KEYNOTE SPEAKERS

Melanie Mitchell

Title: *Why AI is Harder Than We Think*

Abstract: Since its beginning in the 1950s, the field of artificial intelligence has cycled several times between periods of optimistic predictions and massive investment ("AI Spring") and periods of disappointment, loss of confidence, and reduced funding ("AI Winter"). Even with today's seemingly fast pace of AI breakthroughs, the development of long-promised technologies such as self-driving cars, housekeeping robots, and conversational companions has turned out to be much harder than many people expected.

One reason for these repeating cycles is our limited understanding of the nature and complexity of intelligence itself. In this talk I will discuss some fallacies in common assumptions made by AI researchers, which can lead to overconfident predictions about the field. I will also speculate on what is needed for the grand challenge of making AI systems more robust, general, and adaptable—in short, more intelligent.

Speaker Bio: Melanie Mitchell is the Davis Professor of Complexity at the Santa Fe Institute. Her current research focuses on conceptual abstraction, analogy-making, and visual recognition in artificial intelligence systems. Melanie is the author or editor of six books and numerous scholarly papers in the fields of artificial intelligence, cognitive science, and complex systems. Her book Complexity: A Guided Tour (Oxford University Press) won the 2010 Phi Beta Kappa Science Book Award and was named by Amazon.com as one of the ten best science books of 2009. Her latest book is Artificial Intelligence: A Guide for Thinking Humans (Farrar, Straus, and Giroux).

Julie Shah

Title: *Human-Machine Partnerships and Work of the Future*

Speaker Bio: Julie Shah is associate dean of Social and Ethical Responsibilities of Computing at MIT, a Professor of Aeronautics and Astronautics, and director of the Interactive Robotics Group, which aims to imagine the future of work by designing collaborative robot teammates that enhance human capability. She is expanding the use of human cognitive models for artificial intelligence and has translated her work to manufacturing assembly lines, healthcare applications, transportation, and defense. Before joining the faculty, she worked at Boeing Research and Technology on robotics applications for aerospace manufacturing. Prof. Shah has been recognized by the National Science Foundation with a Faculty Early Career Development (CAREER) award and by MIT Technology Review on its 35 Innovators Under 35 list. Her work on industrial human-robot collaboration was also in Technology Review's 2013 list of 10 Breakthrough Technologies. She has received international recognition in the form of best paper awards and nominations from the ACM/IEEE International Conference on Human-Robot Interaction, the American Institute of Aeronautics and Astronautics, the Human Factors and Ergonomics Society, the International Conference on Automated Planning and Scheduling, and the International Symposium on Robotics. She earned degrees in aeronautics and astronautics and in autonomous systems from MIT.

UNIVERSITY OF MICHIGAN SPEAKERS (MICHIGAN AI)

Michał Dereziński

Title: *What is the cost of interpretability in dimensionality reduction?*

Abstract:
Interpretable dimensionality reduction is a core challenge in reasoning about large datasets and machine learning models. It can be used, e.g., to select a representative subset of gene variants from a genetics dataset, or a collection of most informative documents from a text database. When data are represented numerically, they are often described via matrices, in which case linear algebra suggests a natural (and in a certain sense optimal) way of performing dimensionality reduction: find the principal components corresponding to the largest directions of variance. These principal components work well for black-box models that are evaluated only in terms of prediction quality, but they are generally not interpretable in terms of the domain from which the data are drawn. They do not correspond to, say, a particular document or a gene variant, but rather a complex mixture of them. A long-standing challenge has been to find small representations of data which mimic the numerical properties of principal components and which are also interpretable. The added cost of this interpretability can be significant, as shown by prior worst-case analysis. In this talk, I will demonstrate how we can exploit the inherent structure of data to show that, except for pathological examples which we can characterize, the cost of interpretability is far smaller than suggested by prior worst-case analysis, and it is often negligible for real-world problems. To construct the interpretable representations, we used a randomized subset selection technique based on Determinantal Point Processes (DPPs), which provide a probabilistic model of diversity that has emerged across many scientific

domains. I will briefly discuss the key properties of DPPs that make them useful in dimensionality reduction, as well as the efficient algorithms that turned them into a practical tool for large-scale data science.

