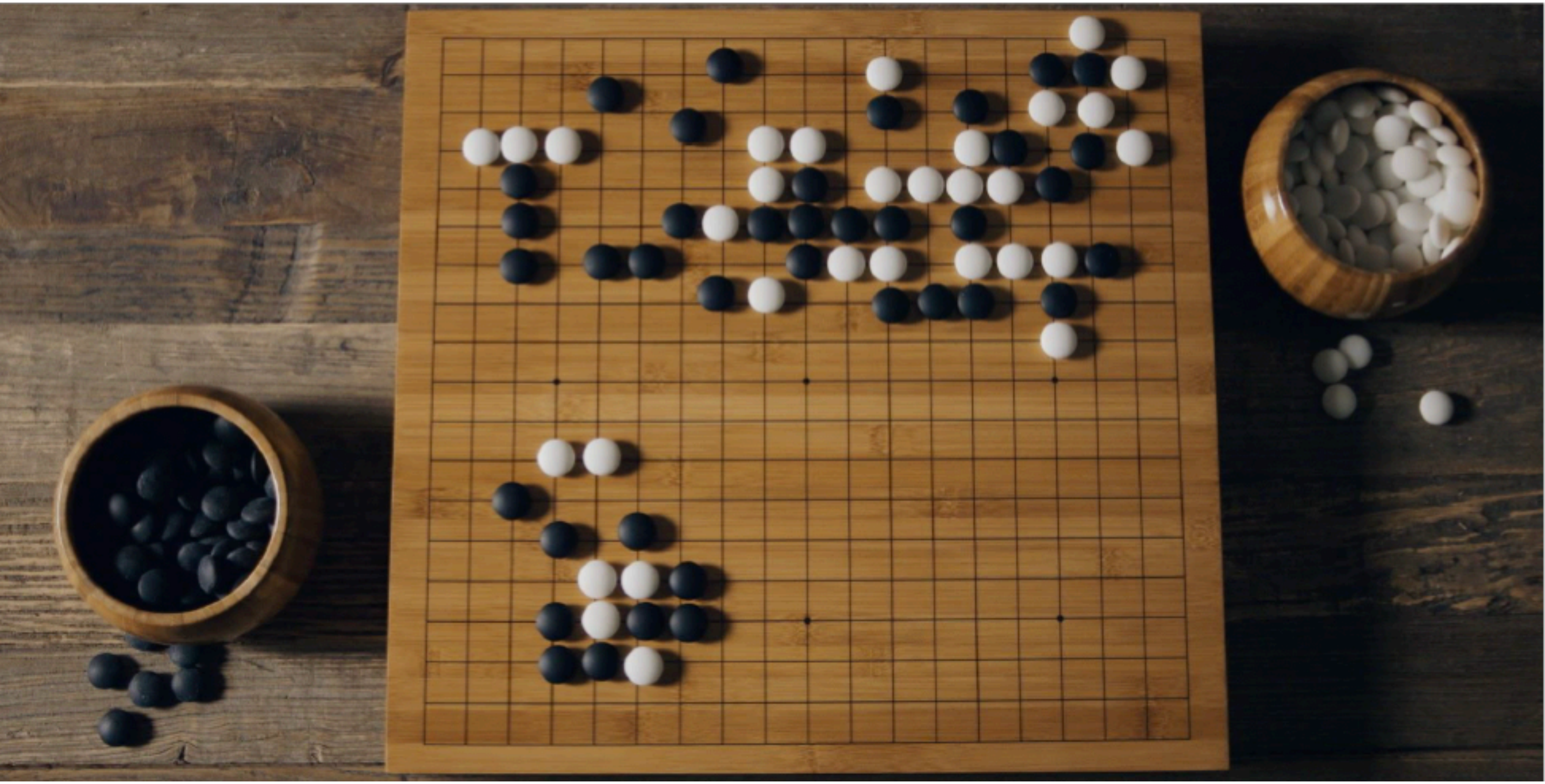
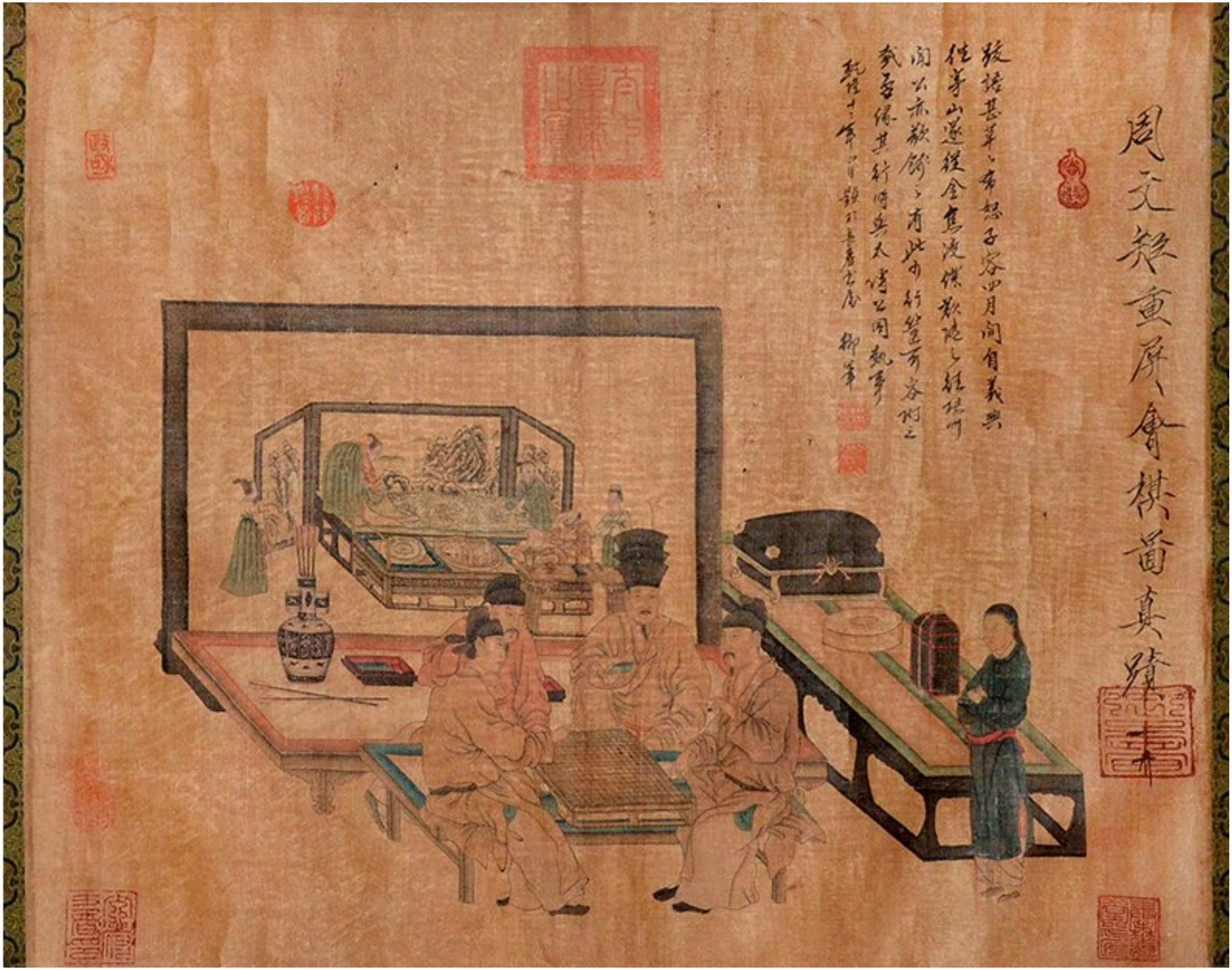
AlphaGo

Go (board game)



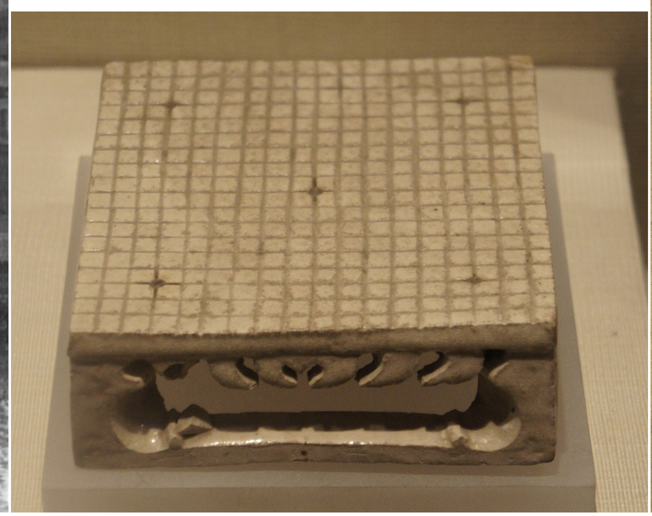
Invented more than 2,000 years ago, in China



Growth in East Asia



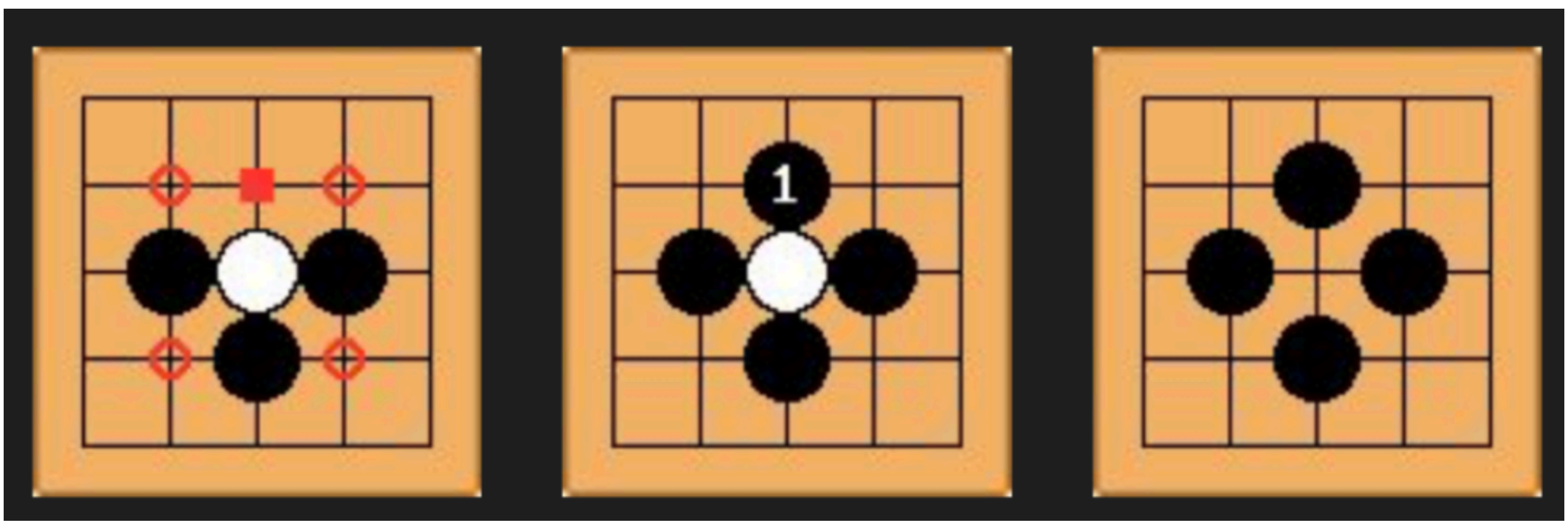





Currently over 40 million players

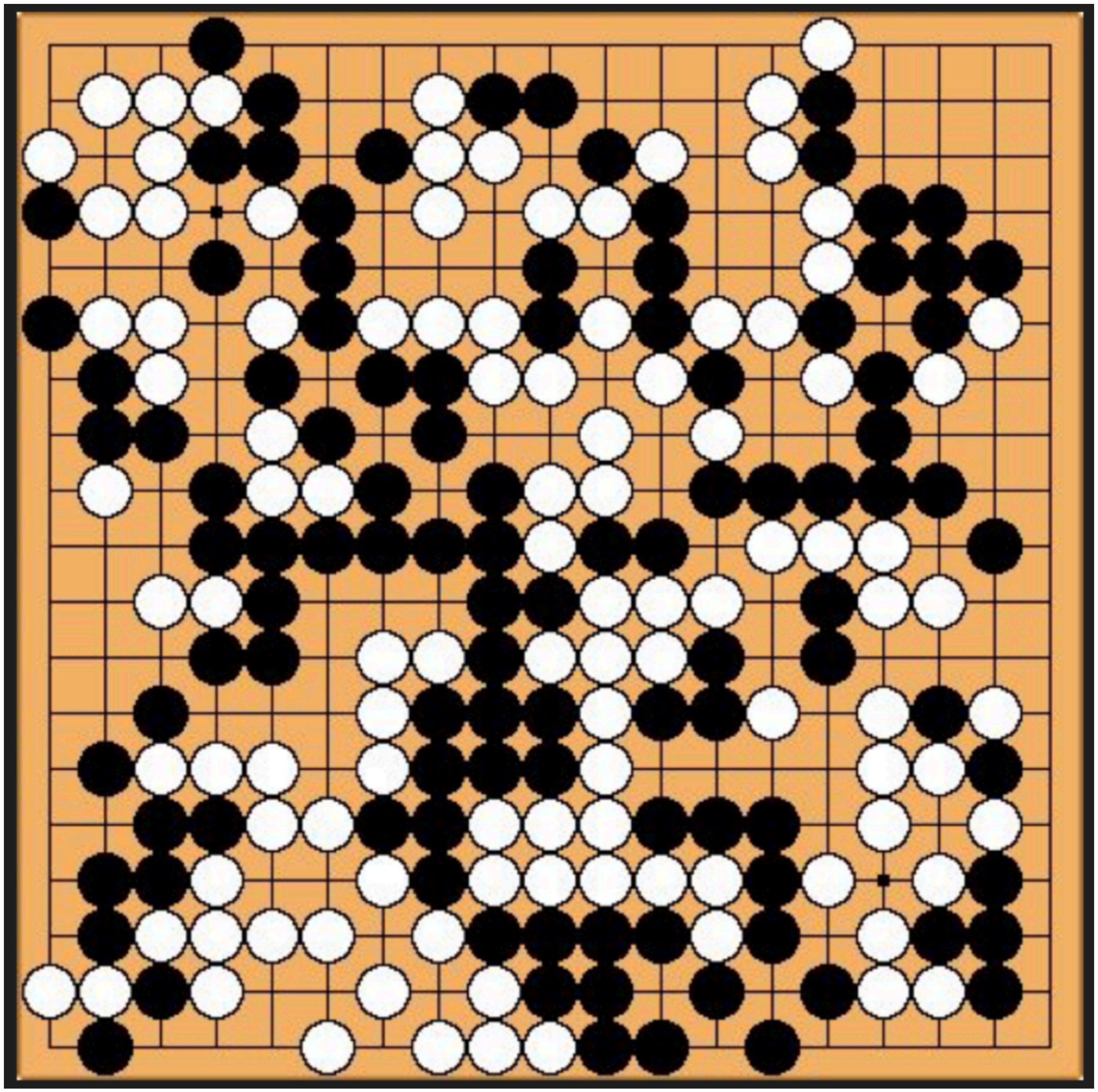
Rule of Go

- Played on a 19x19 board
- Two players, black and white, each place one stone per turn
- Capture the opponent's stones by surrounding them



Rule of Go

- Goal is to control as much territory as possible.



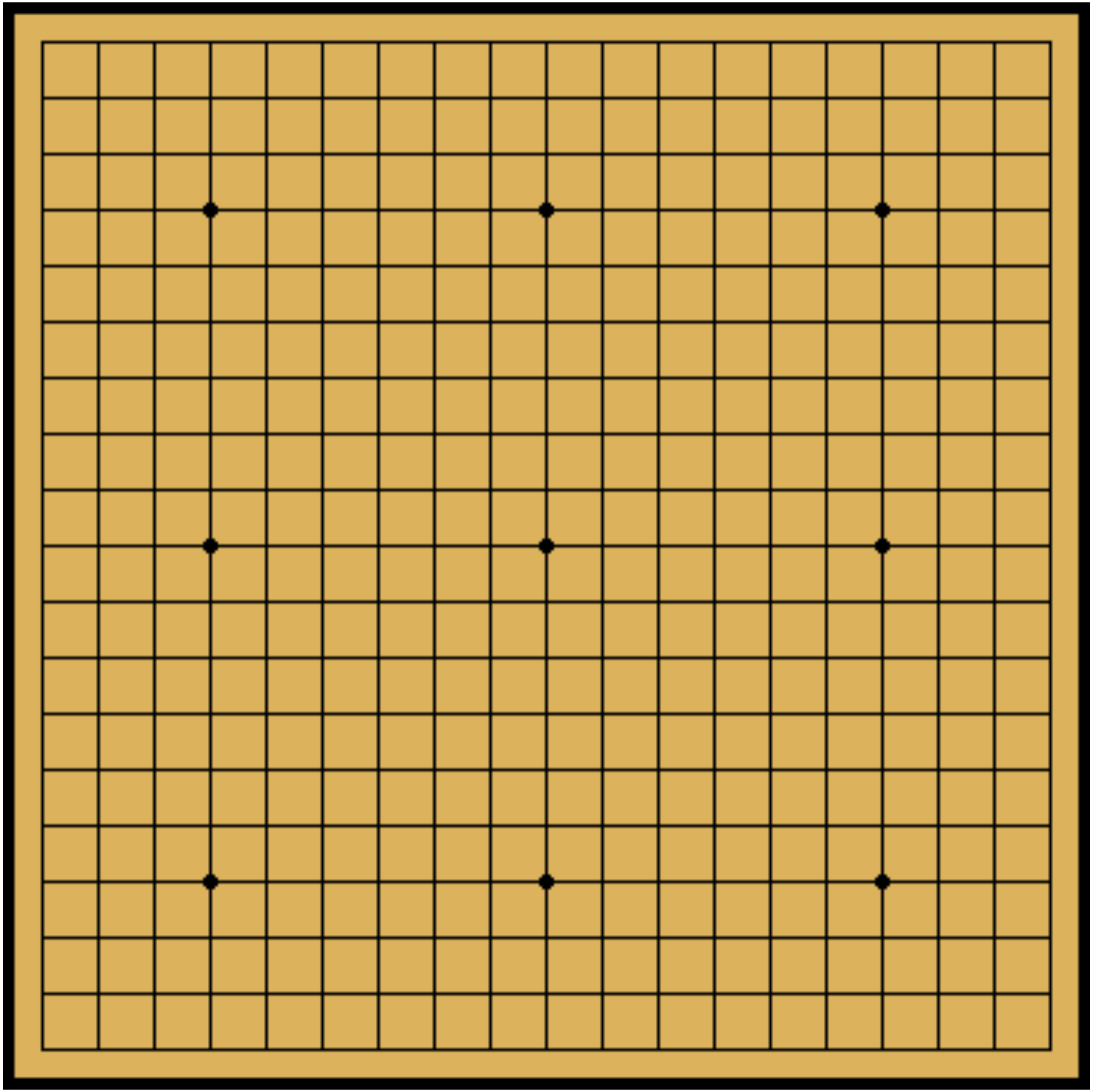
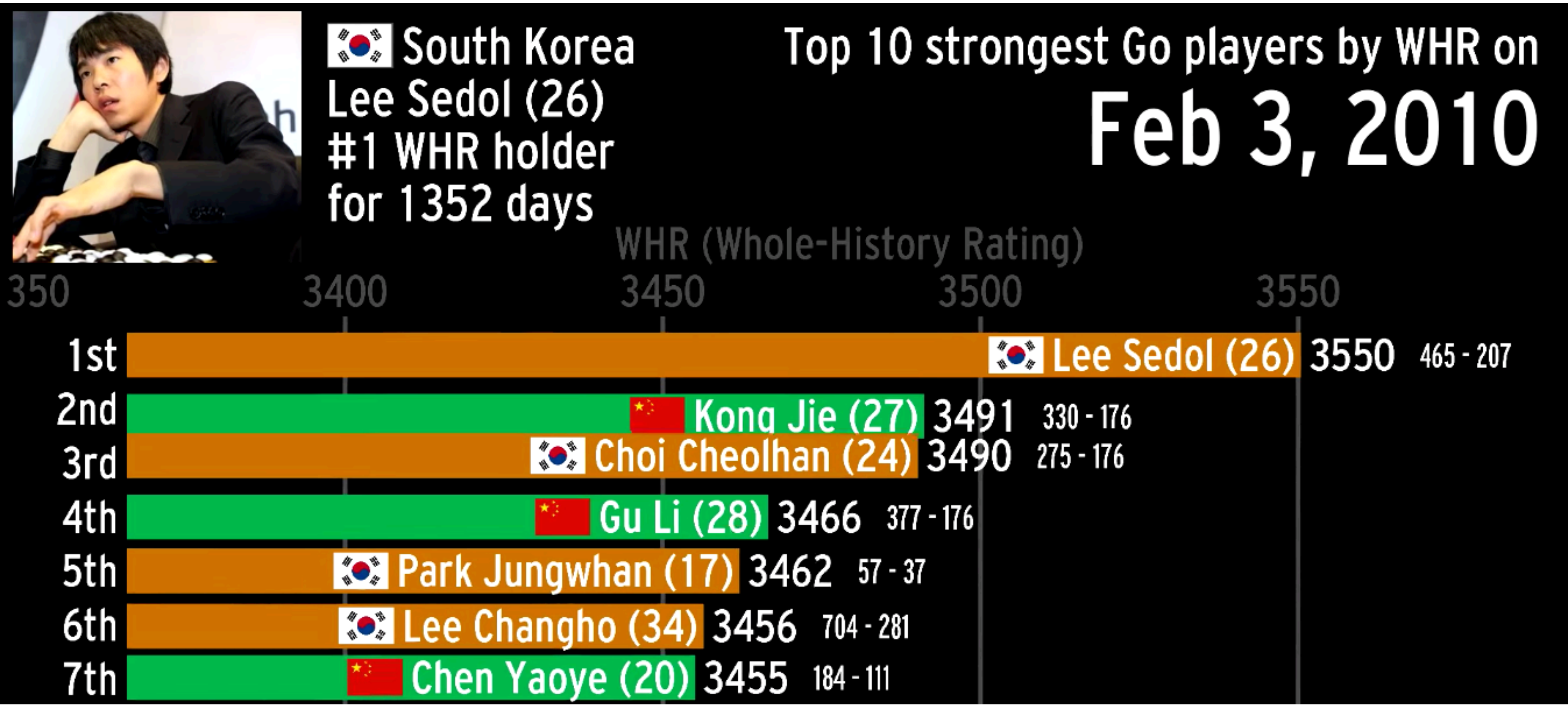


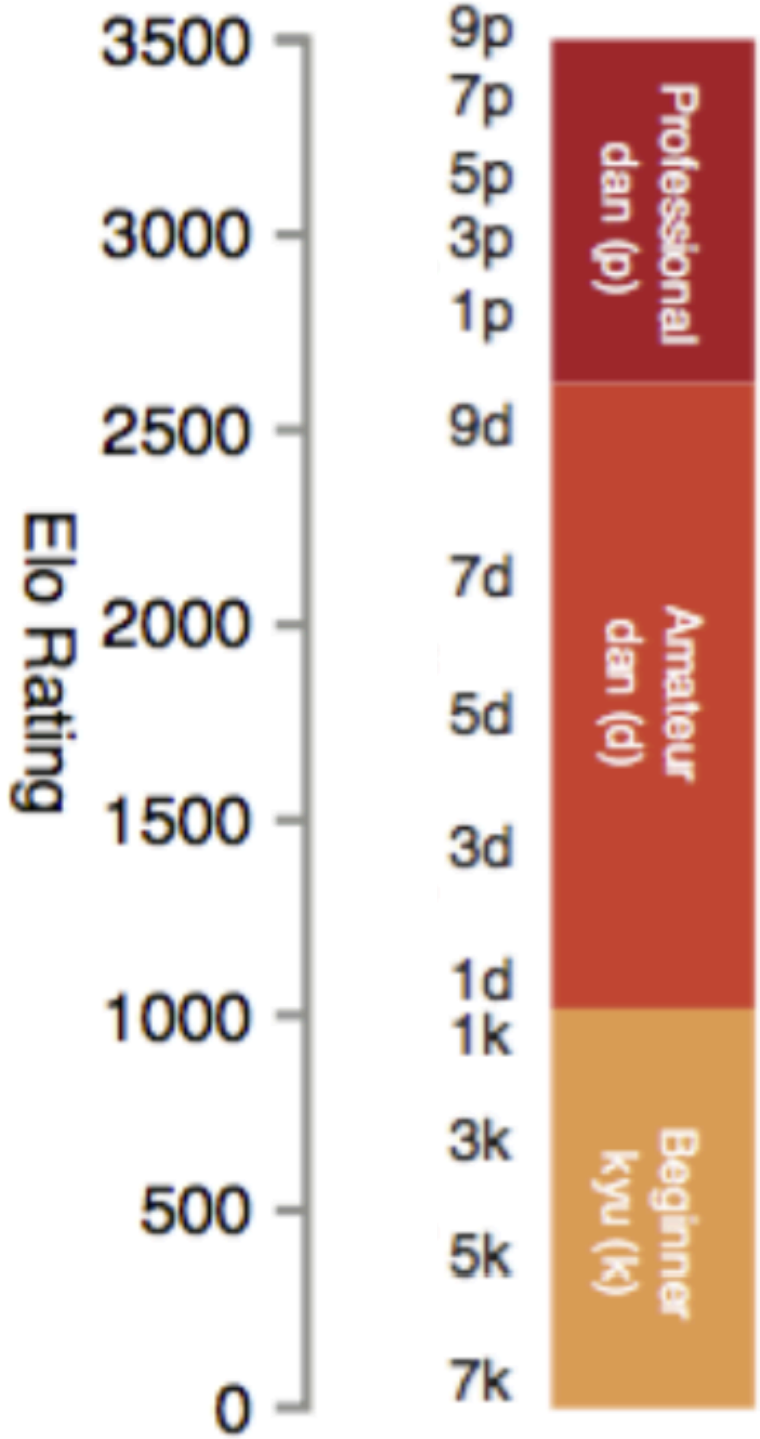
Photo source: Maeil Economics 2013/04

Lee Sedol

Go History



https://www.youtube.com/watch?v=oRvlyEpOQ-8

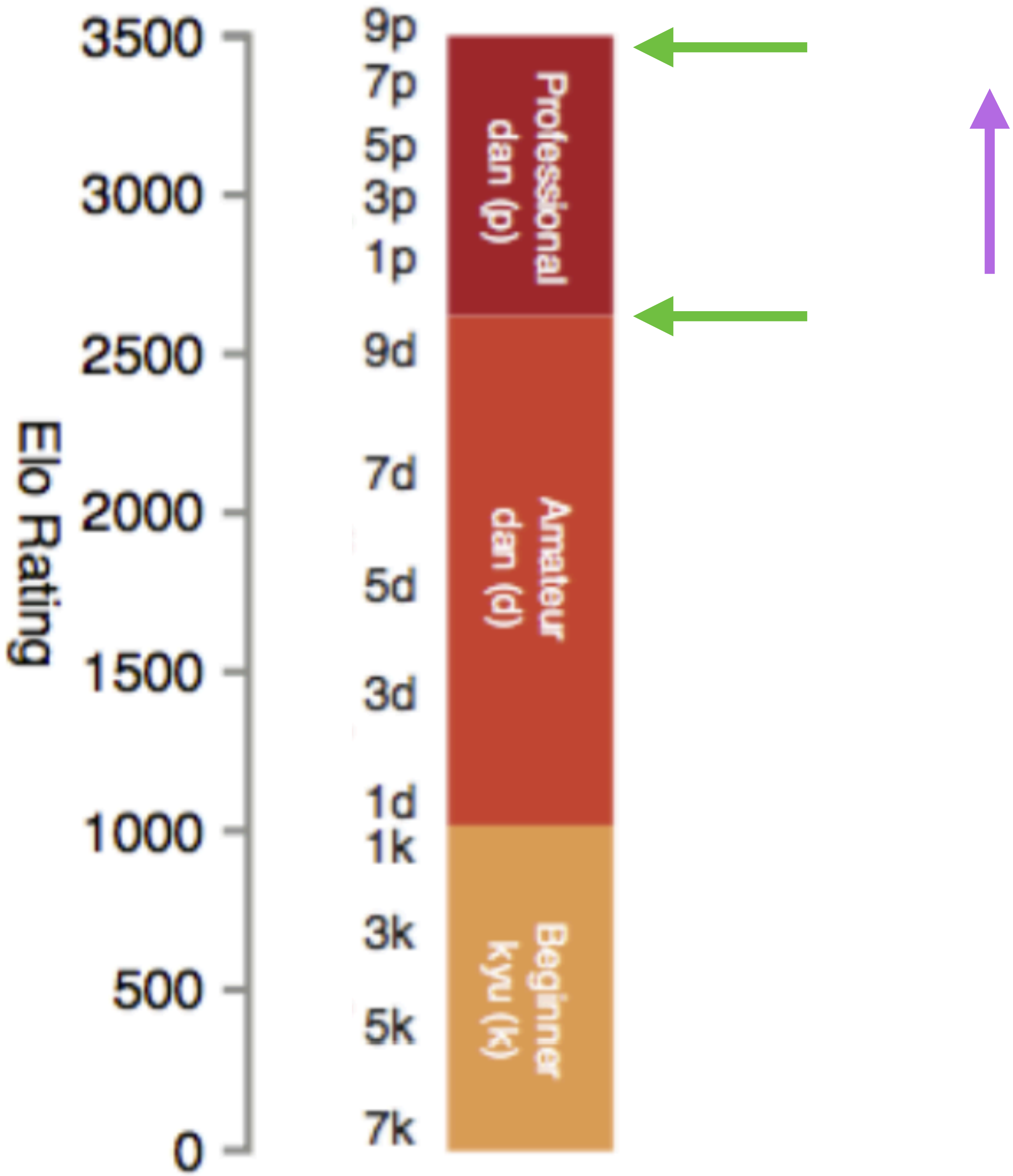


Elo Rating
3500
3000
2500
2000
1500
1000
500
0
9p
7p
5p
3p
1p
9d
7d
5d
3d
1d
1k
3k
5k
7k
Professional dan (p)
Amateur dan (d)
Beginner kyu (k)

Lee Se-dol (1983)

Rank	Date	
1 dan	2 July 1995	Promoted to professional dan rank after passing quali
2 dan	1 January 1998	
3 dan	1 January 2000	
6 dan	28 March 2003	Won the 7th LG Cup against Lee Changho (4 dan and
7 dan	6 May 2003	Runner up in the KT Cup against Yoo Changhyuk.
9 dan	7 July 2003	Won 16th Fujitsu Cup against Song Taekon (8 dan wa

- In 2003, advanced to 9p, became the youngest 9p (21) and the fastest player to reach 9p from 1p (**8 years**).

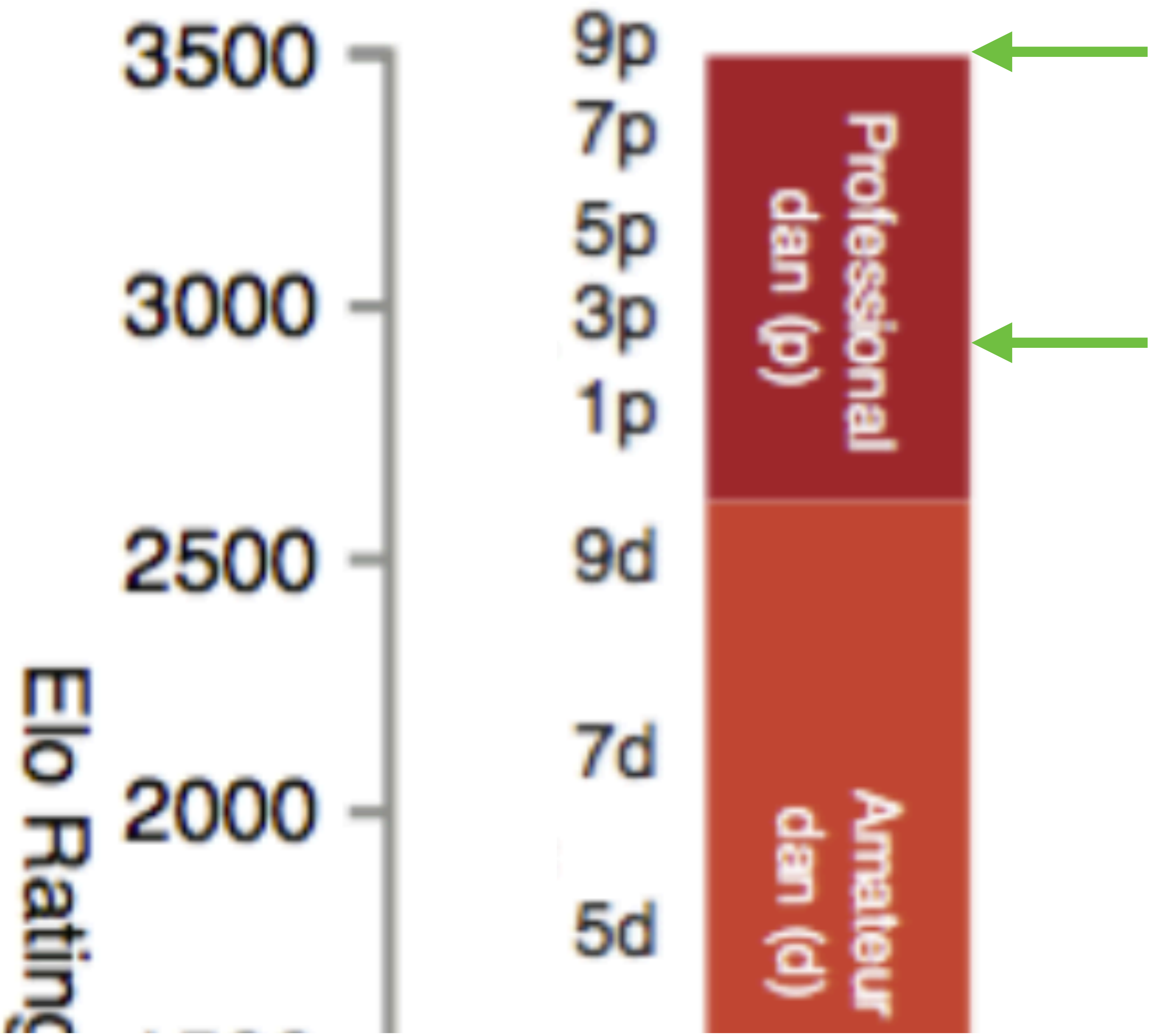


use **8** years

fastest

youngest

2001


3500
3000
2500
2000
Elo Rating
9p
7p
5p
3p
1p
9d
7d
5d
Professional dan (p)
Amateur dan (d)

vs

Lee Chang ho (9p dan)

LG CUP (2001)

9 dan vs 3 dan

LG CUP (2001)

LG CUP (2003)

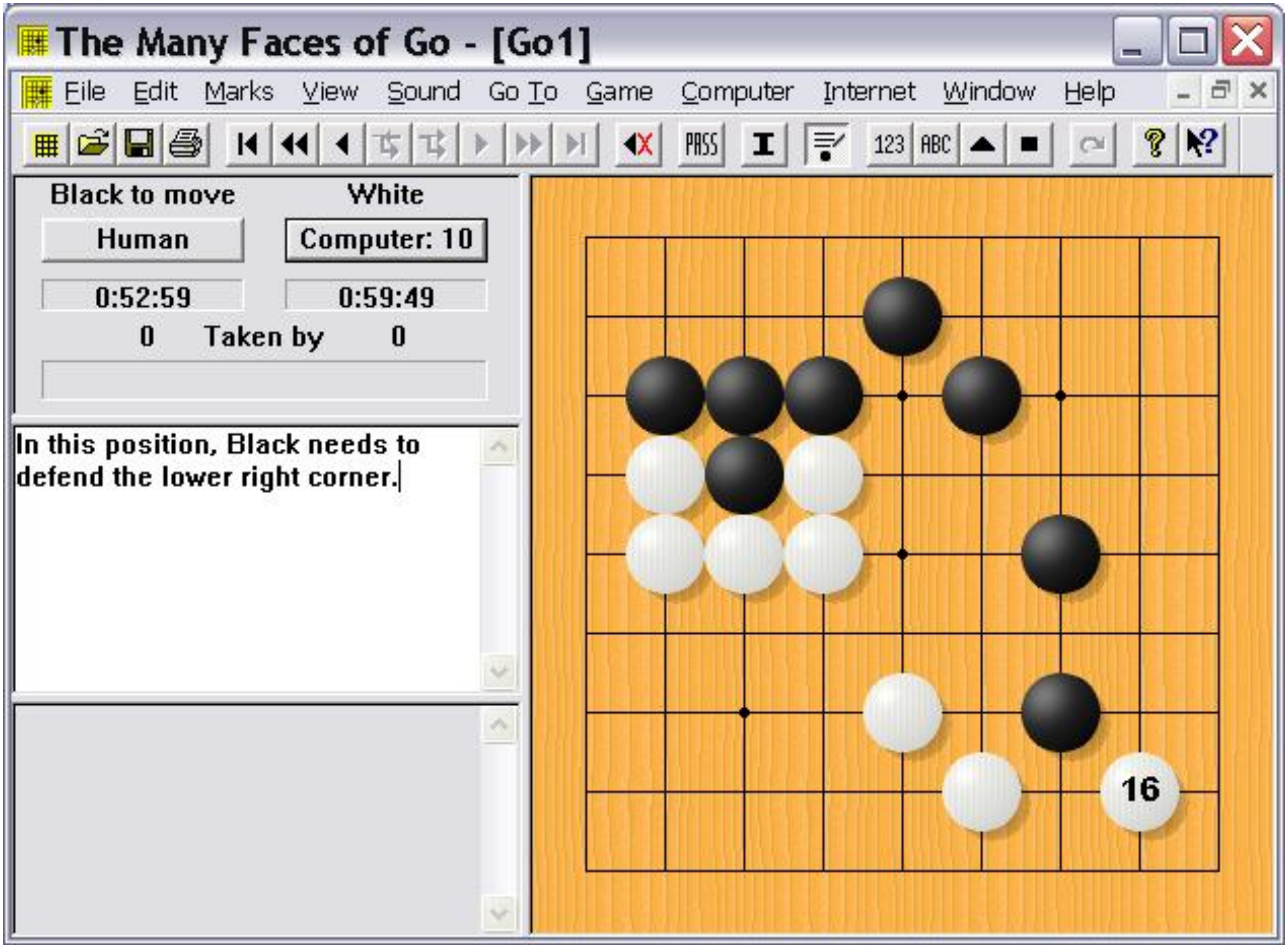
Years	Nat.	Winner	Score	Nat.	Runner-up
1996-1997		Lee Changho	3–0		Yoo Changhyuk
1997-1998		O Rissei	3–2		
1998-1999		Lee Changho	3–0		Ma Xiaochun
1999-2000		Yu Bin	3–1		Yoo Changhyuk
2000-2001		Lee Changho	3–2		Lee Sedol
2001-2002		Yoo Changhyuk	3–2		Cho Hunhyun
2002-2003		Lee Sedol	3–1		Lee Changho

Lee Sedol

wiki

Photo source: Maeil Economics 2013/04

Computer (AI)

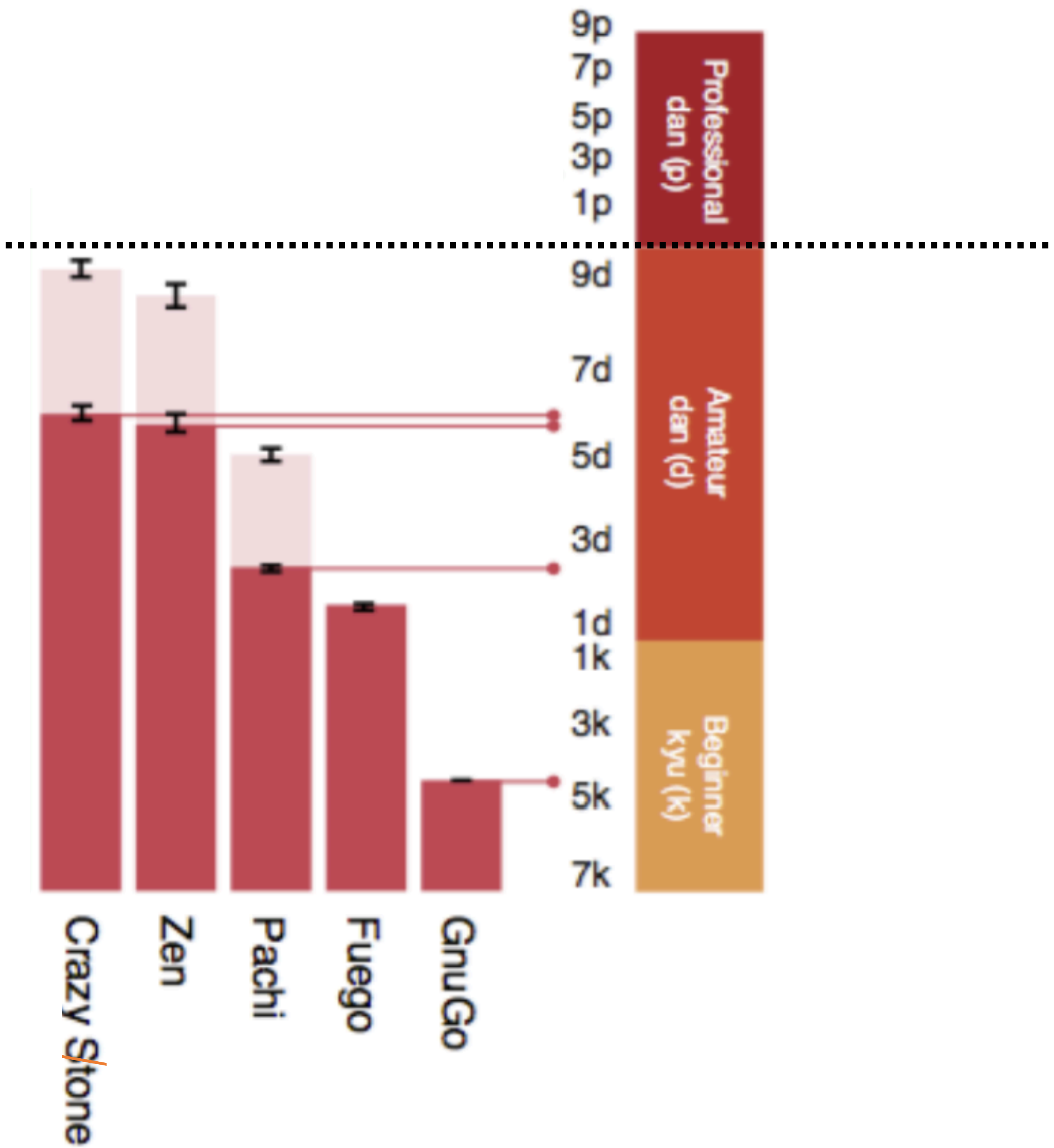

The Many Faces of Go - [Go1]
File Edit Marks View Sound Go To Game Computer Internet Window Help
PASS
123
ABC
Black to move
White
Human
Computer: 10
0:52:59
0:59:49
0 Taken by 0
In this position, Black needs to defend the lower right corner.
16

Human Expert vs Computer AI

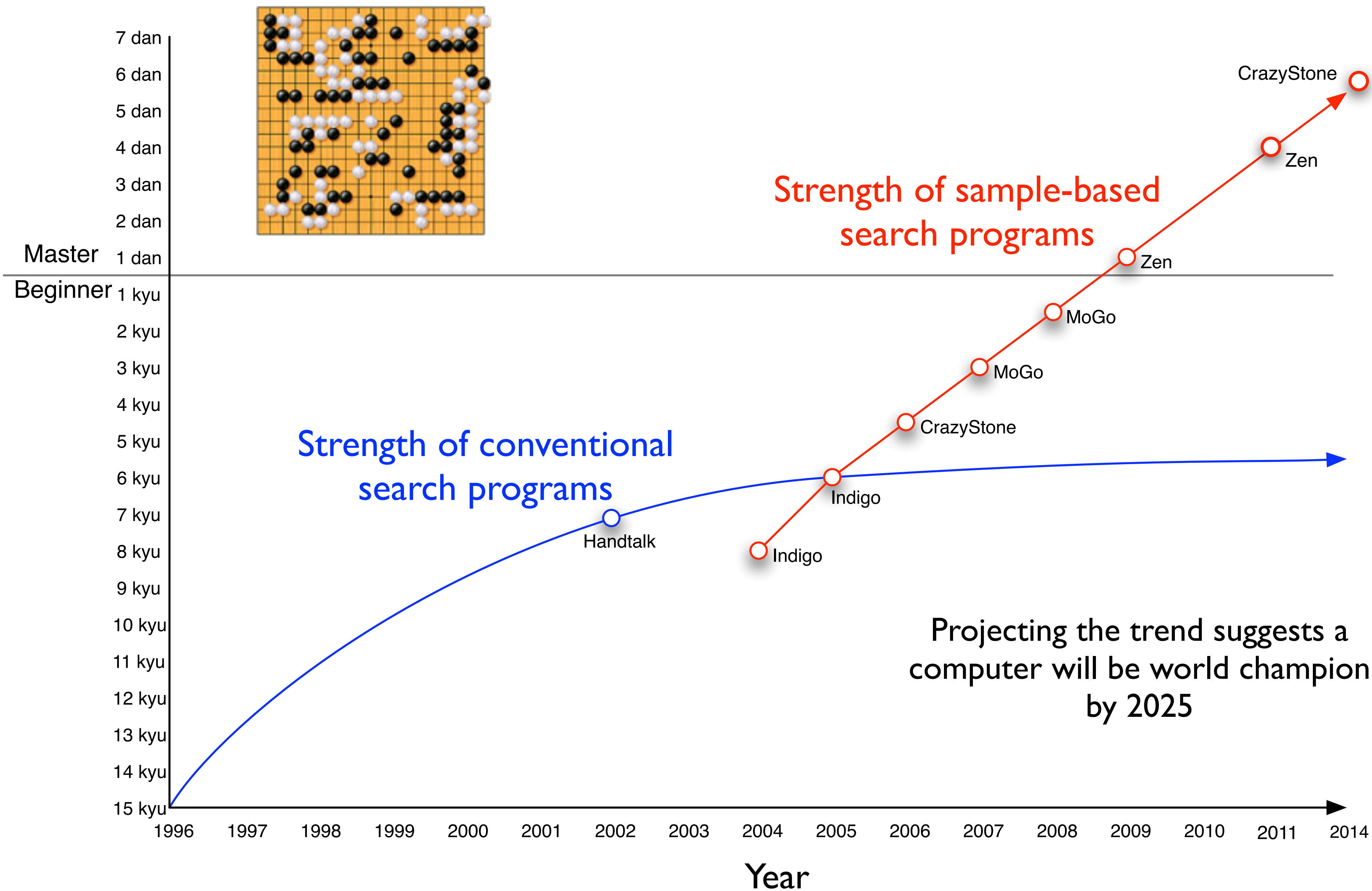
History of AI in Go

- 1997: Super human chess w/ Alpha-Beta + fast computer
- 2005: Computer Go is impossible!
- 2006: Monte-Carlo Tree Search applied to 9x9 Go (bit of learning)
- 2007: Human master level achieved at 9x9 Go (more learning)
- 2008: Human grandmaster level achieved at 9x9 Go (even more learning)
- 2012: Zen program beats former international champion with only 4 stone handicap in 19x19

if 4 handicap moves are given



Steady, exponential improvement (since MCTS, 2005) in the strength of the best computer Go programs



2015


nature
THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE
At last — a computer program that can beat a champion Go player PAGE 484
ALL SYSTEMS GO
CONSERVATION
SONGBIRDS À LA CARTE
Illegal harvest of millions of Mediterranean birds
PAGE 452
RESEARCH ETHICS
SAFEGUARD TRANSPARENCY
Don't let openness backfire on individuals
PAGE 459
POPULAR SCIENCE
WHEN GENES GOT 'SELFISH'
Dawkins's calling card 40 years on
PAGE 462
NATUREASIA.COM
28 January 2016
Vol. 529, No. 7587

AlphaGo vs European Champion (Fan Hui 2-Dan*)

rank

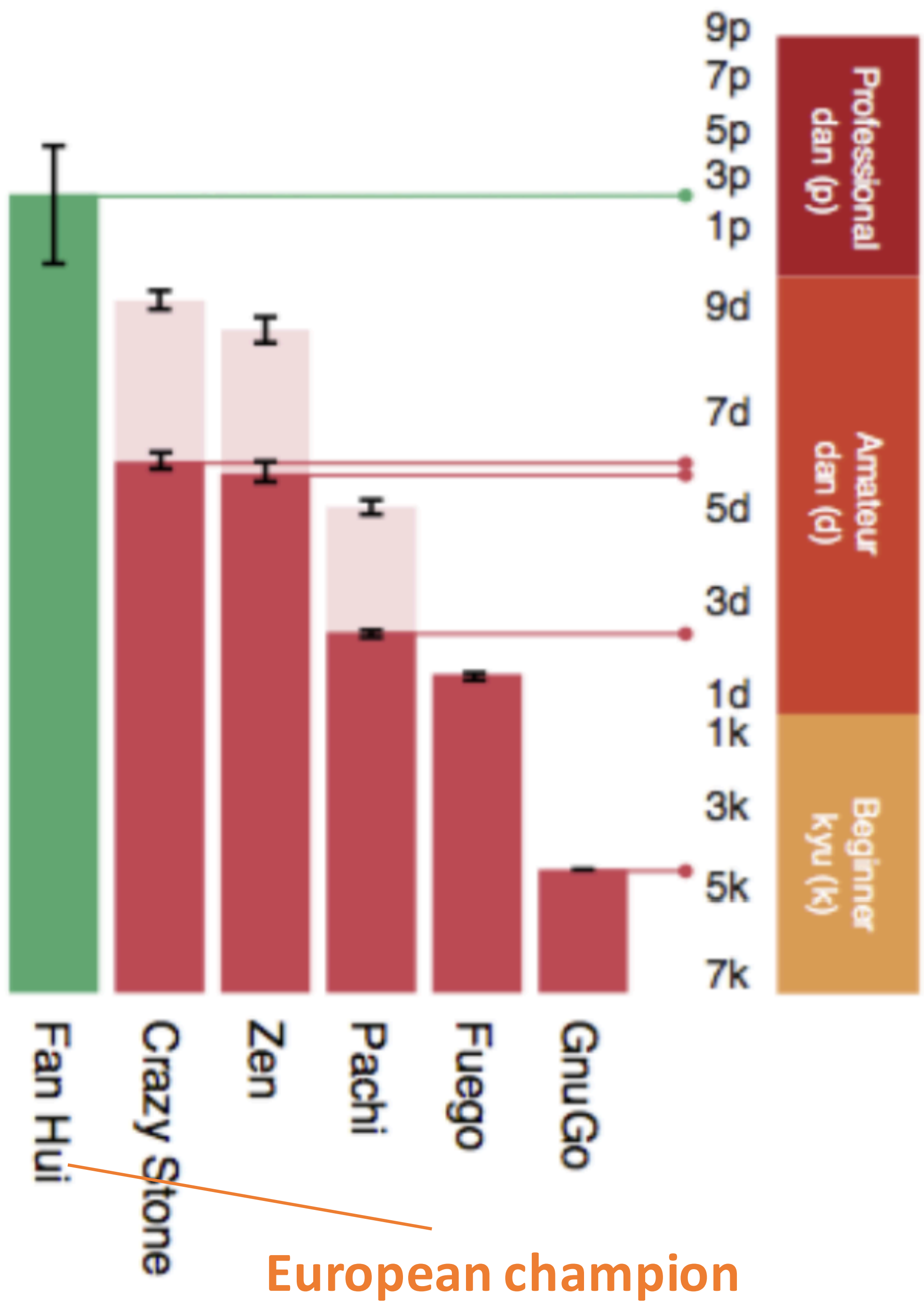
October 5 – 9, 2015

<Official match>

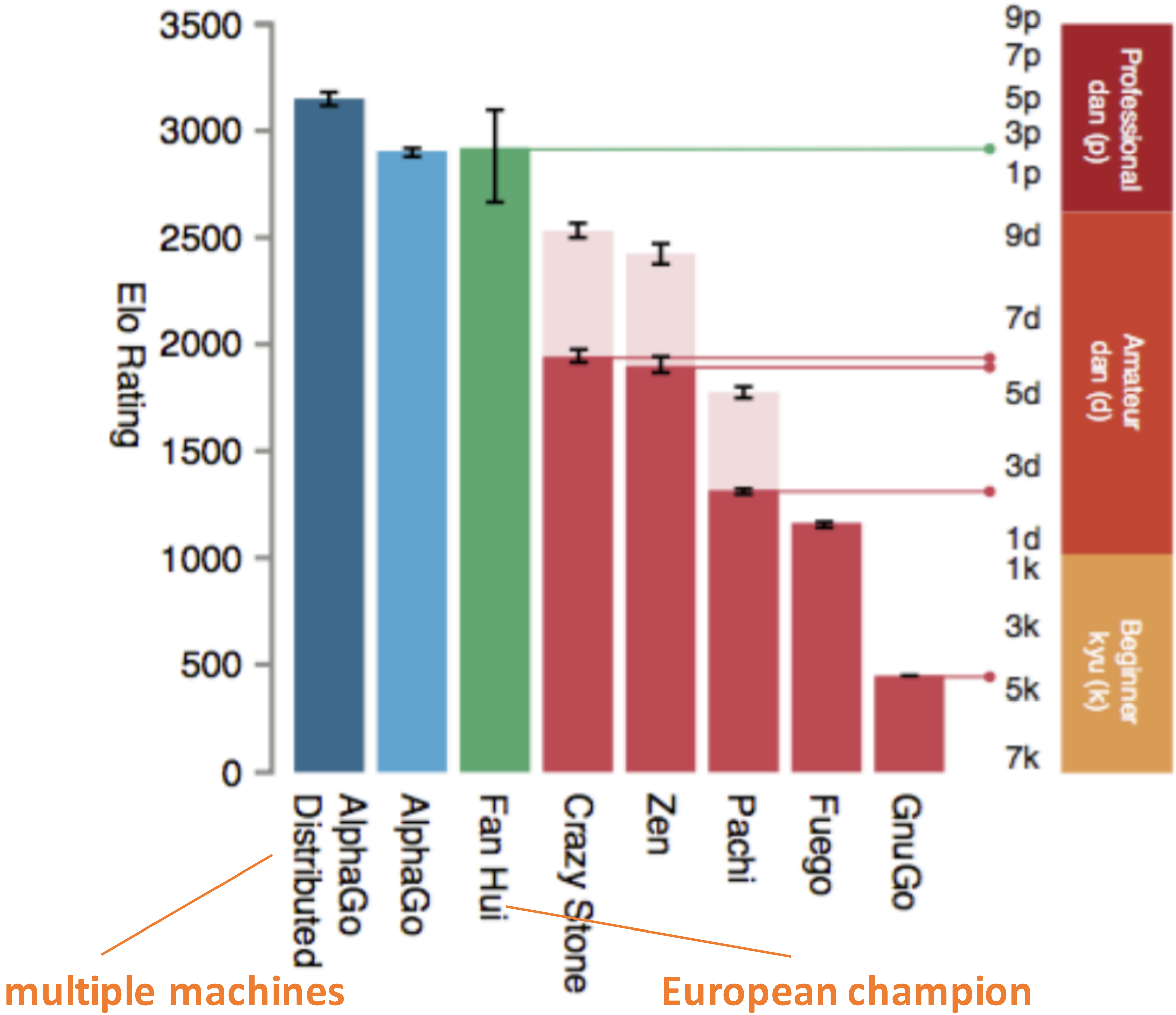
- Time limit: 1 hour
- AlphaGo Wins (5:0)

https://www.youtube.com/watch?v=53YLZBSS0cc

https://www.youtube.com/watch?v=g-dKXOlsf98



9p
7p
5p
3p
1p
9d
7d
5d
3d
1d
1k
3k
5k
7k
Professional dan (p)
Amateur dan (d)
Beginner kyu (k)
Fan Hui
Crazy Stone
Zen
Pachi
Fuego
GnuGo
European champion



