Improving Autonomy by Learning from Failures for a Bolting Task

Elena Galbally, Adrian Piedra, Cynthia Brosque, Yuxiao Chen, and Oussama Khatib

***Abstract*— Even the most robust autonomous behaviors can fail. The goal is to both recover and learn from failures so that they can be prevented in the future. We propose haptic intervention for real-time failure recovery and data collection. The system presented in this paper allows for seamless transitions between autonomous robot behaviors and human intervention, while collecting the necessary sensory information to learn from the human's recovery strategy. We tested the system on two Panda arms with Allegro grippers and where able to successfully perform a bolting task while enhancing our original data set and autonomous behaviors.**

I. Introduction

No behavior is bullet proof neither in terms of task success nor safety. Preprogrammed trajectories will fail as soon as there is any change in the environment. A set of compliant skills executed by a probabilistic plan will be more robust, but also eventually fail in the real world.

We can't predict every failure scenario a priori. Therefore, trying to collect data that will span the entire space of possible scenarios the robot may encounter is pointless and will waste time and resources collecting excessively large data sets. Instead, we devised a system where you only need to collect a few demonstrations (we chose 20 because it proved to be enough for our segmentation module to work well - move this comment to methods) and then . It will collect data and flag failures. This improves both our understanding of the task and the skills that the robot is particularly weak at performing autonomously. and would require us to collect tons of demonstrations

In this paper, we define a *compliant primitive* as a parametrization of a 6-DOF controller defined with respect to a frame attached to the manipulated object. Directions of motion and compliance at the frame are specified such that they result in the desired object behavior. In the directions of compliance we do not control for desired trajectories. Instead we regulate a certain impedance behavior in order to generate a control force. Primitives are accompanied by sensory measurements that are used to establish and evaluate the object state.

Keeping the human in the loop is actually valuable: for safety critical tasks, for keeping jobs (tie this to construction in the paragraph), and for improving autonomy through continued learning. A cool thing is this system allows for one person to oversee a bunch of semi-autonomous robots since it only needs to intervene during failures.

In this paper, we present... describe how the system works

All the authors are part of the Stanford Robotics Lab. Contact: `elenagal@stanford.edu`. For the project website: FILL IN

figures/realHardwareCloseUp.png

Fig. 1: We tackled a steel bolting task using two 7-DOF arms equipped with 4 finger hands and force sensors and haptic device for failure recovery and contact data collection.

The proposed system was tested on ... describe task, hardware, and highlight results

Creating a fully automatic planner and learning the best way to grasp the objects for task success are beyond the scope of this paper.

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

The main contributions are: THINK (make them super clear)

- Use of haptics to recover from failures during contact tasks and record sensor information from the recovery strategy to enhance our originally data set in a targeted way
- Tested method on real hardware on a bimanual setup with force feedback

