Internship proposal, 2014

Deep Learning for Robotics

RESPONSABLE: Michèle Sebag
LABORATOIRE: Laboratoire de Recherche en Informatique CNRS UMR 8623 & INRIA Saclay
ADRESSE: Université Paris-Sud, 91405 Orsay Cedex
MAIL: Michele.Sebag@lri.fr
URL: `http://www.lri.fr/~sebag/Stages/Deep_RL.pdf`
SUMMARY:
Reinforcement learning achievements critically depend on the representation of the state space. High-dimensional state spaces (e.g. described through the many sensors or camera pixels of the robot) hinder the characterization of the value functions. Former attempts rely on function approximations (e.g. to deal with continuous search spaces), feature selection (to cope with high state dimensionality), or the use of models to guide the sampling of the search space.

Basically, RL involves three interdependent problems: modelling the environment and the transition model (a.k.a forward model for a robot, which can be thought of as a simulator, estimating the next state from the current state and the selected action); modelling the environment and the reward (a.k.a. learning the value functions, estimating how much cumulative reward the robot will get from a given state following an improving policy); exploring the action space to support a better modelling of transitions and values.

In the neighbor field of supervised learning, deep learning has revolutionized the learning strategies, suggesting that the unsupervised learning of an information-efficient representation is the primary learning step.

Goal of the internship

The goal is to investigate how the principles of deep learning can be adapted and extended to reinforcement learning and specifically to robotics.

Overview

A possible strategy consists of alternating the physical exploration of the environment, aimed at gathering samples, and using it to extract a better representation of the state space:

1. **Entropy-based exploration**. This phase, resuming earlier work (in simulation), builds controllers maximizing the entropy of the sensori-motor data stream;
2. The sensori-motor samples gathered by the previous controllers are provided to a deep neural net, extracting a **generative model** (restricted Boltzman machine) or achieving **lossy compression** (stacked auto-encoders);
3. The code/generative model can be exploited by an explicit controller for **exploratory action selection** (e.g. selecting the action with most uncertain output);
4. The last phase is concerned with value learning and policy optimization.

The internship study can be conducted along a theoretical or an algorithmic perspective. Excellent programming skills (C++) and a solid mathematical background are required.

References

Reinforcement Learning, An introduction. Richard S. Sutton and Andrew G. Barto. MIT Press, Cambridge, MA, 1998.

Representation Learning: A Review and New Perspectives, Yoshua Bengio, Aaron Courville, Pascal Vincent, Arxiv, 2012.