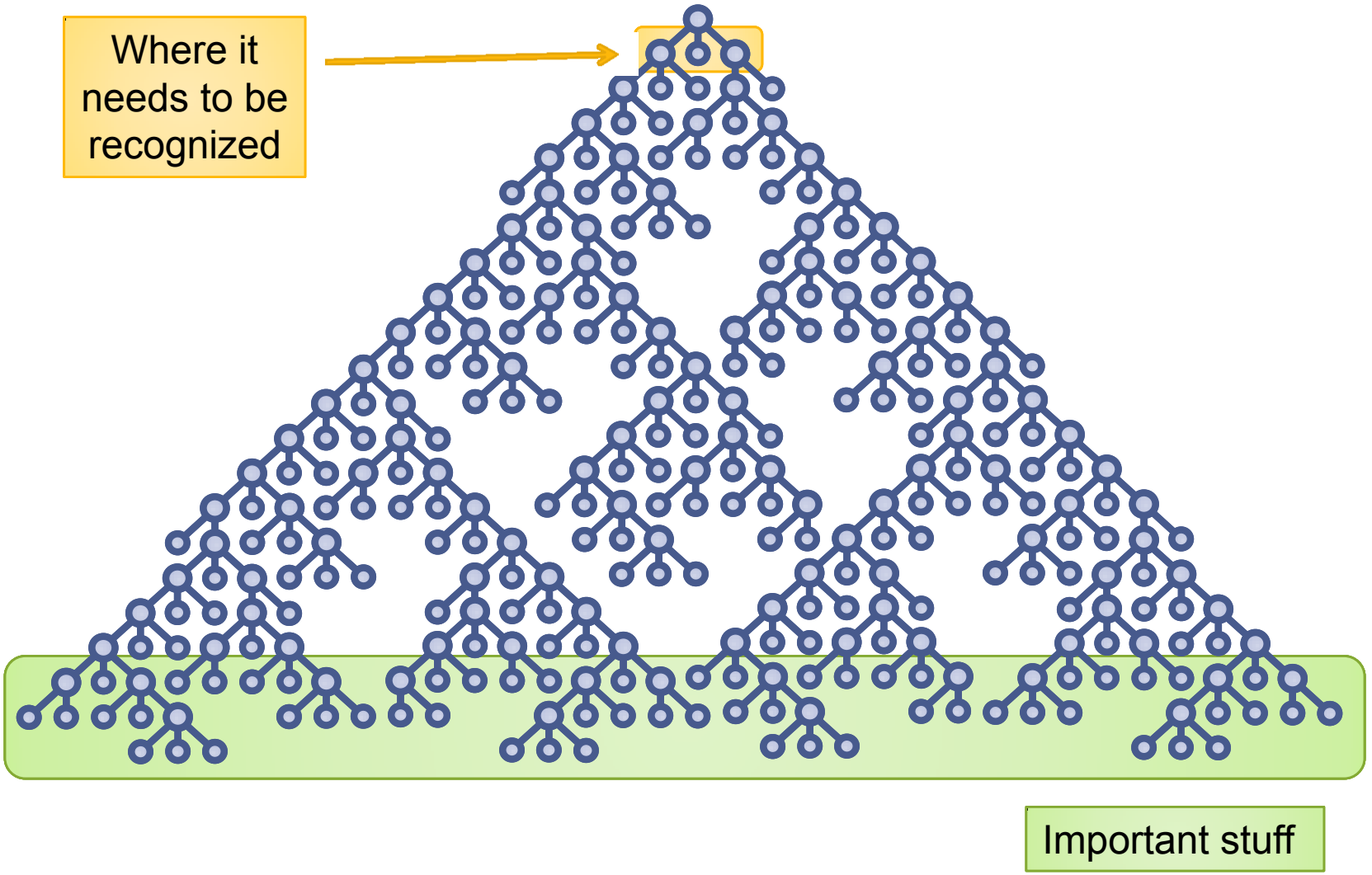
Monte-Carlo Search Algorithms

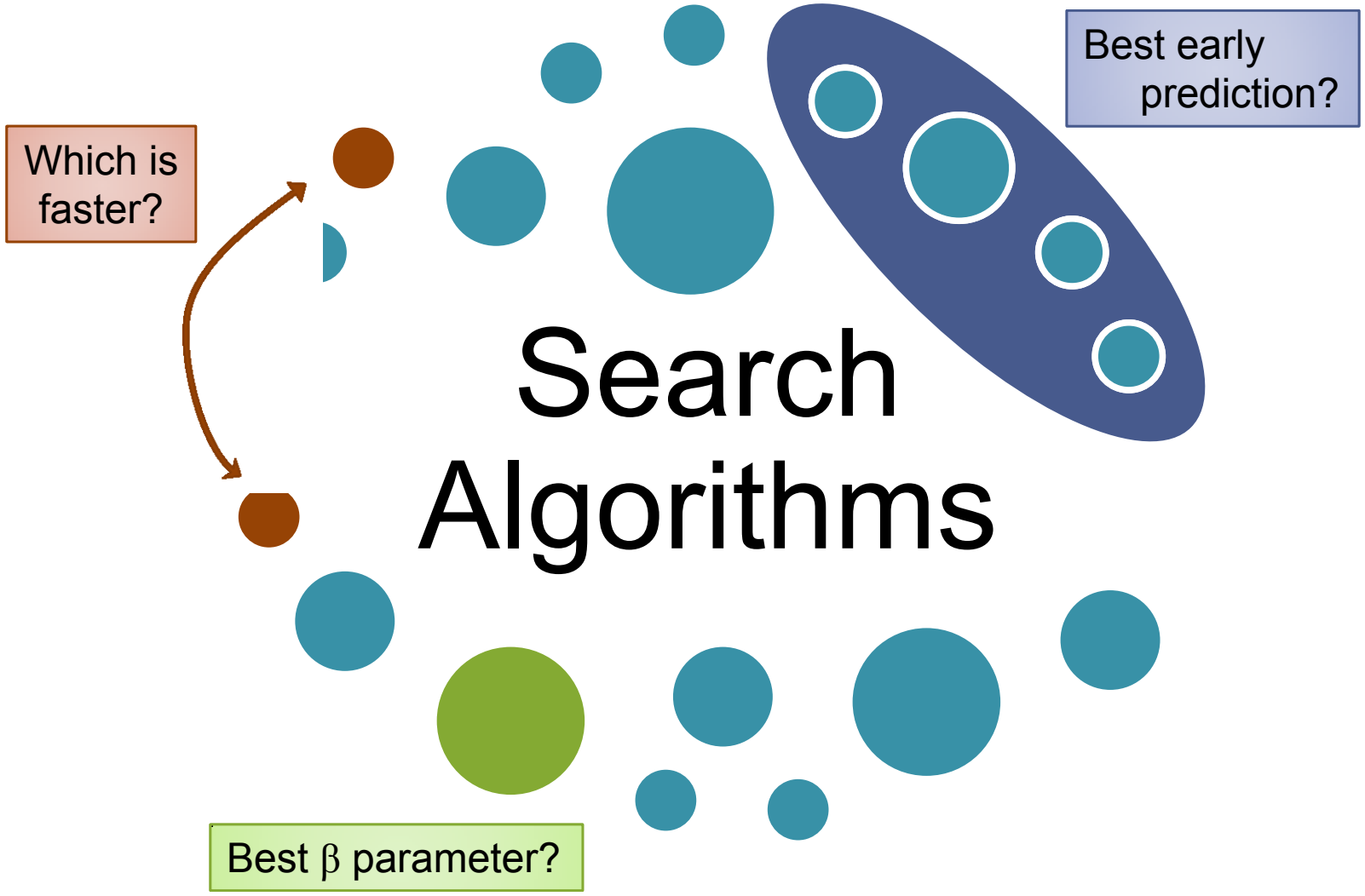
Daniel Bjorge and John Schaeffer

MTA SZTAKI
COMPUTER AND AUTOMATION
RESEARCH INSTITUTE
HUNGARIAN ACADEMY OF SCIENCES

Problem

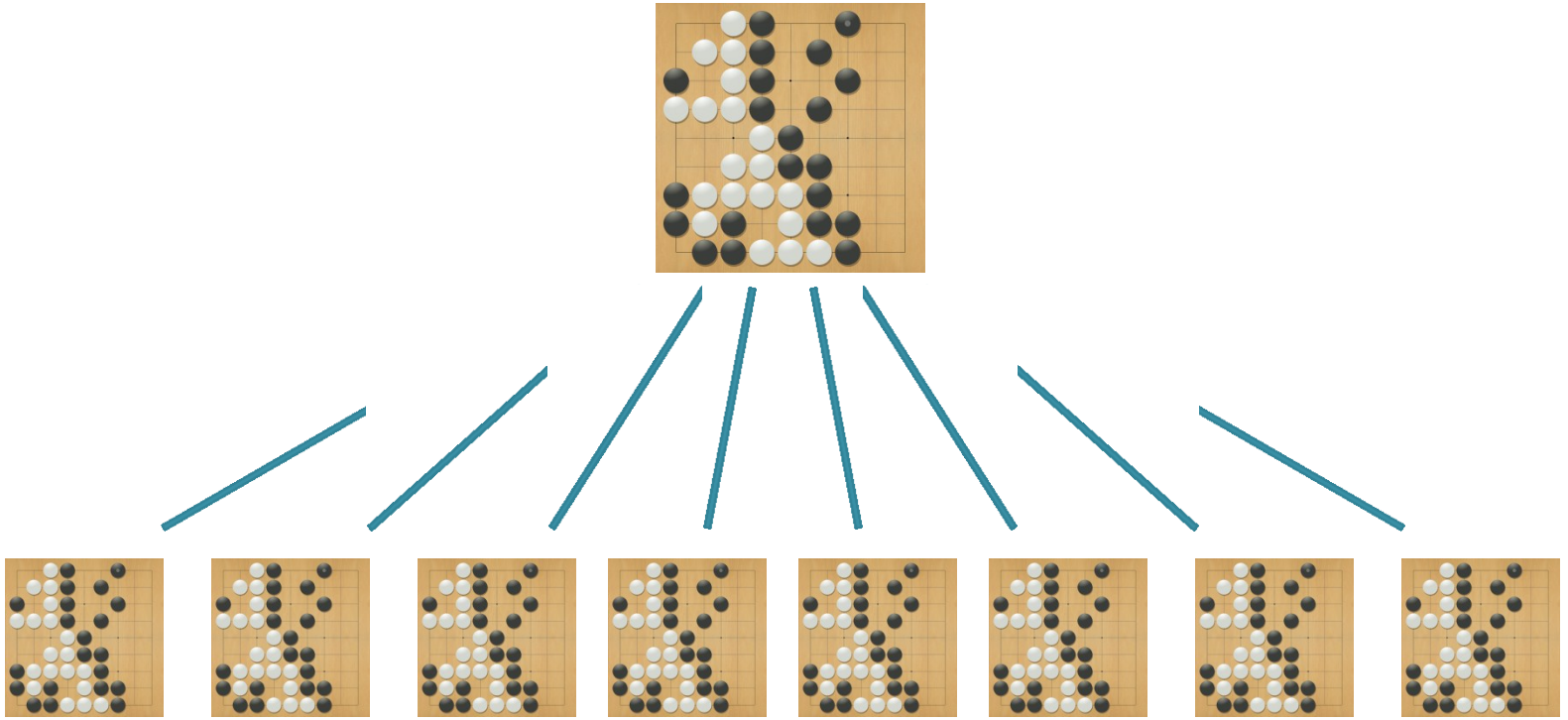


Solution


Best early prediction?
Which is faster?
Search Algorithms
Best β parameter?

