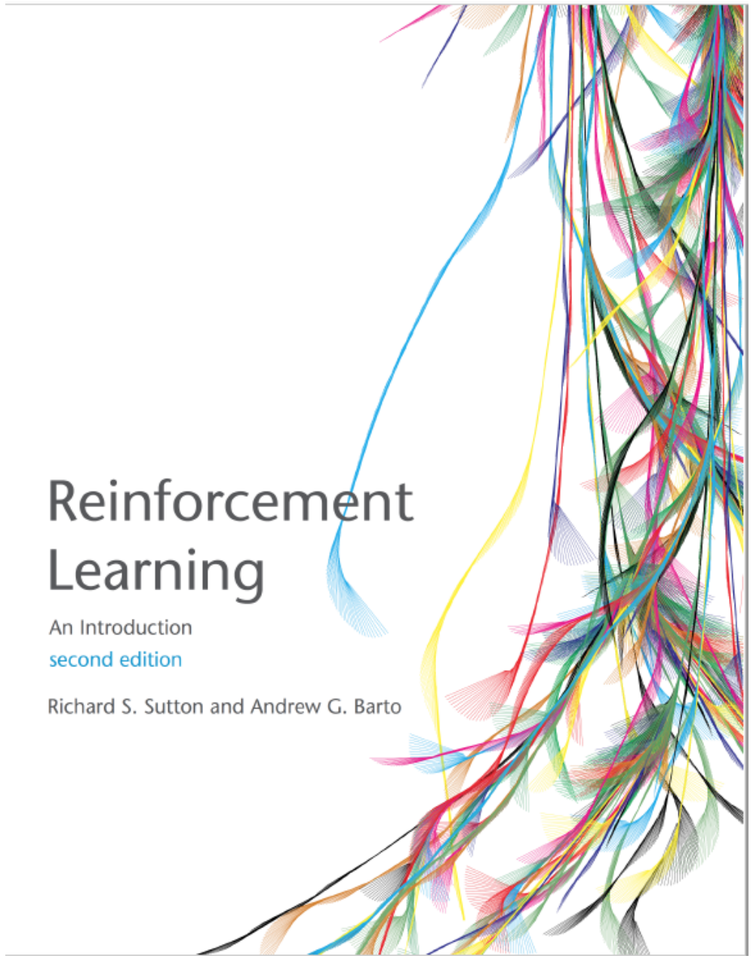
Tree Search

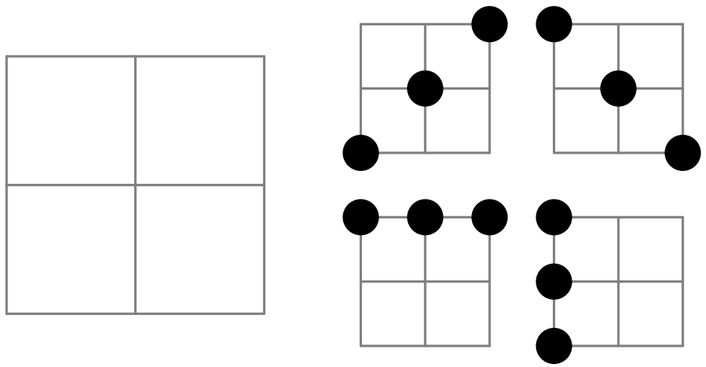


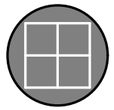
partially

Section 8.11

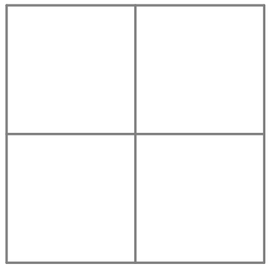
Game Tree

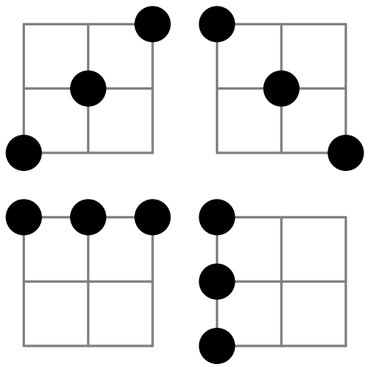
Game Tree

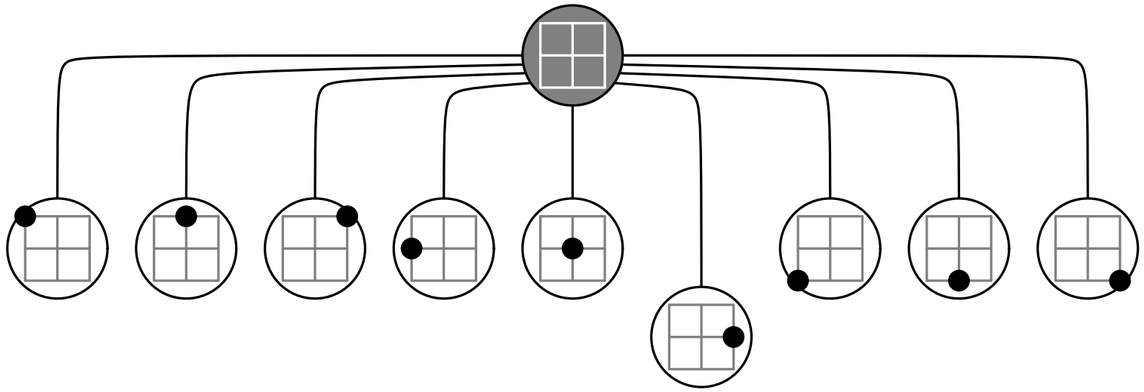




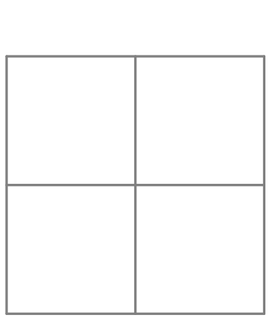
Game Tree

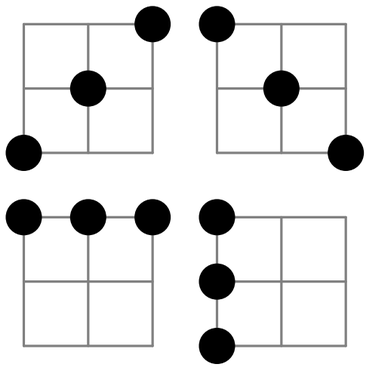




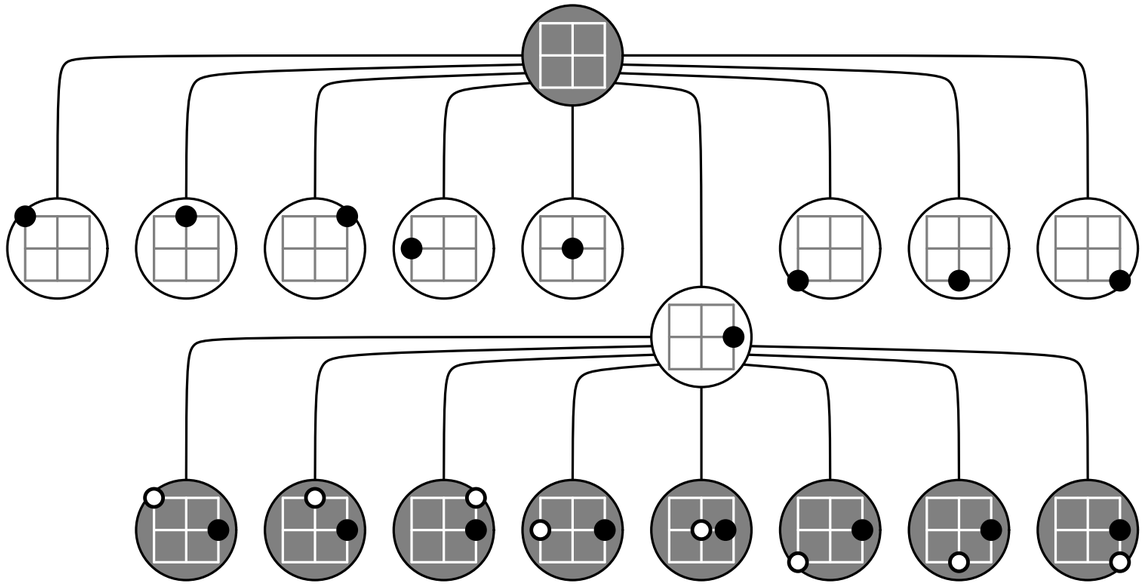


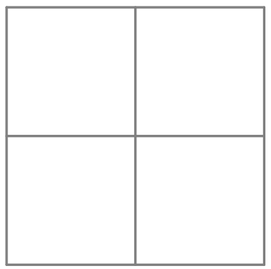
Game Tree

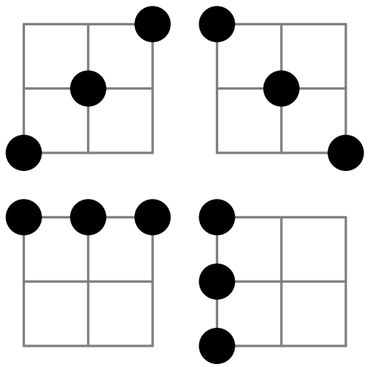




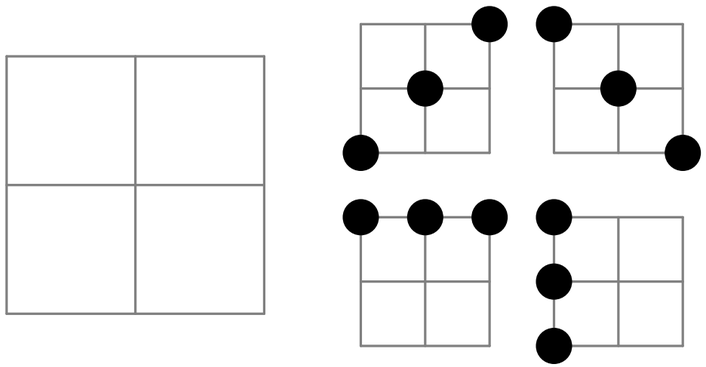
Game Tree







Game Tree



branching factor

states

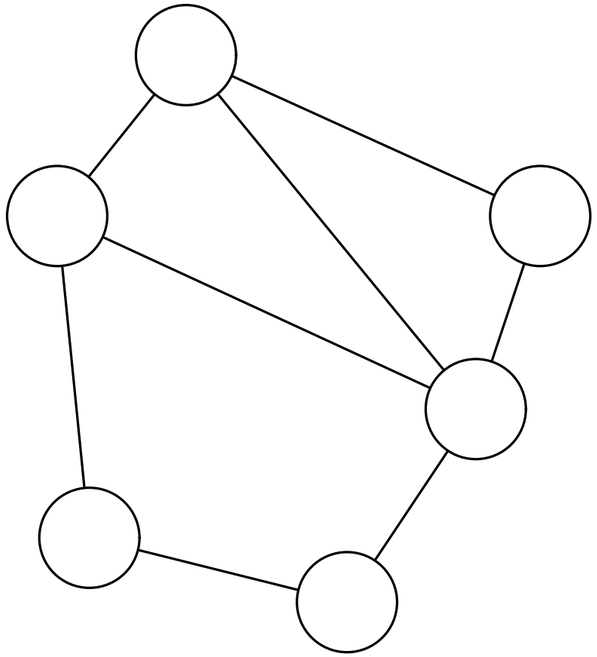
depth

leaves

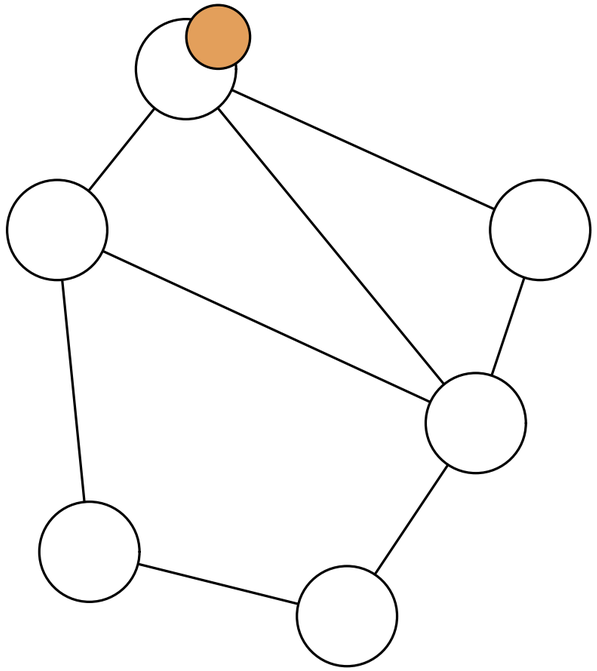
reward

Example

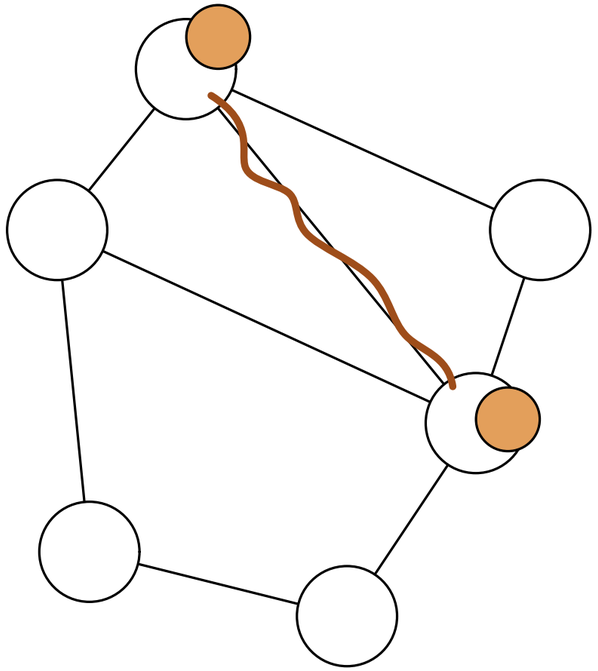
Geography



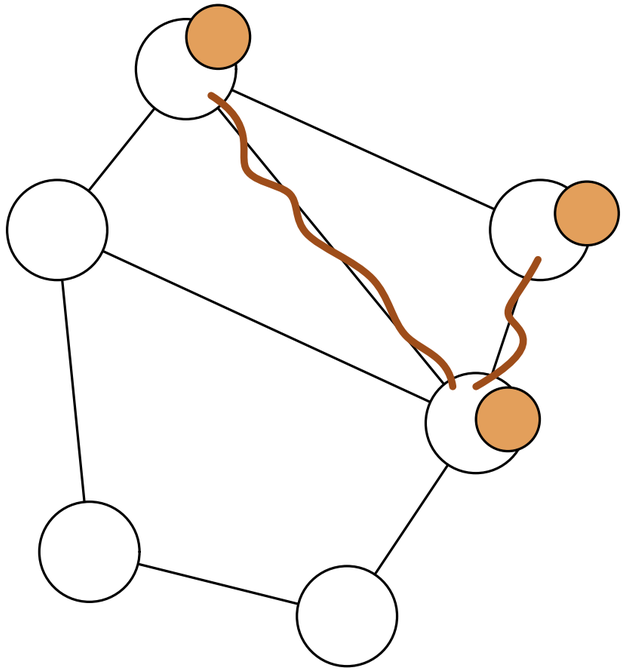
Geography



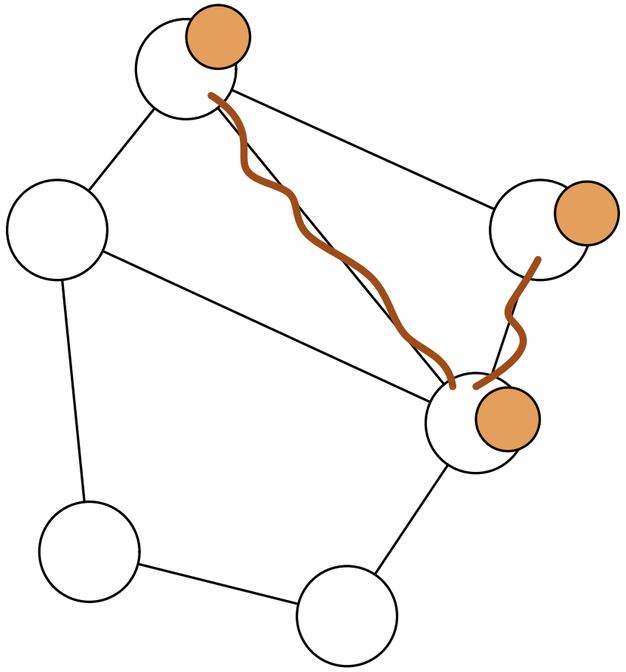
Geography



Geography

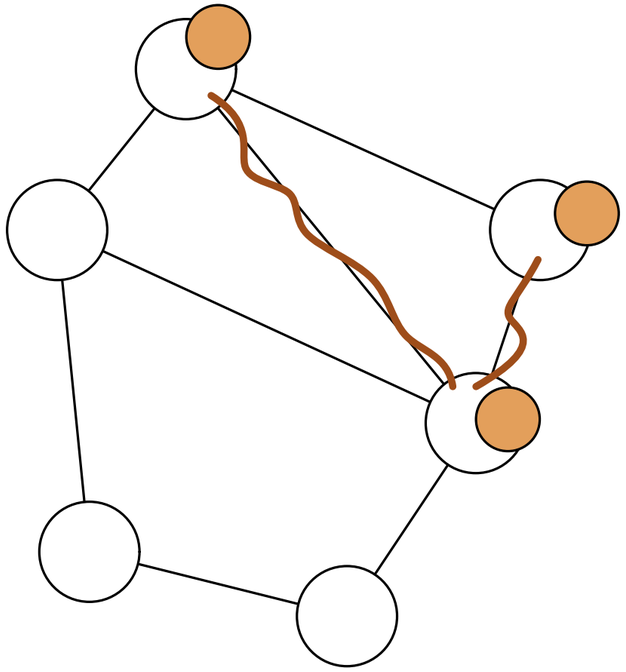


Geography



Lost

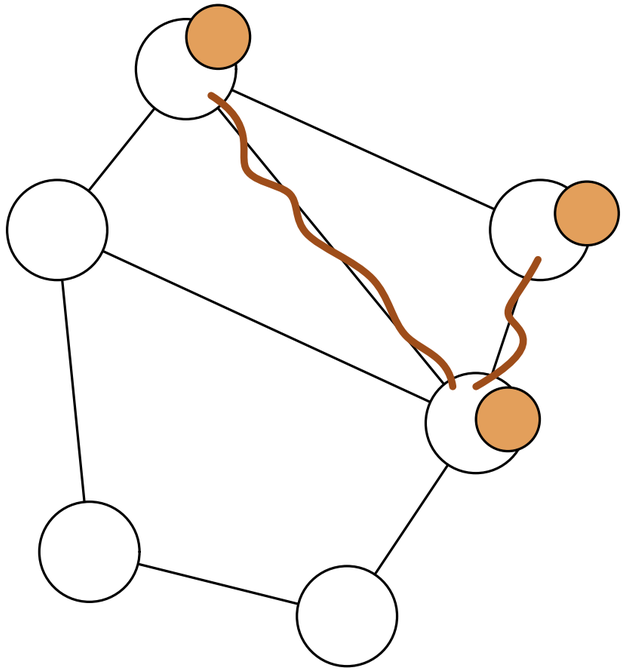
Geography



Lost

Graph with n nodes
max degree d

Geography



Lost

Graph with n nodes
max degree d

states $\leq 2^n \cdot n$
branching factor $\leq$ d
depth $= n$
leaves $=$ “dead-end”
reward $= -1/1$

Assignment 1

Checkers
Nim (5,5,5)

Post on Teams

branching factor

depth

leaves

states

reward

Checkers
Nim (5,5,5)

Post on Teams

branching factor

depth

leaves

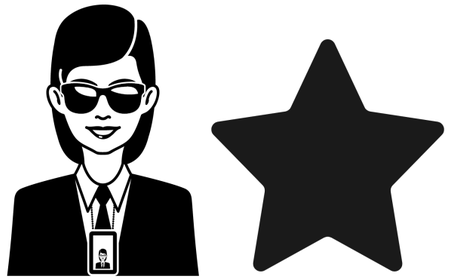
states

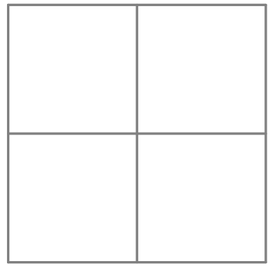
reward

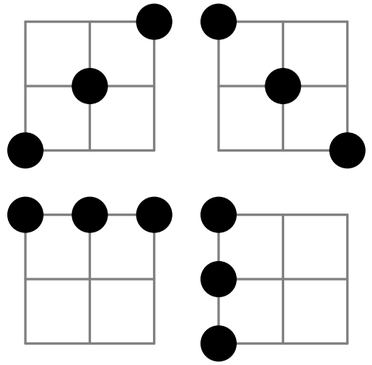
$$\text{State} \longrightarrow \mathbb{R}$$

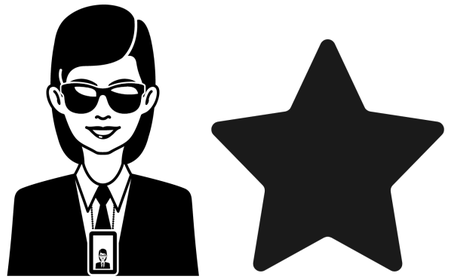
Static Evaluation
Score Function
(Approximate) Value v

