in *"From Motor Learning to Interaction Learning in Robots"*
Springer Verlag, Series: Studies in Computational Intelligence ,
Vol. 264, Sigaud, Olivier; Peters, Jan (Eds.), 2010.

Chapter 1
Abstraction Levels for Robotic Imitation:
Overview and Computational Approaches

Manuel Lopes, Francisco Melo, Luis Montesano and José Santos-Victor

Abstract This chapter reviews several approaches to the problem of learning by imitation in robotics. We start by describing several cognitive processes identified in the literature as necessary for imitation. We then proceed by surveying different approaches to this problem, placing particular emphasys on methods whereby an agent first learns about its own body dynamics by means of self-exploration and then uses this knowledge about its own body to recognize the actions being performed by other agents. This general approach is related to the motor theory of perception, particularly to the mirror neurons found in primates. We distinguish three fundamental classes of methods, corresponding to three abstraction levels at which imitation can be addressed. As such, the methods surveyed herein exhibit behaviors that range from raw sensory-motor trajectory matching to high-level abstract task replication. We also discuss the impact that knowledge about the world and/or the demonstrator can have on the particular behaviors exhibited.

Manuel Lopes
Instituto de Sistemas e Robótica, Instituto Superior Técnico
Lisbon, Portugal e-mail: macl@isr.ist.utl.pt

Francisco Melo
Carnegie Mellon University
Pittsburgh, PA, USA e-mail: fmelo@cs.cmu.edu

Luis Montesano
Universidad de Zaragoza
Zaragoza, Spain e-mail: lmontesa@unizar.es

José Santos-Victor
Instituto de Sistemas e Robótica, Instituto Superior Técnico
Lisbon, Portugal e-mail: jasv@isr.ist.utl.pt

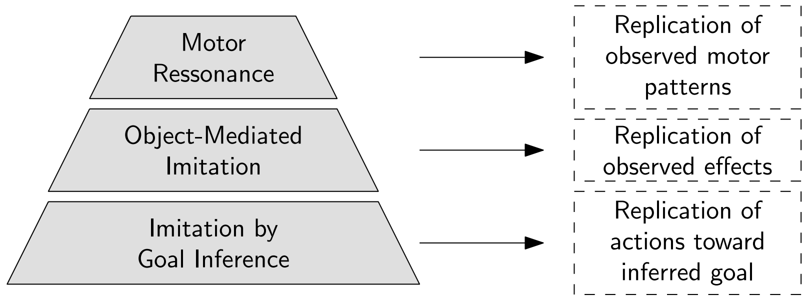


Fig. 1.1: Approaches to imitation at the three levels of abstraction discussed in this chapter.

1.1 Introduction

In this chapter we study several approaches to the problem of imitation in robots. This type of skill transfer is only possible if the robots have several cognitive capabilities that, in turn, pose multiple challenges in terms of modeling, perception, estimation and generalization. Throughout the chapter, we survey several methods that allow robots to learn from a demonstration. Several other surveys cover different aspects of imitation, including [6, 11, 16, 128].

Rather than providing another extensive survey of learning from demonstration, in this chapter we review some recent developments in imitation in biological systems and focus on robotics works that consider self-modeling as a fundamental part of the cognitive processes involved in and required for imitation. Self-modeling, in this context, refers to the learning processes that allow the robot to understand its own body and its interaction with the environment.

In this survey, we distinguish three fundamental classes of methods, each addressing the problem of learning by imitation at different levels of abstraction. Each of these levels of abstraction focuses on a particular aspect of the demonstration, giving rise to different imitative behaviors ranging from motor resonance to a more abstract imitation of inferred goals. This hierarchy of behaviors is summarized in the diagram of Fig. 1.1. It is interesting to note that the approaches at these different levels of abstraction, rather than being mutually exclusive, actually provide a natural hierarchical decomposition, in which approaches at the more abstracted levels can build on the outcome of methods in less abstract levels (see, for example, [80, 84] for an example of such integration).

Why Learn by Imitation?

The impressive research advances in robotics and autonomous systems in the past years have led to the development of robotic platforms of increasingly complex motor, perceptual and cognitive capabilities. These achievements open the way for new applications that require these systems to interact with other robots and/or human users during extended periods of time. Traditional programming methodologies and robot interfaces will no longer suffice, as these systems need to *learn* to execute new complex tasks and improve their performance throughout its lifetime.

Learning by imitation is likely to become one primary form of teaching such complex robots [9, 127]. Paralleling the ability of human infants to learn through (extensive) imitation, an artificial system can retrieve a large amount of task related information simply by observing other individuals, humans or robots, perform that same task. Such a system would ideally be able to observe humans and learn how to solve similar tasks by imitation only. To be able to achieve such capability there are several other skills that must be developed first [84].

The ability to imitate has also been used in combination with other learning mechanisms. For instance, it can speed up learning either by providing an initial solution for the intended task that can then be improved by trial-and-error [109] or by guiding exploration [112, 114]. It also provides more intuitive and acceptable human-machine interactions due to its inherent social component [20,79]. Learning by imitation has been applied before the advent of humanoid robots and in several different applications, including robotics [75], teleoperation [153], assembly tasks [149], game characters [139], multiagent systems [113], computer programming [49] and others.

What Is Imitation?

In biological literature, many behaviors have been identified under the general label of "social learning". Two such social learning mechanisms have raised particular interest among the research community, these being *imitation* and *emulation* [148]. In both the agent tries to replicate the effects achieved by the demonstrator but in imitation the agent also replicates the motor behavior used to achieve such goal, while in emulation only the effects are replicated (the agent achieves the effect by its own means).

In robotic research the word *imitation* is also used to represent many different behaviors and methodologies. Some works seek to clarify and distinguish several such approaches, either from a purely computational point-of-view [84,89,104,127] or taking inspiration in the biological counterparts [25, 79, 140, 143, 146, 148, 154]. The taxonomy depicted in Fig. 1.2 provides one possible classification of different social learning mechanisms that takes into account three sources of information, namely *goals*, *actions* and *effects*.

In this paper, we define imitation in its daily use meaning and use the designations *imitation* and *learning/programming by demonstration* interchangeably. Taking into account the previous taxonomy the works presented may be classified under other labels. Roughly speaking we can consider the three levels as going from mimicking, through (goal) emulation and finally imitation. This division is clear if we consider methods that make an explicit inference about the goal as imitation, but

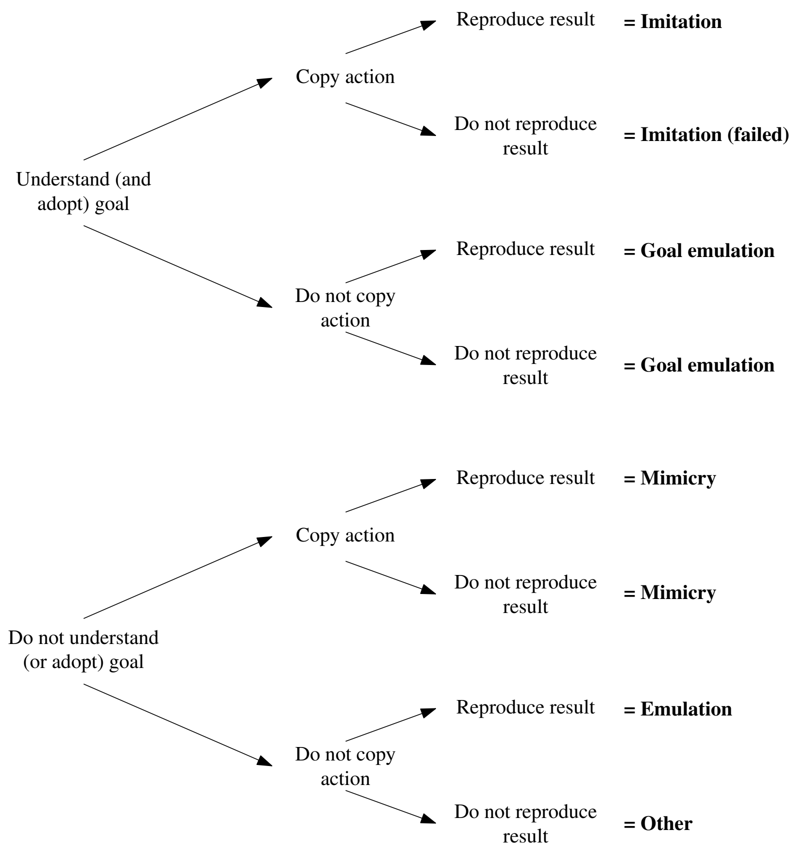


Fig. 1.2: Behavior classification in terms of goals, actions and effects (reproduced from [25]).

not that clear in the cases where the trajectory generalization is performed using an implicit goal inference.

Organization of the Chapter

In the continuation, and before entering into the different computational approaches to imitation, Section 1.2 briefly outlines relevant aspects from psychology and neurophysiology on the topic of imitation in biological systems. Section 1.3 then discusses imitation in artificial systems, by pointing out the main scientific challenges that have been identified and addressed in the literature on imitation learning.

The remainder of the chapter is divided into three main sections, each addressing imitation from a specific perspective:

- Section 1.4 addresses imitation from a motor resonance perspective, namely *trajectory matching and generalization*. It discusses approaches that work at the trajectory level (either joint or task space). These approaches can be interpreted as performing regression (at the trajectory level) using the observed demonstration, and including additional steps to allow the learner to generalize from it.
- Section 1.5 discusses imitation by *replication of observed* (world) *events*. In this section, the learner focuses on replicating observed effects in the world, mainly effects on objects.
- *Goal inference* is finally presented in Section 1.6. We survey approaches in which the learner explicitly tries to infer the goal of the demonstrator and then uses this goal to guide its action-choice.

We note that such division is not strict and some of the approaches share ideas across several of the perspectives above. Also, depending on the application and context, one particular perspective might be more appropriate for imitation than the others. We conclude the paper in Sections 1.7 and 1.8 by discussing other approaches to imitation and providing some concluding remarks.

1.2 Imitation in Natural Systems

The idea of learning by imitation has a clear inspiration in the way humans and other animals learn. Therefore, results from neurophysiology and psychology on imitation in humans, chimpanzees and other primates are a valuable source of information to better understand, develop and implement artificial systems able to learn and imitate. Section 1.2.1 details information from neurophysiology and Section 1.2.2 presents evidence from psychology and biology. Such results illustrate the highly complex task that imitation is. The brief literature review in this section identifies some of the the problems that must be addressed before robots can learn (efficiently) by imitation and some works in the robotic literature that seek to model/test cognitive hypothesis.

1.2.1 Neurophysiology

Neurophysiology identified several processes involved in action understanding that, in turn, contributed differently to the development of learning approaches in robotics. For example, mirror neurons [51, 106] provided an significant motivation for using motor simulation theories in robotics. Similar ideas were suggested in speech recognition [50,77]. Also the existence of forward and backward models in the cerebellum gave further evidence that the production system is involved in the perception, although it is not clear if it is necessary [152]. Most of these methods consider already known actions and not the way such knowledge can be acquired (we refer to [106] for further discussion). For the case of novel actions there is evidence that the mirror system is not sufficient to explain action understanding and a reasoning mechanism must be involved [19]. We now discuss several of these views in more detail.

Motor Theories of Perception

Several theories already claimed that the motor system is involved in perception. An example is the motor theory of speech perception [77]. The three main claims in this theory are: "(a) speech processing is special, (b) perceiving speech is perceiving gestures, and (c) the motor system is recruited for perceiving speech". In [50], the authors revisit such theory taking into account the results from the last 50 years. The authors argue that although claim (a) is likely false, claims (b) and (c) are still likely to be true, although they admit that most of the findings supporting such claims may be explained by alternative accounts.

One evidence in favor of the theory that the motor system is involved in perception is the existence of several mechanisms in the brain involved in motor prediction and reconstruction. One such mechanism depends on the existence of several pairs of forward and backward models in the brain [152]. The forward model codes the perceptual effects of motor actions, while the backward model represents the inverse relation, *i.e.,* the motor actions that might cause a given percept. These models provide the agent with "simulation capabilities" for its own body dynamics, and are thus able to adapt to perturbations. They are also general enough to take into account task restrictions.

Mirror and Canonical Neurons

The discovery of *mirror neurons* [51, 96, 106] fostered a significant interest on the brain mechanisms involved in action understanding. These neurons are located in the F5 area of the macaque's brain and discharge during the execution of hand/mouth movements. In spite of their localization in a pre-motor area of the brain, mirror neurons fire both when the animal performs a specific goal-oriented grasping action and when it sees that same action being performed by another individual. This observation suggests that the motor system responsible for triggering an action is also involved in the recognition of the action. In other words, recognition may also involve motor information, rather than purely visual information. Furthermore, by establishing a direct connection between gestures performed by a subject and similar gestures performed by others, mirror neurons may be related to the ability to imitate found in some species [117], establishing an implicit level of communication between individuals.

Canonical neurons [96] have the intriguing characteristic of responding when objects that afford a *specific* type of grasp are present in the scene, even if the grasp action is not performed or observed. Thus, canonical neurons may encode object affordances, as introduced in [55], and may help distinguishing ambiguous gestures during the process of recognition. In fact, many objects are grasped in very precise ways that allow the object to be used for specific purposes. A pen is usually grasped in a way that affords writing and a glass is held in such a way that we can use it to drink. Hence, by recognizing an object that is being manipulated, it is also possible to attain information about the most likely grasping possibilities (expectations) and hand motor programs, simplifying the task of gesture recognition.

Reasoning Processes

Even if there is strong evidence that the motor system is involved in perception, it is not clear how fundamental it is and many claims on the mirror system are unlikely to hold [56]. For instance, mirror neurons are not strictly necessary for action production as their temporal deactivation does not impair grasping control but only slows it down [47, 106]. On the other hand, more complex mechanisms than mirroring are necessary to understand unexpected behaviors of an agent. In [19] an experiment is presented where a person turns a light on using its knee. Similar demonstrations are shown where the person has the arms occupied with a folder, many folders or none. Results from an *fMRI* scan showed that the mirror mechanism is active during the empty arms situation (expected behavior) but it is not active during the other situation (unexpected behaviour). This and other similar results suggest that action understanding in unexpected situations is achieved by an inference-based mechanism taking the contextual constraints into account. In turn, this indicates that there may exist a reasoning mechanism to understand/interpret the observed behaviors.

