Evaluating the Effect of Aggregation on Value Function Approximation

COMP 767 – Reinforcement Learning

Mathieu Nassif

McGill University

March 10, 2017

Value Function Approximation

Approximate $v(s)$ using a parameter θ and the function $\hat{v}(s, \theta)$.

Aggregation: We aggregate states into groups and compute an aggregated value for each group.

Effect of Groups

We want to evaluate the effect of the aggregation stategy on

- Error in the limit
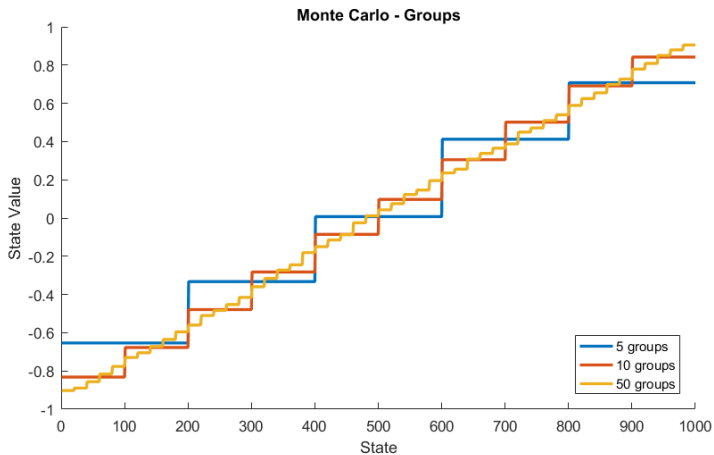- Speed of convergence
- Variance

Description of the MDP

- 1000-state random walk
- 2 actions: left or right
- Size of jump chosen randomly from 1 to 100 uniformly
- +1 for reaching right end, -1 for reaching left end, 0 otherwise
- No discount ($\gamma = 1$)

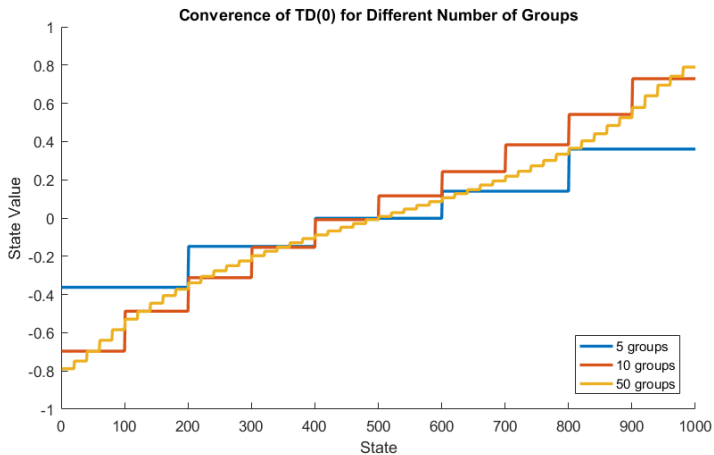- MC and TD(0) algorithms
- For all experiments, $\alpha = 0.0004$

Errors I

- 100 000 episodes
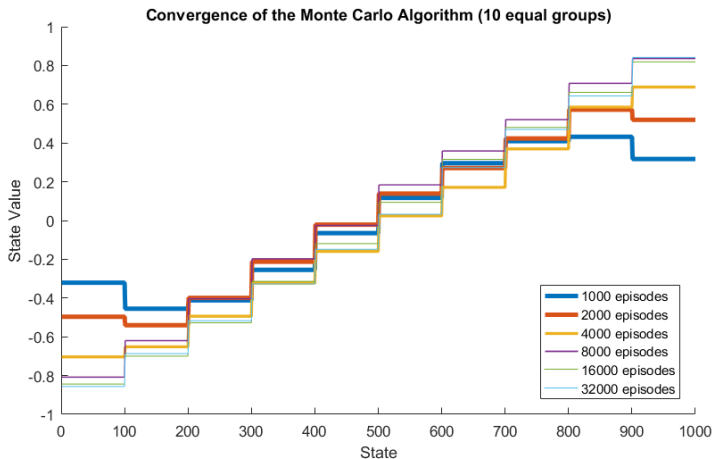- 5, 10, and 50 groups
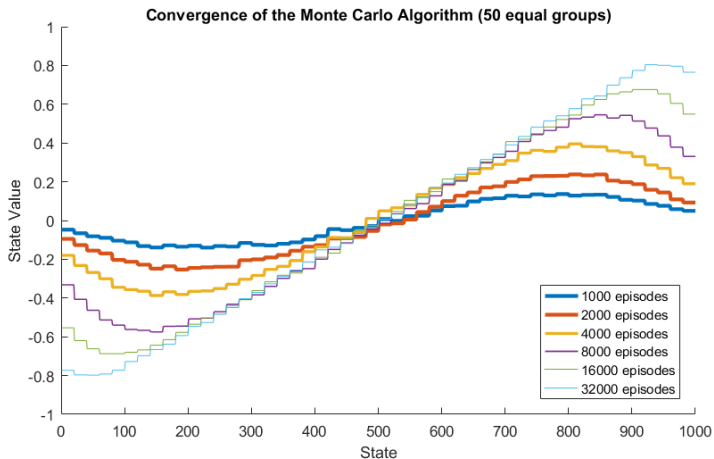
Errors II

