Author's version of the article published in K. Dautenhahn (Eds),
Journal of Interaction Studies, 8(3), 2007. John Benjamins.

What is the Teacher's Role in Robot Programming by Demonstration? Toward Benchmarks for Improved Learning

Sylvain Calinon and Aude G. Billard
Learning Algorithms and Systems Laboratory (LASA)
Ecole Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Lausanne, Switzerland

Abstract

Robot programming by demonstration (RPD) covers methods by which a robot learns new skills through human guidance. We present an interactive, multimodal RPD framework using active teaching methods that places the human teacher in the robot's learning loop. Two experiments are presented in which observational learning is first used to demonstrate a manipulation skill to a HOAP-3 humanoid robot by using motion sensors attached to the teacher's body. Then, putting the robot through the motion, the teacher incrementally refines the robot's skill by moving its arms manually, providing the appropriate scaffolds to reproduce the action. An incremental teaching scenario is proposed based on insights from various fields addressing developmental, psychological, and social issues related to teaching mechanisms in humans. Based on this analysis, different benchmarks are suggested to evaluate the setup further.

In a *robot programming by demonstration* (RPD) framework, a robot learns new skills through the help of a human instructor (Billard & Siegwart, 2004). Traditionally, RPD tends to consider the human user as an expert model who performs a task while the robot *observes passively* the demonstration (Ikeuchi & Suchiro, 1992; Kuniyoshi, Inaba, & Inoue, 1994). However, in humans, teaching is a social and bidirectional process in which teacher and learner are both active. Instead of considering the teacher solely as a model of successful expert behavior, recent work has referred to the teacher-learner couple as a

We gratefully acknowledge Chrystopher L. Nehaniv, Aris Alissandrakis, Joe Saunders, Nuno Otero and Kerstin Dautenhahn for the useful exchanges of thoughts concerning social cues and feedback as well as the user experience and evaluation issues tackled by the *Cogniron* project. We would also like to acknowledge the four anonymous reviewers for their very useful comments on an earlier version of the manuscript.

The work described in this paper was supported in part by the *Secretariat d'Etat a l'Education et la Recherche Suisse* (SER), under Contract FP6-002020, Integrated Project *Cogniron* of the European Commission Division FP6-IST Future and Emerging Technologies, and by the Swiss National Science Foundation, through grant 620-066127 of the SNF Professorships program.

Figure 1. Different modalities are used to convey the demonstrations and scaffolds required by the robot to learn a skill. The user first demonstrates the whole movement using motion sensors (*left*) and then helps the robot refine its skill by kinesthetic teaching (*right*), that is, by *embodying* the robot and putting it through the motion.

team that engages in joint problem solving (Breazeal et al., 2004) and uses active teaching methods to put the human teacher "in the loop" of the robot's learning (Dautenhahn, 1998).

In previous work, we developed a probabilistic framework for extracting the relevant components of a task by observing multiple demonstrations of it (Calinon, Guenter, & Billard, 2007). The system is based on a probabilistic encoding of several demonstrations provided to the robot to generalize the learned skills to different contexts using a *Gaussian mixture model* (GMM). From these experiments, we noticed the importance of adding social components to the teaching paradigm, not only in the user interface but also in the teaching process (Calinon & Billard, 2007). In contrast to the traditional RPD approach, we thus adopted the perspective that the transfer of skills can draw advantages of several psychological factors that the user might encounter during teaching. Thus, the teacher is no longer considered solely a model of expert behavior but becomes an active participant in the learning process (Yoshikawa, Shinozawa, Ishiguro, Hagita, & Miyamoto, 2006). First, we suggest using different modalities to produce the demonstrations (Figure 1). We then suggest following an incremental learning approach to allow the teacher to gradually see the results of the demonstrations.

Related work

Extraction of constraints in robot programming by demonstration

A single demonstration is usually not enough to extract a task's particulars and goals. Compared to batch learning, the benefit of an incremental (Pardowitz, Zoellner, Knoop, & Dillmann, 2006; Saunders, Nehaniv, & Dautenhahn, 2006; Breazeal et al., 2004) or a dynamical learning approach (Ito, Noda, Hoshino, & Tani, 2006; Vijayakumar, D'souza, &

Schaal, 2005) is that the interaction can be performed online, which allows us to observe the results immediately. Different approaches based on multiple observations have been proposed to extract the constraints of a task at a *symbolic level* (Breazeal et al., 2004; Pardowitz et al., 2006; Steil, Röthling, Haschke, & Ritter, 2004; Nicolescu & Mataric, 2003). In Nicolescu and Mataric (2003), a graph-based approach is used to generalize a high-level skill across multiple demonstrations, where generalization takes place at the level of the topological representation of the graph. The advantage of this approach is that high-level tasks consisting of sequences of symbolic cues can be learned efficiently through an interactive process. However, because of its symbolic nature, the method relies on the predetermination of the observed cues and on the efficiency of the segmentation process. Our approach uses a similar paradigm and is complementary with the aim of generalizing a skill represented at a *trajectory level*. The proposed system provides a generic model-based approach in which the generalization and extraction of constraints is performed at a trajectory level through regression.

Scaffolding issues in robot programming by demonstration

Kinesthetic teaching provides a way of supporting the robot in its reproduction of the task (Figure 1). By using scaffolds, the user provides support to the robot by manually articulating a decreasing subset of motors. The scaffolds progressively fade away and the user finally lets the robot perform the task on its own, allowing the robot to experience the skill independently. By taking inspiration from the human tutelage paradigm, Breazeal et al. (2004) showed that a socially guided approach could improve both the human-robot interaction and the machine learning process by taking into account human benevolence. In their work, they highlight the role of the teacher in organizing the skill into manageable steps and maintaining an accurate mental model of the learner's understanding. However, the tasks considered are mainly built on discrete events organized hierarchically. Our work shares similarities with theirs in terms of the tutelage paradigm, but we focus on learning continuous motion trajectories and actions on objects at a trajectory level instead of considering discrete events.

Saunders et al. (2006) provided experiments where a wheeled robot is teleoperated through a screen interface to simulate a *molding* process, that is, by letting the robot experience sensory information when exploring its environment through the teacher's support. Their model uses a memory-based approach in which the user provides labels for the different components of the task to teach hierarchically high-level behaviors. Our work shares similar ideas, but follows a model-based approach. The drawback of using teleoperation is avoided by letting the user directly move the robot's arms while the robot records sensory information through its motors encoders.

Therefore, the teacher uses scaffolding to let the robot gradually generalize the skill for an increasing range of contexts. Knowing that, it may be important to reproduce the acquired skill after each demonstration to help the teacher prepare the following demonstration according to the outcome. This shares similarities with the human process of refining the understanding of a task through practice and corrective feedback, by combining incrementally new information with previous understanding. It thus suggests the use of incremental learning approaches when teaching humanoid robots, allowing them to extend, refine and elaborate their understanding of the skill.

