

## The Cambridge Handbook of Artificial Intelligence

---

Artificial intelligence, or AI, is a cross-disciplinary approach to understanding, modeling, and creating intelligence of various forms. It is a critical branch of cognitive science, and its influence is increasingly being felt in other areas, including the humanities. AI applications are transforming the way we interact with each other and with our environment, and work in artificially modeling intelligence is offering new insights into the human mind and revealing new forms mentality can take. This volume of original essays presents the state of the art in AI, surveying the foundations of the discipline, major theories of mental architecture, the principal areas of research, and extensions of AI such as artificial life. With a focus on theory rather than technical and applied issues, the volume will be valuable not only to people working in AI, but also to those in other disciplines wanting an authoritative and up-to-date introduction to the field.

**Keith Frankish** is Visiting Senior Research Fellow at The Open University UK and Adjunct Professor with the Brain and Mind Program in Neurosciences at the University of Crete. He is the author of *Mind and Supermind* (Cambridge, 2004) and *Consciousness* (2005). He is co-editor of *In Two Minds: Dual Processes and Beyond* (with Jonathan St B. T. Evans, 2009), *New Waves in Philosophy of Action* (with Jesús H. Aguilar and Andrei A. Buckareff, 2010), and *The Cambridge Handbook of Cognitive Science* (with William M. Ramsey, Cambridge, 2012).

**William M. Ramsey** is Associate Professor of Philosophy at the University of Nevada, Las Vegas. He is the author of *Representation Reconsidered* (Cambridge, 2007) and co-editor of *Philosophy and Connectionist Theory* (with David Rumelhart and Stephen Stich, 1991), *Rethinking Intuition* (with Michael DePaul, 1998), and *The Cambridge Handbook of Cognitive Science* (with Keith Frankish, Cambridge, 2012).

Cambridge University Press

978-0-521-87142-6 - The Cambridge Handbook of Artificial Intelligence

Edited by Keith Frankish and William M. Ramsey

Frontmatter

[More information](#)

---

Cambridge University Press  
978-0-521-87142-6 - The Cambridge Handbook of Artificial Intelligence  
Edited by Keith Frankish and William M. Ramsey  
Frontmatter  
[More information](#)

---

# The Cambridge Handbook of Artificial Intelligence

---

EDITED BY

**Keith Frankish and William M. Ramsey**



**CAMBRIDGE**  
 UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

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](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9780521691918](http://www.cambridge.org/9780521691918)

© Cambridge University Press 2014

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2014

Reprinted 2015

Printed in the United Kingdom by Clays, St Ives plc

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

*Library of Congress Cataloguing in Publication data*

The Cambridge handbook of artificial intelligence / edited by Keith Frankish and William M. Ramsey.

pages cm

Includes bibliographical references and index.

ISBN 978-0-521-87142-6 (hardback) – ISBN 978-0-521-69191-8 (paperback)

1. Artificial intelligence – Philosophy. I. Frankish, Keith.

II. Ramsey, William M., 1960– editor of compilation. III.

Title: Handbook of artificial intelligence.

Q335.C26 2014

006.3 – dc23 2013048906

ISBN 978-0-521-87142-6 Hardback

ISBN 978-0-521-69191-8 Paperback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

## Contents

List of figures	<i>page</i> vii
Notes on contributors	ix
Acknowledgments	xii
Introduction	1
Keith Frankish and William M. Ramsey	
<b>Part I Foundations</b>	
1 History, motivations, and core themes	15
Stan Franklin	
2 Philosophical foundations	34
Konstantine Arkoudas and Selmer Bringsjord	
3 Philosophical challenges	64
William S. Robinson	
<b>Part II Architectures</b>	
4 GOFAI	89
Margaret A. Boden	
5 Connectionism and neural networks	108
Ron Sun	
6 Dynamical systems and embedded cognition	128
Randall D. Beer	
<b>Part III Dimensions</b>	
7 Learning	151
David Danks	
8 Perception and computer vision	168
Markus Vincze, Sven Wachsmuth, and Gerhard Sagerer	
9 Reasoning and decision making	191
Eyal Amir	