Comparing Search Algorithms

Option 1: Computer Go



Slow

State expansion

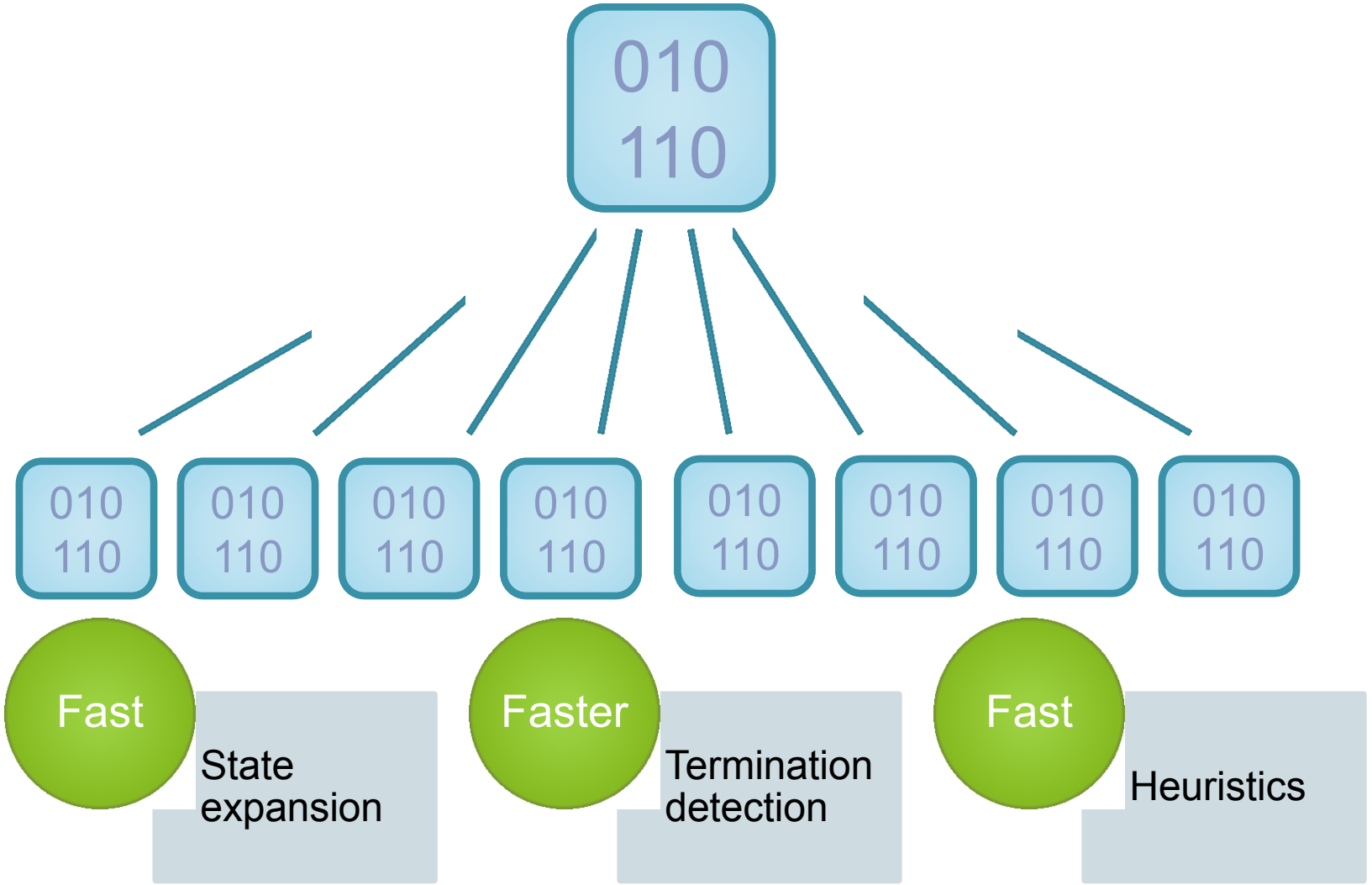
Slower

Termination detection

Lassú

Heuristics

Option 2: Artificial Trees



Existing Frameworks

P-Game

Kocsis

2006

- Artificial
- Fast
- Simple

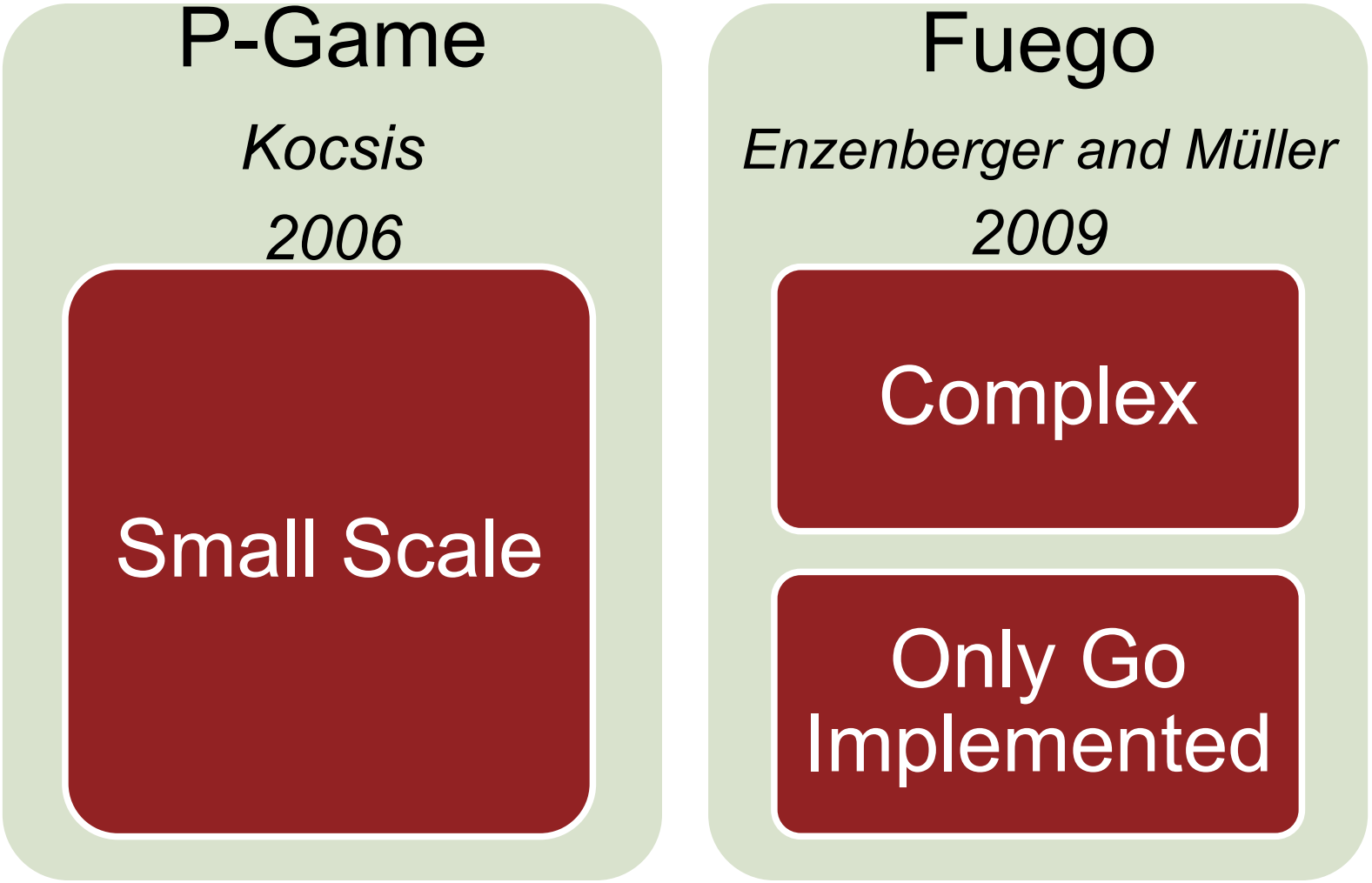
Fuego

Enzenberger and Müller

2009

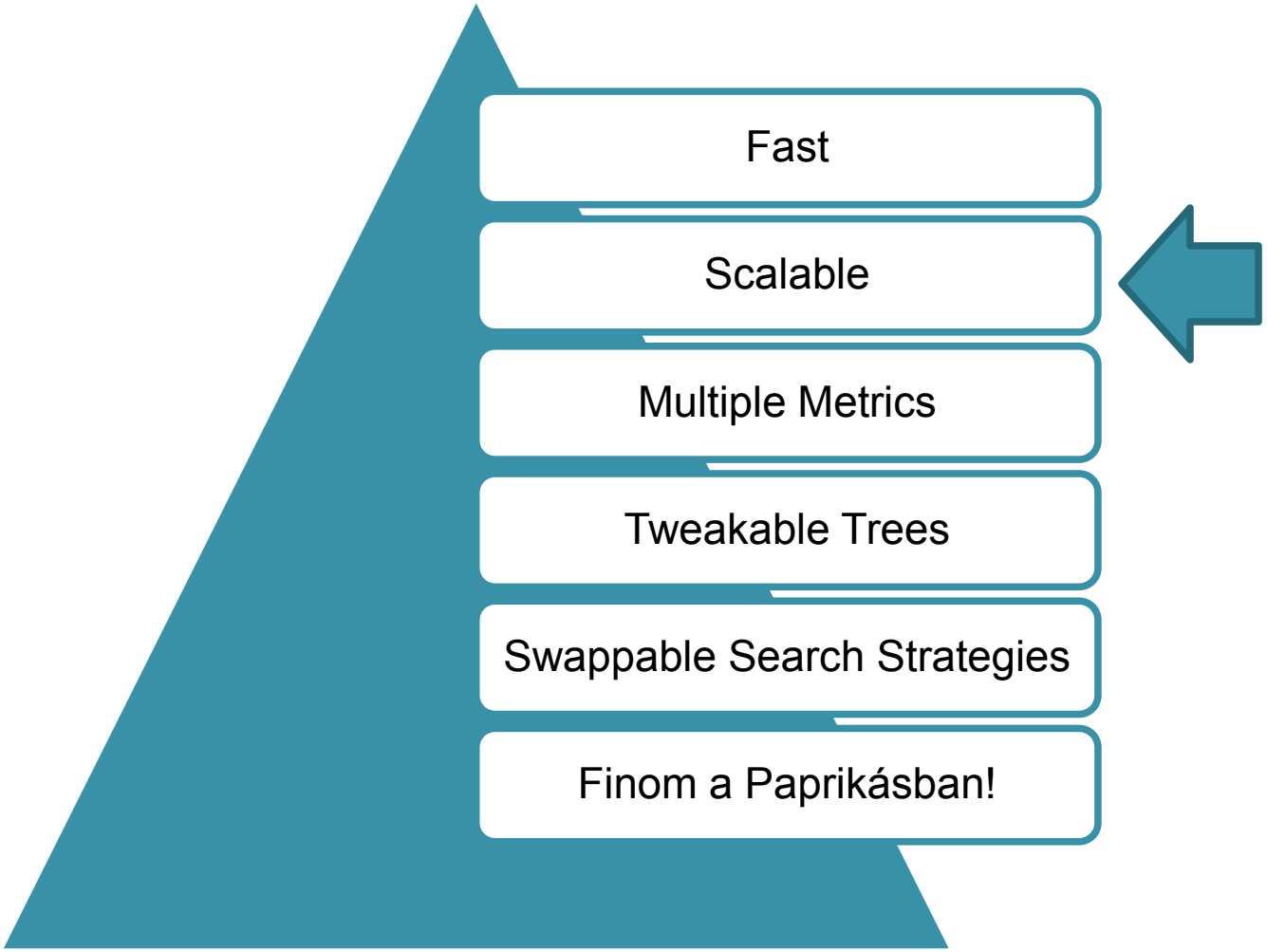
- Generic
- Fast
- Scalable

Existing Frameworks



Gomba Search Framework

(Go Multi-Armed Bandit Analysis)

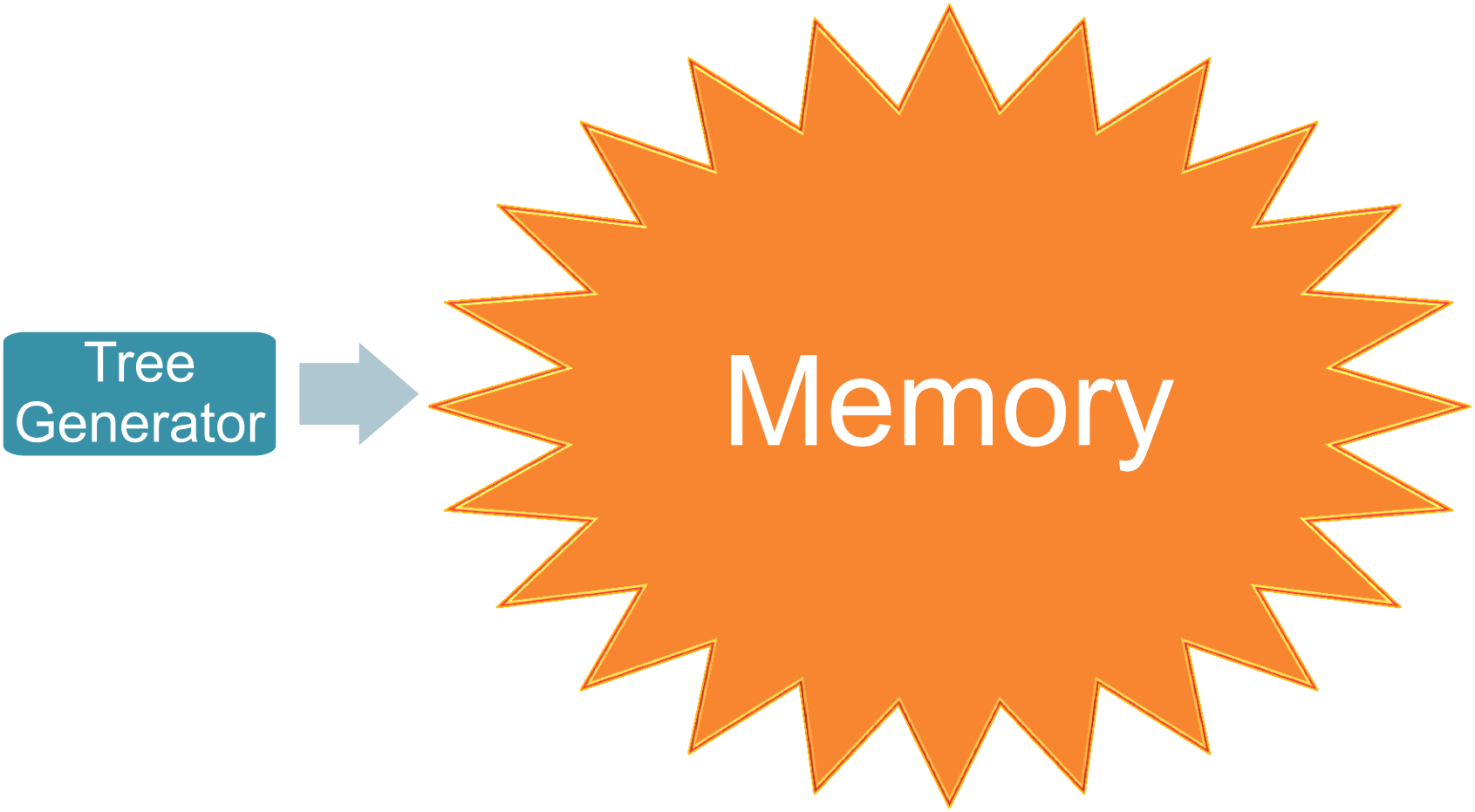


Eager Game Tree Generation

(Small trees)

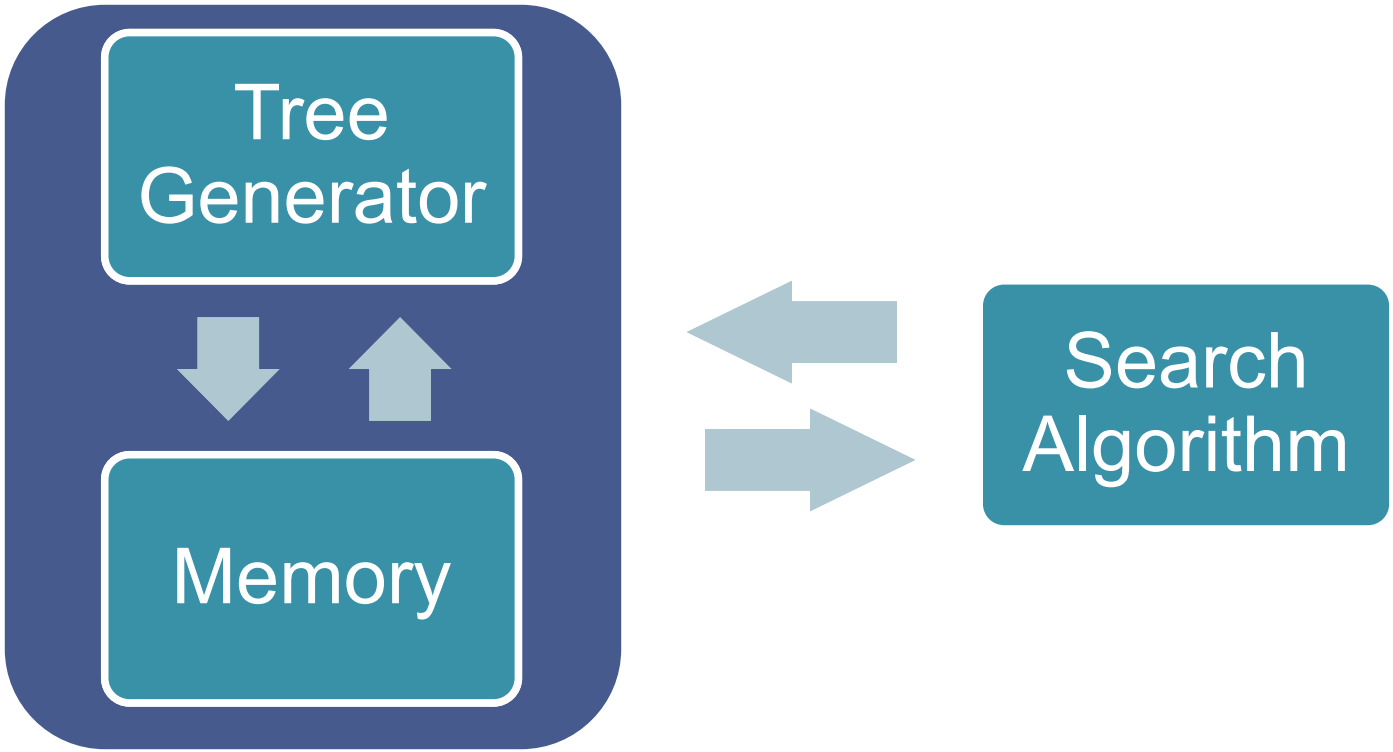
Eager Game Tree Generation

(Big trees)

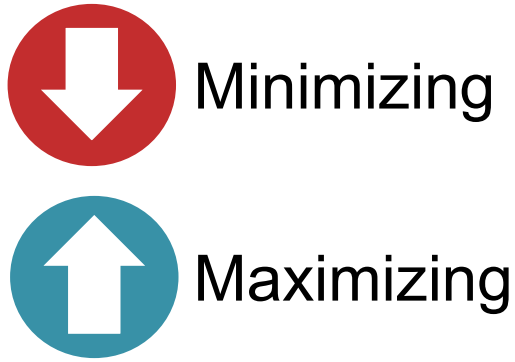


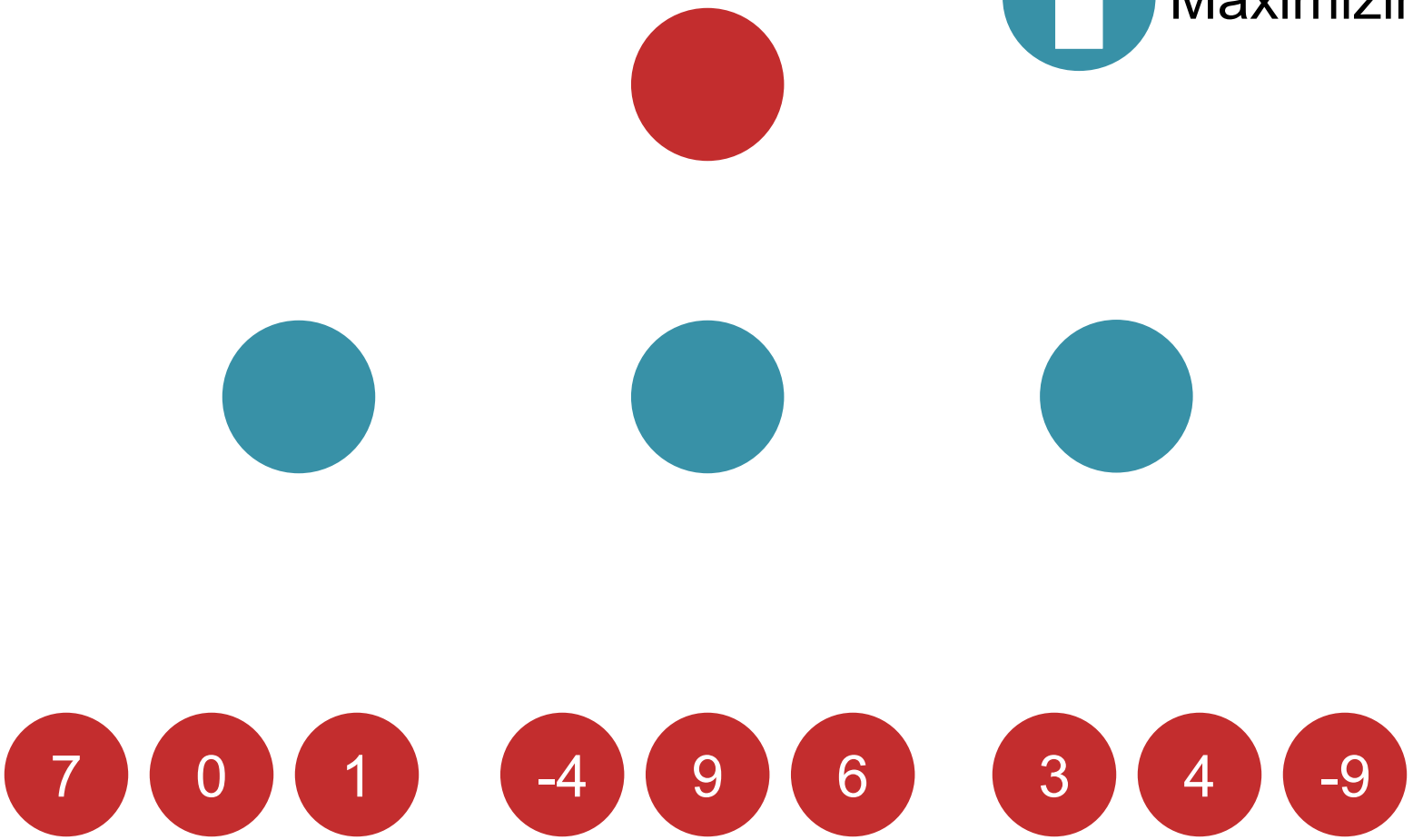
Lazy Game Tree Generation

(Any trees)

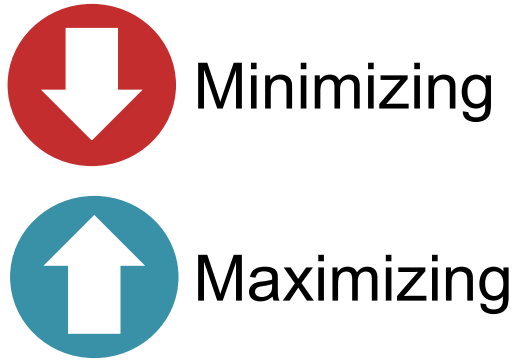


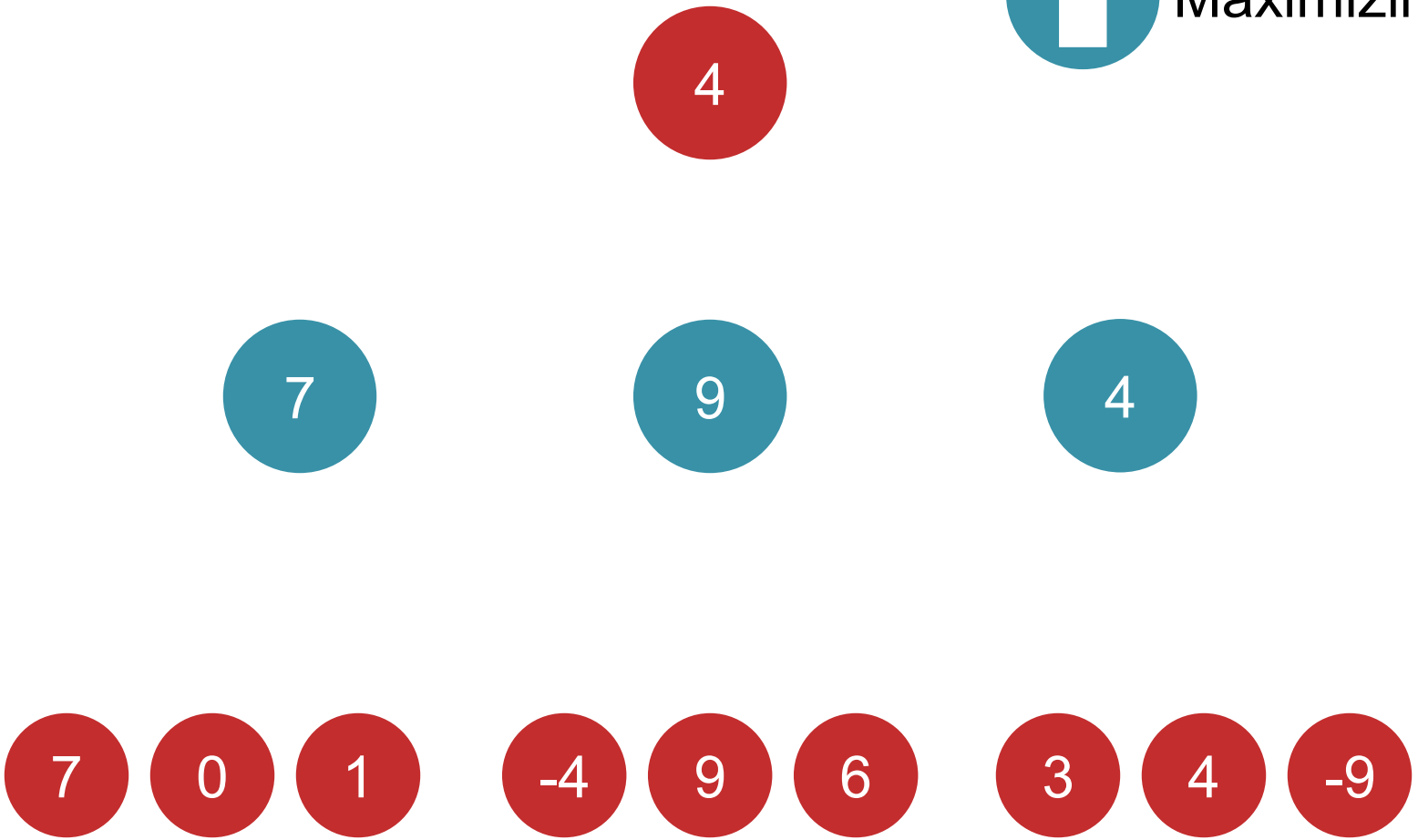
Minimax Values





Minimax Values



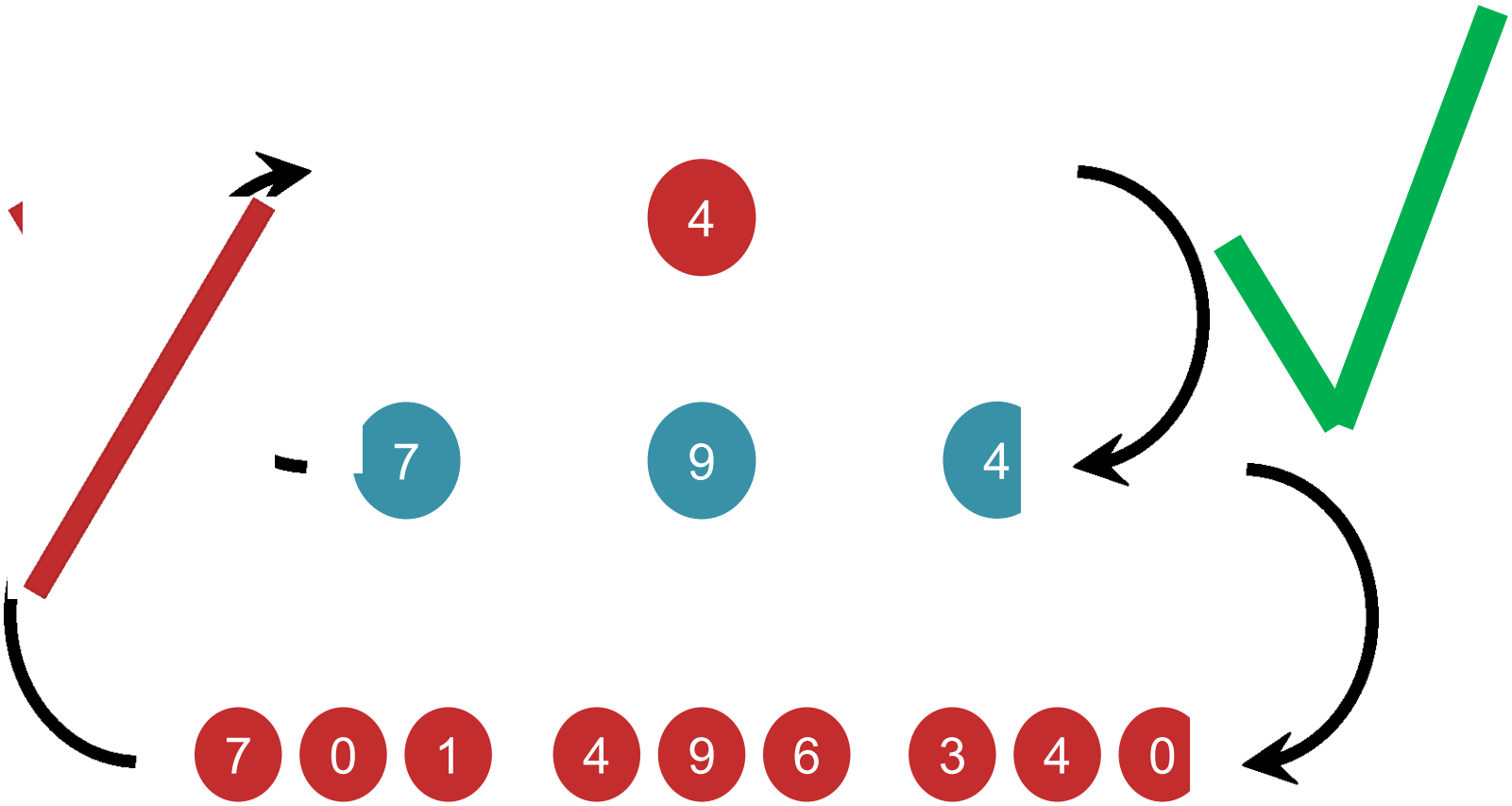


Minimax Values

$$O\left(b^{d/2}\right)$$

Go: $b \approx 200, d \approx 100$

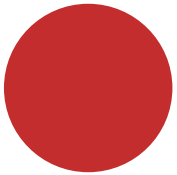
Downward Minimax Evaluation



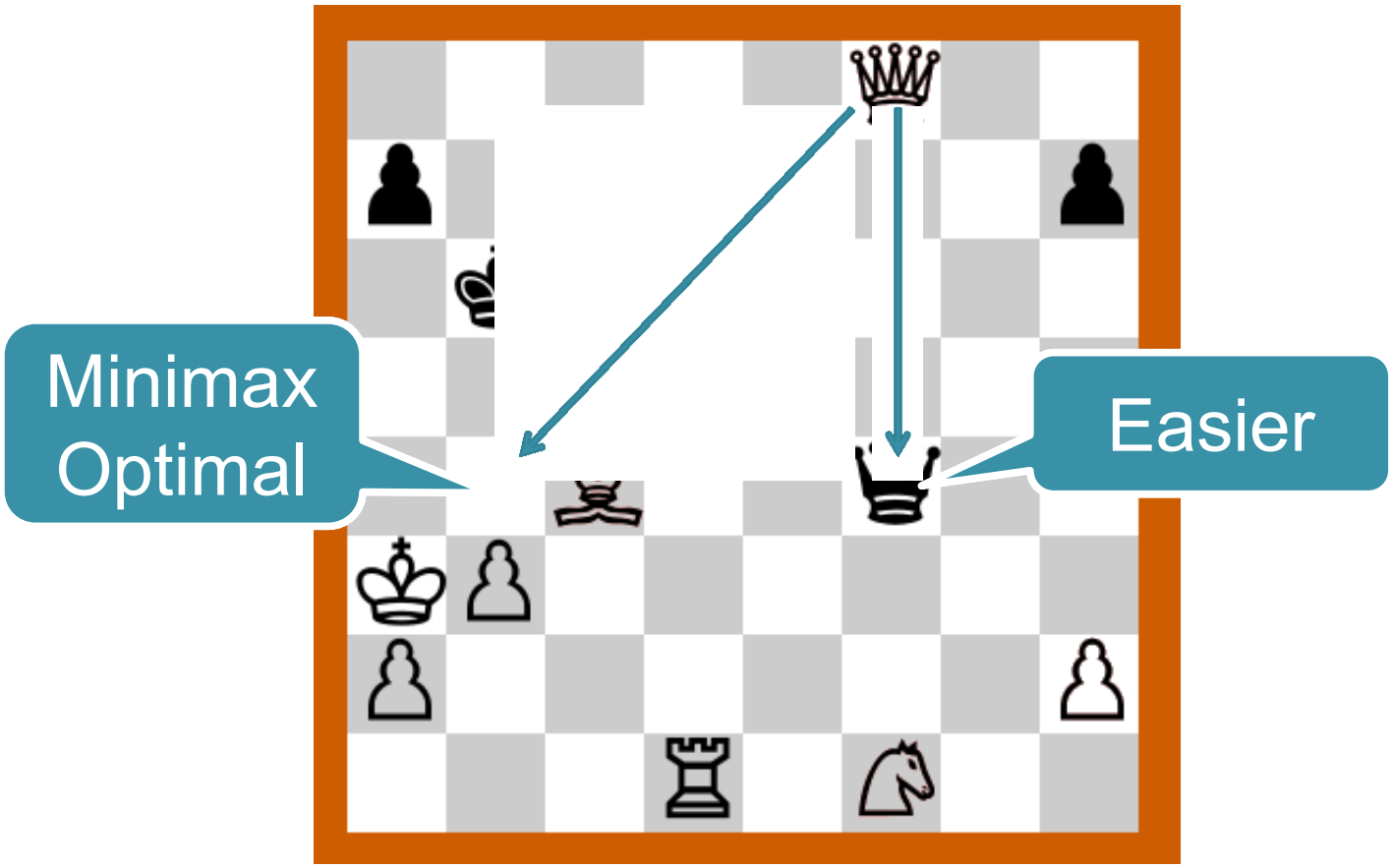
Example

Minimizing

Maximizing

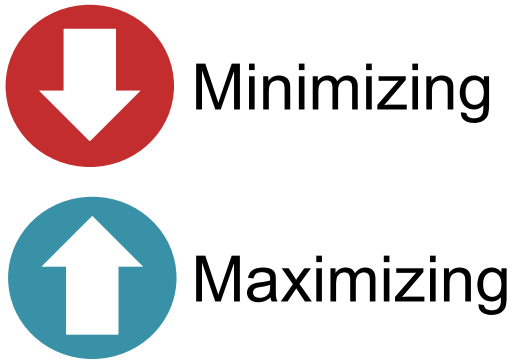


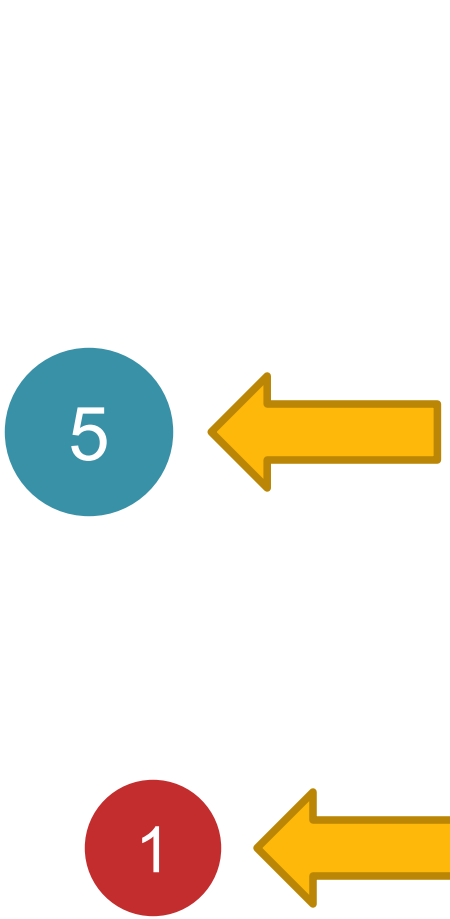
Difficulty



White to play

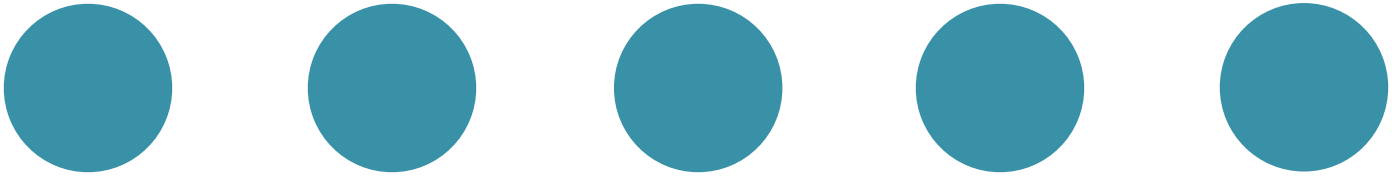
Quantified Difficulty





Application to Search Algorithms

Many-Armed Bandit Solutions

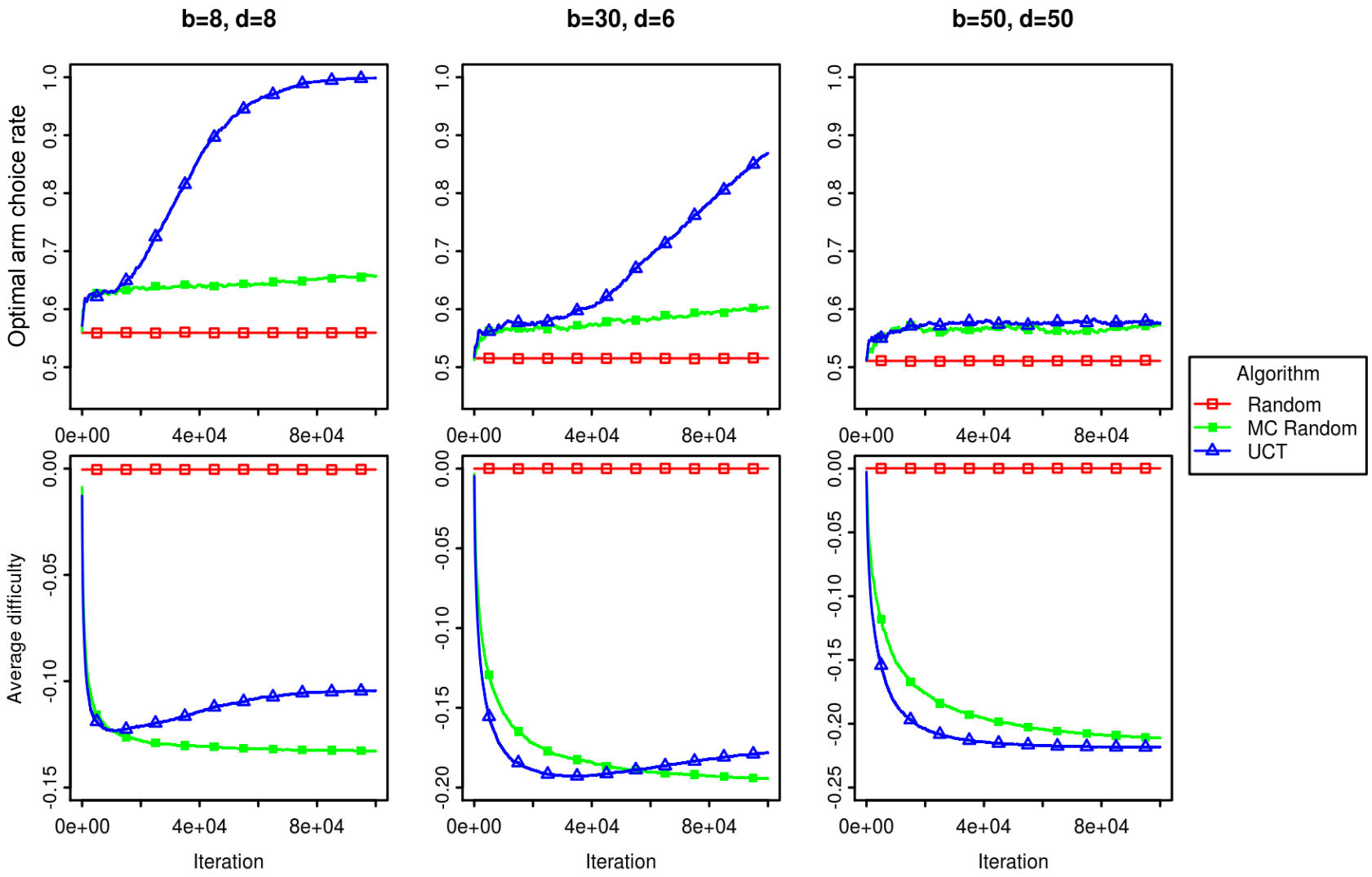


10	0	14	1	-12
8	1	19	2	-4
3	0	-27	2	-2
7	0	-13	2	-8
...	...	...	...	...

UCT

(Kocsis and Szepesvári, 2006)

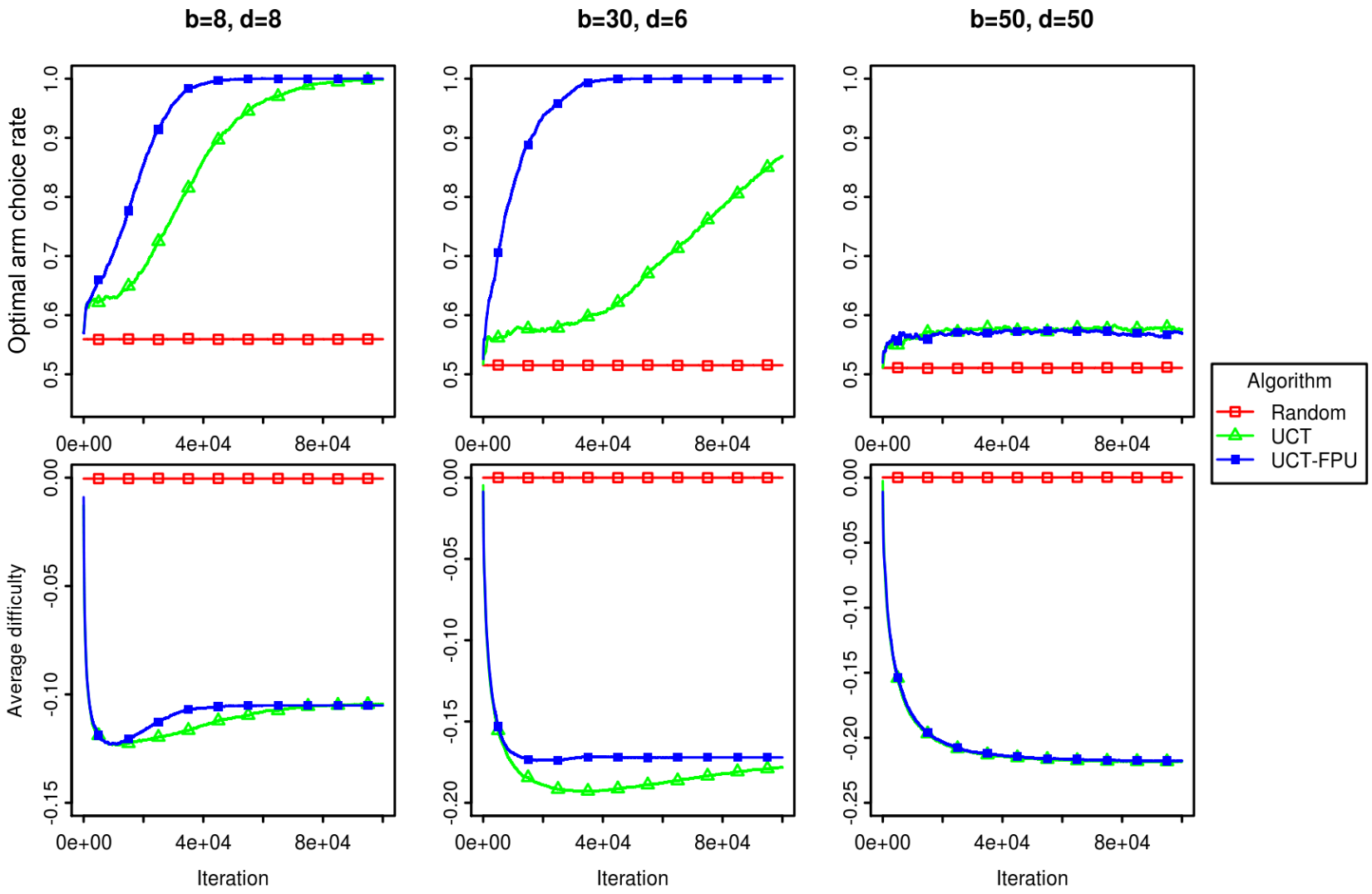
Upper Confidence Bound for Trees



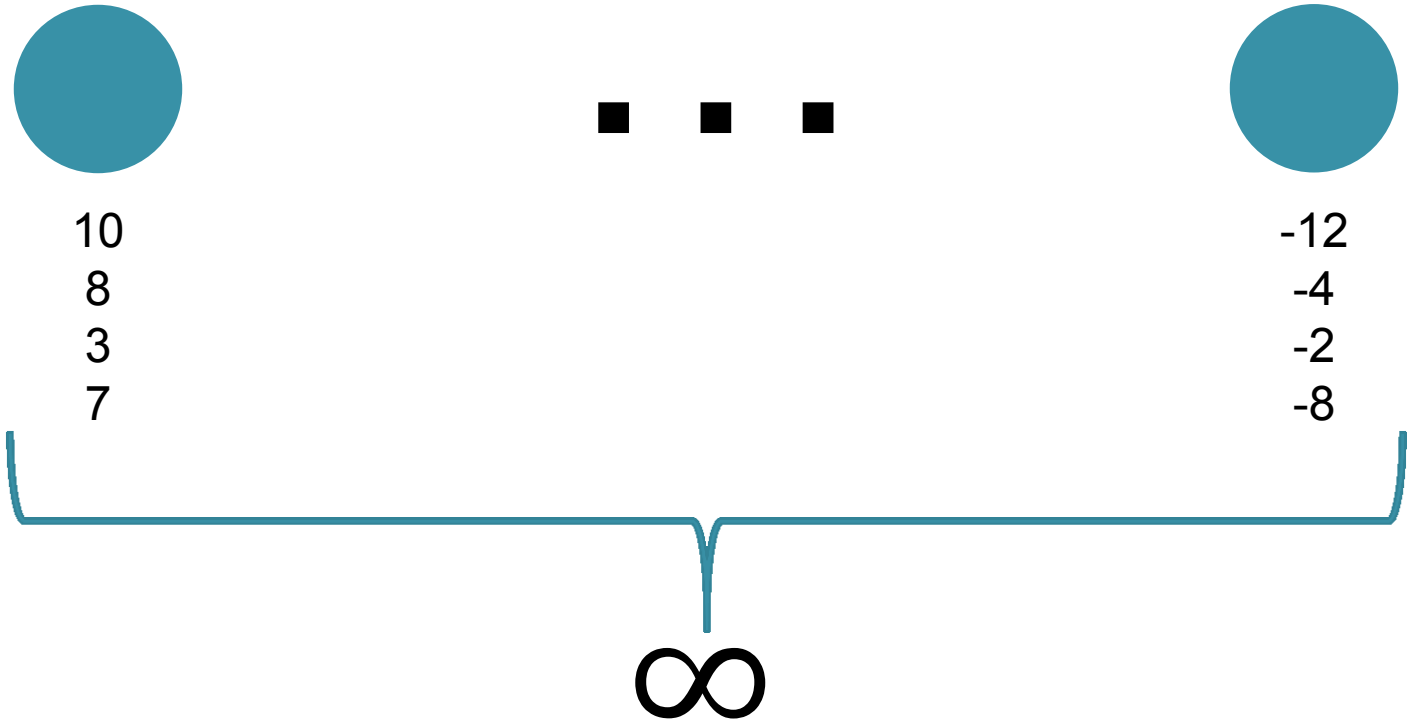
UCT-FPU

UCT with First Play Urgency

(Hartland et al., 2006)

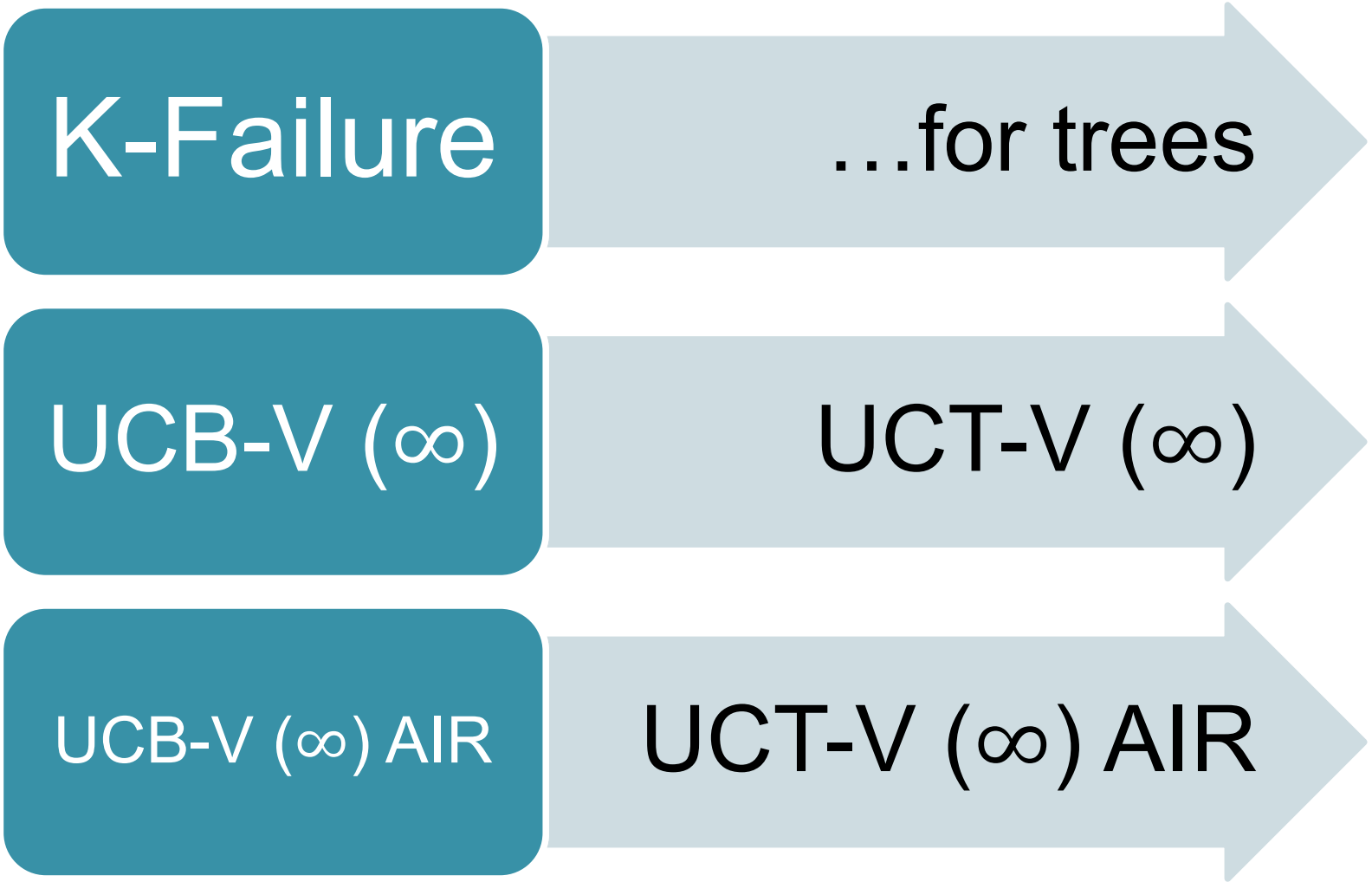


Infinitely-Armed Bandit Solutions



Infinite-Armed Bandit Solutions

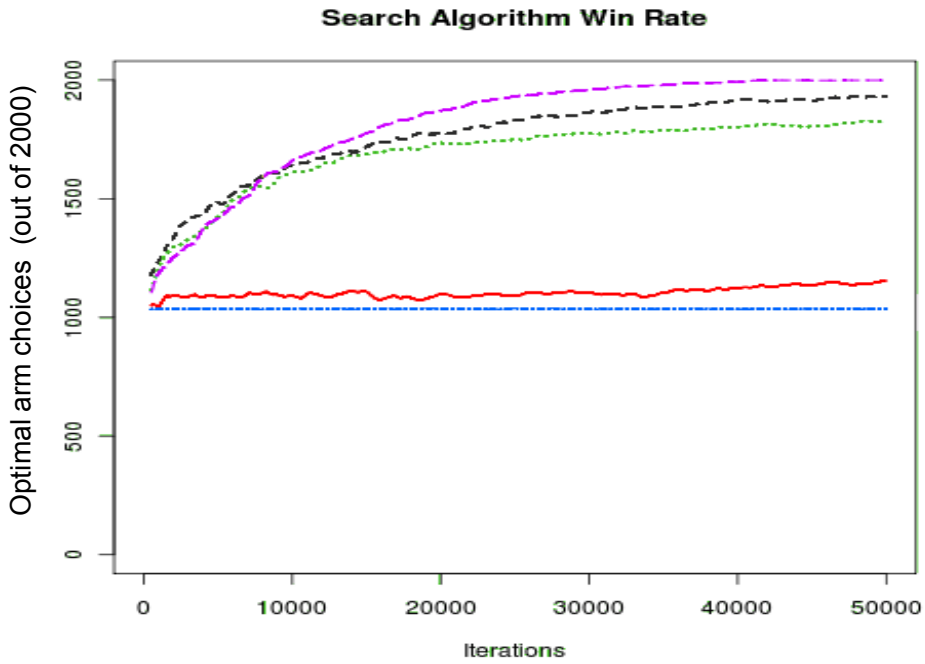
Adapting to Trees

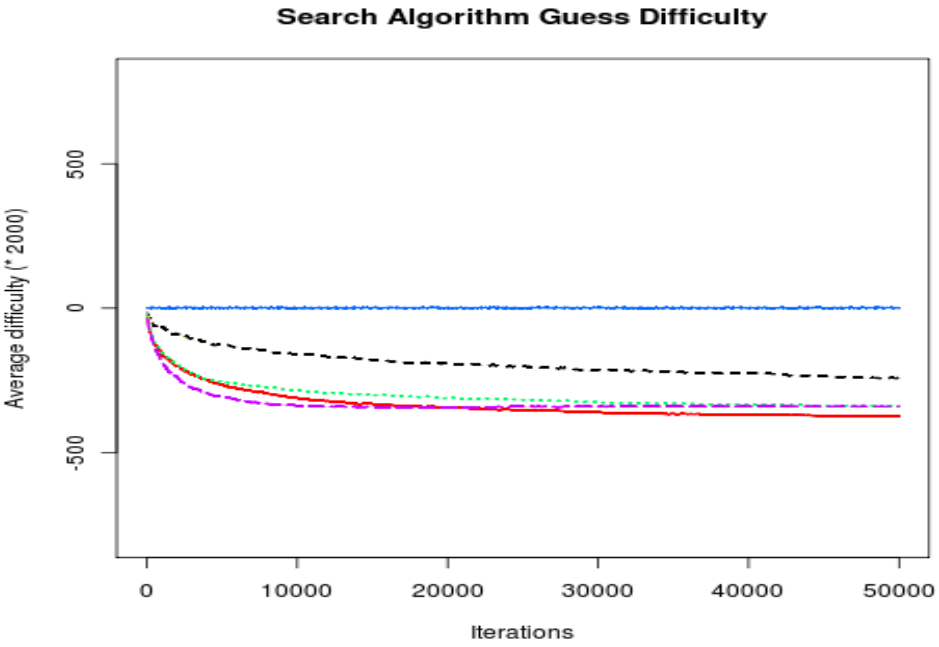


(Wang, Audibert, and Munos, 2006)

