Journal of Combinatorial Theory, Series A 144 (2016) 292–305

Contents lists available at ScienceDirect

Journal of Combinatorial Theory, Series A

www.elsevier.com/locate/jcta

Separation with restricted families of sets ☆

Zsolt Lángi [a], Márton Naszódi [b,c], János Pach [b,d], Gábor Tardos [d], Géza Tóth [d,a]

[a] *Budapest University of Technology, Budapest, Hungary*
[b] *EPFL, Lausanne, Switzerland*
[c] *Eötvös University, Budapest, Hungary*
[d] *Rényi Institute, Budapest, Hungary*

A R T I C L E I N F O

Article history:
Available online 15 July 2016

Keywords:
Search theory
Separation
VC-dimension
Erdős–Szekeres theorem

A B S T R A C T

Given a finite n-element set X, a family of subsets $\mathcal{F} \subset 2^X$ is said to *separate* X if any two elements of X are separated by at least one member of $\mathcal{F}$. It is shown that if $|\mathcal{F}| > 2^{n-1}$, then one can select $\lceil \log n \rceil + 1$ members of $\mathcal{F}$ that separate X. If $|\mathcal{F}| \geq \alpha 2^n$ for some $0 < \alpha < 1/2$, then $\log n + O(\log \frac{1}{\alpha} \log\log \frac{1}{\alpha})$ members of $\mathcal{F}$ are always sufficient to separate all pairs of elements of X that are separated by some member of $\mathcal{F}$. This result is generalized to simultaneous separation in several sets. Analogous questions on separation by families of bounded Vapnik–Chervonenkis dimension and separation of point sets in $\mathbb{R}^d$ by convex sets are also considered.

☆ Zsolt Lángi was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. Márton Naszódi was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the National Research, Development, and Innovation Office, NKFIH Grant PD-104744. János Pach was supported by Swiss National Science Foundation Grants 200020-162884, 200021-165977, and by the Hungarian–Mexican Grant TÉT-12-MX-1-2013-0006. Gábor Tardos was supported by the *Cryptography* "Lendület" project of the Hungarian Academy of Sciences and by the National Research, Development, and Innovation Office, NKFIH Grants K-116769 and SNN-117879. Géza Tóth was supported by the National Research, Development, and Innovation Office, NKFIH Grants K-111827 and K-83767.

E-mail addresses: zlangi@math.bme.hu (Z. Lángi), marton.naszodi@math.elte.hu (M. Naszódi), pach@cims.nyu.edu (J. Pach), tardos.gabor@renyi.mta.hu (G. Tardos), toth.geza@renyi.mta.hu (G. Tóth).

http://dx.doi.org/10.1016/j.jcta.2016.06.002

1. Introduction

For a set X, we say that a subset of X *separates* two elements if it contains one of them and does not contain the other. For a family $\mathcal{F}$ of subsets of X, we say that it *separates* a pair of elements of X if at least one member of $\mathcal{F}$ separates them. Furthermore, $\mathcal{F}$ *separates* X if every pair of distinct elements of X is separated by $\mathcal{F}$.

Suppose your computer is infected by a virus $x \in X$, where X is the set of known computer viruses. You want to perform a number of tests to find out which virus it is. Each test detects a certain set of viruses, which can be associated with the test. Let $\mathcal{F}$ denote the family of subsets of X associated with the tests you can perform. These tests are sufficient to identify the virus if, and only if $\mathcal{F}$ separates X. The number of tests necessary is at least $\log |X|$, where log stands for the base 2 logarithm. On the other hand, there is a family $\mathcal{F} \subset 2^X$ with $|\mathcal{F}| \leq \lceil \log |X| \rceil$ that separates X. This is the starting point of a rich discipline called *combinatorial search theory*; see [1] [2], [4], [10].

Any fixed pair of distinct elements in X is separated by $2^{|X|-1}$ subsets of X, thus a family $\mathcal{F}$ with $|\mathcal{F}| > 2^{|X|-1}$ separates X. Our first theorem states that in this case, even a small subfamily of $\mathcal{F}$ does the job.

Theorem 1. *Let X be a finite set, $\mathcal{F} \subseteq 2^X$ with $|\mathcal{F}| > 2^{|X|-1}$. Then X can be separated by a subfamily $\mathcal{G} \subseteq \mathcal{F}$ of cardinality at most $\lceil \log |X| \rceil + 1$.*

This statement is almost tight, but not completely. Indeed, for $|X| = 5$, Theorem 1 guarantees the existence of a 4-member separating family, but it is easy to verify that 3 sets suffice. In the following generalization we give the best possible bound.

Theorem 2. *Let $X_1, \ldots, X_k$ be pairwise disjoint sets with $|X_i| \leq n$ for $i = 1, 2, \ldots, k$. Let $X = \bigcup_{i=1}^k X_i$. If $\mathcal{F} \subseteq 2^X$ satisfies $|\mathcal{F}| > 2^{|X|-1}$, then $\mathcal{F}$ has a subfamily of cardinality at most $\lceil \log n \rceil + 1$ that separates X_i for every i ($i = 1, 2, \ldots, k$).*

The above bound is tight, that is, the same statement is false for every $n \geq 2$, if we replace $\lceil \log n \rceil + 1$ by $\lceil \log n \rceil$.

We call $|\mathcal{F}|/2^{|X|}$ the *density* of $\mathcal{F}$. If this slips below $1/2$, we cannot guarantee the existence of a small subfamily separating X, as $\mathcal{F}$ itself does not necessarily separate X. But, as claimed by the next theorem, we can still find a small subfamily separating all pairs in X that $\mathcal{F}$ separates. We state this result for simultaneous separation.

Theorem 3. *Let $X_1, \ldots, X_k$ be pairwise disjoint sets with $|X_i| \leq n$ for all $i = 1, 2, \ldots, k$. Let $X = \bigcup_{i=1}^k X_i$ and $\mathcal{F} \subseteq 2^X$. Then $\mathcal{F}$ has a subfamily of size at most $\lceil \log n \rceil + C \log \frac{1}{\alpha} \log \log \frac{1}{\alpha}$ separating every pair in each X_i that is separated by $\mathcal{F}$. Here $\alpha = |\mathcal{F}|/2^{|X|}$ is the density of $\mathcal{F}$ and C is a universal constant.*

As we explain in Section 2 (after the proof of Theorem 2), it is possible that this statement holds without the $\log \log \frac{1}{\alpha}$ factor. However, there exist families $\mathcal{F}$ for which