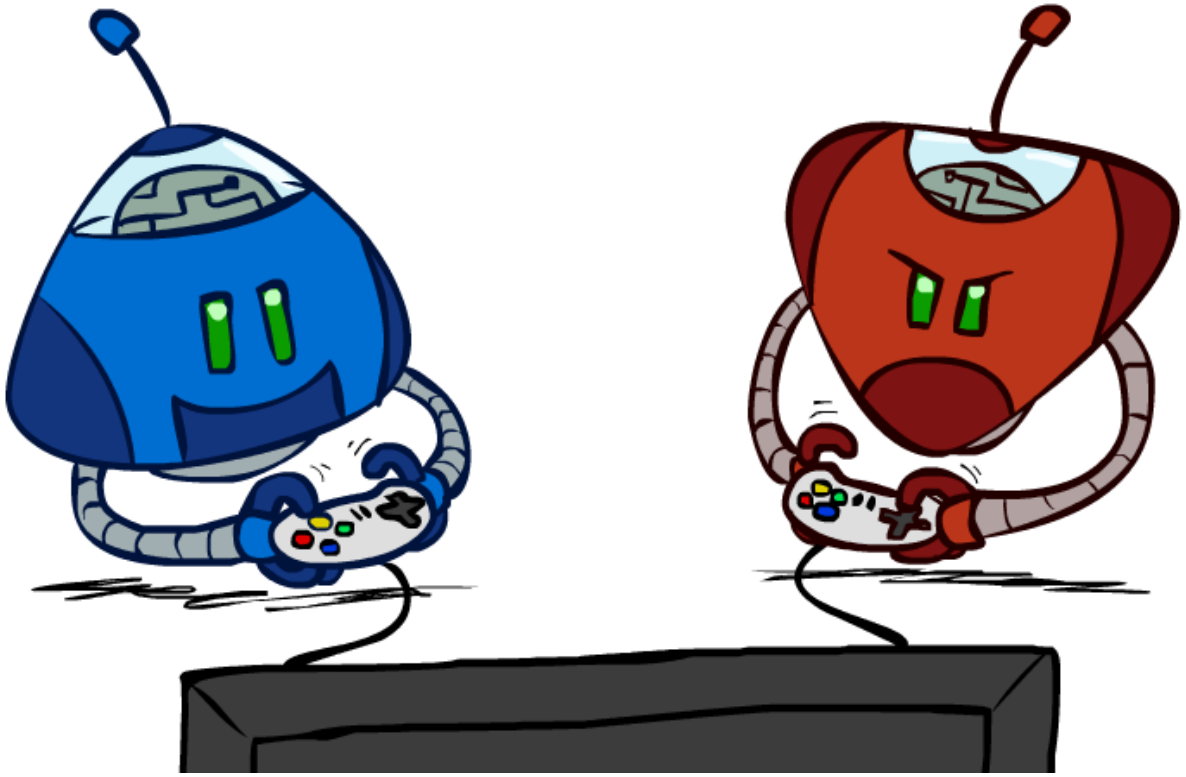
CS 5522: Artificial Intelligence II

Adversarial Search



Instructor: Wei Xu

Ohio State University

[These slides were adapted from CS188 Intro to AI at UC Berkeley.]

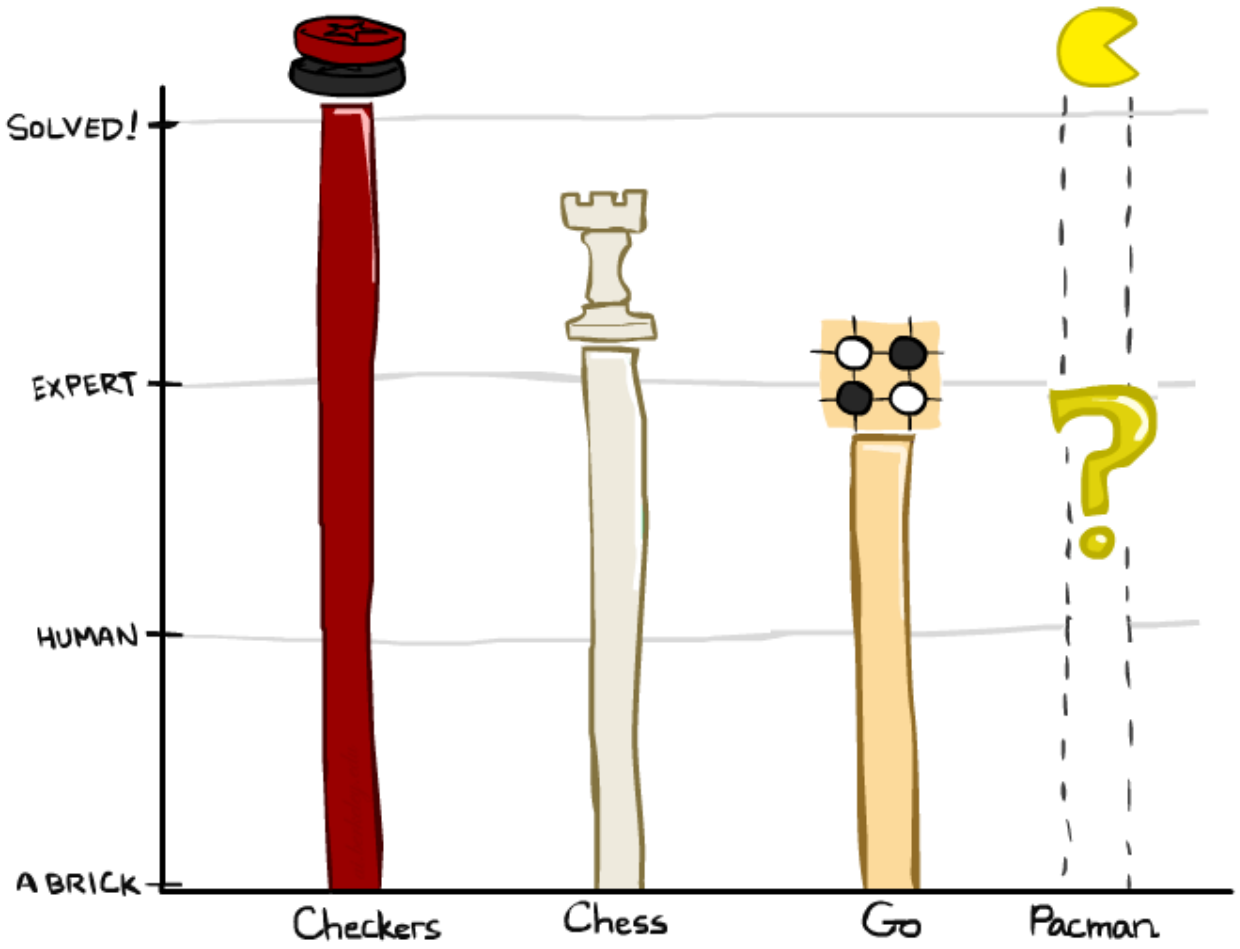
Game Playing State-of-the-Art

- **Checkers:** 1950: First computer player. 1994: First computer champion: Chinook ended 40-year-reign of human champion Marion Tinsley using complete 8-piece endgame. 2007: Checkers solved!

- **Chess:** 1997: Deep Blue defeats human champion Gary Kasparov in a six-game match. Deep Blue examined 200M positions per second, used very sophisticated evaluation and undisclosed methods for extending some lines of search up to 40 ply. Current programs are even better, if less historic.

- **Go:** Human champions are now starting to be challenged by machines, though the best humans still beat the best machines. In go, b > 300! Classic programs use pattern knowledge bases, but big recent advances use Monte Carlo (randomized) expansion methods.

- **Pacman**



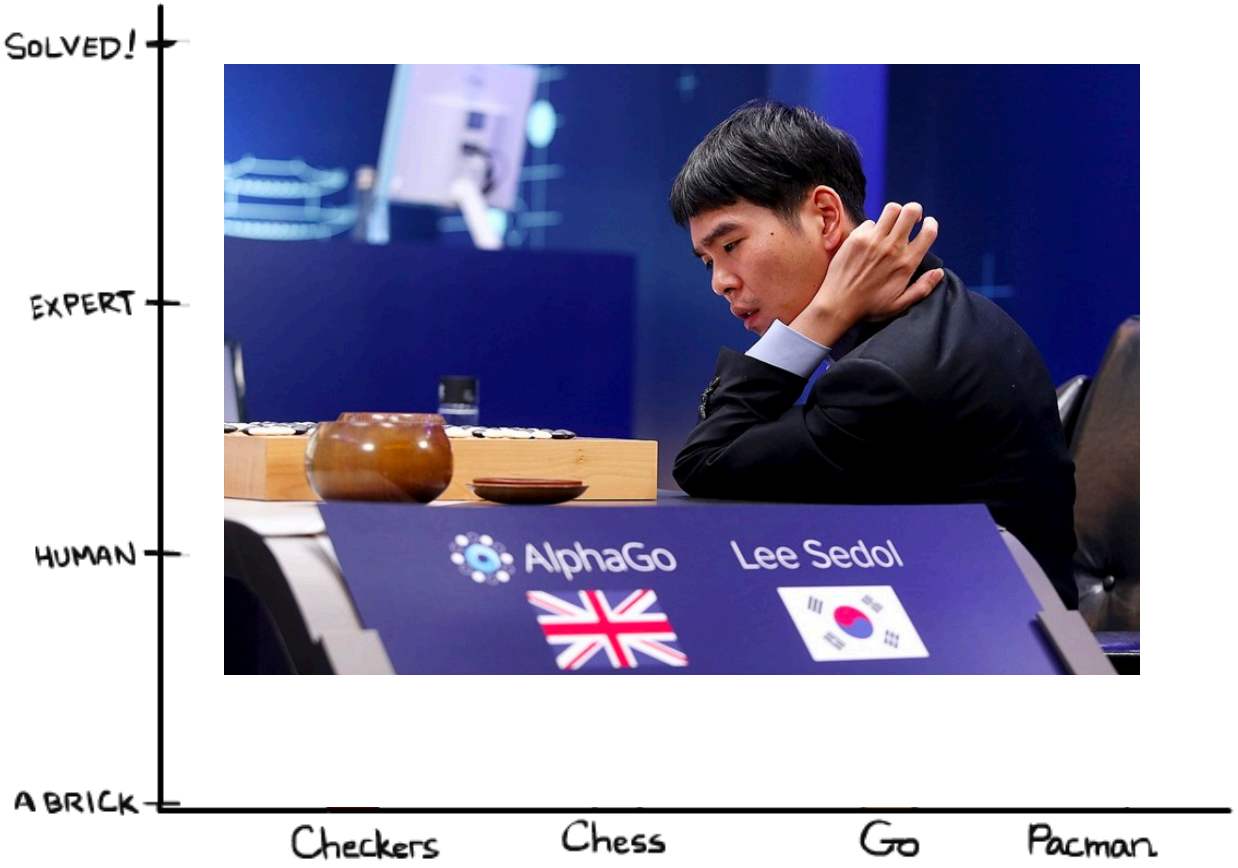
Game Playing State-of-the-Art

- **Checkers:** 1950: First computer player. 1994: First computer champion: Chinook ended 40-year-reign of human champion Marion Tinsley using complete 8-piece endgame. 2007: Checkers solved!

- **Chess:** 1997: Deep Blue defeats human champion Gary Kasparov in a six-game match. Deep Blue examined 200M positions per second, used very sophisticated evaluation and undisclosed methods for extending some lines of search up to 40 ply. Current programs are even better, if less historic.

- **Go:** Human champions are now starting to be challenged by machines, though the best humans still beat the best machines. In go, b > 300! Classic programs use pattern knowledge bases, but big recent advances use Monte Carlo (randomized) expansion methods.

- **Pacman**



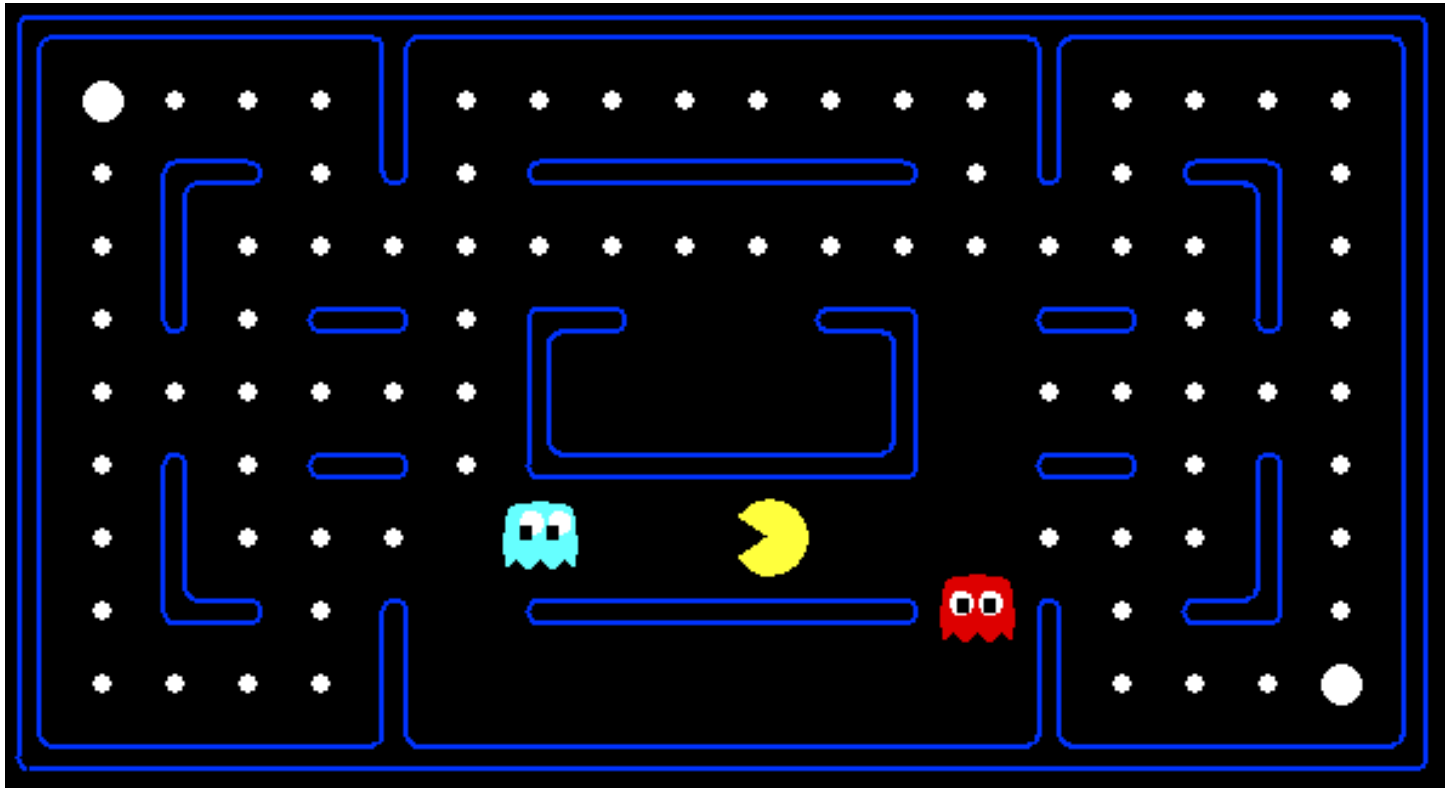
Checkers

Green Lawn Cemetery
Columbus, Franklin County, Ohio, USA

Tinsley (right) and Schaeffer (left) squaring off for the World Checkers Championship (1992)

https://www.youtube.com/watch?v=yFrAN-LFZRU

Behavior from Computation



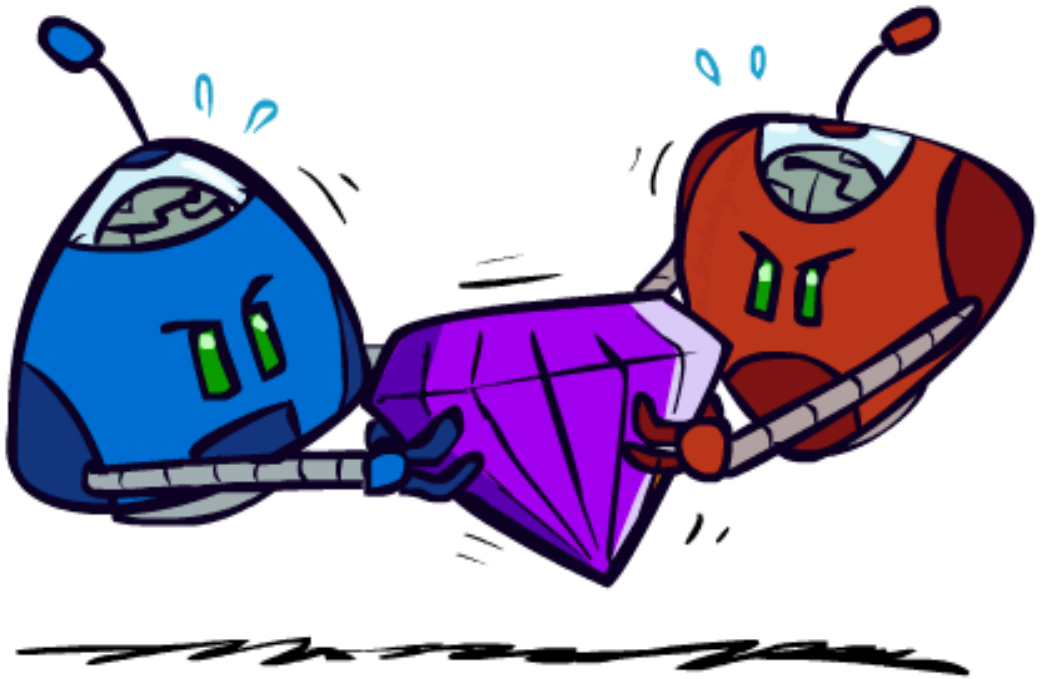
[Demo: mystery pacman (L6D1)]

Pac-Man and AI

https://www.youtube.com/watch?v=w5kFmdkrIuY

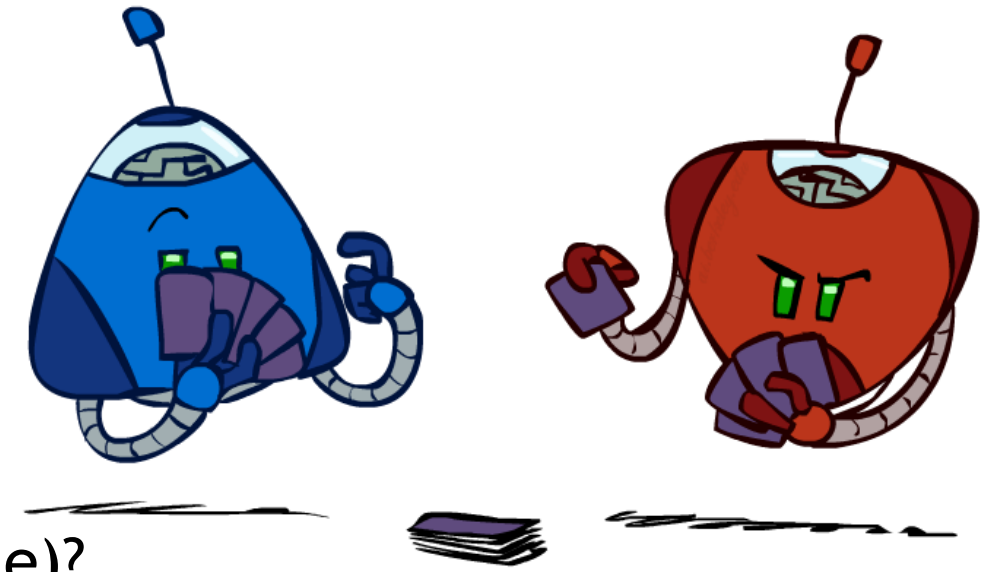
https://www.youtube.com/watch?v=zQyWMHFjewU

Adversarial Games



Types of Games

- Many different kinds of games!

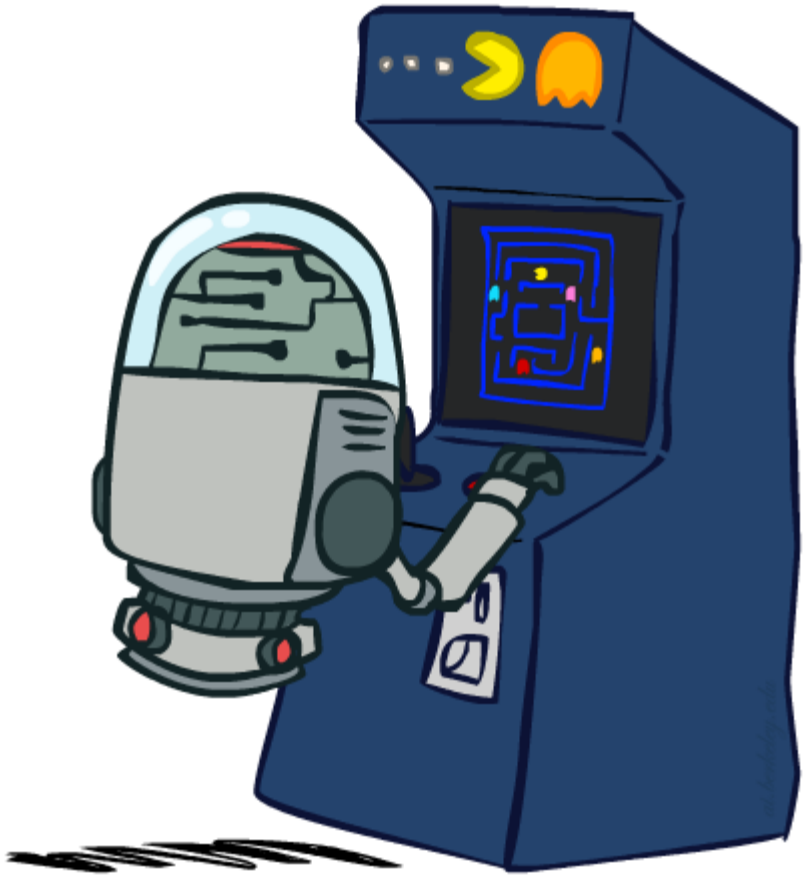


- Axes:
  - Deterministic or stochastic?
  - One, two, or more players?
  - Zero sum?
  - Perfect information (can you see the state)?

