

Involving the motor system in decision making

Reto Wyss^{*}, Peter König and Paul F. M. J. Verschure

*Institute of Neuroinformatics, University/ETH Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
* Author for correspondence (r.wyss@ini.phys.ethz.ch).*

Recd 01.04.03; Acctpd 21.08.03; Online 08.10.03

The control of behaviour is usually understood in terms of three distinct components: sensory processing, decision making and movement control. Recently, this view has been questioned on the basis of physiological and behavioural data, blurring the distinction between these three stages. This raises the question to what extent the motor system itself can contribute to the interpretation of behavioural situations. To investigate this question we use a neural model of sensory motor integration applied to a behaving mobile robot performing a navigation task. We show that the population response of the motor system provides a substrate for the categorization of behavioural situations. This categorization allows for the assessment of the complexity of a behavioural situation and regulates whether higher-level decision making is required to resolve behavioural conflicts. Our model lends credence to an emerging reconceptualization of behavioural control where the motor system can be considered as part of a high-level perceptual system.

Keywords: navigation; decision making; motor system; entropy

1. INTRODUCTION

Recently, physiological and behavioural studies have shown that the borders between sensory and motor systems become increasingly vague (Salinas & Romo 1998; Rizzolatti & Luppino 2001; Kohler *et al.* 2002; Naito *et al.* 2002). Furthermore, there is increasing evidence that motor systems may not only be important in altering the physical state of an organism, but also in its cognitive state (Gardner & Lisberger 2002; Hernandez *et al.* 2002). Carpenter *et al.* (1999) have found direct evidence for task-related modulation of activity in motor systems, which makes it a candidate not only for the execution of movements but also for their planning. Indeed, a recent theoretical study has shown that to satisfy constraints observed in optimal decision making, the motor system should be seen as an integral component of a decision-making process (Verschure & Althaus 2003). The integration of the motor system with decision making and perception raises the question of whether the original triate sensing–deciding–acting captures biological reality (Rizzolatti & Luppino 2001). While the sensory input to an agent is, in general, task- and goal-unspecific, its motor output conforms to physical as well as task-related boundary conditions. Thus, an advantage of involving the motor system in decision making is therefore that decisions

automatically adhere to the constraints given by the morphology of the behaving agent (Rizzolatti & Luppino 2001).

We used a neural model of sensorimotor integration to study the possible role of the motor system in interpreting behavioural situations. An autonomous mobile robot learns to navigate in a maze (figure 1*a,c*), which requires different levels of behavioural control. At a local level, the agent has to learn accurately to follow black lines drawn on the floor using visual information provided by a camera mounted on top of the robot. This results in a stereotyped and reproducible behaviour. At a more global level, however, the agent has to retain a flexible behaviour to explore the maze optimally. We used both a real-world robot, Khepera (K-team, Lausanne, Switzerland, figure 1*a*) and a simulated virtual approximation of the latter to perform the experiments. The robot employs a form of reinforcement learning (Sutton & Barto 1998) to learn the mapping from the sensory input to its motor output (figure 1*b*). The reinforcement learning signal is derived from the visual input and corresponds to the temporal derivative of the summed activity from a small receptive field in the lower centre of the visual input (figure 1*b*). The robot is equipped with three basic pre-wired reflexes: driving straight, turning to the right and turning to the left. Rather than imposing particular motor actions, these reflexes bias the motor map by projecting linear activity ramps onto it.

The robot's behavioural uncertainty that requires decision making is reflected after learning by the entropy of the activity of the motor map. The entropy is low while following lines and increases significantly at crossings, where several valid motor actions are possible. Thus, the motor system may not only change the physical state of an organism but also constitute an integral part of the decision-making system.

