EXPLORING INTERNAL SIMULATION OF PERCEPTION IN MOBILE ROBOTS

Dan-Anders Jirenhed
Department of Computer Science, University of Skövde,
Skövde, Sweden, a99danji@ida.his.se

Germund Hesslow
Department of Physiological Sciences, Lund University,
Lund, Sweden, Germund.Hesslow@mphy.lu.se

Tom Ziemke[1]
Department of Computer Science, University of Skövde,
Skövde, Sweden, tom@ida.his.se

Abstract

Based on a neuroscientific hypothesis, this paper explores the possibility of an 'inner world' based on internal simulation of perception. We present three sets of experiments with a possible minimal model, using a simulated Khepera robot controlled by a simple recurrent connectionist network. Using an evolutionary algorithm the robots are trained on increasingly complex tasks. In the first experiment, serving as a baseline, robots are simply trained to map sensory input to motor output such that they move around in an environment without collisions. In the second experiment robots are additionally trained on predicting the next time step's sensory input. In the third experiment, finally, the robot's own prediction replaces the actual sensory input in order to investigate its capability to act 'blindly', i.e. in the temporary absence of external stimuli. Although only the first two experiments give positive results, we conclude that the experimental framework presented here should turn out useful in the investigation of more complex artificial neural models.

1. Introduction and Background

Introspective observation seems to tell us that we are able to have sensory experiences in absence of external stimuli. This has further been illustrated by experimental results of, e.g., Lee and Thompson [1]. In a series of experiments they demonstrated the accuracy with which humans can guide their behaviour based solely on internally generated sensory experiences. A group of subjects were first allowed to look at their surrounding environment and direct specific attention to certain objects, such as marks on the floor and different obstacles. They were then asked to perform different tasks such as walking to the marked locations, avoiding the obstacles and throwing objects at different targets in the room. All tasks were performed with eyes closed. The subjects performed these tasks almost as accurately with eyes closed as when they were free to look. It thus seems reasonable to assume the existence of an 'inner world' where sensory experiences and consequences of different behaviours may be anticipated.

Evidence from experiments with rats seem to imply a similar interpretation. When rats are allowed to move freely in a maze environment on several trials without receiving food reward, presumably no navigational behaviours are reinforced and thus not learnt. Later, when food is placed at one location for a number of trials, these rats will soon find that location and avoid dead ends just as reliably as a group of continually rewarded rats (a phenomenon known as "latent learning"). The evidence implies that maze topography is actually learnt while moving around although no rewards are given. It has been argued by, e.g., Tolman [2] that such observations should be understood as the result of an internal *cognitive map* being built up in the rat nervous system from sensory experiences. The 'map' is then consulted when navigating to the food location.

The traditional cognitive scientific explanation for how an 'inner world' might be organised, based on symbolic world models, internal maps and planning mechanisms, has been questioned since the mid-1980s by many researchers (e.g. Brooks [3,4]; Suchman [5]) who de-emphasise the role of internal world models in the traditional sense, and instead emphasise the situated nature of intelligence as well as the reactive nature of many of the mechanisms underlying the interaction between agents and their environments (see, e.g., Ziemke [6] for a detailed discussion). Purely reactive

---

[1] Corresponding author.

mechanisms, however, seem ill suited to explain the above capacity to simulate/anticipate sensory experience and to behave appropriately in the absence of external stimuli. Hence, robots or other autonomous agents, controlled by connectionist networks or similar mechanisms, have in many cases exhibited good performance in tasks that could be solved in a more or less reactive fashion. But, it is unclear how such systems could handle tasks that might require planning and action in the absence of sufficient external input.

