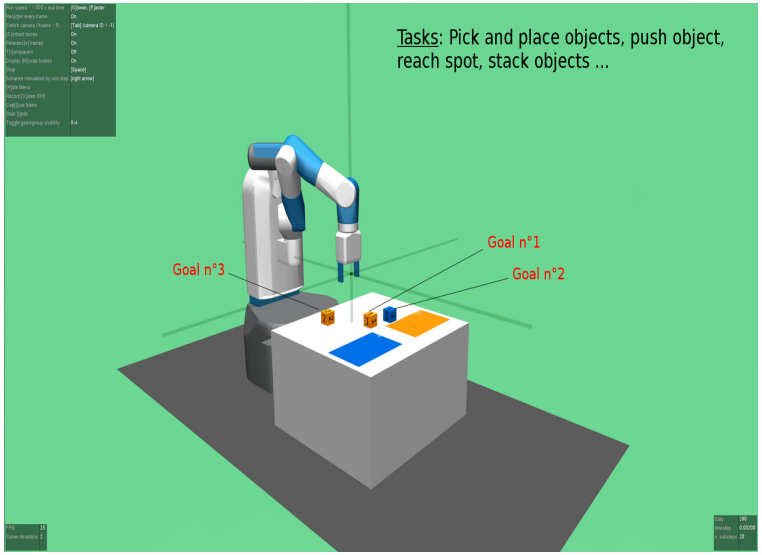
From AlphaNPI to Hilbert

Olivier Sigaud

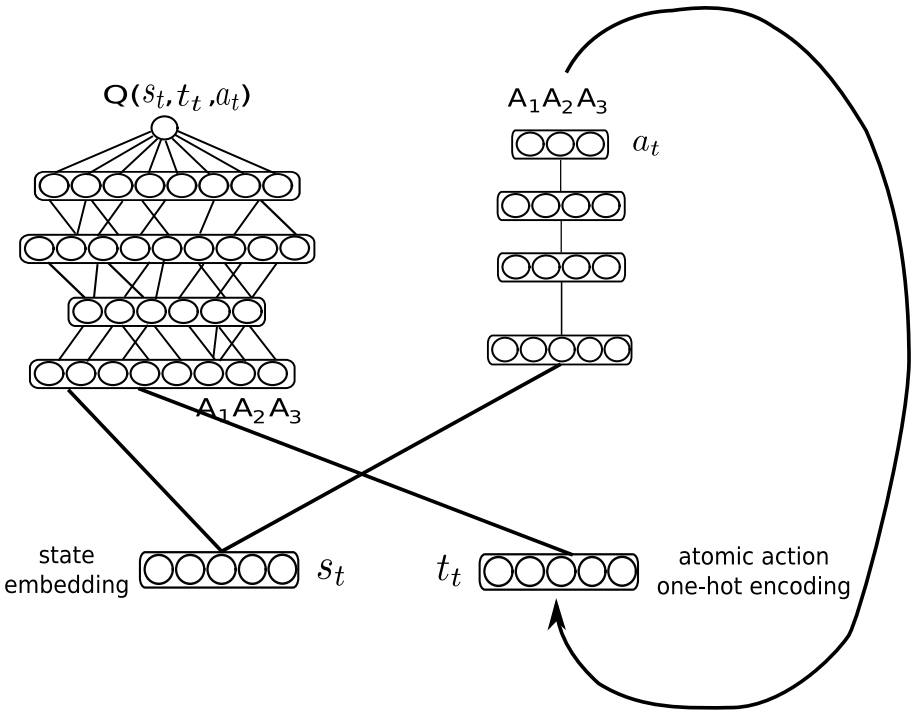
Sorbonne Université
http://www.isir.upmc.fr/personnel/sigaud

Background



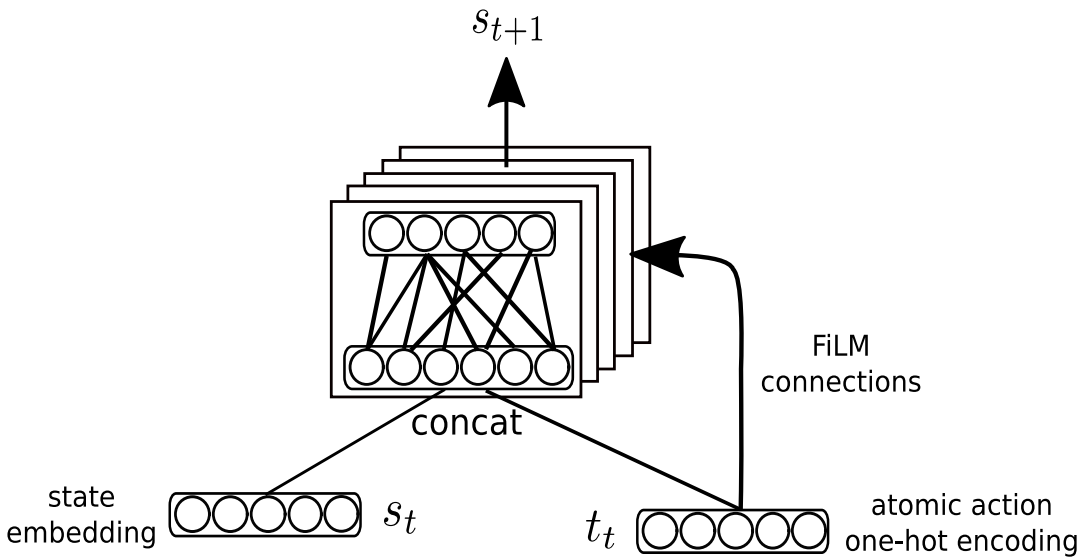
- AlphaNPI is applied to discrete actions
- HILBERT deals with continuous actions
- It learns forward models of the lowest level
- It provides a very sample efficient approach to continuous action HRL

