

Chapter 5

Learning spatio-temporal behavioural sequences

Abstract Living beings are able to adapt their behaviour repertoire to environmental constraints. Among the capabilities needed for such improvement, the ability to store and retrieve temporal sequences is of particular importance. This chapter focuses on the description of an architecture based on spiking neurons, able to learn and autonomously generate a sequence of generic objects or events. The neural architecture is inspired by the insect mushroom bodies already taken into account in the previous chapters as a crucial centre for multimodal sensory integration and behaviour modulation in insects. Sequence learning is only one among a variety of functionalities that coexist within the insect brain computational model. We will propose a series of implementations that can be adopted to obtain these objectives and report the simulation results obtained. We will embed these mechanisms also in roving robots thereby proposing forward-thinking experiments.

5.1 Introduction

Animal brains can be studied by modelling relevant neural structures on the basis of behavioural experiments. This research continuously improves our knowledge about learning mechanisms. The developed architectures have been deeply investigated in the last decades both to understand the sources of the impressive animal capabilities and to design autonomous systems and control strategies able to reach improved levels of autonomy in robot acting in non-structured environments.

A deep analysis of the sequence learning processes existing in living beings is a hard task, however insects can represent an interesting starting point. In fact, in insects, neurobiological evidence is provided for processes that are related to spatio-temporal pattern formation and also learning mechanisms that can be used to solve complex tasks including also sequence learning.

As deeply discussed in chapter 1, there are different types of olfactory receptor neurons found in *Drosophila melanogaster*, whose collective dynamics contributes to the encoding of the features (e.g. odorant components) of the source providing

the stimuli. The antennal lobes (ALs) are the first neuropil in the olfactory path; they consist of glomeruli linked to olfactory receptor neurons coming from the antennal receptors. Information is passed on to projection neurons (PNs), which project to protocerebral areas. [42]. The connection with the large number of MB cells allows a boost in dimensionality useful to improve the representation space [46, 12]. At the same time PNs are connected to the lateral horn (LH). In locusts LH inhibits, after a delay, the activity of the Kenyon cells (KCs) neurons [13]. Therefore the KCs receive a sequence of excitatory and inhibitory waves from the PNs and LH, respectively and are believed to communicate with one-another through axo-axonal synapses [17]. In this chapter as well as in the previous one, we modelled the KC layer as a dynamic spatio-temporal pattern generator extracting the relevant dynamics needed to perform a behavioural task by training multiple read-out maps.

To understand neural circuits in insect brains, several approaches can be taken into account and merged: both behavioural and neurophysiological experiments, and the realization of computational models at different levels of complexity. Several different examples of MB models were recently proposed and are available in the literature. One of the first MB models developed for olfactory associative learning was introduced by [15]. This model is focused on analysing olfactory conditioning and the effect of positive and negative reinforcement signals. On the basis of the experimental biological evidences, other ideas were exploited to design biologically plausible models of the MBs' neural activity and behavioural functionalities [41].

The self-organization properties available in the MBs are fundamental when the sequence learning problem is considered. An interesting analysis of this aspect was presented in [33], where a model based on spiking neurons and synaptic plasticity, distributed through different interacting layers was proposed. In their studies the MBs are assumed to be multi-modal integration centres involving both olfactory and visual inputs. As strongly supported in our model, their system capabilities are independent of the type of information processed in the MBs.

Other works investigated the interaction of MBs and ALs in non-elemental learning processes [49]. Different levels of learning including reinforcement mechanisms were adopted at the level of the KCs with the aim to develop a non-elemental learning strategy. Our proposed architecture considers learning at the KC layer and also plasticity at the level of the AL as suggested by [38] where filtering mechanisms were applied to reduce noise and reconstruct missing features at the beginning of the neural process.

The role of time is also important as emphasised in the locust olfactory processing [48]. To encode complex natural stimuli such as odours we need to consider the precise timing of the neural activity. All these aspects were considered in our model, where the olfactory system and neural circuits are modelled using dynamical systems able to generate different neural activities that can be associated to a series of behaviours. The time evolution is mapped into a space-distributed dynamic which can be adapted to generate a multitude of concurrent behaviours, among which are sequence learning and retrieval.

The problem of modelling biological nervous system functions by neural dynamics was actively investigated [1, 2]. Different kinds of spiking-based networks were taken into account for the development of the proposed architecture [4, 23].

The designed model was developed in different stages that allowed to continuously improve the functionalities included. In the first neural structure that will be presented, the KCs dynamics converges onto a cluster of activity directly related to the input [3, 5]. The network topology, chosen to obtain this behaviour, is inspired by the winner-takes-all solution: it includes in the lattice local excitatory and global inhibitory connections. Successively, a second approach will be also discussed where we decided to model the KCs' activity as a Liquid State Network (LSN), a lattice of connected spiking neurons similar to a Liquid State Machine [28], that contains mainly local synaptic connections (as in a Cellular Nonlinear Network structure [7, 3] already used for locomotion control [8]), resembling axo-axonal communication among the KCs [17]. This lattice modulates sensory information, creating a dynamic map, which can be exploited concurrently both for classification and for motor learning purposes. Taking into account the results obtained in other works [21], the classification task was developed using the sparse dynamics generated within the KC lattice of neurons implemented using a LSN, where an equilibrium in the firing rate is not requested and the neural activity can continuously change in time. In this scenario, multiple read-out maps can exploit this far-from-equilibrium neural activity (as proposed in [36]) to extract the suitable dynamics needed to solve the on-going task.

In the literature there are different MB-inspired models for classification with structures mainly based on several lattices of spiking neurons [32, 40]. In the here presented model we included a new layer, named context layer, needed to develop sequence learning capabilities into the architecture.

To summarize the main capabilities of the developed system, the relevant behaviours that can arise from this unique model are reported in Table 5.1.

Behaviours	Neural Structures involved	Plasticity
Attention	Antennal Lobe (AL) α - β -lobes	(1) STDP from α - β -lobes to AL (2) Memory effect in the α - β -lobes
Delayed match-to-sample	Antennal Lobe (AL) α - β -lobes α' - β' -lobes Sameness Neuron	(1) STDP from α - β -lobes to AL (2) Feedback synapses from α - β -lobes to α' - β' -lobes lobes and vice-versa (3) Activity Detection by the Sameness Neuron
Expectation	Antennal Lobe (AL) α - β -lobes Context layer	(1) STDP between one feature to other features within the AL (2) STDP From α - β -lobes to AL (3) STDP from Context Layer to α - β -lobes
Sequence Learning	Antennal Lobe (AL) α - β -lobes Context Layer Output Layer	(1) STDP From α - β -lobes to AL (2) STDP from Context Layer to α - β -lobes (3) STDP from Context Layer to Output Layer (4) STDP from α - β -lobes to Output Layer
Motor Learning	Central complex (CX) Intrinsic and extrinsic KC Output Layer	(1) Gating function between CX and MBs (2) Read-out maps learning

Table 5.1 Different behaviours that can be obtained using the proposed architecture. For each behaviour the involved neural structures together with the relevant learning aspects are reported.

