

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Ecole Supérieure en Informatique - 08 Mai 1945 - Sidi Bel Abbès



THESIS

To obtain the diploma of **Engineering Degree**

Field: **Computer Science**

Speciality: **Artificial Intelligence and Data Science (IASD)**

Integrating Evolutionary Game Theory in Multi-Agent Reinforcement Learning for Non-Stationary Environments

Presented by: TAHIR Abderrahmane El Mehdi

Supervised by: Dr. Belkacem KHALDI

In front of the jury composed of:

Dr. BENDELLA Mohammed Salih	President
Dr. KHALDI Belkacem	Supervisor
Dr. LOUHAB Zouaoui	Examiner

Submission Date: June 2026

Abstract

One of the main challenges in multi-agent reinforcement learning (MARL) is non-stationarity, which arises from the continuous adaptation of multiple learning agents and the resulting changes in the learning dynamics observed by each agent. This challenge becomes even more significant in environments where the underlying dynamics are themselves non-stationary. Evolutionary game theory (EGT) provides a principled framework for studying the evolution of strategies through evolutionary dynamics, and has been explored as a promising direction for improving stability and adaptation in MARL. However, the integration of evolutionary dynamics with modern deep MARL methods remains relatively underexplored.

This thesis introduces DeepED, a novel evolutionary dynamics-inspired MARL algorithm that combines evolutionary game-theoretic principles with deep reinforcement learning. DeepED maintains an evolutionary target policy generated through evolutionary dynamics, where the fitness of strategies is estimated using a centralized critic, and the actor is optimized to track the evolving target policy.

A theoretical analysis of DeepED is conducted, establishing a bounded Nash tracking error affected by environmental variations and function approximation errors, as well as a recursive contraction behavior of the policy tracking gap. Furthermore, DeepED is evaluated through a comprehensive three-level experimental framework covering matrix games, GridWorld environments, and VMAS tasks. The results demonstrate that DeepED improves equilibrium tracking and adaptation in non-stationary multi-agent settings, while remaining competitive in more complex MARL environments.

Keywords: Multi-agent reinforcement learning, Evolutionary game theory, Non-stationary environments, Evolutionary dynamics.

الملخص

يُعد عدم الثبات أحد التحديات الرئيسية في التعلم المعزز متعدد الوكلاء، حيث ينشأ نتيجة التكيف المستمر لعدة وكلاء متعلمين والتغيرات الناتجة عن ذلك في ديناميكيات التعلم التي يلاحظها كل وكيل. ويصبح هذا التحدي أكثر تعقيداً في البيئات التي تكون فيها الديناميكيات الأساسية نفسها غير ثابتة. توفر نظرية الألعاب التطورية إطاراً منهجياً لدراسة تطور الاستراتيجيات من خلال الديناميكيات التطورية، وقد تم استكشافها كاتجاه واعد لتحسين الاستقرار والقدرة على التكيف في التعلم المعزز متعدد الوكلاء. ومع ذلك، فإن دمج الديناميكيات التطورية مع الأساليب الحديثة للتعلم المعزز متعدد الوكلاء العميق لا يزال مجالاً غير مستكشف بشكل كافٍ.

تقدم هذه الأطروحة خوارزمية جديدة للتعلم المعزز متعدد الوكلاء مستوحاة من الديناميكيات التطورية تحت اسم الديناميكيات التطورية العميقة، حيث تجمع بين مبادئ نظرية الألعاب التطورية والتعلم العميق المعزز. تحافظ الخوارزمية على سياسة هدف تطورية يتم توليدها من خلال الديناميكيات التطورية، حيث يتم تقدير ملاءمة الاستراتيجيات باستخدام ناقد مركزي، بينما يتم تحسين الوكيل الفاعل من أجل تتبع سياسة الهدف المتطورة.

يتم إجراء تحليل نظري للخوارزمية، حيث يتم إثبات وجود حد علوي لخطأ تتبع توازن ناش، والذي يتأثر بتغيرات البيئة وأخطاء تقريب الدوال، بالإضافة إلى إثبات سلوك انكماش تكراري للفجوة بين السياسة وسياسة الهدف. علاوة على ذلك، يتم تقييم الخوارزمية من خلال إطار تجريبي شامل مكون من ثلاثة مستويات يشمل الألعاب المصفوفية، وبيئات العالم الشبكي، ومهام الأنظمة متعددة الوكلاء الافتراضية. تُظهر النتائج أن الخوارزمية تعمل على تحسين تتبع التوازنات والقدرة على التكيف في البيئات متعددة الوكلاء غير الثابتة، مع الحفاظ على أداء تنافسي في بيئات التعلم المعزز متعدد الوكلاء الأكثر تعقيداً.

الكلمات المفتاحية : التعلم المعزز متعدد الوكلاء، نظرية الألعاب التطورية، البيئات غير المستقرة، الديناميكيات التطورية.

Résumé

L'un des principaux défis de l'apprentissage par renforcement multi-agents (MARL) est la non-stationnarité, qui résulte de l'adaptation continue de plusieurs agents apprenants et des changements qui en découlent dans les dynamiques d'apprentissage observées par chaque agent. Ce défi devient encore plus important dans les environnements où les dynamiques sous-jacentes sont elles-mêmes non stationnaires. La théorie évolutionnaire des jeux (EGT) fournit un cadre théorique permettant d'étudier l'évolution des stratégies à travers des dynamiques évolutives, et a été explorée comme une approche prometteuse pour améliorer la stabilité et l'adaptation dans le MARL. Cependant, l'intégration des dynamiques évolutives avec les méthodes modernes de MARL profond reste encore relativement peu explorée.

Cette thèse introduit DeepED, un nouvel algorithme de MARL inspiré des dynamiques évolutives, qui combine les principes de la théorie évolutionnaire des jeux avec l'apprentissage profond par renforcement. DeepED maintient une politique cible évolutive générée à travers des dynamiques évolutives, où la performance des stratégies est estimée à l'aide d'un critique centralisé, tandis que l'acteur est optimisé afin de suivre la politique cible en évolution.

Une analyse théorique de DeepED est réalisée, établissant une borne sur l'erreur de suivi de l'équilibre de Nash, influencée par les variations de l'environnement et les erreurs d'approximation des fonctions, ainsi qu'un comportement contractant récursif de l'écart de suivi de la politique. En outre, DeepED est évalué à travers un cadre expérimental complet en trois niveaux couvrant les jeux matriciels, les environnements GridWorld et les tâches VMAS. Les résultats montrent que DeepED améliore le suivi des équilibres et l'adaptation dans des environnements multi-agents non stationnaires, tout en restant compétitif dans des environnements MARL plus complexes.

Mots-clés : Apprentissage par renforcement multi-agents, Théorie évolutionnaire des jeux, Environnements non stationnaires, Dynamiques évolutives.