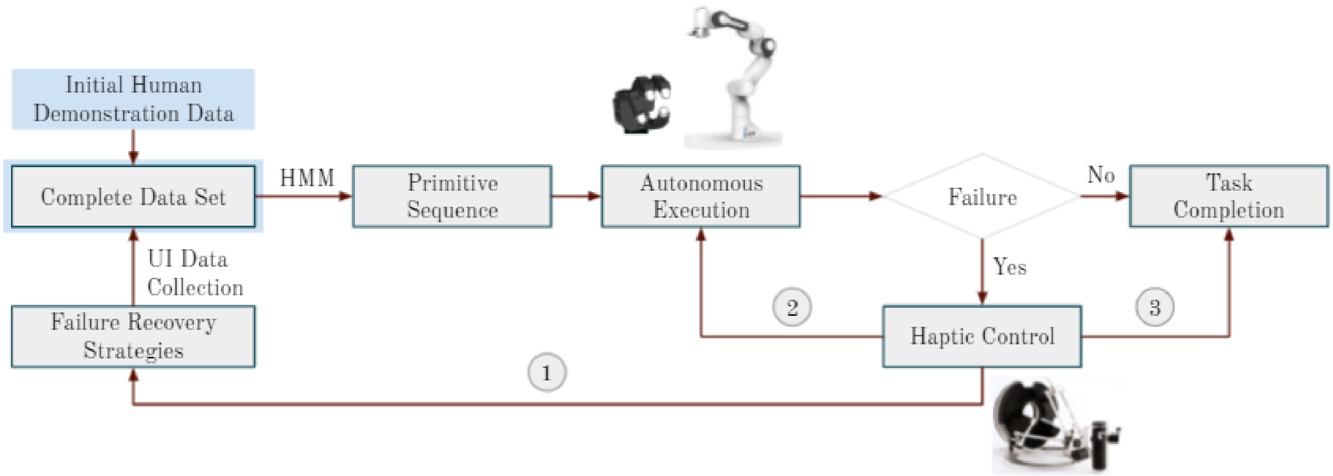

Fig. 2: Concept: We collect a small set of human demonstrations of the complete task. Using prior work we break these demonstrations into a sequence of primitive skills that the robot can perform autonomously. In case of failure we allow for haptic intervention and collect data from the recovery strategy. This data is added to the original data set and used to improve our understanding of the task.

II. Related Work

Manipulation tasks in the real world involve objects of different sizes and shapes in various orientations and positions. As humans, we rely on skills to robustly execute on high level plans but for robots to deal with significant uncertainty we need to rely on a flexible program that is task-driven [1].

A. Contact primitives

Contact primitives are especially useful for object manipulation tasks that involve a sequence of motions such as grasp, move, and release. [2] characterizes each phase as a motion primitive. Each primitive entails a subgoal of the task [2] To optimize these primitives, recent reinforcement learning algorithms have been used problems [3], [4], [5]. [2] highlights that contact is especially challenging because it needs to be adapted to the orientation of the goal object. To enhance existing learning algorithms the researchers proposed a path integrals algorithm to learn the optimal goal of a contact primitive. This study focused on the endpoint of the trajectory generated by the manipulator and equate learning the optimal goal to learning the optimal shape of the motion.

B. Grasping Under Uncertainty

Uncertainty about an object's position can cause failure in manipulation and high-contact tasks. Prior work has deployed strategies to deal with this uncertainty such as sampling motion planners [2]. This approach, however, is costly from a computational perspective and requires an accurate model of the contact object and the environment. Another approach is to use tactile feedback through force and torque sensors to adapt to cases where the object is not located at the expected coordinate [6], [7].

[8] address generalization from human demonstrations to new situations with sequential skills. Imitating a human directly is difficult to match by the robot and also expensive given the need for accurate tracking systems. Hence the authors opt for a kinesthetic teaching approach, in which the human guides the robot motions, combined with reinforcement learning by setting a reward function that favors the robot self exploration. This approach is useful for intuitive programming. This combination helps the authors reduce the number of necessary demonstrations and allows the robot for self-improvement of the skill or primitive.

[8] divide the task in subtasks for ease of implementation and established a dual hierarchy: the lower-level primitives and upper-level sequencing of the primitives. Their approach attempts to answer when to execute each primitive. They labeled the demonstrated data manually according to each skill. One option to determine the primitive sequence is incorporating sensory data in a limited state machine. The following motion is selected by comparing the current sensor data to an expected data and choosing the best match. A k-nearest neighbor classifier has also been used to learn the switching behavior from demonstration data.[9]. The combination of sensed data with primitives has also been described by [1].

C. Robustness and generalization of primitives

Previous approaches to develop generalization capabilities in manipulation control include [10] object-centric task-axes controllers which defines task axis controllers for different subtasks. The axis are set at strategic points of the task objects like the normal of a table or the middle point of a ball and take into account the null space projection. The advantage of object-centric task-axes controllers is the ability to reuse the controllers across different tasks and enhance the robustness of the task execution [10]. More recently, [10] extended this approach by incorporating visual data keypoints to infer the controller parameters directly from visual input instead of relevant 3D positions of the objects and to understand state estimation. This approach is relevant

