

## 11 A Bayesian model of imitation in infants and robots

---

*Rajesh P. N. Rao, Aaron P. Shon and Andrew N. Meltzoff*

### 11.1 Introduction

Humans are often characterized as the most behaviourally flexible of all animals. Evolution has stumbled upon an unlikely but very effective trick for achieving this state. Relative to most other animals, we are born 'immature' and helpless. Our extended period of infantile immaturity, however, confers us with benefits. It allows us to learn and adapt to the specific physical and cultural environment into which we are born. Instead of relying on fixed reflexes adapted for specific environments, our learning capacities allow us to adapt to a wide range of ecological niches, from Alaska to Mica, modifying our shelter, skills, dress and customs accordingly. A crucial component of evolution's design for human beings is imitative learning, the ability to learn behaviours by observing the actions of others.

Human adults effortlessly learn new behaviours from watching others. Parents provide their young with an apprenticeship in how to behave as a member of the culture long before verbal instruction is possible. In Western culture, toddlers hold telephones to their ears and babble into thin air. There is no innate proclivity to treat hunks of plastic in this manner, nor is it due to trial-and-error learning. Imitation is chiefly responsible.

Over the past decade, imitative learning has received considerable attention from cognitive scientists, evolutionary biologists, neuroscientists and robotics researchers. Discoveries in developmental psychology have altered theories about the origins of imitation and its place in human nature. We used to think that humans gradually learned to imitate over the first several years of life. We now know that newborns can imitate body movements at birth (Meltzoff and Moore, 1983, 1997). Such imitation reveals an innate link between observed and executed acts, with important

*Imitation and Social Learning in Robots, Humans and Animals*, ed. Chrystopher L. Nehaniv and Kerstin Dautenhahn. Published by Cambridge University Press.  
© Cambridge University Press 2007.

implications for neuroscience. Evolutionary biologists are using imitation in humans and non-human animals as a tool for examining continuities and discontinuities in the evolution of mind.

Darwin inquired about imitation in non-human animals, but the last 10 years have seen a greater number of controlled studies of imitation in monkeys and great apes than in the previous 100 years. The results indicate that monkey imitation is hard to come by in controlled experiments, belying the common wisdom of 'monkey see monkey do' (Tomasello and Call, 1997; Visalberghi and Frigaszy, 2002; Whiten, 2002). Non-human primates and other animals (e.g. songbirds) imitate, but their imitative prowess is more restricted than that of humans (Meltzoff, 1996). Meanwhile, neuroscientists and experimental psychologists have started investigating the neural and psychological mechanisms underlying imitation, including the exploration of 'mirror neurons' and 'shared neural representations' (e.g. Decety, 2002; Prinz, 2002; Rizzolatti et al., 2002; Meltzoff and Decety, 2003; Jackson et al., 2006).

The robotics community is becoming increasingly interested in robots that can learn by observing movements of a human or another robot. Such an approach, also called 'learning by watching' or 'learning by example', promises to revolutionize the way we interact with robots by offering a new, extremely flexible, fast and easy way of programming robots (Berthouze and Kuniyoshi, 1998; Mataric and Pomplun, 1998; Billard and Dautenhahn, 1999; Breazeal and Scassellati, 2002; Dautenhahn and Nehaniv, 2002; see also *Robotics and Autonomous Systems* (special issue), 2004). This effort is also prompting an increased cross-fertilization between the fields of robotics and human psychology (Demiris et al., 1997; Schaal, 1999; Demiris and Meltzoff, in press).

In this chapter, we set the stage for re-examining robotic learning by discussing Meltzoff and Moore's theory about how infants learn through imitation (Meltzoff, 2005, 2006; Meltzoff and Moore, 1997). They suggest a four-stage progression of imitative abilities: (1) body babbling, (2) imitation of body movements, (3) imitation of actions on objects and (4) imitation based on inferring intentions of others. We formalize these four stages within a probabilistic framework that is inspired by recent ideas from machine learning and statistical inference. In particular, we suggest a Bayesian approach to the problem of learning actions through observation and imitation, and explore its connections to recently proposed ideas regarding the importance of internal models in sensorimotor control. We conclude by discussing two main advantages of a probabilistic approach: (1) the development of robust algorithms for robotic imitation learning in noisy and uncertain environments and (2) the potential for applying Bayesian methodologies (such as manipulation of prior probabilities) and robotic technologies to obtain a deeper understanding

of imitative learning in human beings. Some of the ideas presented in this chapter appeared in a preliminary form in Rao and Meltzoff (2003).

## **11.2 Imitative learning in human infants**

Experimental results obtained by one of the authors (Meltzoff) and his colleagues over the past two decades suggest a progression of imitative learning abilities in infants, building up from 'body babbling' (random experimentation with body movements) in neonates to sophisticated forms of imitation in 18-month-old infants based on inferring the demonstrator's intended goals. We discuss these results below.

### *11.2.1 Body babbling*

An important precursor to the ability to learn via imitation is to learn how specific muscle movements achieve various elementary body configurations. This helps the child learn a set of 'motor primitives' that could be used as a basis for imitation learning. Experiments suggest that infants do not innately know what muscle movements achieve a particular goal state, such as tongue protrusion, mouth opening or lip protrusion. It is posited that such movements are learned through an early experiential process involving random trial-and-error learning. Meltzoff and Moore (1997) call this process 'body babbling'.

In body babbling, infants move their limbs and facial parts in repetitive body play analogous to vocal babbling. In the more familiar notion of vocal babbling, the muscle movements are mapped to the resulting auditory consequence; infants are learning an articulatory-auditory relation (Kuhl and Meltzoff, 1996). Body babbling works in the same way, a principal difference being that the process can begin in utero. What is acquired through body babbling is a mapping between movements and a resulting body part configuration such as: tongue-to-lips, tongue-between-lips, tongue-beyond-lips. Because both the dynamic patterns of movement and the resulting endstates achieved can be monitored proprioceptively, body babbling can build up a 'directory' (an 'internal model') mapping movements to goal states (Meltzoff and Moore, 1997). Studies of fetal and neonatal behaviour have documented self-generated activity that could serve this hypothesized body babbling function (Patrick et al., 1982). Neonates can acquire a rich store of information through such body babbling. With sufficient practice, they can map out an 'act space' enabling new body configurations to be interpolated within this space. Such an interpretation is consistent with the probabilistic notion of forward models and internal models discussed in Section 11.3.1.



**Figure 11.1** Imitative responses in 2- to 3-week-old infants (from Meltzoff and Moore, 1977).

### 11.2.2 *Imitating body movements*

In addition to body babbling, infants have been shown to demonstrate imitative learning. Meltzoff and Moore (1983, 1989) discovered that newborns can imitate facial acts. The mean age of these infants was 36 hours old, the youngest being 42 minutes old at the time of testing. Facial imitation in human infants thus suggests an innate mapping between observation and execution. Moreover, the studies provide information about the nature of the machinery infants use to connect observation and execution, as will be illustrated in the following brief review.

In Meltzoff and Moore (1977), 12- to 21-day-olds were shown to imitate four different gestures, including facial and manual movements. Infants didn't confuse either actions or body parts. They differentially responded to tongue protrusion with lip protrusion and without lip protrusion (Figure 11.1), showing that the specific *body part* can be identified. They also differentially responded to lip protrusion versus lip opening, showing that differential *action patterns* can be imitated with the same body part. This is confirmed by research showing that infants differentially imitate two different kinds of movements with the tongue (Meltzoff

and Moore, 1994, 1997). In all, there are more than 24 studies of early imitation from 13 independent laboratories, establishing imitation for an impressive set of elementary body acts (for a review, see Meltzoff, 2005). This does not deny further development of imitative abilities. Young infants are not as capable as older children in terms of motor skills and the neonate is certainly less self-conscious about imitating than the toddler (Meltzoff and Moore, 1997). The chief question for theory, however, concerns the neural and psychological processes linking the observation and execution of matching acts. How do infants crack the correspondence problem:<sup>1</sup> how can observed body states of a teacher be converted to 'my own body states'? Two discoveries bear on this issue.

First, early imitation is not restricted to direct perceptual-motor resonances. Meltzoff and Moore (1977) put a pacifier in infants' mouths so they couldn't imitate during the demonstration. After the demonstration was complete, the pacifier was withdrawn, and the adult assumed a passive face. The results showed that infants imitated during the subsequent 2.5-minute response period while looking at a passive face. More dramatically, 6-week-olds have been shown to perform deferred imitation across a 24-hour delay (Meltzoff and Moore, 1994). Infants saw a gesture on one day and returned the next day to see the adult with a passive-face pose. Infants stared at the face and then imitated from long-term memory.

Second, infants correct their imitative response (Meltzoff and Moore, 1994, 1997). They converge on the accurate match without feedback from the experimenter. The infant's first response to seeing a facial gesture is activation of the corresponding body part. For example, when infants see tongue protrusion, there is a dampening of movements of other body parts and a stimulation of the tongue. They do not necessarily protrude the tongue at first, but may elevate it or move it slightly in the oral cavity. The important point is that the tongue, rather than the lips or fingers, is energized before the precise imitative movement pattern is isolated. It is as if young infants isolate what part of their body to move before how to move it. Meltzoff and Moore (1997) call this 'organ identification'. Neurophysiological data show that visual displays of parts of the face and hands activate specific brain sites in monkeys and humans (Buccino *et al.*, 2001; Gross, 1992). Specific body parts could be neurally represented at birth and serve as a foundation for infant imitation.

