КОМБИНАТОРНЫЕ ЗАДАЧИ ПЕРЕЧИСЛЕНИЯ В ОЛИМПИАДАХ ПО ПРОГРАММИРОВАНИЮ

М.И. Киндер[a],

[a]E-mail: mkinder@rambler.ru; Казанский Федеральный университет

Аннотация. *В статье обсуждаются классические комбинаторные объекты, для некоторых из них выводятся новые рекуррентные соотношения. Предлагаемые задачи связаны общим подходом, имеют общую алгоритмическую структуру решения. Эти задачи могут быть использованы на олимпиадах по программированию среди студентов и школьников.*

Abstract. M.I. Kinder. Combinatorial problems of enumeration in informatics competitions.

This paper describes a recursive approach to the enumeration of 'non-crossing' geometric configurations built on vertices of a convex polygon in the plane. Some formulae are given for the enumeration of certain types of dissections of the convex polygon by non-crossing diagonals. Consequences are counting results for graphs, dissections and partitions. We give a short and elementary proof of the formulas for classical numbers of polygon dissections. Most tasks can be used in the specific scope of teaching and learning informatics through olympiads and other competitions. We hope that some classes of such tasks would enlarge scope of tasks for use in informatics olympiads at various levels.

Keywords: *dissections of the convex polygon, polygon dissections, olympiads in informatics, tasks.*

Введение. Конструирование и анализ комбинаторных алгоритмов является одной из важнейших задач теоретической информатики. Комбинаторные числа и полиномы, возникающие при решении многих комбинаторных задач, встречаются в различных комбинаторных конфигурациях и являются классическими, хорошо изученными объектами дискретной математики [4]. Предлагаемые в статье перечислительные задачи объединены общим подходом и, по сути, являются задачами, которые можно использовать в олимпиадах по информатике и программированию. Алгоритмы решения опираются на рекуррентные соотношения, которые особенно удобны для применения метода динамического программирования.

Пути Моцкина.

Задача 1. *Необходимо подсчитать количество способов провести произвольное число непересекающихся хорд, соединяющих n точек на окружности.*

Пусть $m[n]$ — количество способов нарисовать непересекающиеся отрезки с вершинами в заданных n точках. Выделим одну из точек, например, первую, и разобьем множество всех способов на два непересекающихся класса: те, которые содержат отрезок, выходящий из точки 1, и те, в которых такого отрезка нет.

Количество способов первого класса, очевидно, равно $m[n-1]$.

Все способы второго класса содержат некоторый отрезок, выходящий из точки 1. Пусть, это будет отрезок, соединяющий точки с номерами 1 и $k+1$. По разные стороны от этого отрезка расположено $k-1$ и $n-k-1$ точек. Количество способов провести непересекающиеся хорды на первой дуге равно $m[k-1]$; количество таких способов на второй дуге равно $m[n-k-1]$. По правилу произведения общее число всех таких расстановок равно $m[k-1]\cdot m[n-k-1]$. Суммируя по всем k от 1 до $n-1$, получим окончательную формулу для подсчета чисел $m[n]$.

Теорема 1. *Количество способов провести непересекающиеся хорды с концами в заданных n точках окружности равно*

$$m[n] = m[n-1] + \sum_{k=1}^{n-1} m[k-1]\cdot m[n-k-1],$$

с начальным условием $m[0] = 1$.

Приведем несколько первых значений последовательности $m[n]$:

$$1,\ 1,\ 2,\ 4,\ 9,\ 21,\ 51,\ 127, \ldots .$$

Последовательность $m[n]$ имеет важную реализацию, связанную с путями Моцкина ([5], [6]) на плоскости. Рассмотрим целочисленную решётку в положительном квадранте. *Путём Моцкина* называется непрерывная ломаная в верхней полуплоскости, составленная из векторов $(1;1)$, $(1;0)$ и $(1;-1)$, начинающаяся в начале координат и заканчивающаяся на оси абсцисс.

