Some Hypergraph Questions Suggested by a Nordic Olympiad Problem

Cecil Rousseau
Department of Mathematical Sciences
The University of Memphis
Memphis, Tennessee, USA

Zsolt Tuza
Computer and Automation Institute
Hungarian Academy of Sciences
Budapest, Hungary

ABSTRACT. A problem about "nine foreign journalists" from a Nordic Mathematical Olympiad is used as the starting point for a discussion of a class of extremal problems involving hypergraphs. Specifically, the problem is to find a sharp lower bound for the maximum degree of the hypergraph in terms of the number of (hyper) edges and their cardinalities.

1 Introduction

The following problem appeared on the First Nordic Mathematical Olympiad.

> Nine foreign journalists meet at a press conference. Each of them speaks at most three different languages, and any two of them can speak a common language. Show that at least five of them speak the same language.

It is natural to express this problem in the language of hypergraphs. We are given a hypergraph with nine pairwise intersecting hyperedges (or simply edges), where multiple edges are allowed. Each edge has cardinality at most three, and need to prove that some vertex is contained in at least five edges. In standard terminology, we are given an *intersecting hypergraph* with *rank* at most three, and are to prove that the *maximum degree* of the hypergraph is at least five. Generalizing this problem, we identify two functions of combinatorial interest.

JCMCC 22 (1996), pp. 183-192

Problem Let $f(r,d)$ denote the smallest integer such that every intersecting hypergraph $\mathcal{H}$ with $f(r,d)$ edges and rank $r(\mathcal{H}) \leq r$ has maximum degree $D(\mathcal{H}) \geq d$. In other words, $f(r,d)$ is the minimum number of pairwise intersecting edges having rank no more than r that necessarily yield a vertex of degree at least d. Let $f_0(r,d)$ be defined the same way, but with $\mathcal{H}$ restricted to be simple (no repeated edges). Find $f(r,d)$ and $f_0(r,d)$.

Thus, the problem from the Nordic Mathematical Olympiad asks for a proof that $f(3,5) \leq 9$. More on this later. First, we establish some general results concerning $f(r,d)$ and $f_0(r,d)$. In particular, we show that $f_0(r,d) = f(r-1,d)$ for $r \geq 5$ and d large enough, provided that a projective plane of order $r-2$ exists.

2 Upper and Lower Bounds

First note that

$$f_0(r+1,d) \geq f(r,d) \geq f_0(r,d) \tag{1}$$

The left hand inequality follows since, if $f(r,d)$ sets are available, we can add a distinct x_i to each E_i and obtain $f(r,d)$ distinct $(r+1)$-element sets.

In order to obtain bounds for f and f_0, we shall need the following hypergraph results.

Theorem 1. (Füredi [3]) *Suppose $\mathcal{H} = \{F_1, F_2, \ldots, F_m\}$ is an intersecting hypergraph with vertex set $V(\mathcal{H}) = X$ having rank $r \geq 3$. Then there exists a labeling $\phi\colon X \to R^+$ such that*

$$\sum_{x \in F_i} \phi(x) \geq 1 \text{ for all } i, \text{ and}$$

$$\sum_{x \in X} \phi(x) \leq r - 1 + \frac{1}{r}.$$

Moreover, if $\mathcal{H}$ does not contain a projective plane of order $r-1$, then $\sum_{x \in X} \phi(x) \leq r-1$ can be ensured.

Corollary 1. *If $\mathcal{H}$ is an intersecting hypergraph with rank $r(\mathcal{H}) \leq r$ and m edges, then*

$$D(\mathcal{H}) \geq \left\lceil \frac{m}{r-1+1/r} \right\rceil.$$

If in addition, $\mathcal{H}$ contains no projective plane of order $r-1$, then

$$D(\mathcal{H}) \geq \left\lceil \frac{m}{r-1} \right\rceil.$$

Proof: The bound on $D(\mathcal{H})$ follows from

$$m \leq \sum_{F_i \in \mathcal{H}} \sum_{x \in F_i} \phi(x) \leq D(\mathcal{H}) \cdot \left(r - 1 + \frac{1}{r}\right),$$

and the same argument (with $r - 1 + 1/r$ replaced by $r - 1$) takes care of the case in which $\mathcal{H}$ contains no projective plane of order $r - 1$. □

