

Homework 12- Due Dec 10

READING: Notes **The Word Probe Model: Membership** and **The Word Probe Model: Evaluating a Polynomial**. (Only the parts we did in class.)

1. (25 points) In this problem we will work with $(U, n; q)$ -WPDSY except that the queries are NON-ADAPTIVE. (So you ask them all at once.)
 - (a) Assume that the SETUP phrase puts the elements into the array in sorted order. Find a function f such that the following is true. If $U \geq f(n)$ then $q \geq n$.
 - (b) Make no assumption on the SETUP phase. Find a function g such that the following is true. If $U \geq g(n)$ then $q \geq n$. Make your bound reasonable (that is, do not use Ramsey's theorem)

ANSWER:

a) We claim $f(n) = 2n + 2$. Assume, by way of contradiction, that there is a $(U, n; n - 1)$ WPDSY with $U \geq 2n + 2$. By restricting the algorithm to $[2n + 2]$ we can assume $U = 2n + 2$.

When the question " $n + 1 \in A$?" is asked, let i be such that $CELL[i]$ is NOT probed. Note that $1 \leq i \leq n$.

Let

$$X = \{1, \dots, i\} \cup \{n + 1\} \cup \{n + 3, \dots, 2n + 1 - i\}.$$

$$Y = \{1, \dots, i\} \cup \{n + 2\} \cup \{n + 3, \dots, 2n + 1 - i\}.$$

If A is stored then we have $CELL[i] = a_i$ where $\{a_1, \dots, a_n\} = A$. If B is stored then we have $CELL[i] = b_i$ where $\{b_1, \dots, b_n\} = B$. Note that $a_1 = b_1, \dots, a_{i-1} = b_{i-1}, a_{i+1} = b_{i+1}, \dots, a_n = b_n$. Hence, when the query $n + 1 \in A$ is made, or $n + 1 \in B$ is made, If they query " $n + 1 \in A$?" is made then if the actual set stored is X or Y the same answers will be returned. This is a contradiction since $n + 1 \in X - Y$.

b) We assume that n divides U . (If not this is a minor modification.) We determine a condition on U that makes the theorem true (the condition will be that $U \geq g(n)$ for some $g(n)$ to be determined.)

Assume that there is a $(U, n; n - 1)$ WPDS (Yao Model). Map each element $u \in [U]$ to the index i such that if “ $u \in A?$ ” is asked then $CELL[i]$ is NOT probed. There will be U/n elements that all map to the same i . Let X be that set of U/n elements.

We map $A \in \binom{X}{n}$ to $[X]^{n-1}$ as follows. Let $A = \{a_1, \dots, a_n\}$. There exists b_1, \dots, b_n such that $\{a_1, \dots, a_n\} = \{b_1, \dots, b_n\}$ and, to store A the SETUP sets $CELL[i] = b_i$. We map $A \in \binom{X}{n}$ to the $n - 1$ -tuple $(b_1, \dots, b_{i-1}, b_{i+1}, \dots, b_n)$.

The domain has size $\binom{U/n}{n}$. The co-domain has size $\frac{(U/n)!}{(U/n - (n-1))!}$.

Let $V = U/n$. Now we have the domain has size $\binom{V}{n}$. The co-domain has size $\frac{V!}{(V-n+1)!}$.

CLAIM: The function is an injection.

PROOF OF CLAIM:

Assume, BWOC, that there are two different sets A_1, A_2 of size n that map to the same thing. Then

- For all $x \in A_1 \cup A_2$, if x is asked about then $CELL[i]$ is not probed.
- $A_1 \neq A_2$.
- A_1 and A_2 lead to the same array except at $CELL[i]$.

Let $x \in A_1 \oplus A_2$. The query algorithm will make a mistake on the query $x \in A$ for either $A = A_1$ or $A = A_2$. This is a contradiction.

END OF PROOF OF CLAIM

Since the map is an injection we must have that the domain is \leq the co-domain. Hence

$$\begin{aligned} \binom{V}{n} &\leq \frac{V!}{(V-n+1)!} \\ \frac{V!}{n!(V-n)!} &\leq \frac{V!}{(V-n+1)!} \\ \frac{1}{n!(V-n)!} &\leq \frac{1}{(V-n+1)(V-n)!} \\ \frac{1}{n!} &\leq \frac{1}{(V-n+1)}. \end{aligned}$$

$$V - n + 1 \geq n!$$

$$V \geq n! + n - 1$$

$$U/n \geq n! + n - 1$$

$$U \geq n(n! + n - 1)$$

Hence we can take $g(n) = n(n! + n - 1)$.