However, Meeden, McGraw and Blank [7] presented experiments with a simple toy-car-like robot controlled by a *Simple Recurrent Network* or *SRN* [8] depicted in Figure 1. The robot's task was to alternate between approach and avoidance of a light source in one corner of its rectangular environment while avoiding obstacles. They showed that through the use of internal feedback (short-term memory) the robot could exhibit behaviour that was "plan-like" in the sense that (a) it associated abstract behavioural goals with sequences of primitive actions, (b) the behaviour could be described in hierarchical terms (e.g., *light seeking* was comprised of the sub-behaviors *orient towards light* and *go to light* which in turn were realised as primitive actions of moving and turning), (c) the robot maintained its overall strategy even when reacting flexibly to the environmental conditions. On the other hand, the behaviour was clearly not plan-like in the traditional sense that the robot would actually ever explicitly anticipate or plan future situations (more than one time step ahead).
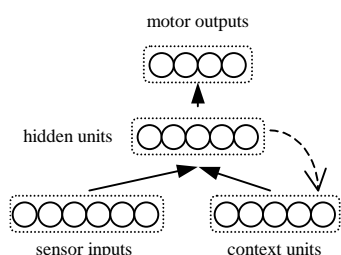

Figure 1: Basic robot control architecture used by Meeden *et al.* [7]. Solid arrows indicate fully connected layers of weights between layers of units (indicated by dotted lines). The dashed arrow represents a 1:1 copy connection.

Meeden *et al.* [7] also conducted experiments in which the robot was additionally trained on predicting the next time step's sensory input. The results showed that this additional training did have a positive effect on the behaviour learning. However, no analysis was presented regarding (a) the quality of the robot's sensor predictions, or (b) the robot's capacity to actually make use of its own prediction instead of external input.

Other investigations of sensor prediction have also been made. Tani and Nolfi [9] investigated how a connectionist robot controller could acquire an internal 'model' of the world through training on sensor prediction while moving around in a two-room environment. After learning, the controller had developed internal dynamics corresponding to environmental features (or concepts) such as 'corner' and 'corridor' and specific sequences of such low-level concepts were also used to form higher-level concepts such as 'room A' and 'room B'.

By drawing on psychological, physiological and neuroanatomical data on the properties of the cerebral cortex Gross, Heinze, Seiler and Stephan [10] presented a computational neural model of perception (called MASIM) in which sensor anticipation/prediction played a central role. In a series of robot experiments the MASIM architecture was used to guide the selection of motor responses based on anticipated sensor flow. Their results showed that anticipating robots displayed far better navigation abilities, with higher speeds and fewer collisions when compared to reactive robots. Most significantly, evasive manoeuvres in order to avoid obstacles could be initiated much sooner.

This paper shares features with the work of Tani and Nolfi [9] and Gross *et al.* [10] and aims to investigate exactly those points 'missing' in Meeden *et al.* [7]. The basic idea is that, if the robot's sensory predictions are sufficiently accurate, then it should be able to use them instead of actual sensory input and thus behave appropriately, at least for some time, in the absence of external stimuli, just as Lee and Thompson's [1] subjects did (cf. above).

This working hypothesis is supported by, among others, Hesslow's [11, 12] account of the 'inner world' in terms of *internal simulation* of perception and behaviour. Neurophysiological findings indicate that the neural structures involved in perception and initiation of *overt behaviour* could also be responsible for *mental imagery* and *covert behaviour*. As illustrated in Figure 2, an agent's overt behaviour when interacting with its environment, could be described, somewhat simplified, as a sequence of stimulus-response pairs. In the initial situation a stimulus, (S) triggers a response (R), which changes the environment from the agent's perspective into a new situation (S´) which functions as a stimulus triggering another response (R´), which in turn causes S´´ and R´´ and so on. All these steps go *via* internal processes where sensory states (s) cause motor response preparations (r).
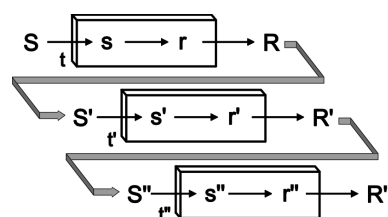

Figure 2: A stimulus-response sequence during overt behaviour.

Let us assume the existence of covert behaviour, i.e. an ability to generate neural motor responses which do not become observable bodily actions but only neural activation patterns which stay internal. Further assume the existence of a *sensor reactivation* or *imagery* mechanism, which allows for internally generated activation of sensor areas in the brain, so as to produce the simulated experience of a stimulus, but without the presence of the external stimulus. Finally, assume the existence of an *anticipation* mechanism, i.e. an ability to predict the sensory consequences of a motor response (a bodily movement). Support for each of these assumptions can be found in the neuroscience literature. Covert behaviour and imagery (sensor reactivation) are discussed by, e.g., Jeannerod [13], and imagery further by Kosslyn, Behrmann and Jeannerod [14]. Finally, several neuroscientists consider the cerebellum a prime candidate module involved in sensory prediction and anticipation (e.g. Miall [15]; Miall and Wolpert [16]; Wolpert, Miall and Kawato [17]; Thach [18]).

With these three mechanisms in place (covert behaviour, sensor reactivation and anticipation), it would be possible, to *internally simulate* the above overt behaviour sequence (cf. Figure 2) as illustrated in Figure 3.
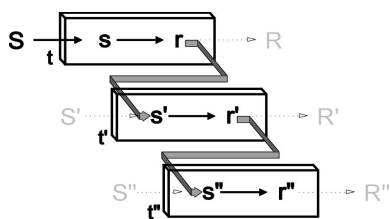

Figure 3: A sequence of internal simulation of perception and behaviour.

In the initial situation, a stimulus (S) causes a sensory activation (s) which triggers a motor response preparation (r). But instead of causing the overt response (R), the motor response preparation could, by influence of other neural activations inhibiting the motor response, cause a new sensory activation (s´) via a mechanism for anticipation of the sensory consequences and, in turn, a sensor reactivation mechanism. This new (internal) sensor activation triggers another motor response preparation (r´), which in turn causes a new sensory activation (s´´) which triggers a motor response preparation (r´´) and so on. Instead of overtly interacting with the environment, the agent internally simulates this interaction.

The rest of this paper is concerned with robot experiments investigating the question to what degree the relatively simple connectionist robot control architecture used in [7] (cf. Figure 1) can serve as the basis for internal simulation of perception.

2. Experiments

Robot and Environment

The experiments documented here have been carried out with a Khepera robot [19] depicted in Figure 4, or to be exact, with a simulator [20] based on sensor and motor measurements obtained from a real Khepera robot.
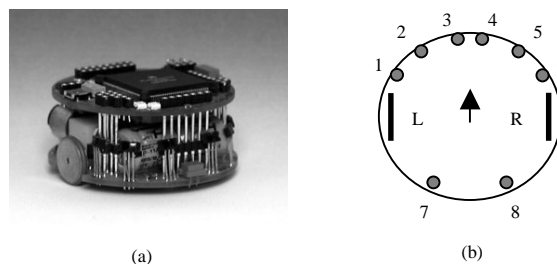

Figure 4: (a) Khepera robot built by K-Team SA (www.k-team.com). (b) Schematic drawing of the robot with infrared proximity sensors (1-8), left and right wheel (controlled by independent motors) . The robot's diameter and sensor range are about 55 mm. The direction of forward motion in the reported experiments is indicated by the arrow.

The experiments have been carried out in the two environments shown in Figure 5, hereafter referred to as the 'h-world' and the 'T-world'.
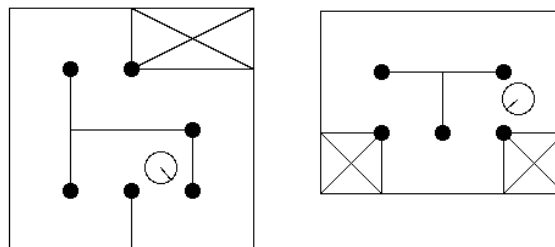

Figure 5: 'h-world' (left) and 'T-world' (right), each containing a simulated Khepera robot inside whose heading is indicated by the line inside the circle.