3500
3000
2500
2000
1500
1000
500
0
Elo Rating
AlphaGo Distributed
AlphaGo
Fan Hui
Crazy Stone
Zen
Pachi
Fuego
GnuGo
9p
7p
5p
3p
1p
9d
7d
5d
3d
1d
1k
3k
5k
7k
Professional dan (p)
Amateur dan (d)
Beginner kyu (k)
multiple machines
European champion

AlphaGo vs World Champion (Lee Sedol 9-Dan)

March 9 – 15, 2016

<Official match>

- Time limit: 2 hours

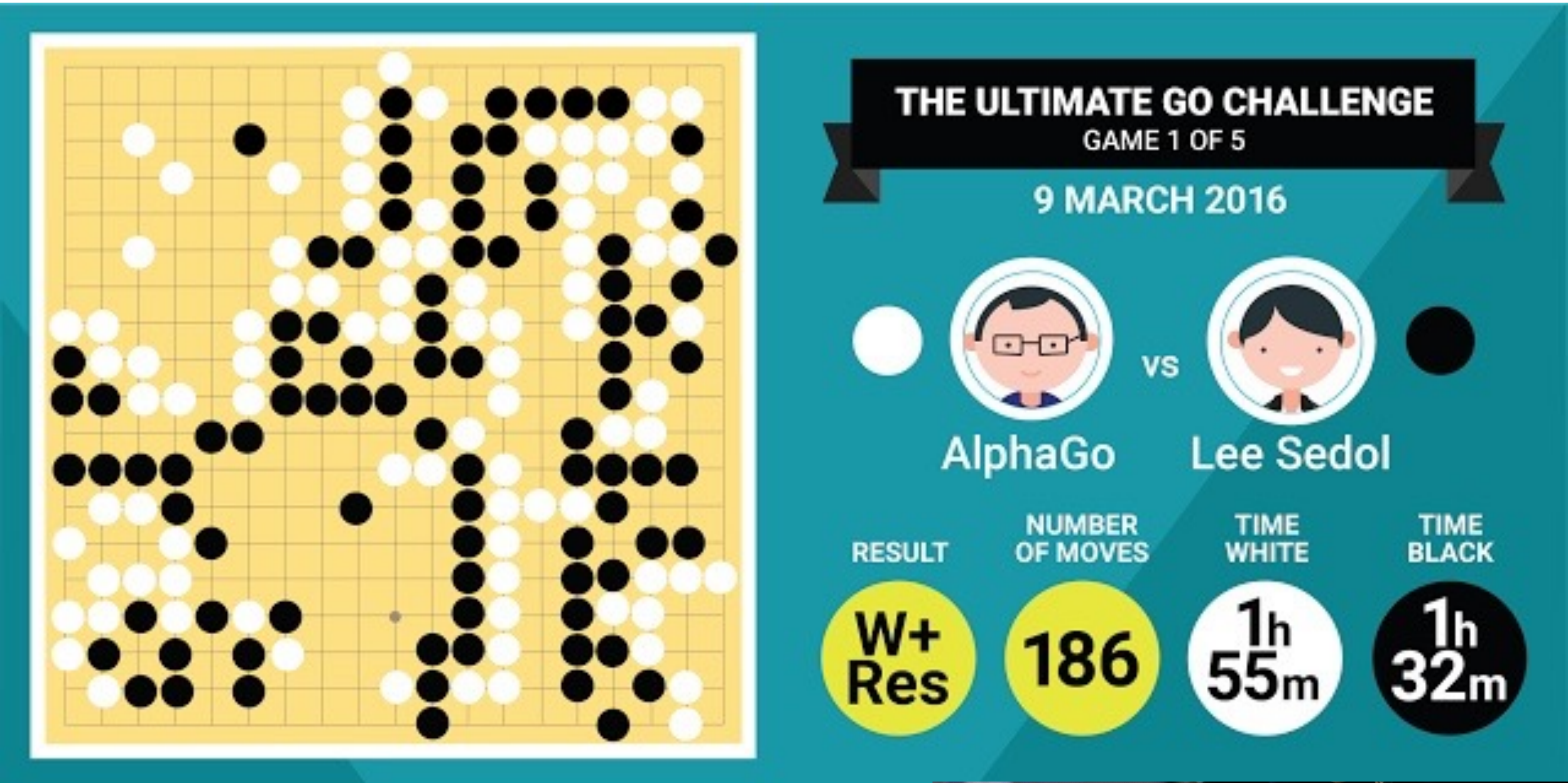
Venue: **Seoul**, Four Seasons Hotel

https://www.youtube.com/watch?v=8tq1C8spV_g


THE ULTIMATE GO CHALLENGE
AlphaGo
Lee Sedol
2015
BORN
1983
~5 dan pro
RANK
9 dan pro
0
TITLES
47

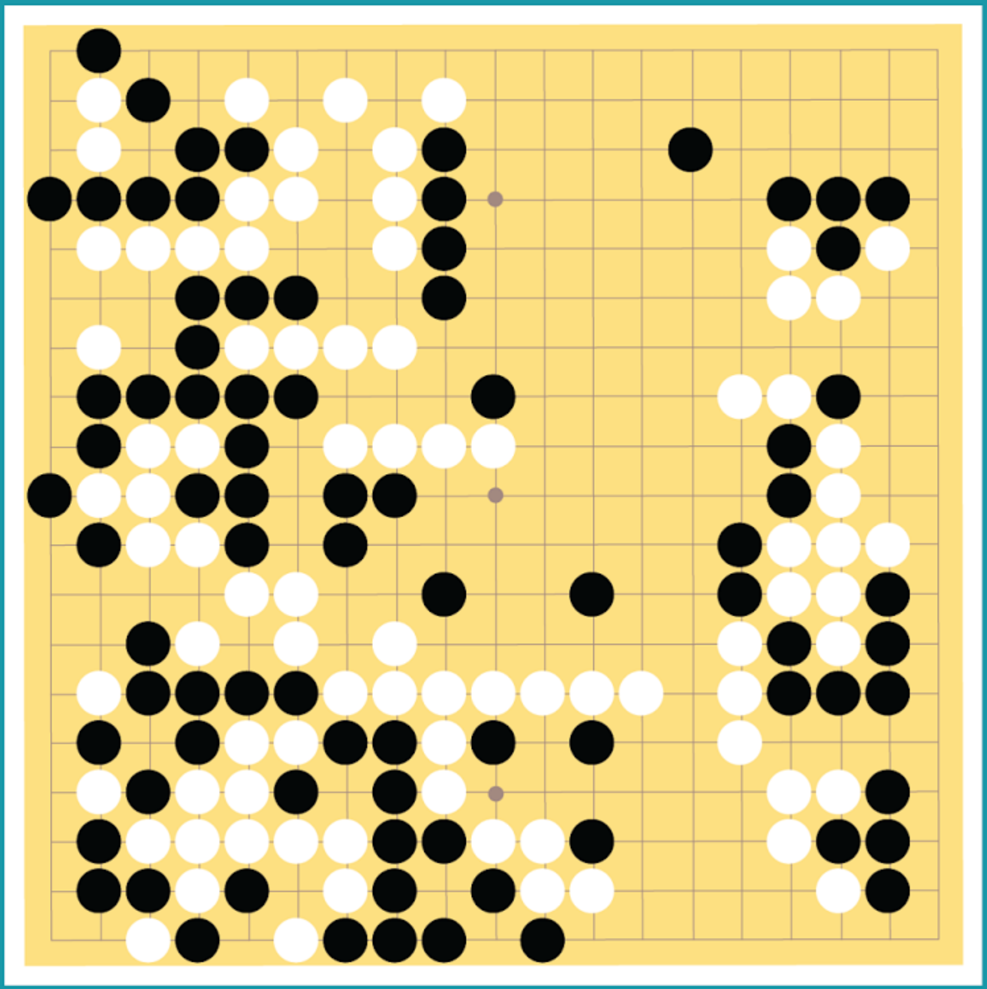
The Game


8 - 15 March 2016
AlphaGo
Lee Sedol
Google DeepMind
Challenge Match
8 - 15 March 2016


THE ULTIMATE GO CHALLENGE
GAME 1 OF 5
9 MARCH 2016
AlphaGo
vs
Lee Sedol
RESULT
W+ Res
NUMBER OF MOVES
186
TIME WHITE
1h 55m
TIME BLACK
1h 32m


ALPHAGO
LEE SEDOL
00:01:00


THE ULTIMATE GO CHALLENGE
GAME 2 OF 5
10 MARCH 2016
VS
AlphaGo
Won 2 of 5
Lee Sedol
Won 0 of 5
RESULT
B+ Res
NUMBER OF MOVES
211
TIME BLACK
2h+
TIME WHITE
2h+



THE ULTIMATE GO CHALLENGE

GAME 3 OF 5

12 MARCH 2016

AlphaGo

Won 3 of 5

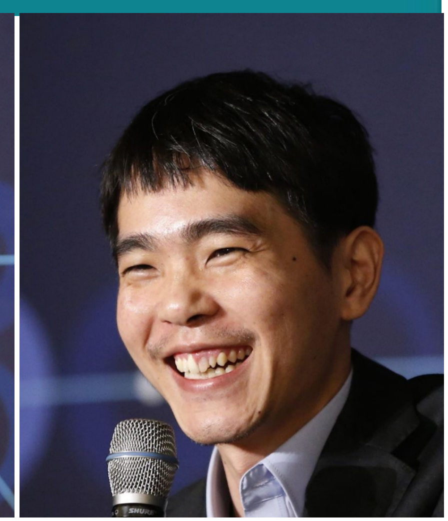
vs

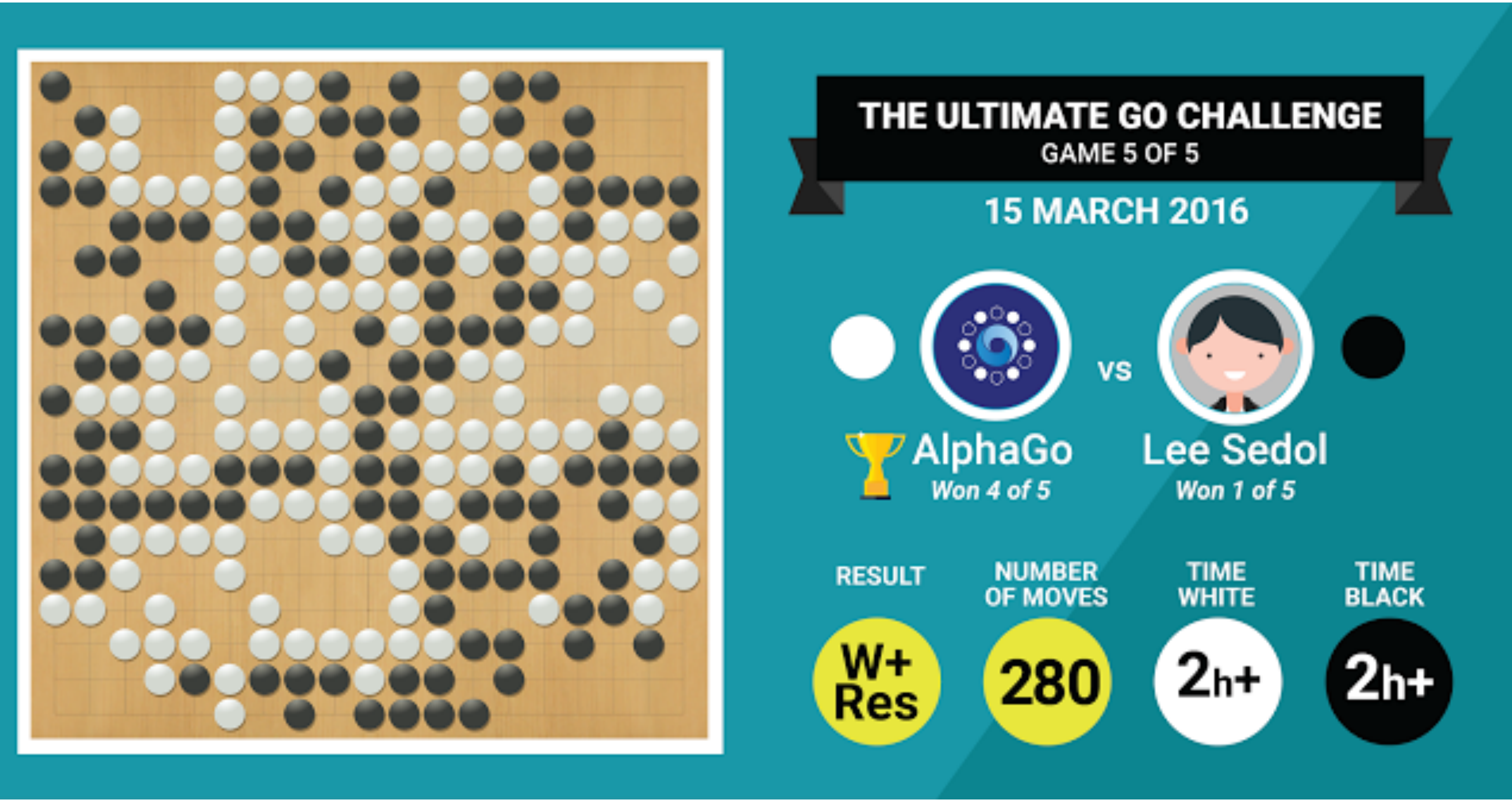
Lee Sedol

Won 0 of 5

RESULT	NUMBER OF MOVES	TIME WHITE	TIME BLACK
W+ Res	176	1h 51m	2h+


© Google
THE ULTIMATE GO CHALLENGE
GAME 4 OF 5
13 MARCH 2016
VS
AlphaGo
Won 3 of 5
Lee Sedol
Won 1 of 5
RESULT
W+ Res
NUMBER OF MOVES
180
TIME BLACK
1h 59m 6s
TIME WHITE
2h+




THE ULTIMATE GO CHALLENGE
GAME 5 OF 5
15 MARCH 2016
AlphaGo
Won 4 of 5
vs
Lee Sedol
Won 1 of 5
RESULT
W+ Res
NUMBER OF MOVES
280
TIME WHITE
2h+
TIME BLACK
2h+

Final Result

Game	Date	Black	White
1	9 March 2016	Lee Sedol	**AlphaGo**
2	10 March 2016	**AlphaGo**	Lee Sedol
3	12 March 2016	Lee Sedol	**AlphaGo**
4	13 March 2016	AlphaGo	**Lee Sedol**
5	15 March 2016	Lee Sedol[note 1]	**AlphaGo**
			Result: **AlphaGo 4 – 1 Lee Sedol**


Challenge Match
8 - 15 March 2016
AlphaGo
Lee Sedol


PRESS PASS

Go Elo Ranking

For older ratings, check the History page. There is also a History of top ladies.

Rank	Name	♂ ♀	Flag	Elo
1	Ke Jie	♂		3622
2	Google DeepMind AlphaGo			3595
3	Park Junghwan	♂		3551
4	Lee Sedol	♂		3532
5	Shi Yue	♂		3529
6	Iyama Yuta	♂		3527
7	Kim Jiseok	♂		3513
8	Zhou Ruiyang	♂		3512
9	Park Yeonghun	♂		3510
10	Mi Yuting	♂		3507
11	Lian Xiao	♂		3504
12	Tang Weixing	♂		3487
13	Tuo Jiaxi	♂		3484

http://www.goratings.org/history/

Lee Sedol VS Ke Jie

60 wins streak

0 loss

Rank	Name	♂ ♀	Flag	Elo
1	Ke Jie	♂		3622
2	Google DeepMind AlphaGo			3595
3	Park Junghwan	♂		3551
4	Lee Sedol	♂		3532
5	Shi Yue	♂		3529
6	Iyama Yuta	♂		3527
7	Kim Jiseok	♂		3513
8	Zhou Ruiyang	♂		3512
9	Park Yeonghun	♂		3510
10	Mi Yuting	♂		3507
11	Lian Xiao	♂		3504
12	Tang Weixing	♂		3487
13	Tuo Jiaxi	♂		3484

2017

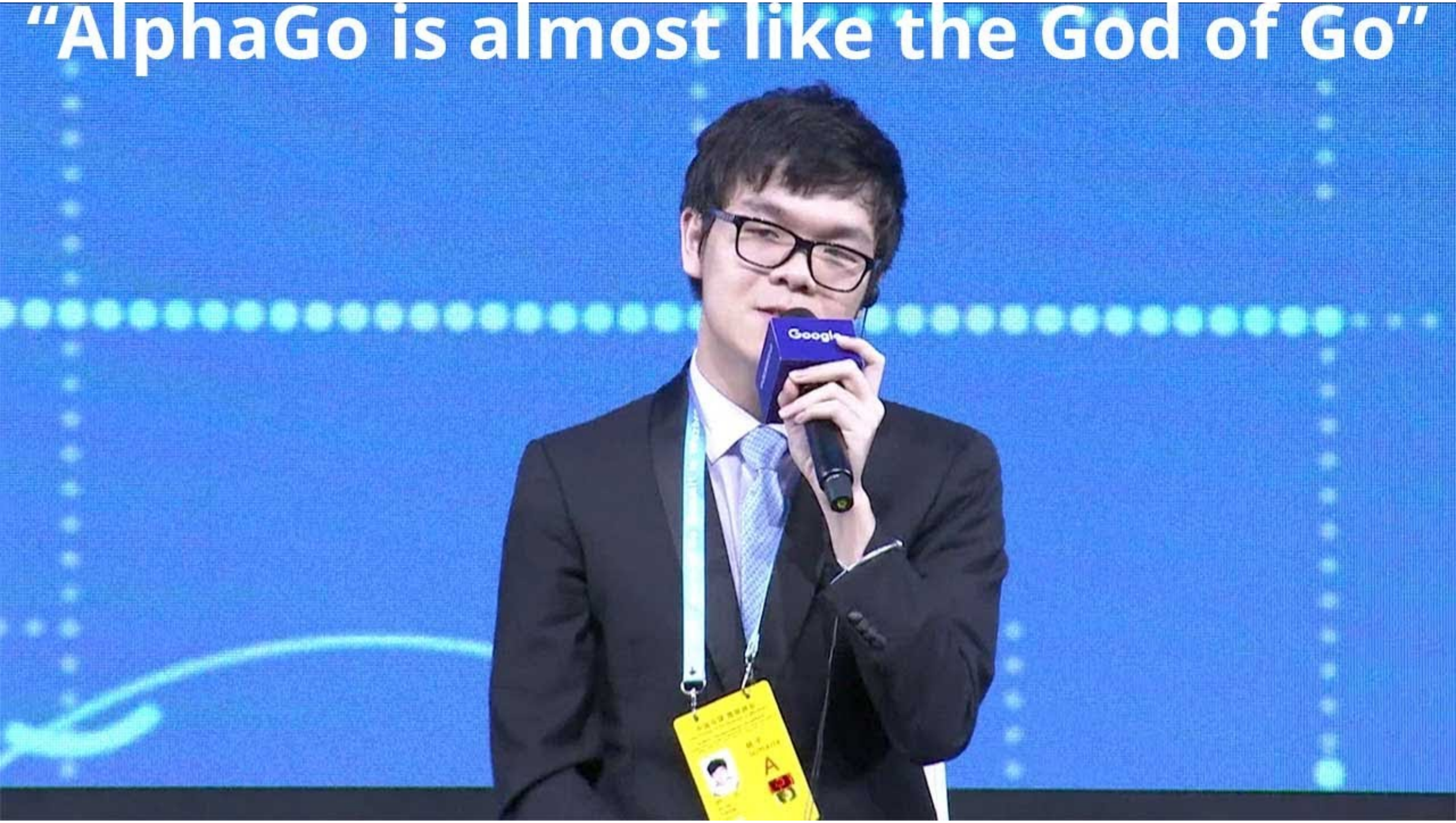
2017

中国围棋协会
Google
浙江省体育局
中国乌镇
棋峰会
The Future of Go
in Wuzhen
KE JIE
03:00:00
ALPHAGO
03:00:00

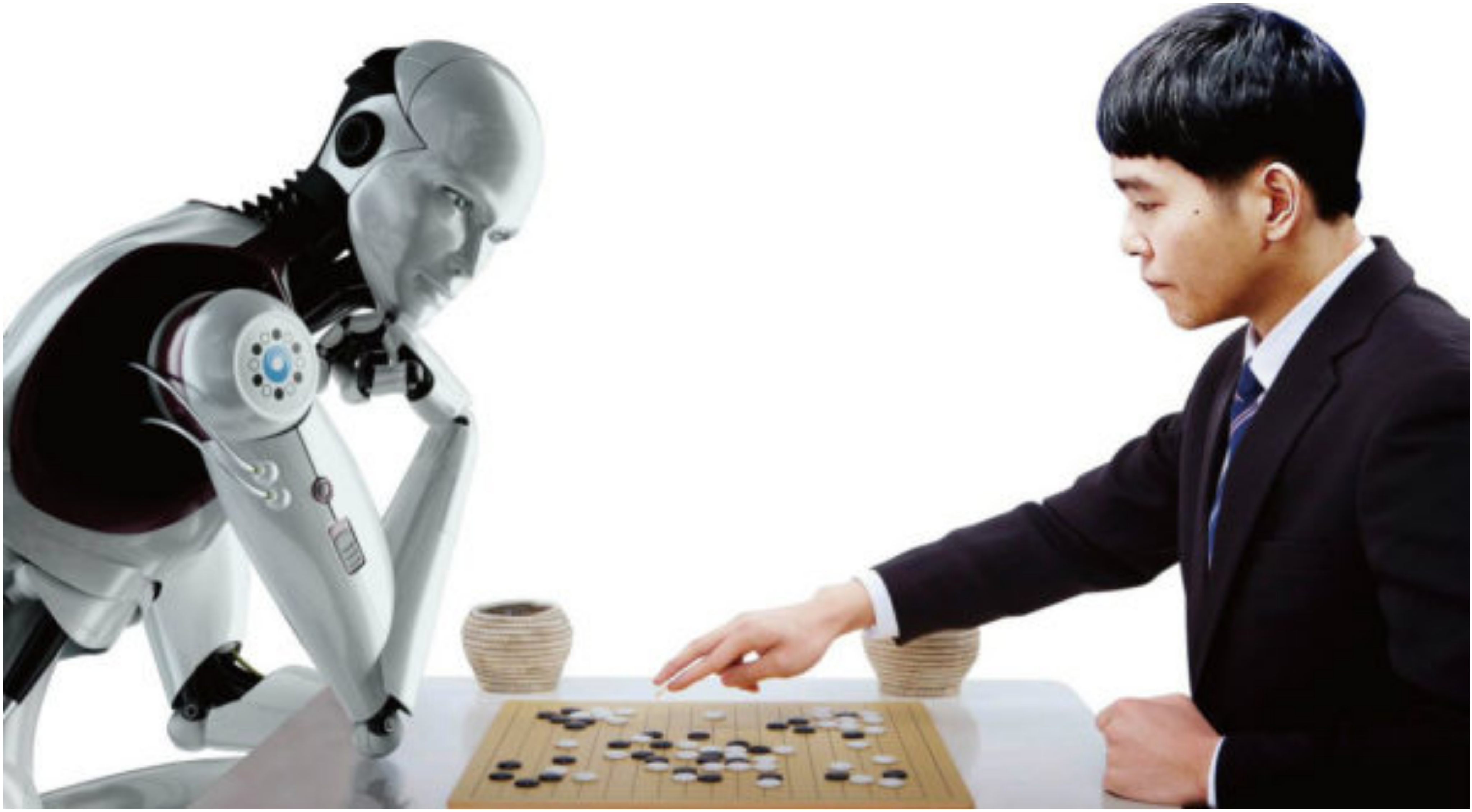

柯洁 KE JIE
00:46:57
ALPHAGO
02:08:18


AlphaGo vs Ke Jie
3 : 0

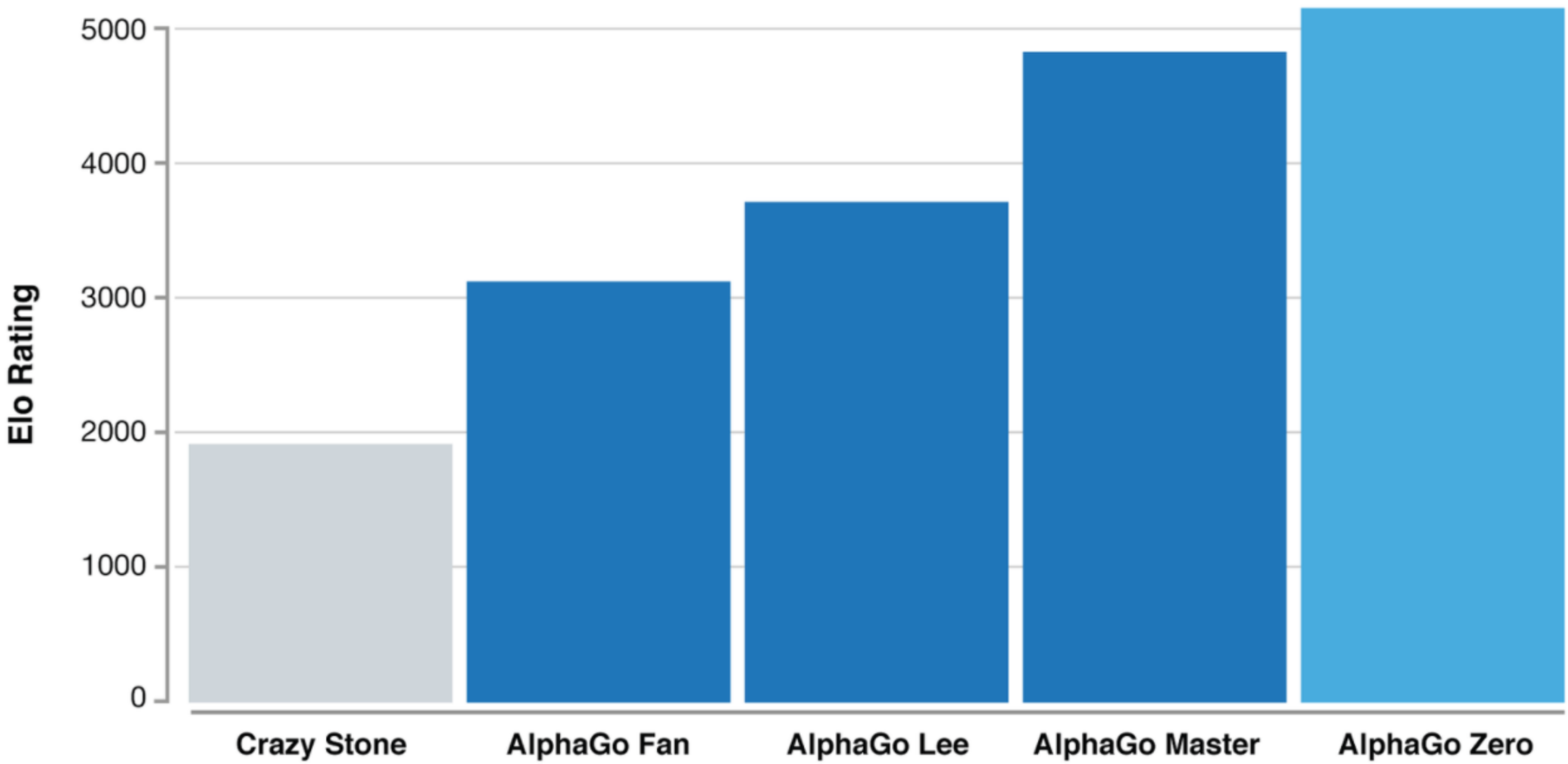
After the game (2017)



Lee Se-dol won 1 game

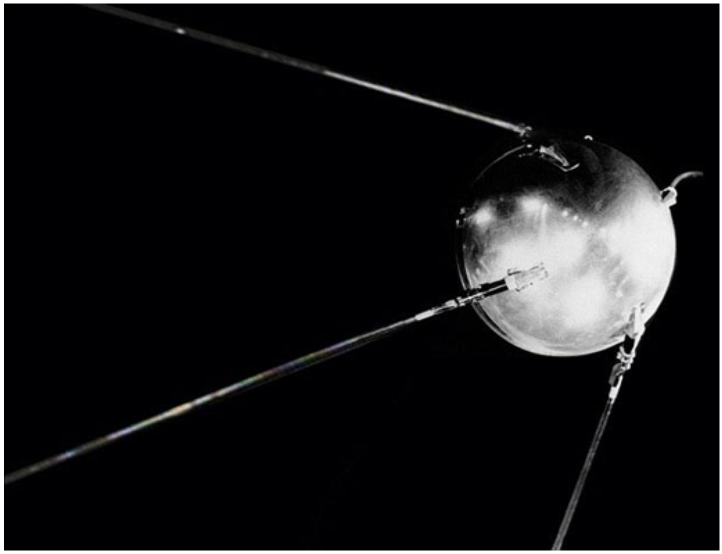


Different versions of AlphaGo



Elo ratings - a measure of the relative skill levels of players in competitive games such as Go - show how AlphaGo has become progressively stronger during its development sputnik moment

1957



turday Evening, October 5, 1957 (UP)—Means United Press Price: Five Cen

Russians Win Race To Launch Earth Satellite

Man On Threshold Of Space Travel

By DANIEL F. GILMORE
United Press Staff Correspondent

LONDON (UP)—The pulsating radio "beep" of the first manmade earth satellite signalled today to the world that man had crossed the threshold into the age of travel through space.

The Soviet Union announced it had won the race into space by launching an earth satellite Friday, a 184-pound, 22-inch globe now orbiting the earth at 18,000 miles an hour, 560 miles up.

Millions of persons throughout the world heard the "beep...beep...beep..." rebroadcast today by local stations and realized that man had taken his first faltering steps into the new era.

Launching of the satellite was a tremendous victory for science. It was a more tremendous victory for Soviet propaganda to be able to trumpet to the world the Russians were the first to break through the frontiers of space.

Bolsters ICBM Claims

It bolstered Russian claims to

— WEATHER —

WEST VIRGINIA—Partly cloudy with highest in the 60s today and Sunday. Lowest tonight 50 west and 40 east portions.

VIRGINIA—Fair with lowest 45 to 50 west and north and 50 to 55 southeast portions tonight, Sunday mostly sunny and a little warmer. Tides on the coast and lower bay will run a foot or two above normal.

How To Spot Satellite

By UNITED PRESS

Here's how to look for the Russian earth satellite which will be whizzing through the sky at 18,000 miles an hour.

The best time to spot it is at dawn or dusk when the sky is semi-dark. There is a chance that it could be seen if it travels across the face of the moon at night.

The best instruments to use are ordinary binoculars or telescopes. Powerful telescopes won't pick it up because of their narrow fields.

Through optical instruments, the satellite will look like the faintest star which can be seen with the naked eye.

Keep a sharp eye out. The satellite travels so fast it may appear on the horizon for only seconds and chances of spotting it have been estimated at one in a hundred.

Epic-Making

U. S. May Speed Up Satellite Program

By JOSEPH L. MYLER
United Press Staff Correspon

WASHINGTON (UP)—Amer scientists, caught flatfooted Russia's epic launching of the man-made moon, indicated t the United States may spee its own earth satellite progra

Leaders of the U.S. satellite gram also said that it ap Russia rocketed its heavy pound satellite into a globe dling orbit with a rocket to" an intercontinental ba missile.

That could mean Russi only has beaten this country frontiers of space, but also t has been called the "ul weapon" for modern day wa ICBM. This country has n tested a successful ICBM.

American diplomats co Russia had scored a notabl

Go Community

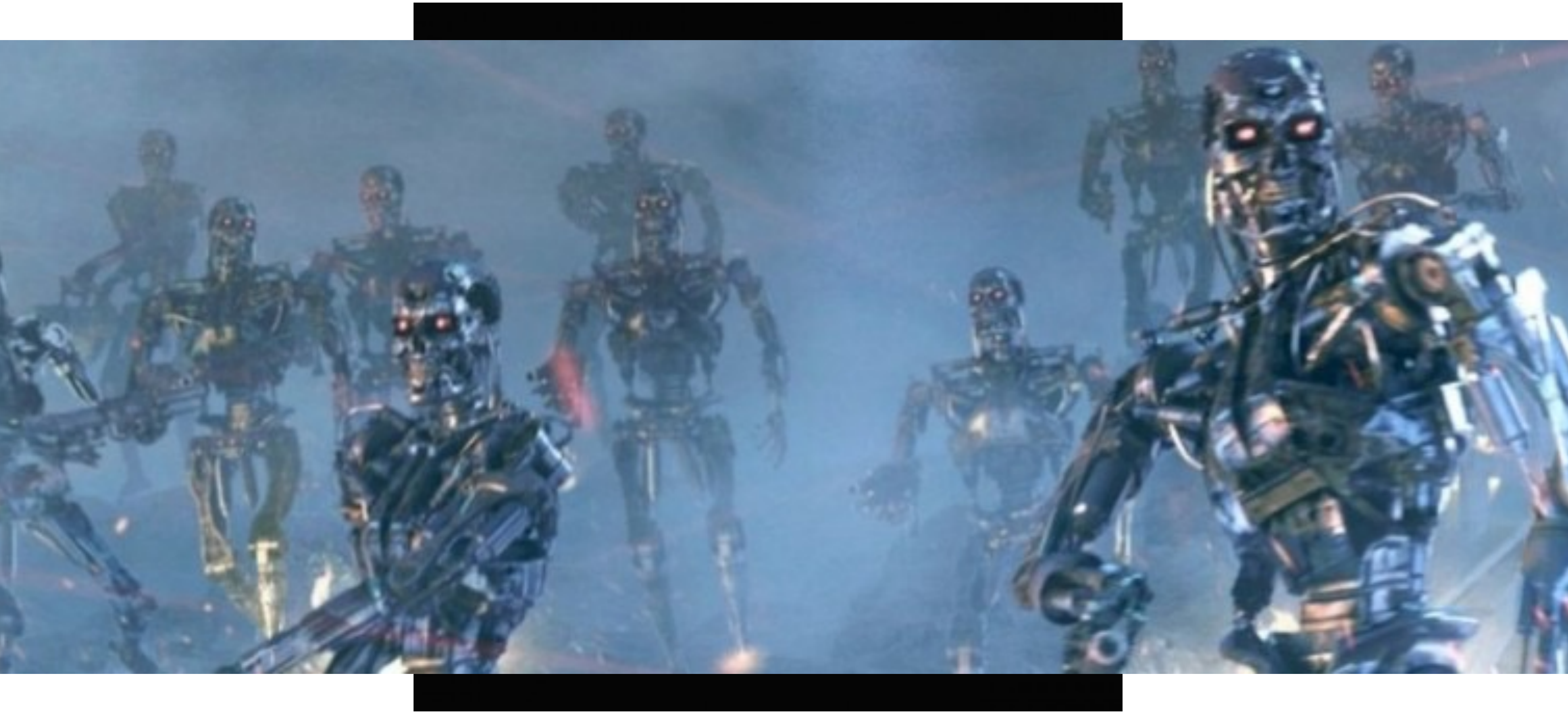
AlphaGo is the Skynet?

SKYNET

NEURAL NET-BASED ARTIFICIAL INTELLIGENCE

CYBERYNE SYSTEMS CORPORATION

AlphaGo is the Skynet?



1994

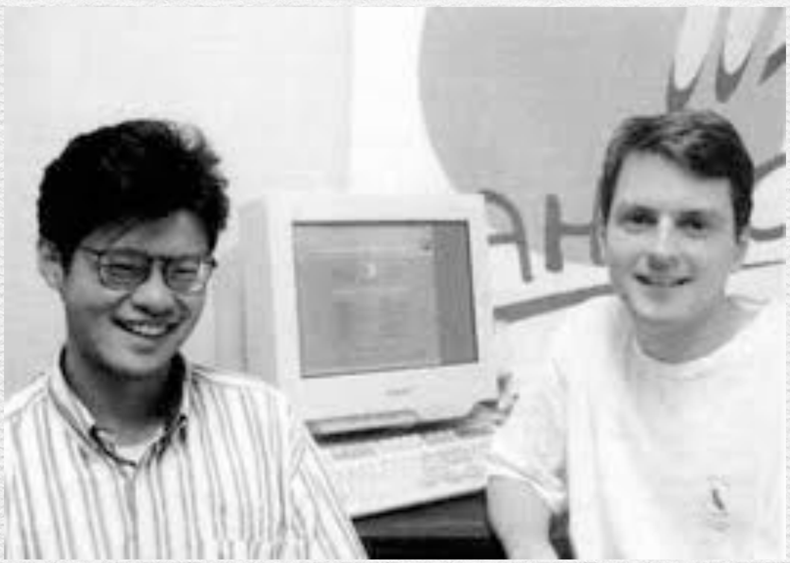
Data

Webpages (URLs)

Problem

Searching Webpage

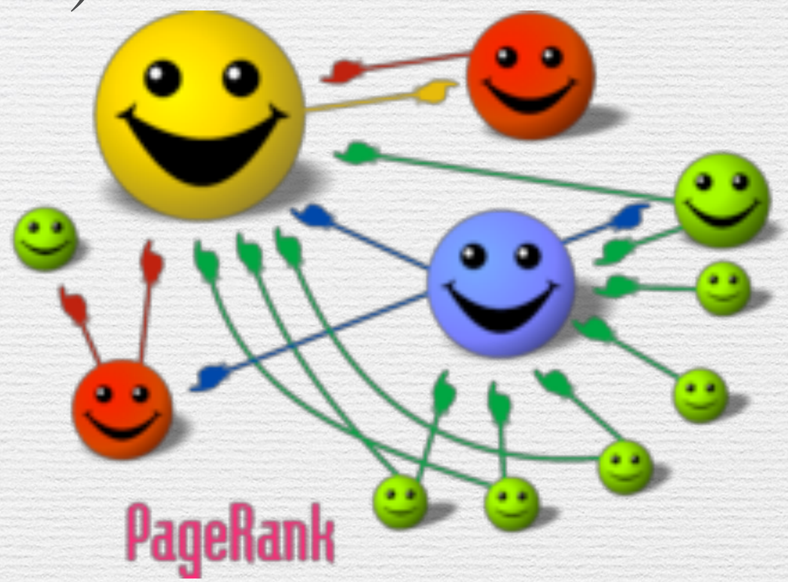
Jerry and David's Guide to the World Wide Web



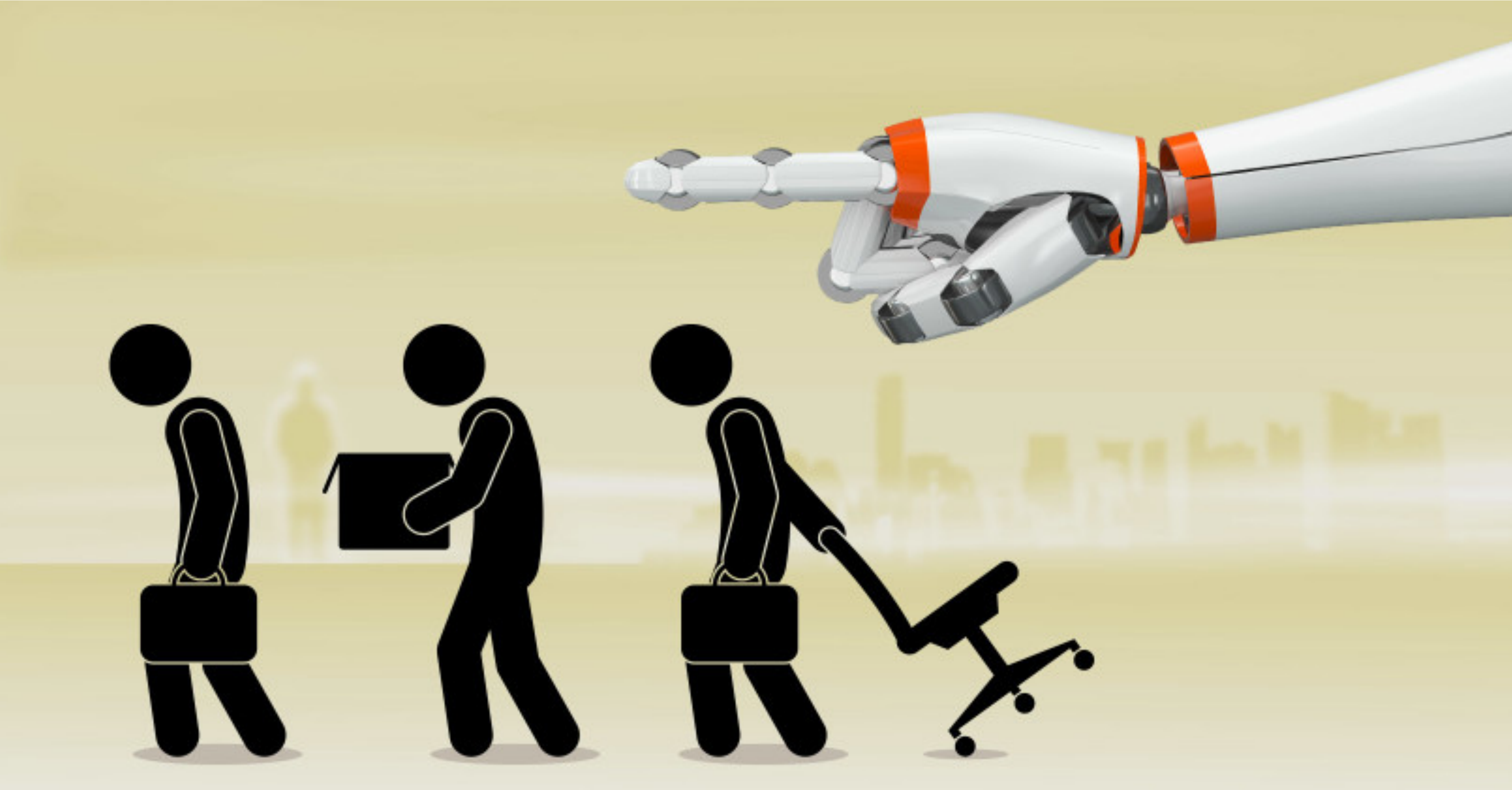
Jerry Yang and David Filo

1997

Data	Webpages (URLs)
Problem	Find URLs
Idea: Algorithm	PageRank



Larry Page and Sergey Brin

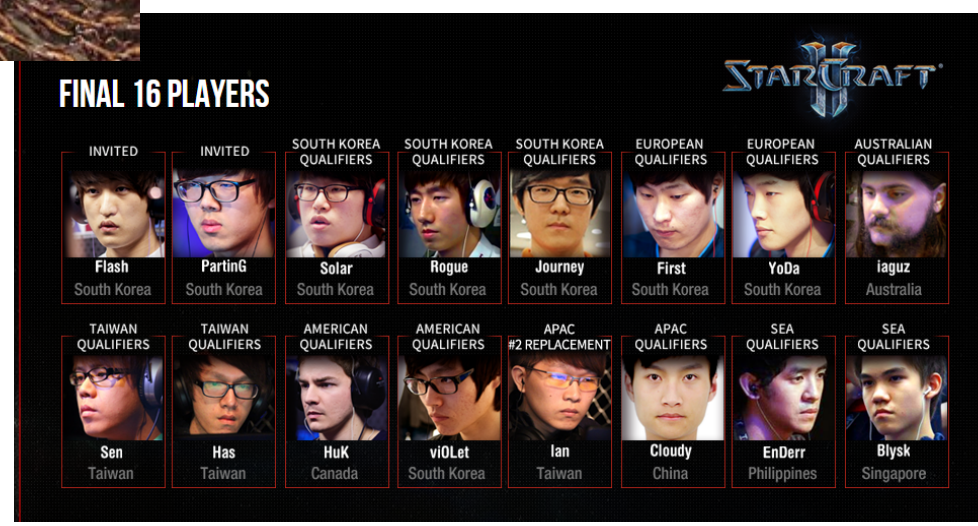


Google
DeepMind

Statement of Financial Position
As at 31 December 2017

	Notes	2017 £
Fixed assets		
Tangible assets	9	1,133,220
Intangible assets	10	1,157,140
Financial assets	11	765,095
		3,055,455
Current assets		
Debtors: amounts falling due within one year	12	99,173,584
Cash and cash equivalents		975,277
		100,148,861
Current Liabilities		
Creditors: amounts falling due within one year	13	(123,939,156)
Net current liabilities		(23,790,295)
Total assets less current liabilities		(20,734,840)
Creditors: amounts falling due after one year	14	(435,318,927)
Net liabilities		(456,053,767)
Capital and reserves		
Called up share capital	15	1,641
Share premium		41,490,727
Retained loss		(497,546,135)
Total shareholders' deficit		(456,053,767)

How about Other Games?

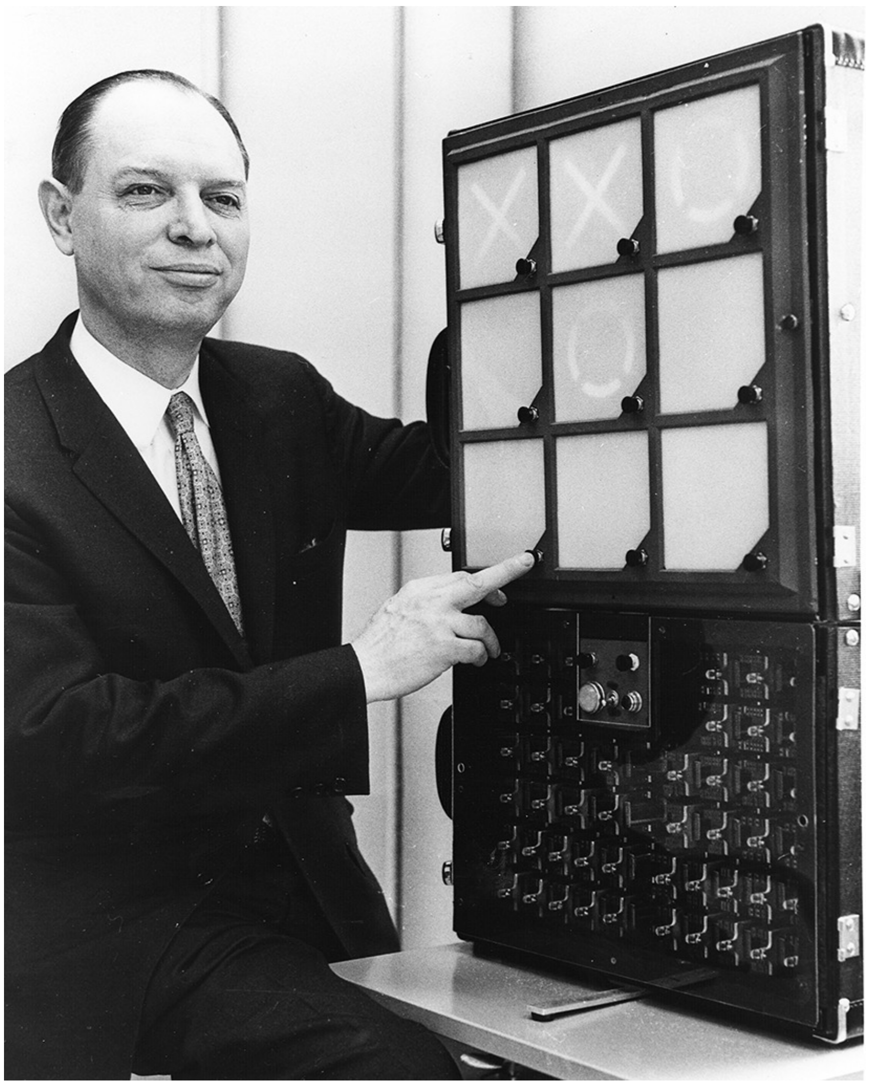


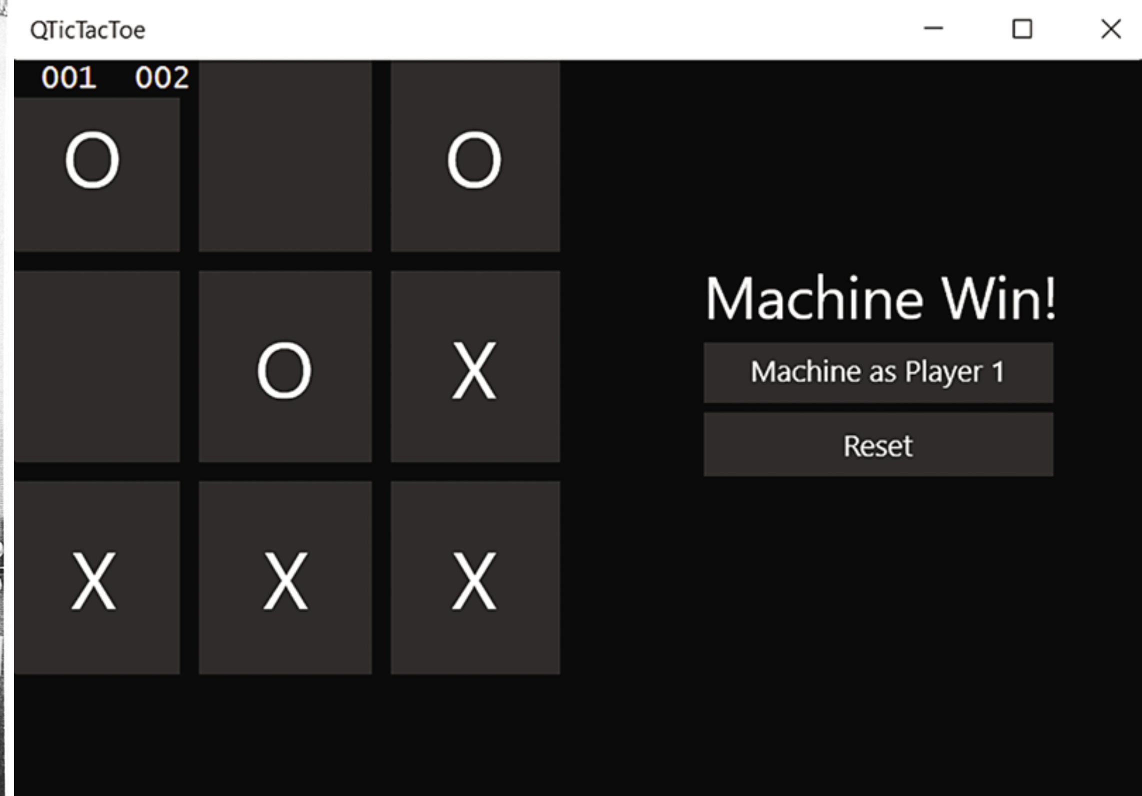
Aug, 2017

https://youtu.be/7U4-wvhgx0w?t=2m10s

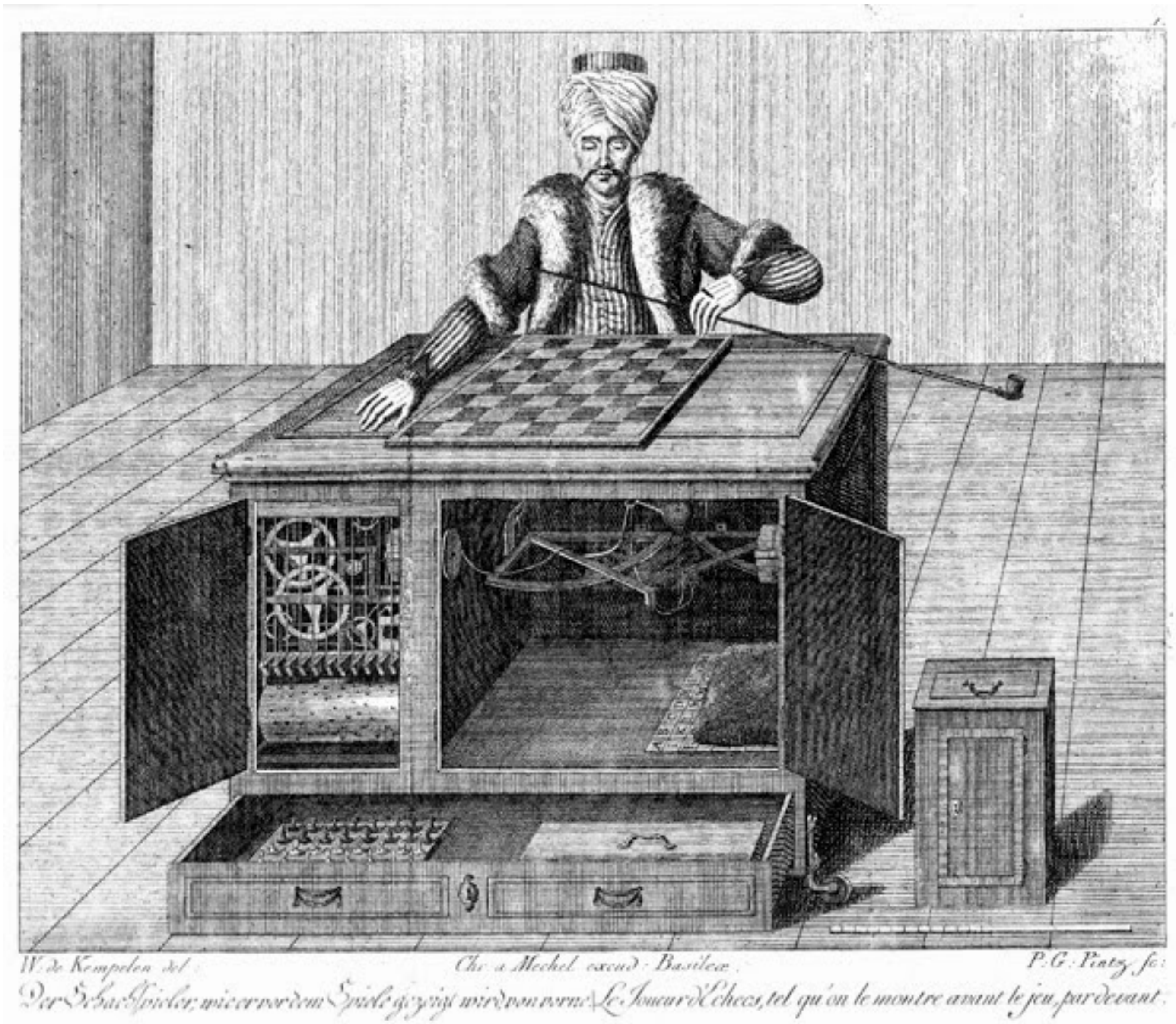
Quiz

Tic Tac Toe




QTicTacToe
001 002
O
O
O
X
X
X
X
Machine Win!
Machine as Player 1
Reset

Chess


W. de Kempelen del.
Chr. a Mechel excud. Basileæ.
P. G. Pintz sc.
Der Schachspieler, wie er vor dem Spiele gezeigt wird, von vorne. Le Joueur d'Echecs, tel qu'on le montre avant le jeu, par devant.

