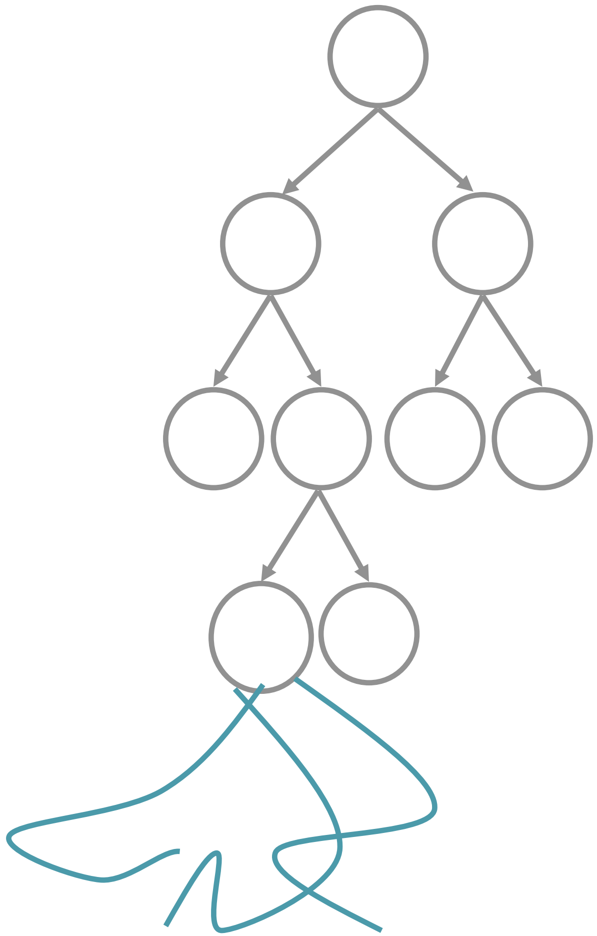


MONTE CARLO TREE SEARCH

ECS170 Spring 2018

Josh McCoy, @deftjams

Too Many Paths

α-β minimax time complexity:

$O\left(moves^{depth/2}\right)$ or $O\left(b^{m/2}\right)$

Monte Carlo Tree Search

- Alpha-Beta Minimax is good for modest branching factor with good heuristic.
- MCTS: no evaluation function/heuristic!
  - Do this many times:
    - Simulate with random moves
    - Score game outcome and keep stats
    - Play move with best stats.

The Tree Part

Statistics tree construction.

Used as a guide to focus on the most promising parts of the search tree.

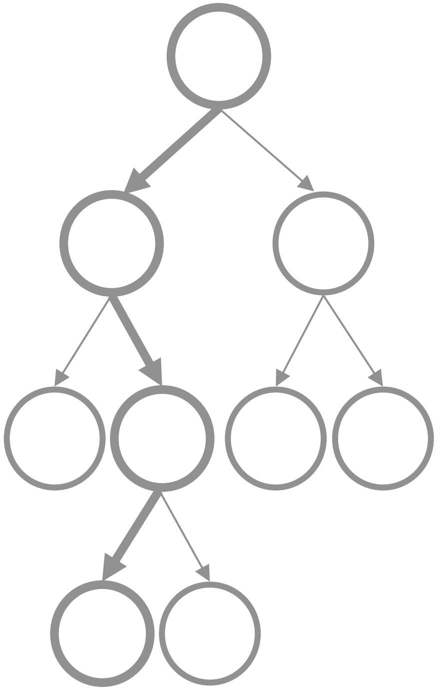
Simulations determine utility/value.

