Towards the Advantages of Hierarchical Anticipatory Behavioral Control

Oliver Herbort (oliver.herbort@stud-mail.uni-wuerzburg.de)
Martin V. Butz (butz@psychologie.uni-wuerzburg.de)
Joachim Hoffmann (hoffmann@psychologie.uni-wuerzburg.de)
Department of Cognitive Psychology, Röntgenring 11
University of Würzburg
Würzburg, 97070, Germany

Abstract

Despite recent successes in control theoretical programs for limb control, behavior-based cognitive approaches for control are somewhat lacking behind. Insights in psychology and neuroscience suggest that the most important ingredients for a successful developmental approach to control are anticipatory mechanisms and hierarchical structures. Anticipatory mechanisms are beneficial in handling noisy sensors, bridging sensory delays, and directing attention and action processing capacities. Moreover, action selection may be immediate using inverse modeling techniques. Hierarchies enable anticipatory influences on multiple levels of abstraction in time and space. This paper provides an overview over recent insights in anticipatory, hierarchical, cognitive behavioral mechanisms, reviews previous modeling approaches, and introduces a novel model well-suited to study hierarchical anticipatory behavioral control in simulated as well as real robotic control scenarios.

Introduction

The autonomous control of the own body is an essential challenge for any cognitive system. Although established behavioral control in animals and humans seems effortless in every day life, many challenges arise. Due to the complex, dynamic, time lagged, noisy, and often nonlinear interactions between body and environment, effective body control in real environments is hard. Movements of different body parts influence each other, actions have different outcomes in different situations, muscle forces are state-dependent, etc. Furthermore, sensory information may be unavailable, as for example in darkness, or may be available to the brain only after a significant time delay. The brain has to learn these complex, often context-dependent, interactions to be able to induce effective adaptive body control.

The notion that most actions are goal directed and that the goal state is represented before the action is performed is labeled the *ideomotor principle* and can be traced back to the 19th century (Herbart, 1825; James, 1890). Although behaviorists later questioned this view, it is now widely accepted that behavior is in most cases goal oriented. Hoffmann (1993) emphasized this insight in his theory of anticipatory behavior control, which theorizes that actions are usually preceded by an anticipatory image of the sensory effects. The image triggers that action(s) that is (are) expected to yield the anticipated effects, considering the current environmental circumstances. Different sensory modalities and sensory aspects can influence action triggering, for example, an external effect, like a tone, or also a proprioceptive effect, like the feeling of bending the fingers or of pressure against the fingertips. To control more complex behavior, actions may be divided into simpler parts. For example, if a piano player wishes to play a tone, the anticipation of the tone causes the anticipation of the feeling of the correct hand position and then the finger pressing the key. Thus, to achieve an overall goal, several successive sub-goals may trigger successive actions.

To be able to generate such complex behavior effectively, hierarchical processes are necessary that generate goals and partition far-reaching goals into suitable sub-goals. However, even if neuroscience shows that brain functions are structured hierarchically (e.g. Poggio & Bizzi, 2004), only few computational arguments exist, *why* such structures are advantageous.

This paper reviews evidence for anticipatory guided human processing and derives design suggestions for cognitive behavior systems. Similarly, we assess evidence for hierarchically structured mechanisms. The gained insights lead us to the development of a simple learning system for studying the potential benefits of hierarchical anticipatory control structures. We introduce the base model and confirm successful behavioral control of a simple arm. In sum, this paper studies anticipatory hierarchically controlled systems that learn effective control structures to guide complex adaptive behavioral patterns.

The remainder of this work is structured as follows. First, we review anticipatory and hierarchical cognitive structures. Next, existing cognitive control models are compared. Finally, we introduce our model revealing its current capabilities, limitations, and potentials. A short discussion concludes the paper.

Anticipatory Hierarchical Structures

In this section, we gather evidence for and benefits of anticipatory and hierarchical structures in learning, behavioral control, and cognition in the broader sense.

