Foundations of Rational Agents

Marcus Hutter
Canberra, ACT, 0200, Australia
http://www.hutter1.net/

ANU RSISE NICTA

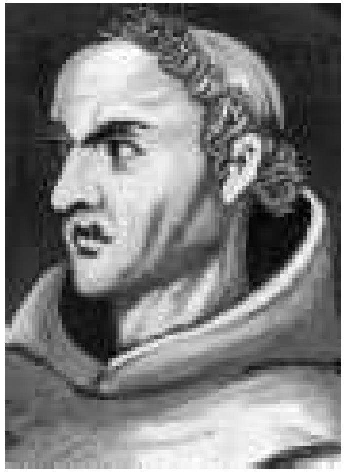
PCAR-2008, 22 September 2008

Abstract

The dream of creating artificial devices that reach or outperform human intelligence is many centuries old. Nowadays most research is more modest, focussing on solving narrower, specific problems, associated with only some aspects of intelligence, like playing chess or natural language translation, either as a goal in itself or as a bottom-up approach. The dual top down approach investigates theories of general-purpose intelligent agents: the power of such theoretical agents, how to scale them down, and the involved key concepts. Necessary ingredients seem to be Occam's razor; Turing machines; Kolmogorov complexity; probability theory; Solomonoff induction; Bayesian sequence prediction; minimum description length principle; agent framework; sequential decision theory; adaptive control theory; reinforcement learning; Levin search and extensions, which are all important subjects in their own right. From a mathematical point of view these concepts also seem to be sufficient.

Contents

Science ≈ Induction ≈ Occam's Razor



- Grue Emerald Paradox:

  Hypothesis 1: All emeralds are green.

  Hypothesis 2: All emeralds found till y2010 are green, thereafter all emeralds are blue.

- Which hypothesis is more plausible? H1! Justification?

- Occam's razor: take simplest hypothesis consistent with data.

  is the most important principle in machine learning and science.

Information Theory & Kolmogorov Complexity

- Quantification/interpretation of Occam's razor:
- Shortest description of object is best explanation.
- Shortest program for a string on a Turing machine T leads to best extrapolation=prediction.

$$K_T(x) = \min_p \{l(p) : T(p) = x\}$$

- Prediction is best for a universal Turing machine U.

Kolmogorov-complexity$(x) \;=\; K(x) \;=\; K_U(x) \qquad \leq K_T(x) + c_T$

Bayesian Probability Theory

Given (1): Models $P(D|H_i)$ for probability of observing data D, when H_i is true.

Given (2): Prior probability over hypotheses $P(H_i)$.

Goal: Posterior probability $P(H_i|D)$ of H_i, after having seen data D.

Solution:

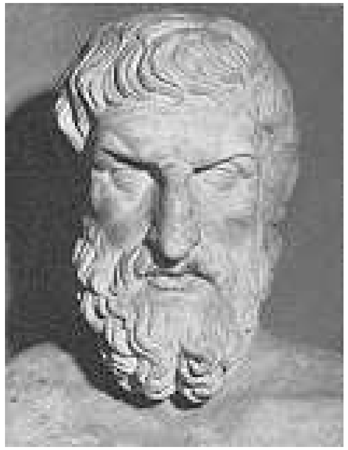
Bayes' rule:

$$\boxed{P(H_i|D) = \frac{P(D|H_i) \cdot P(H_i)}{\sum_i P(D|H_i) \cdot P(H_i)}}$$

(1) Models $P(D|H_i)$ usually easy to describe (objective probabilities)

(2) But Bayesian prob. theory does not tell us how to choose the prior $P(H_i)$ (subjective probabilities)

Algorithmic Probability Theory



- Epicurus: If more than one theory is consistent with the observations, keep all theories.

- $\Rightarrow$ uniform prior over all H_i?

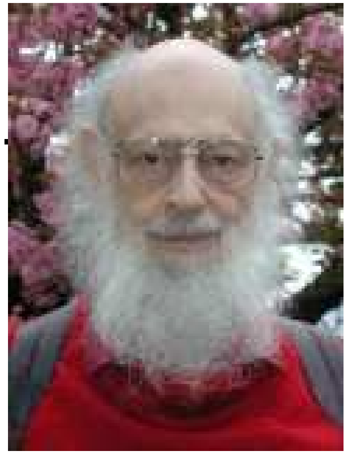
- Refinement with Occam's razor quantified in terms of Kolmogorov complexity:

$$P(H_i) := 2^{-K_{T/U}(H_i)}$$

- Fixing T we have a complete theory for prediction.
  Problem: How to choose T.