Несложно установить взаимно однозначное соответствие между путями Моцкина и количеством расстановок непересекающихся хорд. Поясним это соответствие на примере шести точек, занумерованных числами от 1 до 6. Рассмотрим одну из возможных расстановок из двух непересекающихся хорд 16 и 35. Путь Моцкина длины 6, который соответствует этой расстановке, строится следующим образом. Первый и третий участки пути Моцкина, соответствующие началам хорд, представляют собой вектор $(1;1)$. Шестой и пятый участки, соответствующие концам хорд, — это векторы $(1;-1)$. Наконец, второй и четвёртый участки пути Моцкина, соответствующие точкам окружности, из которых не выходят хорды, — это «горизонтальные» векторы $(1;0)$.

В энциклопедии [1] целочисленных последовательностей количество путей Моцкина заданной длины $n \geq 0$ задаётся последовательностью A001006. Приведем явную формулу для чисел $m[n]$ через биномиальные коэффициенты:

$$m[n] = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{1}{k+1} \, C_n^{2k} \, C_{2k}^{k}.$$

Пути Моцкина высоты k.

Разреженной скобочной последовательностью называется конечная последовательность символов «(», «)» и «0», которая формально определяется следующим образом:

- пустая строка считается разреженной скобочной последовательностью;
- если S и T — разреженные скобочные последовательности, то строки $0S$, $S0$, (S) и ST также являются разреженными скобочными последовательностями.

Глубиной разреженной скобочной последовательности называется максимальная разность между количеством открывающихся и закрывающихся скобок в префиксе последовательности. (Префиксом строки S называется строка, которую можно получить из S удалением некоторого количества последних символов. Например, префиксами строки «$ABCAB$» является пустая строка, а также строки «A», «AB», «ABC», «$ABCA$» и «$ABCAB$»). Так, глубина последовательности «(0)(0())0» равна двум, поскольку среди префиксов есть строка «(0)(0(», которая содержит три открывающиеся и одну закрывающуюся скобки.

Задача 2. *Требуется по заданным значениям n и k вычислить количество разреженных скобочных последовательностей из n символов, которые имеют глубину вложения скобок, равную k.*

Пусть $m[n][k]$ — количество разреженных правильных скобочных последовательностей-строк из n символов с глубиной вложения скобок, *не превышающей* k. Тогда ответом на задачу будет значение выражения $m[n][k] - m[n][k-1]$.

Все рассматриваемые скобочные последовательности длины n разбиваются на два непересекающихся класса: строки, у которых первый символ «0», и строки, у которых этот символ — открывающаяся скобка «(». В первом случае количество таких строк, очевидно, равно $m[n-1][k]$. Скобочные последовательности второго класса начинаются с символа «(», которому соответствует некоторая закрывающая скобка «)». Между этими скобками и вне этих скобок получились две подстроки, состоящие соответственно из i и $n-i-2$ символов, где i принимает все возможные значения от 0 до $n-2$. Каждая подстрока является разреженной правильной скобочной последовательностью, у первой из них глубина вложения скобок не более $k-1$, а у второй — не более k. Общее число всех таких комбинаций равно $m[i][k-1] \cdot m[n-i-2][k]$. Суммируя по всем i от 0 до $n-2$, получим окончательную формулу для подсчета чисел $m[n][k]$.

Теорема 2. *Количество разреженных скобочных последовательностей из n символов с глубиной вложения скобок k равно $m[n][k] - m[n][k-1]$, где*

$$m[n][k] = m[n-1][k] + \sum_{i=0}^{n-2} m[i][k-1] \cdot m[n-i-2][k],$$

при этом $m[i][0] = m[0][j] = 1$ *для всех* $i \geq 0$ *и для всех* $j \geq 0$.

