Proof. Let $\mathbf{A}$ be a set of r random 0-1 vectors of length n, each entry of which takes values 0 or 1 independently and with equal probability 1/2. For every fixed set S of k coordinates and for every fixed vector $v \in \{0,1\}^S$,

$$\text{Prob}\,(v \notin \mathbf{A}\restriction_S) = \prod_{a \in \mathbf{A}} \text{Prob}\,(v \neq a\restriction_S) = \prod_{a \in \mathbf{A}} \left(1 - 2^{-|S|}\right) = \left(1 - 2^{-k}\right)^r.$$

Since there are only $\binom{n}{k}2^k$ possibilities to choose a pair (S, v), the set $\mathbf{A}$ is *not* (n,k)-universal with probability at most $\binom{n}{k}2^k(1-2^{-k})^r$, which is strictly smaller than 1. Thus, at least one set A of r vectors must be (n,k)-universal, as claimed. □

By using the fact that $\binom{n}{k} < (\mathrm{e}n/k)^k$ and $(1-2^{-k})^r \leqslant \mathrm{e}^{-r/2^k}$, and by a special simple construction for $k \leqslant 2$ (cf., for example, Exercise 11.4), it is easy to derive from the last theorem that for every n and k there is an (n,k)-universal set of size at most $k2^k \log n$.

18.6 Cross-intersecting families

A pair of families $\mathcal{A}, \mathcal{B}$ is *cross-intersecting* if every set in $\mathcal{A}$ intersects every set in $\mathcal{B}$. The *degree* $d_{\mathcal{A}}(x)$ of a point x in $\mathcal{A}$ is the number of sets in $\mathcal{A}$ containing x. The *rank* of $\mathcal{A}$ is the maximum cardinality of a set in $\mathcal{A}$.

If $\mathcal{A}$ has rank a, then, by the pigeonhole principle, each set in $\mathcal{A}$ contains a point x which is "popular" for the members of $\mathcal{B}$ in that $d_{\mathcal{B}}(x) \geqslant |\mathcal{B}|/a$. Similarly, if $\mathcal{B}$ has rank b, then each member of $\mathcal{B}$ contains a point y for which $d_{\mathcal{A}}(y) \geqslant |\mathcal{A}|/b$. However, this alone does not imply that we can find a point which is popular in *both* families $\mathcal{A}$ and $\mathcal{B}$. It turns out that if we relax the "degree of popularity" by one-half, then such a point exists.

Theorem 18.6 (Razborov–Vereshchagin 1999). *Let $\mathcal{A}$ be a family of rank a and $\mathcal{B}$ be a family of rank b. Suppose that the pair $\mathcal{A}, \mathcal{B}$ is cross-intersecting. Then there exists a point x such that*

$$d_{\mathcal{A}}(x) \geqslant \frac{|\mathcal{A}|}{2b} \quad \textit{and} \quad d_{\mathcal{B}}(x) \geqslant \frac{|\mathcal{B}|}{2a}.$$

Proof. Assume the contrary and let $\mathbf{A}, \mathbf{B}$ be independent random sets that are uniformly distributed in $\mathcal{A}, \mathcal{B}$ respectively. That is, for each $A \in \mathcal{A}$ and $B \in \mathcal{B}$, $\text{Prob}\,(\mathbf{A} = A) = 1/|\mathcal{A}|$ and $\text{Prob}\,(\mathbf{B} = B) = 1/|\mathcal{B}|$. Since the pair $\mathcal{A}, \mathcal{B}$ is cross-intersecting, the probability of the event "$\exists x (x \in \mathbf{A} \cap \mathbf{B})$" is equal to 1. Since the probability of a disjunction of events is at most the sum of the probabilities of the events, we have

$$\sum_x \text{Prob}\,(x \in \mathbf{A} \cap \mathbf{B}) \geqslant 1.$$

Let X_0 consist of those points x for which

$$\frac{d_{\mathcal{A}}(x)}{|\mathcal{A}|} = \mathrm{Prob}\,(x \in \mathbf{A}) < \frac{1}{2b},$$

and X_1 consist of the remaining points. Note that by our assumption, for any $x \in X_1$,

$$\mathrm{Prob}\,(x \in \mathbf{B}) = \frac{d_{\mathcal{B}}(x)}{|\mathcal{B}|} < \frac{1}{2a}$$

holds. By double counting (see Proposition 1.6), $\sum_x d_{\mathcal{A}}(x) = \sum_{A \in \mathcal{A}} |A|$. Hence,

$$\begin{aligned} &\sum_{x \in X_1} \mathrm{Prob}\,(x \in \mathbf{A} \cap \mathbf{B}) = \sum_{x \in X_1} \mathrm{Prob}\,(x \in \mathbf{A}) \cdot \mathrm{Prob}\,(x \in \mathbf{B}) \\ &< \frac{1}{2a} \cdot \sum_{x \in X_1} \mathrm{Prob}\,(x \in \mathbf{A}) \leqslant \frac{1}{2a} \cdot \sum_x \mathrm{Prob}\,(x \in \mathbf{A}) \\ &= \frac{1}{2a} \cdot \sum_x \frac{d_{\mathcal{A}}(x)}{|\mathcal{A}|} = \frac{1}{2a\,|\mathcal{A}|} \cdot \sum_x d_{\mathcal{A}}(x) = \frac{1}{2a\,|\mathcal{A}|} \cdot \sum_{A \in \mathcal{A}} |A| \leqslant \frac{a\,|\mathcal{A}|}{2a\,|\mathcal{A}|} = \frac{1}{2}. \end{aligned}$$

In a similar way we obtain

$$\sum_{x \in X_0} \mathrm{Prob}\,(x \in \mathbf{A} \cap \mathbf{B}) < \frac{1}{2},$$

a contradiction. □

This theorem has the following application to boolean functions. Recall that a DNF (a disjunctive normal form) is an Or of monomials, each being an And of literals, where a literal is a variable x_i or its negation $\overline{x}_i$. The *size* of a DNF is the number of its monomials, and the *rank* is the maximum length of a monomial in it.

Given a pair F_0, F_1 of DNFs, a decision tree for such a pair is a usual decision tree (see Sect. 10.4) with the exception that this time each leaf is labeled by F_0 or by F_1. Given an input $v \in \{0,1\}^n$ we follow the (unique) path from the root until we reach some leaf. We require that the DNF labeling the leaf so reached must be falsified by v. Clearly, not every pair of DNFs will have a decision tree: we need that, for every input, at least one of the DNFs outputs 0 on it; that is, the formula $F_0 \wedge F_1$ must be not satisfiable; in this case we say that the pair of DNFs is *legal*. As before, the *depth* of a decision tree is the number of edges in a longest path from the root to a leaf. Let $DT(F_0, F_1)$ denote the minimum depth of a decision tree for the pair F_0, F_1.

In particular, if F_1 is a DNF of some boolean function f, and F_0 is a DNF of its negation $\overline{f}$, then their And $F_0 \wedge F_1$ is clearly unsatisfiable, and a decision tree for this pair is precisely the decision tree for f. If, moreover, F_0 has rank a and F_1 has rank b, then we already know (see Theorem 10.12) that $DT(F_0, F_1) \leqslant ab$.

Using Theorem 18.6, we can prove an upper bound which takes into account not only the rank of DNFs but also their size.