Is imitation learning the route to humanoid robots?

Stefan Schaal

This review investigates two recent developments in artificial intelligence and neural computation: learning from imitation and the development of humanoid robots. It is postulated that the study of imitation learning offers a promising route to gain new insights into mechanisms of perceptual motor control that could ultimately lead to the creation of autonomous humanoid robots. Imitation learning focuses on three important issues: efficient motor learning, the connection between action and perception, and modular motor control in the form of movement primitives. It is reviewed here how research on representations of, and functional connections between, action and perception have contributed to our understanding of motor acts of other beings. The recent discovery that some areas in the primate brain are active during both movement perception and execution has provided a hypothetical neural basis of imitation. Computational approaches to imitation learning are also described, initially from the perspective of traditional AI and robotics, but also from the perspective of neural network models and statistical-learning research. Parallels and differences between biological and computational approaches to imitation are highlighted and an overview of current projects that actually employ imitation learning for humanoid robots is given.

In 1921, Karel Capek's play *Rossum's Universal Robots*[1] provided the first concrete vision of how a robot should look: it should look like a human being. Since this time, science fiction stories have created a never ending stream of increasingly sophisticated superhuman machines, but research has not been able to realize a robot that comes even close to Capek's relatively 'simple' artificial humans. What makes it so hard to create a human-like robot? From an engineering point of view, an argument could be made that appropriate materials, motors, power supplies and sensors are missing to achieve the compact, compliant and lightweight design of biology. However, even if we had access to a robotic system that incorporated all these desirable properties, and even if this machine were equipped with a supercomputer, we would still not succeed in creating a humanoid. The problem is that the algorithms required to program this machine such that it achieves the versatility and flexibility of biological systems are not yet available. For the time being, robots can only solve tasks after the task has been carefully analysed and added to the robot program by a human. An impressive example of such a procedure was recently provided by the research laboratories of the Honda Corporation in Japan[2,3]. In about ten years of work, Honda created the first humanoid robot that can walk, climb stairs, and manipulate simple objects (Fig. 1). However, the amount of work it

took to build the Honda robot, and the fact that, besides locomotion, the machine requires teleoperation to perform other tasks, is far from satisfying.

In order to overcome the need for teleoperation and manual 'hard-coding' of every behavior, a learning approach is required. The most general approach to learning control is reinforcement learning[4,5]. Reinforcement learning usually requires an unambiguous representation of **states** and **actions** of the movement system and the existence of a scalar **reward function** (see Glossary). Learning proceeds by trying actions in a particular state, and, based on the received, possibly delayed[6] reward, updating an evaluation function that assigns expected rewards to possible actions. After learning, the action with the highest expected value in each state is chosen to achieve the task goal.

An important element of learning control is the requirement to explore new actions in order to find good, or even optimal, solutions to a given movement task. For a movement system with many degrees of freedom, there is an exponential explosion in the number of actions that can be taken in every state. For example, the Honda robot shown in Fig. 1 has 30 degrees-of-freedom (DOFs), each of which needs a motor command at every instant in time. Even if the command for each DOF has only three possible values (e.g. forward, backward and zero, there is a combination of

S. Schaal is at the Department of Computer Science and Neuroscience, HNB-103, University of Southern California, Los Angeles, CA 90089-2520, USA; and the Kawato Dynamic Brain Project (ERATO/JST), 2-2 Hikaridai, Seika-cho, Soraku-gun, 619-02 Kyoto, Japan.

tel: +1 213 740 9418
fax: +1 213 740 1510/ 5687
e-mail: sschaal@usc.edu
http://www-slab.usc.edu/sschaal

Glossary

Accommodation: The process by which an activated internal perceptual motor representation is adapted to better suit a new task.

Action: All variables (discrete or continuous) that can actively change the state of a system. Usually, actions are motor commands, abbreviated as a vector u.

Action level imitation: The indiscriminate copying of the actions of the teacher without mapping them onto more abstract motor representation.

Assimilation: The mapping of an observed behavior onto an existing perceptual motor representation.

Control policy: A function that maps the state x of a movement system and its environment into an appropriate action u for a particular task, i.e. $u = p(x, t, a)$. As indicated, the function π can directly depend on the time, t, and some additional parameters α that may be useful to adjust the policy for a particular task goal. Movement primitives (c.f.) can be formalized in the form of control policies.

Deferred imitation: Imitation that takes place a certain amount of time after the demonstration was given by the teacher.

Emulation: Goals in the environment become overt due to the actions of others. Afterwards, the observer strives to attain the same goal.

Forward model: A mathematical model that predicts the time evolution of a dynamical system. For example, differential equations of motion in the form of $\dot{x}^Y = f(x,u)$ are prototypical forward models in motor control.

Immediate imitation: Imitation that takes place immediately after the demonstration of the teacher.

Movement primitive: A sequence of actions that can accomplish a certain movement goal. (c.f. Control policy for a more formal definition.)

Object-centered representations: The representation is insensitive to the relative position and orientation of the perceived object to one's own body.

Priming: Stimuli in the environment that co-occur with the actions of the teacher increase the observer's activation of corresponding internal representations in memory. Consequently, the observer's exploratory behavior will be biased by these activations towards receiving similar stimuli.

Program-level imitation: A process by which the structural organization of a behavior is copied from observing a teacher, while the exact details of actions are filled in by individual learning.

Response facilitation: Observed actions enhance corresponding action representations in memory. Afterwards, it is more likely that the observer performs the observed action.

Reward function: A function that provides a scalar (discrete or continuous) value about the goodness of an action u in a state x.

Simultaneous imitation: Imitation that takes place concurrently with the demonstration of the teacher.

Spline approximations: A mathematical concept of curve fitting, usually with low-order polynomials. Complex curves need a sequence of concatenated splines for a good approximation. The points where a spline ends and a new one begins are called spline nodes.

State: All variables (discrete or continuous) that are necessary to model a system. In this article, all states of a system are compactly denoted as a vector x.

State–Action space: The mathematical space spanned by actions u and states x jointly. Solving a movement task can be thought of as finding a path between two points in this state, the initial state and the goal state.

Stimulus enhancement: An object or place becomes more salient due to the actions of the teacher in its vicinity or in conjunction with the object. This enhancement will draw the observer's attention or elicit responses towards this object or place.

Supervised learning: Learning of an input to output mapping under the premise that an explicit error signal can be provided for each output. For instance, a Control policy (c.f.) is a function that can be learned by supervised learning if both the state x and the appropriate target action u are given to the learning system.

Task-level learning: Learning of a task can take place by learning an appropriate Control policy (c.f.) that generates commands u on the actuator level, or by learning a Control policy that generates commands in a more abstract but task related space (e.g. the space of the finger tip). The latter approach is called task-level learning and it requires additional transformations to map the task-level command into actuator space. Usually, errors in performance are more associated with task commands than actuator commands.

Viewer-specific representations: The representation is sensitive to the relative position and orientation of the perceived object to one's own body.

