

Apprentissage par Renforcement sur Processeur Neuromorphique

Proposé par : Alain Dutech (Alain.Dutech@loria.fr),
Bernard Girau (Bernard.Girau@loria.fr)

Problématique

L'équipe BISCUIT, est une équipe du laboratoire Loria qui étudie de nouveaux paradigmes computationnels où les calculs sont adaptatifs, distribués et décentralisés, réalisés par des populations d'unités de calcul simples qui communiquent principalement avec leurs proches voisins. Ces propriétés sont compatibles avec la mise en oeuvre de principes d'auto-organisation non-supervisés, mais guidés, pour s'attaquer à des problèmes difficiles comme le calcul cognitif situé, la robotique autonome, l'allocation adaptative de ressources de calcul, etc.

Ces caractéristiques permettent également d'envisager une meilleure utilisation des processeurs dit "neuromorphiques" qui sont en train d'émerger (IBM Truenorth, Intel Loihi, etc.). Ces processeurs sont basés sur des principes neuro-inspirés qui respectent les contraintes des paradigmes que nous étudions, et peuvent bénéficier des mécanismes d'auto-organisation non supervisée et guidée que nous mettons au point, à la fois en termes d'applications et en termes de gestion des ressources neuromorphiques.

Sujet

L'objectif principal de ce stage de master est d'explorer l'utilisation de mécanismes situés à la croisée du calcul neuro-inspiré et de l'apprentissage par renforcement [1] dans le cadre de ces architectures dites *neuromorphiques*.

En s'inspirant d'un précédent travail de master, nous voulons poursuivre la définition, l'étude et l'implantation un algorithme d'apprentissage par renforcement distribué utilisant des populations de neurones impulsionnels et compatible avec le processeur neuromorphique Loihi d'Intel [2]. Contrairement aux approches principales de la littérature utilisant un signal de récompense dans un cadre neuromorphique, notre algorithme doit être capable d'apprentissage séquentiel : la récompense à l'agent n'est donnée qu'à la fin de la trajectoire. En cela, l'approche de Frémaux est une inspiration intéressante [3]. Nos travaux précédents ont permis de réexprimer cette approche selon les contraintes neuromorphiques du processeur Loihi. Néanmoins, de nombreux aspects restent à étudier, notamment les problèmes de précision des poids synaptiques et de combinaison des traces d'activité synaptique. Enfin, un accent particulier sera mis sur la validation expérimentale du modèle sur des problèmes concrets de contrôle (robotique, saccades attentionnelles, etc).

Compétences

python, Linux, quelques notions de mathématiques (probabilités, équations différentielles), une connaissance préalable de l'apprentissage par renforcement et/ou du calcul neuronal impulsionnel sera appréciée

Conditions de travail

Le stage se déroulera au sein du laboratoire Loria de Nancy. Le développement des modèles se fera principalement en python, y compris pour l'utilisation des bibliothèques spécifiques de programmation sur le processeur neuromorphique Intel Loihi.

Bibliographie

- [1] Sutton, R. and Barto, A. *Reinforcement Learning* Bradford Book, MIT Press, Cambridge, MA. (1998)
- [2] Davies, M., Srinivasa, N., Lin, T., Chinya, G., Cao, Y., Choday, S. H., Dimou, G., Joshi, P., Imam, N., Jain, S., Liao, Y., Lin, C., Lines, A., Liu, R., Mathaikutty, D., McCoy, S., Paul, A., Tse, J., Venkataramanan, G., Weng, Y., Wild, A., Yang, Y., and Wang, H.. *Loihi : A neuromorphic manycore processor with on-chip learning*. IEEE Micro, 38(1) :82-99, (2018).
- [3] Frémaux, Nicolas, Sprekeler, Henning and Gerstner, Wulfram. *Reinforcement learning using a continuous time actor-critic framework with spiking neurons* PLoS Comput Biol, vol 9(4), (2013).