Quixo Project for CPSC 490 (Project Proposal)

Student: Jau Tung Chan (jc3395)

Advisor: Prof. James Glenn

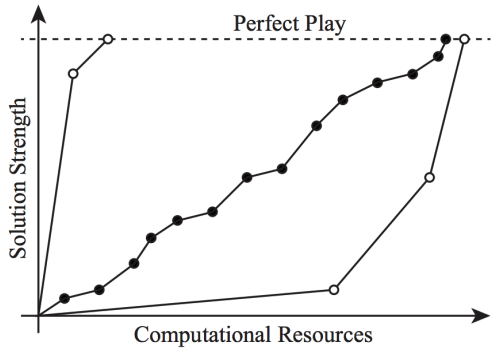
Project Objective

To explore the 'depth' of the game of Quixo, and in particular, including its variants.

Background ('Depth')

The concept of the 'depth' of strategic games is a term that is often used by game designers and players. In loose intuitive terms, this describes how absorbing or profound a game is. As an illustrative first example, chess or Go both have much more 'depth' than tic-tac-toe. Games that are 'deep' in this way typically have large communities of expert players continuously engaging with the game, and yet they still learn more about the game even after playing for a long time.

However, this notion of 'depth' is almost always used informally, and has no formal definition to this date (to my knowledge). In their 2017 paper, Lantz, Isaksen, Jaffe, Nealen, and Togelius explored a semi-formal conceptual model for what 'depth' is and how it can be measured. This conceptual model can be summarized by the following strategy ladder (taken from their paper):



Each strategy ladder is constructed for one particular game (hence three games are represented with the three ladders above). Specifically, each dot on the ladder represents the *best* strategy that can be implemented to play this game, subject to some constraint on computational resources. Here, *best* is measured as solution strength on the y-axis, which can be measured as the strategy's win rate, and/or the quality of their move selection (depending on the game). On the other hand, computational resources on the x-axis can be measured by computational time, amount of pre-computed memory, and/or amount of working memory (depending on the game and the types of strategy being used)[1].

Returning to the three strategy ladders above, the point that Lantz et al. makes is that the middle ladder corresponds to a game with the most 'depth'. Specifically, the ladder on the left corresponds to a game that can be easily solved (e.g. tic-tac-toe), while the ladder on the right corresponds to a game that requires a lot of computational resources to even 'get off the ground' (figuratively). The limited number

[1] These vagueries of the measurements of the two axes is why I view this model as semi-formal rather than formal.

of dots on both of these ladders, compared to the middle ladder, illustrates the limited possibility for players to learn continuously about the game and reach intermediate stages of incremental improvement. In this way, the 'depth' of a game can therefore be measured (or at least approximated) by the number of dots on, and the shape of, its strategy ladder.

Background (Quixo)

Quixo is a board game produced by Gigamic. While the interested reader can visit the provided link for a detailed description of the game rules, here, I will highlight the important properties of Quixo from a computational intelligence perspective.

Firstly, Quixo is a combinatorial game (i.e. it is a two-player game which is turn-based and deterministic, and both players have perfect information about the game tree). Secondly, Quixo is *not* a finite game (i.e. a single game can take arbitrarily many turns). Thirdly, even though Quixo is traditionally played on a 5X5 square board, it is easy to adapt its rules to variants that use a smaller or a larger board (e.g. 3X3, 4X4, 6X6). The crucial relevance of this third property is that the 'depth' of Quixo can be easily adjusted with a board size parameter (specifically, games on larger board sizes are intuitively 'deeper' than those on smaller board sizes), while the core identity of the game itself remains unchanged (or at least, I hope so).

In view of needing to measure the solution strength of different Quixo strategies (i.e. the y-axis on the strategy ladder), it is helpful to have the perfect Quixo player as a benchmark (e.g. for measuring quality of moves). To this end, recently (in 2020), a complete analysis of Quixo was published by Tanaka, Bonnet, Tixeuil, and Tamura. In it, they describe their development of the perfect (computational) Quixo player, using backward induction and value iteration techniques.

The most significant takeaway from Tanaka et al.'s paper is that, using these perfect players, 5X5 Quixo is a draw game (i.e. a game between perfect players never terminates), while 4X4 and 3X3 Quixo are both first-player-win games. This supports the earlier intuition that the 'depth' of Quixo is positively correlated with its board size, since 5X5 Quixo is likely 'deeper' than 4X4 and 3X3 Quixo by this result alone.

It is also worth noting that Tanaka et al.'s complete analysis of 5X5 Quixo took about 2 weeks of single-thread computation time; and it still took about 32 hours of computation time even when using 32 threads in parallel. They (understandably) did not investigate board sizes larger than 5X5, and their code is not publicly available (to my knowledge).

Project Scope

I firstly aim to reproduce the development of the perfect Quixo player based on Tanaka et al.'s paper. Once that is done, I will explore some reinforcement learning techniques (e.g. Q-learning, Monte Carlo tree search) until one of them achieves comparable performance to the perfect player, given reasonable amounts of training (i.e. it does not take too long or require too much computational power).

Then, I will create different versions of this reinforcement learning player, by limiting the amount of its training to varying extents. Using this constrained training as the x-axis of computational resources[2], and

[2] Intentionally vague, since the exact quantification may be tweaked based on empirical results; but I am considering using computational time, or number of training iterations.

using the performance of the resulting player as the y-axis of solution strength[3], I will graph (an approximation of) the strategy ladder of 5X5 Quixo.

Repeating the same process for different board sizes of Quixo will yield (hopefully) different shapes of strategy ladders, and allow for a (at least qualitative) examination of the relationship between 'depth' and the strategy ladder.

Timeline

Finish:	**By:**
Develop the base game engine (including a graphical interface)	Feb 12[4]
Develop the perfect player; optimize computation for time and space	Feb 26
Computation for the perfect player[5]	Mar 5
Initial development of the reinforcement learning player, to decide on the chosen strategy	Mar 26
Develop the reinforcement learning player for experiments (i.e. constrained training)	Apr 2
Initial experiments for 5X5 Quixo, to decide on the specific axes of the strategy ladder	Apr 16
Experiments for different board sizes of Quixo	Apr 30
Tidy up data; write up abstract and final report; create web pages for submission	May 12

Deliverables

- Code base as a GitHub repository, including code for:
  - Base game engine
  - Computation for the perfect player
  - Computation for the reinforcement learning player
  - *[Stretch]* Other players that are empirically interesting, e.g. a heuristic player
  - Experiments for generating the strategy ladders
- Final report, including:
  - Description of choice for the reinforcement learning player strategy
  - Description of choice for the axes of the strategy ladders
  - Strategy ladders of 5X5, 4X4, and 3X3 Quixo
  - *[Stretch]* Other interesting strategy ladders (e.g. with different x- or y- axes)
  - Description and/or analysis of the relationship between 'depth' and the strategy ladders
  - *[Stretch]* Analysis of other empirically interesting results, e.g. performance of a heuristic player

[3] Likewise intentionally vague for the same reason; but I am considering measuring accuracy of move selection compared to the perfect player, or number of moves until it loses (or draws) against the perfect player.
[4] Completed before the submission of this proposal (but after the start of the semester); included for posterity.
[5] Recall that Tanaka et al.'s computation took 2 weeks of single-thread computation time, so this is non-negligible.