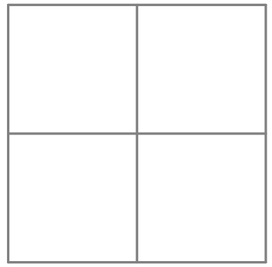
Example

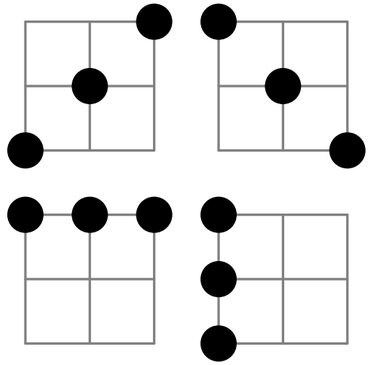


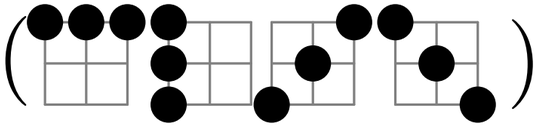


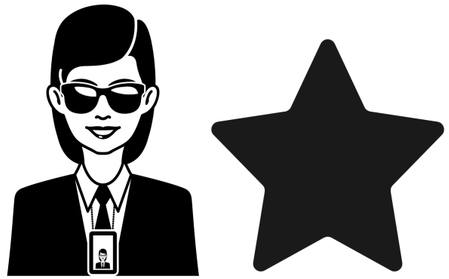


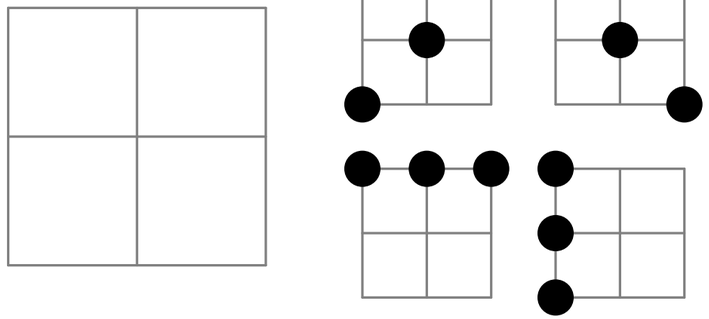






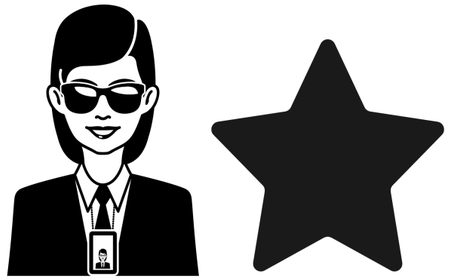
32 ⓗ 

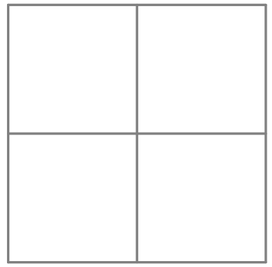


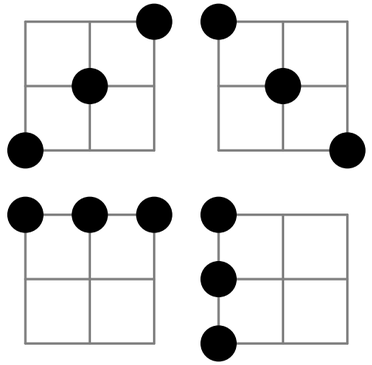


$$32 \circledast \left(\quad \right)$$

$$+ \circledast \left(\quad \cdots \right)$$





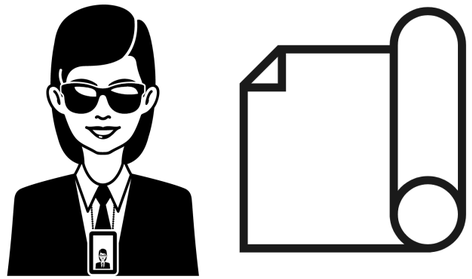


$$32\,\#\,(\quad)$$

$$+\,\#\,(\quad\cdots)$$

$$-32\,\#\,(\quad)$$

$$-\,\#\,(\quad\cdots)$$

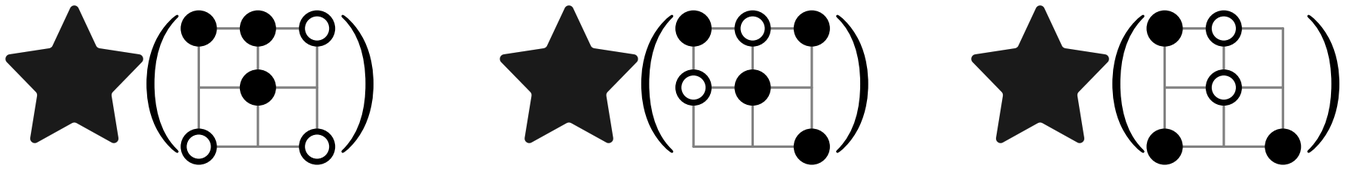


100 roll-outs with random agent.

value = average winrate

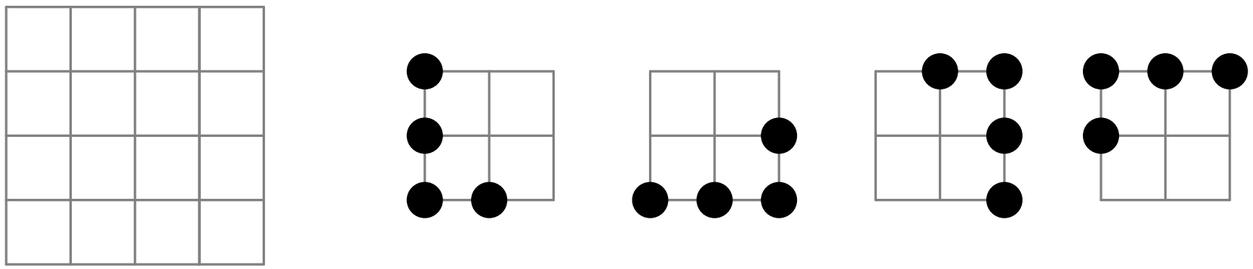
Assignment 2

Compute

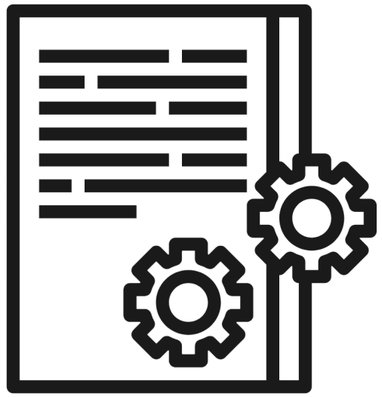


Post on Teams

Describe a score function for the following setup.



(several possible solutions)



Policy

deterministic

State $\longrightarrow$ Action

probabilistic

State $\longrightarrow$ Probability Distribution over Actions

Assignment 3

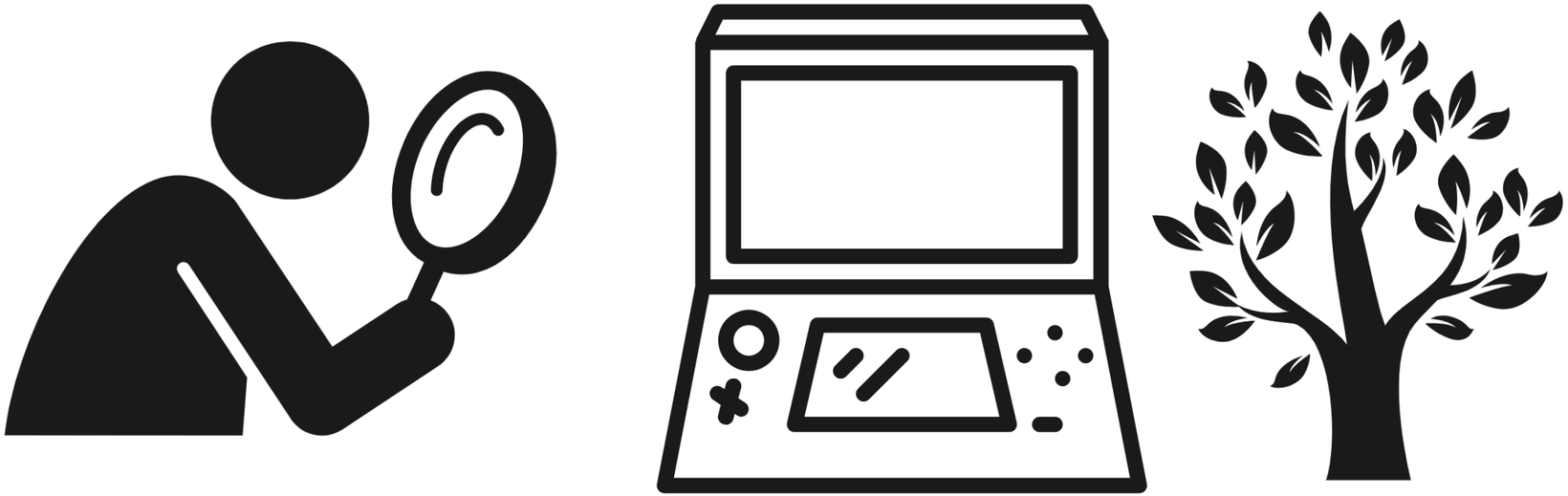
How does a value function determine a policy?

deterministic / random ?

How does a policy define a value function?

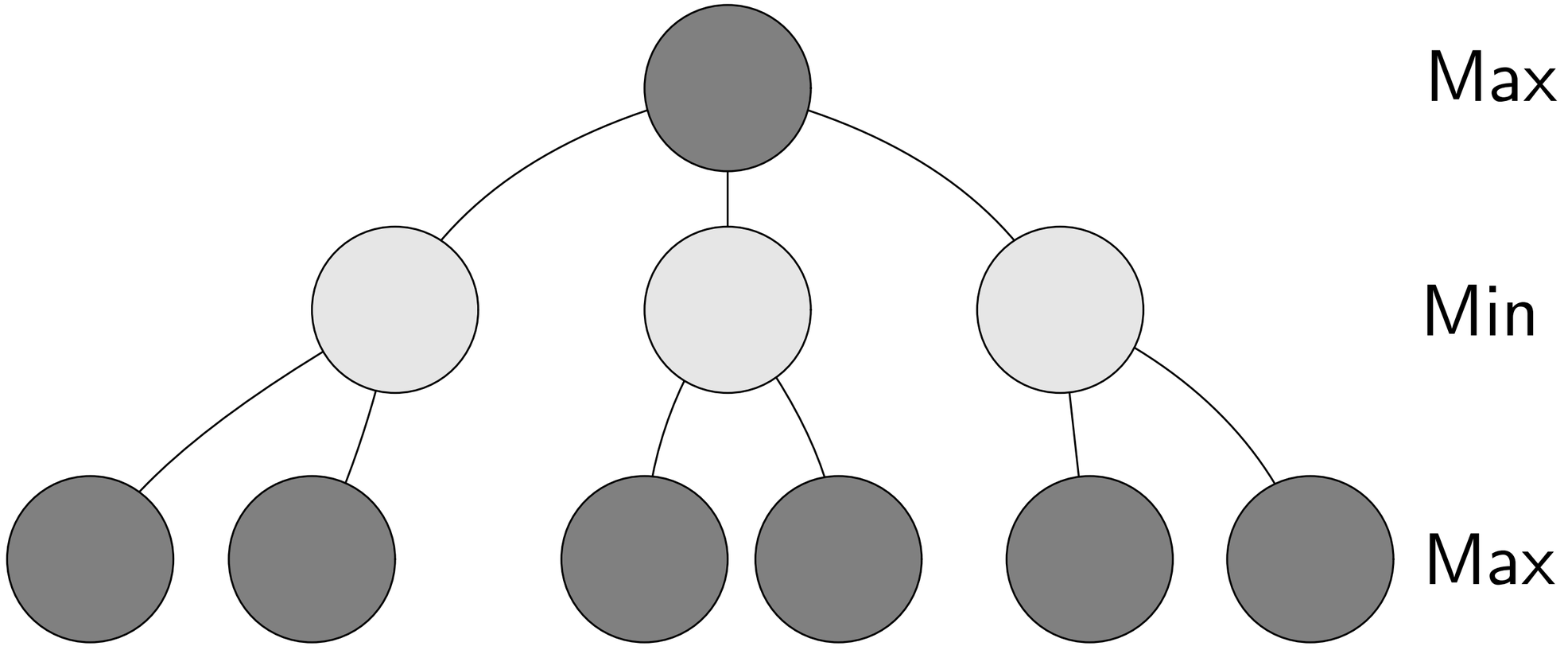
Why might a **random** policy be better?
(for learning)

