A Use of Ramsey Theory in Comm. Comp.

This is from *An Application of Hindman's Theorem to a Problem in Commmunication Complexity* by Pavel Pudlak, which appeared in [1].

1 Introduction

Def 1.1 Let Σ be an alphabet. Let $n \in \mathsf{N}$. Let $f : \Sigma^n \times \Sigma^n \times \Sigma^n \rightarrow D$ be a function. The *Simulataneous Message Complexity of* f, denoted SM(f), is defined as follows. Alice has x, Bob has y, and Carol has z. They will all send a message to THEMAN. This is the only communication. THEMAN then computes $f(x, y, z)$ from the three messagses send to him. SM(f) is the length of the longest message send.

Notation 1.2 Let $\Sigma = \{a, b, c, \eta\}$ throughout. Let $s \in \Sigma^*$. $s_{a\leftarrow\eta}$ means take the string s and replace the a's with η's. Similar for $s_{b\leftarrow\eta}$, $s_{c\leftarrow\eta}$.

Def 1.3 *ABC* is the following function. On input $(s_{a\leftarrow\eta}, s_{b\leftarrow\eta}, s_{c\leftarrow\eta})$ determine if $s \in \Sigma^* a \Sigma^* b \Sigma^* c \Sigma^*$. (If the input is not of the form $(s_{a\leftarrow\eta}, s_{b\leftarrow\eta}, s_{c\leftarrow\eta})$ then the output can be arbitrary.) We'll say things like "Alice gets s" when actually "Alice gets $s_{a\leftarrow\eta}$" is accurate.

2 Upper Bound

Theorem 2.1 $\mathrm{SM}(ABC) = \log\log n + O(1)$.

Proof: Let s be the input. Let P_A be the position of the first a in s. Let P_C be the position of the last c in s. Let P_{aB} be the position of the first b after the first a if it exists. Let P_{Bc} be the position of the last b before the last c if it exists.

If $s \in ABC$ then

$$\mathrm{P}_A < \mathrm{P}_{aB} \leq \mathrm{P}_{Bc} < \mathrm{P}_C.$$

Hence

$$(\mathrm{P}_{aB} - \mathrm{P}_A) + (\mathrm{P}_C - \mathrm{P}_{Bc}) \leq (\mathrm{P}_C - \mathrm{P}_A)$$

If $s \notin ABC$ then

$$\mathrm{P}_{Bc} < \mathrm{P}_A < \mathrm{P}_C < \mathrm{P}_{aB}.$$

Hence

$$(\mathrm{P}_C - \mathrm{P}_{Bc} > \mathrm{P}_C - \mathrm{P}_A) \wedge (\mathrm{P}_{aB} - \mathrm{P}_A > \mathrm{P}_C - \mathrm{P}_A).$$

We rewrite this:

$$s \in ABC \Rightarrow (\mathrm{P}_{aB} - \mathrm{P}_A) + (\mathrm{P}_C - \mathrm{P}_{Bc}) \le (\mathrm{P}_C - \mathrm{P}_A).$$

$$s \notin ABC \Rightarrow (((\mathrm{P}_C - \mathrm{P}_{Bc}) > (Pc - \mathrm{P}_A) \wedge ((\mathrm{P}_{aB} - \mathrm{P}_A) > (\mathrm{P}_C - \mathrm{P}_A))).$$

We will refer to the above statements as THE DICHOTOMY. We will take advantage of THE DICHOTOMY and refer back to it.

1. Alice has $s_{a \leftarrow \eta}$, Bob has $s_{b \leftarrow \eta}$ and Carol has $s_{c \leftarrow \eta}$.

2. Alice computes $\mathrm{P}_C - \mathrm{P}_{Bc}$. Alice then computes the most position of the most significant 1-bit of $\mathrm{P}_C - \mathrm{P}_{Bc}$. This is denoted $\mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_{Bc})$. Alice sends $\mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_{Bc})$ to THEMAN.

3. Bob computes $\mathrm{P}_C - \mathrm{P}_A$. Bob then computes the most position of the most significant 1-bit of $\mathrm{P}_C - \mathrm{P}_A$. This is denoted $\mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$. Bob sends $\mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$ to THEMAN.

