



**HAL**  
open science

# Exploring Sensitization in the Context of Extending the Behavior of an Artificial Agent

Tristan Gillard, Jérémy Fix, Alain Dutech

► **To cite this version:**

Tristan Gillard, Jérémy Fix, Alain Dutech. Exploring Sensitization in the Context of Extending the Behavior of an Artificial Agent. International Conference on Simulation of Adaptive Behavior (SAB22), Sep 2022, Cergy-Pontoise, France. pp.15-27, 10.1007/978-3-031-16770-6\_2. hal-04031636

**HAL Id: hal-04031636**

**<https://hal.science/hal-04031636>**

Submitted on 16 Mar 2023

**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.

# Exploring sensitization in the context of extending the behavior of an artificial agent

Tristan Gillard<sup>1,2</sup>[0000-0002-8469-5828], Jérémy Fix<sup>1,3</sup>[0000-0003-1889-8886], and  
Alain Dutech<sup>1,4</sup>[0000-0001-7549-7988]

<sup>1</sup> Loria, UMR 7503, Nancy, France

<sup>2</sup> Université de Lorraine, Nancy, France

<sup>3</sup> CentraleSupélec, Metz, France

<sup>4</sup> INRIA Nancy Grand-Est, Nancy, France

{tristan.gillard, jeremy.fix, alain.dutech}@loria.fr

**Abstract.** Sensitization, a non-associative learning widely observed across phylogeny, is fundamental for adaptation and, thus, survival of living organisms. This paper investigates one form of sensitization, namely pseudo-conditioning, in order to present a new computational model inspired by its characteristics. We develop this model as part of the Iterant Deformable Sensorimotor Medium (IDSM), a recently developed abstract model of sensorimotor behavior formation. The characteristics of the presented model are studied and analyzed in the light of our long term objective of investigating new unsupervised learning mechanisms for artificial autonomous agents.

**Keywords:** Non-associative learning · sensitization · pseudo-conditioning · artificial agent · adaptive behavior acquisition · embodiment.

## 1 Introduction

What drives the behavior of an organism? That fundamental question has been addressed in the literature for a variety of organisms and through different angles. For some species, part of the answer lies in the associative learning mechanisms but there also exists more widespread non-associative learning mechanisms ([1,2,7]). Behavior, in its simplest expression, is a spatio-temporal sequence of motor responses triggered by sensory signals. Habituation and sensitization, the first translating behaviorally in a response decrease while the second in a response increase, are two non-associative learning mechanisms so widely observed in the phylogeny that it suggests their fundamental role in behavioral learning ([6,16,18,4]).

Sensitization, described in more details in section 2, translates as an increase in response with two flavours. It is defined as either the enhancement of the response to a repeated stimulus (*site-specific sensitization*) or as a generalization of this response to another stimulus (*pseudo-conditioning*). Hence, sensitization, rather than habituation, appears as an interesting mechanism to investigate how an agent can broaden its behavioral responses. Moreover, while site-specific

sensitization is anchored on a single stimulus, pseudo-conditioning transfers the response to other stimuli. We can then see pseudo-conditioning as a first step toward generalization. For example, in the context of an agent exploring an environment, pseudo-conditioning would permit an agent to test actions on previously neutral stimuli, possibly discovering new opportunities.

In the context of autonomous artificial agents, there are several papers modeling sensitization [10,20,21,17,3,13], some of them reviewed in [15]. However, most of them use the word sensitization to designate site-specific sensitization exclusively. As far as we know, only the sensitization model of [3] deals with pseudo-conditioning.

The works of [10,20], which have been released at the same period of the work of Rescorla on classical conditioning [19], are the first computer models of both habituation and site-specific sensitization. The behavioral responses, consistent with biological measurements, are a mix of habituation and site-specific sensitization where habituation is more pronounced for low intensity stimuli, and sensitization for high intensity stimuli.

In a recent review of cognitive architectures ([15]), two models of sensitization are mentioned. In the first, Attentive and Self-Modifying cognitive architecture (ASMO, see [17]), habituation and site-specific sensitization essentially use a "significance" function of the stimulus rather than its intensity, which allows habituation to strong stimuli and site-specific sensitization to weak stimuli. This makes it easier for their robot to stay focused on its target, favoring stimuli close to the target (deemed significant) over others that became more easily ignored. In the second, Self-Aware and Self-Effecting agent (SASE, see [21]), the authors formalize six *mental* architectures from the simplest to the most complex, all based on Markov Decision Processes. Starting with architectures type-4, their conceptual models are said to allow for habituation and site-specific sensitization, but this type-4 architecture is also compatible with classical conditioning, instrumental conditioning, planing and reasoning. Besides, only the simplest architectures (up to type-2) are experimented with, habituation and sensitization both requiring at least type-4 architecture.

