

# A model of non-elemental associative learning in the Mushroom Body neuropil of the insect brain

Jan Wessnitzer, Barbara Webb, and Darren Smith

Institute of Perception, Action and Behaviour,  
University of Edinburgh, Edinburgh EH9 3JZ, UK

[jwessnit@inf.ed.ac.uk](mailto:jwessnit@inf.ed.ac.uk)

<http://www.ipab.inf.ed.ac.uk>

**Abstract.** We developed a computational model of the mushroom body (MB), a prominent region of multimodal integration in the insect brain, and tested the model's performance for non-elemental associative learning in visual pattern avoidance tasks. We employ a realistic spiking neuron model and spike time dependent plasticity, and learning performance is investigated in closed-loop conditions. We show that the distinctive neuroarchitecture (divergence onto MB neurons and convergence from MB neurons, with an otherwise non-specific connectivity) is sufficient for solving non-elemental learning tasks and thus modulating underlying reflexes in context-dependent, heterarchical manner.

## 1 Introduction

Insects are well adapted to their respective ecological niches, but this does not mean (as is often assumed) that they only perform reflexive, hard-wired behaviours. Insects (and other invertebrates) have been shown to have complex and flexible capabilities. For example, honeybees can solve 'delayed match to sample' and 'delayed non-match to sample' tasks [1] and appear to be able to learn concepts such as symmetry [2]. Studying and understanding these competences (which might be considered minimalist solutions for cognition) in invertebrate brains may ultimately help in better understanding more complex vertebrate brains, and in providing useful design methodologies for intelligent robotics.

A large body of evidence suggests that the mushroom body (MB), a distinct region in the insect brain, plays a cardinal role in adaptive behaviour. One of the central functions linked with the MB is olfactory associative learning and memory [3]. However, the MB in many species receives input from a variety of sensory modalities and is involved in multimodal sensory integration. Its roles have been reported to include context generalisation [4] and place memory [5]. Furthermore, there is evidence that MB neurons react differently to self-generated stimuli and other stimuli [6], suggesting proprioceptive or 'efference copy' input. The MB is thus a potential neural substrate for associations and transfer between sensory modalities, underlying context-specific and non-elemental forms of learning. These non-elemental forms of learning constitute the main interest of this paper.

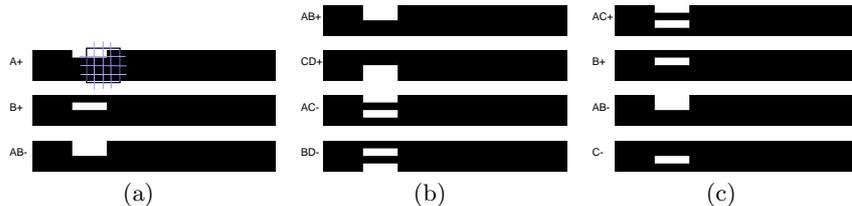
Thus far, computational models of MB function have been restricted to classification of sensory inputs in open-loop conditions [7,8]. In this paper, we develop a MB model that modulates reflexive sensorimotor loops through non-elemental associative learning [9], that is, forms of learning that go beyond simple associations between two stimuli (classical conditioning) or between a stimulus and a response (instrumental conditioning). In non-elemental learning tasks, the stimuli are ambiguously associated with reward or punishment [9]; each stimulus is followed as often by appetitive (+) as aversive (-) reinforcement so that learning requires the context of the stimulus to be taken into account. In negative patterning, the agent has to learn to approach (appetitive action) the single stimuli A and B but retreat (aversive action) from the compound AB. In biconditional discrimination, the agent has to learn to respond appetitively to the compounds AB and CD but aversively to the compounds AC and BD. In feature neutral discrimination, the agent has to learn to respond appetitively to B and AC but aversively to C and the compound AB. In our simulation experiments, we take ‘reinforcement’ and ‘punishment’ to be sensory cues causing different reflex responses (appetitive or aversive); in successful learning, these responses become associated with the appropriate conditioned stimuli.

