Active Teaching in Robot Programming by Demonstration

Sylvain Calinon and Aude Billard

Learning Algorithms and Systems Laboratory (LASA), Ecole Polytechnique Fédérale de Lausanne (EPFL)

CH-1015 Lausanne, Switzerland

{sylvain.calinon, aude.billard}@epfl.ch

Abstract— Robot Programming by Demonstration (RbD) covers methods by which a robot learns new skills through human guidance. In this work, we take the perspective that the role of the teacher is more important than just being a model of successful behaviour, and present a probabilistic framework for RbD which allows to extract incrementally the essential characteristics of a task described at a trajectory level. To demonstrate the feasibility of our approach, we present two experiments where manipulation skills are transferred to a humanoid robot by means of active teaching methods that put the human teacher in the loop of the robot's learning. The robot first observes the task performed by the user (through motion sensors) and the robot's skill is then refined progressively by *embodying* the robot and putting it through the motion (kinesthetic teaching).

Fig. 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 while wearing motion sensors (*left*) and then helps the robot refine its skill by kinesthetic teaching (*right*), that is, by grasping the robot's arms and moving them through the motion.

I. INTRODUCTION

In a *Robot Programming by Demonstration* (RbD) framework, a robot learns new skills through the help of a human instructor [1]. Traditional approaches in RbD tend to consider the human user as an expert model who performs a task while the robot *observes passively* the demonstration, see e.g. [2], [3]. 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, a recent body of work suggested to refer to the teacher-learner couple as a team that engages in joint problem solving [4], and to use active teaching methods to put the human teacher "in the loop" of the robot's learning [5].

In previous work, we developed a probabilistic framework for extracting the relevant components of a task by observing multiple demonstrations of it [6]. The system is based on *Gaussian Mixture Models* (GMMs) encoding a set of trajectories collected by the robot through its sensors. In this work, we extend the use of our RbD framework by adopting the perspective that the teacher is not only a model of expert behavior but becomes an active participant in the learning process.

First, we suggest to follow an incremental learning approach which allows the teacher to watch the robot's reproduction attempts after each demonstration, and thus helps him/her assess the robot's current understanding of the skill and prepare the following demonstration accordingly [7], [8]. This scaffolding process was previously proposed in robotics to let the robot gradually generalize the skill for an increasing range of contexts [4], [9].

We then suggest to use different modalities to produce the demonstrations, similarly to the teaching process where a human teacher would first demonstrate the complete skill to the learner, followed by practice trials performed by the learner under the supervision of the teacher (Fig. 1). The learner can then experience and adapt the skill for his/her particular body capacities, as explored in developmental psychology studies [10]. In our RbD framework, we follow a similar strategy by providing motion sensors and kinesthetic teaching facilities to help the user act pedagogically with the robot. Indeed, to become a *"good teacher"*, the user does not have to simply use his/her knowledge; he must engage in an activity that benefits the learner. We take the perspective that unlike observational learning, *pedagogy* requires a special type of communication to manifest the relevant knowledge of a skill. As discussed by [11], the teacher first needs to analyze his/her knowledge content to emphasize in his/her demonstrations the aspects that are relevant for the learner.

II. EXPERIMENTAL SETUP

A. System overview

We assume that the essential characteristics of a skill can be extracted by observing multiple demonstrations of it. The generalization of a skill to different contexts was previously explored in robotics [9], [12]–[17], but most of the approaches proposed represent the constraints at a *symbolic level*, which requires to segment the skill in a pre-processing phase and, thus, to use pre-determined controllers to let the robot reproduce the skill. Here, we consider

the most general case of representing the constraints at a *trajectory level*, which allows to combine several sets of constraints in a continuous form (e.g., by considering constraints on different modalities or on different objects), and where the constraints can gradually change during the motion. Thus, our framework allows to automatically find a controller for the robot that reproduces smooth generalized trajectories satisfying several constraints simultaneously. The robot gradually builds a model of the skill by observing several demonstrations performed in various contexts (e.g., by demonstrating a manipulation skill with different initial positions of objects). After each demonstration, the robot reproduces a generalized version of the task by combining probabilistically the different constraints extracted. The model is refined incrementally after each demonstration, and the user stops the interaction when the robot has correctly learned the skill.

