ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 89-94

The Problem of Adaptive Control in a Living System or How to Acquire an Inverse Model Without External Help

Karl Theodor Kalveram [1] and Thomas Schinauer

Heinrich-Heine-University, Institute for Experimental Psychology
Section Cybernetical Psychology und Psychobiology
Universitaetsstr.1, 40225 Duesseldorf/Germany
e-mail: kalveram@uni-duesseldorf.de

Keywords: inverse modelling, Hebbian learning, auto-imitation, motor control

Abstract Recent research uncovers that goal directed sensorimotor behaviour is governed by negative feedback of positional error, and by feedforward through inverse modelling of the limb's dynamics. Thereby, forward models seem to provide the kinematic state of the limb. The question addressed in the paper is, how the neural network representing the inverse model can be trained. Because in this case an error based learning algorithm seems to be unavailable, an alternative non error based method called auto-imitation is proposed. It is demonstrated, that, if combining a special type of neural network (the power net) with a modified type of a Hebbian synapse, the inverse dynamics of an onejointed arm can be precisely identified using auto-imitation. This holds for a simulated arm and a real robot arm as well.

1. Introduction

Living on shore confronts with the problem of defending posture against gravition while performing voluntary movements. Nowadays, the general solution suggested for those sensorimotor problems are formulated in terms of control theory which pins negative feedback and feedforward by internal models (inverse and forward models) as the essential units thought to govern movement control. Fig.1 shows a typical diagram characterizing the state of affair (Sabes, 2000).

A very crucial point in this discussion is, how these inverse and forward models can principally be acquired by neural networks. With respect to sensorimotor behaviour, living systems don't have an external teacher. Who else does provide the error necessary to train and update respectively the neural networks representing such internal models? Regarding the forward model, the answer is easy: The actual state of the arm, sensorily picked up, can be taken as the target value, and the difference between this target and the actual output of the respective network is the error, which can be used to train the network by supervised (i.e. error based) learning.

The model of fig.1 which primarily serves to explain principles of motor control seems also to provide a recipe „to drive adaptation of the inverse model" (Sabes 2000 p. 742): The „current motor error" E' as indicated by the dashed arrow crossing through the inverse model appears to determine also the inverse model's updating. However, E' does not represent the true network error, because it is not

[1] Supported by grants 417/18-3 and 417/24-1 from Deutsche Forschungsgemeinschaft (DFG)

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 89-94

expressed in terms of the network‘s output U_{ff}. Furthermore, E‘ is determined not only by the state feedback caused by the output U_{ff} of the - possibly maladapted - inverse model, but also by the signal induced by the negative feedback controller’s output U_{fb}. Therefore, if the network representing the inverse model would be trained using that E‘ it would be trained with a faulty error signal. This cannot be a good solution for the problem to acquire the inverse model, nevertheless there seem to exist training methods, for instance learning by „feedback error“ (Kawato, 1990), which finally – after thousands of training trials – may lead to an acceptable inverse model.

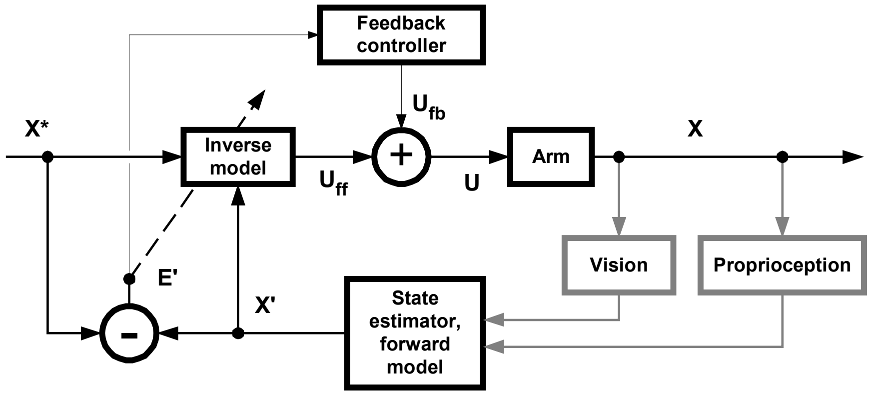

Figure 1: „Schematic diagram of the internal models used in controlling multijoint arm movements. The desired motion of the Arm X* is fed into an inverse model of the arm, which acts as a feedforward controller, producing a command signal U_{ff}. There is also evidence for late influence from a feedback control signal, U_{fb}. The two commands are combined, through simple addition perhaps, to yield the final control signal U. The feedback pathway is illustrated in grey to reflect the fact that it plays a subordinate role as a result of feedback delays. The true state of the arm, X, is estimated with a combination of the visual and proprioceptive feedback. The state estimate, X‘, serves as input to the inverse model and is also compared to the desired state to yield an estimate of the current motor error (E‘). The latter signal is used both in feedback control and to drive adaptation of the inverse model (dashed arrow).“ (Sabes, 2000).

