



HAL
open science

An architecture for online chunk learning and planning in complex navigation and manipulation tasks

Souheil Hanoune, Mathias Quoy, Philippe Gaussier

► **To cite this version:**

Souheil Hanoune, Mathias Quoy, Philippe Gaussier. An architecture for online chunk learning and planning in complex navigation and manipulation tasks. ICDL Epirob IEEE Conference on Development and Learning and Epigenetic Robotics, Nov 2012, San Diego, California, USA, United States. 5 p. <hal-00742030>

HAL Id: hal-00742030

<https://hal.science/hal-00742030v1>

Submitted on 15 Oct 2012

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.



HAL Authorization

An architecture for online chunk learning and planning in complex navigation and manipulation tasks

Souheil Hanoune, Mathias Quoy and Philippe Gaussier

ETIS, CNRS ENSEA University of Cergy-Pontoise, F-95000 Cergy-Pontoise, France

{souheil.hanoune, quoy, gaussier}@ensea.fr

Abstract—When a robot is brought into a new environment, it has a very limited knowledge of what surrounds it and what it can do. Either to navigate in the world, or to interact with humans, the robot must be able to learn complex states, using input information from sensors. For navigation task, visual information are commonly used for localisation. Other signals are usually employed: ultrasounds, lasers and path integration are as many data that can be taken into account. For human-robot interactions the proprioceptive information, like the values of the articulations and the grip state, are additional degrees in the system that can be enrolled in the analysis. All those signals have different dynamics, and the system must manage all those differences. To do so, a solution is to introduce different learning levels. An architecture able to create complex multimodal contexts, to recognise them and to use them afterwards in higher level strategies would give the capacity to resolve such situations.

In this paper, a model has been introduced for complex categorisation problems, patterning and chunk learning. It is used in a complex architecture to create multimodal states for a cognitive map and to resolve ambiguities. Tests were performed in both a simulated navigation task and in a complex arm manipulation and localisation experiment.

I. INTRODUCTION

Autonomous robots have to learn more and more complex categories, either autonomously or with the help of a caregiver. To do so, the robot must be able to recognise new situations, categorise them in order to use them. Commonly, simple associations between input signals are used, which can include ambiguities in the categorisation. Visuo-motor associative learning and a motor controller can provide the basis to let low level immediate imitation emerge as a result of simple mechanisms [1]. The system learns the body configuration using visuo-motor associations, and due to a perception ambiguity, caused by limited visual capabilities, the robot can confuse the hand of the teacher with its own.

In a navigation task [2], the authors used a cognitive map strategy for navigation [3]. The goals were associated with a transition, according to colored place cell. The color and the place cell are nearly correlated, but this correlation is not interpreted in the system.

Transition categories were obtained by combining the two last events received. The time difference estimated in this case is only based on the last event, and every new reception erases the old one [4]. In this situation, a temporal relation between two event, strongly correlated, but separated in time by another

one is not possible. Until now, the signals given to the system were always relevant: visual information, color detection, sound in navigation, proprioceptive information and the grip of the hand in the case of an arm manipulation task. However, if we want to trend to a generic architecture, the large number of signals does not guarantee their pertinence for the event. In such case, using a model able to detect correlations, between signals globally close in time, is necessary.

A solution is to introduce different levels of learning. An architecture able to create complex multimodal contexts, recognise them and use them afterwards in higher level conditioning would give the capacity to resolve such situations.

The aim of this paper is to present a bio-inspired model for generic patterning and chunk learning. The article exposes some preliminary works using the model.

We tested our architecture in two learning tasks. First, a navigation simulation task, exposing the ability of the system to capture correlations between signals close in time even if they are separated by another one. The second one is a robotic experiment, including navigation and arm manipulation. In this experiment, the architecture is used for contexts categorisation. In a human-robot interaction task, the multimodal states are learned based on signals used in the navigation and the proprioceptive information of the arm.

In the next section we present the architecture, the different parts and the learning equations. We also raise the different possible uses of such a model. In section III, we propose two possible application of this model, a navigation experiment and an arm manipulation one are investigated. Finally, the section IV present the perspectives for the next steps to improve the model, and also the future uses for such an architecture.