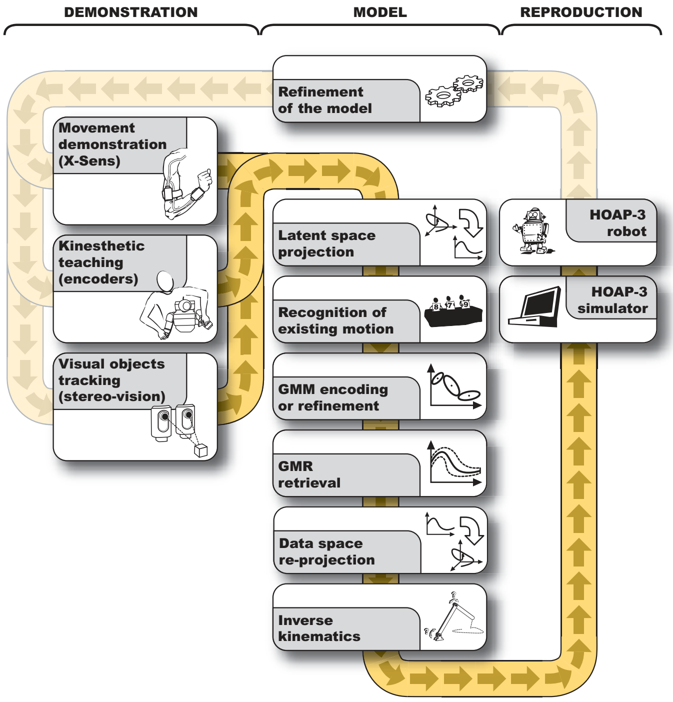

Figure 2. Information flow across the complete system.

Regression issues in robot programming by demonstration

By considering a trajectory as a set of positions ξ observed at time t, the regression problem consists of computing a smooth estimation of $p(\xi|t)$. Among the various regression techniques proposed in statistics, *locally weighted regression* (LWR) has proved appealing for robot control (Atkeson, Moore, & Schaal, 1997) by combining the simplicity of linear least-squares regression and the flexibility of nonlinear regression. As LWR is a memory-based approach, computation also became limited to the memory capacity of the robot when faced with large training sets. To prevent this limitation, further work mainly concentrated on shifting the memory-based approach to a model-based approach and on shifting the batch learning process to an incremental learning strategy (Schaal & Atkeson, 1998; Vijayakumar et al., 2005). Our approach follows a similar trend by using a *Gaussian mixture model* to represent the joint distribution of data $\{t, \xi\}$, and *Gaussian mixture regression* (GMR) to estimate $p(\xi|t)$. This allows the system to deal with encoding, recognition and reproduction issues in a single framework and to use the *expectation-maximization* (EM) algorithm to train the model.

Experimental setup

System overview

In Calinon et al. (2007), we presented an approach based on *principal component analysis* (PCA) and a *Gaussian mixture model* (GMM) to build a probabilistic representation of the movement. As training was performed in batch mode, refinement of the model was not possible without using historical data. To remedy this we recently proposed the use of an incremental learning algorithm to train our model (Calinon & Billard, 2007). Figure 2 presents the principles of the system. The robot builds a model of the task constraints by observing the user demonstrating the manipulation skill using different initial positions of the objects. After each demonstration, the robot reproduces a generalized version of the task by probabilistically combining the various extracted constraints. Watching the robot reproducing the task after each demonstration helps the teacher evaluate the reproduction and generalization capabilities of the robot. The teacher can thus detect the portions of the task that require refinements or that have not been correctly *understood* by the robot, helping the teacher prepare further demonstrations. The model is refined incrementally after each demonstration, and the user decides to stop the interaction when the robot has correctly learned the skill.

By extracting the variations and correlations through GMR, we detect at each time step along the trajectory what the relevant variables composing the task are and how the different variables are correlated. For some tasks, relevant and irrelevant variables are clearly separated (Alissandrakis, Nehaniv, Dautenhahn, & Saunders, 2005), defined a priori, or are constant throughout the task. Here, we consider the most general case where different levels of constraints are allowed, which can freely change during the skill. We believe that representing the task constraints in a binary manner (relevant versus irrelevant features) is not appropriate for continuous movements. Indeed, some goals require different precisions, that is, they can be described with different degrees of invariance. For example, the movement used to drop a piece of sugar in a tiny cup of coffee is more constrained than the movement to drop a bouillon cube in a large pan.

Hardware

The experiments are conducted using a Fujitsu HOAP-3 humanoid robot with 28 degrees of freedom (DOFs), of which only the 16 DOFs of the upper torso are required in the experiments. Two webcams in its head are used to track objects in 3D Cartesian space based on color information. The objects to track are predefined in a calibration phase. Alternatively, the initial positions of the objects can also be recorded by a molding process where the teacher grabs the robot's arm, moves it toward the object and puts the robot's palm around the object. When the object touches its palm, the robot *feels* the object by using a force sensor. It then briefly grasps and releases the object while registering its position in 3D Cartesian space.

Two different modalities are used to convey the demonstrations (Figure 1). First we use motion sensors attached to different body parts of the user. The user's movements are recorded by 8 *X-Sens* motion sensors attached to the torso, upper-arms, lower-arms, hands (at the level of the fingers) and back of the head. Each sensor provides the 3D absolute orientation of each segment by integrating the 3D rate-of-turn, acceleration and earth-magnetic field at a rate of 50 Hz and with a precision of 1.5 degrees. For each joint, a rotation matrix is defined as the orientation of a distal limb segment expressed in the frame of reference of its proximal limb segment. The kinematics motion of each joint is then computed by decomposing the rotation matrix into joint angles.

We then use the motor encoders of the robot to record information while the teacher moves the robot's arms. The teacher selects what motors to control manually by slightly moving the corresponding motors just a few milliseconds before the reproduction starts. The selected motors are set to passive mode, which allows the user to move freely the corresponding degrees of freedom while the robot executes the task. In this way, the teacher can provide partial demonstrations while the kinematics of each joint motion are recorded at a rate of 1000 Hz. The trajectories are resampled to a fixed number of points $T = 100$. The robot is provided with motor encoders for every DOF, except for the hands and the head actuators.

Probabilistic model

To extract constraints from a set of trajectories $\{t, \xi\}$ consisting of temporal and spatial values, we model the joint probability $p(t, \xi)$ with a mixture of Gaussian distributions, trained incrementally by expectation maximization. A generalized version of the trajectories can then be computed by estimating $E[p(\xi|t)]$, where $E[\cdot]$ represents the expectation. The constraints are extracted by estimating $E[cov\,(p(\xi|t))]$, where $cov(\cdot)$ represents the covariance. If multiple constraints are considered (e.g., considering actions $\xi^{(1)}$ and $\xi^{(2)}$ on two different objects), the resulting constraint is computed by first estimating $p(\xi|t) = p(\xi^{(1)}|t) \cdot p(\xi^{(2)}|t)$ and then using $E[p(\xi|t)]$ to reproduce the skill. For a complete description of the algorithms, the interested reader is referred to Calinon et al. (2007) for a detailed description of the extraction of constraints and to Calinon and Billard (2007) for the incremental learning process.

Through the use of Gaussian mixture regression, trajectory constraints take the form of smooth generalized trajectories with associated covariance matrices describing the variations and correlations across the different variables. Trajectory constraints are defined for each object considered in the scenario, which allows the system to model simultaneously gestures and actions on objects. This probabilistic representation of the trajectory constraints is then used to reproduce the task in new conditions, that is, with new positions of objects that have not been used to demonstrate the skill. Thus, extracting not only a generalized movement from the demonstrations, but also variability and correlation information, allows the robot to use its experience in changing environmental conditions. Using GMM, this can be setup in an adaptive way without drastically increasing the complexity of the system when new experiences are provided. Thus, the model does not use historical data and is flexible enough to adapt to new demonstrations (Calinon & Billard, 2007).

