Summer 2014
NSERC USRA Report
Families of Forbidden Configurations
Maxwell Allman

August 29, 2014

This summer, I worked with Dr. Richard Anstee on problems in forbidden configurations, a topic of extremal set theory. The problems involved matrices whose entries are 0 or 1, and for convenience we assume all matrices are of this type. We say a matrix is *simple* if every column of the matrix is unique. For matrices A and F, we say A has F as a *configuration* if there is a submatrix of A which is a row and column permutation of F. Let $\mathcal{F}$ be a family of matrices. We say a matrix $A \in \text{Avoid}(m, \mathcal{F})$ if A has m rows, A is simple, and for each $F \in \mathcal{F}$, A does not have F as a configuration. The central problem is the following: given a certain $\mathcal{F}$, what is the maximum number of columns of a matrix $A \in \text{Avoid}(m, \mathcal{F})$? We denote this value by $\text{forb}(m, \mathcal{F})$. The number of possible unique columns on m rows is 2^m, so $\text{forb}(m, \mathcal{F}) \leq 2^m$, but in fact this function is known to always be bounded above by a polynomial function of m.

In our project, we defined a new operation between two families of k-rowed configurations $\mathcal{F}$ and $\mathcal{G}$. For a matrix A we say $\mathcal{F} + \mathcal{G}$ is a configuration in A if in some k-set of rows of A there is both a configuration from $\mathcal{F}$ and a configuration from $\mathcal{G}$. For a family $\mathcal{F}$ where $\text{forb}(m, \mathcal{F})$ is $\Omega(m^k)$, and a family $\mathcal{G}$ where $\text{forb}(m, \mathcal{G})$ is $O(1)$, we wondered if $\text{forb}(m, \mathcal{F} + \mathcal{G})$ is $O(m^k)$.

Balogh and Bollobás proved the following constant bound where I_k is the $k \times k$ identity matrix, I_k^c is the $k \times k$ (0,1)-complement of I_k and T_k is the $k \times k$ upper triangular (0,1)-matrix with a 1 in row i and column j if and only if $i \leq j$.

Theorem 0.1 *[1] Let k be given. Then there is a constant c_k so that $forb(m, \{I_k, I_k^c, T_k\}) = c_k$.*

We defined the family $\mathcal{G}_{k,t}$ which roughly corresponds to the matrices created by taking k rows of a large identity matrix, k rows of a large identity complement matrix, and k rows from a large triangular matrix. It is known that $\text{forb}(m, \mathcal{F})$ is $O(1)$, following from the Balogh and Bollobás result.

We defined the family $\mathcal{H}_{k,t}$ as the family containing a $k \times t$ matrix of 0's and a $k \times t$ matrix of 1's. We knew $\text{forb}(m, \mathcal{H}_{k,t})$ to be $O(1)$, but this function is unusual in that it does not grow monotonically with m. We found exact results for the maximum of this function over all m for some values of k and t. Interestingly, we also found that for many k-rowed matrices F, $\{F\} + \mathcal{G}_{k,t} = \{F\} + \mathcal{H}_{k,t}$. The latter form is simpler to for the graph theory techniques.

Graph Theory for Two-Rowed Forbidden Families

If a matrix $A \in \text{Avoid}(m, \mathcal{F})$ for some family of two-rowed configurations $\mathcal{F}$, then each pair of rows in A has some relation that causes it to avoid $\mathcal{F}$. We can form a graph G(A) with the rows of A as vertices, and the edges as the relation between the two rows. Here is an example. For a given p, let

$$F = \begin{bmatrix} \overbrace{1\,1 \cdots 1}^{p} \\ 0\,0 \cdots 0 \end{bmatrix}$$

We know $\text{forb}(m, \{F\}) \approx \frac{(p+1)m}{2} + 2$. Now we consider a matrix $A \in \text{Avoid}(m, \{F\} + \mathcal{H}_k)$.

Form the graph G(A), with the edges between rows i and j defined as follows: if F is not a configuration on A restricted to rows i and j, then $i \leftrightarrow j$, and if $\mathcal{H}_k$ is not a configuration on A restricted to rows i and j, then $i == j$. Between any two rows, at least one of these edges must be present, otherwise $\mathcal{F} + \mathcal{H}_k$ is a configuration in A. The edges of G(A) restrict the columns that may be allowed in A, for example, for rows i, j, if $i == j$ then there are at most $2(q-1)$ columns of A where the entry on row i is the same as the entry on row j. We found a list of small graph structure that may not occur in G(A) if the columns of A grows linearly with m. From the absence of these structure we deduced a large structure that must be present on G(A), and used this structure to show the upper bound $\text{forb}(m, \{F\} + \mathcal{H}_k) \leq \frac{(p+1)m}{2} + c$ for some constant c. We then found a construction which shows that $\text{forb}(m, \{F\} + \mathcal{H}_k) \geq \frac{(p+1)m}{2} + (q-p)$, so the maximum number of columns allowed does indeed increase when $\mathcal{H}_k$ is added to $\{F\}$, but only by a constant amount. We used this technique for other two-rowed F with $\text{forb}(m, F)$ being $O(m)$, and found the solution does not differ by more than a constant, unless F contains all 4 possible columns. Then $(m, \{F\} + \mathcal{H}_k) \geq cmq$ for a constant $\frac{1}{2} \leq c \leq 2$.

References

[1] J. Balogh, B. Bollobás, Unavoidable Traces of Set Systems, *Combinatorica*, **25** (2005), 633–643.