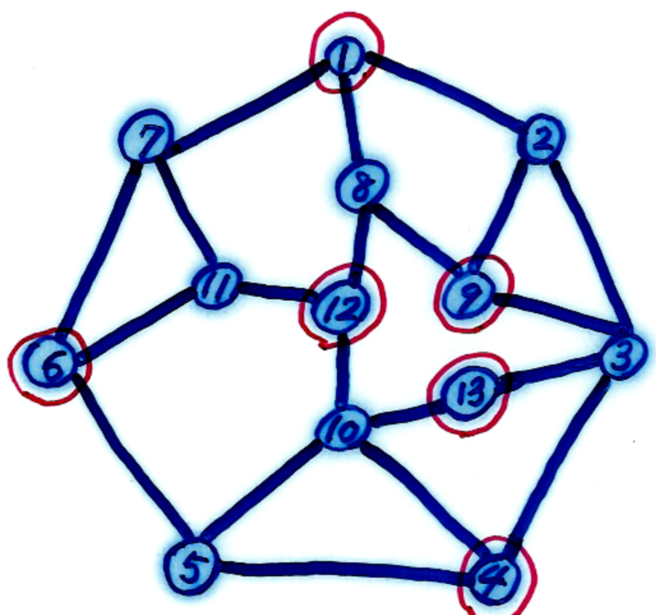
Derandomization

Lecturer: Dr. Hong-Gwa Yeh
Department of Mathematics
National Central University
Taiwan
hgyeh@math.ncu.edu.tw

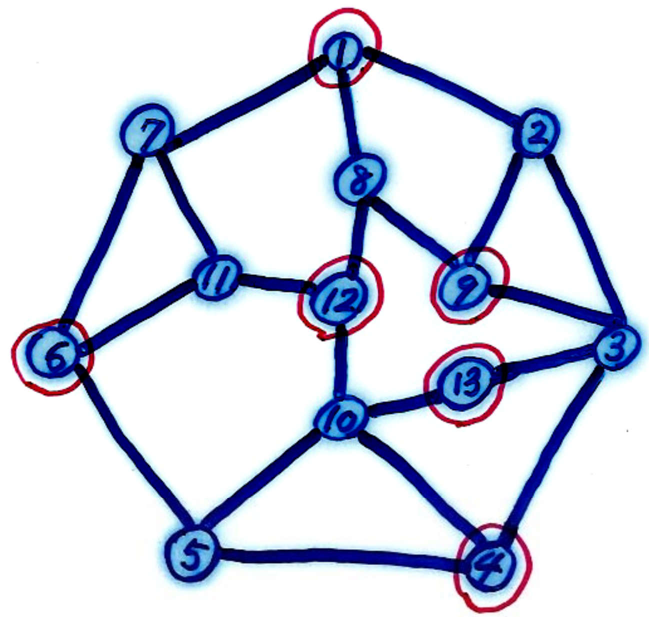
Finding a large independent set via coin tossing

Finding a large independent set of a graph via coin tossing

- Graph: $G=(V,E)$ $V=\{1,2,\cdots,n\}$ $|E|=m$
- $I \subseteq V$ is an *independent set* of G if no two vertices of I are adjacent in G.



- $I \subseteq V$ is an *independent set* of G if no two vertices of I are adjacent in G.



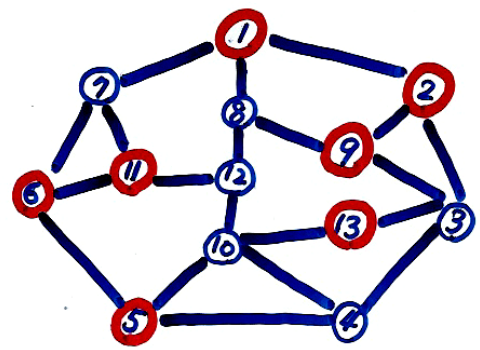
- $\alpha(G) = \max\{|I| : I \text{ is an independent set of } G\}$

Maximum independent set problem $\in \mathcal{NPC}$

- Finding an independent set that is not necessary optimal but has a guaranteed size.

