Foundations of Universal Induction and Intelligent Agents

Marcus Hutter

Canberra, ACT, 0200, Australia

http://www.hutter1.net/

ANU RSISE NICTA

Abstract

The dream of creating artificial devices that reach or outperform human intelligence is many centuries old. This lecture series presents the elegant parameter-free theory, developed in [Hut05], of an optimal reinforcement learning agent embedded in an arbitrary unknown environment that possesses essentially all aspects of rational intelligence. The theory reduces all conceptual AI problems to pure computational questions.

How to perform inductive inference is closely related to the AI problem. The lecture series covers Solomonoff's theory, elaborated on in [Hut07], which solves the induction problem, at least from a philosophical and statistical perspective.

Both theories are based on Occam's razor quantified by Kolmogorov complexity; Bayesian probability theory; and sequential decision theory.

Relation between ML & RL & (U)AI

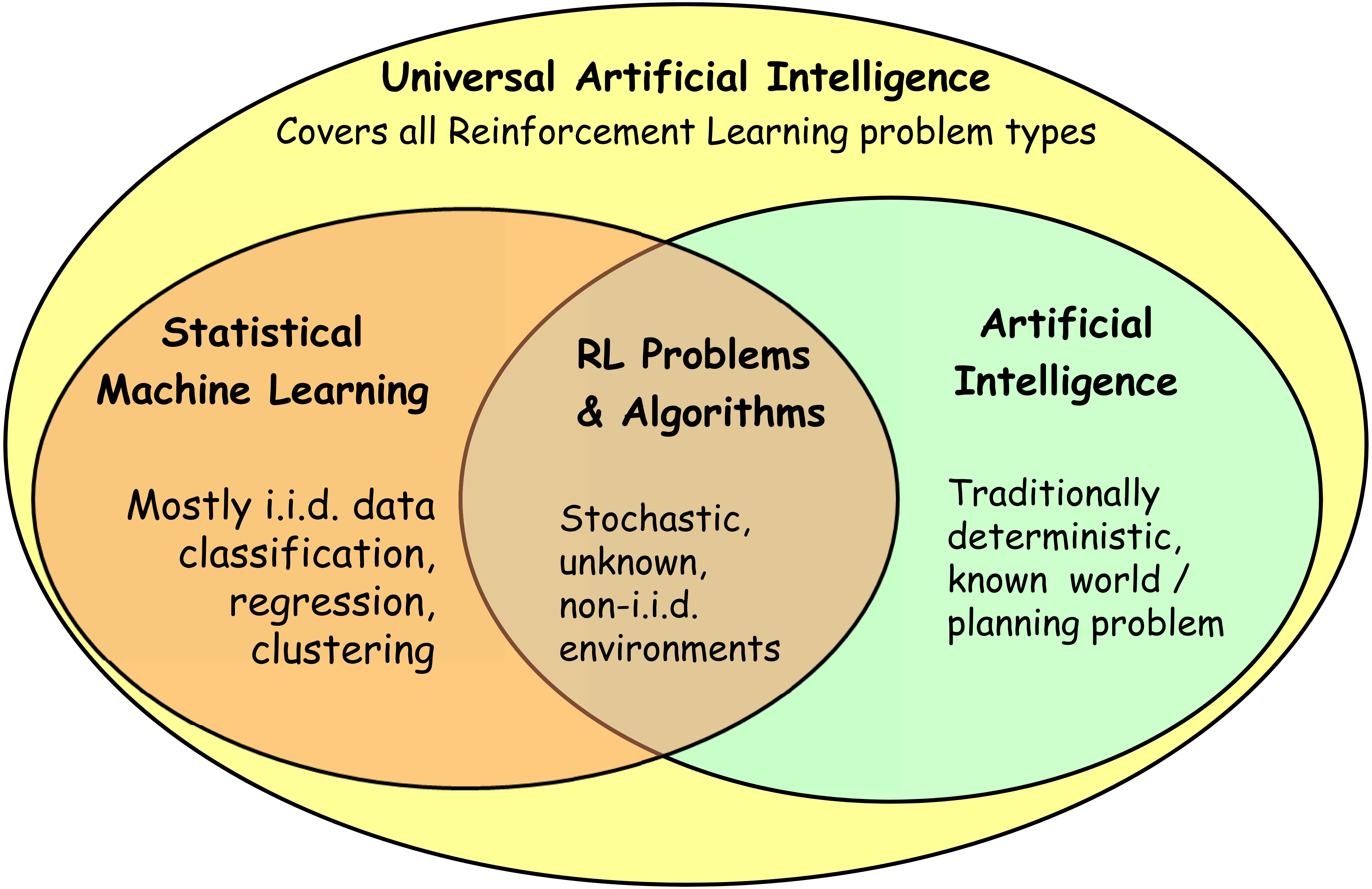


TABLE OF CONTENTS

PHILOSOPHICAL ISSUES

- Philosophical Problems
- What is (Artificial) Intelligence?
- How to do Inductive Inference?
- How to Predict (Number) Sequences?
- How to make Decisions in Unknown Environments?
- Occam's Razor to the Rescue
- The Grue Emerald and Confirmation Paradoxes
- What this Lecture Series is (Not) About

Philosophical Issues: Abstract

I start by considering the philosophical problems concerning artificial intelligence and machine learning in general and induction in particular. I illustrate the problems and their intuitive solution on various (classical) induction examples. The common principle to their solution is Occam's simplicity principle. Based on Occam's and Epicurus' principle, Bayesian probability theory, and Turing's universal machine, Solomonoff developed a formal theory of induction. I describe the sequential/online setup considered in this lecture series and place it into the wider machine learning context.

What is (Artificial) Intelligence?

Intelligence can have many faces ⇒ formal definition difficult

- reasoning
- creativity
- association
- generalization
- pattern recognition
- problem solving
- memorization
- planning
- achieving goals
- learning
- optimization
- self-preservation
- vision
- language processing
- classification
- induction
- deduction
- ...

What is AI?	**Thinking**	**Acting**
humanly	Cognitive Science	Turing test, Behaviorism
rationally	Laws Thought	Doing the Right Thing

Collection of 70+ Defs of Intelligence
`http://www.vetta.org/definitions-of-intelligence/`

Real world is nasty: partially unobservable, uncertain, unknown, non-ergodic, reactive, vast, but luckily structured, ...

Informal Definition of (Artificial) Intelligence

> Intelligence measures an agent's ability to achieve goals in a wide range of environments. [S. Legg and M. Hutter]

Emergent: Features such as the ability to learn and adapt, or to understand, are implicit in the above definition as these capacities enable an agent to succeed in a wide range of environments.

The science of Artificial Intelligence is concerned with the construction of intelligent systems/artifacts/agents and their analysis.

What next? Substantiate all terms above: agent, ability, utility, goal, success, learn, adapt, environment, ...

On the Foundations of Artificial Intelligence

- Example: Algorithm/complexity theory: The goal is to find fast algorithms solving problems and to show lower bounds on their computation time. Everything is rigorously defined: algorithm, Turing machine, problem classes, computation time, ...

- Most disciplines start with an informal way of attacking a subject. With time they get more and more formalized often to a point where they are completely rigorous. Examples: set theory, logical reasoning, proof theory, probability theory, infinitesimal calculus, energy, temperature, quantum field theory, ...

- Artificial Intelligence: Tries to build and understand systems that act intelligently, learn from experience, make good predictions, are able to generalize, ... Many terms are only vaguely defined or there are many alternate definitions.

Induction→Prediction→Decision→Action

Induction infers general models from specific observations/facts/data, usually exhibiting regularities or properties or relations in the latter.

Having or acquiring or *learning* or inducing a model of the environment an agent interacts with allows the agent to make predictions and utilize them in its decision process of finding a good next action.

Example

Induction: Find a model of the world economy.

Prediction: Use the model for predicting the future stock market.

Decision: Decide whether to invest assets in stocks or bonds.

Action: Trading large quantities of stocks influences the market.

Example 1: Probability of Sunrise Tomorrow

What is the probability $p(1|1^d)$ that the sun will rise tomorrow?
($d =$ past # days sun rose, $1 =$sun rises. $0 =$ sun will not rise)

- p is undefined, because there has never been an experiment that tested the existence of the sun *tomorrow* (ref. class problem).
- The $p = 1$, because the sun rose in all past experiments.
- $p = 1 - \epsilon$, where ϵ is the proportion of stars that explode per day.
- $p = \frac{d+1}{d+2}$, which is Laplace rule derived from Bayes rule.
- Derive p from the type, age, size and temperature of the sun, even though we never observed another star with those exact properties.

Conclusion: We predict that the sun will rise tomorrow with high probability independent of the justification.

Example 2: Digits of a Computable Number

- Extend 14159265358979323846264338327950288419716939937?
- Looks random?!
- Frequency estimate: $n =$ length of sequence. $k_i =$ number of occured $i \Longrightarrow$ Probability of next digit being i is $\frac{i}{n}$. Asymptotically $\frac{i}{n} \to \frac{1}{10}$ (seems to be) true.
- But we have the strong feeling that (i.e. with high probability) the next digit will be 5 because the previous digits were the expansion of π.
- Conclusion: We prefer answer 5, since we see more structure in the sequence than just random digits.

Example 3: Number Sequences

Sequence: x_1, x_2, x_3, x_4, x_5, ...
1, 2, 3, 4, ?, ...

- $x_5 = 5$, since $x_i = i$ for $i = 1..4$.
- $x_5 = 29$, since $x_i = i^4 - 10i^3 + 35i^2 - 49i + 24$.

Conclusion: We prefer 5, since linear relation involves less arbitrary parameters than 4th-order polynomial.

Sequence: 2,3,5,7,11,13,17,19,23,29,31,37,41,43,47,53,59,?

- 61, since this is the next prime
- 60, since this is the order of the next simple group

Conclusion: We prefer answer 61, since primes are a more familiar concept than simple groups.

On-Line Encyclopedia of Integer Sequences:
http://www.research.att.com/∼njas/sequences/

Occam's Razor to the Rescue

- Is there a unique principle which allows us to formally arrive at a prediction which
  - coincides (always?) with our intuitive guess -or- even better,
  - which is (in some sense) most likely the best or correct answer?

- Yes! Occam's razor: Use the simplest explanation consistent with past data (and use it for prediction).

- Works! For examples presented and for many more.

- Actually Occam's razor can serve as a foundation of machine learning in general, and is even a fundamental principle (or maybe even the mere definition) of science.

- Problem: Not a formal/mathematical objective principle.
  What is simple for one may be complicated for another.

Grue Emerald Paradox

Hypothesis 1: All emeralds are green.

Hypothesis 2: All emeralds found till y2020 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.

Confirmation Paradox

(i) $R \to B$ is confirmed by an R-instance with property B

(ii) $\neg B \to \neg R$ is confirmed by a $\neg B$-instance with property $\neg R$.

(iii) Since $R \to B$ and $\neg B \to \neg R$ are logically equivalent,
$R \to B$ is also confirmed by a $\neg B$-instance with property $\neg R$.

Example: Hypothesis (o): All ravens are black (R=Raven, B=Black).

(i) observing a Black Raven confirms Hypothesis (o).

(iii) observing a White Sock also confirms that all Ravens are Black, since a White Sock is a non-Raven which is non-Black.

This conclusion sounds absurd! What's the problem?

What This Lecture Series is (Not) About

Dichotomies in Artificial Intelligence & Machine Learning

scope of this lecture series	⇔	scope of other lecture seriess
(machine) learning	⇔	(GOFAI) knowledge-based
statistical	⇔	logic-based
decision ⇔ prediction	⇔	induction ⇔ action
classification	⇔	regression
sequential / non-iid	⇔	independent identically distributed
online learning	⇔	offline/batch learning
passive prediction	⇔	active learning
Bayes ⇔ MDL	⇔	Expert ⇔ Frequentist
uninformed / universal	⇔	informed / problem-specific
conceptual/mathematical issues	⇔	computational issues
exact/principled	⇔	heuristic
supervised learning	⇔	unsupervised ⇔ RL learning
exploitation	⇔	exploration

BAYESIAN SEQUENCE PREDICTION

- Sequential/Online Prediction – Setup
- Uncertainty and Probability
- Frequency Interpretation: Counting
- Objective Uncertain Events & Subjective Degrees of Belief
- Bayes' and Laplace's Rules
- The Bayes-mixture distribution
- Predictive Convergence
- Sequential Decisions and Loss Bounds
- Generalization: Continuous Probability Classes
- Summary

Bayesian Sequence Prediction: Abstract

The aim of probability theory is to describe uncertainty. There are various sources and interpretations of uncertainty. I compare the frequency, objective, and subjective probabilities, and show that they all respect the same rules, and derive Bayes' and Laplace's famous and fundamental rules. Then I concentrate on general sequence prediction tasks. I define the Bayes mixture distribution and show that the posterior converges rapidly to the true posterior by exploiting some bounds on the relative entropy. Finally I show that the mixture predictor is also optimal in a decision-theoretic sense w.r.t. any bounded loss function.

Sequential/Online Prediction – Setup

In sequential or online prediction, for times $t = 1, 2, 3, ...$,

our predictor p makes a prediction $y_t^p \in \mathcal{Y}$

based on past observations $x_1, ..., x_{t-1}$.

Thereafter $x_t \in \mathcal{X}$ is observed and p suffers $\text{Loss}(x_t, y_t^p)$.

The goal is to design predictors with small total loss or cumulative $\text{Loss}_{1:T}(p) := \sum_{t=1}^{T} \text{Loss}(x_t, y_t^p)$.

Applications are abundant, e.g. weather or stock market forecasting.

Example:

$\text{Loss}(x, y)$	$\mathcal{X} = \{$sunny , rainy$\}$	
$\mathcal{Y} = \{$ umbrella	0.1	0.3
sunglasses $\}$	0.0	1.0

Setup also includes: Classification and Regression problems.

Uncertainty and Probability

The aim of probability theory is to describe uncertainty.

Sources/interpretations for uncertainty:

- Frequentist: probabilities are relative frequencies.
  (e.g. the relative frequency of tossing head.)
- Objectivist: probabilities are real aspects of the world.
  (e.g. the probability that some atom decays in the next hour)
- Subjectivist: probabilities describe an agent's degree of belief.
  (e.g. it is (im)plausible that extraterrestrians exist)

Frequency Interpretation: Counting

- The frequentist interprets probabilities as relative frequencies.
- If in a sequence of n independent identically distributed (i.i.d.) experiments (trials) an event occurs $k(n)$ times, the relative frequency of the event is $k(n)/n$.
- The limit $\lim_{n\to\infty} k(n)/n$ is defined as the probability of the event.
- For instance, the probability of the event head in a sequence of repeatedly tossing a fair coin is $\frac{1}{2}$.
- The frequentist position is the easiest to grasp, but it has several shortcomings:
- Problems: definition circular, limited to i.i.d, reference class problem.

