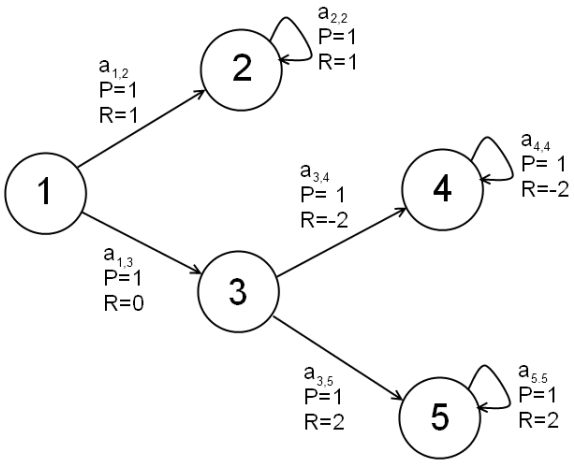
Multi-Agent Learning Systems
Assignment 2: Dynamic Programming

Deadline: 04 November 2010



1 Policy Iteration

Implement the Policy Iteration algorithm[1] in the above model with $\theta = 0.1$, $\gamma = 0.9$ and initial $V(s) = 0$. Start with a policy that in all states selects each action with an equal probability. Do the following:

1. At each iteration show how $V(s)$ changes in the Policy Evaluation step for all states.
2. At each iteration show the improved policy for each state.
3. Discuss the effect of setting the discount factor γ to 1.

Discuss your results!

2 Value Iteration

Implement the Value Iteration algorithm[2] in the same model. At each iteration show the updated policy as well as the updates performed to the values of the states.

Discuss your results!

[1] The pseudo-code can be found at http://webdocs.cs.ualberta.ca/~sutton/book/ebook/node43.html
[2] The pseudo-code can be found at http://webdocs.cs.ualberta.ca/~sutton/book/ebook/node44.html