CS5130 : Mathematical Tools for Theoretical Computer Science

(Scribe Lecture Notes)

Lecturer : JAYALAL SARMA

Department of Computer Science and Engineering
Indian Institute of Technology Madras (IITM)
Chennai, India

Last updated on : September 25, 2020

Preface

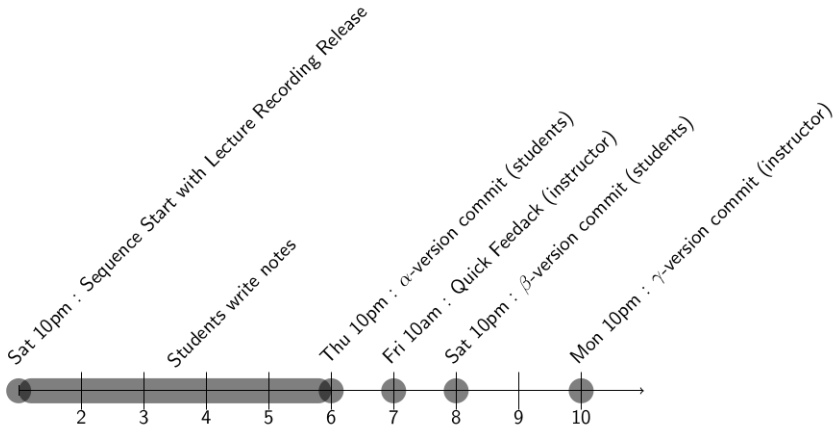
This lecture notes are produced as a part of the course CS5130: Mathematical Tools for Theoretical Computer Science which was a course offered (during the online semester Sep-Dec 2020) at the CSE Department of IIT Madras.

Acknowledgements

We acknowledge the efforts of the scribes and editors of this document.

Scribe status

Each lecture has a field called **status**. It tells which stage of the edit pipeline is the document currently. The scribe notes are due on Thursdays and Saturdays as per the following timeline.



Even after these edits, it is possible that there are still errors in the draft, which may not get noticed. If you find errors still, please report to the instructor.

List of Scribes

Table of Contents

Todo list

Instructor : Jayalal Sarma
Scribe : Jayalal Sarma (*TA:* JS)
Date : Sep 9, 2020
Status : α

Lecture
1

Pigeon Hole Principle and Basic Applications

We start course with the simplest but surprising powerful tool in combinatorial arugments which is the pigeon hole principle. Through this principle as an example, we will also quick review the methods of proof.

1.1 Quick Recap on Proof Techniques

A formal mathematical proof system in our context has axioms about various mathematical objects that we are using, like numbers, graphs which describes them through their properties. Then, there are rules of inferences such as modus ponens, modus tollens, resolution, syllogisms etc which helps us derive new statements from these axioms.

The peculiarity of these rules of inferences are that they "conduct truth" and forms building blocks for huge "truth conducting" structures called mathematical proofs. That is, if for any object[1], the premises of the rules of inference are true, then the conclusion is also true for them. Hence, suppose we derive a statement ϕ starting with the axioms, applying the rules of inferences in various combinations. Since the individual rules of inferences "conduct truth", the resulting structure also conducts truth and is called the mathematical proof of the statement ϕ from the axioms. Note that the truth of the statement ϕ for the object under consideration can be stated on relative to the truth of the axioms that we used. However, this is not a concern, since we are intending to use the mathematical proof systems to derive statements about objects which we know would satisfy the axioms (in fact, we wrote down axioms as properties of those objects.

Curiosity 1.1.1. It is an amusing question to ask, whether there are other objects, which we did not intend to, which also satisfies the axioms that we wrote, by accident. Say for example, we wrote the axioms for graphs, but "strings" also satisfies them. If so, the theorems that we prove for graphs using only those axioms will also be true for strings, automatically !!. Quite interestingly this is true for natural numbers. The mathematical theory of natural numbers is axiomatized by what are called the Peano's axioms. There are numbers that one can define which are different from natural numbers for which any theorem that we prove for natural numbers also are true

[1] a little more formally, the assignment in the propositional logic, and model in general first oder logic

(because they satisfy the Peano's axioms). Then one might ask, are we not trying to represent exactly natural numbers? So should we not augment Peano's axioms with more properties of natural numbers such that we remove such *unwanted* parallel models from satisfying the axioms we write. Even more interestingly, one can argue that this is not even possible. No matter, what extra formula we write the existence of such "parallel models" us inevitable. In fact, not just one "parallel model", there will be infintiely many of them. You should read about *Löwenheim–Skolem theorem*.

Writing down mathematical proofs explicitly by using rules of inference may seem to be a mechanical way of proving statements. While it avoids any chance of mistakes because of the mathematical precision and rigor it affects quick readability and communication of ideas. Hence, one would like to have more "human readable" ways of representing these proofs by writing some of the steps in English, while ensuring that we do not lose the mathematical rigor. This brings in some subjectivity about how "formal" a proof is - that is, how close is it to the formal mathematical framework of rules of inferences in terms of notations, presentation etc. Sometimes, very rigorous proofs tend to hide the intuitive idea behind the proof which one tends to (and sometimes need to) describe separately for easy communication. The more formal your proof is, the less chances of you making a logical error in the proof. It is a good idea to start writing proofs with the mindset of "rigor extremist" and once you are comfortable and see through the mathematically rigorous steps of a statement, you can rely more in English sentences. This course particularly would do it in the latter way, but ensuring that mathematical rigor is kept in tact. The beauty of the combinatorial proofs lies in the elegance and the combinatorial insight and intuition. Balancing the intuition with rigor in presentations and descriptions lies in the art of presentations.

Suppose that we have to prove a statement γ of the form $p \to q$. We quickly recall the different ways of proof in the above described form.

Direct Proof: Assume p and then derive q using the assumption and the axioms by applying the rules of inferences. The is considered as a proof of the statement $p \Rightarrow q$ since it can be associated with a valid argument form by itself.

Indirect Proof: Assume $\neg q$ and then derive $\neg p$. Again, this is also considered as a proof of the statement $p \Rightarrow q$ since it can be associated with a valid argument form by itself. This is also called proof by *contrapositive*.

Proof by Contradiction: A proof by contradiction, assumes the negation of the statement to be proven (that is, $\neg\gamma$) and then defined a statement r (this forms a part of creativitiy of the proof), and then derives $r \wedge (\neg r)$ from the assumption and axioms using the rules of inferences. By an associated valid argument form, this shows that γ must be true, again, by associating the definition of a valid argument form.

In addition, while proving quantified statements, there are a few additional ideas that are used which we quickly review below:

Proof by Exhaustive Cases: Suppose we want to derive a statement Γ of the form $\forall \alpha P(\alpha)$ where α comes from domain of discourse $\mathcal{D}$ (say, for example, α is a natural number, that is, $\mathcal{D} = \mathbb{N}$). We can partition $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2 \ldots \cup \mathcal{D}_k$ into several subdomains and prove the statement $\forall \alpha \in \mathcal{D}_i,\ P(\alpha)$ separately. Each part of the proof $\forall \alpha \in \mathcal{D}_i P(\alpha)$ is said to be a "case" of the proof. The fact that, $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_1 \ldots \cup \mathcal{D}_k$ is what is meant by the statement that the case analysis is *exhaustive*.

Proof by "Counter Example": Suppose we want to disprove statements of the form $\forall \alpha P(\alpha)$. That is, we want to derive $\neg\left(\forall \alpha P(\alpha)\right)$ which is logically equivalent to $\exists \alpha \neg P(\alpha)$. Hence it suffices to demonstrate an α in the domain for which we can show $P(\alpha)$ is false.

Proof by Mathematical Induction: This is a technique to prove statements of the form $\forall \alpha P(\alpha)$ where the domain $\mathcal{D}$ is countably infinite. That is, the domain $\mathcal{D}$ can be put in bijection with the set of natural numbers. The technique forms part of the Peano's axioms that define the natural numbers and hence is a valid proof technique. If $\phi : \mathbb{N} \to \mathcal{D}$ is a bijection, in order to prove $\forall \alpha P(\alpha)$, we can equivalently prove $\forall n \in \mathbb{N}, P(\phi(n))$. In particular, it takes the following form: *If we can prove* $P(\phi(0))$ *and the implication* $[\forall n \in \mathbb{N}, P(\phi(n)) \Rightarrow P(\phi(n+1)]$ *then we can conclude* $\forall n\ P(\phi(n))$. There are versions of this proof techniques such as strong induction, structural induction, spiral induction, double induction etc which are adaptations of the above basic idea.

Most of the proofs that we do in the courses will follow one of the above frameworks. We will not do examples of these techniques since that is already covered in the basic discrete mathematics course.

1.2 The Pigeon Hole Principle (PHP)

With the quick recap done in the previous part, we now plunge into the actual business in this lecture. We first prove the following basic version of the Pigeon hole principle.

THEOREM 1.2.1. *Let $n, k \in \mathbb{N}$, such that $n > k$. Suppose we place n identical balls in k identical bins, then there is a bin that has at least two balls in it.*

Proof. Let $n, k \in \mathbb{N}$ and $n > k$. Assume for the sake of contradiction that when we placed the balls into the bins as indicated in the theorem, there was no bin with at least two balls in it.

As such the bins are identical, but number them from 1 to k now. Using this notation, let us define b_i to be the number of balls that went into the bin number i. Clearly $\forall i, b_i \geq 0$. Since we did distribute all the balls into the bins, we have :

$$\mathcal{R} : \sum_{i=1}^{k} b_i = n$$

Using the assumption, we have that: $\forall i, 0 \leq b_i \leq 1$. Summing up for i: $\sum_{i=1}^{k} b_i \leq \sum_{i=1}^{k} 1 = k < n$. Hence we have derived the statement :

$$\neg\mathcal{R} : \sum_{i=1}^{k} b_i \neq n$$

Hence we have derived $\mathcal{R} \wedge \neg\mathcal{R}$. This is a contradiction and hence the original assumption that we started out with must be false and hence there has to exist a bin which has two balls in it. □

Curiosity 1.2.2. The formal proof of PHP as simple as it sounds is still a subject of substantial research in an area called *proof complexity*. To demonstrate this, let us write the principle itself in more rigorous notations. Let $n > k$, and $\{x_{ij} \mid i \in [n], j \in [k]\}$ be propositional variables (which can be called, say *pigeon hole variables*). Following our original notation, where there are n pigeons and k holes, the basic Pigeon Hole Principle is the following Disjunctive normal form formula :

$$\text{PHP}_k^n \stackrel{\text{def}}{=} \left(\bigvee_{i \in [n]} \bigwedge_{j \in [k]} \overline{x_{ij}} \right) \vee \left(\bigvee_{j \in [k]} \bigvee_{r \neq s \in [n]} (x_{rj} \wedge x_{sj}) \right)$$

To prove this, one possibility is to derive the contradiction from the negation of PHP_k^n. This is an expression in conjunctive normal form, with clauses:

$$\text{For } i \in [n] \text{ the clauses : } Q_i \stackrel{\text{def}}{=} \bigvee_{j=1}^{k} x_{ij}$$

$$\text{and for } s \neq t \in [n], j \in [k] \text{ the clauses } Q_{s,t,j} \stackrel{\text{def}}{=} \overline{x_{sj}} \vee \overline{x_{tj}}$$

Intuitively, these say that there is a function from $[n] \to [k]$ (which is represented by $x_{ij} = 1$ to mean that the function takes i to j) which is well defined (for every i, there exists a j such that $x_{ij} = 1$) and also injective (for two different s and t, it is not the case that x_{si} is 1 and x_{tj}). Since $n > k$, there cannot be an injection, and hence the negation of the conjunction of these clauses PHP_k^n must be true.

