From Spreading of Behavior to Dyadic Interaction—A Robot Learns What to Imitate

E. I. Barakova,[1,*] D. Vanderelst[2]
[1]Faculty of Industrial Design, Eindhoven University of Technology Den Doleh 2 5612 AZ Eindhoven The Netherlands
[2]Active Perception Lab, University Antwerp, Belgium

Imitation learning is a promising way to learn new behavior in robotic multiagent systems and in human-robot interaction. However, imitating agents should be able to decide autonomously which behavior, observed in others, is interesting to copy. This paper shows a method for extraction of meaningful chunks of information from a continuous sequence of observed actions by using a simple recurrent network (Elman Net). Results show that, independently of the high level of task-specific noise, Elman nets can be used for learning through prediction a reoccurring action patterns, observed in another robotic agent. We conclude that this primarily robot to robot interaction study can be generalized to human-robot interaction and show how we use these results for recognizing emotional behaviors in human-robot interaction scenarios. The limitations of the proposed approach and the future directions are discussed. © 2010 Wiley Periodicals, Inc.

1. INTRODUCTION

An organism is able to learn by imitation if it can acquire new behavioral skills by directly copying them from others.[1] Imitation, like other forms of social learning,[2,48,49] has many roles in animal and human societies. It has, potentially, an enormous ecological advantage,[4] by allowing animals to be flexible learners while avoiding the dangers associated with individual learning.[5] The behavior of others has often been already shaped by its consequences and can therefore be assumed to be safe and rewarding to imitate.[3]

Another property of imitation, together with its ecological value, is that it can support the spread of behavior through a population of individuals.[2,6,7] Several observational studies[8] have provided evidence for this in-groups of primates but recently experimental evidence has also been reported. For example, Bonnie[9] taught individual chimpanzees to deposit tokens in a box to receive a reward. Subsequently,

*Author to whom all correspondence should be addressed: e-mail: e.i.barakova@tue.nl

INTERNATIONAL JOURNAL OF INTELLIGENT SYSTEMS, VOL. 00, 1–18 (2010)

View this article online at wileyonlinelibrary.com. • DOI 10.1002/int.20464

these individuals were introduced into a population of naive animals. After some time, the rewarding behavior was copied by the other animals and its frequency in the population increased. Similar findings have since then been reported by Whiten.[10]

Humans and some primates have been found to imitate.[3,8,9,11,12] Experimental studies of imitation have shown that observing adults is a powerful way of learning for children.[13,14] It has been shown that imitation is instrumental in learning to interact with objects[13,14,17–19] and it increases the interaction between the imitator and the imitated agent.[15,16,21,49]

The ecological advantages and the capacity to support the spread of behavior make imitation a potentially interesting mechanism to support learning in multiagent robotic systems.[20] The role of imitation in learning to communicate (via reproduction of gestures and speech), and in stimulating the interaction between the imitator and the imitated agent have made it an interesting topic in human-robot interaction studies.

In a multiagent setting, agents could search simultaneously for a solution for a given problem (e.g. how to pick up food or accomplish a common task). Once a single agent has found a solution, this innovation could be imitated by others and could be propagated through the population. In this way, learning by imitation could drastically reduce the total number of learning trials needed for a population of agents to solve a problem.[20]

Human-robot interaction scenarios often require robots to learn from humans or other robots that are not trained to explicitly demonstrate behaviors. A robot without any *a priori* knowledge about the task does not know which actions of the person are important and necessary for the task, while he/she sometimes produces not only actions directly related to the task but also unrelated ones. Not preconditioned interaction scenarios are preferred in social training of children with autism[21–23,49] where robots have demonstrated to be appropriate educational tools. Moreover, the idea that robots are used in everyday life and they learn skills from human or other robots without being explicitly trained to do that is attractive. The mentioned classes of robot applications, namely spreading of behavior for survival or accomplishment of a common goal, and learning (social) skills from a trustworthy partner share the same problem of identifying and learning sequences of actions that have to be imitated. This problem is also known as learning what to imitate.[24]

We propose a general framework to address this common problem of identifying and learning meaningful sequences of actions for both classes of applications. The framework is as follows. An agent (human, robot or non human primate) sees a continuous sequence of various kinds of actions. The actions which are, from the viewpoint of the observed agents, meaningful and can be divided into exploring and exploiting part, are unordered from the viewpoint of a learning agent. There is no *a priori* way for the learning agent to parse the actions of its conspecifics or of the human companion. Therefore, it cannot know which sequences it should copy and which are to be ignored. This implies that it should copy the exploiting action sequences and ignore the exploring action sequences. Moreover, it cannot know where the exploiting sequences start or end. Even if it is assumed that the imitating agent can detect when another agent is rewarded (i.e., possible end of an exploiting sequence), it cannot know where the sequence of events that lead to the reward started. This problem can also be stated otherwise: Agents cannot act as teachers explicitly signaling to others which actions are interesting to learn. In the absence of teachers that are explicit about what needs to be learned, agents should be able to autonomously decide what is important to copy.

In search for a solution to the problem of identifying and learning meaningful actions we used the analogy to another problem in cognition; that of segmenting a continuous input stream of sounds into words. Elman (1990) presented a simple recurrent neural network that was capable of making good predictions of the unfolding word based on the co-occurrence statistics available in the data. Taking as a basis this network, the imitation of meaningful sequence of behaviors involves prediction of the next movement that constructs the behavior.

While imitation in multiagent settings seems to be a promising learning paradigm, until now little or no research has been done in this area. To the best of our knowledge in the context of embodied autonomous multiagent systems the use of imitation where no explicit teacher is present has been investigated so far only by Belpeame et al. [25]. Most existing research on autonomous imitation in artificial agents focuses on human-machine imitation (see Ref. 26 for an overview) where the human is a teacher who clearly marks the boundaries of the behavior that has to be imitated. In such a setting, there is no need for an agent to detect the boundaries between meaningful chunks of actions since they are marked by the teacher.[25] We intend to make possible for robots to be used also in scenarios where the explicit demonstration of meaningful behaviors is not possible, or even harmful to the interaction process. In particular, we aim to use the method for letting the robot be used for social training of autistic children by robots. In such a scenario, explicit demonstration of behavior may disturb the natural interaction between the robot and the human.

The article is organized as follows. Section 2 outlines the constraints of the problem that is going to be solved in the light of the application domains. It also gives more details on the method and the algorithm that was used. Section 3 gives all necessary details on the data collection, preprocessing, and learning to make possible for the reader to reproduce the experiments. The results are shown and explained in Section 4. Section 5 describe the ongoing work on human-robot interaction that is based on this method. A discussion is provided in Section 6.

