

AITF ANNUAL REPORT 2015

DR. RICHARD SUTTON

REINFORCEMENT LEARNING AND ARTIFICIAL INTELLIGENCE

AITF ANNUAL REPORT MARCH 31, 2015

1. EXECUTIVE SUMMARY

The RLAI research program pursues an approach to artificial intelligence and engineering problems in which they are formulated as large optimal-control problems and approximately solved by reinforcement-learning methods. Reinforcement learning is a new body of theory and techniques for optimal control that has been developed in the last thirty years primarily within the machine learning and operations research communities, and which has separately become important in psychology and neuroscience. Reinforcement learning researchers have developed novel methods to approximate solutions to optimal-control problems that are too large or too ill-defined for classical solution methods such as dynamic programming. For example, reinforcement-learning methods have obtained the best known solutions in such diverse automation applications as helicopter flying, elevator deployment, playing backgammon, and resource-constrained scheduling.

The objectives of the RLAI research program are to create new methods for reinforcement learning that remove some of the limitations on its widespread application and to develop reinforcement learning as a model of intelligence that could approach human abilities. These objectives are pursued through mathematical analysis, through computational experiments, through the development of robotic systems, and through the development and testing of computational models of natural learning processes.

The scope of our research activities increased substantially this year, beginning with a new faculty team member: Assistant Professor Patrick Pilarski of the Department of Medicine at the University of Alberta. The overall research team now consists of about 70 members, 44 of whom are graduate students and, of those, 9 of which are recipients of major scholarships. The output of the research program also increased, with 44 papers published or accepted for publication in refereed archival venues during the reporting period. Two PhD students and three MSc students were graduated.

The primary focus of the research program has been on how intelligent machines represent their knowledge of the world. The key question here is how to organize the knowledge such that it can be verified, learned, and used autonomously without continual tending by human experts. This project has pursued an unusual approach in which knowledge is expressed in terms of the machine's sensors and actuators, thereby enabling it to be compared directly to experiential data. Substantial further progress was made this year toward formalizing the core learning algorithms for doing this.

Highlights of the research program this year include: 1) building a computer system that plays essentially perfectly at one of the most popular forms of Poker, 2) a new algorithmic idea, called *emphasis*, for solving the stability problem of off-policy reinforcement learning, 3) new algorithms for gaining the benefits of *weighted* importance sampling in off-policy reinforcement learning with function approximation, and 4) demonstration of on-line learning in real-time to reduce switching time in smart prosthetic interfaces.

2. RESEARCH PROGRAM OVERVIEW

In the project's proposal, research was divided into three main target areas. The first is extensions of existing reinforcement learning algorithms; there are many open problems in reinforcement learning, and we seek to solve them as opportunities arise. The second area is the extension of reinforcement learning ideas to address the more ambitious goals of artificial intelligence (AI). There is a natural transition from the more advanced reinforcement learning methods to mechanisms for knowledge representation, search, and human-level reasoning. A major goal for the project is to explore, implement, and illustrate these relationships. The third main area of RLAI research is a focus on applications—on designing algorithms and software that are well suited for applied research, and on several specific applications.

The most important development in our research program this year was the budding off of a new research laboratory in the Department of Medicine headed by new faculty team member Patrick Pilarski. Pilarski was previously a postdoctoral fellow in the RLAI project leading our research into adaptive prosthetics in collaboration with the Glenrose Rehabilitation Hospital and others at the University of Alberta. This year he was hired as an assistant professor into the Division of Physical Medicine and Rehabilitation, where he has expanded his work into a 800-square-foot laboratory of his own with multiple 3D printers and robot platforms including one of only two copies of the world's most advanced prosthetic limb systems (the Modular Prosthetic Limb v3.0, CAD\$700,000). Pilarski already supervises four graduate and three undergraduate students.

Our main research focus this year has been on core reinforcement learning algorithms, in particular, on the most classical of all reinforcement-learning algorithms, known as $TD(\lambda)$, which was introduced by Sutton in 1988 and forms the foundation for most modern reinforcement learning algorithms. This year we introduced a new algorithmic idea, called *emphasis*, for extending $TD(\lambda)$ to off-policy learning. Off-policy learning, or learning about one policy from data generated using another, is key to our strategy for scaling reinforcement learning methods to address the more ambitious goals of AI. Our new idea is to vary the strength (emphasis) of the $TD(\lambda)$ update from time step to time step. If some steps are emphasized more than others, in just the right way, then $TD(\lambda)$ can be guaranteed to be stable even with linear function approximation. The only previous methods to achieve this are the gradient-TD methods developed by this project over the last five years. Compared to those methods, the new emphatic $TD(\lambda)$ is substantially simpler, with only one set of weights and one step-size parameter (compared to two of both for gradient-TD methods). Emphatic $TD(\lambda)$ may offer several other advantages over conventional $TD(\lambda)$ which we will be exploring in the coming years.

In the area of applications, this year we obtained a dramatic success in building computer systems that play the popular game of Heads-up Limit Texas Hold'em Poker. We have been working on this game for several years and have dominated most of the international competitions. This year we announced in a *Science* article that we have essentially solved the game—that our program is unbeatable in a statistical sense even over a lifetime of play. This result generated substantial media interest.

3. RESEARCH PROJECTS

This section describes in more detail a few of the research projects making up the research program.

Solving Computer Poker

Poker is representative of a large number of strategic decision problems involving uncertainty and imperfect information. The application of modern AI methods to Poker has become a very active, competitive area of academic research in the last decade, and also one of great interest to the public as the computers have begun to challenge the best human players. RLAI team members, led by faculty team member Michael Bowling, have played a key role in these developments. Bowling's group is one of the original founders in 2006 of the Annual Computer Poker Competition, a common testbed shared by researchers in artificial intelligence and computational game theory. The single most important prior achievement in computer Poker was in 2008 when Polaris, a program developed by Bowling's group, defeated the world's best human players in the game of 2-player limit Texas hold'em Poker at the Man-Machine Poker Competition in Las Vegas. This success demonstrated super-human abilities, leaving just the question of how close computer programs could come to having a perfect, unbeatable Poker strategy, measured by the maximum amount the program could lose, on average, against a perfect adversary. If a program was developed whose maximum loss was at or very near zero, then the game could be considered solved.