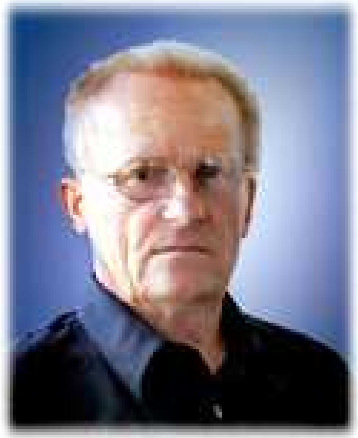
- Choosing U we have a universal theory for prediction.
  Observation: Particular choice of U does not matter much.
  Problem: Incomputable.

Inductive Inference & Universal Forecasting



- Solomonoff combined Occam, Epicurus, Bayes, and Turing into one formal theory of sequential prediction.
- $M(x)$ = probability that a universal Turing machine outputs x when provided with fair coin flips on the input tape.
- A posteriori probability of y given x is $M(y|x) = M(xy)/M(x)$.
- Given $x_1, .., x_{t-1}$, the probability of x_t is $M(x_t|x_1...x_{t-1})$.
- Immediate "applications":
  - Weather forecasting: $x_t \in \{$sun,rain$\}$.
  - Stock-market prediction: $x_t \in \{$bear,bull$\}$.
  - Continuing number sequences in an IQ test: $x_t \in I\!N$.
- Works optimally for everything!

The Minimum Description Length Principle



- Approximation of Solomonoff, since M is incomputable:
- $M(x) \approx 2^{-K_U(x)}$ (quite good)
- $K_U(x) \approx K_T(x)$ (very crude)
- Predict y of highest $M(y|x)$ is approximately same as
- MDL: Predict y of smallest $K_T(xy)$.

The Universal Similarity Metric

- One application among many:
  Determination of composer of music.
- Let $m_1, ..., m_n$ be pieces of music
  of known composer $c = 1, ..., n$.
- Let $m_?$ be (different!) piece of music
  of unknown composer.
- Concatenate each m_i with $m_?$
- Most similarity between pieces of music of same composer
  $\Rightarrow$ maximal compression.
- Guess composer is
  $\hat{i} = \arg\max_i M(m_?|m_i) \approx \arg\min_i [K_T(m_i \circ m_?) - K_T(m_i)]$
- For T choose Lempel-Ziv or bzip(2) compressor.
- No musical knowledge used in this method.

Sequential Decision Theory

Setup: For $t = 1, 2, 3, 4, ...$

Given sequence $x_1, x_2, ..., x_{t-1}$

(1) predict/make decision y_t,

(2) observe x_t,

(3) suffer loss Loss(x_t, y_t),

(4) $t \to t + 1$, goto (1)

Goal: Minimize expected Loss.

Greedy minimization of expected loss is optimal if:

Important: Decision y_t does not influence env. (future observations).

Loss function is known.

Problem: Expectation w.r.t. what?

Solution: W.r.t. universal distribution M if true distr. is unknown.

Example: Weather Forecasting

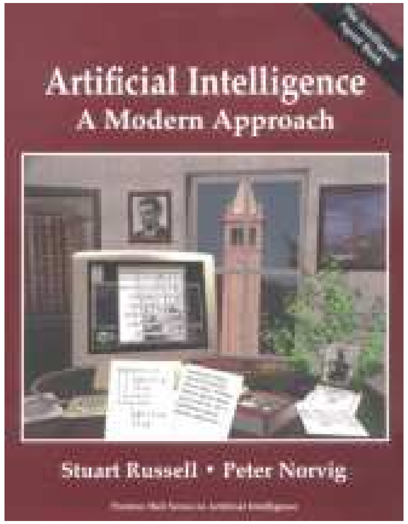
Observation $x_t \in \mathcal{X} = \{$sunny, rainy$\}$