- Generate a random subset $I' \subseteq V$ s.t. $\Pr\{\text{vertex } i \in I'\} = p$, the events $i \in I'$ being mutually independent.

G

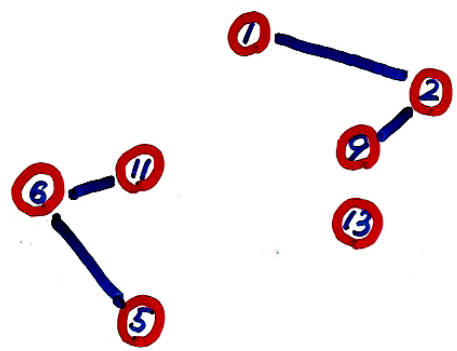


$V = \{1, 2, \cdots, 13\}$

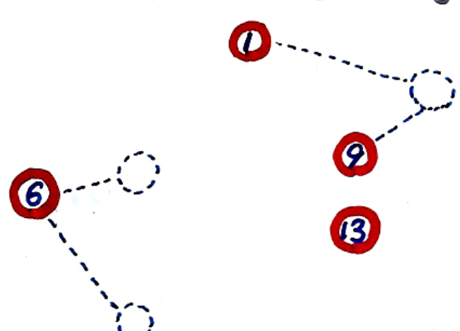
$|E| = 20$

$I' = \{1, 2, 5, 6, 9, 11, 13\}$

$G[I']$



Select one endpoint from each edge in $G[I']$ and remove it from I' to get I.



Let rv $X_i = \begin{cases} 1 & \text{if } i \in I' \\ 0 & \text{o.w.} \end{cases}$

Let $Z = \sum_{i \in V(G)} X_i - \sum_{\overline{ij} \in E(G)} X_i X_j$, $V(G) = \{1, 2, \cdots, n\}$ $|E(G)| = m$

Then $E[Z] = np - mp^2$

Choose $p = \frac{n}{2m} \in (0,1) \longrightarrow E[Z] = \frac{n^2}{4m}$

Theorem A: For any graph G with n vertices and m edges, we have

$$\alpha(G) \geq \frac{n^2}{4m}$$

A Tiny Course on Conditional Probability

$$P\{Y=y \mid A\} \stackrel{\text{def}}{=} \frac{P(\{Y=y\} \cap A)}{P(A)}$$

event

$$E(Y \mid A) \stackrel{\text{def}}{=} \sum_{y} y\, P(Y=y \mid A)$$

discrete case

$$E(Y \mid X) \stackrel{\text{def}}{=} f(X), \text{ where}$$

$$f(x) = E(Y \mid X=x)$$

rv

$$E(Y \mid X_1, X_2, \cdots, X_n) = f(X_1, X_2, \cdots, X_n)$$

where $f(a_1, a_2, \cdots, a_n) = \underbrace{E(Y \mid X_1=a_1, \cdots, X_n=a_n)}_{\bigstar}$

<u>discrete case</u>

$$\bigstar = \sum_y y \, P(Y=y \mid X_1=a_1, \cdots, X_n=a_n)$$

<u>conti. case</u>

$$\bigstar = \int y \, dF(y), \quad F(y) = P(Y \le y \mid X_1=a_1, \cdots, X_n=a_n)$$

Thm $\mathcal{E}\big(\mathcal{E}(Y \mid X_1, \cdots, X_{n+m}) \,\big|\, X_1, \cdots, X_n\big) = \mathcal{E}(Y \mid X_1, \cdots, X_n)$

Pf $n=2$, $m=1$, X_1, X_2, X_3, Y are discrete

$$\mathcal{E}(\bigstar \mid X_1 = a, X_2 = b)$$

$$= \sum_{x_1, x_2, c} \mathcal{E}(Y \mid X_1 = x_1, X_2 = x_2, X_3 = c)\, P(X_1 = x_1, X_2 = x_2, X_3 = c \mid X_1 = a, X_2 = b)$$

$$= \sum_c \sum_y y\, P(Y = y \mid X_1 = a, X_2 = b, X_3 = c)\, P(X_3 = c \mid X_1 = a, X_2 = b)$$

$$= \sum_c \sum_y y\, P(Y = y, X_3 = c \mid X_1 = a, X_2 = b)$$

$$= \sum_y y\, P(Y = y \mid X_1 = a, X_2 = b)$$ done!

