0 Contents

1 Results and Discussion

1.1 Training

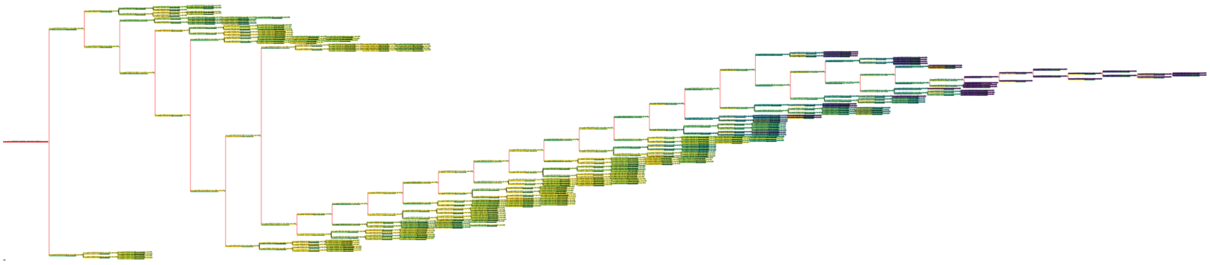

Figure 1.1: An example MCTS tree structure lasting 33 steps with 15 MCTS simulations per step. The neural network has not been trained and is working against an untrained adversary with and average power of 5% of the player. Lighter (yellow) colours represent "better" states.

Figure 1.1 shows the structure of a typical tree search when relying solely on the tree search. This is significantly better than with no tree-search, which typically achieves an episode length of 7.

`Do graph of MCTS sims vs episode length with no training?`

Figure 1.2 shows the max-max nature of the tree search (after negating the action-value for the adversary). also highlights how Nsa is related to V_pred and U, with Nsa almost always being the right one.

1.2 Performance

1.3 Performance Without an Adversary

1.4 Performance Against a Random Opponent

Selected Distributions

Figure 1.2: A zoomed in view of fig. 1.1. The three colours of each node are based on the value of $U = Q(\boldsymbol{x}, u) + ucb$ as in **??**, the number of state-action visits, $N(\boldsymbol{x}, u)$, and the state-cost, c.

do against a constant random opponent do against say off for 30 steps then hit with a +- 5 do against a multinomial distribution with expected value of 0.05? do against

1.5 Performance Against a Trained Adversary

Compare when trained n the adversay against the random opponents? and also against no adversary? Is it robust?