Learning Interactively to Resolve Ambiguity in Reference Frame Selection

Franzese, Giovanni; Celemin, Carlos; Kober, Jens

Publication date
2020

Document Version
Final published version

Published in
Proceedings of Machine Learning Research

Citation (APA)
Franzese, G., Celemin, C., & Kober, J. (2020). Learning Interactively to Resolve Ambiguity in Reference Frame Selection. *Proceedings of Machine Learning Research*, *155*, 1298-1311.

Learning Interactively to Resolve Ambiguity in Reference Frame Selection

Giovanni Franzese **Carlos Celemin** **Jens Kober**
Department of Cognitive Robotics
Delft University of Technology, Netherlands
{g.franzese, c.e.celeminpaez, j.kober}@tudelft.nl

Abstract: In Learning from Demonstrations, ambiguities can lead to bad generalization of the learned policy. This paper proposes a framework called Learning Interactively to Resolve Ambiguity (LIRA), that recognizes ambiguous situations, in which more than one action have similar probabilities, avoids a random action selection, and uses the human feedback for solving them. The aim is to improve the user experience, the learning performance and safety. LIRA is tested in the selection of the right goal of Movement Primitives (MP) out of a candidate list if multiple contradictory generalizations of the demonstration(s) are possible. The framework is validated on different pick and place operations on a Emika-Franka Robot. A user study showed a significant reduction on the task load of the user, compared to a system that does not allow interactive resolution of ambiguities.

Keywords: Learning from Demonstrations, Active Learning, User-friendly Robot Learning, Human Robot Interaction

1 Introduction

In Learning from Demonstrations (LfD) the learning agent requires enough representative demonstrations for understanding the objectives of a task, in order to avoid any possible behavior misassociation. However, the requirement of representative examples could be demanding for a human teacher. Especially for an end-user it might not be clear how many demonstrations are sufficient. In this work, we focus on creating a system that learns from demonstrations, and that is ambiguity-aware. Therefore, it can leverage a teacher's corrective feedback to disassociate the multiple interpretations of the demonstrated behavior. With this awareness, the robot is able to prevent executing wrong (sometimes dangerous) actions, either with active queries before action execution, or enabling kinesthetic corrections by the user during the execution of an ambiguous decision.

Ambiguity is an attribute of any idea or statement where the intended meaning cannot be inferred as there are multiple interpretations. Teaching, as a form of transfer of knowledge, and not only of notions (strictly connected to the context), can be ambiguous. For instance, in a teacher-student scenario at school, if the teacher is explaining something that can have different meanings or interpretations, and the exam question is in the same circumstances of the explanation, a simple transcript of the teacher's exact words will lead to the best score. However, the meaning of learning is not repeating the lecture of the teacher by memory ; a good exam should check whether the student is able to generalize the concept to different situations. Similar situations may show up when robots learn from demonstrations, for instance, a robot arm is shown to go towards a cup placed on a coaster (see Fig. 1), but in a new situation, these two objects are in different positions: Where should the robot go? To the cup or to the coaster? Without any additional information this ambiguous situation could not be solved [1].

The ambiguity investigated in this paper prevents the policy from randomly selecting the dependence of the goal on a reference frame, whenever the demonstrations are not completely informative. The human teacher observes the current policy and by interacting with the environment provides feedback that Learning Interactively to Resolve Ambiguity (LIRA) employs for updating the correct goal. The approach reduces the (unknown) amount of required full demonstrations that a human teacher needs to provide, rather relying on less demanding interactive corrections, as discussed in

4th Conference on Robot Learning (CoRL 2020), Cambridge MA, USA.

Figure 1: Ambiguity of demonstration of grasping a cup on a coaster. The two possible hypothesis for generalizing in future segment are illustrated as two conflicting interpretation of the task.

[2, 3, 4]. Therefore, the system decreases the learning time, the probability of failure, and the workload of the user.

In the next section, related work on LfD, few-shot learning, ambiguity, and reference frame selection is reviewed. In Sec. 3, the LIRA algorithm is explained in detail. Sec. 4 shows the experimental validation of this innovative methodology, and finally, Sec. 5 concludes the work and describes open challenges.

2 Related Work

LfD is an intuitive alternative approach for encoding robot policies, instead of programming them by hand. Having the possibility to show the desired behavior to a robot, and to give corrections or feedback to the demonstrated trajectory has been shown to be a fast and versatile methodology, which is user-friendly for adapting policies to multiple scenarios [5]. One of the challenges in LfD is to extract as much information as possible from the demonstrations, namely, data efficiency, avoiding to burden the human with the responsibility of providing several demonstrations. One-shot imitation learning [6] proposes a Neural Network architecture where the new task is learned with only one new demonstration. Alternatively, the works in [7], [8] and [9] use Movement Primitives for learning new tasks in few shots. The authors of these related works underline the assumption of no ambiguity in the demonstrations remarking a learning failure otherwise.

Ambiguous demonstrations do not allow a unique definition of the decision, then, there is a consequent possibility of multiple valid actions in different scenarios. The origin of the ambiguity could be the uncertanity of the data. In particular, there are approaches that use probabilistic models that have estimates about the uncertainty of the policy [10]. In order to avoid executing unknown or potentially dangerous actions, robot policies could trigger users' demonstrations or feedback, whenever this confidence is below a certain threshold [3]. In navigation applications, novelty detection based on density estimation is used for requesting demonstrations in uncertain situations [11]. Similarly, Gaussian Processes have been used for active robot learners, which request demonstrations based on the policy uncertainty for learning movement primitives [12]. In [13] Confidence Based Autonomy is tested using Gaussian Mixture Models (GMM), Support Vector Machines, and Random Forests.

In this paper, the ambiguity investigated is not due to the inability of the human to provide consistent demonstrations in the current situation, but from the inability to learn a policy for a general situation, given the information contained in the initial demonstrated data, as formalized in [1]. In particular, in this paper we aim to reduce the dimensions of the space of possible generalizations using priors and a new user-friendly interactive methodology in the context of reference frame selection. Similar analysis on user-friendly interactions can be found in the literature about active reward learning [14], reward learning through preferences [15] and designing robots that asks good questions [16].

Reference Frame Selection Actions, (in this context, trajectories) change according to the observations (reference frame positions). In trajectory learning, the use of multiple reference frames allows to encode the movement not with respect to the world coordinates but relative to other points/objects, which increases the chance of matching the demonstrator's intention and allows for natural generalization to location changes of the points/objects. The learning algorithms are in charge of building the relations between observations and actions. When actions are associated to more than one of the frames observed by the robot, multiple demonstrations with a rich variety of initial conditions, that allow to break false associations with redundant frames, are required. However, this is not obvious for a non-expert user, who does not know the required data to have a representative set of demonstrations.

Algorithm 1 LIRA in Reference Frame Selection

Input: Demonstration(s) of the task and list of **valid frame candidates** of each movement primitive
for each movement primitive **do**
    Observe reference frames' positions and orientations
    Rank the frames according to *inter-demonstrations goal variance*.
    Filter out the frames with non-minimum variance and using **priors**
    Project the position of goal respect to each reference frame in the world frame
    Group the goals that are overlapping in the world frame
    for Each group in the list in descending order of dimension **do**
        Move the robot towards the group's goal
        Ask for human feedback
        if Positive Feedback **then**
            Eliminate in the list of **valid frame candidates** all the frames outside the group.
            Break the search and go to next movement primitive.
        else
            Eliminate the **groups** from the list of **valid candidates** according to the **Feedback**.

In [17] and [18], a Task Parametrized GMM is used for describing the set of demonstrations respect each of the reference frames. In a new situation, the resulting trajectory is based on the overlap of the GMM components that are moving according to their reference frames. This approach allows to switch the dependence from different reference frames in each movement segment and to regulate the robot stiffness proportionally to the variance of the model.

Similarly, in [8] a method for inferring the "correct" reference frame, with the computation of a score for each candidate frame, is proposed. The score is inversely proportional to the combination of the inter-demonstration variance derivative and final goal variance. After the segmentation of the trajectory, the method selects the reference frame with the highest score for each movement segment. Alternatively, in [9] the selection of the frame in each segment is performed by clustering the relative final position (of each one of the demonstrations) with respect to each frame in the world frame, and choosing the frame that owns the biggest cluster.

These related works underline the necessity of having more than one demonstration and that those are provided without ambiguities with respect to the way the relevance of each reference frame is computed. Our work is intended to propose a method that deals with ambiguous situations in reference frame selection and solves that with interactive human feedback.

3 LIRA: Learning Interactively to Resolve Ambiguity

In ambiguous situations, the learning system does not completely understand the intention of the task, therefore it has to randomly/arbitrarily choose one out of all the possible candidate interpretations, e.g., either the coaster or the cup of the example in Fig. 1. This makes the robot to behave in an unexpected way, having a negative impact in the user experience, and consequently in the human robot interaction, specifically in factors like engagement, trust or compliance [19, 20]. If users need to be careful about how to demonstrate, the teaching task becomes slower and increases their workload. This creates a user experience in which the robot does not seem to be "intelligent" or to have a certain level of situation "awareness". Without a system that solves ambiguities, a robot learner will request an entire new demonstration with the right decision, trying to understand how to match the *intention* of the teacher with the robot *deduction*.

With LIRA, the robot solves this problem of ambiguous situations by asking the user whether the current choice is correct. If the choice is incorrect, the robot enables the user to guide it, using different possible interaction modalities as shown in the last paragraph of this section.

This work assumes that policies are represented with sequences of movement primitives. During demonstrations, the only recorded data is the positions of the end-effector (EE), and the frames. The output of LIRA is the goal position for each of the movement primitives. LIRA is hence complementary to learning the shape of the movement via many movement primitive representations.

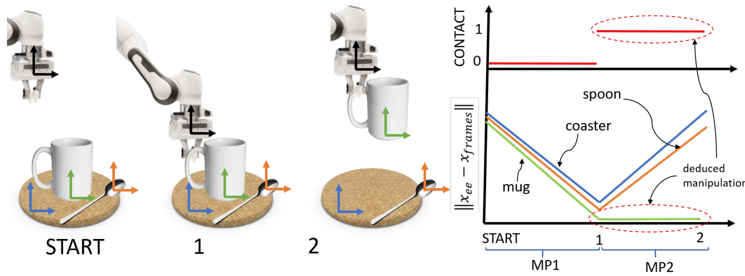

Figure 2: Demonstration of the grasping of a cup on a coaster. Because there is only one demonstration, the selection of the correct frame is ambiguous if no priors are used. In MP2, the grippers are closed (recorded contact) and the measure of the distance between the cup and the end-effector is kept constant. LIRA infers that the frame of the cup is being manipulated, giving priority to it in the resolution of the ambiguity of MP1. In the same way, LIRA removes the cup's frame from the list of candidates of MP2 because that goal would be already satisfied at the beginning of MP2.

LIRA initially takes the goal position from the demonstrations, and transforms it into the local coordinates of each of the n reference frames. The set of all possible single reference frames is

$$F = \{f_i | 1 \leq i \leq n\}$$

and each of its elements f_i has its associated goal ϕ_i contained in the set of goals Φ.

The proposed method is in charge of selecting the right frame f_i for each movement primitive, such that moving in its relative coordinates towards its goal ϕ_i is the correct generalization.Whenever there are ambiguities, the teacher's feedback queried by LIRA helps to reduce the amount of goal candidates, which are linked to the observed frames. The operations that Algorithm 1 is performing in order to take into account multiple demonstrations, priors, overlap of goals and human feedback are explained below.

