MMAT5360 200323

Chance moves

Imperfect information

Bluffing game:

Both players put \$1 (ante) to the pot.

Player I is dealt a card: hidden to player II

Winning card with a probability $1/4$

Losing card with a probability $3/4$

Player I:

Check: the card is checked

Winning: 1 (payoff to I)

Losing: -1

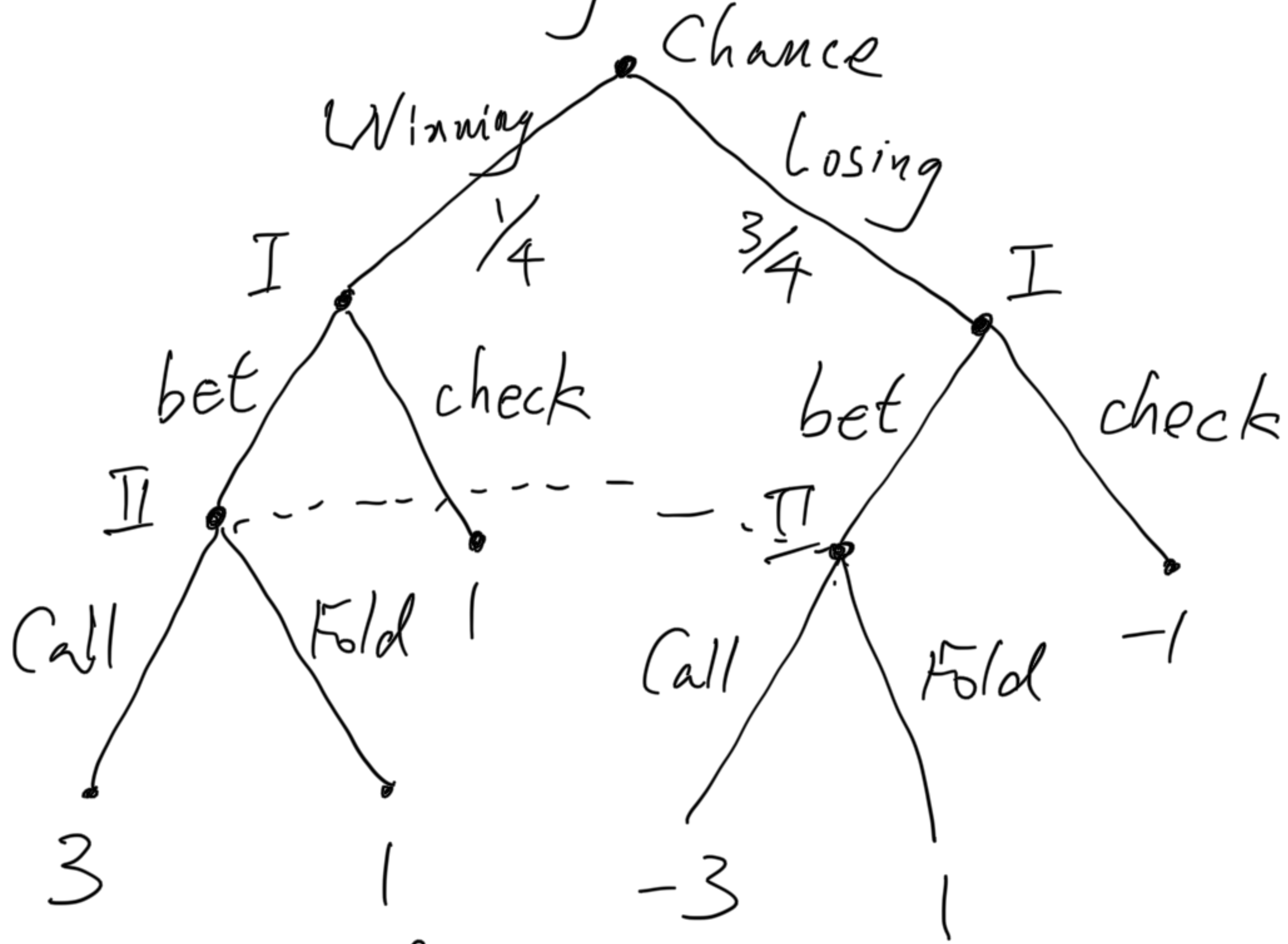
Bet: Puts \$2 to the pot

Player $\underline{II}$:

Fold : 1 (no matter what card I has)

