Name: William Curran

Address: 2600 NW Century Dr. Apt 229

                   Corvallis, Oregon 97330

E-mail: curranw@oregonstate.edu

Phone: 541-670-1382

Home Page: http://oregonstate.edu/~curranw

Gender: Male

Citizenship: United States

Graduation: 2016

Advisor: Professor William Smart

                   School of Mechanical, Industrial and Manufacturing Engineering

Advisor Address: Oregon State University

                   219A Dearborn Hall

                   Corvallis, Oregon 97330

Advisor E-mail   bill.smart@oregonstate.edu

Mentors: George Konidaris

                   Sonia Chernova

                   Andrea Thomaz

                   Michael Littman

                   Maya Cakmak

Keywords: Reinforcement Learning

                   Robotics

                   Dimensionality Reduction

                   Transfer Learning

                   Learning from Demonstration

Robust Learning from Demonstration Techniques and Tools

William Curran
Oregon State University
Corvallis, Oregon
curranw@oregonstate.edu

Abstract

Large state spaces and the curse of dimensionality contribute to the complexity of a task. Learning from demonstration techniques can be combined with reinforcement learning to narrow the exploration space of an agent, but require consistent and accurate demonstrations, as well as the state-action pairs for an entire demonstration. Individuals with severe motor disabilities are often slow and prone to human errors in demonstrations while teaching. My dissertation develops tools to allow persons with severe motor disabilities, and individuals in general, to train these systems. To handle these large state spaces as well as human error, we developed Dimensionality Reduced Reinforcement Learning. To accommodate slower feedback, we will develop a movie-reel style learning from demonstration interface.

Introduction

The main goal of my thesis is to help put robots into real homes to help those with severe physical disabilities. Some individuals have minimal ability to move and speak, and need extended care all day, every day. The strain on families needing to take care of these individuals is enormous. Inexperienced family caregivers use prescription drugs for depression, anxiety, and insomnia two to three times more often than the average population (Gallagher et al. 1989). Robotics can be used to assist those with extreme disabilities and remove much of this burden on the family.

These individuals with disabilities are often non-experts, and currently cannot be part of the development process. Yet, most don't want to wait for someone to program autonomous behaviors for all household tasks. We propose to give disabled non-expert users the tools to teach robots tasks by themselves. This will give the disabled user both independence and personal customization.

There are two key difficulties with learning from demonstration interfaces for individuals with disabilities; resolving these are the key contributions of my dissertation. State-of-the-art personal robots need to perform complex manipulation tasks to be viable in assistive scenarios. These complex robots lead to large dimensional state spaces. Also, the ability to perform good demonstrations and give timely feedback are key assumptions in learning from demonstration

research (Argall et al. 2009). My dissertation addresses these assumptions.

The PR2 robot has two 7 DoF arms. When learning position, velocity, and acceleration control, this leads to a 21 dimensional state space per arm. Learning in these high-dimensional spaces is computationally intractable without approximation techniques. My dissertation introduces Dimensionality Reduced Reinforcement Learning (DRRL) to speed up learning in high-dimensional spaces. DRRL uses demonstrations to compute a projection to a low-dimensional manifold. In each learning iteration, we project the current state onto this manifold, compute and execute an action, project the new state onto the manifold, and perform a reinforcement learning update. This reduces the number of demonstrations and increases the resistance to suboptimal outliers. These are desirable characteristics in general, but especially for non-expert use of the learning algorithm.

Additionally, we plan to develop a movie-reel style interface for detailed, time insensitive feedback. This gives the user as much time as they need for demonstrations and feedback.

Learning from Demonstration in High-Dimensional Spaces

To handle high-dimensional spaces, we performed dimensionality reduction on a set of demonstrated trajectories. We used these demonstrations to compute a projection to a low-dimensional manifold. In each learning iteration, we projected the current state onto this manifold, computed and executed an action, projected the new state onto the manifold, and performed a reinforcement learning update.

We have shown that an agent can learn quickly in the low-dimensional manifold. However, this led to a critical trade-off. By projecting onto a low-dimensional manifold, we discarded low variance, yet potentially important data. By adding DRRL to an existing algorithm, we showed that the agent can quickly converge to a good, yet suboptimal, policy much faster.

