Reinforcement Learning Algorithms in Markov Decision Processes
AAAI-10 Tutorial

Part IV: Take home message

Csaba Szepesvári   Richard S. Sutton

University of Alberta
E-mails: {**szepesva,rsutton**}**@.ualberta.ca**

Atlanta, July 11, 2010

Outline

Main message

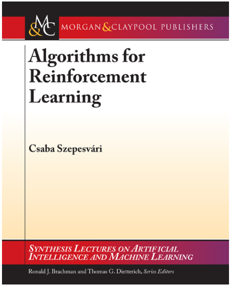
Reinforcement learning, building on a simple, yet powerful theory, provides effective solutions to many AI problems.

Concepts

- Markov decision processes
  - Generalizes shortest path computations
  - Stochasticity, state, action, reward, value functions, policies
  - Bellman (optimality) equations, operators, fixed-points
  - Value iteration, policy iteration
- Value prediction
  - Temporal difference learning unifies Monte-Carlo and bootstrapping
  - Function approximation to deal with large spaces
  - New gradient based methods
  - Least-squares methods
- Control
  - Closed-loop interactive learning: exploration vs. exploitation
  - Q-learning
  - SARSA
  - Policy gradient, natural actor-critic

Literature – books

- Kaelbling et al. – review
- Bertsekas and Tsitsiklis (1996)
- Sutton and Barto (1998)
- Bertsekas (2007a,b)
- Bertsekas (2010) – 160 pages!
- Gosavi (2003)
- Cao (2007) – policy gradient methods
- Powell (2007) – operations research perspective
- Chang et al. (2008) f– adaptive sampling (i.e., simulation-based performance optimization)
- Busoniu et al. (2010) – function approximation
- Szepesvári (2010) – concise, algorithms, ideas (the latest, ...)



Literature – II

- Conferences
  - ICML
  - NIPS
  - UAI, AAAI, IJCAI, COLT, ALT, ..
- Journals
  - MLJ
  - JMLR
  - IEEE TAC
  - MOR
  - NN, Neurocomputing, IEEE TNN

Software

- RL-GLUE: http://glue.rl-community.org
- RL-LIBRARY: http://library.rl-community.org
- CLSquare – http://www.ni.uos.de/index.php?id=70
- PIQLE – http://piqle.sourceforge.net/
- RL Toolbox – http://www.igi.tugraz.at/ril-toolbox/
- JRLF – http://mykel.kochenderfer.com/?page_id=19
- LibPG – http://code.google.com/p/libpgrl/

.. and beyond

- What if the state is not observable?
- Abstractions: time!?
- Knowledge representation (and value functions)
- Automated basis construction, regularization, ...
- Beyond the probabilistic framework

For Further Reading

Bertsekas, D. P. (2007a). *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, Belmont, MA, 3 edition.

Bertsekas, D. P. (2007b). *Dynamic Programming and Optimal Control*, volume 2. Athena Scientific, Belmont, MA, 3 edition.

Bertsekas, D. P. (2010). Approximate dynamic programming (online chapter). In *Dynamic Programming and Optimal Control*, volume 2, chapter 6. Athena Scientific, Belmont, MA, 3 edition.

Bertsekas, D. P. and Tsitsiklis, J. N. (1996). *Neuro-Dynamic Programming*. Athena Scientific, Belmont, MA.

Busoniu, L., Babuska, R., Schutter, B., and Ernst, D. (2010). *Reinforcement Learning and Dynamic Programming Using Function Approximators*. Automation and Control Engineering Series. CRC Press.

Cao, X. R. (2007). *Stochastic Learning and Optimization: A Sensitivity-Based Approach*. Springer, New York.

Chang, H. S., Fu, M. C., Hu, J., and Marcus, S. I. (2008). *Simulation-based Algorithms for Markov Decision Processes*. Springer Verlag.

Gosavi, A. (2003). *Simulation-based optimization: parametric optimization techniques and reinforcement learning*. Springer Netherlands.

Kaelbling, L., Littman, M., and Moore, A. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285.

Powell, W. B. (2007). *Approximate Dynamic Programming: Solving the curses of dimensionality*. John Wiley and Sons, New York.

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Bradford Book. MIT Press.

Szepesvári, C. (2010). *Reinforcement Learning*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers.