Balancing Exploration and Exploitation : A Neurally Inspired Mechanism to Learn Sensorimotor Contingencies

Quentin Houbre, Alexandre Angleraud, and Roel Pieters

Tampere University, Tampere, Finland
quentin.houbre@tuni.fi

Abstract. The learning of sensorimotor contingencies is essential for the development of early cognition. Here, we investigate how such process takes place on a neural level. We propose a theoretical concept for learning sensorimotor contingencies based on motor babbling with a robotic arm and dynamic neural fields. The robot learns to perform sequences of motor commands in order to perceive visual activation from a baby mobile toy. First, the robot explores the different sensorimotor outcomes, then autonomously decides to utilize (or not) the experience already gathered. Moreover, we introduce a neural mechanism inspired by recent neuroscience research that supports the switch between exploration and exploitation. The complete model relies on dynamic field theory, which consists of a set of interconnected dynamical systems. In time, the robot demonstrates a behavior toward the exploitation of previously learned sensorimotor contingencies and thus selecting actions that induce high visual activation.

Keywords: Sensorimotor Contingencies, Dynamic Field Theory, Neural Networks, Developmental Robotics

1 Introduction

The acquisition of early sensorimotor behavior is widely studied in robotics to understand human cognition. In this work, we take insights from neuroscience and developmental psychology to propose a model of the early sensorimotor development driven by neural dynamics [1].

To do so, this paper relies on the field of developmental robotics [2]. Indeed, the principles of developmental processes are a key to better understanding human intelligence. Modelling cognition respecting these principles would theoretically allow a robot to learn and evolve by following the same stages as an infant. In addition, research in developmental psychology and neuroscience are demonstrated to be fundamental for the cognitive abilities of robots. The theory of Sensorimotor Contingency [3] states that sensing is a form of action. The experience of perception (vision, touch, hearing, etc.) is a result of a close interaction with the environment rather than the activation of an internal model of the world through sensing. For example, developmental psychologists such as

Piaget [4] were the first to formulate the "primary circular-reaction hypothesis" where children generate "reflexes" and these reflexes change (even slightly) when they produce an effect on the children's environment. Several models propose a reproduction of this developmental stage through motor babbling. For example, Demiris and Dearden [5] propose to associate motor commands with the sensori outcomes and demonstrate the possibility to use that experience for imitation. Closer to our approach, Mahoor et al. propose a neurally plausible model of reaching in an embodied way through motor babbling [6]. To do so, they use a set of three interconnected neural maps to learn the dynamical relationship between the robot's body and its environment. In order to develop a motor babbling behavior, the robot must be able to autonomously generate motor commands and observe the outcomes in a closed loop. More recent work investigates intermodal sensory contingencies to see if self-perception could lead to causality interpretation [7]. By using a novel hierarchical bayesian system, the researchers were able to combine proprioceptive, tactile and visual cues together in order to infer self-detection and object discovery. In addition to these results, this study states the importance of interacting with the environment as an active process. Indeed, the robot refines its own model of the world and thus can infer more knowledge by continuously interacting with it. In this work, we propose to address the learning of sensorimotor contingencies by extending the previous work [8] and endow the architecture with a mechanism that autonomously switch between exploration and exploitation.

In the literature, the exploration/exploitation architecture has been a challenge for years, spreading beyond the fields of robotics and computer science to become a multidisciplinary issue [9]. The exploration/exploitation trade-off is widely investigated with reinforcement learning. A classical way to deal with this issue is the *greedy* approach, where a probability determines when to explore or exploit [10]. Work related to learning from demonstration proposes to use a confidence metric for learning a new policy [11]. In that case, when the confidence level reaches a certain threshold, the agent asks for a new demonstration. Other research proposed an architecture for learning sensorimotor contingencies based on the past rewards observed by the robot [12]. The action selection algorithm can be seen as an exploration/exploitation trade-off that chooses to explore a new action if this one was never taken before. Then, the algorithm assigns a probability to an action depending on the reward observed. This work rests close to our approach by how they are representing and selecting actions. Despite demonstrating significant results, these attempts rarely take inspiration from the human brain, even less on a neural level. This contribution proposes a method to tackle the exploration/exploitation trade-off based on neural dynamics and is inspired by recent progress in neuroscience. Indeed, Cohen and colleagues [13] suggested that two neuromodulators (acetylcholine and norepinephrine) can be a signal for a source of certainty or uncertainty and thus a factor influencing the trade-off. In recent works, the role of the basal ganglia indicates a modulation of the exploration/exploitation trade-off through dopaminergic control [14]. Interestingly, they advance that the level of dopamine influences the choice of an

action. Specifically, that under certain conditions the increase of the dopamine level decreases the exploration of new actions.