<b>vi</b>	Contents	
	<hr/>	
10	Language and communication Yorick Wilks	213
11	Actions and agents Eduardo Alonso	232
12	Artificial emotions and machine consciousness Matthias Scheutz	247
	<b>Part IV Extensions</b>	
13	Robotics Phil Husbands	269
14	Artificial life Mark A. Bedau	296
15	The ethics of artificial intelligence Nick Bostrom and Eliezer Yudkowsky	316
	Glossary	335
	Index	343

## Figures

1.1	A simple artificial neuron.	<i>page</i> 17
5.1	The overall structure of a connectionist production system.	117
5.2	The CLARION architecture.	121
6.1	An agent and its environment are coupled dynamical systems. The agent in turn is composed of coupled nervous system and body dynamical systems.	137
8.1	An image is given by a 2D pixel array where each pixel measures the amount of light traveling along a ray.	169
8.2	Computer vision as a knowledge-engineering task.	170
8.3	3D scene geometry.	171
8.4	2D scene geometry.	171
8.5	Pattern classification.	172
8.6	Two different perspectives on computer vision.	173
8.7	Matching result based on local descriptors (here SIFT; Lowe, 2004).	176
8.8	Detection results for a dining chair in a home scene (Wohlkinger and Vincze, 2010).	177
8.9	Basic block diagram of visual servoing.	179
8.10	An example of model-based vision.	181
9.1	A Bayesian network graph representing a joint distribution over binary variables <i>Rain</i> , <i>Wet</i> , <i>Sprinkler</i> and multi-valued random variable <i>Money</i> .	198
9.2	A minimax tree of depth 2.	205
13.1	Grey Walter watches one of his tortoises push aside some wooden blocks on its way back to its recharging hutch. Circa 1952.	271
13.2	An articulated industrial robot arm.	272
13.3	Shakey the robot in 1970 reasoning about colored blocks in its environment. Courtesy of SRI International.	274
13.4	Pipeline of functionally decomposed processing used in much classical AI robotics (after Brooks 1986).	274
13.5	The parallel behavioral decomposition for robot control as advocated by Brooks and the behavior-based approach. Decomposition from Brooks (1986).	276

---

13.6	Kismet, a robot designed by Cynthia Breazeal to take part in social interactions with humans. Reproduced with permission.	278
13.7	Key elements of the evolutionary robotics approach. Thanks to Volko Straub for the image.	280
13.8	An early version of the Sussex gantry robot right (b) was a “hardware simulation” of a robot such as that shown left (a).	282
13.9	Schematic diagram of a distributed neural network for the control of locomotion as used by Beer, Chiel, and Sterling (1989) and generalized architecture using a fully connected dynamical network controller for each leg (a), cross-coupled as shown (b).	284
13.10	A fully automatically evolved robot developed by the Golem project (see text for details). Used with permission.	286
13.11	(a) A bee flying down a tunnel with patterned walls (credit: <i>Science</i> , vol. 287) and (b) a robot built by the Research School of Biological Sciences, Australian National University to demonstrate navigation strategies based on observations of the use of optic flow for insects.	287
14.1	A sequence of five steps in the evolution of the Game of Life. Figures produced using Golly 2.1, available on the web at <a href="http://golly.sourceforge.net/">http://golly.sourceforge.net/</a> .	301
14.2	The evolution of the so-called “r-pentomino” shown at times 1, 100, and 200. Figures produced using Golly 2.1, available on the web at <a href="http://golly.sourceforge.net/">http://golly.sourceforge.net/</a> .	302



## Notes on contributors

**Eduardo Alonso** is Reader in Computing at the Department of Computer Science, City University London. He has published his research in journals such as *Knowledge Engineering Review* and *Artificial Intelligence Review*, and in various Springer Lecture Notes in Artificial Intelligence (LNAI) and Lecture Notes in Computer Science (LNCS) volumes, and he has edited a special issue of *International Journal of Autonomous Agents and Multi-Agent Systems* on Multi-Agent Learning.

