Adaptive Sensor-Driven Neural Control for Learning in Walking Machines

Poramate Manoonpong and Florentin Wörgötter

Bernstein Center for Computational Neuroscience (BCCN)
University of Göttingen, D-37073 Göttingen, Germany
{poramate,worgott}@nld.ds.mpg.de
http://www.nld.ds.mpg.de/~poramate/

Abstract. Wild rodents learn the danger-predicting meaning of predator bird calls through the paring of cues which are an aversive stimulus (immediate danger signal or unconditioned stimulus, US) and the acoustic stimulus (predator signal or conditioned stimulus, CS). This learning is a form of pavlovian conditioning. In analogy, in this article a setup is described where adaptive sensor-driven neural control is used to simulate biologically-inspired acoustic predator-recognition learning for a safe escape on a six-legged walking machine. As a result, the controller allows the walking machine to learn the association of a predictive acoustic signal (predator signal, CS) and a reflex infrared signal (immediate danger signal, US). Such that after learning the machine performs fast walking behavior when "hearing" an approaching predator from behind leading to safely escape from the attack.

Keywords: Learned predator recognition, Associative learning, Legged robots, Sound, Reflexes, Prediction, Central pattern generators.

1 Introduction

Animals can effectively avoid dangerous situations including an attack of their predators. Evidence from physiological and ethological studies shows that some prey animals require experience to adaptively respond to predation (learned predator recognition) instead of an innate recognition. For example, wild rodents learn the danger-predicting meaning of predator bird calls through a temporal association of cues which are an aversive stimulus (unconditioned stimulus, US) and the acoustic stimulus (conditioned stimulus, CS) [1]. This mode of learning is a form of classical conditioning which has been discovered by Ivan Pavlov [2]. Generally, it involves presentations of a neutral stimulus (CS) along with a stimulus of some significance (US). Once these two stimuli become associated, animals begin to perform a behavioral response to the CS instead of the US. Inspired by these findings, we simulate here such predator recognition learning for a safe escape on a walking machine. This learned behavior together with the escape response is controlled by adaptive sensor-driven neural control using a correlation based differential Hebbian learning rule. The work presented here extends our previous works [3], [4] by integrating this learning mechanism into the

original modular neural control [3], [4] leading to the adaptive behavior. However, the main purpose of this article is not only to demonstrate the biologically-inspired learning on the walking machine (adaptive behavior) but also to show that the adaptive sensor-driven neural control can be a powerful technique to solve sensorimotor coordination problems of many degrees-of-freedom systems and to effectively provide an online learning capability to the systems.

The following section describes the technical specification of the walking machine platform. Section 3 explains the adaptive sensor-driven neural control for the acoustic predator-recognition learning and escape response. Experiments and results are discussed in Section 4. Conclusion and an outlook on future research are given in the last section.

2   The Walking Machine Platform AMOS

The six-legged walking machine AMOS [4], [5] is a biologically-inspired hardware platform for studying the coordination of many degrees of freedom, for performing experiments with neural controllers and learning and for the development of artificial perception-action systems employing embodied control techniques. It consists of a two-part body connected by one active backbone-joint (BJ), at which six identical legs and one active tail are attached. Each leg has three joints that are controlled by servomotors: the thoraco-coxal (TC-) joint enables forward ($+$) and backward ($-$) movements, the coxa-trochanteral (CTr-) joint enables elevation ($+$) and depression ($-$) of the leg, and the femur-tibia (FTi-) joint enables extension ($+$) and flexion ($-$) of the tibia (see [5] for the leg configuration of AMOS). Each tibia segment has a spring damped compliant element to absorb impact force during walking. All in all, this machine has 20 sensors: six foot contact ($FC_{1,...,6}$) sensors, seven infrared ($IR_{1,...,7}$) sensors, two light dependent resistor ($LDR_{1,2}$) sensors, one gyro (GR) sensor, one inclinometer (IM) sensor, one upside-down detector (UD) sensor, one current sensor (I) and one auditory-wind detector (AW) sensor (see [5] and also http://www.nld.ds.mpg.de/~poramate/ICONIP09/AMOSLearning.mpg for the location of all sensors on AMOS). Here the rear IR_7 and AW sensors installed at the tail are used for learning experiments of acoustic predator recognition which is the main contribution of this article while the use of other sensors in sensor-driven behavioral applications can be found in [4], [5]. The control of this walking machine is programmed into a personal digital assistant (PDA) with the update frequency of $\approx$ 14 Hz (see [4], [5] for more details of AMOS).