In summary, the results suggest that: (1) newborns imitate facial acts that they have never seen themselves perform, (2) there is an innate observation-execution pathway in humans and (3) this pathway is mediated by

<sup>1</sup> For more on the issue of correspondence problems, see Part I, 'Correspondence Problem and Mechanisms' of this volume. -Ed.



**Figure 11.2** A 14-month-old infant imitating the novel action of touching a panel with the forehead (from Meltzoff, 1999).

a representational structure that allows infants to defer imitation and to correct their responses without any feedback from the experimenter.

### 11.2.3 Imitating actions on objects

More sophisticated forms of imitation than facial or manual imitation can be observed in infants who are several months older. In particular, the ability to imitate in these infants begins to encompass actions on objects that are external to the infant's body parts. In one study, toddlers were shown the act of an adult leaning forward and using the forehead to touch a yellow panel (Meltzoff, 1988b). This activated a microswitch, and the panel lit up. Infants were not given a chance for immediate imitation or even a chance to explore the panel during the demonstration session; therefore, learning by reinforcement and shaping was excluded. A one-week delay was imposed. At that point, infants returned to the laboratory and the panel was put out on the table. The results showed that 67% of the infants imitated the head-touch behaviour when they saw the panel. Such novel use of the forehead was exhibited by 0 % of the controls who had not seen this act on their first visit. An example of the head-touch response is shown in Figure 11.2

Successful imitation in this case must be based on observation of the adult's act because perception of the panel itself did not elicit the target behaviour in the naive infants. Moreover, the findings tell us something about what is represented. If the only thing they remembered is that 'the panel lit up' (an

object property), they would have returned and used their hands to press it. Instead, they re-enacted the same unusual act as used by the adult. The absent act had to have been represented and used to generate the behaviour a week later.

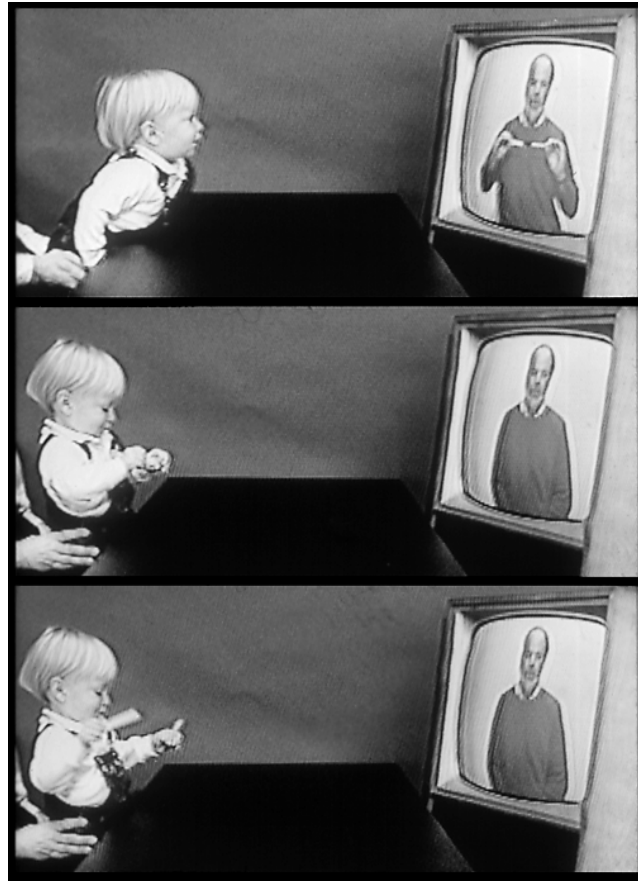
The utility of deferred imitation with 'real world' objects has also been demonstrated. Researchers have found deferred imitation of peer behaviour. In one study, 16-month-olds at a day-care centre watched peers play with toys in unique ways. The next day, an adult went to the infants' house (thereby introducing a change of context) and put the toys on the floor. The results showed that infants played with the toys in the particular ways that they had seen peers play 24 hours earlier (Hanna and Meltzoff, 1993). In another study, 14-month-olds saw a person on television demonstrate target acts toys (Figure 1 1.3). When they returned to the laboratory the next day, they were handed the toys for the first time. Infants re-enacted the events they saw on TV the previous day (Meltzoff, 1988a).

Taken together, these results indicate that infants who are between 1 and 1.5 years old are adept at imitating not only body movements but also actions on objects in a variety of contexts. For imitation to be useful in cultural learning, it would have to function with just such flexibility. The ability to imitate the actions of others on external objects undoubtedly played a crucial role in human evolution by facilitating the transfer of knowledge of tool use and other important skills from one generation to the next.

#### *1 1.2.4 Inferring intentions*

A sophisticated form of imitative learning is that requiring an ability to read below the perceived behaviour to infer the underlying goals and intentions of the actor. This brings the human infant to the threshold of 'theory of mind', in which they not only attribute visible behaviours to others, but develop the idea that others have internal mental states (intentions, perceptions, emotions) that underlie, predict and generate these visible behaviours.

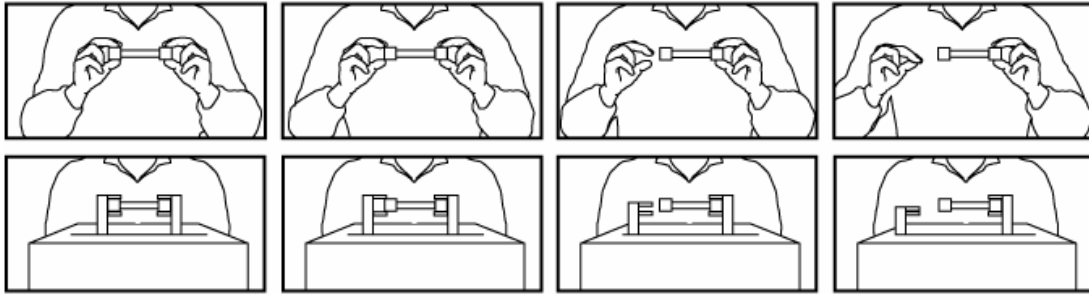
One study involved showing 18-month-old infants an unsuccessful act (Meltzoff, 1995). For example, an adult actor 'accidentally' under- or overshoot his target, or he tried to perform a behaviour but his hand slipped several times; thus the goal-state was not achieved (Figure 1 1.4, top row). To an adult, it was easy to read the actor's intention although he did not fulfill it. The experimental question was whether infants also read through the literal body movements to the underlying goal of the act. The measure of how they interpreted the event was what they chose to re-enact. In this case



**Figure 11.3** Infants as young as 14-months-old can imitate actions on objects as seen on TV (from Meltzoff, 1988a). Experiments have shown infants can also perform deferred imitation based on actions observed on TV the previous day (Meltzoff, 1988a).

the correct answer was not to imitate the movement that was actually seen, but the actor's goal, which remained unfulfilled.

The study compared infants' tendency to perform the target act in several situations: (1) after they saw the full target act demonstrated, (2) after they saw the unsuccessful attempt to perform the act, and after it was neither shown nor attempted. The results showed that 18-month-olds can infer the unseen goals implied by unsuccessful attempts. Infants who saw the unsuccessful attempt and infants who saw the full target act both produced target acts at a significantly higher rate than controls. Evidently, toddlers can understand our goals even if we fail to fulfill them.



**Figure 11.4** Human actor demonstrating an unsuccessful act (top panel) and an inanimate device mimicking the same movements (bottom). Infants attributed goals and intentions to the human but not to the inanimate device (from Meltzoff, 1995).

If infants can pick up the underlying goal or intention of the human act, they should be able to achieve the act using a variety of means. This was tested by Meltzoff (2006) in a study of 18-month-olds using a dumbbell-shaped object that was too big for the infants' hands. The adult grasped the ends of the dumbbell and attempted to yank it apart, but his hands slid off so he was unsuccessful in carrying out his intention. The dumbbell was then presented to the child. Interestingly, infants did not attempt to imitate the surface behaviour of the adult. Instead, they used novel ways to struggle to get the gigantic toy apart. They might put one end of the dumbbell between their knees and use both hands to pull it upwards, or put their hands on inside faces of the cubes and push outwards, and so on. They used different means than the demonstrator in order to achieve the same end. This fits with Meltzoff's (1995) hypothesis that infants had inferred the goal of the act, differentiating it from the surface behaviour that was observed.

People's acts can be goal-directed and intentional but the motions of inanimate devices are not—they are typically understood within the framework of physics, not psychology. In order to begin to assess whether young children distinguish between a psychological vs. purely physical framework, Meltzoff (1995) designed an inanimate device made of plastic, metal and wood. The device had poles for arms and mechanical pinchers for hands. It did not look human, but it traced the same spatio-temporal path that the human actor traced and manipulated the object much as the human actor did (see Figure 11.4). The results showed that infants did not attribute a goal or intention to the movements of the inanimate device. Infants were no more (or less) likely to pull the toy apart after seeing the unsuccessful attempt of the inanimate device as in the baseline condition. This was the case despite the fact that infants pulled the dumbbell apart if the inanimate device successfully completed this act.