Going deeper into details, starting from the basic capabilities of the system, the persistence/distraction mechanisms can be presented. Wild-type insects can follow

a target and thereby avoid fickle behaviour that can arise when distracters are introduced in the scene. Flies with inactivated MBs lose this capability as demonstrated in different experiments with *Drosophila melanogaster*: the attention is continuously switched from the target of interest and the distracter with a considerable worsening in terms of time and energy spent in the process. In the proposed model this attentional capability is performed using feedback connections which produce a memory effect at the level of the KCs in the α - β -lobes. When, in analogy to the MB defective flies, such links are suppressed, the loss of attention is obtained.

Another important capability available in the proposed model is a solution for the delayed match-to-sample task. As illustrated in details in [11], the introduction of the α' - β' -lobes in the architecture allows to identify the presence of two successive presentations of the same element through the detection of an increment in frequency in the α - β -lobes' activity. The acquired information can be also used to elicit, after conditioning, a specific behaviour that can be triggered by a matching detection.

The potentialities of the developed MB-inspired architecture are increased with the introduction of a layer that could be related to the γ -lobe, here called Context layer. This neural structure is used to store information about the sequence of events previously acquired by the system. This capability is relevant for evaluating the neurally encoded causality between consecutively presented objects; expectations on the successive presentation can emerge from this structure [3]. The presence of context layer improves the expectation performance that can be extended from one-step predictions to reproduce sequences of objects, solving also potential ambiguities, exploiting the context that is behind each object.

5.2 Model structure

A first scheme of the proposed architecture is reported in Fig. 5.1. The connections' shape and the weight distribution allow the network to create clusters of activities as shown in Fig. 5.2 where the formation of a cluster of neural activity in the α - β -lobe neurons is shown in different time windows.

5.2.1 Antennal lobe model

Inspired by the insects' ALs, the input layer is able to codify either the odour components (i.e. odorants) or, in a more general scenario, the extracted *features* of presented objects. In the insect ALs each glomerulus receives input from just one type of olfactory receptor; in our model each neuron in the input layer encodes a particular feature related to the object of interest. The AL model contains several neurons organized in groups used to codify a type of feature. The pool of neurons in each group codifies different intensity of the corresponding feature. Within the

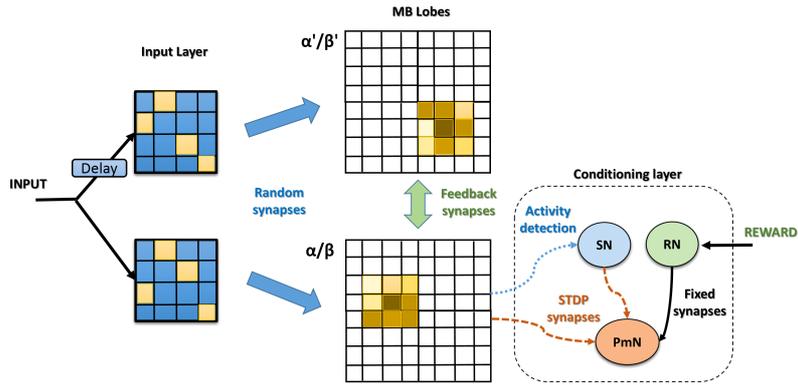


Fig. 5.1 Scheme of a basic computational model of the MB-inspired network. The input layer is randomly connected with the $\alpha - \beta$ and $\alpha' - \beta'$ lobes that are themselves interconnected by synapses subject to learning. The conditioning layer is finally needed to exploit the information embedded in the lobes, through reward-based learning processes.

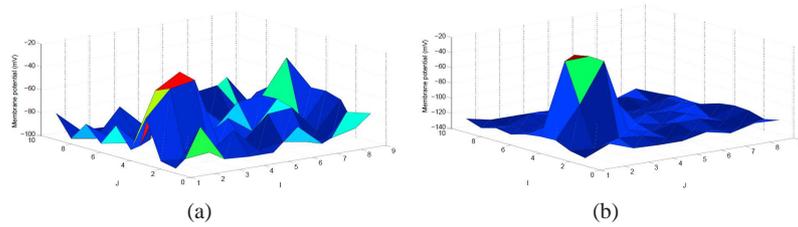


Fig. 5.2 Cluster formation in the $\alpha - \beta$ -lobes. The mean value of the neuron membrane potential is reported for a time window at the beginning of the simulation (a) and at the end when a cluster of activity is established (b).

same group, neurons are linked together through inhibitory synapses to guarantee that only one neuron in each group remains excited (i.e. winner-takes-all solution). Neurons in different groups are connected using plastic synapses that are reinforced when neurons are firing together, according to the STDP mechanism introduced in chapter 2.

In the model, each neuron in the AL layer has a probability $P = 25\%$ of being connected to the KCs. The choice of the sparse connection between the first and the second layer is directly related to the known topology in the biological counterpart [36]. Probability in flies is an average of 4 out of 150 projection neurons per KC.

5.2.2 Initial model of the MB lobes

The KCs in the MBs, as outlined in chapter 1, project through the peduncle into the lobes. The lobes possess roughly the same topology, but are involved in different functionalities. Our first architecture was restricted to model the structure and functions of the $\alpha - / \beta$ -lobes, and the $\alpha' - / \beta'$ -lobes, divided into two distinct neural networks. Each network is able to produce a peculiar dynamics: if excited, the neurons in the AL layer begin a competition that leads to the emergence of one cluster of active neurons.

Each lobe was modelled using a lattice of Izhikevich class I neurons with a toroidal topology. The neurons in this layer are all connected to each other according to the paradigm of local excitation and global inhibition.

The lobes are connected to each other through two sets of synapses, one from the $\alpha - / \beta$ -lobes to the $\alpha' - / \beta'$ -lobes and vice-versa.

We can assume, on the basis of the biological evidences, that information that reaches the $\alpha' - / \beta'$ -lobes is retained there and stored for memory purposes. In particular we hypothesize that the signals coming from the ALs through the calices are delayed while reaching in the $\alpha' - / \beta'$ -lobes.