2. METHODS AND RESULTS

In the following we present experiments performed with the virtual robot and subsequently show that the results generalize to the real world. For detailed descriptions of algorithms and implementation, see electronic Appendix A (available on the The Royal Society's Publications Web site). First, we investigated the robot's trajectories during 10⁶ time-steps for three different sensorimotor mappings (figure 1*d-f*). These mappings were acquired by the robot during 5000, 40 000 and 2 500 000 time-steps of learning, respectively, starting from a randomly initialized mapping. Whereas the robot followed the line only poorly after 5000 time-steps of learning, its performance after 40 000 time-steps was strongly improved and was similar to that after 2 500 000 time-steps. Thus, these results show that the learning mechanism for the acquisition of the sensorimotor mapping is both fast and stable.

On a global level, a flexible behaviour is desirable. This, however, can be achieved only when the constraints given by the local behaviour are relaxed, i.e. at crossings. We hypothesize that the entropy of the activity distribution in the motor map can signal whether behavioural options are available and decision making should be invoked. Therefore, we use a simple criterion for the detection of crossings, namely that the entropy of the motor map exceeds a threshold. Figure 2 shows the probability that the robot detects a crossing based on this criterion as a function of the robot's distance to the next crossing. The distribution peaks at a distance of approximately two body radii, which corresponds to the distance at which the crossings are centred in the robot's visual field. Thus, the entropy in the motor map is a reliable indicator of crossings.

We investigated whether this entropy signal can be used selectively to trigger decision making, allowing for optimal exploration. When the motor-map entropy exceeds the decision-making threshold, the default activity ramp imposed on the motor map for driving forward is replaced by a ramp favouring movements either to the right or the left. Both directions have the same probability of being chosen. We first investigated this principle for one particular crossing that the robot approaches from the bottom left corner (figure 3*a-c*). After

