

Vers une Approche Polyvalente pour les Jeux à Information Imparfaite sans Connaissance de Domaine

Achille Morenville, Éric Piette
ICTEAM, UCLouvain

{achille.morenville, eric.piette}@uclouvain.be

Résumé

Les jeux à information imparfaite représentent un défi majeur pour les agents de General Game Playing (GGP). Pour y remédier, nous proposons d'étendre Ludii, un système GGP état de l'art, aux jeux aux informations cachées, en adoptant un formalisme et une modélisation plus proches de l'apprentissage par renforcement, et en adaptant les algorithmes de recherche pour une meilleure exploitation de la description du jeu. Ces améliorations permettront la simulation et l'analyse d'une large gamme de jeux à information imparfaite.

Mots-clés

General Game Playing, Jeux à Information Imparfaite, Apprentissage par Renforcement

Abstract

Games with imperfect information represent a major challenge for General Game Playing (GGP) agents. To address this, we propose to extend Ludii, a state-of-the-art GGP system, by improving the game description to include hidden information, adopting a formalism and modelling closer to reinforcement learning, and adapting the search algorithms to make better use of the game description. These improvements will make it possible to simulate and analyse a wide range of games with imperfect information.

Keywords

General Game Playing, Imperfect Information, Reinforcement Learning

1 Introduction

Les jeux de société, appréciés depuis des millénaires, offrent une riche histoire et une grande diversité de mécanismes ludiques. Depuis les débuts de l'intelligence artificielle (IA), ils ont été perçus comme une source d'opportunités pour développer des algorithmes innovants et des agents performants. Le *General Game Playing* (GGP) [10] représente l'un de ces champs de recherche. Ce domaine vise à créer des agents capables de s'adapter à un large éventail de jeux, contrairement aux agents spécifiquement conçus pour un jeu unique. Ces agents généraux sont essentiels pour relever les défis du monde réel où l'environnement n'est pas toujours prédictible, tels que dans les do-

maines de la robotique, de la conduite autonome ou même pour l'étude de notre passé [8].

Cependant, malgré des avancées récentes en GGP [14], les jeux à information imparfaite, où des éléments cruciaux sont cachés aux joueurs, représentent encore un défi majeur. Les représentations actuelles de ces jeux font face à des contraintes importantes, limitant leur applicabilité. Elles sont souvent conçues pour des genres spécifiques ou reposent sur des paradigmes complexes, ce qui entrave la modélisation pratique de centaines de jeux différents.

Cet article explore les limites des méthodes actuelles pour les jeux à information imparfaite. Il propose également des pistes de recherche prometteuses pour créer des algorithmes universels, sans connaissance spécifique du domaine, dans l'optique d'atteindre le niveau moyen des joueurs humains.

2 Travaux antérieurs

Dans le cadre des jeux à information imparfaite, les approches actuelles s'appuient souvent sur une connaissance approfondie du domaine, telle que les stratégies de jeu, les tactiques et les hypothèses structurelles, pour modéliser les jeux avec précision et efficacité. Cette dépendance limite leur applicabilité à un nombre restreint de jeux soigneusement analysés, tels que le poker, le Stratego ou le bridge. De plus, les incertitudes dues à l'observabilité partielle nécessitent de modéliser et identifier les états efficacement afin de créer des agents performants.

Ces approches ont obtenu des succès notables dans divers jeux à information imparfaite, avec des agents comme DeepStack [6] et Pluribus [2] surpassant des joueurs de poker professionnels humains dans différentes variantes du Texas Hold'em. Des algorithmes d'apprentissage profond ont également été développés pour des jeux non liés aux cartes, tels que le Stratego ou le Kriegspiel [1]. Cependant, leur adaptabilité à différents jeux pose des défis significatifs, limitant leur applicabilité à un plus large éventail de jeux à information imparfaite. Des frameworks tels qu'OpenSpiel [4] visent à appliquer une grande suite d'agents à un grand nombre de jeux, mais ils exigent l'implémentation de plusieurs fonctions spécifiques pour chaque jeu à information imparfaite, ce qui peut être contraignant pour l'adaptation à de nouveaux jeux.

Les approches basées sur GDL-II [15], telles que l'algorithme HyperPlay-II [11], peinent à inférer automatique-

ment les connaissances nécessaires et manquent d'intégration avec les algorithmes d'apprentissage. De plus, leur complexité limite leur application à des jeux peu complexes.

Ludii [9], développé dans le cadre du projet ERC-Digital Ludeme¹, est un système GGP offrant un accès à une vaste gamme de jeux. Son approche ludémique permet de modéliser de manière complète et courte les jeux à information parfaite. Un article récent [12] a démontré l'universalité de Ludii, indiquant qu'il serait théoriquement possible de décrire n'importe quel jeu à information imparfaite dans Ludii avec certaines extensions.

3 Ludii-II

Cette section présente une série de propositions visant à étendre les fonctionnalités de Ludii pour prendre en charge les jeux à information imparfaite (Ludii-II). Pour atteindre cet objectif, trois axes d'amélioration ont été identifiés : la description des jeux, la modélisation et le formalisme utilisé pour représenter les jeux et les agents qui y jouent.

