

Sorbonne Université

Comparaison entre Augmented Random Search (ARS) et Cross- Entropy Method (CEM)

Joe Chamoun
Jawher Jriby

Overview

- Introduction
- ARS en BBRL
- Comparaison ARS V1 et V2
- Comparaison ARS et CEM
- Conclusion.

Introduction

Cette étude compare les performances de deux algorithmes de renforcement : ARS, développé avec BBRL, et CEM. Nous évaluerons leur efficacité et rapidité de convergence dans divers environnements d'apprentissage pour identifier la méthode la plus performante.



ARS V1

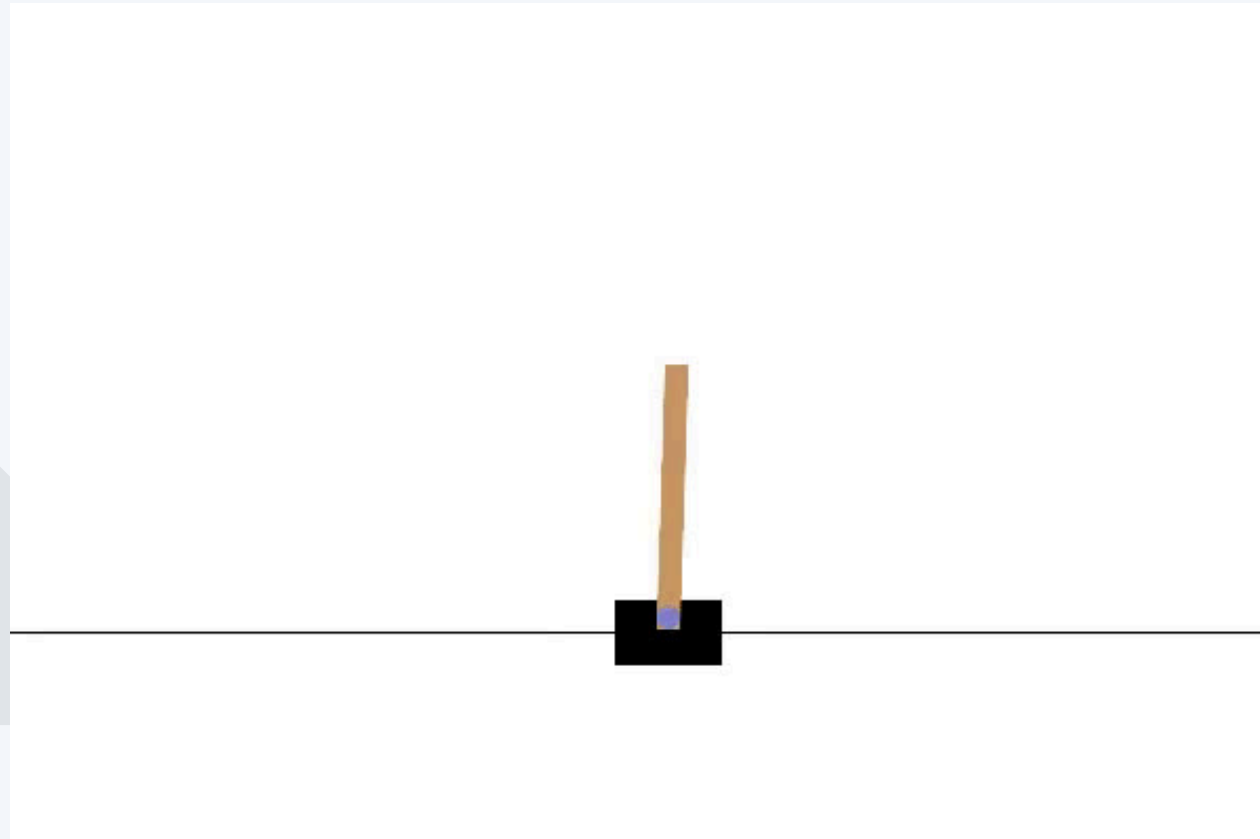
- **Simplicité** : ARS v1 est une méthode d'optimisation simple et efficace pour l'apprentissage par renforcement.
- **Perturbations Aléatoires** : Utilise des perturbations gaussiennes aléatoires pour explorer l'espace des paramètres de la politique.
- **Performances** : Offre des résultats compétitifs dans des environnements discrets et continus.
- **Convergence** : Peut nécessiter un grand nombre d'itérations pour converger vers une politique optimale.