Evidently, infants can pick up certain information from the inanimate device, but not other information: they can understand successes, but not failures. In the case of the unsuccessful attempts, it is as if they see the motions of the machine's mechanical arms as 'physical slippage' but not as an 'effort' or 'intention' to pull the object apart. They appear to make attributions of intentionality to humans but not to this mechanical device. One goal of our current research program is to examine just how 'human' a model must look (and act) in order to evoke this attribution. We plan to test infants' interpretations of the 'intentional' acts of robots.

### **11.3 A probabilistic model of imitation**

In recent years, probabilistic models have provided elegant explanations for a variety of neurobiological phenomena and perceptual illusions (for reviews, see Knill and Richards, 1996; *Rao et al.*, 2002). There is growing evidence that the brain utilizes principles such as probability matching and Bayes theorem for solving a wide range of tasks in sensory processing, sensorimotor control and decision-making. Bayes theorem in particular has been shown to be especially useful in explaining how the brain combines prior knowledge about a task with current sensory information and how information from different sensory channels is combined based on the noise statistics in these channels (see chapters in *Rao et al.*, 2002).

At the same time, probabilistic approaches are becoming increasingly popular in robotics and in artificial intelligence (AI). Traditional approaches to AI and robotics have been unsuccessful in scaling to noisy and realistic environments due to their inability to store, process and reason about uncertainties in the real world. The stochastic nature of most real-world environments makes the ability to handle uncertainties almost indispensable in intelligent autonomous systems. This realization has sparked a tremendous surge of interest in probabilistic methods for inference and learning in AI and robotics in recent years. Powerful new tools known as graphical models and Bayesian networks (Pearl, 1988; Jensen, 2001; Glymour, 2001) have found wide applicability in areas ranging from data mining and computer vision to bioinformatics, psychology and mobile robotics. These networks allow the probabilities of various events and outcomes to be inferred directly from input data based on the laws of probability and a representation based on graphs.

Given the recent success of probabilistic methods in AI/robotics and in modelling the brain, we believe that a probabilistic framework for imitation could not only enhance our understanding of human imitation but also provide new methods for imitative learning in robots. In this section, we explore a

formalization of Meltzoff and Moore's stages of imitative learning in infants within the context of a probabilistic model.

### *11.3.1 Body babbling: learning internal models of one's own body*

Meltzoff and Moore's theory about body babbling can be related to the task of learning an 'internal model' of an external physical system (also known as 'system identification' in the engineering literature). The physical system could be the infant's own body, a passive physical object such as a book or toy, or an active agent such as an animal or another human. In each of these cases, the underlying goal is to learn a model of the behaviour of the system being observed, i.e. to model the 'physics' of the system. Certain aspects of the internal model, such as the structure of the model and representation of specific body parts (such as the tongue), could be innately encoded and refined prior to birth (see Section 11.2.2) but the attainment of fine-grained control of movements most likely requires body babbling and interactions with the environment after birth.

A prominent type of internal model is a forward model, which maps actions to consequences of actions. For example, a forward model can be used to predict the next state(s) of an observed system, given its current state and an action to be executed on the system. Thus, if the physical system being modelled is one's own arm, the forward model could be used to predict the sensory (visual, tactile and proprioceptive) consequences of a motor command that moves the arm in a particular direction.

The counterpart of a forward model is an inverse model, which maps desired perceptual states to appropriate actions that achieve those states, given the current state. The inverse model is typically harder to estimate and is often ill-defined, due to many possible actions leading to the same goal state. A more tractable approach, which has received much attention in recent years (Jordan and Rumelhart, 1992; Wolpert and Kawato, 1998), is to estimate the inverse model using a forward model and appropriate constraints on actions (priors), as discussed below.

Our hypothesis is that the progression of imitative stages in infants as discussed in Section 11.2 reflects a concomitant increase in the sophistication of internal models in infants as they grow older. Intra-uterine and early post-natal body babbling could allow an infant to learn an internal model of its own body parts. This internal model facilitates elementary forms of imitation in Stage 2 involving movement of body parts such as tongue or lip protrusion. Experience with real-world objects after birth allows internal models of the physics of objects to be learned, allowing imitation of actions on such objects as seen in Stage 3. By the time infants are about 1.5 years old, they have interacted extensively with other

humans, allowing them to acquire internal models (both forward and inverse) of active agents with intentions (Meltzoff, 2006). Such learned forward models could be used to infer the possible goals of agents despite witnessing only unsuccessful demonstrations while the inverse models could be used to select the motor commands necessary to achieve the undemonstrated but inferred goals. These ideas are illustrated with a concrete example in a subsequent section.

### 11.3.2 Bayesian imitative learning

Consider an imitation learning task where the observations can be characterized as a sequence of discrete states  $s_1, s_2, \dots, s_N$  of an observed object.<sup>2</sup> A first problem that the imitator has to solve is to estimate these states from the raw perceptual inputs  $I_1, I_2, \dots, I_N$ . This can be handled using state estimation techniques such as the forward-backward algorithm for hidden Markov models (Rabiner and Juang, 1986) and belief propagation for arbitrary graphical models (Pearl, 1988; Jensen, 2001). These algorithms assume an underlying generative model that specifies how specific states are related to the observed inputs and other states through conditional probability matrices. We refer the interested reader to Jensen (2001) for more details. We assume the estimated states inferred from the observed input sequence are in object-centred coordinates.

The next problem that the imitator has to solve is the mechanism problem (Meltzoff and Moore, 1983, 1997) or correspondence problem (Nehaniv and Dautenhahn, 1998; Alissandrakis *et al.*, 2002; Nehaniv and Dautenhahn, 2002): how can the observed states be converted to 'my own body states' or states of an object from 'my own viewpoint'? Solving the correspondence problem involves mapping the estimated object-centred representation to an egocentric representation. In this chapter, for simplicity, we use an identity mapping for this correspondence function but the methods below also apply to the case of non-trivial correspondences (e.g. Nehaniv and Dautenhahn, 2001; Alissandrakis *et al.*, 2002a).

In the simplest form of imitation-based learning, the goal is to compute a set of actions that will lead to the goal state  $s_N$ , given a set of observed and remembered states  $s_1, s_2, \dots, s_N$ . We will treat  $s_t$  as the random variable for the state at time  $t$ . For the rest of the chapter, we assume discrete state and action spaces. Thus, the state  $s_t$  of the observed object could be one of  $M$  different values  $S_1, S_2, \dots, S_M$  while the current action  $a_t$  could be one of  $A_1, A_2, \dots, A_p$ .

<sup>2</sup> We have chosen to focus here on discrete state spaces but Bayesian techniques can also be applied to inference and learning in continuous state spaces (e.g. Bryson and Ho, 1975).

Consider now a simple imitation learning task where the imitator has observed and remembered a sequence of states (for example,  $S_7 \rightarrow S_1 \rightarrow \dots \rightarrow S_{12}$ ). These states can also be regarded as the sequence of sub-goals that need to be achieved in order to reach the goal state  $S_{12}$ . The objective then is to pick the action  $a_t$  that will maximize the probability of taking us from a current state  $s_t = S_i$  to a remembered next state  $s_{t+1} = S_j$ , given that the goal state  $g = S_k$  (starting from  $S_0 = S_7$  for our example). In other words, we would like to select the action  $a_t$  that maximizes:

$$P(a_t = A_i \mid s_t = S_i, s_{t+1} = S_j, g = S_k) \quad (11.1)$$

This set of probabilities constitutes the inverse model of the observed system: it tells us what action to choose, given the current state, the desired next state and the desired goal state.

The action selection problem becomes tractable if a forward model has been learned through body babbling and through experience with objects and agents in the world. The forward model is given by the set of probabilities:

$$P(s_{t+1} = S_j \mid s_t = S_i, a_t = A_i) \quad (11.2)$$

Note that the forward model is determined by the environment and is therefore assumed to be independent of the goal state  $g$ , i.e.:

$$\begin{aligned} P(s_{t+1} = S_j \mid s_t = S_i, a_t = A_i, g = S_k) \\ = P(s_{t+1} = S_j \mid s_t = S_i, a_t = A_i) \end{aligned} \quad (11.3)$$

These probabilities can be learned through experience in a supervised manner because values for all three variables become known at time step  $t + 1$ . Similarly, a set of prior probabilities on actions

$$P(a_t = A_i \mid s_t = S_i, g = S_k) \quad (11.4)$$

can also be learned through experience with the world, for example, by tracking the frequencies of each action for each current state and goal state.

Given these two sets of probabilities, it is easy to compute probabilities for the inverse model using Bayes' theorem. Given random variables  $A$  and  $B$ , Bayes' theorem states that:

$$P(B \mid A) = P(A \mid B)P(B)/P(A) \quad (11.5)$$

This equation follows directly from the laws of conditional probability

$$P(B \mid A)P(A) = P(B, A) = P(A, B) = P(A \mid B)P(B) \quad (11.6)$$

Given any system that stores information about variables of interest in terms of conditional probabilities, Bayes' theorem provides a way to invert the known conditional probabilities  $P(A | B)$  to obtain the unknown conditionals  $P(B | A)$  (in our case, the action probabilities conditioned on states). Our choice for a Bayesian approach is motivated by the growing body of evidence from cognitive and psychophysical studies suggesting that the brain utilizes Bayesian principles for inference and decision-making (Knill and Richards, 1996; Rao *et al.*, 2002; Gopnik *et al.*, 2004).

