

MoHex 2.0: PATTERN-BASED MCTS

huang arneson hayward müller pawlewicz

computing UAlberta hayward@ualberta.ca

CG2013 aug 13

THANK YOU

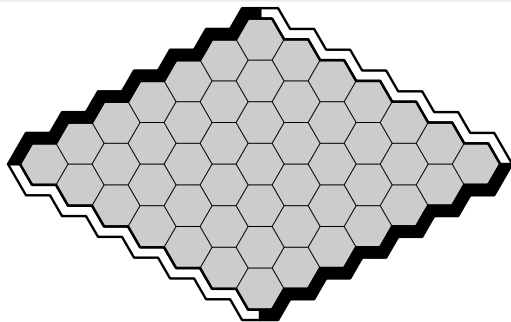
- Natural Sciences and Engineering Research Council of Canada

- 1 HEX
- 2 KNOWLEDGE
- 3 MOHEX
- 4 MOHEX 2.0

1942 HEX

RULES

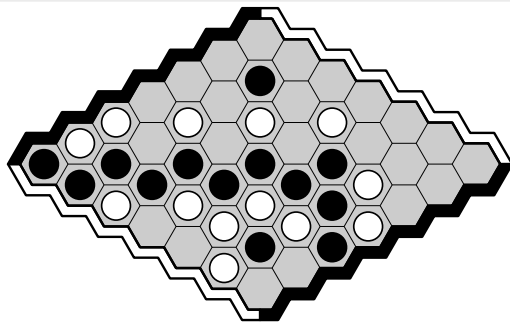
- black v white, alternate moves
- win: connect sides



1942 HEX

RULES

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

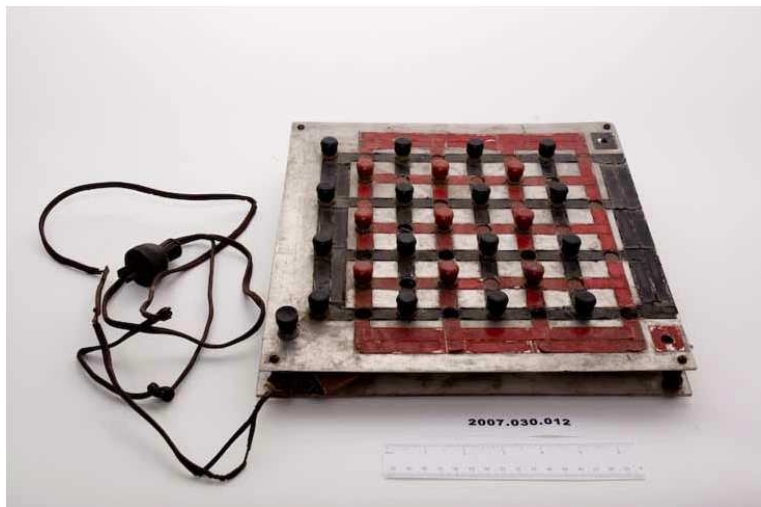


PROPERTIES

PROPERTIES

- no draw
- n -by- n : 1st-player win
- n -by- $(n+k)$: longer-side win
- Pspace-complete