$3^{30} \geq 10^{14}$ different actions that can be taken in every state. As it is impossible to search such huge spaces for what constitutes a good action, it is necessary to either find more compact state–action representations, or to focus learning on those parts of the **state–action space** that are actually relevant for the movement task at hand. In this article, we will review how the latter topic can be approached in the framework of imitation learning, while the former topic – that is, compact state–action representations – will be shown to be a natural prerequisite for imitation learning in the form of movement primitives.

Movement imitation

Movement imitation is familiar to everybody from daily experience: a friend demonstrates a movement, and immediately we are capable of approximately repeating it. For the purpose of this review, only visually mediated imitation will be considered, although, at least in humans, verbal communication can supply important additional information. From the viewpoint of motor learning, a teacher's demonstration as the starting point of one's own learning can significantly speed up the learning process: imitation drastically reduces the size of the state–action space that needs to be explored[7]. With an eye towards computational approaches, we will first review some of the most relevant milestones in the study of imitation learning before we will look into the promises and challenges that imitation learning poses to computational modelers.

Imitation from the viewpoint of behavioral and cognitive sciences

In infant and animal studies, the ability to imitate is usually concluded from the subject's increased tendency to execute a previously demonstrated behavior. However, other causes can equally account for a higher probability of the subject's behavior, in particular **priming**, **emulation**, and **response facilitation** (see Glossary); such causes are not to be mistaken for true imitation[8,9]. True imitation is present only if (1) the imitated behavior is new for the imitator, (2) the same task strategy as that of the demonstrator is employed, and

(3) the same task goal is accomplished[9]. For example, if a movement is not new, response facilitation rather than imitation can account for the imitator's behavior. However, as will be shown below, from a computational point of view, imitative behavior in the form of response facilitation already contains very complex problems. Thus, a formal understanding of the mechanisms of response facilitation would constitute significant scientific progress.

Behavioral and cognitive sciences have long been interested in imitation of movements, from Darwin[10] and Thorndike[11] to Piaget[12]. After Piaget's work, movement imitation did not receive widespread attention anymore, partially due to the prejudice that 'imitating' or 'mimicking' is not an expression of higher intelligence. This attitude changed in the 1970s, to some extent owing to the studies of Meltzoff and Moore[13]. These authors reported on the ability of 12–21 day old and, later, even less than an hour old infants[14] to imitate both facial and manual gestures. Young infants of this age had neither seen their own faces nor been exposed to viewing faces of other humans for any significant amount of time. Thus, the ability to map a perceived facial gesture to their own gestures was concluded to be innate and contradicted Piaget's ontogenetic account of imitation[12]. In the light of this new interest in imitation learning, it was discovered that many animals are unable to learn by imitation[9,15]. These findings contributed to today's view of imitation as an important expression of higher intelligence. In recent years, a significant amount of new work has been published on imitation learning in humans and animals. Examples include comparative studies between human and monkey imitation[16], the interplay between memory and imitation and its implications for learning[17,18], and the focus of attention while observing a demonstration[19]. Imitation is also closely connected to research that investigates the connection between action and perception[20–22].

Behavioral studies can provide several sources of ideas for computational motor control and imitation learning. For instance, the principle of emulation is interesting for learning how to direct the focus of attention towards favorable goals, and priming can be used to bias explorative behavior towards useful stimuli. Thus, both principles help to reduce the number of inputs to a motor system by focusing on important stimuli, but they do not reduce the combinatorial explosion of possible actions that can be explored in a particular state. Alternatively, Meltzoff and Moore's[23] postulate of imitation as a bootstrapping for communication may offer an interesting starting point for computational modelers. The authors argued that given a partially familiar stimulus in the environment (e.g. a face), imitating what this 'stimulus' did during a previous encounter should trigger either a familiar or an unfamiliar response in how the stimulus will change. Such a strategy would help to disambiguate stimuli by acting upon them. Interestingly, we will encounter a similar interpretation of the purpose of imitation from a neurophysiological perspective, and a related computational mechanism for movement recognition based on predictive **forward models** in the section on neural network models to imitation. A more concrete concept of imitation was suggested in Meltzoff and Moore[24]. In their 'Active Intermodal Mapping' (AIM) model, they propose
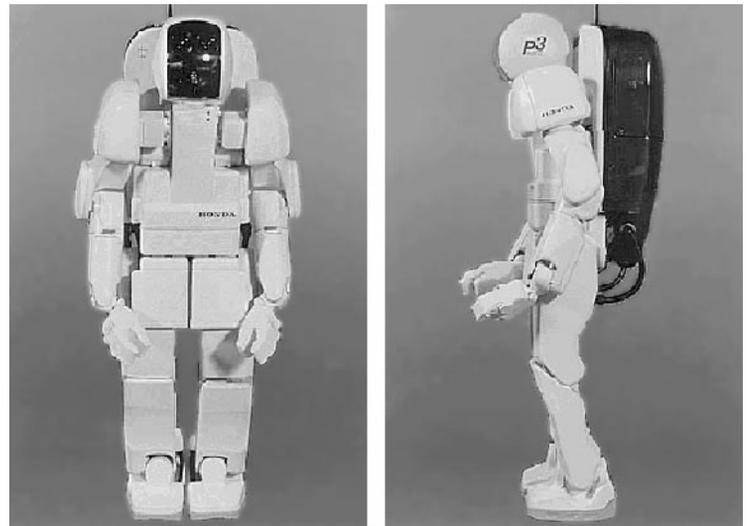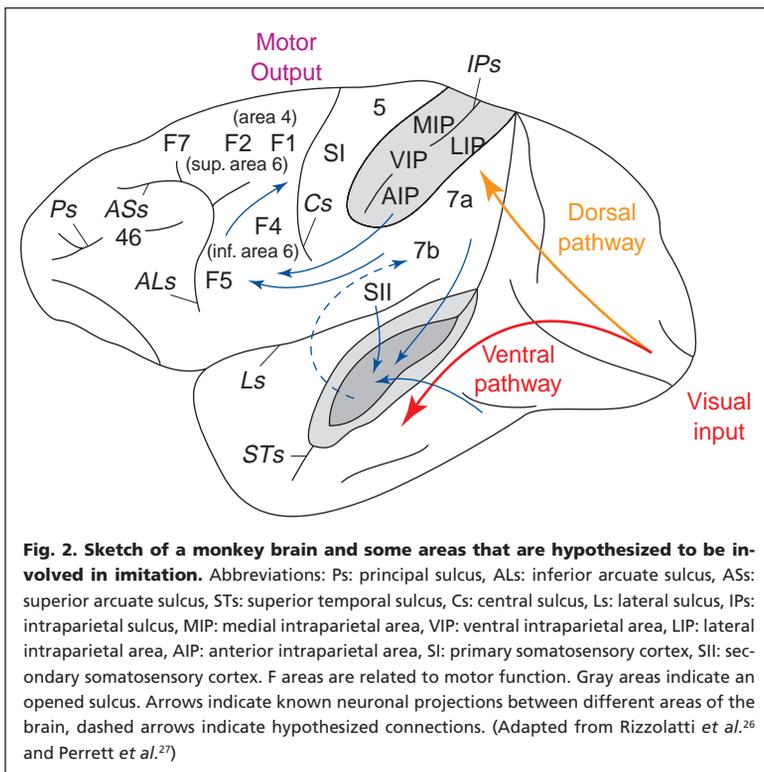

Fig. 1. Honda Humanoid Robot in frontal and side views. (Reproduced with kind permission of Honda.)

that visual perception of the teacher's movement is converted into a higher level representation that can be matched against appropriately transformed proprioceptive information about one's own movement. If this matching space is given, imitation can be seen as learning to achieve a target representation, a problem that can be tackled with techniques from **supervised learning**[25]. This idea can be rephrased in computational terms, as will be discussed below.

