Distributed Algorithms (2)

- $\text{disj}_n(X,Y) = \begin{cases} 1 & \text{iff } X \text{ and } Y \text{ are disjoint} \\ 0 & \text{(i.e. miss an intersection)} \end{cases}$
- $\text{script}_\pi(X,Y) = (m, s_1, m_1, \cdots, s_m, m_m)$
- Matrix of f: rows $X \in \{0,1\}^n$, columns $Y \in \{0,1\}^n$, entry $f(X,Y)$
- Rectangle Property: if scripts are the same for (X,Y) & (X',Y') then they are the same for (X,Y') & (X',Y)
- Prop. 1.3: If F_0, F_1 are fooling set, the number of scripts for EVERY protocols computing F requires $\geq |F_0| + |F_1|$.

Fooling set: $F_0 = \{(x,y) \mid f(x,y) = 0\}$

$\forall (x,y),(x',y') \in F_0 \quad f(x,y') = b \text{ or } f(x',y) = b$

- For disj_n: 2^n-m scripts are needed. Chosen on fooling set $|F_0| = |F_1| = 2^n$

Goal: Make a link between the #scripts and communication complexity.
Const communication cost of π

Recall: $\text{cost}(\pi) = \max_{(x,y)} \sum_{i=0}^{m} |m_i|$

⚠ Com. Complexity cannot be defined as the minimum cost of any protocol computing f.

$\exists \pi_\beta$ that has a cost of 2 bits $\forall f \; \forall (x,y)$.

π_β: Alice & Bob exchange empty msg in turn ($m_i = \varepsilon$). If $x \in \{0\}^n \cup \{1\}^n$ [rounds, then Alice sends 0 when she's alone. Bob, has counted the #rounds, can recover x, and so transmits $f(x,y)$ to Alice.

$\text{script}_\beta(x,y) = (\underset{2x}{\underbrace{\varepsilon, \varepsilon, \varepsilon, \varepsilon}}, \underset{1}{\overset{A}{b}}, \underset{1}{\overset{B}{f(x,y)}})$

$\text{cost}(\pi_\beta) = 2$ bits :((

But, #scripts is 2^{n+2}.

We need an extra property in order to well define Com. Complexity.

$\Rightarrow$ Prefix-free property.

Alice $\xrightarrow{\quad 1100 \quad}$ Bob
Bob $\xrightarrow{\quad 1001 \quad}$ Alice
Alice $\xrightarrow{\quad 1100 \quad}$ Bob

[1]

How Alice can decide if Bob's answer is the last message of Bob?

DEFINITION: Π is prefix-free if $\forall$ round $i \in \mathbb{N}$, the set of all the words up to round i:

$\mathcal{M}_\pi(i) = \bigcup_{(x,y)} \{ m_1(x,y) \# m_2(x,y) \# \cdots \# m_i(x,y) \mid m_1(x,y), \cdots, m_i(x,y) \text{ are msgs for } (x,y) \text{ rounds} \}$

is a prefix-free set.

Ex:
- $\{0, 10, 11\}$ is a prefix-free set
  - $0 \to 0$
  - $1 \to 10$
  - $0 \to 11$
- $\{0,1\}$ is prefix-free
- $\{0, 01, 11\}$ is NOT prefix-free

$\mathcal{M}_{\pi_\beta}(0)$ = set of all possible messages for the first rounds for Alice

Alice $\xrightarrow{\quad m_1 \in \mathcal{M}_{\pi_\beta}(0) \quad}$ Bob

Property of a prefix-free set: It's that any message formed by words $\in \mathbb{Z}$ can be decoded with no errors.

$\mathbb{Z} = \{0, 11\}$
$M = \underbrace{00}_{0}\underbrace{0}_{0}\underbrace{11}_{11}\underbrace{0}_{0}\underbrace{11}_{11}\underbrace{00}_{00}$

Exercice Show that with a prefix-free protocole, Alice & Bob can detect when it is over.

Answer. $S_i = m_0 \# m_1 \# \cdots \# m_i$ (Bob knows S_{-1}, if $S_i = S_{i-1} \# m_i$ & $m_i \in \mathcal{M}_{\pi_\beta}(i)$ and $S_i \notin \mathcal{M}_{\pi_\beta}(i)$ and $m_i \neq \varepsilon$)

round 1: Alice $\xrightarrow{m_1}$ Bob ($m_1 \neq \varepsilon$)
Bob can compute $\mathcal{M}_{\pi_\beta}(1) \ni \mathcal{M}_{\pi_\beta}(i+1)$.
$S_1 \in \mathcal{M}_{\pi_\beta}(1)$: If $S_1 \in \mathcal{M}_{\pi_\beta}(i+1)$ then Π is over for (x,y). Otherwise
$S_1 = m_1 \in \mathcal{M}_{\pi_\beta}(i+1)$ and π continues.

DEFINITION The communication complexity of f is the minimum cost of any prefix-free protocole computing f.

NB. π_β is not prefix-free. (because ε is prefix of any word)

$\mathcal{M}_{\pi_\beta}(0) = \underset{x \neq 0^n \; x \neq 1^n}{\{x, y\}}$ is not prefix-free

Proposition 1.5

Every com. protocole β on $\{0,1\}^n$ with cost k and having at most R rounds can be transformed into a prefix-free protocole with cost $2(k+1)$ and with the same number of scripts.