Theorem 2. (Erdös, Lovász [2]) *Suppose $\mathcal{H} = \{F_1, F_2, \ldots, F_m\}$ is a simple hypergraph satisfying $|F_i| \leq r$ for $1 \leq i \leq m$, and for every F_i and every $x \in F_i$ there is an F_j such that $F_i \cap F_j = \{x\}$. Then $m \leq r^r$.*

In fact, the upper bound r^r can be improved to c'_r where $c'_r/r^r \to 1 - e^{-1}$ as $r \to \infty$ [4], but we shall not need this stronger result in the proofs below (except that it leads to sharper versions of parts (v) and (vi) of the following theorem).

Theorem 3.

(i) *If $r - 1$ is a prime power and $d \equiv 1 \pmod r$, then*

$$f(r, d) \geq (d - 1)\left(r - 1 + \frac{1}{r}\right) + 1.$$

(ii) *If $r - 2$ is a prime power and $d \equiv 1 \pmod{r - 1}$, then*

$$f_0(r, d) \geq (d - 1)\left(r - 2 + \frac{1}{r - 1}\right) + 1.$$

(iii) $f(r, d) \leq \lceil d \cdot (r - 1 + r^{-1}) \rceil$.

(iv) *If a projective plane of order $r - 1$ does not exist, then $f(r, d) \leq d \cdot (r - 1)$.*

(v) *For some $c_r \leq r^r$, independent of d,*

$$f_0(r, d) \leq d \cdot \left(r - 2 + \frac{1}{r - 1}\right) + c_r.$$

(vi) *If a projective plane of order $r - 2$ does not exist, then $f_0(r, d) \leq d \cdot (r - 2) + c_r$.*

Proof:

(i) Take a projective plane of order $r - 1$, and let each line have multiplicity $(d - 1)/r$. Then all degrees are $d - 1$, and the number of r-sets is $(d - 1)(r^2 - r + 1)/r$, any two of them intersecting.

(ii) Use (i) and apply the left hand side of (1).

(iii) See Corollary 1.

(iv) See Corollary 1.

(v) Let $E_1, E_2, \ldots, E_m$ be m distinct intersecting r-sets. If there is an E_i and $x \in E_i$ such that $(E_i \setminus \{x\}) \cap E_j \neq \emptyset$, then replace E_i by $E_i \setminus \{x\}$ and repeat this procedure if possible. (Of course, the maximum degree does not increase.) Finally, we obtain a set system $F_1, \ldots, F_k$ with multiplicities $m_1, m_2, \ldots, m_k$ telling how many E_i were contracted to an F_j. Then $m_1 + \cdots + m_k = m$ and $m_j = 1$ for $|F_j| = r$. Using Theorem 2, we see that $k \leq r^r$. Deleting those F_j with $|F_j| = r$, we obtain an intersecting collection of $m - c_r$ sets of cardinalities $\leq r-1$. Thus the statements follow from (iii) and (iv).

(vi) same as (v)

□

Question. In fact, the preceding argument shows that

$$f_0(r,d) \leq f(r-1,d) + c_r.$$

Is is true that $f_0(r,d) = f(r-1,d)$, for $d \geq d_0(r)$?

Denote by $\nu(\mathcal{H})$ the *matching number* of $\mathcal{H}$ (the maximum number of pairwise disjoint sets in the set system). Define the functions $f(r,d,\nu)$ and $f_0(r,d,\nu)$ in the same way, replacing the intersecting condition by the requirement that the matching number is at most ν. Then

$$f(r-1,d,\nu) \leq f_0(r,d,\nu) \leq f(r-1,d,\nu) + c_{r,\nu}$$

for some constant $c_{r,\nu}$. Moreover, lower bounds are obtained from ν disjoint copies of the constructions given for f and f_0, and the upper bounds remain true if we replace d by νd and c_r by $c_{r,\nu}$.

