in developing and assessing computational explanations of natural cognition. This will likely include cognitive psychologists and cognitive neuroscientists, but also artificial intelligence researchers interested in emulating natural cognition in artificial systems. The book may also be of interest to computer scientists who wish to apply complexity theoretic methodology to cognitive theories. Finally, philosophers of cognitive science may find the meta-scientific considerations laid out in this book of special interest. We hope the book will inspire interesting and productive interdisciplinary collaborations.

In an attempt to give the reader a firm grip on the notion of intractability, we will cover many more formal details than is common in the cognitive science literature. We believe that a proper understanding of tractability, its formalizations, and its role in cognitive science, demands more than a superficial understanding of the mathematical theory of computational complexity. To make these formal details accessible for a broad audience of cognitive scientists we include in our textbook many informal explanations, simple examples, and practical exercises. See also the Section P.5 Guide to the Reader.

P.3 What Not to Expect

Both cognitive science and complexity theory are vast fields of study. We cannot possibly cover everything about these fields in this book. Even if we could, it would detract from our main points. This book focusses on a specific interface between complexity theory and cognitive science: the *intractability* of *computational-level* theories of cognition. Even doing justice to this focus requires us to cover a substantial amount of material in a coherent fashion. Consequently, this book will not cover many topics one may otherwise expect in a computer science book on complexity theory or in a cognitive science book on computational modeling.

For instance, we will not cover complexity classes that distinguish between different types of highly efficient computations (such as the classes inside P). Also, we largely ignore many of the classes that contain NP, because most relevant problems for cognitive science live in NP (or parameterized or probabilistic analogues of NP), and even where they don't it is hardness for NP that matters ultimately for intractability (not membership in any of the other classes that contain it). Although we mention classes such as PSPACE, EXP, etc. in passing, students who become interested in knowing more about the polynomial-time hierarchy and more advanced complexity classes may do well to seek out other, more specialized theoretical computer science sources. In later chapters, we do mention some other classes, such as BPP and BQP for as far as they are relevant for philosophical discussions about the tractability constraint on computational-level theories of cognition.

We also will be relatively silent about the many important advances in algorithmic-level modeling in cognitive science and how these interface with neuroscience research to inform implementational-level explanations. Instead, we will focus on computational-level theories and consider algorithmic- and implementation-level constraints only insofar as they bear on understanding the constraints under which computations postulated at the computational level can or cannot be tractable. As we explain in the more conceptual chapters in this book, this will be largely without loss of generality.

P.4 Organization of the Book

This book is written as a set of technical chapters sandwiched between more conceptual chapters. We start with an accessible introduction to the topic of *Cognition and Intractability* in Chapter 1. We use the problem of selecting pizza toppings as a running example to illustrate the key concepts. We also briefly review conceptual foundations of intractability and computational explanation to lay the groundwork for understanding the need for learning the concepts and proof techniques in Chapters 2–7. The first three of these Chapters (2–4) cover the classical theory of computational complexity, built on notions such as polynomial versus non-polynomial (e.g., exponential) time; polynomial-time reduction; the classes P, NP; and the notions of NP-completeness and NP-hardness. The latter three (Chapters 5, 6, and 7) cover parameterized complexity theory, built on notions such as parameterized problems, fixed-parameter tractability, parameterized reduction, and parameterized complexity classes such as FPT and W[1]. The technical chapters are followed by two chapters that contain conceptual reflections and philosophical elaborations on how the techniques from Chapters 2 to 7 can be used to deal with the problem of intractability as it arises in cognitive science. The first of these (Chapter 8) considers different ways in which cognitive scientists have been trying to deal with intractability and assesses their validity. The second (Chapter 9) presents a set of 14 common objections to complexity analyses and tractability constraints for models of cognition. By considering each such possible objection one by one we aim to make insightful that many objections vanish upon closer inspection because they are typically built on misunderstandings or incorrect suppositions. Next, Chapters 10–12 illustrate the use of the concepts and techniques from Chapters 2 to 7 for computational-level theories in three distinct cognitive domains: constraint satisfaction theory of coherence, structure-mapping theory of analogy, and Bayesian inference theory of communication.

