Tampereen yliopisto
Tampere University

Developmental robotics with DFT : From Sensorimotor Contingencies to Autonomous Goals Discovery

Quentin Houbre, Doctoral Researcher
quentin.houbre@tuni.fi
Tampere University, Tampere, FINLAND
September 3, 2020

Research Context

Autonomous skills discovery in an open-ended learning fashion

Inspired by lifelong learning.

Instead of specifying what the robot has to learn, this one autonomously decide what's interesting to explore and learn.

How ?

Intrinsic Motivation [Oudeyer et al., 2007], Deep Reinforcement Learning [de La Bourdonnaye et al., 2018], Goals-based skills Learning [Mannella et al., 2018] ...

Research Context

Advantages :

- No specification of the task learned by design (reward functions).
- Can lead to the emergence of behaviors, thus demonstrating developmental stages (learning to touch before grasping).

Disadvantages :

- Not always brain inspired (role of memory and attention, interactions between cognitive processes).
- Complex architectures often acting as "black box" models.

Research Context

Advantages :

- No specification of the task learned by design (reward functions).
- Can lead to the emergence of behaviors, thus demonstrating developmental stages (learning to touch before grasping).

Disadvantages :

- Not always brain inspired (role of memory and attention, interactions between cognitive processes).
- Complex architectures often acting as "black box" models.

Goal :

Stepping down the level of modelization (Developmental Stages $\rightarrow$ Neuroscience) by identifying and modelling cognitive process that could lead to Open-Ended Learning Skills.

DFT

Not used to explain and reproduce behavioral datas [Schöner et al., 2016]. More about proposing neural mechanisms as a basis for open-ended learning and skill discovery in robotics.

Learning Sensorimotor Contingencies with DFT

Sensorimotor contingencies via Developmental Robotics

First stage of infant's development : motor babbling (primary circular reaction hypothesis) [Piaget and Cook, 1952].
Enactivism [Degenaar and O'Regan, 2017] :

- Cognition arises from the dynamic interaction with the environment.
- Embodiment by linking perception and motor experience together.
- Homeostasis : self-regulation (circular causality of the sensorimotor experience).

The approach

Model motor babbling behavior by associating actions with the sensori outcomes. Two step process by exploration and exploitation of sensorimotor contingencies.

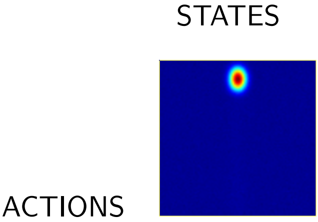
Learning Sensorimotor Contingencies with DFT

How to represent Sensorimotor Contingencies and produce a sequence of actions ?
[Houbre et al., 2020a]
Setting up the experiment inspired by the baby mobile experiment [Watanabe and Taga, 2006].

Exploration

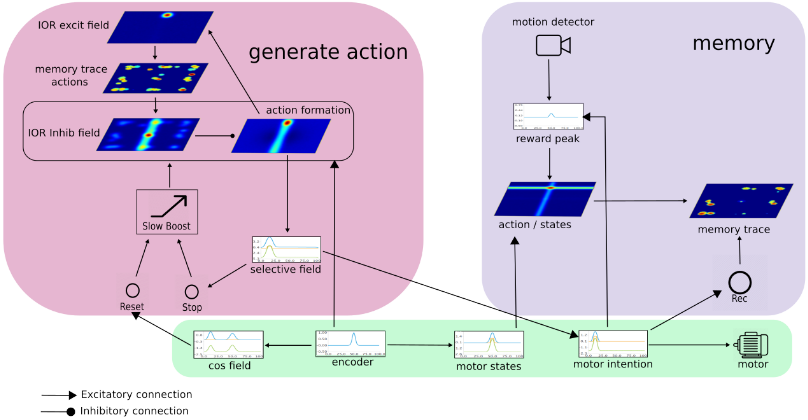
Motor babbling

Generate an action from neural fields. Implementation of an Inhibition of return [Posner et al., 1985] to avoid generating the same action. Fields divided along states (horizontally) and actions (vertically).



