

# Texts in Theoretical Computer Science

## An EATCS Series

Editors: W. Brauer G. Rozenberg A. Salomaa

On behalf of the European Association  
for Theoretical Computer Science (EATCS)

---

Advisory Board: G. Ausiello M. Broy C.S. Calude  
A. Condon D. Harel J. Hartmanis T. Henzinger  
J. Hromkovič N. Jones T. Leighton M. Nivat  
C. Papadimitriou D. Scott

Marcus Hutter

# Universal Artificial Intelligence

Sequential Decisions  
Based on Algorithmic Probability

 Springer

*Author*

Dr. Marcus Hutter  
Istituto Dalle Molle  
di Studi sull'Intelligenza  
Artificiale (IDSIA)  
Galleria 2  
CH-6928 Manno-Lugano  
Switzerland  
marcus@idsia.ch  
www.idsia.ch/~marcus

*Series Editors*

Prof. Dr. Wilfried Brauer  
Institut für Informatik der TUM  
Boltzmannstr. 3, 85748 Garching, Germany  
Brauer@informatik.tu-muenchen.de  
  
Prof. Dr. Grzegorz Rozenberg  
Leiden Institute of Advanced Computer Science  
University of Leiden  
Niels Bohrweg 1, 2333 CA Leiden, The Netherlands  
rozenber@liacs.nl  
  
Prof. Dr. Arto Salomaa  
Turku Centre for Computer Science  
Lemminkäisenkatu 14 A, 20520 Turku, Finland  
asalomaa@utu.fi

Library of Congress Control Number: 2004112980

ACM Computing Classification (1998): I.3, I.2.6, F.0, F.1.3, F.4.1, E.4, G.3

ISBN 3-540-22139-5 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable for prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media  
springeronline.com

© Springer-Verlag Berlin Heidelberg 2005

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: KünkelLopka, Heidelberg  
Typesetting: by the author  
Production: LE-TeX Jelonek, Schmidt & Vöckler GbR, Leipzig  
Printed on acid-free paper 45/3142/YL - 5 4 3 2 1 0

---

## Preface

**Personal motivation.** The dream of creating artificial devices that reach or outperform human intelligence is an old one. It is also one of the dreams of my youth, which have never left me. What makes this challenge so interesting? A solution would have enormous implications on our society, and there are reasons to believe that the AI problem can be solved in my expected lifetime. So, it's worth sticking to it for a lifetime, even if it takes 30 years or so to reap the benefits.

**The AI problem.** The science of artificial intelligence (AI) may be defined as the construction of intelligent systems and their analysis. A natural definition of a *system* is anything that has an input and an output stream. Intelligence is more complicated. It can have many faces like creativity, solving problems, pattern recognition, classification, learning, induction, deduction, building analogies, optimization, surviving in an environment, language processing, and knowledge. A formal definition incorporating every aspect of intelligence, however, seems difficult. Most, if not all known facets of intelligence can be formulated as goal driven or, more precisely, as maximizing some utility function. It is, therefore, sufficient to study goal-driven AI; e.g. the (biological) goal of animals and humans is to survive and spread. The goal of AI systems should be to be useful to humans. The problem is that, except for special cases, we know neither the utility function nor the environment in which the agent will operate in advance. The major goal of this book is to develop a theory that solves these problems.

**The nature of this book.** The book is theoretical in nature. For most parts we assume availability of unlimited computational resources. The first important observation is that this does not make the AI problem trivial. Playing chess optimally or solving NP-complete problems become trivial, but driving a car or surviving in nature do not. This is because it is a challenge itself to well-define the latter problems, not to mention presenting an algorithm. In other words: The AI problem has not yet been well defined. One may view the book as a suggestion and discussion of such a mathematical definition of AI.

**Extended abstract.** The *goal* of this book is to develop a universal theory of sequential decision making akin to Solomonoff's celebrated universal theory of induction. Solomonoff derived an optimal way of predicting future data, given

previous observations, provided the data is sampled from a computable probability distribution. Solomonoff's unique predictor is universal in the sense that it applies to every prediction task and is the output of a universal Turing machine with random input. We extend this approach to derive an optimal rational reinforcement learning agent, called AIXI, embedded in an unknown environment. The *main idea* is to replace the unknown environmental distribution  $\mu$  in the Bellman equations by a suitably generalized universal distribution  $\xi$ . The state space is the space of complete histories. AIXI is a universal theory without adjustable parameters, making no assumptions about the environment except that it is sampled from a computable distribution. From an algorithmic complexity perspective, the AIXI model generalizes optimal passive universal induction to the case of active agents. From a decision-theoretic perspective, AIXI is a suggestion of a new (implicit) "learning" algorithm, which may overcome all (except computational) problems of previous reinforcement learning algorithms.

*Chapter 1.* We start with a survey of the contents and main results in this book.

*Chapter 2.* How and in which sense induction is possible at all has been subject to long philosophical controversies. Highlights are Epicurus' principle of multiple explanations, Occam's razor, and Bayes' rule for conditional probabilities. Solomonoff elegantly unified all these aspects into one formal theory of inductive inference based on a universal probability distribution  $\xi$ , which is closely related to Kolmogorov complexity  $K(x)$ , the length of the shortest program computing  $x$ . We classify the (non)existence of universal priors for several generalized computability concepts.