Exercise

If X, Y are independent then $E(X|Y) = EX$

pf (discrete)

$$LHS = \sum_y E(X|Y=y)\, I_{\{Y=y\}}$$

$$= \sum_y \left(\sum_x x P(X=x \mid Y=y) \right) I_{\{Y=y\}}$$

$$= EX$$

End of the
Tiny Course

Observation:

$$E[Z \mid X_1 = x_1, X_2 = x_2, \dots, X_r = x_r]$$

$$= E[E[Z \mid X_1 = x_1, X_2 = x_2, \dots, X_r = x_r, X_{r+1}]]$$

$$= \underbrace{E[Z \mid X_1 = x_1, \dots, X_r = x_r, X_{r+1} = 1]}_{\text{★}}\, p + \underbrace{E[Z \mid X_1 = x_1, \dots, X_r = x_r, X_{r+1} = 0]}_{\text{☆}}\,(1-p)$$

$$\leq \max\{\text{★}, \text{☆}\}$$

▲ Choose $x_1, \cdots, x_n \in \{0, 1\}$ s.t.

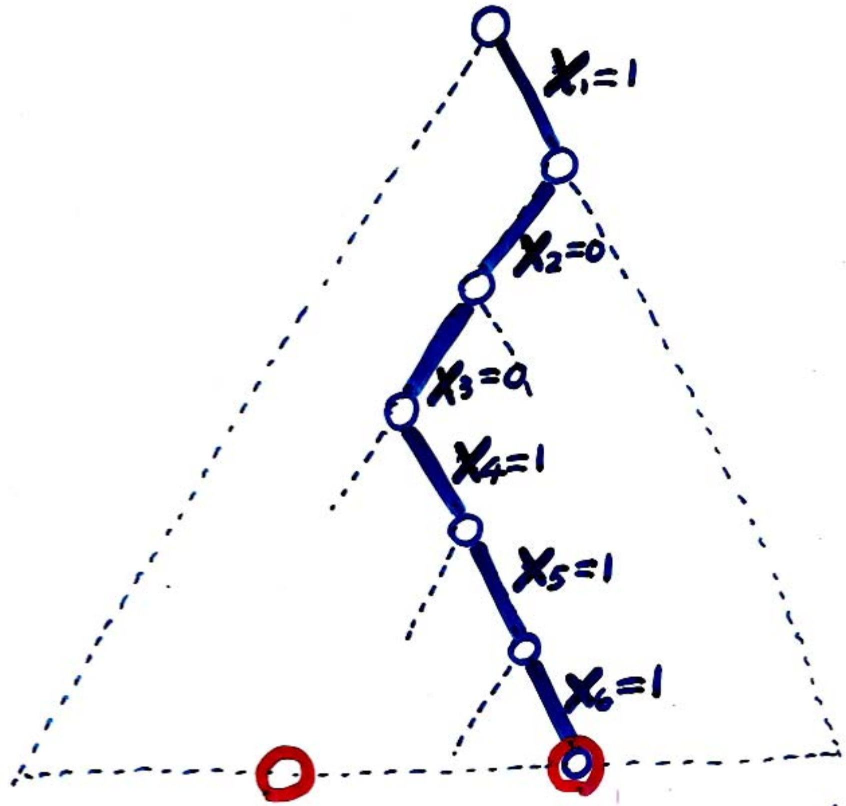
$$E[Z \mid X_1 = x_1] = \max\{E[Z \mid X_1 = 0],\ E[Z \mid X_1 = 1]\}$$

