Journal of Physiology - Paris 104 (2010) 108–117

Contents lists available at ScienceDirect

Journal of Physiology - Paris

journal homepage: www.elsevier.com/locate/jphysparis

Self-control with spiking and non-spiking neural networks playing games

Chris Christodoulou [a,*], Gaye Banfield [b], Aristodemos Cleanthous [a]

[a] Department of Computer Science, University of Cyprus, 75 Kallipoleos Avenue, P.O. Box 20537, 1678 Nicosia, Cyprus
[b] School of Computer Science and Information Systems, Birkbeck, University of London, Malet Street, London WC1E 7HX, United Kingdom

ARTICLE INFO

Keywords:
Self-control
Precommitment
Iterated Prisoner's Dilemma
Reinforcement learning
Spiking neural networks

ABSTRACT

Self-control can be defined as choosing a large delayed reward over a small immediate reward, while *precommitment* is the making of a choice with the specific aim of denying oneself future choices. Humans recognise that they have self-control problems and attempt to overcome them by applying precommitment. Problems in exercising self-control, suggest a conflict between cognition and motivation, which has been linked to competition between higher and lower brain functions (representing the frontal lobes and the limbic system respectively). This premise of an internal process conflict, lead to a behavioural model being proposed, based on which, we implemented a computational model for studying and explaining self-control through precommitment behaviour. Our model consists of two neural networks, initially non-spiking and then spiking ones, representing the higher and lower brain systems viewed as cooperating for the benefit of the organism. The non-spiking neural networks are of simple feed forward multilayer type with reinforcement learning, one with selective bootstrap weight update rule, which is seen as myopic, representing the lower brain and the other with the temporal difference weight update rule, which is seen as far-sighted, representing the higher brain. The spiking neural networks are implemented with leaky integrate-and-fire neurons with learning based on stochastic synaptic transmission. The differentiating element between the two brain centres in this implementation is based on the memory of past actions determined by an eligibility trace time constant. As the structure of the self-control problem can be likened to the Iterated Prisoner's Dilemma (IPD) game in that cooperation is to defection what self-control is to impulsiveness or what compromising is to insisting, we implemented the neural networks as two players, learning simultaneously but independently, competing in the IPD game. With a technique resembling the precommitment effect, whereby the payoffs for the dilemma cases in the IPD payoff matrix are differentially biased (increased or decreased), it is shown that increasing the precommitment effect (through increasing the differential bias) increases the probability of cooperating with oneself in the future, irrespective of whether the implementation is with spiking or non-spiking neural networks.

1. Introduction

Self-control arises out of a desire to control one's behaviour. In psychology, to exercise self-control is to inhibit an impulse to engage in behaviour that violates a moral standard (Morgan et al., 1979). Problems in exercising self-control occur when there is a lack of willingness or motivation to carry out this inhibition. This suggests a cognitive versus a motivational conflict. The motivational problem suggested by problems in exercising self-control can be interpreted as: we know what is good for us (cognition), but we do not do it (motivation). The distinction between cognition and motivation has been likened to the distinction between the higher and lower brain functions representing the frontal lobes and the limbic system respectively (Bjork et al., 2004). This suggests that self-control involves a conflict between cognition and motivation (Rachlin, 1995), a far-sighted planner and a myopic doer (Thaler and Shefrin, 1981), reason and passion, and is not just a case of changing tastes. These two extremes attain different value systems through experience (Scheier and Carver, 1988) and give rise to interpersonal conflict. Self-control problems stem from such a conflict. They also arise from a conflict at any single point in time of the choices we have available now and our future choices, and occur because our preferences for available choices are inconsistent across time (Ainslie, 1975; Loewenstein, 1996). More specifically Rachlin (1995) defines self-control as choosing a large-later (*LL*) reward over a smaller-sooner (*SS*) reward. Studies in self-control have found that increasing the delay of the reward, referred to as the delay of gratification (i.e., waiting for a more appropriate time and place to gain a reward), decreases the discounted value of the reward (Mischel et al., 1989). As the reward *SS* is imminent though, the discounted value of *SS* is greater than the discounted value of *LL*, so the person prefers the reward *SS* over the reward

* Corresponding author. Tel.: +357 22892752; fax: +357 22892701.
E-mail addresses: cchrist@cs.ucy.ac.cy (C. Christodoulou), gaye@dcs.bbk.ac.uk (G. Banfield), aris@cs.ucy.ac.cy (A. Cleanthous).

0928-4257/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jphysparis.2009.11.013

LL and thus we have *reversal of preferences*. In self-control problems the conflict arises out of this reversal of preferences between those choices available immediately (*SS*) and those available at some time later (*LL*). Reversal of preferences is seen in experiments on human subjects (Solnick et al., 1980; Millar and Navarick, 1984). To give an example where self-control behaviour is exercised, consider a student and let the *LL* represent obtaining good grades and the *SS* going to the pub. If we are at the start of the academic year, for most students the value of getting good grades exceeds that of going to the pub. When invited to the pub however, the value of *SS* is higher than their long term goal of getting good grades (*LL*). If the student exercises self-control then he or she will choose study (*LL*) over the pub (*SS*). Self-control behaviour encompasses a resistance to temptation, in this case to go to the pub (*SS*). Even though this view of self-control has been criticised as being a too simplistic representation of self-control in real life, as it models the situation only where the rewards are mutually exclusive and discrete (Mele, 1995; Plaud, 1995), we use it for our modelling in this paper, as it gives a clear preference for one alternative to another. It has to also be noted that the brain's ability to recognise or predict rewards is built in according to experiments by Richmond et al. (2003).