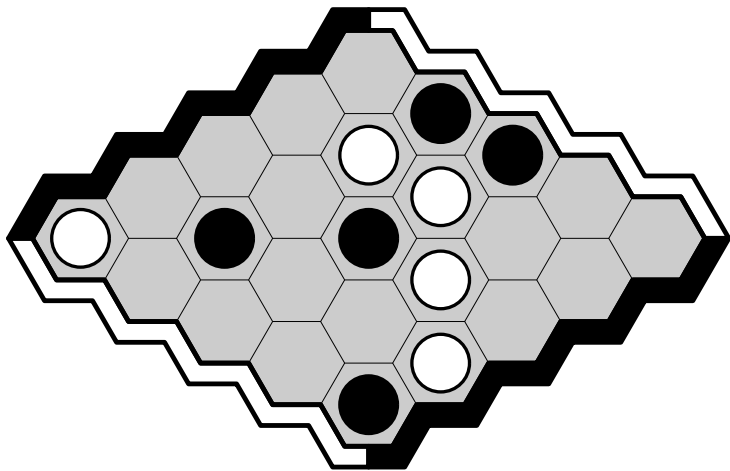
SHANNON'S BIRDCAGE MACHINE



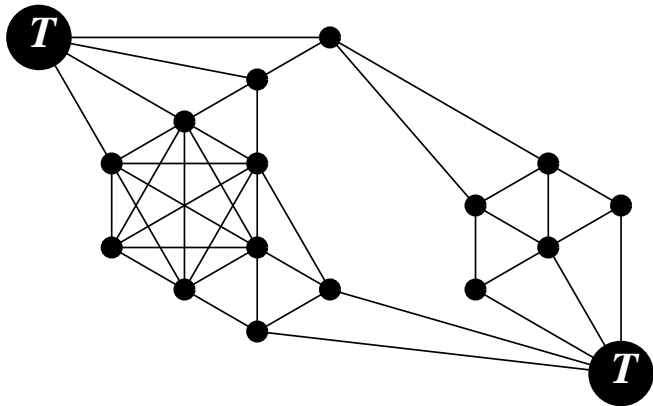
SWITCHING NETWORK

- 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

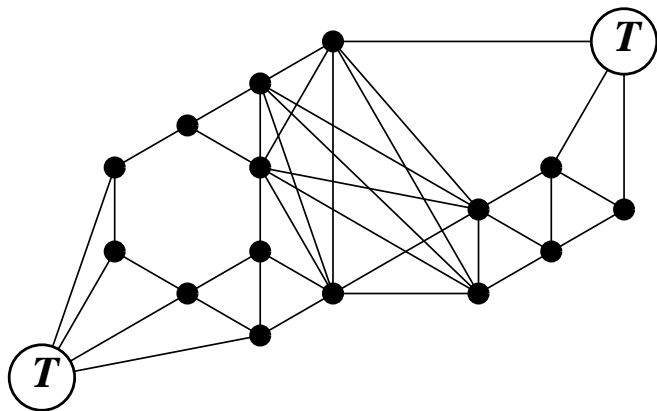
SWITCHING NETWORK



SWITCHING NETWORK



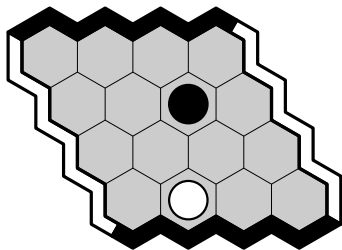
SWITCHING NETWORK



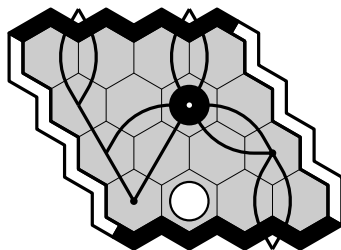
KNOWLEDGE

- virtual connections
- inferior cells

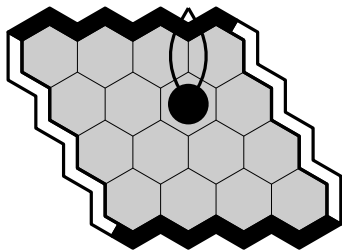
A VIRTUAL CONNECTION



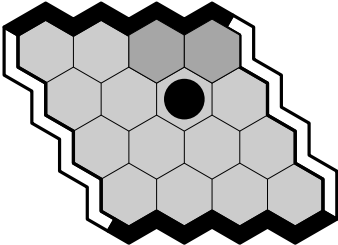
A VIRTUAL CONNECTION



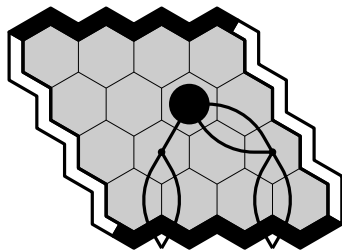
COMBINING RULE: AND (FULL)



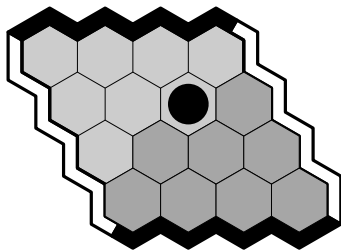
COMBINING RULE: AND (FULL)



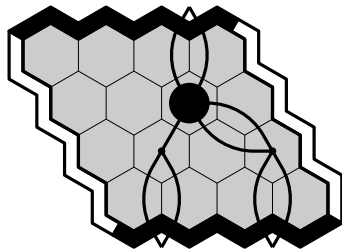
COMBINING RULE: AND (FULL)