Anticipatory Behavior Control

Anticipatory behavior refers to behavior in which currently desired goals precede and trigger the action that usually results in the desired goals. Psychological experiments underline the concept of anticipatory behavior.

A simple experiment confirms the presence of effect representations before action execution. Kunde (2001) paired actions with compatible or incompatible effects, such as the presentation of a bar on the left or on the right compatible or incompatible to a left or right key press. Although the effects were presented only after the key press, reaction times were significantly faster, when the location of target button and visual effect corresponded. Similar effects were found for the modalities of intensity and duration (Kunde, Koch & Hoffmann, 2004). Additionally, (task-irrelevant) stimuli that are incompatible to the expected effects of an action interfere with the selection and initiation of the action (Elsner & Hommel, 2001).

In all cases, it is concluded that anticipatory effect representations interfere with an action code or also with an external stimulus. Thus, goal aspects are represented before action execution in terms of at least some of the sensory effects. Interestingly, it has also been shown that humans acquire such action-effect associations much easier than situation-action relations (Stock & Hoffmann, 2002).

However, the advantages of such anticipatory behavior remain somewhat obscured. What are the benefits of representing effects before or actually for action execution? Other disciplines provide interesting insights in this respect. Artificial intelligence shows that anticipatory representations enable higher flexibility in learning and decision-making. In reinforcement learning (Sutton & Barto, 1998), the DYNA architecture (Sutton, 1990) showed that model-based reinforcement learning mechanisms increase flexibility when goals vary or when the environment is partially dynamic. More recent investigations in relational reinforcement learning have shown similar advantages when the flexible propagation of reinforcement learning is required (Kersting, Van Otterlo & De Raeth, 2004).

In control theory, structures capable of predicting future states yield more powerful controllers. Forward models that predict the effects of action compensate for delayed or noisy feedback (Barlow, 2002; Haykin, 2002; Kalman, 1960; Miall, Weir, Wolpert & Stein, 1993). Additionally, *inverse models* (IMs) that directly determine the action necessary to obtain a desired goal are key components in efficient adaptive controllers (Kawato, Furukawa & Suzuki, 1987).

Thus, cognitive psychology and neuroscience suggest that anticipations are important for effective adaptive learning systems. Artificial intelligence and control theory have shown that anticipatory structures improve learning speed and reliability, behavioral flexibility and execution, and sensory robustness, resulting in effective goal-directed systems.

Hierarchies for Learning and Control

Besides the anticipatory indicators, studies and models suggest that cognitive information is processed hierarchically. Powers (1973) already stressed the importance of hierarchies in behavioral control and consequent computational models of cognitive systems. Just recently, Poggio and Bizzi (2004) pointed out that hierarchical structures are very likely the key to not only sensory processing but also motor control. Available hierarchical models in vision (Giese & Poggio, 2003; Riesenhuber & Poggio, 1999) are suggested to be extended to motor control. Hierarchical top-down influences showed to have advantageous structuring effects (Rao & Ballard, 1999).

Kawato, Furukawa and Suzuki (1987), applied a hierarchical controller to a robot arm. The lowest level contains a simple PD-controller that can in principle handle any task. The controller is not very efficient, because the delayed feedback results in a slow control process. A second layer improves performance. As soon as a forward model of the plant is learned, it updates the control signal using the expected feedback, which is available much faster. However, it is still necessary to adjust the signal iteratively. A third level consists of an inverse model (IM) that calculates a control signal for any given goal. When the IM is accurate, the controller selects a feasible control signal instantly. In case of a failure, the lower levels induce the (slower and less effective) control. The more accurate the models in the higher levels, the more they influence the control signals.

Despite the ubiquitous hints on the importance of hierarchical processing and the first model from Kawato and colleagues, it remains somewhat unclear why hierarchies are advantageous. One advantage may be the general decomposability of our environment due to time and space constrains (Simon, 1969; Gibson, 1979). Computational advantages can be found in artificial intelligence studies.