*Chapter 3.* We prove rapid convergence of  $\xi$  to the unknown true environmental distribution  $\mu$  and tight loss bounds for arbitrary bounded loss functions and finite alphabet. We show Pareto optimality of  $\xi$  in the sense that there is no other predictor that performs better or equal in all environments and strictly better in at least one. Finally, we give an Occam's razor argument showing that predictors based on  $\xi$  are optimal. We apply the results to games of chance and compare them to predictions with expert advice. All together this shows that Solomonoff's induction scheme represents a universal (formal, but incomputable) solution to all *passive* prediction problems.

*Chapter 4.* Sequential decision theory provides a framework for finding optimal reward-maximizing strategies in *reactive* environments (e.g. chess playing as opposed to weather forecasting), assuming the environmental probability distribution  $\mu$  is known. We present this theory in a very general form (called  $AI\mu$  model) in which actions and observations may depend on arbitrary past events. We clarify the connection to the Bellman equations and discuss minor parameters including (the size of) the I/O spaces and the lifetime of the agent and their universal choice which we have in mind. Optimality of  $AI\mu$  is obvious by construction.

*Chapter 5.* Reinforcement learning algorithms are usually used in the case of unknown  $\mu$ . They can succeed if the state space is either small or has ef-

fectively been made small by generalization techniques. The algorithms work only in restricted, (e.g. Markovian) domains, have problems with optimally trading off exploration versus exploitation, have nonoptimal learning rate, are prone to diverge, or are otherwise ad hoc. The formal solution proposed in this book is to generalize the universal prior  $\xi$  to include actions as conditions and replace  $\mu$  by  $\xi$  in the  $\text{AI}\mu$  model, resulting in the AIXI model, which we claim to be universally optimal. We investigate what we can expect from a universally optimal agent and clarify the meanings of *universal*, *optimal*, etc. We show that a variant of AIXI is self-optimizing and Pareto optimal.

*Chapter 6.* We show how a number of AI problem classes fit into the general AIXI model. They include sequence prediction, strategic games, function minimization, and supervised learning. We first formulate each problem class in its natural way for known  $\mu$ , and then construct a formulation within the  $\text{AI}\mu$  model and show their equivalence. We then consider the consequences of replacing  $\mu$  by  $\xi$ . The main goal is to understand in which sense the problems are solved by AIXI.

*Chapter 7.* The major drawback of AIXI is that it is incomputable, or more precisely, only asymptotically computable, which makes an implementation impossible. To overcome this problem, we construct a modified model  $\text{AIXI}t_l$ , which is still superior to any other time  $t$  and length  $l$  bounded algorithm. The computation time of  $\text{AIXI}t_l$  is of the order  $t \cdot 2^l$ . A way of overcoming the large multiplicative constant  $2^l$  is presented at the expense of an (unfortunately even larger) additive constant. The constructed algorithm  $M_p^\varepsilon$  is capable of solving all well-defined problems  $p$  as quickly as the fastest algorithm computing a solution to  $p$ , save for a factor of  $1+\varepsilon$  and lower-order additive terms. The solution requires an implementation of first-order logic, the definition of a universal Turing machine within it and a proof theory system.

*Chapter 8.* Finally we discuss and remark on some otherwise unmentioned topics of general interest. We also critically review what has been achieved in this book, including assumptions, problems, limitations, performance, and generality of AIXI in comparison to other approaches to AI. We conclude the book with some less technical remarks on various philosophical issues.

**Prerequisites.** I have tried to make the book as self-contained as possible. In particular, I provide all necessary background knowledge on algorithmic information theory in Chapter 2 and sequential decision theory in Chapter 4. Nevertheless, some prior knowledge in these areas could be of some help. The chapters have been designed to be readable independently of one another (after having read Chapter 1). This necessarily implies minor repetitions. Additional information to the book (FAQs, errata, prizes, ...) is available at <http://www.idsia.ch/~marcus/ai/uaibook.htm>.

**Problem classification.** Problems are included at the end of each chapter of different motivation and difficulty. We use Knuth's rating scheme for exercises [Knu73] in slightly adapted form (applicable if the material in the corresponding chapter has been understood). In-between values are possible.

C00 *Very easy*. Solvable from the top of your head.

C10 *Easy*. Needs 15 minutes to think, possibly pencil and paper.

C20 *Average*. May take 1–2 hours to answer completely.

C30 *Moderately difficult or lengthy*. May take several hours to a day.

C40 *Quite difficult or lengthy*. Often a significant research result.

C50 *Open research problem*. An obtained solution should be published.

The rating is possibly supplemented by the following qualifier(s):

- i* Especially *interesting* or *instructive* problem.
- m* Requires more or higher *math* than used or developed here.
- o* *Open* problem; could be worth publishing; see web for prizes.
- s* *Solved* problem with published solution.
- u* *Unpublished* result by the author.

The problems represent an important part of this book. They have been placed at the end of each chapter in order to keep the main text better focused.