Post on Teams

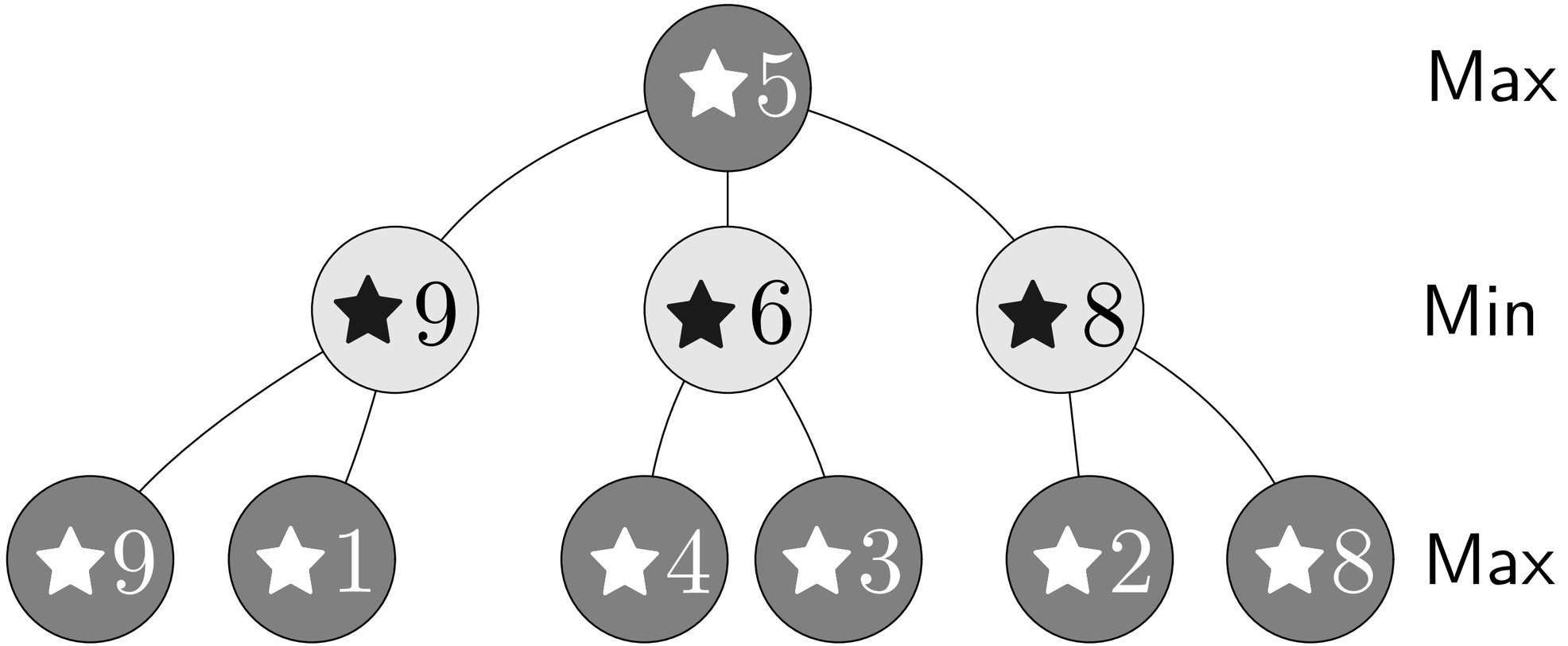


Min Max

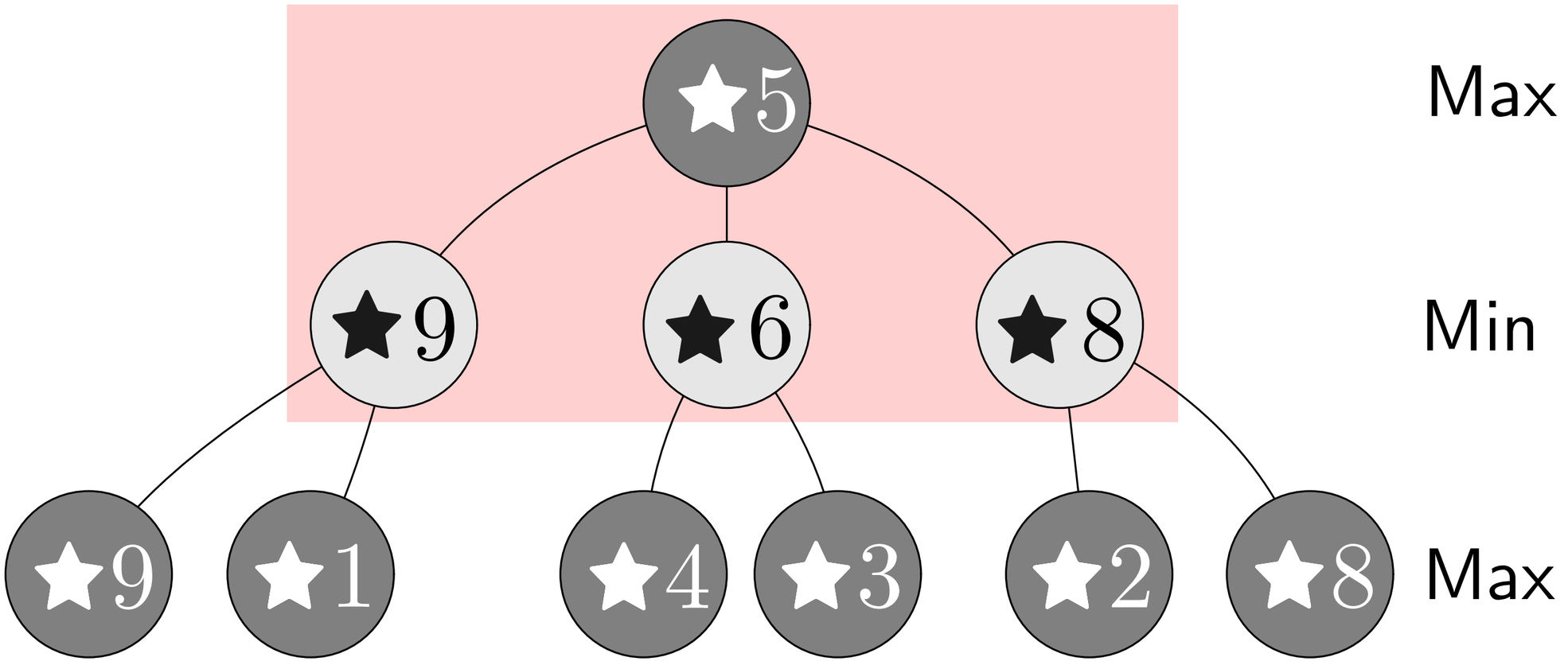
Min Max



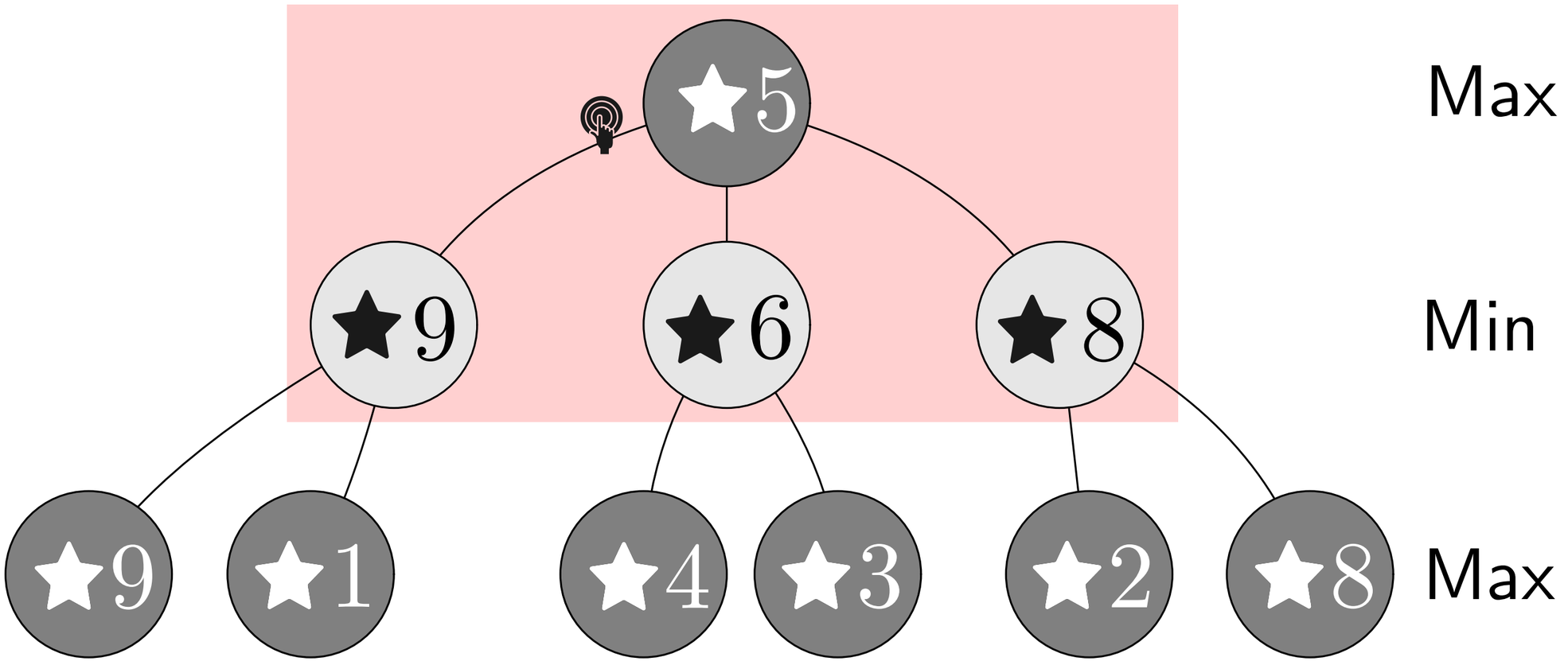
Min Max



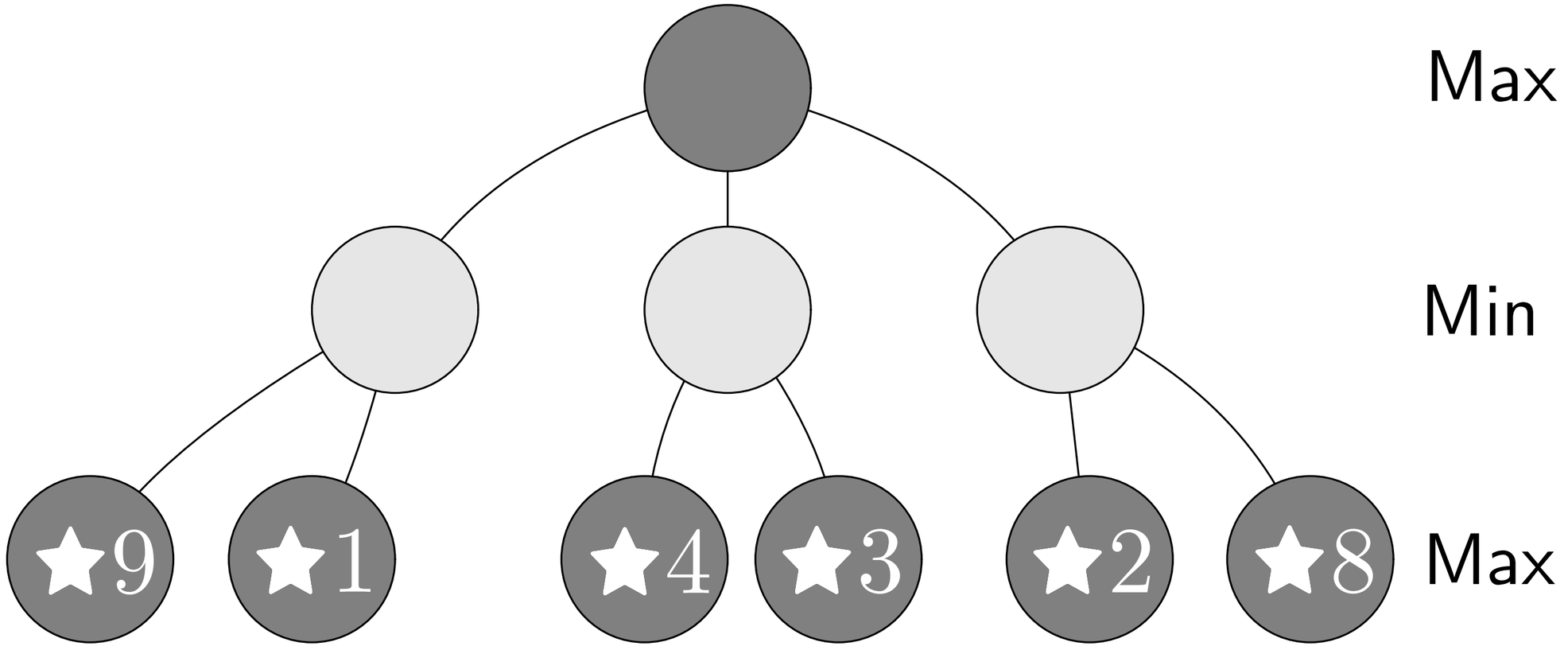
Min Max



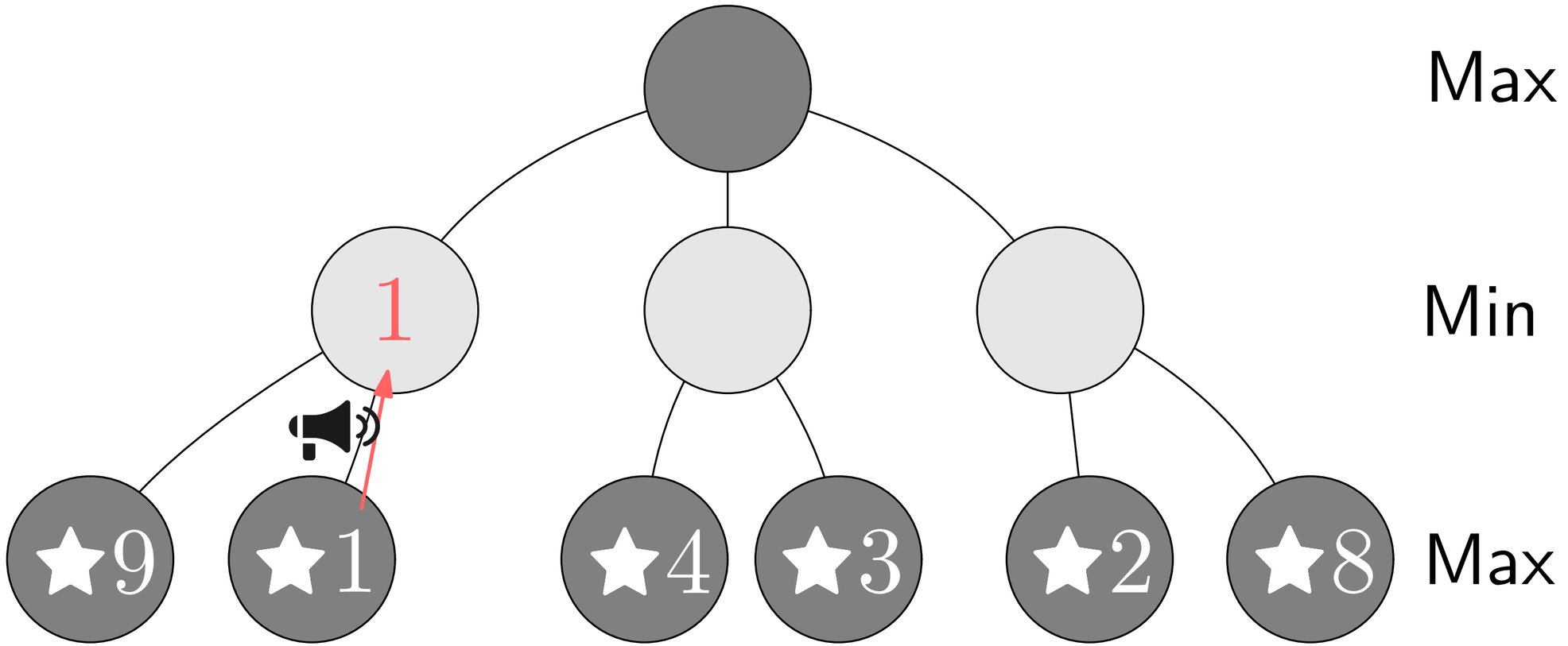
Min Max



Min Max



Min Max

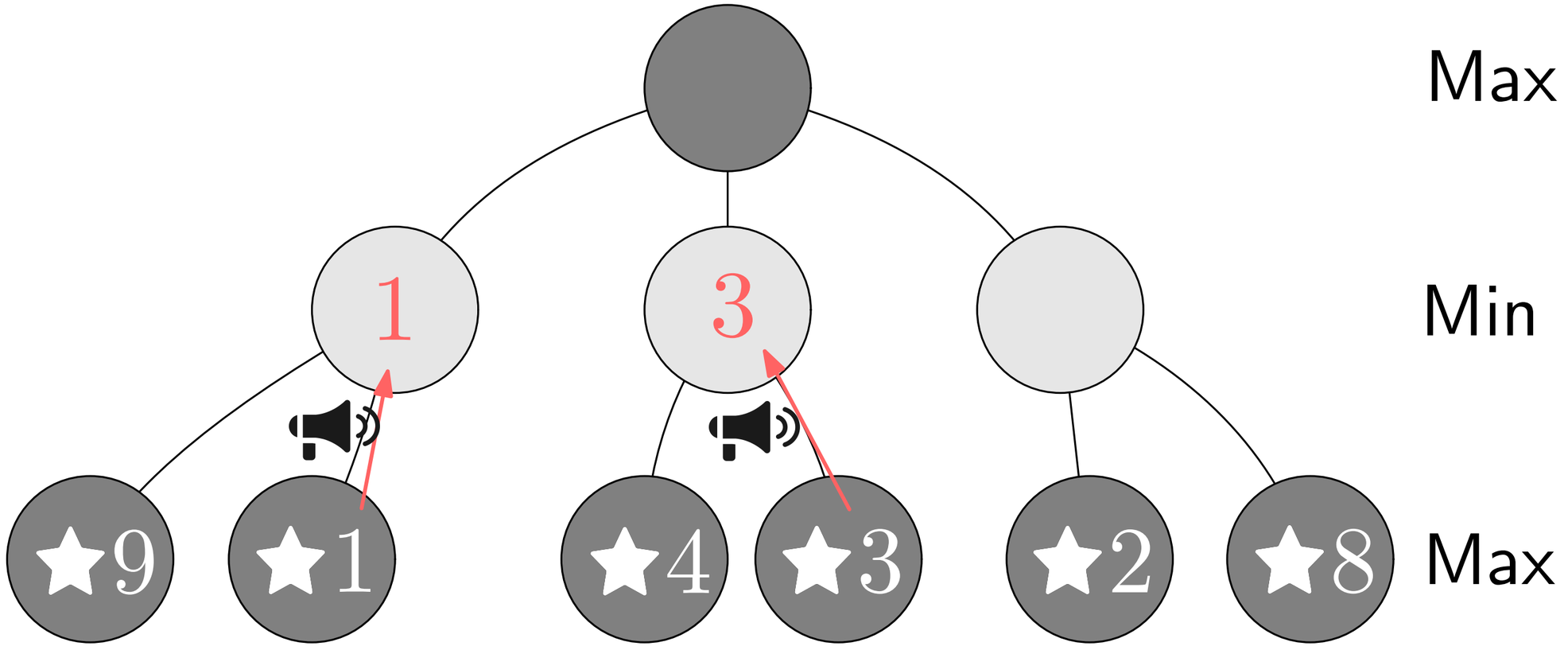


Max

Min

Max

Min Max

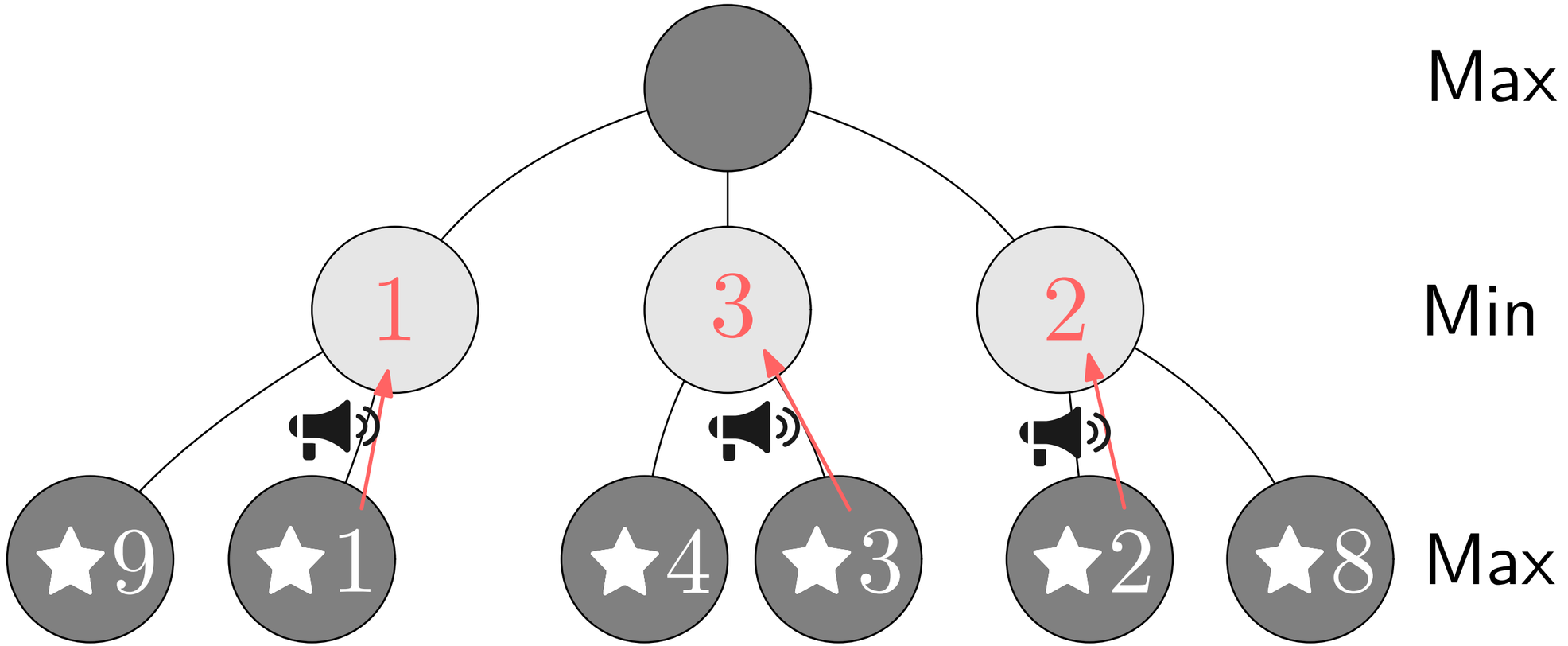


Max

Min

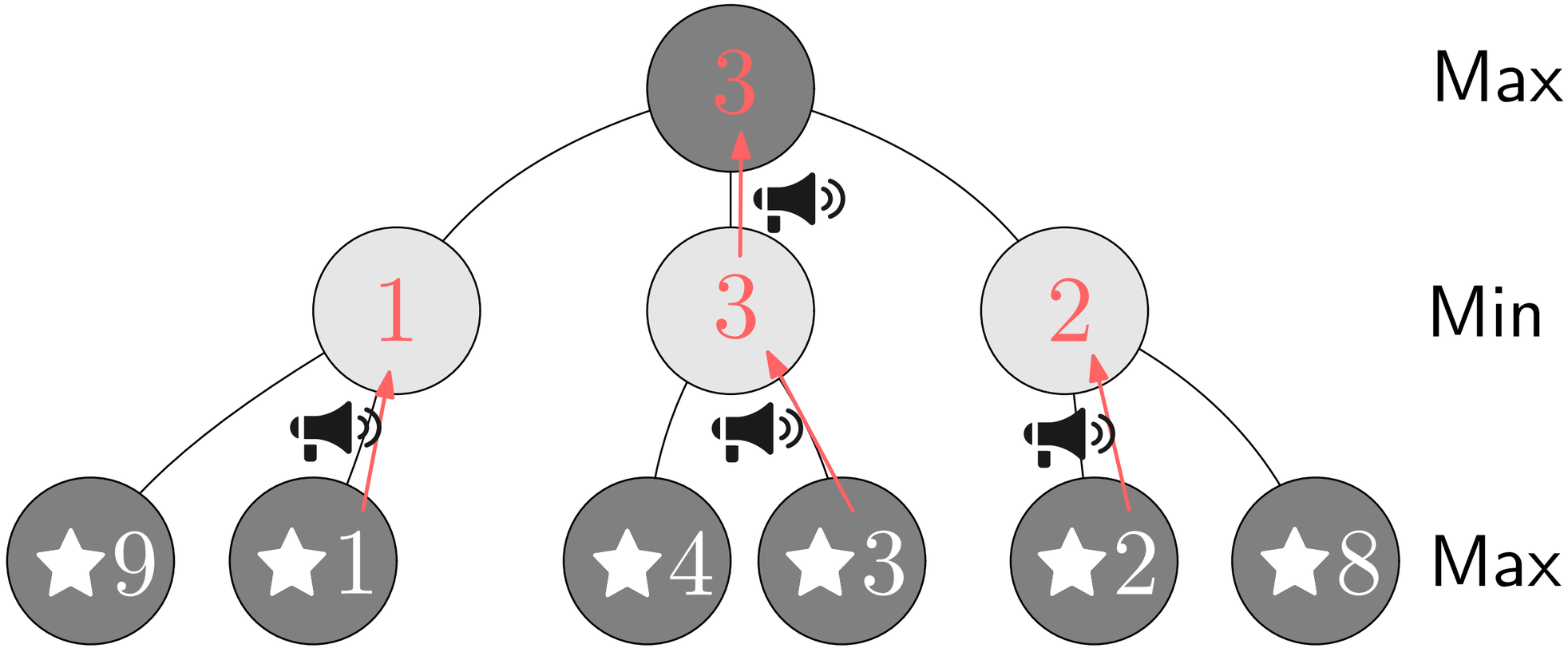
Max

Min Max



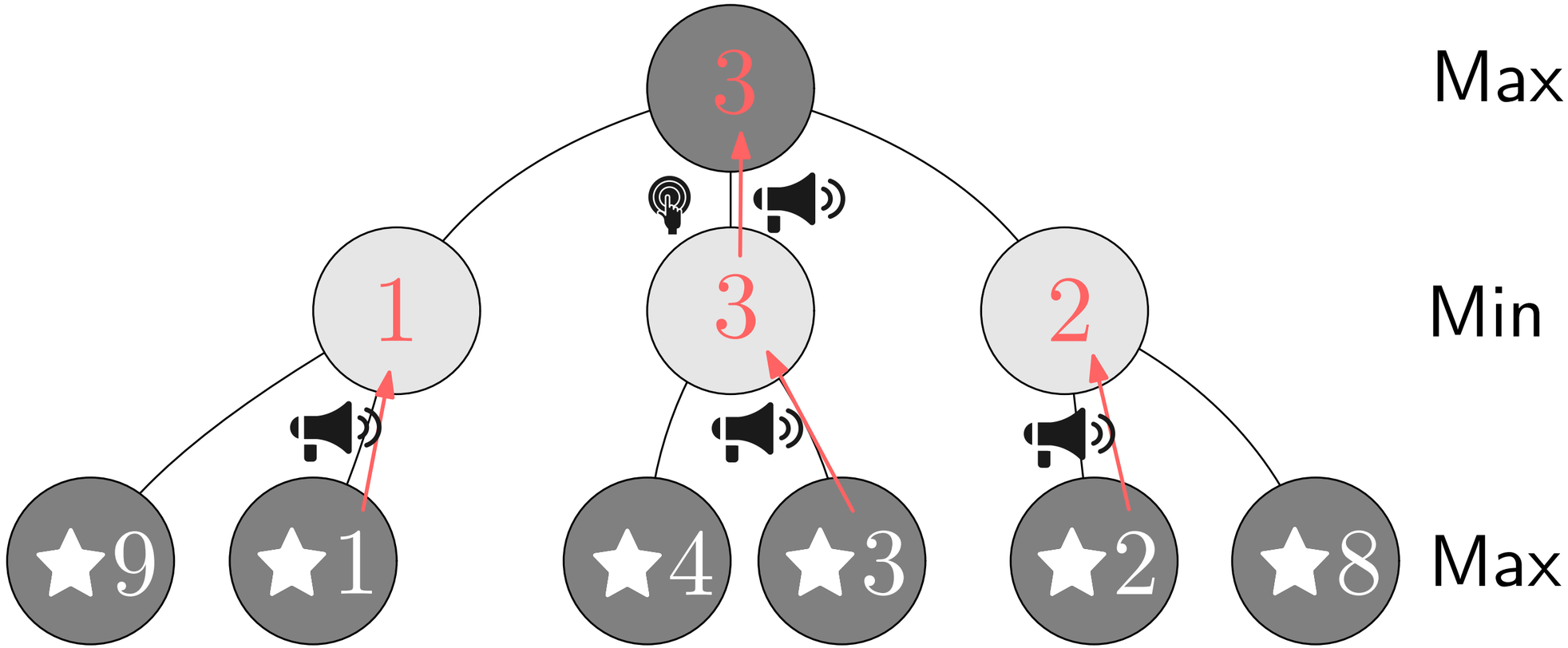
Min Max

Backpropagation



Min Max

Backpropagation



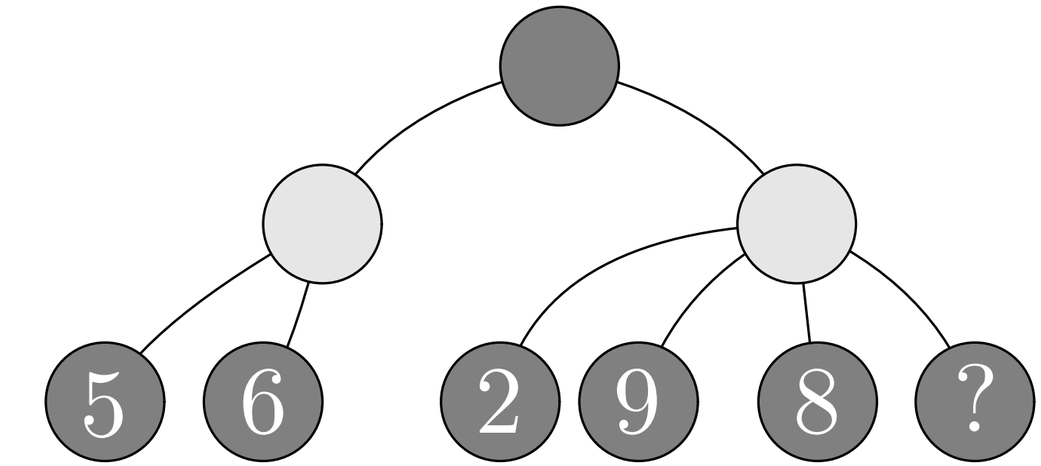
Assignment 4

Write Mini Max recursive Pseudocode

```
function minimax(state depth player)
```

depth means how deep you want to go.