Under these conditions, the winning cluster in the $\alpha - / \beta$ -lobes represents the input presented to the ALs at the actual step, whereas the winning cluster in the $\alpha' - / \beta'$ -lobe represents the input presented to the ALs at the previous time step. The synapses between the lobe systems are reinforced when there are two clusters simultaneously active in different lobes. This structure is able to detect whether the object presented as input is the same for two different subsequent acquisitions. In fact, under these conditions, the plastic synapses between the lobes create a positive loop between the clusters in the two lobe systems: as a consequence the spiking rate of the active neurons is increased. We will assume the presence of a neuron sensitive to the firing activity of the $\alpha - / \beta$ -lobes network. The sequence of the network evolution is reported in Fig. 5.3. In the first step, two subsequent presentations of the same object generate a positive loop between the two lobe systems that correspond to an increment of spiking rate, whereas during the following presentation, a different object is recognized destroying the loop generated and loosing the boosting in the spiking activity within the lobes. In the developed model, the mean spiking activity of the $\alpha - / \beta$ -lobes is encoded in a neuron used to discriminate the matching/no-matching events. It is possible to find a threshold in the neural activity of the $\alpha - / \beta$ -lobes in order to distinguish the activity in the case of loop and no-loop connection as illustrated in Fig. 5.4.

5.2.3 Premotor area

Biological evidences discussed in chapter 1, revealed that context generalisation, visual attention, adaptive termination and decision making are behaviours that involve MBs [27, 44]. Furthermore, MBs have also a role in the control of motor activity.

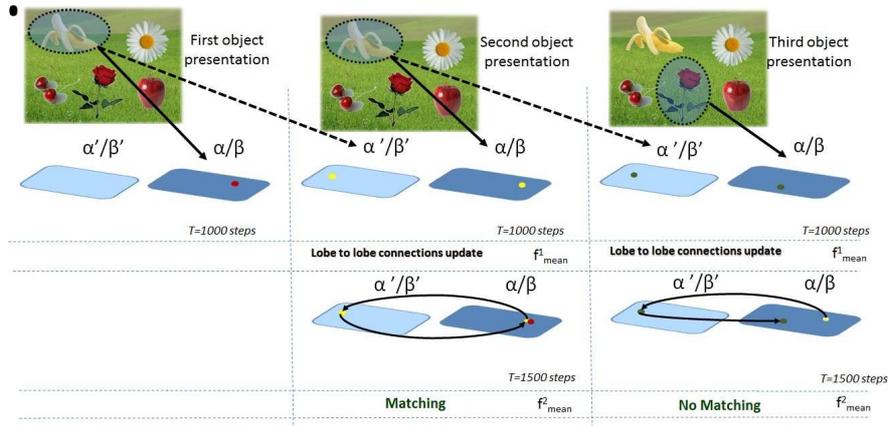


Fig. 5.3 Network evolution due to the presentation of a series of objects. When the first object is presented, its features are processed and a cluster of neural activity arises in the lattice. During the presentation of a second object the $\alpha - \beta$ -lobes behave in a similar way whereas the $\alpha' - \beta'$ -lobes are excited by the lobe-to-lobe connections. If a loop arises, a significant increment in the spiking rate is obtained allowing the matching/no-matching discrimination. This figure was reprinted from [11], Copyright (2013), with permission from Elsevier.

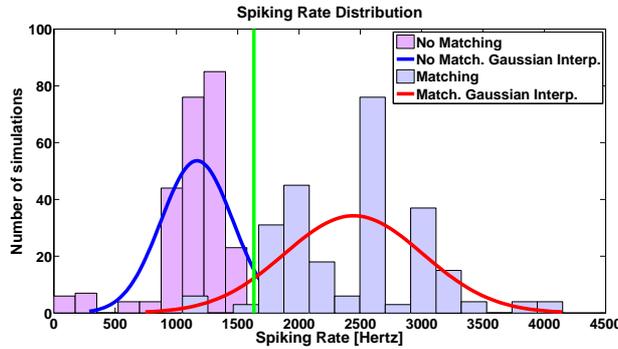


Fig. 5.4 Statistical distribution elaborated using over 500 simulations of the mean spiking activity of the $\alpha - \beta$ -lobes after the reinforcement of the feedback connections between lobes. When a loop is established, the spiking rate is significantly increased and a threshold at $f = 1632\text{Hz}$ can be adopted to distinguish the matching/no-matching of two consecutive presented objects with a processing error of about 3%. This figure was reprinted from [11], Copyright (2013), with permission from Elsevier.

For example, initial motor activity in MB-ablated flies is high, whereas long-term acquisitions show a considerable reduction in motor activity [29].

In the developed architecture the activity of the KCs in the MB-lobes is extracted to determine the system behaviour, realizing a connection with the premotor area devoted to the robot control. The MBs and the premotor area are connected via an associative structure that uses the STDP paradigm (see chapter 2 for details) for a positive/negative-based reinforcement learning.

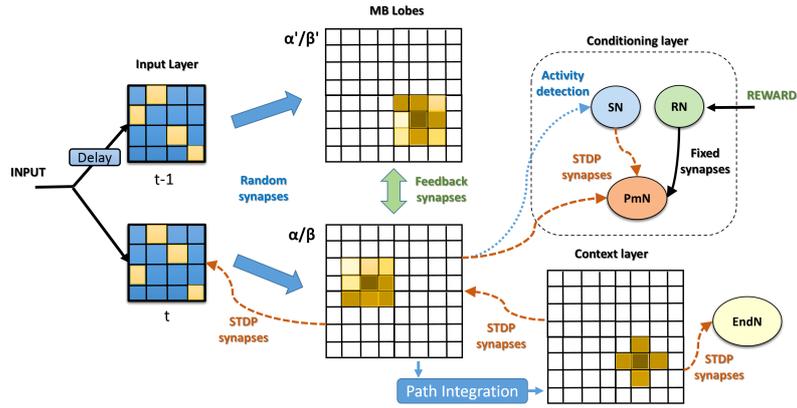


Fig. 5.5 Block scheme of the MB-inspired computational model with the inclusion of the context layer where the history of the sequence is stored using a mechanism similar to path integration. STDP synapses link the active context elements with the next cluster in the $\alpha - \beta$ -lobes and the end neuron is used to identify the end of a sequence.

5.2.4 Context layer

Expectation is the capability of a system to predict the next element on the basis of the last presented one. This one step memory could not be enough to discriminate complex sequences: a memory layer, here called context layer was considered to extend the system capabilities. Experimental results using real data demonstrate that the use of context has three useful consequences: (a) it prevents conflicts during the learning of multiple overlapping sequences, (b) it allows the reconstruction of missing elements in presence of noisy patterns, and (c) it provides a structure for the selective exploration of the learned sequences during the recall phase [14].

A first attempt to develop a Context layer was inspired by the path integration models using the principles of a virtual vectorial sum of the spatial position of the previously emerged clusters creating a spatio-temporal map of contexts. The proposed structure contains a pool of independent neurons spatially distributed in a lattice as illustrated in Fig. 5.5. The horizontal axis indicates the time evolution whereas the vertical axis represents the internal states sequence, forming the context at each time step. The links between the context layer and the $\alpha - \beta$ -lobes are obtained through STDP synapses that realize an all-to-all connections between the two substructures.