Experiment 1: Obstacle avoidance

The first experiment served as a simple test to ensure that Meeden's control architecture could be transferred to the robot and experimental conditions used here. In further experiments with the control architecture illustrated in Figure 1, Meeden [21] found that in experiments where only delayed reinforcement was available training was more successful with an evolutionary algorithm than with a backpropagation algorithm modified for reinforcement learning. Hence, in all experiments documented in this paper robot controllers were trained using a simple genetic

algorithm, similar to the one used by Nolfi [22] or Nolfi and Floreano [23]. A population of 150 individuals ('genotypes') was evolved over 500 generations and the 30 fittest individuals of each generation were selected for reproduction (using mutation only, no crossover) into the next generation. Each artificial genotype/chromosome encoded all connection weights in a recurrent connectionist network of a fixed topology. Each connection weight, ranging between –10.0 and +10.0, was encoded by 8 bits. The control architecture used was very similar to the one shown in Figure 1, except that it had eight input units for the eight infrared proximity sensors, three hidden (and context) units, and two outputs directly controlling the two motors. The fitness function used in experiment 1, similar to the one used in [24], rewarded robots for moving forward as straight and as quickly as possible while avoiding obstacles.

As exemplified in Figure 6, robot controllers evolved collision-free obstacle avoidance behaviour in both environments (in less than 100 generations). This confirms that Meeden's basic architecture (Figure 1) is also suitable for the domain at hand.
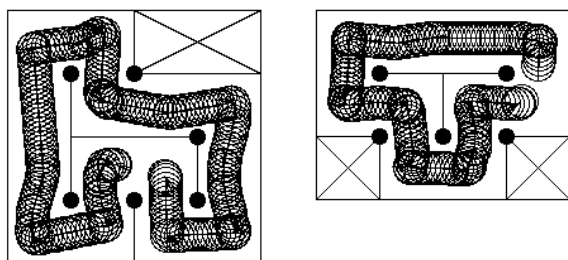

Figure 6: Example trajectories of robots with successful behaviour in the h-world (left; moving forward counter-clockwise) and the T-world (right; moving forward clockwise). Each circle depicts the Khepera robot's position in one time step, with the robot's heading indicated by the line inside each circle.

Experiment 2: Prediction

In experiment 2 the robot controllers were, in addition to the behavioural task of experiment 1, also trained on predicting the activation of their own sensory input units (i.e. the infrared proximity sensors) at the next time step. The robots were thus trained to anticipate the sensory consequences of their actions one time step ahead. The network architecture used here, illustrated in Figure 7, was similar to the one used by Meeden *et al.* [7] in a similar experiment (cf. above). Prediction ability was determined using a prediction fitness function which compared the predicted sensor activation vector with the actual activation vector (at the following time step) and produced a fitness value (reward) according to similarity. Two prediction fitness functions with slightly different scoring procedures were tested, but with no significant differences in the performance achieved. In

order to facilitate the evolution of both behavioural ability *and* prediction capacity, a two-stage selection routine was introduced into the above genetic algorithm. Using the same behavioural fitness function as in experiment 1, the 60 best-behaving individuals were selected in the first step. Then, from these pre-selected individuals the 30 best predictors were selected for reproduction into the next generation.
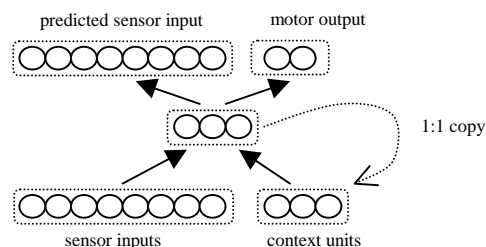

Figure 7: The connectionist controller network architecture used in experiment 2. Solid arrows indicate fully connected layers of weights. The dashed arrow represents a 1:1 copy connection.

Figures 8 and 9 illustrate the performance of representative individuals, hereafter referred to as individuals 1 and 2 (I_1 and I_2), as they evolved in T-world and h-world respectively. Both figures illustrate that the robots still manage to solve the behavioural task, although not necessarily as problem-free as in experiment 1 (cf. Figures 8a and 9a). Logged sensor activations were used to create activation plots, thus giving a description of the world as seen from the perspective of the robot (cf. Figure 8b and 9b). The magnitude of sensor activation at each time step is reflected by the width of the black line in the plot. High sensor activation in a time step is depicted as a wide black line. Plots describing the sensor predictions were also created using the same method (cf. Figure 8c and 9c).