B. Hardware

The experiments are conducted using a Fujitsu HOAP-3 humanoid robot with 28 degrees of freedom (DOFs), of which 16 DOFs of the upper torso are used in the experiments. The initial positions of the objects are recorded through a moulding process where the teacher grabs one of the robot's arm, moves it toward the object, puts the robot's palm around the object and presses its fingers against the object to let the robot *feel* that an object is currently in its hand. When the object touches the palm, a force sensor inside the robot's palm is used to register the object position, i.e., when the force sensor retrieves a value over a given threshold, the robot briefly grasps and releases the object while registering its position in 3D Cartesian space.

The user's movements are first 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 is used to extract joint angle trajectories 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. We then use the motor encoders of the robot to record information while the teacher moves the robot's arms. The teacher first selects the motors that he/she wants to control manually by slightly moving the corresponding motors before reproducing the skill. The selected motors are then set to a passive mode, which allows the user to move freely these corresponding DOFs while the robot executes the rest of the task. In this way, the teacher provides partial demonstrations while the robot acquires proprioceptive information when performing the task in its own environment. The motion of each joint is recorded at a rate of 1000 Hz, and each trajectory is 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.

C. Probabilistic model

We consider a set of trajectories $\{t, \xi\}$, where each datapoint consists of a temporal value t and a spatial vector ξ describing any kind of sensory information (e.g., a posture

defined by joint angles or a position in Cartesian space). The spatial vector ξ can be described either in the original data space or in a latent space of motion (e.g., by projecting linearly the data in a subspace of reduced dimensionality).

To extract constraints from this set of trajectories, we first model the joint probability $p(t, x)$ in a *Gaussian Mixture Model* (GMM), trained incrementally by a modified version of *Expectation-Maximization* (EM) [7]. A generalized version of the trajectories is then computed by estimating $\mathbb{E}[p(\xi|t)]$, with associated constraints determined by $cov\,(p(\xi|t))$. If multiple constraints are considered (e.g., considering actions $\xi^{(1)}$ and $\xi^{(2)}$ on two different objects), the resulting constraints are computed by first estimating $p(\xi|t) = p(\xi^{(1)}|t) \cdot p(\xi^{(2)}|t)$ and then estimating $\mathbb{E}[p(\xi|t)]$ to reproduce the skill.[1] For a complete description of the algorithms, the interested reader is referred to [6] for a detailed description of the extraction of constraints process and to [7] for the incremental learning process.

Regression is used to generalize over the set of demonstrations and to reproduce a smooth trajectory by estimating the conditional probability $p(\xi|t)$ at each time step. Several regression techniques based on *Locally Weighted Regression* (LWR) were previously proposed in robotics to generalize over a set of demonstrations [18], [19]. Our approach follows a similar strategy by using GMMs to represent the joint distribution of the dataset $\{t, x\}$, and using *Gaussian Mixture Regression* (GMR) to estimate $p(x|t)$, as proposed in [20]. By using GMM, it is then possible to deal with encoding, recognition and reproduction issues in a single framework. Apart from retrieving generalized trajectories from the demonstrations, the variations (and correlations) allowed around the generalized trajectories are also retrieved by the model, which are used by the robot to extend the learned skill to different contexts and to find an optimal controller satisfying several constraints simultaneously.

III. EXPERIMENTS

We present two experiments to show that the method is generic and can be used with datasets representing different modalities. In the first experiment, the skill is represented in *joint space*, where the GMM/GMR process is performed in a latent space of motion extracted by *Principal Component Analysis* (PCA). In the second experiment, the skill is represented in *task space*, where the GMM/GMR process acts on the position of the right hand relative to different objects in the scene.

