The distance of $\mathcal{F}$-free families

Balázs Patkós

Department of Mathematics and its Applications
Central European University, Budapest, Hungary
patkos@renyi.hu

Abstract

If $\mathcal{F}$ is a fixed hypergraph, then for two $\mathcal{F}$-free hypergraphs $\mathcal{H}_1 = (V, E_1)$ and $\mathcal{H}_2 = (V, E_2)$ we define their $\mathcal{F}$-free distance by the number of subhypergraphs of $\mathcal{H}_1 \cup \mathcal{H}_2 = (V, E_1 \cup E_2)$ which are isomorphic to $\mathcal{F}$, and we denote this number by $D_{\mathcal{F}}(\mathcal{H}_1, \mathcal{H}_2)$. For a collection $\mathcal{C}$ of hypergraphs the $\mathcal{C}$-free distance of two $\mathcal{C}$-free hypergraphs (that is $\mathcal{F}$-free for all $\mathcal{F} \in \mathcal{C}$) is $D_{\mathcal{C}}(\mathcal{H}_1, \mathcal{H}_2) = \sum_{\mathcal{F} \in \mathcal{C}} D_{\mathcal{F}}(\mathcal{H}_1, \mathcal{H}_2)$. In this paper we will obtain exact results on the maximum distance of K_r-free graphs and Sperner systems and prove upper and lower bounds on the maximum distance of trees.

1 Introduction

Notation. Throughout the paper $[n]$ denotes the set of the first n positive integers, and for any set S, the set of all k-element subsets of S is denoted by $\binom{S}{k}$.

In [3] it is proved that if $\mathcal{F}_1, \mathcal{F}_2$ are two intersecting families of subsets of an n-element underlying set, then the number of disjoint pairs of sets in the union of the two families is at most 3^{n-2}, and this is sharp as is shown by the families $\mathcal{F}_i = \{F \subset [n] : i \in F\}, \mathcal{F}_j = \{F \subset [n] : j \in F\}$ for any distinct i, j. Since any disjoint pair in the union of two intersecting families shows that the union is not an intersecting family, the number of disjoint pairs can be considered as a distance-like concept.

How can we generalize this concept for other hypergraph properties? A heuristic approach is (having two hypergraphs $\mathcal{H}_1 = (V, E_1), \mathcal{H}_2 = (V, E_2)$ both satisfying a property P) to count the subhypergraphs of $\mathcal{H}_1 \cup \mathcal{H}_2$ "showing" that $\mathcal{H}_1 \cup \mathcal{H}_2 = (V, E_1 \cup E_2)$ does not have property P (and denote their number by $D_P(\mathcal{H}_1, \mathcal{H}_2)$). Though generally it is not clear how a family of hyperedges can show that the whole hypergraph does not satisfy a property, in concrete examples it is obvious what we should count.

Examples. Let P be the property of inclusion. That is for any pair of hyperedges H_1, H_2 either $H_1 \subseteq H_2$ or $H_2 \subseteq H_1$. Therefore for two hypergraphs of sets with property

P we should count the pairs of hyperedges H_1, H_2 in the union for which both $H_1 \not\subseteq H_2$ and $H_2 \not\subseteq H_1$ hold. Maximal hypergraphs with property P are saturated chains. (A chain $\mathcal{C} = \{C_0, C_1, ..., C_n\}$ is saturated if for all j, $|C_j| = j$ holds.) The empty set is a subset of each set, and each set is a subset of the whole underlying set $[n]$, so the maximum number of pairs of hyperedges none of them containing the other, where the edges are taken from two chains $\mathcal{C}_1, \mathcal{C}_2$ (this is $D_{\not\subseteq}(\mathcal{C}_1, \mathcal{C}_2)$), is at most $(n-1)^2$. And for any pair of saturated chains of the form $\mathcal{C} = \{C_0 \subseteq C_1 \subseteq C_2 \subseteq ... \subseteq C_{n-1} \subseteq C_n\}$ and $\mathcal{C}' = \{C_n^c \subseteq C_{n-1}^c \subseteq ... \subseteq C_1^c \subseteq C_0^c\}$ where A^c denotes the complement of the set A, we have $D_{\not\subseteq}(\mathcal{C}, \mathcal{C}') = (n-1)^2$.

As another example, let P be the property of (pairwise) disjointness (i.e. for every $F_1, F_2 \in \mathcal{F}$ $F_1 \cap F_2 = \emptyset$). This time maximal hypergraphs are partitions, and $D_{\not\cap}(\mathcal{F}_1, \mathcal{F}_2) = |\{(F_1, F_2) : F_i \in \mathcal{F}_i \text{ and } F_1 \cap F_2 \neq \emptyset\}$. If for every intersecting pair of edges in the union of two partitions we point out an element of the intersection, we get an injective mapping from the non-disjoint pairs to the base set. So the number of such pairs can be at most n. For any partition $\mathcal{F}$ we can create another partition $\mathcal{F}'$ by choosing an element from each hyperedge to form a hyperedge in $\mathcal{F}'$, then again choosing one element from all remaining non-empty edges, and so on to have $D_{\not\cap}(\mathcal{F}, \mathcal{F}') = n$.