Experiments

We present two experiments to show that different data representations can be used to model the skill. For the first experiment, the robot collects the joint angle trajectories of the two arms. The aim of this first experiment is threefold: (1) To show that the method is generic (the joint angles are the lowest-level data that the robot can collect); (2) to show that the system can deal with bimanual coordination by learning the correlations across the joint angles; and (3) to show that the system can efficiently encode high-dimensional data by projecting them in a latent space of motion. For the second experiment, the robot also collects the joint angle trajectories of the right arm, but this time a direct kinematics algorithm is used to convert the joint angles into a 3D Cartesian path for the right hand. Then, the Cartesian path is computed relative to each object. By doing so, a smaller

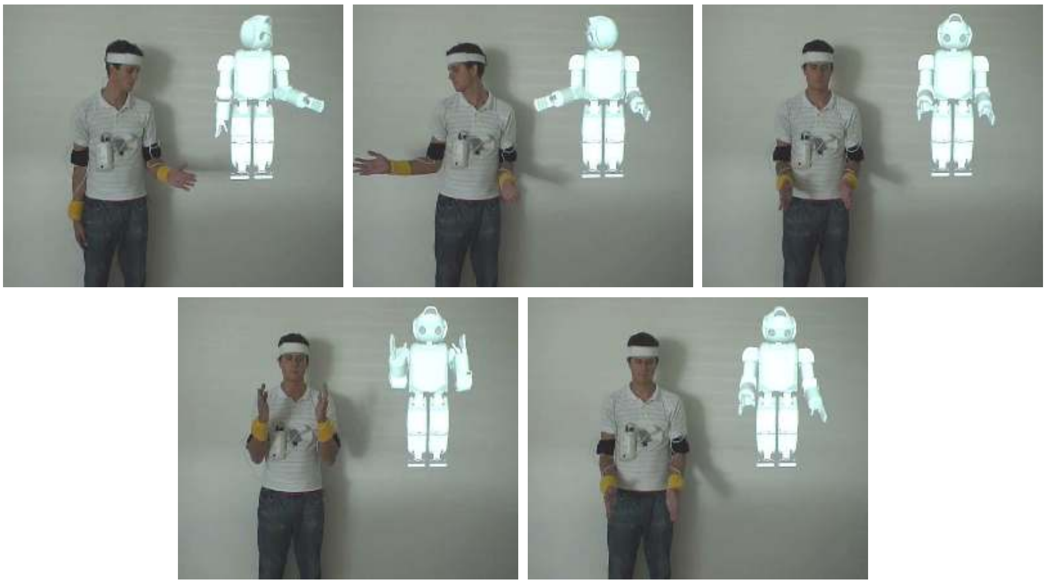

Figure 3. Illustration of the use of motion sensors to acquire gestures from a human model. A simulation of the robot is projected behind the user to show how the robot collects the joint angles during the process. The gesture is similar to the one used in *Experiment 1*, except that the user does not face the robot and only mimics the grasping of the object.

number of demonstrations is required to learn the task. Indeed, for a task involving the manipulation of objects, the position of the hand relative to the objects in the environment is usually highly relevant. The aim of this second experiment is threefold: (1) To show that adding prior knowledge to the representation of the variables has the advantage of speeding up learning (i.e., the system does not need to learn by itself the direct kinematics of the arm); (2) to show that this representation allows the system to learn manipulation skills using multiple objects; and (3) to show that the learned skill can be generalized to various initial positions of objects.

Experiment 1: Learning bimanual coordination gesture

This experiment shows how a bimanual skill can be taught incrementally to the robot in joint angle space using observation and scaffolding. The task consists of grasping and moving a large foam die (Figures 1 and 3). Starting from a rest posture, the left arm is moved first to touch the left side of the die, with the head following the motion of the hand. Then a symmetrical gesture is performed with the right arm. When both hands grasp the object, it is lifted and pulled back on its base, with the robot's head turned toward the object (Figure 3). The teacher wearing the motion sensors performs the first demonstration of the complete task. This allows the user to demonstrate the full gesture by controlling simultaneously 16 joint angles. These joint angles are then projected into a subspace of lower dimensionality using PCA. After observation, the robot reproduces a first generalized version of the motion. This motion is then refined by kinesthetically helping the robot perform the gesture, that is, by physically moving its limbs during the reproduction attempt. The gesture is refined partially by guiding the desired DOFs while

1st reproduction attempt 3rd reproduction attempt 6th reproduction attempt

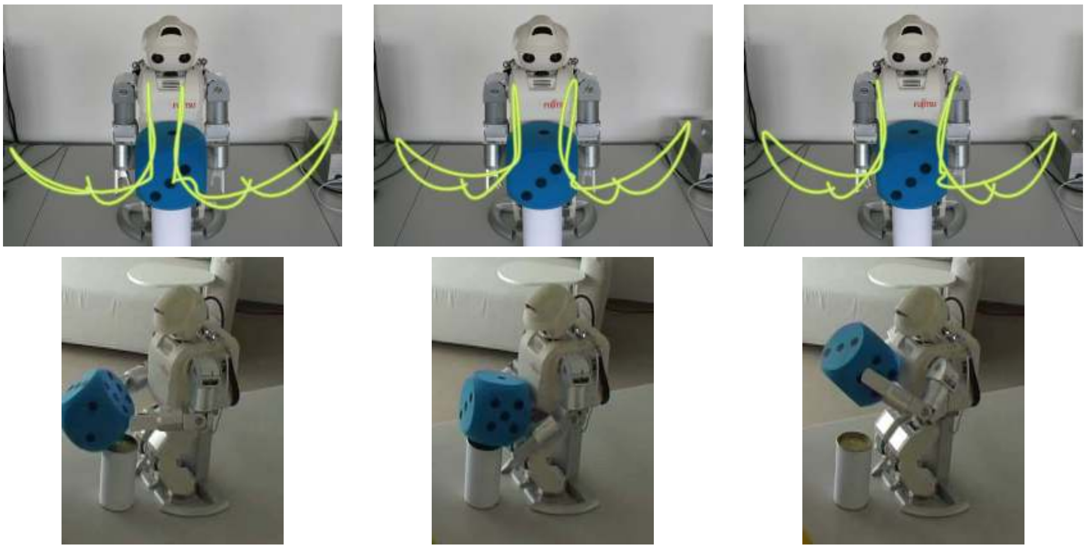

Figure 4. Reproduction of the task after the 1st, 3rd and 6th demonstration. In the first attempt, the robot hits the die when approaching it. In the third attempt, the robot's skill gets better but the grasp is still unstable. In the sixth attempt, the die is correctly grasped. The trajectories of the hands are plotted in the first row, and corresponding snapshots of the reproduction attempts are represented in the second row.

the robot controls the remaining DOFs. By using this method, the teacher can only move a limited subset of DOFs by using his or her two arms. It means that the user can move the two shoulders and the two elbows of the robot simultaneously, but the remaining DOFs (head, wrists and hands) are controlled by the robot. The user first selects the DOFs to be control by moving the corresponding joints. The robot detects the motion and sets the corresponding DOFs to passive mode. Then, the robot reproduces the movement while recording the movement of the limbs controlled by the user.

Results of the experiment are presented in Figures 4 and 5. We see that the resulting paths of the hands are similar to the ones demonstrated by the user (Figure 3), even if training is performed in a subspace of motion where joint angle trajectories have been projected. The system finds five principal components and five Gaussian components to efficiently represent the trajectories in this latent space. After the first demonstration of the movement while wearing motion sensors, the robot can only reproduce a smoothed version of the joint angles produced by the user. Because the user's and robot's bodies differ (the robot is smaller than the user, but the size of the die does not change), the robot approaches the die with its hands too close to grasp it. When trying to reproduce the skill, the robot hits the die by moving its left hand first, making the die fall before moving its right hand. Observing this, the teacher progressively refines the model by providing appropriate scaffolds, that is, by controlling the shoulders and the elbows of the robot while reproducing the skill so that it may grasp the die correctly. In the third reproduction attempt, the robot lifts the die awkwardly. In the sixth attempt, the robot skillfully reproduces the task by itself (Figure 5). Therefore, the user decides to stop the teaching process.