Dynamic neural fields are used in this paper to explore the environment and predict changes by exploiting newly learned associations. Dynamic Field Theory (DFT) is a new approach to understand cognitive and neural dynamics [15]. DFT is suitable to deliver homeostasis [16] to the architecture, providing an intrinsic self-regulation of the system. For the learning of sensorimotor contingencies, the approach allows various ways of learning. The most basic learning mechanism in DFT is the formation of memory traces of positive activation of a Dynamic Neural Field [17]. The use of memory trace fields will support the learning of sensorimotor associations. Usually, the learning of sequences within DFT is achieved by a set of ordinal and intention nodes [18]. This requires to know *a priori* the type of content to learn (intention nodes) and the finite number of actions (ordinal nodes). Due to the nature of sensorimotor contingencies, it is not possible to predict how many actions and which one of them would lead to the highest neural activation. However, it is still possible to implement reinforcement learning within DFT [19] but only by discretizing the actions space into nodes. This contribution proposes to extend the literature of DFT by introducing an exploration/exploitation architecture without knowing the number of actions to learn beforehand.

In this paper, we propose a model to learn sensorimotor contingencies based on a neural mechanism that allow the autonomous switch between the exploration/exploitation stage. We set up a robotic experiment where a humanoid robotic arm [20] is attached to a baby mobile toy with a rubber band. The robot then learns how to move its arm in order to get a visual feedback. Before performing an action, the robot autonomously decides to explore or exploit the sensorimotor experience based on the neural mechanism inspired by recent research in neuroscience. The proposed architecture is self-regulated and is driven by Dynamic Neural Fields in a closed loop, meaning the actions influence future perceptions.

2 Dynamic Field Theory

Dynamic Field Theory is a theoretical framework that provides a mathematical way to model the evolution in time of neural population activity [15]. It demonstrated its ability to model complex cognitive processes [21]. The core elements of DFT are Dynamic Neural Fields (DNF) that represent activation distributions of neural populations. A peak of activation emerges as a result of a supra-threshold activation and lateral interactions within a field. A DNF can represent different features and a peak of activation at a specific location corresponds to the current observation. For instance, a DNF can represent a visual color space (from Red to Blue in a continuous space) and a peak at the "blue location" would mean that a blue object is perceived. Neural Fields are particularly suitable to represent continuous space.

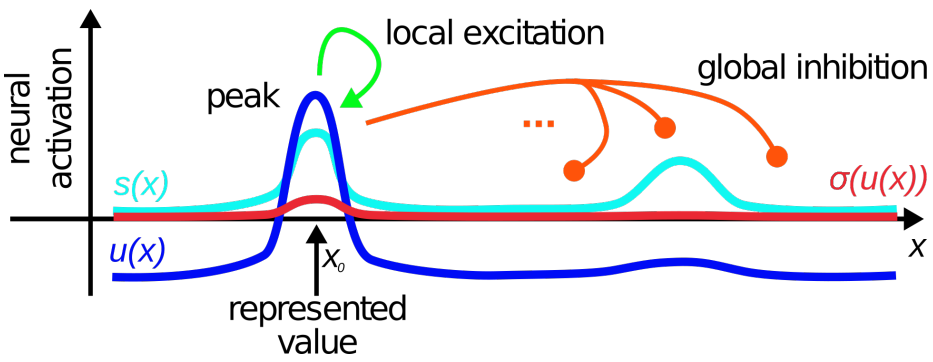

Fig. 1: A dynamic neural field activation spanned accross the feature x.

Dynamic Neural Fields evolve continuously in time under the influence of external inputs and lateral interactions within the Dynamic Field as described by the integro-differential equation :

$$\tau \dot{u}(x,t) = -u(x,t) + h + S(x,t) + \int \sigma(u(x',t))\omega(x-x')dx', \tag{1}$$

where h is the resting level ($h < 0$) and $S(x,t)$ is the external input. $u(x,t)$ is the activation field over a feature dimension x at time t and τ is a time constant. An output signal $\sigma(u(x,t))$ is determined from the activation via a sigmoid function with threshold at zero. This output is then convoluted with an interaction kernel ω that consists of local excitation and surrounding inhibition [22]. The role of the Gaussian kernel is crucial since different shapes influence the neural dynamics of a field. For example, local excitatory (bell shape) coupling stabilizes peaks against decay while lateral inhibitory coupling (Mexican-hat shape) prevents the activation from spreading out along the neural field. Depending on the coupling between local excitation and global inhibition, a neural field can operate on several modes. In a self-stabilized mode, peaks of activation are stabilized against input noise. In a self-sustained mode, the field retains supra-threshold peaks even in the absence of activation. A selective mode is also possible through a lateral inhibition that allows the emergence of a single peak of activation. By coupling or projecting together several neural fields of different features and dimensions, DFT is able to model cognitive processes. While neural fields are the core of the theory, other elements are also essential to our work.