Another important element introduced in the architecture is the *End Neuron* (EndN in Fig. 5.5) used to introduce the information about the length of the sequence during the learning phase. Each context neuron can be connected to the end neuron and through STDP can correlate its activation with the end of the sequence. During the testing phase the context activity is reset when the end neuron emits spikes.

The dynamics generated in the context layer resembles the activity produced during a reaction-diffusion process. The idea is to consider the $\alpha - \beta$ -lobes randomly connected to the context layer that is composed of groups of neurons topologically organized in lines. This adopted topology resembles the linewise arrangement of the MB fibres and recalls the granule cells in the cerebellum that are responsible for encoding both the pattern of activations and also the time variables involved that will be used by the Kenyon cell to generate the suitable output activity [16].

The process involving the context formation starts when the presented input generates a winning cluster in the $\alpha - \beta$ -lobes (time t_0) and consequently the lobes randomly excite the context layer. In this time window, only the first column of context neurons is receptive and a winner-takes-all strategy allows the emergence of a single winning neuron as representative for the current state. After a resetting triggered by the lateral horn, a newly presented element (at time t_1) generates a second cluster in the lobes that randomly excites again the context layer. The previous winner in the context starts a diffusion process with a Gaussian shape toward the second column of neurons. The interactions between these two mechanisms are at the basis of the selection of a second neuron that is related to the history of the previously presented elements. All the neurons in the context are massively connected with the $\alpha - \beta$ -lobes by synapses subject to the STDP learning. Therefore the synapses connecting the active neuron that generates the diffusion process, and the current winner in the $\alpha - \beta$ -lobes are reinforced. Multiple presentations of the same sequence of elements guarantee that the synapses between the context layer and the $\alpha - \beta$ -lobes are strong enough to allow the reconstruction of a learned sequence during the recall phase.

Finally, either rewarding or punishing signals can be linked to the last element of a sequence and this information can determine the selection of the most rewarding sequence to be considered when different possibilities are provided to the system.

5.3 MB-inspired architecture: a step ahead

The previously described architecture is not able to address additional functionalities recently ascribed to the MBs, like motor learning where the time evolution of reference signals needs to be acquired and reproduced [3]. Therefore the neural structure of some key layers of the previously discussed MB model were improved, developing an architecture able to both classify static features and learn time dependent signals used as references for the motor system.

The new functionalities taken into consideration need a different processing layer to correctly learn and reproduce spatio-temporal dynamics. We decided to model the KCs' activity as a Liquid State Network (LSN), a lattice of connected spiking neurons similar to a Liquid State Machine [28], that contains mainly local synaptic connections (as in a Cellular Nonlinear Network structure [7, 3] already used for locomotion control [8]), resembling axo-axonal communication among the KCs [17]. The developed lattice elaborates the sensory information, creating a dynamic map,

which can be exploited concurrently both for classification and for motor-learning purposes.

From the classification point of view, our architecture exploits the complex internal dynamics that is extracted and condensed in periodic signals whose frequency is able to stimulate specific resonant neurons. We introduced an unsupervised growing mechanism that guarantees the formation of new classes when needed and a supervised learning method to train the read-out maps.

The idea is to consider the context layer as a pool of neurons topologically organized in concentric circles. In bees there is biological evidence that indicates the presence of this kind of arrangement: the calyx neuropil is concentrically organized [43]. Moreover, patterns of genetic expression in DM revealed that KC axons projecting into the γ -lobe form the circumference of the peduncle, whereas a quartet of axon bundles form the core of the peduncle and project into the α - and β -lobes [19]: a concentric axon bundle is part of the fly MBs.

In our model we hypothesize that each ring is stimulated when an input is presented and we assume that the neural activity-wave propagates from the inner to the outer ring in time following a diffusion-like process. The context neurons are connected to the resonant neurons where a competition with the current input information is performed to produce the output of the network through non-elemental learning processes [24, 49] that are an important building block for the expectation and sequence learning processes.

The important role of resonant neurons in classification of auditory stimuli was already discussed in a series of works related to other insect species like crickets [47, 37, 6] and also for the classification of mechanical data provided by a bio-inspired antenna system [35]. The proposed architecture permits to introduce also motor learning capabilities within the MB computational model [3]. Therefore we can both classify static features and learn time-dependent signals to be used to modulate the motor activity.

5.3.1 Network behaviour for classification

With regards to classification, each class is represented by a resonant neuron realized through a Morris-Lecar model [31]. A training procedure, using the method introduced in chapter 2, allows associating a different resonant neuron with each individual input. The read-out map is trained in a supervised way: a periodic wave has been chosen as target signal with a frequency able to stimulate the corresponding resonant neuron. The first read-out map is trained to generate a sine-wave with a frequency of $62.5Hz$ when the first input signal is provided. When resonant neurons are not excited by the input, a new resonant neuron with a different frequency is allocated and the corresponding read-out map is learned. The frequency range here adopted spreads from $50Hz$ to $250Hz$. This interval was chosen to allow the coexistence of about five different classes. Lower frequencies cannot be considered because at least five periods are used to have a reliable number of spikes in the

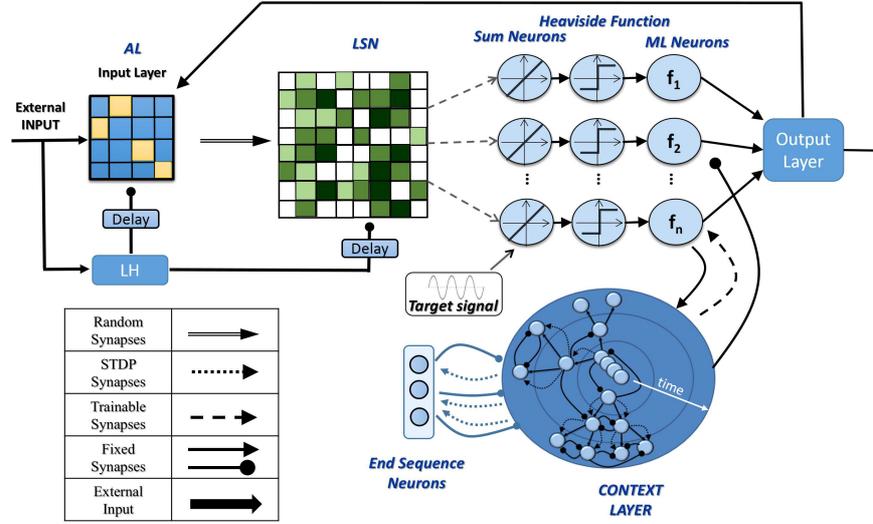


Fig. 5.6 Block scheme of the developed architecture. The external input is processed by the antennal lobes that randomly excite the liquid network; the lateral horn inhibits the lattice activity after a given time window (i.e. 80-100ms). The liquid activity is modulated with multiple read-out maps that are learned in order to excite specific resonant neurons. The active class neuron (i.e. a resonant Morris-Lecar neuron) stimulates the context layer creating a trace of activities that can successively guide the classification of the next stimulus. Feedback from the context layer to the resonant neurons is subject to learning. A pool of End sequence neurons is also employed to reset the context layer activity. The output layer selects the correct behaviour for the system depending on the winning class and can influence the dynamics of the input layer when conditions like sameness and persistence are identified.