3   Adaptive Sensor-Driven Neural Control

The adaptive sensor-driven neural control (Fig. 1) generally generates reactive and adaptive behaviors. The controller based on a modular structure is formed by three main components: a neural preprocessing unit, a modular neural control unit, and a neural learning unit. The neural preprocessing unit filters sensory noise and shapes sensory data to drive corresponding reactive and adaptive behaviors. The modular neural control unit, on the other hand, is used for

locomotion generation of the walking machine. It coordinates leg movements, regulates walking speed, and creates omnidirectional walking. Additionally the neural learning unit allows the walking machine to perform adaptive behavior; i.e., it can learn to respond to a conditioned stimulus, e.g., an acoustic signal which is the main focus of this study. All neurons of the adaptive control are modelled as standard additive non-spiking neurons. Their activity develops according to $a_i(t + 1) = \sum_{j=1}^{n} W_{ij}\sigma(a_j(t)) + \Theta_i$; $i = 1, \ldots, n$, where n denotes the number of units, Θ_i represents a fixed internal bias term together with a stationary input to neuron i, and W_{ij} the synaptic strength of the connection from neuron j to neuron i. The output of the neurons in the neural preprocessing and modular neural control units is given by the standard sigmoid transfer function $\sigma(a_i) = (1 + e^{-a_i})^{-1}$ and the hyperbolic tangent (tanh) transfer function $\sigma(a_i) = \tanh(a_i)$, respectively, while in the neural learning unit is governed by a linear transfer function. Input units are linearly mapped onto the interval [0, 1] for all neurons in the neural preprocessing and learning units and [−1, 1] for those in the modular neural control unit.

3.1   Neural Preprocessing of Sensory Signals

In order to simulate the acoustic predator-recognition learning for escaping a close predator attack from behind (conditioned response) on the walking machine, we use the AW sensor [5] for detecting distant acoustic signals (predator signal) and the rear IR_7 sensor for perceiving too near, hence potentially dangerous, approaching objects from behind. Here we use portable speakers [5] implying a predator and the 300 Hz acoustic signal referring to a predator signal. The speakers are manually moved for creating the predatory attack situation (see http://www.nld.ds.mpg.de/~poramate/ICONIP09/AMOSLearning.mpg).

The raw acoustic and infrared signals coming from the AW and IR_7 sensors, respectively, require preprocessors for eliminating the sensory noise as well as shaping the sensory data. To do so, we utilize the dynamical properties, i.e., hysteresis effects, of a recurrent neuron to construct the preprocessing network of the acoustic signal and the same one is also applied to the infrared signal. Each network consists of a series of single recurrent neurons $H1, 2$. Their output is combined at a threshold output neuron O providing a binary output [0, 1] to the neural learning circuit. The complete neural preprocessing unit is shown in Fig. 1(a) (see [5] for more details). It has a capability to eliminate unwanted noise, shape the sensory data, and also prolong the activation time of the sensory signals. The prolongation of the signals are required in order to obtain the appropriate sensory correlation for a learning mechanism and to allow the walking machine to effectively escape from the predator attack. That is, the action shall persevere for longer than the stimulus itself to ensure a safe escape.