1.2.2 Psychology

Studies in behavioral psychology have evidenced the ability of both children and chimpanzees to use different "imitative" behaviors. Individuals of both species also seem to switch between different such behaviors depending on perceived cues about the world [54, 62]. These cues include, for example, the inferred purpose of the observed actions [13, 14, 67] even when the action fails [67, 91, 145]. Other social learning mechanisms are analyzed in [154] under the more general designation of *social influence/learning*. In the continuation, we discuss some examples of behavior switching identified in the literature.

Imitation Capabilities and Behaviour Switching

Imitation and *emulation* are two classes of social learning mechanisms observed in both children and apes [138, 146, 148].[1] In imitation, the learning individual adheres to the inferred goal of the demonstrator, eventually adopting the same action choice.

[1] Other species, such as dogs, have also been shown to switch strategies after having observed a demonstration, as seen in [118].

In emulation, on the other hand, the individual focuses on the observed *effects* of the actions of the demonstrator, possibly reaching these using a different action choice. The predisposition of an individual to imitate or emulate can thus be confirmed in tasks where the same effect can be achieved using different actions/motor patterns. For example, both chimpanzees and children are able to copy the choice of a push or twist movement in opening a box [147].

Children, in particular, can be selective about which parts of a demonstration to imitate [54, 150], but are generally more prone to imitate than to emulate. For example, children can replicate parts of a demonstration that are clearly not necessary to achieve the most obvious goal – a phenomenon known as *over-imitation* [62]. Over-imitation can be diminished by reducing the social cues or by increasing the urgency of task completion [21, 85, 88]. It has also been argued that over-imitation can occur for a variety of social reasons [103] or because the individuals interpret the actions in the demonstration as causally meaningful [85].

Sensitivity to Task Constraints

Social animals also exhibit some sensitivity to the context surrounding the task execution, particularly task constraints. For example, in [90] 14-month-olds were shown a box with a panel that lit up when the demonstrator touched it with his forehead.The results showed that most infants reproduced the use of the forehead rather than using their hand when presented with the object a week later. This experiment was further extended in [54] by including a condition in which the demonstrator was restricted and could not use her hands. It was observed that only 21% of the infants copied the use of the forehead, against the 69% observed in a control condition replicating the [90] study. It was argued that, in the latter condition, infants recognize no constraints upon the demonstrator and thus encode the use of the forehead as a specific part of the intention. In the restricted case, they recognize the constraint as a extraneous reason for the use of the forehead and do not encode the specific action as part of the intention.

We return to this particular experiment in Section 1.6, in the context of computational models for social learning.

Imperfect Knowledge

Several experiments were conducted to investigate how the knowledge about the world dynamics influences social learning mechanisms. In one archetypical experiment, an individual observes a sequence of actions, not all of which are actually necessary to achieve the outcome. For example, in [62], preschoolers and chimpanzees were presented with two identical boxes, one opaque and one transparent. A demonstrator then inserted a stick into a hole on the top of the box and then into another hole on the front of the box. It was then able to retrieve of a reward from the box. In this experiment, the insertion of the stick into the top hole was unnecessary in order to obtain the reward, but this was only perceivable in the transparent box. The results showed that 3 and 4-year-old children tended to always imitate both actions. On the contrary, chimpanzees were able to switch between emulation and imitation if causal information was available: after having observed demonstrations in a transparent box, the chimpanzees were much less prone to insert the stick into the upper (useless) hole.

Goal Inference

Finally, it has been showed that some species exhibit imitative behavior beyond simple motion mimicry. For example, primates tend to interpret and actually reproduce observed actions in a teleological manner – that is, in terms of the inferred goals of the action [33]. In an experiment designed to test this hypothesis, 3 to 6-year-old children observed a demonstrator reaching across her body to touch a dot painted on a table to one side of her, using the hand on her other side [13]. When prompted to reproduce the observed demonstration, children tended to copy the dot-touching action, but not the use of the contra-lateral hand. However, when the same demonstration was performed without a dot, children tended to imitate the use of the contra-lateral hand. It was argued that, in the first case, children interpreted the dot touching as the intention, choosing their own (easier) way to achieve it, while in the second case there was no clear target of the action but the action itself. As such, children interpreted the use of the contra-lateral hand as the intention and imitated it more faithfully. Results in experiments adapted for older children infants are similar [28].

1.2.3 Remarks

The experiments described above show that imitation behaviors result from several complex cognitive skills such as action understanding, reasoning and planning. Each of them depends on the physical and social context and also the knowledge of the agent. Partial world knowledge and contextual restrictions all influence the way an action is understood and replicated. A robot that is able to imitate in a flexible way should thus be able to consider all of such aspects.

1.3 Imitation in Artificial Systems

Imitation learning brings the promise of making the task of programming robots much easier [127]. However, to be able to imitate, robots need to have several complex skills that must be previously implemented or developed [84].

In Section 1.2 we discussed some of the complexities involved in the process of learning by imitation in natural systems, as well as all the contextual information taken into account when interpreting actions. Now, we replicate this discussion for artificial systems, outlining some of the issues that must be dealt with when developing an artificial system (*e.g.,* a robot) that can learn by imitation.

In [151], the authors identified three subproblems (or classes thereof) to be addressed in developing one such system:

- Mapping the perceptual variables (*e.g.,* visual and auditory input) into corresponding motor variables;
- Compensating for the difference in the physical properties and control capabilities of the demonstrator and the imitator; and
- Understanding the intention/purpose/reason behind an action (*e.g.,* the cost function to be minimized in optimal control that determines the action to be taken in each situation) from the observation of the resulting movements.

If we further take into account that the perceptual variables from the demonstrator must also be mapped from an allo- to an ego- frame of reference, the first of the above subproblems further subdivides into two other sub-problems: viewpoint transformation and sensory-motor matching [8, 22, 82, 83, 123]. The second of the problems referred above is usually known as the *body correspondence problem* [4, 99, 100] and is, in a sense, closely related and dependent on the first problem of mapping between perception and action.

Data Acquisition

The way to address the issues discussed above will largely depend on the context in which imitation takes place. When used for robot programming, it is possible to use *data acquisition* systems that simplify the interpretation and processing of the input data, thus reducing partial observability issues. The latter is important since the learner will seldom be able to unambiguously observe all the relevant aspects of the demonstration. In particular, this can allow more robust and efficient algorithms to tackle the allo-ego transformation, the perception-to-action mapping and the body correspondence. Examples of such systems include exoskeletons, optical trackers or kinesthetic demonstrations [16].

Other applications of imitation occur in less controlled environments, for example as a result of the natural interaction between a robot and a human. In such contexts, perceptual problems must be explicitly addressed. Some authors address this problem adopting a computer vision perspective [82], modeling partial observability [40] or being robust to noise in the demonstration [26, 80, 116].

Mapping of Perceptual Variables and Body Correspondence

Many approaches do not consider a clear separation between data acquisition and learning by demonstration. One way to deal with the lack of information/data is to use prior knowledge to interpret the demonstration. *Action interpretation* strongly depends on the knowledge about how the world evolves as well as on the capabilities of both the learner and the demonstrator to interact with it [54, 62, 79]. This process is closely related with the two first issues pointed out in [151], since it provides a way to map external inputs to internal motor representations (*e.g.,* to robot actions). Therefore, imitation learning algorithms will typically benefit from prior knowledge about the environment, specially when data acquisition cannot provide a full description of the demonstration.

Knowledge about the agent's own body and its interaction with the world simplifies some of the difficulties found in imitation. On one hand, for the perception-to-action mapping, the recognition of others' actions can rely on the learner's model of the world dynamics, *e.g.,* by inferring the most probable state-action sequence given this model. This idea draws inspiration from psychological and neurophysiological theories of motor perception, where recognition and interpretation of behavior are performed using an internal simulation mechanism [5, 51, 83]. As seen in Section 1.2, mirror neurons are one of such mechanisms [51], and a significant amount of research in imitation learning in robotics flourished from this particular discovery.

On the other hand, this type of knowledge also allows action recognition and matching to occur with an implicit body correspondence, even if the bodies of the learner and demonstrator are different. Several works have explored this idea. For example, in [82, 83, 99, 100, 130, 132], action matching is addressed at a trajectory level. In these works, the demonstration is interpreted taking into account the different dynamics of the learner and the demonstrator. In [71, 93, 94], the same problem is addressed at a higher level of abstraction that considers the *effects* on objects.

Goal/Intention Inference

Understanding actions and inferring intentions generally requires a more explicit reasoning process than just a mirror-like mechanism [19]. As discussed in Section 1.2, by further abstracting the process of learning by imitation to *task level* it is possible to additionally include contextual cues [10, 64, 79]. At this level of abstraction the third issue identified in [151] becomes particularly relevant.

Identifying the goal driving a demonstration is a particularly complex inference process, indeed an ill-defined one. What to imitate depends on several physical, social and psychological factors (see Section 1.2.2). One possible way to answer this question relies on the concept of *imitation metrics*. These metrics evaluate "how good" imitation is. Imitation metrics were first explicitly introduced in [101] in order to quantify the quality of imitation, to guide learning and also to evaluate learned behavior. However, it is far from clear what "good imitation" is and, perhaps more important, how variable/well-defined the learned behavior can be. Some studies along this direction have characterized the quality of imitation in humans. In [111], subjects were asked to perform imitation tasks and quantitative results were obtained to assess the effect of rehearsal during observation and repetition of the task.

In any case, it is often not clear whether imitation concerns the motor intention or the underlying goal of that motor intention [17, 23, 79, 127]. In other words, it is often the case that the agent cannot unambiguously identify whether it should imitate the action, its outcome, or the reason driving the demonstrator to do it. In each of the following sections we we discuss in detail each of these three approaches to imitation learning. In particular, we refer to Section 1.6 for a more detailed discussion on the problem of inferring the goal behind a demonstration. In that section we survey several recent works in which a learner does infer the goal of the demonstrator and adopts this goal as its own [2, 10, 64, 80, 119].

The Role of Self-Observation

It is interesting to note that most of the necessary information about the robot's body and the world dynamics can be gathered by self-observation [36,84]. Although slow in many situations, it often allows a greater adaptation to changing scenarios.

Many different techniques can be used by the robot to learn about its own body [41, 84, 108, 129, 144]. For example, several works adopt an initial phase of motor babbling [57, 76, 83, 84]. By performing random motions, a great amount of data becomes available, allowing the robot to infer useful relations about causes and consequences of actions. These relations can then be used to learn a body schema useful in different application scenarios. The specific methods used to learn body models vary, and range from parametric methods [27, 61], neural network methods [76,82,83] to non-parametric regression [41,84,144] and graphical models [57,135]. As for learning about the world dynamics, this is closely related to the concept of learning affordances. Repeated interaction with the world allows the robot to understand how the environment behaves under its actions [46, 94, 95, 122]. As seen in the previous section, the knowledge about the world dynamics and the capabilities of others strongly influences how actions are understood.

$\diamondsuit$

So far in this chapter we presented insights from neurophysiology, psychology and robotics research on the problems involved in learning by demonstration. The next sections will provide an overview of methods that handle such problems. We divide those methods according to the different formalisms or sources of information used, namely (a) trajectory matching and generalization, where sample trajectories are the main source of information from which the learner generalizes; (b) object mediated imitation, where effects occurring on objects are the relevant features of a

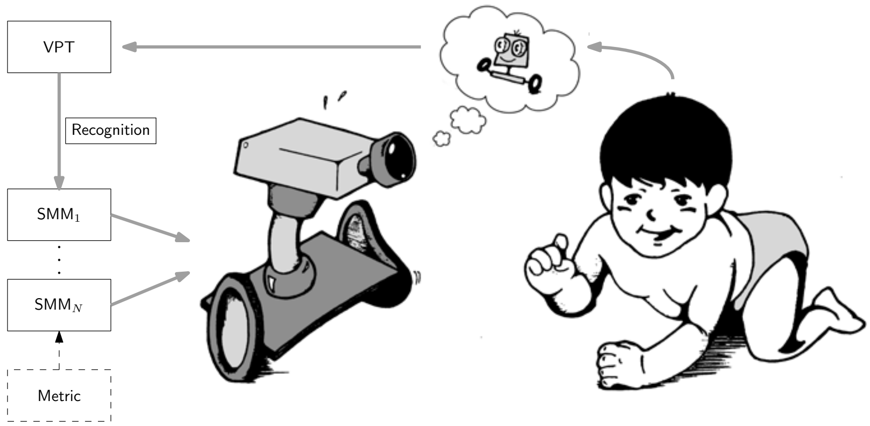


Fig. 1.3: Imitation architecture. Observed actions are first transformed to a ego frame of reference (VPT), where segmentation and recognition take place. After deciding on how to imitate, a correspondence between the two different bodies must be done by selecting the corresponding SMM. Finally, imitation is enacted.

demonstration; and (c) imitation of inferred goals, where there is an explicit estimation of the demonstrator's goal/intention.

1.4 Imitating by Motor Resonance

This section presents several methods that learn by demonstration by first mapping state-action trajectories to the learner's own body and then generalizing them.

Following what is proposed in [83, 151], the imitation process consists of the steps enumerated below and illustrated in Fig. 1.3:

(i) The learner observes the demonstrator's movements;
(ii) A viewpoint transformation (VPT) is used to map a description in the demonstrator's frame *allo-image* to the imitator's frame *ego-image*;
(iii) Action recognition is used (if necessary) to abstract the observed motion; and
(iv) A sensory-motor map (SMM) is used to generate the motor commands that have the higher probability of generating the observed features.

In this section we survey several methods that adopt this general approach. In these methods not all steps enumerated above are explicitly dealt with, but are still implicitly ensured by considering different simplifying assumptions.