Applying Bayes' theorem to the forward model and the prior probabilities given above, we obtain the inverse model:

$$\begin{aligned} P(a_t = A_i / s_t = S_i, s_{t+1} = S_j, g = S_k) \\ = c P(s_{t+1} = S_j / S_t = S_i, a_t = A_i) P(a_t = A_i / s_t = S_i, g = S_k) \end{aligned} \quad (11.7)$$

where  $c = 1/P(s_{t+1} = S_j / S_t = S_i, g = S_k)$  is the normalization constant that can be computed by marginalizing over the actions:

$$\begin{aligned} P(s_{t+1} = S_j / s_t = g = S_k) \\ = \sum_m P(s_{t+1} = S_j / S_t = S_i, a_t = A_m) \\ \times P(a_t = A_m / s_t = S_i, g = S_k) \end{aligned} \quad (11.8)$$

Thus, at each time step, an action  $A_i$  can either be chosen stochastically according to the probability  $P(a_t = A_i / s_t = S_i, s_{t+1} = S_j, g = S_k)$  or deterministically as the one that maximizes:

$$P(a_t = A_i / s_t = S_i, s_{t+1} = S_j, g = S_k) \quad (11.9)$$

The former action selection strategy is known as probability matching while the latter is known as maximum a posteriori (MAP) selection. In both cases, the probabilities are computed based on the current state, the next sub-goal state and the final goal state using the learned forward model and priors on actions (Eq. 11.7). This contrasts with reinforcement learning methods where goal states are associated with rewards and the algorithms pick actions that maximize the total expected future reward. Learning the 'value function' that estimates the total expected reward for each state typically requires a large number of trials for exploring the state space. In contrast, the imitation-based approach as sketched above utilizes the remembered sequence of sub-goal states to guide the action-selection process, thereby significantly reducing the number of trials needed to achieve the goal state. The actual number of trials depends on the fidelity of the learned forward model, which can be fine-tuned

during body babbling and 'play' with objects as well as during attempts to imitate the teacher.

A final observation is that the probabilistic framework introduced above involving forward and inverse models can also be used to infer the intent of the teacher, i.e. to estimate the probability distribution over the goal state  $g$ , given a sequence of observed states  $s_1, s_2, \dots, s_N$  and a sequence of estimated actions  $a_1, a_2, \dots, a_{N-1}$ .

$$\begin{aligned}
 & P(g = S_k / a_t = A_i, s_t = S_i, s_{t+1} = S_j) \\
 &= k_1 P(s_{t+1} = S_j / s_t = S_i, a_t = A_i, g = S_k) \\
 &\quad \times P(g = S_k / s_t = S_i, a_t = A_i) \\
 &= k_2 P(s_{t+1} = S_j / s_t = S_i, a_t = A_i, g = S_k) \\
 &\quad \times P(a_t = A_i / s_t = S_i, g = S_k) P(g = S_k / s_t = S_i) \\
 &= k_3 P(s_{t+1} = S_j / s_t = S_i, a_t = A_i) P(a_t = A_i / s_t = S_i, g = S_k) \\
 &\quad \times P(s_t = S_i / g = S_k) P(g = S_k)
 \end{aligned} \tag{11.10}$$

where the  $k_i$  are normalization constants. The above equations were obtained by repeatedly applying Bayes' rule. The first probability on the right hand side in Eq. (11.10) is the learned forward model and the second is the learned prior over actions. The last two probabilities capture the frequency of a state given a goal state and the overall probability of the goal state itself. These would need to be learned from experience during interactions with the teacher and the environment. It should be noted that the derivation of Eq. (11.10) above uses the remembered state  $s_t$  of the teacher in lieu of the actual state  $s_t$  (as in Equation 11.7) and is based on the assumption that the teacher's forward model is similar to the imitator's model - such an assumption may sometimes lead to inaccurate inferences, especially if the forward model is not sufficiently well-learned or well-matched with the teacher's, or if the observed state estimate itself is not accurate.

### 11.3.3 Example: learning to solve a maze task through imitation

We illustrate the application of the probabilistic approach sketched above to the problem of navigating to specific goal locations within a maze, a classic problem in the field of reinforcement learning. However, rather than learning through rewards delivered at the goal locations (as in reinforcement learning), we illustrate how an 'agent' can learn to navigate to specific locations by combining in a Bayesian manner a learned internal model with observed trajectories from a teacher (see also Hayes and Demiris, 1994). To make the task more realistic, we assume the presence of noise in the environment

leading to uncertainty in the execution of actions.

### 11.3.3.1 Learning a forward model for the maze task

Figure 11.5(a) depicts the maze environment consisting of a 20 x 20 grid of squares partitioned into several rooms and corridors by walls, which are depicted as thick black lines. The starting location is indicated by an asterisk (\*) and the three possible goal locations (Goals 1, 2 and 3) are indicated by circles of different shades. The goal of the imitator is to observe the teacher's trajectory from the start location to one of the goals and then to select appropriate actions to imitate the teacher.

The states  $s_t$  in this example are the grid locations in the maze. The five actions available to the imitator are shown in Figure 11.5(b): North (N), East (E), South (S), West (W) or remain in place (X). The noisy 'forward dynamics' of the environment for each of these actions is shown in Figure 11.5(c) (left panel). The figure depicts the probability of each possible next state  $s_{t+1}$ , that could result from executing one of the five actions in a given location, assuming that there are no walls surrounding the location. The states  $s_{t+1}$ , are given relative to the current state i.e. N, E, S, W, or X relative to  $s_t$ . The brighter a square, the higher the probability (between 0 and 1), with each row summing to 1. Note that the execution of actions is noisy: when the imitator executes an action, for example  $a_t = E$ , there is a high probability the imitator will move to the grid location to the east ( $s_{t+1} = E$ ) of the current location but there is also a non-zero probability of ending up in the location west ( $s_{t+1} = W$ ) of the current location. The probabilities in Figure 11.5(c) (left panel) were chosen in an arbitrary manner; in a robotic system, these probabilities would be determined by the noise inherent in the hardware of the robot as well as environmental noise. When implementing the model, we assume that the constraints given by the walls are enforced by the environment (i.e. it overrides, when necessary, the states predicted by the forward model in Figure 11.5(c)). One could alternately define a location-dependent, global model of forward dynamics but this would result in inordinately large numbers of states for larger maze environments and would not scale well. For the current purposes, we focus on the locally defined forward model described above that is independent of the agent's current state in the maze.

We examined the ability of the imitator to learn the given forward model through 'body babbling' which in this case amounts to 'maze wandering'. The imitator randomly executes actions and counts the frequencies of outcomes (the next states  $s_{t+1}$ ) for each executed action. The resulting learned forward model, obtained by normalizing the frequency counts to yield probabilities, is

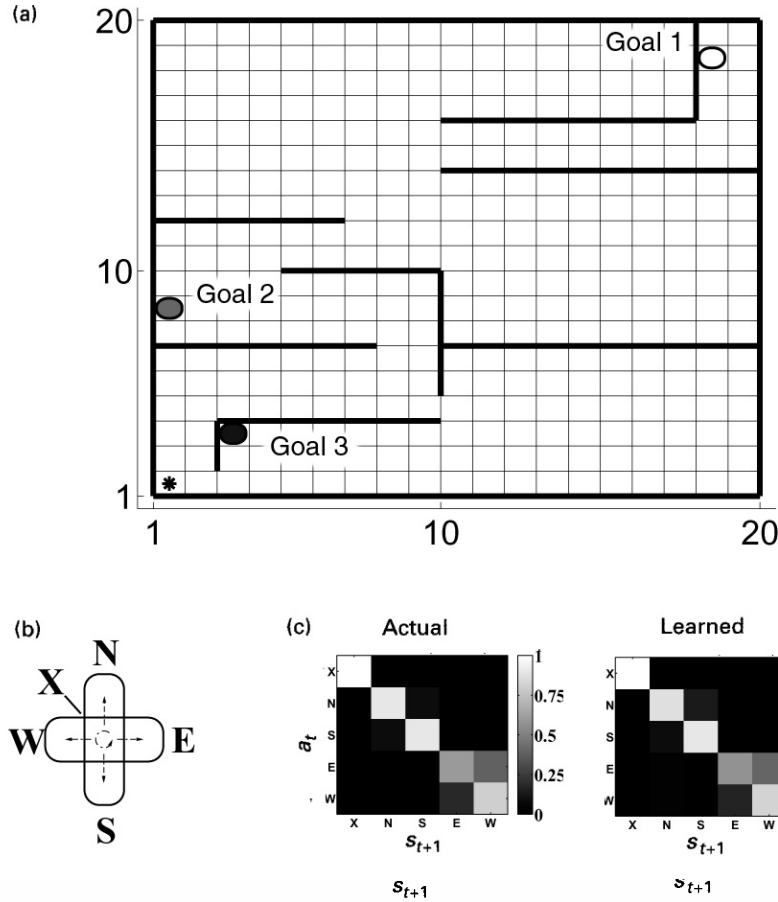


Figure 11.5 Simulated maze environment and learned forward model. (a) Simulated maze environment. Thick lines represent walls. Shaded ovals represent goal states. The instructor and the observer begin each simulated path through the maze at location (1,1), marked by the dark asterisk in the lower left corner of the maze. (b) Five possible actions at a maze location: agents can move north (N), south (S), east (E), west (W), or remain in place (X). (c) Actual and learned probabilistic forward models. The matrix on the left represents the true environmental transition function. The matrix on the right represents an estimated environmental transition function learned through interaction with the environment. Given a current location, each action  $a_t$  (rows) indexes a probability distribution over next states  $s_{t+1}$  (columns). For the states, the labels X, N, S, E, W were used to denote the current location and locations immediately to the north, south, east, and west of the current location respectively. The learned matrix closely approximated the true transition matrix. These matrices assume the agent is not attempting to move through a wall.