Suppose we ask, starting from these clauses as axioms, and applying rules of inferences (say the resolution principle) alone, how many steps of proof does one need to do to derive the contradiction ($r \wedge \neg r$ for some r). [2]. We measure this in terms of n and k which determines the number of variables in the system. The area which studies the complexity of proofs in the above is called *proof complexity theory*. It turns out the the basic PHP itself is one of the tautologies for which one requires exponentially long proofs if we are restricting ourselves to resolution? What if we relax this? The area has several interesting open questions related to this and they have close connections to computational complexity theory too.

[2]Notice that this sounds exactly like computation, how many steps of computation is required in order to certain tasks in terms of input parameters

1.2.1 A Quick Example:

We will now demonstrate the application of the principle itself by a quick example. This is meant to be a revision of the topic from the previous courses and hence it is very much possible that you have seen the application earlier.

THEOREM 1.2.3. *If you consider any five points placed inside the unit square then there must necessarily exist two points are at most* 0.75 *unit away from each other.*

Proof. Firstly, to make it sound less magical, let us comment that theorem is actually true for 0.75 units replaced by 0.707 units which is actually $\frac{1}{\sqrt{2}}$. The application of PHP goes as follows. Consider four small squares which are obtained by the midpoint of the square as one of the corners. These small squares form the bins and the five points that we place forms the balls. By applying PHP, we conclude that there must be two points which falls into the same small square. Now the argument can be completed by the fact that the maximum distance between any two points which are in the same small square is at most $\frac{1}{\sqrt{2}}$ since the sides of the square are $\frac{1}{2}$ each. □

REMARK 1.2.4 (**Tightness)**. Is the above theorem tight? Can it be improved? Improvement can be in terms of two parameters. Firstly, can we make the same claim for 4 points? Secondly, even for 5 points, can we make an improved claim about the minimum distance being, say 0.7 units? The answer to both these questions are no. For the first, we can demonstrate 4 points in which every pair is at least one distance away - the four corners themselves will serve as a counter example. For the second question, we can demonstrate 5 points which are actually only pairwise at least $\frac{1}{\sqrt{2}}$ distance away.

REMARK 1.2.5 (**Glimpse of Extremals in Combinatorics)**. The above example theorem, while is a classical application of Pigeon Hole Principle, it also demonstrates a curious phenomenon. In spirit it says that *if there are large number of objects in a collection, then there must be some structure*. Question is how large? And what is structure? The answers to these vary and forms the foundations of this area. We will see more of this when we see Ramsey Theory.

1.3 Numbers and Remainders

It is customary to do an example of PHP from numbers and division under remainders. We will do a slightly unusual example.

THEOREM 1.3.1. *Consider the infinite sequence* $7, 77, 777, \ldots, 7777777, \ldots$ *- there must necessarily exist a number in this sequence that is divisible by 2003.*

Proof. As weird as it sounds, one might wonder how does PHP play a role. There does not seem to be any place to apply PHP directly in the statement of the problem. Indeed, infinitude seems to indicate that we are allowed to take large numbers in the sequence. A usual trick is the division,

and then consider the remainders.

As a start, consider first 2003 numbers in the sequence. Denote them by $n_1, n_2, \ldots n_{2003}$. Divide them by 2003 and collect the remainders that we see. Denote them by $a_1, a_2, \ldots, a_{2003}$. If any of the a_is are 0, then we are done since that Indeed, we have that $1 \leq a_i \leq 2002$. Clearly, now the pigeons and holes are visible now. The numbers n_is are the pigeons and the reminders are the holes. There are only 2002 holes but there are 2003 pigeons and hence by PHP, there must exists $1 \leq i < j \leq 2003$ such that $a_i = a_j$. This gives:

$$n_i \mod 2003 = n_j \mod 2003 \tag{1.1}$$

$$(n_i - n_j) \mod 2003 = 0 \tag{1.2}$$

$$2003 \text{ divides } (n_i - n_j) \tag{1.3}$$

$$\tag{1.4}$$

That is good progress. We managed to show 2003 divides $(n_i - n_j)$. However, $n_i - n_j$ unfortunately, will not be in the sequence at all. How will this number look like? By the structure of the numbers, this difference will be a number of 7s and then several zeros. More precisely computing these number, we have that:

$$(n_i - n_j) = n_{j-i}10^{j-i}$$

So we have that 2003 divides the product of n_{j-1} and 10^{j-i}. However, 2003 being an odd number which is not a multiple of 5 will not have a common factor with any power of 10. Hence 2003 must necessarily divide n_{j-i} which should be there in the sequence. This completes the proof. □

1.4 Graphs

Our third application is related to problems that can be modelled as graphs.

THEOREM 1.4.1. *In any chess tournament, where there are n participants, at any point of time there must be two participants who finished the same number of games in the tournament.*

It is natural to model this situation as a graph with n vertices where each vertex represents a participant and we put an edge between two vertices if player i and player j have played a game with each other. The number of games played by a player is exactly the degree of the vertex in this graph. Rewriting the above theorem in the new language now:

THEOREM 1.4.2. *In any undirected graph G, there must be two vertices which are having the same degree.*

Proof. The proof is by an exhaustive case analysis. We need to argue the above for all graphs. We divide this domain into two based on whether there is an isolated vertex or not.

Case 1 : G has an isolated vertex - In this case, there is a vertex of degree 0, and hence there cannot be a vertex of degree $n - 1$. Thus we have n vertices, and only $n - 1$ possible degree

values $\{10, , 2, \ldots n-2\}$. By the PHP, we must see two vertices which has the same degree.

Case 2: G does not have an isolated vertex - In this case, there is no vertex of degree 0, and hence the degree values of vertices can only be in the set $\{1, 2, \ldots n-1\}$. Again we have n vertices whose degrees take only $n-1$ possible values. Again, by PHP, we must see two vertices having the same degree.

□

Exercise 1.5 (See Problem Set 1 (Problem 1))**.** A social network is said to be symmetric if the relation between users that is maintained as a part of the network, is symmetric. Consider a symmetric social network and let the symmetric relation maintained be that of "user A and B are *friends*" (like in the case of facebook). A user C is said to be a *mutual friend* of users A and B if, C is a friend of both A and B. Prove that - for any user A of the network who has at least two friends, there must exist two friends of A who has the same number of mutual friends with A.

Comment on whether symmetry is critical for your argument. Take the example of *instagram* where the symmetric relation of *friends* is replaced by *followers*. Generalize the definition of mutual friends to *mutual followers*. Comment on whether a similar statement for followers can be established in this case.

1.6 Discussion Session

Instructor : Jayalal Sarma
Scribe : Narasimha Sai Vempati (*TA:* JS)
Date : Sept 19, 2020
Status : α

Lecture
5

Multichoosing

5.1 Introduction

Consider the definition of *set*. We know that it's a well defined collection of *distinct* objects. From a collection of n distinct symbols, the number of ways to form a *set* of length k is given by $\binom{n}{k}$. Now let's consider the definition of *multi-set*. It's similar to that of a *set*, except that it allows repetition of objects. Now it's natural ask the following question: From a collection of n distinct symbols, what is the number of ways to form a *multi-set* of length k. Multichoosing exactly answers this questions. In this lecture, we explore multichoosing in detail. We discuss several equivalent bijections to this problem and come-up with an algebraic expression for $\left(\binom{n}{k}\right)$ (spelled out as n *multi-choose* k).

5.2 Equivalent bijections

5.2.1 Non-negative solutions

Formally, $\left(\binom{n}{k}\right)$ is the number of ways of choosing k objects from a set of n objects where the order is not important but repetitions are allowed. For all $i = 1, 2, \cdots, n$, if we denote by x_i the number of copies of i^{th} object we choose, then we have the equation

$$x_1 + x_2 + \cdots + x_n = k \tag{5.5}$$

where each $x_i \geq 0$. Therefore, number of *non-negative* integral solutions to this equation gives us the required number of ways of choosing k objects from n objects with given conditions. Let's look at an equivalent problem and establish a bijection between these two.

5.2.2 Voting problem

If n candidates are contesting in an election and there are k voters, how many ways can votes of those k voters be distributed among n candidates?

If we denote by x_i, the number of votes received by i^{th} candidate and there are k voters, we have $x_1 + x_2 + \cdots + x_n = k$ and thus, the number of ways of dividing votes among candidates is the number of non-negative solutions to the equation 5.5. Formally, we can define a bijection f from set of solutions to the equation 5.5 to set of ways of dividing the votes among n candidates.

Definition: f takes the tuple $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ and assign x_i number of votes to i^{th} candidate where $i = 1, 2, \cdots, n$.

Well defined: f is well defined because for every valid tuple $\mathbf{x} = (x_1, x_2, \cdots, x_n)$, we have $x_1 + \cdots + x_n = k$ and thus summing over votes received by i^{th} where $i = 1, 2, \cdots, n$ will be k votes in total.

Injective: f is an injection because for every valid way of dividing the votes among candidates, there's a unique solution tuple in which x_i = number of votes received by i^{th} candidate. In other words, for any two $\mathbf{x}_1 \neq \mathbf{x}_2$, there exists an $i \in [n]$ such that $x_{1_i} \neq x_{2_i}$ and i^{th} candidate gets different votes. Thus $f(\mathbf{x}_1) \neq f(\mathbf{x}_2)$.

Surjective: f is surjective because for every way of dividing k votes among n candidates, there is a pre-image $\mathbf{x} = (x_1, \cdots, x_n)$ which is a valid solution to the equation 5.5 (as there are a total of k voters, sum of number of votes received by each voter must sum up to k).

Thus f is a bijection from the set of non-negative solutions to $x_1 + \cdots + x_n = k$ to the set of ways of dividing k votes among n candidates.

5.2.3 Non-decreasing subsequences

Number of non-decreasing sequences of integers between 1 and n of length k. A non-decreasing sequence is of the form $\{a_1, a_2, \cdots, a_k\}$ where $1 \leq a_1 \leq a_2 \cdots \leq a_k \leq n$. Lets define a bijection f from set of non-negative integral solutions to Eqn. 5.5 to set of non-decreasing sequences between 1 and n of length k.

Definition: f takes $\mathbf{x} = (x_1, \cdots, x_n)$ as input and writes the number i x_i times for all $i = 1, 2, \cdots, n$ to obtain a sequence of length k.