An analysis of I_1 (cf. Figure 8a and 8b) shows that it follows the corridor by approximately more or less 'balancing' the activation values of left- and rightmost sensors 1 and 6 (cf. Figure 4). Figure 8c further illustrates that it also does relatively well at predicting these two sensors (and sensor 2 as well), whereas it basically 'ignores' or fails to predict the other sensors most of the time. I_2, on the other hand, uses a right-hand wall-following strategy (cf. Figure 9), i.e. it keeps right-hand sensors 5 and 6 at an almost constantly high activation level, apparently by turning right when these activation levels drop and turning left as soon as the other sensors get activated more than usual (indicating an obstacle ahead). Not surprisingly, it does relatively well at predicting the usually high activation of sensors 5 and 6 (in the latter case it does in fact predict constant full activation, which is right most of the time).
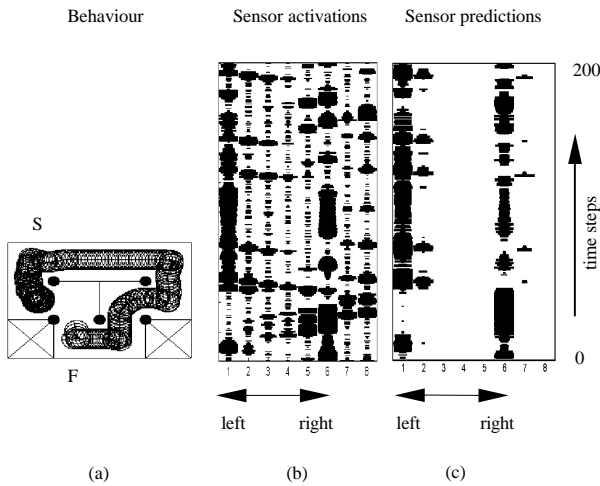
Figure 8: (a) Individual 1's trajectory over 200 time steps in the T-world (starting at S, finishing at F). (b) The robot's sensor state development over time while navigating. (c) The sensor state predictions made by the robot.

However, something that turned out to become problematic in experiment 3 can be pointed out already here: Perhaps due to its 'pre-occupation' with sensors 5 and 6, I_2 completely fails to predict the less frequently active sensors 1-3, although these clearly play a role in controlling its turning behaviour (cf. above). The same applies to I_1 which also failed to predict some of its sensors.
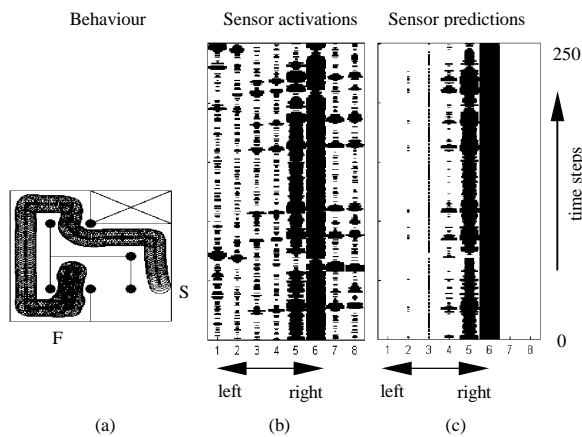

Figure 9: (a) Individual 2's trajectory over 250 time steps in the h-world (starting at S, finishing at F). (b) The robot's sensor state development over time while navigating. (c) The sensor state predictions made by the robot.

