A Multiple Representation Approach To Learning Dynamical Systems

Thomas J. Walsh and Michael L. Littman
RL^3 Laboratory
Department of Computer Science
Rutgers University
Piscataway, NJ USA
{thomaswa,mlittman}@cs.rutgers.edu

Abstract

We examine the problem of learning a model of a deterministic dynamical systems from experience. A handful of representation schemes have been proposed for capturing such systems, including POMDPs, PSRs, EPSRs, Diversity, and PSTs. We argue that no single representation should be expected to be ideal in all situations and describe an approach for learning the most succinct representation of an unknown dynamical system.

Introduction

Problem solving in the AI community has, since its inception, been limited by the difficulty of choosing the "best" individual representation for each environment. In the search or planning setting, an agent given one representation of a task can construct other representations exactly from it while only incurring computational overhead. In the learning setting, where an agent acquires experience of its environment and attempts to refine its model to the point of making perfect predictions, representational problems present even more challenges. An agent cannot swap one representation for another without learning the environment, but it can't learn a representation tractably if its representation scheme is a poor fit (too large). Our goal in this work is to exploit variation between multiple dynamical system representations to learn small models with few mistakes.

Dynamical Systems and Representations

A dynamical system models interaction with an environment over a period of time. We focus in this work on dynamical systems that have a finite set of observations, O ($|O| = k$), and can be influenced by a finite set of actions, A ($|A| = m$). That is, at each timestep $t \geq 0$, the system receives an input $a_t \in A$ and produces some output $o_t \in O$. A *history* is a sequence of zero or more action–observation pairs, $h \in (A \times O)^*$. A deterministic dynamical system f maps histories and actions to next observations. Note that f has in infinite domain, so an explicit representation of the function is infinitely large. For computational tractability, we are interested in finite representation schemes.

For this class of dynamical systems, a number of representations have been proposed in the literature. Several prominent such representations are described briefly in the remainder of the section before we proceed to discussing how choosing between them can affect the efficiency of learning.

POMDPs

The partially observable Markov decision process (POMDP) representation (White & Scherer 1994) is state based. In the deterministic setting, it starts from an initial state and then each action maps the system into a new state, which determines the next observation. It is the most popular, and arguably most natural, finite representation for dynamical systems.

POMDPs are useful for representing both stochastic and deterministic environments. In the deterministic case, an n-state POMDP can be represented with $O(n \log k + nm \log n)$ bits to capture the state transitions and observations.

PSRs

The predictive state representation (PSR) of dynamical systems (Littman, Sutton, & Singh 2002) maintains predictions for a fixed set of tests. In the original (linear) setting, each test is a sequence of actions and observations. For each test in its set, the representation constantly tracks the probability of whether the precise sequence of actions in the test would be obtained *if* the sequence of actions in the test were executed. These updates are made by linear functions and normalizations.

In the deterministic case, predictions are always 0 or 1. So, a set of n tests can capture at most 2^n distinct histories/futures. Since updates in PSRs are made via linear functions, however, it is not clear how to bound the number of PSRs of a certain size nor the size in bits of an n-test PSR.

EPSRs

e-test PSRs (Rudary & Singh 2004), or EPSRs, are a variant of PSRs in which the tests consist of action sequences and a single observation, as opposed to a PSR, which tests an observation after each action. The predictions capture whether that observation will be made if the sequence of observations is carried out.

Diversity

The Diversity representation (Rivest & Schapire 1989) can be viewed as a variant of EPSRs in which the update matrices are captured by graphs. That is, after each action, each test is updated by copying the prediction from some other test. Like POMDPs, an n-test Diversity representation can be written with $O(n \log k + nm \log n)$ bits.

PSTs

Whereas PSRs, EPSRs, and Diversity are grounded in the future, and POMDPs in the current state, looping suffix trees ("predictive" suffix trees) or PSTs make predictions based on past history (Holmes & Isbell 2006). That is, each branch of the suffix tree maps a history sequence and action to the next observation. The inclusion of the possibility of back edges, or loops, in the suffix tree allows PSTs to condition observation choices on actions or observations arbitrarily in the past. A PST with n internal nodes and l leaves can be represented with $O(nmk \log(n + l) + lm \log k)$ bits.