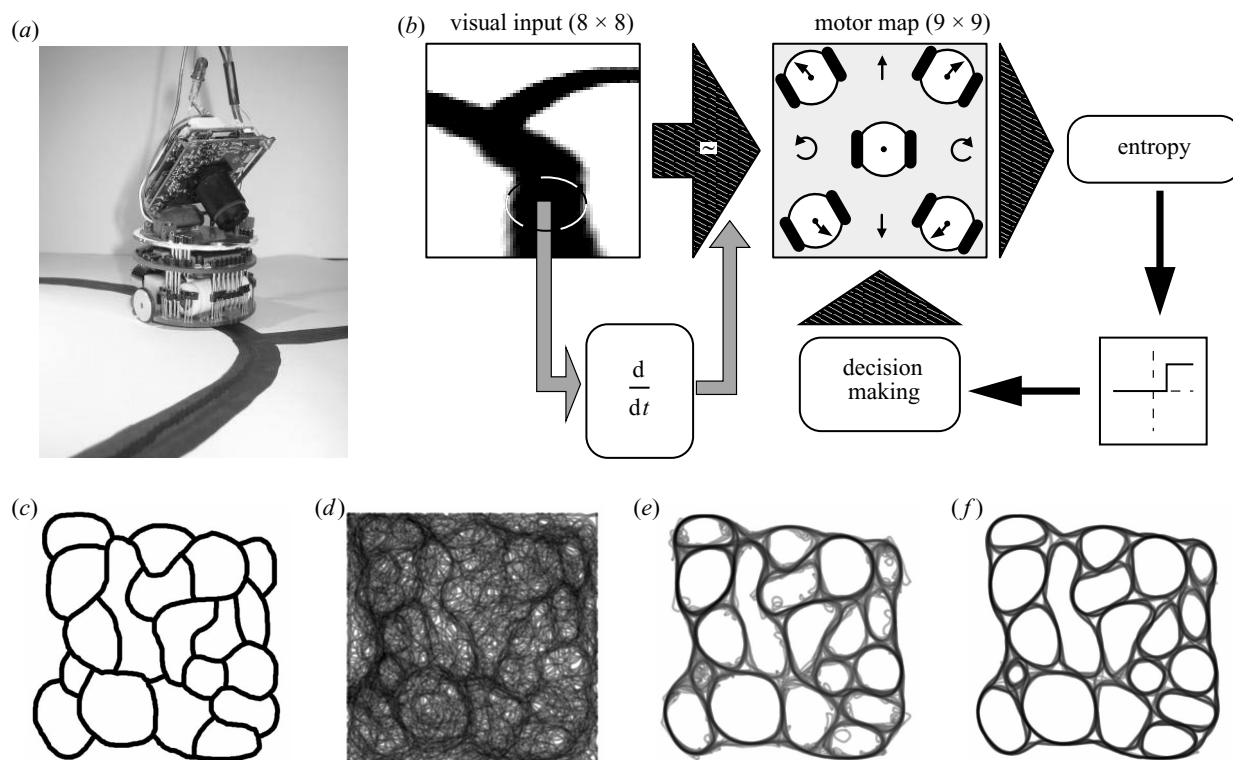


Figure 1. (a) The micro-robot Khepera, following lines drawn on the floor. A camera is mounted on top of the cylindrical body with an inclination angle of *ca.* 45°. (b) The network model. The visual input (8×8 pixels) is mapped onto the motor map consisting of 9×9 units, each corresponding to a pair of speeds for the left and right motors, respectively. Initially random actions are generated. These actions are validated by measuring the change in activity captured by the receptive field in the lower centre of the visual input (dashed circle): black lines moving in/out of the receptive field generate positive/negative learning signals. In parallel, the entropy of the motor map determines the activation of the decision-making mechanism. If the entropy exceeds a threshold, random reflexes with equal probability are triggered causing the robot to bias its actions toward the right or left. (c) The maze of lines for the simulated robot consists of 36 crossings of three lines each. The average thickness of the lines is *ca.* $0.5 \times$ body radius. (d–f) Trajectories of the robot recorded over 106 time-steps after the robot was learning to follow the line for 5000, 40 000 and 2 500 000 time-steps.

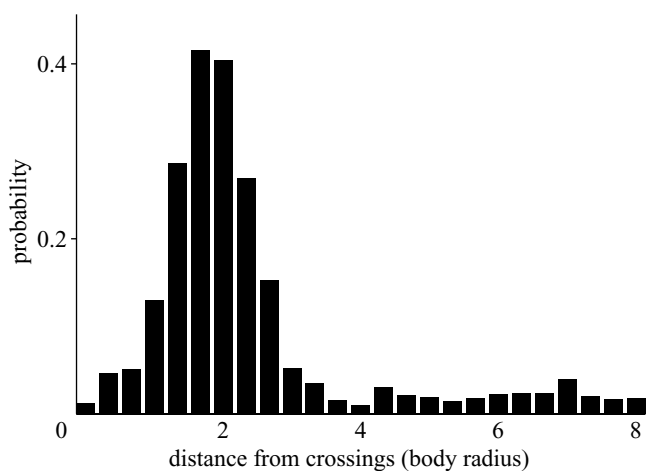


Figure 2. The probability of triggering a crossing event as a function of the robot's distance from the crossing. An event is triggered as soon as the entropy in the motor map exceeds a certain threshold.

10 000 time-steps of learning, the crossing is still a likely source of errors where the robot loses the line (figure 3*a,d*). In addition, the robot follows the right branch more frequently although the decision-making mechanism chooses both directions with equal probability. After further learning (40 000 time-steps), the errors are reduced, but there is still an imbalance in the directions that the robot takes at the crossing (figure 3*b,d*). It is not until a further refinement of the

sensorimotor mapping (2 500 000 time-steps) that the entropy in the motor map becomes a reliable indicator of crossings and the motor map itself allows any of the two branches to be followed with equal probability (figure 3*c,d*). Thus, we exploited the events of high entropy to trigger the transitions from reproducible and stereotyped behaviour at a local level to flexible and explorative behaviour at a global level.