$$E[Z \mid X_1 = x_1,\ X_2 = x_2]$$

$$= \max\{E[Z \mid X_1 = x_1,\ X_2 = 0],\ E[Z \mid X_1 = x_1,\ X_2 = 1]\}$$

$$\vdots$$

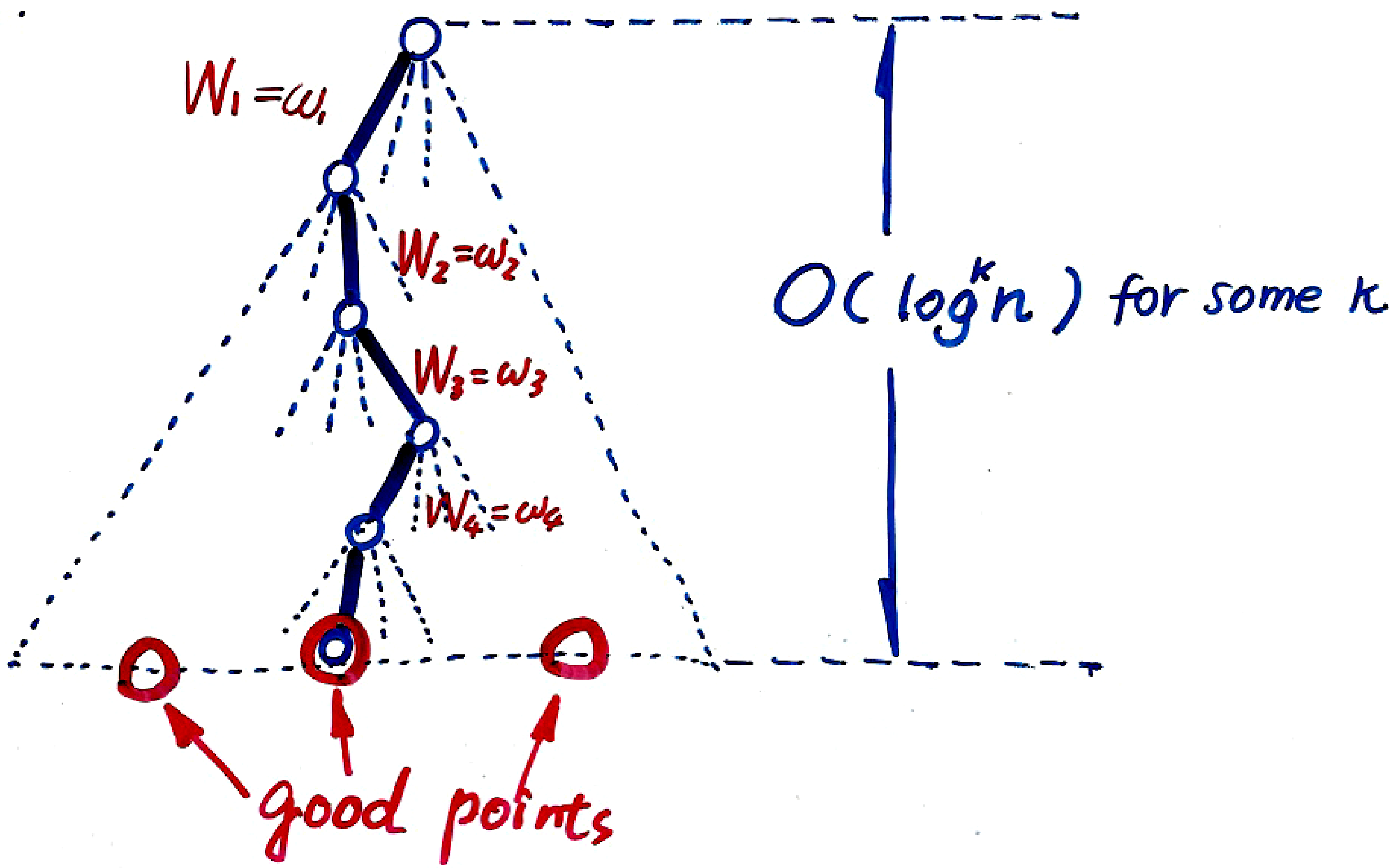
- $$\frac{n^2}{4m} \le E[Z]$$
$$\le E[Z \mid X_1 = x_1]$$
$$\le E[Z \mid X_1 = x_1, X_2 = x_2]$$
$$\vdots$$
$$\le E[Z \mid X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n]$$
- Let $I' = \{ i \in V(G) : x_i = 1 \}$.
Then we can prune I' to get an independent set of size $\geq n^2/4m$