Chess (1996)

Deep Blue (1996)

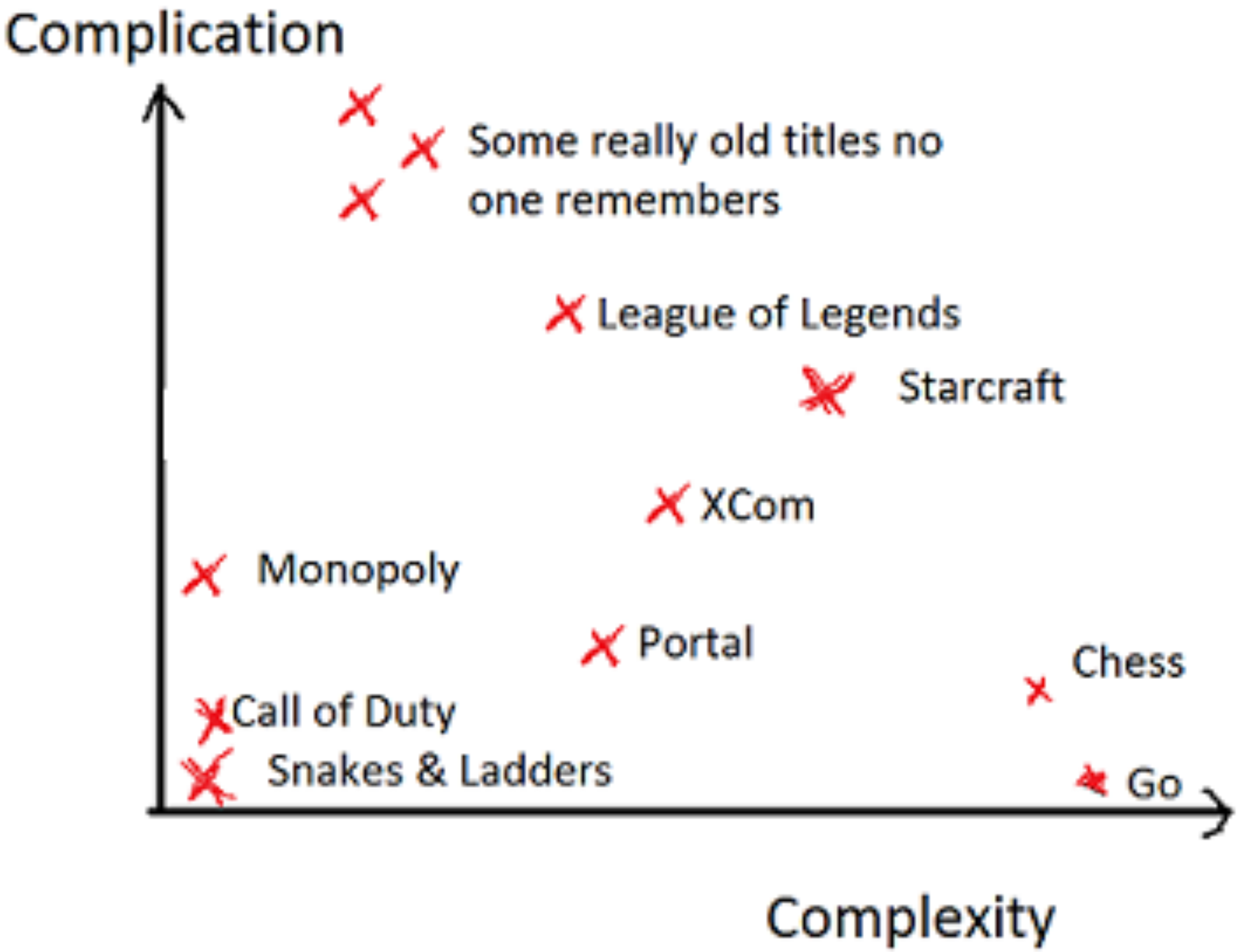
The 1996 match

Game #	White	Black	Result	Comment
1	**Deep Blue**	Kasparov	1-0	
2	**Kasparov**	Deep Blue	1-0	
3	Deep Blue	Kasparov	½–½	Draw by mutual agreement
4	Kasparov	Deep Blue	½–½	Draw by mutual agreement
5	Deep Blue	**Kasparov**	0–1	Kasparov offered a draw after the 23rd move.
6	**Kasparov**	Deep Blue	1–0	
Result: **Kasparov–Deep Blue: 4–2**				

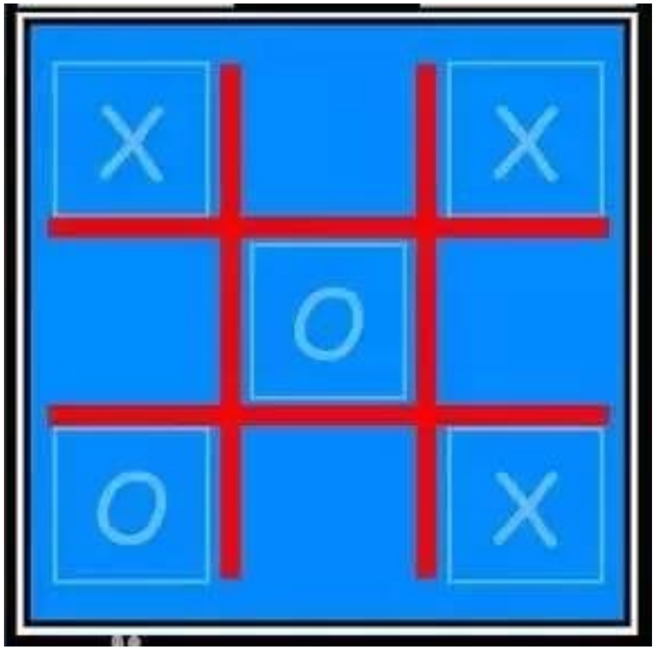
The 1997 rematch

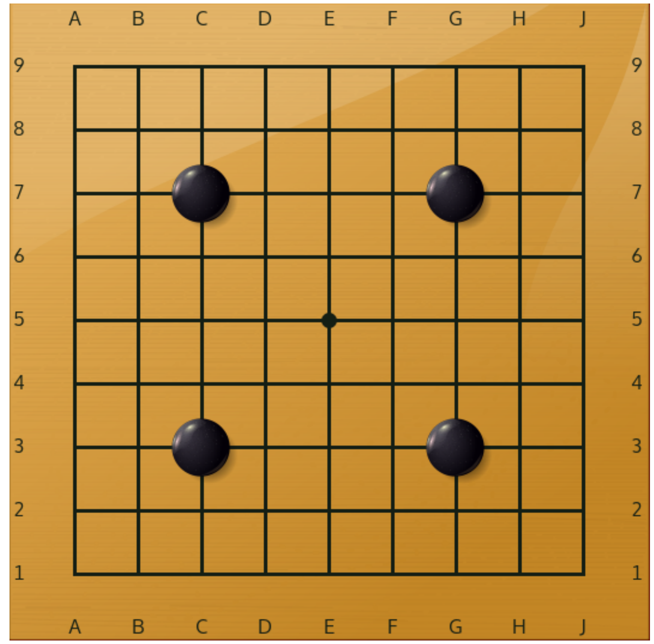
Game #	White	Black	Result	Comment
1	**Kasparov**	Deep Blue	1–0	
2	**Deep Blue**	Kasparov	1–0	
3	Kasparov	Deep Blue	½–½	Draw by mutual agreement
4	Deep Blue	Kasparov	½–½	Draw by mutual agreement
5	Kasparov	Deep Blue	½–½	Draw by mutual agreement
6	**Deep Blue**	Kasparov	1–0	
Result: **Deep Blue–Kasparov: 3½–2½**				

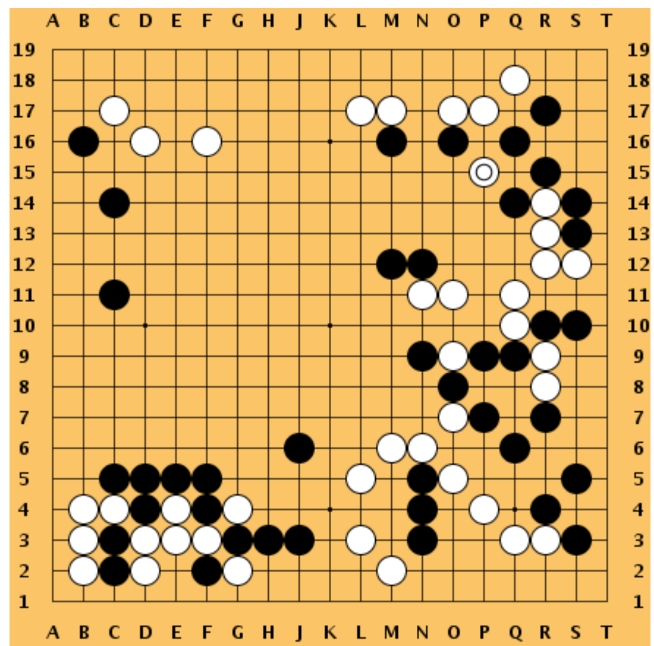
High Complexity



Different Games

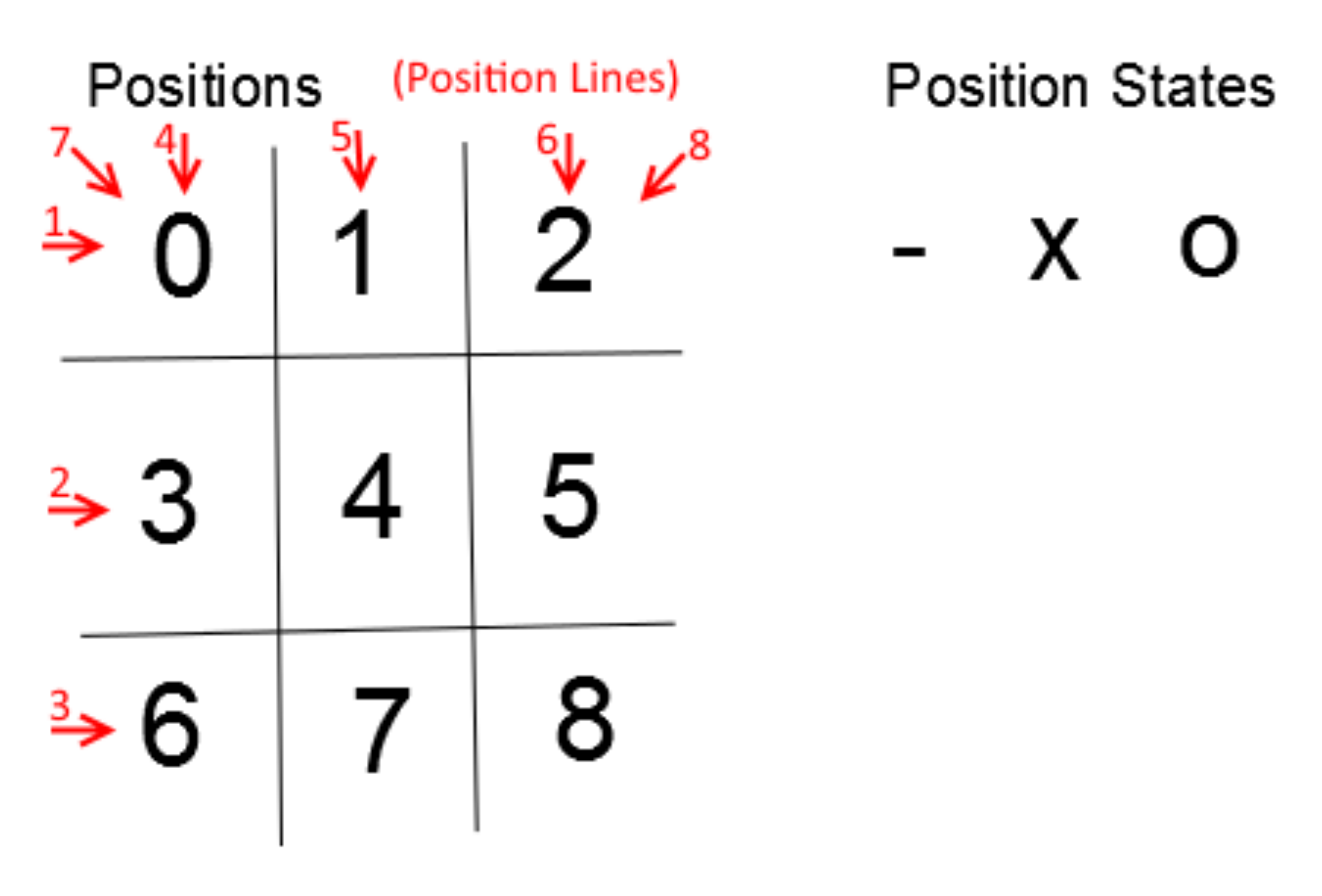


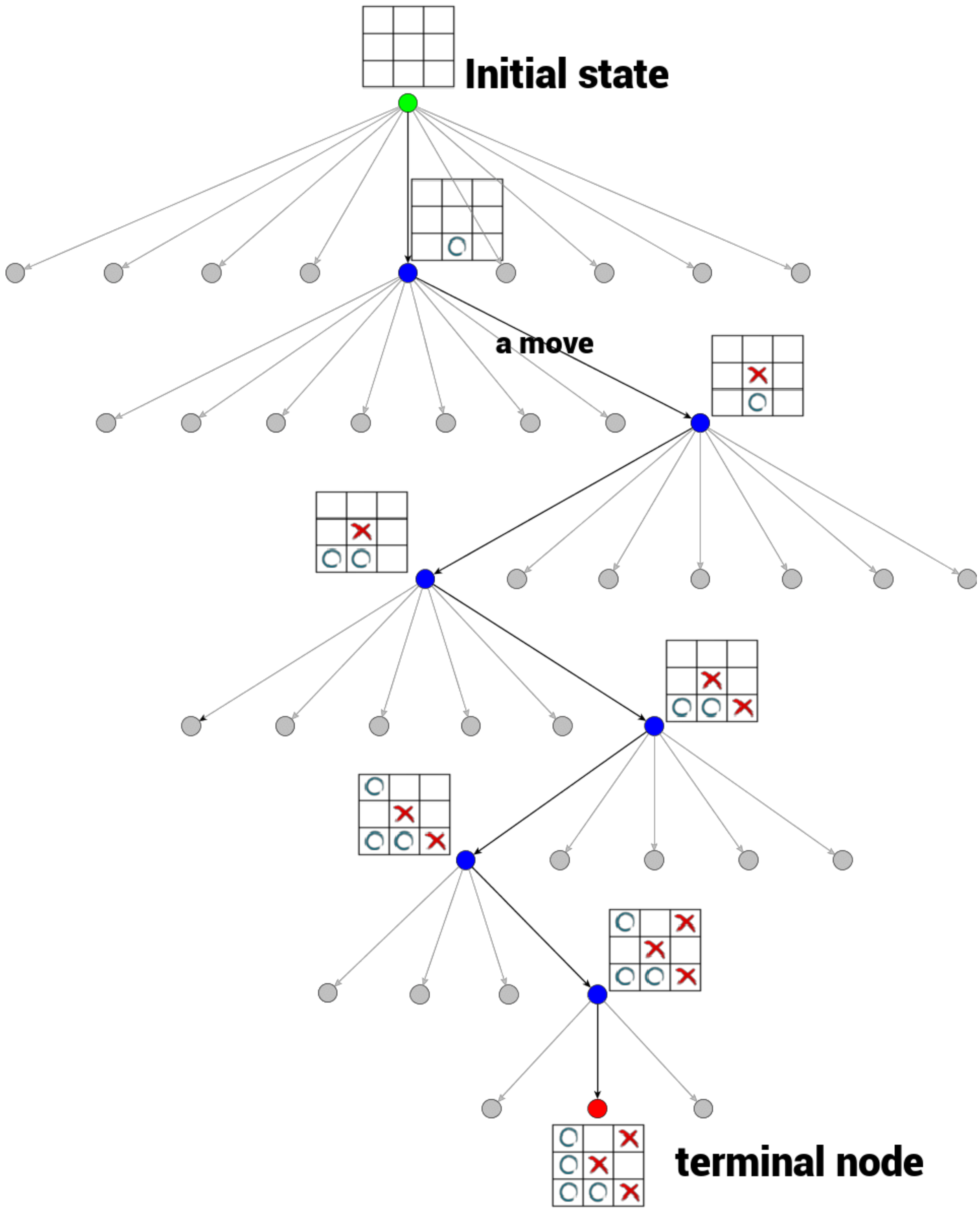




Search Problem
(the search space)

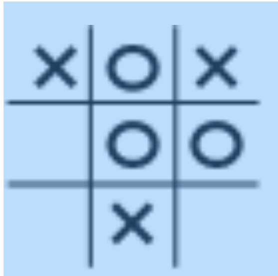
Game State (Tic Tac Toe)




Initial state
a move
terminal node

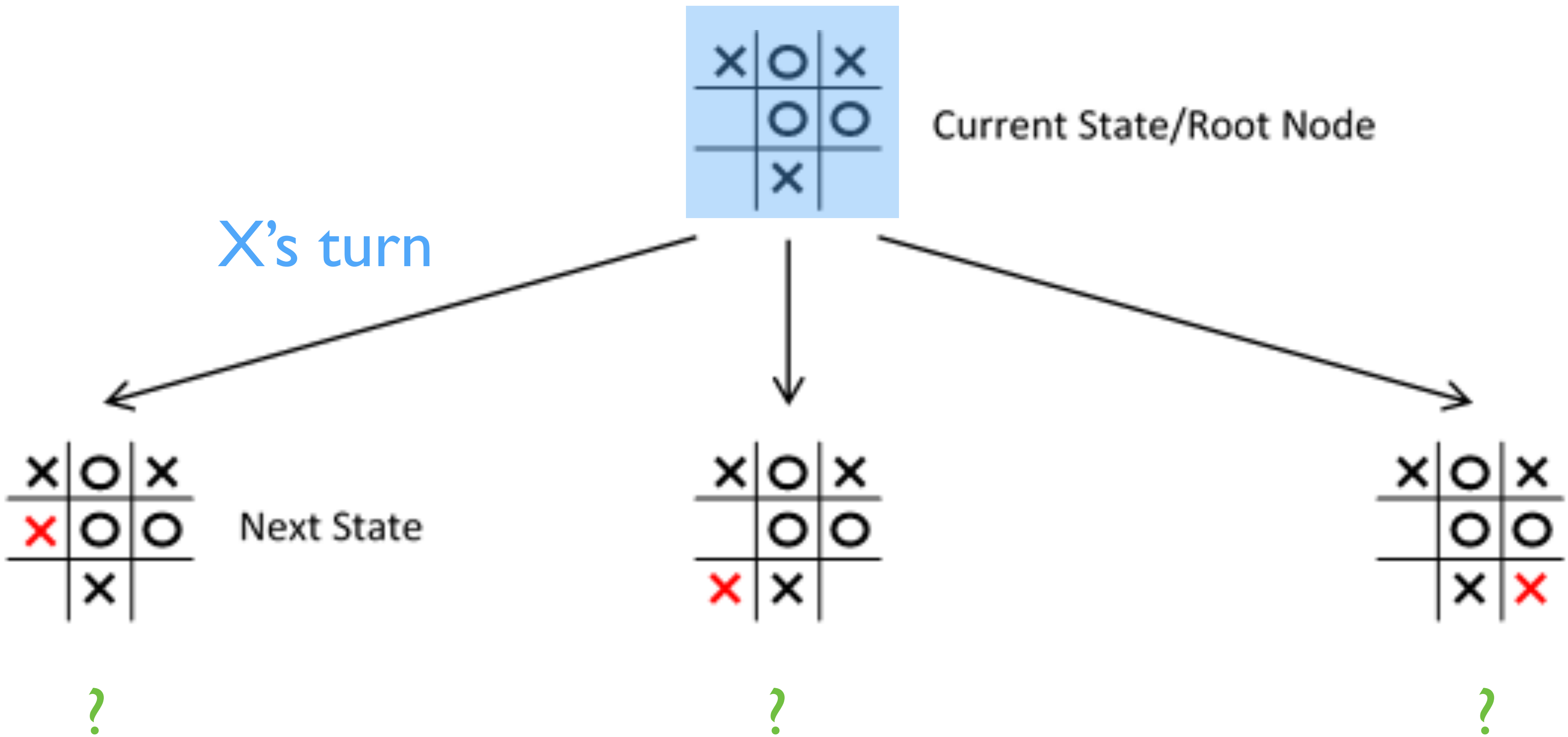
Tic Tac Toe

X’s turn

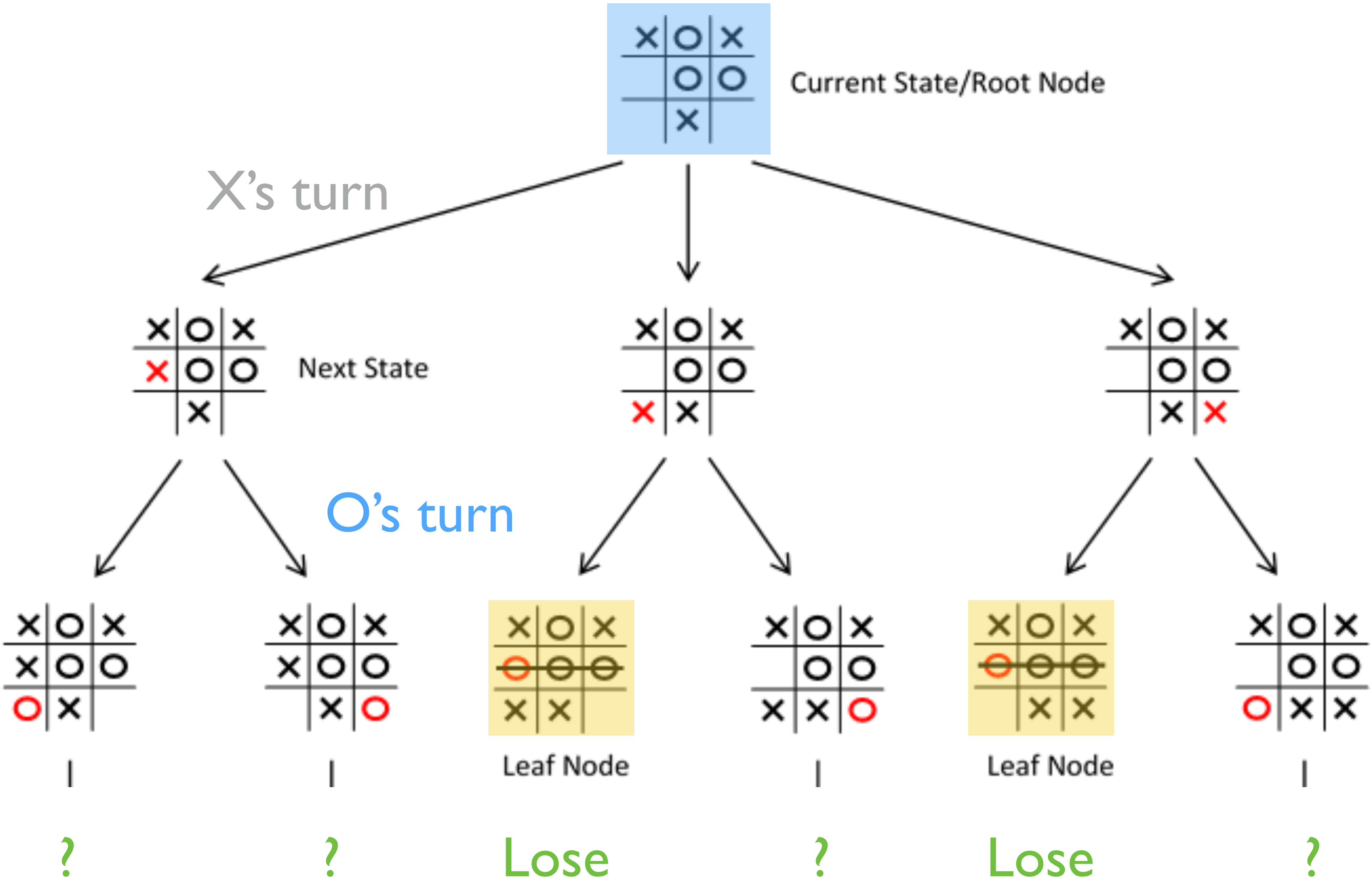


Current State

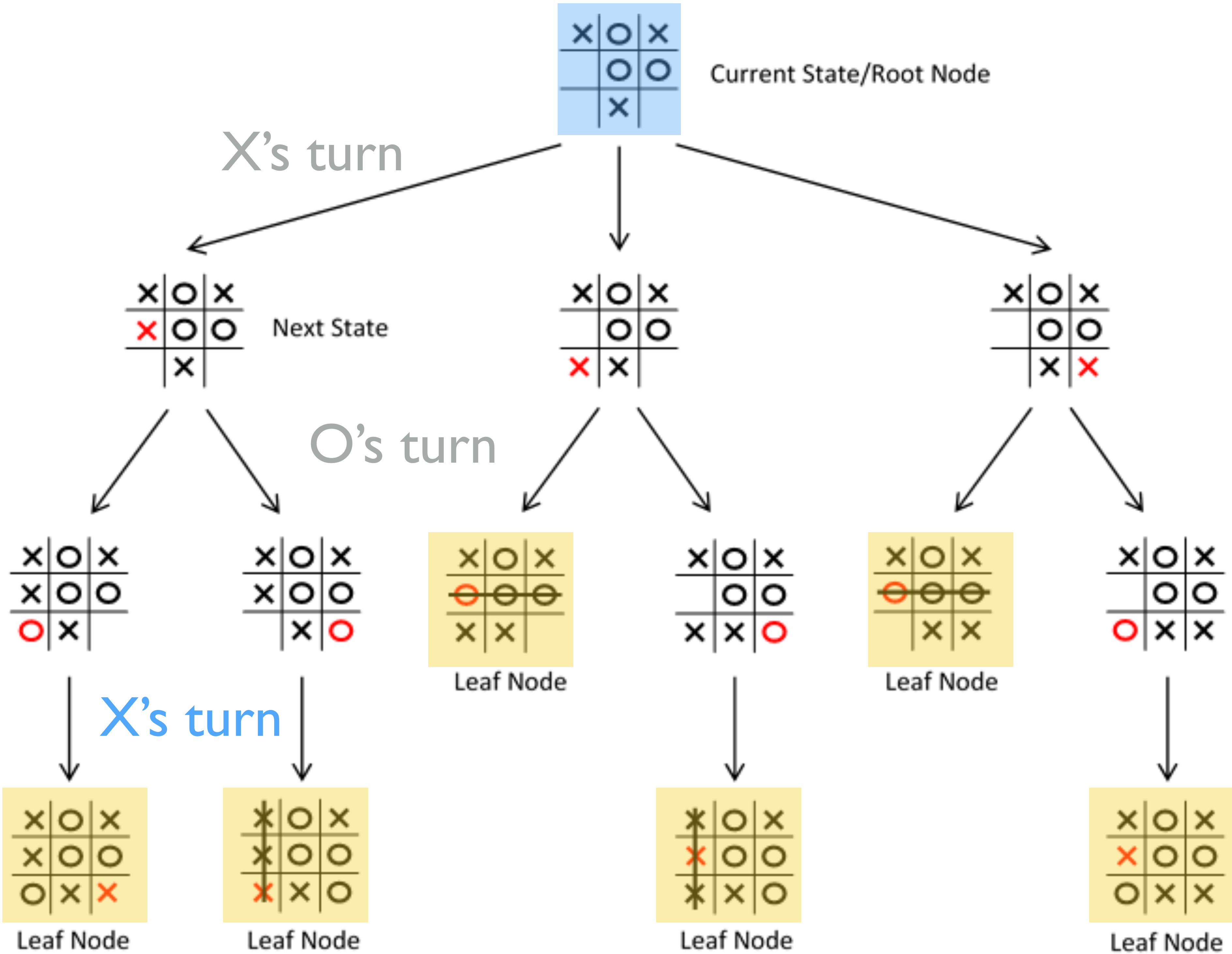
Tic Tac Toe



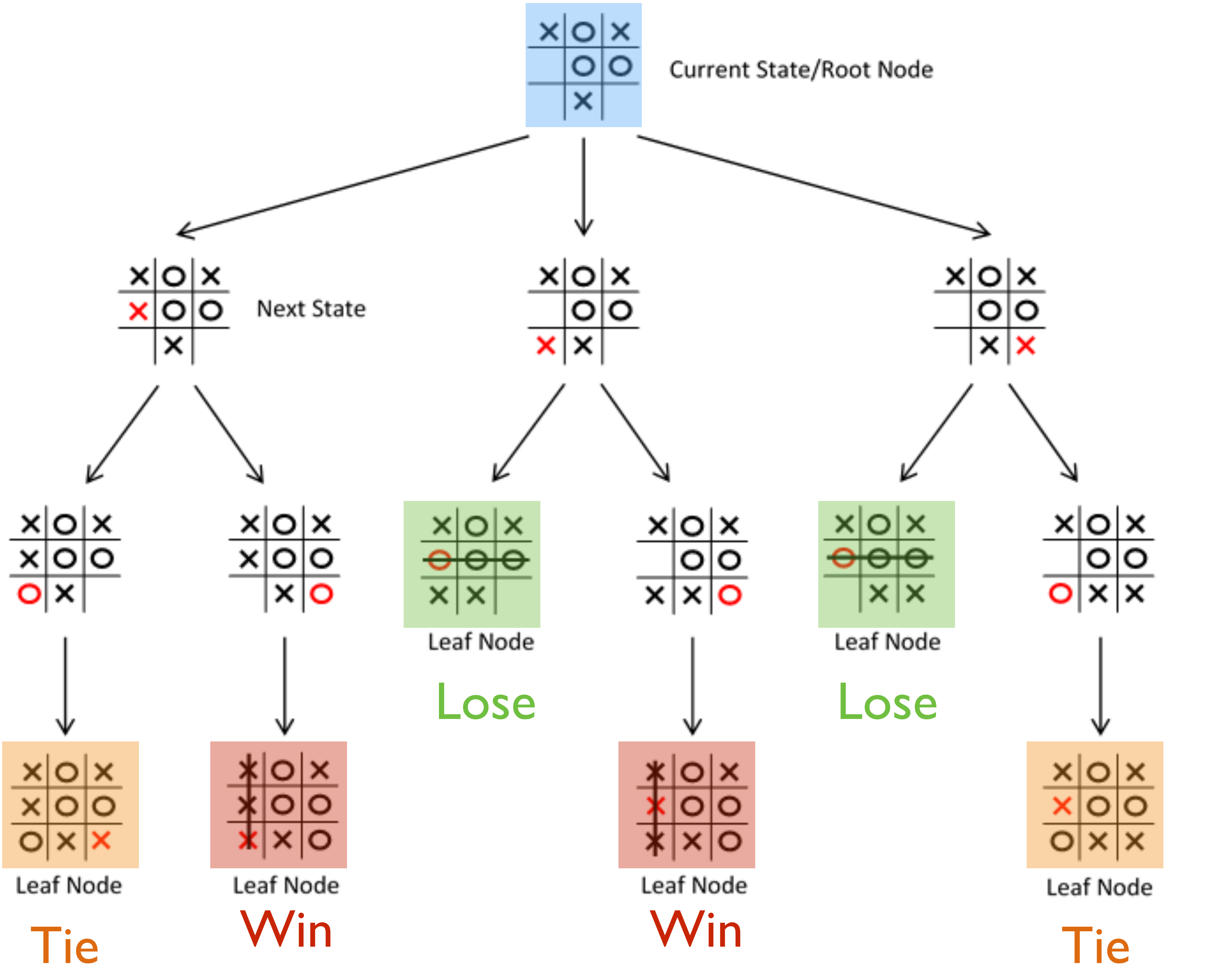
Tic Tac Toe



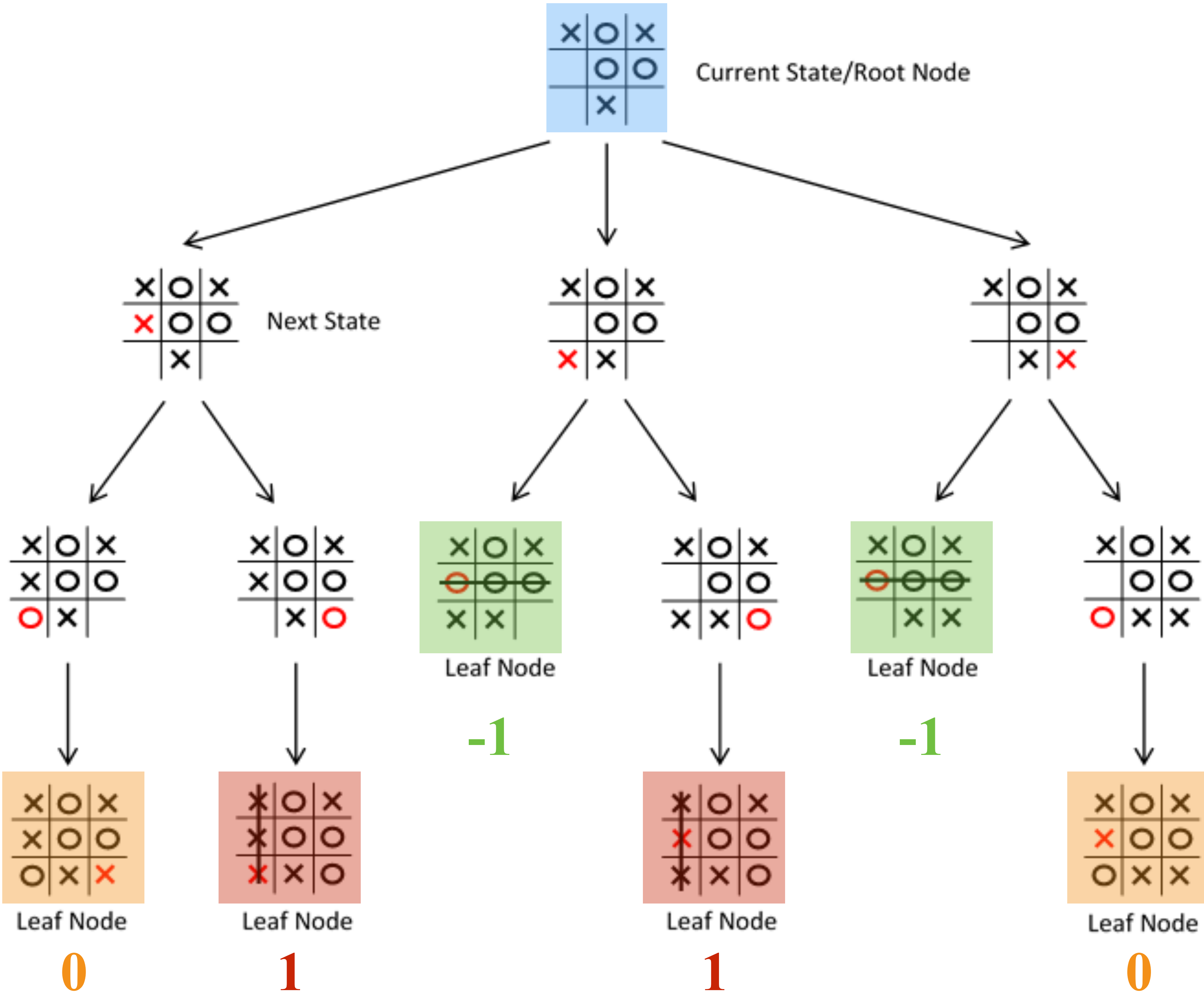
Current State/Root Node
X's turn
Next State
O's turn
Leaf Node
Leaf Node
?
?
Lose
?
Lose
?

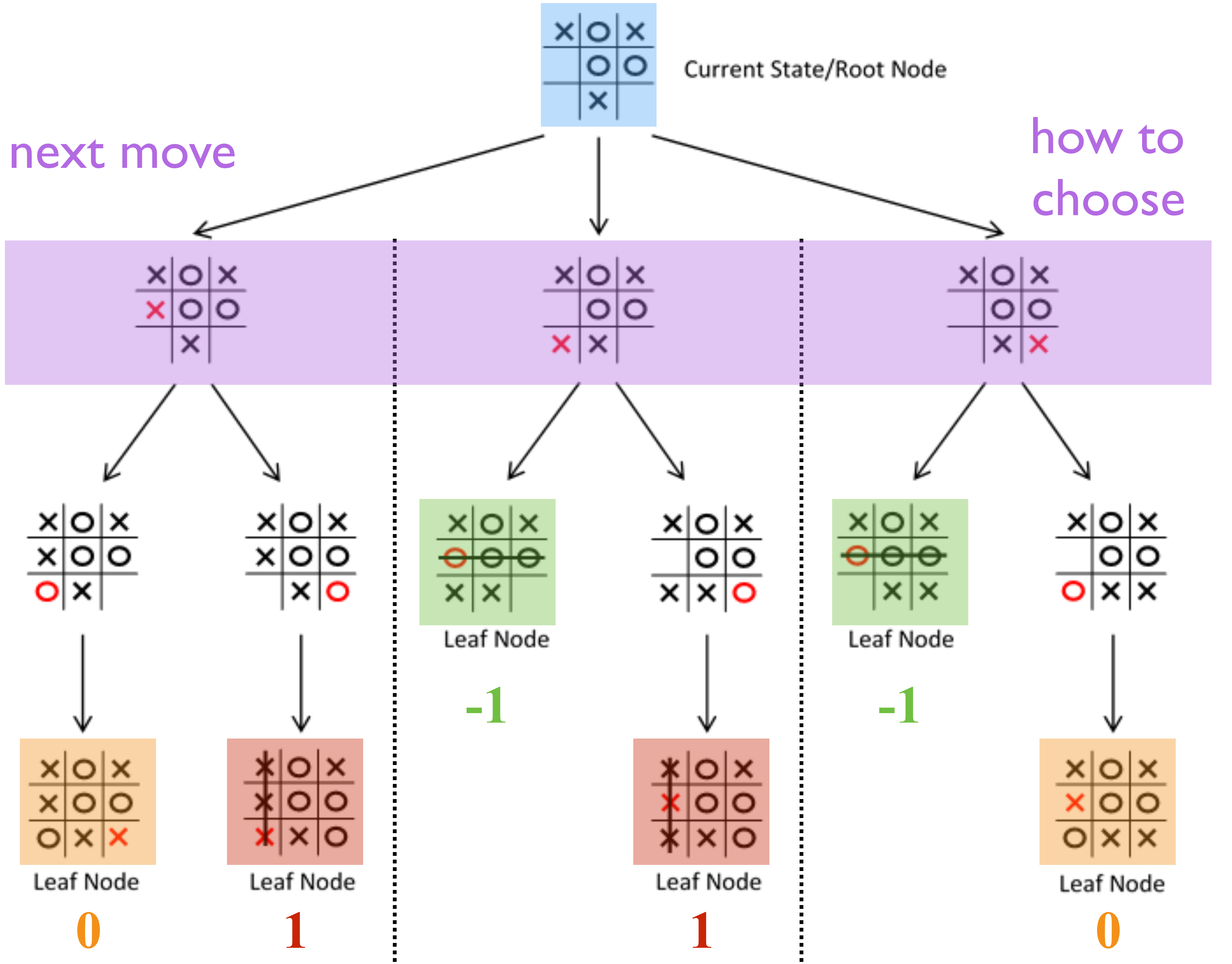


Current State/Root Node
X's turn
Next State
O's turn
X's turn
Leaf Node
Leaf Node
Leaf Node
Leaf Node
Leaf Node
Leaf Node

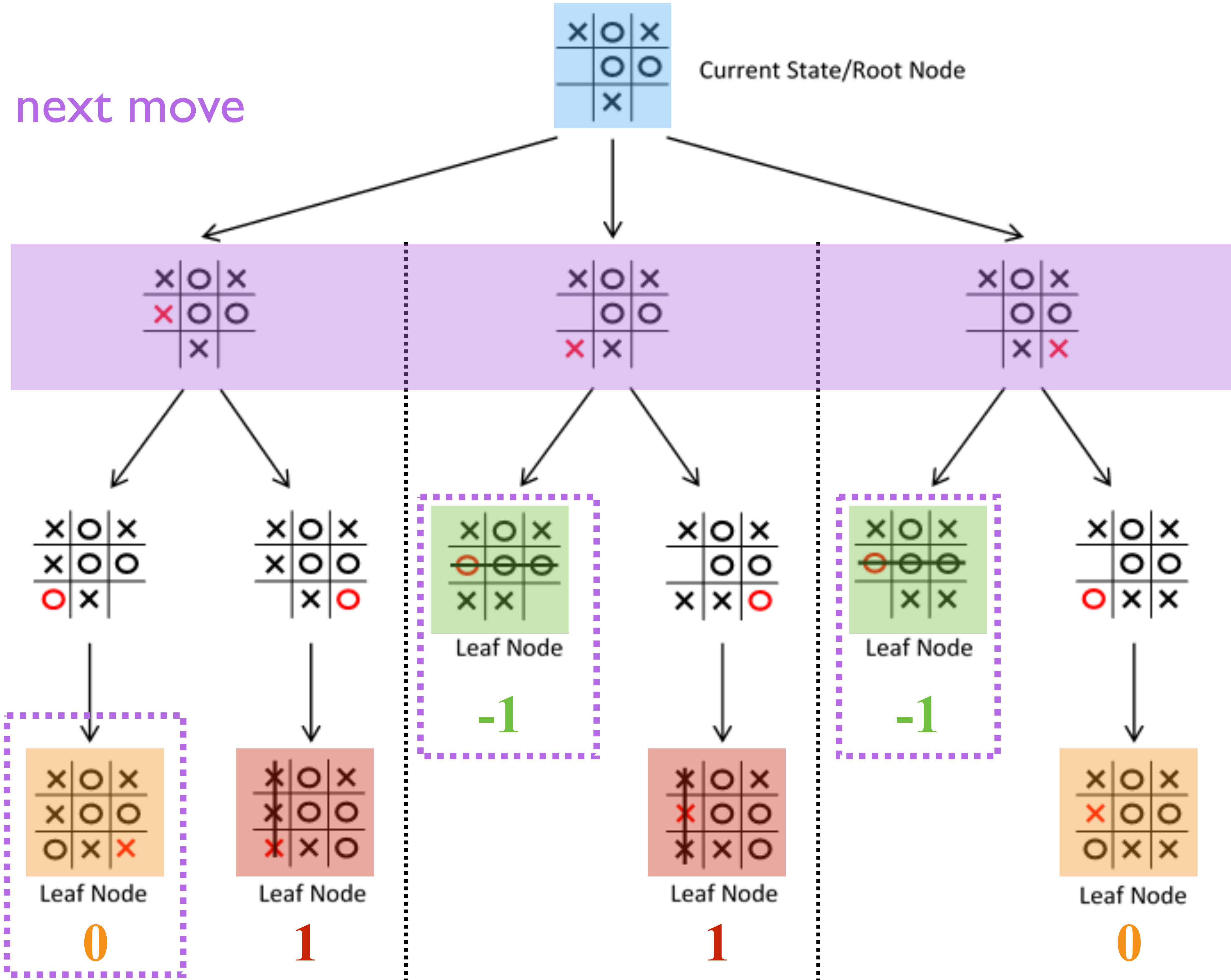


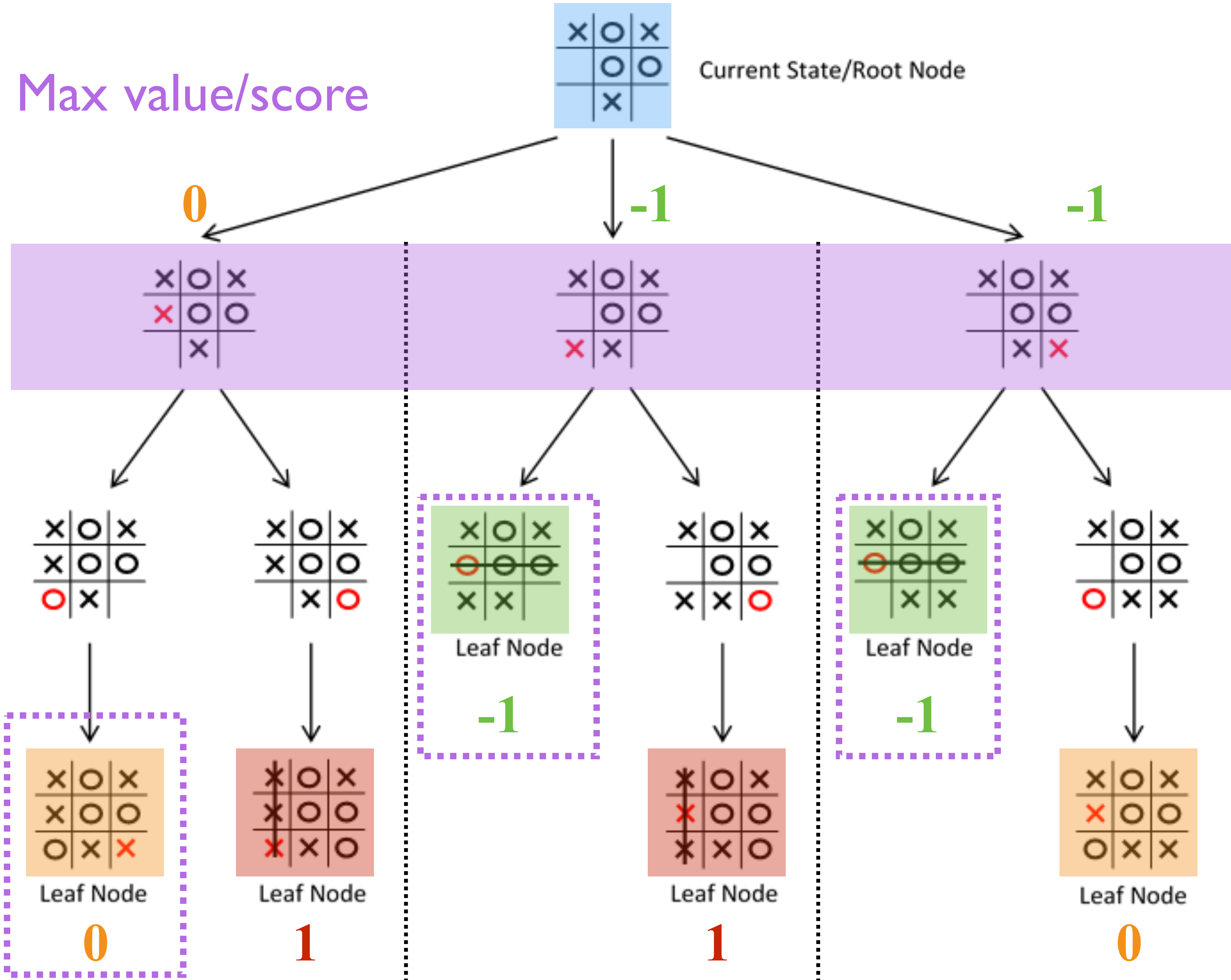
Current State/Root Node
Next State
Leaf Node
Lose
Leaf Node
Lose
Leaf Node
Tie
Leaf Node
Win
Leaf Node
Win
Leaf Node
Tie


Current State/Root Node
Next State
Leaf Node
-1
Leaf Node
-1
Leaf Node
0
Leaf Node
1
Leaf Node
1
Leaf Node
0

