Sequential Decisions based on Algorithmic Probability

Marcus Hutter

Canberra, ACT, 0200, Australia

http://www.hutter1.net/

ANU

RSISE

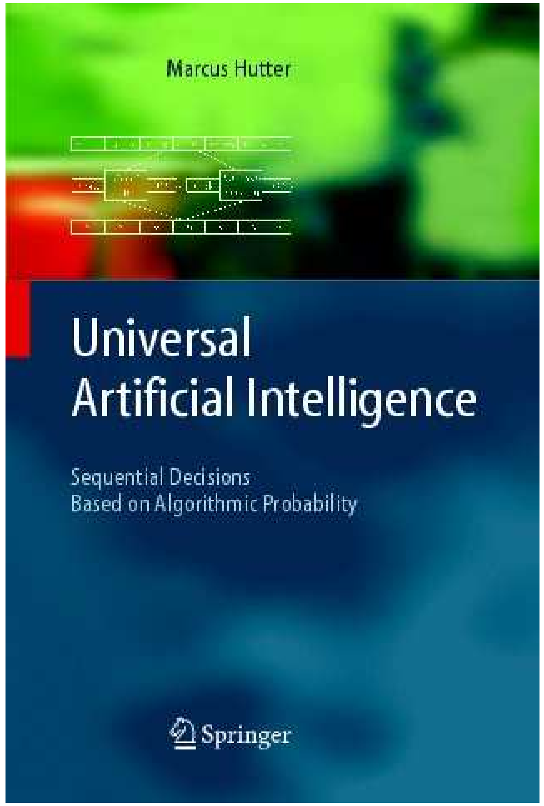
NICTA

Motivation

- 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

A Unified View of Artificial Intelligence

=		=
Decision Theory	=	Probability + Utility Theory
+		+
Universal Induction	=	Ockham + Bayes + Turing



Abstract: Motivation

The dream of creating artificial devices that reach or outperform human intelligence is an old one, however a computationally efficient theory of true intelligence has not been found yet, despite considerable efforts in the last 50 years. Nowadays most research is more modest, focussing on solving more narrow, 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, is to find a mathematical (not computational) definition of general intelligence. Note that the AI problem remains non-trivial even when ignoring computational aspects.

Abstract: Contents

In this course we will develop such an elegant mathematical parameter-free theory of an optimal reinforcement learning agent embedded in an arbitrary unknown environment that possesses essentially all aspects of rational intelligence. Most of the course is devoted to giving an introduction to the key ingredients of this theory, which are important subjects in their own right: Occam's razor; Turing machines; Kolmogorov complexity; probability theory; Solomonoff induction; Bayesian sequence prediction; minimum description length principle; agents; sequential decision theory; adaptive control theory; reinforcement learning; Levin search and extensions.

Table of Contents

1 A SHORT TOUR THROUGH THE COURSE

Introduction

- 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.

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

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

(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 in 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 $\dot{x}_1, .., \dot{x}_{t-1}$, the probability of x_t is $M(x_t|\dot{x}_1...\dot{x}_{t-1})$.
- Immediate "applications":
  - Weather forecasting: $x_t \in \{\text{sun,rain}\}$.
  - Stock-market prediction: $x_t \in \{\text{bear,bull}\}$.
  - 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 example 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) $t \rightarrow t + 1$, 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 $x_t \in \mathcal{X} = \{\text{sunny, rainy}\}$

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

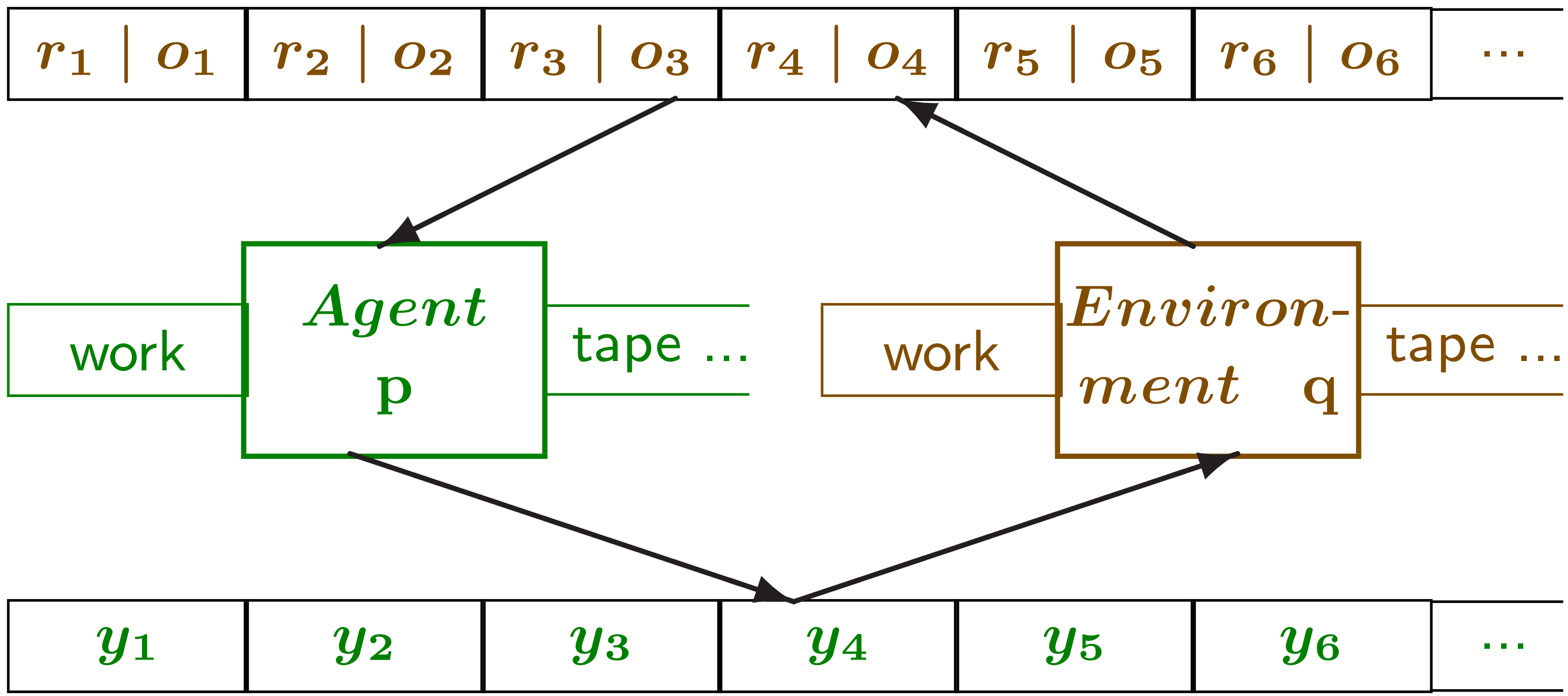
Loss	sunny	rainy
umbrella	0.1	0.3
sunglasses	0.0	1.0

Taking umbrella/sunglasses does not influence future weather

(ignoring butterfly effect)

Agent Model with Reward

if actions/decisions y influence the environment q



Rational Agents in Known Environment

- Setup: Known deterministic or probabilistic environment
- Greedy maximization of reward r $(=-\text{Loss})$ 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.

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.

Discussion at End of Course

- What has been achieved?
- Made assumptions.
- General and personal remarks.
- Open problems.
- Philosophical issues.