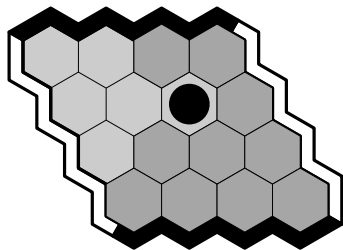
COMBINING RULE: AND (FULL)



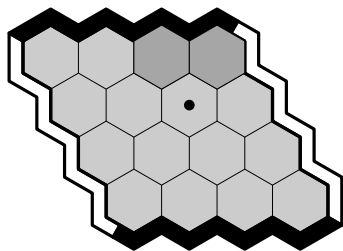
COMBINING RULE: AND (FULL)



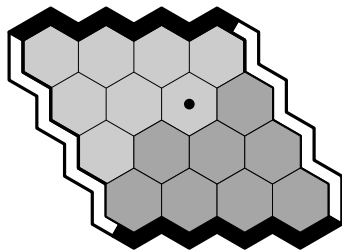
COMBINING RULE: AND (FULL)



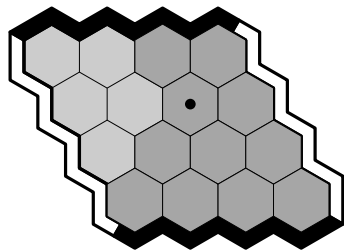
COMBINING RULE: AND (SEMI)



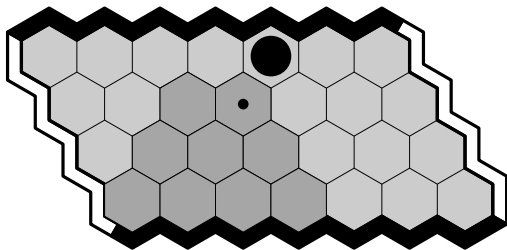
COMBINING RULE: AND (SEMI)



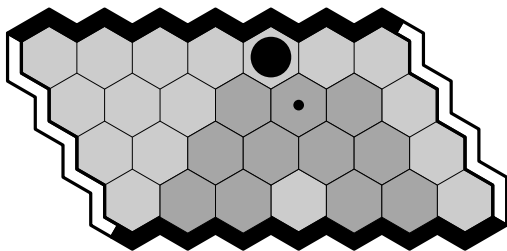
COMBINING RULE: AND (SEMI)



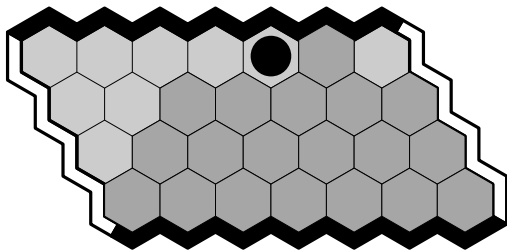
COMBINING RULE: OR



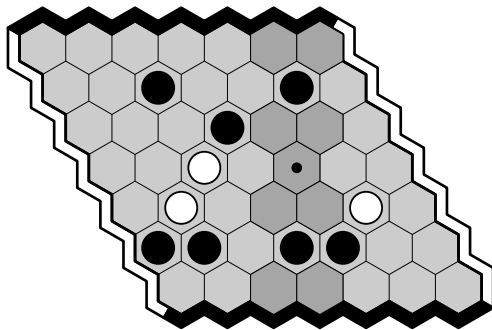
COMBINING RULE: OR



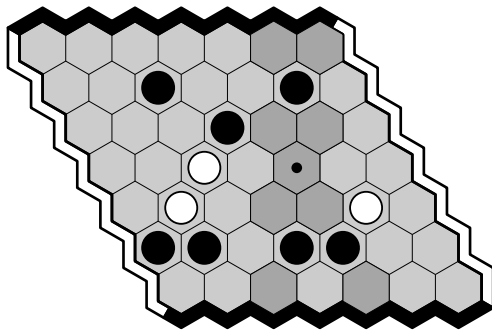
COMBINING RULE: OR



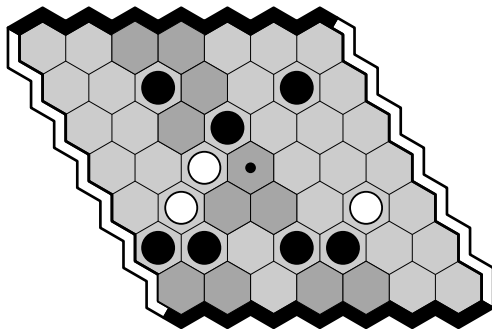
WHERE MUST WHITE PLAY?