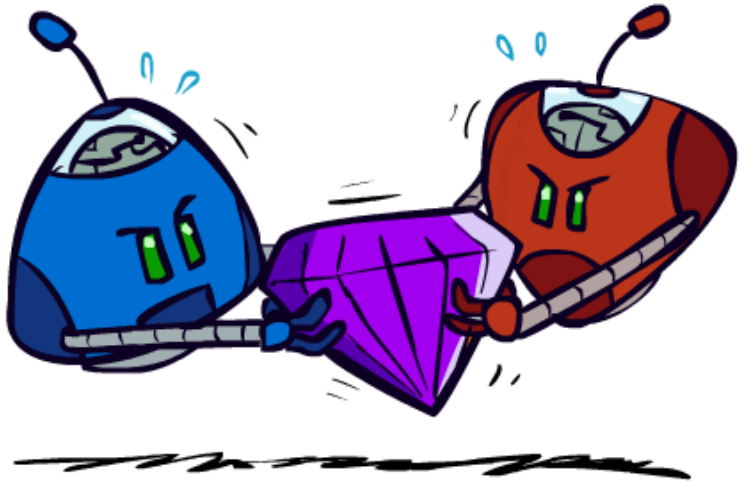
- Want algorithms for calculating a strategy (policy) which recommends a move from each state

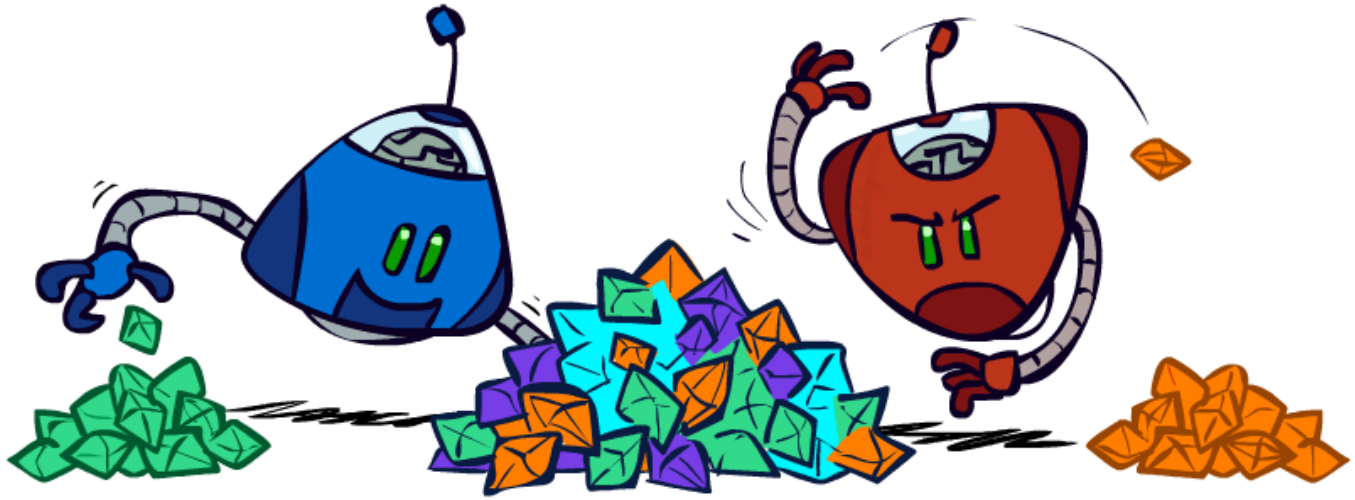
Deterministic Games

- Many possible formalizations, one is:
  - States: S (start at s_0)
  - Players: P={1...N} (usually take turns)
  - Actions: A (may depend on player / state)
  - Transition Function: SxA → S
  - Terminal Test: S → {t,f}
  - Terminal Utilities: SxP → R
- Solution for a player is a policy: S → A



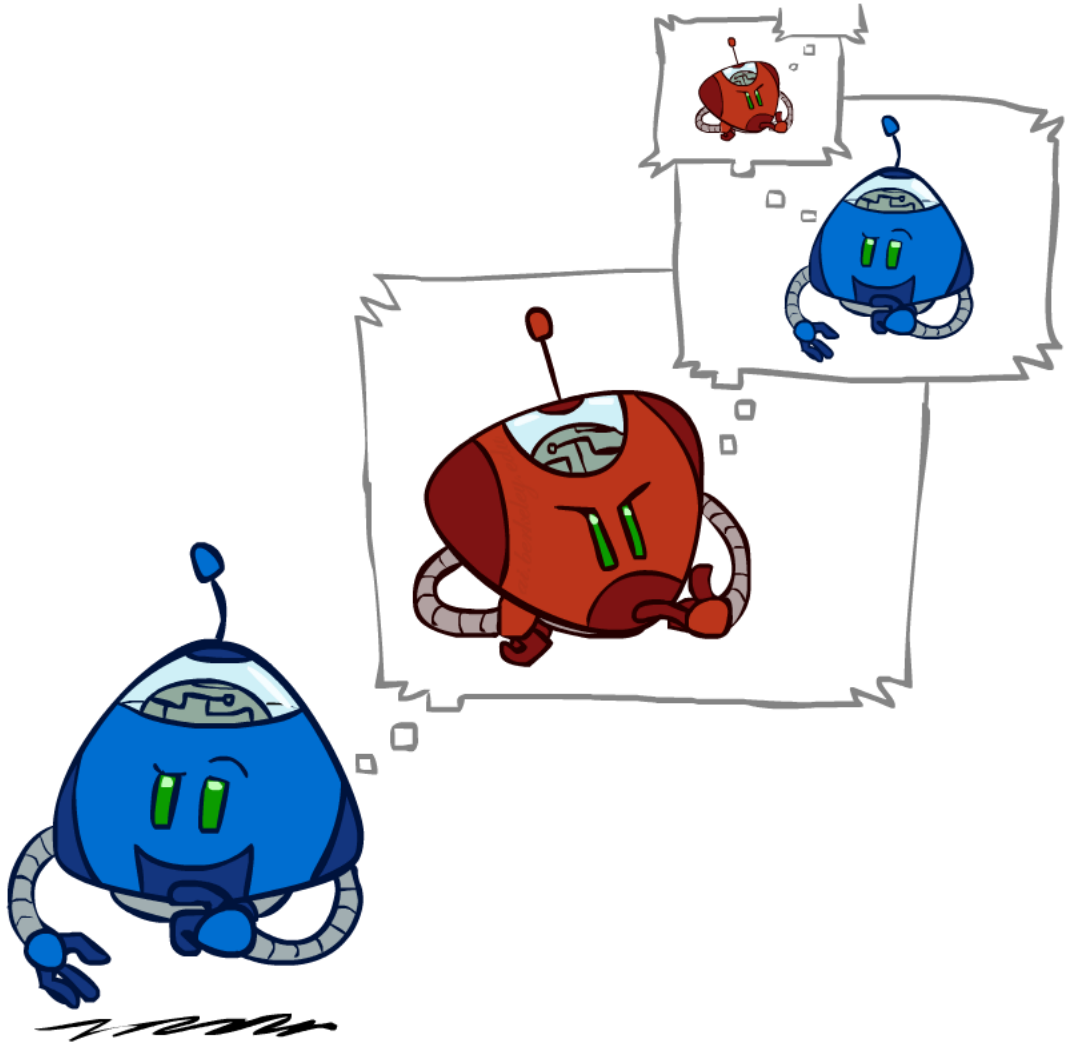
Zero-Sum Games



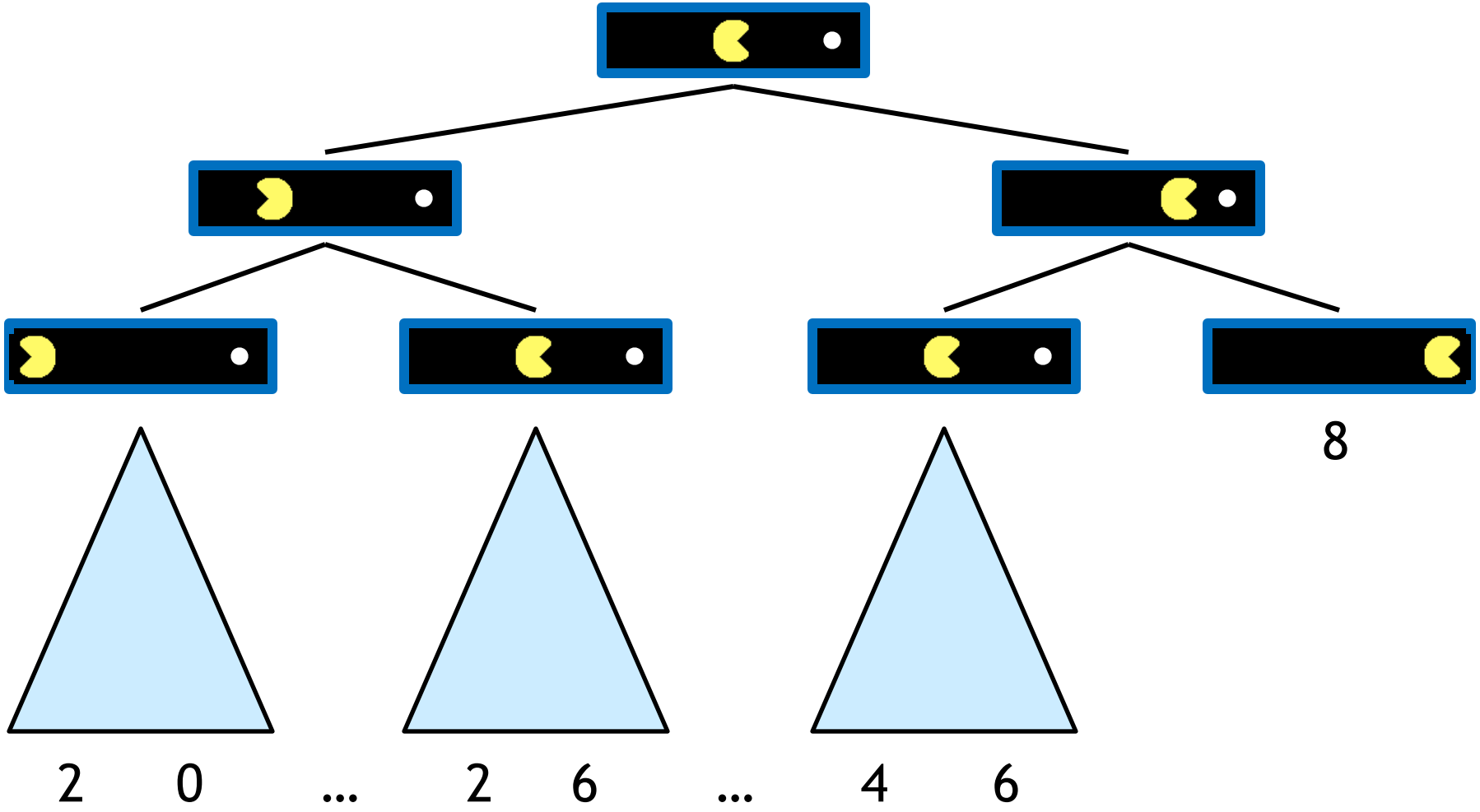


- Zero-Sum Games
  - Agents have opposite utilities (values on outcomes)
  - Lets us think of a single value that one maximizes and the other minimizes
  - Adversarial, pure competition
- General Games
  - Agents have independent utilities (values on outcomes)
  - Cooperation, indifference, competition, and more are all possible
  - More later on non-zero-sum games

Adversarial Search

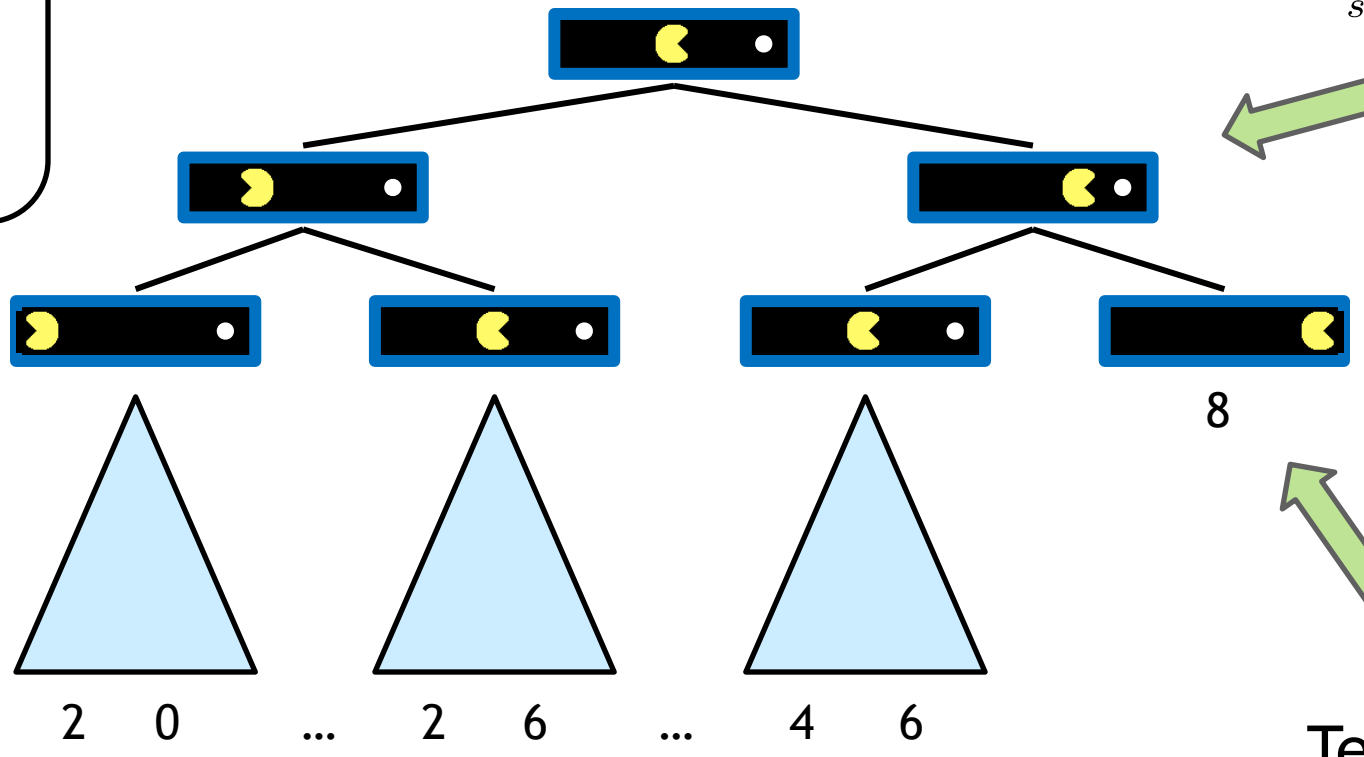


Single-Agent Trees



Value of a State

Value of a state:
The best achievable outcome (utility) from that state



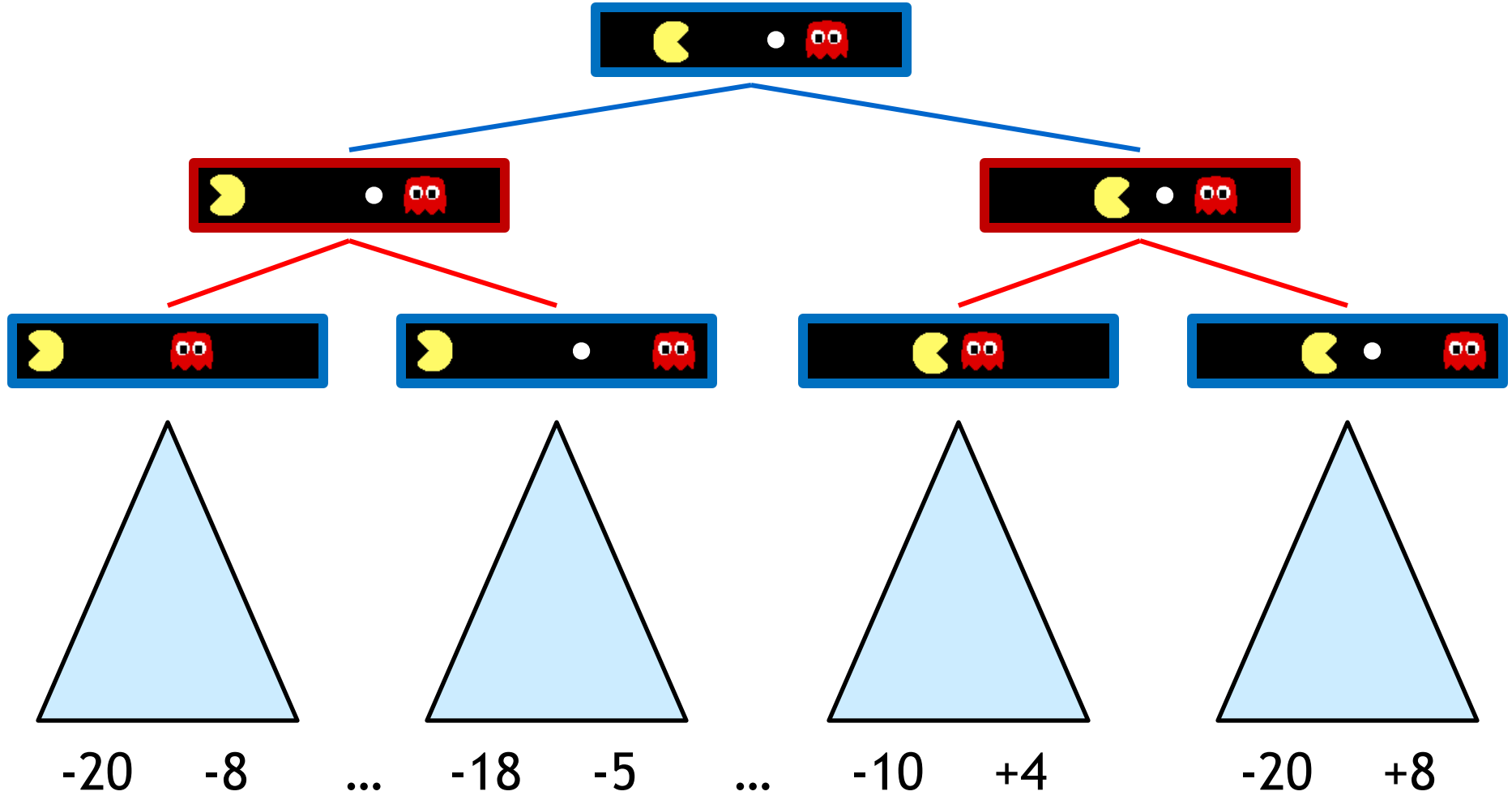
Non-Terminal States:

$$V(s) = \max_{s' \in \text{children}(s)} V(s')$$

Terminal States:

$$V(s) = \text{known}$$

Adversarial Game Trees



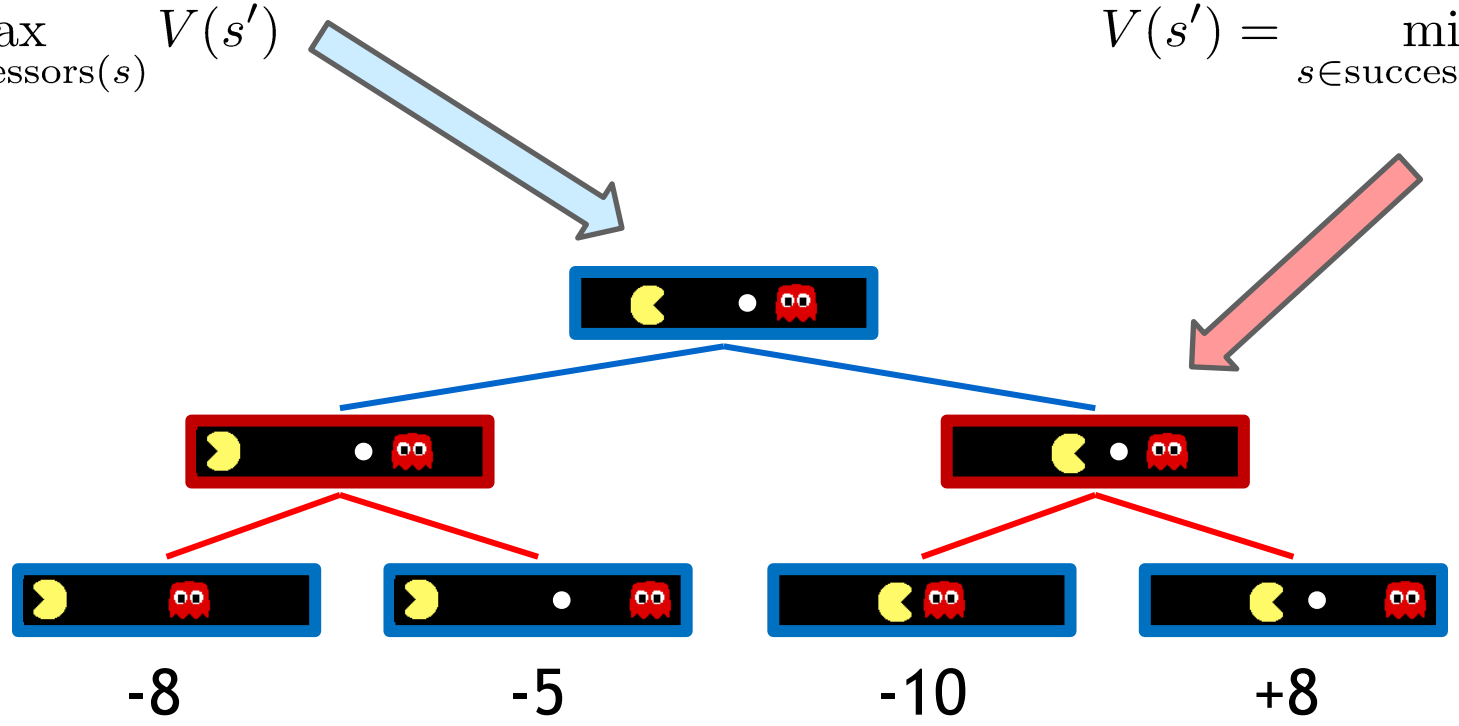
Minimax Values

States Under Agent's Control:

$$V(s) = \max_{s' \in \text{successors}(s)} V(s')$$

States Under Opponent's Control:

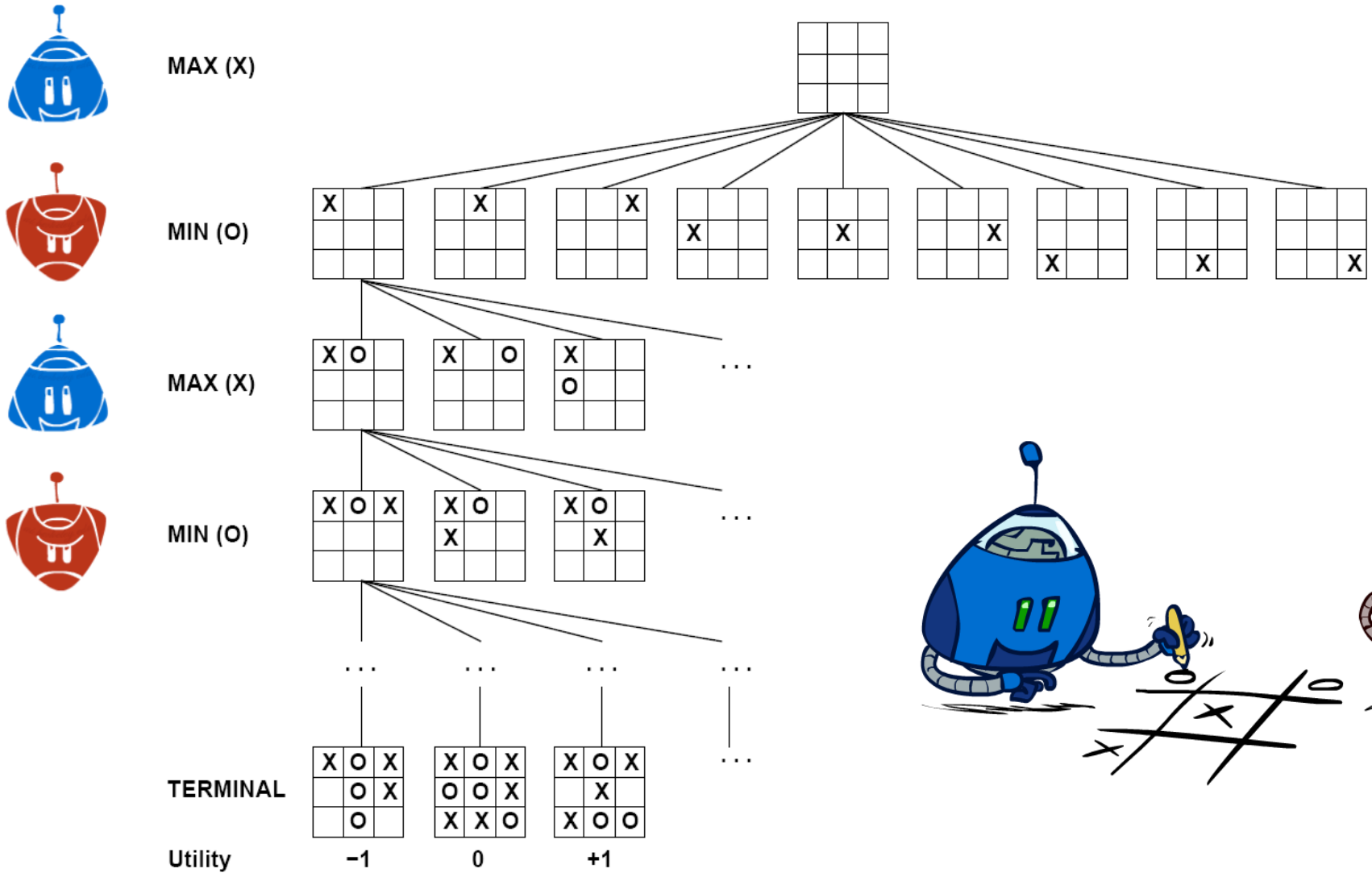
$$V(s') = \min_{s \in \text{successors}(s')} V(s)$$