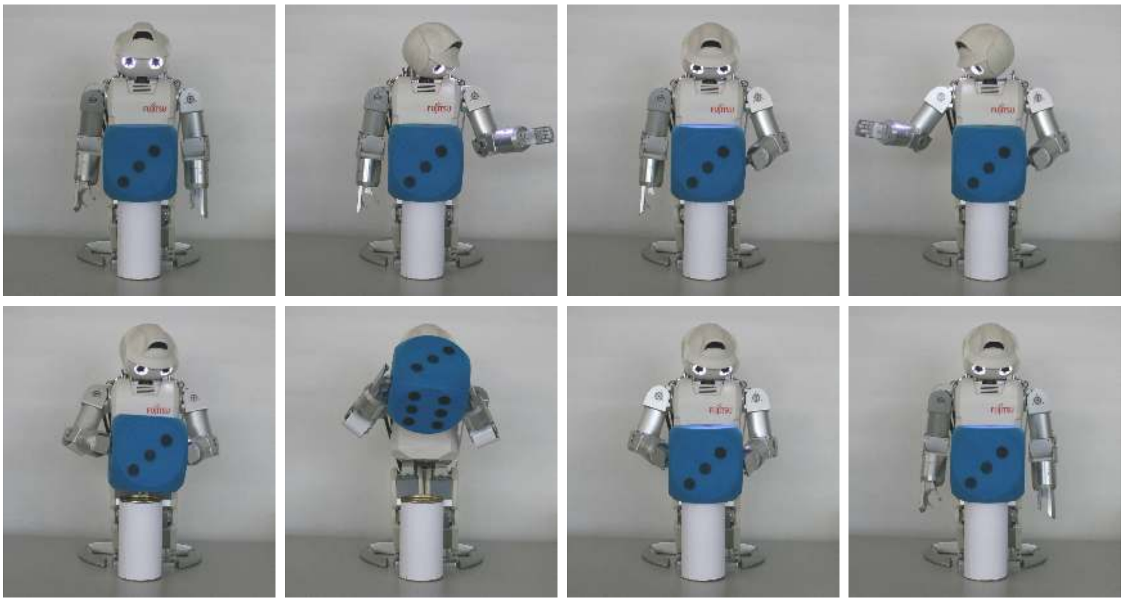

Figure 5. Snapshots of the sixth reproduction attempt.

Experiment 2: Learning the affordances and effectivities of objects

This experiment shows how the system can learn incrementally manipulation skills in 3D Cartesian space. Through the teacher's support, the robot extracts and combines constraints related to different objects in the environment. The aim of the task is to grasp a red cylinder with the right hand and place it on a yellow cube (Figure 6). The robot learns simultaneously the *affordances* of the two objects (i.e., the red cylinder can be stacked on the yellow cube) and the associated *effectivities* (i.e., how the robot should use its body to grasp and bring the cylinder on top of the cube without hitting any object). The specific constraints related to the two objects are extracted by varying the demonstrations provided to the robot, that is, by starting with different initial positions of the objects. Thus, the efficiency of the reproduction mainly relies on the ability of the teacher to provide appropriate scaffolds when demonstrating the task by using sufficient and appropriate variability across the demonstrations. By doing so, the system is able to generalize and reproduce the skill in new situations that have not been used to demonstrate the task. After each demonstration, the robot tries to reproduce the skill in two different situations. This helps the teacher evaluate how well the robot can generalize in new situations.

The system simultaneously learns the 3D Cartesian trajectories relative to the different objects in the environment and the actions used to move these objects. As the robot's arm is kinematically redundant, we define a posture by the position of the right hand in a 3D Cartesian space and by an additional angular parameter α defining the elevation of the elbow (angle formed by the elbow and an horizontal plane). For each demonstration, the hand path and a trajectory of angle α fully describe the gesture. A generalized version of the hand paths and of the α trajectories are then used to compute (by geometrical inverse kinematics) the joint angle trajectories required to control the robot. The advantage of this

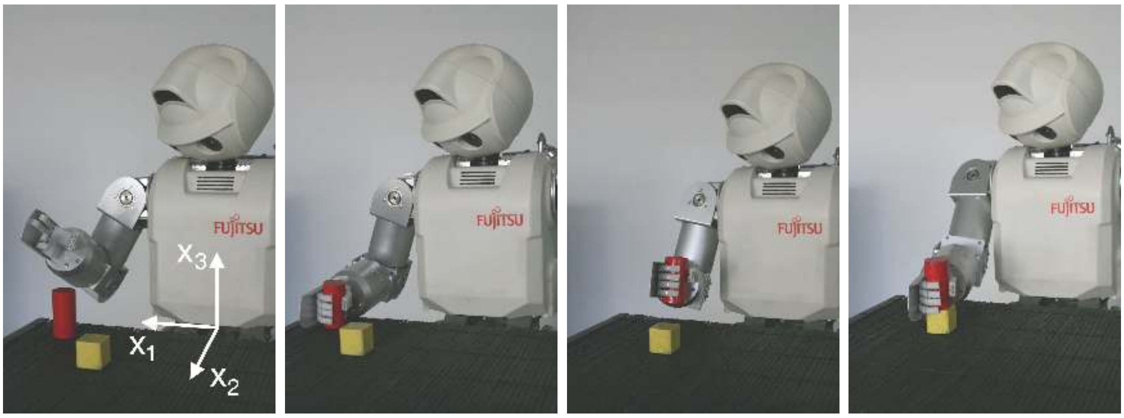

Figure 6. Reproduction of actions on objects. The purpose of the task is to grasp the red cylinder and to bring it on top of the yellow cube. The first snapshot shows the 3D Cartesian frame of reference used in the experiment.

method over other inverse kinematics algorithm is that a constraint based on the observation of the user performing the skill is used as an optimization factor, which produces movements that look natural.

There are $n = 6$ demonstrations of the task where each trajectory is rescaled to $T = 100$ time steps. (After six demonstrations, the user estimated that the robot understood the task correctly.) The total number of observations is thus given by $N = n \times T$. For each demonstration $i \in \{1, \ldots, n\}$, the collected data consists of the initial positions of the M objects $\{o_i^{(h)}\}_{h=1}^{M}$ and of the set of variables $\{x_{i,j}, \alpha_{i,j}\}_{j=1}^{T}$ corresponding to the absolute right hand paths and to the evolution of the parametric angle α. After each demonstration, the trajectories of the right hand relative to the M different objects are computed

$$x_{i,j}^{(h)} \quad = \quad x_{i,j} - o_i^{(h)} \qquad \left| \begin{array}{c} \forall h \in \{1, \ldots, M\} \\ \forall i \in \{1, \ldots, n\} \\ \forall j \in \{1, \ldots, T\}. \end{array} \right.$$

The constraints of the task are extracted from the training set $\{x^{(1)}, \ldots, x^{(M)}, \alpha\}$ to which encoding and generalization are applied separately. By using two objects, a generalized version of the trajectories is thus given by $\{\hat{x}^{(1)}, \hat{x}^{(2)}, \hat{\alpha}\}$ and the associated covariance matrices $\{\hat{\Sigma}^{(1)}, \hat{\Sigma}^{(2)}, \hat{\Sigma}^{\alpha}\}$.

