
**All sessions will be held in the Fairholme Room of the Royal Canadian Lodge in Banff**

**Day 1 (May 14)**

| | | |
|---|---|---|
| 8:00 am | 8:50 am | Breakfast + Registration |
| 8:50 am | 9:00 am | Welcome Message |
| 9:00 am | 10:30 am | Keynote: *Reinforcement Learning as a Sequential Decision Problem using the Universal Modeling Framework*, Warren Powell (Princeton University) – Part 1 |
| 10:30 am | 11:00 am | Break |
| 11:00 am | 12:30 pm | Keynote: *Reinforcement Learning as a Sequential Decision Problem using the Universal Modeling Framework*, Warren Powell (Princeton University) – Part 2 |
| 12:30 pm | 1:30 pm | Lunch (provided) |
| 1:30 pm | 3:00 pm | Technical Session 1 |
| 3:00 pm | 4:00 pm | Break + Poster Session 1 |
| 4:00 pm | 5:30 pm | Technical Session 2 |
| 6:30 pm | 8:00 pm | Dinner (provided) |

**Day 2 (May 15)**

| | | |
|---|---|---|
| 8:00 am | 9:00 am | Breakfast + Registration |
| 9:00 am | 10:30 am | Keynote: *Learning in Stochastic Models: An Asymptotic Approach*, Amy Ward (University of Chicago Booth School of Business) – Part 1 |
| 10:30 am | 11:00 am | Break |
| 11:00 am | 12:30 pm | Keynote: *Learning in Stochastic Models: An Asymptotic Approach*, Amy Ward (University of Chicago Booth School of Business) – Part 2 |
| 12:30 pm | 1:30 pm | Lunch (provided) |
| 1:30 pm | 3:00 pm | Technical Session 3 |
| 3:00 pm | 4:00 pm | Break + Poster Session 2 |
| 4:00 pm | 5:30 pm | Technical Session 4 |
| 5:30 pm | 5:40 pm | Farewell |

Haskayne School of Business | University of Calgary

**Thursday, May 14, 2026**

| 9:00 am – 12:30 pm: Keynote by Warren B. Powell (Princeton University) | |
|---|---|
| | **Reinforcement Learning as a Sequential Decision Problem using the Universal Modeling Framework**<br>Reinforcement learning emerged in the 1980s in computer science for approximating Bellman's equation, building on the framework of Markov decision processes introduced in the 1950s by Bellman. This work paralleled a line of research from the optimal control community starting in 1974 for deterministic control problems, and then the work on approximate dynamic programming in operations research in the 1990s. Ultimately, all three lines of research came to the same conclusion: approximating Bellman's equation is really hard and generally does not work.<br><br>The lines of research emerging from these communities (computer science using "reinforcement learning", optimal control using "adaptive dynamic programming," and operations research using "approximate dynamic programming") would all grow to embrace a wide range of solution methods to meet the needs of the vast range of problems being addressed. Over a series of articles from 2014 to 2019, I formalized the idea that all RL problems are sequential decision problems: decision, information, decision, information, …, which can be modeled using what I call the "universal modeling framework," a derivative of what is used in optimal control.<br><br>I then realized that every method for making decisions could be organized into four classes of policies, only one of which uses Bellman's equation. Policies in each of the four classes can be found throughout the optimal control literature, as well as in the 2018 edition of Sutton and Barto's book Reinforcement Learning. This laid the foundation for my 2022 volume, Reinforcement Learning and Stochastic Optimization: A unified framework for sequential decisions which is written entirely around the universal modeling framework and all four classes of policies.<br><br>I will demonstrate how this framework identifies approaches for any sequential decision problem, spanning the traditional problems with discrete actions up to the high-dimensional problems faced in operations research. All four classes of policies can be applied to stochastic search problems known as multi-armed bandit problems (I like the name "intelligent trial-and-error"), which are arguably the most common decision problem in practice. |

| 1:30 pm – 3:00 pm: Technical Session 1 | |
|---|---|
| | *Solving Assortment Optimization with First-Order Methods and Neural Networks: A Computational Framework and Public Benchmark*<br>**Ningyuan Chen** (University of Toronto) |
| | *Learning for Assortment Decisions: A Decision-Focused Approach under MNL Model*<br>**Yu Sun** (University of Toronto) |
| | *Return Concerned Adaptive Control of Product Offerings*<br>**Recep Bekci** (University of Waterloo) |
| | *Reinforcement Learning for Cost Estimation in Two-Level Combinatorial Decision Problems*<br>**X.Y. Han** (The University of Chicago Booth School of Business) |