II. A BIO-INSPIRED MODEL FOR CHUNKING AND CONTEXT LEARNING

Chunks learning is used on many levels, in low level visuo-motor learning [5], and also in high level interaction like for talking or singing [6]. To encode the information, two different methods can be used: either the input signals categorisation is saved in some neuron [7], or the categorisation is based on population coding or a dynamic field activity [8], [9].

In [10], Schmajuk and DiCarlo presented a model for context learning. The paper treated the functional concept of the model. The principal aspect is that the authors used, in addition

to the direct pathway commonly used for multimodal fusion, a hidden layer in which the correlations between the inputs are learned. The second pathway is used to detect chunks and associate time correlated signals. The final categorisation will be conditioned by either the direct pathway with rough input signals, or the chunks learned previously.

The model we present in this paper is based on Schmajuk's and Dicarlo's architecture introduced in [10], where they presented a functional bio-inspired model, exposing the error calculation, the different brain areas involved pointing out the various signals included in the learning mechanism. For our model, we derived this architecture to obtain a more generic one that can be implied in different robotic application.

The architecture we propose is presented in figure 1.

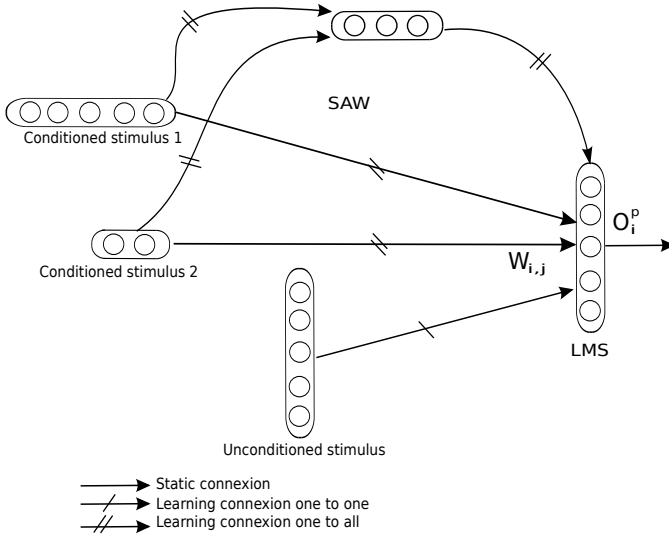


Figure 1. There are two pathways for the inputs in the architecture. The direct pathway is just a direct injection of the inputs into the neurons. The second one, passes through a hidden layer, where the correlations between the inputs are learned.

The direct pathway can only perform rough learnings of the inputs values. This pathway does not take the possible time associations between the inputs. The indirect pathway is used to encode the trace conditioning.

The conditioning of the weights in the hidden layer of the model is implemented based on a simple reactive architecture. We use a network inspired from the Adaptive Resonance Theory [11]. The states are encoded on the weights of the input links of recruited neurons. The computation of the activity (eq. 1) and learning (eq. 2) for those neurons is given by the following equations:

$$A_i = 1 - \frac{1}{N} \sum_j |w_{ij} - e_j| \quad (1)$$

$$\begin{aligned} \Delta w_{ij} &= \epsilon(e_j - w_{ij}) & \text{if } A_i > \text{vigilance} \\ \Delta w_{ij} &= (e_j - w_{ij}) & \text{else} \end{aligned} \quad (2)$$

where N is the dimension of the input layer, A_i is the activity of the i th neuron of the output layer, ϵ the learning rate and w_{ij}

is the weight of the connection between an output neuron i and the j^{th} neuron of the input layer whose activity is e_j . H is the Heaviside function. As in the ART initial model, a vigilance factor is used to tune the granularity of the recruitment, that is to say the decision of recruiting a new chunk, vs averaging the winning (higher A_i) value.