To make things precise we need the following definitions.

Definition: If $\mathcal{F}$ and $\mathcal{G}$ are hypergraphs, then $\mathcal{G}$ is *$\mathcal{F}$-free* if there is no subhypergraph of $\mathcal{G}$ which is isomorphic to $\mathcal{F}$. If $\mathcal{H}_1$ and $\mathcal{H}_2$ are two $\mathcal{F}$-free hypergraphs, then their *$\mathcal{F}$-free distance* $D_{\mathcal{F}}(\mathcal{H}_1, \mathcal{H}_2)$ is the number of subhypergraphs of $\mathcal{H}_1 \cup \mathcal{H}_2$ which are isomorphic to $\mathcal{F}$. If $\mathcal{C}$ is a collection of hypergraphs and $\mathcal{H}_1$ and $\mathcal{H}_2$ are two $\mathcal{C}$-free (i.e. $\mathcal{F}$-free for all $\mathcal{F} \in \mathcal{C}$) hypergraphs, then their *$\mathcal{C}$-free distance* is $D_{\mathcal{C}}(\mathcal{H}_1, \mathcal{H}_2) = \sum_{\mathcal{F} \in \mathcal{C}} D_{\mathcal{F}}(\mathcal{H}_1, \mathcal{H}_2)$.

Note that these definitions cover all properties we have already treated.

2 Sperner Systems

Definition: A hypergraph $\mathcal{F}$ is said to be a *Sperner system* if $F_1 \not\subseteq F_2$ for any distinct $F_1, F_2 \in \mathcal{F}$. So the distance of two Sperner systems is $D_{\subseteq}(\mathcal{F}, \mathcal{G}) = |\{\{A_1, A_2\} : A_i \in \mathcal{F} \cup \mathcal{G} \text{ and } A_1 \subseteq A_2\}|$.

For the precise definition: $\mathcal{C}_{\subseteq} = \{\mathcal{G}_{k,l} : k \in \mathbb{N}, l \in \mathbb{N}\}$ where $\mathcal{G}_{k,l} = \{\{1, ..., k\}, \{1, ..., k+l\}\}$ is the collection of forbidden hypergraphs, and $D_{\subseteq}(\mathcal{F}, \mathcal{G}) = D_{\mathcal{C}_{\subseteq}}(\mathcal{F}, \mathcal{G})$.

Theorem 2.1 *If $\mathcal{F}, \mathcal{G} \subseteq 2^{[n]}$ are two Sperner systems, then*

$$D_{\subseteq}(\mathcal{F}, \mathcal{G}) \leq D_{\subseteq}(\mathcal{F}_0, \mathcal{G}_0)$$

where $\mathcal{F}_0$ is the hypergraph with edge set $\binom{[n]}{k_1}$, $\mathcal{G}_0$ is the hypergraph with edge set $\binom{[n]}{k_2}$ $(k_1 < k_2)$ such that each of $k_1, k_2 - k_1, n - k_2$ differ by at most one. In particular, if 3 *divides n, then $\mathcal{F}_0$ is the hypergraph with edge set $\binom{[n]}{n/3}$ and $\mathcal{G}_0$ is hypergraph with edge set $\binom{[n]}{2n/3}$.*

First proof: W.l.o.g. we can assume that both $\mathcal{F}$ and $\mathcal{G}$ are maximal Sperner systems, since adding new edges to the hypergraphs cannot decrease the distance.

Our goal is to show that by starting with any pair of Sperner systems $(\mathcal{F}, \mathcal{G})$, in finitely many steps $(\mathcal{F}^i, \mathcal{G}^i)$ we can reach $(\mathcal{F}_0, \mathcal{G}_0) = (\mathcal{F}^m, \mathcal{G}^m)$ such that

$$D_{\subseteq}(\mathcal{F}, \mathcal{G}) \leq D_{\subseteq}(\mathcal{F}^1, \mathcal{G}^1) \leq D_{\subseteq}(\mathcal{F}^2, \mathcal{G}^2) \leq ... \leq D_{\subseteq}(\mathcal{F}^m, \mathcal{G}^m) = D_{\subseteq}(\mathcal{F}_0, \mathcal{G}_0).$$