3.2   A Neural Learning Circuit

The preprocessed acoustic and infrared signals described in the previous section are fed into the neural learning circuit to modify a plastic synapse during learning. The learning goal in this study is to enable the walking machine to learn to
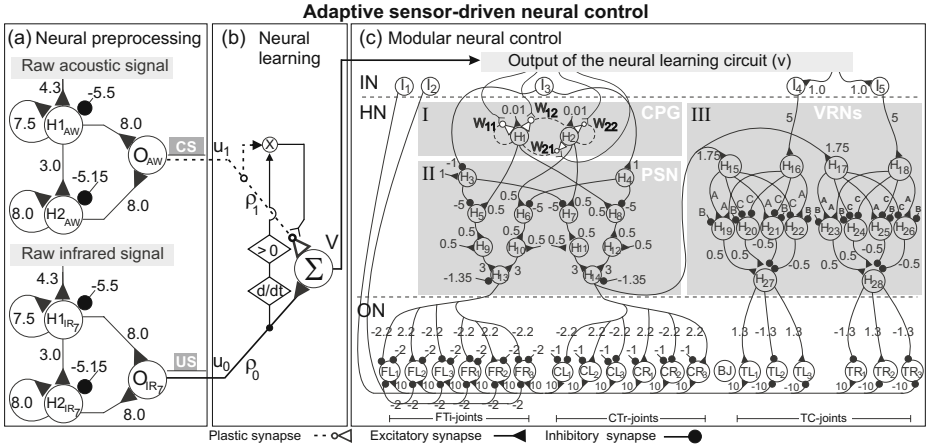
Fig. 1. (a) The neural preprocessing unit consisting of two identical networks with appropriate weights (see [5] for more details of neural parameter design and neurodynamics, i.e., hysteresis effects). These networks are applied to preprocess the acoustic and infrared signals. Their output is fed to the neural learning unit. $H1, 2_{AW}$ and $H1, 2_{IR_7}$ = hidden neurons of the AW and IR_7 sensor networks with the standard sigmoid transfer function, O_{AW} and O_{IR_7} = output neuron of the AW and IR_7 sensor networks with the threshold transfer function, CS = conditioned stimulus, US = unconditioned stimulus. (b) The neural learning circuit where its output activates a fast walking behavior through modular neural control. ρ_0 = synapse of the US input set to 1.0., ρ_1 = plastic synapse of the CS input, $u_{0,1}$ = inputs to a learner neuron, v = output of the learner neuron. (c) The modular neural control consisting of input (IN), hidden (HN), and output (ON) neuron groups. Input neurons I are the neurons used to control walking direction ($I_{2,...,5}$) and to trigger the self-protective reflex (I_1). Hidden neurons H are divided into three modules (CPG, PSN, and VRNs (see [4] for details)). Output neurons (TR, TL, CR, CL, FR, FL) directly command the position of servo motors. The location of the motor neurons on the walking machine AMOS is shown in [5]. BJ = backbone joint, TR(L) = TC-joints of right (left) legs, CR(L) = CTr-joints of right (left) legs, FR(L) = FTi-joints of right (left) legs. All connection strengths together with bias terms are indicated by numbers except those of the VRNs given by A = 1.7246, B = −2.48285, C = −1.7246. W_{11}, W_{22}, W_{12}, and W_{21} are modifiable synapses (see Sect. 3.3).

recognize its acoustic predator signal (300 Hz sound) and then perform a fast walking behavior as soon as it detects the signal. As a consequence, it can avoid the close attack of the predator. To achieve this we apply a correlation based differential Hebbian learning rule [6]. It correlates two kinds of input signals: one is a predictive signal or an early input (conditioned stimulus (CS), see Fig. 1(b)) and the other is a reflex signal or a later input (unconditioned stimulus (US), see Fig. 1(b)). The acoustic signal from the AW sensor is used as the predictive signal while the infrared signal from the IR_7 sensor serves as the reflex one. Both sensory signals are provided to the learning mechanism as shown in Fig. 1(b). At the beginning, the connection between the predictive acoustic signal and a

