A unifying computational framework for motor control and social interaction

Daniel M. Wolpert[1*], Kenji Doya[2,3] and Mitsuo Kawato[2]

[1]*Sobell Department of Motor Neuroscience and Movement Disorders, Institute of Neurology, Queen Square, London WC1N 3BG, UK*
[2]*ATR Human Information Science Laboratories, and* [3]*CREST, Japan Science and Technology Corporation, 2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan*

Recent empirical studies have implicated the use of the motor system during action observation, imitation and social interaction. In this paper, we explore the computational parallels between the processes that occur in motor control and in action observation, imitation, social interaction and theory of mind. In particular, we examine the extent to which motor commands acting on the body can be equated with communicative signals acting on other people and suggest that computational solutions for motor control may have been extended to the domain of social interaction.

Keywords: motor control; social interaction; computational models; internal models; theory of mind

1. INTRODUCTION

Movement is the only way we have of interacting with the world, whether foraging for food or attracting a waiter's attention. Direct information transmission between people, through speech, arm gestures or facial expressions, is mediated through the motor system which provides a common code for communication. From this viewpoint, the purpose of the human brain is to use sensory representations to determine future actions. Moreover, in recent years the motor system has been implicated in many traditionally non-motor domains. An important idea is that the perception of the action of others, including speech, involves the motor system (Liberman & Whalen 2000). The proposal is that others' actions are decoded by activating one's own action system at a sub-threshold level and there appears to be a special neural mechanism for decoding such information. Recently, these ideas have gained empirical support in neuroscience with the finding of 'mirror neurons' that respond to both self-generated actions and the actions of others (Gallese *et al.* 1996; Rizzolatti & Arbib 1998; Gallese 2003). Human neuroimaging and magnetic stimulation studies have also shown that the areas associated with action are also active during imitation and observation (Fadiga *et al.* 1995, 2002; Iacoboni *et al.* 1999; Grezes *et al.* 2001). Moreover, pre-motor systems are activated when subjects view manipulable tools or even action verbs (Martin *et al.* 1996; Grafton *et al.* 1997). Such studies have brought the motor system to the forefront in the investigation of action interpretation and social interaction. In this paper, we explore the parallels between the computations that occur in motor control and in action observation, imitation, social interaction and

theory of mind. In particular, we examine the extent to which motor commands acting on the body can be equated with communicative signals acting on other people. We suggest that computational solutions that developed for motor control could have been extended to the domain of social interaction.

2. THE SENSORIMOTOR AND SOCIAL INTERACTION LOOPS

The study of motor control is fundamentally the study of sensorimotor transformations. We can view the motor system as forming a loop in which motor commands cause muscle contractions, with consequent sensory feedback, which in turn influences future motor commands (Wolpert & Ghahramani 2000) (figure 1*a*). The transformation from motor commands to their sensory consequences is governed by the physics of the musculoskeletal system, the environment and the sensory receptors. The descending motor command generates contractions in the muscles and causes the musculoskeletal system to change its configuration. However, the same motor command can have very different consequences in different situations. For example, the same motor command will generate less muscle contraction when the muscles are fatigued. Moreover, the same motor command can lead to very different changes in body configuration depending on the nature of the physical objects we interact with. To describe the variables that specify the configuration of the body, such as joint angles or hand position, we use the word *state*. In general, a state is a set of variables which vary over time and when taken together with fixed parameters of the system, such as the mass of body segments, and the equations governing the physics of the musculoskeletal system and the world are sufficient to predict the system's future behaviour. In general, the state, for example the set of activations of groups of muscles (synergies) or the position and velocity of the hand, changes rapidly and continuously

593
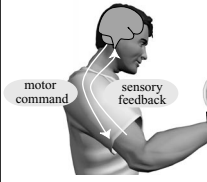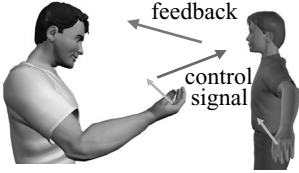
	motor control	social interaction
loop	(a)	(b)
control signal	motor command	communicative actions e.g. speech, gesture
consequences	change in my body's state	change in your mental state
state	configuration of my body	mental state of your mind

Figure 1. The sensorimotor and social interaction loops. The motor control loop (a) involves generating motor commands that cause changes in the state of my own body. Depending on this new state and the outside world I receive sensory feedback. The social interaction loop (b) involves me generating motor commands that cause communicative signals. These signals when perceived by another person can cause changes in their internal mental state. These changes can lead to actions which are, in turn, perceived by me.

within a movement. However, other key parameters change discretely, like the identity of a manipulated object, or, on a slower time-scale, like the mass of a limb. We refer to such discrete or slowly changing parameters as the *context* of the movement. Finally, dependent on sensory feedback the CNS can generate a new motor command or update the current motor command, thereby completing the sensorimotor loop.

For accurate control the CNS has to adapt the motor command to both the current context and state of the body. However, this information is not directly available to the CNS and these variables are refereed to as *hidden variables* in the engineering literature. Instead the CNS has access to sensory feedback from which it may be able to estimate the state of the body. For example, there is no sensory receptor that directly tells us the location of our hand in space, but many proprioceptive and tactile sensors from the arm can be used to make an estimate of this state variable. Similarly the weight of an object to be picked up can be estimated visually on the basis of prior experience and then updated during the handling of the object.

