Learning Agents with Evolving Hypothesis Classes

Peter Sunehag (with Marcus Hutter)

THE AUSTRALIAN NATIONAL UNIVERSITY

2013

Introduction

Universal learning is incomputable. More generally, working with really rich hypothesis classes is costly. This talk is based on two observations about science.

1. Science is only considering a few explicit hypothesis at a time. New hypothesis are generated as old ones are discarded
2. Science has developed by learning laws of limited applicability. The laws are combined to form complete hypothesis about the world

Limited Explicit Hypothesis Classes

- We consider an initial class $\mathcal{M}_0$ of finitely many (reinforcement learning) environments and a hypothesis generating process that over time adds new environments exhausting a countable class in the limit, and an exclusion principle
- $\mathcal{M}_t$ at time t
- Bayesian agent, optimistic agent
- The optimistic agent is simple to define and analyze for this case
- The optimistic agent is more explorative than a Bayesian agent
- We instead view the Bayesian beliefs as being part of the hypothesis generation

The RL Agent

- Consider a class $\mathcal{M}$ of finitely many environments
- We have previously introduced an analyzed an optimistic agent that finds the pair of policy π and environment $\nu \in \mathcal{M}$ that promises the highest reward. The agent then follows π until contradiction
- The number of ε-errors is bounded by $\frac{|\mathcal{M}|}{1-\gamma} \log \frac{1}{\varepsilon(1-\gamma)}$.
- The agent is defined similarly with growing class of environments where we must, however, also switch policy if a newly introduced environment promises more reward
- Assuming the truth is eventually introduced, there is a constant C such that if at time t, we have introduced N_t environments (for all t), then the number of ε-errors is bounded by $C + \frac{N_t}{1-\gamma} \log \frac{1}{\varepsilon(1-\gamma)}$

Combining deterministic laws

- Instead of the hypotheses being complete environments, its more efficient to use a class of laws that makes partial predictions under some circumstances and combines into a huge class of environments
  - Newton's three laws of motion and the law of universal gravitation forms Newton's mechanical universe
  - Contradiction of a law is a contradiction of a lot of environments
  - $|\mathcal{M}|$ is replaced by $|\mathcal{T}|$ in the error bound, i.e. the number of laws instead of the number of environments which can now be uncountably infinite. Extension to growing classes as before

Semi-determinism: Deterministic Laws and Correlations

Class of laws making partial deterministic predictions and separately learnt correlations between the entries in a feature vector (i.e. within a time slice)

- Example: Getting married at time t means you are married at time $t + 1$. 43% of married people are very happy
- However, this does not mean that anyone who gets married have a 43% probability of happiness

- We demand that one predicts ALL that one can with deterministic laws and conditioning on ALL of the predicted features when assigning probabilities for the rest
- Optimistic agents have the same error bounds as before. We only check for contradiction with the deterministic laws

Generating New Hypothesis

Properties of ideal hypothesis generation

1. Simpler hypothesis are more likely
2. Hypothesis that align well with the observed data are more likely

- This can be formalized using algorithmic information theory
- If this process can generate any computable hypothesis we still have a form of universal model
- In reality simplicity is conditional on what has been generated before (conditional Kolmogorov complexity)
- The negation of an existing hypothesis is simple, though still unlikely if the original hypothesis align well with data
- Unclear which algorithm would provide this process for AI. ILP, Logical Probability (Demski last year)?

As a model of science

Our framework with implicit beliefs over a universal class and a small explicit class of hypothesis is a slightly less idealized model of science than universal Bayes

- Either in a passive inference setting, or with knowledge-gain reward (Orseau) or society-gain (e.g. economical)
- Unlike in universal Bayes one can discuss problems like “New Theories and Old Evidence” or the claims that science is irrational due to unconceived alternatives
- In our framework, the old evidence has been used to update the implicit beliefs so if data aligns well with a new hypothesis and its simple it deserves to be introduced as a

Conclusions

- Scaling down universal learners by working with small changing hypothesis classes
- New hypothesis should be generated by a process which is implicitly sampling from a universal distribution
- Sample partial laws that can be combined into full environments, use correlations for other features
- More appropriate as a (still idealized) model of science than pure Bayesian inference