The cerebellum in action: a simulation and robotics study

Constanze Hofstötter,[1] Matti Mintz[2] and Paul F. M. J. Verschure[1]

[1]Institute of Neuroinformatics, University and ETH Zurich, Winterthurerstr. 190, CH-8057 Zurich, Switzerland
[2]Psychobiology Research Unit, Department of Psychology, Tel Aviv University, Israel

Abstract

The control or prediction of the precise timing of events are central aspects of the many tasks assigned to the cerebellum. Despite much detailed knowledge of its physiology and anatomy, it remains unclear how the cerebellar circuitry can achieve such an adaptive timing function. We present a computational model pursuing this question for one extensively studied type of cerebellar-mediated learning: the classical conditioning of discrete motor responses. This model combines multiple current assumptions on the function of the cerebellar circuitry and was used to investigate whether plasticity in the cerebellar cortex alone can mediate adaptive conditioned response timing. In particular, we studied the effect of changes in the strength of the synapses formed between parallel fibres and Purkinje cells under the control of a negative feedback loop formed between inferior olive, cerebellar cortex and cerebellar deep nuclei. The learning performance of the model was evaluated at the circuit level in simulated conditioning experiments as well as at the behavioural level using a mobile robot. We demonstrate that the model supports adaptively timed responses under real-world conditions. Thus, in contrast to many other models that have focused on cerebellar-mediated conditioning, we investigated whether and how the suggested underlying mechanisms could give rise to behavioural phenomena.

Introduction

The classical conditioning paradigm was introduced by Pavlov in the early 20th century to study associative learning (Pavlov, 1927; 1928). If an initially neutral stimulus (conditioned stimulus, CS) is repeatedly presented paired with a motivational stimulus (unconditioned stimulus, US) which elicits a certain response (unconditioned response, UR), the CS will eventually trigger a similar response (conditioned response, CR). During training the CS and US are usually presented with a fixed interstimulus interval (ISI) and the acquired CR reaches its peak amplitude (e.g. measured as strength of the EMG for conditioned motor responses) just before the US is expected to occur. Hence, the CR reflects knowledge about an association between CS and US and their temporal relationship.

Classical conditioning of discrete motor responses, such as the eyeblink response, has been studied extensively over many decades (Gormezano *et al*., 1983), and multiple lines of evidence illustrate the central role of the cerebellum in this type of learning (Thompson *et al*., 1983; Lavond *et al*., 1993). The pathways for CS, US and CR have been identified (Kim & Thompson, 1997; Thompson *et al*., 1998) but the relative contribution of the cerebellar cortex and the deep nuclei remains difficult to assess (Welsh & Harvey, 1989; Attwell *et al*., 2001; Mauk, 1997). Synaptic changes in both sites seem to be involved in classical conditioning, but may affect different properties of the CR (Raymond *et al*., 1996). It has been suggested that plasticity in the cerebellar nuclei is permissive for the CR expression while plasticity in the cerebellar cortex controls the CR timing (Perret, 1998; Ohyama & Mauk, 2001; Bao *et al*., 2002).

In this paper we present a neural model of the cerebellum to investigate whether plasticity in the cortex alone is sufficient to adaptively control the timing of CRs. Our model includes five assumptions:

- Inferior olive, cerebellar cortex and deep nucleus are organized in distinct microcomplexes (Ito, 2001) which constitute negative feedback loops (Thompson *et al*., 1998).

- Plasticity in the cerebellar cortex is controlled by these olivo-cortico-nuclear feedback loops (Hesslow & Ivarsson, 1996).

- Induction of long-term depression (LTD) (Ito and Kano, 1982) and long-term potentiation (LTP) (Salin *et al*., 1996) of the synapses formed between parallel fibres and Purkinje cells depends on US-related climbing fibre and CS-related parallel fibre activation. Purkinje cell dendrites function as asymmetric coincidence detectors for these signals (Fiala *et al*., 1996; Wang *et al*., 2000).

- The learning procedure leads to a pause in Purkinje cell activity following CS presentation (Houk *et al*., 1996; Hesslow & Ivarsson, 1994; Medina and Mauk, 2000).

- A CR is triggered by rebound excitation in the deep nucleus upon release from Purkinje cell inhibition (Hesslow, 1994b; Hesslow, 1994a; Aizenman & Linden, 1999).

In addition, we assume that Purkinje cells operate in two distinct modes: a spontaneous and a CS-driven mode.

The relationship between the model's parameters and its learning performance was studied in simulated conditioning experiments. For behaviourists, learning is defined as a long-lasting experience-dependent change in behaviour (Mackintosh, 1974) and can as such only be inferred from observing the interaction of a behaving system with its environment (Webb, 2000). Therefore, we also evaluated the
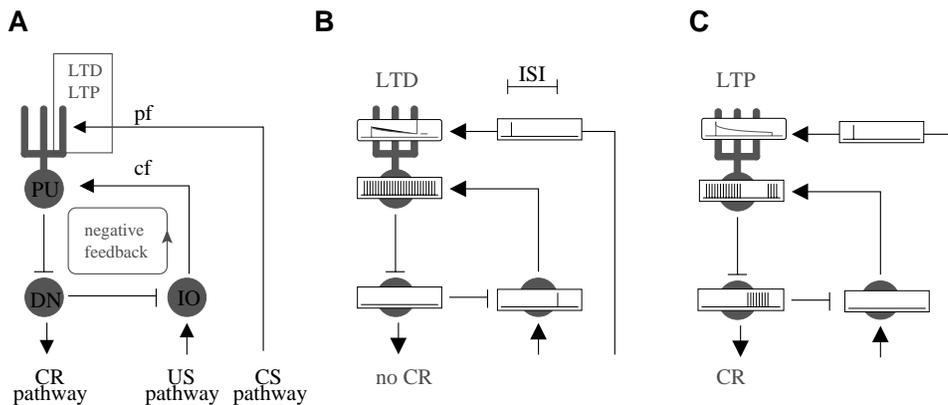
FIG. 1. The learning mechanism embedded in the cerebellar microcircuit. Excitatory connections between cells are indicated with arrows, inhibitory connections are indicated with bars. Boxed traces represent activity in model elements. (A) Purkinje cells (*PU*) receive CS- and US-related input via parallel fibres (*pf*) and climbing fibres (*cf*), respectively. LTD and LTP of synapses formed between *pf* and *PU* depend on the temporal pattern of *pf*- and *cf*-input. *PU* inhibits cells in the deep nucleus (*DN*), which in turn inhibit cells in the inferior olive (*IO*) that give rise to *cf*. Thus, *IO*, *PU* and *DN* form a negative feedback loop. *DN* activity controls the reinforcing pathway by preventing US-related *cf*-activity through *IO* inhibition and it triggers the expression of CRs. (B) In a naive circuit *PU* tonically inhibits *DN* and no CR is expressed. LTD is induced if a *cf*-signal coincides with a prolonged response in the *PU* dendrite due to a previous *pf*-signal. (C) In a trained circuit there is a pause in *PU* spiking following the CS presentation. During this disinhibition *DN* repolarises and triggers a CR. *DN* inhibition of *IO* prevents US-induced *cf*-activation. LTP is induced because the *pf*-signal is not reinforced by a *cf*-stimulus.

learning performance of the model by studying the behaviour of a mobile robot in an obstacle avoidance task. Examining the behaviour of the robot allowed us to determine whether the assumptions embedded in the model can account for associative learning under more realistic conditions (Voegtlin and Verschure, 1999; Verschure, 1998), i.e. when the occurrence of CS and US is not controlled by an experimenter but caused solely by the interaction of the learning system with its environment.

Theoretical background

While it is generally accepted that the cerebellum is critically involved in the classical conditioning of discrete motor responses, the relative contribution of the two sites where CS- and US-related inputs converge, i.e. cerebellar cortex and deep nuclei, is unresolved (Krupa & Thompson, 1997; Mauk, 1997). For instance, Mauk and associates demonstrated that aspiration lesions of the deep nuclei completely abolish CR expression, while lesions of the cortex result in non-adaptive, short-latency CRs (Perrett *et al.*, 1993; Perrett & Mauk, 1995; however, see Llinas *et al.*, 1997; Attwell *et al.*, 2001). Based on these findings it has been suggested that plasticity in the cerebellar cortex and cerebellar deep nuclei serve different functions (Raymond *et al.*, 1996). The 'latent learning' hypothesis states that temporal specific learning occurs initially in the cerebellar cortex, but that a CR can only be elicited after synaptic changes driven by cortical output occurred in the deep nucleus (Ohyama & Mauk, 2001). Most recent findings are consistent with the hypothesis that the cerebellar cortex may play a crucial role in the adaptation of the CR timing (Svensson & Ivarsson; 1999; Hesslow *et al.*, 1999; Bao *et al.*, 2002; Perret, 1998) while synaptic changes in the deep nucleus serve to enhance a CR (Gruart *et al.*, 2000). The goal of our modelling study was to investigate whether synaptic changes within the cerebellar cortex alone are sufficient to support stable CR timing when examined in the context of other aspects of the cerebellar circuit and physiology. This section describes the biological findings that led to the assumptions embedded in our model and illustrates how these different components of cerebellar-mediated learning contribute to adaptive CR timing (Fig. 1).

The negative feedback loop within the cerebellar microcircuit

The most extensively studied example of cerebellar-mediated conditioning is the acquisition and extinction of the eyelid or nictitating membrane response (NMR) (Gormezano *et al.*, 1983). In this conditioning paradigm a tone CS predicts the occurrence of a corneal air puff or periorbital shock US. Parts of lobule HVI, a different region of the cerebellar cortex, anterior parts of the interpositus nucleus and the medial part of rostral dorsal accessory olive are involved in the conditioning of the NMR (Yeo *et al.*, 1985a; b; 1986). Functionally these regions form a microcomplex (Ito, 1984; Ito, 2001). Our model describes the minimal computational unit within such a microcomplex, a microcircuit comprising one Purkinje cell and its peripheral afferents and efferents (see also Barto *et al.*, 1999). For clarity and brevity, we will use abbreviations (listed in Table 1) when referring to model elements. A Purkinje cell (*PU*) receives CS- and US-related signals via parallel fibres (*pf*) and climbing fibres (*cf*), respectively, (Steinmetz *et al.*, 1986; Steinmetz *et al.*, 1989). Activity in the deep nucleus (*DN*) activates motor nuclei via the red nucleus and predicts the amplitude–time course of the motor CR (McCormick & Thompson, 1984; Rogers *et al.*, 2001; however, see Gruart *et al.*, 2000). Distinct cell groups in the inferior olive (*IO*), cerebellar cortex and deep nuclei are interconnected and form negative feedback loops (Hesslow & Ivarsson, 1996; Ito, 2001). It has been demonstrated that inhibition of the inferior olive by the deep nucleus can prevent climbing fibre responses to a US (Kim *et al.*, 1998). Thus, the output of the deep nucleus may control the reinforcing US pathway (Thompson *et al.*, 1998). We address the question of how the negative feedback loop formed between *IO*–*PU*–*DN* affects learning related plasticity in the cerebellar cortex.

Synaptic plasticity in the cortex

While various kinds of plasticity in the cerebellar cortex have been described (for a review see Hansel *et al.*, 2001), the best studied form is the long-term depression (LTD) of the synapse formed between parallel fibres and Purkinje cells (Linden & Conner, 1995; Ito, 2001). It has long been suggested that this type of plasticity plays a crucial role in motor learning (Albus, 1971; Ito & Kano, 1982). More recently, some evidence has been presented for the long-term

potentiation (LTP) of the same synapse (Salin *et al*., 1996; Linden and Ahn, 1991). LTD and LTP of the parallel fibre to Purkinje cell synapse are often assumed to constitute the crucial synaptic mechanism of cerebellar-mediated learning (Mauk *et al*., 1998; Medina *et al*., 2000; Barto *et al*., 1999; Gluck *et al*., 2001) and were therefore included in our model.

Because only a certain range of ISIs leads to behavioural conditioning, the underlying learning mechanism must respond asymmetrically to the occurrence of CS and US (Montague & Sejnowski, 1994). For example, using the delay conditioning paradigm, the NMR can be learned only when the CS precedes the US by ≈ 0.07–3 s (Gormezano *et al*., 1983). It has been shown that the induction of LTD depends on the temporal relationship of parallel fibre and climbing fibre stimulation (Ito, 2001). Although the precise induction properties vary for different preparations, several studies have reported that LTD is strongest if the parallel fibre stimulation precedes the climbing fibre stimulation by 80–250 ms (Schreurs *et al*., 1996; Chen & Thompson, 1995), i.e. intervals which fall within the range of behaviourally effective ISIs.

