

Lecture Graphical Models

https://ml01.zrz.tu-berlin.de/wiki/Main/SS09_GraphicalModels

Machine Learning Group, TU Berlin

Instructors: Dr. Ulf Brefeld, Dr. Marc Toussaint

Tutor: Tobias Lang, lang@cs.tu-berlin.de

Sheet 11

Due: 14 July 2009

1. In the lecture, we introduced the expected discounted future return of a policy π

$$V^\pi = \mathbb{E} \left\{ \sum_{t=0}^{\infty} \gamma^t r_t; \pi \right\}.$$

We define the *value function* as the same quantity, conditioned on a certain start state $s_0 = s$,

$$V^\pi(s) = \mathbb{E} \left\{ \sum_{t=0}^{\infty} \gamma^t r_t | s_0 = s; \pi \right\}$$

That is, $V^\pi(s)$ is the discounted future return that we can expect when following policy π starting in state s .

Prove the following recursive property of the value function:

$$V^\pi(s) = \mathbb{E} \{ r_0 | s_0 = s \} + \gamma \sum_{a, s'} \pi_{as} P(s' | a, s) V^\pi(s')$$