In [13], both habituation and site-specific sensitization are modeled for a neuromorphic processor. A weak, repeated stimulus leads to habituation of the neuron activity, whereas a strong stimulus leads to the site-specific sensitization of the neuron activity. They show that such architecture lead to more robust hardware with better aging properties.

All these works use site-specific sensitization along with habituation, but do not mention pseudo-conditioning. The neuro-physiological models of Hawkins [11] can account for a variety of learning mechanisms including pseudo-conditioning. They bring in the suggestion that non-associative learning forms the breeding ground for the emergence of other forms of learning, especially associative conditioning. In the "Autonomous Robot Based on Inspirations from Biology" (ARBIB, see [3]), habituation and pseudo-conditioning are modeled with a spiking neural network, along with classical conditioning. The neurons encode reflexes, and are implemented on a Khepera robot to observe the resulting behaviors. Conscious

of the daunting complexity of models too faithful to the physiological modeling, Damper et al. relied on spiking neural networks to abstract the physiological mechanisms, but this is still a rather complex model. It is also unclear to which extent their model can exploit non associative learning without relying on the mechanisms needed for classical conditioning.

In this paper, we propose a new computational mechanism inspired by pseudo-conditioning as described by biologists. As such, the proposed mechanism is centered on the response generalization to other stimuli. We plug our proposition to an existing conceptual model of habits formation called Iterant Deformable Sensorimotor Medium (IDSM, see [5]), more conceptual than spiking neural networks but less than models in the tradition of Rescorla-Wagner. We are interested in artificial agents that can acquire new behaviors and the IDSM framework specifically addresses the problem of using a simple kind of basic behaviors: habits. IDSM performs well for *memorizing* new behaviors and reinforcing them so as to *replaying* them later. But it has some limitations when it comes to *generating* new behaviors. Our motivation is to add learning or adaptation mechanisms to the IDSM framework so as to open the possibility for the artificial agent to produce new behaviors.

This paper is structured as follows. We start by recalling the biological definition of sensitization and its behavioral correlates in section 2. The Iterant Deformable Sensorimotor Medium (IDSM) framework is then described in section 3. The proposed model is detailed in section 4 and various experiments are presented and analyzed in section 5. We conclude the paper with a discussion.

## 2 Sensitization

Sensitization is a *non-associative* learning observed within numerous and various species, including insects or single cell organisms ([18,4]). However, it is important to note that in the literature, unlike habituation (*i.e.* decrease of a response to a repetitive stimulus), sensitization refers to different phenomena ([14,12,8,2,4]). Sensitization is defined as the enhancement of a response behavior after the presentation of a stimulus, and can be of two forms: *site-specific* sensitization and *pseudo-conditioning*. By definition, site-specific sensitization is the enhancement of the behavioral response following repeated presentation of the same stimulus.

Pseudo-conditioning is a form of learning allowing a behavioral response to a given *unconditioned* stimulus *US* to generalize, temporarily, as a response to another *conditioned* stimulus *CS*. There are major differences with the associative learning *classical conditioning*. In pseudo-conditioning the presentation of the *US* alone is enough to prepare for the trigger of the unconditioned response to other stimuli. A stimulus triggering this unconditioned response temporarily becomes a *CS*. The triggered response in pseudo-conditioning is weaker than with classical conditioning.

The figure 1 illustrates the behavioral response in case of pseudo-conditioning with either a continuous or repeated and shorter stimulation. As shown on

the figure, when pseudo-conditioning takes place, a previously neutral stimulus triggers the conditioned response.

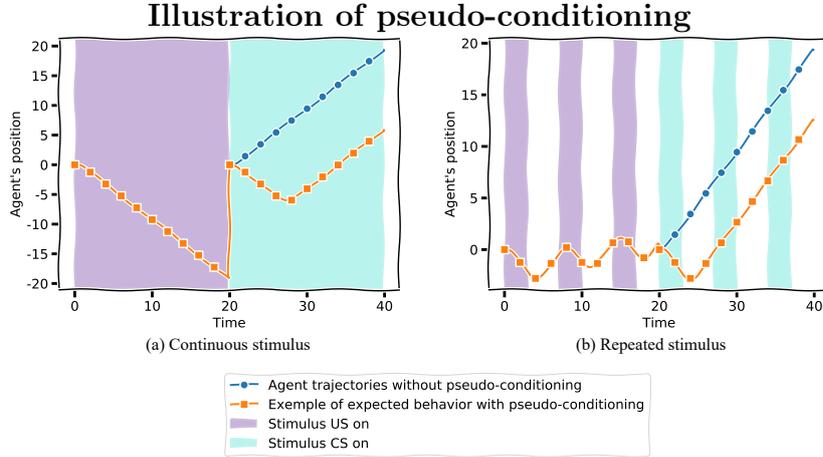


Fig. 1: Illustration of the expected behaviors, with or without pseudo-conditioning learning with continuous stimuli (a) or with repeated stimuli (b, 2 time units on, 2 time units off).