Imitation from the viewpoint of neuroscience and cognitive neuroscience

An essential prerequisite for imitation is a connection between the sensory systems and the motor systems such that percepts can be mapped onto appropriate actions. This mapping is a difficult computational process as visual perception takes place in a different coordinate frame from motor control. This process is also more complex than pure object recognition because it requires integration of multiple objects (i.e. several limbs), their spatial relationships, their relative and absolute movements, and even the intention of these movements. Given the current knowledge about neuroanatomy in primates, such a process is likely to happen in various steps, involving both the ventral (what) and dorsal (where, how) pathways[28] (Fig. 2). From a neurophysiological point of view, a first question that can be addressed is whether there are any particular brain areas and representations that are specialized to facilitate imitation.

Perrett and co-workers[29–31] reported that neurons within the superior temporal sulcus (STs) of macaques (Fig. 2) respond to both form and motion of objects. Interestingly, many cells were sensitive to movements of specific body parts of an observed human. For instance, cell specificity was found for faces with translatory motion in a particular direction, faces that rotated, movement of particular body parts (head, arm, leg, hip) and combinations of body parts, and even movement of the entire body. In the lower bank of the STs, similar phenomena were found for actions of the hand[30]. Two important characteristics of this distributed representation are worth highlighting. First,

Fig. 2. Sketch of a monkey brain and some areas that are hypothesized to be involved in imitation. Abbreviations: Ps: principal sulcus, ALs: inferior arcuate sulcus, ASs: superior arcuate sulcus, STs: superior temporal sulcus, Cs: central sulcus, Ls: lateral sulcus, IPs: intraparietal sulcus, MIP: medial intraparietal area, VIP: ventral intraparietal area, LIP: lateral intraparietal area, AIP: anterior intraparietal area, SI: primary somatosensory cortex, SII: secondary somatosensory cortex. F areas are related to motor function. Gray areas indicate an opened sulcus. Arrows indicate known neuronal projections between different areas of the brain, dashed arrows indicate hypothesized connections. (Adapted from Rizzolatti *et al.*[26] and Perrett *et al.*[27])

most cells were **viewer-specific**, that is, there existed no indication of an object centered representation. From this property, Perrett *et al.*[27] concluded that STs is well suited to extract the attention and goals of others. Second, owing to a connection of somatosensory cortex to STs (Fig. 2), most of the form and motion neurons were insensitive to self-motion due to re-afferent signals[32]. Thus, STs neurons are in the ideal situation to analyse the movement of others without interference from one's own body and seem to be an excellent candidate for a first processing step for imitation.

What is the next processing step in a hypothesized neural pathway for imitation? Area F5 in monkeys could play a crucial role (Fig. 2). Rizzolatti *et al.*[33] found neurons in area F5 that were specific to the execution of goal related movements (e.g. 'reaching', 'bringing-to-the-body', 'grasp-ing-with-the-hand'), as well as selective to a particular type of grasp (e.g. precision grips, finger prehension, or whole-hand-prehension). Interestingly, many neurons in F5 were active during the entire or at least extended parts of a motor act, instead of just a single submovement or muscle activation. Jeannerod *et al.*[34] and Murata *et al.*[35] interpreted this firing pattern as coding complete segments of motor acts, or motor schemas[36]. The possible connection to imitation, however, came with the finding that some of the neurons in F5, called 'mirror neurons', were active both when the monkey *observed* a specific behavior and when it *executed* it itself[37]. Mirror neurons fire highly specifically only to a special motor behavior with a particular object. These results are similar to those in STs (Ref. 30), with the difference that neurons in STs do not respond to executed motor acts, but rather only to perceived ones. From imaging and transcranial magnetic stimulation studies, there is also some evidence that a similar mirror system exists in humans[38–40]. Surprisingly, this system seems to involve Broca's area[41], a brain region

normally only associated with speech production. The possible homology of F5 in monkey and Broca's area in humans led some authors to speculate that the ability to imitate actions and to understand them could have subserved the development of communication skills[41]. This idea is consistent with Meltzhoff and Moore's developmental work and interpretations[23].
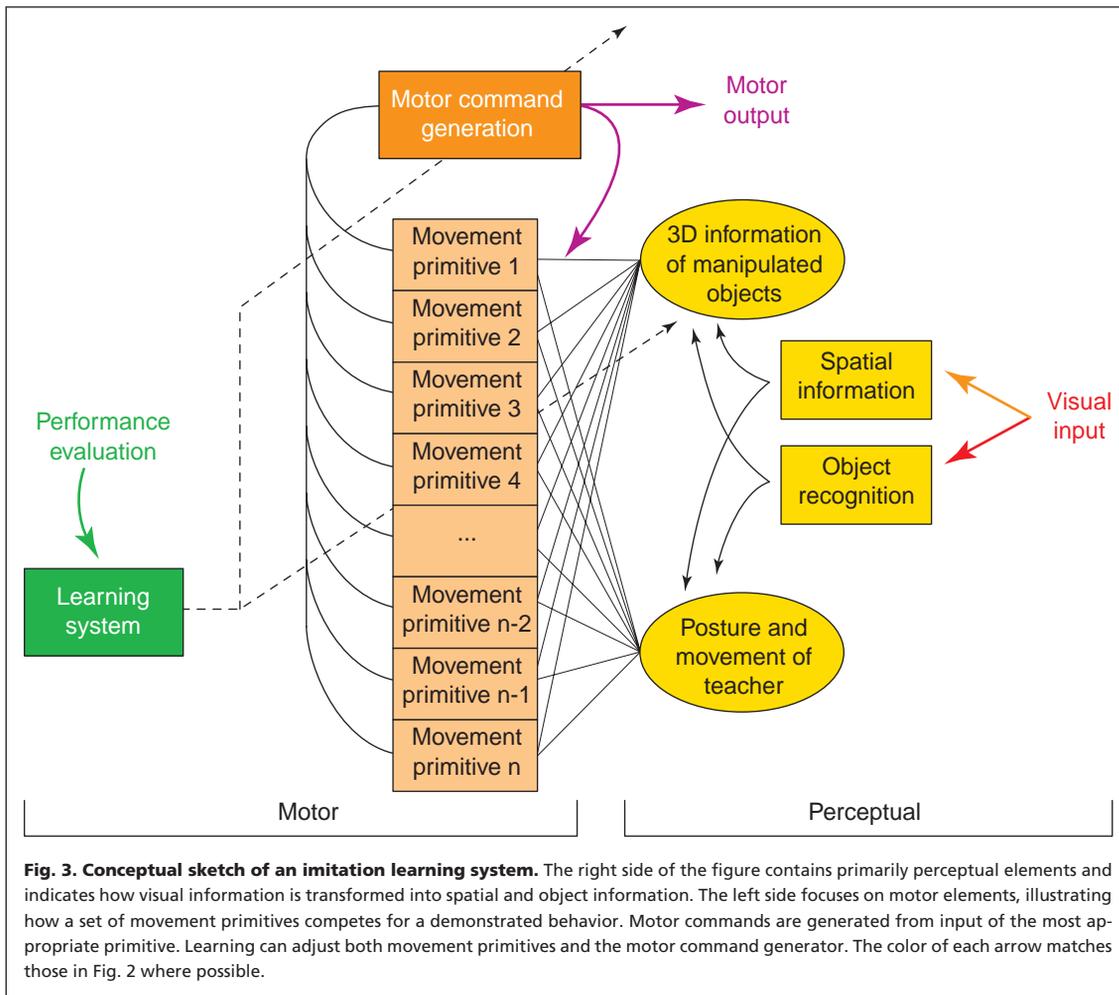
Whether mirror neurons are really part of the imitation system, however, has remained speculative so far. Neither did the neurophysiological experiments on mirror neurons require true imitation behavior nor did they test a movement repertoire that monkeys are known to imitate. Gallese and Goldman rather suggest that mirror neurons participate in 'mind reading', a process accomplished by using one's own mental apparatus to predict the psychological state of others through mental simulations[42]. In computational approaches, mental simulations are based on predictive forward models[43]. As will be discussed in the next section, forward models are likely to be an important component of an imitation learning system. From this viewpoint, simulation theory and imitation learning may share some computational mechanisms.

