

(Doctorat : D₄)
RESUME DE THESE¹

Nom et Prénom du candidat : AMHRAOUI EIMehdi.....

Formation Doctorale : Recherche et Développement en Sciences & Ingénierie

Etablissement de domiciliation : ENSAM-Meknès

Centre d'Etudes Doctorales : Sciences et techniques et sciences médicales

Titre de la thèse	Adaptive Optimal Control in Markov Decision Processes and Markov Games using Reinforcement Learning
Discipline/ Spécialité	Discipline : Sciences de l'ingénieur Spécialité : Mathématiques Appliquées et Intelligence Artificielle
Nom et Prénom du Directeur de thèse	MASROUR Tawfik
Structure de Recherche/Etablissement d'Attache	L2M3S – Laboratory of Mathematical Modeling: Simulation and Smart System
Nom et Prénom du responsable de la Structure de Recherche	MASROUR Tawfik
Nom du Codirecteur de thèse	
Structure de Recherche/Etablissement d'Attache	

Résumé : (150 mots)

L'apprentissage par renforcement est une branche de l'apprentissage automatique qui étudie comment un ensemble d'agents peut apprendre à contrôler un système dynamique pour atteindre un objectif spécifique. Ma recherche de thèse se concentre sur la résolution de certains défis de l'apprentissage par renforcement dans deux systèmes dynamiques stochastiques: les processus décisionnels de Markov et les jeux de Markov coopératifs. Ma thèse comporte trois contributions :

La première présente une nouvelle méthode appelée Least Double Q-learning, conçue pour déterminer une politique optimale dans les processus décisionnels de Markov. Cet algorithme traite le biais de maximisation observé dans le Q-learning lorsqu'il est appliqué à des problèmes stochastiques.

La deuxième est l'Expected Lenient Q-learning, un algorithme qui améliore l'apprentissage lent du Lenient Q-learning standard pour les jeux de Markov coopératifs.

La troisième est le Smooth Q-learning, qui établit un compromis entre les approches du maximum et de la moyenne dans l'apprentissage indépendant.

Mots clés : Apprentissage par renforcement, Contrôle optimal, Processus décisionnels de Markov, Jeux de Markov, Jeux de Markov entièrement coopératifs, Q-learning, Deep Q-network, Biais de maximisation, Systèmes multi-agents, Systèmes multi-agents coopératifs, Approche d'apprentissage indépendant, Facteurs de non-coordination, Problème de sur-généralisation relative, Problème de la stochasticité.

¹ Le présent résumé sera publié conformément à l'article 31 des NSPCD- 2023.