- (a) It is hard to compute conditional expectations.
  (b) There are many instance where there is no known efficient way of computing the required conditional expectation.

  (c) The conditional probability method is inherently sequential and has running time $\geq n$.

(c) The conditional probability method is inherently sequential and has running time $\geq n$.



- **Recall:** rv $X_i = \begin{cases} 1 & \text{if } i \in \text{random subset } I' \subseteq V \\ 0 & \text{otherwise} \end{cases}$

$$Z = \sum_{i=1}^{n} X_i - \sum_{(i,j)\in E} X_i X_j$$

$$\Pr\{X_i = 1\} = p \quad \forall i$$

- **Observation:** If $X_1, \cdots, X_n$ are pairwise independent then also $E[Z] = \frac{n^2}{4m}$.

- **Observation:** If $X_1, \cdots, X_n$ are pairwise independent then also $E[Z] = \frac{n^2}{4m}$.

- Let $W_1, \cdots W_\ell$ be iid rvs s.t.
  $Pr\{W_i = k\} = \frac{1}{2m}$ for any $k \in \mathbb{Z}_{2m}$
  where $\ell = \lceil \log_2 n \rceil$

Redefine

$$X_i = \begin{cases} 1 & \text{if } \sum_{t=1}^{\ell} (i_t \cdot W_t) \pmod{2m} \in H \\ 0 & \text{otherwise} \end{cases}$$

where H is a fixed subset of $\mathbb{Z}_{2m}$ with size n and $\langle i_\ell, \cdots, i_1 \rangle$ is the binary expansion of i.

- Let $W_1, \cdots W_\ell$ be iid rvs s.t.
  $Pr\{W_1 = k\} = \frac{1}{2m}$ for any $k \in \mathbb{Z}_{2m}$
  where $\ell = \lceil \log_2 n \rceil$

  Redefine

  $$X_i = \begin{cases} 1 & \text{if } \sum_{t=1}^{\ell} (i_t \cdot W_t) \pmod{2m} \in H \\ 0 & \text{otherwise} \end{cases}$$

  where H is a fixed subset of $\mathbb{Z}_{2m}$ with size n and $\langle i_\ell, \cdots, i_1 \rangle$ is the binary expansion of i.

- **Observation:**

  $$Z = \sum_{i=1}^{n} X_i - \sum_{(i,j)\in E} X_i X_j = \mathbf{F}(W_1, \cdots, W_\ell)$$

- $E[Z \mid W_1 = \omega_1] = \max\{E[Z \mid W_1 = k] : k \in \mathbb{Z}_{2m}\}$

$$\begin{aligned}\frac{n^2}{4m} &\leq E[Z] \\ &\leq E[Z \mid W_1 = \omega_1] \\ &\leq E[Z \mid W_1 = \omega_1, W_2 = \omega_2] \\ &\vdots \\ &\leq E[Z \mid W_1 = \omega_1, \ldots, W_\ell = \omega_\ell]\end{aligned}$$

- **Main result**: There exists a deterministic Parallel algorithm for finding an independent set of size $\geq \frac{n^2}{4m}$ in a graph of n vertices and m edges.

Which can be implemented on an EREW-PRAM in $O(\log^2 n)$-time by using $O(m^2)$ processors.

- In search of the biggest determinant.
  - $A_n \in M_{n\times n}[\{-1, +1\}]$
  - How big can $|\det A_n|$ be ?

the famous (and unsolved) determinant problem of Hadamard.