| 3:00 pm – 4:00 pm: Poster Session 1 | |
|---|---|
| | *A Theoretical Framework for Auxiliary-Loss-Free Load Balancing of Sparse Mixture-of-Experts in Large-Scale AI Models*<br>**Fausto Errico** (École de technologie supérieure - CIRRELT and GERAD) |
| | *Climate Risk Exposure and Supply Chain Restructuring: A Large Language Model Approach*<br>**Jiong Sun** (Purdue University), **Leyao Tan** (University of British Columbia) |
| | *Methodological Advances in Human-LLM Decision-Making Research: Interactive AI Experimental Design Protocol and Conversational Analysis*<br>**Tracy Jenkin** (Queen's University) |
| | *Generative Artificial Intelligence for Decision Making in Supply Chain*<br>**Mohammad Moshref-Javadi** (University of Illinois at Urbana Champaign) |

| 4:00 pm – 5:30 pm: Technical Session 2 | |
|---|---|
| | *Balancing Optimism and Pessimism in Offline-to-Online Learning*<br>**Ilbin Lee** (University of Alberta) |
| | *Finite-Sample Analysis of Decentralized Q-Learning for Stochastic Games*<br>**Zuguang Gao** (University of California, Irvine) |
| | *Reinforcement Learning for Reasoning in Time Series Analysis*<br>**Kimia Ghobadi** (Johns Hopkins University) |
| | *Reinforcement Learning for Optimizing Physician Effort Allocation in Emergency Departments*<br>**Parvin Malekzadeh** (University of Toronto) |

**Friday, May 15, 2026**

| 9:00 am – 12:30 pm: Keynote by Amy R. Ward (The University of Chicago Booth School of Business) | |
|---|---|
| | **Learning in Stochastic Models: An Asymptotic Approach**<br>We consider queueing models with reneging in which the primitive distributions are unknown and must be learned. The question of interest is how to prioritize arriving individuals for service. When the reneging distribution is non-exponential, the state space is very complex, because we must track the amount of time each customer in queue has been waiting in order to have a Markovian system. As a result, learning an exact optimal policy is very challenging.<br><br>The approach we take is to first ignore the discrete and stochastic nature of arrivals, and to instead assume that the number of individuals to arrive in a time interval of length $t>0$ is proportional to $t$. Then, we use historical (offline) data to formulate a data-driven (fluid) optimization problem (that may incorporate machine learning predictions) whose objective is to maximize long-run average reward. The solution to the data-driven optimization problem motivates a data-driven policy for deciding how to prioritize arriving individuals for service. With an infinite amount of historical data available, the data-driven policy is asymptotically optimal as the arrival rate and service capacity become large. With a finite amount of historical data available, the challenge is to establish regret bounds and/ or finite sample statistical guarantees. |

| 1:30 pm – 3:00 pm: Technical Session 3 | |
|---|---|
| | *Data-Driven Optimal and Myopic Policies for Inventory Systems with Demand Covariates*<br>**Kevin Shang** (Duke University) |
| | *Dynamic Matching for Refugee Resettlement: A Case Study*<br>**Rad Niazadeh** (The University of Chicago Booth School of Business) |
| | *Exploring the Frontier of AI in Optimization Research*<br>**Michael Wagner** (University of Washington) |
| | *Cloud Value Chains in the Age of AI: An Operations Management Perspective*<br>**Shi Chen** (University of Washington) |

| 3:00 pm – 4:00 pm: Poster Session 2 | |
|---|---|
| | *A Distributed Decision-Making Mechanism Integrating Optimization and Reinforcement Learning for Coordinated Restoration of Interdependent Complex Networks*<br>**M. Hadi Amini** (Florida International University) |
| | *Human Effort in the AI Value Chain: Improving Crowdsourced Data Annotation*<br>**Setareh Farajollahzadeh** (McGill University) |
| | *Selective Decision-Making for Human–AI Systems*<br>**Michael Huang** (CUNY Baruch College) |
| | *History-Dependent Fluid Models for Weakly Coupled MDPs: Improving Transition Feasibility*<br>**Seyedeh Parisa Moosavi** (University of Toronto) |
| | *Adaptive Behaviour of Paramedics: The Heterogeneous Impact of Paramedics Workload on Scene Decisions*<br>**Maryam Zakeri** (University of Alberta) |

| 4:00 pm – 5:30 pm: Technical Session 4 | |
|---|---|
| | *Pricing and Competition for Generative AI*<br>**Rafid Mahmood** (University of Ottawa) |
| | *LLM Pricing*<br>**Hamid Arzani** (University of Toronto) |
| | *Improving Behavioral Alignment in LLM Social Simulations via Context Formation and Navigation*<br>**Qianran (Jenny) Jin** (The Chinese University of Hong Kong) |
| | *Why the Best Machine May Not Be the Best: Incentivizing Human-Machine Collaboration*<br>**Shouqiang Wang** (University of Texas at Dallas) |