MCTS Hight Level Process

1. Selection
2. Expansion
3. Play-out
4. Backpropagation

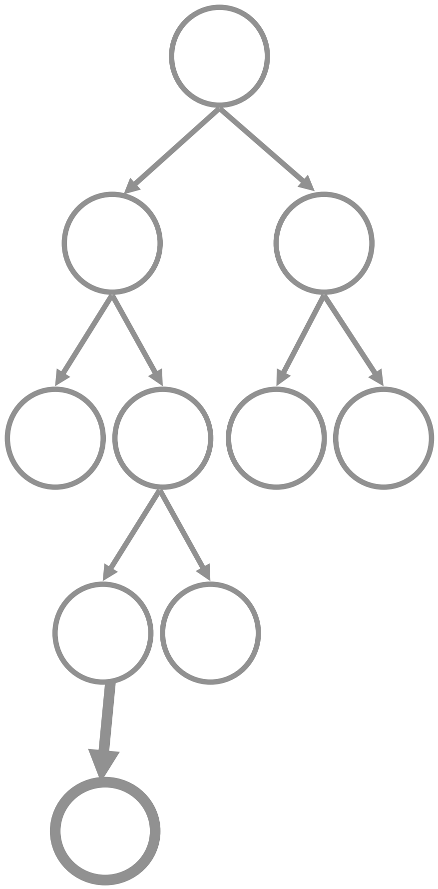
MCTS Hight Level Process

1. Selection
2. Expansion
3. Play-out
4. Backpropagation



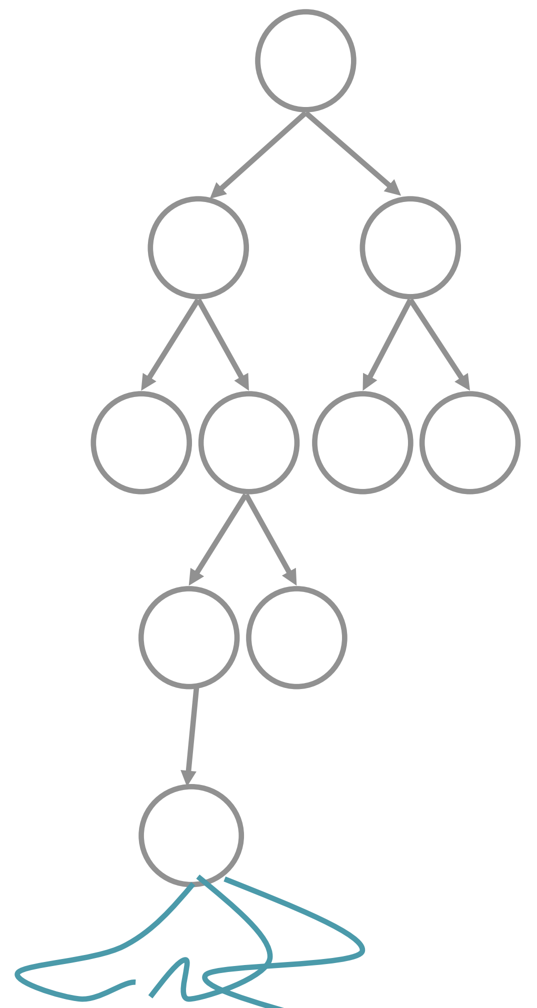
MCTS Hight Level Process

1. Selection
2. Expansion
3. Play-out
4. Backpropagation



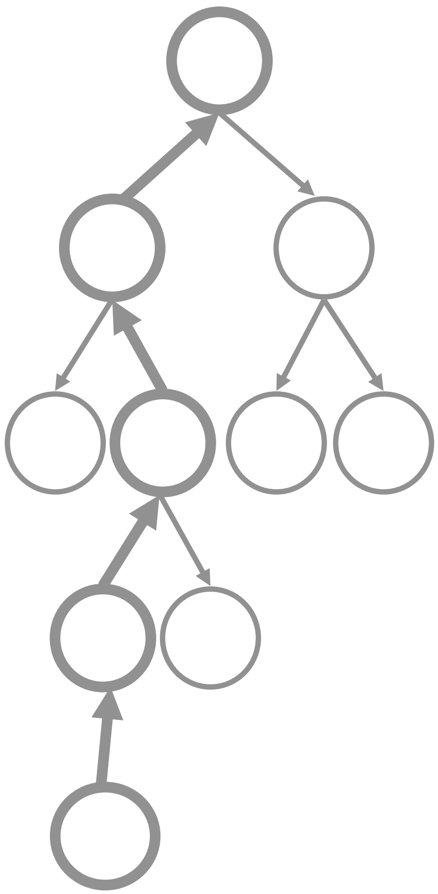
MCTS Hight Level Process

1. Selection
2. Expansion
3. Play-out
4. Backpropagation



MCTS Hight Level Process

1. Selection
2. Expansion
3. Play-out
4. Backpropagation



Upperbound Confidence

Multi-armed Bandit

P(A wins) = 0.4

P(B wins) = 0.55

P(C wins) = 0.7