### 3 IDSM

The Iterant Deformable Sensorimotor Medium (or IDSM for short) of [5] is a conceptual model of habits formation from an artificial agent perspective as self-reinforcing sensorimotor behavioral patterns. It is an intermediate model between neuronal models and macroscopic behavior models.

The IDSM works in the sensorimotor space  $SM$  of the agent, with  $K = \dim(SM)$  the number of sensorimotor dimensions. As depicted on figure 2a, part of the environment state  $\mathcal{S}_E$  is abstracted into a sensorimotor state  $\mathcal{S}_{SM}$ . In the sensorimotor space, the IDSM uses  $\mathcal{S}_{SM}$  and some *nodes* that store preferred sensorimotor alterations to compute a command  $\mu$ . This command  $\mu$ , when applied on the agent in its environment, will alter the environment and the agent.

Formally, a node  $N$  of the IDSM is defined as  $N = \langle \mathbf{p}, \mathbf{v}, w \rangle$  with  $\mathbf{p}$  a vector indicating the node's position in sensorimotor space  $SM$ ,  $\mathbf{v}$  a vector indicating the desired velocity in sensorimotor space  $SM$  at the position  $\mathbf{p}$ , and  $w$  a scalar indicating the node's weight, which balances its influence relatively to the other nodes within the computation of the dynamic of the agent. The sensorimotor space is normalized and each dimension (either sensor or motor) lies in the range  $[0, 1]$ .

As illustrated on figure 2b the influence of each node consists of two components: an attraction component  $\mathbf{A}^i$  and a velocity component  $\mathbf{V}^i$ . These

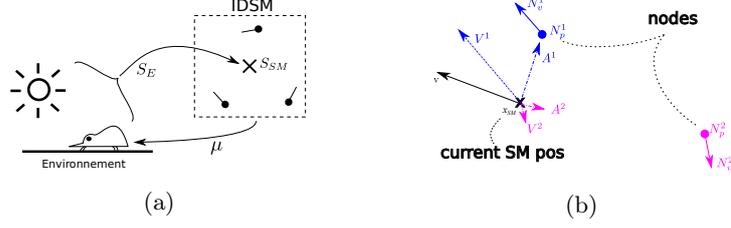


Fig. 2: a) An IDSM works in the sensorimotor space of the agent. According to the current sensorimotor state  $S_{SM}$  and to the various nodes of the IDSM, it produces commands  $\mu$  that control the agent behavior in its environment.  $S_E$  is the complete state of the environment. b) An IDSM is built from a collection of nodes  $N^i$  which influence the agent dynamics with a velocity  $\mathbf{V}^i$  and attraction  $\mathbf{A}^i$  components.

components impact the agent's velocity, the attraction component seeking to attract the agent towards its position  $\mathbf{N}_p$  and the velocity component seeking to align the agent's velocity with  $\mathbf{N}_v$ . The amplitude of the components is modulated by the distance of the agent to the node (the closer the higher) and by the weight of the node  $N_w$  (equations (1)-(8)). The velocity command  $\boldsymbol{\mu}$  of the agent, normalized in the sensorimotor space SM, is given by<sup>5</sup> :

$$\frac{d\boldsymbol{\mu}}{dt} = \begin{cases} \frac{\mathbf{V}(\mathbf{x}_{SM}) + \mathbf{A}(\mathbf{x}_{SM})}{\phi_N(\mathbf{x}_{SM})} & \text{if } \phi_N(\mathbf{x}_{SM}) > \epsilon \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$\mathbf{V}(\mathbf{x}_{SM}) = \sum_N \omega d_N \cdot [\mathbf{N}_v]^\mu \quad (2)$$

$$\mathbf{A}(\mathbf{x}_{SM}) = \sum_N \omega d_N \cdot [\Gamma(\mathbf{N}_p - \mathbf{x}_{SM}, \mathbf{N}_v)]^\mu, \quad (3)$$

$$\phi_N(\mathbf{x}_{SM}) = \sum_N \omega d_N \quad (4)$$

$$\omega d_N = \omega(N_w) \cdot d(\mathbf{N}_p, \mathbf{x}_{SM}) \quad (5)$$

$$\omega(N_w) = \frac{2}{1 + \exp(-k_\omega N_w)} \quad (6)$$

$$d(\mathbf{N}_p, \mathbf{x}_{SM}) = \frac{2}{1 + \exp(k_d \|\mathbf{N}_p - \mathbf{x}_{SM}\|^2)} \quad (7)$$

$$\Gamma(\mathbf{a}, \mathbf{N}_v) = \begin{cases} \mathbf{a} - \left\langle \mathbf{a}, \frac{\mathbf{N}_v}{\|\mathbf{N}_v\|} \right\rangle \frac{\mathbf{N}_v}{\|\mathbf{N}_v\|} & \text{if } \mathbf{N}_v \neq 0 \\ \mathbf{a} & \text{otherwise} \end{cases} \quad (8)$$

