The Correspondence Problem in Social Learning: What Does it Mean for Behaviors to "Match" Anyway?

Chrystopher L. Nehaniv Kerstin Dautenhahn
Adaptive Systems Research Group
University of Hertfordshire
College Lane
Hatfield, Hertfordshire AL10 9AB
United Kingdom

1 Matching and Imitation

The identification of any form of social learning, imitation, copying or mimicry presupposes a notion of *correspondence* between two autonomous agents. Judging whether a behavior has been transmitted socially requires the observer to identify a *mapping* between the demonstrator and the imitator. If the demonstrator and imitator have similar bodies, *e.g.* are animals of the same species, of similar age, and of the same gender, then to a human observer an obvious correspondence is to map the corresponding body parts: left arm of demonstrator maps to left arm of imitator, right eye of demonstrator maps to right eye of imitator, tail of demonstrator maps to tail of imitator. There is also an obvious correspondence of actions: raising the left arm by the model corresponds to raising the left arm by the imitator, production of vocal signals by the model corresponds to the production of acoustically similar ones by the imitator, picking up a fruit by the demonstrator corresponds to picking up a fruit of the same type by the imitator. Furthermore, there is a correspondence in sensory experience: audible sounds, a touch, visible objects and colors, and so on evidently seem to be detected and experienced in similar ways.

What to take as the correspondence seems relatively clear in this case. As humans, we are good at imitating and at recognizing such correspondences. It is also clear that most other animals, robots, and software programs may in fact generally fail to recognize any such correspondences. To judge a produced behavior to be a copy of an observed one, we require at least that it respects some such correspondence. The faithfulness or precision of the behavioral match can obviously vary, and no absolute cutoff or threshold exists defining success as opposed to failure of behavioral matching. But one can study the degree of success using various metrics and measures of correspondence (Nehaniv & Dautenhahn, 2001; also see below).

Moreover, it turns out that the "obvious" correspondences between similar bodies mentioned above are not the only ones possible. Consider a human imitating another one that is facing her: if the demonstrator raises her left arm, should the imitator raise her own left arm? Or should she raise her right, to make a "mirror image" of the demonstrator's actions? If the demonstrator picks up a brush, should an imitator pick up the same brush? Or just another brush of the same type? If the demonstrator opens a container to get at chocolate inside, should the imitator open a similar container in the same way – *e.g.* by unwrapping but not tearing the surrounding paper?, or is it enough just to open the container somehow? The different possible answers to these questions presuppose different correspondences.

If a child watches a teacher solving subtraction problems in arithmetic, and then solves for the first time similar but not identical problems on its own, social learning has occurred. But what type of correspondence is at work here? In China and Japan, the ideographic character for "to imitate" also means "to learn" or "to study". By going through the motions of an algorithm for solving sample problems, students everywhere are able to learn how to solve similar ones, of course without necessarily gaining understanding of why the procedures they have learned work.

In this article, for lack of a better term, we shall use the word "imitator" to refer to any autonomous agent performing a candidate behavioral match. The use of this word here does not entail any particular mechanism of matching or any particular type of social learning. In what follows, we shall describe how different matching phenomena arise depending on the criteria employed in generating the behavior of the imitator. For example, goal emulation, stimulus enhancement, mimicry, and so on, will all be cast as solutions to correspondence problems with different particular selection criteria.

Dissimilar Bodies

A correspondence need not be a one-to-one mapping, but could also be a one-to-many, many-to-one, or many-to-many relation. If the number of *degrees of freedom* (DOFs) in, for example, the joints of two agents' arms are different, then there can be no simple one-to-one correspondence between their actions. A robot might imitate a human nodding or waving successfully even without requiring that it has the same number and type of joints in its head and neck or arms and hands as the human whose behavior it emulates.

In fact, exact copying, even with similar embodiment, is almost never possible: One never has exactly the same agents with exactly the same kinds of bodies in exactly the same setting when the behavior of one agent is said to match that of another, as they must differ at least in their situatedness in time and/or space, not to mention numerous other details (Whiten & Ham, 1992; Nehaniv & Dautenhahn, 2001).

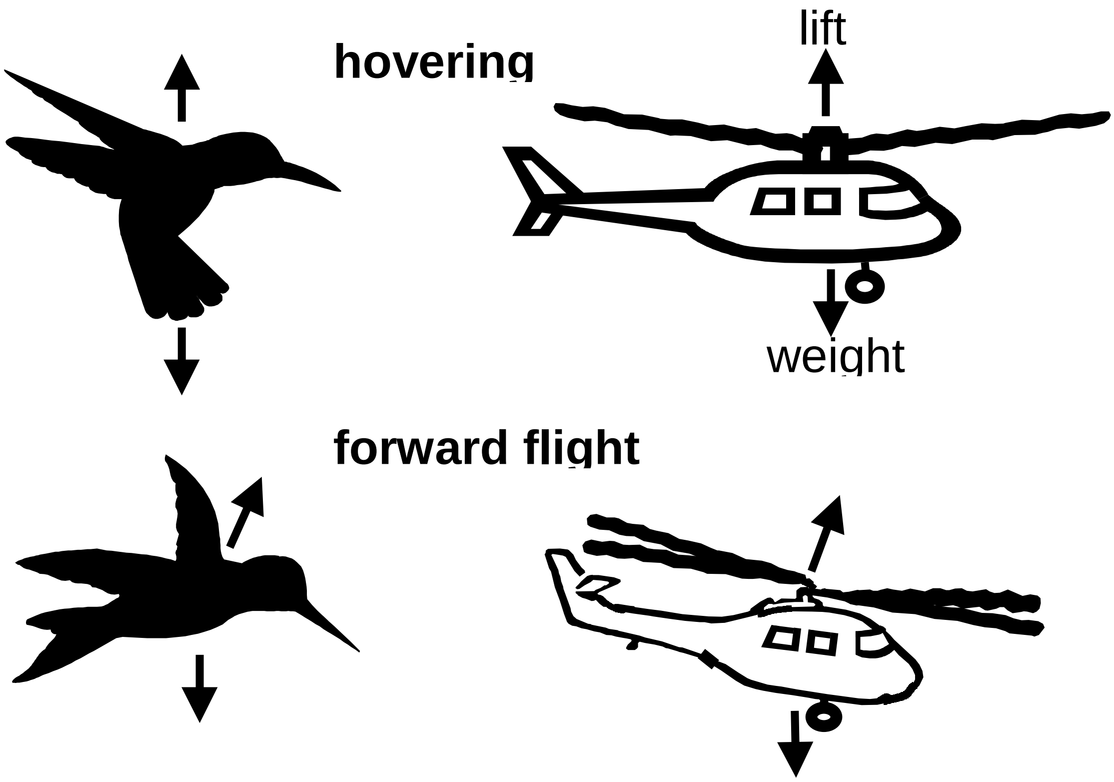
A useful correspondence could also be a *partial* mapping. That is, it need not be defined on all possible states and actions of the model: it may describe corresponding states and actions in the imitator for only some of these. For example, a robot might be able to successfully imitate me waving my arms using a particular correspondence between the angles it observes between my shoulder and arm and at my elbow, but this does not mean that the correspondence gives any information on how to relate my legs to its legs. Indeed, it might not even have legs!