**Eyal Amir** is Associate Professor of Computer Science at the University of Illinois at Urbana-Champaign (UIUC). His research focuses on AI, specifically reasoning, learning, and decision making with logical and probabilistic knowledge. In 2006 Eyal was chosen by the Institute of Electrical and Electronics Engineers (IEEE) as one of the “10 to watch in AI.”

**Konstantine Arkoudas** is an AI research scientist at Applied Communication Sciences, with a focus on reasoning and knowledge engineering. He has published many articles in AI and computer science, as well as in philosophy, on topics ranging from philosophy of mind and cognitive science to epistemology and philosophy of mathematics.

**Mark A. Bedau** is Professor of Philosophy and Humanities at Reed College, and Editor-in-Chief of the journal *Artificial Life*. He has co-authored or co-edited *Emergence* (2008), *Protocells: Bridging Non living and Living Matter* (2009), *The Ethics of Protocells: Moral and Social Implications of Creating Life in the Laboratory* (2009), and *The Nature of Life* (Cambridge, 2010).

**Randall D. Beer** is a Professor of Informatics and Computing in the Cognitive Science program at Indiana University. He is the author of *Intelligence as Adaptive Behavior* (1990) and the editor of *Biological Neural Networks in Invertebrate Neuroethology and Robotics* (1993), as well as numerous articles.

**Margaret A. Boden** is Research Professor of Cognitive Science at the University of Sussex. She is the author of *The Creative Mind* (2004), *Mind as Machine: A History of Cognitive Science* (2006), *Creativity and Art* (2011), and several other books plus many journal articles.

**Nick Bostrom** is Professor in the Faculty of Philosophy at Oxford University and Director of the Future of Humanity Institute within the Oxford Martin

---

School. He is the author of some 200 publications, including *Anthropic Bias* (2002), *Global Catastrophic Risks* (ed., 2008), and *Human Enhancement* (ed., 2009).

**Selmer Bringsjord** is Professor of Cognitive Science and Professor of Computer Science at Rensselaer Polytechnic Institute. He has written numerous books, including *Superminds: People Harness Hypercomputation and More* (with M. Zenzen, 2003) and over 150 refereed papers.

**David Danks** is Associate Professor of Philosophy and Psychology at Carnegie Mellon University. He has developed multiple novel machine learning algorithms and has written numerous articles and book chapters on the intersection between machine learning and cognitive science. He is also the author of *Unifying the Mind: Cognitive Representations as Graphical Models* (2014).

**Keith Frankish** is Visiting Senior Research Fellow at The Open University UK and Adjunct Professor with the Brain and Mind Program in Neurosciences at the University of Crete. He is the author of *Mind and Supermind* (Cambridge, 2004) and *Consciousness* (2005). He is co-editor of *In Two Minds: Dual Processes and Beyond* (with Jonathan St. B. T. Evans, 2009), *New Waves in Philosophy of Action* (with Jesús H. Aguilar and Andrei A. Buckareff, 2010), and *The Cambridge Handbook of Cognitive Science* (with William M. Ramsey, Cambridge, 2012).

**Stan Franklin** is the W. Harry Feinstone Interdisciplinary Research Professor at the University of Memphis and Co-director of its Institute for Intelligent Systems. He is the author of *Artificial Minds* (1995) as well as numerous articles and book chapters on cognitive modeling, artificial general intelligence, and machine consciousness.

**Phil Husbands** is Research Professor of Artificial Intelligence and Co-founder of the Centre for Computational Neuroscience and Robotics at the University of Sussex. He has published numerous papers and edited several books on complex adaptive systems, both natural and artificial, bio-inspired approaches to robotics, modeling of neural systems, history of AI and Cybernetics, machine learning and creative systems.