2. PROBLEM DEFINITION AND APPROACH

2.1. Problem Definition

Imitating agents in a multiagent setting face a number of fundamental problems.[24,25] A problem of crucial importance of them is how agents can autonomously select the behavior that should be copied.[24] Great apes and humans seem to be very good at determining what behavior should be imitated when they observe a demonstrator.[27] However, for robots in the multiagent scenario sketched above, determining what they should imitate is no simple task.[26]

Similarly, the robot that has to take part in realistic human-robot interaction scenarios cannot always be explicitly thought of which actions to imitate for a successful interaction with the human. The problem of social learning or one-directional transfer of skills implies a delicate balance between social mechanisms and motor actions. Detecting social cues that are given implicitly or explicitly by the teacher during training can serve as a way to determine when the imitation process starts or stops. In multiple social interaction scenarios, this is not possible. The robot would try to pick up behavior from the humans and respond to it without the human to explicitly signal for that.

Usually, more than one meaningful behavior has to be learned. Lets imagine a number of agents exploring an artificial world. Different types of food and supplies are available in this artificial world (like in[2]) and different resources need to be approached differently to use them. For example, nuts need to be gathered from the ground and smashed against a rock before they can be eaten while a banana must be picked from a tree and carefully peeled. It is assumed that originally all agents are naive concerning the rules governing the world. When the agents are released in the world they start off generating sequences of behavior in the hope of finding a sequence that gives access to some rewarding food. Agents that find such a sequence will remember it for future use. This means that after a while agents will alternate between generating new behavior (exploration) and exploiting gathered knowledge (when a food for which a known sequence is stored is encountered).[28] Exploitation behavior will consist of fairly fixed action sequences like picking up a nut, smashing it onto a rock and eating it. Analogically, robot agent will encounter more than one meaningful behavior to learn from a human agent.

The current study focuses on a simplified interaction setting. We assume that only a single demonstrating robot is available. This setting mimics the situation in the cited experiments of Bonnie and Whiten[9,10] where a single, well-trained animal is observed by others or in a variety of human-robot interaction tasks where the human master a skill that has to be learned by a robot. A second simplification is that in both tasks, the imitating robot can see unambiguously the actions of the imitated agent. For this purpose, the following experimental scenario was constructed. Two robots operate in rectangular arena (Figure 1). The first robot performs interchangeably exploring (random action sequences) and exploiting (different predetermined action sequences). The imitating robot observes the actions of the first robot through a top view camera. The advantage of top view camera in multiagent setting is that the actions of the first robot will be seen independently from the viewing angle. For a human-robot interaction task, realistic scenarios can be constructed in which the robot-observer sees the movement of the human agent via a static camera.

The agent that tries to observe and imitate the action sequences of the others sees a continuous sequence of various kinds of actions. The actions which are, from the viewpoint of the observed agents, meaningful and can be parsed into exploring and exploiting parts, are unordered from the viewpoint of a learning agent. There is no *a priori* way for the learning agent to parse the actions of its conspecifics. It cannot know which sequences it should copy and which are to be ignored. It should copy the exploiting action sequences and ignore the exploring action sequences. However, it cannot know where the exploiting sequences start or end. Even if it is assumed that the imitating agent can detect when another agent is rewarded (i.e., the end of an exploiting sequence), it cannot know where the sequence of events that lead to the reward started.

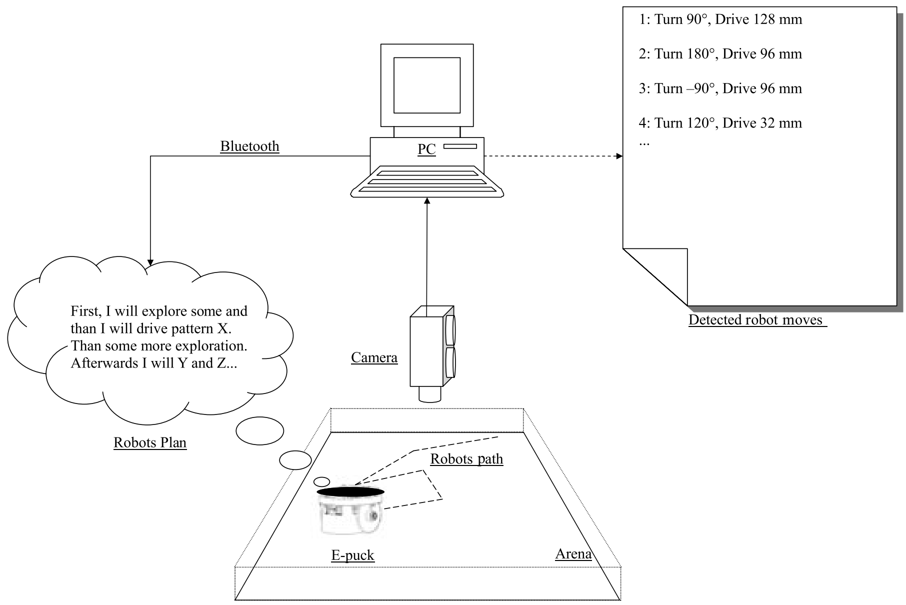

Figure 1. The overall experimental setup.

2.2. Experimental Approach

Parsing a continuous stream of actions is reminiscent of another problem in cognition; that of segmenting a continuous input stream of sounds into words. In a set of seminal articles, Elman[29,30] presented a simple recurrent neural network (Figure 2a) that was able to segment a sequence of letters into words. His network was constructed to take an input letter n and predict the next letter n + 1 in the sequence from n. After some training the network was capable of making good predictions. The network made its predictions based on the co-occurrence statistics available in the data. Furthermore, segmenting the sequence into words was possible using the error signal produced by the network while executing this task. Figure 2b shows an error curve that was produced while the network processed an input stream of letters. Inspecting the error curve, it can be seen that the error is high at the boundaries between words while it drops over the course of a word. This is caused by the fact that while a word unfolds, the next letter becomes more and more predictable with each new letter. On the other hand, at the boundaries between words, the next letter is very hard to predict since it is not determined by the previous one (in the dataset provided to the net). Therefore, the error provides a good clue as to what are recurring sequences in the input, and these correlate highly with words.[31] The network learned which parts of the input should be regarded as meaningful chunks.