Приведем несколько первых значений последовательности $m[n][k]$:

$m[1][0] = 1.$
$m[2][0] = 1;\quad m[2][1] = 1.$
$m[3][0] = 1;\quad m[3][1] = 3;\quad m[3][2] = 0.$
$m[4][0] = 1;\quad m[4][1] = 7;\quad m[4][2] = 1;\quad m[4][3] = 0.$
$m[5][0] = 1;\quad m[5][1] = 15;\quad m[5][2] = 5;\quad m[5][3] = 0;\quad m[5][4] = 0.$
$m[6][0] = 1;\quad m[6][1] = 31;\quad m[6][2] = 18;\quad m[6][3] = 1;\quad m[6][4] = 0;\quad m[6][5] = 0.$

Несложно установить соответствие между путями Моцкина и разреженными скобочными последовательностями. Для этого нужно сопоставить вектору $(1; 1)$ левую скобку, вектору $(1; -1)$ — правую скобку, а вектору $(1; 0)$ — символ «0». Тогда условие того, что путь лежит в верхней полуплоскости и заканчивается на оси абсцисс, и есть в точности условие разреженной скобочной структуры.

Разреженная скобочная последовательность с глубиной вложения скобок k соответствует пути Моцкина, который не поднимается выше уровня k.

В энциклопедии [1] количество путей Моцкина длины $n \geq 0$ и высоты $k \geq 0$ задаётся последовательностью A097862.

Пути Дика высоты k.

Если в разреженной скобочной последовательности удалить все символы «0», то полученная строка будет *правильной скобочной последовательностью.* Для таких строк можно также определить понятие *глубины* правильной скобочной последовательности как максимальной разности между количеством открывающихся и закрывающихся скобок в префиксе последовательности.

Задача 3. *Требуется по заданным значениям n и k вычислить количество правильных скобочных последовательностей из n символов, которые имеют глубину вложения скобок, равную k.*

Теорема 3. *Количество правильных скобочных последовательностей из n символов с глубиной вложения скобок k равно $m[n][k] - m[n][k-1]$, где*

$$m[n][k] = \sum_{i=0}^{n-1} m[i][k-1] \cdot m[n-i-1][k],$$

при этом $m[i][0] = 0$ *для всех* $i > 0$ *и* $m[0][j] = 1$ *для всех* $j \geq 0$.

Приведем несколько первых значений последовательности $m[n][k]$:

$m[1][1] = 1.$
$m[2][1] = 1;\quad m[2][2] = 1.$
$m[3][1] = 1;\quad m[3][2] = 3;\quad m[3][3] = 1.$
$m[4][1] = 1;\quad m[4][2] = 7;\quad m[4][3] = 5;\quad m[4][4] = 1.$
$m[5][1] = 1;\quad m[5][2] = 15;\quad m[5][3] = 18;\quad m[5][4] = 7;\quad m[5][5] = 1.$

Последовательность $m[n][k]$ связана с *путями Дика* на целочисленной решётке, которые определяются так же, как и пути Моцкина, только они не содержат «горизонтальных» векторов $(1; 0)$. Пути Дика также начинаются в начале координат и заканчиваются на оси абсцисс.

Как и в предыдущем случае, несложно установить, что правильная скобочная последовательность с глубиной вложения скобок k соответствует пути Дика, который не поднимается выше уровня k. В энциклопедии [1] количество путей Дика длины n и высоты k $(1 \leq k \leq n)$ задаётся последовательностью A080936.

Замечание. Задача о подсчёте количества правильных скобочных последовательностей с заданной глубиной вложения встречалась на различных олимпиадных соревнованиях по информатике. По-видимому, впервые она появилась на 5-й Балтийской олимпиаде по информатике 1999 г. [3], а затем на Всероссийской олимпиаде по информатике 2007 г.

Пути Моцкина с заданным числом шагов вверх.