We extended DRRL to add an iterative component (IDRRL). Instead of learning entirely in one manifold, we iteratively learned in all manifolds by using transfer learning. The agent can quickly learn in a low-dimensional space d, and transfer that knowledge to the $d + 1$ dimensional
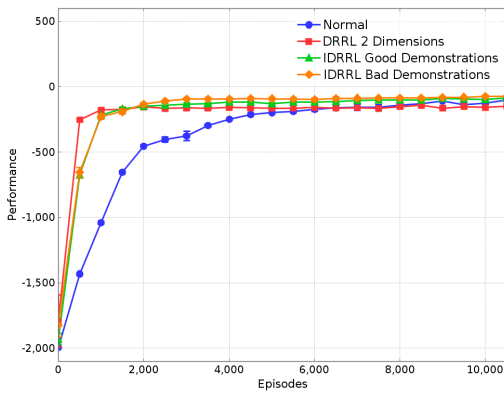
Figure 1: DRRL and IDRRL were combined with standard Q-Learning. DRRL converged faster, but to a slightly suboptimal solution. IDRRL converged quickly to an optimal solution. Additionally, IDRRL is robust to suboptimal demonstrations in Mountain Car 3D.

space using the known mapping between the spaces. We have shown the speed and robustness of this approach when combined with standard Q-Learning in the Mountaincar 3D (Figure 1) domain using good and bad demonstrations. Increasing the robustness to human error in demonstrations and reducing the number of demonstrations are essential. Since we work with individuals with disabilities, physical demonstrations are difficult, and expect suboptimal demonstrations. Since demonstrations are difficult, fewer demonstrations are also a desirable quality.

We plan to extend this work to robotic control. By parameterizing the motor primitives in a low-dimensional manifold, we hypothesize that we will be able to more easily generalize motor skills to similar tasks. This will reduce the number of different demonstrations required. Additionally, dimensionality reduction smooths the lower-dimensional trajectory. We hypothesize that this makes the learning more robust to suboptimal outliers caused by human error.

Learning from Demonstration Interfaces

We need to develop a learning from demonstration interface that allows individuals with disabilities to teach their assistive robot at their own speed. We propose two interfaces: an early interface for super users, and one for users with severe physical disabilities. In this work we will use RViz, a ROS package that assists users in the visualization of robotic movements and sensors (Quigley et al. 2009). Our first focus of research is to extend RViz to be an interface for learning from demonstration.

The interface for users without disabilities will take advantage of the direct first-person control aspect of RViz. We will also use the Oculus Rift and Razer Hydra control systems. The Oculus Rift allows a user to see through the sensors of the robot, while the Hydra controls the robot. The Hydra control system computes the exact location and orientation of controllers in your hands, allowing the user more natural control of the robotic arms. Using these tools the user

can efficiently teach the robot through demonstration.

The interface for users with disabilities will use a less direct approach. The robot will come with a motion library of basic autonomous functions, such as turning knobs and picking up objects. First, the user will give the robot a sequence of these high level actions to perform the task. The robot will simulate itself doing the task in RViz. We will extend RViz with a movie-reel style interface for rewinding and fast forwarding of the robot's simulation. The user can then provide feedback on an action at his or her own speed. This learning from demonstration interface allows the users to build custom routines for the robot without direct knowledge of the reinforcement learning algorithm.

To analyze the initial efficacy of the movie-reel style interface there will be a within-subjects study, in which the users will be able-bodied. We will take objective measurements such as the total task time and the user's cognitive load, as well as subjective measurements with questionnaires. We will then perform another study with disabled individuals that have previously worked with our lab. Comparing the results will give us a better understanding of the efficacy of the interface.

Conclusion

Those suffering from ALS and quadriplegia need assistive robots now. These individuals cannot wait for someone to program autonomous behaviors for all household tasks. This work intends to help those suffering from severe physical disabilities by giving them the ability to teach the robot themselves.

