Algorithms and Networking for Computer Games

Chapter 4: Game Trees

http://www.wiley.com/go/smed

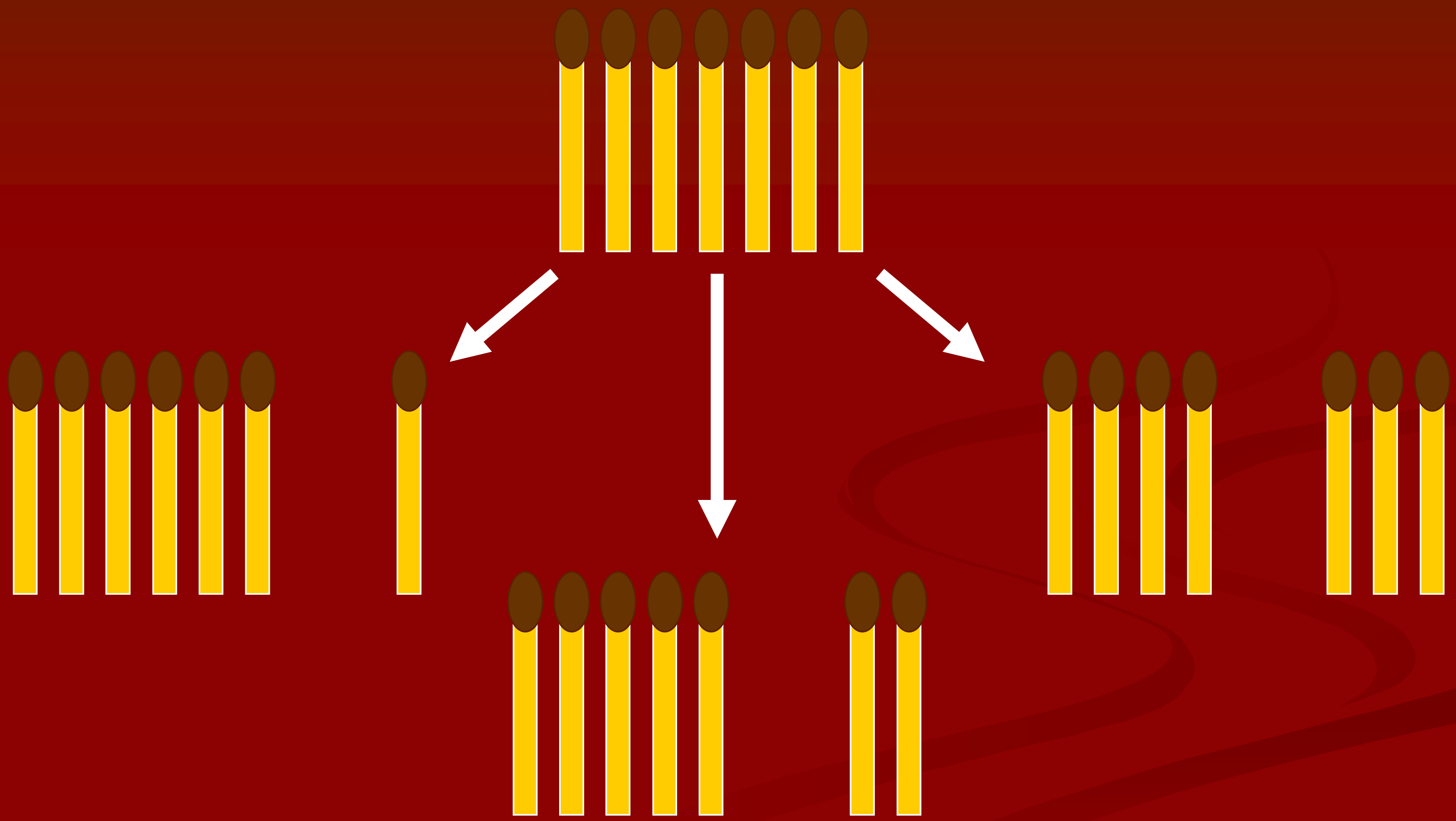
Game types

- perfect information games
  - no hidden information
- two-player, perfect information games
  - Noughts and Crosses
  - Chess
  - Go
- imperfect information games
  - Poker
  - Backgammon
  - Monopoly
- zero-sum property
  - one player's gain equals another player's loss

