Second-Order Conditioning in Mobile Robots

Samuel Benzaquen and Carolina Chang

Grupo de Inteligencia Artificial, Departamento de Computación y TI
Universidad Simón Bolívar
Apartado Postal 89000, Caracas 1080, Venezuela
{96-28234,cchang}@ldc.usb.ve http://www.gia.usb.ve

Abstract. We have proposed a neural network that learns to control avoidance behaviors of a physical mobile robot through classical conditioning and operant conditioning. In this article we test whether our network can acquire second-order conditioning. During training we first associate the activation of the robot's infrared sensors with collisions. Then, the activation of a visual sensor is repeatedly paired with the activation of the infrared sensors. Results show that the robot learns to elicit avoidance responses whenever the visual sensor becomes active.

1 Introduction

Over the past years we have been interested in mimicking biological mechanisms of learning in physical mobile robots. We have proposed and tested a neural network model for classical conditioning and operant conditioning [3], added mechanisms of habituation to the network [2], explored sequence learning [7], and compared our network with related work, such as [1, 8]. In this paper we focus on the phenomenon of classical conditioning known as second-order conditioning.

In the classical or Pavlovian conditioning paradigm [5], learning occurs by repeated association of a Conditioned Stimulus (CS), which normally has no particular significance for an animal, with an Unconditioned Stimulus (UCS), which has significance and always gives rise to an Unconditioned Response (UCR). The response that comes to be elicited by the CS after classical conditioning is known as the Conditioned Response (CR), which is frequently very similar to the UCR. For example, every time a dog eats some food (UCS), it salivates (UCR). If the dog repeatedly hears a bell (CS) before being fed it will eventually begin to salivate (CR) soon after the bell is heard. This type of learning allows an animal to predict the events that take place in its environment.

Pavlov also described the second-order conditioning, which is a two-step process. In the first step, the CS is paired with the UCS (*i.e.* classical conditioning). In the second step, another neutral stimulus CS2 is presented repeatedly before the CS. After enough training, presentation of the CS2 alone produces the CR. Learning takes place even in the absence of the UCS, because the CS acquired UCS-like properties in step 1. Second-order conditioning is a complex phenomenon of animal learning. However, it seems feasible that it is acquired through the same mechanisms of classical (*i.e.*, first-order) conditioning.

In this paper we test the ability of our neural network for acquiring second-order conditioning. As we have done in the past, we let a robot learn through classical conditioning to avoid obstacles in a closed environment. Learning is achieved by repeated association of the robot's infrared sensor activation (CSs) with a collision signal (UCS). After training the robot can manage to avoid obstacles (UR) before a collision takes place. In a second step of conditioning, we use the robot's camera to detect color bricks (CS2) on the walls of the environment. Detection of the color bricks is often followed by the activation of the robot's short range infrared sensors. Hence, if our neural network can explain second-order conditioning, then the robot should learn to avoid impending collisions whenever the color bricks are detected.

2 The Neural Network Model

In recent years we have developed a neural network model that learns simultaneously to produce obstacle avoidance behaviors and light approach behaviors in a mobile robot [3]. The neural network is based on a detailed theory of learning proposed by Grossberg, which was designed to account for a variety of behavioral data on learning in vertebrates [4]. In this paper we have used a simplified version of our network, which focuses only on its mechanisms of classical conditioning. A description of the complete neural network can be found in [3].

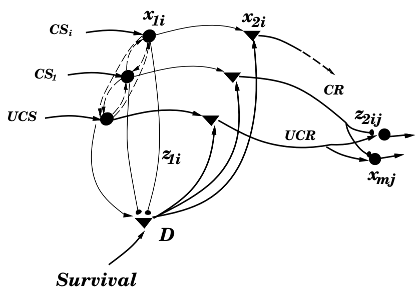


Fig. 1. The neural network model

Figure 1 depicts the neural network. Sensory nodes correspond to the conditioned stimuli, whose activation x_{1i} is given by:

$$x_{1i}(t) = \frac{I_i(t)}{\sum_j I_j(t)} \tag{1}$$

Here I_i represents a sensor value which codes proximal objects with large values, and distal objects with small values.

The drive node D codes the need to release an avoidance behavior. It becomes active when a homeostatic signal (*e.g.* survival instinct) is active in combination with a UCS, or with a CS that has learned to predict the UCS. The activation of drive node D is determined by the weighted sum of all the CS inputs, plus the UCS input, which is presumed to have a large, fixed connection strength:

$$y(t) = \sum_i x_{1i} z_{1i}(t) - T_y + UCS(t) \tag{2}$$

where $y(t)$ is the activation of the drive node D, z_{1i} is the adaptive weight connecting the sensory node x_{1i} to the drive node, T_y is a threshold that controls how easily the drive node is activated, and UCS(t) represents the collision status at time t ($UCS = 1$ if a collision just occurred, and $UCS = 0$ otherwise).