Experiment 3: Internal Simulation

Experiment 3 involved no further evolutionary runs. Instead the 20 best predicting individuals from each environment were chosen, from different stages of the evolutionary training processes, and their capacity for internal simulation, i.e. the capacity to temporarily rely on their own predictions rather than external stimuli, was tested as follows: (I) Starting from a random position in their 'home' environment the robots were allowed to move around for 200 (T-world) or 250 time steps (h-world) *with* external sensory input in order to be able to build up some internal context reflecting their position. (II) After the *context building phase* (I), the robots were 'cut off' from external sensory input and instead had to use their own sensory predictions for 10 time steps. During this *internal simulation phase*, they got no input from the environment (cf. Figure 10), but still their motor output was used to steer the robot in the real environment (this is supposed to roughly correspond to temporarily moving blindly in a previously seen environment, cf. above). (III) A *context-rebuilding phase* of another 40 time steps *with* external input.
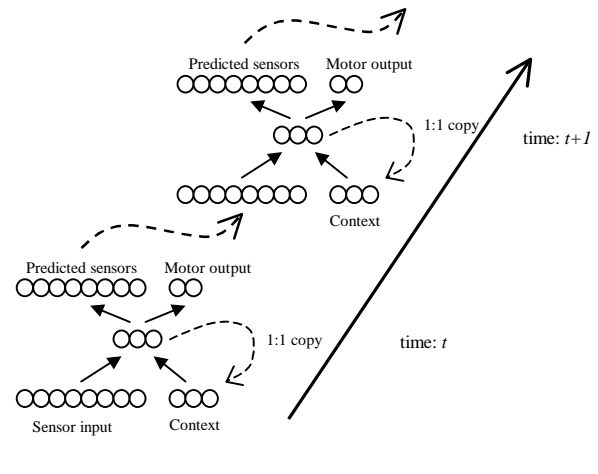

Figure 10: Internal simulation of perception. Instead of real sensor readings the previously predicted sensor activation vector was used as input in each time step. Overt behaviour during a simulation phase (II) was thus solely based on the robot's predictions about sensor states. Solid arrows indicate fully connected layers of weights, dashed arrows indicate 1:1 copy connections.

Phases (II) and (III) were repeated five times, such that altogether the robot was tested for five internal simulation phases of 10 time steps each. Figures 11 and 12 illustrate the behaviour of representative individuals I_1 and I_2 (cf. experiment 2), during the whole process of context building phase (of 200/250 time steps), followed by five repetitions of internal simulation phase (10 time steps, depicted in even numbered frames) plus context re-building phase (40 time steps, depicted in odd numbered frames). The figures also show the hidden and output unit activation levels during each time step.
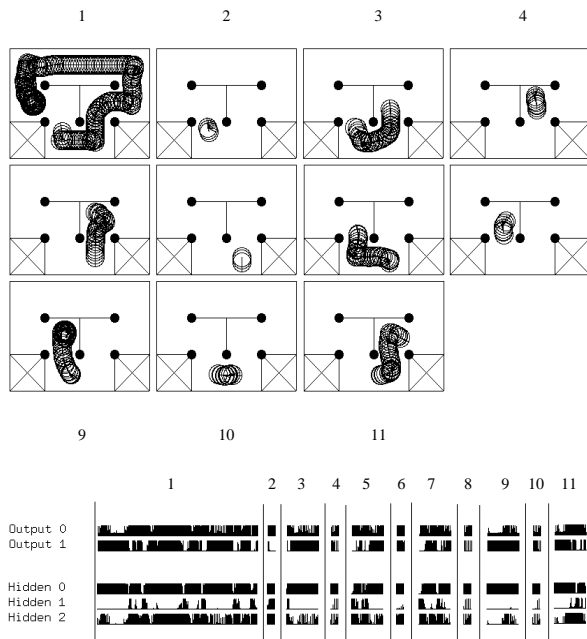
Figure 11: I_1 during context building phase (frame 1; 200 time steps), internal simulation phases (even numbered frames; 10 steps each) and context-rebuilding phases (odd numbered frames; 40 steps each).

Figure 12: I_2 during context building phase (frame 1; 250 time steps), internal simulation phases (even numbered frames; 10 steps each) and context-rebuilding phases (odd numbered frames; 40 steps each).

These figures illustrate quite clearly that both I_1 and I_2, although they do not actually collide more than once (cf. frame 6 in Figure 11), come quite close to the walls during internal simulation phases. More significantly, both of them in most of the cases do not correctly initiate turns when they should have.