In addition to the inputs, the chunks are used to predict the desired signal in a Least Mean Square (LMS) learning group. This group uses a modified Widrow-Hoff learning rule [12] described in eq. 3. The output of the neuron is computed as described in eq. 4

$$dW_{i,j} = \mu(t) \cdot (D_j - O_j^p) \cdot C_i \quad (3)$$

where D_j is the activity of the unconditional stimulus, $\mu(t)$, $W_{i,j}$ the weights for each input, C_i the context signal and O_j^p the predicted output. The equation for the neural activity computing is:

$$X_i(t) = f\left(\sum_j W_{ij} \cdot X_j\right) \quad (4)$$

The architecture will capture correlations between close inputs. If there is no correlation, the hidden layers weights tend to 0; otherwise, those weight heighten up and the direct ones tend to 0.

To illustrate the learning behavior we present a simple example. We simulate three inputs, and according to the inputs configuration we will show how the system will react. In the figure 3 we can see the inputs signals in the 3 first graphics.

The chunks are learned according to the equations presented in 12. The recruiting signal is activated each time the activity in the output decreases under a certain threshold. In Figure 3, we can see the output activity of the chunks in the 4th graphic.

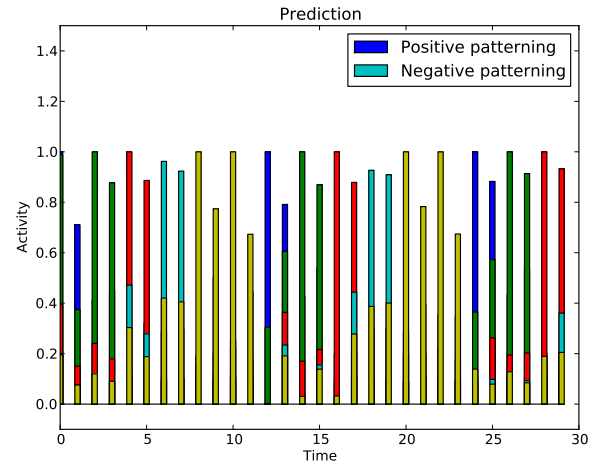


Figure 2. The prediction is exposed in the last graphic (LMS), after few iteration the prediction is stabilised and the response is coherent.

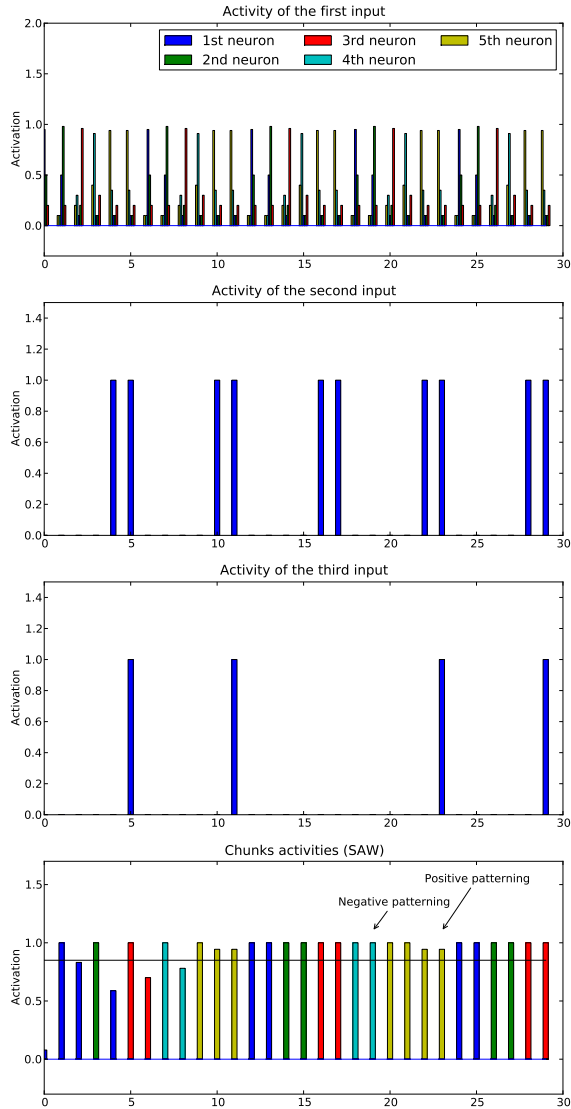


Figure 3. The three first graphics present the inputs. The first one is composed by 5 neurons corresponding to each color. The chunks (SAW) learning is shown in the 4th graphic, we can see an example of a positive patterning (activated when the 2nd and 3rd inputs are activated), and a negative patterning (activated when the inputs are inactivated).

