

Apprentissage par renforcement et Système dynamique : équilibrage d'un double pendule.

Polytech Lyon, MAM 4A

3 mars 2025

1 Description du problème

Dans ce projet, vous devrez entraîner un agent à garder en équilibre un Double Pendule sur un Chariot (DPC) illustré dans la Figure 1. Les interactions de l'agent avec l'environnement consistent à appliquer une force horizontale u sur le chariot. À chaque transition (changement d'état du système suite à l'action de l'agent), l'agent reçoit une récompense suivant le résultat de son action. Lorsque la distance entre la position à l'équilibre ($\theta_1 = 0, \theta_2 = 0$) et la position actuelle dépasse un certain seuil, l'environnement / la simulation est interrompu.

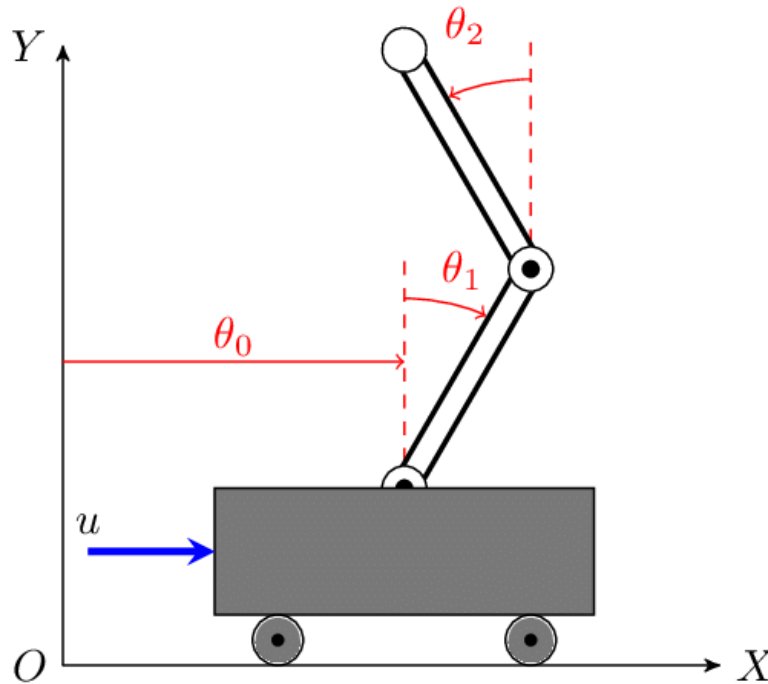


FIGURE 1 – Schéma d'un double pendule sur un chariot mobile

2 Évaluation

Vous devez fournir (i) le code source du projet, ainsi que (ii) un rapport **source** résumant votre travail sur ce projet. Votre rapport devra également détailler les améliorations possibles de votre approche. Veuillez noter que le code source et le rapport sont tous deux obligatoires et que vous devez implémenter **vous-même** les algorithmes d'apprentissage par renforcement pertinents.

Pour aller plus loin :

Si vous souhaitez aller plus loin, vous pouvez essayer de créer votre propre environnement/simulation afin d'essayer de résoudre un problème encore plus complexe : la transition d'un point d'équilibre à un autre (ex : de $[\theta_1 = \pi, \theta_2 = \pi]$ à $[\theta_1 = 0, \theta_2 = 0]$)

3 Environnement

Pour l'aspect de simulation physique de l'environnement, vous utiliserez Gymnasium [1] qui comporte une simulation du DPC. Dans votre rapport, vous devrez formaliser le problème, décrire l'espace d'état (*i.e.* input) et l'espace d'action (*i.e.* output) de votre agent. Vous n'aurez pas besoin d'écouter les équations de la dynamique du système pour cette partie. Cependant, si vous souhaitez aller plus loin et faire votre propre simulation physique, ce travail sera nécessaire.

4 Apprentissage par renforcement

Comme dit précédemment, dans ce projet, vous devrez entraîner un agent à stabiliser le DPC. Pour ce faire, vous allez devoir construire vos propres algorithmes d'apprentissage par renforcement. Vous devrez créer au moins deux algorithmes : un génétique [2] [3] et un de Q-learning [4] [5] [6] [7]. Dans votre rapport, vous devrez détailler le protocole expérimental qui vous permettra d'évaluer les performances et les limites de chaque modèle. Pensez à citer les travaux sur lesquels vos approches sont basées.

Vous trouverez ci-dessous plusieurs références sur les différentes méthodes existantes.

Références

- [1] Mark TOWERS et al.
« Gymnasium : A standard interface for reinforcement learning environments ».
In : *arXiv preprint arXiv :2407.17032* (2024).
- [2] Anupriya SHUKLA, Hari Mohan PANDEY et Deepti MEHROTRA.
« Comparative review of selection techniques in genetic algorithm ».
In : *2015 international conference on futuristic trends on computational analysis and knowledge management (ABLAZE)*. IEEE. 2015, p. 515-519.
- [3] Kenneth O STANLEY et Risto MIKKULAINEN.
« Evolving neural networks through augmenting topologies ».
In : *Evolutionary computation* 10.2 (2002), p. 99-127.
- [4] Jesse CLIFTON et Eric LABER. « Q-learning : Theory and applications ».
In : *Annual Review of Statistics and Its Application* 7.1 (2020), p. 279-301.
- [5] Volodymyr MNIH et al. « Playing atari with deep reinforcement learning ».
In : *arXiv preprint arXiv :1312.5602* (2013).
- [6] Timothy P LILICRAP et al. « Continuous control with deep reinforcement learning ».
In : *arXiv preprint arXiv :1509.02971* (2015).
- [7] Tuomas HAARNOJA et al. « Soft actor-critic : Off-policy maximum entropy deep reinforcement learning with a stochastic actor ».
In : *International conference on machine learning*. Pmlr. 2018, p. 1861-1870.