Candidate Ranking: In case there are multiple demonstrations, the generation of the priority list for the reference frame is based on the measure of goal variance. After the segmentation, the final EE position of each movement primitive is computed with respect to each of the n frames, and the standard deviation of the distribution of inter-demonstration goals is computed for each frame in each movement segment. LIRA prioritizes the frames that have lower inter-demonstration variance, as in [8]. In case only one demonstration is provided, the co-variance cannot be calculated and a default value is set: only priors and human feedback can be used for finding the right candidate.

Reducing Ambiguity using Priors: As described in [1] and [21], the use of priors reduces the dimension of the candidate frames search space, making the ambiguity resolution faster. Various priors can be included to prioritize frames in the selection, e.g., distance to the frame origin (as a proxy for an object/goal being reached) [22] or temporal consistency (discouraging or encouraging frequent switches between frames). In some cases, the selection of the goal of each movement primitive is based on information of the immediate subsequent segments. For example, if a constant grasping contact and a constant distance with respect to a frame is detected in an entire segment, it means that the frame (i.e., the object it is attached to) is being manipulated. Hence in the preceding segment, in case of ambiguity, LIRA will give priority to that frame, encoding that an object needs to be grasped before being manipulated. An example of this situation is illustrated in Fig. 2.

Similarly, the frame being manipulated can be discarded as candidate reference frame for that segment. This frame is temporarily rigidly attached to the end-effector, and the goal for that frame would hence already be satisfied at the beginning of that segment.

Candidate Grouping: After applying the filtering with the priors and selecting the frames with the lowest inter-demonstration variance (in the limit of a threshold), LIRA groups the reference frames according to their current goal positions in the executing moving primitive. The j^{th} group $G_j \subset F$ is defined as a subset of reference frames whose associated goals are overlapping in the global coordinates. Therefore, regardless of the reference frame chosen among the group, the movement primitive would have the same global goal position. Thus, for a specific situation, there is an

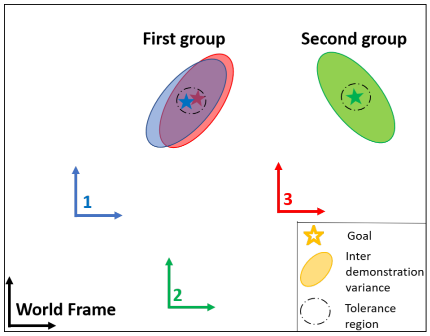

Figure 3: Candidate Grouping: all the goal distribution have the same inter-demonstration variance and they are not filtered away by any prior. Then, LIRA creates groups of goal, checking if they are overlapping within a tolerance.

ambiguity with an amount of c candidate goals, where c is the amount of groups encountered by LIRA. Goals are grouped by the distance between the candidate goals, and a threshold defining a tolerance region. In the example of Fig. 3, the goals of the three frames are organized in two groups, given the drawn tolerance region. Therefore, in this ambiguous situation, LIRA gives priority to the biggest group on the left (line 8 of Algorithm 1).

When the user is labeling one of the groups as correct, all candidates in the other groups are eliminated. All members of the selected group are saved to the memory and become valid candidates for future iterations. Due to the human feedback, there is no longer an ambiguity about the goal in this specific situation. Nonetheless, those multiple selected candidates might result in an ambiguity in future situations if the goals of the group do not overlap anymore.

Human Feedback: LIRA moves the robot towards the goal of largest group, i.e., maximizing the probability to find the right frame. Then it queries feedback from the teacher (line 10 in Algorithm 1) in these three possible ways:

- *Labelling correctness (right/wrong)*: The algorithm asks whether the current goal that was reached is correct or not. This query could be done through a graphic interface, or with a sound signal, whereas the teacher's response could be obtained with a keyboard, remote control, voice, etc. When the feedback is negative, in the **Eliminate** step (line 15 of Algorithm 1), LIRA only discards the frames belonging to the current selected group.
- *End-effector directional kinesthetic perturbation*: In order to avoid the dependency on a human-computer interface, another proposed solution is to make the robot compliant and to enable the human to correct the candidate choice simply by pushing the robot towards the direction of the correct goal (Fig. 4.2.b,1), i.e., LIRA assumes positive feedback (line 10 in Algorithm 1) when there is no kinesthetic perturbation, and negative otherwise. This interaction mode is richer in information than the previous one, as in case of a wrong guess of reference frame, the user interaction does not only make LIRA discard the wrongly selected goal, but also all the potential candidates which are in the opposite direction of the correction, therefore, the ambiguity resolution would be faster.
- *End-effector kinesthetic movement*: Similarly, in case of ambiguity, LIRA allows the user to move the EE close to the right goal (Fig. 5.2.b,1). If no correction happens, LIRA assumes positive feedback (line 10 in Algorithm 1). Otherwise, in case of EE movements it assumes negative feedback (line 15) and labels the closest group's goal to the final EE position after the interaction as correct, discarding all the others. This feedback modality is the most information rich and should be considered the best option when there are many perceived reference frames. However, due to the longer interaction with the user, this option would be expected to be more mentally and physically demanding compared to the other two options.

