К-1. Базовые понятия комбинаторики

5 октября 2024 г. 12:57

Понятие о комбинаторике. Биекции.

★ Комбинаторика — наука, которая работает с дискретными объектами и отвечает на два основных вопроса:

- существует ли объект с заданными свойствами?
- сколько существует объектов с заданными свойствами?

Основные комбинаторные объекты

- натуральные числа
- множества
- функции
- графы
- слова
- алгебры
- геометрические фигуры и конфигурации

- Есть такие понятия, как
  - комбинаторная алгебра
  - комбинаторная геометрия
  - комбинаторные алгоритмы
  - комбинаторная теория графов
  - комбинаторика слов
  - ...
  - ★ в русскоязычной математической литературе есть еще термин комбинаторный анализ (синоним комбинаторики, к матанализу отношения не имеет)

Основной инструмент комбинаторики - построение биекций.

Биекции. 2 простых примера

Пример: на плоскости даны 5 точек с целочисленными координатами; доказать, что середина некоторого отрезка с концами в этих точках имеет целочисленные координаты

Решение: воспользуемся принципом Дирихле (кролики и клетки)

- ★ если $f : A \to B$ — функция и $|B| < |A| < \aleph_0$, то $\exists a_1, a_2 \in A : f(a_1) = f(a_2)$
- пусть A — заданное множество точек, $B = \{(1,1), (1,0), (0,1), (0,0)\}$, $f(x, y) = (x \bmod 2, y \bmod 2)$

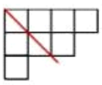
$\Rightarrow$ по принципу Дирихле найдутся (x_1, y_1) и (x_2, y_2) с одинаковым образом

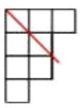
$\Rightarrow$ $x_1 + x_2$ и $y_1 + y_2$ — четные $\Rightarrow$ $\frac{x_1+x_2}{2}$ и $\frac{y_1+y_2}{2}$ целые □

решили задачу с помощью отсутствия биекции.

Пример: натуральное число n представляют в виде суммы меньших натуральных чисел разными способами (порядок слагаемых не важен); каких представлений больше — тех, в которых ровно k слагаемых или тех, в которых максимальное слагаемое равно k?

Решение: такие представления натуральных чисел (их называют разбиениями) обычно визуализируют при помощи диаграмм Ферре:


8=4+3+1


8=3+2+2+1

- Пусть f отображает каждое разбиение числа n с диаграммой D в разбиение с «отраженной» диаграммой D'
- f — биекция множества разбиений n на k слагаемых на множество разбиений n с максимальным слагаемым k □

крутой пример.

Построили биекцию, => мощность множеств одинакова.

Рекурентные формулы. Асимптотика.

Ещё о разбиениях натуральных чисел.

Рассмотрим более сложную задачу о разбиениях:

Задача: дано натуральное число n, вычислить число $P(n)$ различных разбиений n на натуральные слагаемые

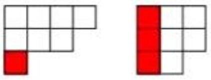
Возможное решение: найти рекуррентную формулу, выражающую $P(n)$ через $P(1), \ldots, P(n-1)$

- такая формула действительно есть: $P(n) = \sum_{k=1}^{n-1} \alpha_k P(k)$, где $\alpha_k \in \{-1, 0, 1\}$
- значение коэффициента α_k определяется пентагональной теоремой Эйлера
  - https://en.wikipedia.org/wiki/Pentagonal_number_theorem

- Попробуем найти формулу попроще
  - пусть $P_k(n)$ — число разбиений n на ровно k слагаемых; тогда $P(n) = \sum_{k=1}^{n} P_k(n)$

★ $P_k(n) = P_{k-1}(n-1) + P_k(n-k)$ при начальном условии $P_0(0) = 1$

Доказательство на картинке:

 - разбиения числа 8

- Разбиение n на k слагаемых либо содержит слагаемое 1 (левая диаграмма), либо нет (правая)
  - Закрашенные квадраты объясняют, почему первых разбиений $P_{k-1}(n-1)$, а вторых — $P_k(n-k)$ □

строк осталось столько же.