To achieve a behavioral match, of whatever sort, a correspondence must explicitly or implicity be present. In trying to imitate or learning socially by observing, an autonomous agent must somehow solve the *correspondence problem*:

Informal Statement of the Correspondence Problem:
Given an observed behavior of the model, which from a given starting state leads the model through a sequence (or hierarchy) of subgoals - in states, action, and/or effects, while possibly responding to sensory stimuli and external events, find and execute a sequence of actions using one's own (possibly dissimilar) embodiment, which from a corresponding starting state, lead through corresponding subgoals - in corresponding states, actions, and/or effects, while possibly responding to corresponding events.

Thus such a correspondence can require relating aspects of actions of the model to actions that one can carry out, or states of the model in its environment to states of one's own body and environment, or both. It may also require relating sensory and external events of the demonstrator to those experienced by the imitator. A simple correspondence recognized between the hummingbirds and helicopters is given in the example of figure 1 (Nehaniv & Dautenhahn, 2001).

When creatures with less similar embodiments, such as dolphins, parrots, orangutans, chimpanzees and bonobos, exhibit vocal or motor or goal-oriented behaviors matching those of human demonstrators, the correspondences between the bodies of the animals and the humans become more abstract than those between similarly embodied model and imitator. Nevertheless, these animals, at least when



enculturated with humans, do display behaviors matching those of the human models (Herman, 2002; Pepperberg, 2002; Russon and Galdikas, 1993, 1995; Tomasello *et al*., 1993). This means that either they – or at least the human experimenters observing their behavior – recognize correspondences between the bodies, sensory and effector systems of these animals and the bodies, sensory and effector systems of their human models. Such correspondence may indeed be attributed to be present by the experimenters. While we may not conclude that the animals in question are necessarily aware of the correspondence, nevertheless, these animals at least are able to act in a manner strongly suggesting that they have solved a partial correspondence problem between their own bodies and that of a human demonstrator.

Referring to Louis Herman's results on dolphin imitation of humans (Herman, 2002), the dolphins were tested for imitation of human demonstrators using an unstated correspondence and showed that they are largely able to imitate the humans according to it: human waving forearms with bent elbow relate to wiggling pectoral fins; a human propelling her body partially out of water corresponds to a leap out of the water by the dolphin; human and dolphin underwater somersaults correspond; human head and dolphin head correspond, but human legs correspond to the dolphin's tail when raising them out of the water or slapping them on its surface. One could speculate that possibly the dolphin understands how its body-plan relates to that of a human. Referring to Irene Pepperberg's data on Grey parrots, 2002., Alex and other parrots show acoustical production of human-like speech sounds using very different vocal apparatus. Both they and we have recognized the correspondence between the sounds although these do not have identical acoustical formants and spectra. Moreover, these parrots use their speech to refer to and describe properties of objects in a manner related to how we as humans use similar sounds.

Furthermore, by the use of more “exotic” correspondences, it is possible to set up unusual temporal synchronizations, of perceived behavior and action. For example, Dautenhahn (1999) describes experimental couplings between hand movements of a human and the behavior of a "dancing" mobile robot moving on the ground.

2 Successful Correspondence?

When does a candidate behavior actually match an observed behavior? In stating the correspondence problem, we required that a sequence or hierarchy of corresponding subgoals be attained. This notion of *subgoal* in the statement of the correspondence problem should not be taken to necessarily imply any intentionality on the part of the demonstrator or imitator. For biological and autonomous agents, it is useful to accept a notion of “on behalf of” (S. Kaufmann, pers. comm.). Biological agents engage in behavior that is generally somehow beneficial to them. *E. coli* follow a gradient likely to lead them to food, and growing plants may turn toward sunlight. These behaviors are on behalf of the agent, helping it attain its goals. Having goals does not imply any intentional mental state. It is in this sense that *goal* is used here. A subgoal here is thus either a state of affairs that would promote the “on behalf of” the autonomous agent or which is observer-attributed as such. Goals do not arise independently of autonomous agents. (See also the agent-based discussions of *meaningful information* as information which is *useful* for an autonomous agent in achieving its goals (Nehaniv, 1999; Nehaniv *et al.*,1999)).

Whether or not behavior is judged as matching seems to be very much a subjective issue. Different observers of a candidate matching behavior may attribute a vast range of differing mechanisms and goals to the imitator and/or the demonstrator. The imitator itself might use yet another mechanism and unknown criteria in responding to stimuli and perceptions of its own surrounding world (*Umwelt*) in order to generate the behavior. This does not mean that it is hopeless to endeavor to formalize what is meant by such terms as imitation, mimicry, etc., but it does point to the central role of the observer (who might coincide with the demonstrator, the learner, or be a third party) in deciding whether or not an exhibited behavior matches that of a model. (See also the discussion of the role of the observer in (Dautenhahn & Nehaniv, 2002).)

Clearly different observers may have different answers to the question of whether or not a corresponding behavior has occurred: Suppose for example that a certain fictional species of bird learns to produce an ultra-sonic signal by repeated exposure to the sound of a dog whistle. A naive human observer would hear no sound when observing this bird produce this call and could not say to what extent the call was similar or not to the sound the whistle produces. But an observer equipped with listening equipment to transpose the sounds to the human-audible range would be able to give some evaluation of how well the bird’s call matched that of the whistle.