Dynamic neural nodes are essentially a 0-dimensional neural field and follow the same dynamic:

$$\tau \dot{u}(x,t) = -u(x,t) + h + c_{uu} f(u(t)) + \sum S(x,t). \tag{2}$$

The terms are similar to a Neural Field except for c_{uu} which is the weight of a local nonlinear excitatory interaction. A node can be used as a boost to another Neural Field. By projecting its activation globally, the resting level of the neural field will rise allowing to see the rise of activation peaks.

Finally, the memory trace is another important component of DFT:

$$\dot{v}(t) = \frac{1}{\tau_+}(-v(t) + f(u(t)))f(u(t)) + \frac{1}{\tau_-}(-v(t)(1 - f(u(t))), \tag{3}$$

with $\tau_+ < \tau_-$. A memory trace in DFT has two different time scales, a build up time τ_+ that corresponds to the time for an activation to rise in the memory and a decay time τ_- which is the time decay of an activation.

3 Model

In this paper, the model autonomously adopts a motor babbling behavior in order to learn sensorimotor contingencies. The system explores the motor space by linking together motor commands and observed outcomes. Then, it autonomously decides when to balance the exploratory and exploitatory behaviors. For more clarity, we split the explanation of the model between the exploration, the exploitation process and the switch component. In this paper a single degree of freedom is considered (the upper arm roll joint) to activate the robot's arm. Each two dimensional field is divided by states and actions of that joint along the horizontal and vertical dimension respectively. Representing neural fields that way allows to represent the current state of the upper arm roll horizontally and the action to be selected (future state of the joint) vertically. Each dimension is defined between the interval [0;100] and represents a motor angle within a range of [-1;1]. For instance, if a peak of activation emerges at position [25;75], this means at state 25 (motor angle of -0.5) the action 75 (angle of 0.5) is selected. The use of a single degree of freedom is a current limitation of the model, although we will discuss about the possibility to use the complete arm kinematics in section 5.

3.1 Exploration

In order to explore the environment, the model must first generate motor commands and associate them with the perceived outcomes.

Action Generation Regarding the formation of motor commands, the model relies on neural dynamics. Since the two dimensional neural fields are represented by states (horizontally) and actions (vertically), the principle is the following : a zero dimensional memory trace (slow boost module) slowly increases the resting level of a neural field (action formation field) until a peak of activation emerges (Figure 2). This particular memory trace (Equation 4) rises activation when the node *bExplore* is active, and resets the activation when an action has been performed (CoS field).

$$\dot{v}(t) = \frac{1}{\tau_+}(-v(t) + f(u(t)))f(u(t)) + \sigma(n_{cos})\big[\frac{1}{\tau_-}(-v(t)(1 - f(u(t)))\big]. \tag{4}$$