The book also contains a set of Appendices. Appendix A gives primers on mathematical concepts, definitions, and notation that is assumed throughout the book. Readers who miss some of this background are advised to carefully study this Appendix before reading Chapters 2–7. Appendix B presents an alphabetically ordered list of *all* the computational problems referenced in illustrations, practices, and exercises in this book. Appendix C presents a compendium of existing classical and parameterized complexity results for a set of computational-level theories in cognitive science, spanning a variety of cognitive domains, such as perception, planning, decision-making, language processing, and higher-order thinking and reasoning.

P.5 Guide for the Reader

Whether you are a student or otherwise, we believe that the material in this book is best learned when taking an active approach to learning. We stimulate this in three distinct ways: (1) Stop-and-think boxes, (2) Practices, and (3) Exercises.

1. **Stop-and-think boxes** sometimes appear in the text where the reader is invited to contemplate an idea, topic, or a problem before continuing reading. After such a reflection, we always describe the main message that we think the reader should take away from these reflections and explain how and why.
2. In **Practices** we leave it to the reader to find a way to successfully perform those practices. The build-up of practices throughout the book is such that

a reader should be able to perform them – possibly with non-negligible effort – after having studied and mastered all the material up to that point. Especially practices that require the reader to come up with a proof can take some effort. This should not be taken as discouragement, but it can be seen as a natural consequence of the fact that coming up with proofs is a creative process. Acquiring this competence requires regularly trying to come up with proofs and then learning from your mistakes and successes. In our experience, this learning experience can benefit from peer and teacher feedback on written proofs. A good understanding of the material in this book does not require that one is able to perform all practices. Whether you are a student, teacher, or other reader, you can make a wise selection depending on your interests and the challenges you may have in understanding the material.
3. At the end of each chapter, we include **Exercises**. These are not unlike Practices, but are made available as additional ways for students to quiz themselves about the material and as inspiration for teachers as examination questions.

We think all readers will benefit from reading the chapters in the order that they appear in the book, with the exception of the Appendices (which are best referred to on a need-to basis), the starred subsections in Chapters 3, 4, and 6 (which consist of advanced optional material), and Chapters 8 and 9 (in Part III, Reflections and Elaborations). It is our experience that especially readers with a prior cognitive science training and background may have all kinds of concerns when reading Chapter 1 about how "intractability" is relevant or irrelevant for cognitive science. For instance, they may be concerned that intractability is not a real issue for real brains as it is for computers, because brains are very different from Turing machines or contemporary neural network models. Similarly, they may be concerned that human cognition does not typically operate by exact, serial, or deterministic algorithms but rather adopts all kinds of heuristics, approximate, parallel, and probabilistic computations. Although these concerns are understandable, they need not arise once one has a proper understanding of the formal foundations of intractability. Therefore, Chapters 8 and 9 can probably best be appreciated after having rigorously studied the concepts and techniques in Chapters 2–7. Nevertheless, readers may find it useful to occasionally have a peek at Chapters 8 and 9 whenever concerns arise, so that they can hopefully see that these conceptual issues will be dealt with in due course. Moreover, it is our hope that given the knowledge in this book, the reader may even be able to address concerns that are not present in this book. If you do, then please write us about them!

P.6 Reuse of Materials

The book as it lies before you has grown from the collective effort of four authors, each with their own styles, focus, and expertise. Besides lecture notes that we had developed over the years, the text in this book builds on some of our previously published work. This is most visible in Chapter 9, which reiterates a substantial part of section 6 of van Rooij, I. (2008). The Tractable Cognition thesis. *Cognitive Science, 32*, 939–984. This paper also formed the basis of Section 1.2.2 in Chapter 1 and parts of Section 8.2 in Chapter 8, and its appendix was reused and updated to form Sections A.1 and A.3 in Appendix A.

In other chapters reuse of materials is more implicit, as ways of phrasings and footnotes have been reused from some of our collaborative papers in an integrated manner. Here we give a complete list of which materials this includes.

