KI-Programmierung

Game Playing

Bernhard Beckert

UNIVERSITÄT KOBLENZ-LANDAU

Winter Term 2007/2008
Outline

- Perfect play
- Resource limits
- $\alpha$-$\beta$ pruning
- Games of chance
- Games of imperfect information
Games vs. Search Problems

Game playing is a search problem

Defined by

– Initial state
– Successor function
– Goal test
– Path cost / utility / payoff function
Games vs. Search Problems

Game playing is a search problem

Defined by

- Initial state
- Successor function
- Goal test
- Path cost / utility / payoff function

Characteristics of game playing

- “Unpredictable” opponent:
  Solution is a strategy specifying a move for every possible opponent reply

- Time limits:
  Unlikely to find goal, must approximate
Game Playing

Plan of attack

- Computer considers possible lines of play  
  [Babbage, 1846]
- Algorithm for perfect play  
  [Zermelo, 1912; Von Neumann, 1944]
- Finite horizon, approximate evaluation  
  [Zuse, 1945; Wiener, 1948; Shannon, 1950]
- First chess program  
  [Turing, 1951]
- Machine learning to improve evaluation accuracy  
  [Samuel, 1952–57]
- Pruning to allow deeper search  
  [McCarthy, 1956]
### Types of Games

<table>
<thead>
<tr>
<th>Perfect Information</th>
<th>Deterministic</th>
<th>Chance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chess, checkers, go, othello</td>
<td>chess, checkers, go, othello</td>
<td>Backgammon, monopoly</td>
</tr>
<tr>
<td>Imperfect Information</td>
<td>Bridge, poker, scrabble, nuclear war</td>
<td>Bridge, poker, scrabble, nuclear war</td>
</tr>
</tbody>
</table>
Game Tree: 2-Player / Deterministic / Turns

MAX (X)

MIN (O)

MAX (X)

MIN (O)

TERMINAL

Utility

-1 0 +1
Minimax

Perfect play for deterministic, perfect-information games

Idea

Choose move to position with highest minimax value, i.e., best achievable payoff against best play
Minimax: Example

2-ply game

MAX

MIN

B. Beckert: KI-Programmierung – p.8
Minimax Algorithm

function MINIMAX-DECISION(game) returns an operator

for each op in OPERATORS[game] do
    VALUE[op] ← MINIMAX-VALUE(APPLY(op, game), game)
end
return the op with the highest VALUE[op]

function MINIMAX-VALUE(state, game) returns a utility value

if TERMINAL-TEST[game](state) then
    return UTILITY[game](state)
else if MAX is to move in state then
    return the highest MINIMAX-VALUE of SUCCESSORS(state)
else
    return the lowest MINIMAX-VALUE of SUCCESSORS(state)
Properties of Minimax

- Complete
- Optimal
- Time
- Space
Properties of Minimax

**Complete**
Yes, if tree is finite  (chess has specific rules for this)

**Optimal**

**Time**

**Space**
## Properties of Minimax

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>Yes, if tree is finite (chess has specific rules for this)</td>
</tr>
<tr>
<td>Optimal</td>
<td>Yes, against an optimal opponent. Otherwise??</td>
</tr>
<tr>
<td>Time</td>
<td></td>
</tr>
<tr>
<td>Space</td>
<td></td>
</tr>
</tbody>
</table>
Properties of Minimax

**Complete**    Yes, if tree is finite   (chess has specific rules for this)

**Optimal**    Yes, against an optimal opponent.  Otherwise??

**Time**    $O(b^m)$     (depth-first exploration)

**Space**
Properties of Minimax

**Complete**  Yes, if tree is finite  (chess has specific rules for this)

**Optimal**  Yes, against an optimal opponent.  Otherwise??

**Time**  \( O(b^m) \)  (depth-first exploration)

**Space**  \( O(bm) \)  (depth-first exploration)
Properties of Minimax

**Complete**
Yes, if tree is finite  (chess has specific rules for this)

**Optimal**
Yes, against an optimal opponent.  Otherwise??

**Time**
\( O(b^m) \)  (depth-first exploration)

**Space**
\( O(bm) \)  (depth-first exploration)

**Note**

Finite strategy can exist even in an infinite tree
Resource Limits

Complexity of chess

\[ b \approx 35, \ m \approx 100 \] for “reasonable” games

Exact solution completely infeasible
Resource Limits

Complexity of chess

\[ b \approx 35, \ m \approx 100 \text{ for “reasonable” games} \]