**Acknowledgements.** I would like to thank all those people who in one way or another have contributed to the success of this book. For interesting discussions I am indebted to Jürgen Schmidhuber, Ray Solomonoff, Paul Vitányi, Peter van Emde Boas, Richard Sutton, Leslie Kaelbling, Leonid Levin, Peter Gács, Wilfried Brauer, and many others. Shane Legg, Jan Poland, Viktor Zhumatiy, Alexey Chernov, Douglas Eck, Ivo Kwee, Philippa Hutter, Paul Vitányi, and Jürgen Schmidhuber gave valuable feedback on drafts of the book. Thanks also collectively to all other IDSIANies and to the Springer team for the pleasant working atmosphere and their support. This book would not have been possible without the financial support of the SNF (grant no. 2000-61847.00). Thanks also to my father, who taught me to think sharply and to my mother who taught me to do what one enjoys. Finally, I would like to thank my wife and children who patiently supported my decision to write this book.

---

# Contents

<b>0</b>	<b>Meta Contents</b>	iii
	Preface	v
	Contents	ix
	Tables, Figures, Theorems, ...	xv
	Notation	xvii
<b>1</b>	<b>A Short Tour Through the Book</b>	1
1.1	Introduction	2
1.2	Simplicity & Uncertainty	3
1.2.1	Introduction	3
1.2.2	Algorithmic Information Theory	4
1.2.3	Uncertainty & Probabilities	5
1.2.4	Algorithmic Probability & Universal Induction	6
1.2.5	Generalized Universal (Semi)Measures	7
1.3	Universal Sequence Prediction	7
1.3.1	Setup & Convergence	8
1.3.2	Loss Bounds	8
1.3.3	Optimality Properties	9
1.3.4	Miscellaneous	10
1.4	Rational Agents in Known Probabilistic Environments	11
1.4.1	The Agent Model	11
1.4.2	Value Functions & Optimal Policies	11
1.4.3	Sequential Decision Theory & Reinforcement Learning	12
1.5	The Universal Algorithmic Agent AIXI	13
1.5.1	The Universal AIXI Model	13
1.5.2	On the Optimality of AIXI	14
1.5.3	Value-Related Optimality Results	15
1.5.4	Markov Decision Processes	17
1.5.5	The Choice of the Horizon	18
1.6	Important Environmental Classes	18
1.6.1	Introduction	18
1.6.2	Sequence Prediction (SP)	19
1.6.3	Strategic Games (SG)	19
1.6.4	Function Minimization (FM)	19
1.6.5	Supervised Learning from Examples (EX)	19
1.6.6	Other Aspects of Intelligence	20
1.7	Computational Aspects	20



1.7.1	The Fastest & Shortest Algorithm for All Problems . . . .	20
1.7.2	Time-Bounded AIXI Model . . . . .	22
1.8	Discussion . . . . .	24
1.9	History & References . . . . .	26
<b>2</b>	<b>Simplicity &amp; Uncertainty</b> . . . . .	<b>29</b>
2.1	Introduction . . . . .	30
2.1.1	Examples of Induction Problems . . . . .	30
2.1.2	Ockham, Epicurus, Hume, Bayes, Solomonoff . . . . .	31
2.1.3	Problem Setup . . . . .	32
2.2	Algorithmic Information Theory . . . . .	33
2.2.1	Definitions and Notation . . . . .	33
2.2.2	Turing Machines . . . . .	34
2.2.3	Kolmogorov Complexity . . . . .	36
2.2.4	Computability Concepts . . . . .	38
2.3	Uncertainty & Probabilities . . . . .	40
2.3.1	Frequency Interpretation: Counting . . . . .	40
2.3.2	Objective Interpretation: Uncertain Events . . . . .	41
2.3.3	Subjective Interpretation: Degrees of Belief . . . . .	43
2.3.4	Determining Priors . . . . .	45
2.4	Algorithmic Probability & Universal Induction . . . . .	45
2.4.1	The Universal Prior $M$ . . . . .	45
2.4.2	Universal Sequence Prediction . . . . .	47
2.4.3	Universal (Semi)Measures . . . . .	48
2.4.4	Martin-Löf Randomness . . . . .	54
2.5	History & References . . . . .	55
2.6	Problems . . . . .	60
<b>3</b>	<b>Universal Sequence Prediction</b> . . . . .	<b>65</b>
3.1	Introduction . . . . .	67
3.2	Setup and Convergence . . . . .	68
3.2.1	Random Sequences . . . . .	68
3.2.2	Universal Prior Probability Distribution . . . . .	69
3.2.3	Universal Posterior Probability Distribution . . . . .	70
3.2.4	Convergence of Random Sequences . . . . .	71
3.2.5	Distance Measures between Probability Distributions . .	72
3.2.6	Convergence of $\xi$ to $\mu$ . . . . .	74
3.2.7	Convergence in Martin-Löf Sense . . . . .	76
3.2.8	The Case where $\mu \notin \mathcal{M}$ . . . . .	80
3.2.9	Probability Classes $\mathcal{M}$ . . . . .	81
3.3	Error Bounds . . . . .	82
3.3.1	Bayes Optimal Predictors . . . . .	82
3.3.2	Total Expected Numbers of Errors . . . . .	82
3.3.3	Proof of Theorem 3.36 . . . . .	84
3.4	Loss Bounds . . . . .	86