Задача 4. *Пусть на окружности задано n точек. Необходимо подсчитать количество способов провести ровно k непересекающихся линий, каждая из которых соединяет ровно две точки.*

Пусть $m[n][k]$ — количество способов нарисовать k непересекающихся линий с вершинами в заданных n точках. Выделим одну из точек, например, первую, и разобьем все способы на два непересекающихся класса: те, которые содержат соединительную линию, выходящую из точки 1, и те, в которых нет такой линии.

Количество способов первого класса, очевидно, равно $m[n-1][k]$.

Все способы второго класса содержат некоторую линию, выходящую из точки 1. Пусть, это будет линия, соединяющая точки с номерами 1 и $i+2$. По разные стороны от этой линии расположено i и $n-i-2$ точки. Предположим, что на дуге, содержащей i точек, проведено j соединительных линий, и значит, на дуге с $n-i-2$ точками проведено $k-j-1$ таких линий. (Учтём, что мы уже провели одну линию, соединяющую точки 1 и $i+2$.) Количество способов провести j непересекающихся линий на первой дуге равно $m[i][j]$; количество таких способов на второй дуге равно $m[n-i-2][k-j-1]$. По правилу произведения общее число всех таких комбинаций равно $m[i][j] \cdot m[n-i-2][k-j]$. Суммируя по всем i от 0 до $n-2$ и по всем j от 0 до $k-1$, получим окончательную формулу для подсчета чисел $m[n][k]$.

Теорема 4. *Количество способов провести k непересекающихся хорд с концами в заданных n точках окружности равно*

$$m[n][k] = m[n-1][k] + \sum_{i=0}^{n-2} \sum_{j=0}^{k-1} m[i][j] \cdot m[n-i-2][k-j-1],$$

при этом $m[i][0] = 1$ для всех $i \geq 0$ и $m[0][j] = 0$ для всех $j \geq 1$.

Приведем несколько известных значений последовательности $m[n][k]$:

$$m[n][1] = \frac{1}{2} n(n-1)$$

для всех $n \geq 1$, так как имеется $\frac{1}{2}n(n-1)$ способов выбрать две точки и провести между ними одну линию.

$$m[2n][n] = C_n = \frac{1}{n+1} C_{2n}^{n} = \frac{(2n)!}{(n+1)!\, n!}$$

для всех $n \geq 1$, где C_n — n-ое число Каталана, так как для $2n$ точек и n соединительных хорд мы получаем известную интерпретацию чисел Каталана.

В энциклопедии [1] ненулевые числа $m[n][k]$ образуют последовательность A080159, они совпадают также с количествами путей Моцкина длины n, у которых имеется ровно k шагов вверх [2].

Приведем явную формулу для ненулевых чисел $m[n][k]$ (при $n \geq 2k$):

$$m[n][k] = \frac{n!}{(n-2k)!\, k!\, (k+1)!}.$$

Разрезание выпуклых многоугольников на k частей.

Задача 5. *Сколько существует способов разрезать выпуклый n-угольник ровно на k частей с помощью диагоналей, непересекающихся внутри многоугольника?*

Пусть $m[n][k]$ — количество способов разрезать n-угольник на k частей.

Занумеруем вершины многоугольника числами от 1 до n в порядке его обхода по часовой стрелке. Выделим одно из ребер, например, ребро 1-2, и рассмотрим ту часть K, которая содержит это ребро. Разобьем все способы разрезания на два непересекающихся класса:

1) те, у которых эта часть содержит ребро 2-j, выходящее из вершины 2;
2) и те, у которых в части K нет ребра из вершины 2.

