Constraining the Size Growth of the Task Space with Socially Guided Intrinsic Motivation using Demonstrations

Sao Mai Nguyen, Adrien Baranes and Pierre-Yves Oudeyer

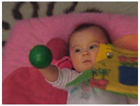
Flowers Team, INRIA Bordeaux - Sud-Ouest, France

IJCAI 2011

Workshop on Agents Learning Interactively from Human Teachers

- Why combine *Intrinsically Motivated* and *Socially Guided* Exploration? limitations for complex environments.
- **S**ocially **G**uided **I**ntrinsic **M**otivation (SGIM) algorithm
- Fishing skill learning
- SGIM learns efficiently in complex environments

Combining *Intrinsically Motivated* and *Socially Guided* Exploration

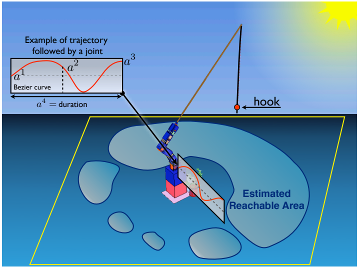
Exploration	strategies	Advantages	Limitations
	Curiosity Driven, Intrinsic Motivation	independent from human, broad task repertoire	unlearnability, unboundedness
	Socially Guided	transfer knowledge from human to robot	teacher's patience, ambiguous human input, ...

Learning skills

- Humans provide an example task demonstration through teleoperation *[Peters and Schaal, 2008; Kormushev et al., 2010]*: **role of the teacher restricted to the initialization phase, learn a specific preset task**.
- Socially Guided Exploration *[Thomaz and Breazeal, 2008; Thomaz, 2006]* : **the representation of the environment is discrete, limited and preset world, few primitive actions possible**.

Learn open set of *skills* in a *large*, *non-preset* and *continuous* environments.

Fishing Experiment Setup



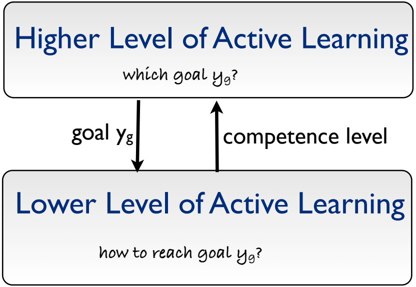
C : configuration/actuator space
Y : operational/task space
A : action space

- 6 dof robot
- $a = (a^1, a^2, ...a^{24})$ define the Bezier curve of the joint angles of the robot. A is 24-D .
- Learn the forward and inverse models of

$$M : A \to Y$$
$$a_t \mapsto y_{t+1}$$

Continuous, complex, large, non-preset environment
Highly **redundant** robot
mapping between **high dimensions**
Fishing rod **can hardly be modeled** mathematically

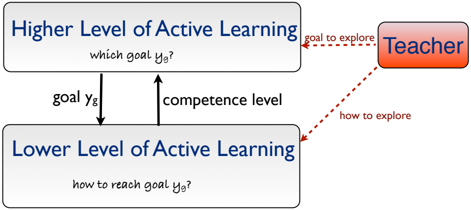
A 2-Levels Architecture



- **Higher level:** actively decides which goal y_g is interesting to explore. Longer time scale.
- **Lower level:** active learning for reaching the goal y_g. Shorter time scale.

Interest level ← entropy, variance, progress...
Ex of intrinsic motivation: SAGG-RIAC algorithm *[Baranes and Oudeyer, 2010]*

Human teacher



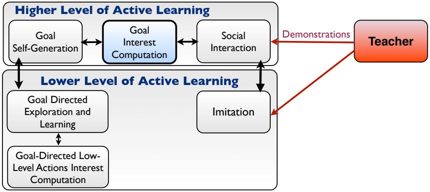
The human teacher can intervene in either level:

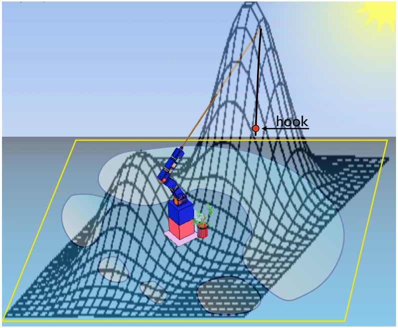
- **Higher level:** bias the robot in its goal exploration
- **Lower level:** helps building a local model, increase competence

Demonstrations:

- of actions and goals
- of unreached goals
- every 150 actions of the learner

Socially Guided Intrinsic Motivation : Higher level

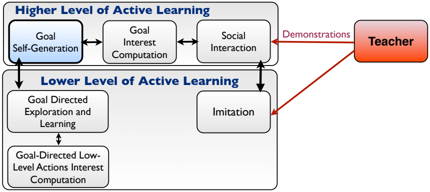


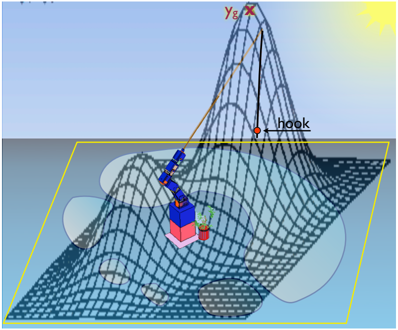


Goal Interest Computation: maps the interest level in the task space Y

.

