



HAL
open science

Gated spiking neural network using Iterative Free-Energy Optimization and rank-order coding for structure learning in memory sequences (INFERNO GATE)

Alexandre Pitti, Mathias Quoy, Catherine Lavandier, Sofiane Boucenna

► **To cite this version:**

Alexandre Pitti, Mathias Quoy, Catherine Lavandier, Sofiane Boucenna. Gated spiking neural network using Iterative Free-Energy Optimization and rank-order coding for structure learning in memory sequences (INFERNO GATE). *Neural Networks*, 2020, 121, pp.242-258. 10.1016/j.neunet.2019.09.023 . hal-02309379

HAL Id: hal-02309379

<https://hal.science/hal-02309379>

Submitted on 11 Oct 2019

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.

Gated Spiking Neural Network using Iterative Free-Energy Optimization and rank-order coding for structure learning in memory sequences (INFERNO GATE)

Alexandre Pitti^{1,*}, Mathias Quoy¹, Catherine Lavandier¹, Sofiane Boucenna¹
Laboratoire ETIS UMR 8051, Universit Paris-Seine, Universit de Cergy-Pontoise, ENSEA, CNRS, France

Abstract

We present a framework based on iterative free-energy optimization with spiking neural networks for modeling the fronto-striatal system (PFC-BG) for the generation and recall of audio memory sequences. In line with neuroimaging studies carried out in the PFC, we propose a genuine coding strategy using the gain-modulation mechanism to represent abstract sequences based solely on the rank and location of items within them. Based on this mechanism, we show that we can construct a repertoire of neurons sensitive to the temporal structure in sequences from which we can represent any novel sequences. Free-energy optimization is then used to explore and to retrieve the missing indices of the items in the correct order for executive control and compositionality. We show that the gain-modulation mechanism permits the network to be robust to variabilities and to have long-term dependencies as it implements a gated recurrent neural network. This model, called Inferno Gate, is an extension of the neural architecture Inferno standing for Iterative Free-Energy Optimization of Recurrent Neural Networks with Gating or Gain-modulation. In experiments performed with an audio database of ten thousand MFCC vectors, Inferno Gate is capable of encoding efficiently and retrieving chunks of fifty items length. We

*Corresponding author

¹surname.name@u-cergy.fr

then discuss the potential of our network to model the features of working memory in the PFC-BG loop for structural learning, goal-direction and hierarchical reinforcement learning.

Keywords: Free-energy, Rank-Order Coding, Structural learning, Compositionality, Gated spiking networks, Predictive Coding, Prefrontal cortex, Language development

1. Introduction

1.1. Proposal framework for sequence learning

In this paper, we propose to use the neural architecture Inferno, standing for Iterative Free-Energy Optimization in Recurrent Neural Network, for the learning of temporal patterns and the serial recall of sequences [1, 2]. We originally proposed this neuronal architecture to model the cortico-basal ganglia loop [1] for retrieving motor and audio primitives using Spike Timing-dependent Plasticity (STDP) within the framework of predictive coding and free-energy minimization [3, 4, 5]. Here, we propose to implement a similar free-energy minimization network but this time in the prefrontal-basal ganglia loop for the serial recall of memory sequences and for the learning of temporal pattern primitives, using gain-modulation instead of STDP. Since this working memory uses gain-modulated or gating cells instead of STDP, we propose to name it Inferno Gate in order to disambiguate this architecture from our original network.

Gain-modulation will serve to model neurons salient to the temporal order of items and their sequential organization. As we will explain further, prefrontal units depend crucially on this type of coding for serial recall. They support a gain-modulated mechanism to jointly encode items and rank-order information in a sequence [6]. This mechanism of gain-modulation is also described as a gating or conjunctive function in other research [7], placing more emphasis on the properties of filtering out or holding on to information.

We will show that Inferno Gate is capable of learning temporal primitives sensitive to the serial order of items within sequences, coding abstract temporal

sequences *without* information about items, and accurately retrieving sequences
25 of items with respect to the given serial order information *only*. We think its
architecture makes it robust for structural learning, model-based reinforcement
learning and compositionality.

Our main contributions are to propose a neuro-computational architecture
of the PFC and a novel mechanism to encode temporal sequences in an efficient
30 way for language processing. The neuro-computational architecture Inferno
Gate learns, recognizes and retrieves missing elements in memory sequences
based on an original encoding mechanism that represents sequences with a
distributed neural population of temporal primitives learnt. These temporal
primitives are abstract patterns constructed from information about the rank
35 order of the items within sequences, *without* their index per se. This new brain-
inspired encoding based on spikes makes the representation of sequences more
compact, the learning faster, and the retrieval of missing items more efficient
than the encoding performed in conventional neural networks and possibly deep
networks [8].

40 The paper is organized as follows. We will present first the developmental
and neural foundations of our neural architecture and its purpose. In second,
we will present a state of art of prefrontal models and justify how our model is
original in comparison to them.

We will detail then the neural mechanisms used. We explain how an analog
45 gating can be created with spiking neurons and how gain-modulated neurons can
represent a compact code for sequences. In comparison to other gain-modulation
architectures that require a one-to-one conversion matrix necessary for multi-
plicative binding –, which consumes neurons for this computation,– we discov-
ered that a rank-order coding algorithm can model gain-modulation in a more
50 efficient manner with spiking neurons.

We apply this network for the learning of temporal primitives from audio
sequences. These primitives are then used for representing and recalling these
audio sequences with a length of one second (1000 milliseconds), corresponding
to chunks of 50 items' length, despite information about the items' index (their

55 content or their identity) being lost.

We then discuss the originality of our approach and implications in terms of computation for modeling sequences, extracting temporal tree structure-like patterns, and compressive coding of grammar-like models, recursive representation, compositionality and transfer of learning.

60 1.2. Developmental and neural foundations

During early development, infants are keen on grasping structure in several core domains [9, 10], inferring causal models and making hypotheses like little scientists [11, 12]. They rapidly develop knowledge about numerosity, space, physics and psychology but it is only at around 8 months that they gain the
65 aptitude to make complex sequences and to retain structural information in their environment.

In language acquisition, this skill is central for word segmentation and for detecting grammatical and ungrammatical sentences [13, 14]. For instance, infants are sensitive to the temporal order of events in spoken words and in
70 music so that they can be surprised if one syllable is changed or if one sound is removed, violating their prior expectations [15].

It is at this period, too, that the prefrontal cortex (PFC) develops. The prefrontal circuits comprise a working memory for executive control and planning that evaluates sequences online based on uncertainty [16], and select/unselect
75 them according to the current context, or create new ones if any are satisfying [17, 18, 19].

More than any other brain areas, the PFC can extract abstract rules and parametric information within structured data in order to carry out a plan [20, 21, 22]. This aspect makes it particularly important for problem-solving tasks,
80 language and maths [23, 24, 25, 26].

Experiments carried out on subjects performing hierarchical tasks such as drawing a geometrical figure [27, 28] or detecting temporal patterns within action sequences [29, 30] have permitted identification of some properties of PFC neurons for binding features and for higher-order sequence planning. In series of

85 observations done on PFC neurons, a critical finding was that sequences were en-
coded through a conjunctive code, which crosses items and serial orders [31, 32].
In similar experiments performed by Inoue and Mikami, some PFC neurons were
found to modulate their amplitude level with respect to the position of items
during the sequential presentation of two visual shape cues [33]. The PFC neu-
90 rons displayed graded activity with respect to their ordinal position within the
sequence and to the visual shapes; e.g. first-ranked items, or second-ranked
items. In more complex tasks, PFC neurons were found to fire at particular
moments within the sequence [21]; e.g. the beginning, the middle, the end, or
even throughout the evolution of the sequence.

95 Despite these findings, the precise role played by conjunctive cells in the PFC
and the mechanisms behind the process are still under investigation. In contrast,
the conjunctive cells in the Parietal Cortex have been studied more frequently
and many neurocomputational models explain how they contribute to spatial
representation [34, 35], coordinate transformation [36, 37] and numerosity capa-
100 bilities [38]. In most research, conjunctive cells or gain-modulation neurons in
parieto-motor neurons are seen as a way of binding different received informa-
tion (e.g. in vision and proprioception) for preparing an action (e.g. reaching
a target). In [39, 37], Pouget proposes that gain-modulated conjunctive cells
in the Parietal Cortex can serve as radial basis functions for constructing any
105 spatial metric; e.g., a hand-centred relative metric [40, 41], a head-centred re-
lative metric [34]. Similar to the role played by conjunctive cells in the spatial
domain in the Parietal Cortex, we suggest that the conjunctive cells in the PFC
play the role of radial basis functions in the temporal domain to decompose and
code sequences. Gain-modulation in the PFC may serve to extract temporal
110 patterns and to represent them as primitives for encoding existing sequences or
for generating new ones, see Fig. 1 a-b).

This idea is in line with comparative neuroanatomical studies which attribute
similar functions to the parietal cortex and to the prefrontal cortex, representing
relative metrics or conjunctive representations [42] such as order with relative
115 duration, and order with relative distance; but only the PFC is in a position to

generate goal-based aims in context [43]. This is also suggested by Botvinick and Watanabe in [6] that these cells in the PFC describe a compressive representation of sequences without items. Gain-modulated conjunctive cells can give an insight into how the PFC manages to plan sequences and encode them [25].
120 For instance, they may be seen as a solution to disentangle the features (items) from the sequence (ordinal information) in planning. In line with this idea, they may gate information at particular moments – i.e., not only predicting which action to perform but also knowing *when* to do so within a sequence [44, 45]. Their role may be in line with other frameworks in which neuronal “pointers” or
125 “fillers” or “timestamp” neurons, are proposed for binding or gating information with respect to the current inputs [46, 47, 48, 49].

Furthermore, this mechanism may serve to construct a basis for composing any sequences, recombining items in different orders and generating novel sequences with different items, see Fig. 1 b). This capacity for combinatorial
130 re-use is particularly robust and specific to human behavior, which corresponds to the capacity for compositionality, hierarchy and systematicity found in languages and structured grammars [50, 51, 52, 53, 54, 55, 25].

Since the seminal work of Broca, we know that the circuits in the left cortical hemisphere and in the pre-frontal area implement language with the Broca
135 area forming a syntax-sensitive system (grammar) for perception and production of semantic and rule-based behaviors. Although the mechanism behind its functional organization is relatively unknown, we suggest that gain-modulated conjunctive cells may shed lights on it.

2. State of the art and model justification

140 2.1. Prefrontal models and gated neural networks

In prefrontal models for sentence processing, Dominey proposed earlier versions of echo-state networks to model the associative memory in the corticostriatal system for language processing and sequence learning [56, 57, 58].

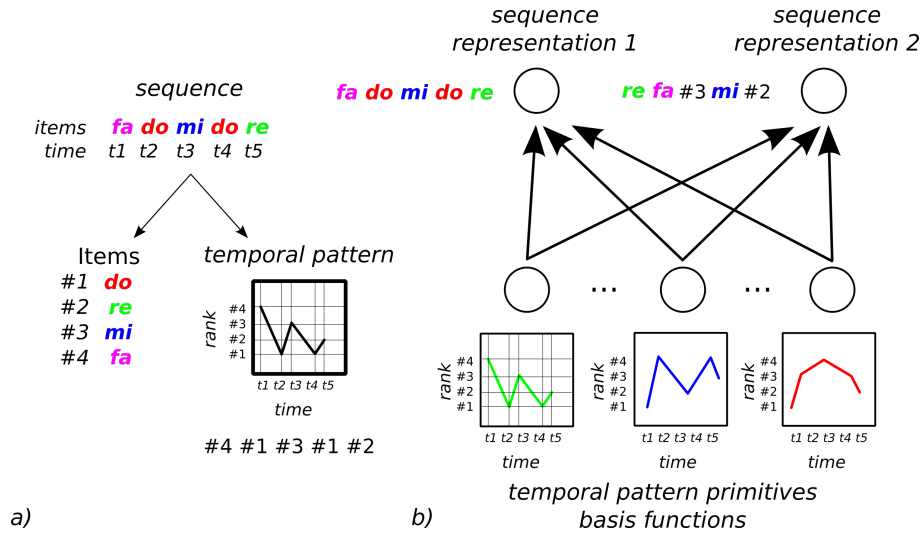


Figure 1: Gating operation for feature and structural separation in sequence learning. In a), we can discriminate the items' index (rank #) from their position (order) to represent one sequence. By separating the two, we can extract the temporal pattern and arrange items in a different order. Hence, the coding of the temporal pattern can make it robust to variability and can represent many sequences (generalization). This process is operated by a gain-modulation or gating mechanism explained later. In b), the combination of these temporal patterns can serve to compose any novel temporal pattern in the same fashion as radial basis functions would do.

These architectures differ from other types of recurrent neural networks
 145 for serial recall and online activation, which can nonetheless show complex sequences of ordering [59] and error-based predictive coding [49, 60, 61].