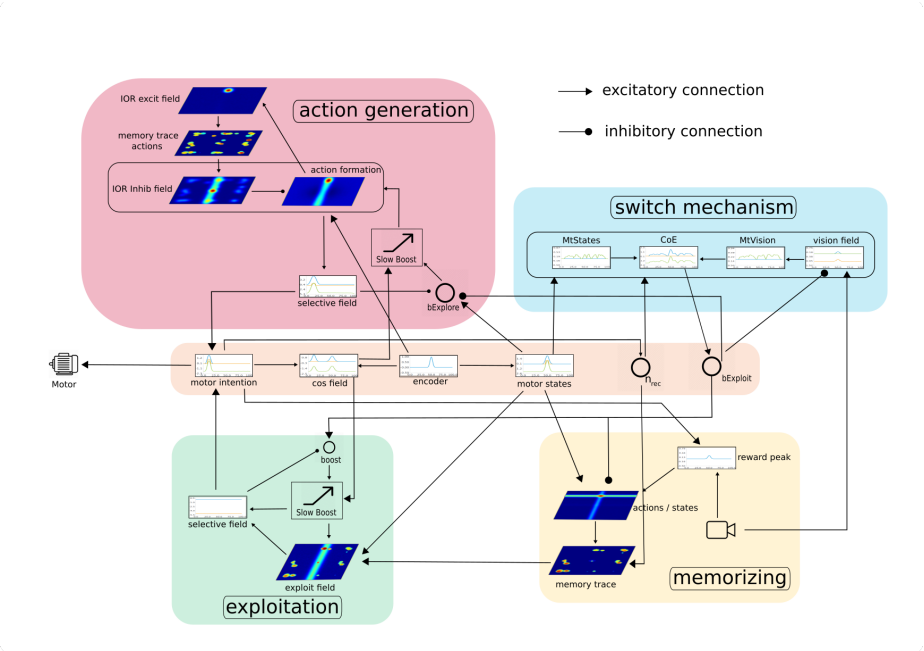

Fig. 2: General Architecture. Here, the model follows an exploration phase as seen by the activation within the different neural fields. The connections point directly to a field or to a group of fields for more clarity. However, the Condition of Exploitation (CoE) field does not receive input from the motor states field. The neural dynamics drive entirely the exploration/exploitation of sensorimotor contingencies.

The Condition of Satisfaction (CoS) field signals when an action is over, in other words, it indicates when the motor state corresponds to the action just taken. For clarity, Figure 2 only shows the connection between the CoS field and the slow boost without adding the (n_{cos}) node. But in practice, the CoS field projects activation to (n_{cos}) then activates the dynamics of the two slow boost memory traces. During the rise of an activation, an inhibition of return takes place in order to avoid generating the same action twice. This mechanism is well studied, especially regarding visual attention [23], [24], where immediately after an event at a peripheral location, there is facilitation for the processing of other stimuli near that location. Therefore, when a peak reaches the threshold of activation within the AF field, the stimuli is projected and recorded to a memory trace before being projected again as an inhibitory input. Following, the activation within the AF field is transmitted to a motor intention field via a selective field. The memory section of our model associates a visual stimuli to the action performed.

Memory Association The perception of visual stimuli is done through the camera inside the robot's head. A motion detector substracts two consecutive images and applies a threshold to observed the changed pixels. The result is then scaled from 0 to 1 and serves as input to the reward peak module. This gathers the actions being executed with the value of the visual stimuli. In practice, it forms a Gaussian curve centered at the action location within the motor intention field with an amplitude corresponding to the visual stimuli currently perceived. If the stimulus is strong enough, a peak of activation appears within the actions/states field. During the execution of an action, a memory trace keeps track of perceived stimuli (Equation 5).

$$\dot{v}(t) = \sigma(n_{rec})\left[\frac{1}{\tau_+}(-v(t) + f(u(t)))f(u(t)) + \frac{1}{\tau_-}(-v(t)(1 - f(u(t)))\right]. \quad (5)$$

This last memory trace slightly differs from the Slow Boost since the dynamics evolve only when the n_{rec} node is active. This allows the storing of perceptions only during an action, when a peak appears in the motor intention field. Without the presence of a n_{rec} node to control the activation, and due to the nature of the experiment, the memory trace would store stimuli that do not necessarily correspond to the action currently performed. The next part describes the exploitation of the sensorimotor associations.

3.2 Exploitation

The exploitation behavior select an action according to the current motor state. Given a motor position, the model encoded the result of actions taken during exploration. Here, a choice is made by selecting the action with the highest peak encoded in memory. Then, the exploitation of the sensorimotor contingencies is straightforward : the model follows the "path" of high activation along the memory trace and executes the corresponding actions. To do so, the exploit field receives input from the memory trace and the current motor state. A slow boost (Equation 4) rises the resting level in that field until a peak reaches the supra-threshold activation (Figure 3).

Following, rising the resting level of the exploit field triggers the emergence of the best action for the current motor state. As presented in the previous section, the best action is the one producing the biggest/most important changes in the environment. So, the model executes the best action, updates its position (motor state), rises the resting level and executes the best action again. By doing so, a pattern appears and produces the same sequence of actions which generates the highest visual neural activation for the robot. The last part of the model introduces the balance mechanism, and how it enables autonomously switching between exploration and exploitation.