Objective Interpretation: Uncertain Events

- For the objectivist probabilities are real aspects of the world.
- The outcome of an observation or an experiment is not deterministic, but involves physical random processes.
- The set Ω of all possible outcomes is called the sample space.
- It is said that an event $E \subset \Omega$ occurred if the outcome is in E.
- In the case of i.i.d. experiments the probabilities p assigned to events E should be interpretable as limiting frequencies, but the application is not limited to this case.
- (Some) probability axioms:
  $p(\Omega) = 1$ and $p(\{\}) = 0$ and $0 \leq p(E) \leq 1$.
  $p(A \cup B) = p(A) + p(B) - p(A \cap B)$.
  $p(B|A) = \frac{p(A \cap B)}{p(A)}$ is the probability of B given event A occurred.

Subjective Interpretation: Degrees of Belief

- The subjectivist uses probabilities to characterize an agent's degree of belief in something, rather than to characterize physical random processes.
- This is the most relevant interpretation of probabilities in AI.
- We define the plausibility of an event as the degree of belief in the event, or the subjective probability of the event.
- It is natural to assume that plausibilities/beliefs $\text{Bel}(\cdot|\cdot)$ can be repr. by real numbers, that the rules qualitatively correspond to common sense, and that the rules are mathematically consistent. $\Rightarrow$
- Cox's theorem: $\text{Bel}(\cdot|A)$ is isomorphic to a probability function $p(\cdot|\cdot)$ that satisfies the axioms of (objective) probabilities.
- Conclusion: $\boxed{\text{Beliefs follow the same rules as probabilities}}$

Bayes' Famous Rule

Let D be some possible data (i.e. D is event with $p(D) > 0$) and $\{H_i\}_{i\in I}$ be a countable complete class of mutually exclusive hypotheses (i.e. H_i are events with $H_i \cap H_j = \{\}\ \forall i \neq j$ and $\bigcup_{i\in I} H_i = \Omega$).

Given: $p(H_i)$ = a priori plausibility of hypotheses H_i (subj. prob.)

Given: $p(D|H_i)$ = likelihood of data D under hypothesis H_i (obj. prob.)

Goal: $p(H_i|D)$ = a posteriori plausibility of hypothesis H_i (subj. prob.)

Solution: $$p(H_i|D) = \frac{p(D|H_i)p(H_i)}{\sum_{i\in I} p(D|H_i)p(H_i)}$$

Proof: From the definition of conditional probability and

$$\sum_{i\in I} p(H_i|...) = 1 \quad \Rightarrow \quad \sum_{i\in I} p(D|H_i)p(H_i) = \sum_{i\in I} p(H_i|D)p(D) = p(D)$$

Example: Bayes' and Laplace's Rule

Assume data is generated by a biased coin with head probability θ, i.e. H_θ :=Bernoulli(θ) with $\theta \in \Theta := [0,1]$.

Finite sequence: $x = x_1 x_2 ... x_n$ with n_1 ones and n_0 zeros.

Sample infinite sequence: $\omega \in \Omega = \{0,1\}^\infty$

Basic event: $\Gamma_x = \{\omega : \omega_1 = x_1, ..., \omega_n = x_n\}$ = set of all sequences starting with x.

Data likelihood: $p_\theta(x) := p(\Gamma_x | H_\theta) = \theta^{n_1}(1-\theta)^{n_0}$.

Bayes (1763): Uniform prior plausibility: $p(\theta) := p(H_\theta) = 1$

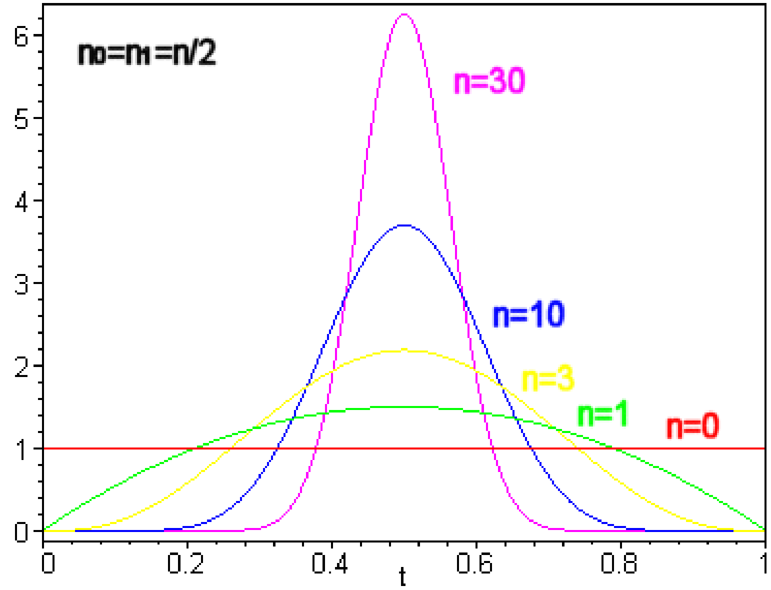
$$\left(\int_0^1 p(\theta)\, d\theta = 1 \text{ instead } \sum_{i\in I} p(H_i) = 1\right)$$

Evidence: $p(x) = \int_0^1 p_\theta(x) p(\theta)\, d\theta = \int_0^1 \theta^{n_1}(1-\theta)^{n_0}\, d\theta = \frac{n_1! n_0!}{(n_0+n_1+1)!}$

Example: Bayes' and Laplace's Rule

Bayes: Posterior plausibility of θ after seeing x is:

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{(n+1)!}{n_1!n_0!}\theta^{n_1}(1-\theta)^{n_0}$$



.

Laplace: What is the probability of seeing 1 after having observed x?

$$p(x_{n+1}=1|x_1...x_n) = \frac{p(x1)}{p(x)} = \frac{n_1+1}{n+2}$$

Laplace believed that the sun had risen for 5000 years = 1'826'213 days, so he concluded that the probability of doomsday tomorrow is $\frac{1}{1826215}$.

Exercise: Envelope Paradox

- I offer you two closed envelopes, one of them contains twice the amount of money than the other. You are allowed to pick one and open it. Now you have two options. Keep the money or decide for the other envelope (which could double or half your gain).

- Symmetry argument: It doesn't matter whether you switch, the expected gain is the same.

- Refutation: With probability $p = 1/2$, the other envelope contains twice/half the amount, i.e. if you switch your expected gain increases by a factor 1.25=(1/2)*2+(1/2)*(1/2).

- Present a Bayesian solution.

Notation: Strings & Probabilities

Strings: $x = x_1 x_2 ... x_n$ with $x_t \in \mathcal{X}$ and $x_{1:n} := x_1 x_2 ... x_{n-1} x_n$ and $x_{<n} := x_1 ... x_{n-1}$.

Probabilities: $\sigma(x_1 ... x_n)$ is the probability that an (infinite) sequence starts with $x_1 ... x_n$.

Conditional probability:

$$\sigma_n := \sigma(x_n | x_{<n}) = \sigma(x_{1:n}) / \sigma(x_{<n}),$$

$$\sigma(x_1 ... x_n) = \sigma(x_1) \cdot \sigma(x_2 | x_1) \cdot ... \cdot \sigma(x_n | x_1 ... x_{n-1}).$$

True data generating distribution: μ

The Bayes-Mixture Distribution ξ

- Assumption: The true (objective) environment μ is unknown.
- Bayesian approach: Replace true probability distribution μ by a Bayes-mixture ξ.
- Assumption: We know that the true environment μ is contained in some known countable (in)finite set $\mathcal{M}$ of environments.
- The Bayes-mixture ξ is defined as

$$\xi(x) := \sum_{\nu\in\mathcal{M}} w_\nu \nu(x) \quad \text{with} \quad \sum_{\nu\in\mathcal{M}} w_\nu = 1, \quad w_\nu > 0 \; \forall\nu$$

- The weights w_ν may be interpreted as the prior degree of belief that the true environment is ν, or $k^\nu = \ln w_\nu^{-1}$ as a complexity penalty (prefix code length) of environment ν.
- Then $\xi(x)$ could be interpreted as the prior subjective belief probability in observing x.

Relative Entropy

Relative entropy: $D(\boldsymbol{p}||\boldsymbol{q}) := \sum_i p_i \ln \frac{p_i}{q_i}$

Properties: $D(\boldsymbol{p}||\boldsymbol{q}) \geq 0$ and $D(\boldsymbol{p}||\boldsymbol{q}) = 0 \Leftrightarrow \boldsymbol{p} = \boldsymbol{q}$

Instantaneous relative entropy: $d_t(x_{<t}) := \sum_{x_t \in \mathcal{X}} \mu(x_t|x_{<t}) \ln \frac{\mu(x_t|x_{<t})}{\xi(x_t|x_{<t})}$

$\boxed{\text{Total relative entropy: } D_n := \sum_{t=1}^n \mathbf{E}[d_t] \leq \ln w_\mu^{-1}}$

$\mathbf{E}[f]$ =Expectation of f w.r.t. the *true* distribution μ, e.g.

If $f : \mathcal{X}^n \to I\!R$, then $\mathbf{E}[f] := \sum_{x_{1:n}} \mu(x_{1:n}) f(x_{1:n})$.

Proof based on dominance or universality: $\xi(x) \geq w_\mu \mu(x)$.

Proof of the Entropy Bound

$$D_n \equiv \sum_{t=1}^{n} \sum_{x_{<t}} \mu(x_{<t}) \cdot d_t(x_{<t}) \stackrel{(a)}{=} \sum_{t=1}^{n} \sum_{x_{1:t}} \mu(x_{1:t}) \ln \frac{\mu(x_t|x_{<t})}{\xi(x_t|x_{<t})} =$$

$$\stackrel{(b)}{=} \sum_{x_{1:n}} \mu(x_{1:n}) \ln \prod_{t=1}^{n} \frac{\mu(x_t|x_{<t})}{\xi(x_t|x_{<t})} \stackrel{(c)}{=} \sum_{x_{1:n}} \mu(x_{1:n}) \ln \frac{\mu(x_{1:n})}{\xi(x_{1:n})} \stackrel{(d)}{\leq} \ln w_\mu^{-1}$$

(a) Insert def. of d_t and used chain rule $\mu(x_{<t}) \cdot \mu(x_t|x_{<t}) = \mu(x_{1:t})$.

(b) $\sum_{x_{1:t}} \mu(x_{1:t}) = \sum_{x_{1:n}} \mu(x_{1:n})$ and argument of log is independent of $x_{t+1:n}$. The t sum can now be exchanged with the $x_{1:n}$ sum and transforms to a product inside the logarithm.

(c) Use chain rule again for μ and ξ.

(d) Use dominance $\xi(x) \geq w_\mu \mu(x)$.

Predictive Convergence

$\boxed{\text{Theorem: } \xi(x_t|x_{<t}) \to \mu(x_t|x_{<t}) \text{ rapid w.p.1 for } t \to \infty}$

Proof: $D_\infty \equiv \sum_{t=1}^{\infty} \mathbf{E}[d_t] \le \ln w_\mu^{-1}$ and $d_t \ge 0$

$$\Longrightarrow \quad d_t \stackrel{t\to\infty}{\longrightarrow} 0 \quad \Longleftrightarrow \quad \xi_t \to \mu_t.$$

Fazit: ξ is excellent universal predictor if unknown μ belongs to $\mathcal{M}$.

How to choose $\mathcal{M}$ and w_μ? Both as large as possible?! More later.

Sequential Decisions

A prediction is very often the basis for some decision. The decision results in an action, which itself leads to some reward or loss.

Let $\text{Loss}(x_t,y_t)\in[0,1]$ be the received loss when taking action $y_t\in\mathcal{Y}$ and $x_t\in\mathcal{X}$ is the t^{th} symbol of the sequence.

For instance, decision $\mathcal{Y}=\{$umbrella, sunglasses$\}$ based on weather forecasts $\mathcal{X}=\{$sunny, rainy$\}$.

Loss	sunny	rainy
umbrella	0.1	0.3
sunglasses	0.0	1.0

The goal is to minimize the μ-expected loss. More generally we define the Λ_σ prediction scheme, which minimizes the σ-expected loss:

$$y_t^{\Lambda_\sigma} \;:=\; \arg\min_{y_t\in\mathcal{Y}}\sum_{x_t}\sigma(x_t|x_{<t})\text{Loss}(x_t,y_t)$$

Loss Bounds

- Definition: μ-expected loss when Λ_σ predicts the t^{th} symbol:
  $\text{Loss}_t(\Lambda_\sigma)(x_{<t}) := \sum_{x_t}\mu(x_t|x_{<t})\text{Loss}(x_t,y_t^{\Lambda_\sigma})$

- $\text{Loss}_t(\Lambda_{\mu/\xi})$ made by the informed/universal scheme $\Lambda_{\mu/\xi}$.
  $\text{Loss}_t(\Lambda_\mu)\leq\text{Loss}_t(\Lambda)\ \forall t,\Lambda$.

- $\boxed{\text{Theorem: } 0\leq\text{Loss}_t(\Lambda_\xi)-\text{Loss}_t(\Lambda_\mu)\leq\sum_{x_t}|\xi_t-\mu_t|\leq\sqrt{2d_t}\stackrel{w.p.1}{\longrightarrow}0}$

- Total $\text{Loss}_{1:n}(\Lambda_\sigma):=\sum_{t=1}^n\mathbf{E}[\text{Loss}_t(\Lambda_\sigma)]$.

- $\boxed{\text{Theorem: } \text{Loss}_{1:n}(\Lambda_\xi)-\text{Loss}_{1:n}(\Lambda_\mu)\leq 2D_n+2\sqrt{\text{Loss}_{1:n}(\Lambda_\mu)D_n}}$

- Corollary: If $\text{Loss}_{1:\infty}(\Lambda_\mu)$ is finite, then $\text{Loss}_{1:\infty}(\Lambda_\xi)$ is finite, and $\text{Loss}_{1:n}(\Lambda_\xi)/\text{Loss}_{1:\infty}(\Lambda_\mu)\to 1$ if $\text{Loss}_{1:\infty}(\Lambda_\mu)\to\infty$.

- Remark: Holds for any loss function $\in[0,1]$ with no assumptions (like i.i.d., Markovian, stationary, ergodic, ...) on $\mu\in\mathcal{M}$.

