Learning from Demonstration for Distributed, Encapsulated Evolution of Autonomous Outdoor Robots

Travis DeVault
University of Idaho
Moscow, Idaho 83844
zerill@gmail.com

Seth Forrest
University of Idaho
Moscow, Idaho 83844
forr9430@vandals.uidaho.edu

Ian Tanimoto
University of Idaho
Moscow, Idaho 83844
tani0633@vandals.uidaho.edu

Terence Soule
University of Idaho
Moscow, Idaho 83844
tsoule@uidaho.edu

Robert B. Heckendorn
University of Idaho
Moscow, Idaho 83844
heckendo@uidaho.edu

ABSTRACT

In learning from demonstration (LfD) a human trainer demonstrates desired behaviors to a robotic agent, creating a training set that the agent can learn from. LfD allows non-programmers to easily and naturally train robotic agents to perform specific tasks. However, to date most LfD has focused on single robot, single trainer paradigms leading to bottlenecks in both the time required to demonstrate tasks and the time required to learn behaviors. A previously untested, approach to addressing these limitations is to use distributed LfD with a distributed, evolutionary algorithm. Distributed on-board learning is a model for robust real world learning without the need for a central computer. In the distributed LfD system presented here multiple trainers train multiple robots on different, but related, tasks in parallel and each robot runs its own on-board evolutionary algorithm. The robots share the training data, reducing the total time required for demonstrations, and exchange promising individuals as in typical island models. These experiments compare robotic performance on a task after distributing the behaviors or the simple demonstrations to performance using a non-distributed LfD model receiving complex demonstrations. Our results show a strong improvement to behavior when using distributed simple demonstrations.

1. INTRODUCTION

In this paper we present a distributed, evolutionary approach to learning from demonstration (LfD) for teaching mobile COTSBOTS[2]. In LfD [1] a trainer performs the target actions and the agent records the both the current state and the demonstrated actions to build a training set of state-action pairs that it learns from.

[1]Also known as learning by demonstration and closely related to imitation learning and apprenticeship learning.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

GECCO '15 July 11-15, 2015, Madrid, Spain

© 2015 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-3488-4/15/07.
DOI: http://dx.doi.org/10.1145/2739482.2764721

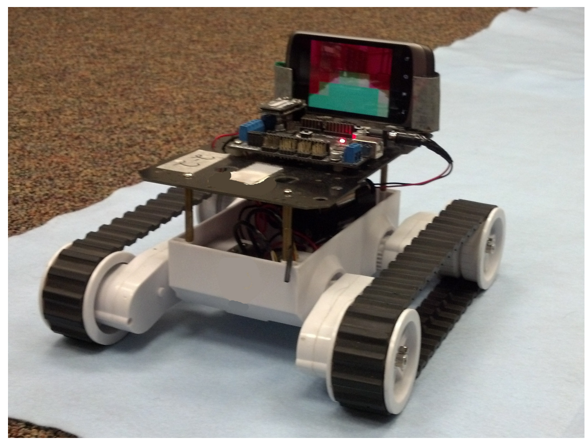

Figure 1: The robot. The smart phone receives wireless commands from the user and forwards the commands via bluetooth to an Arduiono type microcontroller that controls the robot's motors. The smart phone can perform image processing and onboard learning.

In LfD the learning process is typically on-line, i.e. the agent is learning during the demonstrations so that it can perform the task immediately, or at least shortly, after the demonstrations are complete. However, as the complexity of the behavior increases so does the learning time required, which makes it more difficult for agents to learn behaviors as they are demonstrated[1]. This can be a particular issue in LfD because the process is often cyclic: the trainer demonstrates a task or behavior, the agent attempts to replicate it.

To overcome these two difficulties: longer training times and more difficult search spaces; we present a *distributed*, evolutionary LfD approach in which multiple trainers train multiple robots in parallel, each robot runs its own on-board evolutionary algorithm following the work done by Soule[3] and exchanges elite individuals with the other robots, thereby

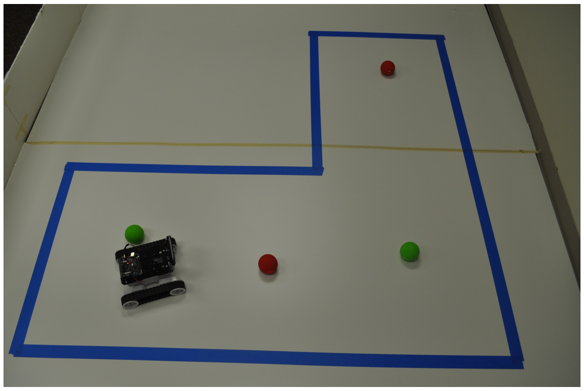

Figure 2: The robot was trained by collecting demonstrations in this training environment. For demonstrations requiring only red and green balls, the robot was positioned to not see the blue tape.