Results for encoding, generalization and extraction of constraints are presented in Figure 7 (Figure 6 for the x_1, x_2, x_3 directions). For the first object, we see that the trajectory $\hat{x}^{(1)}$ is highly constrained when grasping the object between time steps 20 and 40 (a tight envelope representing the constraints around the generalized trajectory). The constraints are tighter for the first two variables $\hat{x}_1^{(1)}$ and $\hat{x}_2^{(1)}$, defining the movement with respect to the surface of the table. This is consistent with the shape of the object to grasp (a cylinder placed vertically on the table). Indeed, the form and orientation of the object enable it to be grasped with more variability on the third axis x_3, defining the vertical movement. When observing the constraints associated with $\hat{x}_3^{(1)}$, we also see that between time steps 40 and 70, the generalized trajectories of the hand holding the object follows

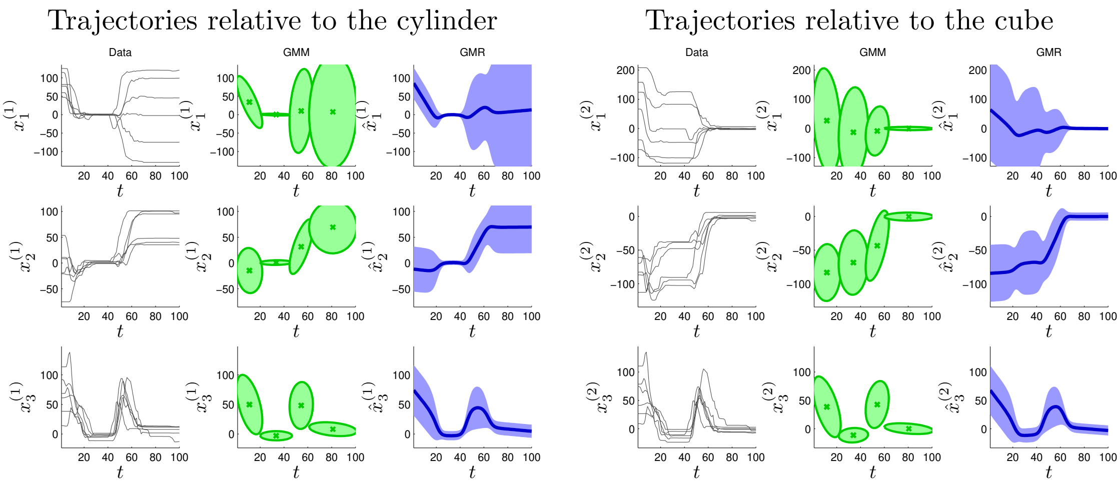

Figure 7. Generalization and extraction of the constraints for the Cartesian trajectories relative to the two objects, after six demonstrations of the skill. For each object, we have *Left:* The six demonstrations observed consecutively; *Middle:* The *Gaussian mixture model* (GMM) of four components used to incrementally build the model; and *Right:* A representation of the generalized trajectories and associated constraints extracted by *Gaussian mixture regression* (GMR).

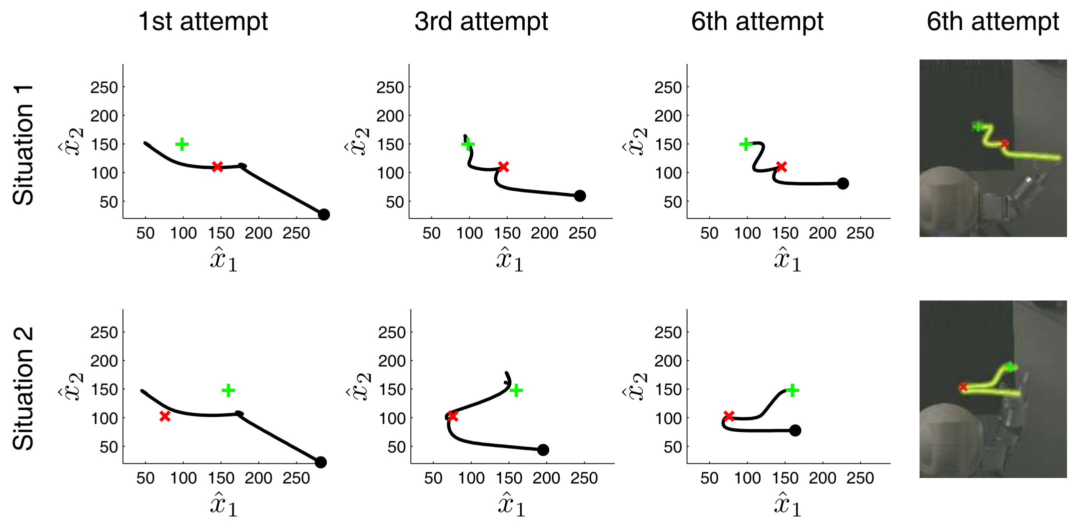

Figure 8. Reproduction of the task after the first, third, and sixth demonstrations for two new initial positions of the objects. A solid line represents the hand path with a point marking the starting position. A cross and a plus sign represent the cylinder and the cube, respectively.

a bell-shaped path, which is consistent with the demonstrations provided. For the second object, we see that the trajectory $\hat{x}^{(2)}$ is highly constrained at the end of the task, when placing the cylinder on top of the cube.

The reproduction of the task in two new situations are presented in Figure 8. After the first demonstration, the system is not able to generalize the skill (a similar motion is used for the two reproduction conditions). After the third demonstration, the constraints are already roughly extracted. We see that the robot correctly grasps the cylinder, but still has some difficulty in placing it on the cube. By watching the robot reproduce the skill in the two different situations, the teacher detects at the third reproduction attempt that the second part of the task is not fully understood. The teacher then provides appropriate scaffolds by kinesthetically demonstrating the task with an increased variability in the initial positions of the cube. Thus, the constraint of placing the object on top of the cube becomes more salient, and at the sixth attempt, the different constraints are finally fulfilled for the different reproduction conditions; namely, grabbing the cylinder, moving it with an appropriate bell-shaped movement, and placing it on top of the cube.

Toward evaluation of the proposed imitation scenario

The two teaching scenarios presented in the previous section have been inspired by one of the key experiments developed in the *Cogniron* project. This European project aims at the development of a cognitive robot companion for use in a domestic environment. One study in this project is the human perspective of how robots could be useful in domestic environments; in particular the roles, tasks, and social behavior that will be necessary for robots to integrate into normal domestic situations. The last phase of the project consists in evaluating the performance of the different *human-robot interaction* (HRI) setups developed by the different partners involved in the project through the use of key experiments.[1] The evaluation is performed at different levels: (1) Through technical evaluation of the algorithms; and (2) through user evaluation studies. Following this idea, user experience studies have already been proposed within the consortium (Woods, Walters, Koay, & Dautenhahn, 2006; Alissandrakis, Nehaniv, Dautenhahn, & Saunders, 2006). The principal aim of these studies is to bridge the gap between insights from the user's perspective and the actual development of algorithms that allow the robot to learn in the context of HRI.