Well defined: As f constructs the sequence in increasing order from 1 to n by writing i x_i times, the resulting sequence will be non-decreasing. Therefore, f is well defined.

Injective: For every $\mathbf{x}_1 \neq \mathbf{x}_2$, there exists an i such that $x_{1_i} \neq x_{2_i}$ and thus in the resulting sequences, number i is written different number of times. Therefore, f is injective.

Surjective: Every non-decreasing sequence of integers between 1 and n of length k has a pre-image $\mathbf{x} = (x_1, \cdots, x_n)$ which is a valid solution to equation 5.5 (where x_i is the number of times the number i is present in the sequence and as length of sequence is k, all x_i's where $i = 1, 2, \cdots, n$ sum up to k).

Thus f is a bijection.

5.2.4 Stars and bars problem

There are k stars placed horizontally. Find the number of ways to place $n-1$ bars in between those k stars. Lets define a bijection f from set of non-negative integral solutions to Eqn. 5.5 to set of ways of placing $n-1$ bars among k stars.

Definition: f takes $\mathbf{x} = (x_1, \cdots, x_n)$ as input and place x_i number of stars between $(i-1)^{th}$ bar and i^{th} bar. We leave it as an exercise to prove that f is well-defined, injective and surjective.

1: Jayalal says: Todo - Prove that f is a bijection

5.3 Algebraic expression

So far in Sec. 5.2, we have established bijections between *non-negatives integral* solutions of Eq. 5.5 and various other problems and argued that number of ways of solving any particular problem is equal to the number of non-negative integral solutions to Eq. 5.5. In this section, we are interested in coming up with a concrete expression for $\left(\binom{n}{k}\right)$ by solving it's equivalent bijection.

Method 1 Let's solve the *stars and bars* problem defined in Sec. 5.2.4. Let's use the fact that any placement of $n-1$ bars among k stars can be equivalently thought of as a string of length $n+k-1$ over the alphabet $\{\star, |\}$ with k $\star$'s. Therefore,

$$\begin{aligned}\text{number of ways of placing } n-1 \text{ bars among } k \text{ stars } &= \text{number of such strings} \\ &= \binom{n+k-1}{k}\end{aligned}$$

Method 2 Let's solve the *Non-decreasing subsequences* problem defined in Sec. 5.2.1. Let's establish a bijection f from set β of non-decreasing subsequences of integers between 1 and n of length k to a set Γ of strictly increasing subsequences of integers between 1 and $n+k-1$ of length k. A strictly increasing subsequence is of the form $1 \leq b_1 < b_2 < \cdots < b_k \leq n+k-1$

Definition: f takes as input a non-decreasing subsequence $(a_1, a_2, \cdots, a_k)$ between 1 and n and for all $i = 1, 2, \cdots, k$ set $b_i = a_i + i - 1$ and output the sequence $(b_1, b_2, \cdots, b_k)$

Well defined: For any $(a_1, a_2, \cdots, a_k) \in \beta$, we have for all $i = 1, 2, \cdots, k-1$,

$$\begin{aligned} a_i &\leq a_{i+1} \\ a_i + i &\leq a_{i+1} + i \\ a_i + i - 1 &< a_{i+1} + i \\ b_i &< b_{i+1} \end{aligned}$$

Therefore, the subsequence $(b_1, \cdots, b_k)$ is strictly increasing subsequence and thus f is well defined.

Injective: For every non-decreasing subsequence $(a_1, \cdots, a_k)$, there's a unique strictly increasing subsequence $(b_1, \cdots, b_k)$ where for all $i = 1, \cdots, k$, $b_i = a_i + i - 1$. Therefore f is injective.

Surjective: For every strictly increasing subsequence $(b_1, \cdots, b_k)$, there's a pre-image $(a_1, \cdots, a_k)$ which is non-decreasing where for all $i = 1, \cdots, k$, $a_i = b_i - i + 1$

Therefore, f is a bijection. The number of ways of choosing a strictly increasing subsequence $(b_1, \cdots, b_k)$ between integers 1 and $n+k-1$ is just choosing k integers from first $n+k-1$ integers and arrange them in one way(increasing order). Therefore number of ways = $\binom{n+k-1}{k}$. As f is a bijection, therefore, the number of non-decreasing subsequences between 1 and n of length k are $\binom{n+k-1}{k}$

Method 3 Let's solve the *Voting* problem defined in Sec. 5.2.2. Let's ask a slightly modified question.

Question: How many ways to distribute m votes among n candidates such that each candidate gets at least one vote.

Answer 1: As every candidate gets at least one vote, let's first distribute one vote each to each of the n candidate and the distribute the remaining $m - n$ votes among n candidates. By the bijection defined in Sec. 5.2.2, the number of ways of distributing $m - n$ votes among n candidates is $\left(\binom{n}{m-n}\right)$

Answer 2: Let's interpret votes as $\star$ s. Then the question essentially reduces to placing $n - 1$ bars (since there are n candidates, we divide by placing $n - 1$ bars) among m stars (since there are m voters). i^{th} candidate gets votes equal to number of stars between $(i-1)^{th}$ | and i^{th} |. However, there are two additional constraints

1. A bar cannot be placed in the beginning or in the end (if not then either the first candidate or the last candidate gets 0 votes)
2. We cannot place two | s between same two $\star$ s (if we place $(i-1)^{th}$ | and i^{th} | between same two $\star$ s, the i^{th} candidate gets 0 votes)

Hence, we have to choose $n - 1$ gaps among the $m - 1$ gaps (because we have $m + 1$ gaps and by cond. 1 we remove two) to place $n - 1$ | s without repetitions (because repeating violates cond. 2). Therefore, there are $\binom{m-1}{n-1}$ ways of doing it. Thus $\left(\binom{n}{m-n}\right)$=$\binom{m-1}{n-1}$ and by substituting $m = n + k$, we have

$$\left(\binom{n}{k}\right) = \binom{n+k-1}{n-1} = \binom{n+k-1}{k}$$

5.4 Identities

In this section, we discuss some identities on $\left(\binom{n}{k}\right)$ and argue their proofs using the idea of either double counting or bijections.

Identity 1

$$\left(\binom{n}{k}\right) = \left(\binom{k+1}{n-1}\right)$$

Proof. Let's use the bijection method to prove this. Formally, lets define sets S_1 and S_2 and count their cardinalities independently and then establish a bijection from S_1 to S_2 proving that $|S_1| = |S_2|$.

S_1: Configuration of k $\star$ s and $n-1$ | s as described in Sec. 5.2.4. By the bijection defined in it, $|S_1| = \left(\binom{n}{k}\right)$

S_2: Configuration of $n-1$ $\star$ s and k | s as described in Sec. 5.2.4. Again, by the bijection defined in it, $|S_2| = \left(\binom{k+1}{n-1}\right)$

Bijection: Let's define a bijection f from S_1 to S_2. f takes a configuration from S_1 as input and interpret $\star$ s as | s and | s as $\star$ s. Therefore it ends up with a configuration with $n-1$ $\star$ s and k | s which is a configuration is S_2. It's easy to observe that f is a bijection.

As f is a bijection from S_1 to S_2, we have $|S_1| = |S_2|$. This completes the proof ∎

Identity 2

$$k\left(\binom{n}{k}\right) = n\left(\binom{n+1}{k-1}\right)$$

Proof. Let's use the method of double counting to prove this.

Question: In how many ways can we construct a non-decreasing sequence $1 \leq a_1 \leq a_2 \cdots \leq a_k \leq n$ and mark one element?

Asnwer 1: By the bijection established in Sec. 5.2.3 we have $\left(\binom{n}{k}\right)$ number of non-decreasing subsequences and for every such subsequence, we can mark any one of the k elements choose. Thus the answer is $k\left(\binom{n}{k}\right)$

Answer 2: Firstly, determine the value in $[n]$ which is to be marked. Let r be this value. Now, consider a non-decreasing subsequence between 1 and $n+1$ with $k-1$ elements. Using r and the non-decreasing sequence chosen, we construct a unique non-decreasing sequence between 1 and n of length k with r as marked in the following way:

Let $(b_1, b_2, \cdots, b_{k-1})$ with $1 \leq b_1 \leq b_2 \leq \cdots \leq b_{k-1} \leq n+1$ be the chosen sequence,

- Insert marked-r in the right most position so that the resulting sequence is still sorted.

- As long as there's an $n+1$ in the sequence, remove it and add it as r to the right of marked-r in the sequence

Therefore, number of required sequences

$$\begin{aligned} &= \text{number of ways to choose } r \times \overset{\text{number of non-decreasing sequences of length } k-1 \text{ between 1 and } n+1}{\left(\!\binom{n+1}{k-1}\!\right)} \\ &= n \times \left(\!\binom{n+1}{k-1}\!\right) \end{aligned}$$

This completes the proof ∎

Exercise 5.5.

Prove the following by combinatorial arguments

$$\left(\!\binom{n}{k}\!\right) = \sum_{m=1}^{n} \left(\!\binom{m}{k-1}\!\right)$$

Hint: Look for bijection to number of non-decreasing subsequences

Prove the following by combinatorial arguments

$$\sum_{k=0}^{m} \left(\!\binom{n}{k}\!\right) = \left(\!\binom{n+1}{m}\!\right)$$

Hint: Look for bijection to Voting problem

Prove the following by combinatorial arguments

$$\left(\!\binom{n}{k}\!\right) = \sum_{m=0}^{n} \binom{n}{m} \left(\!\binom{m}{k-m}\!\right)$$

Instructor : Jayalal Sarma
Scribe : Anshu and Narasimha Sai (*TA:* JS)
Date : Sept 19, 2020
Status : α

Lecture
6

Catlan Bijections

6.1 Introduction

One of the classic examples to demonstrate the power of bijections is *Catlan numbers*. The Catlan numbers form a sequence of natural numbers that occur in various counting problems and occurs in several seemingly different contexts. Historically, *Euler* is the first person to study them. He was interested in counting the number of ways of dividing a polygon into triangles by drawing non-overlapping diagonals. Catlan numbers got their name from *Eugene Catlan* when he used them to answer the *Parenthesisation problem* which is the following: Consider a sequence $(a_1, a_2, \cdots, a_{n+1})$ of $n+1$ numbers, If we have to perform a binary operations $\odot$ n times among them, how many number of ways are there to parenthesise (or bracket) them using n parenthesis of single type (say $'()'$). In this lecture, we will see a few equivalent problems to this and then arrive at an explicit expression of Catlan numbers.