2. Auto-Imitation as a learning algorithm not based on error

An alternative solution to that of fig.1 is to apply a non error based algorithm for learning the inverse model. An appropriate algorithm, as proposed by Kalveram (1981) or Jordan (1988), can be developed in the framework of the re-afference principle (Holst and Mittelstaedt 1950). The algorithm, later on also called „auto-imitation“ (Kalveram, 1992), is outlined in fig.2 using an onejointed arm movable in a plane as given in (1). In the learning phase, the output of the network going to learn the inverse model is not connected to the input u of the plant. The plant‘s input are arbitrary torques u_{fb} caused by the pattern generator operating like a "blind teacher". This makes the arm attain the related values of angular acceleration, velocity and position x at the elbow joint according to the differential equation

$$\boldsymbol{J}\cdot\ddot{\boldsymbol{x}}(\boldsymbol{t})+\boldsymbol{B}\cdot\dot{\boldsymbol{x}}(\boldsymbol{t})-\boldsymbol{C}\cdot\sin(\boldsymbol{x}(\boldsymbol{t})-\boldsymbol{x}_g)=\underbrace{\boldsymbol{K}\cdot\left(\boldsymbol{x}_0(\boldsymbol{t})-\boldsymbol{x}(\boldsymbol{t})\right)}_{\boldsymbol{u}_{fb}}+\boldsymbol{Q}(\boldsymbol{t}),\ \text{with}\ \ \boldsymbol{C}=\boldsymbol{m}\cdot\boldsymbol{g}\cdot\boldsymbol{a} \quad (1)$$

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 89-94

with **J**= moment of inertia, related to the joint, **B**= coefficient of viscous damping, **K**= joint stiffness of the arm, $\mathbf{x_0}$= angle defining the momentary mechanical equilibrium point, **m**= mass of the arm, **a**= distance between center of mass and joint, **Q**= disturbing torque from non-systematic external sources, **g**= graviational constant (9,81 m/s^2 resp. N/Kg), $\mathbf{x_g}$= angle between direction of gravitation and body axis, here assumed as zero .

Because of the sine in (1) the arm exhibits a non-linear behaviour.

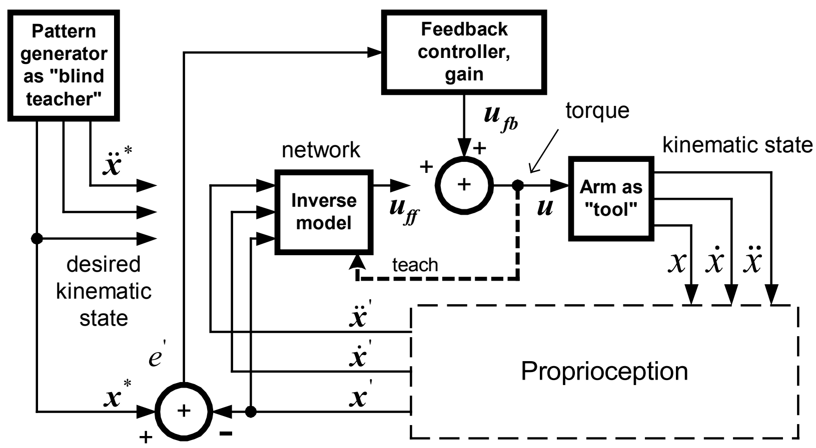

Figure 2: Neural controller learning an inverse model of the plant (onejointed arm) by **auto-imitation**. The plant is given by its tool transformation (see equation (1)); i.e., the rule governing the behaviour of the plant in the forward direction (from torques to kinematic states). Notice that auto-imitative learning does not require a forward model, and obviously needs also no movement error.

The kinematic outputs (angular position, velocity and acceleration) are proprioceptively measured and fed back into the network's regular inputs. The torques, also measured proprioceptively, are simultaneously offered also to the teaching input (dashed arrow). These torques function as targets to be attained after learning. This setup will enable the synapses of the network (see below) to attain the right strengths. After training, the output u_{ff} of the network is linked with the summing junction, and the output of the feedback controller is disconnected from the teaching input of the network, but remains connected to the arm’s input via the summing junction. Also position, velocity and acceleration signals are disconnected from the network and connected to the related desired outputs of the pattern generator. Now, the pattern generator’s output, when passing the network, will cause torques forcing the arm to attain the desired movement. So long as the inverse model being learned in this manner is correct, the negative feedback loop is uncommitted, and the movement is controlled solely by feedforward generated through the inverse model.