A similar solution might be used to let agents autonomously decide what to imitate (see[32] for a related suggestion). An imitating agent could notice that some sequences of actions are consistently executed in the same order. These rewarding

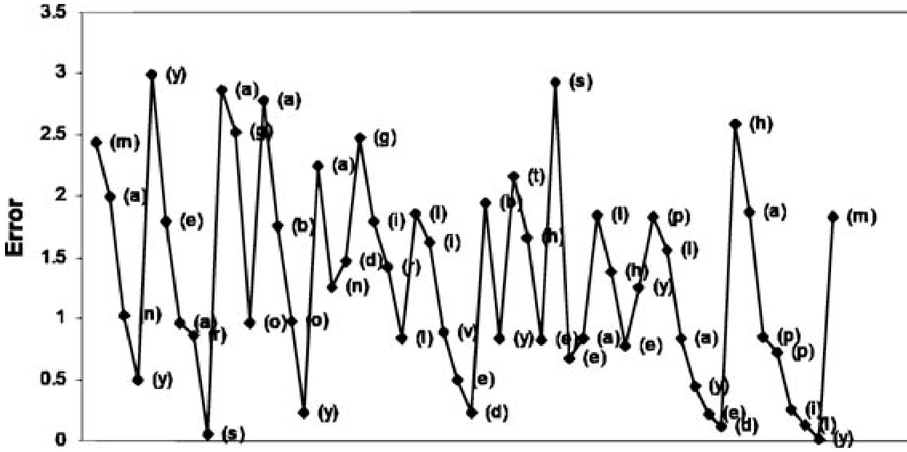

Figure 2. (a) Schematic representation of an Elman net. On each feed-forward sweep, the activation of the Hidden Nodes is copied to the Context Nodes. The Context Nodes are used as an additional set of input nodes on the next feed-forward sweep. (b) The error curve produced by an Elman network that was trained to predict the next letter in a sequence from the previous one. Here, the sequence “many-years-ago-a-boy-and-a-girl-lived-by-the-sea-they-played-happily” is given to the network.

action sequences will be repeated often by the demonstrating robot. Using an Elman network, an observing agent could try to predict the actions of its fellow agents. After a given amount of training the observer could use the error curve to isolate the segments of the input that are interesting to imitate (or a least to evaluate before attempting imitation). Exploiting sequences will be characterized by being predicable (low predicting error).

3. EXPERIMENTAL SCENARIO

3.1. Data Collection and Preprocessing

All experiments reported in this article were conducted using the e-puck robot platform (http://www.e-puck.org). The e-puck is a small mobile robot measuring 70 mm in diameter and 55 mm in height. The robot is equipped with infrared distance sensors that are located around the body at 10°, 45°, 90°, 270°, 315°, and 350° with respect to the heading direction of the robot. Two sensors located at the back of the robot were not used in the reported experiments. The robot was controlled by a personal computer through a Bluetooth interface. A rectangular arena was constructed for the robot which measured about 100 cm × 70 cm. The arena was fenced by cardboard walls which were about 10 cm high. The robots movements were filmed by a Logitech QuickCam camera (http://www.logitech.com) suspended about 160 cm above the floor of the arena. The camera captured the entire arena using 320 × 240 pixels at 10 Hz. The floor of the arena was white. The robot was fitted with a black cap for maximal contrast so that tracking the robot was easy. All

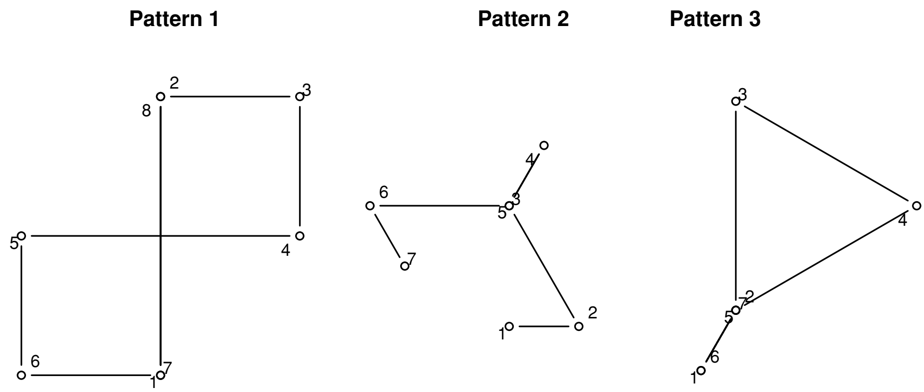

Figure 3. These are the three predefined patterns that could be driven by the robot. Numbers signify the order of execution. See Table I for details about these patterns.

image processing and tracking of the robot was done using RoboRealm software (http://roborealm.com). Processing the images of the camera included correcting for radial distortion. The tracking software provided the approximate location of the center of the robot in each camera frame and its speed.

In this study, the robot could only drive straight on and to turn in place. The robot was allowed to drive four fixed driving distances: 32 mm, 64 mm, 96 mm, and 128 mm. Also 10 turning angles were fixed: $-120°$, $-90°$–$60°$, $-30°$, $0°$, $30°$, $60°$, $90°$, $120°$, and $180°$. Negative values denote counterclockwise turns. Given the constraints imposed on the movements, all movement of the robot was an alternation between turning in place conform to one of the fixed angles, followed by driving one of the fixed distances. The robot iteratively selected a turning angle and a traveling distance to execute.

The e-puck executed two different kinds of behavior. First, the robot could execute exploration behavior. In this mode a turning angle and a traveling distance were selected at random on each iteration. Second, after each turn and drive action the robot could, with a probability of 0.3, select at random one of three patterns to execute. The patterns are depicted in Figure 3. Detailed information on the patterns can be found in Table I. While the robot was driving, the distance sensors were probed each 200 ms to determine whether it was about to hit the walls of the arena. If the robot detected a wall, it aborted its current action and moved away from the wall. In case the robot was executing one of the predetermined patterns of action, the pattern was aborted and a random move was initiated after the avoidance maneuver.

In the experiment reported here, the robot executed 1500 moves consisting of turning and driving. This amounted to about 120 minutes and 74702 image frames. The robot path captured by the camera was preprocessed using R-software.[33] Preprocessing aimed at reconstructing the actions of the robot from the camera images as an observing agent could do. Figure 4 illustrates the preprocessing steps. Note that only a small subset of data has been used for these plots. Plotting an entire dataset would result in graphs that are too cluttered.

Table I. This table lists the turning angles and the driving distances that made up the three predefined patterns plotted in Figure 3.