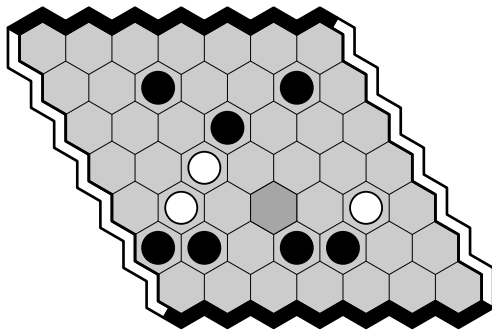
WHERE MUST WHITE PLAY?



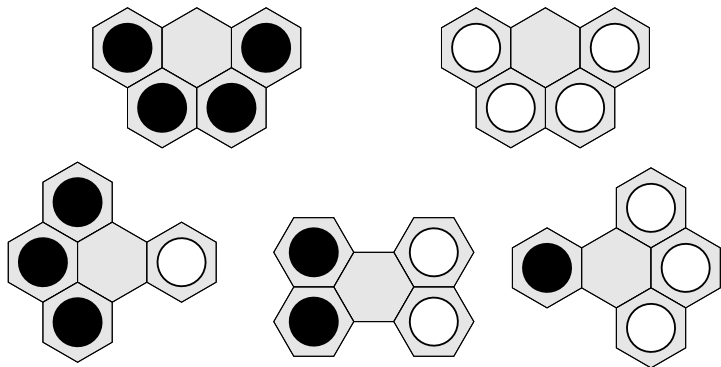
WHERE MUST WHITE PLAY?



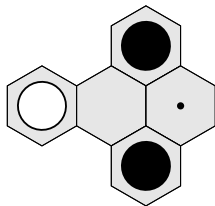
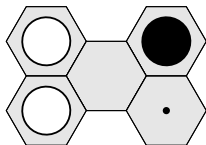
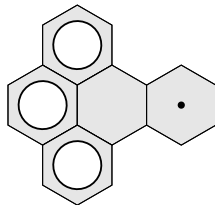
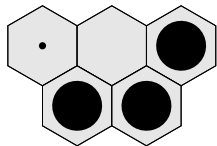
WHERE MUST WHITE PLAY?



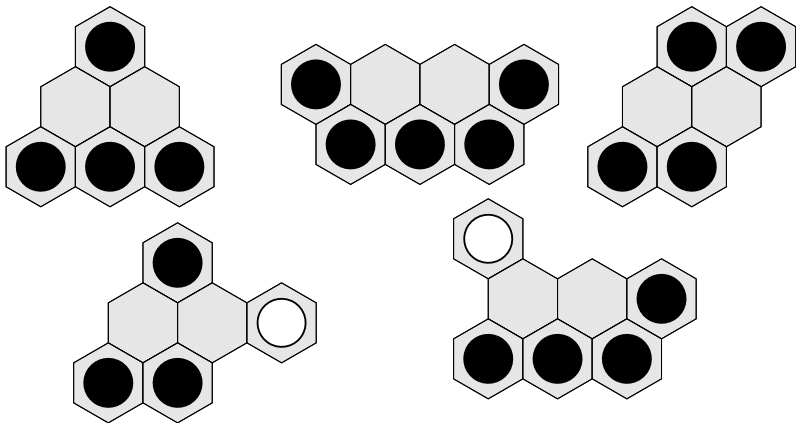
DEAD



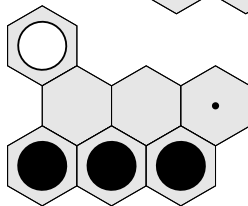
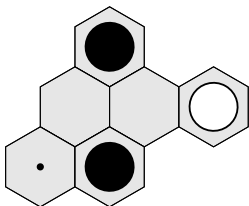
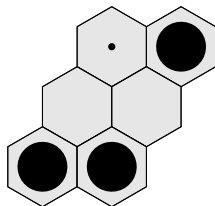
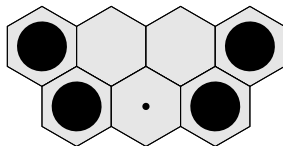
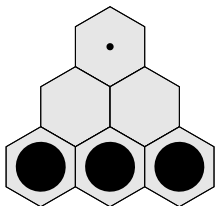
BLACK-DOMINATED (DOT SUPERIOR)