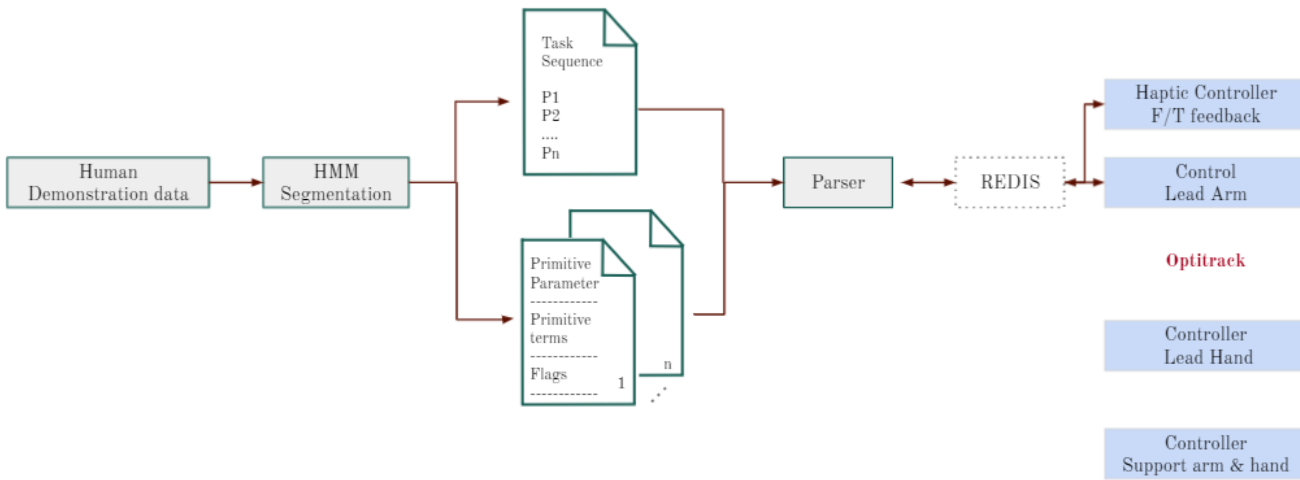

Fig. 3: Data segmentation and robot controls: blablabal

especially in cases in which object geometry data is not known or easily available.

D. Learning from failure

[11], [12] explored what can be learned when the humans do not provide successful examples by developing probabilistic approaches that avoid replication of previous failures.

CITE TOKI

III. Methods

Brief summary of the complete controls. Everything communicates through redis. Briefly compare to other methods. Maybe reference back to related work. Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

A. From Human Data to a Controller Parametrization

- Mention pre, post conditions in here. - Weakly supervised segmentation algorithm (cite IROS) Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

B. The Basic Primitive: Lead Arm Controller

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

C. Regrasping

Cite Marco's thesis Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

D. Contact Driven Control: The Support Arm

Cite Mikael paper (con elena) Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

E. Dual-proxy Haptic Controller

Cite Mikael thesis Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

F. UI for Failure Recovery Data Collection

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

IV. EXPERIMENTS

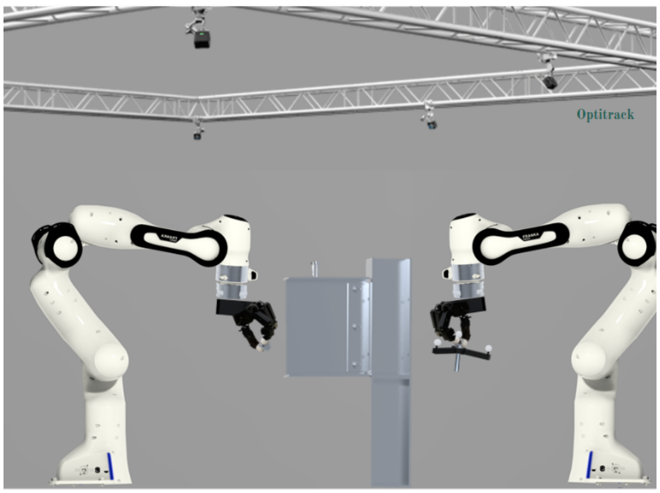

Fig. 4: 3D render of all the components

Why we selected the steel bolting? - Contact task - Requires a primitive sequence - May need to feel your way to an occluded hole making our force sensor approach make more sense than in other scenarios - Routinely performed in construction and very valuable to automate. However, construction environments are very unstructured and dynamic making a fully automated approach very challenging. You can also keep workers involved but performing less physically taxing tasks.