6.2 Equivalent Bijections

In this section, we see a few equivalent problems of the *parenthesisation* problem and argue that answer to each of them is also the *catlan number*

Full binary trees If we observe the Parenthesisation problem carefully, we notice that every valid parentesisation of those $n+1$ numbers form a *full binary tree* (a binary tree in which every node have either two children or no children) of $n+1$ leaves and n internal nodes where leaves represents the numbers $a_1, \cdots, a_{n+1}$ and each internal node corresponds to one operation. Therefore, there's an implicit bijection between the set of valid parenthesisations and full binary trees with n internal nodes. Therefore,

$$\begin{matrix}\text{number of valid parenthesisations of} \\ n+1 \text{ elements}\end{matrix} = \begin{matrix}\text{number of full binary trees with} \\ n \text{ internal nodes}\end{matrix} \tag{6.6}$$

Balanced parenthesised strings A balanced parenthesised string of length $2n$ is a string consists of n left brackets $'('$ and n right brackets $')'$ in which every prefix of the string has number of left brackets $'('$ $\geq$ number of right brackets $')'$. One can easily observe the bijection from set of balanced paranthesised string to valid parenthesisations of $n+1$ numbers

Euler's problem Find the number of ways of triangulating a polygon with $n+2$ edges

Handshaking problem Consider a scenario where $2n$ people are sitting around a table. How many ways they can shake hands with each other without crossing hands. We leave it as an exercise to establish bijections from *Euler's* problem to *Full binary tree* problem and *handshaking* problem to *balanced parenthesised strings* problem.

2: Jayalal says: Todo - Establish bijections from *Euler's* problem to *Full binary tree* problem and *handshaking* problem to *balanced parenthesised strings* problem

6.3 Algebraic Expression

In this section, we are interested in arriving at a concrete expression of the n^{th} *catlan number* (denoted by c_n). Let's solve another problem and then, by establishing a bijection to one of the above problems, we can arrive at an expression for c_n.

6.3.1 Monotone walk on $n \times n$ grid

Suppose we have a grid of size $n \times n$. How many ways are there to go from $(0,0)$ to (n,n) by using only downward edges or right edges. A sample path is represented in Fig. 6.1. We observe that each step can increment the value of exactly one of the co-ordinates by 1. Since we have to move from $(0,0)$ to (n,n), we have to increase the value of both the co-ordinates by n and n and thus irrespective of the path you take, the length of a path from $(0,0)$ to (n,n) must be of length $n+n=2n$.

If we represent each right move as R and each downward move as D, one can observe that there's a bijection f from the set of paths to set of strings of length $2n$ over the alphabet $\{D,R\}$ with number of D's = number of R's = n. Formally, if $(u_0,v_0),(u_1,v_1),\cdots,(u_{2n},v_{2n})$ represents the path where $(u_0,v_0)=(0,0)$ and $(u_{2n},v_{2n})=(n,n)$, and $b=b_1b_2\cdots b_{2n}$ represents the string where each b_i is either D or R, our bijection f takes a path as input and sets b_i as

$$b_i = \begin{cases} D & \text{if } u_i = u_{i-1}+1 \\ R & \text{if } v_i = v_{i-1}+1 \end{cases}$$

Well defined: As we have exactly n x co-ordinate increments and n y co-ordinate increments, we will have exactly n D's and n R's in our string and thus f is well defined.

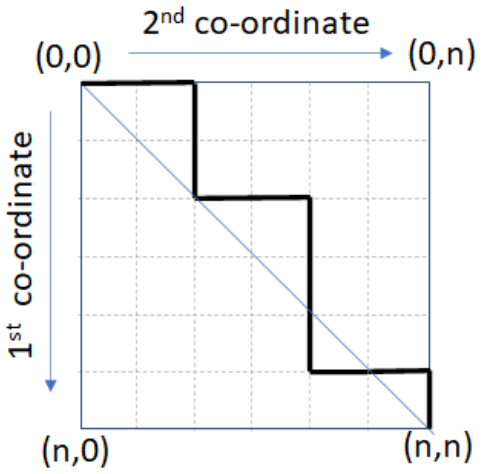


Figure 6.1: A path from $(0, 0)$ to (n, n) using downward and right edges

Injective: Two different paths from $(0, 0)$ to (n, n) will different in at least one (u_{i-1}, v_{i-1}) to (u_i, v_i) transition where $i = 1, 2, \cdots, 2n$, their corresponding strings under f will differ in at least i^{th} position and thus f is injective.

Surjective: Every string over $\{D, R\}$ of length $2n$ with equal number of D's and R's has a pre-image under f which is defined by $(u_0, v_0) = (0, 0)$ and (u_i, v_i) is $(u_{i-1} + 1, v_{i-1})$ if $b_i = R$ and $(u_{i-1}, v_{i-1} + 1)$ if $b_i = D$. As there will be n D's and n R's, $(u_{2n}, v_{2n}) = (n, n)$ and thus f is surjective .

Thus f is bijection. As we have number of string over $\{D, R\}$ of length $2n$ with equal number of D's and R's equal to $\binom{2n}{n}$ (select n positions out of $2n$ available and fill them with D's and the rest with R's). Thus the number of paths from $(0, 0)$ to (n, n) with only downward and rightward movements is $\binom{2n}{n}$.

Lets ask a slightly question. How many ways are there to go from $(0, 0)$ to $(n + 1, n - 1)$ using only downward or right edges.Using a similar arguments as above, we can come up with a bijection to set of string over $\{D, R\}$ of length $2n$ with $n + 1$ D's and $n - 1$ R's. Therefore number of required paths are $\binom{2n}{n+1} = \binom{2n}{n-1}$

6.3.2 Diagonal avoiding paths and Catlan numbers

In this section we explore the connection between the above paths that we discussed and the Catalan number. Let us ask this question: How many paths are there in the grid from $(0, 0)$ to (n, n) that avoids crossing the diagonal?

We first define what *crossing the diagonal* means. The diagonal consists of the points of the form (i, i), $i \in \{0, \ldots, n\}$. A path $((u_0, v_0), \ldots, (u_{2n}, v_{2n}))$ is said to be crossing the diagonal if it *intersects* through the diagonal and goes to some point below the diagonal. Mathematically, a path is a diagonal crossing path if $\exists\ i$ such that $u_i > v_i$. In particular, $\exists i : u_i = v_i + 1$ (refer fig.

6.2 for example. Any diagonal crossing path must necessarily pass through one of the red dots). Equivalently, in a diagonal avoiding path $\forall i \in \{0, \ldots, 2n\}, v_i \geq u_i$. A sample *diagonal-avoiding path* is shown in the fig. 6.3

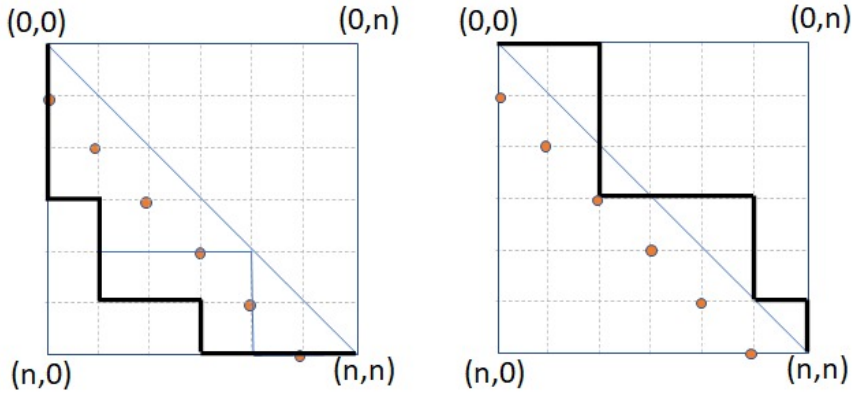


Figure 6.2: Diagonal crossing paths. Note that path in (a) is crossing the diagonal at $(0, 0)$

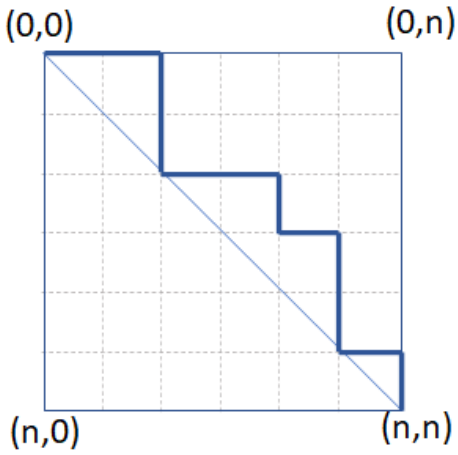


Figure 6.3: A diagonal avoiding path. Observe that it can still touch the diagonal

Before computing this number, an obvious question is what is the connection between such restricted paths and Catalan number. It turns out that the set of diagonal avoiding paths from $(0, 0)$ to (n, n) is in bijection with the set of balanced paranthesized strings of length $2n$. Hence, to count the number of balanced paranthesized strings of length $2n$, which is also the Catalan number, we only need to count the diagonal avoiding paths from $(0, 0)$ to (n, n). Let us first establish the bijection between the two.

6.3.3 Bijection from Diagonal avoiding paths to Balanced parenthesisation problem

Intuitively, the bijection can be defined as follows: for any given balanced parenthesized string $w = w_1 w_2 \ldots w_{2n}$, the corresponding path from $(0, 0)$ to (n, n) is obtained by starting from position

$(0,0)$, and scanning the string from left to right. Take right move whenever '(' is encountered and a down move for ')'. Formally we define the bijection as follows:

Defining the bijection: Let P be the set of diagonal avoiding paths from $(0,0)$ to (n,n) and B be the set of balanced paranthesized strings of length $2n$ over the alphabets $\{(,)\}$. Define the bijection $\phi : B \to P$ as follows:
For $w = w_1 w_2 \dots w_{2n} \in B$, $\phi(w) = (u_0, v_0), (u_1, v_1), \dots, (u_i, v_i), \dots, (u_{2n}, v_{2n})$, where

1. $(u_0, v_0) = (0,0)$
2. $\forall i \in \{1, 2, \dots, 2n\}$

$$(u_i, v_i) = \begin{cases} (u_{i-1} + 1, v_{i-1}) & \text{if } w_i =) \\ (u_{i-1}, v_{i-1} + 1) & \text{if } w_i = (\end{cases}$$

Proof of bijection

Well-defined: From the above description, given any string w, $\phi(w)$ is uniquely defined. Further, for any string $w \in B$, since the number of $'('$ is same as the number of $')' = n$, the corresponding path has n right and n down moves and hence it ends at (n,n). Also, since the number of left brackets is greater than or equal to the number of right brackets in any prefix of w, for all $i \in [2n]$, $v_i \geq u_i$. This shows that $\forall w \in B, \phi(w) \in P$. Hence, ϕ is well-defined.

Injective: Let w, w' be two different strings in set B. Then $\exists$ an index $i \in [2n]$ where $w_i \neq w'_i$. Hence $\phi(w)$ and $\phi(w')$ also differ at the ith step, where one of the paths takes one step right while the other takes one step down.

Surjective: Given any path $((0,0), (u_1, v_1), \dots, (u_{2n}, v_{2n}))$ the corresponding string $w \in B$ is defined as follows:
$\forall i \in [2n]$

$$w_i = \begin{cases} \text{`(`} & \text{if } (u_i, v_i) = (u_{i-1}, v_{i-1} + 1) \\ \text{`)`} & \text{if } (u_i, v_i) = (u_{i-1} + 1, v_{i-1}) \end{cases}$$

We can verify that the string w indeed is in set B, because firstly, for any path in P, $\forall i, v_i \geq u_i$ and hence by definition, number of left brackets '(' in w is greater than or equal to number of right brackets, '(' in any prefix of w. Secondly, for any path to reach from $(0,0)$ to (n,n) it must have n right moves (increase in 2nd coordinate) and n down moves (increase in 1st coordinate) and hence w must have n left brackets and n right brackets.