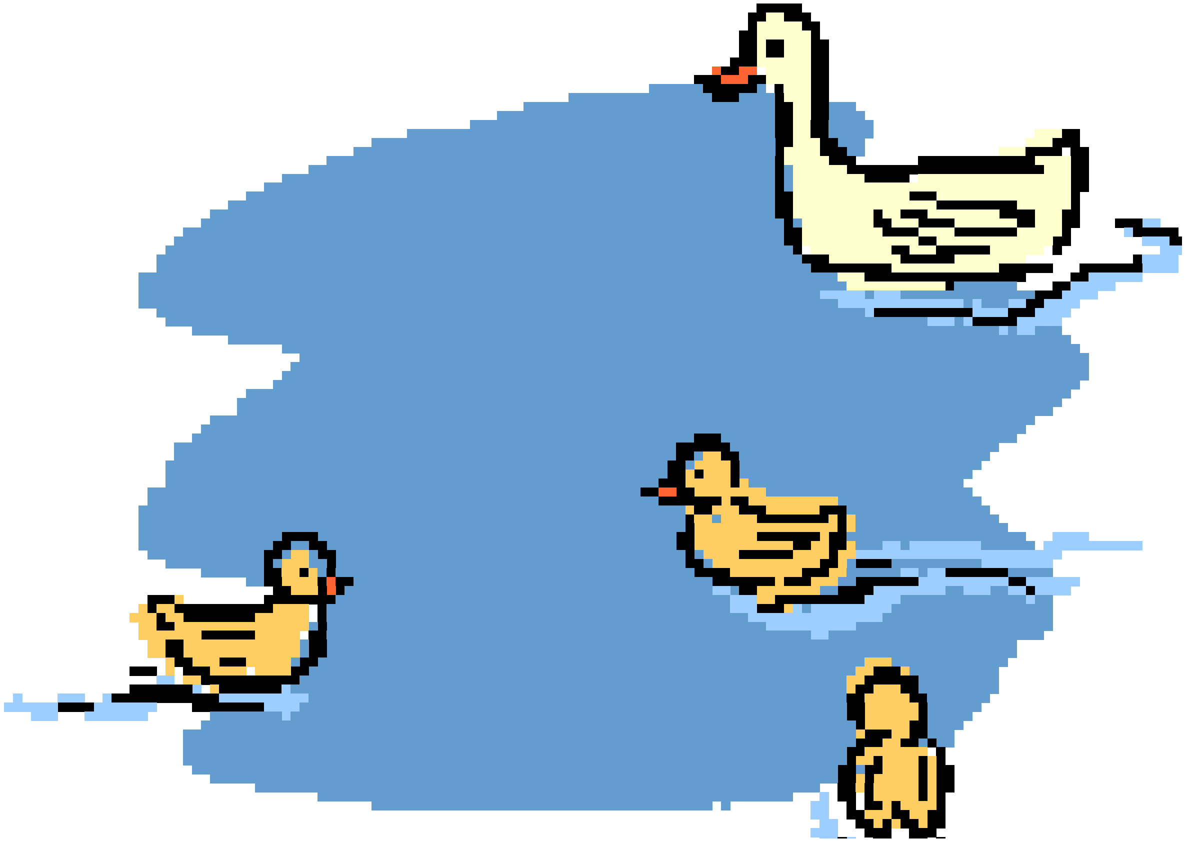
  - Fact: $|\det A_n| \leq n^{n/2}$

corollary of Hadamard's determint thm

- Consider $E[(\det A_n)^2]$
- (M. Kac) $E[(\det A_n)^2] = n!$ and hence there exists an $n \times n$ matrix of ± 1's whose determinant is $\geq (n!)^{1/2}$
- No one know how to construct one efficiently.

$\Pr(A) > 0$ says that

there is a juicy fish in the lake !



Can we find the fish efficiently ?

G. C. Rota said that

- "It is widely conjecture that an algorithm should exist that would transform an existence proof obtained by Erdos's probabilistic method into an ordanary constructive logical proof" (1996)