<sup>5</sup> vectors are depicted in boldface (e.g.  $\mathbf{v}$ ) or between square brackets (e.g.  $[f, 0, 0]$ ),  $[\mathbf{v}]^\mu$  is the projection of  $\mathbf{v}$  on the motor dimensions, and  $[\mathbf{v}]^\sigma$  is the projection on the sensor dimensions. Furthermore,  $\mathbf{1}$  is the vector where all components are 1 and  $\langle \dots \rangle$  denotes the dot product.

where  $k_w = 0.025$  and  $k_d = 30$ . The parameter  $k_d$  determines the radius and the amplitude of influence a node while  $k_w$  determines the influence of a node relative to the others. The differential equation (1) is integrated with forward Euler using  $\Delta t = 0.1$ . The function  $\Gamma(\mathbf{x}, \mathbf{y})$  keeps only the component of  $\mathbf{x}$  which is perpendicular to  $\mathbf{y}$ . In our work, we consider a simplified IDSM where the nodes weights  $N_w$  are not adapted.

#### 4 Model of pseudo-conditioning through generalization of nodes

Pseudo-conditioning is a form of learning allowing a behavioral response to a given stimulus to be expressed as response to another stimulus. As such, it is a kind of generalization of this response to other stimuli. The following mechanism aims at extending the response dynamics from one sensory dimension of the IDSM to another. We propose to equip every node with an adaptive mask which determines the extend of the influence of a node in the sensorimotor space. Although initially specific to the sensorimotor state of the node, pseudo-conditioning adapts the mask to widen the influence of that node. This adaptation allows the node to become "responsive" to sensorimotor states in which it was not previously responsive.

For the experiments on this model, there are three dimensions in the sensorimotor space  $SM$ : one motor dimension and two sensor dimensions: sensor  $s1$  perceiving the unconditioned stimulus  $US$  and sensor  $s2$  perceiving the conditioned stimulus  $CS$ . The nodes in the sensorimotor space  $SM$  are shown on Figure 3a, without representation of the second sensor dimension for simplicity (we only plot the subspace  $m \times s_1$ , which is independent on the specific  $s_2$  value). The two nodes are defined by:

- $N_1 = \langle \mathbf{p} = [0.5, 1, 0], \mathbf{v} = [-0.5, 0, 0], w = 0 \rangle$ ,
- $N_2 = \langle \mathbf{p} = [0, 0, 0], \mathbf{v} = [0.5, 0, 0], w = 0 \rangle$ .

Our objective with these two nodes is to produce a default behavior of moving upward without stimuli and downward when the unconditioned stimulus  $US$  (perceived by sensor  $s1$ ) is on. Since conditioned stimulus  $CS$  is neutral, the dynamics is unchanged on the dimension of sensor  $s2$ . Thus the dynamics boils down to two blocks, one above the plan  $s1 = 0.5$ , with the downward behavior activated by the  $US$ , and the other below the plan  $s1 = 0.5$ , with the default upward behavior.

We add a mask component  $\mathbf{N}_m$  of dimension  $K$  with components in  $[0, 1]$  to every node  $N$ .  $\mathbf{N}_m$  is used to compute a virtual *masked* distance  $dm$ , between the node  $N$  and the agent sensorimotor position  $\mathbf{x}_{SM}$ :

$$dm(\mathbf{N}_p, \mathbf{x}_{SM}) = \frac{2}{1 + \exp(k_d \|\mathbf{N}_p - \mathbf{x}_{SM}\|_m^2)} \quad (9)$$

$$\|\mathbf{N}_p - \mathbf{x}_{SM}\|_m^2 = \sum_i \mathbf{N}_{m_i}^2 (\mathbf{N}_{p_i} - \mathbf{x}_{SM_i})^2 \quad (10)$$

