# Polynomal-size Frege proofs of Bollobás' theorem on the trace of sets

By Akihiro NOZAKI*), Toshiyasu ARAI**) and Noriko H. ARAI***)

(Communicated by Heisuke HIRONAKA, M.J.A., Sept. 12, 2008)

**Abstract:** In this note we show that Bollobás' theorem on the trace of sets has polynomial-size Frege proofs.

**Key words:** Lengths of proofs; Frege system; trace of sets.

**1. Frankl's theorem on the trace of sets.** In seeking natural, combinatorial problems that are candidates for separating Frege and extended Frege proof systems, Bonet, Buss and Pittasi [2] concluded that P. Frankl's theorem on the trace of sets [3] is the only example they found which are known to have polynomial-size extended Frege proofs and for which they have no reason to suspect that they have subexponential-size Frege proofs.

Let us introduce some notations to state Frankl's theorem. $|X|$ denotes the cardinality of finite sets $X$. Throughout this paper *matrices* are matrices over $\{0,1\}$, i.e., entries of matrices are either 0 or 1. Let $A = (a_{ij})$ be an $m \times n$-matrix. $\boldsymbol{a}_i$ denotes the $i$-th row $(a_{i1} \dots a_{in})$. Each vector $\boldsymbol{a}_i$ is identified with the set $X(\boldsymbol{a}_i) := \{j : a_{ij} = 1\}$, and the matrix $A$ with the family $\mathcal{F}(A) := \{X(\boldsymbol{a}_i) : 1 \leq i \leq m\}$ of subsets of $[n] := \{1, \dots, n\}$. For a subset $Y$ of $[n]$, the trace $\mathcal{F}(A)|_Y$ of $\mathcal{F}(A)$ on $Y$ is defined to be the family of subsets $\{X(\boldsymbol{a}_i) \cap Y : 1 \leq i \leq m\}$ of $Y$. In other words, $\mathcal{F}(A)|_Y = \mathcal{F}(A|_Y)$ where $A|_Y$ denotes the $m \times n$-matrix erasing all 1-entries on the columns not in the set $Y$: $(i,j)$-entry of $A|_Y$ is 1 iff $a_{ij} = 1$ & $j \in Y$.

Then the arrow notation

$$(m,n) \to (r,s)$$

designates that for any $m \times n$-matrix $A$ of distinct rows (, i.e., $|\mathcal{F}(A)| = m$), there exists a set $Y \subseteq [n]$ of columns such that $|Y| = s$ and at least $r$ rows differ from each other on the columns $Y$: $|\mathcal{F}(A|_Y)| \geq r$. Observe that

(a) if $(m,n) \to (r,s)$ and $m' > m$, then $(m',n) \to (r,s)$.

(b) If $(m,n) \to (m-r,s)$ and $m' < m \leq 2^n$, then $(m',n) \to (m'-r,s)$.

Let $F(n,t)$ denote the maximum $m$ for which $(m,n) \to (m - 2^{t-1} + 1, n-1)$ holds. Thus $F(n,t) \geq m$ iff for any $m \times n$-matrix $A$ of distinct rows we can find a column such that, if this column is deleted, the resulting $m \times (n-1)$-matrix will contain at most $2^{t-1} - 1$ pairs of equal rows.

P. Frankl [3] showed the following theorem.

**Theorem 1.** $F(n,t) \geq n\frac{(2^t-1)}{t}$.

The simplest case $t = 1$ of Theorem 1 is Bondy's Theorem, and the next case $t = 2$ is due to Bollobás [5].

A brief outline of his proof is as follows: Call a matrix $A$ of distinct rows *hereditary* if erasing any 1 entry causes two rows to become identical. Namely $\forall i,j[a_{ij} = 1 \Rightarrow \exists k \neq i \forall l \neq j(a_{il} = a_{kl})]$.

