# RGame: Embodied Gaming for Robot Learning by Demonstration

**Daniel H Grollman** and **Odest Chadwicke Jenkins** and **Jesse Butterfield**
Brown University Department of Computer Science
Providence, RI 02912-1910
{dang,cjenkins,jbutterf}@cs.brown.edu

## Overview

We here demonstrate robot learning from demonstration using interactive tutelage. Our Dogged Learning architecture (introduced in (Grollman and Jenkins 2007a) and shown in Figure 1) combines concepts of teleoperative demonstration, mixed-initiative control, feedback and transparency, and real time policy inference. It is designed to be abstract, applicable to many platforms (robots), demonstration interfaces, and learning algorithms.

Briefly, a robot platform provides perceptual inputs to the system, which are displayed to a demonstrator, as well as analyzed by a learned approximate policy. Both the demonstrator and the learner generate desired action outputs, as well as confidence values. The outputs can be NULL, and confidences 0, if either is not present. Using the confidences, the system arbitrates between the two outputs by, for example, taking the more confident output. Results of arbitration are displayed to the user, and the chosen action is passed to the robot, which performs as commanded. Additionally, if the commanded action is from the demonstrator, the action is paired with the perception that led to it, and the pair is used by the learner to update its policy approximation.

The DL architecture is embedded within the RGame program, which allows for users who are not co-located with the robots to generate demonstration data for them. Further, the system collects data for learning, and manages the learned policies. It is a goal of our work that remote users can log in, train robots to perform tasks, and then continue to monitor and improve their performance over time.

## Learning

Within DL, we use regression to directly approximate the mapping from perception to actuation. Both spaces are usually multidimensional and continuous, although discrete spaces are also allowed. In order to obtain realtime performance, which is necessary for interactive tutelage, we have thus been exploring incremental, sparse algorithms.

Initial experiments (Grollman and Jenkins 2008) used Locally Weighted Projection Regression (Vijayakumar, D'Souza, and Schaal 2005) and Sparse Online Gaussian Processes (Csató 2002) to learn basic robot soccer skills on a Sony AIBO robotic dog. However, issues of perceptual aliasing cause violations of the underlying assumption that observed actions are unimodally distributed around the true desired one. That is, due to hidden state, a particular perception may lead to different actions. Extending the perception to include this state is one method to alleviate this issue (Grollman and Jenkins 2007b), but is not generally applicable.

Instead, we have developed a nonparametric Bayesian technique (Wood et al. 2008) for regression with multimodal output distributions, or multimap (one-to-many) policies. We consider a possibly infinite number of modes of the distribution, and estimate both the number of modes and their properties incrementally. We liken this approach to addressing the issues of model selection and subtask learning in finite state machine estimation. Using this learner, we hope to learn policies dependant upon hidden state without requiring explicit representations of that state.

## Interface

A snapshot of the user interface is shown in Figure 2, along with several possible interface devices. Our system is based on consumer electronics such as the Nintendo Wii (Lapping-Carr et al. 2008), and video games (Byers et al. 2008) in order to be approachable by general users. Using this interface, users teleoperate robots to perform tasks in the context of a game, such as soccer. When a user ceases to control the robot, an approximate policy based on the user's commanded controls can be used instead, for continuing robot behavior.

In our system, users can select from a variety of control devices and robots in order to compare the relative benefits of each for themselves. Further, collected data can be reused for empirical comparisons of learning techniques and the development of new approaches.

## References

Byers, D.; Lapping-Carr, M.; Kumar, J.; Hinkle, T.; Grollman, D.; and Jenkins, C. 2008. Game-based learning. AAAI2008 Video Competition.

Csató, L. 2002. *Gaussian Processes - Iterative Sparse Approximations*. Ph.D. Dissertation, Aston University.

Grollman, D. H., and Jenkins, O. C. 2007a. Dogged learning for robots. In *International Conference on Robotics and Automation*, 2483 – 2488.