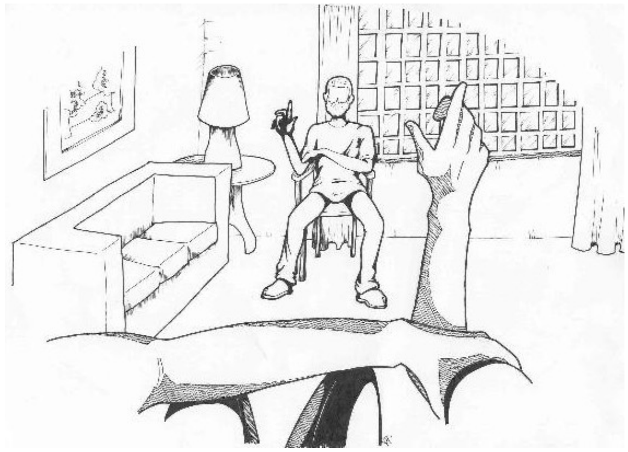


Fig. 1.4: Perceptual difference between the same gesture having an ego- or an allo-perspective.

1.4.1 Visual Transformations

The same motor action can have very distinctive perceptual results if one considers different points-of-view. For example when a person gestures goodbye, she can see the back of her hand, while when someone else is doing the same she will see the palm of the other's hand. This poses a problem when mapping actions from the demonstrator to the learner, as already identified in [22] and depicted in Fig. 1.4.

The typical solutions for this problem is to perform a complete three-dimensional reconstruction. However, several works proposed alternative approaches that consider simplifications to such problem. In [8, 12, 52, 82, 83] several transformations are discussed, ranging from a simple image transformation (*e.g.,* mirroring the image) to a partial reconstruction assuming an affine camera and the full three dimensional reconstruction. These works also point out that such transformations can be seen as imitation metrics, because the depth information in some gestures can indeed change the meaning of the action. Such transformations can also be done using neural networks [123].

1.4.2 Mimicking Behaviors and Automatic Imitation

Several works on imitation seek to transfer behaviors by a simple motor resonance process. The observation of perceptual consequences of simple motor actions can elicit an automatic mimicking behavior, and several initial works on imitation adopted this approach [34, 35, 53, 59]. In these works, the relation between changes in perceptual channels caused by a demonstrator are directly mapped into motor actions of the learner. Particularly in [34], the orientation of a mobile robot is controlled according to the observed position and orientation of a human head.

In these approaches the model about the own body is not explicit, the relations between action and perception is learned without concern about the true geometry and dynamics of the body.

Using the visual transformations introduced earlier it is possible to generate more complex behaviors. In [82, 83], the learner starts by acquiring a correspondence between its own perceptions and actions. The mimicking behavior results from mapping the perception of the demonstrator to its own using a view-point transformation, and then activating an inverse sensory-motor map. Different visual transformations result in different types of behavior.

Examples of other methods posing imitation within this visuo-somatic perspective include [7,53]. An interesting approach is proposed in [18] where facial expression are learned by interacting with a person and creating a resonance of expressions.

1.4.3 Imitation through Motor Primitives

All previous approaches consider a simple perception-action loop in imitation. However, when considering a learner equipped with several motor primitives, it is possible to achieve more complex interactions. The "recognition" block in Fig. 1.3 represents the translation of observed trajectories in terms of such motor primitives. A motor sequence is perceived, maybe after some visual processing, and recognized as a specific motor primitive [86, 87]. This general approach can be used to perform human-machine interaction [20] or to learn how to sequence such primitives [23].

In [23] a string parsing mechanism is proposed to explain how apes are able to learn by imitation to process food. The string parsing mechanism is initialized with several sequences of primitives. Learning and generalization are performed by extracting regularities and sub-sequences. This approach can be seen as a grammatical inference process.

Other approaches use hidden Markov models to extract such regularities and filter behavior, usually in the context of tele-operation. The main goal of such approaches is to eliminate sluggish motion of the user and correct errors. We refer to [153] for an application of one such approach to the control of an Orbit Replaceable Unit. In this work, a robot observes an operator performing a task and builds a hidden Markov model that describes that same task. In [63], a similar approach is used in assembly tasks. Such models can also be used to detect regularities in human motion [24].

Regularities of human motion can be represented in low-dimensions using principal component analysis [29], clustering [74] or other non-linear manifold learning techniques [65].

Some authors rely on completely separated methods to recognize and to generate motor primitives. Others combine both processes, thus exploring the self-modeling phase [37, 38, 66, 83]. As seen in Section 1.2.1, this process can be explained by the use of motor simulations of some kind. For example in both [38] and [151], action recognition is performed using the idea of coupled forward/backward models discussed in [152]. The former decouples each of the action primitives and is thus able to deal with a larger variety of tasks, while the latter is able to combine several primitives and deal with complex motor skills in a robust way. In [43,44] dynamic neural networks are used to recognize actions goals taking into account task restrictions. A single neural network able to encode several behaviors was introduced in [137] and performs similar computations.

1.4.4 Learning of New Task Solutions

In some cases the learner has an explicit goal. However, it might be very difficult to plan how to reach such goal. This is specially important in complex environments or in situations involving highly redundant robots (*e.g.,* humanoid robots). One of the main motivations behind imitation learning is that it provides an easy way to program robots. Therefore, most approaches to imitation consider the learning of new actions. In such approaches two main trends have been adopted: one considers many demonstrations of the same task and tries to find invariants in the observed motions [17, 119]. The other uses only the observed trajectories as an initialization and then improves and generalizes further (*e.g.,* using reinforcement learning) [109].

In [112], imitation is used to speed up learning and several metrics were defined for evaluating the improvement in learning when using imitation. Imitation is used in [129] to learn dynamic motor primitives. As argued in this work, "*Movement primitives are parameterized policies that can achieve a complete movement behavior*". From this perspective, motor primitives can be seen as dynamical systems that generate different complex motions by changing a set of parameters. The authors also suggest the use of data from a demonstration to initialize such parameters. The parameters can then be optimized using, for example, policy gradient [109]. We point out that these methods consider classes of parameterized policies, namely the parameters of the dynamical system.

One of the few approaches taking uncertainty into account is proposed in [57]. The approach in this work starts by learning a Gaussian mixture model as a forward model, using self-observation [130]. The demonstration is taken as observation of a probabilistic process and the goal is to find a sequence of actions that maximizes the likelihood of such evidence. The work focuses on copying motions taking into account the dynamics of the robot and, as such, uses as observations the estimated state trajectory and ignores the dynamic information. It achieves body correspondence by inferring the most probable trajectory using the imitator's body. Extra task restrictions can also be included. In [29] walking patterns are transferred from humans to robots after adapting it for different kinematics using low level representations.

Finally, several other methods are agnostic as to what is exactly the goal of the demonstrated task and aim only at learning the observed course of action. For example, in [31], learning from demonstration is formulated as a classification problem and solved using support-vector machines. These methods, in a sense, disregard the effects and social context of the demonstration, and focus only on replicating in each situation the demonstrated course of action (see Section 1.6 for a more detailed discussion on this). One disadvantage of this general approach is that it places excessive confidence on the demonstrator. Furthermore, the course of action learned is specific to the context and environment of the learner and, as such, is not generalizable to different environments.

1.5 Object Mediated Imitation

In the previous section, we have discussed imitation from a motor perspective. From this perspective, context is mainly provided by the body parts and corresponding motion. In this section we discuss a more abstract approach to imitation, where the context is enlarged to accommodate objects. In other words, the learner is now aware of the interaction with objects and, consequently, takes this information into account during learning. The most representative example of this type of behavior is *emulation* (see Fig. 1.2). In contrast with the motor resonance mechanisms discussed previously, which could perhaps be best described as mimicry, emulation focuses on copying (replicating) the results/effects of actions.

This abstraction from low-level control to higher-level representations of actions also facilitates reasoning about causality of actions, *i.e.,* how to induce specific changes to the environment. Consider the simple case of piling two objects. To learn this task, motor resonance alone does not suffice, as the learner must take into account which object can be placed on top of which depending on their sizes, shapes and other features. Another illustrative example is opening a door, where the shape of the handle provides meaningful information about how to perform the action. In the motor-resonance-based approaches, body correspondence addressed problems such as different number of degrees of freedom, or different kinematics and dynamics. In this section we discuss correspondence in terms of the usage that different objects have to different agents.

A biologically inspired concept related to the previous discussion is that of *affordances* [55]. Developed in the field of psychology, the theory of affordances states that the relation between an individual and the environment is strongly shaped by the individual's perceptual-motor skills. Back in the 70s, this theory established a new paradigm where action and perception are coupled at every level. Biological evidence of this type of coupling is now common in neuroscience [51] and several experiments have shown the presence of affordance knowledge based on the perception of heaviness [141] or traversability [69].

Affordances have also been widely studied in robotics. In this section we discuss this concept of affordances in the context of imitation learning in robots, as well as the inclusion of object information and properties in the learning process. A thorough review on these topics can be found in [122], with special emphasis placed in affordance-based control.

We generally define affordances as mappings that relate *actions*, *objects* and *consequences* (effects). This very general concept can be modeled using different formalisms including dynamical systems [131], self-organizing maps [32], relational learning [58] and algebraic formulations [100].

However, independently of the selected representation or formalism, there are two core challenges to achieve affordance-based imitation: acquiring the model of affordances and exploiting this model. The latter depends heavily on the representation, but usually resorts to some type of action selection. For instance, if a dynamical system is used to encode a forward model, then the agent emulates by selecting the action that best matches the desired effect. It is worth mentioning that the approaches in this section are strongly dependent on a self-modeling phase as the concept of affordances is strictly a self-modeling idea.

The required data to infer the affordances model may be acquired either by *observation* or by *experimentation*. In the case of self-experimentation there is no body correspondence or visual transformation required, but such capability is important when interacting with other agents. When learning by observation such problem is immediately present. An advantage of *object mediated imitation* is that the match only occurs in the effects on object and so the specific body kinematics and action dynamics are not considers, thus simplifying several problems in imitation.

Affordances as Perception-Action Maps

A simple way of describing effects is to learn mappings from a set of predefined object features to changes in these features. This approach was used in a manipulation task to learn by experimentation the resulting motion directions as a function of the object shape and the poking direction [46]. Once the mapping has been learned, emulating an observed motion direction can be achieved by simply selecting the appropriate poking direction for the object. A similar approach was proposed in [71], where the imitation is also driven by the effects. In this case, the demonstrator's and imitator's actions are grouped according to the effect they produce in an object, irrespectively of their motor programs. Given an observed object and effect pair, the appropriate action (or sequence of actions) can then be easily retrieved. Another example is the use of invariant information [32, 48, 133]. In this case, the system learns invariant descriptions across several trials of the same action upon an object. Depending on the parameterization, the learned invariant descriptors may represent object characteristics or effects. Although this type of information has been usually applied in robot control, it is possible to use the invariants for emulation under the assumption that they are invariant to the viewpoint and that they capture the effects.

Grasping is a paradigmatic example of affordance knowledge that has been widely studied in the literature. Perception-action maps have appeared in several forms such as the Q-function of a reinforcement learning algorithm [155]; a pure regression map from object features to image points based labelled examples [125, 126] and self-experimentation [39, 92]; or as the correct position of a mobile robot to trigger a particular grasping policy [134].

Affordances as Dynamical Models

An alternative approach consists in modeling the dynamical system composed by the agent (demonstrator or imitator) and the environment. In [131], a hidden Markov model is used to encode the state of the agent and objects. In order to train the forward model, reflective markers were placed on the demonstrator and on the objects and tracked by a capture system. Viewpoint transformation then uses linear transformations between the demonstrator's and the imitator's body poses. Emulation is casted as a Bayesian decision problem over the Markov model, *i.e.,* selecting the maximum a posteriori action for each transition of the Markov chain. Interestingly, the model is able to modify the selected behavior with its own experience and refine the model previously learned solely by observation. Again, this is due to the fact that emphasis is placed on achieving the same effects instead of copying the action.

Dynamical systems have also been used for goal directed imitation in [43, 44]. The proposed architecture contains three interconnected layers corresponding to different brain areas responsible for the observed action, the action primitives and the goal. Each layer is implemented using a dynamic field that evolves with experience and its able to incorporate new representations using a correlation learning rule between adjacent neurons.

1.5.1 Bayesian Networks as Models for Affordances

Affordances can be seen as statistical relations between actions, objects and effects, modeled for example using Bayesian networks. One such approach was proposed in [94], in which the nodes in the network represent actions, object features or measured effects. As in standard Bayesian networks, the absence of a vertex between two nodes indicates conditional independence. Self-experimentation provides most of the data necessary to learn such relations. If a robot exerts its actions upon different objects, it can measure the effects of such actions. Even in the presence of noise the robot is able to infer that some actions have certain effects that depend on some of the object features. Also, the existence of irrelevant and redundant features is automatically detected.

Based on such prior experience, structure learning [60] can be used to distinguish all such relations. Once these dependencies are known, one can query the network to provide valuable information for several imitation behaviors. Table 1.1 summarizes the different input-output combinations. In particular, for emulation purposes, the ability to predict the effect conditioned on a set of available objects and the possible actions directly provides the robot with a way to select the appropriate action in a single-step Bayesian decision problem. It is interesting to note that at this abstraction level the same mechanism is used for prediction and control, giving a mirror-like behavior.

This approach is halfway between learning specific maps and a full dynamical system description. If specific maps are learned, it is not easy to consider demon-

Table 1.1: Affordances as relations between actions (A), objects (O) and effects (E) that can be used for different purposes: predict the outcome of an action, plan actions to achieve a goal or recognize objects or actions

Inputs	Outputs	Function
(O,A)	E	Predict effect
(O,E)	A	Action recognition and planning
(A,E)	O	Object recognition and selection

strators with different dynamics nor to explicitly consider task restrictions. This approach is not as applicable as learning a full dynamical system description, because it is does not easily allow encoding long-term plans. It provides predictions for incremental state changes. For a more detailed discussion, we refer to [80].

