ROBOT LEARNING

Summary of the post-NIPS workshop
Vail, Colorado, Dec 5th, 1992

David A. Cohn (MIT) Tom Mitchell (CMU) Sebastian Thrun (CMU)
cohn@psyche.mit.edu mitchell@cs.cmu.edu thrun@cs.cmu.edu

Robot learning has grasped the attention of many researchers over the past few years. Previous robotics research has demonstrated the difficulty of manually encoding sufficiently accurate models of the robot and its environment to succeed at complex tasks. Recently a wide variety of learning techniques ranging from statistical calibration techniques to neural networks and reinforcement learning have been applied to problems of perception, modeling and control. Robot learning is characterized by sensor noise, control error, dynamically changing environments and the opportunity for learning by experimentation.

The goal of this workshop was to provide a forum for researchers active in the area of robot learning and related fields. Due to the limited time available, we attempted to focus discussion around two major issues: 1) How can current learning robot techniques scale to more complex domains, characterized by massive sensor input, complex causal interactions, and long time scales? 2) Where are the new ideas in robot learning coming from? Below, we briefly recount some of the points brought up and discussed by the speakers and participants.

Session 1: Scaling up and the State of the Art

The key to success in scaling up, as proposed by participants in this session, can best be described in a pair of catch phrases: "Be lazy," and "Two (strategies) are better than one." In brief, the former approach (expounded by members of the MIT AI Lab) has the learning robot computing/exploring/committing as little as possible in order to keep large tasks tractable. The latter strategy (expounded by Siemens, CMU and UMass Amherst researchers) has the robot learner mixing several representations or learning algorithms, and allowing the robot to make use of whichever one is most appropriate for a given situation. Fortunately, these approaches are not inherently incompatible with each other, so we were able to avoid undue violence during the course of discussion.

Chris Atkeson of the MIT AI Lab began the session with a keynote talk introducing a number of "essential questions" for robot learning, which we have included here:

1) Which to use: direct or indirect methods?
2) How to choose inputs and outputs?
3) How to guide exploration?
4) How to handle uncertainty?
5) Is learning simply function approximation?
6) Which to use: local or distributed representations?
7) Can we achieve commonality across tasks?

8) Is there life after dynamic programming?
9) When will a robot do more than one task?

Chris then briefly described a philosophy for robot learning systems which should be most robust in the face of more and more complex problem domains. Some points of this philosophy include:

"Learn models, compute actions." Don't compute actions directly. This supports the reminder that "representations $\neq$ algorithms;" the representation is your model, the algorithm specifies how you build it and what you do with it once it's built. Why learn models? It allows "mental simulation," and maintains flexibility in the case of changing goals.

"Be lazy." Never compute/commit to/explore anything you don't have to. A corollary of this rule follows: Never throw data away – you may have to re-compute/re-explore it again later. A corollary that follows this one concerns domain-specific knowledge – if you've got it, use it.

Andrew Moore, also of the MIT AI Lab, briefly described his recent work on robot exploration in continuous state spaces. Andrew's strategy uses the "be lazy" approach on two fronts: representation and computation. For representation, the "parti-game" algorithm breaks its state space into equivalence classes of varying resolutions, grouping together states from which a given action produces indistinguishable results. For computation of actions, Andrew has developed a "prioritized sweeping" algorithm which is both data-efficient and computation-efficient. When new information is added (by the robot making an action and observing the result), the information is incorporated in a prioritized fashion, forcing computation only on the state transitions which are likely to change the most as a result.

Steve Hanson, of Siemens, was the first of the speakers to discuss a "two are better than one" approach to scaling up. He described two navigation strategies used by humans, map building and landmark identification, and suggested that they may be used in concert with each other. He briefly described current work on "RatBot," a mobile robot with which these mixed navigation strategies are being tested.

Satinder Singh, of the University of Massachusetts at Amherst, discussed a different form of mixed strategies: the composition of a high-level strategy out of "primitive" policies. These primitive policies (such as avoiding walls, directly approaching the target, etc.), form a basis for all the actions of a robot, and produce an effective reduction in the "dimensionality" of the problem to the dimensionality of the space of primitives.

