Chapter 7. PROBABILISTIC METHOD

The probabilistic method is a powerful tool to demonstrate the existence of combinatorial objects.

In some cases this method can be used also to show the existence of algorithms.

two Basic arguments of the probabilistic method:

1. The expectation argument Any random variable V assumes at least one value that is not smaller than its expectation EV, and at least one value that is not greater than its expectation EV.

2. The sampling argument If an object chosen randomly from a universe U satisfies a property P with a positive probability, then there must be an object in U that satisfies the property P.

The above two simple ideas have a surprising power.

Their power comes from our ability to reformulated, in various ways, counting arguments in the language of probability and then *to apply* various tools of probability theory.

Example: **One can show that for every $n \times n$ 0-1-matrix A, and for any randomly chosen vector $b \in \{-1,+1\}^n$, it holds**

$$\|Ab\| \leq 4\sqrt{n \ln n}$$

with probability at least $1 - \frac{2}{n}$.

From that we may conclude that for every such matrix A, there always exists a vector $b \in \{-1,+1\}^n$ such that $\|Ab\| \leq 4\sqrt{n \ln n}$.

BASIC IDEA and an EXAMPLE

Probabilistic method consists of two stages.

1. **A "thought experiment" $\mathcal{E}$ is designed in which a carefully chosen random process (called usually as *an experiment* - say dice tossing) P plays a key role.**
2. **The random process P is first analyzed and then some conclusions are made that are, or look as, independent of the experiment $\mathcal{E}$.**

EXAMPLE - MAX-CUT PROBLEM

Max-cut-problem: Given an undirected graph $G = (V, E)$ with $n = |V|$ vertices and $m = |E|$ edges, we need to partition the vertices in V into two sets A and B in such a way that maximizes number of edges (u, v), with $u \in A, v \in B$.

Theorem: For any undirected graph $G = (V, E)$ with n vertices and m edges, there is a partition of the vertex set V into two sets A and B such that

$$|\{(u, v) \in E \,|\, u \in A, v \in B\}| \geq \frac{m}{2}.$$

Proof: Let us consider the following experiment: Each vertex of G is independently and equiprobably assigned to either A or B.

For any edge (u, v), the probability that its end-vertices are in different sets is $\frac{1}{2}$.

By linearity of expectations, the expected number of edges with end-vertices in different sets is thus

$$\mathbf{E}[\text{cut-size}] = \frac{m}{2}.$$

That implies that there must be a partition satisfying the theorem.

DERANDOMIZATION

- **In some cases the proof of existence obtained by the probabilistic method can be converted into an efficient randomized algorithm.**
- **In some other cases the existence proof obtained by the probabilistic method can be converted even to an efficient deterministic algorithm - such a process is called derandomization.**

FROM PROOF of EXISTENCE to LAS VEGAS ALG. for MAX-SAT

We show how to transform the argument about the existence of a partition with at least $\frac{m}{2}$ of edges to a Las Vegas algorithm.

Let us design a partition $C(A, B)$ using randomization described above and denote

$$p = Pr(C(A, B) \geq \frac{m}{2})$$

Then

$$\frac{m}{2} = \mathbf{E}[C(A, B)] = \sum_{i \leq m/2-1} iPr(C(A, B) = i) + \sum_{m/2 \leq i \leq m} iPr(C(A, B) = i) \leq (1-p)\left(\frac{m}{2} - 1\right) + pm$$

which implies that

$$p \geq \frac{1}{m/2 + 1}$$

Therefore, the expected number of samples before finding a cut with value at least $\frac{m}{2}$ is therefore just $\frac{m}{2} + 1$.

Since we can test in polynomial time whether the value of the cut determined by the sample is at least $m/2$, by counting the edges crossing the cut, we have Las Vegas algorithm for finding the cut.

HAMILTONIAN PATHS in TOURNAMENTS

A **tournament** is a complete directed graph.

A **Hamiltonian path** in a graph $G = (V, E)$ is a path that visits each vertex (representing a player) of V exactly once.

Theorem Every tournament has a Hamiltonian path.

Proof will be by induction on the number n of vertices in a tournament.

Induction step. Suppose that every tournament with at most n vertices has a Hamiltonian path and let a tournament $T = (V, E)$ with $n + 1$ vertices is given.

Choose any vertex v and define two sets of vertices

$$A = \{u \mid (u, v) \in E\} \quad B = \{u \mid (v, u) \in E\}.$$

Two subgraphs induced by these two sets of vertices form tournaments. By induction both of them have Hamiltonian paths.

By connecting these two paths through the node v we get a Hamiltonian path for T.

NUMBER of PATHS in TOURNAMENTS

Theorem For any n there exists a tournament of size n with $n!/2^{n-1}$ Hamiltonian paths.

Proof Generate a random tournament $T = (V, E)$, $V = \{1, \ldots, n\}$, by randomly choosing direction for all edges of K_n - of a complete graph of n vertices.

For each permutation σ on V let X_σ be a random variable defined as follows:

$$X_\sigma = \begin{cases} 1, & \sigma \text{ defines a path on } T \\ 0 & \text{otherwise} \end{cases}$$

For all σ, $\Pr(X_\sigma = 1) = (\frac{1}{2})^{n-1} = \mathbf{E}[X_\sigma]$.