1.5.2 Experiments

In this section we provide several experimental results obtained with a Bayesian network model for affordances. In particular, we describe both how the model of affordances can be learned by the robot and then used to attain affordance-based emulation behaviors.

For all experiments we used BALTAZAR [78], a robotic platform consisting of a humanoid torso with one anthropomorphic arm and hand and a binocular head (see Fig. 1.5). The robot is able to perform a set of different parameterized actions, namely $\mathscr{A} = \{a_1 = grasp(\lambda), a_2 = tap(\lambda), a_3 = touch(\lambda)\}$, where λ represents the height of the hand in the 3D workspace when reaching an object in the image. It also has implemented an object detector that extracts a set of features related to the object properties and the effects of the action.

Affordance Learning

We now describe the process by which the robot learned the affordance network used in the emulation behaviors. We recorded a total of 300 trials following the protocol depicted in Fig. 1.6. At each trial, the robot was presented with a random object of one of two possible shapes (round and square), four possible colors and three possible sizes (see Fig. 1.5 for an illustration of the objects used). BALTAZAR then randomly selects an action and moves its hand toward the object using pre-learned action primitives [84, 94]. When the reaching phase is completed, the robot then performs the selected action ($grasp(\lambda)$, $tap(\lambda)$ or $touch(\lambda)$) and finally returns the hand to the initial position. During the action, object features and effects are recorded.

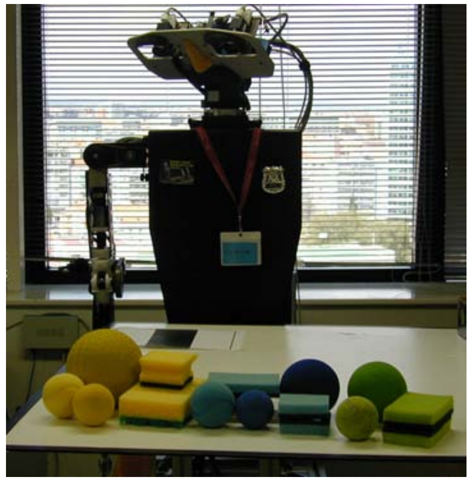


Fig. 1.5: Robot playground used in the experiments.

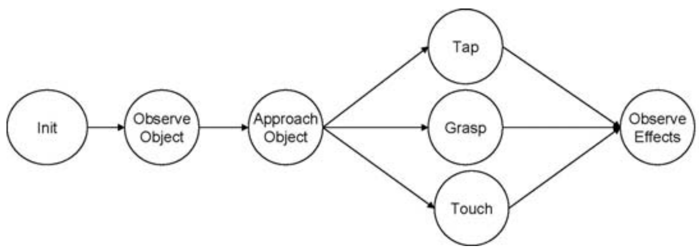


Fig. 1.6: Protocol used in the experiments. The object used in each trial is selected manually and the robot then randomly selects an action to interact with it. Object properties are recorded from the Init to the Approach states, when the hand is not occluding the object. The effects are recorded in the Observe state. Init moves the hand to a predefined position in open-loop.

Visual information is automatically clustered using the *X-means algorithm* [107]. The resulting classes constitute the input for the affordance learning algorithm. The features and their discretization are shown in Table 1.2. Summarizing, shape descriptors (*e.g.,* compactness and roundness) provided two different classes, size was discretized in 3 different classes and color in four. Based on this data, the robot adjusted the parameter λ for each action and then learned an affordance model as described above.

The learned model is shown in Fig. 1.7. The network was learned using Monte Carlo sampling with BDeu priors for the graph structure and a random network initialization. The dependencies basically state that color is irrelevant for the behavior of the objects under the available actions. In addition to this, a successful grasp requires the appropriate object size, while the velocity of the object depends on its

Table 1.2: Random variables in the network and possible values.

Symbol	Description	Values
A	Action	*grasp, tap, touch*
C	Color	*green*$_1$,*green*$_2$, *yellow, blue*
Sh	Shape	*ball, box*
S	Size	*small, medium, large*
OV	Object velocity	*small, medium, large*
HV	Hand velocity	*small, medium, large*
Di	Object-hand velocity	*small, medium, large*
Ct	Contact duration	*none, short, long*

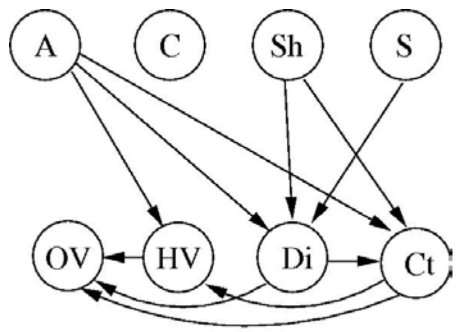


Fig. 1.7: Affordance network representing relations between actions, object features and the corresponding effects. Node labels are shown in Table 1.2.

shape and the action. We refer the reader to [94] for further details on the affordance learning.

Emulation

We now present the results obtained in several basic interaction games using the affordance network. The games proceed as follow. The robot observes a demonstrator performing an action on a given object. Then, given a specific imitation metric, it selects an action and an object to interact with so as to imitate (emulate) the demonstrator. Figure 1.8 depicts the demonstration, the different objects presented to the robot and the selected actions and objects for different metrics.

In the experiments we used two different demonstrations, a tap on a small ball (resulting in high velocity and medium hand-object distance) and a grasp on a small square (resulting in small velocity and small hand-object distance). Notice that contact information is not available when observing others.

The goal of the robot is to replicate the observed effects. The first situation (Fig. 1.8a) is trivial, as only tapping has a non-zero probability of producing high velocity. Hence, the emulation function selected a tap on the single object available.

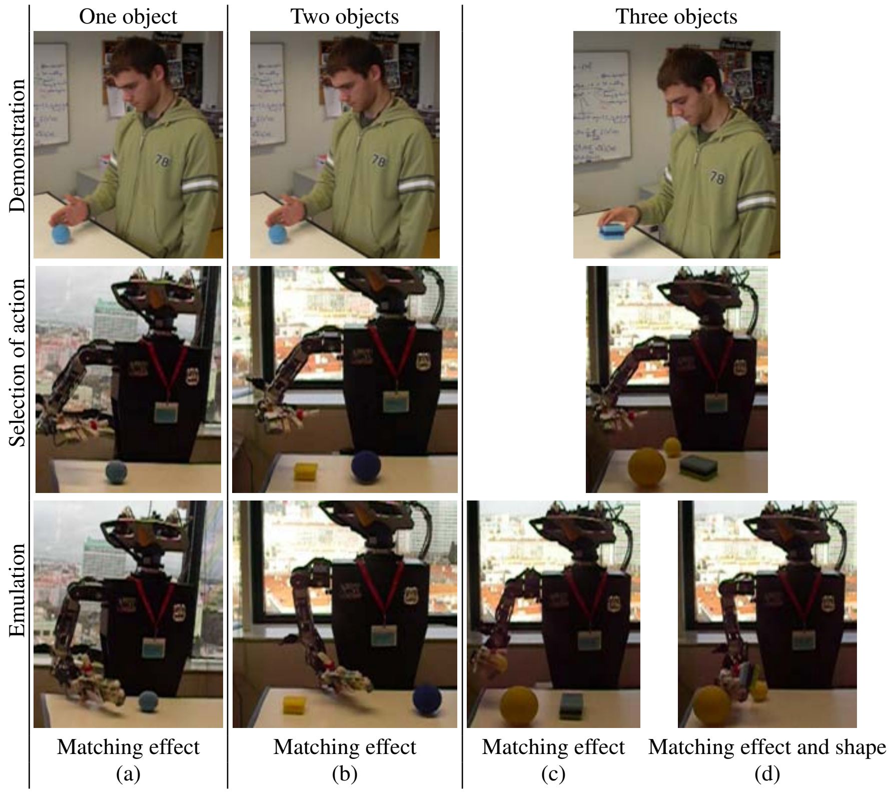


Fig. 1.8: Different emulation behaviors. Top row: Demonstration; Middle row: Set of potential objects; Bottom row: Emulation. Situations (a) through (d) represent: (a) emulation of observed action, (b) replication of observed effect, (c) replication of observed effect, and (d) replication of observed effect considering the shape of the object.

In Fig. 1.8b the demonstrator performed the same action, but the robot now has to decide between two different objects. Table 1.3 shows the probabilities for the desired effects given the six possible combinations of actions and objects. The robot selected the one with highest probability and performed a tap on the ball.

Table 1.3: Probability of achieving the desired effect for each action and the objects of Fig. 1.8b.

Obj \ Action	Grasp	Tap	Touch
Large blue ball	0.00	0.20	0.00
Small yellow box	0.00	0.06	0.00

Figures 1.8c and 1.8d illustrate how including the object features in the metric function produce different behaviors. After observing the grasp demonstration, the robot has to select among three objects: large yellow ball, small yellow ball and small blue box. In the first case the objective was to obtain the same effects. The probability of grasping for each of the objects is 0.88, 0.92 and 0.52, respectively, and the robot grasped the small yellow ball even if the same object is also on the table (Fig. 1.8c). Notice that this is not a failure, since it maximizes the probability of a successful grasp which is the only requirement of the metric function.

We conclude by noting that other criteria can include more complex information, such as similarly shaped objects. When also taking this new criterion into account, the robot selected the blue box instead (see Fig. 1.8d).

1.6 Imitating by Inferring the Goal

So far in this chapter we have discussed learning by imitation at the motor level and at the effect level. The former focuses on replicating the exact movement observed, in a sense disregarding the effect on the environment and the social context in which the movement is executed. The latter addresses imitation at a higher level of abstraction, focusing on replicating the effects produced by the observed action in the environment, ignoring to some extent the exact motor trajectory executed. As seen in Section 1.2, knowledge about the world, the demonstrator and the learner's own body all influence the way a goal is inferred.

In this section we discuss imitation learning at a yet higher level of abstraction, approaching the concept of "imitation" according to the taxonomy in Fig. 1.2. Concretely, we discuss the fundamental process by which a learner can infer the *task* to be learned after observing the demonstration by another individual (*e.g.,* a human). We discuss several approaches from the literature that address the problem of inferring the goal of a demonstration at different levels. We then discuss in detail a recent approach to this problem that provides a close relation and potentially interesting insights into imitation in biological contexts.

1.6.1 Goal Inference from Demonstration

Inferring the goal behind a demonstration is, in general, a hard problem, as it requires some form of common background for the learner and the demonstrator. In social animals, this common background greatly depends on the social relation between the demonstrator and the learner. For example, social cues were found important in promoting imitation in infants [21, 91]. Several other studies address the general problem of understanding the process of inferring the goal/intention behind a demonstration [14, 67, 91]. Most such studies also address the related problem of understanding the process of perceiving *unfulfilled intentions*.

Translating this complex social learning mechanism into artificial systems usually requires the common background to be provided by the designer, who "imprints" in the system whatever of her own background knowledge it determines to be relevant for the particular environmental context of the system. As such, it is hardly surprising that different researchers address the problem of "goal inference" from perspectives as distinct as their own backgrounds and lines of work. For example, in [17] goal inference is cast as an optimization problem. Motor theories of perception are used in [38] to build better imitation systems. These works essentially seek to determine which specific elements of the demonstration are relevant, seeking a *hard* answer to the fundamental problem of "What to imitate" discussed in Section 1.3.

Other recent works have adopted a fundamentally different approach, in which the learning agent chooses among a library of possible goals the one most likely to lead to the observed demonstration. For example, in [64] the problem of imitation is tackled within a planning approach. In this setting, the learner chooses between a pool of possible goals by assessing the optimality of the demonstration (viewed as a plan). Evaluative feedback from the demonstrator is also used to disambiguate between different possible goals.

One significant difficulty in inferring the goal behind a demonstration is that the same observed behavior can often be explained by several possible goals. Goal inference is, therefore, an ill-posed problem, and many approaches adopt a probabilistic setting to partly mitigate this situation [10, 119, 142]. For example, in [10], the authors address the problem of action understanding by children. To this purpose, they propose the use of a Bayesian model that matches observed inferences in children facing new tasks or environmental constraints. Similar ideas have been applied to robots in different works [80, 119, 142]. In [119], the goals of the robot are restricted to shortest-path problems while in [80, 142] general goals are considered. In [156], a maximum entropy approach is used to infer the goal in navigation tasks. The paper computes a distribution over "paths to the goal" that matches the observed empirical distributions but otherwise being as little "committed" as possible. Optimization is performed by a gradient-based approach. All these approaches handle the body correspondence problem by performing the recognition in terms of a self-world model.

In a sense, all the aforementioned approaches interpret the demonstration as providing "implicit" information about the goal of the demonstrator, a *soft* answer to the problem of "What to imitate". In other words, while the approaches in [17, 38] seek to single out a particular aspect of the demonstration to replicate, the latter assumes that the actual goal of the demonstration drives the action choice in the demonstration, but needs not be "contained" in it. This makes the latter approach more flexible to errors in the demonstration and non-exhaustive demonstrations. Another way of looking at the distinction between the two classes of approaches outlined above is by interpreting the latter as providing models and methods that allow the agent to extract a general *task description* from the demonstration, rather than a specific mapping from situations to actions that may replicate, to some extent, the observed behavior. This approach is closer to imitation in the biological sense, as defined in [25]. Finally, several recent works have proposed general models that contrast with those referred above in that they are able to generate *multiple* social-learning behaviors [79, 89].

In the remainder of this section we describe in detail the approach in [79]. Following the taxonomy in [25], our model takes into account several possible sources of information. Concretely, the sources of influence on our model's behavior are: beliefs about the world's possible states and the actions causing transitions between them; a baseline preference for certain actions; a variable tendency to infer and share goals in observed behavior; and a variable tendency to act efficiently to reach the observed final states (or any other salient state).

1.6.2 Inverse Reinforcement Learning as Goal Inference

We now introduce the general formalism used to model the interaction of both the learning agent and the demonstrator with their environment. This approach shares many common concepts with those in [10, 119, 142], in that it infers a goal from the behaviors using bayesian inference to deal with noise and to disambiguate the set of possible goals.

