

An english description of the phd topic is presented below.

Titre : Algorithmes de recherche dans les arbres : au-delà de Monte-Carlo Tree Search

Encadrants

Encadrants universitaires

- Bruno Scherrer, HDR, IECL, INRIA (Larsen), <https://scherrer.perso.math.cnrs.fr/index.html>, <mailto:bruno.scherrer@inria.fr>
- Amine Boumaza, LORIA, INRIA (Larsen), <https://members.loria.fr/ABoumaza/>, <mailto:amine.boumaza@loria.fr>

IECL, LORIA, Centre INRIA de l'université de Lorraine.

Collaborateurs

- Michal Valko, Google DeepMind / INRIA / enseignant au MVA ENS PS, <https://misovalko.github.io/>, michal.valko@gmail.com
- Pierre Ménard, postdoctoral researcher at Institut für Mathematische Stochastik (IMST) in the Otto-von-Guericke-Universität Magdeburg, <https://menardpr.github.io/>, pierre.menard.public@gmail.com

Problématique

Ces dernières années, le succès d'AlphaZero dans la maîtrise de jeux complexes tels que les échecs et le Go a démontré la puissance de l'apprentissage par renforcement couplé à des algorithmes de recherche dans les arbres (de type MCTS) et des approximateurs de fonction puissants (réseaux de neurones profonds). Cependant, même avec ces réalisations remarquables, il reste des questions théoriques intrigantes insuffisamment explorées. Cette thèse concerne l'étude des fondements théoriques des algorithmes de jeu, en se concentrant spécifiquement sur les jeux déterministes à deux joueurs. Les principaux objectifs comprennent l'amélioration des algorithmes existants de recherche dans les arbres, le développement de modèles probabilistes de jeux déterministes plus sophistiqués que l'état de l'art, les algorithmes d'inférence pour ces modèles, et le développement de benchmarks pour ces problèmes.

Sujet

La première voie d'exploration consiste à faire progresser les **algorithmes de recherche dans les arbres** en particulier ceux concernant des **arbres MIN-MAX adaptés aux jeux déterministes** à deux joueurs. En s'appuyant sur des algorithmes établis comme A^* [14], AlphaBeta [7] et MTD(f) [13], cette recherche vise à proposer des algorithmes de recherche nouveaux. Les améliorations potentielles peuvent inclure des stratégies d'élagage optimisées et des approches innovantes de l'exploration de l'espace de recherche comme l'utilisation d'hypothèses de régularité de l'approche optimiste (algorithmes DOO/SOO [10, 11, 2], SEQUOOL [1]). L'analyse théorique se penchera sur les propriétés mathématiques des algorithmes proposés, en établissant leurs forces et leurs limites.

Le deuxième axe de recherche se concentre sur le fait d'identifier et analyser des **modèles probabilistes de jeux déterministes**, qui soient moins triviaux que le modèle de Pearl [12]. En

développant des modèles plus sophistiqués, l'objectif est de fournir une représentation plus riche de la complexité inhérente aux jeux déterministes. Cela inclut l'incorporation d'éléments tels que la dépendance par le biais d'ancêtres d'arbres communs [3, 6, 4] afin de mieux refléter les scénarios du monde réel. L'analyse théorique examinera les algorithmes d'inférence pour mettre à jour les paramètres du modèle en fonction des observations acquises pendant la recherche et leur *qualité de modélisation* pour un ensemble varié de jeux déterministes.

Sur la base de ces modèles probabilistes, un troisième objectif est d'étudier la **complexité moyenne** correspondante de plusieurs algorithmes de pointe, d'explorer des stratégies alternatives pour la prise de décision dans les jeux déterministes. Les connaissances théoriques acquises contribueront à une meilleure compréhension de la manière dont les modèles probabilistes peuvent guider les processus de décision dans le contexte des jeux (pour l'instant, la plupart de ces algorithmes sont comparés sur le modèle de Pearl [12] et sont essentiellement indifférentiables... étant tous optimaux dans ce cadre).

