

University College Dublin

Lecture for *AI for Games and Puzzles*

Strategic Potential in Games

Cameron Browne
Maastricht University
8/11/2018



Outline

Prior work:

- Automated Game Design
- Measuring Game Quality

Current work:

- Reconstruction of Ancient Games

Importance of Strategic Potential

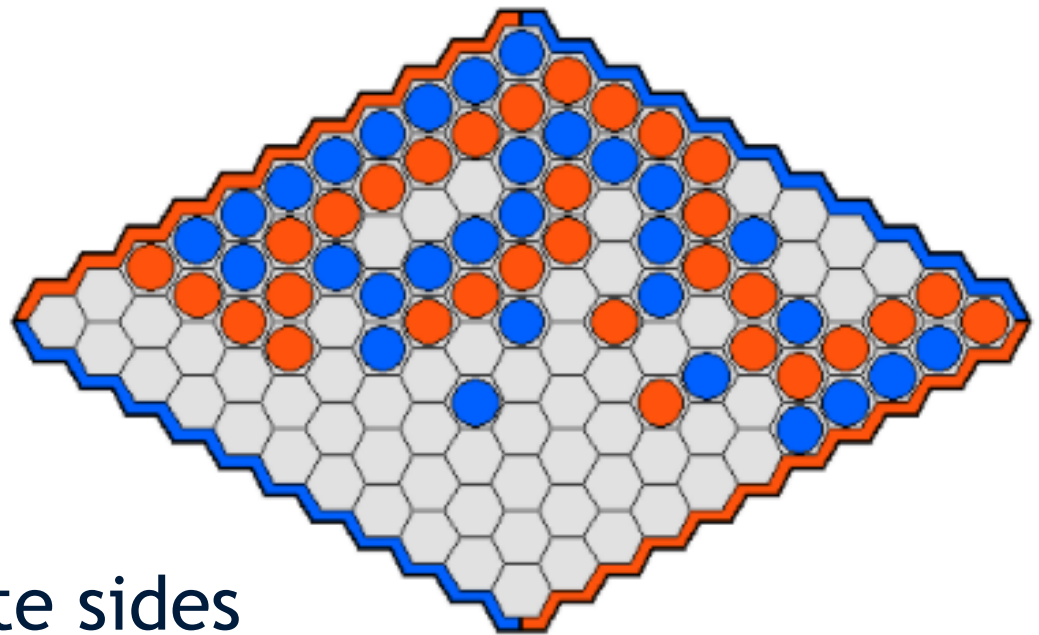
Connection Games

Games in which players aim to complete a given type of connection with their pieces:

- Connecting goals,
- Completing a loop,
- Gathering pieces,
- etc.

Hex:

- Invented 1940s
- Connect own opposite sides
- Source of all connection games



Red won

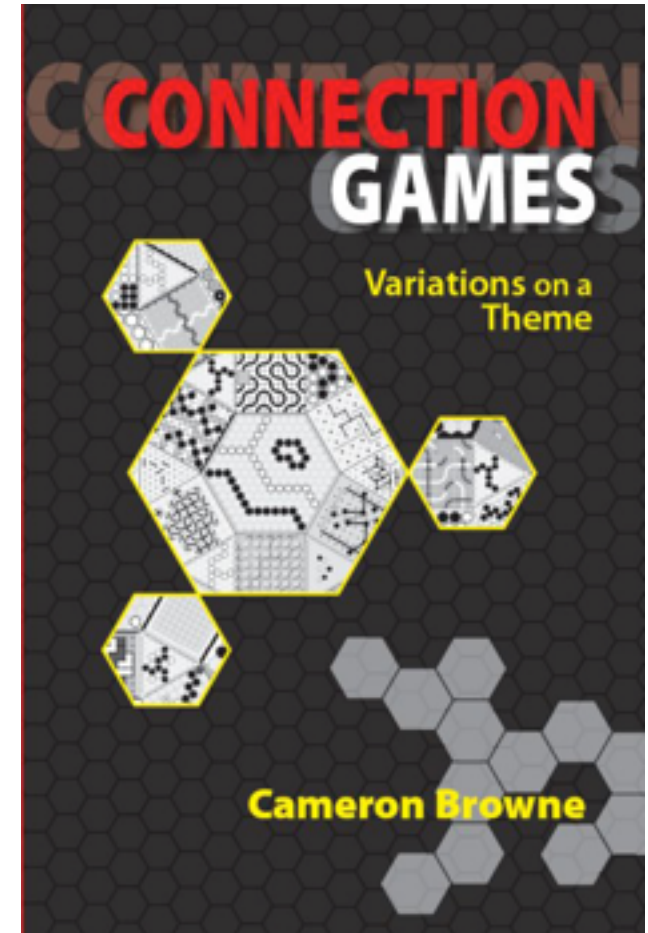
Connection Games

Most players know a few:

- Hex, Havannah, TwixT, Bridg-It, Y, Star, ...

Connection Games: Variations on a Theme (2005):

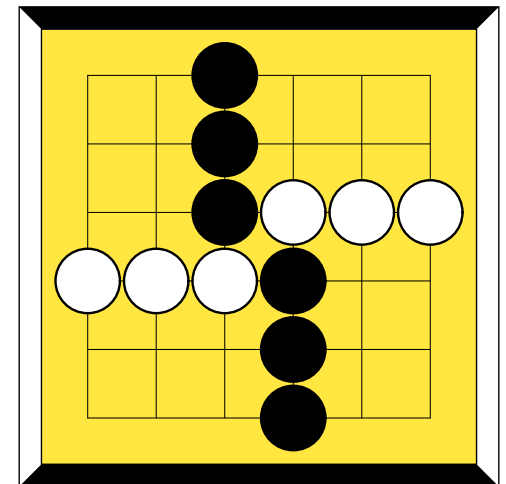
- Survey of genre
- Found over 200
- Most games involve some aspect of connection



Connection Games

Almost all are combinations of same few rules:

- **Hex:** Connect own opposite sides
- **Chameleon:** Connect any opposite sides
- **Jade:** Connect opposite sides or all sides
- **Havannah:** Connect three sides or corners or form loop
- **Unlur:** Connect two or three sides
- **Y:** Connect three sides
(triangular board)
- **TwixT:** Connect opposite sides
(square board with bridge moves)
- **Quax:** Connect opposite sides
(square board with bridge moves)
- **Gonnect:** Connect opposite sides
(square board with surround capture)



All Games

Almost all games are new combinations of existing ideas

Truly revolutionary ideas in games are rare:

- e.g. Connection as a goal
 - Inspired by a mathematical problem
- “Nothing new under the sun”

Obvious question:

- *Why not evolve rule sets to create new games?*

Evolutionary Game Design

2009 PhD Thesis:

- *Automatic Generation and Evaluation of Recombination Games*

```
(game Tic-Tac-Toe
  (players White Black)
  (board
    (tiling square)
    (shape square)
    (size 3 3)
  )
  (end (All win (in-a-row 3))))
```

Defined games as *ludemes*:

- Units of game-related information
- LISP-like *s-expressions*

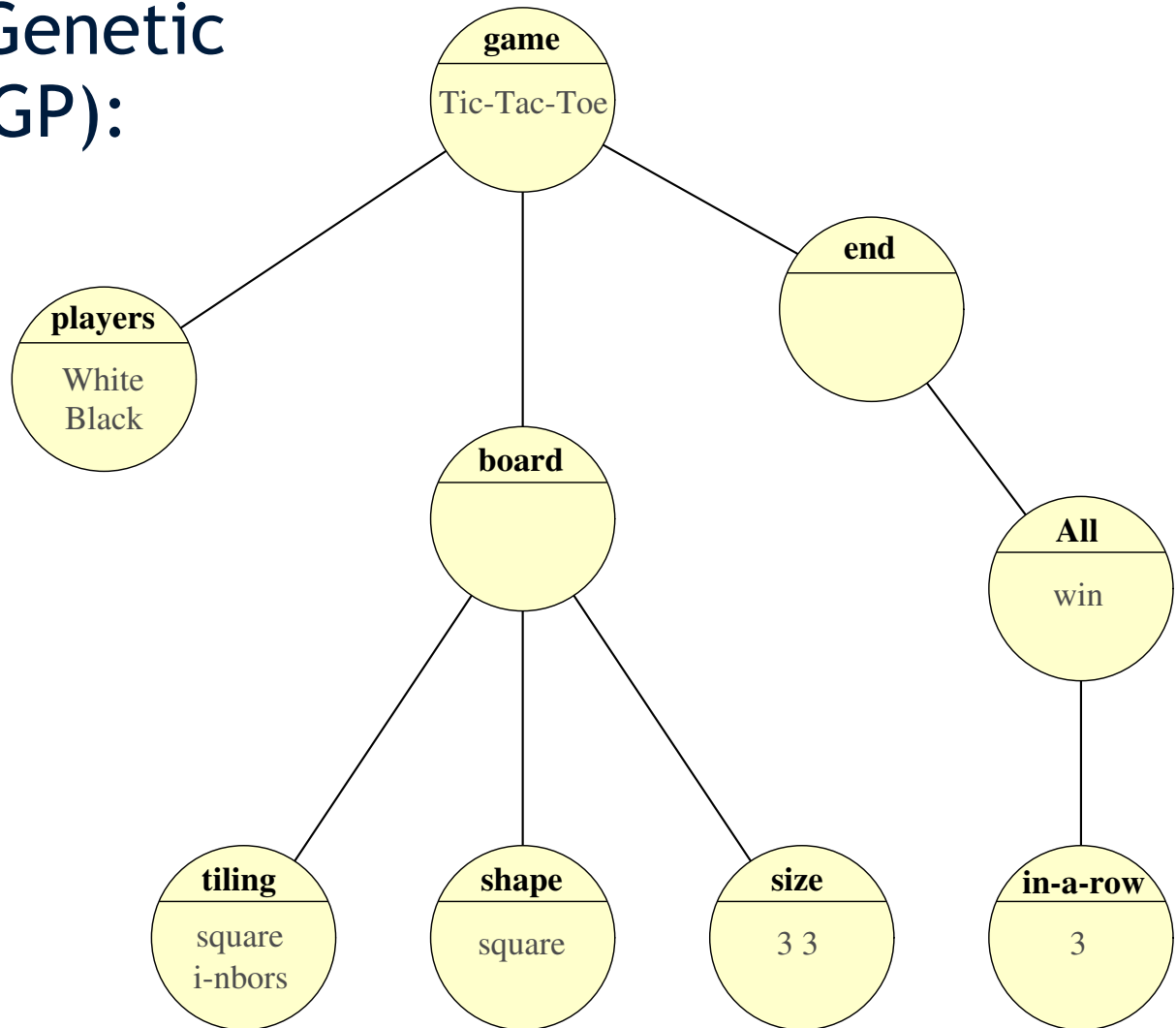
Evolved rule sets

- Measured for quality

Ludeme Trees

Evolved using Genetic Programming (GP):

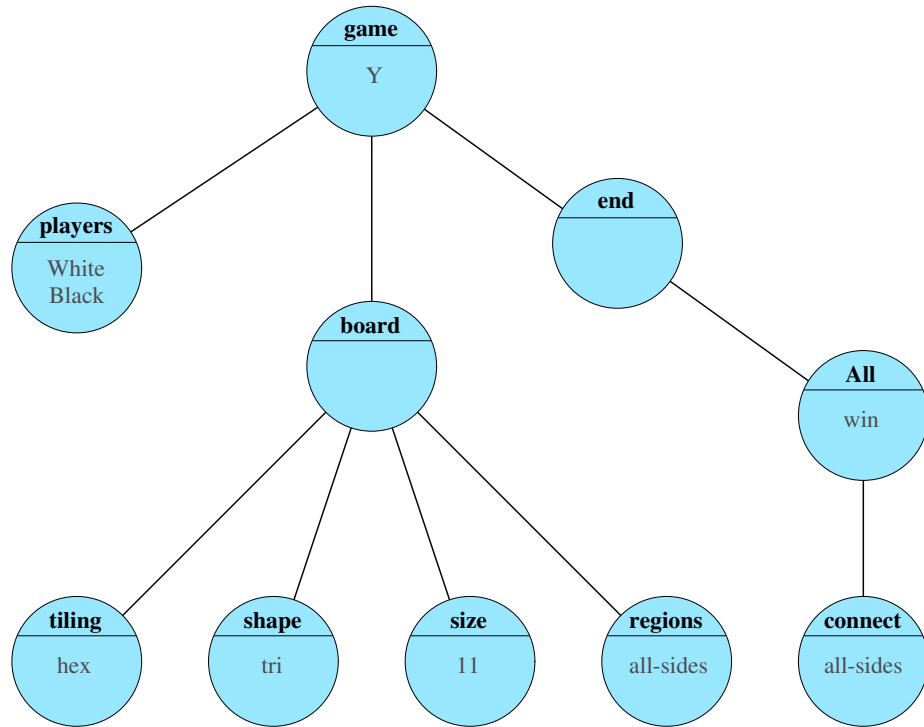
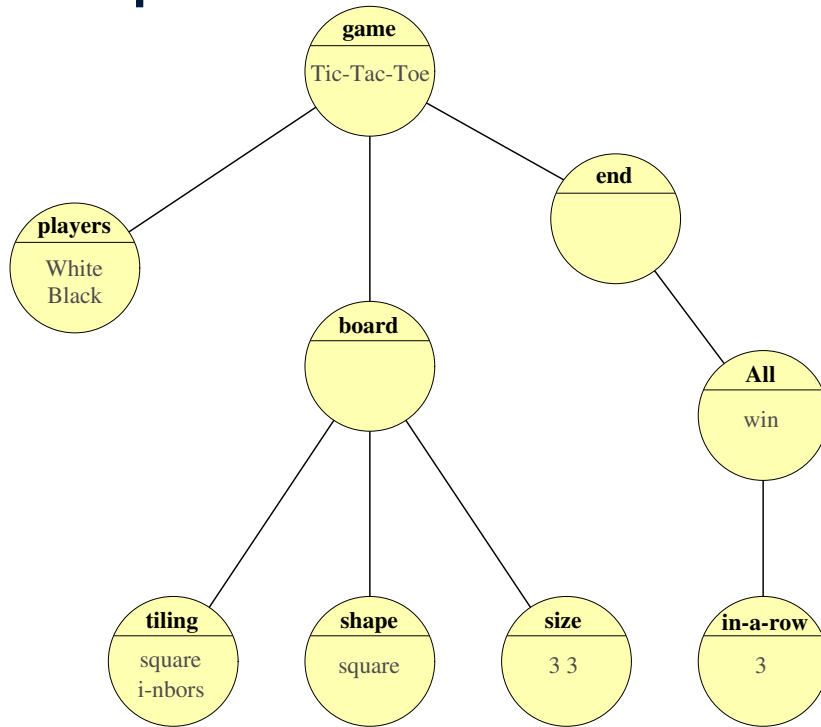
- Crossover
- Mutation



Evolving Games

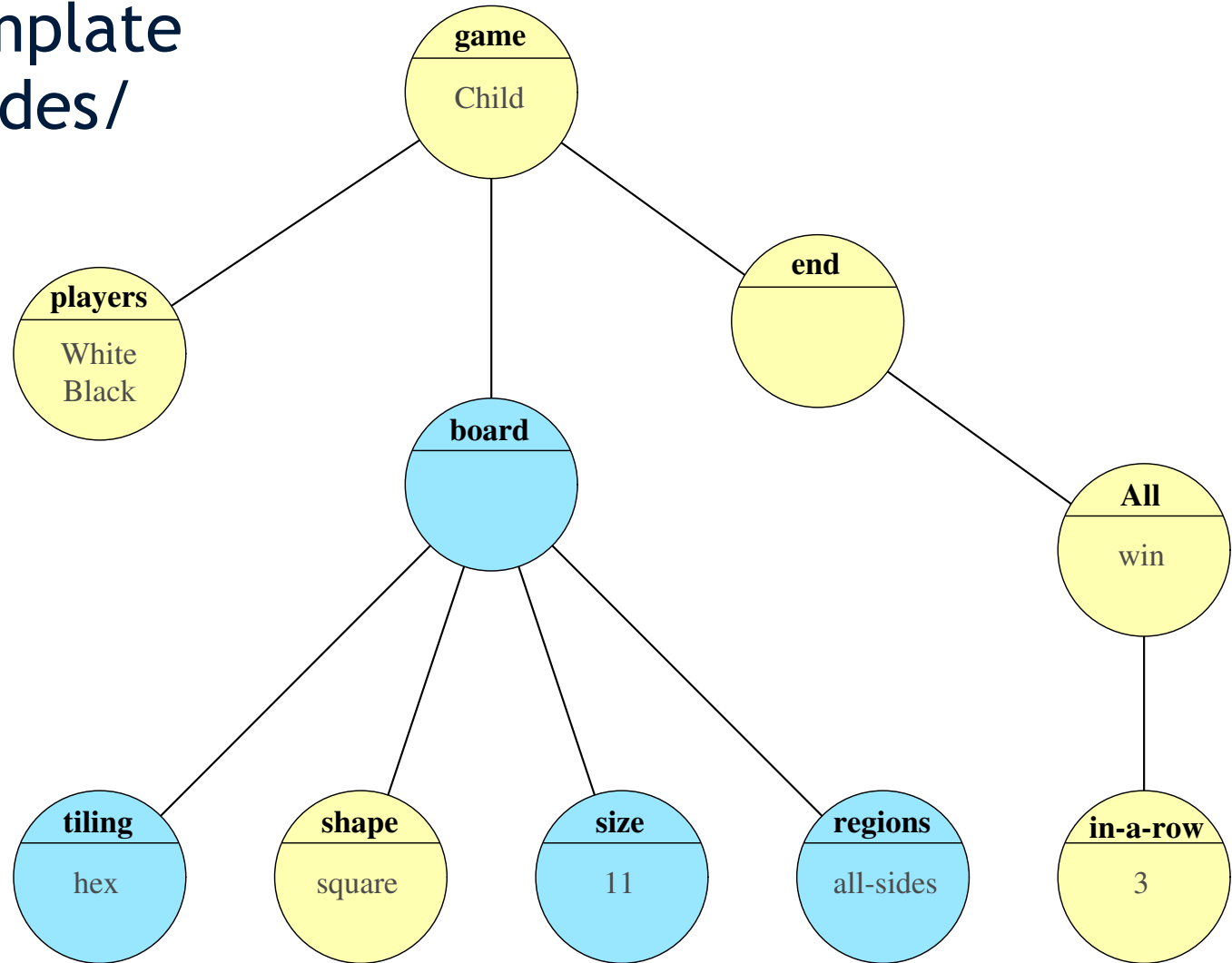
Parent games:

- Tic-Tac-Toe
- Y



Crossover

1. Choose template
2. Swap in nodes/
branches



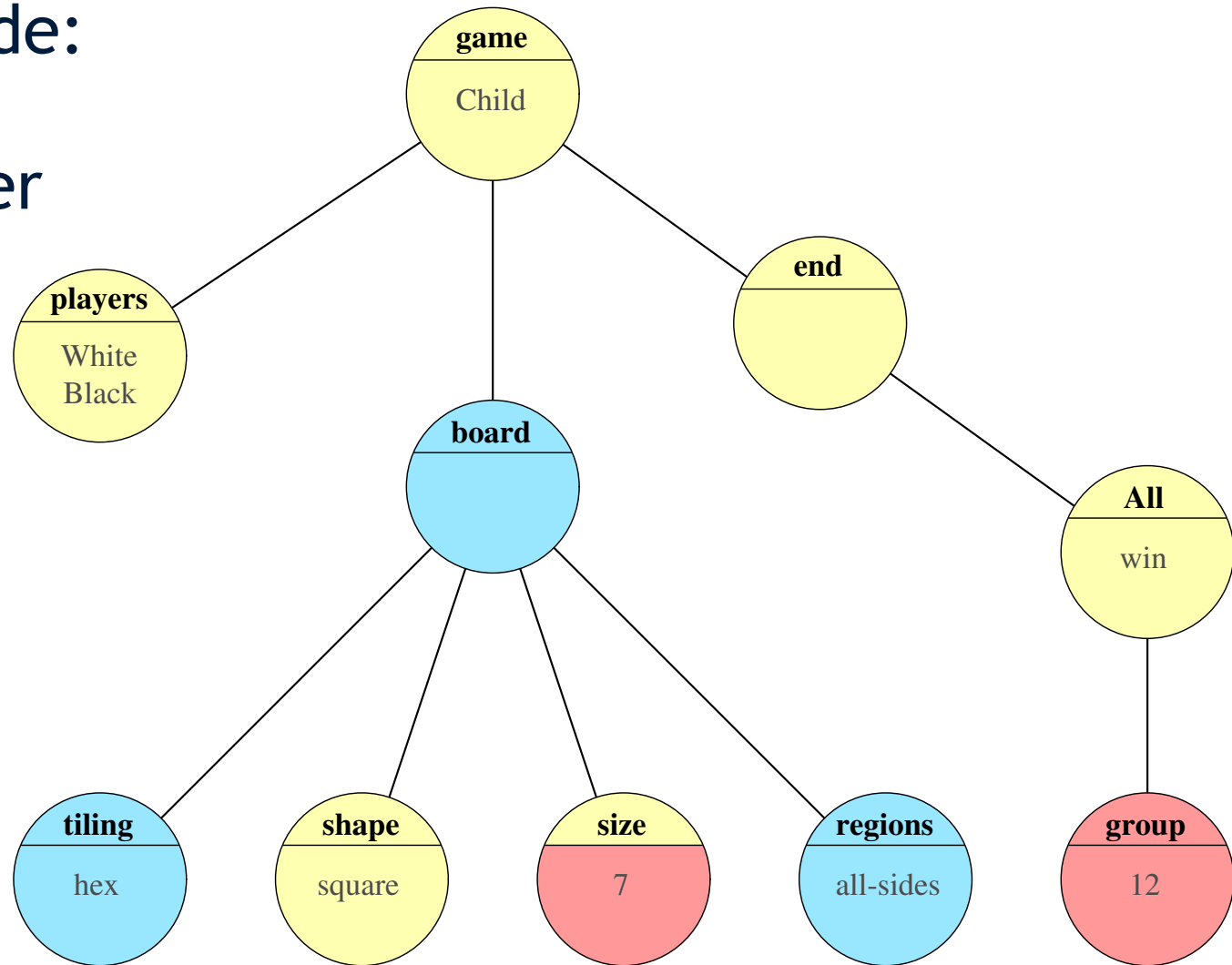
Mutation

1. Change node:

- Type
- Parameter

2. Add node

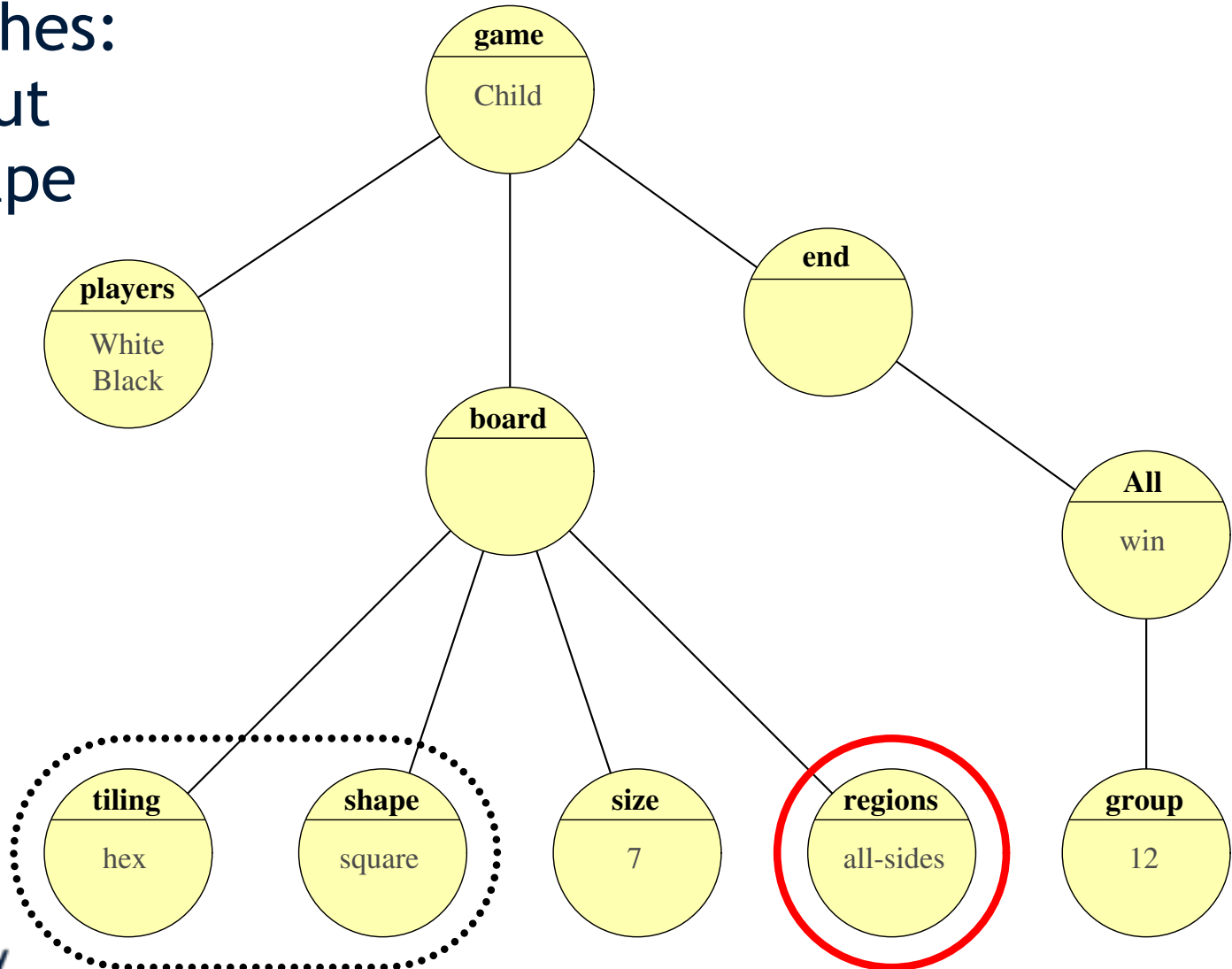
3. Remove node



Check Conflicts

Ludeme clashes:

- Hex grid but square shape

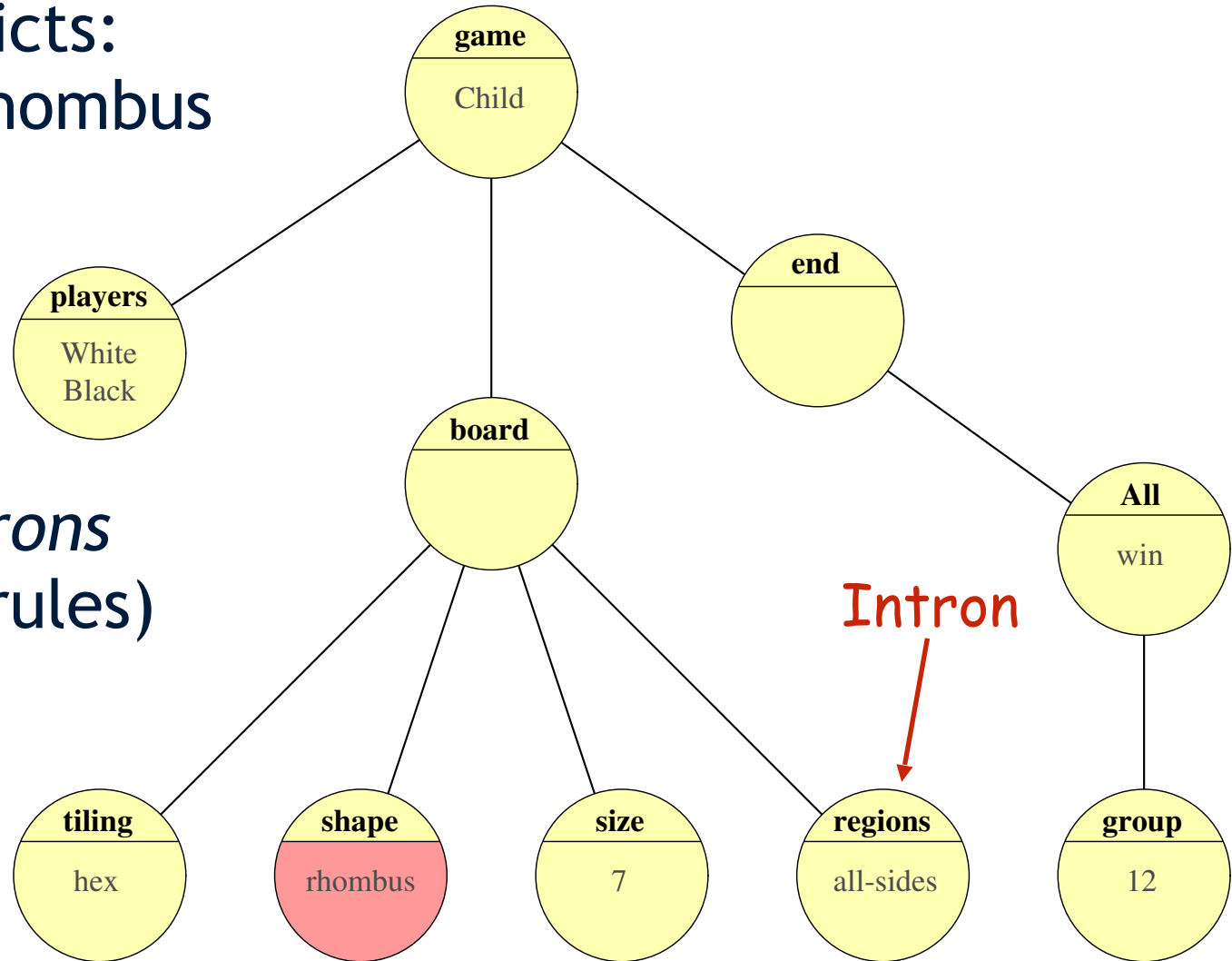


Repair

Resolve conflicts:

- Square \Rightarrow rhombus

But leave *introns*
(superfluous rules)



Introns

Unused genetic material:

- Present in *genotype* (i.e. rules)
- No effect on *phenotype* (i.e. play)

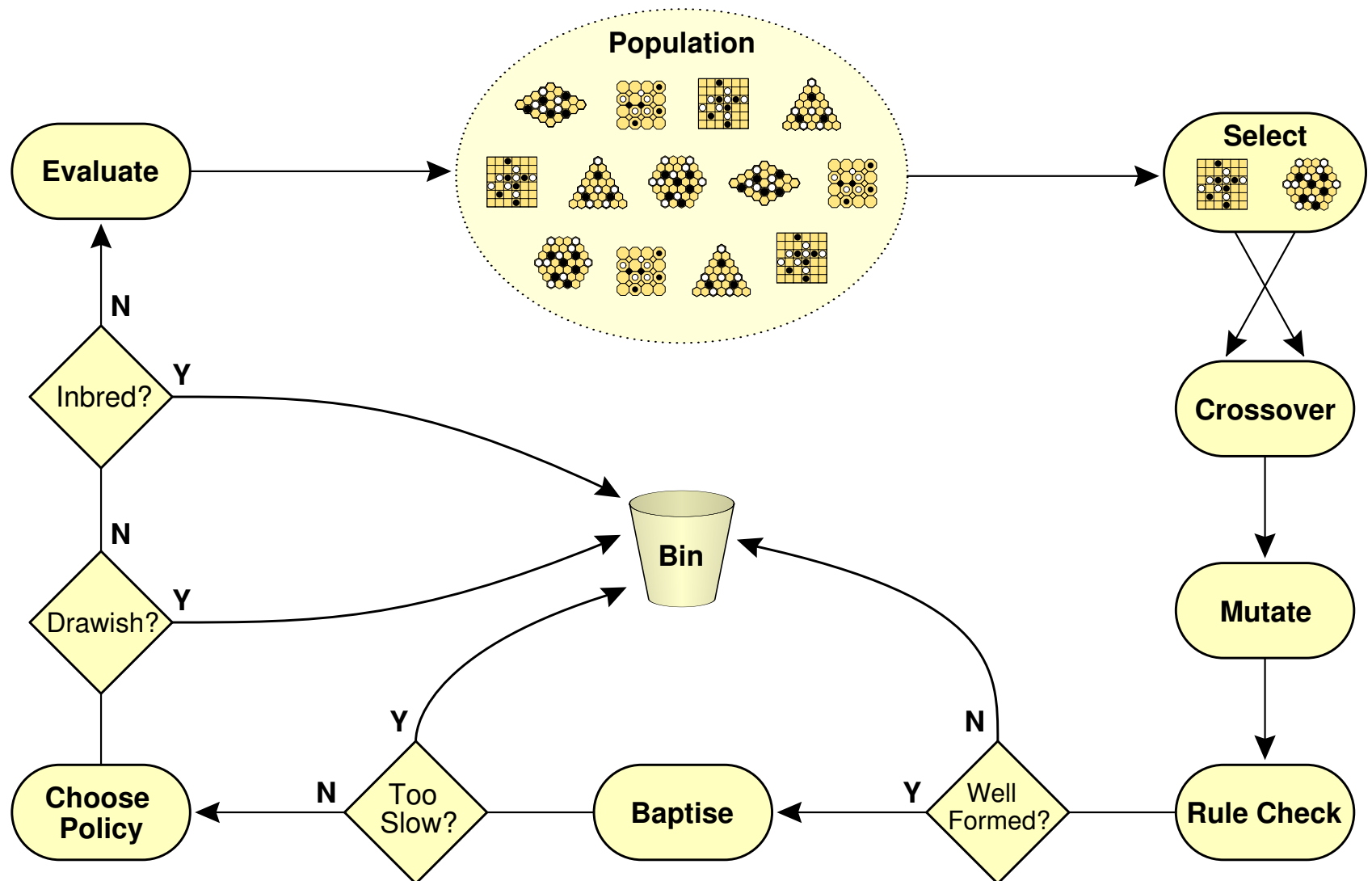
Removing introns:

- Simplest, elegant rule sets
- Actually hinders evolution!
 - Homogenous results

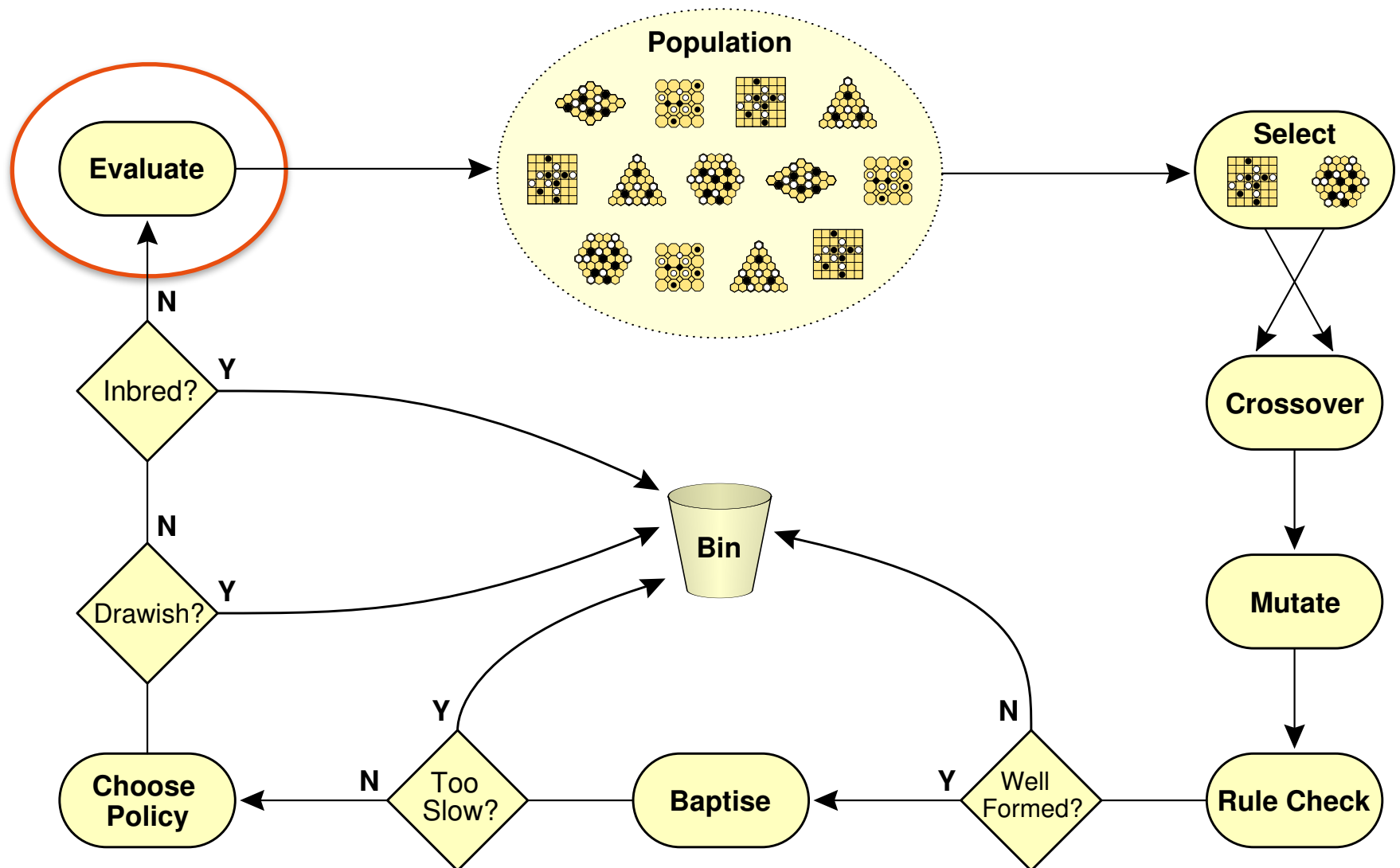
Keeping introns:

- Dormant rules become active later in new contexts
- More emergent surprises

Method



Method



Measuring Game Quality

Q. How to measure games for quality?

- Run N self-play trials between AI agents
- Look for trends in play
- Correlate with human player ranking of games

