Verifiable Autonomous Systems

How can we provide guarantees of behaviours for autonomous systems such as driverless cars? This tutorial text for professionals, researchers, and graduate students explains how autonomous systems, from intelligent robots to driverless cars, can be programmed in ways that make them amenable to formal verification. The authors review specific definitions, applications, and the unique future potential of autonomous systems, along with their impact on safer decisions and ethical behaviour. The topics discussed in this book include the use of rational cognitive agent programming from the Beliefs–Desires–Intentions paradigm to control autonomous systems, and the role of model-checking in verifying properties of this decision-making component. Several case studies concerning both the verification of autonomous systems and extensions to the framework beyond the model-checking of agent decision-makers are included, along with complete tutorials for the use of the freely available verifiable cognitive agent toolkit Gwendolen, written in Java.

DR LOUISE A. DENNIS is leader of the Autonomy and Verification research group at The University of Manchester and conference coordinator for the ACM Special Interest Group for Artificial Intelligence (AI). She studied mathematics and philosophy at the University of Oxford and received her PhD from the University of Edinburgh in using AI techniques to prove mathematical theorems; her interest in the overlap between mathematics, philosophy, and AI has continued ever since. Her current research encompasses the programming of autonomous systems, the development of agent programming languages, reasoning about systems and programs via formal mathematical techniques, and the ethical implications of AI. Beyond the university setting, Dr Dennis is active in public engagement and spends a lot of time taking Lego Robots into schools to introduce robotics programming to children.

DR MICHAEL FISHER is a professor of Computer Science at The University of Manchester. He holds a Royal Academy of Engineering Chair in Emerging Technologies and is a fellow of both the British Computer Society and the Institution of Engineering and Technology. He was previously a professor of logic and computation in the Department of Computing & Mathematics at Manchester Metropolitan University and a professor of computer science at University of Liverpool. Dr Fisher's research concerns autonomous systems, particularly software engineering, formal verification, safety, responsibility, and trustworthiness. He has been involved in over 200 journal and conference papers and authored the book *An Introduction to Practical Formal Methods Using Temporal Logic* (Wiley) in 2011.

Verifiable Autonomous Systems

Using Rational Agents to Provide Assurance about Decisions Made by Machines

LOUISE A. DENNIS
University of Manchester

MICHAEL FISHER
University of Manchester

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, 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

103 Penang Road, #05-06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment, a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108484992
DOI: 10.1017/9781108755023

First published 2023

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

Library of Congress Cataloging-in-Publication Data
Names: Dennis, Louise, author. | Fisher, Michael, 1962- author.
Title: Verifiable autonomous systems : using rational agents to provide assurance about decisions made by machines / Louise A. Dennis, University of Manchester, Michael Fisher, University of Manchester.
Description: Cambridge, United Kingdom ; New York, NY, USA : Cambridge University Press, 2023. | Includes bibliographical references and index.
Identifiers: LCCN 2023006704 | ISBN 9781108484992 (hardback) | ISBN 9781108755023 (ebook)
Subjects: LCSH: Robust control. | Automatic control. | Machine learning. | Computer programs–Verification.
Classification: LCC TJ217.2 .D46 2023 | DDC 629.8–dc23/eng/20230302
LC record available at https://lccn.loc.gov/2023006704

ISBN 978-1-108-48499-2 Hardback

This book is dedicated to Gwendolen Sellers.
The original, you might say.

Contents

Acknowledgements

Thanks are due to the many people who worked with us on the material in this book: Davide Ancona, Martin Mose Bentzen, Rafael Bordini, Paul Bremner, Neil Cameron, Rafael C. Cardoso, Marie Farrell, Angelo Ferrando, Mike Jump, Maryam Kamali, Nick Lincoln, Felix Lindner, Alexei Lisitsa, Matt Luckuck, Viviana Mascardi, Owen McAree, Marija Slavkovik, Sandor Veres, Matt Webster, and Alan F. Winfield. Thanks also to Chris Anderson for proof reading the early chapters.