Terminal States:

$$V(s) = \text{known}$$

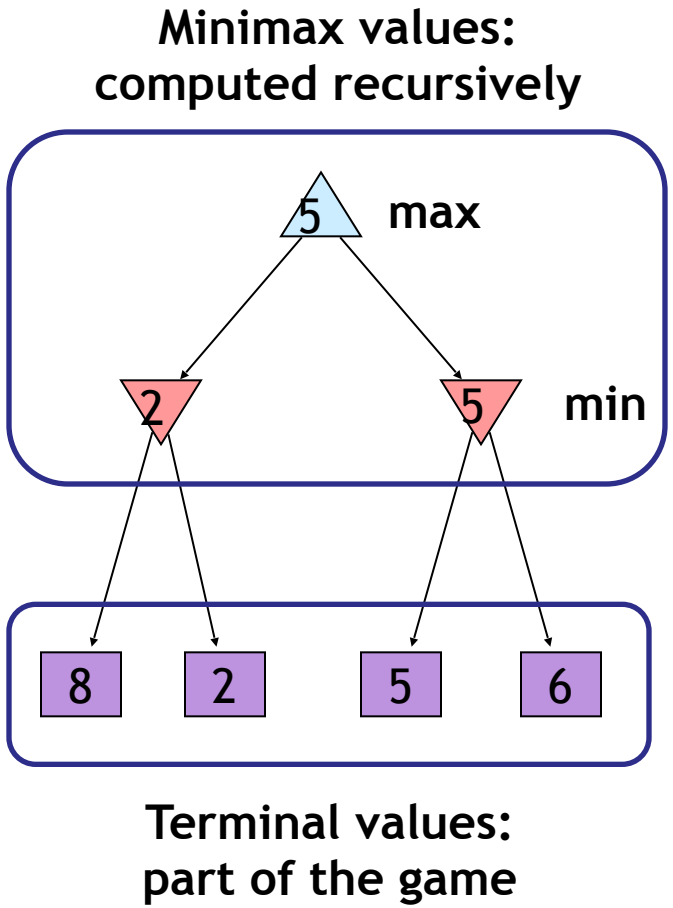
Tic-Tac-Toe Game Tree



Adversarial Search (Minimax)

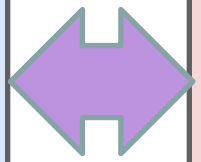
- Deterministic, zero-sum games:
  - Tic-tac-toe, chess, checkers
  - One player maximizes result
  - The other minimizes result

- Minimax search:
  - A state-space search tree
  - Players alternate turns
  - Compute each node's minimax value: the best achievable utility against a rational (optimal) adversary



Minimax Implementation

```
def max-value(state):
    initialize v = -∞
    for each successor of state:
        v = max(v, min-value(successor))
    return v
```



```
def min-value(state):
    initialize v = +∞
    for each successor of state:
        v = min(v, max-value(successor))
    return v
```

$$V(s) = \max_{s' \in \text{successors}(s)} V(s')$$

$$V(s') = \min_{s \in \text{successors}(s')} V(s)$$

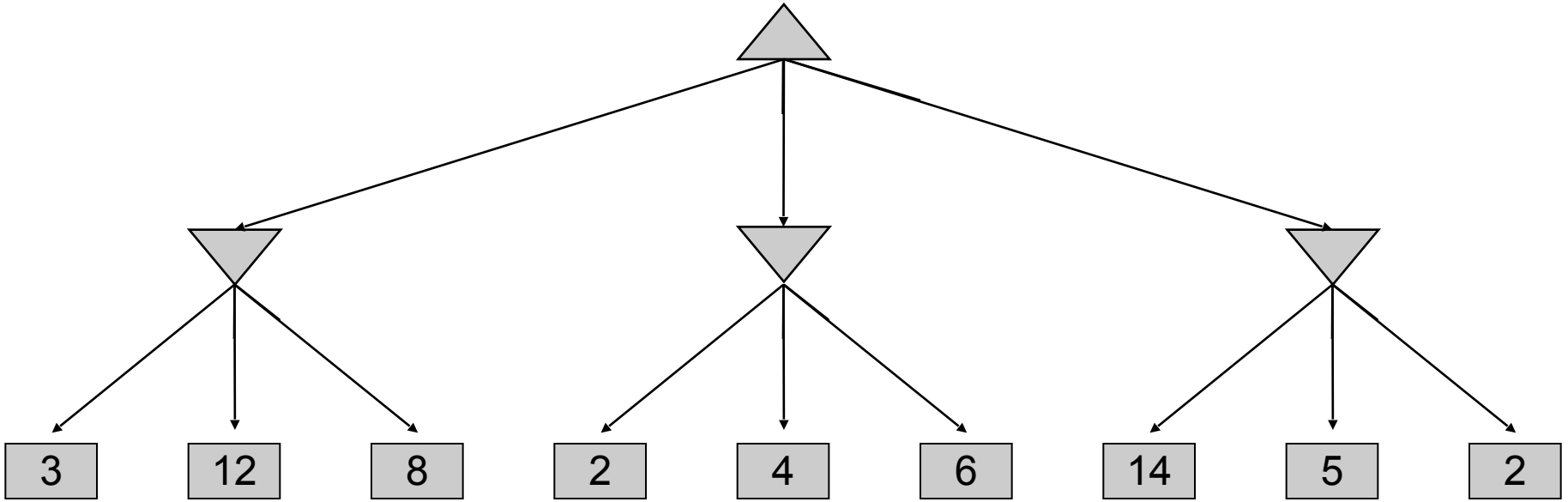
Minimax Implementation (Dispatch)

```
def value(state):
    if the state is a terminal state: return the state's utility
    if the next agent is MAX: return max-value(state)
    if the next agent is MIN: return min-value(state)
```

```
def max-value(state):
    initialize v = -∞
    for each successor of state:
        v = max(v, value(successor))
    return v
```

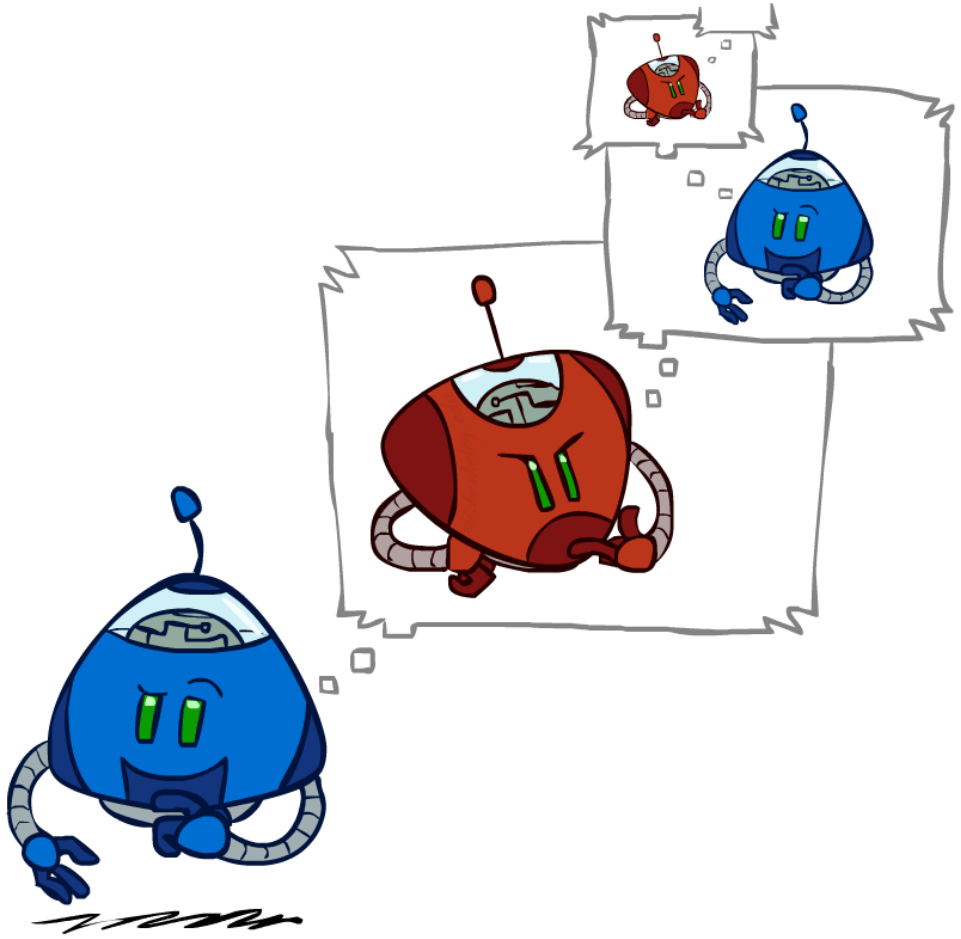
```
def min-value(state):
    initialize v = +∞
    for each successor of state:
        v = min(v, value(successor))
    return v
```

Minimax Example

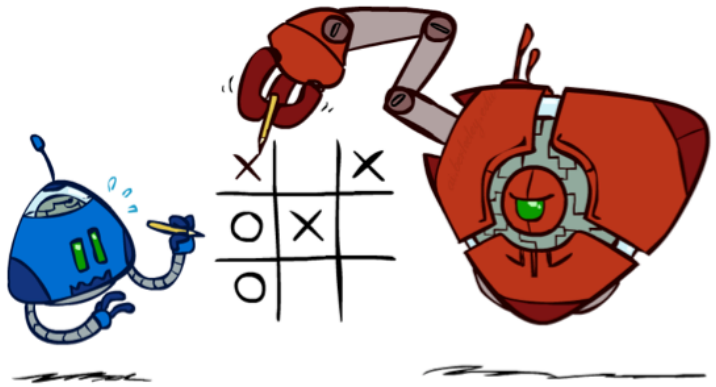


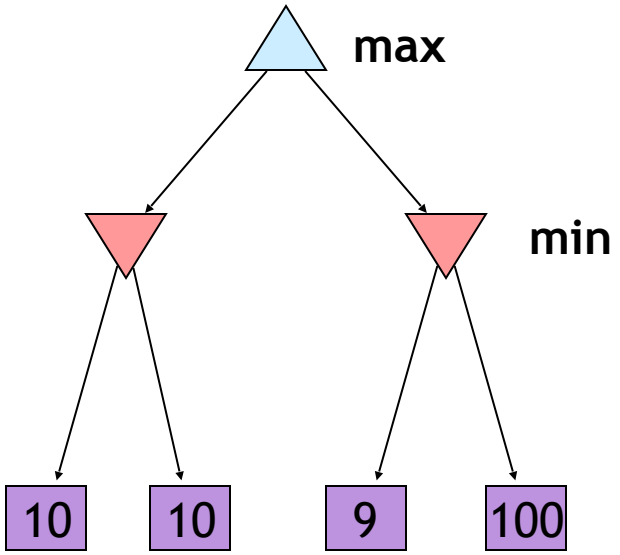
Minimax Efficiency

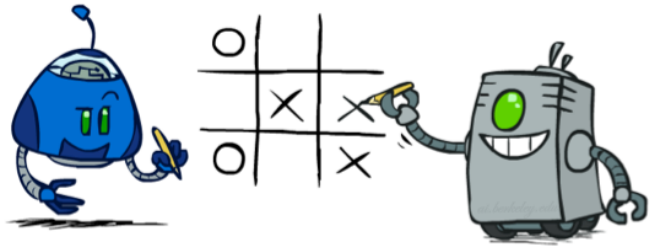
- How efficient is minimax?
  - Just like (exhaustive) DFS
  - Time: $O(b^m)$
  - Space: O(bm)
- Example: For chess, $b \approx 35$, $m \approx 100$
  - Exact solution is completely infeasible
  - But, do we need to explore the whole tree?



Minimax Properties



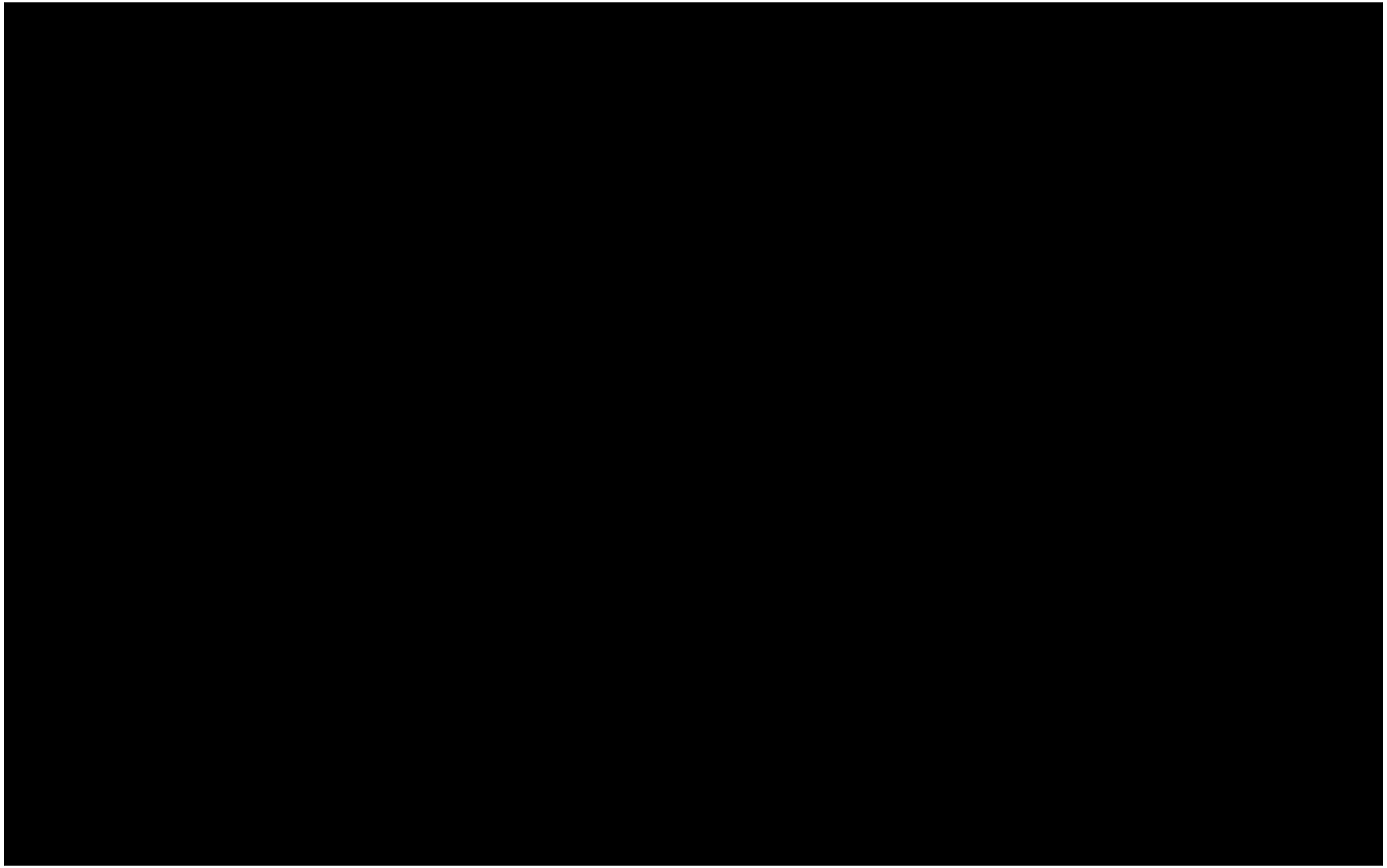




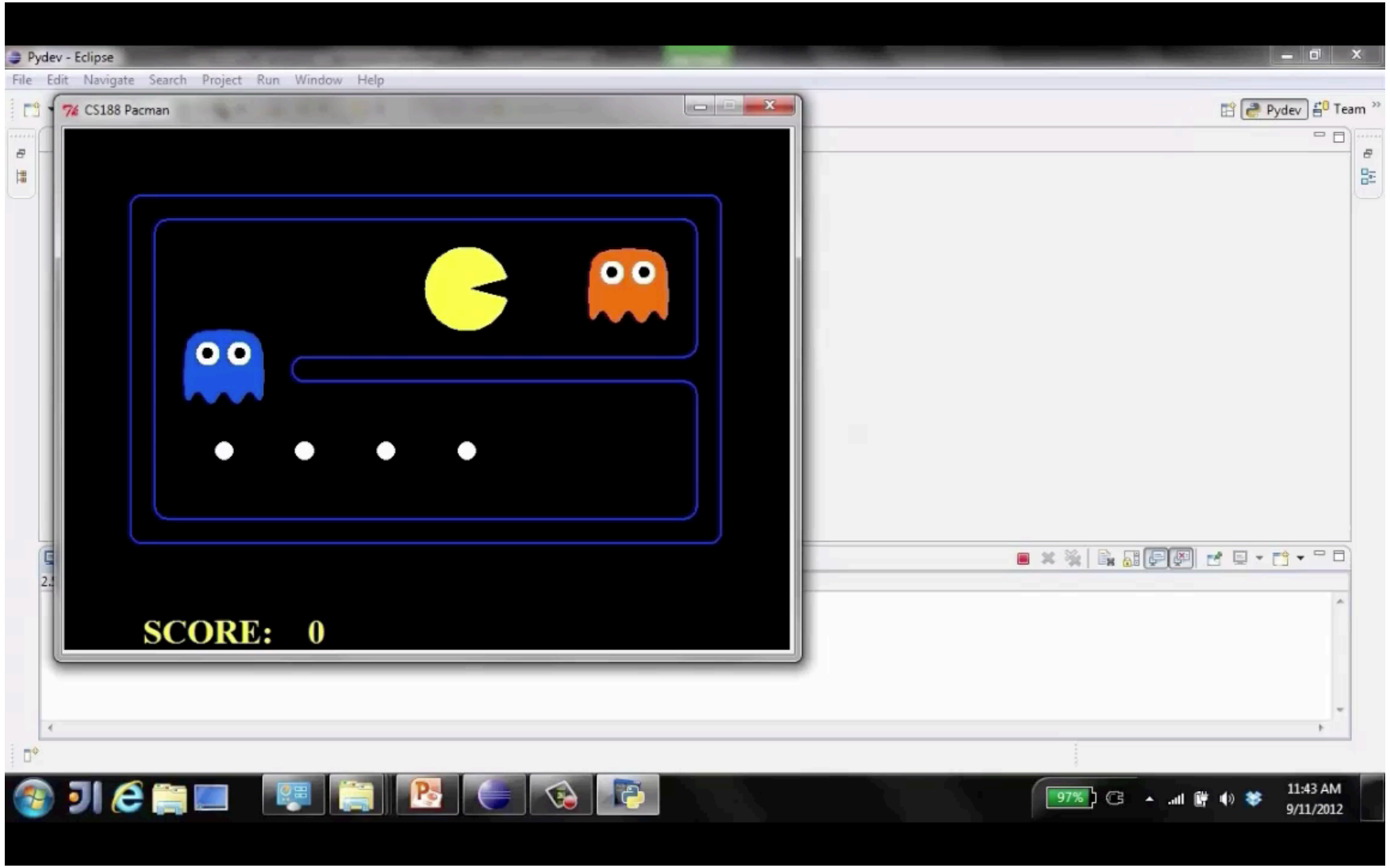
Optimal against a perfect player. Otherwise?

[Demo: min vs exp (L6D2, L6D3)]

Video of Demo Min vs. Exp (Min)



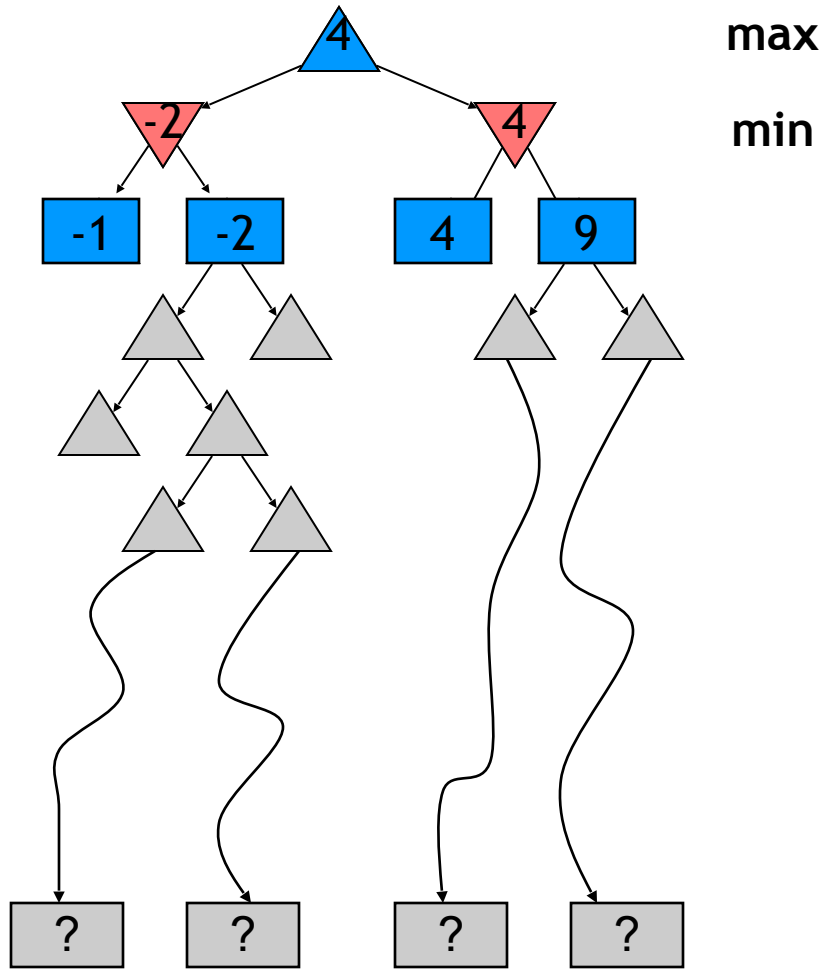
Video of Demo Min vs. Exp (Exp)



Resource Limits

Resource Limits

- Problem: In realistic games, cannot search to leaves!
- Solution: Depth-limited search
  - Instead, search only to a limited depth in the tree
  - Replace terminal utilities with an evaluation function for non-terminal positions
- Example:
  - Suppose we have 100 seconds, can explore 10K nodes / sec
  - So can check 1M nodes per move
  - α-β reaches about depth 8 – decent chess program
- Guarantee of optimal play is gone
- More plies makes a BIG difference
- Use **iterative deepening** for an anytime algorithm

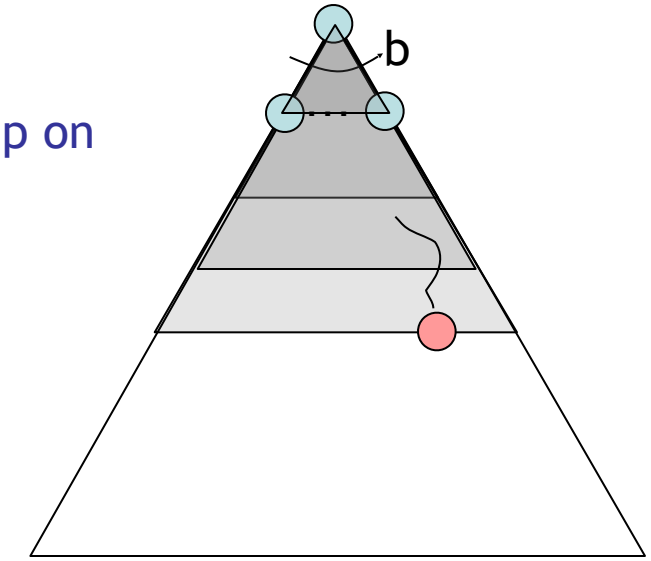


Iterative Deepening

Iterative deepening uses DFS as a subroutine:

1. Do a DFS which only searches for paths of length 1 or less. (DFS gives up on any path of length 2)
2. If “1” failed, do a DFS which only searches paths of length 2 or less.
3. If “2” failed, do a DFS which only searches paths of length 3 or less.

….and so on.

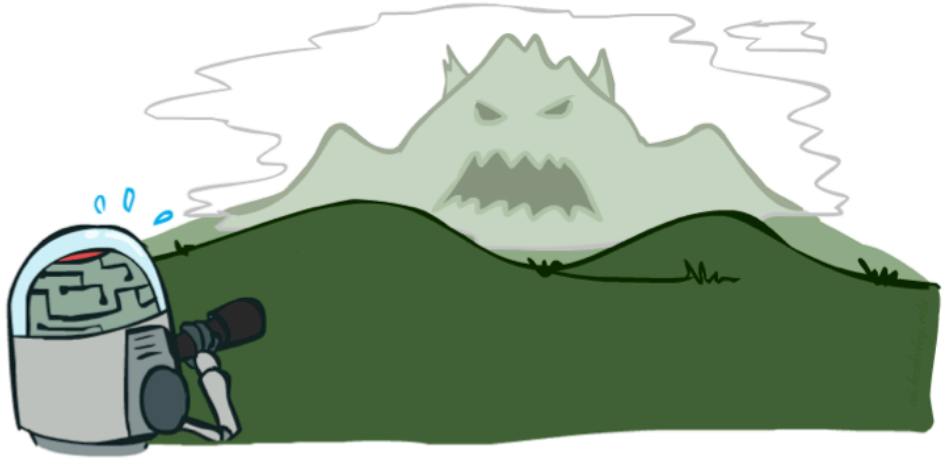


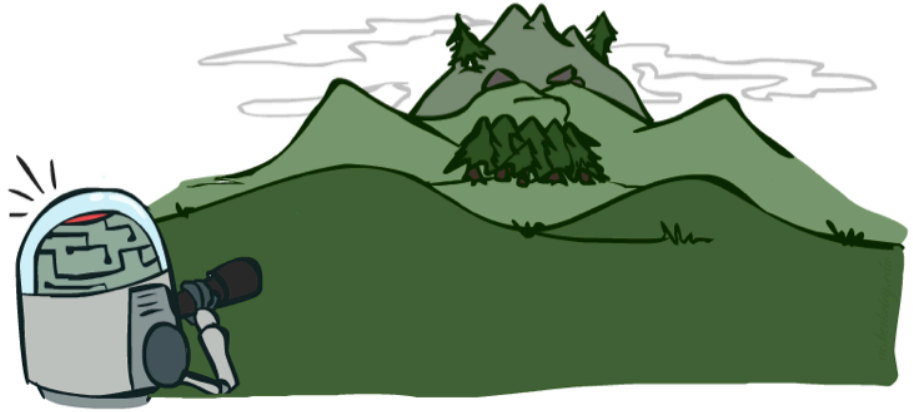
Why do we want to do this for multiplayer games?