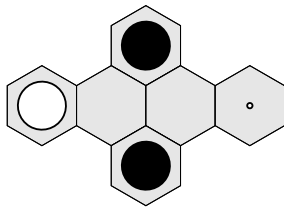
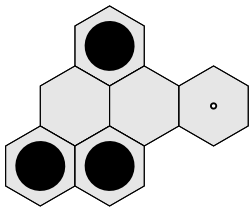
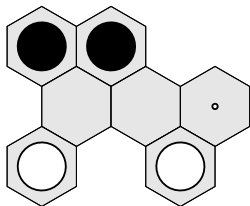
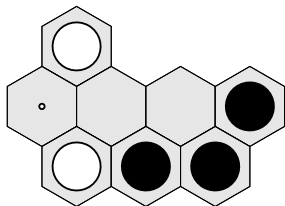
BLACK-CAPTURED



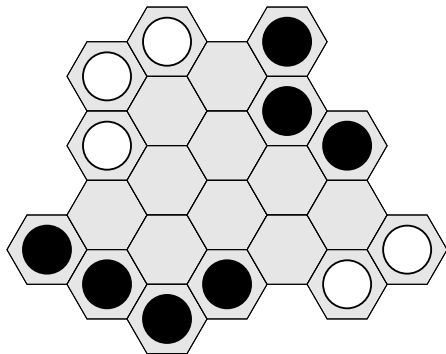
BLACK-DOMINATED (DOT SUPERIOR)



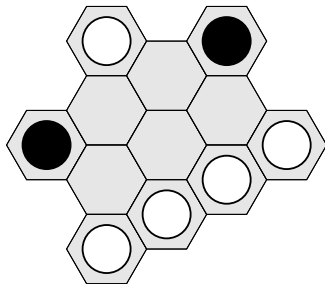
BLACK-CAPTURE-REVERSIBLE (TO WHITE DOT)