Re-considering reinforcement learning, it has become clear that hierarchical processing mechanisms are mandatory for effective reward propagation and flexible learning (Barto & Mahadevan, 2003). Hierarchical structures are formed that can trigger abstract action representations including goals. Most recent publications have shown that such hierarchical representations may be learned by using simple statistics of the environment searching for decomposable sub-structures (Butz, Swarup & Goldberg, 2004; Simsek & Barto, 2004).

Thus, hierarchical control enables the discovery and representation of more distant and abstract dependencies as well as increases flexibility in behavioral learning and decision making, as well as in sensory processing at different levels of abstraction in time and space.

Merging Both

As we have seen, cognitive processing is guided by anticipations that improve sensory processing and behavioral control. Hierarchies yield more flexible representations for anticipatory learning and behavior. Thus the combination of anticipatory and hierarchical structures may be a promising approach to understand and model human motor learning and control.

The review suggests several requirements for a cognitive controller. First, the controller must represent a goal in terms of desired sensory inputs. Partial, underspecified, and even contradicting goals may be represented in different sensory modalities. Second, goal representations should not only be modular but also hierarchical. Higher level goal

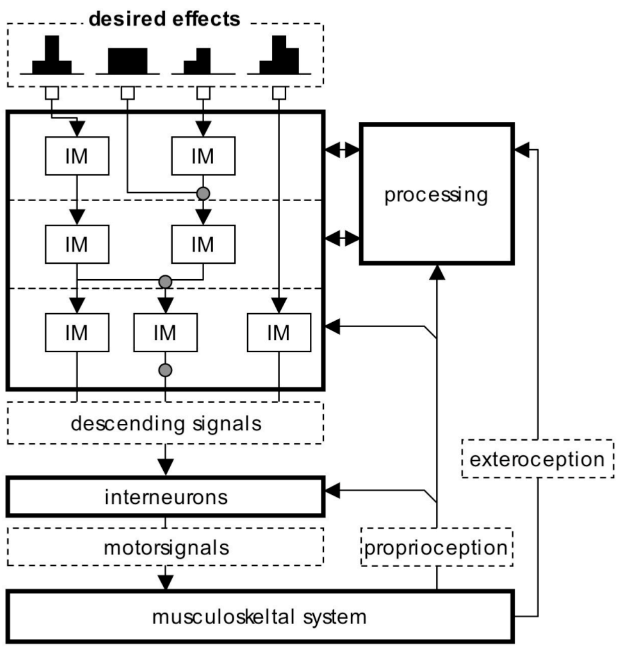

Figure 1: The left side of the chart shows how the desired effects are transformed into motor signals using a hierarchy of inverse models (IM). The motor signals cause changes in the arm. The changes are fed back to the controller (right side). Perceptions can be used directly or after processing.

representations are usually more abstract in time and space and trigger lower level, more concrete, sensory dependent goal representations. Third, the representations should be learned by interacting with the environment. Learning architecture and learning biases, however, are provided in advance.

Before we introduce our model, which can satisfy these constraints, we review other related systems.

Cognitive Movement Controllers

Numerous computational models for motor learning and control have been proposed. Most of them address specific stages of movement generation, for example trajectory formation (Cruse, Steinkühler & Burkamp, 1998; Hirayama, Kawato & Jordan, 1993) or coordinate transformation (Salinas & Abbott, 1995). Others are tracking reference signals, relying on IMs and feedback controllers (Kalveram, 2004; Kawato, Furukawa & Suzuki, 1987), which might be combined in a single control structure (Stroeve, 1997). Some approaches gate a number of single control structures to be able to quickly adapt to changing limb properties (Wolpert & Kawato, 1998; Haruno, Wolpert & Kawato, 2001).

While each model has interesting properties on its own, none match all the suggested cognitive systems requirements. The described hierarchical model of Kawato, Furukawa and Suzuki (1987) contains three different levels but does not accept goals in arbitrary modalities. Other controllers (e.g. Cruse, Steinkühler & Burkamp, 1998) accept underspecified goals but do not include hierarchical layers. Many models contain neural networks that learn by cognitively implausible mechanisms like back-propagation. Our model intends to bridge the respective drawbacks effectively creating a hierarchical, anticipatory cognitive model that is suitable to process any goal representation flexibly and hierarchically.