Exact solution completely infeasible

Standard approach

- Cutoff test
  - e.g., depth limit (perhaps add quiescence search)

- Evaluation function
  - Estimates desirability of position
Evaluation Functions

Estimate desirability of position

Black to move
White slightly better

White to move
Black winning
Evaluation Functions

Typical evaluation function for chess

Weighted sum of features

$$\text{Eval}(s) = w_1 f_1(s) + w_2 f_2(s) + \cdots + w_n f_n(s)$$
Evaluation Functions

Typical evaluation function for chess

Weighted sum of features

\[ \text{Eval}(s) = w_1 f_1(s) + w_2 f_2(s) + \cdots + w_n f_n(s) \]

Example

\[
\begin{align*}
  w_1 &= 9 \\
  f_1(s) &= \text{(number of white queens)} - \text{(number of black queens)}
\end{align*}
\]
Digression: Exact Values Do Not Matter

Behaviour is preserved under any monotonic transformation of \textsc{Eval}.

Only the order matters: payoff in deterministic games acts as an ordinal utility function.
Cutting Off Search

Does it work in practice?

\[ b^m = 10^6, \quad b = 35 \quad \Rightarrow \quad m = 4 \]
Cutting Off Search

Does it work in practice?

\[
b^m = 10^6, \quad b = 35 \quad \Rightarrow \quad m = 4
\]

Not really, because . . .

4-ply \quad \approx \quad \text{human novice} \quad (\text{hopeless chess player})

8-ply \quad \approx \quad \text{typical PC, human master}

12-ply \quad \approx \quad \text{Deep Blue, Kasparov}
\( \alpha-\beta \) Pruning Example

MAX

MIN

\[
\begin{array}{c}
3 \\
12 \\
8
\end{array}
\]

\( \geq 3 \)
$\alpha\beta$ Pruning Example

MAX

MIN

3 12 8

3 2

$\geq 3$

$\leq 2$

X X
α-β Pruning Example

MAX

MIN

3 ≥ 3

3 12 8

≤ 2

2 14

X X

3 12 8 2 14
\( \alpha-\beta \) Pruning Example
α-β Pruning Example
Properties of $\alpha$-$\beta$

Effects of pruning

- Reduces the search space
- Does not affect final result
Properties of $\alpha$-$\beta$

Effects of pruning

- Reduces the search space
- Does not affect final result

Effectiveness

Good move ordering improves effectiveness

Time complexity with “perfect ordering”: $O(b^{m/2})$

Doubles depth of search

For chess:
Can easily reach depth 8 and play good chess
The Idea of $\alpha-\beta$

$\alpha$ is the best value (to $\text{MAX}$) found so far off the current path.

If value $x$ of some node below $V$ is known to be less than $\alpha$,

then value of $V$ is known to be at most $x$, i.e., less than $\alpha$,

therefore $\text{MAX}$ will avoid node $V$.

**Consequence**

No need to expand further nodes below $V$. 
The $\alpha$-\(\beta\) Algorithm

```plaintext
function $\text{MAX-VALUE}(\text{state}, \text{game}, \alpha, \beta)$ returns the minimax value of $\text{state}$

inputs: $\text{state}$ /* current state in game */
        $\text{game}$ /* game description */
        $\alpha$ /* the best score for $\text{MAX}$ along the path to $\text{state}$ */
        $\beta$ /* the best score for $\text{MIN}$ along the path to $\text{state}$ */

if $\text{CUTOFF-TEST}(\text{state})$ then return $\text{EVAL}(\text{state})$

for each $s$ in $\text{SUCCESSORS}(\text{state})$ do
    $\alpha \leftarrow \text{MAX}(\alpha, \text{MIN-VALUE}(s, \text{game}, \alpha, \beta))$
    if $\alpha \geq \beta$ then return $\beta$
end
return $\alpha$
```
The $\alpha$-$\beta$ Algorithm

