

Démarche

Mehdi Mounsif

11 avril 2018

1 Récap : The day before

- Préparation de Reacher GAIL
- Lectures sur les quaternions duals pour représentation de la position et de l'orientation
- Lecture de [1]
- GANs : génération de points autour d'une gaussienne. Echec. Essayer en 1D

2 GAIL : Echec discret

La mise en place de GAIL dans un environnement Reacher Discret avec 8 actions possible a échoué. La politique semble directement aller en mode collapse : une action écrase les autres et les épisodes d'observations montrent que l'agent reste souvent figé dans le même coin de l'espace. La fonction discriminatrice semble affecter une récompense basse à l'agent, ce qui est assez cohérent. Pour la suite :

- Incorporer entropie à la fonction de coût de la politique, pour conserver l'exploration.
- Curriculum Learning
- PPO avec espace d'action continu
- Soft Actor Critic ?

3 Noisy Grid Reacher

En attendant une prise de position sur l'amélioration de SuperBot, j'ai évalué le comportement d'un agent face à des perceptions bruitées. Les procédés appliqués sont rudimentaires, puisque aucune donnée n'est disponible sur la forme, l'intensité ou la magnitude de ce bruit.

- Appliquer directement du bruit aux inputs. On considère $s_\epsilon = s + \epsilon$ avec $\epsilon \sim \mathcal{N}(0, \sigma)$. C'est s_ϵ qui est fourni à la fonction plutôt que s . Pour des variances relativement importantes (par rapport à l'échelle de l'environnement), le comportement reste similaire
- Utilisation d'une couche de Dropout après chaque couche cachée pendant l'entraînement. Performances similaires.

Améliorations possibles :

- La précision de la connaissance des positions des obstacles ou cible pourrait être relative à la distance.

4 SAC ?

Références

- [1] GOODFELLOW, I., BENGIO, Y., AND COURVILLE, A. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.