Call : Winning : 3

Losing : −3



	Information set	Strategies
Player I	Winning, Losing	bb, bc, cb, cc
Player II	Bet	c, f

		Player II	
		c	f
Player I	bb	$\frac{1}{4}\times 3+\frac{3}{4}(-3)=-\frac{3}{2}$	
	bc		$\frac{1}{4}\times 1+\frac{3}{4}(-1)=-\frac{1}{2}$
	cb	$\frac{1}{4}\times 1+\frac{3}{4}(-3)=-2$	
	cc		

$$\begin{pmatrix} -\frac{3}{2} & 1 \\ 0 & -\frac{1}{2} \\ \cancel{-2} & \cancel{1} \\ \cancel{-\frac{1}{2}} & \cancel{-\frac{1}{2}} \end{pmatrix}$$

bb : bluffing strategy

bc : honest strategy

The matrix reduces to

$$\begin{array}{c} bb \\ bc \end{array} \begin{pmatrix} -\frac{3}{2} & 1 \\ 0 & -\frac{1}{2} \end{pmatrix}$$

Always bet for winning
Bet 1/6 for Losing

Maximin strategy : $\vec{p} = (1/6, 5/6, 0, 0)$

Minimax strategy : $\vec{q} = (1/2, 1/2)$

value : $v = -\frac{1}{4}$

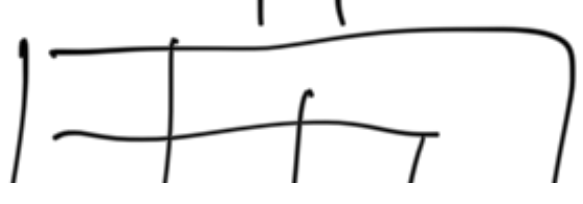
Monte Carlo tree search (MCTS)