The connection between neurons in STs and F5, however, remains unclear at present. Area F5 does receive input from SII, 7b, and AIP (Fig. 2), and sends projections to F1 (i.e. primary motor cortex), the spinal cord, and also back to AIP (Ref. 35). AIP neurons were shown to be either sensitive to object specific properties (shape, size, orientation) or to movements and grasps executed by the monkey, or to both visual and motor stimuli[44]. Following the additional observation that many AIP neurons were modulated during ongoing grasping movements, Jeannerod *et al.* hypothesized that AIP might provide the necessary spatial object information to F5 for continuous guidance of grasping movement[34]. Because AIP receives input from other parietal areas, the connection to visual information through the dorsal stream would seem to be complete. However, this visual information is more concerned with object properties, rather than information about the movement of others. Arbib and Rizzolatti speculate that a connection from STs to area 7b and then to F5 provides the input about the movements of others[45]. Such statements are partially supported by imaging studies that demonstrated the simultaneous activation of the inferior area 6 and STs in humans[46,47]. The STs→7b→F5 connection would complete a first hypothesis for a neurophysiological pathway of imitation.

Imitation from the viewpoint of robotics, artificial intelligence and neural computation

Symbolic approaches to imitation learning

At the beginning of the 1980s, the idea of imitation learning started to find increasing attention in the field of manipulator robotics as it seemed to be a promising route to automate the tedious manual programming of robots. Inspired by the ideas of artificial intelligence, symbolic reasoning was commonly chosen to approach imitation, as outlined in the following discussion (e.g. Refs 48–52). During a training phase, several example movements were generated under manual robot control that achieved a given task with a robot. Sensor readings, for example, position and force, were stored

Fig. 3. Conceptual sketch of an imitation learning system. The right side of the figure contains primarily perceptual elements and indicates how visual information is transformed into spatial and object information. The left side focuses on motor elements, illustrating how a set of movement primitives competes for a demonstrated behavior. Motor commands are generated from input of the most appropriate primitive. Learning can adjust both movement primitives and the motor command generator. The color of each arrow matches those in Fig. 2 where possible.

throughout the demonstration together with the positions and orientations of obstacles and the goal state. For imitation, the example movements were first segmented into subgoals and appropriate primitive actions to attain these subgoals. Primitive actions were commonly chosen to be the simple point-to-point movement that industrial robots employed at this time. Subgoals could be the robot's gripper orientation and position – defined in a geometrical relationship to the goal – after each primitive action, or they were labeled manually by the demonstrator[50]. Consequently, the demonstrated task was segmented into a sequence of state–action–state transitions. However, given some uncertainty in the environment and some variability between several demonstrated movements, it was necessary to consolidate all demonstrated movements. For this purpose, the state–action–state sequence was converted into symbolic 'if–then' rules, for instance, expressing the state in terms of 'aligned', 'in contact', 'near-to', and actions as 'move-to', 'grasp-object', 'move-above', etc. Appropriate numerical definitions of these symbols (e.g. at which distance threshold an object is near) were provided as prior knowledge. Such abstraction resulted in graph-based representations, each state becoming a graph node and each action a directed link between two nodes. Symbolic reasoning could now unify different graphical representations for the same task by merging and deleting nodes[49] such that the final program for imitating the demonstrated movement was achieved.

In essence, many recent robotics approaches to imitation learning have remained closely related to the example above. New elements include the use of visual input of the teacher and performance of movement segmentation from computer vision algorithms[53–55]. Other projects used data gloves[56] or marker-based observation systems as input for imitation learning[56]. More recently, research on imitation learning has been influenced increasingly by non-symbolic learning tools, for instance, artificial neural networks, fuzzy logic, and so on[57–59], thus entering the category of neural computation described in the following section.

Inductive approaches to imitation learning
The major components of an imitation learning system are shown in Fig. 3. Most research projects have either focused on the perceptual side of imitation by investigating movement systems with low complexity[54,60,61] (e.g. artificial oculomotor systems, mobile robots, pick-and-place industrial robots), or on the motor end by assuming the existence of all necessary perceptual information. In the following, we will primarily focus on projects employing the latter strategy.

After spatial information about the teacher's movement and object information has been extracted, one of the major questions becomes how such information should be converted into action. For this purpose, Fig. 3 alludes to the concept of **movement primitives**, also called 'movement schemas', 'basis behaviors', 'units of action', 'macro actions', etc.[36,62,63] Movement primitives are sequences of

action that accomplish a complete goal-directed behavior. A movement primitive can be as simple as an elementary action in the symbolic approaches to imitation, for example, 'go forward', 'go backward', etc. However, as discussed in the introduction of this review, such low-level representations do not scale well to learning in systems with many degrees-of-freedom. Thus, it is useful for a movement primitive to code complete temporal behaviors, like 'grasping a cup', 'walking', 'a tennis serve', etc. This coding results in a compact state–action representation where only a few parameters need to be adjusted for a specific goal. For instance, in reaching movements, the target state and movement duration are such parameters, or in a rhythmic movement, frequency and amplitude need to be specified[64]. Using such primitives dramatically reduces the number of parameters that need to be learned for a particular movement. The drawback is that the possible movement repertoire becomes more restricted.