Let X be a random variable counting the number of Hamiltonian paths in T

$$X = \sum_{\sigma \in \mathrm{Perm}(n)} X_\sigma.$$

Theorem now follows from the following calculations:

$$\mathbf{E}[X] = \mathbf{E}[\sum_\sigma X_\sigma] = \sum_\sigma \mathbf{E}[X_\sigma] = n!(\frac{1}{2})^{n-1}.$$

TOURNAMENTS WITH PROPERTY S_k

In a tournament, if there is an edge from a node A to a node B, then we say that the player A beats B,

A tournament T is said to have property S_k, for an integer k, if for any set of k players there is one player that beats all of them.

Theorem If $\binom{n}{k}(1-2^{-k})^{n-k} < 1$, then there is a tournament on n vertices with property S_k.

Proof Consider a random tournament with a set of n nodes. For any subset K of k vertices let A_K be the event that there is no player/node that beats all players/nodes in K. Clearly,

$$\Pr[A_K] = (1-2^{-k})^{n-k}.$$

Therefore

$$\Pr[\bigvee_{K\subset V, |K|=k} A_K] \le \sum_{K\subset V,|K|=k} \Pr[A_K] = \binom{n}{k}(1-2^{-k})^{n-k} < 1$$

Therefore, with a positive probability no event A_K occurs. That is, there is a tournament on n vertices that has the property S_k.

COMMENT

If K is a set on k players, then probabiity that a player P not in K beats all of them is 2^{-k} and probbaility is $1-2^{-k}$ that he does not beat all of them.
Since there are $n-k$ players outside the group K, the probability that none of them beats all players in K is $(1-2^{-1})^{n-k}$.

MAX-SAT - PROBLEM

Given are m clauses in conjunctive normal form over n variables. Find assignment (of truth values to variables) that maximizes the number of satisfied clauses.

Theorem: For any set of m clauses there is a truth assignment that satisfies at least $\frac{m}{2}$ clauses.

Proof: Suppose each variable is set to 0 or 1 independently and equiprobably and for $1 \leq i \leq m$, let a random variable $Z_i = 1$ if the i-th clause is satisfied.

The probability that a clause with k literals is not satisfied by this random assignment is 2^{-k}.

The probability that a clause with k literals is satisfied is $1 - \frac{1}{2^k} \geq \frac{1}{2}$ what implies that $\mathbf{E}[Z_i] \geq \frac{1}{2}$ for all i.

The expected number of clauses satisfied by a random assignment is

$$\sum_{i=1}^{m} \mathbf{E}[Z_i] \geq \frac{m}{2}$$

Therefore, there exists at least one assignment for which

$$\sum_{i=1}^{m} Z_i \geq \frac{m}{2}.$$

FROM a PROOF of EXISTENCE to an ALGORITHM

We show now that a variant of the probabilistic proof of existence in the last theorem can be turned into an approximation algorithm.

Notation:

- I - an input instance - a set of clauses.
- $m_*(I)$ - the maximum number of clauses of I that can be satisfied.
- $m_A(I)$ - the number of clauses of I satisfied by the algorithm A.
- **performance ratio** of the algorithm A : $\inf_I \frac{m_A(I)}{m_*(I)}$.

If A achieves a performance ratio α, we say that A is an α-approximation algorithm.

If A is a randomized algorithm, then $m_A(I)$ is a random variable and in such a case $m_A(I)$ is replaced by $\mathbf{E}[m_A(I)]$ in the definition of the performance ratio.

We now show a randomized algorithm for MAXSAT with performance ratio $\frac{3}{4}$.

The procedure in the proof of the last theorem actually yields a randomized algorithm whose guaranteed performance is $1 - 2^{-k}$, provided every clause contains at least k literals.

As a consequence, *we have a randomized $\frac{3}{4}$-approximation algorithm for instances of MAX-SAT in which every clause contains at least 2 literals.*

We now show another algorithm that performs especially well when there are (many) clauses consisting of a single literal.

Finally, we show that on any input instance, one of the two designed algorithms yields a randomized $\frac{3}{4}$-approximation algorithm.

BASIC IDEA

is similar as in the case of the global wiring problem.

1. **Reformulate the problem as a 0-1 linear programming problem.**
2. **Solve the corresponding rational linear programming problem.**
3. **Use the randomized rounding technique.**

Notation: ***With each clause C_j, in the given input formula, we associate an indicator variable $z_j \in \{0,1\}$, that indicates whether or not the clause C_j is satisfied.***

Moreover, to each variable x_i we assign an indicator variable y_i defined by

$$x_i = true \Longleftrightarrow y_i = 1 \qquad .$$

C_j^+ **- set of indices of variables that appear uncomplemented in** C_j
C_j^- **- set of indices of variables that appear complemented in** C_j

0-1 linear programming reformulation of the MAX-SAT problem. Find

$$y_i, z_j \in \{0,1\} \quad (\forall i, j) \qquad\qquad (\star)$$

such that the sum

$$\sum_{j=1}^{m} z_j$$

is maximized and

$$\sum_{i \in C_j^+} y_i + \sum_{i \in C_j^-} (1 - y_i) \geq z_j \quad (\forall j). \tag{1}$$

Rational linear programming problem is then obtained by replacing the condition $(\star)$ by the condition

$$y_i, z_j \in [0,1] \quad (\forall i, j).$$

Let $\hat{y}_i$ $(\hat{z}_i)$ be the value of variable y_i (z_i) obtained by solving the rational linear programming problem. Clearly, $\sum_{i=1}^{n} z_j \leq \sum_{j=1}^{n} \hat{z}_j$.

We first show that using the randomized rounding method we obtain a truth assignment for which the expected number of satisfied clauses is at least $(1 - \frac{1}{e}) \sum_j \hat{z}_j$.