3.3 Balancing Exploration and Exploitation

As presented in the introduction, the exploration/exploitation trade-off is not trivial to approach. In this work, we propose a neural mechanism inspired by

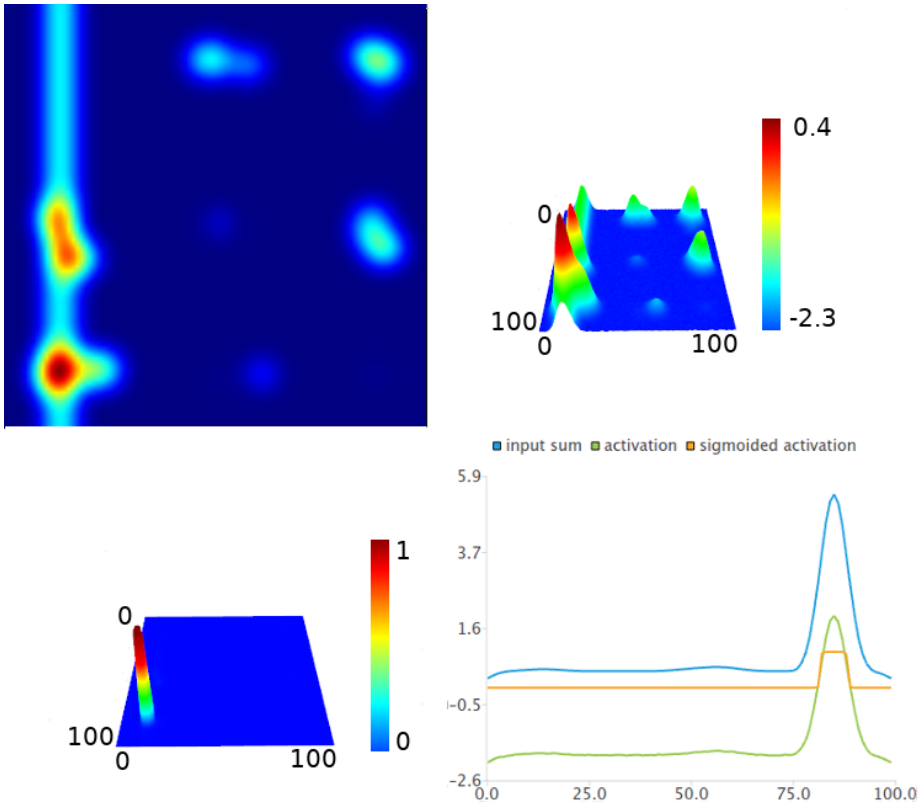

Fig. 3: Snapshot of the exploitation stage when rising the resting level of the exploit field. Top row represents the activation within the exploit field. Bottom left is a 3D view of the sigmoid activation. Bottom right is the selective field for the case where multiple activations would appear.

neuroscience to address this issue. It is already known that the basal ganglia plays a major role in learning [25]. Moreover, recent discoveries [14] suggest that the basal ganglia influences the decision to explore or exploit one's own experiences. More precisely, a moderate and regular level of dopamine leads to a more exploitative behavior. Two functions of the basal ganglia have been developed here: its role as a reinforcing signal and its influence on the choice of a strategy.

More precisely, the reinforcing signal is seen as an excitatory peak of activation when the robot explores an action with a high visual outcome. To do so, we use a memory trace (MtVision) that takes as input the vision field (suprathreshold activation at the current state location only when a visual stimuli happens) and the n_{rec} node. By doing so, an activation peak rises at the location of the current motor state when an action is being performed. The principal

advantage is to signal at which state the robot perceived a high stimulus. For example, if the robot goes through the same state many times, but does not perform any meaningful action from that state, then there is no activation at that location.

To select a strategy, the model delivers a small excitatory signal each time an action is explored. The goal here is not to accurately model the findings from [14] but rather see if a regular input of dopamine could effectively lead toward an exploitative behavior. A memory trace (MtStates) imitates a regular and moderate flow of dopamine to keep track of the number of times a state has been visited. This field receives input from the motor state briefly before selecting a new action. Independently from any visual change, an activation peak slowly rises during an action (n_{rec} active) at the current state. If a state has been visited several times, then the activation at that location will be high.

So, MtVision delivers a punctual activation at a current state location when a visual stimuli happens and MtStates regularly increases the activation at the current state location. We then project these two memory traces to the Condition of Exploitation field (CoE). When a peak emerges within that field, then the *bExploit* node is active and triggers the exploitation process.

To resume the switch mechanism :

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

The rest of the processing is rather simple : when *bExploit* is active, it activates the boost from exploitation. Simultaneously, *bExplore* receives inhibition to avoid generating an action. The field actions/states from the memory part is also inhibited to bypass recording the exploited action. The same process takes place for the vision field in the confidence section. The exploited action must not influence MtVision by increasing an activation. Indeed, an experience with new stimuli is considered highly rewarding and strengthened when it is first encountered. To be closer to reality, a decay mechanism could be introduced when the same stimuli is processed, however, this did not bring significant changes to the results. The next section presents the experimental results of our model with a humanoid robotic arm.