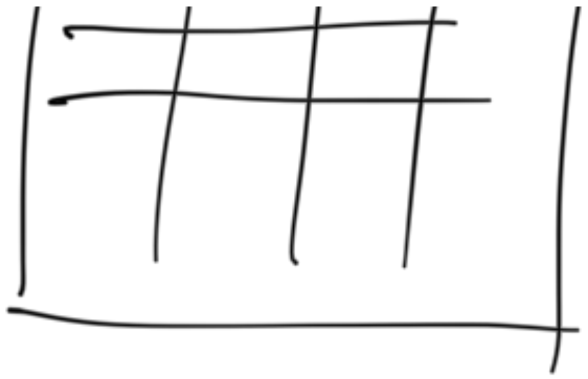
AlphaGo vs Lee Sedol 2016 4:1

Deep Blue 1996 Chess game

Minimax Algorithm

19





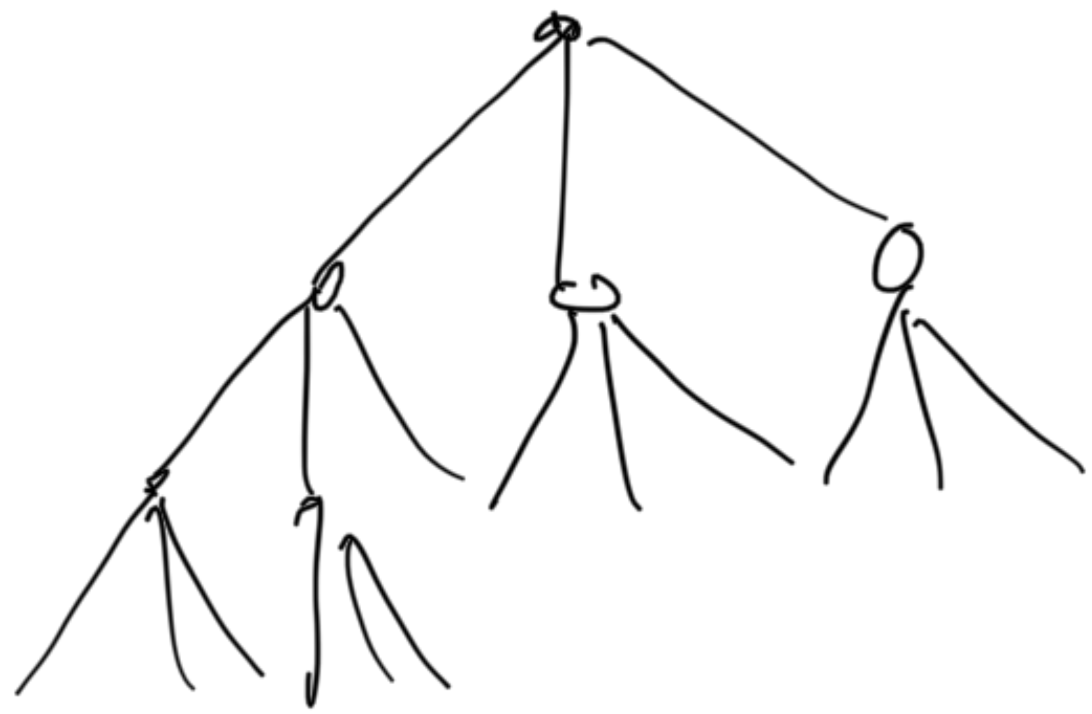
19

$19 \times 19 = 361$ squares

$361! \simeq 10^{768}$

4 steps of MCTS

1. Selection
2. Expansion
3. Simulation
4. Backpropagation



Choose the move which have the highest upper confidence bounded UCB

$$\frac{S_c}{n_c} + K\sqrt{\frac{2\ln n_p}{n_c}}$$

Dollar note game:

\$10 note \$20 note

Terran:

Left pocket	1	1
Right pocket	1	3

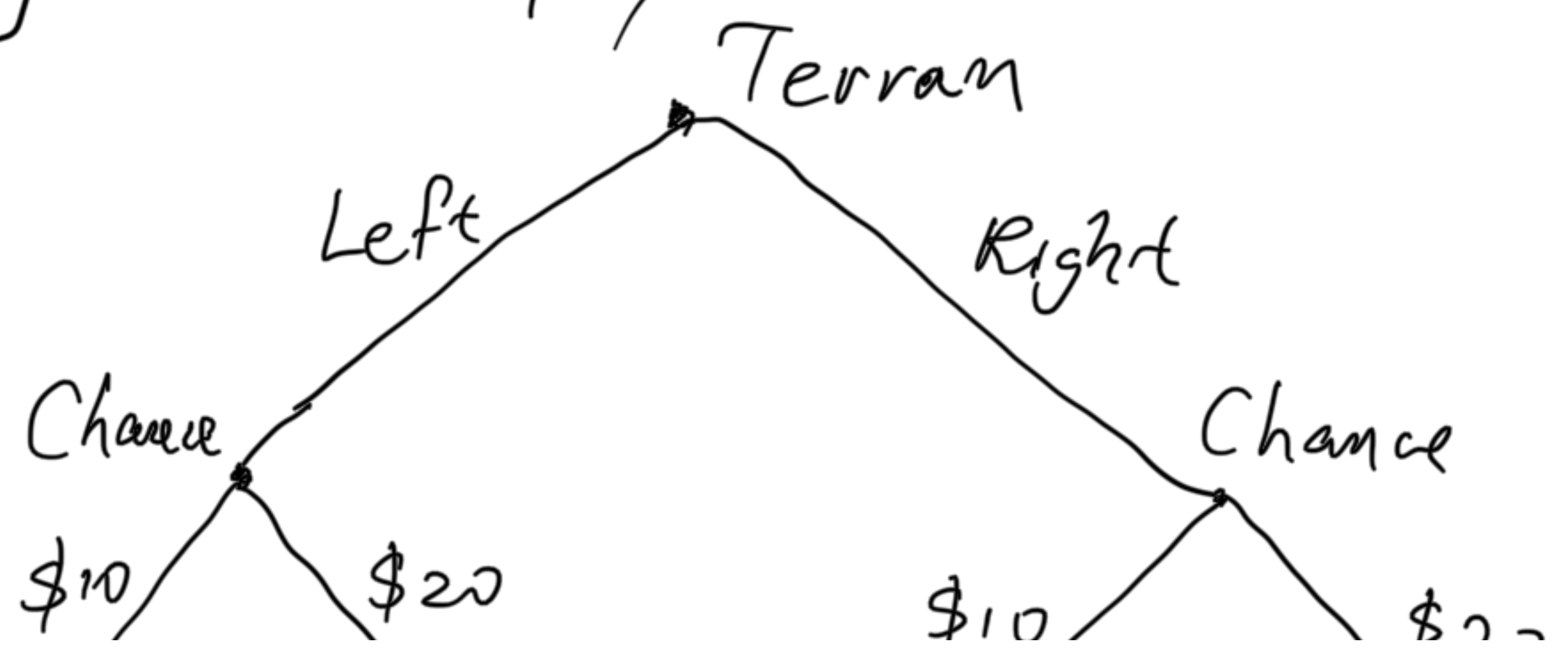
Terran chooses one pocket, pick a role at random and reveal it to Martian