3 Correspondences in Actions, States, and Goals

In imitation, emulation, mimicry, stimulus enhancement and other behavioral areas of learning and matching, we take account of mapping (and more generally relations) in

STATE of the system (body), objects and environment
ACTION (and sequences of actions) which transform the state, including internally generate actions, and external ones (sensory stimuli and other events)
GOALS – the configuration of state (and/or possibly action sequences) that meet an external or internal criterion

The aspects have also been identified by Call and Carpenter, 2002. as three major sources of information in social learning, which researchers may use to categorize different types of matching behavior and social learning. (They also use the term "result" as an alternative to "state", but we do not in order to avoid the suggestion of a goal implicit in the word "result".)

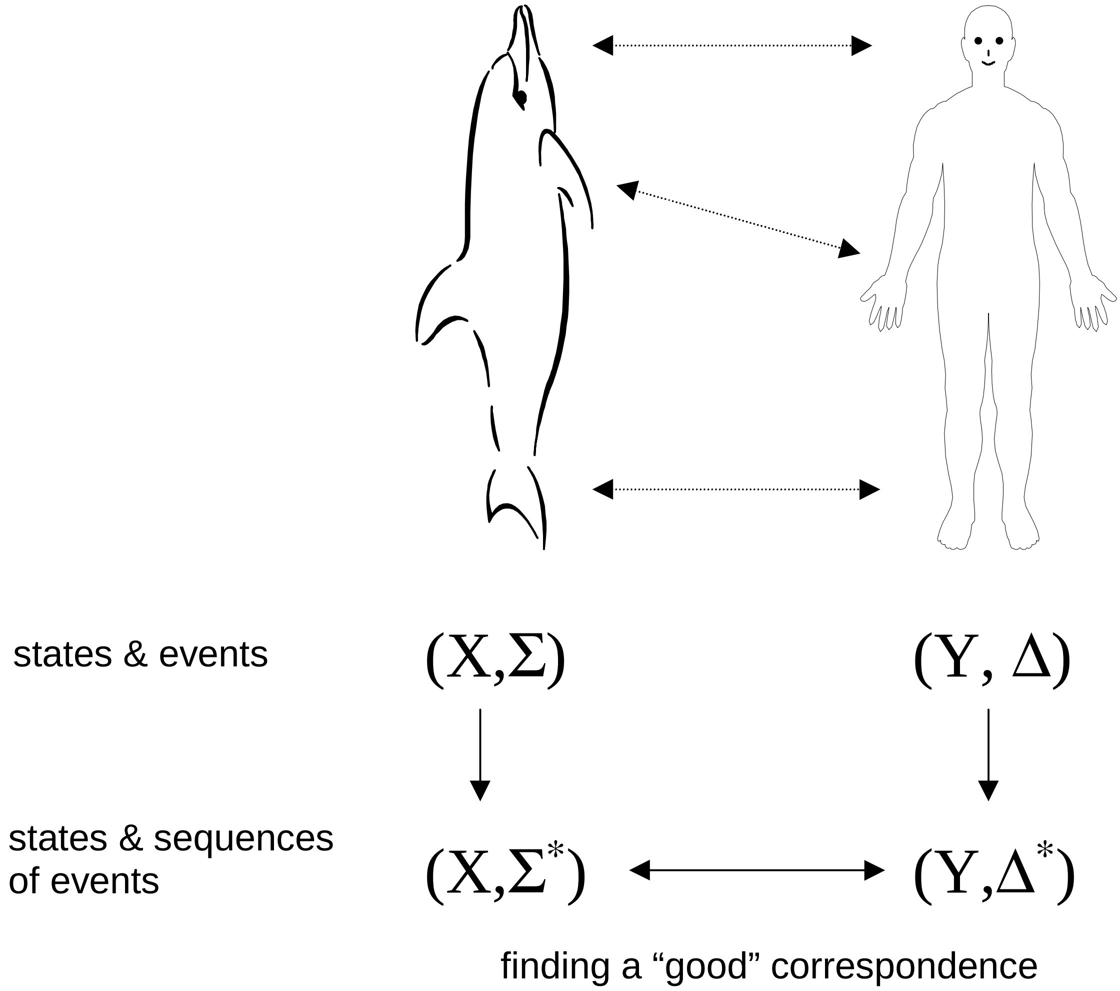
Our formalization of correspondences takes all of the above into account. Degree of success of attempted matching behavior can be formalized by metrics on states and actions with respect to attainment of sequences of subgoals as we explain in the next section. Different types of error measures, variously emphasizing state, actions, or goals, and granularity, in this formalization describe different types of observational behavior matching.

4 Formalization of Correspondences

In this section we give a simple mathematical framework to describe the notion of correspondence between two autonomous agents in a rigorous way. We have tried to do this in manner that will be useful for an interdisciplinary reader. No specialized mathematical training is required to benefit from this discussion, which provides a broadly applicable framework useful in understanding various solutions of the correspondence problem.

States and Events

For two autonomous agents, animals, robots, or software systems, identify the set of each one’s possible states and actions. Denote the states (of the body and environment) of the first agent as X and that states of the second as Y. Denote the set of elementary actions of the first agent by Σ and those of the second agent by Δ. The set of all finite sequence of actions of the first agent is denoted Σ^* and the set of all finite sequences of action of the second is denoted Δ^*. Effects on the environment will be reflected in state. One may also speak of particular properties of action/events or of sequences of action/events. From the description of an agent by a set of states and set of



action/events (X, Σ), we derive an extended description with the same states and all finite sequences of events (X, Σ*).[*] Using sequences of events rather than such individual ones allows for the fact that a single action, for example, of the demonstrator might correspond in the imitator to several actions, or *vice versa*. This situation is illustrated in figure 2.

Although we shall not go into this in article, the identification (collapsing) of sequences of actions that always have the same effects allows one to pass from the description of states and elementary action/events (X, Σ) to an algebraic invariant of the agent called its *transformation semigroup* (X,S). This will not be described further here, but we wish only to note that correspondences can be constructed at this level. For details, see (Nehaniv & Dautenhahn, 2001). Mathematical techniques for solving the problem of making a correspondence many-to-one (rather than many-to-many) are given by Nehaniv (1996).

Correspondences as Relational Mapping

Formally, a *correspondence* (or *relational mapping*) between the two autonomous agents is a relation of states $\Phi \subseteq X \times Y$ and a relation of sequences of actions $\Psi \subseteq \Sigma^* \times \Delta^*$, satisfying:

$$\text{For all } x \in X \text{ and } y \in Y, \text{ if } (x,y) \in \Phi \text{ and } (s,t) \in \Psi \text{ then } (x\,s,\ y\,t) \in \Phi.$$

Let us consider what this means in plainer language. A state relation Φ consists of those pairs of states of the two agents which are said to correspond. Similarly, an action relation Ψ consists of those pairs of sequences of elementary actions which are said to correspond. In words, the above condition says: "If state x of the first system corresponds to state y of the second system, and action sequence s in the first is related to action sequence t in the second, then the corresponding resulting states must correspond". More briefly: "When starting with corresponding states, corresponding actions lead to corresponding states."