This will follow from the Lemma shown on next slide in case we use the following: Randomized rounding: each y_i is, independently, set to 1 with probability $\hat{y}_i$.

Analysis: For an integer k denote $\beta_k = 1 - (1 - \frac{1}{k})^k > 1 - \frac{1}{e}$.

Lemma: Let C_j be a clause with k literals. The probability that it is satisfied by the randomized rounding is at least $\beta_k \hat{z}_j > (1 - \frac{1}{e})\hat{z}_j$.

Proof: Without loss of generality we can assume that

$$C_j : \quad x_1 \vee \ldots \vee x_k$$

By constrain (**??**) $\qquad \hat{y}_1 + \ldots + \hat{y}_k \geq \hat{z}_j$.

Observe that the clause C_j remains unsatisfied by randomized rounding method only if every variable y_i is rounded to 0. Since each variable is rounded fully independently, this occurs with probability

$$\prod_{i=1}^{k} (1 - \hat{y}_i).$$

It remains to show that

$$\underbrace{1 - \prod_{i=1}^{k} (1 - \hat{y}_i)}_{\text{This is minimized if } \hat{y}_i = \frac{\hat{z}_j}{k}} \qquad \geq \ \beta_k \hat{z}_j$$

It suffices to show $1 - (1 - \frac{z}{k})^k \geq \beta_k z$ *for all* $0 < z < 1$.

Since $f(x) = 1 - (1 - \frac{x}{k})^k$ is a concave function it suffices to verify the above inequality for $x = 0$ and $x = 1$.

FINAL RESULT I

From the last Lemma, and from the linearity of expectations, it follows:

Theorem: Given an instance of MAX-SAT, the expected number of clauses satisfied by linear programming and randomized rounding is at least $(1 - \frac{1}{e})$ time the maximum number of clauses that can be satisfied on that instance.

Comparison of peformances of our two algorithms for MAX-SAT

k	$1-2^{-k}$	β_k
1	0.5	1.0
2	0.75	0.75
3	0.875	0.704
4	0.938	0.684
5	0.969	0.672

We now show that on any instance one of the algorithms is a $\frac{3}{4}$ - approximation algorithm:

Let n_1 denote the expected number of clauses that are satisfied when each variable is independently set to 1 with probability $\frac{1}{2}$ (what corresponds to the first algorithm).

Let n_2 denote the expected number of clauses that are satisfied when we use the linear programming followed by the randomized rounding (what corresponds to the second algorithm).

FINAL RESULT II

Theorem:

$$\max\{n_1, n_2\} \geq \frac{3}{4}\sum_j \hat{z}_j.$$

Proof: It suffices to show that $(\frac{n_1+n_2}{2}) \geq \frac{3}{4}\sum_j \hat{z}_j$.

Let S_k denote the set of clauses that contain k literals. We know that

$$n_1 = \sum_k \sum_{C_j \in S_k} (1-2^{-k}) z_j \geq \sum_k \sum_{C_j \in S_k} (1-2^{-k}) \hat{z}_j,$$

$$n_2 \geq \sum_k \sum_{C_j \in S_k} \beta_k \hat{z}_j.$$

Thus

$$\frac{n_1+n_2}{2} \geq \sum_k \sum_{C_j \in S_k} \frac{(1-2^{-k})+\beta_k}{2} \hat{z}_j.$$

Since $(1-2^{-k})+\beta_k \geq \frac{3}{2}$ for all k, we get

$$\frac{n_1+n_2}{2} \geq \frac{3}{4}\sum_k \sum_{C_j \in S_k} \hat{z}_j = \frac{3}{4}\sum_j \hat{z}_j.$$

RAMSEY NUMBER PROBLEM

The Ramsey number $R(k, l)$ is the smallest integer n such that in any 2-coloring of the edges of a complete graph K_n, on n nodes, by red and blue, there either is a red K_k (i.e. a complete subgraph on k vertices with edges coloured red), or there is a blue K_l.

Ramsey (1930) showed that $R(k, l)$ is finite for any two integers k and l.

We use the probabilistic method to show a lower bound on $R(k, k)$.

LOWER BOUND on the RAMSEY NUMBER

Theorem **If $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$, then $R(k,k) > n$.**

Corollary **Since $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$ for $n = 2^{k/2}$ (that is if $k = 2\lg n$), the above theorem implies that $R(k,k) > 2^{k/2}$ for all $k \geq 3$.**
Proof of Theorem **Let n, k satisfy the assumption of the theorem. Consider a random 2-coloring of the edges of K_n by red or blue.**

For any fixed set R of k vertices, let A_R be the event that the induced subgraph of K_n on R is monochromatic (i.e. that either all its edges are red or they are all blue).

Clearly, $Pr(A_R) = 2\frac{1}{2^{\binom{k}{2}}} = 2^{1-\binom{k}{2}}$. Since there are $\binom{n}{k}$ possible choices for R, the probability that at least one of the events A_R occurs is at most $\binom{n}{k}2^{1-\binom{k}{2}} < 1$.

Thus, with positive probability, no event A_R occurs and therefore there is a 2-coloring of K_n without a monochromatic K_k, that is, $R(k,k) > n$.

Note that if $k \geq 3$ and $n = \lfloor 2^{k/2} \rfloor$, then

$$\binom{n}{k} 2^{1-\binom{k}{2}} < \frac{2^{1+k/2}}{k!} \frac{n^k}{2^{k^2/2}} < 1$$

and hence $R(k,k) > 2^{k/2}$ for all $k \geq 3$.