т.е. мы P(n) разбили на два класса и построили к ним биекцию.

Асимптотика числа разбиений

- Основной недостаток рекуррентной формулы — трудно оценить значение функции при большом n, не вычисляя все предыдущие значения

★ Часто на рекуррентных формулах не останавливаются, а выводят из них асимптотические оценки для функции

- Например, асимптотика функции $P(n)$ задается так:
  - $P(n) \sim \frac{1}{4\sqrt{3}\,n} e^{\pi\sqrt{2n/3}}$ (формула Харди–Рамануджана–Успенского)

... Как-то так выглядит «взрослая» комбинаторика

Биномиальные коэфициенты

- ★ База многих комбинаторных подсчетов — биномиальные коэффициенты $\binom{n}{k}$
  - в России принято писать «цешки» C_n^k
  - в англоязычной терминологии говорят «n choose k»; я буду говорить «n по k»

Определение 1: $\binom{n}{k}$ есть число k-элементных подмножеств n-элементного множества
  - говорят также «число способов выбрать k элементов из набора в n элементов»

Определение 2: $\binom{n}{k}$ есть коэффициент при $x^k y^{n-k}$ многочлена $(x+y)^n$
  - термин биномиальные коэффициенты как раз отсюда

Определение 3: $\binom{n}{k} = \Delta(n,k)$, где $\Delta(n,k)$ задана рекуррентным соотношением
$\Delta(n,k) = \Delta(n-1,k-1) + \Delta(n-1,k)$; $\Delta(n,0) = \Delta(n,n) = 1$
  - таблица значений функции Δ называется треугольником Паскаля

★ Эквивалентность определений:

★ перемножив n скобок $(x+y)(x+y)\cdots(x+y)$, получим сумму одночленов
одночлен $x^k y^{n-k}$ получится, если из k скобок выбрать x, а из остальных — y

⇒ коэффициент при $x^k y^{n-k}$ равен числу способов выбрать k скобок из n

⇒ определение 2 задает $\binom{n}{k}$

★ $\binom{n}{0} = 1$ (пустое подмножество), $\binom{n}{n} = 1$ (само множество)
k-элементные подмножества разобьем на 2 группы:
содержащие n-й элемент и не содержащие n-й элемент

- первых $\binom{n-1}{k-1}$, вторых $\binom{n-1}{k}$

⇒ $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$

⇒ $\binom{n}{k} = \Delta(n,k)$, т.е. определение 3 задает $\binom{n}{k}$ □

Свойства:

Св-во 1. (= определение 4) $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Д-во: определим способ выбора k-элементного подмножества:
зафиксируем перестановку на $[1..n]$ (одну из $n!$) и возьмём первые k её элементов.

порядок первых k элементов / последних $(n-k)$ элементов не влияют на выбор.

⇒ каждое подмножество будет выбрано $k!(n-k)!$ раз.

⇒ всего подмн-в: $\frac{n!}{k!(n-k)!}$

Следствие: $\Delta(n,k) = \binom{n}{n-k}$ (умножение коммутативно)

Свойство 2.

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

Д-во:

$$2^n = (1+1)^n = \sum_{k=0}^{n} \binom{n}{k} 1^k \cdot 1^{n-k} = \sum_{k=0}^{n} \binom{n}{k} \boxtimes$$

Свойство 3.

$$\sum_{k \text{ чёт}} \binom{n}{k} = \sum_{k \text{ нечет}} \binom{n}{k} \quad (\text{при } n > 0)$$

Д-во:

$$0 = (-1+1)^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^k = \sum_{k \text{ чет}} \binom{n}{k} - \sum_{k \text{ нечёт}} \binom{n}{k} \quad \blacksquare$$