3.1 Description des jeux

Les jeux Ludii sont décrits à l'aide de ludèmes, qui sont des unités conceptuelles d'informations relatives au jeu [13]. Les ludèmes capturent les concepts clés du jeu et leur donnent des étiquettes explicites. Cette approche ludémique facilite la compréhension et la description des jeux, permettant l'utilisation de ce système par le plus grand nombre, en proposant un langage proche des mécanismes de jeu.

Toutefois, pour décrire les informations cachées à chaque joueur, Ludii utilise un ludème spécifique permettant de rendre visible ou de cacher à un joueur une information dans la description du jeu [7]. Cette approche, similaire à celle utilisée dans GDL-II, souffre de plusieurs inconvénients. Comme chaque information cachée doit être spécifiée séparément, la description des jeux devient dense et longue, ce qui complique la lecture et la compréhension. De plus, cette méthode manque de généralité, car elle nécessite une expression explicite des informations cachées pour chaque jeu, même s'ils peuvent être similaires. Cela expose également la représentation interne du système, ce qui éloigne la description des règles de jeu.

Une solution consiste à ce que les informations cachées soient implicitement décrites au travers des ludèmes dédiés aux mécanismes les plus communs dans les jeux à information imparfaite. Par exemple, pour la distribution de cartes aux joueurs, le ludème correspondant pourrait tout simplement cacher implicitement l'information aux joueurs qui ne reçoivent pas la carte. Cette approche permettrait de décrire les jeux de manière plus concise et flexible, tout en masquant les détails internes du modèle.

3.2 Modélisation et formalisme

Ludii, initialement conçu pour les jeux à information parfaite, utilise le formalisme des jeux à forme extensive (EFG) pour représenter les jeux sous forme d'arbre, où les nœuds

représentent les états du jeu et les branches représentent les actions des joueurs. Ce formalisme peut être étendu aux jeux à information imparfaite en incluant la notion d'ensembles d'états, qui regroupent les états qui ne sont pas distinguables par un joueur en raison des informations cachées. Bien que cette approche permettrait théoriquement de représenter tous les jeux à information imparfaite dans Ludii [12], elle ne permet pas de représenter toutes les informations disponibles pour les joueurs.

En effet, EFG ne permet pas de stocker dans les états les informations qui ont été reçues, qui les a reçues et à quel moment, ce qui est crucial pour les agents jouant à des jeux à information imparfaite. De plus, pour jouer efficacement, il faut pouvoir prendre en compte quelles actions les opposants pourraient prendre et donc savoir quelles informations ils possèdent. Or, bien qu'il soit possible d'étendre EFG aux ensembles d'information, il n'est pas possible d'extraire ces notions depuis ce formalisme. Par conséquent, les agents représentant leurs connaissances sous forme EFG doivent utiliser des connaissances spécifiques aux jeux plutôt que de s'appuyer uniquement sur la représentation utilisée.

Kovavrik et al. [3] proposent d'utiliser un formalisme plus proche de ce qui est utilisé dans le domaine de l'apprentissage par renforcement à multi-agents, qui est basé sur la notion d'observation : le *Factored-Observation Stochastic Games* (FOSG). FOSG stocke les observations que les agents reçoivent de l'environnement dans une séquence publique et des séquences privées pour chaque joueur, représentant ainsi toutes les informations manquantes dans EFG. De plus, une représentation EFG peut être extraite à partir d'une FOSG, permettant d'appliquer facilement les algorithmes d'apprentissage par renforcement et les résultats théoriques issus du formalisme EFG [3].

L'intégration de l'information imparfaite nécessite une refonte de la communication au sein de l'implémentation. Une approche consiste à ajouter un intermédiaire entre l'état et l'agent, qui serait responsable de la communication entre l'état et l'agent et de filtrer les informations pour ne fournir que les informations nécessaires à l'agent. Cette approche permettrait également d'éviter que les agents ne puissent obtenir des informations en demandant simplement si une action est valide ou non, contrairement à l'implémentation actuelle de Ludii.

3.3 Agents

Les méthodes hors-ligne comme CFR [16] et ses variantes, bien qu'elles convergent vers un équilibre de Nash, présentent des limites en termes de rapidité et d'applicabilité dans le domaine du GGP. Ceci s'explique par le temps de calcul important nécessaire avant de pouvoir jouer au jeu. De plus, leur dépendance à l'état et aux informations fournies restreint leur efficacité lorsque peu d'informations sont disponibles et qu'aucune connaissance humaine n'est ajoutée à la représentation de l'état. En revanche, les méthodes de recherche parfaite appliquées aux jeux à information imparfaite, telles que PIMC [5], ont démontré de bons résultats sur certains jeux, malgré leurs déficiences théoriques. Dans le contexte du GGP, où l'objectif n'est pas néces-

1. Digital Ludeme Project : <https://ludeme.eu/>

sairement d’obtenir une stratégie optimale, il pourrait être intéressant de se tourner vers des approches qui ne se limitent pas à la recherche de l’équilibre de Nash. Une approche prometteuse consisterait à adapter les techniques de recherche en information parfaite, pour tirer parti de la description du jeu fournie et inférer des informations à partir des règles du jeu. Cette méthode, combinée à un formalisme basé sur les observations, offre une direction intéressante pour le développement d’agents performants dans le domaine du GGP.