Theorem 4. *Suppose that $f(r-1,d) \geq d \cdot (r-2+(r-1)^{-1}) + c$ for some constant $c = c(r)$ (i.e. a projective plane of order $r-2$ exists.) Then $f_0(r,d) = f(r-1,d)$ for $r \geq 5$ and $d \geq d_0(r)$.*

Proof: Consider the set system $F_1, F_2, \ldots, F_k$ obtained in the proof of (v) and (vi) of Theorem 3, and assume $|F_i| \leq r-1$ for $1 \leq i \leq$ ' and $|F_i| = r$ for $l+1 \leq i \leq k$. Note that $k - l \leq c_r \leq r^r$. If the sets $F_1, F_2, \ldots, F_l$ do not form a projective plane of order $r-2$, then a variant of Theorem 1 yields $l \leq (d-1)(r-2)$ (if the minimum degree is $\leq d-1$). Thus for $d > d_0$,

$$l + c_r < d\left(r - 2 + \frac{1}{r-1}\right) - c$$

would hold, contradicting our assumptions. Consequently, $F_1, F_2, \ldots, F_l$ is a projective plane of order $r-2$, and the proof will be finished if we prove $l = k$.

Otherwise, there is a set F_k that meets all lines of the plane but does not contain any line (because then F_k should be replaced by that line). Now $|F_k| = r$, and thus F_r meets $r(r-1)$ lines, counting multiplicity. (Each point has degree $r-1$ in the plane.) The number of lines is $(r-2)^2+(r-2)+1 = r^2-3r+3$, so it is enough to show that

$$\sum_{L_j}(|F_k \cap L_j| - 1) \geq 2r-2,$$

where the L_j are the lines of the projective plane. Set $H_j = F_k \cap L_j$. Each pair of elements of F_k is contained in precisely one H_j, and by our assumptions $|F_k| = r$ and $2 \leq |H_j| \leq r-2$. The Erdős-de Bruijn theorem [1] shows that we have at least r sets H_j. Put $t_j = |H_j|$. Then under the assumptions

$$\sum_{j=1}^{r'} \binom{t_j}{2} = \binom{r}{2}, \; 2 \leq t_j \leq r-2, \; r' \geq r,$$

the minimum of $\sum_{j=1}^{r'}(t_j - 1)$ is achieved when $r' = r$, $t_1 = t_2 = \cdots = t_{r-2} = 2$, $t_r = r-2$ and

$$t_{r-1} = \left\lceil \sqrt{2r - \frac{7}{4}} + \frac{1}{2} \right\rceil .$$

Thus, $\sum_{j=1}^{r'}(t_j - 1) \geq 2r-2$ for $r > 4$. □

3 Exact Values

If $r-1$ is a prime power and d is sufficiently large ($d \geq r(r-3)(r+1)$ will do), there is an intersecting hypergraph $\mathcal{H}$ of rank $\leq r$ with at least $(r-1)(d-1)+1$ edges and maximum degree $D(\mathcal{H}) \leq d-1$. Just take the edges to be the lines of a projective plane of order $r-1$, with each line repeated $\lfloor (d-1)/r \rfloor$ times. In fact, any intersecting hypergraph with these parameters must have the lines of a projective plane of order $r-1$ as its distinct edges, and this fact reduces the determination of $f(r,d)$ to a relatively straightforward computation.

Lemma 1. *Suppose that $r-1$ is a prime power. Let $\mathcal{H}$ be an intersecting hypergraph with at least $(r-1)(d-1)+1$ edges, rank at most r and maximum degree at most $d-1$. Then $\mathcal{H}$ is r-uniform and its distinct edges are lines of a projective plane of order $r-1$.*

Proof: If $\mathcal{H}$ contains an edge E with $|E| < r$, then some element of E has degree at least $\lceil |\mathcal{H}|/(r-1) \rceil \geq d$, a contradiction. Hence $\mathcal{H}$ must be r-uniform. If $\mathcal{H}$ contains no projective plane of order $r-1$, then by Corollary 1,

$$D(\mathcal{H}) \geq \left\lceil \frac{|H|}{r-1} \right\rceil \geq d,$$

a contradiction. Now it is easy to see that there is no set E that meets each of the lines of the projective plane but is not identical with any of them. Thus the distinct edges of $\mathcal{H}$ are the lines of the given projective plane. □