A look at the hidden and output unit values (cf. Figures 11 and 12) shows that it is *not* the case that the robot controllers simply lack internal dynamics and just keep doing the same thing over and over again. In fact, both of them *do* exhibit varying behaviour and internal dynamics, but nevertheless they obviously fail to correctly predict, during internal simulation, the significant sensory dynamics/changes that help to control their behaviour while using external sensory input. I_2, for example, as already pointed out in the discussion of experiment 2 (cf. Figure 9), simply fails to predict the crucial changes in the activation of sensors 1-3, and thus cannot predict when it will need to initiate a turn.

3. Summary and Conclusion

Inspired by a neuroscientific hypothesis about internal simulation of perception (and behaviour) as the basis of an 'inner world' that allows humans to behave and anticipate the future even in the (temporary) absence of external sensory stimuli, we have presented a series of robot experiments with the aim to investigate a possible minimal artificial neural model.
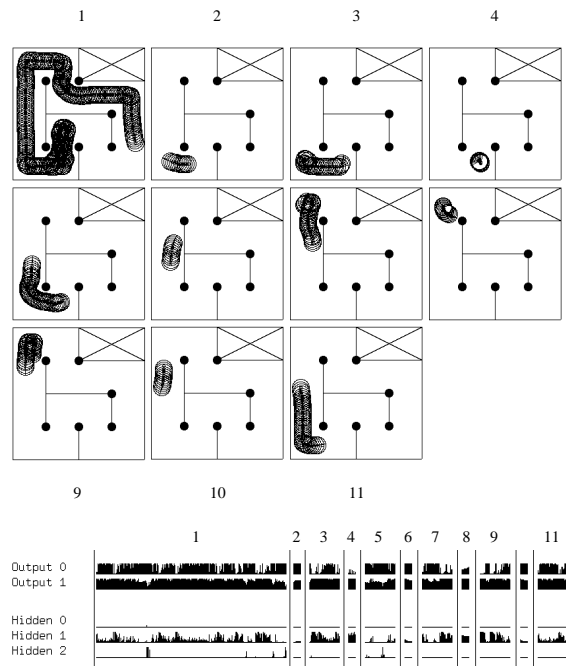
These experiments and the recurrent connectionist robot control mechanisms used here can be considered a further development of the work of Meeden (*et al.*) [7, 21], inspired by the work of Tani & Nolfi [9] and Gross *et al.* [10], in the sense that predictions are used to realise a minimal 'inner world' and also to control overt behaviour.

The results, however, show that although robots evolve to solve the behavioural task (experiment 1 and 2) and also seem to be able to roughly capture some of the relevant sensory dynamics (experiment 2), their prediction capacity is not sufficiently developed to rely on during extended periods of time (experiment 3). Of course, this can neither be interpreted as evidence for or against the internal simulation hypothesis as such. However, the somewhat minimalistic architecture and training regime used in the experiments presented here obviously have their limitations, nevertheless they might serve as a useful starting point for experiments with other architectures.

For future investigations it might be useful to give the controller network more freedom to develop suitable internal dynamics by adding hidden nodes, since the low number of hidden units in the experiments reported here may have contributed to the negative results for internal simulation. Another possible cause may have been the robot's limited sensor range. In the reported experiments the robots did not have much of a chance to anticipate obstacles since they could not detect them until they were quite close. Nothing beyond a distance

of 55 mm could be sensed. That means, the robot's task was not at all comparable to that of the human subjects in Lee and Thompson's experiments who had seen a complete view of the room they had to navigate blindly. Using a robot equipped with long-range sensors and/or a camera would allow for richer and more continuous sensor activation and smoother changes in the perceived world while behaving. This may have a positive effect on prediction learning and internal simulation of perception.

In conclusion, although the results presented here are not as good and as conclusive as one might have hoped for, we believe that the general approach and experimental framework presented in this paper will turn out useful for later investigations of the internal simulation hypothesis, using more complex artificial neural models.