An attempted correspondence might be a one-to-one function, or many-to-many relation as discussed above. It need not be fully defined. Of course there will in general exist many correspondences between two autonomous agents. Whether or not a particular candidate relation is a good solution to the correspondence problem depends on the evaluation of how appropriate the correspondence is from the viewpoint of some observer. This might be an external observer, or a demonstrator or an imitator.

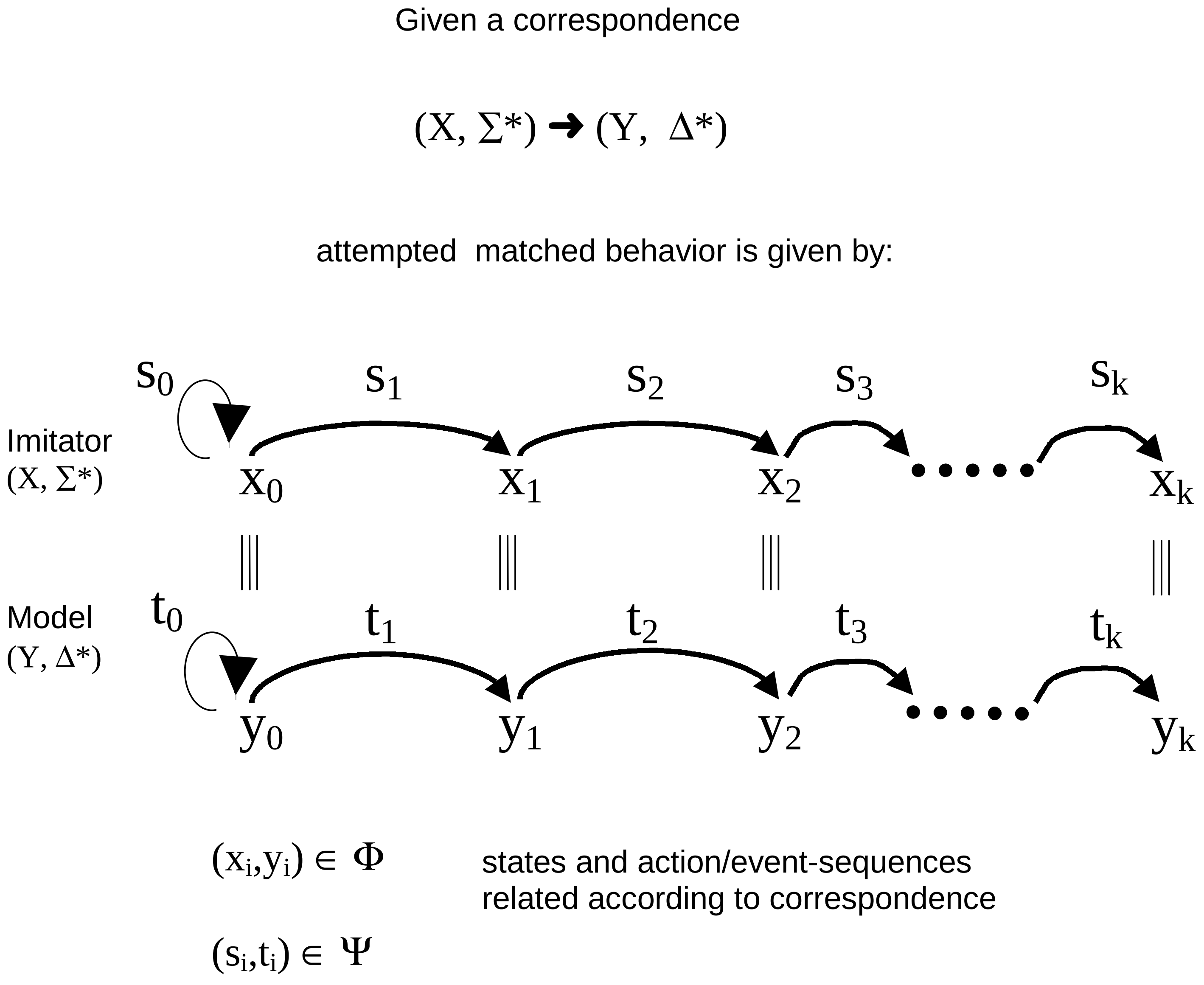
[*] Mathematically, this can be described as deriving, in a canonical manner, a free transformation semigroup action from an automaton.

Using the concepts and notation given above, a simple formalization of the correspondence problem is given by the following:

Formal Statement of the Correspondence Problem:
Given an observed behavior of the model, parse this behavior into action sequences $t_1, ..., t_k$, starting in state y_0 with states $y_1, ..., y_k$ the states which successively result from carrying out the t_i in order. Find a correspondence, Φ on states and Ψ on action sequences, relating one's own actions and states to those of the observed model, and from a corresponding state x_0 execute corresponding behaviors $s_1, ..., s_k$ such that the resulting behavior is as similar as possible, where the degree of dissimilarity is measured by summing an error measure $\mathbf{d}((y_i,t_i),(x_i,s_i))$ over i.

In fact, the "actions" in this formulation are actually "action/events", but we have tried to make the statement as concise as possible and so have included the events within the notion of action: Experiencing an event such as a particular sensory perception is taken to be a particular kind of action.

Behavior matching according to a correspondence is schematically illustrated in figure 3. (Also in the figure, s_0 and t_0 are "null



actions", included for mathematical convenience in computing dissimilarity as the error sum with index i going from 0 to k.) Note that a correspondence that works for generating one behavioral match may or may not work well for matching other observed behaviors. But a good, detailed correspondence, such as the dolphins' evident identification of actions of its body parts with those of human trainers mentioned above, provides reusable knowledge that applies across many situations to yield successful imitative behaviors.

