PERGAMON

Neural Networks 15 (2002) 305–307

Neural Networks

www.elsevier.com/locate/neunet

Current opinions

Three creatures named 'forward model'

A. Karniel*

Department of Physiology, Northwestern University Medical School, 303 East Chicago Avenue, Chicago, IL 60611-3008, USA

Accepted 10 January 2002

Abstract

It has recently been suggested that the nervous system employs forward models for the purpose of motor control. The evidence for this hypothetical computational structure comes from various sources, theoretical and experimental. The purpose of this commentary is to distinguish between three different structures that are used to support the forward model hypothesis: (1) output predictor, (2) state estimator, and (3) distal teacher. It is possible that the brain employs all these structures for control. However, the general term forward model could be misleading since evidence for one structure could be misinterpreted as supporting evidence for the other structures. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Motor control; Forward model; Output predictor; State estimator; Distal teacher

1. Introduction

The main pillar of any model for motor adaptation is the simple fact that the brain knows something about its motor apparatus and the controlled environment. This knowledge is frequently described by the notion of internal models (Kawato, 1999). Fig. 1(a) shows a feed-forward control system, where the controller generates motor command to the plant, which determines the output. For example, the motor command could be a neural command to the muscles, and the output could be the position of the hand in the extrinsic workspace. If one assumes that the controller generates its command based on a desired output signal then it is obvious that the controller is essentially an inverse of the controlled system.

A forward model is a structure that predicts the output of the system based on its input (Fig. 1(b)). In the last decade, many studies presented experimental and theoretical arguments to support the hypothesis that the brain employs forward models (Blakemore, Wolpert, & Frith, 2000; Flanagan & Wing, 1997; Guenther & Micci Barreca, 1997; Jordan & Rumelhart, 1992; Miall & Wolpert, 1996). Unfortunately, the forward model hypothesis is not always supported by direct evidence for the structure in Fig. 1(b).

2. Output predictor

Many studies that support the forward model hypothesis present evidence for output predictor (Fig. 1(c)). The observation that we cannot tickle ourselves is vivid evidence for the assertion that the brain can predict (and cancel) the results of its own action (Blakemore et al., 2000). Another clear example is the ability to generate the necessary grip force during object manipulation (Flanagan & Wing, 1997). Evidence for output predictor (Fig. 1(c)) indeed supports the hypothesis that a forward model (Fig. 1(b)) exists, however, it is not the only way to predict the output, in some cases, a simpler way could be using the desired output as a predictor to the actual output. In ideal conditions, when a perfect inverse model generates the control signal and a perfect forward model predicts the output, the desired output and the predicted output are the same (see, e.g. fig. 1c in Kawato, 1999). Clearly, noise and uncertainties always exist, and therefore the prediction could be better than the desired output, in addition, the use of previous outputs could further improve the prediction. However, once we introduce additional inputs we deviate from the simple definition of Fig. 1(b).

3. State estimator

A dynamic system is typically described by a vector of the internal state, x, a state equation, $\dot{x} = f(x, u)$, and an output equation, $y = g(x, u)$. The state equation describes

* Tel.: +1-312-503-1407; fax: +1-312-503-5101.
E-mail address: karniel@northwestern.edu (A. Karniel).

PII: S0893-6080(02)00020-5

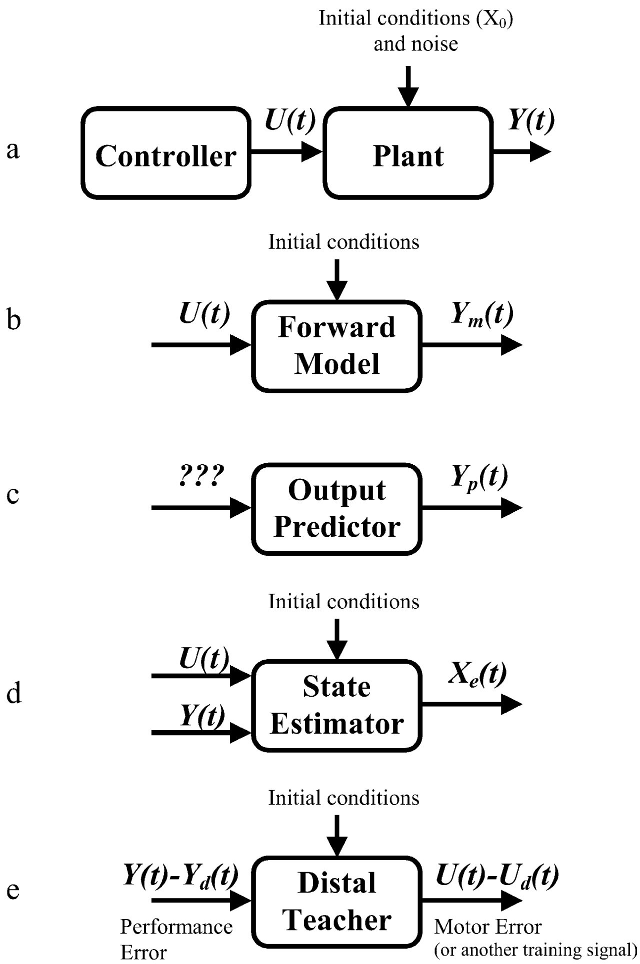

Fig. 1. The control scheme, a forward model, and three structures that use a forward model. (a) The controller and the controlled system (the plant). (b) A forward model of the plant. (c) An output predictor. (d) A state estimator. (e) A distal teacher that use performance error in order to train the controller. The hypothesis that structure b exists is typically supported by the evidence for structures c, d or e.