Different PFC neural architectures from dynamical systems theory have been
 proposed to code and retrieve memory sequences based on phase synchronisation for feature binding such as the LISA architecture [62], on chaotic networks
 150 as in [63, 64, 65, 66, 67], or on reservoir computing networks [68]. For instance, echo-state networks have been utilized for modeling the learning of structure and the acquisition of a grammar of rules [58, 69, 70, 71]. Despite having many advantages as a dynamical system to embed attractors, the learning phase is almost always carried out offline with supervised learning for labelling the pat-
 155 terns. Meanwhile, the process needs to be properly initialized to be effective,

and the way structural information (topology) is embedded within these networks is also not clearly defined as it is often used as a black box. Furthermore, it is not clear how they can support other coding strategies such as the gating mechanism or other learning mechanisms such as reinforcement learning.

160 Jun Tani proposed neural architectures to manage dynamics with recurrent neural networks at multiple temporal scales (MTRNN) and with parametric bias (RNNPB) for learning the attractors (abstract temporal structures) in sequences [51, 72, 54]. In neurorobotic experiments, within the framework of dynamical systems and chaos theory, he showed how recurrent neural networks
165 can embed several dynamics as symbolic units (rhythmical and sequential) for robot control, imitation and social interaction, giving rise to compositionality. These networks have some links with the predictive coding framework. For instance, the parameter bias in RNNPB can be seen as a control variable to select one specific sequence while the hierarchical control done in MTRNN permits to
170 encode and retrieve sequences at multiple temporal scales. The extension into a complete predictive coding framework was proposed recently with the embedding of error prediction signals with the network P-MSTRNN [73]. Similar ideas on hierarchical reinforcement learning and executive control in PFC can be found in recent works by [74, 75].

175 In contrast to these algorithms, another family of PFC models is based on a gating or gain-modulation mechanism. In the literature, this corresponds to the Long-Short Term Memory (LSTM) [76, 77, 22], the gated prefrontal networks with stripes by O'Reilly and colleagues [78, 18], the SPAWN architecture with neuronal pointers by Eliasmith and colleagues [47, 79], or the prefrontal
180 architectures that explicitly use gain-modulation as in [6] and in [7].

O'Reilly and colleagues attribute to the PFC neurons the role of variable binders to identify rules in the sentences (subjects, verbs, complements) and to process new ones by filling the holes with current values (e.g. binding new words) [19, 80]. Current models of the PFC show the importance of gating
185 networks [76, 18, 19, 7, 22, 47]. Gated information is particularly useful to maintain contextual variables for several cycles in order to reuse them later

or to process new memories from them. For some models, such as the long short-term memory (LSTM) networks [76], these algorithms have proved their robustness in spite of their lack of accessibility and biological plausibility. In literature, the gating mechanism is mostly understood as an on/off switch for maintaining or shunting memories. In comparison, the gain-modulation mechanism is very similar to a gating mechanism except that it places more emphasis on the binding of the signals from each other. For instance, this *analog* gating can serve to bind the relative order of items within temporal sequences and to retrieve them as suggested in [30, 6, 25].

In more general frameworks without close bio-inspiration, we can cite the works by Kemp, Lake and Tenenbaum who proposed several architectures based on Bayesian theory for probabilistic encoding and compositional capabilities [52, 12, 81]. Using a Bayesian framework for generating probabilistic models, their model could extract primitives from motor sequences to construct new symbols of the same types, differently combined.

In this line of research, neural networks with attention mechanisms have been proposed to overcome the limitations of current memory networks by learning both content and location of the information in the input sequence to decode later the output sequence. The attention mechanism allows the memory networks to learn where to pay attention in the input sequence for each item in the output sequence. We can cite the Neural Turing Machine, which is basically a Turing machine (read and write heads on a memory block) with a LSTM (or sometimes simple neural networks) based controller [82, 83], and other memory-augmented neural networks for neural attention [84, 85].

Novel memory networks under the predictive coding paradigm extend this attention mechanism to incorporate control and Bayesian inference based on error prediction at various hierarchical (temporal) levels. We can cite the memory networks REPRISSE [86], P-MSTRNN [73] and DRAW as well [87]. This ability makes these attentional and predictive networks good candidates for hierarchical reinforcement learning, as they are capable of both long-term storage via slow up-dates of their weights, and short-term storage via an external memory

module or via an attention buffer (the working memory).

Our neural architecture Inferno Gate embeds similar features found in those
220 networks namely the gating mechanism, structural learning, compositionality as
well as a generative and attentional mechanism. To our knowledge it is the first
time that a gated spiking recurrent neural network is proposed with results on
sequence learning comparable with deep recurrent networks. We will develop
hereinafter its neuro-biological foundations.

225 *2.2. Prefrontal functional organization for model-based reinforcement learning*

We justify our neural architecture from the work done on several neural
structures in the PFC identified for serial recall and the temporal organization
of behavior [88, 42, 58, 23, 24].

Functional imaging studies suggest that the PFC provides top-down support
230 for organizing the orderly activation of lower stages of the executive hierarchy
in sequences of actions (e.g., goal representations). Koechlin and colleagues
propose that the PFC subserves executive control and decision making in the
service of adaptive behavior [89, 23, 24].

In order to sustain such adaptive behaviors, it has been proposed that the
235 working memory in the PFC has to embed mechanisms for flexibility [90, 91]
to maintain memory sequences over a long time range in a hierarchical manner
and to explore new behavioral strategies. Such mechanisms have typically been
proposed within the inferential Bayesian theory [17, 16, 92] and within a rein-
forcement learning framework [88, 93, 23]. Such an approach has been extended
240 to PFC models based on predictive coding, and to free-energy minimization.

In line with these studies and the models proposed in [23, 93], we present
a detailed neural architecture in Fig. 2 in which we identify the Brodmann
area B45, the lateral PFC (IPFC), the dorsolateral PFC (dlPFC), the Orbito-
Frontal Cortex (OFC) and the Anterior Cingulate Cortex (ACC) to participate
245 in a model-based RL system for the active inference of memory sequences.

First, we suggest that the group constituted by B45, the IPFC and the
dlPFC are associated with the representation of the temporal organization of

sequences. Second, we suggest how the group constituted by the dlPFC, ACC
and OFC interact for decision-making, executive control and problem-solving
tasks.

We base our assumptions principally on the review papers in [44, 94, 58, 42,
23, 93]. For instance, in our schema, the rostral lateral PFC in both monkeys
and humans (typically Brodman areas 9/46 in [58, 42]) has been identified as
grouping actions for each other, for episodic or temporal control. We therefore
associate the Brodman area B45, part of the Broca area, with the chunking of
raw memory sequences coming from parietal and striatal areas.

At the second stage, the lateral PFC (lPFC) appears to proactively build
actor task sets from long-term memory that match the context in which the
individual is acting [44, 42]. Neurons in lPFC represent the long-term memory
of behavioral sequences and the plans or schemas of action [44]. These actor
task sets correspond in our view to temporal patterns or sequence prototypes
as presented in Fig. 1.

At the third stage, the dorsal lateral PFC (dlPFC) appears implicated in the
temporal integration of information for the attainment of prospective behavioral
goals [44, 23]. Reports suggest the involvement of the dlPFC for order memory
in term of choosing the correct sequence from among several. In our schema,
the dlPFC combines the temporal primitives of the lPFC to have an estimate
of the most suitable sequence.

The group constituted by the dlPFC, ACC and OFC networks appears in-
volved in a model-based RL working memory for which the anterior cingulate
cortex (ACC) seems to be active in the motivation to perform goal-directed
action (the task context units in the dlPFC), whereas the orbitofrontal cor-
tex is involved in value-based decision-making in novel choices. According to
Fuster, the orbital PFC, which is well connected with the brainstem and limbic
formations, plays a major role in the control of basic drives [44]. The OFC
might realize the downward trend or cascade of the processing of decided goal-
directed actions (concrete sequences) that have been decided upon, and repre-
sent option-specific state values [42]. The OFC might involve competition for

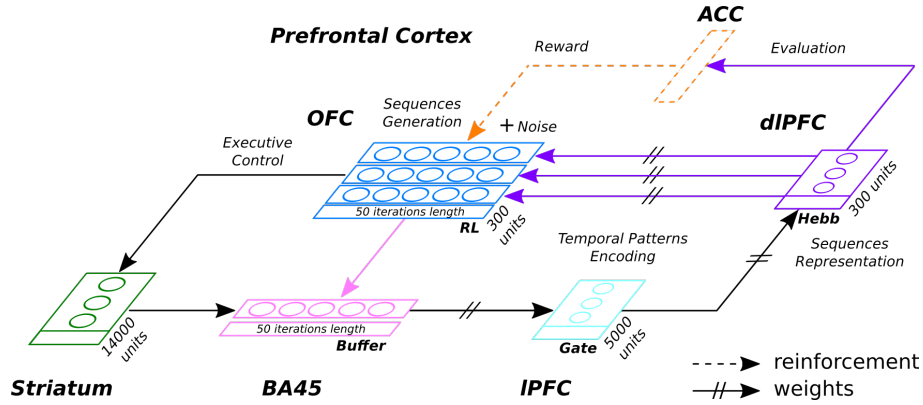


Figure 2: Framework for sequence learning based on iterative optimization in Fronto-Striatal (PFC-BG) loop. Our putative architecture follows the models proposed by [88, 23]. The Striatum (STR) represents the action/sound units that are assembled into a sequence at the Brodman area B45, part of the Broca area, in order to form a chunk. This chunk is read by the lateral PFC layer (IPFC), which learns and recognizes the temporal patterns via a gain-modulation/gating mechanism. The different temporal patterns in the IPFC are assembled in the dorso-lateral PFC layer (dIPFC), whose units represent the temporal structure within the sequence at a higher level. Based on the evaluation of the dIPFC, the Orbito-Frontal layer (OFC) and the Anterior Cingulate layer (ACC) generate and select sequences that follow the temporal patterns in the dIPFC in order to retrieve the lost indices of STR units for executive control. This reinforcement learning stage corresponds to a free-energy minimization process to reduce error prediction. The framework follows the Inferno architecture proposed in [1, 2] for the cortico-basal ganglia (CX-BG) loop.

280 decision-making among multiple choices. The ensemble is organized for the serial order encoding of sequences in the dIPFC and the exploration and recall of sequences in the OFC.

3. Methods

In section 3.1, we present the neural architecture INFERNO Gate used for serial recall in audio sequences associated with the PFC-Basal Ganglia loop. We then describe in section 3.2 the coding mechanism used for learning the serial order of items within sequences using the rank-order coding algorithm for modeling the gain-modulation mechanism with spiking neurons. Finally

285

in section 3.3, we define the experimental setup and the parameters used for accurate recall of long-range speech sequences.

290 *3.1. The network architecture Inferno Gate*

The neural architecture Inferno Gate reproduces the main configuration of the original Inferno network [1] with two coupled learning systems that minimize their mutual prediction error (the free-energy), see Fig. 2. The two networks both use spiking neurons and the difference between the two comes from the types of coding employed to represent temporal events: the original Inferno network employs the spike timing-dependent plasticity mechanism whereas the second one uses gain-modulation.

Considering the global architecture in Fig. 2, the two learning systems (lPFC/dlPFC and dlPFC/OFC) correspond to two associative networks of spiking neurons (SNNs) similar to radial basis functions. Bidirectionally coupled, the first SNN (lPFC/dlPFC) implements a forward model of the incoming signals while the second SNN (dlPFC/OFC) implements an inverse model aimed at retrieving and controlling those signals. The two learning systems can be viewed as an inverse-forward controller that can be modeled with the function $Y_{out} = f(I)$ for the first SNN and with the function $I = g(Y_{out})$ for the second one, in which I is the input vector and Y_{out} are the output dynamics. I is a sequence of Striatal units over time.

In order to minimize error, the second network (dlPFC/OFC) generates intrinsic noise I_{noise} to control the dynamics of the first one (lPFC/dlPFC) following a reinforcement learning (RL) mechanism. In Fig. 2, this role is devolved to the ACC for error evaluation. The activity level of one unit in dlPFC, $Y = Y_{out}$, is compared to its maximum amplitude level Y_{max} in order to compute the error E between Y_{max} and Y_{out} . The current input $I(t) = I_{noise}$ is kept for the next step $I(t+1) = I(t) + I_{noise}$, if and only if it diminishes the gradient ΔE . Over time, I converges to I_{opt} , its optimum sequence vector, and Y_{out} converges to Y_{max} its maximal value. This scheme is in line with predictive coding algorithms and its organization is similar to novel architectures combining two

or more competitive neural networks such as auto-encoders or the generative adversarial networks.

320 We showed in [1] that this variational process is similar to a stochastic descent gradient algorithm performed iteratively. We here add a more sophisticated gradient descent algorithm corresponding to a simulated annealing mechanism in order to account for the neuromodulators involved in decision-making in the PFC for uncertainty and surprise [16, 95].

325 As proposed by [88], adding temporal structure to RL can ease the scaling problem in the exploration process. In the original version of Inferno, we found that STDP helped to learn and retrieve temporal chains. Thereinafter, we will show that the gain-modulation can even go further for abstracting temporal sequences, and can be more robust to variability.

330 3.2. Gain-modulation mechanism based on Rank-Order Coding

The rank-order coding (ROC) algorithm has been proposed by Thorpe and colleagues to model the information processing performed in the Visual Cortex by feedforward integrate-and-fire neurons [96]. We have expanded their use to recurrent neural network models in [97, 1] replicating the Spike Timing-
335 Dependent Plasticity learning mechanism.