In the next section we will test the model in two robotic experiments in order to explore the usability of such an architecture.

A. Complex categorisation in autonomous navigation task

This a navigation simulation task. The simulated robot is able to learn its environment using a strategy based on hippocampal place cells presented in [13]. This approach

discriminates the space into different "places" recognised independently, from one another, during the navigation. The architecture uses the place cell in order to form transition cells, based on the model of Banquet et al. [14], successfully used in the past by Cuperlier et al. [3] and by Hirel et al. [2] for robotic navigation tasks. Transitions between visual place cells were learned and used in navigation with a planning strategy based on a cognitive map of transitions. However, no timing was learned for the spatial transitions between places. The cognitive map provided information about the available paths between spatial states, but not about the time needed to shift from one state to the next. In order to use the cognitive map for the navigation we need to activate goals to reach, so the activation is propagated through the cognitive map to find out the fastest path.

In the simulation, we use a multimodal goal. It is a combination of a red colored place cell and a sound, activated 300ms after the robot enters in the red place cell. The time delay is modeled by Granular cells (DG) activations. Those granular cells are commonly used in time prediction models, for instance [15] and [16] [17].

To test the architecture, we use a square environment, in which two places are red, one in the south-east and the other on the north-west of the arena, but only the first one can activate the sound.

The architecture used is described in figure 4.

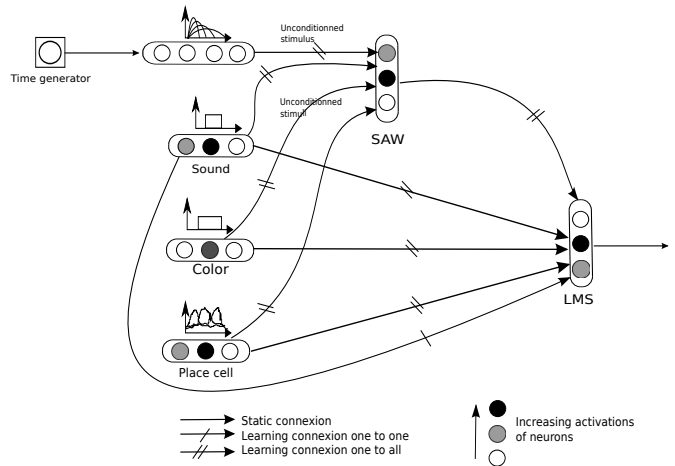


Figure 4. Architecture used for the multimodal goal learning. When the robot enters in a red colored place cell, waits for 300ms, a sound is perceived giving the reward. The timing is generated by the temporal activities in the granular cells. The red-time-sound association is then learned to obtain the reward.

Our architecture allows the robot to learn the correlation between the red place cell and the time delayed sound. The condition based learning between the sensory inputs and the reward is known as positive patterning. It is largely studied by Schmajuk and DiCarlo in [10]. Such a conditioning occurs when the two stimuli, in addition to the condition are presented, if only one of the stimuli is present, the condition is not satisfied.

In figure 5, we can see the arena with the two goals.

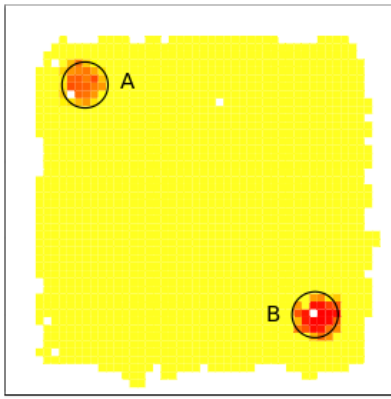


Figure 5. Both goals A and B are red colored place cells. The red color is the activation of the neurons predicting the occurrence of the sound.

As exhibit, both A and B fires the prediction neurons. This result is obtained in the early stages of the learning. We can see that, even if the neurons fires in A, the red color is less intense. The association detection resulting in a reward, we see in figure 6 the result after some time.