To evaluate an imitation attempt, a metric of imitation performance can be technically defined, but is not used in the same way as a metric comparing, for example, a visual tracking system. Indeed, using a probabilistic framework, it is straightforward to evaluate quantitatively the efficiency of the system in terms of encoding, recognition and generalization performance, for a particular dataset. Indeed, the probabilistic GMM representation can be used to recognize new gestures as well as to evaluate a reproduction attempt. A metric of imitation is then defined by estimating the likelihood that an observed trajectory could have been generated by the model, that is, by computing the log-likelihood of a model

[1]The duration of the project is four years, with the remaining year devoted essentially to the technical evaluation of the developed setups and to the planning of user studies evaluating the acceptability of the developed setups.

when testing a set of N data points ξ_j

$$\mathcal{L} = \sum_{j=1}^{N} \log\left(p(\xi_j)\right),$$

where $p(\xi_j)$ is the probability that the model generates ξ_j.

This metric can be used easily to determine the imitation and generalization performance of the model with respect to a specific dataset ξ. However, it is trickier to measure the *pedagogical* quality of the data provided. This issue led us to the design of a teaching scenario mimicking the human process of teaching. As the different components of the scenario are tightly intertwined, it is not straightforward to measure separately the advantage of each component composing the complete system. Indeed, teaching involves multiple perceptual modalities that cannot be clearly separated. Thus, we think that the evaluation should not focus on trying to isolate each components of the system but rather check whether the combination of these components is satisfying for the end-user and whether it is computationally advantageous in terms of learning.

To design efficient teaching systems and appropriate benchmarks, we highlight the importance of taking psychological and social factors into consideration. Even if the learning efficiency can be measured quantitatively (i.e., measuring how well the system reproduces the skill and generalizes it with respect to a specific dataset), this evaluation is not sufficient, because the quality of the dataset also depends on the teaching abilities of the user. It is thus important to consider the skill transfer process at the user level, by checking whether the complete HRI setup makes efficient use of the teaching abilities of the user.

We first present several insights that guided us to the design of the current setup, from various fields of research covering psychology, pedagogy, developmental sciences, sociology and sports science. For further evaluation of the setup, we then suggest to group these social and psychological factors into three benchmarks to evaluate the skill transfer from a user's perspective.

Insights from psychology

Considering the robot as a peer learner and watching the evolution of its understanding of the skill are important psychological factors for a successful interaction. By drawing parallels with a caregiver-child interaction (and the associated human social aptitude at transmitting culture), we see that by watching the evolution and the outcomes of teaching, the teacher can feel psychologically more implicated in the teaching teamwork. In psychology, benchmarks such as autonomy, moral accountability and reciprocity have been proposed to evaluate HRI setups (Kahn, Ishiguro, Friedman, & Kanda, 2006). For teaching applications, the robot's capacity to generalize over different contexts depends on the number of demonstrations provided to the robot, but more importantly on the *pedagogical* quality of these demonstrations (gradual variability of the situations and exaggerations of the key features to reproduce). To succeed, it is therefore crucial that the teacher feels implicated in the teamwork. A possible benchmark to measure such success would be to test whether the teacher understands his or her role in the interaction and whether he or she naturally becomes a good teacher for the robot by making use of humanlike teaching skills.

Indeed, an essential feature of human beings is our natural ability and desire to transmit skills to the next generation. This transfer problem is complex and involves a combination of social mechanisms such as speech, gestures, imitation, observational learning, instructional scaffolding, as well as physical interaction such as molding or kinesthetic teaching (Salter, Dautenhahn, & Boekhorst, 2006). The humanoid shape of the robot and the humanlike properties of the teaching system aim at helping the teacher consider the robot as a peer. As noted by MacDorman and Cowley (2006), a humanoid robot is the best equipped platform for reciprocal relationships with a human being. When considering a robot conforming to humanlike appearance and learning behaviors, a psychological factor related to altruism and reciprocity naturally appears. This may be particulary important when considering teaching interactions. Creating a humanlike teaching scenario could reinforce the user's feeling that his or her role is to pass on knowledge to the robot, just as another human might help him or her out in a similar situation. In other words, the user would act as he or she would like to be treated.

The active teaching process probably evokes some of the feelings normally attributed to a caregiver-child dyad, in which the robot would elicit a certain form of attention. It would thus be important to consider benchmarks that look at whether the teacher feels humanly involved in the interaction, and check whether he or she is pushed by a desire to see the robot student progress. Indeed, watching the robot progress through the user's support may be psychologically relevant in HRI setups. This could be checked by discerning when the user may actually feel enjoyment during the scenario. The aim would be to discover whether the user feels enjoyment during the whole training process or only at the final reproduction phase, and to find out whether this feeling is related to the robot's progress. Indeed, watching the increased understanding of the skill by progressively lowering the scaffolds may be an important psychological factor for the teacher. By defining benchmarks in imitation, it would thus be useful to determine whether the teaching interaction can be related to the natural moral disposition of caregivers to teach children, and whether the user feels self-esteem when teaching the less-knowledgeable robot.

As shown by Lee, Hope, and Witts (2006), behavioral synchronization is also a powerful communicative experience, which suggests the use of a relational closeness benchmark in HRI. Through their experiments, the authors show that in human dyadic interactions, close relationship partners are more engaged in joint attention to objects in the environments and to their respective body parts than strangers. Moreover, they are more engaged in *discovering* and *showing* activities, sharing experience and exploring the environment together. Indeed, by contrasting a humanoid robot learning system with a computer language interpreting commands executed by a programmer, we see that both processes involve a transfer of information to the machine by incrementally watching the outcome of the transfer to help the user refine his or her teaching strategy. However, in the computer language situation, the machine considered is more of a tool to process the user's strategy rather than a peer learner what the user might wish to instruct.

In robotics, Thomaz, Berlin, and Breazeal (2005) presented HRI experiments in which the user can guide the robot's understanding of the objects in its environment. The interaction creates a form of closeness where the user explores together with the robot its environment. Toward this view, it would be important to check whether a kinesthetic teaching approach also provides such a close relationship with the robot by physically guiding the robot's arms to let it collect kinesthetic experiences. This could be performed by comparing an unsocial process of data acquisition (e.g., acquisition through motion sensors without facing the robot) with a social process of data acquisition such as kinesthetic teaching.

Insights from pedagogy

Gergely and Csibra (2005) contrasted observational learning to pedagogy. The authors noted that although such animals as chimpanzees can use tools to achieve a goal, they tend to discard the tools after using them. Humans however have developed the ability to recursively use tools (use of a tool to create another tool), attaching functions to objects that do not directly involve outcomes. This could have led to the development of pedagogical skills as powerful social learning mechanisms to enable the transmission of not just observable behaviors but also unobservable knowledge. Unlike observational learning, pedagogy requires an active participation of the teacher which is achieved by a special type of communication that aims at manifesting the relevant knowledge of a skill. The teacher does not simply use his or her knowledge but engages in an activity that benefits the learner. To highlight which features of a skill are relevant, the teacher must first recognize the features by analyzing his or her knowledge relative to the knowledge of the learner. Indeed, one does not need to be aware of knowledge content to generate an appropriate behavior (e.g., riding a bike may appear easy for a person, even when describing and teaching the skill to another person is not obvious).

