The Probabilistic Method

Camilo Espinosa

May 16, 2018

Contents

Abstract

In this paper, we will examine a group of techniques that fall under the umbrella of the Probabilistic Method and will solve several problems in order to illustrate how the method is used. We will first introduce the probabilistic and algebraic concepts necessary to understand the method, followed by a description of the workflow the method entails. We will then examine applications of the method to prove statements about Ramsey Numbers, specific colorings of hypergraphs, and Hamiltonian paths in directed graphs. Afterwards, we will delve into advanced applications of the method by stating and proving the Symmetric Lovasz Local Lemma, which we will use to show an interesting result on cycles in directed graphs. We will then briefly discuss some of the newer results that have followed the Lovasz Local Lemma, specifically the Algorithmic Lovasz Local Lemma and a problem that can be solved in polynomial time because of this theorem.

1 Background

1.1 Probability Basics

Definition 1.1.1: A *Finite Probability Space* is defined as a pair $(\Omega, \mathbf{P})$ where Ω is a set of elements ω called *elementary events* and $\mathbf{P}$ is a *Probability Function* from Ω to the interval $[0, 1]$ such that

$$\sum_{\omega \in \Omega} \mathbf{P}(\omega) = 1$$

In particular, we call a subset A of Ω an *event*, and we define $\mathbf{P}(A)$ as

$$\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\omega)$$

Another identity that follows directly from the definition of probability is that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

We also define $\overline{A}$, the *complement* of an event A, and its associated probability as

$$\overline{A} = \Omega / A$$
$$P(\overline{A}) = 1 - P(A)$$

Lemma 1.1.2 (Union Bound): Given a set $\{A_1, A_2, \ldots, A_n\}$ of events in a probability space $(\Omega, \mathbf{P})$, we must have that

$$P\Big[\bigcup_{i=1}^{n} A_i\Big] \leq \sum_{i=1}^{n} P[A_i]$$

Definition 1.1.3 (Conditional Probability): Given two events A and B in a probability space $(\Omega, \mathbf{P})$, we define the conditional probability $P(A|B)$ as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

We can intuitively understand this as the probability that event A will occur if we know that event B occurs; hence, we constrain on the probability of the intersection and compare this to the probability of event B happening.

Definition 1.1.4 (Independence): We call two events A and B in a probability space $(\Omega, \mathbf{P})$ independent from each other if we have that

$$P(A|B) = P(A)$$

Note that this is equivalent to $P(B|A) = P(B)$, and, more generally, that $P(A \cap B) = P(A)P(B)$.

Definition 1.1.5 (Random Variables): A random variable on a probability space $(\Omega, \mathbf{P})$ is a function $\mathbf{X}$ from Ω to $\mathbb{R}$. Generally, a random variable will assay a specific property of the combinatorial object we are examining, like the number of edges in a random graph, so that we can use properties of random variables and probabilty to deduce properties of the combinatorial object.

Definition 1.1.6 (Expectation): Given a random variable $\mathbf{X}$ in a probability space $(\Omega, \mathbf{P})$, we define the expectation $\mathbb{E}[X]$ as

$$\mathbb{E}[X] = \sum_{\omega \in \Omega} \mathbf{X}(\omega) P(\omega)$$

The core idea behind the Expectation is to assay for, roughly, the *average* value the random variable will take when evaluated across all of the events that we are considering.

Lemma 1.1.7 (Linearity of Expectation): One of the most useful results regarding the expectation of random variables talks about how random variables that are themselves sums of simpler random variables behave. Just like with events, we can have that random variables are dependent or independent from each other. However, regardless of their dependence, given random variables $\mathbf{X}$ and $\mathbf{Y}$ in a probability space $(\Omega, \mathbf{P})$ and integers a and b, we have that

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$$

This result is called *Linearity of Expectation* and can be used to prove many other properties about random variables.