4. Carol computes $\mathrm{P}_{aB} - \mathrm{P}_A$. Carol then computes the most position of the most significant 1-bit of $\mathrm{P}_{aB} - \mathrm{P}_A$. This is denoted $\mathrm{SIG}(\mathrm{P}_{aB} - \mathrm{P}_A)$. Carol sends $\mathrm{SIG}(\mathrm{P}_{aB} - \mathrm{P}_A)$ to THEMAN.

5. (a) If $\mathrm{SIG}(\mathrm{P}_{aB} - \mathrm{P}_A) > \mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$ or $\mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_{Bc}) > \mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$ then REJECT.
   (b) If $\mathrm{SIG}(\mathrm{P}_{aB} - \mathrm{P}_A) < \mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$ or $\mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_{Bc}) < \mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$ then ACCEPT.
   (c) If $\mathrm{SIG}(\mathrm{P}_{aB} - \mathrm{P}_A) = \mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$ and $\mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_{Bc}) = \mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$ (the only remaining case) then REJECT.

How many bits are communicated? The numbers $\mathrm{P}_A, \mathrm{P}_C, \mathrm{P}_{aB}, \mathrm{P}_{Bc}$ are all $\le n$ so they take $\log n + O(1)$ bits. Hence a position in them takes $\log \log n + O(1)$ bits.

Why does the protocol work?

Case 1: $\mathrm{SIG}(\mathrm{P}_{aB} - \mathrm{P}_A) > \mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$. From this we easily see

$$(\mathrm{P}_{aB} - \mathrm{P}_A) > (\mathrm{P}_C - \mathrm{P}_A).$$

Hence

$$(\mathrm{P}_{aB} - \mathrm{P}_A) + (\mathrm{P}_C - \mathrm{P}_{aB}) \not\le \mathrm{P}_C - \mathrm{P}_A.$$

By THE DICHOTOMY $s \notin ABC$.

Case 2: $\mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_{Bc}) > \mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$. Similar to Case 1.

Case 3: $\mathrm{SIG}(\mathrm{P}_{aB} - \mathrm{P}_A) < \mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$ From this we easily see that

$$(\mathrm{P}_{aB} - \mathrm{P}_A) < (\mathrm{P}_C - \mathrm{P}_A).$$

By the DICHOTOMY statement we get $s \in ABC$. (Look at the $s \notin ABC$ part and take the contrapositive.)

Case 4: $\mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_{Bc}) < \mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$ Similar to Case 3.

Case 5: $\mathrm{SIG}(\mathrm{P}_{aB} - \mathrm{P}_A) = \mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$ and $\mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_{Bc}) = \mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A)$. From this we easily see $\mathrm{SIG}((\mathrm{P}_{aB} - \mathrm{P}_A) + (\mathrm{P}_C - \mathrm{P}_{Bc})) = \mathrm{SIG}(\mathrm{P}_C - \mathrm{P}_A) + 1$ Hence $(\mathrm{P}_{Bc} - \mathrm{P}_A) + (\mathrm{P}_C - \mathrm{P}_{aB}) > \mathrm{P}_C - \mathrm{P}_A$. By THE DICHOTOMY $s \notin ABC$. ∎

Can we do better?

3 Lower Bound

We use the following version of Ramsey's Theorem.

Def 3.1 $[n]^2$ is the set of all unordered pairs of elements of $\{1, \ldots, n\}$. Let $COL \in \mathsf{N}$, $COL \geq 2$. Let $f : [n]^2 \rightarrow \{1, \ldots, COL\}$. A set $H \subseteq \{1, \ldots, n\}$ is *Homogenous with respect to* f if there exists a color $i \in COL$ such that, $(\forall x, y \in H)[f(x, y) = i]$.

Lemma 3.2 *Let $COL \in \mathsf{N}$, $COL \geq 2$. Let $f : [n]^2 \rightarrow \{1, \ldots, COL\}$. There exists a homogenous set H such that $|H| \geq \frac{1}{2}\frac{\log n}{\log COL}$. (The $\frac{1}{2}$ can be replaced with any $\delta < 1$.)*

Theorem 3.3 $\mathrm{SM}(ABC) = \Omega(\log\log n)$.

