Chandrashekar Lakshminarayanan

Mobile: +1(587)-778-2035
E-mail: chandrurec5@gmail.com

Short Summary

My research work deals with developing Approximate Dynamic Programming (ADP) and Reinforcement Learning (RL) algorithms with provable performance and guaranteed convergence. I am also interested in applications of RL and stochastic optimization to problems arising in practice. Currently, I am a post-doctoral research fellow in the Department of Computing Science, University of Alberta, and I obtained my PhD from the Department of Computer Science and Automation, Indian Institute of Science.

Interests

Markov Decision Processes, Approximate Dynamic Programming, Reinforcement Learning, Bandits, Machine Learning, Learning Theory, Stochastic Optimization, Stochastic Approximation and Stochastic Control.

Relevant Courses

- Stochastic Models and Applications, Stochastic Processes and Queuing Theory, Stochastic Approximation Algorithms
- Convex Optimization, Game Theory, Dynamics of Linear Systems
- Data mining, Pattern Recognition, Detection and Estimation Theory
- Real Analysis, Measure Theory, Topology (point-set), Linear Algebra

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.

Publications

- Chandrashekar, L. and Bhatnagar, S. “A Stability Criterion for Two Timescale Stochastic Approximation Schemes”, Accepted for publication in Automatica, 2017.
- 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.

Work under consideration

- Chandrashekar, L.; Bhatnagar, S and Szepesvári C., “A Generalized Reduced Linear Program for Markov Decision Processes”.

Work in Progress

- “A stochastic gradient approach to parameter tuning in Hadoop”.
- “Robust Temporal Difference Learning Algorithms”.

Patents

- Methods and Systems for Crowdsourcing of Tasks (US 20160071048 A1).
- System and method for improving dynamic performance of a circuit (US 7538701 B2).
- System and method for reducing power dissipation in an analog to digital converter (US 7821436 B2).

Industry Experience

Analog Design Engineer, Cosmic Circuits Pvt Ltd, Bangalore, (2005 – 2008).
High Speed Analog to Digital Converters (ADC) Team.

- Design of blocks such as
  - Amplifiers with state-of-the-art feedback compensation schemes.
  - Comparators.
  - Digital Error Correction Logic.
- Full system simulation, testing and debugging of 10 to 12 bit ADCs.
- Worked in CMOS technologies varying from 0.360μm to 65nm.

Computing Skills Familiar with C, C++, Matlab/Octave, Linux/Windows.

Teaching Assistantship

- Linear Algebra and Applications, Aug-Dec, 2012- 2014.
- Foundation of Data Sciences, Jan-Apr 2014.

Talks and Posters

- *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.
- *A Generalized Reduced Linear Program for Markov Decision Processes*, Twenty-Ninth AAAI Conference on Artificial Intelligence, January $25-30$, 2015, Austin Texas, USA.
- *A Generalized Reduced Linear Program for Markov Decision Processes*, 3^{rd} IKDD Conference on Data Science, March $13-16$, 2016, Pune, India.
- *Introduction to Reinforcement Learning*, I-CARE Big Data Analytics and Cognitive Computing winter school, Bangalore 2014.
- *Introduction to Probability*, Computer Science and Automation Undergraduate Summer School, Bangalore 2013.
- *Introduction to Linear Algebra*, Computer Science and Automation Undergraduate Summer School, Bangalore 2014.

Academics

Ph.D.[2010-2015]
Department of Computer Science and Automation (CSA),
Indian Institute of Science (IISc), Bangalore, India.

M.E. [2008 - 2010]
Systems Science and Automation,
Department of Electrical Engineering (EE),
Indian Institute of Science (IISc), Bangalore, India.
C.G.P.A: 7.2/8

B.Tech. [2001 - 2005]
Instrumentation and Control Engineering,
National Institute of Technology, Tiruchirapalli.
C.G.P.A: 8.61/10

Accomplishments

1. I was All India Rank 1 in the Graduate Aptitude Test in Engineering conducted in the year 2008 with a score of 1000/1000.
2. Presented with Medal of Honor for Best Pass out student in Systems Science and Automation, Indian Institute of Science.
3. Best research presentation award in the IIST-Research Scholars Day held at Indian Institute of Space Technology, Trivandrum, December 2013.
4. Best research presentation award in the Electrical Engineering and Computer Science Colloquium held at IISc, Bangalore, February 2015.

