"Abstractions and Hierarchies for Learning and Planning"

Maintained by: Lihong Li
Last updated: 2005-08-22

A. Learning and Planning with Hierarchies/Abstractions:

1. Thomas L. Dean and Shieu-Hong Lin. Decomposition techniques for planning in stochastic domains. In Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95), pages 1121–1129, 1995.

2. Nicolas Meuleau, Milos Hauskrecht, KeeEung Kim, Leonid Peshkin, Leslie Pack Kaelbling, Thomas L. Dean, and Craig Boutilier. Solving very large weakly coupled Markov decision processes. In Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98), pages 165–172, 1998.

3. Milos Hauskrecht, Nicolas Meuleau, Craig Boutilier, Leslie Pack Kaelbling, and Tom Dean. Hierarchical solution of Markov decision processes using macro-actions. In Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-98), pages 220–229, 1998.

4. Thomas G. Dietterich. Hierarchical reinforcement learning with the MAXQ value function decomposition. Journal of Artificial Intelligence Research, 13:227–303, 2000.

5. Ronald Parr and Stuart J. Russell. Reinforcement learning with hierarchies of machines. In Advances of Neural Information Processing Systems 10 (NIPS-97), pages 1043–1049, 1998.

6. Doina Precup and Richard S. Sutton. Multi-time models for temporally abstract planning. In Advances of Neural Information Processing Systems 10 (NIPS-97), pages 1050–1056, 1998.

7. Richard S. Sutton, Doina Precup, and Satinder P. Singh. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. Artificial Intelligence, 112:181–211, 1999.

8. Satinder P. Singh. Transfer of learning by composing solutions of elemental sequential tasks. Machine Learning, 8:323–339, 1992.

9. Satinder P. Singh. Reinforcement learning with a hierarchy of abstract models. In Proceedings of the Tenth National Conference on Artificial Intelligence (AAAI-92), pages 202–207, 1992.

10. Satinder P. Singh. Scaling reinforcement learning algorithms by learning variable temporal resolution models. In Proceedings of the Ninth International Conference on Machine Learning (ICML-92), pages 406–415, 1992.

11. Satinder P. Singh and David Cohn. How to dynamically merge Markov decision processes. In Advances in Neural Information Processing Systems 10 (NIPS-97), pages 1057–1063, 1998.

12. Andrew Barto and Sridhar Mahadevan. Recent advances in hierarchical

reinforcement learning. Discrete Event Dynamic Systems: Theory and Applications, 13(1–2):41–77, 2003.

13. Peter Dayan and Geoffrey E. Hinton. Feudal reinforcement learning. In Advances in Neural Information Processing Systems 5 (NIPS-92), pages 271–278, 1993.

14. Richard S. Sutton. TD models: Modeling the world at a mixture of time scales. In Proceedings of the Twelfth International Conference on Machine Learning (ICML-95), pages 531–539, 1995.

15. David Andre and Stuart J. Russell. Programmable reinforcement learning agents. In Advances in Neural Information Processing Systems 13 (NIPS-00), pages 1019–1025, 2001.

16. Shieu-Hong Lin. Exploiting Structure for Planning and Control. PhD thesis, Brown University, Providence, RI, 1997.

17. Thomas L. Dean, Robert Givan, and Sonia M. Leach. Model reduction techniques for computing approximately optimal solutions for Markov decision processes. In Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI-97), pages 124–131, 1997.

18. Craig Boutilier and Richard Dearden. Approximating value trees in structured dynamic programming. In Proceedings of the Thirteenth International Conference on Machine Learning (ICML-96), pages 54–62, 1996.

19. Thomas G. Dietterich. The MAXQ method for hierarchical reinforcement learning. In Proceedings of the Fifteenth International Conference on Machine Learning (ICML-98), pages 118–126, 1998.

20. Marco Wiering and Jurgen Schmidhuber. HQ-learning. Adaptive Behavior, 6(2):219–246, 1997.

21. Georgios Theocharous and Sridhar Mahadevan: Approximate planning with hierarchical partially observable Markov decision process for robot navigation. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), 2002.

22. Georgios Theocharous: Hierarchical Learning and Planning in Partially Observable Markov Decision Processes. Ph.D. Thesis, Michigan State University, 2002.