A. Learning bi-manual gestures

This experiment shows how a bimanual skill can be taught incrementally to the robot in joint space using different modalities (observational learning and scaffolding). The task consists of grasping and moving a large foam die (Figs. 1 and 2). 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

---

[1]The notations $\mathbb{E}[\cdot]$ and $cov(\cdot)$ are used respectively to express expectation and covariance.
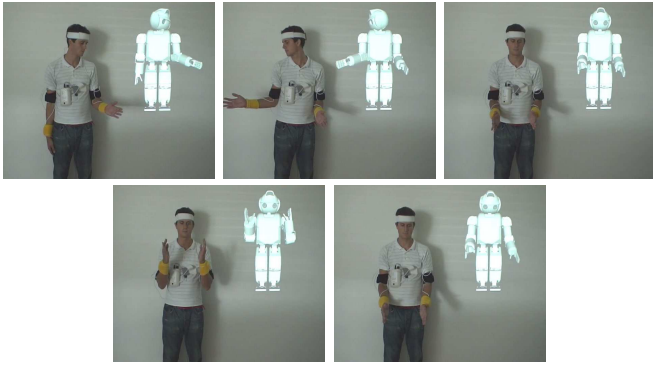
Fig. 2. Illustration of the use of motion sensors to acquire gestures from a human model. A simulation of the robot is projected behind the user to show the gesture observed by the robot. The gesture depicted here is similar to the one used in the experiment, except that in these snapshots the user does not face the robot and only mimics the grasping of the object.
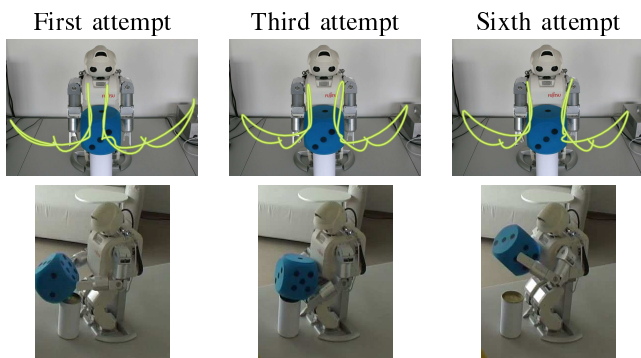

First attempt   Third attempt   Sixth attempt

Fig. 3. Reproduction of the task after the first, third and sixth demonstration. In the first attempt, the robot hits the die when approaching it, making it fall. In the third attempt, the robot's skill gets better but the grasp is still unstable. In the sixth attempt, the die is grasped correctly. The trajectories of the hands are plotted in the first row. The second row shows the corresponding snapshots of the reproduction attempts.

performed with the right arm. When both hands grasp the object, it is lifted and pulled back on its base, with the head turned toward the object (Fig. 2). The teacher performs the first demonstration of the complete task while wearing motion sensors. He can thus demonstrate the full gesture by controlling simultaneously the 16 joint angles, which are then projected into a subspace of lower dimensionality determined by PCA. After observation, the robot reproduces a first generalized version of the motion. This motion is then refined by physically moving the robot's limbs during the reproduction attempt, that is, by *embodying* the robot and putting it through the motion. Note that the gesture can only be refined partially by guiding the desired DOFs while the robot controls the remaining DOFs. Indeed, the teacher can only move a limited subset of DOFs by using his or her two arms. Thus, as discussed in [8], the combination of observational learning (through motion sensors) and kinesthetic teaching techniques allow to leverage the respective drawbacks of the two methods.

Results of the experiment are presented in Figs. 3 and 4, where the resulting hand paths are similar to the ones
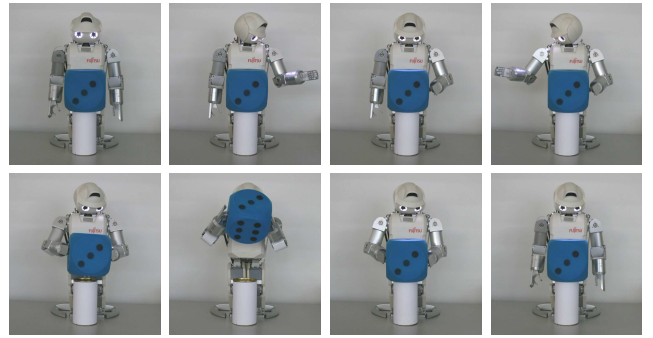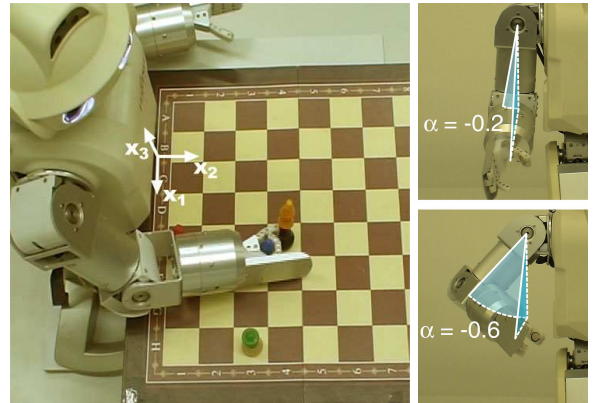