	Pattern 1		Pattern 2		Pattern 3	
Step	Angle (Degrees)	Distance (mm)	Angle (Degrees)	Distance (mm)	Angle (Degrees)	Distance (mm)
1	0	128	90	32	30	32
2	90	64	−120	64	−30	96
3	90	64	60	32	120	96
4	90	128	180	32	120	96
5	−90	64	60	64	−30	32
6	−90	64	−120	32	180	32
7	−90	128				

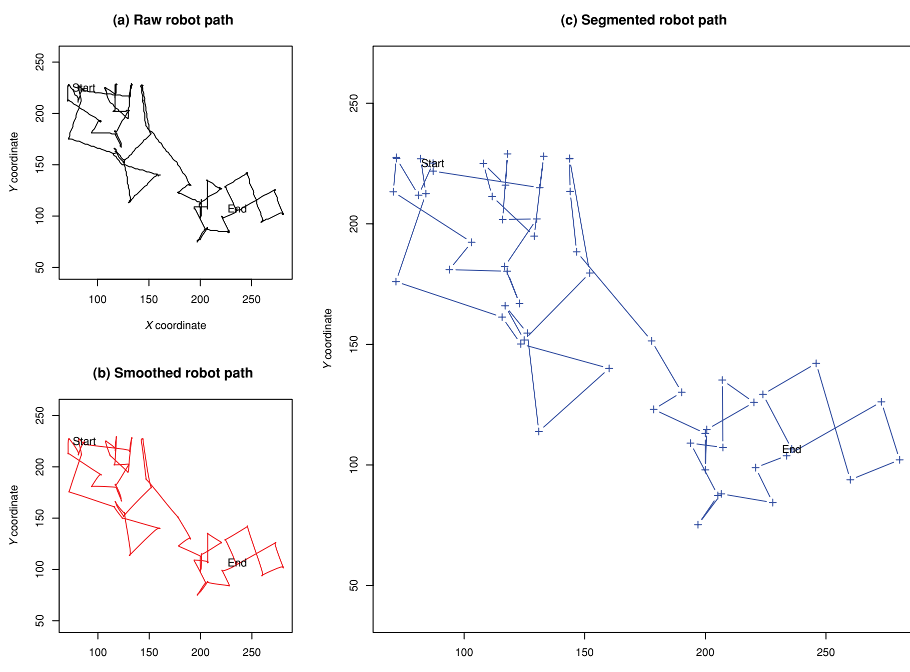

Figure 4. These plots depict the raw (a), smoothed (b), and segmented (c) path the robot drove in experiment 1. For reasons of clarity only a small subset moves of the robot have been plotted.

As can be seen in plot 4a–c, the path of the robot consists of random sequences interwoven with a number of predefined patterns. A cubic smoothing spline was fitted to the raw robot path. The smoothed path is plotted in Figure 4b. The smoothed track was segmented in order to reconstruct the moves the robot executed. Segmenting the track was done based on the detected speed of the robot. Because the top of the e-puck robot is perfectly round, it looks as if it stands completely still, from the

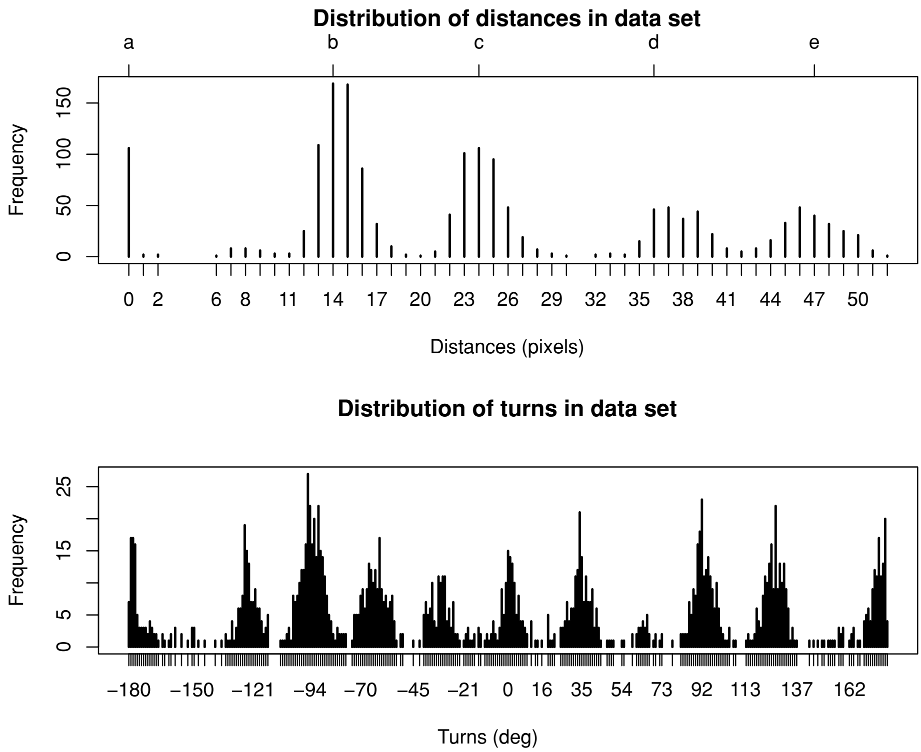

Figure 5. (a) The distribution of traveling distances detected by the camera in the dataset reported in the Results section. Letters a-e refer to five clusters present in this distribution. (b) The distribution of the turning angles in the same data.

viewpoint of the overhead camera, when it turns in place. Therefore, local minima in the speed curve signify points in time when the robot was (probably) executing a turn. The smoothed robot path was segmented at these points in time. Figure 4c depicts the segmented version of the robot path. Next, the segmentation points were connected by straight lines because the robot could only drive straight on between two turning points. From the segmented path, it was trivial to calculate the sequence of the approximate angles the robot turned and the distances it drove.

The final result of the preprocessing step is an approximate reconstruction of the program that has been executed by the robot. This is a record of the actions of the robot as perceived by the camera (or any other onlooker).

Figure 5 depicts the distribution of the angles and distances that were detected by the camera. As can be seen in this graph, the execution and the perception of the moves of the robots was liable to noise. The distribution of the detected distances shows five clusters labeled as a–e. Cluster b–e correspond to the 4 distances the robot could drive in the experiment. Cluster "a" contains very short traveling distances. These are caused by oversegmentation of the robot path and by instances in which the driving of the robot was aborted due to the detection of the walls. The distribution of the turning angles also shows overlapping clusters. The fact that both distributions are characterized as a number of overlapping clusters indicates that the level of noise was relatively high.

Because of the noise, some further processing of the perceived moves was necessary before the data could be fed to an Elman network. The turning angles and the traveling distances were made discrete. Turning angles were mapped onto the nearest 30°. After this, the data contained 12 different turning angles (360/12) instead of the 10 used by the robot. Traveling distances were mapped onto the nearest cluster center (0, 14, 24, 36 or 47). Moves that were classified as belonging to the first cluster (i.e., 0), were discarded from further processing. These final data cleaning steps can be considered as reflecting categorical perception.