Step 1

Let $\mathcal{C} = \mathcal{F} \cap \mathcal{G}$ and partition $\mathcal{F}$ and $\mathcal{G}$ by

$$\mathcal{F} = \mathcal{C} \cup \mathcal{F}_1 \cup \mathcal{F}_2 \text{and} \quad \mathcal{G} = \mathcal{C} \cup \mathcal{G}_1 \cup \mathcal{G}_2$$

where $\mathcal{F}_1 = \{F \in \mathcal{F} : \text{there is } G \in \mathcal{G}\ G \subsetneq F\}$, $\mathcal{F}_2 = \{F \in \mathcal{F} : \text{there is } G \in \mathcal{G}\ G \supsetneq F\}$ and $\mathcal{G}_1, \mathcal{G}_2$ defined similarly. Note, that any $F \in \mathcal{F}$ contains or is contained in some $G \in \mathcal{G}$, because otherwise we could add it to $\mathcal{G}$, which would contradict the maximal property of $\mathcal{G}$, and no $F \in \mathcal{F}$ belongs to both $\mathcal{F}_1, \mathcal{F}_2$, otherwise there exist $G_1, G_2 \in \mathcal{G}$ such that $G_1 \subsetneq F \subsetneq G_2$ contradicting the Sperner property of $\mathcal{G}$. So $\mathcal{C}, \mathcal{F}_1, \mathcal{F}_2$ is really a partition of $\mathcal{F}$.

Now let $\mathcal{F}^1 = \mathcal{C} \cup \mathcal{F}_1 \cup \mathcal{G}_1$ and $\mathcal{G}^1 = \mathcal{C} \cup \mathcal{F}_2 \cup \mathcal{G}_2$. It is easy to check that both $\mathcal{F}^1$ and $\mathcal{G}^1$ are Sperner systems. The fact that $D_{\subseteq}(\mathcal{F}, \mathcal{G}) = D_{\subseteq}(\mathcal{F}^1, \mathcal{G}^1)$ follows from the fact that $E(\mathcal{F}) \cup E(\mathcal{G}) = E(\mathcal{F}^1) \cup E(\mathcal{G}^1)$.

By the above change of the systems there is no $F \in \mathcal{F}^1$ for which there exists a $G \in \mathcal{G}^1$ with $F \subseteq G$, so we can refer to $\mathcal{F}^1$ as the upper Sperner system, and to $\mathcal{G}^1$ as the lower system.

From now on in any even step we replace some of the edges of the upper Sperner system by other edges of larger size, and in any odd step we do the same to some edges of the lower system.

Step 2

Partition $\mathcal{F}^1$ into two subsystems: $\mathcal{F}_1^1 = \{F \in \mathcal{F}^1 : |F| > n/2\}, \mathcal{F}_2^1 = \{F \in \mathcal{F}^1 : |F| \leq n/2\}$. Put $\mathcal{F}^2 = \mathcal{F}_1^1 \cup \{F \in \binom{[n]}{\lceil n/2 \rceil} : \text{ there is } F' \in \mathcal{F}_2^1 \text{ such that } F' \subseteq F\}$ and $\mathcal{G}^2 = \mathcal{G}^1$. It is clear that $\mathcal{F}^2$ is a Sperner system.

$D_{\subseteq}(\mathcal{F}^1, \mathcal{G}^1) \leq D_{\subseteq}(\mathcal{F}^2, \mathcal{G}^2)$ follows from Sperner's lemma [5] stating, that if $\mathcal{G}$ is a k-uniform hypergraph with $k \leq \frac{n}{2}$, then $|\nabla \mathcal{G}| \geq |\mathcal{G}|$, where $\nabla \mathcal{G} = \{G' \subset [n] : |G'| = k+1 \text{ and there is } G \in \mathcal{G} \text{ such that } G \subseteq G'\}$.

Step 3

Now we want to push the lower system up, so we replace all the small sets.

$$\mathcal{G}_2^2 = \{G \in \mathcal{G}^2 : |G| < \lceil \frac{\lceil \frac{n}{2} \rceil}{2} \rceil\}; \ \ \mathcal{G}_1^2 = \mathcal{G}^2 \setminus \mathcal{G}_2^2$$

$$\mathcal{G}^3 = \mathcal{G}_1^2 \cup \{G \in \binom{[n]}{\lceil \frac{\lceil \frac{n}{2} \rceil}{2} \rceil} : \text{there is } G' \in \mathcal{G}_2^2 \ \text{ with } G' \subseteq G\}; \ \mathcal{F}^3 = \mathcal{F}^2$$

Just as in the argument in Step 2 $\mathcal{G}^3$ is a Sperner system, and using the original proof of Sperner's theorem one can verify that for any fixed $F \in \mathcal{F}^3 = \mathcal{F}^2$ the number of edges in $\mathcal{G}^3$ contained by F is at least the number of edges in $\mathcal{G}^2$ contained by F.

