The 21st IEEE International Conference on Machine Learning and Applications (IEEE ICMLA 2022)

Keynote talk: **Behavior design**

December 12-14, 2022, The Bahamas, Caribbean

https://www.icmla-conference.org/icmla22/

Geoff Gordon

We often need to teach an agent to follow a desired behavior in a given environment — for example when creating an intelligent software assistant, a mobile robot, or a character in a video game. This sort of behavior design problem is harder than supervised learning, for several reasons. First, sequential decisions are harder than decisions made one at a time: we have to deal with compounding errors, distribution shift, exploration, and the influence of function approximation. Second, it can be hard for designers to communicate their intentions, so we may need to refine our behavior iteratively. This sort of iterative refinement means that we need our learner to be legible (the designer can predict how it will act), interpretable (the designer can understand why it acts a given way), and responsive (able to change behavior based on feedback). Finally, we have to handle many different kinds of feedback, such as demonstrations, constraints, comparisons, rewards, and nudges. Despite the difficulties, there are many kinds of tools that can help us with behavior design, including reinforcement learning, imitation learning, model learning, and representation learning. We'll cover several of these tools, and present them through examples.