This work introduces a new learning from demonstration technique that utilizes human demonstration from non-experts. This requires robots to learn from human demonstrations, even when those demonstrations are suboptimal. By learning in a lower-dimensional space, we hypothesize that it will be easier to generalize between similar tasks, as well as ease the learning of trajectories with outliers.

With these new interfaces and learning advances, individuals with disabilities will be able to do day-to-day tasks without help from others. The lack of a human assistant performing the task and the addition of positive experiences like teaching, doing it yourself, and being more independent increases the quality of life for people with extreme disabilities.

References

Argall, B. D.; Chernova, S.; Veloso, M.; and Browning, B. 2009. A survey of robot learning from demonstration. *Robot. Auton. Syst.* 57(5):469–483.

Gallagher, D.; Rose, J.; Rivera, P.; Lovett, S.; and Thompson, L. W. 1989. Prevalence of depression in family caregivers. *The Gerontologist* 29(4):449–456.

Quigley, M.; Conley, K.; Gerkey, B. P.; Faust, J.; Foote, T.; Leibs, J.; Wheeler, R.; and Ng, A. Y. 2009. ROS: An open-source robot operating system. In *ICRA Workshop on Open Source Software*.

William Curran

Contact Information	William Curran 2600 NW Century Dr. Apt 229 Corvallis, OR 97330 *Mobile:* 541-670-1382 *E-mail:* curranw@oregonstate.edu *Website:* http://oregonstate.edu/~curranw

Research Interests

Machine Learning, Multiagent Reinforcement Learning, Learning-Based Controls, Human-Robot Interaction.

Education

Oregon State University, Corvallis, OR

PhD, School of Mechanical, Industrial and Manufacturing Engineering, Expected 2016
- Adviser: Professor Bill Smart
- Area of Study: Human-Robot Interaction and Learning from Demonstration

M.S., School of Electrical Engineering and Computer Science, June 2012
- Adviser: Professor Kagan Tumer
- Area of Study: Multi-Agent Systems and Learning Based Controls

B.S., School of Electrical Engineering and Computer Science, June 2010
- Emphasis on Machine Learning and Artificial Intelligence
- *Cum Laude*

Conference Publications

[1] Curran, W., Brys, T., Aha, D., Taylor, M, and Smart, B. Dimensionality Reduced Reinforcement Learning. Submitted To: *AAAI Conference on Artificial Intelligence 2016* February 12–17, 2016

[2] Curran, W., Thornton, T., Arvey, B, and Smart, B. Evaluating Impact in the ROS Ecosystem. In: *IEEE International Conference on Robotics and Automation 2015* May 26–30, 2015

[3] Colby, M., Curran, W., Rebhuhn, C., and Tumer, K. Approximating Difference Evaluations with Local Knowledge. In: *International Conference on Autonomous Agents and Multiagent Systems 2015* May 4–8, 2015

[4] Curran, W. Developing Learning from Demonstration Techniques for Individuals with Physical Disabilities. In: *International Conference on Human-Robot Interaction. HRI Pioneers Workshop* March 2–5, 2015

[5] Curran, W., Agogino, A., Tumer, K. Agent Partitioning with Reward/Utility-Based Impact In: *AAAI Conference on Artificial Intelligence 2015. Multiagent Interaction without Prior Coordination Workshop* Jan 25–30, 2015

[6] Curran, W., Arvey, B., Thornton, T. and Smart, B. The ROS Ecosystem: Impacts, Insights and Improvements. In: *ROSCon 2014* September 12–13, 2014

[7] Lazewatsky, D., Bowie C., Curran, W., LaFortune, J., Narin, B., Nguyen, D., Wyman, A., Smart, W. Wearable Computing to Enable Robot Microinteractions. In: *Robot and Human Interactive Communication 2014* August 25–29, 2014

[8] Curran, W., Agogino, A., Tumer, K. Hierarchical Simulation for Complex Domains: Air Traffic Flow Management. In: *Genetic and Evolutionary Computation Conference 2014* July 12–16, 2014

[9] Curran, W., Agogino, A., Tumer, K. Using Reward/Utility Based Impact Scores in Partitioning. In: *International Conference on Autonomous Agents and Multiagent Systems 2013* May 5–9, 2014