Proof: Assume $\mathrm{SM}(ABC) = m = c \log\log n$. (will detemine c later).

For $1 \leq i < j \leq n$ let s_{ij} denote the string that has η at all positions except the ith, where it has an a and the jth where it has a c. We have a picture of it below which serves as an example of how we will denote strings throughout this proof.

$$\begin{array}{ccccc} \cdots & i & \cdots & j & \cdots \\ \eta\cdots\eta & a & \eta\cdots\eta & c & \eta\cdots\eta \end{array}$$

We color all unordered pairs $\{i, j\}$ with the message Bob sends if he sees s_{ij}. This uses 2^m colors. By Lemma 3.2 there is a homogenous set of size $\frac{1}{2}\frac{\log n}{\log 2^m} = \frac{\log n}{2m}$. We remove every other element so there is at least one number between every element. This set, H, has size $h = \frac{\log n}{4m}$. We can assume h is even.

$$H = \{i_1 < i_2 < \cdots < i_h\}.$$

We create the string $MAIN$ pictured below

$$\begin{array}{ccccccc} \cdots & i_1\cdots i_2 & \cdots & i_3\cdots i_4 & \cdots & i_5\cdots i_6 & \cdots \\ b\cdots b & \eta\cdots\eta & b\cdots b & \eta\cdots\eta & b\cdots b & \eta\cdots\eta & \cdots \end{array}$$

For all odd j, $1 \le j \le h-1$, let s_j be the string obtained by taking $MAIN$ and replacing the η in the i_jth position with an a and the η in the i_{j+1}th position with a c. We show s_3 below:

$$\begin{array}{ccccccc} \cdots & i_1\cdots i_2 & \cdots & i_3\cdots i_4 & \cdots & i_5\cdots i_6 & \cdots \\ b\cdots b & \eta\cdots\eta & b\cdots b & a\eta\cdots\eta c & b\cdots b & \eta\cdots\eta & \cdots \end{array}$$

Given j, form s_j, and map to the ordered pair formed by taking (1) first component is message Alice would send to THEMAN on input s_j, (2) second component is message Carol would send to THEMAN on input s_j.

This map has domain of size $\frac{h}{2}$ and range of size 2^{2m}. Unravelling the definitions this has domain of size

$$\frac{h}{2} = \frac{\log n}{8m} = \frac{\log n}{8c\log\log n},$$

and range of size

$$2^{2m} = 2^{2c\log\log n} = 2^{\log(\log n)^{2c}} = (\log n)^{2c}.$$

We take $c < \frac{1}{2}$. Now the domain is larger than the range so there exists $j,k \in \mathsf{N}$ such that the strings s_j and s_k that map to the same ordered pair. Hence we have the following.

1. Alice transmits the same message given s_j or s_k. Call this message m_A.
2. Carol transmits the same message given s_j or s_k. Call this message m_C.

Note that we also know the following

1. Since $i_j, i_{j+1}, i_k, i_{k+1} \in H$ Bob transmits the same message given s_j or s_k. Call this message m_B.
2. Since $s_j \notin ABC$, when THEMAN gets (m_A, m_B, m_C) he will reject.

We form a NEW string s which is $MAIN$ except that we have an a in i_jth position and a c in i_{k+1}th position. The following picture summarizes what we have. The notation $\cdots b\cdots\eta\cdots$ means a string of b's and η's. When it appears in the same column it means the same string. When the notation $\cdots\eta\cdots\eta\cdots$ appears in that same column it means that we take that string of b's and η's and replace all the b's with η's.