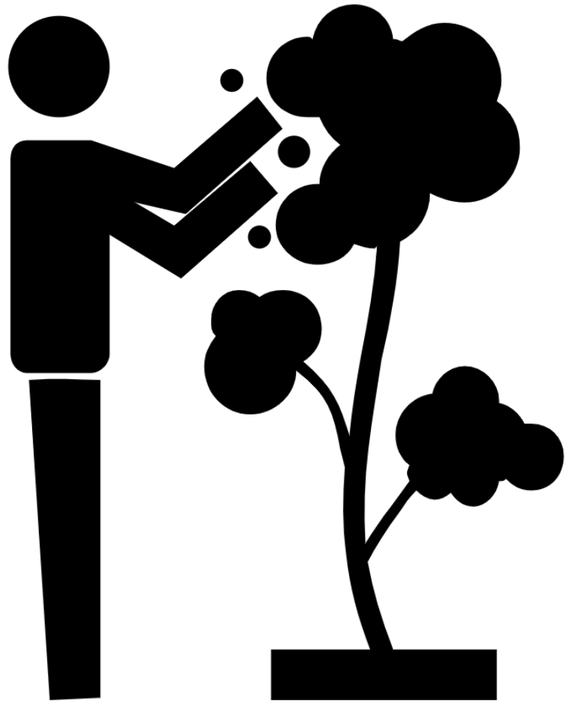
You can access all children



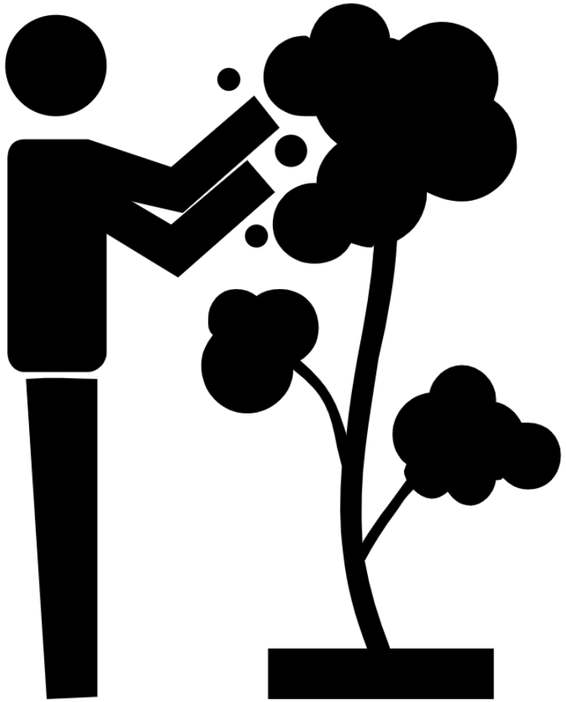
Evaluate

Observations?

Post on Teams



Alpha-Beta Pruning



Don't look, if you don't have to!