6.3.4 Counting the number of diagonal avoiding paths

Having established the bijection between Catalan number and diagonal avoiding paths, we get

$$C_n = \#\text{ of diagonal avoiding paths from } (0,0) to (n,n) \tag{6.7}$$

So, our next task is to count the number of diagonal avoiding paths from $(0,0)$ to (n,n). To count this, we take following approach. Let us call the diagonal avoiding paths as *good* paths and diagonal crossing paths as *bad* paths. Then,

$$\begin{array}{c}\text{\# of diagonal avoiding paths}\\ \text{from } (0,0) \text{ to } (n,n)\end{array} = \begin{array}{c}\text{\# of paths}\\ \text{from } (0,0) \text{ to } (n,n)\end{array} - \begin{array}{c}\text{\# of diagonal crossing paths}\\ \text{from } (0,0) \text{ to } (n,n)\end{array} \tag{6.8}$$

So, now our revised goal is to count the number of diagonal crossing paths from $(0,0)$ to (n,n). How do we do that? Here again bijection plays an important role. The idea is to translate diagonal crossing paths into different kind of paths which are easy to count.

Let us define the following path translation: Let $\pi = (0,0), (u_1, v_1), \ldots, (u_{2n}, v_{2n})$ be a diagonal crossing path. Then there must exist i such that $u_i = v_i + 1$. There can be many such indices as the path can cross the diagonal multiple times. Choose i to be the least such index. Let $u_i = \ell$, then the first co-ordinate after crossing the diagonal is $(\ell, \ell - 1)$. Let us call this point P (refer fig. 6.4(a)). Then to find the translated path we reflect the part of the path π after point P w.r.t. the main diagonal.

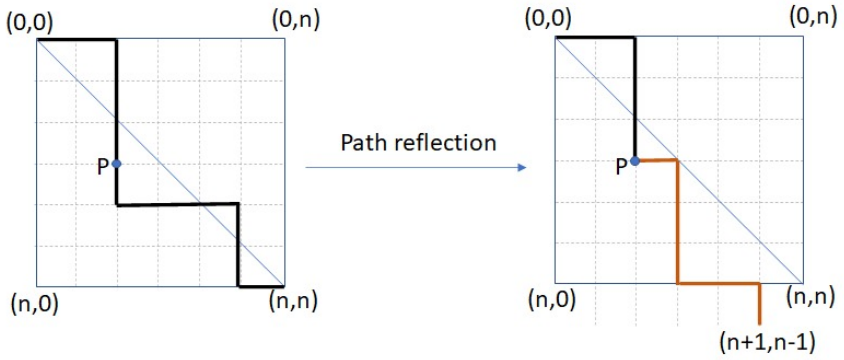


Figure 6.4: Point P in a diagonal crossing path and the reflected path after P

More precisely, we can divide the diagonal crossing path into two stretch S_1, S_2, where S_1 is the part of the path between $(0,0)$ to P and S_2 is the part of the path between P to (n,n). Then to translate π into a new path, replace S_2 with S_2' to get a new path $\pi' = S_1 S_2'$. The replacement S_2' is defined as follows:

- replace downward edges with right edges and
- replace right edges with downward edges.

Refer fig. 6.4(b) We can observe that the new path π' described in this way is always between $(0,0)$ to $(n+1, n-1)$. The argument for this goes as follows:

Originally (in S_2), $(\ell, \ell-1)$ goes to (n,n) which means it takes $(n-\ell)$ downward moves and $(n-\ell+1)$ right moves. Since, we are swapping the right and downward moves to get S_2' from S_2, there are $(n-\ell+1)$ downward moves and $(n-\ell)$ right moves from point $P = (\ell, \ell-1)$ in S_2'.

Thus, S'_2 goes from $(\ell, \ell-1)$ to $(\ell+n-\ell+1, \ell-1+n-\ell) = (n+1, n-1)$ and hence, $\pi' = S_1 S'_2$ is a path from $(0,0)$ to $(n+1, n-1)$.

Thus we have established that any diagonal crossing path from $(0,0)$ to (n,n) maps to a path from $(0,0)$ to $(n+1, n-1)$ after applying the transformation described above. The converse is also true, i.e., given any path from $(0,0)$ to $(n+1, n-1)$, we can translate it back to a diagonal crossing path from $(0,0)$ to (n,n) by using the same reflection technique. Thus, we get a bijection between the set of diagonal crossing paths from $(0,0)$ to (n,n) to the set of paths from $(0,0)$ to $(n+1, n-1)$. We formally define the translation and prove that it is indeed a bijection.

Bijection: Let A be the set of diagonal crossing paths from $(0,0)$ to (n,n) and B be the set of paths from $(0,0)$ to $(n+1, n-1)$. Then the mapping $\phi : A \to B$ is formally defined as follows:
Let $\pi = (0,0), (u_1, v_1), \ldots, (u_{2n}, v_{2n})$ and (u_i, v_i) be the first point when π crosses the diagonal. Then $\phi(\pi) = \pi' = (0,0), (u'_1, v'_1), \ldots, (u'_{2n}, v'_{2n})$ is given by:

1. $\forall 1 \leq j \leq i, (u'_j, v'_j) = (u_j, v_j)$
2. $\forall i+1 \leq j \leq 2n$,

$$(u'_j, v'_j) = \begin{cases} (u'_{j-1}+1, v'_{j-1}) & \text{if } (u_j, v_j) = (u_{j-1}, v'_{j-1}+1) \\ (u'_{j-1}, v'_{j-1}+1) & \text{if } (u_j, v_j) = (u_{j-1}+1, v'_{j-1}) \end{cases}$$

Well-defined: We already observed that any path $\pi \in A$ from $(0,0)$ to (n,n) maps to a path $(0,0)$ to $(n+1, n-1)$. Hence ϕ is well defined.

Injection: Consider two different diagonal crossing paths π_1 and π_2. Let $\pi_1 = S_{1,1} S_{1,2}$ and $\pi_2 = S_{2,1} S_{2,2}$, where the two components $S_{i,1}$ and $S_{i,2}$ for $i \in \{1,2\}$ are as defined before. Then following two cases are possible:

- Case1: $S_{11} \neq S_{21}$. Then $\pi'_1 \neq \pi'_2$, because the first component is copied as it is in the translation, i.e. $\pi'_1 = S_{1,1} S'_{1,2}$ and $\pi'_2 = S_{2,1} S'_{2,2}$.
- Case2: $S_{11} = S_{21}$, but $S_{12} \neq S_{22}$. In this case $S'_{12} \neq S'_{22}$ because of the way it is defined, i.e. for every right move there is a downwards move and vice-versa. Hence, $\pi'_i \neq \pi'_2$.

Surjective: Given any path π' from $(0,0)$ to $(n+1, n-1)$, we can construct the corresponding path π from $(0,0)$ to (n,n), such that $\phi(\pi) = \pi'$, as follows.
Let $\pi' = (0,0), (u'_1, v'_1), \ldots, (u'_{2n}, v'_{2n})$. Since π' goes to $(n+1, n-1)$ which is below the diagonal there must exist i such that (u'_i, v'_i) is below the diagonal. Again, there can be many such indices. Take i to be the first such index. Same as before, let $\pi' = S'_1 S'_2$, where S'_1 is the path from $(0,0)$ to (u'_i, v'_i) and S'_1 is the path from (u'_i, v'_i) to (u'_{2n}, v'_{2n}). Then $\pi = S'_1 S_2$ where S_2 is obtained from S'_2 by swapping the right and downwards moves. Mathematically, let $\pi = (0,0), (u_1, v_1), \ldots, (u_{2n}, v_{2n})$. Then

1. $\forall j \leq i$, $(u_j, v_j) = (u'_j, v'_j)$

2. $\forall\ i+1 \le j \le 2n$

$$(u_j, v_j) = \begin{cases} (u_{j-1}+1, v_{j-1}) & \text{if } (u'_j, v'_j) = (u'_{j-1}, v'_{j-1}+1) \\ (u_{j-1}, v_{j-1}+1) & \text{if } (u'_j, v'_j) = (u'_{j-1}+1, v'_{j-1}) \end{cases}$$

Again by the same argument as before it can be verified that π is a diagonal crossing path from $(0,0)$ to (n,n). We write it here for completeness. Let $(\ell, \ell-1)$ be the first point when π' crosses the diagonal. Then since the path from $(0,0)$ to $(\ell, \ell-1)$ remains as it is in π, it is a diagonal crossing path. Further since π' is path from $(0,0)$ to $(n+1, n-1)$, it takes $n+1-\ell$ downward steps and $n-\ell$ right steps from $(\ell, \ell-1)$. Hence, π takes $n+1-\ell$ right and $n-\ell$ downward steps from $(\ell, \ell-1)$. Thus, π ends at $(\ell+n-\ell, \ell-1+n+1-\ell) = (n,n)$.

Thus, we have established a bijection between the set of diagonal crossing paths from $(0,0)$ to (n,n) and the set of paths from $(0,0)$ to $(n+1, n-1)$. Hence,

$$\begin{aligned} \#\text{of diagonal crossing paths from } (0,0) \text{ to } (n,n) &= \#\text{of paths from } (0,0) \text{ to } (n+1,n-1) \\ &= \binom{2n}{n+1} \end{aligned}$$

Hence, from (6.7),(6.8),

$$\begin{aligned} C_n &= \#\text{of diagonal avoiding paths from} (0,0) \text{ to } (n,n) \\ &= \begin{matrix}\#\text{ of paths} \\ \text{from } (0,0) \text{ to } (n,n)\end{matrix} - \begin{matrix}\#\text{ of diagonal crossing paths} \\ \text{from } (0,0) \text{ to } (n,n)\end{matrix} \\ &= \binom{2n}{n} - \binom{2n}{n+1} \\ &= \binom{2n}{n} - \frac{n}{n+1}\binom{2n}{n} \\ &= \frac{1}{n+1}\binom{2n}{n} \end{aligned}$$

Here, in the second last line, we have used the identity:

$$\binom{2n}{n+1} = \frac{n}{n+1}\binom{2n}{n}.$$

Exercise 6.4.

Try to establish a bijection between the set of different possible polygon triangulation in a polygon of $n+2$ nodes and the set of binary trees with n internal nodes.

Hint: associate each internal node with a triangle in a triangulation. Then, each internal node will have degree three, which is the case for full binary tree, except for the leaves. Leaves will correspond to those triangles whose one of the edge is the boundary of the polygon.

Instructor : Jayalal Sarma
Scribe : Anshu Yadav (*TA:* JS)
Date : Sept 19, 2020
Status : α

Lecture 7

From Bijections to PIE

7.1 Introduction

In this lecture, we will continue with the use of bijections and use it in formally proving the two identities that we discussed in class and then see their relationship to the Principal of Inclusion and Exclusion.