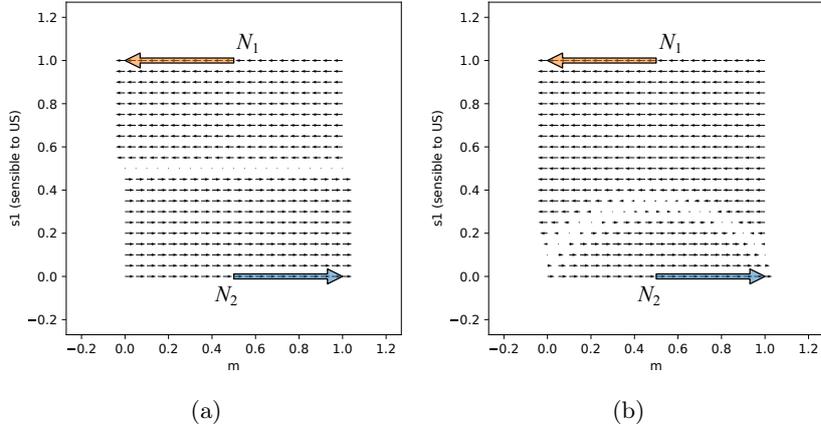


Fig. 3: Motor command variation ( $d\mu$ ) in the sensorimotor subspace  $m \times s_1$  of an IDSM with two nodes used in pseudo-conditioning experiments. Large arrows are nodes, with the base on the  $\mathbf{N}_p$  and the length of the arrow depicting  $\mathbf{N}_v$ . a) with a mask  $\mathbf{N}_m = [1, 1, 1]$ . b) with a mask  $\mathbf{N}_m = [0.5, 0.5, 0.5]$  for  $N_1$  leading to its stronger influence over  $N_2$ .

This distance  $dm$ , which artificially decrease the sensorimotor distance between the agent and the nodes, is used instead of  $d$  in equations 2.

It follows that for a given dimension  $i$ , the most specific the associated mask component can be is  $\mathbf{N}_{m_i} = 1$  (*i.e.* the distance between  $\mathbf{N}_{p_i}$  and  $\mathbf{x}_{SM_i}$  is fully taken into account). And the most general the associated mask component can be is  $\mathbf{N}_{m_i} = 0$  (*i.e.* the distance between  $\mathbf{N}_{p_i}$  and  $\mathbf{x}_{SM_i}$  is completely ignored, as if  $\mathbf{N}_{p_i} = \mathbf{x}_{SM_i}$ ).

A node starts with its mask equal to  $\mathbf{1}$ , the most specific possible. Then the mask is modified following:

$$f_{gen} = (\sqrt{K^\sigma} - \|\mathbf{x}_{SM}\|^\sigma - \|\mathbf{N}_p\|^\sigma) \beta_m \omega d_{mN} \quad (11)$$

$$\frac{\partial \mathbf{N}_m}{\partial t}(\mathbf{x}_{SM}, t) = \gamma_m \cdot \mathbf{1} - f_{gen} \cdot \mathbf{N}_m \quad (12)$$

Parameters  $\beta_m$  and  $\gamma_m$  adjust respectively the generalization and the specialization of the masks and are set to  $\beta_m = 3$  and  $\gamma_m = 0.05$ . Each of the mask components is kept in  $[0, 1]$ .

Thus, the smaller the masked distance between a node  $N$  and the sensorimotor position of the agent  $\mathbf{x}_{SM}$ , the more the mask of the node  $\mathbf{N}_m$  is decreased (and so becomes generalized). Specialization (*i.e.* increase) of the mask occurs slightly over time. Fig. 3b shows the modification of the subspace  $m \times s_1$  dynamics when the mask of node  $N_1$  is generalized to  $\mathbf{N}_m = [0.5, 0.5, 0.5]$ . The block of the default behavior is partly reduced to a quarter of ellipsoid with  $N_2$  as its center.

## 5 Experiments and results

### 5.1 Setup

In order to measure if our models are capable of pseudo-conditioning, we use the following protocol. What we seek to evaluate is the ability of the model to produce a generalized response to a previously neutral stimulus (the conditioned stimulus  $CS$ , perceived by sensor  $s2$ ), after continuous or repeated presentation of a non-neutral stimulus (the unconditioned stimulus  $US$ , perceived by sensor  $s1$ ).

The experiment is split into three phases of twenty time units. During the first phase, the  $CS$  is presented, to observe the agent behavior prior to pseudo-conditioning. The  $CS$  is then replaced by the  $US$  during the second phase. Lastly, the  $CS$  is presented again during the third phase. When either the  $CS$  or  $US$  is active, its value is set to 1. To limit the differences in experimental conditions, at the beginning of each phase, the agent is reset to a null velocity starting position ( $\mathbf{x}_E = \dot{\mathbf{x}}_E = \ddot{\mathbf{x}}_E = 0$ ) and a sensorimotor state  $\mathbf{x}_{SM} = [0.5, 0, 0]$ . Here, the choice of forcing the sensor value to 0 is made to have identical initial conditions between phases, in particular to compare the first and third phases. This can be interpreted as the insertion of a time unit without stimulus between each phase. As explained in section 4, the IDSM is modified accordingly.