Alpha-Beta Pruning

Monte Carlo Tree Search

Selection

Expansion

Simulation

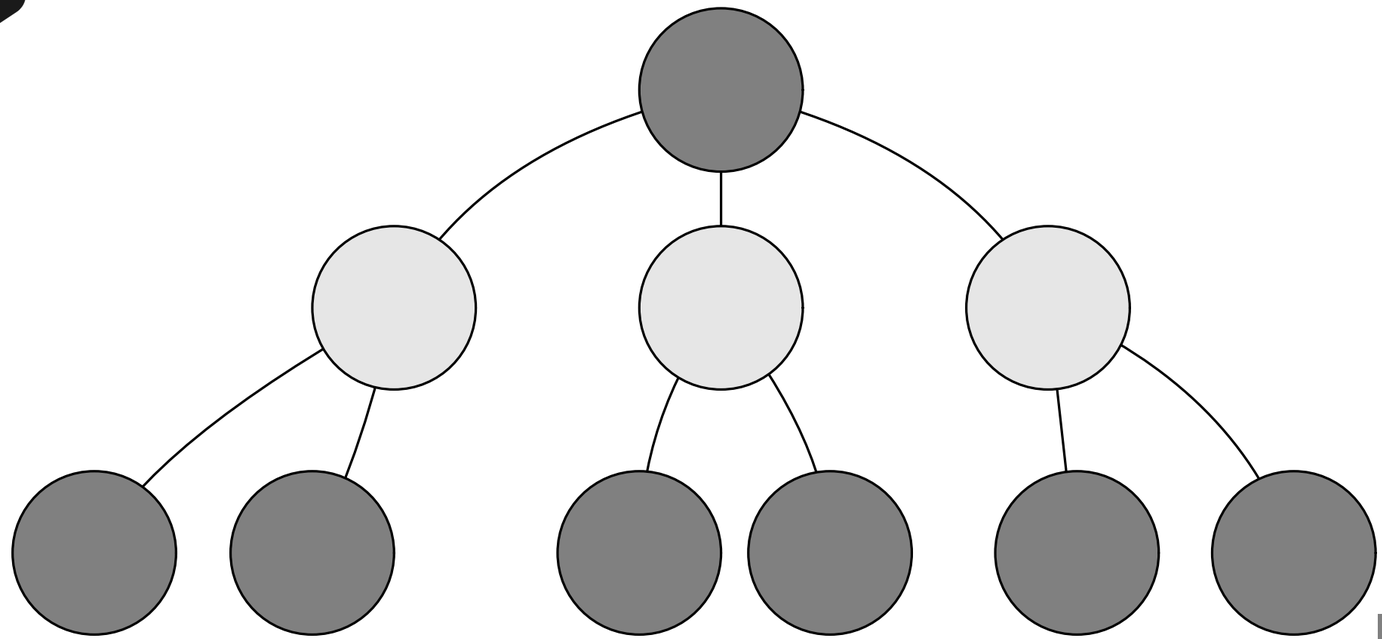
Backpropagation

Selection

Expansion

Simulation

Backpropagation

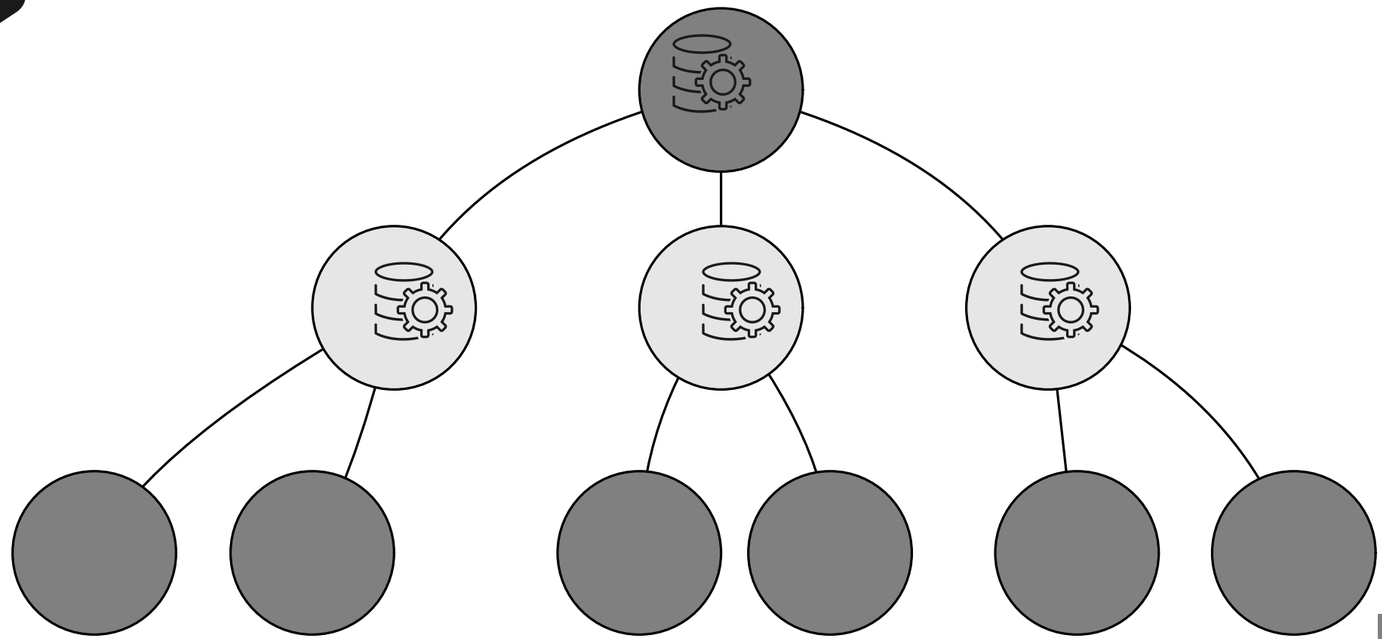


Selection

Expansion

Simulation

Backpropagation

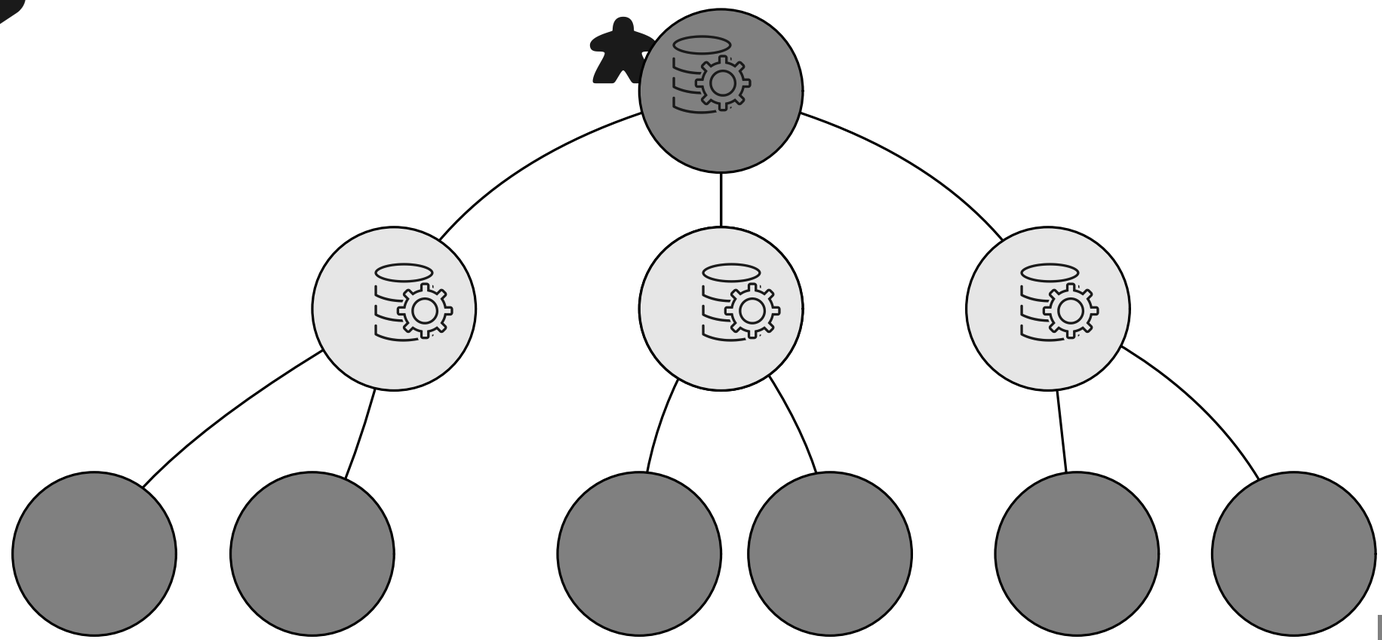


Selection

Expansion

Simulation

Backpropagation

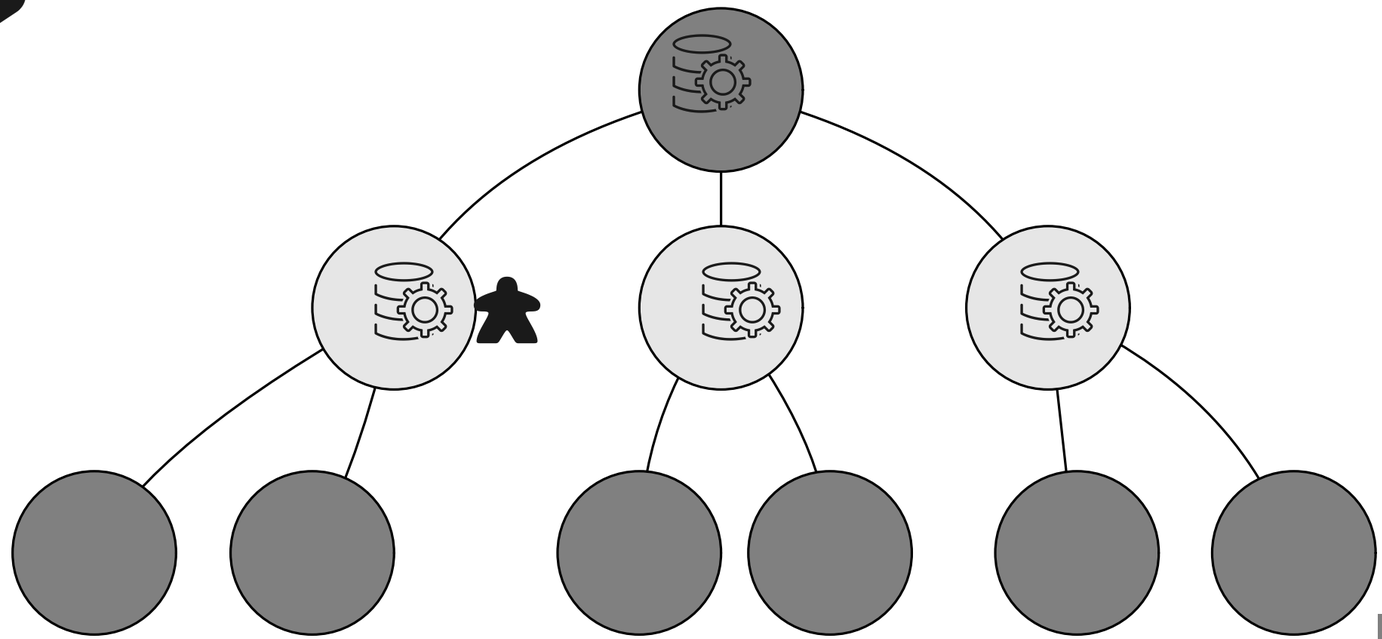


Selection

Expansion

Simulation

Backpropagation

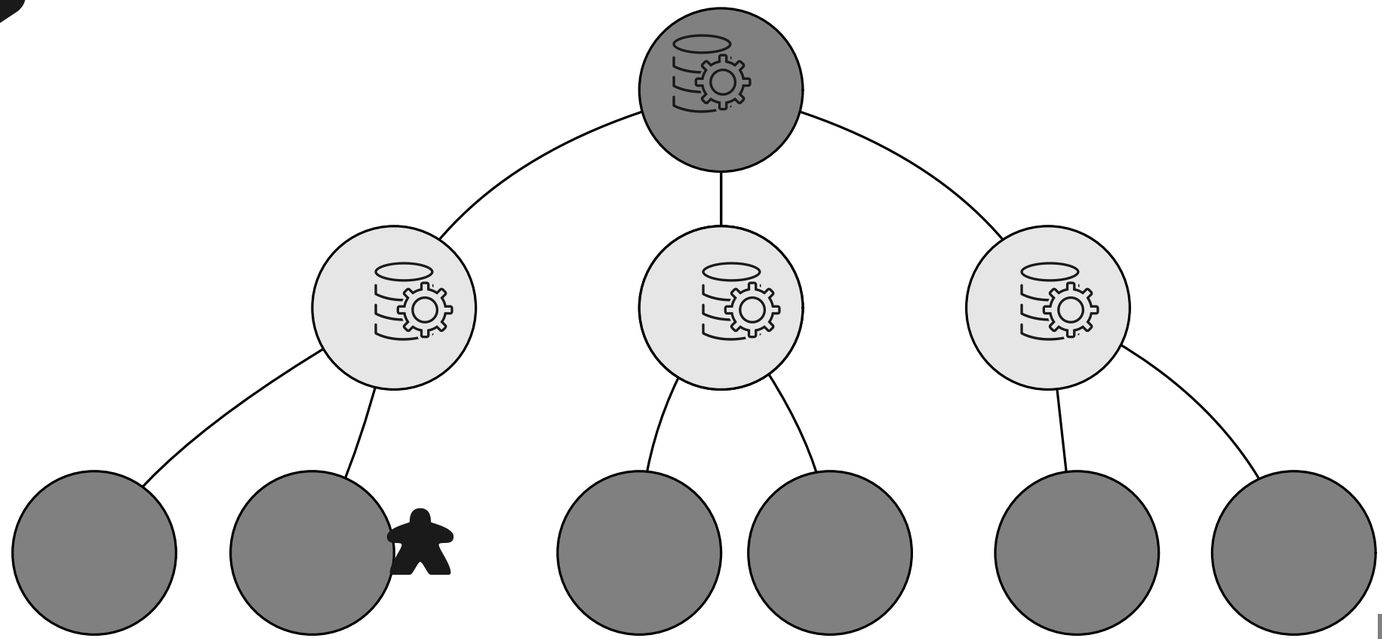


Selection

Expansion

Simulation

Backpropagation

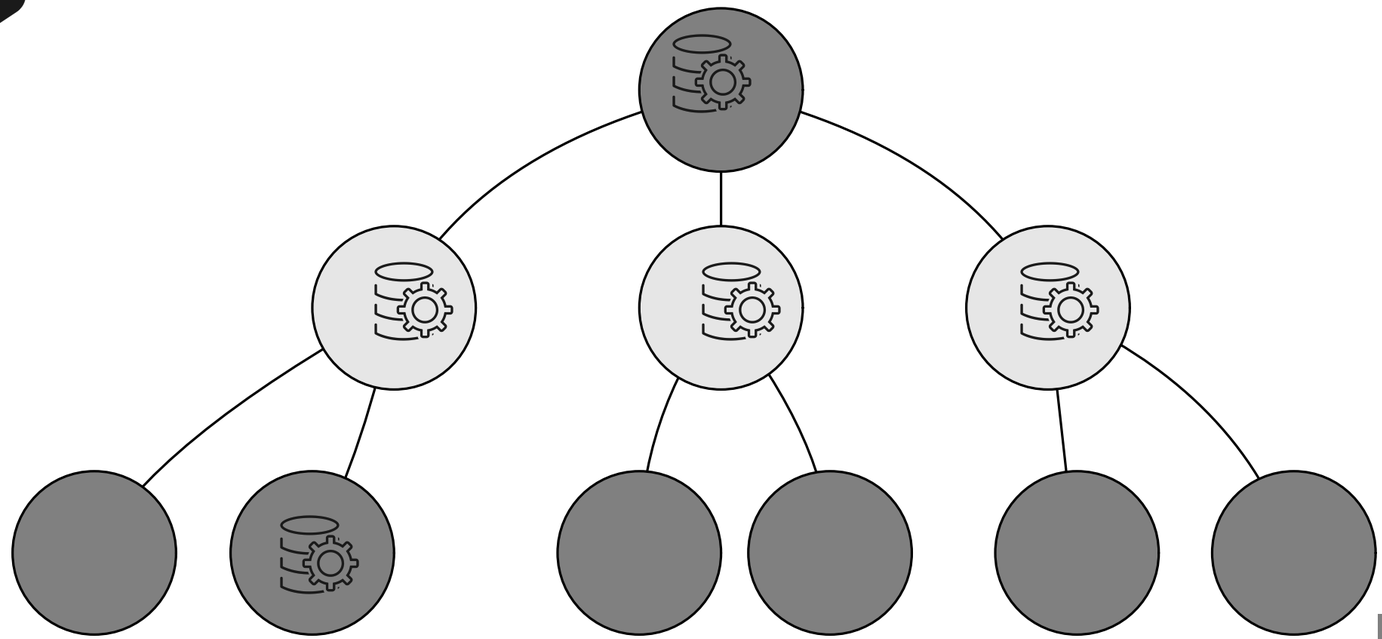


Selection

Expansion

Simulation

Backpropagation

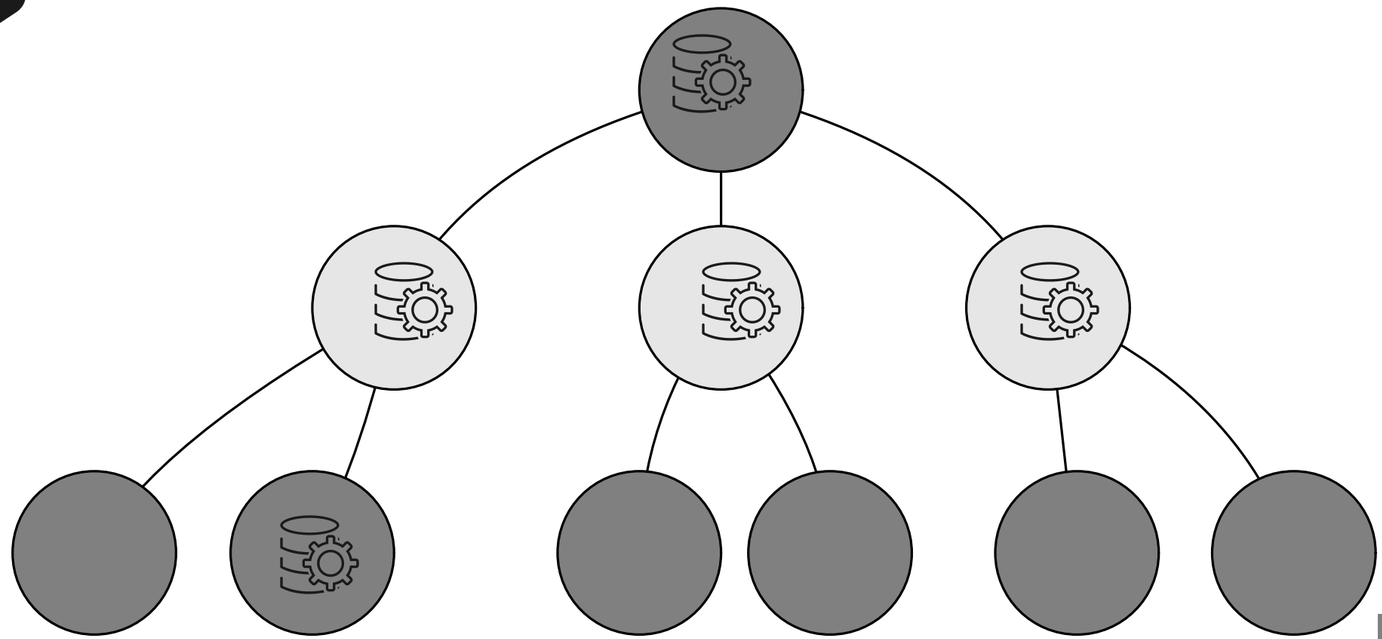


Selection

Expansion

Simulation

Backpropagation

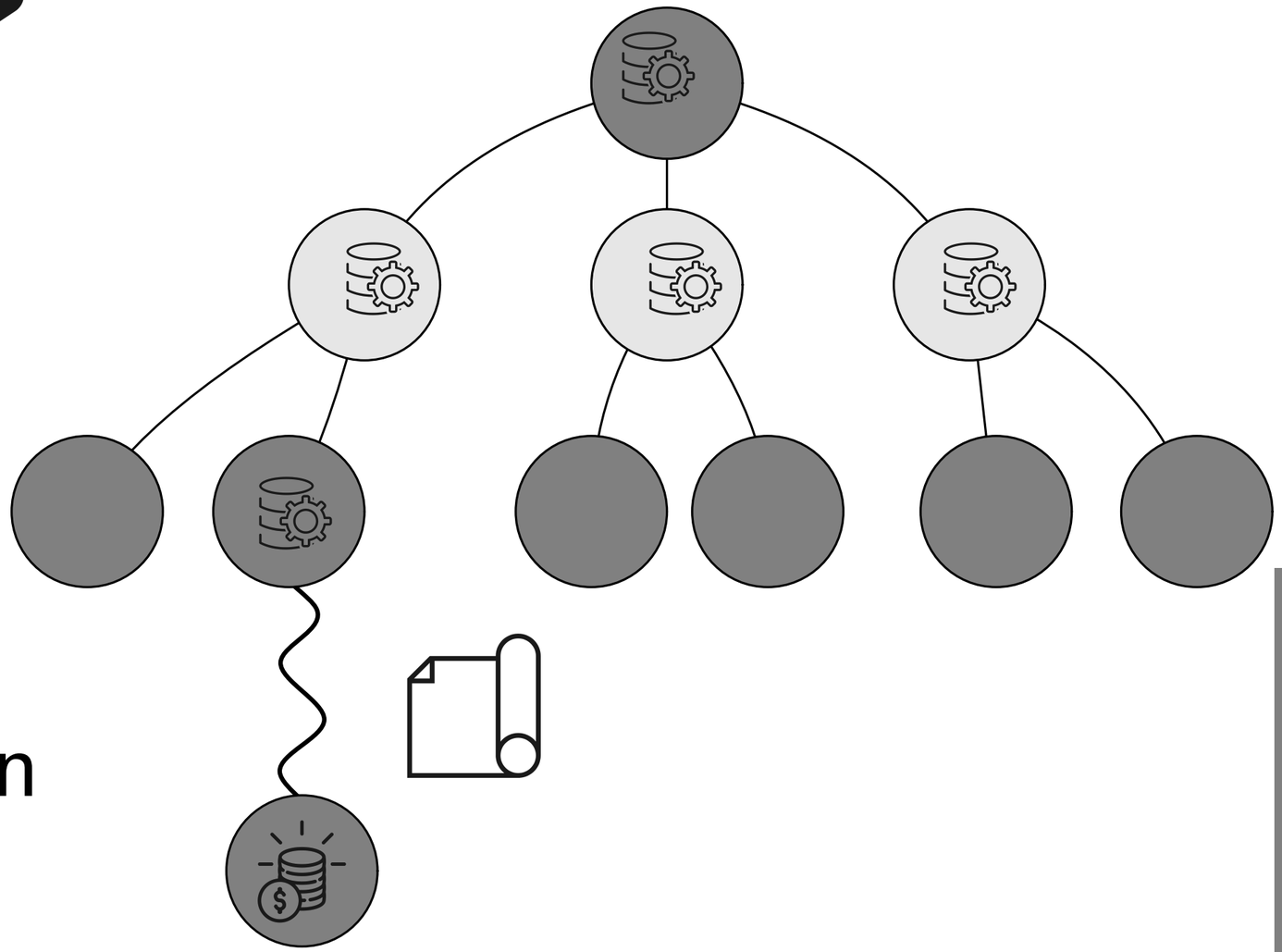


Selection

Expansion

Simulation

Backpropagation

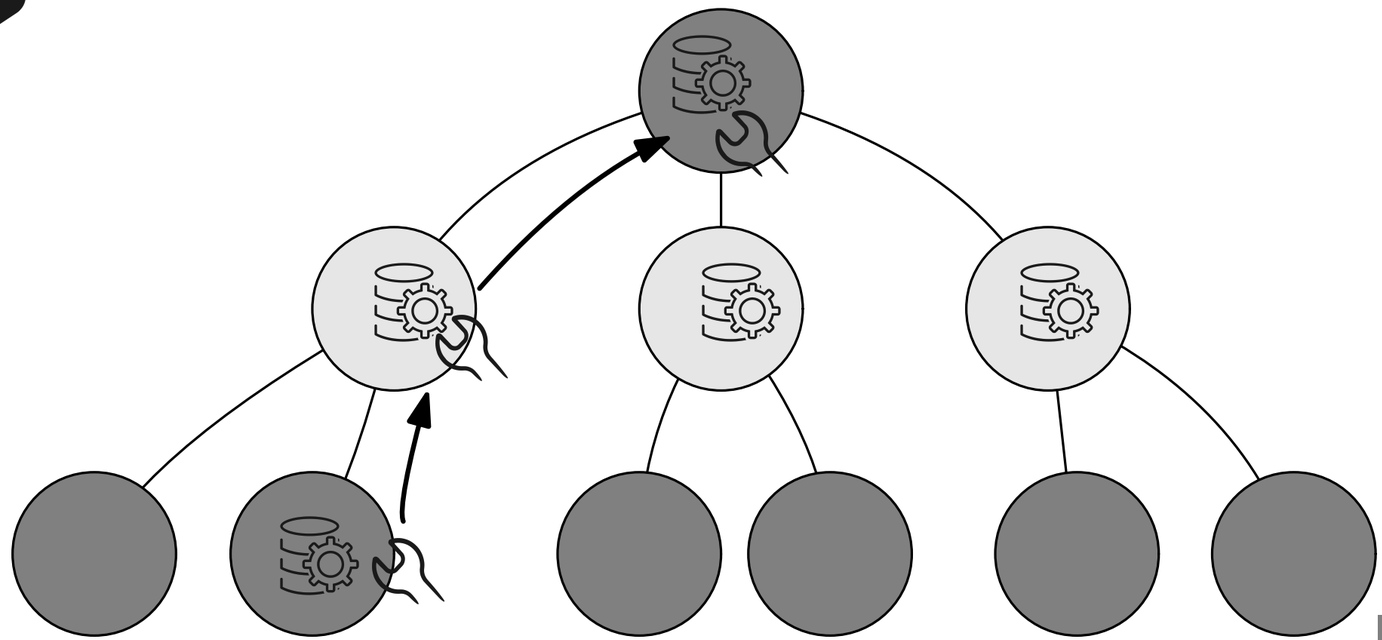


Selection

Expansion

Simulation

Backpropagation

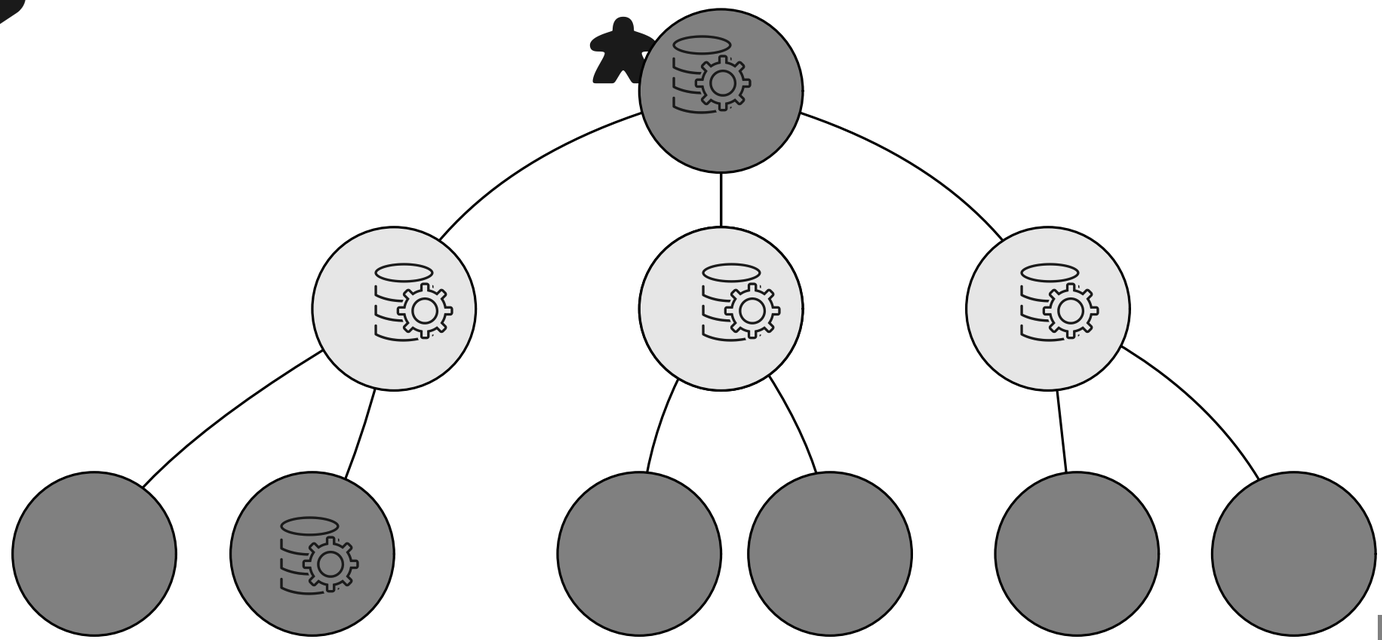


Selection

Expansion

Simulation

Backpropagation

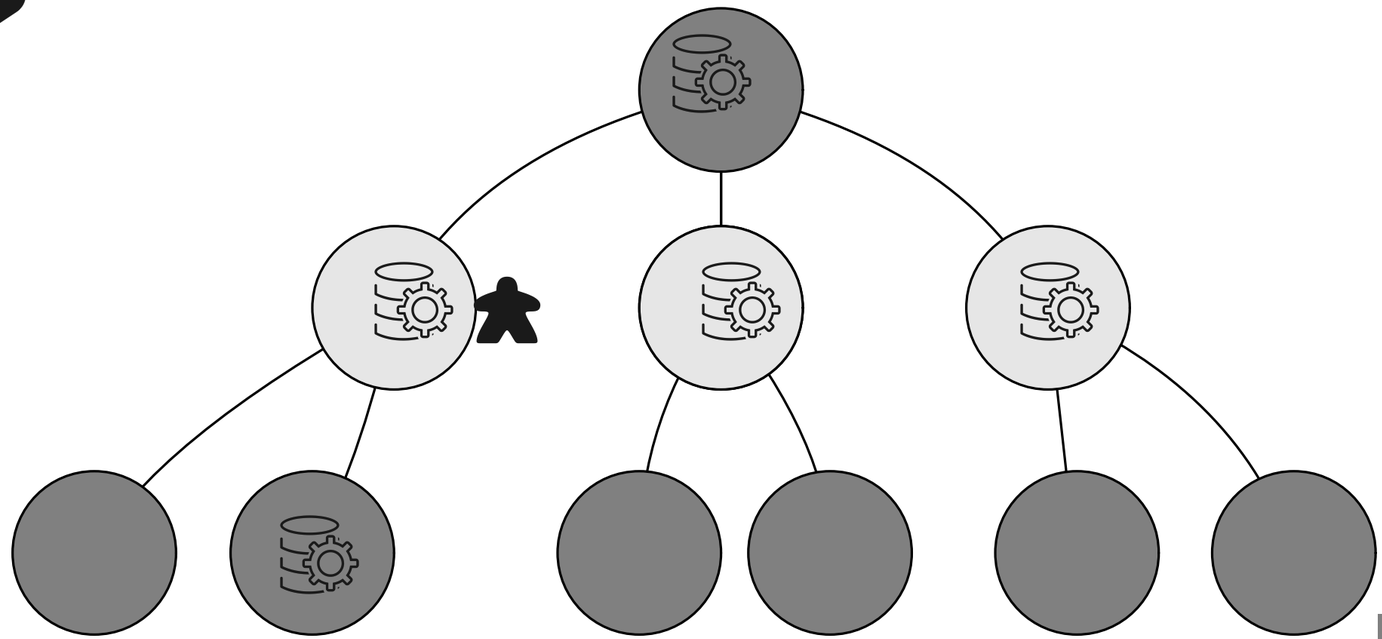


Selection

Expansion

Simulation

Backpropagation

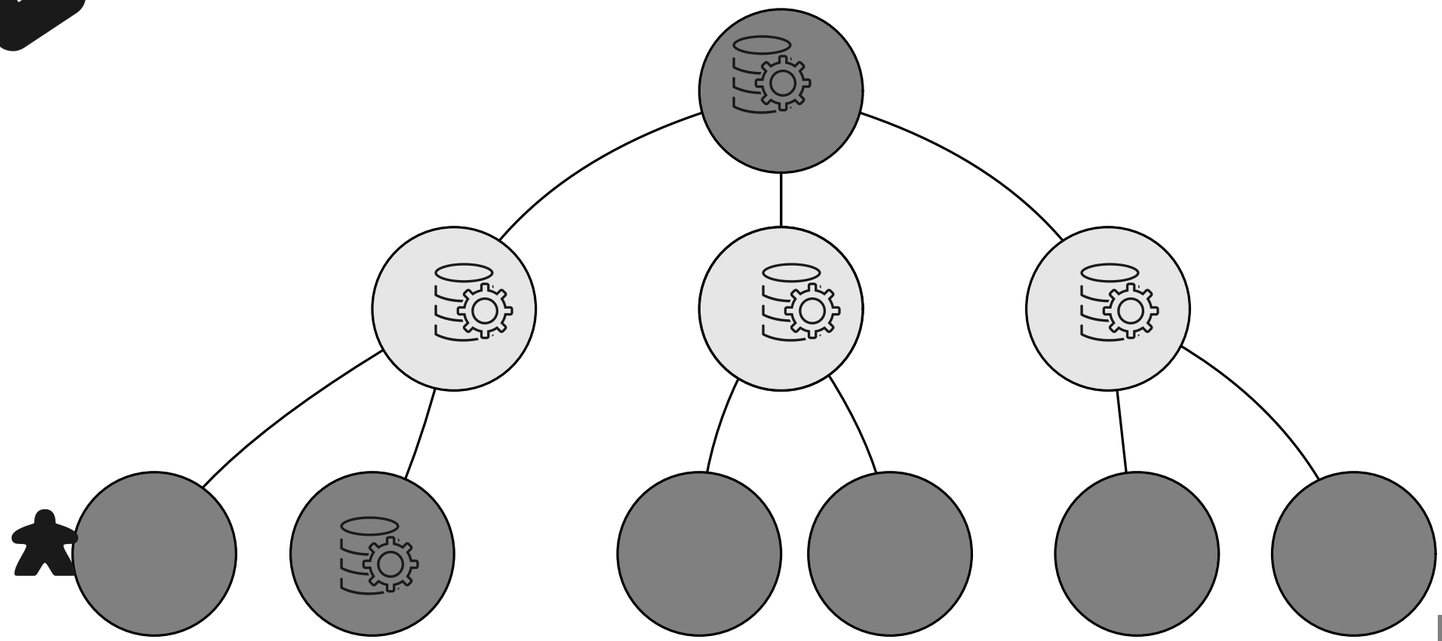


Selection

Expansion

Simulation

Backpropagation

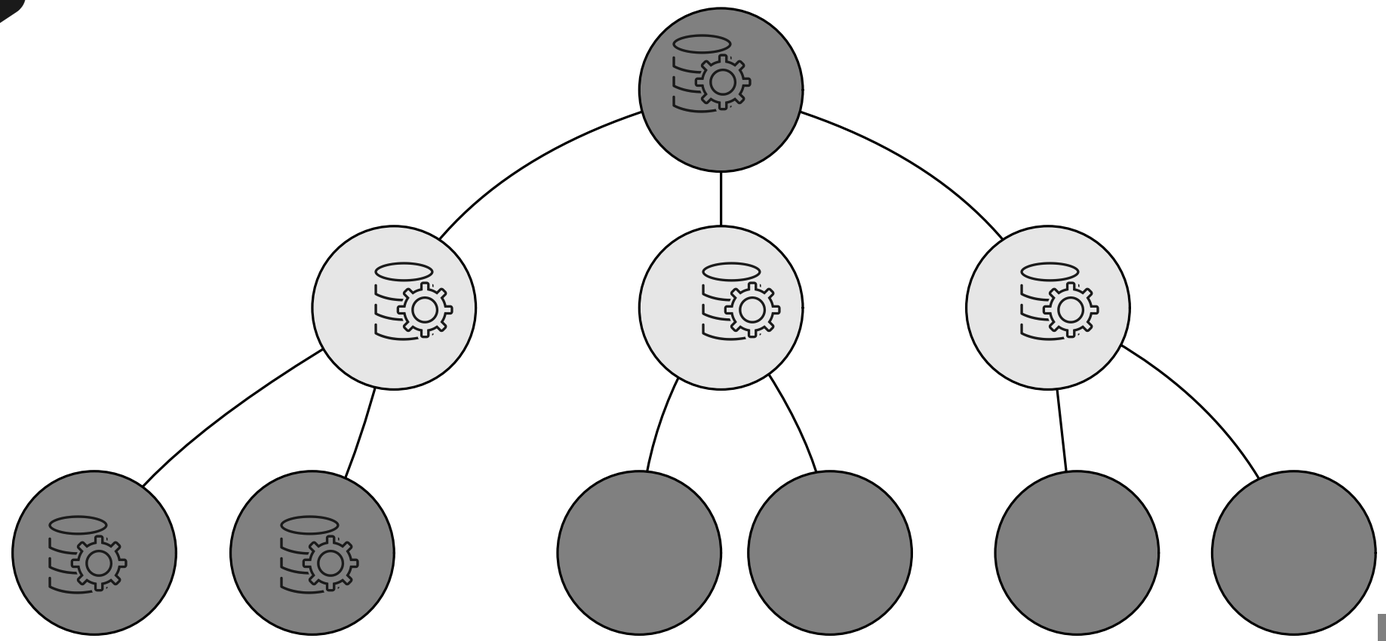


Selection

Expansion

Simulation

Backpropagation

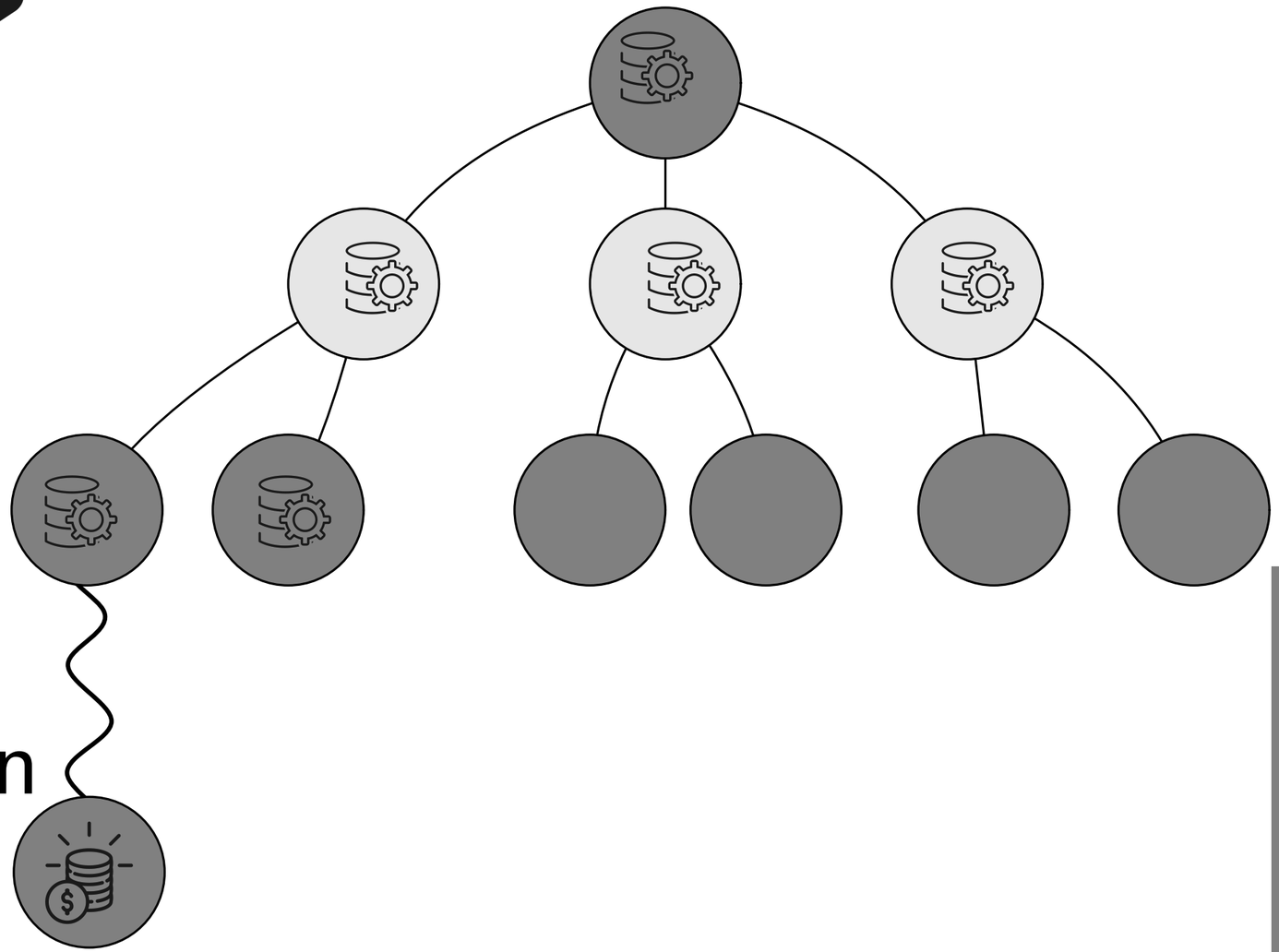


Selection

Expansion

Simulation

Backpropagation

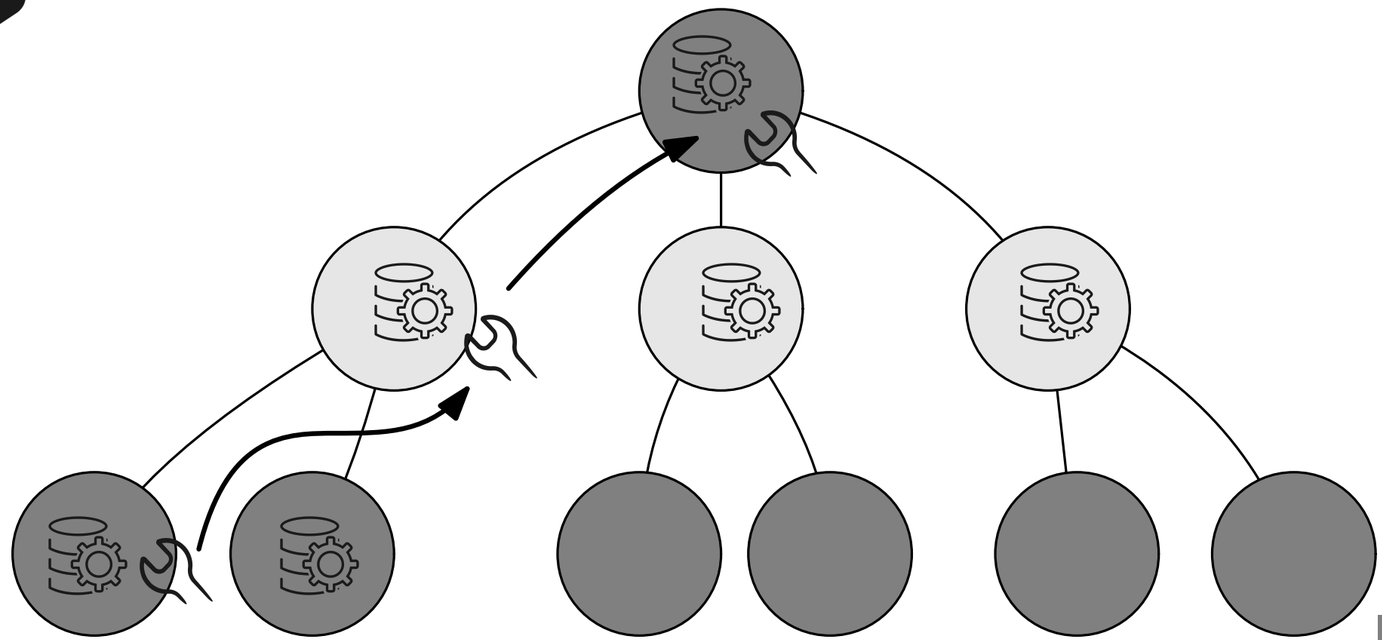


Selection

Expansion

Simulation

Backpropagation



Selection

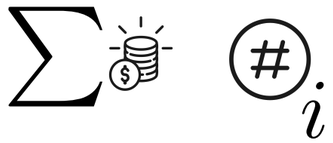
Selection UCB ε-greedy

Selection UCB ε-greedy

Selection

UCB

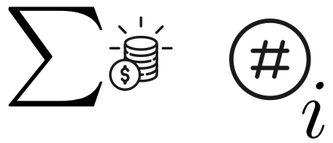
ε-greedy



Selection

UCB

ε-greedy