1.2 Useful Approximations

In this paper, we will use several approximations that will be useful in proving some of the theorems we will cover. The two key approximations that we will use are:

1. Stirling's Approximation: Given $n \in \mathbb{N}$, we have that

$$(\frac{n}{e})^n \leq n! \leq en(\frac{n}{e})^n$$

2. Approximating with *e*: Give $n \in \mathbb{N}$ and a small p with $p > 0$, we have

$$(1-p)^n \leq e^{-np}$$

With these tools in hand, we can proceed to learn about the simplest form of the probabilistic method and use it to solve a few relevant problems.

2 The Basic Probabilistic Method

2.1 Description of the Method

The idea behind the probabilistic method is to attack problems in which we wish to prove the existence of a specific combinatoric object (whether it be a graph, coloring, or bound) but can't easily show a construction of the object. So, instead of presenting a successful object, we randomize over all the possible configurations in order to show that the probability that a successful ubject exists is non-zero. Hence, while we can't actually show the graph or coloring that exhibits the desired property, it must exist, and that is usually enough for our purposes.

Now, there are two different ways in which we can show that the probability of a successful object existing is non-zero: we can either explicitly find that the probability our object exists is greater than 0 or we can show that the probability of a bad object existing is less than 1. To illustrate the general flow of the method, we will start with a simple example involving Ramsey Numbers.

2.2 Ramsey Numbers

Ramsey theory is a branch of combinatorics which aims to find conditions in which specific regular properties of structures are unavoidable, like the existence of cliques in a graph or monochromatic progressions in a random coloring of numbers. The first object of study of Ramsey Theory was the number of vertices needed so that, given numbers k and l, there will always be either a k-clique or independent set size l regardless of how we draw the edges of the graph.

Definition 2.2.1: The Ramsey number $R(k, l)$ is the minimum n such that any graph of at least n vertices will have either a k-clique or an independent set size l regardless of how the edges of the graph are drawn.

Proposition 2.2.2: $R(k, k) > 2^{\frac{1}{2}k-1}$

Proof. Suppose our graph G has n vertices, and we randomly draw edges between each pair of vertices with probability 1/2. The probability that a given set of k vertices has either all or none of the edges in it (and is therefore a k-clique or independent set size k) is $2 * (1/2)^{\binom{k}{2}}$. Looking at the probability over the entire graph and using the Union Bound, we have that

$$P(G \text{ contains } k\text{-clique or independent set size } k) \leq \binom{n}{k} * 2^{1-\binom{k}{2}}$$

Since we want to look at the event that at least one graph on n vertices does not have either a k-clique or independent set size k, we want to see that this probability is < 1. Hence we have

$$\binom{n}{k} * 2^{1-\binom{k}{2}} < 1 \Leftrightarrow \binom{n}{k} < 2^{\binom{k}{2}-1}$$

And using the approximation $\binom{n}{k} \leq n^k$ we get that

$$n \leq 2^{\frac{k}{2}-1} \Rightarrow \binom{n}{k} \leq n^k \leq 2^{\frac{1}{2}k^2-k} < 2^{\binom{k}{2}-1}$$

This means that, for $n \leq 2^{\frac{k}{2}-1}$, there is at least one graph for which there are no k-cliques or independent sets size k, so that $R(k, k) > 2^{\frac{k}{2}-1}$, as wanted. □

Now we will examine a more complicated example.

2.3 Hypergraph Colorings

Definition 2.3.1: A k-uniform hypergraph is a graph $G = (V, E)$ such that each edge $e \in E$ is a k-tuple $(v_1, v_2, \ldots, v_k)$. On n vertices, there are $\binom{n}{k}$ possible edges in G.

Proposition 2.3.2: Suppose we color each vertex of a k-uniform hypergraph G one of two colors. Let c_k be the smallest integer such that if G has c_k edges then it has at least one monochromatic edge. Then $c_k \geq 2^{k-1}$.