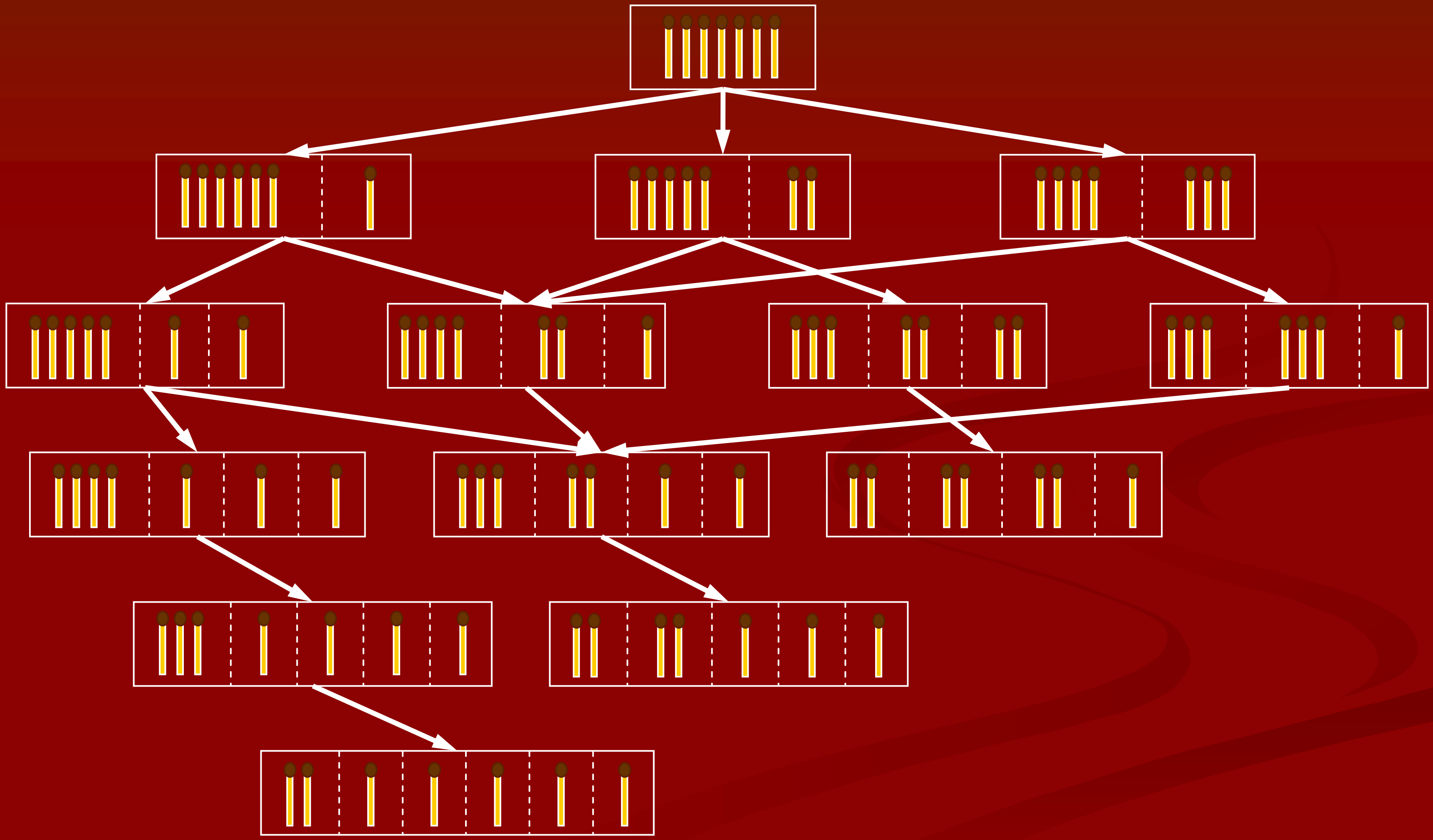
Game tree

- all possible plays of two-player, perfect information games can be represented with a game tree
  - nodes: positions (or states)
  - edges: moves
- players: MAX (has the first move) and MIN
- ply = the length of the path between two nodes
  - MAX has even plies counting from the root node
  - MIN has odd plies counting from the root node