4 Results

Experiments were conducted with a humanoid robotic arm [20] and a camera. The robot is a 3D printed arm with 7 degrees of freedom (+2 for the head). In this settings, only a single degree of freedom is utilized (upper arm roll). A

rubber band is attached from the palm of the hand to one of the moving toys in the baby mobile. The camera (intel RealSense D435) mounted inside the custom designed head [26] is used for visual perception (i.e. motion detection). The toys hanging on the baby mobile are within the visual field of the camera whereas the arm is out of sight. The experiments consist of a set of 10 trials lasting 350 seconds each. Each memory trace is cleared before launching a new trial.

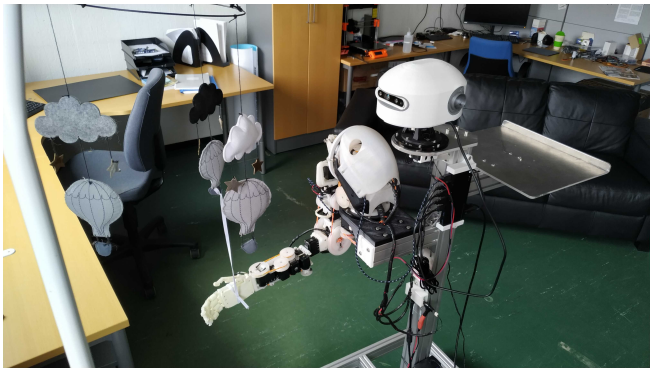

Fig. 4: Set up of the GummiArm, with the rubber band linking the robot's palm with the babymobile toy.

Regarding the visual neural activation (Figure 5-left), the linear regression shows a rise after 150 seconds. This is approximately the moment when the architecture begins to exploit the experience already gathered. At that time, the robot already visits states with high value, meaning when actions with a high visual activation are selected. On a neural level, this is the moment when the Condition of Exploitation field emits a peak to activate the *bExploit* node and thus inhibits the *bExplore* node.

The activation of these nodes in time provides a clear representation of when the robot is exploring the environment or exploiting the gathered experience (Figure 5-right). Despite the fact that the activation in time of these nodes is averaged between 10 trials, there is almost no overlap (no activation from both nodes at the same instant). The time activation demonstrates a clear tendency toward an exploitation behavior after 250 seconds. Most importantly, the frequency at which the *bExploit* node is active corresponds to the increase of visual neural activation seen before. Indeed, the exploitation phase does lead to a gain of visual reward and the switch between both behaviors prevents the robot from being blocked on a specific state. For example, the robot could decide to exploit a state without having discovered a significant action. In that case the model would be blocked because there would be no action to exploit in that state. Despite that risk, the robot produces a sequence of actions without finishing in a "dead" state.

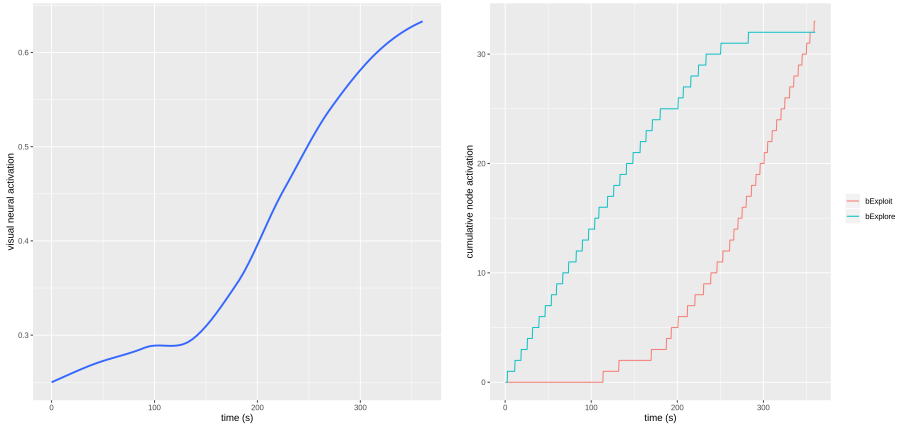

Fig. 5: Average results for 10 experiments. Left : the average visual neural activation over time of 10 experiments is represented by a linear regression. The curve shows an increase of visual activation when the model begins to exploit the sensorimotor contingencies. Right : the sum of the activation nodes *bExplore* and *bExploit* (respectively when Exploring and Exploiting) over time for the 10 experiments. A decision for exploration (at the beginning) and a trend for exploitation (starting around 190 seconds) can be observed.