Proof. We will proceed by showing that any k-uniform hypergraph G with $m \leq 2^{k-1}$ edges must be colorable with 2 colors so that no edge is monochromatic. To do so, randomly assign each edge to a color, red or blue, with probability 1/2. We have that the probability that a given edge is monochromatic is $2 * (1/2)^k$. Therefore, the probability that the graph has at least one monochromatic edge is the union of the events that each edge is monochromatic which is less than the sum of the probabilities that each individual edge is monochromatic by the Union Bound. Hence we have that

$$P(G \text{ has a monochromatic edge}) \leq m * 2^{1-k} < 1 \Leftrightarrow m < 2^{k-1}$$

This means that, since $m < 2^{k-1}$, the probability that G has a monochromatic edge is less than 1, so there must be at least one coloring of vertices in which there are no monochromatic edges at all. Therefore, $c_k \geq 2^{k-1}$ as wanted. □

With this next problem, we will see how we can use expectation and random variables together with the probabilistic method to prove properties about a specific type of graph.

2.4 Hamiltonian Paths

The question of whether we can walk around the graph and visit either all the edges or all the vertices has been around since the historical"Seven Bridges of Königsberg Problem" was solved by Eunhard Euler in 1736. This problem focused on the question of whether one could traverse all the bridges of Königsberg without repeating any on the way, a question equivalent to traversing all the edges on a multigraph in what is know as a Eulerian path. A Hamiltonian path, on the other hand, seeks to traverse all the vertices.

Definition 2.4.1: Given a graph G, a Hamiltonian Path through G is a sequences of moves through edges of the graph in which we visit every vertex exactly once.

Now we will proceed to state and prove the last result in this section.

Proposition 2.4.2: Given $n \geq 1$, there exists a complete directed graph G on n vertices that has at least $\frac{n!}{2^{n-1}}$ Hamiltonian paths.

Proof. We are going to start with a graph G on n vertices and then we are going to randomly orient every edge one way or another with probability 1/2. Each possible Hamiltonian Path in the graph can be seen as a permutation of the numbers $1, 2, \ldots, n$, since it must pass by every vertex exactly once and we know that there is an edge between every pair of vertices. The question, however, is if a given permutation σ_k represents a valid Hamiltonian Path in G after randomly orienting the edges.

We are now going to introduce, for each possible Hamiltonian Path as represented by a permutation σ, an indicator random variable $\mathbf{H}_\sigma$ that will show if the permutation σ is a valid Hamiltonian Path. This means that, in our probability space $(\Omega, \mathbf{P})$ where each elementary event represents an orientation of G, our random variable $\mathbf{H}_\sigma$ will equal 1 in the events in which σ is a valid path and 0 in those in which it doesn't.

Because of the definition of expectation, we have that

$$\mathbb{E}[\mathbf{H}_\sigma] = \sum_{\omega \in \Omega} \mathbf{H}_\sigma p(\omega) = \sum_{\sigma \text{ valid in } \omega} p(w)$$

Now, we want to count the number of cases in which a permutation σ will represents valid Hamiltonian path in G. To do so, we can see that for every pair of vertices that are adjacent in σ, there is a $1/2$ chance that the edge will be oriented in the direction that will permit it being valid. Since there's $n-1$ of these edges that need to be oriented correctly, out of the $2^{\binom{n}{2}}$ possible orientations of G, we have that the probability one of them will have σ as a valid path is $1/2^{n-1}$. Hence, we have that

$$\mathbb{E}[\mathbf{H}_\sigma] = \sum_{\sigma \text{ valid in } \omega} p(w) = \frac{1}{2^{n-1}}$$

We have established the expectation that a given Hamiltonian path will appear in an orientation of G is $1/2^{n-1}$. Since there's $n!$ possible permutations, each with their associated indicator random variable, we have that the expected number of Hamiltonian paths $\mathbf{H}$ is such that

$$\mathbb{E}[\mathbf{H}] = \mathbb{E}\left[\sum_{\sigma \text{ a permutation}} \mathbf{H}_\sigma\right] = \sum_{\sigma \text{ a permutation}} \mathbb{E}[\mathbf{H}_\sigma] = \frac{n!}{2^{n-1}}$$

using linearity of expectation. What this means is that, *roughly*, the average number of Hamiltonian paths that we can expect in an orientation of G is $n!/2^{n-1}$. This, more importantly, means that at least one orientation of G must have at least $\mathbb{E}[\mathbf{H}] = \frac{n!}{2^{n-1}}$ Hamiltonian paths, which is what we wanted to prove. □

3 The Lovasz Local Lemma

In this section, we will state and prove two versions of a lemma that will allow us to use the Probabilistic Method in more complicated problems and in problems where we use the method in particularly clever ways. The main technique outlined in this section will be the Lovasz Local Lemma.