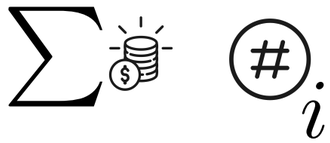


Simulation

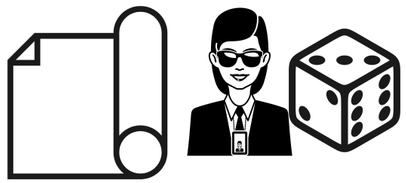
Selection

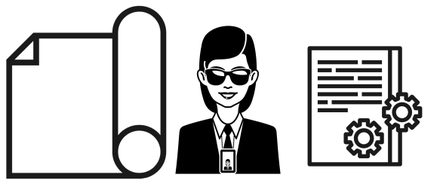
UCB

ε-greedy

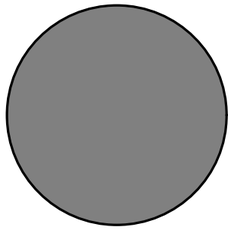


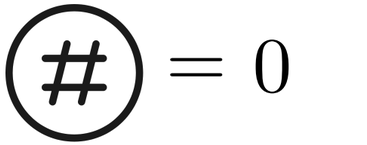
Simulation

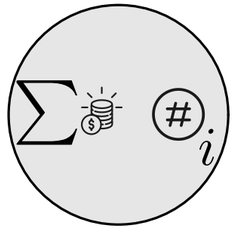




Selection
Expansion
Simulation
Backpropagation





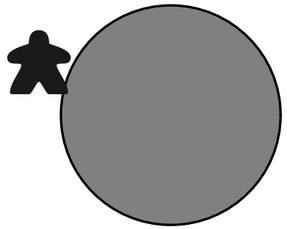


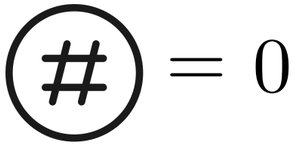
Selection

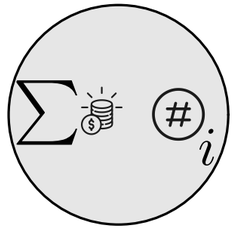
Expansion

Simulation

Backpropagation





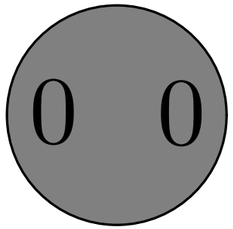


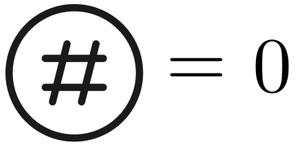
Selection

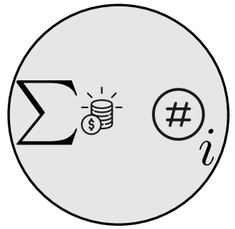
Expansion

Simulation

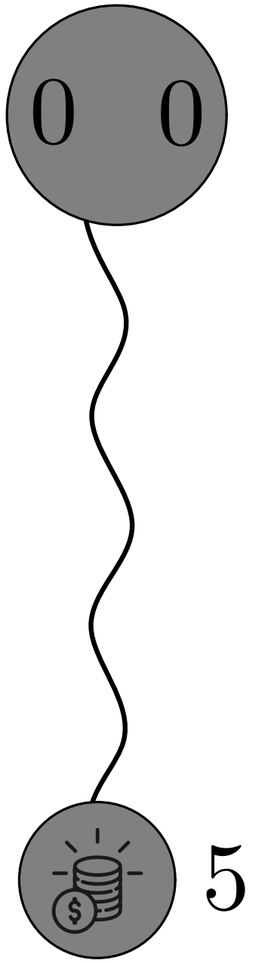
Backpropagation

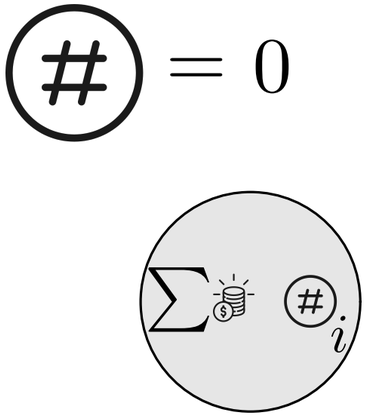




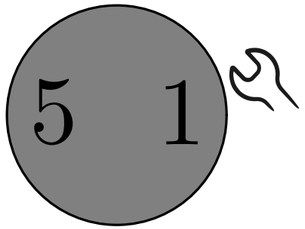


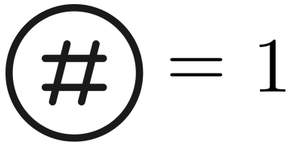
Selection
Expansion
Simulation
Backpropagation

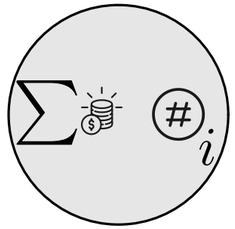




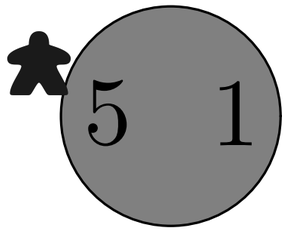
Selection
Expansion
Simulation
Backpropagation

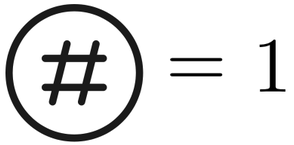


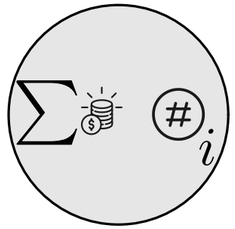




Selection
Expansion
Simulation
Backpropagation





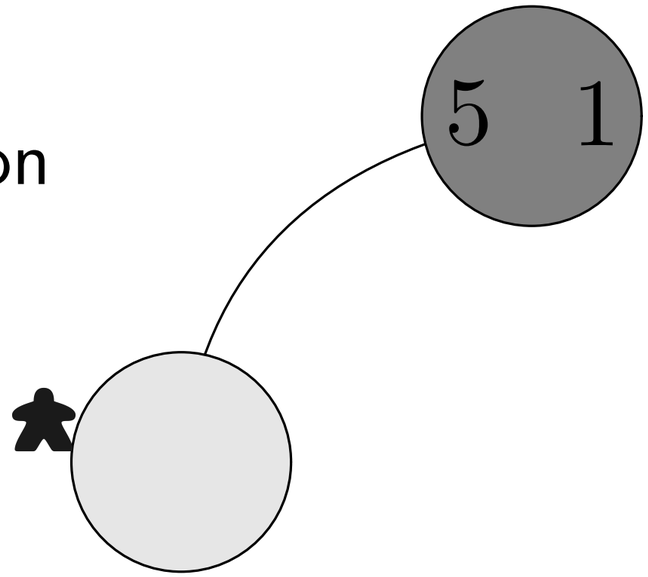


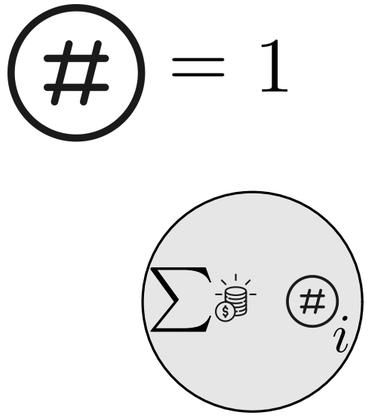
Selection

Expansion

Simulation

Backpropagation



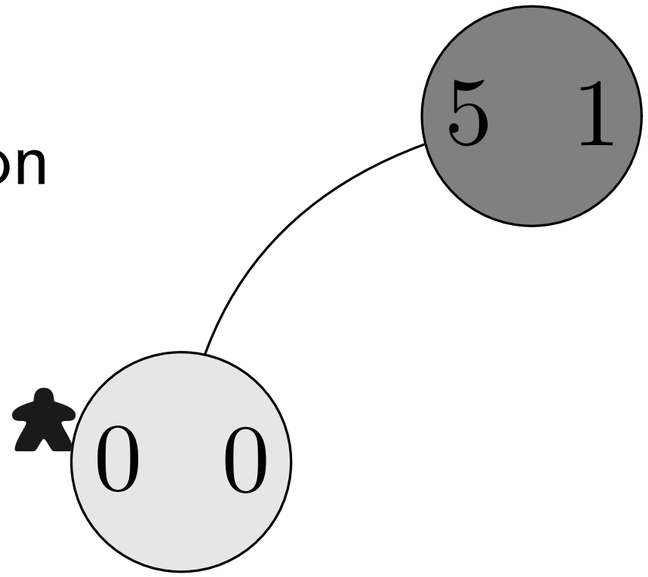


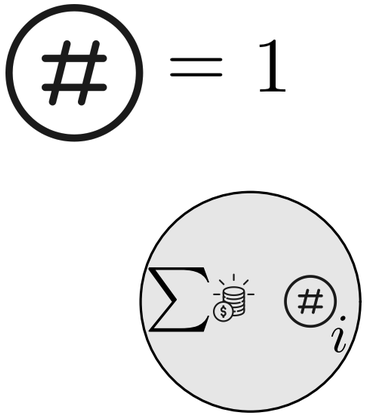
Selection

Expansion

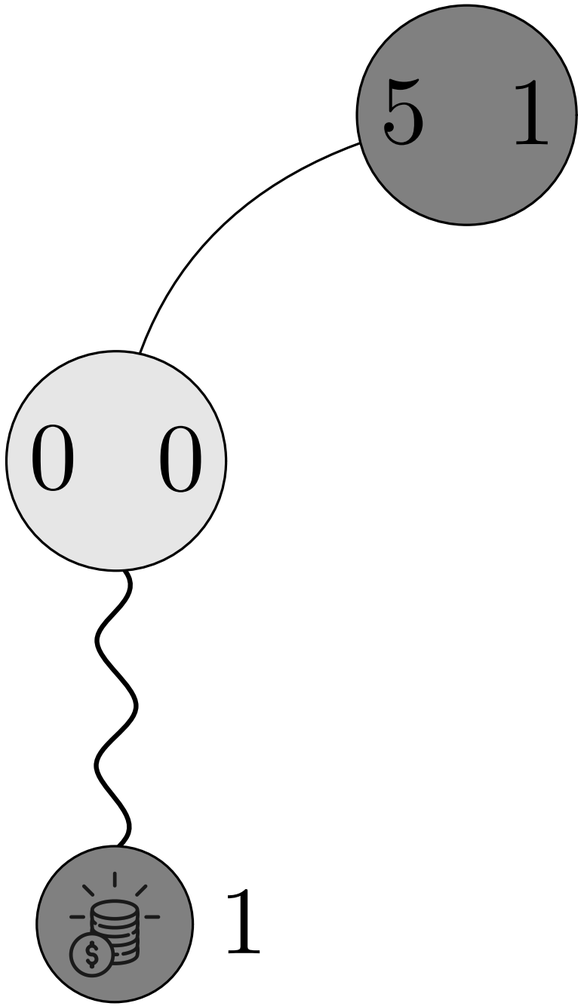
Simulation

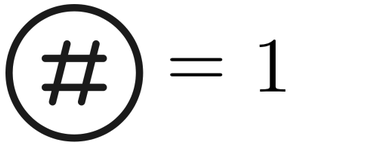
Backpropagation

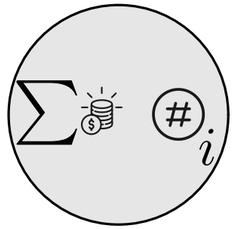




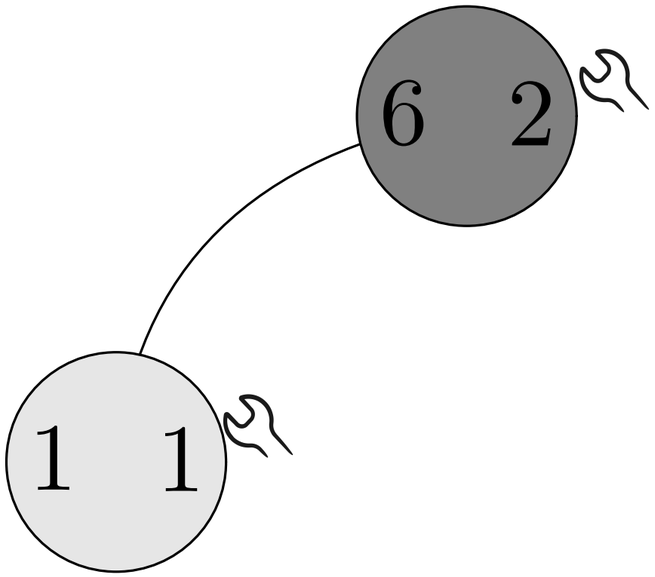
Selection
Expansion
Simulation
Backpropagation

