



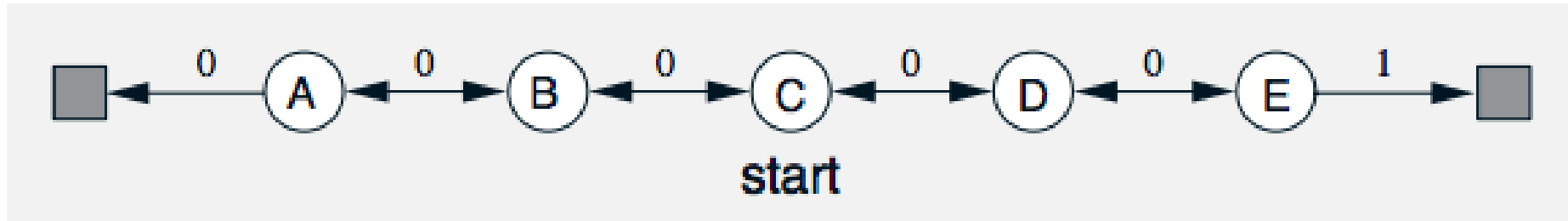
Reinforcement Learning

HW1. Random Walk

Department of Software Convergence

Hyoseok Hwang

Random Walk Example



- 5 states {A, B, C, D, E}
- Init State: C
- Uniform random transition (50% left, 50% right)
- Initial values are $v(s) = 0.0$, for all s
- Reward: 1 for tight terminal, 0 for others
- Discount factor: 1

Random Walk Example



- TODO
 - MC, Incremental MC, TD 구현하여 Value Table(function) 작성
 - `def run_td_episode(v_func, s0=STARTING_STATE, alpha=None, gamma=1.0):`
 - `def run_mc_episode(v_func,num_episodes, s0=STARTING_STATE, gamma=1.0):`
 - `def run_incre_mc_episode(v_func,num_episodes, s0=STARTING_STATE, gamma=1.0,alpha=None):`

Random Walk Example



- Total: 100
 - 파일 제출: 10
 - TD Error < 0.01 - 30
 - MC Error < 0.01 - 30
 - Incremental MC Error < 0.01 - 30
- 다른 방식으로 계산: -20 (예: MC인데 TD로 계산)
- 허용 함수 외 수정: 1글자당 -10
- DELAY 감점: 1분당 1점
- Cheating: -300점

Random Walk Example



- Episode는 10만 번 이상
- Discount factor: 1.0
- 다음 함수 이외에는 절대 수정하지 말 것
 - `def run_td_episode(v_func, s0=STARTING_STATE, alpha=None, gamma=1.0):`
 - `def run_mc_episode(v_func,num_episodes, s0=STARTING_STATE, gamma=1.0):`
 - `def run_incre_mc_episode(v_func,num_episodes, s0=STARTING_STATE, gamma=1.0,alpha=None):`