Fig. 4. Snapshots of the sixth reproduction attempt.

Fig. 5. *Left:* Experimental setup with frame of reference. *Right:* Definition of a posture by the position of the hand and an angular parameter α.

demonstrated by the user (Fig. 2).[2] The system finds automatically that five principal components and five Gaussian components are sufficient to encode the motion. After the first demonstration, the robot can only reproduce a smoothed version of the gesture 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 first 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 robot's skill 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 (Fig. 4). Therefore, the user decides to stop the teaching process.

B. Learning to move objects

In this experiment, the robot learns how to move different chess pieces by considering the hand-objects relationships at a trajectory level. The experiment draws insights from the teaching processes described by [10] in developmental psychology. It shows that teaching manipulation skills to the robot can be achieved through a scaffolding process where the user gradually highlights the *affordances* of different

---

[2]Note that the skill is represented as joint angle trajectories projected in a latent space of motion of lower dimensionality.
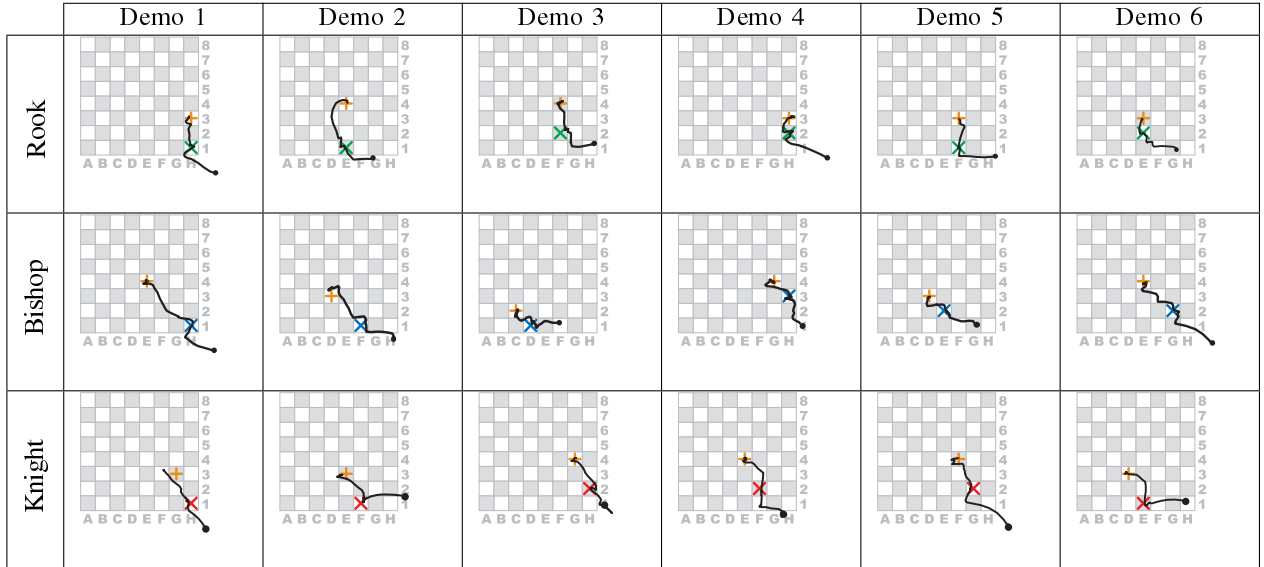