Division Nim with seven matches



Game tree for Division Nim



Problem statement

Given a node *v* in a game tree

find a winning strategy for MAX (or MIN) from *v*

or (equivalently)

show that MAX (or MIN) can force a win from *v*

Minimax

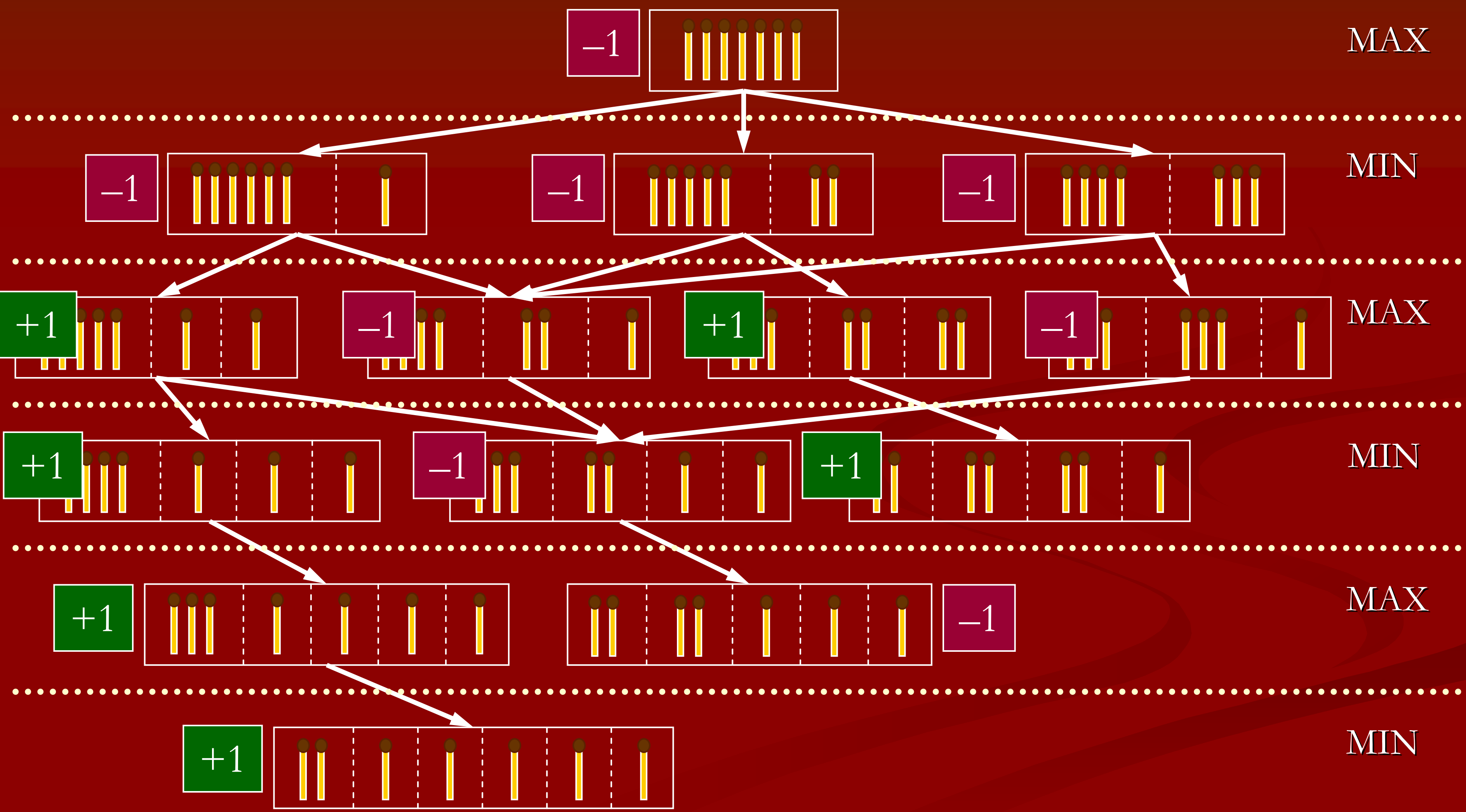
- assumption: players are rational and try to win
- given a game tree, we know the outcome in the leaves
  - assign the leaves to win, draw, or loss (or a numeric value like +1, 0, –1) according to MAX's point of view
- at nodes one ply above the leaves, we choose the best outcome among the children (which are leaves)
  - MAX: win if possible; otherwise, draw if possible; else loss
  - MIN: loss if possible; otherwise, draw if possible; else win
- recurse through the nodes until in the root

Minimax rules

1. If the node is labelled to MAX, assign it to the maximum value of its children.
2. If the node is labelled to MIN, assign it to the minimum value of its children.