Motor control is, therefore, concerned with inputs and outputs from a controlled object (e.g. the arm) that is part of our own body. When interacting with another person we can think of an analogous social interaction loop in which the controlled object is the other person rather than part of our own body (figure 1b). Again, our motor commands cause muscle contractions and these lead to motor consequences which generate communicative signals, such as speech or gestures. When perceived by another person these can have influences on their hidden (mental) state, which constitutes the set of parameters that determine their behaviour. We can regard the other person as having a state in the same way that our own body has a state. If we know the state of someone else and have a model of their behaviour, we should be able to predict their response to a given input that we or the environment provides. Given the other person's state, the motor command we have generated, and the context provided by the environment, the other person will generate motor commands causing consequences. We can perceive these consequences and these can be used to determine our next motor command, thereby closing a social interaction loop.

Therefore, in social interactions, by controlling someone else rather than our own body, we can estimate their hidden state including their mental state rather than the state of our own body.

3. WHAT MAKES MOTOR CONTROL AND SOCIAL INTERACTION DIFFICULT?

There are several features of the neural circuitry and musculoskeletal system that significantly complicate our ability to produce accurate and fast movements. First, there are considerable *time delays* in both the transduction and transport of sensory signals to the CNS. For example, visual feedback can take *ca*. 100 ms to be processed. When this sensory delay is combined with efferent delays associated with movement, the combined delay is appreciable. As a consequence, sensory information cannot be used to guide the initial part of a movement and skilled performance requires feed-forward control. However, there is still a problem of co-registering actions with their consequences in time as these signals can be separated by several hundred milliseconds. In addition to delays, the sensory inputs and motor commands suffer from intrinsic *neural noise*, or randomness, which limits the ability of the system to perform rapid and accurate movements simultaneously (Harris & Wolpert 1998).

Not only are motor and sensory signals delayed and noisy, but the relationship between the motor commands and sensory consequences can be very complicated. The equations relating the force produced by muscles and the ensuing motion of the body are highly complex. For example, the equations that determine the effect that a single muscle acting on the elbow has on the subsequent change in elbow angle will, owing to interactions between body segments, have terms that depend in complex ways on factors such as the orientation of the body with respect to gravity, the rotation of the body in space and the rotational velocity of the shoulder joint. Moreover, the complexity of the musculoskeletal system is made worse because it has *nonlinear* properties. Linear systems are ones in which if you know how the system responds to two different sequences of force acting on it, then it is very easy to predict what will happen when the two series of

forces are added and applied together. For example, a ball on a table acted on by forces is a linear system. A sequence of forces acting on the ball will cause the ball to take up a sequence of positions on the table. Another sequence of forces acting on the ball will cause the ball to take up a different sequence of positions. If we add the two sequences of forces and applied these to the ball it would follow a path determined by the sum of the positions from each sequence individually. However, the musculoskeletal system is nonlinear and this makes motor control difficult as knowing the consequence of a variety of motor commands does not allow us easily to generalize to what will happen to combinations of these motor commands. Moreover, the relationship between motor commands and ensuing movement changes every time we interact with a novel object. This property of being ever-changing is known as *non-stationarity*. This requires that the command sent to our body be tailored to the changing interactions with the world.

Finally, the motor system has a high-dimensional state (dimension refers to number of parameters required to define the state). For example, the final control must be exerted on the 600 or so muscles in the human body. Even if we consider each, as being, for extreme simplicity, either contracted or relaxed, this leads to 2^{600} possible motor activation patterns, more than the number of atoms in the known universe. When trying to represent such high-dimensional data we run into the problem of the 'curse of dimensionality' (Bellman 1957). It is implausible that the CNS represents all possible configurations so it must instead find simplifying rules during control and learning.

When considering the social interaction loop and regarding another person as the controlled object, we encounter similar, but usually more severe, problems. First, the time delays between our action and the consequences on our own body are of the order of hundreds of milliseconds, whereas with other people the consequences can be of the order of seconds to minutes or even days. Moreover, the response of a person to our actions is not easily predicted. There is usually a complex, noisy and nonlinear relationship between our actions and the consequences. In a similar way to the nonlinearity of the arm, knowing how someone will respond to two separate actions we perform does not allow us to predict accurately the response to both actions performed simultaneously. Moreover, in the same way that motor command and sensory feedback are corrupted by noise we can regard the other person as a nonlinear system with noise. There is noise in both their perception of our actions and our perception of their response. But moreover, there may be a stochastic element in their response to the same action. Part of this is due to their internal state to which we do not have access, and part can be considered as a stochastic element in their choice of response. In addition, whereas the state of the human body has perhaps several hundred degrees of freedom, the possible degrees of freedom of another person's brain are likely to be far greater.

Finally, in the same way that the motor system has to deal with multiple contexts, such as multiple tools, social interaction requires us to interact with multiple people. Different tools have different dynamics, that is, different response to forces we apply to them. Similarly, different people will react in different ways to the same input.

Therefore both control and social interaction have to take into account the context, whether it is the identity of a tool or the identity of another person.

However, although the behaviour of others given our actions are more noisy, nonlinear, delayed and of higher dimension than the response of our arm to our motor command, they may not be fundamentally different in terms of computational requirements.

4. INTERNAL MODELS OF THE LOOP TRANSFORMATIONS

On the basis of computational studies it has been proposed that the CNS internally simulates aspects of the sensorimotor loop in planning, control and learning (Kawato *et al.* 1987; Jordan 1995; Miall & Wolpert 1996; Wolpert & Flanagan 2001). The neural circuits within the CNS that perform such transformations are termed *internal models* as they are internal to the CNS and model aspects of the sensorimotor loop. Internal models that predict the sensory consequences of a motor command are known as *forward* models as they model the causal (forward) relationship between actions and their consequences. A forward model, therefore, can be used to predict how the motor system's state changes in response to a given motor command. Therefore, whereas the descending motor command acts on the actual sensorimotor system, a copy of this motor command, termed *efference copy* can pass into a forward model which acts as a neural simulator of the musculoskeletal system and environment. A forward model can, therefore, be used as a predictor or simulator of the consequences of an action. An inverse model performs the opposite transformation to a forward model, determining the motor command required to achieve some desired outcome. Here, we will use predictor and controller synonymously with forward and inverse models, respectively.

