

COMP7704 Dissertation

Detailed Dissertation Proposal

Name: Ng Argens

Supervisor(s): Dr. Dirk Schnieders

Dissertation Title: To Play and Cooperate in Imperfect Information Games with
Machine Learning

Planned Submission Semester: 2018-2019 Semester 2

Aim

The goal of my dissertation is to explore the limitation of machine learning with yet another type of game, contract bridge. This game is chosen for having a wide array of features that can be added or removed depending on the research progress, including information hiding and discovery, variable reward, cooperation, as well as other rule modification that complicates or simplifies the game. Ideally, the aim is to solve the game of bridge under the system regulation of WBF, such that the resulting agent would be eligible to compete in World Computer-Bridge Championship in the future.

Brief Literature Review

Deep reinforcement learning has been having great success in gaming in general. In 2013, Google DeepMind developed an agent which outperformed all previous approaches on six of the seven Atari 2600 games with no change in architecture and hyperparameter (except one which happens to coincide with the flash rate of laser in Space Invader) [1]. Then in 2017, Google DeepMind used similar techniques to play Go at a superhuman performance, winning 100-0 against Alpha Go, another program which had defeated a world champion in 2016 [2].

Both successes mark the versatility and power of deep reinforcement learning in computer gaming. However, one field that has yet been extensively researched is multiplayer imperfect information game, in which another technique, counterfactual regret minimization (CFR), has a head start [3, 4]. One drawback of CFR and its variations though, is it requiring the storage of (game state, action) pair values or (information set, action) pair values, subject to the level of abstraction with human intervention.

A brief look at CFR literatures suggests that the main benefit of it being the ability to greatly reduce the storage required in regret minimization. However, most abstractions are still done by human, and the power does not seem comparable to that of neural networks in the game of Go. Therefore, the hope is high that deep q network (DQN) might be able to provide a better solution in games other than two-player zero-sum complete-information games and one-player environment-interaction games.

1. V. Mnih et al., "Playing Atari with deep reinforcement learning", arXiv:1312.5602v1 [cs], Dec. 2013.
2. D. Silver et al., "Mastering the game of Go without human knowledge," Nature, vol. 550, no. 7676, pp. 354–359, Oct. 2017.
3. R. Gibson et al., "Generalized sampling and variance in counterfactual regret minimization", in AAAI-12: Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence, Jul. 2012, Toronto, Canada.
4. N. A. Risk & D. Szafron, "Using counterfactual regret minimization to create competitive multiplayer poker agents", in AAMAS '10: Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems, May 2010, Toronto, Canada.

Proposed Methodology

As mentioned in the aim of the dissertation, contract bridge has the benefit of having many features that can be added or removed to slightly modify the game as the research progresses.

The research will start by creating a double dummy analyzer that can solve no-trump contracts with all cards known to all players, a two-player constant-sum complete-information game. I plan to do it by looking at existing literature, possibly creating a computer program capable of doing so due to the lack of such programs with computer interface.

After that, trump suits would be introduced, followed by contract, which adds complexity to the rules and card-play. Then the suitable infrastructure for training bidding agents would be created, namely knowledge base and logic engine. The former would involve lots of human input for the knowledge representation of bidding and card suits.

Then we would move on to the card play aspect of bridge. With another knowledge base dedicated to information required in card play. We would then train two similar agents to play against each other. To avoid overtraining, we would use random sample of past agents at each iteration. We would focus on the exchange of information between the defending agents and observe the reaction from declarer agent at this stage.

At each stage, we would compare the results with existing programs and experts. We would also lastly invite former and current Hong Kong youth team members to play against the program to both quantitatively and qualitatively evaluate the program.

Milestones

<i>Week</i>	<i>Tasks</i>	<i>Time Period</i>	<i>Learning Hours</i>	<i>Concurrent Activity</i>
1 - 4	Literature review (game theory, deep learning, reinforcement learning)	1/8 – 24/8	101	COMP 7904 Examination. LSK Hall Admission
5	Reserved for Hall Orientation			
6	Obtaining of double dummy analyzer either from past research or self-construction	3/9 – 7/9	40	Nil
7 - 8	Construction of bidding infrastructure (knowledge base, logic engine)	10/9 – 20/9	72	Nil
9 - 10	Training of bidding agent	24/9 – 5/10	80	Nil
9	Competitive multi-agent reinforcement learning of bidding	8/10 – 19/10	80	Nil
12	Comparison with other existing programs in the aspect of bidding	22/10 – 26/10	40	Nil
12	Construction of card-play infrastructure (environment, knowledge base)	29/10 – 9/11	80	Nil
12	Training of Agent	12/11 – 23/11	80	Nil
13 – 15	Comparison with other existing programs in the aspect of card-play	26/11 – 30/11	40	Nil
16 – 17	Writing of dissertation	3/12 – 21/12	120	Nil
			Total: 653	

Deliverables

<i>Items</i>	
1	Website
2	Report
3	Gaming agent
4	Bidding agent
5	Table manager (Environment)
6	Database (for experience replay)
7	Visualizer (for demonstration of agents' play)

Resource Needed (Expected)

<i>Items</i>		<i>Hours Needed</i>	<i>Application Done Before</i>
1	Gridpoint / HPC 2015	~ 30	15/10