resonators for a robust classification. More classes can be learned if the time window defined for the target signal is increased accordingly. Moreover, using a wider frequency spectrum, the time constants used in the LSN should be tuned to generate the frequencies needed. A minimum number of spikes (larger than 50% of the maximum allowable depending on the signal frequency) has been considered to determine, whether the resonator is in an active state. The signal coming from the sum neuron is filtered using a Heaviside function, before entering into the resonator. As illustrated in Fig. 5.6, the context layer structure is constituted of concentric rings (only the first three rings were reported for the sake of simplicity). The first ring contains a number of neurons equal to the current number of classes (N_c). In the second ring, for each neuron of the first ring, there are N_c neurons, and the structure develops like a tree in the successive rings. This implies that, within each ring, there are $N_c^{N_r}$ neurons, where N_r is the ring number. The potentially large number of neurons building-up the context layer is justified, because we are simultaneously learning both sequences and sub-sequences. This possibility boosts the capabilities of the structure, much beyond the simple sequence learning. The number of rings present in the context defines the maximum sequence length. Lateral inhibition among neu-

rons of the same ring generates a competition, filtering out potential disturbances [24].

Sequence learning takes place through different stages (i.e. epochs) characterized by a neural activity either stimulated by an external input or an internal input, generated to recall a missing element stored in the network. During each epoch information propagates one ring ahead, from the inner to the outermost.

In every epoch, the winning neuron in the outermost ring and the winner neuron of the previous ring are subject to an STDP learning process which modulates their connection weights as introduced in chapter 2. In our model, this process could cover larger time scales than the standard STDP. This is required to create correlations among consecutive objects, presentation of which does not happen within the usual STDP time window. Theoretical discussions are presented in [21] whereas further biological evidences are reported in [45] where these learning processes can be modelled using memory traces and reverberation mechanisms [20, 25].

5.3.2 *End sequence neurons*

Each neuron in the context layer is also linked to an end sequence neuron with STDP synapses. In details, all neurons in each ring are connected with the corresponding end sequence neuron arranged in an end sequence vector with a length equal to the number of rings in the context. A rewarding signal, at the end of a sequence, activates the end sequence neuron for the outermost active ring in the context layer. The synapses connecting this end sequence neuron with the winner neuron in the outermost ring will be reinforced accordingly. In our model the reward signal acts as a dopaminergic stimulus on the end sequence neuron to reward the sequence just completed and reset the activity in the context layer for the learning of a new sequence [25].

5.3.3 *Neural models and learning mechanisms*

Different neuron models were used in the architecture to generate the suitable dynamics needed in the subsystems: the Izhikevich's spiking neurons and the Morris-Lecar model (ML) as introduced in Chapter 2. We adopted the Izhikevich Tonic spiking model in the ALs, context layer and end sequence neurons whereas the Class I model was exploited in the LSN.

A decay rate has been introduced to consider dynamically changing environments where the learned sequences could be forgotten if no longer rewarding. Details on applications of this learning paradigm to biorobotics are illustrated in [9]. Spiking neurons in the KC lattice are fully connected to the sum neurons via plastic trainable synapses. A simple supervised learning method based on the pseudo-inverse algorithm [26] has been adopted and compared with an incremental learning

rule as illustrated in [30] where the spiking activity of the neurons is transformed in continuous signals using different functions allowing the evaluation of an error needed for the learning. Different supervised learning methods, based on back propagation for spiking networks [22], have not been adopted due to the presence of recurrent connections in the lattice. Although different learning approaches could be taken into consideration [15], we introduced a simple incremental learning strategy based on the least mean square algorithm, that adapts the synaptic weight using the computed error and the local activity generated by the pre-synaptic neuron, working with the synaptic response (i.e. continuous variables) instead of with the spike train.

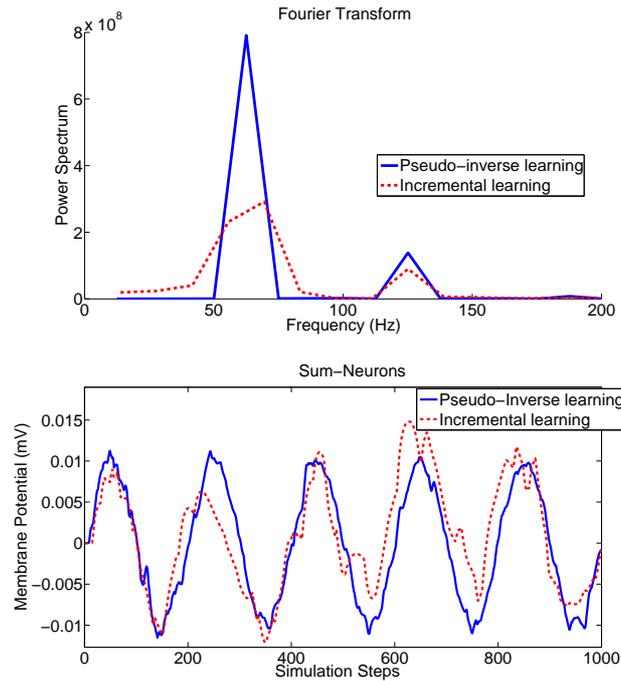


Fig. 5.7 Comparison between the results of the read-out map obtained using a pseudo-inverse (solid line) and an incremental (dashed line) learning method. The power spectrum of the signals (a) and the cumulative output of the sum neuron (b) are reported.

The exit condition for the learning process is obtained monitoring the spikes emitted by the ML neurons: when a given number of spikes is correctly emitted, the learning is stopped and the weights of the read-out map are stored in the architecture. To evaluate the performance of the learning process, a comparison between the pseudo-inverse (standard solution) and the biologically more plausible step-by-step method was performed for a simple interpolation task. The results, shown in Fig. 5.7, illustrate that using an incremental procedure the learning process converges on an oscillatory signal able to properly stimulate the corresponding ML neuron;

in fact, the maximum amount of allowable spikes is emitted. It can be noticed that even if the output signal of the sum neuron does not exactly match the sinusoidal target, the classification phase is successfully accomplished.