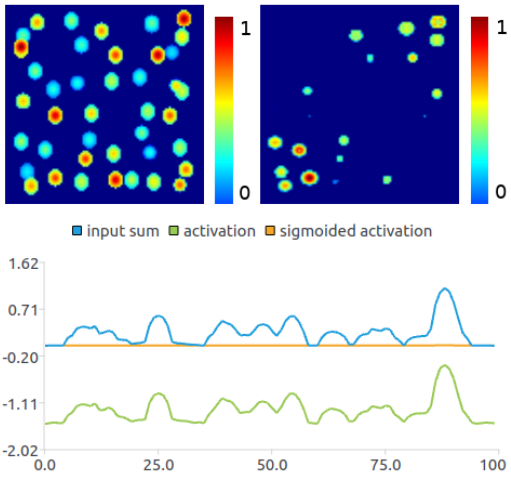

Fig. 6: Neural activations after one experiment. Top-left are the neural activation gathering the actions performed and sent as inhibition (memory trace actions). Top-right are the actions with a high visual outcomes (memory trace). Bottom part are the activations within the Condition of Exploitation Field (without the current input from the state of the arm).

Figure 6 depicts the neural activations of the actions taken as well as the Condition of Exploitation field. The architecture does not need to explore the complete actions space to reach a stable sequence of actions because of the switch mechanism. The last section will conclude this paper, state the current limitation and discuss future work.

5 Conclusion and Discussion

This paper introduces a model for learning sensorimotor contingencies with a humanoid robotic arm based on neural dynamics. The architecture takes insights from human development by performing motor babbling in a closed loop. The learning occurs when generating motor commands, and associating them with the changes induced in the environment. An inhibition of return prevents the model from generating the same action twice. At any moment, the system can decide whether to explore the environment or to exploit the sensorimotor associations. Indeed, the main contribution rests on a neural switch mechanism that dynamically balances between both behaviors. Results demonstrate an increase of visual neural activation when the robot begins to exploit its knowledge. In addition, the time course of both exploratory and exploitative behavior shows a tendency toward using the sensorimotor knowledge after a certain time. Finally, the switch mechanism allows the robot to avoid exploring the complete sensorimotor space.

However, only a single degree of freedom is utilized to demonstrate the advantages of the switch mechanism. The setup of the experiment is voluntary simple to keep a track on the rewards in time (visual neural activation). Indeed, due to the complexity of the model, this setting allows also to study and validate with clarity the behavior of the neural fields and memory traces composing the switch mechanism. To address this issue, the future work will use the whole GummiArm in an inverse kinematic mode with a three-dimensional neural field representing the robot's end-effector.

Finally, we intend to develop the model toward goal directed actions in a richer environment. In order to model higher-order goals, we will adapt the method of researcher regarding the gain modulation of multimodal cues [27] to dynamic neural fields . A novelty detector based on the three layer model [28] will be used as a dynamic neural mechanism delivering rewards by peaks of activation in case of "novel" events, avoiding to specify an external reward by design.

Then, the robot will generate and learn to reach goals with the help of this exploration/exploitation behavior. The switch mechanism introduced here will exploit the goals with the highest rewards to discover other potential goals. The complete architecture would represent perceptions as hierarchically organized, and sequences of goals will lead to more complex perception over time. With the possibility to represent perceptions as probabilities with peaks of activation, a particular attention will be given to possibly apply inference processes [29, 30] on these stimuli.

6 Appendix

Wiki, set of parameters, source code and architecture files to reproduce the experiment are available at https://github.com/rouzinho/neural-switch-dft/wiki.