The main assumption of the ROC algorithm is that spiking neurons perform a quantization of the variable inputs occurring in time discretized with respect to their temporal delays, see Fig. 3 b). The temporal order of the inputs are transcribed into a rank code that is translated into weight value and summed
340 at the neuron level. The more similar the temporal order of the incoming signals, the higher the amplitude level of the ROC neurons. Conversely, the less similar the sequence order of the incoming signals, the lower the amplitude level of the ROC neurons. Although this mechanism can encode discretized temporal sequences as shown in our previous work, it does not retranscribe a
345 gain-modulation mechanism, a sensitivity to a rank-order within a sequence *independently* of the neurons' index. We suggest here, as a novel coding strategy using gain-modulation, that we can construct "compressive codes" of tempo-

ral sequences by removing the index of the neurons (their identity) within the temporal sequence and keeping just their ranking order, see Fig. 3 c).

350 *STDP coding strategy* – If the ROC algorithm computes the neurons activity based on the discretized temporal delays of incoming events, the coding strategy resembles the STDP learning mechanism with ROC neurons becoming sensitive to the temporal contingency of incoming inputs; see Fig. 3 b). In this coding scheme, ROC units are thus considered to be contingency detectors inducing
355 phase synchronization [98].

Gain-modulation coding strategy – A second coding strategy consists in discretizing the serial order of units both in time *and* space, see Fig. 3 c). Here, the indices of the neurons (or their identities) are no longer preserved and it is their rank within the sequence that is taken into account; e.g. first, second or n-th
360 in the sequence. This strategy reduces drastically the amount of information to process, which makes possible the discovery of an abstract temporal structure disregarding the units indices; eg the sequence becomes a template. This coding mechanism is described as compressive representation by [6]. Hence, since the units index is no longer present in the temporal code, it is sensitive to any novel
365 sequences that preserve the global temporal structure.

For instance, in Fig. 3, the temporal encoding of two sequences following the same spatio-temporal pattern is constructed successively by first dismissing the temporal information and then the identity information with the rank-coding algorithm first on the time axis and then on the index axis.

370 The problem’s dimensionality for temporal sequences of M elements is reduced from a continuous time \times space dimension in \mathbb{R}^{2M} to an intermediate representation of $\mathbb{N}^M \times \mathbb{R}^M$ and then to a compressive representation of \mathbb{N}^{2M} . Although the reduction of complexity does not appear important when looking at the dimensionality of the vector quantization, it permits us to represent in
375 a compact way an infinity of varying spatio-temporal sequences that follow the same structure, which corresponds well to the variable binder property found in PFC neurons.

In comparison to other methods used to code the gain-modulation mech-

anism as in [6, 61, 99, 100], this one has the advantage of not projecting the
380 rank code and the position code into an intermediate 2D matrix of complexity
 $O(M^2)$. It does not need to also separate the ranking information into separate
modules or stripe codes as proposed in [19, 80].

This property of identity masking appears similar to the idea of variable
binding in [80], of timestamp neurons in [101, 60] or of neuronal pointers in [47]
385 as these neurons can fill out any new variables in the correct rank in the se-
quence. This coding strategy thus requires two types of units, one for maintain-
ing the input information (variable values) and one for recalling the sequential
order. Both are found in the prefrontal cortex for the maintenance of the units
activity and for the learning of a task-related activity. The ensemble constitutes
390 the behavior of one Working Memory.

One advantage of the gating strategy compared with the STDP one is that
the temporal information is learned separately from the inputs, which enables
the network to learn long-range dependencies at an abstract level and to prevent
it losing information more rapidly within a temporal horizon; this corresponds
395 to the so-called vanishing gradient effect in deep networks. As a remark, feed-
forward (deep) networks, standard recurrent neural networks (with/out STDP)
or hidden Markov models will easily lose accuracy after several iterations due
to the accumulated errors because any errors, noise or, delays within a sequence
and sensitivity to duration, will disrupt the sequence. One explanation for why
400 any error introduced into the network will make conventional neural networks
brittle is that the state and the temporal information are coded together. This
is not the case in neural models with a gating mechanism like PBWMs [80],
SPAWN [47] or LSTMs because the temporal information of a sequence can be
learned in memory cells separately from the variable values that can be retrieved
405 online or maintained dynamically over an indefinite amount of time.

The equations of the rank-order coding algorithm that we used are as follows.
The neurons' output Y is computed by forming the dot product between the
function $rank()$ sensitive to a specific rank ordering within the input signal
vector I and the synaptic weights w ; $w \in [0, 1]$. For a vector signal of dimension

410 M and for a population of N neurons (M afferent synapses), we have:

$$Y_n = \sum_m^M \text{rank}(I_m) w_{nm}, \forall n \in N \quad (1)$$

We implement the rank function $\text{rank}()$ as a power law of the `argsort()` function normalized between $[0, 1]$ for modeling the gain-modulation mechanism applied twice on the time axis and on the rank axis. This guarantees that the density distribution is limited and that the weight matrix is sparse, which makes
 415 the rank-order coding neurons similar to radial basis functions. This attribute permits us to use them as receptive fields so that the more distant the input signal is to the receptive field, the lower is its activity level. The updating rule of the weights is similar to the winner-takes-all strategy in Kohonen networks [102] with an adaptive learning rate $\alpha_n, \forall n \in N$. For the best neuron Y_b , we have:

$$\Delta w_{bm} = \alpha_b (\text{rank}(I_m) - w_{bm}), \forall m \in M \quad (2)$$

$$\alpha_b = 0.9 \alpha_b \quad (3)$$

420 3.3. Experimental Setup

We give the implementation details about the striatum-prefrontal working memory modeled by the Inferno Gate architecture. We ascribe to it the role of learning temporal patterns and representing audio memory sequences, see Fig. 2. The audio database used as input consists of a small audio dataset of
 425 2 minutes length with a native French woman speaker repeating five sentences three times. The audio .wav file is translated into MFCC vectors (dimension 12) sampled at 25ms each and tested with a stride of 10ms. The whole sequence represents 14.000 MFCC vectors, the number of units in the Striatum layer not encoded in the temporal order.

430 The Inferno Gate architecture is based on the same principle of the Inferno architecture –, the use of noise and reinforcement learning to control a spiking network,– except that the coding strategy now exploits compressive ranks (the

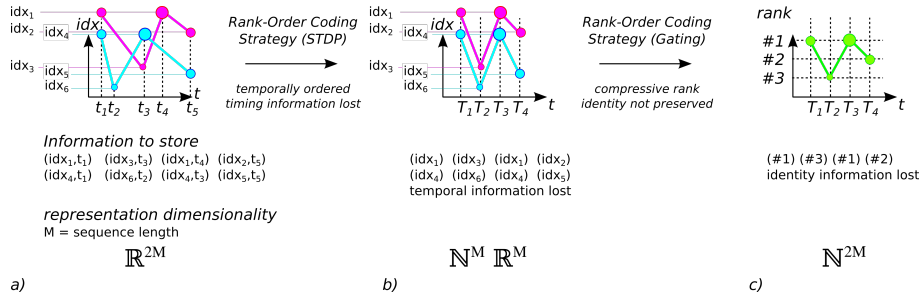


Figure 3: Rank-order algorithm for compressive rank representation. We describe the two-step process carried out with the rank-order coding algorithm to model the Spike Timing-dependent Plasticity rule and the Gating mechanism. In a), two sequences in cyan and magenta are represented with different neuron indices (idx), different timing but the same temporal structure (up-down-up-down). In b), the rank-order coding algorithm is used to quantify any sequence in the temporal domain with discrete timing; e.g. first ranked, second ranked. This is a rough approximation of the STDP rule. The indices of the neurons are kept and only the temporal information is lost. In c), we can for a second time use the rank-order coding algorithm to now suppress the neurons' indices (their identity) within the sequence in order to keep only their rank ($\#$) within the sequence. This second process makes possible a temporal pattern, a compressive representation of the two sequences in which only the rank order is kept. This second process reduces drastically the amount of information to encode any sequence, irrespective of the neurons' index and their precise timing. For any sequence of length M , the problem dimensionality is reduced to $\mathbb{R}^{2M} \rightarrow \mathbb{N}^{2M}$.

unit's index is not preserved) and temporal order (sensitivity to the position
in the sequence). Here, the B45 area is modeled as a buffer of 50 units length
435 receiving the indices ordered in time of the Striatum layer consisting of 14000
units; the number of coded MFCC in STR. Therefore, each chunk in B45 rep-
resents a sequence of 50 MFCCs, corresponding to a chunk of 1250 ms length.

Then, the IPFC layer encodes the ordinal information from the B45 buffer.
The IPFC layer consists of 5000 units for which each unit encodes a specific
440 temporal pattern through gain-modulation. Each IPFC unit learns the temporal
pattern that follows the serial order within the sequence of 50 units in B45,
independent of their true index.

At the next stage, the dIPFC layer combines together the IPFC units to rep-
resent abstract sequences. The dIPFC layer consists of 300 contextual neurons.

445 Each unit encodes a compositional representation from the 5000 basis functions
in the IPFC. The strong dimensionality reduction in the IPFC as explained in
section 3.2 and in Fig. 3 makes it possible to learn and rapidly generalize tempo-
ral patterns within sequences in the dlPFC and to rapidly explore new sequence
solutions at high speed.

450 Finally, the dlPFC units are evaluated by the ACC and a prediction error
signal is processed to search for and retrieve the optimal sequence in the OFC.
The OFC layer consists of 300 vectors of 50 iterations' length, one vector for each
unit in the dlPFC. Each vector is generated to retrieve back the corresponding
sequence of 50 iterations' length with the retrieved STR index values. The OFC
455 vectors are used for the executive control of the Striatal units.

4. Results

This section explains the two different experiments carried out to model the
striatum-prefrontal working memory for learning abstract temporal patterns
and for retrieving audio sequences through an exploration process.

460 In the first experiment presented in section 4.1, we explain the control carried
out from the striatum on the prefrontal layers to learn temporal patterns and to
represent abstract sequences with information flow $STR \rightarrow B45 \rightarrow IPFC \rightarrow dlPFC$.
In the second experiment presented in section 4.2, we explain the exploration
process carried out to generate audio memory sequences using evaluation of
465 prediction error from the encoded abstract sequences previously learned; the
information flow goes this time from $dlPFC \rightarrow OFC \rightarrow B45 \rightarrow IPFC \rightarrow dlPFC$.

In the third experiment presented in section 4.3, we explain the integration
of the two process into the complete Inferno Gate neural network for generating
structured sequences in a self-supervised manner. In the fourth experiment
470 presented in section 4.4, we compare the performances between Inferno Gate
and the standard LSTM recurrent neural network in classification and sequences
retrieving on our data.

4.1. Experiment 1 – Learning temporal patterns and encoding abstract sequences

4.1.1. Model-based gated control in striatum-prefrontal loops

475 Experiment 1 explains the information flow $\text{STR} \rightarrow \text{B45} \rightarrow \text{IPFC} \rightarrow \text{dlPFC}$ in order to learn the temporal patterns in the IPFC layer and the abstract sequences in the dlPFC layer respectively.

We present in Fig. 4 a) the dynamics of IPFC and dlPFC layers during the learning stage. The activity level of each neuron indicates their sensitivity to the temporal pattern found in the sequence. The neurons in the IPFC present
480 a compressive code of the B45 input sequences, which are then combined in the dlPFC layer at a more abstract level. In Fig. 4 b) in the top chart, we plot the snapshot of the IPFC population activity taken at one time step and sorted (black line) and in the bottom chart the temporal pattern of the most active
485 neuron in the IPFC (red line) and the rank order at the population level (blue line).

The activity level of the IPFC units indicate their saliency to one specific rank order within input sequences. The sorted activity in the black line indicates that the coding representation at the population level is not sparse but many
490 neurons are necessary to code the sequences. For instance, because each IPFC unit encodes one temporal pattern at the unit level, the decomposition of one sequence can only be partial as seen in Fig. 4 b) bottom chart, whereas at the population level, the sequence can be represented and discriminated.

4.1.2. Analysis of the sequence encoding in the IPFC network

495 After the learning stage of the IPFC and dlPFC units, we can analyze how the ranking information is encoded in the weight matrices. We plot in Fig. 5 a) the weights of the IPFC units (5000 units in the Y axis) reordered with respect to their sensitivity to specific positions within B45 sequences (50 iterations length in the X axis). The amplitude level of the synaptic weights in the diagonal
500 indicates that each IPFC neuron is sensitive to different positions within the sequences. For instance, neurons with high weight values in the beginning of the sequence will be more sensitive to forthcoming events occurring within the

sequence, and will serve as a prospective template. Conversely, neurons with high weight values in the ending of the sequence will be more sensitive to the past events in the sequence, they will serve as a retrospective template. This behavior has been observed in PFC neurons for sequence retrieval [103, 30] as well as in the prefrontal-hippocampal loop in T-Maze tasks [104, 105].

