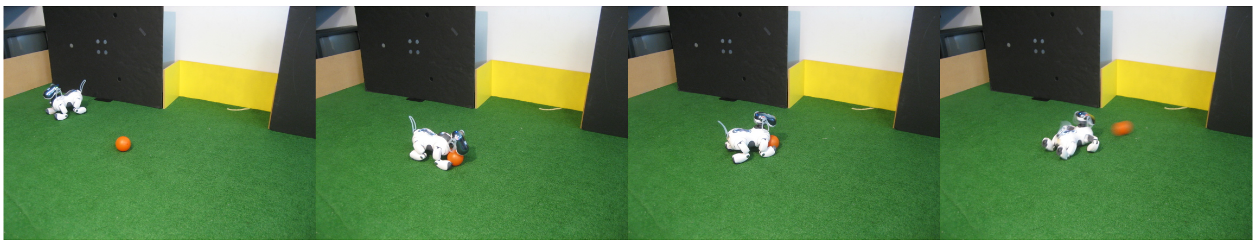
Multimap Regression for Perceptual Aliasing in Learning Finite State Machine Robot Controllers from Interactive Demonstration

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



(a) Seek (b) Trap (c) Aim (d) Kick

Figure 1: The goal scoring behavior that we wish to learn from demonstration. Standard regression algorithms assume a many-to-one mapping from perception to actuation. However, due to perceptual aliasing, that assumption is false for this task. Multimap regression algorithms may be one way to enable learning of finite state machine controllers such as these directly from perception-actuation tuples.

We argue for the development and use of multimap regression algorithms for interactive **Robot Learning from Demonstration** (RLfD) as a step towards learning **Finite State Machine** (FSM) based policies. Specifically, current techniques learn FSMs by combining known subtasks into overall behaviors [7] or learning individual subtasks given the transitions between them [8]. What is needed are techniques for discovering the number of subtasks and their policies from demonstrated input-output pairs. This estimation can be viewed as inferring multiple overlapping many-to-one mappings.

Regression algorithms are appropriate for RLfD, since from demonstration we get pairs of robot perception ($\hat{\mathbf{s}}$) and actuation ($\mathbf{a}$), and seek to learn the policy or mapping $\pi(\hat{\mathbf{s}}) \rightarrow \mathbf{a}$. Interactive training, or tutelage, requires a computationally efficient algorithm that is capable of updating an approximated policy as data arrives, and generating control signals to drive the robot as needed. Due to our desire for interactive capabilities, we have thus considered incremental, sparse regression algorithms such as Sparse Online Gaussian Processes [2] and Locally Weighted Projection Regression [9] for policy estimation [4].

These approaches work well when the underlying mapping is many-to-one. However, in FSM based controllers such as the robot soccer goal scorer shown in Figure 1, the underlying mapping is one-to-many (a *multimap*). That is, in certain world states, such as when the robot is checking the location of the goal in Figure 1c, the perceptual information available to the robot does not uniquely determine the correct action to perform[1]. Demonstration data may thus have multiple possible correct actions associated with a given perception. If the observed data is assumed to be unimodally distributed around the true action, as it is in standard regression, the learned policy may average these observations, resulting in inappropriate robot behavior.

This issue, where one state requires different actions based on unobservable information, is known as *perceptual aliasing*. It can arise when there is state that the robot cannot perceive (hidden state [5]), inherent ambiguity in the task (equivalent actions [1]), or an alteration in task goals (subtask switch). The last of these is what occurs in FSMs, when in addition to the perception, a controller must also know the current subtask being performed in order to behave correctly.

Multimap regression can be used to directly approximate multimaps from perceptually aliased data. One approach is an infinite mixture of experts model where new experts (subtasks) are continually hypothesized as needed to approximate the observed data [6]. An incremental, sparse extension may be appropriate for robot tutelage and be able to learn subtasks as demonstration occurs. By incorporating techniques that infer the transitions between subtasks from the same data [3], full FSMs may be inferable.

[1]The robot cannot simultaneously view the goal and the ball, as the camera is located in the nose.

References

[1] Sonia Chernova and Manuela Veloso. Learning equivalent action choices from demonstration. In *International Conference on Intelligent Robots and Systems*, Nice, France, September 2008.

[2] Lehel Csató. *Gaussian Processes - Iterative Sparse Approximations*. PhD thesis, Aston University, March 2002.

[3] Emily B. Fox, Erik B. Sudderth, Michael I. Jordan, and Alan S. Willsky. An hdp-hmm for systems with state persistence. In *International Conference on Machine Learning*, pages 312–319, Helsinki, Finland, July 2008.

[4] Daniel H Grollman and Odest Chadwicke Jenkins. Sparse incremental learning for interactive robot control policy estimation. In *International Conference on Robotics and Automation*, pages 3315–3320, Pasadena, CA, USA, May 2008.

[5] Andrew Kachites Mccallum. *Reinforcement learning with selective perception and hidden state*. PhD thesis, The University of Rochester, 1996.

[6] Edward Meeds and Simon Osindero. An alternative infinite mixture of gaussian process experts. In *Neural Information Processing Systems*, pages 883–890, Whistler, CAN, December 2006.

[7] Monica Nicolescu and Maja J Matarić. Natural methods for robot task learning: Instructive demonstration, generalization and practice. In *International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 241–248, Melbourne, AUS, July 2003.

[8] Peter Stone and Manuela M. Veloso. Layered learning. In *European Conference on Machine Learning*, volume 1810, pages 369–381, Barcelona, Catalonia, Spain, May 2000. Springer, Berlin.

[9] Sethu Vijayakumar, Aaron D'Souza, and Stefan Schaal. Lwpr: A scalable method for incremental online learning in high dimensions. Technical Report EDI-INF-RR-0284, Edinburgh, 2005.