Fig. 6. Trajectories for the six consecutive kinesthetic demonstrations and for the three different chess pieces (only x_1 and x_2 are represented, corresponding to the plane of the chessboard). The cross and the plus sign represent respectively the chess piece to grasp and the opponent chess piece.
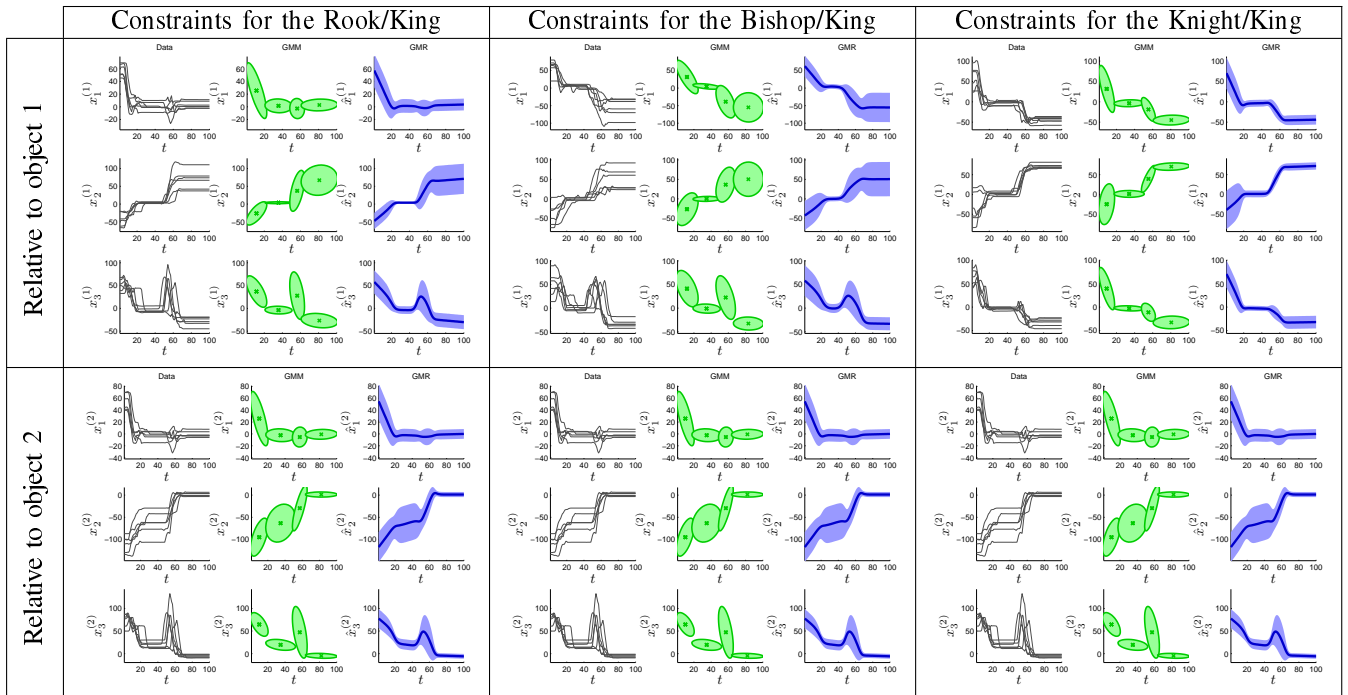

Fig. 7. Extraction of the hand-objects position constraints (position $x^{(1)}$ of the right hand relative to the Rook/Bishop/Knight and position $x^{(2)}$ of the right hand relative to the King). In each cell, the first column shows the trajectories, the second column shows the GMM and the last column shows the extraction of the generalized trajectories and associated constraints through GMR. The parts of the motion showing a thin envelope around the generalized trajectories are locally highly constrained, while the parts with a large envelope allow a loose reproduction of the trajectories.

objects in its environment and the *effectivities* of the body required to perform actions on these objects. A chess game is used in this experiment to explore how a robot can learn the *affordances* of different chess pieces (the particular relations and rules associated with the objects) and the associated *effectivities* (how the robot should use its body to manipulate and displace them without hitting other chess pieces). The chess paradigm was also used by [21] to explore the correspondence problem in imitation learning, that is, to know

how to reproduce a particular motion on the chessboard by considering different chess pieces (different *embodiments*).

