AI Nanodegree Project 2: Build a Game-Playing Agent

Research Review

Social Learning Methods in Board Games

Vukosi N. Marivate and Tshilidzi Marwala

Introduction

The paper **Social Learning Methods in Board Games** reasons and evaluates the benefits of training an agent/agents using ‘non-stationary opponents’. Where ‘non-stationary opponents’ is intended to capture the dynamics of how people learn from each other, taking inspiration from traditional person to person board game clubs where juniors are paired with mentors to accelerate learning. The authors argue that the existing approach, especially playing against itself, doesn’t expose the agent to enough experience (game states).

Reinforcement learning

Reinforcement learning is used in this paper, a machine learning technique that is said to model how humans learn, through trial and error. Simply put, given the state of an environment, we perform an action and observe it’s reward (or punishment) and resulting state. Using the TD-Lambda algorithm (Temporal Difference-Lambda Algorithm), the reward is associated to the given state-action pair Q(s,a) and propagated across the previous steps.

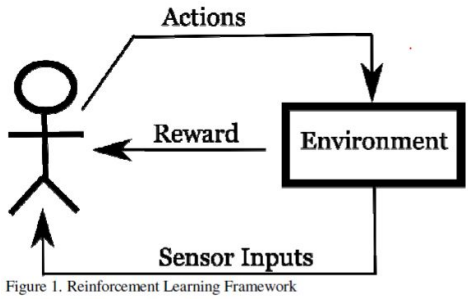

Figure 1. Reinforcement Learning Framework

Social Learning

As mentioned above; the motivation behind this paper is to evaluate the effectiveness of using Social Learning to increase the effectiveness of the agents learning. This is inspired by how humans learn from each other, when playing someone better we observe and learn from them. The example given in the paper is based on an agent competing against a logic based intelligence computer program, using reinforcement learning, the agent would learn a policy that would optimally beat this 'logical' opponent. The goal of this paper is to simulate this phenomenon.

Technique

Two *general* groups of agents are trained, one to simulate social learning and the other used as a control group. The authors used the game of Tic Tac Toe where each player is designated noughts or crosses and the winner is the player who sets three in a row (horizontally, vertically, or diagonally).

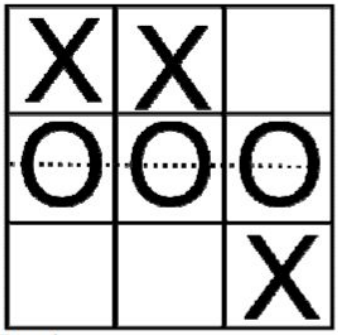

Figure 4. Tic Tac Toe Board

Social learning is simulated by training a set of agents using two game configurations, one using a modified Swiss competition configuration, where losers are pitched against winner and visa-versa, and the other using round-robin. Three groups of agents were trained, each varying in population sizes, including: 4, 6 and 8.

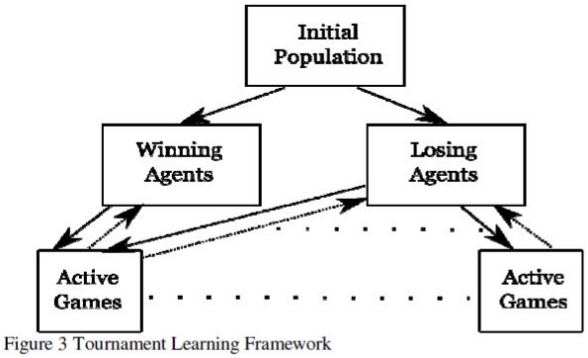

Figure 3 Tournament Learning Framework

The other set was a control group and trained by itself (self-play).

Once trained over 5000 epochs, the agents are pitched against each other in a series of play tests and a full league. Play tests included 10 board configurations with varying difficulty, 5 beginner, 2 intermediate, and 3 hard. Each agent had to choose the right move to score a point. The following figure shows the results of these play tests for each group.

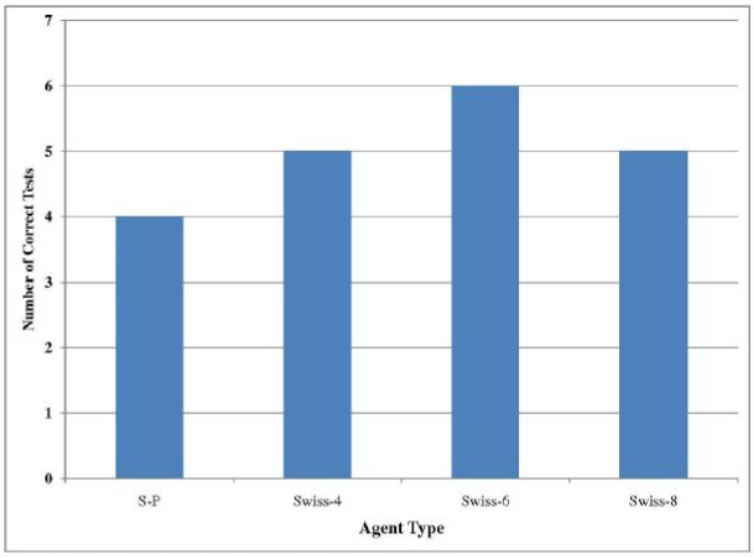

Figure 8. Board test results SP vs. Swiss Self Play

The results show the Self-Playing (SP) agents scored 4 correct showing beginner capabilities, Swiss Social agent 4 (trained with a population of 4) scored 5 and Swiss Social agent (training with a population of 6) scoring 6 correct.

The results of the league showed the social agents on average beat the self-play agents 50 times or more.

Conclusion

The authors acknowledge that all agents played at a beginner level but have showed that introducing non-stationary increases their effectiveness (independent of training cycles) and suggest using larger population sizes is likely to result in supervisor agents.