However, such temporal induction properties for LTD require some persistent trace of the parallel fibre stimulation to allow for coupling with climbing fibre stimulation (Linden & Conner, 1995; Sutton & Barto, 1990). A prolonged metabolic second-messenger response in Purkinje cells following parallel fibre stimulation could constitute such a trace (Fiala *et al*., 1996). The induction of LTD may depend on the coincidence of responses to the second messenger inositol 1,4,5-triphosphate as a result of parallel fibre activity and postsynaptic depolarization or Ca^{2+} entry following climbing fibre activity (Miyata *et al*., 2000). Single spines of Purkinje cells could serve as coincidence detectors of CS- and US-related responses (Wang *et al*., 2000). It has been suggested that there is a negative correlation between the postsynaptic Ca^{2+} concentration and the changes in the efficacy of the synapses formed between parallel fibres and Purkinje cells following parallel fibre stimulation (Hartell, 2002; Shibuki & Okada, 1992). Hence, functionally, LTD and LTP can be interpreted as antagonistic processes. However, because in the cerebellum LTD is a postsynaptic mechanism while LTP occurs presynaptically, it is unclear to what extent LTD and LTP can truly reverse each other at the molecular level (Ito, 2001; Kitazawa, 2002). In our model we implemented a functional approximation of these mechanisms underlying LTD and LTP at the parallel fibre to Purkinje cell synapse (see Fig. 1B and C), where LTD results from coincident *pf*- and *cf*-activation, while LTP results from *pf*-stimulation alone (Hansel *et al*., 2001).

Nuclear release from cortical inhibition

Purkinje cell inhibition forms the sole output of the cerebellar cortex. Thus, if the adaptive timing of the CR depends on the cerebellar cortex there needs to be a mechanism that allows the inhibitory Purkinje cell output to control the timing of excitatory responses in the deep nucleus. Purkinje cells are spontaneously active (Raman & Bean, 1999), supplying tonic inhibition to the deep nucleus. The activity pattern of Purkinje cells is altered during conditioning (Schreurs *et al*., 1998) and some cells show a reduced firing rate to a CS following conditioning (Hesslow and Ivarsson, 1994). Cells in the deep nuclei show a characteristic feature called rebound excitation following their disinhibition (Hesslow, 1994b). In particular, a hyperpolarizing pulse deinactivates low-threshold voltage-gated Ca^{2+} channels (Aizenman *et al*., 1998). Upon the release of inhibition these channels mediate the repolarisation of the membrane potential which is followed by a burst of Na^+ spikes. This suggests that a timed pause in Purkinje cell spiking could control the timing of rebound

excitation (Aizenman & Linden, 1999). Thus, synaptic changes affecting Purkinje cell responses to a CS could control the timing of motor CRs (Gauk & Jaeger, 2000; Medina *et al*., 2000). Based on this assumption, we implemented synaptic changes in our model which alter the excitatory response of *PU* to a presented CS. During a pause in *PU* spiking, *DN* repolarises and can trigger a CR.

Stimulus representation and CS trace

To allow for the expression of a CR adapted to the ISI the temporal relationship of CS and US needs to be encoded. Little is known about the mechanism used in the nervous system to encode time. A recent study demonstrated that individual animals can produce differently timed CRs when stimulation of the same subset of mossy fibres (CSs) is paired with an eye-shock US at various ISIs (Perret, 1998). Even a single-pulse CS can lead to the acquisition of differently timed CRs. These findings indicate that precerebellar pathways are not necessary for maintaining the neural activity to elicit a CR and that the representation of time does not require the recruitment of different mossy fibres by the CS. Mossy fibre activity related to the onset of the CS may thus be sufficient to support correct CR timing. Similarly, climbing fibres fire at ultra-low frequencies (≈ 1 Hz) and could not encode the duration of the US (De Zeeuw *et al*., 1998; Kuroda *et al*., 2001). Based on these findings, we therefore only represented the onset of CS and US as neural activity in our model. Thus, in contrast to many other models of cerebellar information processing (see, e.g., Kawato, 1999; Huang *et al*., 2000; Medina *et al*., 2000) our model does not depend on complex assumptions regarding the encoding of the sensory stimuli.

Various models have been proposed for the putative mechanism underlying the representation of time by the cerebellum, such as tagged elements or delay lines (Braitenberg & Atwood, 1958; Moore & Choi, 1997), a time-varying representation of the CS within the cortex (Mauk, 1997), or oscillatory signals (Gluck *et al*., 2001) and delayed reverberation (Kistler & van Hemmen, 1999). For most of these hypotheses there is no profound physiological evidence, or it has been demonstrated that the biological substrate would not meet the requirements for the suggested mechanism. A signal travelling along parallel fibres will not be delayed by up to several seconds constitute delay lines (MacKay & Murphy, 1976). In our model the prolonged responses in *PU* dendrites constitute a CS as required for parallel fibres to trace allowing for the association of CS and US, which is supported by physiological studies (Wang *et al*., 2000). This internal representation of the CS has a different time course from the CS presented externally and relates to the concept of a stimulus trace suggested by Hull (1939). The notion of an 'eligibility trace' has been used (Sutton & Barto, 1990; Barto *et al*., 1999) to describe that synapses which have been activated by a CS-related input remain eligible to US-induced weight changes for some period of time (see Discussion).

The model

Our neural model is based on the anatomy of the cerebellar microcircuit (see Fig. 2) and requires minimal assumptions on the encoding of stimuli and the representation of time. The goal of the model design was not to resemble the physiological mechanisms involved in cerebellar-mediated learning as closely as possible, but to find an abstract description of the underlying principles. In this section we describe how the implementation of the model function-ally accounts for the assumptions on cerebellar information process-
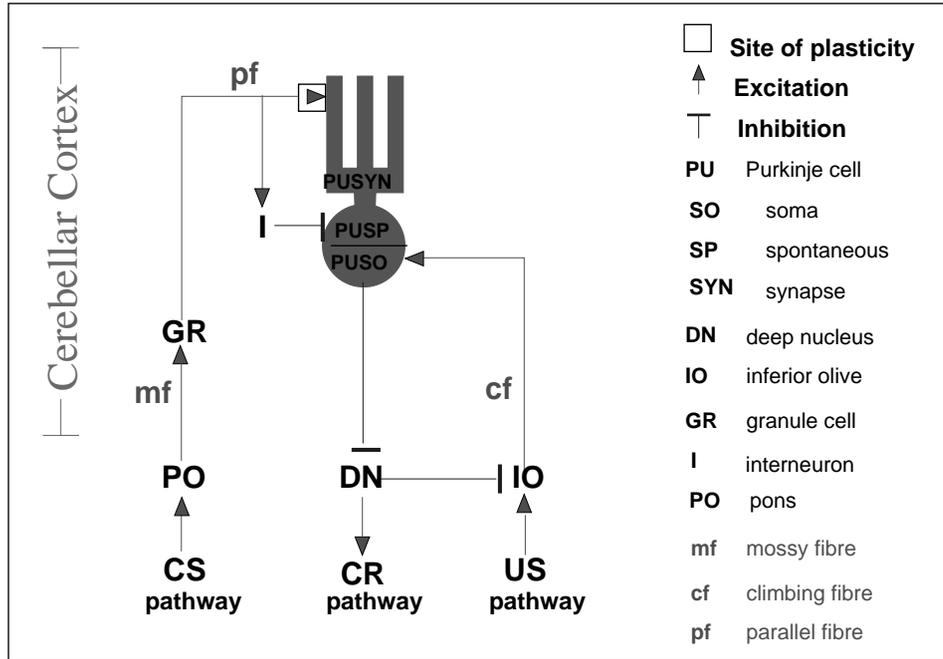
FIG. 2. The anatomy of the model circuit. As discussed in Fig. 1, CS- and US-related input converge on *PU,* and the *IO*, *PU* and *DN* form a negative feedback loop. The model elements *PO*, *GR* and *I* represent pontine nucleus, granule cells and inhibitory interneurons in the cerebellar cortex. The model elements *CSpathway*, *USpathway* and *CRpathway* stand for the pathways upstream of the pontine nucleus and the inferior olive and downstream from the deep nucleus, respectively. *PU* comprises three model compartments: *PU–SYN*, *PU–SP* and *PU–SO*, accounting for responses in synaptic regions, the spontaneous activity and the somatic regions of Purkinje cells, respectively. The CS-related input is conveyed to GR via mossy fibres (mf) originating in PO. Parallel fibres (*pf*) form excitatory synapses at *PU–SYN*, while *PU–SO* receives US-related input via climbing fibres (*cf*). *I* receives input from *pf* and inhibit *PU–SP*. Because we focused on the functional role of the cerebellar cortex only, CS- and US-related signals, conveyed to the deep nucleus via collaterals of mossy fibres and climbing fibres, respectively, were omitted.

ing introduced above. More details on soft- and hardware and lists of model parameters (Tables 1–3) are given in the Appendix.

Model equations

The model elements are based on a generic type of integrate-and-fire neuron. The summed excitatory input of neuron i at time t, $E_i(t)$, is defined as:

$$E_i(t) = \gamma^E \sum_{j=0}^{N} A_j(t) w_{ij}(t) \qquad (1)$$

where γ^E is the excitatory gain of the input, N is the number of afferent projections, $A_j(t)$ is the activity of presynaptic neuron $j \in N$, at time t, and w_{ij} is the efficacy of the connection between the presynaptic neuron j and postsynaptic neuron i.

The summed inhibitory input of neuron i at time t, $I_i(t)$, is defined as:

$$I_i(t) = -\gamma^I \sum_{j=0}^{N} A_j(t) w_{ij}(t) \qquad (2)$$

where γ^I is the inhibitory gain of the input.

The membrane potential of neuron i at time $t + 1$, $V_i(t + 1)$, is given by:

$$V_i(t + 1) = \beta V_i(t) + E_i(t) + I_i(t) \qquad (3)$$

where $\beta \in [0,1]$ is the persistence of the membrane potential which defines the speed of decay towards the resting state.

The activity of an integrate-and-fire neuron i at time t, $A_i(t)$, is given by:

$$A_i(t) = H(V_i(t) - \theta^A) \qquad (4)$$

where θ^A is the firing threshold and H is the Heaviside function:

$$H(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \qquad (5)$$

If V_i of integrate-and-fire neuron i exceeds θ^A, the neuron is active and emits spikes. The duration of a spike is 1 simulation timestep (ts) and is followed by a refractory period of 1 timestep. The model elements *PO*, *GR*, *I* and *IO*, representing pons, granule cells, inhibitory interneurons and inferior olive, respectively, are constructed with such generic integrate-and-fire neurons.

DN *rebound excitation*

A variant of a generic integrate-and-fire neuron was used to model rebound excitation of the deep nucleus, *DN*. The membrane potential of *DN* neuron i at time $t + 1$, $V_i(t + 1)$, is given by:

$$V_i(t+1) = \beta V_i(t) + [H(V_i(t) - \theta^R) H(\theta^R - V_i(t-1)) \cdot \mu] + I_i(t) \quad (6)$$

where θ^R is the rebound threshold and μ is the rebound potential. The potential of *DN* is kept below θ^R by tonic inhibitory input from *PU*. However, when *DN* is disinhibited, its membrane potential slowly repolarises. If the rebound threshold θ^R is reached, the membrane potential is set to a fixed rebound potential μ. Subsequently the

potential of *DN* decays and spikes are emitted as long as it is above the spiking threshold θ^A (equation 4). The output of *DN* activates the CR pathway downstream of the deep nucleus (*CRpathway*) and inhibits *IO*. This inhibition outweighs a US-related excitatory input to *IO*. As a result, US-related *cf*-activity during *DN* rebound activity is prevented.

A three-compartment model of the Purkinje cell

The modelled Purkinje cell, *PU*, is composed of three different compartments: *PU–SP* accounts for the tonic, spontaneous activity of the Purkinje cell, *PU–SO* represents the soma and *PU–SYN* represents the dendritic regions where synapses are formed with parallel fibres. *PU–SP* is a spontaneously active element which emits spikes unless inhibited by *I*. *PU–SO* is a generic integrate-and-fire element and receives excitatory input from *PU–SP* and *PU–SYN*, and *IO*, reflecting *cf*-input. The activity of *PU–SO* forms the output of the modelled *PU*, inhibiting *DN* and activating *CRpathway*. The metabolic postsynaptic responses in Purkinje cell dendrites to parallel fibre stimulation are represented by *PU–SYN*. Unlike a generic integrate-and-fire neuron, *PU–SYN* does not emit spikes but shows continuous dynamics according to:

$$A_i(t) = H(V_i(t) - \theta^A) \, V_i(t) \tag{7}$$

Due to a high persistence, which will be referred to as β^{SYN}, *PU–SYN* shows a prolonged response to a CS-related *pf*-input according to equations 1 and 3. In our model this prolonged response constitutes the intrinsic memory trace for the CS and the eligibility trace for synaptic changes.

The weight of the synapse formed between parallel fibre and Purkinje cell, the *pf–PU*-synapse, is defined as the connection strength between *PU–SYN* and *PU–SO* and can be altered by LTD and LTP. Synaptic changes can only be induced while there is an active stimulus trace of a CS (i.e. $A_{PU\text{-}SYN} > 0$). In this case, at any simulation timestep the induction of LTD and LTP depends on the summed excitatory input to *PU–SO* (see equation 1) according to:

$$w_{ij}(t+1) \begin{cases} \epsilon w_{ij}(t) & \text{if } E_i \in [E_{min}^{LTD}, E_{max}^{LTD}] \\ w_{ij}(t) & \text{otherwise} \end{cases} \tag{8}$$

$$w_{ij}(t+1) \begin{cases} w_{ij}(t) + \eta(w_{ij}^{max} - w_{ij}(t)) & \text{if } E_i \in [E_{min}^{LTP}, E_{max}^{LTP}] \\ w_{ij}(t) & \text{otherwise} \end{cases} \tag{9}$$

The values defining the ranges in which LTD is triggered, E_{min}^{LTD} and E_{max}^{LTD}, and the rate constant for LTD, ϵ, were chosen to allow exactly one strong depression event as the result of a *cf*-stimulus following a *pf*-input. The values defining the ranges in which LTP is triggered, E_{min}^{LTP} and E_{max}^{LTP}, and the rate constant for LTP, η, were chosen to allow several weak potentiation events following a *pf*-input. As a result, *pf* stimulation alone leads to a weak net increase in the connection strength of the *pf–PU*-synapse, while *pf*-stimulation followed by *cf*-stimulation leads to a large net decrease. Thus LTD and LTP, respectively, decrease and increase the excitatory drive from *PU–SYN* to *PU–SO*.

