Advances in Interaction Studies

7

Future Robots

Towards a robotic science of human beings

Domenico Parisi

John Benjamins Publishing Company

Future Robots

Towards a robotic science of human beings

Domenico Parisi
Institute of Cognitive Sciences and Technologies,
National Research Council, Rome

John Benjamins Publishing Company
Amsterdam / Philadelphia

The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Parisi, Domenico, author.
Future Robots : Towards a robotic science of human beings / Domenico Parisi.
p. cm. (Advances in Interaction Studies, ISSN 1879-873X ; v. 7)
Includes bibliographical references and index.
1. Robotics--Human factors. 2. Androids. 3. Artificial intelligence. I. Title.
TJ211.49.P37 2014
629.8'92--dc23 2014008326
ISBN 978 90 272 0461 5 (Hb ; alk. paper)
ISBN 978 90 272 7008 5 (Eb)

© 2014 – John Benjamins B.V.
No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

Future Robots

Advances in Interaction Studies (AIS)

Advances in Interaction Studies (AIS) provides a forum for researchers to present excellent scholarly work in a variety of disciplines relevant to the advancement of knowledge in the field of interaction studies. The book series accompanies the journal *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems.*

The book series allows the presentation of research in the forms of monographs or edited collections of peer-reviewed material in English.

For an overview of all books published in this series, please see
http://benjamins.com/catalog/ais

Editors

Kerstin Dautenhahn
The University of Hertfordshire

Angelo Cangelosi
University of Plymouth

Editorial Board

Henrik Christensen
Georgia Institute of Technology

Harold Gouzoules
Emory University

Takayuki Kanda
ATR, Kyoto

Tetsuro Matsuzawa
Kyoto University

Giorgio Metta
IIT, Genoa

Adam Miklosi
Eötvös Loránd University

Robert W. Mitchell
Eastern Kentucky University

Chrystopher L. Nehaniv
University of Hertfordshire

Stefano Nolfi
CNR, Rome

Pierre-Yves Oudeyer
INRIA, Bordeaux

Irene M. Pepperberg
Harvard University & Brandeis University

Kerstin Severinson Eklundh
KTH. Stockholm

Stefan Wermter
University of Hamburg

Volume 7

Future Robots. Towards a robotic science of human beings
by Domenico Parisi

试读结束：需要全本请在线购买： www.ertongbook.com

To my wife Cristina

Preface

Finish what you began, and we shall be born
(Brunella Antomarini, The Maiden Machine)

Human beings are the greatest challenge for science. They are the most complicated of all entities that make up reality. They are embarrassing entities because they seem to be made of both physical matter and something else which is not physical matter. They are difficult to study with the necessary detachment which is required by science because scientists also are human beings. This explains why science knows and understands nature much better than human beings. One might think that one must only wait and, someday, science will understand human beings as it understands all other phenomena of reality. But this may not be true unless the science of human beings changes radically.

The goal of this book is to outline a new science of human beings: a robotic science of human beings. The premise on which this science is based is that we will really understand and explain human beings and their societies only if we succeed in constructing robots that behave like human beings and live in societies which are like human societies. If we succeed in constructing these robots, the theory – or, rather, the blueprint – which has been used to construct the robots captures what underlies human behaviour and human societies and explains them.

But the robots of a robotic science of human beings must be *human*, not *humanoid*, robots. Today's humanoid robots have the external appearance of human beings and do some simple things that human beings do such as grasping an object with their hand or walking on two legs. Human robots must progressively reproduce all that we know about human beings: their body, their brain, their genes, their environment, their evolutionary origins, how they develop during their life, how they acquire new behaviours through learning and imitation, their motivations and emotions, their mental life, their families, their cultures, their economic and political institutions, and how they and their societies have changed and continue to change in the course of time.

