Final Report: Hybrid Computational Models for Skill Acquisition

Thomas G. Dietterich
Prasad Tadepalli

Department of Computer Science
Oregon State University
Corvallis, Oregon 97331

Office of Naval Research
800 North Quincy Street, Ballston Tower One
Arlington, VA 22217-5660

The goal of this research was to develop a hybrid real-time problem-solving architecture that couples symbolic planning methods with connectionist reinforcement learning methods. The advantage of this hybrid architecture is that it can immediately achieve reasonable performance, because the symbolic planning system can quickly develop an acceptable control policy, but it can also gradually achieve optimal real-time performance, because the reinforcement learning system will eventually converge on a near-optimal policy. Many DoD problems would benefit from the ability to perform near-optimal real-time control of complex systems.

real-time problem solving, machine learning, reinforcement learning, planning
Final Report:
Hybrid Computational Methods for Skill Acquisition
ONR Grant Number N00014-95-1-0557

Thomas G. Dietterich
Prasad Tadepalli

Department of Computer Science
303 Dearborn Hall
Oregon State University
Corvallis, OR 97331

1 Administrative Information

Title: Hybrid Computational Methods for Skill Acquisition
Grant Number: N00014-95-1-0557
Grantee: Oregon State University, Corvallis, OR 97331
Start Date: March 1, 1995
End Date: February 28, 1998
Principal Investigators: Thomas G. Dietterich
Department of Computer Science
303 Dearborn Hall
Oregon State University
Corvallis, OR 97331
503-737-5559 (voice) 503-737-3014 (fax) tgd@cs.orst.edu (email)

Prasad Tadepalli
Department of Computer Science
303 Dearborn Hall
Oregon State University
Corvallis, OR 97331
503-737-5552 (voice) 503-737-3014 (fax) tadepalli@cs.orst.edu (email)
August 28, 1998

Dr. Michael Shneier
Office of Naval Research
800 North Quincy Street
Ballston Tower One
Arlington, VA 22217-5660

Dear Michael Shneier:

Please find enclosed two copies of our final report for N00014-95-1-0557. I have also sent one copy to the Defense Technical Information Center. If there are other people who should receive copies, please let me know. This report contains the same information that we sent in earlier via electronic mail.

All publications supported by this grant are available for downloading from our web pages.

Thank-you very much for your support of this research. It has been a very successful project.

Sincerely,

Thomas G. Dietterich
Professor

Telephone: (541) 737-5559
Internet: tgd@research.cs.orst.edu
2 Project Summary

The goal of this research was to develop a hybrid real-time problem-solving architecture that couples symbolic planning methods with connectionist reinforcement learning methods. The advantage of this hybrid architecture is that it can immediately achieve reasonable performance, because the symbolic planning system can quickly develop an acceptable control policy, but it can also gradually achieve optimal real-time performance, because the reinforcement learning system will eventually converge on a near-optimal policy. Many DoD problems would benefit from the ability to perform near-optimal real-time control of complex systems.

3 Accomplishments

- Developed the ALERT hybrid architecture which combines symbolic (DRULE) planner with hierarchical reinforcement learning
- Showed experimentally that the DRULE planner could achieve human-level performance on the Kanfer-Ackerman air traffic control (ATC) task.
- Developed two learning algorithms for DRULES: one based on random examples and queries, and the other based on exercises.
- Showed experimentally that both learning algorithms could achieve intermediate performance on the ATC task.
- Proved that both learning algorithms are correct and computationally feasible. This involved proving a new result on learning of Horn clause logic programs.
- Developed a new, hierarchical method for reinforcement learning, the MAXQ method.
- Proved that MAXQ can represent any hierarchical policy.
- Developed the MAXQ-Q learning algorithm for hierarchical reinforcement learning.
- Proved that MAXQ-Q converges to a recursively optimal policy asymptotically.
- Demonstrated experimentally that MAXQ-Q attains optimal performance on a simplified task that shares many properties with the ATC task.

4 Transitions

We are currently working with i2 Technologies (Dallas, Texas) to apply our reinforcement learning methods to supply chain scheduling and optimization.

5 ONR-Funded Publications for entire grant period


6 Online Information Available

Postscript files for all papers are available via WWW from the following URL's:

http://www.cs.orst.edu/~tgd/
http://www.cs.orst.edu/~tadepall/