ALGORITHMIC CONSEQUENCES

Last theorem implies that there is an edge two-coloring of K_n without a monochromatic $K_{2\lg_2 n}$. It is therefore natural to ask whether we can find efficiently such a coloring. Since there are $2^{\binom{n}{2}}$ possible colorings, an exhaustive search cannot be efficient.

However, a closer look at the proof of the last theorem shows that the proof can be used to produce effectively a coloring that is very likely to be good. This is due to the fact for large k if $n = \lfloor 2^{k/2} \rfloor$, then

$$\binom{n}{k} 2^{1-\binom{k}{2}} < \frac{2^{1+\frac{k}{2}}}{k!} \left(\frac{n}{2^{k/2}}\right)^k \leq \frac{2^{1+\frac{k}{2}}}{k!} << 1$$

because $\binom{n}{k} \leq \frac{n^k}{k!}$. Hence, a random coloring of K_n is very likely not to contain a monochromatic $K_{2\lg_2 n}$.

As a consequence, if we need to find a two-coloring of edges of K_{1024} without a monochromatic K_{20} we can simply produce a random two-coloring and then the probability that it contains a monochromatic K_{20} is less than $\frac{2^{11}}{20!}$ what is much, much less than probability of error in any proof that a certain coloring is good.

DELETION METHOD

So called deletion method can be useful in some cases when it seems to be difficult to apply the probabilistic method directly.

The proof, by the deletion method, that a certain combinatorial object $\mathcal{O}$ exists, consists, conceptually, of two stages:

- **It is first shown that with positive probability an object $\mathcal{O}'$ exists that is very close, in some sense, to $\mathcal{O}$.**
- **Secondly, $\mathcal{O}'$ is changed, to obtain $\mathcal{O}$, and it is shown that the probability of the existence of $\mathcal{O}$ remains positive.**

MIN-MAX TRIANGLE PROBLEM

Let $S = \{p_1, \ldots, p_n\}$ be a set of points located in a unit square of the plane.

Consider all triangles whose vertices are points of S and let $T(S)$ be the area of such a triangle with the minimum area.

The following theorem asserts that for any n there is a set S of n points such that $T(S)$ is not too small. Namely, it holds:

Theorem For any n there is a set S of n points in a unit square such that $T(S) \geq \frac{1}{100n^2}$.

PROOF. BASIC IDEA: $2n$ points are chosen randomly in the unit square. All triangles created by these points are tested and those that are "too small" are eliminated - by deleting one vertex from each of them until only n of nodes are left. The idea is that this way we will be left with enough points, and with no too-small-area triangle.

PROOF Choose uniformly $2n$ points in the unit square.

For any three points p, q, r let $A(p, q, r)$ denote the area of the triangle these three points create.

For any real numbers $0 \leq b$ and $\Delta b \leq 1$ it holds

$$\Pr[b \leq ||q - p|| \leq b + \Delta b] \leq \pi(b + \Delta b)^2 - \pi b^2 = 2\pi b \Delta b + \pi(\Delta b)^2,$$

where $||p - q||$ denote the Euclidean distance between p and q.

(Observe that inequality follows from the fact that ring defined by b and $b + \Delta b$ may not be completely contained in the unit square.)

Let (p, q) be the base of the triangle (p, q, r) and let $||p - q|| = b$.

We show now how to estimate $\Pr[A(p, q, r) \leq \varepsilon]$ for any $\varepsilon > 0$ and $p.q.r$.

Such an event happens when the height h of the triangle is $\leq \frac{2\varepsilon}{b}$ and therefore r is not farther than $\frac{2\varepsilon}{b}$ from the line of points p and q.

The probability that this happens is less than $\frac{2\sqrt{2}\varepsilon}{b}$, because r has to be in a strip of width $\frac{2\varepsilon}{b}$ and length less than $\sqrt{2}$. Hence,

$$\begin{aligned}\Pr[A(p,q,r) \leq \varepsilon] &= \int_{b=0}^{\sqrt{2}} \Pr[b \leq ||p-q|| \leq b + \Delta b] \times \Pr[\text{triangle. h.} \leq \frac{2\varepsilon}{b}] \\ &\leq \int_{b=0}^{\sqrt{2}} (4\pi b) \frac{2\sqrt{2}\varepsilon}{b} \Delta b = 16\pi\varepsilon.\end{aligned}$$

Let us now compute the expected number of triangles with the area $\leq \varepsilon = \frac{1}{100n^2}$.

Let S' be a set of $2n$ points uniformly distributed in the unit square. For each triple (p_i, q_i, r_i) in S' let X_{p_i,q_i,r_i} be the indicator variable having value 1 if the area of the triangle determined by (p_i, q_i, r_i) is less than $\varepsilon = \frac{1}{100n^2}$.

The probability that the area of some specific triangle is less than $\frac{1}{100n^2}$ is less than $\frac{16\pi}{100n^2} \leq \frac{0.6}{n^2}$ – This is also the expected value of X_{p_i,q_i,r_i}.

If X denotes the number of triangles with area less than $\frac{1}{100n^2}$, then

$$\mathbf{E}[X] = \sum_{p,q,r \in S'} \mathbf{E}[X_{p_i,q_i,r_i}] \leq \binom{2n}{3} 0.6n^{-2} \leq n.$$

Finally, by throwing away an arbitrary vertex from each of such "small area triangles", we are left with a new set S'' of points the expected size of which (of S''), is not less than n, in which no small-area-triangles exist.

Therefore, there exists a set S'', of size n, such that $T(S'') \geq \frac{1}{100n^2}$.