The activation of the drive node and of the sensory nodes converge upon a population of polyvalent cells. Each polyvalent cell receives input from only one sensory node, and all polyvalent cells also receive input from the drive node. We denote by x_{2i} the activation of the ith polyvalent cell:

$$x_{2i}(t) = x_{1i}(t) f(y(t)) \tag{3}$$

where $f(y(t)) = 1$ if $y(t) > 0$, and $f(y(t)) = 0$ otherwise. The multiplication of $x_{1i}(t)$ and $f(y(t))$ in Eq. 3 codes the need for joint activation from the sensory nodes (CS) and the drive node, in order for the polyvalent cells to become active.

The learning laws are given by:

$$z_{1i}(t) = E_1 z_{1i}(t-1) + F_1 x_{1i}(t) f(y(t)) \tag{4}$$

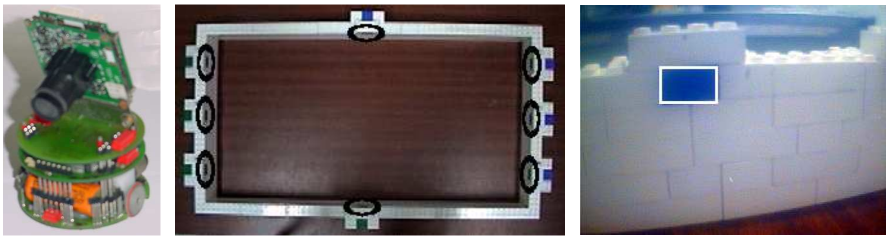
$$z_{2ij}(t) = E_2 z_{2i}(t-1) + F_2 x_{2i}(t) x_{mj}(t) \tag{5}$$

where E_i is the weight decay rate, and F_i is the learning rate. The adaptive connections z_{1i} grow as the CSs learn to predict the arrival of the UCS. On the other hand, connections z_{2ij} allow the network to learn to produce a CR, which is similar to the UCR elicited at the motor neurons x_{mj}. The motor neurons have binary activation, indicating whether each of the robot's motors is moving forward or backwards.

3 Robot and Environment

We have implemented the neural network of section 2 on a Khepera miniature mobile robot (K-Team SA, Préverenges, Switzerland). Khepera is a 55mm diameter differential drive robot (see figure 2(a)). It has eight infrared proximity sensors, and a color video turret. We situate the robot in a 48x26cm rectangular environment made out of white LEGO bricks, as shown in figure 2(b).Panel (d) is an image of the environment from the robot's view. Note that the walls occasionally have some color bricks, which are the neutral stimuli (CS2) used during second-order conditioning.

We use the *Khepera Integrated Testing Environment* (KITE) [6] to process the visual input. KITE is a tool for the evaluation of navigation algorithms for



(a) (b) (c)

Fig. 2. The Khepera miniature robot. **(a)** Khepera has 8 infrared sensors and a CCD color camera. **(b)** The environment is a 48x26cm rectangle made out of LEGO bricks. The picture shows the location of the color bricks on the walls. **(c)** The environment as seen from the robot's camera.

mobile robots. In particular, it features a segmentation module that detects and segments colored objects, automatically and in real time. For instance, in figure 2(c) KITE has segmented a blue brick from the background. The figure shows the bounding box of the segmented object.

4 Experiments and Results

The goal of this work is to achieve second-order conditioning in our mobile robot. To this end, we designed the following experiment. By default the robot always moves forward. We use information from the robot's encoders to detect collisions (UCS). The UCS always elicits an avoidance response (UCR). Whenever the robot collides with a wall, it turns to a side and then continues moving forward.

Training is done in two steps. In the first step the walls of the environment are made out of white LEGO bricks only. Readings from the robot's six front infrared sensors are neutral stimuli (CSs) that initially elicit no responses. The Khepera's infrared sensors can detect the proximity of objects within a range of 2.5cm. Learning through classical conditioning takes place as the robot navigates its environment because activation of any of the infrared sensors is often followed by a collision with a wall. Once the first-order conditioning is well established, activation of the infrared sensors alone elicits avoidance responses (CR). Therefore, after learning no more collisions take place.

In the second step of training we add some color bricks to the walls of the environment, as shown in figure 2(b). Initially the color bricks are regarded by the robot as neutral stimuli (CS2). The color bricks are detected by the robot's camera from an average distance of 8cm. Second-order conditioning develops when the robot navigates, due to the repeated association of the detection of the color bricks with the activation of the infrared sensors.

We trained the robot 10 times. In all trials the robot learned to avoid obstacles using its infrared sensors and camera. Table 1 shows the number of collisions required to learn to avoid obstacles through classical conditioning. It also shows the distance from the wall at which the responses are elicited, after first- and second-order conditioning. Second-order conditioning allowed the robot to predict impending collisions more ahead of time, resulting in a smoother navigation.