End of Proof

2. (25 points) Show that if $U = n + 2$ then there is an $O(1)$ probe Data Structure in the Yao Model. You can assume $U \geq 20$. (Give a SIMPLE algorithm- so do not use the one in the Yao Paper, which we did not cover in class.)

ANSWER:

SETUP: We want to store A . Let $x, y \in [n+2] - A$ be such that $x < y$.
Let

$$A_1 = A \cap \{1, \dots, \lfloor \frac{U}{2} \rfloor\},$$

$$A_2 = A \cap \{\lfloor \frac{U}{2} \rfloor + 1, \dots, U\},$$

There are two cases. We will put at $CELL[1]$ an element from A_1 or A_2 to tell the user which case we are in.

CASE 1: $x + 1 \bmod U \in A$. (So x, y are not adjacent.) Let

$$\begin{aligned} CELL[1] &= \text{least element of } A_1 - \{x + 1 \bmod U, y + 1 \bmod U\} \\ CELL[2] &= x + 1 \bmod U \\ CELL[3] &= y + 1 \bmod U \\ CELL[4..n] &= \text{the rest of the elements of } A \text{ in sorted order} \end{aligned}$$

CASE 2: $x = y + 1$. (So x, y are adjacent.)

$$\begin{aligned} CELL[1] &= \text{least element of } A_2 - \{x + 1 \bmod U, y + 1 \bmod U\} \\ CELL[2] &= x - 1 \bmod U \\ CELL[3] &= y + 1 \bmod U \\ CELL[4..n] &= \text{the rest of the elements of } A \text{ in sorted order} \end{aligned}$$

QUERY ALGORITHM: To determine if $a \in A$ do the following: Query $CELL[1], CELL[2], CELL[3]$. There are two cases.

CASE 1: $CELL[1] \in A_1$. Then we know that case 1 was used in the SETUP to store A . Hence we know that $CELL[2] = x + 1 \pmod U$ and $CELL[3] = y + 1 \pmod U$. Hence we know x, y . If $a \in \{x.y\}$ then output NO, else YES.

CASE 2: $CELL[1] \in A_2$. Then we know that case 2 was used in the SETUP to store A . Hence we know that $CELL[2] = x - 1$ and $CELL[3] = y + 1$. Hence we know x, y . If $a \in \{x.y\}$ then output NO, else YES.

3. (25 points) **DEFINITION:** Let $f : A \times B \times C \rightarrow D$. Let $a, b, c, t \in N$. A t -round (a, b, c) -Alice-Bob-Carol protocol for f is a protocol with the following properties.

- (a) Alice goes first. Then Bob. Then Carol.
- (b) There are t rounds. That means that Alice broadcasts/Bob broadcasts/Carol broadcasts happens t times. (Note also that we are assuming that when someone talks they all listen.)
- (c) Alice only sends elements from $[a]$. Bob only sends elements from $[b]$. Carol only sends elements from $[c]$.
- (d) At the end Alice knows the answer (Bob and Carol do not have to).

PROVE THE FOLLOWING LEMMA.

LEMMA: Let $f : A \times B \times C \rightarrow D$. Let $a, b, c, t \in N$. Assume there is a t -round (a, b, c) -Alice protocol for f . Then there exists $X \subseteq A$, $Y \subseteq B$, $Z \subseteq C$ such that the following are true.

- (a) $|X| \geq |A|/a^t$.
- (b) $|Y| \geq |B|/b^t$.
- (c) $|Z| \geq |C|/c^t$.
- (d) For all $x \in X$, for all $y, y' \in Y$, for all $z, z' \in Z$,

$$f(x, y, z) = f(x, y, z') = f(x, y', z) = f(x, y', z').$$

(So on $X \times Y \times Z$ $f(x, y, z)$ only depends on x .)

ANSWER:

We prove this by induction on t .

Base Case: If $t = 0$ then there is no communication but Alice knows the answer. All Alice has is x . Since Alice knows $f(x, y, z)$, $f(x, y, z)$ cannot depend on y or z . We take $X = A$ and $Y = B$ and $Z = C$.