Suppose we achieved in Step $2k$ that the hyperedges in the upper system have size at least $c_k n$, and in Step $2k+1$ that all the edges in the lower system have size at least $d_k n$. Then in Step $2(k+1)$ we will show that all edges in the upper system have size at least $c_{k+1}n = d_k n + \lceil \frac{1-d_k}{2} n \rceil$, and in Step $2(k+1)+1$ that the sets of the lower system have size at least $d_{k+1}n = \lceil \frac{c_{k+1}n}{2} \rceil$. Formally Step $2(k+1)$ and Step $2(k+1)+1$ are defined as follows:

Step $2(k+1)$

Let $\mathcal{F}^{2k+1} = \mathcal{F}_1^{2k+1} \cup \mathcal{F}_2^{2k+1}$, where $\mathcal{F}_1^{2k+1} = \{F \in \mathcal{F}^{2k+1} : |F| > d_k n + \lceil \frac{1-d_k}{2} n \rceil\}$ and $\mathcal{F}_2^{2k+1} = \mathcal{F}^{2k+1} \setminus \mathcal{F}_1^{2k+1}$. Then let

$$\mathcal{F}^{2(k+1)} = \mathcal{F}_1^{2k+1} \cup \{F' \in \binom{[n]}{d_k n + \lceil \frac{1-d_k}{2} n \rceil} : \exists F \in \mathcal{F}^{2k+1} \ \text{ such that } F \subseteq F'\}$$

and let

$$\mathcal{G}^{2(k+1)} = \mathcal{G}^{2k+1}.$$

Step $2(k+1)+1$

Let us partition $\mathcal{G}^{2(k+1)}$ into two subhypergraphs: $\mathcal{G}_1^{2(k+1)} = \{G \in \mathcal{G}^{2(k+1)} : |G| > \lceil \frac{c_{k+1}n}{2} \rceil\}$ and $\mathcal{G}_2^{2(k+1)} = \mathcal{G}^{2(k+1)} \setminus \mathcal{G}_1^{2(k+1)}$. Then put

$$\mathcal{G}^{2(k+1)+1} = \mathcal{G}_1^{2(k+1)} \cup \{G' \in \binom{[n]}{\lceil \frac{c_{k+1}n}{2} \rceil} : \exists G \in \mathcal{G}^{2(k+1)} \ \text{ such that } G \subseteq G'\}$$

and

$$\mathcal{F}^{2(k+1)+1} = \mathcal{F}^{2(k+1)}.$$

The fact that during Step $2(k+1)$ and Step $2(k+1)+1$ the distance of our systems cannot decrease follows just as in the case of Step 2 and Step 3. (Note that in Step $2(k+1)$ we apply Sperner's lemma to the posets $\mathbb{P}_G = \{H \setminus G : H \supseteq G\}$, where G ranges over the edges in $\mathcal{G}^{2k+1}$, while in Step $2(k+1)+1)$ to the posets $\mathbb{P}_F = \{H : H \subseteq F\}$, where F ranges over the edges in $\mathcal{F}^{2(k+1)}$ and this is the reason why we c_k and d_k have to be defined in the way we did.) The statement about c_{k+1} and d_{k+1} is true by definition.

So (forgetting the ceiling signs for a moment) $c_{k+1} = \frac{1}{2}c_k + \frac{1-\frac{1}{2}c_k}{2} = 1/2 + c_k/4$ (and $d_{k+1} = c_{k+1}/2$). As for any $x \in [0; 2/3)$ $x < 1/2 + x/4$, in finitely many steps (by virtue of the ceiling sign) we can achieve that all the edges in the upper system have size at least $\lceil 2n/3 \rceil$, and all the edges in the lower system have size at least $\lceil n/3 \rceil$.

To finish the proof we need the observation that the complement system of a Sperner system is a Sperner system, and that (denoting the complement system of $\mathcal{F}$ by $\mathcal{F}^c = \{[n] \setminus F : F \in \mathcal{F}\}$), we have $D_{\subseteq}(\mathcal{F}, \mathcal{G}) = D_{\subseteq}(\mathcal{F}^c, \mathcal{G}^c)$.

In the complement systems of the above pair, all edges have size at most $\lfloor n/3 \rfloor$ or $\lfloor 2n/3 \rfloor$, and after the same pushing up procedure we get one of the optimal pairs. □

Second proof: By Step 1 of the previous proof we reduce the problem to Sperner systems $\mathcal{F}, \mathcal{G}$ where for all $F \in \mathcal{F}$ there is a $G \in \mathcal{G}$ with $F \subseteq G$. Then we are done by the following theorem of Katona.