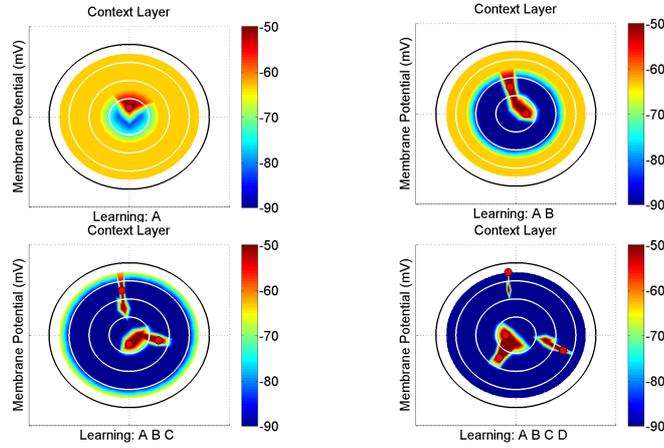


Fig. 5.8 Activity of the Context layer when the sequence ABC is provided and the sub-sequences are learned by the system. The red circles represent the active neurons, multiple chains are generated to trace the subsequences BCD , CD and D acquired in the Context layer.

5.3.4 Decision-making process

The architecture is able to store and consequently retrieve multiple sequences that can be followed in a decision-making task. A rewarding signal, provided to the system at the end of each sequence, is used to evaluate the importance of each stored sequence. In our model the level of reward associated with the generated sequence is directly related to the number of spikes emitted by the end neuron. As an example after learning two sequences $ABBA$ and $BDDC$, the system needs to choose which sequence to generate in front of the presentation of the first objects of the two sequences (A and B). This is a typical case of landmark sequence following to reach a target place. The system internally simulates one and then the other sequence and compares the number of spikes of the end sequence neuron in both cases. In the reported example the sequence $BDDC$ is the most rewarding one and then the behaviours related to this sequence are performed.

5.3.5 Learning sequences and sub-sequences

The developed architecture, thanks to the Context layer topology, allows to learn not only a sequence of elements but also the sub-sequences included therein. We can exploit this capability offered by the model, allowing each presented element to activate, besides a neuron included in the ring after the last active one, also a neuron in the first ring of the Context.

This procedure allows to consider each element of a sequence as a starting point for a new one. This leads to generation and storage of multiple chains of context activity: under presence of very noisy and complex conditions this strategy could be useful to retain only statistically relevant sub-sequences that are reinforced depending on the presentation frequency that is used to improve the sequence reliability.

An example of learning for sub-sequences is shown in Fig. 5.8, where the evolution of the context layer is shown, while learning the sequence *ABCD*. Each element of the sequence creates a trace in the Context layer that starts from the inner ring and propagates towards the outer ones. The memory trace in the Context layer is incrementally strengthened through multiple presentations of the same sequence.

5.4 Robotic experiments

To evaluate the performance of the designed architecture, a series of experiments were carried out with a roving platform. The robot is equipped with a PC on board that communicates with a series of micro-controller-based boards used for the motor control. The sensory system consists of two ultrasound sensors used to detect frontal obstacles and an omnidirectional camera included to identify the presence of visual targets with specific shapes in the environment (i.e. objects of interest). A first experiment is directly related to the persistence behaviour as shown in Fig. 5.9 (a). The robot is attracted by the objects shown on the monitors placed in the arena and is able either to filter out the distracter, replicating the attention capability present in the wild-type insect, or to switch among the presented objects, like in the MB-defective case, as reported in Fig. 5.9 (b), where the robot is continuously distracted by the presentation of the conflictual stimuli.

Honeybees are able to deal with a maze by using symbolic cues [18, 50], and ants are able to navigate following routes [13], therefore we evaluate our control architecture on a roving robot facing with scenarios where the available paths from the entrance to the exit of a maze, have to be acquired and evaluated [51]. The objective is to identify from the visual input the landmarks of interest and process the acquired landmark features using a spiking neural network to learn multiple sequences of events/actions and the corresponding expected rewards.

The information acquired through the visual system is pre-processed using simple segmentation libraries, and then processed by the network to identify the suitable action (or more complex behaviours) to be performed. In the proposed experiment, the robot can perform a turning action followed by a forward motion to reach a new

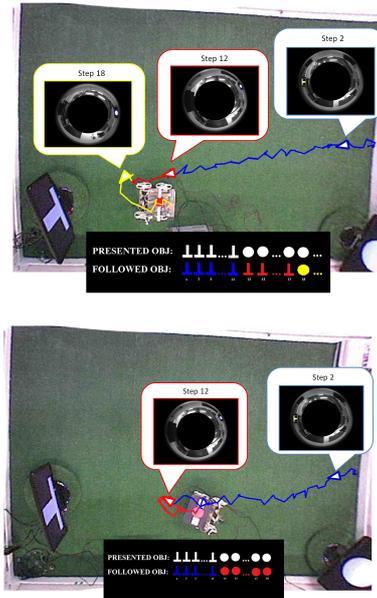


Fig. 5.9 (a) Persistence experiment with a roving robot. Trajectory followed by the robot during the experiment. The images acquired from an on-board fish-eye camera are reported for three different processing steps. The robot is able to persistently move in the direction of the inverted T also when a distracter is presented (i.e. the circle). After multiple presentations, the memory associated with the inverted T fades out and the robot follows the circle. (b) Distraction experiment with a rover. The presentation of a new object at step 12 is enough to change the robot behaviour independently of the time/energy already invested in following the previous target. This figure was reprinted from [5], Copyright (2015), with permission from Elsevier.

branch of the maze. Two monitors are used to simulate the presence of landmarks in each branch of the multiple T-maze. To guide the robot in the maze we considered a double reversed T to indicate a right turn and a double circle for a left turn. During the learning phase, the correct actions to be performed are provided showing only one landmark each time on the monitor positioned in the correct turning direction (see Fig. 5.10 (a)). During the testing phase only one monitor is used and the structure of the maze is modified to demonstrate that the robot is able to solve it using the knowledge acquired in the previous phase (Fig. 5.10 (b)).