These phenomena are traditionally studied by separate scientific disciplines – biology, neuroscience, psychology, anthropology, sociology, economics, political science, history – and this is a problem because scientific disciplines divide reality into separate pieces but reality is not divided into separate pieces. Reality is a large ensemble of different phenomena which are all connected together and, very often, the phenomena studied by one discipline can only be understood and explained by taking into consideration the phenomena studied by another discipline. A robotic science of human beings is a non-disciplinary science of human beings. One and the same robot and one

and the same collection of robots must reproduce all the different human phenomena which today are studied by separate scientific disciplines.

Today's robots are mostly technologies with practical applications and almost all research money goes to robots as practical applications. Our robots do not have practical applications but they are purely scientific tools. The only reason why we want to construct them is that, by constructing them, we can better understand human beings and human societies. They are mirrors through which human beings will be able to see themselves. This means that this book is addressed more to students of human behaviour and human societies than to roboticists – although roboticists may find suggestions on how to construct new robots with practical applications.

But our robots do have one practical application, and this may be the most important application of robots: they can help human beings to better understand the difficult problems they face today and will face in the future and, perhaps, to find solutions to these problems.

This is an ambitious program of research and this is why human robots are future robots. The robots described in the book reproduce in a very simplified way some very limited aspects of human behaviour and human societies. And writing the book has made very clear to its author how many phenomena concerning human beings and human societies still remain to be reproduced by constructing robots. But it is possible to make the first steps towards the realization of the program, and this book describes these first steps.

Acknowledgements

The robots described in this book have been constructed by people who have worked or are still working at the Institute of Cognitive Sciences and Technologies, National Research Council, Rome. Here is a list of these people: Alberto Acerbi, Gianluca Baldassarre, Paolo Bartolomeo, Valerio Biscione, Anna Borghi, Daniele Caligiore, Angelo Cangelosi, Giovanni Sirio Carmantini, Federico Cecconi, Antonio Cerini, Federico Da Rold, Tommasino Ferrauto, Dario Floreano, Simone Giansante, Onofrio Gigliotta, Nicola Lettieri, Henrik Hautop Lund, Filippo Menczer, Orazio Miglino, Marco Mirolli, Francesco Natale, Stefano Nolfi, Fabio Paglieri, Giancarlo Petrosino, Fabio Ruini, Filippo Saglimbeni, Matthew Schlesinger, Massimiliano Ugolini, Alberto Venditti.

I thank Cristina Delogu, Nicola Lettieri, and Marco Mirolli for reading some of the chapters of the book and suggesting how to improve them. I am also very grateful to Cristina Delogu, Emanuela Parisi, and Giuseppe Polegri for their great help in preparing the book's figures.

Table of contents

Preface xi

CHAPTER 1

Robots as theories of behaviour 1

1. The problem with scientific theories of human beings 3
2. Theories as artefacts 6
3. Robots as practical applications and robots as science 13
4. One robot, many phenomena 16
5. Human, not humanoid, robots 24
6. This book 26

Appendix 28

CHAPTER 2

Robots that have motivations and emotions 33

1. The cognitive level and the motivational level of behaviour 33
2. Today's robots do not have motivations 36
3. Robots that have motivations 38
4. Robots that have emotions 55
5. Motivations, emotions, and covert attention 64
6. Internal robotics 68
7. Robots with a biological clock 75
8. The two halves of the mind 79

CHAPTER 3

How robots acquire their behaviour 81

1. Why learning? 81
2. The neural network of robots that learn 85
3. The genetic predictability of the environment 89
4. Living in genetically predictable environments 91
5. Living in genetically unpredictable environments 98
6. Learning in the experimental laboratory 100
7. Imprinting and learning from one's mother 107
8. The influence of learning on learning 112
9. Learning with an evolved neural architecture 113
10. On the limits of the robots described in this Chapter 116

CHAPTER 4
Robots that have language 121
1. The cognitive consequences of having language 121
2. Meaning as co-variation between sounds and nonlinguistic experiences 122
3. Classes of linguistic sounds 126
4. Language helps human beings to categorize their environment 131
5. The invention of language 141
6. Asymmetries between language production and language understanding 146
7. Robots that count and measure 149
8. On the limits of our robots that have language 154