Decision $y_t \in \mathcal{Y} = \{$umbrella, sunglasses$\}$

Loss	sunny	rainy
umbrella	0.1	0.3
sunglasses	0.0	1.0

Taking umbrella/sunglasses does not influence future weather

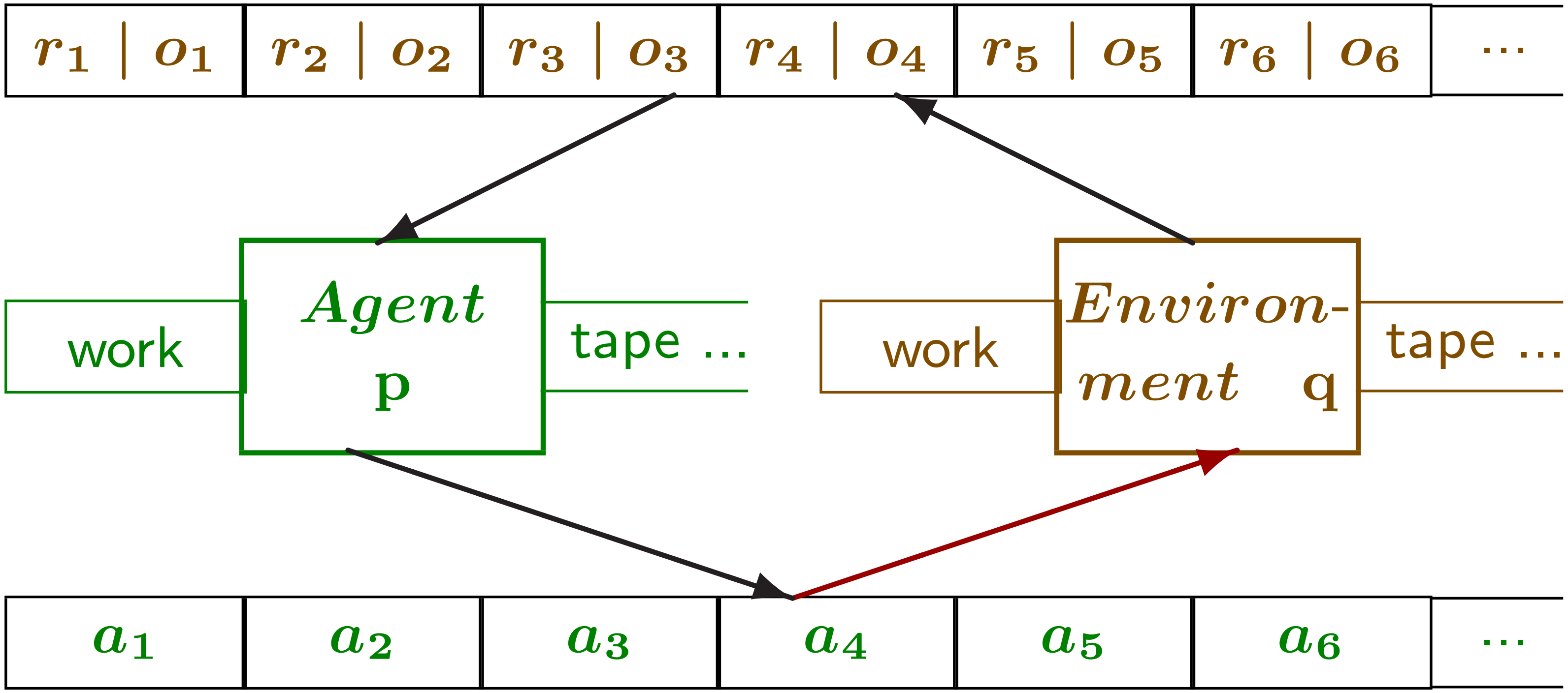
(ignoring butterfly effect)



Agent Model with Reward

if actions/decisions a

influence the environment q



Rational Agents in Known Environment

- Setup: Known deterministic or probabilistic environment
- Greedy maximization of reward r (=−Loss) no longer optimal.
  Example: Chess
- Exploration versus exploitation problem.

  $\Rightarrow$ Agent has to be farsighted.
- Optimal solution: Maximize future (expected) reward sum, called value.
- Problem: Things drastically change if environment is unknown

Rational Agents in Unknown Environment

Additional problem: (probabilistic) environment unknown.

Fields: reinforcement learning and adaptive control theory

Bayesian approach: Mixture distribution.

1. What performance does Bayes-optimal policy imply?
   It does not necessarily imply self-optimization
   (Heaven&Hell example).