P(D wins) = 0.3

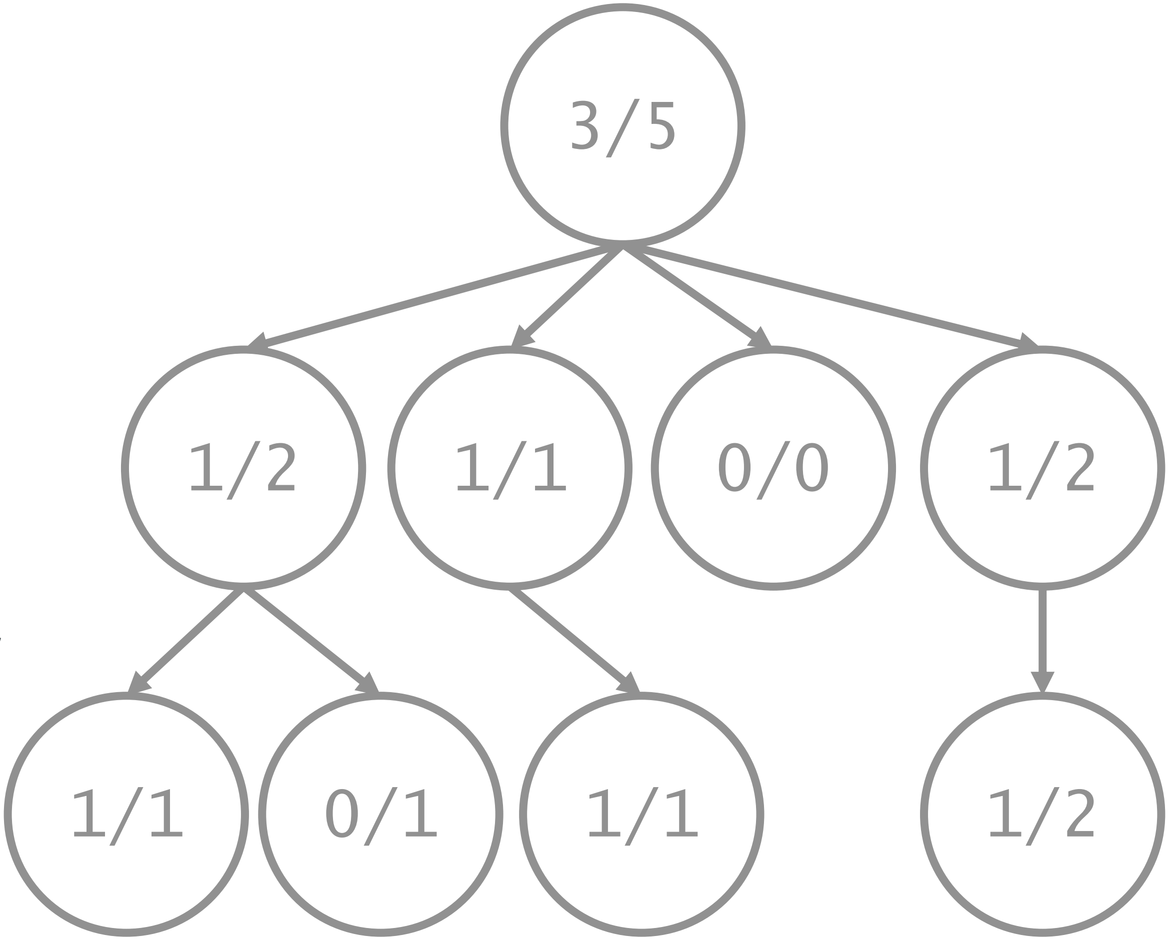
Each play has win/lose binary result.

Selection

We need:

- Tree of win statistics
- Selection policy

(still alternating player and opponent at each depth)



Upper Bound Confidence

$$\underbrace{v_i}_{\text{value/ utility}} = \underbrace{x_i}_{\text{mean reward}} + \sqrt{\frac{\ln(\overbrace{N}^{\text{total trials}})}{\underbrace{n_i}_{\text{trials for subtree}}}}$$

- Higher value estimate is better
- Gives an estimate of true value.
- More trials makes better estimate

```
v_i = x_i + math.sqrt(math.log(N)/n_i)
```

AAAI-14 Games Tutorial, by Martin Muller:

Upper Bound Confidence

$$v_i = x_i + \sqrt{\frac{\ln(N)}{n_i}}$$

total trials

trials for subtree

value/ utility

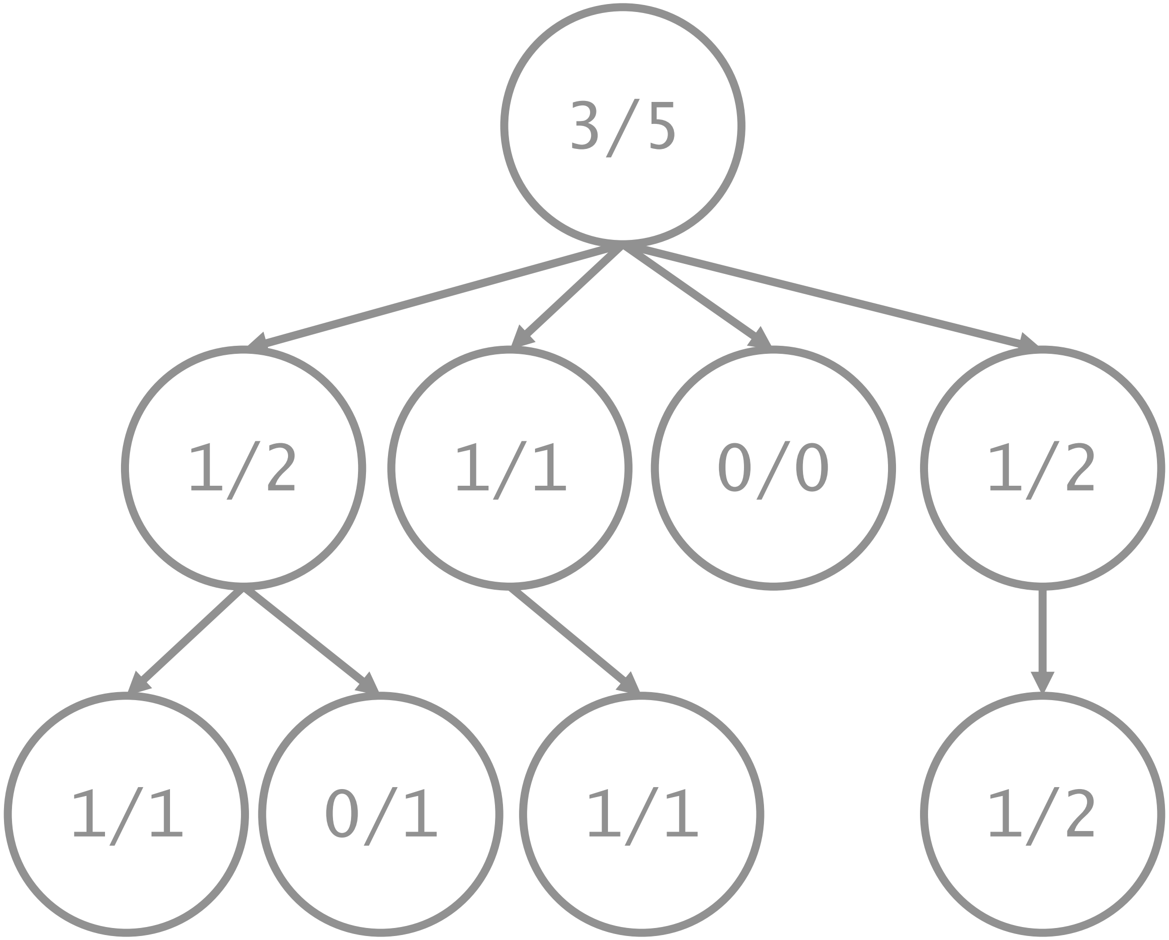
mean reward

- Confidence interval is large when n_i is small, shrinks in proportion to $\sqrt{n_i}$
  - High uncertainty about move, larger **exploration** term
- Explore: if trials is less than total trials.
- Tolerant to adversarial situations.

AAAI-14 Games Tutorial, by Martin Muller:

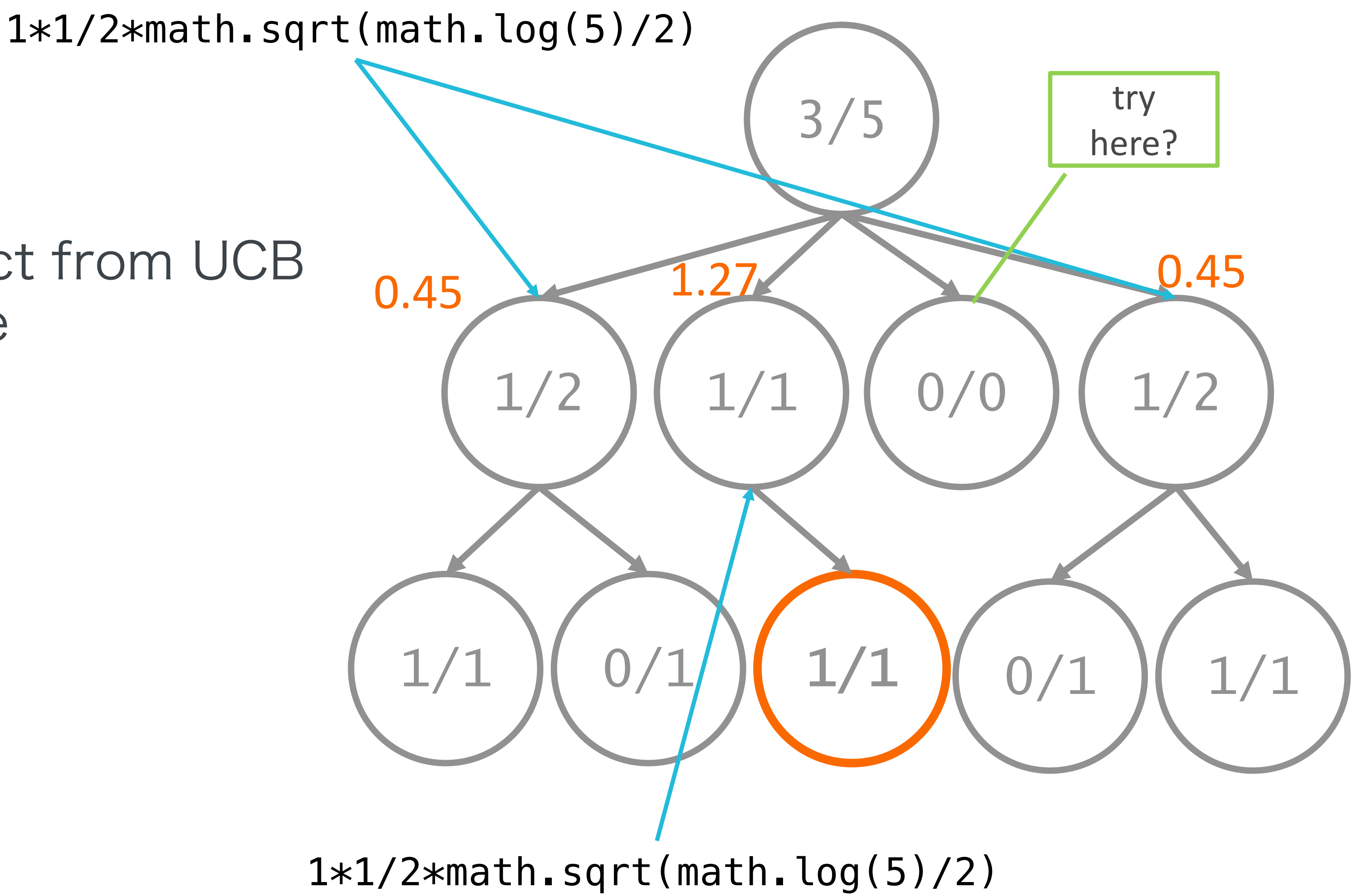
Selection

Select from UCB value



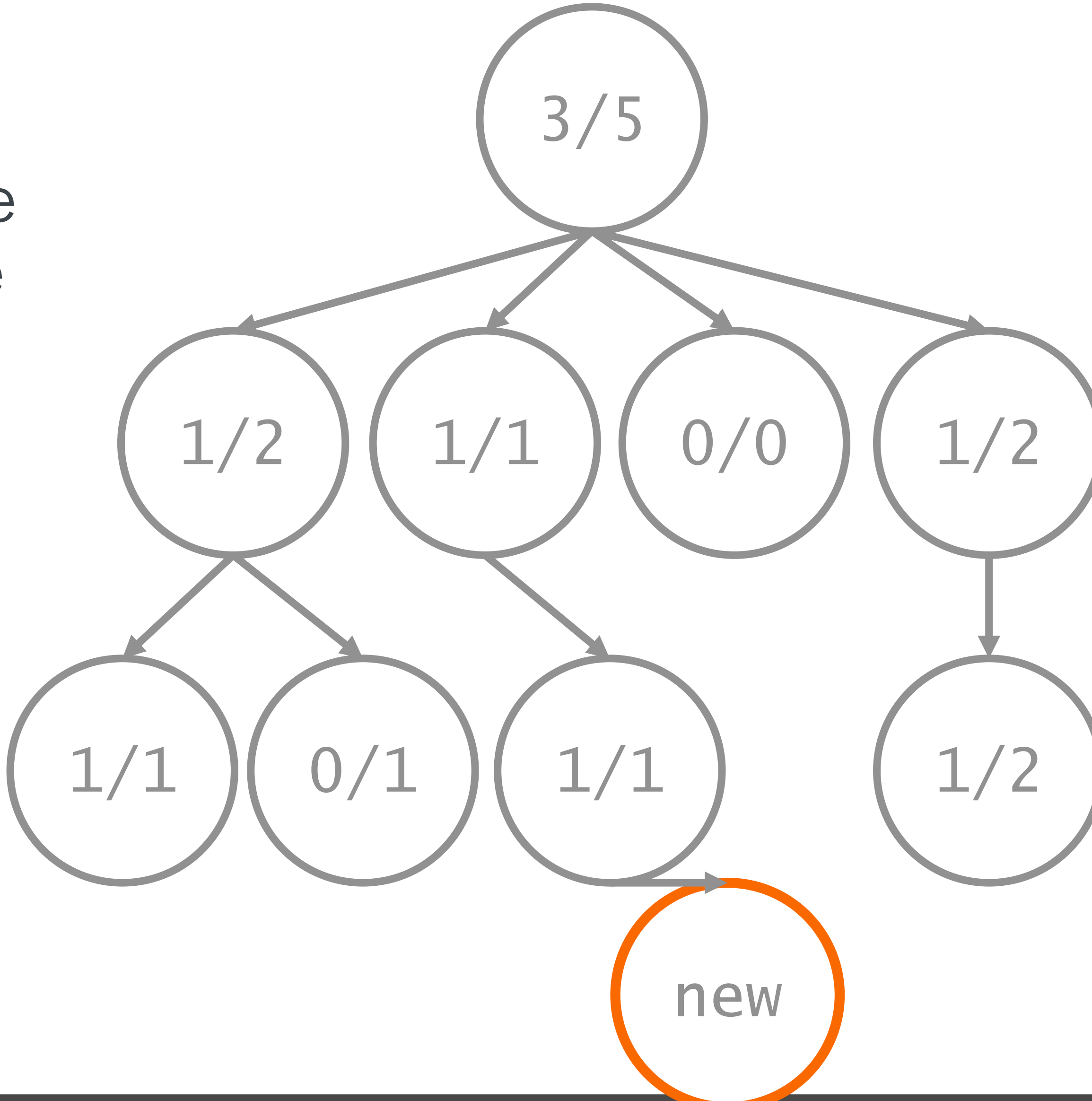
Selection