Non-elemental learning task	Stimuli-reward combinations
Negative patterning	A+ B+ AB-
Biconditional discrimination	AB+ CD+ AC- BD-
Feature neutral discrimination	AC+ C- AB- B+

We propose a minimalist architecture able to modulate reflex behaviours in closed-loop conditions (where the system’s output influences the system’s inputs) for non-elemental learning tasks. In this paper, we show that the general neuroarchitecture of the MB (fan-out and fan-in) is sufficient for explaining the above forms of non-elemental learning. Section 2 describes the simulation framework and Section 3 outlines the general architecture (as suggested by neurobiology). Section 4 describes the neural model in detail. The MB model uses a biologically plausible neuron model and synapses obey a local spike-time dependent plasticity rule. Section 5 presents the simulation results. Section 6 concludes and discusses directions for future work.

## 2 Experimental setup for non-elemental learning

Our experimental set-up was inspired by conditioning paradigms for visual pattern avoidance in flies, in which the animal in a flight simulator learns an appropriate yaw response to a particular visual pattern which is associated with an unpleasant heat beam. In our simulation, the agent has a limited field of view (45 degrees) on a ‘wallpaper’ (of 360 degrees total width) that displays different patterns. In the absence of any action by the agent, the field of view is moved gradually to the left, at 1.5 degrees per millisecond. If it reaches the left edge the agent is “punished” - this generates a reflex action, which moves the field of view back 180 degrees to the right. Before it reaches the edge it will



**Fig. 1.** The wallpapers used for the non-elemental learning tasks: (a) negative patterning, (b) biconditional discrimination, and (c) feature neutral discrimination. The agent’s field of view is represented by the 4-by-4 grid. As the field of view is gradually moved to the left, the visual patterns predict what the agent will experience (+ or -) when it reaches the left edge. Refer to text for further explanations.

encounter a visual pattern, which can thus be used to predict that the edge will be encountered. The aim is to learn to associate the reflex action with the visual pattern and execute it before encountering the edge, thus avoiding punishment. This anticipatory or conditioned reflex, if executed, will move the field of view 21 degrees to the right.

In the non-elemental learning tasks, there are two reflexes, X0-V0 and X1-V1 (these could be called ‘appetitive’ and ‘aversive’, but in fact they have the same effective result of turning the agent back to 180 degrees). There are two corresponding modes for the simulator, i.e. when the field of view reaches the left edge, the agent experiences either X0 or X1, and will execute the corresponding reflex, V0 or V1. Which experience will occur is predicted by the visual pattern, according to the schemes illustrated in figure 1; for example, in negative patterning, the patterns A and B predict X0(+), and the pattern AB predicts X1(-). The agent must learn to execute the correct reflex (V0 or V1) when it sees a particular visual pattern, which will move it away from the edge. If it executes the wrong reflex, then it is instead moved further towards the edge (i.e. 21 degrees to the left).

The field of view contains 45-by-45 pixels, which are mapped onto a 4-by-4 set of sensory neurons (see network description below). White areas on the wallpaper excite these neurons, thus stimulus A will excite the first row of neurons, B the second row and so on. A typical simulation run lasts 50 seconds, during which the wallpapers are switched every 0.5 seconds. The simulation timestep is 0.25ms.

### 3 The mushroom body model

The neural architecture for the agent is based on the insect brain [10], in particular, on evidence that the MB is involved in modulating more basic, reflexive behaviours [11] [12] and thus acts as a neural substrate for associations underlying context-specific and non-elemental forms of learning. However, the goal of the implemented model design was not to imitate physiological mechanisms

involved in MB-mediated learning as closely as possible, but rather to find an abstract description of the underlying principles, able to reproduce associative learning in closed-loop conditions. Yet, we aim to use realistic models of the biological components as more realistic models can be quantitatively and qualitatively different from more abstract connectionist approaches [7].