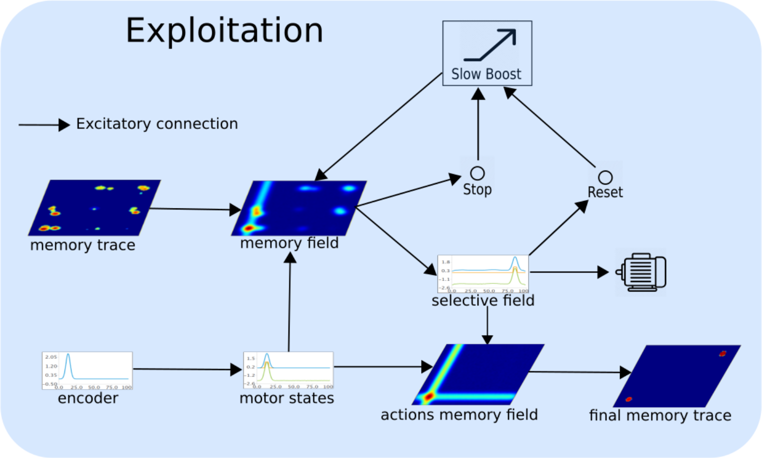
Record the visual outcomes as a peak within a memory trace.

Exploration



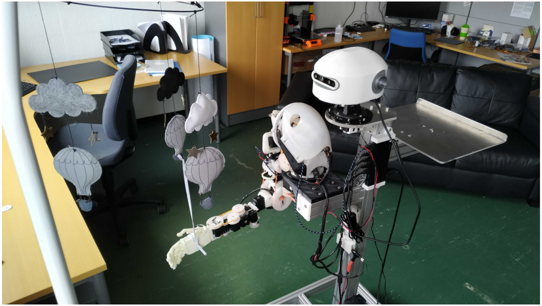
Exploitation

Follow high activation until reaching a stabilized sequence of actions.

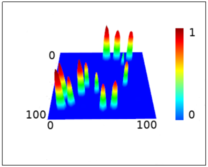


Experiment

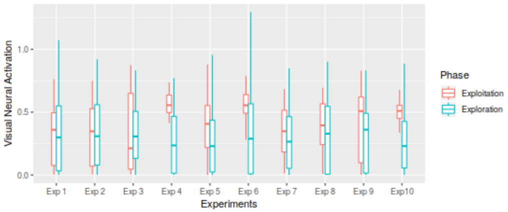
GummiArm [Stoelen et al., 2016]



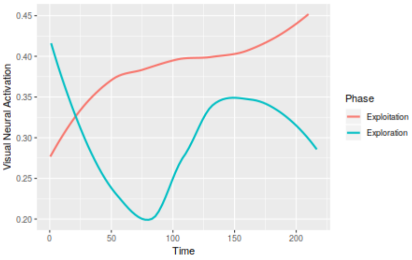
Upper arm roll motor angle [-1;1] spanned over Neural Field [0;100]



Results



Visual neural activation per experiment



Average visual activation in time

Coding

https://github.com/rouzinho

Wiki of the experiment :
https://github.com/rouzinho/DynamicExploration/wiki

EXAMPLES : SlowBoost

Influence from the inhibition-of-return mechanism

Investigate the influence of the strength of the IOR
[Houbre et al., 2020b]

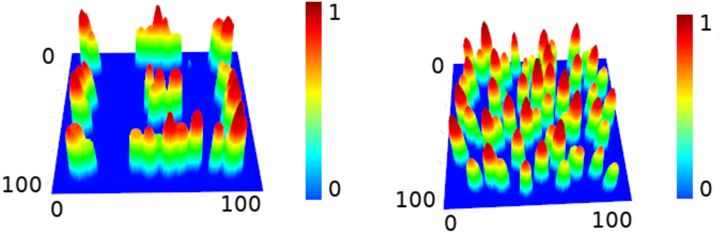


Figure: Left - Memory Trace actions for an exploratory behavior with a strong I-O-R. Right - Exploratory behavior with a weak I-O-R.