ARS V2

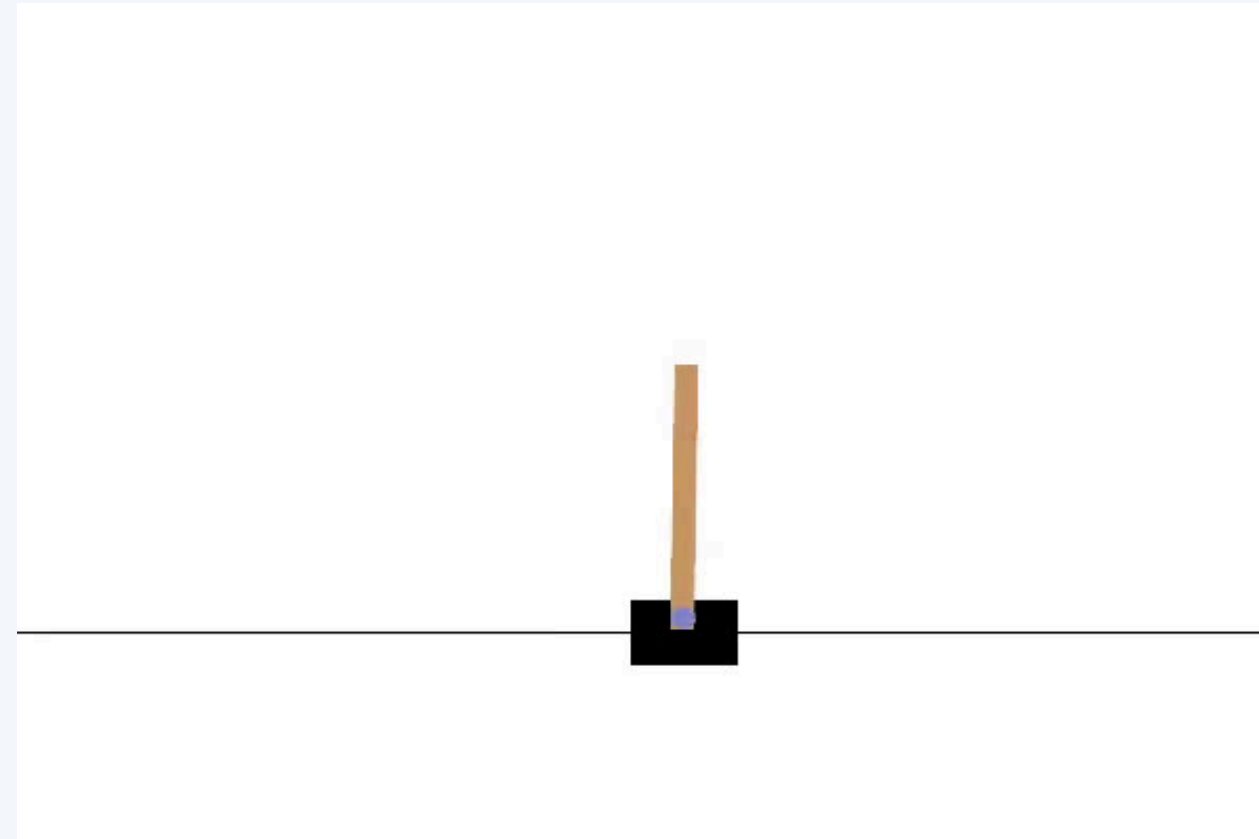
- **Normalisation** : Introduit la normalisation par l'écart type, ce qui permet des ajustements plus précis des paramètres.
- **Stabilité Améliorée** : La normalisation réduit la variance des récompenses entre les épisodes, augmentant la stabilité des performances.
- **Convergence Accélérée** : Converge généralement plus rapidement vers des récompenses élevées par rapport à ARS v1.
- **Efficacité** : Montre une meilleure efficacité dans l'exploration de l'espace des politiques, surtout dans des environnements complexes.