3.1 The General Lovasz Local Lemma

The core idea behind the Lovasz Local Lemma is that we want to look at the probability that the intersection of the complements of a set of events is nonzero. In this case, the events will probably be dependent on each other to a degree. However, the fact that we take this into account is part of the reason we arrive at better bounds and stronger results than in the previous sections.

When using the probabilistic method in the past couple of problems, we have looked at the simplest applications of probability (like the Union Bound and Linearity of Expectation) to produce our results. While useful, these techniques are in no way refined. Here, by setting our events to be the situations in which what we don't want to happen happens (the *bad* events), we will use the lemma to show that there is at least one case in which none of these events happen. To do so, we will first prove a General form of the Lemma, after which we will show the Symmetric form.

Theorem 3.1.1 (The General Lovasz Local Lemma): In a probability space $(\Omega, \mathbf{P})$, let I be a set of labels such that we have events A_i for $i \in I$. Define I_k to be the set of labels such that the event A_k is not independent of A_j if and only if $j \in I_k$. Suppose there exist numbers $0 < p_k < 1$ that satisfy the condition

$$P(A_k) \leq p_k \prod_{i \in I_k} (1 - p_i)$$

Then

$$P(\bigcap_{i\in I}\overline{A_i}) \geq \prod_{i\in I}(1-p_i) > 0$$

Proof. In order to prove the Lovasz Local Lemma, we will first prove a related statement through induction. What we will show is that, given a set $S \subset I$ we must have that

$$P(A_k|\bigcap_{i\in S}\overline{A_i}) \leq p_k$$

We will proceed with induction on the size of S.

Base Case: $|S| = 0$ In this case, we have, from the assumptions of the lemma, that

$$P(A_k|\bigcap_{i\in S}\overline{A_i}) = P(A_k) \leq p_k \prod_{i\in I_k}(1-p_i) \leq p_k$$

Hypothesis of Induction: We will assume that, for all sets S' such that $|S'| < |S|$, the statement in the lemma holds.

Inductive Step: We will show that the result holds for S. To do so, we will first consider decomposing the set S into two sets, S_1 and S_2, such that $S_1 = I_k \cap S$ and $S_2 = S/S_1$. If $S_1 = \emptyset$, that means that all the events A_i with labels $i \in S$ are independent from A_k. This implies that

$$P(A_k|\bigcap_{i\in S}\overline{A_i}) = P(A_k) \leq p_k \prod_{i\in I_k}(1-p_i) \leq p_k$$

as wanted. If $|S_1| \geq 1$, then we must have $|S_2| < |S|$ so that our inductive hypothesis holds for S_2. Note that

$$P(A_k|\bigcap_{i\in S}\overline{A_i}) = \frac{P(A_k \bigcap_{i\in S_1}\overline{A_i}|\bigcap_{j\in S_2}\overline{A_j})}{P(\bigcap_{i\in S_1}\overline{A_i}|\bigcap_{j\in S_2}\overline{A_j})}$$

Let us now find bounds on the numerator and denominator of this fraction in order to show that the probability we want is less than p_k. First lets look at the numerator. Because A_k is independent of all the events A_j with $j \in S_2$ by definition, we have that

$$P(A_k \bigcap_{i\in S_1}\overline{A_i}|\bigcap_{j\in S_2}\overline{A_j}) \leq P(A_k)|\bigcap_{j\in S_2}\overline{A_j}) = P(A_k) \leq p_k \prod_{i\in I_k}(1-p_i)$$