Errors III

Speed of Convergence I

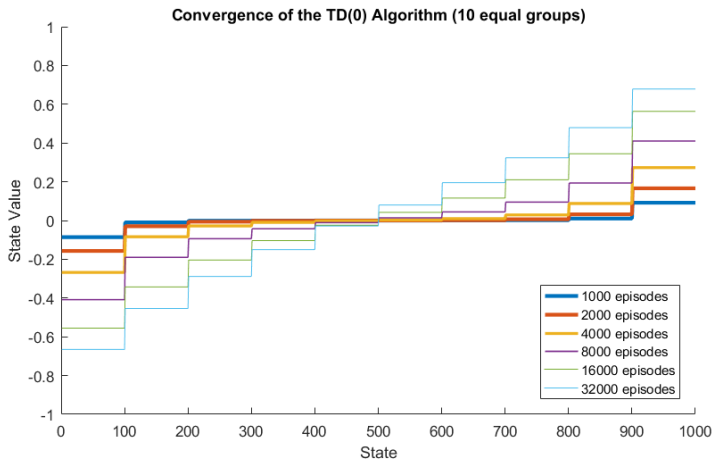- 32000 episodes (shown after first 1000, 2000, 4000, 8000, 16000, 32000)
- 10 and 50 groups
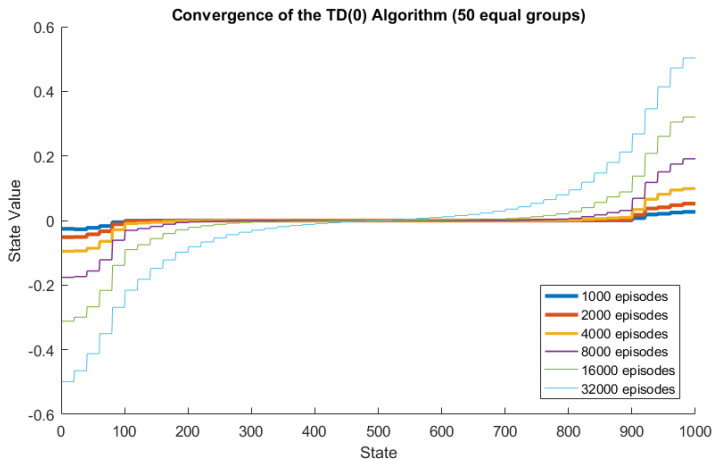
Speed of Convergence II
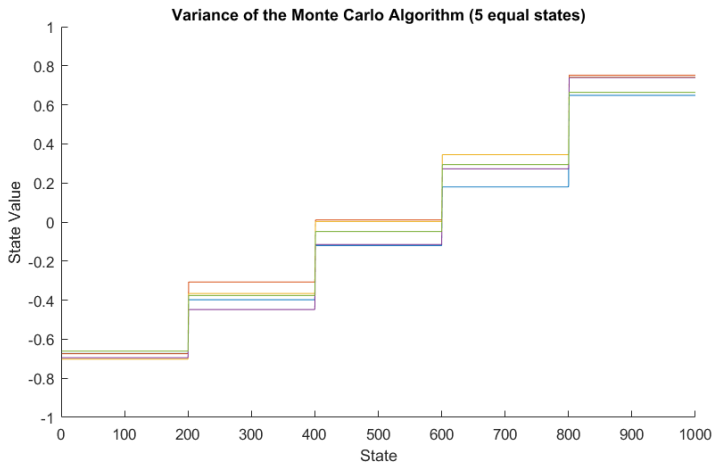
Speed of Convergence III

Speed of Convergence IV

Convergence of the TD(0) Algorithm (10 equal groups)

Speed of Convergence V

Convergence of the TD(0) Algorithm (50 equal groups)

