Improvement Proposals to Intrinsically Motivationnal Robotics

Fabien Hervouet, Eric Bourreau

Improvement Proposals to Intrinsically Motivational Robotics

Fabien Hervouet and Eric Bourreau

LIRMM – Université Montpellier 2 – France

Introduction: why intrinsically motivated robots?

- Key challenge is to identify and implement **low-level mechanisms** that allow a **long-term development**.
- The more low-level these mechanisms are, the more the system can be considered as relevant.
- Inspect and design **scalable task-independent mechanims** that may involve the robot in a **self autonomous skill practice**.

Zoom on SAGG-RIAC

SAGG-RIAC

- Deeply anchored at a sensorimotor level and allows low-level action selection in the **high-dimensional sensorimotor space** for a robot.
- Explores the **self competence acquisition paradigm**: choose sensory regions where it wants to return to instead of sensorimotor regions where it comes from.
- Sensory learning guided by a goal which consists in **mixing exploitation phases and local exploration phases**.
- The purpose of reaching phases is to test the **reliability of the forward motor model** while the purpose of exploration phases is to **improve the inverse model** of the system.
- Exploration phases are triggered when the **reliability is too low**.

$$\kappa(\sigma_i, \gamma, \sigma_f) = max\left(-\sum_{i=1}^{i=|S|} \frac{|\sigma_f.S_i - \gamma.S_i|}{|\sigma_i.S_i - \gamma.S_i|}, \kappa_{max}\right)$$

Curious Developmental Living Loop

input: ξ_r: raw experiments; ξ_g: goal experiments; σ: states;

while *True* **do**
- $start \leftarrow \sigma_t$
- $R \leftarrow argmax(\rho(R_i))$
- $\gamma \leftarrow R$.randomGoal()
- $actions \leftarrow \emptyset$
- **repeat**
  - $action \leftarrow$ getNextAction(σ_t, γ)
  - $actions \leftarrow actions \cup action$
  - execute($action$)
  - $\xi_r \leftarrow \xi_r \cup < \sigma_{t-1}, action, \sigma_t >$
  - **if** $\kappa_t \leq \kappa_{max}$ **then**
    - **for** $i \in \{1..explorationTrials\}$ **do**
      - $action \leftarrow$ randomAction(σ_t)
      - execute($action$)
      - $\xi_r \leftarrow \xi_r \cup < \sigma_{t-1}, action, \sigma_t >$
    - **end**
  - **end**
- **until** $\kappa_t \geq \kappa_{min}$ *or timeout exceeded*

end

$\xi_g \leftarrow \xi_r \cup < start, \gamma, actions, \sigma_t >$

R.reorganizeMemory()

Improvements proposals to SAGG-RIAC

Although we keep the overall operation of the motivationnal living algorithm **SAGG-RIAC** *we draw some improvements we describe here.*

Interest measure

$$\rho(R_i) = LP(R_i) + UCT(R_i)$$

- **Timestamped derivative** that tends to reduce the interest by flattening the interest curve when experiments are very infrequent.

$$LP(R_i) = \frac{\sum_{j=0}^{|R_i|/2} c_j - \sum_{j=|R_i|/2}^{|R_i|} c_j}{\sum_{j=0}^{|R_i|/2} t_j - \sum_{j=|R_i|/2}^{|R_i|} t_j}$$

- **UCT based diversification** measure taking into account the number of experiments conducted in the current region relative to the total number of experiments.

$$UCT(R_i) = c \times \sqrt{\frac{\ln n}{n_i}}$$

Next action to reach a goal

- Using **k-nearest-neighbour** experiments among previously acquired experiments maximizing two criteria.

$$v(\xi_k) = \sum_{j=1}^{j=|S|} |\xi_k.\sigma_i.S_j - \sigma.S_j| + |\xi_k.\sigma_f.S_j - \gamma.S_j|$$

- Generating a **mean action** with respect to actions performed in these experiments.

$$v(a_i) = \frac{\sum_{j=1}^{j=|\xi|} \xi_j.action[i]}{|\xi|}$$

Memory restructuring

- Upgrading **splitting condition** so as to make it dynamic, i.e. correlated with the development of the agent
- Introducing a mechanism for **merging regions** to allow a subsequent restructuring.
- Trying to maximize the absolute value of the difference between the learning progress in the two subregions relative to the current learning progress in the mother region.

$$\mu(R_1, R_2, R) = \frac{|LP(R_1) - LP(R_2)|}{LP(R)}$$

Future works

- Discovering process is underlined by a better **dynamic splitting condition**.
- Forgetting is only about **parts of the segmentation** of the sensorimotor space that used to make sense with a lack of information but that seems inappropriate with more experience.
- We are willing to run precise **parametric experiments** to compare our proposals.
- We are also working at new ways of **defining, evaluating and comparing** performance between different developmental trajectories.

Quick References

- Baranes, A., and Oudeyer, P.-Y. Intrinsically-motivated goal exploration for active motor learning in robots: A case study. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010)* (2010).
- Kocsis, L., and Szepesvári, C.Bandit based monte-carlo planning. In *In: ECML-06. Number 4212 in LNCS* (2006), Springer, pp. 282–293.
- Oudeyer, P.-Y., and Kaplan, F.Intelligent adaptive curiosity: a source of self-development. In *Proceedings of the 4th International Workshop on Epigenetic Robotics* (2004), L. Berthouze, H. Kozima, C. G. Prince, G. Sandini, G. Stojanov, G. Metta, and C. Balkenius, Eds., vol. 117, Lund University Cognitive Studies, pp. 127–130.