The setup of the experiment is presented in Fig. 5. Six consecutive kinesthetic demonstrations are provided to show how to grasp a Rook, a Bishop or a Knight and bring it to the King of the adversary (Figs. 6). Three different models are created for the Rook, the Bishop and the Knight. The setup is simplified by using only two chess pieces at the same time (attention scaffolding is used to let the robot recognize only

these two chess pieces), and we do not take into account the generalization across different frames of reference. It means that in our setup, the Rook is moved only in a forward linear direction, the Bishop is moved only in a forward-left diagonal direction, and the Knight is moved only two squares forward followed by one square to the left, forming an inverse "L" shape.[3]

Through the teacher's support, the robot extracts and combines the constraints related to different objects. After each demonstration, the robot tries to reproduce the skill with the chess pieces placed in a new configuration that has not been observed during the demonstrations. The user can thus test the ability of the robot to generalize the skill over different situations. The constraints extracted after the sixth demonstration are presented in Fig. 7, showing the displacement constraints with respect to the first object (respectively the Rook, the Bishop and the Knight) and the second object (the King). For the Rook, we see that the trajectory is highly constrained for $x_1^{(1)}$ from time step 30 while the hand grasps the Rook, i.e., the hand-object relationship allows only low variability during this part of the skill. The Rook is then moved in a straight line, i.e., the direction in x_1 remains constant after grasping of the Rook. However, its final position can change in amplitude, i.e., the Rook is moved along a straight line but its final position on this straight line can vary, which is reflected by the constraints extracted for $x_2^{(1)}$ (larger envelope after time step 70). For the Bishop, we see that the generalized trajectories (relative to Object 1) follows a diagonal. The direction is highly constrained but the final position is not. For the knight, the generalized trajectories (relative to Object 1) are more constrained. Indeed, for a given initial position of the Knight, only one final position is allowed in the proposed setup. This is reflected by the constraints for $x_1^{(1)}$ and $x_2^{(1)}$ (and complementary for $x_1^{(2)}$ and $x_2^{(2)}$), where the path with respect to the initial position of the Knight is highly constrained (the followed path is quasi invariant across all demonstrations). For the three chess pieces, the constraints for the vertical axis $x_3^{(1)}$ share similarities, showing that the user grasps the chess piece from above, and displaces the chess piece following a bell-shaped trajectory in a vertical plane. We observe that for each chess piece the constraints for the first object are correlated with the constraints for the second object. Indeed, the positions of the two objects have important dependencies for the skill (it is important to reach the King of the opponent with the chess piece).

By considering the constraints relative to the two objects, the reproduction of the hand path is then computed for new initial positions of the objects. To do so, the absolute constraint for each object is computed by adding the new initial position of each object to the corresponding relative constraint (represented as a varying mean and associated co-variance matrix along the motion). The final hand path used for reproduction is then computed by multiplying at each

[3]Note that the basic rules remain the same and that the directions only depend on the frame of reference considered.
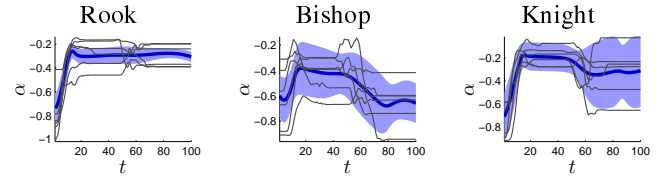