CHAPTER 5
Robots with a mental life 159
1. Mental life as the self-generation of sensory inputs 160
2. Mental images 161
3. Robots that predict 163
4. Predicting and anticipating 167
5. Evaluating the predicted consequences of one's actions 169
6. Freedom of the will 171
7. Predicted sensory inputs replace missing sensory input 174
8. Other consequences of the ability to predict 179
9. Talking to oneself 182

CHAPTER 6
Social robots 187
1. There is no social robotics today 187
2. Living together 190
3. Why *not* live together 196
4. Socially damaging behaviours and how to contain them 199
5. Why live together: Groups as information centres 211
6. Living in small communities and living in large communities 213
7. The social environment is very different from the non-social environment 215

CHAPTER 7
Robotic families 221
1. Genetic families and social families 221
2. Mothers and daughters 223
3. Grandmothers 227
4. Sisters 231

试读结束：需要全本请在线购买： www.ertongbook.com

5. Males and females 235
6. Homes 253
7. Conclusions 257

CHAPTER 8
Robots that learn from other robots and develop cultures and technologies 259
1. Learning from others 259
2. The cultural emergence of behaviours 263
3. Staying near to others in order to learn from them 266
4. Should adolescents learn from adults or from other adolescents? 269
5. The evolution of artefacts 274
6. Cultures 280
7. The diminishing role of space in cultural processes 286
8. Cultural globalization 289
9. The expansion of Indo-European languages 296

CHAPTER 9
Robot that own things 301
1. External stores 302
2. Individual stores 303
3. Family stores 313
4. Central stores 330

CHAPTER 10
Political robotics 339
1. Chief robots 340
2. Chiefs and leaders 344
3. No chiefs 350
4. Historical changes in human settlements in ancient Southern Etruria 356

CHAPTER 11
Robotic economies 361
1. Goods 361
2. Value 363
3. Specialization and the exchange of goods 372
4. Merchant robots 383
5. The emergence of money 386
6. Money and the price of goods 390
7. Entrepreneur robots and worker robots 394
8. The increase in the number of things that are goods for human beings 401

CHAPTER 12
Individually different robots and robots with pathologies 407
1. Current robotics ignores inter-individual differences 407
2. Fitness is not enough 410
3. Motivations as a source of inter-individual differences 416
4. Learning and inter-individual differences 417
5. Why construct robots that have pathologies? 419
6. Neurological pathologies and psychiatric pathologies 420
7. Prediction, diagnosis, prognosis, and therapy 424

CHAPTER 13
Robots that have art, religion, philosophy, science, and history 427
1. Explaining this Chapter 428
2. Robots that have art 428
3. Robots that have religion 435
4. Doing metaphysics with robots 436
5. Robots that do science 439
6. Robots that have history 444

CHAPTER 14
Human robots are future robots 451
1. A new science of human beings and its problems 451
2. What has still to be done 452

CHAPTER 15
How human robots can be useful to human beings 461
1. Human robots should be practically useful to human beings 461
2. Robots make it possible for human beings to know themselves without being influenced by their desires and fears 462
3. Robots can help human beings to identify possible conflicts between biology and culture and the consequences of these conflicts or their happiness 463
4. Robots can help human beings to understand the impact of science on their life 467
5. Robotic societies as tools for voters 472
6. Difficult Problems 473

References and additional readings 479

Index 489

CHAPTER 1

Robots as theories of behaviour

Martian Eye (ME) is a scientist from Mars which one day arrives on the Earth and, of all the things which exist on the Earth, it decides to study human beings. There are no particular reasons why ME chooses to study human beings rather than some other species of animals except that human beings are particularly complex animals and, therefore, they represent a greater challenge for ME. But, to tell the truth, there is one reason why ME chooses to study human beings. Human beings have many desires and many fears concerning themselves and these desires and fears obscure the knowledge that they have of themselves. They think they are what they would like to be and they are not what they don't like to be. And they are inevitably anthropocentric. They see themselves as the centre of the world. But they are not the centre of the world and their anthropocentrism distorts the vision that they have of themselves. Scientists are human beings and, for them too, it is difficult to see human beings as only one of the many things that make up reality and to study human beings with the necessary detachment which is required by science. ME is not a human being and it believes that this is very important to understand human beings in the way in which science understands every other phenomenon of reality.