23. David Andre, Stuart J. Russell: State Abstraction for Programmable Reinforcement Learning Agents. AAAI/IAAI 2002: 119-125.

24. Fausto Giunchiglia and Toby Walsh: A Theory of Abstraction. Artificial Intelligence, 57(2-3): 323-389, 1992.

25. Ravindran, B. and Barto, A. G. (2003) "Relativized Options: Choosing the Right Transformation". In the Proceedings of the Twentieth International Conference on Machine Learning(ICML 2003), pp. 608-615.

26. Ravindran, B. (2004) "An Algebraic Approach to Abstraction in Reinforcement Learning". Doctoral Dissertation, Department of Computer Science, University of Massachusetts, Amherst MA.

27. Ravindran, B. and Barto, A. G. (2003) "SMDP Homomorphisms: An Algebraic Approach to Abstraction in Semi Markov Decision Processes". In the Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI 03), pp. 1011-1016.

28. Craig Boutilier and Richard Dearden: Using Abstractions for Decision-Theoretic Planning with Time Constraints. AAAI'94.

29. Carlos Guestrin and Geoffrey Gordon: Distributed Planning in Hierarchical Factored MDPs. UAI'02.

30. Leslie Pack Kaelbling: Hierarchical Learning in Stochastic Domains: Preliminary Results. ICML'93.

31. H.S. Chang, P.J. Fard, S.I. Marcus, and M. Shayman. Multitime scale Markov decision processes. IEEE Transactions on Automatic Control, 48(6):976-987, 2003.

B. Automated Hierarchy/Abstraction Discovery

1. Amy McGovern and Andrew G. Barto. Automatic discovery of subgoals in reinforcement learning using diverse density. In Proceedings of the Eighteenth International Conference on Machine Learning (ICML-01), pages 361–368, 2001.

2. Bernhard Hengst. Discovering hierarchy in reinforcement learning with HEXQ. In Proceedings of the Nineteenth International Conference on Machine Learning (ICML-02), pages 243–250, 2002.

3. Sebastian Thrun and Anton Schwartz. Finding structure in reinforcement learning. In Advances in Neural Information Processing Systems 7 (NIPS-94), pages 385–392, 1995.

4. Craig A. Knoblock. Automatically Generating Abstractions for Problem Solving. PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 1991.

5. Andrew W. Moore, Leemon C. Baird, and Leslie Pack Kaelbling. Multi-value-functions: Efficient automatic action hierarchies for multiple goal MDPs. In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI-99), pages 1316–1323, 1999.

6. Amy McGovern. Autonomous Discovery of Temporal Abstractions from Interaction with an Environment. Ph.D. Thesis, University of Massachusetts, MA, 2002.

7. Nicholas K. Jong and Peter Stone. State Abstraction Discovery from Irrelevant State Variables. In Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI-05), August 2005.

8. Ozgür Simsek, Andrew G. Barto: Using relative novelty to identify useful temporal abstractions in reinforcement learning. ICML 2004.

9. Balaraman Ravindran, Andrew G. Barto: SMDP Homomorphisms: An algebraic approach to abstraction in semi-Markov decision processes. IJCAI 2003: 1011-1018.

10. Jonsson, A., and Barto, A. G. 2001. Automated state abstraction for options using the U-tree algorithm. In Advances in Neural Information Processing Systems: Proceedings of the 2000 Conference, Cambridge, MA: MIT Press, pp. 1054-1060.

11. Shie Mannor, Ishai Menache, Amit Hoze, Uri Klein: Dynamic abstraction in reinforcement learning via clustering. ICML 2004.

12. Bram Bakker, Jurgen Schmidhuber: Hierarchical Reinforcement Learning Based on Subgoal Discovery. NCI 2004.

13. Amy McGovern: Autonomous Discovery of Abstractions through Interaction with an Environment. SARA 2002: 338-339.

14. Ozgür Simsek, Alicia P. Wolfe, and Andrew G. Barto: Identifying useful subgoals in reinforcement learning by local graph partitioning. In ICML 2005.

15. Ishai Menache, Shie Mannor, Nahum Shimkin: Q-Cut - Dynamic Discovery of Sub-goals in Reinforcement Learning. ECML 2002: 295-306.