As an example, we plot in Fig. 5 b) the weight distribution of two IPFC units. We select them because these two units have their highest weight value for the position located at the middle of the sequence (position #24), and the other weights located at other positions within the sequence have a lower value. The two circles indicate two positions where the IPFC units have the same weights value: at positions #4 and #24. Based on these weight distributions, we can reconstruct back the temporal patterns for which the IPFC units are the most sensitive, as each weight's value corresponds to one rank within the sequence. We plot in Fig. 5 c) the two reconstructed temporal patterns. Since the two IPFC units have the same weights at the circle positions, the reconstructed sequences code the same neuron's rank at those positions within the sequence.

This is how in our framework the IPFC neurons retranscribe the gain-modulation mechanism: with respect to the position of one item within a sequence, the IPFC activity level will be modulated with respect to a weight's value depending where the corresponding item (the neuron's rank) within the sequence is located.

Depending on the activity level of the different IPFC units for a specific sequence, a decomposition in the IPFC space is represented at the dIPFC level, as in see Fig. 1 b). The decomposition in the IPFC layer permits us to represent at a more abstract level, and in a more compact fashion, the compressive rank of the sequence at the dIPFC layer.

4.2. dIPFC-OFC Iterative free-energy exploration-optimization

4.2.1. Retrieving memory sequences from incomplete information

We present in this section the iterative optimization process carried out at the dIPFC level for retrieving memory sequences at the OFC level with

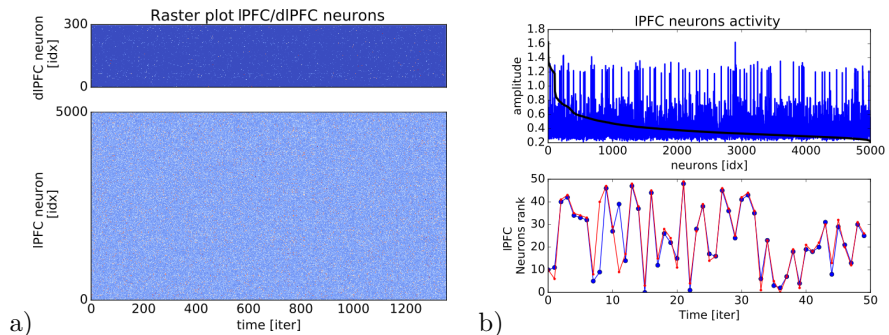


Figure 4: Encoding of temporal patterns in IPFC/dIPFC layers. In a), each IPFC unit encodes a temporal pattern from the incoming sequences in B45 (bottom chart, during 1400 iterations). The decomposition at the IPFC layer is a representation of the temporal structure in the sequences that is then learned at the dIPFC layer (top chart). In b) we give a snapshot of the IPFC layer at one particular timing (top chart). The black line indicates the sorted activity of the IPFC units. The red sequence in the bottom chart indicates the temporal pattern for which the most active IPFC neuron is the most sensitive; in this example, the neuron nearby the index number 2800. The blue sequence indicates the temporal pattern coded at the population level (for all IPFC neurons combined). This graph indicates that only one unit is not discriminative enough to represent any input sequence and that the coding at the population level is more robust and precise.

the error rate computed at the ACC. The information flow corresponds to $dIPFC \rightarrow OFC \rightarrow B45 \rightarrow IPFC \rightarrow dIPFC$, see Fig. 2.

535 In order to understand better the global process, we display in Fig. 6 a-c) and in d-f) the iterative optimization process carried out during 10.000 iterations for two dIPFC units, the dynamics of the IPFC layer and of the ACC unit are showed in a) and d), the final retrieved sequence in the OFC with respect to the one represented by the dIPFC units is displayed in b) and e), the raster plot of the iterative search of exact sequences in OFC and B45 is shown resp. c) and
 540 f).

In Fig. 6 a) and d) the ACC unit in the top chart represents the error rate, which is the inverse of the activity level of the dIPFC units. The raster plot of the IPFC dynamics is displayed in the bottom chart. The desired sequences we
 545 want to reconstruct in the OFC are presented at Fig 6 b) and e) in the top chart.

IPFC Analysis

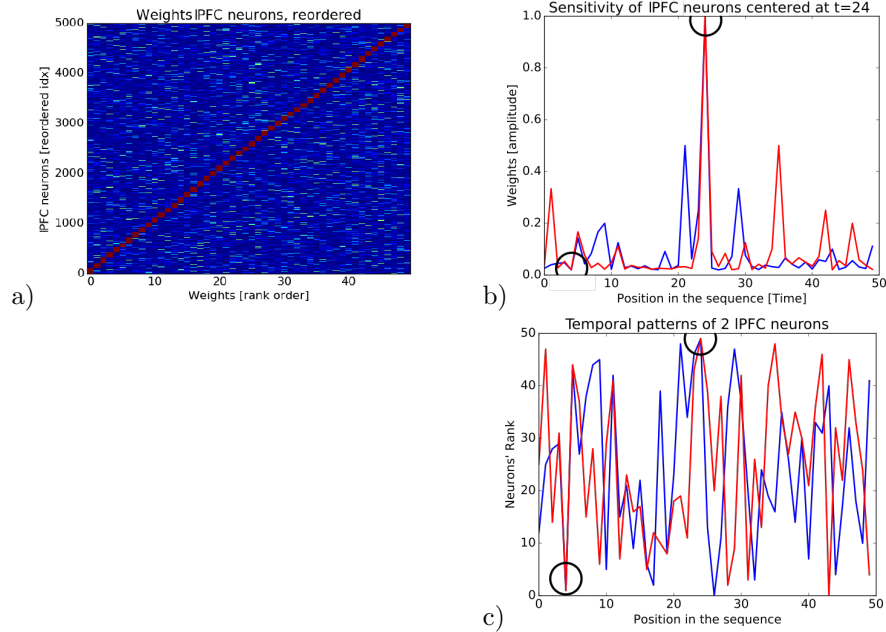


Figure 5: IPFC coding analysis. In a), raster plot of the weights of the IPFC units (Y axis) reordered with respect to their sensitivity to the location (e.g., rank order) within the sequence (X axis). Each neuron within the IPFC network has learned a weight distribution centered at one position within the sequence; e.g. the beginning, middle or end of the sequence. Neurons that code well the beginning of the sequence can help to predict the rest of the sequence (prospective memory). Neurons that code well the end of the sequence are salient to the elements during the whole sequence (retrospective memory). In b), weight distribution for two IPFC units centered at location #24 within the sequence is shown. The circles indicate the locations where the weight values overlap. In c), as the weight index indicates the location within the sequence and the weight value indicates the neuron's rank at that position, it is possible to reconstruct the temporal pattern for which the IPFC neurons are the most sensitive. The circles indicate weights with the same index and the same value encoding a temporal pattern with the same location and the same rank. The two temporal patterns cross at these locations.

The raster plot of the reconstructed OFC/B45 dynamics is shown in Fig 6 c) and f).

The exploration search is performed after the learning stage done in the previous section 4.2. Over time, a sequence in the OFC is explored iteratively

550 using noise so that the dlPFC activity is at maximal activity level and the ACC reaches a minimal value.

We display in Fig 6 b) and e), in red, the retrieved OFC sequence in the top chart and the serial order for which the two different dlPFC neurons are the most sensitive in the bottom chart.

555 In the top chart, we can observe that the reconstructed OFC sequences in red follow a similar pattern to the ones in blue although the index of the neurons is not completely preserved. Nonetheless, we can see that the ordinal information in the bottom chart is matched, which means that the proposed sequence in the top chart follows the temporal pattern encoded in the lPFC and in the dlPFC
560 layers.

Hence, despite the indices in the sequence being lost in the encoding process, the system is capable of retrieving the memory sequences from incomplete information (due to compressive rank) with small error.

4.2.2. Performance analysis from incomplete information

565 In order to analyze the accuracy of the Inferno Gate network, we plot in Fig. 7 the Euclidean error normalized between $[0, 1]$ made by the network during recall with respect to the number of items given as input vector resp. a) and with respect to the position within the sequence, resp. b). The exploration stage was limited to 10.000 iterations for each experiment and we plot the retrieved
570 sequences from 0 to 40 items given out of 50 items to retrieve, resp. in Fig. 8 a-e). The grey areas indicate the part of the sequence given to the system to reconstitute the missing part.

The fewer units there are to search, the more accurate is the recall, this can be seen if we provide 20% of the items in the sequence or 80%. In this later
575 case, the error on the items to retrieve is particularly small and almost error free. Moreover, the error made is not related to the temporal position of the items to retrieve but to the amount of missing information.

For 0% of information given, which corresponds to the previous situation in which the system has to retrieve the whole sequence from scratch, the root error

580 is 0.08 corresponding to 8% error with large variance, see Fig. 8 a).

For 20% of information given, which corresponds to ten items given out of fifty as displayed in Fig. 8 b), the accuracy is not necessarily higher, as we would expect, and the error rate reaches 0.07 with a higher variance in comparison with the previous case. We analyse this result as the difficulty the system has
585 in starting out from local minima with such small constraints added, which is different from the previous situation in which Inferno could freely search for solutions.

For 40% of information given, which corresponds to twenty items given out of fifty as displayed in Fig. 8 c), we can observe a strong decrease in both the
590 error rate of around 0.02 and the variance. This means that the network can retrieve 60% of the missing information with good accuracy.

Error rate continues to diminish below 0.01 if we provide 60% of the information (30 units), see Fig. 8 d), and serial recall is almost error-free if the network has to retrieve the index of ten units out of fifty (80% of the information given).
595 We can observe that the order position, to which the PFC neurons are sensitive, are all retrieved in the bottom chart although there is some slight errors in the reconstruction.

To resume, Fig. 8 shows that the more information is given to the system, the easier is the explorative search to retrieve the missing units index. Although
600 the rank order in the temporal patterns of the units in the IPFC is respected, this does not guarantee that the units index is retrieved correctly in the OFC sequences.

From additional studies that we did not present here, we observed that it was possible to complete the serial recall for all the cases with an error rate below
605 0.01 if we continue the exploration search for 50.000 to 100.000 iterations. These results indicate the generalization capabilities of Inferno separating linearly the input dynamics as we can achieve error-free retrieval.

4.3. Integrated Inferno Gate network, generating ongoing sequences autonomously

The two processes presented in sections 4.1 and 4.2 correspond to the two-
610 step mechanism of the complete Inferno Gate network for learning and gener-
ating ongoing sequences autonomously in a self-supervised manner. On the one
hand, the first circuit categorizes the incoming input sequences within a reper-
toire of learnt sequences in the dlPFC layer. On the other hand, the second
circuit generates novel sequences of items on the fly that follow the temporal
615 structure found by the first circuit using the reinforcement learning mechanism
in the ACC and the OFC layers.

We present in Fig. 9 a-b) the raster plots of the network dynamics at the
beginning and at the end of this error minimization process during ten thousand
iterations and in Fig. 9 c) the sorted error rate in the dlPFC layer at different
620 periods of time, which represents several million of iterations. At the beginning
of the self-organization process in Fig. 9 a), the error prediction in the ACC layer
takes long time-steps to reach small values and Inferno Gate converges slowly
both to learn the structure within the sequences and to generate new sequences
that follow these structures; the sorted error rate in ACC in Fig. 9 c) is high.
625 At the end of the self-organization process in Fig. 9 b), the error prediction
in the ACC layer takes shorter time-steps to switch to different sequences and
Inferno Gate generates coherent and structured sequences faster in a model-
based manner; the sorted error rate in ACC in Fig. 9 c) reaches a limit in the
optimization process.

630 Through time, Inferno Gate self-supervises its learning and the discovery of
novel sequences. This stage corresponds to the free-energy minimization of the
error prediction within the two sub-systems. The whole process may be thought
as a developmental stage in which the global network self-organizes its dynamics
following a model-based RL approach; achieving plasticity and stability within
635 the memory network in order to cumulate information while not diverging.

4.4. Performance comparison with LSTM networks

In order to assess the performances of the Inferno Gate network, we propose to compare the results in prediction with the state of art LSTM network used on our database. In comparison with standard time series, ours is difficult for deep
640 learning networks because there is no or few redundancy within the data: the instances in the dataset are seen few times and the dataset is large. We remind from section 3.3 that our time series are constituted of 1350 sequences of 50 iterations length each, for a corpus of 13.598 MFCCs. There is a fundamental trade-off in deep learning approach between the number of trials necessary to
645 learn one model and the number of samples to generalize correctly.

We present in table 4.4 the performances of the LSTM network in five experiments for various size of our time series, computing either the items index as input data or their rank-within the sequence. Experiment #1 indicates the performance of the LSTM network with 512 units to retrieve an input vector of
650 ten items length; the length of the time series is of 132 iterations, with the same number of units. We remark that for a prediction problem of small dimension with respect to the number of hidden units, the LSTM network learns perfectly the time series with 100% accuracy.

If we augment the size of the time series ten time over with a dataset of
655 1359 units (experiment #2) or one hundred time over with a dataset of 13598 units (experiment #3), then the network is not capable anymore to sustain the learning and the accuracy drops to 0. These results show that the LSTM network was overfitting its learning stage in experiment #1.