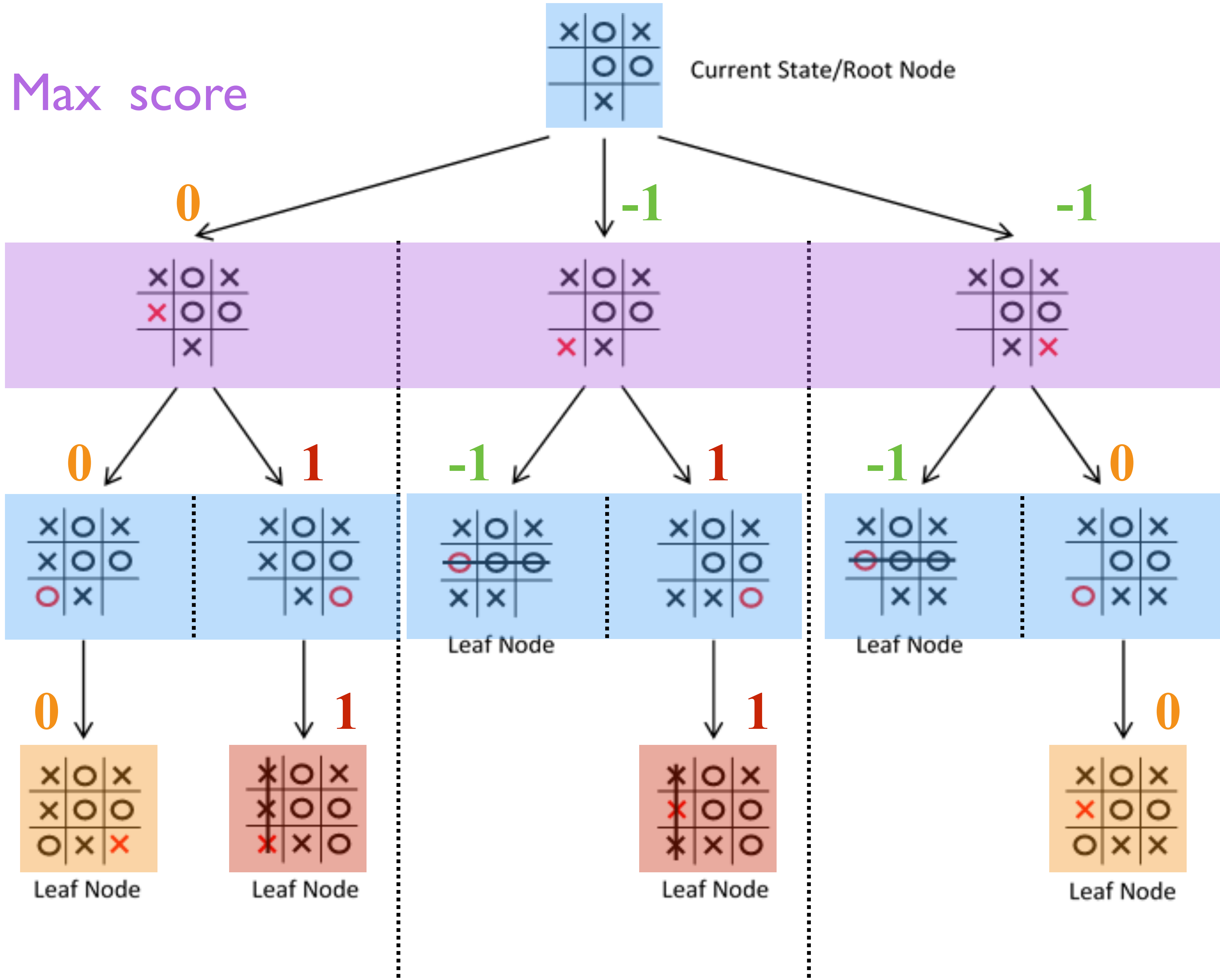


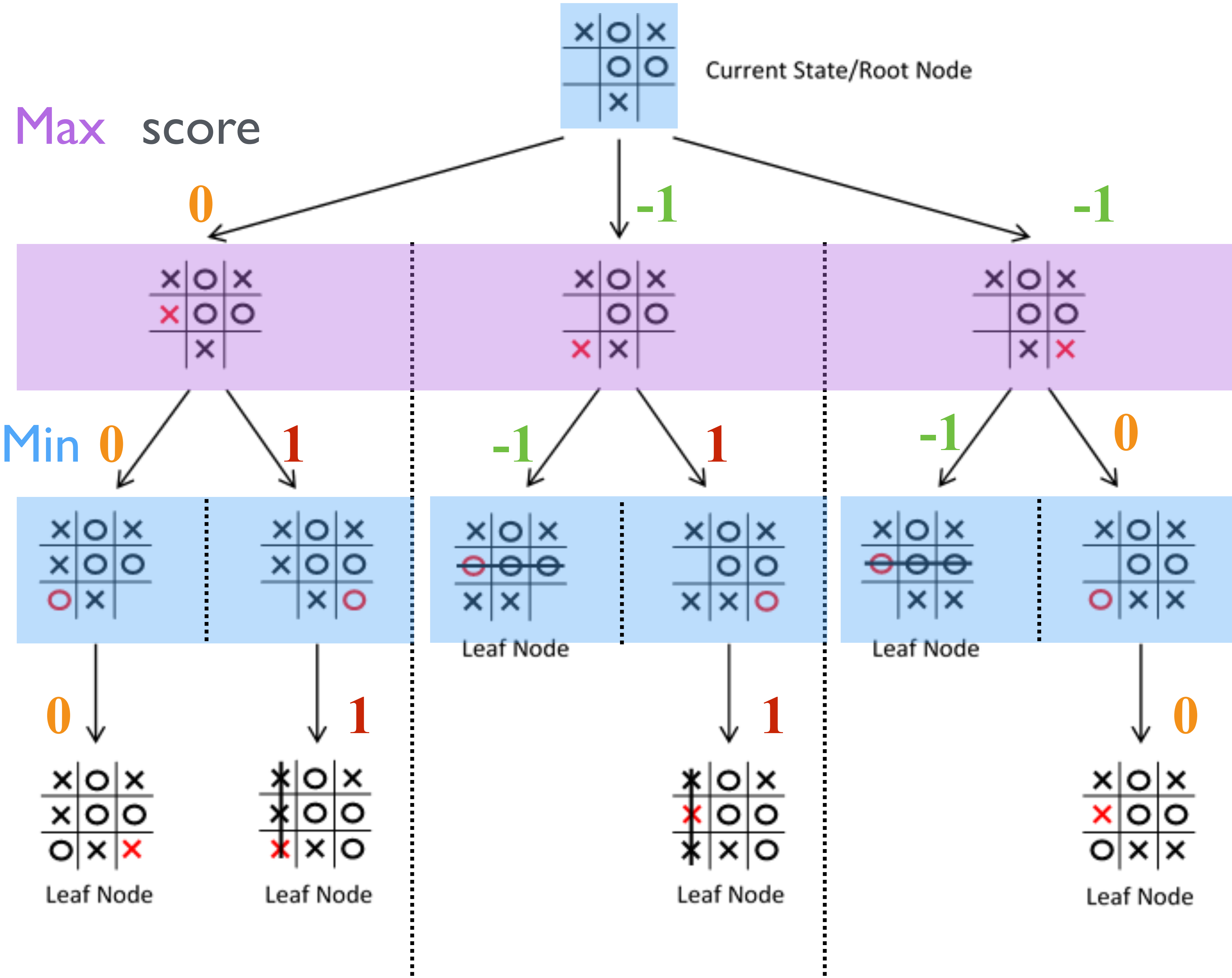
Current State/Root Node
next move
how to choose
Leaf Node
-1
Leaf Node
-1
Leaf Node
0
Leaf Node
1
Leaf Node
1
Leaf Node
0

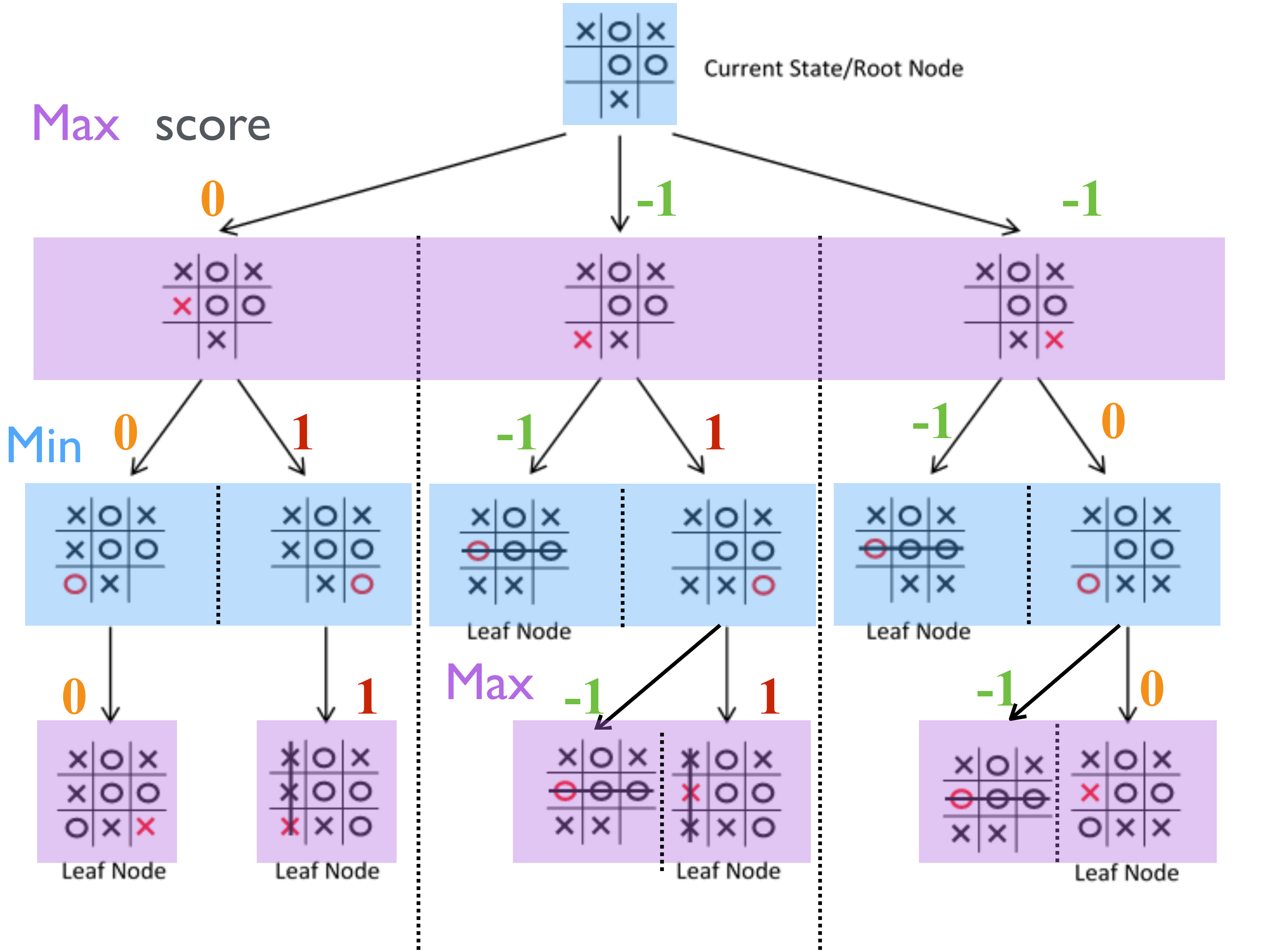

next move
Current State/Root Node
Leaf Node
0
Leaf Node
1
Leaf Node
-1
Leaf Node
1
Leaf Node
-1
Leaf Node
0



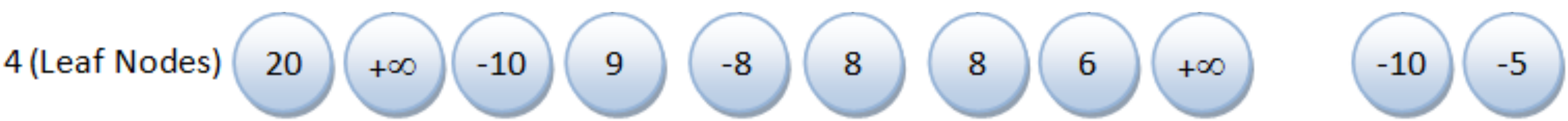
Max value/score
Current State/Root Node
0
-1
-1
Leaf Node
-1
Leaf Node
-1
Leaf Node
0
Leaf Node
1
Leaf Node
1
Leaf Node
0


Max score
Current State/Root Node
0
-1
-1
0
1
-1
1
-1
0
Leaf Node
Leaf Node
0
1
1
0
Leaf Node
Leaf Node
Leaf Node
Leaf Node

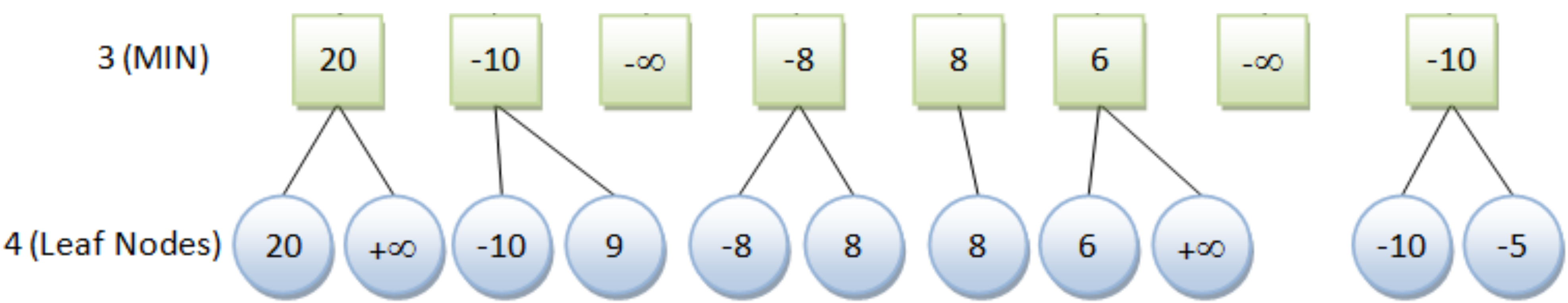

Current State/Root Node
Max score
0
-1
-1
Min 0
1
-1
1
-1
0
Leaf Node
Leaf Node
0
1
1
0
Leaf Node
Leaf Node
Leaf Node
Leaf Node


Current State/Root Node
Max score
0
-1
-1
Min
0
1
-1
1
-1
0
Leaf Node
Leaf Node
Max
0
1
-1
1
-1
0
Leaf Node
Leaf Node
Leaf Node
Leaf Node

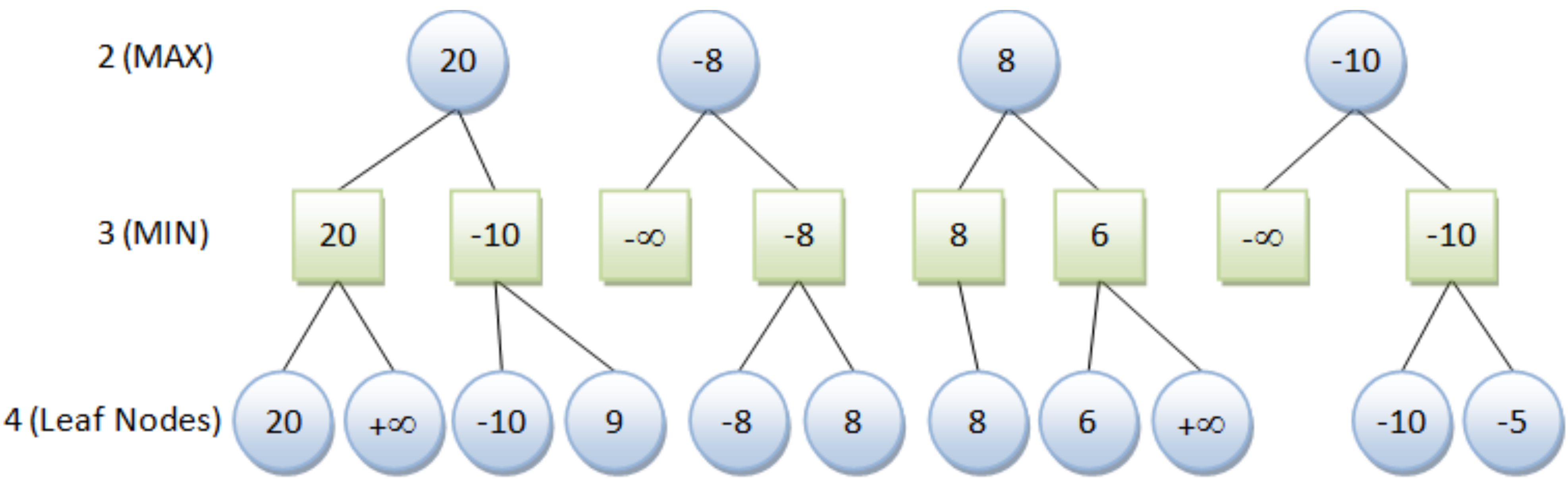
MiniMax



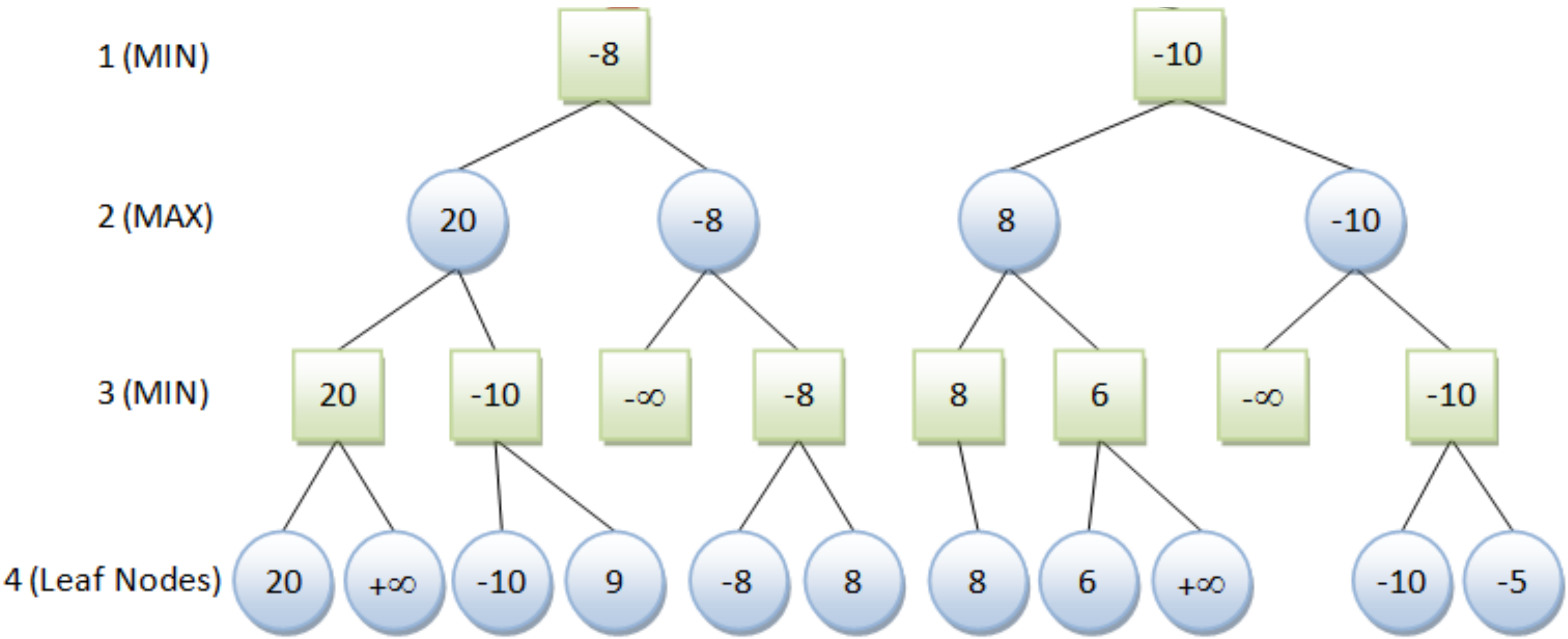
MiniMax



MiniMax

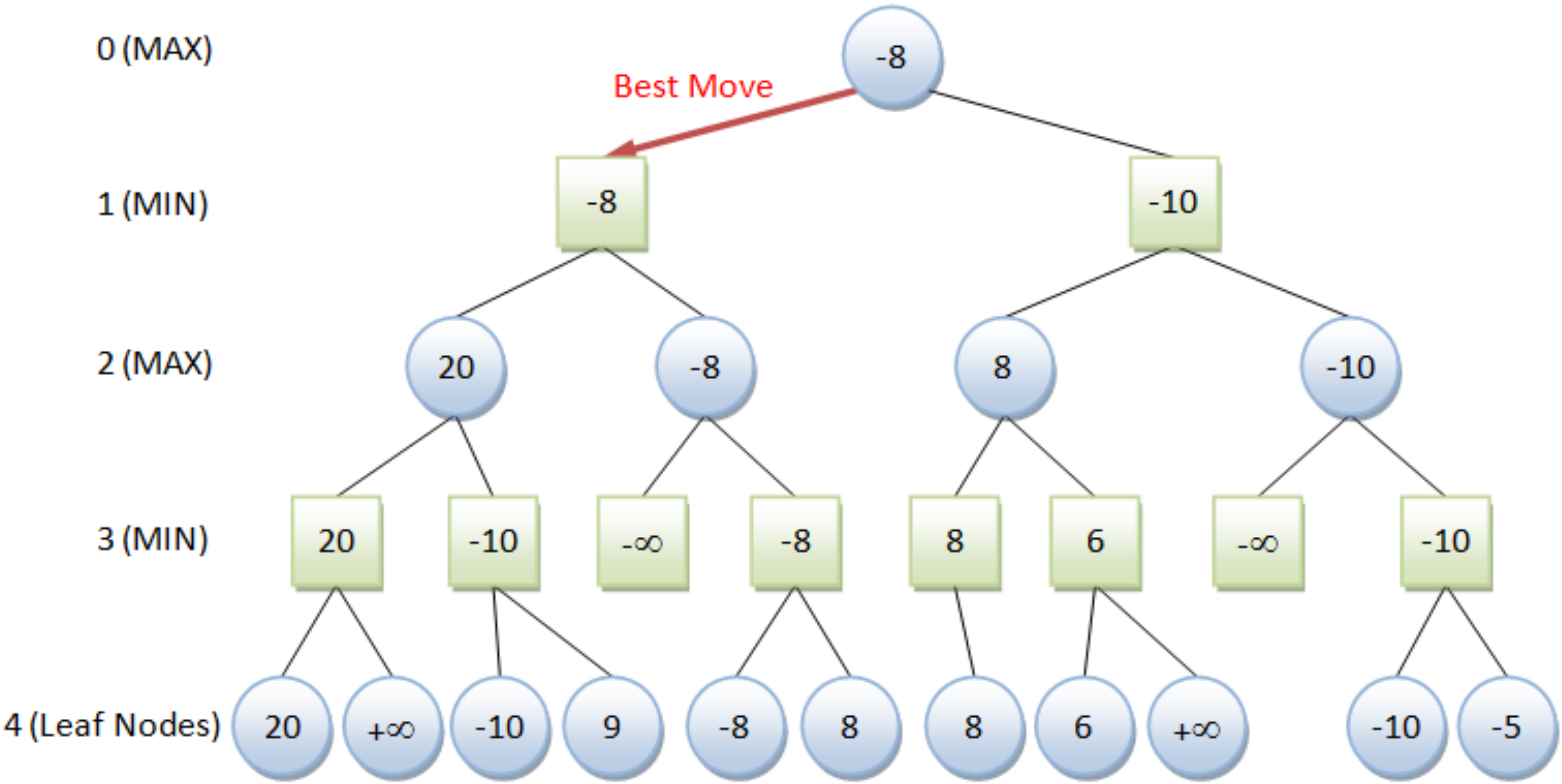


MiniMax



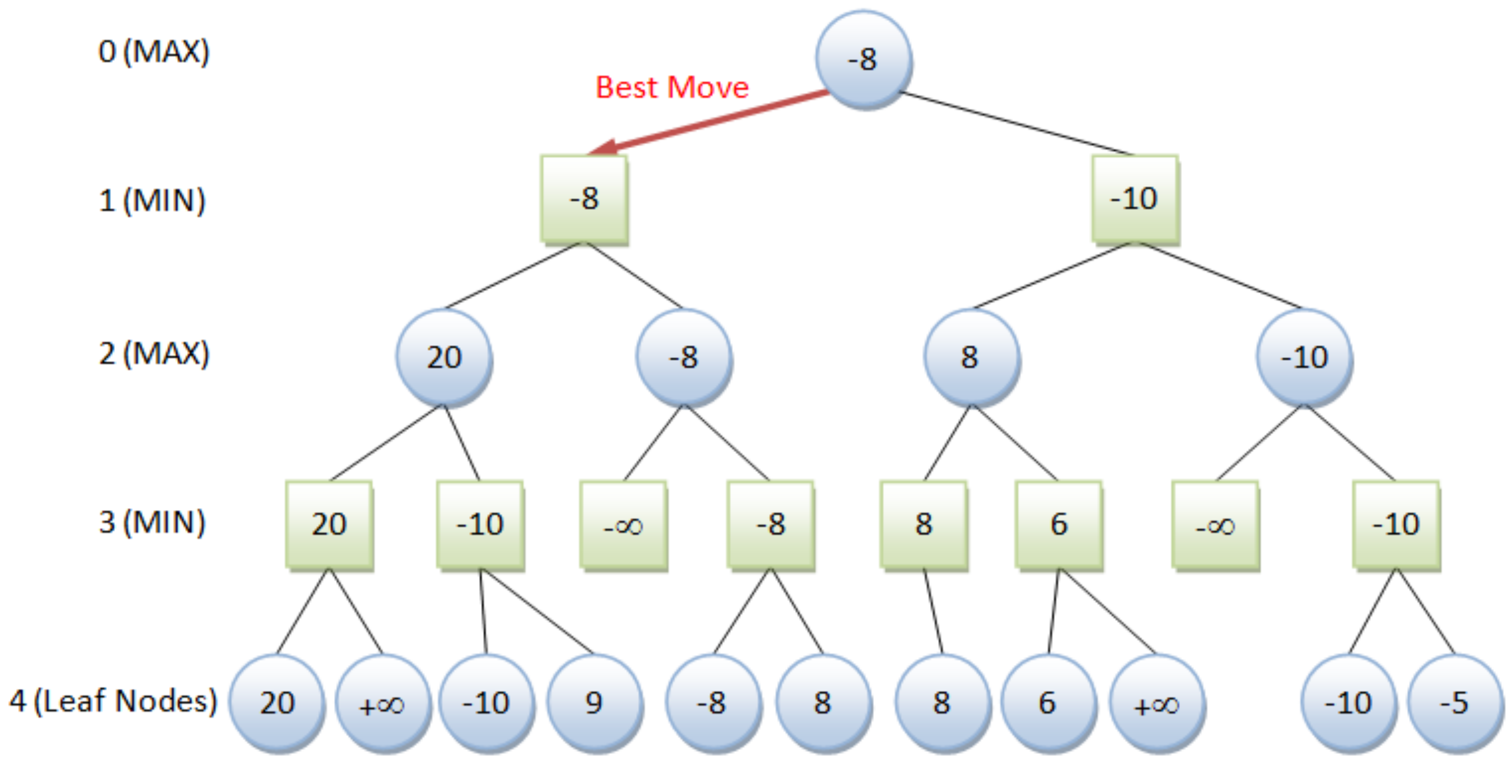
1 (MIN)
2 (MAX)
3 (MIN)
4 (Leaf Nodes)
-8
-10
20
-8
8
-10
20
-10
-∞
-8
8
6
-∞
-10
20
+∞
-10
9
-8
8
8
6
+∞
-10
-5

MiniMax

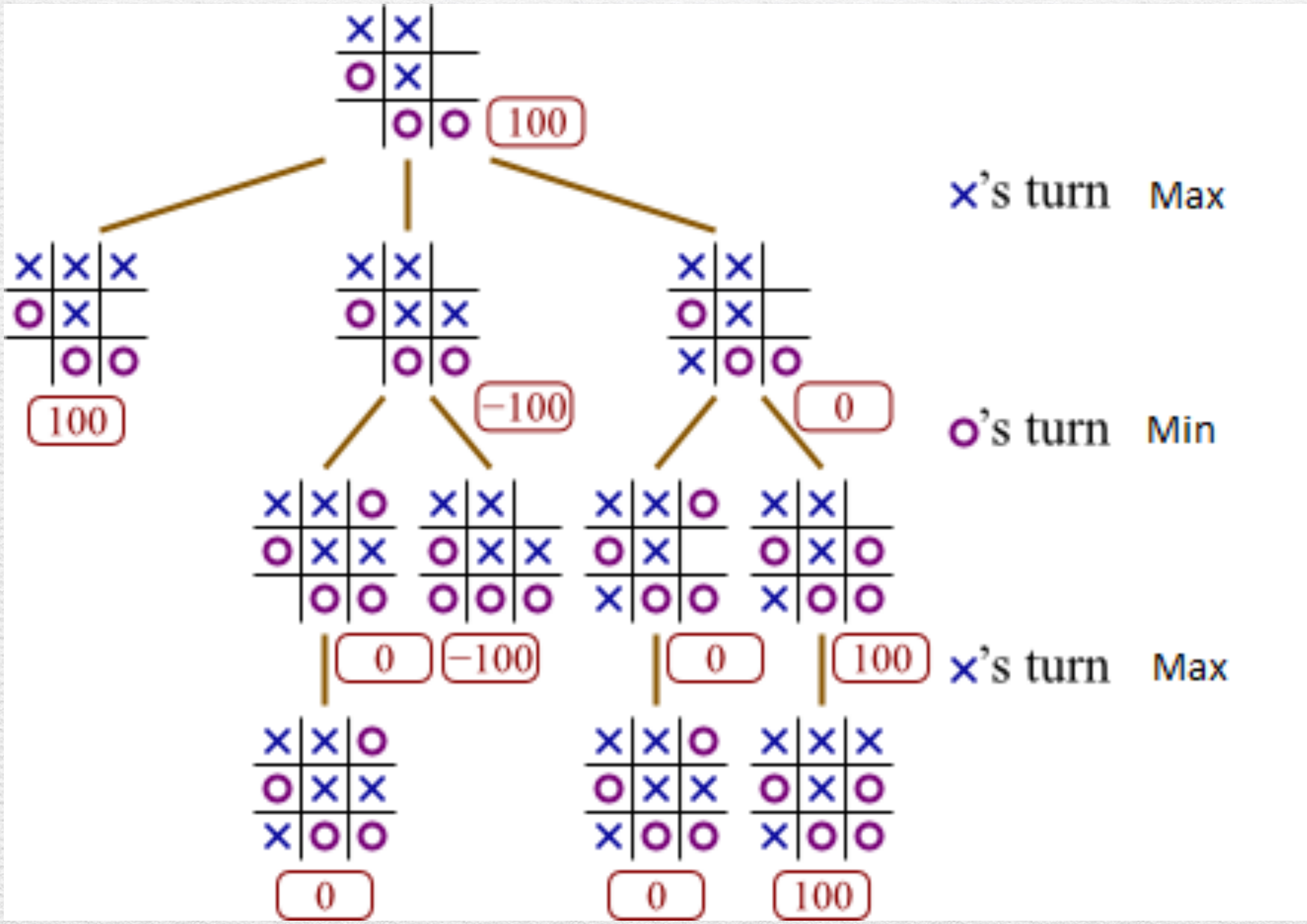

0 (MAX)
1 (MIN)
2 (MAX)
3 (MIN)
4 (Leaf Nodes)
Best Move
-8
-8
-10
20
-8
8
-10
20
-10
-∞
-8
8
6
-∞
-10
20
+∞
-10
9
-8
8
8
6
+∞
-10
-5

Practical Tips: MiniMax

	Max Case	Min Case
X Player		
O Player		

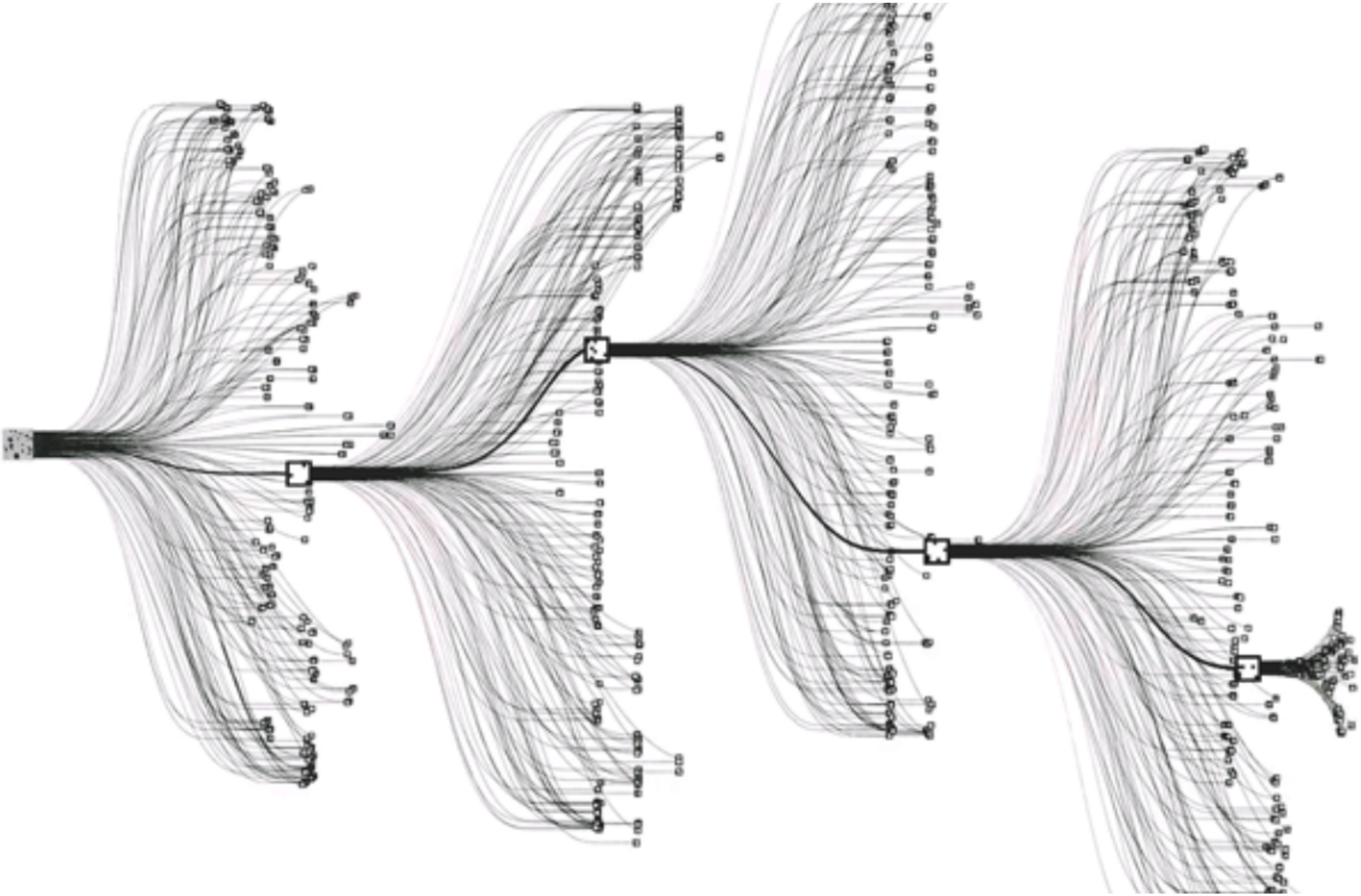


1. Generate the Search Tree
2. use MinMax Search



What is the problem?

The "Tree" of Go Game



The Size of the Tree

Tic Tac Toe:

$b = 9, d = 9$

Chess:

$b = 35, d = 80$

Go:

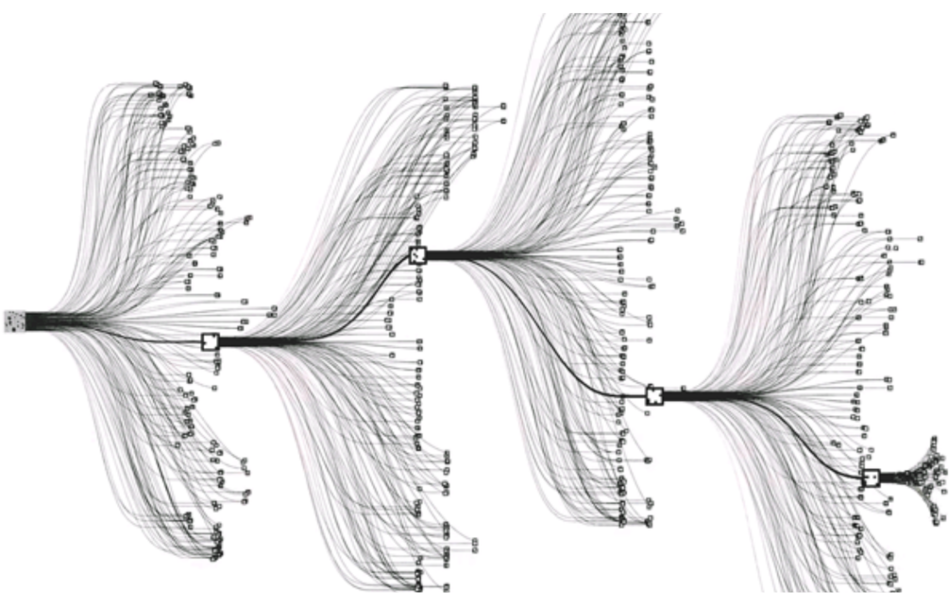
$b = 250, d = 150$

b : number of legal move per position

d : its depth (game length)

Size of the Tree in GO

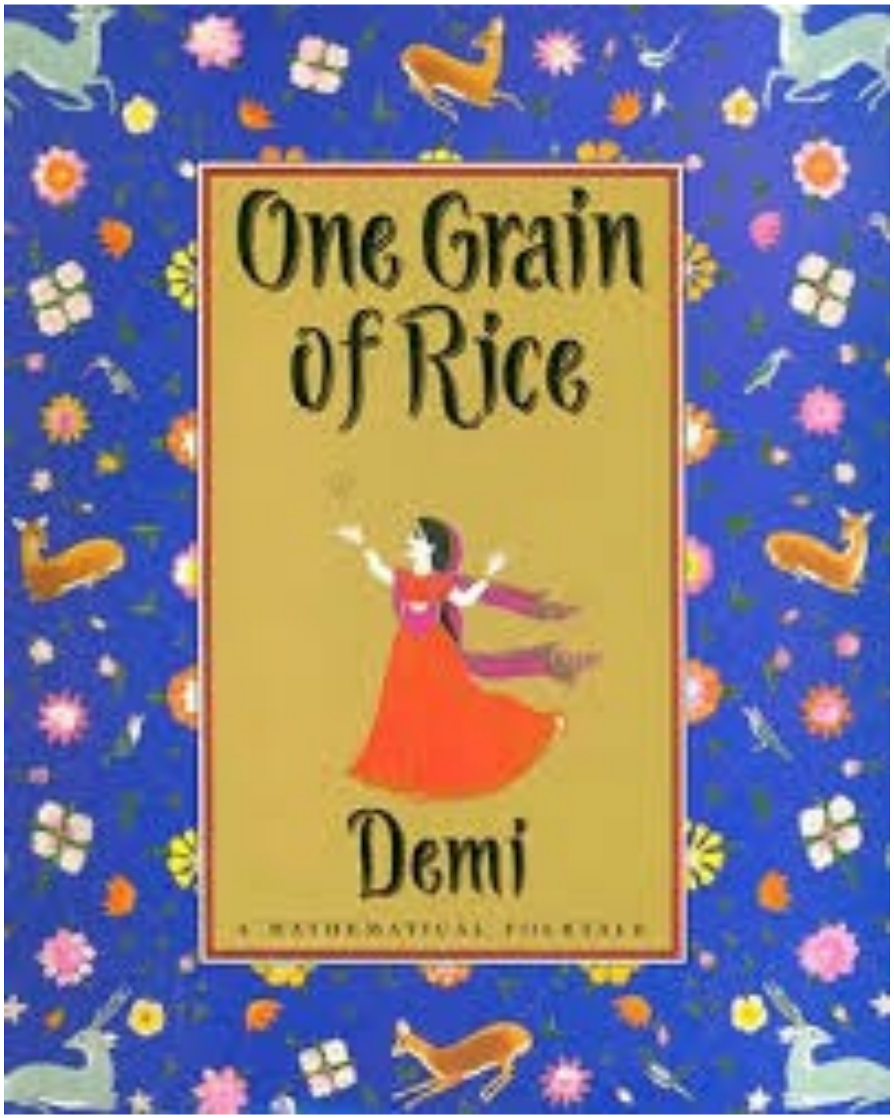
$$250^{150} \approx 10^{300}$$



$$\approx 2^{1200}$$

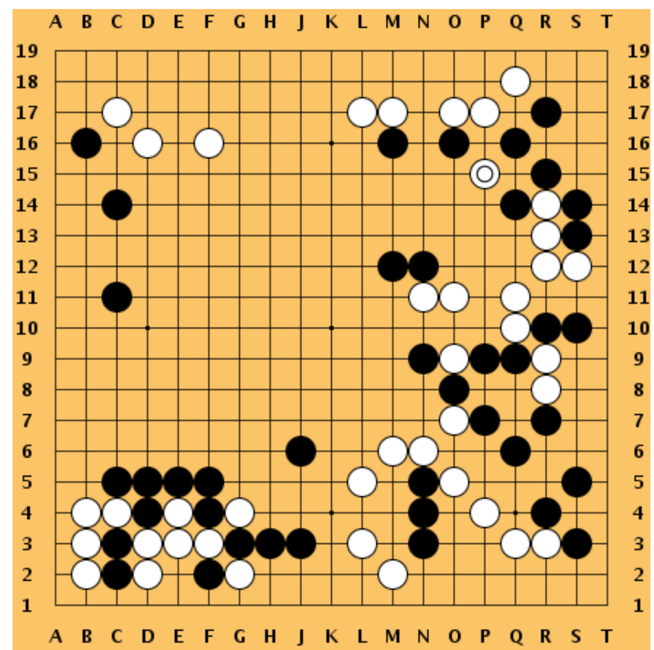
One Grain of Rice

https://www.youtube.com/watch?v=byk3pA1GPgU

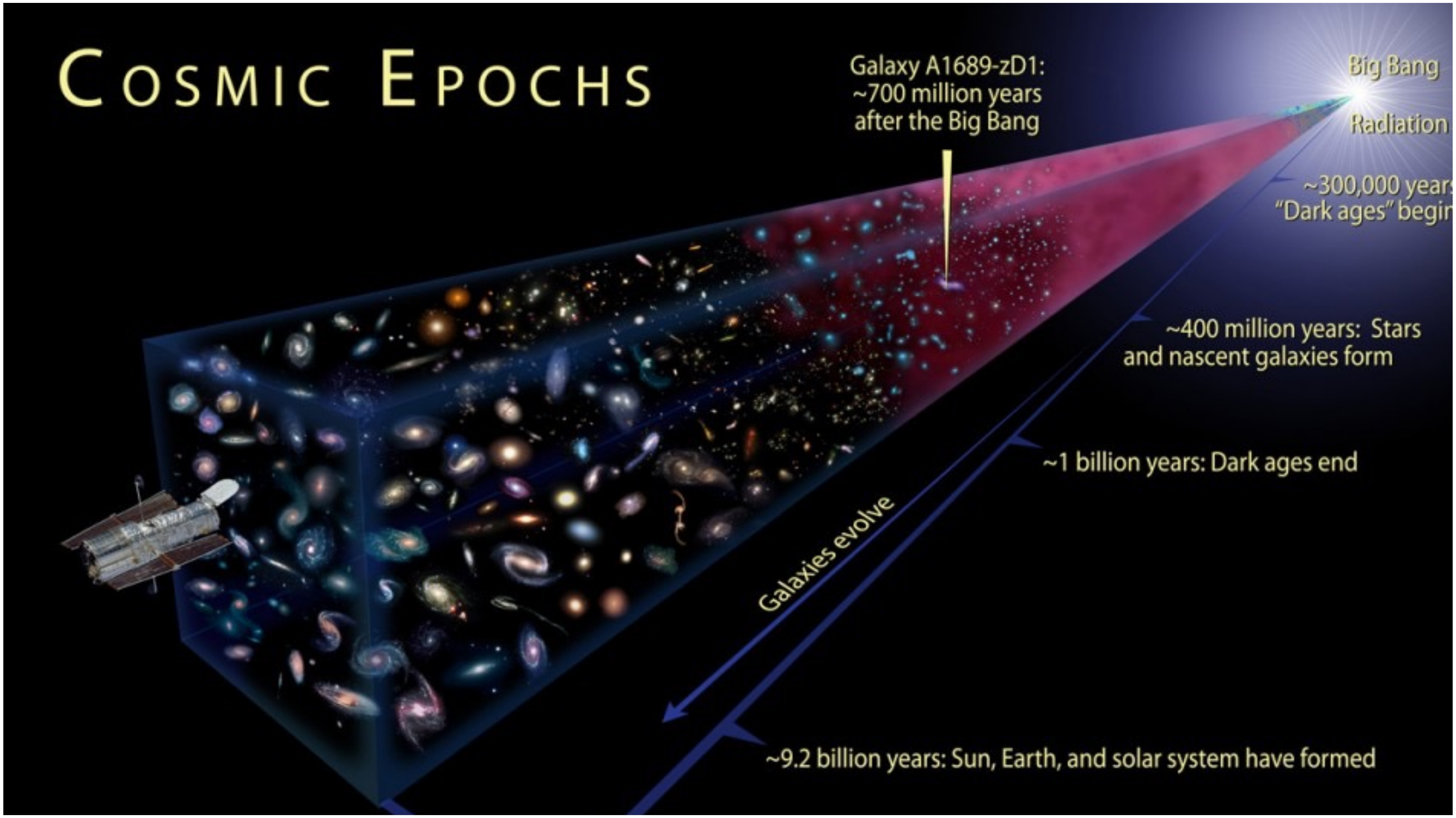


Size of the Rice in Chess

$2^{64} \rightarrow 2^{1200}$



The "Space" of GO Game



between 10^78 to 10^82 atoms in the known, observable universe

Numbers

1 trillion dollar	10^18
num Atoms in Universe	10^78 to 10^82
num settings in GO game	250^150 = 10^300
	(10^222 **universes**)

What's the current state of Computing Power?

$$N_A = 6.02 \times 10^{23}$$

Avogadro constant

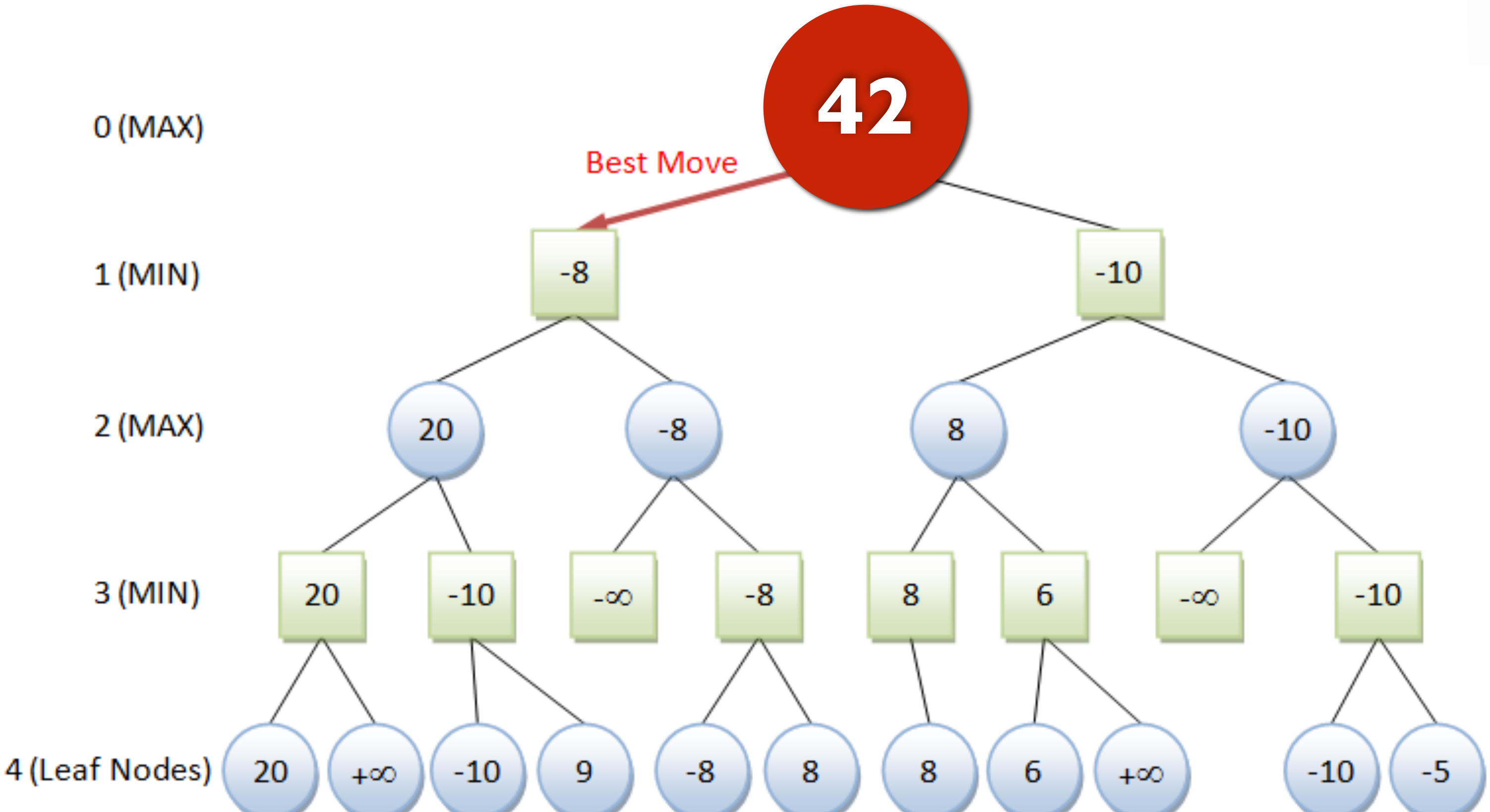
1 mol of water

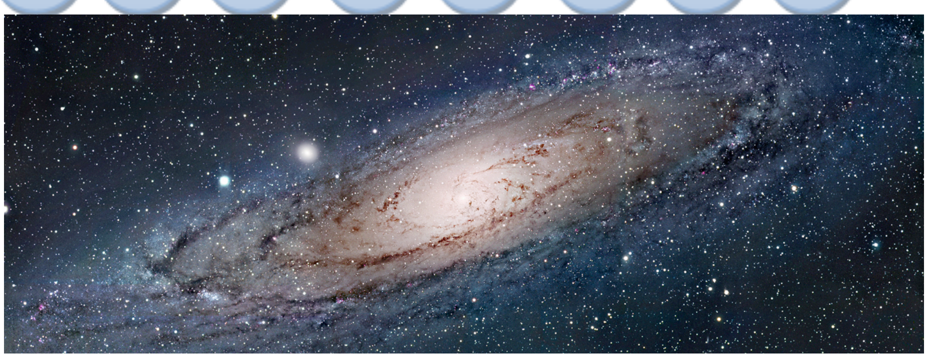
Numbers

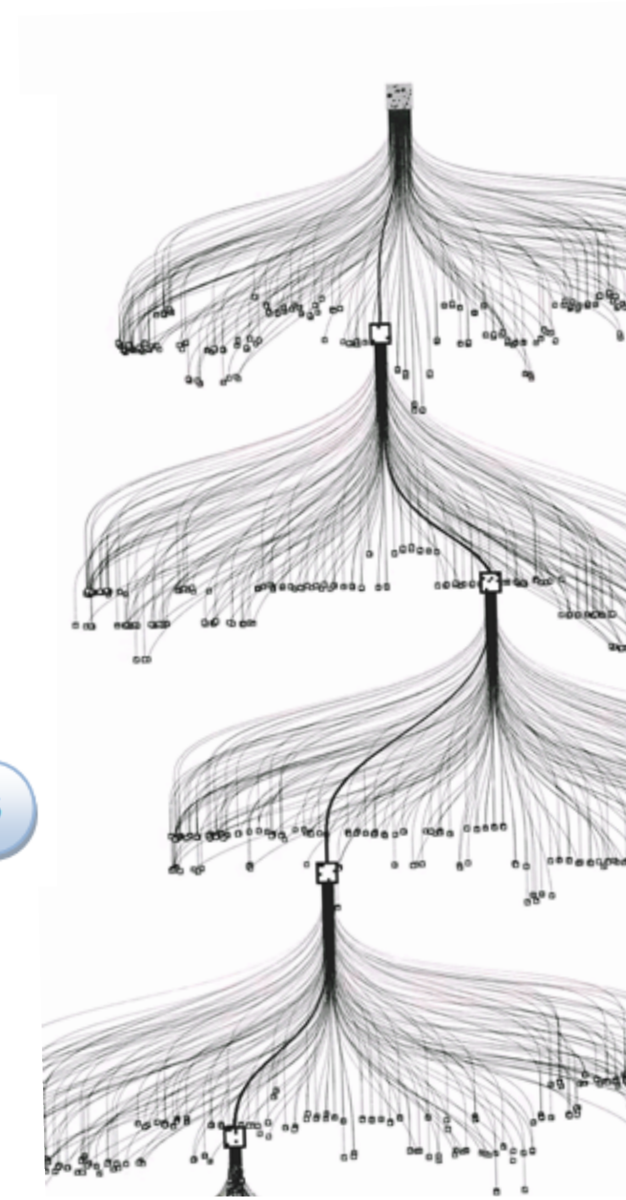
1 trillion dollar	10^18
num of molecules in 1 mol of Water (computing power)	10^23
num Atoms in Universe	10^78 to 10^82
num settings in GO game	250^150 = 10^300 (10^222 **universes**)

Computers are also Hopeless to find the ultimate answer of GO

BTW: the ultimate Answer to Life is 42 https://youtu.be/aboZctrHfK8?t=2s imagine that computing this score needs 8M years with a super computer that has yet been invented





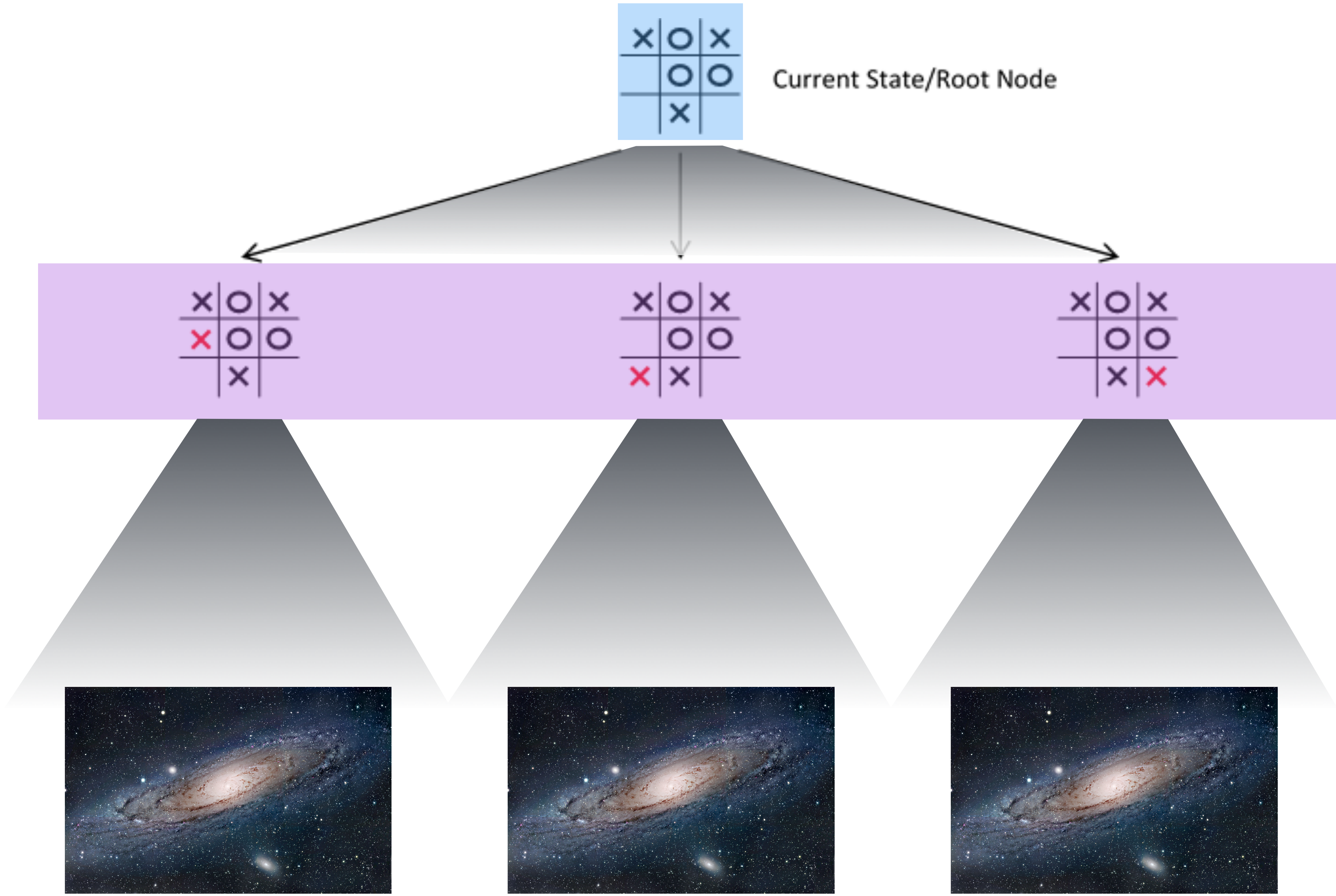


How to solve?

