CS540 Introduction to Artificial Intelligence
Lecture 22

Young Wu
Based on lecture slides by Jerry Zhu and Yingyu Liang

August 8, 2019

Coverage

- Unsupervised Learning
- Search

Unsupervised Learning

- Hierarchical Clustering
- K Means Clustering (no Gaussian mixture)

Single Complete Linkage Example

	A	B	C	D
A	0	~~x~~	2	3
B		0	4	5
C			0	~~6~~
D				0

single linkage

merge A, B.

	AB	C	D
AB	0	2	3
C		0	6
D			0

single linkage dist between AB, CD ?

$\Downarrow$ $\min\{2, 3, 4, 5\} = 2$

Principal Component Analysis

- Basic Linear (Matrix) Algebra
- Compute Projection and Variance (use formula on the formula sheet)
- Feature Reconstruction

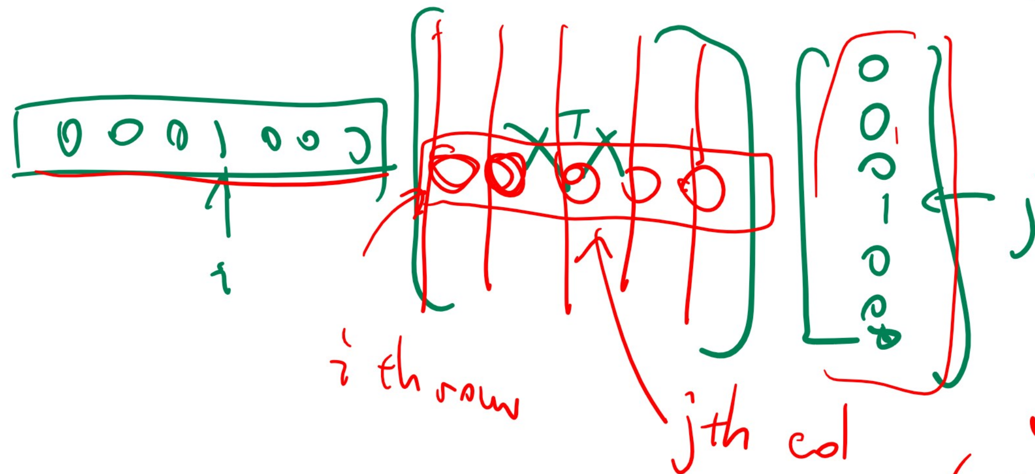
Matrix Multiplication Example

$X^T X$ entry (i, j) ?

$e_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ 0 \end{pmatrix} \leftarrow$ position

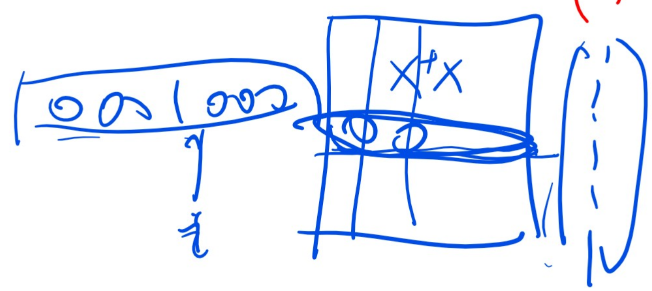
$e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