The perceived action of the teacher is mapped onto a set of existing primitives in an **assimilation** phase. Subsequently, the most appropriate primitives are adjusted by learning to improve the performance in an **accommodation** phase. Figure 3 indicates such a process by highlighting the better-matching primitives with increasing line widths. If no existing primitive is a good match for the observed behavior, a new primitive must to be generated. This concept of movement primitives is closely related to the interpretation of mirror neurons in the previous section: mirror neurons are thought to code complete motor acts, that is, primitives. However, mirror neurons seem to be a more high-level indicator of which movement primitive is appropriate rather than directly the place of motor command generation. Movement primitives could also form the 'supramodal' representation system of Meltzoff and Moore in which the authors assume that matching between the perceived and one's own movement takes place[24]. Those authors suggested that in early infancy, matching is based on goal states for the motor system, while, at a higher developmental stage, matching may be based on a temporal sequence of goal states, or a transition between goal states. Goal states are some of the natural parameters of a movement primitive[64], and transitions between goal states are the primitive's spatio-temporal signature to which Meltzoff and Moore appeal as a candidate to achieve matching between different modes of sensory information.

The conceptual sketch also allows distinguishing true imitation from response facilitation. For explaining response facilitation, every movement primitive has to keep a frequency count of its past activations over a restricted temporal window. A demonstration of a familiar movement primitive by a teacher will change this frequency distribution, that is, it will be biased towards the observed primitive. If spontaneous action is generated according to the probability distribution that is formed by the frequency counts, the demonstrated behavior is more likely to occur. The described process can readily be modeled by using Bayesian statistics[25]. For true imitation, no existing primitive is a good match for the demonstrated behavior such that learning is required to either adapt an existing primitive or to generate a new one. From this viewpoint, response

facilitation and true imitation share largely the same circuitry and have to tackle similar computational problems.

Imitation learning of novel behaviors
Several research projects have focused on imitation as a method of creating novel behaviors. Three major approaches can be distinguished:

(1) *Learning a control policy directly*: The demonstrated behavior can be used to learn the appropriate **control policy** directly by supervised learning. For this purpose, the **state**, *x*, and the **action**, *u*, of the teacher need to be observable and identifiable. This prerequisite, shared by all forms of imitation learning, imposes a serious constraint because, normally, motor commands and the internal variables of the teacher are hidden from the observer. Thus, a movement primitive needs to be defined in coordinate frame based on variables that can be perceived (e.g. the acceleration of the fingertip in the task of pole balancing instead of the commands sent to the motor neurons). Aboaf *et al.* called such an approach **task-level learning**[65], and by analogy, 'task-level imitation' can be used in the context of imitation learning. Task-level imitation requires prior knowledge of how a task-level command (e.g. the desired acceleration of the fingertip) can be converted into an actuator-level command. Motor control needs to be modular for this purpose; that is, separate processes for, at least, movement planning and execution need to be assumed[66,67].

Direct-policy learning was conducted for the task of pole balancing with a computer-simulated pole[68,69]. For this purpose, a neural network was trained on task-level data recorded from a human demonstration. Similarly, several mobile robotics groups adopted imitation by direct-policy learning using a 'robot teacher'[60,70–72]. For example, the 'robot student' followed the 'robot teacher's' movements in a specific environment, mimicked its actions, and learned to associate which action to choose in which state. Afterwards, the robot student had the same competence as the teacher in this environment. Importantly, in all these direct-policy learning approaches, there is no need for the student to know the goal of the teacher. Imitation learning is greatly simplified in this manner. However, the student will not be able to undergo self-improvement unless an explicit reward signal, usually generated from an optimization criterion, is provided to the student, as in the following approaches.

(2) *Learning from demonstrated trajectories*: A second approach to learning novel behaviors is based on building policies out of demonstrated trajectories. This idea was explored with an anthropomorphic robot arm for dynamic manipulation tasks, for instance, learning a tennis forehand and the game of kendama ('ball-in-the-cup')[73,74]. At the outset, a human demonstrated the task, and his/her movement was recorded with marker-based optical recording equipment. This process resulted in data about the movement of the manipulated object in Cartesian coordinates, as well as the movement of the actuator (arm) in terms of joint angle coordinates. For imitation learning, a hybrid strategy was chosen. Initially, the robot aimed at indiscriminate imitation in task space based on position data of the endeffector, while trying to use an arm posture as similar as possible to the demonstrated posture of the teacher. Afterwards,

based on manually provided knowledge of the task goal in the form of an optimization criterion, the robot's performance improved by trial and error learning until the task was accomplished. For this purpose, the desired end-effector trajectory of the robot was approximated by **splines**, and the spline nodes, called via-points, were adjusted by supervised learning until the task was fulfilled. Using this method, the robot learned to manipulate a stochastic, dynamic environment within a few trials.

(3) *Model-based imitation learning*: A third approach to learning a novel primitive employs model-based learning[75,76]. From the demonstrated behavior, the dynamics of the task is approximated in the form of a predictive forward model (see Ref. 76). Given knowledge of the task goal, the task-level policy of the movement primitive can be computed with reinforcement learning procedures based on the learned model. For example, Schaal and Atkeson showed how the model-based approach allowed an anthropomorphic robot arm to learn the task of pole-balancing in just a single trial, and the task of a 'pendulum swing-up' in only three to four trials[75,77]. These authors also demonstrated that task-level imitation based on direct-policy learning, augmented with subsequent self-learning, can be rather fragile and does not necessarily provide significant learning speed improvement over pure trial-and-error learning without a demonstration.

Implications for computational models of imitation learning
The approaches discussed in the previous paragraphs illustrated some computational ideas for how novel behaviors can be learned by imitation. Interesting insights into these methods can be gained by analysing the process of how a perceived behavior is mapped onto a set of existing primitives. Two major questions (see Ref. 24) become (1) What is the matching criterion for recognizing a behavior?, and (2) In which coordinate frame does matching take place?

If only the control policy of the movement primitive exists, finding a matching criterion becomes difficult. One solution would be to try a primitive, observe its outcome in task space, and generate a performance criterion based on the similarity between the executed and the teacher's behavior. This procedure needs to be repeated for every primitive in the repertoire and is thus quite inefficient. Another possibility arises if the primitive outputs task-level commands that can be compared directly with the teacher's performance. In this case, the movement primitive acts simultaneously as a forward model[64], an approach that is described in more detail below.

The via-point method[73] can easily be adapted for movement recognition. Via-points are a parsimonious representation of a movement and may be used for classification as well. For example, a demonstrated movement can be transformed into via-points, and the number and location of via-points can be compared against those of existing movement primitives in order to choose the best match, as has been demonstrated for handwriting and character recognition[78]. Despite the fact that there are various open issues with regard to translation, scale and rotation invariance in the via-point approach, the suggested bi-directional interaction between perception and action is noteworthy.

Movement recognition is directly accomplished with the movement generating mechanism. This concept is compatible with what has been observed in mirror neurons, and it also ties into other research projects that emphasize the bi-directional interaction of generative and recognition models in unsupervised learning[79].