Theorem 5. *Suppose that $r-1$ is a prime power and d is appropriately large. Let* **A** *denote the incidence matrix of the projective plane of order $r-1$. Then*

$$f(r,d) = \max \mathbf{1}^T \mathbf{q} + 1,$$

where **1** *denotes the vector of all ones, and the maximum is taken over all vectors* **q** *with positive integer coordinates satisfying $\mathbf{A}\mathbf{q} \leq (d-1)\mathbf{1}$.*

Proof: Since d is appropriately large, there is an intersecting hypergraph $\mathcal{H}$ with at least $(r-1)(d-1)+1$ edges having rank $r(\mathcal{H}) \leq r$ and maximum degree $D(\mathcal{H}) \leq d-1$. Let $\mathcal{H}$ be such a hypergraph with $f(r,d)-1$ edges. By Lemma 1, $\mathcal{H}$ is r-uniform and its distinct edges are the lines of a projective plane of order $r-1$. Enumerate the lines of the projective plane $L_1, L_2, \ldots, L_n$ where $n = r^2 - r + 1$. To specify $\mathcal{H}$ it is only necessary to list the corresponding multiplicities $q_1, q_2, \ldots, q_n$. The fact that $D(\mathcal{H}) \leq d-1$ is expressed by $\mathbf{A}\mathbf{q} \leq (d-1)\mathbf{1}$, and since $\mathcal{H}$ is extremal, the number of edges, $q_1 + q_2 + \cdots + q_n = \mathbf{1}^T\mathbf{q}$, has the maximum possible value. □

Theorem 6. *For every $d \geq 5$,*

$$f(3,d) = \begin{cases} \frac{7d}{3} - 2 & \text{if } d \equiv 0 \pmod 3, \\ \frac{7(d-1)}{3} + 1 & \text{if } d \equiv 1 \pmod 3, \\ \frac{7(d-2)}{3} + 2 & \text{if } d \equiv 2 \pmod 3. \end{cases}$$

Proof: We take the incidence matrix of the projective plane of order two to be

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

First, we give the required constructions. Throughout this discussion, k is a positive integer.

(i) $d = 3k$. Set $q_1 = q_2 = q_6 = q_7 = k$ and $q_3 = q_4 = q_5 = k - 1$. Then $\mathbf{Aq} \leq (3k-1)\mathbf{1}$ and $\mathbf{1}^T\mathbf{q} = 7k - 3$.

(ii) $d = 3k + 1$. With $q_1 = q_2 = \cdots = q_7 = k$ we have $\mathbf{Aq} \leq 3k\mathbf{1}$ and $\mathbf{1}^T\mathbf{q} = 7k$.

(iii) $d = 3k + 2$. Set $q_1 = k + 1$ and $q_2 = q_3 = \cdots = q_7 = k$. Then $\mathbf{Aq} \leq (3k+1)\mathbf{1}$ and $\mathbf{1}^T\mathbf{q} = 7k + 1$.

Next, we prove that these are best possible. Suppose in each case that there is an intersecting hypergraph with rank $r(\mathcal{H}) \leq 3$, maximum degree $D(\mathcal{H}) \leq d - 1$ and at least $7d/3 - 2$ edges if $d \equiv 0 \pmod 3$ etc. Then $|\mathcal{H}| \geq 2d - 1$, the distinct edges of $\mathcal{H}$ are the lines of a projective plane, and $\mathcal{H}$ is completely described by the number of times each line occurs. By assumption, the multiplicity vector $\mathbf{q}$ satisfies $\mathbf{Aq} \leq (d-1)\mathbf{1}$.

(i) $d = 3k$. Since $\mathbf{1}^T\mathbf{q} \geq 7k - 2$, we may assume $q_1 \geq k$. Adding the inequalities involving q_1, we obtain

$$2k + (7k - 2) \leq 3q_1 + q_2 + q_3 + \cdots + q_7 \leq 3(3k - 1),$$

a contradiction.