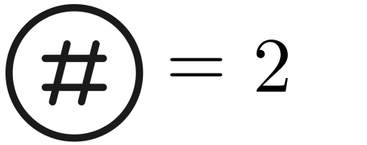


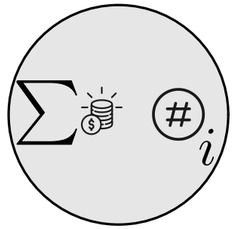




Selection
Expansion
Simulation
Backpropagation





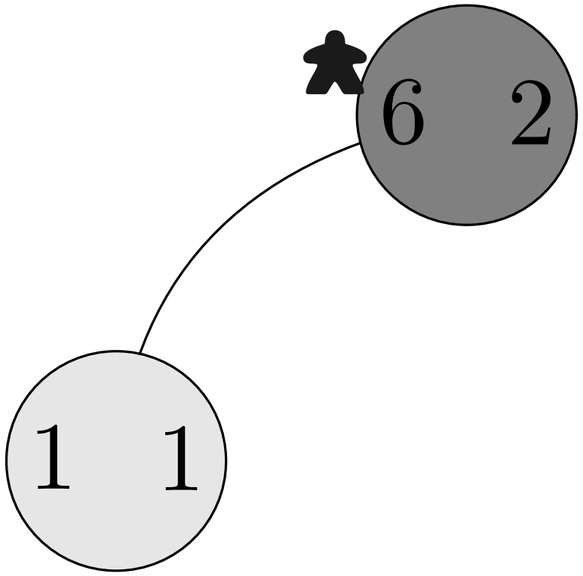


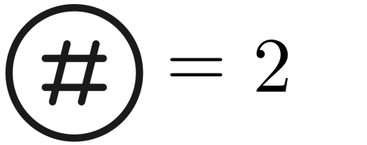
Selection

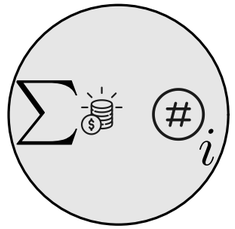
Expansion

Simulation

Backpropagation





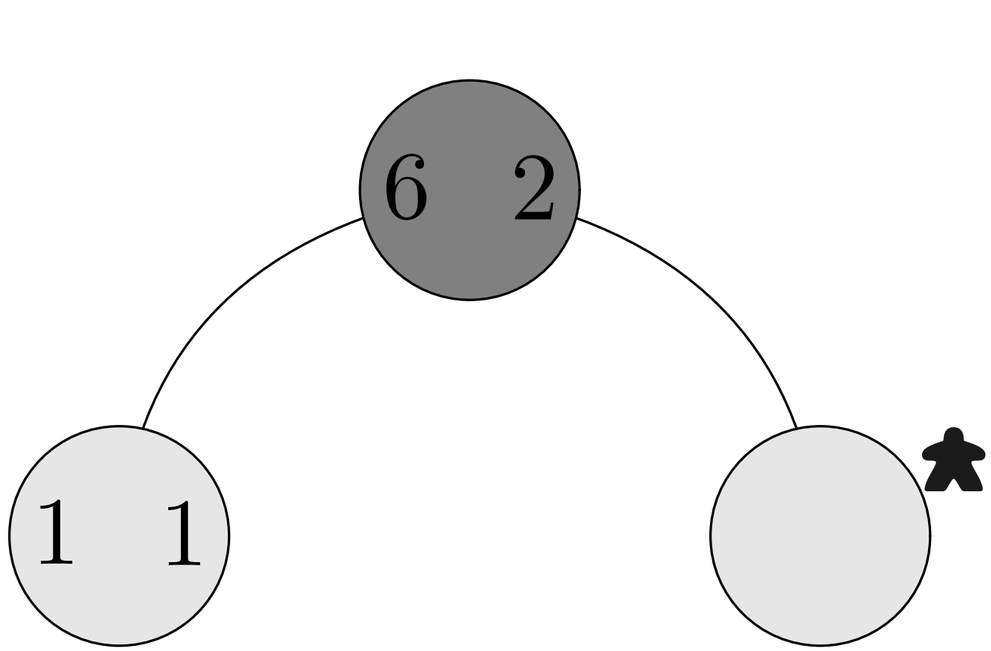


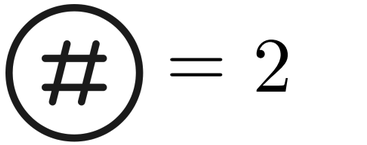
Selection

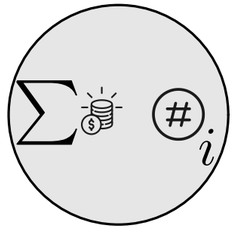
Expansion

Simulation

Backpropagation





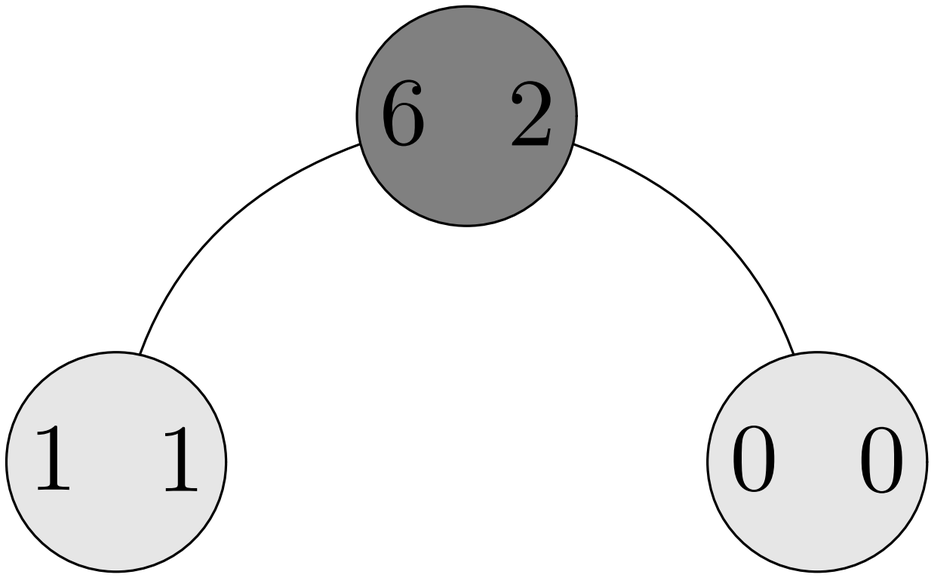


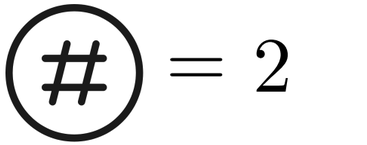
Selection

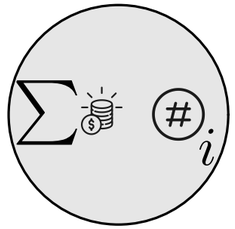
Expansion

Simulation

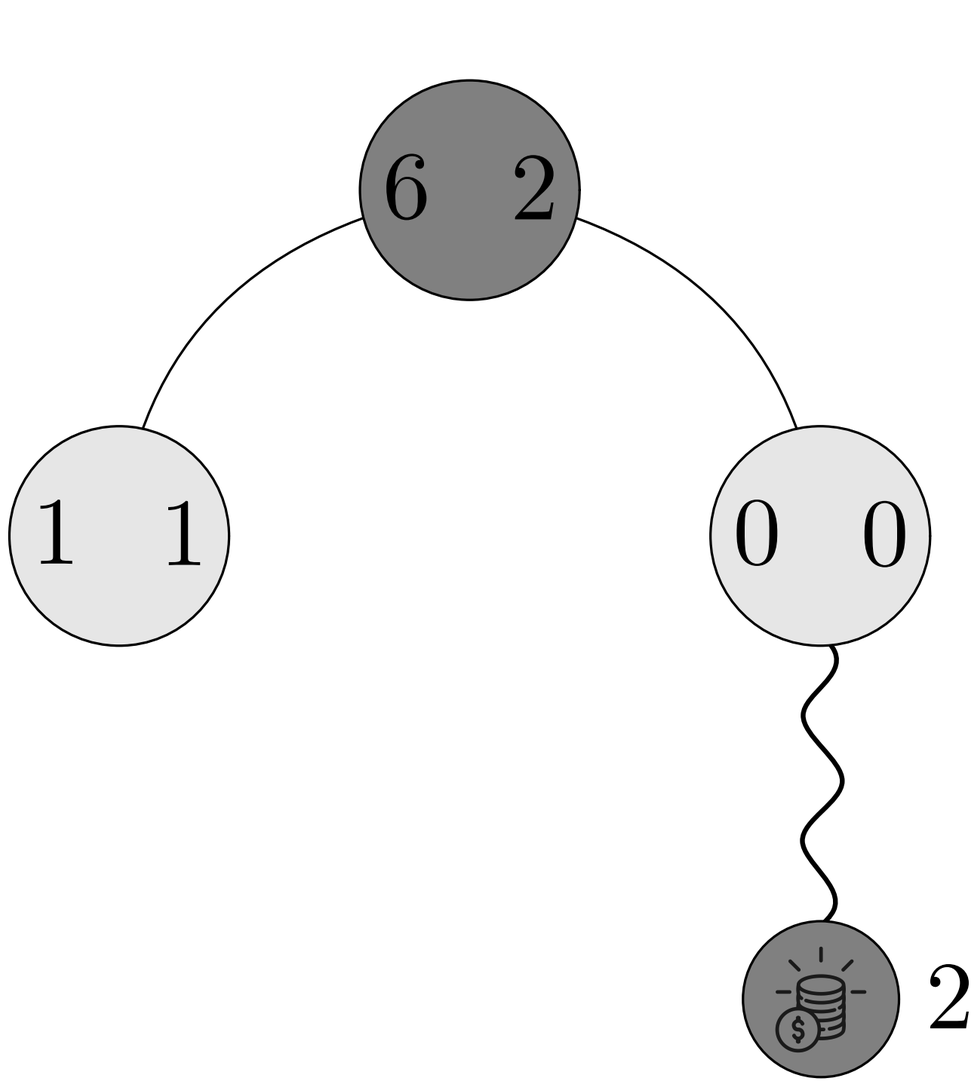
Backpropagation

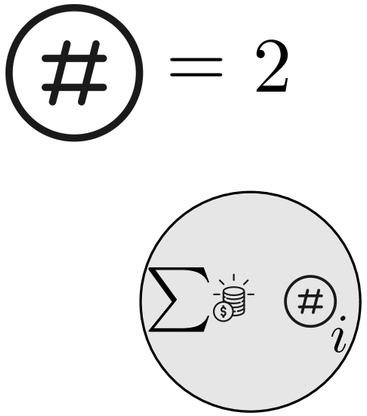




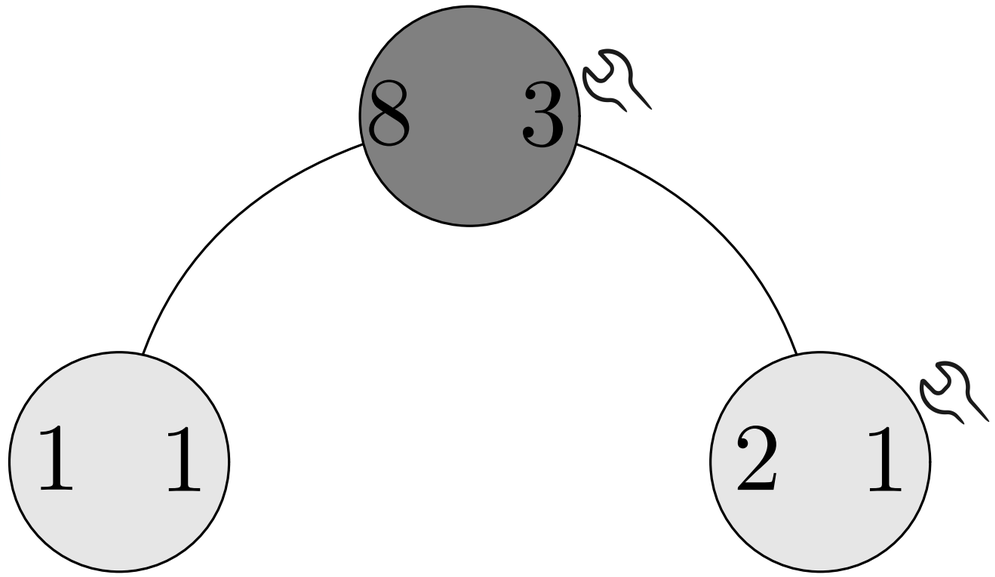


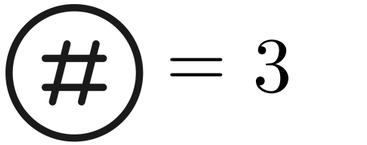
Selection
Expansion
Simulation
Backpropagation