Proof of Instantaneous Loss Bounds

Abbreviations: $\mathcal{X}=\{1,...,N\}$, $N=|\mathcal{X}|$, $i=x_t$, $y_i=\mu(x_t|x_{<t})$,
$z_i=\xi(x_t|x_{<t})$, $m=y_t^{\Lambda_\mu}$, $s=y_t^{\Lambda_\xi}$, $\ell_{xy}=\text{Loss}(x,y)$.

This and definition of $y_t^{\Lambda_\mu}$ and $y_t^{\Lambda_\xi}$ and $\sum_i z_i\ell_{is}\leq\sum_i z_i\ell_{ij}\,\forall j$ implies

$$\text{Loss}_t(\Lambda_\xi)-\text{Loss}_t(\Lambda_\mu)\equiv\sum_i y_i\ell_{is}-\sum_i y_i\ell_{im}\stackrel{(a)}{\leq}\sum_i(y_i-z_i)(\ell_{is}-\ell_{im})$$

$$\leq\sum_i|y_i-z_i|\cdot|\ell_{is}-\ell_{im}|\stackrel{(b)}{\leq}\sum_i|y_i-z_i|\stackrel{(c)}{\leq}\sqrt{\sum_i y_i\ln\frac{y_i}{z_i}}\equiv\sqrt{2d_t(x_{<t})}$$

(a) We added $\sum_i z_i(\ell_{im}-\ell_{is})\geq 0$.

(b) $|\ell_{is}-\ell_{im}|\leq 1$ since $\ell\in[0,1]$.

(c) Pinsker's inequality (elementary, but not trivial)

Generalization: Continuous Classes $\mathcal{M}$

In statistical parameter estimation one often has a continuous hypothesis class (e.g. a Bernoulli(θ) process with unknown $\theta \in [0,1]$).

$$\mathcal{M} := \{\nu_\theta : \theta \in I\!R^d\}, \quad \xi(x) := \int_{I\!R^d} d\theta\, w(\theta)\, \nu_\theta(x), \quad \int_{I\!R^d} d\theta\, w(\theta) = 1$$

Under weak regularity conditions [CB90,H'03]:

$$\boxed{\text{Theorem: } D_n(\mu||\xi) \;\leq\; \ln w(\mu)^{-1} + \tfrac{d}{2}\ln\tfrac{n}{2\pi} + O(1)}$$

where $O(1)$ depends on the local curvature (parametric complexity) of $\ln \nu_\theta$, and is independent n for many reasonable classes, including all stationary (k^{th}-order) finite-state Markov processes ($k=0$ is i.i.d.).

$D_n \propto \log(n) = o(n)$ still implies excellent prediction and decision for most n. [RH'07]

Bayesian Sequence Prediction: Summary

- The aim of probability theory is to describe uncertainty.
- Various sources and interpretations of uncertainty: frequency, objective, and subjective probabilities.
- They all respect the same rules.
- General sequence prediction: Use known (subj.) Bayes mixture $\xi = \sum_{\nu \in \mathcal{M}} w_\nu \nu$ in place of unknown (obj.) true distribution μ.
- Bound on the relative entropy between ξ and μ.

$\Rightarrow$ posterior of ξ converges rapidly to the true posterior μ.

- ξ is also optimal in a decision-theoretic sense w.r.t. any bounded loss function.
- No structural assumptions on $\mathcal{M}$ and $\nu \in \mathcal{M}$.

UNIVERSAL INDUCTIVE INFERENCE

- Foundations of Universal Induction
- Bayesian Sequence Prediction and Confirmation
- Convergence and Decisions
- How to Choose the Prior – Universal
- Kolmogorov Complexity
- How to Choose the Model Class – Universal
- The Problem of Zero Prior
- Reparametrization and Regrouping Invariance
- The Problem of Old Evidence / New Theories
- Universal is Better than Continuous Class
- More Bounds / Stuff / Critique / Problems
- Summary / Outlook / Literature

Universal Inductive Inference: Abstract

Solomonoff completed the Bayesian framework by providing a rigorous, unique, formal, and universal choice for the model class and the prior. I will discuss in breadth how and in which sense universal (non-i.i.d.) sequence prediction solves various (philosophical) problems of traditional Bayesian sequence prediction. I show that Solomonoff's model possesses many desirable properties: Strong total and weak instantaneous bounds , and in contrast to most classical continuous prior densities has no zero p(oste)rior problem, i.e. can confirm universal hypotheses, is reparametrization and regrouping invariant, and avoids the old-evidence and updating problem. It even performs well (actually better) in non-computable environments.

Induction Examples

Sequence prediction: Predict weather/stock-quote/... tomorrow, based on past sequence. Continue IQ test sequence like 1,4,9,16,?

Classification: Predict whether email is spam.
Classification can be reduced to sequence prediction.

Hypothesis testing/identification: Does treatment X cure cancer?
Do observations of white swans confirm that all ravens are black?

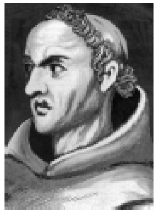
These are instances of the important problem of inductive inference or time-series forecasting or sequence prediction.

Problem: Finding prediction rules for every particular (new) problem is possible but cumbersome and prone to disagreement or contradiction.

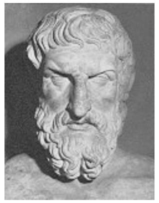
Goal: A single, formal, general, complete theory for prediction.

Beyond induction: active/reward learning, fct. optimization, game theory.

Foundations of Universal Induction



Ockhams' razor (simplicity) principle
Entities should not be multiplied beyond necessity.

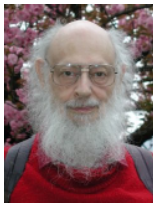


Epicurus' principle of multiple explanations
If more than one theory is consistent with the observations, keep all theories.

Bayes' rule for conditional probabilities
Given the prior belief/probability one can predict all future probabilities.

Turing's universal machine
Everything computable by a human using a fixed procedure can also be computed by a (universal) Turing machine.

Kolmogorov's complexity
The complexity or information content of an object is the length of its shortest description on a universal Turing machine.



Solomonoff's universal prior=Ockham+Epicurus+Bayes+Turing
Solves the question of how to choose the prior if nothing is known.
$\Rightarrow$ universal induction, formal Occam, AIT,MML,MDL,SRM,...

Bayesian Sequence Prediction and Confirmation

- Assumption: Sequence $\omega \in \mathcal{X}^\infty$ is sampled from the "true" probability measure μ, i.e. $\mu(x) := \mathrm{P}[x|\mu]$ is the μ-probability that ω starts with $x \in \mathcal{X}^n$.
- Model class: We assume that μ is unknown but known to belong to a countable class of environments=models=measures $\mathcal{M} = \{\nu_1, \nu_2, ...\}$. [no i.i.d./ergodic/stationary assumption]
- Hypothesis class: $\{H_\nu : \nu \in \mathcal{M}\}$ forms a mutually exclusive and complete class of hypotheses.
- Prior: $w_\nu := \mathrm{P}[H_\nu]$ is our prior belief in H_ν

$\Rightarrow$ Evidence: $\xi(x) := \mathrm{P}[x] = \sum_{\nu\in\mathcal{M}} \mathrm{P}[x|H_\nu]\mathrm{P}[H_\nu] = \sum_\nu w_\nu \nu(x)$ must be our (prior) belief in x.

$\Rightarrow$ Posterior: $w_\nu(x) := \mathrm{P}[H_\nu|x] = \frac{\mathrm{P}[x|H_\nu]\mathrm{P}[H_\nu]}{\mathrm{P}[x]}$ is our posterior belief in ν (Bayes' rule).

When is a Sequence Random?

a) Is 0110010100101101101001111011 generated by a fair coin flip?

b) Is 1111111111111111111111111111 generated by a fair coin flip?

c) Is 1100100100001111110110101010 generated by a fair coin flip?

d) Is 0101010101010101010101010101 generated by a fair coin flip?

- Intuitively: (a) and (c) look random, but (b) and (d) look unlikely.
- Problem: Formally (a-d) have equal probability $(\frac{1}{2})^{length}$.
- Classical solution: Consider hypothesis class $H := \{\text{Bernoulli}(p) : p \in \Theta \subseteq [0,1]\}$ and determine p for which sequence has maximum likelihood $\Longrightarrow$ (a,c,d) are fair Bernoulli($\frac{1}{2}$) coins, (b) not.
- Problem: (d) is non-random, also (c) is binary expansion of π.
- Solution: Choose H larger, but how large? Overfitting? MDL?
- AIT Solution: A sequence is **random** *iff* it is **incompressible**.

What does Probability Mean?

Naive frequency interpretation is circular:

- Probability of event E is $p := \lim_{n\to\infty} \frac{k_n(E)}{n}$,
  $n = \#$ i.i.d. trials, $k_n(E) = \#$ occurrences of event E in n trials.
- Problem: Limit may be anything (or nothing):
  e.g. a fair coin can give: Head, Head, Head, Head, ... $\Rightarrow p = 1$.
- Of course, for a fair coin this sequence is "unlikely".
  For fair coin, $p = 1/2$ with "high probability".
- But to make this statement rigorous we need to formally know what "high probability" means. Circularity!

Also: In complex domains typical for AI, sample size is often 1.
(e.g. a single non-iid historic weather data sequences is given).
We want to know whether certain properties hold for this *particular* seq.

How to Choose the Prior?

The probability axioms allow relating probabilities and plausibilities of different events, but they do not uniquely fix a numerical value for each event, except for the sure event Ω and the empty event $\{\}$.

We need new principles for determining values for at least some basis events from which others can then be computed.

There seem to be only 3 general principles:

- The principle of indifference — the symmetry principle
- The maximum entropy principle
- Occam's razor — the simplicity principle

Concrete: How shall we choose the hypothesis space $\{H_i\}$ and their prior $p(H_i)$ –or– $\mathcal{M} = \{\nu\}$ and their weight w_ν.

Indifference or Symmetry Principle

Assign same probability to all hypotheses:

$p(H_i) = \frac{1}{|I|}$ for finite I

$p(H_\theta) = [\mathrm{Vol}(\Theta)]^{-1}$ for compact and measurable Θ.

$\Rightarrow p(H_i|D) \propto p(D|H_i) \stackrel{\wedge}{=}$ classical Hypothesis testing (Max.Likelihood).

Prev. Example: H_θ =Bernoulli(θ) with $p(\theta) = 1$ for $\theta \in \Theta := [0,1]$.

Problems: Does not work for “large” hypothesis spaces:

(a) Uniform distr. on infinite $I = \mathbb{N}$ or noncompact Θ not possible!

(b) Reparametrization: $\theta \leadsto f(\theta)$. Uniform in θ is not uniform in $f(\theta)$.

Example: “Uniform” distr. on space of all (binary) sequences $\{0,1\}^\infty$:

$p(x_1...x_n) = (\frac{1}{2})^n \, \forall n \forall x_1...x_n \Rightarrow p(x_{n+1} = 1|x_1...x_n) = \frac{1}{2}$ always!

Inference so not possible (No-Free-Lunch myth).

Predictive setting: All we need is $p(x)$.

Occam's Razor — The Simplicity Principle

- Only Occam's razor (in combination with Epicurus' principle) is general enough to assign prior probabilities in *every* situation.

- The idea is to assign high (subjective) probability to simple events, and low probability to complex events.

- Simple events (strings) are more plausible a priori than complex ones.

- This gives (approximately) justice to both Occam's razor and Epicurus' principle.

Prefix Sets/Codes

String x is (proper) prefix of y $:\Longleftrightarrow$ $\exists\, z(\neq \epsilon)$ such that $xz = y$.

Set $\mathcal{P}$ is prefix-free or a prefix code $:\Longleftrightarrow$ no element is a proper prefix of another.

Example: A self-delimiting code (e.g. $\mathcal{P} = \{0, 10, 11\}$) is prefix-free.

Kraft Inequality

For a prefix code $\mathcal{P}$ we have $\sum_{x\in\mathcal{P}} 2^{-\ell(x)} \leq 1$.

Conversely, let l_1, l_2, ... be a countable sequence of natural numbers such that Kraft's inequality $\sum_k 2^{-l_k} \leq 1$ is satisfied. Then there exists a prefix code $\mathcal{P}$ with these lengths of its binary code.

Proof of the Kraft-Inequality

Proof $\Rightarrow$: Assign to each $x\in\mathcal{P}$ the interval $\Gamma_x := [0.x, 0.x+2^{-\ell(x)})$.

Length of interval Γ_x is $2^{-\ell(x)}$.

Intervals are disjoint, since $\mathcal{P}$ is prefix free, hence

$$\sum_{x\in\mathcal{P}} 2^{-\ell(x)} = \sum_{x\in\mathcal{P}} \text{Length}(\Gamma_x) \leq \text{Length}([0,1]) = 1$$

$\Leftarrow$: Idea: Choose l_1, l_2, ... in increasing order. Successively chop off intervals of lengths 2^{-l_1}, 2^{-l_2}, ... from left to right from $[0,1)$ and define left interval boundary as code.

Priors from Prefix Codes

- Let Code(H_ν) be a prefix code of hypothesis H_ν.
- Define complexity $Kw(\nu)$:=Length(Code(H_ν))
- Choose prior $w_\nu = p(H_\nu) = 2^{-Kw(\nu)}$
  $\Rightarrow \sum_{\nu\in\mathcal{M}} w_\nu \leq 1$ is semi-probability (by Kraft).
- How to choose a Code and hypothesis space $\mathcal{M}$?
- Praxis: Choose a code which is reasonable for your problem and $\mathcal{M}$ large enough to contain the true model.
- Theory: Choose a universal code and consider "all" hypotheses ...