Skilled motor behaviour relies on accurate predictive models of both our own body and external objects and environments. As the dynamics of our body changes during development, and as we experience tools that have their own intrinsic dynamics, we constantly need to acquire new models and update existing models. Thus, forward models are not fixed entities but must be learned and updated through experience. Learning a predictive model is relatively straightforward. By comparing the predicted and actual outcome of a motor command a prediction error can be generated. Well-established computational learning rules can be used to translate these errors in prediction into changes in synaptic weights that will improve any future predictions of a forward model. We can consider a similar forward or predictive model for social interaction. In this case another person's response to my motor commands or communicative behaviour is modelled. Again, discrepancies between anticipated and actual behaviour can be used to refine such a model. Therefore, by monitoring one's own action and the response of others it is possible to learn a predictive model of the likely behaviour of someone in response to our actions.

Inverse models or controllers are in general more difficult to learn. Additional transformations may have to be applied to the error signal before it can be used to train a

controller. For example, when we throw a dart, the error we receive is in visual coordinates. This sensory error must be converted into motor command errors suitable for updating the inverse model. The two principal methods proposed in the motor control literature for solving this problem are 'distal supervised learning' (Jordan & Rumelhart 1992) and 'feedback error learning' (Kawato 1990). Distal supervised learning uses a predictive model of the system to convert from sensory errors to the required changes to the motor command, whereas feedback error learning uses a simple feedback controller to achieve a similar conversion of errors. In motor control a controller often tries to achieve some desired state of the motor system. Similarly, an inverse social model could be used to try to achieve some hidden mental state, and hence behaviour, in another person. Again, learning such a model is difficult in social interaction, as a discrepancy between another person's internal state and/or behaviour and what you wanted does not directly allow you to determine how to change your communicative signals to get nearer to the desired outcome. As with motor control, a forward social model could be used to determine the appropriate change in our actions to achieve our desired result.

Although we can phrase the forward and inverse social models in the same computational framework as motor control this should not hide several differences which makes learning such social models immensely more difficult. First, when the brain models (either forward or inverse) the motor apparatus, regardless of noise, delay, nonlinearity, the degrees of freedom are relatively small, and although some states can be considered as hidden, the depth to which they are hidden is not severe. This is because our sensory system provides us with ample information to determine the state of our arm and we have relatively limited set of control parameters that we can apply to our 600 or so muscles. Alternatively, when trying to learn an internal model of another person, the degrees of freedom are enormous, and the hidden variables are more deeply hidden. We usually need to estimate inputs and outputs of a system to model it. The brain's inputs and outputs are sensory feedbacks and motor commands. Those of the other person's brain are not available. My communication signal transmitted to you and your perceived communication signals may be too superficial to train a good internal model of you. If these signals were sufficient for a general algorithm to learn, then we would expect there to be nothing special to human communication when compared with learning an internal model of a pet dog or a humanoid robot. So, if exactly the same computational algorithms as those used in motor control are applied for communication problems, we believe the task would be excessively difficult to solve. Another problem in terms of learning is that when learning how a system responds to a set of inputs you normally want to explore a large range of inputs to see the range of outputs. Although this is possible when trying out commands on your arm, you cannot give an arbitrary battery of inputs to another person for system identification purposes, as unlike your arm another person has the option to withdraw communication once you have provided a 'bad' input (except, perhaps, in the case of infants and their mothers).

We propose that the reason we are able to solve the problem of learning internal models of other people is because of the similarity of brains across people. We propose that the uniqueness of human communication relies on our brains being similar. This allows the brain to use this fact to train a good internal model of another person's brain. We will review how having a similar motor system (brain and musculoskeletal system) between two people enables us to use the mappings between our actions and our own mental states as *a priori* information to bootstrap any learning of another person's internal models. We will illustrate these principles for a model of motor control: the MOSAIC model that we have developed.