3.4.1	Unit Loss Function . . . . .	86
3.4.2	Loss Bound of Merhav & Feder . . . . .	88
3.4.3	Example Loss Functions . . . . .	89
3.4.4	Proof of Theorem 3.48 . . . . .	89
3.4.5	Convergence of Instantaneous Losses . . . . .	91
3.4.6	General Loss . . . . .	92
3.5	Application to Games of Chance . . . . .	93
3.5.1	Introduction . . . . .	93
3.5.2	Games of Chance . . . . .	94
3.5.3	Example . . . . .	95
3.5.4	Information-Theoretic Interpretation . . . . .	95
3.6	Optimality Properties . . . . .	96
3.6.1	Lower Error Bound . . . . .	96
3.6.2	Pareto Optimality of $\xi$ . . . . .	99
3.6.3	Balanced Pareto Optimality of $\xi$ . . . . .	101
3.6.4	On the Optimal Choice of Weights . . . . .	102
3.6.5	Occam's razor versus No Free Lunches . . . . .	103
3.7	Miscellaneous . . . . .	103
3.7.1	Multistep Predictions . . . . .	104
3.7.2	Continuous Probability Classes $\mathcal{M}$ . . . . .	106
3.7.3	Further Applications . . . . .	108
3.7.4	Prediction with Expert Advice . . . . .	108
3.7.5	Outlook . . . . .	110
3.8	Summary . . . . .	111
3.9	Technical Proofs . . . . .	112
3.9.1	How to Deal with $\mu=0$ . . . . .	112
3.9.2	Entropy Inequalities (Lemma 3.11) . . . . .	113
3.9.3	Error Inequality (Theorem 3.36) . . . . .	115
3.9.4	Binary Loss Inequality for $z \leq \frac{1}{2}$ (3.57) . . . . .	116
3.9.5	Binary Loss Inequality for $z \geq \frac{1}{2}$ (3.58) . . . . .	117
3.9.6	General Loss Inequality (3.53) . . . . .	117
3.10	History & References . . . . .	119
3.11	Problems . . . . .	119
4	<b>Agents in Known Probabilistic Environments</b> . . . . .	125
4.1	The $\text{AI}\mu$ Model in Functional Form . . . . .	126
4.1.1	The Cybernetic Agent Model . . . . .	126
4.1.2	Strings . . . . .	128
4.1.3	AI Model for Known Deterministic Environment . . . . .	128
4.1.4	AI Model for Known Prior Probability . . . . .	130
4.2	The $\text{AI}\mu$ Model in Recursive and Iterative Form . . . . .	132
4.2.1	Probability Distributions . . . . .	132
4.2.2	Explicit Form of the $\text{AI}\mu$ Model . . . . .	133
4.2.3	Equivalence of Functional and Explicit AI Model . . . . .	134
4.3	Special Aspects of the $\text{AI}\mu$ Model . . . . .	135

4.3.1	Factorizable Environments . . . . .	135
4.3.2	Constants and Limits . . . . .	138
4.3.3	Sequential Decision Theory . . . . .	139
4.4	Problems . . . . .	140
<b>5</b>	<b>The Universal Algorithmic Agent AIXI</b> . . . . .	<b>141</b>
5.1	The Universal AIXI Model . . . . .	142
5.1.1	Definition of the AIXI Model . . . . .	142
5.1.2	Universality of $M^{\text{AI}}$ and $\xi^{\text{AI}}$ . . . . .	144
5.1.3	Convergence of $\xi^{\text{AI}}$ to $\mu^{\text{AI}}$ . . . . .	145
5.1.4	Intelligence Order Relation . . . . .	146
5.2	On the Optimality of AIXI . . . . .	147
5.3	Value Bounds and Separability Concepts . . . . .	149
5.3.1	Introduction . . . . .	149
5.3.2	(Pseudo) Passive $\mu$ and the HeavenHell Example . . . . .	149
5.3.3	The OnlyOne Example . . . . .	150
5.3.4	Asymptotic Learnability . . . . .	151
5.3.5	Uniform $\mu$ . . . . .	152
5.3.6	Other Concepts . . . . .	152
5.3.7	Summary . . . . .	153
5.4	Value-Related Optimality Results . . . . .	153
5.4.1	The $\text{AI}\rho$ Models: Preliminaries . . . . .	153
5.4.2	Pareto Optimality of $\text{AI}\xi$ . . . . .	154
5.4.3	Self-Optimizing Policy $p^\xi$ w.r.t. Average Value . . . . .	156
5.5	Discounted Future Value Function . . . . .	159
5.6	Markov Decision Processes (MDP) . . . . .	165
5.7	The Choice of the Horizon . . . . .	169
5.8	Outlook . . . . .	172
5.9	Conclusions . . . . .	173
5.10	Functions $\leadsto$ Chronological Semimeasures . . . . .	173
5.11	Proof of the Entropy Inequality . . . . .	175
5.12	History & References . . . . .	177
5.13	Problems . . . . .	178
<b>6</b>	<b>Important Environmental Classes</b> . . . . .	<b>185</b>
6.1	Repetition of the $\text{AI}\mu/\xi$ Models . . . . .	186
6.2	Sequence Prediction (SP) . . . . .	187
6.2.1	Using the $\text{AI}\mu$ Model for Sequence Prediction . . . . .	188
6.2.2	Using the $\text{AI}\xi$ Model for Sequence Prediction . . . . .	190
6.3	Strategic Games (SG) . . . . .	192
6.3.1	Introduction . . . . .	192
6.3.2	Strictly Competitive Strategic Games . . . . .	193
6.3.3	Using the $\text{AI}\mu$ Model for Game Playing . . . . .	193
6.3.4	Games of Variable Length . . . . .	195
6.3.5	Using the $\text{AI}\xi$ Model for Game Playing . . . . .	195