the change in the state variables as a function of the current state and the input. The output equation describes the output as a function of the state and possibly of the input. In this framework, a forward model should first calculate the internal state based on the initial conditions and the input, and then calculate the output. In the presence of noise and uncertainties, the first step is called state estimation. The estimation of the state could significantly improve with some knowledge of the output. The problem of state estimation is shown in Fig. 1(d), i.e. find the best estimate of the internal state using the inputs as well as outputs (partial or complete). The study of Wolpert, Ghahramani, and Jordan (1995) demonstrates remarkable similarity between simulated optimal state estimator (a Kalman filter) and human motor performance. Since a Kalman filter includes a forward model of the system, this study is frequently cited to support the forward model hypothesis, nevertheless, the salient point in optimal state estimation is the proper combination of input and output information, an issue that has nothing to do with the forward model hypothesis as shown in Fig. 1(b).

Another study that is cited in support of the forward model hypothesis describes a detailed computational model that includes both inverse and forward models in order to account for experimental data of motor adaptation (Bhushan & Shadmehr, 1999). In that study, the forward model uses the control signal as well as a delayed output signal in order to predict the output. It is important to carefully consider the role of the delayed output, since this branch is actually a form of feedback control. The fact that a delayed feedback can contribute to the explanation of the observed data should not be considered as evidence for the forward model hypothesis. Indeed a forward model could help in the proper usage of input and output information. Nevertheless, the use of delayed output signal is not consistent with the basic forward model hypothesis as shown in Fig. 1(b).

4. Distal teacher

A forward model was also proposed as a means for learning the inverse model controller. The important element in this argument is the transformation of the error from the output coordinates to the control coordinates (Fig. 1(e)). There are at least two good reasons to use the output error (also called performance error) in order to train the controller. One is the non-convexity problem that typically occurs in redundant systems (Jordan & Rumelhart, 1992; Mussa-Ivaldi & Hogan, 1991). Another problem occurs at the presence of noise. In this case, certain types of direct inverse learning could yield wrong result even for a non-redundant system (Karniel, Meir, & Inbar, 2001). One solution could be an indirect learning, that is, learning a forward model and than inverting it analytically. This method is sometimes referred to as indirect self-tuning adaptive control method. Another method is to learn a forward model and then backpropogate the error from the performance (output) error to the motor error (Jordan & Rumelhart, 1992). This computational argument suggests that the structure of a forward model should be either amenable to inversion or to backpropogation of error. The arguments for proper usage of error (Fig. 1(e)) supports the forward model hypothesis, nevertheless, the problems described above might be solved without a forward model.

5. Final remark

A prominent scientific strength of any internal model hypothesis should be the possibility to predict the range of learnable tasks. Therefore, it is suggested that any mention of the term internal model would accompany a clear definition of three elements: (i) the input space, (ii) the output space, and (iii) the structure of the model. Once these elements are defined it is worthwhile to discuss evidence for this internal model in terms of neuronal pathways, applications, and most importantly generalization capabilities, which is an essential property of any learning system, artificial and natural.

Acknowledgements

This commentary was first conceived and presented at a Robotics lab meeting in the Rehabilitation Institute of Chicago. It was written after numerous discussions with many colleagues that provided useful comments. In particular, the author wishes to thank Sandro Mussa-Ivaldi. The author was supported by NIH grant 5 P50 MH48185.

References

Bhushan, N., & Shadmehr, R. (1999). Computational nature of human adaptive control during learning of reaching movements in force fields. *Biological Cybernetics*, *81*, 39–60.

Blakemore, S. -J., Wolpert, D. M., & Frith, C. (2000). Why can't you tickle yourself? *NeuroReport*, *11* (11), R11–R16.

Flanagan, J. R., & Wing, A. M. (1997). The role of internal models in motion planning and control: Evidence from grip force adjustments during movements of hand-held loads. *Journal of Neuroscience*, *17* (4), 1519–1528.

Guenther, F. H., & Micci Barreca, D. (1997). Neural models for flexible control of redundant systems. In V. Sanguineti, *Self-organization, computational maps, and motor control* (pp. 383–421). North Holland: Elsevier.

Jordan, M., & Rumelhart, D. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science*, *16*, 307–354.

Karniel, A., Meir, R., & Inbar, G. F. (2001). Best estimated inverse versus inverse of the best estimator. *Neural Networks*, *14*, 1153–1159.

Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, *9*, 718–727.

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks*, *9* (8), 1265–1279.

Mussa-Ivaldi, F. A., & Hogan, N. (1991). Integrable solutions of kinematic redundancy via impedance control. *The International Journal of Robotics Research*, *10* (5), 481–491.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, *269* (5232), 1880–1882.