In a given experiment, the stimulation follows one of the four patterns below. Note that for every experiment, all the stimulation (whether  $CS$  or  $US$ ) follow the same pattern.

- (a): continuous stimulus;
- (b): repetitions of 3 time units on / 1 time unit off;
- (c): repetitions of 2 time units on / 2 time units off;
- (d): repetitions of 1 time unit on / 3 time units off.

### 5.2 Results

The results can be seen on figs. 4 and 5. For the control agent, the default behavior (upward motion) is observed during the first and third phase, with presentation of the  $CS$ . During the second phase, with presentation of the  $US$ , a response behavior (downward motion) is observed.

Each figure is composed of four subplots, one for each stimulation pattern. Each subplot has the same layout. The top plot displays the learning agent trajectory and the control agent trajectory. The second plot from the top gives information regarding perception of both stimuli by the agent. The third plot from the top shows the normalized motor value of the agent in the sensorimotor space (the first component of  $\mathbf{x}_{SM}$ ), and the velocity command  $\boldsymbol{\mu}$ . In order not to overload the figure unnecessarily, the last plot shows the evolution of the nodes masks reduced only to their motor component (the mask of other dimensions, *i.e.* first sensor sensible to the  $US$  and second sensor sensible to the  $CS$ , having exactly the same values).

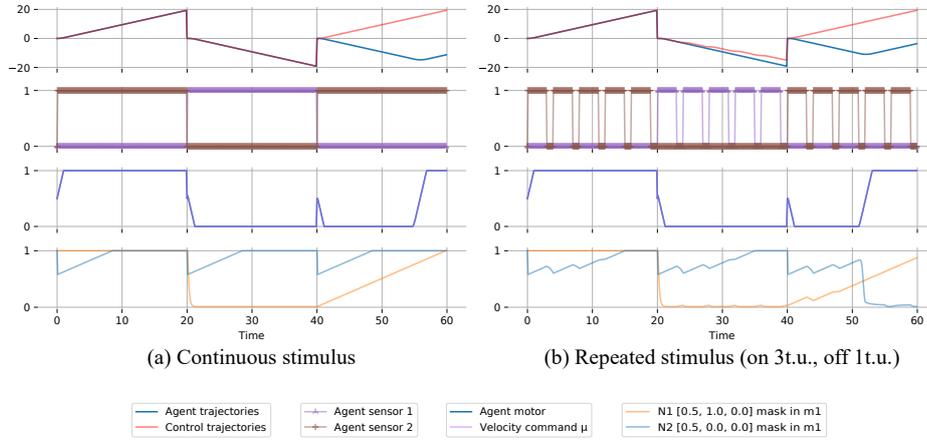


Fig. 4: Results for activation patterns (a) and (b). From top to bottom: agents trajectories, sensory activation, motor command and masks. After time 40, the stronger generalisation of  $N_1$  compared to  $N_2$  (see masks plot), leads to pseudo-conditioned behavior of the agent, seen as "learning" to move downward when the  $CS$  is perceived.

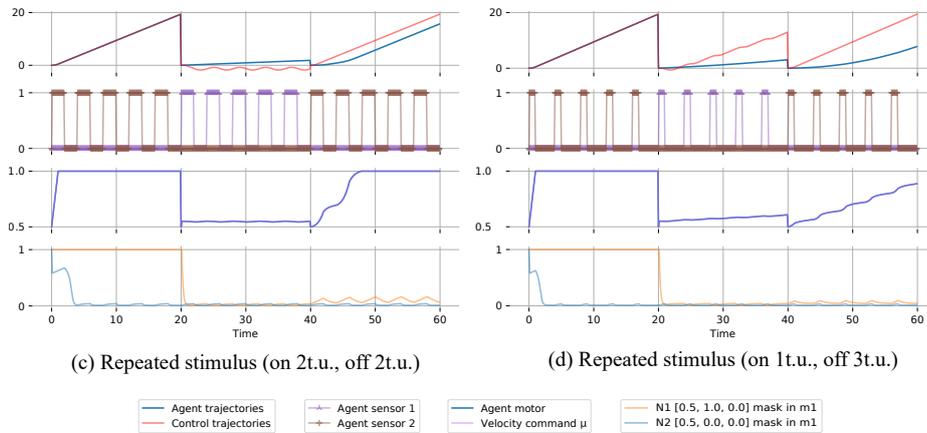


Fig. 5: Results for activation patterns (c) and (d) where the activation time equal or superior to the deactivation time. From top to bottom: agents trajectories, sensory activation, motor command and masks. These patterns highlight that in phase 2 the generalization of masks leads to equal contributions of the 2 nodes and so, to the immobility of the agent. In phase 3, masks values strongly inhibit the expression of the moving down behavior.