Frankl first shows that it suffices to prove Theorem 1 for hereditary matrices: starting from a given family of sets violating Theorem 1, iterating the *down-shift*, cf. [4], produces a hereditary familly of sets violating the same theorem. He then gives a proof of the theorem for hereditary matrices based on a corollary to the Kruskal-Katona theorem.

Bonet, Buss and Pittasi [2] show that the (propositional tautologies translating the) corollary to the Kruskal-Katona theorem have polynomial-size Frege proofs. Thus there are polynomial-size Frege proofs of Theorem 1 for hereditary matrices.

Moreover they mentioned two special cases of Theorem 1, $t = 1, 2$. Bondy's theorem had been suggested by J. Krajíček as a candiate for an exponential separation between Frege and extended Frege systems. However the fact $F(n,1) \geq n$ was

2000 Mathematics Subject Classification. Primary 03F20; Secondary 05D05.

*) Corresponding address: 2-944-1 Kabe, Ome, Tokyo 198-0036, Japan.
**) Graduate School of Engineering, Kobe University, 1-1 Rokko-dai, Kobe, Hyogo 657-8501, Japan.
***) National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan.

shown to have polynomial-size Frege proofs in [2]. After that they wrote

> The second is when $t = 2$ and $m \le 3n/2$: we have not been able to find subexponential-size Frege proofs even for this case. (p.40, [2])

Also they noted

> In fact, the only good combinatorial candiates we have found are based on Frankl's theorem (even for the $t = 2$ case). However, in the past a similar state of affiars has held for the pigeonhole principle and for Bondy's theorem and, subsequently, polynomial-size Frege proofs for these have been found. Thus, it is not unlikelly that further progress will find polynomial-size Frege proofs of the tautologies based on Frankl's theorem. (p.53, [2])

In this note we show the following theorem.

**Theorem 2.** *The fact* $F(n, 2) \ge 3n/2$ *has polynomial-size Frege proofs.*

In section 2, we give an elementary proof of the fact $F(n, 2) \ge 3n/2$. The proof is based on an idea due to the first author, and is seen readily formalizable in the bounded arithmetic *AID*, [1], which yields a polynomial-size Frege proofs of the case.

The idea is to divide the set $[m] = \{1, \ldots, m\}$ of rows into finer equivalence classes step by step.

**Definition 1.** Let $A = (a_{ij})$ be an $m \times n$-matrix, and $Y$ a set of columns, i.e., $Y \subseteq [n]$. Define an equivalence relation $i \equiv_Y k$ on the set $[m]$ of rows as follows:

$$i \equiv_Y k :\Leftrightarrow \forall j \in Y(a_{ij} = a_{kj}).$$

$A/Y$ denotes the set of equivalence classes.

Thus for example, $A/\emptyset = \{[m]\}$, $A/[n] = [m]$, and $A/Z$ is a refinement of $A/Y$ if $Z \supseteq Y$.

## 2. The case $t = 2$.

**2.1. An elementary proof of the case $t = 2$.** In this subsection we give an elementary proof of the fact $F(n, 2) \ge 3n/2$. Put $m = \lceil 3n/2 \rceil$, and let $A = (a_{ij})$ be an $m \times n$-matrix of distinct rows. We have to find a column such that, if this column is deleted, the resulting $m \times (n-1)$-matrix will contain at most a pair of equal rows.

Suppose contrarily that there is no such column. This means for any column $j$ we can find two pairs $\{\langle E(j,i,0), E(j,i,1)\rangle : i = 0, 1\}$ of rows such that the $E(j, i, 0)$-th row differs only on the $j$-th column from the $E(j, i, 1)$-th row: $a_{E(j,i,k),j} = k$ for $k = 0, 1$ and $E(j, i, 0) \equiv_Z E(j, i, 1)$ for $Z = [n] - \{j\}$.

**Lemma 3.** *For any* $Y \subseteq [n]$ *and any* $j \notin Y$, *if* $E(j, 0, 0) \not\equiv_Y E(j, 1, 0)$, *then* $|A/(Y \cup \{j\})| - |A/Y| \ge 2$, *and if* $E(j, 0, 0) \equiv_Y E(j, 1, 0)$, *then* $|A/(Y \cup \{j\})| - |A/Y| \ge 1$.

First consider the case when $E(j, 0, 0) \not\equiv_Y E(j, 1, 0)$. This means that the columns in $Y$ have already distinguished the row $E(j, 0, k)$ from the row $E(j, 1, k)$. Then there are two equivalent classes under $\equiv_Y$ one of which contains rows $\{E(j, 0, k) : k = 0, 1\}$, and the other contains $\{E(j, 1, k) : k = 0, 1\}$, but each class splits up into two equivalence classes under $\equiv_{Y\cup\{j\}}$. Hence $|A/(Y \cup \{j\})| - |A/Y| \ge 2$.

Next consider the case when $E(j, 0, 0) \equiv_Y E(j, 1, 0)$. There is one equivalent class under $\equiv_Y$ which contains rows $\{E(j, i, k) : i, k = 0, 1\}$. This class splits up into two equivalence classes under $\equiv_{Y\cup\{j\}}$. Hence $|A/(Y \cup \{j\})| - |A/Y| \ge 1$. This shows Lemma 3.

Now let us divide the set $[n]$ into two sets

$$Y_\ell := \{j \in [n] : E(j, 0, 0) \not\equiv_{[j-1]} E(j, 1, 0)\}$$
$$Y_r := \{j \in [n] : E(j, 0, 0) \equiv_{[j-1]} E(j, 1, 0)\}$$

for $[j - 1] = \{1, \ldots, j - 1\}$.

Obviously $|Y_\ell| + |Y_r| = n$.

Let $]k[ := \{k, \ldots, n\}$. Suppose $j \in Y_r$. This means that the entries of the $E(j, 0, k)$-th row coincide with ones of $E(j, 1, k)$-th row on the columns $[j - 1]$. Since these rows are distinct, we can find another column $p \notin [j]$, i.e., $p \in ]j + 1[$ which distinguishes the row $E(j, 0, k)$ from the row $E(j, 1, k)$ simultaneously: $a_{E(j,0,k),p} \ne a_{E(j,1,k),p}$ for any $k = 0, 1$. Thus we have shown

$$j \in Y_r \Rightarrow E(j, 0, 0) \not\equiv_{]j+1[} E(j, 1, 0).$$

Hence by Lemma 3, if $j \in Y_\ell$, then $|A/[j]| - |A/[j - 1]| \ge 2$ and $|A/]j[| - |A/]j + 1[| \ge 1$. Moreover if $j \in Y_r$, then $|A/[j]| - |A/[j - 1]| \ge 1$ and $|A/]j[| - |A/]j + 1[| \ge 2$.

Therefore

$$|A/[n]| - |A/[0]| \ge 2|Y_\ell| + |Y_r| \tag{1}$$

and

$$|A/]1[| - |A/]n + 1[| \ge |Y_\ell| + 2|Y_r|. \tag{2}$$

Thus we get a contradiction $2(m - 1) = (|A/[n]| - |A/[0]|) + (|A/]1[| - |A/]n + 1[|) \ge 3(|Y_\ell| + |Y_r|) = 3n$ since $m \le 3n/2$.

**Remark.** In fact the bound is tight, i.e.,

$F(n,2) = \lceil 3n/2 \rceil$. As Frankl [3] noted that $F(n,t) = n\frac{(2^t-1)}{t}$ holds if $t$ divides $n$. For readers' convenience let us reproduce his counterexmple. Put $m = n\frac{(2^t-1)}{t}$ and define $(m+1) \times n$-matrix $A$ as follows: Let $K$ denote the $(2^t-1) \times t$-matrix which is obtained from a complete binary tree of depth $t$ by deleting the zero branch. Namely the $i$-th row denotes the number in binary notation for $1 \le i < 2^t$. Make $n/t$ copies of $K$, and arrange these on the diagonal of the $n/t$ square zero matrix, and finally append the zero row to the lowest:

$$A = \begin{pmatrix} K & O & \cdots & O \\ O & K & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & K \\ O_{1t} & O_{1t} & \cdots & O_{1t} \end{pmatrix}$$

where $O_{1t}$ denotes the $1 \times t$-zero matrix.