Lastly, a third alternative should be considered, based on forward models[75,76], but put into a multiple-model competition[67,80,81]. If every movement primitive has a forward model, all primitives can simultaneously attempt to predict the teacher's behavior in the form of a mental simulation loop, indicated by the 'efference copy' arc in Fig. 3. The motor command will be generated by the primitives that make the most accurate prediction. As in the via-point approach, movement recognition is based on the movement generating mechanism (as, for example, in Refs 67, 81). Such an approach is particularly easy if movement primitives are coded in task space because the prediction of a primitive will be directly comparable to the teacher's performance. If the primitive operates in actuator space, additional coordinate transformations and, possibly, inverse models are needed before a task level comparison can be accomplished[67]. Because forward models have various additional advantages for motor control[43,67], movement recognition based on forward models could offer the most general and powerful solution to the perception to action mapping. Interestingly, movement recognition based on forward models integrates smoothly with the simulation theory of mind reading[42]. It can also provide a computational mechanism for Meltzoff and Moore's ideas of imitation as a basic communication skill and their 'Active Intramodal Matching' model of imitation[23,24].

One final issue concerns the imitation of complex motor acts that involve learning a sequence of primitives and when to switch between them. In this context, Fagg and Arbib provided a model of reaching and grasping based on the known anatomy of the fronto-parietal circuits, including the mirror neuron system[82]. Essentially, their model employed a recurrent neural network that sequenced and switched between motor schemas based on sensory cues. In a robotic study, Pook and Ballard used hidden Markov models to learn appropriate sequencing from demonstrated behavior[58]. There is also large body of literature in the field of time series segmentation[83–85] that employed competitive learning and forward models for recognition and sequencing in a way that is easily adapted for imitation learning as illustrated in Fig. 3.

Is imitation learning the route to humanoid robots?
In the introduction of this article, we appealed to a pragmatic view of imitation learning as a means to speed up learning in complex high dimensional motor systems, such as humanoid robots. This view emerged from the lack of theories of motor learning that are able to work efficiently in high dimensional spaces. Interestingly, the apparently simple idea of imitation opened a Pandora's box of important computational questions in perceptual motor control. None of the approaches described in this article could provide satisfying answers to questions of appropriate perceptual representations for imitation, motor representations,
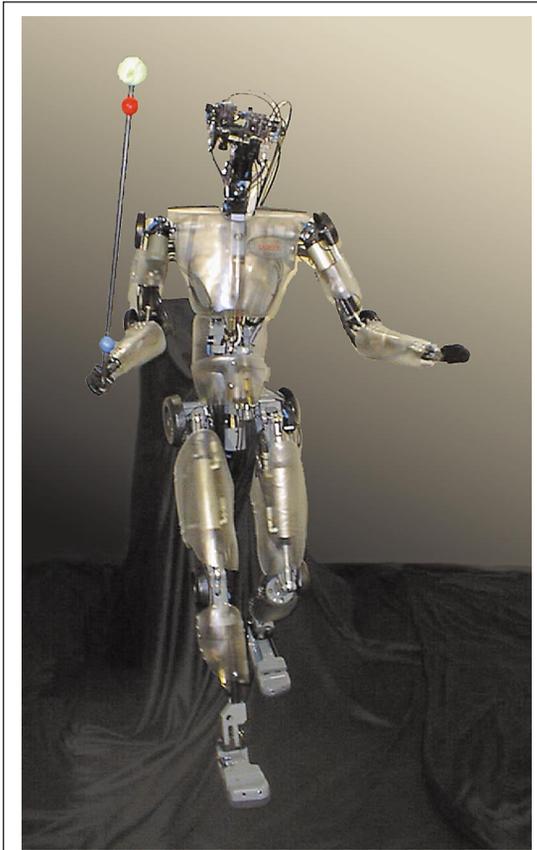
Fig. 4. Humanoid robot at the Dynamic Brain project at the ATR Labs in Japan.

and the learning of these representations. However, consensus may be reached that research into imitation needs to include a theory of motor learning, of compact state–action representations or movement primitives, and of the interaction of perception and action. It is likely that these three components cannot be studied in isolation: perceptual representations serve motor representations, motor representations facilitate perception, and learning provides the mutual constraints between them. Learning theories based on such reciprocal interactions are currently under investigation in computational neuroscience[79,86]. On this view, instead of being an idiosyncratic research topic, imitation learning could be conceived of as a research strategy

that channels investigations in computational motor control towards the important topic of action-perception coupling.

While it seems fair to say that a formal understanding of imitation would certainly be a major step towards creating humanoid robots, biomimetic robotic systems, in particular humanoid robots, have also become a new tool to investigate cognitive and biological questions[87]. For instance, the Cog Project at the Massachusetts Institute of Technology investigates how far a humanoid robot could become 'cognitive' via a bottom-up approach[88]. The Dynamic Brain Project at the ATR Laboratories in Japan has been conducting research with an anthropomorphic robot arm for several years[73,75] and is currently working with a novel humanoid robot (Fig. 4) to study theories of computational neuroscience and imitation learning. Other humanoid robot projects include a research group at the University of Tokyo[89], and the humanoid robotics project at Waseda University[90]. The blending of psychology, neuroscience and engineering in such research projects seems to be a new trend that will be beneficial for advancing knowledge in both technological and biological sciences.

References

1 Sciavicco, L. and Siciliano, B. (1996) *Modeling and Control of Robot Manipulators*, McGraw-Hill

2 Hirai, K. (1998) Current and future perspective of Honda humanoid robot, in *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robot and Systems*, Grenoble, France, pp. 500–508, IEEE

3 Hirai, K. *et al*. (1998) The development of Honda humanoid robot, in *IEEE International Conference on Robotics and Automation*, Leuven, Belgium, pp. 1321–1326, IEEE

4 Kaelbling, L.P., Littman, M.L. and Moore, A.W. (1996) Reinforcement learning: a survey *J. Artif. Intell. Res.* 4, 237–285

5 Sutton, R.S. and Barto, A.G. (1998) *Reinforcement Learning: An Introduction*, MIT Press

6 Sutton, R.S. (1988) Learning to predict by the methods of temporal differences *Machine Learn.* 3, 9–44

7 Atkeson, C.G. and Schaal, S. (1997) Learning tasks from a single demonstration, in *IEEE International Conference on Robotics and Automation (ICRA97)*, pp. 1706–1712, IEEE

8 Byrne, R.W. and Russon, A.E. Learning by imitation: a hierachical approach *Behav. Brain Sci.* (in press)

9 Tomasello, M., Kruger, A.C. and Ranter, H.H. (1993) Cultural learning *Behav. Brain Sci.* 16, 495–552

10 Romanes, G.J. and Darwin, C. (1884) *Mental Evolution in Animals*, Appleton & Co.

11 Thorndike, E.L. (1898) Animal intelligence: an experimental study of the associative process in animals *Psychol. Rev. Monogr.* 2, 551–553