Insights from developmental sciences

Zukow-Goldring (2004) presented experiments in developmental psychology to understand the methods used by caregivers to assist infants as they gradually learn new skills by engaging in new activities. By analyzing how a child learns manipulation and assembling skills (using pop beads), the authors show that observation alone is not sufficient for the child to learn a new activity. The caregiver first focuses the attention of the child to the *affordances* of the objects (the possibility to assemble the two objects). Because of the lack of information concerning the paths to follow to correctly assemble the objects, the child first fails at reproducing the actions. The caregiver then shows what the child's body should do to assemble the objects with the required orientations and paths to follow, demonstrating the *effectivities* required to assemble the objects. The caregiver assists the child by partially demonstrating the action and lets the child resolve progressively the skill by herself. Thus, the caregiver's gestures gradually provide perceptual information that guide the infant to perform the skill. The task is first simplified and the child is then progressively put in various situations to experience and generalize the skill. *Embodying* and putting the child through the motions draws her attention to the coordination of *affordances* of the two objects and *effectivities* of the body required to connect the two objects. Thus, teaching is structured to let the child gather information on the characteristics of the objects and actions specific to each of them.

By providing bodily experience, the caregiver provides the infant the opportunity to see and feel the solutions to the correspondence problem (i.e., detecting the match between self and other). Similarly to the teaching process presented in these developmental psychology experiments, our proposed human-robot teaching scenario starts with the robot first observing the task performed by the user (through motion sensors). The robot learner then begins to reproduce the task through the teacher's assistance, gradually performing parts of the task independently (by selecting the motors to control manually, *embodying* the robot and putting it through the motion). In our experiment, the *affordances* and *effectivities* respectively refer to the correct way of assembling the objects and the correct movements that the robot's body must adopt to manipulate the objects appropriately.

Rohlfing, Fritsch, Wrede, and Jungmann (2006) also highlighted the importance of having multimodal cues to reduce the complexity of human-robot skill transfer. In their work, they consider multimodal information as an essential element to structure the demonstrated tasks. Through experiments, the authors show that humans transfer their knowledge in a social interaction by recognizing what current knowledge the learner lacks. They are thus sensitive to the cognitive abilities of their interaction partner. The authors then suggest taking insights from these studies to reduce the learning complexity of current RPD frameworks; thus, sharing human adaptability with the less knowledgeable becomes a central issue when designing social robots. Therefore, they hypothesize that a human teacher can also adapt naturally to a robot equipped with specific abilities. We adopt a similar strategy in our learning framework and show that the skill transfer process could benefit from the user's capacity to adapt his or her teaching strategies to the particular context.

Insights from sociology

Vygotsky (1978) introduced the *zone of proximal development* (ZPD) as a general term to define the gap between what the learner already knows, namely, the learner's *zone of current development* (ZCD), and what the learner can acquire through the teacher's assistance (Wood & Wood, 1996). An efficient teaching strategy consists of exploring and being familiar with the ZPD of the learner to evaluate what the learner is able to acquire given his or her current ability. This general paradigm can also be applied to human-robot interaction, where the role of the teacher is to ascertain the current ZCD of the robot learner (by testing the robot's current understanding of the skill), and where its ZPD lies (i.e., to ascertain what can be achieved by the robot with assistance). Searching for what the robot already knows can appear time-consuming, but it is also an important psychological factor for the teacher. It helps the teacher feel involved in collaborative human-robot teamwork.

When designing a human-robot teaching scenario, it is thus important to allow the teacher to acquire the robot's ZPD through interaction to provide individualized support to the robot. This is achieved by anticipating the problems that the robot might encounter, providing the appropriate scaffolds and gradually dismantling these scaffolds as the learner progresses (and eventually constructing further scaffolds for the next stage of learning). Similar to *molding* behaviors, moving the robot kinesthetically in its own environment provides a social way of feeling the robot's capacities and limitations when interacting with the environment.

Insights from sports science

To transfer a skill between two human partners, different ways of performing demonstrations can be used depending on the motor skill that must be transferred. Several methodologies have been investigated for skill acquisition in sports with the aim of providing advice to sport coaches on how to transfer a motor skill efficiently and how to measure success depending on the capacities of individual athletes (Horn & Williams, 2004).

Coaching is the learning support aimed at improving the performance of the learner to carry out the task by providing directions and feedback; it is highly interactive and requires that one continuously analyzes the learner performance. To do so, the coach needs to be receptive to the learner's current level of performance. Different modalities are traditionally used by sport coaches to help the athletes acquire the skill (Horn & Williams, 2004), where the visual observation of a skilled model completing the entire task provides a good basis for movement production. This also applies to other types of skills, see, for example, training by expert surgeons (Custers, Regehr, McCulloch, Peniston, & Reznick, 1999). Similarly, in our system, the principal aim of observing the performance of an expert through motion sensors is to provide a complete and temporally continuous demonstration of the skill. By using motion sensors, the human expert can freely perform the task while the robot *observes* the full-body motion. However, as noted by Hodges and Franks (2002), optimal movement templates do not always generalize well across individuals. Thus, individually-based templates may be more appropriate in refining and achieving consistency in a skill. This supports the further use of kinesthetic learning to guide the robot's arms physically to let the robot experience the skill.

This body of work also states that multiple exposures provide the opportunity to discern the structure of the modeled task. It allows the teacher to organize and verify what the learner knows and to focus attention on problematic aspects in subsequent exposures. This is achieved until a *saturation point* is reached, determined by the coach, at which additional demonstrations of the task do not yield learning benefits. Similarly, after each reproduction by the robot, the user decides whether the robot has correctly acquired the skill or whether further refinement is required. Several factors affect the amount of additional benefit that can be derived from multiple observations of a model. Among them is the isomorphism of the various demonstrations. Indeed, repetition of identical demonstrations may be of limited utility, whereas extremely diverse demonstrations may generate conflicts or confusion. To encourage flexibility and adaptability, the coaches often manipulate task properties (e.g., by using different situations) to provide appropriate variability depending on the learner's capacities. Similarly, in our teaching scenario, the user is instructed to displace progressively the objects after each demonstration to provide variability in the exposures of the skill.

Defining benchmarks for further evaluation

The insights presented above come from various fields of research but highlight the importance of having a multimodal and incremental learning system to help the robot experience various situations and refine its skill gradually. To plan the future evaluations of our setup, we have organized the different psychological and social factors used to design our teaching scenario into three different benchmarks.

The first benchmark considers the multimodal cues used to provide the demonstrations. Its aim is to test the practical and psychological advantages of using different teaching processes and associated modalities (observational learning through the help of motion sensors; molding and scaffolding techniques through the use of kinesthetic teaching). We plan to test this by teaching a task using both motion sensors and kinesthetic learning, and using either one or the other method separately. We then suggest measuring the duration of the teaching process to attain a certain degree of generalization.

The second benchmark considers the relevance of the incremental process by contrasting it to a batch learning process. The generalization capabilities of the model can be used to evaluate how the incremental process improves the user's teaching skills. Then, to evaluate how the reproductions helps the user assess the robot's current understanding of the skill, we plan to check if the user stops the experiment after a correct number of demonstrations. Finally, to measure to what extent the user refines his or her demonstrations according to the reproduction attempts, we plan to use a *Wizard-of-Oz* procedure. Before the experiment, we could record a set of reproductions that do not fulfill some constraints of the task. Then, by replacing the real reproduction attempts with these trajectories, we could check whether the user refines his or her demonstrations accordingly.

The third benchmark proposes to check whether the user's awareness of the robot's cognitive capabilities alters the transfer of skill. Indeed, if the user knows how the robot combines the different demonstrations to learn the skill, he or she may change teaching styles accordingly. Similarly, the user may teach differently if he or she knows the physical capabilities of the robot in terms of manipulation, speed, degrees of freedom, or range of motion. Indeed, in the previous experiments performed with our robot, we observed that users were showing better teaching and pedagogical skills if they knew in advance what were the hardware/body and the software/cognitive capabilities of the robot. We plan to set up a more systematic evaluation by first checking whether the gestures used to teach the robot are different from the gestures used to simply apply the skill. Then, we plan to assess the minimum level of instruction required by a naive user to become a "good teacher."