Induction Step: Assume the theorem holds for t . Assume there is a $(t + 1)$ -round (a, b, c) -Alice protocol for f . Alice goes first.

Let $m_A : A \rightarrow [a]$ be defined by

$m_A(x)$ is what Alice broadcasts in the first round if she has x .

For each $i \in [a]$ let

$$X_i = \{x \mid m_A(x) = i\}.$$

The map m_A has domain of size $|A|$ and range of size a . Hence there exists some $i_0 \in [a]$ such that $|X_{i_0}| \geq |A|/a$.

Let $m_B : B \times [a] \rightarrow [b]$ be defined by

$m_B(y, i)$ is what Bob broadcasts in the first round if he has y and sees i from Alice.

For each $j \in [b]$ let

$$Y_j = \{y \mid m_B(y, i_0) = j\}.$$

The map $m_B(-, i_0)$ has domain of size $|Y|$ and range of size b . Hence there exists some $j_0 \in [b]$ such that $|Y_{j_0}| \geq |B|/b$.

Let $m_C : B \times [a] \times [b] \rightarrow [c]$ be defined by

$m_C(z, i, j)$ is what Carol broadcasts in the first round if she sees i from Alice and j from Bob.

For each $k \in [c]$ let

$$Z_k = \{z \mid m_C(z, i_0, j_0) = k\}.$$

The map $m_B(-, i_0, j_0)$ has domain of size $|Z|$ and range of size c . Hence there exists some $k_0 \in [b]$ such that $|Z_{k_0}| \geq |C|/c$.

Look at f restricted to $X_{i_0} \times Y_{j_0} \times Z_{k_0}$. If Alice and Bob and Carol all know i_0, j_0, k_0 ahead of time then they can use these values, and the $(t + 1)$ -round (a, b, c) Alice-protocol, to obtain a t -round (a, b, c) Alice-protocol for f restricted to $X_{i_0} \times Y_{j_0} \times Z_{i_0}$. Inductively apply the theorem to this protocol. Hence there exists $X \subseteq X_{i_0}, Y \subseteq Y_{j_0}, Z \subseteq Z_{k_0}$ such that

- (a) $|X| \geq |X_{i_0}|/a^t \geq |A|/a^{t+1}$.
- (b) $|Y| \geq |Y_{j_0}|/b^t \geq |B|/b^{t+1}$.
- (c) $|Z| \geq |Z_{k_0}|/c^t \geq |C|/c^{t+1}$.
- (d) If $x \in X, y, y' \in Y, z, z' \in Z$ then $f(x, y, z) = f(x, y', z') = f(x, y, z')$.

4. (25 points) State and prove a lemma of the following form. *If there exists $(U, n; q)$ -WPDSY for membership then communication complexity problem XXX has complexity YYY.* (For your own good: Try to use it to prove a theorem about $(U, n; q)$ -WPDSY using the lemma and see what happens.)

ANSWER:

DEFINITION: Let $U, n \in N$. The (U, n) -membership problem is the following communication complexity problem.

- Alice has $u \in U$.
- Bob has $A \in \binom{[U]}{n}$.
- At the end of the protocol Alice knows if $u \in A$. (Bob need not know this.)

END OF DEFINITION

LEMMA: Assume there is a $(U, n; q)$ -WPDSY. Then there is a q -round (n, U) Alice-protocol for (U, n) -membership problem.

PROOF:

- (a) Alice has u , Bob has $A \in \binom{[U]}{n}$.
- (b) Bob sets up cells $\text{CELL}[1], \dots, \text{CELL}[n]$ as the data structure dictates.
- (c) Alice and Bob simulate the data structure as follows:
 - i. If the data structure asks about cell i , Alice sends $i \in \{1, \dots, n\}$ to Bob.
 - ii. Bob sends back $\text{CELL}[i]$ (which is an element of $[U]$).

Note that since the data structure allows q probes, the protocol will go q rounds. At the end of the simulation Alice knows if $u \in A$ or not.

END OF PROOF

5. EXTRA CREDIT: Find a reasonable function f (do not use Ramsey) such the following is true:

For almost all U , if there is an ADAPTIVE $(U, n; 2)$ then $U \leq f(n)$.