Kolmogorov Complexity K(x)

K. of string x is the length of the shortest (prefix) program producing x:

$$\boxed{K(x) \;:=\; \min_p\{l(p) : U(p) = x\}, \quad U = \text{universal TM}}$$

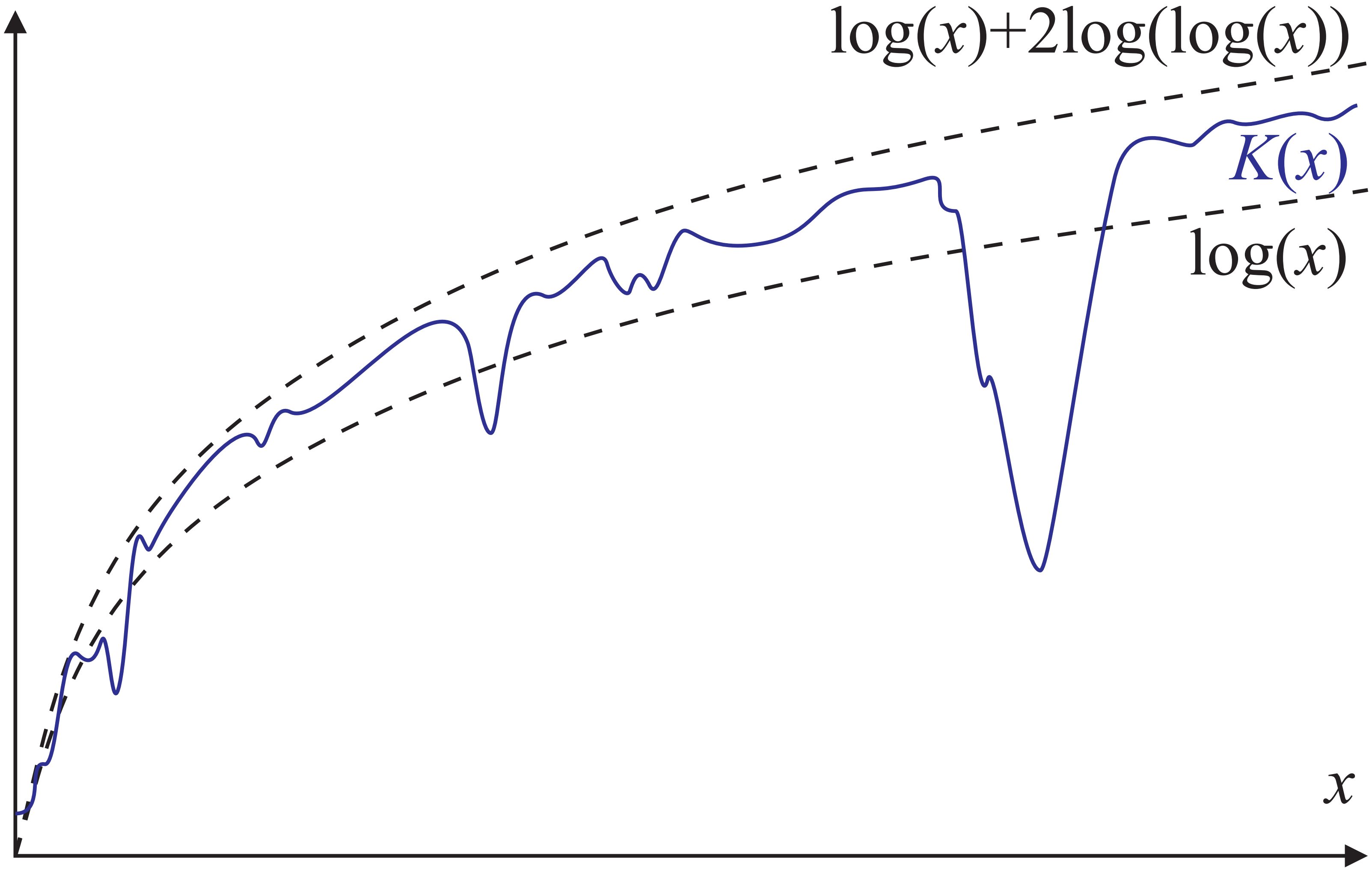
For non-string objects o (like numbers and functions) we define $K(o) := K(\langle o\rangle)$, where $\langle o\rangle \in \mathcal{X}^*$ is some standard code for o.

+ Simple strings like $000...0$ have small K, irregular (e.g. random) strings have large K.
- • The definition is nearly independent of the choice of U.
+ K satisfies most properties an information measure should satisfy.
+ K shares many properties with Shannon entropy but is superior.
– $K(x)$ is not computable, but only semi-computable from above.

Fazit: K is an excellent universal complexity measure, suitable for quantifying Occam's razor.

Schematic Graph of Kolmogorov Complexity

Although $K(x)$ is incomputable, we can draw a schematic graph



The Universal Prior

- Quantify the complexity of an environment ν or hypothesis H_ν by its Kolmogorov complexity $K(\nu)$.
- Universal prior: $w_\nu = \boxed{w_\nu^U := 2^{-K(\nu)}}$ is a decreasing function in the model's complexity, and sums to (less than) one.

$\Rightarrow$ $D_n \leq K(\mu)\ln 2$, i.e. the number of ε-deviations of ξ from μ or l^{Λ_ξ} from l^{Λ_μ} is proportional to the complexity of the environment.

- No other semi-computable prior leads to better prediction (bounds).
- For continuous $\mathcal{M}$, we can assign a (proper) universal prior (not density) $w_\theta^U = 2^{-K(\theta)} > 0$ for computable θ, and 0 for uncomp. θ.
- This effectively reduces $\mathcal{M}$ to a discrete class $\{\nu_\theta \in \mathcal{M} : w_\theta^U > 0\}$ which is typically dense in $\mathcal{M}$.
- This prior has many advantages over the classical prior (densities).

The Problem of Zero Prior

$=$ the problem of confirmation of universal hypotheses

Problem: If the prior is zero, then the posterior is necessarily also zero.

Example: Consider the hypothesis $H = H_1$ that all balls in some urn or all ravens are black (=1) or that the sun rises every day.

Starting with a prior density as $w(\theta) = 1$ implies that prior $\mathrm{P}[H_\theta] = 0$ for all θ, hence posterior $P[H_\theta|1..1] = 0$, hence H never gets confirmed.

3 non-solutions: define $H = \{\omega = 1^\infty\}$ | use finite population | abandon strict/logical/all-quantified/universal hypotheses in favor of soft hyp.

Solution: Assign non-zero prior to $\theta = 1 \quad \Rightarrow \quad \mathrm{P}[H|1^n] \to 1$.

Generalization: Assign non-zero prior to all "special" θ, like $\frac{1}{2}$ and $\frac{1}{6}$, which may naturally appear in a hypothesis, like "is the coin or die fair".

Universal solution: Assign non-zero prior to all comp. θ, e.g. $w_\theta^U = 2^{-K(\theta)}$

Reparametrization Invariance

- New parametrization e.g. $\psi = \sqrt{\theta}$, then the ψ-density $w'(\psi) = 2\sqrt{\theta}\, w(\theta)$ is no longer uniform if $w(\theta) = 1$ is uniform
  $\Rightarrow$ indifference principle is not reparametrization invariant (RIP).

- Jeffrey's and Bernardo's principle satisfy RIP w.r.t. differentiable bijective transformations $\psi = f^{-1}(\theta)$.

- The universal prior $w_\theta^U = 2^{-K(\theta)}$ also satisfies RIP w.r.t. simple computable f. (within a multiplicative constant)

Regrouping Invariance

- Non-bijective transformations:
  E.g. grouping ball colors into categories black/non-black.

- No classical principle is regrouping invariant.

- Regrouping invariance is regarded as a very important and desirable property. [Walley's (1996) solution: sets of priors]

- The universal prior $w_\theta^U = 2^{-K(\theta)}$ is invariant under regrouping, and more generally under all simple [computable with complexity $O(1)$] even non-bijective transformations. (within a multiplicative constant)

- Note: Reparametrization and regrouping invariance hold for arbitrary classes and are not limited to the i.i.d. case.

Universal Choice of Class $\mathcal{M}$

- The larger $\mathcal{M}$ the less restrictive is the assumption $\mu \in \mathcal{M}$.
- The class $\mathcal{M}_U$ of all (semi)computable (semi)measures, although only countable, is pretty large, since it includes all valid physics theories. Further, ξ_U is semi-computable [ZL70].
- Solomonoff's universal prior $M(x) :=$ probability that the output of a universal TM U with random input starts with x.
- Formally: $\boxed{M(x) := \sum_{p\,:\,U(p)=x*} 2^{-\ell(p)}}$ where the sum is over all (minimal) programs p for which U outputs a string starting with x.
- M may be regarded as a $2^{-\ell(p)}$-weighted mixture over all deterministic environments ν_p. ($\nu_p(x) = 1$ if $U(p) = x*$ and 0 else)
- $M(x)$ coincides with $\xi_U(x)$ within an irrelevant multiplicative constant.

The Problem of Old Evidence / New Theories

- What if some evidence $E\widehat{=}x$ (e.g. Mercury's perihelion advance) is known well-before the correct hypothesis/theory/model $H\widehat{=}\mu$ (Einstein's general relativity theory) is found?
- How shall H be added to the Bayesian machinery a posteriori?
- What should the "prior" of H be?
- Should it be the belief in H in a hypothetical counterfactual world in which E is not known?
- Can old evidence E confirm H?
- After all, H could simply be constructed/biased/fitted towards "explaining" E.

Solution of the Old-Evidence Problem

- The universal class $\mathcal{M}_U$ and universal prior w_ν^U formally solves this problem.

- The universal prior of H is $2^{-K(H)}$ independent of $\mathcal{M}$ and of whether E is known or not.

- Updating $\mathcal{M}$ is unproblematic, and even not necessary when starting with $\mathcal{M}_U$, since it includes all hypothesis (including yet unknown or unnamed ones) a priori.

Universal is Better than Continuous $\mathcal{M}$

- Although $\nu_\theta()$ and w_θ are incomp. for cont. classes $\mathcal{M}$ for most θ, $\xi()$ is typically computable. (exactly as for Laplace or numerically)

$\Rightarrow$ $\boxed{D_n(\mu||M) \stackrel{+}{<} D_n(\mu||\xi)+K(\xi)\ln 2 \text{ for all } \mu}$

- That is, M is superior to all computable mixture predictors ξ based on any (continuous or discrete) model class $\mathcal{M}$ and weight $w(\theta)$, save an additive constant $K(\xi)\ln 2 = O(1)$, even if environment μ is not computable.

- While $D_n(\mu||\xi) \sim \frac{d}{2}\ln n$ for all $\mu \in \mathcal{M}$, $D_n(\mu||M) \leq K(\mu)\ln 2$ is even finite for computable μ.

Fazit: $\boxed{\text{Solomonoff prediction works also in non-computable environments}}$

Convergence and Loss Bounds

- Total (loss) bounds: $\sum_{n=1}^{\infty} \mathbf{E}[h_n] \overset{\times}{<} K(\mu)\ln 2$, where $h_t(\omega_{<t}) := \sum_{a\in\mathcal{X}}(\sqrt{\xi(a|\omega_{<t})} - \sqrt{\mu(a|\omega_{<t})})^2$.
- Instantaneous i.i.d. bounds: For i.i.d. $\mathcal{M}$ with continuous, discrete, and universal prior, respectively:
  $\mathbf{E}[h_n] \overset{\times}{<} \frac{1}{n}\ln w(\mu)^{-1}$ and $\mathbf{E}[h_n] \overset{\times}{<} \frac{1}{n}\ln w_\mu^{-1} = \frac{1}{n}K(\mu)\ln 2$.
- Bounds for computable environments: Rapidly $M(x_t|x_{<t}) \to 1$ on every computable sequence $x_{1:\infty}$ (whichsoever, e.g. 1^∞ or the digits of π or e), i.e. M quickly recognizes the structure of the sequence.
- Weak instantaneous bounds: valid for all n and $x_{1:n}$ and $\bar{x}_n \neq x_n$: $2^{-K(n)} \overset{\times}{<} M(\bar{x}_n|x_{<n}) \overset{\times}{<} 2^{2K(x_{1:n}*)-K(n)}$
- Magic instance numbers: e.g. $M(0|1^n) \overset{\times}{=} 2^{-K(n)} \to 0$, but spikes up for simple n. M is cautious at magic instance numbers n.
- Future bounds / errors to come: If our past observations $\omega_{1:n}$ contain a lot of information about μ, we make few errors in future: $\sum_{t=n+1}^{\infty} \mathbf{E}[h_t|\omega_{1:n}] \overset{+}{<} [K(\mu|\omega_{1:n})+K(n)]\ln 2$

More Stuff / Critique / Problems

- Prior knowledge y can be incorporated by using "subjective" prior $w_{\nu|y}^U = 2^{-K(\nu|y)}$ or by prefixing observation x by y.
- Additive/multiplicative constant fudges and U-dependence is often (but not always) harmless.
- Incomputability: K and M can serve as "gold standards" which practitioners should aim at, but have to be (crudely) approximated in practice (MDL [Ris89], MML [Wal05], LZW [LZ76], CTW [WSTT95], NCD [CV05]).
- The Minimum Description Length Principle: $M(x) \approx 2^{-K_U(x)} \approx 2^{-K_T(x)}$. Predict y of highest $M(y|x)$ is approximately same as MDL: Predict y of smallest $K_T(xy)$.

Universal Inductive Inference: Summary

Universal Solomonoff prediction solves/avoids/meliorates many problems of (Bayesian) induction. We discussed:

+ general total bounds for generic class, prior, and loss,
+ i.i.d./universal-specific instantaneous and future bounds,
+ the D_n bound for continuous classes,
+ indifference/symmetry principles,
+ the problem of zero p(oste)rior & confirm. of universal hypotheses,
+ reparametrization and regrouping invariance,
+ the problem of old evidence and updating,
+ that M works even in non-computable environments,
+ how to incorporate prior knowledge,
− the prediction of short sequences,
− the constant fudges in all results and the U-dependence,
− M's incomputability and crude practical approximations.

Outlook

- Relation to Prediction with Expert Advice
- Relation to the Minimal Description Length (MDL) Principle
- Generalization to Active/Reinforcement learning (AIXI)

Open Problems

- Prediction of selected bits
- Convergence of M on Martin-Loef random sequences
- Better instantaneous bounds for M
- Future bounds for Bayes (general ξ)

UNIVERSAL RATIONAL AGENTS

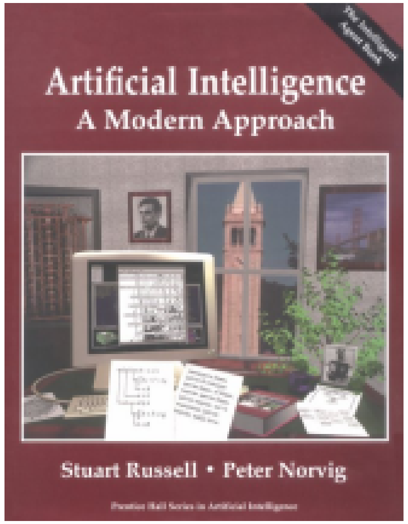
- Rational agents
- Sequential decision theory
- Reinforcement learning
- Value function
- Universal Bayes mixture and AIXI model
- Self-optimizing and Pareto-optimal policies
- Environmental Classes
- The horizon problem
- Computational Issues

Universal Rational Agents: Abstract

Sequential decision theory formally solves the problem of rational agents in uncertain worlds if the true environmental prior probability distribution is known. Solomonoff's theory of universal induction formally solves the problem of sequence prediction for unknown prior distribution.