4 Experiments and Results

LIRA has been proposed to solve the ambiguity in selection of reference frame in an interactive way, reducing the total number of required full demonstrations. Experimental evaluations have been carried out for measuring the effectiveness of LIRA. Three validation scenarios of pick and place tasks, and a user study were run in order to illustrate the operation of the system, and to evaluate its

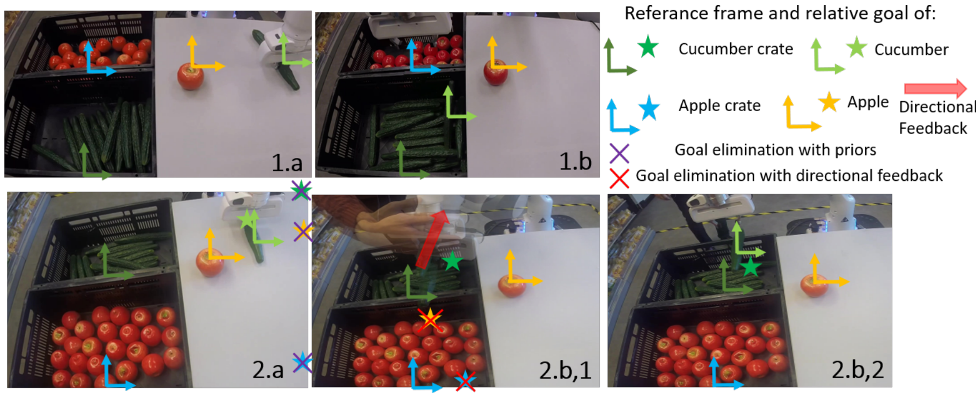

Figure 4: Fruit pick and place 1. The first row shows the policy execution in the same scenario of the demonstration (1.a-b), i.e. no ambiguity can arise. After a frame rearrangement, in the second row, LIRA uses the manipulation prior for the picking operation (2.a) but it asks the human feedback (2.b,1) for solving the ambiguity in the placing one (2.b,2).

performance compared to a system that is not ambiguity-aware. All the experiments were conducted with a 7 DoF Franka Emika Panda robot arm with a parallel gripper. The robot is controlled with Cartesian impedance control, which allows to have compliance for the interactions with the user. The robot is used in gravity compensation mode for recording the kinesthetic demonstrations. The demonstrated trajectories are segmented in movement primitives, and represented with linear attractors to their goals, with constant velocity. For the validation of LIRA, only one fully kinesthetic demonstration is provided as input to the policy.

4.1 Validation Tasks

Fruit sorting in crates: In the case depicted in Fig. 4, there are apples and cucumbers in the workspace that need to be picked up and placed in their respective crates (on the left). In the sequence shown in the top row (1.a-b) of Fig. 4, the scenario of the demonstration is reproduced, whereas in the bottom row (2.a-b,1-2), the position of the cucumber is different and the position of the crates is swapped (the scenario differs from the demonstration and it is potentially ambiguous as a consequence). Due to the priors, the robot is able to go and pick up the cucumber in the new scenario, however it requires the disambiguation for the goal of the second segment of the task (selection of the basket for placing the cucumber). Therefore, the robot moves towards the basket of apples, but stops for requesting the teacher's feedback before placing the object. Then, the simple physical perturbation towards the right goal is enough to solve the ambiguity. Without the ambiguity-awareness system, the user would have to demonstrate the whole trajectory in the second scenario (after observing a flawed object placement), instead of giving one correction during the autonomous execution.

Another validation example is depicted in Fig. 5, where the user is showing the restocking operation of bananas in the crates of a supermarket. After one demonstration (1.a-b), when the environment is modified (2.a), LIRA faces an ambiguity in the selection of the goal for the banana placement (2.b,1). As LIRA is implemented with the kinesthetic movement feedback in this example, the robot becomes compliant on the EE and the user moves it towards the right goal (2.b,1), solving the ambiguity (2.b,2).

Box stacking The third validation case is a box stacking task, as shown in Fig. 6. The sequence of pictures on top is showing the demonstrated scenario, where the stack of boxes is organized with the green one on bottom, the orange in the middle, and the black on top, each of which is perceived as a reference. For the second scenario (the bottom sequence of pictures), the use of the manipulation prior eliminates ambiguities during the grasping segment. However, in the placing operation, an ambiguity is found when the user disturbs the environment. With this disturbance, the goals associated to the frames on the orange and green boxes are not overlapping anymore. With the information obtained with the first demonstration, the robot learner faces an ambiguous situation, in which the goal for placing the black box could be with respect to either the orange or green box.

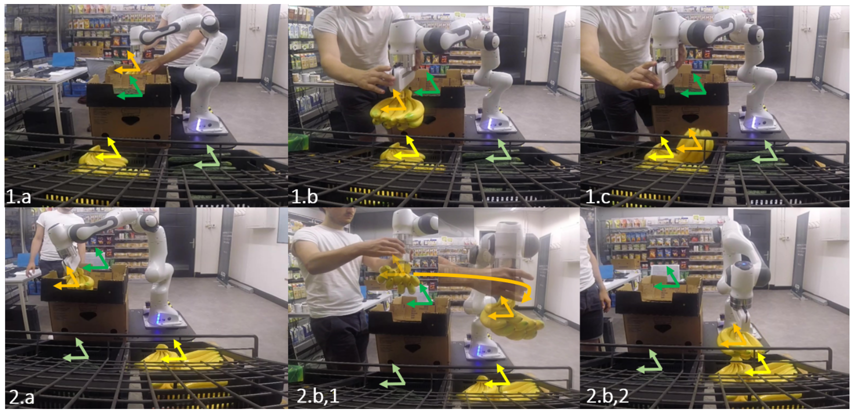

Figure 5: Fruit pick and place 2. The first row shows the user demonstrating the task with Kinesthetic Teaching (1.a-b-c). After the frame rearrangement, LIRA uses the manipulation prior for the picking operation (2.a) but it asks for feedback in the placing one (2.b,1) allowing the user to move the EE close to the desired goal (2.b,2).

