Past Research

Chandrashekar L

Thesis Title:"Approximate Dynamic Programming and Reinforcement Learning - Algorithms, Analysis and an Application"
Advisor: Prof. Shalabh Bhatnagar (shalabh@csa.iisc.ernet.in), Department of Computer Science and Automation, Indian Institute of Science, Bangalore - 560012

Abstract: MDP is a useful mathematical framework to cast a variety of optimal sequential decision making problems under uncertainty in domains such as engineering, science and economics. However, computing optimal value function and optimal policy is difficult in practice because either the state space is too large or the model information is not available.

The primary investigations in the thesis were:

- *Approximate Dynamic Programming* refers to a gamut of methods that compute an approximate value function and a sub-optimal policy. The thesis investigated a widely used ADP method namely the Approximate Linear Programming (ALP) formulation. In particular, analytical tools were developed to bound the performance degradation that occurs when the constraint of the ALP are reduced or approximated. The analysis is based on ideas of monotone projections in tropical linear algebra.
- *Reinforcement Learning* algorithms are stochastic approximation (SA) schemes and solve the MDP by making use of sample trajectories. Actor-Critic algorithms are two timescale SA schemes since they make use of different step-size schedules. The thesis investigated the conditions under which two timescale SA schemes are stable and convergent.
- *Crowd Sourcing* is a new mode of organizing work in multiple groups of smaller chunks of tasks and outsourcing them to a distributed and large group of people in the form of an open call. An important task attribute that affects the completion time of a task is its price, and incorrect pricing leads to task starvation. In the thesis, the pricing problem is formulated in the MDP framework to compute a pricing policy that achieves predictable completion times in simulations as well as real world experiments.

I. Publications (from the thesis)

A. Journal

- Chandrashekar, L. and Bhatnagar, S. "A Stability Criterion for Two Timescale Stochastic Approximation Schemes", Accepted for publication in Automatica, 2017.
- Chandrashekar, L.; Bhatnagar, S and Szepesvári C., "A Generalized Reduced Linear Program for Markov Decision Processes" (under Review in IEEE Transactions on Automatic Control).

B. International Conferences

- Chandrashekar, L.; Bhatnagar, S., "Approximate Dynamic Programming with $(\min, +)$ linear function approximation for Markov Decision Processes," 53rd IEEE Annual Conference on Decision and Control (CDC), December $15 - 17$, 2014, Los Angeles California, USA.
- Chandrashekar, L.; Bhatnagar, S., "A Generalized Reduced Linear Program for Markov Decision Processes," Twenty-Ninth AAAI Conference on Artificial Intelligence, January $25 - 30$, 2015, Austin Texas, USA.
- Chandrashekar, L.; Dubey, A.; Bhatnagar, S. and Chithralekha, B., "A Markov Decision Process framework for predictable job completion times on crowdsourcing platforms", Proceedings of the

Second AAAI Conference on Human Computation and Crowdsourcing, HCOMP 2014, Pittsburgh, Nov. 2-4, 2014.

- Maity, R. K.; Chandrashekar, L.; Padakandla, S.; Bhatnagar, S., " Shaping Proto-Value Functions Using Rewards", European Conference on Artificial Intelligence 2016.