A Hierarchically Anticipatory Model of Motor Control

We devise a new computational model for motor learning and control (Figure 1). To account for anticipatory control, a central controller responds to many different goal modalities. IMs are used to store action-effect relationships. The controller does not directly address the musculoskeletal system but the interneurons of the spinal cord that interact with each other, for example by inhibiting antagonistic motoneurons. The interneurons activate alpha and gamma motoneurons that innervate certain muscles attached to a mechanical arm model. This causes sensors for muscle length, velocity and tension to change their output. These proprioceptions are linked to spinal interneurons and to the central controller and are available after a short time delay or after some more processing. For example, length of several muscles that embrace a joint can be used to represent the joint angle. As the state of the limb can be perceived visually, too, information like the hand position in external space is also available to the controller.

The Controller

The task of the inverse controller is to reach or maintain a certain state. A set of desired sensory goals is provided to the controller and must be transformed in descending signals that innervate the spinal cord.

The general controller consists of a set of IMs. Given a desired goal, each IM calculates an action dependent on the current situation. An action can either be an actual control signal sent to the spinal cord or a goal input signal for other IMs. Thus, a single IM connects with three sets of signals. The first set describes the desired state or goal. The second one codes relevant sensory signals consisting of direct proprioceptions, exteroceptions, and/or pre-processed, internal state signals. The third set sends information to subsequent IMs or to the spinal cord. For example, to press a key, visual signals may provide a desired fingertip position in external space. The signal then may be transformed to appropriate joint angles, which are then transformed to signals that descend to the spinal cord. Several signals may be merged using weighting mechanisms according to estimated importance.

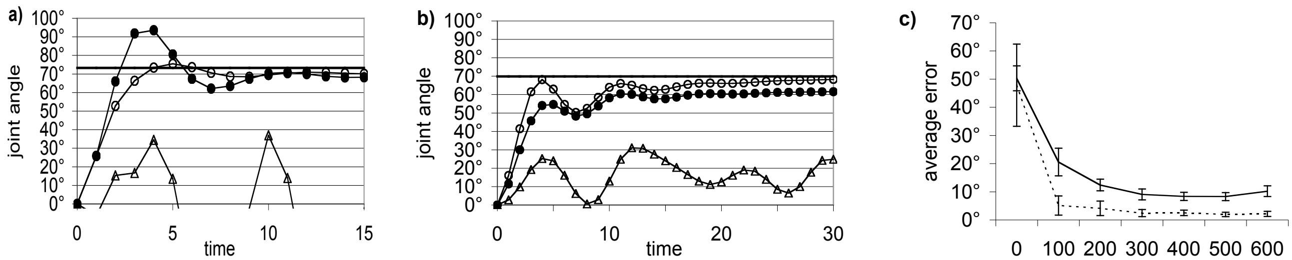

Figure 2: a) Inverse models that compute position and speed can reach a specific angle (in this example: 70°) fast improving the approximation with learning (white triangles: naïve; black circles: after 100 examples; white circles: 600 examples). b) Inverse models that process error signals approximate a target slower but more precise (same learning steps). c) On average (n=10, error bars indicate standard deviations), the error-based model (dotted line, setting as in (b)) is more accurate than the model that generates fast movements (straight line, setting as in (a)).

Each IM learns by observing which effects arbitrary random actions have dependent on the current sensory state. Thus those situation-action-effect relations are stored, which occur during the application of initially random motor signals. The mechanism is sometimes referred to as *direct inverse modeling* (Jordan & Rumelhart, 1992) or *autoimitation* (Kalveram, 2004). Initial random commands must be sufficient to learn coarse IMs. Once the controller is sufficiently accurate, control signals can be used to systematically explore and improve relevant movements. Note that the time interval may vary between a situation, an action and the effect that is linked to it. Thus, some IMs can be used to represent what will happen shortly after a command signal is changed while others store longer term dependencies. IMs that bridge longer term intervals can produce control signals for subsequent shorter term IMs.

