

Report on the outcomes of a Short-Term Scientific Mission¹

Action number: CA22145

Grantee name: David Milec

Details of the STSM

Title: Reducing sample complexity in learning games

Start and end date: 06/06/2025 to 14/07/2025

Description of the work carried out during the STSM

Description of the activities carried out during the STSM. Any deviations from the initial working plan shall also be described in this section.

We explored DeepCFR and relevant extensions in literature such as SingleDeepCFR and ESCHER. We explored OpenSpiel with different implementations of DeepCFR and chose one. We picked Kuhn poker for quick small experiments. Leduc poker for larger validation and Liars Dice and Goofspiel for additional testing.

We first chose a neural network architecture that attempted to capture interaction between input features in DeepCFR and we attempted to find subset of features which need to interact together while using only a simple linear or functional layer to combine them. However, the pursuit was unsuccessful as we were not able to find any grouping that would lead to reasonable results in Leduc, showing that we need the interaction between all the features. We were also trying to see how many features need to interact at once and produced strong results with maximal three features interaction (Leduc has 11 features in OpenSpiel), but every feature needs to interact with some other feature. We continued by analysing a converged optimal strategy in Leduc, whether we can infer some groupings from that, but we were unable to do so even in this simple static settings. Therefore, the results from this part were unfortunately mainly negative.

We continued by replacing the network with LGBM optimizer and we observed superior performance in all the games. To pursue the interpretability, we decided to follow what Spyros already used on games with perfect information and introduced multiple optimizers which have the necessary combination of private and public information as input and there is different optimizer for sequence of actions (betting sequence in pokers). We call the new approach public table, and it outperformed both the DeepCFR and

¹ This report is submitted by the grantee to the Action MC for approval and for claiming payment of the awarded grant. The Grant Awarding Coordinator coordinates the evaluation of this report on behalf of the Action MC and instructs the GH for payment of the Grant.

the LGBM improvement in Leduc. Due to time constraints, we were not able to test the approach on other games but there the sequence would be the calls observed in liars dice and the win/loss sequence in Goofspiel.

We were also testing the generalization capabilities of the approaches by leaving a specific states from the observations altogether and testing if the algorithm can still learn to play reasonably well. We observed that public table worsened about three times and got to the level of DeepCFR in the full game while the DeepCFR worsened about five times.

Finally, we prepared necessary tools to get plots from the results so that they can be added to a paper and easily modified all at once.

Description of the STSM main achievements and planned follow-up activities

Description and assessment of whether the STSM achieved its planned goals and expected outcomes, including specific contribution to Action objective and deliverables, or publications resulting from the STSM. Agreed plans for future follow-up collaborations shall also be described in this section.

We showed possible path leading to generalization in game playing settings, while also enhancing interpretability by introducing multiple smaller optimizers which can be implemented as decision trees or splines and directly observed.

We plan to write a technical report of the work done, with detail and we are also adding the results produced to a paper which was already in progress, and we plan to submit to transactions on games during September.

We will present the results in the next meetup of the COST working group for the possibility of future collaboration.

Finally, we would like to continue the evaluation of the public table approach on the other games and also attempt to play full Limit Texas Holdem using the public table approach to test the generalization ability even in larger game and see if we can produce competitive performance with cheaper sampling (due to the smaller optimizer compared to big neural network) and interpretability we are given by multiple smaller optimizers.