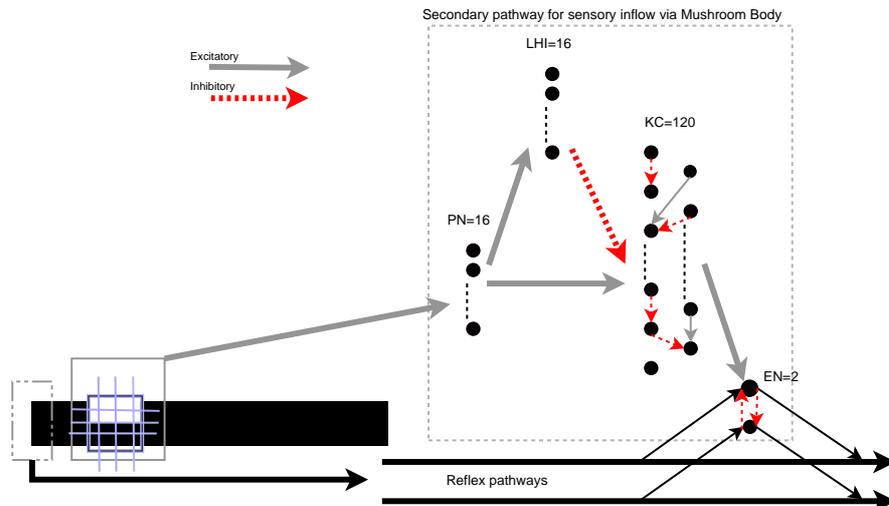
Detailed discussion of insect brain architecture is provided in [10]. The main idea is that of parallel pathways, with sensory inputs forming direct reflex loops, but also feeding into secondary routes that are used to place information from various sensory modalities or other domain-specific sensorimotor loops into context. The system can thus improve on reflexive behaviours by learning to adapt and anticipate reflex-causing stimuli. This adaptation process is assumed to occur in the MB, which form such a parallel pathway for sensory inputs in the insect brain (see figure 2).

The mushroom bodies in insects have a characteristic neuroarchitecture: namely a tightly-packed, parallel organisation of thousands of neurons, the Kenyon cells. The mushroom bodies are further subdivided into several distinct regions: the calyces (input), the pedunculus, and the lobes (output). The dendrites (inputs) of the Kenyon cells have extensive branches in the calyces, and the axons (outputs) of the Kenyon cells run through the pedunculus before extending to form the lobes. Synaptic interconnections between Kenyon cell axons have been reported [13]. Note that there is considerable divergence (1:50) from a small number of sensory projection neurons (PN) onto the large number of Kenyon cells (KC), and considerable convergence (100:1) from the Kenyon cells onto extrinsic output neurons (EN) (these ratios are estimates based on data from [14]). KC receive direct excitatory input from PN neurons, but also indirect inhibitory inputs from the same neurons via lateral horn interneurons (LHI), arriving shortly after the excitation. These connections are illustrated in figure 2.

It is hypothesised that the MB help disentangle spatio-temporal input patterns by operating as coincidence detectors selective to particular correlations in the input spike trains [15]. The mapping of sensory neurons onto MB neurons shows high divergence which can serve for the recognition of unique relationships in primary sensory channels. In our model, nonlinear transformation, separating the activity patterns in the PN layer into sparse activity patterns in the KC layer, is implemented by a randomly determined connectivity matrix between these layers. EN linearly classify the KC activity patterns. Plasticity of KC-EN synapses is achieved with a spike timing dependent plasticity rule. The EN output mediates conditioned responses by activating the appropriate reflex responses. The inhibition from the LHI, quickly following excitation from the PN, limit integration time for the KC to short time windows, making them highly sensitive to precise temporal correlations.

### 3.1 Model description

We chose the neuron model proposed by Izhikevich [16] since it exhibits biologically plausible dynamics, similar to Hodgkin-Huxley-type neurons, but is computationally less expensive and thus, suitable for large-scale simulation:



**Fig. 2.** The implemented MB network receives sensory cues from the visual field via projection neurons (PN), which make direct excitatory connections, and indirect inhibitory connections (via the lateral horn interneurons (LHI)) to the Kenyon cells (KC). The MB output converges on a small number of extrinsic neurons (EN), which are also excited by the underlying direct reflex pathways, and can activate these pathways. Learning occurs between the KC and EN, allowing anticipation of the reflex responses due to associations with particular visual patterns.

