UNIVERSITY OF ALBERTA

Prof. Csaba Szepesvári
Department of Computing Science
3-11 Athabasca Hall
Edmonton, Alberta, Canada T6G 2E8

T: (416) 710-9610
F: (780) 492-6393
http://www.ualberta.ca/~szepesva
csaba.szepesvari@ualberta.ca

Report on the PhD thesis "Search in Imperfect Information Games" by Martin Schmid

Csaba Szepesvári

1 Summary

The main topic of Martin Schmid's thesis is the design of efficient and effective algorithms for finding good strategies for imperfect information games.

In imperfect information games the players have partial information about the "state" of the game which they jointly control, while they aim to maximize the reward that they receive. The purpose of algorithm design in the thesis is to find "good" strategies for the players and as the thesis focuses on the two-player case, this means finding approximate Nash equilibria. While the thesis is concerned with games that people play, imperfect information games capture many real world scenarios when multiple agents interact. There are many exampled in economics, control and more broadly engineering (e.g., auctions with asymmetric information, market design, design of networking algorithms, etc.). As such, the problem itself is highly important.

The thesis lists the following major contributions:

(i) Bounding support sizes in imperfect information games

(ii) Factored-observation formalism

(iii) A variance reduced version of Monte Carlo counterfactual regret minimization

(iv) Sound and safe sub-games refinement in large imperfect information games

(v) Formalizing ε-sound search with special attention to online planning

(vi) DeepStack, a program that plays a variant of poker using the methods described in the thesis

(vii) AIVAT, a technique to reduce the variance of payoff estimates

I found that all seven contributions are significant and novel. I found contributions *(iii)-(vi)* to be the most significant of these. Notably, DeepStack is the first program that consistently outperform professional human poker players and has been the basis of considerable future work.

I was very happy with the carefully constructed modular structure of the thesis. In general, the quality of writing is very high.

UNIVERSITY OF ALBERTA

Prof. Csaba Szepesvári
Department of Computing Science
3-11 Athabasca Hall
Edmonton, Alberta, Canada T6G 2E8

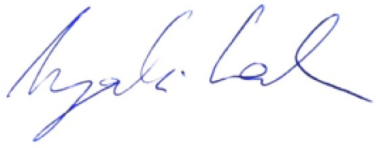
T: (416) 710-9610
F: (780) 492-6393
http://www.ualberta.ca/~szepesva
csaba.szepesvari@ualberta.ca

Overall, the thesis clearly proves the Martin Schmid's ability for creative scientific work.

Some minor questions, comments are as follows:

- I expected to see at least a formal statement on the support size (Section 14.5.1).
- The acronym EFG is used without being introduced first.
- It is quite surprising that lemma 13, which states that perfect recall extensive form games have an equivalent normal form representation, is from 1994. One suspects that Kuhn, who proved his result in 1953, may have also known this already. Bernhard [1992] seems quite relevant here, although this is also from a similar era as the paper cited.
- Concerning Definition 20: Are the rewards part of the observation? That is, O_i is information that player i observes in addition to observing R_i, or is O_i all the information that a player observes? I am guessing that the rewards are not observed (as they are not mentioned as being observed), but this is worth emphasizing as in POMDPs the reward is often by default assumed to be observed. The paper mentioned is also quite relevant for the factored-observation formalism.
- Section 17.8.1: I suspect that the last formula should have $|c|/2$ in it and not $c^2/2$.

20 July, 2021



Csaba Szepesvári

References

P. Bernhard. Information and strategies in dynamic games. *SIAM J. Control Optim.*, 30:212–228, 1992.