Environment Model

At each time instant t, the environment can be described by its *state*, a random variable that takes values in a finite set of possible states (the state-space). The transitions between states are controlled to some extent by the actions of the agent (the demonstrator during the demonstration and the learner otherwise). In particular, at each time instant the agent (be it the learner or the demonstrator) chooses an action from its (finite) repertoire of action primitives and, depending on the action particular action chosen, the state evolves at time $t+1$ according to some transition probabilities $\mathbf{P}[X_{t+1} \mid X_t, A_t]$.

We assume that the learner has knowledge of its world, in the sense that it knows the set of possible states of the environment, its action repertoire and that of the demonstrator, and the world dynamics, *i.e.,* how both his and the demonstrator's actions affect the way the state changes (the transition probabilities). Note that we do not assume that this world knowledge is *correct*, in the sense that the agent may not know (or may know incorrectly) the transitions induced by certain actions. In any case, throughout this section we assume this knowledge as fixed – one can imagine the approach described herein eventually to take place after a period of self-modeling and learning about the world.[2] In this section, the modeled agent does not learn new actions, but instead learns how to combine known actions in new ways. In this sense, it is essentially distinct from the approach surveyed in Section 1.4.

[2] To our knowledge, no work exists that explores knowledge acquisition simultaneously with learning by imitation, but we believe that such approach could yield interesting results.

Finally, in a first approach, we assume that the agent is able to recognize the actions performed by the demonstrator. In this section we do not discuss how this recognizer can be built, but refer that it can rely, for example, on the affordance models discussed in Section 1.5. Toward the end of the section we briefly discuss how the ability to recognize the demonstrator's actions affects the ability of the learner to recover the correct task to be learned (see also [80]).

In our adopted formalism, we "encode" a general task as a function r mapping states to real values that describes the "desirability" of each particular state. This function r can be seen as a *reward* for the learner and, once r is known, the problem falls back into the standard framework of Markov decision processes [115]. In fact, given the transition probabilities and the reward function r, it is possible to compute a function Q_r that, at each possible state, provides a "ranking" of all actions detailing how useful each particular action is in terms of the overall goal encoded in r. From this function $Q_r(x,a)$ it is possible to extract an optimal decision rule, henceforth denoted by π_r and referred as the *optimal policy for reward r*, that indicates the agent the best action(s) to choose at each state,

$$\pi_r(x) = \arg\max_a Q_r(x,a)$$

The computation of π_r, or equivalently Q_r, given r, is a standard problem and can be solved using any of several standard methods available in the literature [115].

Bayesian Recovery of the Task Description

We consider that the demonstration consists of a sequence $\mathscr{D}$ of state-action pairs

$$\mathscr{D} = \{(x_1,a_1),(x_2,a_2),\ldots,(x_n,a_n)\}.$$

Each pair (x_i,a_i) exemplifies to the learner the expected action (a_i) in each of the states visited during the demonstration (x_i). In the formalism just described, the goal inference problem lies in the estimation of the function r from the observed demonstration $\mathscr{D}$. Notice that this is closely related to the problem of *inverse reinforcement learning* as described in [1]. We adopt the method described in [89], which is a basic variation of the *Bayesian inverse reinforcement learning* (BIRL) algorithm in [116], but the same problem could be tackled using other IRL methods from the literature (see, for example, [102, 136]).

For a given reward function r, we define the *likelihood of a state-action pair*, (x,a), as

$$L_r(x,a) = \mathbf{P}\left[(x,a) \mid r\right] = \frac{e^{\eta Q_r(x,a)}}{\sum_b e^{\eta Q_r(x,b)}},$$

where η is a user-defined *confidence parameter* that we describe further ahead. The value $L_r(x,a)$ translates the plausibility of the choice of action a in state x when the underlying task is described by r. Therefore, the likelihood of a demonstration sequence $\mathscr{D}$ can easily be computed from the expression above. We use MCMC to estimate the distribution over the space of possible reward functions given the demonstration (as proposed in [116]) and choose the maximum *a posteriori*.

We note that it may happen that the transition model available is *inaccurate*. In this situation, the learner should still be able to perceive the demonstrated task, given that the "errors" in the model are not too severe. We also note that, in the process of estimating this maximum, the learner uses the knowledge concerning the action repertoire and world dynamics *of the demonstrator*. After the task description (the reward function r) is recovered, the learning agent then uses its own world model to compute the right policy for the recovered task in terms of its own world dynamics and action repertoire.

A Model for Social Learning Mechanisms

Following the taxonomy in Fig. 1.2, we include in our model of social learning three sources of information to be used by the learner in determining the behavior to adopt [79]. The sources of influence on our model's behavior are baseline preferences for certain actions; a tendency to infer and share goals in observed behavior; and a tendency to act efficiently to reach rewarding states. Each of the three sources of information is quantified in terms of a utility functions Q_A, Q_I and Q_E, respectively. The learner will adhere to the decision-rule obtained by combining the three functions. In particular, the learner will adhere to the decision-rule associated with the function

$$Q^* = \lambda_A Q_A + \lambda_E Q_E + \lambda_I Q_I, \tag{1.1}$$

with $\lambda_A + \lambda_I + \lambda_E = 1$. By resorting to a convex combination as in Eq. 1.1, there is an implicit trade-off between the different sources of information (see also Fig. 1.9). It remains to discuss how Q_A, Q_I and Q_E are computed from the demonstration.

- The first source of information is the learner's *preference between actions*. This preference can be interpreted, for example, as representing a preference for "easier" actions than "harder" actions, in terms of the respective energetic costs. This preference corresponds to *natural inclinations* of the learner, and is independent of the demonstration. The preference is translated in the corresponding utility function Q_A, whose values are pre-programmed into the agent.[3]
- The second source of information corresponds to the desire of the learner to replicate the *effects* observed in the demonstration. For example, the learner may wish to reproduce the change in the surroundings observed during the demonstration, or to replicate some particular transition experienced by the teacher. This can be translated in terms of a utility function Q_E by considering the reward function that assigns a positive value to the desired effect and then solving the obtained Markov decision process for the corresponding Q-function. The latter is taken as Q_E.

[3] The exact values of Q_A translate, at each state, how much a given action is preferred to any other. The values are chosen so as to lie in the same range as the other utility functions, Q_I and Q_E.

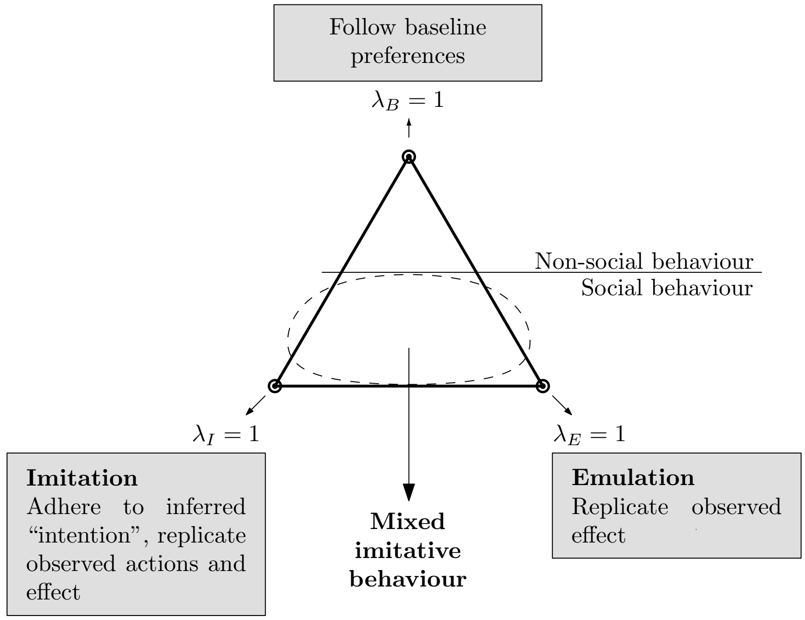


Fig. 1.9: Combination of several simple behaviors: Non-social behavior, emulation and imitation. The line separates the observed social vs non-social behavior, and does not correspond to the agent's reasoning (reproduced with permission from [79]).

- The third and final source of information is related to the desire of the learner to pursue the same *goal* as the teacher. Given the demonstration, the learner uses the Bayesian approach outlined before, inferring the underlying intention of the teacher. Inferring this intention from the demonstration is thus achieved by a teleological argument [33]: the goal of the demonstrator is perceived as the one that more rationally explains its actions. Note that the goal cannot be reduced to the final effect only, since the means to reach this end effect may also be part of the demonstrator's goal. We denote the corresponding utility function by Q_I.

It is only to be expected that the use of different values for the parameters λ_A, λ_E and λ_I will lead to different behaviors from the learner. This is actually so, as illustrated by our experiments. We also emphasize that Q_E greatly depends on the world model of the *learner* while Q_I also depends on the world model of the *teacher*.[4]

[4] Clearly, the world model of the learner includes all necessary information relating the action repertoire for the learner and its ability to reproduce a particular effect. On the other hand, the world model of the teacher provides the only information relating the decision-rule of the teacher and its eventual underlying goal.

1.6.3 Experiments

In this section we compare the simulation results obtained using our proposed model with those observed in a well-known biological experiment in children. We also illustrate the application of our imitation-learning framework in a task with a robot.

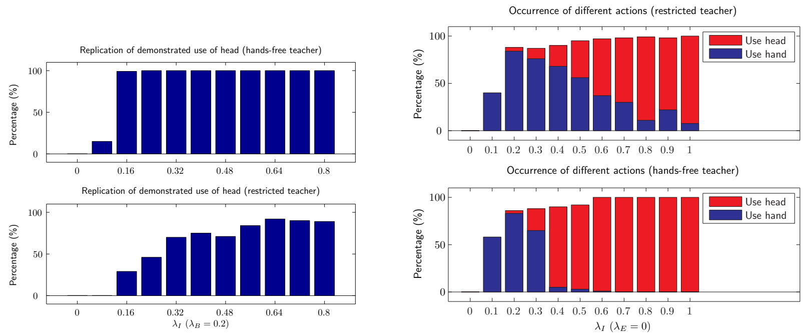
Modeling Imitation in Humans

In a simple experiment described in [90], several infants were presented with a demonstration in which an adult turned a light on by pressing it with the head. One week later, most infants replicated this peculiar behavior, instead of simply using their hand. Further insights were obtained from this experiment when, years later, a new dimension to the study was added by including task constraints [54]. In the new experiment, infants were faced with an adult turning the light on with the head but having the hands restrained/occupied. The results showed that, in this new situation, children would display a more significant tendency to use their hands to turn the light on. The authors suggest that infants understand the goal and the restriction and so when the hands are occupied they emulate because they assume that the demonstrator did not follow the "obvious" solution because of the restrictions. Notice that, according to Fig. 1.2, using the head corresponds to *imitation* while using the hand corresponds to (goal) *emulation*.

We applied our model of social learning to an abstracted version of this experiment, evaluating the dependence of the behavior by the learner on the parameters λ_A, λ_E and λ_I in two distinct experiments. In the first experiment, we fixed the weight assigned to the baseline preferences (*i.e.,* we set $\lambda_A = 0.2$) and observed how the behavior changed as as λ_I goes from 0 to 1 (*i.e.,* as the learner increasingly adheres to the inferred goal of the demonstration). The results are depicted in Figure 10(a). Notice that, when faced with a restricted teacher, the learner switches to an "emulative" behavior much sooner, replicating the results in [54].

On a second experiment, we disregarded the observed effect (*i.e.,* we set $\lambda_E = 0$) and observed how the behavior of the learner changes as it assigns more importance to the demonstration and focuses less on its baseline preferences (*i.e.,* as λ_I goes from 0 to 1). The results are depicted in Figure 10(b). Notice that, in this test, we set λ_E to zero, which means that the agent is not explicitly considering the observed effect. However, when combining its own interests with the observed demonstration (that includes goals, actions and effects), the learner tends to *replicate the observed effect* and disregard the observed actions, thus displaying emulative behavior. This is particularly evident in the situation of a restricted teacher.

We emphasize that the difference in behavior between the restricted and non-restricted teacher is due only to the *perceived difference on the ability of the teacher to interact with the environment*. We refer to [79] for further details.



(a) (b)

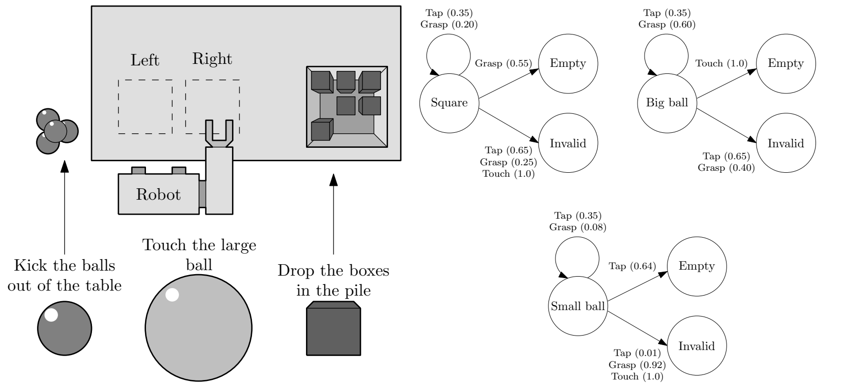
Fig. 1.10: Percentages of replication of demonstrated action. (a) Percentage of runs in which the learner replicates the demonstrated use of head. Whenever the action was not performed with the head, it was performed with the hand. (b) Rates of occurrence of the different actions. When none of the two indicated actions is performed, no action is performed. In both plots, each bar corresponds to a trial of $2,000$ independent runs.