shown in Figure 11.5(c) (right panel). By comparing the learned model with the actual forward model, it is clear that the imitator has succeeded in learning the appropriate probabilities  $P(s_{t+1} | s_t, a_t)$  for each value of  $a_t$ , and  $s_{t+1}$  ( $s_t$  is any arbitrary location not adjacent to a wall). The 'body babbling' in this simple case, while not directly comparable to the multi-stage developmental process seen in infants, still serves to illustrate the general concept of learning a forward model through experimentation and interactions with the environment.

### 11.3.3.2 Imitation using the learned forward model and learned priors

Given a learned forward model, the imitator can use Eq. (11.7) to select appropriate actions to imitate the teacher and reach the goal state. The learned prior model  $P(a_t = A_i | s_t = S_i, g = S_k)$ , which is required by Eq. (11.7), can be learned through experience, for example, during earlier attempts to imitate the teacher or during other goal-directed behaviours. The learned prior model provides estimates of how often a particular action is executed at a particular state, given a fixed goal state. For the maze task, this can be achieved by keeping a count of the number of times each action (N, E, S, W, X) is executed at each location, given a fixed goal location.

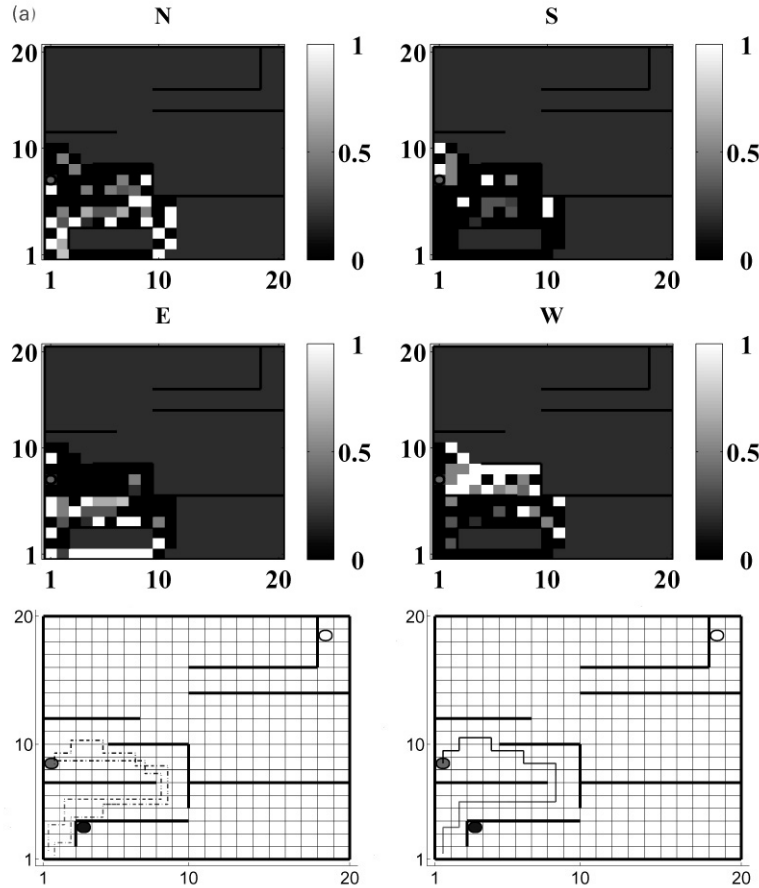
Figure 11.6(a) shows the learned prior model:

$$P(a_t = A_i | s_t = S_i, g = S_k) \quad (11.11)$$

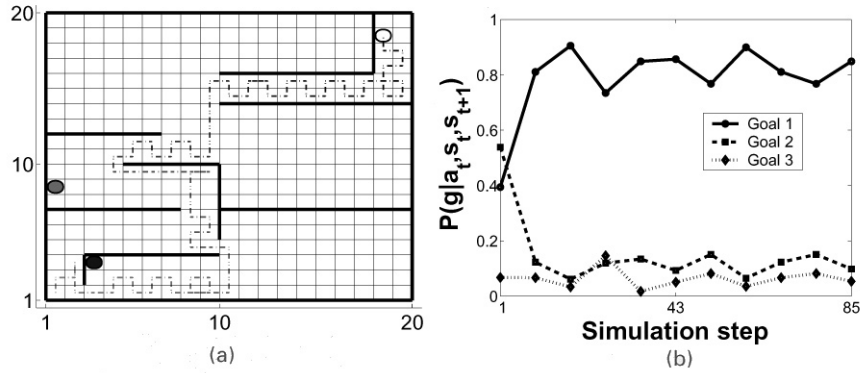
for an arbitrary location  $S_i$  in the maze for four actions  $A_i = N, S, E$ , and  $W$  when the goal state  $g$  is the location (1,8) (Goal 2 in Figure 11.5(a)). The probability for a given action at any maze location (given Goal 2) is encoded by the brightness of the square in that location in the maze-shaped graph for that action in Figure 11.6(a). The probability values across all actions (including X) sum to one for each maze location.

It is clear from Figure 11.6(a) that the learned prior distribution over actions given the goal location points in the correct direction for the maze locations near the explored trajectories. For example, for the maze locations along the bottom-left corridor (from (2,5) to (7,5)), the action with the highest probability is E while for locations along the middle corridor (from (2,8) to (8,8)), the action with the highest probability is W. Similar observations hold for sections of the maze where executing N and S will lead the imitator closer to the given goal location. The priors for unexplored regions of the maze were set to uniform distributions for these simulations (dark regions in Figure 11.6(a)).

The learned forward model in Figure 11.5(c) can be combined with the learned prior model in Figure 11.6(a) to obtain a posterior distribution over actions as specified by Eq. (11.7). Figure 11.6(c) shows an example of the trajectory



**Figure 11.6** Learned priors and example of successful imitation: (a) learned prior distributions  $P(a_t | s_b, s_G)$  for the four directional actions (north, south, east, and west) for Goal 2 (map location (1,s)) in our simulated maze environment. Each location in the maze indexes a distribution over actions (the brighter the square, the higher the probability), so that the values across all actions (including X - not shown) sum to one for each maze location. (b) Trajectories (dashed lines) demonstrated by the instructor during training. The goal location here is Goal 2 depicted by the grey circle at map location (1,8). Trajectories are offset within each map cell for clarity; in actuality, the observer perceives the map cell occupied by the instructor at each time step in the trajectory. So, for example, both trajectories start at map cell (1, 1). Time is encoded using grayscale values, from light grey (early in each trajectory) to black (late in each trajectory). (c) Example of successful imitation. The observer's trajectory during imitation is shown as a solid line with grayscale values as in (b). Imitation is performed by combining the learned forward and prior models, as described in the text, to select an action at each step.



**Figure 11.7** Inferring the intent of the teacher. (a) Dashed line plots a testing trajectory for intent inference. Greyscale values show the progression of time, from light grey (early in the trajectory) to black (late in the trajectory). The intended goal of the instructor was Goal 1 (the white circle at the top right). (b) Inferred intent, shown as a distribution over goal states. Each point in the graph represents the output of the intent inference algorithm, averaged over eight individual simulation steps (the final data point is an *average* over five simulation steps). Note that the instructor's desired goal, Goal 1, is correctly inferred as the objective for all points on the graph except the first. Potential ambiguities at different locations are not obvious in this graph due to averaging and unequal priors for the three goals (see text for details).

followed by the imitator after observing the two teacher trajectories shown in Figure 11.6(b). Due to the noisy forward model as well as limited training data, the imitator needs more steps to reach the goal than does the instructor on either of the training trajectories for this goal, typically involving backtracking over a previous step or remaining in place. Nevertheless, it eventually achieves the goal location as can be seen in Figure 11.6(c).

