Information Theoretic Approaches for Predictive Models

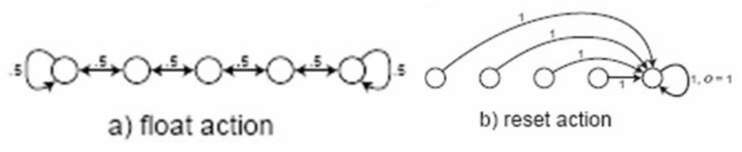
Results and Analysis

Monica Dinculescu   Supervised by Doina Precup

Undergraduate Summer Research Symposium
August 31, 2006

Problem Definition

- Partially observable system, discrete time
- Hidden states: you can't see a state, but you can see an observation



[Littman, Sutton and Singh, 2001]

Problem Definition (cont'd)

Definitions

$x_t^{past} \in X$: **histories** (pasts)

- finite length sequences of action:observation pairs
  $x_t^{past} = [a_{t-k} : o_{t-k}, ..., a_{t-1} : o_{t-1}]$

$y_t^{fut} \in Y$: **future observations**

- Want to be able to predict what the outcome of an action will be: i.e. predict $p(y_t^{fut} | a_t, x_t^{past})$

Solution I

Partially Observable Markov Decision Processes

- POMDP [Sondik, 1971]

Definition

A **belief state** is used to keep track of the probabilities of being in each of the hidden states. [Sondik, 1971]

Problem

- Computationally expensive
- Depends on a good model of underlying states

Solution II

Predictive State Representations

- PSR [Littman, Sutton and Singh, 2001]
- Constructs a state representation based only on actions and observations

Definitions

A **test** is a sequence of **future** action:observation pairs

- $q = [a_t : o_t, ..., a_{t+l} : o_{t+l}]$

A **history** is a sequence of **past** action:observation pairs

- $h = x_t^{past} = [a_{t-k} : o_{t-k}, ..., a_{t-1} : o_{t-1}]$

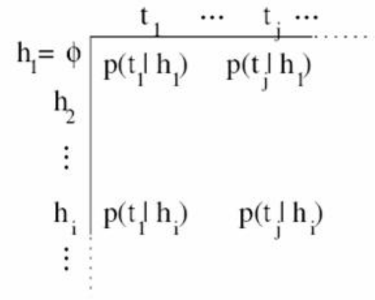
- PSR representation predicts $p(q|h)$

Solution II (cont'd)

Predictive State Representations

Definition

- A **System Dynamics Matrix** is an ordering over all possible tests and histories.[Singh, James et al, 2004]
- Infinite matrix but with a finite number of linearly independent columns called **core tests.**



[Singh, James et al, 2004]

Problem

- Less restrictive model but very data expensive
- No good learning algorithms

Motivation

- Flexible model based on finite length histories
- Data efficient learning algorithm
- Computation/memory affordable
- Good predictions

Information Theoretical Approach

- Based on the Active Learning algorithm developed by S.Still and W.Bialek, 2004

Definition

We define an internal representation $s_t \in S$ such that:

1. There is a **lossy compression** of the information from x_t^{past}
2. It has a good **predictive power**

Optimization Principle

Definition

$$F = max_{p(s_t|x_t^{past})}[I(\{s_t, a_t\}, y_t^{fut}) - \lambda I(s_t, x_t^{past})]$$

- First term: maximize predictive information about the future
- Second term: compress information about the past
- λ is a constant that trades 1) and 2) off

Solution

Theorem

The $s_t \leftarrow x_t^{past}$ assignment is:

$$p(s_t|x_t^{past}) \sim exp(\frac{-1}{\lambda} \sum_a p(a_t|x_t^{past}) \cdot D_{KL}[p(y_t^{fut}|a_t, x_t^{past})||p(y_t^{fut}|a_t, s_t)]$$

- The D_{KL} compares how different the future prediction as given by the state is compared to the future prediction as given by the entire history
- A better state assignment $\longrightarrow$a prediction more similar to the one returned by the history
- λ acts as a temperature parameter; as $\lambda \longrightarrow 0$, the $s_t \leftarrow x_t^{past}$ assignment becomes deterministic

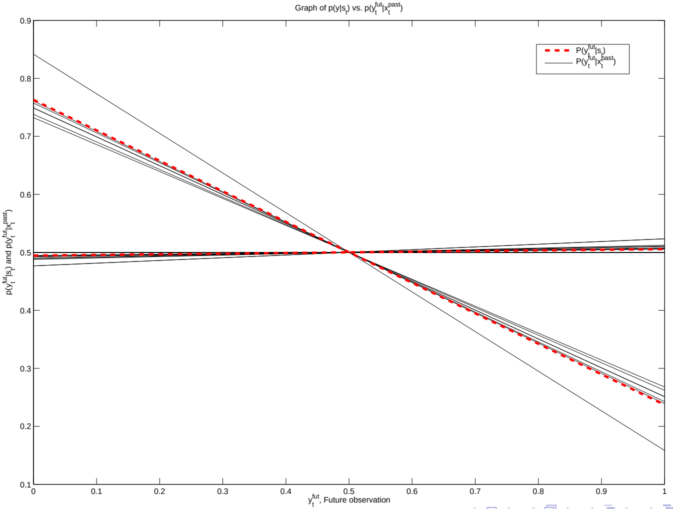
Algorithm

Input: # states s_t , length of x_t^{past}, length of initial trajectory, λ
Output: $p(s_t|x_t^{past})$