Robot Learning by Imitation

We now present an application of our imitation learning model in a sequential task using BALTAZAR [78]. To test the imitation learning model in the robot we considered a simple recycling game, where the robot must separate different objects according to their shape (Figure 1.11). We set $\lambda_E = \lambda_A = 0$ and used only the imitation module to estimate the intention behind the demonstration. In front of the robot are two slots (Left and Right) where 3 types of objects can be placed: Large Balls, Small Balls and Boxes. The boxes should be dropped in a corresponding container and the small balls should be tapped out of the table. The large balls should be touched upon, since the robot is not able to efficiently manipulate them. Every time a large ball is touched, it is removed from the table by an external user. Therefore, the robot has available a total of 6 possible actions: Touch Left (TcL), Touch Right (ThR), Tap Left (TpL), Tap Right (TpR), Grasp Left (GrL) and Grasp Right (GrR).

For the description of the task at hand, we considered a state-space consisting of 17 possible states. Of these, 16 correspond to the possible combinations of objects in the two slots (including empty slots). The 17th state is an invalid state that accounts for the situations where the robot's actions do not succeed (for example, when the robot drops the ball in an invalid position in the middle of the table).

We first provided the robot with an error-free demonstration of the optimal behavior rule. As expected, the robot was successfully able to reconstruct the optimal policy. We also observed the learned behavior when the robot was provided with *two* different demonstrations, both optimal. The results are described in Table 1.4. Each state is represented as a pair (S_1, S_2) where each S_i can take one of the values "Ball"



(a) (b)

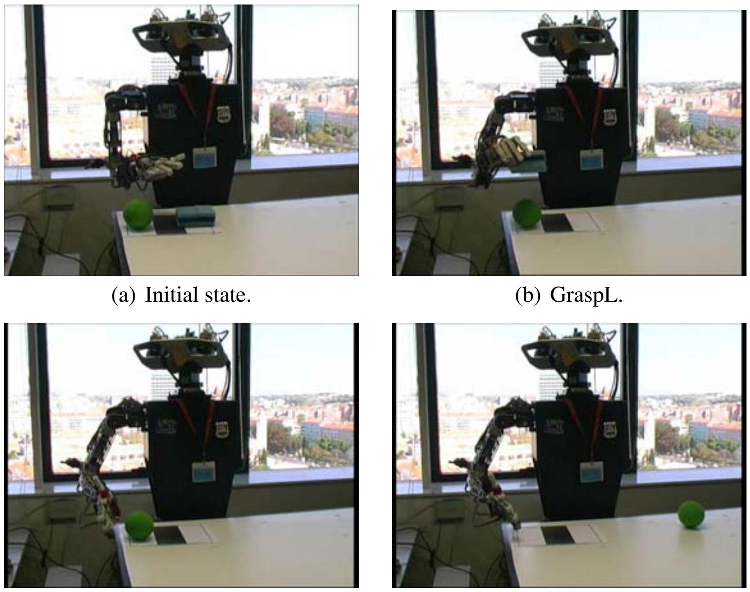
Fig. 1.11: Illustration of the recycling game. (a) The setup. (b) Transition diagrams describing the transitions for each slot/object.

(Big Ball), "ball" (Small Ball), "Box" (Box) or ∅ (empty). The second column of Table 1.4 then lists the observed actions for each state and the third column lists the learned policy. Notice that, as before, the robot was able to reconstruct an optimal policy, by choosing one of the demonstrated actions in those states where different actions were observed.

Table 1.4: Demonstration 1: Error free demonstration. Demonstration 2: Inaccurate and incomplete demonstration, where the boxed cells correspond to the states not demonstrated or in which the demonstration was inaccurate. Columns 3 and 5 present the learned policy for Demo 1 and 2, respectively.

State	**Demo 1**	**Learned Pol.**	**Demo 2**	**Learned Pol.**
(∅, Ball)	TcR	TcR	$\boxed{\text{-}}$	TcR
(∅, Box)	GrR	GrR	GrR	GrR
(∅, ball)	TpR	TpR	TpR	TpR
(Ball, ∅)	TcL	TcL	TcL	TcL
(Ball, Ball)	TcL,TcR	TcL,TcR	$\boxed{\text{GrR}}$	TcL
(Ball, Box)	TcL,GrR	GrR	TcL	TcL
(Ball, ball)	TcL	TcL	TcL	TcL
(Box, ∅)	GrL	GrL	GrL	GrL
(Box, Ball)	GrL,TcR	GrL	GrL	GrL
(Box, Box)	GrL,GrR	GrR	GrL	GrL
(Box, ball)	GrL	GrL	GrL	GrL
(ball, ∅)	TpL	TpL	TpL	TpL
(ball, ball)	TpL,TcR	TpL	TpL	TpL
(ball, Box)	TpL,GrR	GrR	TpL	TpL
(ball, ball)	TpL	TpL	TpL	TpL

We then provided the robot with an *incomplete and inaccurate* demonstration. As seen in Table 1.4, the action at state (∅, Ball) was never demonstrated and the action at state (Ball, Ball) was *wrong*. The last column of Table 1.4 shows the learned policy. Notice that in this particular case the robot was able to recover the *correct policy*, even with an incomplete and inaccurate demonstration.



(a) Initial state. (b) GraspL.

(c) TapR. (d) Final state.

Fig. 1.12: Execution of the learned policy in state (Box, SBall).

In Figure 1.12 we illustrate the execution of the optimal learned policy for the initial state (Box, SBall).[5]

To assess the sensitivity of the imitation learning module to the action recognition errors, we tested the learning algorithm for different error recognition rates. For each error rate, we ran 100 trials. Each trial consists of 45 state-action pairs, corresponding to three optimal policies. The obtained results are depicted in Figure 1.13.

As expected, the error in the learned policy increases as the number of wrongly interpreted actions increases. Notice, however, that for small error rates ($\leq 15\%$) the robot is still able to recover the demonstrated policy with an error of only 1%. In particular, if we take into account the fact that the error rates of the action recognition method used by the robot are between 10% and 15%, the results in Figure 1.13 guarantee a high probability of accurately recovering the optimal policy.

We conclude by remarking that a more sophisticated model can be used in which observation noise is taken into account. This may allow more insensitivity to the

[5] For videos showing additional experiences see http://vislab.isr.ist.utl.pt/baltazar/demos/

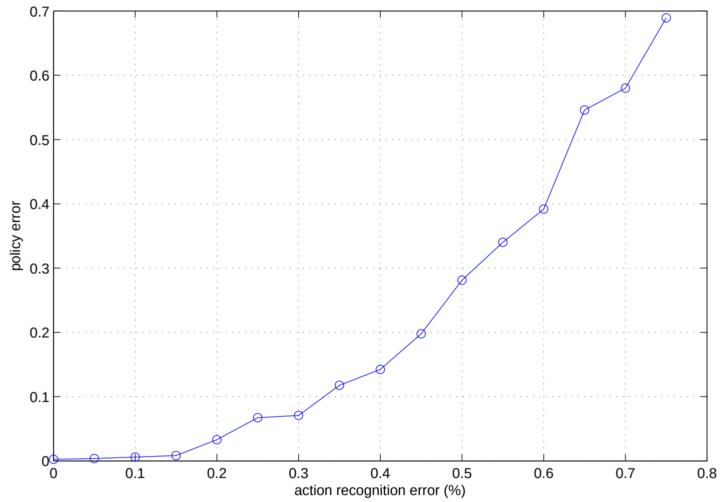


Fig. 1.13: Percentage of wrong actions in the learned policy as the action recognition errors increase.

noise, by including it explicit in the inference module that estimates the reward representing the goal of the demonstrator.

1.7 Other Imitation Settings

Imitation learning goes far beyond programming a single robot to perform a task, and has been used in many other settings.

For example, in [120], learning from demonstrated behavior somewhat resembles transfer learning, in which an agent observes demonstrations in different scenarios and uses this knowledge to recover a reward function that can be used in yet other scenarios. This problem is addressed as a max-margin structured learning problem to recover a reward function from a set of demonstrations. In order to simplify the problem and to leverage fast solution methods, the paper formulates the problem as a non-linear (non-differentiable) unconstrained optimization problem, that is tackled using subgradient techniques that rely on the solution structure of the (embedded) constraints.

In [45,97], robots able to imitate have been used to interact with autistic children. On a related application, the Infanoid project [70, 72] deals with gesture imitation [71], interaction with people [73], and joint attention [98]. The robots in this project are used in human-robot interaction scenarios with particular emphasis on people with special needs. Although the results seem promising, and in short term people seem to react well, care must be taken in ensuring that the robots are used to promote socialization with other people, and not a stronger focus on the machine itself [121].

Some authors have also addressed imitation in multiagent scenarios, considering multiple demonstrators [132], multiple learners [30] and human-robot joint work [42]. In the presence of multiple demonstrators, these may be performing different tasks and the agent must actively select which one to follow. In [132], this observation led the authors to call their approach *active imitation*. Active learning approaches applied to imitation are very recent [81, 132]. Typically, the burden of selecting informative demonstrations has been completely on the side of the mentor. Active learning approaches endow the learner with the power to select which demonstrations the mentor should perform. Several criteria have been proposed: game theoretic approaches [132], entropy [81] and risk minimization [40].

Computational models of imitation have also been proposed to understand biology by synthesis. Examples include models of language and culture [3, 15], curiosity drives resulting in imitation behaviors [68], behavior switching in children and chimpanzees [79]. There have also been studies of imitation deficits relying on models of brain connections [110, 124]

We also note that there are other social learning mechanisms that fall outside the "imitation realm" in biological research. Often imitation is seen as a fundamental mental process for acquiring complex social skills but other mechanisms, although cognitively simpler, may have their own evolutionary advantages [89, 104, 105].

1.8 Discussion

In this chapter we presented an overview of imitation learning from two different perspectives. First, we discussed evidence coming from research in biology and neurophysiology and identified several cognitive processes required for imitation. We particularly emphasized two ideas that have a direct impact on imitation: 1) the action-perception coupling mechanisms involved, for instance, in the mirror system; and 2) the different social learning mechanisms found in infants and primates, not all of which can be classified as "true imitation". We also pointed out the importance of contextual cues to drive these mechanisms and to interpret the demonstration. As the bottom line, we stress that social learning happens at different levels of abstraction, from pure mimicry to more abstract cognitive processes.

Taking this evidence into account, we then reviewed imitation in artificial systems *i.e.,* methods to learn from a demonstration. As a result of the advances on this topic, there is currently a vast amount of work. Following the three main challenges identified in [151], we surveyed several methods that take advantage of the information provided by a demonstration in many different ways: as initial conditions for self-exploration methods (including planning), as exploration strategies, as data from which to infer world models, or as data to infer what the task is. These methods are being used with many different goals in mind, either to speed up robot programming, to develop more intuitive human-robot interfaces or to study cognitive and social skills of humans. In addition to this, we provide experimental results of increasing abstract imitation behaviors, from motor resonance to task learning.

An open question, and one we only addressed in an empirical way, is how all these methods are related or could be combined to achieve complex imitation behaviors. Indeed, different approaches usually tailor their formalisms to a particular domain of application. It is still not clear how different they are and if they can be used in several domains. If it becomes clear that they are indeed different, it would be interesting to understand how to switch between each mechanism, and eventually understand if there is a parallel in the brain.

Another important aspect that requires further research is related to perception. Although this is not specific to imitation, it plays a crucial role when interpreting the demonstration. Currently, robots are still unable to properly extract relevant information and perceive contextual restrictions from a general purpose demonstration. Due to this difficulty, having a robot companion that learns by imitation is still beyond our technological and scientific knowledge.

Nevertheless, most of these problems can be somewhat reduced when robot programming is conducted by skilled people that can handle more intrusive sensory modalities. In the chapter, we analyzed more in detail an alternative path to imitation which relies on previously learned models for the robot and the environment that help the understanding of the demonstration.

Using prior knowledge may simplify the interpretation of the demonstration, but requires the acquisition of good motor, perceptual and task descriptions. Most approaches consider predefined feature spaces for each of these entities. When considering object-related tasks, this problem is even more important than when addressing pure motor tasks. A given world state may be described in terms of object locations, object-object relations, robot-object relations, among many others, but it is not easy to automatically extract, or choose, among the correct representations.

Finally, a recent trend in imitation learning tries to learn task abstractions from demonstrations. The rationale is that, once the robot has understood a task in an abstract manner, it can easily reason about the contextual cues that drive imitation behaviors, include them in future plans and, as a result, generalize better to other situations. In our experimental results, we showed how to combine multiple task descriptions to switch between different social learning behaviors through a biologically-inspired computational imitation model. Also, having such a representation opens the door to more general cognitive imitation architectures for robots.

Future applications of imitation will handle human-robot collaboration in cooperative settings (with several robots or people) and active strategies for interaction with the demonstrator.

We conclude this review by stating our belief that imitation and learning by demonstration will become one of the capabilities that future fully autonomous robots will extensively use, both to acquire new skills and to adapt to new situations in an efficient manner. The path to this objective is still full of exciting research challenges and fascinating links to the way we, humans, develop and learn.

Acknowledgements This work was (partially) supported by: the Portuguese Fundação para a Ciência e a Tecnologia, the Information and Communications Technologies Institute (under the Carnegie Mellon-Portugal Program), the Programa Operacional Sociedade de Conhecimento (POS_C) that includes FEDER funds, and the project PTDC/EEA-ACR/70174/2006), and by the EU Projects (IST-004370) RobotCub and (EU-FP6-NEST-5010) Contact.