The IMs are realized by associative networks using Hebbian learning (Hebb, 1949). The real-valued sensory inputs are split up using radial basis function (RBF) input neurons, which restrict the receptive field of each neuron. A neuron is activated if the actual signal values are close to a preferred combination of values. Hebbian learning strengthens connections between situations, actions and coincidental sensory effects. If later a goal is chosen, only the desired effect has to be fed into the network consequently activating the appropriate action.

Model Evaluation

To consider the feasibility of our approach, we evaluate the performance of a single IM on a simple 1-dof-arm model. The arm is dampened and a nonlinear restoring force pulls it to its initial position. To move the arm, a motor signal is proportionally transformed into a torque that is applied to the joint. Thus, applying a constant torque signal moves the arm to a certain equilibrium position after some oscillations. Two different types of IMs were tested. The first type converts desired joint angles and velocities into a torque considering current joint angle and torque. Receptive fields of sensory states (desired or actual) are distant 10° of joint angle from each other and 5°/time-unit of joint velocity. Standard deviations are set to half the distance. The receptive fields of the torque signal are .1 apart (a torque change of .1 causes a change in joint angle of about 7°-13° depending on the arm position). To convert the activation of RBF-neurons into a torque signal, a modified winner-takes-all mechanism is used that takes the activation of the winner-neighboring RBF-neurons into account.

Each learning example consisted of a torque signal, the current sensory state and the subsequent sensory state. Learning examples were generated by setting new random torque signals that were evenly distributed between -1 and 1 (allowing an arm range from about -80° to 80°). Every four time-units the signal changed. After five changes the arm was set back to its initial position. To test the IM, it had to reach different joint angles between -60° and 60°. The data shows that after a few examples, the IM is able to reach the vicinity of a desired goal very fast but its endpoint position is not very accurate (Figure 2a, c).

Another type of IM learns how signal changes were related to joint angle changes. The difference between actual and desired state (error) were provided as goals, the current absolute joint angle and torque signal were used as sensory input and the control signal was used to change the torque signal (receptive fields covering 10° for the error, 20° for absolute joint angle, .4 for absolute torque signal, .2 for the torque change). The same learning and testing procedures were used as for the first type. The data shows, that this type of IM is capable of reaching a goal very precisely, despite the widths of the receptive fields (Figure 2b, c).

Model Capabilities, Potentials, and Challenges

The results show that the two investigated connectivities have two interesting properties that may be combined. While the first controller was able to generate very fast goal-directed movements, the second was able to induce very accurate, albeit slower movements. The combination of both may yield an effective two-staged controller that approaches the target location quickly and then stops accurately at the exact target position (Hirayama, Kawato & Jordan, 1993). The integration of the combined model into the more sophisticated musculoskeletal system and the spinal cord model outlined above promises additional stabilizing effects.

A major point of critique on the current model may be the method used for learning the IM. Since the method is not goal-directed, the mapping is nowhere guaranteed to converge to the optimum (Jordan & Rumelhart, 1992). However, we believe that it is not necessary to obtain an optimally accurate mapping between action and effects in the general sense. Action execution is usually noisy and easily perturbed so that sensory feedback control is expected to be generally necessary to reach a precise goal. Additionally, trained movements become progressively more optimal, which is the case in our model.

Another concern is that the chosen RBF encoding is not very suitable for generalization. Using different layers of RBFs with a combination of larger and smaller receptive fields may solve this problem. However, the encoding also has advantages. The representation facilitates dealing with uncertainty (Knill & Pouget, 2004) and allows very flexible goal representations. A goal does not need to be exactly specified but a range of acceptable goal states or goal features can be presented to the network. This feature increases flexibility, which is advantageous for control (Todorov & Jordan, 2002). Additionally, the representation allows the encoding of many-to-many relationships. A final RBF-related concern may be the curse of dimensionality and the consequently exploding number of RBF neurons. However, an adaptive distribution of the RBF neurons and separating comparatively independent parts of the sensory space in different networks can reduce the amount of required neurons (Urban, 1998).

