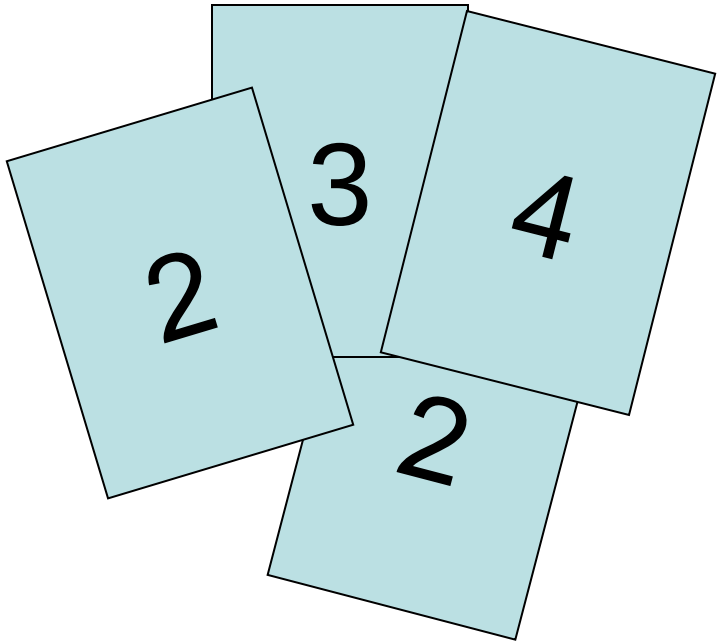
- Project 2, question 5:
  - Grade based on avg. score for won games

- layout
- commands.txt

Example: High-Low

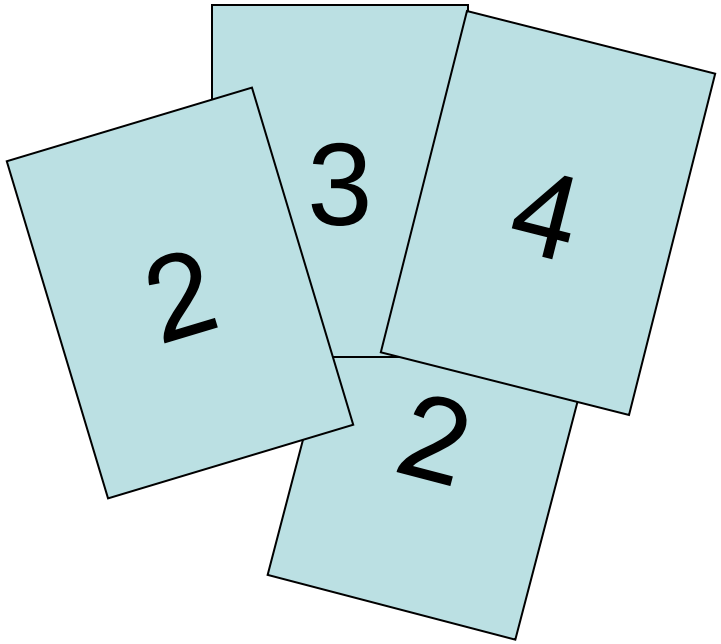
- Three card types: 2, 3, 4
- Infinite deck, twice as many 2's
- Start with 3 showing
- After each card, you say "high" or "low"
- New card is flipped
- If you're right, you win the points shown on the new card
- Ties are no-ops
- If you're wrong, game ends