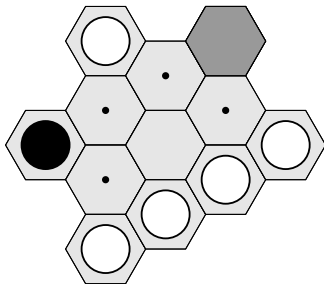
BLACK FILL DECOMPOSITION



STAR DECOMPOSITION

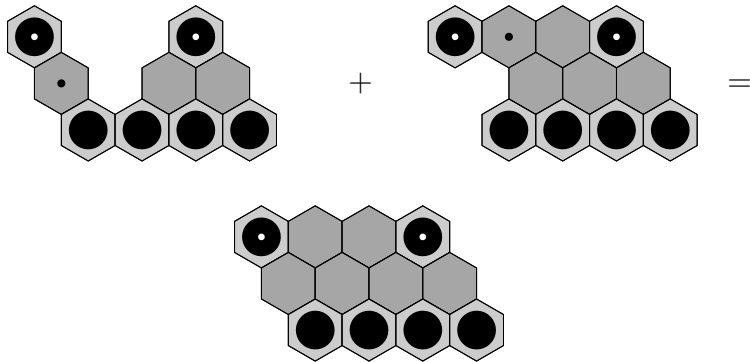


BLACK STAR DECOMP DOMINATION



modify H-search

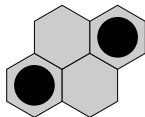
- and/or combining rules + capture



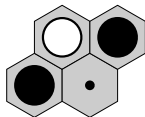
MOHEX FRAMEWORK

- 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

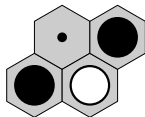
MOHEX SIMULATION PATTERN



MOHEX SIMULATION PATTERN



MOHEX SIMULATION PATTERN



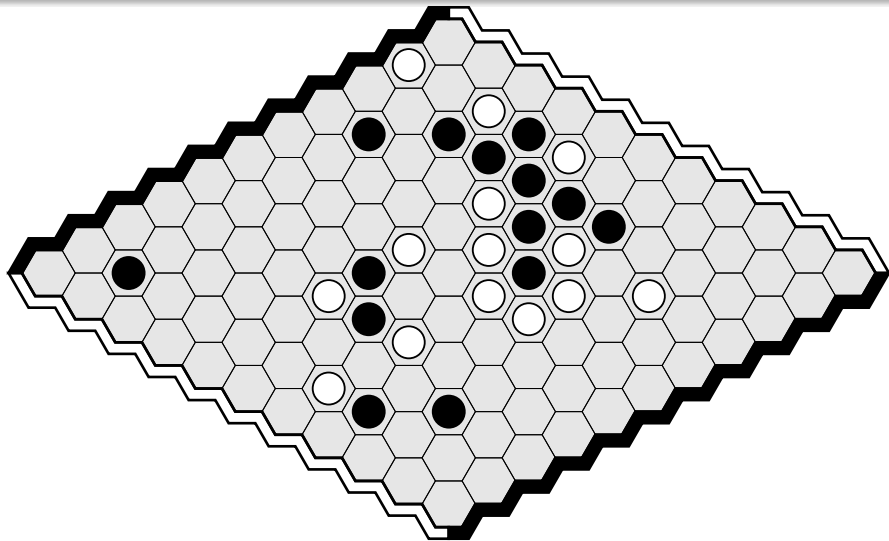
ALL MOVES AS FIRST

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

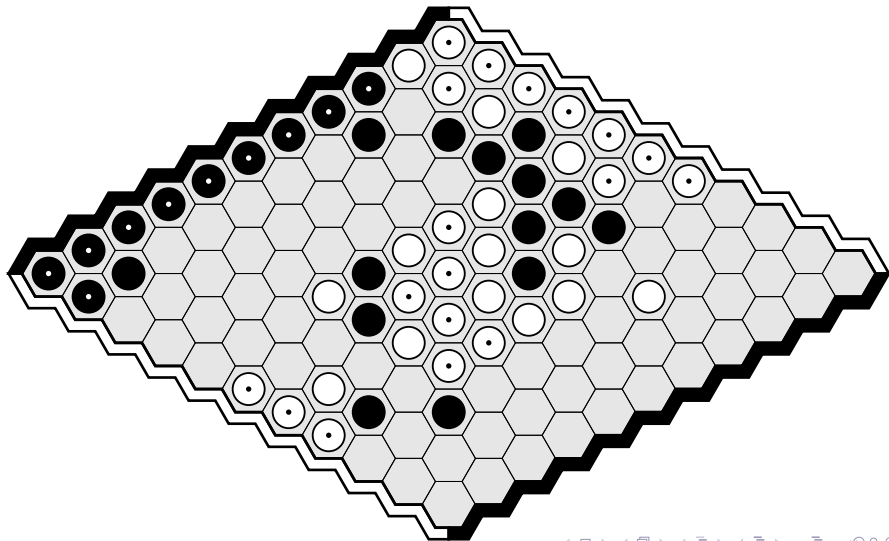
ICE/VCE PRUNING

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

LAZY DELETE OBSOLETE SUBTREE

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 : UCT mean (wins/visits)

R : RAVE mean (wins/visits)

n : parent visit count

n_j : node visit count

c_b : constant

w : RAVE term weight (decays ~ 1 to 0 with n_j)

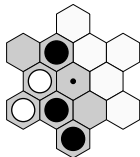
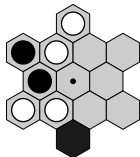
E : UCT exploration formula $c_b \times \sqrt{\frac{\ln n}{n_j}}$

$$\text{score}(j) = (1 - w) \times (U + E) + w \times R$$

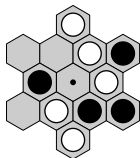
PATTERNS

- 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: $\gamma \rightarrow 1e-5, 1e-4$

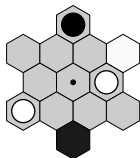
PATTERNS



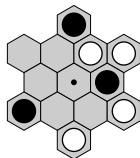
$(\gamma, p, a) = (886, 439, 479) (754, 179, 194)$



$(754, 179, 194)$

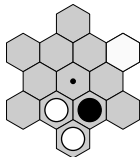


$(321, 48, 64)$

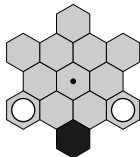


$(213, 52, 65)$

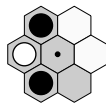
PATTERNS



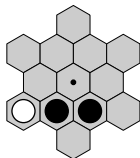
(194,2247,3259)



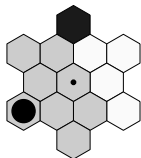
(100,86,182)



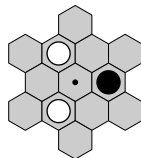
(98,94,191)



(.04,0,10190)



(.05,3,14270)



(.05,6,17351)

ESTIMATING PRIOR KNOWLEDGE

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

PROGRESSIVE BIAS

following Mango, ...

$$Score(j) = (1 - w) \times (U + E) + w \times R + PB$$

following Castro, ...