This year we have achieved this result, essentially solving the game of 2-player limit Texas hold'em Poker. This was achieved by combining counterfactual regret minimization (CFR⁺) with techniques for minimizing the search space, and then applying massive computational power. The ultimate CFR⁺ implementation ran on a cluster of 200 nodes each with 24 2.1 Ghz cores, 32 GB of RAM, and a 1-TB local disk. The computation ran for more than 68 days, consuming a total of 900 core-years. The final Poker strategy was not shown to be optimal, but was shown to be so close that any difference from optimality would be impossible to measure, in a statistical sense, in a lifetime of playing Poker (estimated at 200 hands per hour, 12 hours a day, for 70 years). This is a landmark result in the field of computational game theory, as it is the first time a human-scale game of imperfect information has been essentially solved.

This result is also a landmark in AI and computer science. It was announced this January in a research article *Science*, and resulted in followon news articles in *The New York Times*, *The Washington Post*, *The Guardian*, *The Economist*, and the BBC. Associated Press and Reuters had syndicated stories that appeared in many other newspapers. Articles also appeared in diverse magazines such as *Popular Mechanics*, *Cigar Aficionado*, and *Maxim*. *CBC Radio*, (US) *National Public Radio*, and *CBS News Radio* also ran stories. Michael Bowling gave 52 interviews over the phone or in person, and others on the team gave at least ten more. The project's web page had over a million hits from 250,000 distinct users in the first week after the announcement. The *Association for the Advancement of Artificial Intelligence* (AAAI) included this achievement as one of only five stories to highlight about AI in a recent press conference.

Emphatic TD(λ) and Off-policy Learning

Off-policy learning refers to learning about one decision-making policy from data obtained while following a different policy. The policy learned about is called the *target policy*, and the policy used to generate the data is called the *behaviour policy*. The special case in which the two policies are the same, called the *on-policy* case, is an easier learning problem, but the general, off-policy case must be addressed to obtain the full power of reinforcement learning. For example, suppose you are observing a chemical plant while it is controlled by a person (the behaviour policy) but hoping to learn a better policy (the target policy) that you would use if you were allowed to take control of the plant. This is a case of off-policy learning. In general, off-policy learning is important whenever you want to take full advantage of the available data without interfering with an existing controller—a very common case indeed. Unfortunately, reinforcement learning methods for the off-policy case are often unstable when combined with function approximation. This is arguably the biggest technical problem in modern reinforcement learning, and much of the RLAI project's research in recent years has been devoted to solving it (that is, to finding off-policy learning methods that are compatible with function approximation). In past years, we have developed the gradient-TD family of learning algorithms which have been the best known solution to this problem.

This year we have developed a new solution idea, called *emphasis*. The idea is to vary the intensity of the update of classical reinforcement learning methods such as TD(λ) from step to step based on the sequence of recent actions. If the recent actions were unlikely to occur under the target policy, then the update is de-emphasized. On the other hand, if the recent actions were more likely to occur under the target policy than the behavior policy, then the updates are emphasized. We have been able to prove mathematically that if the emphasis is varied in just the right way, then the overall expected update is stable, and convergence of the learning algorithm, called *emphatic TD(λ)*, is guaranteed. Moreover, this is achieved with only a trivial increase in computational complexity (only one additional scalar memory variable is required). This is a substantial improvement in computational complexity over gradient-TD methods, which require a whole second vector of memory parameters and a second step-size parameter. Nevertheless, there are numerous other practical issues that could favor one class of solution method over another, and we have yet to perform significant empirical comparisons. There is a rich set of new conceptual issues arising with the emphatic idea and we expect it to be a focus of core algorithmic research in the project in future years.

Weighted Importance Sampling

Off-policy learning methods rely on a special statistical technique called *importance sampling*. The idea of this technique is to scale the observed outcomes according to the ratio of the probabilities of the trajectory of states and actions under the target and behavior policies. When a trajectory occurs that is more likely under the target policy than the behavior policy, then the outcome is scaled up accordingly. And if a trajectory occurs that is less likely under the target policy than the behavior policy, then the outcome is scaled down (towards zero). Importance sampling appears to be essential to any method for off-policy learning. However, the ratios can sometimes become very

large, causing high-variance updates and very slow learning. To some extent this may be unavoidable, but we want to reduce variance as much as possible. In conventional statistical applications of importance sampling considerable reductions in variance are often obtained using a variation known as *weighted* importance sampling (WIS).

The idea of WIS is very simple. Whereas an ordinary importance sampling (OIS) estimate is a simple average of the scaled outcomes, a WIS estimate is a weighted average of the scaled outcomes using the scaling ratios as the weights. WIS has much lower variance than OIS, albeit at the cost of some bias, and is almost universally preferred and used in practical statistical applications. We would like to obtain the benefits of WIS in our off-policy learning algorithms, but unfortunately there is no established effective way of combining WIS with function approximation, which is essential for the large applications that we target in RLAI research.

This year we have made significant progress overcoming this problem by combining WIS with linear function approximation. First, we established an ideal way of doing this in supervised learning if we relax concerns about computational complexity and allow computation that is quadratic or even cubic in the number of features and weights. This work convinced us that a linear complexity solution that exactly reproduced the ideal solution was not possible. Second, after much experimentation, we found a linear complexity algorithm that closely approximates the ideal solution in many cases. We are currently seeking to extend these ideas to general temporal-difference and reinforcement-learning settings.

Real-time Autonomous Switching in Prosthetic Interfaces

Assistive biomedical devices augment the abilities of amputees and other patients with lost physical or cognitive function. Specifically, these devices replace abilities lost due to traumatic injury, disease, aging, or congenital complications. This project focuses on one representative class of assistive robots: powered artificial limbs. Powered prostheses monitor electromyographic (EMG) signals produced by muscle tissue in a patient's body, and use these signals to control the movement of a robotic appendage with one or more controllable dimensions. Despite the potential for improved function with myoelectric control, many patients reject the use of powered artificial limbs. Recent needs-assessment surveys of prosthesis users point out three key principal reasons for patient rejection: lack of intuitive control, insufficient functionality, and insufficient feedback from the myoelectric device. RLAI research led by faculty team member Patrick Pilarski, in collaboration with the Glenrose Rehabilitation Hospital (GRH), the Division of Physical Medicine and Rehabilitation, and the Department of Mechanical Engineering, aims to remove these barriers and increase the independence and ability of amputees through the use of adaptive machine learning techniques.