Two response modes of the circuit

To support the learning mechanism outlined in the previous section the model needs to account for the acquisition of a pause in Purkinje cell activity following a CS, required to allow rebound excitation in the deep nucleus (see Fig. 1). For such a pause to occur, the tonic activity of Purkinje cells must be suppressed. In our model, *PU–SP* is suppressed by *I*, where *I* represents the inhibitory interneurons in the cerebellar cortex (see Fig. 2). One central assumption of our model is
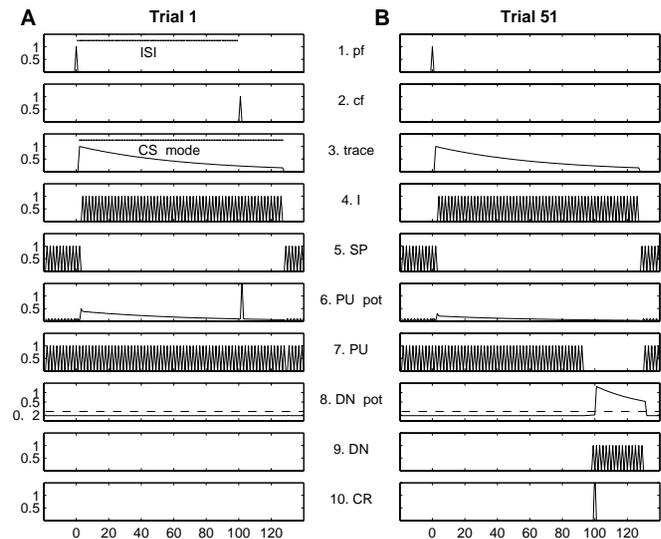

FIG. 3. Learning-related response changes in the model. The most relevant neural responses to a CS–US pair (ISI of 100 ts) are presented for (A) a trial before significant learning occurred and (B) when a correctly timed CR is expressed. A CS-related *pf*-signal (1) evokes a prolonged response in *PU–SYN*, the CS-trace (3). While there is an active CS-trace, *I* is active (4) and inactivates *SP* (5). In the CS-mode the membrane potential of *PU* (6) is driven by the relayed CS-input. Only in this CS-mode can synaptic changes be induced and a pause in *PU* spiking occur. (A) In Trial 1 the US-related *cf*-input (2) occurs while there is an active CS-trace (3), in this case following the CS-related *pf*-input (1) with an ISI of 100 ts. These conditions induce LTD (not illustrated in the figure). Because the *PU* membrane potential remains above spiking threshold, *PU* is active (7) and supplies constant inhibition to *DN* while in the CS-mode. Thus, *DN* cannot repolarise (8) and remains inactive (9) so that no CR is triggered (10). (B) Later in the experiment (Trial 51), the synaptic weight of the *pf–PU*-synapse has been reduced due to previous LTD. As a result, the *PU* potential (6) falls below the spiking threshold, which leads to a pause in *PU* spiking (7). *DN* membrane potential repolarises (8), so that rebound spikes are emitted (9). This rebound excitation triggers a CR (10). *DN* inhibition of *IO* prevents US-related *cf*-activity (2). Thus, although a US signal is still presented to the circuit, the reinforcing US pathway is blocked. These conditions induce LTP (not illustrated in the figure). The direction of change in the synaptic efficacy following a CS is thus dependent on the occurrence of a US as well as on the activity in the olivo-cortical-nuclear loop.

that the inhibitory effect of *I* onto *PU* following *pf*-activation is matched to the duration of the CS-trace, i.e. as long as there is ongoing activity at *PU–SYN*, *I* inhibits *PU–SP*. Whenever there is no active CS-trace the output of *PU* is determined by *PU–SP* and supplies tonic inhibition to *DN*. A *pf*-signal switches *PU* into a mode where its output is no longer driven by *PU–SP* but only by the CS-related input. Thus, the *PU* operates in two modes: a default, *spontaneous mode* and a *CS-mode*. Only in the CS-mode can synaptic changes be induced, can rebound excitation of the *DN* occur and thus can a CR be triggered. The properties of the model components are illustrated by cell traces before (Fig. 3A) and after (Fig. 3B) a CR is expressed. In summary, we based the design and implementation of our model on several assumptions on the functioning of cerebellar components proposed in the current literature. In simulated conditioning experiments and robot studies we evaluated our hypothesis that the combination of these assumptions supports associative learning

Simulated conditioning experiments

The aim of the simulated conditioning experiments was to understand how central circuit parameters influence the learning performance of
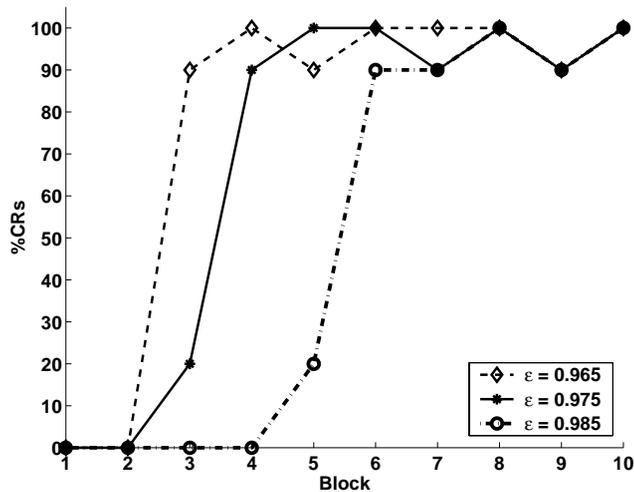
FIG. 4. The rate constant for LTD, ε, determines the speed of the acquisition of a CR. Comparison of the acquisition behaviour of circuits with values for ε of 0.965 (dashed line, diamonds), 0.975 (solid line, stars) and 0.985 (dashed–dotted line, circles). The percentage of CRs is plotted over 10 blocks of 10 CSs of an acquisition experiment for an ISI of 100 ts.

FIG. 5. The rate constant for LTP, η, determines the speed of the extinction of a CR. Comparison of the extinction behaviour of circuits with values for η of 0.0001 (dashed line, diamonds), 0.0002 (solid line, stars) and 0.0004 (dashed–dotted line, circles). Initially, the weight of the *pf–PU*-synapse was set to 0.2052 allowing for the expression of a CR for a training ISI of 100 ts. The percentage of CRs is plotted over 10 blocks of unreinforced CSs.

the model. The effect of the relative strength of LTD or LTP was studied varying the rate constant for LTD, ε, or the rate constant for LTP, η, respectively. Furthermore, circuits with different persistence of the CS-trace, i.e. varying in their value of *PU–SYN*, β^{SYN}, were tested. CS and US were represented as short activations ($A = 1$ for 1 simulation timestep) of *CSpathway* (CS pathway upstream of pons) and *USpathway* (US pathway upstream of inferior olive), respectively. Responses in *CRpathway* were counted as CRs if they occurred before the presentation of the US.

Acquisition experiments

An acquisition experiment consisted of 10 blocks of 10 trials each. In the first nine trials of a block the CS preceded a US with a fixed ISI (CS–US trials). In the last trial of a block the CS was not followed by a US (CS-alone trial).

The acquisition performance of the model is illustrated in Fig. 4 for an ISI of 100 simulation ts. The three curves allow a comparison of the learning performance of model circuits varying only in their value of the rate constant for LTD, ε. The acquisition curves of the three circuits show several common properties. None of the circuits elicits a CR within the first two blocks. Over the next few blocks there is a rapid increase in the percentage of elicited CRs. All circuits reach a maximum value of 100% CRs per block. Once this maximal value is reached the percentage of CRs per block fluctuates between 90 and 100%. The depicted curves approximately follow an S-shape similar to acquisition curves typically obtained in behavioural conditioning experiments (Bower & Hillgard, 1981).

When comparing the three graphs in detail it is apparent that the value of ε has a strong impact on the acquisition behaviour. The maximal percentage of CRs per block (and the first CR, not explicitly shown) is expressed earlier in the training procedure for lower values of ε. Thus, the lower the chosen value of ε, the stronger the LTD (see equation 8) and the faster the acquisition of a correctly timed CR.

Extinction experiments

In acquisition experiments using an ISI of 100 ts, we determined that a circuit elicits the maximal rate of CRs if the synaptic weight of the
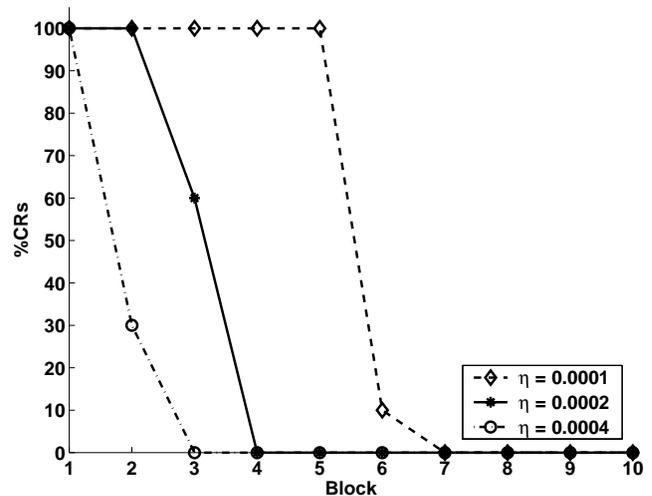
pf–PU-synapse has been reduced by LTD to 0.2052 (data not shown). In extinction experiments this value was used as initial synaptic weight and 10 blocks of 10 CS-alone trials were presented.

The extinction performance of the model is illustrated in Fig. 5 for an ISI of 100 ts. The three curves allow a comparison of the extinction behaviour for three model circuits varying only in their rate constant for LTP, η. Initially 100% CRs per block are expressed but, over a number of blocks, the percentage of CRs decreases significantly until no more CRs are elicited. When comparing the three graphs in detail it is apparent that the value of η strongly influences the extinction behaviour. The smaller the value of η, the later a block with less than 100% CRs occurs and the later the CR is completely abolished. Thus, the higher the chosen value of η, the stronger the LTP (see equation 9) and the faster the extinction of a CR.

Adaptation of PU response duration

The learning mechanism underlying acquisition (and extinction) can be studied in more detail by examining the adaptation of the *PU* response to a CS (Fig. 6). As discussed above (see Fig. 3), the persistence of the CS-trace, β^{SYN}, defines how long *PU* is in the CS-mode following *pf*-activity. When in the CS mode, *PU* in the naïve circuit constantly inhibits *DN* but training leads to a gradual shortening of the *PU* response duration until a pause in *PU* spiking allows for *DN* rebound excitation to trigger a CR. Thus, the acquisition and timing of the CR is reflected by changes in the *PU* response duration following the CS.

The change in *PU* response duration is illustrated for an acquisition experiment with a training ISI of 60 ts for three circuits varying only in their persistence of the CS-trace, β^{SYN} (Fig. 6A). The initial excitatory response duration to a CS is longer for higher values of β^{SYN} (i.e. it lasts 65, 127 and 187 ts for the circuits with a value of β^{SYN} of 0.970, 0.985 and 0990, respectively). All curves initially show a significant reduction in the *PU* response duration (label 1 in Fig. 6A). Subsequently the response duration stabilises (label 2). The three curves eventually converge to the same response duration, i.e.
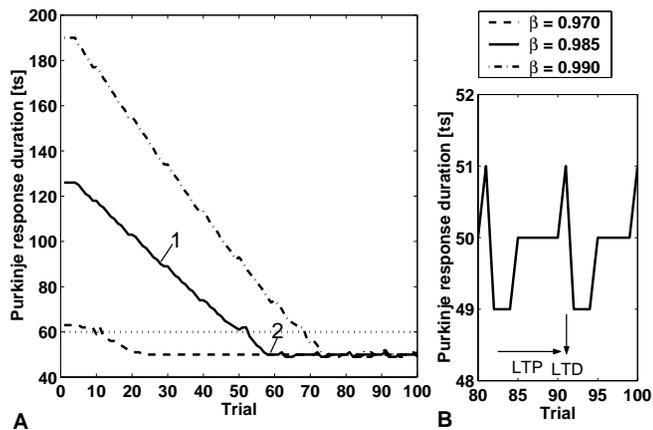
FIG. 6. Change of the *PU* response duration to a CS. (A) The excitatory response duration (in simulation timesteps, ts) is plotted over 100 trials of an acquisition experiment for an ISI of 60 ts (indicated with dotted line). Three circuits varying in the persistence of their CS-trace, i.e. β^{SYN} of 0.970 (dashed line), 0.985 (solid line) and 0.990 (dashed–dotted line), are compared. For explanation of labels 1 and 2 see text. (B) Close-up of fluctuations in *PU* response duration after a CR has been acquired ($\beta^{SYN} = 0.985$). See text for explanation.

50 ts, which allows rebound excitation of *DN* just before the US is presented. This reflects the fact that the same anticipatory CR timing is acquired by all three circuits.