1. create an initial trajectory by taking random actions
2. estimate $p(x_t^{past})$ and $p(y_t^{fut}|x_t^{past}, a_t)$
3. **for** i←1 **to** t
4. **while** $p(s_t|x_t^{past})$ does not converge
5. iteratively solve :
$$p^{(j+1)}(y_t^{fut}|a_t,s_t) \sim \frac{\sum_h p(y_t^{fut}|a_t,x_t^{past})p(a_t|x_t^{past})p^{(j)}(s_t|x_t^{past})p(x_t^{past})}{\sum_h p(a_t|x_t^{past})p^{(j)}(s_t|x_t^{past})p(x_t^{past})}$$
$$p^{(j)}(s_t|x_t^{past}) \sim exp(\frac{-1}{\lambda}\sum_a p(a_t|x_t^{past}) \cdot D_{KL}[p(y_t^{fut}|a_t,x_t^{past})||p^{(j)}(y_t^{fut}|a_t,s_t)]$$
6. take a random action a_t and update $p(x_t^{past})$ and $p(y_t^{fut}|x_t^{past}, a_t)$

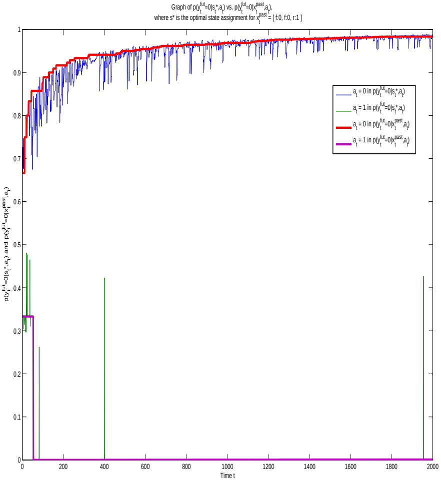
Lossy compression of available pasts

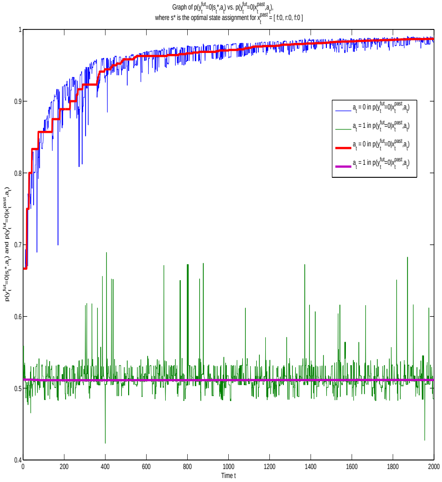
The internal representation is a sufficient statistic of the system



Good Predictive States

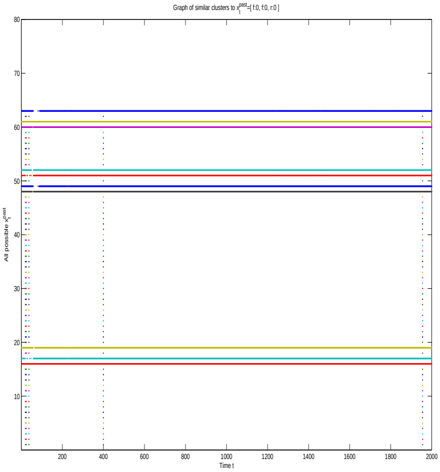
The internal representation retains good predictive powers

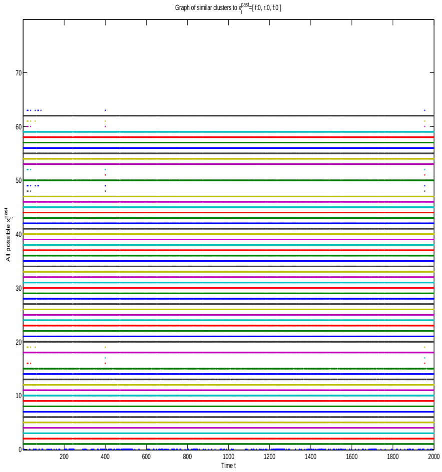




Consistent clustering

Consistent $s_t \leftarrow x_t^{past}$ assignment





Conclusions and Future Work

- The algorithm learns a predictive model with a limited amount of data
- Predictions are consistent

Future Work:

- Compare predictive model with PSRs
- Learn optimal action policies