This contrasts with the results found with Inferno Gate in section 4.2 and
660 summarized in experiment #4 in table 4.4. In our model, the accuracy level varies from 92% to 99% depending on the number of items to retrieve and for a number of iterations fixed to ten thousand, precisely 13.598 items. Furthermore, because our network is an instance of a model-based RL algorithm, it required less units to encode sequences and less time also to retrieve missing elements
665 (results not shown here). In comparison, the number of parameters to optimize within the LSTM deep network makes the learning slow and proportional to the

size of the database.

Instead, if we give as input data the rank order of the items within the sequences to the LSTM network as in experiment #5, then the accuracy jumps
670 back to 99% accuracy as it is with Inferno Gate in experiment #4. In this experiment, the LSTM network is now sensitive to the data temporal structure but information about the content (item’s index) is lost.

For both Inferno Gate and LSTM networks in experiments #4 and #5, the encoding space using the rank-order information between $[0, 50]$ is now smaller
675 than the dimension of the input space between $[0, 13498]$, which makes the encoding simpler. However, because the LSTM network does not have any attention mechanism as Inferno Gate does, it cannot retrieve back the items index within the sequence; ie, the MFCC index. At reverse, Inferno Gate possesses such mechanism with the ACC and OFC components, which permits it
680 to recover back the hidden content.

5. Discussion

We have presented a novel neural architecture Inferno Gate based on free-energy minimization and model-based RL using recurrent spiking neural networks for learning and retrieving temporal sequences and modeling the fronto-striatal (PFC-BG) loop. This network extends our original neural architecture
685 Inferno in [1], which exploits model-free RL, aiming at modeling the cortico-basal ganglia (CX-BG) loop for learning motor primitives. In [2], we have shown its effectiveness in the more challenging tasks of speech recognition and production. Although the two networks are similar in their functional organization, the
690 encoding type is different. The Inferno network uses a model-free approach with the STDP mechanism for learning temporal correlations between spiking events whereas Inferno Gate uses a model-based approach with a gating mechanism for binding the item’s rank and its position within a sequence. This separation of functions between lexicon and grammar by the two networks follows the
695 declarative/procedural model on language proposed by Ullman [94, 106] with

case	time series' length (=nb of items)	network model	hidden neurons	input length	index loss / accuracy(%)	rank-order loss / accuracy(%)
#1	135 (w id)	LSTM	512	10	0.0 / 100	x
#2	1359 (w id)	LSTM	512	10	7.07 / 0.80	x
#3	13598 (w id)	LSTM	2500	10	37 / 0	x
#4	13598 (w id)	Inferno Gate	5000+300	50	10^{-8} - 10^{-4} / 92-99	10^{-8} - 10^{-4} / 92-99
#5	13598 (w/o id)	LSTM	2500	50	x	2.10^{-3} / 99

Table 1: Performances table between LSTM and Inferno Gate on sequence retrieving on a time series of rare and few redundant events. The first three experiments present the results of the LSTM network for various lengths of the our time series depending on its length for different size of the network. Because the time series possess rare and few redundant events within it, it is complex for the LSTM to generalize from it. Conversely, it is easier to learn the structure within it when the content is removed and replaced by the location of the items' indices within the sequence (their rank-order) as it is the case in experiment #4 for Inferno Gate and experiment #5 for LSTM. But only Inferno Gate can retrieve the items' indices because of the attentional mechanism it possesses.

similar memory circuits: the cortical layer for the lexicon memory system and the prefrontal-basal ganglia layer for the grammar memory system.

By discriminating content (i.e. which sound to play) and contextual information (i.e. when to play it in the sequence), we have shown that Inferno Gate
700 is capable of robustly learning the temporal structure within sequences and retrieving the items' index in the correct order. In a complementary paper, we have described the architecture Inferno for modeling the CX-BG structure for retrieving the sound primitives [2]. The BG network explored and retrieved sound vectors by testing them stochastically through the CX layer. The more
705 the Striatal units recognize and predict the CX output, the stronger they reinforce their link with the sound vectors encoded in the Globus Pallidus layer, which constitutes one sound repertoire at the end of the optimization process.

Although a stable activity can be retrieved in a self-organized manner within the CX-BG network, the top-down control of a precise temporal sequence is devoted to another structure, the PFC-BG loop, which selects and influences the
710 first system. In order to model the PFC-BG loop, we reuse the same neural architecture Inferno but with a different temporal coding to assess the property of the PFC neurons. In contrast to STDP, PFC neurons employ a gain-modulation mechanism to bind multiple information at once; e.g. the relative position of an
715 item within a sequence for instance.

Gain modulation – gain-modulated units learn the order and the rank of one item within a sequence so that any misplacement of it will reduce its activity level. Furthermore, the capacity to encode the items' localization follows an inverse power-law scale due to the rank-order coding algorithm: therefore the
720 precision of encoding is nonlinear. From a computational viewpoint, this inverse power-law scale is one important property in order to construct radial basis functions and to have orthogonal (discriminative) representations coded at the population level. Such coding is apparent to a nonlinear gating as the activity of these representations is modulated by the occurrence of multiple information.
725 In return, the population coding permits us to have a compressive code that can help the exploration search in a reinforcement learning framework.

This property permits to use the IPFC units as “fillers” or “pointers”, which adds some variability in the encoding, but in a structured way. Thus, we think that this property of gain-modulation is interesting in order to learn abstract and temporal information about structures within sequences such as AAB or ABA patterns to which infants are sensitive to at a very early age. Because the activity level of IPFC units in our system is modulated by multiple information, the item’s rank and time order, we suggest that it can be used for representing other relative metrics as proposed in [30, 42] and that gain-modulation can be assimilated to the nonlinear-mixed selectivity mechanism proposed by [107]. For instance, the conjunctive cells in the parietal cortex are found to encode relative metrics based on multimodal binding to infer non-trivial information about space and numerosity [108, 6, 109].

In previous neurorobotic researches, we have modeled these parietal gain-modulated neurons for visuomotor coordination and for body representation using a more standard coding strategy based on multiplication [110, 61, 99, 100, 111]. However, the number of units necessary to process gain-modulation evolves quadratically with respect to the problem’s dimensionality. We think spiking neural networks along with the gain-modulation learning mechanism have the potential to represent multimodal information in a more compact manner, perhaps even more efficiently than do conventional multi-layer feedforward networks, as there is no loss of structural information in the encoding, which is not the case in deep networks due to the max-pooling operation. Therefore, we can envisage some tasks in multimodal integration, which are still difficult to realize with spiking neural networks.

Retrospective and prospective encoding – The temporal coding carried out in Inferno Gate extends the STDP mechanism with extra information, the position of an item within a sequence, making it more nonlinear and abstract in the sense that the neurons’ receptive field encodes structural information about the sequence and not the sequence itself. In our experiments, IPFC neurons code for a position in the sequence either at the beginning, middle or end of it. They are nevertheless sensitive to other positions but with less strength.

As some IPFC neurons are sensitive to items at the beginning of a sequence or at the end, this behavior reflects well the PFC’s behavior, firing in line with retrospective or anticipatory events. Retrospective neurons are firing depending on the previous events. Conversely, prospective neuronal firing depends on future events. These properties were showed in the hippocampal cells of rats in W-shaped tracks with alternate trajectories [105] or in a T-Maze with return arms [104]. These results and many others mainly refer to hippocampal activities. However, it has been shown that these activities may be modulated by prefrontal information [112, 103]. We here make the prediction that some IPFC neurons exhibit these kinds of activities, and that this information is used as a global context for driving neuronal firing in the hippocampus [113]; e.g., during the replay of memory sequences [114] and during the pre-play of trajectories toward future events [115].

Compositionality– the dlPFC units combine the IPFC temporal patterns in a similar way to a radial basis function network: as this layer embeds a variety of temporal primitives, it can rapidly encode any novel sequences. Hence, we suggest that this mechanism of gain-modulation is potentially important in infants for fast inference and for learning abstract patterns with few samples; the capability of learning-to-learn [52, 12]. The learning and error minimization processing of IPFC/dlPFC temporal rules may be viewed as an inference mechanism, which attempts to catch up with rules in one domain. As these temporal rules are abstracted (no information about the content), they may be applied to other domains for grasping other observations; for instance, learning the temporal patterns during motion sequences, like planning, drawing or solving a task. As a beneficial effect, learning higher-order hypotheses may massively accelerate learning in other domains [12, 81]. It may help to learn abstract or contextual words, which is still a challenge that AI architectures have yet to meet.

To summarize, we suggest that this system presents some capabilities suited for learning linguistic systems (e.g., a grammar of rules) and timely ordered behaviors. Since Inferno Gate encodes temporal patterns in an abstract manner, like AAB or ABA patterns, we can expect that by adding another layer to the

model, presumably the Polar Frontal Cortex as proposed in [26], it may be possible to create sequences of sequences such as ((AAB)B(AAB)) with A=(AAB) and the ABA pattern, mixing two or more temporal patterns in an iterative fashion. In this way, our network may be extended to fractal-coding to have a hierarchical representation of sequences at any depth.

Long-term dependencies – In experiment 2 in section 4.2, we have shown that it was possible to retrieve accurately in the OFC layer the index of units and their order in long-range sequences (fifty iterations length) although this information (the neuron’s index) was not encoded in the dlPFC units. With no external information, the system requires a long period of time, above ten thousand iterations, to search for the items’ index as well as their position. All generated sentences display the same temporal pattern as the one we want to retrieve, although there is some variability present in them.

This exploration process may be seen as a babbling period in infant development. The explorative search is based on a free-energy minimization process of the OFC vectors and on the evaluation of the dlPFC units computed at the ACC level. We have showed that error minimization is fast and that it requires only several hundred iterations to retrieve the missing items within the sequence. If items are furnished and imposed on the sequence to be retrieved (a sequence ‘à trous’ [with holes]), the search is even faster and accurate with respect to the number of items given, below two hundred iterations.

The generative property of the system shows that it incorporates computational capabilities of robustness to noise and can retain long-term dependencies for sequences of fifty iterations length as no or little information is required to retrieve which item has to be performed and when. This property is advantageous in comparison to standard recurrent neural networks. For instance, Inferno Gate behaves differently from classical recurrent networks, including the gated ones, which are directed graphes that attempt to predict the next items depending on the past ones in a Markovian fashion. Our network may extend the idea of vector codes or of vector symbolic architectures by [116] or of neural pointers by [47, 46] or of the “merge” function proposed by Chomsky for en-

820 coding nested structures in language [53, 25]. Furthermore, we have tested the
performances and computational power of the Inferno Gate network in compar-
ison to the popular LSTM gated networks. We have shown that only Inferno
Gate could retrieve back items’ in sequences, even from a dataset of rarely ob-
served items. This is an advantage over statistical methods and deep networks
825 that require many observations to learn correlations between the items.

Developmental learning – Infants skill at learning event sequences ex-
tends to different stimulus types in different modalities speech, tones, or pic-
tures [117]. According to Saffran, young children are equipped with a statistical
learning device for syntactical acquisition [13, 14]. In a series of experiments,
830 Saffran and colleagues observed that after exposure to the speech streams for
two minutes, infants reacted differently to hearing pseudowords as opposed to
nonwords from the speech stream, where nonwords were composed of the same
syllables that the infants had been exposed to, but in a different *order*. They
suggest that infants are able to learn statistical relationships between syllables
835 even with very limited exposure to a language.

Baillargeon and colleagues discovered similar behaviors in the visual do-
main [118, 119] and proposed also that infants may possess a Physical Reasoning
System to endow them with a grasp of intuitive physics about objects. Import-
antly, they discovered that young infants reason at an abstract level about
840 objects categories (inanimated, animated), physics (occlusion, shadows, rigid
3D objects dont distort or disappear) and properties (soft, round) to simulate
what will happen next. Similar to grammar in language, this Physical Reason-
ing System may allow infants to reason about causal and physical events and
to detect several physical violations occurring in the visual scene. For instance,
845 infants were surprised when an object placed behind a screen disappeared and
a different object reappeared after. Young infants possess therefore some ex-
pectations about physical knowledge, such as object permanence and object
occlusion.

Although they don’t identify any neural correlates of these processings, we
850 suggest that the PFC is involved in extracting the structure within sound and

visual sequences.

We propose also that our model may provide a computational basis to these results. Because the content (the items' index) has been removed from the temporal information within the sequence, Inferno Gate can learn temporal
855 patterns. The learning and error minimization processing of these rules may be similar to an inference process that attempts to capture the rules in one domain. As these temporal rules are abstracted, amodal, they may be applied to other domains as well or to other modalities; for instance, during action observation for grasping other observations.

860 In this line, brain theorists such as Arbib and Fadiga suggested that the Broca area in PFC plays a more general role, beyond language, as being a supra-modal Syntax Engine in the broaden sense, to encode hierarchical structures and grammars in other core domains and modalities [120, 121, 122, 123]; for instance, in music and in action representation. They propose that motor articulation in
865 speech is deeply rooten to the syntactic and structural organization of actions. They suggest that the neural pathways for syntactic representation in speech and action are located in the sample places, respectively in the Broca area in human and in the Mirror Neurons System (F5 motor area) in monkeys, the two presenting potentially similar functional organization and similar mechanisms
870 for the acquisition of conceptual knowledge and hierarchical representation [124].