7.2 The Identities

Recall that we proved following two identities in one of the discussion sessions

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k} = 0 \tag{7.9}$$

$$\sum_{k=0}^{m}(-1)^k\binom{n}{k} = (-1)^m\binom{n-1}{m} \tag{7.10}$$

In this section, we will see the proofs for the above equations is detail

7.2.1 Proof for Eqn. (7.9)

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k} = 0$$

Proof. The LHS counts the number of even sized subsets of $[n]$ with positive sign and odd size subsets with negative sign. Then we proved the result using bijection between even sized and odd sized subsets of $[n]$. Hence, we get 0 on RHS. Let us formally define the bijection here.

Let E be the set of all even sized subsets of $[n]$ and O be the set of all odd sized subsets of $[n]$. Then the bijection $\phi_i : E \to O$ is defined with respect to an element $i \in [n]$ as follows.

Let $X \subseteq [n]$, such that $|X|$ is even. Then

$$\phi_i(X) = \begin{cases} X \setminus \{i\} & \text{if } i \in X \\ X \cup \{i\} & \text{if } i \notin X \end{cases}$$

Proof of bijection:

Well-defined: Given any even sized subset X, there are two possibilities: (i) $i \in X$, (ii) $i \notin X$. In first case, i is removed from X, hence its size reduces by one and becomes odd. In the second case, i is added, hence the size of the subset increases by one and becomes odd. Hence, ϕ is well defined.

Injective: Let X and X' be two distinct subsets of $[n]$. Then $\exists j \in [n]$ such that j is present in exactly one of the two subsets. Wlog, let $j \in X$ and $j \notin X'$. Now, if $j \neq i$, then $j \in \phi(X)$ and $j \notin \phi(X')$ and hence $\phi(X) \neq \phi(X')$. On the other hand, if $j = i$, then $j \notin \phi(X)$ and $j \in \phi(X')$. Hence, $\phi(X) \neq \phi(X')$.

Surjective: Let $Y \in O$ be an odd sized subset of $[n]$. From Y, we can recover X such that $\phi(X) = Y$ by the same operation as in ϕ. That is,

$$X = \begin{cases} Y \setminus \{i\} & \text{if } i \in Y \\ Y \cup \{i\} & \text{if } i \notin Y \end{cases}$$

It can easily be verified that in both the cases, X is an even sized subset of $[n]$.

This completes the proof. □

7.2.2 Proof for Eqn. (7.10)

$$\sum_{k=0}^{m} (-1)^k \binom{n}{k} = \binom{n-1}{m}$$

Proof. Now we look at the second identity which is even more interesting. To prove this identity we use *almost bijection* where the bijection is between a set and subset of another set.

In words, the identity to prove, can be described as

$$\underset{\text{of size atmost } m}{\#\text{ of even sized subsets of } [n]} - \underset{\text{of size atmost } m}{\#\text{ of odd sized subsets of } [n]} = (-1)^m \binom{n-1}{m}.$$

Clearly, there cannot be a bijection between the two sets (even sized subsets and odd sized subsets) in this case, since their difference is non-zero. This is where we use almost bijection.

We use following case analysis.

Case1: m is even: Then the identity to prove is:

$$\sum_{k=0}^{m}(-1)^k\binom{n}{k} = \binom{n-1}{m} \tag{7.11}$$

This can be interpreted as

$$\sum_{\substack{k=0,\\ k \text{ is even}}}^{m}\binom{n}{k} - \sum_{\substack{k=1,\\ k \text{ is odd}}}^{m-1}\binom{n}{k} = \binom{n-1}{m} \tag{7.12}$$

Let E be the set of all the even sized subsets of $[n]$ of size at most m and O be the set of odd sized subsets of $[n]$ having size at most $m-1$. Then, Eqn. (7.12) can intuitively interpreted as follows: there is a subset $E' \subseteq E$, such that E' is in bijection with O and $|E \setminus E'| = \binom{n-1}{m}$. Thus, we have three tasks at hand

- identify the set E', and
- define and prove the bijection between E' and O.
- prove that $|E \setminus E'| = \binom{n-1}{m}$

Defining the set E': Set E' is the union of two sets:

$$E' = \{X \subseteq [n] : |X| \text{ is even and } |X| \leq m-2\} \cup \{X \subseteq [n] : i \in X \text{ and } |X| = m\}$$

Defining the bijection: The bijection $\phi : E' \to B$ is defined in the same way as we defined it for first identity. That is, for $X \in E'$,

$$\phi(X) = \begin{cases} X \setminus \{i\} & \text{if } i \in X \\ X \cup \{i\} & \text{if } i \notin X \end{cases}$$

Proof of bijection

Well-defined: Let $X \in E'$, then (i) if $|X| \leq m-2$, then $|\phi(X)|$ is odd and $|\phi(X)| \leq m-1$, (ii) if $|X| = m$, then $i \in X$, hence $\phi(X) = X \setminus \{i\}$. This implies $|\phi(X)| = m-1$. Thus, in both the cases $\phi(X) \in O$.

Injective: Since, the function is same as in the previous case, the same argument for injectivity works.

Surjective: Let $Y \in O$ be an odd sized subset of $[n]$. From Y, we can recover $X \in E'$ such that $\phi(X) = Y$ by the same operation as in ϕ. That is,

$$X = \begin{cases} Y \setminus \{i\} & \text{if } i \in Y \\ Y \cup \{i\} & \text{if } i \notin Y \end{cases}$$

It can easily be verified that in both the cases, $|X|$ is even. In first case, since $|Y| \leq m-1, |X| \leq m-2$, hence $X \in E'$. In second case, since $i \notin Y$ and $|Y| \leq m-1, |X| \leq m$ and $i \in X$. Hence $X \in E'$, by definition.

This proves the bijection between E' and O.

Proof for: $|E \setminus E'| = \binom{n-1}{m}$

From the above definitions, $E \setminus E' = \{X \subseteq [n] : |X| = m, i \notin X\}$. This can be interpreted as $E \setminus E' = \{X \subseteq [n] \setminus \{i\} : |X| = m\}$. Hence, $|E \setminus E'| = \binom{n-1}{m}$.

Case2: m is odd: In this case the identity to prove is:

$$\sum_{k=0}^{m} (-1)^k \binom{n}{k} = -\binom{n-1}{m} \tag{7.13}$$

This can be interpreted as

$$\sum_{\substack{k=0,\\ k \text{ is even}}}^{m-1} \binom{n}{k} - \sum_{\substack{k=1,\\ k \text{ is odd}}}^{m} \binom{n}{k} = -\binom{n-1}{m} \tag{7.14}$$

Equivalently,

$$\sum_{\substack{k=1,\\ k \text{ is odd}}}^{m} \binom{n}{k} - \sum_{\substack{k=0,\\ k \text{ is even}}}^{m-1} \binom{n}{k} = \binom{n-1}{m} \tag{7.15}$$

This time the set of odd sized subsets of $[n]$ of size at most m is bigger than the even sized subsets of $[n]$ of size at most m. The proof is same as that for the case of even m. Let E be the set of all the even sized subsets of $[n]$ of size at most $m-1$ (since m is odd) and O be the set of odd sized subsets of $[n]$ having size at most m. Then (7.15) can be interpreted as follows: there is a subset $O' \subseteq O$, such that E is in bijection with O' and $|O \setminus O'| = \binom{n-1}{m}$.

Thus, we have two task at hand

- identify the set O', and
- define and prove the bijection between E and O'.
- prove that $|O \setminus O'| = \binom{n-1}{m}$

Defining the set O': Set O' to be the union of two sets:

$$O' = \{Y \subseteq [n] : |Y| \text{ is odd and } |Y| \leq m-2\} \cup \{Y \subseteq [n] : i \in Y \text{ and } |Y| = m\}$$

Defining the bijection: The bijection $\phi : E \to O'$ is defined in the same way as we defined it

for first identity. That is, for $X \in E$,

$$\phi(X) = \begin{cases} X \setminus \{i\} & \text{if } i \in X \\ X \cup \{i\} & \text{if } i \notin X \end{cases}$$

Proof of bijection

Well-defined: Let $X \in E$, then $\phi(X)$ is of odd size because either an element is added or removed from X, which is of even size. Now, (i) if $i \in X$, then $\phi(X) = X \setminus \{i\}$. Hence, $|\phi(X)| \leq m - 2$ (because $|X| \leq m - 1$) which implies $\phi(X) \in O'$ (ii) if $i \notin X$, then, $\phi(X) = X \cup \{i\}$. This implies $|\phi(X)| \leq m$. But since, $i \in \phi(X)$, $\phi(X) \in O'$. This proves that ϕ is well- defined.

Injective: Since, the function is same as in sub section 7.2.1, the same argument for injectivity works.

Surjective: Let $Y \in O'$ be an odd sized subset of $[n]$. From Y, we can recover $X \in E$ such that $\phi(X) = Y$ by the same operation as in ϕ. That is,

$$X = \begin{cases} Y \setminus \{i\} & \text{if } i \in Y \\ Y \cup \{i\} & \text{if } i \notin Y \end{cases}$$

It can easily be verified that in both the cases, $|X|$ is even. In first case, $|Y| \leq m$ and hence $|X| \leq m - 1$. So, $X \in E$. In second case, since $i \notin Y$, $|Y| \leq m - 2$ (by definition) and hence $|X| \leq m - 1$. Hence $X \in E$.

This proves the bijection between E and O'.

Proof for: $|O \setminus O'| = \binom{n-1}{m}$
From the above definitions, $O \setminus O' = \{Y \subseteq [n] : |Y| = m, i \notin Y\}$. This can be interpreted as $O \setminus O' = \{Y \subseteq [n] \setminus \{i\} : |Y| = m\}$. Hence, $|O \setminus O'| = \binom{n-1}{m}$.

This completes the proof □

This proves both the identities.

7.3 Principle of Inclusion and Exclusion

Suppose we are given n sets $A_1, A_2, \ldots, A_n \subseteq G$, where G is some ground set. We are interested in finding the size of $A = A_1 \cup A_2 \cup \ldots \cup A_n$. This is very abstract scenario and we will see specific examples later, but here we are going to see classic use of the above identities in deriving this number.

So, we are interested in finding $|A| = |A_1 \cup A_2 \cup \ldots \cup A_n|$.

So, here is a thought process - Clearly, we can add the size of individual sets as $|A| = |A_1| + |A_2| + \ldots + |A_n|$, but this will over-count if there are some elements present in more than one sets. So, for that we need to subtract the double counting. For e.g. if $x \in A_1$ and $x \in A_2$, then it gets counted twice and to compensate for that we need to subtract $|A| = |A_1 \cap A_2|$ and we might attempt $|A| = |A_1| + |A_2| + \ldots + |A_n| - \sum_{1 \le i < j \le n} |A_i \cap A_j|$. But then, if x is present in A_1, A_2 and A_3, the it is under-counted (added thrice and subtracted thrice). So, again we need to compensate for that by adding $\sum_{1 \le i \le j \le k \le n} |A_i \cap A_j \cap A_k|$ in the above expression and this sequence goes on for any element being present in $k \le n$ sets and finally we get the expression for $|A|$ as follows

$$|A| = |A_1| + \cdots + |A_n| - \sum_{1 \le i < j \le n} |A_i \cap A_j| + \sum_{1 \le i < j < k} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n+1} |A_1 \cap A_2 \cap \cdots \cap A_n| \tag{7.16}$$

For $n = 2$, the above expression gives

$$|A| = |A_1| + |A_2| - |A_1 \cap A_2|$$

which we all must have seen before and can easily prove using Venn diagram.