References

Available on request.

National Institute of Technology

TIRUCHIRAPPALLI - 620 015 INDIA

hereby confers the degree of

Bachelor of Technology

in

INSTRUMENTATION AND CONTROL ENGINEERING

on

CHANDRASHEKAR L (IC10107)

for successfully completing the prescribed programme of study
and having been placed in

First Class with Distinction.

Given this day the Twentythird of July 2005
under the seal of the Institute.

Registrar	*Director*	*Chairman,* *Board of Governors*

Indian Institute of Science Bangalore.

Faculty of Engineering

The Council on the recommendation of the Senate of the Institute hereby confers on

Chandrashekar L

the degree of

Master of Engineering

(First Class with Distinction)

in

SYSTEMS SCIENCE AND AUTOMATION

Chairman of the Council

Veeranna Kammar
Assistant Registrar (Academic)

INDIAN INSTITUTE OF SCIENCE
BANGALORE 560 012, INDIA
Telephone: 091-80-22932706/2333/2233/2937
Telefax: 091-80-23600861/23601413
E mail: thesisengg@academic.admin.iisc.ernet.in

R(II)005/CSA/6254

14th September, 2016

CERTIFICATE

This is to certify that **Dr.Chandrashekar L, SR No:4716-120-101-07396, Doctor of Philosophy (Ph D) Student** in the **Department of Computer Science and Automation** under **Faculty of Engineering** at the **Indian Institute of Science,** has been awarded Ph D Degree by the Council at its meeting held on **25th June 2016.**

The title of his Ph D thesis is: **"Approximate Dynamic Programming and Reinforcement Learning-Algorithms, Analysis and an Application"**

14/9/16

Assistant Registrar
Academic

INDIAN INSTITUTE OF SCIENCE
BANGALORE

Teaching Plan

Chandrashekar L

I. Background

I did my Bachelors in Instrumentation and Control Engineering at the National Institute of Technology. After my Bachelors, I worked for a period of 3 years in Cosmic Circuits Private Limited as an Analog Design Engineer. I was involved in designing various blocks such as multi-stage amplifiers with feedback compensation, comparators and digital error correction logic circuits in cmos technologies ranging from 360nm to 65nm. I pursued my Masters in Systems Sciences and Automation at the Indian Institute of Science, Bangalore. I also continued my PhD at the Indian Institute of Science specializing in Reinforcement Learning and Approximate Dynamic Programming.

II. Experience

- **Teaching Assistantship:** At the Indian Institute of Science, I was a TA for courses such as Linear Algebra and Applications (Aug-Dec 2012, 2013, 2014) and Foundations of Data Sciences (Jan-May 2014). My duties as a TA involved conducting tutorials, setting quizzes and evaluating scripts.
- **Summer School:** I was a member of the organizing committee for the Undergraduate Summer School [1] conducted by the Department of CSA, IISc. The main aim of the summer school was to introduce undergraduate students to cutting-edge research in computer science via talks, demos, and hands-on sessions. I have given expository talks on linear algebra [2] and probability at the UG summer school. As a member of the organizing committe I was involved in the identification of interesting topics/speakers and selection of candidates.
- **Research Presentations:** I have presented my research on internal forums such as the 53^{rd} IEEE Conference of Decision and Control held at Los Angeles, California USA, 2014 [8] and the 29^{th} AAAI conference held at Austin, Texas USA, 2015 [6]. I have also received awards for *best research presentation* from Indian Institute of Space Technology, Trivandrum and Indian Institute of Science, Bangalore.
- **Project Guidance:** I have guided masters students of the Stochastic Systems Lab (SSL) at Dept. of CSA, IISc. My work in crowdsourcing [7] was joint effort with Masters students of the SSL.

III. Teaching Goals

My course, work experiece and research training enables me to offer a variety of courses related to Electrical Engineering and Computer Science at the undergraduate level such as courses on Data Structures and Algorithms [9], Probability [20], Linear Algebra [11], Engineering Mathematics [14], Digital Design [15], Microprocessors [12], Network/Circuit Theory [13], Analog Electronics [21], Linear Integrated Circuits [10], Signals and Systems [17], Classical and Modern Control Theory [16]. I also feel that courses in algorithms, circuits, mircro-processors need to be supplemented by carefully designed *laboratory* sessions and *mini-projects* so that the students can learn to apply theory in practice. I am also interested in designing advanced undergraduate courses such as ‘Introduction to Machine Learning’, where, in addition to the examinations, students will be encouraged to read and present research *papers*.