We propose that our architecture may account for the modeling of syntactic representation found in those structures.

Transfer of learning – As the content of the data is not encoded within the PFC neurons, we can expect that they are more robust to variability within the
875 inputs such as distorted voices or voices with different tones (high and low pitch). In this line, we can expect Inferno Gate to find the same temporal patterns between different modalities, visual and auditory, for instance as during lip-reading accentuating the pronounced sound with the visual input. This gating mechanism may make it possible to express another way of solving the binding
880 problem across modalities and to perform transfer learning. For instance, some experiments performed with babies have found how at an early age they bind

the tactile texture of a protruder and its visual shape although they have not seen it before [125]. Other experiments have shown such a binding effect being created between sound and vision in which babies correlated an ascending sound with a more intense light following the same temporal pattern.

Advantages and disadvantages of the proposed architecture – In comparison to model-free RL and deep neural-based approaches that require some redundancy within the signals for learning correlations, the learning and discovery stage of a structure with model-based RL such as Inferno Gate is extremely fast with few examples seen in one or few-shot learning [75, 74]. On the one hand, deep networks require thousands of iterations during the offline learning stage to converge, they classify almost instantaneously any data afterwards. On the other hand, neural networks with an attention mechanism and/or a working memory like Inferno Gate learn a structure extremely fast with few examples only [82, 84], while they recognize and retrieve online one given sequence in few hundred iterations using a read/write address memory.

As a result, model-free algorithms like Inferno have difficulty to generalize and infer from unseen data or rare events. The compositionality of Inferno Gate instead permits to accelerate the learning stage and to discover a sequence even with elements not seen before. Therefore, in our case the dimensionality of the input space of 14.000 MFCC index is not seen as a major problem in the retrieval stage. For instance, we have shown that it was possible to retrieve sentences of 50 iterations length and to create new ones on the fly. Furthermore, there is not a great sensitivity of the sequences' length on the retrieval stage as long as a certain number of entries are given for discriminating the sentences from each other. Therefore, there is no vanishing gradient effect as seen in other deep or gated networks such as the LSTMs. This is because the index of the input (their identity) is not learned.

Our next steps will be to combine the networks Inferno and Inferno Gate in order to obtain a meta-learning architecture [74]. The dual working-memory will integrate both model-free RL and model-based RL to design the interaction between the cortico-basal loop and the striatal-prefrontal loop [74, 95, 75] under

frameworks of free-energy, predictive coding [3, 126, 127] and intrinsic motivation [128, 129]. This will follow the declarative/procedural model proposed by
915 Ullman for language [106]. We will study how our model will achieve structural learning in core domains with robots.

Acknowledgements

We would like to thank Mathieu Lagrange (LS2N, UMR 6004 CNS, Ecole Centrale de Nantes) for providing the audio database. This work was partially
920 funded by EQUIPEX-ROBOTEX (CNRS), chaire d'excellence CNRS-UCP and project Labex MME-DII (ANR11-LBX-0023-01).

References

- [1] A. Pitti, P. Gaussier, M. Quoy, Iterative free-energy optimization for recurrent neural networks (inferno), PLoS ONE 12 (3) (2017) e0173684.
- 925 [2] A. Pitti, , M. Quoy, C. Lavandier, S. Boucenna, Feature and structural learning of memory sequences with recurrent and gated spiking neural networks using free-energy: application to speech perception and production i, under review.
- [3] K. Friston, Learning and inference in the brain, Neural Networks 16 (9)
930 (2003) 13251352.
- [4] K. Friston, J. Kilner, L. Harrison, A free energy principle for the brain, Journal of Physiology-Paris 100 (1-3) (2006) 7087.
- [5] G. Keller, T. Mrsic-Flogel, Predictive processing: A canonical cortical computation, Neuron 100 (2) (2018) 424435.
- 935 [6] M. Botvinick, T. Watanabe, From numerosity to ordinal rank a gain-field model of serial order representation in cortical working memory, The Journal of Neuroscience 27 (32) (2007) 8636–8642.

- [7] M. Hasselmo, C. Stern, A network model of behavioural performance in a rule learning task, *Phil. Trans. R. Soc. B* 373 (2018) 20170275.
- 940 [8] N. Kasabov, Time-space, spiking neural networks and brain-inspired artificial intelligence, Heidelberg: Springer.
- [9] E. Spelke, What makes us smart? core knowledge and natural language, In: *Language in mind: Advances in the Investigation of language and thought*, ed. D. Gentner & S. Goldin-Meadow. MIT Press. (2003) 277311.
- 945 [10] E. Spelke, K. Kinzler, Core knowledge, *Developmental Science* 10 (1) (2007) 8996.
- [11] A. Gopnik, A. Meltzoff, P. Kuhl, *The scientist in the crib what early learning tells us about the mind*, Publisher William Morrow Paperbacks.
- [12] J. Tenenbaum, C. Kemp, T. Griffiths, N. Goodman, How to grow a mind
950 statistics, structure, and abstraction, *Science* 331 (6022) (2011) 1279–1285.
- [13] J. Saffran, R. Aslin, E. Newport, Statistical learning by 8-month-old infants, *Science* 274 (1996) 19261928.
- [14] J. Saffran, D. Wilson, From syllables to syntax: Multilevel statistical
955 learning by 12-month-old infants, *Infancy* 4 (2003) 273284.
- [15] A. Basirat, S. Dehaene, G. Dehaene-Lambertz, A hierarchy of cortical responses to sequence violations in three-month-old infants, *Cognition* 132 (2014) 137150.
- [16] A. Yu, P. Dayan, Uncertainty, neuromodulation, and attention, *Neuron*
960 46 (2005) 681–692.
- [17] N. Daw, Y. Niv, P. Dayan, Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control, *Nat. Neurosci.* 8 (2005) 1704–1711.

- [18] N. Rougier, R. O'Reilly, Learning representations in a gated prefrontal cortex model of dynamic task switching, *Cognitive Science* 26 (2002) 503520.
965
- [19] F. M. O'Reilly, RC, Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia, *Neural Comput* 18 (2) (2006) 283328.
- [20] R. Romo, C. Brody, A. Hernandez, L. Lemus, Neuronal correlates of parametric working memory in the prefrontal cortex, *Nature* 399 (6735) (2018) 470.
970
- [21] J. Tanji, E. Hoshi, Behavioral planning in the prefrontal cortex, *Curr. Opin. Neurobiol.* 11 (2001) 164170.
- [22] J. X. Wang, Z. Kurth-Nelson, D. Kumaran, D. Tirumala, H. Soyer, J. Z. Leibo, D. Hassabis, M. Botvinick, Prefrontal cortex as a meta-reinforcement learning system, *Nature Neuroscience* 21 (2018) 860868.
975
- [23] E. Koechlin, An evolutionary computational theory of prefrontal executive function in decision-making, *Phil. Trans. R. Soc. B* 369 (2014) 20130474.
- [24] E. Koechlin, Prefrontal executive function and adaptive behavior in complex environments, *Current Opinion in Neurobiology* 37 (2016) 1–6.
980
- [25] S. Dehaene, F. Meyniel, C. Wacogne, L. Wang, C. Pallier, The neural representation of sequences from transition probabilities to algebraic patterns and linguistic trees, *Neuron* 88 (2015) 2–19.
- [26] M. Rouault, E. Koechlin, Prefrontal function and cognitive control: from action to language, *Current Opinion in Behavioral Sciences* 21 (2018) 106111.
985
- [27] B. Averbeck, M. Chafee, D. Crowe, G. A.P., Neural activity in prefrontal cortex during copying geometrical shapes. i. single cells encode shape, sequence, and metric parameters, *Exp Brain Res.* 150 (2) (2003) 127–41.

- 990 [28] B. Averbeck, D. Crowe, M. Chafee, G. A.P., Neural activity in prefrontal cortex during copying geometrical shapes. ii. decoding shape segments from neural ensembles, *Exp Brain Res.* 150 (2) (2003) 143153.
- [29] K. Shima, M. Isoda, H. Mushiake, J. Jun Tanji, Categorization of behavioural sequences in the prefrontal cortex, *Nature* 445 (2007) 315–318.
- 995 [30] J. Tanji, K. Shima, H. Mushiake, Concept-based behavioral planning and the lateral prefrontal cortex, *Trends in Cognitive Sciences* 11 (12) (2007) 528–534.
- [31] P. Barone, J. J.P., Prefrontal cortex and spatial sequencing in macaque monkey, *Exp Brain Res* 78 (2018) 447 464.
- 1000 [32] Y. Ninokura, H. Mushiake, J. Tanji, Integration of temporal order and object information in the monkey lateral prefrontal cortex, *J Neurophysiol* 91 (2004) 555–560.
- [33] M. Inoue, A. Mikami, Prefrontal activity during serial probe reproduction task: encoding, mnemonic and retrieval processes, *J Neurophysiol* 95 (2018) 1008 1041.
- 1005 [34] R. Andersen, V. Mountcastle, The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex, *J. Neuroscience* 3 (1983) 532–548.
- [35] R. Andersen, G. Essick, R. Siegel, Encoding of spatial location by posterior parietal neurons, *Science* 230 (1985) 450–458.
- 1010 [36] R. Andersen, Multimodal integration for the representation of space in the posterior parietal cortex, *Philos Trans R Soc Lond B Biol Sci* 353 (1997) 1421–1428.
- [37] A. Pouget, L. Snyder, Computational approaches to sensorimotor transformations, *Nature Neuroscience* 3 (2000) 1192–1198.
- 1015

- [38] E. Hubbard, M. Piazza, P. Pinel, S. Dehaene, Interactions between number and space in parietal cortex, *Nat. Rev. Neurosci.* 6 (2005) 435438.
- [39] A. Pouget, L. Snyder, Spatial transformations in the parietal cortex using basis functions, *J. of Cog. Neuro.* 3 (1997) 1192–1198.
- 1020 [40] A. Georgopoulos, H. Merchant, T. Naselaris, B. Amirkian, Mapping of the preferred direction in the motor cortex, *Proc Natl Acad Sci USA.* 104 (26) (2007) 11068–72.
- [41] S. Kakei, D. Hoffman, P. Strick, Sensorimotor transformations in cortical motor areas, *Neuroscience Research* 46 (2003) 110.
- 1025 [42] A. Genovesio, S. Wise, R. Passingham, Prefrontalparietal function: from foraging to foresight, *Trends in Cognitive Sciences* 18 (2) (2014) 7281.
- [43] A. Genovesio, S. Tsujimoto, S. Wise, Feature- and order-based timing representations in the frontal cortex, *Neuron* 63 (2009) 254266.
- [44] J. Fuster, The prefrontal cortexan update time is of the essence, *Neuron* 30 (2001) 319–333.
- 1030 [45] J. Paton, D. Buonomano, The neural basis of timing: Distributed mechanisms for diverse functions, *Neuron* 98 (2) (2018) 687705.
- [46] A. Zylberberg, S. Dehaene, P. Roelfsema, M. Sigman, The human turing machine a neural framework for mental programs, *Trends in Cognitive Science* 15 (7) (2011) 293–300.
- 1035 [47] C. Eliasmith, T. Stewart, X. Choo, T. Bekolay, T. DeWolf, Y. Tang, D. Rasmussen, A large-scale model of the functioning brain, *Science* 338 (6111) (2012) 12021205.
- [48] A. Zylberberg, L. Paz, P. Roelfsema, S. Dehaene, M. Sigman, A neuronal device for the control of multi-step computations, *Papers in Physics* 5 (2013) 050006.
- 1040

- [49] C. Wacongne, E. Labyt, V. van Wassenhove, T. Bekinschtein, L. Naccache, S. Dehaene, Evidence for a hierarchy of predictions and prediction errors in human cortex, *Proc. Natl. Acad. Sci. USA* 108 (2011) 2075420759.
- 1045 [50] N. Chomsky, Three models for the description of language, *IEEE Trans. Inf. Theory* 2 (1956) 113124.
- [51] Y. Sugita, T. Tani, Learning semantic combinatoriality from the interaction between linguistic and behavioral processes, *Adapt. Behav.* 13 (1) (2005) 33–52.
- 1050 [52] T. Griffiths, N. Chater, C. Kemp, A. Perfors, J. Tenenbaum, Probabilistic models of cognition exploring representations and inductive biases, *Trends in Cognitive Sciences* 14 (8) (2010) 357–364.
- [53] N. Chomsky, Problems of projection, *Lingua* 130 (2013) 3349.
- [54] T. Tani, Self-organization and compositionality in cognitive brains: A
1055 neurorobotics study, *Proceedings of the IEEE* 102 (4) (2014) 586–605.
- [55] B. Lake, C. Lee, J. Tenenbaum, One-shot learning of generative speech concepts, in: *Proceedings of the 36th Annual Conference of the Cognitive Science Society*, 2014.
- [56] P. Dominey, Complex sensory-motor sequence learning based on recurrent
1060 state representation and reinforcement learning, *Biol. Cybern.* 73 (1995) 265–274.
- [57] P. Dominey, M. Arbib, J. Joseph, A model of corticostriatal plasticity for learning oculomotor associations and sequences, *J. Cogn. Neurosci.* 7 (25).
- [58] P. Dominey, Recurrent temporal networks and language acquisition from
1065 corticostriatal neurophysiology to reservoir computing, *Front. Psy.* 4 (500) (2013) 1–14.
- [59] M. Botvinick, D. Plaut, Short-term memory for serial order: A recurrent neural network model, *Psychological Review* 113 (2) (2006) 201233.