Figure 6A shows a direct relationship between the initial response duration of *PU* to a CS and the number of trials required to elicit a correctly timed CR. The higher the persistence of the CS-trace, β^{SYN}, the longer the initial *PU* response duration to a CS and the more the synaptic weight has to be depressed before a CR can be triggered. Therefore, if all other parameters are kept constant, the higher the value of β^{SYN}, the more trials are needed before the response duration of *PU* is adapted to a given ISI.

After the acquisition of a correctly timed CR, the response duration of *PU* does not have a fixed value but fluctuates slightly (Fig. 6B). Whenever a CS–US pair is presented and no CR is expressed, *pf*- and *cf*-activation of *PU* coincide and trigger LTD. The resulting decrease in synaptic weight is reflected by a strong decrease in *PU* response duration (see arrow labelled LTD in Fig. 6B). However, if a CR is expressed, *DN* rebound excitation which triggers the CR also inhibits *IO*. By this negative feedback mechanism US-related *cf*-activation is prevented. Consequently, the CS–US pair does not lead to coincident *pf*- and *cf*-activation of *PU*, but to *pf*-activation alone. Thus, if a correctly timed CR is expressed, the induction properties for LTP, not those for LTD, are met. LTP results in a slight increase in the synaptic weight which is reflected by a gradual increase in *PU* response duration over trials (see arrow labelled LTP in Fig. 6B). Only when, after several trials, *DN* activity is too late to prevent US-related *cf*-activity, LTD is induced again. Thus, after a correctly timed CR has been acquired, there are regular fluctuations in CR latency caused by the ongoing interaction of LTD and LTP. This mechanism implies that a US cannot always be prevented although the CR timing is adapted overall to the training ISI.

Robot associative learning experiments

To analyse the performance of the model at the behavioural level, a mobile robot was interfaced to the circuit. The learning behaviour of the model was evaluated by observing the behaviour of the robot in an unsupervised obstacle avoidance task. In robot experiments CS and US occurrences were not controlled by an experimenter; the stimuli occurred purely as a result of the interaction of the learning system with its environment. In standard classical conditioning experiments, as well as in our simulated conditioning experiments, CS and US are presented in a fixed temporal relationship. It is unclear to what extent this constant ISI approximates the temporal relationship between CS and US under realistic conditions where fluctuations in their temporal coupling can be expected. Our robot experiments allowed us to investigate this issue.

Experimental set-up

The experiments were performed using a Khepera microrobot (K-team, Lausanne, Switzerland; Fig. 7A). The robot was equipped with a set of basic unconditioned reflexes triggered by stimulation of the (proximal) infra-red (IR) sensors. Activation of these sensors (US) due to the collision with an obstacle triggered a turn (UR) of $\approx 110°$ in the opposite direction. In addition the camera mounted on the robot constituted a distal sensor. In a separate process, cells with receptive fields much like complex cells in the visual cortex responded to certain spatial frequencies in the camera picture. The activity of a cell responding to a spatial frequency of ≈ 0.16 periods per degree signalled the initially neutral CS. Visual CSs and collision USs (stimulation of two frontal IR sensors) were conveyed to *CSpathway* and *USpathway* of the cerebellar circuit, respectively (each stimulus resulting in an activity of $A = 1$ for 1 ts). The activation of *CRpathway* (activity of $A = 1$ for 1 ts) triggered a specific motor CR, i.e. a strong ($\approx 150°$) turn.

The obstacle avoidance task

The robot was placed in a circular arena with a striped border (Fig. 7B) exploring its environment with a constant speed. Thus, when the behaviour is purely UR-driven, collisions with the wall of the arena occur regularly. A spatial frequency CS was detected at some distance when the robot approached the wall (Fig. 8A). Shortly afterwards, the collision with the wall stimulated the frontal IR sensors triggering a US (Fig. 8B). Hence, as in conditioning experiments, the CS was correlated with and predicted the US.

However, the ISIs of these stimuli were not constant but showed some variability under these uncontrolled conditions, e.g. due to noise in sensor sampling as well as the precise angle at which the robot approached the wall (Fig. 9). The aim of the robot studies was to determine whether the model circuit could support behavioural associative learning, i.e. reduce the number of collisions by adaptively timed CRs, under these less controlled, noisy, real-world conditions.

Alterations of the model circuit

Two elements needed to be added to alleviate shortcomings of the model under real-world conditions. Firstly, in the early learning phase, a CR could be expressed after a US had already been triggered. In previous simulation studies these long-latency responses would also occur, but were by definition excluded from the analysis. Their impact in robot experiments led to undesirable behaviour that was prevented by adding an inhibitory connection from *IO* to *CRpathway*.

Secondly, the response timing acquired at the circuit level had to be matched to the required timing at the behavioural level. To compensate for transduction delays from the occurrence of a CR signal in *CRpathway* of the cerebellar circuit to the completion of the CR (turning response) a delay of 15 ts was added to the inhibition from *DN* to *IO*. These examples illustrate how the real-world evaluation of a model may reveal limitations of a model not identified
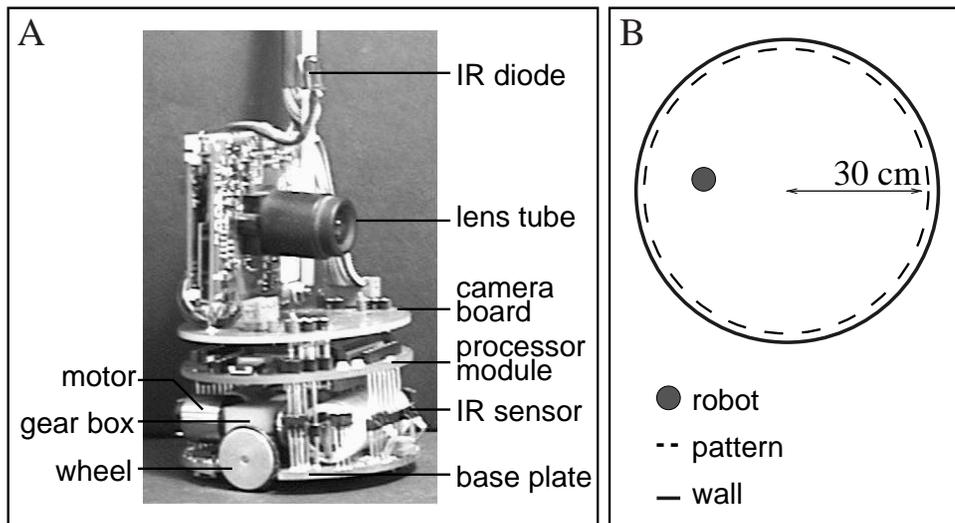
FIG. 7. The robot and its environment. (A) Khepera microrobot. The robot consists of a base plate with two motors and eight IR sensors, the processor module and an additional camera board carrying a colour camera. The robot has an diameter of 50 mm and a height of 80 mm. It was tethered to the host PC via a 2-m-long cable. To allow the recording of its trajectories, an IR diode was mounted on the robot. (B) The robot environment. The robot experiments were performed in a circular arena with a diameter of 60 cm, surrounded by a 15-cm-high Plexiglas wall. Paper with a pattern of vertical, equally sized black and white bars was placed against the wall. Patterns with bar widths of 9.5, 11 and 12.5 mm were used in the experiments.

in simulations. Possible analogous neural substrates are explored further in the Discussion.

Observable change in behaviour

Associative learning mediated by the cerebellar model significantly altered the robot's behaviour in the obstacle avoidance task (Fig. 10). At the beginning of the experiment, when the behaviour was determined by URs, the robot drove forward until it collided with the wall and only then performed a turn (Fig. 10A). Later in the experiment the robot usually turned just before it would collide with the wall so that regions close to the wall were avoided (Fig. 10B). Thus, associative learning reduced the number of collisions in the obstacle avoidance task. The robot performed significantly more successfully when exploiting the associative properties of the model than when relying purely on its URs.

The position of the robot when CSs, USs and CRs occurred in these two periods of the experiment can be examined at the circuit level (Fig. 10C and D). The first 10 CSs were all followed by a US which triggered a UR (Fig. 10C). While the first seven CSs did not trigger CRs, later CSs, i.e. CSs 8–10, elicited CRs at the circuit level. However, at the behavioural level these first CRs did not avoid the occurrence of a US. They were triggered with a long latency with respect to the CS (indicated by dotted lines), i.e. when the robot was already so close to the wall that a CR-turn could not prevent the collision. After 30 CS-events (Fig. 10D), all CSs triggered CRs. The latencies of these CRs with respect to the CSs were much shorter than those of the first CRs. Thus, most turns were performed some distance from the wall, preventing a subsequent collision. For those antici-patory responses the CS was not followed by a US. Note that the CR was not triggered immediately after a CS was detected, but rather the CR latency was adjusted with respect to the CS in such a way that the CR-turn just prevented the collision-US.

The timing of CRs

The increase in the performance supported by the model can be explained by the changes in ISI (or CS–US interval) and CR latency (or CS–CR interval) over the course of an experiment on a trial-by-
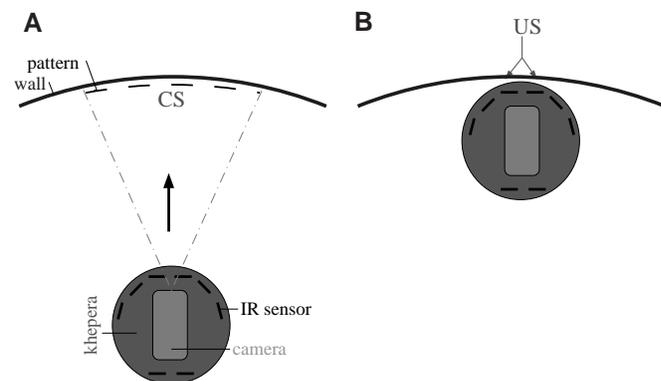

FIG. 8. Detection of CS and US. Eight IR sensors constitute the proximal sensors of the robot, the camera constituting the distal sensor. (A) A CS (a specific spatial frequency in the visual field of the camera) is detected a certain distance from the wall. (B) A US (the stimulation of the two frontal IR sensors) is detected directly at the wall.

trial basis (Figs 11 and 12). At the beginning of an experiment, when the behaviour of the robot was purely driven by URs, the CS was generally followed by the US and the ISIs ranged from 58 to 65 ts (Fig. 12A). The first CR was triggered with a long latency with respect to the CS, i.e. 63 ts. Because such long-latency CRs could not avoid the collision with the wall, the next few CSs were still followed by USs. As a result of LTD caused by these reinforced CSs, the CR latency further decreased, until most CRs prevented the occurrence of a US. Thus, associative learning lead to the expression of CRs, which prevented the occurrence of most USs. The strong decrease in CR latency at the beginning of the experiment was followed by slight fluctuations in the CS–CR interval (Fig. 11). If consecutive CSs were not followed by a US, there was a gradual increase in the CS–CR interval (Fig. 12B). This increase was due to the accumulative effect of LTP induced by several unreinforced CSs (as discussed for Fig. 6B). However, if a CS was followed by a US there was a strong

decrease in the CS–CR interval (Fig. 12C) due to the induced LTD. Thus, the CR latency was constantly adapted to the current experiences in the environment. The ISIs occurring later in the experiment were generally shorter than those in the beginning of the
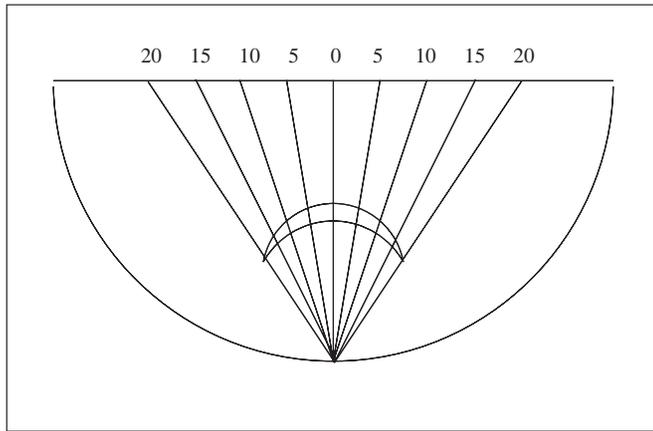

FIG. 9. Region in which a CS can be detected for the pattern with a bar width of 9.5 mm. Half of the environment is displayed; numbers indicate the distance (in cm) from the centre of the environment on the midline. Whether and where a CS can be detected (indicated by banana-shaped region) depends on the angle of approach of the robot to the wall. When approaching the wall with an angle steeper than $\approx 60°$ no CS is detected. The further from the wall the CS occurs, the longer the ISI.

experiment. The timing of the CRs was at first suitable to avoid USs for long ISIs, which occurred primarily when the robot drove perpendicularly toward the wall (i.e. if its trajectory formed an angle of $\approx 90°$ with the tangent of the arena wall; see Fig. 9). USs for shorter ISIs, i.e. those occurring with more shallow angles of approach, could only be avoided with lower synaptic weights and thus required more LTD. It was mostly the few remaining USs occurring under short ISIs that maintained the short CR latency via the induction of LTD.