6.4	Function Minimization (FM) . . . . .	197
6.4.1	Applications/Examples . . . . .	197
6.4.2	The Greedy Model $FMG_{\mu}$ . . . . .	198
6.4.3	The General $FM_{\mu/\xi}$ Model . . . . .	199
6.4.4	Is the General Model Inventive? . . . . .	201
6.4.5	Using the AI Models for Function Minimization . . . . .	202
6.4.6	Remark on TSP . . . . .	203
6.5	Supervised Learning from Examples (EX) . . . . .	204
6.5.1	Applications/Examples . . . . .	204
6.5.2	Supervised Learning with the $AI_{\mu/\xi}$ Model . . . . .	204
6.6	Other Aspects of Intelligence . . . . .	206
6.7	Problems . . . . .	207
<b>7</b>	<b>Computational Aspects</b> . . . . .	<b>209</b>
7.1	The Fastest & Shortest Algorithm for All Problems . . . . .	210
7.1.1	Introduction & Main Result . . . . .	210
7.1.2	Levin Search . . . . .	212
7.1.3	Fast Matrix Multiplication . . . . .	213
7.1.4	Applicability of the Fast Algorithm $M_{p^*}^{\varepsilon}$ . . . . .	214
7.1.5	The Fast Algorithm $M_{p^*}^{\varepsilon}$ . . . . .	215
7.1.6	Time Analysis . . . . .	216
7.1.7	Assumptions on the Machine Model . . . . .	218
7.1.8	Algorithmic Complexity and the Shortest Algorithm . . . . .	218
7.1.9	Generalizations . . . . .	220
7.1.10	Summary & Outlook . . . . .	220
7.2	Time-Bounded AIXI Model . . . . .	221
7.2.1	Introduction . . . . .	221
7.2.2	Time-Limited Probability Distributions . . . . .	222
7.2.3	The Idea of the Best Vote Algorithm . . . . .	224
7.2.4	Extended Chronological Programs . . . . .	224
7.2.5	Valid Approximations . . . . .	225
7.2.6	Effective Intelligence Order Relation . . . . .	226
7.2.7	The Universal Time-Bounded $AIXI_{tl}$ Agent . . . . .	226
7.2.8	Limitations and Open Questions . . . . .	227
7.2.9	Remarks . . . . .	228
<b>8</b>	<b>Discussion</b> . . . . .	<b>231</b>
8.1	What has been Achieved . . . . .	232
8.1.1	Results . . . . .	232
8.1.2	Comparison to Other Approaches . . . . .	234
8.2	General Remarks . . . . .	235
8.2.1	Miscellaneous . . . . .	235
8.2.2	Prior Knowledge . . . . .	236
8.2.3	Universal Prior Knowledge . . . . .	237
8.2.4	How $AIXI_{tl}$ Deals with Encrypted Information . . . . .	237

8.2.5	Mortal Embodied Agents . . . . .	238
8.3	Personal Remarks . . . . .	239
8.3.1	On the Foundations of Machine Learning . . . . .	239
8.3.2	In a World Without Occam . . . . .	240
8.4	Outlook & Open Questions . . . . .	241
8.5	Assumptions, Problems, Limitations . . . . .	242
8.5.1	Assumptions . . . . .	243
8.5.2	Problems . . . . .	244
8.5.3	Limitations . . . . .	244
8.6	Philosophical Issues . . . . .	245
8.6.1	Turing Test . . . . .	245
8.6.2	On the Existence of Objective Probabilities . . . . .	245
8.6.3	Free Will versus Determinism . . . . .	246
8.6.4	The Big Questions . . . . .	248
8.7	Conclusions . . . . .	248
<b>Bibliography</b> . . . . .		251
<b>Index</b> . . . . .		265