[10] Curran, W., Agogino, A., Tumer, K. Partitioning Agents and Shaping Their Evaluation Functions in Air Traffic Problems with Hard Constraints. In: *Genetic and Evolutionary Computation Conference 2013* July 6–10, 2013

[11] Curran, W., Agogino, A., Tumer, K. Addressing Hard Constraints in the Air Traffic Problem through Partitioning and Difference Rewards. In: *International Conference on Autonomous Agents and Multiagent Systems 2013* May 6–10, 2013

Experience

Oregon State University, Corvallis, OR

Personal Robotics Lab *September 2013 to Current*
- Developing learning from demonstration algorithms for severely disabled individuals with amyotrophic lateral sclerosis and quadriplegia. This involves looking into effective interfaces for those with severe disabilities, developing shared autonomy learning algorithms, employing multi-robot coordination for efficient task execution, and how to learn from highly suboptimal human demonstrations.
  - Dr. Bill Smart, Oregon State, MIME Department

Autonomous Agents and Distributed Intelligence *Winter 2011 to June 2013*
- Currently working on Air Traffic Control with a focus on multi agent reinforcement learning. This involves learning and using the air traffic simulator FEATS with the goal to apply reinforcement learning techniques in an effort to alleviate the air traffic congestion problem. Current work involves removing congestion and optimizing delay in the national airspace by controlling ground delay imposed upon aircraft.
  - Dr. Kagan Tumer, Oregon State, MIME Department

Navy Research Lab, Washington, DC

NRL Student Contractor *June 2015 to September 2015*
- Developed the Dimensionality Reduced Reinforcement Learning (DRRL) algorithm for learning in high-dimensional spaces. It uses dimensionality reduction on demonstrations to compute a low-dimensional manifold in which to learn in.
  - Dr. David Aha, Navy Research Lab, Navy Center for Applied Research in Artificial Intelligence

NASA AMES Moffet Field, Mountainview, CA

NASA AMES Intern *June 2013 to September 2013*
- Developed a new simulation approach leveraging the concept of a multi-fidelity simulation. Computationally complex simulation is used only when agents are in a potential separation violation, while a low fidelity simulation is used for navigation. The key research focus is when to change from one fidelity to another, as small nuanced issues can easily occur that impedes agent learning.
  - Dr. Adrian Agogino, NASA AMES, Moffett Field

NASA AMES Intern *June 2012 to September 2012*
- Extended my work on Air Traffic Control with adviser Adrian Agogino. Here I extended previous Air Traffic Control work by simulating the entire NAS (35,000 planes) rather than a small subset of the NAS, and developed an automated partitioning system for aircraft to learn quickly and efficiently.
  - Dr. Adrian Agogino, NASA AMES, Moffett Field

September 22, 2015

To the AAAI-16 Doctoral Consortium Committee,

Attending AAAI-16 gives me an excellent opportunity to learn how the AI community can assist in my dissertation topic. During the next year I will be preparing my proposal, and by attending AAAI-16 I have the opportunity to discuss my research directions with experts in the field.

My dissertation develops robot learning from demonstration techniques and tools for individuals with Amyotrophic Lateral Sclerosis (ALS). This requires working with robots in high-dimensional state spaces. Learning in these high-dimensional state spaces with suboptimal and infrequent demonstrations are the challenges in this work. The scalability and robustness requirement in my research is applicable in many artificial intelligence applications outside of my own. The algorithm I have developed leverages knowledge in transfer learning, learning from demonstration, reinforcement learning and dimensionality reduction. This gives me a broad understanding in many AI topics. In the doctoral consortium I can contribute by presenting and discussing a new algorithm, as well as assist others with my background knowledge.

I usually submit to the International Conference on Robotics and Automation (ICRA), the International Conference on Autonomous Agents and Multiagent Systems (AAMAS), the Conference on Human-Robot Interaction (HRI), and this conference. I am not submitting to any other doctoral consortium this year, and I have not attended one in the past.

Thank you,
William Curran