V1 & V2

Environnement Continu (CartPoleContinuous-v1)



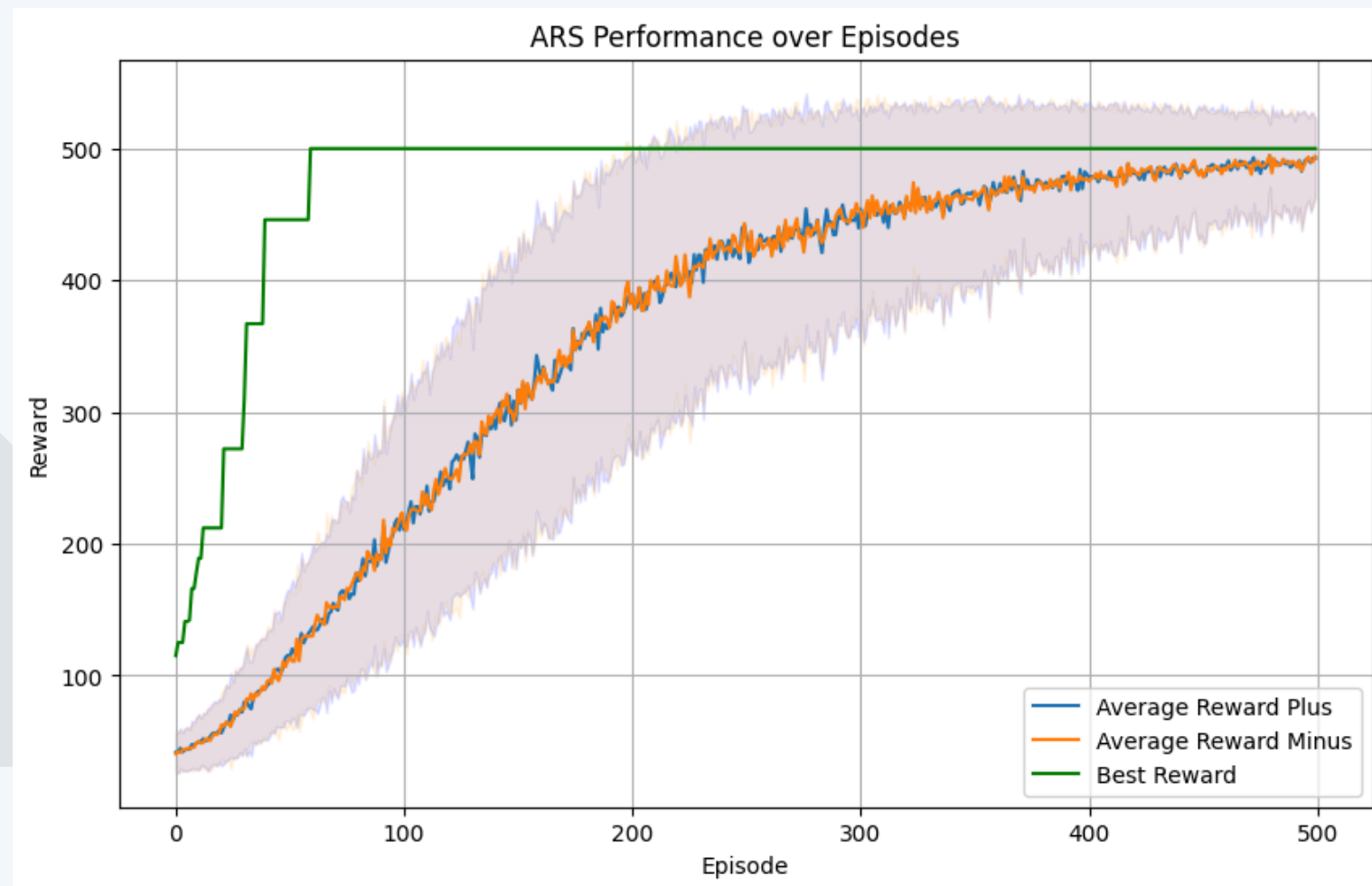
ARS v1 CartPoleContinuous-v1