Avoidance of USs

In Figs 11 and 12 all CRs are plotted that were triggered at the circuit level. Although after the rapid learning phase every CS generally elicited a CR, some CRs did not prevent the occurrence of a US. To obtain a more meaningful evaluation of the performance in the obstacle avoidance task, we introduced the measure of 'effective CRs', which includes only CRs that prevent the US occurrence (Fig. 13). In unconditioned reflex-driven behaviour, generally every CS is followed by a US. The percentage of effective CRs is equivalent to the percentage of USs that could be prevented compared to UR-driven behaviour. After the rapid learning phase the percentage of effective CRs per block fluctuated between 70 and 100%, meaning that between 7 and 10 out of 10 expected USs could be anticipated. Over the whole experiment $\approx 80\%$ of the USs expected for UR-driven behaviour were successfully anticipated by the model. As in simulated experiments, the learning curve is of S-shape, indicating a rapid initial learning phase followed by fluctuations around asymptotic performance.
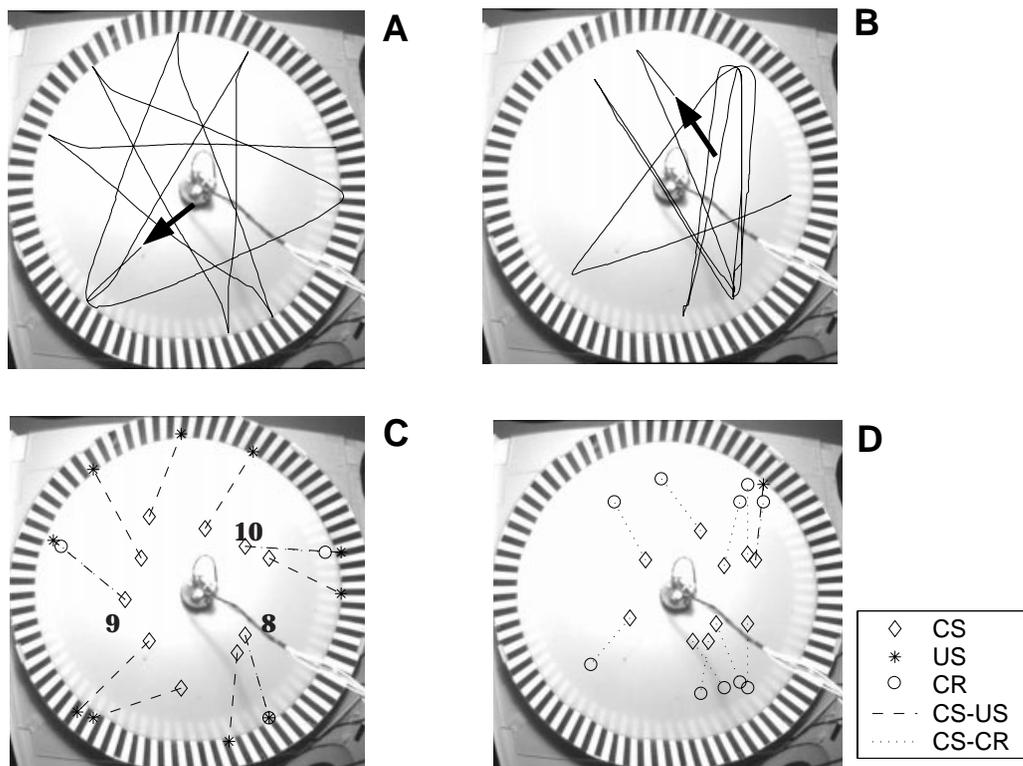

FIG. 10. Learning performance of the robot. (Top row) Trajectories of the robot, arrows indicating beginning of trajectories. (A) Beginning of the experiment (CS 1–10, occurring after ≈ 0–5 min). (B) Later in the experiment (CS 31–40, occurring after ≈ 15–20 min). (Bottom row) The same periods of the experiment examined at the circuit level. The symbols indicate the position of the robot where CSs (diamonds), USs (stars) and CRs (circles) occurred. CS–US pairs are connected by dashed lines, CS–CR pairs by dotted lines. (C) CS 1–10. (D) CS 31–40.
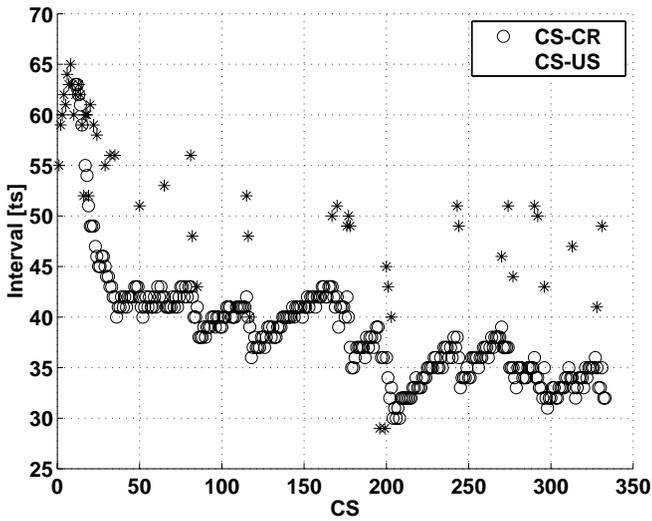
FIG. 11. Changes in the ISI and CR latency during an acquisition experiment. The experiment ran for 60 000 ts (≈ 120 min) using a bar width of 12.5 mm. The CS–US interval (ISI, stars) and the CS–CR interval (CR latency, circles) are measured in simulation timesteps (ts) and plotted over the 340 CSs that occurred in the experiment.
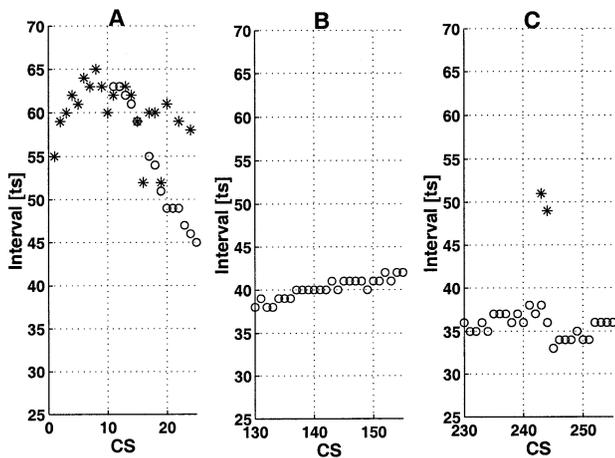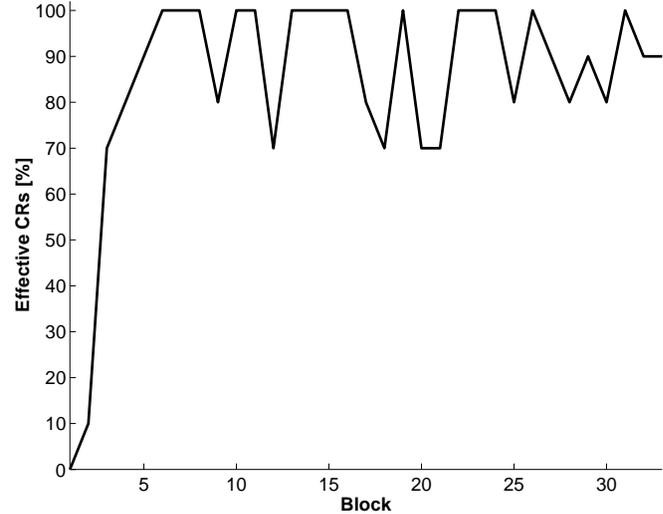

FIG. 13. Effective CRs (CRs that prevent the occurrence of a US). Percentage of effective CRs per block of 10 CSs during the experiment illustrated in Figs 11 and 12.
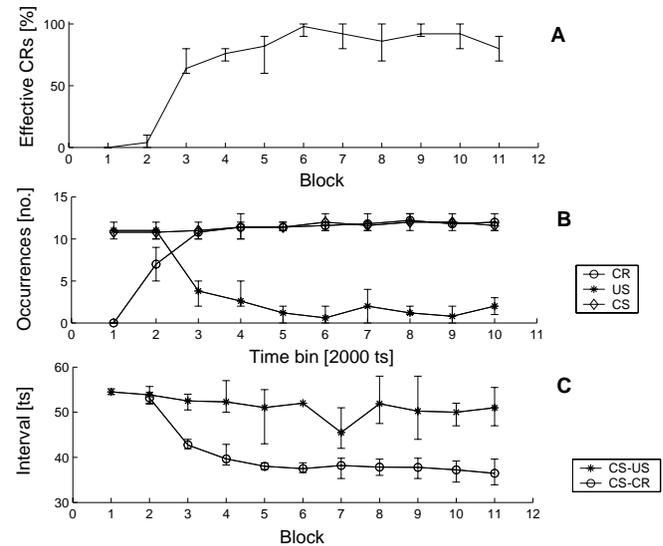

FIG. 12. Detailed plots of the same experimental data as depicted in Fig. 11. CS–US intervals (ISIs) are indicated with stars, CS–CR intervals (CR latencies) are indicated with circles. (A) Initial rapid-learning phase. (B) Period in which no USs occurred. (C) Effect of US occurrences.

FIG. 14. Average of five experiments to illustrate trends in learning behaviour. Experiments lasted 20 000 ts (≈ 40 min) in an environment with a bar width of 12.5 mm and led to a minimum of 110 CSs. Error bars give the minimum and maximum values of five data sets. (A) Average percentage of effective CRs (CRs that prevented the occurrence of a US) over 10 blocks of 10 CSs. (B) Number of CS occurrences (diamonds), US occurrences (stars) and all CR occurrences (circles) plotted over time bins of 2000 ts. (C) ISI (CS–US interval, stars) and CR latency (CS–CR interval, circles) over blocks of 10 CSs.

Generalisation of results

Simulated conditioning experiments are deterministic and give exactly the same learning curves when repeated. Successive runs of a robot experiment in the real world differ, in particular due to the variations that occur in the ISI distribution. To assess whether there were clear trends in the learning performance we studied the averaged data of five experiments (Fig. 14).

The average number of effective CRs over blocks (of 10 CSs) shows an initial fast learning phase followed by a stable phase with much higher values of the percentage of effective CRs (Fig. 14A). On average ≈ 75% of the USs can be anticipated after the rapid learning phase. The changes in the average CS, US and all CR occurrences over time similarly illustrate an increase in the number of CR occurrences followed by a decrease in the number of US occurrences

(Fig. 14B). After an initial rapid learning phase, every CS triggered a CR and USs occurred on average at an approximately constant rate which is again significantly lower than when behaviour was reflex-driven. The robot travelled less distance until the detection of the next CS because areas close to the wall were avoided. This caused a slight increase in the number of CS occurrences over time. To illustrate average changes in CR latency and ISI, the CS–CR interval and the CS–US interval were plotted over blocks of 10 CSs (Fig. 14C). As previously discussed, the average ISI decreased slightly over the
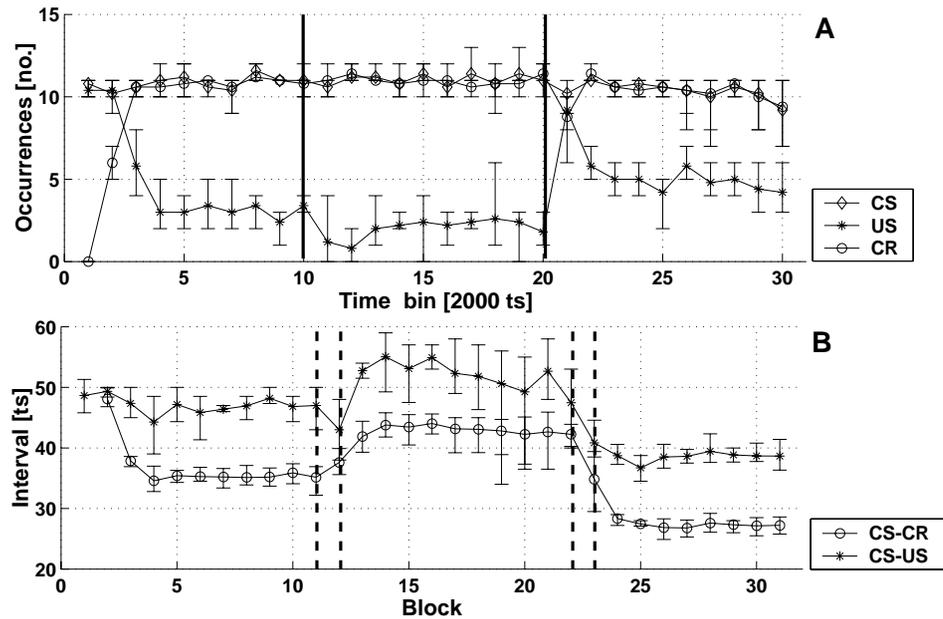
FIG. 15. Adaptation to changes in the ISI range. Experiments lasted 60 000 ts (≈ 120 min). Error bars give the minimum and maximum values of the five data sets. (A) Number of CS occurrences (diamonds), US occurrences (stars) and CR occurrences (circles) over time bins of 2000 ts. The pattern of the environment was exchanged twice (indicated by solid, vertical lines): from a bar width of 11 to 12.5 mm after 20 000 ts and from a bar width of 12.5 to 9.5 mm after 40 000 ts. (B) ISI (CS–US interval, stars) and CR latency (CS–CR interval, circles) over blocks of 10 CSs measured in simulation timesteps (ts). For all five experiments the transition of patterns took place between blocks 11 and 12 and blocks 22 and 23 (indicated by dashed vertical lines).

course of the experiment, because those USs that would occur for long ISIs could be avoided most effectively. The CR latency with respect to the CS decreased quickly at the beginning of the experiment, but stabilised after a few blocks. The final adapted CS–CR interval is ≈ 15 ts shorter than the ISI of the remaining CS–US pairs. This interval is equivalent to the time required from the occurrence of a CR at the circuit level to the performance of a complete CR turn. The averaged variables plotted show little variation (in their minimum and maximum values), indicating that the course and the overall effect of learning was similar for different runs of an experiment. Thus, the learning performance previously discussed for one long experiment (Figs 11–13) can be generalized.