With all the patterns of stimulation, the behavior of the learning agent during the third phase (*i.e.* presentation of the *CS* following presentation of the *US*) is a mix between its behaviors of the first and second phases, starting as a "*US* influenced" behavior and ending as a "*CS* influenced" behavior (default behavior).

In figure 4a, pseudo-conditioning has no influence in first and second phases. The pseudo-conditioning occurs as expected in the last phase, allowing the agent to experiment a "new" behavior (*i.e.* temporarily modified dynamics in the sensorimotor space *SM*): going downward while perceiving only the *CS*. The evolution of the masks shows that only the node  $N_1$ , responsible for the response to the *US*, is maximally generalized. At all times, the mask of  $N_1$  is more generalized than the mask of  $N_2$ . Constant specialization of the mask of  $N_1$  allows the default behavior to be expressed again at the end of the third phase. The node  $N_2$ , responsible for default downward behavior, only generalizes slightly and temporarily following the reset of the agent at the start of each phase.

In all repeated patterns (see figs. 4b and 5), the mask of node  $N_2$  generalizes each time there is no active stimulus. In the pattern **(b)**, with the shorter deactivation time between stimulus activation, the learning agent in the second phase does not respond to deactivation of the stimulus since  $N_1$  stays more generalized than  $N_2$  and behaves as if the *US* was activated continuously during the first two deactivation. However, the balance between nodes generalisation during the third deactivation allows  $N_2$  to become "stronger" than  $N_1$ , hence the switch in the expressed behavior. In the patterns **(c)** and **(d)**, with the longer deactivation time, generalization of node  $N_2$  leads to a tie between the two nodes influences, which results in weaker motor commands and extends the time needed for the agent to move.

In all patterns, we observe that the resulting behavior in third phase is the result of the ratio of influence between the generalization of  $N_1$  when *US* is perceived, and the generalization of  $N_2$  when no stimuli are present. We thus have a mechanism of generalization of *behavior* which is activated more or less strongly according to the time of expression of the behavior in a certain sensorimotor context.

## 6 Discussion

We have proposed a model to provide pseudo-conditioning (with nodes masks) to an agent based on the IDSM. Since IDSM is a conceptual model, our model is easier to read than models based on neural networks.

The generalization of the nodes through the masks results in a transfer of the behavioral response from the stimulus initially triggering it to other stimuli. Nevertheless, our model also allows for the generalization of masks of nodes which are "linked" to no stimuli (sensori components of the position of such nodes are all set to 0). We do not know if it makes sense from the point of view of biology, since by definition pseudo-conditioning occurs after a stimulus has induced a response behavior.

The mask mechanism we propose for pseudo-conditioning leads to a widespread generalization: once a node is fully generalized it impacts the behavior of the agent equally everywhere in the sensorimotor space. This generalized behavior is only counterbalanced by the existence of other nodes (through the usual IDSM equations). This "competition" between nodes may not be sufficiently strong to preserve other existing behaviors. In such a case, it may be needed to introduce a specialization mechanism shaping the sensorimotor regions where the node should actually generalize, or restrict the dimensions on which to do so. For example, specialization could be triggered by the discrepancy between the component  $\mathbf{N}_v$  of node  $N$  and the last movement of the agent in the sensorimotor space  $SM$ , since such difference would mean that the motor command stored in  $N$  is not in line with the current environment. Adding another non-associative learning like habituation (*i.e.* inhibition of the response to a repeatedly presented stimulus) could also help prevent over-generalization by minimising the impact of the repetitions of low-intensity stimuli. We started to work on those non-associative learning ([9]). Moreover the dual-process theory (see [10]) and the behavioral homeostasis theory (see [8,6]) are works studying how non-associative learning fit together and what could be the impacts of such association.

Alternatively to the masks model, even in the context of the IDSM, we may have proposed other mechanisms allowing the same pseudo-conditioning as experimented in this paper. As the aim is to adapt the sensorimotor region of influence of a node, we could have simply defined and adapted one  $k_d$  per node, which would have resulted in similar results. However, the mask mechanism is strictly more general in the sense that it allows (although not illustrated here) to generalize along certain dimensions while still being specific along others.

Note that the results of our experiments depend strongly on many parameters. The number of nodes, their positions, the various IDSM parameters are very inter-dependant and were carefully chosen. We considered a very simple IDSM for experimenting the mask mechanism and more complex settings with more nodes and more complex behaviors remain to be studied. In addition, the original IDSM model includes an evolution of the weights and the generation of new nodes which were here removed to ease the analysis. The idea is that the mask mechanism should allow the agent to experiment known behaviors in unknown situations, which will eventually be memorized by the creation of new nodes.