In previous years we showed in pre-clinical trials that RLAI prediction-learning algorithms can significantly improve the amputee-prosthesis interface. In one common EMG setup, the amputee uses a myoelectric signal from a stump muscle to control joint at a time of the prosthetic arm, switching between joints using a separate toggle switch. The time spent switching through joints can consume up to half the total time to perform many tasks. We showed that this wasted time can be significantly reduced if the joints

were presented in an intelligent, user- and task-specific order. These results were on data from amputee and able-bodied subjects, but the learning was done off-line, after the subjects were disconnected from the system.

This year Pilarski has extended his work in two major ways. The first is to perform the learning in real-time, during the subjects' interaction with the prosthesis, and actually present the subjects with joints in the learned order. Thus we have gone beyond just showing that learning *could* have saved the subject time; now our learning algorithms actually *do* save that time. The second major extension is still in the early stages, but we have demonstrated prototypes in which the subjects do not even have to explicitly ask for a switch of the controlled joint. If they merely pause in their motion, and our algorithm is confident enough in its prediction of which joint the subject wants next, then it will switch to it automatically without being asked, saving even more time.

4. OBJECTIVES FOR NEXT YEAR

With regard to core reinforcement learning algorithms, we envision taking several further steps over the next year. One is to resolve certain long-standing issues in off-policy Monte Carlo learning. A second is to complete a thorough empirical study of the true-online TD(λ) algorithm introduced last year. A third is to explore the hypothesis that the emphasis idea helps not just with the stability of off-policy learning, but also with the asymptotic solution quality of on-policy learning. More ambitiously, we hope that the theoretical understanding that we have obtained from that algorithm might enable us to find yet a new stable and efficient off-policy algorithm by controlling stability in the feature dimensions. We also seek various combinations of previously obtained innovations: emphasis with true-online, WIS with gradient-TD, and so forth.

With regard to the larger ambitions of AI, work will proceed on several fronts. One is to apply new algorithms such as emphatic TD(λ) to learn extensive dynamics knowledge in robotics domains. We seek to learn models of these domains that can be used for planning. We are now developing a new architecture in which planning of action can be combined with direct reinforcement learning of action in a *cascade* architecture in which planning can give early fast improvements in performance while reinforcement learning enables optimal long-term performance. In this way we hope to avoid the common problematic outcome in which structural inaccuracies of the model, which are always present, permanently debilitate the learning agent's long-run performance.

With regard to the adaptive prosthetics effort, next year we will be utilizing the newly purchased (using CFI funds) \$700,000 Modular Prosthetic Limb (MPL v3, one of only two in the world). We hope to scale the parallel learning architecture Horde (introduced a few years ago) by at least two orders of magnitude, using specialized GPU (Graphic Processing Unit) hardware, and apply it to the MPL. We are also planning pre-clinical studies of automatic switching and of what we call *direct predictive collaborative control*, in which the system learns to control many joints simultaneously in collaboration with the amputee user.

5. RESEARCH TEAM MEMBERS AND CONTRIBUTIONS

a. Team Leader

Name	Role	Awards / Special Info
Richard Sutton	Principal Investigator	AICML, NSERC

b. Faculty Team Members

Michael Bowling	Faculty team member	AICML, NSERC
Dale Schuurmans	Faculty team member	AICML, CRC II Chair, NSERC, MITACS
Csaba Szepesvári	Faculty team member	AICML, NSERC
Patrick Pilarski	Faculty team member	AICML, NSERC
András György	Adjunct Faculty Member; Research Associate	Computer and Research Institute of the Hungarian Academy of Sciences
Martin Müller	Associated faculty member	NSERC

c. Post Doctoral Fellows & Research Associates

Name	Role	Awards / Special Info
Joseph Modayil	Research Associate	
Huizhen Yu	Research Associate	
Patrick Pilarski	Post doctoral fellow, adjunct faculty member	Now a faculty team member at the University of Alberta, Department of Medicine
Harm van Seijen	Post doctoral fellow	
Hado van Hasselt	Post doctoral fellow	Now employed at Google Deepmind in London
Tor Lattimore	Post doctoral fellow	
Karim Abou-Moustafa	Post doctoral fellow	
Mohamed Elgendi	Post doctoral fellow	Now postdoctoral fellow at UBC, Biomedical Engineering
Ramki Gummadi	Post doctoral fellow	Now data scientist at Facebook in Menlo Park, CA

d. PhD Candidates

Name	Role	Scholarships / Awards / Special Info
Gabor Balazs	PhD candidate	
Nolan Bard	PhD candidate	
Neil Burch	PhD candidate	AIF, NSERC CGS,
Katherine Chen	PhD candidate	
Kenneth Dwyer	PhD candidate	
Pooria Joulani	PhD candidate	
Ruitong Huang	PhD candidate	
Michael Johanson	PhD candidate	
Anna Koop	PhD candidate	
Ashique Mahmood	PhD candidate	
Ozlem Aslan	PhD candidate	
Bernardo Ávila Pires	PhD candidate	2014 Computing Science GPA Award
James Neufeld	PhD candidate	
Chris Rayner	PhD candidate	
Adam White	PhD candidate	
Martha White	PhD candidate	NSERC CGS D, Honorary Izaak Walton Killam Memorial Scholarship
Farzaneh Mirzazadeh	PhD candidate	NSERC PGSD2, Computing Science GPA Award, Queen Elizabeth II Scholarship
Hengshuai Yao	PhD candidate	
Mohammad Ajallooeian	PhD candidate	
Junfeng Wen	PhD candidate	
Marlos Machado	PhD candidate	AITF Graduate Student Scholarship

Leah Hackman	PhD candidate	
Mahdi Karami	PhD candidate	
Kiarash Shaloudegi	PhD candidate	

e. MSc Candidates

Name	Role	Scholarships / Awards / Special Info
Trevor Davis	MSc candidate	Alberta Innovates Graduate Student Scholarship, University of Alberta Doctoral Recruitment Scholarship
Kavosh Asadi Atui	MSc candidate	
Bing Xu	MSc candidate	
Ujjwal Das Gupta	MSc candidate	
Roshan Shariff	MSc candidate	
Sriram Srinivasan	MSc candidate	
Travis Dick	MSc candidate	Queen Elizabeth II Graduate Scholarship Masters Level, Alexander Graham Bell Scholarship M
Joshua Davidson	MSc candidate	
Shun Jie Lau	MSc candidate	
Kory Mathewson	MSc candidate	
Ann Edwards	MSc candidate	
Craig Sherstan	MSc candidate	Alberta Innovates Graduate Student Scholarship
Sanjeev Sharma	MSc candidate	
Brendan Bennett	MSc candidate	
Tian Tian	MSc candidate	Queen Elizabeth Scholarship II
Yifan Wu	MSc candidate	Jeffrey R Sampson Scholarship