---

## Tables, Figures, Theorems, ...

Table 2.2 ((Prefix) coding of natural numbers and strings) . . . . .	34
Thesis 2.3 (Turing) . . . . .	34
Thesis 2.4 (Church) . . . . .	34
Assumption 2.5 (Short compiler) . . . . .	34
Definition 2.6 (Prefix/Monotone Turing machine) . . . . .	35
Theorem 2.7 (Universal prefix/monotone Turing machine) . . . . .	36
Definition 2.9 (Kolmogorov complexity) . . . . .	37
Theorem 2.10 (Properties of Kolmogorov complexity) . . . . .	37
Figure 2.11 (Kolmogorov Complexity) . . . . .	38
Definition 2.12 (Computable functions) . . . . .	38
Theorem 2.13 ((Non)computability of Kolmogorov complexity) . . . . .	39
Axioms 2.14 (Kolmogorov's axioms of probability theory) . . . . .	41
Definition 2.15 (Conditional probability) . . . . .	42
Theorem 2.16 (Bayes' rule 1) . . . . .	42
Axioms 2.17 (Cox's axioms for beliefs) . . . . .	43
Theorem 2.18 (Cox's theorem) . . . . .	43
Theorem 2.19 (Bayes' rule 2) . . . . .	44
Definition 2.22 ((Semi)measures) . . . . .	46
Theorem 2.23 (Universality of $M$ ) . . . . .	46
Theorem 2.25 (Posterior convergence of $M$ to $\mu$ ) . . . . .	48
Theorem 2.28 (Universal (semi)measures) . . . . .	49
Table 2.29 (Existence of universal (semi)measures) . . . . .	50
Theorem 2.31 (Martin-Löf random sequences) . . . . .	54
Definition 2.33 ( $\mu/\xi$ -random sequences) . . . . .	54
Definition 3.8 (Convergence of random sequences) . . . . .	71
Lemma 3.9 (Relations between random convergence criteria) . . . . .	71
Lemma 3.11 (Entropy inequalities) . . . . .	72
Theorem 3.19 (Convergence of $\xi$ to $\mu$ ) . . . . .	74
Theorem 3.22 ( $\mu/\xi$ -convergence of $\xi$ to $\mu$ ) . . . . .	76
Theorem 3.36 (Error bound) . . . . .	83
Theorem 3.48 (Unit loss bound) . . . . .	87
Corollary 3.49 (Unit loss bound) . . . . .	88
Theorem 3.59 (Instantaneous loss bound) . . . . .	91
Theorem 3.60 (General loss bound) . . . . .	92
Theorem 3.63 (Time to win) . . . . .	94
Theorem 3.64 (Lower error bound) . . . . .	97

Definition 3.65 (Pareto optimality) .....	99
Theorem 3.66 (Pareto optimal performance measures) .....	99
Theorem 3.69 (Balanced Pareto optimality w.r.t. $L$ ) .....	101
Theorem 3.70 (Optimality of universal weights) .....	102
Theorem 3.74 (Continuous entropy bound) .....	106
Definition 4.1 (The agent model) .....	126
Table 4.2 (Notation and emphasis in AI versus control theory) .....	127
Definition 4.4 (The $\text{AI}\mu$ model) .....	130
Definition 4.5 (The $\mu/\text{true}/\text{generating}$ value function) .....	130
Figure 4.13 (Expectimax tree/algorithm for $\mathcal{O}=\mathcal{Y}=\mathbb{B}$ ) .....	133
Theorem 4.20 (Equivalence of functional and explicit AI model) .....	134
Theorem 4.25 (Factorizable environments $\mu$ ) .....	137
Assumption 4.28 (Finiteness) .....	138
Claim 5.12 (We expect AIXI to be universally optimal) .....	146
Definition 5.14 (Intelligence order relation) .....	147
Definition 5.18 ( $\rho$ -Value function) .....	153
Definition 5.19 (Functional $\text{AI}\rho$ model) .....	153
Theorem 5.20 (Iterative $\text{AI}\rho$ model) .....	154
Theorem 5.21 (Linearity and convexity of $V_\rho$ in $\rho$ ) .....	154
Definition 5.22 (Pareto optimal policies) .....	155
Theorem 5.23 (Pareto optimality of $p^\xi$ ) .....	155
Theorem 5.24 (Balanced Pareto optimality) .....	155
Lemma 5.27 (Value difference relation) .....	156
Lemma 5.28 (Convergence of averages) .....	157
Theorem 5.29 (Self-optimizing policy $p^\xi$ w.r.t. average value) .....	157
Definition 5.30 (Discounted $\text{AI}\rho$ model and value) .....	159
Theorem 5.31 (Linearity and convexity of $V_\rho$ in $\rho$ ) .....	160
Theorem 5.32 (Pareto optimality w.r.t. discounted value) .....	160
Lemma 5.33 (Value difference relation) .....	160
Theorem 5.34 (Self-optimizing policy $p^\xi$ w.r.t. discounted value) .....	161
Theorem 5.35 (Continuity of discounted value) .....	162
Theorem 5.36 (Convergence of universal to true value) .....	163
Definition 5.37 (Ergodic Markov decision processes) .....	165
Theorem 5.38 (Self-optimizing policies for ergodic MDPs) .....	165
Corollary 5.40 ( $\text{AI}\xi$ is self-optimizing for ergodic MDPs) .....	168
Table 5.41 (Effective horizons) .....	170
Theorem 7.1 (The fastest algorithm) .....	211
Theorem 7.2 (The fastest & shortest algorithm) .....	219
Definition 7.8 (Effective intelligence order relation) .....	226
Theorem 7.9 (Optimality of $\text{AIXI}tl$ ) .....	227
Table 8.1 (Properties of learning algorithms) .....	234

---

## Notation

The following is a list of commonly used notation. The first entry is the symbol itself, followed by its meaning or name (if any) and the page number where the definition appears. Some standard symbols like  $\mathbb{R}$  are not defined in the text. There appears a \* in place of the page number for these symbols.