$$C \frac{dv}{dt} = k(v - v_r)(v - v_t) - u + I + [\xi \sim N(0, \sigma)] \quad (1)$$

$$\frac{du}{dt} = a(b(v - v_r) - u) \quad (2)$$

where  $v$  is the membrane potential and  $u$  is the recovery current.  $a = 0.3$ ,  $b = -0.2$ ,  $c = -65$ ,  $d = 8$ , and  $k = 2$  are model parameters.  $C = 100$  is the capacitance,  $v_r = -60$  is the resting potential, and  $v_t = -40$  is the instantaneous threshold potential.  $\xi$  is a Gaussian noise term with standard deviation  $\sigma = 1$ . The variables  $v$  and  $u$  are reset if  $v \geq +35\text{mV}$ :

$$\begin{cases} v \leftarrow c \\ u \leftarrow u + d \end{cases} \quad (3)$$

Synaptic inputs are modelled by:

$$I(t + \Delta t) = gS(t)(v_{\text{rev}} - v(t)) \quad (4)$$

where  $v_{\text{rev}}$  is the reversal potential of the synapse ( $v_{\text{rev}} = 0\text{mV}$  for excitatory and  $v_{\text{rev}} = -90\text{mV}$  for inhibitory synapses) and  $g$  is the maximal synaptic

conductance.  $S(t)$  is the amount of neurotransmitter active at the synapse at time  $t$  and is updated as follows:

$$S(t + \Delta t) = \begin{cases} S(t)e^{\frac{-\Delta t}{\tau_{\text{syn}}}} + \delta & , \text{ if presynaptic spike} \\ S(t)e^{\frac{-\Delta t}{\tau_{\text{syn}}}} & , \text{ otherwise} \end{cases} . \quad (5)$$

where  $\delta = 0.5$  is the amount of neurotransmitter released when a presynaptic spike occurred and  $\tau_{\text{syn}}$  is the synaptic timescale. The simulation timestep  $\Delta t$  is set to 0.25ms.

### 3.2 Network geometry

The network geometry as shown in figure 2 retains proportional dimensions to the MB system in insects but is smaller in size. A strategy based entirely on random connectivity and self-organisation through local learning and competition is explored. Each neuron pair X-X is connected with probability  $p_{X,X}$ . The system implements non-specific connectivity with the exception of full inhibitory connectivity between EN (c.f., [8]). We describe the various network layers, their parameters, and their roles below. Learning occurs only through modulation of the KC-EN connections. We report in section 4 the effects on learning performance of changing connectivity between the LHI and KC layers, and varying the size of the KC layer.

**PN layer.** This layer receives sensory input. The layer consists of 16 neurons (the agent’s 45-by-45-pixels FoV is divided into a 4-by-4 grid - c.f., section 2). The input to a single PN neuron is calculated as follows:

$$\text{IPN} = \frac{\frac{\text{sum of pixel values}}{\text{number of pixels}}}{255} \quad (6)$$

The neurotransmitter released at each timestep is calculated as follows:

$$S(t + \Delta t) = S(t) + \text{IPN} \times \delta. \quad (7)$$

Black areas have a pixel value of 0 whereas white areas have pixel values of 255, thus only white areas in the FoV excite the network.

**KC layer.** The KC layer consists of  ${}^{16}C_2 = 120$  neurons. Each KC will act as a coincidence detector and receive inputs from a small number of PNs ( $p_{\text{PN,KC}} = 0.1$ ). The synaptic timescale  $\tau_{\text{PN,KC}}$  is set to 2 ms. This parameter needs to be small in order to make the KC neurons very sensitive to the relative timing of incoming input from the PN layer. This setup allows the KC neurons to act as coincidence detectors. The synaptic strength of PN-KC synapses needed to be carefully adjusted ( $g_{\text{PN,KC}}$  are initialised uniformly at random in [20,30]). In addition to this we add an uniformly distributed jitter to the synaptic strengths. We implemented excitatory and inhibitory KC-KC connections ( $p_{\text{KC,KC}} = 0.1$ ,  $\tau_{\text{KC,KC}} = 5\text{ms}$ ) with equal probability ( $g_{\text{KC,KC}}$  are initialised uniformly at random in [5,10]).