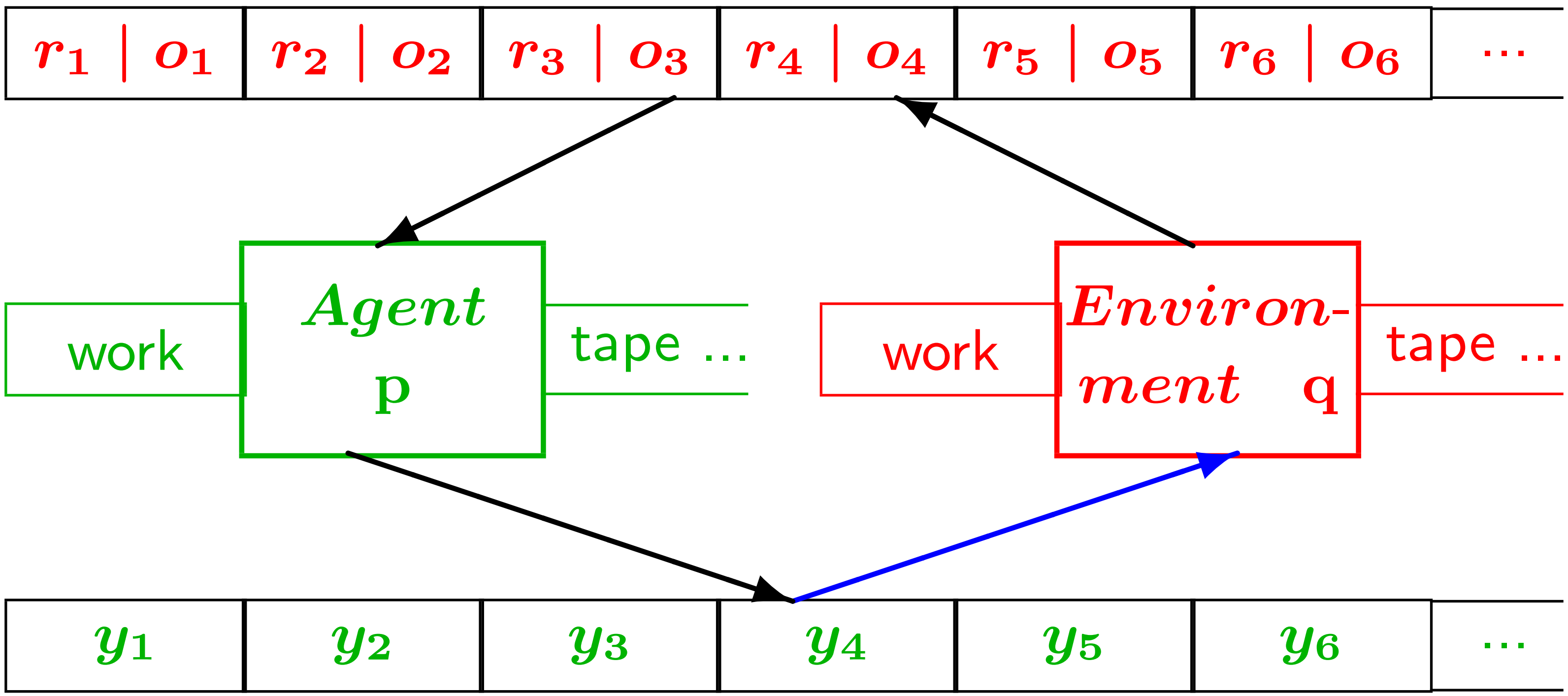
Here we combine both ideas and develop an elegant parameter-free theory of an optimal reinforcement learning agent embedded in an arbitrary unknown environment that possesses essentially all aspects of rational intelligence. The theory reduces all conceptual AI problems to pure computational ones.

There are strong arguments that the resulting AIXI model is the most intelligent unbiased agent possible. Other discussed topics are relations between problem classes, the horizon problem, and computational issues.



The Agent Model

Most if not all AI problems can be formulated within the agent framework



Rational Agents in Deterministic Environments

- $p:\mathcal{X}^*\to\mathcal{Y}^*$ is deterministic policy of the agent,
  $p(x_{<k})=y_{1:k}$ with $x_{<k}\equiv x_1...x_{k-1}$.
- $q:\mathcal{Y}^*\to\mathcal{X}^*$ is deterministic environment,
  $q(y_{1:k})=x_{1:k}$ with $y_{1:k}\equiv y_1...y_k$.
- Input $x_k\equiv r_k o_k$ consists of a regular informative part o_k
  and reward $r_k\in[0..r_{max}]$.
- Value $V_{km}^{pq}:=r_k+...+r_m$,
  optimal policy $p^{best}:=\arg\max_p V_{1m}^{pq}$,
  Lifespan or initial horizon m.

Agents in Probabilistic Environments

Given history $y_{1:k}x_{<k}$, the probability that the environment leads to perception x_k in cycle k is (by definition) $\sigma(x_k|y_{1:k}x_{<k})$.

Abbreviation (chain rule)

$$\sigma(x_{1:m}|y_{1:m}) \;=\; \sigma(x_1|y_1)\cdot\sigma(x_2|y_{1:2}x_1)\cdot\,...\,\cdot\sigma(x_m|y_{1:m}x_{<m})$$

The average value of policy p with horizon m in environment σ is defined as

$$V_\sigma^p \;:=\; \tfrac{1}{m}\sum_{x_{1:m}}(r_1+\,...\,+r_m)\sigma(x_{1:m}|y_{1:m})_{|y_{1:m}=p(x_{<m})}$$

The goal of the agent should be to maximize the value.

Optimal Policy and Value

The σ-optimal policy $p^\sigma := \arg\max_p V_\sigma^p$ maximizes $V_\sigma^p \leq V_\sigma^* := V_\sigma^{p^\sigma}$.

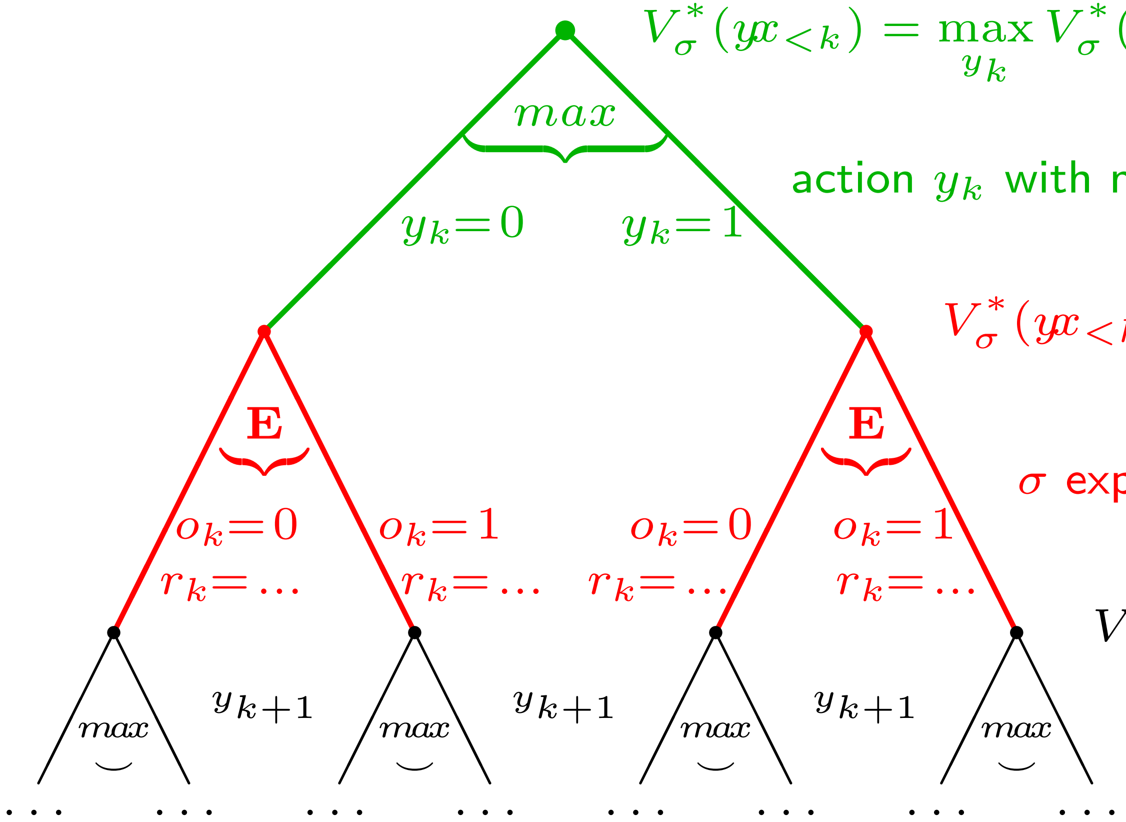
Explicit expressions for the action y_k in cycle k of the σ-optimal policy p^σ and their value V_σ^* are

$$y_k \;=\; \arg\max_{y_k}\sum_{x_k}\max_{y_{k+1}}\sum_{x_{k+1}}\;...\;\max_{y_m}\sum_{x_m}(r_k+...+r_m)\cdot\sigma(x_{k:m}|y_{1:m}x_{<k}),$$

$$V_\sigma^* \;=\; \textstyle\frac{1}{m}\displaystyle\max_{y_1}\sum_{x_1}\max_{y_2}\sum_{x_2}\;...\;\max_{y_m}\sum_{x_m}(r_1+...+r_m)\cdot\sigma(x_{1:m}|y_{1:m}).$$

Keyword: Expectimax tree/algorithm.

Expectimax Tree/Algorithm



$$V_\sigma^*(y\!x_{<k}) = \max_{y_k} V_\sigma^*(y\!x_{<k}y_k)$$

action y_k with max value.

$$V_\sigma^*(y\!x_{<k}y_k) = \sum_{x_k}[r_k + V_\sigma^*(y\!x_{1:k})]\sigma(x_k|y\!x_{<k}y_k)$$

σ expected reward r_k and observation o_k.

$$V_\sigma^*(y\!x_{1:k}) = \max_{y_{k+1}} V_\sigma^*(y\!x_{1:k}y_{k+1})$$

Known environment μ

- Assumption: μ is the true environment in which the agent operates
- Then, policy p^μ is optimal in the sense that no other policy for an agent leads to higher μ^{AI}-expected reward.
- Special choices of μ: deterministic or adversarial environments, Markov decision processes (MDPs), adversarial environments.
- There is no principle problem in computing the optimal action y_k as long as μ^{AI} is known and computable and $\mathcal{X}$, $\mathcal{Y}$ and m are finite.
- Things drastically change if μ^{AI} is unknown ...

The Bayes-mixture distribution ξ

Assumption: The true environment μ is unknown.

Bayesian approach: The true probability distribution μ^{AI} is not learned directly, but is replaced by a Bayes-mixture ξ^{AI}.

Assumption: We know that the true environment μ is contained in some known (finite or countable) set $\mathcal{M}$ of environments.

The Bayes-mixture ξ is defined as

$$\xi(x_{1:m}|y_{1:m}) := \sum_{\nu\in\mathcal{M}} w_\nu \nu(x_{1:m}|y_{1:m}) \quad \text{with} \quad \sum_{\nu\in\mathcal{M}} w_\nu = 1, \quad w_\nu > 0\ \forall\nu$$

The weights w_ν may be interpreted as the prior degree of belief that the true environment is ν.

Then $\xi(x_{1:m}|y_{1:m})$ could be interpreted as the prior subjective belief probability in observing $x_{1:m}$, given actions $y_{1:m}$.

Questions of Interest

- It is natural to follow the policy p^ξ which maximizes V_ξ^p.
- If μ is the true environment the expected reward when following policy p^ξ will be $V_\mu^{p^\xi}$.
- The optimal (but infeasible) policy p^μ yields reward $V_\mu^{p^\mu} \equiv V_\mu^*$.
- Are there policies with uniformly larger value than $V_\mu^{p^\xi}$?
- How close is $V_\mu^{p^\xi}$ to V_μ^*?
- What is the most general class $\mathcal{M}$ and weights w_ν.

A universal choice of ξ and $\mathcal{M}$

- We have to assume the existence of some structure on the environment to avoid the No-Free-Lunch Theorems [Wolpert 96].
- We can only unravel effective structures which are describable by (semi)computable probability distributions.
- So we may include all (semi)computable (semi)distributions in $\mathcal{M}$.
- Occam's razor and Epicurus' principle of multiple explanations tell us to assign high prior belief to simple environments.
- Using Kolmogorov's universal complexity measure $K(\nu)$ for environments ν one should set $w_\nu \sim 2^{-K(\nu)}$, where $K(\nu)$ is the length of the shortest program on a universal TM computing ν.
- The resulting AIXI model [Hutter:00] is a unification of (Bellman's) sequential decision and Solomonoff's universal induction theory.

The AIXI Model in one Line

complete & essentially unique & limit-computable

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

k=now, $a$ction, $o$bservation, $r$eward, Universal TM, $p$rogram, m=lifespan

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: Strategic Games, Function Optimization, Supervised Learning, Sequence Prediction, Classification, ...

In the following we consider generic $\mathcal{M}$ and w_ν.

Pareto-Optimality of p^ξ

Policy p^ξ is Pareto-optimal in the sense that there is no other policy p with $V_\nu^p \geq V_\nu^{p^\xi}$ for all $\nu \in \mathcal{M}$ and strict inequality for at least one ν.

Self-optimizing Policies

Under which circumstances does the value of the universal policy p^ξ converge to optimum?