In this section, we will formally prove the above expression for general n using the two identities we proved in previous section.

Proof. Consider any $x \in A_1 \cup A_2 \cup \cdots \cup A_n$. Let x appears in k of the A_i's. Then let us see how x gets counted

- $|A_1| + |A_2| + \cdots + |A_n|$: counts x k times (added)
- $\sum_{1 \le i < j \le n} |A_i \cap A_j|$: counts x $\binom{k}{2}$ times (subtracted)
- $\sum_{1 \le i < j < k \le n} |A_i \cap A_j \cap A_k|$: counts x $\binom{k}{3}$ times (added)
- and so on ...

Notice that in terms involving intersection of more than k sets, x never appears.

Thus,

$$\begin{aligned}
\substack{\#\text{of times } x \\ \text{gets counted}} &= k + \binom{k}{2} - \binom{k}{3} + \cdots + (-1)^{k+1} \binom{k}{k} \\
&= -\binom{k}{0} + \binom{k}{1} + \binom{k}{2} - \binom{k}{3} + \cdots + (-1)^{k+1} \binom{k}{k} + \binom{k}{0} \\
&= -\sum_{i=0}^{k} (-1)^i \binom{k}{i} + \binom{k}{0} \\
&= \binom{k}{0} \qquad \text{from (7.9)} \\
&= 1
\end{aligned}$$

Thus, irrespective of the value of k, any element $x \in A_1 \cup A_2 \cup \cdots \cup A_n$ is counted exactly once. Hence, every $x \in A_1 \cup A_2 \cup \cdots \cup A_n$ is counted exactly once in RHS in (7.16).

This proves the PIE $\square$

Now let us look at the application of second identity that we derived. This identity is used in deriving a version of PIE which appears very naturally in several context. Let us look at one such example.

PIE says that if we want to derive $|A_1 \cup A_2 \cup \cdots \cup A_n|$, then the following expression does not give the correct count.

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = |A_1| + |A_2| + \cdots + |A_n|$$

But we can ask, does this expression gives a lower or an upper bound? As we saw, this does over-counting, hence we can write

$$|A_1 \cup A_2 \cup \cdots \cup A_n| \leq |A_1| + |A_2| + \cdots + |A_n|$$

Now, suppose we include the next component, i.e.

$$|A_1| + |A_2| + \cdots + |A_n| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j|$$

Again from PIE we know that this also does not give the correct count. But we ask the same question again - does it give any lower or upper bound. And as we saw that this term can do some over-subtraction and hence we can say that this expression gives the lower bound. That is,

$$|A_1 \cup A_2 \cup \cdots \cup A_n| \geq |A_1| + |A_2| + \cdots + |A_n| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j|$$

Similarly,

$$|A_1 \cup A_2 \cup \cdots \cup A_n| \leq |A_1| + |A_2| + \cdots + |A_n| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j| + \sum_{1 \leq i < j < k \leq n} |A_i \cap A_j \cap A_k|$$

and we continue like this.

Let us now formally establish this observation. We use the same technique that we used in the proof of PIE.

Let x appears in k of the sets in $A_1, A_2, \ldots, A_n$. Suppose we cut off the PIE after $m \leq n$ sized

intersections. Then

$$\begin{aligned}
\#\text{of times } x \text{ gets counted} &= \binom{k}{1} - \binom{k}{2} + \cdots + (-1)^{m+1}\binom{k}{m} \\
&= -\sum_{i=0}^{m}(-1)^i\binom{k}{i} + \binom{k}{0} \\
&= 1 + (-1)^{m+1}\binom{k-1}{m} \qquad \text{from (7.10)}
\end{aligned}$$

Thus, x is over counted or under counted depending on whether the second term on RHS is positive or negative. Let us analyze this for two cases.

Case1: $k \leq m$

Since, x appears in only $k \leq m$ sets and we are cutting down only after m, then this means that all possible intersections of this particular x are added and subtracted and x can not appear in any of the intersections of more than k sets. Hence, x is neither under counted nor over counted. In the expression, $\binom{k-1}{m} = 0$ Hence,

$$\#\text{of times } x \text{ is counted } = 1$$

Case2: $k > m$

In this case, x can be under counted or over counted depending upon whether m is even or odd. If m is odd then x is over counted.

If m is even then x is under counted.

Notice that either all $x \in A_1 \cup \mathsf{A}_2 \cup \cdots \cup A_n$ are correctly counted or under counted or all x are correctly counted or over counted based on the parity of m. Thus, whether a PIE cut down after m intersections gives lower bound or upper bound depends only on the parity of m. This principle is also called the *Bon Ferroni's inequality*.

REMARK 7.3.1. We used the equality in (7.11) to prove PIE. We can actually do the other way round as well, i.e. we can use PIE to prove this equality too.

This completes this lecture. In the next lecture we will look at some applications of PIE.

7.4 Discussions

Bijection from Euler's problem to Binary Trees As we have already established a bijection from set of balanced parenthesisations to set of full binary trees and established that number of full binary trees with n internal nodes is the catlan number C_n, in this section, let's establish a bijection from the *Euler's Problem* to set of full binary trees to establish that the solution to *Euler's problem* is also catlan number C_n.

Lets recall *Euler's problem* first. Consider a convex polygon with $n+2$ edges. Euler's problem is the number of ways of triangulating it (partition the polygon into triangles) by drawing non-crossing diagonals. (Refer fig. 7.5). We know that number of non-crossing diagonals in a polygon of $n+2$ edges is $n-1$ (proof follows from a simple induction) and from those $n-1$ non-crossing diagonals, we have our polygon partitioned into n triangles. Let's associate each of the triangles with a vertex (green dots in the fig. 7.5). Observe that if two triangles share an edge, it must be one of the diagonals (no two triangles can share an edge because of non-crossing diagonals). Now, let's connect the vertices whose corresponding triangles share an edge. Any edge connecting two of these vertices crosses a diagonal. Now, consider a polygon edge e. For every polygon edge

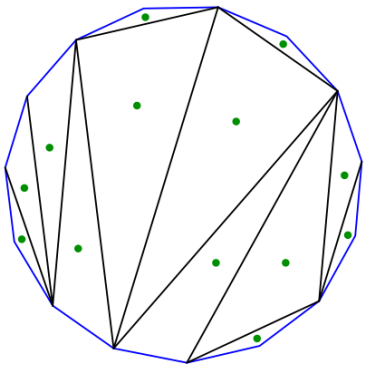


Figure 7.5: Partitioning a polygon into triangles by non-crossing diagonals. Observe that green dots in each triangle associates the triangle with a vertex

surrounding a vertex (other than e), add an open-edge originating from that vertex (see fig. 7.6). We arrive at the following claim.

CLAIM 7.4.1. *If we remove the underlying triangles (which are formed with polygon edges and diagonals), from fig. 7.6, the resulting graph obtained (see fig. 7.7) is a full binary tree with the vertices as internal nodes.*

Proof. We observe that degree of every vertex other than the vertex surrounded by edge e is 2. This vertex will act as root to our full binary tree. All other vertices have degree 3 because each vertex is surrounded by a triangle and if a side is a diagonal, it will be connected to vertex which is surrounded by triangle that shares the diagonal and if the side is a polygon edge, then there will be an open edge corresponding to it originating from the vertex. Therefore the resulting graph formed is a full binary tree with our vertices as n internal nodes and vertices corresponding to open edges are $n+1$ leaves (because there are $n+2$ edges and one edge is under consideration). This completes the description of bijection. □

We leave it as an exercise to the reader to prove that the mapping defined above is indeed a

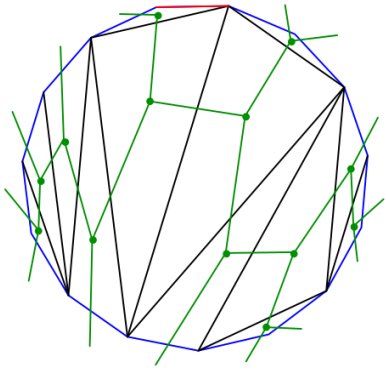


Figure 7.6: Polygon with vertices connected to form a tree

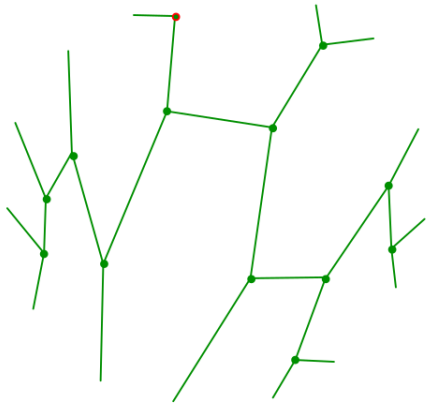


Figure 7.7: Tree formed by connecting vertices

bijection.

Bijection from binary trees to full binary trees In this section we are interested in connection between binary and full binary trees. Recall that a full binary tree is one in which each node has either 0 or two children. On the other hand, when we say binary tree then it only means that each node can have at most two children. We want to find a bijection between set of binary trees with n internal nodes and set of full binary trees with certain number of internal nodes.

First of all lets try to see how to convert a given binary tree into a full binary tree so that we can reverse the process, i.e. recover the original (binary) tree back from the full binary tree without ambiguity.

Here is the first attempt:

Attempt 1: First natural approach can be to add a leaf node to all non-full (internal nodes having only one child) nodes, as shown in figure 7.8

But notice that this transformation is not injective. For example, it can be observed that both

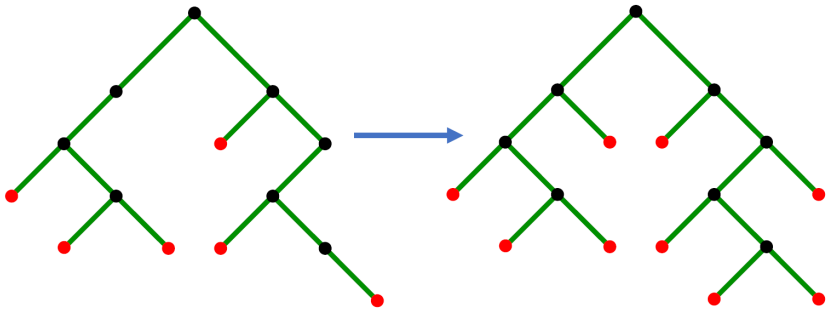


Figure 7.8: Binary to full binary tree attempt1: adding a child node to each non full node

the trees in figure 7.9 map to same full binary tree.

