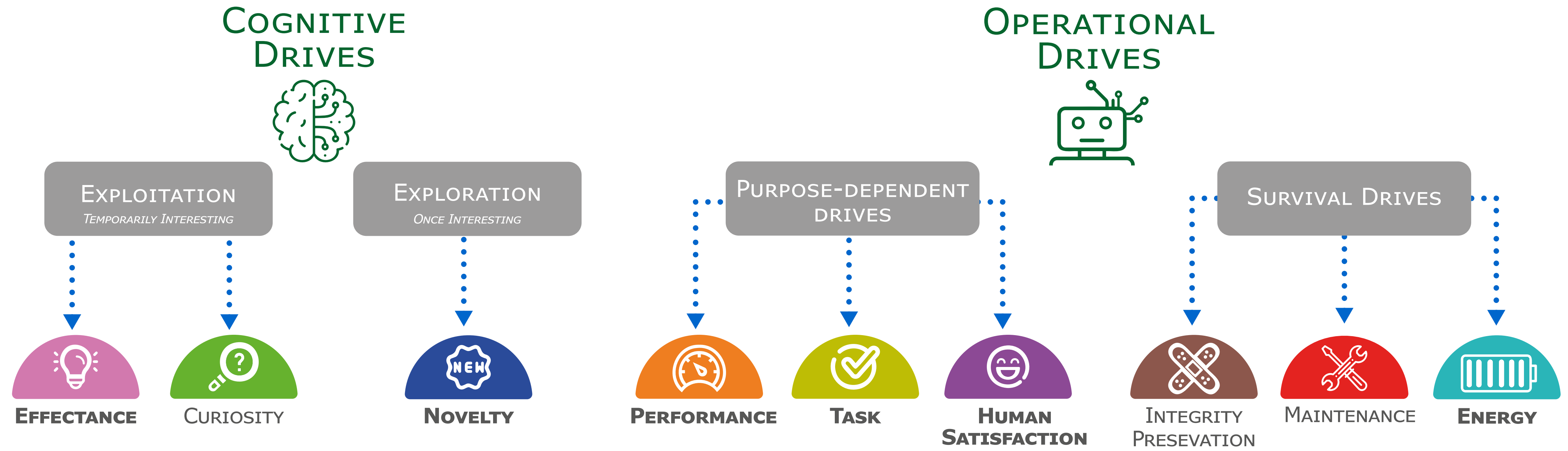
An Autonomous Drive Balancing Strategy for the Design or Purpose in Open-ended Learning Robots

Alejandro Romero, Francisco Bellas, Richard J. Duro / *GII, CITIC Research Center, Universidade da Coruña. Spain*

Introduction

- This paper is concerned with **designing purpose** in autonomous robots for open-ended learning settings [1].
- An approach to the appropriate design and engineering of motivational structures to endow robots with a particular purpose is proposed and tested.
- This approach focuses on a drive structure ($\vec{D}$) and how it can be made to autonomously adapt to changing circumstances.
- We define a simple evolutionary strategy for the **autonomous regulation of multiple drives** in order to optimize long-term operation (adjustment of the drive weights ($\vec{c_i}$)).

Drives and Balancing Strategy



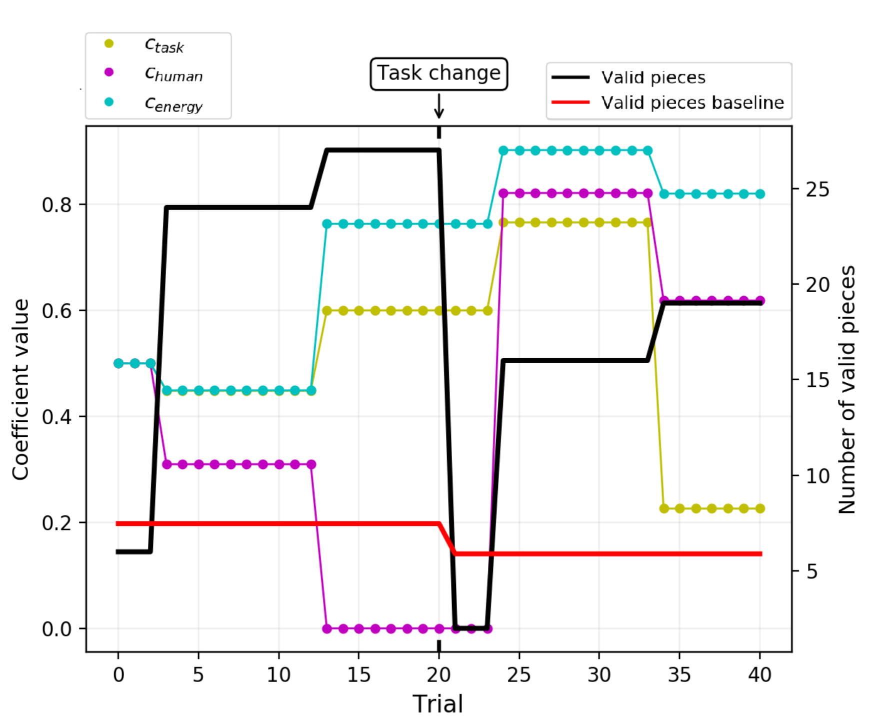
$$\vec{D}=\{c_{effectance}\cdot {}_{cg}D_{effectance},\ c_{novelty}\cdot {}_{cg}D_{novelty},\ c_{performance}\cdot {}_{op}D_{performance},\ c_{task}\cdot {}_{op}D_{task},\ c_{human}\cdot {}_{op}D_{human},\ c_{energy}\cdot {}_{op}D_{energy}\}$$