Select from UCB value



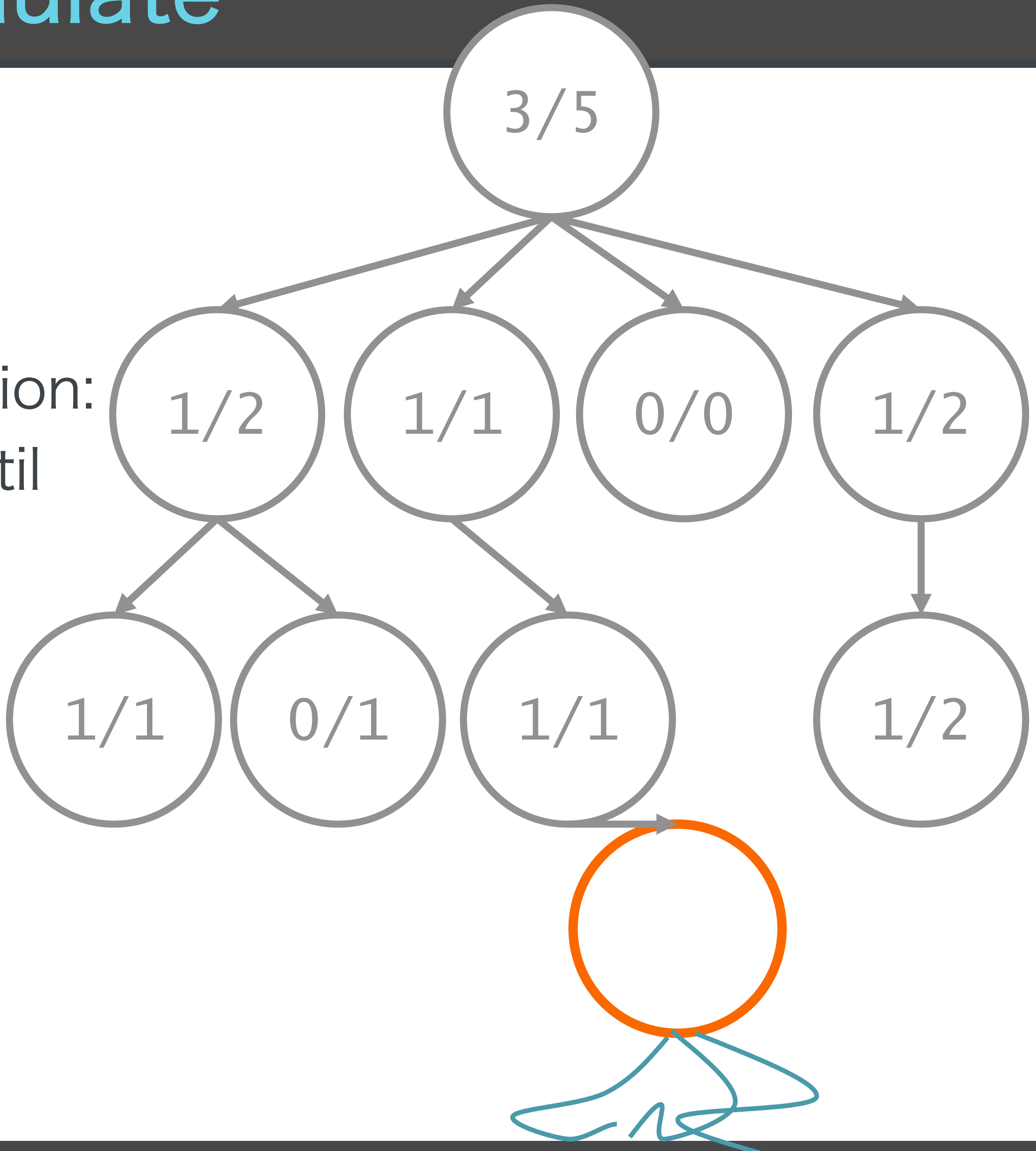
Expansion

Create a new node on the search tree based off UCB1.