5. MULTIPLE INTERNAL MODELS FOR ACTION PRODUCTION AND IMITATION

Humans demonstrate a remarkable ability to generate accurate and appropriate motor behaviour under many different and often uncertain environmental conditions. It has been proposed that the CNS uses a modular approach in which multiple controllers coexist and are selected based on the movement context or state (Jacobs *et al.* 1991; Narendra *et al.* 1995; Narendra & Balakrishnan 1997; Ghahramani & Wolpert 1997). Therefore, when we pick up an object with unknown dynamics we need to identify the context and select the appropriate controller. One possible solution to this identification and selection problem has been proposed in the form of the MOSAIC model (Wolpert & Kawato 1998; Haruno *et al.* 2001; Doya *et al.* 2002). The idea is that the brain simultaneously runs multiple forward models that predict the behaviour of the motor system to determine the current dynamics of the body which will change when interacting with different objects. Consider a very simple example in which there are only two contexts: that a teapot to be lifted is either full or empty (figure 2). When a motor command is generated, an efference copy of the motor command is used to simulate the sensory consequences under the two possible contexts. The predictions based on an empty teapot suggest that lift-off will take place early compared with a full teapot and that the lift will be higher. These predictions are compared with actual feedback. As the teapot is, in fact, empty the sensory feedback matches the predictions of the empty teapot context. This leads to a high likelihood for the empty teapot and a low likelihood of the full teapot. Each predictor can, therefore, be regarded as a hypothesis tester for the context that it models. The smaller the error in prediction, the more likely the context. Moreover, each predictor is paired with a corresponding controller forming a predictor–controller pair. The MOSAIC model is able to learn a set of predictors to cover the experienced behaviours and also ensures that the each paired controller is the appropriate controller to use in the context for which paired predictor is tuned (Haruno *et al.* 2001). If the prediction of one of the forward models closely matches the actual sensory feedback, then its paired controller will be selected and used to determine subsequent motor commands. In computational terms, the sensory prediction error from a given forward model is represented as a probability; if the error is small then the probability that the forward model is appropriate is high. The set of probabilities, termed *responsibilities*, from
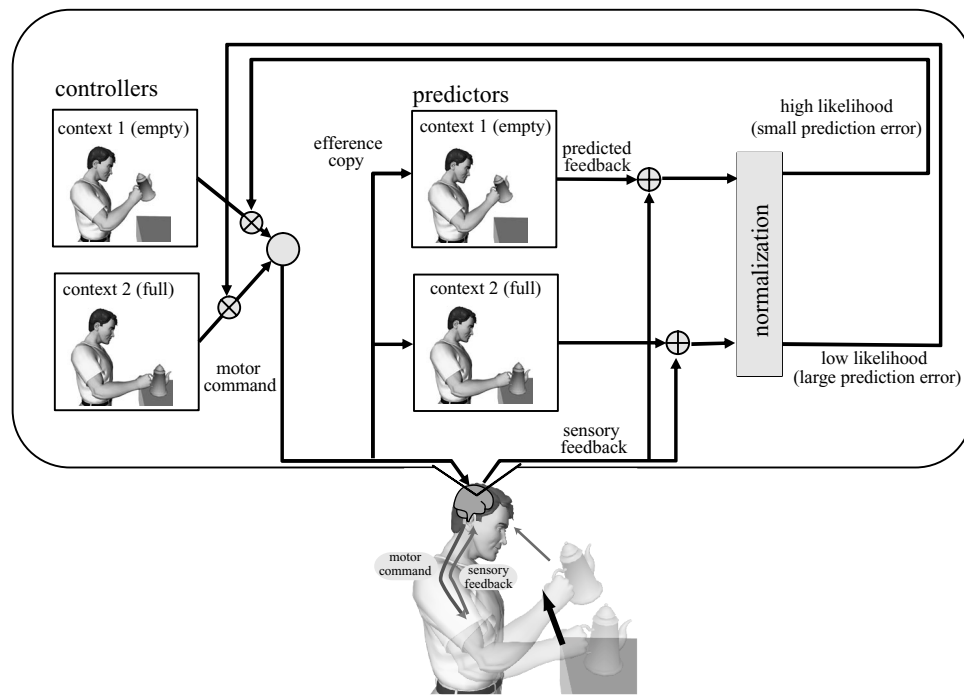
Figure 2. The MOSAIC architecture. A schematic of context estimation with just two contexts: that a teapot is empty or full. In this highly simplified example, a module consists of a controller–predictor pair. In this case two controller–predictor pairs exist: one tuned for a full teapot and one for an empty teapot. The outputs of the controllers are weighted by the likelihood that each is appropriate, to determine the final motor command. When this motor command is generated, an efference copy of the motor command is used to simulate, using the two predictors, the sensory consequences under the two possible contexts. The predictions based on an empty teapot suggest that lift-off will take place early compared with a full teapot and that the lift will be higher. These predictions are compared with actual feedback and the errors are normalized to turn them into likelihood or responsibilities. As the teapot is, in fact, empty the sensory feedback matches the predictions of the empty teapot context. This leads to a high likelihood for the empty teapot and a low likelihood of the full teapot. These responsibilities are used to adjust the weightings of the controllers so as to generate motor commands appropriate for an empty teapot. In addition, the responsibilities are used to gate the learning of the predictors and controllers (not shown).

an array of forward models is used to weight the outputs of the paired controllers.

Learning by imitation is an essential part of human motor behaviour and seems very limited in other animals, even chimpanzees. Although seemingly a trivial task of 'copying' somebody's action, learning by imitation poses a series of computational challenges including:

(i) how to map the perceptual variables (e.g. visual and auditory input) into corresponding motor variables;
(ii) how to compensate for the difference in the physical properties and control capability of the demonstrator and imitator; and
(iii) how to understand the intention of action (e.g. objective function in optimal control) from observation of the resulting movements (see Schaal *et al.* 2003).

In the MOSAIC model the consequences of a movement are compared with multiple predictions as a form of hypothesis testing as to the dynamics of the current state or context. Each predictor tests the hypothesis that the current dynamic is well captured by the predictor. The set of errors are transformed into responsibilities (probabilities) and provide rich information about the likely state the system is in. A natural extension of the model is to compare the predictions, not with one's own state, but with the state of a system that is being observed.

We hypothesize that, in this way, during action observation the motor system can be used to understand the actions of others. This could be an efficient process because our CNS has learned to predict the consequences of actions on our own body and this can be used to make accurate prediction about others. The use of our own motor system in understanding actions could underlie our extraordinary ability to detect and identify biological motion (Johansson 1973).