Note: wrongness of eval functions matters less and less the deeper the search goes!

Depth Matters

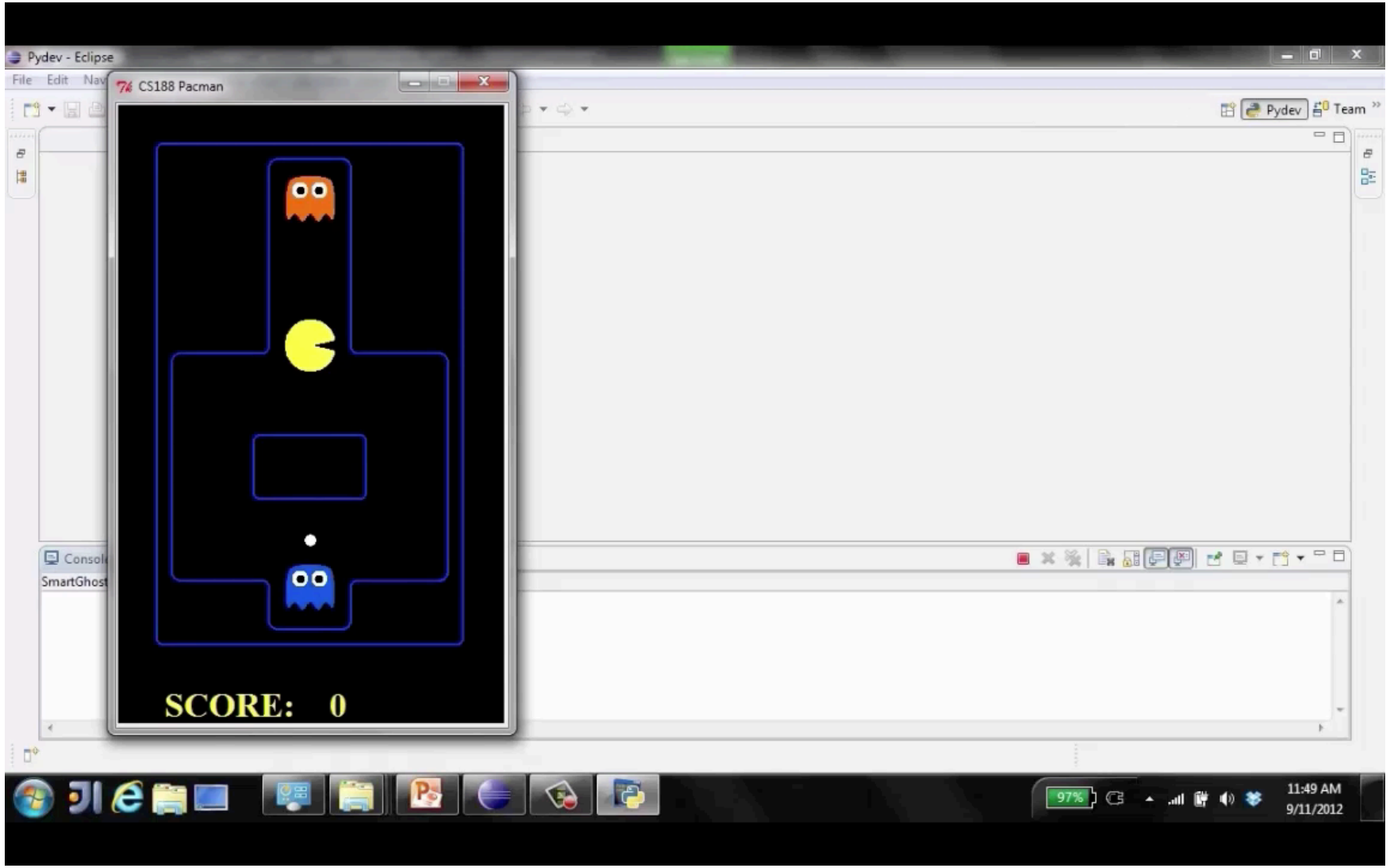
- Evaluation functions are always imperfect
- The deeper in the tree the evaluation function is buried, the less the quality of the evaluation function matters
- An important example of the tradeoff between complexity of features and complexity of computation



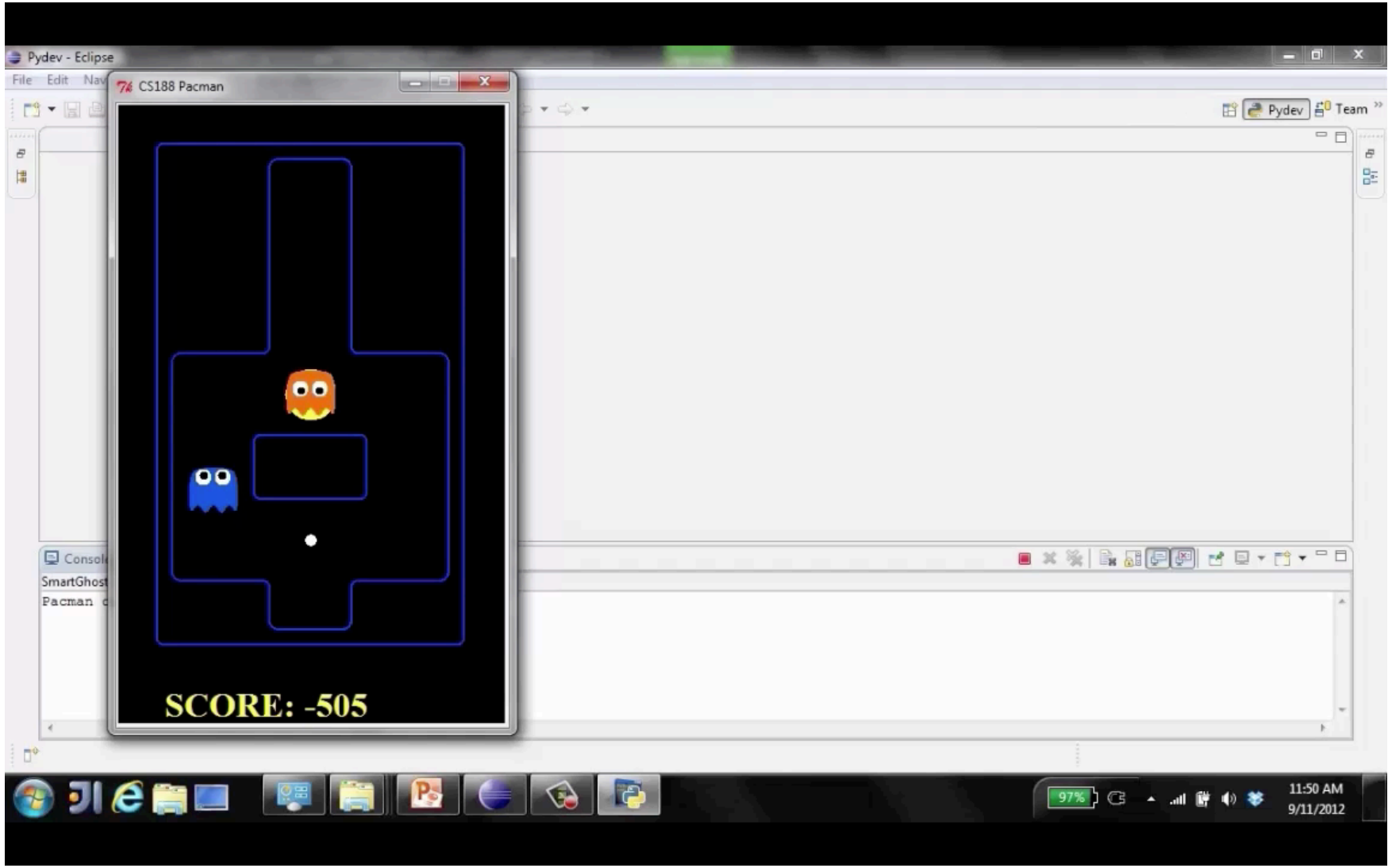


[Demo: depth limited (L6D4, L6D5)]

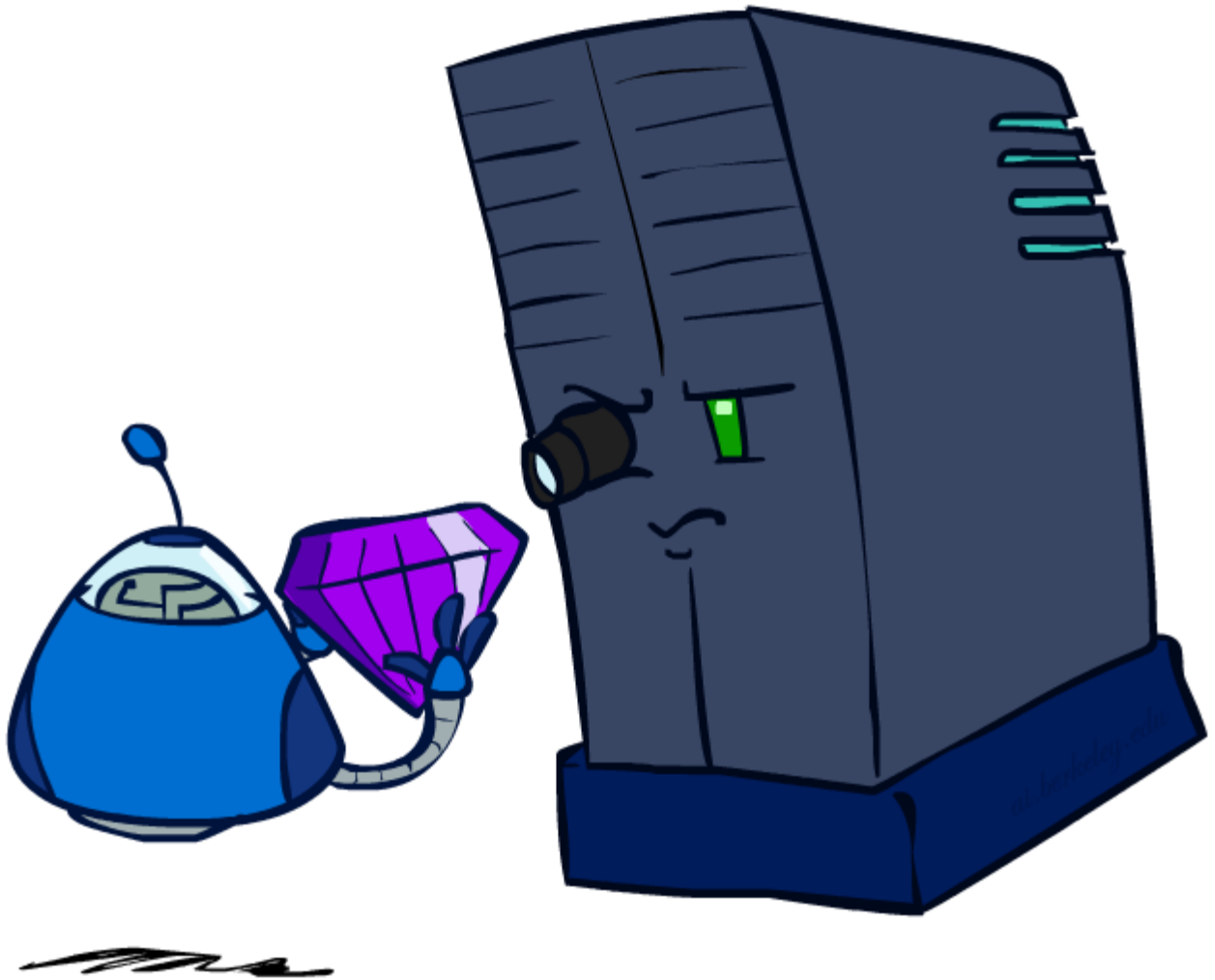
Video of Demo Limited Depth (2)



Video of Demo Limited Depth (10)

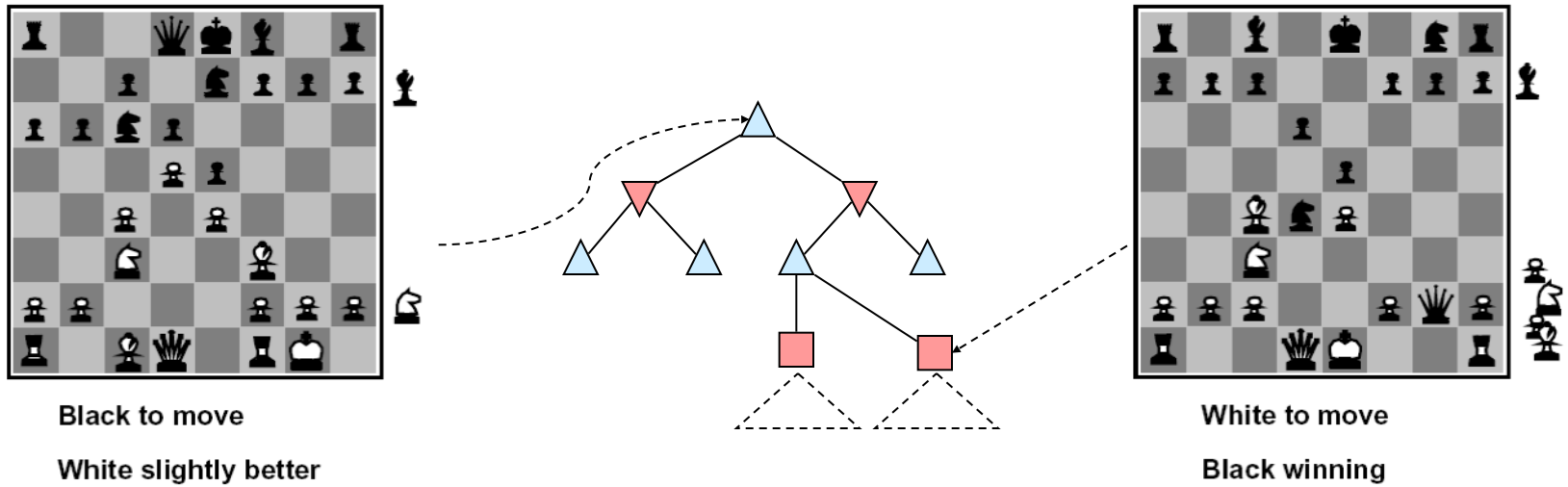


Evaluation Functions



Evaluation Functions

- Evaluation functions score non-terminals in depth-limited search

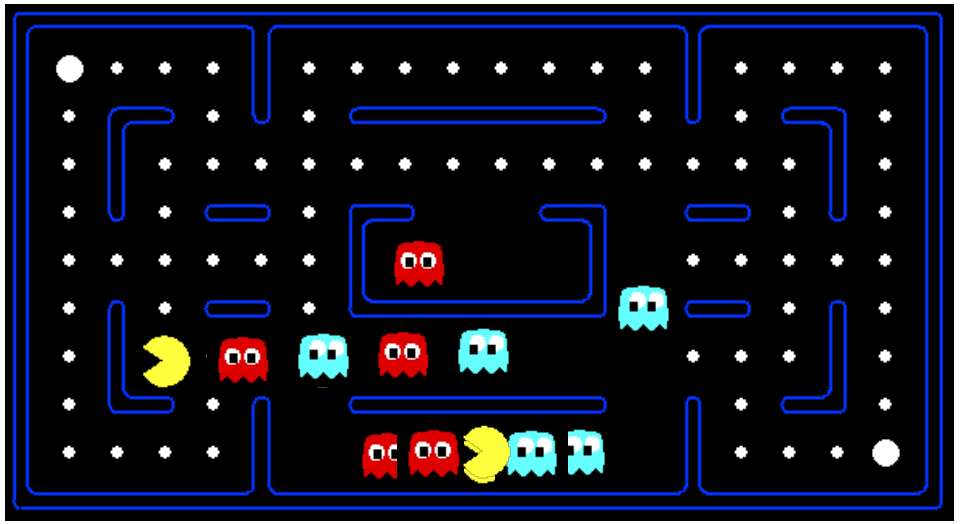


- Ideal function: returns the actual minimax value of the position
- In practice: typically weighted linear sum of features:

$$Eval(s) = w_1 f_1(s) + w_2 f_2(s) + \ldots + w_n f_n(s)$$

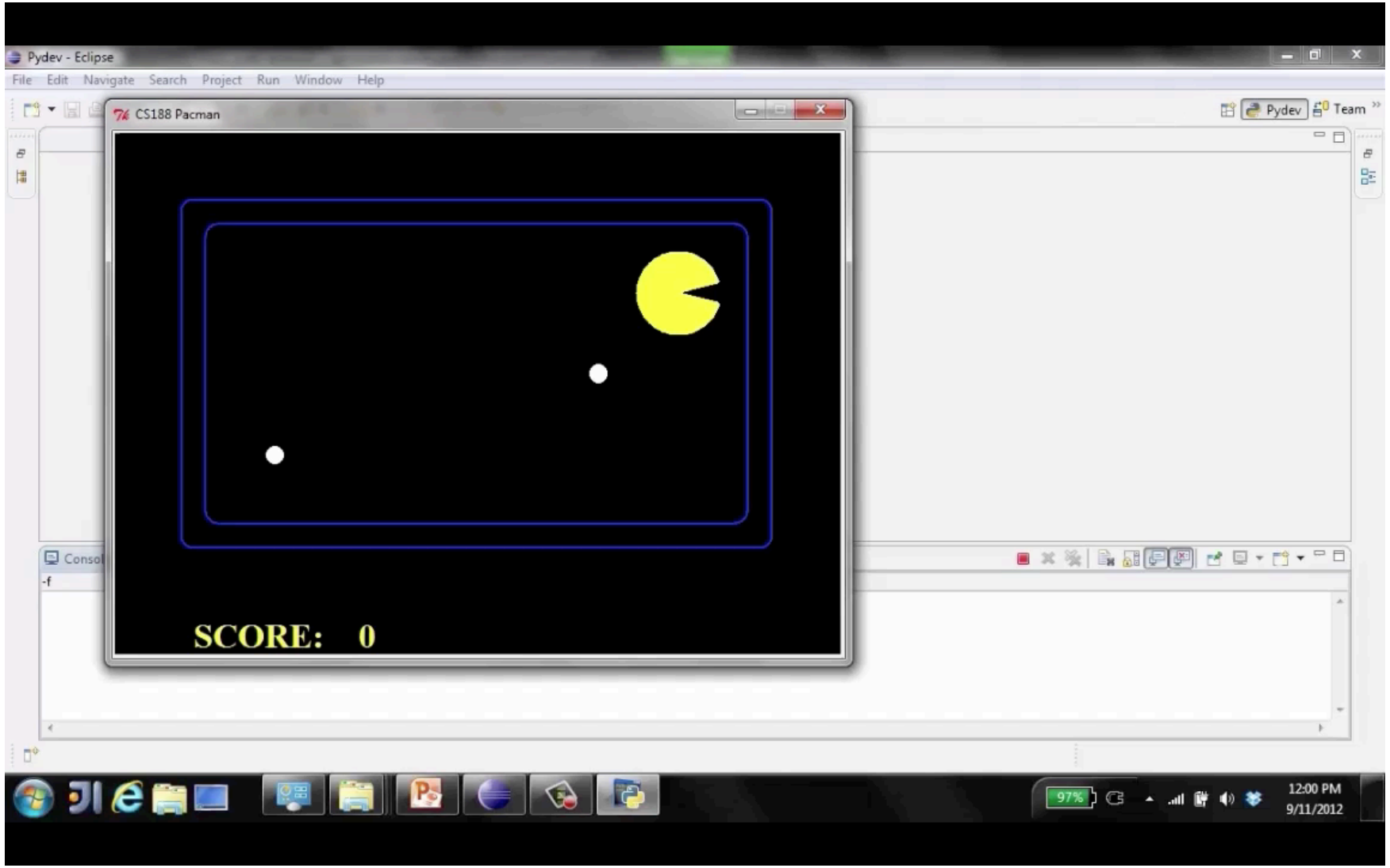
- e.g. $f_1(s)$ = (num white queens – num black queens), etc.

Evaluation for Pacman

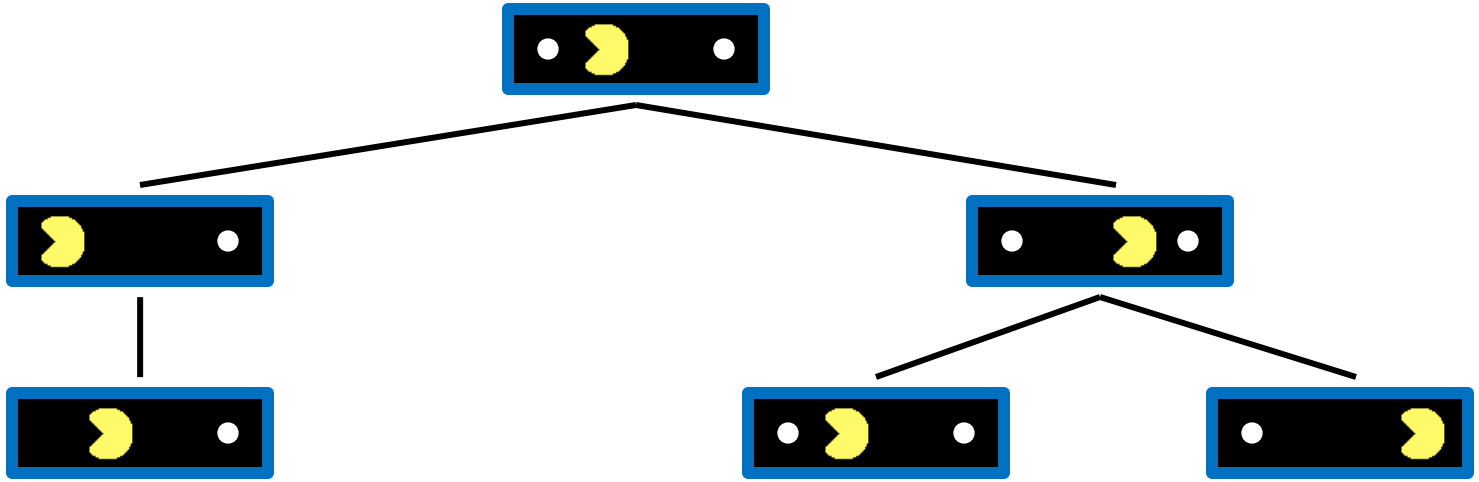


[Demo: thrashing d=2, thrashing d=2 (fixed evaluation function), smart ghosts coordinate

Video of Demo Thrashing (d=3)