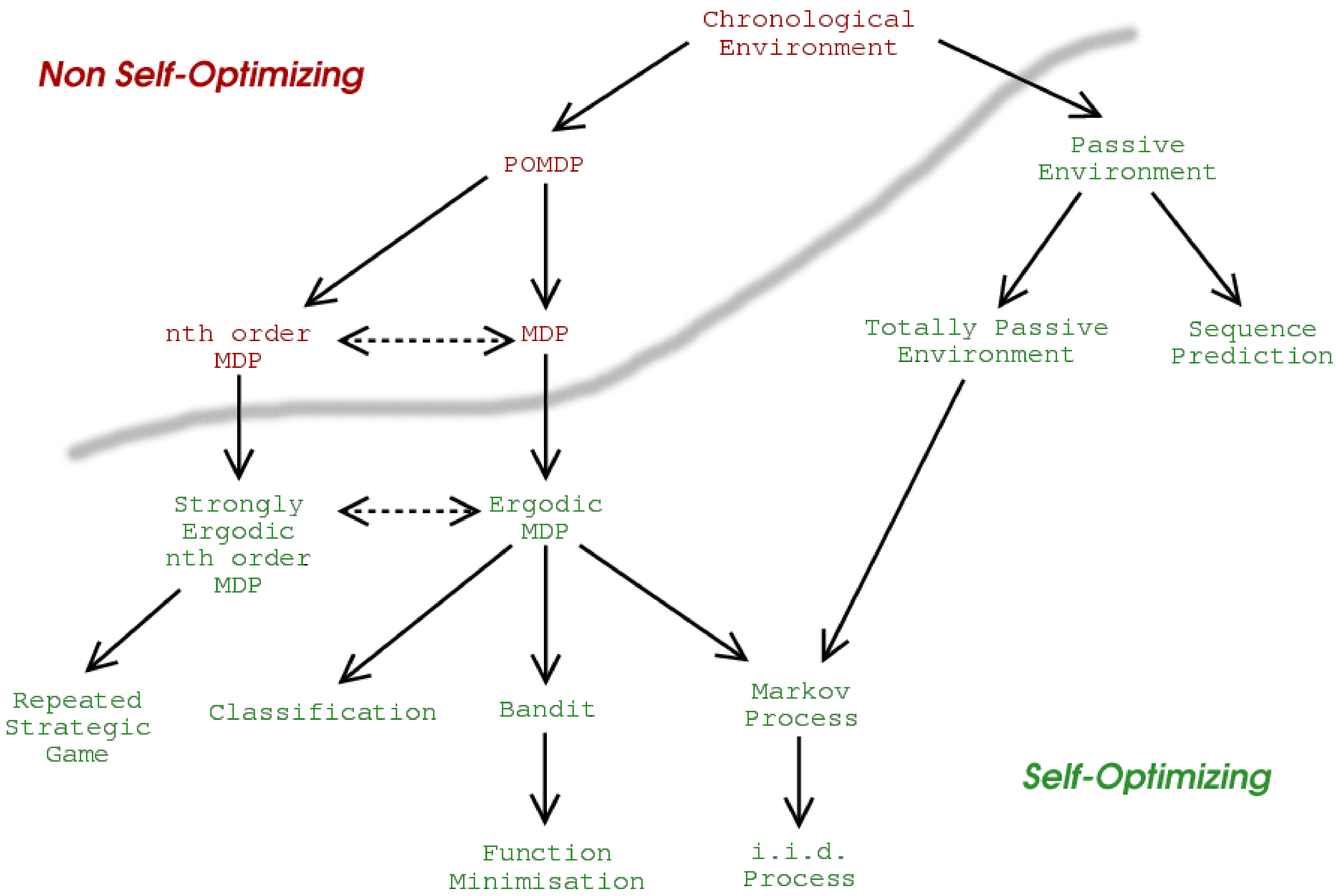
$$V_\nu^{p^\xi} \to V_\nu^* \quad \text{for horizon} \quad m \to \infty \quad \text{for all} \quad \nu \in \mathcal{M}. \qquad (1)$$

The least we must demand from $\mathcal{M}$ to have a chance that (1) is true is that there exists some policy $\tilde{p}$ at all with this property, i.e.

$$\exists \tilde{p}: \; V_\nu^{\tilde{p}} \to V_\nu^* \quad \text{for horizon} \quad m \to \infty \quad \text{for all} \quad \nu \in \mathcal{M}. \qquad (2)$$

Main result: $(2) \Rightarrow (1)$: The necessary condition of the existence of a self-optimizing policy $\tilde{p}$ is also sufficient for p^ξ to be self-optimizing.

Environments w. (Non)Self-Optimizing Policies



Particularly Interesting Environments

- Sequence Prediction, e.g. weather or stock-market prediction.
  Strong result: $V_\mu^* - V_\mu^{p^\xi} = O(\sqrt{\frac{K(\mu)}{m}})$, $m =$horizon.

- Strategic Games: Learn to play well (minimax) strategic zero-sum games (like chess) or even exploit limited capabilities of opponent.

- Optimization: Find (approximate) minimum of function with as few function calls as possible. Difficult exploration versus exploitation problem.

- Supervised learning: Learn functions by presenting $(z, f(z))$ pairs and ask for function values of z' by presenting $(z', ?)$ pairs.
  Supervised learning is much faster than reinforcement learning.

AIξ quickly learns to predict, play games, optimize, and learn supervised.

Future Value and the Right Discounting

- Eliminate the arbitrary horizon parameter m by discounting the rewards $r_k \rightsquigarrow \gamma_k r_k$ with $\Gamma_k := \sum_{i=k}^{\infty} \gamma_i < \infty$ and letting $m \to \infty$:
$V_{k\gamma}^{\pi\sigma} :=$
$$\frac{1}{\Gamma_k} \lim_{m\to\infty} \sum_{x_{k:m}} (\gamma_k r_k + ... + \gamma_m r_m) \sigma(x_{k:m}|y_{1:m} x_{<k})_{|y_{1:m}=p(x_{<m})}$$
- If there exists a self-optimizing policy for $\mathcal{M}$, then p^ξ is self-optimizing: If $\exists \tilde{\pi}_k \forall \nu : V_{k\gamma}^{\tilde{\pi}_k \nu} \stackrel{k\to\infty}{\longrightarrow} V_{k\gamma}^{*\nu} \Rightarrow V_{k\gamma}^{p^\xi \mu} \stackrel{k\to\infty}{\longrightarrow} V_{k\gamma}^{*\mu}$.
- Standard geometric discounting: $\gamma_k = \gamma^k$ with $0 < \gamma < 1$.
Problem: Most environments do not possess self-optimizing policies under this discounting, since effective horizon h_k^{eff} is finite.
- Power discounting: $\gamma_k = k^{-2} \;\Rightarrow h_k^{\mathit{eff}} \sim k =$ agent's age.
Universal discounting: $\gamma_k = 2^{-K(k)} \Rightarrow h_k^{\mathit{eff}} \sim$ Ackermann(k)
- Result: Policy p^ξ is self-optimizing for ergodic MDPs if $\frac{\gamma_{k+1}}{\gamma_k} \to 1$.

Universal Rational Agents: Summary

- Setup: Agents acting in general probabilistic environments with reinforcement feedback.
- Assumptions: True environment μ belongs to a known class of environments $\mathcal{M}$, but is otherwise unknown.
- Results: The Bayes-optimal policy p^ξ based on the Bayes-mixture $\xi = \sum_{\nu\in\mathcal{M}} w_\nu \nu$ is Pareto-optimal and self-optimizing if $\mathcal{M}$ admits self-optimizing policies.
- Application: The class of ergodic MDPs admits self-optimizing policies.
- New: Policy p^ξ with unbounded effective horizon is the first purely Bayesian self-optimizing consistent policy for ergodic MDPs.
- Learn: The combined conditions $\Gamma_k < \infty$ and $\frac{\gamma_{k+1}}{\gamma_k} \to 1$ allow a consistent self-optimizing Bayes-optimal policy based on mixtures.

Universal Rational Agents: Remarks

- We have developed a parameterless AI model based on sequential decisions and algorithmic probability.
- We have reduced the AI problem to pure computational questions.
- AIξ seems not to lack any important known methodology of AI, apart from computational aspects.
- There is no need for implementing extra knowledge, as this can be learned by presenting it in o_k in any form.
- The learning process itself is an important aspect of AI.
- Noise or irrelevant information in the inputs do not disturb the AIξ system.
- Philosophical questions: relevance of non-computational physics (Penrose), number of wisdom Ω (Chaitin), consciousness, social consequences.

Universal Rational Agents: Outlook

- Continuous classes $\mathcal{M}$.
- Restricted policy classes.
- Non-asymptotic bounds.
- Tighter bounds by exploiting extra properties of the environments, like the mixing rate of MDPs.
- Search for other performance criteria [Hutter:00].
- Instead of convergence of the expected reward sum, study convergence with high probability of the actually realized reward sum.
- Other environmental classes (separability concepts, downscaling).

APPROXIMATIONS & APPLICATIONS

- Universal Similarity Metric
- Universal Search
- Time-Bounded AIXI Model
- Brute-Force Approximation of AIXI
- A Monte-Carlo AIXI Approximation
- Feature Reinforcement Learning
- Comparison to other approaches
- Future directions, wrap-up, references.

Approximations & Applications: Abstract

Many fundamental theories have to be approximated for practical use. Since the core quantities of universal induction and universal intelligence are incomputable, it is often hard, but not impossible, to approximate them. In any case, having these "gold standards" to approximate (top$\rightarrow$down) or to aim at (bottom$\rightarrow$up) is extremely helpful in building truly intelligent systems. The most impressive direct approximation of Kolmogorov complexity to-date is via the universal similarity metric applied to a variety of real-world clustering problems. A couple of universal search algorithms ((adaptive) Levin search, FastPrg, OOPS, Goedel-machine, ...) that find short programs have been developed and applied to a variety of toy problem. The AIXI model itself has been approximated in a couple of ways (AIXItl, Brute Force, Monte Carlo, Feature RL). Some recent applications will be presented. The Lecture Series concludes by comparing various learning algorithms along various dimensions, pointing to future directions, wrap-up, and references.

Conditional Kolmogorov Complexity

Question: When is object=string x similar to object=string y?

Universal solution: x similar y ⇔ x can be easily (re)constructed from y
⇔ Kolmogorov complexity $K(x|y) := \min\{\ell(p) : U(p,y) = x\}$ is small

Examples:

1) x is very similar to itself ($K(x|x) \stackrel{+}{=} 0$)

2) A processed x is similar to x ($K(f(x)|x) \stackrel{+}{=} 0$ if $K(f) = O(1)$).
   e.g. doubling, reverting, inverting, encrypting, partially deleting x.

3) A random string is with high probability not similar to any other string ($K(\text{random}|y) =$length(random)).

The problem with $K(x|y)$ as similarity=distance measure is that it is neither symmetric nor normalized nor computable.

The Universal Similarity Metric [CV'05]

- Symmetrization and normalization leads to a/the universal metric d:

$$0 \leq d(x,y) := \frac{\max\{K(x|y),K(y|x)\}}{\max\{K(x),K(y)\}} \leq 1$$

- Every effective similarity between x and y is detected by d
- Use $K(x|y) \approx K(xy) - K(y)$ and $K(x) \equiv K_U(x) \approx K_T(x)$ (coding) T $\Longrightarrow$ computable approximation: Normalized compression distance:

$$d(x,y) \approx \frac{K_T(xy) - \min\{K_T(x),K_T(y)\}}{\max\{K_T(x),K_T(y)\}} \lesssim 1$$

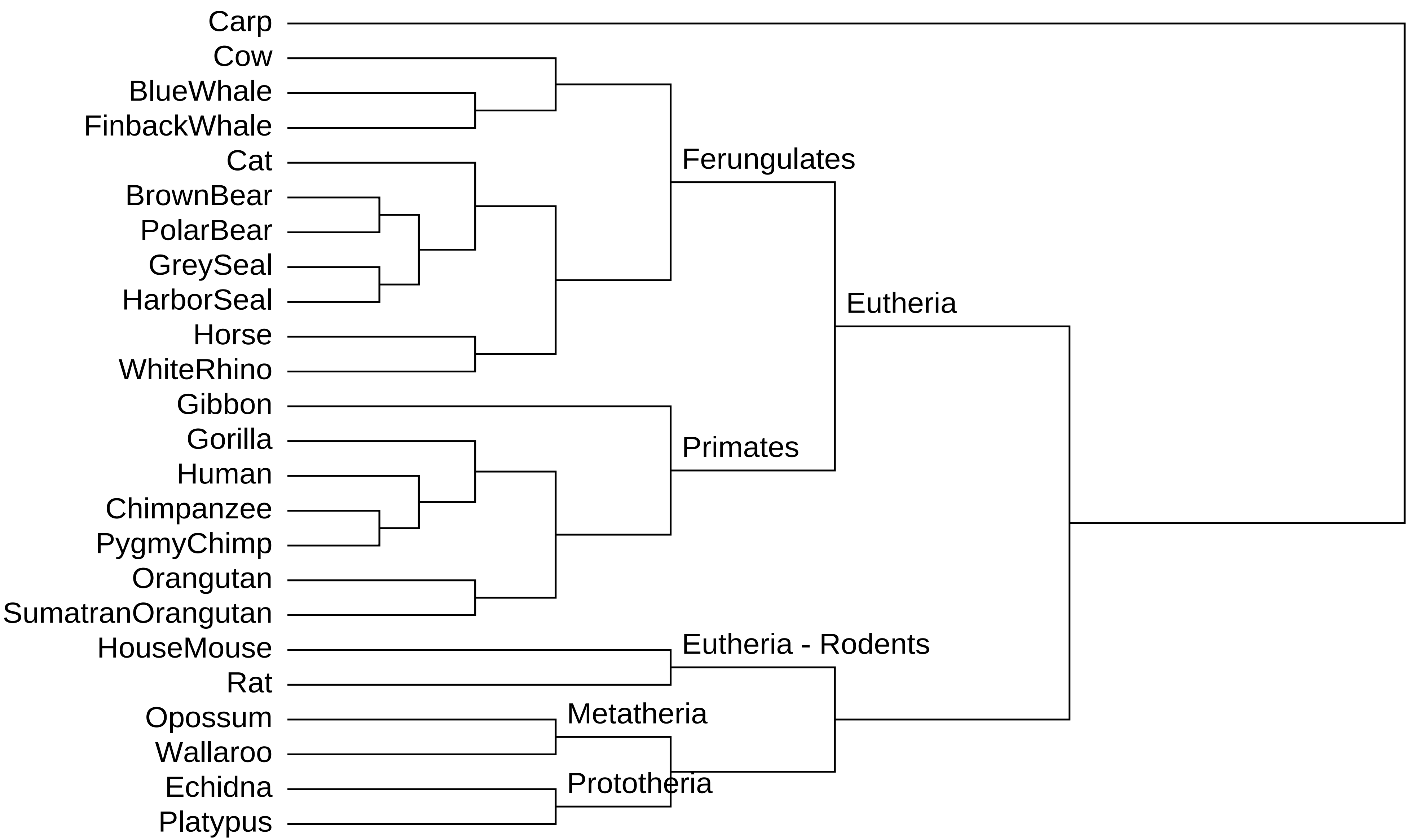
- For T choose Lempel-Ziv or gzip or bzip(2) (de)compressor in the applications below.
- Theory: Lempel-Ziv compresses asymptotically better than any probabilistic finite state automaton predictor/compressor.

