

# Études et implémentations de modèles d'apprentissage par renforcement pour les jeux stochastiques de grande taille

**Thématique** : Apprentissage par renforcement, Théorie des Jeux, Processus de Décision Markovien Factorisés

**Équipe d'accueil** : Département DESIR : Équipe RO - Équipe DECISION

**Laboratoire d'accueil** : LIP6 Sorbonne Université

**Lieu** : Paris

**Encadrants** : Pierre-Henri Wuillemin (<http://www-desir.lip6.fr/~phw/>)

Emmanuel Hyon (<http://www-sysdef.lip6.fr/~hyon/index.html>)

**Contacts** : emmanuel.hyon@lip6 pierre-henri.wuillemin@lip6.fr

**Gratification** : Selon critères légaux ( $\approx 550$  euros / mois)

## Contexte

Les jeux stochastiques sont une généralisation des modèles de jeux statiques qui permettent de représenter un grand nombre de situations : en économie, en sécurité des réseaux, en partage de ressources ou encore en modélisation de systèmes multi agents. En effet, ils permettent de modéliser finement les interactions entre les agents (les joueurs) au cours du temps dans un environnement qui évolue aléatoirement. Ce type de jeu se déroule en plusieurs étapes durant lesquelles l'état du système varie suivant les actions des agents et de l'environnement probabiliste. Résoudre de tels jeux consiste à trouver ce qu'on appelle des *équilibres de Nash*. La résolution de tels problèmes est complexe, c'est pourquoi la recherche d'algorithmes efficaces constitue un sujet d'investigation tout aussi bien théorique que pratique.

Nous nous plaçons dans des contextes dans lesquels la totalité de l'espace est difficilement explorable et nous nous intéressons aux méthodes de résolutions basées sur des techniques d'apprentissages par renforcement. Dans ces contextes, les temps de convergence des algorithmes d'apprentissage classiques sont irréalistes. Pour autant, lorsque la totalité de l'espace est difficilement explorable, les processus de décision markoviens qui sont soumis à ce même genre de problèmes ont fait naître des techniques pour prendre en compte cette explosion combinatoire : il s'agit de techniques d'agrégation ou de factorisation qui permettent de décomposer le problème et amènent une possibilité de le représenter de manière très compacte ce, même si le problème est de très grande taille. C'est pourquoi, nous voulons adapter ces techniques de factorisation ou d'agrégation au cadre plus complexe du jeu stochastique.

## Sujet

Le but de ce projet vise à adapter les techniques de factorisation ou d'agrégation au cadre du jeu stochastique. Il s'agira dans un premier temps d'étudier le domaine du jeu Markovien en grande taille et les techniques de factorisation et désagrégation des processus de Décision Markoviens. Dans un second temps, on cherchera aussi à travailler sur la coordination dans l'apprentissage par renforcement qui permet l'accélération des calculs à chaque étape. Enfin, une implémentation et analyse des principaux algorithmes de l'état de l'art est attendue.

La programmation s'effectuera majoritairement en python.

## Ressources

Michael Littman,

Markov games as a framework for multi-agent reinforcement learning

ICML 1994

G. Chalkiadakis, and C. Boutilier,

Coordination in Multiagent Reinforcement Learning: A Bayesian Approach

AAMAS 2003

K. Zhang, and Z. Yang, and H. Liu, and T. Zhang, and T. Basar

Fully Decentralized Multi-Agent Reinforcement Learning with Networked Agents

International Conference on Machine Learning (ICML 2018),

Andriy Brurkov and Brahim Chaib-Draa

Une introduction aux jeux stochastiques

Processus Décisionnels de Markov en Intelligence Artificielle,

Hermes 2008