string	$\cdots$	$i_j \cdots i_{j+1}$	$\cdots$	$i_k \cdots i_{k+1}$	$\cdots$
s_j	$\cdots b \cdots \eta \cdots$	$a\eta \cdots \eta c$	$\cdots b \cdots \eta \cdots$	$\eta\eta \cdots \eta\eta$	$\cdots b \cdots \eta \cdots$
Alices's s_j	$\cdots b \cdots \eta \cdots$	$\eta\eta \cdots \eta c$	$\cdots b \cdots \eta \cdots$	$\eta\eta \cdots \eta\eta$	$\cdots b \cdots \eta \cdots$
Bob's s_j	$\cdots \eta \cdots \eta \cdots$	$a\eta \cdots \eta c$	$\cdots \eta \cdots \eta \cdots$	$\eta\eta \cdots \eta\eta$	$\cdots \eta \cdots \eta \cdots$
Carol's s_j	$\cdots b \cdots \eta \cdots$	$a\eta \cdots \eta\eta$	$\cdots b \cdots \eta \cdots$	$\eta\eta \cdots \eta\eta$	$\cdots b \cdots \eta \cdots$
s_k	$\cdots b \cdots \eta \cdots$	$\eta\eta \cdots \eta\eta$	$\cdots b \cdots \eta \cdots$	$a\eta \cdots \eta c$	$\cdots b \cdots \eta \cdots$
Alices's s_k	$\cdots b \cdots \eta \cdots$	$\eta\eta \cdots \eta\eta$	$\cdots b \cdots \eta \cdots$	$\eta\eta \cdots \eta c$	$\cdots b \cdots \eta \cdots$
Bob's s_k	$\cdots \eta \cdots \eta \cdots$	$\eta\eta \cdots \eta\eta$	$\cdots \eta \cdots \eta \cdots$	$a\eta \cdots \eta c$	$\cdots \eta \cdots \eta \cdots$
Carol's s_k	$\cdots b \cdots \eta \cdots$	$\eta\eta \cdots \eta\eta$	$\cdots b \cdots \eta \cdots$	$a\eta \cdots \eta\eta$	$\cdots b \cdots \eta \cdots$
s	$\cdots b \cdots \eta \cdots$	$a\eta \cdots \eta\eta$	$\cdots b \cdots \eta \cdots$	$\eta\eta \cdots \eta c$	$\cdots b \cdots \eta \cdots$
Alices's s	$\cdots b \cdots \eta \cdots$	$\eta\eta \cdots \eta\eta$	$\cdots b \cdots \eta \cdots$	$\eta\eta \cdots \eta c$	$\cdots b \cdots \eta \cdots$
Bob's s	$\cdots \eta \cdots \eta \cdots$	$a\eta \cdots \eta\eta$	$\cdots \eta \cdots \eta \cdots$	$\eta\eta \cdots \eta c$	$\cdots \eta \cdots \eta \cdots$
Carol's s	$\cdots b \cdots \eta \cdots$	$a\eta \cdots \eta\eta$	$\cdots b \cdots \eta \cdots$	$\eta\eta \cdots \eta\eta$	$\cdots b \cdots \eta \cdots$

Lets look at the string s very carefully.

1. Alices's view of s and s_k are the same. Hence on s Alice sends m_A.

2. Carol's view of s and s_j are the same. Hence on s Carol sends m_C.

3. Since $i_j, i_{j+1}, i_k, i_{k+1}$ are in the homogenous set Bob transmits the same message on s_j, s_k, s. Hence Bob transmits m_B.

4. By the above Alice, Bob, and Carol transmit to THEMAN the same info on s_j, s_k, s. But $s \in ABC$ so this is a contradition.

∎

4 What Else is Known?

We looked at the ABC problem with 3 people. We could also look at the $ABCD$ problem with 4 people, the $ABCDE$ problem with 5 people, etc.

1. SM(ABC) = $\Theta(\log\log n)$.

2. SM($ABCD$) is not constant. No non-trivial upper bound is known.

3. SM($ABCDE$) is not constant. No non-trivial upper bound is known.

4. SM($ABCDEF$) and beyond. Nothing is known. It has been conjectured that it is not constant.

The proof that SM($ABCD$), SM($ABCDE$) are not constant uses a Ramsey-Type theorem called Hindman's theorem.

References

[1] P. Pudlak. An application of hindman's theorem to a problem on communication complexity. *Combinatorics, Probability and Computing*, 12, 2003.