For the actor, at a given time only one or a small set of modules generates a motor output (figure 3*a*). To use MOSAIC to imitate movements requires three stages. First, the visual information of the actor's movement must be converted into a format that can be used as inputs to the system such as the motor system. This requires that the visual processing system obtains something akin to state (e.g. joint angles) over time which can then be used by the MOSAIC (we do not deal with this visual problem here). The second stage is that each controller in the observer generates the motor command which it would produce given the observed trajectory and current state of the actor. Rather than these commands acting on the observer's own musculoskeletal system, the output of each controller forms the input to its paired predictor, thereby generating a prediction of the next likely state (figure 3*b*). Therefore the observer uses his own multiple modules to try to simulate the observed percept. This next state prediction can be compared with the actor's next state to pro-
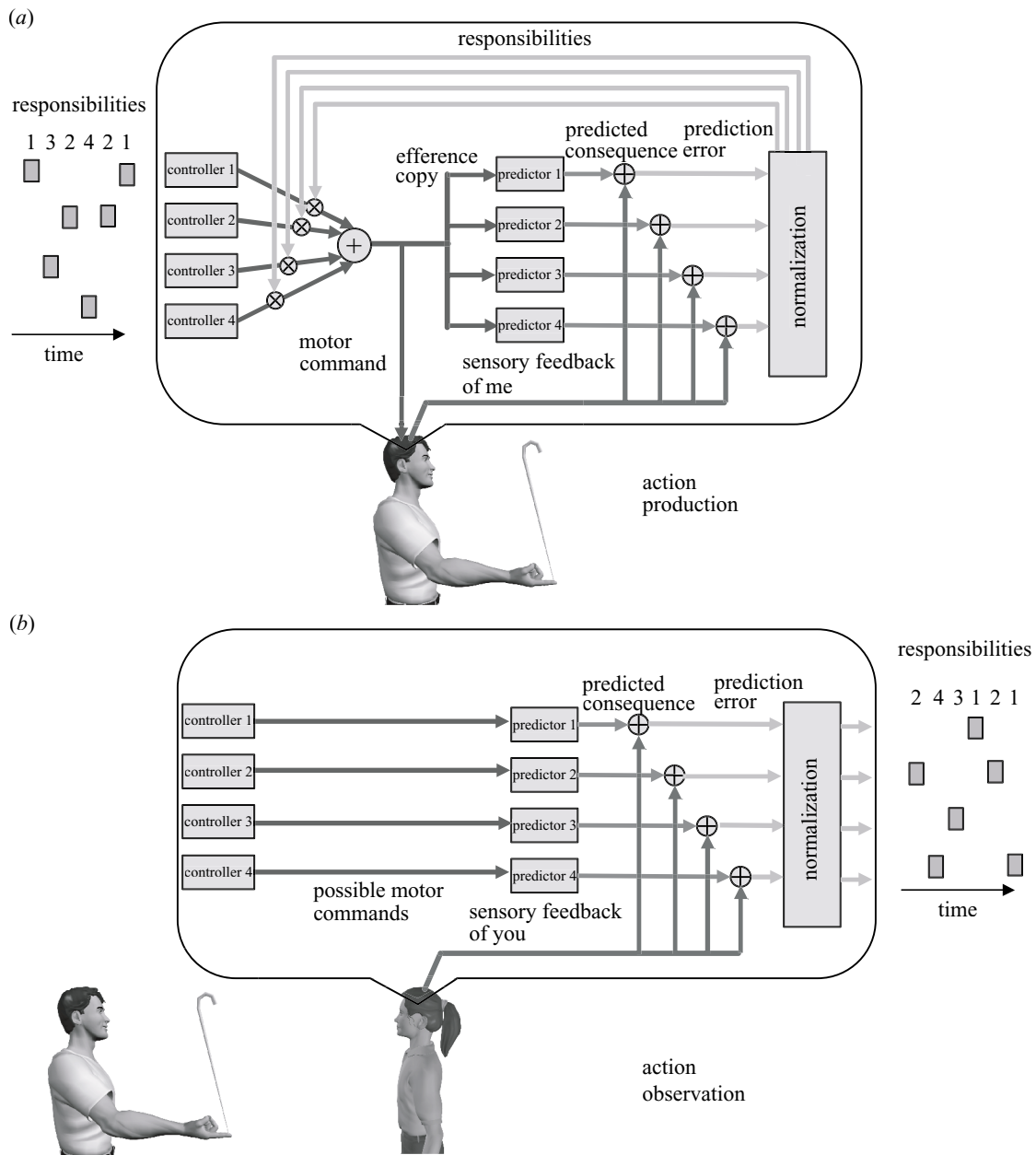
Figure 3. The MOSAIC for action observation. During action production (*a*), at a given time only one or a small set of modules generates a motor output. In this example of balancing a walking stick on a finger, the modules are activated in a particular sequence such as 1 → 3 → 2 → 4 → 2 → 1. For action production the outputs of the controllers are combined and predictions of the consequences of the motor command are compared with sensory feedback from my own body to determine future control. For action observation (*b*) each controller in the observer generates the motor command that it would produce given the observed trajectory and current state of the observed person. Rather than these commands acting on the observer's own musculoskeletal system, the output of each controller forms the input to its paired predictor, thereby generating a prediction of the likely next state. Therefore, the observer uses her own multiple modules to try to simulate the observed percept. These predictions are compared with the observed next state of the performer, leading to the likelihood that each of the observer's controllers would have generated the behaviour. Therefore, the observer encodes this as a symbolic stream, for example 2 → 4 → 3 → 1 → 2 → 4, representing the sequence of modules that needs to be used to generate the observed behaviour. The observer can use this information in imitation either by replacing their usual sequence of module activation or by biasing the selection.

duce prediction errors. Again, these prediction errors can be converted into responsibilities determining which of my controllers has to be active to generate the motion I see you perform. Therefore, the identities of the modules which best account for the percept form a symbolic code of the hidden state of the actor. When the actor generates a continuous trajectory (by activating modules 2 → 1→3 → 1→4 …), the observer encodes this as a sym-

bolic stream (e.g. module 1 → 3→4 → 2→1 …) representing which module needs to be used to generate the observed behaviour. This symbolic representation captures a representation of the observed movement, which has fewer dimensions than would be needed to store the entire trajectory. Moreover, the movement is represented in the observer's private lexicon. If the MOSAIC of the actor and observer are identical (which is never likely to

be the case) then the symbolic representations should be identical. The more different the MOSAICs the harder it may be for the observer to represent the actor's behaviour. The final stage is for the observer to use the symbolic sequence in imitation. By using the extracted symbolic sequence of module activations to activate her modules over time she is able to generate the behaviour. This information can either replace the observer's usual sequence of module activation or be used to bias it towards a better action. Preliminary simulations show that the MOSAIC can be used in this way to learn a simple acrobot task (swinging up a jointed stick to the vertical) through action observation and imitation (Doya *et al.* 2000). Therefore, the MOSAIC architecture could form the basis of a system for action production and action imitation.