learner neuron converges with zero strengths (ρ_1, dashed line in Fig. 1(b)). In this situation, the fast walking behavior will be controlled only by the reflex infrared signal. That is, the machine generally walks with normal walking speed. Detecting the acoustic signal and afterwards the objects (portable speakers) from behind here implies the attack of the predator. This situation leads to the correlation between the predictive and reflex signals, such that the modifiable synapse ρ_1, which connects the predictive signal with the learner neuron, will grow. Consequently, after 2-3 attacks (depending signal correlation) during the learning phase, the fast walking behavior will finally be driven by the predictive signal instead. The used learning algorithm has the property that learning will stop when the reflex signal is zero [6]; i.e., when the portable speakers (sound source) does not get too near to the walking machine. As the weight is stored, after learning the next time the acoustic signal is detected the fast walking behavior will immediately be activated such that triggering an earlier escape reaction. Eventually the machine will return to its normal walking speed when the acoustic stimulus disappears implying that it is far enough from the sound source or the predator.

Learning Algorithm: The correlation based differential Hebbian learning rule (ICO-learning [6]) for the weight change of ρ_1 is given by $d\rho_1/dt = \mu u_1 du_0/dt$, where μ is the learning rate which will define how fast a system can learn, e.g., 0.35. One could consider μ as the susceptibility for a synaptic change, which in a biological agent will be defined by its evolutionary development, which determines the agent's ability to learn a certain task. How and if these values could also be influenced (possibly by mechanisms of meta-plasticity), changing learning susceptibility, goes beyond the scope of this article. In this learning rule, only the plastic synapse ρ_1 is allowed to change while the synapse of the reflex input ρ_0 is set to a positive value, e.g., 1.0. Note that here the weight change takes place only at a positive derivative otherwise it remains unchanged. Formally we have $v = \rho_0 u_0 + \rho_1 u_1$ as the learner neuron output driven by inputs $u_{0,1}$. Here we set u_0 and u_1 to the infrared signal of the IR_7 sensor (US) and the 300 Hz acoustic signal of the AW sensor (CS), respectively.

3.3   Modular Neural Control

Actions or walking behaviors of the machine are generated through modular neural control. This modular controller consists of three subordinate networks (colored boxes I, II, III in Fig. 1(c)): a neural oscillator network (I), a phase switching network (PSN, II), and two velocity regulating networks (VRNs, III). Here, we discuss only main functions of the networks (see [4] for a complete description). The neural oscillator network, serving as a central pattern generator (CPG) [7], generates periodic output signals. These signals are provided to all CTr-joints and FTi-joints only indirectly passing through all hidden neurons of the PSN. TC-joints are regulated via the VRNs. Thus, the basic rhythmic leg movement is generated by the neural oscillator network and the steering capability of the walking machine is realized by the PSN and the VRNs.

The neural oscillator network consists of two neurons $H_{1,2}$ with full connectivity. Its synaptic weights and bias terms are selected in accordance with the dynamics of the 2-neuron system [8] staying near the Neimark-Sacker bifurcation where the quasi-periodic attractors occur. The example of periodic output signals having different frequencies resulting from different weights can be seen at [4]. The network has the capability to generate various sinusoidal outputs depending on the weights. Here we use the output signal v of the neural learning circuit (Fig. 1) to modify all weights of the network determined by $w_{11,22} = 0.75v+1.125$, $w_{12} = 1.5v-0.35$, and $w_{21} = -1.5v+0.35$. As a consequence, walking frequency of the machine will be increased by the activation of v, resulting in a fast walking speed during escaping from the predator attack.

The PSN is a generic feed-forward network, which reverses the phase of the periodic signals driving the CTr- and FTi-joints (Fig. 1(c), see also [4]). These periodic signals can be switched to lead or lag behind each other by $\pi/2$ in phase in accordance with the given input I_3 (Fig. 1(c)). The PSN has been implemented to allow for sideways walking, e.g., for obstacle avoidance (see [4] for more details on parameters and experiments on sensor-driven sideways walking).

