Lecture 4: Random Networks I - The method of positive probability

Prof. Sotiris Nikoletseas

University of Patras
and CTI

ΥΔΑ ΜΔΕ, Patras
2020 - 2021

The Probabilistic Method - major applications (I)

A powerful tool used in many applications in different topics:

I) Study of random graph models ($G_{n,p}$, $G_{n,R}$, $G_{n,k}$ etc) which are:

- typical instances for average case analysis of graph algorithms and
- abstract models of modern networks (sensor networks, social networks etc.)

Note. Time efficiency of an algorithm: best case, worst case, average case. Average case analysis requires random inputs (an input distribution).

The Probabilistic Method - major applications (II)

II) Design and analysis of randomized algorithms:

- evolution based on random choices (not deterministic decisions based on the input)
- solutions provided a) either are always correct but their running time is a random variable (Las Vegas algorithms) b) or may be erroneous but are correct w.h.p. (Monte Carlo algorithms)
- trade-off performance (faster, simpler than deterministic algorithms) with very small, controlled error probability.

The core of the method

The Probabilistic Method

- uses simple techniques
  - the Basic Method
  - Linearity of Expectation
- as well as complex ones
  - the Local Lemma
  - Martingales
  - Markov Chains

but there is a common, underlying concept:

The core of the method

Non-constructive proof of existence of combinatorial structures that have certain desired properties.

The Basic Method (method of positive probability)

- Construct (by using abstract random experiments) an appropriate probability sample space of combinatorial structures (thus, the sample points correspond to the combinatorial structures whose existence we try to prove).
- Prove that the probability of the desired property in this space is positive (i.e. non-zero).

$$\Downarrow$$

There is at least one point in the space with the desired property.

$$\Downarrow$$

There is at least one combinatorial structure with the desired property.

Characteristics of the P.M.

- comprehensible, pretty short proofs
- simple (basic knowledge of Probabilistic Theory, Graph Theory, Combinatorics suffices)
- elegant
  - qualitative ideas, subtle notions
  - not lengthy, mechanical operations
- still very powerful (use to resolve extremely difficult problems)

Examples in this lecture

1. Monochromatic arithmetic progressions (Van der Waerden property)
2. Ramsey Numbers
3. Tournaments

(I) Van der Waerden property

Definition 1

$W(k)$ is the smallest natural number n, such that for any two-coloring of the numbers $1, 2, ..., n$ there is a monochromatic arithmetic progression of k terms.

Theorem 1

$$W(k) > 2^{\frac{k}{2}}$$

Proof of Theorem 1 (1/3)

- We construct a probability space by two-coloring the numbers $1, 2, ..., n$ at random, equiprobably for the two colors and independently for every number. Clearly, the sample points of this space are random two-colorings of the n numbers.
- Let S be any fixed arithmetic progression of k terms.
- Define the event $M_S := \{S \text{ is monochromatic}\}$.
  - i.e, all terms of S must have the same color.
- Compute the probability $\Pr[M_S]$.
  - every term is colored red (or blue) with probability $1/2$
  - all k terms are red-colored (or blue-colored) with probability $(\frac{1}{2})^k$

$$\Pr[M_S] = \left(\frac{1}{2}\right)^k + \left(\frac{1}{2}\right)^k = 2^{1-k}$$

Proof of Theorem 1 (2/3)

- Define the event $M := \{\exists$ at least one monochromatic arithmetic progression of k terms $\} \Rightarrow M = \bigcup_{|S|=k} M_S$.
- An arithmetic progression of k terms is defined uniquely by its two first terms $\Rightarrow$ There are at most $\binom{n}{2}$ arithmetic progressions $\Rightarrow \#(S : |S| = k) \leq \binom{n}{2}$
- Using Boole's inequality we can compute $\Pr[M]$

$$\Pr[M] = \Pr\left\{\bigcup_{|S|=k} M_S\right\} \leq \sum_{|S|=k} \Pr[M_S] \leq \binom{n}{2} 2^{1-k}$$

Proof of Theorem 1 (3/3)

- We easily get:

$$\Pr[M] < \frac{n^2}{2} 2^{1-k} = \frac{n^2}{2^k}$$

- If $n < 2^{\frac{k}{2}}$ then $\Pr[M] < 1 \Rightarrow \Pr[\overline{M}] > 0$.
- Hence, there is a two-coloring without a monochromatic arithmetic progression of k terms when $n < 2^{\frac{k}{2}}$.
- Thus, $W(k) > 2^{\frac{k}{2}}$.