4 Discussion

Dans cet article, nous avons exploré les limites des méthodes actuelles pour les jeux à information imparfaite et proposé des pistes pour créer des algorithmes généraux dans le cadre du General Game Playing (GGP). En étendant Ludii, un système GGP état de l’art, nous proposons d’améliorer la description des jeux, d’adopter un formalisme et une modélisation plus proches de l’apprentissage par renforcement, et d’adapter les algorithmes de recherche pour mieux exploiter la description du jeu.

Ces améliorations permettraient à Ludii de simuler et d’analyser une large gamme de jeux à information imparfaite, tout en assurant une intégration transparente avec les concepts fondamentaux de Ludii. En outre, l’adoption d’un formalisme plus proche de l’apprentissage par renforcement ouvrirait la voie à l’utilisation d’un plus grand nombre de méthodes d’apprentissage pour les agents.

Cette approche permettrait d’élargir l’analyse et la simulation d’une grande variété de jeux à information imparfaite, contribuant ainsi à l’avancement du domaine du GGP et à une meilleure compréhension des jeux.

Remerciements

Cet article est basé sur les travaux de l’Action COST CA22145 - GameTable, soutenue par COST (European Cooperation in Science and Technology).

Références

- [1] N. Brown, A. Bakhtin, A. Lerer, and Q. Gong. Combining deep reinforcement learning and search for imperfect-information games. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*, volume 33, pages 17057–17069. Curran Associates, Inc., 2020.
- [2] N. Brown and T. Sandholm. Superhuman AI for multiplayer poker. *Science*, 365(6456) :885–890, 2019.
- [3] V. Kovář, M. Schmid, N. Burch, M. Bowling, and V. Lisỳ. Rethinking formal models of partially observable multiagent decision making. *Artificial Intelligence*, 303 :103645, 2022.
- [4] 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. De Vylder, 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. <https://arxiv.org/abs/1908.09453>, 2019.
- [5] J. Long, N. Sturtevant, M. Buro, and T. Furtak. Understanding the success of perfect information monte carlo sampling in game tree search. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 24, pages 134–140, 2010.
- [6] M. Moravčík, M. Schmid, N. Burch, V. Lisỳ, D. Morrill, N. Bard, T. Davis, K. Waugh, M. Johanson, and M. Bowling. Deepstack : Expert-level artificial intelligence in Heads-up No-limit Poker. *Science*, 356(6337) :508–513, 2017.
- [7] É. Piette, C. Browne, and D. J. N. J. Soemers. Ludii game logic guide. <https://arxiv.org/abs/2101.02120>, 2021.
- [8] É. Piette, W. Crist, D. J. N. J. Soemers, L. Rougetet, S. Courts, T. Penn, and A. Morenville. Gametable cost action : Kickoff report. *ICGA Journal*, 46(1) :1–17, 2024.
- [9] É. Piette, D. J. N. J. Soemers, M. Stephenson, C. F. Sironi, M. H. M. Winands, and C. Browne. Ludii – the ludemic general game system. In G. De Giacomo, A. Catala, B. Dilkina, M. Milano, S. Barro, A. Bugarián, and J. Lang, editors, *Proceedings of the 24th European Conference on Artificial Intelligence (ECAI 2020)*, volume 325 of *Frontiers in Artificial Intelligence and Applications*, pages 411–418. IOS Press, 2020.
- [10] J. Pitrat. Realization of a general game-playing program. In *IFIP Congress (2)*, pages 1570–1574, 1968.
- [11] M. Schofield and M. Thielscher. General game playing with imperfect information. *Journal of Artificial Intelligence Research*, 66 :901–935, 2019.
- [12] D. J. N. J. Soemers, É. Piette, M. Stephenson, and C. Browne. The Ludii game description language is universal. <https://arxiv.org/abs/2205.00451>, 2022.
- [13] M. Stephenson, É. Piette, D. J. N. J. Soemers, and C. Browne. An overview of the ludii general game system. In *Proceedings of the 2019 IEEE Conference on Games*, pages 864–865. IEEE, 2019.
- [14] M. Świechowski, H. Park, J. Mańdziuk, and K.-J. Kim. Recent advances in general game playing. *The Scientific World Journal*, 2015, 2015.
- [15] M. Thielscher. A general game description language for incomplete information games. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*, pages 994–999. AAAI, 2010.
- [16] M. Zinkevich, M. Johanson, M. Bowling, and C. Piccione. Regret minimization in games with incomplete information. *Advances in neural information processing systems*, 20, 2007.