Preface

"Only a small community has concentrated on general intelligence. No one has tried to make a thinking machine ...
The bottom line is that we really haven't progressed too far toward a truly intelligent machine. We have collections of dumb specialists in small domains; the true majesty of general intelligence still awaits our attack. ...
We have got to get back to the deepest questions of AI and general *intelligence..."*

– Marvin Minsky
as interviewed in *Hal's Legacy*, edited by David Stork, 2000.

Our goal in creating this edited volume has been to fill an apparent gap in the scientific literature, by providing a coherent presentation of a body of contemporary research that, in spite of its integral importance, has hitherto kept a very low profile within the scientific and intellectual community. This body of work has not been given a name before; in this book we christen it "Artificial General Intelligence" (AGI). What distinguishes AGI work from run-of-the-mill "artificial intelligence" research is that it is explicitly focused on engineering general intelligence in the short term. We have been active researchers in the AGI field for many years, and it has been a pleasure to gather together papers from our colleagues working on related ideas from their own perspectives. In the Introduction we give a conceptual overview of the AGI field, and also summarize and interrelate the key ideas of the papers in the subsequent chapters.

Of course, "general intelligence" does not mean exactly the same thing to all researchers. In fact it is not a fully well-defined term, and one of the issues raised in the papers contained here is how to define general intelligence in a way that provides maximally useful guidance to practical AI work. But,

nevertheless, there is a clear qualitative meaning to the term. What is meant by AGI is, loosely speaking, AI systems that possess a reasonable degree of self-understanding and autonomous self-control, and have the ability to solve a variety of complex problems in a variety of contexts, and to learn to solve new problems that they didnt know about at the time of their creation. A marked distinction exists between practical AGI work and, on the other hand:

- Pragmatic but specialized "narrow AI" research which is aimed at creating programs carrying out specific tasks like playing chess, diagnosing diseases, driving cars and so forth (most contemporary AI work falls into this category.)
- Purely theoretical AI research, which is aimed at clarifying issues regarding the nature of intelligence and cognition, but doesnt involve technical details regarding actually realizing artificially intelligent software.

Some of the papers presented here come close to the latter (purely theoretical) category, but we have selected them because the theoretical notions they contain seem likely to lead to such technical details in the medium-term future, and/or resonate very closely with the technical details of AGI designs proposed by other authors.

The audience we intend to reach includes the AI community, and also the broader community of scientists and students in related fields such as philosophy, neuroscience, linguistics, psychology, biology, sociology, anthropology and engineering. Significantly more so than narrow AI, AGI is interdisciplinary in nature, and a full appreciation of the general intelligence problem and its various potential solutions requires one to take a wide variety of different perspectives.

Not all significant AGI researchers are represented in these pages, but we have sought to bring together a multiplicity of perspectives, including many that disagree with our own. Bringing a diverse body of AGI research together in a single volume reveals the common themes among various researchers work, and makes clear what the big open questions are in this vital and critical area of research. It is our hope that this book will interest more researchers and students in pursuing AGI research themselves, thus aiding in the progress of science.

In the three years that this book has been in the making, we have noticed a significant increase in interest in AGI-related research within the academic AI community, including a number of small conference workshops with titles related to "Human-Level Intelligence." We consider this challenge to the overwhelming dominance of narrow-AI an extremely positive move; however, we submit that "Artificial General Intelligence" is a more sensible way to conceptualize the problem than "Human-Level Intelligence." The AGI systems and approaches described in these pages are not necessarily oriented towards emulating the human brain; and given the heterogeneity of the human mind/brain and its highly various levels of competence at various sorts of tasks, it seems very difficult to define "Human-Level Intelligence" in any way that is generally

applicable to AI systems that are fundamentally non-human-like in conception. On the other hand, the work of Hutter and Schmidhuber reported here provides a reasonable, abstract mathematical characterization of general intelligence which, while not in itself providing a practical approach to AGI design and engineering, at least provides a conceptually meaningful formalization of the ultimate goal of AGI work.

The grand goal of AGI remains mostly unrealized, and how long it will be until this situation is remedied remains uncertain. Among scientists who believe in the fundamental possibility of strong AI, the most optimistic serious estimates we have heard are in the range of 5-10 years, and the most pessimistic are in the range of centuries. While none of the articles contained here purports to present a complete solution to the AGI problem, we believe that they collectively embody meaningful conceptual progress, and indicate clearly that the direct pursuit of AGI is an endeavor worthy of significant research attention.

Contents

http://www.springer.com/978-3-540-23733-4

Artificial General Intelligence
Goertzel, B.; Pennachin, C. (Eds.)
2007, XVI, 509 p., Hardcover
ISBN: 978-3-540-23733-4