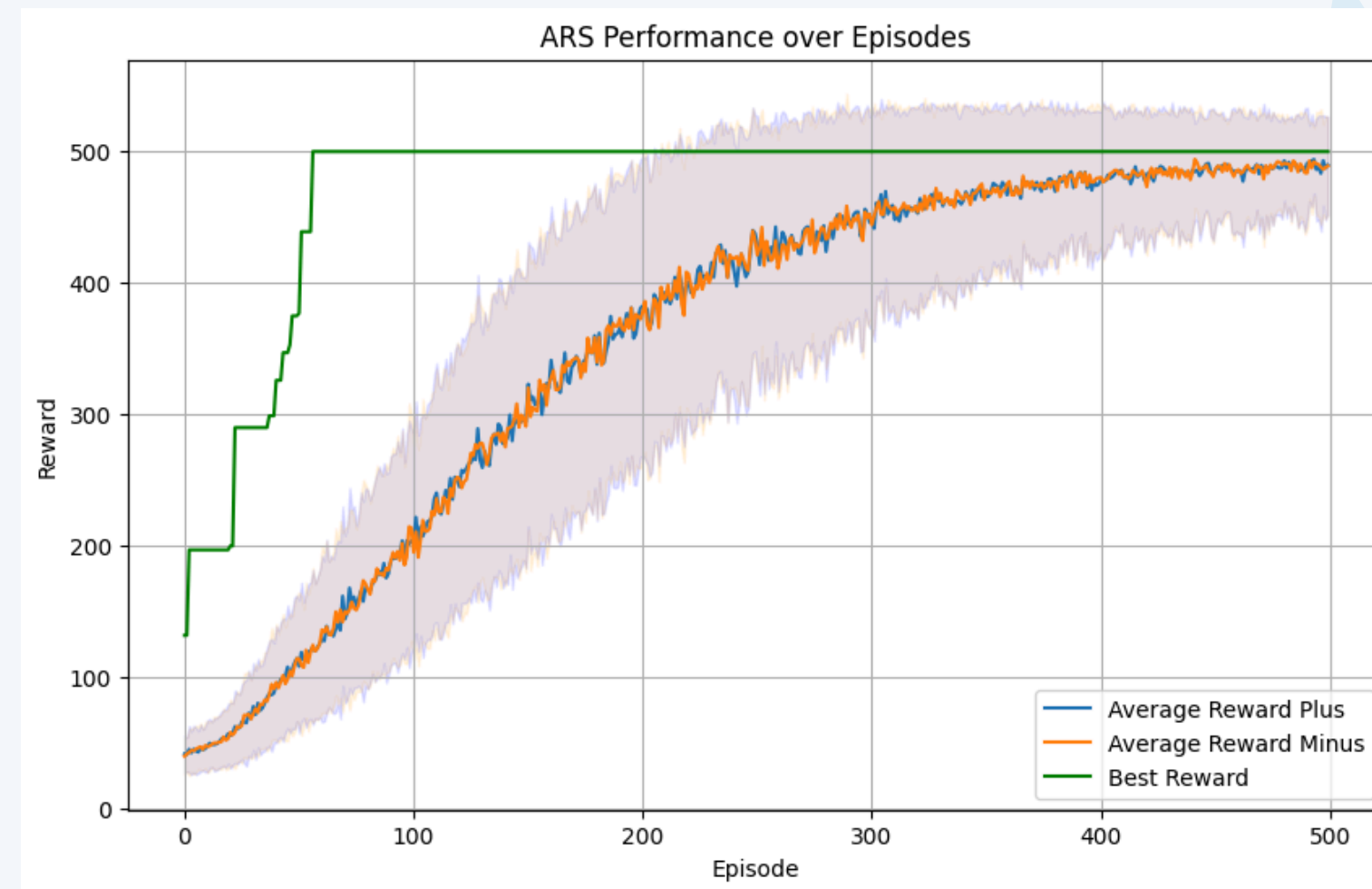
ARS v2 CartPoleContinuous-v1

V1 & V2

Environnement Continu (CartPoleContinuous-v1)