Levels of Evaluation

1. Viability (Playability) Filter

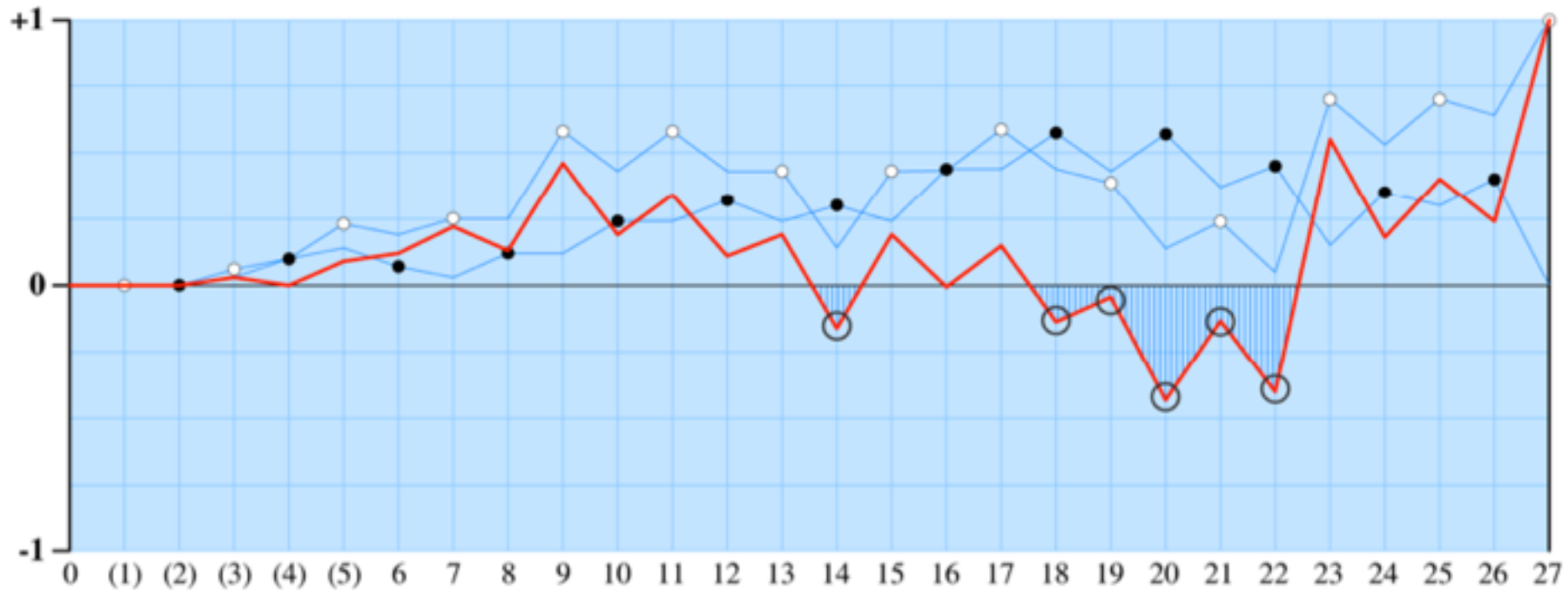
- a) *Balance*: No bias due to colour or play order
- b) *Completion*: More victories than draws
- c) *Duration*: Games not too short nor too long

2. Aesthetic Criteria

- More subtle
- Much longer to compute
- Attempt to model player experience:
 - Depth, drama, decisiveness, clarity, etc.
- 57 implemented

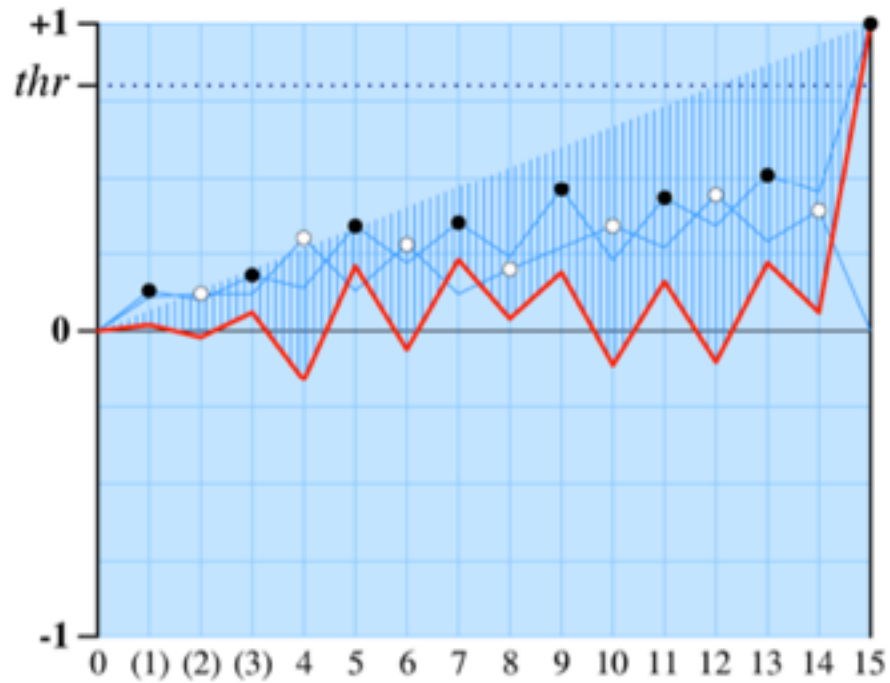
Drama

A dramatic recovery by White

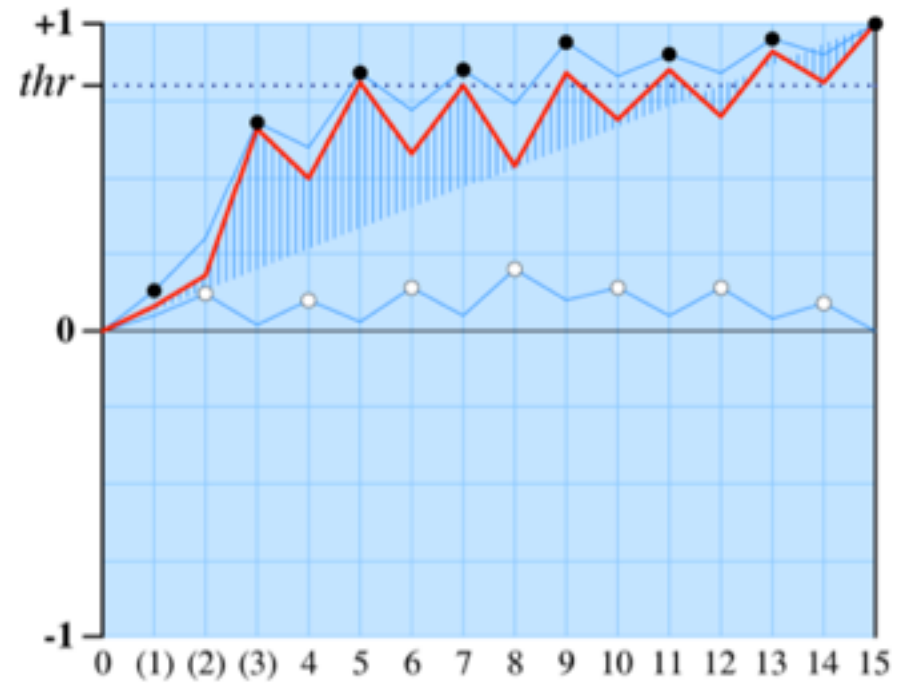


Uncertainty

Uncertain game



Certain game

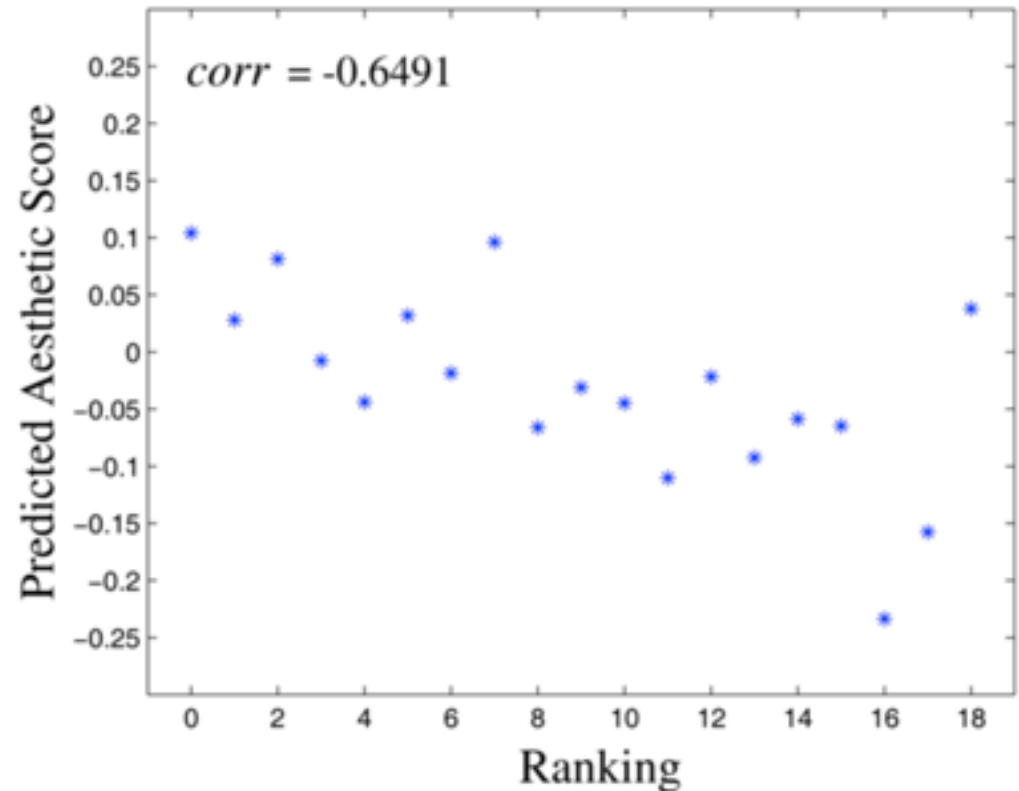


Results

- 79 source games
- 3 machines x 3 weeks
- 1,389 survived tests
- 19 deemed viable
- 0.6491 correlation

Good result, but:

- Subset of games
- Subset of population

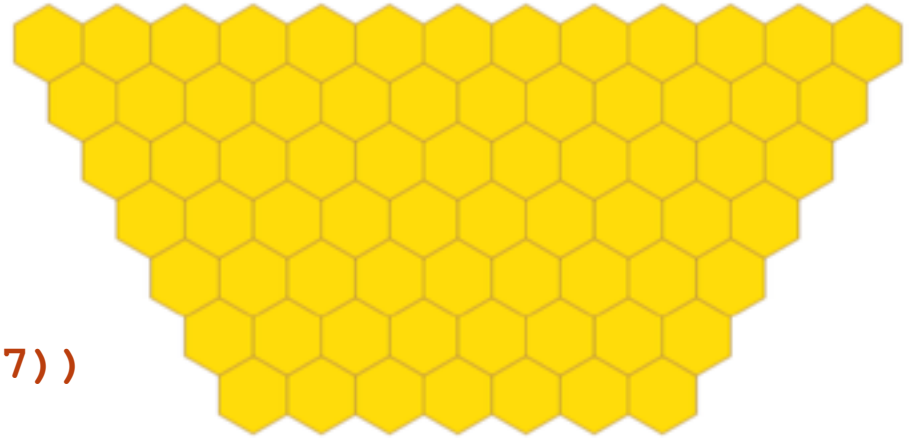


Two evolved games proved especially interesting

Game #1: Ndengrod

Computer ranking: #1

```
(game Ndengrod
  (players White Black)
  (board (tiling hex)
    (shape trapezium) (size 7 7))
  (pieces
    (Piece All (moves
      (move
        (pre (empty to))
        (action (push))
        (post (capture surround))
      )))
    (end (All win (in-a-row 5)))
  )
```



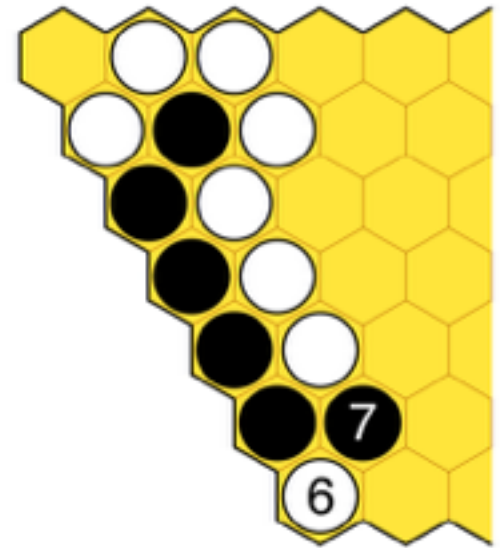
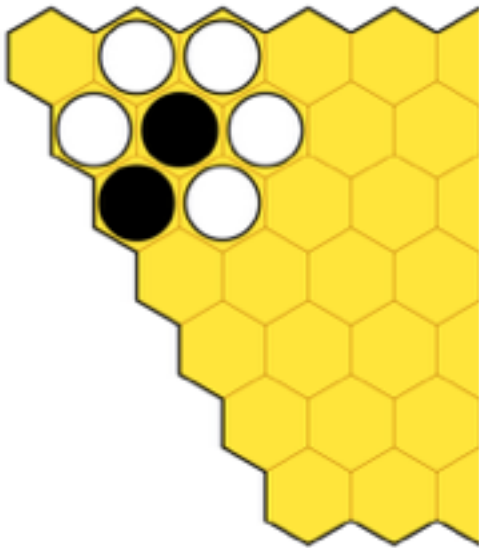
5-in-a-row with Go-like surround capture

Game #1: Ndengrod

Nice game:

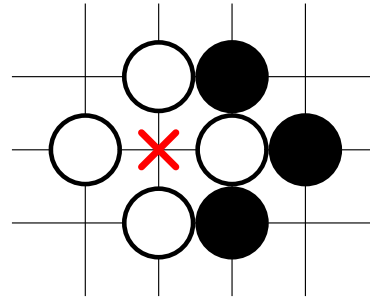
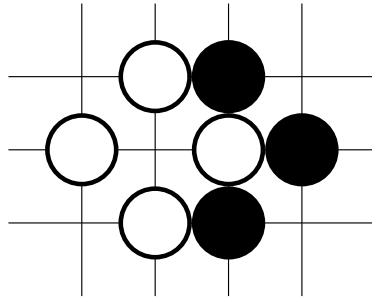
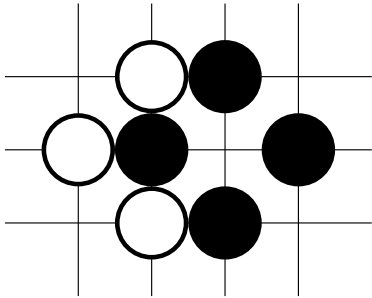
- Actually a connection game!

Interesting tactical plays



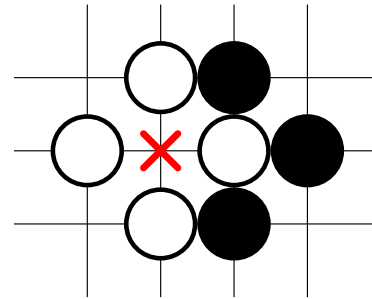
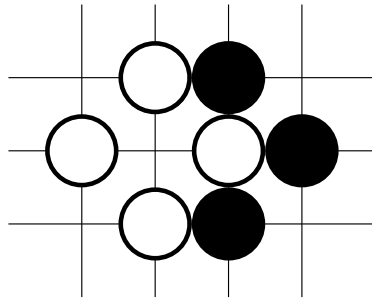
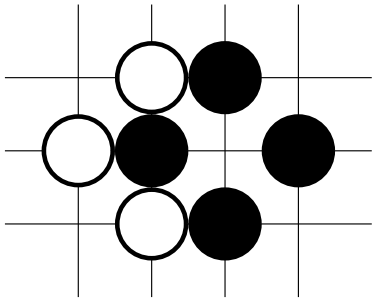
Game #1: Ndengrod

No *ko* rule:

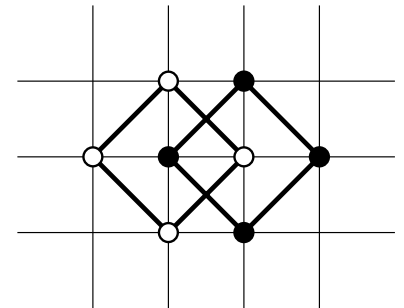
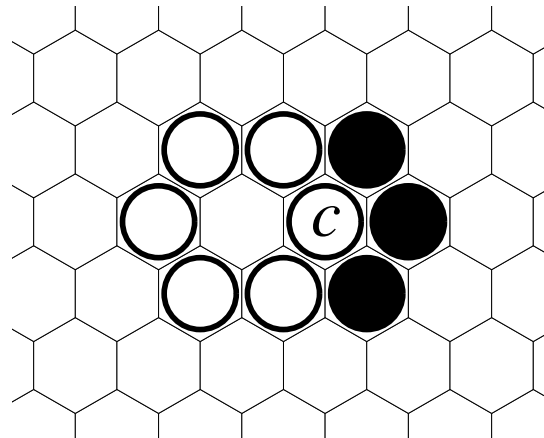
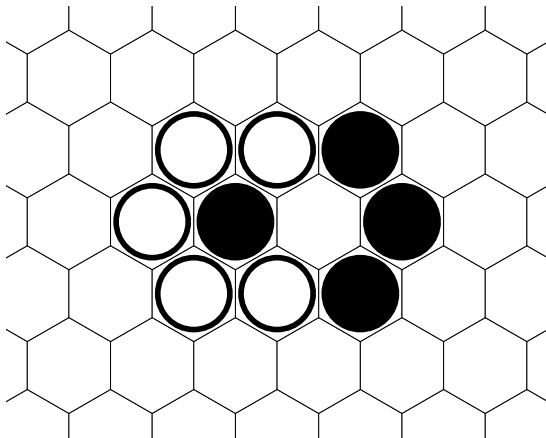


Game #1: Ndengrod

No *ko* rule:



Not needed on hexagonal grid!



Game #1: Ndengrod

Renamed as “Pentalath”

- Published in 2010
- Almost no interest from players!
- But human testers & AI agents liked it...

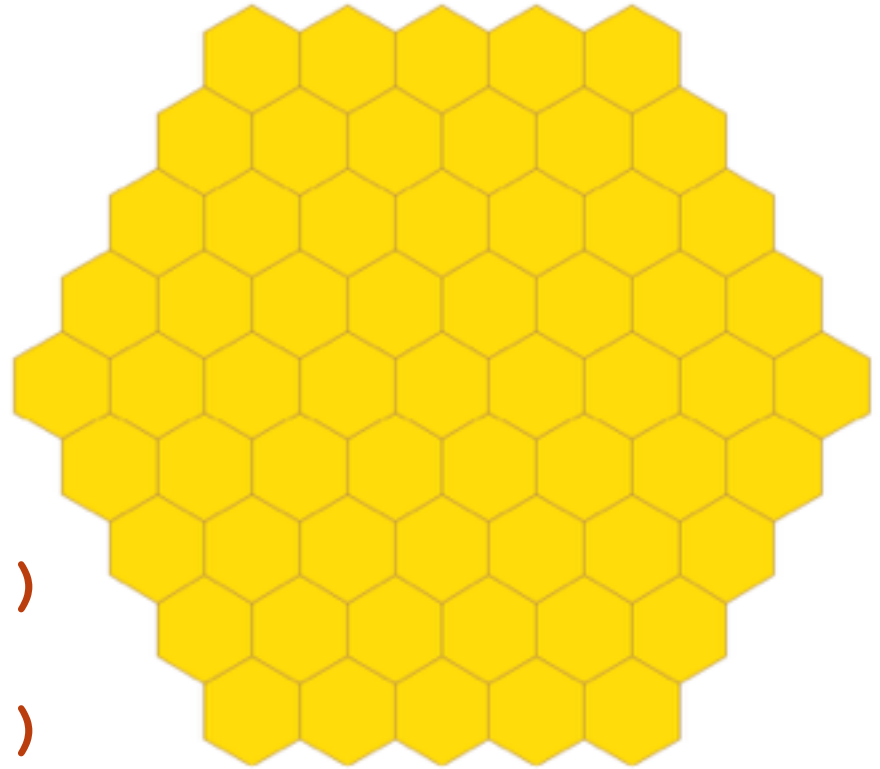
Why?



Game #2: Yavalath

Computer ranking: #4

```
(game Yavalath
  (players White Black)
  (board (tiling hex)
    (shape hex) (size 5))
  (All win (in-a-row 4))
  (All lose (in-a-row 3))
)
```

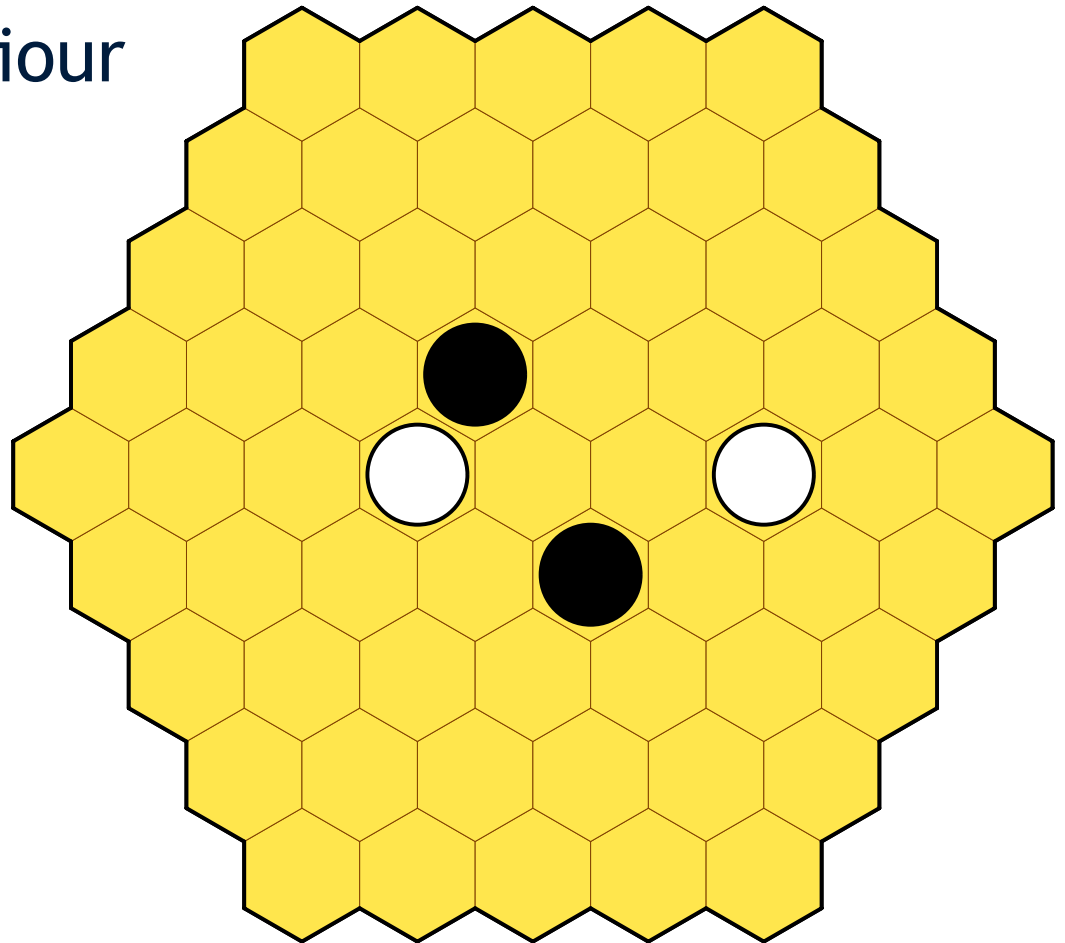


Win by making 4-in-a-row,
Lose by making 3-in-a-row beforehand

Game #2: Yavalath

Win with 4-in-a-row, lose with 3-in-a-row – *What?*

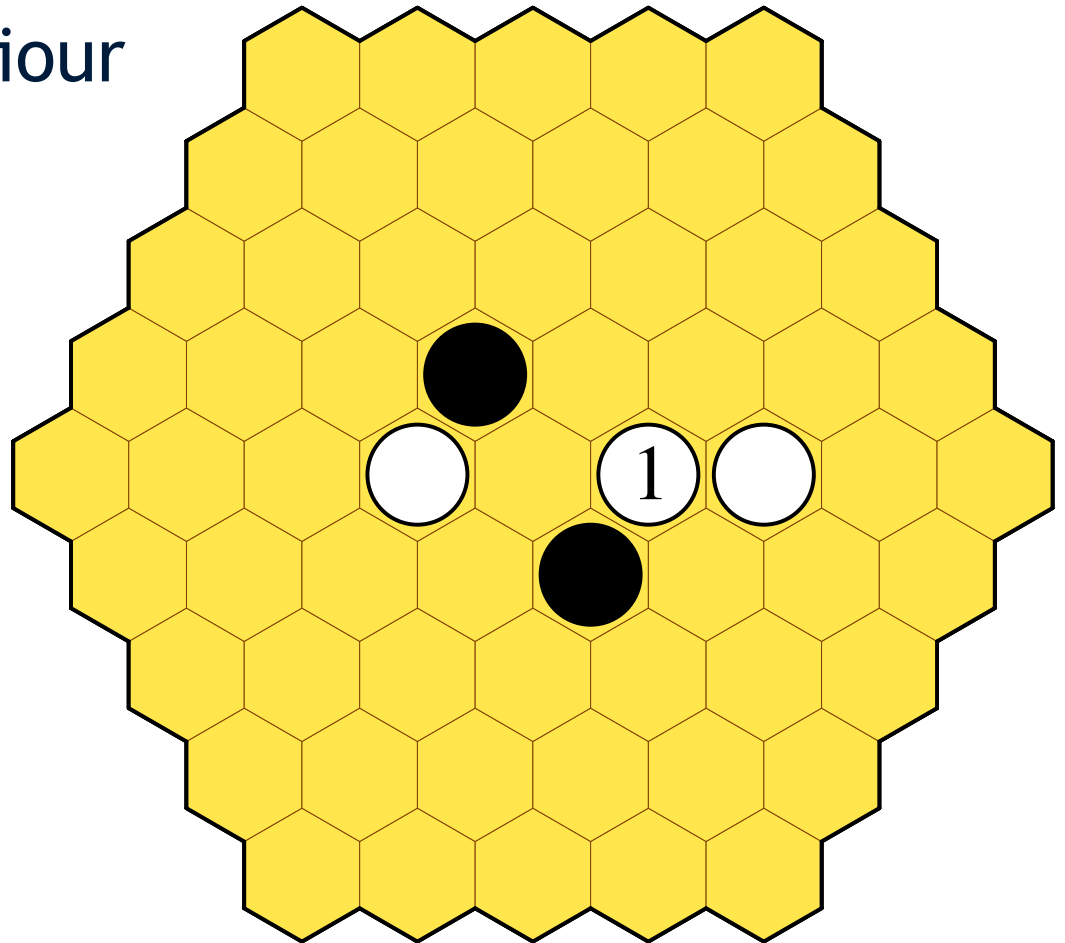
Nice emergent behaviour



Game #2: Yavalath

Win with 4-in-a-row, lose with 3-in-a-row – *What?*

Nice emergent behaviour



Game #2: Yavalath

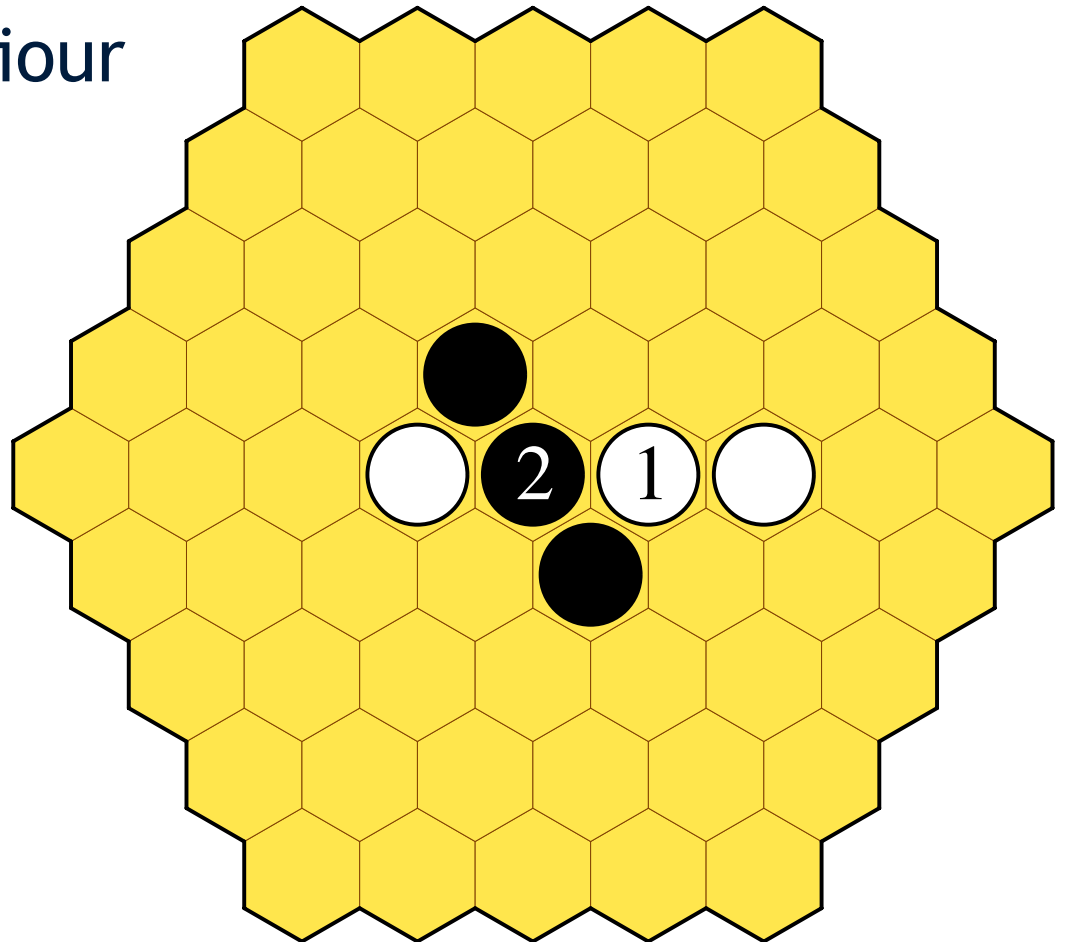
Win with 4-in-a-row, lose with 3-in-a-row — *What?*

Nice emergent behaviour

“Aha!” moments:

- Forcing moves
- Can force a win!

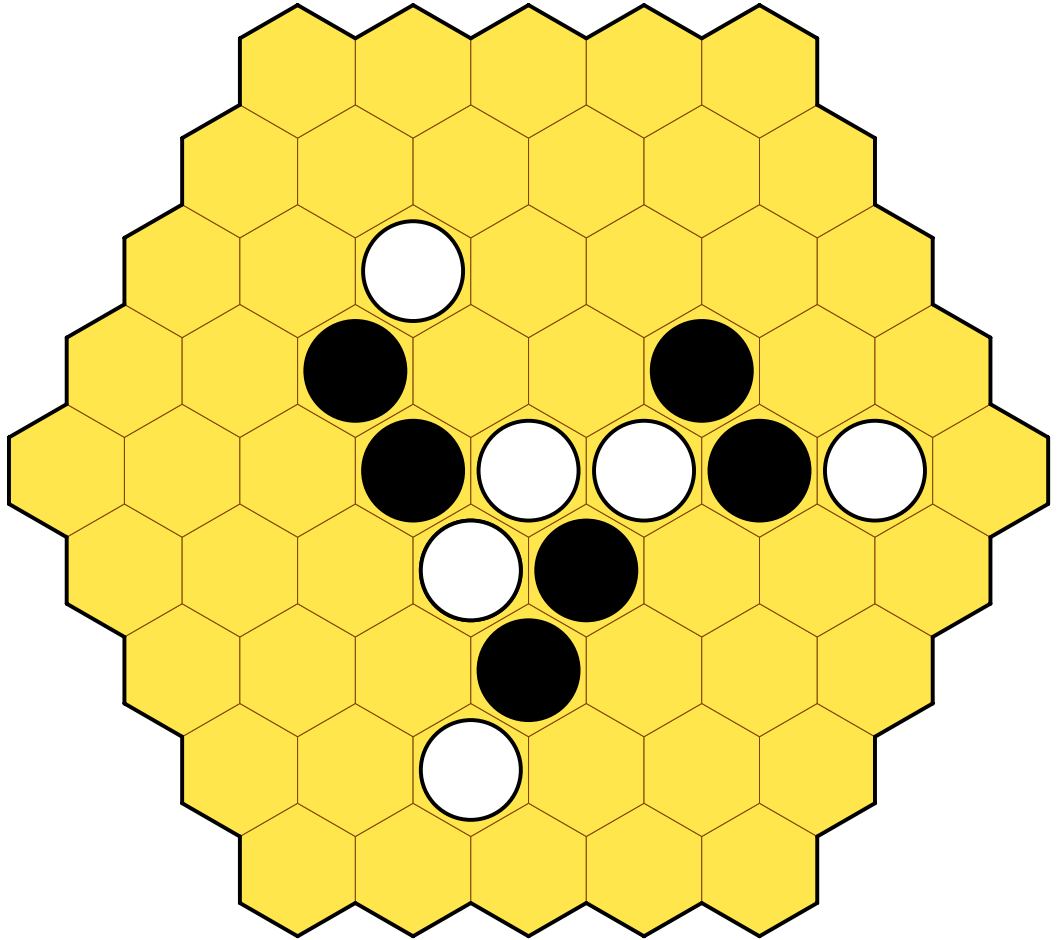
Would not have
existed without
introns



Game #2: Yavalath

Puzzle:

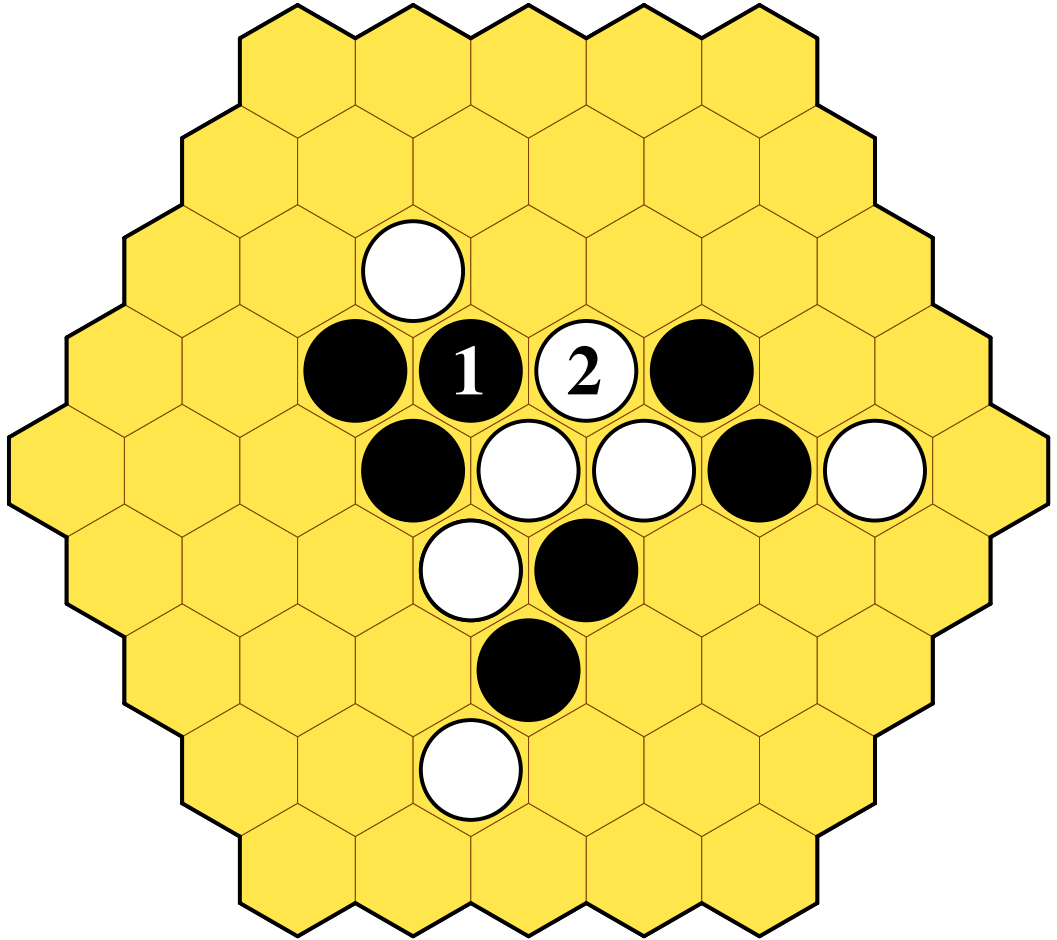
- White to play



Game #2: Yavalath

Puzzle:

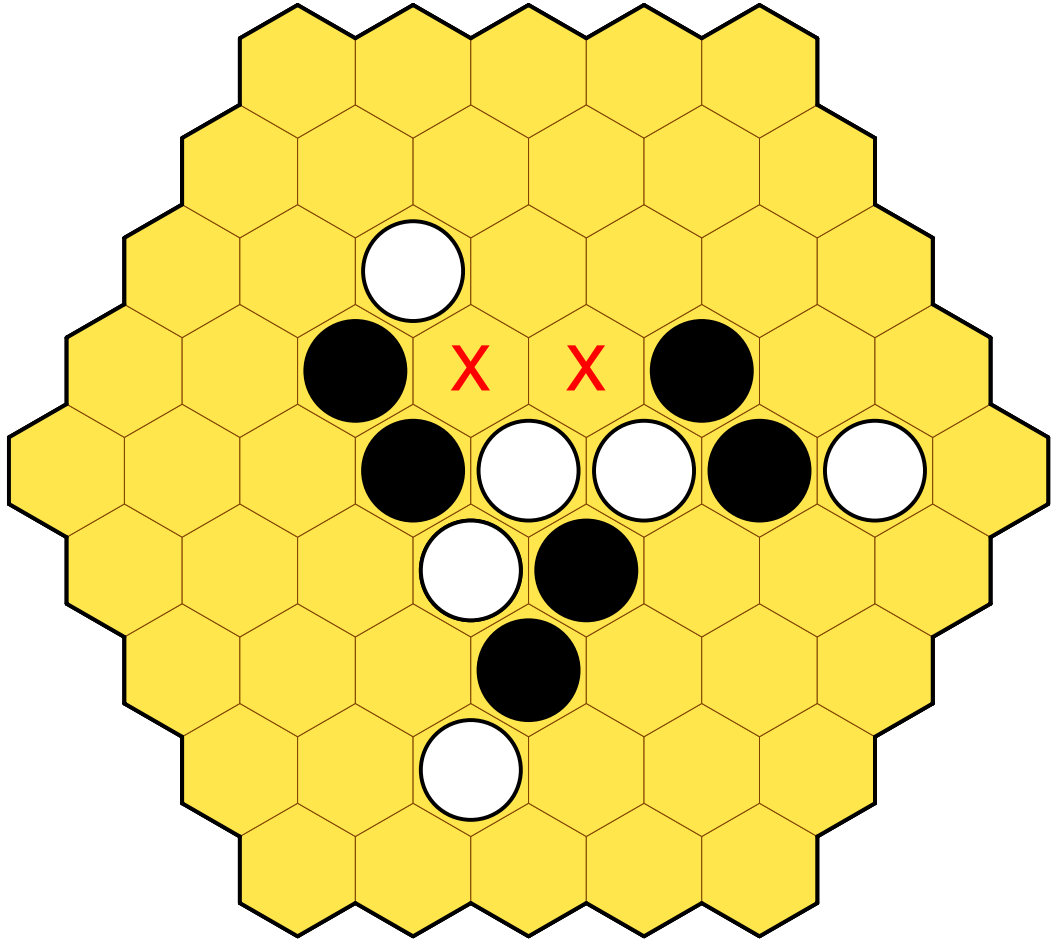
- White to play



Game #2: Yavalath

Puzzle:

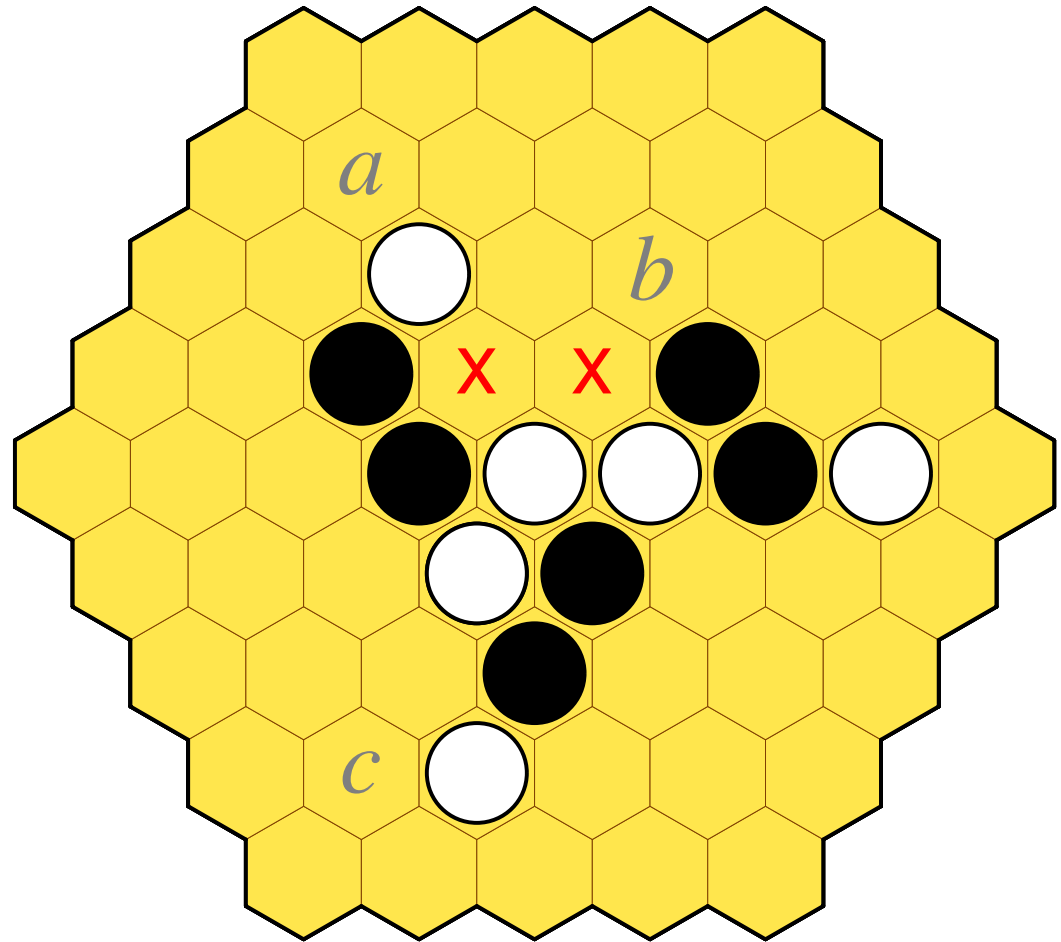
- White to play



Game #2: Yavalath

Puzzle:

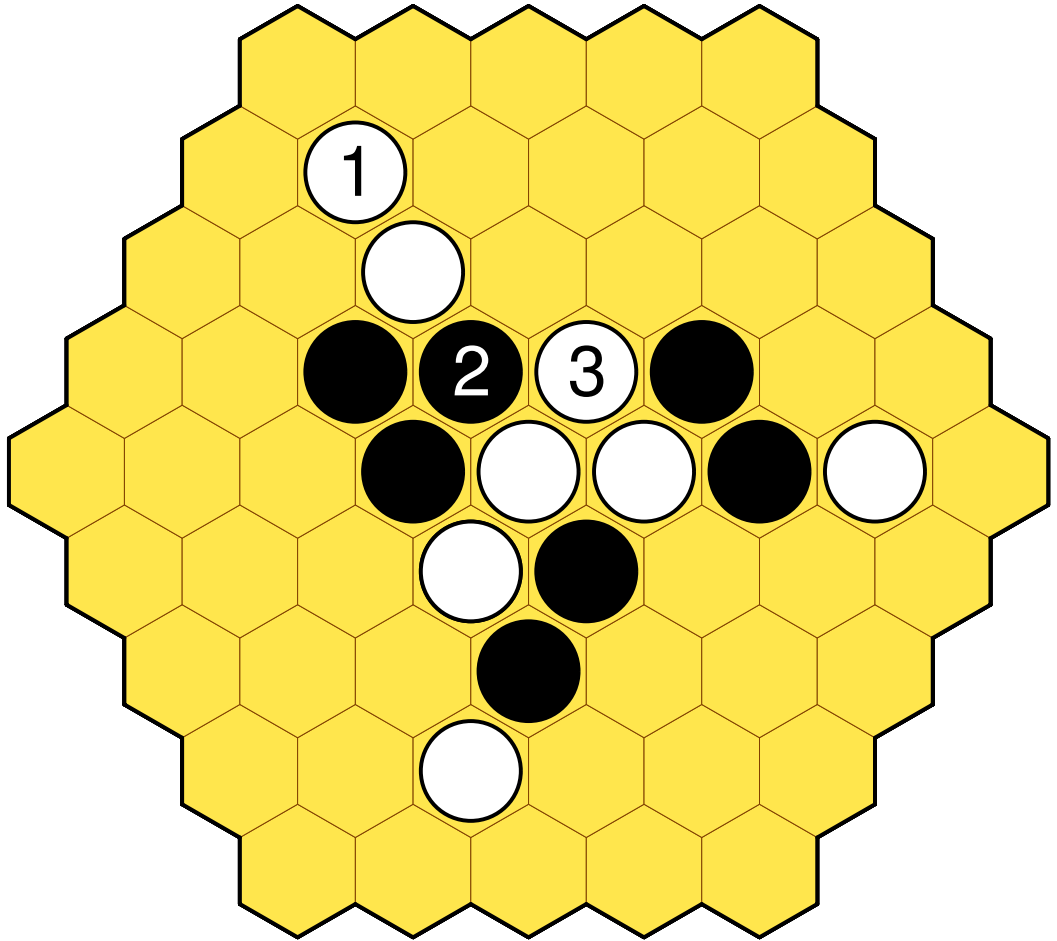
- White to play



Game #2: Yavalath

Puzzle:

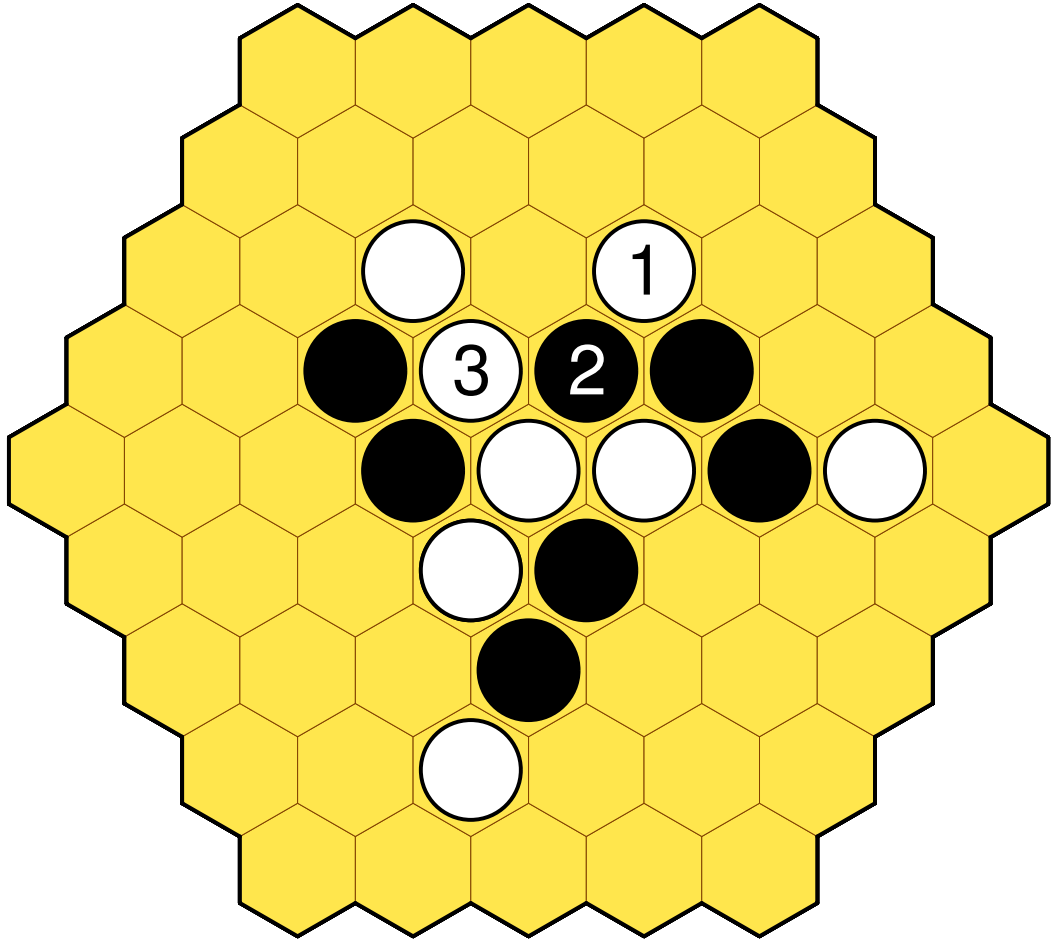
- White to play



Game #2: Yavalath

Puzzle:

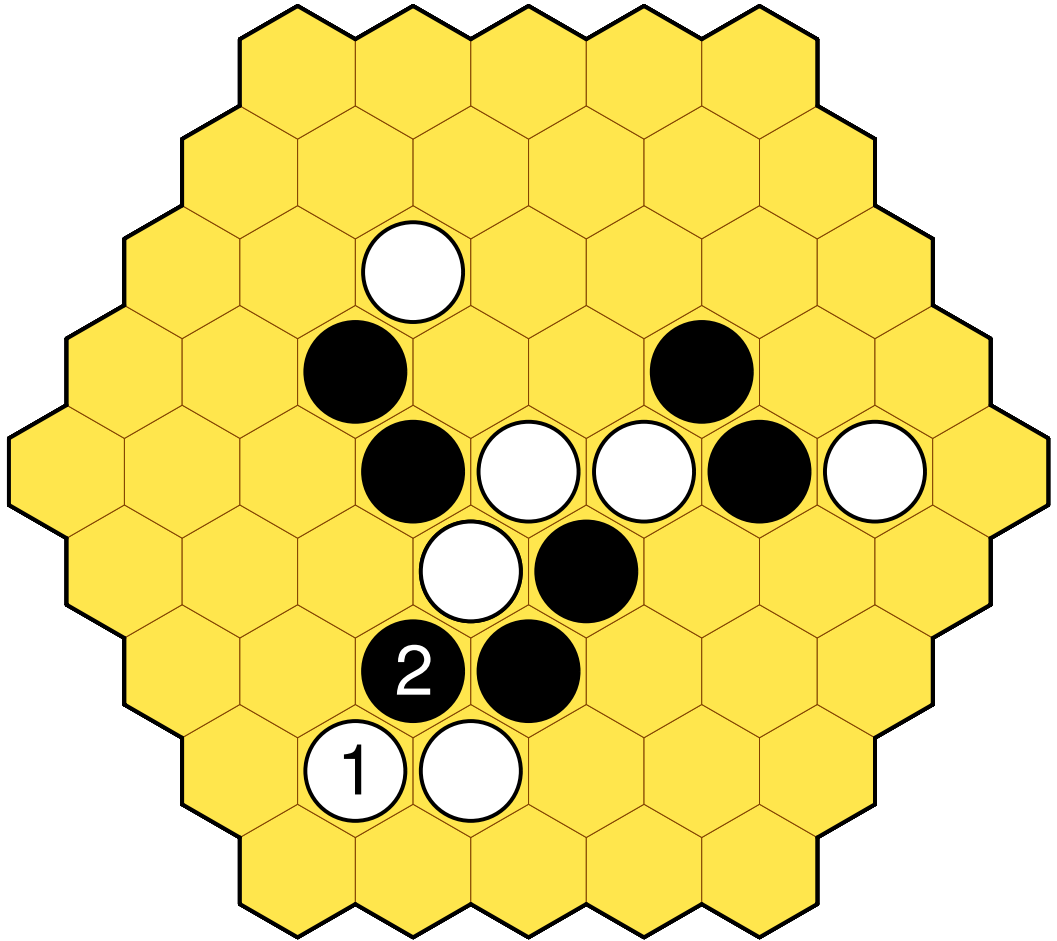
- White to play



Game #2: Yavalath

Puzzle:

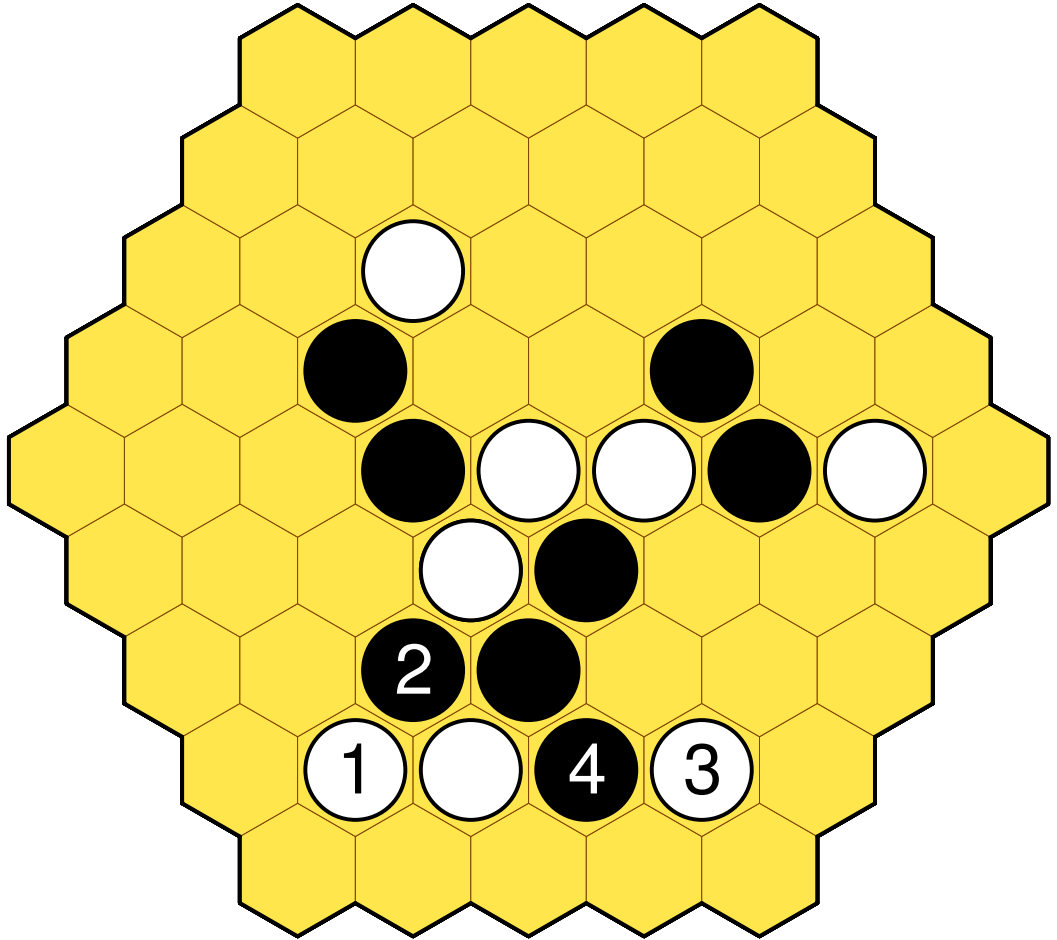
- White to play



Game #2: Yavalath

Puzzle:

- White to play



Game #2: Yavalath

Puzzle:

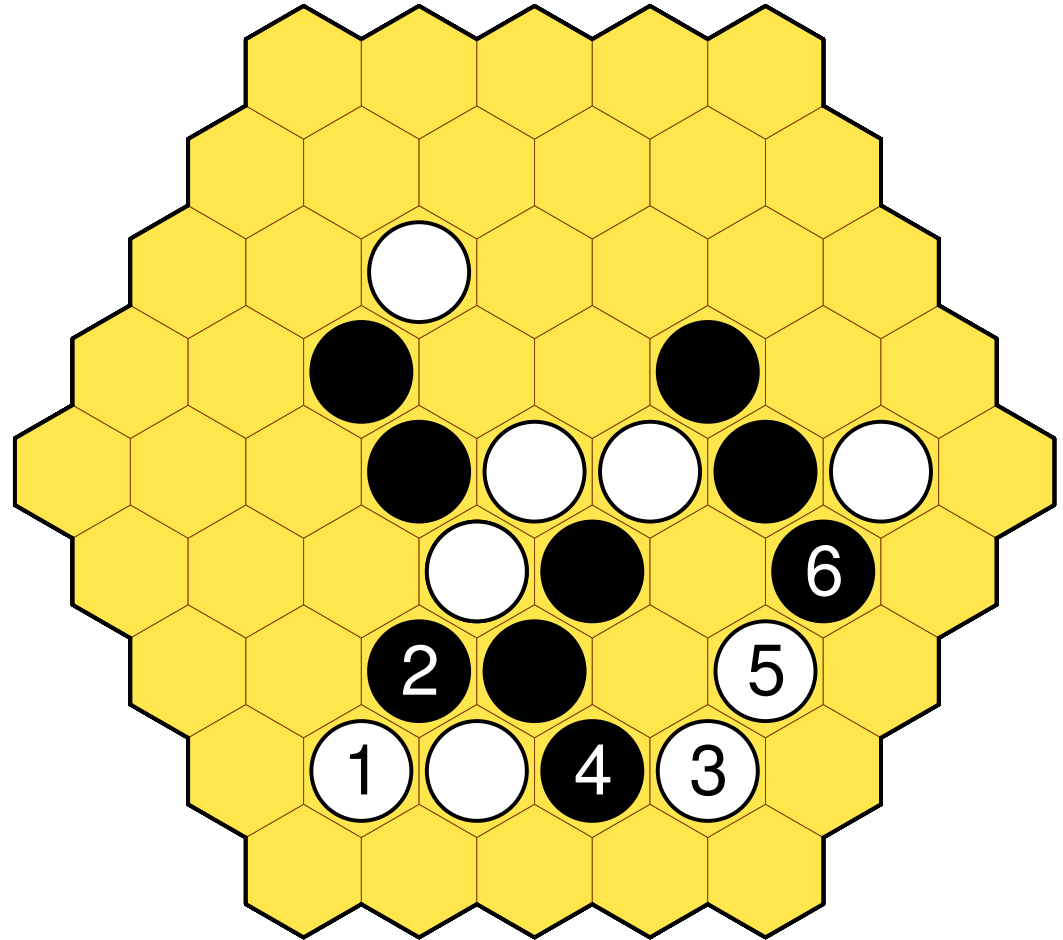
- White to play

Chains of forcing moves:

- Interesting

Puzzle from actual game:

- Sign of a good game



Game #2: Yavalath

Published in 2009:

- Still flagship product for publisher

Three-player version:

- Works well



Game #2: Yavalath

Popular with players:

- Easy to learn
- Familiar
- Unexpected twist

#3	Go
#45	Chess
#99	Yavalath
#112	Backgammon
#267	Othello
#539	Mastermind
#546	Chinese Checkers

2011: Ranked top #100 abstract board games (BGG):

- Top 2.5% of all abstract games
- Higher than any of my games!