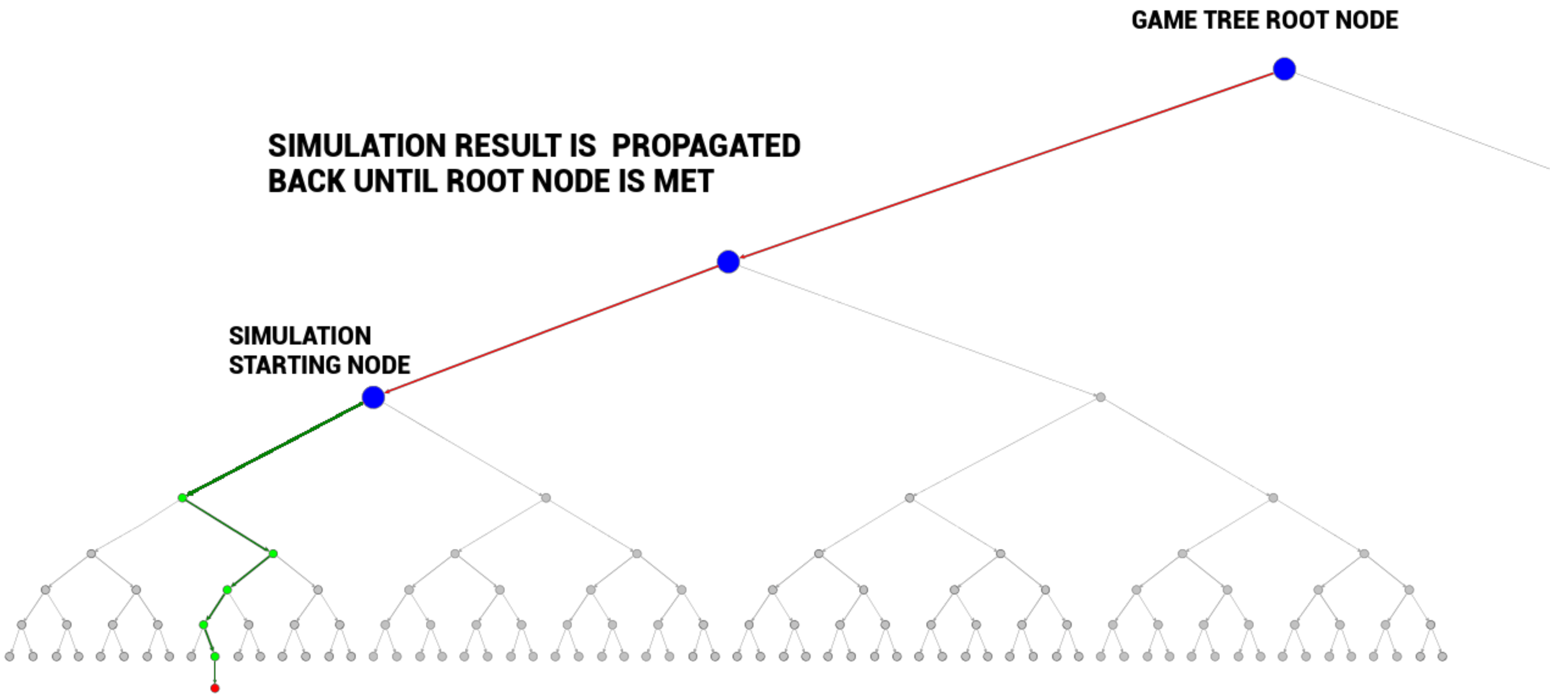
Good News
It's all about Winning
Human Players


Start with the current state

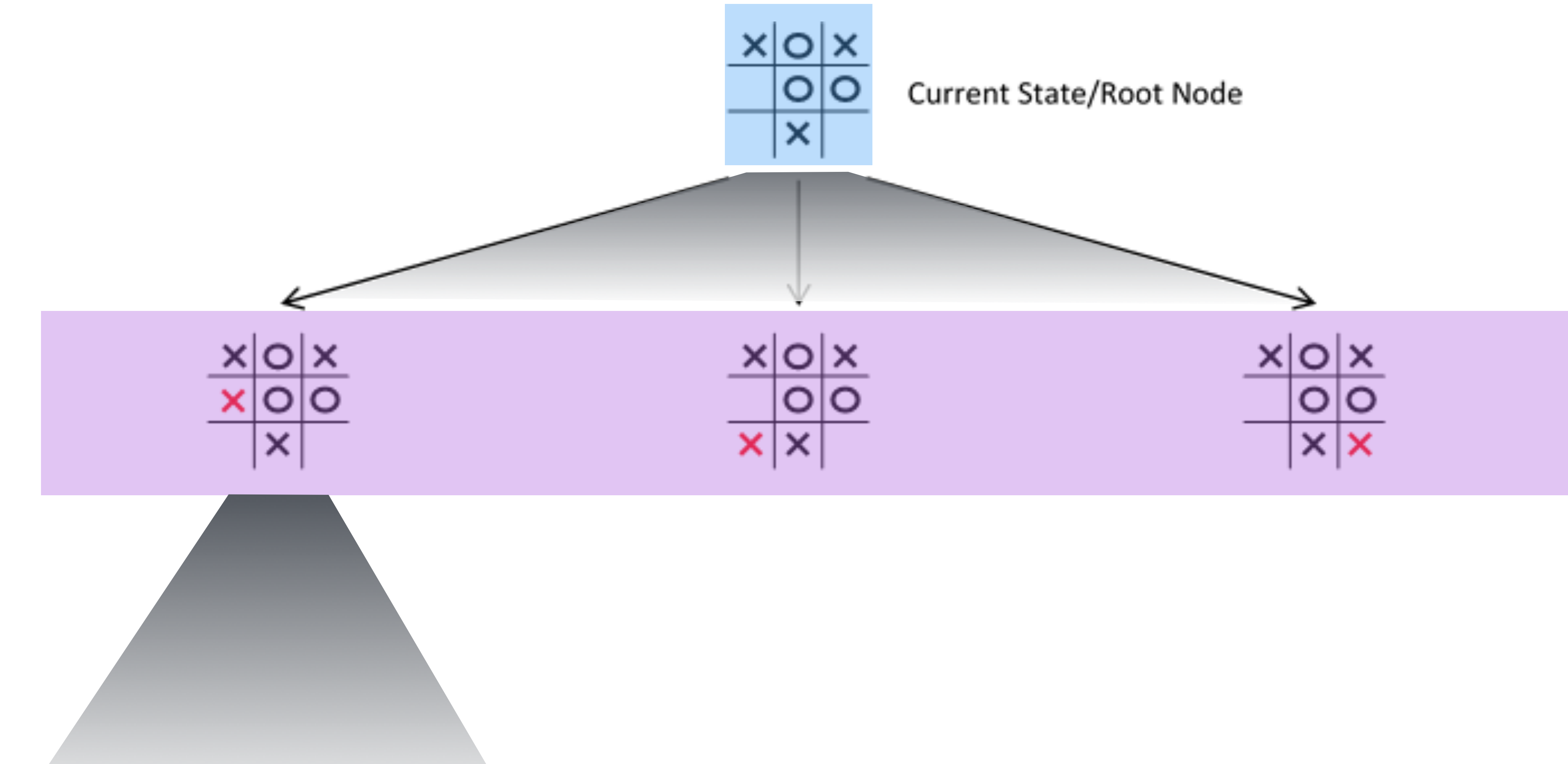
how to estimate the score of each move, without minimax search over each tree



Monte Carlo Tree Search (2005)

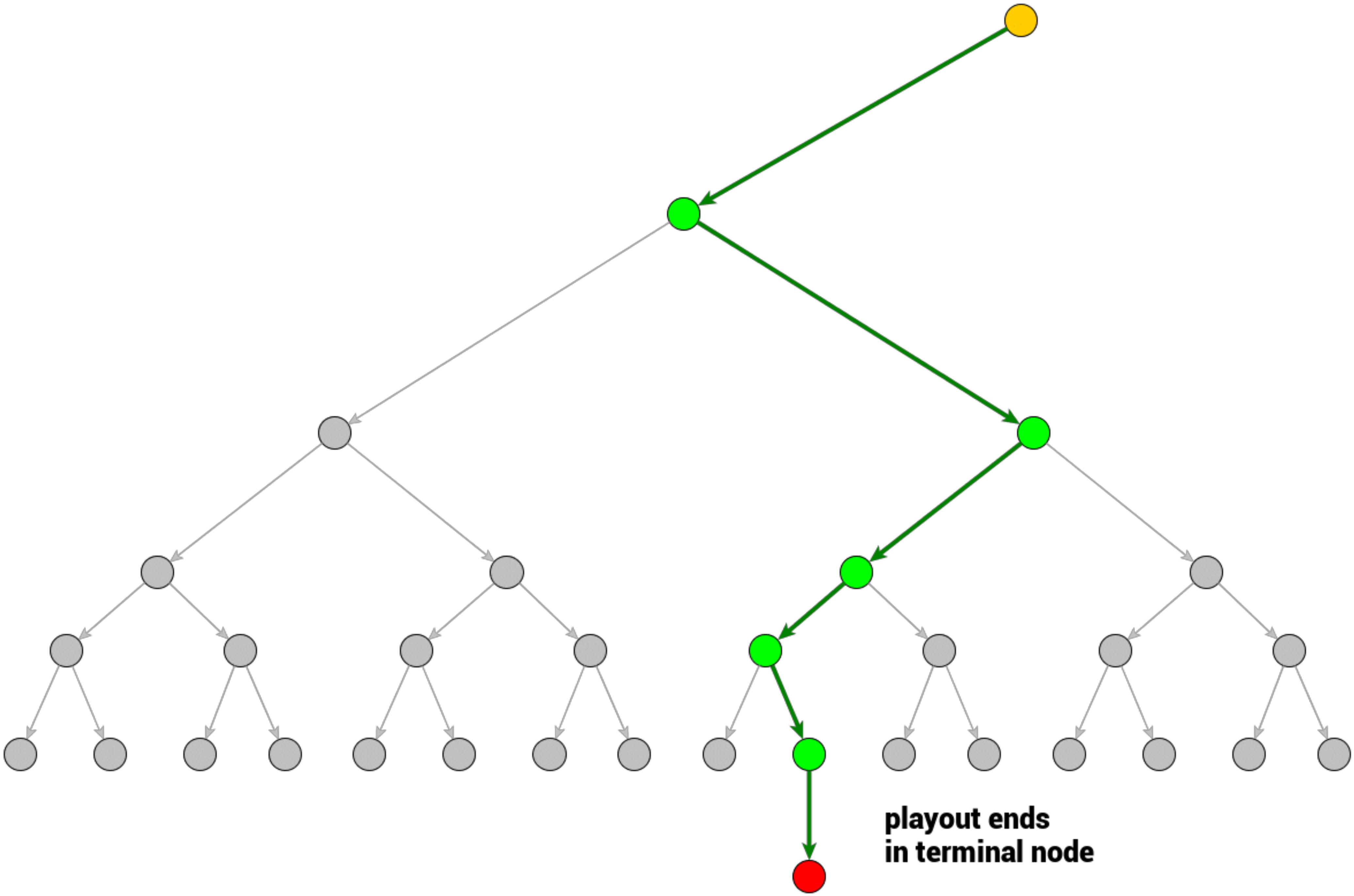


Sampling



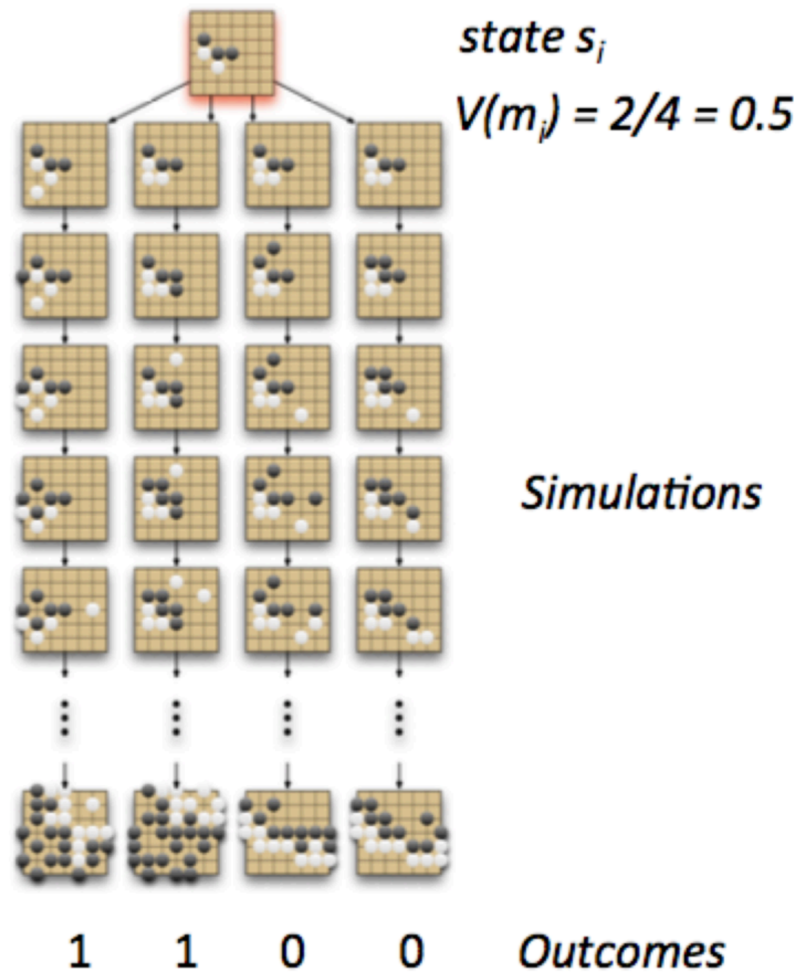
Get samples from the tree and estimate the winning probability

Monte Carlo Simulation

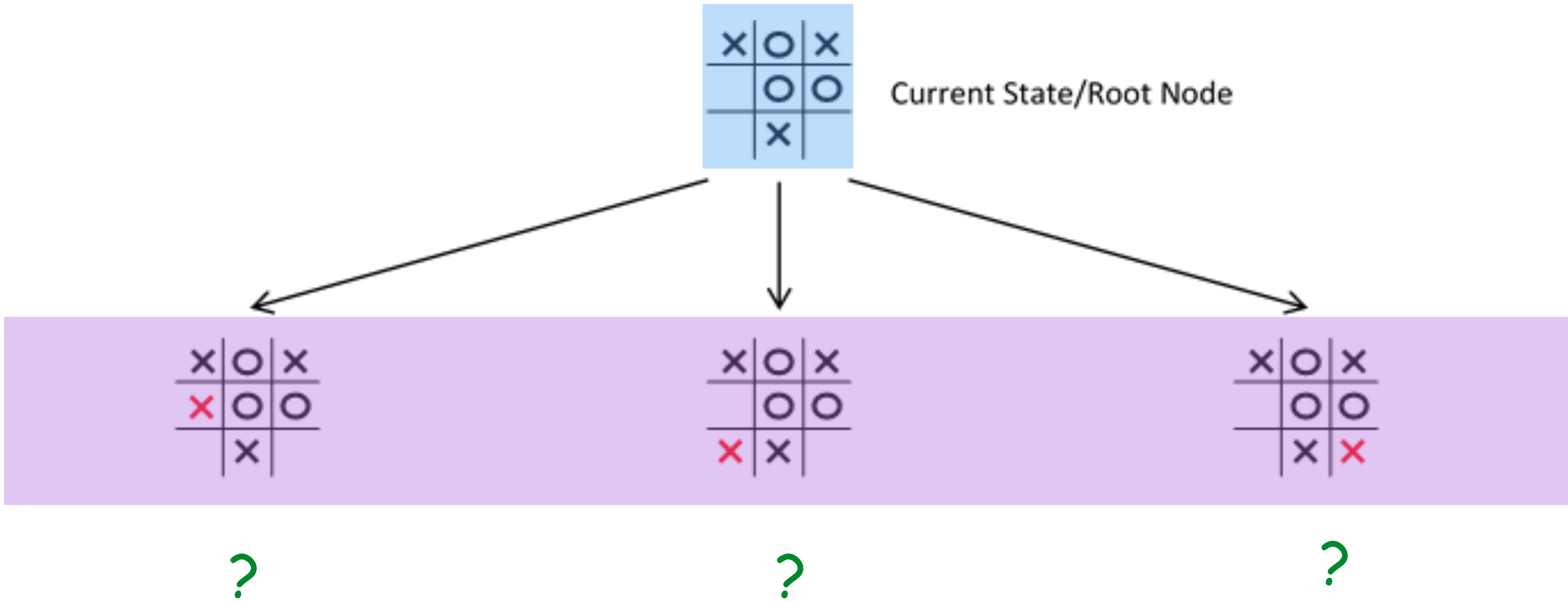


Basic Idea

- No evaluation function?
  - Simulate game using random moves
  - Score game at the end, keep winning statistics
  - Play move with best winning percentage
  - Repeat



A Simpler Problem

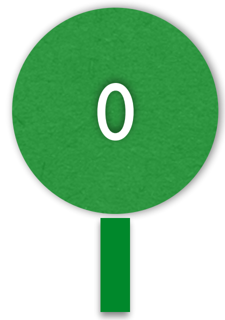


each choice of move has a probability to win

how choose to maximize the pay-off in the long term

Slot Machine

an arm to pull



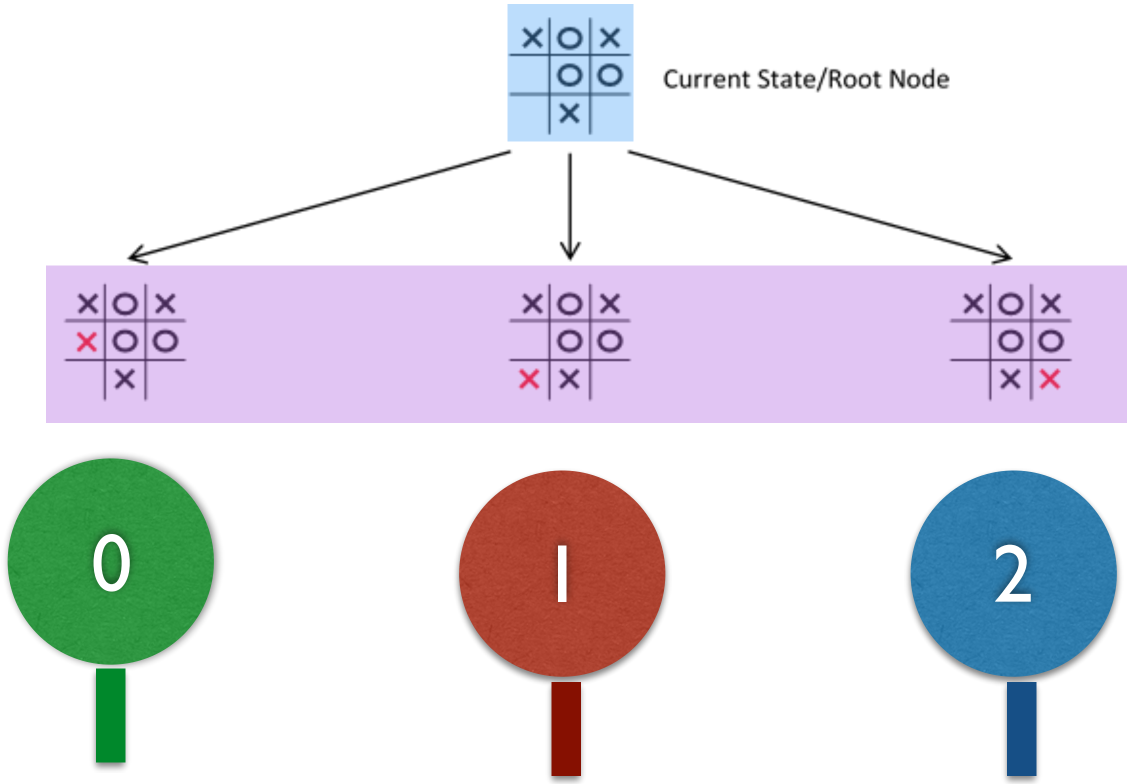
The machine has a probability to win say 0.4

When a player pull the arm, return reward 1 or 0 randomly

according to the probability of winning

https://www.youtube.com/watch?v=9qedY_Ujs1Y
https://www.youtube.com/watch?v=LPDI3uuRWCg

Multi-Armed Bandit Problem



n arms to pull

each arm has a probability to win

$p_0 = 0.8$ $p_1 = 0.5$ $p_2 = 0.2$

When a player pull an arm, return reward 1 or 0 randomly according to the probability of winning

Multi-Armed Bandit Problem

n arms to pull　　0　　1　　2

each arm has a probability to win (hidden from player)

?　　?　　?

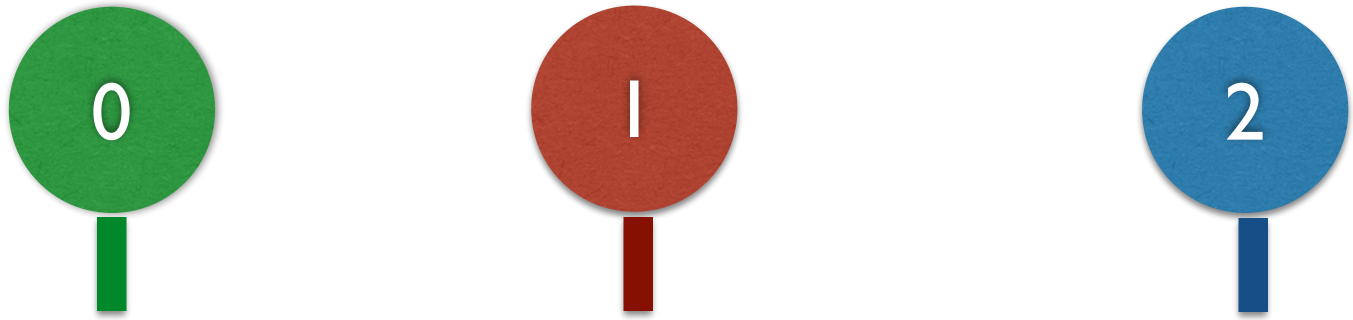
Multi-Armed Bandit Problem

- Assumptions
  - Choice of several arms
  - Each arm pull is independent of other pulls
  - Each arm has fixed, unknown average payoff
- Which arm has the best average payoff?

How to Solve ?

Suppose p is known

n arms to pull



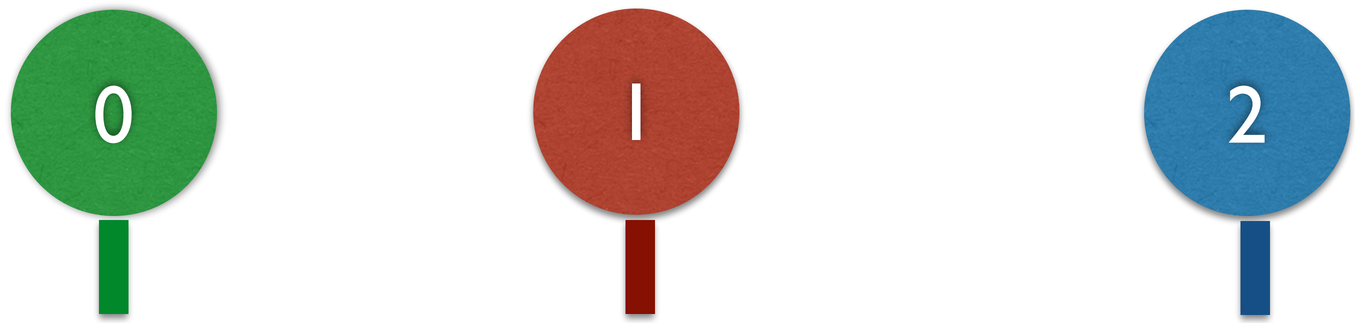
each arm has a probability to win

$p_0 = 0.5$ $\quad p_1 = 0.7$ $\quad p_2 = 0.2$

$$\boxed{\text{arm to pull} = \operatorname{argmax}_i \, p_i}$$

What if p is unknown

n arms to pull



each arm has a probability to win

$p_0 = 0.5$ $p_1 = 0.7$ $p_2 = 0.2$

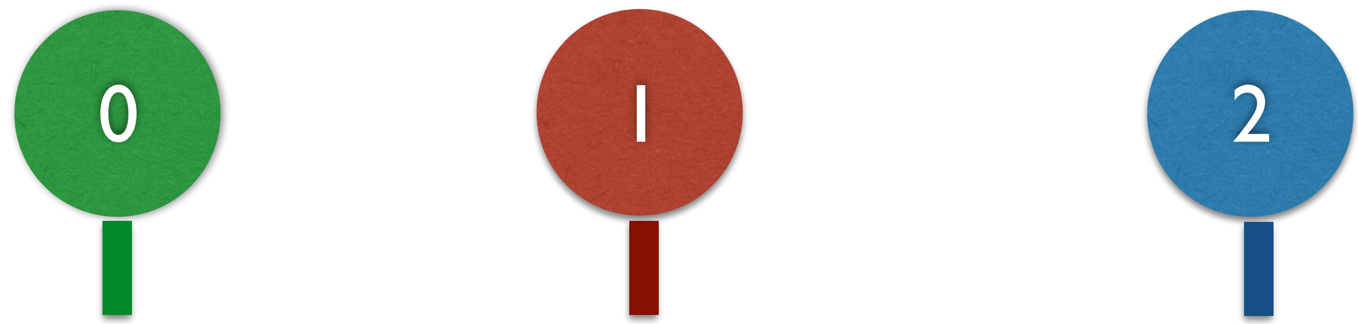
Player's estimation of winning probability:

$Q_0 = 0.4$ $Q_1 = 0.6$ $Q_2 = 0.3$

Each estimation is based upon the ratio of winning after pulling each arm

Strategy 1: choose the arm with the largest estimated reward

n arms to pull



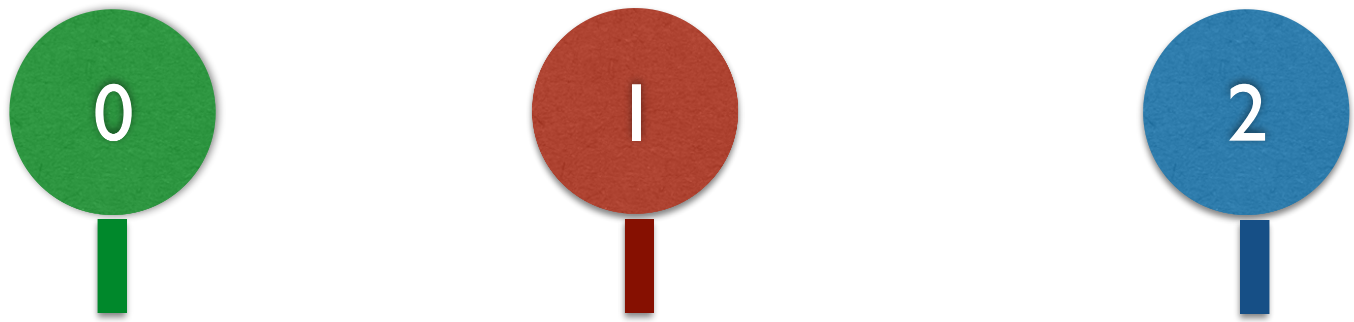
Player's estimation of winning probability:

$Q_0 = 0.4$ $Q_1 = 0.6$ $Q_2 = 0.3$

$$\boxed{\text{arm to pull} = \operatorname{argmax}_i Q_i}$$

Problem

How to initialize Q?

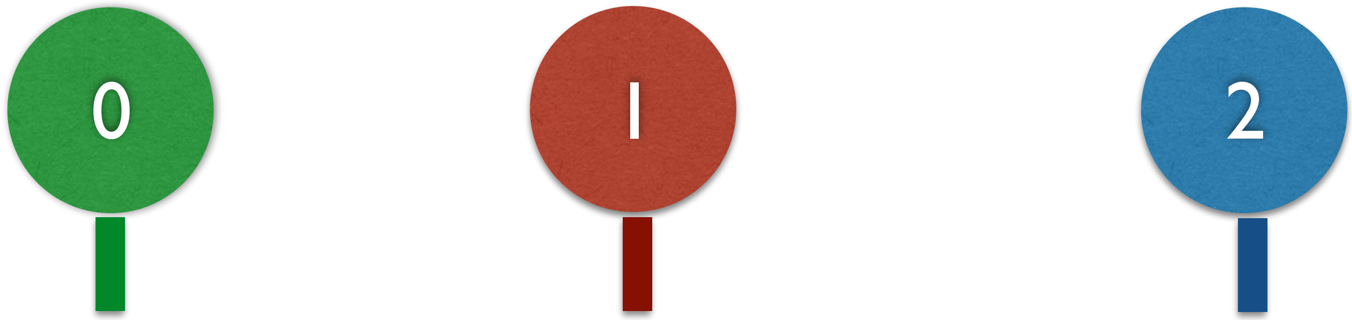


$Q_0 = 0$ $Q_1 = 0$ $Q_2 = 0$

Pull arm 0

Problem

How to update Q?

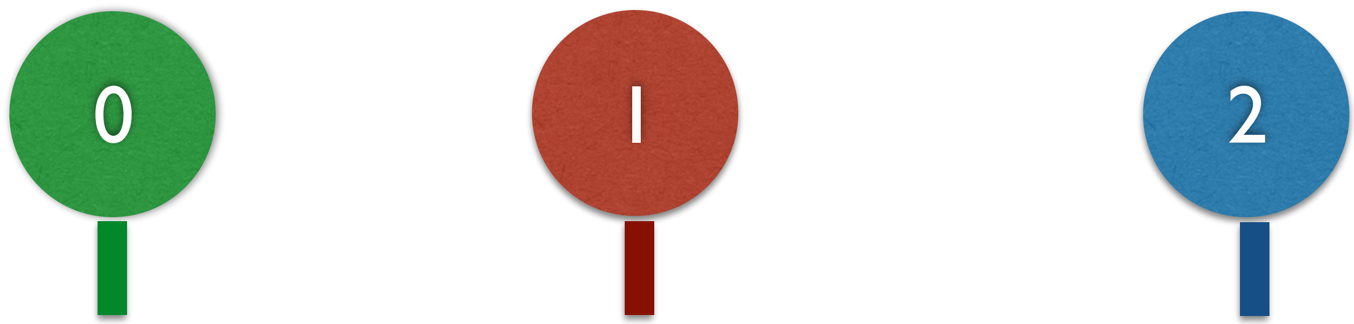


$Q_0 = 0$ $Q_1 = 0$ $Q_2 = 0$

Pull arm 0
reward=1 (Win)

Problem

How to update Q?

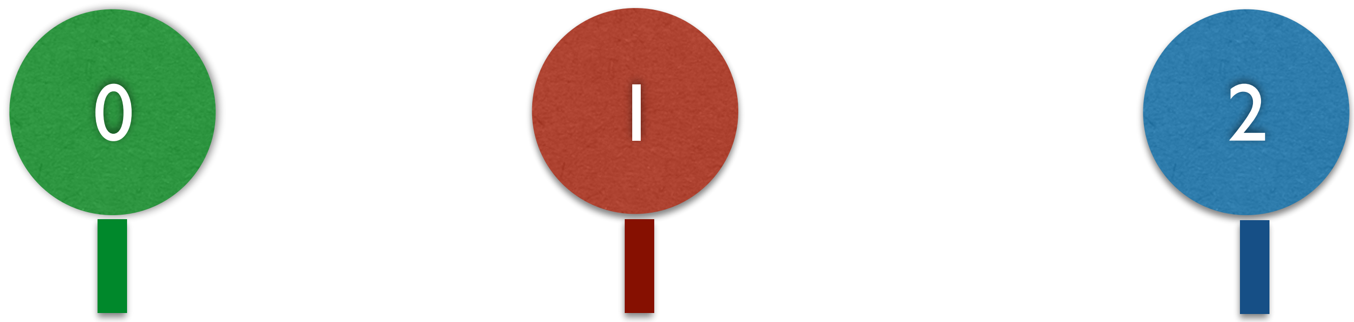


$Q_0 = 1.0$ $Q_1 = 0$ $Q_2 = 0$

Pull arm 0
reward=1 (Win)

Problem

How to update Q?

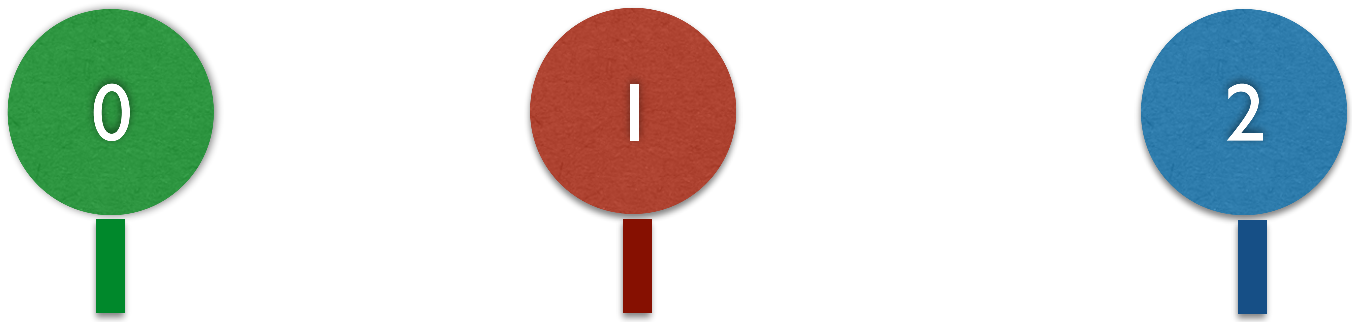


$Q_0 = 1.0$ $Q_1 = 0$ $Q_2 = 0$

Pull arm 0
reward=1 (Win)
Pull arm 0
reward=0 (Lose)

Problem

How to update Q?



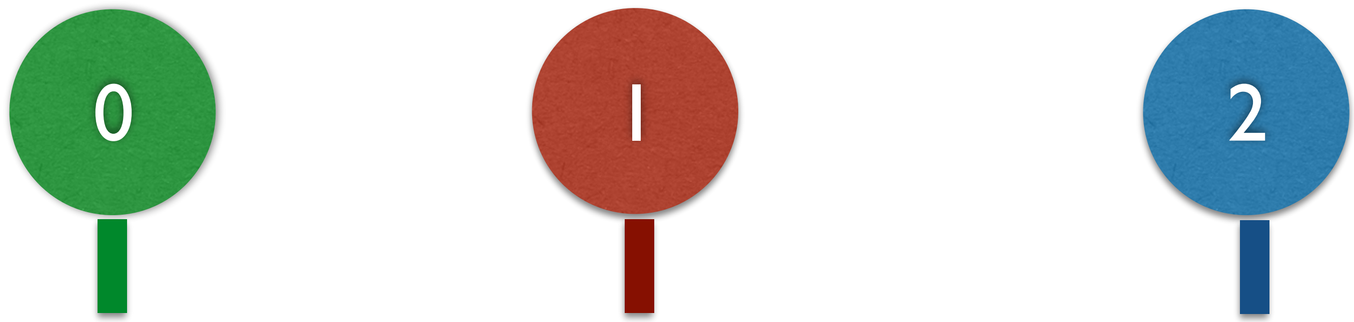
$Q_0 = 0.5$ $Q_1 = 0$ $Q_2 = 0$

Pull arm 0
reward=1 (Win)
Pull arm 0
reward=0 (Lose)
Pull arm 0

reward=1 (Win)

Problem

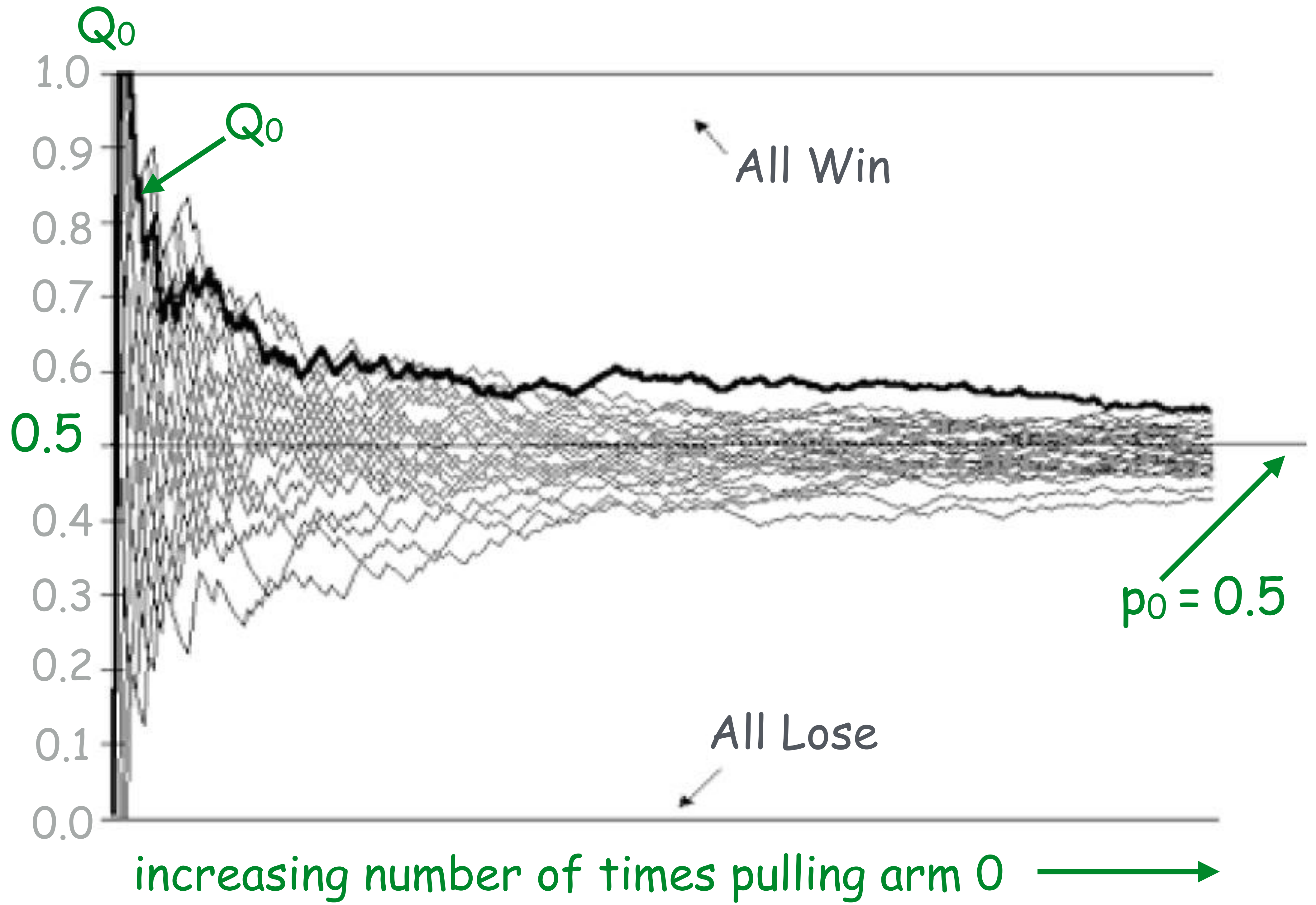
How to update Q?



$Q_0 = 0.67$ $Q_1 = 0$ $Q_2 = 0$

Pull arm 0
reward=1 (Win)
Pull arm 0
reward=0 (Lose)
Pull arm 0

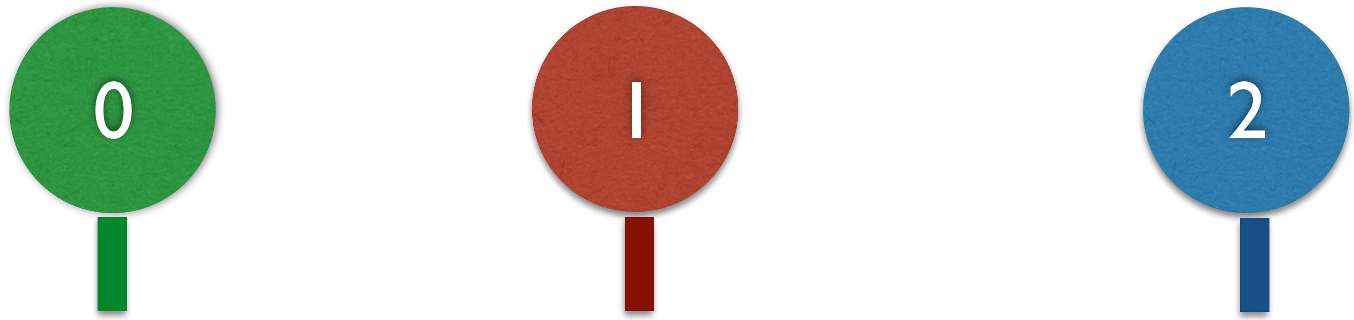
reward=1 (Win)



Q0
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Q0
All Win
All Lose
p0 = 0.5
increasing number of times pulling arm 0

Strategy 2: choose an arm randomly

n arms to pull



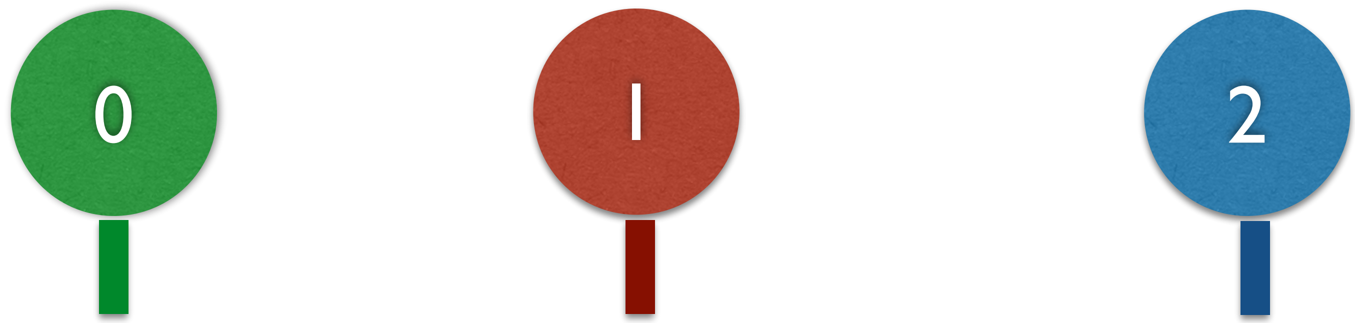
probability to pull: 1/3 1/3 1/3

$$\boxed{\text{arm to pull} = \text{random}}$$

ignoring Q values

Pros and Cons

n arms to pull



each arm has a probability to win

$p_0 = 0.5$ $p_1 = 0.7$ $p_2 = 0.2$

Player's estimation of winning probability:

$Q_0 = 0.5$ $Q_1 = 0.7$ $Q_2 = 0.2$

If we pull many times, the estimated probability Q will be very close to p

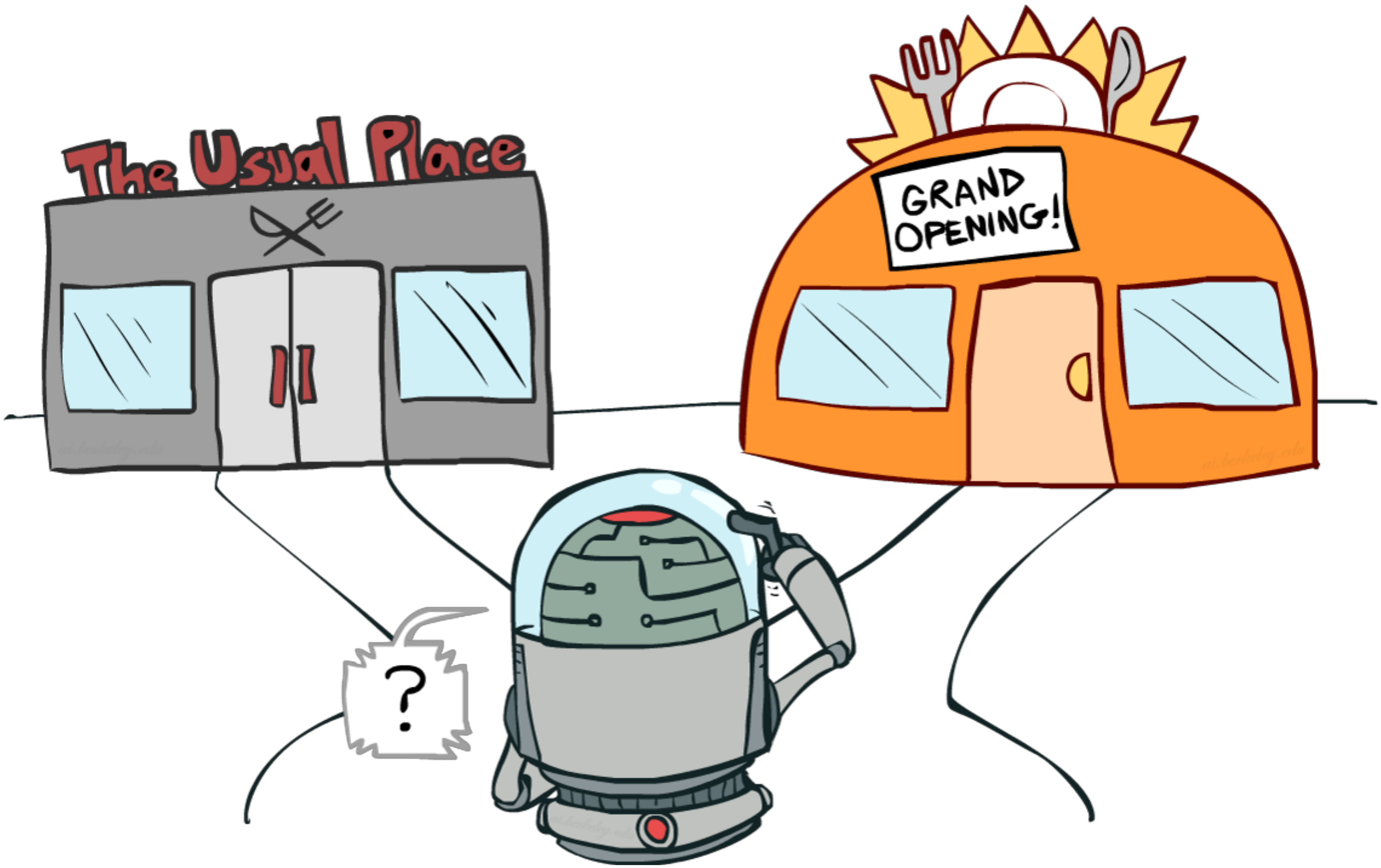
Exploration vs Exploitation

(Exploitation)

$\text{argmax}_i\ Q_i$

(Exploration)

random



Upper Confidence Bound

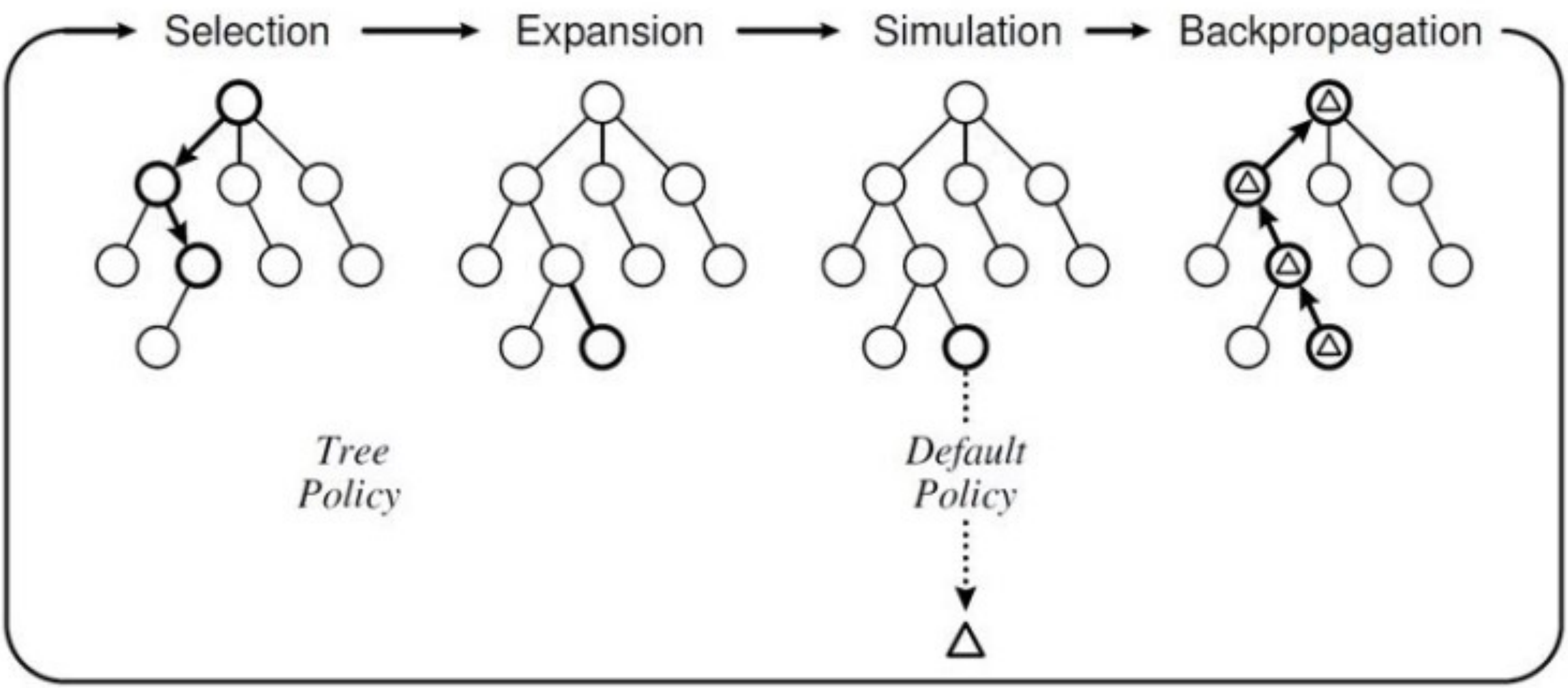
- Choose the arm that maximizes formula:

$$\underbrace{v_i}_{\text{value estimate}} + \underbrace{C}_{\text{tunable parameter}} \times \sqrt{\frac{\ln(\overbrace{N}^{\text{total number of trials}})}{\underbrace{n_i}_{\text{num trials for arm i}}}}$$