We emphasize here that our formalization is a method of description helpful in modeling an imitator matching an observed behavior. We are certainly not claiming that animals are using such a framework to guide their own matching behavior, but only that the framework provides a useful description of behavioral matching that is independent of the particular mechanisms actually employed. A designer of an autonomous agent that is trying to imitate another agent could have this agent use this framework for selecting among possible correspondences. An observer, for example, a scientist studying behavioral matching in animals, can usefully employ this framework to describe the particular correspondence in any matching observed, in particular, whether aspects of state, actions, or goals are most descriptive (or predictive) of observed imitative behavior and which type of dissimilarity measure best describes (or predicts) the aspects that are matched. In addition, by casting results on matching behavior in this framework, different research on imitation and matching behavior can be compared.

Metrics and Measures of Success in Behavior Matching

Metrics and *error measures*, can capture the notion of (1) the difference of performed actions from desired actions, (2) can measure the difference of attained states from desired states, or (3) can measure the difference of sequences of both. Such measures can take be discrete or continuous values. This is the role of the dissimilarity measure **d** in the formal statement of the correspondence problem above.

Evaluation of a candidate for behavioral matching depends in the formalization very much on the error measure **d**. Different kinds of measures result in different types of matching and would lead to types of learning of behaviors that ethologists and psychologists would classify as different types of social learning or copying. For example, if the error measure **d** in the above formulation of the correspondence problem ignores the state component, then only actions are salient to the success of attempted matching behavior.

Granularity refers to the fineness of the imitation, *e.g.* in the number of states, actions, or subgoals matched. If the measure **d** ignores the action component, then only the sequence of states attained is salient. The number k in the formal statement of the correspondence problem gives the granularity of the attempted behavioral match. If k=1, then only the overall end-result is salient for determining successful correspondence. The measure of dissimilarity can require that certain states or actions are attained or closely matched; thus it can be used to evaluate whether a sequence of subgoals has been attained or not, and to what degree. Matching the "right" subgoals can be forced for example by making the measure take value zero when they are achieved but a high error value otherwise.

Types of Behavior Matching

Learning a correspondence means learning a piece of such a relational mapping. Metrics and measures on state, actions, and goals can guide this learning. The type of social learning (or any of the related phenomena involving matching) depends on the metrics used and which aspect or aspects of a behavior they measure. More details are in (Nehaniv & Dautenhahn, 2001). For a review of mechanisms of social learning and simpler phenomena see (Call & Carpenter, 2002; Noble & Todd, 2002; Zentall, 2001). Various areas of matching and learning can be classified according to different aspects of the relational mapping in a particular attempted correspondence:

Construction of correspondence in actions and their sequencing using an action relation Ψ required by the metric characterizes "*mimicry*", or "*copying*" (without goal- or state-matching).

Construction of correspondences in actions Ψ and states Φ at high granularity matching attributed goals of demonstrator characterizes forms of *"imitation"* such as "*action-level imitation*" (Byrne & Russon, 1998) or "*string parsing*" (Byrne 1999) (with metric reflecting subgoal salience).

Construction of correspondence in states required by the metric restricting the state relation Φ characterizes "*emulation*" (Tomasello, 1990; see discussion by Call & Carpenter, 2002).

Construction of correspondence using state relation Φ via metrics requiring matching of an attributed goal characterizes "*goal emulation*" (Whiten & Ham, 1992) (with low granularity) or "*stimulus enhancement*" (Spence, 1937) (with low granularity and only state matching).

Construction of a correspondence in the action relation Φ with effect information in the states and metric reflecting subgoal salience characterizes learning "*tool affordances*" or "*action affordances*" (Tomasello & Call, 1997).

The above types of correspondence in which matching is required for hierarchically structured tasks, subtasks, involving behavioral loops and conditionals characterizes various forms of "*program-level imitation*", "*hierarchical procedure imitation*", *"procedural matching"* or *"programming by example"* (Byrne & Russon 1998; Whiten, 2002; Cypher, 1993; Lieberman, 2002; Furse, 2001),

Understanding and matching higher level structures characterize various other forms of observational learning and matching involving "*theory of mind" or "mind-reading"* (Premack & Woodruff, 1978; Byrne & Whiten, 1988; Povinelli & Preuss, 1995; Whiten & Byrne, 1997; V. Gallese & A. Goldman, 1998), "*empathy*" (O'Connell, 1995; Dautenhahn, 1997, 2000), or a "*body-plan correspondence*" (Nehaniv & Dautenhahn, 2002).

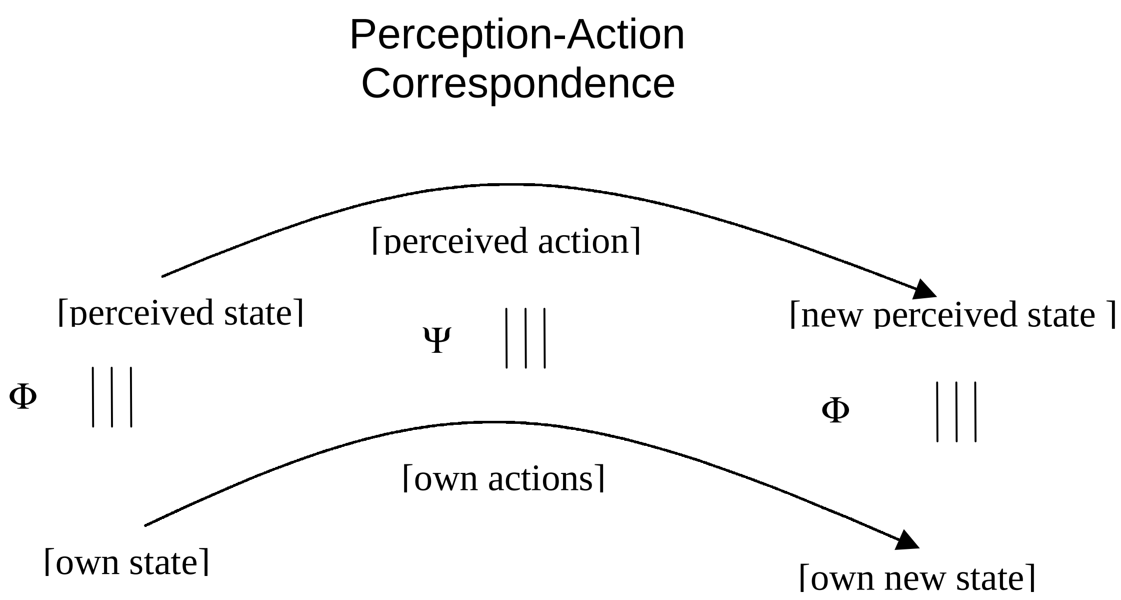
The reader may like to compare a similar classification of behavioral matching and types of social learning given by Call and Carpenter, 2002. As they have emphasized, in studying such behavior in animals, it is a good idea to try to identify which types of information the subject animal is using. Competing predictive models could be constructed using the above formalizations and different error measures reflecting various types of observational behavior matching. To help identify the mechanism being used by the animal, these competing models could then be evaluated by comparing their predictions with experimentally observed behavioral matching.