Created new subclass of “ N but not sub- N ” games:

- Tritt, Cross, Tailath, Morro, Epsilon, Manalath, ...

Comparison



RANK: OVERALL 3,995 ABSTRACT 111

7.2 **Yavalath** (2007)
185 Ratings & 80 Comments · GeekBuddy Analysis

2-3 Players
Community: 2-3 — Best: 2


10 Min
Playing Time

Age: --
(Not provided by publisher)
Community: 8+

Weight: 2.32 / 5
'Complexity' Rating ⓘ

Designer: Cameron Browne
Artist: Néstor Romeral Andrés
Publisher: (Web published), nestorgames

Settings Share Edit



RANK: OVERALL 9,655 ABSTRACT 428

6.7 **Pentalath** (2009)
40 Ratings & 15 Comments · GeekBuddy Analysis

2 Players
Community: 2 — Best: 2

20 Min
Playing Time

Age: 8+
Community: 8+

Weight: 2.00 / 5
'Complexity' Rating ⓘ

Designer: Cameron Browne
Artist: Néstor Romeral Andrés
Publisher: (Web published), nestorgames

Settings Share Edit

Comparison

N-in-a-row:

Yavalath

+

Pentalath

+

Emergent behaviour:

Branching factor:

Game length:

Familiarity:

Depth:

Barrier to entry:

Comparison

N-in-a-row:

Yavalath

+

Pentalath

+

Emergent behaviour:

+

+

Branching factor:

Game length:

Familiarity:

Depth:

Barrier to entry:

Comparison

	Yavalath	Pentalath
<i>N</i> -in-a-row:	+	+
Emergent behaviour:	+	+
Branching factor:	+	+
Game length:	+	+
Familiarity:		
Depth:		
Barrier to entry:		

Comparison

N-in-a-row:

Yavalath

+

Pentalath

+

Emergent behaviour:

+

+

Branching factor:

+

+

Game length:

+

+

Familiarity:

High

Low

Depth:

Medium

High

Barrier to entry:

Low

High

Current Work

Digital Ludeme Project:

- 5-person, 5-year project
- ERC Consolidator Grant (€2m)
- Started April at Maastricht University



Digital
Ludeme
Project

Computational study:

- World's traditional strategy games
- Recorded human history

Aim:

To improve our understanding of ancient games through modern AI techniques



European Research Council
Established by the European Commission



Objectives

1. ***Model***: The world's traditional strategy games in a single (playable) digital database
2. ***Reconstruct***: Missing knowledge about ancient games with improved accuracy
3. ***Map***: The spread of games and associated mathematical ideas throughout history

Scope: ~1,000 games from ~3,500BC to ~1900AD

Problem

The rules for ancient games are almost always lost

Modern understanding based on (often flawed) reconstructions

Q. How to evaluate reconstructions of ancient games for quality?



Ancient Game Evaluation

Player preferences:

- Vary by culture, period, individual, etc.
- No universal indicators of quality

BUT

- Flaws are universal!
- Viability filter still valid:
 - Bias
 - Drawishness
 - Game length
- Can eliminate bad rule sets



Example: Hnefatafl

- Vikings, Scandinavia, c.400BC
- No rules recorded



1732: Linnaeus saw Tablut played
Recorded rules in Latin

1811: Smith translated (badly) to English:
“likewise the king” => “except the king”
King’s side always wins

1913: Murray published biased rules
Became de facto, corrected ever since

Example: Hnefatafl

Could the biased rule set be the correct one?

No!

We have the original transcription and can easily see the translation error, and why it was made

The original game was un(or at least less)biased

In general, games should provide a fair contest



Example: Mu Torere

- Maori, New Zealand, 18thC
 1. Move to adjacent empty cell
 2. Lose if no moves
- Opening rule:
 - 1a. First piece moved must be adjacent to an enemy piece
- Most historical accounts include it:
 - Two do not



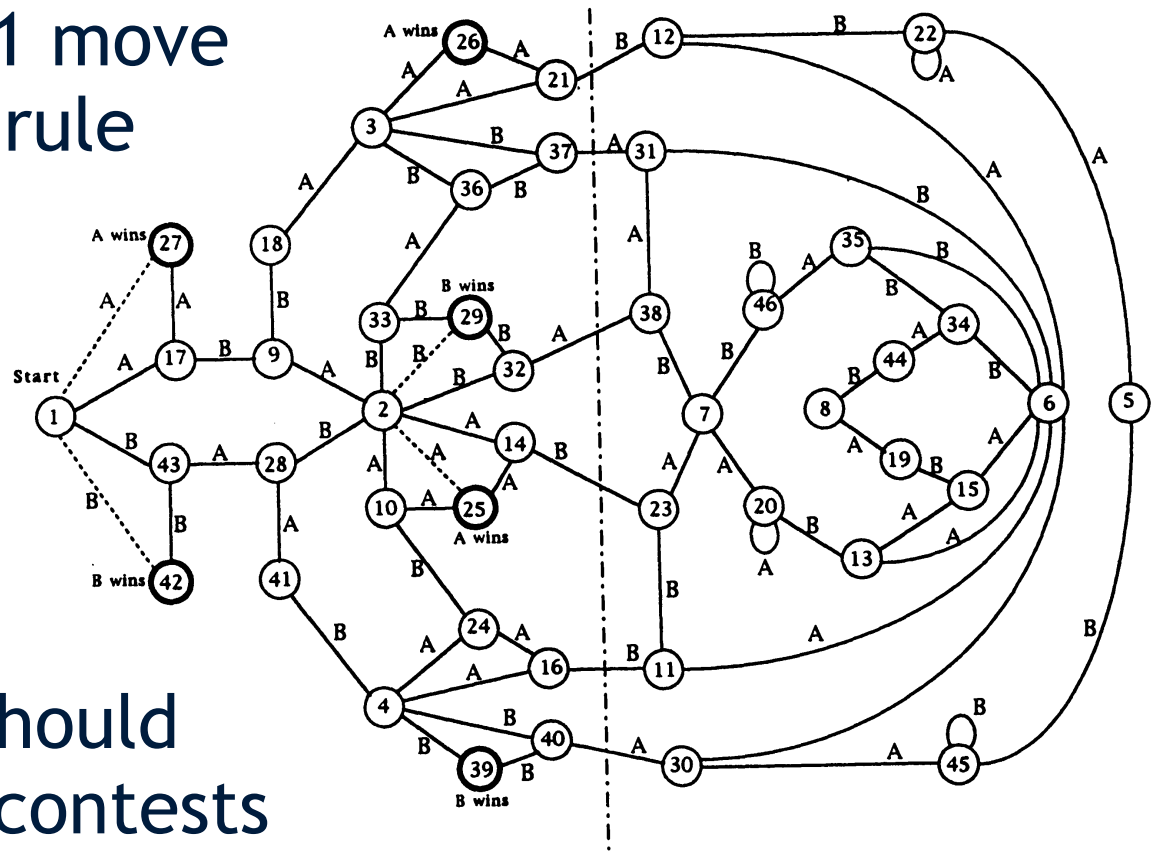
Who is right?

Example: Mu Torere

Straffin (1995):

- Full game tree expansion
- Game ends after 1 move without opening rule

Obviously wrong!



In general, games should provide non-trivial contests

Playability

Even ancient games should be:

- Non-biased
- Non-drawish
- Non-trivial but finite

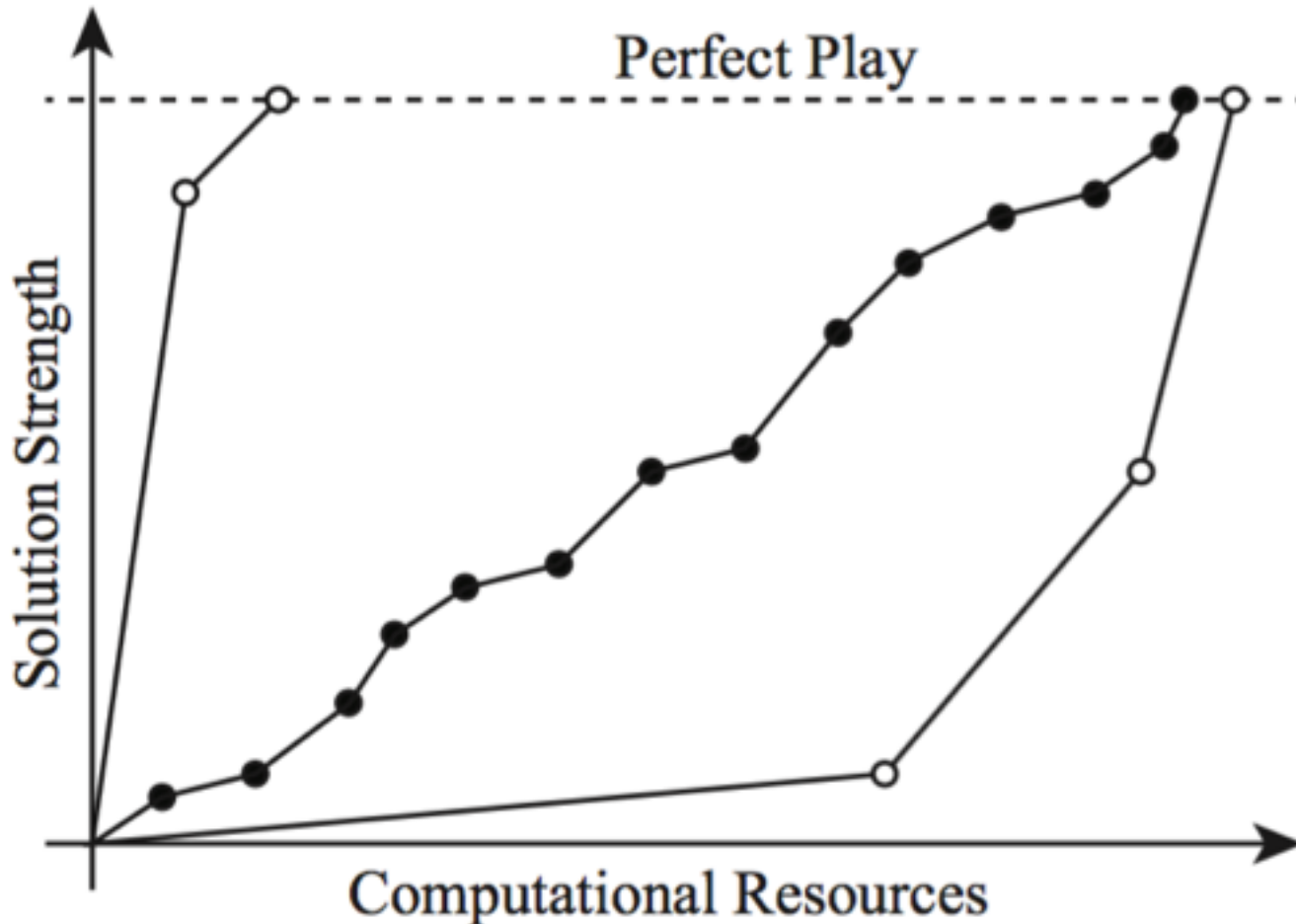
Q. Are there more subtle indicators?

Allis *et al.* (1991) “Games Solved:
Now and in the Future”:

- “...intellectual challenge neither too simple nor too hard.”

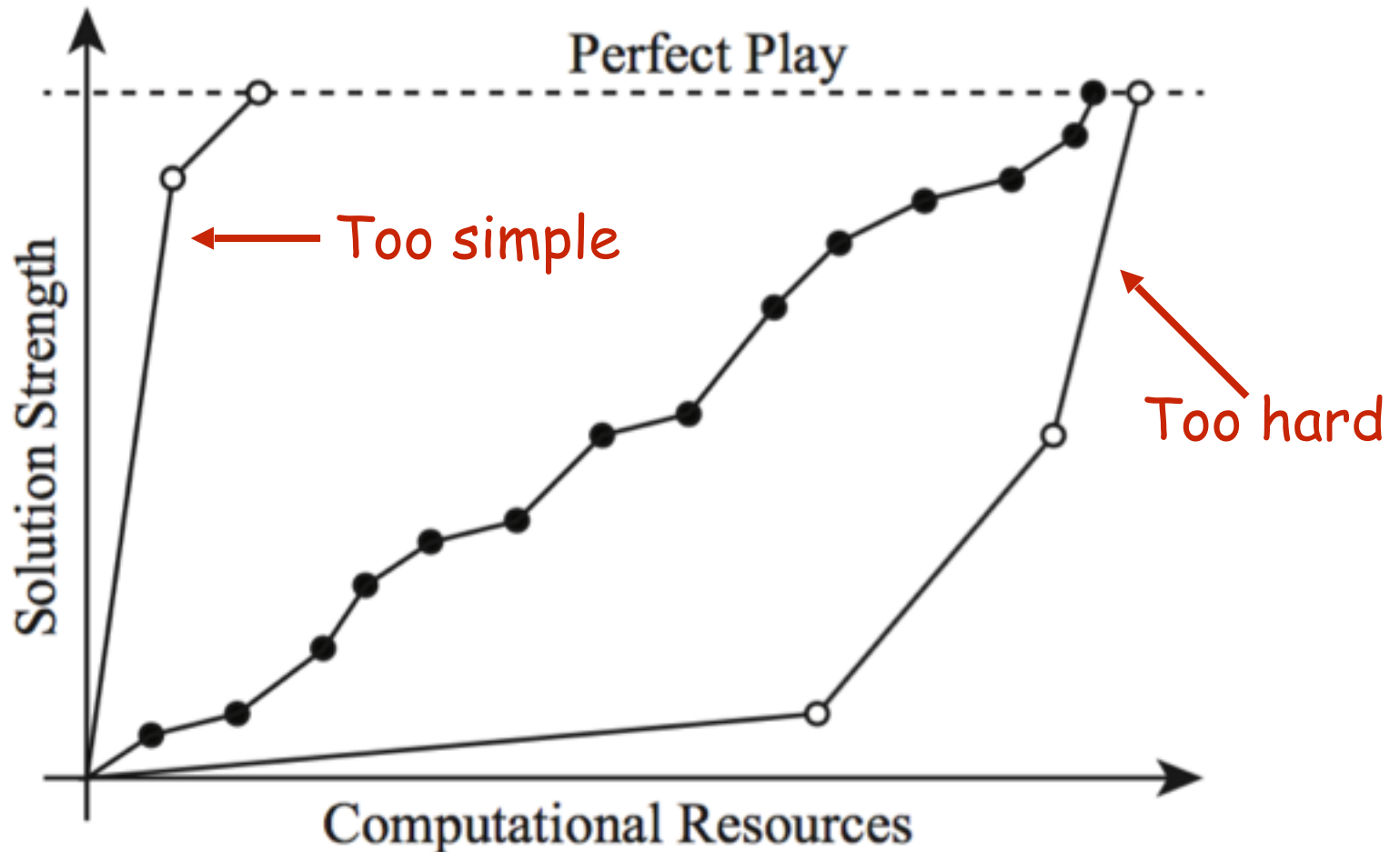
Strategy Ladder

Lantz *et al.* (2017) “Depth in Strategic Games”



Strategy Ladder

Lantz *et al.* (2017) “Depth in Strategic Games”



Strategic Potential

Strategic potential = Potential to provide ongoing series of interesting learnable strategies for players

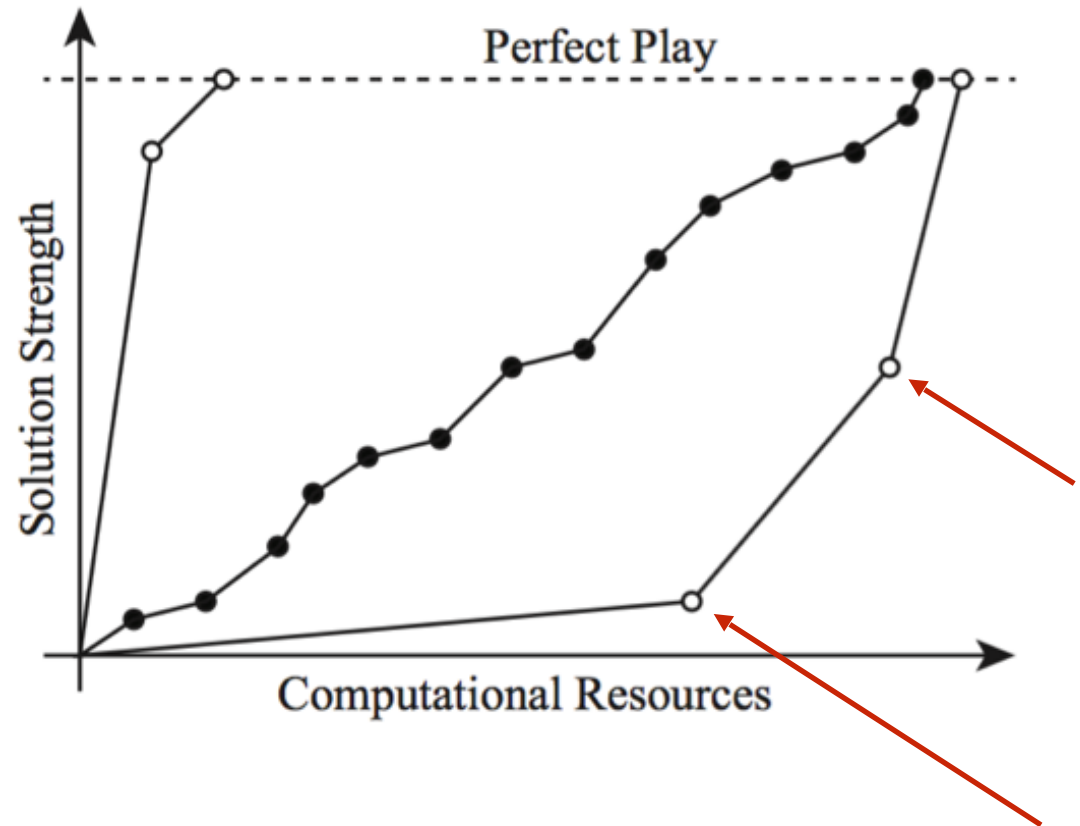
- Simple strategies to start with:
 - Low barrier to entry
- Increasingly complex latent strategies:
 - Strategic depth
 - Build on existing knowledge
 - Maximises replayability

“Minute to learn, a lifetime to master”

Game #1: Ndengrod

Strategies:

- Difficult
- Unusual
- No hints



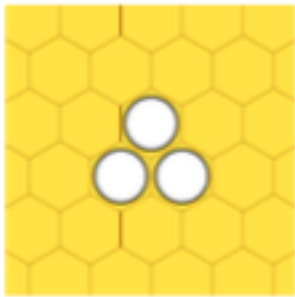
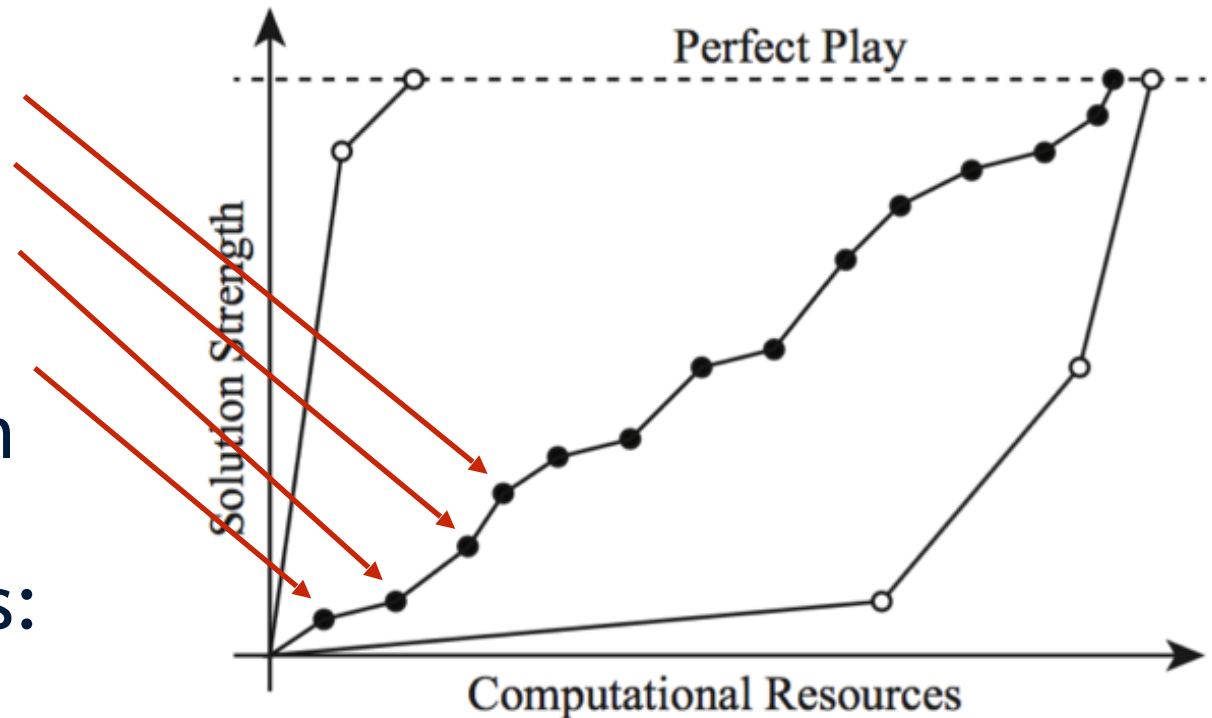
Too strange and difficult!

Game #2: Yavalath

Strategies:

- Immediate
- Dependent
- Linear accumulation

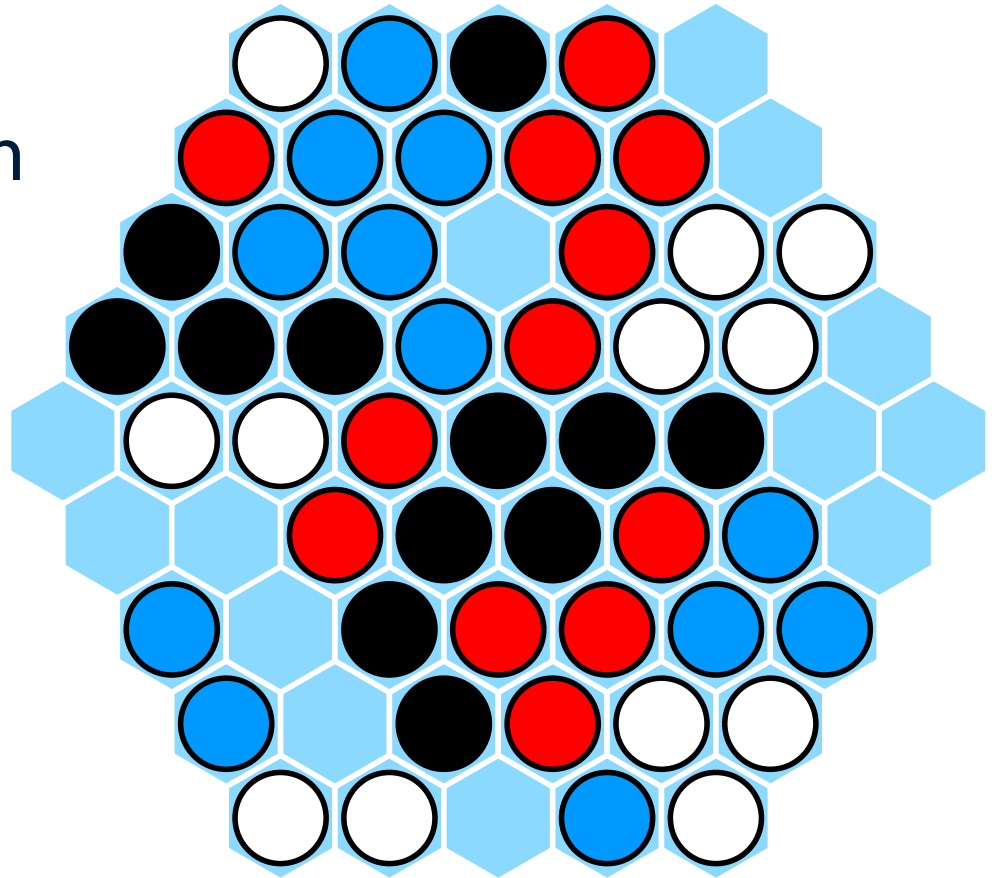
Strong patterns:



Importance of Strategies

Omega (2010):

- Place 1 piece of each colour per turn
- Score = product of own group sizes



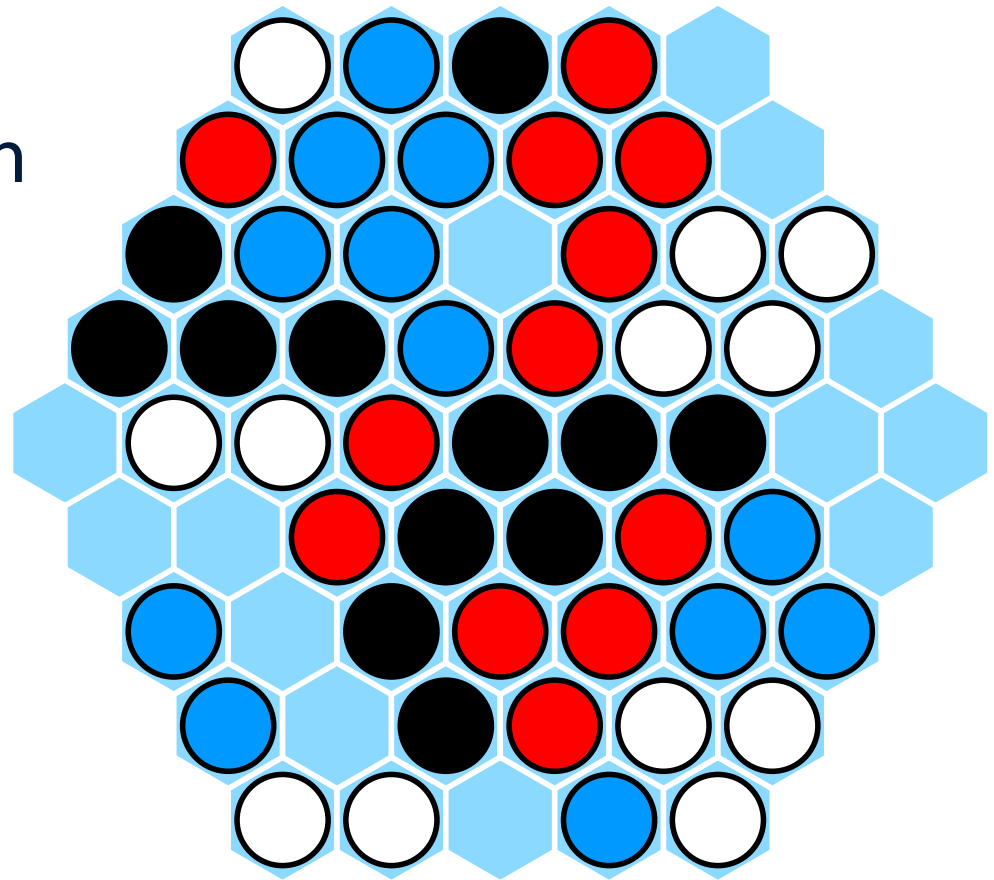
Importance of Strategies

Omega (2010):

- Place 1 piece of each colour per turn
- Score = product of own group sizes

Mental Bookkeeping

- Confusing, opaque
- Planning difficult
- Random moves
- No tension
- Boring and unpopular



White: $1 \times 2 \times 2 \times 3 \times 4 = 48$

Red: $1 \times 2 \times 4 \times 5 = 40$

Blue: $1 \times 2 \times 3 \times 6 = 36$

Black: $1 \times 4 \times 7 = 28$

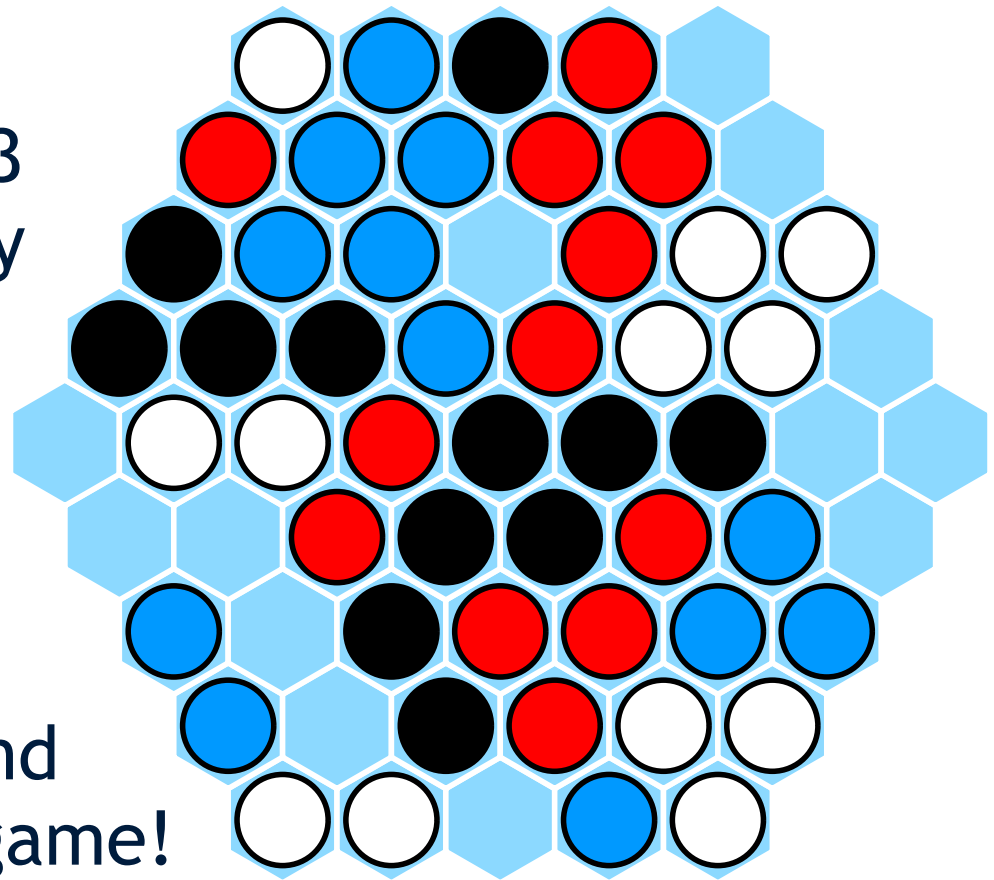
Importance of Strategies

Winning Strategy

- Form groups of size 3
- Observed in UCT play
- Mathematical proofs

Result

- Concrete strategy
- Intuitive, clear, fun
- Both a connection and an anti-connection game!



Simple strategy transformed
this game

White: $1 \times 2 \times 2 \times 3 \times 4 = 48$

Red: $1 \times 2 \times 4 \times 5 = 40$

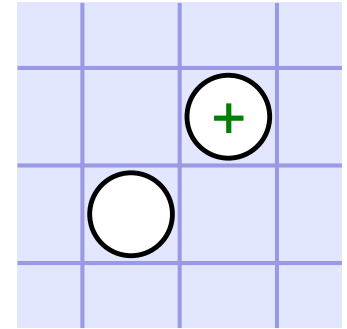
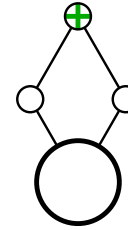
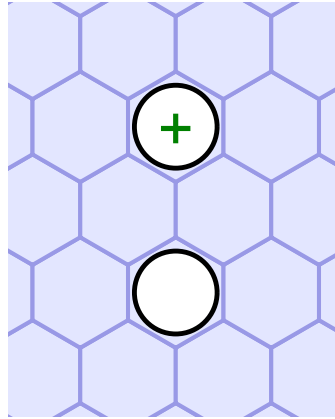
Blue: $1 \times 2 \times 3 \times 6 = 36$

Black: $1 \times 4 \times 7 = 28$

Strategic Features

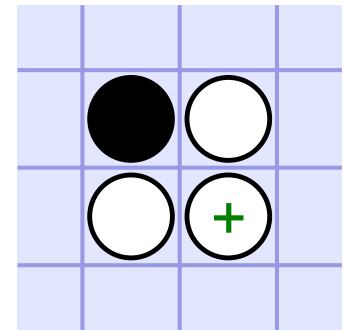
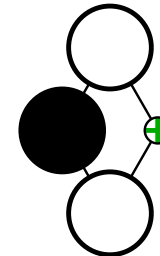
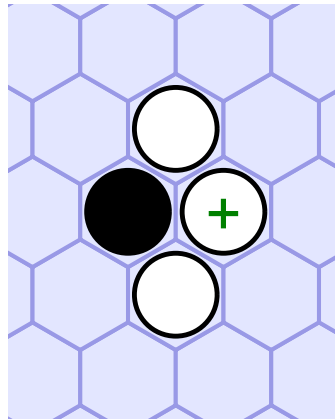
MCTS move planning:

- Bias playouts
- Features (geometric piece patterns)
- Learnt by self-play



e.g. Hex patterns:

- Effective: 55% \Rightarrow 85%



Geometry independent

- Transfer to other bases

Strategic Features

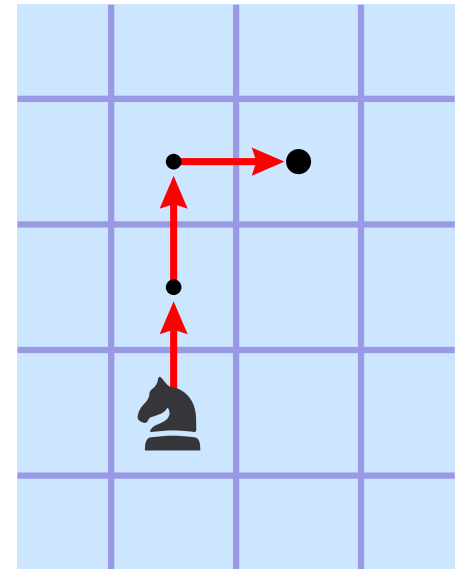
Representation

- Relative cell locations
- Graph of underlying board
- Turtle steps through adjacent cells:
 - F, B, L, R

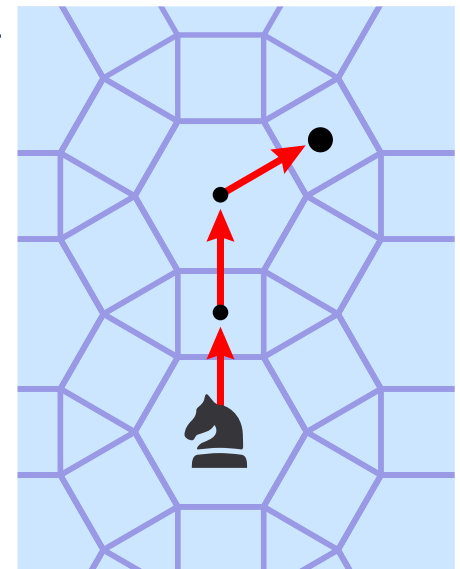
e.g. Knight move: {F, F, R, F} = {0, 0, 1}

Advantages

- Transfer between geometries
- Efficient + small memory footprint
- Human-comprehensible explanations



$$P_k = \{0,0,1\}$$



Strategic Features

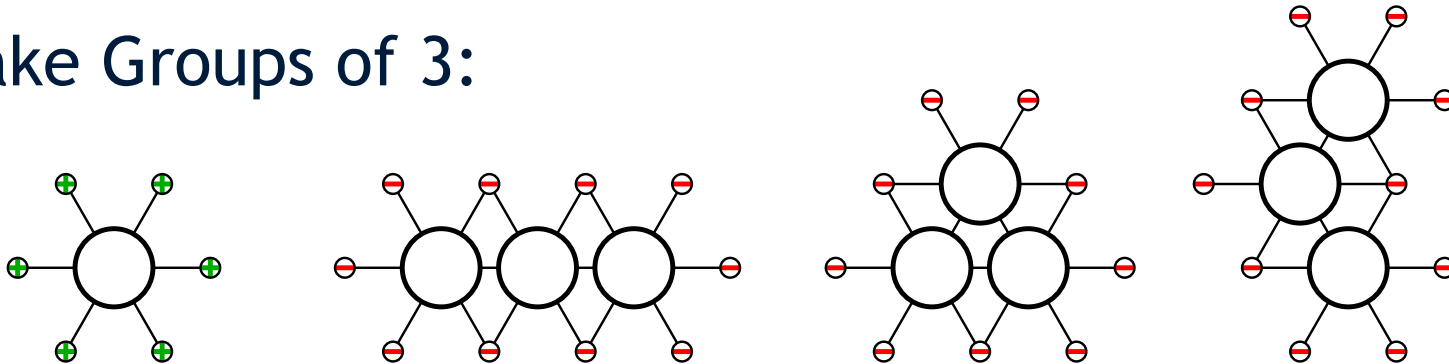
Make Lines of 4:



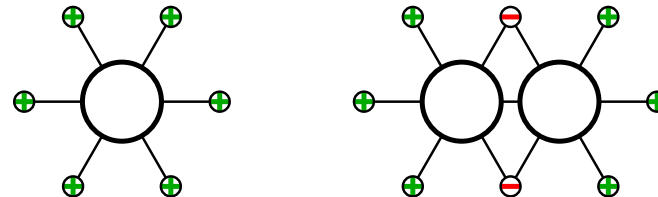
Avoid Lines of 3:



Make Groups of 3:



Make Long Thin Groups:



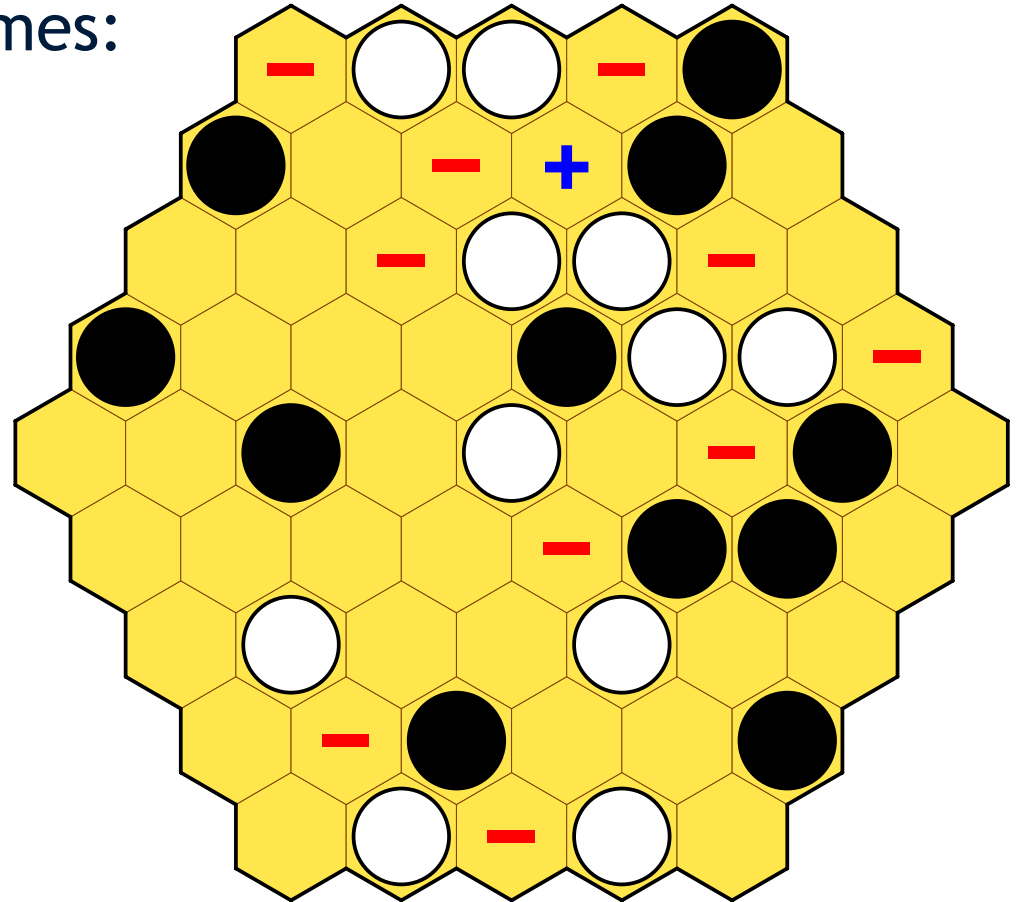
Hypothesis: *Features indicate strategic potential*

Monte Carlo Resistant Games

Monte Carlo Resistant Games:

- Random playouts give misleading results
- e.g. Yavalath:
 - Losing moves > winning moves
 - Random play 10 x more likely to make losing move

Human player would never make losing move!



Monte Carlo Resistant Games

MCTS assumption:

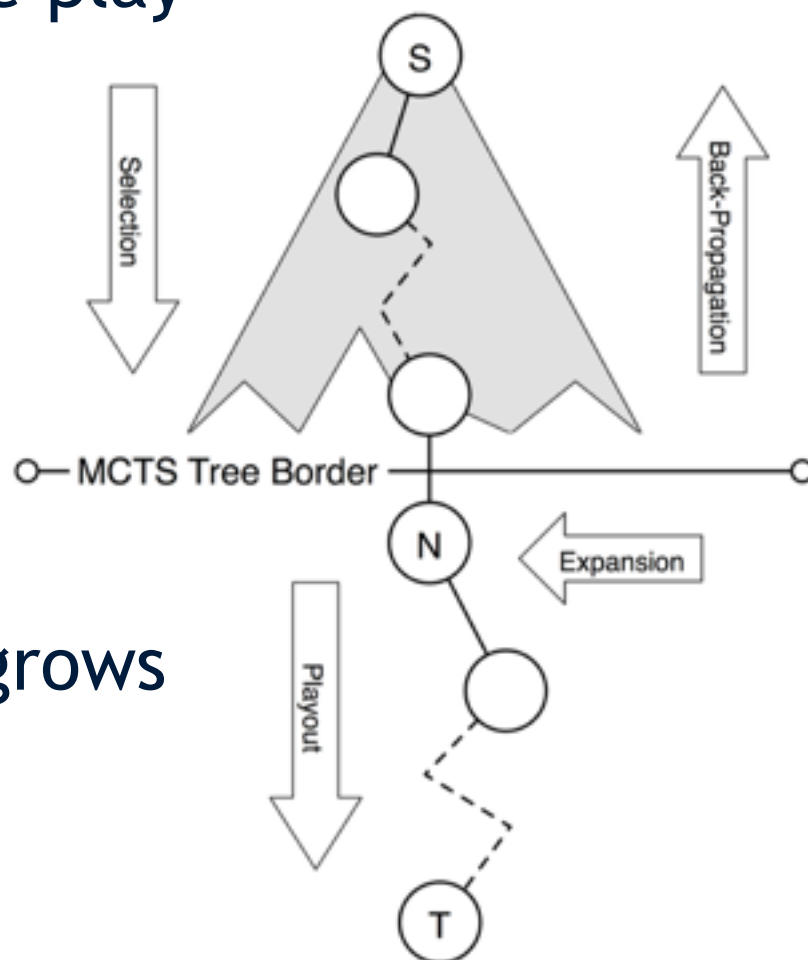
- MC simulations approximate play
- Not always true!

Misleading playouts give bad rewards below tree:

- Tree must grow to correct

Initially bad moves:

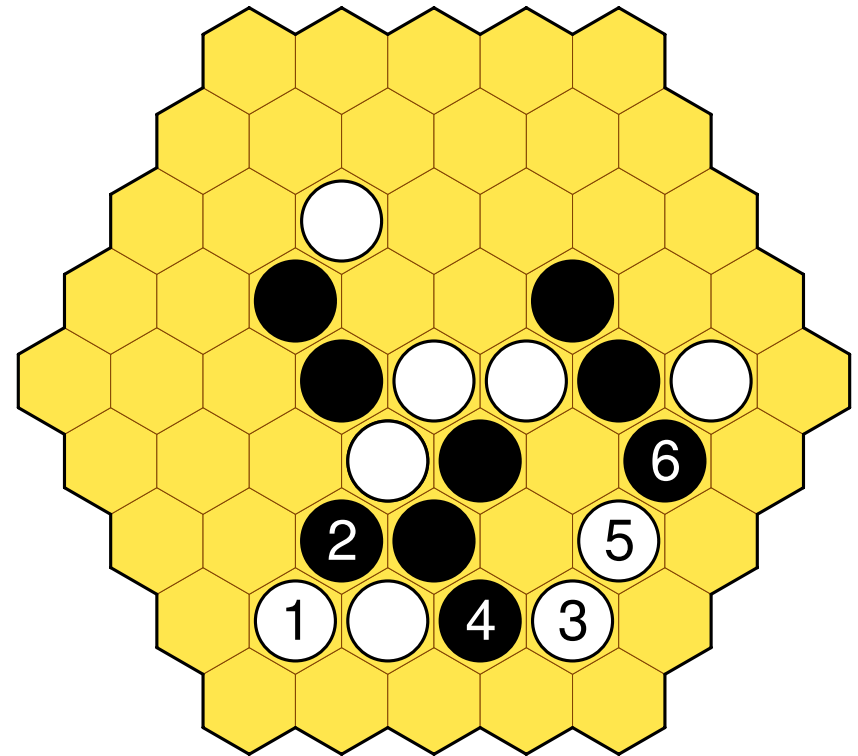
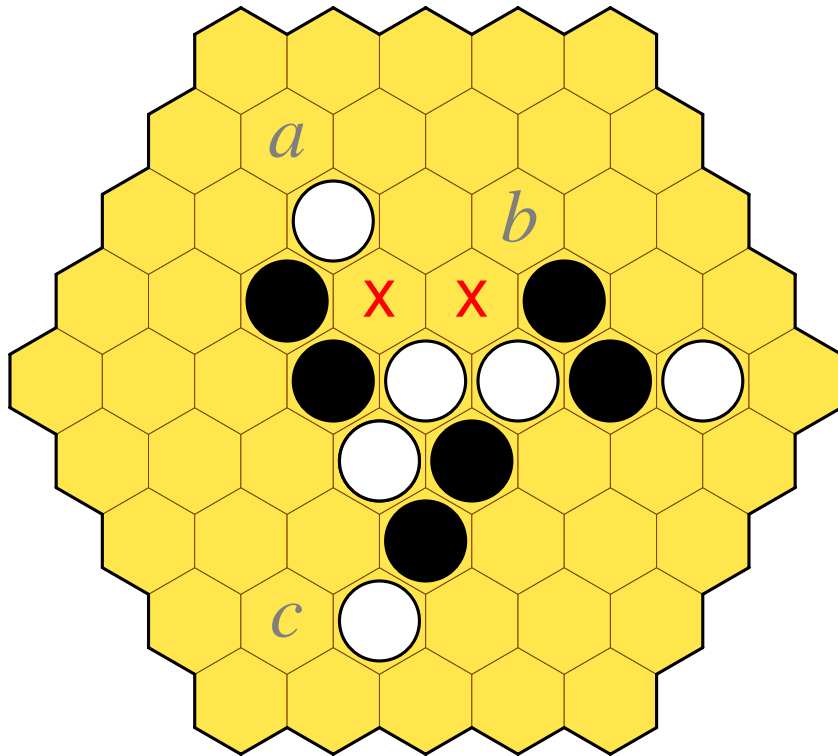
- Improve over time as tree grows



Application

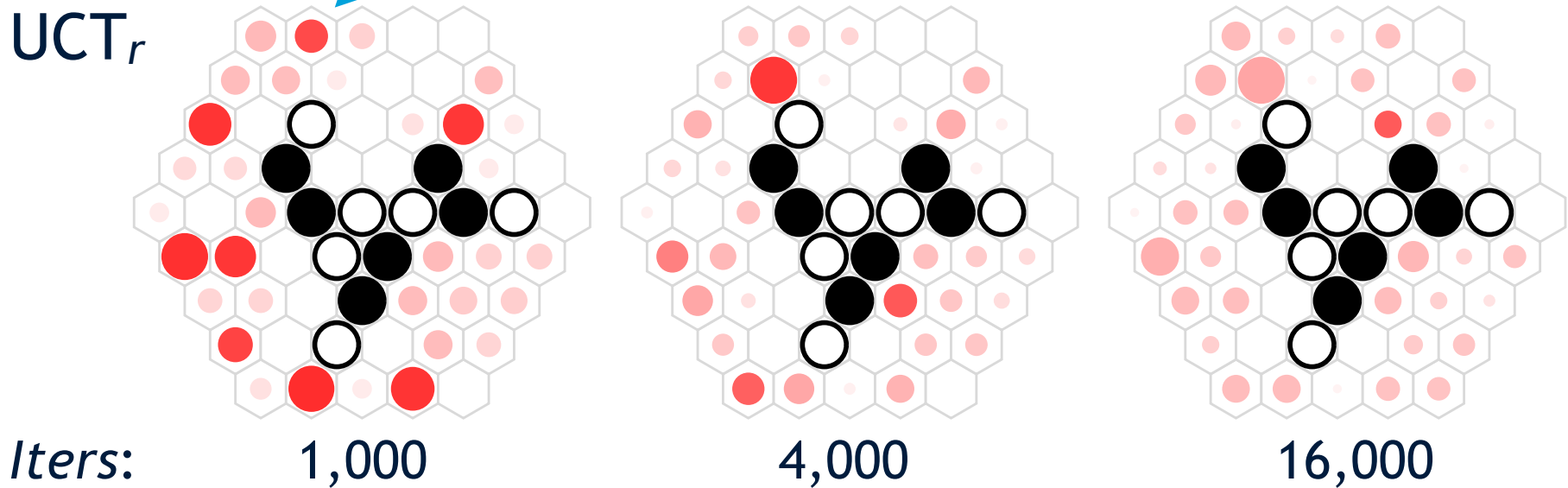
Recall Yavalath puzzle:

- How does MCTS perform?



Problem

UCT_r Boundary moves



How does MCTS with random playouts perform?

- Badly
- 50,000 - 100,000 iterations required to solve

Win/Loss Filters

Make Lines of 4:  

Win filter: Encourages winning moves

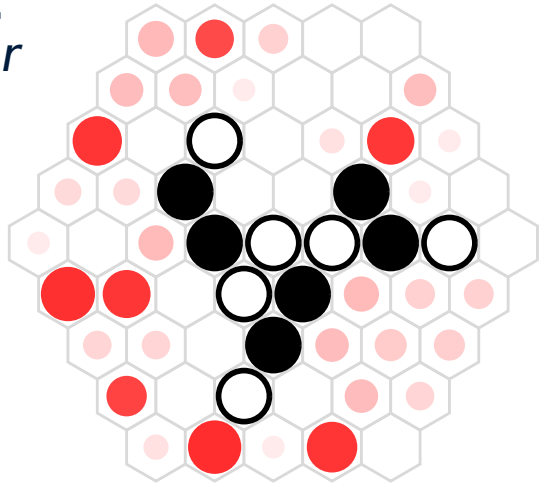
Avoid Lines of 3:  

Loss filter: Discourages losing moves

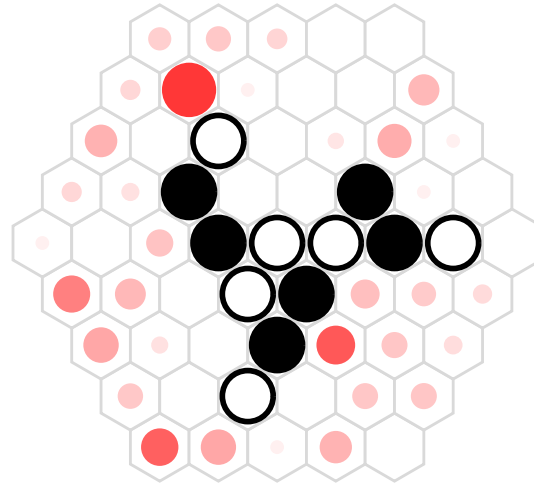
Similar to “decisive” and “anti-decisive” moves
used for Havannah

Solution

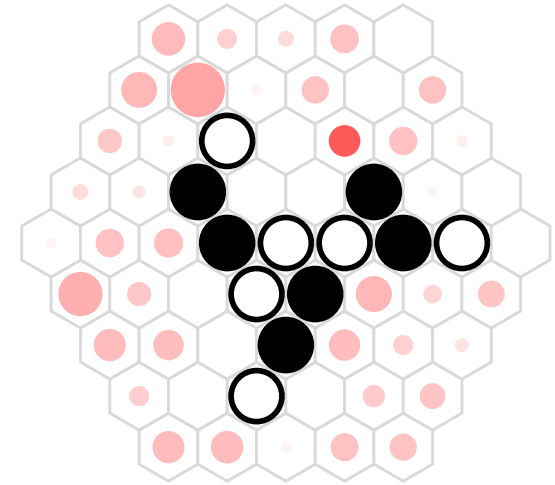
UCT_r



1,000



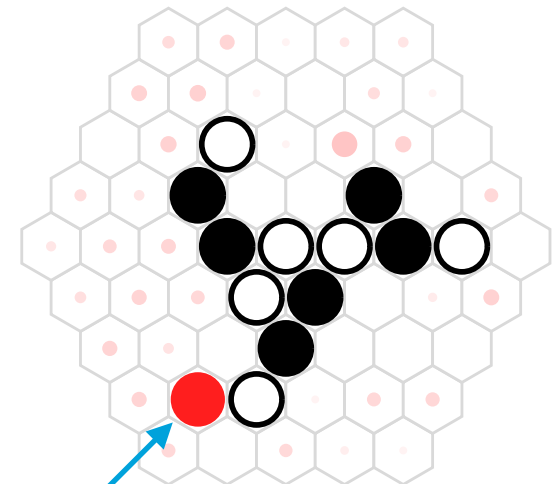
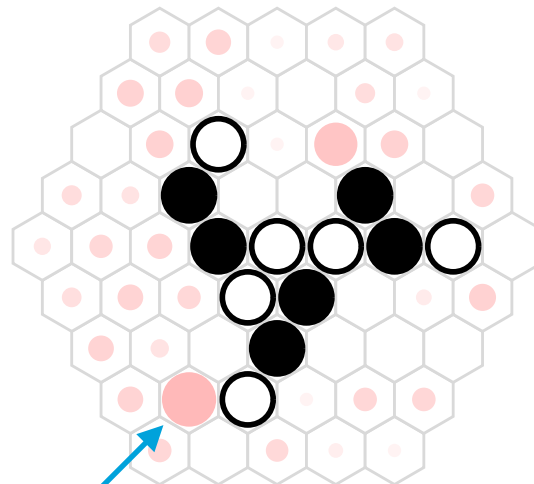
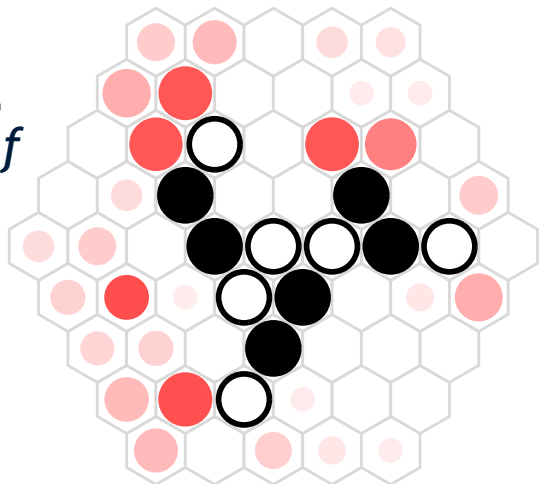
4,000



16,000

Iters:

UCT_f



Conclusion

Almost all games are composed from existing ideas

Evolution of new games is easy:

- Evaluating them automatically is hard!

Universal indicators of playability:

- Balance
- Completion
- Duration

Universal indicator of quality:

- Strategic potential?