Tree-Based Clustering [CV'05]

- If many objects $x_1, ..., x_n$ need to be compared, determine the Similarity matrix: $M_{ij} = d(x_i, x_j)$ for $1 \leq i, j \leq n$
- Now cluster similar objects.
- There are various clustering techniques.
- Tree-based clustering: Create a tree connecting similar objects,
- e.g. quartet method (for clustering)
- Applications: Phylogeny of 24 Mammal mtDNA,
  50 Language Tree (based on declaration of human rights),
  composers of music, authors of novels, SARS virus, fungi,
  optical characters, galaxies, ... [Cilibrasi&Vitanyi'05]

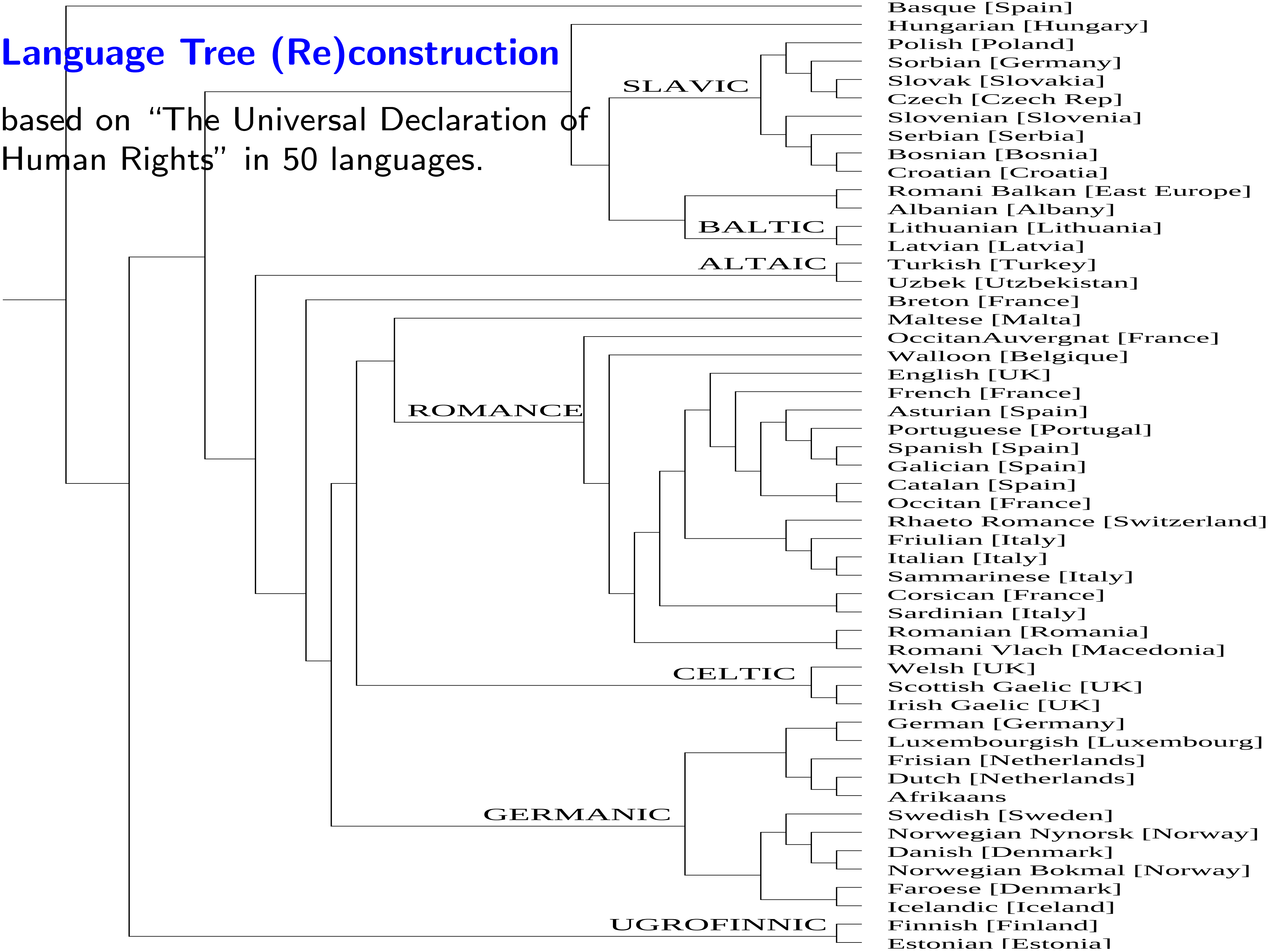
Genomics & Phylogeny: Mammals [CV'05]

Evolutionary tree built from complete mammalian mtDNA of 24 species:



Language Tree (Re)construction

based on “The Universal Declaration of Human Rights” in 50 languages.



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 (OOPS) 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€)

The Time-Bounded AIXI Model (AIXItl)

An algorithm p^{best} has been constructed for which the following holds:

- Let p be any (extended chronological) policy
- with length $\ell(p)\leq\tilde{l}$ and computation time per cycle $t(p)\leq\tilde{t}$
- for which there exists a proof of length $\leq l_P$ that p is a valid approximation.
- Then an algorithm p^{best} can be constructed, depending on $\tilde{l}$,$\tilde{t}$ and l_P but not on knowing p
- which is effectively more or equally intelligent according to $\succeq^c$ than any such p.
- The size of p^{best} is $\ell(p^{best})=O(\ln(\tilde{l}\cdot\tilde{t}\cdot l_P))$,
- the setup-time is $t_{setup}(p^{best})=O(l_P^2\cdot 2^{l_P})$,
- the computation time per cycle is $t_{cycle}(p^{best})=O(2^{\tilde{l}}\cdot\tilde{t})$.

Brute-Force Approximation of AIXI

- Truncate expectimax tree depth to a small fixed lookahead h.
  Optimal action computable in time $|\mathcal{Y}\times\mathcal{X}|^h\times$ time to evaluate ξ.

- Consider mixture over Markov Decision Processes (MDP) only, i.e.
  $\xi(x_{1:m}|y_{1:m})=\sum_{\nu\in\mathcal{M}} w_\nu \prod_{t=1}^m \nu(x_t|x_{t-1}y_t)$. Note: ξ is *not* MDP

- Choose uniform prior over w_μ.
  Then $\xi(x_{1:m}|y_{1:m})$ can be computed in linear time.

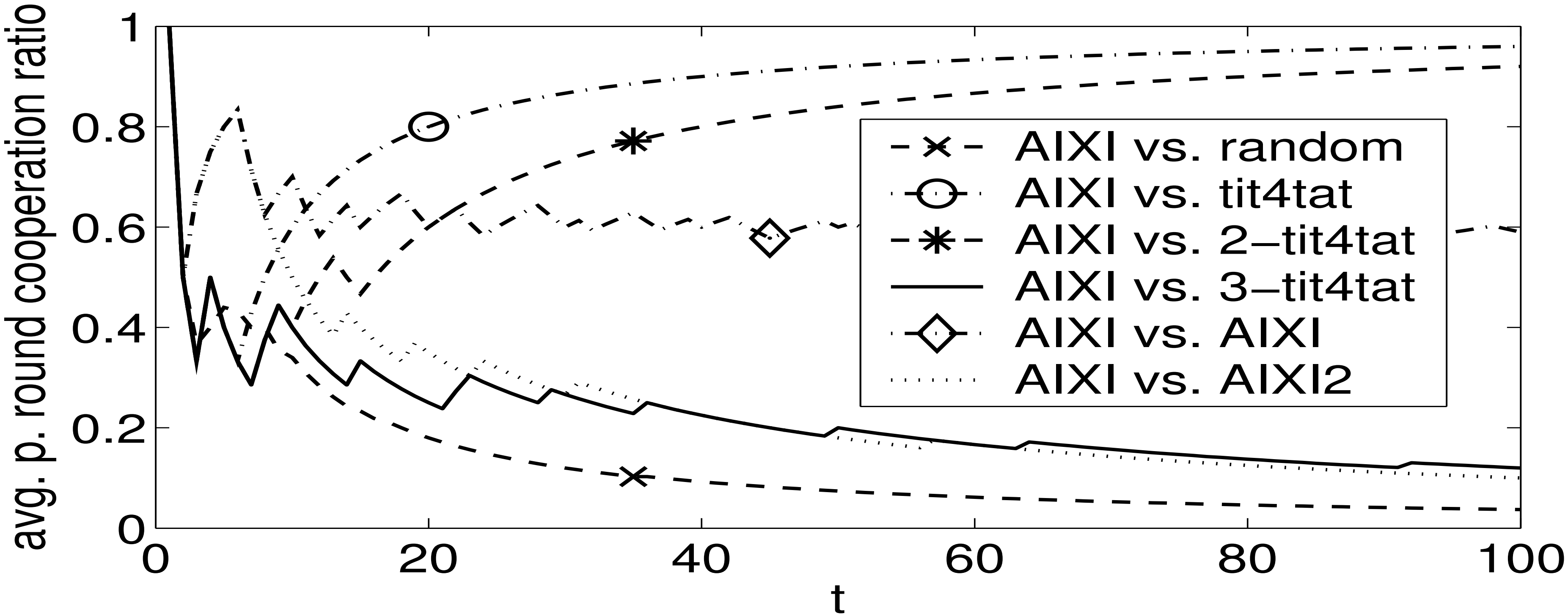
- Consider (approximately) Markov problems
  with very small action and perception space.

- Example application: 2×2 Matrix Games like Prisoner'S Dilemma,
  Stag Hunt, Chicken, Battle of Sexes, and Matching Pennies. [PH'06]

AIXI Learns to Play 2×2 Matrix Games

- Repeated prisoners dilemma.
- Game unknown to AIXI. Must be learned as well
- AIXI behaves appropriately.

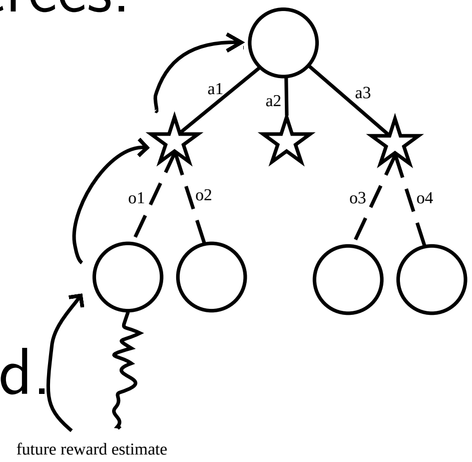
Loss matrix	cooperates	defects
cooperates	=0.3 years	=1 year
defects	free	=0.7 years



A Monte-Carlo AIXI Approximation

Consider class of Variable-Order Markov Decision Processes.

The Context Tree Weighting (CTW) algorithm can efficiently mix (exactly in essentially linear time) all prediction suffix trees.



Monte-Carlo approximation of expectimax tree:

Upper Confidence Tree (UCT) algorithm:

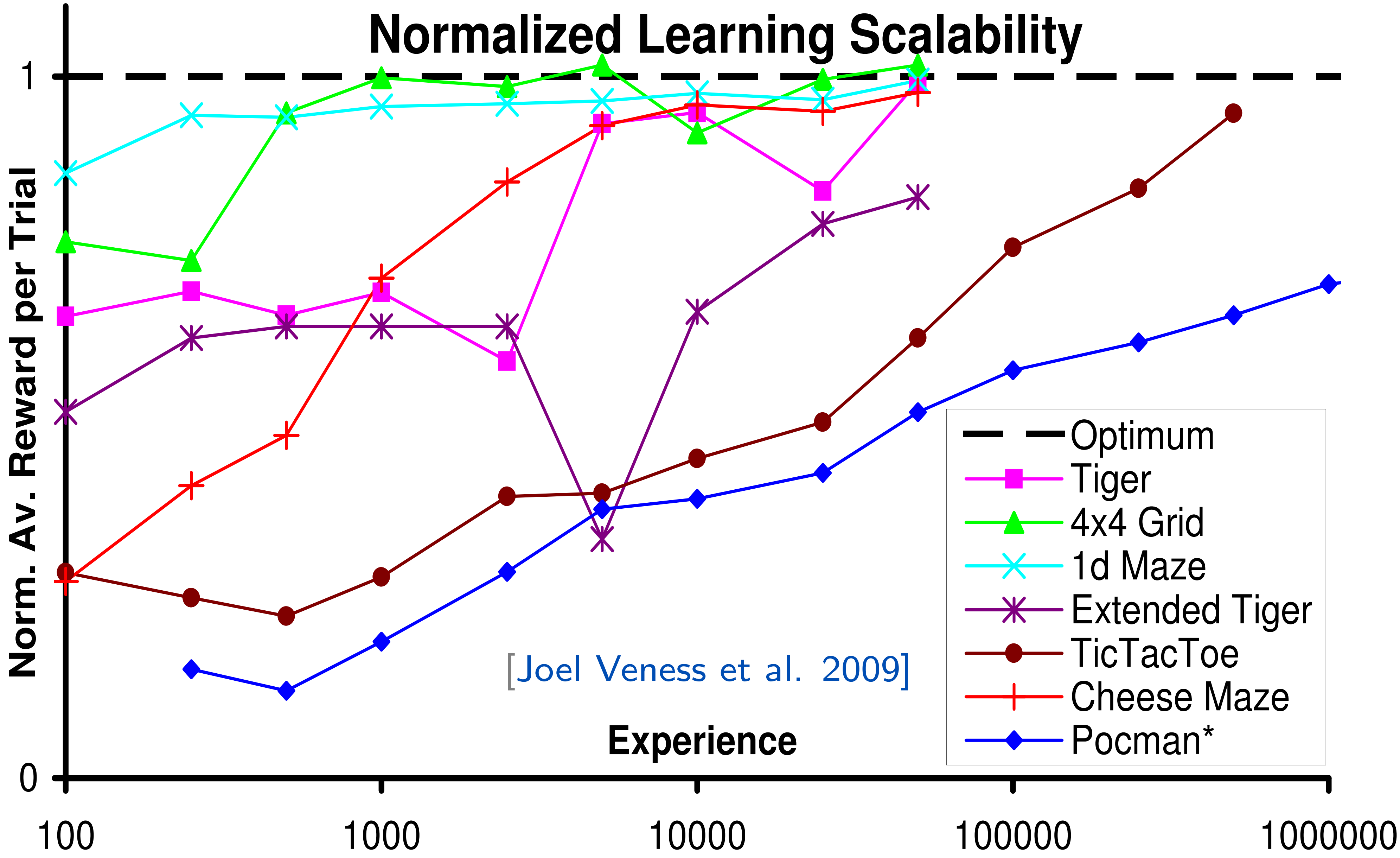
- Sample observations from CTW distribution.
- Select actions with highest upper confidence bound.
- Expand tree by one leaf node (per trajectory).
- Simulate from leaf node further down using (fixed) playout policy.
- Propagate back the value estimates for each node.

Repeat until timeout. [VNHS'09]

Guaranteed to converge to exact value.

Extension: Predicate CTW not based on raw obs. but features thereof.