Conclusion

Throughout this work, we highlighted the importance of designing and evaluating a RPD framework not only by considering the learning system itself but the interaction scenario. We proposed an incremental and interactive RPD framework that we tested in two experiments to teach manipulation skills to a humanoid robot. We emphasized the active role of the user in the design and evaluation of the system, and showed the advantages of putting the user in the loop of the robot's learning. By considering several social and psychological factors induced by teaching mechanisms, we finally suggested possible benchmarks for the further evaluation of the current setup.

References

Alissandrakis, A., Nehaniv, C., Dautenhahn, K., & Saunders, J. (2005). An approach for programming robots by demonstration: Generalization across different initial configurations of manipulated objects. In *Proceedings of the IEEE international symposium on computational intelligence in robotics and automation* (pp. 61–66).

Alissandrakis, A., Nehaniv, C., Dautenhahn, K., & Saunders, J. (2006, March). Evaluation of robot imitation attempts: Comparison of the system's and the human's perspectives. In *Proceedings of the ACM SIGCHI/SIGART conference on human-robot interaction (HRI)* (pp. 134–141). Salt Lake City, Utah, USA.

Atkeson, C., Moore, A., & Schaal, S. (1997). Locally weighted learning for control. *Artificial Intelligence Review*, *11*(1-5), 75–113.

Billard, A., & Siegwart, R. (2004). Special issue on robot learning from demonstration. *Robotics and Autonomous Systems*, *47*(2-3), 65–67.

Breazeal, C., Brooks, A., Gray, J., Hoffman, G., Kidd, C., Lee, H., et al. (2004). Tutelage and collaboration for humanoid robots. *Humanoid Robots*, *1*(2), 315–348.

Calinon, S., & Billard, A. (2007, March). Incremental learning of gestures by imitation in a humanoid robot. In *Proceedings of the ACM/IEEE international conference on human-robot interaction (HRI).* Arlington, Virginia, USA.

Calinon, S., Guenter, F., & Billard, A. (2007). On learning, representing and generalizing a task in a humanoid robot. *IEEE Transactions on Systems, Man and Cybernetics, Part B. Special issue on robot learning by observation, demonstration and imitation*, *37*(2).

Custers, E., Regehr, G., McCulloch, W., Peniston, C., & Reznick, R. (1999). The effects of modeling on learning a simple surgical procedure: See one, do one or see many, do one? *Advances in Health Sciences Education*, *4*(2), 123–143.

Dautenhahn, K. (1998). The art of designing socially intelligent agents: Science, fiction, and the human in the loop. *Applied Artificial Intelligence*, *12*(7-8), 573–617.

Gergely, G., & Csibra, G. (2005). The social construction of the cultural mind: Imitative learning as a mechanism of human pedagogy. *Interaction Studies*, *6*, 463–481.

Hodges, N., & Franks, I. (2002). Modelling coaching practice: the role of instruction and demonstration. *Sports Sciences*, *20*(10), 793–811.

Horn, R., & Williams, A. (2004). Observational learning: Is it time we took another look? In A. Williams & N. Hodges (Eds.), *Skill acquisition in sport: Research, theory and practice* (pp. 175–206). Routledge.

Ikeuchi, K., & Suchiro, T. (1992, May). Towards an assembly plan from observation, part i: Assembly task recognition using face-contact relations (polyhedral objects). In *Proceedings of the IEEE international conference on robotics and automation* (Vol. 3, pp. 2171–2177).

Ito, M., Noda, K., Hoshino, Y., & Tani, J. (2006). Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model. *Neural Networks*, *19*(3), 323–337.

Kahn, P., Ishiguro, H., Friedman, B., & Kanda, T. (2006). What is a human? - toward psychological benchmarks in the field of human-robot interaction. In *Proceedings of the IEEE international symposium on robot and human interactive communication (RO-MAN).*

Kuniyoshi, Y., Inaba, M., & Inoue, H. (1994). Learning by watching: Extracting reusable task knowledge from visual observation of human performance. *IEEE Transactions on Robotics and Automation*, *10*(6), 799–822.

Lee, B., Hope, G., & Witts, N. (2006). Could next generation androids get emotionally close? relational closeness from human dyadic interactions. In *Proceedings of the IEEE international symposium on robot and human interactive communication (RO-MAN).*

MacDorman, K., & Cowley, S. (2006). Long-term relationships as a benchmark for robot personhood. In *Proceedings of the IEEE international symposium on robot and human interactive communication (RO-MAN).*

Nicolescu, M., & Mataric, M. (2003). Natural methods for robot task learning: Instructive demonstrations, generalization and practice. In *Proceedings of the international joint conference on autonomous agents and multiagent systems (AAMAS)* (pp. 241–248).

Pardowitz, M., Zoellner, R., Knoop, S., & Dillmann, R. (2006). Incremental learning of tasks from user demonstrations, past experiences and vocal comments. *IEEE Transactions on Systems, Man and Cybernetics, Part B. Special issue on robot learning by observation, demonstration and imitation*, *36*(5). (In press)

Rohlfing, K., Fritsch, J., Wrede, B., & Jungmann, T. (2006). How can multimodal cues from child-directed interaction reduce learning complexity in robots? *Advanced Robotics*, *20*(10), 1183–1199.

Salter, T., Dautenhahn, K., & Boekhorst, R. te. (2006). Learning about natural human-robot interaction styles. *Robotics and Autonomous Systems*, *54*(2), 127–134.

Saunders, J., Nehaniv, C. L., & Dautenhahn, K. (2006, March). Teaching robots by moulding behavior and scaffolding the environment. In *Proceedings of the ACM SIGCHI/SIGART conference on human-robot interaction (HRI)* (pp. 118–125).

Schaal, S., & Atkeson, C. (1998). Constructive incremental learning from only local information. *Neural Computation*, *10*(8), 2047–2084.

Steil, J. J., Röthling, F., Haschke, R., & Ritter, H. (2004). Situated robot learning for multi-modal instruction and imitation of grasping. In *Robotics and autonomous systems* (Vol. 47, pp. 129–141).

Thomaz, A., Berlin, M., & Breazeal, C. (2005). Robot science meets social science: An embodied computational model of social referencing. In *Workshop toward social mechanisms of android science (CogSci).*

Vijayakumar, S., D'souza, A., & Schaal, S. (2005). Incremental online learning in high dimensions. *Neural Computation*, *17*(12), 2602–2634.

Vygotsky, L. (1978). *Mind in society - the development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Cambridge, Harvard University Press.

Wood, D., & Wood, H. (1996). Vygotsky, tutoring and learning. *Oxford Review of Education*, *22*(1), 5–16.

Woods, S., Walters, M., Koay, K., & Dautenhahn, K. (2006, September). Methodological issues in HRI: A comparison of live and video-based methods in robot to human approach direction trials. In *Proceedings of the IEEE international symposium on robot and human interactive communication (RO-MAN)* (pp. 51–58). Hatfield, UK.

Yoshikawa, Y., Shinozawa, K., Ishiguro, H., Hagita, N., & Miyamoto, T. (2006, August). Responsive robot gaze to interaction partner. In *Proceedings of robotics: Science and systems.* Philadelphia, USA.

Zukow-Goldring, P. (2004). Caregivers and the education of the mirror system. In *Proceedings of the international conference on development and learning (ICDL)* (pp. 96–103). La Jolla, California, USA.