Benjamin Fish
Title: *The value gap in human-algorithm interaction*

Abstract:
Algorithms and other formal models incorporating human values like fairness have grown increasingly popular in computer science. Yet formal approaches to incorporating values in algorithmic systems rarely seem to guarantee these values will hold in practice, particularly when these systems involve interactions with and guidance by people. In response to this “value gap” between how we design for values and how they emerge in practice, designers and researchers have taken widely divergent positions on how formal models incorporating aspects of human values should be used: encouraging their use, moving away from them, or ignoring the normative consequences altogether. In this talk, I will seek to resolve these divergent positions by identifying the main conceptual limits of formal modeling, and develop four reflexive values--value fidelity, accuracy, value legibility, and value contestation--vital for incorporating human values adequately into formal models. This comes from joint work with Luke Stark.

Odest Chadwicke Jenkins
Title: *Distributed teaching collaboratives for addressing systemic disparities in artificial intelligence*

Abstract:
We can start the fun stuff in AI for first-year undergraduates across higher education. Distributed Teaching Collaboratives is one emerging model to make this a reality. The idea of a Distributed Teaching Collaborative (DTC) is for faculty at multiple universities to collaborate in the teaching of a course. Michigan Robotics is designing a Distributed Teaching Collaboratives model to minimize the overhead for faculty to update a current course offering, create a new course, or improve the efficiency of a course to engage in other scholarly activities (such as for research). This DTC model is also intended to provide focused collaborations between R1 universities and Minority Serving Institutions with clear goals, which can serve as a foundation for larger partnerships and improved pipelines to graduate programs. I will describe our current progress and successes in Michigan Robotics offering DTC courses, such as Computational Linear Algebra (before Calculus) in collaboration with Morehouse College.

Maggie Makar
Title: *Causally motivated shortcut removal using auxiliary labels*

Abstract:
Robustness to certain forms of distribution shift is a key concern in many ML applications. Often, robustness can be formulated as enforcing invariances to particular interventions on the data generating process. I will present a flexible, causally-motivated approach to enforcing such invariances, paying special attention to shortcut learning, where a robust predictor can achieve optimal iid generalization in principle, but instead it relies on spurious correlations or shortcuts in practice. I will present our approach, which uses auxiliary labels, typically available at training time, to enforce conditional independences between the latent factors that determine these labels. I will show theoretical and empirical that arguments causally-motivated regularization schemes (a) lead to more robust estimators that generalize well under distribution shift, and (b) have better finite sample efficiency compared to usual regularization schemes, even in the absence of distribution shifts. The work I will present highlights important theoretical properties of training techniques commonly used in causal inference, fairness, and disentanglement literature.

Lu Wang
Title: *Convince Me If You Can: natural language generation for argumentation*

Abstract:
Understanding, evaluating, and generating arguments are crucial elements of the decision-making and reasoning process. A multitude of arguments and counter-arguments are constructed on a daily basis to persuade and inform us on a wide range of issues. However, constructing persuasive arguments is a challenging task for both human and computers, as it requires credible evidence, rigorous logical reasoning, and sometimes emotional appeals.

In this talk, I will introduce our argument generation framework. It consists of a powerful retrieval system and a novel two-step generation model, where a text planning decoder first decides on the main talking points and argument structure, then a content realization component constructs an informative and fluent argument. Our argument generation framework will enable many compelling applications, including providing balanced perspectives on complex issues, debate coaching, and essay writing tutoring. Our framework is also generic and has been applied to various text generation problems, e.g., writing news and Wikipedia articles.

Xu Wang

Title: *Sourcing student open-ended solutions to create scalable learning opportunities*

Abstract:
A challenge to meet the demand on higher education and professional development is to scale these educational opportunities while maintaining their quality. My work tackles this challenge by harnessing examples from existing resources to enable the creation of scalable and quality educational experiences. I will share insights about developing effective learning at scale systems by leveraging the complementary strengths from peers, experts, and machine intelligence, differentiating it from existing systems that solely rely on machine or crowds of peers.
Specifically, I'll talk about one technique UpGrade, which uses student solution examples to semi-automatically generate multiple-choice questions for deliberate practice of higher order thinking in varying contexts. From experiments in authentic college classrooms, I show that UpGrade helps students gain conceptual understanding more efficiently and helps improve students' authentic task performance.