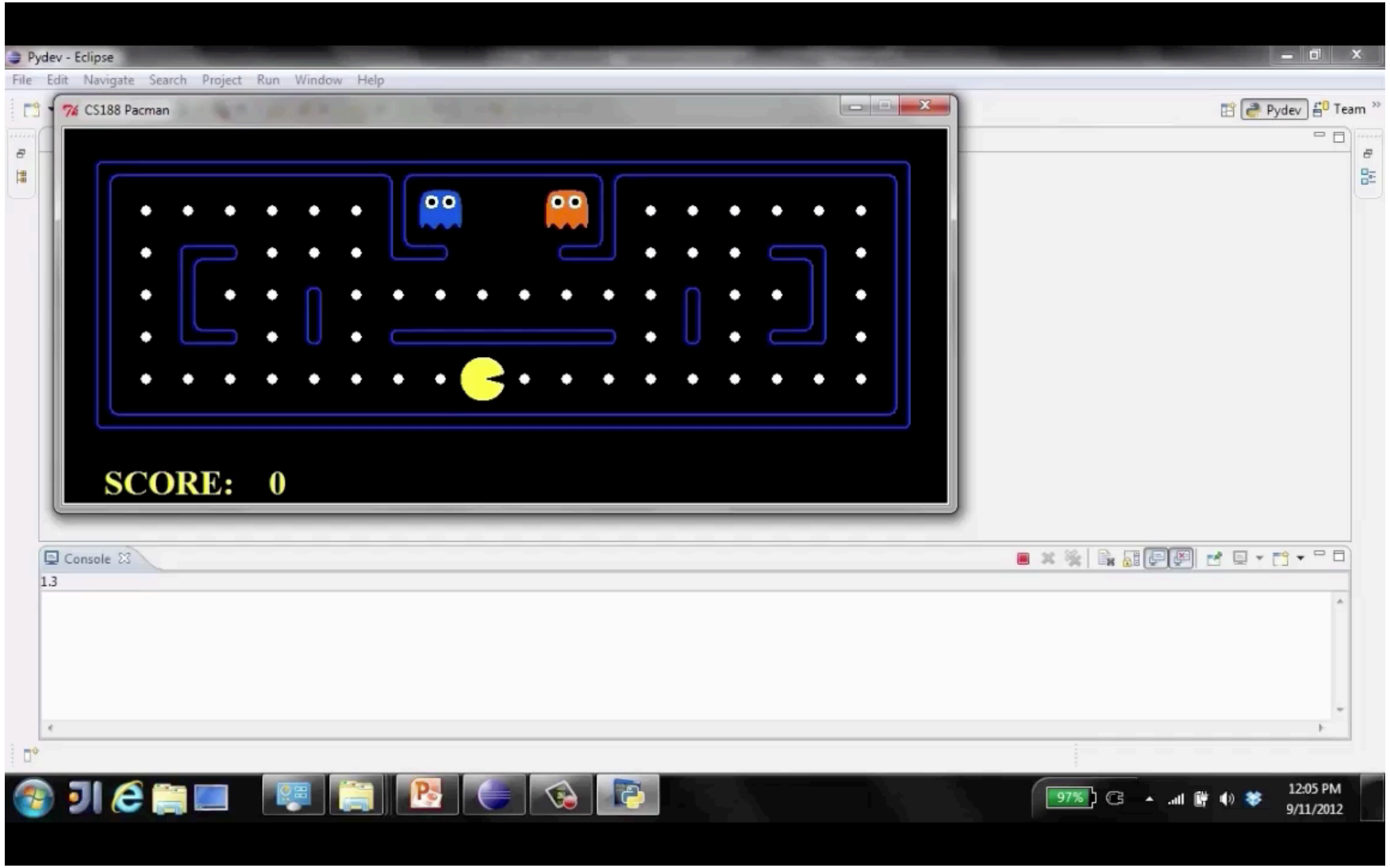


Why Pacman Starves

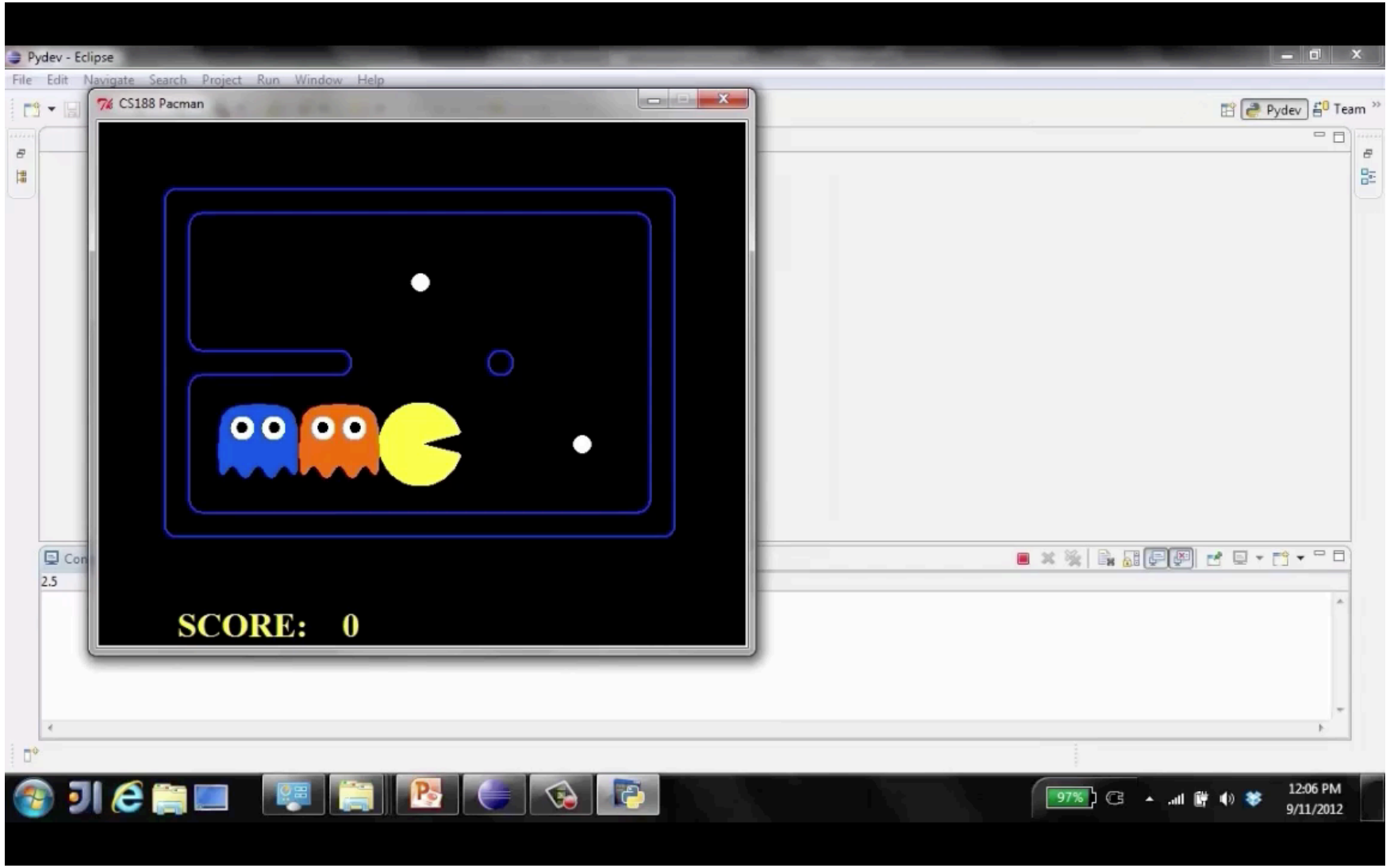


- A danger of replanning agents!
  - He knows his score will go up by eating the dot now (west, east)
  - He knows his score will go up just as much by eating the dot later (east, west)
  - There are no point-scoring opportunities after eating the dot (within the horizon, two here)
  - Therefore, waiting seems just as good as eating: he may go east, then back west in the next round of replanning!

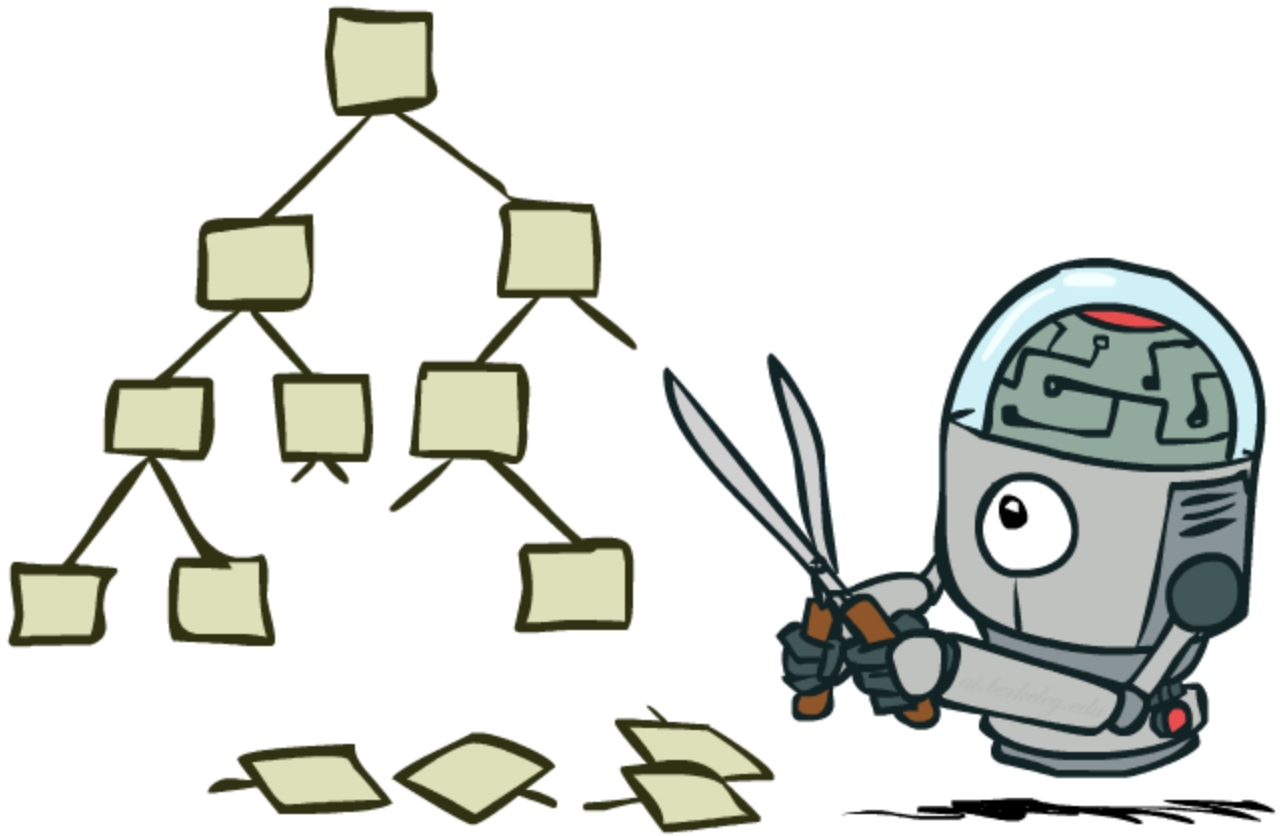
Video of Demo Smart Ghosts (Coordination)



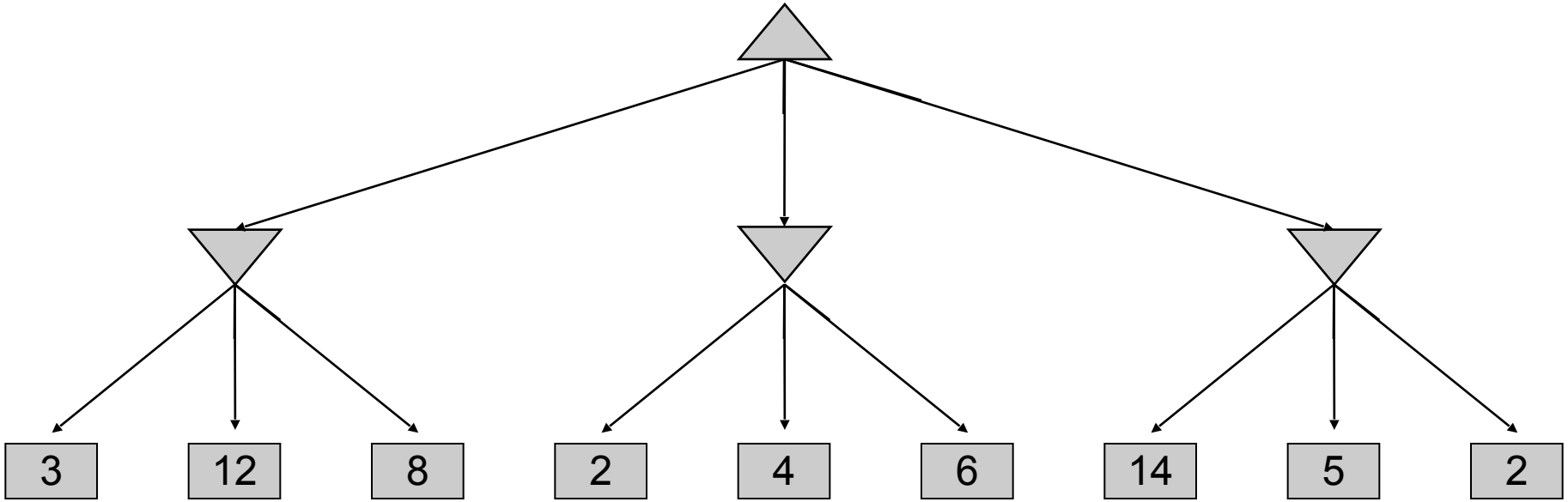
Video of Demo Smart Ghosts (Coordination) – Zoomed In



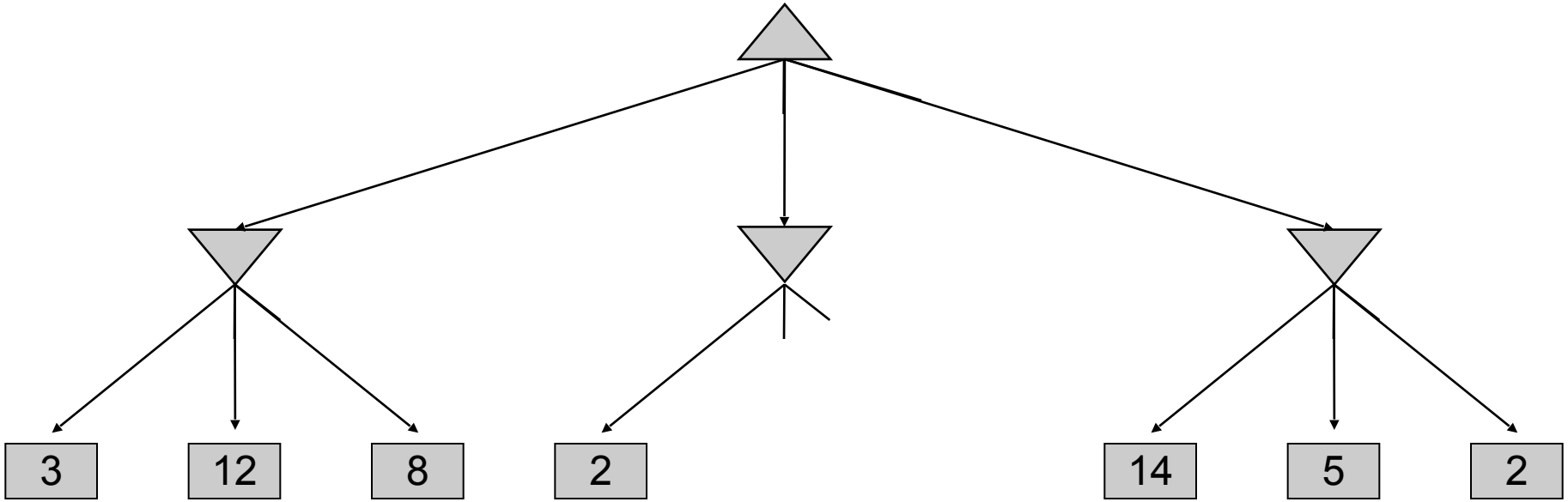
Game Tree Pruning



Minimax Example

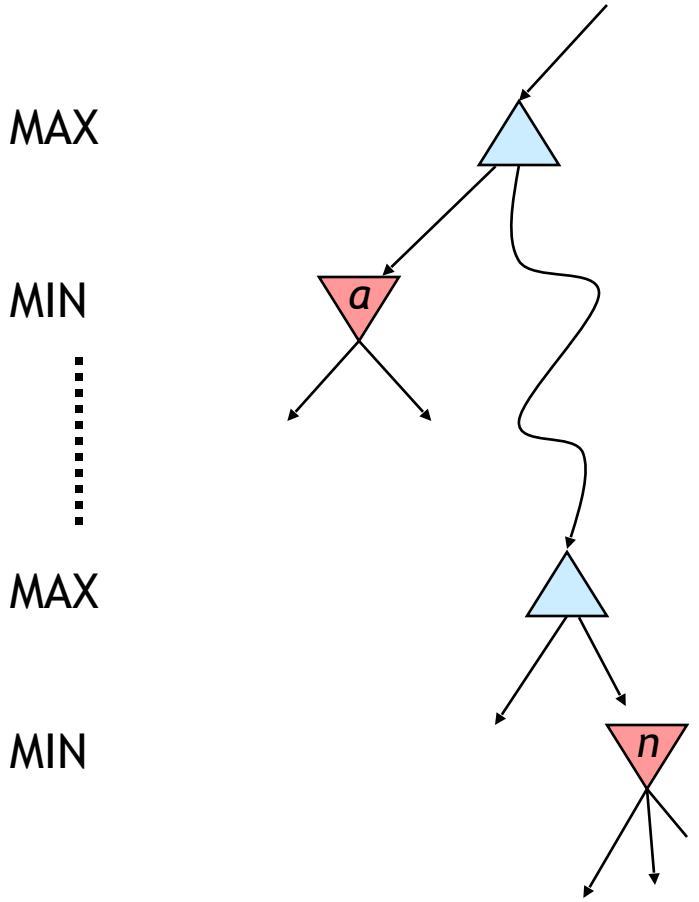


Minimax Pruning

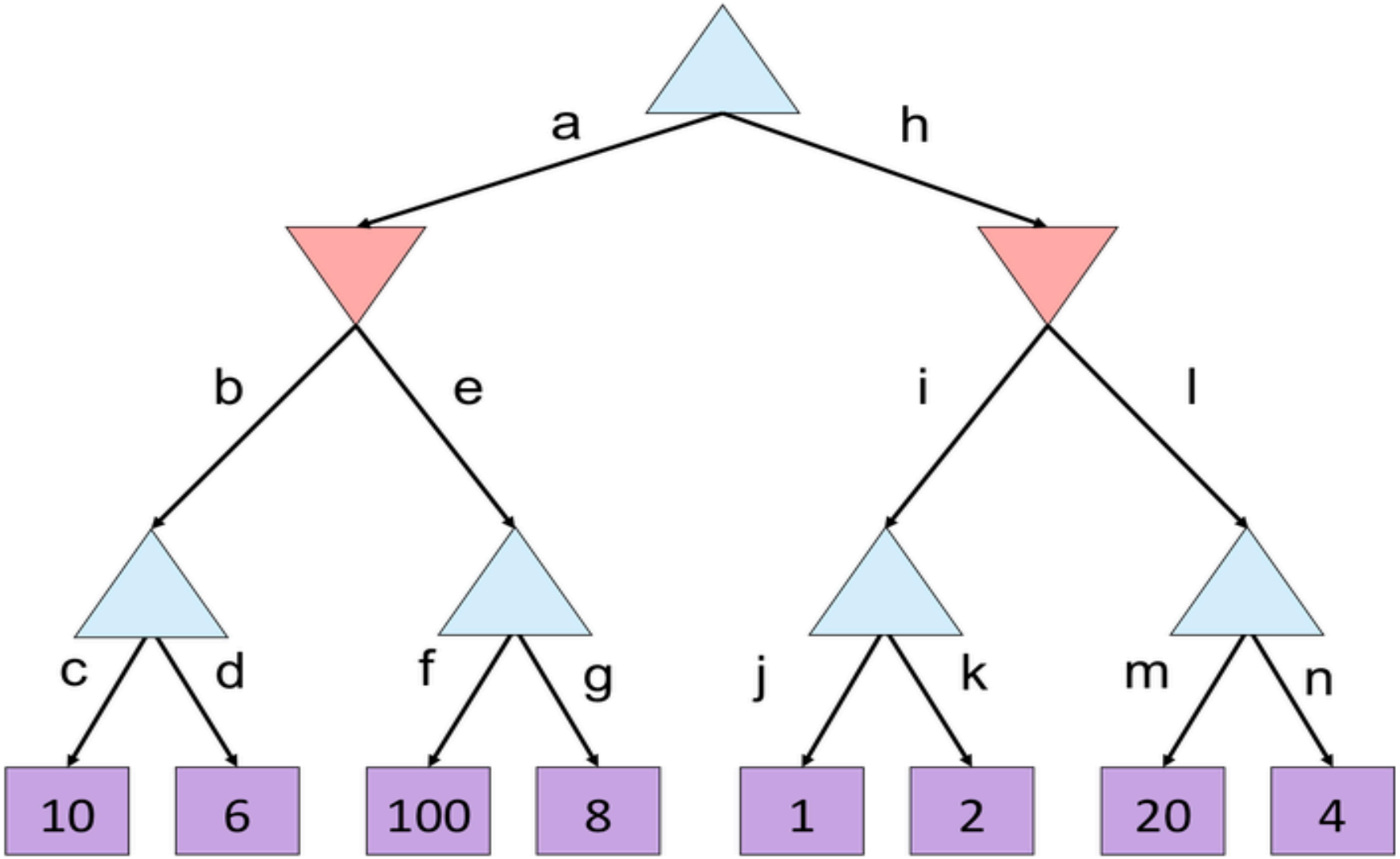


Alpha-Beta Pruning

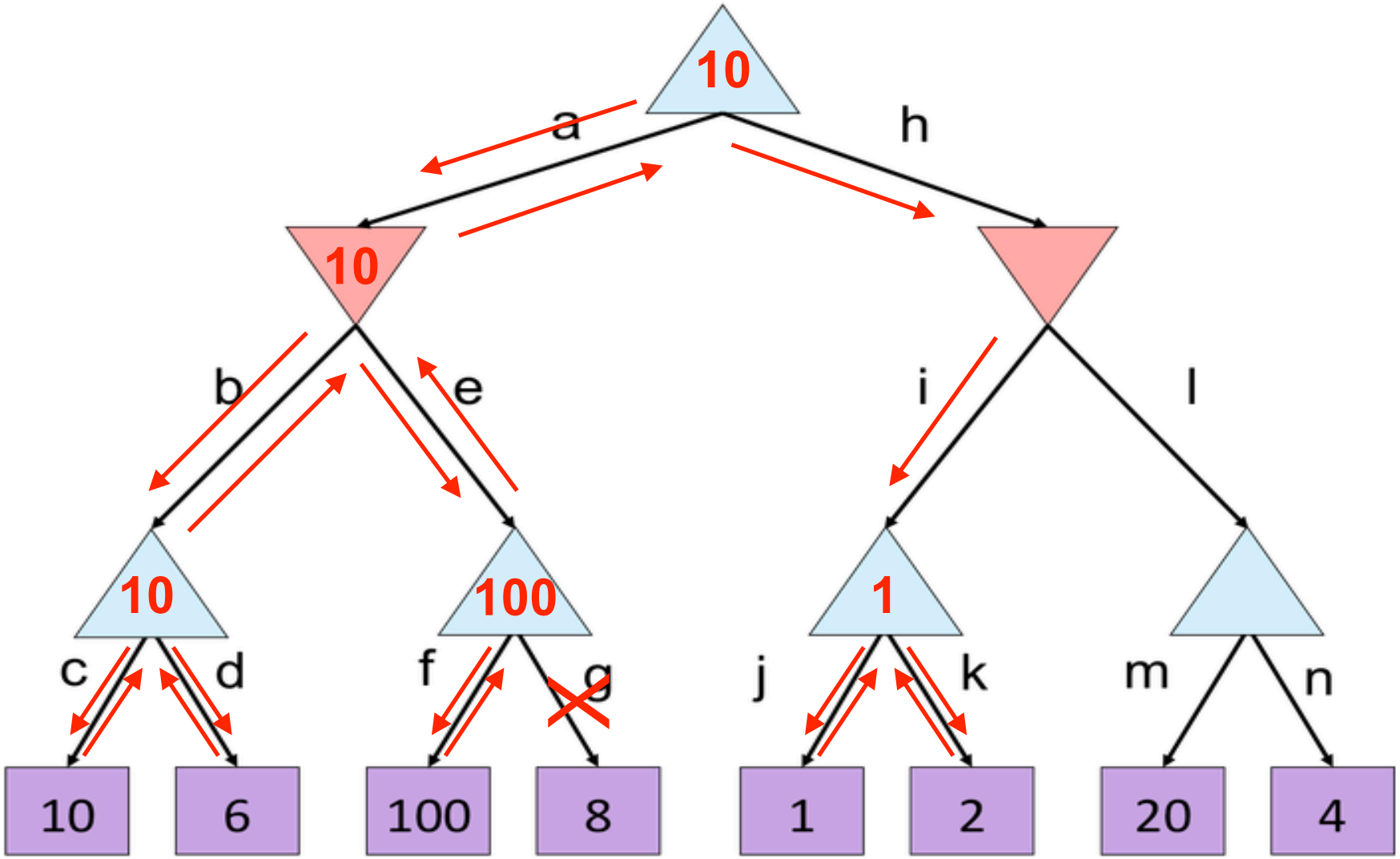
- General configuration (MIN version)
  - We're computing the MIN-VALUE at some node *n*
  - We're looping over *n*'s children
  - *n*'s estimate of the childrens' min is dropping
  - Who cares about *n*'s value? MAX
  - Let *a* be the best value that MAX can get at any choice point along the current path from the root
  - If *n* becomes worse than *a*, MAX will avoid it, so we can stop considering *n*'s other children (it's already bad enough that it won't be played)
- MAX version is symmetric