Monte-Carlo AIXI Applications



Feature Reinforcement Learning (FRL)

Goal: Develop efficient general purpose intelligent agent.

State-of-the-art: (a) AIXI: Incomputable theoretical solution.
(b) MDP: Efficient limited problem class.
(c) POMDP: Notoriously difficult. (d) PSRs: Underdeveloped.

Idea: ΦMDP reduces real problem to MDP automatically by learning.

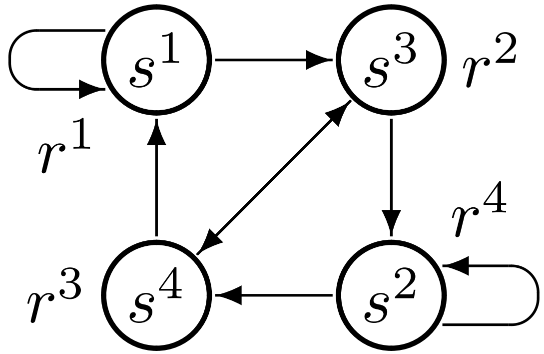
Accomplishments so far: (i) Criterion for evaluating quality of reduction.
(ii) Integration of the various parts into one learning algorithm.
(iii) Generalization to structured MDPs (DBNs)

ΦMDP is promising path towards the grand goal & alternative to (a)-(d)

Problem: Find reduction Φ efficiently (generic optimization problem?)

Markov Decision Processes (MDPs)

a computationally tractable class of problems



- MDP Assumption: State $s_t := o_t$ and r_t are probabilistic functions of o_{t-1} and a_{t-1} only.
- Further Assumption: State=observation space $\S$ is finite and small.
- Goal: Maximize long-term expected reward.
- Learning: Probability distribution is unknown but can be learned.
- Exploration: Optimal exploration is intractable but there are polynomial approximations.
- Problem: Real problems are not of this simple form.

Map Real Problem to MDP

Map history $h_t := o_1 a_1 r_1 ... o_{t-1}$ to state $s_t := \Phi(h_t)$, for example:

Games: Full-information with static opponent: $\Phi(h_t) = o_t$.

Classical physics: Position+velocity of objects = position at two time-slices: $s_t = \Phi(h_t) = o_t o_{t-1}$ is (2nd order) Markov.

I.i.d. processes of unknown probability (e.g. clinical trials $\simeq$ Bandits), Frequency of obs. $\Phi(h_n) = (\sum_{t=1}^n \delta_{o_t o})_{o\in\mathcal{O}}$ is sufficient statistic.

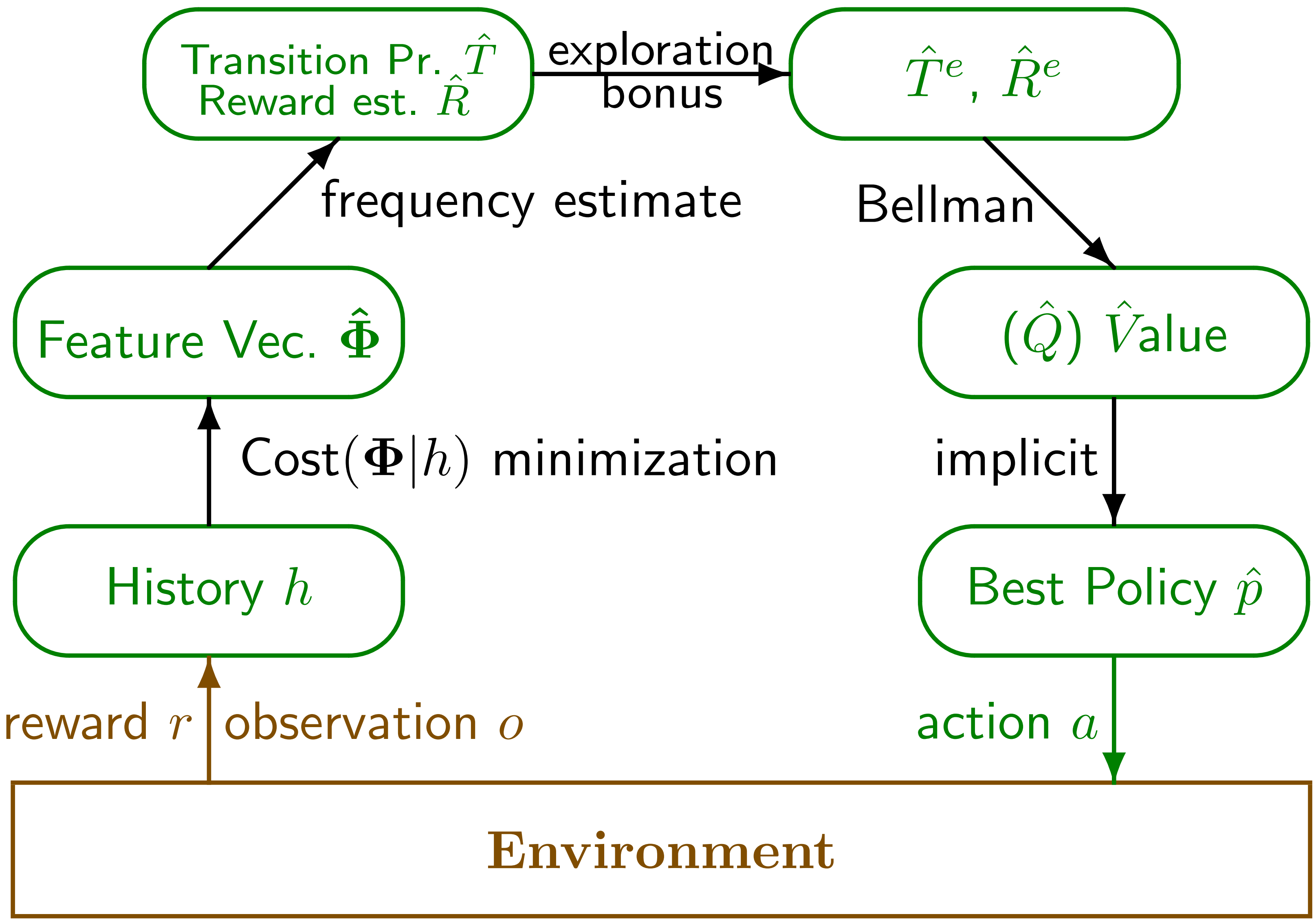
Identity: $\Phi(h) = h$ is always sufficient, but not learnable.

Find/Learn Map Automatically

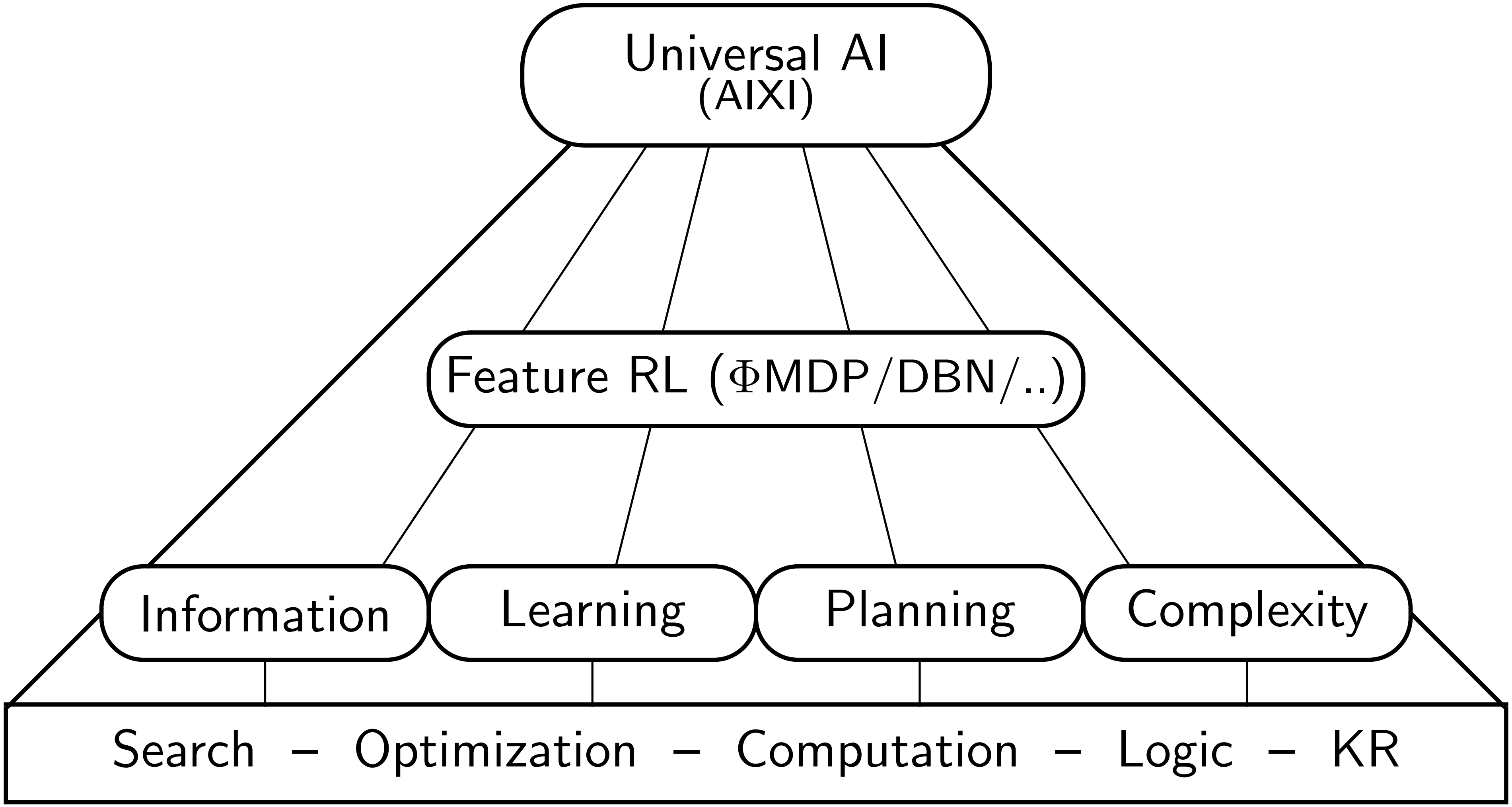
$$\Phi^{best} := \arg\min_\Phi \text{Cost}(\Phi|h_t)$$

- What is the best map/MDP? (i.e. what is the right Cost criterion?)
- Is the best MDP good enough? (i.e. is reduction always possible?)
- How to find the map Φ (i.e. minimize Cost) efficiently?

ΦMDP: Computational Flow



Intelligent Agents in Perspective



Agents = General Framework, Interface = Robots, Vision, Language

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
MC-AIXI-CTW	yes/no	yes	YES	YES	yes	NO	yes/no	YES	YES
Feature RL	yes/no	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

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.

The Big Questions

- Is non-computational physics relevant to AI? [Penrose]
- Could something like the number of wisdom Ω prevent a simple solution to AI? [Chaitin]
- Do we need to understand consciousness before being able to understand AI or construct AI systems?
- What if we succeed?

Wrap Up

- Setup: Given (non)iid data $D = (x_1, ..., x_n)$, predict x_{n+1}
- Ultimate goal is to maximize profit or minimize loss
- Consider Models/Hypothesis $H_i \in \mathcal{M}$
- Max.Likelihood: $H_{best} = \arg\max_i p(D|H_i)$ (overfits if $\mathcal{M}$ large)
- Bayes: Posterior probability of H_i is $p(H_i|D) \propto p(D|H_i)p(H_i)$
- Bayes needs prior(H_i)
- Occam+Epicurus: High prior for simple models.
- Kolmogorov/Solomonoff: Quantification of simplicity/complexity
- Bayes works if D is sampled from $H_{true} \in \mathcal{M}$
- Bellman equations tell how to optimally act in known environments
- Universal AI $=$ Universal Induction $+$ Sequential Decision Theory
- Practice $=$ approximate, restrict, search, optimize, knowledge

Literature

[CV05] R. Cilibrasi and P. M. B. Vitányi. Clustering by compression. *IEEE Trans. Information Theory*, 51(4):1523–1545, 2005.

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

[Hut07] M. Hutter. On universal prediction and Bayesian confirmation. *Theoretical Computer Science*, 384(1):33–48, 2007.

[LH07] S. Legg and M. Hutter. Universal intelligence: a definition of machine intelligence. *Minds & Machines*, 17(4):391–444, 2007.

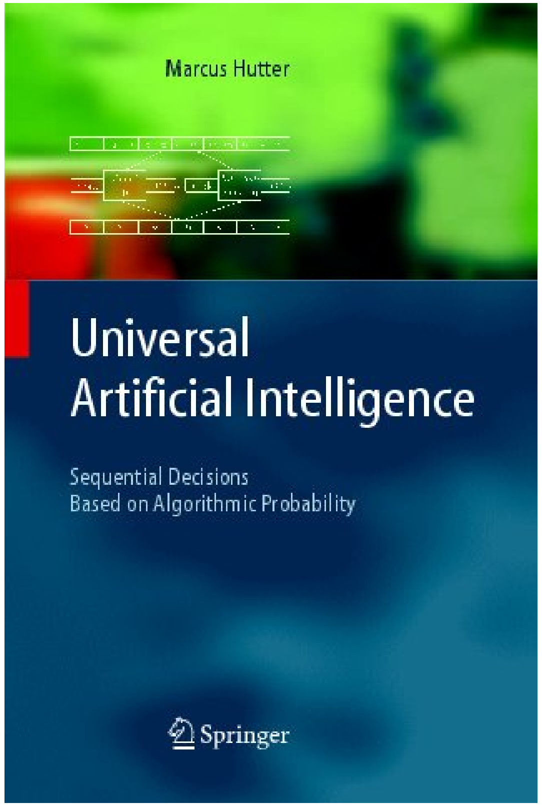
[Hut09] M. Hutter. Feature reinforcement learning: Part I: Unstructured MDPs. *Journal of Artificial General Intelligence*, 1:3–24, 2009.

[VNHS09] J. Veness, K. S. Ng, M. Hutter, and D. Silver. A Monte Carlo AIXI approximation. *Technical report, NICTA, Australia, 2009.*

Thanks! Questions? Details:

Jobs: PostDoc and PhD positions at RSISE and NICTA, Australia

Projects at http://www.hutter1.net/



A Unified View of Artificial Intelligence

= =

Decision Theory = Probability + Utility Theory

\+ +

Universal Induction = Ockham + Bayes + Turing

Open research problems at www.hutter1.net/ai/uaibook.htm

Compression contest with 50'000€ prize at prize.hutter1.net