Fig. 8. Extraction of the constraints on the gesture (modeled by angle α) used to move the 3 different chess pieces.
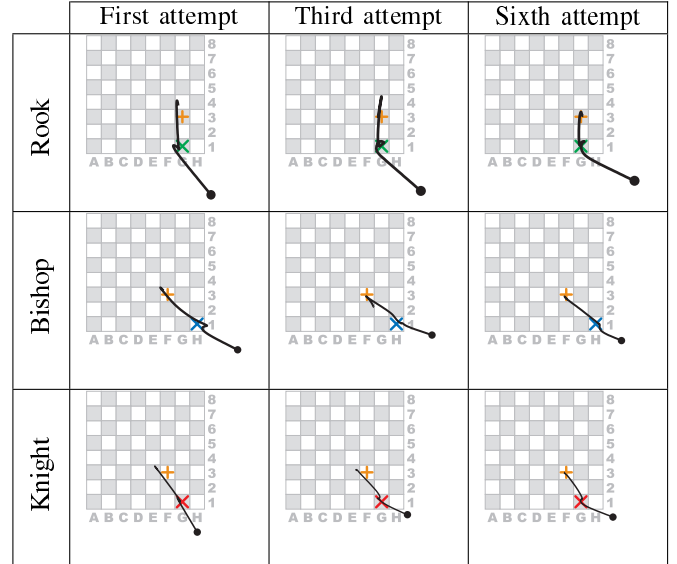

Fig. 9. Reproduction of the task after observation of the first, third and sixth demonstration. The cross and the plus sign represent respectively the chess piece to grasp and the opponent chess piece.

time step the two Gaussian distributions characterizing the absolute constraints for the two objects (i.e., by computing the probability of two events in conjunction). This hand path is converted to joint angles by using a geometrical inverse kinematics algorithm parameterized by the angle α between the elbow and a vertical plane (Fig. 5).

A generalization of the angle α observed during the demonstrations is used to reproduce the task (Fig. 8). This allows to reproduce naturally-looking gestures by providing an additional constraint to the redundant inverse kinematics problem. For the three chess pieces, we see that the gesture used to reach for the chess piece share similarities. Indeed, the α trajectories start with a negative value and progressively converge to zero. It means that the elbow is first elevated outward the body and is progressively lowered, approaching the body until the arm and forearm are almost in a vertical plane. This allows the user (and the robot) to approach the chess piece carefully without hitting the other chess piece. Indeed, when experiencing the skill together with the robot through a scaffolding process, the user quickly notices that when the robot is close to the chess piece, its elbow has to be lowered to grasp the chess piece correctly. This is mainly due to the missing DOF at the level of the wrist, which constrains the grasping posture of HOAP-3 to a value of angle α close to zero. When the chess piece is grasped, we see that two different movement strategies are

adopted depending on the path to follow. As the Rook is moved forward, the angular configuration does not need to be changed after grasping. As the Bishop is moved in a diagonal (forward-left), the user helps the robot adopt a correct posture to avoid hitting its own body when performing the move (i.e., learning *effectivities*). This is reflected by the decreasing negative value of angle α. A similar strategy is employed for the Knight, but the amplitude of change in angle α is lower due to the shorter path followed by the Knight. We also see that there is a slight tendency to first keep the arm in a vertical plane (α close to zero), and to finally use a posture with a slight elevation of the elbow (negative α value). This behaviour is probably due to the inadvertent decomposition of the "L" shape by the user when helping kinesthetically the robot displace the Knight.

Finally, Fig. 9 shows different reproduction attempts for initial configurations that have not been observed by the robot during the teaching process. We see that the ability of the robot increases, i.e., each demonstration helps the robot refine its model of the skill and ability to generalize across different situations.

IV. Conclusion

We presented a probabilistic RbD framework extracting incrementally the constraints of a task in a continuous form, which allows to generalize and reproduce the task in new contexts. We used two experiments to highlight the advantages of designing a human-robot interaction scenario mimicking the human process of teaching and taking advantages of the teaching abilities of the user. The first experiment showed the importance of having multimodal cues to structure the demonstrated tasks and to reduce the complexity of the skill transfer process. Indeed, we showed that while observational learning is a useful strategy to demonstrate a skill in its integrity, the use of kinesthetic teaching gives the opportunity to see and feel the solutions to the correspondence problem (detecting the match between self-generated and other motion). The second experiment showed that teaching manipulation skills to the robot can be achieved through a scaffolding process, where the user gradually highlights the *affordances* of different objects in its environment and the *effectivities* of the body required to perform grasping and moving actions on these objects.

Acknowledgment

References

[1] A. Billard and R. Siegwart, "Special issue on robot learning from demonstration," *Robotics and Autonomous Systems*, vol. 47, no. 2-3, pp. 65–67, 2004.