**LHI layer.** Feed-forward inhibition by lateral horn interneurons (LHI) dampens KC activity in the MB. Thus, the integration time for the KC neurons is limited to short time windows, making them highly sensitive to precise temporal correlations. This was implemented through 16 LHIs receiving their inputs from the PN layer and inhibiting activity in the KC layer ( $p_{\text{PN,LHI}} = 0.2$ ,  $\tau_{\text{PN,LHI}} = 5\text{ms}$ ,  $g_{\text{PN,LHI}}$  are initialised uniformly at random in  $[20,30]$ ,  $p_{\text{LHI,KC}} = 0.1$ ,  $\tau_{\text{LHI,KC}} = 5\text{ms}$ ,  $g_{\text{LHI,KC}}$  are initialised uniformly at random in  $[20,30]$ ).

**EN layer.** Every KC-EN pair is connected ( $p_{\text{KC,EN}} = 1$  and  $\tau_{\text{KC,EN}} = 5\text{ms}$ ). However, the synaptic conductance  $g_{\text{KC,EN}}$  for all synapses is initialised to 0, and is subsequently modified by STDP as described below. The ENs also receive excitatory input from the underlying reflex pathways, thus the learning reflects the coincidence of activity in these pathways and particular patterns of KC activity.

### 3.3 Spike Time-Dependent Plasticity

Synapses are modified using Spike Time-Dependent Plasticity (STDP) which has been observed in biological neural systems (e.g., [17]). In STDP, synaptic change depends on the relative timing of pre- and post-synaptic action potentials. Synaptic conductances are adapted as follows:

$$\Delta g = \begin{cases} A_+ e^{\frac{t_{\text{pre}} - t_{\text{post}}}{\tau_+}} - \frac{g_{\text{max}}}{r} & , \text{ if } t_{\text{pre}} - t_{\text{post}} < 0 \\ A_- e^{\frac{-(t_{\text{pre}} - t_{\text{post}})}{\tau_-}} & , \text{ if } t_{\text{pre}} - t_{\text{post}} \geq 0 \end{cases} . \quad (8)$$

where  $t_{\text{pre}}$  and  $t_{\text{post}}$  are the spiking times of the pre- and postsynaptic neuron respectively.  $A_+ = 2$ ,  $A_- = -1$ ,  $\tau_+ = 50\text{ms}$ , and  $\tau_- = 5\text{ms}$  are parameters. We modified the STDP rule proposed by [18] by adding an additional term  $-\frac{g_{\text{max}}}{r}$  if  $t_{\text{pre}} - t_{\text{post}} < 0$  where  $r = 10^3$  is a parameter. This means that if postsynaptic spikes are not matched with presynaptic ones, the synaptic conductance between them is decreased by this term. If this modification rule of synaptic conductances  $g$  pushes the values out of the allowed range  $0 \leq g \leq g_{\text{max}}$ ,  $g$  is set to the appropriate limiting value ( $g_{\text{max}} = 30$ ).

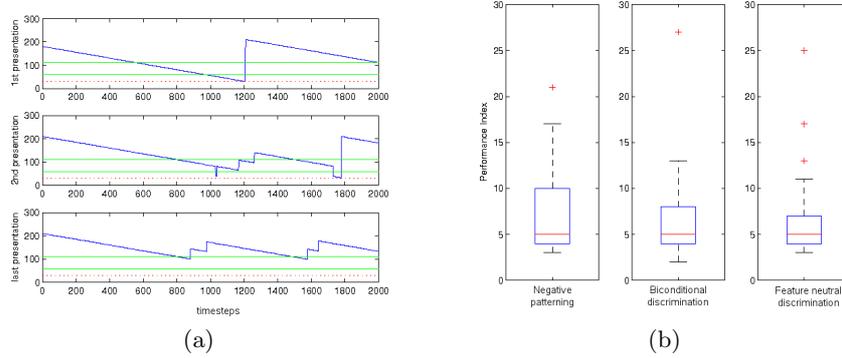
A ‘forgetting’ factor is introduced in the form of a slow decay of  $g$ :

$$g(t + \Delta t) = g(t) e^{\frac{-\Delta t}{\tau_{\text{decay}}}} \quad (9)$$

where  $\tau_{\text{decay}} = 10^5$ .