- 1070 [60] C. Wacongne, J. Changeux, , S. Dehaene, A neuronal model of predictive coding accounting for the mismatch negativity, *J. Neurosci.* 32 (2012) 36653678.
- [61] A. Pitti, R. Braud, S. Mahé, M. Quoy, P. Gaussier, Neural model for learning-to-learn of novel task sets in the motor domain, *Frontiers in Psychology* 4 (771).
- 1075 [62] B. J. Knowlton, R. G. Morrison, J. E. Hummel, K. J. Holyoak, A neurocomputational system for relational reasoning, *Trends in Cognitive Sciences* 16 (M) (2012) 373–381. doi:10.1016/j.tics.2012.06.002.
URL <http://dx.doi.org/10.1016/j.tics.2012.06.002>
- [63] R. Laje, D. Buonomano, Robust timing and motor patterns by taming chaos in recurrent neural networks, *Nature Neuroscience* 16 (7) (2013) 925–935.
- 1080 [64] D. Sussillo, Neural circuits as computational dynamical systems, *Current opinion in neurobiology* 25 (2014) 156163.
- [65] I. Tsuda, Chaotic itinerancy and its roles in cognitive neurodynamics, *Current Opinion in Neurobiology* 31 (2015) 6771.
- 1085 [66] E. Dauté, M. Quoy, B. Cessac, B. Doyon, M. Samuelides, Self-organization and pattern-induced reduction of dynamics in recurrent networks, *Neural Networks* 11 (1998) 521–533.
URL <https://hal.archives-ouvertes.fr/hal-01416782>
- 1090 [67] E. Dauté, M. Quoy, B. Doyon, Resonant spatio-temporal learning in large random neural networks, *Biological Cybernetics* 87 (2002) 185–198.
- [68] F. Mannella, M. Mirolli, G. Baldassare, Goal-directed behavior and instrumental devaluation: A neural system-level computational model, *Front. Behav. Neurosci.* 10 (181) (2016) 575–595.

- 1095 [69] X. Hinaut, M. Petit, G. Pointeau, P. Dominey, Real-time parallel processing of grammatical structure in the fronto-striatal system a recurrent network simulation study using reservoir computing, *PLoS ONE* 8 (2) (2013) e52946.
- [70] X. Hinaut, M. Petit, G. Pointeau, P. Dominey, Exploring the acquisition and production of grammatical constructions through human-robot
1100 interaction with echo state networks, *Front. Neurobot.* 8 (16).
- [71] F. Mannella, G. Baldassare, Selection of cortical dynamics for motor behaviour by the basal ganglia, *Biological Cybernetics* 109 (2015) 575–595.
- [72] Y. Yamashita, T. Tani, Emergence of functional hierarchy in a multiple
1105 timescale neural network model: a humanoid robot experiment, *PLoS Comput. Biol.* 4 (11) (2008) e1000220.
- [73] A. Ahmadi, J. Tani, How can a recurrent neurodynamic predictive coding model cope with fluctuation in temporal patterns? robotic experiments on imitative interaction, *Neural Networks* 92 (2017) 3–16.
- 1110 [74] M. Botvinick, S. Ritter, J. X. Wang, Z. Kurth-Nelson, C. Blundell, D. Hassabis, Reinforcement learning, fast and slow, *Trends in Cognitive Sciences* 23 (5) (2019) 408 – 422. doi:<https://doi.org/10.1016/j.tics.2019.02.006>.
URL <http://www.sciencedirect.com/science/article/pii/S1364661319300610>
1115 S1364661319300610
- [75] E. O. Neftci, B. B. Averbeck, Reinforcement learning in artificial and biological systems, *Nature Machine Intelligence* 1 (2019) 133143.
- [76] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural Comput.* 9 (1997) 17351780.
- 1120 [77] F. Gers, J. Schmidhuber, F. Cummins, Learning to forget: Continual prediction with lstm, *Neural Comput.* 12 (10) (2000) 24512471.

- [78] M. Frank, B. Loughry, R. O'Reilly, Interactions between the frontal cortex and basal ganglia in working memory: a computational model, *Cogn Affect Behav Neurosci* 1 (2001) 137–160.
- 1125 [79] C. Eliasmith, *How to build a brain A neural architecture for biological cognition*, Oxford University Press, New York, NY, 2013.
- [80] T. Kriete, D. C. Noelle, J. D. Cohen, R. C. O'Reilly, Indirection and symbol-like processing in the prefrontal cortex and basal ganglia, *Proc Natl Acad Sci U S A.* 110 (41) (2013) 16390–5.
- 1130 [81] B. M. Lake, T. D. Ullman, J. B. Tenenbaum, S. J. Gershman, Building machines that learn and think like people, *Behavioral and Brain Sciences* 40 (2017) e253. doi:10.1017/S0140525X16001837.
- [82] A. Graves, G. Wayne, I. Danihelka, *Neural Turing Machines*, arXiv 1410.541v2 (2014) 1–26.
- 1135 [83] A. e. a. Graves, Hybrid Computing Using a Neural Network with Dynamic External Memory, *Nature* 538 (2016) 471476.
- [84] A. Santoro, S. Bartunov, M. Botvinick, D. Wierstra, T. P. Lillicrap, One-shot learning with memory-augmented neural networks, *CoRR* abs/1605.06065. arXiv:1605.06065.
- 1140 URL <http://arxiv.org/abs/1605.06065>
- [85] T. Rocktschel, E. Grefenstette, K. M. Hermann, T. Koisk, P. Blunsom, Reasoning about entailment with neural attention, *ICLR* abs/1509.06664. URL <https://arxiv.org/abs/1509.06664>
- [86] M. V. Butz, D. Bilkey, D. Humaidan, A. Knott, S. Otte, Learning, planning, and control in a monolithic neural event inference architecture, *CoRR* abs/1809.07412. arXiv:1809.07412.
- 1145 URL <http://arxiv.org/abs/1809.07412>

- 1150 [87] K. Gregor, I. Danihelka, A. Graves, D. Rezende, D. Wierstra, Draw: A recurrent neural network for image generation, International Conference on Machine Learning (2015) 14621471.
- [88] M. Botvinick, Y. Niv, A. Barto, Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective, *Cognition* 113 (3) (2009) 262280.
- 1155 [89] E. Koechlin, C. Ody, F. Kouneilher, The architecture of cognitive control in the human prefrontal cortex, *Science* 302 (14) (2003) 1181–1185.
- [90] T. Buschman, E. Miller, Goal-direction and top-down control, *Phil. Trans. R. Soc. B* 369 (2014) 20130471.
- [91] E. Miller, The "working" of working memory, *Dialogues Clin Neurosci.* 15 (4) (2015) 411–418.
- 1160 [92] W. Ma, J. Beck, P. Latham, A. Pouget, Bayesian inference with probabilistic population codes, *Nat Neurosci* 9 (11) (2006) 14321438.
- [93] M. Botvinick, A. Weinstein, Model-based hierarchical reinforcement learning and human action control, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369 (2014) 20130480.
- 1165 [94] M. T. Ullman, A neurocognitive perspective on language: the declarative/procedural model, *Nat. Rev. Neurosci.* 2 (2001) 717726.
- [95] A. Cortese, B. D. Martino, M. Kawato, The neural and cognitive architecture for learning from a small sample, *Current Opinion in Neurobiology* 55 (2019) 133 – 141, machine Learning, Big Data, and Neuroscience. doi:<https://doi.org/10.1016/j.conb.2019.02.011>.
1170 URL <http://www.sciencedirect.com/science/article/pii/S0959438818301077>
- [96] S. Thorpe, A. Delorme, R. Van Rullen, Spike-based strategies for rapid processing, *Neural Networks* 14 (2001) 715–725.

- 1175 [97] A. Pitti, R. Braud, S. Mahé, M. Quoy, P. Gaussier, Neural model for learning-to-learn of novel task sets in the motor domain, *Frontiers in Psychology* 4 (771).
- [98] E. Izhikevich, Polychronization computation with spikes, *Neural Computation* 18 (2006) 245–282.
- 1180 [99] S. Mahé, P. Braud, R. Gaussier, M. Quoy, A. Pitti, Exploiting the gain-modulation mechanism in parieto-motor neurons application to visuomotor transformations and embodied simulation, *Neural Networks* 62 (2015) 102–111.
- [100] J. Abrossimoff, A. Pitti, P. Gaussier, Visual learning for reaching and
1185 body-schema with gain-field networks, *Joint IEEE International Conferences on Development and Learning and Epigenetic Robotics (ICDL-Epirob)* (2018) 191–197.
- [101] D. Jin, N. Fujii, A. Graybiel, Neural representation of time in cortico-basal ganglia circuits, *Proc. Natl. Acad. Sci. USA* 106 (2009) 1915619161.
- 1190 [102] T. Kohonen, Self-organized formation of topologically correct feature maps, *Biological Cybernetics* 43 (1982) 59–69.
- [103] S. Funahashi, Prefrontal cortex and working memory processes, *Neuroscience* 139 (1) (2006) 251–261.
- [104] E. Wood, P. Dudchenko, R. J. Robitsek, H. Eichenbaum, Hippocampal
1195 neurons encode information about different types of memory episodes occurring in the same location, *Neuron* 27 (3) (2000) 623–633.
- [105] L. Frank, B. E.N., M. Wilson, Trajectory encoding in the hippocampus and entorhinal cortex, *Neuron* 27 (1) (2000) 169 – 178.
- [106] M. T. Ullman, Contributions of memory circuits to language: the declarative/procedural model, *Cognition* 92 (2004) 231270.
1200

- [107] M. Rigotti, O. Barak, M. Warden, X. Wang, N. Daw, E. Miller, S. Fusi, The importance of mixed selectivity in complex cognitive tasks, *Nature* 497 (7451) (2013) 585–590.
- [108] E. Salinas, P. Thier, Gain modulation a major computational principle of the central nervous system, *Neuron* 27 (2000) 15–21.
1205
- [109] G. Blohm, J. Crawford, Fields of gain in the brain, *Neuron* 64 (2009) 598–600.
- [110] A. Pitti, A. Blanchard, M. Cardinaux, P. Gaussier, Gain-field modulation mechanism in multimodal networks for spatial perception, 12th IEEE-RAS International Conference on Humanoid Robots Nov.29-Dec.1, 2012. Business Innovation Center Osaka, Japan (2012) 297–302.
1210
- [111] Brain-inspired coding of robot body schema through visuo-motor integration of touched events, *Frontiers in NeuroRobotics* 13.
- [112] S. Fujisawa, A. Amarasingham, M. Harrison, G. Buzsaki, Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex, *Nature Neuroscience* 11 (7) (2008) 823–33.
1215
- [113] J. Ainge, M. van der Meer, R. Langston, W. E.R., Exploring the role of context-dependent hippocampal activity in spatial alternation behaviour, *Hippocampus* 17 (2007) 9881002.
- [114] M. Wilson, B. McNaughton, Reactivation of hippocampal ensemble memories during sleep, *Science* 265 (5172) (1994) 676679.
1220
- [115] G. Dragoi, S. Tonegawa, Preplay of future place cell sequences by hippocampal cellular assemblies, *Nature* 469 (7330) (2011) 397401.
- [116] P. Smolensky, Tensor product variable binding and the representation of symbolic structures in connectionist systems, *Artif. Intell.* 46 (1990) 159216.
1225

- [117] E. M. Ellis, M. R. Gonzalez, G. O. Dek, Visual prediction in infancy: What is the association with later vocabulary?, *Language Learning and Development* 00 (2013) 1–15.
- 1230 [118] R. Baillargeon, Physical reasoning in young infants: Seeking explanations for unexpected events, *British Journal of Developmental Psychology* 12 (1994) 933.
- [119] R. Baillargeon, J. Li, W. Ng, S. Yuan, An account of infants physical reasoning, A. Woodward & A. Needham (Eds.), *Learning and the infant mind*. New York: Oxford University Press (2009) 66116.
- 1235 [120] G. Rizzolatti, A. Arbib, Language within our grasp, *Trends in Neuroscience* 21 (1998) 188–194.
- [121] M. Arbib, From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics, *Behavioral and Brain Sciences* 28 (2005) 105167.
- 1240 [122] J. Schwartz, M. Sato, L. Fadiga, The common language of speech perception and action: a neurocognitive perspective, *Revue Franciase de Linguistique Applique*.
- [123] L. Fadiga, L. Craighero, A. DAusilio, Brocas area in language, action, and music, *The Neurosciences and Music IIDisorders and Plasticity: Ann. N.Y. Acad. Sci.* 1169 (2009) 448458.
- 1245 [124] V. Gallese, G. Lakoff, The brains concepts: The role of the sensory-motor system in concpetual knowlegde, *Cognitive Neuropsychology* 21 (0).
- [125] A. Meltzoff, R. Borton, Intermodal matching by human neonates, *Nature* 282 (1979) 403–404.
- 1250 [126] G. Pezzulo, M. van der Meer, C. Lansink, C. Pennartz, Internally generated sequences in learning and executing goal-directed behavior, *Trends in Cognitive Sciences* 18 (2014) 647657.