References

1. Tekülve, J., Fois, A., Sandamirskaya, Y., Schöner, G.: Autonomous sequence generation for a neural dynamic robot: Scene perception, serial order, and object-oriented movement. Frontiers in Neurorobotics 13, 95 (2019)
2. Cangelosi, A., Schlesinger, M.: Developmental Robotics: From Babies to Robots. The MIT Press (2014)
3. O'Regan, J.K., Noë, A.: A sensorimotor account of vision and visual consciousness. Behavioral and Brain Sciences 24(5), 939–973 (2001)
4. Piaget, J., Cook, M.: The origins of intelligence in children, vol. 8. International Universities Press New York (1952)
5. Demiris, Y., Dearden, A.: From motor babbling to hierarchical learning by imitation: a robot developmental pathway. In: International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems. vol. 123, pp. 31–37 (2005)
6. Mahoor, Z., MacLennan, B.J., McBride, A.C.: Neurally plausible motor babbling in robot reaching. In: Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob). pp. 9–14 (2016)
7. Lanillos, P., Dean-Leon, E., Cheng, G.: Yielding self-perception in robots through sensorimotor contingencies. IEEE Transactions on Cognitive and Developmental Systems 9(2), 100–112 (2016)
8. Houbre, Q., Angleraud, A., Pieters, R.: Exploration and exploitation of sensorimotor contingencies for a cognitive embodied agent. In: ICAART (2). pp. 546–554 (2020)
9. Berger-Tal, O., Nathan, J., Meron, E., Saltz, D.: The exploration-exploitation dilemma: a multidisciplinary framework. PloS one 9(4) (2014)
10. Sutton, R.S., Barto, A.G.: Introduction to reinforcement learning, vol. 135. MIT press Cambridge (1998)
11. Chernova, S., Veloso, M.: Interactive policy learning through confidence-based autonomy. Journal of Artificial Intelligence Research 34, 1–25 (2009)
12. Maye, A., Engel, A.K.: A discrete computational model of sensorimotor contingencies for object perception and control of behavior. In: 2011 IEEE International Conference on Robotics and Automation. pp. 3810–3815. IEEE (2011)
13. Cohen, J.D., McClure, S.M., Yu, A.J.: Should I stay or should I go? how the human brain manages the trade-off between exploitation and exploration. Philosophical Transactions of the Royal Society B: Biological Sciences 362(1481), 933–942 (2007)
14. Humphries, M., Khamassi, M., Gurney, K.: Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. Frontiers in Neuroscience 6, 9 (2012)
15. Schöner, G., Spencer, J., Group, D.F.T.R.: Dynamic Thinking: A Primer on Dynamic Field Theory. Oxford University Press (2016)
16. Cannon, W.B.: Organization for physiological homeostasis. Physiological reviews 9(3), 399–431 (1929)

17. Perone, S., Spencer, J.P.: Autonomy in action: Linking the act of looking to memory formation in infancy via dynamic neural fields. Cognitive Science 37(1), 1–60 (2013)
18. Sandamirskaya, Y., Schöner, G.: Serial order in an acting system: A multidimensional dynamic neural fields implementation. In: 2010 IEEE 9th International Conference on Development and Learning. pp. 251–256 (2010)
19. Kazerounian, S., Luciw, M., Richter, M., Sandamirskaya, Y.: Autonomous reinforcement of behavioral sequences in neural dynamics. In: The 2013 International Joint Conference on Neural Networks (IJCNN). pp. 1–8 (2013)
20. Stoelen, M.F., Bonsignorio, F., Cangelosi, A.: Co-exploring actuator antagonism and bio-inspired control in a printable robot arm. In: From Animals to Animats 14. pp. 244–255. Springer International Publishing, Cham (2016)
21. Spencer, J.P., Perone, S., Johnson, J.S.: The dynamic field theory and embodied cognitive dynamics. Toward a unified theory of development: Connectionism and dynamic systems theory re-considered pp. 86–118 (2009)
22. Amari, S.I.: Dynamics of pattern formation in lateral-inhibition type neural fields. Biological Cybernetics 27(2), 77–87 (1977)
23. Posner, M.I., Rafal, R.D., Choate, L.S., Vaughan, J.: Inhibition of return: Neural basis and function. Cognitive Neuropsychology 2(3), 211–228 (1985)
24. Tipper, S.P., Driver, J., Weaver, B.: Short report: Object-centred inhibition of return of visual attention. The Quarterly Journal of Experimental Psychology Section A 43(2), 289–298 (1991)
25. Bar-Gad, I., Morris, G., Bergman, H.: Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. Progress in neurobiology 71(6), 439–473 (2003)
26. Netzev, M., Houbre, Q., Airaksinen, E., Angleraud, A., Pieters, R.: Many faced robot-design and manufacturing of a parametric, modular and open source robot head. In: 2019 16th International Conference on Ubiquitous Robots (UR). pp. 102–105. IEEE (2019)
27. Mahé, S., Braud, R., Gaussier, P., Quoy, M., Pitti, A.: Exploiting the gain-modulation mechanism in parieto-motor neurons: Application to visuomotor transformations and embodied simulation. Neural Networks 62, 102–111 (2015)
28. Johnson, J.S., Spencer, J.P., Luck, S.J., Schöner, G.: A dynamic neural field model of visual working memory and change detection. Psychological science 20(5), 568–577 (2009)
29. Cuijpers, R.H., Erlhagen, W.: Implementing bayes' rule with neural fields. In: International Conference on Artificial Neural Networks. pp. 228–237. Springer (2008)
30. Gepperth, A., Lefort, M.: Latency-based probabilistic information processing in recurrent neural hierarchies. In: International Conference on Artificial Neural Networks. pp. 715–722. Springer (2014)