Theorem 2.2 [2] *(Iterated Sperner theorem) Let $A_1, ..., A_m$ be subsets of an n element set satisfying $A_j \not\subseteq A_k$ $1 \le j, k \le m, j \ne k$. For each $i = 1, ..., m$, suppose $B_{i,1}, ..., B_{i,m_i}$ are subsets of A_i satisfying $B_{i,j} \not\subseteq B_{i,k}$ $1 \le j, k \le m_i$. Then*

$$\sum_{i=1}^{m} m_i \le \binom{n}{\lfloor \frac{2n}{3} \rfloor} \binom{\lfloor \frac{2n}{3} \rfloor}{\lfloor \frac{n}{3} \rfloor}.$$

□

Remark: Theorem 2.2 (besides Step 1) is stronger than Theorem 2.1 (since in Theorem 2.2 we do not require that the Bs form a Sperner system), but Katona's proof of Theorem 2.2 uses a generalization of the LYM-inequality, while our first proof uses only Sperner's original idea of his well-known theorem.

3 K_r-free Graphs

We denote by K_r the complete graph on r vertices. A graph is K_r-free, if it does not contain any copy of K_r. The K_r-free distance of two K_r-free graphs (G_1, G_2) on the same underlying set V is

$$D_{K_r}(G_1, G_2) = |\{\{x_1, ..., x_r\} : x_i \in V \text{ for all } i, \text{ and any } (x_i, x_j) \in E(G_1 \cup G_2)\}|$$

In all the cases we have already treated (intersecting, pairwise disjoint and Sperner families, chains), the structure of the hypergraphs in the optimal pair (the pair with the maximum distance) was very similar to that of the optimal family in the original problem (what is the largest family with the desired property). Therefore it is quite natural to conjecture that Turán graphs will come into sight. (Turán's well-known theorem [6] says that a K_r-free graph on n vertices with the most possible number of edges must be isomorphic to the complete $r-1$-partite graph, where the sizes of any two partition classes may differ by at most one. These graphs are called the Turán graphs.)

Though it is not true that if $D_{K_r}(G_1, G_2)$ is maximal, then both G_1, G_2 should be Turán-graphs, still Turán graphs will play an important role in the proof of the next theorem. First we need to introduce some notation.

$T(n, r)$ is the usual notation for the r-partite Turán graph on n vertices and $t(n, r)$ denotes the number of edges in the graph. Now let $k_s(G)$ denote the number of s-cliques in G. (So $t(n, r) = k_2(T(n, r))$.)

The Ramsey number $R(k)$ denotes the least integer n for which any E_0, E_1 partition of the edges of K_n, there is a sample of K_k either in E_0 or in E_1.

Let us write furthermore $D_r^n := \max\{D_{K_r}(G_1, G_2) : G_1, G_2$ are K_r-free on the same vertex set $[n]\}$ and put $m = R(r) - 1$.

Theorem 3.1 $D_r^n = k_r(T(n, m))$

Proof: For the $\geq$ part we need a construction.

Let us fix a partition E_0, E_1 of the edges of K_m such that there is no K_r neither in E_0 nor in E_1. We want to define G_0, G_1 two K_r-free graphs on $[1, .., n]$. So we have to decide which edges we want to put into G_0 and which into G_1. To do this, for any $1 \leq i < j \leq n$ write $i = l_i m + i'$, $j = l_j m + j'$ where $1 \leq i', j' \leq m$.

Now put (i, j) into $E(G_0)$ iff $(i', j') \in E_0$, and into $E(G_1)$ iff $(i', j') \in E_1$. Since (i, j) is an edge if and only if $i \neq j \mod m$, therefore $G_0 \cup G_1$ is just $T(n, m)$ and the classes are just the congruency classes modulo m. We have to check that G_0, G_1 are both K_r-free. If not, then $i_1, i_2, .., i_r$ form a K_r in, say, G_0. But then $i'_1, i'_2, ..., i'_r$ should be all distinct, and should form a K_r in E_0 - a contradiction.

For the $\leq$ part of the proof, note that $G_0 \cup G_1$ cannot contain a $K_{R(r)}$ as otherwise G_0 or G_1 would contain a K_r. So the following result of Sauer (its $s = 2$ case is exactly Turán's theorem) completes the proof.

Lemma 3.2 (Sauer [4] see also [1]) *If $s < p$ and G is a K_p-free graph on n vertices, then the number of K_ss in G is at most $k_s(T(n, p-1))$.* □

Remark: If m divides n, then $k_r(T(n, m)) = \binom{m}{r}(\frac{n}{m})^r$, so the problem of giving the exact value of D_r^n for large enough n is equivalent to giving the exact value of $R(r)$.

4 Trees

Trees are circle-free graphs, so $C_{\text{cycle}} = \{C_k : k \geq 3\}$ where C_k is the cycle of length k. Therefore this time the question is, how many cycles we can have in the union of two trees on the same n-element vertex set. $D_{\text{cycle}}(T_1^n, T_2^n)$ will denote the tree-distance (the number of cycles in the union) of two trees T_1^n and T_2^n. D_{cycle}^n will denote the maximum tree-distance of two trees on the same n vertices.