Conversely, in building artifacts and employing appropriate measures of dissimilarity on different aspects of state, actions, goals, and granularity of desired matching behavior, one can implement various forms of behavioral matching and social learning. For example, in agent-based computer simulations, Aris Alissandrakis and the authors have studied the results of applying different metrics and different granularities to generate different corresponding behaviors in the imitation of sequences of moves by differently embodied agents. These agents are embodied as chess pieces, whose movement constraints provide a well-known example of dissimilar "bodies" in a simple, discrete shared world, namely, a chess-board. A chess-world knight or bishop for example may imitate a zig-zag of three moves by a demonstrator queen, but each such agent is subject to the constraints on how it is allowed to move (its embodiment) and so can only roughly approximate the queen's precise behavior. Nevertheless, the knight and bishop or other pieces can successfully imitate the behavior to varying degrees with respect particular granularities (*e.g.* end result or trajectory matching) and various metrics. Varying the granularity and metrics for assessing attempts at imitation in this experimental setting illustrates the profound influence of these factors on the qualitative features of the resulting imitative behavior (Alissandrakis *et al.*, 2000). Similarly, differently embodied animals, robots, or other autonmomous agents whose behavior matches those of others can be modeled as using correspondences between their own states and actions and those of a demonstrator, using various granularities and measures of dissimilarity.

5 Perception-Action Correspondences

Another particular application of the relational mapping formalism is to perception-action correspondence. A perceived state corresponds to one's own state under the state component (Φ) of a relational mapping, perceptual stimuli from an action correspond under an action component (Ψ) to one's own action. The new perceived state – that is, the perceived state following the perceived action — should correspond to one's own new state after one's action. This is similar to inter-agent correspondence but applies it to the case that the first system is given by one's perceptions of one's own state and of one's own actions – for example by sensory kinaesthetic and proprioreceptive feedback (or also via external feedback, *e.g.* via mirrors or from others responding by imitating one's behavior) – while the second system is given by one's own state and actions. Solving the correspondence problem in this case means learning to control one's own body in the sense of relating one's sensors and effectors to the perceptual feedback they produce. Such a correspondence is schematized in figure 4.

Associative Sequence Learning Theory. For example, in the Associative Sequence Learning Theory (ASL) of C. M. Heyes and E. D. Ray (2000), see also (Heyes, 2002), associations between sensory data and motor representations of behavior ("vertical" links) can be viewed as parts of a correspondence in the formal sense described above. Learning these sensory-motor associations corresponds to building up a solution to the correspondence problem between perceived stimuli (generated by either one's own action or another's action) and one's own motor actions. Complementing this, representation of patterns of behavior is encoded during the learning of temporal sequencing of actions or of higher level behavioral programs which make use of perception-action correspondence. Currently ASL does not address the effect of actions on objects or the environment, but we believe that it can be fruitfully extended to incorporate these.



The Mirror System. An example of a candidate natural mechanism at the neurological level for solving the correspondence problem is the mirror system present the brains of at least some primate species, in which certain neurons fire both when performing a particular motor action such as grasping a piece of food and when seeing another performing such an act. Some of these neurons appear to encode the particular affordances of movements in relation to objects in the animal's environment (Gallese *et al.*, 1996, Gallese & Goldman, 1998; Hari *et al.*, 1998; Rizzolatti *et al.*, 1998; Rizzolatti & Arbib, 1998; Arbib, 2002). As such, the mirror system could mediate a relational mapping between the action of an agent on its environment and those of others it observes. This has also inspired the use of a similar mechanism in robots (Demiris, 1999; Demiris & Hayes, 2002).

Mechanisms, Representations and the Structure of Socially Learned Information

The structure of imitated behavior can often reveal something of the way in which observed behavior is represented in the imitator. This realization is apparent in the animal imitation literature in such distinctions as action- vs. program-level imitation (Byrne & Russon, 1998), sequential vs. hierarchical structuring of observationally acquired behavior (Whiten, 2002), and statistical string parsing without intentionality (Byrne, 1999).

The methods for extracting socially observed information are by no means unique and the spectrum of representations for organizing observationally acquired information into procedural representations is broad. Some representational formats, techniques, and paradigms used in sciences of the artificial include finite state automata (*e.g.* Hopcroft & Ullman, 1979), epsilon machines (Crutchfield 1994), generalized Hebbian temporal learning (Billard, 2002), behavioral cloning and rule extraction (Sammut *et al.*, 1992, 2002), programming by example (Cypher 1993; Lieberman, 2002), case-based reasoning (*e.g.* Kolodner, 1993), and stories (Nehaniv, 1997; Goodenough, 2002).

6 Discussion

Whatever form of behavioral matching one considers, regardless of the mechanisms behind it, at its heart must lie a notion of correspondence between autonomous agents. Together with such correspondences between the possible states and sequences of action/events of the respective agents, measures of dissimilarity in matching behavior are enough to classify these phenomena into different classes used by ethologists and psychologists. Various forms of imitation and related phenomena such as emulation, mimicry, blind copying, and social learning correspond to different aspects of constructing a relational mapping between possibly dissimilar bodies (or autonomous agents), in particular to whether actions, states, or goals, or some combination of these are required to correspond in the course of the behavior. If an animal learns to match the behavior of another, the type of correspondence problem being solved may be indicated in which aspects the matching actually occurs. This can be used to distinguish one form of matching from another, e.g. mimicry from goal emulation, etc. Metrics encoding goals relate to intentionality, tool affordances, reinforcement, and the judgement of whether situations are equivalent or not. The structuring of behavior in sequences of actions, hierarchies, or behavioral programs is relevant for studying how solutions of the corrrespondence problem are utilized by an animal. The representation of procedural knowledge as a program, string of actions, hierarchically organized collection of subgoals, etc., can be used in artificial systems to make use of attempted correspondences and perception-action relations in applications such as programming by example, behavioral cloning, and robot and agent social learning.