ANOTHER EXAMPLE - INDEPENDENT SETS OF VERTICES

We show a lower bound on the size of the largest independent set of vertices in certain graphs.

Definition An independent set of a graph $G = (V, E)$ is a subset of vertices of V such that no two vertices in the set are adjacent. Denote by $\alpha(G)$ the size of the largest independent set of a graph G.

Our aim is to prove the following lower bound:

Theorem If $|V| = n$ and $|E| = nk/2$ for a graph $G = (V, E)$, then $\alpha(G) \geq \frac{n}{2k}$.

Proof ideas:

- To choose randomly a subset of vertices that would be a candidate for an independent set.
- To show, using the probabilistic argument, that there is a subset of the chosen set of vertices that has many more vertices than edges.
- By deleting one vertex from each of such edges, to produce an independent set.

PROOF

Create a set $S \subset V$ by putting into S each vertex independently with probability p (to be specified later). It therefore holds for the average size of S:

$$\mathbf{E}[|S|] = np.$$

Let G_S be the subgraph of $G = (V, E)$, induced by S.

For any $e \in E$ let Y_e be the indicator variable that has value 1 if $e \in E(G_S)$ - the set of edges of G_S - and 0 otherwise.

$\mathbf{E}[Y_e] = p^2$ because an edge belongs to $E(G_S)$ iff both if its endpoints are in S, what happens with probability p^2. Let $Y = |E(G_S)|$. It holds

$$\mathbf{E}[Y] = \mathbf{E}[\sum_{e \in E} Y_e] = \sum_{e \in E} \mathbf{E}[Y_e] = \frac{nk}{2}p^2.$$

After deleting all edges from G_S, by dropping a vertex from each of such edges, it remains a set S' of the expected size $\mathbf{E}[|S'|] = \mathbf{E}[|S| - Y]$, and therefore

$$\mathbf{E}[|S| - Y] = \mathbf{E}[|S|] - \mathbf{E}[Y] = np - \frac{nk}{2}p^2.$$

The last expression has the largest value for $p = \frac{1}{k}$ and in such a case

$$\mathbf{E}[|S^*|] = \mathbf{E}[|S| - Y] = \frac{n}{2k}.$$

EXPANDING GRAPHS

Informally, an expanding graph is a graph in which the number of neighbors of any sufficiently small set of vertices S is larger than $c|S|$ for some positive constant $c > 1$.

OR-concentrators are a special type of expanding graphs.

Definition:An (n, d, α, c) OR-concentrator is a bipartite multigraph $G = (L, R, E)$ with independent sets of vertices L and R, each of cardinality n, such that

1. **every vertex in L has degree at most d;**
2. **for any subset S of vertices from L such that $|S| \leq \alpha n$ there are more than $c|S|$ neighbors in R.**

In most of the applications, it is desirable to have d as small and c as large as possible.

Of particular interest is to study OR-concentrators in which α, c and d are constants fixed independently of n, with $c > 1$.

Finding an explicit construction of OR-concentrators is a non-trivial task. However, the probabilistic method can be used to show the existence of such concentrators.

Theorem: There is an integer n_0 such that for all $n > n_0$ there is an $(n, 18, \frac{1}{3}, 2)$ OR-concentrator.

Proof: The first part of the proof will be for all (n, d, c, α) OR-concentrators.

Consider a random bipartite graph with two disjoint sets of vertices, L and R, each of n vertices, in which each vertex of L chooses randomly and independently d vertices from R as neighbours.

For any integer s **let ξ_s denote the event that a (bad) subset of s vertices of L has fewer than cs neighbours in R**.

We first derive an upper bound on $Pr[\xi_s]$ and then an upper bound on the sum of $\Pr[\xi_s]$ over all $s \leq \alpha n$. **This way we obtain a non-trivial upper bound on the probability that a random graph fails to be an OR-concentrator with parameters we seek.**

Fix any subset $S \subseteq L$ of size s, and any subset $T \subseteq R$ of size cs.

(There are $\binom{n}{s}$ ways of choosing S, and $\binom{n}{cs}$ ways of choosing T.)

The probability that T contains all of at most ds neighbours of the vertices in S is $\left(\frac{cs}{n}\right)^{ds}$.

The probability of the event that all ds edges going out from some s vertices of L fall within any cs vertices of R is bounded by

$$\Pr[\xi_s] \le \binom{n}{s}\binom{n}{cs}\left(\frac{cs}{n}\right)^{ds}.$$

Using the inequality $\displaystyle\binom{n}{s} \le \left(\frac{ne}{s}\right)^s$ we obtain

$$\begin{aligned}\Pr(\xi_s) &\le \binom{n}{s}\binom{n}{cs}\left(\frac{cs}{n}\right)^{ds}\\ &\le \left(\frac{ne}{s}\right)^{s}\left(\frac{ne}{cs}\right)^{cs}\left(\frac{cs}{n}\right)^{ds}\\ &= \left[\left(\frac{s}{n}\right)^{d-c-1} e^{1+c}\, c^{d-c}\right]^{s}.\end{aligned}$$