The previous findings, based on the single crossing generalize to the whole maze. In total, the maze contains 36 crossings that can each be approached from three different branches. Thus, the maze consists of 108 oriented branches, which should be visited equally often given that the robot makes unbiased decisions at each crossing (F. Roth and R. Wyss, personal communication). In an experiment over 106 time-steps, we investigated the rates at which the robot visited the individual oriented branches with and without the entropy-driven decision-making mechanism (figure 3*e*). In the case without decision making (control condition), the maze coverage was very uneven, e.g. some oriented branches were not visited at all, while others were visited over three times more often than expected from a uniform distribution. With the decision-making mechanism, this imbalance was greatly reduced and, in particular, there were no more unvisited branches. We quantified the difference between the two distributions by calculating the standard deviation over the rates at which the oriented branches were visited for the two conditions (figure 3*f*). The coverage of the maze was more homogeneous using the decision-making mechanism.

We explored with the Khepera robot whether these results generalize to the real world. After 1 hour of learning (*ca.* 90 000 time-steps), the precision of line-tracking reached a high level (data not shown), while the homogeneity of maze coverage was comparable with that achieved with the simulated robot (figure 3*f*). Thus, the behaviour of the real-world robot matched the performance of the simulated system on the local as well as on the global level.

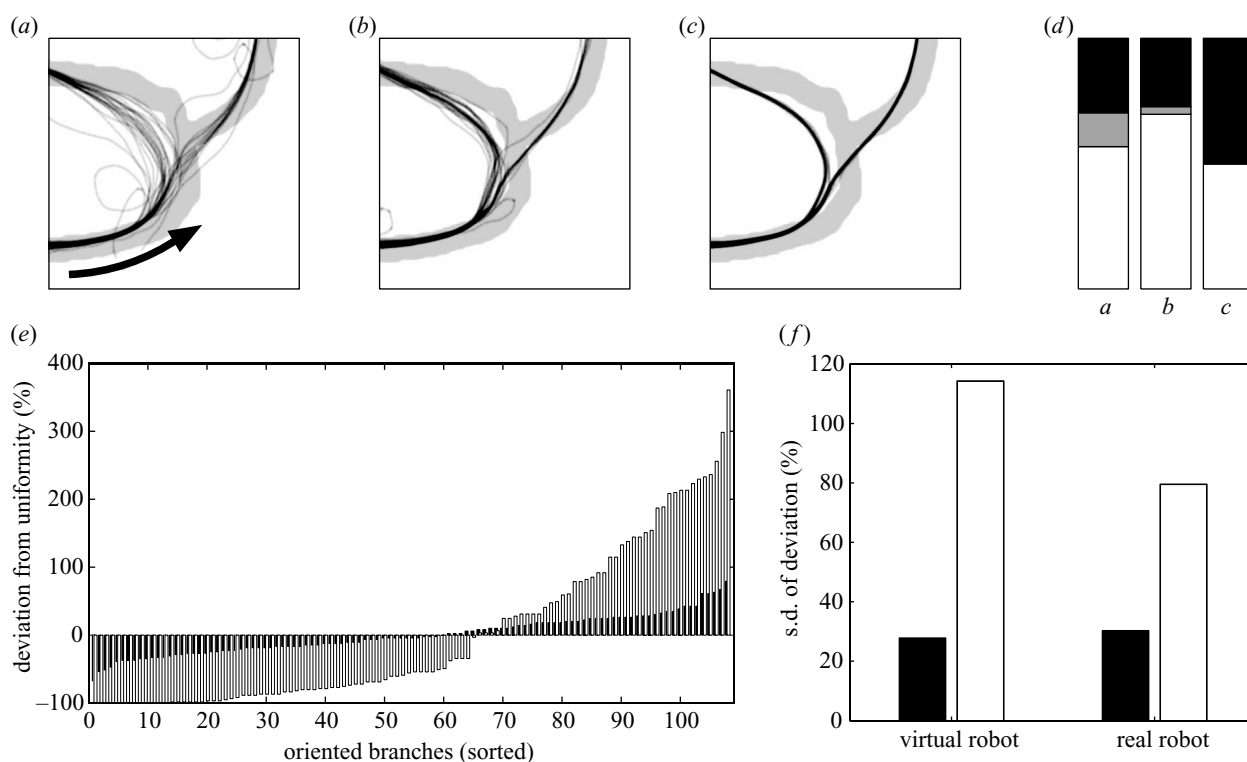


Figure 3. (a–c) Trajectories of the robot approaching one crossing from the lower-left branch (see arrow) after 10 000, 40 000 and 2 500 000 time-steps of learning, respectively. (d) Fractions of left turns (black filled), right turns (grey filled) and errors (open) at the crossing for the situations shown in (a–c) with the number of trajectories being 37, 33 and 32, respectively. (e) The deviation from a uniform maze coverage for a robot using the decision-making mechanism (filled bars) and a control where decision making is disabled (open bars). For both distributions, the oriented branches were sorted according to their deviation from uniformity. The sorting order is not necessarily identical for the two distributions. (f) Inhomogeneity for the virtual and the real-world robot with and without decision making (filled and open bars, respectively). Inhomogeneity is defined as the standard deviation of the rates at which the differently oriented branches are visited in percentage of the uniform rate. Please note that this measure is independent of the total number of crossings.

3. DISCUSSION

