A NEW CONSTRUCTION FOR CANCELLATIVE FAMILIES OF SETS

James B. Shearer
IBM Research Division
T.J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598
email: jbs@watson.ibm.com

Submitted: March 25, 1996; Accepted: April 20, 1996

Abstract. Following [2], we say a family, H, of subsets of a n-element set is cancellative if $A \cup B = A \cup C$ implies $B = C$ when $A, B, C \in H$. We show how to construct cancellative families of sets with $c2^{.54797n}$ elements. This improves the previous best bound $c2^{.52832n}$ and falsifies conjectures of Erdös and Katona [3] and Bollobas [1].

AMS Subject Classification. 05C65

We will look at families of subsets of a n-set with the property that $A \cup B = A \cup C \Rightarrow B = C$ for any A, B, C in the family. Frankl and Füredi [2] call such families cancellative. We ask how large cancellative families can be. We define $f(n)$ to be the size of the largest possible cancellative family of subsets of a n-set and $f(k, n)$ to be the size of the largest possible cancellative family of k-subsets of a n-set.

Note the condition $A \cup B = A \cup C \Rightarrow B = C$ is the same as the condition $B \triangle C \subseteq A \Rightarrow B = C$ where $\triangle$ denotes the symmetric difference.

Let F_1 be a family of subsets of a n_1-set, S_1. Let F_2 be a family of subsets of a n_2-set, S_2. We define the product $F_1 \times F_2$ to be the family of subsets of the $(n_1 + n_2)$-set, $S_1 \cup S_2$, whose members consist of the union of any element of F_1 with any element of F_2.

It is easy to see that the product of two cancellative families is also a cancellative family $((A_1, A_2) \cup (B_1, B_2) = (A_1, A_2) \cup (C_1, C_2) \Rightarrow (A_1 \cup B_1, A_2 \cup B_2) = (A_1 \cup C_1, A_2 \cup C_2) \Rightarrow A_1 \cup B_1 = A_1 \cup C_1$ and $A_2 \cup B_2 = A_2 \cup C_2 \Rightarrow B_1 = C_1$ and $B_2 = C_2 \Rightarrow (B_1, B_2) = (C_1, C_2))$. Hence $f(n_1 + n_2) \geq f(n_1)f(n_2)$. Similarly $f(k_1 + k_2, n_1 + n_2) \geq f(k_1, n_1)f(k_2, n_2)$.

It is easy to show that $f(n_1+n_2) \geq f(n_1)f(n_2)$ implies that $\lim_{n\to\infty} \frac{1}{n} lg(f(n))$ exists (lg means log base 2). Let this limit be α. Note that $\alpha \geq \frac{1}{n} lg\ (f(n))$ for any fixed n.

Clearly $f(1,n) = n$ as we may take all the 1-element sets. Let H_n be the family of all 1-element sets of a n-set. It had been conjectured that the largest cancellative families could be built up by taking products of the families H_n. For example Bollobas conjectured [1] that

$$f(k,n) = \prod_{i=1}^{k}[(n+i-1)/k] \tag{1}$$

which comes from letting $n = n_1 + \cdots + n_k$ where the n_i are as nearly equal as possible and considering the family $H_{n_1} \times \cdots \times H_{n_k}$. When $k = 2$ determining $f(2,n)$ is the same as determining how many edges a triangle-free graph can contain. So in this case (1) follows from Turan's theorem. Bollobas [1] proved (1) for $k = 3$. Sidorenko [4] proved (1) when $k = 4$. Frankl and Füredi [2] proved (1) for $n \leq 2k$. However, we will show below that (1) is false in general.

Also Erdös and Katona conjectured (see [3]) that (for $n > 1$) the families achieving $f(n)$ could be built up as products of H_3 and H_2 taking as many H_3's as possible. So for example

$$f(3m) = 3^m. \tag{2}$$

This would mean $\alpha = \frac{lg3}{3} = .52832+$. However, as we will see this conjecture is false as well. In fact we show $\alpha \geq .54797+$.

We now describe the construction which is the main result of this paper. Fix $m \geq 3$. Chose $m-1$ integers $n_1, \ldots, n_{m-1}$ from $\{0,1,2\}$ so that $n_1 + \cdots + n_{m-1} \equiv 0\ mod\ 3$. Chose an integer h from $\{1, \ldots, m\}$. Clearly these choices can be made in $m3^{m-2}$ ways. We now form a cancellative family of subsets of a $3m$-set containing $m3^{m-2}$ elements as follows. Identify subsets of a $3m$-set with 0,1 vectors of length $3m$ in the usual way. Let the $3m$ vectors consist of m subvectors of length 3. Let $v_0 = (100), v_1 = (010), v_2 = (001)$ and $w = (111)$. Form a $3m$-vector from our choices above as follows. Let the hth 3-subvector be w. Let the remaining $m-1$ 3-subvectors be $v_{n_1}, \ldots, v_{n_{m-1}}$ in order. Let F be the family consisting of all $3m$-vectors we can form in this way. Clearly each of the $m3^{m-2}$ choices gives a different vector so F contains $m3^{m-2}$ elements. We claim F is a cancellative family. For let B, C be two different vectors in F and look at $B \triangle C$. We claim $B \triangle C$ contains at least two 3-subvectors with two 1's. There are two cases. If the 3-subvector w is in different positions in B and C then the 3-subvectors in $B \triangle C$ in these positions contain two 1's. Alternatively, if the 3-subvector w is in the same position in B and C then the condition $n_1 + \cdots + n_{m-1} \equiv 0$ mod 3 insures that at least two of the n_i differ between B and C (assuming B and C are distinct) and the 3-subvectors in these positions of $B \triangle C$ contain two 1's. However, this means $B \triangle C \subseteq A \in F$ is impossible (unless $B = C$) because all elements of F contain only one 3-subvector containing two or more 1's.

Hence we have

$$f(3m) \geq m3^{m-2} \tag{3}$$

$$f(m+2,3m) \geq m3^{m-2}. \tag{4}$$

Clearly (3) is better than (2) for $m > 9$. We also have $\alpha \geq \frac{1}{3m} lg(m3^{m-2})$. This is maximized for $m = 24$ giving $\alpha \geq .54797+$. So we have counter examples to the Erdös and Katona conjecture.

Furthermore (4) is better than (1) for $m \geq 8$. So the Bollobas conjecture fails for $k \geq 10$.

The idea of the above construction which improves on products of H_3 can be applied to products of other families as well. For example, we can do better than (1) starting with products of H_k for any $k > 3$ as well. Or we can start with the families F constructed above. This will allow a very slight improvement in the lower bound found for α above.

The best upper bound known for α, $\alpha < lg(3/2) = .58496+$, is due to Frankl and Füredi [2].

The author thanks Don Coppersmith for bringing this problem to his attention.

References

[1] B. Bollobás, *Three-Graphs without two triples whose symmetric difference is contained in a third*, Discrete Mathematics **8** (1974), 21–24.

[2] P. Frankl and Z. Füredi, *Union-free Hypergraphs and Probability Theory*, European Journal of Combinatorics **5** (1984), 127–131.

[3] G.O.H. Katona, *Extremal Problems for Hypergraphs*, Combinatorics, Mathematical Centre Tracts part 2 **56** (1974), 13–42.

[4] A.F. Sidorenko, *The Maximal Number of Edges in a Homogeneous Hypergraph containing no prohibited subgraphs*, Mathematical Notes **41** (1987), 247-259 (translation from Mathematicheskie Zametki **41** (1987), 433-455).