**William M. Ramsey** is Associate Professor of Philosophy at the University of Nevada at Las Vegas. He is the author of *Representation Reconsidered* (Cambridge, 2007) and co-editor of *Philosophy and Connectionist Theory* (with David Rumelhart and Stephen Stich, 1991), *Rethinking Intuition* (with Michael DePaul, 1998), and *The Cambridge Handbook of Cognitive Science* (with Keith Frankish, Cambridge, 2012).

**William S. Robinson** is Emeritus Professor of Philosophy at Iowa State University. He is the author of *Computers, Minds & Robots* (1992), and two

---

more recent books: *Understanding Phenomenal Consciousness* (Cambridge, 2004) and a work for a general audience, *Your Brain and You: What Neuroscience Means for Us* (2010).

**Gerhard Sagerer** is Professor of Computer Science at the University of Bielefeld, Germany and head of the research group for Applied Informatics. He is author, co-author, or editor of several books and articles on image and speech understanding and the application of pattern understanding methods to natural science domains as well as social robotics.

**Matthias Scheutz** is Professor of Computer and Cognitive Science in the Department of Computer Science at Tufts University and Adjunct Professor in the Department of Psychology. He has over 200 peer-reviewed publications in artificial intelligence, artificial life, agent-based computing, natural language processing, cognitive modeling, robotics, human–robot interaction, and foundations of cognitive science.

**Ron Sun** is Professor of Cognitive Sciences at Rensselaer Polytechnic Institute. He is the author of *Integrating Rules and Connectionism for Robust Commonsense Reasoning* (1994) and *Duality of the Mind* (2001), and the editor of *The Cambridge Handbook of Computational Psychology* (Cambridge, 2008) and many other books, in addition to publishing numerous articles and book chapters.

**Markus Vincze** leads a group of researchers in the Vision for Robotics laboratory at Technische Universität Wien (TUW). With Gregory Hager he edited a book on Robust Vision for IEEE and is author or co-author of over 250 papers, specializing in particular on computer vision techniques for robotic systems situated in real-world environments.

**Sven Wachsmuth** is a lecturer at Bielefeld University and is currently heading the Central Lab Facilities of the Center of Excellence Cognitive Interaction Technology (CITEC). He is author or co-author of over 100 papers on human–robot interaction, especially high-level computer vision, as well as system integration and evaluation aspects.

**Yorick Wilks** is Professor of Artificial Intelligence at the University of Sheffield and Senior Research Scientist at Florida Institute for Human and Machine Cognition (IHMC). He is the co-editor (with Derek Partridge) of *The Foundations of Artificial Intelligence* (Cambridge, 1990), the editor of *Close Engagements with Artificial Companions* (2010), and the author of some 200 papers.

**Eliezer Yudkowsky** is Research Fellow at the Machine Intelligence Research Institute. His publications include “Timeless decision theory” and “Artificial intelligence as a positive and negative factor in global risk.”

## Acknowledgments

Many people have contributed to the production of this volume. First, we would like to thank the contributors for their expertise, hard work, patience, and responsiveness to our often exacting requirements. We are also grateful to Hilary Gaskin at Cambridge University Press for inviting us to compile the volume and for her patience when the project took longer than expected; to David Mackenzie for his careful management of the volume's production process; and to Anna Oxbury for her scrupulous and thoughtful copy-editing of the text.

We also wish to thank our home institutions. Keith is grateful to The Open University for support during some difficult personal times; to the Department of Philosophy and Social Studies at the University of Crete for welcoming him as a Visiting Researcher; to Professor Adonis Moschovakis, Professor Helen Savaki, and other members of the Brain and Mind Program at the University of Crete for helping him feel part of the cognitive science community in Crete; and, most importantly, to his partner Maria Kasmirli for her patience, support, and excellent advice on philosophical matters. William is grateful to both the University of Notre Dame and the University of Nevada at Las Vegas, and especially would like to thank Notre Dame for its Associative Professor's Special Leave, which helped support work on this project.