3.2. Learning

A generic Elman network was implemented (see[29,31] for more details about the structure of Elman networks). The network had 12 input nodes coding the turning angles and four inputs that coded each of the four traveling distances present in the data. The hidden layer of the network consisted of 30 nodes. All neurons had a sigmoid activation function. The network was trained by presenting it with a turning angle and a traveling distance by setting the corresponding input nodes to one (other nodes were assigned an activation value of 0). So, an input vector consisted of 16 values of which two were set to one to signify the current angle and driving distance. Importantly, each turning angle and traveling distance was assigned a coding input neuron at random. In this way, the coding of the moves was completely abstract. Thus, although the predefined patterns were visually symmetrical (Figure 3), from the viewpoint of the network they were not. The visual symmetry could not be exploited by the network to learn to recognize the patterns. After the presentation of an angle and distance, the network predicted the next turning angle and driving distance the robot would execute. Simple gradient descent (Error Backpropagation) was used to adapt the connection weights of the network after each presented input. After updating the network connections, the next turning angle and traveling distance was presented to the network. In this way, the whole data set was presented 10 times to the network (i.e., 10 epochs of training). This amounted to about 15,000 training trials.

4. RESULTS

The experiment and the training of the network were replicated several times using slightly different parameter settings. However, qualitatively the results were always similar to the ones reported in this section.

Data about the moves executed by the robot can be found in Table II. These data show that a large proportion of the patterns were not completed. This introduced additional noise into the training data. Figure 5 depicts the most important training results. Plot 5a shows the change in the prediction error by the network in the form

Table II. This table lists the number of patterns the robot drove while collecting the data reported in the Results section. Also the number of random moves is listed. The number of patterns and random moves that were terminated because the robot detected a wall are listed separately.

	Termination		
	Normal	Stopped	Total
Pattern 1	39	19	58
Pattern 2	47	21	68
Pattern 3	37	17	54
Total	123	57	180
Random	274	46	320

of a density plot. One can see that after some initial training, there is a bifurcation in the error. After about 2000 trials, most moves of the robot are well predicted (low error) while others are not (high error). This binomial distribution of the error is also clearly visible in plot 5b. A Gaussian Mixture Model[34] with two components was fitted to the error distribution across all trianing trials to obtain an objective threshold to separate trials for which the prediction error was low and trials for which the error was high. At about a value of 0.9, an error value had an equal probability of belonging to either of the two clusters (assuming equal priors for both components). This value is indicated by an arrow in Figure 6b. A Gaussian Mixture Model[35] with two components was fitted to the error distribution to obtain an objective threshold to separate trials for which the prediction error was low. At about a value of 0.9, an error value had an equal probability of belonging to either of the two clusters. This value is indicated by an arrow in Figure 6b. This value was used as a cutoff to identify trials in which the network had made a good prediction. So, trials in which the network predicted the next step with an error lower than 0.9, were considered as trials with a low error. Subsequently, sequences of trials longer than two steps, in which the prediction was better than the cutoff, were identified in the data. The cumulative frequency of all identified sequences is plotted in Figure 6c as a function of the number of training trials. As can be seen, a lot of different sequences were identified ($n = 141$). However, most of these are identified only a few times. Only 10 patterns were recognized 20 times or more while 54 were encountered only once. As can be seen in plot 6c, three patterns are clearly recognized more often than all others. These are indeed the patterns the network was supposed to learn (see labels in plot 6c). So, by analyzing the error curve produced by the Elman network during training, the three predefined patterns could be isolated as being high frequent sequences the network could predict very well.

The fact that the network reliably isolated the three predefined patterns is further analyzed in Figure 7. This plot depicts the frequency of the patterns that were detected more than 19 times over the course of the training. As stated before, the most frequent patterns were the three predefined patterns. However, of equal interest is the fact that the other detected patterns were parts of the goal patterns.

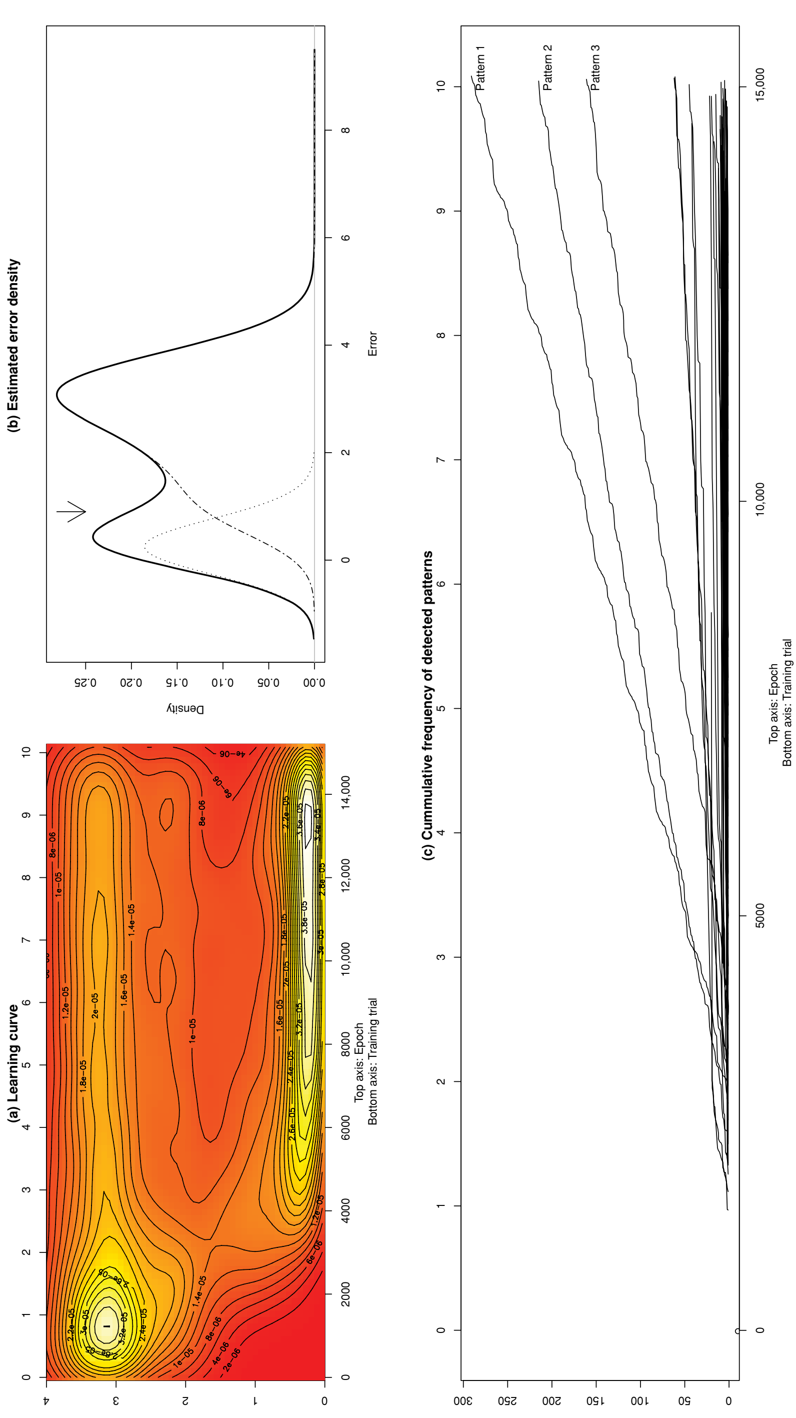

Figure 6. These plots depict the training results of the experiment. (a) A density estimation of the prediction error of the Elman network in individual learning trials in function of the number of trials. (b) A density estimation of the prediction error collapsed across all learning trials. (c) The cumulative frequency of all different sequences of actions that were well predicted by the network (see text for details).