### 11.3.3.3 Inferring the intent of the teacher

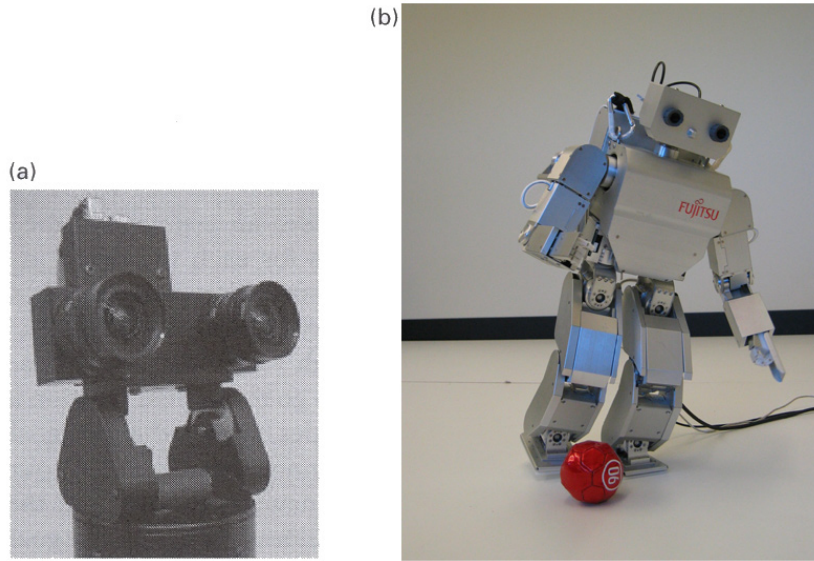
After training on a set of trajectories for each goal (one to two trajectories per goal for the simple example illustrated here), the imitator can attempt to infer the intent of the teacher based on observing some or all of the teacher's actions. Figure 11.7(a) depicts an example trajectory of the teacher navigating to the goal location in the top right corner of the maze (Goal 1 in Figure 11.5(a)). Based on this observed trajectory of 85 total steps, the task of the imitator in this simple maze environment is to infer the probability distribution over the three possible goal states given the current state, the next state and the action executed

at the current state. The trajectory in Figure 11.7(a) was not used to train the observer; instead, this out-of-sample trajectory was used to test the intent inference algorithm described in the text. Note that the desired goal with respect to the prior distributions learned during training is ambiguous at many of the states in this trajectory.

The intent inference algorithm provides an estimate of the distribution over the instructor's possible goals for each time step in the testing trajectory. The evolution of this distribution over time is shown in Figure 11.7(b) for the teacher trajectory in (a). Note that the imitator in this case converges to a relatively high value for Goal 1, leading to a high certainty that the teacher intends to go to the goal location in the top right corner. Note also that the probabilities for the other two goals remain non-zero, suggesting that the imitator cannot completely rule out the possibility that the teacher may in fact be navigating to one of these other goal locations. In this graph, the probabilities for these other goals are not very high even at potentially ambiguous locations (such as location (9,9)) because: (1) the plotted points represent averages over five simulation steps and (2) Eq. (11.10) depends on  $P(g = S_k)$ , the prior probabilities of goals, which in this case involved higher values for Goal 1 compared to the other goals. Other choices for the prior distribution of goals (such as a uniform distribution) can be expected to lead to higher degrees of ambiguity about the intended goal at different locations. The ability of the imitator to estimate an entire probability distribution over goal states allows it to ascribe degrees of confidence to its inference of the teacher's intent, thereby allowing richer modes of interaction with the teacher than would be possible with purely deterministic methods for inferring intent (see Verma and Rao, 2006).

#### 11.3.3.4 Summary

Although the maze task above is decidedly simplistic, it serves as a useful first example in understanding how the abstract probabilistic framework proposed in this chapter can be used to solve a concrete sensorimotor problem. In addition, the maze can be regarded as a simple 2-D example of the general sensorimotor task of selecting actions that will take an agent from an initial state to a desired goal state, where the states are typically high-dimensional variables encoding configurations of the body or a physical object rather than a 2-D maze location. However, because the states are assumed to be completely observable, the maze example by itself does not provide an explanation for the results obtained by Meltzoff (1995) showing that an infant is able to infer intent from an unsuccessful human demonstration but not an unsuccessful mechanical demonstration.



**Figure 11.8** Robotic platforms for testing Bayesian imitation models. (a) A binocular pan-tilt camera platform ('Biclops') from Metrica, Inc. (b) A miniature humanoid robot (HOAP-2) from Fujitsu Automation, Japan. Both robotic platforms are currently being used to test the Bayesian framework sketched in this chapter.

(Figure 11.4). An explanation for such a phenomenon would require the correspondence problem to be addressed as well as a sophisticated forward model of finger and hand movements, topics that we intend to address in future modelling studies.

#### 11.3.4 Further applications in robotic learning

We are currently investigating the applicability of the probabilistic framework described above to the problem of programming robots through demonstration of actions by human teachers (Demiris *et al.*, 1997; Berthouze and Kuniyoshi, 1998; Mataric and Pomplun, 1998; Schaal, 1999; Billard and Dautenhahn, 1999; Breazeal and Scassellati, 2002; Dautenhahn and Nehaniv, 2002). Two robotic platforms are being used: a binocular robotic head from Metrica, Inc. (Fig 11.8(a)), and a recently acquired Fujitsu HOAP-2 humanoid robot (Fig 11.8(b)).

In the case of the robotic head, we have investigated the use of 'oculomotor babbling' (random camera movements) to learn the forward model probabilities  $P(s_{t+1} = S_j / s_t = S_i, a_t = A_i)$ . The states  $S_i$  in this case are the feedback from the motors ('proprioception') and visual information (for example, positions of object features). The learned forward model for the robotic head can be used in the

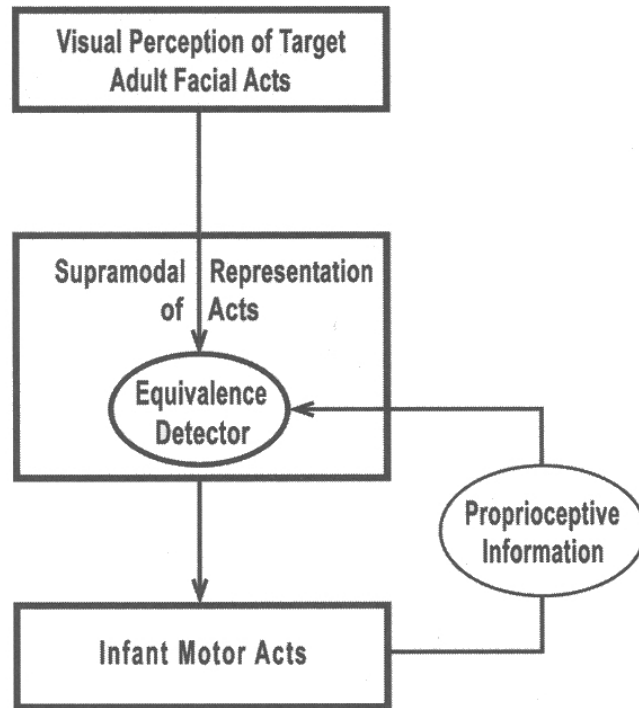
manner described in Section 11.3.2 to solve head movement imitation tasks (Demiris et al., 1997). In particular, we intend to study the task of robotic gaze following. Gaze following is an important component of language acquisition: to learn words, a first step is to determine what the speaker is looking at, a problem solved by the human infant by about one year of age (Brooks and Meltzoff, 2002, 2005). We hope to design robots with a similar capability (see Hoffman et al., 2006, for progress).

Other work will focus on more complex imitation tasks using the HOAP-2 humanoid robot, which has 25 degrees of freedom, including articulated limbs, hands and a binocular head (Fig 11.8 (b)). Using the humanoid, we expect to be able to rigorously test the strengths and weaknesses of our probabilistic models in the context of a battery of tasks modeled after the progressive stages in imitative abilities seen in infants (see Section 11.2). Preliminary results can be found in Grimes et al., 2006.

### *11.3.5 Towards a probabilistic model for imitation in infants*

The probabilistic framework sketched above can also be applied to better understand the stages of infant imitation learning described by Meltzoff and Moore. For example, in the case of facial imitation, the states could encode proprioceptive information resulting from facial actions such as tongue protrusion or at a more abstract level, 'supramodal' information about facial acts that is not modality-specific (visual, tactile, motor, etc.). Observed facial acts would then be transformed to goal states through a correspondence function, which has been hypothesized to be innate (Meltzoff, 1999). Such an approach is consistent with the proposal of Meltzoff and Moore that early facial imitation is based on active intermodal mapping (AIM) (Meltzoff and Moore, 1977, 1994, 1997). Figure 11.9 provides a conceptual schematic of the AIM hypothesis. The key claim is that imitation is a matching-to-target process. The active nature of the matching process is captured by the proprioceptive feedback loop. The loop allows infants' motor performance to be evaluated against the seen target and serves as a basis for correction. One implementation of such a match-and-correction process is the Bayesian action selection method described above with both visual and proprioceptive information being converted to supramodal states. The selection of actions according to Eq. (11.7) followed by subsequent matching of remembered and actual states could implement the closed-loop matching process in the AIM model.

As a second example of the application of the probabilistic framework, consider imitation learning of actions of objects. In this case, the states to be



**Figure 11.9** Meltzoff and Moore's AIM model of facial imitation (from Meltzoff and Moore, 1997).

encoded are the states of the object ('joined together', 'pulled apart', etc. for the dumbbell-shaped object mentioned above). The forward model to be used would presumably be one that has been learned from experience with similar objects ('objects that can be pulled apart'). This, along with the learned priors for various actions, would allow appropriate actions to be selected based on the observed sequence of object states.