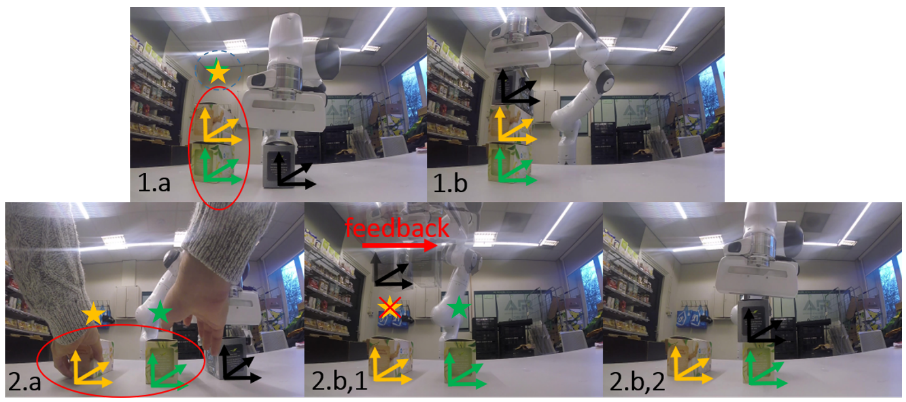

Figure 6: Frames overlap ambiguity. The first row shows the demonstration (1.a-b). Because the green and yellow boxes are stacked, then the relative position of their frames is never changing in new scenario and their goal will always be grouped in the placing segment of the black box (1.b). When the overlap is broken (2.a), an ambiguity is faced and LIRA asks for feedback (2.b,1): LIRA infers the user's preference of the dependence of the goal on the green frame (2.b,2).

The physical correction when the user pushes the end-effector makes the robot to understand that the black box's position should be with respect to the green one.

4.2 User Study

The performance of LIRA has been evaluated and compared with a system that learns from kinesthetic demonstrations without ambiguity-awareness: in case of frames with the same inter-demonstration variance, the selection is random and, if wrong, it requires a new complete kinesthetic demonstration. The intention of this study is to compare the amount of time reduction by the proposed interactive system for this specific task, along with the workload required from the user, both, with respect to the system without ambiguity-awareness. In this regard, a user study was conducted with twelve participants, who had to interact with both of the aforementioned systems, for teaching the same task. The type of feedback used in LIRA for the experiments was the directional perturbation because it was considered to be the best compromise between efficiency and user friendliness. The participants were students of an engineering faculty who had no prior experience of kinesthetic teaching, and their age ranged between 22 and 30 years. In order to have a measurement of the user workload during the teaching process, the NASA Task Load Index (NASA-TLX) questionnaire [23] was performed after the participants finished to interact with the system. This questionnaire has six

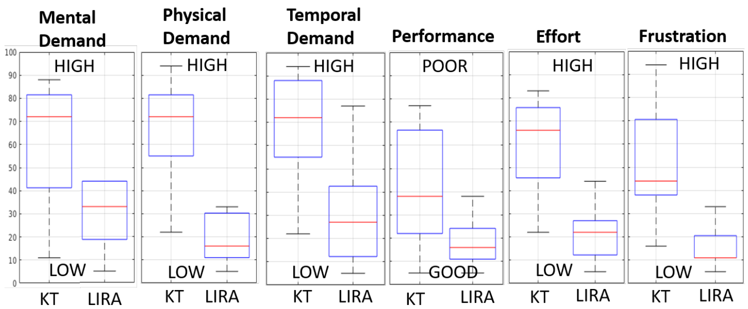

Figure 7: Statistical results of the workload NASA-TLX questionnaire that compares Kinesthetic teaching (KT) with LIRA on different aspects.

questions related to mental demand, physical demand, temporal demand, performance, effort, and frustration. The task was about stacking pairs of boxes with a predefined order. There were six boxes numbered from 1 to 6 on a table, and the objective was to place box number 1 on top of 2, 3 on top of 4, and 5 on top of 6 (see details in the supplementary material). At the end of the execution of each scenario, the objects were rearranged according to a predefined order. In the experiment with only kinesthetic teaching, the robot recorded the first demonstration, and then it was tested for the other cases. However, whenever the system failed, it would receive a full kinesthetic demonstration of the correct execution of the task for that scenario from the user. After 4 demonstrations, all the users managed to teach the task successfully, i.e., the ambiguities were solved.

For evaluating LIRA, a full kinesthetic demonstration was required only for the first scenario. For the other scenarios, LIRA requested the user to push the end-effector towards the right direction of the goal, in case of a detected ambiguity.

The results of the experiments showed the learning time was reduced by at least a half with LIRA, since it allows corrections during execution time, and does not request the user to record entire new trajectory demonstrations. In each scenario, LIRA never performed a wrong task reproduction, as the ambiguity-awareness prevented the robot from executing a mistaken decision, with the support of the human teacher.

Fig. 7 reports the results of the NASA-TLX questionnaire, which show how demanding the teaching task was with each of the methods. It is possible to observe that for all the questions, the participants reported better results with LIRA. For all the questions, the difference of the mean scores are considerable, and with LIRA the variance is lower in general. This was expected since the interactive system with ambiguity-awareness eliminated the flawed executions, along with the need of entire new kinesthetic demonstrations, which took about two minutes each time. Rather, LIRA requested shorter interactions of corrective feedback that required only approximately one second of physical contact with the robot.

5 Conclusion and Future Work

A system that is able to give awareness to the robot about ambiguities in the selection of reference frames for the goal of a movement primitive has been developed. This awareness is useful for implementing active robot learners, which can prevent the execution of mistaken actions produced by multiple potential interpretations of the demonstrated examples. The robots are able to request local physical corrections or signals of (dis)approval from the user, in order to feed its knowledge-base, and to eliminate the wrong associations. With this interactive approach, the workload of the teachers is reduced, since the user robot interaction is decreased in both cognitive and physical levels. This approach was validated with a user study, wherein the obtained results have shown an improvement in the user experience, with respect to a system that does not have ambiguity-awareness and an interactive disambiguation modality. This contribution intends to have more friendly robots which are able to share spaces and activities with their end-user counterparts. These robots need to be more and more adaptable to people who are not robot experts or do not have technical background, since they will become part of our daily life supporting basic activities. In the future, some extensions of the framework are going to be studied: ambiguity due to the inconsistency in the demonstrations, along with the ambiguity in the force-position modality selection, ambiguity in multiple sensor fusion and ambiguity in the null-space configuration of manipulators when there is a control only on the end-effector.