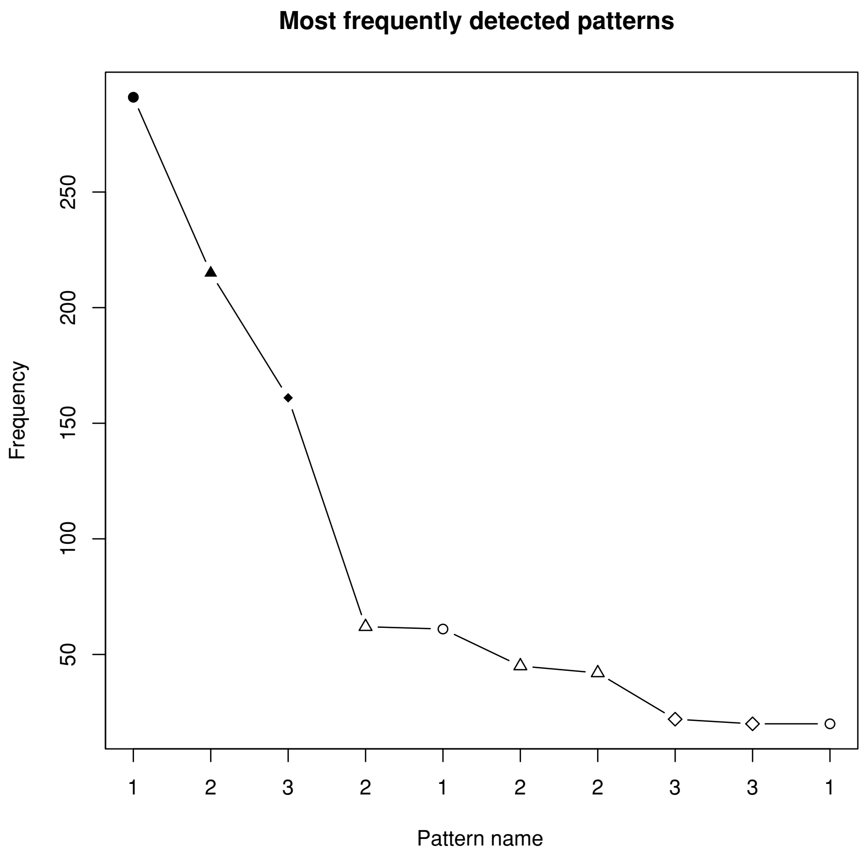

Figure 7. Total frequency of detection of the most frequently detected patterns of actions; pattern 1, pattern 2, pattern 3. Opened markers signify subpatterns of the patterns denoted with a similar shape.

5. WORK IN PROGRESS

At present, simulation studies to support human-robot interaction experiments are carried out. The aim of these experiments is teaching of socially relevant behaviors to children through games with robots. Especially, we target autistic children, since they are marked by delays in social development[21] and will benefit most of such training.

There is evidence that individuals with autistic spectrum disorders do not interpret social messages that motion conveys as typical people do. Moore, Hobson, and Lee[36] found that 14-year-old individuals with autism have deficits in perceiving emotion-related attitudes and subjective states, given the motion cues of a point-light-walker display.[36,37] This finding reveals a deficit in perceiving mental states based on motion cues. Klin[37] found autism-specific differences in people's descriptions of the Heider and Simmel task.[38] Heider and Simmel[38] showed to subjects actions of simple geometric figures. The subjects were asked to narrate the perceived actions. The subjects reported to see that geometric figures had goals, desires, intentions, and emotions. Klin[39] conducted an experiment with 60 participants with autism that were asked to provide narratives describing Heider and Simmel's animation. He found differences at interpreting social motion in autistic people, suggesting the important developmental question of whether people with autism would have typical precursors to this ability to perceive social information in motion cues.

A study by Pierno et al[40] concludes that visuomotor priming proceeds normally in children with autism when primed by a robot. This finding is consistent with other results demonstrating that people with autism perform at normal to superior levels at tasks presented in a repeatable and predictable formats established by a robot or a computer.[41,42]

To involve autistic and typical children in social training with robots we created emotional behaviors and behavioral analysis system that determines the emotional state of an agent[50,51]. For this purpose, the Laban movement analysis method was used[50,51]. The Laban Movement Analysis[35] is a well-established, effective method for observing, describing, annotating, and interpreting human movement. It provides descriptors for the content of human body movements in terms of four factors. We identified the effort factor as being related to the dynamic characteristics of the movement. These characteristics can be translated to measurable movement characteristics such as trajectory, velocity, and acceleration. When three Laban motion factors are combined, they fall into categories called externalized drives. Of special interest is the so-called Passion drive, which combines Time, Weight, and Flow. Using these motion determinants the movement that body takes in space gives a way to detect, express, and model emotional actions.

This method was used to both analyze human movement patterns and create emotional movements for robots. We designed several emotional behaviors and these behaviors were tested with a group of 42 typically developing children. The outcome of the tests showed a good recognition of these basic emotions by the children. For a human observer, the trajectory, the direction with respect to the observer, the speed, and the acceleration of the movement were the determining factors to conclude on the kind of emotion that the robot is expressing. For the response behavior of a robot observer, acceleration profile and the intensity of the movement is essential[50]. The trajectories of two emotional movement behaviors, expressing anger and fear are shown in Figure 8. The trajectories of movements that represented different emotional states, were more complicated and unsuitable to be learned with Elman net. The trajectories resemble the simplicity of the patterns in Figure 3, and the recognition by the robot was a straightforward application of the method proposed in this article.

At present, we are trying to extend the results to trajectories recorded by a three-dimensional human movement using a combination of Wiimote controller and a camera that is positioned in front of the human. These behaviors are to be recognized and imitated by a humanoid robot.