ARS v1 CartPoleContinuous-v1



ARS v2 CartPoleContinuous-v1

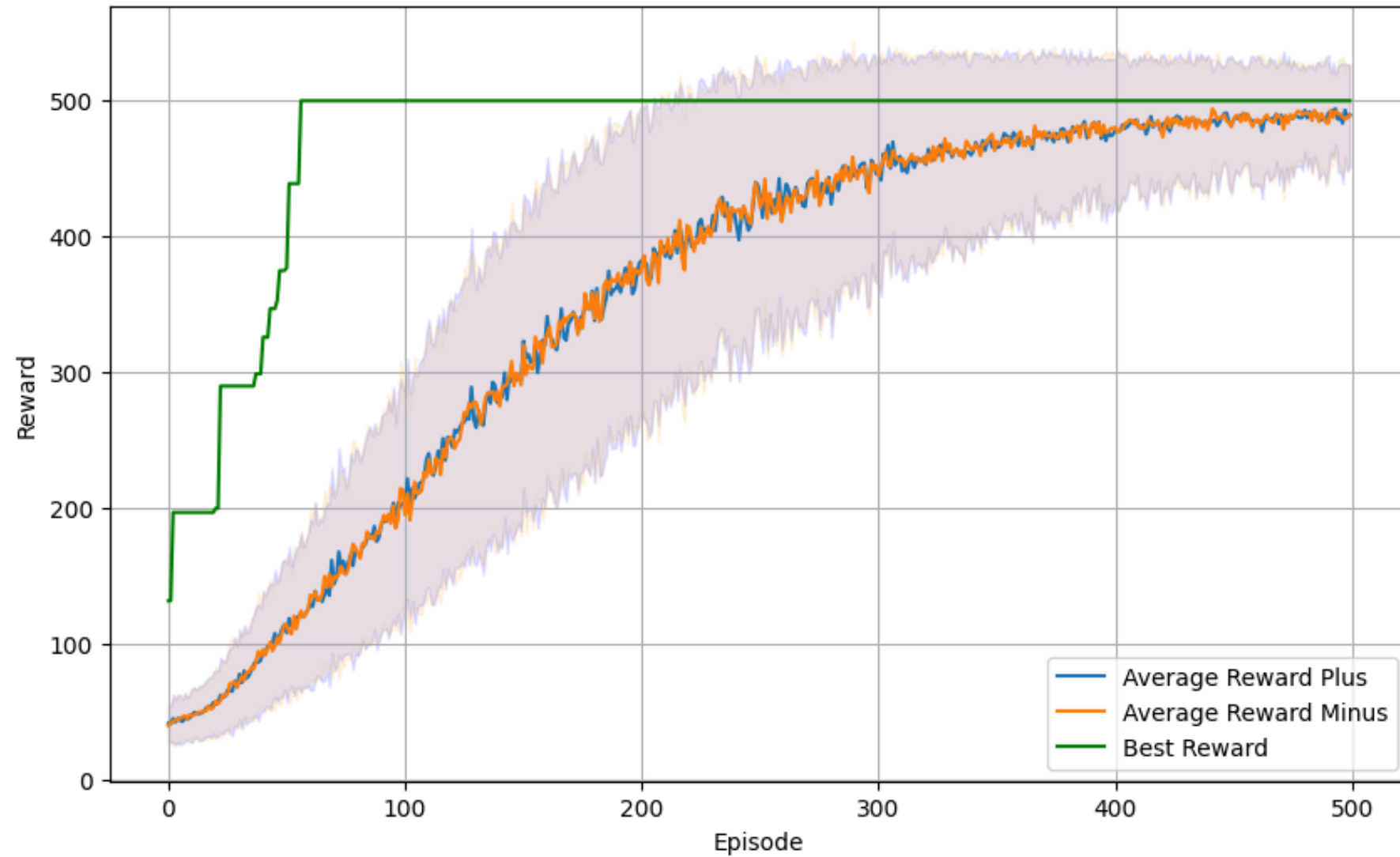
V1 & V2

- ARS v2 normalise les récompenses :
 - Aide à gérer différentes échelles de récompenses.
 - Rend le processus d'apprentissage plus stable.
 - Convergence plus rapide vers des politiques à haute récompense.
- ARS v2 normalise les observations :
 - Assure une moyenne nulle et une variance unitaire.
 - Conduit à un apprentissage plus robuste.