Prefers higher payoff arm

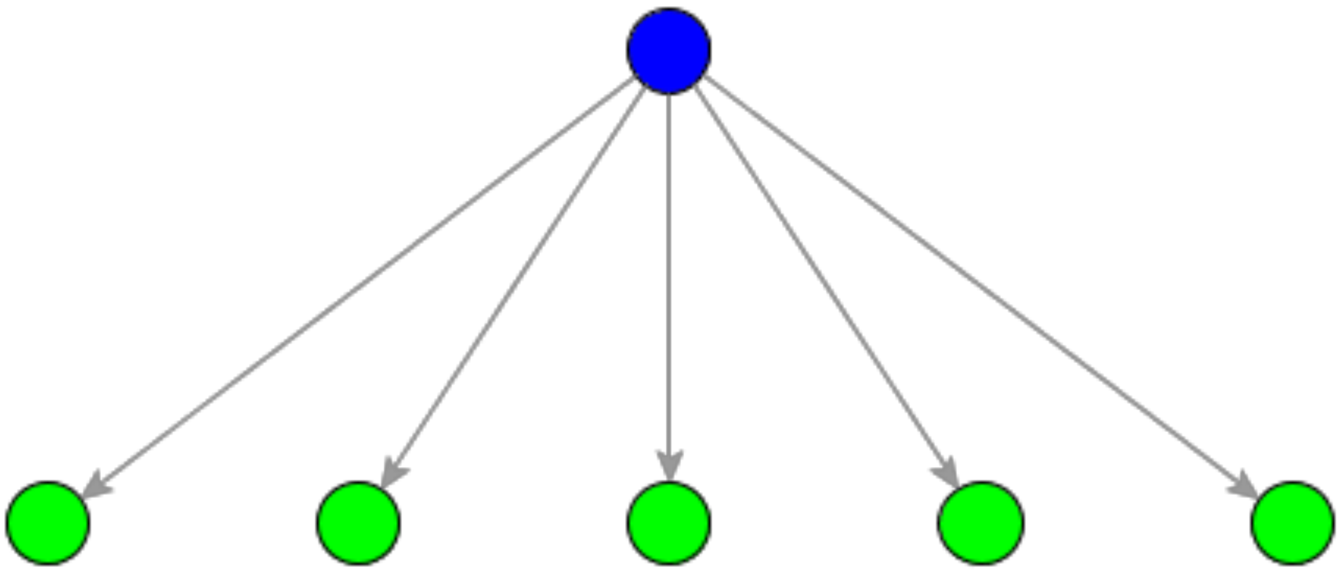
Prefers less played arm

Monte-Carlo Tree-Search



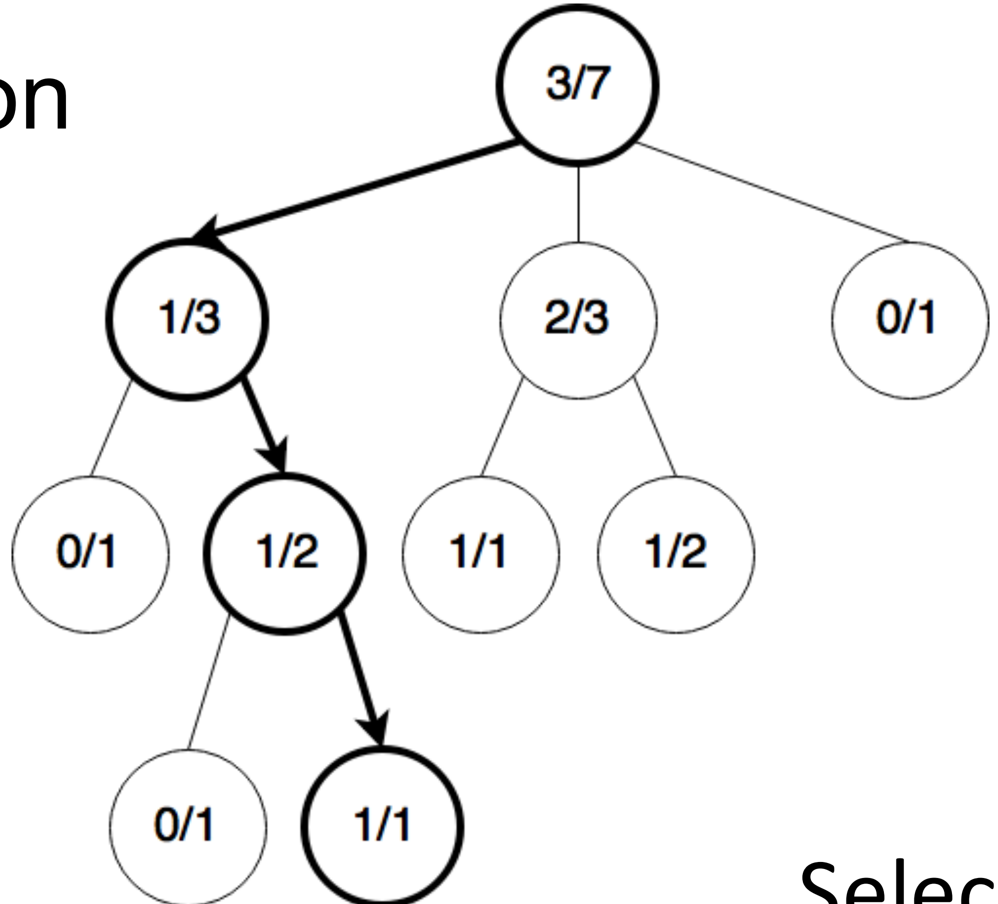
Expanding a Node

Expanding a Node



Monte Carlo Tree Search

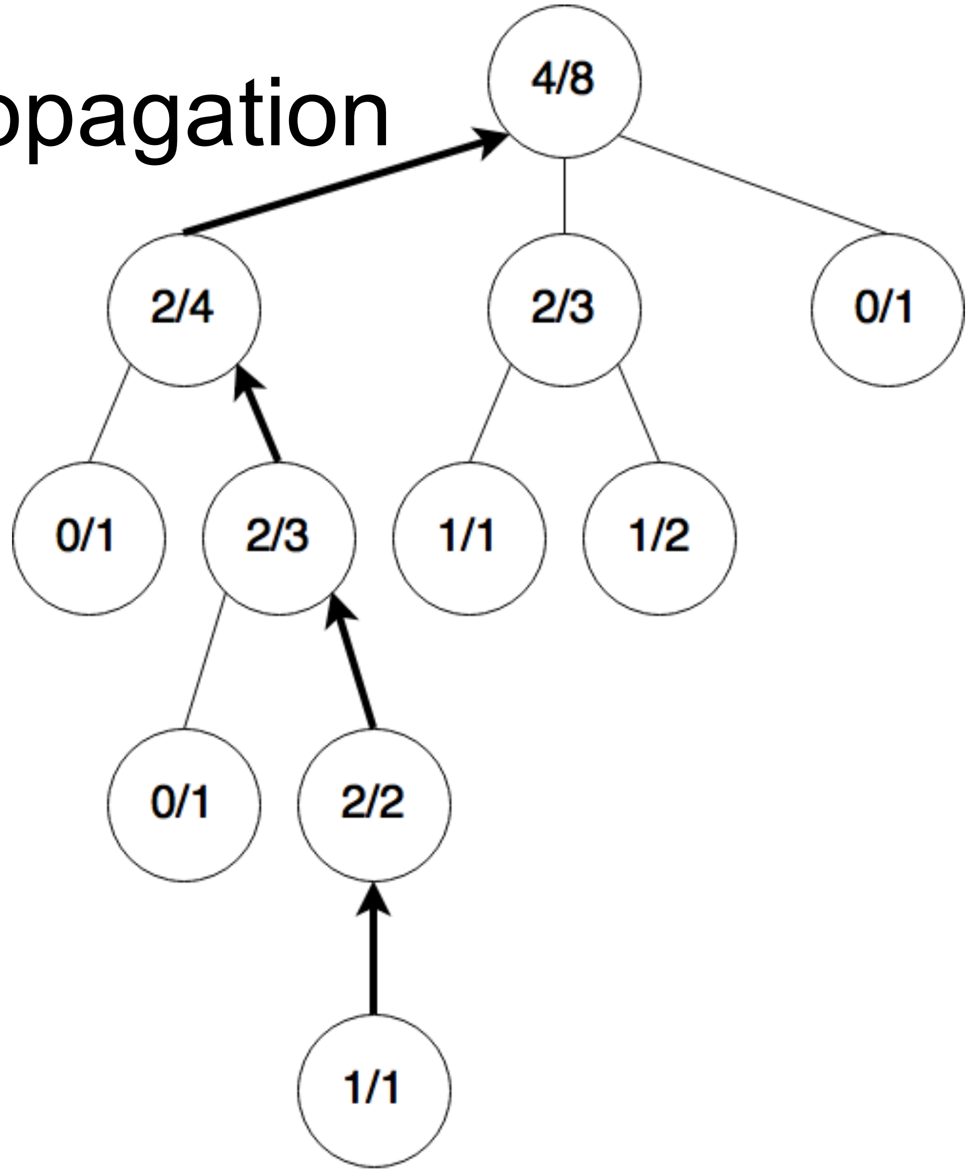
(1) Selection

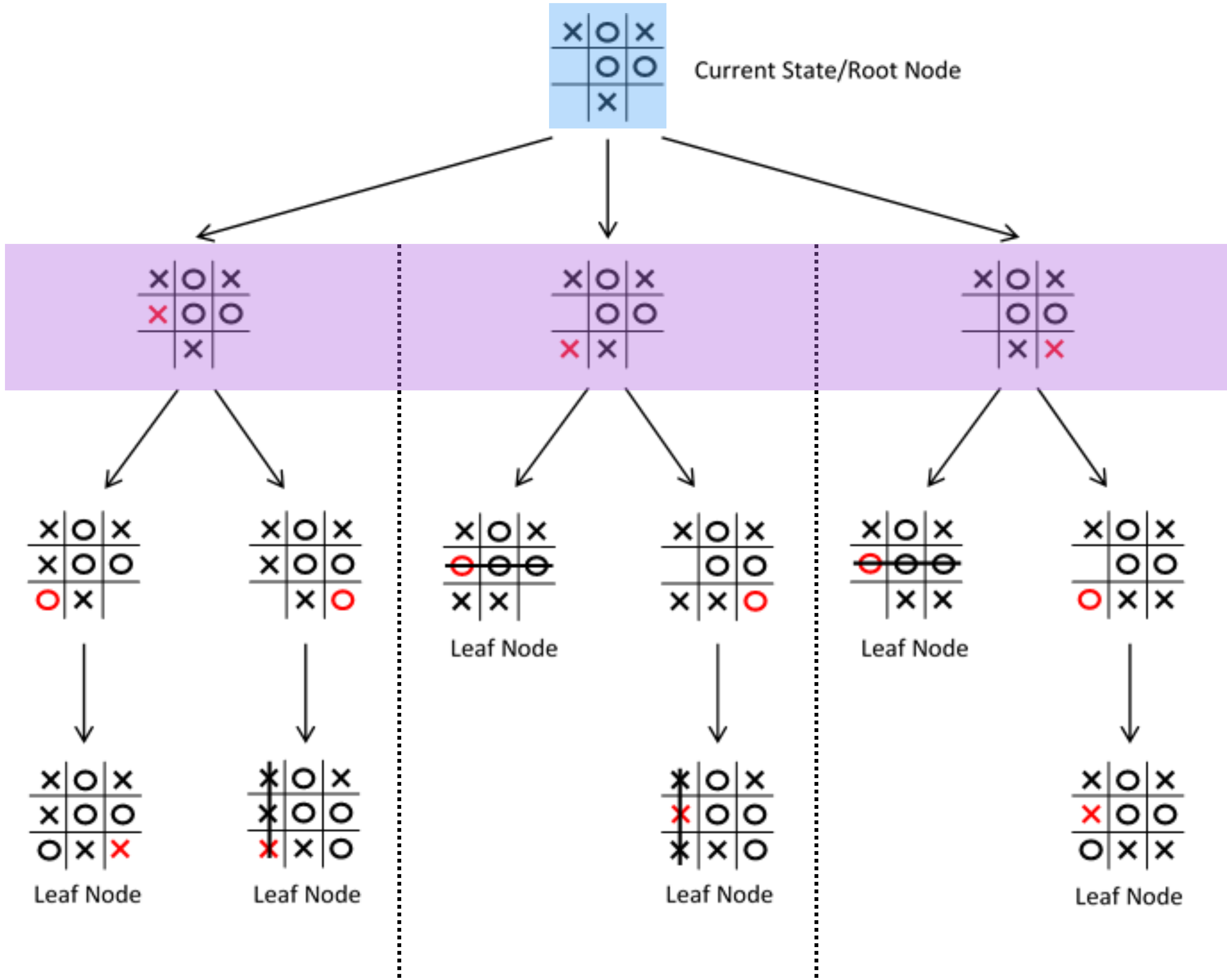


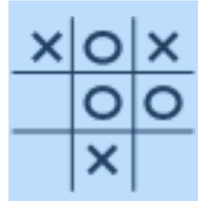
Selection policy is applied recursively until a leaf node is reached

Monte Carlo Tree Search

(4) Backpropagation

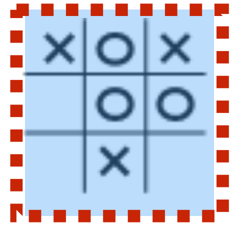



Current State/Root Node
Leaf Node
Leaf Node
Leaf Node
Leaf Node
Leaf Node
Leaf Node

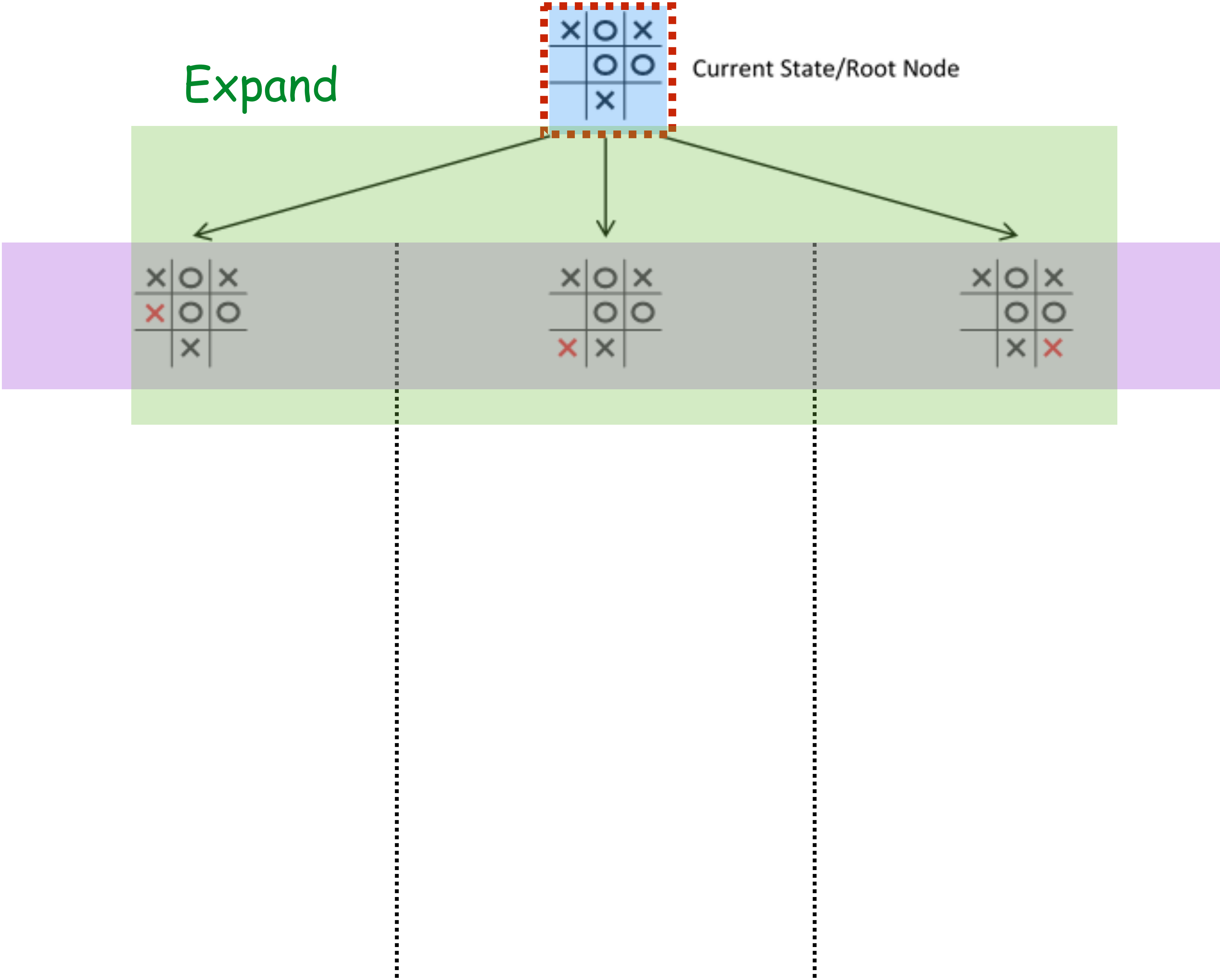


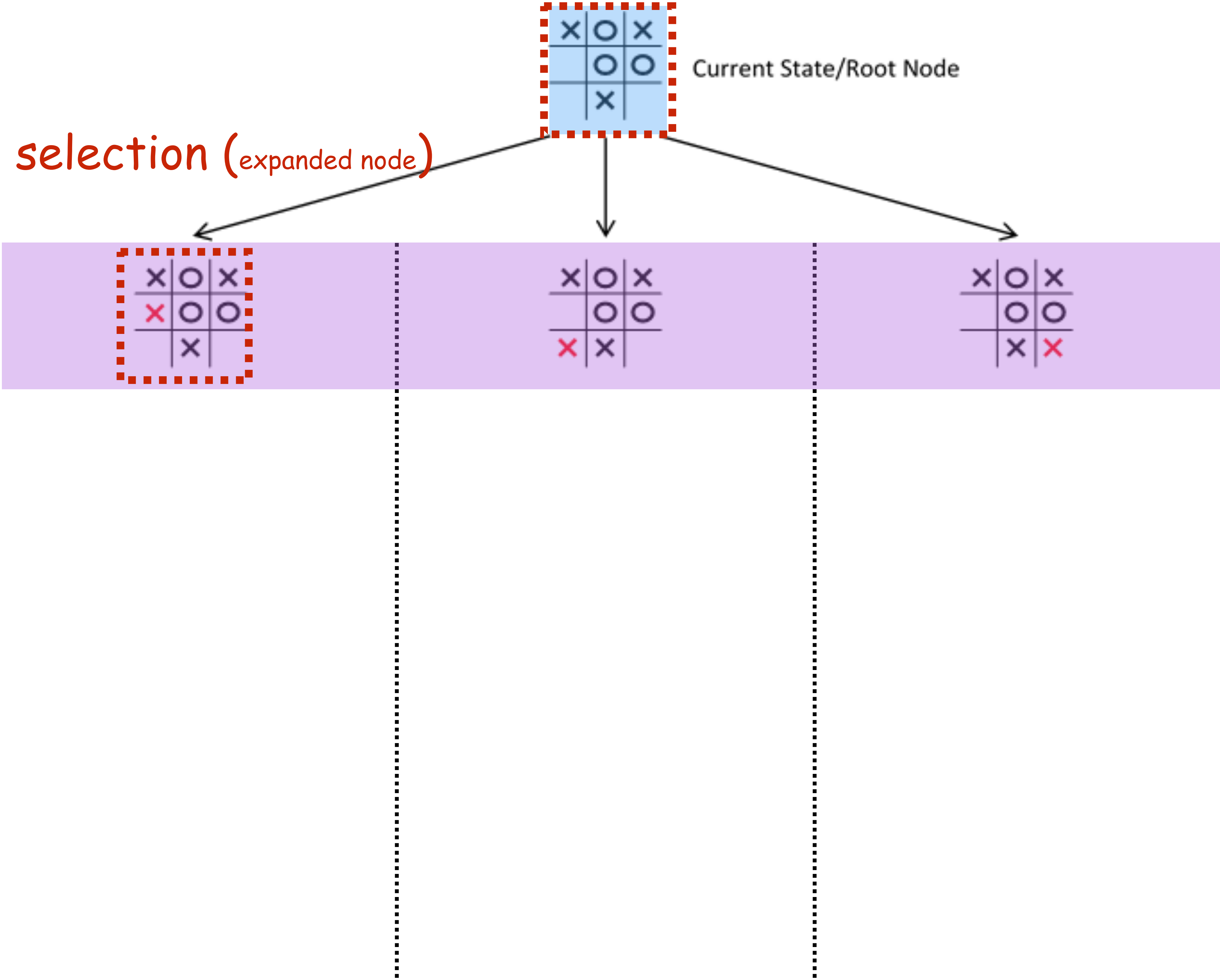
Current State/Root Node

Selection (from root)

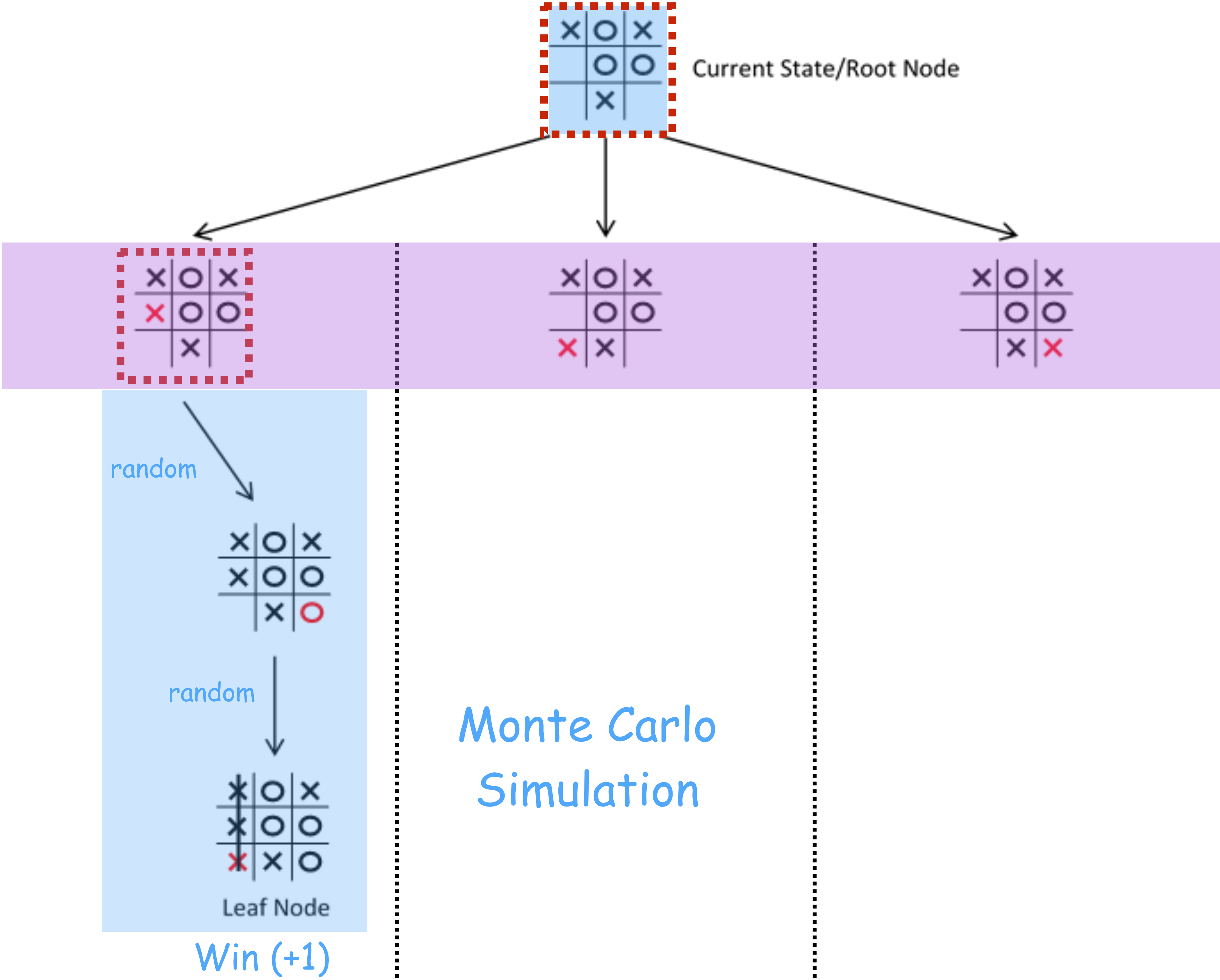


Current State/Root Node

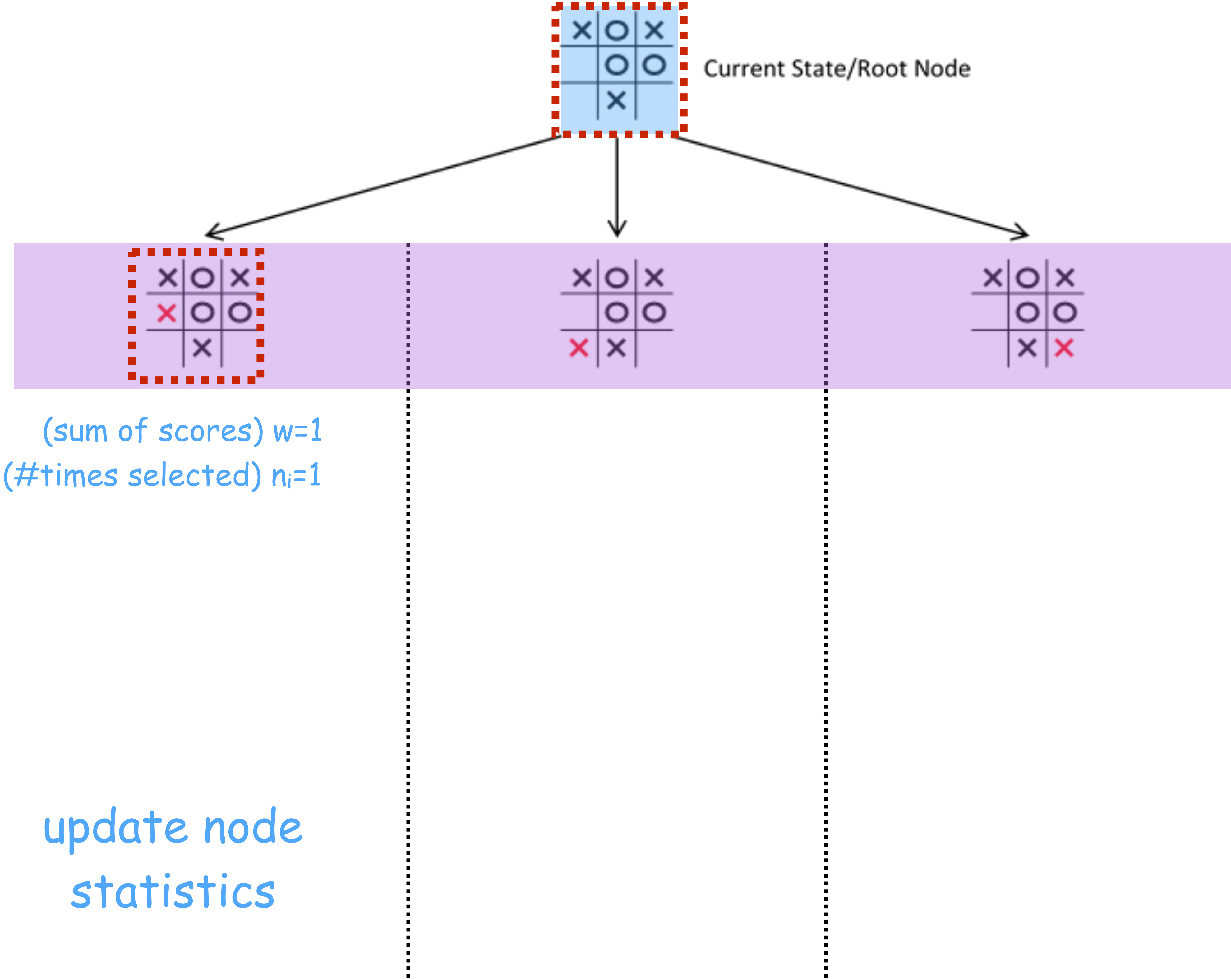

Expand
Current State/Root Node



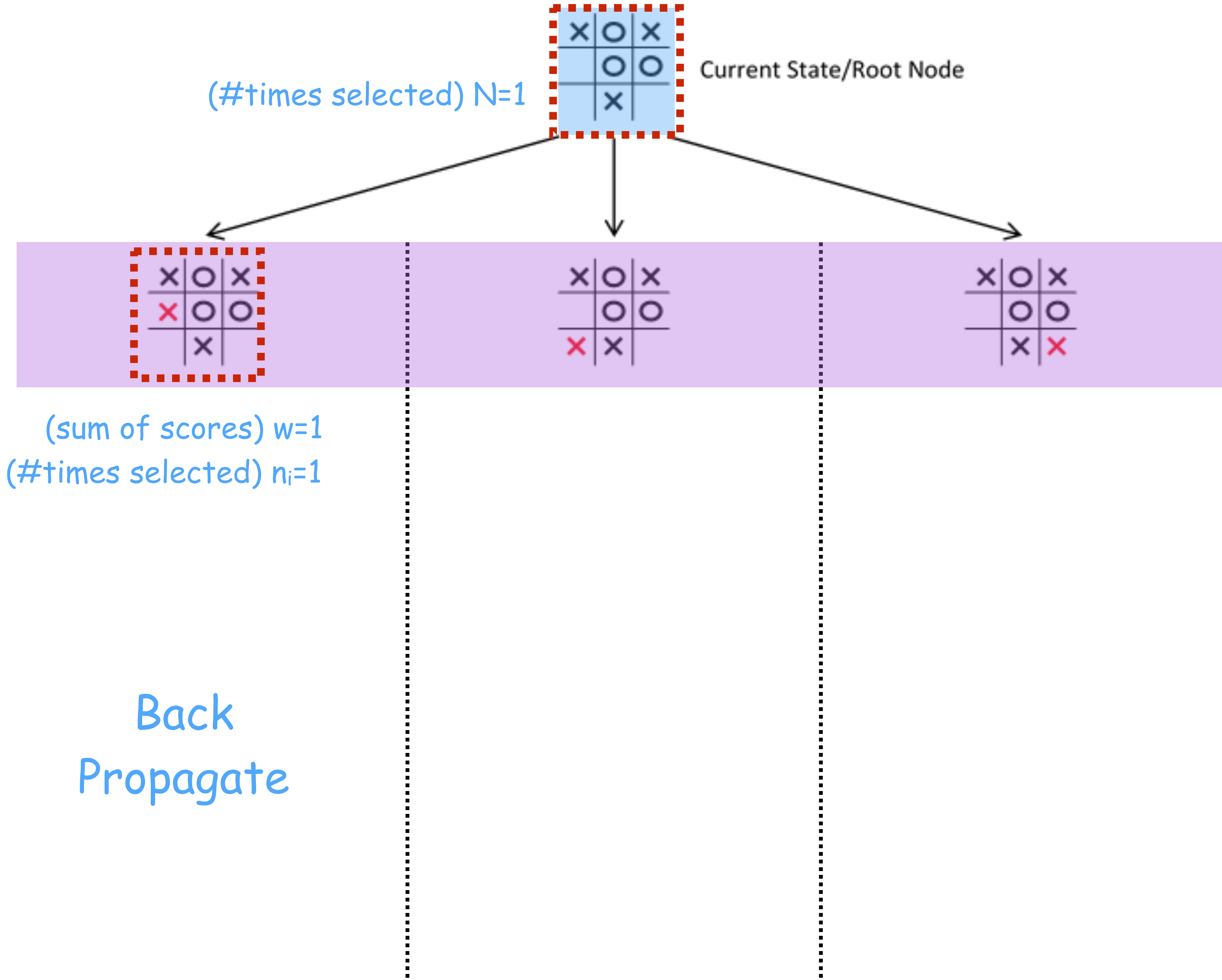
Current State/Root Node
selection (expanded node)

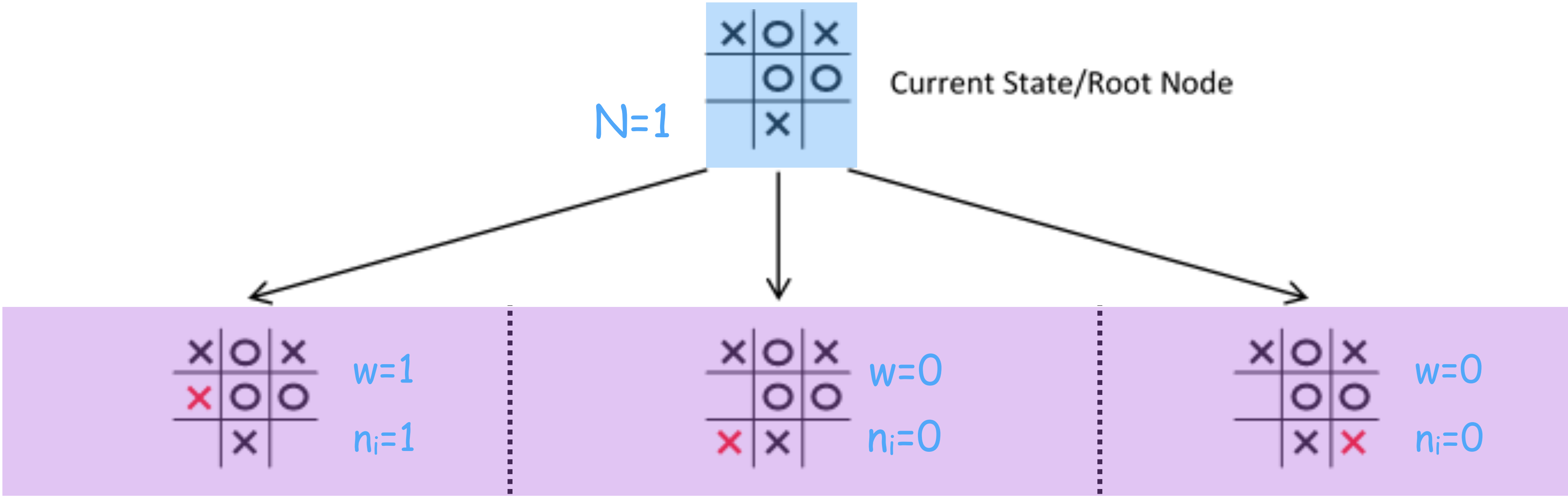


Current State/Root Node
random
random
Leaf Node
Win (+1)
Monte Carlo Simulation



Current State/Root Node
(sum of scores) w=1
(#times selected) n_i=1
update node statistics

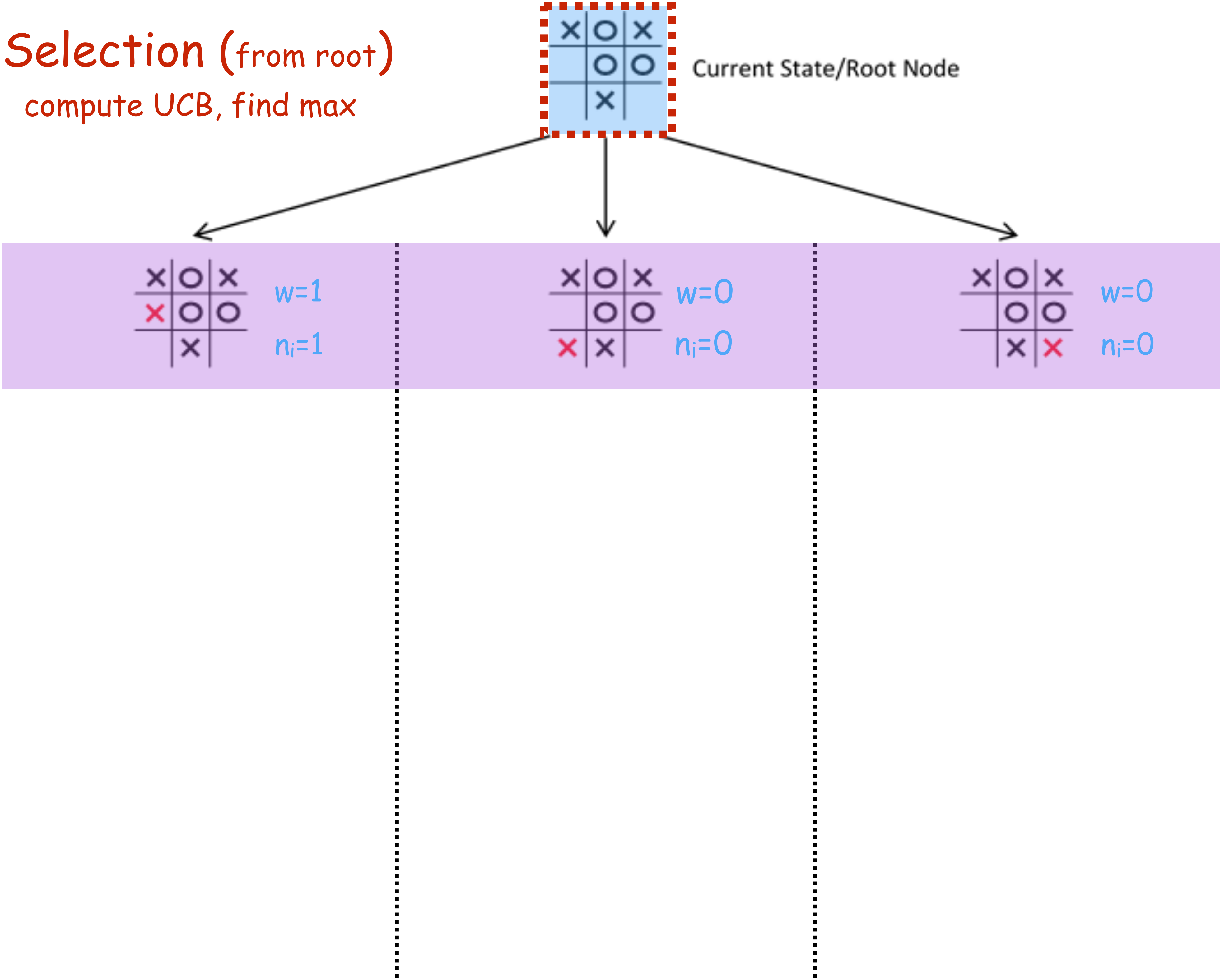

(#times selected) N=1
Current State/Root Node
(sum of scores) w=1
(#times selected) n_i=1
Back
Propagate



After 1 iteration

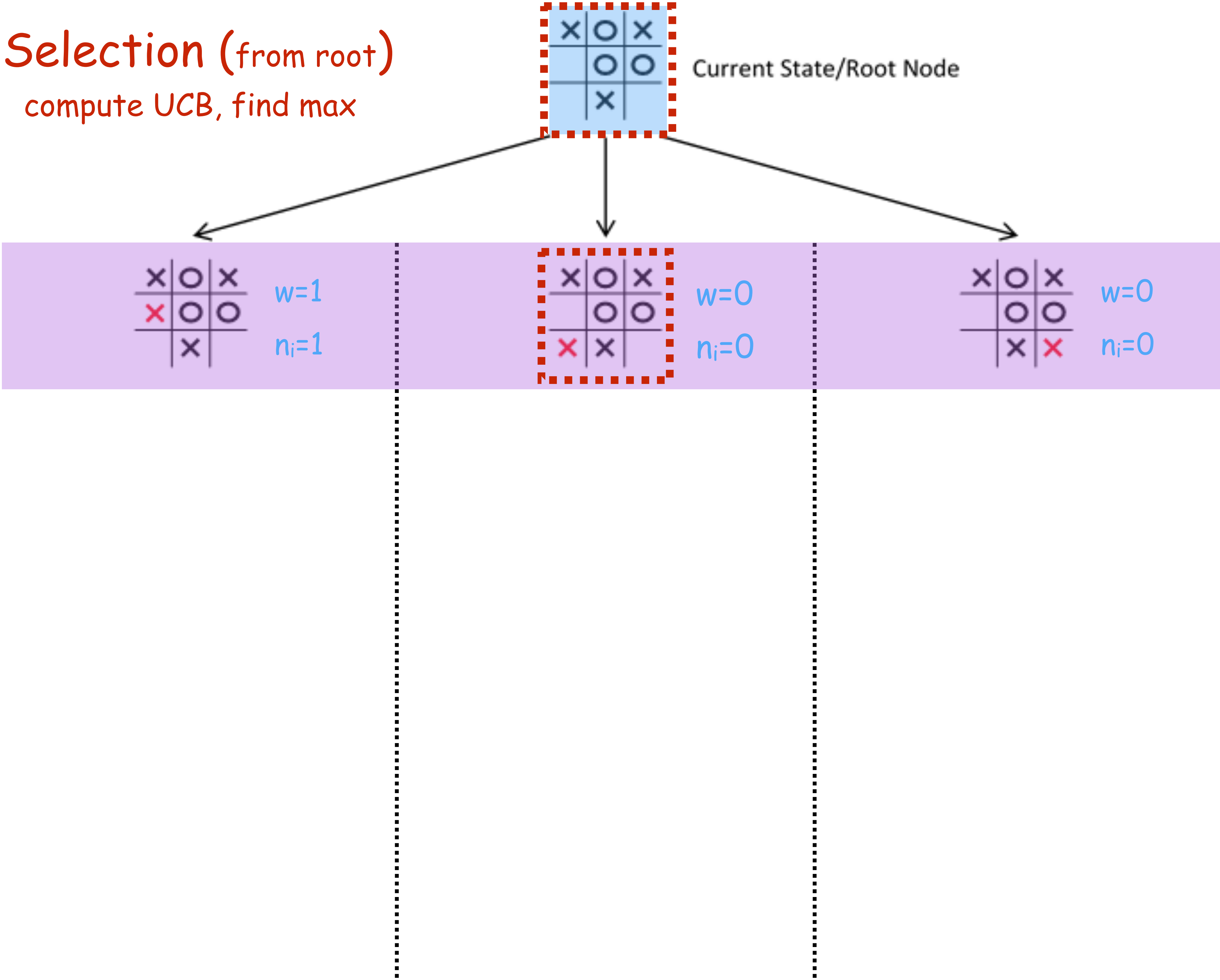
Selection (from root)

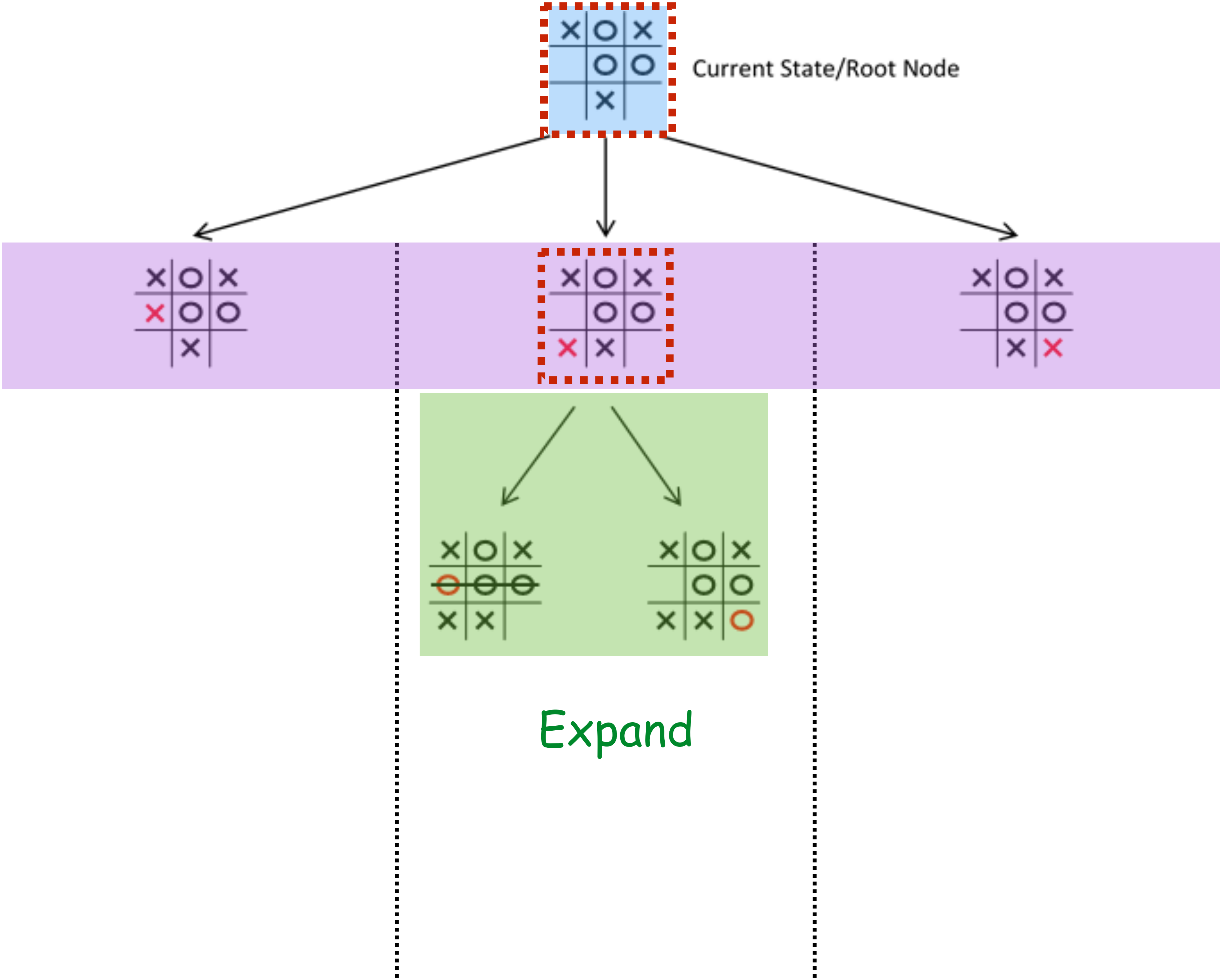
compute UCB, find max

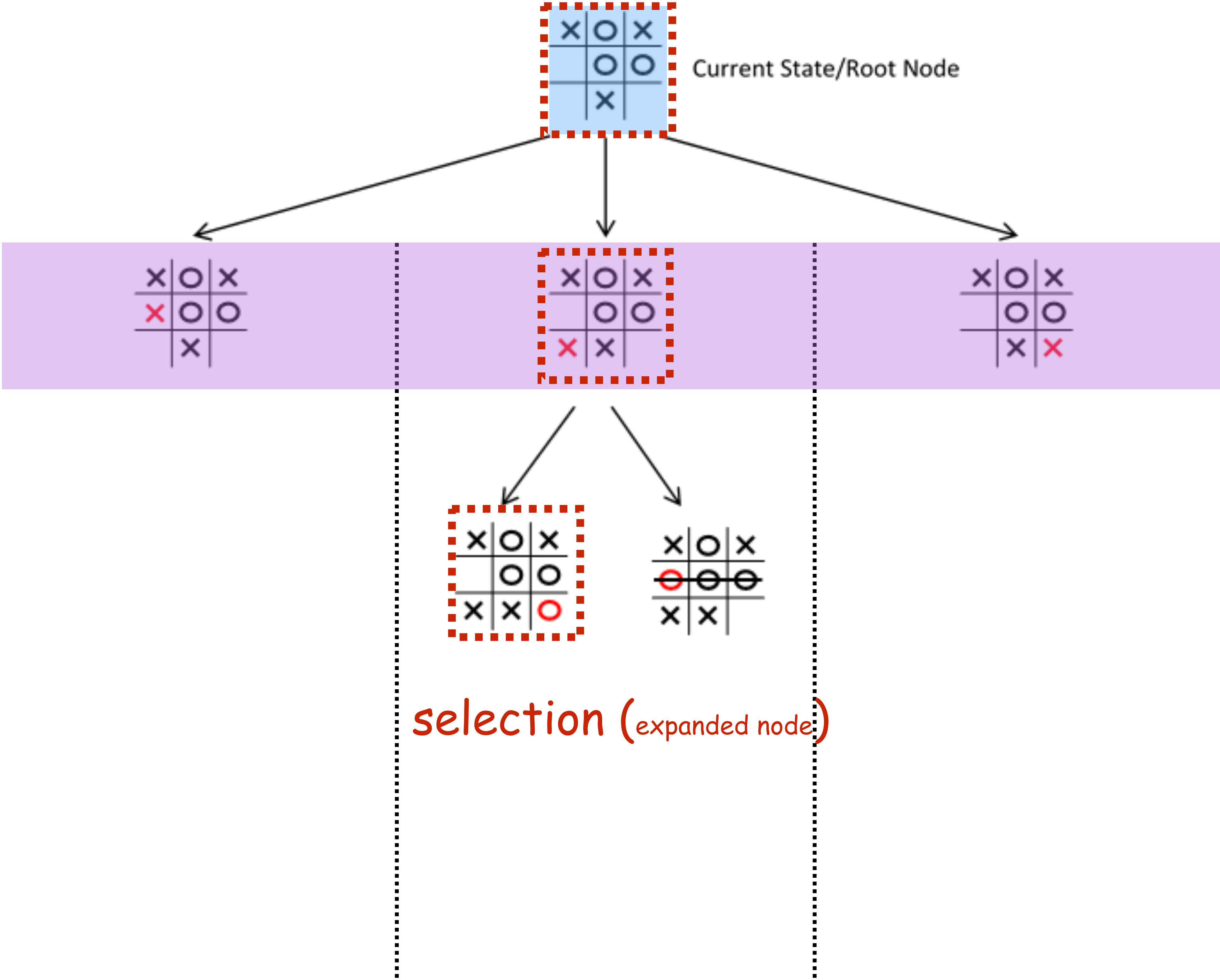


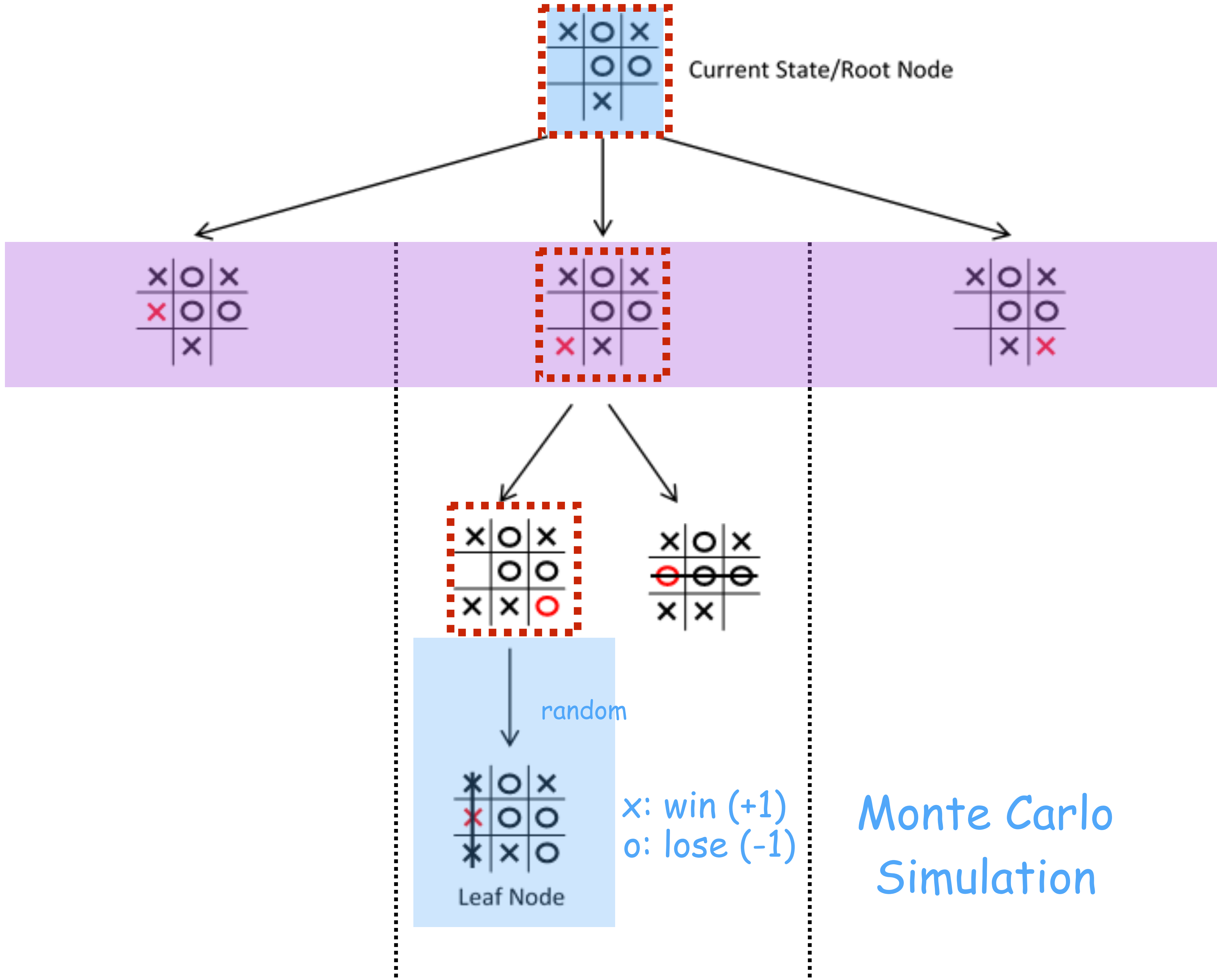
Selection (from root)

compute UCB, find max

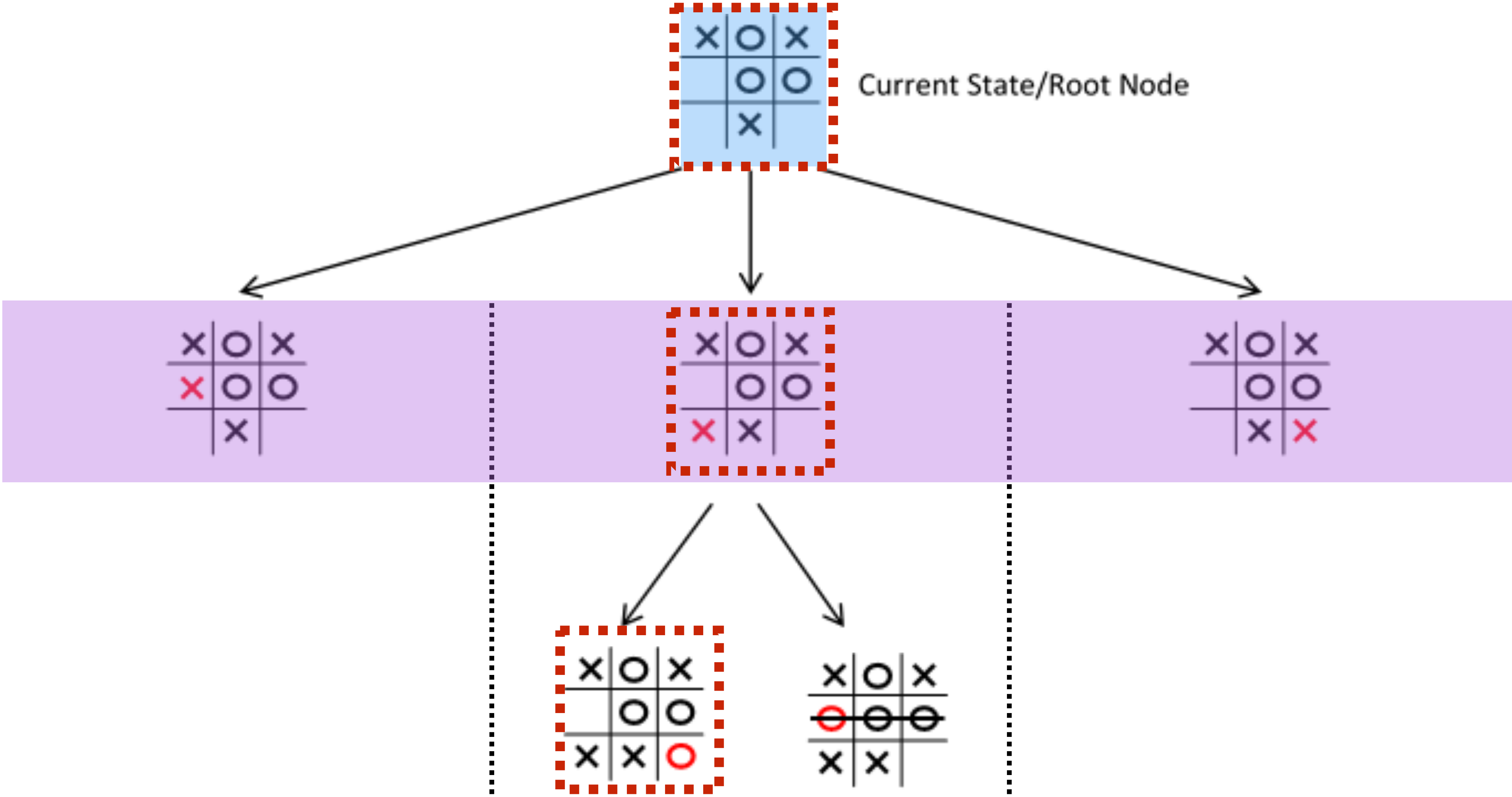



Current State/Root Node
Expand


Current State/Root Node
selection (expanded node)