A big challenge arises considering the need to learn and execute motor programs. Currently the system state only changes, if the desired effects or sensory inputs change. Thus, very fast or complex movements are not possible. Two ways exist to integrate motor programs. First, it has been shown that neural circuits exist in the spinal cord of animals that generate specific motor signals to coordinate simple rhythmic behavior, like walking or swimming (Dietz, 2003). Thus, the model of the spinal cord could be extended to include such rhythmic pattern generators. Additional representations would be necessary to code the behavior caused by the pattern generators, such as representations of walking or moving forward, to be able to address the behavior with anticipations. A second way to include motor programs would be to delegate this task to higher structures that send continuously changing desired effects to the controller. The combination of both features may be able to learn rhythmic behavior combined with consecutive behavioral pattern changes, as appropriate.

In this paper we only presented results for one single controller. Experiments are in progress combining multiple controllers as outlined above. Two approaches need to be distinguished: parallel, modular combinations and hierarchical, abstracting combinations. Shadmehr and Brashers-Krug (1997) showed that human subjects are able to store many different controllers for different situations. For example, one controller could be trained for moving light objects and another for heavy objects. The weighted combination of both controllers then enables fast adaptation to movements with different weights. This feature can be added by using an array of controllers that are experts for a specific situation and are weighted accordingly (Haruno, Wolpert & Kawato, 2001). Hierarchically connected IMs might prove advantageous when different objects need to be moved. Although different weighting of lower level IMs is necessary to calculate descending commands from the desired joint angles, the relationship between external coordinates and joint angles stay constant. Thus, only parts need to be adapted to the current situation. Additionally, longer time delays in higher layers may be compensated for by lower level control structures. On the other hand, the different times integrated by different models may be used to facilitate more complex, longer term movements.

Besides the combination and extension of IMs, strongly noisy signals will require more elaborate processes. Forward models can be included in the processing of the sensory inputs to bridge temporary misperceptions, sensory failure, or noisy sensory inputs akin to Kalman filtering.

Summary and Conclusion

This paper has reviewed indications and benefits of anticipatory mechanisms and hierarchical structures in control processes. Both mechanisms are involved in human motor learning and control. While anticipatory mechanisms lead to direct action selections in inverse models and effective filtering mechanisms in forward models, the modular and hierarchical combination of such models promises to yield a more effective environmental representation increasing behavioral flexibility, adaptivity, and decision making.

The gathered potentials of combining both mechanisms into artificial cognitive systems promise fruitful future research. The proposed model provides a novel, integrative approach for studying such combinations. The generality of the proposed associative structures enables direct modular and hierarchical combinations. Future research will investigate the suitability and extendibility of our approach for the simulation of efficient cognitive learning systems in simulated and real robotic environments. Moreover, future research will further study the benefits of hierarchical, anticipatory behavior control, learning, behavior, and cognition in general using and extending the proposed model.

Acknowledgments

This work was sponsored by the German research foundation (DFG) under grant DFG HO1301/4-3 as well as from the Euopean commission contract no. FP6-511931.

References

Barlow, J. S. (2002). *The cerebellum and adaptive control*. New York, NY: Cambridge University Press.

Barto, A. G., & Mahadevan, S. (2003). Recent advances in hierarchical reinforcement learning. *Discrete Event Dynamic Systems*, 13, 341-379.

Butz, M. V., Swarup, S., & Goldberg, D.E. (2004). *Effective online detection of task-independent landmarks*. (IlliGAL report 2004002). University of Illinois at Urbana-Champaign, IL.

Cruse, H., Steinkühler, U., & Burkamp, C. (1998). MMC – a recurrent neural network which can be used as manipulable body model. *From Animals to Animats 5 (SAB 1998),* 381-389.

Dietz, V. (2003). Spinal cord pattern generators for locomtion. *Clinical Neurophysiology*, 114, 1379-1389.

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 29-50.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience*, 4, 179-192.