Random Walk Example



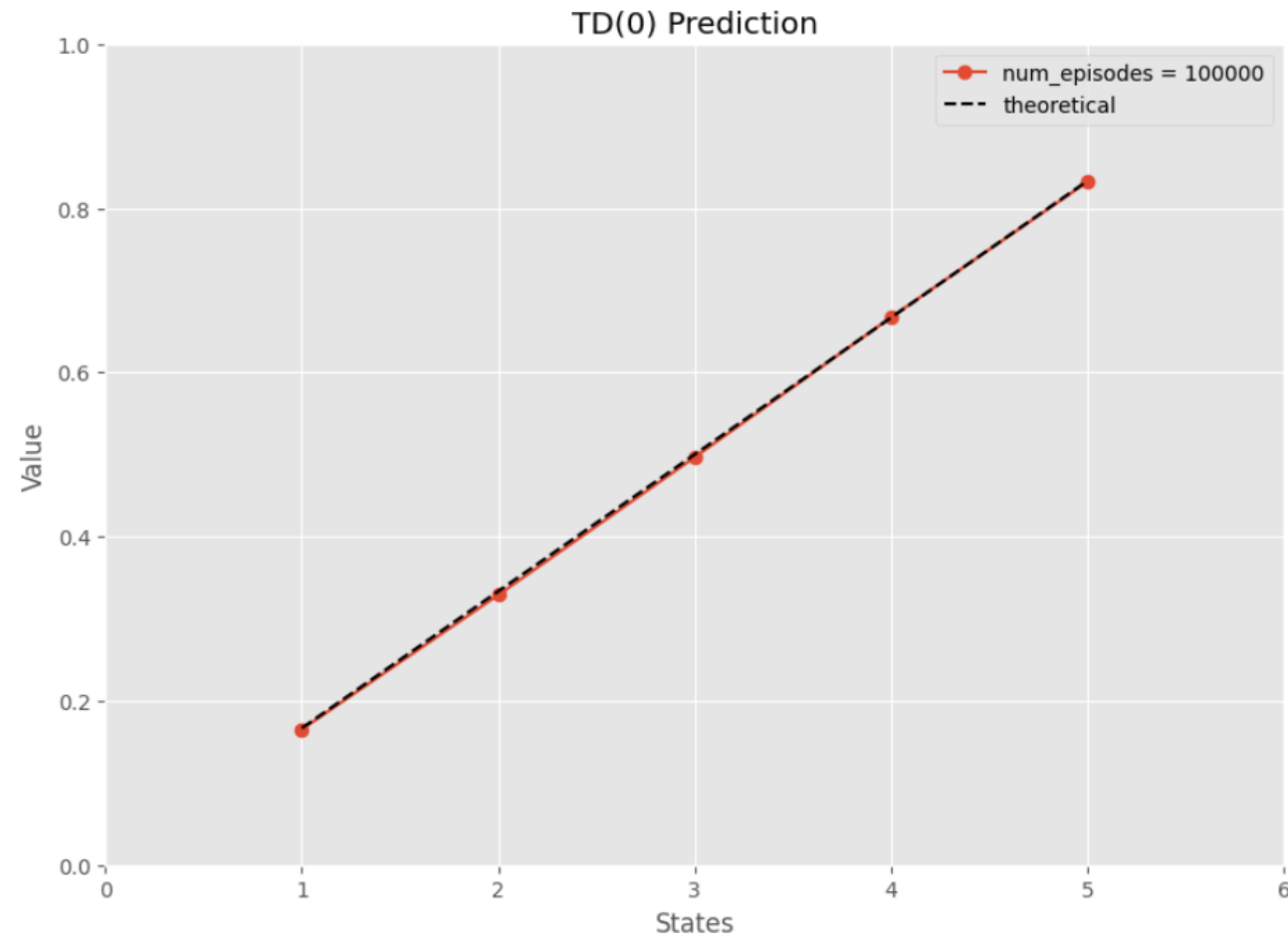
- Episode는 10만 번 이상
- Discount factor: 1.0
- 제출 방식: 과제 및 평가 > HW1
 - **ipynb, py 파일 제출**
- 제출 기한: ~4월 14일 오후 11시 59분 59초
- DELAY 감점: 1분당 1점
- Cheating: -300점
- 문의: TA (김동욱, g9896@khu.ac.kr)

참조



- Result example

working on 100000
Error: 7.1183949354947285e-06
(0.0, 1.0)