Influence from the inhibition-of-return mechanism

Results

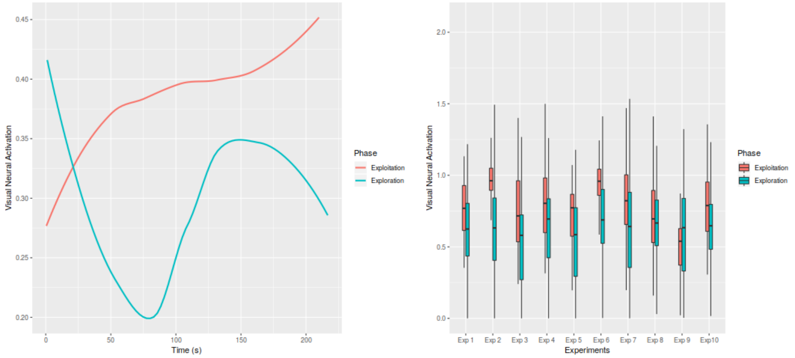


Figure: Left - average visual neural activation for a weak I-O-R (10 experiments). Right - Visual neural activation for a strong I-O-R.

Influence from the inhibition-of-return mechanism

Results

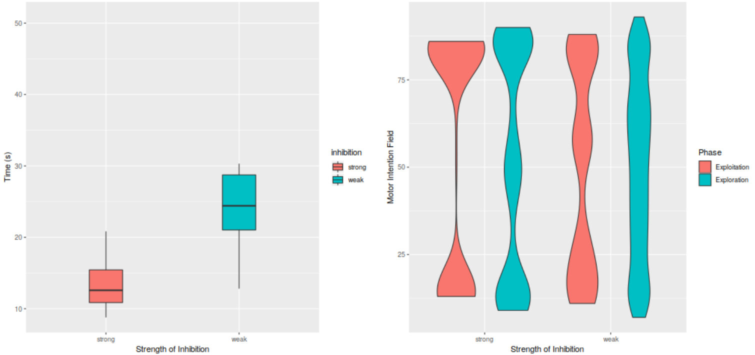


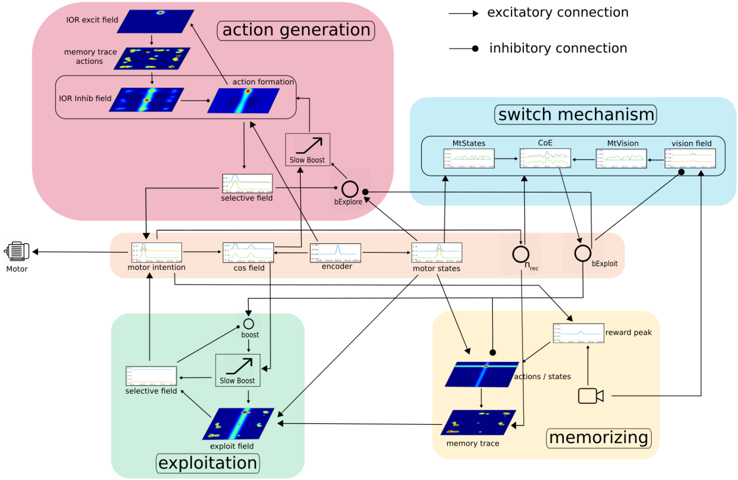
Figure: Left - Elapsed time before reaching a stable sequence of action during the exploitation of 10 explorations with a weak IOR as well as the exploitation of 10 exploration with a strong IOR. Right - Motor distribution during exploration and exploitation for both exploratory behavior.

Autonomous switch between exploration/exploitation

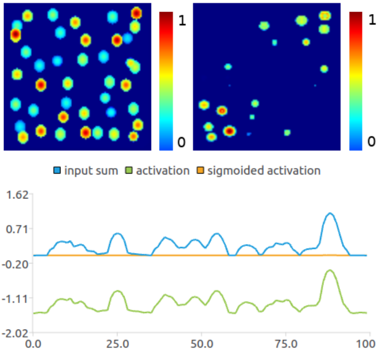
Switch mechanism inspired by recent neuroscience research [Humphries et al., 2012]. The role of basal ganglia for the modulation of the exploration/exploitation stages.
Under certain condition, the increase of dopamine decreases the exploration of new actions :

- A moderate and regulate level of dopamine reduce the exploratory behavior.
- The role of dopamine as a reinforcing signal.

Architecture



Condition of Exploitation



- When a state has never been visited and no reward action was performed, there is no peak of activation within CoE.
- If a state was visited only a few times but a high reward action was performed, a peak emerges from CoE and trigger the exploitation.
- A state visited multiple times with no meaningful action produced will activate the CoE node.

Results

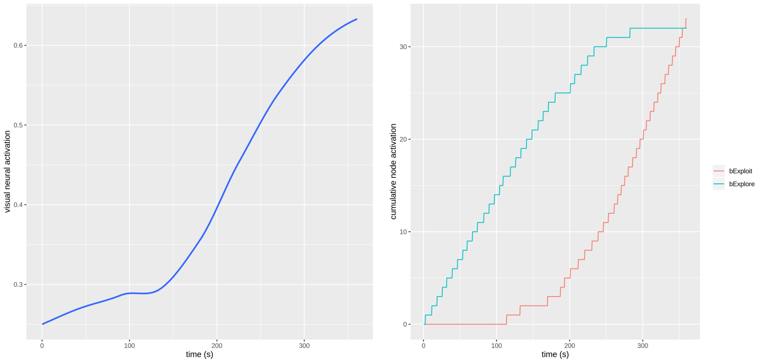
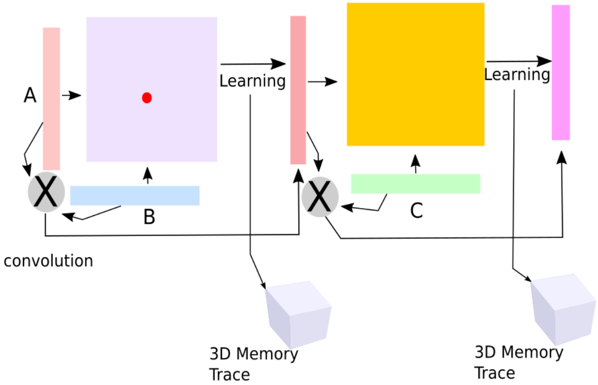


Figure: Average results for 10 experiments. Left : the average visual neural activation over time of 10 experiments is represented by a linear regression. The curve shows an increase of visual activation when the model begins to exploit the sensorimotor contingencies. Right : the sum of the activation nodes bExplore and bExploit (respectively when Exploring and Exploiting) over time for the 10 experiments.

Future Work

Toward the learning of higher-order goals. Formation of multimodal goals by a gain modulation. Inspired by [Schneegans and Schöner, 2012] and [Mahé et al., 2015]



Future Work

How to detect novel goals ? The three layers architecture [Johnson et al., 2009]