We used a neural model of sensorimotor integration to study to what extent the motor system itself can contribute to the interpretation of behavioural situations. The proposed neural network controls a behaving robot that learns to explore a maze optimally. We show that the population response of the motor system provides a substrate for the categorization of behavioural situations in terms of the behavioural options that they allow. Based on this categorization, higher-level decision making is engaged to resolve behavioural conflicts.

The linkage between sensory input and behaviour involving interpretation and behavioural selection is usually referred to as a decision process. Currently, we do not know of any dedicated brain structure that contains an abstract representation of interpretations or decisions that is not related to a motor system or dependent on continuous sensory stimulation (Leon & Shadlen 1998). A theoretical study suggests that sensory neurons can encode the variables required to compute a decision, but they neither carry out this computation nor represent its outcome (Shadlen *et al.* 1996). Here, we propose that the motor system itself is a likely candidate for the detection of behaviourally relevant situations.

Acknowledgements

This work was supported by the European Community/Bundesamt für Bildung und Wissenschaft grant no. IST-2001-33066, the Volkswagen Foundation (to P.V.) and BBW grant no. 01.0208-1, IST-

2000-28127 (to P.K.). The authors thank F. Roth and P. Knüsel for valuable discussions and contributions to this study.

- Carpenter, A., Georgopoulos, A. & Pellizzer, G. 1999 Motor cortical encoding of serial order in a context-recall task. *Science* **283**, 1752–1757.
- Gardner, J. & Lisberger, S. 2002 Serial linkage of target selection for orienting and tracking eye movements. *Nature Neurosci.* **5**, 892–899.
- Hernandez, A., Zainos, A. & Romo, R. 2002 Temporal evolution of a decision-making process in medial premotor cortex. *Neuron* **33**, 842–844.
- Kohler, E., Keysers, C., Umiltà, M., Fogassi, L., Gallese, V. & Rizzolatti, G. 2002 Hearing sounds, understanding actions: action representation in mirror neurons. *Science* **297**, 846–848.
- Leon, M. & Shadlen, M. 1998 Exploring the neurophysiology of decisions. *Neuron* **21**, 669–672.
- Naito, E., Roland, P. & Ehrson, H. 2002 I feel my hand moving: a new role of the primary motor cortex in somatic perception of limb movement. *Neuron* **36**, 979–988.
- Rizzolatti, G. & Luppino, G. 2001 The cortical motor system. *Neuron* **31**, 889–901.
- Salinas, E. & Romo, R. 1998 Conversion of sensory signals into motor commands in primary motor cortex. *J. Neurosci.* **18**, 499–511.
- Shadlen, M., Britten, K., Newsome, W. & Movshon, J. 1996 A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* **16**, 1486–1510.
- Sutton, R. & Barto, A. 1998 *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press.
- Verschure, P. & Althaus, P. 2003 A real-world rational agent: unifying old and new AI. *Cogn. Sci.* **27**, 561–590.

Visit <http://www.pubs.royalsoc.ac.uk> to see an electronic appendix to this paper.