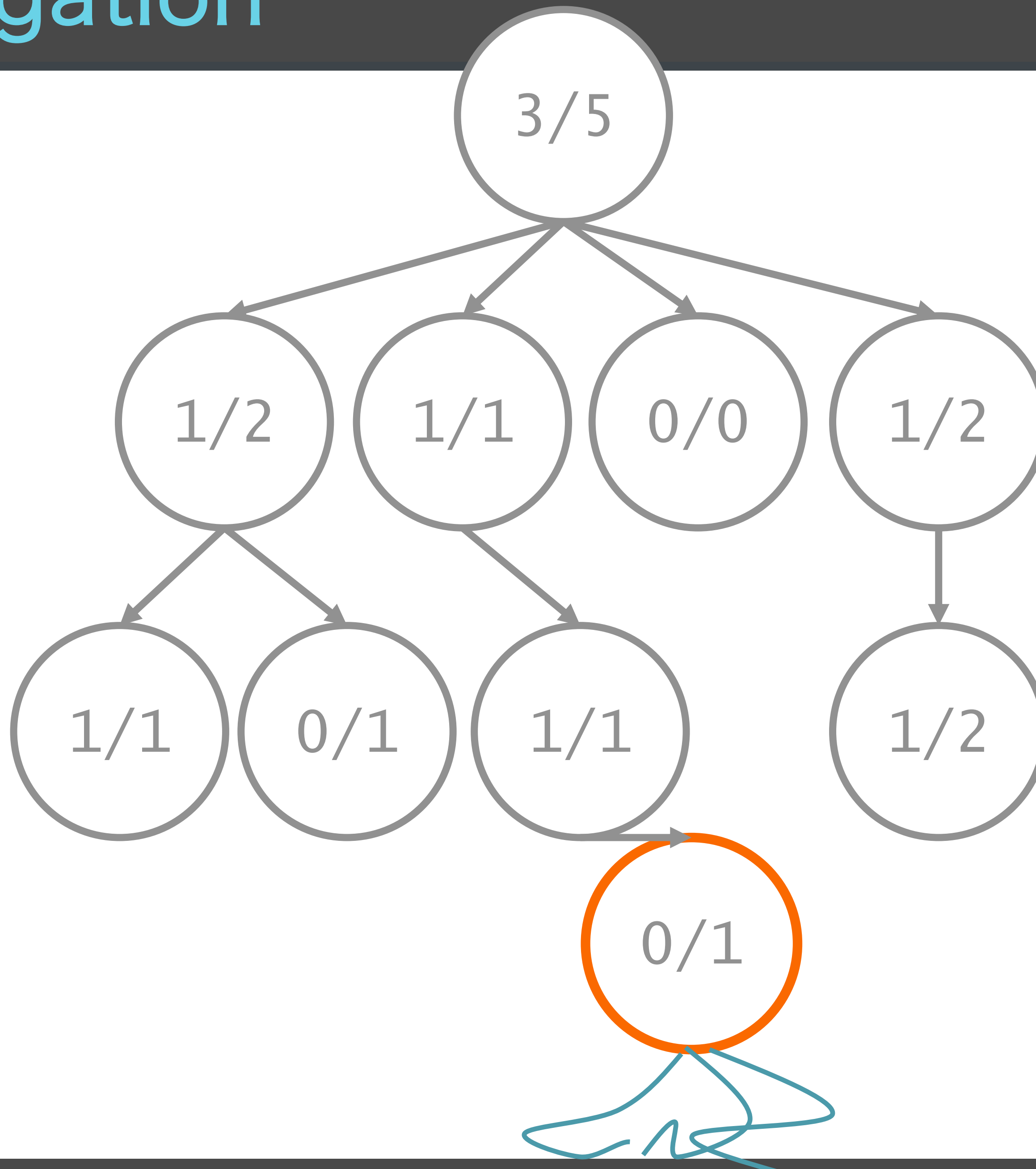
Playout/Simulate

Stochastic simulation: random moves until payout.