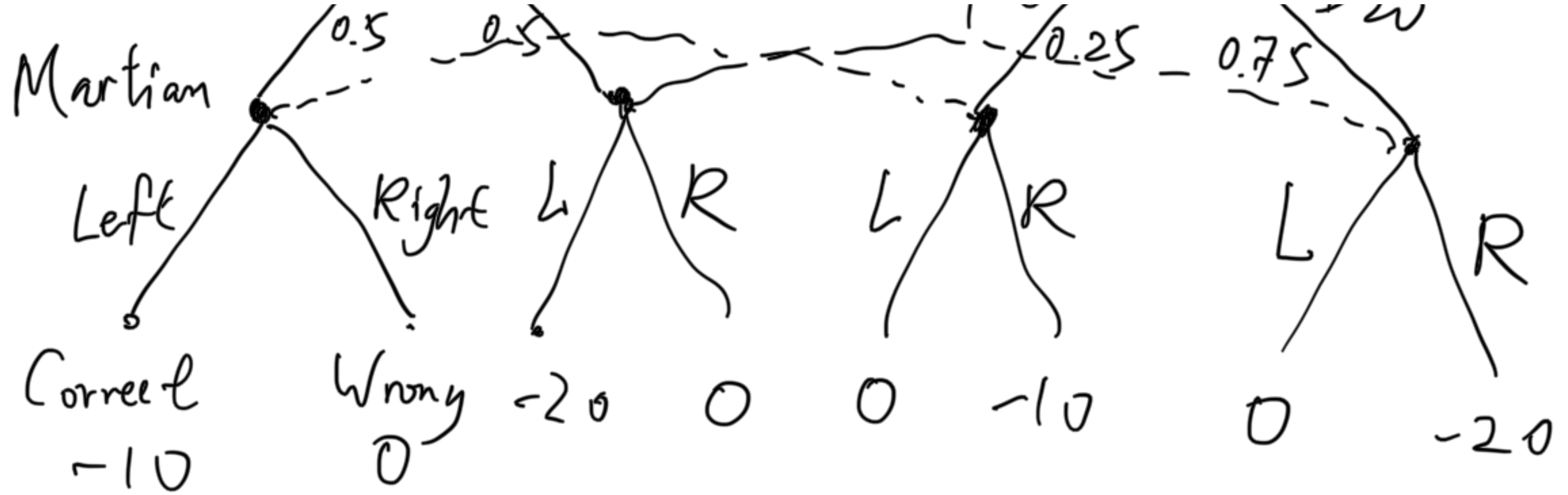
Martian does not know which pocket Terran has chosen but knows the value of the role.