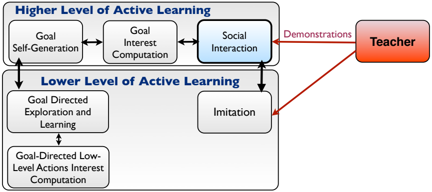
Socially Guided Intrinsic Motivation : Higher level

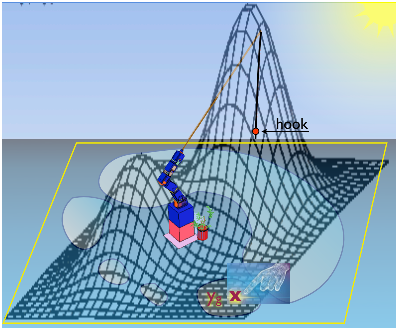




Goal Self-Generation: splits Y into subspaces according to their levels of interest and selects the region where future goals will be chosen.

Socially Guided Intrinsic Motivation : Higher level

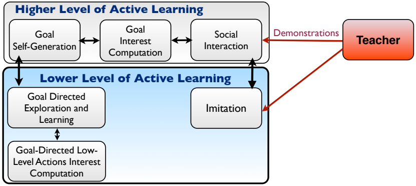




Social Interaction: manages the interaction with the human teacher. The teacher shows an interesting goal to reach.

.

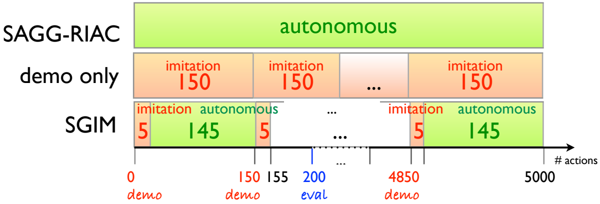
Socially Guided Intrinsic Motivation : Lower Level

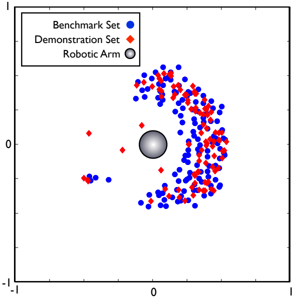


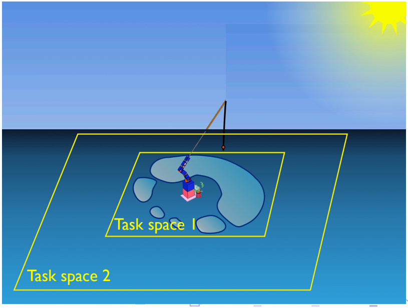
Lower level: active learning for reaching the goal y_g. Shorter time scale.

- *Goal Directed Low-Level Exploration and Learning:* chooses action a_t to reach the goal y_g and creates a model of the world.
- *Goal-Directed Low Level Actions Interest Computation:* measures the competence of the system to reach y_g
- *Imitation:* tries small variations to explore in the locality of a_{demo}

Experimental Protocol

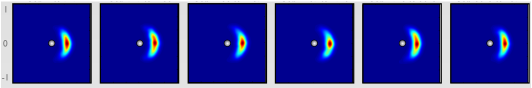




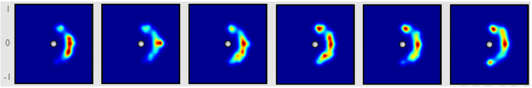


Experimental Results: points explored

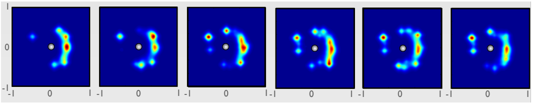
- RANDOM : a natural position $(0.5, 0)$



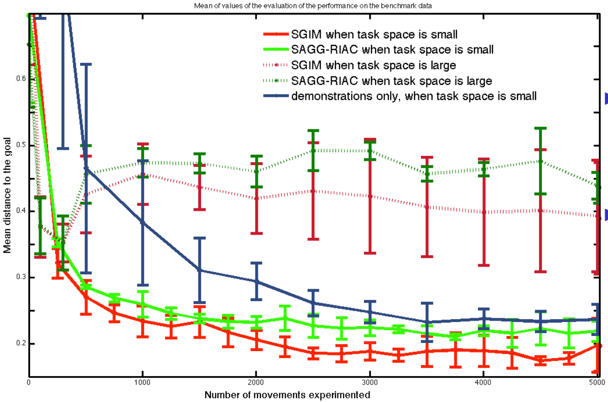
- SAGG-RIAC : increased the explored space, explores more uniformly the reachable space



- SGIM : Isolated subspace explored



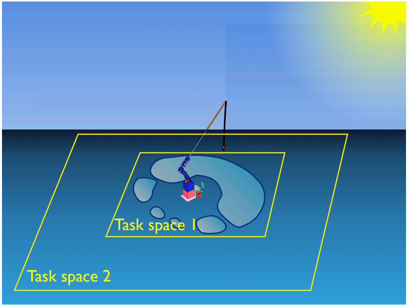
Experimental Results: Precision



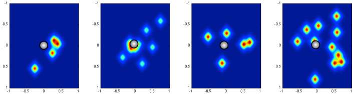
- SGIM > SAGG-RIAC > demo only
- improvement is most significant when the task space size increases

Experimental Results : self-generated goals

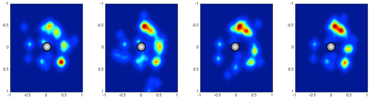
Histogram of the self-generated goals explored by the robot



- SAGG-RIAC: goal points appear disorganised



- SGIM: most goals are within the reachable space



SGIM can differentiate the reachable subspaces

Conclusion

- Efficient knowledge transfer from teacher to learner even in large task space
- Identify the reachable space
- Little intervention from the teacher
- Learn a set of skills, continuous and complex environment

Future work:

- Experiment with physical robots
- Natural interaction between teacher and learner