

Sujet de stage M2 Recherche

Sujet du stage	Application des techniques d'AlphaZero à la résolution de problème de planification automatique
Encadrants	Damien Pellier, Humbert Fiorino, Jean-Pierre Chevallet

Durée : 5 mois

Laboratoire : Laboratoire d'Informatique de Grenoble, 700 avenue Centrale, 38058 Grenoble cedex 9

Mots clé : Intelligence artificielle, deep learning, jeux, planification automatique

1. Contexte

Deep Mind a développé en 2018 un programme appelé AlphaZero [1] qui est une version généraliste d'AlphaGo Zero [2], un logiciel de go qui a été adapté pour jouer aux échecs et au shogi (échecs japonais). Selon DeepMind, AlphaZero a atteint en 24 heures un niveau de jeu supérieur aux humains au jeu d'échecs, au shogi et au go en battant les programmes champions du monde Stockfish (échecs), Elmo (shogi) et la version d'AlphaGo Zero ayant eu trois jours d'apprentissage. AlphaZero possède désormais un réseau de neurones mis à jour de manière continue et possède des règles codées pour l'établissement de recherche d'hyperparamètres qui lui permet d'explorer de manière efficace un espace de recherche très grand. Ces résultats sont impressionnants mais se limitent à des problèmes relativement simples tels que les jeux.

2. Objectifs

L'objectif de ce sujet de tenter d'appliquer les techniques développées dans le cadre d'AlphaZero à la résolution de problèmes plus généraux de planification automatique [3]. La planification automatique (automated planning en anglais) ou plus simplement planification, vise à développer des algorithmes pour produire des plans, typiquement pour une exécution par un robot. Les logiciels de planification qui incorporent ces algorithmes s'appellent des *planificateurs*. La difficulté du problème de planification dépend des hypothèses de simplification que l'on tient pour acquises par exemple : un temps atomique, un temps déterministe, une observabilité complète, etc. Un planificateur typique manipule trois entrées décrites dans un langage formel (PDDL) qui utilise des prédicats logiques :

- une description de l'état initial d'un monde,
- une description d'un but à atteindre et
- un ensemble d'actions possibles (parfois appelés opérateurs).

Chaque action est spécifiée par des préconditions qui doivent être satisfaites dans l'état actuel pour qu'elle puisse être appliquée, et des postconditions (effets sur l'état actuel). L'espace de recherche à parcourir pour trouver un plan solution est très grand et repose généralement sur l'utilisation d'une fonction heuristique *ad hoc*.

L'objectif du stage est alors d'étudier et de proposer des algorithmes inspirées par AlphaZero pour apprendre ces fonctions heuristiques de manière automatique en utilisant des techniques fondées sur les réseaux de neurones.

3. Profil du candidat recherché

Le candidat devra :

- être inscrit en Master 2;
- avoir des compétences avancées en programmation (conception et implémentation), notamment en Java, et python;
- avoir un bon niveau académique attestant de sa capacité à allier pratique et théorie;
- posséder un niveau d'anglais professionnel oral et écrit;
- posséder des connaissances générales dans le domaine de l'intelligence artificielle est un plus

4. Procédure de contact

Envoyer à Damien.Pellier@imag.fr, Humbert.Fiorino@imag.fr et Jean-Pierre.Chevallet@imag.fr :

- Vos notes de master 1
- Votre CV

Les candidatures sont gérées au fil de l'eau. Vous serez prévenu rapidement par mail de la recevabilité de votre candidature et si vous êtes invité à un premier entretien.

5. Références

[1] <https://www.youtube.com/watch?v=MgowR4pq3e8>

[2] Silver, David; Huang, Aja; Maddison, Chris J.; Guez, Arthur; Sifre, Laurent; Driessche, George van den; Schrittwieser, Julian; Antonoglou, Ioannis; Panneershelvam, Veda; Lanctot, Marc; Dieleman, Sander; Grewe, Dominik; Nham, John; Kalchbrenner, Nal; Sutskever, Ilya; Lillicrap, Timothy; Leach, Madeleine; Kavukcuoglu, Koray; Graepel, Thore; Hassabis, Demis (28 January 2016). "Mastering the game of Go with deep neural networks and tree search". *Nature*. 529 (7587): 484–489.