Algorithm 1 Autonomous drive regulation

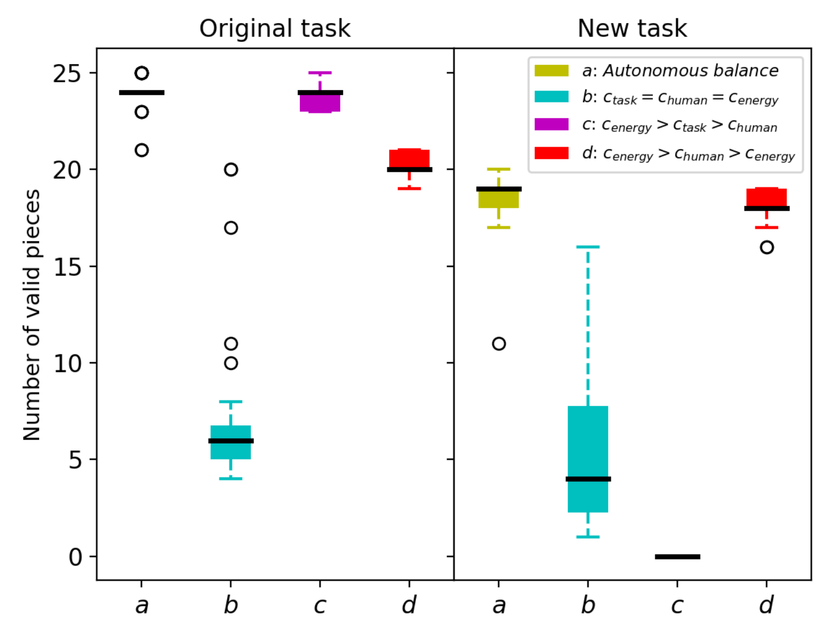
```
e: current experiment
E: set of experiments to execute
pe: probability of exploring versus exploding
v: drive coefficients vector
v': candidate drive coefficients vector
c: current drive coefficient
Pv: performance value associated to v
Pv': performance value associated to v'
 1: pe is initially set by the designer
 2: v is initially set by the designer
 3: Pv ← EvaluatePerformance(v)
 4: for e ϵ E do
 5:     exploration ← ChooseBetweenExplorationExploitation(pe)
 6:     if exploration is True then
 7:         v' is initially set to v
 8:         for c ϵ v' do
 9:             v'[c] ← ChooseNewCoefficientValue()
10:             Pv' ← EvaluatePerformance(v')
11:         if Pv' > Pv then
12:             Pv is set to Pv'
13:             v is set to v'
```