Acknowledgments

This research is founded by European Research Council Starting Grant TERI "Teaching Robots Interactively", project reference 804907.

The authors would like to thank the reviewers for useful feedback on presentation and contextualization of the current work. The experiments were conducted in the AIRLab of Robo-House, Delft. The authors want to share their gratitude with the members of RoboHouse for their hospitality and availability.

References

[1] S. Bensch and T. Hellström. On ambiguity in robot learning from demonstration. In *International Conference on Intelligent Autonomous Systems*, 2010.

[2] R. Pérez-Dattari, C. E. Celemin, G. Franzese, J. Ruiz-del Solar, and J. Kober. Interactive learning of temporal features for control: Shaping policies and state representations from human feedback. *IEEE Robotics & Automation Magazine*, 27(2):46–54, 2020.

[3] M. Kelly, C. Sidrane, K. Driggs-Campbell, and M. J. Kochenderfer. Hg-dagger: Interactive imitation learning with human experts. In *International Conference on Robotics and Automation (ICRA)*, pages 8077–8083, 2019.

[4] A. Bajcsy, D. P. Losey, M. K. O'Malley, and A. D. Dragan. Learning robot objectives from physical human interaction. *Proceedings of Machine Learning Research*, 78:217–226, 2017.

[5] B. D. Argall, S. Chernova, M. Veloso, and B. Browning. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469–483, 2009.

[6] Y. Duan, M. Andrychowicz, B. Stadie, O. Jonathan Ho, J. Schneider, I. Sutskever, P. Abbeel, and W. Zaremba. One-shot imitation learning. In *Advances in Neural Information Processing Systems*. 2017.

[7] S. Schaal, J. Peters, J. Nakanishi, and A. Ijspeert. Learning movement primitives. In *11th International Symposium on Robotics Research*, 2005.

[8] J. Kober, M. Gienger, and J. J. Steil. Learning movement primitives for force interaction tasks. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2015.

[9] S. Niekum, S. Osentoski, G. Konidaris, and A. G. Barto. Learning and generalization of complex tasks from unstructured demonstrations. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2012.

[10] M. Ewerton, G. Neumann, R. Lioutikov, H. B. Amor, J. Peters, and G. Maeda. Learning multiple collaborative tasks with a mixture of interaction primitives. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 1535–1542, 2015.

[11] D. Silver, J. A. Bagnell, and A. Stentz. Active learning from demonstration for robust autonomous navigation. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2012.

[12] G. Maeda, M. Ewerton, T. Osa, B. Busch, and J. Peters. Active incremental learning of robot movement primitives. In *1st Annual Conference on Robot Learning (CoRL)*, 2017.

[13] S. Chernova and M. Veloso. Interactive policy learning through confidence-based autonomy. *Journal of Artificial Intelligence Research*, 34:1–25, 2009.

[14] C. Daniel, O. Kroemer, M. Viering, J. Metz, and J. Peters. Active reward learning with a novel acquisition function. *Autonomous Robots*, 39(3):389–405, 2015.

[15] D. Sadigh, A. D. Dragan, S. Sastry, and S. A. Seshia. Active preference-based learning of reward functions. In *Robotics: Science and Systems*, 2017.

[16] M. Cakmak and A. L. Thomaz. Designing robot learners that ask good questions. In *7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 17–24, 2012.

[17] S. Calinon. A tutorial on task-parameterized movement learning and retrieval. *Intelligent Service Robotics*, 9(1):1–29, 2016.

[18] J. Hoyos, F. Prieto, G. Alenyà, and C. Torras. Incremental learning of skills in a task-parameterized Gaussian mixture model. *Journal of Intelligent & Robotic Systems*, 82(1):81–99, 2016.

[19] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich. Common metrics for human-robot interaction. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*, pages 33–40, 2006.

[20] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. De Visser, and R. Parasuraman. A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, 53(5): 517–527, 2011.

[21] B. Fonooni, T. Hellström, and L.-E. Janlert. Priming as a means to reduce ambiguity in learning from demonstration. *International Journal of Social Robotics*, 8(1):5–19, 2016.

[22] S. Alexandrova, M. Cakmak, K. Hsiao, and L. Takayama. Robot programming by demonstration with interactive action visualizations. In *Robotics: Science and Systems*, 2014.

[23] S. G. Hart and L. E. Staveland. Development of NASA-TLX (task load index): Results of empirical and theoretical research. In *Advances in Psychology*, volume 52, pages 139–183. Elsevier, 1988.

A Dependence on Two Frames: Increase the Search Space for Solving the Ambiguity

In the main paper on LIRA, different possibilities on how to use the human feedback and priors for selecting the right candidate reference frame (and solving the ambiguity) are introduced. However, sometimes, the intention of the demonstration is not dependent on only one reference frame but in the activation of two at the same time. For example, in Fig. 8 on the left side, the user shows the placement of a cup in the middle of a spoon and a box of cereals. In a different situation with a bigger table, on the right side of Fig. 8, referring to one of the two frames would result in mismatching the human will. This problem is solved creating $\binom{n}{2}$ other goals that are dependent on each pair of the n frames. The position of the goal is obtained preserving the proportion of the distance from each frame along their connecting line while, on the perpendicular direction, the distance is kept the same, see Fig. 9. Although this choice is arbitrary, we believe it matches the intention of the user in many use cases.

The downside is the increased number of goal candidates. This makes the algorithm slower in the selection of the right candidate but it also allows the teaching of a broader range of tasks.