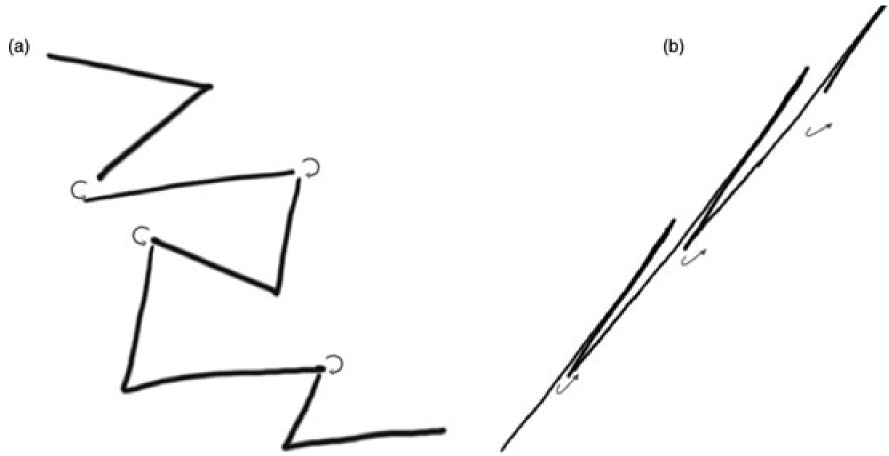

Figure 8. The trajectories of two sample emotional robot behaviors. The upper plot denotes the trajectory of a joyful or happy movement, while the lower plot is derived from fearful behavior.

6. DISCUSSION

We proposed a practical solution for the problem of what to imitate, i.e. which part of the behavior of an agent is relevant and must be learned by another agent, in our case a robot. After revising the animal studies we found out that repetition of a behavior in one or group of agents is a good indication for the importance of the behavior and we show how a simple recurrent network can learn to predict the re-occurence of simple behavioral sequences. We also showed that behavioral sequences with such a simplicity are for instance the movement trajectories of persons and robots expressing some emotional behaviors through whole body motion. Our method is especially suited to work with behavioral sequences that are divided by noisy or undefined behavioral data between the meaningful behavioral sequences. We have used very well defined behavioral sequences, but such can be learned beforehand by a clustering algorithm. The experimental work completed so far showed that an Elman network can be used to reliably identify reoccurring sequences of actions, which have not explicitly been indicated beforehand. In contrast to the simulation studies of Elman,[29–31] the data collected by observing the robot contained substantial amounts of noise (Figure 5). Nevertheless, given enough learning trials, the error curve could be used to reconstruct the three goal patterns. Therefore, it is possible for an observer to extract the interesting parts of the behavior of a demonstrator by using a simple recurrent network.

However, while the proposition of the study was confirmed by the data, their main valor might be to suggest new lines of research and improvements of the current approach. To this end some issues of the study will be discussed.

A first issue is the learning speed demonstrated in the current study. The Elman network needs many trials before a reliable extraction of the patterns is possible (something that was also experienced by Elman in his original studies). This makes the mechanism, exactly as it is implemented in this paper, an unlikely candidate to be used in a multiagent setting or in human-robot interaction scenarios. At present, we are implementing more advanced recurrent learning algorithms. Recent developments by the personal computers ensure substantial computing power of several tera-floating point operations per second (TFLOPS), if graphical processing units (GPUs) are used. Therefore, we are working on a GPU-based parallel implementation of these algorithms.

In the described experiments, one robot is watching another agent (robot or human) continuously (from a favorable perspective). When implementing the current setup with different demonstrators and observers, the noise levels in the perception of each and every robot will rise. In a multiagent task, a robot will have to choose between a number of agents to observe. Some of these will perform very well while others might be still learning themselves. If the robot chooses to imitate an under trained co-agent, it will not be able to learn the task. Instead, it will extract any reoccurring sequences that are coincidently demonstrated. Furthermore, in absence of any reoccurring pattern in the demonstrator's behavior, nothing will be learned. In short, because any learner has access to the behavior of trained as well as untrained individuals, the noise in the perception will increase and the chance of mastering the task will decrease. To avoid such a scenario, more flexible learning mechanisms, adapted to multiagent settings, must be researched. One obvious way to extend the current mechanism is to add reinforcement (or a similar mechanism)—a source of information also used by animals and humans.

Adding reinforcement to the learning mechanism will also reduce the number of learning trials needed. A robot could try out each (partial) sequence of actions it discovers in the behavior of others. If the action sequence is successful according to some measure, the behavior should be consolidated. If not, it should be discarded until further training changes it in some respect. Such a mechanism will speed up learning drastically since a target sequence has to be isolated only once. Furthermore, this way the number of demonstrations of nonadaptive behavior in a set of robots is kept at bay which reduces the noise level.[7]

Acknowledgments

The authors would like to thank the students Frank de Jong, Marco van Beers, Yening Jin, Anh Khoa Nguyen, and Joep Wijnands who worked on the design and testing of robot emotional behaviors.

References

1. Thorndike E. Animal intelligence: an experimental study of the associative process in animals. Psychological Review, Monograph Supplements 1889;2:551–553.
2. Noble J, Franks DW. Social learning mechanisms compared in a simple environment. In: Proceedings of the Eighth International Conference on Artificial Life 2002.
3. Zentall TR. Imitation: definitions, evidence, and mechanisms. Anim Cogn 2006; 9(4):335–353.
4. Leadbeater E, Chittka L. Social learning in insects–from miniature brains to consensus building. Curr Biol 2007;17(16):R703–R713.