Current State/Root Node
random
x: win (+1)
o: lose (-1)
Leaf Node
Monte Carlo Simulation

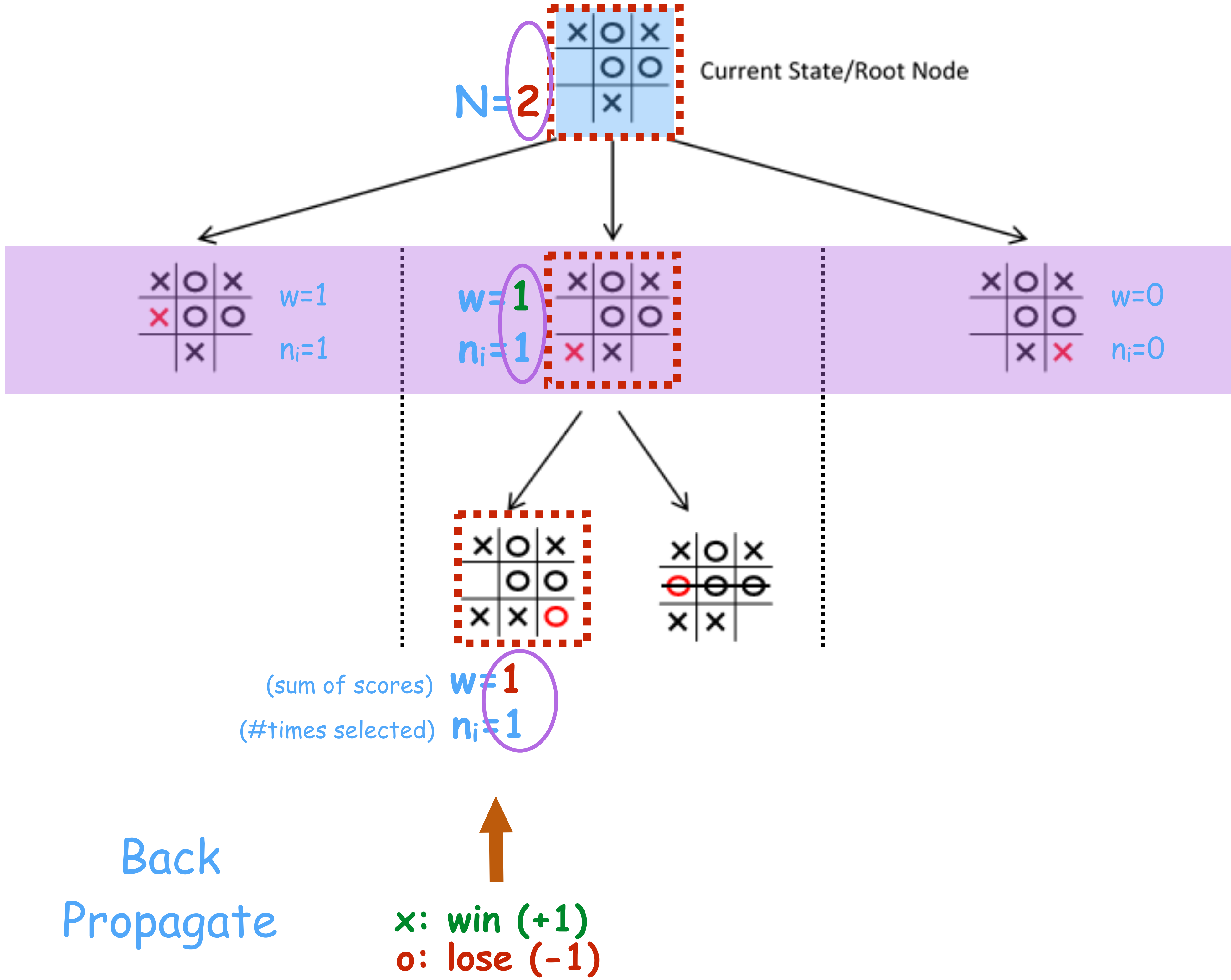


(sum of scores) $w=\mathbf{1}$

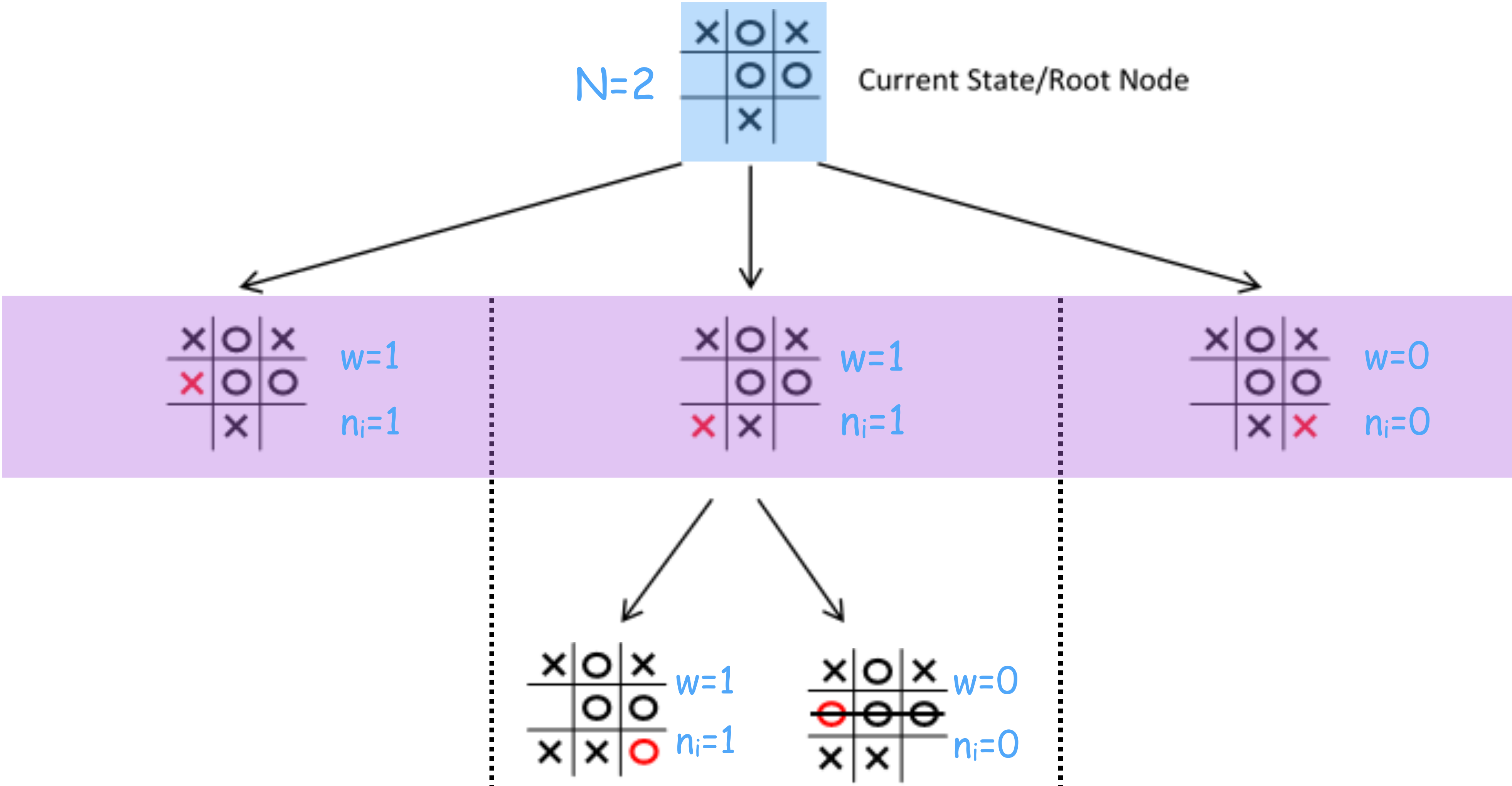
(#times selected) $n_i=1$

x: win (+1)

update node statistics

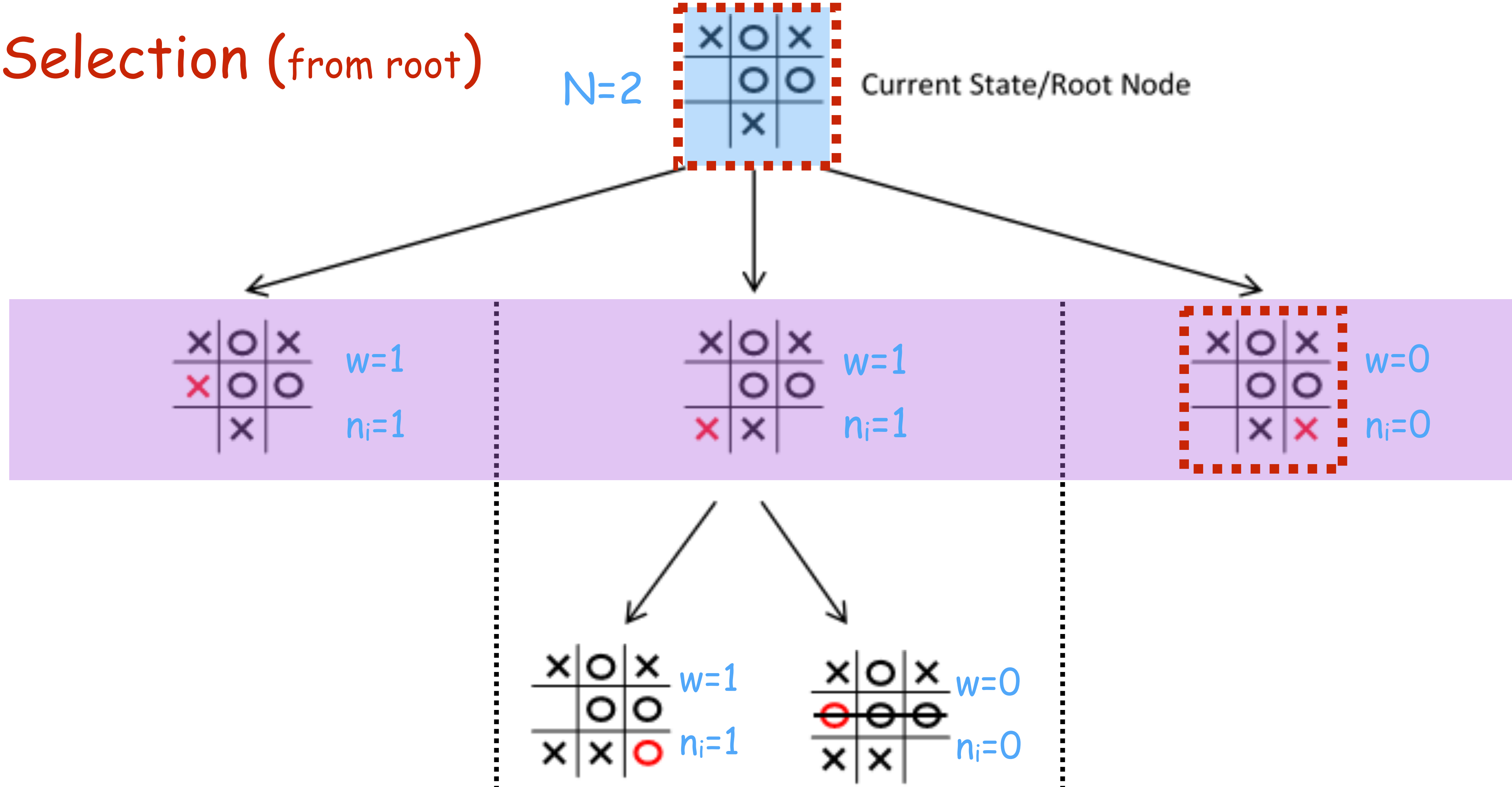


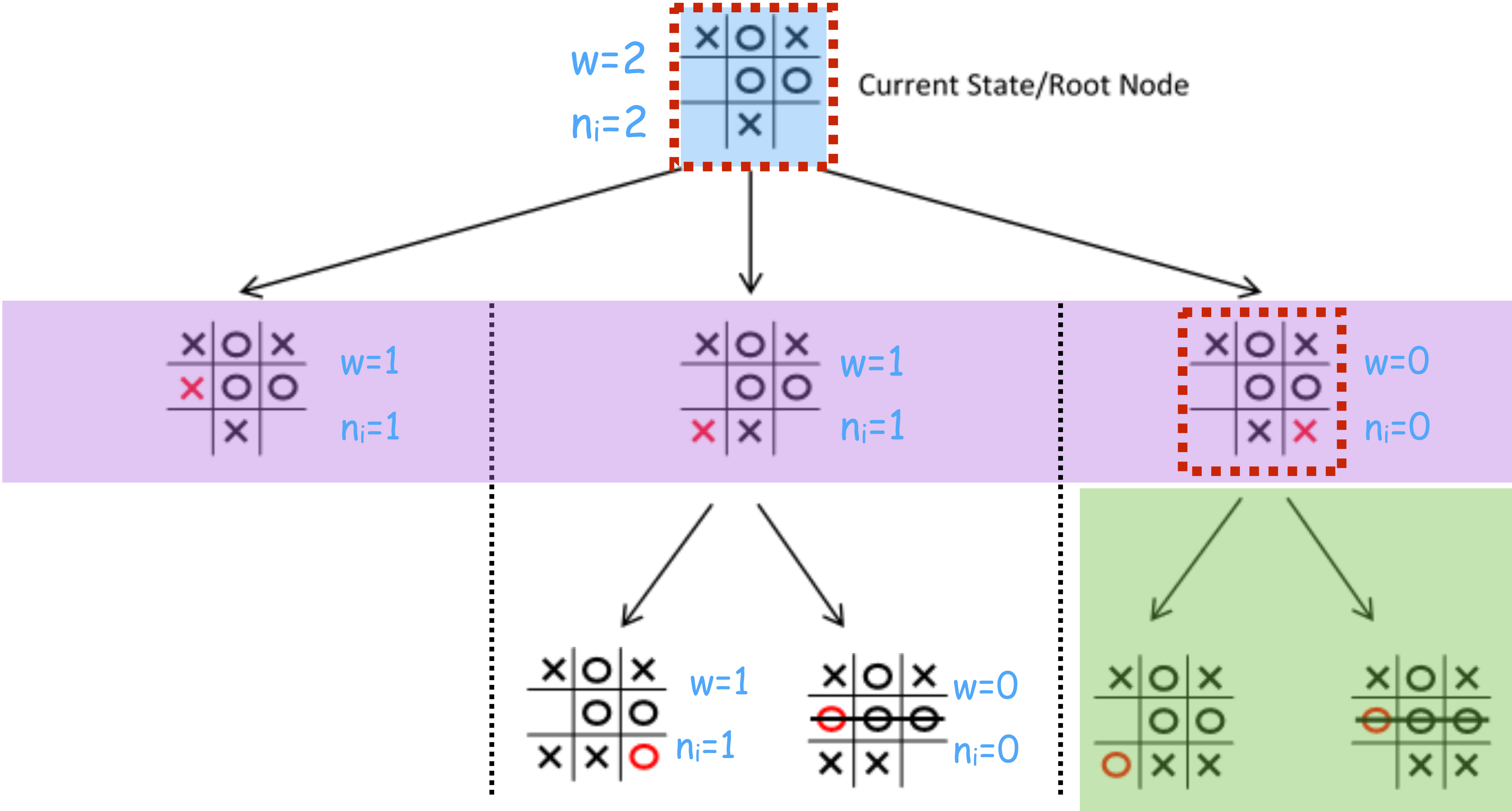
N=2
Current State/Root Node
w=1
n_i=1
w=1
n_i=1
w=0
n_i=0
(sum of scores) w=1
(#times selected) n_i=1
Back Propagate
x: win (+1)
o: lose (-1)



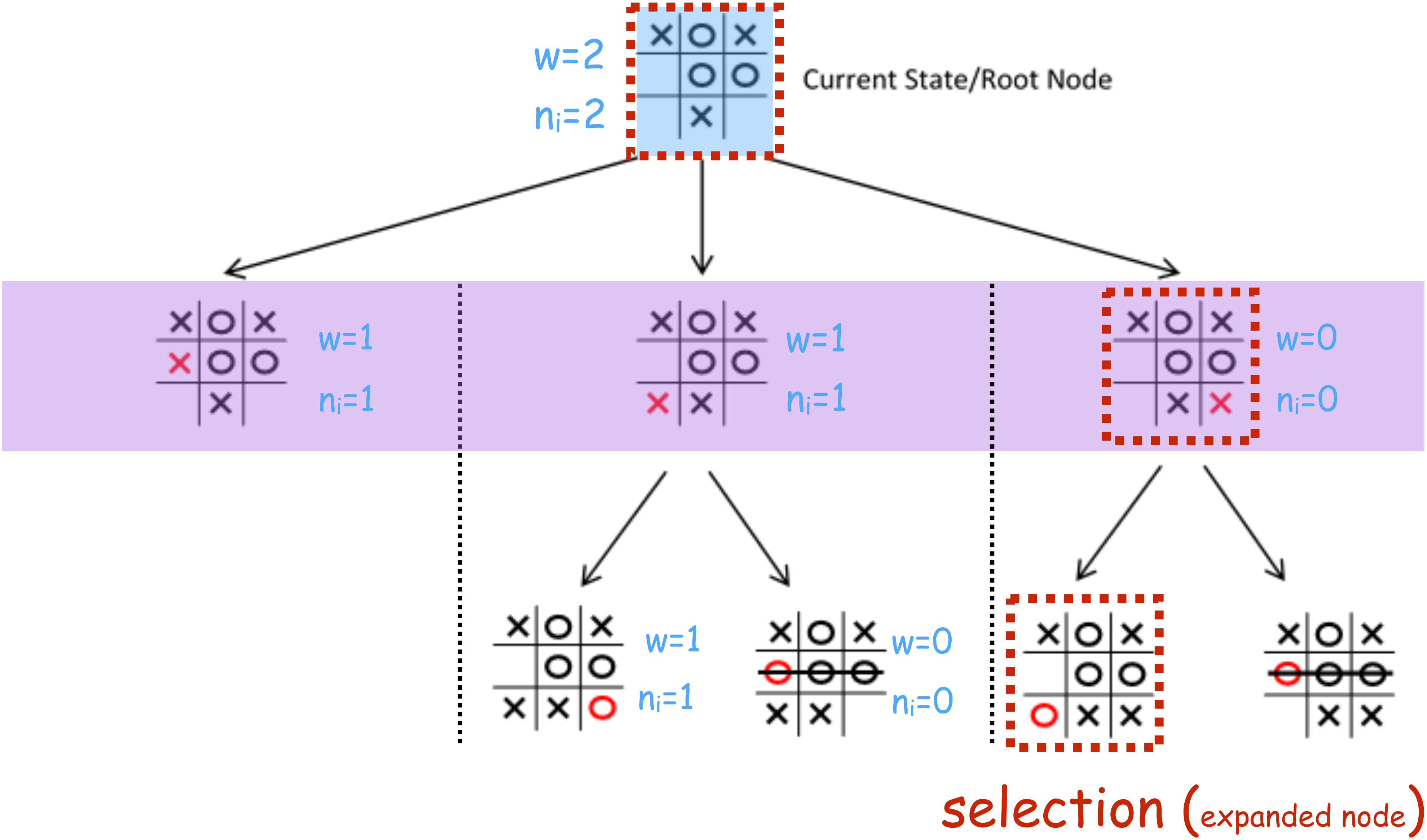
After 2 iterations

Selection (from root)

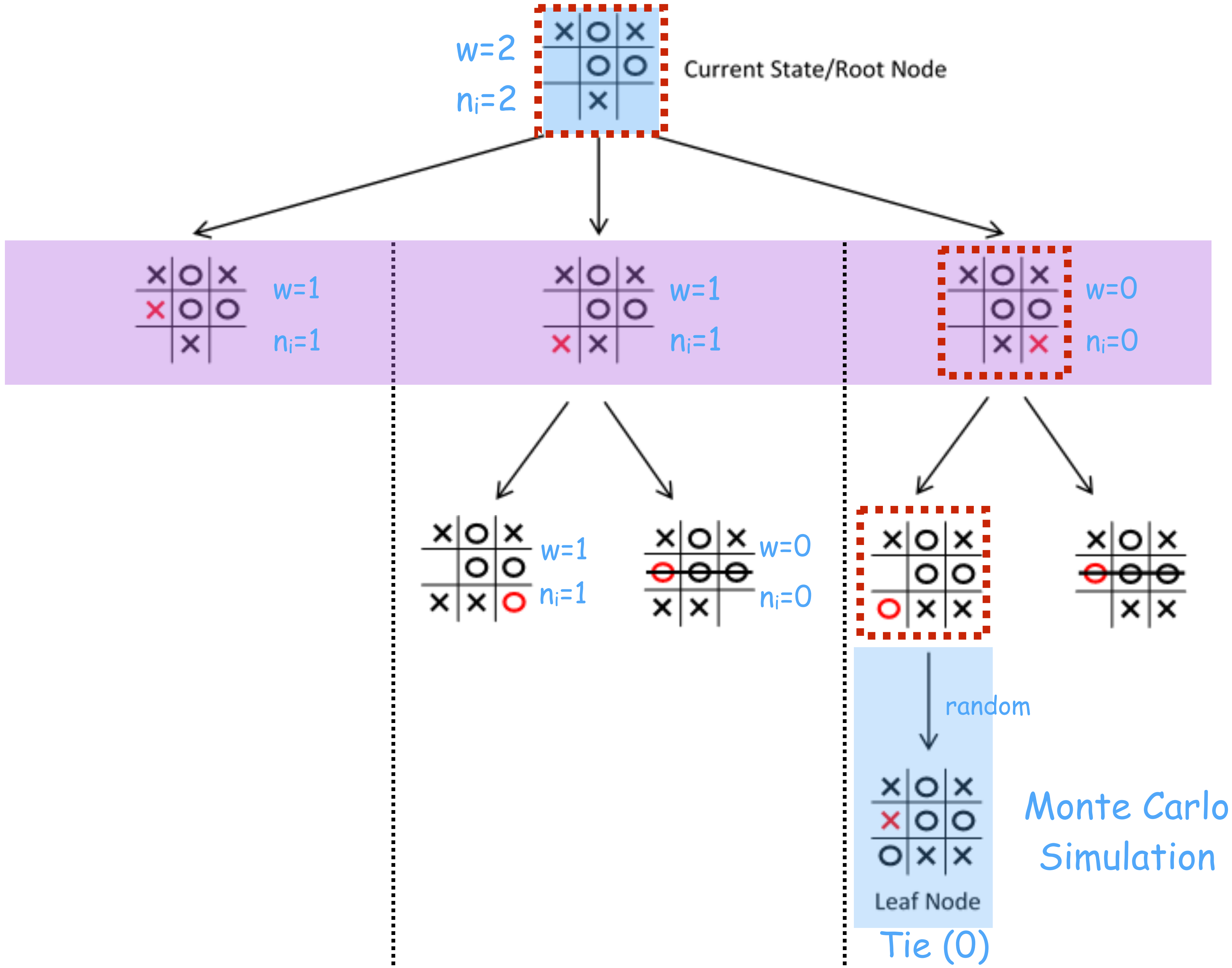



w=2
n_i=2
Current State/Root Node
w=1
n_i=1
w=1
n_i=1
w=0
n_i=0
w=1
n_i=1
w=0
n_i=0

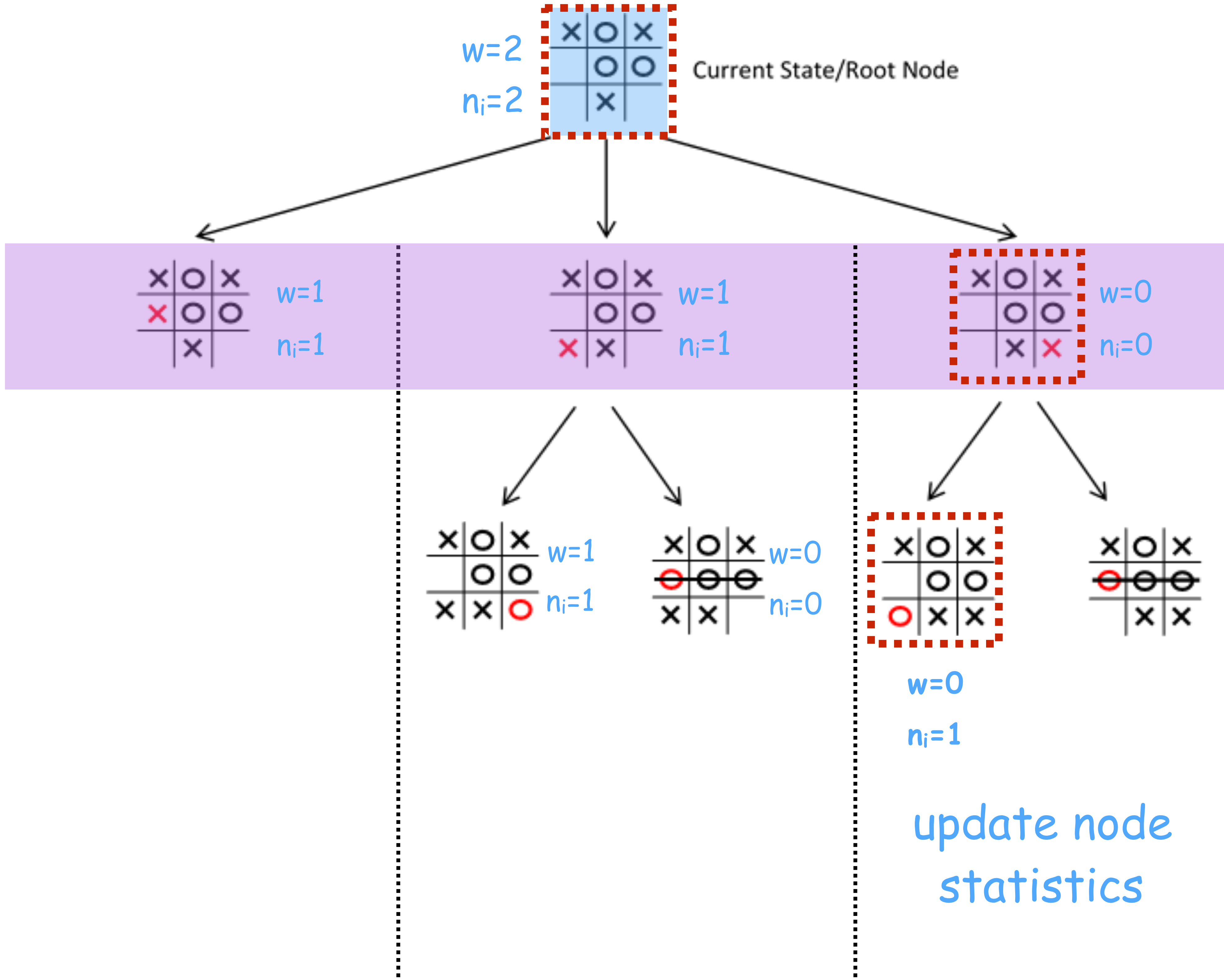
Expand



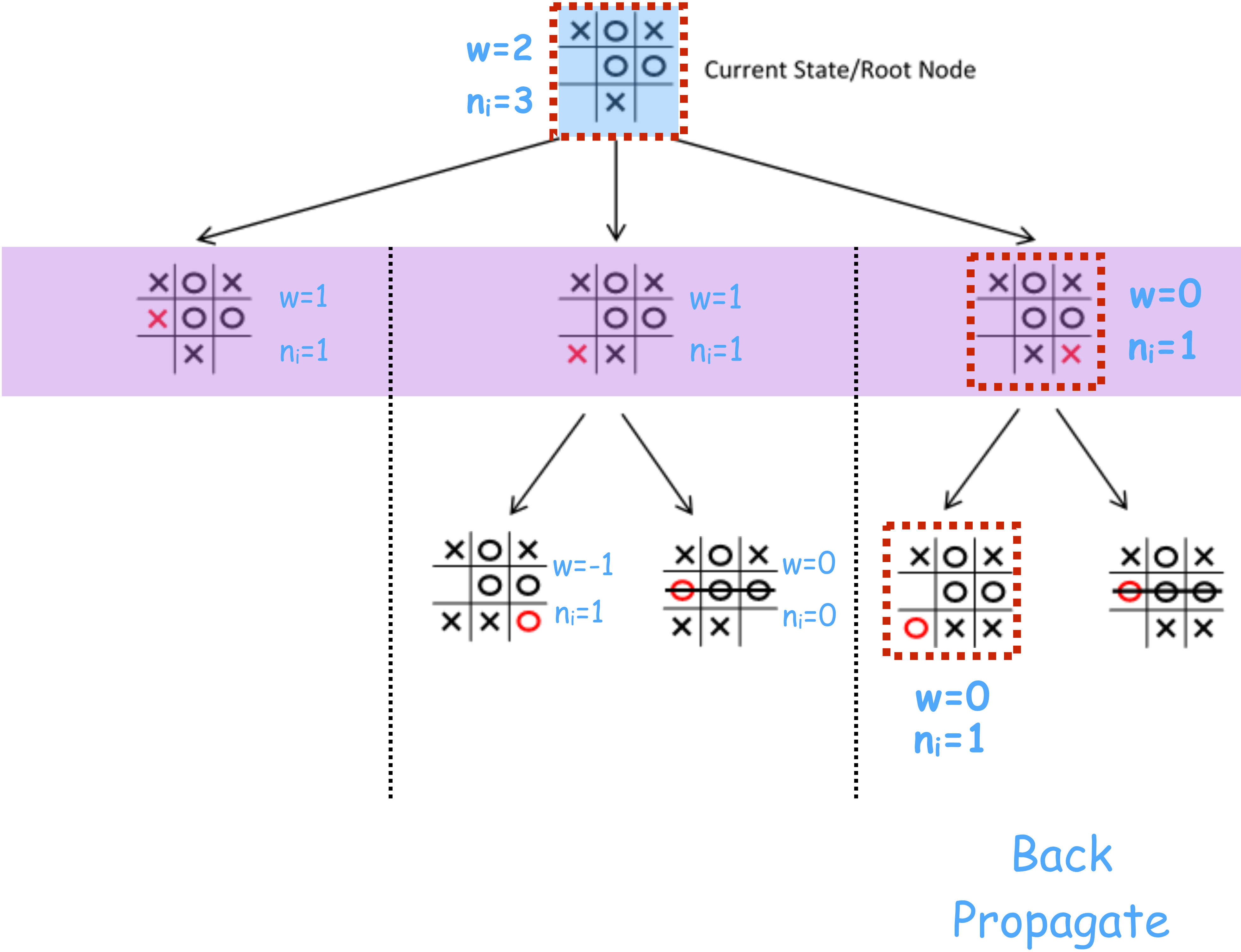
w=2
ni=2
Current State/Root Node
w=1
ni=1
w=1
ni=1
w=0
ni=0
w=1
ni=1
w=0
ni=0
selection (expanded node)



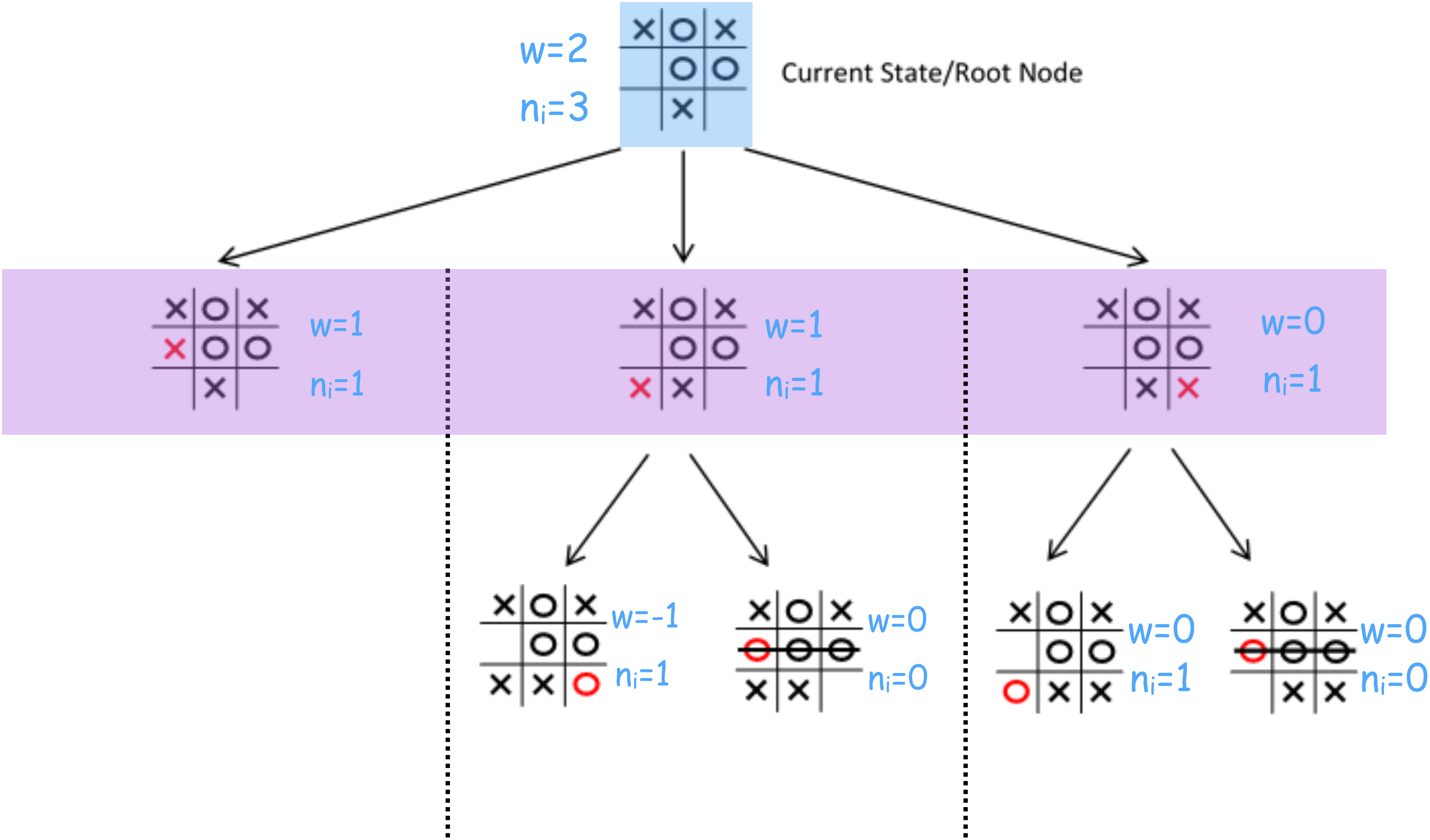
w=2
nᵢ=2
Current State/Root Node
w=1
nᵢ=1
w=1
nᵢ=1
w=0
nᵢ=0
w=1
nᵢ=1
w=0
nᵢ=0
random
Leaf Node
Tie (0)
Monte Carlo Simulation



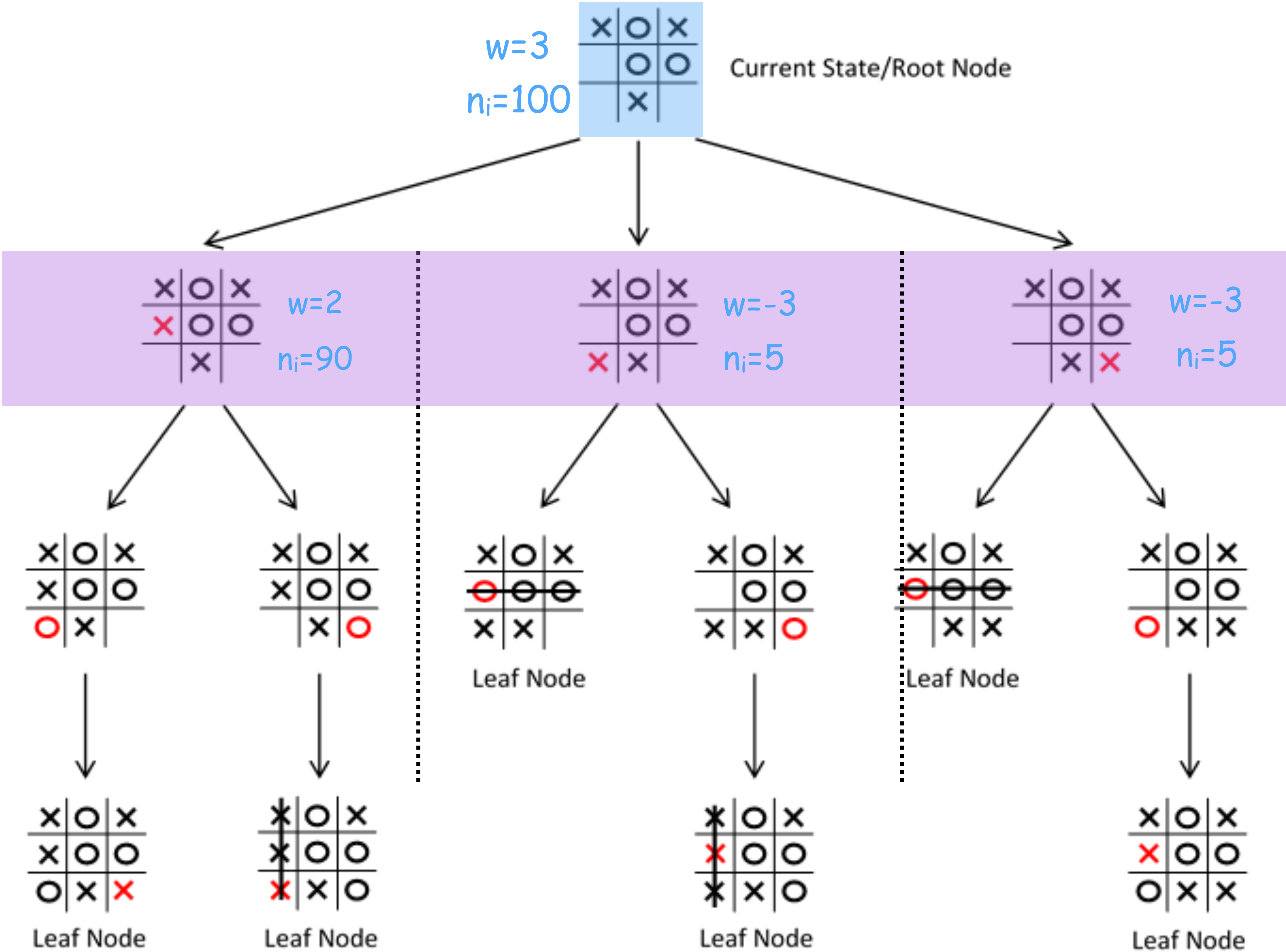
w=2
ni=2
Current State/Root Node
w=1
ni=1
w=1
ni=1
w=0
ni=0
w=1
ni=1
w=0
ni=0
w=0
ni=1
update node statistics



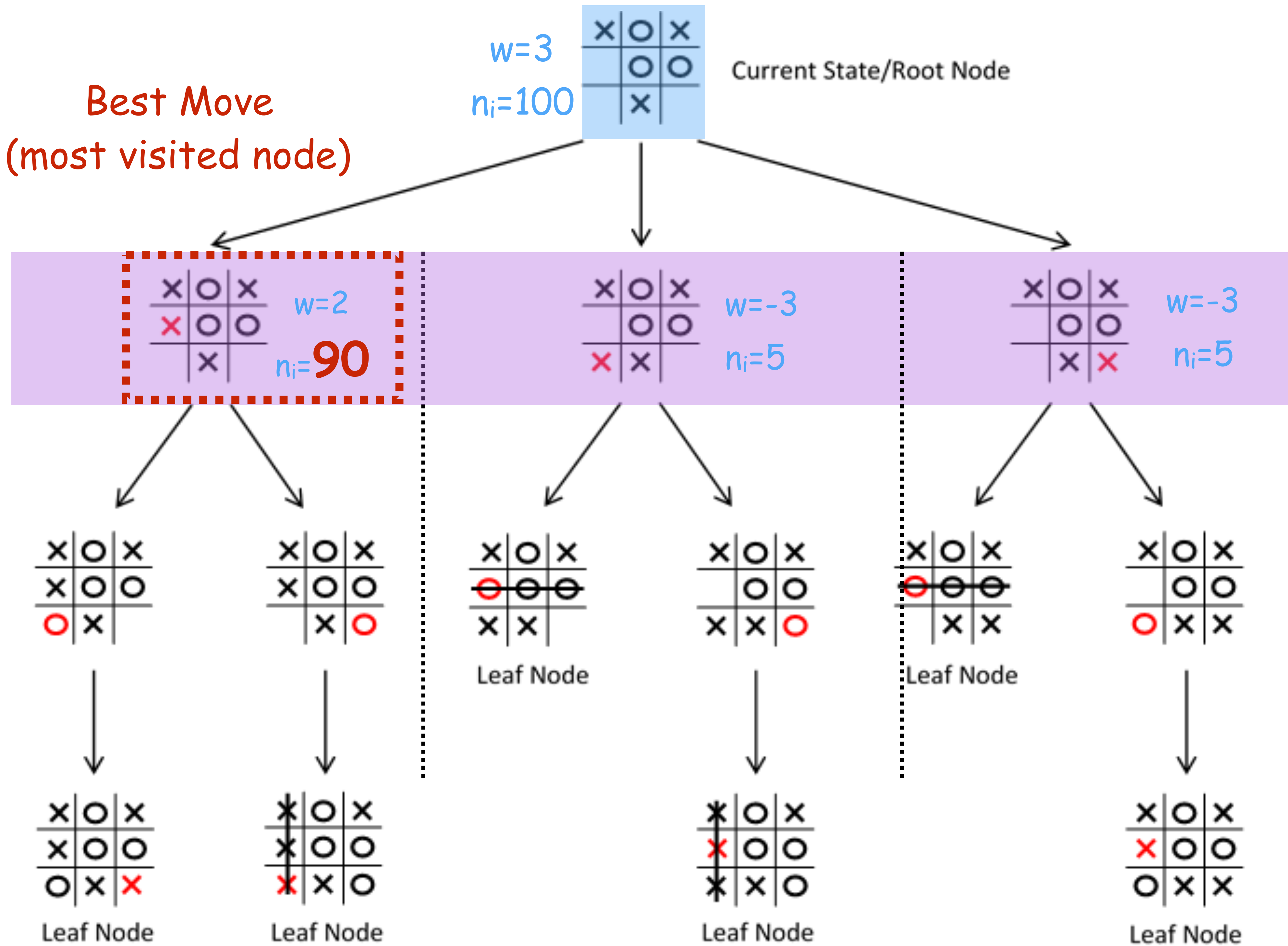
w=2
n_i=3
Current State/Root Node
w=1
n_i=1
w=1
n_i=1
w=0
n_i=1
w=-1
n_i=1
w=0
n_i=0
w=0
n_i=1
Back
Propagate



After 3 iterations

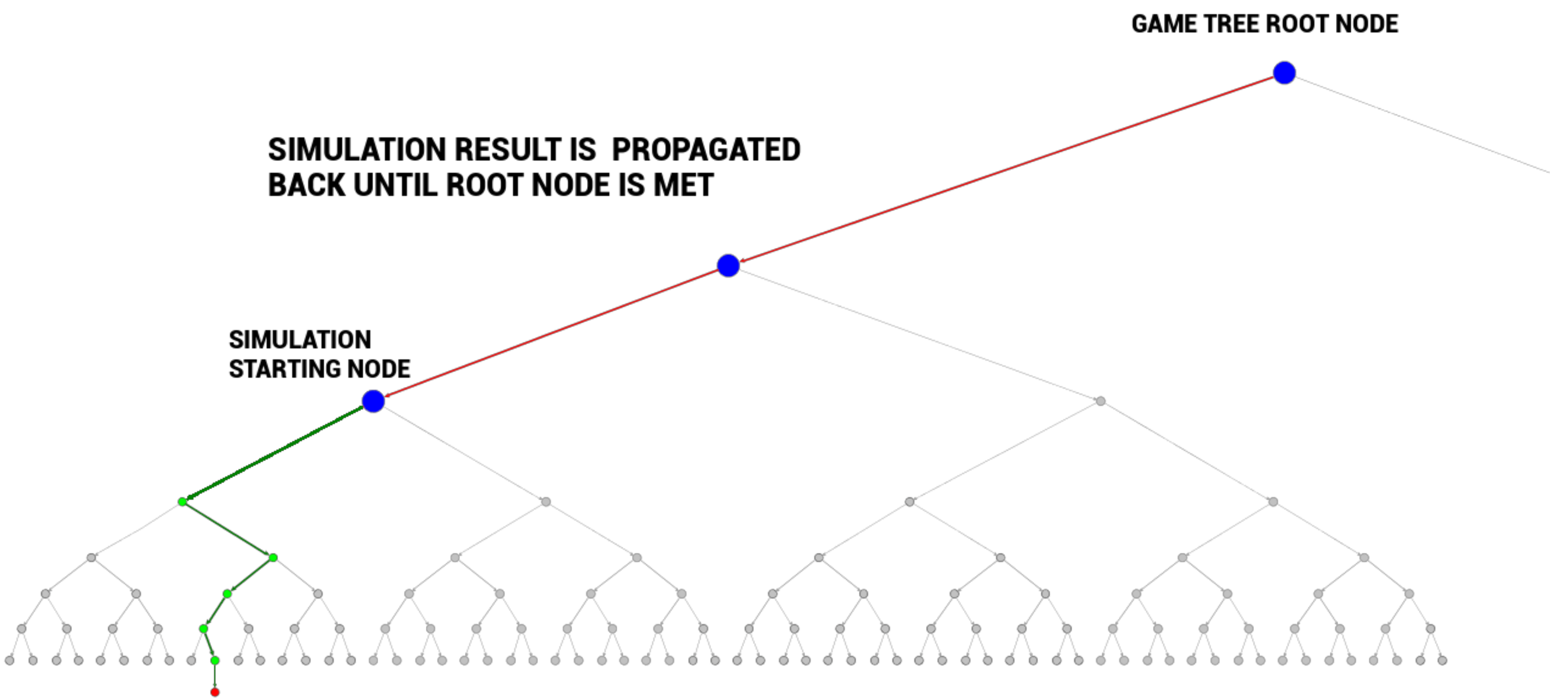


After 100 iterations

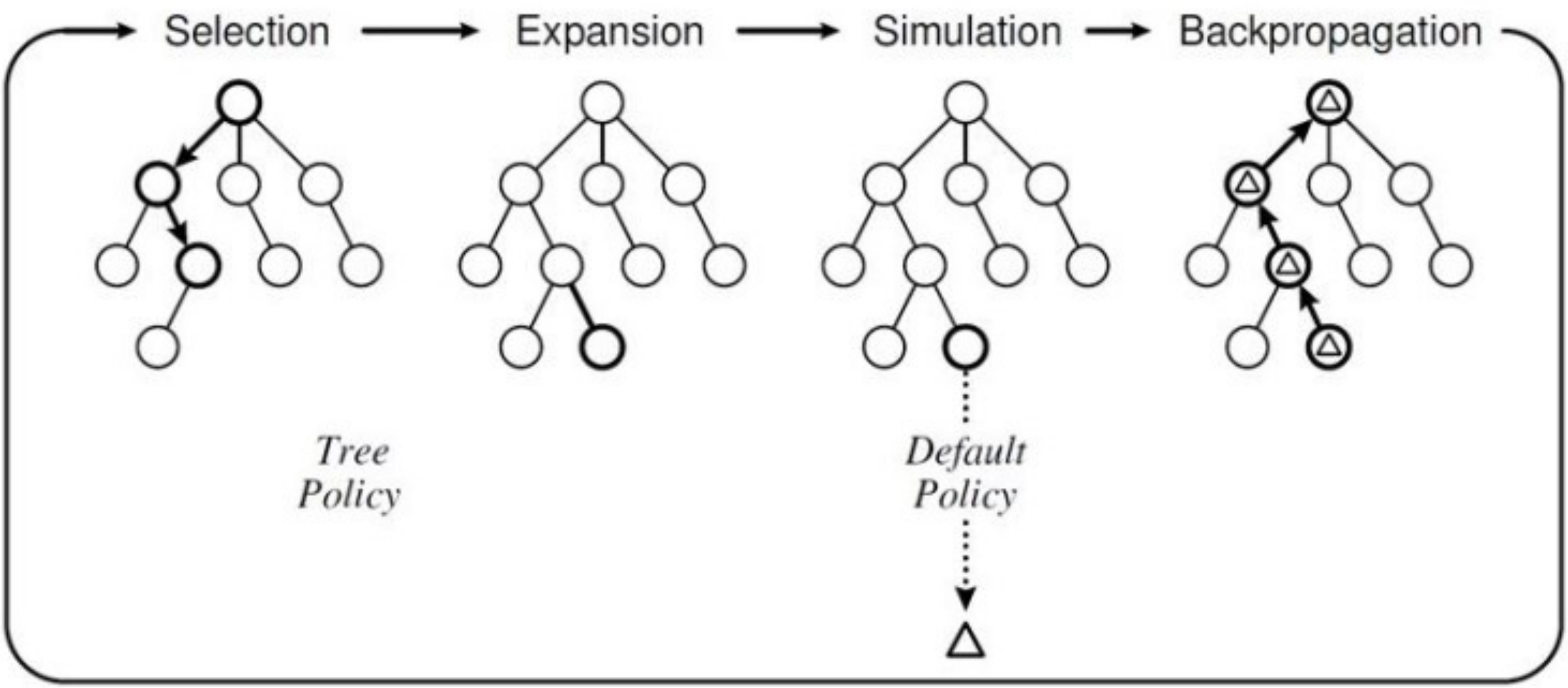


Best Move
(most visited node)
w=3
nᵢ=100
Current State/Root Node
w=2
nᵢ=90
w=-3
nᵢ=5
w=-3
nᵢ=5
Leaf Node
Leaf Node
Leaf Node
Leaf Node
Leaf Node
Leaf Node