(ii) $d = 3k+1$. By assumption, $\mathbf{1}^T\mathbf{q} \geq 7k+1$. Adding all of the constraint inequalities, we obtain

$$3(7k + 1) \leq 3(q_1 + q_2 + \cdots + q_7) \leq 7(3k),$$

a contradiction.

(iii) $d = 3k+2$. Since $\mathbf{1}^T\mathbf{q} \geq 7k+2$, we may assume $q_1 \geq k+1$. Adding the constraint inequalities involving q_1, we again obtain a contradiction. This time, it is

$$2(k + 1) + 7k + 2 \leq 3q_1 + q_2 + \cdots + q_7 \leq 3(3k + 1).$$

Thus the proposed formula for $f(3, n)$ holds in all cases. □

Oh yes, for $d = 5$ we have $f(3, 5) = 7(5 - 2)/3 + 2 = 9$. Nine journalists suffice!

4 As a Student Might Have Done It

We have used the "nine foreign journalists" problem from the Nordic Olympiad to motivate a discussion of $f(r, d)$ and $f_0(r, d)$, and point to interesting problems and results concerning these two functions. However, since we have used results due to de Bruijn, Erdös, Füredi and Lovász to obtain the solution, this may leave the reader wondering how a high school student from Sweden (for example) is supposed to solve the problem.

The solution given below is one which a student might have given. Our hypothetical student finds a proof along the lines of the one already given, but discovers the relevance of the projective plane of order two "from scratch" and needs nothing more than the pigeonhole principle, good instincts and a little patience to carry out the proof.

The student's proof. Assume that no language is spoken by five or more journalists. Then we have the following sequence of conclusions.

(i) *Each journalist must speak exactly three languages.*

Proof: Suppose one of the journalists speaks at most two languages. Then since each of the remaining eight journalists must speak a language in common with the one who is at most bilingual, the pigeonhole principle yields a language spoken by at least $\lceil 8/2 \rceil + 1 = 5$ journalists, a contradiction.

For $i = 1, 2, \ldots, 9$, let S_i denote the set of languages spoken by journalist i, and let $H_1, H_2, \ldots, H_m$ denote the distinct sets among the S_i. We shall refer to the H_i as *lines*.

(ii) *No two of the S_i share precisely two elements.*

Proof: Suppose that $\{1,2,3\}$ and $\{1,2,4\}$ are two of the S_i and let p, q, r, s denote the following cardinalities:

$$p = |\{i \mid 1 \in S_i, 2 \in S_i\}|, \quad q = |\{i \mid 1 \in S_i, 2 \notin S_i\}|,$$
$$r = |\{i \mid 1 \notin S_i, 2 \in S_i\}|, \quad s = |\{i \mid 1 \notin S_i, 2 \notin S_i\}|.$$

Then $p \geq 2$ and $p + q + r + s = 9$. Since no language is spoken by five or more journalists, $p + q \leq 4$ and $p + r \leq 4$. Suppose $\{1, 2, \ell\}$ is one of the S_i where $\ell \neq 3, 4$. Then either $p = 3$ and $q, r \leq 1$ or $p = 4$ and $q = r = 0$; in each case $p + q + r \leq 5$ so $s \geq 4$. It follows that $\{1, 2, \ell\}$ and at least four other S_i contain ℓ, a contradiction. Hence each of the p sets containing both 1 and 2 must contain either 3 or 4, and each of the s sets containing neither 1 nor 2 must contain both 3 and 4 since each such set must intersect both $\{1,2,3\}$ and $\{1,2,4\}$. We thus conclude that either 3 or 4 must be contained in at least $s + p/2$ of the S_i and thus $s + p/2 \leq 4$. Adding $p + q \leq 4$, $p + r \leq 4$ and $s + p/2 \leq 4$, and using $p + q + r + s = 9$, we obtain $3p/2 \leq 3$ and thus $p = 2$. Now $q \leq 2$ and $r \leq 2$ so $s \geq 3$. But $s > 3$ is impossible since then $\{1,2,3\}$ and four other sets would contain 3. Hence we are left with $p = q = r = 2$ and $s = 3$ and we may assume

$$S_1 = \{1,2,3\}, \quad S_2 = \{1,2,4\}, \quad S_3 = \{1,-,-\},$$
$$S_4 = \{1,-,-\}, \; S_5 = \{2,-,-\}, \; S_6 = \{2,-,-\},$$
$$S_7 = \{3,4,-\}, \quad S_8 = \{3,4,-\}, \quad S_9 = \{3,4,-\}.$$

Without loss of generality, $S_7 = \{3,4,5\}$. Since each of S_3, S_4, S_5, S_6 intersects S_7, one of them intersects $\{3,4\}$ or else each contains 5. In either case, there is a language spoken by five journalists, a contradiction.

