



## Stage de 6 mois dans l'équipe Flowers, INRIA Bordeaux Sud-Ouest

**Titre:** Apprentissage de modèles par exploration de buts intrinsèquement motivée et réseaux de neurones profonds

Ce stage aura pour objectif d'étudier comment des algorithmes d'apprentissage intrinsèquement motivés, et basés sur la génération autonome de buts (Intrinsically Motivated Goal Exploration Processes), peuvent être utilisés pour apprendre efficacement des modèles neuronaux profonds de la dynamique d'un environnement. Il s'agira d'abord de prendre en main des algorithmes d'apprentissage de modèles basés sur des réseaux neuronaux profonds qui apprennent à prédire la distribution des états  $s(t+1)$  résultant de la réalisation de l'action  $a(t)$  dans l'état  $s(t)$  (Ha and Schmidhuber, 2018), et qui peuvent être utilisés pour optimiser des séquences d'actions/politiques qui maximisent une fonction de récompense (Chua et al., 2018). Ces architectures seront ensuite utilisées dans le cadre des IMGEPs modulaires (Forestier et al., 2016, 2017) afin de guider l'exploration de l'environnement, et la collecte de données permettant d'entraîner ces modèles. La dynamique de cette architecture pourra être comparée à des algorithmes d'apprentissage par renforcement model-free, comme DDPG+HER (Andrychowicz et al., 2017), et leur combinaison avec des IMGEPs (Colas et al., 2018).

**Encadrants :** Pierre-Yves Oudeyer et Cédric Colas

### Références:

Andrychowicz, M., Wolski, F., Ray, A., Schneider, J., Fong, R., Welinder, P., ... & Zaremba, W. (2017). Hindsight experience replay. In *Advances in Neural Information Processing Systems*(pp. 5048-5058).

Colas, C., Sigaud, O., & Oudeyer, P. Y. (2018). GEP-PG: Decoupling Exploration and Exploitation in Deep Reinforcement Learning Algorithms. *arXiv preprint arXiv:1802.05054*.

Ha, D., & Schmidhuber, J. (2018). Recurrent World Models Facilitate Policy Evolution. *arXiv preprint arXiv:1809.01999*.

Kurtland Chua, Roberto Calandra, Rowan McAllister, and Sergey Levine. Deep reinforcement learning in a handful of trials using probabilistic dynamics models. *Advances in Neural Information Processing Systems (NIPS)*, 2018.

Forestier S, Oudeyer P-Y. 2017. Intrinsically Motivated Goal Exploration Processes with Automatic Curriculum Learning. *arXiv*.

Forestier S, Oudeyer P-Y. 2016. Modular Active Curiosity-Driven Discovery of Tool Use. 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).

Forestier S, Oudeyer P-Y. 2016. Modular Active Curiosity-Driven Discovery of Tool Use. 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).