In a special case, for $\alpha = \frac{1}{3}$ and $s \leq \alpha n$ we have

$$\begin{aligned} \Pr(\xi_s) &= \left[\left(\frac{1}{3}\right)^{d-c-1} e^{1+c}\, c^{d-c} \right]^s \\ &\leq \left[\left(\frac{c}{3}\right)^{d} (3e)^{c+1} \right]^s, \end{aligned}$$

and, in addition, for $c = 2,\ d = 18$

$$\Pr(\xi_s) = \left[\left(\frac{2}{3}\right)^{18} (3e)^3 \right]^s.$$

Since:

$$r = \left(\frac{2}{3}\right)^{18} (3e)^3 \leq \frac{1}{2}$$

we have

$$\sum_{s \geq 1} Pr(\xi_s) \leq \sum_{s \geq 1} r^s = \frac{r}{1-r} < 1.$$

EXISTENCE of CONCENTRATORS - II

Theorem: For n sufficiently large, there is a bipartite graph $G = (L, R, E)$ with $|L| = n$, $|R| = 2^{\lg^2 n}$ such that:

1. Every subset of $\frac{n}{2}$ vertices of L has more than $2^{\lg^2 n} - n$ neighbors in R;
2. No vertex in R has more than $12 \lg^2 n$ neighbors in L.

Proof: **Consider a random graph in which each vertex in L chooses, independently and uniformly, $d = 2^{\lg^2 n}(4\lg^2 n/n)$ neighbors in R.**

We show now that such a random graph violates each of the two properties with a positive probability smaller than $\frac{1}{2}$.

From that it follows that such a random graph satisfies both of the above properties with a positive probability.

Similarly, as in the proof of previous theorem, the probability that there is a set of $\frac{n}{2}$ vertices in L having fewer than $2^{lg^2 n} - n$ neighbors in R is at most

$$P = \binom{n}{\frac{n}{2}} \binom{2^{\lg^2 n}}{2^{\lg^2 n} - n} \left(1 - \frac{n}{2^{\lg^2 n}}\right)^{\frac{dn}{2}}.$$

From that it can be shown, using inequalities

$$\binom{g}{k} \leq \left(\frac{ge}{k}\right)^k, \; 1 - \frac{n}{2^{\lg^2 n}} \leq e^{-\frac{n}{2^{\lg^2 n}}}$$

that the probability P, that Property 1 is violated, is less than $\frac{1}{2}$

Concerning property 2: Note that the expected number of neighbors of a vertex in R is $\frac{nd}{2^{\lg^2 n}} = \frac{n2^{\lg^2 n}(4\lg^2 n/n)n}{2^{\lg^2 n}} = 4\lg^2 n$.

Using the bound $F^+(\mu,\delta) < \left(\frac{e}{1+\delta}\right)^{(1+\delta)\mu}$ with $\mu = 4\lg^2 n,\ \delta = 2$ we get that the probability of exceeding $(1+\delta)\mu = 12\lg^2 n$ neighbors is less than

$$\left(\frac{e}{3}\right)^{12\lg^2 n}.$$

Since R contains $2^{\lg^2 n}$ vertices and $(\frac{e}{3})^{12} < \frac{1}{2}$, we have

$$2^{\lg^2 n}\left(\frac{e}{3}\right)^{12\lg^2 n} \to 0$$

and therefore

$$2^{\lg^2 n}\left(\frac{e}{3}\right)^{12\lg^2 n} < \frac{1}{2},$$

for sufficiently large n.

Therefore, it is guaranteed that the probability that the property 2 is violated at any node of R is also less than $\frac{1}{2}$.

OBLIVIOUS ROUTING REVISITED

Another example will now be shown how the probabilistic method can be used to prove the existence of an algorithm.

We focus on the minimal number of random bits needed by randomized oblivious routing algorithms. We derive the following surprising result.

1. We prove that our randomized oblivious routing algorithm, that uses $d2^d$ random bits to route a d-dimensional hypercube, uses much too much random bits.
2. **The proof of the existence (by the probabilistic method) of a randomized routing algorithm that uses (within a constant factor) only $3d$ random bits to route d-dimensional hypercube.**

To remember: $\sqrt{\frac{2^d}{d}}$ is a lower bound on oblivious deterministic routings on a d-dimensional hypercube H_d.

Question: *How much randomness (how many random bits) is (are) needed to have a routing algorithm with the expected running time $O(d)$?*

Observation I: A randomized oblivious algorithm for permutation routing is a probability distribution on a set of deterministic oblivious routing algorithms.

Observation II. Each deterministic oblivious algorithm for a 2^d-node network is a set of $2^{2d} = 2^d \times 2^d$ routes, one for each source-target pair.

Note: Every randomized oblivious algorithm can be expressed by sequences

$$(A_1, \ldots, A_r), \quad (p_1, \ldots, p_r),$$

where each A_j is a deterministic oblivious routing algorithm and each p_j is the probability that we use A_j on a run of the randomized routing algorithm.

Note: We know:

1. **With 0 random bits the expected running time of any oblivious routing algorithm on H_d is $\Omega\left(\sqrt{\frac{2^d}{d}}\right)$;**
2. **For the randomized oblivious routing the expected running time is $O(d)$ and $d\,2^d$ random bits are used (each of 2^d nodes chooses a random d-bit auxiliary goal).**

Are so many random bits indeed necessary for efficient randomized routing?

Theorem: Consider any randomized oblivious algorithm for the permutation routing on the hypercube with $n = 2^d$ nodes. If this algorithm uses k random bits, then its expected running time is $\Omega\left(2^{-k}\sqrt{\frac{2^d}{d}}\right)$.

Proof: Since a randomized oblivious algorithm is a probability distribution on deterministic oblivious algorithms, if only k random bits are used, then at least one of these deterministic algorithms, say A_1, is chosen with probability at least 2^{-k}, on any execution.

By the lower bound theorem, there is an input that requires time $\Omega\left(\sqrt{\frac{2^d}{d}}\right)$ on the deterministic algorithm A_1.