- [127] G. Pezzulo, F. Rigoli, K. Friston, Hierarchical active inference: a theory
1255 of motivated control, *Trends Cogn. Sci.* 22 (4) (2018) 294–306.
- [128] F. Kaplan, P.-Y. Oudeyer, In search of the neural circuits of intrinsic
motivation, *Frontiers in Neuroscience* 1 (1) (2007) 225–236.
- [129] S. Singh, R. Lewis, A. Barto, J. Sorg, Intrinsically motivated reinforcement
1260 learning an evolutionary perspective, *IEEE Transactions on Autonomous Mental Development* 2 (2) (2010) , 70–82.

Reconstructed OFC Dynamics for two dlPFC units

dlPFC unit #1

dlPFC unit #2

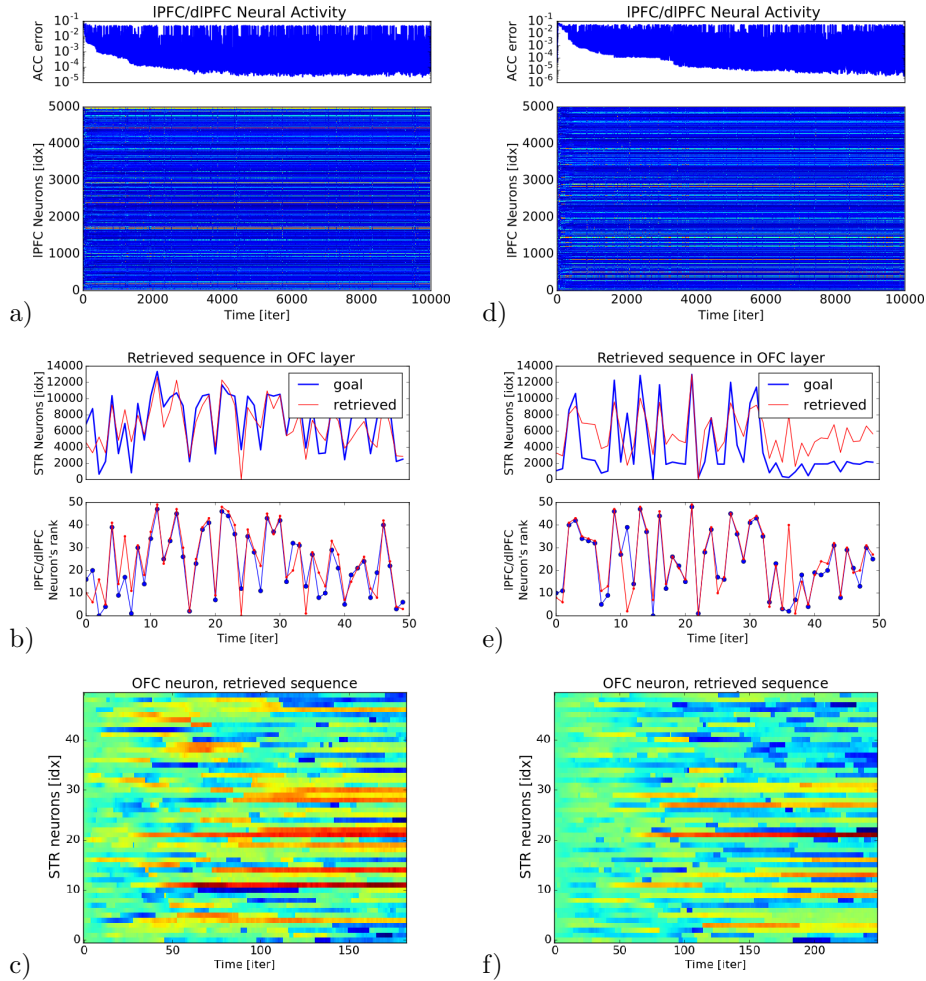


Figure 6: Free-energy optimization for retrieving sequences in OFC layer for two dlPFC units. In a) and d), we show error minimization in ACC unit (top chart) and optimization of dlPFC activity through exploration in OFC and observed in IPFC layer (bottom chart). The desired sequences we want to reconstruct in the OFC are presented in b) and e). We show the final sequence retrieved in red in the OFC layer with the neurons index of the STR layer between [0, 14.000] with respect to a goal sequence in blue (top chart). In c) and f), we show the raster plot of the reconstructed OFC/B45 dynamics. The OFC layer manages to retrieve a temporal sequence with some uncertainty but the global structure of the sequence and the rank orders are mostly respected.

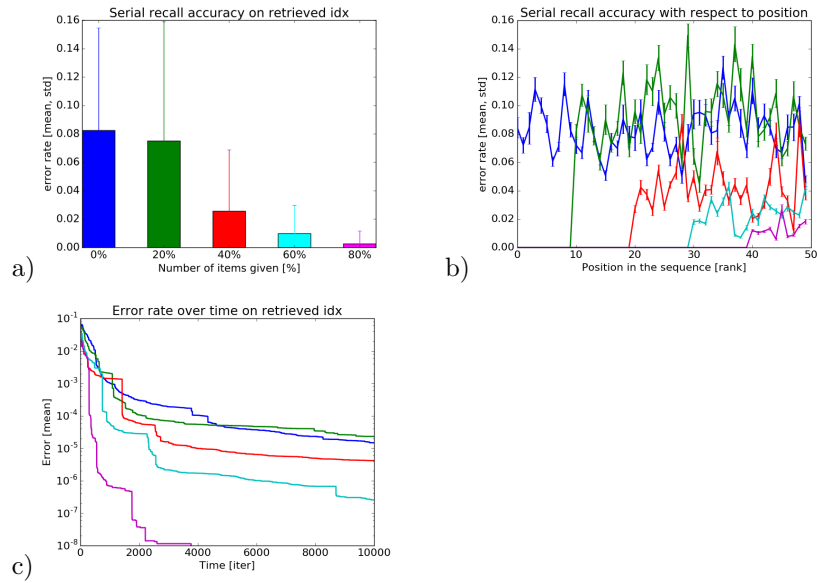


Figure 7: Performance analysis of the Inferno gate architecture for retrieving sequences with respect to the amount of information given. In a), we show the error rate computed from the goal sequence encoded in IPFC/dIPFC and retrieved sequences in the OFC layer with respect to the amount of items given from 0 to 80% of the sequence given at the B45 level. The fewer units there are to search, the more accurate is the recall. If we provide 40% of the items of the sequence we want to retrieve, the error on the neurons' index is particularly small and almost error free if 80% of the neurons are given. In b), we show the distribution of the error rate with respect to the position within the sequence and the amount of information provided. The error made by the network in retrieving the sequences is not related to the temporal position within the sequence but to the amount of information furnished to it. In c), we show the convergence rate of the network with respect to the amount of information provided. The blue color corresponds to 0% of items given, the green corresponds to 20% of items given, the red corresponds to 40% of items given, cyan corresponds to 60% of items given and purple corresponds to 80% of items given.

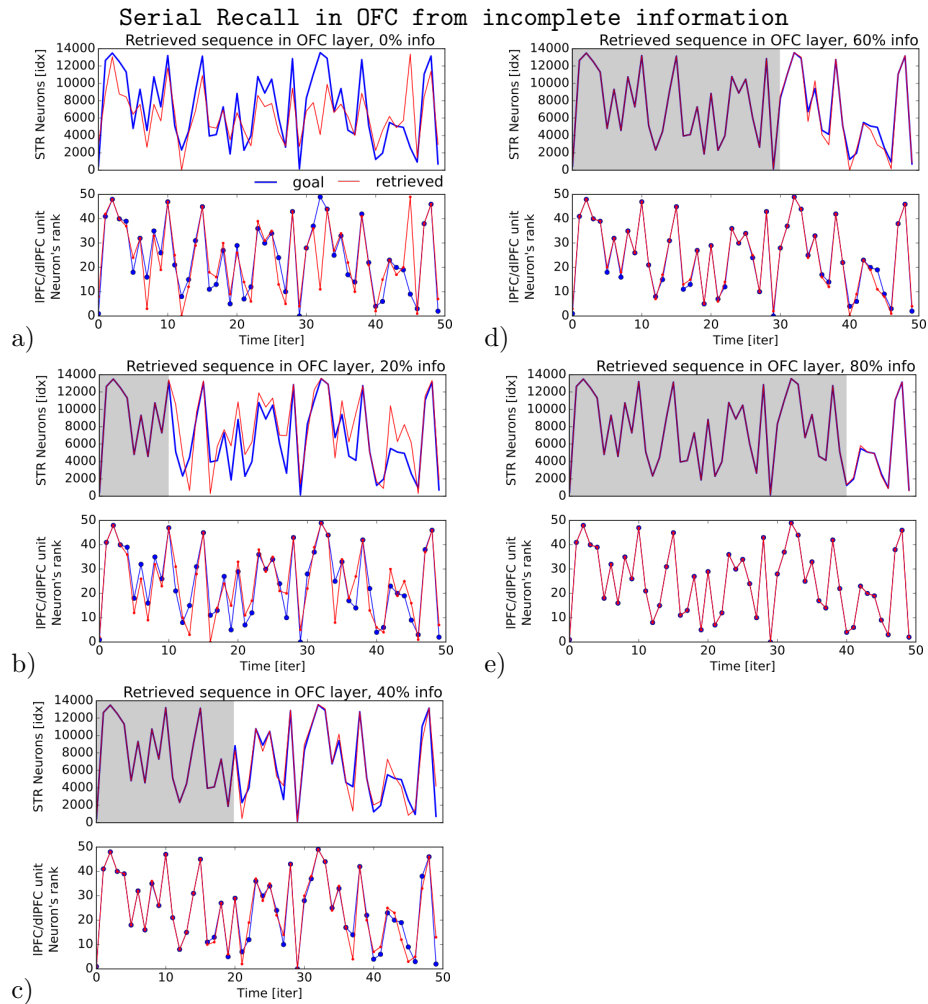


Figure 8: Serial recall in OFC layer from incomplete information. Retrieved goal sequence when 0%, 20%, 40%, 60%, 80% information is furnished to the system, respectively a-e). In the top charts, the generated sequences in the OFC layer with index of the STR neurons are displayed in red with the goal sequences to be retrieved in blue. The more information is given to the system, the easier is the explorative search to retrieve the missing units' index. In the bottom chart, although the rank order in the temporal patterns of the units in the IPFC is respected, this does not guarantee that the units' index is retrieved correctly in the OFC sequences.

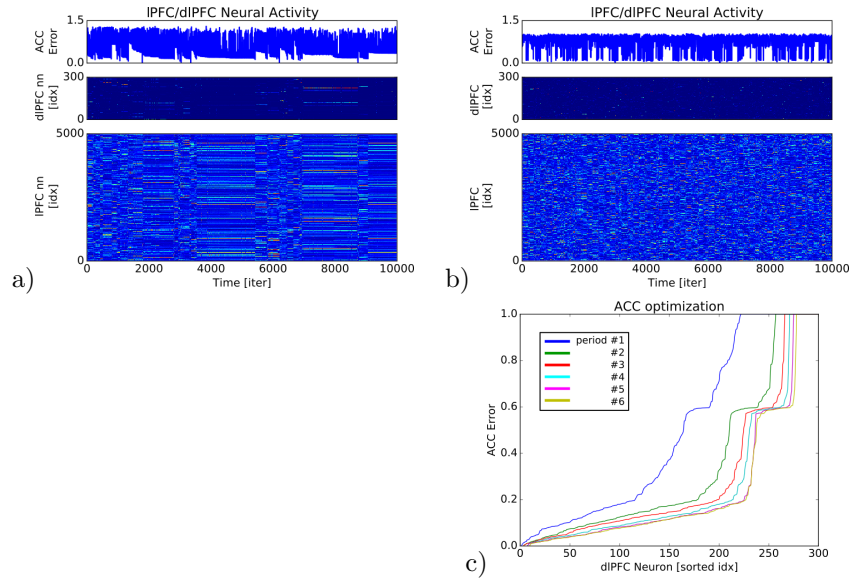


Figure 9: Raster plots during self-supervising at two different learning period. In a) and b), the bottom chart represents the IPFC dynamics, the middle chart represents the dIPFC dynamics and the top chart represents the ACC error rate. The IPFC dynamics Inferno Gate at period #0 learns to categorize ongoing sequences with respect to their structure. In a), error rate in the ACC layer is slow to converge as Inferno Gate generates novel sequences that have to conform to the corresponding temporal order of the selected dIPFC neuron. In b), over time, Inferno Gate has minimized error prediction within the different networks to recognize and retrieve structured sequences online. In c), the graph shows the sorted Acc error presented over different learning periods corresponding to several million of iterations. This graph shows that Inferno Gate achieves to combine stability and plasticity within the network dynamics with stable convergence rate to a minimal error probability density distribution and variability with constrained dynamics.