MoHex 2.0: pattern-based MCTS

huang arneson hayward müller pawlewicz

computing UAlberta hayward@ualberta.ca

CG2013 aug 13
THANK YOU

- Natural Sciences and Engineering Research Council of Canada
1. HEX
2. KNOWLEDGE
3. MOHEX
4. MOHEX 2.0
1942 Hex

**RULES**

- black v white, alternate moves
- win: connect sides
1942 Hex

**RULES**

- black v white, alternate moves
- win: connect sides
PROPERTIES

- no draw
- n-by-n: 1st-player win
- n-by-(n+k): longer-side win
- Pspace-complete
SHANNON’S BIRDCAGE MACHINE
SWITCHING NETWORK

- play on any graph
- two marked vertices
- black move: ‘short’ any vertex (make nbrs clique)
- white move: ‘cut’ any vertex (delete)
- black wins iff two marked vertices are shorted (connected)
- generalizes Hex
SWITCHING NETWORK
SWITCHING NETWORK
SWITCHING NETWORK
virtual connections
inferior cells
A VIRTUAL CONNECTION
A VIRTUAL CONNECTION
COMBINING RULE: AND (FULL)
COMBINING RULE: AND (FULL)
COMBINING RULE: AND (FULL)
COMBINING RULE: AND (FULL)
COMBINING RULE: AND (FULL)
COMBINING RULE: AND (FULL)
COMBINING RULE: AND (SEMI)
COMBINING RULE: AND (SEMI)
COMBINING RULE: AND (SEMI)
COMBINING RULE: OR
COMBINING RULE: OR
COMBINING RULE: OR
COMBINING RULE: OR
WHERE MUST WHITE PLAY?
WHERE MUST WHITE PLAY?
WHERE MUST WHITE PLAY?
WHERE MUST WHITE PLAY?
DEAD

MoHex 2.0: pattern-based MCTS
BLACK-DOMINATED (DOT SUPERIOR)
BLACK-CAPTURTIRED
BLACK-DOMINATED (DOT SUPERIOR)
BLACK-CAPTURE-REVERSIBLE (TO WHITE DOT)
BLACK FILL DECOMPOSITION
STAR DECOMPOSITION
BLACK STAR DECOMP DOMINATION
modify H-search

- and/or combining rules + capture
• while time remains:
  • traverse tree (repeat: select child, move to child)
  • expand: leaf $\rightarrow$ node
  • evaluate node: simulation
  • update info: traverse from node back to root
• select most-visited root-child as move
MOHEX SIMULATION PATTERN
MOHEX SIMULATION PATTERN
MOHEX SIMULATION PATTERN
ALL MOVES AS FIRST

- use RAVE, an AMAF heuristic
- set exploration multiplier to 0 (so not UCT)
ICE/VCE PRUNING

during traversal:
  if node becomes heavy
  apply ICE/VCE
  prune inferior cells
  prune non-mustplay
ICE PRUNING
ICE PRUNING
MOHEX FLAWS

- weak without VCE, ICE
- weak playouts
IMPROVEMENTS

- extend on unstable search
- lazy delete obsolete subtrees
- improved RAVE formula
- patterns
  - estimate prior knowledge
  - progressive bias
  - probabilistic simulations
- experiments
- future work
move becomes obsolete?

1) mark child obsolete

2) in traversal, before moving to a child, check whether obsolete: yes? mark as proven loss
**IMPROVED RAVE FORMULA**

\[ U: \text{UCT mean (wins/visits)} \]

\[ R: \text{RAVE mean (wins/visits)} \]

\[ n: \text{parent visit count} \]

\[ n_j: \text{node visit count} \]

\[ c_b: \text{constant} \]

\[ w: \text{RAVE term weight (decays } \sim 1 \text{ to } 0 \text{ with } n_j \text{)} \]

\[ E: \text{UCT exploration formula} \quad c_b \times \sqrt{\frac{\ln n}{n_j}} \]

\[ \text{score}(j) = (1 - w) \times (U + E) + w \times R \]
supervised learning minorization-maximization
15 000 11x11 mohex-wolve games (ignore 1st move)
20 000 13x13 strong little golem games
consider 6- 12- 18-cell patterns
65 900 global 6-,12-patterns (30 600 prunable)
11 600 local 6-,12-patterns (3 700 prunable)
prunable dead/captured, dominated: $\gamma \rightarrow 1e-5, 1e-4$
MoHex 2.0: pattern-based MCTS

\[
(\gamma, p, a) = (886, 439, 479) (754, 179, 194)
\]

\[
(754, 179, 194)
\]

\[
(321, 48, 64)
\]

\[
(213, 52, 65)
\]
MoHex 2.0: pattern-based MCTS
ESTIMATING PRIOR KNOWLEDGE

- check pattern of every available move
- prunable \( ? \) move not considered
- non-prunable \( ? \) \( \rho \leftarrow \) relative global + local \( \gamma \) sum
- unvisited node: RAVE score, count \( \leftarrow .5, 8 \)
PROGRESSIVE BIAS

following Mango, ...

\[ \text{Score}(j) = (1 - w) \times (U + E) + w \times R + PB \]

following Castro, ...

\[ PB = c_{pb} \times \rho / \sqrt{n_j + 1} \]

from CLOP

\[ c_{pb} = 2.47 \]
PROBABILISTIC SIMULATIONS

use weights, generate moves stochastically via softmax

cap global $\gamma_{\text{max}} \leftarrow .157$, by CLOP
MoHex 2.0: pattern-based MCTS
**EXPERIMENTS**

- all openings
- each player: 4 cores, 1.5Gb, 1-3-5 min/game
- 3000 $13 \times 13$ games, each player 3-min/game
  - M-W ($0.587 \pm 0.008$) M2-W ($0.854 \pm 0.006$) 245 Elo
- 1000 games M2-M:

<table>
<thead>
<tr>
<th>board size</th>
<th>1 min</th>
<th>3 min</th>
<th>5 min</th>
</tr>
</thead>
<tbody>
<tr>
<td>$11 \times 11$</td>
<td></td>
<td>$0.811 \pm 0.010$</td>
<td></td>
</tr>
<tr>
<td>$13 \times 13$</td>
<td>$0.853 \pm 0.006$</td>
<td>$0.852 \pm 0.006$</td>
<td>$0.856 \pm 0.010$</td>
</tr>
</tbody>
</table>
FAILURES

- hand-crafted patterns
  savebridge + breakbridge + ladder
  win rate .6/10K .5/100K
- degrade RAVE by distance to last move
- move criticality
- ...
FUTURE WORK

W: Panoramex  B: MoHex (2011 Olympiad)
Natural Sciences and Engineering Research Council of Canada