12 Piaget, J. (1951) *Play, Dreams, and Imitation in Childhood*, W.W. Norton

13 Meltzoff, A.N. and Moore, M.K. (1977) Imitation of facial and manual gestures by human neonates *Science* 198, 74–78

14 Meltzoff, A.N. and Moore, M.K. (1983) Newborn infants imitate adult facial gestures *Child Dev.* 54, 702–709

15 Byrne, R. (1995) *The Thinking Ape*, Oxford University Press

16 Tomasello, M., Savage-Rumbaugh, S. and Kruger, A.C. (1993) Imitative learning of actions on objects by children, chimpanzees, and

Outstanding questions

- *Learning perceptual representations*: How can appropriate representations of the identity and movement of others be developed in an automated fashion from visual input? Is it necessary that such representations develop simultaneously with the motor representations?
- *Movement primitives*: Is there a basic set of primitives that can initialize imitation learning? How complex are the most elementary primitives in this set? How can new primitives be learned, and old primitives be combined to form higher level movement primitives? How is sequencing and the recognition of sequences of movement primitives accomplished?
- *Movement recognition through movement generation*: Is the movement generating mechanism directly employed for movement recognition? What representation allows such a dual use of the motor system? Are movement primitives simultaneously predictive forwards models?
- *Understanding task goals*: How can the intention of a demonstrated movement be recognized and converted to the imitator's goal?

enculturated chimpanzees *Child Dev.* 64, 1688–1705

17 Abravanel, E. (1991) Does immediate imitation influence long-term memory for observed actions? *J. Exp. Child Psychol.* 51, 235–244

18 Weeks, D.L., Hall, A.K. and Anderson, L.P. (1996) A comparison of imitation strategies in observational learning of action patterns *J. Motor. Behav.* 28, 348–358

19 Mataric, M.J. and Pomplun, M. (1998) Fixation behavior in observation and imitation of human movement *Cognit. Brain Res.* 7, 191–202

20 Turvey, M.T. (1977) Preliminaries to a theory of action with reference to vision, in *Perceiving, Acting, and Knowing: Toward an Ecological Psychology* (Shaw, R. and Bransford, J., eds), pp. 253–257, Erlbaum

21 Kugler, P.N. and Turvey, M.T. (1987) *Information, Natural Law, and the Self-Assembly of Rhythmic Movement*, Erlbaum

22 Bertenthal, B.I. (1996) Origins and early development of perception, action, and representation *Annu. Rev. Psychol.* 47, 431–459

23 Meltzoff, A.N. and Moore, M.K. (1995) Infant's understanding of people and things: from body imitation to folk psychology, in *The Body and The Self* (Bermúdez, J.L., Marcel, A. and Eilan, N., eds), pp. 43–69, MIT Press

24 Meltzoff, A.N. and Moore, M.K. (1997) Explaining facial imitation: a theoretical model *Early Development and Parenting* 6, 179–192

25 Bishop, C.M. (1995) *Neural Networks for Pattern Recognition*, Oxford University Press

26 Rizzolatti, G. *et al.* (1996) Premotor cortex and the recognition of motor actions *Cognit. Brain Res.* 3, 131–141

27 Perrett, D.I. *et al.* (1990) Social signals analysed at the single cell level: someone is looking at me, something touched me, something moved! *Int. J. Comp. Psychol.* 4, 25–55

28 van Essen, D.C. and Maunsell, J.H.R. (1983) Hierarchical organization and functional streams in the visual cortex *Trends Neurosci.* 6, 370–375

29 Perrett, D.I. *et al.* (1985) Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: a preliminary report *Behav. Brain Res.* 16, 153–170

30 Perrett, D.I. *et al.* (1989) Frameworks of analysis for the neural representation of animate objects and actions *J. Exp. Biol.* 146, 87–113

31 Oram, M.W. and Perrett, D.I. (1996) Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey *J. Neurophysiol.* 76, 109–129

32 Hietanen, J.K. and Perrett, D.I. (1996) Motion sensitive cells in the macaque superior temporal polysensory area: response discrimination between self-generated and externally generated pattern motion *Behav. Brain Res.* 76, 155–167

33 Rizzolatti, G. *et al.* (1988) Functional organization of inferior area 6 in the macaque monkey: II. Area F5 and the control of distal movements *Exp. Brain Res.* 71, 491–507

34 Jeannerod, M. *et al.* (1995) Grasping objects: the cortical mechanisms of visuomotor transformation *Trends Neurosci.* 18, 314–320

35 Murata, A. *et al.* (1997) Object representation in the ventral premotor cortex (area F5) of the monkey *J. Neurophysiol.* 78, 2226–2230

36 Arbib, M.A. (1981) Perceptual structures and distributed motor control, in *Handbook of Physiology, Section 2: The Nervous System* (*Vol. II, Motor Control, Part 1*) (Brooks, V.B., ed.), pp. 1449–1480, American Physiological Society

37 di Pellegrino, G. *et al.* (1992) Understanding motor events: a neurophysiological study *Exp. Brain Res.* 91, 176–180

38 Decety, J. *et al.* (1994) Mapping motor representations with positron emission tomography *Nature* 371, 600–602

39 Fadiga, L. *et al.* (1995) Motor facilitation during action observation: a magnetic stimulation study *J. Neurophysiol.* 73, 2608–2611

40 Decety, J. (1996) Do imagined and executed actions share the same neural substrate? *Cognit. Brain Res.* 3, 87–93

41 Rizzolatti, G. and Arbib, M.A. (1998) Language within our grasp *Trends Neurosci.* 21, 188–194

42 Gallese, V. and Goldman, A. (1998) Mirror neurons and the simulation theory of mind-reading *Trends Cognit. Sci.* 2, 493–501

43 Miall, R.C. and Wolpert, D.M. (1996) Forward models for physiological motor control *Neural Networks* 9, 1265–1285

44 Sakata, H. *et al.* (1992) Hand-movement related neurons of the posterior parietal cortex of the monkey: their role in visual guidance of hand movements, in *Control of Arm Movement in Space* (Caminiti, R., Johnson, P.B. and Burnod, Y., eds), pp. 185–198, Springer-Verlag

45 Arbib, M.A. and Rizzolatti, G. (1996) Neural expectations: a possible evolutionary path from manual skills to language *Commun. Cognit.* 29, 393–424

46 Rizzolatti, G. *et al.* (1996) Localization of grasp representations in humans by PET: 1. Observation versus execution *Exp. Brain Res.* 111, 246–252

47 Grafton, S.T. *et al.* (1996) Localization of grasp representations in humans by positron emission tomography: 2. Observation compared with imagination *Exp. Brain Res.* 112, 103–111

48 Lozano-Pérez, T. (1982) Task-planning, in *Robot Planning: Motion and Control* (Brady, M. *et al.*, eds), pp. 473–498, MIT Press

49 Dufay, B. and Latombe, J.C. (1984) An approach to automatic robot programming based on inductive learning *Int. J. Robotics Res.* 3, 89–101

50 Levas, A. and Selfridge, M. (1984) A user-friendly high-level robot teaching system, in *International Conference on Robotics and Automation*, Altanta, GA, pp. 413–416

51 Segre, A.B. and DeJong, G. (1985) Explanation-based manipulator learning: acquisition of planning ability through observation, in *International Conference on Robotics and Automation*, pp. 555–560