According to Ariely and Wertenbroch (2002) and Rachlin (2000), we recognise that we have self-control problems and try to solve them by *precommitment* behaviour. Precommitment behaviour can be seen as a desire by people to protect themselves against a future lack of willpower. Results by Ariely and Wertenbroch (2002) from a series of experiments on college students showed that we recognise that we have self-control problems, and attempt to control them by setting costly deadlines. These deadlines help to control procrastination, but are not as effective as externally imposed deadlines. *Precommitment* is more formally defined as making a choice now with the specific aim of denying (or at least restricting) oneself future choices (Rachlin, 1995). A typical example of precommitment is putting an alarm clock away from your bed, to force you to get up to turn it off. There are different levels of precommitment, which determine how successful the precommitment will be. According to Nesse (2001) precommitment is either (i) *conditional*, e.g., a threat, or (ii) *unconditional*, e.g., a promise. As he states, the carrying out of precommitment or not depends on how it is enforced. If the precommitment behaviour is *secured*, i.e., is enforced by the situation or a third party, then there is a greater degree of certainty that the behaviour will be carried out. If the precommitment behaviour is *unsecured*, i.e., it depends on the individual's emotion or reputation, then it is less certain that the precommitment behaviour will be carried out.

The internal process conflict suggested by self-control as described above, lead to a behavioural model being proposed (Rachlin, 2000), based on which, we implemented a computational model for studying and explaining self-control through precommitment behaviour (Banfield and Christodoulou, 2005). Our original model consisted of two simple feed forward multilayer perceptron type neural networks with reinforcement learning, representing the higher and lower brain systems viewed as cooperating for the benefit of the organism. In the latest version of the model, which is also presented in this paper, the feed forward multilayer perceptron type neural networks are replaced with two networks of leaky integrate-and-fire (LIF) neurons using a learning scheme based on reinforcement of stochastic synaptic transmission (Seung, 2003). As the structure of the self-control problem can be likened to the Iterated Prisoner's Dilemma (IPD) game, firstly in that cooperation is to defection what self-control is to impulsiveness (Brown and Rachlin, 1999) and secondly in that an interpretation of the IPD is that it demonstrates interpersonal conflict (Kavka, 1991), we implemented the neural networks as two players, learning simultaneously but independently, competing in the IPD game. The IPD was also used to model cooperation behaviour (Axelrod and Hamilton, 1981). Moreover based on our developed technique resembling precommitment, whereby the payoffs for the dilemma cases in the IPD payoff matrix are differentially biased, the relationship between precommitment behaviour and the value systems is also investigated.

2. Methods

2.1. The general model

In the viewpoint of modern cognitive neuroscience, self-control as an internal process can be represented in a highly schematic way as in Fig. 1a (based upon Rachlin, 2000). Arrow **1** in Fig. 1a denotes information coming into the cognitive system located in the higher centre of the brain, which represents the frontal lobes associated with rational behaviour such as planning and control. This information combines with messages from the lower brain, representing the limbic system (including memory from the hippocampus) that is associated with emotion and action selection (O'Reilly and Munakata, 2000; Rachlin, 2000). This travels back down to the lower brain and finally results in behaviour (Arrow **2** in Fig. 1a), which is rewarded or punished by stimuli entering the lower brain (Arrow **3** in Fig. 1a). In this paper, we implement the simple model of Fig. 1a as an architecture of two interacting networks of neurons (Fig. 1b). We also make the theoretical premise that the higher and lower brain functions cooperate, i.e., work together, which is in contrast to the traditional view of the higher brain functioning as a controller overriding the lower brain. From this viewpoint, a

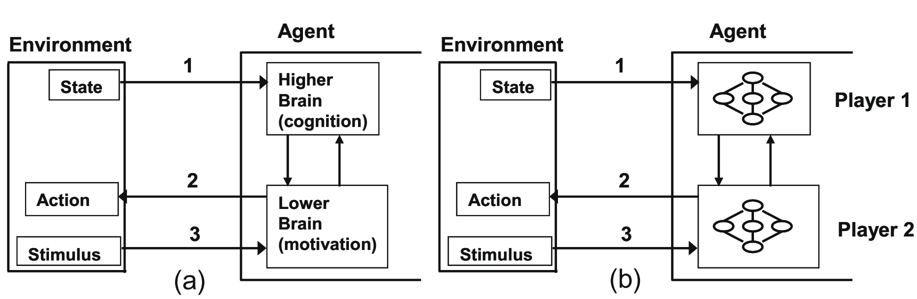

Fig. 1. A model of self-control behaviour. (a) Self-control as an internal process, from the viewpoint of modern cognitive neuroscience. Information of the temptation (*SS*) comes into the cognitive system (Arrow **1**). This combines with the messages from the lower brain and memory of our larger-later reward (*LL*). A choice is made, either LL or SS, which results in *behaviour* (Arrow **2**). We are then rewarded with *SS* or *LL* (Arrow **3**) (based upon Rachlin, 2000). (b) Our proposed computational model of self-control with the higher and lower brain centres modelled as two neural network players competing in the Iterated Prisoner's Dilemma. The *State* (Arrow **1**) summarises information both past and current about the environment; the *Action* (Arrow **2**) is the emergent behaviour of the Agent (the combined networks), and the *reinforcer* (Arrow **3**) is a global reward or penalty signal as a response to the *Action* (Arrow **2**).