The two VRNs are also simple feed-forward networks (see [4]). Each VRN controls the three ipsilateral TC-joints on one side (Fig. 1(c), see also [4]). Because the VRNs behave qualitatively like a multiplication function [4], they have capability to increase or decrease the amplitude of the periodic signals by the magnitude of the inputs $I_{4,5}$. Consequently, the walking speed of the machine will be regulated, i.e., the higher the amplitude of the signal the faster it walks. Therefore, we also apply the neural learning output v to $I_{4,5}$ (see Fig. 1(c) and equations below). Such that the amplitude of the TC-joint signals will be amplified for the fast walking behavior while escaping. Moreover these VRNs can be used to achieve more walking directions, like forward and backward movement (sign inversion of the multiplication) or turning left or right where the directions are driven by other preprocessed infrared and light dependent resistor sensor signals through also $I_{4,5}$ (not shown in the current set of experiments but see [4]).

Figure 1(c) shows the complete network structure together with the synaptic weights of the connections between the controller and the corresponding motor neurons as well as the bias term of each motor neuron. These synaptic weights and all bias terms were manually constructed and adjusted to obtain an optimal gait; i.e., a typical tripod gait [4]. Taken together, this modular neural control can generate more than 10 different walking patterns including various reactive behaviors controlled via the input neurons $I_{1,...,5}$ (Fig. 1(c)). Here, we only focus on the acoustic predator-recognition learning for the safe escape; i.e., the walking machine will perform fast forward walking behavior as soon as it recognizes sound (after learning) or detects the close objects (before/during learning) from behind. Thus, the input neurons $I_{1,...,3}$ are here set to $I_1 = 0.0$, $I_2 = 1.0$, $I_3 = 1.0$, $I_4 = -1.0 + v$, $I_5 = -1.0 + v$ allowing the machine to walk forward only in order to clearly observe the difference between the slow walking speed under normal condition and the fast one during escaping the attack.

4 Experiments and Results

In this section, we illustrate the acoustic predator-recognition learning experiments. As shown in Fig. 2, usually before learning AMOS walks fast (≈ 17 cm/s) only if it detects the close objects (here, portable speakers manually moved) from behind through the rear IR_7 sensor. During learning, detecting the 300 Hz acoustic signal via the AW sensor and afterwards the objects from behind hints here danger from the predator attack. This situation leads to the correlation between the acoustic signal (predictive signal or "conditioned stimulus" (CS)) and the infrared signal (reflex signal or "unconditioned stimulus" (US)), such that the weight of the plastic synapse ρ_1 (Fig. 1(b)) increases. While the weight grows influencing the output v of the learner neuron, AMOS starts to walk fast driven by the CS instead of the US. After learning, it immediately performs fast walking when "hearing" the approaching predator (learned acoustic predator recognition) this way triggering an earlier escape reaction resulting in the safe predator avoidance. The learning process stops when the
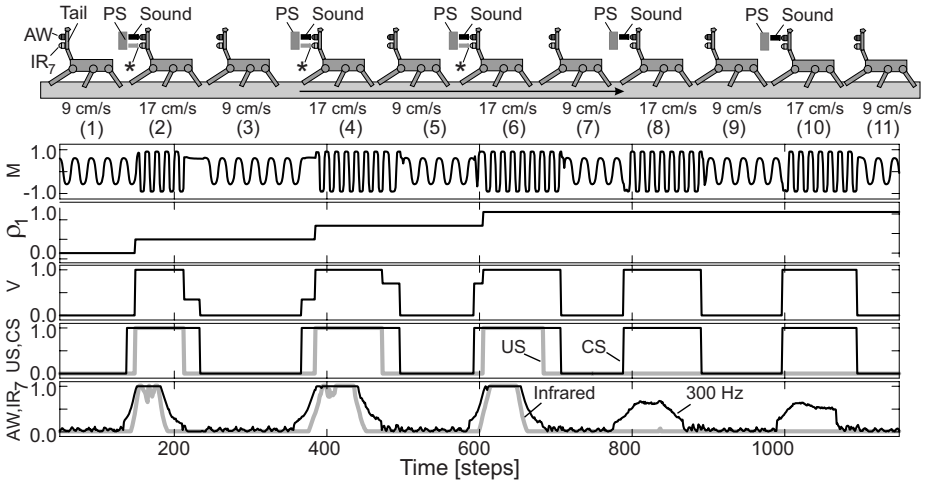