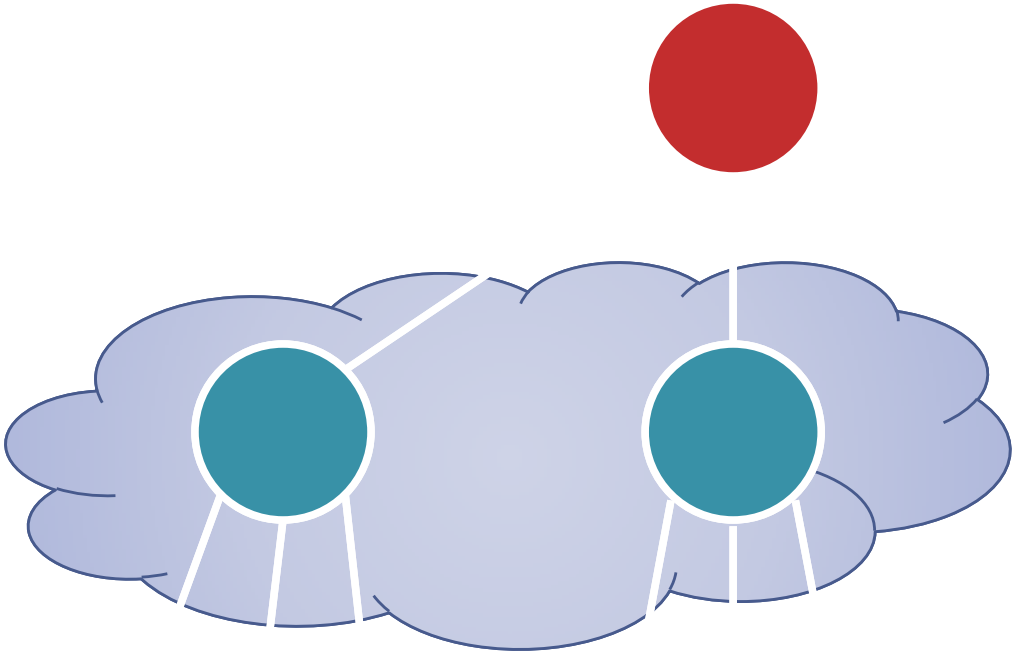
K-Failure

- Random
- UCT-FPU
- 1-Failure
- 3-Failure
- 10-Failure



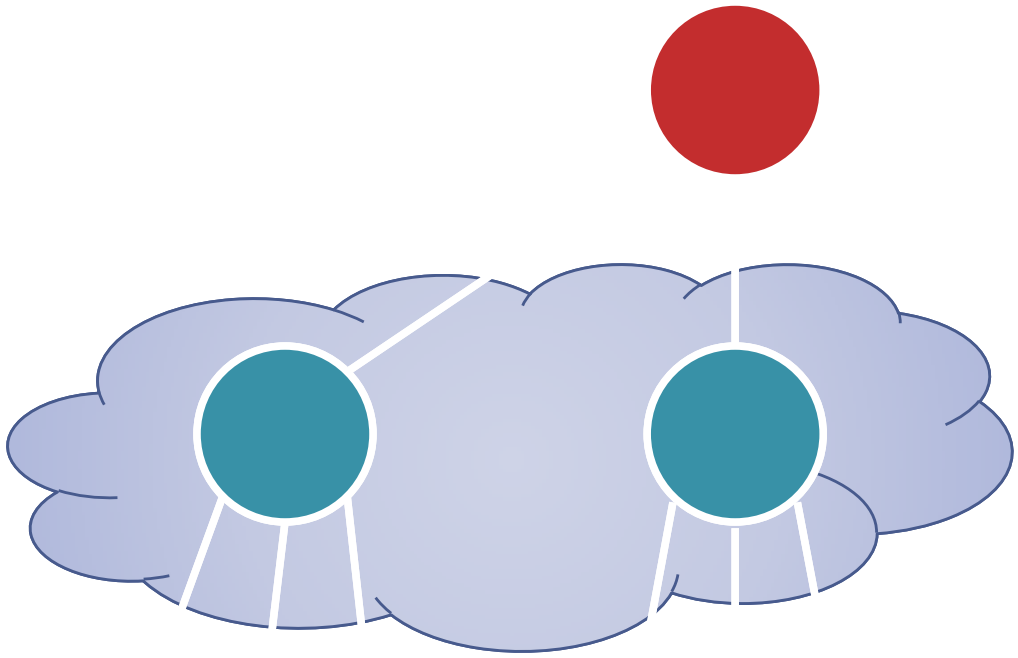


UCT-V (∞)



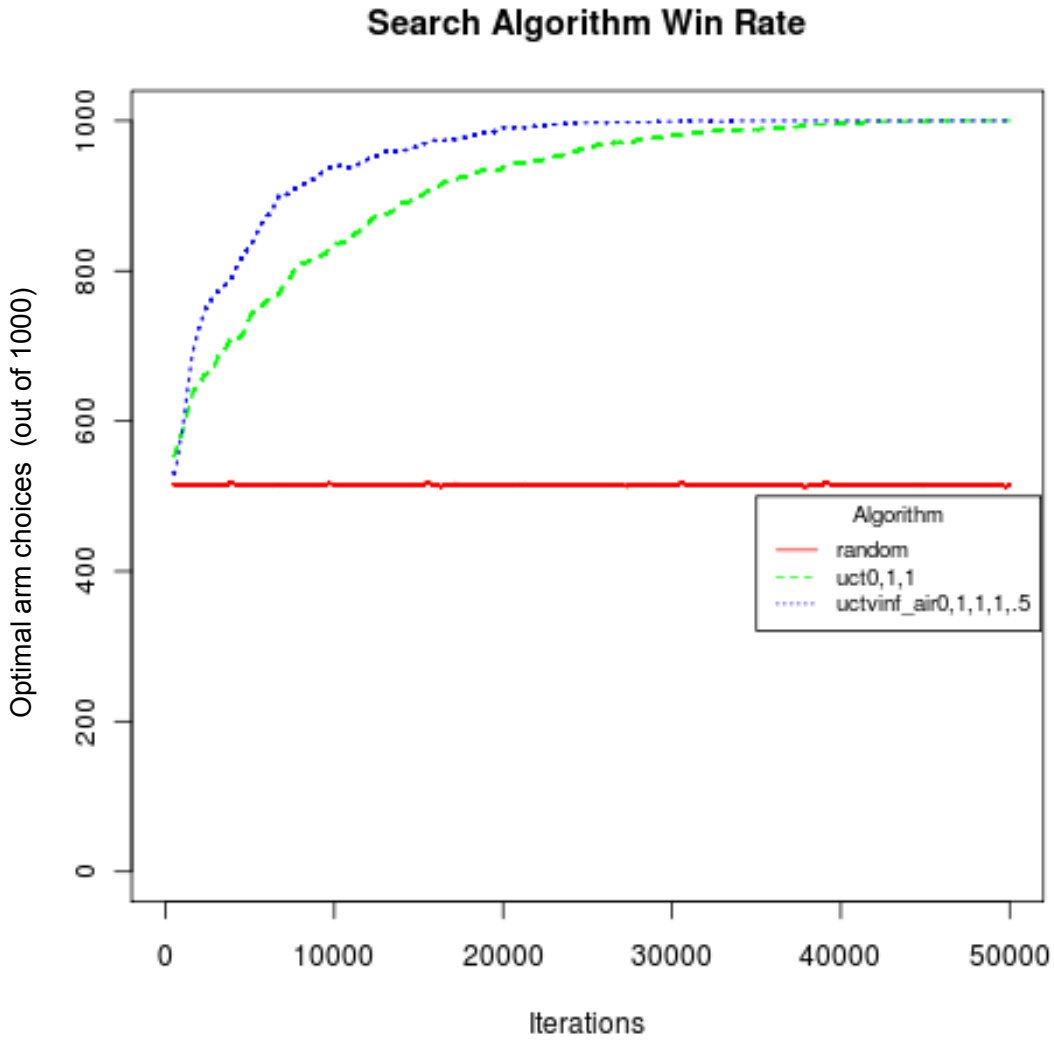
… … … … … …

UCT-V (∞) with AIR



… … … … … … … … …

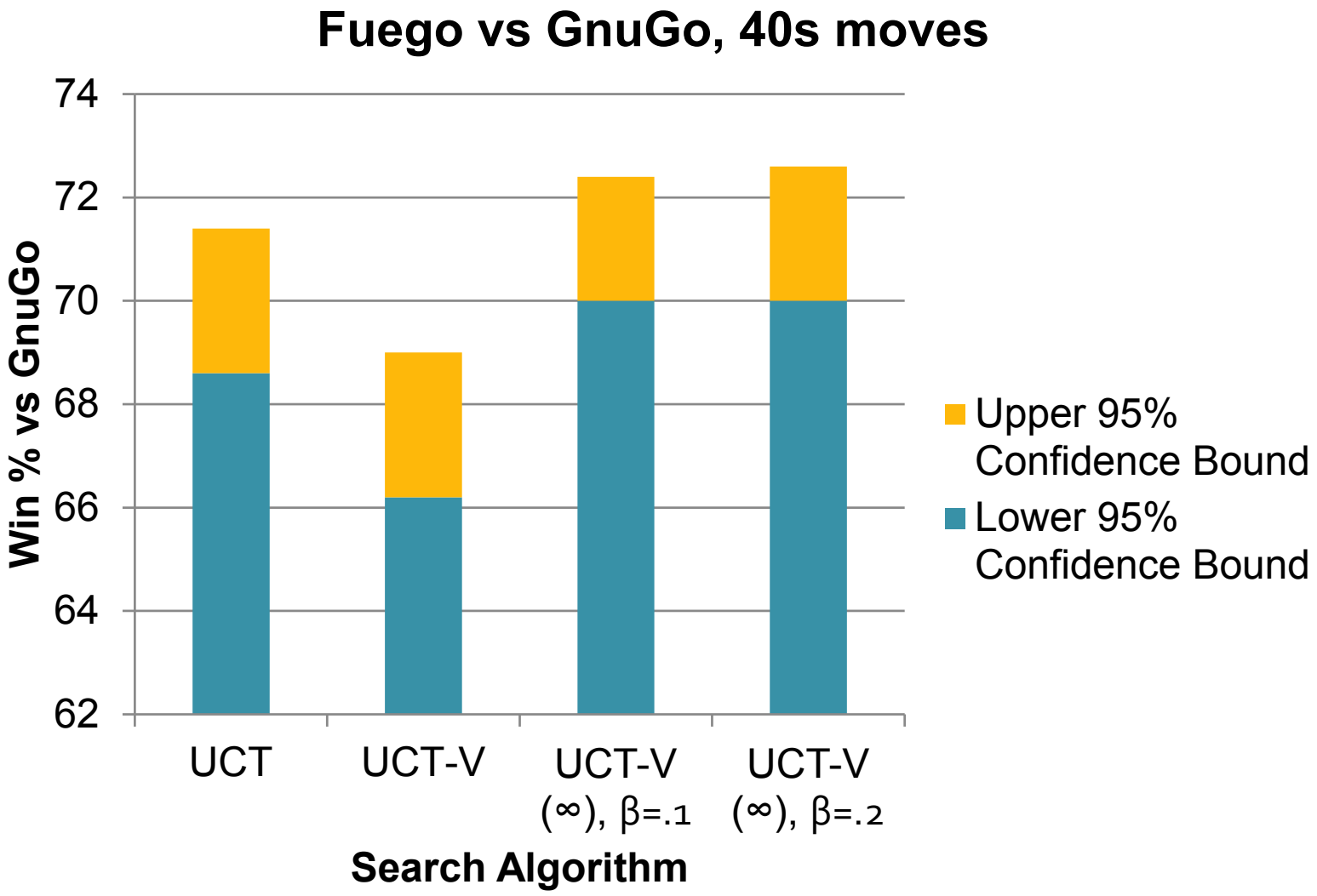
UCT-V (∞) with AIR: Gomba



$b = 30$

$d = 6$

In Fuego


Fuego vs GnuGo, 40s moves
Win % vs GnuGo
74
72
70
68
66
64
62
UCT
UCT-V
UCT-V (∞), β=.1
UCT-V (∞), β=.2
Search Algorithm
Upper 95% Confidence Bound
Lower 95% Confidence Bound

Wrapping Up

Gomba

- Fast
- Scalable
- Reasonably accurate
- Simple
- Interchangeable

Infinite Bandit-Based Solutions

- Improvement

Questions?

Köszönjük szépen!

Advisors

- Kocsis Levente
- Sárközy Gábor
- Stanley Selkow

SZTAKI Info Lab

- Benczúr András
- Recski Gábor
- Zséder Attila
- Kornai András
- Erdélyi Miklós

Fuego Development Team