52 Segre, A.M. (1988) *Machine Learning of Robot Assembly Plans*, Kluwer Academic Publishers

53 Kuniyoshi, Y., Inaba, M. and Inoue, H. (1989) Teaching by showing: generating robot programs by visual observation of human performance, in *Proceedings of the International Symposium of Industrial Robots*, pp. 119–126

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

55 Ikeuchi, K., Kawade, M. and Suehiro, T. (1993) Assembly task recognition with planar, curved and mechanical contacts, in *Proceedings IEEE International Conference on Robotics and Automation*, Atlanta, GA, pp. 688–693, IEEE

56 Tung, C.P. and Kak, A.C. (1995) Automatic learning of assembly task using a DataGlove system, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Pittsburgh, PA, pp. 1–8, IEEE/RSJ

57 Dillmann, R., Kaiser, M. and Ude, A. (1995) Acquisition of elementary robot skills from human demonstration, in *International Symposium on Intelligent Robotic Systems (SIRS'95)*, Pisa, Italy, pp. 1–38

58 Pook, P.K. and Ballard, D.H. (1993) Recognizing teleoperated manipulations, *Proceedings IEEE International Conference on Robotics and Automation*, Atlanta, GA, pp. 913–918

59 Hovland, G.E., Sikka, P. and McCarragher, B.J. (1996) Skill acquisition from human demonstration using a hidden Markov Model, in *IEEE International Conference on Robotics and Automation*, Minneapolis, MN, pp. 2706–2711, IEEE

60 Grudic, G.Z. and Lawrence, P.D. (1996) Human-to-robot skill transfer using the SPORE approximation, in *IEEE International Conference on Robotics and Automation*, Minneapolis, MN, pp. 2962–2967, IEEE

61 Demiris, J. *et al.* (1997) Deferred imitation of human head movements by an active stereo vision head, in *IEEE International Workshop on Robot Human Communication*, Sendai, Japan, pp. 45–51, IEEE

62 Sternad, D. and Schaal, D. (1999) Segmentation of endpoint trajectories does not imply segmented control *Exp. Brain Res.* 124, 118–136

63 Sutton, R.S. *et al.* Improved switching among temporally abstract actions, *Advances in Neural Information Processing Systems* (Vol. 11), MIT Press (in press)

64 Schaal, S. and Sternad, D. (1998) Programmable pattern generators, in *Proceedings of 3rd International Conference on Computational Intelligence in Neuroscience*, Research Triangle Park, NC, pp. 48–51

65 Aboaf, E.W., Drucker, S.M. and Atkeson, C.G. (1989) Task-level robot learing: juggling a tennis ball more accurately, in *Proceedings of IEEE Interational Conference on Robotics and Automation*, Scottsdale, Arizona, pp. 331–348, IEEE

66 Wolpert, D.M. (1997) Computational approaches to motor control *Trends Cognit. Sci.* 1, 209–216

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

68 Widrow, B. and Smith, F.W. (1964) Pattern recognizing control systems, in *1963 Computer and Information Sciences (COINS) Symposium Proceedings*, pp. 288–317, Spartan

69 Nechyba, M.C. and Xu, Y. (1995) Human skill transfer: neural networks as learners and teachers, in *IEEE/RSJ International Conference on Intelligence Robots and Systems*, Pittsburgh, PA, pp. 314–319, IEEE/RSJ

70 Lin, L-J. (1991) Programming robots using reinforcement learning and teaching, in *Proceedings of theNinth National Conference on Artificial Intelligence*, pp. 781–786

71 Hayes, G. and Demiris, J. (1994) A robot controller using learning by imitation, in *Proceedings of the 2nd International Symposium on Intelligent Robotic Systems*, Grenoble, France (Borkowski, A. and Crowley. J.L., eds), pp. 198–204, LIFTA–IMAG

72 Dautenhahn, K. (1995) Getting to know each other: artificial social intelligence for autonomous robots *Robotics and Autonomous Systems* 16, 333–356

73 Miyamoto, H. *et al*. (1996) A Kendama learning robot based on bi-directional theory *Neural Networks* 9, 1281–1302

74 Miyamoto, H. and Kawato, M. (1998) A tennis serve and upswing learning robot based on bi-directional theory *Neural Networks* 11, 1331–1344

75 Schaal, S. (1997) Learning from demonstration, in *Advances in Neural Information Processing Systems* (Vol. 9) (Mozer, M.C., Jordan, M. and Petsche, T., eds), pp. 1040–1046, MIT Press

76 Wolpert, D.M., Miall, R.C. and Kawato, M. (1998) Internal models in the cerebellum *Trends Cognit. Sci.* 2, 338–347

77 Atkeson, C.G. and Schaal, S. (1997) Robot learning from demonstration, in *International Conference on Machine Learning*, pp. 11–73

78 Wada, Y. and Kawato, M. (1995) A theory for cursive handwriting based on the minimization principle *Biol. Cybern.* 73, 3–13

79 Dayan, P. *et al*. (1995) The Helmholtz machine *Neural Comput.* 7, 889–904

80 Murray-Smith, R. and Johanson, T.A. (1998) *Multiple Model Approaches to Modelling and Control*, Taylor & Francis

81 Demiris, J. and Hayes, G. (1999) Active and passive routes to imitation in *Proceedings of the AISB '99 Symposium of Imitation in Animals and Artifacts*, pp. 81–87, AISB

82 Fagg, A.H. and Arbib, M.A. (1998) Modeling parietal–premotor interactions in primate control of grasping *Neural Networks* 11, 1277–1303

83 Weigend, A.S., Mangeas, M. and Srivastava, A.N. (1995) Nonlinear gated experts for time series: discovering regimes and avoiding overfitting *Int. J. Neural Syst.* 6, 373–399

84 Cacciatore, T.W. and Nowlan, S.J. (1994) Mixtures of controllers for jump linear and non-linear plants, in *Advances in Neural Information Processign Systems* (Vol. 6) (Cowen, J.D., Tesauro, G. and Alspector, J., eds), pp. 719–726, Morgan Kaufmann

85 Pawelzik, K., Kohlmorgen, J. and Müller, K.R. (1996) Annealed competition of experts for a segmentation and classification of switching dynamics *Neural Comput.* 8, 340–356

86 Kawato, M. (1996) Bi-directional theory approach to integration, in *Attention and Performance, Vol. XVI* (Konczak, J. and Thelen, E., eds), pp. 335–367, MIT Press

87 Beer, R.D. *et al*. (1998) Bio-robotic approaches to the study of motor systems *Curr. Opin. Neurobiol.* 8, 777–782

88 Brooks, R.A. *et al*. The Cog Project: building a humanoid robot, *Lecture Notes in Computer Science*, Springer-Verlag (in press)

89 Korno, A. *et al*. (1997) Development of a humanoid robot Saika, in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'97)*, pp. 805–810, IEEE

90 Kikuchi, H. *et al*. (1998) Controlling gaze of humanoid in communication with human, *International Conference on Intelligent Robots and Systems*, Victoria, BC, pp. 255–260