Martian guesses: Left or Right

Correct: Martian gets the role Terran

Wrong: No payoff





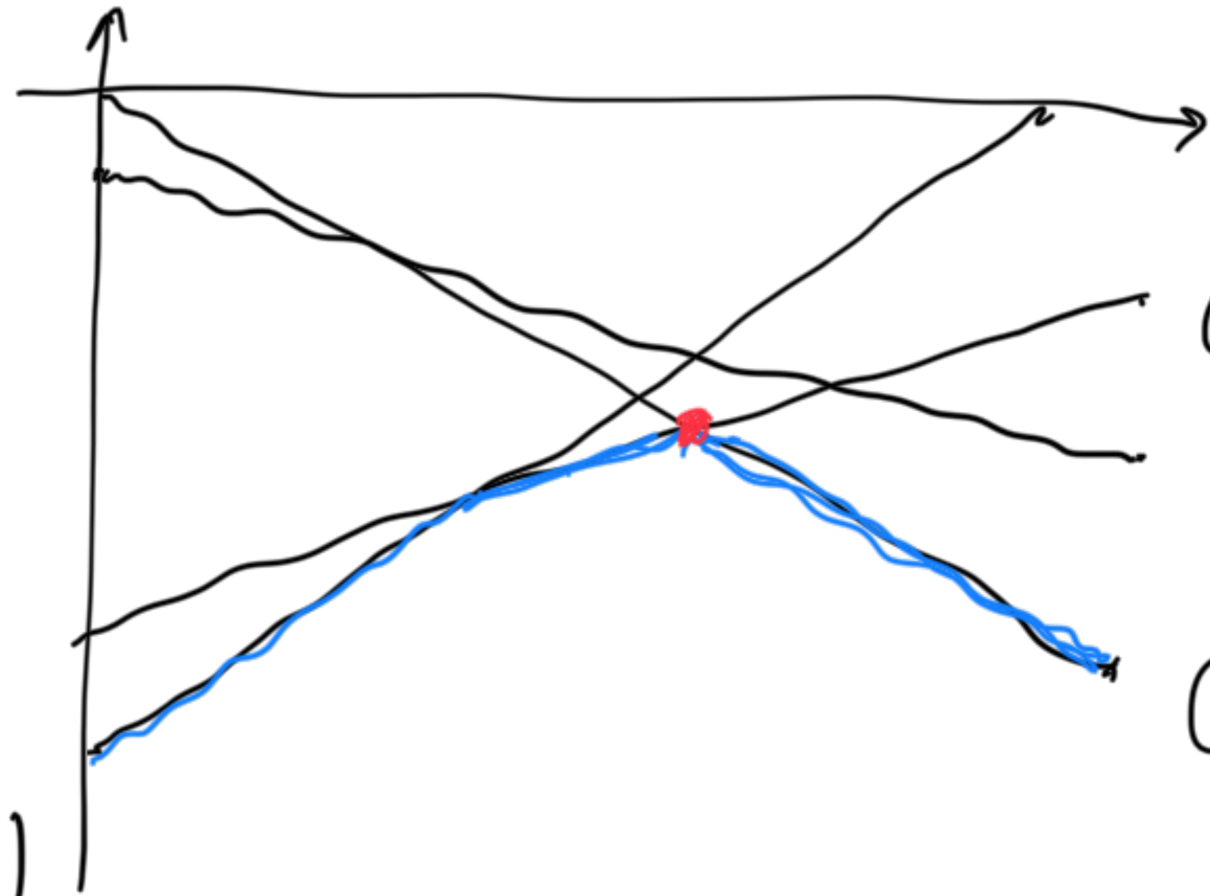
payoff to Terran

	Information set	Strategies
Terran	Initial	L, R
Martian	\$10, \$20	LL, LR, RL, RR

	LL	LR	RL	RR
L	$0.5(-10)+0.5(-20)=-15$		$0.5(0)+0.5(-20)=-10$	
R		$0.25(0)+0.75(-20)=-15$		

$$\begin{pmatrix} -15 & -5 & -10 & 0 \\ 0 & -15 & -2.5 & -17.5 \end{pmatrix}$$

$$\begin{pmatrix} -15 & -5 \\ 0 & -15 \end{pmatrix}$$



Maximin strategy: $\vec{p} = (0.6, 0.4)$

Minimax strategy: $\vec{q} = (0.4, 0.6, 0, 0)$

value : $v = -9$

Always Left if \$10

Left, Right with probabilities 0.4, 0.6 if \$20

Recursive games

$G_1 = \begin{pmatrix} 0 & 3 \\ 2 & -1 \end{pmatrix}$, $G_2 = \begin{pmatrix} 0 & 1 \\ 4 & 3 \end{pmatrix}$

2×2 zero sum games

$$G = \begin{pmatrix} G_1 & 4 \\ 5 & G_2 \end{pmatrix} \qquad \begin{array}{l} v_1 = v(G_1) = 1 \\ v_2 = v(G_2) = 3 \end{array}$$

G is equivalent to $\begin{pmatrix} 1 & 4 \\ 5 & 3 \end{pmatrix}$

$v(G) = 17/5$.

Inspection game:

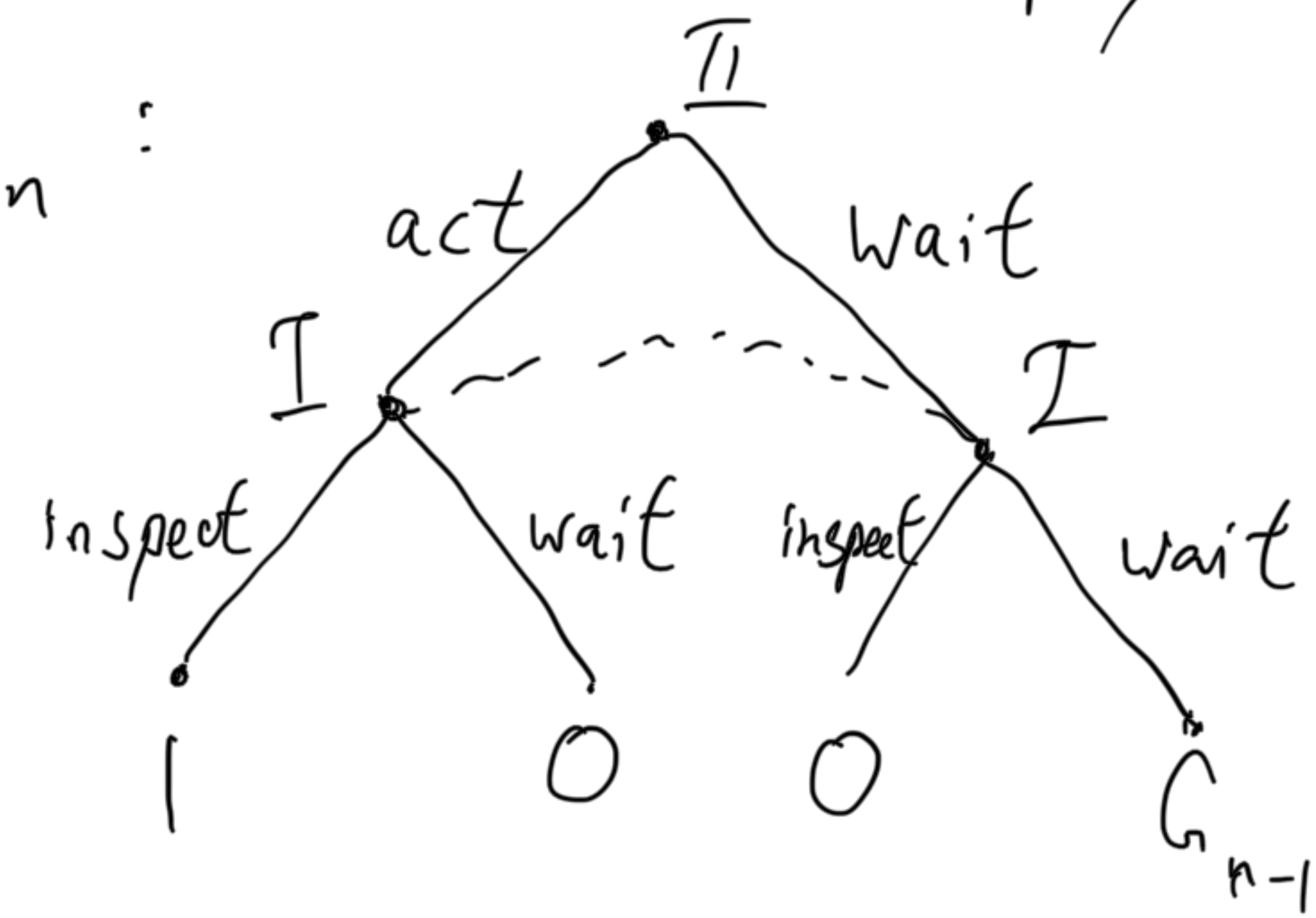
Player II must perform an action is the next n time periods.

Player I is allowed to inspect once in n time periods.

Player II is caught by Player I : payoff $= 1$

Player I missed : No payoff

G_n :



I : inspect , wait

II : act , wait

$$G_n = \begin{pmatrix} 1 & 0 \\ 0 & G_{n-1} \end{pmatrix}$$

$$G_1 = (1) \qquad v_1 = v(G_1) = 1$$

Let $v_n = v(G_n)$

Let $v_n = v(G_n)$ $\qquad \frac{\;}{1+\frac{1}{2}} = \frac{1}{2+1} = \frac{1}{3}$

$$G_n = \begin{pmatrix} 1 & 0 \\ 0 & V_{n-1} \end{pmatrix}$$

$$V_n = v(G_n) = \frac{1}{\frac{1}{1} + \frac{1}{V_{n-1}}} = \frac{V_{n-1}}{1+V_{n-1}}$$

n	1	2	3	4
V_n	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$

In fact $v_n = \frac{1}{n}$

$$G_n = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{n} \end{pmatrix}$$

maximin strategy: $\vec{p} = \left(\frac{1}{n}, \frac{n-1}{n}\right)$

minimax strategy: $\vec{q} = \left(\frac{1}{n}, \frac{n-1}{n}\right)$