At the graduate level, I wish to offer courses on Stochastic Control [3], Reinforcement Learning [22], Stochastic Approximation [4], Queuing Theory [19], Optimization [5], Detection and Estimation Theory [18].

REFERENCES

[1] CSA Undergraduate Summer School, http://events.csa.iisc.ernet.in/summerschool2014/.

[2] Introduction to Linear Algebra, https://www.youtube.com/watch?v=qr-icg9wrz8.

[3] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume II. Athena Scientific, Belmont, MA, 4th edition, 2013.

[4] V. S. Borkar. *Stochastic Approximation: A Dynamical Systems Viewpoint*. TRIM, 2008.

[5] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[6] L. Chandashekar and S. Bhatnagar. A generalized reduced linear program for markov decision processes. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA.*, pages 2722–2728, 2015.

[7] L. Chandrashekar, A.Dubey, S. Bhatnagar, and B. Chithralekha. A markov decision process framework for predictable job completion times on crowdsourcing platforms. In *Proceedings of the Seconf AAAI Conference on Human Computation and Crowdsourcing, HCOMP 2014, November 2-4, 2014, Pittsburgh, Pennsylvania, USA*, 2014.

[8] L. Chandrashekar and S. Bhatnagar. Approximate dynamic programming with (min; +) linear function approximation for markov decision processes. In *53rd IEEE Conference on Decision and Control, CDC 2014, Los Angeles, CA, USA, December 15-17, 2014*, pages 1588–1593, 2014.

[9] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to algorithms*, volume 6. MIT press Cambridge, 2001.

[10] S. Franco. *Design with operational amplifiers and analog integrated circuits*. McGraw-Hill, 2001.

[11] Gilbert G. S. Strang. *Introduction to Linear Algebra*, volume 3. Wellesley-Cambridge Press Wellesley, MA, 1993.

[12] Ramesh S Gaonkar. *Microprocessor architecture, programming, and applications with the 8085*. Prentice-Hall, Inc., 1995.

[13] W. H. Hayt, J. E. Kemmerly, and S. M. Durbin. *Engineering circuit analysis*. McGraw-Hill New York, 1986.

[14] E. Kreyszig. *Advanced engineering mathematics*. John Wiley & Sons, 1988.

[15] M. M. Mano. *Digital design*. EBSCO Publishing, Inc., 2002.

[16] Katsuhiko Ogata. *Modern control engineering*. Prentice Hall PTR, 2001.

[17] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab. *Signals and systems*. Pearson, 2014.

[18] H. V. Poor. *An introduction to signal detection and estimation*. Springer Science & Business Media, 2013.

[19] S. M. Ross. *Stochastic processes*, volume 2. John Wiley & Sons New York, 1996.

[20] Sheldon M Ross. *Introduction to probability models*. Academic press, 2014.

[21] A. S. Sedra and K. C. Smith. *Microelectronic circuits*. Holt, Rinehart and Winston, 1982.

[22] R. S. Sutton and A. G. Barto. *Introduction to Reinforcement Learning*. MIT Press, Cambridge, MA, USA, 1st edition, 1998.

Research Proposal

Real world is often uncertain, staring from the time taken one waits for a bus to the amount of rainfall in a given year. Planning in the face of such uncertainty is hence crucial if we are interested in designing practical systems with good performance. The framework of stochastic control serves as a useful mathematical tool to model and solve problems of planning under uncertainty. Some common instances of stochastic control problems are

- Networking: Control of Network of Queues, Traffic Control.
- Business: Inventory Control.
- Artificial Intelligence: Terrain Exploration by an autonomous Robot
- Energy: Optimal energy harvesting in Sensor Networks, Control of Smart Grids.

In general several dynamic resource management problems can be cast in the framework of stochastic control. Some of the interesting research directions related to theory and practice of stochastic control are as under.

1 Approximate Dynamic Programming

A large number of stochastic control problems occurring in practice belong to the sub-class called Markov Decision Processes (MDPs) [2]. The Markovian assumption in the system dynamics enables us to solve MDPs using the principle of Dynamic Programming (DP). Exact solution methods for MDPs are built on the idea DP are not applicable to MDPs occurring in practice which typically have large number of states.