## 4 Non-elemental discrimination performance

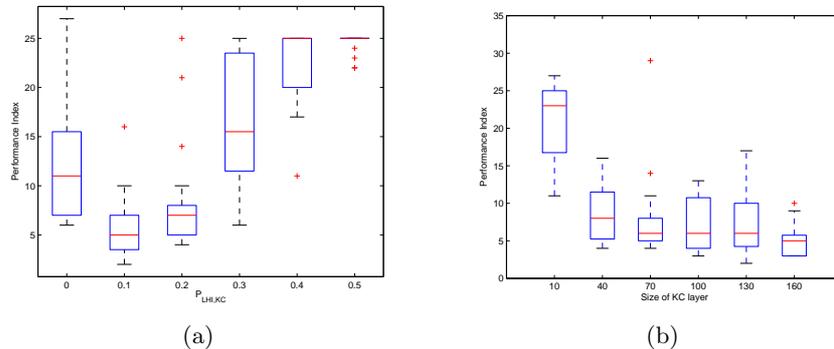
The system was able to learn each of the non-elemental associations shown in figure 1. As the system learns to respond to the visual patterns, the reflex responses to encountering the edge are executed less often and the MB drives the agent’s behaviour (as shown in figure 3(a)). The performance index used in this



**Fig. 3.** (a) Agent behaviour in response to first (0ms), second (2000ms) and final (18000ms) presentations of the same wallpaper during a 20s simulation run. At the start it reaches the edge position (lower dotted line) and performs a reflex turn. In the next presentation it responds to the visual stimulus (between the upper dotted lines) but the response is sometimes incorrect. By the final presentation it reliably responds to the visual stimulus and thus successfully avoids the edge. (b) Boxplots of performance for different learning tasks (1) negative patterning, (2) biconditional discrimination, and (3) feature neutral discrimination. The agent encounters a median of 5 punishments before successfully using the visual patterns to anticipate and avoid it. The simulation runs lasted 50s.

paper is simply the number of times the reflexes are executed. As shown in figure 3(a), the naive system will respond with one reflex action during one presentation of one wallpaper. In the biconditional discrimination setup, for example, there are  $N_w = 4$  wallpapers which are interchanged every  $t_c = 0.5$ s during  $t_T = 50$ s. Thus,  $t_T/(N_w \times t_c) = 25$  activations would mean that a run was unsuccessful. Figure 3(b) shows boxplots of the number of times reflex pathways were active over 30 simulation runs (each lasting 50 seconds) for each of the three conditioning paradigms. All simulation runs for negative patterning were successful and only one simulation run for biconditional and feature neutral discrimination each was unsuccessful. The agent learnt after a median of 5 activations of the reflex pathways which reflex to use, in response to which visual patterns, to successfully avoid the edge.

Figure 4(a) shows the learning performance with varying probability of connectivity  $p_{LHI,KC}$ . As the connectivity between LHI-KC neurons increases (and with it the inhibition from this layer), the learning performance becomes maximal at  $p_{LHI,KC} = 0.1$ . As the inhibition increases further, the performance drops off. With increasing inhibition by the LHI neurons, the activity in the KC layer becomes sparser. Figure 4(b) shows the network performance with varying KC-layer size with probability of connectivity  $p_{LHI,KC} = 0.0$ . The performance tends to improve with increasing KC layer size.



**Fig. 4.** Boxplots of (a) performance with varying probability of connectivity  $p_{LHI,KC}$  between the LHI-KC layers. Performance with varying KC layer size (10,40,70,100,130, and 160 neurons) for a probability of connectivity  $p_{LHI,KC} = 0.0$ .

## 5 Discussion

Our aim in this paper is to explore the capabilities of an insect-inspired brain architecture, consisting of reactive behaviours which are modulated by the MB. We adapted and simulated a widely used conditioning paradigm (the *Drosophila* flight simulator) for non-elemental tasks of visual pattern avoidance, and tested the role of the neuroarchitecture of the MB for this task. The distinct neuroarchitecture of the MB, where input (PN) neurons diverge onto large numbers of mushroom body (KC) neurons and output converges onto a relatively small number of output (EN) neurons, is able to recognise unique relationships of excitations in the primary sensory channels indicating particular stimulus situations. The output (EN) neurons, receiving the output of mushroom body neurons, mediate conditioned responses.