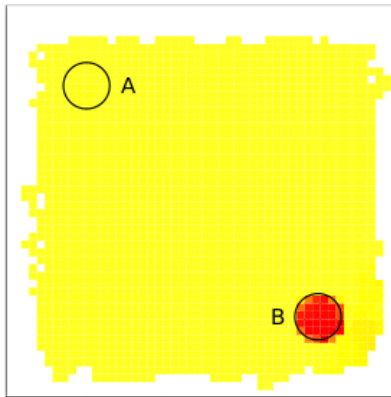


Figure 6. The A goal have disappeared, the model learned that the association red-sound is present only in place B.

In figure 6, we see that even if, at the start of the learning, the robot was associating the place A with the reward prediction, during time, this association is lost and only place B is taken into consideration for the reward prediction. This is the result of the positive patterning. The hidden layer in the model learns the association between the time delay from DG and the red place cell, when at the same time, it loses the direct association learned in place A, because of the reward was not obtained.

B. Context capture in arm manipulation behaviour

This experiment is derived from De Rengervé's previous work described in [18]. It is a combination of the work done by Lagarde et al in [1] and by Calinon et al. in [19], for learning a navigation task in association with a guided arm manipulation for future reproduction.

The experiment goes like this: the robot navigates to a certain point in the environment, must wait for an object to be placed in its gripper, then takes the object to another point to be released. There are different actions possible depending of the objects size. Figure 7 shows a global view of the experiment.

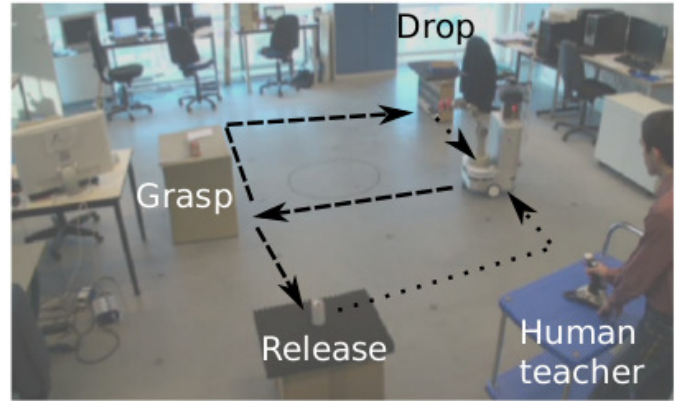


Figure 7. Overview of the task. At the grasping point, the robot chooses what action to do (drop or release) depending on the objects size. The navigation is based on vision, enabling a robust self-localisation.

The actions are taught to the robot by the human teacher. The aim of the experiment is to use the model presented in section II to learn arm contexts, according to the different inputs (grip, localisation, arm position,...) so that the robot knows when to drop or the release the object. The adaptation of the architecture is presented in figure 8.

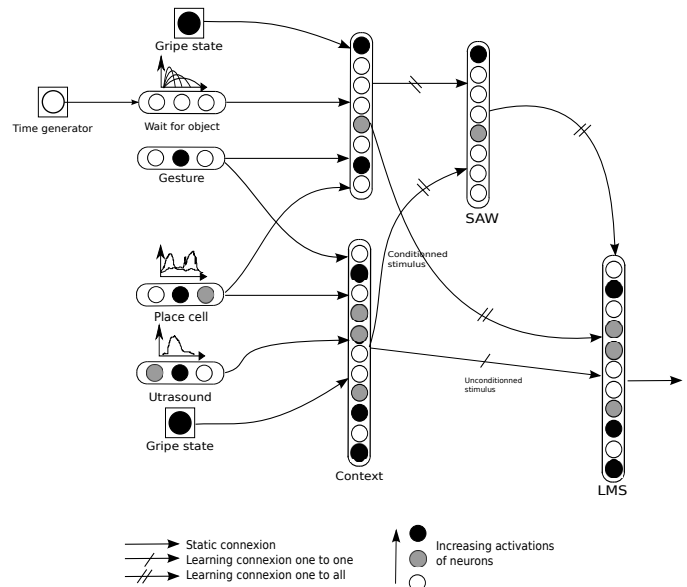


Figure 8. Architecture used for the arm context learning. Each state is an association based on the grip, the place cell, the amount of time the robot waited and the gesture taught by the human. According to the encoded association, the robot will be able to identify this context according to the ultrasound sensors, the place cell, the action and the grip state.