[3] Damien Pellier, Humbert Fiorino: PDDL4J: a planning domain description library for java. *J. Exp. Theor. Artif. Intell.* 30(1): 143-176 (2018)

Internship subject M2 Research

Internship subject	Application of AlphaZero techniques to solve automated planning problems
Supervisors	Damien Pellier, Humbert Fiorino, Jean-Pierre Chevillet

Duration : 5 mois

Research laboratory : Laboratoire d'Informatique de Grenoble, 700 avenue Centrale, 38058 Grenoble cedex 9

Keywords : Artificial intelligence, deep learning, games, automated planning

1. Context

Deep Mind developed in 2018 a program called AlphaZero [1] which is a generalist version of AlphaGo Zero [2], a go software that has been adapted to play chess and shogi (Japanese chess). According to DeepMind, AlphaZero reached a higher level of play than humans in chess, shogi and go in 24 hours by beating the world champion programs Stockfish (chess), Elmo (shogi) and the AlphaGo Zero version that had three days of learning. AlphaZero now has a continuously updated neural network and coded rules for the hyperparameter search establishment that allow it to efficiently explore a very large search space. These results may be impressive, but they are limited to relatively simple problems such as games.

2. Objectives

The objective of this topic is to attempt to apply the techniques developed in AlphaZero to solve more general automated planning problems [3]. Automated planning (automated planning) or simply planning, aims to develop algorithms to produce plans typically for execution by a robot. Planning software that incorporates these algorithms is called planners. The difficulty of the planning problem depends on the simplification assumptions that are taken for granted, such as atomic time, deterministic time, full observability, etc. A typical planner handles three inputs described in a formal language (PDDL) that uses logical predicates:

- a description of the initial state of a world,
- a description of a goal to be achieved and
- a set of possible actions (sometimes called operators).

Each action is specified by preconditions that must be met in the current state for it to be applied, and post-conditions (effects on the current state).

The search space to browse in order to find a solution plan is very large and generally relies on the use of an ad hoc heuristic function. The objective of the internship is to study and propose algorithms inspired by AlphaZero to learn these heuristic functions automatically using techniques based on neural networks.

3. Profile of the candidate

The candidate must:

- be registered in Master 2
- have advanced programming skills (design and implementation), especially in Java
- own knowledge of how to take users into account in interactive systems
- have a good academic level attesting to his ability to combine practice and theory
- have a level of professional oral and written English
- own general knowledge in the fields of data analysis and artificial intelligence is a plus

4. Contact procedure

Send to Damien.Pellier@imag.fr and Humber.Fiorino@imag.fr and Jean-Pierre.Chevallet@imag.fr :

- Your Master marks 1,
- Your CV.

Applications are managed on a case-by-case basis. You will be informed promptly by email of the admissibility of your application and if you are invited to a first interview.

5. References

[1] <https://www.youtube.com/watch?v=MgowR4pq3e8>

[2] Silver, David; Huang, Aja; Maddison, Chris J.; Guez, Arthur; Sifre, Laurent; Driessche, George van den; Schrittwieser, Julian; Antonoglou, Ioannis; Panneershelvam, Veda; Lanctot, Marc; Dieleman, Sander; Grewe, Dominik; Nham, John; Kalchbrenner, Nal; Sutskever, Ilya; Lillicrap, Timothy; Leach, Madeleine; Kavukcuoglu, Koray; Graepel, Thore; Hassabis, Demis (28 January 2016). "Mastering the game of Go with deep neural networks and tree search". *Nature*. 529 (7587): 484–489.

[3] Damien Pellier, Humbert Fiorino: PDDL4J: a planning domain description library for java. *J. Exp. Theor. Artif. Intell.* 30(1): 143-176 (2018)