DISSERTATION DEFENSE

Max Smith

Efficient Game Solving through Transfer Learning

Wednesday, July 26, 2023
4:00pm – 6:00pm
3725 Beyster
Hybrid – Zoom

ABSTRACT: Game-solving with reinforcement learning entails considerable computational cost due to training agents with or against a series of other-agent strategies. Each round of training brings us closer to the game's solution, but training an agent can require data from millions of games played. The cost of game solving reflects the cumulative data cost of repeatedly training agents. This cost is also a result of treating each training as an independent problem. However, these problems share elements that reflect the nature of the game-solving process. These similarities present an opportunity for an agent to transfer learning from previous problems to aid in solving the current problem.

I introduce game-solving algorithms that are based on new methods for transfer learning to capitalize on the shared elements between these problems, reducing costs. I explore two types of transferable knowledge: strategic and world. Strategic knowledge describes knowledge that depends on the other agents. In the simplest case, strategic knowledge may be encapsulated in a policy that was trained to play, with or against, fixed other agents. To facilitate transfer of this kind of strategic knowledge, I propose Q-Mixing, a technique that constructs a policy to play against a distribution of other agents by combining strategic knowledge regarding each agent in the distribution. I provide a practical approximate version of Q-Mixing that features another type of strategic knowledge: a learned belief in the distribution of the other agents. I then develop two game-solving algorithms, Mixed-Oracles and Mixed-Opponents. These algorithms use Q-Mixing to shift the learning focus from interacting with a distribution of other agents to concentrating on a single other agent. This transition results in a significantly easier and, therefore, less costly learning problem. Complementary to strategic knowledge, world knowledge is independent of the other agents. I demonstrate that co-learning a world model along with game solving allows the world model to benefit from more strategically diverse training data. It also renders game solving more affordable through planning. I realize both of these benefits in a new game-solving algorithm Dyna-PSRO. Overall, this dissertation introduces techniques to substantially lessen the cost of game solving.

CHAIR: Prof. Michael P. Wellman