Свойство 4:

$$\sum_{k=0}^{n} k \binom{n}{k} = n \cdot 2^{n-1}$$

Д-во:

$$\sum_{k=0}^{n} \underbrace{k \binom{n}{k}}_{\substack{\uparrow \\ \text{суммируем} \\ \text{все эл-ты} \\ \text{во всех подмн-вах}}} = \sum_{B \subseteq [1..n]} |B| = \sum_{B \subseteq [1..n]} |\overline{B}| = \sum_{B \subseteq [1..n]} \frac{|B| + |\overline{B}|}{2} = \frac{1}{2} \sum_{B \subseteq [1..n]} n = \frac{1}{2} n \underbrace{2^n}_{\substack{\text{число} \\ \text{всех подмн-в}}} = n 2^{n-1}$$

Биномиальные коэф-ы. Асимптотика

Вопрос: чему равен $\binom{n}{k}$ при больших n, k, т.е. асимптотически?

- сумма $(n+1)$ биномиальных коэффициентов равна 2^n, т.е. наибольший из них имеет порядок между $\frac{2^n}{n}$ и 2^n
- если k — константа, то $\binom{n}{k}$ — полином k-й степени

★ Для более точных оценок есть формула Стирлинга: $n! \sim \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n$

- доказательство — матан с двумя красивыми интегралами

- Оценим центральный биномиальный коэффициент:
  - n четное, при нечетном n аналогичную формулу для $\binom{n}{(n-1)/2}$ выведите сами

$$\binom{n}{n/2} = \frac{n!}{\left(\frac{n}{2}\right)!\left(\frac{n}{2}\right)!} \sim \frac{\sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n}{\sqrt{\pi n} \cdot \left(\frac{n/2}{e}\right)^{n/2} \sqrt{\pi n} \cdot \left(\frac{n/2}{e}\right)^{n/2}} = \sqrt{\frac{2}{\pi n}} \cdot 2^n$$

- Еще одна оценка:

$$\binom{n}{n/3} = \frac{n!}{\left(\frac{n}{3}\right)!\left(\frac{2n}{3}\right)!} \sim \frac{\sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n}{\sqrt{\frac{2}{3}\pi n} \cdot \left(\frac{n/3}{e}\right)^{n/3} \sqrt{\frac{4}{3}\pi n} \cdot \left(\frac{2n/3}{e}\right)^{2n/3}} = \frac{3}{2}\sqrt{\frac{1}{\pi n}} \cdot \left(\frac{3}{\sqrt[3]{4}}\right)^n$$

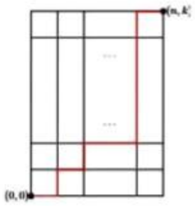
  - заметим, что $\left(\frac{3}{\sqrt[3]{4}}\right)^n \approx 1.89^n \ll 2^n$

- Более подробно про асимптотику биномиальных коэффициентов — в теории вероятностей при изучении схемы Бернулли и биномиального распределения

Числа Каталана

Пути в целочисленной решетке

- Рассмотрим специальный вид графа — целочисленную решетку (grid)
  - вершины (n, k)-решетки — целочисленные точки (x, y) на плоскости, $0 \leqslant x \leqslant n, 0 \leqslant y \leqslant k$
  - каждая пара вершин, находящихся на расстоянии 1, соединена ребром:



Вопрос: Сколько существует кратчайших путей из $(0, 0)$ в (n, k)?

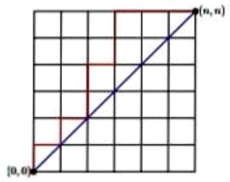
Решение: кратчайший путь имеет длину $n + k$ и состоит из n горизонтальных и k вертикальных ребер

$\Rightarrow$ Путь записывается словом длины $n+k$ в алфавите $\{\uparrow, \rightarrow\}$

$\Rightarrow$ Число путей равно числу способов выбрать позиции символов $\rightarrow$, т.е. $\binom{n+k}{n}$ $\square$

Числа Каталана

Путь в (n, n)-решетке — верхний, если он не опускается ниже диагонали $y = x$:



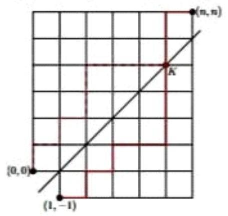
- Число Каталана C_n — это число верхних путей в (n, n)-решетке
- $C_0 = C_1 = 1, C_2 = 2, C_3 = 5, C_4 = 14, C_5 = 42, C_6 = 132, \dots$

Теорема

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

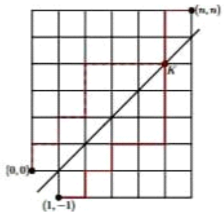
Доказательство: Поскольку количество всех путей в (n, n)-решетке известно из предыдущего слайда и равно $\binom{2n}{n}$, вместо верхних путей будем подсчитывать все остальные (назовем их неверхними) $\Longrightarrow$

Добавим к (n, n)-решетке один ряд снизу и проведем нижнюю диагональ $y = x-1$:



- Рассмотрим $(n-1, n+1)$-решетку с углами $(1, -1)$ и (n, n)
  - любой путь P из $(1, -1)$ в (n, n) начинается ниже нижней диагонали, а заканчивается выше ее
  - $\Rightarrow$ P пересекает эту диагональ; рассмотрим первую (нижнюю) точку касания K
  - фрагмент P ниже K отразим относительно нижней диагонали
    - остальная часть P не изменяется
  - получим путь P' из $(0, 0)$ в (n, n)
  - ★ P' — неверхний, потому что проходит через точку K
- ★ Положив $f(P) = P'$ для всех P, получили функцию из множества путей в $(n-1, n+1)$-решетке во множество неверхних путей в (n, n)-решетке

★ Положив $f(P) = P'$ для всех P, получили функцию из множества путей в $(n-1, n+1)$-решетке во множество неверхних путей в (n, n)-решетке



- ★ f — инъекция:
  - пути P_1 и P_2 отличаются фрагментом до точки K или фрагментом после нее
  - ⇒ $f(P_1)$ и $f(P_2)$ тоже отличаются этим фрагментом
- ★ f — сюръекция:
  - для неверхнего пути P' возьмем самую левую точку K на диагонали $y = x - 1$
  - отразим фрагмент P' до точки K относительно этой диагонали
  - ⇒ полученный путь P будет прообразом P' относительно f

⇒ f — биекция

- ⇒ число неверхних путей в (n, n)-решетке равно числу путей в $(n-1, n+1)$-решетке

⇒ Число верхних путей в (n, n)-решетке есть

$$C_n = \binom{2n}{n} - \binom{2n}{n-1} = \frac{(2n)!}{n!n!} - \frac{(2n)!}{(n-1)!(n+1)!} = \frac{(n+1)(2n)! - n(2n)!}{n!(n+1)!} = \frac{1}{n+1}\binom{2n}{n}$$ □

C_n — это...

★ Число $Par(n)$ правильных расстановок n пар скобок

- правильность расстановки проверяется простым правилом:
- ★ в расстановке левых скобок столько же, сколько правых, а в любом префиксе расстановки — не меньше чем правых
- ! докажите достаточность этого условия

Доказательство: верхний путь в (n, n)-решетке, закодированный как слово над $\{\uparrow, \rightarrow\}$, удовлетворяет тому же самому условию, если положить $\uparrow = ($, $\rightarrow =)$

⇒ значит, между этими множествами есть биекция и $Par(n) = C_n$ □

★ Число $Tr(n+1)$ полных бинарных корневых деревьев с $n{+}1$ листьями

- полное означает, что каждая вершина в дереве имеет 0 либо 2 детей

Доказательство: начиная с корня, обойдем дерево в глубину, спускаясь из каждой вершины вначале в левого ребенка, а затем в правого

- запишем порядок обхода ребер, кодируя левое ребро как (, а правое — как)
- в бинарном дереве с $n{+}1$ листьями $2n$ ребер (n правых и n левых)
- по расстановке скобок можно однозначно восстановить дерево

⇒ $Tr(n+1) = Par(n) = C_n$ □

- Дальнейшие примеры — на практике
  - желающие могут заглянуть на `https://oeis.org/A000108` (не для слабонервных)