Min Yang	MSc candidate	
Xiaowei Hu	MSc candidate	
Tim Yee	MSc candidate	
Dustin Morril	MSc candidate	Alberta Innovates-Tech Futures, NSERC CGSM

f. Other Members

Name	Role
Beverly Balaski	Program administrator
Alexandra Kearney	Undergrad researcher
Adam Parker	Undergrad researcher
Dylan Brenneis	Undergrad researcher
Jaden Travnik	Undergrad researcher
Adeline Wang	High school summer student, July-August 2014
Nicholas Westbury	High school summer student, July-August 2014
Pavlo Malynin	High school summer student, July-August 2014
Jason Yaskowich	High school summer student, July-August 2014
Maharsh Patel	High school summer student, July-August 2014
Aishwarya Venkitachalam	High school summer student, July-August 2014
David Pankratz	High school summer student, July-August 2014

g. Visitors

Name	Institution
Tom Schaul	Google Deepmind
Nicola Catenacci Volpi	University of Hertfordshire
David Krueger	University of Montreal
Shiva Kasiviswanathan	General Electric Global Research Center, San Ramon, CA
Yasin Abbasi-Yadkori	Queensland University of Technology Brisbane, Australia
Lihong Li	Microsoft Research Redmond, WA
Andras Antos	Senior Research Scientist from Hungary
Elizabeth Guevara	CICIPN, Mexico
Balazs Kegl	CNRS, Linear Accelerator Laboratory (France)
Wouter M. Koolen	Queensland University of Technology, Australia
Mark Schmidt	Department of Computer Science at the University of British Columbia
Marc G. Bellemare	Research Scientist at Google DeepMind in London, UK
Kevin Waugh	Carnegie Mellon University, Pittsburgh, PA
Yang Wang	University of Manitoba

6. COLLABORATIONS

Provincial	
Participants	Nature of Collaboration
Alberta Ingenuity Center for Machine Learning (AICML)	R. Sutton, D. Schuurmans, Cs. Szepesvári, Patrick Pilarski, and M. Bowling are among the ten principle investigators for this center at the University of Alberta. Total annual funding for AICML is \$2M

<p>Patrick M. Pilarski and Richard Sutton (RLAI), Linda Pilarski (U of A, Faculty of Medicine), and Carina Debes Marun (Cross Cancer Institute)</p>	<p>Collaboration with researchers from the AHFMR Team Microfluidics, Alberta Innovates Health Solutions, University of Alberta, and Cross Cancer Institute to explore intelligent biomedical image analysis methods to facilitate rapid lab-on-a-chip diagnostics. Part of the technology transfer focused “nanoBridge Research Grant RES-NAN-07-G10: Development of Fluorescence In Situ Hybridization (FISH) Platform, Chip, and Analysis Software” (\$145K)</p>
<p>Patrick Pilarski and Richard Sutton (RLAI) and Dr. Kim Adams University of Alberta Rehab Med and Dr. Mahdi Tavakoli Department of Electrical and Computer Engineering University of Alberta</p>	<p>Collaborative work on the use of assistive robots to facilitate the development, skill acquisition, and cognitive assessment of young children and infants with severe motor impairments (e.g., children with cerebral palsy); work involves the use of machine intelligence to enable the automatic adaptation of a robot control system to match the needs and abilities of individual children. Pilarski and Adams co-supervise visiting graduate students from a partner institution in Mexico city. (CHRP \$333K over three years)</p>
<p>Patrick Pilarski (RLAI) and Martin Ferguson-Pell (Dean, Faculty of Rehabilitation Medicine, University of Alberta), Liping Qi (University of Alberta, Rehabilitation Medicine), Simon Grange (University of Calgary / University of Alberta)</p>	<p>Collaboration between the RLAI and the Rehabilitation Robotics Sandbox Laboratory (Rehabilitation Medicine, U. Alberta); this collaboration focuses on using new machine learning methods to predict fatigue in wheelchair users and enable novel muscle stimulation paradigms.</p>
<p>Patrick Pilarski and Richard Sutton (RLAI), and AHFMR SMART Neuroprosthetics Team, University of Alberta</p>	<p>Preliminary work on the use of machine learning to adapt and optimize neural interfaces and assistive robotic devices for people with motor disabilities.</p>
<p>L.M. Pilarski (Dept. Oncology, University of Alberta), L.M. McMullen (Agriculture, University of Alberta), M. Gaenzle (Agriculture, University of Alberta), Patrick Pilarski (RLAI), X. Yang (Lacombe Research Centre)</p>	<p>The meat processing industry needs a portable and rapid platform to detect pathogens during meat processing. We work with end users and colleagues to develop inexpensive computer systems that enable the automation of the testing system. This work includes software and hardware automation, biomedical pattern analysis, and machine intelligence.</p>
<p>Dr. Ian Adatia, Director of the Pediatric Cardiac Critical and Intermediate Care Program, Stollery Children's Hospital and Roger Zemp, Department of Electrical and Computer Engineering University of Alberta</p>	<p>Joint research with Dale Schuurmans on adaptive signal processing techniques for heart disease diagnosis, focusing on predicting events with ECG and PCG signals.</p>

Joseph Modayil and Gary Faulkner, Jim Raso, Michael Cimolini from the Glenrose Rehabilitation Hospital	Collaboration with the “Robotics in the Glenrose Courage Centre.” This project is studying the interaction between people and autonomous robots in a hospital setting.
--	--

National	
Participants	Nature of Collaboration

International	
Participants	Nature of Collaboration
Zinhua Zhang, NICTA	Joint research with Dale Schuurmans on representation learning and optimization.
Yuhong Guo, Temple University	Joint research with Dale Schuurmans on machine learning in bioinformatics and machine learning methods for large scale multi-target prediction.
Yaoliang Yu, Carnegie Mellon University	Joint research with Dale Schuurmans on efficient optimization methods for machine learning, convex reformulations of representation learning problems, and algorithmic techniques for exploiting structured sparsity.
Shie Mannor, Technion, Haifa, Israel	Joint research with Csaba Szepesvári on online collaborative filtering.
Andras Antos, MTA SZTAKI, Hungary	Joint research with Csaba Szepesvári on active learning.
Lihong Li, Microsoft Research, Redmond	Joint research with Csaba Szepesvári on off- policy learning.
Branislav Kveton, Adobe Research, San Jose	Joint research with Csaba Szepesvári on bandit combinatorial optimization.
Koby Crammer, Technion, Haifa, Israel	Joint research with Csaba Szepesvári on bandit resource allocation.
Mohammad Ghavamzadeh, Adobe Research, San Jose	Joint research with Csaba Szepesvári on classification calibration.
David Pal, Yahoo Research, New York	Joint research with Csaba Szepesvári on online learning.
Yasin Abbasi-Yadkori, Queensland University of Technology, Brisbane, Australia	Joint research with Csaba Szepesvári on online learning in control.