ME knows that objective and quantitative data are the basis of science but data is not enough. Science must answer the question "What?" but it must also answer the question "Why?", and to answer the question "Why?" science needs theories which explain the data. Human scientists express their theories by using words or mathematical symbols. ME is a Martian scientist and it has its own way of doing science. To express its theories, ME constructs artefacts that incorporate its theories and behave like what the theories must explain. ME's science is based on the principle "Whatever X is, to understand X you must reproduce X in an artefact". Only if its artefacts do what human beings do, ME is satisfied that it has understood and explained human beings.

When it arrives on the Earth, ME finds that human scientists already construct physical artefacts that resemble human beings and other animals, and they call these artefacts "robots". But human scientists mostly construct robots because robots have practical application and economic value. ME's robots are purely scientific tools that must make it possible to better understand what human beings are. Human scientists call their robots which resemble human beings "humanoid". To distinguish its robots from "humanoid" robots, ME calls its robots "human".

Constructing human robots that behave like human beings and live in societies which are like human societies is a very difficult task and ME is well aware that its robots often resemble much simpler animals and that really human robots are future robots. But ME's program of research is clear.

Another thing which makes ME's science different from how science is done on the Earth is that for ME the divisions among the different scientific disciplines do not exist. Reality is a large ensemble of different phenomena but these phenomena are all linked together and, very often, to understand the phenomenon studied by one discipline it is necessary to take into consideration the phenomena studied by other disciplines. ME's science is a non-disciplinary science of human beings. The robots constructed by ME must reproduce all aspects of human behaviour, from its physical and biological bases to all the individual and social phenomena which, on the Earth, are studied by psychologists, anthropologists, sociologists, economists, and political scientists.

ME also believes that to understand human beings it is necessary to know how they have become what they are: how they have evolved from nonhuman ancestors, how they mature and develop during their life, what they learn from their experiences, how human cultures, economies, and political institutions were born and how they have changed and continue to change. This implies that ME cannot design or program its robots but they must autonomously become what they are.

ME is well aware that, like all theories, its robots simplify with respect to real human beings and real human societies but it also knows that, to be useful, scientific theories must necessarily simplify with respect to reality. The problem is that they must make the appropriate simplifications, and ME believes it is more probable that its robots make the appropriate simplifications if one and the same robot and one and the same robotic society reproduce not one single phenomenon but as many phenomena as possible about human beings and human societies.

ME's research "philosophy" can be summarized in three principles:

- If you want to understand human beings, construct robots that behave like human beings.
- If you want to understand human beings, construct robots that not only reproduce human beings but also reproduce how human beings have become what they are.
- If you want to understand human beings, construct robots that reproduce as many different phenomena as possible concerning human beings.

ME is not an imaginary scientist like the Martian scientists of science fiction and thought experiments. ME is a real scientist and it constructs real robots – although, for now, most of its robots are simulated in a computer. And, being a real scientist, ME wants to discuss and collaborate with human scientists.

1. The problem with scientific theories of human beings

Imagine you are interested in how human beings behave and how their societies function and that you want to understand human behaviour and human societies as science knows and understands all other phenomena of reality. What must you do? You must collect objective and possibly quantitative data on the phenomena that interest you because, for science, reality is, first of all, empirical data. But empirical data is not enough. You must also formulate theories that explain the data. Traditionally, scientific theories of human behaviour and human societies are formulated in words but words have serious limitations as tools for expressing scientific theories. Words have unclear meanings, they have different meanings for different people, and they are often value-based and emotionally charged – and these properties of words are not good for science. A crucial requirement of scientific theories is that from a scientific theory it must be possible to derive specific and unambiguous predictions to be compared with the empirical data. Verbally expressed theories are generally unable to generate specific and unambiguous empirical predictions.