3. The power network used to represent the inverse model

The network representing the inverse model must operate very precisely. For this reason, the "power network" (Kalveram, 1993) can be used. This is a three layer SIGMA-PI network with feedforward architecture, fixed synaptic weights in the hidden layer, and plastic weights in the output layer. The hidden neurons multiply potentiated variables coming from the input layer, and the output neurons then compute a weighed sum of these products. This feature maps that the central nervous system is both additive and multiplicative, as given in sensory gating and descending modulation (Gossard and Rossignol, 1990). If M=number of output neurons and K=number of input neurons, the power network represents M abbreviated K-dimensional power series, or even Taylor series, known to approximate any function to every required degree of precision solely by choosing appropriate coefficients, respectively by determining the plastic synaptic weights. For that, a modified type of Hebbian learning can be administered which guaranties convergence of the weights (Kalveram 1997, 2000). In the present investigation, such a power network was used, established with 3 input neurons, 15 hidden nodes representing the powers 1 to 5, and one output neuron, connected with the hidden neurons by modified Hebbian synapses. An outline of the network is given in fig.4 representing a simplified type of power net (mixed terms are omitted).

4. The modified type of Hebbian learning

The modified type of Hebbian learning (Kalveram 2000) is outlined in fig.3 in terms of the neuron model of McCulloch and Pitts. The transfer characteristic between the postsynaptic potential of the neuron and its output is assumed to be linear.

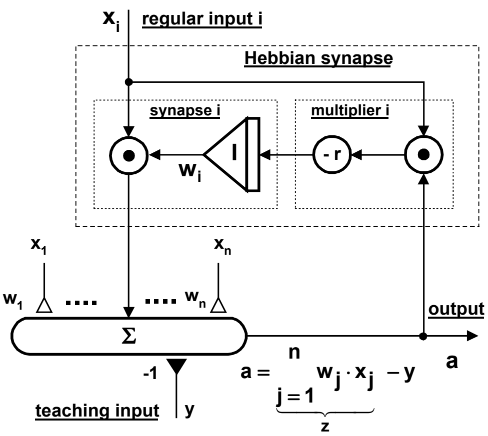

Figure 3: Modified Hebbian synaptic learning. During learning the integrator **I** accumulates the products put out by the multiplier i. After learning, the connection between the output a and the synapse i is cut, and y is set to zero. Now **I** holds the output value w_i representing the synaptic strengths. All other n synaptic units operate in the same manner. If learning ends with $a \rightarrow 0$, for all combinations of values of the input variables $x_1, ..., x_n$ the output **a** now automatically attains the desired values of **y**.

The teaching input neither enforces the neuron to attain a post-synaptic potential equal to y, nor reinforces the synaptic state in case of a positive outcome signalled by a supervisor, as proposed by Pennartz (1997). The intracellular postsynaptic potential itself mirrors the error z-y which has to be minimized (relaxation principle). Referring to adaptive signal processing, the neuron in Fig.3 is an "adaptive linear combiner" capable of "adaptive linear filtering" (Widrow and Stearns, 1985). When using finite steps, the adaptation minimizes the error step by step in a least mean squares steepest-descent manner, with r regulating the decrease of error per step.

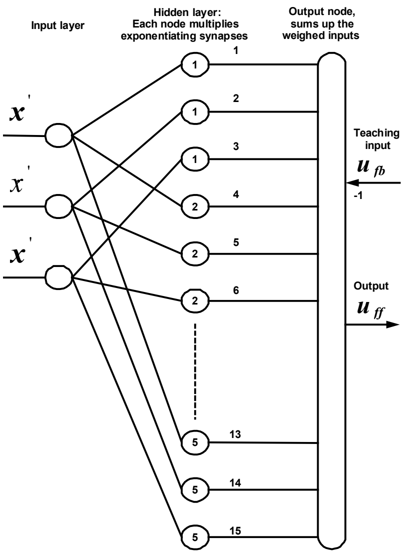

Figure 4: Simplified power network realizing modified Hebbian learning. The numbers in the nodes of the hidden layer indicate the exponents. The 15 weights a_{ip} of the synaptic contacts to the output neuron are to be determined by the learning rule indicated in fig.3.

Denoting estimated angular acceleration , velocity and position by x_1, x_2, x_3, the network mirrors the equation

$$u_{ff} = \sum_{p=1}^{5} \sum_{i=1}^{3} a_{ip} x_i^p - u_{fb} ,$$

where $\{a_{ip}\}$ is a 3 by 5 matrix of coefficients respresenting the synaptic weights. That means, that all mixed terms of typ

$$x_i^p \cdot x_j^q \cdot x_k^r \;, 0 \le p+q+r \le 5,\; 1 \le i,j,k$$

with different indices p, q, r are discarded.

5. Results of simulated and real arm movements

In a simulation with Matlab/Simulink software, the arm was represented by (1). The network of fig.4 was trained according to the auto-imitation algorithm of fig.2 and modified Hebbian learning of fig.3. The result was a nearly perfect inverse model. The same procedure was applied to a real robot arm. Fig.5 shows that also in this case the physical properties of the arm were reproduced very well by the neural controller.