16. Marc Pickett, Andrew G. Barto: PolicyBlocks: An Algorithm for Creating Useful Macro-Actions in Reinforcement Learning. ICML 2002: 506-513.

17. Glenn A. Iba: A heuristic approach to the discovery of macro-operators. Machine Learning, 3, 285-317.

18. Bruce L. Digney: Learning hierarchical control structure for multiple tasks and changing environments. In From Animals to Animats 5: The Fifth Conference on the Simulation of Adaptive Behavior, 1998.

19. Martin Stolle and Doina Precup. Learning options in reinforcement learning. In Proceedings of the Fifth International Symposium on Abstraction, Reformulation and Approximation (SARA-02), volume 2371 of Lecture Notes in Computer Science, pages 212–223, 2002.

20. Vadim Bulitko, Nathan Sturtevant and Maryia Kazakevich. Speeding Up Learning in Real-time Search via Automatic State Abstraction. AAAI 2005.

21. Anders Jonsson and Andrew Barto. A Causal Approach to Hierarchical Decomposition of Factored MDPs. ICML 2005.

22. Craig A. Knoblock. Automatically generating abstractions for planning. Artificial Intelligence, 68(2), 1994.

23. Satinder Singh, Andrew G. Barto and Nuttapong Chentanez: Intrinsically

Motivated Reinforcement Learning, NIPS'04.

24. Norm Ferns, Prakash Panangaden, and Doina Precup: Metrics for Finite Markov Decision Processes. UAI, 2004.

C. State Abstraction and Aggregation (some have beeb listed above)

(for Tom: 3, 4, 8, 11, 12, 13, 14, 15, 16, 18)

1. David Andre, Stuart J. Russell: State Abstraction for Programmable Reinforcement Learning Agents. AAAI/IAAI 2002: 119-125.

2. Fausto Giunchiglia and Toby Walsh: A Theory of Abstraction. Artificial Intelligence, 57(2-3): 323-389, 1992.

3. B. Ravindran and A. G. Barto. SMDP Homomorphisms: An Algebraic Approach to Abstraction in Semi Markov Decision Processes. In the Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence (IJCAI'03), pp. 1011-1016.

4. Craig Boutilier and Richard Dearden: Using Abstractions for Decision-Theoretic Planning with Time Constraints. AAAI'94.

5. Bernhard Hengst. Discovering hierarchy in reinforcement learning with HEXQ. In Proceedings of the Nineteenth International Conference on Machine Learning (ICML-02), pages 243–250, 2002.

6. Craig A. Knoblock. Automatically Generating Abstractions for Problem Solving. PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 1991.

7. Craig A. Knoblock. Automatically generating abstractions for planning. Artificial Intelligence, 68(2), 1994.

8. Nicholas K. Jong and Peter Stone. State Abstraction Discovery from Irrelevant State Variables. In Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI-05), August 2005.

9. Jonsson, A., and Barto, A. G. 2001. Automated state abstraction for options using the U-tree algorithm. In Advances in Neural Information Processing Systems: Proceedings of the 2000 Conference, Cambridge, MA: MIT Press, pp. 1054-1060.

10. Vadim Bulitko, Nathan Sturtevant and Maryia Kazakevich. Speeding up learning in real-time search via automatic state abstraction. AAAI 2005.

11. Eyal Evan-Dar and Yishay Mansour: Approximate equivalence of Markov decision processes. COLT 2003.

12. Robert Givan, Thomas Dean, and Matthew Greig: Equivalence notions and model minimization in Markov decision processes. Artificial Intelligence, 147:163-223, 2003.

13. Thomas Dean, Robert Givan, and Sonia Leach: Model reduction techniques for computing approximately optimal solutions for Markov decision processes. UAI 1997.

14. Thomas Dean and Robert Givan: Model minimization in Markov decision processes, AAAI 1997.

15. Dimitri P. Bertsekas and David A. Castanon. Adaptive aggregation methods for infinite horizon
dynamic programming. IEEE Transactions on Automatic Control, 34(6):589–598, 1989.

16. James C. Bean, John R. Birge, and Robert L. Smith. Aggregation in dynamic programming. Operations Research, 35(2), 1989.