**function** $\text{MIN-VALUE}(state, game, \alpha, \beta)$ **returns** the minimax value of $state$

**inputs:**
- $state$ /* current state in game */
- $game$ /* game description */
- $\alpha$ /* the best score for $\text{MAX}$ along the path to $state$ */
- $\beta$ /* the best score for $\text{MIN}$ along the path to $state$ */

if $\text{CUTOFF-TEST}(state)$ then return $\text{EVAL}(state)$

for each $s$ in $\text{SUCCESSORS}(state)$ do
  $\beta \leftarrow \text{MIN}(\beta, \text{MAX-VALUE}(s, game, \alpha, \beta))$
  if $\beta \leq \alpha$ then return $\alpha$

end

return $\beta$
Checkers

Chinook ended 40-year-reign of human world champion Marion Tinsley in 1994. Used an endgame database defining perfect play for all positions involving 8 or fewer pieces on the board, a total of 443,748,401,247 positions.
Deterministic Games in Practice

Checkers

Chinook ended 40-year-reign of human world champion Marion Tinsley in 1994. Used an endgame database defining perfect play for all positions involving 8 or fewer pieces on the board, a total of 443,748,401,247 positions.

Chess

Deep Blue defeated human world champion Gary Kasparov in a six-game match in 1997. Deep Blue searches 200 million positions per second, uses very sophisticated evaluation, and undisclosed methods for extending some lines of search up to 40 ply.
Deterministic Games in Practice

Checkers

Chinook ended 40-year-reign of human world champion Marion Tinsley in 1994. Used an endgame database defining perfect play for all positions involving 8 or fewer pieces on the board, a total of 443,748,401,247 positions.

Chess

Deep Blue defeated human world champion Gary Kasparov in a six-game match in 1997. Deep Blue searches 200 million positions per second, uses very sophisticated evaluation, and undisclosed methods for extending some lines of search up to 40 ply.

Go

Human champions refuse to compete against computers, who are too bad. In go, \( b > 300 \), so most programs use pattern knowledge bases to suggest plausible moves.
Nondeterministic Games: Backgammon
Nondeterministic Games in General

Chance introduced by dice, card-shuffling, etc.

Simplified example with coin-flipping
Algorithm for Nondeterministic Games

**EXPECTMINIMAX gives perfect play**

if \textit{state} is a \texttt{Max} node then
return the highest \texttt{EXPECTMINIMAX} value of \texttt{SUCCESSORS}(\textit{state})

if \textit{state} is a \texttt{Min} node then
return the lowest \texttt{EXPECTMINIMAX} value of \texttt{SUCCESSORS}(\textit{state})

if \textit{state} is a chance node then
return average of \texttt{EXPECTMINIMAX} value of \texttt{SUCCESSORS}(\textit{state})
Pruning in Nondeterministic Game Trees

A version of $\alpha$-$\beta$ pruning is possible
Pruning in Nondeterministic Game Trees

A version of $\alpha$-$\beta$ pruning is possible

![Game Tree Diagram](image)
A version of $\alpha$-$\beta$ pruning is possible

![Game Tree Diagram]

- **Root Node**
  - Value range: $[-\infty, +\infty]$ (gray)
  - Probability: 0.5

- **Left Subtree**
  - Value range: $[2, 2]$ (red)
  - Probability: 0.5

- **Right Subtree**
  - Value range: $[-\infty, +\infty]$ (gray)
  - Probability: 0.5

- **Leaves**
  - Value range: $2$ (gray)
  - Probability: 0.5
  - Value range: $2$ (gray)
  - Probability: 0.5

---

*B. Beckert: KI-Programmierung – p.25*
Pruning in Nondeterministic Game Trees

A version of $\alpha$-$\beta$ pruning is possible

![Game Tree Diagram]

- Upper node: $[\infty, 2]$
- Lower nodes:
  - Left: $[2, 2]$ with values $2, 2$
  - Right: $[\infty, +\infty]$ with values $0.5, 0.5$

B. Beckert: KI-Programmierung – p.25
A version of $\alpha$-$\beta$ pruning is possible
Pruning in Nondeterministic Game Trees

A version of $\alpha$-$\beta$ pruning is possible

\[
\begin{array}{c}
\left[ 1.5, 1.5 \right] \\
0.5 \\
\left[ 2, 2 \right] \\
2
\end{array}
\begin{array}{c}
\left[ 1, 1 \right] \\
0.5 \\
2
\end{array}
\begin{array}{c}
\left[ -\infty, +\infty \right] \\
0.5 \\
\left[ -\infty, 0 \right] \\
0
\end{array}
\begin{array}{c}
\left[ -\infty, +\infty \right] \\
0.5 \\
\left[ -\infty, +\infty \right]
\end{array}
\]
Pruning in Nondeterministic Game Trees