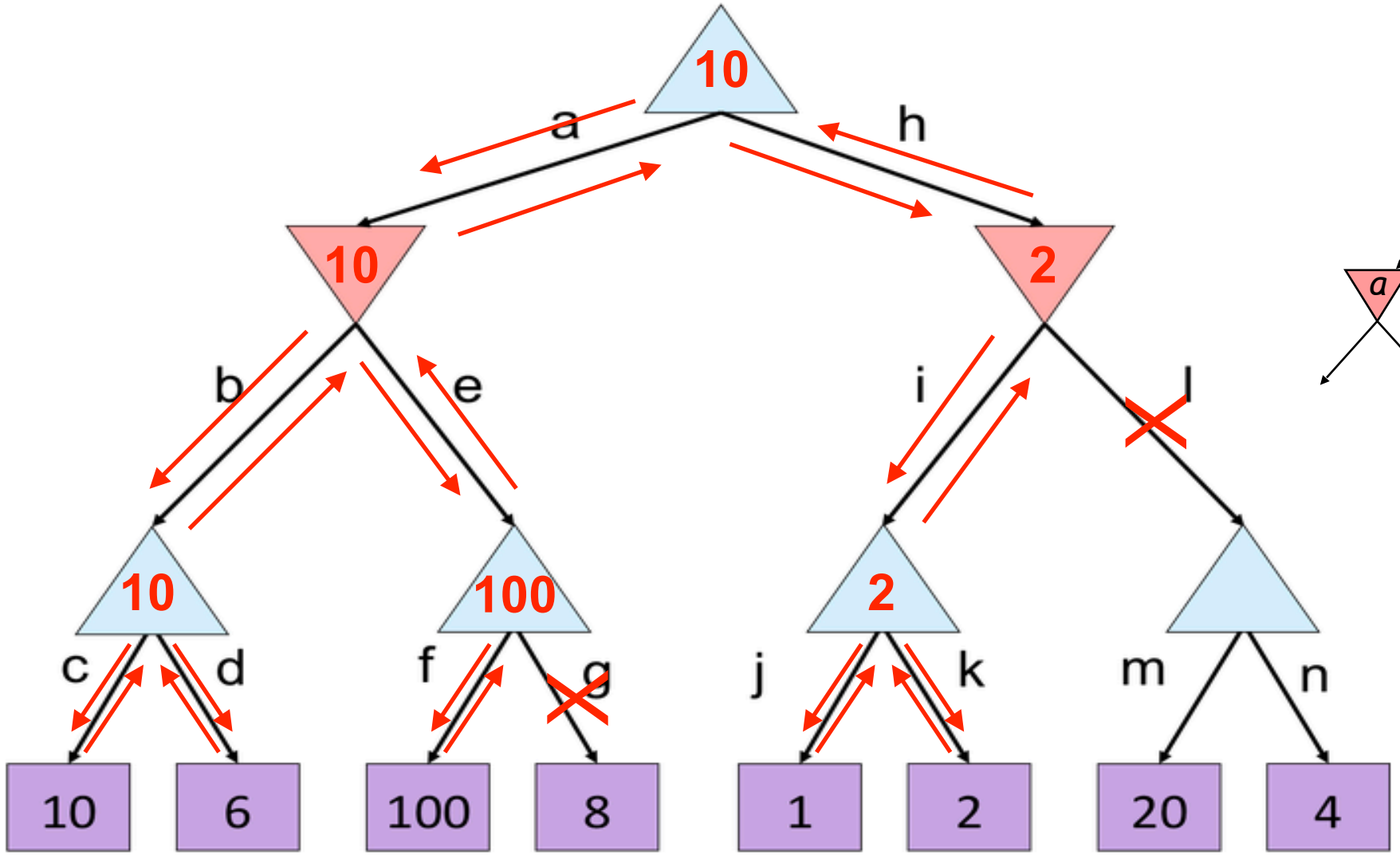
Alpha-Beta Quiz

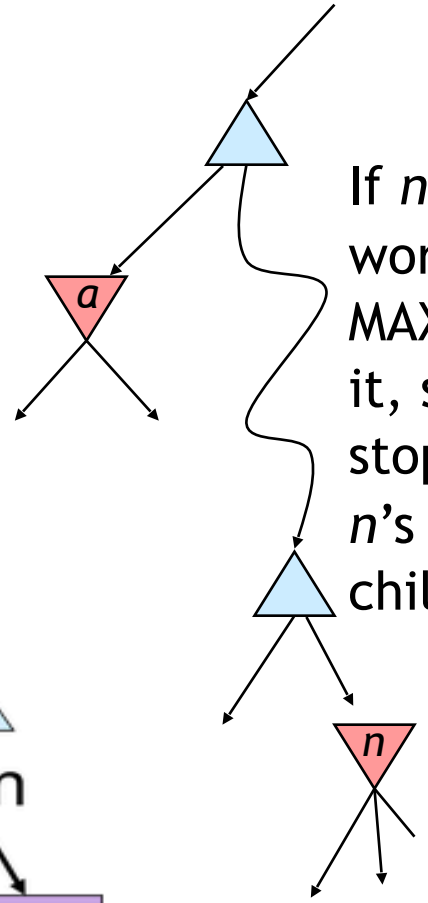


Alpha-Beta Quiz



Alpha-Beta Quiz





If *n* becomes worse than *a*, MAX will avoid it, so we can stop considering *n*'s other children

Minimax Implementation (Recap)

```
def value(state):
    if the state is a terminal state: return the state's utility
    if the next agent is MAX: return max-value(state)
    if the next agent is MIN: return min-value(state)
```

```
def max-value(state):
    initialize v = -∞
    for each successor of state:
        v = max(v, value(successor))
    return v
```

```
def min-value(state):
    initialize v = +∞
    for each successor of state:
        v = min(v, value(successor))
    return v
```

Alpha-Beta Implementation

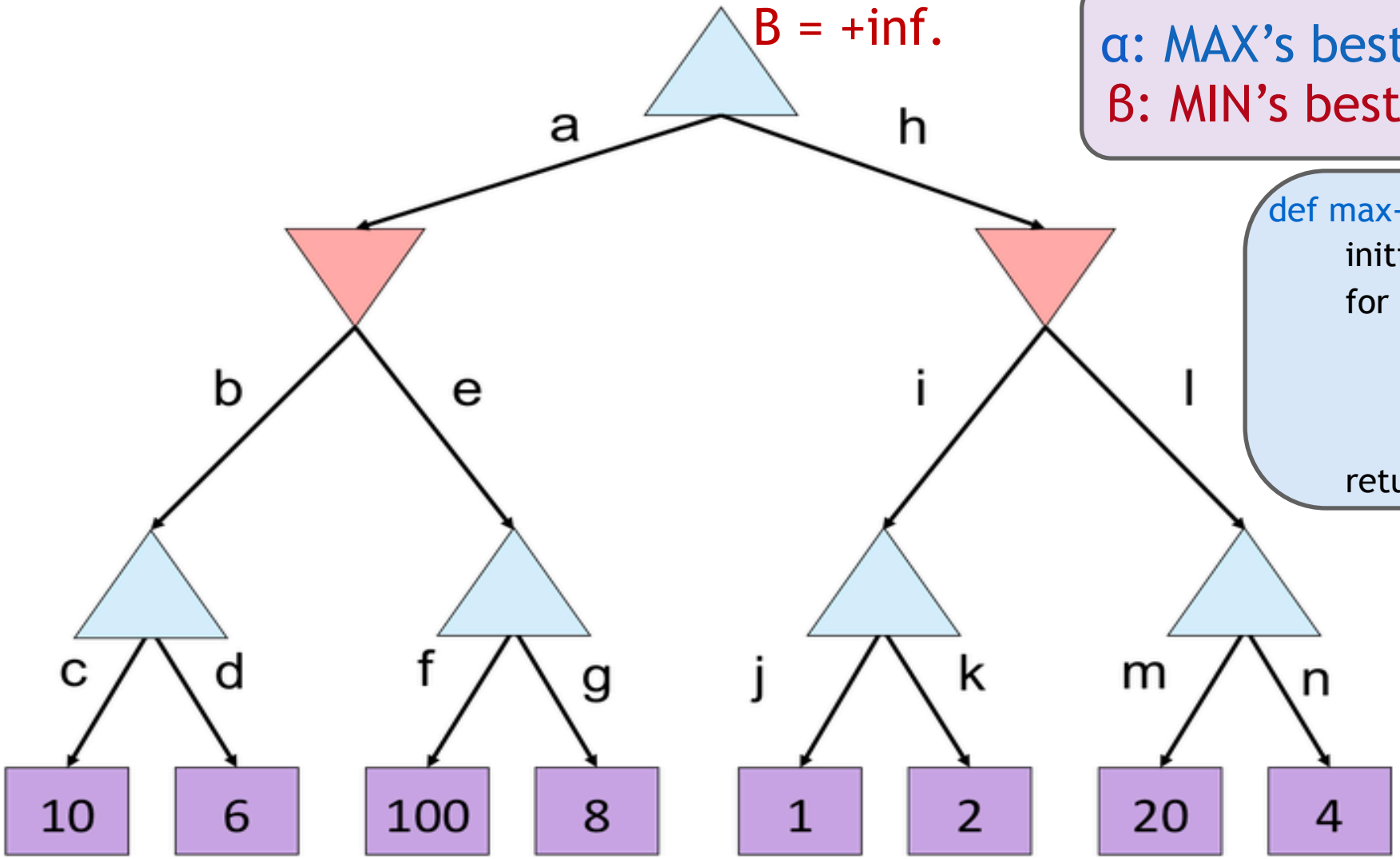
```
def value(state):
    if the state is a terminal state: return the state's utility
    if the next agent is MAX: return max-value(state)
    if the next agent is MIN: return min-value(state)
```

```
def max-value(state, α, β):
    initialize v = -∞
    for each successor of state:
        v = max(v, value(successor, α, β))
        if v ≥ β return v
        α = max(α, v)
    return v
```

```
def min-value(state , α, β):
    initialize v = +∞
    for each successor of state:
        v = min(v, value(successor, α, β))
        if v ≤ α return v
        β = min(β, v)
    return v
```

Alpha-Beta Quiz

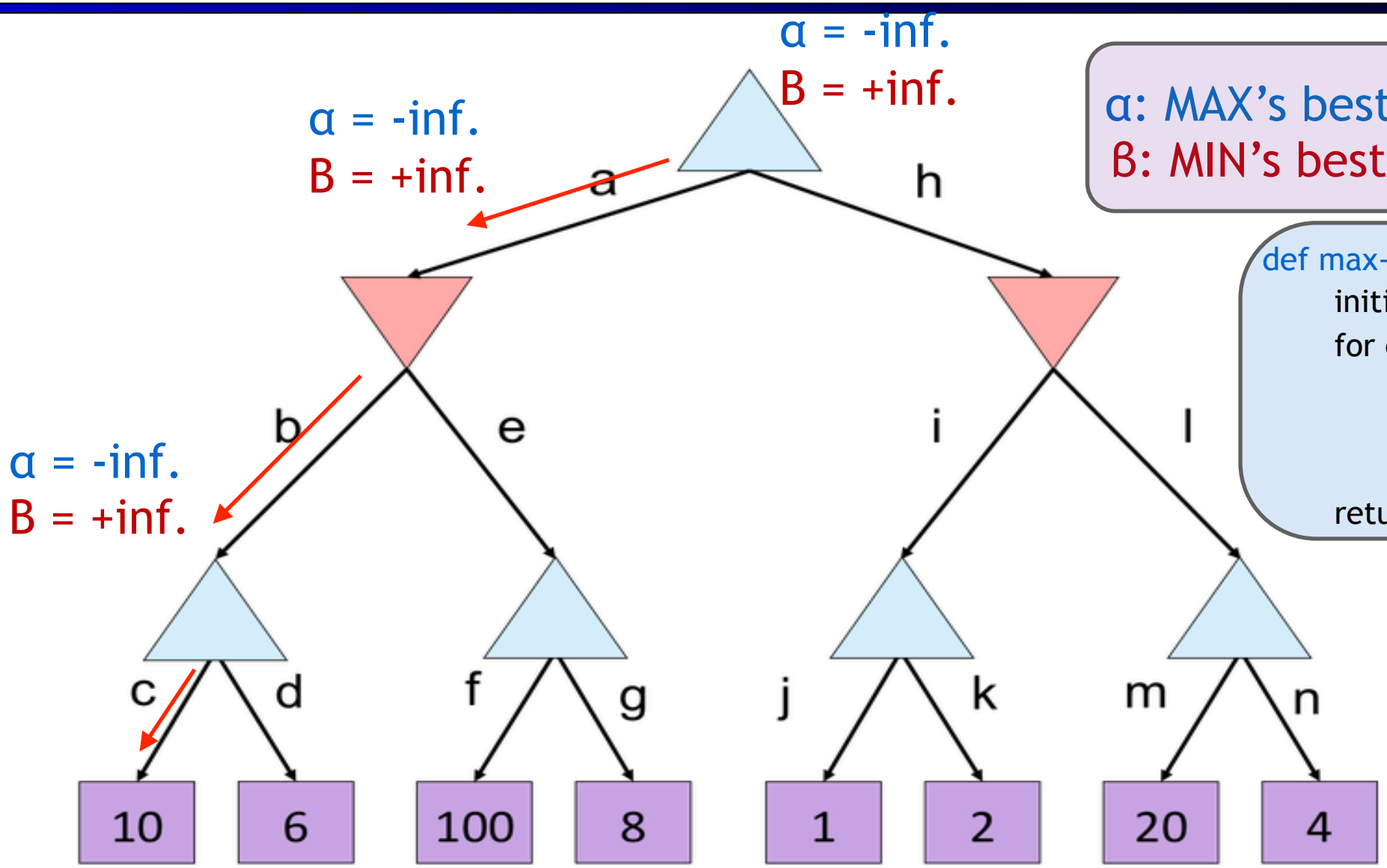
α = -inf.
Β = +inf.



α: MAX's best option on path to root
Β: MIN's best option on path to root

```
def max-value(state, α, Β):
    initialize v = -∞
    for each successor of state:
        v = max(v, value(successor, α, Β))
        if v ≥ Β return v
        α = max(α, v)
    return v
```

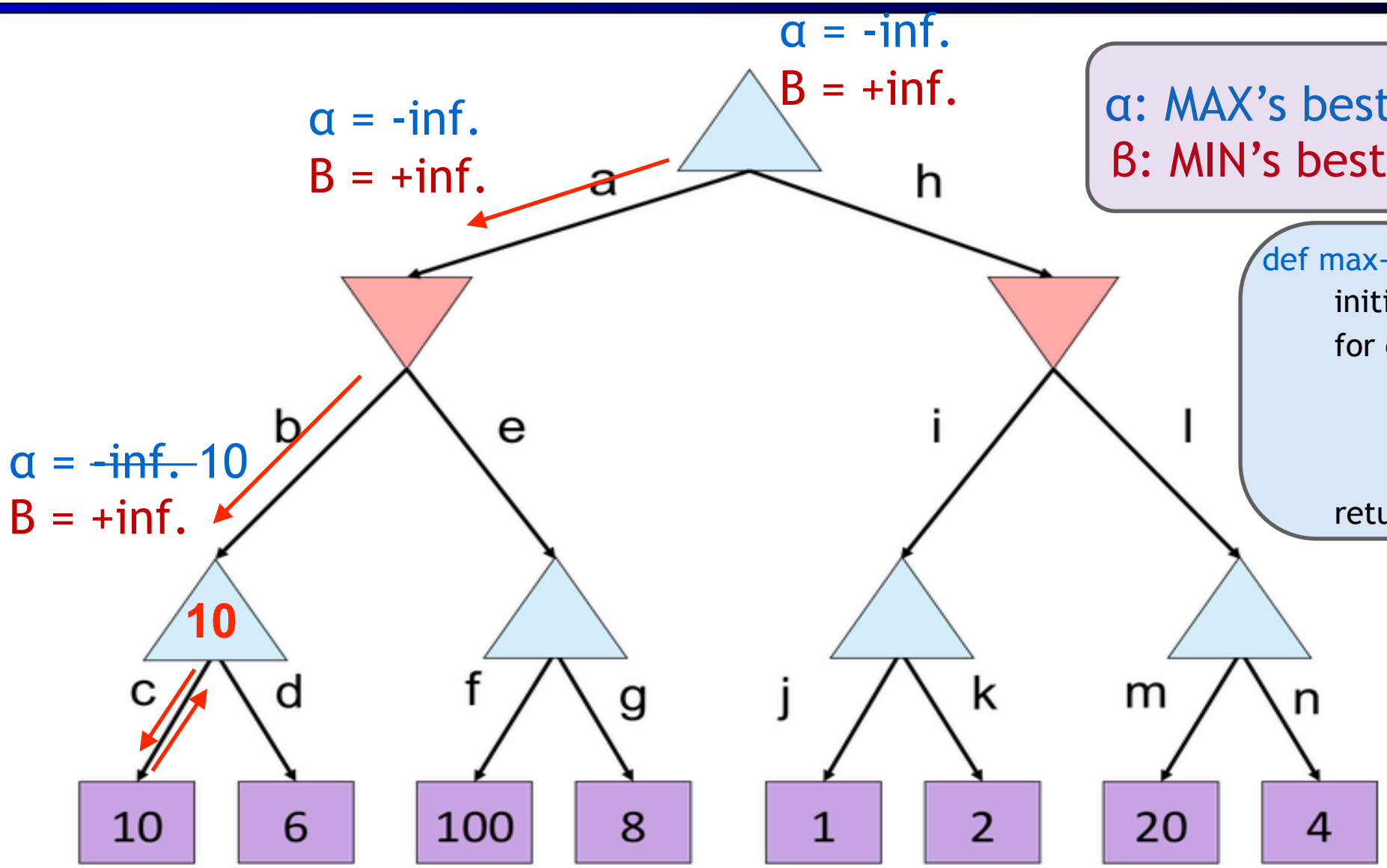
Alpha-Beta Quiz



α: MAX's best option on path to root
Β: MIN's best option on path to root

```
def max-value(state, α, Β):
    initialize v = -∞
    for each successor of state:
        v = max(v, value(successor, α, Β))
        if v ≥ Β return v
        α = max(α, v)
    return v
```

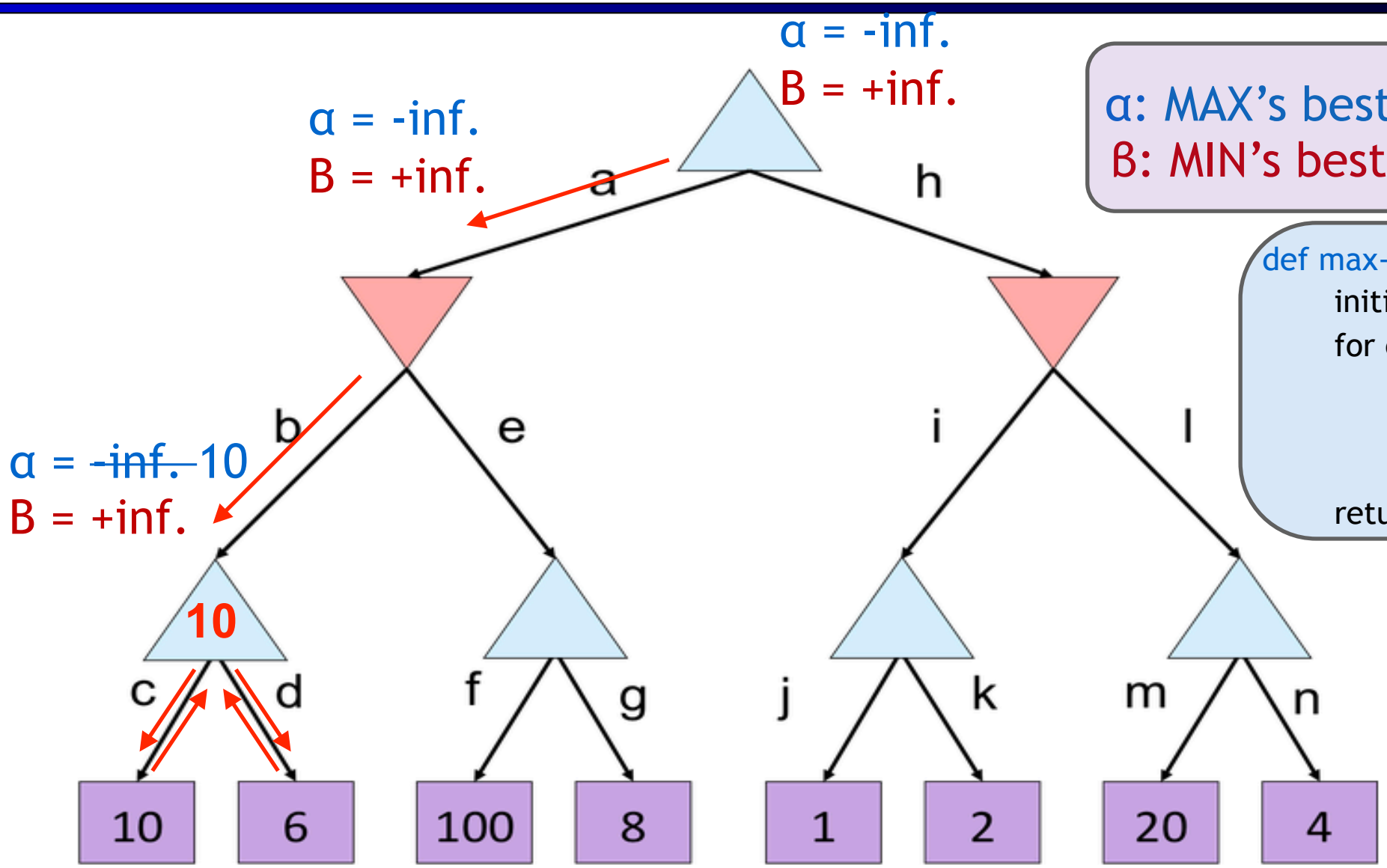
Alpha-Beta Quiz



α: MAX's best option on path to root
Β: MIN's best option on path to root

```
def max-value(state, α, Β):
    initialize v = -∞
    for each successor of state:
        v = max(v, value(successor, α, Β))
        if v ≥ Β return v
        α = max(α, v)
    return v
```

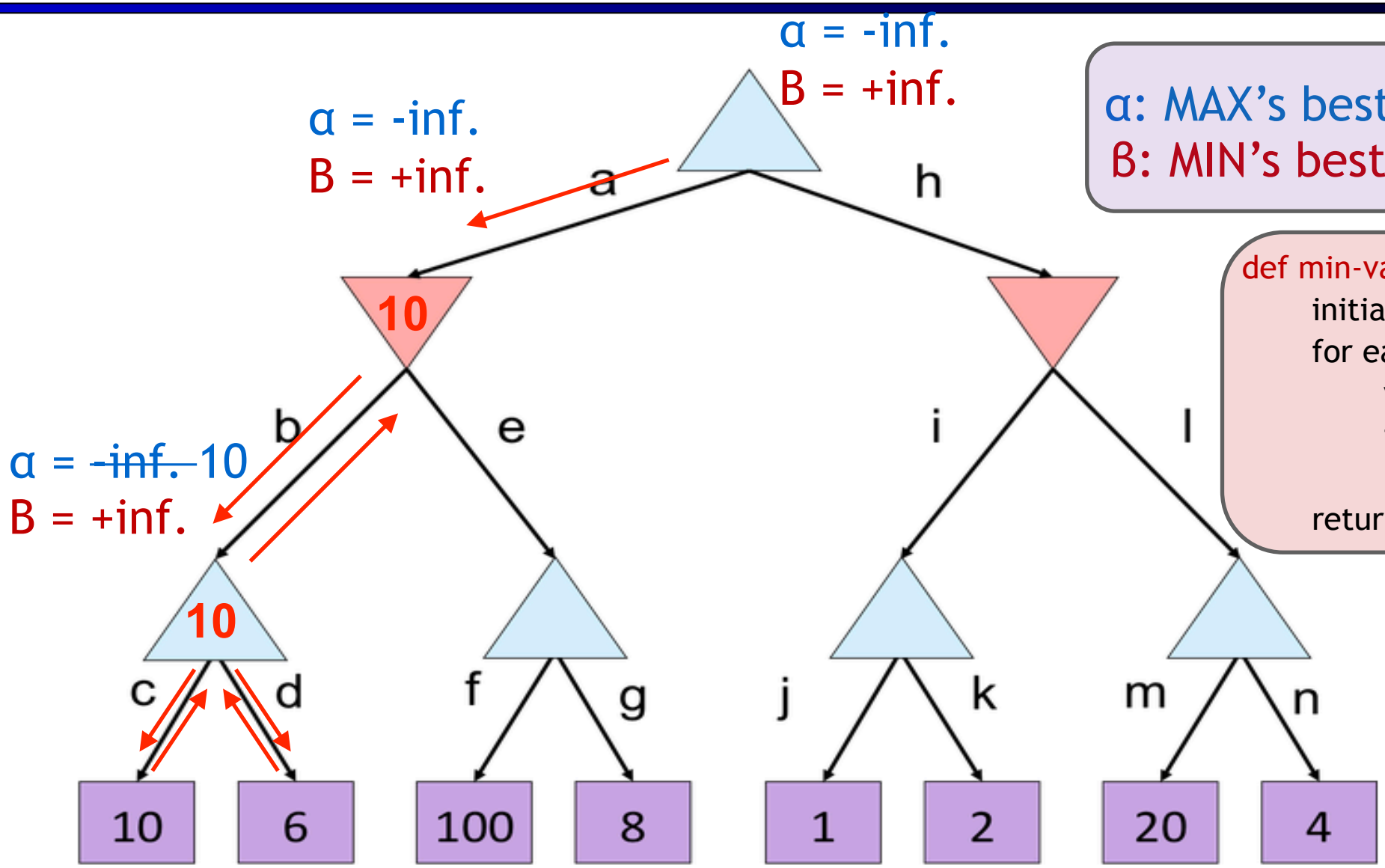
Alpha-Beta Quiz



α: MAX's best option on path to root
Β: MIN's best option on path to root

```
def max-value(state, α, Β):
    initialize v = -∞
    for each successor of state:
        v = max(v, value(successor, α, Β))
        if v ≥ Β return v
        α = max(α, v)
    return v
```