Idea: Each msg is transformed in Π' by transformation δ (double each bit but) & transformation σ (remove the last one).

$\delta(x_1 x_2 x_3 \cdots) = x_1 x_1 x_2 x_2 x_3 x_3 \cdots$ (if msg $\neq \varepsilon$)

$\sigma(m_i) = \begin{cases} \varepsilon \to m_i & \text{if } i = \pi(x,y) = \text{last round} \\ 01 & \text{if } \sigma(\pi(x,y)) \neq m_i = \varepsilon \\ x_1 x_2 \cdots \delta(x_n) & \text{otherwise } (m_i \neq \varepsilon) \end{cases}$

We can show $\mathcal{M}_{\pi'}(i)$ is prefix-free set. ☐

Notation The minimum number of scripts of any protocole computing f is $\mathcal{C}(f) = \min_{\Pi} \{ \text{script}_\pi(x,y) : (x,y) \in \{0,1\}^{2n} \}$

Observation. The minimum is taken over all protocoles Π (prefix-free or not)

Proposition

The com. complexity for f is $CC(f) = \log_2 \mathcal{C}(f)$

Idea: - we can restrict the minimum protocole is prefix-free.

- Assume the scripts are
  (0000, $\varepsilon\varepsilon$), (0001, $\varepsilon\varepsilon$), (0010, $\varepsilon\varepsilon$), (0011, $\varepsilon\varepsilon$), (0100, $\varepsilon\varepsilon$), (0101, $\varepsilon\varepsilon$), (0110, $\varepsilon\varepsilon$), (0111, $\varepsilon\varepsilon$), (1000, $\varepsilon\varepsilon$), (1001, $\varepsilon\varepsilon$), (1010, $\varepsilon\varepsilon$), (1011, $\varepsilon\varepsilon$), (1100, $\varepsilon\varepsilon$), (1101, $\varepsilon\varepsilon$), (1110, $\varepsilon\varepsilon$), (1111, $\varepsilon\varepsilon$)
  $\mapsto$ (0,100), (0,101), (0,110), (0,111), (1,00), (1,01), (1,10), (1,11)
  (Bob knows Alice's input)
- Kraft's inequality that applies to prefix-free set:

$\mathbb{Z} = \{0, 10, 11, 100\} \to \mathbb{Z}' = \{0, 10, 11\}$

$\sum_{z \in \mathbb{Z}} 2^{-\text{length}(z)} \leq 1$ ☐

Corollary

$CC(\text{disj}_n) = n+1$

Proof. $CC(f) \leq n+2$ because of the trivial protocole C which is prefix-free
we have seen that $\delta(\text{disj}_n) \geq 2^n$
$CC(\text{disj}_n) = \log_2\left(\sum_{\text{script} \in \mathcal{M}} 2^{-n}\right) > n$
$CC(\text{disj}_n) \geq n+1$ ☐

If remaining to show how to reduce the problem of detecting a C_4/C_6 in a graph from SET-DISJOINTNESS Problem?

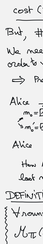
Take a graph H with m vertices m edges with no C_3 and C_4 cycles

We assume that H has the max. number of edges

We select 2 sets $A, B \subseteq E(H)$

$H_0 = H[A]$ $H_1 = H[B]$

$G = H \sqcup K_2(\ell)$



$G_{A,B}$ (ET4) Assume that $A, B \subseteq H$ are such that $CC(\text{disj}_{m'}) \geq m'+1$ ($m' = \text{number of edges in } H$)

$G_{A,B}$ has a C_4 or C_6 iff $A \cap B \neq \emptyset$

Why? If $\exists$ $e_H \in A \cap B$ $H_0[A]$, $H_1[B]$ have C_4.

Assume A is a distributed algorithm that detects C_4/C_6 in any G with n vertices. t = #rounds (A) = time(A, G).

By simulating A on $G_{A,B}$ on their part (H_0, H_1), Alice & Bob have a protocole for SET-DISJOINTNESS. Once Alice or Bob decide a C_4 (C_6 not) a last msg is sent to the other part.

Cost of this protocole obtained by simulating A on $G_{A,B}$? [b-CONGEST]

$\underbrace{CC(\text{disj}_{m'})}_{m'+1} \leq \text{COST} \leq \underbrace{t}_{\text{#rounds}} \times \underbrace{b \times 2m}_{\text{#msg in 2 rounds}}$ (b msg bound between Alice/Bob)

$m+1 \leq t \cdot b \cdot 2m$

$\Downarrow$

$t \geq \frac{m+1}{2mb} = \Omega\left(\frac{m}{mb}\right)$.

m? m = #edges in H.

Magic argument: For every m, there exists a graph H with m vertices & $\geq \frac{1}{2}m^{\frac{3}{2}}$ edges. where H has no C_4/C_6.

$t = \Omega\left(\frac{m^{3/2}}{mb}\right) = \Omega\left(\frac{\sqrt{n}}{b}\right)$

We have proved that $\text{TIME}(C_4/C_6) = \Omega(\sqrt{n})$ in the b-CONGEST model.

That's All