5. Boyd R, Richardson PJ. An evolutionary model of social learning: the effect of spatial and temporal variation. In Zentall RR, Galef BJ eds. Social Learning: Psychological and Biological Perspectives. Erlbaum, Hillsdale, NJ 1988;29–48.
6. Nakamaru M, Levin SA. Spread of two linked social norms on complex interaction networks. J Theor Biol 2004; 230(1):57–64.
7. Parisi D. Cultural evolution in neural networks. IEEE Expert: Intelligent Systems and Their Applications 1997;12(4):9–11.
8. Whiten A. The second inheritance system of chimpanzees and humans. Nature 2005: 437(7055):52–55.
9. Bonnie KE, Horner V, Whiten A, de Waal FBM. Spread of arbitrary conventions among chimpanzees: a controlled experiment. Proceedings in Biological Science 2006; 274(1608):367–372.
10. Whiten A, Spiteri A, Horner V, Bonnie KE, Lambeth SP, Schapiro SJ, de Waal FBM. Transmission of multiple traditions within and between chimpanzee groups. Curr Biol Jun 2007;17(12):1038–1043.
11. Price E, Caldwell CA. Artificially generated cultural variation between two groups of captive monkeys, colobus guereza kikuyenis. Behav processes 2007;27(1):13–20.
12. Zentall TR. Imitation: definitions, evidence, and mechanisms. Anim Cogn 2006; 9(4):35–353.
13. Meltzoff AN. Infant imitation after a 1-week delay: long-term memory for novel acts and multiple stimuli. Developmental Psychology 1988a;24:470–476.
14. Meltzoff AN. Infant imitation and memory: nine month olds in immediate and deferred tests. Child Dev (1988b);59:217–225.
15. Fouts GT, Waldner DN, Watson MW. Effects of being imitated and counterimitated on the behaviour of preschool children. Child Dev 1976;47:172–177.
16. Grusec JE, Abramovitch R. Imitation of peers and adults in a natural setting: a functional analysis. Child Development 1982;53:636–642.
17. Abravanel E, Gingold H. Learning via observation during the second year of life. Developmental Psychology 1985;21:614–623.
18. Barr R, Dowden A, Hayne H. (1996). Developmental changes in deferred imitation by 6- to 24-month-old infants. Infant Behavior and Development, 19, 159–171.
19. Meltzoff, AN. Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. Developmental Psychology 1995;31:838–850.
20. Pini G, Tuci E, Dorigo M.: Evolution of social and individual learning in autonomous robots. Ecal Workshop: Social Learning in Embodied Agents 2007.
21. Barakova EI, Gillessen J, Feijs L. Social training of autistic children with interactive intelligent agents. Journal of Integrative Neuroscience, 8(1):23-34, 2009.
22. Billard B, Robins K, Dautenhahn J. Nadel. Building Robota, a Mini- Humanoid Robot for the Rehabilitation of Children with Autism. RESNA Assistive Technology Journal 2007;19(1):37–49.
23. Nadel J. "Early Imitation and a Sense of Agency," Proc. 4th Intl. Workshop on Epigenetic Robots 2004.
24. Dautenhahn K, Nehaniv C. An agent-based per- spective on imitation. In: Dautenhahn K, Nehaniv CL, editors. Imitation in Animals and Artifacts. MIT Press; 2007. MIT Press, Cambridge, MA, USA.
25. Belpaeme T, de Boer B, Jansen B. The dynamic emergence of categories trough imitation. In: Dautenhahn K, Nehaniv CL, editors. Imitation in Animals and Artifacts. MIT Press; 2007. In Imitation in animals and artifacts, Kerstin Dautenhahn and Chrystopher L. Nehaniv (Eds.). MIT Press, Cambridge, MA, USA.
26. Breazeal C, Scassellati B. Robots that imitate humans. Trends Cogn Sci 2002;6(11):481–487.
27. Carpenter M, Call J. The question of what to imitate: inferring goals and intentions from demonstrations. In: Dautenhahn K, Nehaniv CL, editors. Imitation in Animals and Artifacts. MIT Press; 2007. MIT Press, Cambridge, MA, USA.

28. Thrun SB. Efficient exploration in reinforcement learning. Technical report, Pittsburgh, PA, USA 1992.
29. Elman JL. Learning and development in neural networks: the importance of starting small. Cognition 1993;48(1):71–99.
30. Elman JL. Distributed representations, simple recurrent networks, and grammatical structure. Machine Learning 1991;7:195–225.
31. Elman JL. Finding structure in time. Cogn Sci 1990;14(2):179–211.
32. Moga S, Gaussier P. Artificial neural network for sequence learning. In: International Joint Conference on Artificial Intelligence 2003. Proceedings of the 18th international joint conference on Artificial intelligence Acapulco, Mexico 1507-1508 2003 Morgan Kaufmann Publishers Inc.
33. Team RDC. R: A language and environment for statistical computing (2007) Proceedings of the 18th international joint conference on Artificial intelligence %C Acapulco, Mexico %P 1507-1508 %D 2003 %I Morgan Kaufmann Publishers Inc. ISBN 3–900051-07–0.
34. Fraley C, Raftery AE. MCLUST version 3 for R: Normal mixture modeling and model-based clustering. Technical Report 504, University of Washington, Department of Statistics, 2006.
35. Laban R, Lawrence FC. Effort. London: Macdonald & Evans Ltd; 1974.
36. Moore D, Hobson RP, Lee A. Components of person perception: An investigation with autistic, non-autistic retarded and typically developing children and adolescents. Br J Dev Psychol 1997;15:401–423.
37. Blake R, Turner LM, Smoski MJ, Pozdol SL, Stone, WL Visual recognition of biological motion is impaired in children with autism. Psychol Sci 2003;14:151–157.
38. Heider F, Simmel M. An experimental study of apparent behavior. Am J Psychol 1944;57:243–259.
39. Klin A. Attributing social meaning to ambiguous visual stimuli in higher-functioning autism and Asperger syndrome: The Social Attribution Task. J Child Psychol Psychiatry 2000;41:831–846.
40. Pierno AC, Morena Mari, Dean Lusher, Umberto Castiello, Robotic movement elicits visuomotor priming in children with autism Neuropsychologia 2008;46(2):448–454.
41. Hobson P. The Cradle of Thought. Oxford:Oxford University Press; 2004.
42. Hobson P. Understanding persons: The role of affect. In: S Baron-Cohen, Tager-Flusberg H, Cohen DJ, editors. Understanding other minds: Perspectives from autism. Oxford, UK: Oxford University Press;1993. pp 204–227.
43. Dieter Vanderelst, Rene M.C. Ahn, Emilia I. Barakova Simulated Trust: A cheap social learning strategy. Theoretical Population Biology 2009;76(3):189–196.
44. Brok JCJ, Barakova EI Engaging autistic children in imitation and turn-taking games with multiagent system of interactive lighting blocks. In H S Yang et al. (Eds.): ICEC 2010, LNCS 6243, pp 115–126, 2010.
45. Emilia I. Barakova and Tino Lourens Expressing and interpreting emotional movements in social games with robots. Personal and Ubiquitous Computing 2010;14:457–467.
46. Tino Lourens, Roos van Berkel, and Emilia Barakova Communicating emotions and mental states to robots in a real time parallel framework using Laban movement analysis. Robotics and Autonomous Systems 2010;58:1256–1265.
47. Dieter Vanderelst , Rene M.C. Ahn, Emilia I. Barakova Simulated Trust: A cheap social learning strategy. Theoretical Population Biology, Vol 76, no 3, pp 189–196 November 2009.
48. J. C. J. Brok, E. I. Barakova Engaging autistic children in imitation and turn-taking games with multiagent system of interactive lighting blocks.in H.S. Yang et al. (Eds.): ICEC 2010, LNCS 6243, pp 115–126, 2010.
49. Emilia I. Barakova and Tino Lourens Expressing and interpreting emotional movements in social games with robots. Personal and Ubiquitous Computing 2010;14:457–467.
50. Tino Lourens, Roos van Berkel, and Emilia Barakova Communicating emotions and mental states to robots in a real time parallel framework using Laban movement analysis. Robotics and Autonomous Systems 2010;58:1256–1265.