To conclude, the presented mechanisms could be seen as a first step to allow a reflex-based agent to "experiment" different behaviors in various situations depending on its recent experience. Further studies are needed to explore to which extend such non-associative mechanisms allow for an agent to extend its behavioral repertoire. Our work is a proof of concept since we considered very simple environments and we still need to explore the applicability of these mechanisms in more complex and challenging environments.

## References

1. Alex, J.Y., Rankin, C.H.: Nonassociative learning in invertebrates. In: The Oxford Handbook of Invertebrate Neurobiology. Oxford University Press (2017)

2. Byrne, J.H., Hawkins, R.D.: Nonassociative learning in invertebrates. *Cold Spring Harbor perspectives in biology* **7**(5), a021675 (2015)
3. Damper, R.I., French, R.L., Scutt, T.W.: Arbib: An autonomous robot based on inspirations from biology. *Robotics and Autonomous systems* **31**(4), 247–274 (2000)
4. Dussutour, A.: Learning in single cell organisms. *Biochemical and Biophysical Research Communications* **564**, 92–102 (2021)
5. Egbert, M.D., Barandiaran, X.E.: Modeling habits as self-sustaining patterns of sensorimotor behavior. *Frontiers in human neuroscience* **8**, 590 (2014)
6. Eisenstein, E.M., Eisenstein, D.L., Sarma, J.S., Knapp, H., Smith, J.C.: Some new speculative ideas about the “behavioral homeostasis theory” as to how the simple learned behaviors of habituation and sensitization improve organism survival throughout phylogeny. *Communicative & Integrative Biology* **5**(3), 233–239 (2012)
7. Eisenstein, E., Brunder, D., Blair, H.: Habituation and sensitization in an aneural cell: Some comparative and theoretical considerations. *Neuroscience & Biobehavioral Reviews* **6**(2), 183–194 (1982)
8. Eisenstein, E., Eisenstein, D., Smith, J.C.: The evolutionary significance of habituation and sensitization across phylogeny: A behavioral homeostasis model. *Integrative Physiological & Behavioral Science* **36**(4), 251–265 (2001)
9. Gillard, T., Dutech, A., Fix, J.: Non-Associative Learning and the Iterant Deformable Sensorimotor Medium. Research Report 9472, Inria Nancy - Grand Est (May 2022)
10. Groves, P.M., Thompson, R.F.: Habituation: a dual-process theory. *Psychological review* **77**(5), 419 (1970)
11. Hawkins, R.D.: A biologically based computational model for several simple forms of learning. In: Hawkins, R.D., Bower, G.H. (eds.) *Computational Models of Learning in Simple Neural Systems, Psychology of Learning and Motivation*, vol. 23, pp. 65–108. Academic Press (1989)
12. Hawkins, R.D., Kandel, E.R.: Is there a cell-biological alphabet for simple forms of learning? *Psychological review* **91**(3), 375 (1984)
13. Hong, Q., Yan, R., Wang, C., Sun, J.: Memristive circuit implementation of biological nonassociative learning mechanism and its applications. *IEEE Transactions on Biomedical Circuits and Systems* **14**(5), 1036–1050 (2020)
14. Kimble, G.A.: *Hilgard and Marquis Conditioning and learning*. Prentice-Hall, Englewood Cliffs, 2nd edn. (1961)
15. Kotseruba, I., Tsotsos, J.K.: 40 years of cognitive architectures: core cognitive abilities and practical applications. *Artificial Intelligence Review* **53**(1), 17–94 (2020)
16. McSweeney, F.K., Murphy, E.S.: Sensitization and habituation regulate reinforcer effectiveness. *Neurobiology of learning and memory* **92**(2), 189–198 (2009)
17. Novianto, R., Johnston, B., Williams, M.A.: Habituation and sensitisation learning in asmo cognitive architecture. In: *International Conference on Social Robotics*. pp. 249–259. Springer (2013)
18. Perry, C.J., Barron, A.B., Cheng, K.: Invertebrate learning and cognition: relating phenomena to neural substrate. *Wiley Interdisciplinary Reviews: Cognitive Science* **4**(5), 561–582 (2013)
19. Rescorla, R.A.: A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Current research and theory* pp. 64–99 (1972)
20. Stanley, J.C.: Computer simulation of a model of habituation. *Nature* **261**(5556), 146–148 (1976)
21. Weng, J.: On developmental mental architectures. *Neurocomputing* **70**(13-15), 2303–2323 (2007)