□

(II) Ramsey Numbers

Definition 2

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

Difficulty of computation:

- Ramsey (1930) proved that $R(k,l)$ is finite
- Greenwood and Gleason (1955) computed $R(3,3)=6$ and $R(4,4)=18$
- since then there is no notable progress - $R(4,5)$ is still unknown
- Erdös suggested that $R(6,6)$ is too difficult to be computed

Ramsey Numbers

Definition 3

$R(k,k)$: diagonal Ramsey number (a monochromatic K_k is required).

Theorem 2 (Erdös, 1947)

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

Proof of Theorem 2 (1/3)

- Construct a probability sample space by two-coloring at random, equiprobably (for the two colors) and independently (for the edges) every edge of K_n.
- Let S be any fixed set of k vertices and consider the edges induced.
- Define the event $M_S := \{S \text{ is monochromatic}\}$.
  - i.e. all $\binom{k}{2}$ edges in S have the same color.
- Compute the probability $\Pr[M_S]$.
  - every edge is colored red (or blue) with $1/2$ probability

$$\Pr[M_S] = \left(\frac{1}{2}\right)^{\binom{k}{2}} + \left(\frac{1}{2}\right)^{\binom{k}{2}} = 2^{1-\binom{k}{2}}$$

Proof of Theorem 2 (2/3)

- Define the event $M := \{\exists$ at least one monochromatic set of k vertices$\}$.
- Hence, $M = \bigcup_{|S|=k} M_S$.
- Using Boole's inequality we can compute the $\Pr[M]$

$$\Pr[M] \leq \sum_{|S|=k} \Pr[M_S] = \binom{n}{k} 2^{1-\binom{k}{2}}$$

Proof of Theorem 2 (3/3)

- If $\Pr[M] < 1 \Rightarrow \Pr[\overline{M}] > 0$
  $\Rightarrow$ if $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$ then there is a point in the sample space without $M \Rightarrow$ there is a monochromatic K_k.
- Hence, it must be $R(k,k) > n$.

□

Lower Bound of Ramsey Numbers

- We proved that if $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ then $R(k,k) > n$
- If $\binom{n}{k}2^{1-\binom{k}{2}} \sim 1$ then we can find the best possible lower bound for $R(k,k)$ (with this derivation).
- By using Stirling's formula and binomial approximation we obtain:

$$\frac{n^k}{k!} \cdot 2^{1-\binom{k}{2}} \sim \frac{n^k}{\sqrt{2\pi k}\left(\frac{k}{e}\right)^k} \cdot 2^{-\frac{k^2}{2}} \sim 1$$

$$\Rightarrow n^k \sim \sqrt{2\pi k} \cdot \left(\frac{k}{e}\right)^k \cdot 2^{\frac{k^2}{2}}$$

$$\Rightarrow R(k,k) > n \sim \frac{k}{e\sqrt{2}} 2^{\frac{k}{2}}$$

(III) Tournaments

Definition 4

A **tournament** T_n is a complete directed graph on n vertices i.e., for every pair (i, j), there is either an edge from i to j or from j to i, but not both.

Why do we call these graphs tournaments?

- Each vertex corresponds to a team playing at some tournament.
- The directed edge (i, j) means that team i wins team j.
- all teams play against each other.

The S_k Property

Definition 5

A tournament T_n is said to have **property** S_k if for any set of k vertices in the tournament, there is some vertex that has a directed edge to each of those k vertices.

Theorem 3 (Erdös, 1963)

$\forall k, \exists$ *a tournament* T_n *that has the property* S_k.

Proof of Theorem 3 (1/2)

- Construct a probability sample space with points random tournaments by choosing the direction of each edge at random, equiprobably for the two directions and independently for every edge.
- Let S be any fixed set of k teams and define the event $M_S := \{\nexists$ a team that wins all teams in $S\}$.
- For any team, the probability to win all teams in S is $(\frac{1}{2})^k$.
- Hence, the probability of not winning at least one of them is $1 - (\frac{1}{2})^k$.
- The probability that this is happening for all $n-k$ teams that don't belong in S is:

$$\Pr[M_S] = \left(1 - \left(\frac{1}{2}\right)^k\right)^{n-k}$$

Proof of Theorem 3 (2/2)

- Define the event $M := \{\exists$ a set S of k teams such that $\nexists$ a team $u : u \notin S$ that wins all teams in $S\}$.
- $M = \bigcup_S M_S$
- Using Boole's inequality we can compute $\Pr[M]$

$$\Pr[M] \leq \sum_{S,|S|=k} \Pr[M_S] = \binom{n}{k} \left(1 - \left(\frac{1}{2}\right)^k\right)^{n-k}$$

- If $\binom{n}{k} \left(1 - \left(\frac{1}{2}\right)^k\right)^{n-k} < 1$ then $\Pr[M] < 1 \Rightarrow \Pr[\overline{M}] > 0$.
- Hence, there is a tournament with property S_k.