If such an input is put into the randomized algorithm and A_1 is chosen, then the running time is at least $\Omega\left(\sqrt{\frac{2^d}{d}}\right)$. The expected running time of the randomized algorithm is therefore $\Omega\left(2^{-k}\sqrt{\frac{2^d}{d}}\right)$.

Corollary: Any randomized oblivious algorithm for permutation routing on the hypercube H_d must use $k = \Omega(d)$ random bits in order to achieve the expected running time $O(d)$. (Actually, $k = \frac{d}{2} - \frac{3}{2}\lg d$)

Theorem: For every d there exists a randomized oblivious scheme (algorithm) for a permutation routing on the hypercube with $n = 2^d$ nodes that uses only $3d$ random bits and still runs in the expected time $15d$ at most.

Proof: Notation We say that a set $\mathcal{B} = \{B_1, \ldots, B_t\}$ of deterministic oblivious permutation routing algorithms on H_d is an efficient routing scheme, if for any input instance, the expected number of steps using a randomly chosen algorithm from $\mathcal{B}$ is at most $15d$.

To prove the theorem, we show that for every $n = 2^d$ there is an efficient routing scheme for H_d with $t = 2^{3d} = n^3$.

Our resulting randomized routing scheme will randomly choose n^3 of n^n possible deterministic oblivious routing algorithms. (n^n is due to the fact that there are n sources and for each one we can choose from n possible intermediate destinations.)

Let us denote such deterministic algorithms by

$$A_j,\ 1 \leq j \leq n^n.$$

On an n-node network there are $n!$ distinct possible instances of the permutation routing problem, one for each permutation on $\{1, 2, \ldots, n\}$.

For a permutation π_i, $1 \leq i \leq n!$, let us call a deterministic oblivious routing algorithm A_j **good** if A_j routes π_i in $14d$ or fewer steps, and **bad** otherwise.

By our randomized routing result: (with probability at least $1 - \frac{1}{n}$ every packet reaches its destination in $14n$ or fewer steps) for any particular π_i a fraction of at most $\frac{1}{n}$ of the algorithms A_j are bad - which of them are bad may differ from instance to instance.

Experiment: Choose n^3 indices $i_1, \ldots, i_{n^3}$, randomly, independently and uniformly from the set $\{1, \ldots, n^n\}$. We show that the set of deterministic algorithms

$$\mathcal{A} = \{\mathcal{A}_{i_1}, \ldots, \mathcal{A}_{i_{n^3}}\}$$

is an efficient routing scheme with a positive probability. This will imply that an efficient routing scheme exists for any $n = 2^d$.

For any π_i, a fraction of at most $\frac{1}{n}$ of the algorithms $\mathcal{A}_1, \ldots, \mathcal{A}_{n^n}$ is bad. Therefore, the expected number of algorithms in $\mathcal{A}$ that are bad for π_i is at most $n^3 \cdot \frac{1}{n} = n^2$.

Let the indicator variable X_j be 1 if A_{i_j} is bad, $1 \leq j \leq n^3$, and 0 otherwise. Thus $\mathbf{E}[\sum_j X_j] \leq n^2$. Since X_j are independent, we may apply Chernoff bound

$F^+(\mu,\delta) < e^{-\frac{\mu\delta^2}{4}}$ (Chapter 6, page 10), to obtain (for $\mu = n^2, \delta = 1$) that the probability that more than $2n^2$ of the algorithms in $\mathcal{A}$ are bad for π_i is $\leq e^{\frac{-n^2}{4}}$.

Let $\mathcal{B}_i$ denote the **bad** event that more than $2n^2$ algorithms in $\mathcal{A}$ are bad for π_i. Then, for $n \geq 4$

$$Pr[\bigcup_{i=1}^{n!} \mathcal{B}_i] \leq \sum_{i=1}^{n!} Pr(\mathcal{B}_i) \leq n!\, e^{-\frac{n^2}{4}} < 1 \text{ (by Stirling's formula for } n!)$$

Therefore, with the positive probability, not more than $2n^2$ of the algorithms in $\mathcal{A}$ are bad for any π_i. Therefore, there exists a subset of n^3 algorithms from $\{\mathcal{A}_1, \ldots, \mathcal{A}_{n^n}\}$ with the property that at most $2n^2$ in this subset are bad for any π_i.

Let us denote this subset $\mathcal{B} = \{B_{i_1}, \ldots, B_{i_{n^3}}\}$. $\mathcal{B}$ is an efficient routing scheme: for any π_i a randomly chosen algorithm form $\mathcal{B}$ fails to route π_i within $14n$ steps with probability at most $\frac{2n^2}{n^3} = \frac{2}{n}$.

From that one can deduce that the expected number of steps using an algorithm randomly chosen from $\mathcal{B}$ is less than $15n$.

THE LOVÁSZ LOCAL LEMMA - MOTIVATION I

This is one of the most elegant and useful tools to apply the probabilistic method.

Let $E_1, \ldots, E_n$ be a set of bad events in some probabilistic space. Lovasz lemma will be used to show that, under certain conditions, there is an element that is not included in any of these bad events.

This can be shown easily if events $E_1, \ldots, E_n$ are independent. Indeed, in such a case, for any subset $I \subseteq [1, n]$,

$$\boldsymbol{Pr}(\bigcap_{i \in I} E_i) = \prod_{i \in I} \boldsymbol{Pr}(E_i)$$

In such a case also complements of events $E_1, \ldots E_n$ are independent and therefore if $\boldsymbol{Pr}(E_i) < 1$ for all i, then

$$\boldsymbol{Pr}(\bigcap_{i=1}^{n} \bar{E}_i) = \prod_{i=1}^{n} \boldsymbol{Pr}(\bar{E}_i) > 0$$

and therefore there is an element of the sample space that is not included in any bad event.