Based on the proposed architecture and the presented simulation experiments, the following claims can be made. First, the neuroarchitecture is suited for pattern recognition and was successfully demonstrated for non-elemental learning tasks. Note that these tasks essentially require the agent to learn to associate different patterns of stimuli, rather than single stimuli, with the correct action. Hence the pattern recognition properties of the MB could indeed be a suitable substrate for this form of learning (c.f., [9]). Second, coincidence detection and sparse coding are useful for learning. We could show that learning performance is maximal for small levels of inhibition from the LHI layer. Sparse coding of sensory inputs helps “make neurons more selective to specific patterns of input, hence making it possible for higher areas to learn structure in the data” [19]. Future work will further investigate these issues for more specific (i.e. non random) connectivity between layers and for modulating larger collections of reflex behaviours.

## Acknowledgments

We acknowledge support from EU grant IST-004690 SPARK.

## References

- [1] Giurfa, M., Zhang, S., Jennett, A., Menzel, R., & Srinivasan, M. (2001) The concepts of sameness and difference in an insect. *Nature*. **410**:930-933.
- [2] Giurfa, M., Eichmann, B., & Menzel, R. (1996) Symmetry perception in an insect. *Nature* **382**:458-461.
- [3] Brembs, B. (2000) An analysis of associative learning in *Drosophila* at the flight simulator. *PhD thesis*. Julius Maximilians Universität Würzburg.
- [4] Liu, L., Wolf, R., Ernst, R., & Heisenberg, M. (1999) Context generalisation in *Drosophila* visual learning requires the Mushroom Bodies. *Nature* **400**:753-756.
- [5] Mizunami, M., Okada, R., Li, Y., & Strausfeld, N. (1998) Mushroom bodies of the cockroach: their participation in place memory. *J. Comp. Neurol.* **402**:520-537.
- [6] Mizunami, M., Okada, R., Li, Y., & Strausfeld, N. (1998) Mushroom bodies of the cockroach: activity and identities of neurons recorded in freely moving animals. *J. Comp. Neurol.* **402**:501-519.
- [7] Nowotny, T., Rabinovich, M., Huerta, R., & Abarbanel, H. (2003) Decoding temporal information through slow lateral excitation in the olfactory system of insects. *J. Comput. Neurosci.* **15**:271-281.
- [8] Nowotny, T., Huerta, R., Abarbanel, H., & Rabinovich, M. (2005) Self-organization in the olfactory system: one shot odor recognition in insects. *Biol. Cybern.* **93**:436-446.
- [9] Giurfa, M. (2003) Cognitive neuroethology: dissecting non-elemental learning in a honeybee. *Curr. Opin. Neurobiol.* **13**:726-735.
- [10] Wessnitzer, J. & Webb, B. (2006) Multimodal sensory integration in insects - towards insect brain control architectures. *Bioinspir. Biomim.* **1**:63-75.
- [11] Okada, R., Sakura, M., & Mizunami, M. (2003) Distribution of dendrites of descending neurons and its implications for the basic organisation of the cockroach brain. *J. Comp. Neurol.* **458**:158-174.
- [12] Menzel, R. & Giurfa, M. (2001) Cognitive architecture of a mini-brain: the honeybee. *Trends Cogn. Sci.* **5**:62-71.
- [13] Heisenberg, M. (1998) What do the mushroom bodies do for the insect brain? An introduction. *Learning Memory* **5**:1-10.
- [14] Laurent, G. (2002) Olfactory network dynamics and the coding of multidimensional signals. *Nature* **3**:884-895.
- [15] Perez-Orive, J., Bazhenov, M., & Laurent, G. (2004) Intrinsic and circuit properties favor coincidence detection for decoding oscillatory input. *J. Neurosci.* **24**:6037-6047.
- [16] Izhikevich, E. (2006) Dynamical systems in neuroscience: the geometry of excitability and bursting. The MIT Press.
- [17] Dan, Y. & Poo, M. (2004) Spike timing dependent plasticity of neural circuits. *Neuron* **44**:23-30.
- [18] Song, S., Miller, K., & Abbott, L. (2000) Competitive Hebbian learning through spike-timing dependent synaptic plasticity. *Nature Neurosci.* **3**:919-926.
- [19] Olshausen, B. & Field, D. (2004) Sparse coding of sensory inputs. *Curr. Opin. Neurobiol.* **14**:481-487.