ARS & CEM

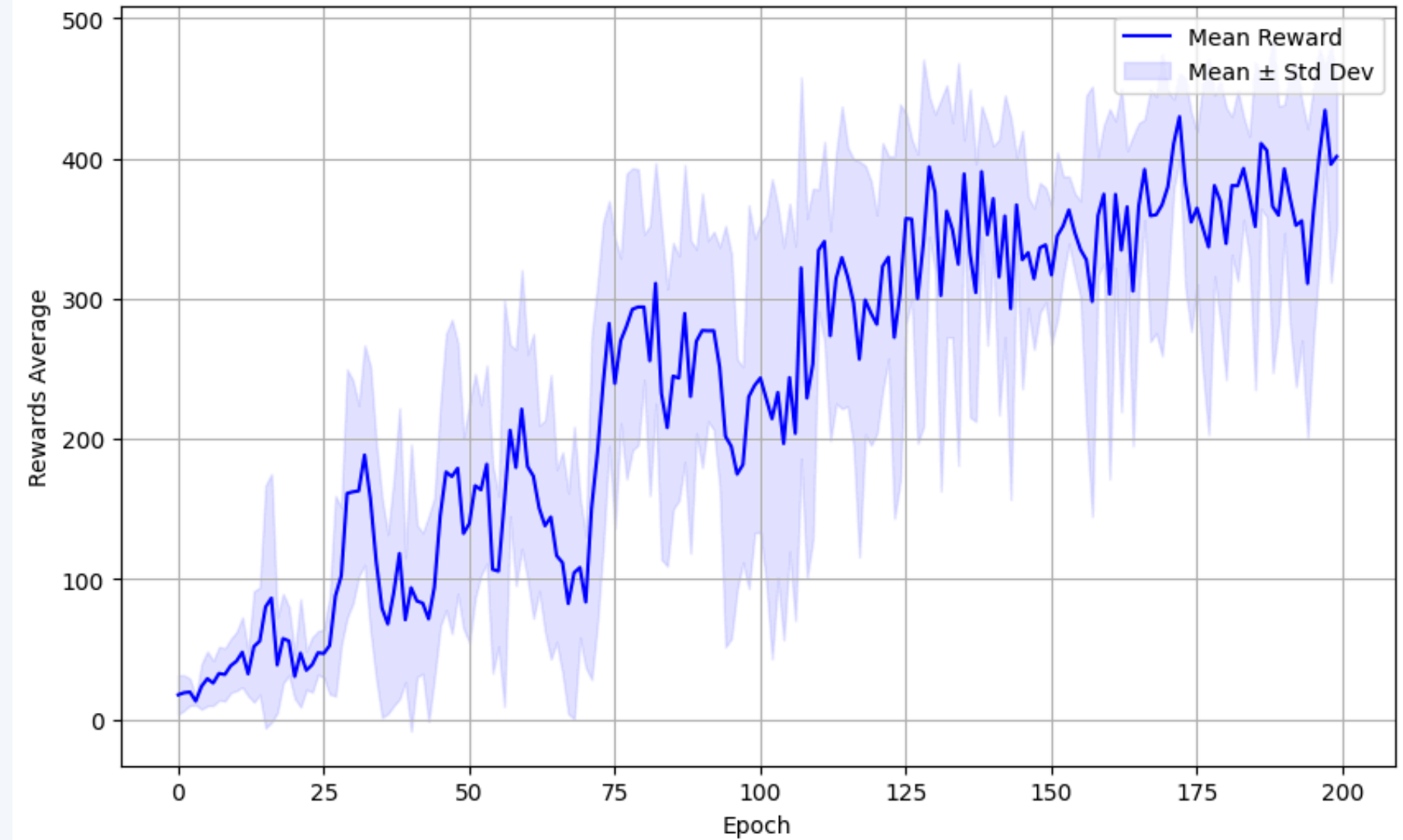
Environnement Continu (CartPoleContinuous-v1)

ARS Performance over Episodes



ARS v2 CartPoleContinuous-v1

Rewards Across Epochs (Mean \pm Std Dev)



CEM CartPoleContinuous-v1

ARS & CEM

Environnement Continu (CartPoleContinuous-v1)

- Normalisation:
 - **ARS** : Normalise les récompenses et les observations pour une stabilité et une robustesse accrues.
 - **CEM** : Ne normalise pas les récompenses, se concentre sur l'optimisation directe des politiques.
- Stabilité:
 - **ARS** : Convergence plus stable grâce à la gestion des différentes échelles de récompenses.
 - **CEM** : Peut être moins stable sans normalisation, dépend fortement des paramètres initiaux.
- Rapidité de Convergence:
 - **ARS** : Convergence rapide vers des politiques à haute récompense.
 - **CEM** : Rapide mais peut nécessiter un réglage précis des paramètres.
- Robustesse:
 - **ARS** : Moins affecté par les variations des caractéristiques d'entrée.
 - **CEM** : La performance peut varier en fonction des distributions initiales.

Conclusion

Les résultats montrent que ARS, notamment sa version 2, présente une stabilité supérieure et une convergence plus rapide dans la plupart des cas. Cependant, bien qu'avec une variance plus importante, CEM peut atteindre des niveaux de performance élevés. En fin de compte, le choix de la méthode appropriée dépend des besoins et des contraintes spécifiques de l'application concernée.





Sorbonne Université

THANK YOU

For Your Attention