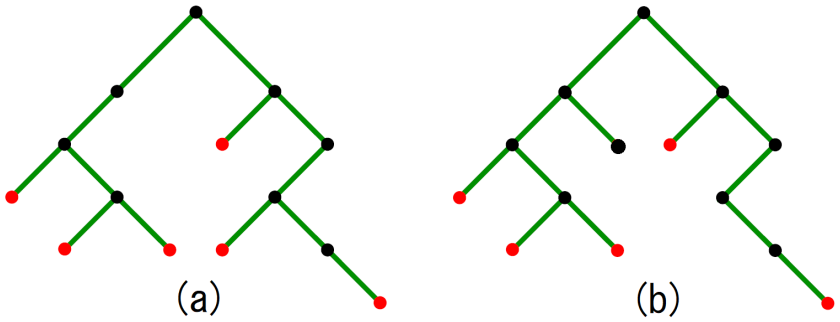


Figure 7.9: Two different binary trees that map to same full binary tree

Attempt 2(correct) Lets try a slightly different approach. Given a binary tree, do the following:

- to each leaf node, add two children
- to each internal node having only one child, add another child

Figure 7.10 shows the full binary tree constructed in this way for the same binary tree as in Figure 7.8. We can see that this solution addresses the issue in the first attempt. Intuitively because of following argument: in the previous attempt the problem was that given a full binary tree, it was hard to decide if a leaf node was originally present in the binary tree or added during transformation. Now, in the current solution, this issue does not arise, because for any leaf node originally present in the binary tree, we add two new leaves as its children. Thus, it can be observed that all the leaf nodes (and only these nodes) are added during transformation.

To see that this translation is well-defined, we can see that the transformed tree is full binary tree by construction itself. Surjectivity is also easy to prove. To recover a binary tree from any given full binary tree, simply remove all the leaf nodes. We discussed injection informally. To give a formal argument, we first need to identify how to characterize two different binary trees? One of the hint as given during the discussion is to assign address to the nodes in the form of binary string, where 0-1 represents left or right child.

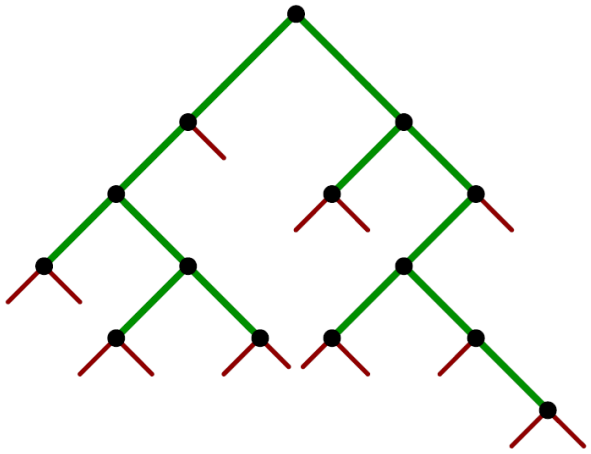


Figure 7.10: Full binary tree for the (non-full) binary tree given in fig 7.8. Notice that all the leaf nodes are added during transformation

Here we argued the bijection only intuitively and there are many things to be worked out formally. For example, proof for injection is not formally argued. Also, to argue surjection, we need to fix the number of nodes in full binary tree. Once we figure out this number, the argument for transformation being well-defined also need to take that into account.

Writing a complete formal proof of bijection is left as homework exercise.

Bijection between plane trees and full binary trees A plane tree is a rooted tree with an ordering among the children. A plane tree can have more than two children. Figure 7.11 shows a plane tree.

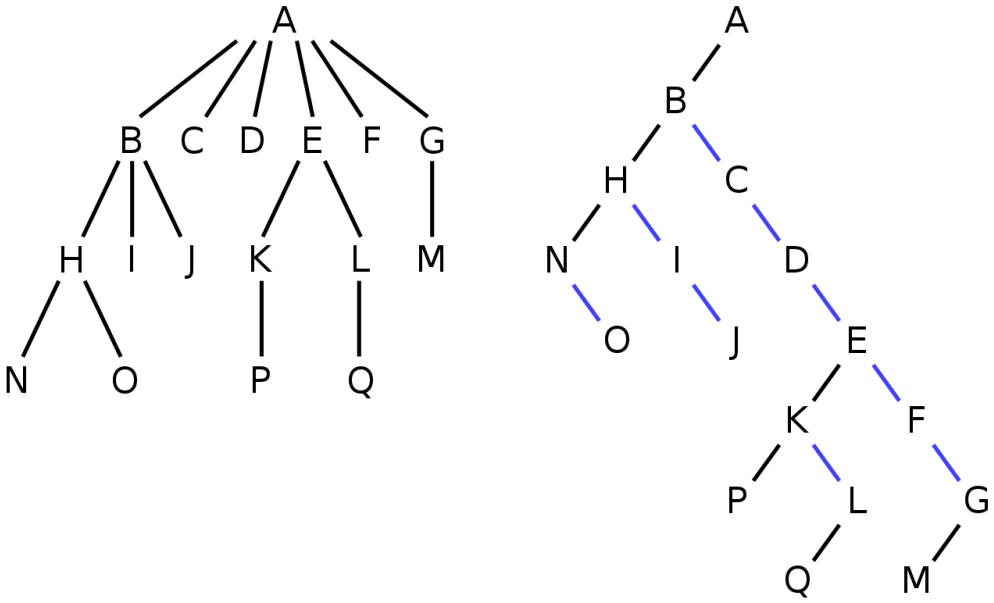


Figure 7.11: An example of plane trees and its transformation to a binary tree

We are interested in studying the connection between plane trees and binary trees. The number of plane trees with n nodes is equal to the number of binary trees with n nodes. Thus, there is bijection between set of plane trees with n nodes and the set of binary trees with n nodes.

Here we define the bijection function.

The Bijection: Given any plane tree, do the following

- For each node in the tree,
  - add its first child in plane tree as its left child in binary tree
  - add its immediate sibling on right as its right child in binary tree.

  child in the binary tree.

By following the above rule, we get a binary tree from given plane tree.

Observe that in the binary tree thus obtained, root node has only one child, while in general, in a binary tree the root can have both its children. Hence, we won't include the root as part of the binary tree.

Writing formal argument for all the properties is left as homework excercise.

Chapter 8

Supplementary Material

8.1 Curiosity Collection

Here we list down all the "out of curious" questions that we discussed (sometimes even not discussed) in the class (and hence in this document).

Curiosity 8.1.1. It is an amusing question to ask, whether there are other objects, which we did not intend to, which also satisfies the axioms that we wrote, by accident. Say for example, we wrote the axioms for graphs, but "strings" also satisfies them. If so, the theorems that we prove for graphs using only those axioms will also be true for strings, automatically !!. Quite interestingly this is true for natural numbers. The mathematical theory of natural numbers is axiomatized by what are called the Peano's axioms. There are numbers that one can define which are different from natural numbers for which any theorem that we prove for natural numbers also are true (because they satisfy the Peano's axioms). Then one might ask, are we not trying to represent exactly natural numbers? So should we not augment Peano's axioms with more properties of natural numbers such that we remove such *unwanted* parallel models from satisfying the axioms we write. Even more interestingly, one can argue that this is not even possible. No matter, what extra formula we write the existence of such "parallel models" us inevitable. In fact, not just one "parallel model", there will be infintiely many of them. You should read about *Löwenheim–Skolem theorem*.

Curiosity 8.1.2. The formal proof of PHP as simple as it sounds is still a subject of substantial research in an area called *proof complexity*. To demonstrate this, let us write the principle itself in more rigorous notations. Let $n > k$, and $\{x_{ij} \mid i \in [n], j \in [k]\}$ be propositional variables (which can be called, say *pigeon hole variables*). Following our original notation, where there are n pigeons and k holes, the basic Pigeon Hole Principle is the following Disjunctive normal form formula :

$$\text{PHP}_k^n \stackrel{\text{def}}{=} \left(\bigvee_{i \in [n]} \bigwedge_{j \in [k]} \overline{x_{ij}} \right) \vee \left(\bigvee_{j \in [k]} \bigvee_{r \neq s \in [n]} (x_{rj} \wedge x_{sj}) \right)$$

To prove this, one possibility is to derive the contradiction from the negation of PHP_k^n. This is an expression in conjunctive normal form, with clauses:

$$\text{For } i \in [n] \text{ the clauses : } Q_i \stackrel{\text{def}}{=} \bigvee_{j=1}^{k} x_{ij}$$

$$\text{and for } s \neq t \in [n], j \in [k] \text{ the clauses } Q_{s,t,j} \stackrel{\text{def}}{=} \overline{x_{sj}} \vee \overline{x_{tj}}$$

Intuitively, these say that there is a function from $[n] \to [k]$ (which is represented by $x_{ij} = 1$ to mean that the function takes i to j) which is well defined (for every i, there exists a j such that $x_{ij} = 1$) and also injective (for two different s and t, it is not the case that x_{si} is 1 and x_{tj}). Since $n > k$, there cannot be an injection, and hence the negation of the conjunction of these clauses PHP_k^n must be true.

Suppose we ask, starting from these clauses as axioms, and applying rules of inferences (say the resolution principle) alone, how many steps of proof does one need to do to derive the contradiction ($r \wedge \neg r$ for some r). [1]. We measure this in terms of n and k which determines the number of variables in the system. The area which studies the complexity of proofs in the above is called *proof complexity theory*. It turns out the the basic PHP itself is one of the tautologies for which one requires exponentially long proofs if we are restricting ourselves to resolution? What if we relax this? The area has several interesting open questions related to this and they have close connections to computational complexity theory too.

[1] Notice that this sounds exactly like computation, how many steps of computation is required in order to certain tasks in terms of input parameters

8.2 Exercises

Exercise 8.3. A social network is said to be symmetric if the relation between users that is maintained as a part of the network, is symmetric. Consider a symmetric social network and let the symmetric relation maintained be that of "user A and B are *friends*" (like in the case of facebook). A user C is said to be a *mutual friend* of users A and B if, C is a friend of both A and B. Prove that - for any user A of the network who has at least two friends, there must exist two friends of A who has the same number of mutual friends with A.

Comment on whether symmetry is critical for your argument. Take the example of *instagram* where the symmetric relation of *friends* is replaced by *followers*. Generalize the definition of mutual friends to *mutual followers*. Comment on whether a similar statement for followers can be established in this case.

8.4 Problem Sets

8.4.1 Problem Set #1

(1) (See Exercise 1) A social network is said to be symmetric if the relation between users that is maintained as a part of the network, is symmetric. Consider a symmetric social network and let the symmetric relation maintained be that of "user A and B are *friends*" (like in the case of facebook). A user C is said to be a *mutual friend* of users A and B if, C is a friend of both A and B. Prove that - for any user A of the network who has at least two friends, there must exist two friends of A who has the same number of mutual friends with A.

Comment on whether symmetry is critical for your argument. Take the example of *instagram* where the symmetric relation of *friends* is replaced by *followers*. Generalize the definition of mutual friends to *mutual followers*. Comment on whether a similar statement for followers can be established in this case.