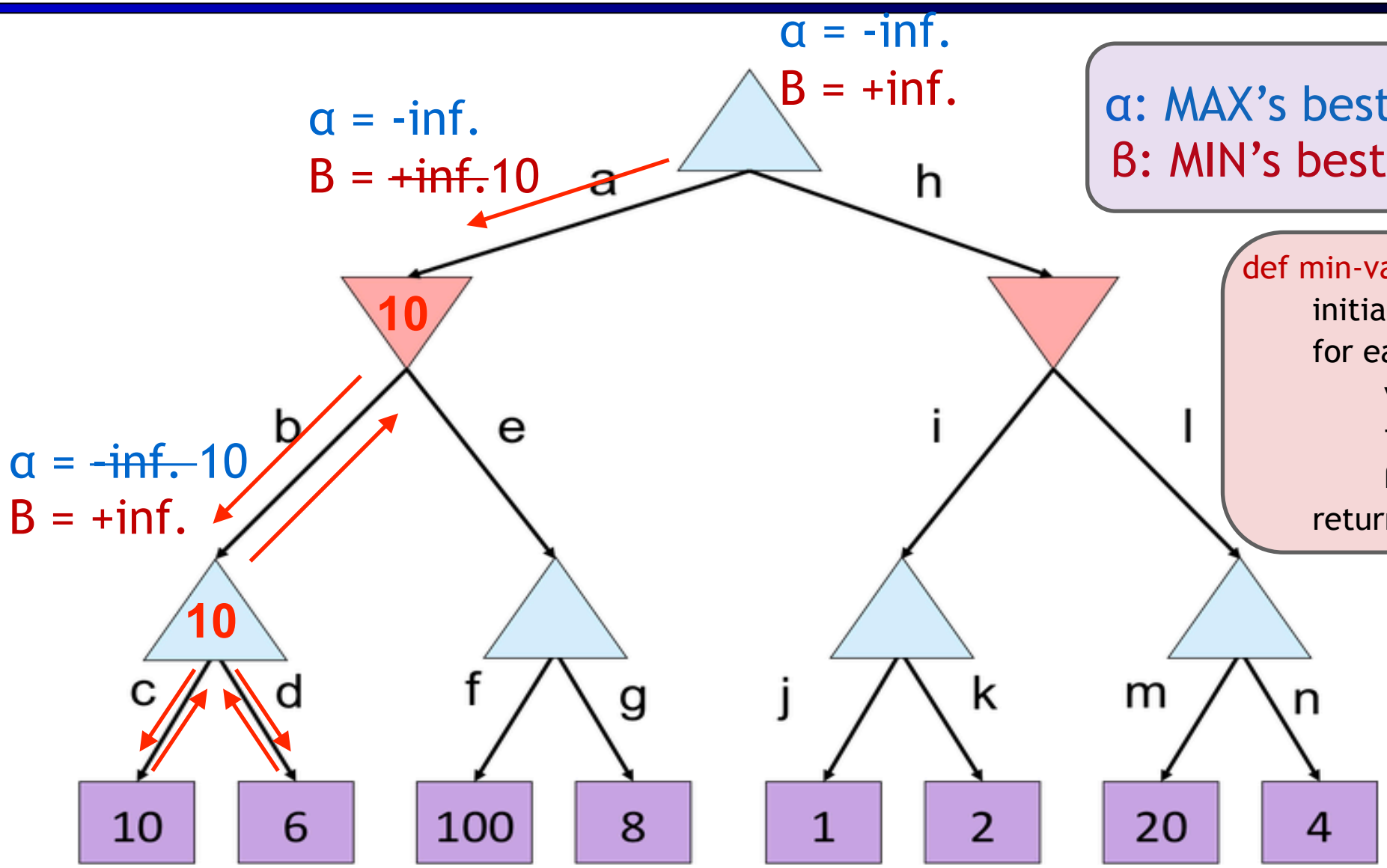
Alpha-Beta Quiz



α: MAX's best option on path to root
Β: MIN's best option on path to root

```
def min-value(state , α, Β):
    initialize v = +∞
    for each successor of state:
        v = min(v, value(successor, α, Β))
        if v ≤ α return v
        Β = min(Β, v)
    return v
```

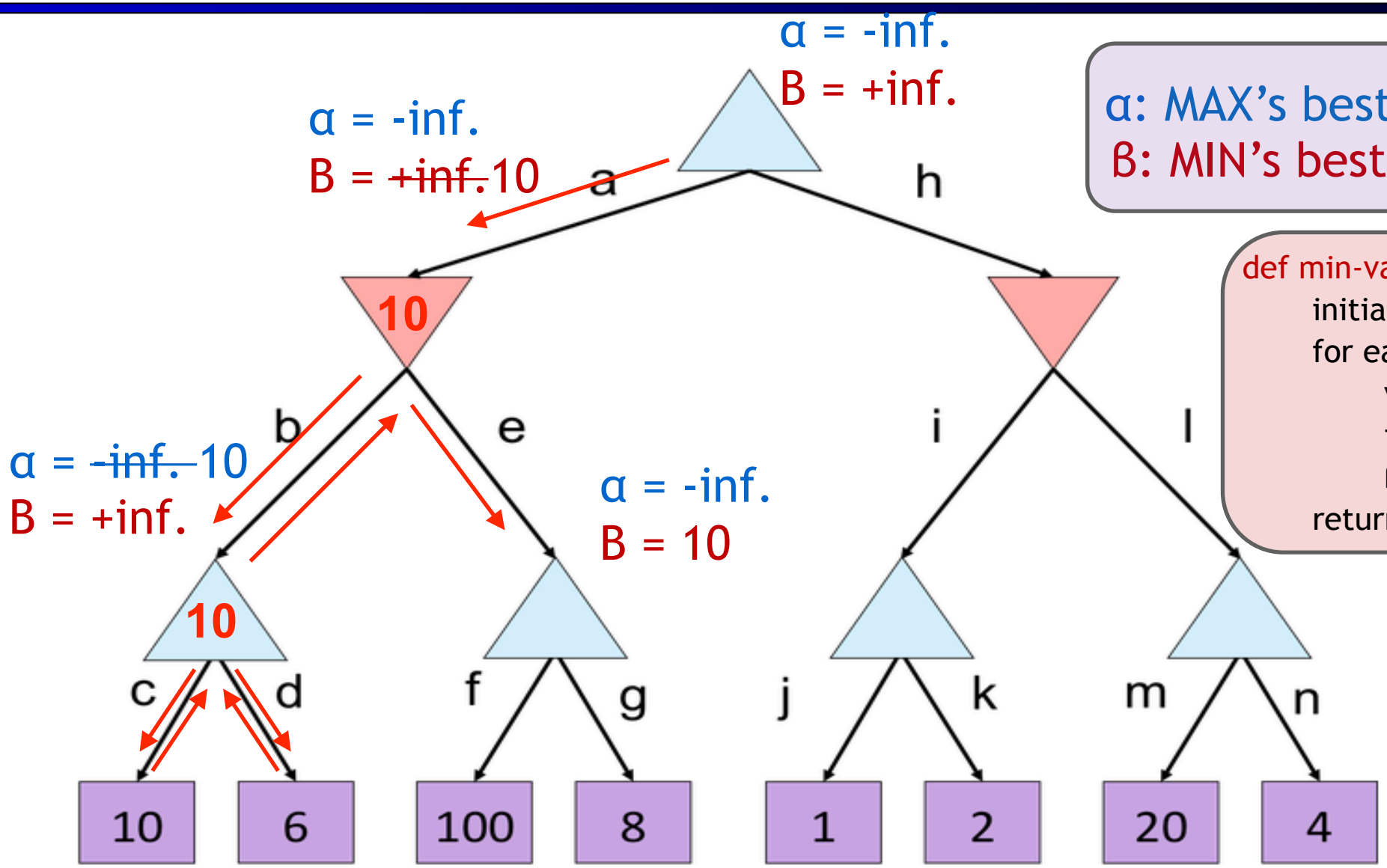
Alpha-Beta Quiz



α: MAX's best option on path to root
Β: MIN's best option on path to root

```
def min-value(state , α, Β):
    initialize v = +∞
    for each successor of state:
        v = min(v, value(successor, α, Β))
        if v ≤ α return v
        Β = min(Β, v)
    return v
```

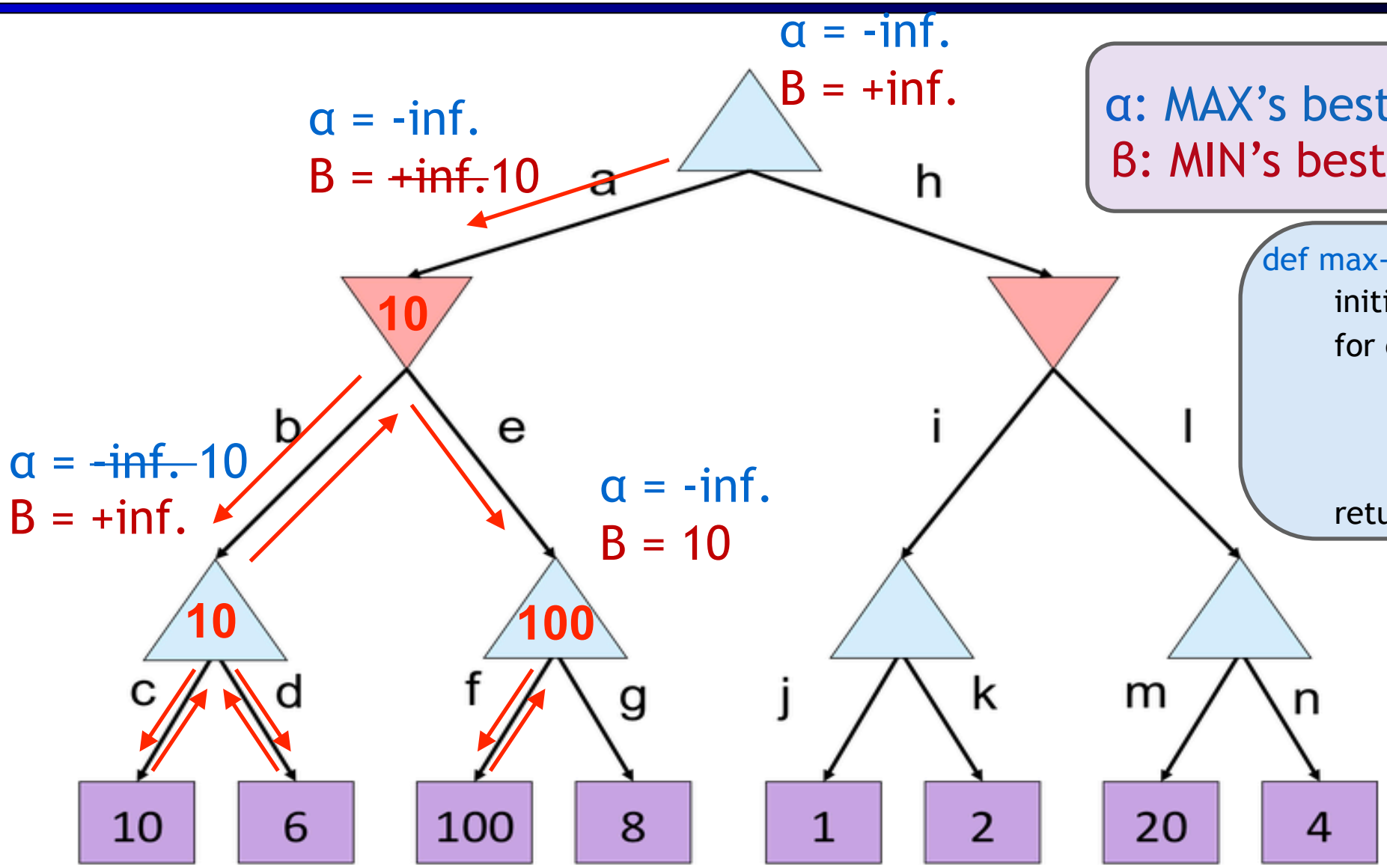
Alpha-Beta Quiz



α: MAX's best option on path to root
Β: MIN's best option on path to root

```
def min-value(state , α, Β):
    initialize v = +∞
    for each successor of state:
        v = min(v, value(successor, α, Β))
        if v ≤ α return v
        Β = min(Β, v)
    return v
```

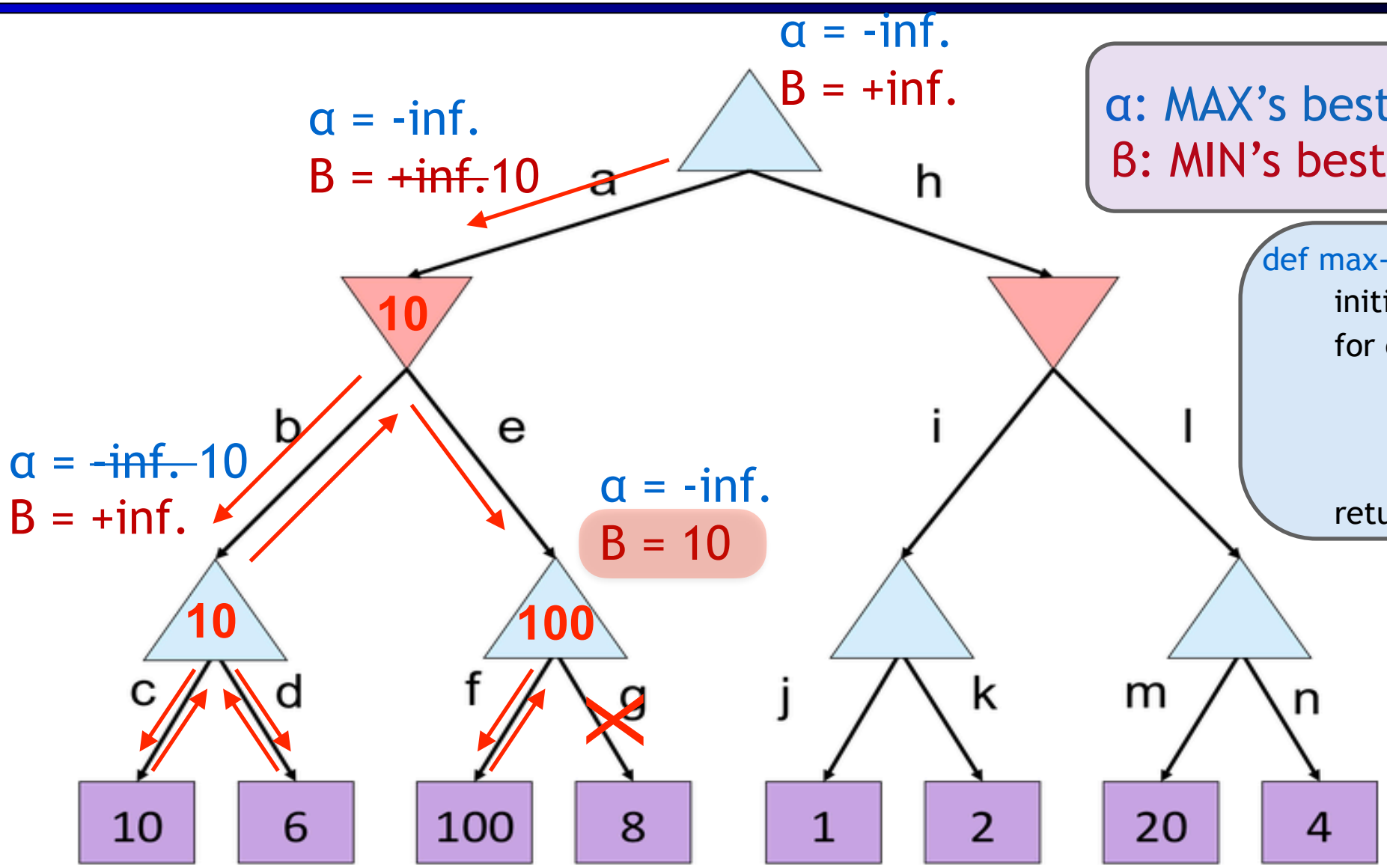
Alpha-Beta Quiz



α: MAX's best option on path to root
Β: MIN's best option on path to root

```
def max-value(state, α, Β):
    initialize v = -∞
    for each successor of state:
        v = max(v, value(successor, α, Β))
        if v ≥ Β return v
        α = max(α, v)
    return v
```

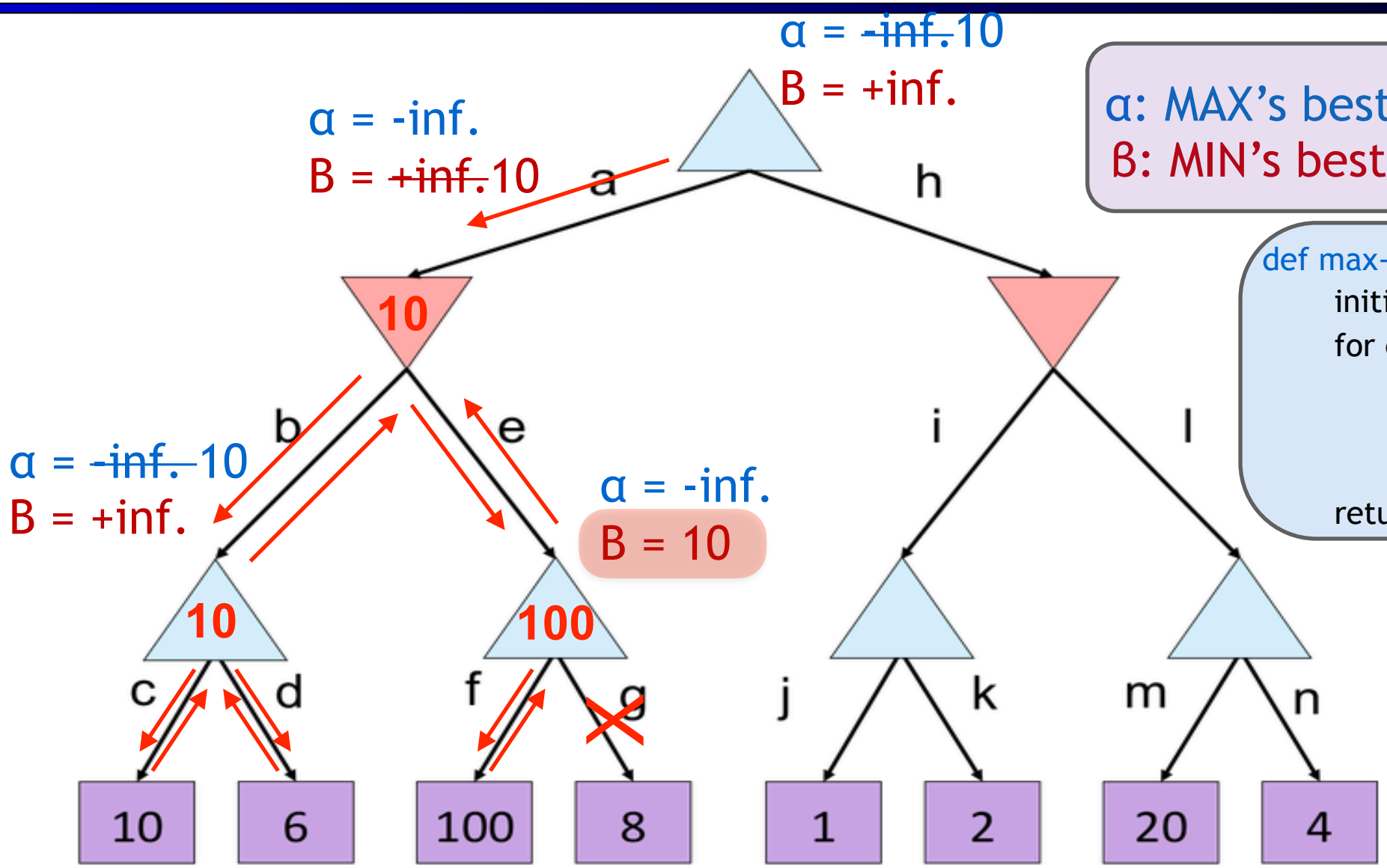
Alpha-Beta Quiz



α: MAX's best option on path to root
Β: MIN's best option on path to root

```
def max-value(state, α, Β):
    initialize v = -∞
    for each successor of state:
        v = max(v, value(successor, α, Β))
        if v ≥ Β return v
        α = max(α, v)
    return v
```

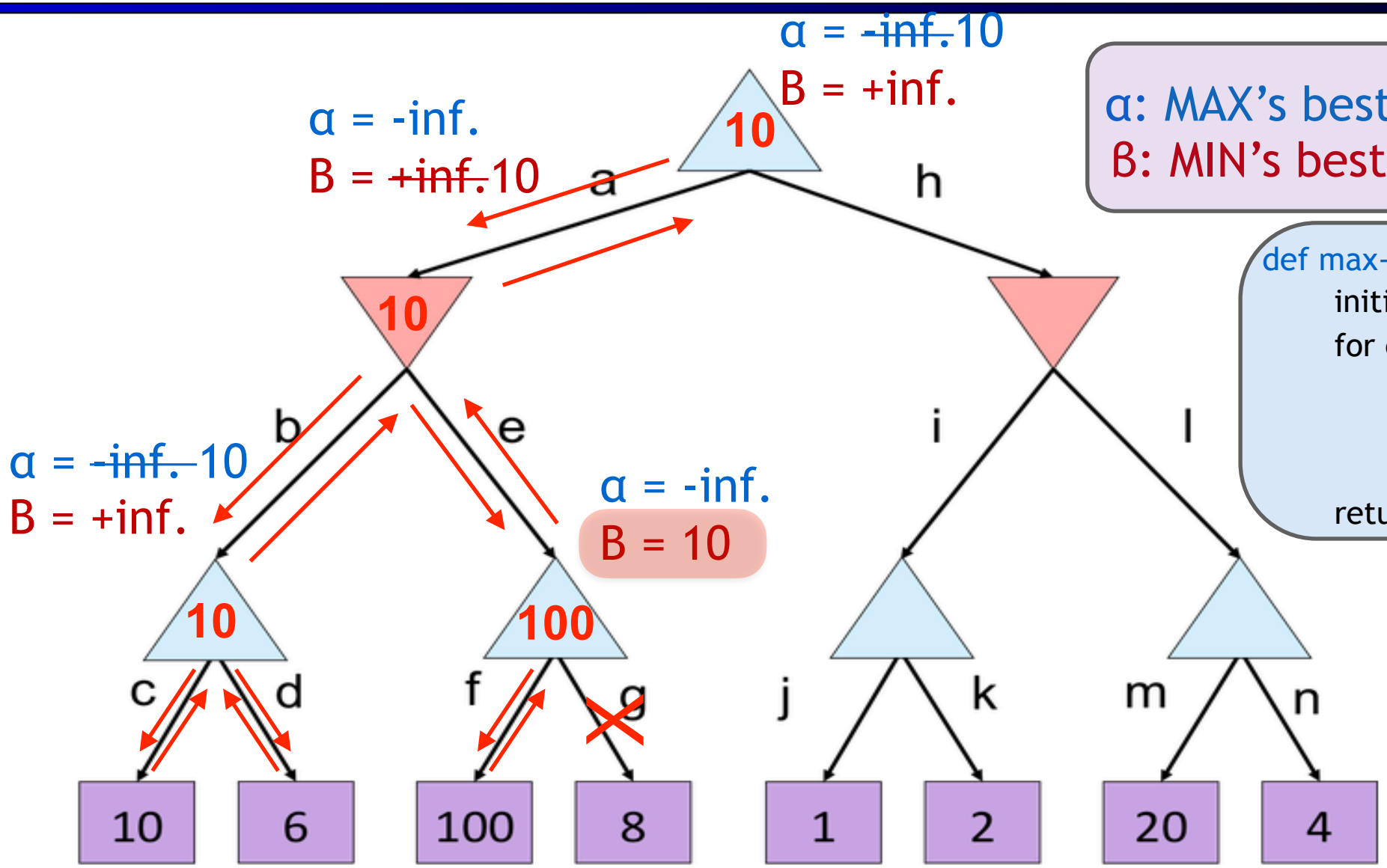
Alpha-Beta Quiz



α: MAX's best option on path to root
Β: MIN's best option on path to root

```
def max-value(state, α, Β):
    initialize v = -∞
    for each successor of state:
        v = max(v, value(successor, α, Β))
        if v ≥ Β return v
        α = max(α, v)
    return v
```