Low-level controller: GC-RL



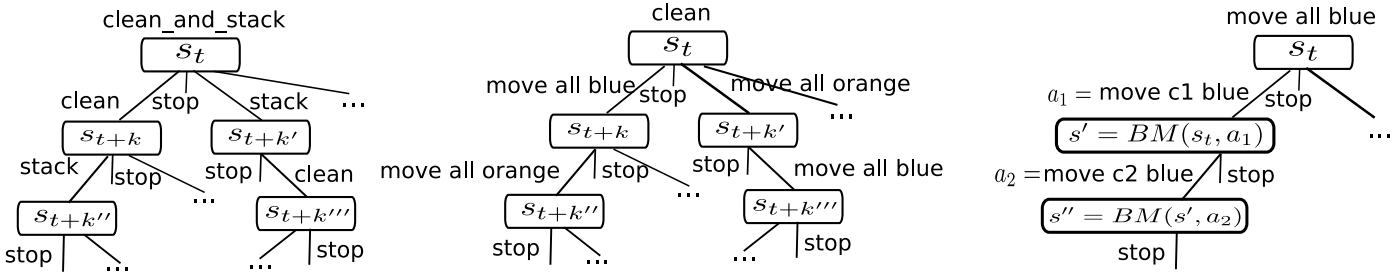
- GC-RL using DDPG (or SAC) + HER

Learning a behavioral model of low level controller



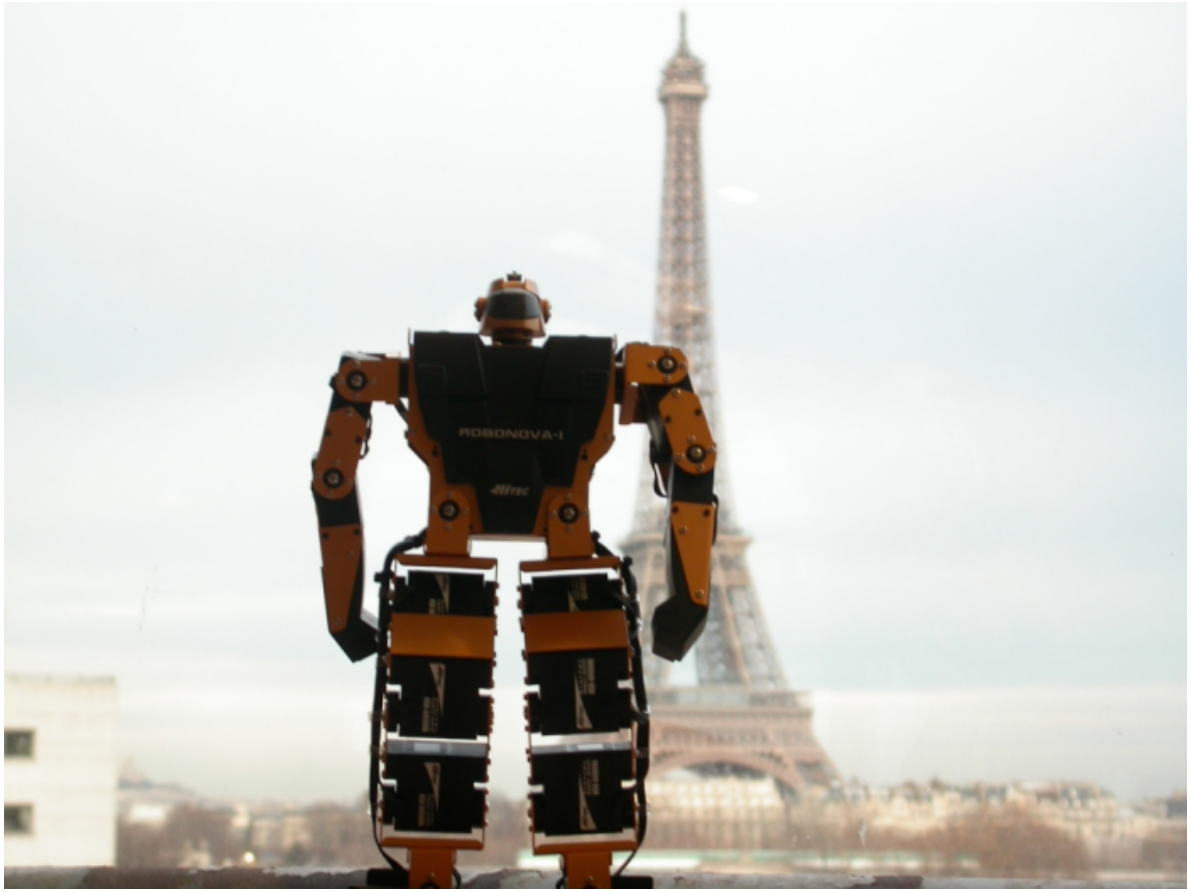
- This is a supervised learning problem
- The FiLM layer improves accuracy

Recursive tree search: continuous action case



- The lowest level stops the recursion
- HILBERT can perform hierarchical planning without rolling the low-level policy
- By using the behavioral model, higher level planning is learned without sampling
- Extremely sample efficient search approach

Any question?



Send mail to: Olivier.Sigaud@upmc.fr

ISIR
INSTITUT
DES SYSTÈMES
INTELLIGENTS
ET DE ROBOTIQUE