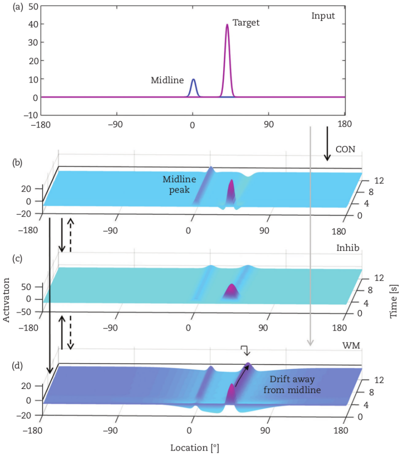
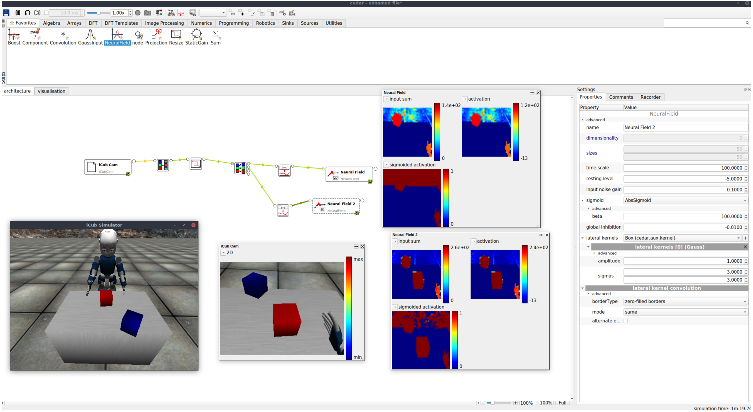


Figure: The three layers architecture, that can act as a novelty detector. Figure taken from [Schöner et al., 2016]

Future Work

Code

Because of current situation : iCub Simulator



Plugins also available to control the iCub end-effector (left or right arm) from a 2D or 3D Neural Field.

Thank You !

de La Bourdonnaye, F., Teulière, C., Triesch, J., and Chateau, T. (2018).
Learning to touch objects through stage-wise deep reinforcement learning.
In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1–9.

Degenaar, J. and O'Regan, J. K. (2017).
Sensorimotor Theory and Enactivism.
Topoi, 36(3):393–407.

Houbre, Q., Angleraud, A., and Pieters, R. (2020a).
Exploration and exploitation of sensorimotor contingencies for a cognitive embodied agent.
In *ICAART (2)*, pages 546–554.

Houbre, Q., Angleraud, A., and Pieters, R. (2020b).
An inhibition of return mechanism for the exploration of sensorimotor contingencies.

In *1st International Conference on Human-Machine System (2)*.

Humphries, M., Khamassi, M., and Gurney, K. (2012).
Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia.
Frontiers in Neuroscience, 6:9.

Johnson, J. S., Spencer, J. P., Luck, S. J., and Schöner, G. (2009).
A dynamic neural field model of visual working memory and change detection.
Psychological science, 20(5):568–577.

Mahé, S., Braud, R., Gaussier, P., Quoy, M., and Pitti, A. (2015).
Exploiting the gain-modulation mechanism in parieto-motor neurons: Application to visuomotor transformations and embodied simulation.
Neural Networks, 62:102–111.

Mannella, F., Santucci, V. G., Somogyi, E., Jacquey, L., O'regan, K. J., and Baldassarre, G. (2018).
Know your body through intrinsic goals.
Frontiers in Neurorobotics, 12(July):1–17.

Oudeyer, P.-Y., Kaplan, F., and Hafner, V. V. (2007).
Intrinsic motivation systems for autonomous mental development.
IEEE transactions on evolutionary computation, 11(2):265–286.

Piaget, J. and Cook, M. (1952).
The origins of intelligence in children, volume 8.
International Universities Press New York.

Posner, M. I., Rafal, R. D., Choate, L. S., and Vaughan, J. (1985).
Inhibition of return: Neural basis and function.
Cognitive Neuropsychology, 2(3):211–228.

Schneegans, S. and Schöner, G. (2012).

A neural mechanism for coordinate transformation predicts pre-saccadic remapping.
Biological cybernetics, 106(2):89–109.

Schöner, G., Spencer, J., and Group, D. F. T. R. (2016).
Dynamic Thinking: A Primer on Dynamic Field Theory.
Oxford University Press.

Stoelen, M. F., Bonsignorio, F., and Cangelosi, A. (2016).
Co-exploring actuator antagonism and bio-inspired control in a printable robot arm.
In *From Animals to Animats 14*, pages 244–255, Cham. Springer International Publishing.

Watanabe, H. and Taga, G. (2006).
General to specific development of movement patterns and memory for contingency between actions and events in young infants.
Infant Behavior and Development, 29(3):402–422.