Finally, consider the case where an infant learns from unsuccessful demonstrations by inferring the intention of a human demonstrator. In this case, forward models could be put to good use to infer intention. By using a forward model of a human manipulating an object, the consequences of attempted actions by the human demonstrator can be predicted. For example, in the case of the dumbbell-shaped object used by Meltzoff (1995), the forward model, presumably learned through prior experience with pulling apart objects, would predict that when a person is applying forces at the two ends in opposite directions (away

from the centre), there is a high probability for the state where the object has been pulled apart into two halves. This state could in turn be adopted as the desired goal state and the appropriate action that maximizes the probability of achieving this state could be selected in the Bayesian manner described above. On the other hand, in the case of an inanimate device demonstrating a failed action (Figure 11.4), the infant presumably does not apply its forward model to predict the consequences of the attempted action, perhaps because the inanimate device is not deemed to be human enough to use the infant's forward model (cf. the 'like me' hypothesis of Meltzoff, 2006, 2007). The important question of what type of forward model is chosen for predicting outcomes and intentions under specific circumstances remains a topic for further study.

The Bayesian model sketched in this chapter is only a first step towards more sophisticated probabilistic models of imitation. For example, it does not explain why human infants are superior in their imitation abilities when compared to apes and other animals that could potentially possess many of the same components of the Bayesian model as an infant might possess. Addressing such distinctions between imitative abilities will likely require a more in-depth study of the types of internal models used by infants and animals, and an examination of their ability to generalize to novel circumstances when using these internal models for imitation.

### **11.4 Prospects for developmental robotics**

Humans at birth do not have the full set of skills and behaviours exhibited by adults. Human beings are not 'turn key' systems that function perfectly out of the box. There are at least four sources of behavioural change in human development: (1) maturational changes in the sensory, motor and cognitive system, (2) reinforcement learning, (3) independent invention and discovery, often called 'insight' and (4) imitative learning. The first three have been widely celebrated: maturation has been studied by neuroscientists; reinforcement learning by Skinner (1953) and generations of learning theorists; and independent invention and solitary discovery by Piaget (1954) and others. The imitative competence of young infants has only recently been discovered, and its enormous impact on human development and learning only recently sketched (e.g. see Meltzoff, 2005, 2006, 2007).

Imitative learning is more flexible and responsive to cultural norms than maturation; it is safer for the child than Skinnerian trial-and-error learning; and it is faster than relying on Piagetian solitary discoveries. These advantages of imitation learning apply equally well to robots and other autonomous agents. In particular, learning through imitation offers substantial benefits over

other leading robotic learning methods (such as reinforcement learning) by: (1) overcoming the need for a huge number of learning trials and (2) avoiding the need for risky and dangerous experimentation during learning.

At the same time, unlike supervised learning methods, imitative learning does not require a human to provide the exact motor signals needed to accomplish each task - the robot deduces these based only on observing a human or robotic demonstrator.

In this chapter, we discussed some of the main results obtained from studies of imitative learning in infants. These results suggest a four-stage progression of imitative learning abilities: (1) body babbling, (2) imitation of body movements, (3) imitation of actions on objects and (4) imitation based on inferring intentions of others. We took a first step towards formalizing these stages within a probabilistic framework inspired by recent ideas from machine learning, and provided an example demonstrating the application of the framework to the imitation learning problem.

The probabilistic approach is well-suited to imitation learning in real-world robotic environments which are noisy and uncertain. The success of recent approaches to robotic navigation and control can be attributed to the use of probabilistic techniques such as Kalman filtering and particle filtering for handling uncertainty (Blake and Yuille, 1992; Fox *et al.*, 2001). Similarly, techniques based on statistical learning form the back-bone of several recent successful computer vision systems for tracking and recognizing persons (for example, see Jojic and Frey, 2001). We are optimistic that a probabilistic approach to robotic imitation learning will offer many of the advantages of these preceding systems, including the ability to handle missing data, robustness to noise, ability to make predictions based on learned models, etc. We are currently testing our ideas on a binocular robotic head and a humanoid robot. An important issue we intend to explore is the scalability of the proposed approach. The Bayesian model requires both a forward model as well as a prior model to be learned. In the case of the maze example, the forward and prior models were learned using a relatively small number of trials due to the small size of the state space. A more realistic scenario involving, for example, a humanoid robot would presumably require a larger number of trials due to the greater degrees of freedom available; however, the problem may be alleviated to some extent by learning forward and prior models only for the parts of the state space that are selectively utilized in imitative sessions (e.g., Grimes *et al.*, 2006). Hierarchical state-space models may also help in this regard. We hope to explore the issue of scalability in future studies involving the humanoid robot.

The probabilistic approach also opens up the possibility of applying Bayesian methodologies such as manipulation of prior probabilities of task alternatives to obtain a deeper understanding of imitation in humans. For example, one could explore the effects of biasing a human subject towards particular classes

of actions (e.g. through repetition) under particular sets of conditions, a manipulation that would in effect test the contributions of the prior model in Eq. (11.7). One could also manipulate the forward model used by subjects with the help of virtual reality environments. Such manipulations have yielded valuable information regarding the type of priors and internal models that the adult human brain uses in perception (see chapters in Rao *et al.*, 2002) and in motor learning (Wolpert *et al.*, 1995).

We believe that the application of such methodology to imitation could shed new light on the problem of how infants acquire internal models of the people and things they encounter in the world. Conversely, we believe that biologically inspired models will help shape the architecture and algorithms used to solve imitative learning problems in robots (see Demiris *et al.*, 1997; Schaal, 1999; Demiris and Hayes, 2002). For example, Meltzoff and Moore's four stages of imitation in infants suggests a hierarchical approach to robotic imitation, starting from learning internal models of self motion to more sophisticated models of interactions with active behaving agents.

Imitation is an especially fruitful domain for interdisciplinary collaboration between robotics and developmental science. It is a perceptual-motor activity of great adaptive value and a channel for learning that lends itself to computational modelling. Additionally, it presents an interesting challenge to robotics and offers an extremely versatile and flexible way to program robots. Such interdisciplinary collaborations could eventually allow us to investigate interactions between young children and robots. Do young children, prior to developing the prejudices and philosophies of adults, think that robots have subjective beliefs, desires, emotions and intentions? Experiments addressing such questions will not only provide new insights into how humans develop a 'theory of other minds', but at the same time, will allow us to use unprejudiced humans (children) as judges of robots in a new form of the celebrated 'Turing test' for autonomous machines. We therefore look forward to rich bi-directional benefits emerging from collaborations between developmental science and robotics in the coming years.

### Acknowledgments

We would like to thank the reviewers for their constructive comments and suggestions. This work is supported by grants to RPNR from ONR (grant no. N00014-03-1-0457), NSF (AICS-0413335) and the Packard Foundation, and grants to ANM from NIH (HD-225 14) and NSF (SBE-0354453).

## REFERENCES

- Alissandrakis, A., Nehaniv, C. L. and Dautenhahn, K. (2002). Do as I do: correspondences across different robotic embodiments. In D. Polani, J. Kim, and T. Martinetz (eds.), *Proceedings of the 5th German Workshop on Artificial Life*, 143-52.
- Alissandrakis, A., Nehaniv, C. L. and Dautenhahn, K. (2002a). Imitation with ALICE: learning to imitate corresponding actions across dissimilar embodiments, *IEEE Transactions on Systems, Man, and Cybernetics*, Part A, **32**(4), 482-96.
- Berthouze, L. and Kuniyoshi, Y. (1998). Emergence and categorization of coordinated visual behaviour through embodied interaction. *Machine Learning*, **31**, 187-200.
- Billard, A. and Dautenhahn, K. (1999). Experiments in learning by imitation - grounding and use of communication in robotic agents. *Adaptive Behaviour*, **7**, 415-38.
- Blake, A. and Yuille, A. (eds.) (1992). *Active Vision*. Cambridge, MA: MIT Press.
- Breazeal, C. and Scassellati, B. (2002). Challenges in building robots that imitate people. In K. Dautenhahn and C. L. Nehaniv (eds.), *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press, 363-90.
- Brooks, R. and Meltzoff, A. N. (2002). The importance of eyes: how infants interpret adult looking behavior. *Developmental Psychology*, **38**, 958-66.
- Brooks, R. and Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, **8**, 535-43.
- Bryson, A. E. and Ho, Y. C. (1975). *Applied Optimal Control*. Bristol, PA: Taylor and Francis.
- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G. and Freund, H.J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European Journal of Neuroscience*, **13**, 400-4.
- Dautenhahn, K. and Nehaniv, C. L. (2002). *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press.
- Decety, J. (2002). Is there such a thing as functional equivalence between imagined, observed, and executed action? In A. N. Meltzoff and W. Prinz (eds.), *The Imitative Mind: Development, Evolution, and Brain Bases*. Cambridge, UK: Cambridge University Press, 291-310.
- Demiris, J. and Hayes, G. (2002). Imitation as a dual-route process featuring predictive and learning components: A biologically plausible computational model. In K. Dautenhahn and C. Nehaniv (eds.), *Imitation in animals and artifacts*. Cambridge, MA: MIT Press, 327-61.
- Demiris, Y. and Meltzoff, A. N. (in press). The robot in the crib: A developmental analysis of imitation skills in infants and robots. *Infant and Child Development*.
- Demiris, J., Rougeaux, S., Hayes, G. M., Berthouze, L. and Kuniyoshi, Y. (1997). Deferred imitation of human head movements by an active stereo vision head. *Proceedings of the 6th IEEE International Workshop on Robot and Human Communication (Sept 29-Oct 1, Sendai, Japan)*, 88-93.

