

Compte Rendu de Réunion 4

POINTS CLÉS

BBRL et méthodologie.

- Complétion des notebooks BBRL, intérêt des multiples environnements gymnasium pour la suite
- Hyperparamètres : phase de compréhension manuelle, puis phase de comparaison avec budget équitable

Revue des implémentations.

- SAC avec BBRL : revue de la théorie en détail et revue de l'implémentation
- AFU avec BBRL : performances qui semblent inférieures à la version de M. Perrin, poursuite du développement.

Résultats expérimentaux.

- Implémentation AFU sous-optimale
- Convergence inattendue des trois algorithmes en off-policy
- Stabilité post-convergence en off-policy

POUR LA SEMAINE PROCHAINE

Technique.

- Tests avec code de M. Perrin
- Extension de `set_state` (pendulum)
- Vérification uniformité échantillonnage
- Amélioration implémentation AFU
- Étude IQL

Organisation.

- Cahier des charges
- Communication résultats intermédiaires