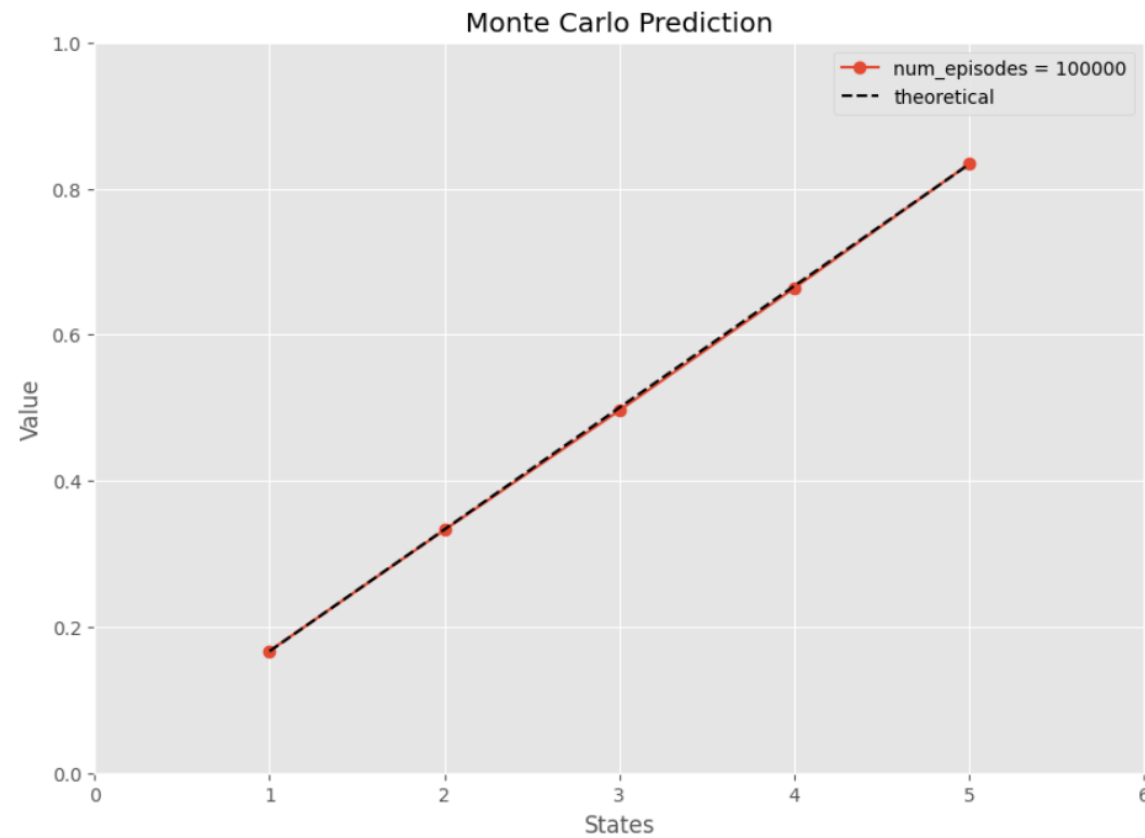


참조



- Result example

```
working on 100000  
Error: 4.7291828259822976e-06  
<ipython-input-14-9d31ea3c15e0>:32: RuntimeWarning: invalid value encountered in true_divide  
  v_func=s_v/(n_v)  
(0.0, 1.0)
```

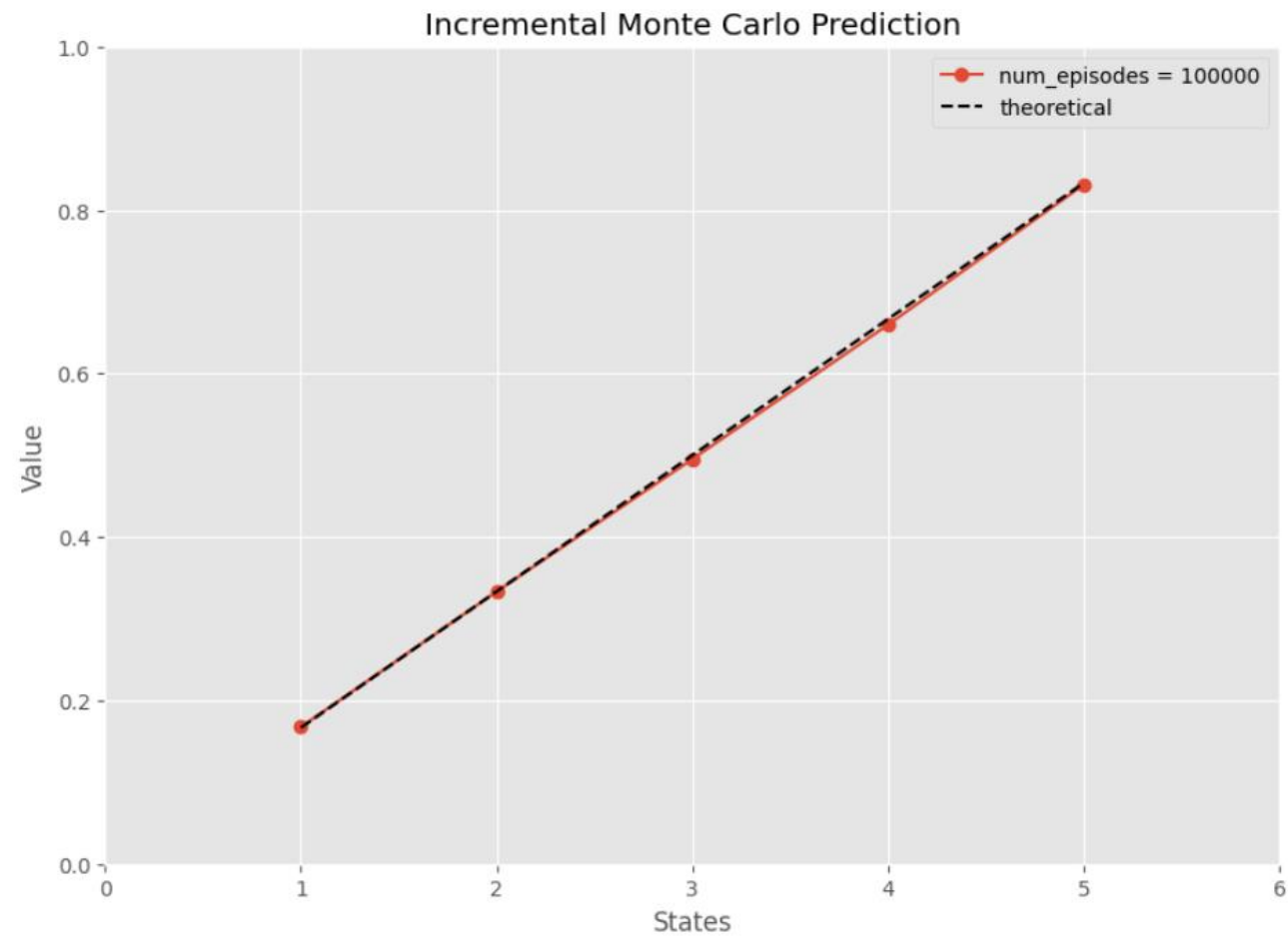


참조



- Result example

```
working on 100000  
Error: 1.498161123674598e-05  
(0.0, 1.0)
```





End of class

Q&A