Approximate Dynamic Programming (ADP) algorithms [6] tackle the issue of large number of states by cleverly combining approximate representations and the DP principle. While several algorithms [2] exist in literature, not all of them have guaranteed performance. An important and fundamental research question is to understand conditions under which ADP algorithms can be guaranteed to perform well in practice.

2 Reinforcement Learning

In most practical scenarios, the model information of the underlying MDP is not available. However, the system is available only in the form of a simulator or samples can be obtained via direct interaction. In such a scenario, we need to *learn* using the samples. Reinforcement Learning (RL) [9] algorithms are sample trajectory based solution methods for MDP. RL algorithms such as Q-learning [11], temporal difference learning [10] have been successful in domains such as Backgammon, elevator control etc. The aim is to develop sample efficient, stable and convergent RL algorithms for real world systems.

3 Stochastic $\max$-Plus Systems

The class of discrete event systems where elementary components interact via 'synchronization' are called Stochastic $\max -$ Plus (SMPL) Systems [4]. Examples of such systems include railway networks, production chain, scheduling, queuing and digital systems. The inherent randomness in the different interacting components leads to the stochastic behavior. Developing approximate algorithms to derive control strategies for SMPL systems is of interest.

4 Bandit Algorithms

The classical multi-arm-bandit problem [1] is an example of trade off between exploration (trying out the arms to find the best arm) and exploitation (playing the arm currently known to be the best). Several variants of the problems include linear bandits, convex bandits, contextual bandits, bandits with complex or partial feedback have been considered in literature to model various scenarios. Exploring newer variants of the bandit problems is important.

5 Optimal Resource Allocation Problems

An interesting line of research is to empirically test and tune the performance of the state of the art ADP, RL and bandit based algorithms in various application domains. The following are the practical application domains that are of interest in the immediate future.

5.1 Pricing

Optimal pricing of resources is a common problem in various domains such as crowd-sourcing, smart grids and inventory systems. For instance, in crowd sourcing the price offered for a task affects its completion time [3]. In the case of smart grid [7], optimal pricing is key to make profits in a *Tariff* market setting. Dynamic pricing of items is important in the context of inventory management.

5.2 Sensor Network

Optimal resource allocation in wireless sensor networks has attracted attention in recent times [8, 5]. The nodes harvest energy and are self-powered. The network consists of one centralized controller that collates information from other nodes, and the center also has control over the *sleep-wake* schedules of all the other nodes. The aim here is to come up with a policy that keeps the minimum number of sensors awake per unit time while meeting the communication objective. A novel research direction would be to employ methods that exploit the topology and nature of object movement by computing the Laplacians of the associated graphs.

5.3 Expected Milestones

Quantifiable:
$2 - 3$ papers per year in top tier conferences and 1 paper per year in top tier journals.
Non- Quantifiable:

- Get involved in Research Projects relvant to IIT-Dharwad.
- Collaborate with colleagues and play vital role in the research and teaching activities at IIT-Dharwad.

References

[1] Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2-3):235–256, 2002.

[2] D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume II. Athena Scientific, Belmont, MA, 4th edition, 2013.

[3] S. Faradani, B. Hartmann, and P. G. Ipeirotis. What's the right price? pricing tasks for finishing on time. In *Human Computation*, 2011.

[4] S. S. Farahani. Approximation methods in stochastic max-plus systems. 2012.

[5] Jason A Fuemmeler and Venugopal V Veeravalli. Energy efficient multi-object tracking in sensor networks. *Signal Processing, IEEE Transactions on*, 58(7):3742–3750, 2010.

[6] W. B. Powell. *Approximate Dynamic Programming: Solving the curses of dimensionality*, volume 703. John Wiley & Sons, 2007.

[7] Prashant P Reddy and Manuela M Veloso. Strategy learning for autonomous agents in smart grid markets. 2011.

[8] Vinod Sharma, Utpal Mukherji, Vinay Joseph, and Shrey Gupta. Optimal energy management policies for energy harvesting sensor nodes. *Wireless Communications, IEEE Transactions on*, 9(4):1326–1336, 2010.

[9] R. S. Sutton and A. G. Barto. *Introduction to Reinforcement Learning.* MIT Press, Cambridge, MA, USA, 1st edition, 1998.

[10] John N. Tsitsiklis and Benjamin Van Roy. An analysis of temporal-difference learning with function approximation. 1997.

[11] Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine learning*, 8(3-4):279–292, 1992.