2. Computationally very hard problem.

3. Choice of horizon? Immortal agents are lazy.

Universal Solomonoff mixture $\Rightarrow$ universal agent AIXI.

Represents a formal (math., non-comp.) solution to the AI problem?

Most (all?) problems are easily phrased within AIXI.

The AIXI Model in one Line

complete & essentially unique & limit-computable

$$\text{AIXI:}\quad a_k \;:=\; \arg\max_{a_k}\sum_{o_k r_k} \ldots \max_{a_m}\sum_{o_m r_m}[r_k+\ldots+r_m]\sum_{q\,:\,U(q,a_1..a_m)=o_1 r_1..o_m r_m} 2^{-\ell(q)}$$

$a$ction, $r$eward, $o$bservation, Universal TM, $q$rogram, k=now

AIXI is an elegant mathematical theory of AI

Claim: AIXI is the most intelligent environmental independent, i.e. universally optimal, agent possible.

Proof: For formalizations, quantifications, and proofs, see [Hut05].

Applications: Robots, Agents, Games, Optimization, Supervised Learning, Sequence Prediction, Classification, ...

Computational Issues: Universal Search

- Levin search: Fastest algorithm for inversion and optimization problems.
- Theoretical application:
  Assume somebody found a non-constructive proof of P=NP, then Levin-search is a polynomial time algorithm for every NP (complete) problem.
- Practical applications (J. Schmidhuber)
  Maze, towers of hanoi, robotics, ...
- FastPrg: The asymptotically fastest and shortest algorithm for all well-defined problems.
- AIXItl: Computable variant of AIXI.
- Human Knowledge Compression Prize: (50'000€)

Properties of Learning Algorithms

Comparison of AIXI to Other Approaches

Algorithm \ Properties	time efficient	data efficient	exploration	convergence	global optimum	generalization	POMDP	learning	active
Value/Policy iteration	yes/no	yes	–	YES	YES	NO	NO	NO	yes
TD w. func.approx.	no/yes	NO	NO	no/yes	NO	YES	NO	YES	YES
Direct Policy Search	no/yes	YES	NO	no/yes	NO	YES	no	YES	YES
Logic Planners	yes/no	YES	yes	YES	YES	no	no	YES	yes
RL with Split Trees	yes	YES	no	YES	NO	yes	YES	YES	YES
Pred.w. Expert Advice	yes/no	YES	–	YES	yes/no	yes	NO	YES	NO
OOPS	yes/no	no	–	yes	yes/no	YES	YES	YES	YES
Market/Economy RL	yes/no	no	NO	no	no/yes	yes	yes/no	YES	YES
SPXI	no	YES	–	YES	YES	YES	NO	YES	NO
AIXI	NO	YES	YES	yes	YES	YES	YES	YES	YES
AIXItl	no/yes	YES	YES	YES	yes	YES	YES	YES	YES
Human	yes	yes	yes	no/yes	NO	YES	YES	YES	YES

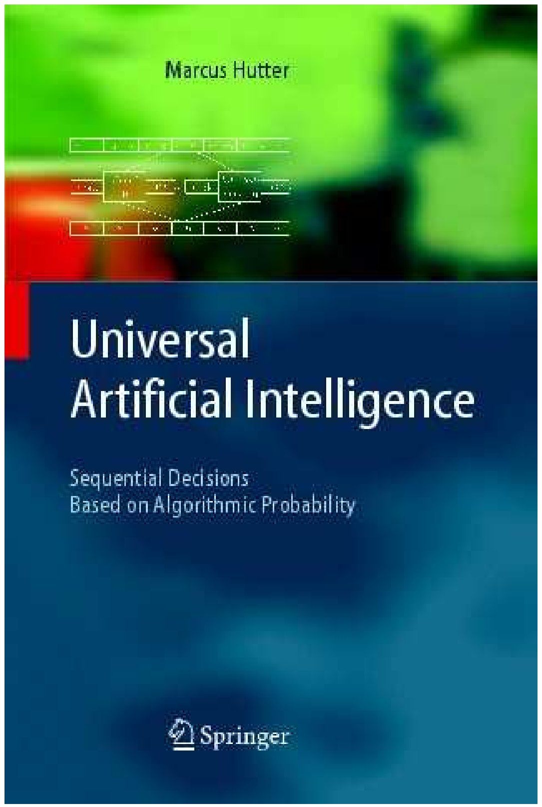
Machine Intelligence Tests & Definitions