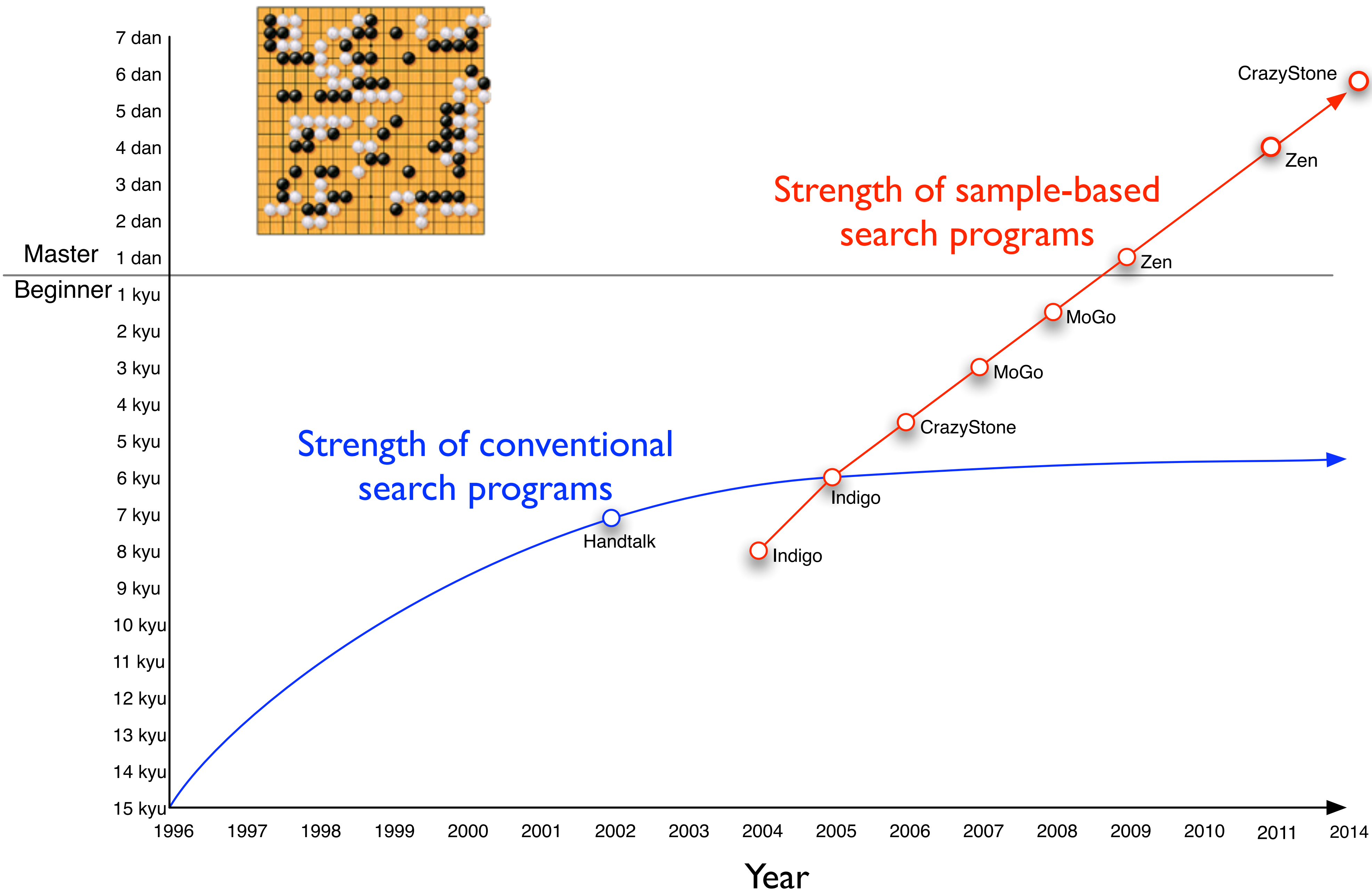
Monte Carlo Tree Search (2005)



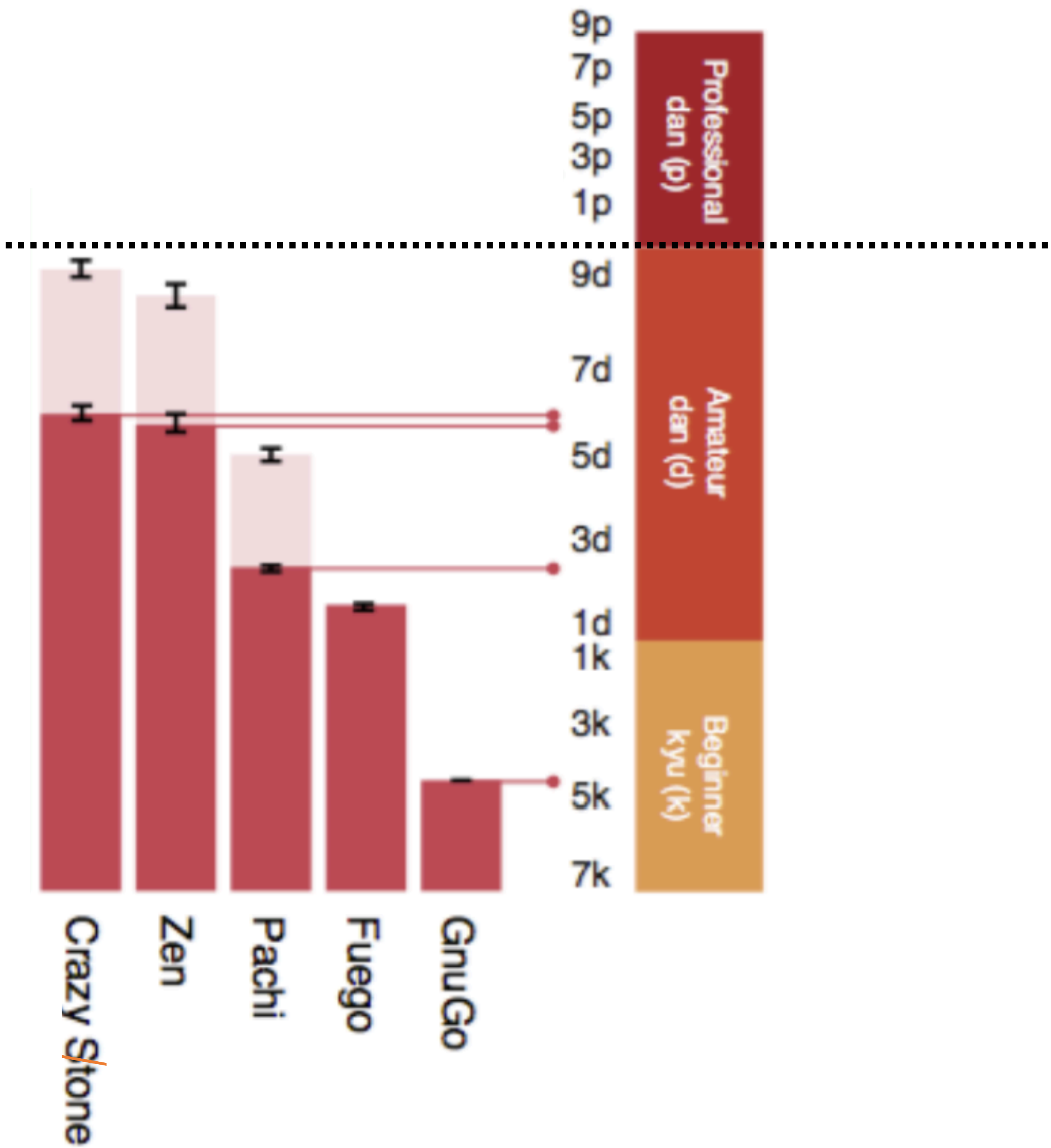
Monte-Carlo Tree-Search



Steady, exponential improvement (since MCTS, 2005) in the strength of the best computer Go programs

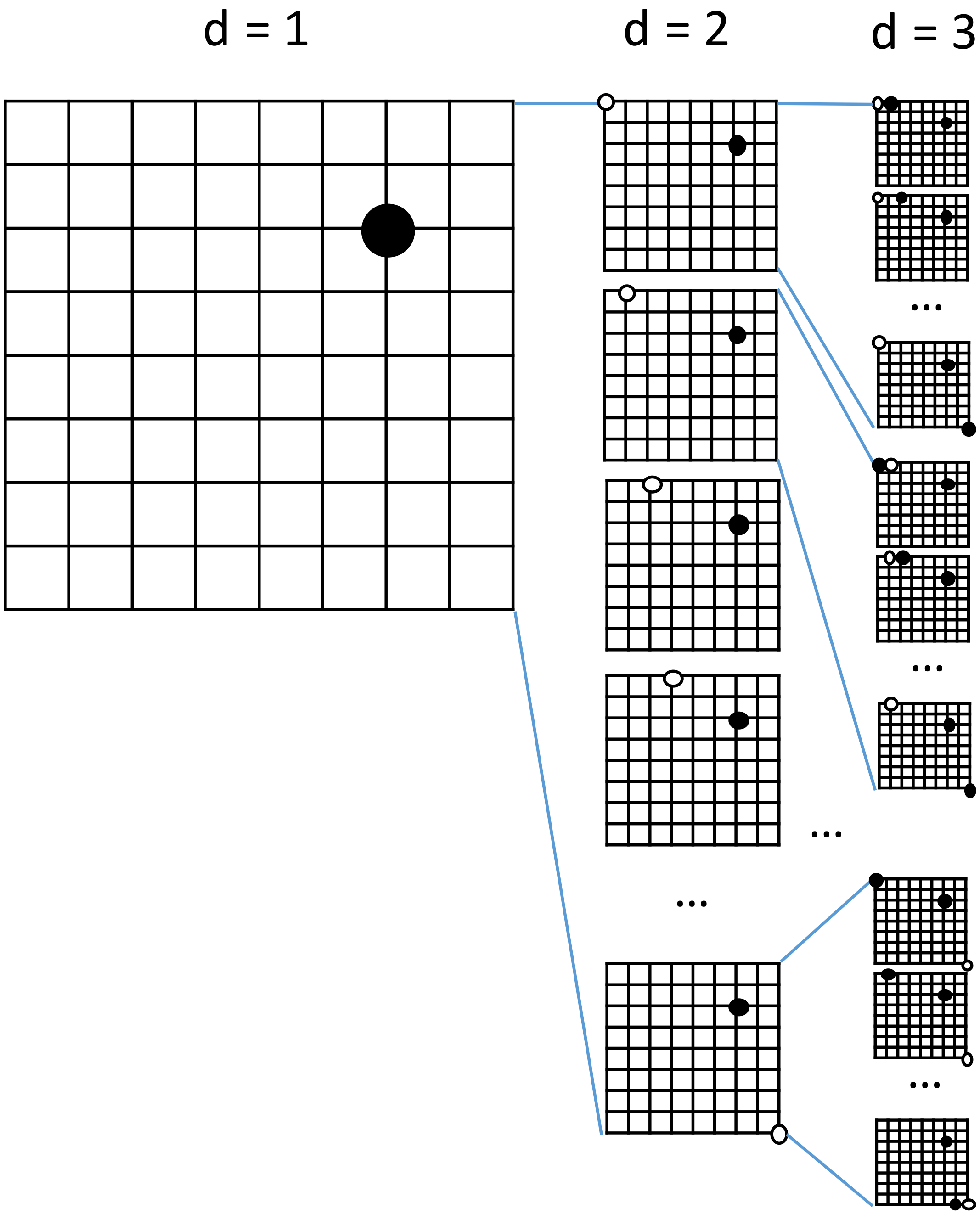


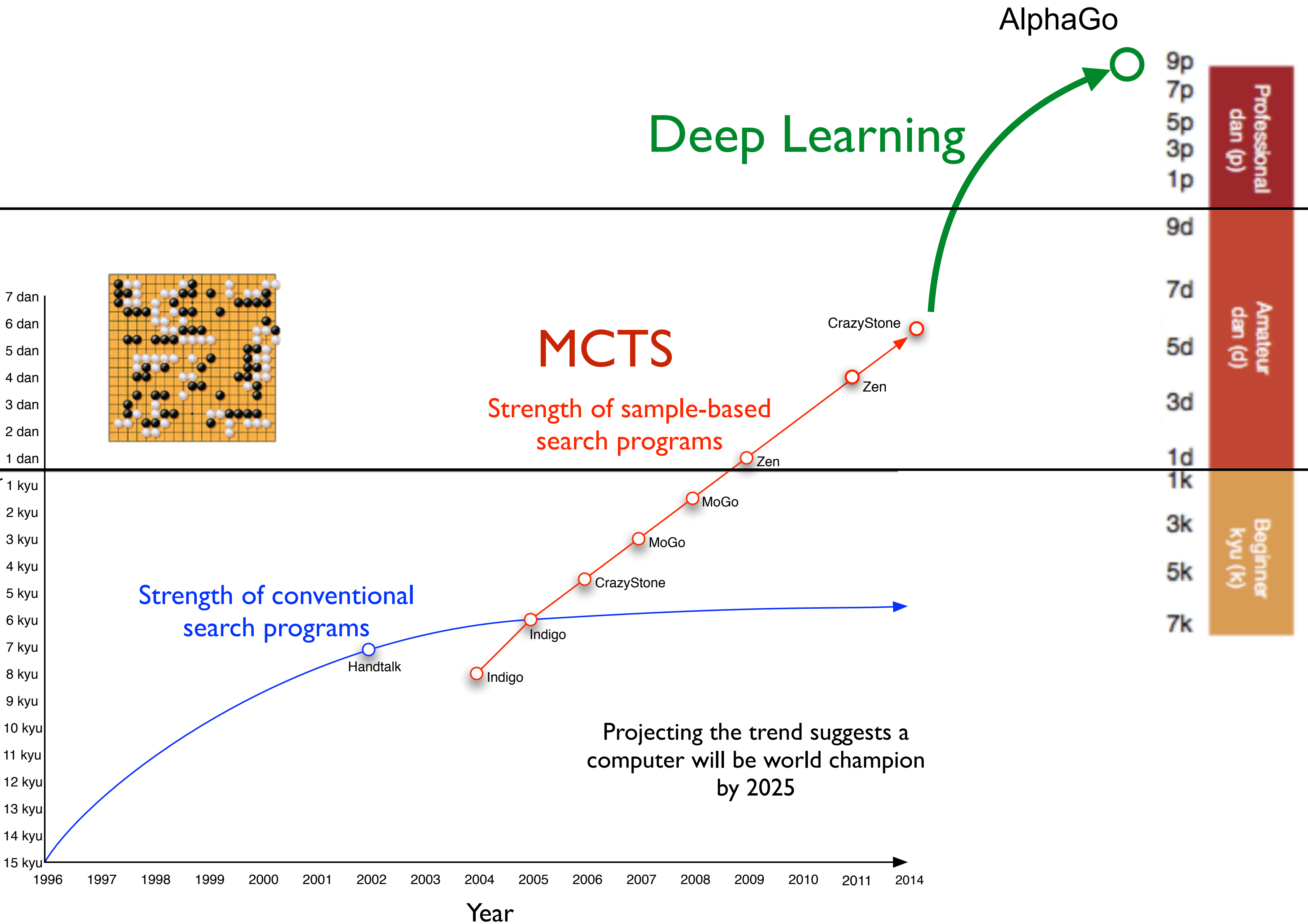
if 4 handicap moves
are given



Professional Player vs MCTS

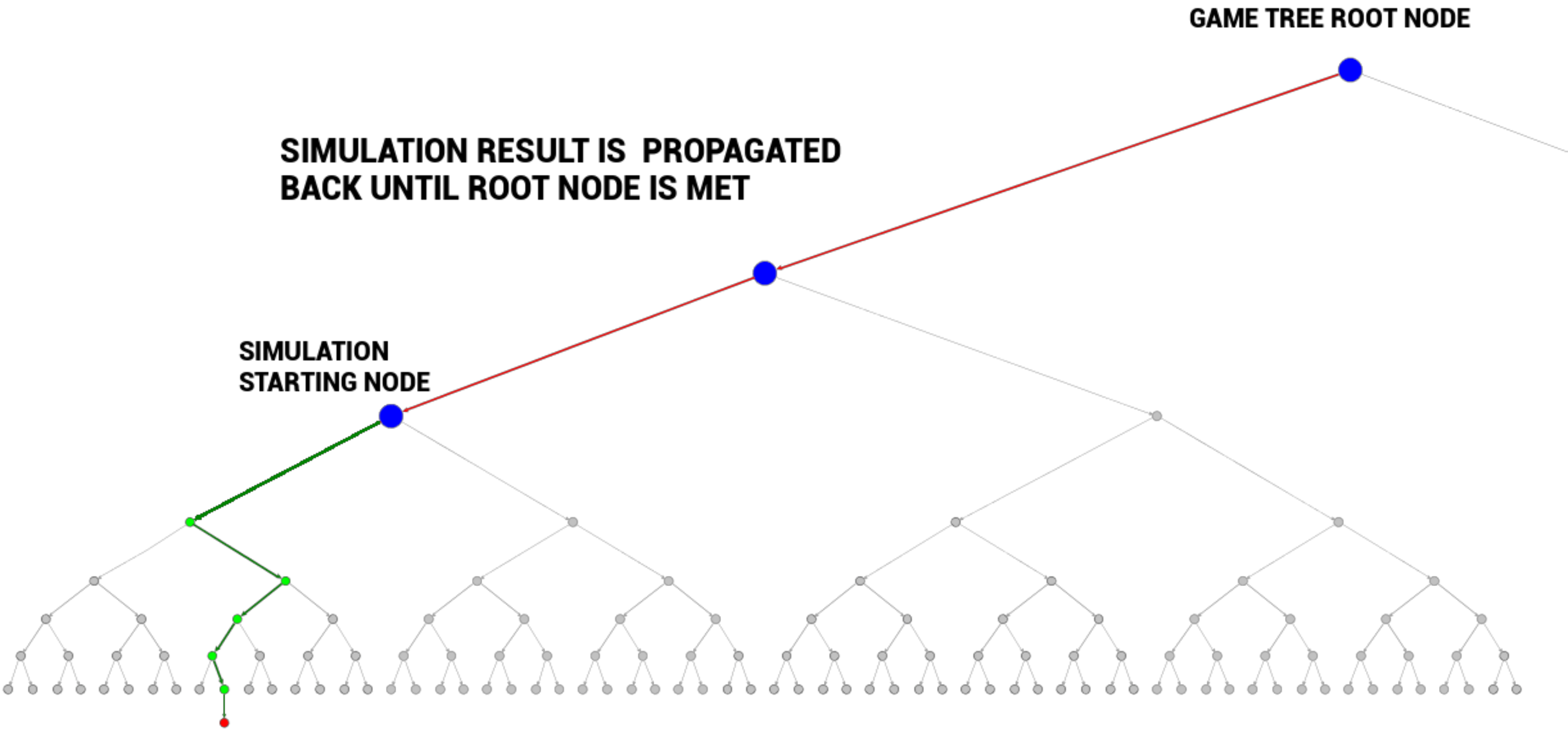
The Huge Tree

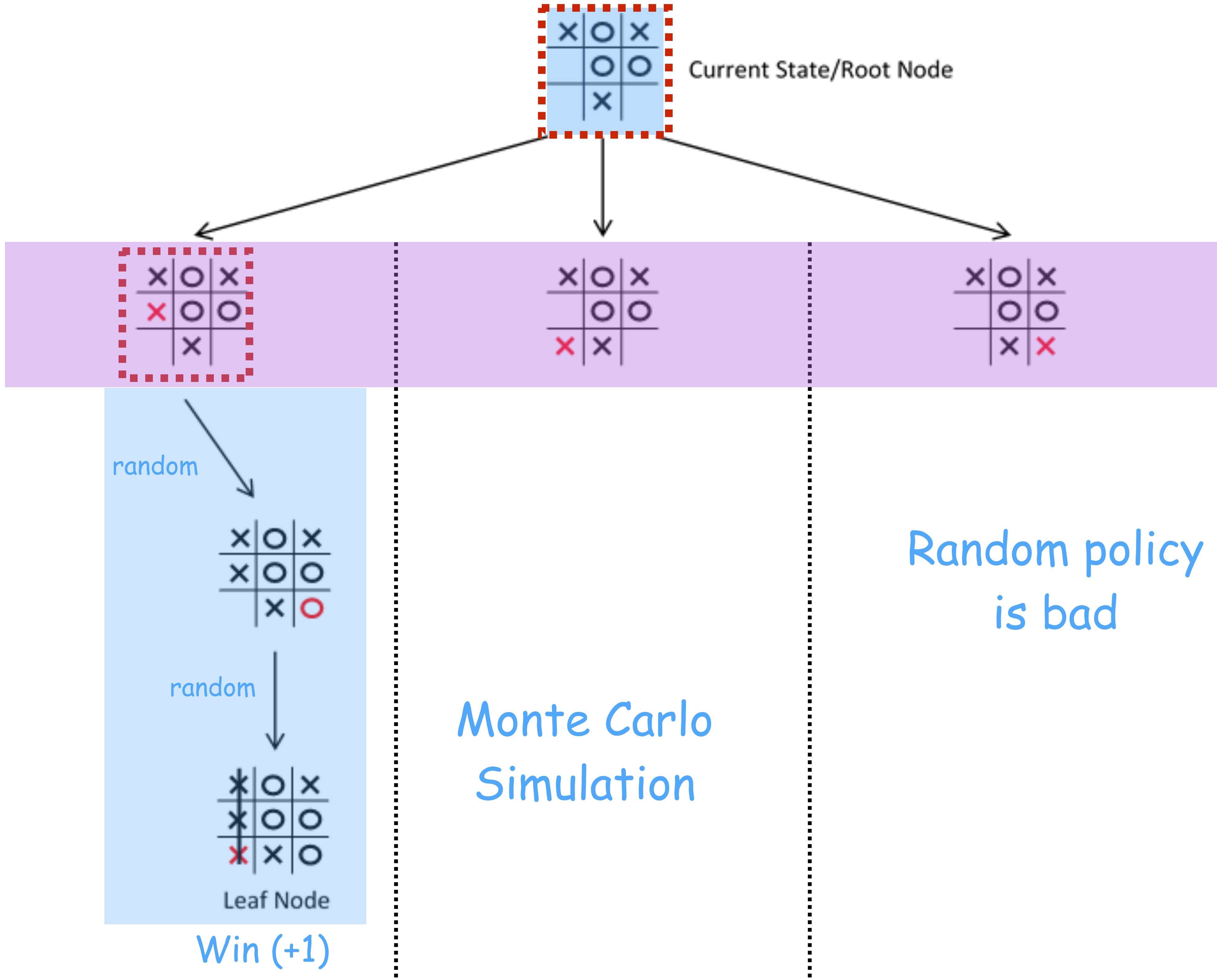




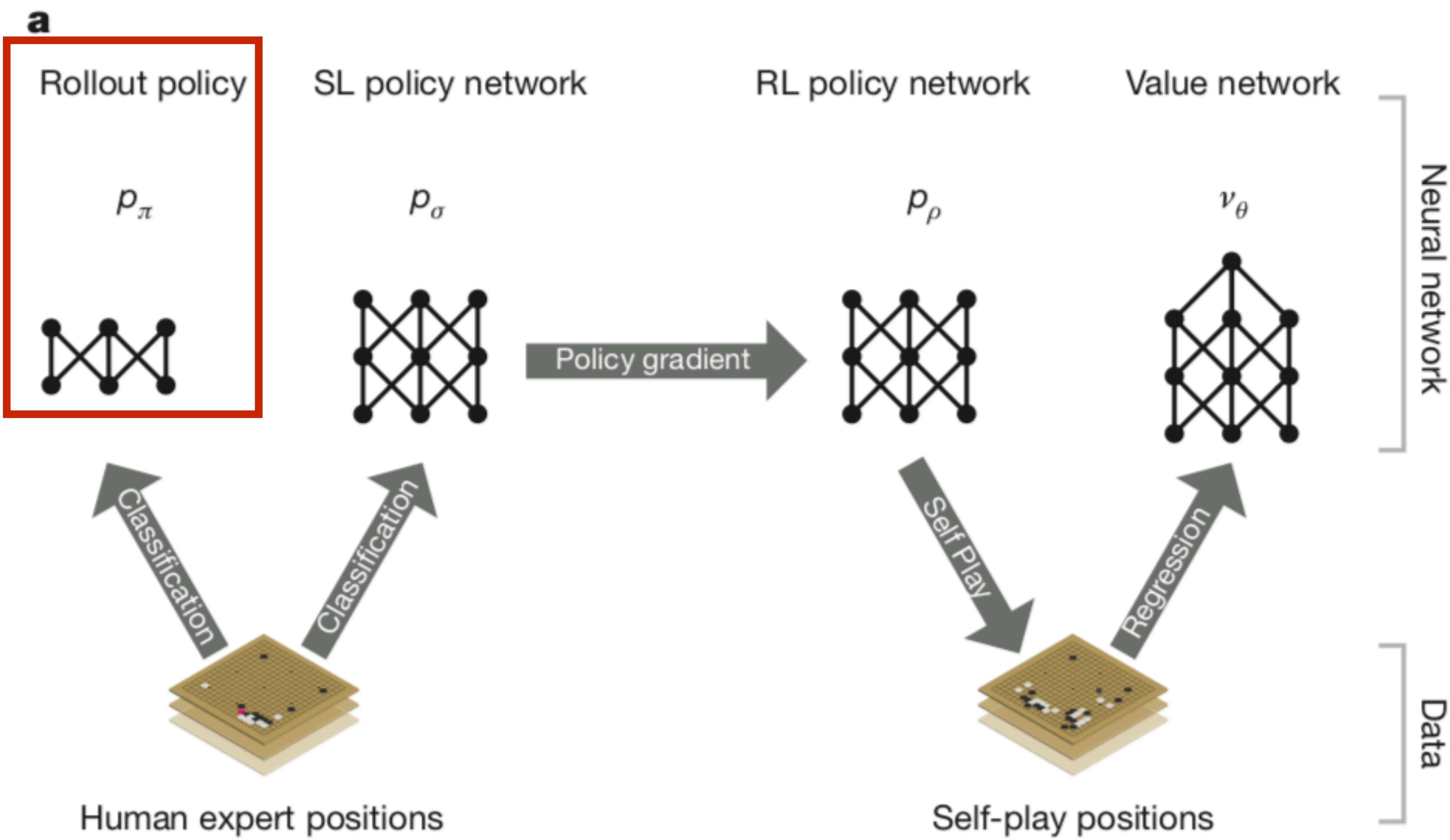
AlphaGo
Deep Learning
MCTS
Strength of sample-based search programs
Strength of conventional search programs
Projecting the trend suggests a computer will be world champion by 2025
CrazyStone
Zen
Zen
MoGo
MoGo
CrazyStone
Indigo
Indigo
Handtalk
7 dan
6 dan
5 dan
4 dan
3 dan
2 dan
1 dan
1 kyu
2 kyu
3 kyu
4 kyu
5 kyu
6 kyu
7 kyu
8 kyu
9 kyu
10 kyu
11 kyu
12 kyu
13 kyu
14 kyu
15 kyu
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2014
Year
9p
7p
5p
3p
1p
9d
7d
5d
3d
1d
1k
3k
5k
7k
Professional dan (p)
Amateur dan (d)
Beginner kyu (k)

What to improve in MCTS ?



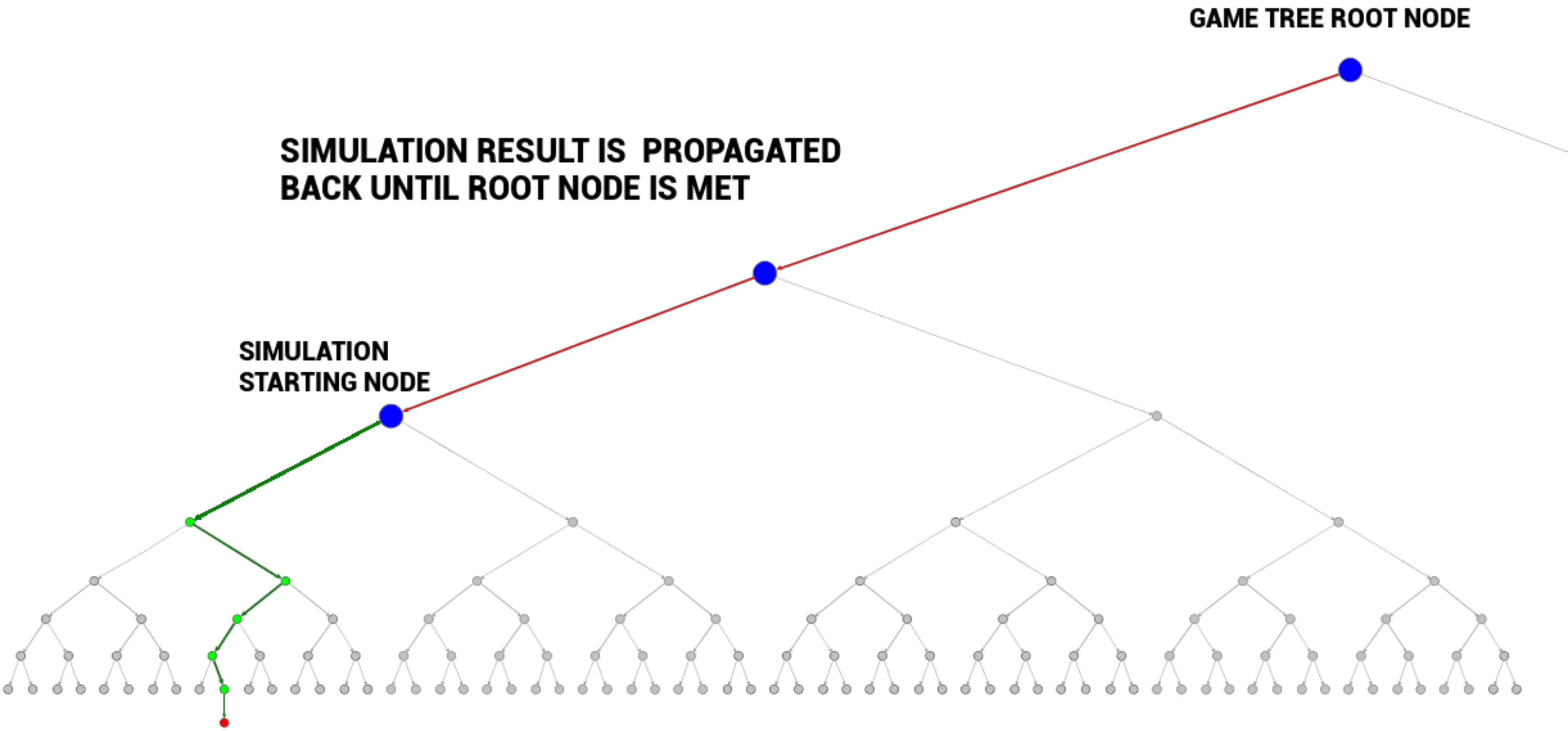

Current State/Root Node
random
random
Leaf Node
Win (+1)
Monte Carlo Simulation
Random policy is bad

AlphaGo



a
Rollout policy
SL policy network
RL policy network
Value network
p_π
p_σ
p_ρ
v_θ
Policy gradient
Neural network
Classification
Classification
Self Play
Regression
Data
Human expert positions
Self-play positions

Selection

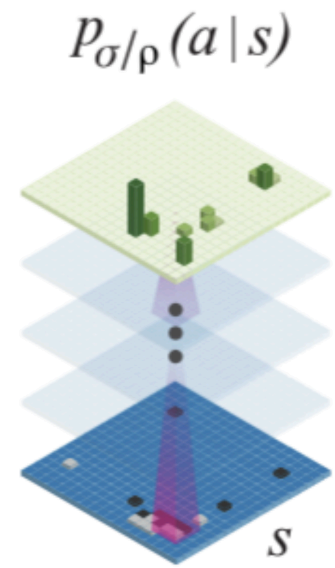


GAME TREE ROOT NODE
SIMULATION RESULT IS PROPAGATED BACK UNTIL ROOT NODE IS MET
SIMULATION STARTING NODE

Reducing Search Space

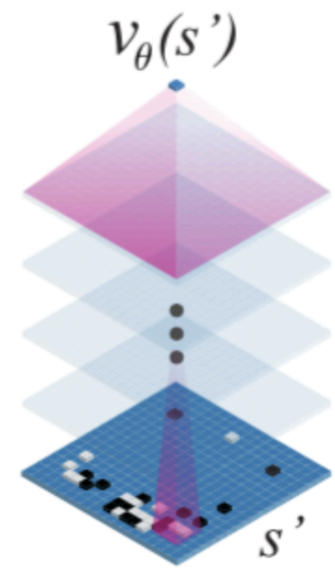
1. **Reducing "action candidates" (Breadth Reduction)**

   Policy Network



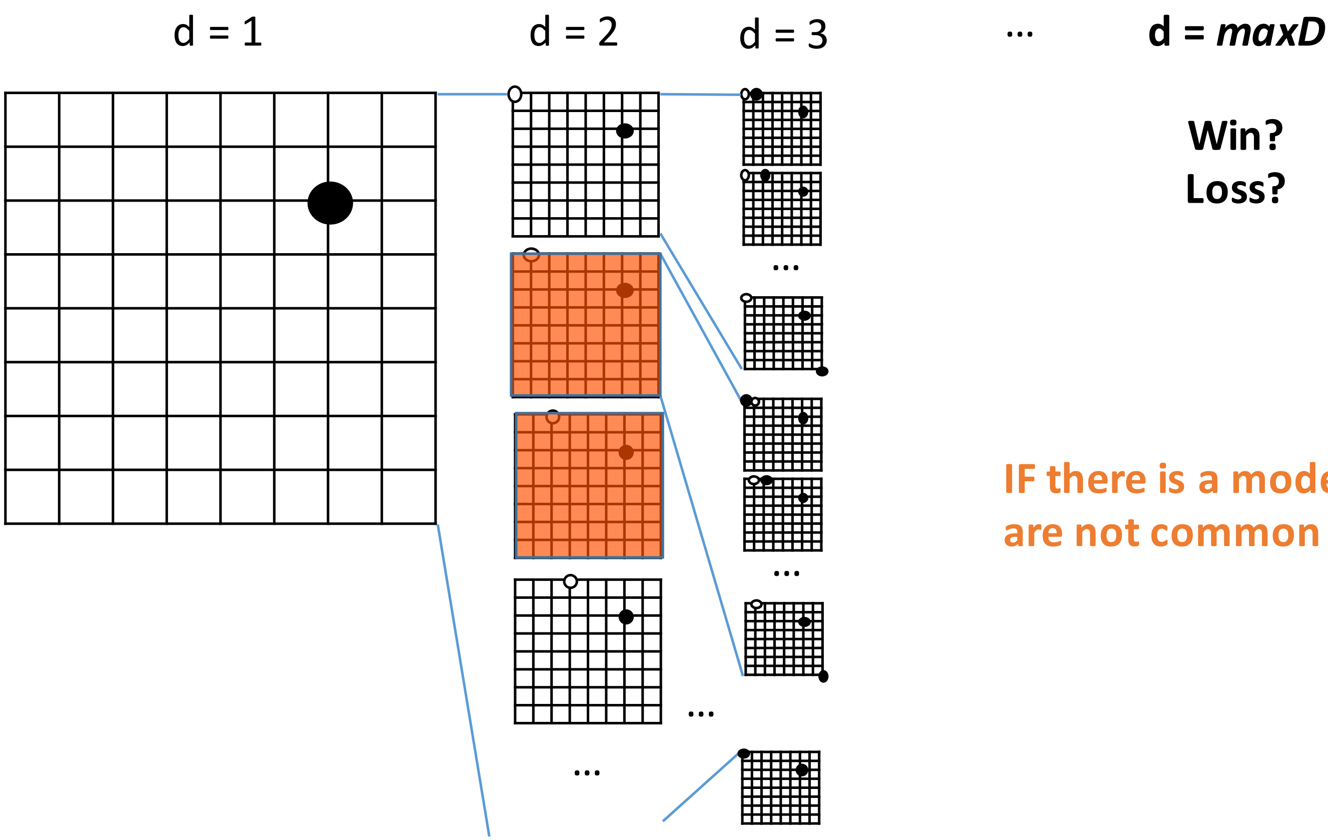
2. **Board Evaluation (Depth Reduction)**

   Value Network



Reducing num of Choices in each decision

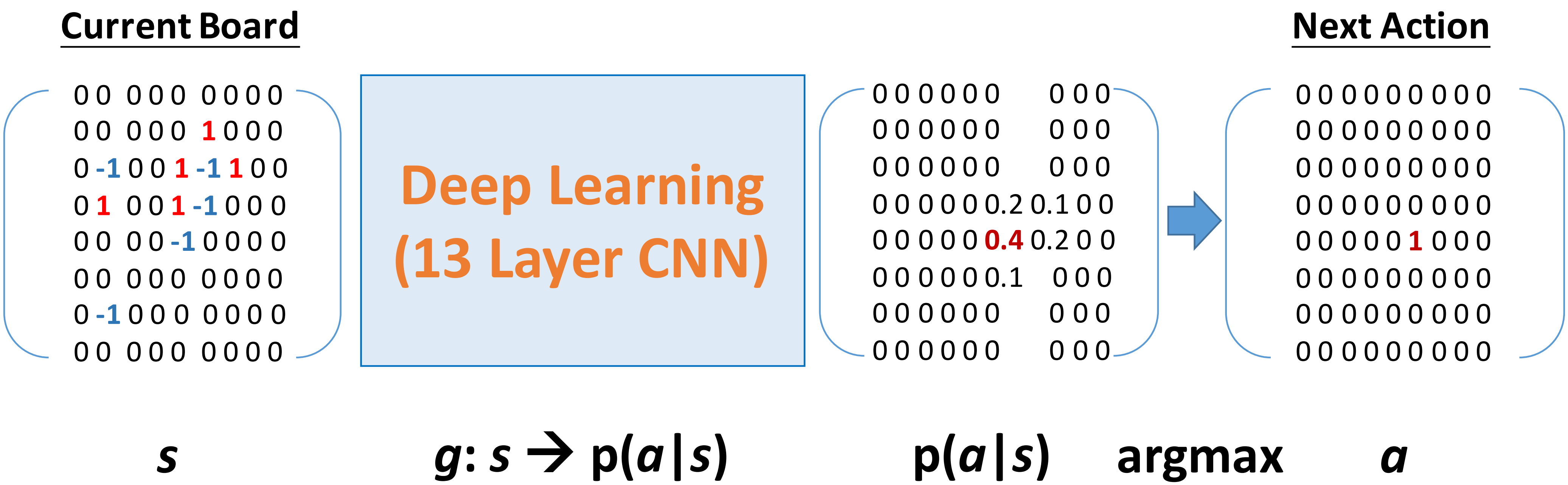
1. Reducing “action candidates” (Breadth Reduction)



IF there is a model that can tell you that these moves are not common / probable (e.g. by experts, etc.) …

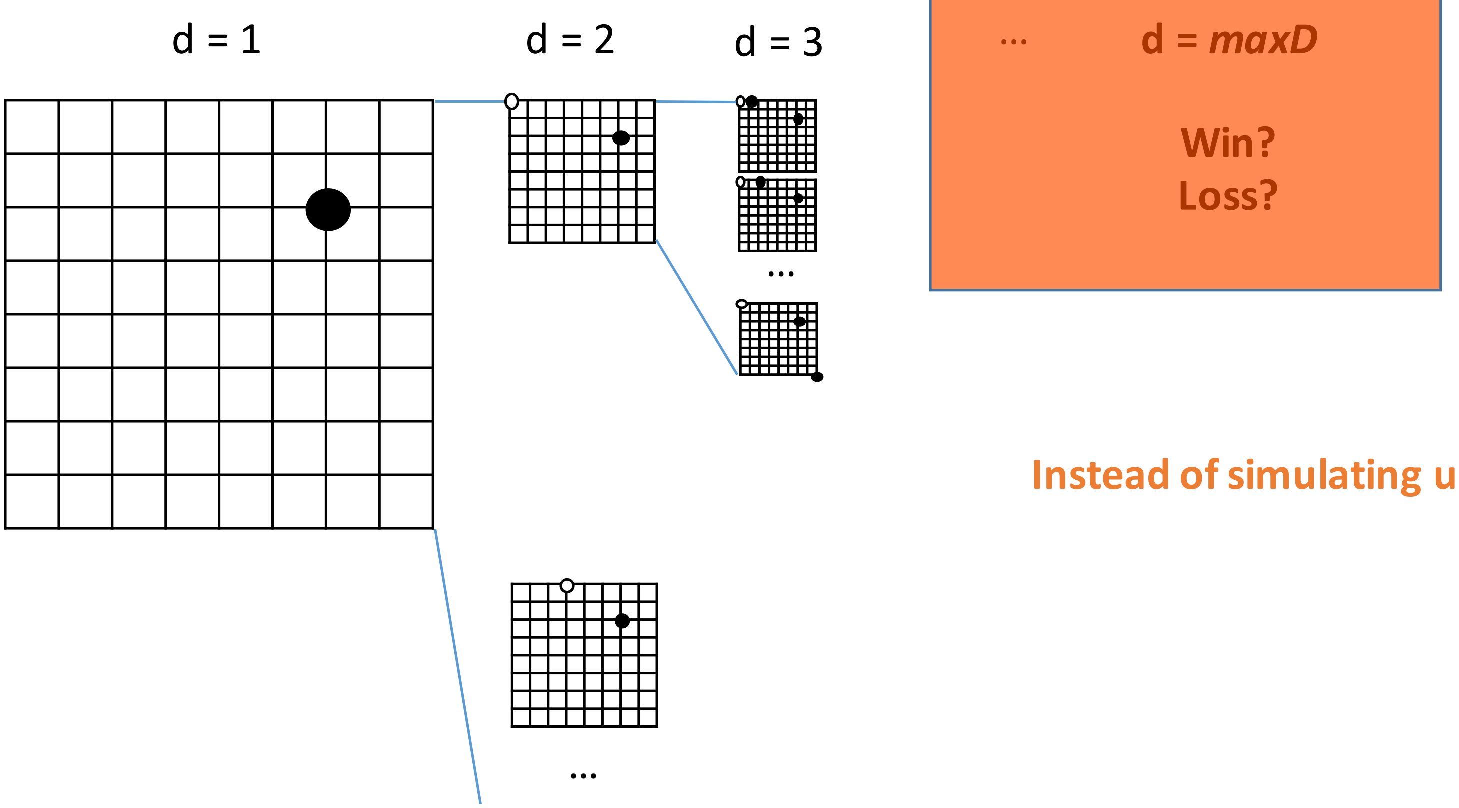
1. Reducing “action candidates”

(1) Imitating expert moves (supervised learning)



Reducing Depth of Thought

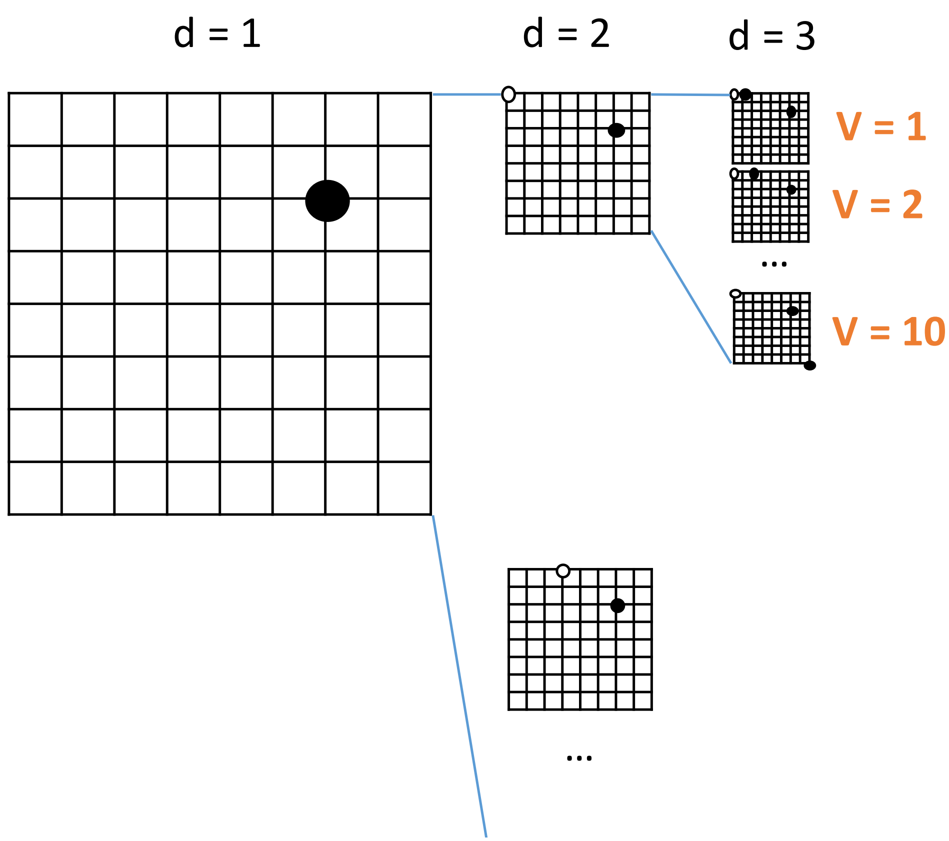
2. Position evaluation ahead of time (Depth Reduction)



Instead of simulating until the maximum depth ..

Reducing Depth of Thought

2. Position evaluation ahead of time (Depth Reduction)



IF there is a function that can measure:
V(*s*): "board evaluation of state *s*"

Game records

For other sources of game records, see the list in the links section.

On this page, you can download game records of top amateur games played on the K Go Server (KGS, formerly known as the Kiseido Go Server). I am grateful to Bill Shubert, who created KGS, for the permission to use these files, and for making them available to me in an easy way.

The games in the archives below are those where either one of the players (or both) is 7d or stronger, or both are 6d. All comments are stripped from these games, and all games with variations are omitted. They are suitable for use with Kombilo.

Need still more games? Have a look at the KGS games played by 4d+ players.

There are several versions of each archive, compressed in different ways: a .tar.gz, .tar.bz2, and .zip version; please choose the one which is most suitable for you. The content of the uncompressed archives is completely identical.

Source	Time period	Number of games	Archive format	File size	Link
KGS	2017_01	733	.zip	0.7 MB	Download
			.tar.gz	0.3 MB	Download
			.tar.bz2	0.2 MB	Download
	2016_12	1208	.zip	1.0 MB	Download
			.tar.gz	0.5 MB	Download
			.tar.bz2	0.3 MB	Download
	2016_11	980	.zip	0.9 MB	Download
			.tar.gz	0.4 MB	Download






https://u-go.net/gamerecords/

https://github.com/hughperkins/kgsgo-dataset-preprocessor

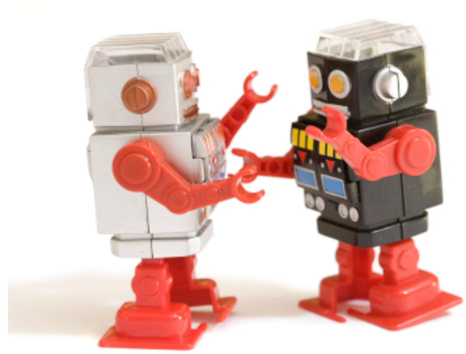
Winning Human Players

1. Reducing "action candidates"

(2) Improving through self-plays (reinforcement learning)

Improving by playing against itself

Expert Moves Imitator Model (w/ CNN)	**VS**	**Expert Moves Imitator Model (w/ CNN)**



1. Reducing “action candidates”

(2) Improving through self-plays (reinforcement learning)

Updated Model ver 1.3 **VS** **Updated Model ver 1.7**

Return: board positions, win/lose info

1. Reducing "action candidates"

(2) Improving through self-plays (reinforcement learning)

Updated Model ver 1.5

VS

Updated Model ver 2.0

Return: board positions, win/lose info

1. Reducing "action candidates"

(2) Improving through self-plays (reinforcement learning)

Updated Model ver 3204.1 **VS** **Updated Model ver 46235.2**

Return: board positions, win/lose info

1. Reducing “action candidates”

(2) Improving through self-plays (reinforcement learning)

Expert Moves Imitator Model	VS	Updated Model ver 1,000,000

The final model wins 80% of the time when playing against the first model


AlphaGo Zero
Starting from scratch

5000
4000
3000
2000
1000
0
−1000
−2000
Elo Rating
0
5
10
15
20
25
30
35
40
40 days
AlphaGo Zero surpasses all other versions of AlphaGo and, arguably, becomes the best Go player in the world. It does this entirely from self-play, with no human intervention and using no historical data.
AlphaGo Zero 40 blocks
AlphaGo Lee
AlphaGo Master

Lee Sedol 9-dan vs AlphaGo Energy Consumption

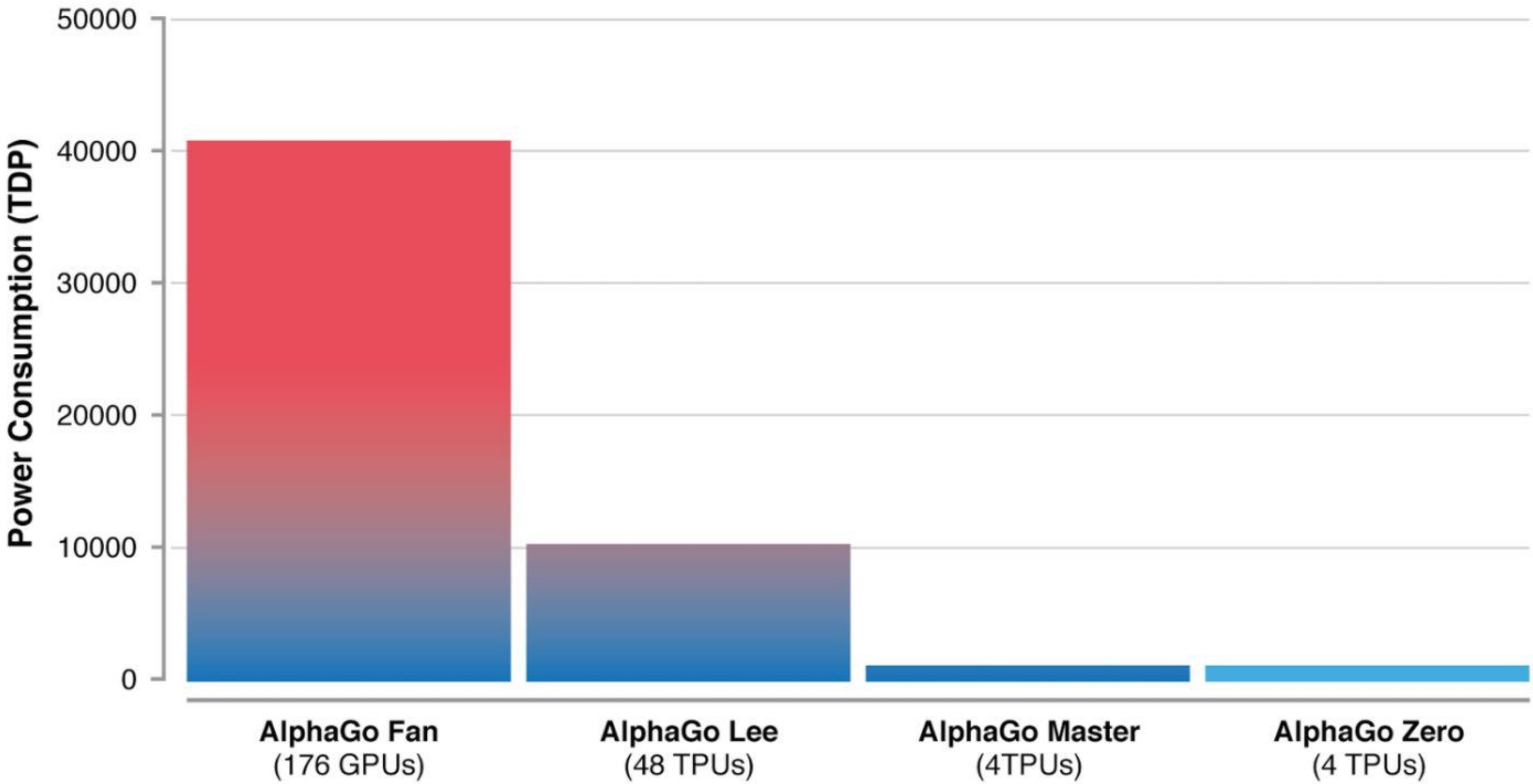
Lee Sedol	AlphaGo
- Recommended calories for a man per day : ~2,500 kCal - Assumption: Lee consumes the entire amount of per-day calories in this one game 2,500 kCal * 4,184 J/kCal **~= 10M [J]**	- Assumption: CPU: ~100 W, GPU: ~300 W - **1,202** CPUs, **176** GPUs 170,000 J/sec * 5 hr * 3,600 sec/hr **~= 3,000M [J]**

A very, very rough calculation ;)

Seoul AlphaGo

Deep Reinforcement Learning (as Nature AlphaGo)

- Improved value network
- Improved policy network
- Improved search
- Improved hardware (TPU vs GPU)


Power Consumption (TDP)
50000
40000
30000
20000
10000
0
AlphaGo Fan
(176 GPUs)
AlphaGo Lee
(48 TPUs)
AlphaGo Master
(4TPUs)
AlphaGo Zero
(4 TPUs)