Combinatorial Multi-armed Bandits for Real-Time Strategy Games

Abdullah alGhaznawi, Cairo University

I. REAL-TIME STRATEGY GAMES

Strategy (RTS) games are genre of video games where players compete on a battlefield by gathering resources to build up their bases along with units and defenses in real time. Using AI to design agents capable of effectively playing RTS games against human is challenging due to huge decision and state spaces comparing to e.g. Chess and Go, as well as the huge number of decision cycle and simultaneous actions from all players. The paper experiment with Micro RTS simulator, a simplified, deterministic and fully observable RTS game.

II. MONTE CARLO TREE SEARCH IN RTS GAMES

Monte Carlo Tree Search (MCTS) samples the decision space rather than fully exploring it. It starts from the root node (initial state) selecting an action that leads to one of the children nodes (next state) following a tree policy until it reaches a state not known by the tree policy. It then start to simulate based on some random actions and averaging the score. Finally, it propagates the information back up to the roots to upgrade the reward. When MCTS reaches a certain depth, it selects an action that leads the a child of the root node which has the best score.

III. COMBINATORIAL MULTI-ARMED BANDITS

Combinatorial multi-armed bandits problem definition is as follows: A set of n variables $X = \{X1, ..., Xn\}$, where different values can be assigned from a set of values to Xi. An unknown reward distribution $u : X -> R$ over each combination. A function $L : X -> \{true, false\}$ that determines which macro-arms are legal.