A trivial upper bound on D_{cycle}^n is 4^{n-1}, since the union of two trees may contain at most $2(n-1)$ edges, so the number of subsets of the edge set of the union is clearly an upper bound for the number of cycles.

The following recursive construction (fig.1) shows that D_{cycle}^n does have an exponential growth. Suppose we have T_1^n, T_2^n two trees on n vertices, and an edge e (with endpoints x and y) in their union, through which there are c_n cycles in $T_1^n \cup T_2^n$. Likewise suppose we have T_1^m, T_2^m two trees on m vertices (with vertex set disjoint from that of T_1^n and T_2^n), and an edge f (with endpoints u and v) in their union, through which there are c_m cycles in $T_1^m \cup T_2^m$. Let $T_1^{n+m} = T_1^n \cup T_1^m \cup \{xu\}$ and $T_2^{n+m} = T_2^n \cup T_2^m \cup \{yv\}$. We claim that in $T_1^{n+m} \cup T_2^{n+m}$ there are $(c_n+1)(c_m+1)$ cycles through the edge xu. Indeed, there are c_n paths from x to y in $T_1^n \cup T_2^n$ plus the edge $xy = e$, then we have to go through the edge yv, then choose among the c_m paths in $T_1^m \cup T_2^m$ from v to u (or the edge $vu = f$), and then we finish off the cycle using the edge xu. Since we just took the sum of the number of vertices, and the number of cycles got multiplied, this is really of exponential growth. To be more concrete: we can cover K_4 by two (edge-disjoint) paths, so we have $D^4_{cycle} = 7$. By the above recursive bound we get that $D^{4n}_{cycle} \geq 7^n = (7^{1/4})^{4n}$. ($7^{1/4} = 1.625...$)

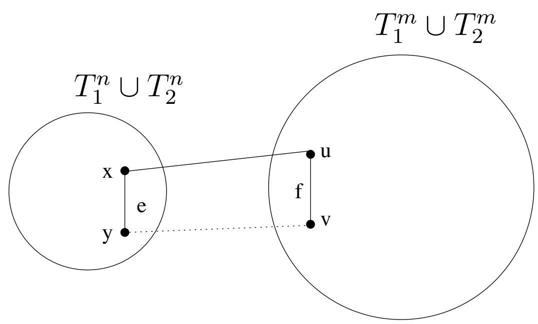

Figure 1: The recursive construction showing the exponential growth of D^n_{tree}

In the next two subsections we prove the following lower and upper bounds on D^n_{cycle}:

Theorem 4.1 *There exists a constant* c *for which the following inequalities hold*

$$cx_0^n \leq D^n_{cycle} \leq \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2i} = 2^{n-1} - 1,$$

where x_0 *is the unique real root of the equation* $x^3 - x^2 - x - 1 = 0$ *(*$x_0 = 1.8392...$*).*

4.1 Lower Bound on D^n_{cycle}

In this subsection we will give a "real" construction for the lower bound on D^n_{cycle} (see fig.2). Both of the trees in the construction are paths, and we will refer to them as the blue tree (denoted by B_n) and the red tree (denoted by R_n). The vertices of the trees are the integers from $-k$ up to k if $n = 2k+1$ and the integers from $-k$ to $k-1$ if $n = 2k$. Two integers are adjacent in B_n if and only if they are consecutive. If $n = 2k+1$, then the edge set of R_n is $\{(-l, l) : 1 \leq l \leq k\} \cup \{(l, -(l+1)) : 1 \leq l < k\} \cup \{(k, 0)\}$. If $n = 2k$ is even, one just drops the vertex k and the edges incident to it, and add the edge $(-k, 0)$ to the red tree.

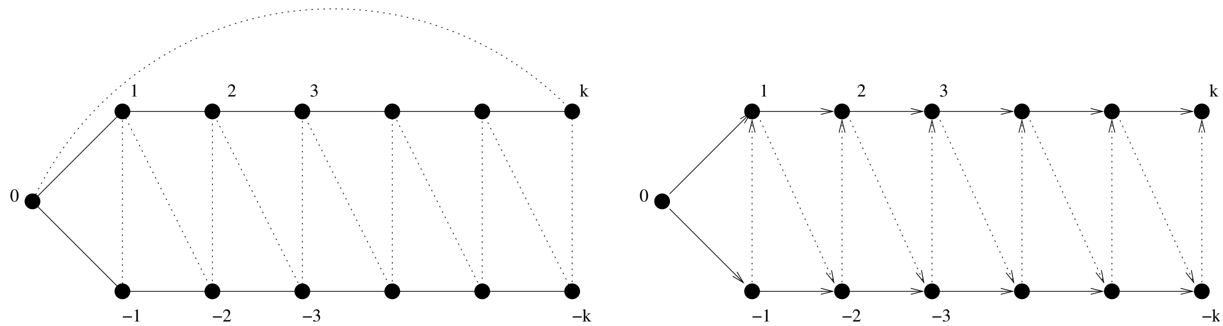

Figure 2: The “real construction” $R_n \cup B_n$ and the orientation of its edges

