

Systemes autonomes et apprenants

GRAVOUIL, Kevin, kevin.gravouil@delfox.net, Delfox, Mérignac (Orateur)

JAUREGUIBERRY, Xavier, xavier.jaureguiberry@delfox.net, Delfox, Mérignac

REY, Maxime, maxime.rey@delfox.net, Delfox, Mérignac

Thématique : Intelligence Artificielle

Résumé : Delfox développe des systèmes autonomes et apprenants entraînés par renforcement. Ce *framework* permet à l'utilisateur final d'exprimer sa problématique sous forme de récompenses pour entraîner des IA en simulation. Celles-ci peuvent ensuite être déployées sur des vecteurs physiques et accomplir dans le monde réel la mission pour laquelle elles ont été entraînées.

Mots clés : *deep reinforcement learning, intelligence artificielle, autonomie*

1. Introduction

Delfox est une startup centrée sur la technologie de l'apprentissage profond par renforcement. Notre objectif est de développer des systèmes autonomes et apprenants pour le contrôle ou l'aide à la décision et qui pourront s'adapter à des environnements changeants où une stratégie prédéfinie est rapidement obsolète voire sans solution pratique.

2. Méthodologie

L'entraînement de nos IA repose sur l'apprentissage par renforcement. Ce *framework* nécessite trois composantes : l'environnement de simulation, les réseaux de neurones et l'algorithme d'apprentissage par renforcement.

Nous développons des environnements dédiés aux problématiques de nos clients, incluant notamment la simulation (physique ou logique) et les signaux de récompenses décrivant les missions souhaitées (*reward shaping*, Wiewiora *et al.*, 2011).

Nous structurons l'entraînement des IA par des "plans d'apprentissage" (Bengio *et al.*, 2009) pour aborder le problème de manière didactique et permettre à l'IA d'atteindre ses objectifs en réduisant drastiquement la durée des entraînements.

Nous portons une attention particulière aux applications multi-agents avec une approche "*centralized learning / decentralized execution*" (Lowe *et al.*, 2017) ainsi qu'à l'adéquation entre les réseaux de neurones, la simulation et les hyperparamètres relatifs à l'apprentissage.

3. Originalité / perspective

L'apprentissage profond par renforcement est une technologie de rupture dans l'industrie ASD qui permet (i) d'aborder l'autonomie des systèmes par une formulation des problématiques sous forme de signaux de récompenses et (ii) d'apporter une capacité d'adaptation aux variations imprévues de l'environnement grâce à la puissance d'abstraction des réseaux de neurones profonds.

Nous avons ainsi pu déployer *in situ* une flotte de drones simulant un combat aérien après un entraînement exclusivement en simulation en ne formulant qu'un système de récompenses pour guider l'apprentissage.

Nous développons activement un produit logiciel centré sur l'apprentissage profond par renforcement pour permettre à des utilisateurs de produire simplement des IA pour des problématiques d'autonomie.

L'adoption de nos IA nécessite naturellement une confiance en leur capacité décisionnelle. La suite de nos travaux s'articule donc autour de l'incorporation de mécanismes d'explicabilité pertinent au sein du renforcement.

Références

- Lowe *et al.*, 2017, <https://arxiv.org/abs/1706.02275>
- Bengio *et al.*, 2009, <https://doi.org/10.1145/1553374.1553380>
- Wiewiora *et al.*, 2011, <https://doi.org/10.1007/978-0-387-30164-8>