Relationships Between Representations

Let D be the set of deterministic dynamical systems that can be represented by a finite POMDP. We note, without proof, that D is also precisely the set of deterministic dynamical systems that can be represented finitely by EPSRs, by PSRs, by Diversity, and by PSTs. The fact that these definitions all coincide is not a trivial fact. Indeed, the same is not true for non-looping suffix trees (Holmes & Isbell 2006) or k-order Markov models (Littman, Sutton, & Singh 2002), which represent finitely a proper subset of D.

Although the representations can capture the same environments finitely, they vary in the number of bits they need to express them. For example, there are families of dynamical systems that require exponentially larger Diversity representations than POMDPs and vice versa (Rivest & Schapire 1989). However, these two representations are never more than exponentially far apart in size (Rivest & Schapire 1989). Similar minimal-size comparisons for EPSRs, POMDPs, PSRs and Diversity are also known (Rudary & Singh 2004), and show that there are settings where each of these representations may provide an exponentially smaller model than the others.

No relative size bounds for PSTs have been published. We have derived several relevant bounds. We can show a double exponential upper bound on the size of a PST compared to its size as a POMDP. We have also identified a dynamical system consisting of relatively prime loops that requires an exponentially larger representation as a PST compared to a POMDP. We conjecture that the reverse can also be accomplished.

Enumerative Search

We now consider a primitive learning algorithm that exploits the lack of a general "best" representation by learning multiple representations of a given dynamical system. Our algorithm is a search procedure that enumerates representations in order of their size. That is, it considers all representations of size B, and if and only if all these models are invalidated through experience, moves on to representations of size $B+1$. The "pure" representations (POMDPs, Diversity, and PSTs) are the best matches for this approach. PSRs and EPSRs can be enumerated, but because their update functions can include arbitrary numbers, some additional results are needed to ensure a scheme that is sound and complete.

The approach is reminiscent of Levin search (Wiering & Schmidhuber 1996) in that it considers representations in size order. It also accomplishes the stated goal of learning with a small number of mistakes when compared to the same basic search strategy restricted to a single representation. In the worst case, the multiple representations strategy learns the correct model while making a number of prediction mistakes that is exponential in the size of the *smallest* representation, which we have seen in the last section may be exponentially smaller than some a priori chosen representation.

Conclusion

Our exploration of dynamical system learning with multiple representations is still in its preliminary stages, but the facts reported above evince clear advantages over learning with an a priori fixed representation. We are currently endeavoring to fill in several of the "gaps" in the size comparisons concerning PSTs discussed above, and considering how to integrate PSR and EPSR representations into the ensemble. Overall, the approach seeks to leverage the parallel updating of models into an exponential speedup in the learning process and the corresponding exponential reduction in mistakes made during learning.

References

Holmes, M., and Isbell, C. 2006. Looping suffix tree-based inference of partially observable hidden state. In *Proceedings of the Twenty-Third International Conference on Machine Learning (ICML)*.

Littman, M. L.; Sutton, R. S.; and Singh, S. 2002. Predictive representations of state. In *Advances in Neural Information Processing Systems 14*, 1555–1561.

Rivest, R. L., and Schapire, R. E. 1989. Inference of finite automata using homing sequences. In *Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing*, 411–420.

Rudary, M. R., and Singh, S. 2004. A nonlinear predictive state representation. In *Advances in Neural Information Processing Systems 16 (NIPS)*, 855–862.

White, III, C. C., and Scherer, W. T. 1994. Finite-memory suboptimal design for partially observed Markov decision processes. *Operations Research* 42(3):439–455.

Wiering, M., and Schmidhuber, J. 1996. Solving POMDPs with Levin search and EIRA. In *Machine Learning: Proceedings of 13th International Conference*, 534–542.