Now we will look at the denominator. If we think of the labels in S_1 as $\{i_1, i_2, \ldots, i_{|S_1|}\}$ We can express it as

$$P(\bigcap_{i\in S_1}\overline{A_i}|\bigcap_{j\in S_2}\overline{A_j})$$

$$= P(\overline{A_{i_1}}|\bigcap_{j\in S_2}\overline{A_j}) * P(\overline{A_{i_2}}|\overline{A_{i_1}} \cap \bigcap_{j\in S_2}\overline{A_j}) * \ldots * P(\overline{A_{i_{|S_1|}}}|\bigcap_{m<|S_1|}\overline{A_{i_m}}\bigcap_{j\in S_2}\overline{A_j})$$

From our hypothesis of induction, and since we saw that $|S_1| > 0$, we have that each of the terms on the right-hand side of the expression is $\geq (1-p_{i_m})$, so that our entire expression is such that

$$P(\bigcap_{i\in S_1}\overline{A_i}|\bigcap_{j\in S_2}\overline{A_j})\geq\prod_{i\in S_1}(1-p_i)\geq\prod_{i\in I_k}(1-p_i)$$

Combining both bounds, we will have that

$$P(A_k|\bigcap_{i\in S}\overline{A_i})=\frac{P(A_k\bigcap_{i\in S_1}\overline{A_i}|\bigcap_{j\in S_2}\overline{A_j})}{P(\bigcap_{i\in S_1}\overline{A_i}|\bigcap_{j\in S_2}\overline{A_j})}$$

$$\leq\frac{p_k\prod_{i\in I_k}(1-p_i)}{\prod_{i\in I_k}(1-p_i)}=p_k$$

as wanted. This concludes the proof of our small lemma.

We will now use what we just showed in order to prove the General Lovasz Local Lemma. To see this, it suffices to notice that

$$P(\bigcap_{i\in I}\overline{A_i})$$

$$=P(\overline{A_1})*P(\overline{A_2}|\overline{A_1})*\ldots*P(\overline{A_{|I|}}|\bigcap_{m<|I|}\overline{A_m})$$

$$\geq\prod_{i\in I}(1-p_i)>0$$

This concludes our proof of the General Lovasz Local Lemma □

3.2 The Symmetric Lovasz Local Lemma

In most cases in which the General Lovasz Local Lemma can be used to solve a problem, the Symmetric form of the lemma, which is simpler to use, suffices. Thus, it is of worth to note and proof the Symmetric Lovasz Local Lemma, since it is what we will be using in our example of how to use this new machinery alongside the Probabilistic Method.

Theorem 3.2.1 (The Symmetric Lovasz Local Lemma): In a probability space $(\Omega,\mathbf{P})$, let I be a set of labels such that we have events A_i for $i\in I$ such that:

1. Each A_i is independent of the rest except of at most d other events.

2. $P(A_i)\leq p$ for every $i\in I$.

Then, if $ep(d+1)\leq 1$, we must have that

$$P(\bigcap_{i\in I}\overline{A_i})>0$$

Proof. We will use the General Local Lemma in order to prove the Symmetric form. By setting $p_k=\frac{1}{d+1}$ in the general form, we have that

$$p_k\prod_{i\in I_k}(1-p_i)\geq\frac{1}{d+1}(1-\frac{1}{d+1})^d\geq\frac{1}{d+1}e^{-\frac{d}{d+1}}\geq\frac{1}{(d+1)e}\geq p\geq P(A_k)$$

This means that, by our choice of p_k, we have that our conditions for the General Lemma to apply are satisfied. So, we must have that

$$P(\bigcap_{i\in I}\overline{A_i}) \geq \prod_{i\in I}(1-p_i) = (1-\frac{d}{d+1})^{|I|} > 0$$

And our Symmetric Lovasz Local Lemma is proven. □

Now we will examine an application of the Lemma in a problem about cycles in a graph.

3.3 Cycles in a Directed Graph

In this problem, we will examine conditions in which we can find cycles in a directed graph of length divisible by a given k. This will establish a surprising link between graphs and number theory.