Scientists can agree on these limitations of words as scientific tools and they can try to overcome them by defining the words used in their theories or by specifying in which sense they use them. Or they may create new words and define them by using existing words. But this, clearly, is a circular strategy. Words are defined or re-defined by using other words, and these other words also have unclear meanings, mean different things to different people, and are value-based and emotionally charged. So, expressing scientific theories by using words remains a problem for science because verbally formulated theories leave us with what the English poet Thomas Stearns Eliot called "the intolerable wrestle with words and their meanings".

There is another property of words which is a problem for science. Given a word, there is a tendency to think that there exists an entity which corresponds to the word. This may be all good for the everyday use of language but it is not good for science. Given a word of the common language, it is not automatic that the explanatory apparatus of science should include the entity designated by the word. For example, the words "belief" and "goal" may be perfectly appropriate for everyday life. But a good scientific theory of human behaviour and human societies may *not* need these words and may better explain the phenomena to which in everyday life we refer by using these words by *not* using these words.

Expressing scientific theories by using words explains why the paradigmatic science of human behaviour, psychology, is only a half-science. The human mind has been studied by philosophers for thousands of years, and philosophers have studied the mind in their characteristic way: by doing conceptual analyses, by reasoning and presenting arguments, and by discussing with colleagues. More than one century ago, psychology was born, and psychology was a revolution in the study of the mind. Psychologists investigate the mind by observing behaviours in the controlled

conditions of an experimental laboratory and in other objective and quantitative ways. Therefore, they can claim they have created a *science* of the mind which uses the same methods of the natural sciences and is not confined, like philosophy, to conceptual analyses, argumentation, and discussion with colleagues. But science does not need only objective and quantitative data. It also needs theories that explain the data. The problem with psychology is that while its empirical methods are those of science, its theoretical vocabulary is still a philosophical vocabulary. If you read psychological books and papers, you find the same vocabulary of ancient or recent philosophical treatises, or the vocabulary that people use in their everyday life to talk about their behaviour and the behaviour of others. Here is a sample of this vocabulary: sensation, perception, attention, memory, thinking, reasoning, predicting, planning, motivation, emotion, all the different motivations and emotions, representation, concept, category, meaning, object, property, action, intention, goal, awareness, consciousness. These words cannot be the vocabulary of a science because they have unclear meanings, they have different meanings for different scientists and, even when one scientist provides a precise definition of one particular word, the definition is not generally adopted by his or her colleagues. Therefore, using words to express psychological theories causes endless and often useless discussions and it is a serious obstacle to the progress of psychology. Science is based on a virtuous circle between theories and data: theories must predict data, and data must confirm theories. Psychologists find it difficult to establish this virtuous circle. Theories are formulated and discussed but it is difficult to arrive to a consensus on which theory is the right one. Data are collected, often in the controlled conditions of an experimental laboratory, but they are rarely really illuminated and explained by a theory. This is why psychology is only a scientific half-revolution. Psychologists adopt the empirical methods of science but they continue to talk about human behaviour and the human mind by using the vocabulary of philosophy or everyday life.

Psychology is the paradigmatic science of human behaviour but many other disciplines study human beings: anthropology, sociology, economics, political science, history. These disciplines are called the social sciences and with the social sciences the problems are even more serious. Not only these sciences have verbally formulated theories but, since they generally cannot study the phenomena that interest them in the controlled conditions of the experimental laboratory, their empirical data are often unclear – or nonexistent. Therefore, for the social sciences, the necessary dialogue between theories and facts which is the foundation of science cannot be established because neither of the two interlocutors actually exists. (The science of economics is the only exception to this rule but the science of economics has its own problems. See later in this section.)

Another weakness of the social sciences is that they often resemble philosophy more than science. Much of the work of social scientists is an exegesis of what other social scientists have written and social scientists give the impression that they do not

试读结束：需要全本请在线购买： www.ertongbook.com