Then if any column in $A$ is deleted, the resulting $m \times (n-1)$-matrix will contain $2^{t-1}$ pairs of equal rows. Note that the matrix $A$ is hereditary.

Next consider the case when $t = 2$ and $n = 2k+1$. Then $\lceil 3n/2 \rceil = 3k+2$. First let

$$K = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and let $B$ denote the following matrix:

$$B = \begin{pmatrix} K & O & \cdots & O & O_{3,1} \\ O & K & \cdots & O & O_{3,1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & \cdots & K & O_{3,1} \\ O_{1,2} & O_{1,2} & \cdots & O_{1,2} & 0 \\ O_{1,2} & O_{1,2} & \cdots & O_{1,2} & 1 \\ A_{k+3,1} & O_{1,2} & \cdots & O_{1,2} & 1 \end{pmatrix}$$

where $O$ denote the $3 \times 2$-zero matrix, $O_{3,1}$ the $3 \times 1$-zero, $O_{1,2}$ the $1 \times 2$-zero matrix and $A_{k+3,1}$ is the vector $(1\ 0)$. Then if any column in $B$ is deleted, the resulting matrix will contain two pairs of equal rows. Note that the matrix $B$ is again hereditary.

**2.2. Formailizability in *AID*.** In this subsection we briefly discuss the formalizability of the proof given in the subsection 2.1.

Any $m \times n$-matrix over $\{0,1\}$ is coded by a natural number in such a way that its $(i,j)$-entry is a bit of the number. Then the arrow notation $(m,n) \to (m - 2^{t-1} + 1, n-1)$ can be expressed by a $\forall\Sigma_0^b$-sentence in *AID* since bounded vector summation of any $\Sigma_0^b$-bitdefinable function, and bounded counting of any $\Sigma_0^b$-formula are $\Sigma_0^b$-bitdefinable in *AID*. Likewise for any given set $Y$ of columns of a matrix $A$, the number $|A/Y|$ of equivalence classes under the equivalent relation $\equiv_Y$ is $\Sigma_0^b$-bitdefinable by counting, e.g., the number of rows which are not equivalent to any preceeding rows.

Thus bounded vector summations suffice to prove Lemma 3, (1), (2), and to deduce the contradiction.

Therefore our proof is formalizable in *AID*, and this yields polynomial-size Frege proofs of tautologies derived from the case $t = 2$ of the Frankl's Theorem 1 $F(n,2) \ge 3n/2$.

**Acknowledgments.** We would like to thank the anonymous referee for pointing out an error and giving useful comments and suggestions.

## References

[ 1 ] T. Arai, A bounded arithmetic *AID* for Frege systems, Ann. Pure Appl. Logic **103** (2000), nos. 1–3, 155–199.

[ 2 ] M. L. Bonet, S. R. Buss and T. Pitassi, Are there hard examples for Frege systems?, in *Feasible mathematics,* II (*Ithaca, NY,* 1992), 30–56, Birkhäuser, Boston, Boston, MA, 1994.

[ 3 ] P. Frankl, On the trace of finite sets, J. Combin. Theory Ser. A **34** (1983), no. 1, 41–45.

[ 4 ] P. Frankl, Extremal set systems, in *Handbook of combinatorics, Vol.* 1, 2, 1293–1329, Elsevier, Amsterdam, 1995.

[ 5 ] L. Lovász, *Combinatorial problems and exercises*, North-Holland, Amsterdam, 1979.