Richard Yee, also of the University of Massachusetts at Amherst, discussed another approach to scaling up that incorporated mixed strategies. Richard's work explores a learner's transition from using rote memorization to using abstractions as the amount of available data increases. By maintaining both representations, the learner may exploit whichever is most appropriate for its immediate task.

Session 2: Where are the new ideas coming from?

Every field of science and engineering is driven by an interplay of problems and solutions. In the morning, we concentrated on the problem of scaling up in robot learning and sought solutions. In the afternoon we turned our attention to the question of where, in general, the new solutions in robot learning were coming from. We asked for anything from hard-and-fast results to wild speculation and were well rewarded with both.

We began the afternoon session with a tutorial by **Andrew Barto**, of the University of Massachusetts at Amherst. Reinforcement learning is one of the most pervasive techniques in robot learning, and was a central theme throughout the workshop. In order to give participants an even footing on the subject, Andy introduced the theory behind reinforcement learning, and discussed how that theory has translated into practice. He described how reinforcement learning can be viewed as an incremental real-time stochastic approximation of dynamic programming. He raised questions about the appropriateness of combining functional approximation methods (such as neural networks) with dynamic-programming-based reinforcement learning.

During the panel discussion which followed, participants suggested a number of sources of new ideas in robot learning. We have reproduced this list below, along with some interspersed commentary.

Tom Mitchell (CMU):

1) Shifting from single task to life-long learning. Reusing learned behaviors. This should lead to the evolution of "generally useful" representations (as opposed to representations that are useful only for a specific task).

2) Focusing on architecture-level issues, such as "what is the correct I/O representation?"

Chris Atkeson (MIT):

3) The new ideas are coming from the 50's and 60's. Many people are rediscovering relevant work done years ago in related fields.

4) Hierarchies. Separating levels of a system which do separate things that are complementary. We need a task which motivates this.

5) Integrating planning and control. In the past, with planning problems, the control components have been trivial, and in control problems, the planning components have been trivial.

6) Toy stores. Some of the best new ideas (at least at the MIT AI Lab) are motivated by clever toys.

Dean Pomerleau (CMU):

7) Including handcrafted knowledge. If you have prior information about the domain, build it in.

8) Partitioning knowledge into experts. More and more systems are making use of subsystems which are experts in some specific part of a domain, and allowing the robot to choose among these experts for guidance.

Steve Suddarth (UMD):

9) Convincing the control theorists that neural networks are "okay." This needs to be approached on a number of fronts. In *system identification*, we need to demonstrate that we can get models on the fly using limited model information. In *dynamics*, we need to demonstrate that the emerging rule extraction technology should allow plant characterization. In *optimal control*, we need to demonstrate that neural networks doing reinforcement learning are trying to do optimal control. We need to convince the control theorists that we're "on their team."

10) Looking the other direction as well: we need to address control issues of stability and robustness in neurocontrol research.

From the field:

11) Look at literature on adaptive control, rather than optimal control only. Many issues studied in neurocontrol could benefit.

12) Using information/knowledge transfer. This hearkens back to the morning session, and to points 1 and 7 of the afternoon session. There is a growing trend of building knowledge into a new network, either by incorporating human knowledge or by transferring information from a previously-learned model.

13) Looking at biological systems. This, after all, was the inspiration of the field, and there is no lack of interesting problems, nor of interesting solutions which evolution has found for them.

14) Look at the psychological literature.

The most persistent theme of the afternoon was that of reducing the amount of real-world experimentation needed to learn. Many of the above sources for "new ideas" are being tapped for just those purposes. The approaches being followed can perhaps be divided into two broad categories: a) partitioning the problem into experts domains, and b) making use of as much domain knowledge as possible. The latter approach is being pursued both by the inclusion of "handcrafted" knowledge and the transfer of domain knowledge from previously-trained systems. **Sebastian Thrun** and **Tom Mitchell's** recent work this year demonstrates transfer from action model networks for a robotics task to a Q-learning task. In the workshop, Sebastian discussed solving multiple tasks and making use of synergy between them. He demonstrated this with his approach to learning sensor interpretations for a mobile robot.

Lori Pratt also discussed her work on neural network transfer, which improves backpropagation learning by using weights from networks trained on related tasks. She pointed out that a variety of techniques now exist for both extracting rule from and inserting rules into artifical neural networks, making "information transplants" feasible.