$e_i^T X^T X e_j$



$(X^T X)_{ij}$

$\boxed{\text{Sum of ith row}}$ $e_i^T X^T X e$

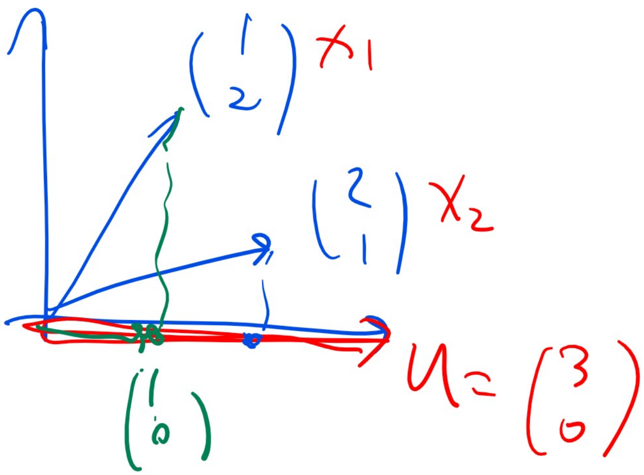


Sum of jth col $e^T X^T X e_j$

Sum of all entries: $e^T X^T X e$

Projected Variance Example

variance of magnitude of projection



proj X_1 onto u

$$\frac{x_1^T u}{u^T u}\, u$$

$$= \frac{\begin{pmatrix}1\\2\end{pmatrix}\cdot\begin{pmatrix}3\\0\end{pmatrix}}{\begin{pmatrix}3\\0\end{pmatrix}\cdot\begin{pmatrix}3\\0\end{pmatrix}}\begin{pmatrix}3\\0\end{pmatrix}$$

$$= \frac{3}{9}\cdot\begin{pmatrix}3\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}$$

length = 1

proj X_2 onto u = $\begin{pmatrix}2\\0\end{pmatrix}$

length = 2

(1, 2)

$$\hat{\sigma}^2 = \begin{cases} \text{MLE} & \frac{1}{n}\sum (x_i - \hat{\mu})^2 \\ \text{unbiased} & \frac{1}{n-1}\sum (x_i - \hat{\mu})^2 \end{cases}$$

$$\text{MLE} = \frac{1}{2}\left((1-1.5)^2 + (2-1.5)^2\right) = \frac{1}{2}\left(0.5^2 + 0.5^2\right)$$

Search

- Uniformed: no hueristic
- Informed: hueristic
- Local: optimization
- Adverserial: sequential move game
- Equilibrium: simultaneous move game

Uniformed Search

- BFS
- DFS
- IDS
- UCS

Informed Search

- Greedy
- A (or A Star)

Uniformed Search Counting Example

States $1, 2, 3 \ldots 1024 = 2^{10}$, successors of i

$2i$ and $2i+1$

initial 1, goal 1024

states expanded during BFS, DFS?

