

Preface

Welcome to the second volume of *Engineering General Intelligence*! This is the second half of a two-part technical treatise aimed at outlining a practical approach to engineering software systems with general intelligence at the human level and ultimately beyond.

Our goal here is an ambitious one and not a modest one: machines with flexible problem-solving ability, open-ended learning capability, creativity and eventually, their own kind of genius.

Part 1 set the stage, dealing with a variety of general conceptual issues related to the engineering of advanced AGI, as well as presenting a brief overview of the CogPrime design for Artificial General Intelligence. Now here in Part 2 we plunge deep into the nitty-gritty, and describe the multiple aspects of CogPrime with a fairly high degree of detail.

First we describe the CogPrime software architecture and knowledge representation in detail; then we review the “cognitive cycle” via which CogPrime perceives and acts in the world and reflects on itself. We then turn to various forms of learning: procedural, declarative (e.g., inference), simulative, and integrative. Methods of enabling natural language functionality in CogPrime are then discussed; and the volume concludes with a chapter summarizing the argument that CogPrime can lead to human-level (and eventually perhaps greater) AGI, and a chapter giving a “thought experiment” describing the internal dynamics via which a completed CogPrime system might solve the problem of obeying the request “Build me something with blocks that I haven’t seen before.”

Reading this book before *Engineering General Intelligence, Part 1* first is not especially recommended, since the prequel not only provides the context for this one, but also defines a number of specific terms and concepts that are used here without explanation (for example, Part 1 has an extensive Glossary). However, the impatient reader who has not mastered Part 1, or the reader who has finished Part 1 but is tempted to hop through Part 2 nonlinearly, might wish to first skim the final chapter, and then return to reading in linear order.

While the majority of the text here was written by the lead Author Ben Goertzel, the overall work and underlying ideas have been very much a team effort, with major input from the secondary authors Cassio Pennachin and Nil

Geisweiller, and large contributions from various other contributors as well. Many chapters have specifically indicated Co-authors; but the contributions from various collaborating researchers and engineers go far beyond these. The creation of the AGI approach and design presented here is a process that has occurred over a long period of time among a community of people; and this book is in fact a quite partial view of the existent body of knowledge and intuition regarding CogPrime. For example, beyond the ideas presented here, there is a body of work on the OpenCog wiki site, and then the OpenCog codebase itself.

More extensive introductory remarks may be found in the Preface of Part 1, including a brief history of the book and acknowledgements to some of those who helped inspire it.

Also, one brief comment from the Preface of Part 1 bears repeating: At several places in this volume, as in its predecessor, we will refer to the “current” CogPrime implementation (in the OpenCog framework); in all cases this refers to the OpenCog software system as of late 2013.

We fully realize that this book is not “easy reading,” and that the level and nature of exposition varies somewhat from chapter to chapter. We have done our best to present these very complex ideas as clearly as we could, given our own time constraints, and the lack of commonly understood vocabularies for discussing many of the concepts and systems involved. Our hope is that the length of the book, and the conceptual difficulty of some portions, will be considered as compensated by the interest of the ideas we present. For, make no mistake—for all their technicality and subtlety, we find the ideas presented here incredibly exciting. We are talking about no less than the creation of machines with intelligence, creativity, and genius equaling and ultimately exceeding that of human beings.

This is, in the end, the kind of book that we (the authors) all hoped to find when we first entered the AI field: a reasonably detailed description of how to go about creating thinking machines. The fact that so few treatises of this nature, and so few projects explicitly aimed at the creation of advanced AGI, exist, is something that has perplexed us since we entered the field. Rather than just complain about it, we have taken matters into our own hands, and worked to create a design and a codebase that we believe capable of leading to human-level AGI and beyond.

We feel tremendously fortunate to live in times when this sort of pursuit can be discussed in a serious, scientific way.

Online Appendices

Just one more thing before getting started! This book originally had even more chapters than the ones currently presented in Parts 1 and 2. In order to decrease length and increase focus, however, a number of chapters dealing with peripheral—yet still relevant and interesting—matters were moved to online appendices.

These may be downloaded in a single PDF file at http://goertzel.org/engineering_general_intelligence_appendices_B-H.pdf. The titles of these appendices are:

- Appendix A: Possible Worlds Semantics and Experiential Semantics
- Appendix B: Steps Toward a Formal Theory of Cognitive Structure and Dynamics
- Appendix C: Emergent Reflexive Mental Structures
- Appendix D: GOLEM: Toward an AGI Meta-Architecture Enabling Both Goal Preservation and Radical Self-Improvement
- Appendix E: Lojban++: A Novel Linguistic Mechanism for Teaching AGI Systems
- Appendix F: Possible Worlds Semantics and Experiential Semantics
- Appendix G: PLN and the Brain
- Appendix H: Propositions About Environments in Which CogPrime Components Are Useful

None of these are critical to understanding the key ideas in the book, which is why they were relegated to online appendices. However, reading them will deepen your understanding of the conceptual and formal perspectives underlying the CogPrime design. These appendices are referred to here and there in the text of the main book.

September 2013

Ben Goertzel

Engineering General Intelligence, Part 2
The CogPrime Architecture for Integrative, Embodied
AGI

Goertzel, B.; Pennachin, C.; Geisweiller, N.
2014, XXII, 562 p. 42 illus., 9 illus. in color., Hardcover
ISBN: 978-94-6239-029-4
A product of Atlantis Press