- MIN minimizes, MAX maximizes → minimax

Game tree with valued nodes



Analysis

- simplifying assumptions
  - internal nodes have the same branching factor b
  - game tree is searched to a fixed depth d
- time consumption is proportional to the number of expanded nodes
  - 1 — root node (the initial ply)
  - b — nodes in the first ply
  - b^2 — nodes in the second ply
  - b^d — nodes in the dth ply
- overall running time $O(b^d)$

Rough estimates on running times when $d = 5$

- suppose expanding a node takes 1 ms
- branching factor b depends on the game
- Draughts ($b \approx 3$): $t = 0.243$ s
- Chess ($b \approx 30$): $t = 6\frac{3}{4}$ h
- Go ($b \approx 300$): $t = 77$ a
- alpha-beta pruning reduces b

Controlling the search depth

- usually the whole game tree is too large
  → limit the search depth
  → a partial game tree
  → partial minimax
- *n*-move look-ahead strategy
  - stop searching after *n* moves
  - make the internal nodes (i.e., frontier nodes) leaves
  - use an evaluation function to ‘guess’ the outcome

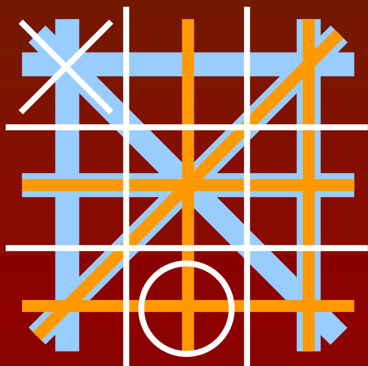
Evaluation function

- combination of numerical measurements $m_i(s, p)$ of the game state
  - single measurement: $m_i(s, p)$
  - difference measurement: $m_i(s, p) - m_j(s, q)$
  - ratio of measurements: $m_i(s, p) \,/\, m_j(s, q)$
- aggregate the measurements maintaining the zero-sum property

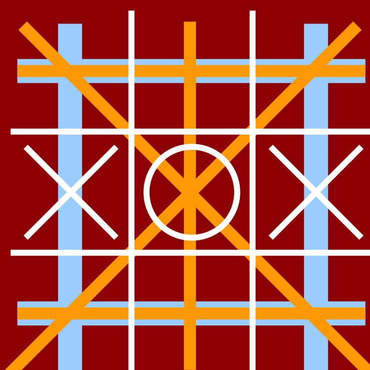
Example: Noughts and Crosses

- heuristic evaluation function *e*:
  - count the winning lines open to MAX
  - subtract the number of winning lines open to MIN
- forced wins
  - state is evaluated $+\infty$, if it is a forced win for MAX
  - state is evaluated $-\infty$, if it is forced win for MIN

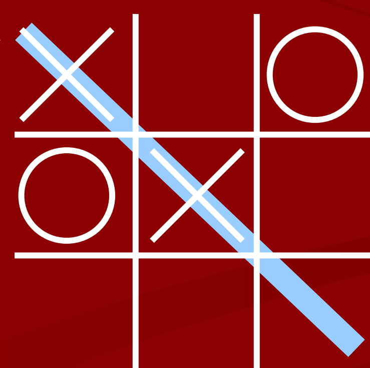
Examples of the evaluation



$$e(\bullet) = 6 - 5 = 1$$



$$e(\bullet) = 4 - 5 = -1$$



$$e(\bullet) = +\infty$$

Drawbacks of partial minimax

- horizon effect
  - heuristically promising path can lead to an unfavourable situation
  - staged search: extend the search on promising nodes
  - iterative deepening: increase n until out of memory or time
  - phase-related search: opening, midgame, end game
  - however, horizon effect cannot be totally eliminated
- bias
  - we want to have an estimate of minimax but get a minimax of estimates
  - distortion in the root: odd plies → win, even plies → loss

The deeper the better…?