These results show that the robot is able to negotiate a maze by using symbolic cues as shown by honeybees [51]. We then included the sequence learning capabilities considering a more complex scenario. The robot initially learns two sequences of actions to be performed to solve the maze following two different routes that guarantee, at the end, reward signals with different values. The first sequence (i.e. *Inverted – T*, *Circle*, *Inverted – T*) is associated to the left, right and then left maze-branch selection, whereas the second learned sequence (i.e. *T*, *inverted – T*) is associated to a right and then left turning action. The reward level that modu-

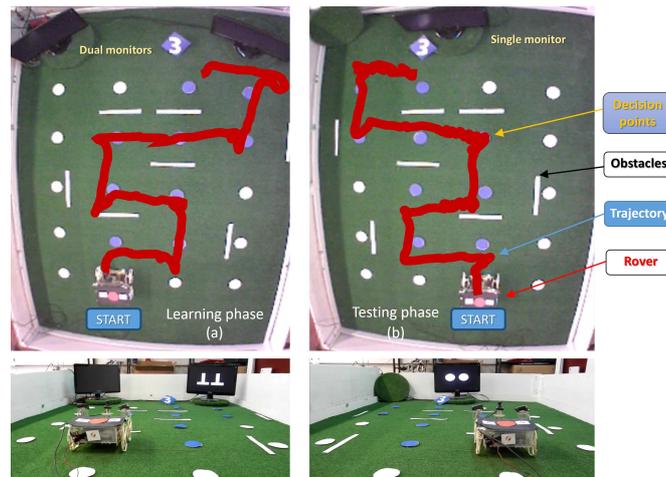


Fig. 5.10 Robotic experiments performed using different maze configurations for the learning (a) and the testing (b) phase. Both the top and lateral view acquired from the starting point are reported. The solid line represents the trajectory followed by the robot. The scenario is equivalent to the biological experiments performed with honeybees [51]. During the learning phase the robot memorizes the correct actions to be associated with each presented element. During the testing phase (b) the robot can solve a different maze using the landmark projected on a single monitor.

lates the stimulation of the end sequence neuron, is lower for the first sequence than for the second one that should be preferred. During the testing phase the robot is placed in front of two concurrent stimuli (i.e. T and $inverted - T$) to perform a decision-making process. In this way the architecture can internally retrieve different sequences, depending on the initial stimulus, and selects the most performing one depending on the spiking activity of the end neuron, that encodes the cumulative level of rewards obtained, for that sequence, during the learning process. The control structure internally simulates the outcome of the two possible sequences to be followed and, on the basis of the expected reward, selects the most rewarding one (Fig. 5.11).

5.5 Conclusions

The ability to understand the environment is a fundamental skill for living beings and needs to be acquired as a dynamic process. The context in which events occur cannot be ignored, in fact, sometimes it is more important than the events themselves. Starting from the biological evidences concerning insect capabilities to learn sequences of events and the known facts on the neural structures responsible for these processes, in this work a neural-based architecture for sequence representation and learning is proposed. The proposed model is therefore based on experimental evidences on insect neurodynamics and on specific hypotheses on the mechanisms



Fig. 5.11 After learning two sequences, the robot selects the most rewarding one performing a decision-making process. The learned sequences are depicted on the two sides; the solid line represents the selected trajectory during the testing phase.

involved in the processing of time-related events. Starting from basic capabilities like attention, expectation and others, the MB-inspired model was extended to include sequence learning with the addition of a context layer. A series of implementations is proposed: simulation and experimental results using a roving robot, demonstrate the effectiveness of the proposed architecture that represents a key structure for the development of a complete insect brain computational model. In conclusion, the model architecture discussed here represents a fundamental building block toward an artificial neural processing structure unifying different functionalities, and performing different behaviours, that biological counterparts are able to show.

References

1. Adeli, H., Park, H.: A neural dynamics model for structural optimization - theory. *Computers and Structures* **57**(3), 383–390 (1995)
2. Ahmadkhanlou, F., Adeli, H.: Optimum cost design of reinforced concrete slabs using neural dynamics model. *Engineering Applications of Artificial Intelligence* **18**(1), 65–72 (2005)
3. Arena, P., Caccamo, S., Patané, L., Strauss, R.: A computational model for motor learning in insects. In: *IJCNN*, pp. 1349–1356. Dallas, TX (2013)
4. Arena, P., Calí, M., Patané, L., Portera, A.: A fly-inspired mushroom bodies model for sensory-motor control through sequence and subsequence learning. *International Journal of Neural Systems* **24**(5), 1–16 (2016)
5. Arena, P., Calí, M., Patané, L., Portera, A., Strauss, R.: Modeling the insect mushroom bodies: application to sequence learning. *Neural Networks* **67**, 37–53 (2015)
6. Arena, P., Fortuna, L., Frasca, M., Ganci, G., Patané, L.: A bio-inspired auditory perception model for amplitude-frequency clustering. In: *Proceeding of SPIE*, vol. 5839, pp. 359–368 (2005)

7. Arena, P., Fortuna, L., Frasca, M., Patané, L.: Sensory feedback in CNN-based central pattern generators. *International Journal of Neural Systems* **13**(6), 349–362 (2003)
8. Arena, P., Fortuna, L., Frasca, M., Patané, L.: A CNN-based chip for robot locomotion control. *IEEE Transactions on Circuits and Systems I* **52**(9), 1862–1871 (2005)
9. Arena, P., Fortuna, L., Frasca, M., Patané, L.: Learning anticipation via spiking networks: application to navigation control. *IEEE Transactions on Neural Networks* **20**(2), 202–216 (2009)
10. Arena, P., Patané, L., Termini, P.: Learning expectation in insects: a recurrent spiking neural model for spatio-temporal representation. *Neural Networks* **32**, 35–45 (2012)
11. Arena, P., Stornanti, V., Termini, P., Zaepf, B., Strauss, R.: Modeling the insect mushroom bodies: Application to a delayed match-to-sample task. *Neural Networks* **41**, 202–211 (2013)
12. Aso, Y., Hattori, D., Yu, Y., Johnston, R.M., Iyer, N.A., Ngo, T.T., Dionne, H., Abbott, L., Axel, R., Tanimoto, H., Rubin, G.M.: The neuronal architecture of the mushroom body provides a logic for associative learning. *eLife* **3** (2014). DOI 10.7554/eLife.04577
13. Baddeley, B., Graham, P., Husbands, P., Philippides, A.: A model of ant route navigation driven by scene familiarity. *PLoS Computational Biology* **8**(8), e1002336 (2012). DOI 10.1371/journal.pcbi.1002336
14. Berthouze L. Tijsseling, A.: A neural model for context dependent sequence learning. *Neural Processing Letters* **23**(1), 27–45 (2006)
15. Brea, J., Senn, W., Pfister, J.: Matching recall and storage in sequence learning with spiking neural networks. *The Journal of Neuroscience* **33**(23), 9565–9575 (2013)
16. Buonomano, D.V., Mauk, M.D.: Neural network model of the cerebellum: Temporal discrimination and the timing of motor responses. *Neural Computation* **6**(1), 38–55 (1994). DOI 10.1162/neco.1994.6.1.38. URL <http://dx.doi.org/10.1162/neco.1994.6.1.38>
17. Cassenaer, S., Laurent, G.: Hebbian STDP in mushroom bodies facilitates the synchronous flow of olfactory information in locusts. *Nature* **448**(7154), 709–713 (2007)
18. Collet, T., Fry, S., Wehner, R.: Sequence learning by honeybees. *Journal of Comparative Physiology A* **172**(6), 693–706 (1993)
19. Davis, R., Han, K.: Neuroanatomy: Mushrooming mushroom bodies. *Current Biology* **6**, 146–148 (1996)
20. Drew, P.J., Abbott, L.F.: Extending the effects of spike-timing-dependent plasticity to behavioral timescales. *Proceedings of the National Academy of Sciences USA* **103**(23), 8876–8881 (2006)
21. Friedrich, J., Urbanczik, R., Senn, W.: Code-specific learning rules improve action selection by populations of spiking neurons. *International Journal of Neural Systems* **24**(05), 1450,002 (2014). DOI 10.1142/S0129065714500026. URL <http://www.worldscientific.com/doi/abs/10.1142/S0129065714500026>
22. Ghosh-Dastidar, S., Adeli, H.: Improved spiking neural networks for eeg classification and epilepsy and seizure detection. *International Journal of Neural Systems* **14**(03), 187–212 (2007)
23. Ghosh-Dastidar, S., Adeli, H.: Spiking neural networks. *International Journal of Neural Systems* **19**(4), 295–308 (2009)
24. Giurfa, M.: Cognitive neuroethology: dissecting non-elemental learning in a honeybee brain. *Current Opinion in Neurobiology* **13**(6), 726–735 (2003)
25. Izhikevich, E.M.: Solving the distal reward problem through linkage of stdp and dopamine signaling. *Cerebral Cortex* **17**, 2443–2452 (2007)
26. Jaeger, H.: Short term memory in echo state networks. *GMD-Report German National Research Institute for Computer Science* **152** (2002)
27. Liu, L., Wolf, R., Ernst, R., Heisenberg, M.: Context generalization in *Drosophila* visual learning requires the mushroom bodies. *Nature* **400**, 753–756 (1999)
28. Maass, W., Natschlaeger, T., Markram, H.: Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation* **14**(11), 2531–2560 (2002)
29. Martin, J.R., Ernst, R., Heisenberg, M.: Mushroom bodies suppress locomotor activity in *Drosophila melanogaster*. *Learning & Memory* **5**, 179–191 (1998)