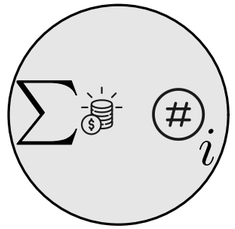




Selection
Expansion
Simulation
Backpropagation





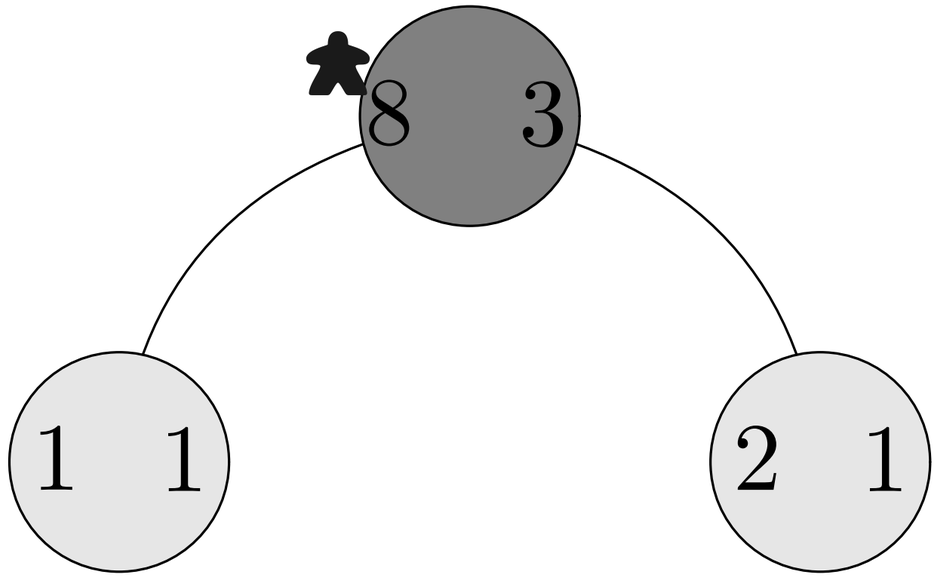


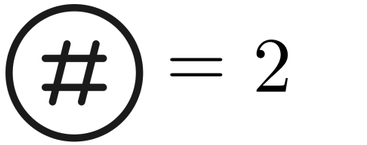
Selection

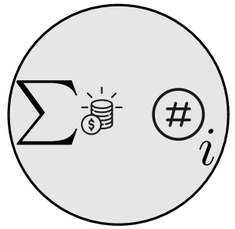
Expansion

Simulation

Backpropagation





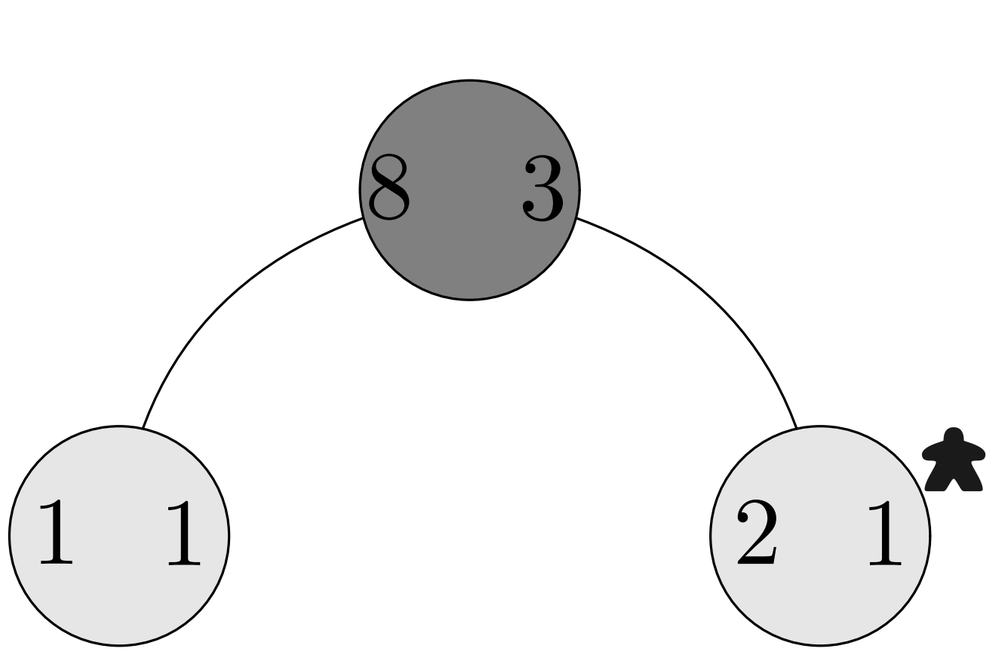


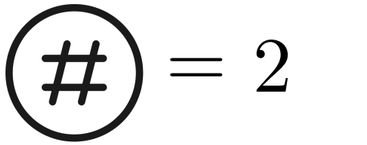
Selection

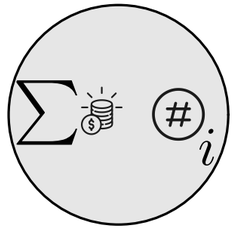
Expansion

Simulation

Backpropagation





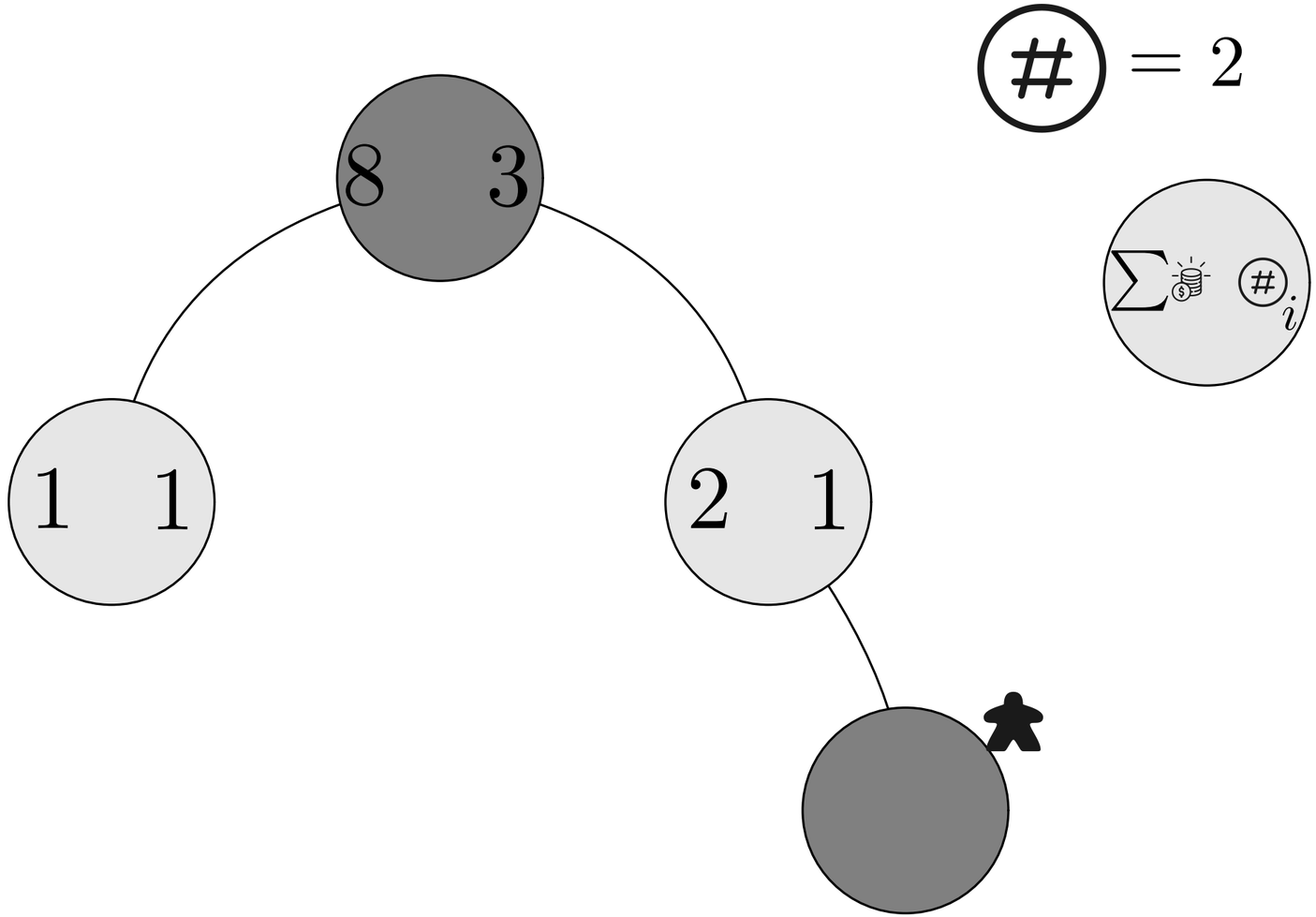


Selection

Expansion

Simulation

Backpropagation

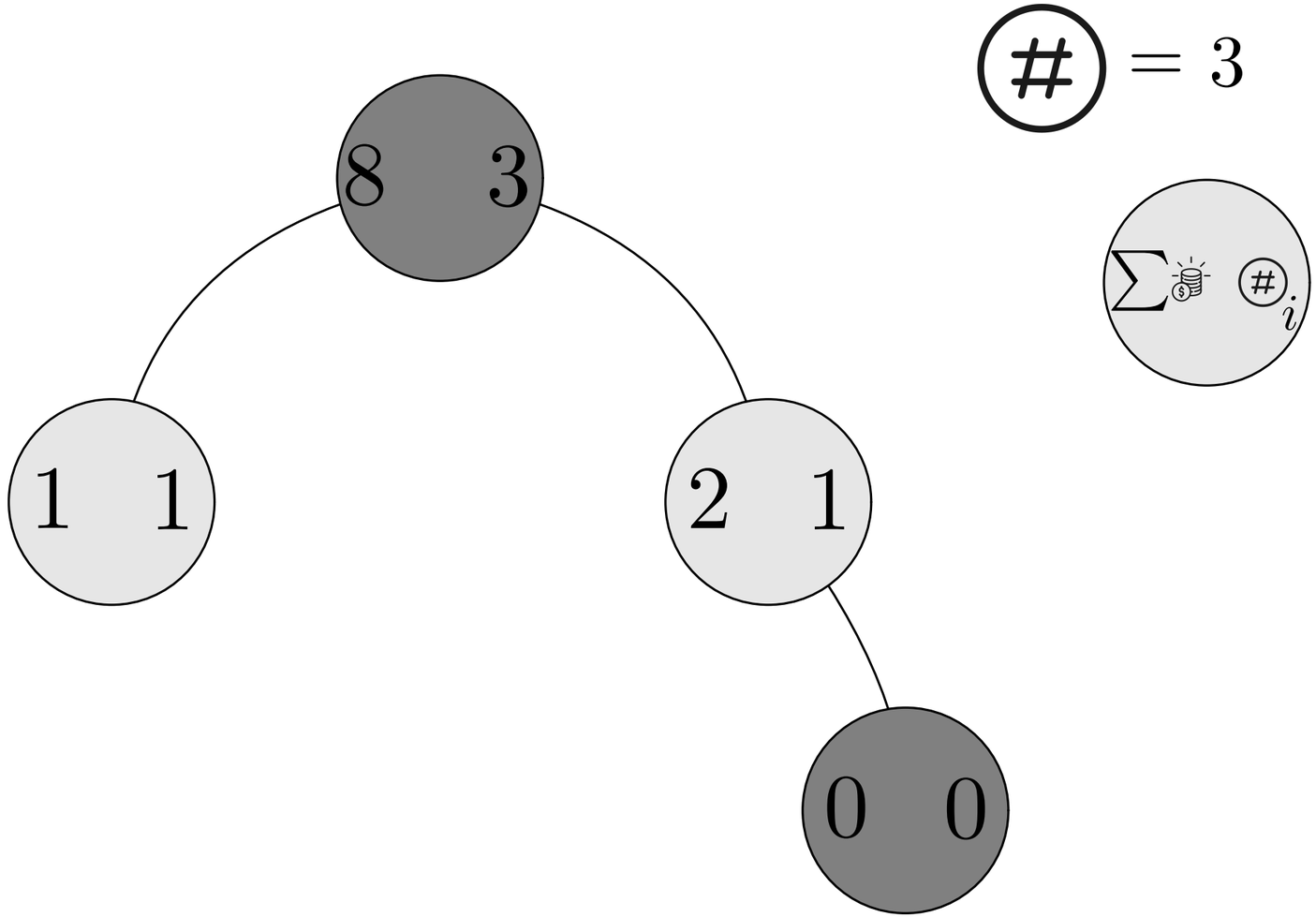


Selection

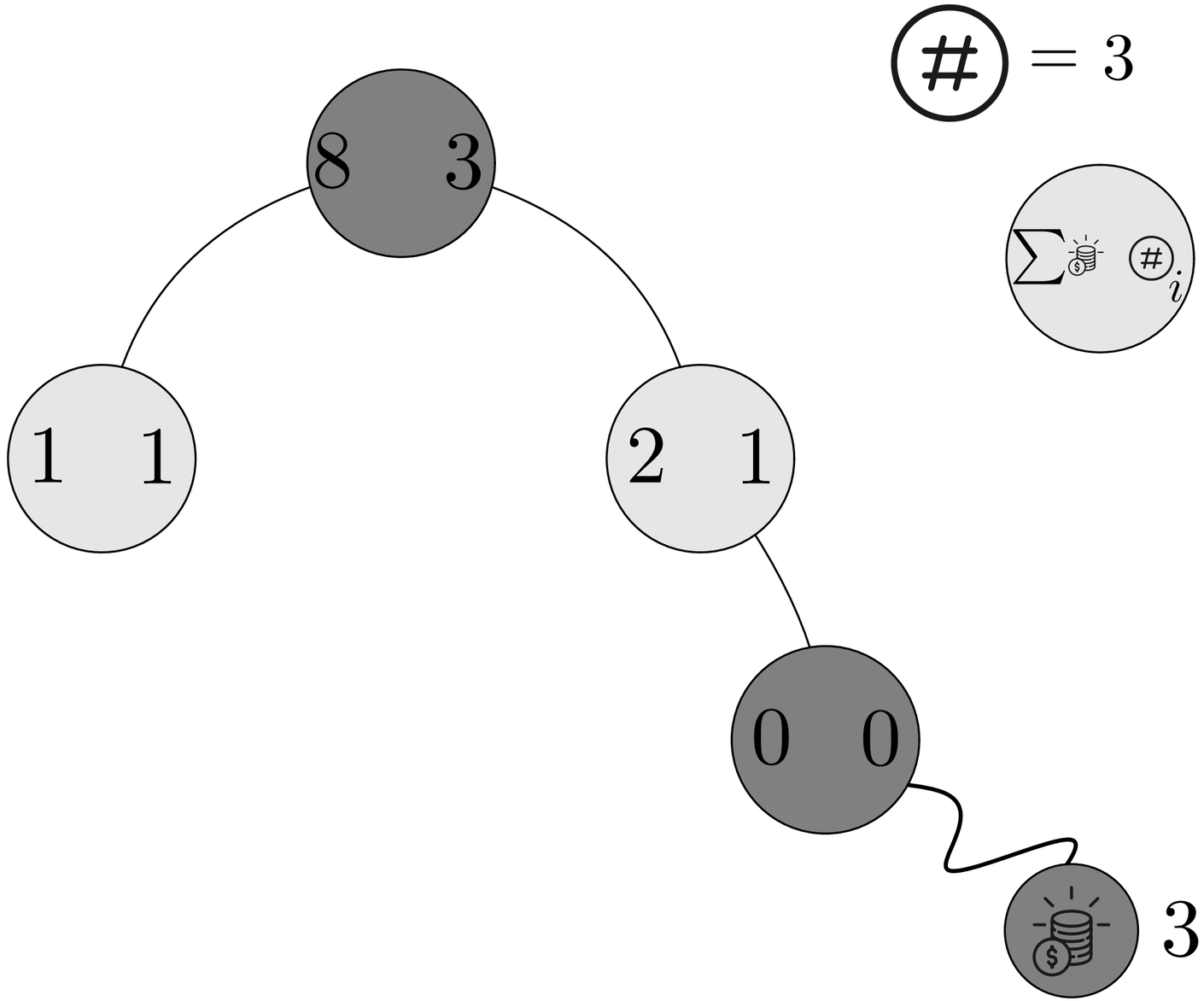
Expansion

Simulation

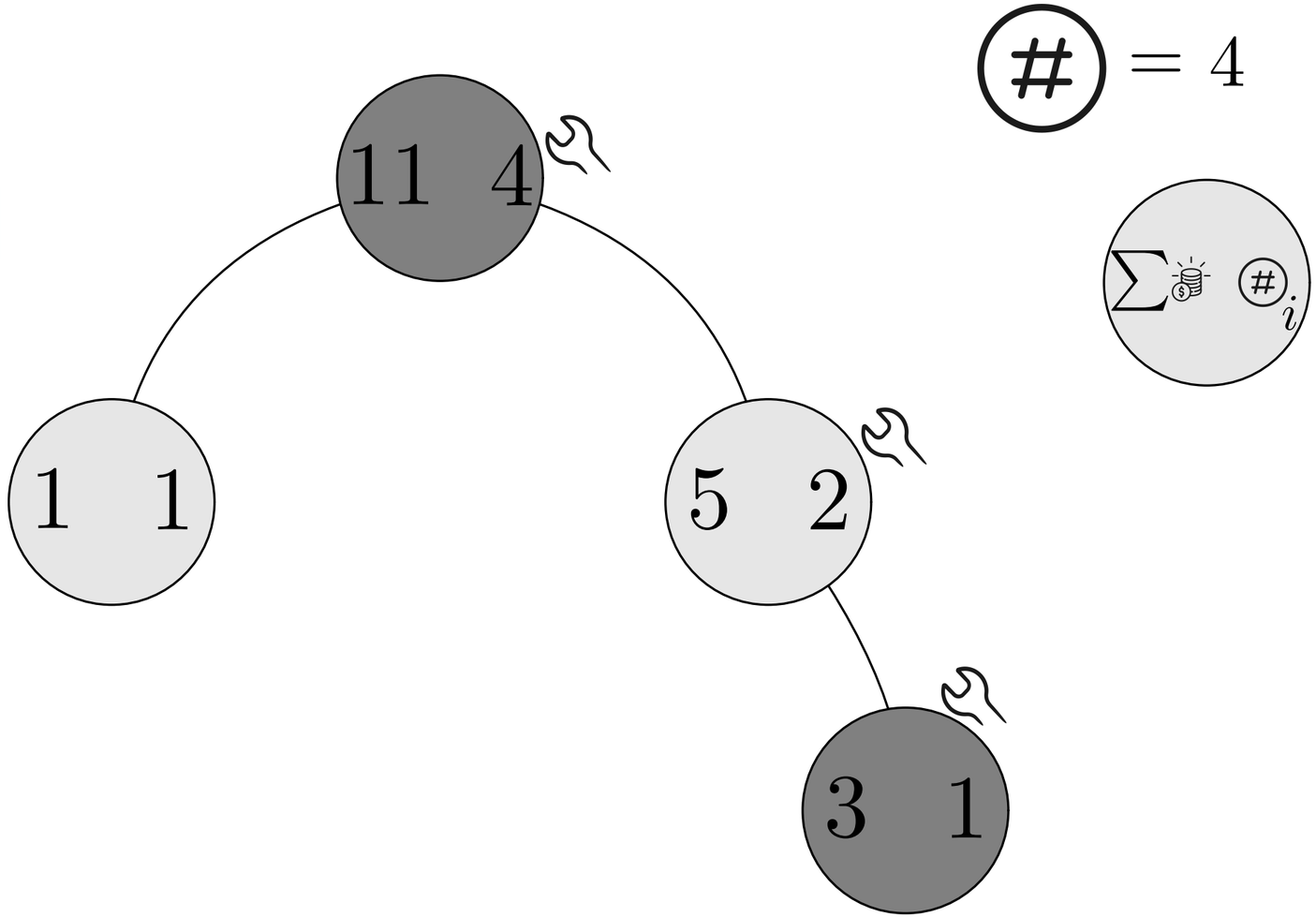
Backpropagation



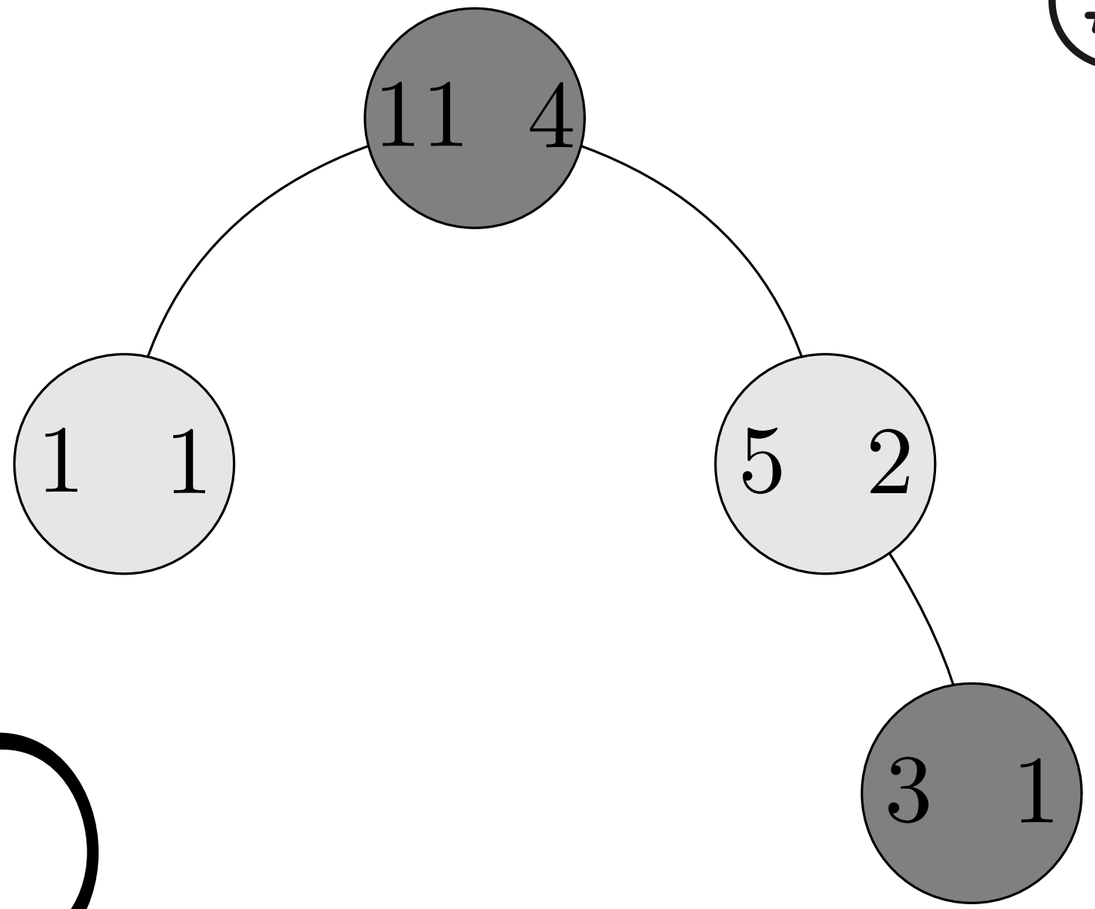
Selection
Expansion
Simulation
Backpropagation

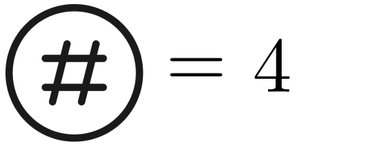


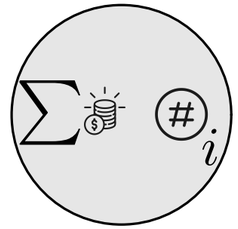
Selection
Expansion
Simulation
Backpropagation

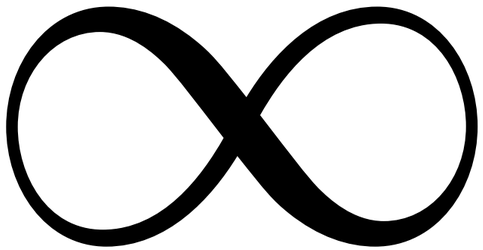


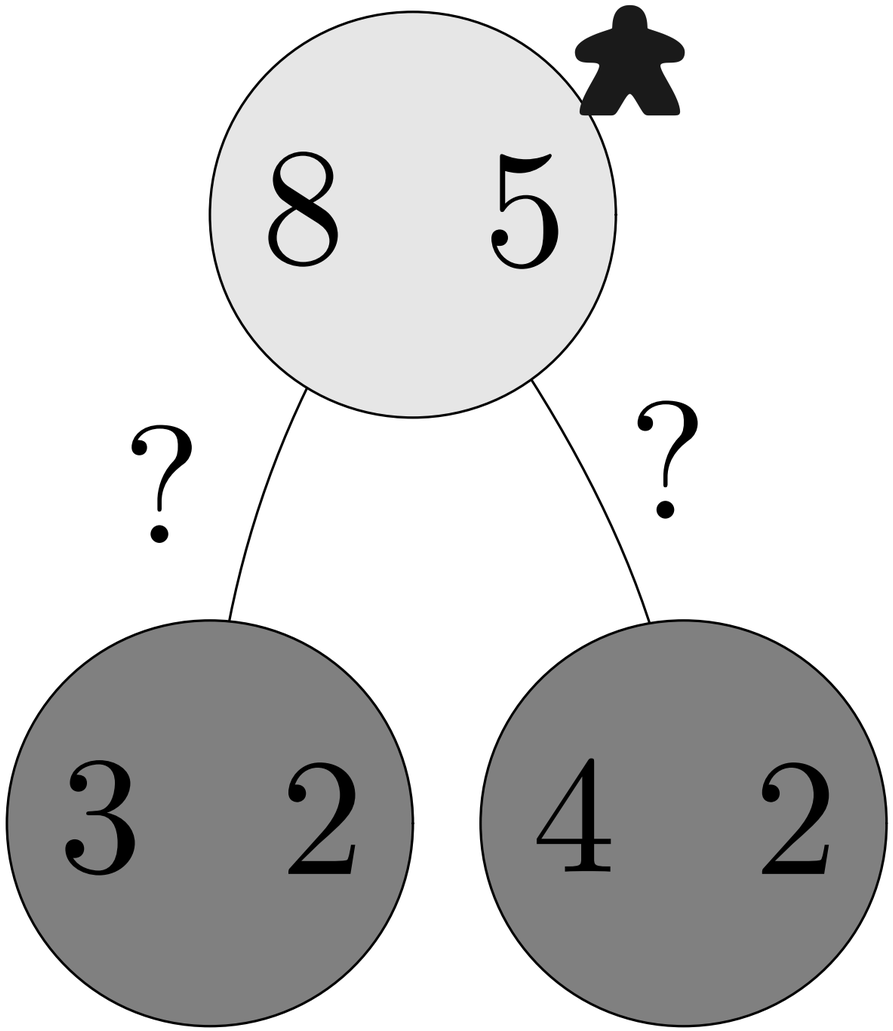
Selection
Expansion
Simulation
Backpropagation

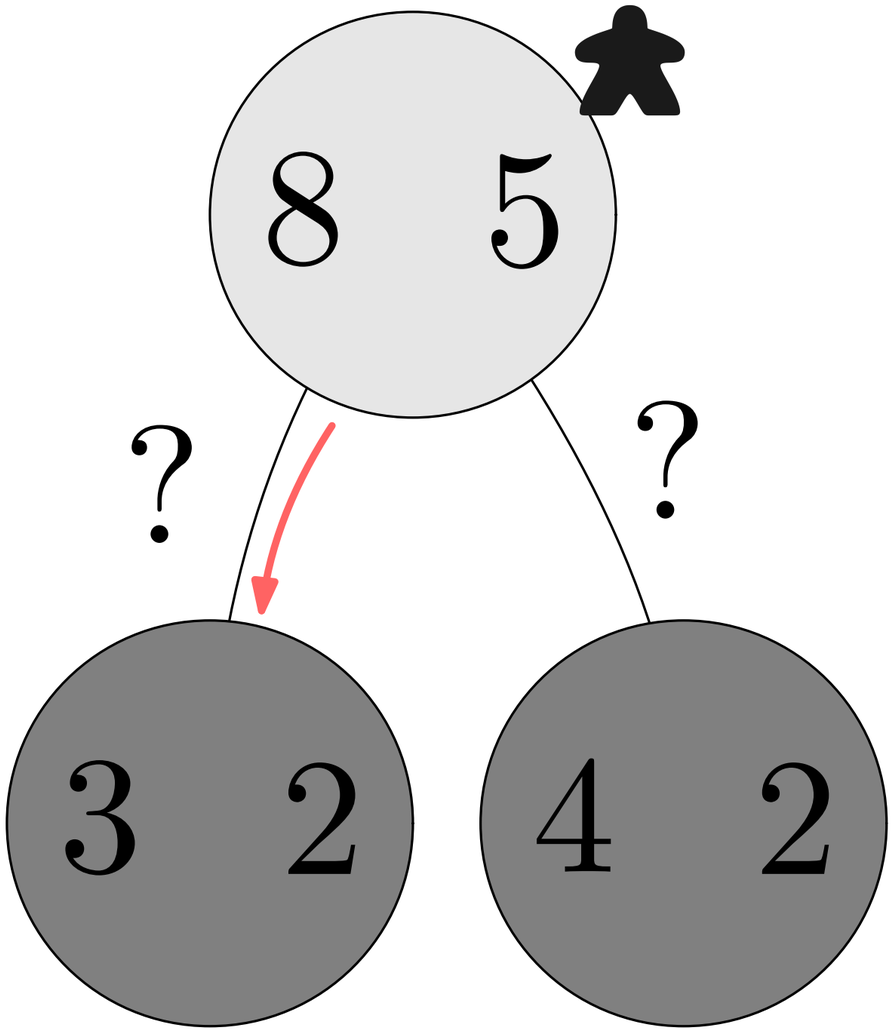


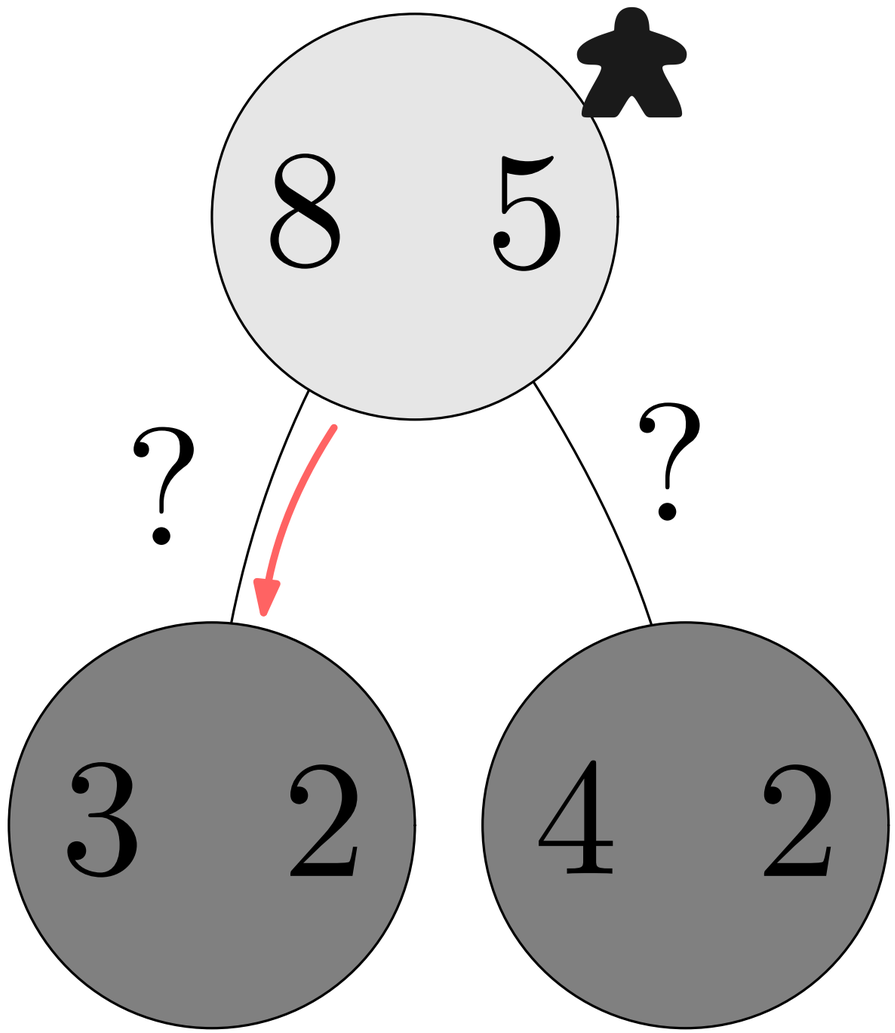




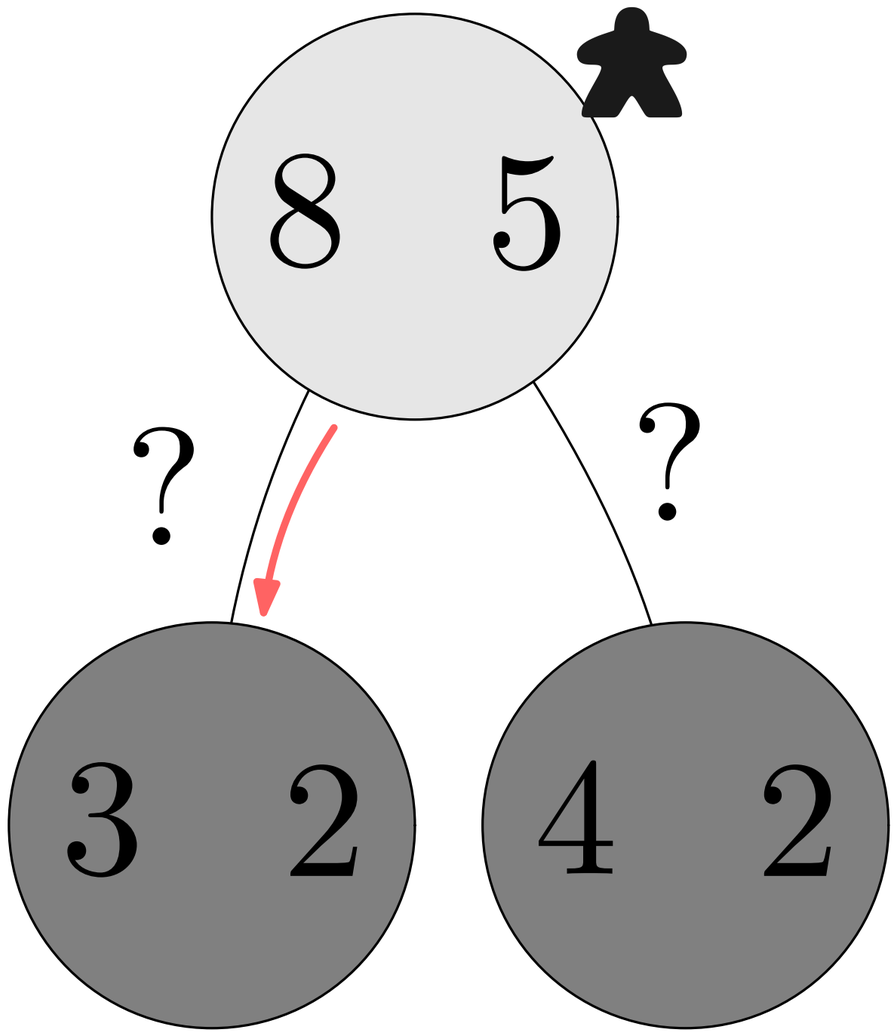



8 5
?
?
3 2
4 2


8 5
?
?
3 2
4 2



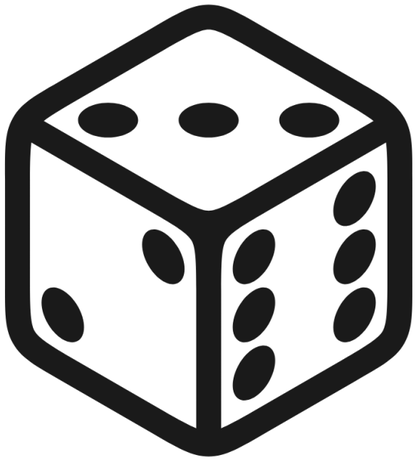
$$\max \frac{\sum \text{💰}_i}{\text{\#}_i} + c\sqrt{\log \text{\#} / \text{\#}_i}$$



$$\max \frac{\sum_i}{\#_i} + c\sqrt{\log \#/\#_i}$$

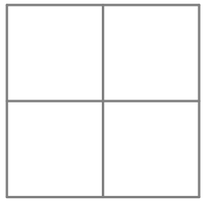
$$\max \frac{-\sum_i}{\#_i} + c\sqrt{\log \#/\#_i}$$

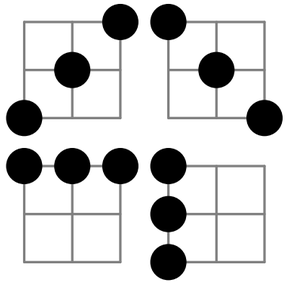
≠



Assignment 5

Tic Tac Toe





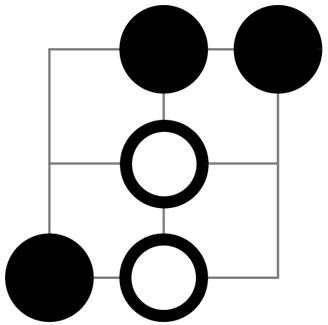
Selection Upper Confidence Bound (c = 1)

Simulation  as earlier

Post on Teams

10 steps final tree

Root



tip: use a computer :p