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

from CLOP

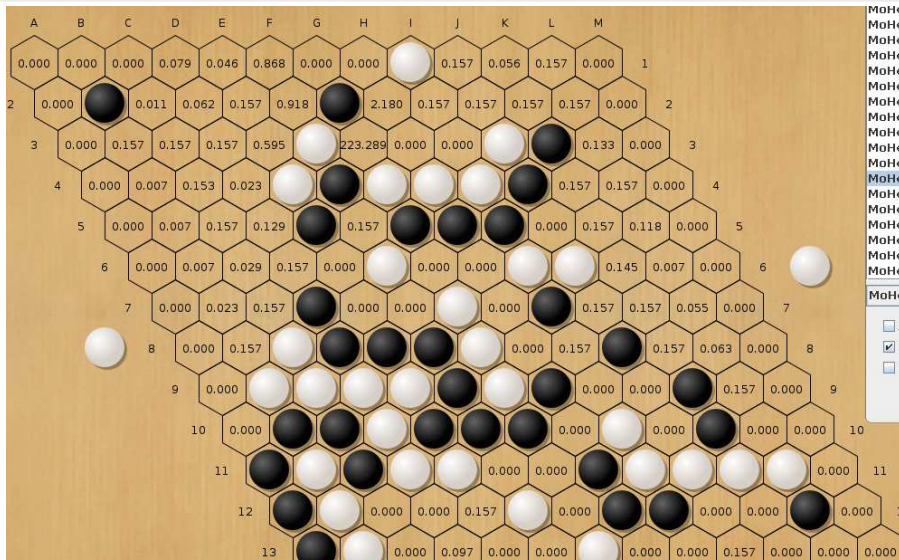
$$c_{pb} = 2.47$$

PROBABILISTIC SIMULATIONS

use weights, generate moves stochastically via softmax

cap global γ max \leftarrow .157, by CLOP

PROBABILISTIC SIMULATIONS



EXPERIMENTS

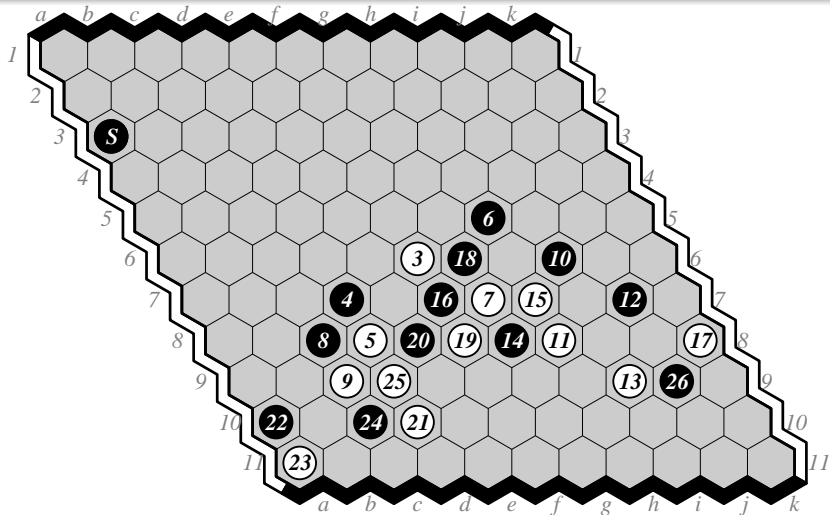
- all openings
- each player: 4 cores, 1.5Gb, 1-3-5 min/game
- 3000 13×13 games, each player 3-min/game
 M-W ($.587 \pm .008$) M2-W ($.854 \pm .006$) 245 Elo
- 1000 games M2-M:

	time/player		
board size	1 min	3 min	5 min
11×11		$.811 \pm .010$	
13×13	$.853 \pm .006$	$.852 \pm .006$	$.856 \pm .010$

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)

THANK YOU

- Natural Sciences and Engineering Research Council of Canada