References

1. Abbeel, P., Ng, A.Y.: Apprenticeship learning via inverse reinforcement learning. In: Proceedings of the 21st International Conference on Machine Learning (ICML'04), pp. 1–8 (2004)
2. Alissandrakis, A., Nehaniv, C.L., Dautenhahn, K.: Imitating with alice: Learning to imitate corresponding actions across dissimilar embodiments. IEEE Transactions on Systems, Man, & Cybernetics, Part A: Systems and Humans **32**(4), 482–496 (2002)
3. Alissandrakis, A., Nehaniv, C.L., Dautenhahn, K.: Towards robot cultures? - learning to imitate in a robotic arm test-bed with dissimilarly embodied agents. Interaction Studies **5**(1), 3–44 (2004)
4. Alissandrakis, A., Nehaniv, C.L., Dautenhahn, K.: Action, state and effect metrics for robot imitation. In: 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 06), pp. 232–237. Hatfield, United Kingdom (2006)
5. Arbib, M.A., Billard, A., Iacoboni, M., Oztop, E.: Synthetic brain imaging: grasping, mirror neurons and imitation. Neural Networks **13**, 975–997 (2000)
6. Argall, B., Chernova, S., Veloso, M.: A survey of robot learning from demonstration. Robotics and Autonomous Systems **57**(5), 469–483 (2009)
7. Asada, M., Ogino, M., Matsuyama, S., Ooga, J.: Imitation learning based on visuo-somatic mapping. In: O.K. Marcelo, H. Ang (eds.) 9th Int. Symp. Exp. Robot.,, vol. 21, pp. 269–278. Springer-Verlag, Berlin, Germany (2006)
8. Asada, M., Yoshikawa, Y., Hosoda, K.: Learning by observation without three-dimensional reconstruction. In: Intelligent Autonomous Systems (IAS-6), pp. 555–560 (2000)
9. Atkeson, C.G., Schaal, S.: Robot learning from demonstration. In: 14th International Conference on Machine Learning, pp. 12–20. Morgan Kaufmann (1997)
10. Baker, C.L., Tenenbaum, J.B., Saxe, R.R.: Bayesian models of human action understanding. Advances in Neural Information Processing Systems **18** (2006)
11. Bakker, P., Kuniyoshi, Y.: Robot see, robot do : An overview of robot imitation. In: AISB'96: Workshop on Learning in Robots and Animals, pp. 3–11 (1996)
12. Bandera, J.P., Marfil, R., Molina-Tanco, L., Rodríguez, J.A., Bandera, A., Sandoval, F.: Robot Learning by Active Imitation, pp. 519–544. ARS Publ. (2007)
13. Bekkering, H., Wohlschläger, A., Gattis, M.: Imitation of gestures in children is goal-directed. Quarterly J. Experimental Psychology **53A**, 153–164 (2000)
14. Bellagamba, F., Tomasello, M.: Re-enacting intended acts: Comparing 12- and 18-month-olds. Infant Behavior and Development **22**(2), 277–282 (1999)
15. Billard, A.: Imitation: A means to enhance learning of a synthetic proto-language in an autonomous robot. In: Imitation in Animals and Artifacs, pp. 281–311. MIT Press (1999)
16. Billard, A., Calinon, S., Dillman, R., Schaal, S.: Handbook of Robotics, chap. Chapter 59: Robot Programming by Demonstration. Springer (2007)
17. Billard, A., Epars, Y., Calinon, S., Cheng, G., Schaal, S.: Discovering optimal imitation strategies. Robotics and Autonomous Systems **47:2-3** (2004)
18. Boucenna, S., Gaussier, P., Andry, P.: What should be taught first: the emotional expression or the face? In: 8th International conference on Epigenetic Robotics. Brighton, UK (2008)
19. Brass, M., Schmitt, R.M., Spengler, S., Gergely, G.: Investigating action understanding: Inferential processes versus action simulation. Current Biology **17**(24), 2117–2121 (2007)
20. Breazeal, C.: Imitation as social exchange between humans and robots. In: AISB Symp. Imitation in Animals and Artifacts, pp. 96–104 (1999)

21. Brugger, A., Lariviere, L.A., Mumme, D.L., Bushnell, E.W.: Doing the right thing: Infants' selection of actions to imitate from observed event sequences. Child Development **78**(3), 806–824 (2007)
22. Bruner, J.: Nature and use of immaturity. American Psychologist **27**, 687–708 (1972)
23. Byrne, R.W.: Imitation without intentionality. using string parsing to copy the organization of behaviour. Animal Cognition **2**, 63–72 (1999)
24. Calinon, S., Billard, A.: Learning of gestures by imitation in a humanoid robot. In: Imitation and Social Learning in Robots, Humans and Animals: Behavioural, Social and Communicative Dimensions, K. Dautenhahn and C.L. Nehaniv edn., pp. 153–177. Cambridge University Press (2007)
25. Call, J., Carpenter, M.: Three sources of information in social learning. In: Imitation in animals and artifacts. MIT Press, Cambridge, MA, USA (2002)
26. Cantin-Martinez, R., Lopes, M., Melo, F.: Inverse reinforcement learning with noisy observations. Tech. rep., Institute for Systems and Robotics, Lisbon, Portugal (2009)
27. Cantin-Martinez, R., Lopes, M., Montesano, L.: Active body schema learning. Tech. rep., Institute for Systems and Robotics, Lisbon, Portugal (2009)
28. Carpenter, M., Call, J., Tomasello, M.: Twelve- and 18-month-olds copy actions in terms of goals. Developmental Science **1**(8), F13–F20 (2005)
29. Chalodhorn, R., Rao, R.P.N.: From Motor to Interaction Learning in Robots, chap. Learning to imitate human actions through Eigenposes. Springer (2009)
30. Chernova, S., Veloso, M.: Teaching collaborative multi-robot tasks through demonstration. In: IEEE-RAS International Conference on Humanoid Robots. Korea (2008)
31. Chernova, S., Veloso, M.: Interactive policy learning through confidence-based autonomy. J. Artificial Intelligence Research **34**, 1–25 (2009)
32. Cos-Aguilera, I., Cañamero, L., Hayes, G.: Using a SOFM to learn object affordances. In: Workshop of Physical Agents (WAF). Girona, Spain (2004)
33. Csibra, G., Gergely, G.: "Obsessed with goals": Functions and mechanisms of teleological interpretation of actions in humans. Acta Psychologica **124**, 60–78 (2007)
34. Demiris, J., Hayes, G.: Imitative learning mechanisms in robots and humans. In: European Workshop on Learning Robots, pp. 9–16 (1996)
35. Demiris, J., Rougeaux, S., Hayes, G.M., Berthouze, L., Kuniyoshi, Y.: Deferred imitation of human head movements by an active stereo vision head. In: 6th IEEE Int. Workshop on Robot Human Communication, pp. 88–93 (1997)
36. Demiris, Y., Dearden, A.: From motor babbling to hierarchical learning by imitation: a robot developmental pathway. In: EPIROB-2005, pp. 31–37. Japan (2005)
37. Demiris, Y., Hayes, G.: Imitation as a dual-route process featuring predictive and learning components: a biologically-plausible computational model. In: K. Dautenhahn, C. Nehaniv (eds.) Imitation in Animals and Artifacts. MIT Press, Cambridge, MA, USA (2002)
38. Demiris, Y., Khadhouri, B.: Hierarchical attentive multiple models for execution and recognition of actions. Robotics and Autonomous Systems **54**, 361–369 (2006)
39. Detry, R., Baseski, E., , M.P., Touati, Y., Kruger, N., Kroemer, O., Peters, J., Piater, J.: Learning object-specific grasp affordance densities. In: IEEE 8TH International Conference on Development and Learning (2009)
40. Doshi, F., Pineau, J., Roy, N.: Reinforcement learning with limited reinforcement: using bayes risk for active learning in pomdps. In: Proceedings of the 25th international conference on Machine learning (ICML'08), pp. 256–263 (2008)
41. D'Souza, A., Vijayakumar, S., Schaal, S.: Learning inverse kinematics. In: International Conference on Intelligent Robots and Systems, pp. 298–303. Hawaii, USA (2001)
42. Erlhagen, W., Mukovskiy, A., Chersi, F., Bicho, E.: On the development of intention understanding for joint action tasks. In: 6th IEEE International Conference on Development and Learning (ICDL 2007). London, UK (2007)
43. Erlhagen, W., Mukovsky, A., Bicho, E.: A dynamic model for action understanding and goal-directed imitation. Brain Research **1083**, 174–188 (2006)

44. Erlhagen, W., Mukovsky, A., Bicho, E., Panin, G., Kiss, C., Knoll, A., van Schie, H., Bekkering, H.: Goal-directed imitation in robots: a bio-inspired approach to action understanding and skill learning. Robotics and Autonomous Systems **54**(5), 353–360 (2006)
45. Field, T., Field, T., Sanders, C., Nadel, J.: Children with autism become more social after repeated imitation sessions. Autism pp. 317–324 (2001)
46. Fitzpatrick, P., Metta, G., Natale, L., Rao, S., Sandini., G.: Learning about objects through action: Initial steps towards artificial cognition. In: IEEE International Conference on Robotics and Automation. Taipei, Taiwan (2003)
47. Fogassi, L., Gallese, V., Buccino, G., Craighero, L., Fadiga, L., Rizzolatti, G.: Cortical mechanism for the visual guidance of hand grasping movements in the monkey: A reversible inactivation study. Brain **124**(3), 571–586 (2001)
48. Fritz, G., Paletta, L., Breithaupt, R., Rome, E., Dorffner, G.: Learning predictive features in affordance based robotic perception systems. In: IEEE/RSJ International Conference on Intelligent Robots and Systems. Beijing, China (2006)
49. Furse, E.: A model of imitation learning of algorithms from worked examples. Cybernetics and Systems **32**, 121–154 (2001)
50. Galantucci, B., Fowler, C.A., Turvey, M.T.: The motor theory of speech perception reviewed. Psychonomic Bulletin & Review **13**(3), 361–377 (2006)
51. Gallese, V., Fadiga, L., Fogassi, L., Rizolaatti, G.: Action recognition in the premotor cortex. Brain **119**, 593–609 (1996)
52. Gardner, M.: Imitation and egocentric perspective transformation. In: Virtual Poster associated with Perspectives on Imitation Conference (2002)
53. Gaussier, P., S. Moga, J.B., Quoy, M.: From perception-action loops to imitation processes: A bottom-up approach of learning by imitation. Applied Artificial Intelligence: An Internation Journal **1**(7) (1997)
54. Gergely, G., Bekkering, H., Király, I.: Rational imitation in preverbal infants. Nature **415**, 755 (2002)
55. Gibson, J.J.: The Ecological Approach to Visual Perception. Houghton Mifflin, Boston (1979)
56. Gopnik, A.: Cells that read minds? what the myth of mirror neurons gets wrong about the human brain. Slate: Brains: A special issue on neuroscience and neuroculture. (2007)
57. Grimes, D.B., Rashid, D.R., Rao, R.P.: Learning nonparametric models for probabilistic imitation. In: Advances in NIPS (2007)
58. Hart, S., Grupen, R., Jensen, D.: A relational representation for procedural task knowledge. In: Proceedings of the 2005 American Association for Artificial Intelligence (AAAI) Conference. Pittsburgh, USA (2005)
59. Hayes, G., Demiris, J.: A robot controller using learning by imitation. In: Int. Symp. Intelligent Robotic Systems, pp. 198–204 (1994)
60. Heckerman, D., Geiger, D., Chickering, M.: Learning bayesian networks: the combination of knowledge and statistical data. Machine Learning (1995)
61. Hersch, M., Sauser, E., Billard, A.: Online learning of the body schema. International Journal of Humanoid Robotics **5**(2), 161–181 (2008)
62. Horner, V., Whiten, A.: Causal knowledge and imitation/emulation switching in chimpanzees (Pan troglodytes) and children (Homo sapiens). Animal Cognition **8**, 164–181 (2005)
63. Hovland, G., Sikka, P., McCarragher, B.: Skill acquisition from human demonstration using a hidden markov model. In: IEEE International Conference on Robotics and Automation, pp. 2706–2711. Minneapolis, MN (1996)
64. Jansen, B., Belpaeme, T.: A model for inferring the intention in imitation tasks. In: 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN06). Hatfield, UK (2006)
65. Jenkins, O., Bodenheimer, R., Peters, R.: Manipulation manifolds: Explorations into uncovering manifolds in sensory-motor spaces. In: International Conference on Development and Learning (ICDL 2006). Bloomington, IN, USA (2006)

66. Jenkins, O.C., Matarić, M.J., Weber, S.: Primitive-based movement classification for humanoid imitation. In: IEEE International Conference on Humanoid Robots (Humanoids 2000) (2000)
67. Johnson, S., Booth, A., O'Hearn, K.: Inferring the goals of a nonhuman agent. Cognitive Development **16**(1), 637–656 (2001)
68. Kaplan, F., Oudeyer, P.Y.: The progress drive hypothesis: an interpretation of early imitation. In: Models and Mechanims of Imitation and Social Learning: Behavioural, Social and Communication Dimensions, pp. 361–377. Cambridge University Press., Cambridge (2007)
69. Konczak, J., Meeuwsen, H., Cress, M.: Changing affordances in stair climbing: The perception of maximum climbability in young and older adults. Journal of Experimental Psychology: Human Perception & Performance **19**, 691–7 (1992)
70. Kozima, H.: Infanoid: An experimental tool for developmental psycho-robotics. In: International Workshop on Developmental Study. Tokyo, Japan (2000)
71. Kozima, H., Nakagawa, C., Yano, H.: Emergence of imitation mediated by objects. In: 2nd Int. Workshop on Epigenetic Robotics (2002)
72. Kozima, H., Nakagawa, C., Yano, H.: Attention coupling as a prerequisite for social interaction. In: IEEE International Workshop on Robot and Human Interactive Communication. Millbrae, USA (2003)
73. Kozima, H., Yano, H.: Designing a robot for contingency-detection game. In: International Workshop on Robotic and Virtual Agents in Autism Therapy. Hertfordshire, England (2001)
74. Kulić, D., Nakamura, Y.: From Motor to Interaction Learning in Robots, chap. Incremental Learning of Full Body Motion Primitives. Springer (2009)
75. Kuniyoshi, Y., Inaba, M., Inoue, H.: Learning by watching: Extracting reusable task knowledge from visual observation of human performance. IEEE Trans. on Robotics and Automation **10(6)**, 799–822 (1994)
76. Kuniyoshi, Y., Yorozu, Y., Inaba, M., Inoue, H.: From visuo-motor self learning to early imitation-a neural architecture for humanoid learning. In: IEEE Int. Conf. Robotics and Automation, vol. 3, pp. 3132–3139 (2003)
77. Liberman, A.M., Mattingly, I.G.: The motor theory of speech perception revised. Cognition **21**, 1–36 (1985)
78. Lopes, M., Beira, R., Praça, M., Santos-Victor, J.: An anthropomorphic robot torso for imitation: design and experiments. In: International Conference on Intelligent Robots and Systems. Sendai, Japan (2004)
79. Lopes, M., Melo, F., Kenward, B., Santos-Victor, J.: A computational model of social-learning mechanisms. Adaptive Behavior (to be published)
80. Lopes, M., Melo, F.S., Montesano, L.: Affordance-based imitation learning in robots. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1015–1021. USA (2007)
81. Lopes, M., Melo, F.S., Montesano, L.: Active learning for reward estimation in inverse reinforcement learning. In: European Conference on Machine Learning (ECML/PKDD). Bled, Slovenia (2009)
82. Lopes, M., Santos-Victor, J.: Visual Transformations in Gesture Imitation: What you see is what you do. In: IEEE Int. Conf. Robotics and Automation (2003)
83. Lopes, M., Santos-Victor, J.: Visual learning by imitation with motor representations. IEEE Trans. Systems, Man, and Cybernetics - Part B: Cybernetics **35**(3) (2005)
84. Lopes, M., Santos-Victor, J.: A developmental roadmap for learning by imitation in robots. IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics **37**(2) (2007)
85. Lyons, D.E., Young, A.G., Keil, F.C.: The hidden structure of overimitation. Proceedings of the National Academy of Sciences **104**(50), 19,751–19,756 (2005)
86. Maistros, G., Marom, Y., Hayes, G.: Perception-action coupling via imitation and attention. In: AAAI Fall Symp. Anchoring Symbols to Sensor Data in Single and Multiple Robot Systems (2001)
87. Matarić, M.J.: Sensory-motor primitives as a basis for learning by imitation:linking perception to action and biology to robotics. In: Imitation in Animals and Artifacts. MIT Press (2002)