Haruno, M., Wolpert, D. M., & Kawato, M. (2001). MOSAIC model for sensorimotor learning and control. *Neural Computation*, 13, 2201-2220.

Haykin, S. (2002). *Adaptive filter theory* (Fourth Edition). Upper Saddle River, NJ: Prentice Hall

Hebb, D. O. (1949). *The organization of behavior*. New York: John Wiley.

Herbart, J. F. (1825) *Psychologie als Wissenschaft neu gegründet auf Erfahrung, Metaphysik und Mathematik. Zweiter analytischer Teil.* München, Germany: August Wilhelm Unzer.

Hirayama, M., Kawato, M., & Jordan, M. I. (1993). The cascade neural network model and a speed accuracy trade-off of arm movement. *Journal of Motor Behavior, 25*(3), 162-174.

Hoffmann, J. (1993). *Vorhersage und Erkenntnis*, Göttingen, Germany: Hogrefe.

James, W. (1981/1890). *The principles of psychology. Volume 2*. Cambridge, MA: Havard University Press.

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

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME-Journal of Basic Engineering,* 82, 35-45.

Kalveram, K. T. (2004). The inverse problem in cognitve, perceptual and proprioceptive control of sensorimotor behaviour: Towards a biologically plausible model of the control of aiming movements. *International Journal of Sport and Exercise*, 2, 255-273.

Kawato, M., Furukawa, K., & Suzuki, R. (1987). A hierarchical neural-network model for control and learning of voluntary movement. *Biological Cybernetics*, 57, 169-185.

Kersting, K., Van Otterlo, M., & De Raedt, L. (2004). Bellman goes relational. *Proceedings of the Twenty-First International Conference on Machine Learning*, 465-472.

Knill, D. C., & Pouget, A. (2004). The bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27, 712-719.

Kunde, W. (2001). Response-effetc compatibility in manual choice reaction tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 387-394

Kunde, W., Koch, I., & Hoffmann, J. (2004). Anticipated action effects affect the selection, initiation, and execution of actions. *The Quarterly Journal of Experimental Psychology. Section A: Human Experimental Psychology*, 57, 87-106.

Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein, J. F. (1993). Is the cerebellum a smith predictor? *Journal of Motor Behavior*, 25(3), 203-216.

Poggio, T., & Bizzi, E. (2004). Generalization in vision and motor control. *Nature*, 431, 768-774.

Powers, W. T. (1973). *Behavior: The Control of Perception*. Hawthorne, NY: Aldine de Gruyter.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2, 79-87.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience, 2*(11), 1019-1025.

Salinas, E., & Abbott, L. F. (1995). Transfer of coded information from sensory to motor networks. *Journal of Neuroscience, 15*(10), 6461-6474.

Shadmehr, R., & Brashers-Krug, T. (1997). Functional stages in the formation of human long term motor memory. *Journal of Neuroscience, 17*(1), 409-419.

Simon, H. A. (1969). *Sciences of the Artificial*. Cambridge, MA: MIT Press.

Simsek, Ö., & Barto, A. G. (2004). Using relative novelty to identify useful temporal abstractions in reinforcement learning. *Proceedings of the Twenty-first International Conference on Machine Learning*. 751-758.

Stock, A., & Hoffmann, J. (2002). Intentional fixation of behavioral learning or how R-E learning blocks S-R learning. *European Journal of Cognitive Psychology*, 14, 127-153.

Stroeve, S. (1997). A learning feedback and feedforward neuromuscular control model for two degrees of freedom human arm movements. *Human Movement Science*, 16, 621-651.

Sutton, R. S. (1990). Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. *Proceedings of the Seventh International Conference on Machine Learning*, 216-224.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Campridge, MA: MIT Press.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, *5(*11), 1226-1235.

Urban, J.-P., Bruessler, J. L., & Gresser J. (1998). Neural networks for visual servoing in robotics. In L. C. Jain & T. Fukuda (Eds.): *Soft computing for intelligent robotic systems*. Berlin, Germany: Springer-Verlag.

Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks, 11*(7-8), 1317–1329.