Symbol	Explanation	Page
[C35s]	classification of problems	viii
[Hut04b]	paper, book or other reference	*
(5.3)	label/reference for a formula/theorem/definition/...	*
$\infty$	infinity	*
$\{a, \dots, z\}$	set containing elements $a, b, \dots, y, z$ . $\{\}$ is the empty set	*
$[a, b)$	interval on the real line, closed at $a$ and open at $b$	*
$\cap, \cup, \setminus, \in$	set intersection, union, difference, membership	*
$\wedge, \vee, \neg$	Boolean conjunction (and), disjunction (or), negation (not)	*
$\subseteq, \subset$	subset, proper subset	*
$\Rightarrow$	implies	*
$\Leftrightarrow$	equivalence, if and only if, iff	*
$\square$	q.e.d. (Latin), which was to be demonstrated	*
$\forall, \exists$	for all, there exists	*
$\approx, \lesssim, \gtrsim$	approximately equal, less equal, greater equal	33
$\ll, \gg$	much smaller/greater than	*
$\equiv$	equivalent, identical, equal by definition	*
$\cong$	isomorphic	*
$:=$	define as	*
$\hat{=}$	corresponds to, informal equality	*
$\sim$	asymptotically proportional to	33
$\propto$	proportional to	*
$=, \neq$	equal to, not equal to	*
$+, -, \cdot, /$	standard arithmetic operations: sum, difference, product, ratio	*
$\sqrt{\phantom{x}}$	square root	*
$\leq, \geq, <, >$	standard inequalities	*
$ \mathcal{S} ,  a $	size/cardinality of set $\mathcal{S}$ , absolute value of $a$	*



$\rightarrow$	mapping, approaches, Boolean implication	*
$\rightarrow$	converge to each other	33
$\lim_{n \rightarrow \infty}$	limiting value of argument for $n$ tending to infinity	*
$\leadsto$	replace with	*
$\lceil x \rceil$	ceiling of $x$ : smallest integer larger or equal than $x$	*
$\lfloor x \rfloor$	floor of $x$ : largest integer smaller or equal than $x$	*
$\delta_{ab}$	Kronecker symbol, $\delta_{ab} = 1$ if $a = b$ and 0 otherwise	*
$\sum_{k=1}^n$	summation from $k=1$ to $n$	*
$\sum'_x$	summation over $x$ for which $\mu(x) \neq 0$	69
$\prod_{k=1}^n$	product from $k=1$ to $n$	*
$\int, \int_a^b dx$	Lebesgue integral, integral from $a$ to $b$ over $x$	*
$i, k, n, t$	natural numbers	33
$x, y, z$	finite strings	33
min/max	min-/maximal element of set: $\min_{x \in \mathcal{X}} f(x) = \min\{f(x) : x \in \mathcal{X}\}$	*
argmin	$\operatorname{argmin}_x f(x)$ is the $x$ minimizing $f(x)$ (ties broken arbitrarily)	*
l.h.s.	left-hand side	*
r.h.s.	right-hand side	*
w.r.t.	with respect to	*
e.g.	exempli gratia (Latin), for example	*
i.e.	id est (Latin), that is	*
etc.	et cetera (Latin), and so forth	*
cf.	confer (Latin, imperative of conferre), compare with	*
et al.	et alii (Latin), and others	*
q.e.d.	quod erat demonstrandum (Latin), which was to be shown	*
i.i.d.	independent identically distributed (random variables)	*
iff	if and only if	*
w.p.1/i.p.	with probability 1 / in probability	71
i.m./i.m.s.	in the mean / in mean sum	71
log	logarithm to some basis	*
$\log_b$	logarithm to basis $b$	*
ln	natural logarithm to basis $e = 2.71828\dots$	*
e	base of natural logarithm $e = 2.71828\dots$	*
$\mathbb{R}$	set of real numbers	*
$\mathbb{R}^+$	set of nonnegative real numbers	*
$\mathbb{N}$	set of natural numbers $\{1, 2, 3, \dots\}$	33
$\mathbb{N}_0$	set of natural numbers including zero $\{0, 1, 2, 3, \dots\}$	33
$\mathbb{Z}$	set of integers $\{\dots, -2, -1, 0, 1, 2, 3, \dots\}$	*
$\mathbb{Q}$	set of rational numbers $\{\frac{n}{d}\}$	*