Let $c(n)$ denote the number of cycles through the edge $(k, 0)$ (that is the number of paths from the vertex 0 to the vertex k) if $n = 2k + 1$ and the number of cycles through $(-k, 0)$ if $n = 2k$. We claim that the following recurrence holds: $c(n) = c(n-1) + c(n-2) + c(n-3)$ where $c(1) = c(2) = 1,\ c(3) = 2$.

To see this let us consider the graph $B_n \cup R_n \setminus \{(0, k)\}$ ($B_n \cup R_n \setminus \{(0, -k)\}$ if $n = 2k$) as a directed graph with the following orientation of the edges (fig.2): all edges are directed from the vertex of smaller absolute value to the vertex of bigger absolute value. The red edges of type $(-l, l)$ are directed from the vertex $-l$ toward the vertex l. In a path $0 = x_0, x_1, ..., x_{l-1}, x_l = k$ the edge (x_j, x_{j+1}) is a *backward edge* if x_j is the endpoint and x_{j+1} is the starting point of the edge in the above orientation. Other edges will be called *forward edges*.

First note, that there can be no blue backward edges in a path from 0 to k. Since if there was, let us take the “rightmost” one (x_j, x_{j+1}) (i.e. the one with an endpoint of greatest absolute value). Assume $x_j = -(l+1), x_{j+1} = -l$ (the case $x_j = l+1, x_{j+1} = l$ is similar). Then because this is the rightmost backward blue edge in the path, x_{j-1} cannot be $-(l+2)$ ($(-(l+2), -(l+1))$ would be a backward blue edge "further to the right"). Therefore x_{j-1} is either $l+1$ or l. In both cases the vertex $x_{j+1} = -l$ is cut from the vertex k by the edge (x_{j-1}, x_j), so we cannot finish the path in this way.

How about backward red edges? (fig.3) If $x_j = -(l+1)$, $x_{j+1} = l$, then x_{j-1} cannot be $-(l+2)$ as $(-(l+2), -(l+1))$ would be a backward blue edge. x_{j-1} cannot be $l+1$ either for the edge (x_{j-1}, x_j) would cut x_{j+1} from the vertex k. So x_{j-1} must be $-l$. x_{j+2} cannot be $l-1$ (backward blue edge), so x_{j+2} is $l+1$. In the same manner one can see, if (x_j, x_{j+1}) is a backward red edge with $x_j = l$, $x_{j+1} = -l$, then x_{j-1} should be $l-1$ and x_{j+2} should be $-(l+1)$. So if we add the directed edges $\{(l, -(l+2)) :\ 0 \le l \le k-2\} \cup \{(-l, l+1) :\ 1 \le l \le k-1\}$ to the directed graph $R_n \cup B_n$, then in this new graph, the number of directed paths from 0 to k is equal to the number of non-directed paths from 0 to k in the non-directed graph $R_n \cup B_n$. (In fact we constructed a bijection among the non-oriented and the oriented paths of the two graphs: whenever a non-oriented path of the original graph uses a backward edge, the corresponding new edge should be used in the new graph to create an oriented

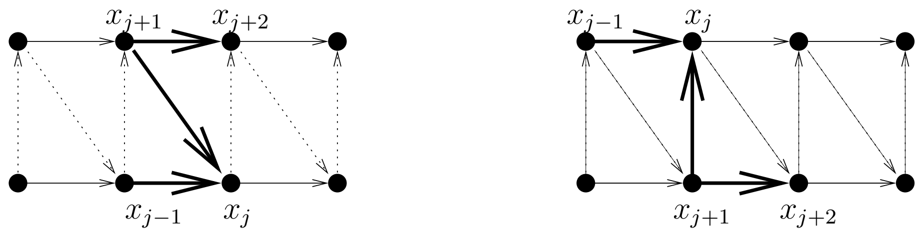

Figure 3: Backward red edges in oriented paths

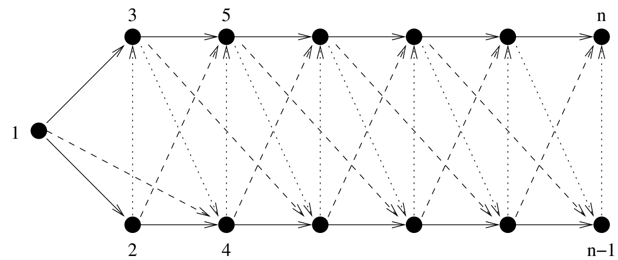

Figure 4: $R_n \cup B_n$ with the added oriented edges

path, and vice versa.)