THE LOVASZ LOCAL LEMMA - FORMULATION

The Lovasz local lemma generalizes the preceding argument to the case where events are not mutually independent, but the dependence is limited.

In order to formulate lemma we need the following definition of mutual independence. An event E is mutually independent of events $E_1, \ldots, E_n$, if for any $I \subseteq [1, n]$,

$$\boldsymbol{Pr}\left(E \mid \bigcap_{i \in I} E_i\right) = \boldsymbol{Pr}(E).$$

Dependency between events can be represented in terms of a dependency graph.

Definition Dependency graph for a set of events $E_1, \ldots, E_n$ is a graph $G = (V, E)$ such that $V = \{1, \ldots, n\}$ and for $i = 1, \ldots, n$, the event E_i is mutually independent of the events $\{E_j \mid (i, j) \notin E\}$.

Lovasz Local Lemma **Let $E_1, \ldots, E_n$ be events and p be a fixed number such that**

- **for all i, $\boldsymbol{Pr}(E_i) \leq p$;**
- **the degree of the dependency graph given by $E_1, \ldots, E_n$ is bounded by a d;**
- **$4dp \leq 1$.**

Then

$$\boldsymbol{Pr}\left(\bigcap_{i=1}^{n} \bar{E}_i\right) > 0$$

APPLICATION - EDGE-DISJOINT PATHS

- **Assume that n pairs of users need to communicate using edge-disjoint paths on a given network.**
- **Assume that each ith pair of users can choose a path from a collection F_i of m paths for a fixed m.**
- **We show, using simple version of Lovasz local lemma that, if the possible paths do not share too many edges, then there is a way to choose n edge-disjoint paths connecting the n given pairs.**

Theorem If, for a fixed k, any path in any F_i shares edges with no more than k paths in any $F_j, j \neq i$, and $8nk/m \leq 1$, then there is a way to choose n edge-disjoint paths connecting the n given pairs.

PROOF of THEOREM

- **Let each ith pair chooses a path from F_i randomly, with probability $\frac{1}{m}$.**
- **Let $E_{i,j}$ be the event that paths chosen by pairs i and j share at least one edge.**
- **Since for any i and $j \neq i$ the path chosen from F_i shares edges with at most k paths in any $F_j, j \neq i$, we have**

$$p = \mathbf{Pr}(E_{i,j}) \leq \frac{k}{m}.$$

- **Let d be the degree of the dependency graph of all events $E_{,j}$. Since the event $E_{i,j}$ is independent of all events $E_{i',j'}$, when $i' \notin \{i,j\}$ and $j' \notin \{i,j\}$, we have $d < 2n$.**

- **Since**

$$4dp < \frac{8nk}{m} \leq 1$$

all conditions of the Lovacz local lemma are satisfied and therefore

$$\boldsymbol{Pr}\left(\bigcap_{i \neq j} \bar{E}_{i,j}\right) > 0$$

and therefore there is such a choice of n paths that are edge disjoint.

THE LOVÁSZ LOCAL LEMMA - GENERAL CASE

Suppose we have n events each of which occur with probability at most $\frac{1}{2}$. Let each of these events corresponds to one of n ways in which a probabilistic experiment could fail.

If the event were independent, we could then assert that at such an experiment with probability at least 2^{-n}, none of the events occurs.

The Lovász Local Lemma generalizes the above setting and result to the case where each of the events is independent of all but small number of other events.

Lovász Local Lemma: Let $G = (V, E)$ be the dependency graph for events $\xi_1, \ldots, \xi_n$ in a probability space. Suppose that there exist $x_i \in [0, 1]$, for $1 \leq i \leq n$, such that

$$\Pr[\xi_i] \leq \prod_{(i,j)\in E} (1 - x_j).$$

Then

$$\Pr[\bigcap_{i=1}^{n} \bar{\xi}_i] \geq \prod_{i=1}^{n}(1 - x_i).$$

PAUL ERDÖS

- Paul Erdös, a Hungarian mathematician, died in September 1996 at the age of 83.
- He was the most prolific mathematicians of the twentieth century with over 1500 papers written and more than 490 collaborators.
- He can be seen as the founder (in 1947) of the probabilistic method.

THOUGHTS ABOUT ERDÖS

- In our century, in which mathematics is so strongly dominating by "theory constructors" Erdös remained the prince of problem solvers and the absolute monarch of problem posers. (E. Strauss)
- Erdös had no job; he worked constantly. He had no home, the world was his home. Possessions were a nuisance, money a bore. He lived on a web of trust, traveling ceaselessly from Center to Center, spreading his mathematical pollen. His enormous talents and energies were given entirely to the Temple of Mathematics.

STORIES about ERDÖS

- It is six in the morning. The house is sleeping. I prove and conjecture.

Paul Erdös in a letter to Vera Sós

STORY about the BOOK

- Story about The Book. Erdös liked to talk about The Book, that contains all theorems of mathematics and for each of them one proof - beautiful, aesthetic and insightful - so called The Book proof. Each time one of his conjectures was resolved in "an ugly way" Erdös congratulated the prover, but added "let us now look for a Book proof".
- In 1985 he started his lecture in a math camp by saying: "You do not have to believe in the God, but you should believe in The Book".