BFS = 1024

DFS = 10+1 = 11

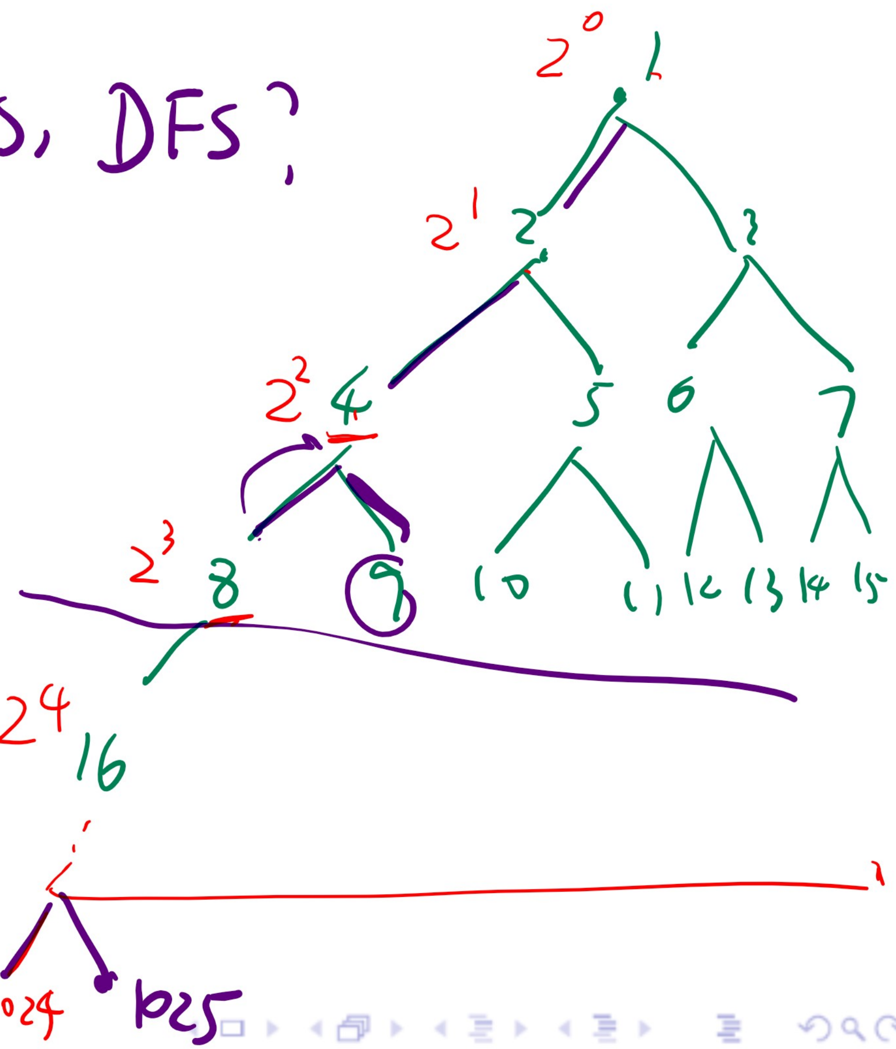
states $0 \ldots 1023 = 2^{10} - 1$

goal = 1023

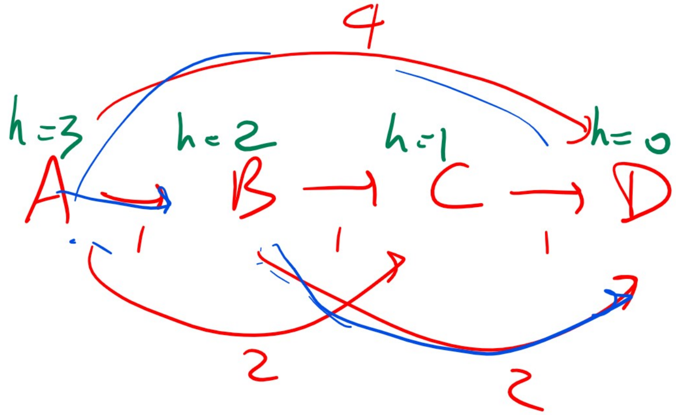
BFS = 1023

DFS = 1023

1025 : # BFS = 1025, # DFS = 12



Informed Search Complete Graph Example



① UCS

replaced D_A 4

PQ: ~~A~~ ~~B~~ from A ~~C~~ A ~~D~~ B

g = 0 1 2 3

expansion path: A, B, C, D

② Greedy PQ ~~A~~ ~~D~~ from A C B

h = 0 1 2

expansion path: A D

D from B

③ A* PQ ~~A~~ ~~B~~ from A ~~C~~ D^{3+0}

g+h 0+3 1+2 2+1 ~~4+0~~

expansion path: A, B, C, D

Local Search

- Hill Climbing
- Simulated Annealing
- Genetic Algorithm

SAT Local Search Example

SAT $\underbrace{(\underline{x} \vee y)}_{\text{Clause } 1} \wedge \underbrace{(\underline{x} \vee z)}_{2} \wedge \underbrace{(y \vee z)}_{3}$

⓪ initialize $x = F,\ y = F,\ z = F \longrightarrow S_0 \qquad f(S_0) = 0$

① random successor (neighbor) $x = T,\ y = F,\ z = F \longrightarrow S_1,\ f(S_1) = 2$

SA : $\begin{cases} \text{better} \rightarrow \text{prob } 1 \\ \text{worse} \rightarrow \text{prob } e^{-\frac{|f(s') - f(s)|}{Temp}} \end{cases} \Longrightarrow \text{prob} = 1$

② random successor. $x = F,\ y = F,\ z = F \longrightarrow S_2,\ f(S_2) = 0$

$Temp_0 = 1$

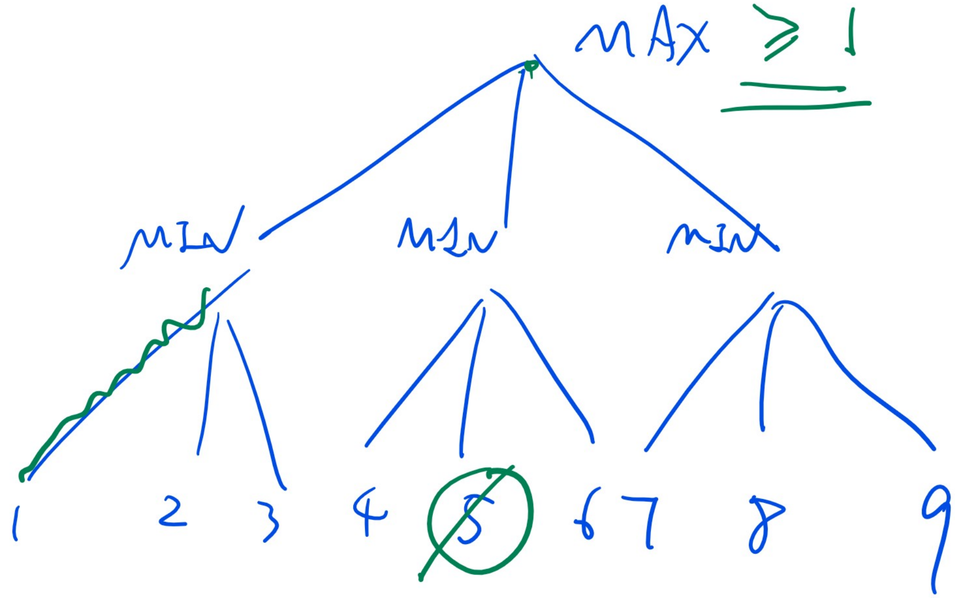
$Temp_t = Temp_{t-1} \cdot 0.9$

$$e^{-\frac{|2-0|}{Temp}} = e^{-\frac{2}{0.9^2}}$$

~~Adverseial~~ Search

- Minimax
- Alpha Beta

Alpha Beta Max Pruning



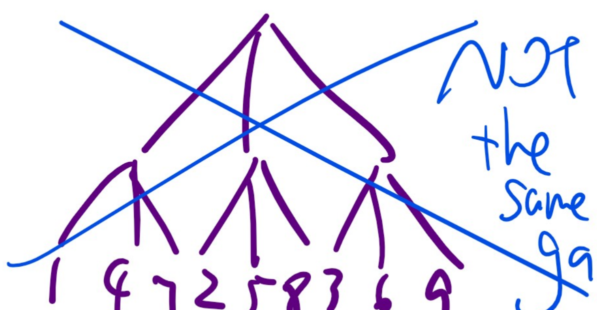
states you can prune during α-β search

DFS left most branch first.

↳ if this is 2 ⟹ 2 can be the value of game

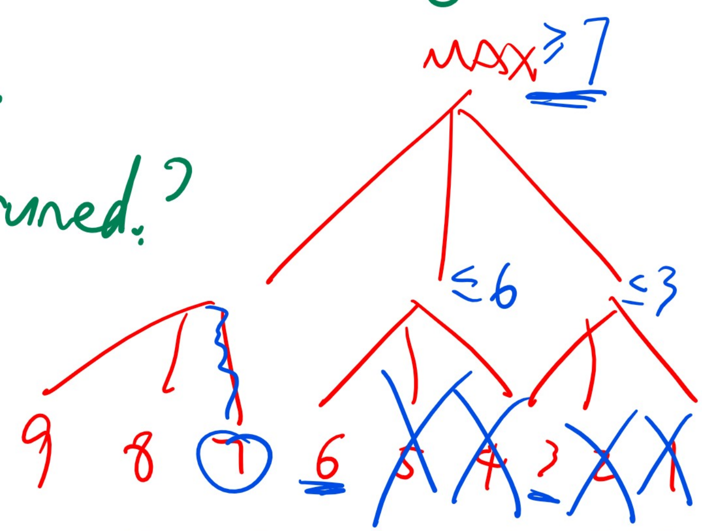
⟹ cannot be pruned.

reorder tree → max # of branches pruned?