This method of action observation contrasts with previous methods of imitation learning that use several heuristic methods for storing features of movement patterns, for example, points of high curvature or discontinuity (Kuniyoshi *et al.* 1994; Wada *et al.* 1995; Miyamoto *et al.* 1996). The current approach could provide a more general principle for segmenting continuous movement patterns: a local trajectory that is well predicted by a pair of controllers and predictors could be regarded as a primitive motion.

Although action observation and understanding could be achieved by purely sensory approaches we suggest that there are computational benefits to using the motor system in approaches such as with the MOSAIC model. For example, HMMs have been used extensively for automatic segmentation of motion capture data of full body motion. Multiple HMMs have the same probabilistic and modular architecture as MOSAIC, and a long history of moderately successful application to fields such as speech recognition. The essential difference between MOSAIC and HMMs is that controllers are involved. Inclusion of controllers may be beneficial for two reasons. First, the communication signals such as speech, facial expressions or body language, are generated by controllers. Thus, MOSAIC is a better model than HMM as a generative model of these communicative signals. Second, given the similarity of brains within the human species, my MOSAIC should be a much better approximation than any arbitrary recurrent or feed-forward neural network or HMM as a model of another person's brain.

6. HIERARCHY FOR THE CONTROL AND EXTRACTION OF INTENTIONS

Hierarchy plays a key role in human motor control. We can generate a variety of motor sequences in a very coherent manner despite the different conditions and contexts in which we have to act. For example, the kinematics of writing is preserved when using different effectors and when the dynamics of the pen are varied. This suggests that high-level representations of the characters may exist and that the lower levels are concerned with compensating for different dynamics. An interesting question is how such hierarchical motor control can be learned and used?

A feature lacking in the current formulation of the MOSAIC model is the hierarchical and bi-directional control of the modules' activity. To incorporate such control, we have proposed a new conceptual architecture, the

HMOSAIC consisting of several layers of MOSAIC (Haruno *et al.* 2003) (figure 4). Bi-directional information processing between layers of HMOSAIC can be phrased within a Bayesian statistical framework. The input of higher-level modules is the (bottom-up) responsibility signals (posterior probability) from the subordinate modules, which represent the currently selected modules given the current behavioural situation. The output of higher-level modules is a set of (top-down) prior probabilities of the subordinate modules, which act to prioritize lower-level module selection. More precisely, the higher control model learns to output the prior probabilities to lower modules given the current behavioural situation and possibly an abstract (symbolic) desired behaviour. By contrast, the higher predictive model learns to anticipate the posterior probability of the lower level at the next time step. The precision of the prediction is used to weight the outputs from control models as well as the learning signal for both predictive and control models. Thus, the lower- and higher-level modules interact bidirectionally during learning and control of hierarchically organized movements. The HMOSAIC architecture can learn both elementary movements (lower-level chunking) and their hierarchical temporal order (higher-level sequencing) through sensorimotor learning. Simulations have shown that the HMOSAIC can learn how to control multiple objects and learn how the object is likely to change over time, thereby learning temporal sequences (Haruno *et al.* 2003). The hierarchical architecture embodies a way of reconciling top-down plans and bottom-up constraints. This is a fundamental problem in hierarchical decision systems, often called a 'symbol grounding' problem.

Conceptually the lowest level in the hierarchy learns the elements of control for different contexts or states. The next level up learns how to put elemental sequences together: for example, learning how to control transitions between the modules, thereby learning elemental sequence patterns. Progressively higher levels learn more abstract representations, with the higher levels learning goals or intentions. Therefore the activations of a higher-level goal such as to get a drink of water, would activate lower levels in such a way as to finally generate the appropriate commands to reach for a glass of water. An important feature of the hierarchy is the tree-like structure so that higher levels could have multiple paths to activating lower levels, and the choice of path, or way of achieving a goal, can be biased by higher-level factors. By including recurrent networks within the modules at higher levels in HMOSAIC, the architecture should be able to generate arbitrary combinations of the lowest primitives, using a finite set of primitives to generate a possibly vast repertoire of actions, in a similar way to the role of recursion in language (Hauser *et al.* 2002).