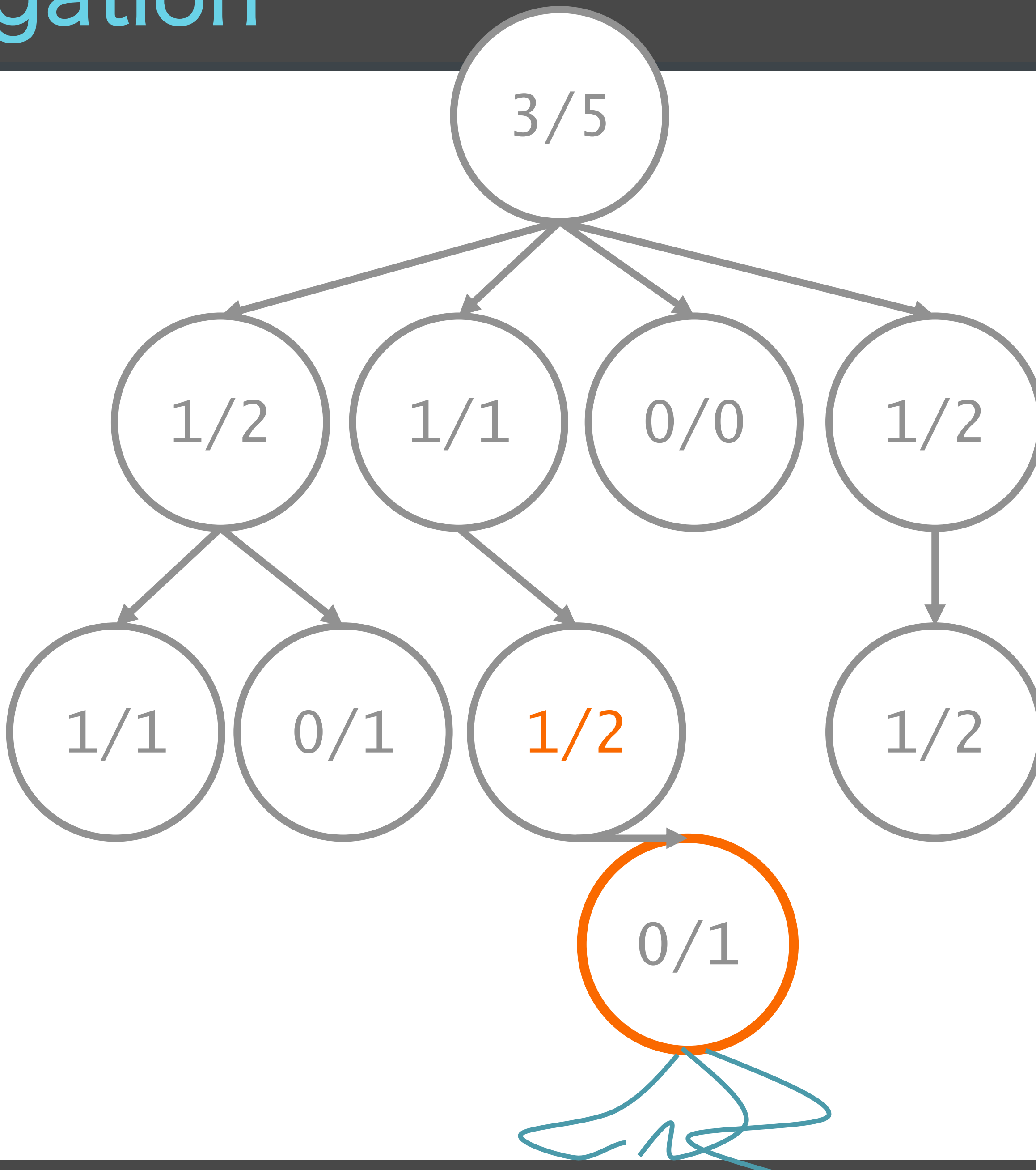
Back Propagation

Update the tree based off simulation results.



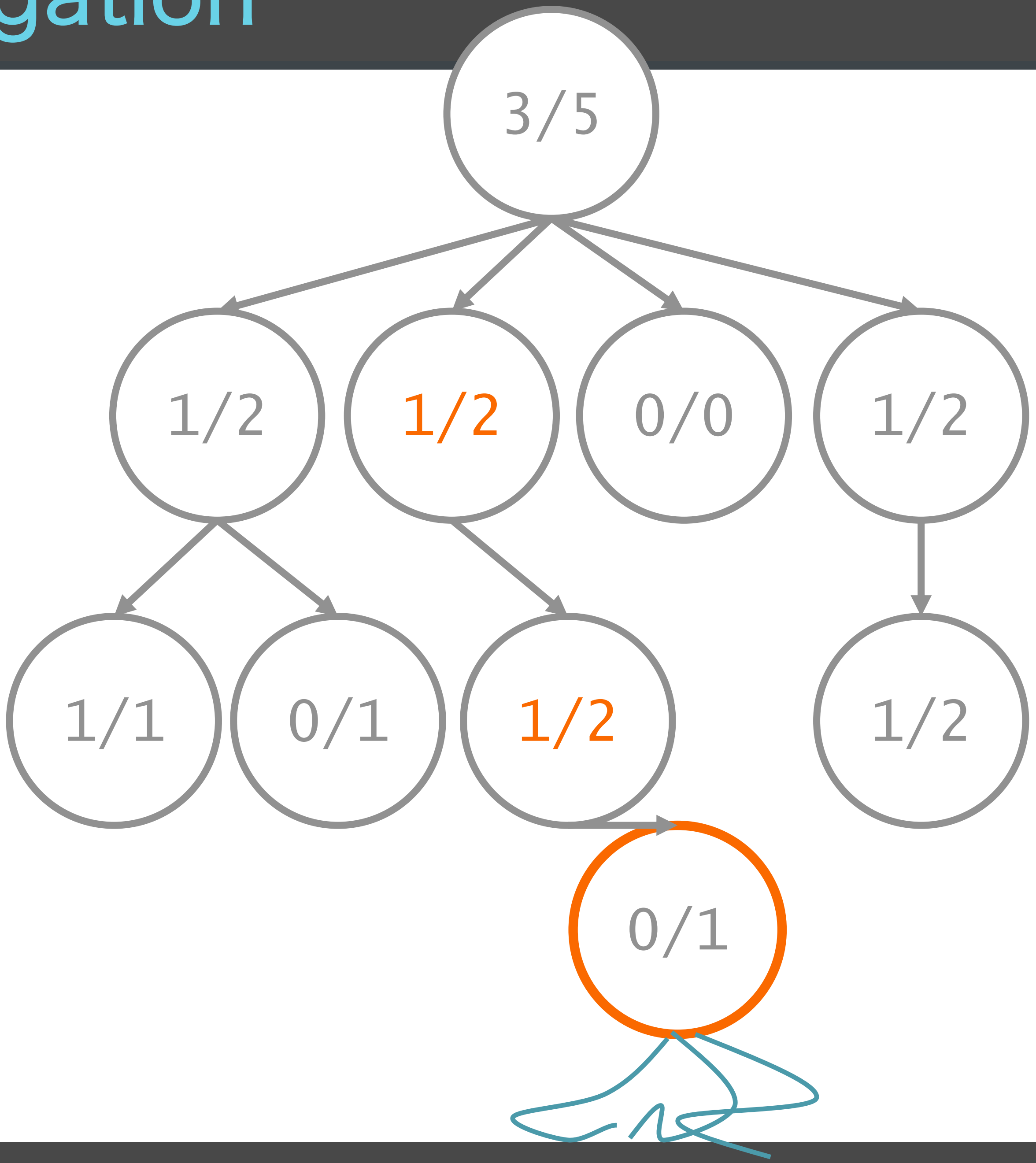
Back Propagation

Update the tree based off simulation results.



Back Propagation

Update the tree based off simulation results.



Back Propagation

Update the tree based off simulation results.