In this case, the model learns the contexts based on conditional associations. When the robot is in a certain place cell (to

take an object for example), the teacher provides the guideline for this object. When the robot arrives to the destination (to release it for instance), a conditional learning is triggered. According to a concatenation of the action done, the actual place cell, the grip state (in fact, the size of the object) and the time passed waiting, the system makes the association with the concatenation of the place cell, the ultrasound sensors, the grip state and the action. First, the possible correlations between those inputs are learned, then the global context is encoded giving the ability to recognise it in the future.

III. DISCUSSION AND CONCLUSION

In this paper, we presented a neural architecture for multimodal states categorisation in a multi-task problem. The system has been tested on a simulated of a navigation task, and a robotic arm manipulation associated with a navigation task, in order to show the different contexts where it can be used. The architecture allows to create contexts in order to use them in a long term learning like in the cognitive map [3] for example. With this architecture, stable states can be built, from correlations between multimodal signals. If we add new modalities to categorise, we will obtain more possible situations, but also an explosion in the system's complexity. The presented model deals with such a problem by including the correlation between in the categorisation, in addition to the inputs themselves.

The error in the model is automatically calculated depending on the reward. In [20], Giovannangeli exposed some bases for auto-evaluation in autonomous robotics. Coupling such a system with the given model could give an interesting advance in non-supervised learning tasks.

This work exposed some preliminary work in the tests of the given architecture. The analysis needs to be more accurate in order to have the limitations for such categorisation and to explore new possible improvements.

A next step is to introduce a loop in the learning process, in order to reuse the past learned states as possible inputs for next categorisation. A system with this ability can be built by including a memory describing the past learned contexts. This modification obliges us to study the robustness and the stability of such a system. Another possible direction is the ability to forget unused categories. If the system can categorise using past learned states, the obsolete one have to be less implied in future categories. Those are possible improvements for our architecture. As we have seen in the arm-navigation task, even if complex states are learned by the system, the human is still a key in the learning process. Allowing the human to directly correct the robot gives a good recognition contexts for the demanded experience and the teaching is less tedious for the human. The implication of the human in the leaning process or in any other interaction, implies a more complex analysis, because of all the synchronies and the timings introduced in any interactions, and in this case, the model can be a solution.

In [21], the author notices the lack of stability in long term learnings. The system is most likely to lose its stability and the generalisation is less efficient. A system that consolidates

its memory would be more precise: during sleep, short term memory is re-encoded in the neocortex to become long term memory. A system including such a regulation will use past experience to develop new states, causing an improvement in generalisation while being more robust to noise.

ACKNOWLEDGMENTS

This work was supported by the Neurobot project ANR-BLAN-SIMI2-LS-100617-13-01.