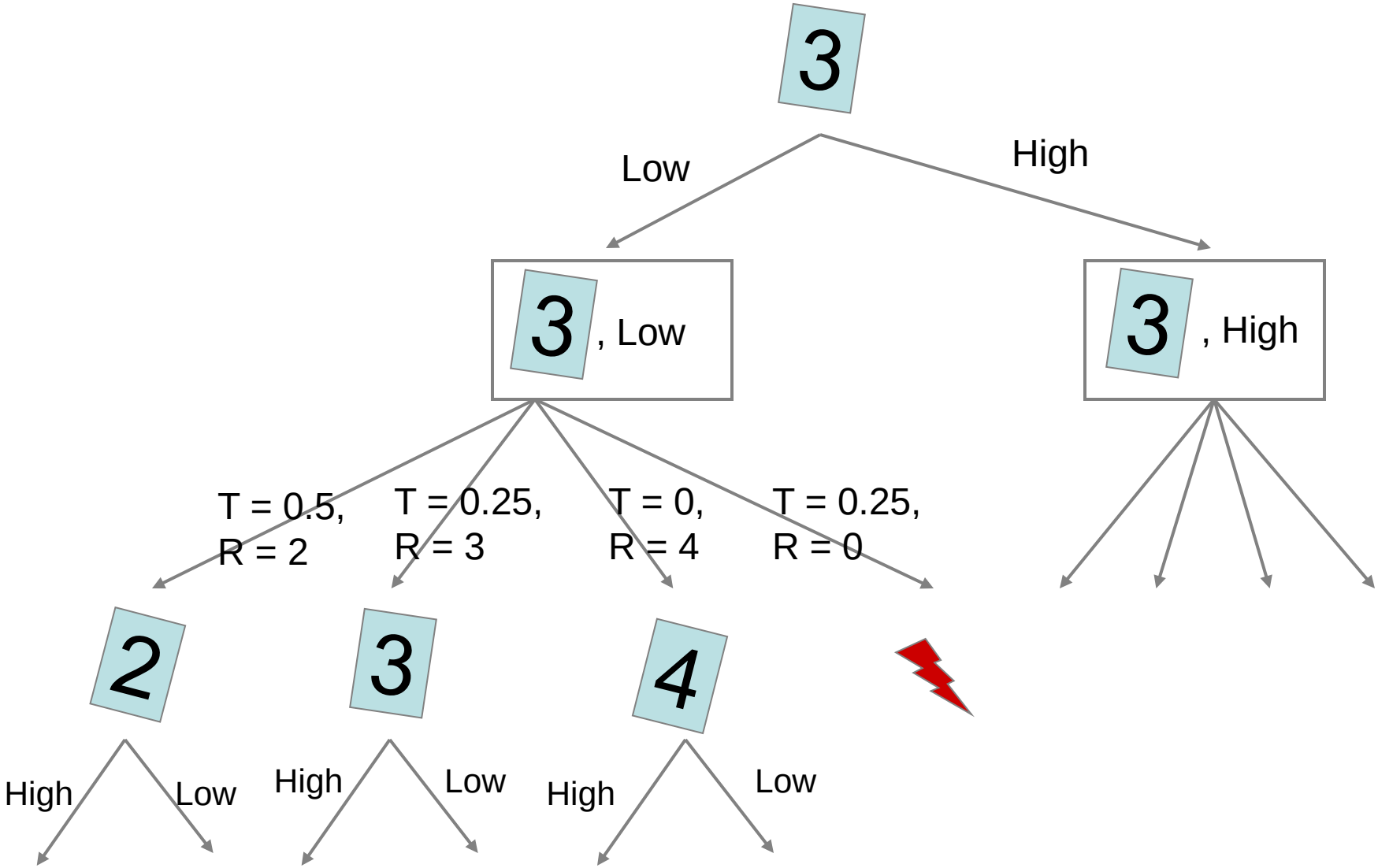
- Why not use expectimax?
  - #1: get rewards as you go
  - #2: you might play forever!

High-Low as an MDP

- States: 2, 3, 4, done
- Actions: High, Low
- Model: T(s, a, s'):
  - P(s'=4 | 4, Low) = 1/4
  - P(s'=3 | 4, Low) = 1/4
  - P(s'=2 | 4, Low) = 1/2
  - P(s'=done | 4, Low) = 0
  - P(s'=4 | 4, High) = 1/4
  - P(s'=3 | 4, High) = 0
  - P(s'=2 | 4, High) = 0
  - P(s'=done | 4, High) = 3/4
  - …
- Rewards: R(s, a, s'):
  - Number shown on s' if s $\neq$ s'
  - 0 otherwise
- Start: 3