Thus we have shown that any two distinct lines H_i, H_j have exactly one common element.

(iii) *Each language belongs to precisely three lines.*

Proof: Suppose language 1 belongs to one, two, or more than three lines. If $\{1,2,3\}$ is the only line containing 1, then each of the nine journalists speaks either language 2 or language 3 so there is a language spoken by five journalists, a contradiction. If $\{1,2,3\}$ and $\{1,4,5\}$ are the only two lines containing 1, and these occur with multiplicity p and q, respectively, among $S_1, S_2, \ldots, S_9$, then $p+q \leq 4$, either language 2 or language 3 is spoken by at least $p+(9-(p+q))/2$ journalists, and either language 4 or language 5 is spoken by at least $q+(9-(p+q))/2$ journalists. Thus language 2, 3, 4 or 5 is spoken by at least $\lceil 9/2 \rceil = 5$ journalists, a contradiction. Clearly, 1 doesn't belong to four of the lines since these four sets must be otherwise pairwise disjoint, and an S_j not containing 1 cannot intersect each of them.

(iv) *Among the nine sets of languages spoken by the journalists, there are seven distinct ones, and these may be taken to be*

$$H_1 = \{1,2,3\},\ H_2 = \{1,4,5\},\ H_3 = \{1,6,7\},\ H_4 = \{2,4,6\}, \tag{2}$$

$$H_5 = \{2,5,7\},\ H_6 = \{3,4,7\},\ H_7 = \{3,5,6\}. \tag{3}$$

Proof: Without loss of generality, we may assume that the lines containing 1 are

$$\{1,2,3\},\ \{1,4,5\},\ \{1,6,7\},$$

and the two additional lines containing 2 are

$$\{2,4,6\},\ \{2,5,7\}.$$

Since the pairs $\{1,7\}$, $\{2,7\}$, $\{5,7\}$ and $\{6,7\}$ occur in the above lines and any line must intersect $\{1,2,3\}$ and $\{1,4,5\}$, the unique choice for the third line containing 7 is $\{3,4,7\}$. Similarly, the unique choice for the third line containing 5 is $\{3,5,6\}$. Clearly, there can be no more lines.

Completion of the proof: Each of the seven lines occur as language sets of the journalists, and no other language set is possible. Since there are nine language sets S_i and only seven lines H_j, some line must occur at least twice, and we may assume that the language sets of our nine journalists are

$$\{1,2,3\},\ \{1,2,3\},\ \{1,4,5\},\ \{1,6,7\},\ \{2,4,6\},$$
$$\{2,5,7\},\ \{3,4,7\},\ \{3,5,6\},\ \{a,b,c\},$$

where $\{a,b,c\}$ is one of the seven lines in (2)–(3) so $\{a,b,c\} \cap \{1,2,3\} \neq \emptyset$. This implies one of languages 1, 2, 3 is spoken by five journalists, a contradiction. □

References

[1] N.G. de Bruijn and P. Erdös, On a combinatorial problem, *Nederrl. Akad. Wetensch. Indag. Math.*, **51** (1948), 1277–1279.

[2] P. Erdös and L. Lovász, Problems and results on 3-chromatic hypergaphs and some related questions, in *Infinite and Finite Sets*, Proc. Colloq Math. Soc. Janos Bolyai, Vol. 10, 1979, North-Holland, 609–627.

[3] Z. Füredi, Maximum degree and fractional matchings in uniform hypergraphs, *Combinatorica*, 1 (1981), 155–162.

[4] Zs. Tuza, Inequalities for minimal covering sets in set systems of given rank, *Discrete Applied Math.*, **51** (1994), 187–195.