Fig. 2. Acoustic predator-recognition learning. The 300 Hz acoustic signal is detected by the auditory-wind detector (AW) sensor while the objects (portable speakers, PS) are detected by the rear infrared (IR_7) sensor. In the learning process, AMOS initially walks with its normal speed (1) and then speedily escapes from the too-near objects (2). As soon as it is far enough from the objects it then returns to its normal walking speed (3). Such learning situations repeatedly occur (4-7). Simultaneously, the correlation between the acoustic and infrared signals is being learned. Finally after 3 attacks (1-7) AMOS can immediately react to the acoustic signal as soon as it detects (8-11). Symbols: M = motor signal of a TC-joint, ρ_1 = plastic synapse, v = output of a learner neuron, US = unconditioned stimulus (infrared signal), CS = conditioned stimulus (acoustic signal), $*$ = detecting the PS (sound source). Note that one-trial learning could be also achieved by simply increasing a learning rate.

US is no longer triggered which happens when the fast walking behavior is driven by the CS. The video clip of the learning experiments can be seen at http://www.nld.ds.mpg.de/~poramate/ICONIP09/AMOSLearning.mpg.

5   Conclusion

In this study, we simulate acoustic predator-recognition learning on a six-legged walking machine such that the machine can learn the correlation between the predictive acoustic signal ("conditioned stimulus") and the reflex infrared signal ("unconditioned stimulus"). As a consequence, after learning, it performs an earlier escape reaction resulting in the safe predator avoidance as soon as it hears the approaching predator. These learned acoustic predator recognition and escape behavior are achieved under adaptive sensor-driven neural control (Fig. 1) which is modular structure-based design. It consists of three main modules: 1) neural preprocessing unit utilizing recurrent neurodynamics for preprocessing sensory signals, 2) neural learning unit using a correlation based differential Hebbian learning rule for associative learning capabilities, and 3) modular neural control employing a central pattern generator for basic locomotion generation. More demanding tasks will be related to implement neural memory to enhance learning capability through long-term memorization of walking relevant events.

Acknowledgements

References

1. Kindermann, T., Siemers, B.M., Fendt, M.: Innate or Learned Acoustic Recognition of Avian Predators in Rodents? J. Exp. Biol. 212, 506–513 (2009)
2. Pavlov, I.P.: Conditioned reflexes. Oxford University Press, London (1927)
3. Manoonpong, P.: Neural Preprocessing and Control of Reactive Walking Machines: Towards Versatile Artificial Perception-Action Systems. In: Cognitive Technologies. Springer, Heidelberg (2007)
4. Manoonpong, P., Pasemann, F., Wörgötter, F.: Sensor-Driven Neural Control for Omnidirectional Locomotion and Versatile Reactive Behaviors of Walking Machines. Robot. Auton. Syst. 56(3), 265–288 (2008)
5. Manoonpong, P., Pasemann, F., Wörgötter, F.: Neural Preprocessing of Auditory-Wind Sensory Signals and Modular Neural Control for Auditory- and Wind-Evoked Escape Responses of Walking Machines. In: 2008 IEEE International Conference on Robotics and Biomimetics, pp. 786–793. IEEE Press, Los Alamitos (2008)

6. Porr, B., Wörgötter, F.: Strongly Improved Stability and Faster Convergence of Temporal Sequence Learning by Using Input Correlations Only. Neural Comput. 18(6), 1380–1412 (2006)
7. Ijspeert, A.J.: Central Pattern Generators for Locomotion Control in Animals and Robots: A Review. Neural Networks 21, 642–653 (2008)
8. Pasemann, F., Hild, M., Zahedi, K.: SO(2)-Networks as Neural Oscillators. In: Mira, J., Álvarez, J.R. (eds.) IWANN 2003. LNCS, vol. 2686, pp. 144–151. Springer, Heidelberg (2003)