Une dernière dimension de cette recherche, qui est transversale à toutes les dimensions précédentes, implique la constitution d'**une base benchmark de fonctions et jeux** (dans un esprit similaire aux bases [9, 5] pour l'optimisation de fonctions). On pourra s'appuyer sur la base [8]. Des problèmes de référence soigneusement sélectionnés fourniront une évaluation complète des méthodes proposées. Pour les jeux, on s'attachera à identifier les faiblesses potentielles d'AlphaZero et comment les algorithmes que nous proposerons peuvent exploiter ces vulnérabilités.

Enfin, selon l'avancement sur les dimensions ci-dessus, des perspectives du travail pourront inclure le cas de problèmes avec information partielle, ou la mise en œuvre dans un schéma de programmation dynamique approchée similaire à AlphaZero.

Title : Tree Search Algorithms : Beyond Monte Carlo Tree Search

Supervisors

Academic supervisors

- Bruno Scherrer, HDR, IECL, INRIA (Larsen), <https://scherrer.perso.math.cnrs.fr/index.html>, <mailto:bruno.scherrer@inria.fr>
- Amine Boumaza, LORIA, INRIA (Larsen), <https://members.loria.fr/ABoumaza/>, <mailto:amine.boumaza@loria.fr>

IECL, LORIA, INRIA Center of Université de Lorraine.

Collaborators

- Michal Valko, Google DeepMind / INRIA / MVA ENS PS lecturer, <https://misovalko.github.io/>, michal.valko@gmail.com
- Pierre Ménard, postdoctoral researcher at Institut für Mathematische Stochastik (IMST) in the Otto-von-Guericke-Universität Magdeburg, <https://menardprr.github.io/>, pierre.menard.public@gmail.com

Problem

In recent years, AlphaZero's success in mastering complex games such as chess and Go has demonstrated the power of reinforcement learning coupled with tree search algorithms (such as MCTS) and powerful function approximators (deep neural networks). However, even with these

remarkable achievements, intriguing theoretical questions remain unexplored. This thesis concerns the study of the theoretical foundations of game algorithms, focusing specifically on deterministic two-player games. The main objectives include improving existing tree search algorithms, the development of more sophisticated probabilistic models of deterministic game, the corresponding inference algorithms, and the development of benchmarks.

Phd topic

The first avenue of exploration is to advance the **tree search algorithm** literature, in particular those concerning **MIN-MAX trees adapted to two-player games**. Building on established algorithms such as A^* [14], AlphaBeta [7] and MTD(f) [13], The aim of this research is to propose new search algorithms. Potential improvements include optimized pruning strategies and innovative approaches to the exploration of the search space such as the use of optimistic approach through regularity assumptions (DOO/SOO [10, 11, 2], SEQUOOL [1]). The theoretical analysis will focus on the mathematical properties of the proposed algorithms, establishing their strengths and limitations.

The second line of research focuses on identifying and analyzing probabilistic models of deterministic games, which are less trivial than Pearl’s model [12]. By developing more sophisticated models, the aim is to provide a richer representation of the complexity inherent in deterministic games. This includes incorporating elements such as dependency through common tree ancestors [3, 6, 4] to better reflect real-world scenarios. Theoretical analysis will examine inference algorithms to update model parameters based on observations acquired during the research and their *modeling quality for a varied set of deterministic games*.

On the basis of these probabilistic models, a third objective is to study the corresponding **average complexity** of several state-of-the-art algorithms, to explore alternative strategies for decision decision making in deterministic games. The theory acquired will contribute to a better understanding of how probabilistic models can guide decision-making processes in the context of games (at present, most of these algorithms are compared on the Pearl model and are essentially indifferentiable... being all optimal in this framework).

A final dimension of this research, which cuts across all the previous dimensions, involves the constitution of a **benchmark database for functions and games** (in a spirit similar to what is done for function approximation [9, 5]). Carefully selected problems will provide a comprehensive evaluation of the proposed methods. For games, the focus will be on identifying AlphaZero’s potential weaknesses, and propose algorithms that may exploit these vulnerabilities.