17. Theodore J. Lambert III, Marina A. Epelman, and Robert L. Smith. Aggregation in Stochastic Dynamic programming. Techical report, University of Michigan, 2004.

18. Benjamin van Roy. Performance loss bounds for approximate value iteration with state aggregation. To appear in Mathematics of Operations Research. A short version is published in NIPS 2005.

19. A. A. Al-Falou and J. H. van Schuppen. Aggregation in hierarchical discrete-event systems. Discrete Event Dynamic Systems: Theory and Applications, vol 13, 2003.

20. Satinder P. Singh, Tommi Jaakkola, and Michael I. Jordan: Reinforcement learning with soft state aggregation. NIPS 1994: 361-368.

21. D.F. Rogers, R.D. Plante, R.T. Wong, and J.R. Evans. Aggregation and disaggregation techniques and methodology in optimization. Operations Research, 38:553-582, 1991.

22. Xi-Ren Cao, Zhiyuan Ren, Shalabh Bhatnagar, Michael Fu, and Steven Marcus. A time aggregation approach to Markov decision processes. Automatica, 28:929-943, 2002.

23. H.S. Chang, P.J. Fard, S.I. Marcus, and M. Shayman. Multitime scale Markov decision processes. IEEE Transactions on Automatic Control, 48(6), 2003.

D. Related Techniques: Factorization, Shaping, Inverse RL, Least-Squares, Off-policy learning etc.

1. Michael J. Kearns and Daphne Koller. Efficient reinforcement learning in factored MDPs. In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI-99), pages 740–747, 1999.

2. Daphne Koller and Ronald Parr. Computing factored value functions for policies in structured MDPs. In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI-99), pages 1332–1339, 1999.

3. Lawrence Saul and Michael I. Jordan. Mixed memory Markov models: Decomposing complex stochastic processes as mixtures of simpler ones. Machine Learning, 37(1):75–87, 1999.

4. Andrew Y. Ng, Daishi Harada, and Stuart J. Russell. Policy invariance under reward transformations: Theory and application to reward shaping. In Proceedings of the Sixteenth International Conference on Machine Learning (ICML-99), pages 278–287, 1999.

5. Andrew Y. Ng and Stuart J. Russell. Algorithms for inverse reinforcement learning. In Proceedings of the Seventeenth International Conference on Machine Learning (ICML-00), pages 663–670, Stanford, California, 2000. Morgan Kaufmann.

6. Justin A. Boyan. Least-squares temporal difference learning. Machine Learning, 49(2):233–246, November 2002.

7. Michail G. Lagoudakis and Ronald Parr. Model-free least squares policy iteration. In Advances in Neural Information Processing Systems 14 (NIPS-01), pages 1547–1554, 2002.

8. Doina Precup, Richard S. Sutton, and Satinder P. Singh. Eligibility traces for off-policy policy evaluation. In Proceedings of the Seventeenth Interational Conference on Machine Learning (ICML-00), pages 759–766, 2000.

9. Doina Precup, Richard S. Sutton, and Sanjoy Dasgupta. Off-policy temporal-difference learning with funtion approximation. In Proceedings of the Eighteenth Conference on Machine Learning (ICML-01), pages 417–424, 2001.

10. Doina Precup and Richard S. Sutton: Exponentiated gradient methods for reinforcement learning. Proceedings of the 14th International Conference on Machine Learning (ICML'97), pp. 272-277.

11. Sridhar Mahadevan, "Samuel Meets Amarel: Automating Value Function Approximation using Global State Space Analysis" , Proceedings of the National Conference on Artificial Intelligence (AAAI-2005), Pittsburgh, PA, 2005.

12. Craig Boutilier, Richard Dearden, and Moises Goldszmidt: Exploiting Structure in Policy Construction. IJCAI'95.

13. Robert Givan, Sonia Leach, and Thomas Dean: Bounded-parameter Markov decision processes. Artificial Intelligence, 122:71-109, 2000.

14. Carlos Guestrin, Daphne Koller, Ronald Parr, and Shobha Venkataraman. Efficient solution algorithms for factored MDPs. Journal of Artificial Intelligence Research, volume 19, pages 399-468, 2003.