A.1 Validation Task: Sorting Fruit in a Box with Different Dimensions

The validation example considers the case where the robot has to place objects in the middle of a box. In Fig. 10 on the left, the robot places a cucumber in the center of a small box (between peppers and apples), and there are reference frames over the left and right edges of the box. Therefore, there are three candidate goals, one associated to the left frame, another to the right frame, and another to the combination of both frames. After a single demonstration, LIRA cannot disambiguate the correct goal dependency using priors in the placing operation. On the right of Fig. 10, the new scenario has a bigger box, consequently, the candidate goals do not overlap, i.e., there is an ambiguity with three

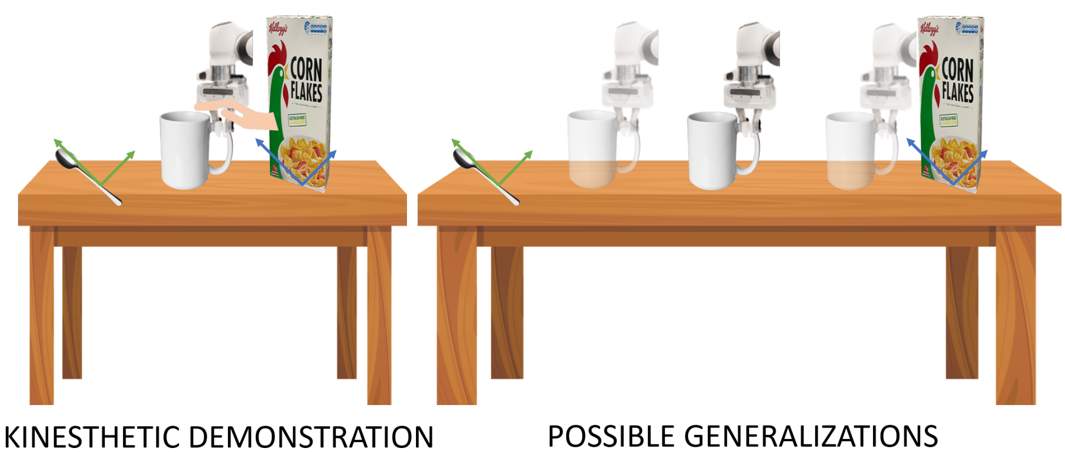

Figure 8: On the left, the demonstration of placing a cup in the middle of the two frames of a spoon and a box of cereals. On the right, in a new scenario with a bigger table, the dependency of the goal from each frame (transparent) and from both (solid).

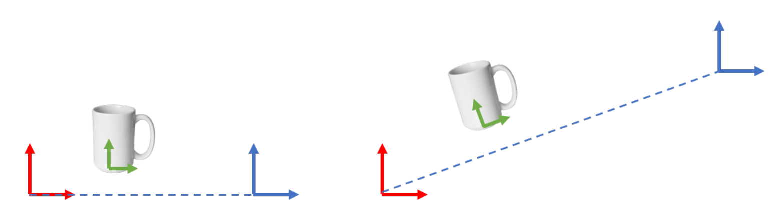

Figure 9: Graphical representation on how the general position and orientation of the goal is calculated as a function of two frames in two example situations.

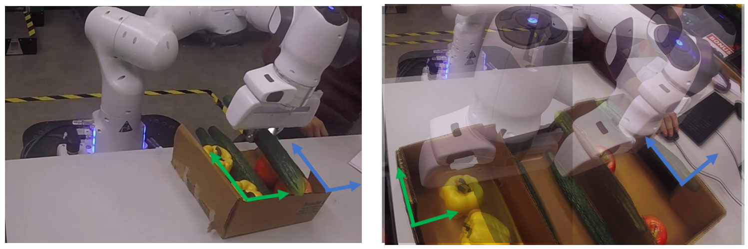

Figure 10: Two frame dependency. The two boxes are defined by two frames on each side of the box. The generalization with respect to two frames gives the possibility to reproduce the trajectories for objects of different scale.

different candidate solutions. In this situation, as shown in the video[1], the robot goes to the frame on the right, and requests feedback from a human, who replies: "wrong goal", through a user interface on a PC, according to the first type of feedback described in the main paper. Then the robot goes to the frame on the left, obtaining the same answer. With the second correction, the robot goes to the goal that activates both frames and the ambiguity is completely solved: for a new situation no more user interaction would be required.

B User Study: Extra Details

The video of the experiments shows the teaching of the task of placing of box 1 on 2, 3 on 4, and 5 on 6, when the only recorded data is the end-effector position with respect to each of the reference frames (attached to each box). As depicted in Fig. 11, in case of ambiguity, LIRA creates groups and uses the positive/negative feedback of the human for selecting the right candidate. The manipulation prior (introduced in the paper) automatically solves the ambiguity in the picking operations (MP 1,3,5) of the boxes 1,3,5. Similarly, in the placing operations (MP 2,4,6) the prior deletes the current manipulated frame, respectively 1,3,5. However, with only one full demonstration in Scenario 1 (Fig. 12) and with the use of priors, there is still not enough information for uniquely finding the dependence of the goal from a single reference frame, in new scenarios, for MP 2,4,6.