If we reindex the vertices as in fig.4 the recurrence formula above follows, as any vertex l is adjacent to an incoming edge from $l-3, l-2$ and $l-1$. Solving this formula we get that $D^n_{\text{cycle}} \geq cx_0^n$ for some constant c, where x_0 is the unique real root of the equation $x^3 - x^2 - x - 1 = 0$, $x_0 = 1.8392...$ (Note that the number of cycles in $R_n \cup B_n$ not containing the edge $(0, k)$ grows polynomially in n, so it is negligible compared to number of cycles containing $(0.k)$.)

4.2 Upper Bound on D^n_{cycle}

Proof of the upper bound in Theorem 4.1: To establish the inequality let us consider two trees B_n and R_n (a blue one and a red one) on the same n-element vertex set. In a cycle in $R_n \cup B_n$ there are consecutive red edges, then consecutive blue edges, then red edges again, and so on. (Edges that are both red and blue will be considered as red.) A maximal path of consecutive red edges will be called a *red segment* (a blue segment is defined similarly. The number of blue segments in a cycle clearly equals the number of red segments, and since each segment contain at least one edge, the number of red segments is at most $\lfloor \frac{n}{2} \rfloor$ and is at least 1 (for a cycle without red segment is a blue cycle, which is impossible, since B_n is a tree).

We will count the cycles in $R_n \cup B_n$ partitioning them according to the number of red segments. So we have to show that there are at most $\binom{n}{2i}$ cycles having i red segments. To do this first note that in a fixed cycle the set of the endpoints of the red segments and the set of the endpoints of the blue segments are just the same.

Lemma 4.2 *Given a tree and $2i$ vertices of its vertex set, then there is at most one way to choose i vertex-disjoint paths in the tree with the given vertices as endpoints.*

Proof: By induction on i. If $i = 1$ then clearly the statement holds, for in a tree there is exactly one path from any vertex to any other vertex.

Let $i > 1$. A set of vertex-disjoint paths defines naturally a matching on the set of endpoints. Notice, that an edge in such a matching corresponds always to the same path (for there is one single path between any two vertices of a tree). Let us suppose to the contrary, that there are two different sets of paths satisfying the statement of the lemma. If there exists a common edge in the corresponding matchings, then we remove this common edge (and its endpoints) and by induction we arrive at a contradiction. If there is no such edge, then the two matchings have together $2i$ edges on $2i$ vertices, so there should be a cycle involving these edges, that is there should be a cycle in the corresponding paths, which contradicts the fact that our graph is a tree. □

To finish the proof of the upper bound, observe that by Lemma 4.2, the mapping which maps any cycle in $R_n \cup B_n$ to the set of endpoints of the segments of the cycle is injective. The statement of the theorem follows. □

Remark: It is easy to see, that in the statement of Lemma 4.2, "at most" is necessary if the tree is not a path. Hence we know that the upper bound for D^n_{cycle} holds with strict inequality for non-path trees. Since the graphs of the construction in the previous subsection were paths, one may conjecture, that trees with maximal distance are paths. But even if this conjecture is false, the question that how many cycles we can have in the union of two paths is a distance-type question. To see this we just have to figure out what the path-distance of two paths is. Since a path is a cycle-free (connected) graph in which all vertices have degree at most 2, the forbidden collection of subgraphs consists of the cycles and the 3-star (i.e. the graph consisting of the edges $(1,2),(1,3),(1,4)$). But since any vertex has degree at most 4 in the union of two paths, any vertex can be the middle vertex of at most $\binom{4}{3} = 4$ 3-stars, therefore there can be at most $4n$ 3-stars in the union of two paths on n vertices. As $4n$ is negligible compared to the exponentially growing number of cycles, $D_{path}(P_1^n, P_2^n) = \Theta(D_{cycle}(P_1^n, P_2^n))$ for the sequence of optimal pairs of paths.

References

[1] B. Bollobás, On complete subgraphs of different order, Math. Proc. Camb. Philos. Soc., 80 (1976) 19-24

[2] G.O.H. Katona, Two applications (for search theory and truth functions) of Sperner type theorems, Period. Math. Hung., 3 (1973) 19-26

[3] B. Patkós, How different can two intersecting families be?, Electronic Journal of Combinatorics, 12 (2005) R24

[4] N. Sauer, A generalization of a theorem of Turán, Res P. 61, Department of Mathematics, Calgary, 1968

[5] E. Sperner, Ein Satz über Untermenge einer endlichen Menge, Math Z., 27 (1928) 544-548

[6] P. Turán, An extremal problem in graph theory, Mat. Fiz. Lapok, 48 (1941) 436-452 (in Hungarian)