

Projet de Recherche Doctoral
Labex SMART

Titre : **Robots constructing representations from experience**

Encadrants :

- *directeur de thèse* : Stéphane Doncieux, Professeur à l'ISIR (doncieux@isir.upmc.fr)
- *co-directeur de thèse* : Nicolas Perrin, CR2 CNRS à l'ISIR (perrin@isir.upmc.fr)

Laboratoire d'accueil : **ISIR**

Axes thématiques :

- *Interfaces and interactions & Humans at the convergence of digital and real environments* (axes 2 et 3) : Le but de cette thèse est de permettre à un robot de créer, à partir de ses expériences sensori-motrices, des représentations internes permettant d'améliorer ses prises de décisions. En utilisant ces représentations dans un contexte d'apprentissage par renforcement, le robot pourra adapter efficacement son comportement afin d'optimiser sa capacité à obtenir une récompense définie par l'utilisateur. Ce type d'aptitude ouvre la voie à des interfaces simples et intuitives où l'humain peut enseigner de nouvelles tâches au robot très facilement, sans avoir à programmer en détails les comportements attendus. Puisque les représentations internes sont construites à partir d'une perception de l'environnement réel sur lequel le robot peut agir, elles sont le fruit d'une interaction forte entre les algorithmes utilisés et le monde réel, interaction dans laquelle l'humain a un rôle central.
- *Modeling of humans* (axe 1) : L'approche proposée s'inspirera en partie de la façon dont les humains construisent et reconstruisent leurs propres représentations. Par exemple, des expériences ont pu montrer que le sommeil engendre une restructuration des connaissances, favorisant ainsi la résolution de problèmes [1]. La psychologie développementale a étudié de nombreuses trajectoires développementales d'enfants et a pu également mettre en évidence le rôle crucial de la redescription des représentations [2, 3]. Les résultats obtenus permettront éventuellement de faire un pas vers la modélisation computationnelle de ces phénomènes.

Sujet de thèse :

Designing algorithms that provide robots or agents with autonomy is often a complicated and tedious endeavor. To make it easier, many learning algorithms exist that can optimize over time the parameters of the mapping from perceptions to actions, with the objective to increase some measure of the quality of the behavior, like for instance the ability to collect a reward.

However, the mapping from perceptions to actions has most of the time a fixed structure, with some internal representations that are not modified during learning. The success of the learning process depends largely on these representations, and it would be desirable for learning algorithms to have the ability to infer appropriate representations directly from experience.

The goal of the proposed PhD is to explore different research directions for the automated construction of representations via an a posteriori analysis of sensori-motor traces.

The idea is to start with a classical learning algorithm with fixed representations, and, based on the results obtained, extract interesting features from the sensori-motor recordings (the traces), and construct new representations that seem appropriate for the tasks the robot is trying to perform. The usefulness of the proposed representations will then be quantitatively assessed, following two methods that consist in answering the following questions :

- Do the new representations yield a speed-up when trying to relearn from scratch ?
- Do the new representations yield better transferability, when, after achieving the learning process, the robot is suddenly put in a new environment ?

Two main research axes will be considered :

1. Skill extraction, organization and refining.

Skills are perception-action feedback loops. Instead of learning a unique unconstrained map from perceptions to actions, it is in general easier to use a finite set of skills and learn how to make appropriate switches.

Humans and animals tend to train isolated skills and then compose with them to generate rich arrays of movements. Skills help to articulate low-level reactive behaviors and higher-level decisions.

Our specific objective will be to design skills by analysing traces of behaviors generated via a first learning phase with “naive” representations and no skills.

A large body of relevant work exists on skill chaining and on the related concept of options [4, 5, 6], which have also ties with subgoal extraction methods [7, 8]. However, many questions are still open, especially in the case of continuous sensory inputs and only partially observable states. For instance :

- What are the most adequate ways to represent skills ?
- Is it useful to differentiate several categories of skills ?
- How to validate the relevance of a given skill with respect to some well-defined task ?
- How to perform high-level reasoning to orchestrate the composition of skills ?
- For efficient skill construction, how to cope with the problems of delayed reward and perceptual aliasing ?

Such questions will be central in the proposed PhD thesis.