Variance I

- 5000 episodes
- 5 and 10 groups
- Repeated 5 times each

Variance II

Variance III

Variance of Monte Carlo Algorithm (10 equal groups)
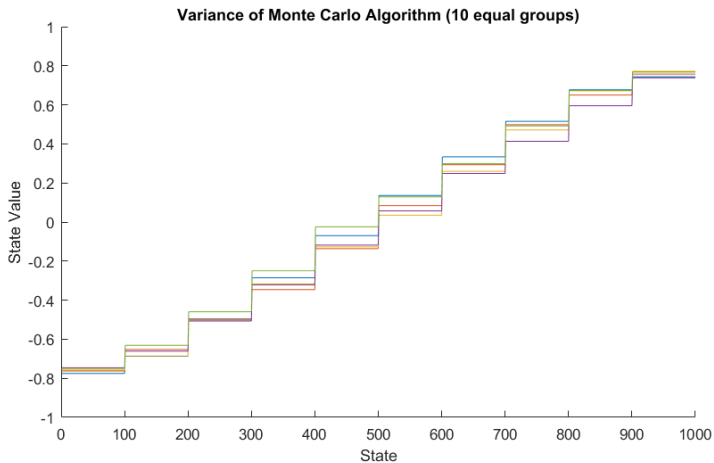
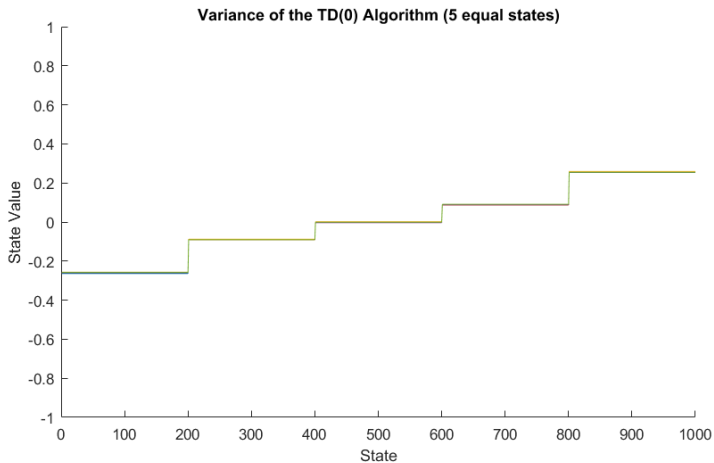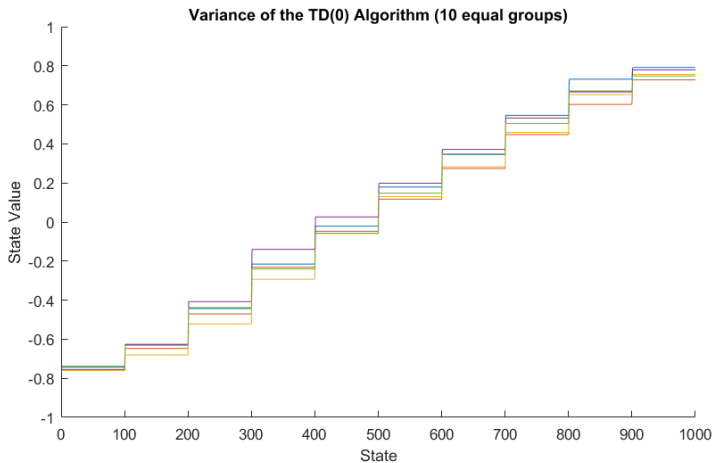# Variance IV

Variance V

Difference Grouping Strategies I
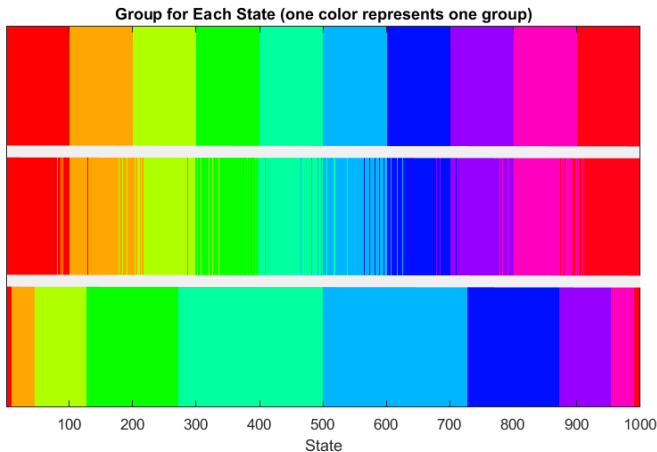
Grouping Strategies

- Constant size
- Constant size with noise

  A state near the limit of the group have a non-zero probability of being in the previous/next group.

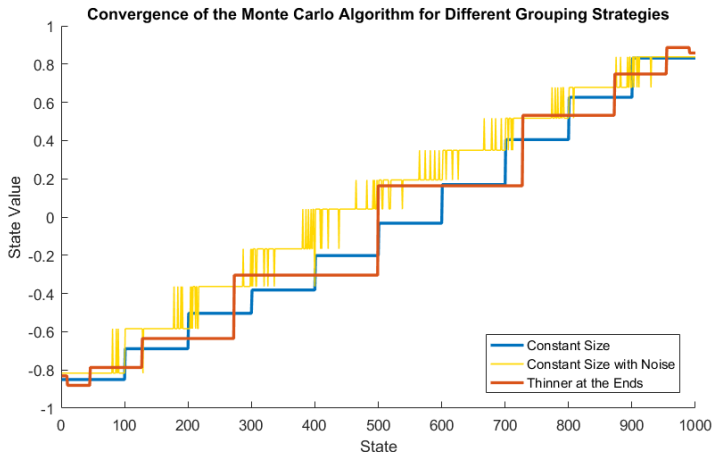- Larger groups in the center, thinner near the ends

Difference Grouping Strategies II

Visualization of the groups

(from top to bottom: Constant, Constant with noise, Thinner at the ends)

Group for Each State (one color represents one group)

Difference Grouping Strategies III

Difference Grouping Strategies IV