In § 5 we proposed that the flat MOSAIC could be used for low-level imitation of the modules which would directly reproduce the kinematics (trajectory) of a movement. However, using the HMOSAIC we could propagate up the responsibility signal during action observation to estimate which module at the various levels of the HMOSAIC would need to be active to generate the observed behaviour. Using such an architecture it may be possible to have several representations of the observed action, from the low-level kinematics of movement (which mod-
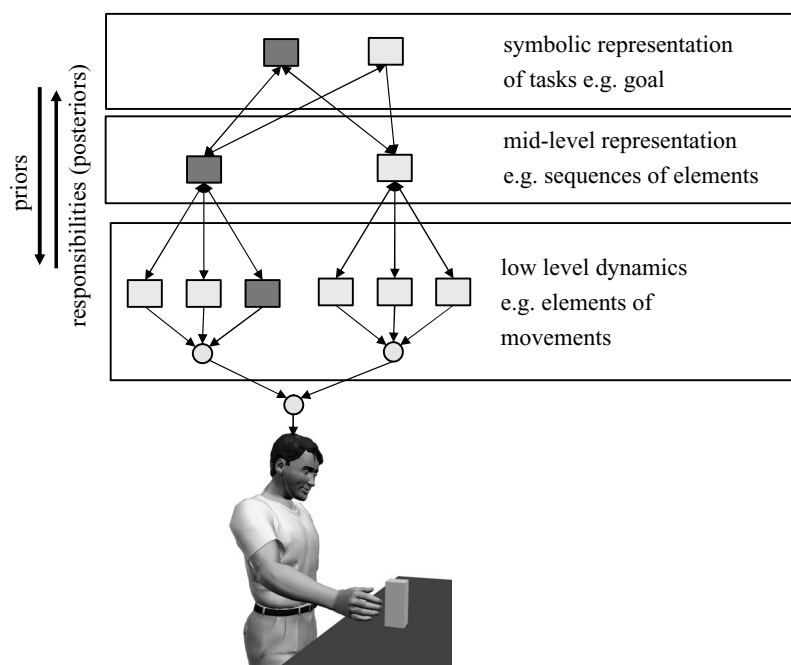
Figure 4. The hierarchical MOSAIC for action generation. Three layers of a simple HMOSAIC are shown in which each block is a module, representing a predictor–controller pair, with the lowest levels represented by two flat MOSAIC structures. The input of higher-level modules is the (bottom-up) responsibility signals (posterior probability) from the subordinate modules, which represent the currently selected modules given the current behavioural situation. The output of higher-level modules is a set of (top-down) prior probabilities of the subordinate modules, which act to prioritize lower-level module selection. The HMOSAIC architecture can learn both elementary movements (lower-level chunking) and their hierarchical temporal order (mid-level sequencing) through sensorimotor learning. Progressively higher levels learn more abstract representations, with the higher-level learning goals or intentions. Therefore the activations of a higher-level goal, such as to pick up an object, would activate lower-level modules (dark) in such a way as to finally generate the appropriate commands to reach for an object.

ules are active in the lowest level), to representing sequences of actions (intermediate levels) to the goal (highest level). The degree to which propagation up the hierarchy is possible depends on the extent to which a coherent account of an observed action can be made using the observer's HMOSAIC. The more similar the observer's HMOSAIC is to the actor's HMOSAIC the easier it will be to make coherent, and unique interpretations at higher levels. Therefore, a movement that has a clear goal (which is also a goal that I have represented in my HMOSAIC) could be understood at all levels and imitation of the goal, even with different effectors, would be possible. However, a meaningless movement, or one for which the observer does not have a goal, could be understood only at lower levels, with imitation slavishly replicating kinematics or sequences (Wohlschläger *et al.* 2003). The key idea is that having similar computational structures to generate movement, such as HMOSAIC, dramatically reduces the computational problems in action understanding. We have yet to simulate the hierarchical action understanding.

7. COMMUNICATION AS CLOSING THE LOOP

So far we have discussed the use of MOSAIC and HMOSAIC in an unidirectional manner, in that the actor pays no attention to the observer's actions. In true communication the actor (the transmitter) is responsive to misperception by the observer (the receiver). One way to close the communication loop is as follows. The transmit-

ter uses his internal symbolic stream to generate a series of motor commands that in turn cause movements. The receiver decodes the movements he sees into his internal symbols and then also generates a series of motor commands (attempting to imitate the transmitter). The transmitter then sees these imitative movements and interprets them back into his own symbols. He can then compare the symbols he wished to transmit with the symbols he believes he has transmitted. This discrepancy error can then be used by the transmitter to determine a new sequence of motor commands in an attempt to get the receiver to internalize these symbols more accurately. So, for example, if the symbols were responsibilities he could generate an action using the original responsibilities augmented with the error. Alternatively, to learn the internal structure of the MOSAIC of others we could use the discrepancy error to update the structure of our own MOSAIC to more closely match those of others.

One of the necessary conditions for exact and rigorous communication at symbolic levels is to have an identity mapping in the closed loop of my symbols → my action → your symbols → your imitation → my perception → my interpretation of your symbols = my original symbols.

There is, therefore, no need in principle why your MOSAIC and my MOSAIC should have similar structures. However, we expect that if they have identical structures you and I will be able to communicate anything we wish. The more dissimilar the structure the more things we will get confused about during communication. In

either case there is no need for the modules to be numbered or related—just the fact that you have a module somewhere that does the same job as one of mine is good enough. If discrepancies exist between the responsibilities used for generation and the responsibilities during perception, these could be used to update my MOSAIC, to make it more like yours.

Analogous to the state of our own system is the state of someone else's mind, being the set of parameters that are required to predict the behaviour of the person given inputs and their dynamics. Although in the case of our own arm we may be able to monitor fully the inputs of the system, for another person we may only know some of the inputs. Knowing the system dynamics requires us to learn how, given a particular internal state and input, the other persons will respond. A default is, as described already, to use our own HMOSAIC to estimate other people's hidden states. This allows us to use a single system to interpret the actions of all other people. However, there are situations in which it is inappropriate to assign the same set of internal state to action mappings to everyone. An alternative is to learn a new HMOSAIC for other people. One possibility is that our own HMOSAIC could be augmented by structures that aim to model the difference between our HMOSAIC and others. Such a system would allow a representation of others' internal mental state separately from our own HMOSAIC structure and may therefore form a basis for theory of mind.

8. CONCLUSION

We have explored the computational parallels between the computations that occur in motor control and in social interaction. In particular we examined how models of motor control, such as the HMOSAIC, could be used for action observation, imitation, social interaction and theory of mind. We suggest that using our motor system in action understanding is an efficient mechanism for performing the computations needed in social interaction.

REFERENCES

Bellman, R. 1957 *Dynamic programming*. Princeton University Press.

Doya, K., Katagiri, K., Wolpert, D. M. & Kawato, M. 2000 Recognition and imitation of movement patterns by a multiple predictor–controller architecture. *Technical Rep. IEICE* **TL2000-11**, 33–40.