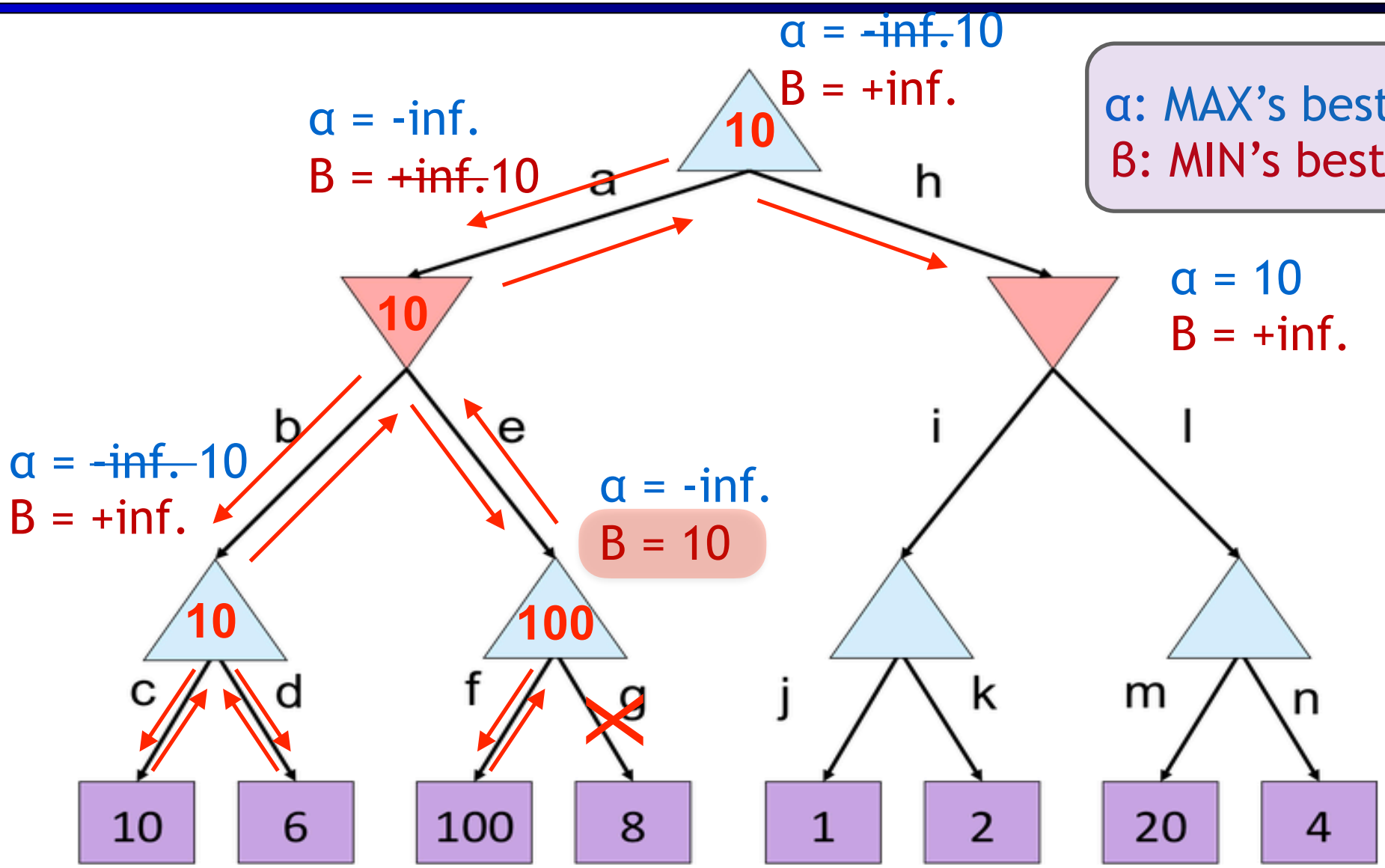
Alpha-Beta Quiz



α: MAX's best option on path to root
Β: MIN's best option on path to root

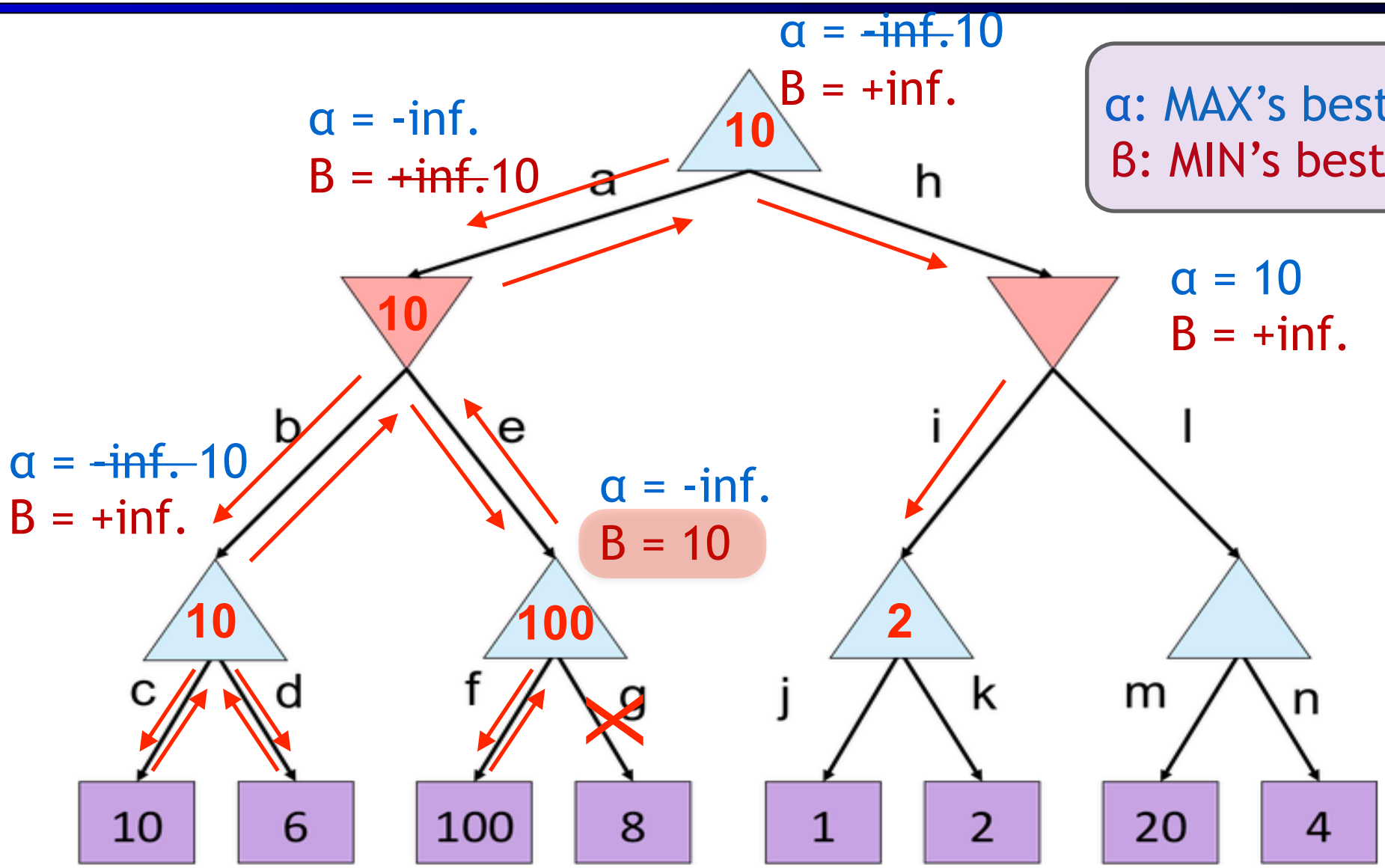
```
def max-value(state, α, Β):
    initialize v = -∞
    for each successor of state:
        v = max(v, value(successor, α, Β))
        if v ≥ Β return v
        α = max(α, v)
    return v
```

Alpha-Beta Quiz



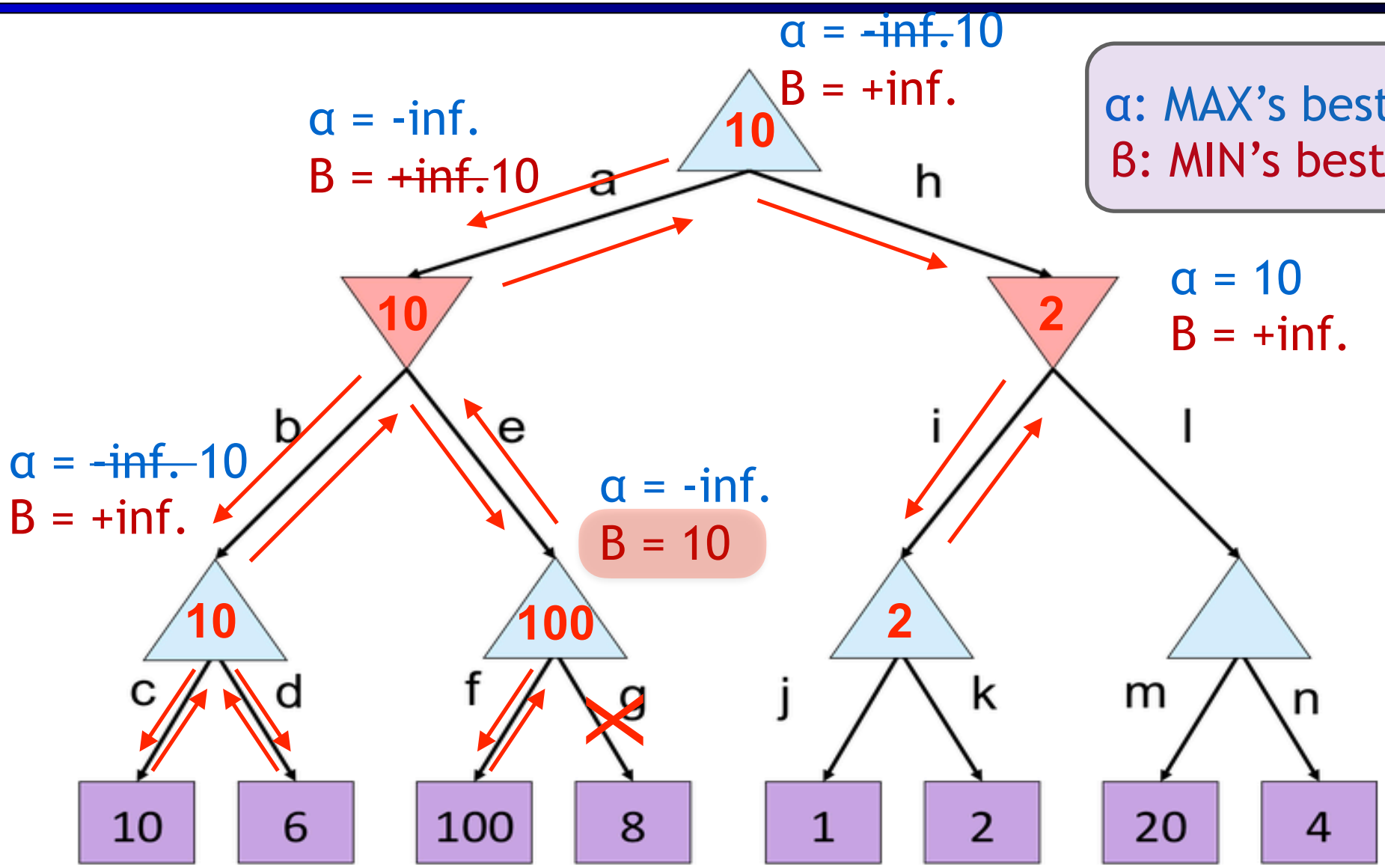
α: MAX's best option on path to root
Β: MIN's best option on path to root

Alpha-Beta Quiz



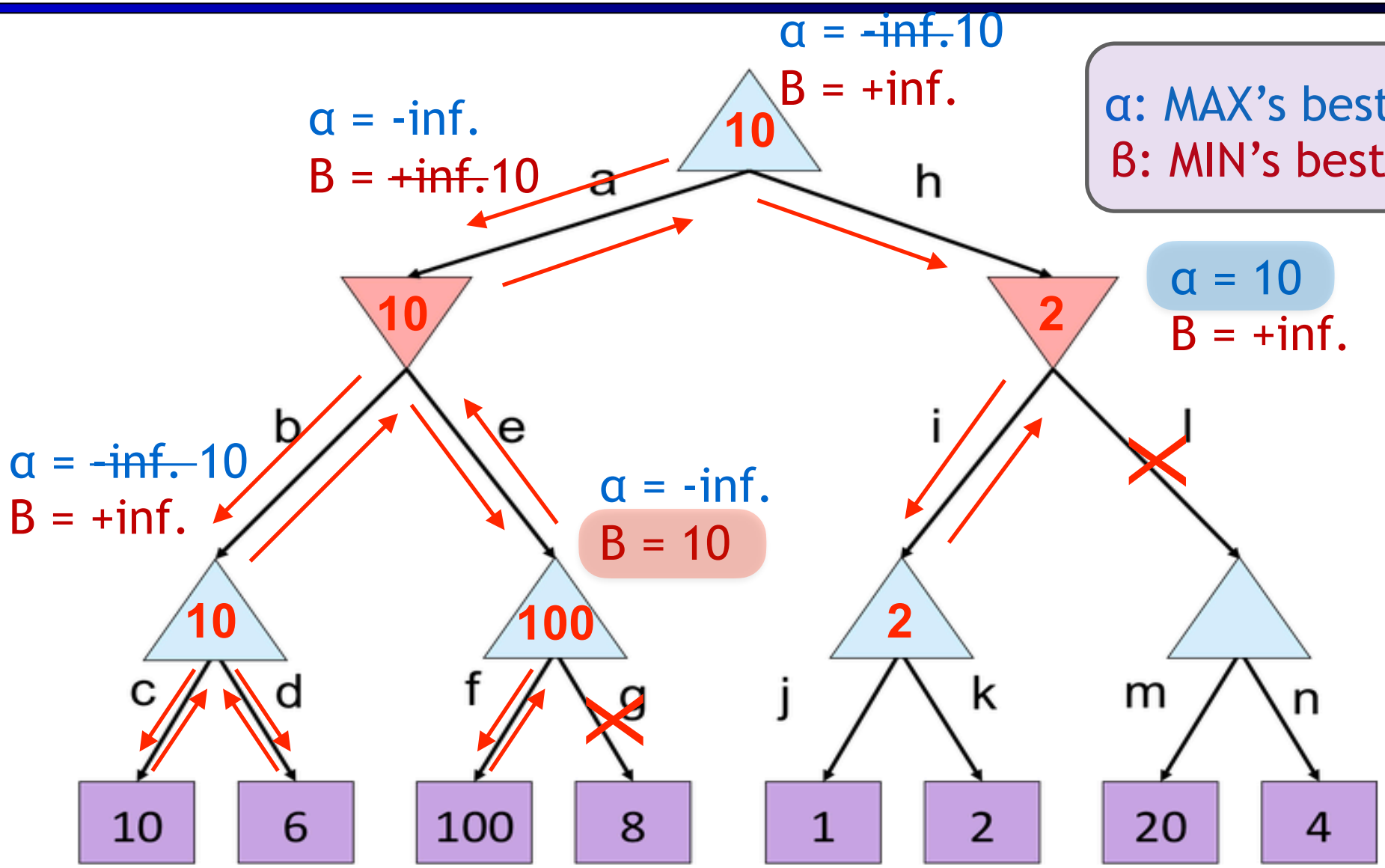
α: MAX's best option on path to root
Β: MIN's best option on path to root

Alpha-Beta Quiz



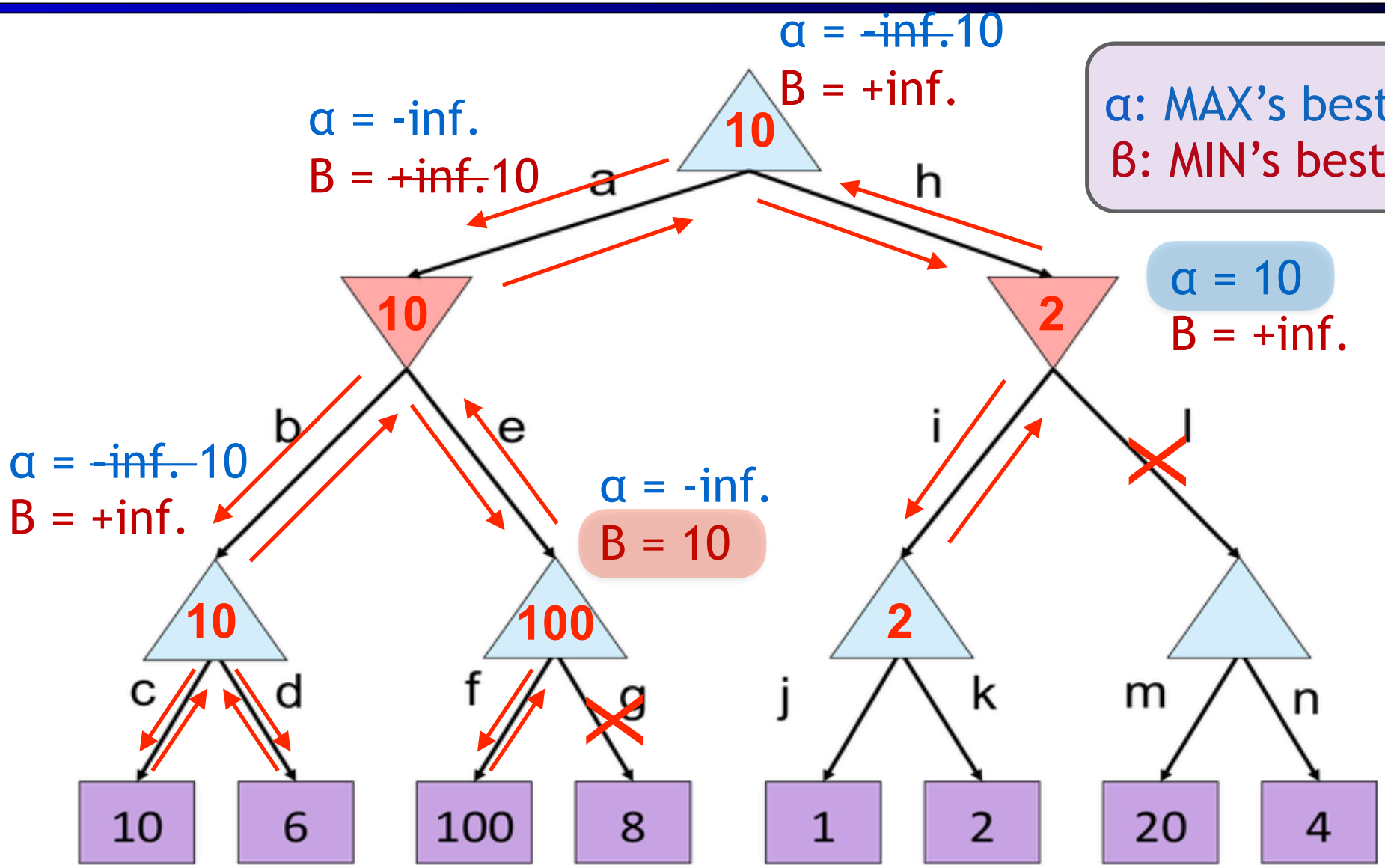
α: MAX's best option on path to root
Β: MIN's best option on path to root

Alpha-Beta Quiz



α: MAX's best option on path to root
Β: MIN's best option on path to root

Alpha-Beta Quiz



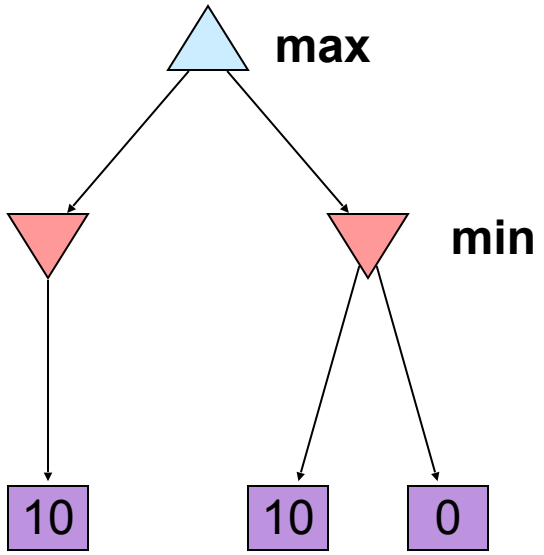
α: MAX's best option on path to root
Β: MIN's best option on path to root

Alpha-Beta Pruning Properties

- This pruning has no effect on minimax value computed for the root!

- Values of intermediate nodes might be wrong
  - Important: children of the root may have the wrong value
  - So the most naïve version won't let you do action selection

```
def min-value(state , α, β):
    initialize v = +∞
    for each successor of state:
        v = min(v, value(successor, α, β))
        if v ≤ α return v
        β = min(β, v)
    return v
```

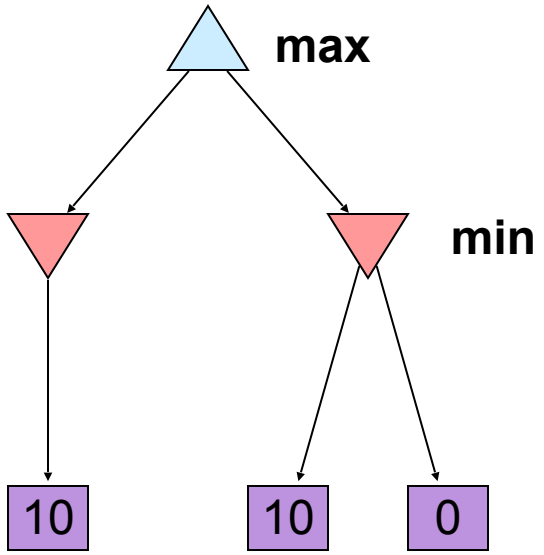


Alpha-Beta Pruning Properties

- This pruning has no effect on minimax value computed for the root!

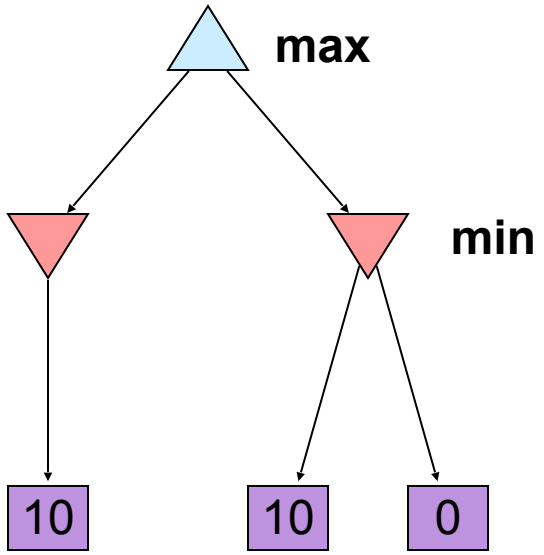
- Values of intermediate nodes might be wrong
  - Important: children of the root may have the wrong value
  - So the most naïve version won't let you do action selection

```
def min-value(state , α, β):
    initialize v = +∞
    for each successor of state:
        v = min(v, value(successor, α, β))
        if v < α return v
        β = min(β, v)
    return v
```



Alpha-Beta Pruning Properties

- This pruning has no effect on minimax value computed for the root!
- Values of intermediate nodes might be wrong
  - Important: children of the root may have the wrong value
  - So the most naïve version won't let you do action selection
- What if we ask what action to take? Have to be careful!!
  - Soln. 1: separate alpha-beta for each child of the root node, and we continue to prune with equality
  - Soln. 2: prune with inequality
  - Soln. 3: alter alpha-beta just at the root to only prune with inequality



Next Time: Uncertainty!