Shifts in ISI distribution

The next experiment was designed to test whether the timing of CRs could be adapted to environmental changes that cause major shifts in occurring ISI ranges (Fig. 15). These ISI shifts were induced by changing the bar widths of the pattern of the arena border. The wider the bar width of the pattern, the further from the wall the spatial frequency CS is detected and thus the longer the occurring ISIs. In particular, bar widths of 9.5, 11 and 12.5 mm resulted in average ISIs of ≈ 39, 48 and 56 ts, respectively. A change from a bar width of 11 to 12.5 mm was used to cause a large increase in the average ISI after 20 000 ts. A change from a bar width of 12.5 to 9.5 mm resulted in a large decrease in the average ISI after 40 000 ts.

The average results of five experiments show three distinct plateau values in the average ISI resulting from the three different patterns used (Fig. 15B). After less than four blocks the CR latency was adapted to the new ISI range, i.e. ≈ 10–15 ts shorter than the average CS–US interval. The sudden increase in the average ISI after 20 000 ts resulted in a short-term decrease in the number of USs detected per time bin (Fig. 15A). The circuit was previously trained to shorter ISIs and thus, immediately after the pattern transition, CRs

were triggered early with respect to the new ISI. Thus the first CRs following the pattern transition were executed a significant distance from the wall, preventing almost all USs. However, the consecutive unreinforced CSs induced LTP. This lead to a slow increase in the CR latency with respect to the CS until again a stable number of USs occurred per time bin. In principle, this process is equivalent to a slow extinction of the previously acquired timing of a CR. If the ISIs suddenly decrease (after 40 000 ts), the learning process is equivalent to that during acquisition. Immediately after the pattern transition, the latency of the previously acquired CR is not sufficient to prevent the occurrence of USs, because CRs would be triggered too late. As a result, immediately after the second pattern transition the number of detected USs increased. However, reinforced CSs trigger LTD, leading to a decrease in the CR latency with respect to the CS. These results show that the continuous interaction of LTD and LTP leads to a circuit which can, within a few blocks of training, readapt the timing of the CR to major changes in the environment.

Discussion

We have presented a neural model which includes anatomical and physiological constraints of the cerebellar microcircuit while constituting a sufficiently reduced description to allow real-time simulations combined with real-world devices. In simulated classical conditioning experiments we have demonstrated that the model supports acquisition and extinction of accurately adaptively timed CRs. This performance is controlled by a limited number of model parameters, i.e. the rate constants for LTD and LTP and persistence of the CS-trace. Moreover, using a mobile robot we have demonstrated that the model microcircuit can support effective associative learning under less controlled, real-world conditions.

Models of cerebellar information processing

Since Marr (1969) and Albus (1971) proposed functional models of the cerebellum based on its unique anatomical organization more than 30 years ago, many computational models of cerebellar-mediated classical conditioning have been suggested (for a recent review see Medina & Mauk, 2000). These models range from algorithmic, functional top-down models (e.g. Moore & Choi, 1997) to detailed bottom-up models (e.g. Medina *et al.*, 2000; Fiala *et al.*, 1996). Our model constitutes a compromise between these two approaches, in that it represents an abstract and functional implementation of assumptions on cerebellar-mediated learning that are strictly based on biological findings. Our goal was to design a minimal model which extracts the principles of cerebellar information processing and to understand how certain aspects of the underlying learning mechanism affect the learning performance. Thus, the purpose of our studies was neither to mimic detailed dynamics or mechanisms of the underlying cellular responses as closely as possible nor to generate learning curves that match exactly those obtained in behavioural studies. This explains why we did not attempt to match simulation time to time measured in biological experiments.

Many of the assumptions embedded in our model have also been included in other models of cerebellar-mediated learning. Both bidirectional synaptic changes of the synapses formed between parallel fibres and Purkinje cells and negative feedback control of these synaptic changes are common to several recent models in this context (Spoelstra *et al.*, 2000; Medina *et al.*, 2000; Gluck *et al.*, 2001).

In particular, a model proposed by Barto and colleagues, which also focuses on a single Purkinje cell, shows at first glance several similarities to the model described in this paper (Barto *et al.*, 1999). Initially, Barto and Sutton applied reinforcement-learning algorithms established in the field of machine learning to models of classical conditioning (Barto & Sutton, 1982). In their more recent work, they attempt to match functional aspects of temporal difference algorithms to the cerebellar circuitry (Sutton & Barto, 1990; Barto *et al.*, 1999). As in our model, an error signal is signalled by the climbing fibres and drives a temporally asymmetric form of plasticity at the synapses formed between parallel fibres and Purkinje cells. Based on earlier work by Klopf (1982), this asymmetry is mediated by an eligibility trace in parallel fibres, similar to our CS-trace. However, while the amplitude of the eligibility trace in the model of Barto and colleagues defines the amount of synaptic change induced, the CS-trace in our model only determines whether the learning rules for LTD and LTP are activated. The model suggested by Barto and colleagues focuses at the cerebellar control of movements and includes the cerebellar interaction with a (formalised) premotor network. The model requires complex and predefined parallel fibre inputs which encode specific variables, such as the current position and resting position of the simulated limb. The inversion of the inhibitory output of Purkinje cells to a motor command signal is achieved by a cortico-rubro-cerebellar network while, in our model, this inversion is caused by the rebound depolarization in the deep nucleus. The model output is stabilised, not due to the activity in the negative feedback loop formed between inferior olive, cerebellar cortex and deep nucleus as in our model, but by bipolar climbing fibre error signals, i.e. positive and negative signals induce left- and rightward corrections of movements, respectively. In our model the climbing fibre signal is more simple, signalling the occurrence of an unpredicted event. Reinforcement models are popular among learning theorists, but a predefined discrete input space and a bipolar error signal are strong assumptions (Schultz *et al.*, 1997; Schultz & Dickinson, 2000). Our model shows that the cerebellar component of classical conditioning can be accounted for without making these assumptions.

Interaction of components

Although we did not include plasticity within the cerebellar nucleus in our model, our studies demonstrate that in order to understand cerebellar-mediated timing of CRs both cortical and nuclear components need to be taken into account. The expression of the CR depends on the rebound excitation in *DN*. The timing of this *DN* response is defined by the onset of the acquired pause in *PU* activity following CS presentation which is modulated by LTD and LTP of the *pf–PU*-synapse. However, in the absence of *DN–IO* inhibition, these synaptic changes would not lead to stable CR timing. Presentation of reinforced CSs after the acquisition of a correctly timed CR would continue to induce LTD and further reduce the CR latency. Hence, to stabilize the correct duration of the acquired pause in *PU* activity, LTD must be prevented once a correctly timed CR is expressed. This is achieved by the negative feedback from *DN* to *IO* because LTP is induced if US-related *cf*-activity is prevented. Thus, the negative feedback loop formed between microcomplexes of interconnected groups of cells in the inferior olive, Purkinje cells and cells in the deep nucleus is crucial for the stabilization of the CR timing.

Implication of the learning mechanism

The implemented learning mechanism is based on the continuous interaction of LTD and LTP. This leads to two important features of the asymptotic learning performance of the model. Firstly, while the timing of CRs is in general adapted to the training ISI, there will always be slight fluctuations. Secondly, some USs will intermittently be anticipated incorrectly. If the synaptic weight changes due to LTD are relatively strong in comparison to LTP, learning is fast and few CSs will fail to elicit a CR. However, this also means that extinction of a previously learned CR and readaptation to long ISIs can only be acquired slowly. Thus, there is a trade-off between the adaptability of the circuit and the stability of its performance.

A functional analysis of model elements

Our modelling approach allows us to ascribe functional roles to certain model elements. In our model, *cf*-signals the occurrence of a prediction error to *PU* (as suggested in Bloedel & Bracha, 1998; Schultz & Dickinson, 2000), while activity in the deep nucleus can be interpreted as an internal prediction of the occurrence of a US. The integration of signals pertaining to the prediction and occurrence of the US takes place in the *IO*. It has been argued that such integration of inhibitory deep nucleus and excitatory US-related signals could occur at the level of the olivary glomeruli (De Zeeuw *et al.*, 1998). This is in accord with findings that US-evoked activity in the inferior olive decreases during the acquisition of a CR and increases again following the extinction of a CR (Hesslow & Ivarsson, 1996; Sears and Steinmetz, 1991; Medina *et al.*, 2002). LTP and LTD serve functionally as a continuous check for the current internal US prediction and can only constitute a highly adaptive functional learning mechanism under the control of the nuclear-olivo-cortical feedback loop (Hesslow and Ivarsson, 1996). The control of the reinforcing signal by the negative feedback from the deep nucleus to the inferior olive could explain why cerebellar-mediated classical conditioning does not obey the law of effect (Thorndike, 1898; Gardner & Gardner, 1988). Because the US-related signal is blocked by deep nuclear inhibition of the inferior olive, the timing of an acquired eye-blink CR is stabilised not only when the behavioural CR actively prevents the aversive US (i.e. eyelid closure to avoid airpuff

US) but also if the US cannot be avoided (i.e. electric shock US) (Gormezano *et al.*, 1983).

Associative learning vs. simulated conditioning experiments

The strength of our approach lies in the possibility of elucidating how the suggested underlying mechanisms at different levels of description could interact and scale up to behavioural learning. Models of cerebellar learning have been evaluated using real-world devices in the past (Van der Smagt, 2000), e.g. to study trajectory planning (Kawato, 1999) or smooth pursuit (Huang *et al.*, 2000; Coenen *et al.*, 2001). However, in these studies, the cerebellar circuitry modulates motor commands that are generated elsewhere and parallel fibre signals carry complex, prewired information. In our model of cerebellar-mediated classical conditioning, the cerebellum receives simple CS representations and triggers the expression of motor responses. This is in accord with data demonstrating that motor CRs can be learned in decerebrated animals (Mauk & Thompson, 1987).

Associative learning, when CS and US occurrences arise from the interaction of the learning system with its environment, differs in at least two important aspects from (simulated) classical conditioning experiments. Firstly, the ISIs vary. This variability can be attributed to the interaction of the robot with its environment (i.e. the angle at which the robot approached the wall), to intrinsic properties of the behaving system (e.g. noise in sensor sampling), and to the learning process itself. In the simulated classical conditioning experiments, the required CR timing was well defined, i.e. the CS has to trigger a CR with a latency just smaller than the ISI. However, in robot studies, successful learning was defined by the avoidance of collision USs and, as a result of unpredictable ISI variations, the required CR timing was less clearly defined. Secondly, in simulation studies, the reinforcing *cf*-signal is entirely controlled by the inhibition of *IO* by *DN*. However, in robot studies the expression of a CR turn may, in addition, actively prevent the occurrence of a collision US at the behavioural level. Despite these differences in the learning conditions, the learning behaviour of the circuit in robot experiments is in many ways similar to that observed in simulated conditioning experiments. There is an initial rapid learning phase after which the performance reaches asymptotic levels. However, in robot studies, strong and irregular fluctuations in the CR timing can be observed. These fluctuations illustrate how the internal US prediction is constantly adapted to recent experiences in the environment. This allows the adaptation of the CR timing to minor variability or noise and to major shifts in the ISI distribution. A gradual decrease in the ISIs on nonsuccessful trials was observed over time. This can be understood in terms of the learning mechanisms implemented in our model. Due to the gradual acquisition of a pause in *PU* activity, at first only US occurrences with long ISIs can be prevented. Hence, only after these events are captured by the learning system can shorter ISIs be acquired. In summary, our robot experiments demonstrate that the mechanisms of cerebellar conditioning in the context of the nuclear-olivo-cortical loop support robust and adaptive learning in a real-world system which imposes variable ISIs. The adaptation of the CR latency is predominantly driven by short-latency USs that were not anticipated correctly.

Alterations to the model circuit

In order to adapt our model of the cerebellar microcircuit to the performance in the real world, two functional elements had to be added. Firstly, we observed that in the initial learning phase CRs could be expressed after a US had already triggered a UR. However, animal studies indicate that the expression of CR and UR may overlap (Gormezano *et al.*, 1983), but not that a CR occurs after a

UR. This suggests that the cerebellar microcircuit and/or its periphery includes mechanisms that prevent the expression of a CR after a UR. In our model this issue was solved functionally by including an inhibitory connection from *IO* to *CRpathway* preventing the expression of CRs after US-related *cf*-activation. This solution is in accord with the notion that climbing fibre activity may prevent rebound excitation of cells in the deep nucleus (Hesslow, 1994b). In particular, it has been shown that single-pulse cortical stimulation following a train stimulation can inhibit the expression of a CR. Although a complex spike initially triggers a short temporary pause in simple spike activity, this pause is, in most cells, followed by a significant increase in simple spike activity (Sato *et al.*, 1992). It remains to be studied whether climbing fibre activation could thereby interfere with rebound depolarization of cells in the deep nucleus, before a Na$^+$ burst is elicited.