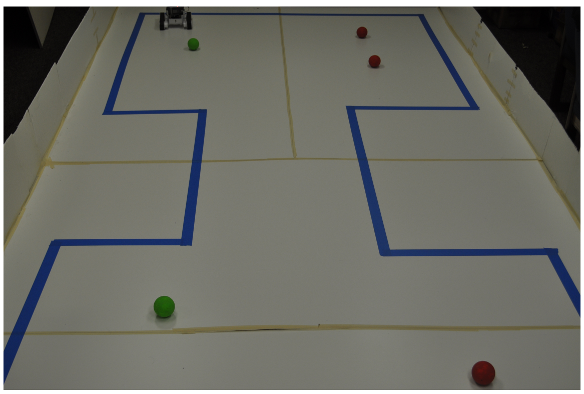

Figure 3: The robot was tested in an environment different from it's training area. During the testing phase, the robot autonomously navigates the testing environment while human observers watch the robot's behavior.

creating a distributed island model. Showing that such an approach is feasible significantly increases Evolutionary Computation's (EC) suitability as a learning technique for LfD by mitigating issues of efficiency of search.

2. EXPERIMENT DESCRIPTION

In order to show distributed evolutionary computation's suitability in training a LfD system, we created an environment to test the robots behaviors in motion, and in interaction with objects. The environment has borders which the robot is trained not to cross, and different colored balls which the robot was trained to search for and stop at once the robot moved close enough to the ball.

Our experiment uses three types of evolution.

Evolution 1: These controllers are evolved with no distribution. The robot is trained given demonstrations including the entire training environment. The robot is able to see borders, and different colored balls together.

Evolution 2: Three robots are trained on individual behaviors. One robot was only given demonstrations with the red balls, another robot received demonstrations involving the green balls, and the last robot was only given demonstrations with the blue border of the environment. Robots in these trials were evolving by distributing their best individuals during every generation of evolution. Evolution was also done on these demonstrations without using distribution to allow analysis of the behavior.

Evolution 3: Demonstrations for these robots were collected by combining the demonstration data from *Evolution 2*. This is the case of distributed demonstration collection. The robot does not receive behavior data from other robots.

The robot was given 90 demonstrations in each evolution case. Demonstrations are composed of the input the neural network would receive and the action taken by the trainer. Demonstrations are collect when the trainer puts the robot into LfD mode and then drives the robot by giving it discrete commands from the remote control. The commands that the trainer has access to are: Left, Forward, Right, Backward, and Stop.

To test the robot's behavior we used a blind survey similar to how behavioral scientist judge the behavior of animals in the wild. A human observer was gives a description of the behaviors the robot was trained to perform, and then without know what type of training was used, score the robot's behavior between 1-5. 1 meaning that the robot did not appear to demonstrate that behavior at all, and 5 meaning the robot demonstrated the behavior every time, in multiple positions.

3. CONCLUSIONS

Our results show that distributed LfD, using multiple robots, is a practical approach to reducing the time required to produce sufficiently rich training sets for learning. In addition, leveraging the multiple robots to create a distributed, island model, evolutionary algorithm will not, by itself, allow for behavior migration. We also show that collecting specific examples of behavior to create a demonstration set or distributed training, provides more robust behavior than a demonstration set create in a noisy environment.

Future work will include using learning from demonstration to navigate the robots in outdoor fields to seek out invasive species in lawns and agricultural fields.

4. ACKNOWLEDGMENTS

This material is based in part upon work supported by the National Science Foundation under Cooperative Agreement No. DBI-0939454. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

5. REFERENCES

[1] S. Chernova and M. Veloso. Confidence-based multi-robot learning from demonstration. *International Journal of Social Robotics*, 2(2):195–215, 2010.

[2] T. Soule and R. B. Heckendorn. Cotsbots: computationally powerful, low-cost robots for computer science curriculums. *Journal of Computing Sciences in Colleges*, 27(1):180–187, 2011.

[3] T. Soule and R. B. Heckendorn. A practical platform for on-line genetic programming for robotics. In *Genetic Programming Theory and Practice X*, pages 15–29. Springer, 2013.