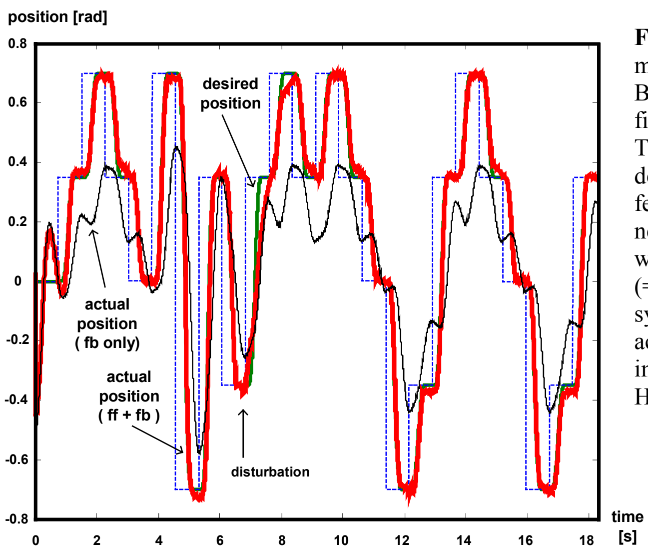

Figure 5: Movement of a mechanical lever (J=0.081, B=0.3) in a gravitational force field, driven by a torque motor. Torques -and trajectories- are determined by negative feedback (=fb) only, or by negative feedback combined with the levers's inverse model (=ff+fb) whose parameters (i.e. synaptic weights) have been acquired previously by auto-imitation plus modified Hebbian learning.

In fig.5, the dotted steps represented the goal sequence which was converted into the desired positional trajectory by the pattern generator. At the beginning and after about 7s the inverse controlled movement was shortly disturbed. This activated the negative feedback controller, which ironed out the distortion in less than 2s. Except for the disturbations, desired and actual positions in the ff+fb condition nearly coincided.

At the moment, an experiment is conducted with a twojointed robot arm. We expect, that, according to the theoretical solution given in Kalveram (2000), also this limb system can be physically identified and controlled by the methods outlined above.

References

Gossard, J.P. and S. Rossignol, 1990. Phase-dependent modulation of dorsal root potentials evoked by peripheral nerve stimulation during fictive locomotion in the cat. Brain Research 537, 1-13

Holst, E. von and H. Mittelstaedt, 1950. Das Reafferenzprinzip. Naturwissenschaften 37, 464-476

Jordan, M.I. (1988) Supervised learning and systems with excess degrees of freedom. COINS Technical Report 88/27, 1-41

Kalveram, K.Th., 1981. Erwerb sensumotorischer Koordinationen unter störenden Umwelteinflüssen: Ein Beitrag zum Problem des Erlernens von Werkzeuggebrauch. (Engl. titel: Acquisition of sensorimotor co-ordinations under environmental disturbations. A contribution to the problem of learning to use a tool) In: L. Tent (Ed.): Erkennen, Wollen, Handeln. Festschrift für Heinrich Düker (S. 336-348). Göttingen: Hogrefe

Kalveram, K.Th., 1992. A neural network model rapidly learning gains and gating of reflexes necessary to adapt to an arm's dynamics. Biological Cybernetics 68, 183-191

Kalveram K.Th., 1993. Power series and neural-net computing. Neurocomp. 5, 165-174

Kalveram K. Th., 1998. Wie das Individuum mit seiner Umwelt interagiert. Psychologische, biologische und kybernetische Betrachtungen über die Funktion von Verhalten (engl. titel: How the individual interacts with its environment. Psychological, biological and cybernetical considerations about the function of behaviour). Lengerich: Pabst

Kalveram K. Th. 1999. A modified model of the Hebbian synapse and its role in motor learning. Human Movement Science, 18, 185-199.

Kalveram, K.Th., 2000. Sensorimotor Sequential Learning by a Neural Network Based on Redefined Hebbian Learning. In H. Malmgren, M. Borga & L. Niklasson (Eds) Artificial Neural Networks in Medicine and Biology. London: Springer, pp.271-276

Kawato, M. 1990. Feedback-error-learning neural network for superwised learning. In R. Eckmiller (Ed.), Advanced neural computers (pp. 365-372). Amsterdam: North-Holland.

Pennartz, C.M.A. 1997. Reinforcement learning by Hebbian synapses with adaptive thresholds. Neuroscience 81, 303-319

Sabes, P.N., 2000. The planning and control of reaching movements. Current Opinion in Neurobiology, 10, 740-746

Widrow, B. and S.D. Stearns, 1985. Adaptive signal processing. Englewood Cliffs: Prentice-Hall