[1]https://youtu.be/tSQJP8Hpmbk

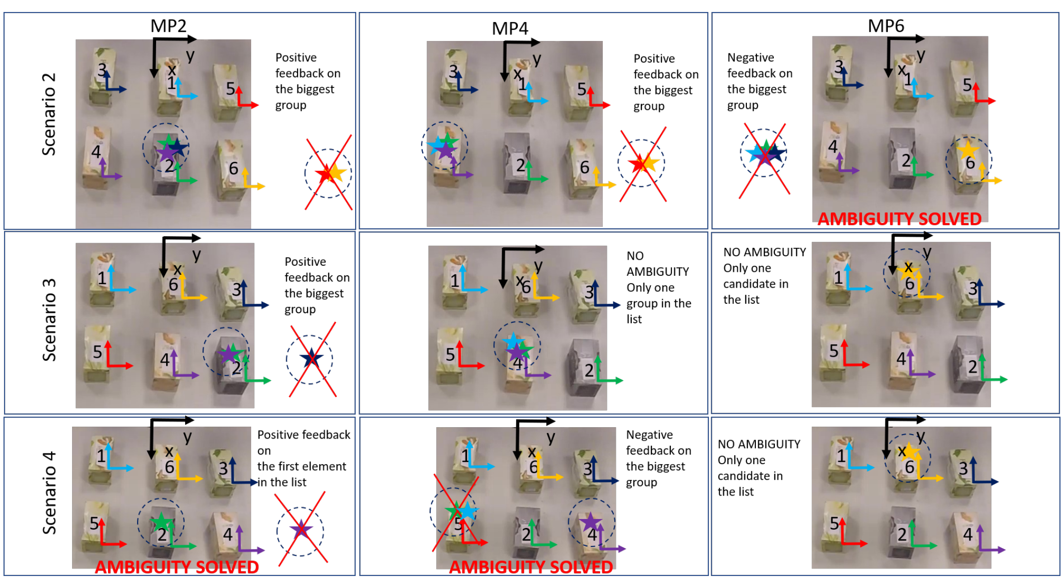

Figure 11: Details on the disambiguation procedure with LIRA and the human feedback in the Scenarios 2, 3, 4 after one complete Kinesthetic Demonstration in the Scenario 1 of Fig. 12

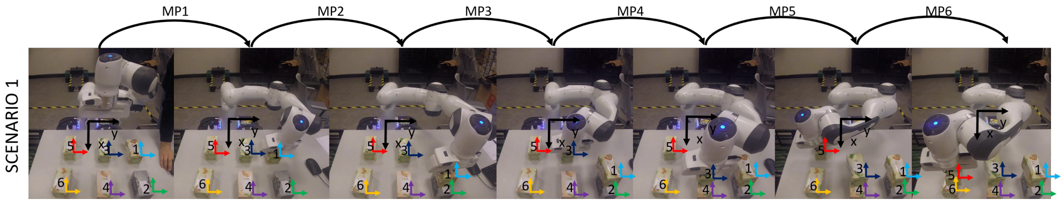

Figure 12: User study frames of the Kinesthetic Demonstration in the first scenario. The teaching task is stacking box 1 on 2, 3 on 4, and 5 on 6. The task is segmented in six movement primitives: three picking operations and three placing operations.

In Fig. 11, the position of the goals with respect to each reference frame, learned in the first demonstration, are projected in the new scenarios for each of the 'placing' movement primitives, i.e., where there are multiple valid candidates in the list and a potential ambiguity. The groups are formed according to the grouping operation explained in the paper. LIRA is always going to the biggest group first and if the user does not give any perturbation to the end-effector within a time limit[2], it is labeled as correct, i.e., all the elements of other groups are removed from the list of valid candidates. Alternatively, in case of a perturbation, the selected group is labeled as wrong and all its elements are removed from the list of valid candidates. The different rows of Fig. 11 show how the user, through positive and negative feedback, removes all the redundant candidates from the list until there is only one candidate left for each MP and no ambiguity can arise in future different scenarios.

C Mathematical Implementation of the Method

This section provides the mathematical formulation of the LIRA algorithm.

C.1 Inter-demonstration Variance

In case multiple full kinesthetic demonstrations occur, the computed inter-demonstration variance index is the trace of the covariance matrix of the normal distribution $\mathcal{N}(\phi_m^{RF}, \Sigma_m^{RF})$ that describes the goal position of each segment m with respect to each reference frame RF:

$$V_m^{RF} = \mathrm{tr}\left(\Sigma_m^{RF}\right).$$

This means that for every segment of the trajectory m, there are going to be as many co-variance matrices as there are reference frames. This value is used for ranking the relevance of the candidate frame. In case only one demonstration is provided, the resulting variance with respect to each reference frame would be 0. Hence the priority selection can only be done via the priors.

C.2 Manipulation Prior

An indicator that a object is being manipulated in a segment m, is that the goal with respect to the reference frame attached to the object is constant within segment m. The object is moving with the end-effector and hence not changing its relative position with respect to the end-effector. Hence, a manipulation frame MF is detected if

$$\mathrm{dist}\left(\phi_{m-1}^{MF}, \phi_m^{MF}\right) < d_{manip},$$

where ϕ_{m-1}^{MF} is the relative goal position at the start of segment m (which is equivalent to the goal of segment $m-1$), ϕ_m^{MF} is the relative goal position at the end of segment m, and d_{manip} is a tolerance parameter.

The manipulation prior uses manipulation frames to determine the priority within the immediately preceding segment $m-1$, if an object is grasped there (i.e., when a contact force is detected or the gripper is being closed). As it is likely that the object being manipulated in segment m corresponds to the one being grasped in segment $m-1$, the manipulation frame MF of segment m will have priority in the goal ranking of segment $m-1$.

[2]Please notice the pause in the execution of the task (in the video) when LIRA is unsure about the correctness of the goal and waits for the eventual feedback of the user with EE perturbation. LIRA labels that goal as correct if no feedback occurs.

C.3 Grouping Operation

The grouping operation is a clustering of goals according to the relative Euclidean distance. Taking the average ϕ_m of all demonstrations' goals with respect to each of the reference frames, and projecting them to *global robot* coordinates, two goals of the reference frames j and k would belong to the same group, in the segment m, if

$$\mathrm{dist}\left(\mathrm{proj}_{global}\left(\phi_m^j\right), \mathrm{proj}_{global}\left(\phi_m^k\right)\right) < D_{grouping},$$

where $D_{grouping}$ is the maximum diameter of a group.