$\mathcal{B} = \{0,1\}$	binary alphabet	*
$y_t \in \mathcal{Y}$	action (output of agent) in cycle $t$ , followed by ...	128
$x_t \in \mathcal{X}$	perception (feedback/input to agent) in cycle $t$	45, 128
$o_t \in \mathcal{O}$	informative input/observation in cycle $t$	128
$r_t \in \mathcal{R} \subset \mathbb{R}$	reward in cycle $t$	128
$\varepsilon$	some small positive real number	*
$\epsilon$	empty string	33
$*$	wildcard for some string (prefix, finite, or infinite)	33
$x_{1:n}$	$= x_1 \dots x_n =$ string of length $n$	45, 68, 128
$x_{<t}$	$= x_1 \dots x_{t-1} =$ string of length $t-1$	45, 68, 128
$y_{k:n}$	action-perception sequence $y_k x_k \dots y_n x_n$	128
$\dot{y}_{<k}$	actually realized action-perception sequence $\dot{y}_1 \dot{x}_1 \dots \dot{y}_{k-1} \dot{x}_{k-1}$	130
$\omega$	infinite sequence, elementary event	33
$\Omega$	sample space	42, 68
$\Gamma_{x_{1:n}}$	$= \{\omega : \omega_{1:n} = x_{1:n}\} =$ cylinder set	46, 68
$\ell(x)$	length of string $x$	33
$\langle o \rangle$	coding of object $o$	33
$\langle x, y \rangle$	uniquely decodable pairing of $x$ and $y$	33
$x'$	prefix coding of $x$	33
$O(), o()$	big and small oh-notation	33
$a \stackrel{\pm}{\leq} b$	less within an additive const., i.e. $a \leq b + O(1)$ . Similarly $\stackrel{\pm}{\geq}$	33
$a \stackrel{\times}{\leq} b$	less within a multiplicative const., i.e. $a = O(b)$ . Similarly $\stackrel{\times}{\geq}$	33
$K(x)$	prefix Kolmogorov complexity of string $x$	37
$Km(x_{1:n})$	monotone (Kolmogorov) complexity of string $x_{1:n}$	47, 190
$K(o_1 o_2)$	Kolmogorov complexity of object $o_1$ , given object $o_2$	37
$M \stackrel{\cong}{=} \xi_U$	Solomonoff-Levin's universal semimeasure	46, 48
$\mathcal{M} = \{\nu\}$	(usually countable) set of (semi)measures	48, 81
EC	$\in \{\text{AI, SP, FM, EX, SG, ...}\}$ is an environmental class	*
AI	artificial or algorithmic intelligence,	2
	most general computational environmental class	130, 154
SP	sequence prediction	187
CF	classification	108
SG	strategic two-player informed zero-sum games	192
FM	function minimization	197
EX	supervised learning (by examples)	204
pd	probability density function / distribution / measure	*
$\rho(x_{1:n})$	probability of string/sequence starting with $x_{1:n}$	46, 68
$\mu \in \mathcal{M}$	true generating environmental pd	68

<b>E</b>	expectation value, usually w.r.t. the true distribution $\mu$	68
<b>P</b>	probability, usually w.r.t. the true distribution $\mu$	68
$\mu(x_1x_2x_3x_4)$	$\mu$ probability that the $2^{nd}$ and $4^{th}$ symbols of a string are $x_2$ and $x_4$ , given the $1^{st}$ and $3^{rd}$ symbols are $x_1$ and $x_3$	132
$\nu \in \mathcal{M}$	any pd in $\mathcal{M}$	70
$\rho$	any pd not necessarily in $\mathcal{M}$ usually specifying a policy	68
$\xi$	$= \sum_{\nu \in \mathcal{M}} w_\nu \nu =$ mixture (belief) pd	48, 70
$w_\nu$	prior degree of belief in $\nu$ –or– weight of $\nu$	48, 70
$\rho^{EC}$	pd of environmental argument type EC	185
$\xi^{EC}$	mixture distribution of type EC for class EC	185
$\ell_{x_t y_t}$	incurred loss when predicting $y_t$ and $x_t$ is next symbol	86
$l_{t\nu}^\Lambda$	$\nu$ -expected instantaneous loss in step $t$ of predictor $\Lambda$	99, 87
$L_{n\nu}^\Lambda$	$\nu$ -expected cumulative loss of steps $1 \dots n$ of predictor $\Lambda$	100
$\Theta_\rho$	predictor with minimal number of $\rho$ -expected errors	82
$\Lambda_\rho$	predictor that minimizes the $\rho$ -expected loss	87
$e_{t\nu}^\Theta$	$\nu$ -probability that $\Theta$ -predictor errs in step $t$	83
$E_{n\nu}^\Theta$	$\nu$ -expected number of errors in steps $1 \dots n$ of predictor $\Theta$	83
$L_n^\Lambda \equiv L_{n\mu}^\Lambda$	abbreviation for true $\mu$ -expected loss	86
$V_{km}^{p\nu}(\dot{y}_{<k})$	value of policy $p$ in environment $\nu$ given history $\dot{y}_{<k}$	153
$y_t^\Lambda$	prediction/decision/action of predictor $\Lambda$ in step $t$	87
$y_k^p$	action of policy $p$ in cycle $k$	*
$\gamma_k$	discounting sequence	159
$\Gamma_k$	value function normalization ( $\sum_{i=k}^\infty \gamma_i$ )	159
$m, h$	agent's lifespan, horizon	129, 169
$p$	agent's policy	126
$q$	deterministic environment	126
$p^\nu$	policy that maximizes value $V_\nu^p$	130
$V_\mu^* \equiv V_{1m}^{p^\mu \mu}$	true or generating value	130
$V_\xi^* \equiv V_{1m}^{p^\xi \xi}$	universal value	146
$D_n \equiv D_{n\mu}^\xi$	relative entropy between $\mu$ and $\xi$ for the first $n$ cycles	73