



HAL
open science

Improvement Proposals to Intrinsically Motivational Robotics

Fabien Hervouet, Eric Bourreau

► **To cite this version:**

Fabien Hervouet, Eric Bourreau. Improvement Proposals to Intrinsically Motivational Robotics. ICDL-EpiRob: Conference on Development and Learning and Epigenetic Robotics, Nov 2012, San Diego, CA, United States. IEEE, pp.1-2, 2012. lirmm-01066388v2

HAL Id: lirmm-01066388

<https://hal-lirmm.ccsd.cnrs.fr/lirmm-01066388v2>

Submitted on 7 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



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 mechanisms** 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**.

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

Curious Developmental Living Loop

```

input:  $\xi_r$ : raw experiments;  $\xi_g$ : goal experiments;  $\sigma$ : states;
while True do
  start  $\leftarrow \sigma_t$ 
  R  $\leftarrow \text{argmax}(\rho(R_i))$ 
   $\gamma \leftarrow R.\text{randomGoal}()$ 
  actions  $\leftarrow \emptyset$ 
  repeat
    action  $\leftarrow \text{getNextAction}(\sigma_t, \gamma)$ 
    actions  $\leftarrow \text{actions} \cup \text{action}$ 
    execute(action)
     $\xi_r \leftarrow \xi_r \cup \langle \sigma_{t-1}, \text{action}, \sigma_t \rangle$ 
    if  $\kappa_t \leq \kappa_{max}$  then
      for  $i \in \{1..explorationTrials\}$  do
        action  $\leftarrow \text{randomAction}(\sigma_t)$ 
        execute(action)
         $\xi_r \leftarrow \xi_r \cup \langle \sigma_{t-1}, \text{action}, \sigma_t \rangle$ 
      end
    end
  until  $\kappa_t \geq \kappa_{min}$  or timeout exceeded
end
 $\xi_g \leftarrow \xi_r \cup \langle \text{start}, \gamma, \text{actions}, \sigma_t \rangle$ 
R.reorganizeMemory()

```

Improvements proposals to SAGG-RIAC

Although we keep the overall operation of the motivational 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.

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

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

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

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.