- assumptions:
  - *n*-move look-ahead
  - branching factor *b*, depth *d*,
  - leaves with uniform random distribution
- minimax convergence theorem:
  - *n* increases → root value converges to *f*(*b*, *d*)
- last player theorem:
  - root values from odd and even plies not comparable
- minimax pathology theorem:
  - *n* increases → probability of selecting non-optimal move increases (← uniformity assumption!)

Alpha-beta pruning

- reduce the branching factor of nodes
- alpha value
  - associated with MAX nodes
  - represents the worst outcome MAX can achieve
  - can never decrease
- beta value
  - associated with MIN nodes
  - represents the worst outcome MIN can achieve
  - can never increase

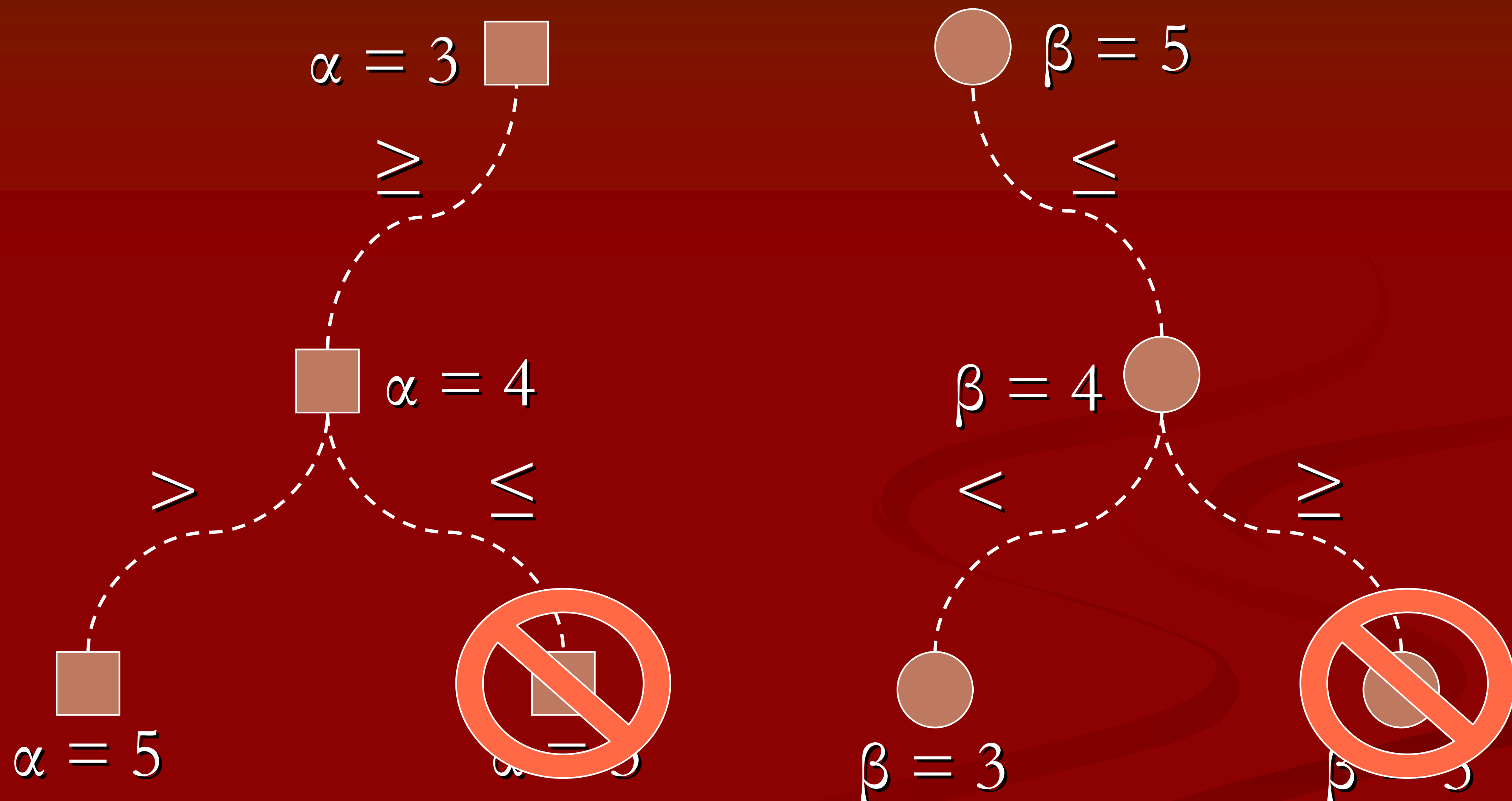
Example

- in a MAX node, $\alpha = 4$
  - we know that MAX can make a move which will result at least the value 4
  - we can omit children whose value is less than or equal to 4
- in a MIN node, $\beta = 4$
  - we know that MIN can make a move which will result at most the value 4
  - we can omit children whose value is greater than or equal to 4

Ancestors and α & β

- alpha value of a node is never less than the alpha value of its ancestors
- beta value of a node is never greater than the beta value of its ancestors

Once again



Rules of pruning

1. Prune below any MIN node having a beta value less than or equal to the alpha value of any of its MAX ancestors.
2. Prune below any MAX node having an alpha value greater than or equal to the beta value of any of its MIN ancestors

Or, simply put: If $\alpha \geq \beta$, then prune below!

Best-case analysis

- omit the principal variation
- at depth $d - 1$ optimum pruning: each node expands one child at depth d
- at depth $d - 2$ no pruning: each node expands all children at depth $d - 1$
- at depth $d - 3$ optimum pruning
- at depth $d - 4$ no pruning, etc.
- total amount of expanded nodes: $\Omega(b^{d/2})$

Principal variation search

- alpha-beta range should be small
  - limit the range artificially → aspiration search
  - if search fails, revert to the original range
- game tree node is either
  - α-node: every move has $e \leq \alpha$
  - β-node: every move has $e \geq \beta$
  - principal variation node: one or more moves has $e > \alpha$ but none has $e \geq \beta$

Principal variation search (cont'd)

