

Temporal Difference (TD) Updating without a Learning Rate



Marcus Hutter
www.hutter1.net



Poster T7



Shane Legg
www.vetta.org



- Reinforcement learning TD update:
$$V_s^{t+1} = V_s^t + \tilde{\beta}_t(s, s_{t+1})(r_t + \gamma V_{s_{t+1}}^t - V_{s_t}^t)$$
- We derived learning rate $\tilde{\beta}_t$ for TD with eligibility traces from statistical principles.
- In every setting that we have tested,
superior performance & fewer parameters to tune

