Learning to Recall

by Jaldert O. Rombouts, Pieter R. Roelfsema and Sander M. Bohte

From the infinite set of routes that you could drive to work, you have probably found a way that gets you there in a reasonable time, dealing with traffic conditions and running minimal risks. Humans are very good at learning such efficient sequences based on very little feedback, but it is unclear how the brain learns to solve such tasks. At CWI, in collaboration with the Netherlands Institute for Neuroscience (NIN), we have developed a biologically realistic neural model that, like animals, can be trained to recall relevant past events and then to perform optimal action sequences, just by rewarding it for correct sequences of actions. The model explains neural activations found in the brains of animals trained on similar tasks.

Neuroscientists have prodded the inner workings of the brain to determine how this vast network of neurons is able to generate rewarding sequences of behaviour, in particular when past information is critical in making the correct decisions. To enable computers to achieve similarly good behaviour, computer scientists have developed algorithmic solutions such as dynamic programming and, more recent, reinforcement learning (Sutton and Barto 1998).

We applied the insights from reinforcement learning to biologically plausible models of neural computation. Concepts from reinforcement learning help resolve the critical credit assignment problem of determining which neurons were useful in obtaining reward, and when they were useful.

Learning to make rewarding eye movements

Animal experiments have shown that in some areas of the brain, neurons become active when a critical cue is shown, and stay active until the relevant decision is made. For example, in a classical experiment by Gnadt & Andersen (1988) a macaque monkey sits in front of a screen with a central cross (Figure 1). The monkey should fixate its eyes on the central cross and, while it is fixating, a cue is briefly flashed to the left or right of the cross. Then, after some delay, the fixation mark disappears. This indicates that the monkey should make an eye movement to where the cue was flashed. The monkey only receives a reward, usually a sip of fruit juice, when it executes the whole task correctly.

To solve the task, the monkey must learn to fixate on the correct targets at the correct times, and it must learn to store the location of the flashed cue in working memory, all based on simple reward feedback. The critical finding in these experiments was that, after learning, neurons were found that "remembered" the location of the flashed cue by maintaining persistently elevated activations until the animal had to make the eye movement.

Neural network model

We designed a neural network model that is both biologically plausible and capable of learning complex sequential tasks (Rombouts, Bohte, and Roelfsema 2012).A neural network model is a set of equations that describes the computations in a network of artificial neurons, which is an abstraction of the computations in real neurons.We incorporated three innovations in our neural model:

1. Memory neurons that integrate and maintain input activity, mimicking the persistently active neurons found in animal experiments.
2. Synaptic tags as a neural substrate for maintaining traces of an input's past activity, corresponding to eligibility traces in reinforcement learning (Sutton and Barto 1998).
3. We let the neural network predict the expected reward for different possible actions at the next time step: action values. At each time step, actions are chosen stochastically, biased towards actions with the highest predicted values.

A plausible learning rule then adjusts the network parameters to have the action values better approximate the amount of reward that is expected for the remainder of the trial. This learning rule is implemented through a combination of feedback activity in the network, and a global reward signal analogous to the function of the neurotransmitter dopamine in the brain.

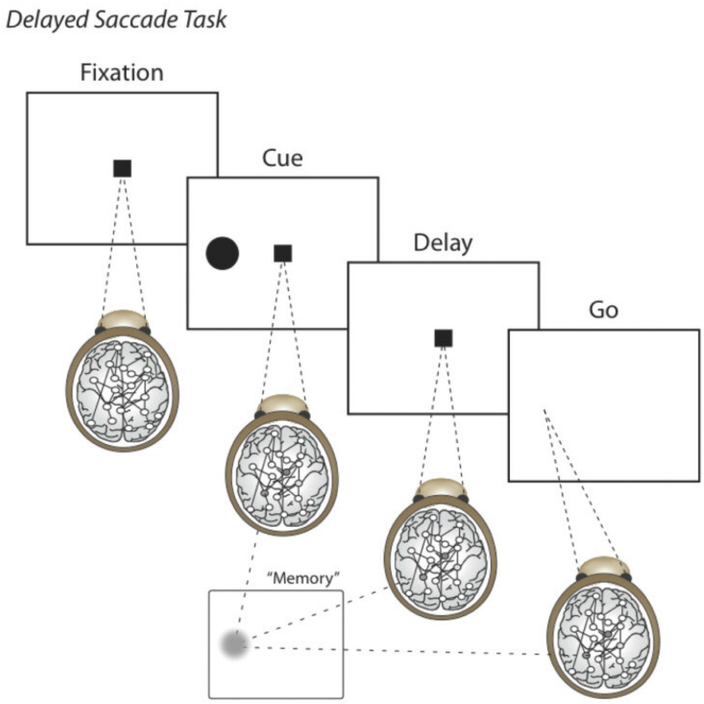

Figure 1: Delayed Saccade Task (Source: CWI)

When this model was applied to complex sequential tasks like the eye-movement task described above, we find that activity in the artificial neurons closely mimics the activity found in real neurons. In the example task, integrating neurons learn to code the cue that indicates the correct action as persistent activity, effectively learning to form a working memory. Thus, the model learns a simple algorithm by trial and error: fixate on the fixation mark, store the location of the flashed cue, and then make an eye-movement towards it when the fixation mark turns off.

The neural network model solves the problem of disambiguating state information: while driving to work, some streets look very similar; remembering the sequence of turns taken provides the information to determine your position. Mathematically, problems where instantaneous state information is aliased with other states are known as non-Markovian. Learning to extract and store information to disambiguate states is a challenge and an open problem. The neural model suggests how brains may solve some of these problems.

Link:
http://homepages.cwi.nl/~rombouts

References:
1. J.W. Gnadt, R.A. Andersen: "Memory Related motor planning activity in posterior parietal cortex of macaque", Experimental brain research 70(1):216–220, 1988

2. J.O. Rombouts, S.M. Bohte, P.R. Roelfsema: "Neurally Plausible reinforcement learning of working memory tasks", to Appear in Advances in Neural Information Processing (NIPS) 25, Lake Tahoe, USA, 2012

3. R.S. Sutton, A.G. Barto: "Introduction to Reinforcement Learning" MIT Press, 1998

Please contact:
Sander Bohte, Jaldert Rombouts
CWI, The Netherlands
E-mail: S.M.Bohte@cwi.nl, J.O.Rombouts@cwi.nl

Pieter Roelfsema
Netherlands Institute for Neuroscience (NIN), Amsterdam
E-mail: p.roelfsema@nin.knaw.nl

The Computer and the Brain, Synergies and Robots

by Martin Nilsson

Compared to a contemporary robot, the human body comprises a large number of actuators and sensors. Nevertheless, the central nervous system can efficiently extract just the right low-dimensional subsets of these for fast and precise motion control. How is this achieved? In the EU FP7-ICT project THE, scientists from neurophysiology, physics, computer science, and robotics are working together in order to try to answer this question. SICS' role is to try to understand and model some of the functioning of the mammalian central nervous system in order to apply it to adaptive control of robot limbs.

The Synergy: a clever brain trick?

The human hand-arm system has on the order of 102 degrees of freedom, but studies [1] have shown that just a small number of combinations — "motor synergies" — of elemen-

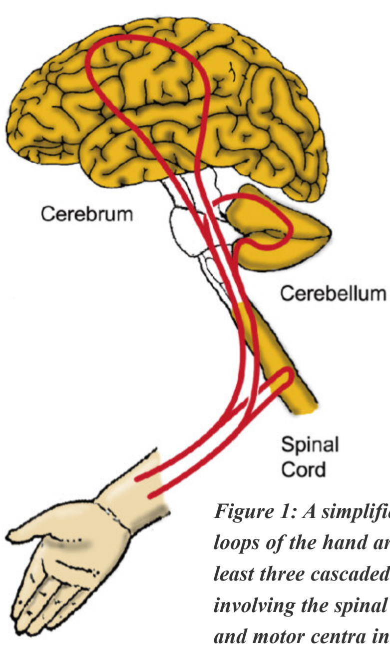

Figure 1: A simplified view of motor control loops of the hand arm system: It contains at least three cascaded feedback loops, involving the spinal cord, the cerebellum, and motor centra in the cerebral cortex.

tary movements account for most of the motion repertoire. How and why are such synergies formed? The hand-arm system also has in the order of 104 sensors. How does the brain "know" which sensor combination, "sensor synergy", best represents which motor? Knowing the answers to these questions would enable us to improve the design and control of robots. It is currently a real challenge to achieveanything approximating human agility in robots.

In the EU FP7-ICT project THE (www.thehandembodied.eu), scientists from neurophysiology, physics, computer science, and robotics are working together in order to try to find the answers. Considering the complexity of the mammalian central nervous system (CNS), this may appear to be an impossible task. However, we are beginning to see indications that much of the observed complexity in the adult is due to experience, while some of the fundamental, underlying mechanisms may be simpler than previously thought. Could it be that nature originally provides a relatively simple, general-purpose substrate, on which our interaction with the environment builds and optimizes the control circuitry? As our picture of the low-level machinery is crystallizing, some surprising properties are revealed.

von Neumann: ahead of his time

It is well known that when John von Neumann wrote the seminal "First Draft of a Report on the EDVAC" in 1945, he was deeply impressed by Alan Turing, but it is seldom mentioned that he was also much inspired by McCulloch and Pitts' work in neuroscience. In fact, references to the nervous system abound in the report, and the only publication referred to explicitly is their 1943 paper "A logical calculus of the ideas immanent in nervous activity". Although, for many years, there has been public debate on whether the brain can be compared to a computer, or even be understood at all, von Neumann himself considered the brain and the computer two kinds of automata. In his last work, "The Computer and the Brain", written in 1956, von Neumann compares computers with the brain, and many of von Neumann's observations are amazingly on target, more than 50 years later.