Finally, depending on the progress made on the above dimensions, future perspectives of the work may include the case of problems with partial information, or the implementation of a dynamic programming scheme similar to AlphaZero.

Références

- [1] P. L. Bartlett, V. Gabillon, and M. Valko. A simple parameter-free and adaptive approach to optimization under a minimal local smoothness assumption. In A. Garivier and S. Kale, editors, *Algorithmic Learning Theory, ALT 2019, 22-24 March 2019, Chicago, Illinois, USA*, volume 98 of *Proceedings of Machine Learning Research*, pages 184–206. PMLR, 2019.

- [2] L. Busoniu, R. Munos, and E. Páll. An analysis of optimistic, best-first search for minimax sequential decision making. In *Adaptive Dynamic Programming and Reinforcement Learning (ADPRL), 2014 IEEE Symposium on*, page 1–8. IEEE, IEEE, 2014.
- [3] L. Devroye and O. Kamoun. Random minimax game trees. In D. Aldous and R. Pemantle, editors, *Random Discrete Structures*, pages 55–80, New York, NY, 1996. Springer New York.
- [4] J. Grosse, C. Zhang, and P. Hennig. Probabilistic dag search. In C. de Campos and M. H. Maathuis, editors, *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, volume 161 of *Proceedings of Machine Learning Research*, pages 1424–1433. PMLR, 27–30 Jul 2021.
- [5] N. Hansen, D. Brockhoff, O. Mersmann, T. Tutar, D. Tutar, O. A. ElHara, P. R. Sampaio, A. Atamna, K. Varelas, U. Batu, D. M. Nguyen, F. Matzner, and A. Auger. COmparing Continuous Optimizers : numbbo/COCO on Github, Mar. 2019.
- [6] P. Hennig, D. Stern, and T. Graepel. Coherent inference on optimal play in game trees. In Y. W. Teh and M. Titterton, editors, *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, volume 9 of *Proceedings of Machine Learning Research*, pages 326–333, Chia Laguna Resort, Sardinia, Italy, 13–15 May 2010. PMLR.
- [7] D. E. Knuth and R. W. Moore. An analysis of alpha-beta pruning. *Artificial Intelligence*, 6(4) :293–326, 1975.
- [8] M. Lanctot, E. Lockhart, J.-B. Lespiau, V. Zambaldi, S. Upadhyay, J. Pérolat, S. Srinivasan, F. Timbers, K. Tuyls, S. Omidshafiei, D. Hennes, D. Morrill, P. Muller, T. Ewalds, R. Faulkner, J. Kramár, B. D. Vyllder, B. Saeta, J. Bradbury, D. Ding, S. Borgeaud, M. Lai, J. Schrittwieser, T. Anthony, E. Hughes, I. Danihelka, and J. Ryan-Davis. OpenSpiel : A framework for reinforcement learning in games. *CoRR*, abs/1908.09453, 2019.
- [9] W. Li, H. Li, J. Honorio, and Q. Song. Pyxab – a python library for \mathcal{X} -armed bandit and online blackbox optimization algorithms, 2023.
- [10] R. Munos. Optimistic optimization of a deterministic function without the knowledge of its smoothness. In J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 24. Curran Associates, Inc., 2011.
- [11] R. Munos. From bandits to monte-carlo tree search : The optimistic principle applied to optimization and planning. *Foundations and Trends® in Machine Learning*, 7(1) :1–129, 2014.
- [12] J. Pearl. Asymptotic properties of minimax trees and game-searching procedures. *Artificial Intelligence*, 14(2) :113–138, 1980.
- [13] A. Plaat, J. Schaeffer, W. Pijls, and A. de Bruin. Best-first fixed-depth minimax algorithms. *Artificial Intelligence*, 87(1) :255–293, 1996.
- [14] S. Russell and P. Norvig. *Artificial Intelligence : A Modern Approach*. Prentice Hall, 3 edition, 2010.