88. McGuigan, N., Whiten, A., Flynn, E., Horner, V.: Imitation of causally opaque versus causally transparent tool use by 3- and 5-year-old children. Cognitive Development **22**, 353–364 (2007)
89. Melo, F., Lopes, M., Santos-Victor, J., Ribeiro, M.I.: A unified framework for imitation-like behaviors. In: 4th International Symposium on Imitation in Animals and Artifacts. Newcastle, UK (2007)
90. Meltzoff, A.N.: Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. Developmental Psychology **24**(4), 470–476 (1988)
91. Meltzoff, A.N.: Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. Developmental Psychology **31**(5), 838–850 (1995)
92. Montesano, L., Lopes, M.: Learning grasping affordances from local visual descriptors. In: IEEE 8TH International Conference on Development and Learning. China (2009)
93. Montesano, L., Lopes, M., Bernardino, A., Santos-Victor, J.: Affordances, development and imitation. In: IEEE - International Conference on Development and Learning. London, UK (2007)
94. Montesano, L., Lopes, M., Bernardino, A., Santos-Victor, J.: Learning object affordances: From sensory–motor coordination to imitation. IEEE Transactions on Robotics **24**(1), 15–26 (2008)
95. Montesano, L., Lopes, M., Bernardino, A., Santos-Victor, J.: Learning object affordances: From sensory–motor coordination to imitation. IEEE Transactions on Robotics **28**(1) (2008)
96. Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., Rizzolatti, G.: Object representation in the ventral premotor cortex (area f5) of the monkey. Journal of Neurophysiology **78**(4), 2226–2230 (1997)
97. Nadel, J.: Imitation and imitation recognition: functional use in preverbal infants and nonverbal children with autism. In: The imitative mind. Cambridge University Press (2002)
98. Nagai, Y., Asada, M., Hosoda, K.: A developmental approach accelerates learning of joint attention. In: International Conference on Development and Learning (2002)
99. Nakaoka, S., Nakazawa, A., Yokoi, K., Hirukawa, H., Ikeuchi, K.: Generating whole body motions for a biped humanoid robot from captured human dances,. In: IEEE International Conference on Robotics and Automation. Taipei, Taiwan (2003)
100. Nehaniv, C., Dautenhahn, K.: Mapping between dissimilar bodies: Affordances and the algebraic foundations of imitation. In: European Workshop on Learning Robots (1998)
101. Nehaniv, C.L., Dautenhahn, K.: Like me? - measures of correspondence and imitation. Cybernetics and Systems **32**, 11–51 (2001)
102. Neu, G., Szepesvri, C.: Apprenticeship learning using inverse reinforcement learning and gradient methods. In: Uncertainty in Artificial Intelligence (UAI), pp. 295–302 (2007)
103. Nielsen, M.: Copying actions and copying outcomes: Social learning through the second year. Developmental Psychology **42**(3), 555–565 (2006)
104. Noble, J., Franks, D.W.: Social learning mechanisms compared in a simple environment. In: Artificial Life VIII: Proceedings of the Eighth International Conference on the Simulation and Synthesis of Living Systems, pp. 379–385. MIT Press, Cambridge, MA, USA (2002)
105. Oudeyer, P.Y.: The Social Formation of Acoustic Codes with "Something Simpler". In: K. Dautenham, C. Nehaniv (eds.) Second International Workshop on Imitation in Animals and Artefacts, AISB'03. Aberystwyth, Wales (2003)
106. Oztop, E., Kawato, M., Arbib, M.: Mirror neurons and imitation: A computationally guided review. Neural Networks **19**(3), 254–271 (2006)
107. Pelleg, D., Moore, A.W.: X-means: Extending k-means with efficient estimation of the number of clusters. In: International Conference on Machine Learning. San Francisco, CA, USA (2000)
108. Peters, J., Schaal, S.: Reinforcement learning of motor skills with policy gradients. Neural Networks pp. 682–697 (2008)
109. Peters, J., Vijayakumar, S., Schaal, S.: Natural Actor-Critic. In: Proc. 16th European Conf. Machine Learning, pp. 280–291 (2005)
110. Petreska, B., Billard, A.: A Neurocomputational Model of an Imitation Deficit following Brain Lesion. In: ICANN'2006, vol. vol. 4131, pp. pp. 770–779 (2006)

111. Pomplun, M., Matarić, M.J.: Evaluation metrics and results of human arm movement imitation. In: IEEE-RAS Int. Conf. Humanoid Robotics (2000)
112. Price, B.: Accelerating reinforcement learning with imitation. Ph.D. thesis, University of British Columbia (2003)
113. Price, B., Boutilier, C.: Implicit imitation in multiagent reinforcement learning. In: Proc. ICML, pp. 325–334 (1999)
114. Price, B., Boutilier, C.: Accelerating reinforcement learning through implicit imitation. J. Artificial Intelligence Research **19**, 569–629 (2003)
115. Puterman, M.: Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley & Sons, Inc. (1994)
116. Ramachandran, D., Amir, E.: Bayesian inverse reinforcement learning. In: 20th Int. Joint Conf. Artificial Intelligence. India (2007)
117. Ramachandran, V.: Mirror neurons and imitation learning as the driving force behind the great leap forward in human evolution. Edge **69** (2000)
118. Range, F., Viranyi, Z., Huber, L.: Selective imitation in domestic dogs. Current Biology **17**(10), 868–872 (2007)
119. Rao, R., Shon, A., Meltzoff, A.: Imitation and social learning in robots, humans, and animals, chap. A Bayesian model of imitation in infants and robots. Cambridge University Press (2007)
120. Ratliff, N., Bagnell, J., Zinkevich, M.: Maximum margin planning. In: Proc. 23rd Int. Conf. Machine Learning, pp. 729–736 (2006)
121. Robins, B., Dautenhahn, K., Dubowski, J.: Robots as isolators or mediators for children with autism? a cautionary tale. In: AISB05: Social Intelligence and Interaction in Animals, Robots and Agents, pp. 12–15. Hatfield, UK (2005)
122. Sahin, E., Cakmak, M., Dogar, M., Ugur, E., Ucoluk, G.: To afford or not to afford: A new formalization of affordances towards affordance-based robot control. Adaptive Behavior **15**(5), 447–472 (2007)
123. Sauser, E., Billard, A.: View sensitive cells as a neural basis for the representation of others in a self-centered frame of reference. In: 3rd Int. Symp. Imitation in Animals & Artifacts (2005)
124. Sauser, E., Billard, A.: Parallel and Distributed Neural Models of the Ideomotor Principle: An Investigation of Imitative Cortical Pathways. Neural Networks **19(3)**, 285–298 (2006)
125. Saxena, A., Driemeyer, J., Ng, A.Y.: Robotic grasping of novel objects using vision. International Journal of Robotics Research (IJRR) (2008)
126. Saxena, A., Wong, L., Ng, A.Y.: Learning grasp strategies with partial shape information. In: AAAI (2008)
127. Schaal, S.: Is imitation learning the route to humanoid robots. Trends in Cognitive Sciences **3(6)**, 233–242 (1999)
128. Schaal, S., Ijspeert, A., Billard, A.: Computational approaches to motor learning by imitation. Phil. Trans. of the Royal Society of London: Series B, Biological Sciences **358**(1431) (2003)
129. Schaal, S., Peters, J., Nakanishi, J., Ijspeert, A.: Learning movement primitives. In: International Symposium on Robotics Research (ISRR2003) (2003)
130. Shon, A., Grochow, K., Rao, R.: Robotic imitation from human motion capture using gaussian processes. In: IEEE/RAS International Conference on Humanoid Robots (Humanoids), (2005)
131. Shon, A.P., Joshua, Storz, J., Rao, R.P.N.: Towards a real-time bayesian imitation system for a humanoid robot. In: ICRA'07 (2007)
132. Shon, A.P., Verma, D., Rao, R.P.N.: Active imitation learning. In: AAAI (2007)
133. Stoytchev, A.: Behavior-grounded representation of tool affordances. In: International Conference on Robotics and Automation. Barcelona, Spain (2005)
134. Stulp, F., Fedrizzi, A., Beetz, M.: Learning and performing place-based mobile manipulation. In: IEEE 8TH International Conference on Development and Learning (2009)
135. Sturm, J., Plagemann, C., Burgard, W.: Adaptive body scheme models for robust robotic manipulation. In: RSS - Robotics Science and Systems IV. Zurich, Switzerland (2008)

136. Syed, U., Schapire, R., Bowling, M.: Apprenticeship learning using linear programming. In: Proceedings of the 25th International Conference on Machine Learning (ICML'08), pp. 1032–1039 (2008)
137. Tani, J., Ito, M., Sugita, Y.: Self-organization of distributedly represented multiple behavior schemata in a mirror system: Reviews of robot experiments using rnnpb. Neural Networks **17**(8/9), 1273–1289 (2004)
138. Tennie, C., Call, J., Tomasello, M.: Push or pull: Imitation vs. emulation in great apes and human children. Ethology **112**(12), 1159–1169 (2006)
139. Thurau, C., Paczian, T., Sagerer, G.: Bayesian imitation learning in game characters. Int. J. Intelligent Systems Technologies and Applications **2**(2/3) (2007)
140. Tomasello, M., Kruger, A.C., Ratner, H.H.: Cultural learning. Behavioral and Brain Sciences **16**(3), 495–511 (1993)
141. Turvey, M., Shockley, K., Carello, C.: Affordance, proper function, and the physical basis of perceived heaviness. Cognition **73** (1999)
142. Verma, D., Rao, R.: Goal-based imitation as probabilistic inference over graphical models. In: Advances in NIPS 18 (2006)
143. Visalberghi, E., Fragaszy, D.: "Do monkeys ape?": ten years after. In: Imitation in Animals and Artifacts. MIT Press, Cambridge, MA, USA (2002)
144. Wang, J.M., Fleet, D.J., Hertzmann, A.: Gaussian process dynamical models for human motion. Trans. PAM pp. 283–298 (2008)
145. Want, S., Harris, P.: Learning from other people's mistakes: Causal understanding in learning to use a tool. Child Development **72**(2), 431–443 (2001)
146. Want, S.C., Harris, P.L.: How do children ape? Applying concepts from the study of non-human primates to the development study of "imitation" in children. Developmental Science **5**(1), 1–13 (2002)
147. Whiten, A., Custance, D., Gomez, J.C., Teixidor, P., Bard, K.A.: Imitative learning of artificial fruit processing in children (Homo sapiens) and chimpanzees (Pan troglodytes). Journal of Comparative Psychology **110**, 3–14 (1996)
148. Whiten, A., Horner, V., Litchfield, C.A., Marshall-Pescini, S.: How do apes ape? Learning & Behavior **32**(1), 36–52 (2004)
149. Williams, T.G., Rowland, J.J., Lee, M.H.: Teaching from examples in assembly and manipulation of snack food ingredients by robot. In: 2001 IEEE/RSJ, International Conference on Intelligent Robots and Systems, pp. 2300–2305 (2001)
150. Williamson, R.A., Markman, E.M.: Precision of imitation as a function of preschoolers' understanding of the goal of the demonstration. Developmental Psychology **42**(4), 723–731 (2006)
151. Wolpert, D.M., Doya, K., Kawato, M.: A unifying computational framework for motor control and social interaction. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences **358**(1431), 593–602 (2003)
152. Wolpert, D.M., Kawato, M.: Multiple paired forward and inverse models for motor control. Neural Networks **11**(7-8), 1317–1329 (1998)
153. Yang, J., Xu, Y., Chen, C.: Hidden markov model approach to skill learning and its application to telerobotics. IEEE Transations on Robotics and Automation **10**(5), 621–631 (1994)
154. Zentall, T.R.: Imitation in animals: Evidence, function, and mechanisms. Cybernetics and Systems **32**(1), 53–96 (2001)
155. Zhang, J., Rössler, B.: Self-valuing learning and generalization with application in visually guided grasping of complex objects. Robotics and Autonomous Systems **47**, 117–127 (2004)
156. Ziebart, B., Maas, A., Bagnell, J., Dey, A.: Maximum entropy inverse reinforcement learning. In: Proc. 23rd AAAI Conf. Artificial Intelligence, pp. 1433–1438 (2008)