- Some footnotes in Chapter 1 have been adapted from: van Rooij, I. (2003). *Tractable cognition: Complexity theory in cognitive psychology.* PhD thesis, University of Victoria, Canada.
- Parts of Section 8.2 in Chapter 8 have been adapted from: van Rooij, I. (2015). How the curse of intractability can be cognitive science's blessing. In Noelle, D. C., Dale, R., Warlaumont, A. S., Yoshimi, J., Matlock, T., Jennings, C. D., & Maglio, P. P. (Eds.). In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Parts of Section 9.14 in Chapter 9 have been adapted from: van Rooij, I., Wright, C., Kwisthout, J., & Wareham, T. (2018). Rational analysis, intractability, and the prospects of 'as if'-explanations. *Synthese, 195*(2), 491–510.
- Parts of Section 8.2.3 in Chapter 8 have been adapted from: van Rooij, I., Wright, C., & Wareham, H. T. (2012). Intractability and the use of heuristics in psychological explanations. *Synthese, 187*, 471–487.
- Parts of Chapter 12 have been adapted from: Blokpoel, M., Kwisthout, J., Wareham, T., Haselager, P., Toni, I., & van Rooij, I. (2011). The computational costs of recipient design and intention recognition in communication. In L. Carlson, C. Holscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 465–470). Austin, TX: Cognitive Science Society.

Where figures and tables are used or adapted from previously published work this is indicated in the captions.

P.7 Acknowledgments

This book would not have come into existence without the help, support, inspiration, and enthusiasm of many people.

First off, we would like to thank our students for having been willing participants in our experimental approach to teaching the materials in this book. The very first version of a course on this topic was taught by Iris in 2008, with several guest lectures from Todd. The course had only three enrolled students (one of the students was Mark, now a co-author of this book). Over the last decade, the group of students has grown to a yearly group of 20–40 students. This is excepting those students and researchers that we have taught in guest lectures, workshops, and tutorials. We thank the Cognitive Science Society, the International Conference for Computational Modeling, and the Interdisciplinary College for funding and hosting tutorials and workshops for international audiences in the period 2012–2013. We also thank Frank Jäkel for hosting a three-day block course on the topic of this book for Bachelor, Master, and PhD students at the University of Osnabrück, Germany, in 2012 and Daniël Lakens for hosting a workshop for PhD students from across The Netherlands at Eindhoven University of Technology in 2013.

From all students we have learned invaluable things – both what works and what doesn't when teaching computational complexity to cognitive science students with little background in complexity theory or to complexity theorists with little background in cognitive science. We have also learned what are typical concerns or misunderstandings to be prevented or fixed. We hope that we have been able to transform those learning experiences into a book with helpful illustrations, practices, and exercises that take the learner step by step through the conceptual foundations of intractability and computational modeling in cognitive science to the concepts and proof techniques from classical and parameterized complexity.

We next would like to thank our departments and institutes – the Department of Artificial Intelligence and the Donders Institute for Brain, Cognition and Behaviour at the Radboud University (Nijmegen, the Netherlands) and the Department of Computer Science at the Memorial University of Newfoundland (St. John's, Canada) – for their generous support of the time needed for writing this book. The Internationalisation fund of the Radboud University has furthermore made it possible to fund yearly visits by Todd to the Radboud University to co-teach and collaborate with Iris, Johan, and Mark and further develop our ideas for this book. We also thank our teaching assistants, Tobias Winner and Nils Donselaar, for helping us evaluate and develop practices over the years.

In the final phases of writing this book we had the pleasure to be supported by excellent proof readers. We thank Antonina Kolokolova and Mesam Timmar for their careful reading and useful feedback on early versions of the book, and we thank Iris van de Pol and Ronald de Haan for checking and contributing to some entries in the compendium. Special thanks goes to Nils Donselaar whose rigor on many dimensions far surpasses ours, which has taught us some invaluable lessons in humility. We are indebted to him for helping us fix numerous mistakes and oversights in earlier versions of this book. We remain, of course, solely responsible for any remaining errors.

Last but not least, we thank our families – in particular Denise, Ilaij, Zinzi, Steffie, Judith, Koen, and Linda – for their loving support.