[2] K. Ikeuchi and T. Suchiro, "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 (ICRA)*, vol. 3, no. 12-14, May 1992, pp. 2171–2177.

[3] Y. Kuniyoshi, M. Inaba, and H. Inoue, "Learning by watching: Extracting reusable task knowledge from visual observation of human performance," *IEEE Transactions on Robotics and Automation*, vol. 10, no. 6, pp. 799–822, 1994.

[4] C. Breazeal, A. Brooks, J. Gray, G. Hoffman, C. Kidd, H. Lee, J. Lieberman, A. Lockerd, and D. Chilongo, "Tutelage and collaboration for humanoid robots," *Humanoid Robots*, vol. 1, no. 2, pp. 315–348, 2004.

[5] K. Dautenhahn, "The art of designing socially intelligent agents: Science, fiction, and the human in the loop," *Applied Artificial Intelligence*, vol. 12, no. 7-8, pp. 573–617, 1998.

[6] S. Calinon, F. Guenter, and A. Billard, "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*, vol. 37, no. 2, pp. 286–298, 2007.

[7] S. Calinon and A. Billard, "Incremental learning of gestures by imitation in a humanoid robot," in *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, March 2007, pp. 255–262.

[8] ——, "What is the teacher's role in robot programming by demonstration? - Toward benchmarks for improved learning," *Interaction Studies. Special Issue on Psychological Benchmarks in Human-Robot Interaction*, vol. 8, no. 3, 2007.

[9] J. Saunders, C. Nehaniv, and K. Dautenhahn, "Teaching robots by moulding behavior and scaffolding the environment," in *Proceedings of the ACM SIGCHI/SIGART conference on Human-Robot Interaction (HRI)*, March 2006, pp. 118–125.

[10] P. Zukow-Goldring, "Caregivers and the education of the mirror system," in *Proceedings of the International Conference on Development and Learning (ICDL)*, 2004, pp. 96–103.

[11] G. Gergely and G. Csibra, "The social construction of the cultural mind: Imitative learning as a mechanism of human pedagogy," *Interaction Studies*, vol. 6, pp. 463–481, 2005.

[12] S. Muench, J. Kreuziger, M. Kaiser, and R. Dillmann, "Robot programming by demonstration (RPD) - Using machine learning and user interaction methods for the development of easy and comfortable robot programming systems," in *Proceedings of the International Symposium on Industrial Robots (ISIR)*, 1994, pp. 685–693.

[13] M. Nicolescu and M. Mataric, "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)*, 2003, pp. 241–248.

[14] J. Steil, F. Röthling, R. Haschke, and H. Ritter, "Situated robot learning for multi-modal instruction and imitation of grasping," *Robotics and Autonomous Systems*, vol. 47, no. 2-3, pp. 129–141, 2004.

[15] A. Alissandrakis, C. Nehaniv, K. Dautenhahn, and J. Saunders, "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*, 2005, pp. 61–66.

[16] S. Ekvall and D. Kragic, "Learning task models from multiple human demonstrations," in *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, September 2006, pp. 358–363.

[17] M. Pardowitz, R. Zoellner, S. Knoop, and R. Dillmann, "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*, vol. 37, no. 2, pp. 322–332, 2007.

[18] S. Schaal and C. Atkeson, "Constructive incremental learning from only local information," *Neural Computation*, vol. 10, no. 8, pp. 2047–2084, 1998.

[19] S. Vijayakumar, A. D'souza, and S. Schaal, "Incremental online learning in high dimensions," *Neural Computation*, vol. 17, no. 12, pp. 2602–2634, 2005.

[20] Z. Ghahramani and M. Jordan, "Supervised learning from incomplete data via an EM approach," in *Advances in Neural Information Processing Systems*, J. D. Cowan, G. Tesauro, and J. Alspector, Eds., vol. 6. Morgan Kaufmann Publishers, Inc., 1994, pp. 120–127.

[21] A. Alissandrakis, C. Nehaniv, and K. Dautenhahn, "Imitation with ALICE: Learning to imitate corresponding actions across dissimilar embodiments," *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, vol. 32, no. 4, pp. 482–496, 2002.