Erik Talvitie, Franklin and Marshall College	Joint research and supervision of graduate students with Michael Bowling, developing new algorithms for intelligent exploration in domains with sparse rewards.
Nathan Sturtevant, University of Denver	Joint research with Michael Bowling on multiplayer game theory and automating heuristic construction for search.
Joel Veness and Marc Bellemare, Google DeepMind	Joint research with Michael Bowling on developing the Arcade Learning Environment, and algorithms for using compression in game theory.
Patrick Pilarski and Dr. Xuantao Su, Shandong University, China	Ongoing collaboration to explore the use of machine intelligence, pattern analysis, and optical lab-on-chip technology for the assessment of disease-related nanostructures in human cells.
Patrick M. Pilarski (RLAI) and Sophia Adamia (Dana Farber Cancer Institute, Harvard Medical School, USA).	Investigation of cancer-related genetic markers in patients. Collaboration involving machine learning and biomedical data mining methods.
Patrick Pilarski and Richard Sutton (RLAI), Jason P. Carey (Department of Mechanical Engineering, University of Alberta), Michael R. Dawson, Jacqueline S. Hebert, K. Ming Chan (Glenrose Rehabilitation Hospital).	“Reinforcement Learning for Adaptive Prosthetics;” this collaboration investigates the use of reinforcement learning and real-time machine learning to enable adaptive, intuitive control of myoelectric prostheses and other assistive robotic devices.
Gergely Neu, INRIA Lille	Joint research with Andras Gyorgy on online learning, Markov decision problems, and adaptive source coding.
Sandor Laki, Eötvös Loránd University, Budapest, Hungary and Béla Hullár, ETH Zürich, Switzerland	Joint research with Andras Gyorgy on Internet traffic classification based on packet content.

7. GRADUATES

Name	Degree	Research topic	Current Position
Ujjwal das Gupta	MSc	Adaptive Representation for Policy Gradient	Google, Mountain View CA.
Srinivasan Sriram	MSc	State generalization in UCT using local manifolds	Amazon
Travis Dick	MSc	Policy Gradient Reinforcement Learning Without Regret	PhD Student at Carnegie Mellon University
Martha White	PhD	Regularized Factor Models	Assistant professor at Indiana University
Chris Rayner	PhD	Optimization for Heuristic Search	Search and Business Intelligence Developer at Dissolve, Calgary, AB

8. INTELLECTUAL PROPERTY

Intellectual Property	Status	Short Description
PATENTS	none	
LICENSES		
Spinoff Companies	none	

9. PUBLICATIONS

REFEREED JOURNAL PUBLICATIONS

K. Abou-Moustafa, F. De La Torre, and F. Ferrie, "Pareto Models for Discriminative Multiclass Linear Dimensionality Reduction," *Pattern Recognition*, Jan. 2015, 1, pp. 1863-1877.

K. Abou-Moustafa, and F. Ferrie, "Local Generalized Quadratic Distant Metrics for Query Based Operations," *IEEE Transaction on Cybernetics*, Feb. 2015.

G. Bartok, D. Foster, D. Pal, A. Rakhlin, and Cs. Szepesvari, "Partial Monitoring -- Classification, Regret Bounds, and Algorithms," *Mathematics of Operations Research*, Dec. 2014.

M. Bowling, N. Burch, M. Johanson, and O. Tammelin, "Heads-up Limit Hold'em Poker is Solved," *Science*, Jan. 2015, pp. 145-149.

C. Castellini, P. Artemiadis, M. Wininger, A. Ajoudani, M. Alimusaj, A. Bicchi, B. Caputo, W. Craelius, S. Dosen, K. Englehart, D. Farina, A. Gijsberts, S. Godfrey, L. Hargrove, M. Ison, T.A. Kuiken, M. Markovic, P.M. Pilarski, R. Rupp, and E.J. Scheme, "Proceedings of the First Workshop on Peripheral Machine Interfaces: Going Beyond Traditional Surface Electromyography," *Frontiers in Neurorobotics*, Aug. 2014, pp. 1-17.

A. Edwards, M.R. Dawson, J.S. Hebert, C. Sherstan, R.S. Sutton, M. Chan, and P.M. Pilarski, "Application of Real-Time Machine Learning to Myoelectric Prosthesis Control: A Case Study in Adaptive Switching," *Prosthetics and Orthotics International*, to appear.

M. Elgendi, "Detection of c, d, and e Waves in the Acceleration Photoplethysmogram," *Computer Methods and Programs in Biomedicine*, Aug. 2014.