All these aspects – action and state correspondences, metrics encoding goals, and the structuring of behavior – need to be considered in studying observational learning and imitation. One should ask, what type of novelty occurs in each of these aspects? (See also the discussion of novelty in (Dautenhahn & Nehaniv, 2002).) In which of these aspects is there a consistent correspondence? In regard to animals, psychologists and ethologists are interested in uncovering the structure of observational learning, while, in the realm of artifacts, engineering such structure and tinkering with its various aspects affords the possibility to introduce biologically-inspired social learning mechanisms into artificial systems.

Solutions to the correspondence problem are a result of successful attempts at imitating (*trying to imitate* (Dautenhahn, 1994)) or at mimicry or learning socially. In contrast, in building artifacts, or sometimes even in education (rote or pattern learning), it can be useful to harness *learning by imitation*: Here the correspondence problem is solved at the outset, imitation has been engineered in, but by imitating a teacher, a learner agent comes to experience situations where learning of other skills takes place. See (Hayes & Demiris, 1994; Demiris & Hayes, 1996; Billard & Dautenhahn, 1997, 1998, 1999; Billard, 2002) for examples of such learning that is facilitated by imitation. See (Nehaniv & Dautenhahn, 2001) for further discussion of the distinction between trying to imitate and learning by imitation.

We have endeavored to provide a framework with solid mathematical foundations in which one can address matching phenomena in both natural and artificial systems. The algebraic framework presented for the correspondence problem was initiated by Nehaniv and Dautenhahn (1998). We hope this will be useful in revealing hidden assumptions and observer-dependent aspects of criteria for judging whether or not social learning, imitation, or matching has taken place. Moreover, the notions of relational mapping and metrics provide one with a toolkit for studying what type of correspondence is being constructed, whether one is observing it in controlled experiments with humans and animals, or in other settings involving the matching of behaviors. For workers building artificial systems that can learn by observing, this framework provides a language and mathematical tools for analyzing the aspects of any artificial system that should learn by observing another agent, whether human, animal, robotic or of other type. By studying which components of a relational mapping need to be engineered – what aspects of a correspondence are to be built in and what aspects are to learned – a designer may approach the various aspects of a social learning artifact in a systematic manner. Introducing this common framework for correspondence to the study of imitation in animals and artifacts also gives us a way to compare, evaluate, and relate research on imitation and behavior-matching in widely disparate studies ranging from animal studies involving various different species to research studies on robots and software agents that seek to engineer and harness such social learning phenomena into artifacts.

References

R. M. Alexander, *Exploring Biomechanics: Animals in Motion*, Scientific American Library, Ch. 5, 1992.

A.Alissandrakis, C. L. Nehaniv, & K. Dautenhahn, Learning how to do things with imitation, *Proceedings AAAI Fall Symposium* on *Learning How to Do Things*, American Association for Artificial Intelligence Press, 2000. On the web at http://homepages.feis.herts.ac.uk/~nehaniv/chessworld.ps

M. Arbib, The mirror system, imitation, and the evolution of language. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts.* MIT Press, 2002.

A. Billard, Imitation: A means to enhance learning of a synthetic proto-language in autonomous robots. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.
A. Billard & K. Dautenhahn, Grounding communication in situated, social robots, In *Proceedings TIMR UK 97 Towards Intelligent Mobile Robots*, Technical Report Series of the Department of Computer Science, Manchester University, 1997.
A. Billard & K. Dautenhahn, Grounding communication in autonomous robots: an experimental study. *Robotics and Autonomous Systems*, 24(1-2):71-81, 1998.
A. Billard & K. Dautenhahn, Experiments in social robotics: grounding and use of communication in autonomous agents, *Adaptive Behavior Journal*, special issue on "Simulation Models of Social Agents", 7(3-4):415-438, 1999.
C. Breazeal & B. Scassellati, Issues in building robots that imitate people. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.
R. W. Byrne, Imitation without intentionality. Using string parsing to copy the organization of behaviour, *Animal Cognition*, 2:63-72, 1999.
R. W. Byrne & A. E. Russon. Learning by imitation: a hierarchical approach. *Behavioral and Brain Sciences*, 21, 667-709, 1998.
R. W. Byrne & A. Whiten (Eds.), *Machiavellian Intelligence*, Clarendon Press, 1988.
J. Call & M. Carpenter, Three sources of information in social learning. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.
J. Crutchfield, Observing complexity and the complexity of observation. In H. Atmanspacher (Ed.), *Inside versus Outside,* Springer Verlag, pp. 234-272, 1994.
A. Cypher (Ed.), *Watch What I Do: Programming by Demonstration*, MIT Press, 1993.
K. Dautenhahn, Trying to Imitate — a step towards releasing robots from social isolation. In P. Gaussier and J.-D. Nicoud (Eds.), *Proc. From Perception to Actions Conference (Lausanne, Switzerland)*, IEEE Computer Society Press, pp. 290-301, 1994.
K. Dautenhahn, I could be you – the phenomenological dimension of social understanding, *Cybernetics and Systems*, 25(8): 417-453, 1997.
K. Dautenhahn, Embodiment and interaction in socially intelligent life-like agents. In C. L. Nehaniv, ed. *Computation for Metaphors, Analogy, and Agents*, Lecture Notes in Artificial Intelligence, Vol. 1562. Springer Verlag, pp 102-142, 1999.
K. Dautenhahn, Socially Intelligent Agents and the Primate Social Brain: Towards a Science of Social Minds, Proceedings AAAI Fall Symposium *Socially Intelligent Agents - The Human in the Loop*, American Association for Artificial Intelligence Press, 2000.
K. Dautenhahn & C .L. Nehaniv, The Agent-based Perspective on Imitation. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.
J. Demiris. *Movement Imitation Mechanisms in Robots and Humans*. PhD Thesis, University of Edinburgh, Scotland, U.K., 1999.
J. Demiris & G. Hayes, Imitative learning mechanisms in robots and humans, *Proceedings of the 5th European Workshop on Learning Robots*, Bari, Italy, July 1996.
J. Demiris & G. Hayes, Imitation as a dual-route process featuring predictive and learning components: A biologically-plausible computational model. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.
E. Furse, A model of imitation learning of algorithms from worked examples, *Cybernetics and Systems*, Special issue "Imitation in Natural and Artificial Systems", 32(1-2), 2001.
V. Gallese, L. Fadiga, L. Fogassi & G. Rizzolatti, Action recognition in the premotor cortex, *Brain* 119:593-609, 1996.
V. Gallese & A. Goldman, Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12): 493-501, December 1998.
O. R. Goodenough, Information replication in culture: Three modes for the transmission of culture elements through observed action. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.
R. Hari, N. Forss, S. Avikainan, S. Salenius, & G. Rizzolatti. Activation of human primary cortex during action observations: a neuromagnetic study. *Proceedings of the National Academy of Sciences*, 95, 15061-15065, 1998.
G. Hayes & J. Demiris, A robot controller using learning by imitation. In A. Borkowski & J. L. Crowley (Eds.), *SIRS-94, Proceedings of the 2nd International Symposium on Intelligent Robotic Systems*, LIFIA-IMAG, Grenoble, France, pp. 198-204, July 1994.
L. M. Herman, Vocal, social, and self imitation by bottlenosed dolphins. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.
C. M. Heyes, Transformational and associative theories of imitation. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.
C. M. Heyes & E. D. Ray, What is the significance of imitation in animals? *Advances in the Study of Behavior*, 29:215-245, 2000.
H. Lieberman, Art imitates life: Programming by example as an imitation game. In K. Dautenhahn & C. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.
J. E. Hopcroft & J. D. Ullman, *Introduction to Automata Theory, Languages and Computation, Addison-Wesley*, 1979.
S. Kauffman, remarks on goals and purpose in biological agents during a lecture delivered to the Working Group on Evolvability, Santa Fe Institute, New Mexico, pers. comm., April 2000.
J. Kolodner, *Case-Based Reasoning,* Morgan Kaufmann Publishers, 1993.
C. L. Nehaniv, From relation to emulation: The Covering Lemma for transformation semigroups, *Journal of Pure and Applied Algebra*, 107:75-87, 1996.
C. L. Nehaniv, What's your story? – Irreversibility, algebra, autobiographic agents. In K. Dautenhahn (Ed.), *Socially Intelligent Agents: Papers from the 1997 AAAI Fall Symposium* (November 1997, MIT, Cambridge, Massachusetts), American Association for Artificial Intelligence Press, pp. 150-153, 1997.
C. L. Nehaniv, Meaning for Observers and Agents, *IEEE International Symposium on Intelligent Control / Intelligent Systems and Semiotics (ISIC/ISAS'99)*, Cambridge, MA, Institute of Electronic and Electrical Engineers Press, pp. 435-440, 1999.
C. Nehaniv & K. Dautenhahn, Mapping between dissimilar bodies: Affordances and the algebraic foundations of imitation. In J. Demiris and A. Birk (Eds.), *Proc. European Workshop on Learning Robots 1998 (EWLR-7)*, Edinburgh, Scotland, 1998.
C. L. Nehaniv & K. Dautenhahn, Like me? - Measures of correspondence and imitation *Cybernetics and Systems*, Special issue on "Imitation in Natural and Artificial Systems", 32(1-2), 2001.
C. L. Neha niv & K. Dautenhahn, The Correspondence Problem. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.
C. L. Nehaniv, K. Dautenhahn, & M. J. Loomes, Constructive biology and approaches to temporal grounding in post-reactive robotics. In G. T. McKee & P. Schenker (Eds.), *Sensor Fusion and Decentralized Control in Robotics Systems II,* Proceedings of SPIE Vol. 3839, pp. 156-167, 1999.