В первом случае часть K содержит ребра 1-2 и 2-j. Все такие разбиения можно разбить на два множества:

1a) те, в которых K — это треугольник 1-2-j. Удалив эту часть из разбиения, получим аналогичные задачи для многоугольников с числом сторон $(j-1)$ и $(n-j+2)$. Если в первом из них i частей, то второй многоугольник содержит $(k-i-1)$ частей. Количество разбиений этих многоугольников равно

$$m[j-1][i]\cdot m[n-j+2][k-i-1].$$

1*б*) те, в которых K не является треугольником 1-2-j. Тогда можно «слить» ребра 1-2 и 2-j в одно ребро 1-j. Это означает, что из части K можно удалить треугольник 1-2-j, получив при этом аналогичную задачу для $(j-1)$-угольника 2-3- … -j и $(n-j+2)$-угольника 1-j- … -n. Если в первом из них i частей, то второй многоугольник содержит $(k-i)$ частей. Количество разбиений теперь равно

$$m[j-1][i]\cdot m[n-j+2][k-i].$$

Осталось просуммировать найденные количества способов по всем значениям j от 4 до n и всем i от 0 до $k-1$. (Учтём, что в $(j-1)$-многоугольнике количество частей не может быть равно k, так как по другую сторону от ребра 2-j есть по крайней мере одна часть.) Имеем

$$\sum_{j=4}^{n}\sum_{i=0}^{k-1} m[j-1][i]\cdot(m[n-j+2][k-i-1]+m[n-j+2][k-i]).$$

Во втором случае, кроме ребра 1-2, часть K содержит также и ребро 2-3. Все такие разбиения можно разбить на два множества:

2*а*) те, в которых K — это треугольник 1-2-3. Удалив эту часть из разбиения, получим аналогичную задачу для $(n-1)$-угольника, который требуется разрезать на $(k-1)$ частей. Количество таких разбиений равно $m[n-1][k-1]$.

2*б*) те, в которых K не является треугольником 1-2-3. Тогда можно слить ребра 1-2 и 2-3 в одно ребро 1-3. Это означает, что из части K можно удалить треугольник 1-2-3, получив при этом аналогичную задачу для $(n-1)$-угольника 1-3-4- … -n, который нужно разрезать на k частей. Количество таких разбиений равно $m[n-1][k]$. Значит, количество способов во втором случае равно $m[n-1][k-1]\cdot m[n-1][k]$.

Теорема 5. *Количество способов разрезания выпуклого n-угольника ровно на k частей с помощью непересекающихся внутри диагоналей равно*

$$m[n][k]=\sum_{j=3}^{n}\sum_{i=0}^{k-1} m[j-1][i]\cdot(m[n-j+2][k-i-1]+m[n-j+2][k-i]),$$

при этом $m[n][1]=1$; $m[n][0]=0$ *при* $n\neq 2$ *и* $m[2][0]=1$.

Приведем несколько первых значений последовательности $m[n][k]$:

$m[2][0]=1$.
$m[3][0]=0$; $m[3][1]=1$.
$m[4][0]=0$; $m[4][1]=1$; $m[4][2]=2$.
$m[5][0]=0$; $m[5][1]=1$; $m[5][2]=5$; $m[5][3]=5$.
$m[6][0]=0$; $m[6][1]=1$; $m[6][2]=9$; $m[6][3]=21$; $m[6][4]=14$.

В энциклопедии [1] последовательность $m[n][k]$ имеет номер A033282.

Приведём явную формулу для чисел $m[n][k]$ через биномиальные коэффициенты:

$$m[n][k]=\frac{1}{k}\,C_{n-3}^{k-1}\cdot C_{n+k-2}^{k-1},\ \text{где } n\geq 3 \text{ и } 1\leq k\leq n-2.$$

Литература

[1] http://oeis.org — on-line энциклопедия целочисленных последовательностей.

[2] http://oeis.org/A055151.

[3] http://www.ut.ee/boi/?item=boi.about.

[4] Липский В. *Комбинаторика для программистов.* — М.: Мир, 1988. — 213 с.

[5] Donaghey R., Shapiro L. *Motzkin Numbers*, Journal of Combinatorial Theory. 1977, Series A, 23, pp. 291-301.

[6] Стенли Р. *Перечислительная комбинаторика. Деревья, производящие функции и симметрические функции.* — М.: Мир, 2009. — 767 с. (Задача 6.38, с. 306.)