- M. Elgendi, P. Bobhate, S. Jain, L. Guo, J. Rutledge, Y. Coe, R. Zemp, D. Schuurmans, and I. Adatia, "Time-domain Analysis of Heart Sound Intensity in Children With and Without Pulmonary Artery Hypertension: A Pilot Study Using a Digital Stethoscope," *Pulmonary Circulatio*, Jul. 2014.
- M. Elgendi, F. Picon, N. Magnenat-Thalmann, and D. Abbott, "Arm Movement Speed Assessment Via Kinect Camera: A Preliminary Study in Healthy Subjects," *Biomedical Engineering OnLine*, Jun. 2014.
- A. Gyorgy, and G. Neu, "Near-Optimal Rates for Limited-Delay Universal Lossy Source Coding," *IEEE Transactions on Information Theory*, May 2014, pp. 2823-2834.
- B. Hullar, S. Laki, and A. Gyorgy, "Efficient Methods for Early Protocol Identification," *IEEE Journal on Selected Areas of Communications*, Oct. 2014, pp. 1907-1918.
- E.J. Kehoe, E.A. Ludvig, and R.S. Sutton, "Time Course of the Rabbit's Nictitating Membrane During Acquisition, Extinction, and Reacquisition," *Learning and Memory*, Oct. 2014.
- T. Lattimore, and M. Hutter, "Near-optimal PAC Bounds for Discounted MDPs," *Theoretical Computer Science*, Nov. 2014.
- T. Lattimore, and M. Hutter, "On Martin-Löf (Non-)convergence of Solomonoff's Universal Mixture," *Theoretical Computer Science*, Dec. 2014.
- T. Le, R. Zheng, and Cs. Szepesvari, "Sequential Learning for Multi-channel Wireless Network Monitoring with Channel Switching Costs," *IEEE Transactions on Signal Processing*, Oct. 2014.
- P.M. Pilarski, A. Edwards, and K.M. Chan, "Novel Control Strategies for Arm Prostheses: A Partnership Between Man and Machine," *The Japanese Journal of Rehabilitation Medicine*, Mar. 2015, pp. 91-95.
- H. Yu, "On Convergence of Value Iteration for a Class of Total Cost Markov Decision Processes," *SIAM Journal on Control and Optimization*, to appear.
- H. Yu, and D.P. Bertsekas, "A Mixed Value and Policy Iteration Method for Stochastic Control with Universally Measurable Policies," *Mathematics of Operations Research*, to appear.

HIGHLY REFEREED ARCHIVAL CONFERENCE PROCEEDINGS

O. Aslan, X. Zhang, and D. Schuurmans, “Convex Deep Learning via Normalized Kernels,” *Proc. Conf. Advances in Neural Information Processing Systems* (NIPS 2014), Dec. 2014, 24% acceptance.

G. Balazs, Cs. Szepesvari, and A. Gyorgy, “Near-optimal Max-affine Estimators for Convex Regression,” *Proc. Int’l Conf. on Artificial Intelligence and Statistics* (AISTATS 2015), Feb. 2015, pp. 56–64, 29% acceptance.

N. Bard, M. Johanson, and M. Bowling, “Asymmetric Abstractions for Adversarial Settings,” *Proc. 13th Int’l Conf. on Autonomous Agents and Multiagent Systems* (AAMAS 2014), May 2014, pp. 501-508, 24% acceptance.

N. Bard, D. Nicholas, Cs. Szepesvari, and M. Bowling, “Decision-theoretic Clustering of Strategies,” *Proc. 14th Int’l Conf. on Autonomous Agents and Multiagent Systems* (AAMAS 2015), to appear, 25% acceptance.

N. Burch, M. Johanson, and M. Bowling, “Solving Imperfect Information Games Using Decomposition,” *Proc. 28th AAAI Conf. on Artificial Intelligence* (AAAI 2014), Jul. 2014, 28% acceptance.

U. Das Gupta, E. Talvitie, and M. Bowling, “Policy Tree: Adaptive Representation for Policy Gradient,” *Proc. 29th AAAI Conference on Artificial Intelligence* (AAAI 2015), Jan. 2015.

T. Davis, N. Burch, and M. Bowling, “Using Response Functions to Measure Strategy Strength,” *Proc. 28th AAAI Conf. on Artificial Intelligence* (AAAI 2014), Jul. 2014, 28% acceptance.

I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative Adversarial Nets,” *Proc. Conf. Advances in Neural Information Processing Systems* (NIPS 2014), Jun. 2014.

R. Huang, and Cs. Szepesvari, “A Finite-Sample Generalization Bound for Semiparametric Regression: Partially Linear Models,” *Proc. 17th Int’l Conf. on Artificial Intelligence and Statistics* (AISTATS 2014), Apr. 2014.

T. Lattimore, K. Crammer, and Cs. Szepesvari, “Optimal Resource Allocation with Semi-Bandit Feedback,” *Proc. 30th Conf. on Uncertainty in Artificial Intelligence* (UAI 2014), Jul. 2014.

T. Lattimore, A. Gyorgy, and Cs. Szepesvari, “On Learning the Optimal Waiting Time,” *Proc. Int’l Conf. on Algorithmic Learning Theory* (ALT 2014), Oct. 2014, pp. 200-214.

T. Lattimore, and M. Hutter, “Bayesian Reinforcement Learning with Exploration,” *Proc. Int’l Conf. Algorithmic Learning Theory (ALT 2014)*, Oct. 2014.

T. Lattimore, and R. Munos, “Bounded Regret for Finite-Armed Structured Bandits,” *Proc. Conf. Advances in Neural Information Processing Systems (NIPS 2014)*, Dec. 2014.

A. Mahmood, H. van Hasselt, and R.S. Sutton, “Weighted Importance Sampling for Off-policy Learning With Linear Function Approximation,” *Proc. Conf. Advances in Neural Information Processing Systems (NIPS 2014)*, Dec. 2014.

A. Mahmood, and R.S. Sutton, “Off-policy Learning Based on Weighted Importance Sampling With Linear Computational Complexity,” *Proc. 31st Conf. on Uncertainty in Artificial Intelligence (UAI 2015)*, Mar. 2015.

F. Mirzazadeh, Y. Guo, and D. Schuurmans, “Convex Co-embedding,” *Proc. 28th AAAI Conf. on Artificial Intelligence (AAAI 2014)*, Jul. 2014, pp. 1989-1996, 28% acceptance.

J. Neufeld, D. Schuurmans, and M. Bowling, “Variance Reduction via Antithetic Markov Chains,” *Proc. Int’l Conf. on Artificial Intelligence and Statistics (AISTATS 2015)*, Feb. 2015, 28% acceptance.

R. Shariff, Cs. Szepesvari, and A. Gyorgy, “Exploiting Symmetries to Construct Efficient MCMC Algorithms with an Application to SLAM,” *Proc. Int’l Conf. on Artificial Intelligence and Statistics (AISTATS 2015)*, Feb. 2015, pp. 866–874, 29% acceptance.

S. Sriram, E. Talvitie, and M. Bowling, “Improving Exploration in UCT Using Local Manifolds,” *Proc. 29th AAAI Conf. on Artificial Intelligence (AAAI 2015)*, Jan. 2015.

O. Tammelin, N. Burch, M. Johanson, and M. Bowling, “Solving Heads-up Limit Texas Hold'em,” *Proc. Int’l Joint Conf. on Artificial Intelligence (IJCAI 2015)*, to appear.