Acknowledgements

References

[1] Lee, D. N. & Thompson, J. A. I. (1982). Vision in action: The control of locomotion. In D. Ingle, M. A. Goodale & R. J. W. Mansfield (Eds.). *Analysis of visual behaviour*. Cambridge, MA: MIT Press.

[2] Tolman, E. C. (1948). Cognitive maps in rats and men. *The Psychological Review*, 55(4), 189-208.

[3] Brooks, R. A. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, 2, 14-23.

[4] Brooks, R. A. (1991). Intelligence without Representation. *Artificial Intelligence*, 47, 139-159.

[5] Suchman, L. A. (1987). *Plans and Situated Action: The Problem of Human-Machine Communication*. New York: Cambridge University Press.

[6] Ziemke, T. (2001). The Construction of 'Reality' in the Robot: Constructivist Perspectives on Situated Artificial Intelligence and Adaptive Robotics. *Foundations of Science*, 6(1), 163-233.

[7] Meeden, L., McGraw, G. & Blank, D., (1993). Emergent Control and Planning in an Autonomous Vehicle. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 735-740). Hillsdale, NJ: Lawrence Erlbaum.

[8] Elman, J., (1990). Finding structure in time. *Cognitive Science*, 14, 179-214.

[9] Tani, J. and Nolfi, S., (1999). Learning to perceive the world as articulated: an approach for hierarchical learning in sensory-motor systems. *Neural Networks*, 12, 1131-1141.

[10] Gross, H.-M., Heinze, A., Seiler, T., Stephan, V. (1999). Generative Character of Perception: A Neural Architecture for Sensorimotor Anticipation. *Neural Networks,* 12, 1101-1129.

[11] Hesslow, G. (1994). Will Neuroscience Explain Consciousness?. *Journal of Theoretical Biology*, 171, 29-39.

[12] Hesslow, G. (2001). *Consciousness as simulation of behaviour and perception: Lecture notes*. URL: http://www.mphy.lu.se/avd/nf/hesslow/index.html. (Accessed 2001-07-30).

[13] Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioural and Brain Sciences*, 17, 187-245.

[14] Kosslyn, S. M., Behrmann, M. & Jeannerod, M. (1995). The cognitive neuroscience of mental imagery. *Neurophysiologia*, 33(11), 1335-1344.

[15] Miall, R. C. (1997). Sequences of sensory predictions. *Behavioural and Brain Sciences*, 20, 258-259.

[16] Miall, R. C., and Wolpert, D. M. (1996). Forward Models for Physiological Motor Control. *Neural Networks*, 9(8), 1265-1279.

[17] Wolpert, D. M., Miall, R. C. & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, 2(9), 338-347.

[18] Thach, W. T. (1998). What is the role of the cerebellum in motor learning and cognition? *Trends in Cognitive Sciences*, 2(9), 331-337.

[19] Mondada, F., Franzi, E., and Ienne, P. (1993). Mobile robot miniaturization: A tool for investigating in control algorithms. In *Third International Symposium on Experimental Robotics*, Kyoto, Japan.

[20] Miglino, O., Lund, H. H. & Nolfi, S. (1995). Evolving Mobile Robots in Simulated and Real Environments. *Artificial Life*, 2(4), 417.

[21] Meeden, L., (1996). An incremental approach to developing intelligent neural network controllers for robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 26(3), 474-485.

[22] Nolfi, S., (1997). Using Emergent Modularity to Develop Control Systems for Mobile Robots. *Adaptive Behaviour*, 5(3-4), 343.

[23] Nolfi, S., & Floreano, D. (2000). *Evolutionary Robotics – The Biology, Intelligence, and Technology of Self-Organizing Machines*. Cambridge, MA: MIT Press.

[24] Floreano, D., and Mondada, F., (1996). Evolution of Homing Navigation in a Real Mobile Robot. *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, 26(3), 396-407.

[25] Jirenhed, D.-A. (2001). *Exploring Internal Simulation of Perception in a Mobile Robot*. Master's Dissertation LIU-KOGVET-D-0047-SE. Department of Computer Science, University of Linköping, Sweden.