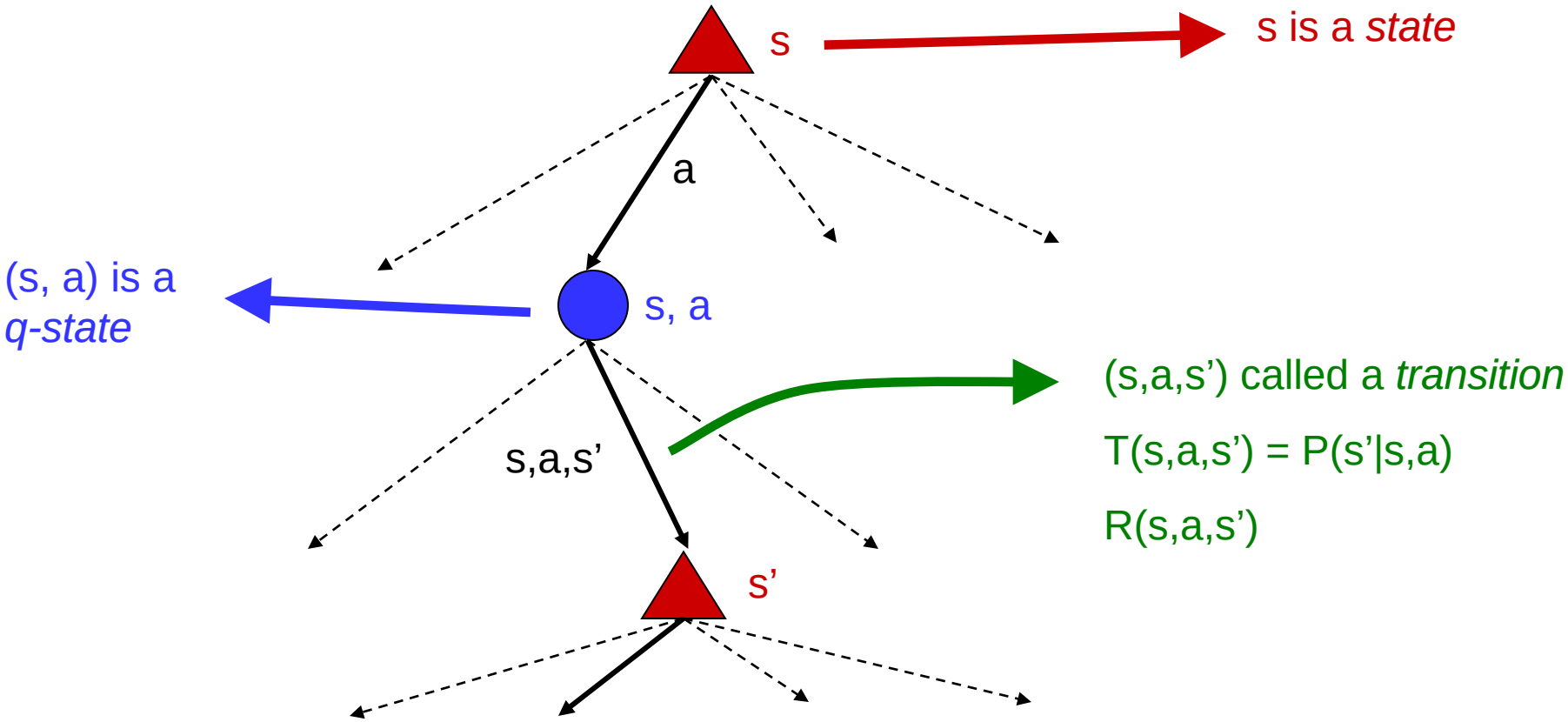


Example: High-Low



MDP Search Trees

- Each MDP state gives an expectimax-like search tree



17.2.1:

Anthony, Aaron, Greg, Kevin T.

17.2.2:

Bobby, Kumera, Maytee, Sam, Tong

17.2.3:

Chris, Dave, Nick, Tim

17.3:

Kayo, Kevin M., Ryan, Shannon