★= yes, ·= no, •= debatable, ? = unknown. Intelligence Test	Valid	Informative	Wide Range	General	Dynamic	Unbiased	Fundamental	Formal	Objective	Fully Defined	Universal	Practical	Test vs. Def.
Turing Test	•	·	·	·	•	·	·	·	·	•	·	•	T
Total Turing Test	•	·	·	·	•	·	·	·	·	•	·	·	T
Inverted Turing Test	•	•	·	·	•	·	·	·	·	•	·	•	T
Toddler Turing Test	•	·	·	·	•	·	·	·	·	·	·	•	T
Linguistic Complexity	•	★	•	·	·	·	·	•	•	·	•	•	T
Text Compression Test	•	★	★	•	·	•	•	★	★	★	•	★	T
Turing Ratio	•	★	★	★	?	?	?	?	?	·	?	?	T/D
Psychometric AI	★	★	•	★	?	•	·	•	•	•	·	•	T/D
Smith's Test	•	★	★	•	·	?	★	★	★	·	?	•	T/D
C-Test	•	★	★	•	·	★	★	★	★	★	★	★	T/D
AIXI	★	★	★	★	★	★	★	★	★	★	★	·	D

Summary

- Sequential Decision Theory solves the problem of rational agents in uncertain worlds if the environmental probability distribution is known.
- Solomonoff's theory of Universal Induction solves the problem of sequence prediction for unknown prior distribution.
- Combining both ideas one arrives at

$$
\begin{array}{ccc}
\multicolumn{3}{c}{\text{A Unified View of Artificial Intelligence}} \\
= & & = \\
\text{Decision Theory} & = & \text{Probability} + \text{Utility Theory} \\
+ & & + \\
\text{Universal Induction} & = & \text{Ockham} + \text{Bayes} + \text{Turing}
\end{array}
$$



Common Criticisms

- AIXI is obviously wrong.
  (intelligence cannot be captured in a few simple equations)
- AIXI is obviously correct. (everybody already knows this)
- Assuming that the environment is computable is too strong.
- All standard objections to strong AI also apply to AIXI.
  (free will, lookup table, Lucas/Penrose Goedel argument)
- AIXI doesn't deal with X or cannot do X.
  (X = consciousness, creativity, imagination, emotion, love, soul, etc.)
- AIXI is not intelligent because it cannot choose its goals.
- Universal AI is impossible due to the No-Free-Lunch theorem.

See [Legg:08] for refutations of these and more criticisms.

General Murky & Quirky AI Questions

- Does current mainstream AI research has anything todo with AI?
- Are sequential decision and algorithmic probability theory all we need to well-define AI?
- What is (Universal) AI theory good for?
- What are robots good for in AI?
- Is intelligence a fundamentally simple concept? (compare with fractals or physics theories)
- What can we (not) expect from super-intelligent agents?
- Is maximizing the expected reward the right criterion?
- Isn't universal learning impossible due to the NFL theorems?

Next Steps

- Address the many open theoretical questions (see Hutter:05).
- Bridge the gap between (Universal) AI theory and AI practice.
- Explore what role logical reasoning, knowledge representation, vision, language, etc. play in Universal AI.
- Determine the right discounting of future rewards.
- Develop the right nurturing environment for a learning agent.
- Consider embodied agents (e.g. internal$\leftrightarrow$external reward)
- Analyze AIXI in the multi-agent setting.

Introductory Literature

[HMU01] J. E. Hopcroft, R. Motwani, and J. D. Ullman. *Introduction to Automata Theory, Language, and Computation*. Addison-Wesley, 2nd edition, 2001.

[RN03] S. J. Russell and P. Norvig. *Artificial Intelligence. A Modern Approach*. Prentice-Hall, Englewood Cliffs, NJ, 2nd edition, 2003.

[LV97] M. Li and P. M. B. Vitányi. *An Introduction to Kolmogorov Complexity and its Applications*. Springer, Berlin, 2nd edition, 1997.

[SB98] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.

[Leg08] S. Legg. *Machine Super Intelligence*. PhD Thesis, Lugano, 2008.

[Hut05] M. Hutter. *Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability*. Springer, Berlin, 2005.