H. van Hasselt, A. Mahmood, and R. S. Sutton, “Off-policy TD(λ) with a true online equivalence,” *Proc. 30th Conf. on Uncertainty in Artificial Intelligence (UAI 2014)*, Jul. 2014.

K. Waugh, D.R. Morrill, A. Bagnell, and M. Bowling, “Solving Games with Functional Regret Estimation,” *Proc. 29th AAAI Conf. on Artificial Intelligence (AAAI 2015)*, Jan. 2015.

J. Wen, C. Yu, and R. Greiner, “Robust Learning under Uncertain Test Distributions: Relating Covariate Shift to Model Misspecification,” *Proc. 31st Int’l Conf. on Machine Learning (ICML 2014)*, Jun. 2014.

M. White, J. Wen, M. Bowling, and D. Schuurmans, “Optimal Estimation of Multivariate ARMA Models,” *Proc. 29th AAAI Conf. on Artificial Intelligence*, (AAAI 2015), Jan. 2015.

H. Yao, Cs. Szepesvari, R.S. Sutton, S. Bhatnagar, and J. Modayil, “Universal Option Models,” *Proc. Conf. Advances in Neural Information Processing Systems* (NIPS 2014), Dec. 2014, 23% acceptance.

H. Yu, “On Convergence of Emphatic Temporal-Difference Learning,” *Conf. on Learning Theory* (COLT 2015), to appear.

OTHER CONFERENCE AND WORKSHOP PROCEEDINGS

B. Ávila Pires, and Cs. Szepesvari, “Pathological Effects of Variance on Classification-Based Policy Iteration,” *29th AAAI Conf. Learning for General Competency in Video Games Workshop* (AAAI 2015), Workshop paper, Jan. 2015.

N. Bard, D. Nicholas, Cs. Szepesvari, and M. Bowling, “Decision-theoretic Clustering of Strategies,” *AAAI Workshop on Computer Poker and Imperfect Information* (AAAI 2015), Workshop paper, Jan. 2015.

M.R. Dawson, C. Sherstan, J.P. Carey, J.S. Hebert, and P.M. Pilarski, “Development of the Bento Arm: An Improved Robotic Arm for Myoelectric Training and Research,” *Proc. of Myoelectric Controls Symposium* (MEC 2014), Aug. 2014, pp. 60-64.

A. Edwards, M.R. Dawson, J.S. Hebert, R.S. Sutton, K.M. Chan, and P.M. Pilarski, “Adaptive Switching in Practice: Improving Myoelectric Prosthesis Performance through Reinforcement Learning,” *Proc. Myoelectric Controls Symposium* (MEC 2014), Aug. 2014, pp. 69-73.

M. Elgendi, “Preliminary Study for Localizing c, d, and e Waves in Photoplethysmogram Signals,” *Proc. 36th Annual Int’l Conf. of the IEEE Engineering in Medicine and Biology Society* (IEEE 2014), Aug. 2014.

T. Everitt, T. Lattimore, and M. Hutter, “Free Lunch for Optimisation under the Universal Distribution,” *Proc. IEEE Congress on Evolutionary Computation*, Jul. 2014.

R. Gummadi, “Resampled Belief Networks for Variational Inference,” *NIPS 2014 Workshop on Advances in Variational Inference*, Workshop Paper, Dec. 2014.

J. Gunther, P.M. Pilarski, G. Helfrich, H. Shen, and K. Diepold, “First Steps Toward an Intelligent Laser Welding Architecture Using Deep Neural Networks and Reinforcement Learning,” *Proc. 2nd Joint Int’l Conf. on System-integrated Intelligence: New Challenges for Product and Production Engineering* (SysInt 2014), Jul. 2014, pp. 481–490.

A. Koop, A. Kearney, M. Bowling, and P.M. Pilarski, "Dealing with Changing Contexts in Myoelectric Control," *Proc. Myoelectric Controls Symposium (MEC 2014)*, Aug. 2014, pp. 117-120.

P. Kumar, F. Mahmood, D. Menoth, K. Wong, A. Agrawal, M. Elgendi, R. Shukla, J. Dauwels, and A. Chan, "On the Effect of Subliminal Priming on Subjective Perception of Images: A Machine Learning Approach," *Proc 36th Annual Int'l Conf. of the IEEE Engineering in Medicine and Biology Society (IEEE 2014)*, Aug. 2014.

M. Machado, S. Srinivasam, and M. Bowling, "Domain-Independent Optimistic Initialization for Reinforcement Learning," *29th AAAI Conf. Learning for General Competency in Video Games workshop (AAAI 2015)*, Workshop paper, Jan. 2015.

J. Modayil, and R.S. Sutton, "Prediction Driven Behavior: Learning Predictions that Drive Fixed Responses," *AAAI Workshop on AI and Robotics Workshop (AAAI 2014)*, Workshop paper, Jul. 2014.

C. Sherstan, and P.M. Pilarski, "Multilayer General Value Functions for Robotic Prediction and Control," *2014 IROS Workshop on AI and Robotics*, Sep. 2014.

A. White, J. Modayil, and R.S. Sutton, "Surprise and Curiosity for Big Data Robotics," *AAAI Workshop on Sequential Decision Making with Big Data*, Jul. 2014, Workshop paper.

M. White, "A Factorization Perspective for Learning Representations in Reinforcement Learning," *AAAI Workshop on Sequential Decision-Making with Big Data*, Jul. 2014, Workshop paper.