Proposition 3.3.1: Suppose we have a directed graph G on n vertices with minimum outdegree α and maximum indegree β. Then, for $k \in \mathbb{N}$ such that

$$k \leq \frac{\alpha}{1+ln(1+\alpha\beta)}$$

there must exist a cycle in G of length which is divisible by k.

Proof. We will begin by creating a graph G' such that the outdegree of every vertex is exactly α. We can do this by just erasing edges from the vertices with outdegree greater than α, and it's clear that if we prove the desired property on our new graph G', it must also be true of G.

We will now randomly color each vertex one of k colors as represented by a number from 1 to k. Now, we will define our events A_k as the event that, for a vertex v, none the vertices dependent on v are color $f(v)+1$ $(mod\ k)$, where $f(v)$ is the color of vertex v.

Now we want to see what events A_i are not independent of an event A_k. The immediate ones that pop out are the events A_i which relate to the vertices that are dependent on v. Other ones that we have to consider are those vertices w that share dependent vertices wit v, since they both force conditions on the same vertex. Of note, however, is that vertices that have v as a dependent vertex are not necessarily dependent events with their coloring.

With these considerations, we can affirm that each event A_k is dependent of at most $\alpha + (\beta-1)\alpha = \alpha\beta$ others. Since each event happens with probability $(1-\frac{1}{k})^{\alpha}$, we have that

$$e(1-\frac{1}{k})^{\alpha}(\alpha\beta+1) \leq e^{1-\frac{\alpha}{k}}(\alpha\beta+1) \leq e^{-ln(\alpha\beta+1)}(\alpha\beta+1) = 1$$

So that the conditions of the Symmetric Lovasz Local Lemma are satisfied and the intersection of the complements of the events will have probability greater than zero.

So, we have proven that there is a coloring in which there is always a vertex w dependent on v of color $f(v)+1(mod\ k)$. Now, what we will do is start on any vertex and randomly walk around the graph. Since there is only a finite number of vertices, I must eventually repeat a vertex. Hence, there is a cycle in G'. Because we are going down the list of colors with every step and we have arrived at a vetex we passed, we must have gone through the list an integer number of times. This means the length of the cycle must be divisible by k, which shows that there exists a cycle in G which is length multiple of k as wanted. □

4 Further Work

The probabilistic method has allowed us to provide bounds for problems in algorithms and graph theory. However, one of the issues with using the probabilistic method to prove the existence of combinatorial objects that fulfill certain properties is that this method is non-constructive. In a way, the fact that it is non-constructive is part of the reason it works so well - we don't need to know how to find this object, we just do a survey over all possible objects to determine that at least one must exist that satisfies our needs.

This problem was addressed and solved by Robin Moser and Gábor Tardos with the introduction of an algorithmic version of the Lovasz Local Lemma, which provides a method to construct objects that are show to exist using the lemma. This way, we can not only survey for the existence of an object but, under certain conditions, actually show an example of an object that satisfies the desired property.

This version of the lemma can be used to show the satisfiability of CNF-formulas (logic formulas that are in conjunctive-normal form) that are within certain bounds in their literals and clauses. Furthermore, using the algorithm, we can find the solution that satisfies the CNF in polynomial time.

Proposition 4.1.1: Let Φ be a CNF formula on n variables with n clauses and at least k literals in each clause. If each variable appears in at most $\frac{2^k}{ke}$ clauses, then Φ is satisfiable and a solution can be found in polynomial time.

Using the Symmetric Local Lovasz Lemma, we can prove the satisfiability of Φ, and with the Algorithmic version we can actually provide a construction of the solution.

This concludes our treatment of the Probabilistic Method.

5 References

[1] Alon, N; Spencer, J. (2000). *The Probabilistic Method.* New York: Wiley-Interscience.

[2] Havet, F. (2011). *Introduction to the Probabilistic Method.* INRIA Sophia Antipolis: Projet Mascotte

[3] Matousek, J; Vonrak, J. (2008). *The Probabilistic Method: Lecture Notes.* Charles University: Department of Applied Mathematics.

[4] Moser, R. (2008). *A constructive proof of the Lovasz Local Lemma.* ETH Zurich: Institute for Theoretical Computer Science.