Doya, K., Samejima, K., Katagiri, K. & Kawato, M. 2002 Multiple model-based reinforcement learning. *Neural Comput.* **14**, 1347–1369.

Fadiga, L., Fogassi, L., Pavesi, G. & Rizzolatti, G. 1995 Motor facilitation during action observation: a magnetic stimulation study. *J. Neurophysiol.* **73**, 2608–2611.

Fadiga, L., Craighero, L., Buccino, G. & Rizzolatti, G. 2002 Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eur. J. Neurosci.* **15**, 399–402.

Gallese, V. 2003 The manifold nature of interpersonal relations: the quest for a common mechanism. *Phil. Trans. R. Soc. Lond.* B **358**, 517–528. (DOI 10.1098/rstb.2002.1234.)

Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. 1996 Action recognition in the premotor cortex. *Brain* **119**, 593–609.

Ghahramani, Z. & Wolpert, D. M. 1997 Modular decomposition in visuomotor learning. *Nature* **386**, 392–395.

Grafton, S. T., Fadiga, L., Arbib, M. A. & Rizzolatti, G. 1997 Premotor cortex activation during observation and naming of familiar tools. *NeuroImage* **6**, 231–236.

Grezes, J., Fonlupt, P., Bertenthal, B., Delon-Martin, C., Segebarth, C. & Decety, J. 2001 Does perception of biological motion rely on specific brain regions? *NeuroImage* **13**, 775–785.

Harris, C. M. & Wolpert, D. M. 1998 Signal-dependent noise determines motor planning. *Nature* **394**, 780–784.

Haruno, M., Wolpert, D. M. & Kawato, M. 2001 Mosaic model for sensorimotor learning and control. *Neural Comput.* **13**, 2201–2220.

Haruno, M., Wolpert, D. & Kawato, M. 2003 Hierarchical MOSAIC for movement generation. In *Excepta Medica International Congress Series*, vol. 1250 (ed. T. Ono, G. Matsumoto, R. R. Llinas, A. Bethoz, R. Norgren, H. Nishijo & R. Tamura). Amsterdam: Elsevier Science.

Hauser, M. D., Chomsky, N. & Fitch, W. T. 2002 The faculty of language: what is it, who has it, and how did it evolve? *Science* **298**, 1569–1579.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C. & Rizzolatti, G. 1999 Cortical mechanisms of human imitation. *Science* **286**, 2526–2528.

Jacobs, R. A., Jordan, M. I., Nowlan, S. J. & Hinton, G. E. 1991 Adaptive mixture of local experts. *Neural Comput.* **3**, 79–87.

Johansson, G. 1973 Visual perception of biological motion and a model for its analysis. *Perception Psychophys.* **14**, 201–211.

Jordan, M. I. 1995 Computational aspects of motor control and motor learning. In *Handbook of perception and action: motor skills* (ed. H. Heuer & S. Keele). New York: Academic.

Jordan, M. I. & Rumelhart, D. E. 1992 Forward models: supervised learning with a distal teacher. *Cogn. Sci.* **16**, 307–354.

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

Kawato, M., Furawaka, K. & Suzuki, R. 1987 A hierarchical neural network model for the control and learning of voluntary movements. *Biol. Cybern.* **56**, 1–17.

Kuniyoshi, Y., Inaba, M. & Inoue, H. 1994 Learning by watching: extracting reusable task knowledge from visual observation of human performance. *IEEE Trans. Robotics Automation* **10**, 799–822.

Liberman, A. M. & Whalen, D. H. 2000 On the relation of speech to language. *Trends Cogn. Sci.* **4**, 187–196.

Martin, A., Wiggs, C. L., Ungerleider, L. G. & Haxby, J. V. 1996 Neural correlates of category-specific knowledge. *Nature* **379**, 649–652.

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

Miyamoto, H., Schaal, S., Gandolfo, F., Koike, Y., Osu, R., Nakano, E., Wada, Y. & Kawato, M. 1996 A Kendama learning robot based on bi-directional theory. *Neural Networks* **9**, 1281–1302.

Narendra, K. S. & Balakrishnan, J. 1997 Adaptive control using multiple models. *IEEE Trans. Automatic Control* **42**, 171–187.

Narendra, K. S., Balakrishnan, J. & Ciliz, M. K. 1995 Adaptation and learning using multiple models, switching, and tuning. *IEEE Control Systems Mag.* **15**, 37–51.

Rizzolatti, G. & Arbib, M. A. 1998 Language within our grasp. *Trends Neurosci.* **21**, 188–194.

Schaal, S., Ijspeert, A. & Billard, A. 2003 Computational approaches to motor learning by imitation. *Phil. Trans. R. Soc. Lond.* B **358**, 537–547. (DOI 10.1098/rstb.2002.1258.)

Wada, Y., Koike, Y., Vatikiotis-Bateson, E. & Kawato, M. 1995 A computational theory for movement pattern recognition based on optimal movement pattern generation. *Biol. Cybern.* **73**, 15–25.

Wohlschläger, A., Gattis, M. & Bekkering, H. 2003 Action generation and action perception in imitation: an instance of the ideomotor principle. *Phil. Trans. R. Soc. Lond.* B **358**, 501–515. (DOI 10.1098/rstb.2002.1257.)

Wolpert, D. M. & Flanagan, J. R. 2001 Motor prediction. *Curr. Biol.* **11**, R729–732.

Wolpert, D. M. & Ghahramani, Z. 2000 Computational principles of movement neuroscience. *Nature Neurosci.* **3**(Suppl.), 1212–1217.

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

GLOSSARY

CNS: central nervous system

HMM: hidden Markov model

HMOSAIC: hierarchical modular selection and identification for control

MOSAIC: modular selection and identification for control