J. Noble & P. Todd, Imitation or something simpler? Modeling simple mechanisms for social information processing. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.

S. M. O'Connell, Empathy in chimpanzees: Evidence for theory of mind?, *Primates* 36(3):397-410, 1995.

I. M. Pepperberg, Allospecific referential speech acquistion in Grey parrots (*Psittacus erithacus)*: Evidence for multiple levels of avian vocal imitation. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.

D. J. Povinelli & T.M. Preuss, Theory of mind: evolutionary history of a cognitive specialization, *Trends in Neurosciences*, 18(9):418-424, 1995.

D. Premack & G. Woodruff, Does the chimpanzee have a theory of mind?, *Behavioral and Brain Sciences,* 4:515-526, 1978.

G. Rizzolatti & M. A. Arbib, Language within our grasp, *Trends in Neurosciences*, 21(5):188-194, 1998.

G. Rizzolatti, L. Fadiga, V. Gallese, & L. Fogassi. Premotor cortex and the recognition of motor actions. *Cognition and Brain Research*, 3: 131-141, 1998.

A. Russon and B. M. F. Galdikas, Imitation in ex-captive orangutans, *Journal of Comparative Psychology*, 107:147-161, 1993.

A. Russon and B. M. F. Galdikas, Contraints on great apes' imitation: model and action selectivity in rehabilitant orangutan (*Pongo pygmaeus*) imitation. *Journal of Comparative Psychology*, 109:5-17, 1995.

C. Sammut, S. Hurst, D. Kedzier, & D. Michie, Learning to fly, *Proc. 9th International Machine Learning Conference (ML'92*), Morgan Kaufmann, 1992. Preprinted in K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.

K. W. Spence, Experimental studies of learning and higher mental processes in infra-human primates, *Psychological Bulletin*, 34:806-850, 1937

M. Tomasello, Cultural transmission and tool use and communicatory signaling of chimpanzees? In S. Parker & K. Gibson (Eds.), *Language and Intelligence in Monkeys and Apes: Comparative Develoopmental Perspectives*, Cambridge Unviersity Press, 1990.

M. Tomasello & J. Call, *Primate Cognition*. Oxford University Press, 1997. M. Tomasello, E. S. Savage-Rumbaugh, & A. C. Kruger, Imitative learning of actions on objects by children, chimpanzees, and enculturated chimpanzees*.* *Child Development,* 64:1688-1705, 1993.

A. Whiten, Imitation of sequential and hierarchical structure in action: Experimental studies with children and chimpanzee. In K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts,* MIT Press, 2002.

A. Whiten & RW Byrne*, Machiavellian Intelligence II: Extensions and Evaluations*. Cambridge University Press, 1997.

A. Whiten & R. Ham. On the nature and evolution of imitation in the animal kingdom: Reappraisal of a century of research. In P. J. B. Slater, J. S. Rosenblatt, C. Beer, & M. Milinski (Eds.), *Advances in the Study of Behavior*, *21*, Academic Press, pp. 239-283, 1992.

T. R. Zentall, Imitation and Other Forms of Social Learning in Animals: Evidence, Function, and Mechanisms, *Cybernetics and Systems*, Special issue on "Imitation in Natural and Artificial Systems", 32(1-2), 2001.