	Minimum	Average	Maximum
Number of collisions	34	45	57
Distance from wall at which the CR is elicited by the IRs	1.5cm	1.9cm	2.25cm
Distance from wall at which the CR is elicited by the visual input	3.2cm	5.1cm	7cm

Table 1. Results of 10 training trials.

Figure 3 shows the trajectories followed by the robot after first-order and second-order conditioning. The images are taken from a top-mounted camera. A tracking algorithm depicts the trajectory of the robot, identified by a pink mark. We have used white floor only to improve visualization of the pictures. The top row of the figure (panels (a), (b), and (c)) shows some trajectories after classical conditioning. The bottom row (panels (d), (e), and (f)) shows the trajectories after second-order conditioning, when navigating from the same initial positions and directions as above. The walls are completely white in panels (a), (b), and (c), while color bricks have been added to the walls of panels (d), (e), and (f).

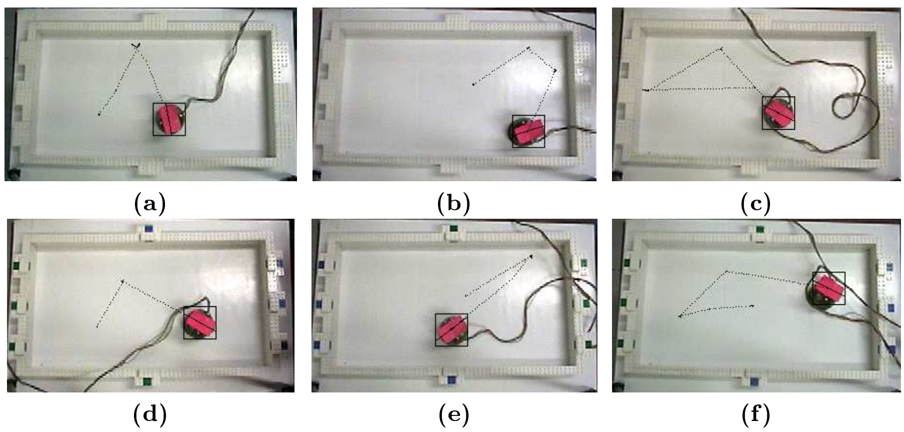


Fig. 3. First- and second-order conditioning. Panels (a), (b) and (c) show robot's trajectories after the infrared sensors have learned to predict impending collisions. Panels (d), (e) and (f) show the trajectories from the same initial points as before but after second-order conditioning. Avoidance behaviors are produced with more anticipation.

In figure 3(a), the infrared sensors become active when the robot gets close to the wall, which elicits a conditioned avoidance response. In (d) the response is generated at a farther distance because the robot has detected the color brick on the wall. Panel (b) shows the trajectory at a corner. The robot detects the walls twice, producing two turns to the right. When the robot detects the first wall in (e), it elicits a turn to the right as in (b). However, while the robot is turning, the camera detects a color brick on the other wall. Hence, the robot does not move forward but keeps turning. Panels (c) and (f) show once again that the responses are elicited with more anticipation after second-order conditioning.

5 Discussion

In this paper we have shown that our neural network can acquire second-order conditioning. The unsupervised network was trained on real time to control the reactive navigation of a physical mobile robot. In the past we have used this network to let a robot learn through first-order conditioning. Now we did not need to introduced any new neural mechanisms in order to account for second-order conditioning. This is one of the properties of the conditioning circuit [4], which is the theoretical foundation of this network. Our neural network could, in principle, acquire third- and higher-order conditioning.

Acknowledgments

We would like to thank Dr. Paolo Gaudiano, who donated the Khepera robot that we have used in this project.

References

1. Baloch, A., Waxman, A. M.: Visual learning, adaptive expectations, and behavioral conditioning of the mobile robot. *Neural Networks*, *4*, (1991) 271–302.
2. Chang, C.: Improving hallway navigation in mobile robots with sensor habituation. Proceedings of the 2000 IEEE International Joint Conference on Neural Networks, (2000) 143–147.
3. Chang, C., Gaudiano, P.: Application of biological learning theories to mobile robot avoidance and approach behaviors. *J. of Complex Systems*, *1*(1), (1998) 79–114.
4. Grossberg, S., Levine, D. Neural dynamics of attentionally modulated Pavlovian conditioning: blocking, interstimulus interval, and secondary reinforcement. *Applied Optics*, *26*, (1987) 5015–5030.
5. Pavlov, I. P. Conditioned Reflexes. Oxford University Press (1927).
6. Şahin E., and Gaudiano P. : KITE: The Khepera Integrated Testing Environment. Proceedings of the First International Khepera Workshop. Paderborn, Germany. (1999) 199–208.
7. Quero, G., and Chang, C.:Sequence Learning in Mobile Robots using Avalanche Neural Networks. *Lecture Notes in Computer Science 2085*, 508-515.
8. Verschure, P. F. M. J., Kröse, Ben J. A., Pfeifer, R.: Distributed adaptive control: The self-organization of structured behavior *Robotics and Autonomous Systems*, 9, (1992) 181–196.