Secondly, when examining the performance of the circuit at the behavioural level, transduction delays between the robot and the model circuit had to be taken into account. In particular, there is a significant delay between the CR triggered at the level of *DN* and the action of the effectors, i.e. the complete execution of a motor CR. Similarly, physiological recordings show that activity of cells in the deep nucleus occurs ≈ 40–60 ms before a behavioural CR is observed (McCormick & Thompson, 1984). Hence, the internal timing acquired at the circuit level has to be matched to the peripheral timing required at the behavioural level. In our model this was achieved functionally by adding a delay in the projections from *DN* to *IO*. Spoelstra *et al.* (2000) also added a delay to the inhibition of the deep nucleus to the inferior olive in their model of cerebellar-mediated control of movements to align the cerebellar output with the error signal within the olivo-cortico-nuclear loop. What the biological substrate for such a delay could be is unclear and needs further experimental analysis.

Two modes of operation

One critical assumption embedded in our model is the operation of *PU* in two modes: a default spontaneous mode and a CS-mode. During the spontaneous mode, *DN* is tonically inhibited by the spontaneous activity of *PU*, while CS-related input drives *PU* output in the CS-driven mode. Recently, Thompson and associates studied the effect of two different types of Purkinje cell inhibition onto cells in the deep nucleus, namely basal and stimulus-activated inhibition (Bao *et al.*, 2002). The learning mechanism embedded in our model is in good accord with their findings that basal inhibition modulates CR expression, while CS-activated inhibition is required for the proper timing of the CR. While there is no physiological evidence that Purkinje cells may switch between two modes of operation, dual response properties have recently been described for other cells, such as deep nuclear cells (Aizenman *et al.*, 1998; Aizenman & Linden, 1999) and thalamic relay cells (Sherman, 2001). It remains to be studied whether the activity of cortical interneurons or some intrinsic mechanism could mediate such a switch in Purkinje cell operation following a CS.

Future work

Our model focused on the description of a single microcircuit in a microcomplex. In future work we will investigate how a collection of these microcircuits can constitute a complete microcomplex. This requires that the peripheral signals feeding into a microcomplex are more explicitly modelled (Hansel *et al.*, 2001). In our current model a representation of the CS is readily available to the circuit. However, under natural learning conditions such a representation of the CS may first have to be acquired or enforced. In the study of classical

conditioning two distinct phases of learning are distinguished (Lennartz & Weinberger, 1992, Mintz & Wang-Ninio, 2001). In a first, fast, nonspecific phase an emotional response is acquired by a process involving the amygdala. In a second, slow, phase discrete motor responses specific to the US are acquired in the cerebellum. We have proposed that the early nonspecific learning phase serves to identify a CS (Verschure & Voegtlin, 1999). For instance, it has been shown that the primary auditory cortex quickly enlarges the representation of tones that are paired with US-derived signals (Kilgard & Merzenich, 1998). We have previously presented a biophysically detailed model that can account for this aspect of auditory conditioning (Sanchez-Montanes *et al.*, 2000). In future work we will interface our model of CS identification with the model presented here to realize a more complete account of the neuronal systems implementing the two-phase theory of classical conditioning.

Acknowledgements

Abbreviations

CR, conditioned response; CS, conditioned stimulus; IR, infra-red; ISI, inter-stimulus interval; LTD, long-term depression; LTP, long-term potentiation; NM(R), nictitating membrane (response); pf, parallel fibre; pf–PU synapse, synapse formed between parallel fibre and Purkinje cell; PU, Purkinje cell; ts, simulation timesteps; UR, unconditioned response; US, unconditioned stimulus.

References

Aizenman, C.D. & Linden, D.J. (1999) Regulation of the rebound depolarization and spontaneous firing patterns of deep nuclear neurons in slices of rat cerebellum. *J. Neurophysiol.*, **82**, 1697–1709.

Aizenman, C.D., Manis, R. & Linden, D.J. (1998) Polarity of long-term synaptic gain change is related to postsynaptic spike firing at a cerebellar inhibitory synapse. *Neuron*, **21**, 827–835.

Albus, J.S. (1971) A theory of cerebellar function. *Math. Biosci.*, **10**, 25–61.

Attwell, P.J.E., Rahman, S. & Yeo, C.H. (2001) Acquisition of eyeblink conditioning is critically dependent on normal function in cerebellar cortical lobule HVI. *J. Neurosci.*, **21**, 5715–5722.

Barto, A.G. & Sutton, R.S. (1982) Simulation of anticipatory responses in classical conditioning by a neuron-like adaptive element. *Behav. Brain Res.*, **4**, 221–235.

Bao, S., Chen, L., Kim, J.J & Thompson, R.F. (2002) Cerebellar cortical inhibition and classical eyeblink conditioning. *Proc. Natl Acad. Sci. USA*, **99**, 1592–1597.

Barto, A.G., Fagg, A.H., Sitkoff, N. & Houk, J.C. (1999) A cerebellar model of timing and prediction in the control of reaching. *Neural Computation*, **11**, 565–594.

Bernardet, U., Blanchard, M. J. & Verschure, P.F.M.J. (2001) IQR: a distributed system for real-time real-world neuronal simulation. *Neurocomputing, 44–46*, 1043–1048.

Bloedel, R.J. & Bracha, V. (1998) Current concepts of climbing fiber function. *Anat. Rec.*, **253**, 118–126.

Bower, G.H. & Hillgard, E.R. (1981) *Theories of Learning*, 5th edn. Prentice Hall Inc., Englewood Cliffs, NJ.

Braitenberg, A. & Atwood, R.P. (1958) Morphological observations on the cerebellar cortex. *J. Comp. Neurol.*, **109**, 1–27.

Chen, C. & Thompson, R.F. (1995) Temporal specificity of long-term depression in parallel fiber-Purkinje synapses in rat cerebellar slice. *Learn. Mem.*, **2**, 185–198.

Coenen, O.J., Arnold, M.P., Jabrin, M.A. & Sejnowski, T.J. (2001) Parallel fiber coding in the cerebellum for life-long learning. *Autonomous Robots*, **11**, 291–297.

De Zeeuw, C.L., Simpson, J.L., Hoogenraad, C.G., Gaqjart, N., Kaekkoek,

S.K.E. & Ruigrok, T. (1998) Microcircuitry and function of the inferior olive. *Trends. Neurosci.*, **21**, 391–400.

Fiala, J.C., Grossberg, S. & Bullock, D. (1996) Metabotropic glutamate receptor activation in cerebellar Purkinje cells as substrate for adaptive timing of the classically conditioned eye-blink response. *J. Neurosci.*, **16**, 3760–3774.

Gardner, R.A. & Gardner, B.T. (1988) Feedforward versus feedback: An ethological alternative to the law of effect. *Behav. Brain. Sci.*, **11**, 429–493.

Gauk, V. & Jaeger, D. (2000) The control of rate and timing of spikes in the deep nuclei by inhibition. *J. Neurosci.*, **20**, 3006–3016.

Gluck, M., Allen, M.T., Myers, C.E. & Thompson, R.F. (2001) Cerebellar substrates for error correction in motor conditioning. *Neurobiol. Learn. Mem.*, **76**, 314–341.

Gormezano, L., Kehoe, E.J. & Marshall, B.S. (1983) Twenty years of classical conditioning with the rabbit. *Prog. Psychobiol. Physiol. Psychol.*, **10**, 197–263.

Gruart, A., Guillazo-Blanch, G., Fernandez-Mas, R., Jimenez-Diaz, L. & Delgado-Gaxcia, J.M. (2000) Cerebellar posterior interpositus nucleus as an enhancer of classically conditioned eyelid responses in alert cats. *J. Neurophysiol.*, **84**, 2680–2690.

Hansel, C., Linden, D.J. & D'Angelo, E. (2001) Beyond parallel fiber LTD: the diversity of synaptic and nonsynaptic plasticty in the cerebellum. *Nature Neurosci.*, **4**, 467–475.

Hartell, N.A. (2002) Parallel fiber plasticity. *Cerebellum*, **1**, 3–18.

Hesslow, G. (1994a) Correspondence between climbing fibre input and motor output in eyeblink-related areas in the cat cerebellar cortex. *J. Physiol. (Lond.)*, **476**, 229–244.

Hesslow, G. (1994b) Inhibition of classically conditioned eyeblink responses by stimulation of the cerebellar cortex in the decerebrate cat. *J. Physiol. (Lond.)*, **476**, 245–256.

Hesslow, G. & Ivarsson, M. (1994) Suppression of cerebellar Purkinje cells during conditioned responses in. *Neuroreport*, **5**, 649–652.

Hesslow, G. & Ivarsson, M. (1996) Inhibition of the inferior olive during conditioned response in the decerebrate ferret. *Exp. Brain Res.*, **110**, 36–46.

Hesslow, G., Svensson, P. & Ivarsson, M. (1999) Learned movements elicited by direct stimulation of cerebellar mossy fiber afferents. *Neuron*, **24**, 179–185.

Houk, J.C., Buckingham, J.T. & Barto, A.G. (1996) Models of the cerebellum and motor learning. *Behav. Brain Sci.*, **19**, 368–383.

Huang, J., Jabri, M.A., Coenen, O.J. & Sejnowski, T.J. (2000) Models of basal ganglia and cerebellum for sensimotor integration and predictive control. 7–8 September 2000, Humanoids 2000: First IEEE-RAS International Conference on Humanoid Robotics, MIT, MA, USA.

Hull, C.L. (1939) The problem of stimulus equivalence in behavior theory. *Psycholog. Rev.*, **46**, 9–30.

Ito, M. (1984) The modifiable neuronal network of the cerebellum. *Jpn. J. Physiol.*, **5**, 781–792.

Ito, M. (2001) Cerebellar long-term depression: Characterization, signal transduction, and functional roles. *Physiol. Rev.*, **81**, 1143–1195.

Ito, M. & Kano, M. (1982) Long-lasting depression of parallel fiber-Purkinje cell transmission induced by conjunctive stimulation of parallel fibers and climbing fibers. *Neurosci. Lett.*, **33**, 253–258.

Kawato, M. (1999) Internal models for motor control and trajectory planning. *Curr. Opin. Neurobiol.*, **9**, 718–727.

Kilgard, M.P. & Merzenich, M.M. (1998) Cortical map reorganization enabled by nucleus basalis activity. *Science*, **279**, 1714–1718.

Kim, J.J., Krupa, D.J. & Thompson, R.F. (1998) Inhibitory cerebello-olivary projections and blocking effect in classical conditioning. *Science*, **279**, 570–573.

Kim, J.J. & Thompson, R.F. (1997) Cerebellar circuits and synaptic mechanisms involved in classical eyeblink conditioning. *Trends Neurosci.*, **20**, 177–181.

Kistler, W.M. & van Hemmen, J.L. (1999) Delayed reverberation through time windows as a key to cerebellar funtion. *Biol. Cybern.*, **81**, 373–380.

Kitazawa, S. (2002) Neurobiology: Ready to unlearn. *Nature*, **416**, 270–273.

Klopf, A. (1982) The Hedonistic Neuron: A theory of memory, learning and intelligence. Hemisphere, Washington DC.

Krupa, D.J. & Thompson, R.F. (1997) Reversible inactivation of the cerebellar interpositus nucleus completely prevents acquisition of the classically conditioned eye-blink response. *Learn. Mem.*, **3**, 545–556.

Kuroda, S., Yamamoto, K., Miyamoto, H., Doya, K. & Kawato, M. (2001) Statisttical characteristics of climbing fiber spikes necessary for efficient cerebellar learning. *Biol. Cybern.*, **84**, 183–192.

Lavond, D.G., Kim, J.J., & Thompson, R.F. (1993) Mammalian brain

substrates of aversive classical conditioning. *Annu. Rev. Psychol.*, **44**, 317–342.

Lennartz, R.C. & Weinberger, N.M. (1992) Analysis of response systems in Pavlovian conditioning reveals rapidly versus slowly acquired conditioned responses: support for two factors, implications for behavior and neurobiology. *Psychobiology*, **20**, 93–119.

Linden, D.J. & Ahn, S. (1991) Activation of the presynaptic cAMP-dependent protein kinase is required for induction of cerebellar long-term potentiation. *J. Neurosci.*, **19**, 10221–10227.

Linden, D.L. & Conner, J.A. (1995) Long-term synaptic depression. *Annu. Rev. Neurosci.*, **18**, 319–357.

Llinas, R.R., Lang, E.J. & Welsh, J.P. (1997) The cerebellum, LTD, and memory: Alternative views. *Learn. Mem.*, **3**, 445–455.

MacKay, W.A. & Murphy, J.T. (1976) Integrative versus delay line characteristics of cerebellar cortex. *Can. J. Neurol. Sci.*, **3**, 85–97.

Mackintosh, N.J. (1974) *The Psychology of Animal Learning*. Academic Press, London.

Marr, D. (1969) A theory of cerebellar cortex. *J. Neurophysiol.*, **202**, 437–470.

Mauk, M.D. (1997) Roles of cerebellar cortex and nuclei in motor learning: Contradictions or clues? *Neuron*, **18**, 343–346.