REFERENCES

- [1] P. Andry, P. Gaussier, M. Lagarde, and S. Boucenna, "Proprioception and imitation: on the road to agent individuation," in *IEEE/RSJ International Conference on Robots and Systems (IROS 2008)*, Session, 2008, to appear.
- [2] J. Hirel, P. Gaussier, and M. Quoy, "Biologically inspired neural networks for spatio-temporal planning in robotic navigation tasks," in *Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on*, dec. 2011, pp. 1627–1632. [Online]. Available: <http://dx.doi.org/10.1109/ROBIO.2011.6181522>
- [3] N. Cuperlier, M. Quoy, and P. Gaussier, "Neurobiologically inspired mobile robot navigation and planning," *Frontiers in neurorobotics*, vol. 1, no. 3, p. 15, 2007.
- [4] J. Hirel, "Codage hippocampique par transitions spatio-temporelles pour l'apprentissage autonome de comportements dans des tâches de navigation sensori-motrice et de planification en robotique," Ph.D. dissertation, Université de Cergy-Pontoise, 2011.
- [5] Y. Kuniyoshi, Y. Yorozu, M. Inaba, and H. Inoue, "From visuo-motor self learning to early imitation -a neural architecture for humanoid learning," in *ICRA'03*, 2003, pp. 3132–3139.
- [6] Y. Yamashita, M. Takahasi, T. Okumura, M. Ikebuchi, H. Yamada, M. Suzuki, K. Okanoya, and J. Tani, "Developmental learning of complex syntactical song in the bengalese finch: A neural network model," *Neural Netw.*, vol. 21, no. 9, pp. 1224–1231, Nov. 2008.
- [7] Y. Sandamirskaya, "Learning in dynamic neural fields model for sequence generation," pp. 15–17.
- [8] G. Zhou and J. Su, "Error-driven hmm-based chunk tagger with context-dependent lexicon," in *Proceedings of the 2000 Joint SIGDAT conference on Empirical methods in natural language processing and very large corpora: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics - Volume 13*, ser. EMNLP '00. Stroudsburg, PA, USA: Association for Computational Linguistics, 2000, pp. 71–79.
- [9] B. Durán, Y. Sandamirskaya, and G. Schöner, "A dynamic field architecture for the generation of hierarchically organized sequences," in *ICANN (1)*, 2012, pp. 25–32.
- [10] N. A. Schmajuk and J. J. DiCarlo, "Stimulus configuration, classical conditioning, and hippocampal function," *Psychological Review*, vol. 99, no. 2, p. 268–305, apr 1992, PMID: 1594726.
- [11] G. Carpenter and y. . . S. Grossberg, title = Adaptive resonance theory.
- [12] B. Widrow and M. E. Hoff, "Adaptive switching circuits," in *1960 IRE WESCON Convention Record, Part 4*. New York: IRE, 1960, pp. 96–104. [Online]. Available: <http://isl-www.stanford.edu/~widrow/papers/c1960adaptiveswitching.pdf>
- [13] J. O'Keefe and J. Dostrovsky, "The hippocampus as a spatial map. preliminary evidence from unit activity in the freely-moving rat." *Brain Research*, vol. 34, no. 1, pp. 171–175, 1971. [Online]. Available: <http://discovery.ucl.ac.uk/95537/>
- [14] J. P. Banquet, P. Gaussier, J. C. Dreher, C. Joulain, A. Revel, and W. Gunther, "Space-time, order and hierarchy in fronto-hippocampal system: A neural basis of personality," in *Cognitive Science Perspectives on Personality and Emotion*, pp. 123–189, 1997.
- [15] S. Grossberg and N. A. Schmajuk, "Neural dynamics of adaptive timing and temporal discrimination during associative learning," *Neural Networks*, vol. 2, no. 2, pp. 79–102, 1989.
- [16] S. Moga, "Apprendre par imitation : une nouvelle voie d'apprentissage pour les robots autonomes," Ph.D. dissertation, Université de Cergy-Pontoise, 2000.

- [17] P. Gaussier, S. Moga, M. Quoy, and J. P. Banquet, "From perception-action loops to imitation processes: A bottom-up approach of learning by imitation," *Applied Artificial Intelligence*, vol. 12, no. 7-8, pp. 701–727, 1998.
- [18] F. D'halluin, A. de Rengervé, M. Lagarde, P. Gaussier, A. Billard, and P. Andry, "A state-action neural network supervising navigation and manipulation behaviors for complex task reproduction," in *Proceedings of the tenth international conference on Epigenetic Robotics*, Orenas Slott, Sweden, December 2010, pp. 165–166.
- [19] S. Calinon, F. Dhalluin, D. Caldwell, and A. Billard, "Handling of multiple constraints and motion alternatives in a robot programming by demonstration framework," *Proceedings of 2009 IEEE International Conference on Humanoid Robots*, 2009, p. 582 - 588, 2009.
- [20] C. Giovannangeli, "Navigation autonome bio-inspirée en environnement intérieur et extérieur: Apprentissages sensori-moteurs et planification dans un cadre interactif." Ph.D. dissertation, 2007.
- [21] T. Ogata, S. Sugano, and J. Tani, "Open-end human robot interaction from the dynamical systems perspective: mutual adaptation and incremental learning," in *Proceedings of the 17th international conference on Innovations in applied artificial intelligence*, ser. IEA/AIE'2004. Springer Springer Verlag Inc, 2004, pp. 435–444. [Online]. Available: <http://dx.doi.org/10.1007/b97304>