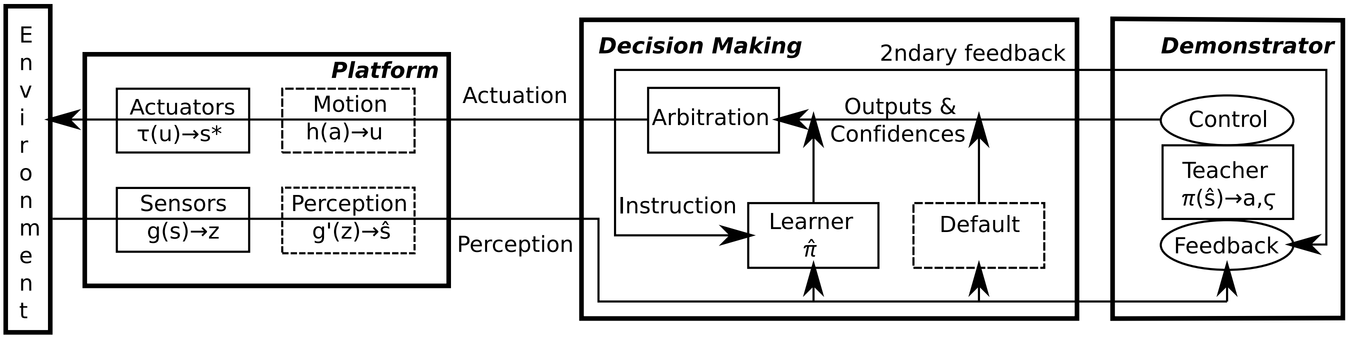
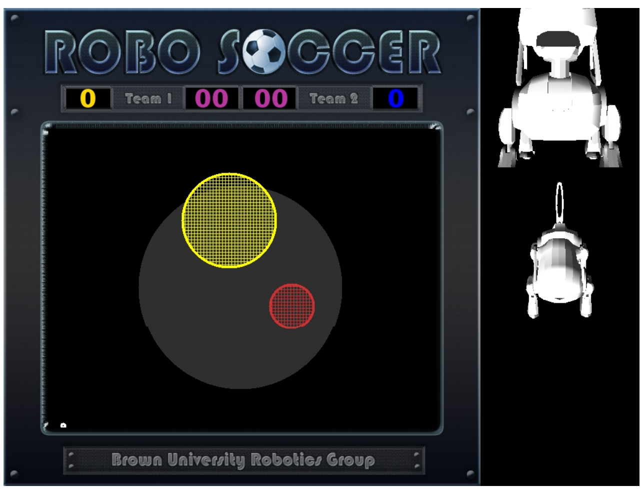



Figure 1: The Dogged Learning architecture, which underlies RGame. It is abstract, allowing for multiple possible platforms (robots), learning algorithms, and demonstration devices.




(a) RGame Display

(b) Control Interfaces

Figure 2: Above: The RGame interface for remotely demonstrating tasks to a robot. Here the robot is an AIBO, and the perceptual space consists of color blobs and motor positions. Below: Possible control interfaces for use by the human demonstrator.

Grollman, D. H., and Jenkins, O. C. 2007b. Learning robot soccer skills from demonstration. In *International Conference on Development and Learning*, 276–281.

Grollman, D. H., and Jenkins, O. C. 2008. Sparse incremental learning for interactive robot control policy estimation. In *International Conference on Robotics and Automation*, 3315–3320.

Lapping-Carr, M.; Jenkins, O. C.; Grollman, D. H.; Schwertfeger, J. N.; and Hinkle, T. R. 2008. Wiimote interfaces for lifelong robot learning. In *AAAI Spring Symposium*.

Vijayakumar, S.; D'Souza, A.; and Schaal, S. 2005. Incremental online learning in high dimensions. *Neural Computation* 17(12):2602–2634.

Wood, F.; Grollman, D. H.; Heller, K. A.; Jenkins, O. C.; and Black, M. 2008. Incremental Nonparametric Bayesian Regression. Technical Report CS-08-07, Brown University Department of Computer Science.