Mauk, M.D. & Thompson, R.F. (1987) Retention of classically conditioned eyelid responses following acute decerebration. *Brain Res.*, **403**, 89–95.

Mauk, M.D., Garcia, K.S., Medina, J.F. & Steele, P.M. (1998) Does cerebellar LTD mediate motor learning? Towards a resolution without a smoking gun. *Neuron*, **20**, 359–362.

McCormick, D.A. & Thompson, R.F. (1984) Neuronal response of the rabbit cerebellum during acquisition and performance of a classically conditioned nictitating membrane-eyelid response. *J. Neurosci.*, **4**, 2811–2822.

Medina, J.F., Garcia, K.S., Nores, W.L., Taylor, X.M. & Mauk, M.D. (2000) Timing mechanisms in the cerebellum: Testing predictions of a large-scale computer simulation. *J. Neurosci.*, **20**, 5516–5525.

Medina, J.F. & Mauk, M.D. (2000) Computer simulation of cerebellar information processing. *Nature Neurosci.*, **19**, 1205–1211.

Medina, J.F., Nores, W.L. & Mauk, M.D. (2002) Inhibition of climbing fibres is a signal for the extinction of conditioned eyelid responses. *Nature*, **416**, 330–333.

Mintz, M. & Wang-Ninio, Y. (2001) Two-stage theory of conditioning: involvement of the cerebellum and the amygdala. *Brain Res.*, **897**, 150–156.

Miyata, M., Finch, E.F., Khiroug, L., Hashimoto, K., Hayasaka, S., Oda, S.L., Inouye, X.I., Takagishi, Y., Augustine, G.J. & Kano, M. (2000) Local calcium release in dendritic spines required for long-term synaptic depression. *Neuron*, **28**, 233–244.

Montague, P.R. & Sejnowski, T.J. (1994) The predictive brain: temporal coincidence and temporal order in synaptic learning mechanism. *Learn. Mem.*, **1**, 1–33.

Moore, J.W. & Choi, J.S. (1997) Conditioned response timing and integration in the cerebellum. *Learn. Mem.*, **4**, 116–129.

Ohyama, T. & Mauk, M.D. (2001) Latent acquisition of timed responses in cerebellar cortex. *J. Neurosci.*, **21**, 682–690.

Pavlov, I.P. (1927) *Conditioned Reflexes*. Oxford University Press, Oxford.

Pavlov, I.P. (1928) *Lectures on Conditioned Reflexes: Twenty-Five Years of Objective Study of the Higher Nervous Ability (Behavior) of Animals*. International Publishers, New York.

Perret, S.P. (1998) Temporal discrimination in the cerebellar cortex during conditioned eyelid responses. *Exp. Brain. Res.*, **121**, 115–124.

Perrett, S.P. & Mauk, M.D. (1995) Extinction of conditioned eyelid responses requires the anterior lobe of cerebellar cortex. *J. Neurosci.*, **15**, 2074–2080.

Perrett, S.P., Ruiz, B.R. & Mauk, M.D. (1993) Cerebellar cortex lesions disrupt learning-dependent timing of conditioned eyelid responses. *J. Neurosci.*, **13**, 1708–1718.

Raman, I.M. & Bean, B.P. (1999) Ionic currents underlying spontaneous action potentials in isolated cerebellar Purkinje neurons. *J. Neurosci.*, **19**, 1663–1674.

Raymond, J.L., Lisberger, S.G. & Mauk, M.D. (1996) The cerebellum: a neuronal learning machine. *Science*, **272**, 1126–1130.

Rogers, R.F., Britton, G.B. & Steinmetz, J.E. (2001) Learning-related interpositus activity is conserved across species as studied during eyeblink conditioning in the rat. *Brain. Res.*, **905**, 171–177.

Salin, P.A., Malenka, R.G. & Roger, A. (1996) Cyclic AMP mediates a presynaptic form of LTP at cerebellar parallel fiber synapses. *Neuron*, **16**, 797–803.

Sanchez-Montanes, M.A., Verschure, P.F. & Koenig, P. (2000) Local and global gating of synaptic plasticity. *Neural. Comput.*, **12**, 519–529.

Sato, Y., Miura, A., Fushiki, H. & Kawasaki, T. (1992) Short-term modulation of cerebellar Purkinje cell activity after spontaneous climbing fiber input. *J. Neurophysiol.*, **68**, 2051–2062.

Schreurs, B.G., Oh, M.M. & Alkon, D.L. (1996) Pairing-specific long-term depression of Purkinje cell excitatory postsynaptic potentials results from a classical conditioning procedure in the rabbit cerebellar slice. *J. Neurophysiol.*, **75**, 1051–1060.

Schreurs, B.G., Pavel, A.G., Tomsic, D., Alkon, D.L. & Shi, T. (1998) Intracellular correlates of acquisition and long-term memory of classical conditioning in Purkinje cell dentrites in slices of rabbit cerebellar lobule HVI. *J. Neurosci.*, **16**, 5498–5507.

Schultz, W., Dayan, P. & Montague, P.R. (1997) A neural substrate of prediction and reward. *Science*, **275**, 1593–1599.

Schultz, W. & Dickinson, A. (2000) Neuronal coding of prediction errors. *Annu. Rev. Neurosci.*, **23**, 473–500.

Sears, L.L. & Steinmetz, J.E. (1991) Dorsal accessory inferior olive activity diminishes during acquisition of the rabbit classically conditioned eyelid response. *Brain Res.*, **545**, 114–122.

Sherman, S.M. (2001) Tonic and burst firing: dual modes of thalamocortical relay. *Trends Neurosci.*, **24**, 122–126.

Shibuki, K. & Okada, D. (1992) Cerebellar long-term potentiation under suppressed postsynaptic Ca^{2+} activity. *Neuroreport*, **3**, 231–234.

Spoelstra, 1., Schweighofer, N. & Arbib, M.A. (2000) Cerebellar learning of accurate predicive control for fast-reaching movements. *Biol. Cybern.*, **82**, 312–333.

Steinmetz, J.E., Lavond, D.G. & T.R. (1989) Classical conditioning in rabbits using pontine nucleus stimulation as a conditioned stimulus and inferior olive stimulation as an unconditioned stimulus. *Synapse*, **3**, 225–233.

Steinmetz, J.E., Rosen, D.J., Chapman, P.F., Lavond, D.G. & Thompson, T.R. (1986) Classical conditioning of the rabbit eyelid response with a mossy-fiber stimulation CS: I. pontine nuclei and middle cerebellar peduncle stimulation. *Behav. Neurosci.*, **10**, 878–887.

Sutton, R.S. & Barto, A.G. (1990) Time-derivative models of Pavlovian reinforcement. In Editor, A. & Editor, B. (eds), *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, Chapter 12. MIT press, Location, pp 497–537.

Svensson, P. & Ivarsson, M. (1999) Short-lasting conditioned stimulus applied to the middle cerebellar peduncle elicits delayed conditioned eye blink responses in the decerebrate ferret. *Eur. J. Neurosci.*, **11**, 4333–4340.

Thompson, R.F., Berger, T.W. & Madden, I.V. (1983) Cellular processes of learning and memory in the mammalian CNS. *Annu. Rev. Neurosci.*, **6**, 447–491.

Thompson, R.F., Thompson, J.K., Kim, J.J, Krupa, D.J & Shinkman, P.G. (1998) The nature of reinforcement in cerebellar learning. *Neurobiol. Learn. Mem.*, **70**, 150–176.

Thorndike, E.L. (1898) Animal intelligence: an experimental study of the associative process in animals. *Psychol. Monogr.*, **2**.

Van der Smagt, P. (2000) Benchmarking cerebellar control. *Robotics and Autonomous Systems*, **32**, 237–251.

Verschure, P.F.M.J. (1998) Synthetic epistomology: The acquisition, retention and expression of knowledge in natural and synthetic systems. In *Proceedings of the 1998 World Congress on Computational Intelligencee, WCCI98*, IEEE Press, Anchorage, Alaska, pp. 147–153.

Voegtlin, T. & Verschure, P.F.M.J. (1999) What can robots tell us about brains? A synthetic approach towards the study of learning and problem solving. *Rev. Neurosci.*, **10**, 291–310.

Wang, S.S.H, Denk, W. & Häusser, M. (2000) Coincidence detection in single dendritic spines mediated by calcium release. *Nature Neurosci.*, **3**, 1266–1273.

Webb, B. (2000) What does robotics offer animal behaviour? *Anim. Behav.*, **60**, 545–558.

Welsh, J.P. & Harvey, J.A. (1989) Cerebellar lesions and the nictitating membrane reflex: Performance deficits of the conditioned and unconditioned response. *J. Neurosci.*, **9**, 299–311.

Yeo, C.H., Hardiman, M.J. & Glickstein, M. (1985a) Classical conditioning of the nictitating membrane response of the rabbit. I. lesions of the cerebellar nuclei. *Exp. Brain Res.*, **60**, 87–98.

Yeo, C.H., Hardiman, M.J. & Glickstein, M. (1985b) Classical conditioning of the nictitating membrane response of the rabbit. II: lesions of the cerebellar cortex. *Exp. Brain Res.*, **60**, 99–113.

Yeo, C.H., Hardiman, M.J. & Glickstein, M. (1986) Classical conditioning of the nictitating membrane response of the rabbit. IV: lesions of the inferior olive. *Exp. Brain Res.*, **63**, 81–92.

Appendix

Soft- and hardware

The model was implementing using the neural network simulation software IQR421 (Bernardet *et al.*, 2002; noncommercial, developed by P. Verschure) running in a Linux environment (Redhat 6.1). The simulated conditioning experiments ran on one computer (Intel Pentium III, 650 MHz; USA). In robot experiments the processes controlling the cerebellar model, the robot, the camera mounted on the robot and the tracking camera were running distributed on four such machines.

Connectivity

Each described model element constitutes a group of 100 cells. All cell groups were connected one-to-one, except for *PU-S*0 for which each element received input from all *PU–SYN* elements. Thus, each *PU* receives input from only one *cf* but multiple *pf*. While the set of experiments described in this paper could have been mediated by a single *PU* cell with its afferents and efferents, the complexity of the circuit has been exploited in subsequent experiments involving multiple CS–US pairs.

Parameter specification

Abbreviations used for model elements are set out in Table 1. The values of cell parameters used are given in Tables 2 and 3. The cell types refer to possible settings in IQR421. A rebound threshold, θ^R, of –0.19 and a

rebound potential, μ, of 1.5 were used for the simulation of the *DN*. The hyperpolarized membrane potential was clamped at a value of –0.23. With these parameters *DN* emits spikes 8 ts after its disinhibition.

The value of W_{ij}^{max}, denoting the maximal (and initial) strength of the connection between the parallel fibre and Purkinje cell, was set to 0.4. The values E_{min}^{LTD} and E_{max}^{LTD} were set to 1.1 and 1.5 and the values for E_{min}^{LTP} and E_{max}^{LTP} were set to 0.05 and 0.054. If not explicitly denoted otherwise the values of the β of *PU–SYN*, the LTD rate ε and the LTP rate η used in the circuit tested in stimulated classical conditioning experiments were 0.985, 0.985 and 0.0002, respectively. Having established the relationship between the model's parameters and performance in simulation studies, the model's parameters could be chosen to allow for the desired fast acquisition and extinction characteristics for a given ISI. For the robot experiments presented the values of the β of *PU–SYN*, the LTD rate ε and the LTP rate η were 0.972, 0.95 and 0.006 (Figs 10–14) and 0.972, 0.94 and 0.002 (Fig. 15), respectively.

TABLE 1. Model elements

Abbreviation	Representing
cf	Climbing fibre
CRpathway	CR pathway downstream of deep nucleus
CSpathway	CS pathway upstream of pons
DN	Deep nucleus
GR	Granule cell
I	Cerebellar cortical interneurons
IO	Inferior olive
mf	Mossy fibre
pf	Parallel fibre
PO	Pons
PU	Purkinje cell
PU-SO	*PU* compartment: soma
PU-SP	*PU* compartment: spontaneous activity
PU-SYN	*PU* compartment: synaptic regions
USpathway	US pathway upstream of inferior olive
β^{SYN}	Model parameter: persistence of CS trace
ε	Model parameter: rate constant for LTD
η	Model parameter: rate constant for LTP

TABLE 2. Model parameters

Model element	Type	w	γ^E	γ^I	θ^A	β
PO	IntegrateFire	1	1	–	0.01	0.1
GR	IntegrateFire	1	1	–	0.01	0.1
IO	IntegrateFire	1.5	1	3	0.01	0.5
I	IntegrateFire	1	1	–	0.175	0.1
DN	Rebound	1	–	1	0.5	0.97

Description of the neuron types and the values of parameters. w, weight to connecting cell groups; γ^E, excitatory gain; γ^I, inhibitory gain; θ^A, firing threshold; β, membrane persistence.

TABLE 3. Model parameters of Purkinje cell

Model element	Type	w	γ^E	γ^I	θ^A	β
PU-SP	Spontaneous	0.1	–	3	0.01	0.5
PU-SO	IntegrateFire	1	1	–	0.06	0.1
PU-SYN	Linear threshold	Changing	1	–	0.15	Var.

Description of the neuron types and the values of parameters used for implementation of the three compartments of the modelled Purkinje cell. w, weight to connecting cell groups; γ^E, excitatory gain; γ^I, inhibitory gain; θ^A, firing threshold; β, membrane persistence; Var., variable.