Real Robot Experiment

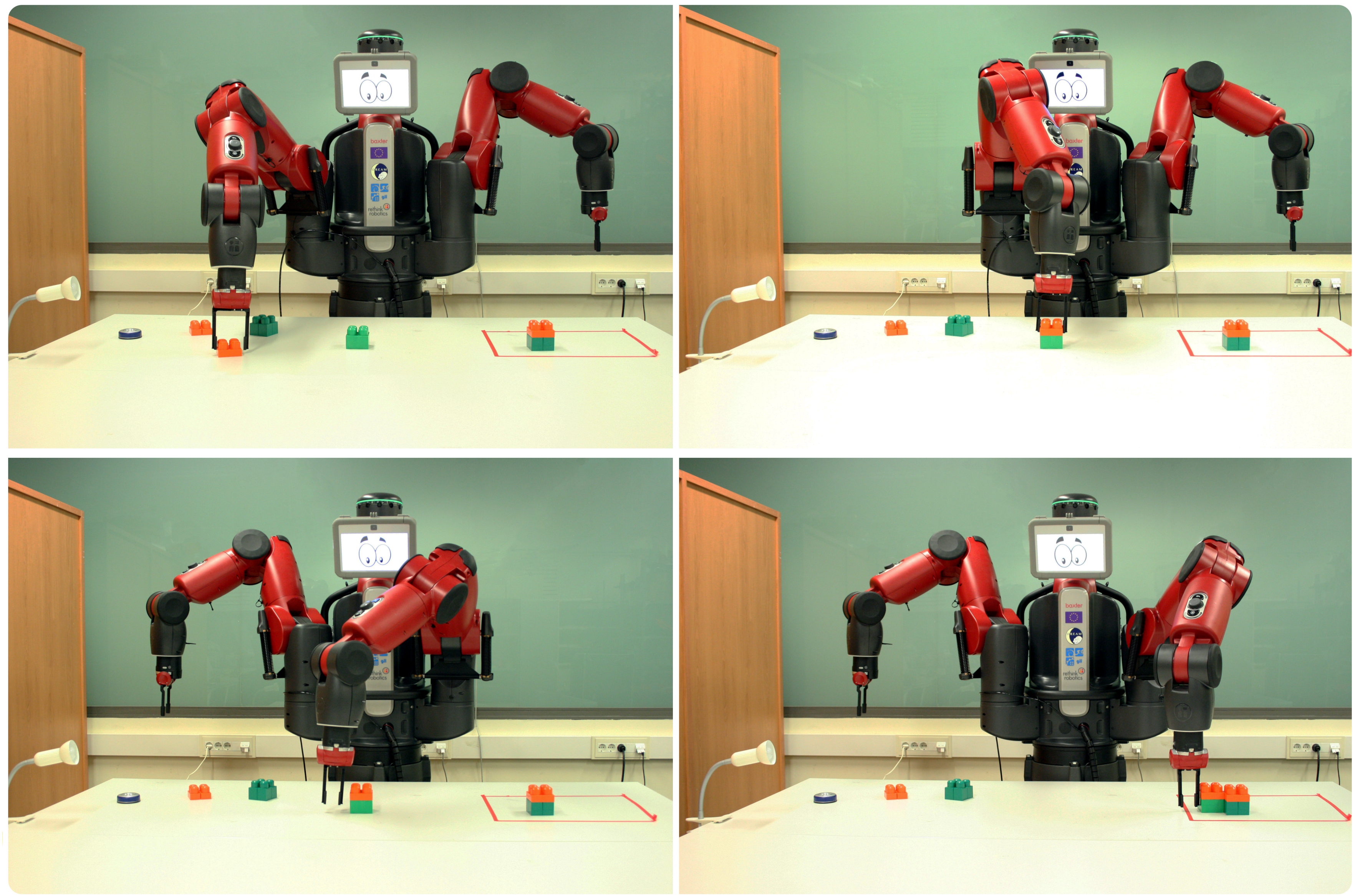
- The drive balancing strategy is applied in a real setup, where a Baxter robot has the PURPOSE of **assembling the maximum number of valid parts** in an environment where human supervisors can interfere with its behavior [2]. The experiment was carried out for 20000 time steps. The proposed strategy acts every 500 time steps, leading to 40 trials associated with them.
- TASK CHANGE: From only assembling parts to assembling parts and showing them to the human supervisor.



Evolution of the coefficient values and the number of valid pieces obtained after one execution of the experiment.



Statistical representation of the robot's performance after the execution of 30 independent experiments.



Experimental setup with the Baxter robot. Assembly task.

Conclusions

- In this paper we have presented the results of the application of an on-line evolutionary mechanism for autonomously and dynamically balancing sets of drives in a real robot setup.
- This approach allows robots that must operate in open-ended learning settings to optimize their satisfaction over time while they adapt to changes in task definitions.

References

[1] S. Doncieux et al. 2018.Open-ended learning: a conceptual framework based on representational redescription. Front. Neurorobot., vol. 12.

[2] A.Romero, F. Bellas, J.A. Becerra, R.J. Duro. 2020. Motivation as a Tool for Designing Lifelong Learning Robots. Integr. Comput. Aided. Eng., vol. 27, no. 4, pp. 353-372.

Acknowledgements

This work was funded by, Ministerio de Ciencia, Innovación y Universidades of Spain/FEDER and Xunta de Galicia by grants RTI2018-101114-B-I00, ED431C 2017/12, CITIC research center grant ED431G 2019/01, and the FPU grant of Alejandro Romero.

CONTACT: alejandro.romero.montero@udc.es