SPECIAL/INVITED PRESENTATIONS

Person	Title	Venue
R. Sutton	Representation Learning: Learning Slow to Enable Learning Fast	International Conference on Representation Learning, Banff
R. Sutton	Foundations, Trends, and Fashions in Artificial Intelligence Research	Google Deepmind, London
R. Sutton	Reinforcement Learning and Psychology: A Personal Story	Department of Psychology Colloquium, Warrick University, UK
R. Sutton	Reinforcement Learning and Psychology: A Personal Story	Keynote at the annual meeting of the Society for Mathematical Psychology, Quebec City
R. Sutton	New Ideas for Reinforcement Learning in the Era of Big Data	AAAI Workshop on Sequential Decision-Making with Big Data, Quebec City
R. Sutton	Temporal-difference Learning and the Coming of Artificial Intelligence	School of Computer Science, McGill University, Montreal
R. Sutton	Temporal-difference Learning and the Coming of Artificial Intelligence	Distinguished Lecture Series, School of Computer Science, University of Massachusetts
R. Sutton	Multi-step Prediction	Team meeting of the CIFAR project on Neural Computation for Adaptive Perception, Montreal
R. Sutton	Artificial Intelligence as a Large-Scale MDP Problem	NIPS Workshop on Large-scale Reinforcement Learning, Montreal
R. Sutton	Challenges and Opportunities for Representation Learning in Reinforcement Learning	NIPS Workshop on Representation and Learning Methods for Complex Outputs, Montreal
R. Sutton	Creating Human-level AI: How and When?	Future of Artificial Intelligence Conference (invited plenary), Puerto Rico
R. Sutton	Introduction to Reinforcement Learning	Machine Learning Summer School, Austin, Texas (4.5 hours)
R. Sutton	The Grand Ambition of Reinforcement Learning and Artificial Intelligence, New Algorithms for Off-policy Temporal-Difference Learning, Independence of Span	Workshop on Recent Advances in Reinforcement Learning, Indian Institute of Technology, Madras (4.5 hours)
D. Schuurmans	Convex Methods for Representation Learning	SDM 2014 Workshop on Heterogeneous Data Mining
D. Schuurmans	Convex Methods for Representation Learning	University of BC Okanagan

D. Schuurmans	Convex Methods for Representation Learning	Fields Institute, Toronto
Cs. Szepesvári	Effective and Efficient Planning and Learning Algorithms for Model-Based Reinforcement Learning	Neural Information Processing Systems Conference (NIPS 2014)
Cs. Szepesvári	Online Learning under Delayed Feedback	Optimal Cooperation, Communication, and Learning in Decentralized Systems, Banff Workshop
Cs. Szepesvári	Adaptive Monte-Carlo via Bandit Allocation	Rutgers University, New Brunswick, NJ
Cs. Szepesvári	Sparse Stochastic Bandits	Google, Mountain View, CA
Cs. Szepesvári	Adaptive Monte-Carlo via Bandit Allocation	Adobe Research
Cs. Szepesvári	Sparse Stochastic Bandits	Stanford University
Cs. Szepesvári	Learning to Make Better Decisions: Challenges and Opportunities for the 21st Century	University of British Columbia
Cs. Szepesvári	Online-to-Confidence-Set Conversions and Application to Sparse Stochastic Bandits	Georgia Institute of Technology, Atlanta
Cs. Szepesvári	Online Learning	Indian Institute of Technology (IIT) Madras
M. Bowling	Adversaries, Abstractions, and Algorithms	Algorithmic Game Theory Seminar, CMU
M. Bowling	Games and Intelligence of the Artificial Kind	EPL Staff Development Day
M. Bowling	von Neumann's Dream	Google DeepMind, London, UK
M. Bowling	von Neumann's Dream	AAAI Conference on Artificial Intelligence, Austin, Texas
M. Bowling	von Neumann's Dream	INRIA Lille, France
M. Bowling	von Neumann's Dream	Google Mountain View, CA
M. Bowling	von Neumann's Dream	Samsung Research America, San Jose, CA
M. Bowling	Games and Intelligence of the Artificial Kind	Bay Area Alumni Event
M. Bowling	von Neumann's Dream	AAAI Spring Symposium on Applied Computational Game Theory, Palo Alto
A. Gyorgy	Online Learning in Markov Decision Processes with Changing Reward Sequences	Microsoft Research, Seattle,
A. Gyorgy	Online Learning in Adversarial Markov Decision Processes	Imperial College, London, UK

A. Gyorgy	Online Learning in Markov Decision Processes With Changing Reward Sequences	BIRS Workshop on Optimal Cooperation, Communication, and Learning in Decentralized Systems, Banff
P. M. Pilarski	Recent Advances in Upper-limb Prosthetic Technology	Queen Elizabeth Hospital, Hong Kong
P. M. Pilarski	Novel Control Strategies for Arm Prostheses - a Partnership Between Man and Machine	51st Japanese Association for Rehabilitation Medicine Annual Scientific Meeting
P. M. Pilarski	Knowledgeable Artificial Limbs	Centro de Investigación en Computación, Instituto Politécnico Nacional, Mexico
P. M. Pilarski	Automating Life: The Automated Past, Present, and Future of Human Life on Earth	Telus World of Science
H. van Hasselt	Efficient Off-policy Learning	Google DeepMind, London, UK

AWARDS

P.M. Pilarski: Myoelectric Control Symposium: Best Paper Award (Signal Processing Track)

THESES

Ujjwal Das Gupta, MSc, “Adaptive Representation for Policy Gradient,”
November 2014.

Srinivasan Sriram, MSc, “State Generalization in UCT using Local Manifolds,”
November 2014.

Travis Dick, MSc, “Policy Gradient Reinforcement Learning Without Regret,”
June 2015.

Chris Rayner, PhD, “Optimization for Heuristic Search,” November 2014.

Martha White, PhD, “Regularized Factor Models,” December 2014.

10. OUTREACH

Csaba Szepesvari supervised two high school students in July and August as part of the High School Internship Program and one WISEST student at the Department of Computing Science.

Richard Sutton gave a 10-minute radio interview on the Tencer and Grose Show (630 CHED) on the Google self-driving car (May 29, 2014).

Richard Sutton gave two interviews with grade-school classrooms as part of the VROC Partners in Research program (October 28, 2014 and August 20, 2014).

Patrick M. Pilarski gave a Laboratory tour and discussion to 9 WISEST students.

Patrick M. Pilarski led a High School Mentorship program at AJM High School

Joseph Modayil mentored the High School Internship student, (HIP student) Adeline Wang.

Joseph Modayil gave a demo of some research in the RLAI lab to the DiscoverE's Girl's Coding Club. The group had about 20 girls in grades 6-9 plus supervisors. They saw the critterbot, and the Nao. They also saw reinforcement learning demos with the Atari Learning Environment and with the Create robot.

Joseph Modayil gave a high-level talk on AI and robots, describing some recent advances in AI and how they might impact society in the future at St. Andrews United Church.

Harm van Seijen mentored two High School Internship students, (HIP student) Nicholas Westbury and Jason Yaskowich.

Ashique Mahmood mentored the High School Internship student, (HIP student) Maharsh Patel.

Anna Koop led a High School Mentorship program at WP Wagner High School.

Katherine Chen gave a presentation on robotics “How to survive a robot apocalypse” to Computer Science Summer camps.