30. Moheemmed, A., Schliebs, S., Matsuda, S., Kasabov, N.: Span: Spike pattern association neuron for learning spatio-temporal spike patterns. *International Journal of Neural Systems* **22**(04), 1250,012 (2012). DOI 10.1142/S0129065712500128. URL <http://www.worldscientific.com/doi/abs/10.1142/S0129065712500128>
31. Morris, C., Lecar, H.: Voltage oscillations in the barnacle giant muscle fiber. *Biophysics Journal* **35**, 193–213 (1981)
32. Mosqueiro, T.S., Huerta, R.: Computational models to understand decision making and pattern recognition in the insect brain. *Current Opinion in Insect Science* **6**, 80 – 85 (2014)
33. Nowotny, T., Huerta, R., Abarbanel, H., Rabinovich, M.: Self-organization in the olfactory system: one shot odor recognition in insects. *Journal of Computational Neuroscience* **93**, 436–446 (2005)
34. Nowotny, T., Rabinovich, M., Huerta, R., Abarbanel, H.: Decoding temporal information through slow lateral excitation in the olfactory system of insects. *Journal of Computational Neuroscience* **15**, 271–281 (2003)
35. Patané, L., Hellbach, S., Krause, A.F., Arena, P., Duerr, V.: An insect-inspired bionic sensor for tactile localisation and material classification with state-dependent modulation. *Frontiers in Neurorobotics* **6**(8) (2012). DOI 10.3389/fnbot.2012.00008. URL <http://www.frontiersin.org/neurorobotics/10.3389/fnbot.2012.00008/abstract>
36. Perez-Orive, J., Mazor, O., Turner, G., Cassenaer, S., Wilson, R., Laurent, G.: Oscillations and sparsening of odor representations in the mushroom body. *Science* **297**, 359–365 (2002)
37. Russo, P., Webb, B., Reeve, R., Arena, P., Patané, L.: Cricket-inspired neural network for feedforward compensation and multisensory integration. In: *IEEE Conference on Decision and Control* (2005)
38. Sachse, S., Galizia, C.: Role of inhibition for temporal and spatial odor representation in olfactory output neurons: a calcium imaging study. *J. Neurophysiol.* **87**, 1106–1117 (2002)
39. Scherer, S., Stocker, R., Gerber, B.: Olfactory learning in individually assayed *Drosophila* larvae. *Learning & Memory* **10**, 217–225 (2003)
40. Schmuker, M., Pfeil, T., Nawrot, M.P.: A neuromorphic network for generic multivariate data classification. *Proceedings of the National Academy of Sciences USA* **111**(6), 2081–2086 (2014). DOI 10.1073/pnas.1303053111. URL <http://www.pnas.org/content/111/6/2081.abstract>
41. Smith, D., Wessnitzer, J., Webb, B.: A model of associative learning in the mushroom body. *Biological Cybernetics* **99**, 89–103 (2008)
42. Stocker, R., Lienhard, C., Borst, A.: Neuronal architecture of the antennal lobe in *Drosophila melanogaster*. *Cell Tissue Research* pp. 9–34 (1990)
43. Strausfeld, N.J.: Organization of the honey bee mushroom body: Representation of the calyx within the vertical and gamma lobes. *The Journal of Comparative Neurology* **450**(1), 4–33 (2002). DOI 10.1002/cne.10285. URL <http://dx.doi.org/10.1002/cne.10285>
44. Tang, S., Guo, A.: Choice behavior of *Drosophila* facing contradictory visual cues. *Science* **294**, 1543–1547 (2001)
45. Tanimoto, H., Heisenberg, M., Gerber, B.: Experimental psychology: Event timing turns punishment to reward. *Nature*. **430**, 983 (2004)
46. Turner, G., Bazhenov, M., Laurent, G.: Olfactory representations by *Drosophila* mushroom body neurons. *Journal of Neurophysiology* pp. 734–746 (2008)
47. Webb, B., Wessnitzer, J., Bush, S., Schul, J., Buchli, J., Ijspeert, A.: Resonant neurons and bushcricket behaviour. *Journal of Comparative Physiology A* **193**(2), 285–288 (2007). DOI 10.1007/s00359-006-0199-1. URL <http://dx.doi.org/10.1007/s00359-006-0199-1>
48. Wehr, M., Laurent, G.: Odor encoding by temporal sequences of firing in oscillating neural assemblies. *Nature* **384**, 162–166 (1996)
49. Wessnitzer, J., Young, J., Armstrong, J., Webb, B.: A model of non-elemental olfactory learning in *Drosophila*. *J. Neurophysiol.* **32**, 197–212 (2012)
50. Zhang, S., Bartsch, K., Srinivasan, M.: Maze learning by honeybees. *Neurobiology of Learning and Memory* **66**(3), 267–282 (1996)
51. Zhang, S., Si, A., Pahl, M.: Visually guided decision making in foraging honeybees. *Frontiers in Neuroscience* **6**(88), 1–17 (2012)