2. Definition of internal states.

This axis is complementary to the first one, and aims at defining internal representations, or internal states, that depend on the perceptions and possibly on the memory.

The role of such internal states is to facilitate the selection of actions, often via an explicit or implicit dimensionality reduction. Indeed, these internal states should provide a compressed representation of the information that can be inferred from the sensory inputs. They should also be robust to noise (thus leading to more robust behaviors), and yield a better generalization than with actions expressed as functions of the raw input data. When these internal states are discrete, they can allow symbolic reasoning, and therefore their construction is closely related to the symbol grounding problem [9].

Again, several open questions will be examined :

- How to decide the amount of memory that should be used in internal states specifically built for a given task ?
- How to dynamically modify internal representations ?
- How to store and decide the reactivation of former internal representations ?
- How to take rewards into account when clustering the perception space ?
- How can relevant discrete internal states be constructed from a completely continuous experience ?

The scientific literature on these questions contains a large range of potential answers, usually without consensus, and many tools exist that consider the problems from

different angles. Various approaches and various data mining and machine learning algorithms will be considered to solve the aforementioned questions (e.g. reinforcement learning & evolutionary algorithms, random forests, self-organizing maps, etc.). One tool that will be of particular interest is *unsupervised deep learning*. It has been shown that deep networks (e.g. Deep Beliefs Networks [10] or Stacked Autoencoders [11]) have the ability to create neurons, in the hidden layers, that fire upon detection of somewhat high-level, abstract features of the input. Such neurons could be interesting building blocks for the construction of abstract internal states, but as of now there exist no method to do it directly. Some progress is required, for instance on the assessment of the “abstractness” of a neuron, or the characterization of the usefulness of a neuron with respect to a given task.

The PhD will be done in the context of the H2020-FETPROACT collaborative project DREAM [12]. A large effort will be put on the implementation of the ideas, with systematic validation on robotics experiments or simulations that have been precisely defined by the DREAM consortium.

Références :

- [1] Wagner, U., Gais, S., Haider, H., Verleger, R., Born, J. (2004). Sleep inspires insight. *Nature*, 427, pp. 352-355.
- [2] Karmiloff-Smith, A. (1990). Constraints on representational change : Evidence from children’s drawing. *Cognition*, 34, pp. 57-83.
- [3] Karmiloff-Smith, A. (1995) *Beyond Modularity : A Developmental Perspective on Cognitive Science*. MIT Press.
- [4] Sutton, R. S., Precup, D. & Singh, S. (1998). Between MDPs and Semi-MDPs : Learning, Planning, and Representing Knowledge at Multiple Temporal Scales. Technical report, University of Massachusetts.
- [5] Konidaris, G., Kuindersma, S., Grupun, R. & Barto, A. (2012). Robot learning from demonstration by constructing skill trees. *The International Journal of Robotics Research*, 31(3), pp. 360-375.
- [6] Stulp, F., Herlant, L., Hoarau, A. & Raiola, G. (2014). Simultaneous On-line Discovery and Improvement of Robotic Skill Options. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS’14)*.
- [7] McGovern, A. & Barto, A. (2001). Automatic discovery of subgoals in reinforcement learning using diverse density. *Proceedings of the 18th International Conference on Machine Learning (ICML’01)*.
- [8] Şimşek, Ö., Wolfe, A. P., & Barto, A. (2005). Identifying useful subgoals in reinforcement learning by local graph partitioning. *Proceedings of the 22nd International Conference on Machine learning (ICML’05)*
- [9] Harnad, S. (1990). The symbol grounding problem. *Physica D : Nonlinear Phenomena*, 42(1), pp. 335-346.
- [10] Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, 18(7), pp. 1527-1554.
- [11] Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y. & Manzagol, P. A. (2010). Stacked denoising autoencoders : Learning useful representations in a deep network with a local denoising criterion. *The Journal of Machine Learning Research*, 11, pp. 3371-3408.
- [12] [http ://www.robotsthatdream.eu/](http://www.robotsthatdream.eu/)

CV détaillé du candidat et lettre de motivation : le candidat est Omar Mohammed. Voir les documents joints pour son CV et sa lettre de motivation, ainsi que des lettres de recommandations.

Notes du master, UEs validées, classement et moyenne du candidat : voir documents joints.