The workshop drew to a close amid lively discussion, as participants broke up into small groups to further pursue their specific interests. It is our belief that this workshop served its intended purpose: clarifying some of the questions that need to be addressed in robot learning, and providing researchers with pointers to how others are attempting to address these questions. We hope to coordinate a similar workshop again next year, in which we will expect to see some of the germinal ideas from this session in full bloom.

Participants

Below we have included the names of workshop participants, along with the names of several people who, although unable to attend the workshop, wished to be added to this "contact list."

C. Atkeson	cga@ai.mit.edu	Yann LeCun	yann@neural.att.com
Andrew Barto	barto@cs.umass.edu	Alexander Linden	Alexander.Linden@gmd.de
Yoshua Bengio	yoshua@neural.att.com	Enno Littmann	littmann@techfak.uni-bielefeld.de
David Bisset	dlb@ukc.ac.uk	Kevin Markey	markey@cs.colorado.edu
Steve Bradtke	bradtke@cs.umass.edu	Ryusuke Masuoka	ryusuke.masuoka@cs.cmu.edu
Neil Burgess	n.burgess@ucl.ac.uk	Tom Mitchell	Tom_Mitchell@cs.cmu.edu
Ah Chung Tsoi	act@s1.elec.uq.oz.au	Jennifer Myers	jmeyers@nwu.edu
Olivier Coenen	olivier@helmholtz.sdsc.edu	Steve Nowlan	nowlan@helmholtz.sdsc.edu
David Cohn	cohn@psyche.mit.edu	John P.H. Steele	jsteele@mines.colorado.edu
Peter Dayan	dayan@helmholtz.sdsc.edu	Gerhard Paass	paass@gmd.edu
Dave DeMers	demers@cs.ucsd.edu	Dean Pomerleau	pomerleau@cs.cmu.edu
Gerald Fahner	gerald@nero.uni-bonn.edu	Larry Pyeatt	pyeatt@texaco.com
Sid Fels	ssfels@ai.toronto.edu	Hamid R. Berenji	berenji@ptolemy.arc.nasa.gov
B. Fritzke	fritzke@icsi.berkeley.edu	Robin R. Murphy	rmurphy@mines.colorado.edu
C. Ganesh	cganesh@mines.colorado.edu	Philip Sabes	sabes@psyche.mit.edu
Zoubin Gharamani	zoubin@psyche.mit.edu	Peter Sandon	sandon@cs.dartmouth.edu
Hiroaki Gomi	gomi@atr-hr.atr.co.jp	Terry Sanger	tds@ai.mit.edu
Aaron Gordon	aaron@ouray.mines.colorado.edu	Stephan Schaal	sschaal@ai.mit.edu
Vijaykumar Gullapalli	vijaykumar@cs.umass.edu	Terry Sejnowski	tsejnowski@ucsd.edu
Eric Hartman	hartman@pav.mcc.com	Jude Shavlik	shavlik@cs.wisc.edu
Ryoichi Hashimoto	ryoichi@psyche.mit.edu	Satinder Singh	singh@cs.umass.edu
Charles Higgins	chuck@micro.caltech.edu	Steve Suddarth	suddarth@afosr.af.mil
Jurgen Hollatz	jh@morla.zfe.siemens.de	Granger Sutton	grange@bengal.tigr.org
John Houde	houde@psyche.mit.edu	Sebastian Thrun	thrun@cs.cmu.edu
Larry Hunter	hunter@nlm.nih.gov	David Tseng	tseng@aic.hrl.hac.com
Michael I. Jordan	jordan@psyche.mit.edu	Yoji Uno	uno@atr-hr.atr.co.jp
Thomas J. Lee	tomlee@starbase.mitre.edu	Andrew W. Moore	awm@ai.mit.edu
Kevin Kittka	kkittka@isi.edu	Ron Williams	rjw@ccs.northeastern.edu
Yasaharu Koike	koike@atr-hr.atr.co.jp	Daniel Wolpert	wolpert@psyche.mit.edu
Anya L. Simons	VY0112@bingvaxa.bitnet	Lorien Y. Pratt	lpratt@mines.colorado.edu
Ken Lang	akl@cs.cmu.edu	Richard Yee	yee@cs.umass.edu