Describe the experimental setup in detail. Maybe mention specifics about the communication rates of different things when explaining all the hardware components involved. Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

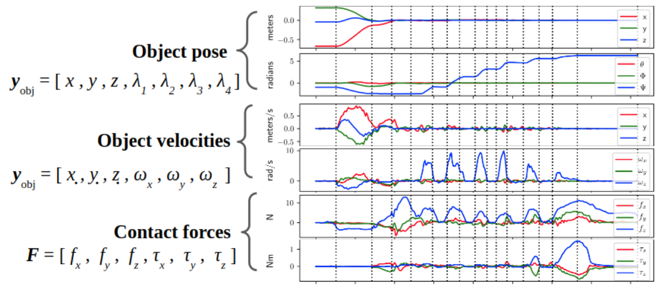

Fig. 5: Sensor data recorded during task execution

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Here is a sample table:

TABLE I: Sample Table

Drywall	**Task Productivity**	**Preparation/day (h)**	**Workers**
Layout	0.46 min/m	0	2
First side drywall (including studs)	6.36 min/m^2	0.5	1
Electrical installation	6.09 min/m^2	0.25	1
Second side drywall	8.4	0.5	1

V. CONCLUSIONS AND FUTURE WORK

Our experiments show that our [framework] is effective at enabling manipulators to learn new tasks with a limited number of demonstrations

- Try other tasks that combine the same primitives - Try - Study the robustness of primitives based on the flagged failures, recovery strategies, and variety of initial conditions.

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Acknowledgments

Thank you to everyone in the robotics lab! Particularly Mikael for his invaluable guidance and Marco Speziali for his help with rendering and grasping advice.

References

[1] D. Morrow and P. Khosla, "Manipulation Task Primitives for Composing Robot Skills," in *Proceedings of the 1997 IEEE International Conference on Robotics and Automation*. Albuquerque, New Mexico: IEEE, 4 1997. [Online]. Available: https://www.ri.cmu.edu/pub_files/pub2/morrow_james_1997_1/morrow_james_1997_1.pdf

[2] E. A. S. S. Stulp, Freek; Theodorou, "Reinforcement Learning With Sequences of Motion Primitives for Robust Manipulation." [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6295672

[3] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler, and A. Ng, "ROS: an open-source Robot Operating System." [Online]. Available: http://stair.stanford.edu

[4] E. Theodorou, "Iterative path integral stochastic optimal control: Theory and applications to motor control," Ph.D. dissertation, 2011.

[5] A. Ude, A. Gams, T. Asfour, and J. Morimoto, "Task-specific generalization of discrete and periodic dynamic movement primitives," *IEEE Transactions on Robotics*, vol. 26, no. 5, pp. 800–815, 10 2010.

[6] A. J. Ijspeert, J. Nakanishi, and S. Schaal, "Movement imitation with nonlinear dynamical systems in humanoid robots," *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2, pp. 1398–1403, 2002.

[7] S. Gomez-Gonzalez, G. Neumann, B. Schölkopf, and J. Peters, "Adaptation and Robust Learning of Probabilistic Movement Primitives," 2017.

[8] S. Manschitz, J. Kober, M. Gienger, and J. Peters, "Learning to sequence movement primitives from demonstrations," *IEEE International Conference on Intelligent Robots and Systems*, no. Iros, pp. 4414–4421, 2014.

[9] S. Niekum, S. Chitta, A. Barto, B. Marthi, and S. Osentoski, "Incremental Semantically Grounded Learning from Demonstration," 1 2016.

[10] M. Sharma and O. Kroemer, "Generalizing Object-Centric Task-Axes Controllers using Keypoints," *arxiv*, 2021.

[11] D. H. Grollman and A. G. Billard, "Robot Learning from Failed Demonstrations," *International Journal of Social Robotics 2012 4:4*, vol. 4, no. 4, pp. 331–342, 6 2012. [Online]. Available: https://link.springer.com/article/10.1007/s12369-012-0161-z

[12] D. H. Grollman and A. Billard, "Donut as I do: Learning from failed demonstrations."