- Fox, D., Thrun, S., Dellaert, F. and Burgard, W. (2001). Particle filters for mobile robot localization. In A. Doucet, N. de Freitas and N. Gordon (eds.), *Sequential Monte Carlo Methods in Practice*. New York: Springer Verlag, 401-28.
- Glymour, C. (2001). *The Mind's Arrows: Bayes Nets and Graphical Causal Models in Psychology*. Cambridge, MA: MIT Press.
- Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T. and Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological Review*, **111**, 1-30.
- Grimes, D. B., Chalodhorn, R. and Rao, R. P. N. (2006). Dynamic imitation in a humanoid robot through nonparametric probabilistic inference. *Proceedings of the 2006 Robotics Science and Systems Conference* (Aug. 16-19, University of Pennsylvania, Philadelphia, PA).
- Gross, C. G. (1992). Representation of visual stimuli in inferior temporal cortex. In V. Bruce, A. Cowey and A. W. Ellis (eds.), *Processing the Facial Image*. New York: Oxford University Press, 3-10.
- Hanna, E. and Meltzoff, A. N. (1993). Peer imitation by toddlers in laboratory, home, and day-care contexts: implications for social learning and memory. *Developmental Psychology*, **29**, 701-10.
- Hayes, G. and Demiris, J. (1994). A robot controller using learning by imitation. *Proceedings of the 2nd International Symposium on Intelligent Robotic Systems, Grenoble, France*, 198-204.
- Hoffman, M., Grimes, D. B., Shon, A. P. and Rao, R. P. N. (2006). A probabilistic model of gaze imitation and shared attention. *Neural Networks*, **19**, 299-310.
- Jackson, P. L., Meltzoff, A. N. and Decety, J. (2006). Neural circuits involved in imitation and perspective-taking. *NeuroImage*, **31**, 429-39.
- Jensen, F. V. (2001). *Bayesian Networks and Decision Graphs*. New York: Springer-Verlag.
- Jojic, N. and Frey, B. (2001). Learning flexible sprites in video layers. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Jordan, M. I. and Rumelhart, D. E. (1992). Forward models: supervised learning with a distal teacher. *Cognitive Science*, **16**, 307-54.
- Knill, D. C. and Richards, W. (eds.) (1996). *Perception as Bayesian Inference*. Cambridge, UK: Cambridge University Press.
- Kuhl, P. K. and Meltzoff, A. N. (1996). Infant vocalizations in response to speech: vocal imitation and developmental change. *Journal of the Acoustical Society of America*, **100**, 2425-38.
- Mataric, M. J. and Pomplun, M. (1998). Fixation behaviour in observation and imitation of human movement. *Cognitive Brain Research*, **7**, 191-202.
- Meltzoff, A. N. (1988a). Imitation of televised models by infants. *Child Development*, **59**, 1221-9.
- Meltzoff, A. N. (1988b). Infant imitation after a 1-week delay: long-term memory for novel acts and multiple stimuli. *Developmental Psychology*, **24**, 470-6.

- Meltzoff, A. N. (1995). Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, **31**, 838-50.
- Meltzoff, A. N. (1996). The human infant as imitative generalist: a 20-year progress report on infant imitation with implications for comparative psychology. In C. M. Heyes and B. G. Galef (eds.), *Social Learning in Animals: The Roots of Culture*. New York: Academic Press, 347-70.
- Meltzoff, A. N. (1999). Origins of theory of mind, cognition, and communication. *Journal of Communication Disorders*, **32**, 251-69.
- Meltzoff, A. N. (2005). Imitation and other minds: the 'Like Me' hypothesis. In S. Hurley and N. Chater (eds.), *Perspectives on Imitation: From Cognitive Neuroscience to Social Science*: Vol. 2. Cambridge: MIT Press, 55-77.
- Meltzoff, A. N. (2006). The 'like me' framework for recognizing and becoming an intentional agent. *Acta Psychologica*, doi:10.1016/j.actpsy.2006.09.005.
- Meltzoff, A. N. (2007). 'Like me': a foundation for social cognition. *Developmental Science*, **10**, 126-34.
- Meltzoff, A. N. and Decety, J. (2003). What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society B*, **358**, 491-500.
- Meltzoff, A. N. and Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, **198**, 75-8.
- Meltzoff, A. N. and Moore, M. K. (1983). Newborn infants imitate adult facial gestures. *Child Development*, **54**, 702-9.
- Meltzoff, A. N. and Moore, M. K. (1989). Imitation in newborn infants: exploring the range of gestures imitated and the underlying mechanisms. *Developmental psychology*, **25**, 954-62.
- Meltzoff, A. N. and Moore, M. K. (1994). Imitation, memory, and the representation of persons. *Infant Behavior and Development*, **17**, 83-99.
- Meltzoff, A. N. and Moore, M. K. (1997). Explaining facial imitation: a theoretical model. *Early Development and Parenting*, **6**, 179-92.
- Nehaniv, C. L. and Dautenhahn, K. (1998). Mapping between dissimilar bodies: affordances and the algebraic foundations of imitation. In J. Demiris and A. Birk (eds.), *Proceedings of the European Workshop on Learning Robots*.
- Nehaniv, C. L. and Dautenhahn, K. (2001). Like me? - measures of correspondence and imitation. *Cybernetics and systems*, **32(1-2)**, 11-51.
- Nehaniv, C. L. and Dautenhahn, K. (2002). The correspondence problem. In K. Dautenhahn and C. L. Nehaniv (eds.), *Imitation in Animals and artifacts*. Cambridge, MA: MIT Press, 41-61
- Patrick, J., Campbell, K., Carmichael, L., Natale, R. and Richardson, B. (1982). Patterns of gross fetal body movement over 24-hour observation intervals during the last 10 weeks of pregnancy. *American Journal of Obstetrics and Gynecology*, **142**, 363-71.
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. San Francisco, CA: Morgan Kaufman.
- Piaget, J. (1954). *The Construction of Reality in the Child*. New York: Basic Books.

- Prinz, W. (2002). Experimental approaches to imitation. In A. N. Meltzoff and W. Prinz (eds.), *The Imitative Mind: Development, Evolution, and Brain Bases*. Cambridge, UK: Cambridge University Press, 143-62.
- Rabiner, L. R. and Juang, B. H. (1986). An introduction to hidden Markov models. *IEEE ASSP Magazine*, **3**(1), 4-16.
- Rao, R. P. N. and Meltzoff, A. N. (2003). Imitation learning in infants and robots: towards probabilistic computational models. *Proceedings of the 2003 Artificial Intelligence and Simulation of Behaviour (AISB) Convention: Cognition in Machines and Animals, UK.*
- Rao, R. P. N., Olshausen, B. A. and Lewicki, M. S. (eds.) (2002). *Probabilistic Models of the Brain: Perception and Neural Function*. Cambridge, MA: MIT Press.
- Rizzolatti, G., Fadiga, L., Fogassi, L. and Gallese, V. (2002). From mirror neurons to imitation, facts, and speculations. In A. N. Meltzoff and W. Prinz (eds.), *The Imitative Mind: Development, Evolution and Brain Bases*. Cambridge, UK: Cambridge University Press, 247-66.
- Robotics and Autonomous Systems* (2004). Special Issue: Robot Learning from Demonstration, **47**(2-3).
- Schaal, S. (1999). Is imitation learning the route to humanoid robots? *Trends in Cognitive Sciences*, **3**, 233-42.
- Skinner, B. F. (1953). *Science and Human Behaviour*. New York: Macmillan
- Tomasello, M. and Call, J. (1997). *Primate Cognition*. New York: Oxford University Press.
- Verma, D. and Rao, R. P. N. (2006). Goal-based imitation as a probabilistic inference over graphical models. In Y. Weiss, B. Schölkopf and J. Platt (eds.), *Advances in Neural Information Processing Systems 18*. Cambridge, MA: MIT Press, 1393-400.
- Visalberghi, E. and Frigaszy, D. (2002). 'Do monkeys ape?' Ten years after. In K. Dautenhahn and C. L. Nehaniv (eds.), *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press, 471-99.
- Whiten, A. (2002). The imitator's representation of the imitated: ape and child. In A. N. Meltzoff and W. Prinz (eds.), *The Imitative Mind: Development, Evolution, and Brain Bases*. Cambridge, UK: Cambridge University Press, 98-121.
- Wolpert, D. M., Ghahramani, Z. and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, **269**, 1880-2.
- Wolpert, D. M. and Kawato, M. (1998). Multiple paired forward and inverse models for motor control, *Neural Networks* **11**(7-8), 1317-29.