Selecting concise sets of samples for a reinforcement learning agent.

Damien Ernst

Department of Electrical Engineering and Computer Science
University of Liège

CIRAS - Singapore - December 14-16, 2005

.

Find slides: http://montefiore.ulg.ac.be/∼ernst/

Reinforcement learning and generalization

What is reinforcement learning: Reinforcement learning is learning what to do, how to map states to actions, from the information acquired from interaction with a system.

An open problem in reinforcement learning: Generalization of the information

An extremely promising approach to this problem: To solve a sequence of batch mode supervised learning problem (algorithm known as fitted Q iteration).

Problem with the fitted Q iteration algorithm

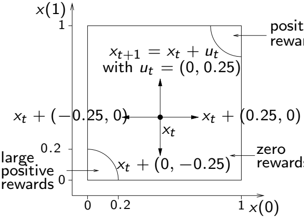
Training sets for the different supervised learning problems contain a number of elements equal to the number of samples the reinforcement learning agent has acquired from interaction with the system.
After a certain time of interaction, these samples may become so numerous that this framework may become computationally impractical.
To reduce the computational burdens, we propose to select a concise set of sufficiently rich representatives of the samples.

Concise set selection: our proposed algorithm

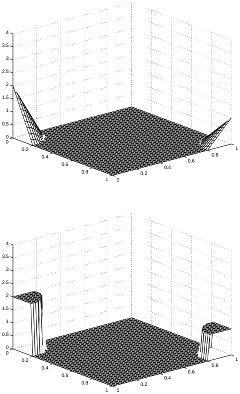
Main charactistics of our algorithm: the algorithm works iteratively by associating to the solution computed from the already selected samples an error function and by selecting the sample for which this error function is the largest.

1. we select a first sample and we add it to the empty set $\mathcal{F}$
2. we run the fitted Q iteration algorithm on $\mathcal{F}$
3. we compute an error function
4. we select from the remaining samples the one that leads to the largest value of the error function
5. we add this sample to $\mathcal{F}$ and go back to step 2 if size of $\mathcal{F}$ is not equal to size concise set targeted

Simulation results



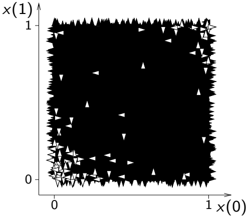
System dynamics



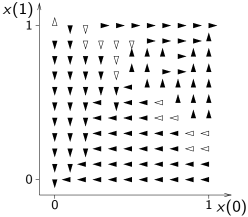
Reward functions

Figure: Two test problems having the same dynamics but different reward functions

Continuous case



Triangles represent the samples.
White triangle = concise set
White + Black triangle =
original set



Triangle represent the policy.
White triangles =
place where the policy is
not optimal when considering the
concise set as input
of the fitted Q iteration algorithm

	size set	suboptimality
Concise sets	50	0.599
	100	0.399
	200	0.399
	500	0.199
	1000	0.199
Random sets (average values)	50	1.939
	100	1.823
	200	1.625
	500	1.330
	1000	0.903
	2000	0.679
The whole set	10000	0.199

Table: Performances of different sets

Discontinuous case

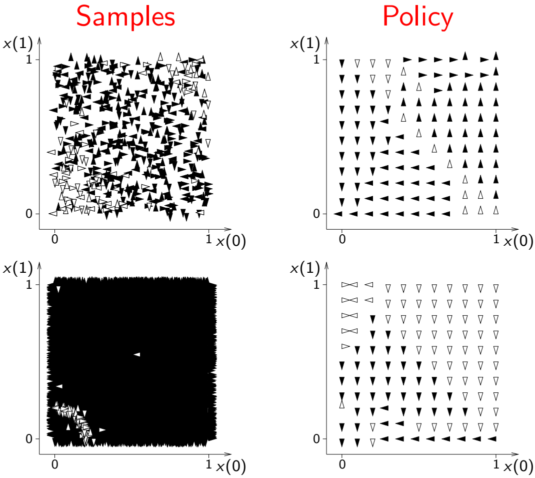


Figure: Original set has either 500 (top) or 100,000 samples (bottom)

Conclusions

We have proposed an algorithm to select concise sets of samples for a reinforcement learning agent.
Main motivation for this work: to lighten the computational burdens required to run batch-mode reinforcement learning algorithms on the whole set of samples
The results were mitigated. Two main problems:

1. does not work for discontinuous systems since too many samples selected alongside the discontinuities.
2. identification of the concise set may take more time than running fitted Q iteration on the whole set of samples.