A version of $\alpha$-$\beta$ pruning is possible
Pruning in Nondeterministic Game Trees

A version of $\alpha$-$\beta$ pruning is possible

```
          [ 1.5 , 1.5 ]
         /    /
0.5    0.5

[ 2 , 2 ]    [ 1 , 1 ]

2    2

[ -\infty , 0.5 ]
         /    /
0.5    0.5

[ 0 , 0 ]    [ -\infty , 1 ]

0    1

1
```
Pruning Continued

More pruning occurs if we can bound the leaf values

```
[ -2, 2 ]
```

```
0.5

[ -2, 2 ]
```

```
0.5

[ -2, 2 ]
```

```
[ -2, 2 ]
```

```
[ -2, 2 ]
```

```
[ -2, 2 ]
```

```
[ -2, 2 ]
```

```
[ -2, 2 ]
```
More pruning occurs if we can bound the leaf values
More pruning occurs if we can bound the leaf values.
More pruning occurs if we can bound the leaf values
Pruning Continued

More pruning occurs if we can bound the leaf values
More pruning occurs if we can bound the leaf values
Nondeterministic Games in Practice

Problem

\(\alpha - \beta\) pruning is much less effective

Dice rolls increase \(b\)

21 possible rolls with 2 dice

Backgammon

\(\approx 20\) legal moves

\[
\text{depth 4} = 20^4 \times 21^3 \approx 1.2 \times 10^9
\]

TDGAMMON

Uses depth-2 search + very good EVAL \(\approx\) world-champion level
Digression: Exact Values DO Matter

Behaviour is preserved only by positive linear transformation of EVAL.

Hence EVAL should be proportional to the expected payoff.
Games of Imperfect Information

Typical examples

Card games: Bridge, poker, skat, etc.

Note

Like having one big dice roll at the beginning of the game
Games of Imperfect Information

Idea for computing best action

Compute the minimax value of each action in each deal, then choose the action with highest expected value over all deals.

Requires information on probability the different deals

Special case

If an action is optimal for all deals, it’s optimal.
Games of Imperfect Information

Idea for computing best action

Compute the minimax value of each action in each deal, then choose the action with highest expected value over all deals.

Requires information on probability the different deals.

Special case

If an action is optimal for all deals, it’s optimal.

Bridge

GIB, current best bridge program, approximates this idea by

– generating 100 deals consistent with bidding information
– picking the action that wins most tricks on average
Commonsense Example

Day 1

Road A leads to a small heap of gold pieces 10 points

Road B leads to a fork:
– take the left fork and you’ll find a mound of jewels 100 points
– take the right fork and you’ll be run over by a bus −1000 points

Best action: Take road B (100 points)
Commonsense Example

Day 1

Road A leads to a small heap of gold pieces 10 points

Road B leads to a fork:
- take the left fork and you’ll find a mound of jewels 100 points
- take the right fork and you’ll be run over by a bus –1000 points

Best action: Take road B (100 points)

Day 2

Road A leads to a small heap of gold pieces 10 points

Road B leads to a fork:
- take the left fork and you’ll be run over by a bus –1000 points
- take the right fork and you’ll find a mound of jewels 100 points

Best action: Take road B (100 points)
Commonsense Example

Day 3

Road A leads to a small heap of gold pieces  (10 points)

Road B leads to a fork:
– guess correctly and you’ll find a mound of jewels  100 points
– guess incorrectly and you’ll be run over by a bus  −1000 points

Best action:  Take road A  (10 points)

NOT:  Take road B  \( \left( \frac{-1000 + 100}{2} = -450 \text{ points} \right) \)
Proper Analysis

Note

Value of an actions is NOT the average of values for actual states computed with perfect information.

With partial observability, value of an action depends on the information state the agent is in.
Proper Analysis

Note

Value of an actions is NOT the average of values for actual states computed with perfect information

With partial observability, value of an action depends on the information state the agent is in

Leads to rational behaviors such as

- Acting to obtain information
- Signalling to one’s partner
- Acting randomly to minimize information disclosure
Summary

- Games are to AI as grand prix racing is to automobile design
- Games are fun to work on (and dangerous)
- They illustrate several important points about AI
  - perfection is unattainable, must approximate
  - it is a good idea to think about what to think about
  - uncertainty constrains the assignment of values to states