NOT the same game

max # pruned = 4

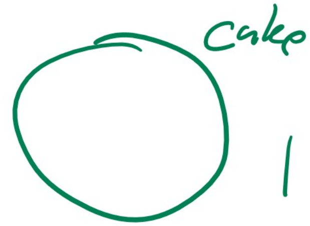


Equilibrium Search

- IESDS
- Best Responses (Nash)
- Fixed Point (Nash)

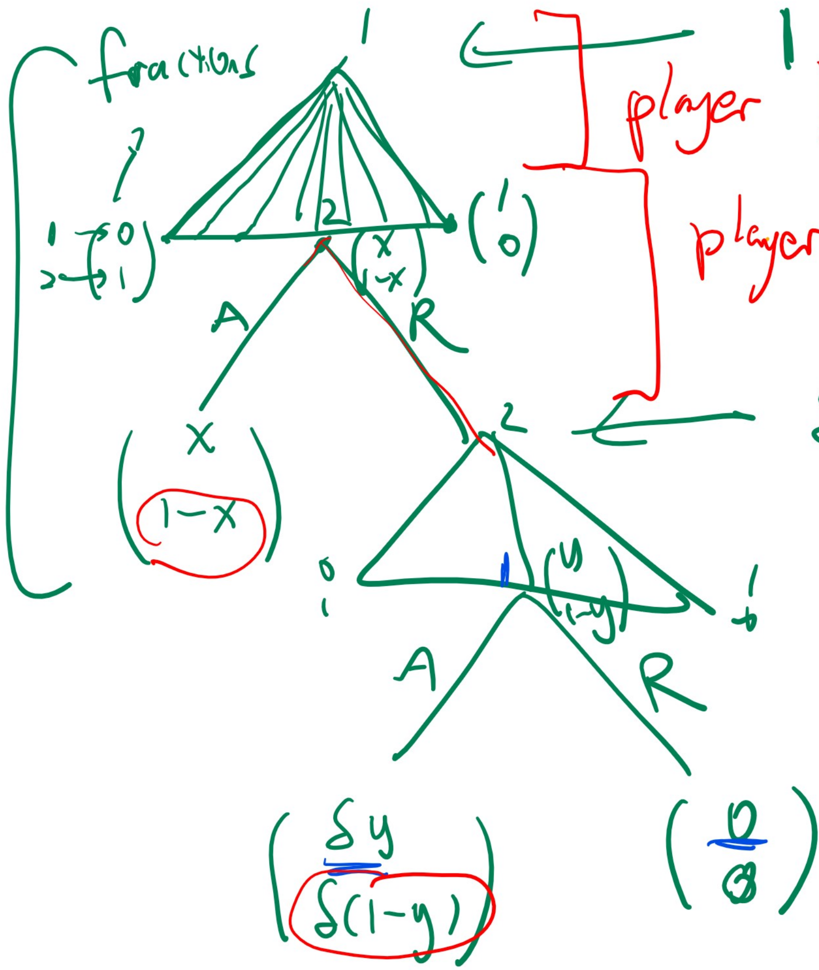
Bargaining Example

2 kids divide a cake



each disagree $\Rightarrow$ dad eats δ portion of cake

2 rounds



1 size

player 1 propose $x = \boxed{1-\delta}$

player 2 accept iff $1 - x \geq \delta(1-y) = \delta$

δ size

player2 propose $y = 0$

player 1 accept iff $\delta y \geq 0$

$N=2$: player 1 propose $\begin{pmatrix}1-\delta\\ \delta\end{pmatrix}$

player 1 propose $\boxed{1-\delta+\delta^2 z}$

player 2 accept

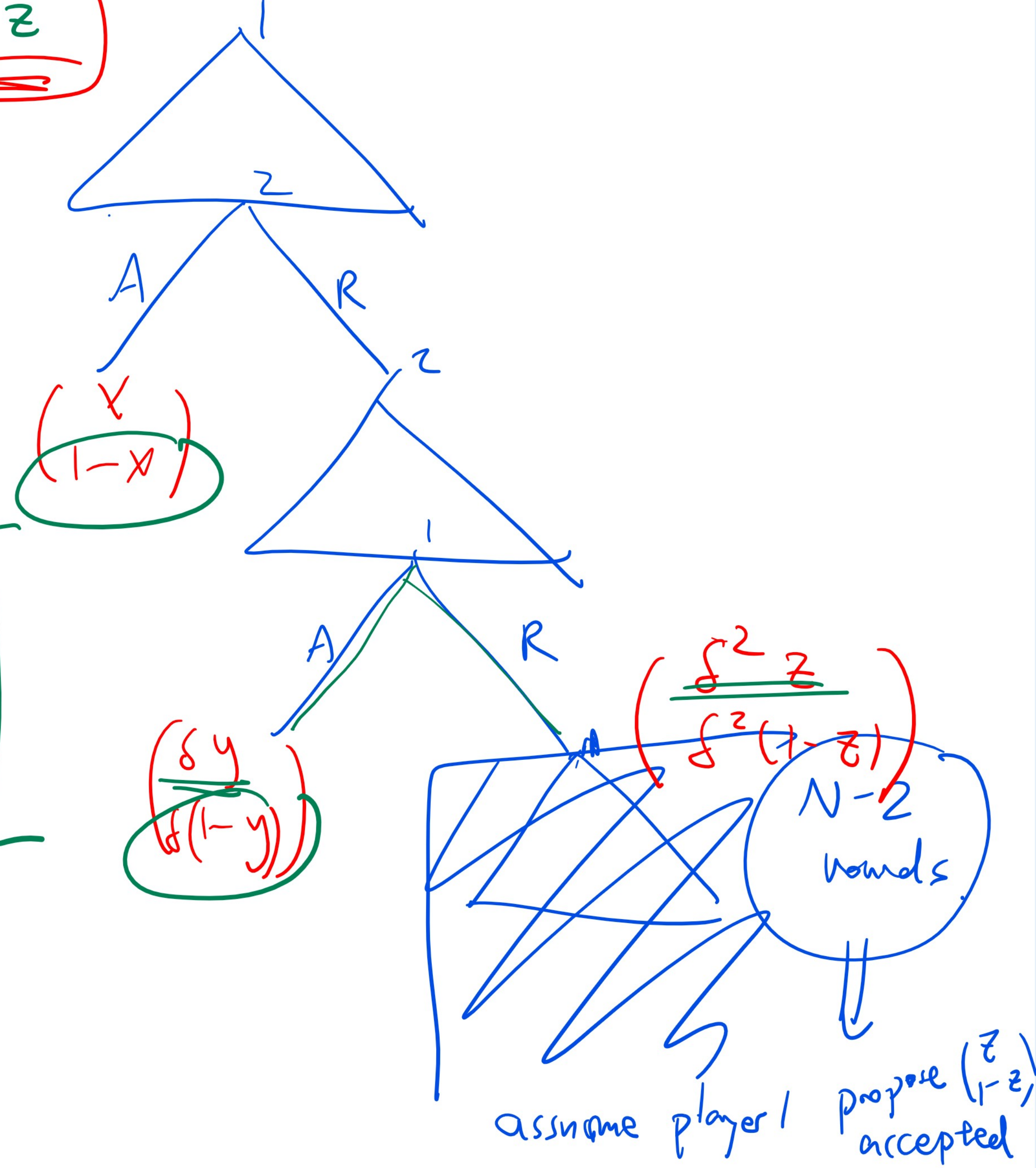
$1-x \geqslant \delta(1-\delta z)$

player 2 propose

$y=\delta z$

player 1 accept

$\delta y \geqslant \delta^2 z$



$N=2 \Rightarrow$ P1 propose $\begin{pmatrix} 1-\delta \\ \delta \end{pmatrix} \rightarrow z$

$N=4 \Rightarrow$ P1 propose $\begin{pmatrix} 1-\delta+\delta^2(1-\delta) \\ 1-\cancel{\delta} \end{pmatrix}$

$\begin{pmatrix} 1-\delta+\delta^2-\delta^3 \\ \\ \end{pmatrix}$

$N=6 \Rightarrow$ P1 propose $\left(1-\delta+\delta^2(z) \right.$

$1-\delta+\delta^2-\delta^3+\delta^4-\delta^5$

$\vdots$

$N=\infty$ $\quad 1-\delta+\delta^2-\delta^3+\delta^4-\delta^5+\delta^6-\delta^7+\ldots$

$$(1-\delta)\underbrace{\left(1+\delta^2+\delta^4+\delta^6+\ldots\right)}$$

$$(1-\delta)\frac{1}{1-\delta^2} = \frac{1}{1+\delta}$$

$$1+x+x^2+\ldots\ldots = \frac{1}{1-x}$$

∞ rounds game

kid 1 propose $\left(\frac{1}{1+\delta}\ \ \frac{\delta}{1+\delta}\right)$

kid 2 accept.