- if we find a principal variation move (i.e., between α and β), assume we have found a principal variation node
  - search the rest of nodes the assuming they will not produce a good move
    - assume that the rest of nodes have values $< \alpha$
    - null window: $[\alpha, \alpha + \varepsilon]$
  - if the assumption fails, re-search the node
  - works well if the principal variation node is likely to get selected first
    - sort the children?

Non-zero sum game: Prisoner's dilemma

- two criminals are arrested and isolated from each other
- police suspects they have committed a crime together but don't have enough proof
- both are offered a deal: rat on the other one and get a lighter sentence
  - if one defects, he gets free whilst the other gets a long sentence
  - if both defect, both get a medium sentence
  - if neither one defects (i.e., they co-operate with each other), both get a short sentence

Prisoner's dilemma (cont'd)

- two players
- possible moves
  - co-operate
  - defect
- the dilemma: player cannot make a good decision without knowing what the other will do

Payoffs for prisoner A

Prisoner B's move *Prisoner A's move*	**Co-operate: keep silent**	**Defect: rat on the other prisoner**
Co-operate: keep silent	Fairly good: *6 months*	Bad: *10 years*
Defect: rat on the other prisoner	Good: *no penalty*	Mediocre: *5 years*

Payoffs in Chicken

Driver B's move *Driver A's move*	**Co-operate: swerve**	**Defect: keep going**
Co-operate: swerve	Fairly good: *It's a draw.*	Mediocre: *I'm chicken...*
Defect: keep going	Good: *I win!*	Bad: *Crash, boom, bang!!*

Payoffs in Battle of Sexes

Wife's move *Husband's move*	**Co-operate: boxing**	**Defect: opera**
Co-operate: opera	Wife: Very bad Husband: Very bad	Wife: Good Husband: Mediocre
Defect: boxing	Wife: Mediocre Husband: Good	Wife: Bad Husband: Bad

Iterated prisoner's dilemma

- encounters are repeated
- players have memory of the previous encounters
- R. Axelrod: *The Evolution of Cooperation* (1984)
  - greedy strategies tend to work poorly
  - altruistic strategies work better—even if judged by self-interest only
- Nash equilibrium: always defect!
  - but sometimes rational decisions are not sensible
- Tit for Tat (A. Rapoport)
  - co-operate on the first iteration
  - do what the opponent did on the previous move