Дендрокомбинаторика

А. ЗАБОЛОТСКИЙ

Продолжаем начатое в 7-м и 8-м номерах «Кванта» за прошлый год исследование Онлайн-энциклопедии целочисленных последовательностей – OEIS.

Огромный пласт интересных последовательностей в OEIS связан с подсчетом количества тех или иных сущностей, которые конструируются из заданного количества элементов. Особенно интересно, когда одна и та же последовательность имеет несколько разных комбинаторных интерпретаций. В одних случаях мы умеем проводить подсчет только прямым ручным или компьютерным перебором, в других же удается задействовать всю мощь науки о подсчетах – перечислительной комбинаторики.

Деревья и леса

Возьмем, к примеру, деревья. Здесь *дерево* – это связный граф без циклов. Несвязное объединение нескольких деревьев (или мультимножество деревьев, т.е. множество, в котором элементы могут повторяться) образует *лес*. Если же в дереве выделить одну вершину и назвать ее корнем, получится *дерево с корнем* (рис. 1).

Последовательность A000055 подсчитывает количество различных деревьев с n (неразличимыми) вершинами. Для небольших n можно аккуратным перебором найти все деревья; они изображены на рисунке 2. Получаем, что существует 1, 1, 1, 2,

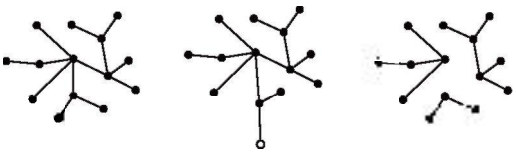

Рис. 1

DOI: https://doi.org/10.4213/kvant20230303

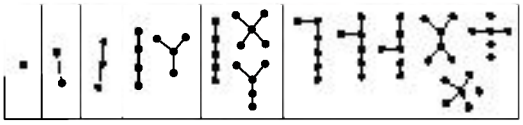

Рис. 2

3, 6, 11, 23,... деревьев с 1, 2, 3, 4, 5, 6, 7, 8,... вершинами.

Комбинируя между собой деревья и учитывая, что *пустой граф* с 0 вершинами тоже является лесом, можно убедиться, что лесов с 0, 1, 2, 3, 4, 5, 6, 7,... вершинами существует 1, 1, 2, 3, 6, 10, 20, 37,... штук – это последовательность A005195. Можно ли как-то связать ее с предыдущей? Ответ на это дает равенство, которое в OEIS называют преобразованием Эйлера.

Преобразование Эйлера

Предположим, имеется n неразличимых элементов (например, отдельных вершин графа) и есть a_n способов придать этим элементам некую структуру (например, собрать из вершин дерево). Тогда количество способов разбить n элементов на сколько-то групп и придать структуру элементам в каждой группе (в данном случае – сделать лес из n вершин) – обозначим это число b_n – связано с числами a_n преобразованием Эйлера

$$b_n = \frac{1}{n}\sum_{k=1}^{n} b_{n-k}\sum_{d|k} d a_d. \qquad (1)$$

Здесь внутреннее суммирование идет по всем делителям d числа k. При этом всегда $b_0 = 1$.

В этом подсчете группы считаются различными лишь тогда, когда различаются структуры, приданные элементам этих групп. В примере с лесами это означает, что важно лишь то, из каких деревьев состоит лес; ни то, в каком порядке они

записаны, ни какая-либо другая информация не учитывается.

Именно равенство (1) связывает A000055 и A005195. Более общо, если a_n подсчитывает *связные* графы на n вершинах, обладающие неким свойством, то b_n подсчитывает *все* графы, обладающие тем же свойством. Обобщая еще сильнее, если a_n подсчитывает количество неких структур на n элементах, то b_n подсчитывает количество *мультимножеств* таких структур.

Другой пример: пусть структурой, о которой идет речь, является множество не равных, но неразличимых элементов без всякой дополнительной информации. Тогда $a_n = 1$ при всех n, поскольку множества из неразличимых элементов отличаются лишь их количеством и для любого n существует лишь одно множество из n элементов. В этом случае преобразование Эйлера дает последовательность A000041 – количество разбиений числа n в сумму натуральных чисел (порядок слагаемых не имеет значения) или, что то же самое, количество разбиений множества из n неразличимых элементов на непересекающиеся подмножества: 1, 1, 2, 3, 5, 7, 11, 15, ...

Равенство, связывающее количество связных и несвязных графов, приведено в книге Харари и Палмера [1], различные «родственники» его обсуждаются в книге Стенли [2], а почти точно в таком виде равенство (1) приведено во вводной части еще бумажной «Энциклопедии целочисленных последовательностей» [3]. Обычно это равенство получают с помощью аппарата производящих функций, повсеместно используемого в перечислительной комбинаторике. Мы докажем его иначе.[1]

В этом нам поможет некий трюк, который мы продемонстрируем сначала на числах Каталана C_n (см. [4]). Числа Каталана образуют последовательность A000108: 1, 1, 2, 5, 14, 42, 132, 429, ... Помимо прочего они подсчитывают количество правильных скобочных строк, т.е. таких строк из n пар скобок, что каждый открывающей скобке однозначно соответствует закрывающая. Например, скобочная строка $s = ((())())(()()())$ – правильная, а $w =)(()$ – неправильная. Выделяем в правильной скобочной строке крайнюю правую закрывающую скобку вместе с соответствующей открывающей и обращаем внимание на скобочную строку s_1 до этой пары скобок и s_2 внутри нее, например: $s = ((())())(()()()) = s_1(s_2)$, где $s_1 = ((())())$ и $s_2 = ()()()$. Каждая непустая правильная скобочная строка единственным образом разлагается на s_1 и s_2 как $s_1(s_2)$, и из каждых двух правильных скобочных строк s_1 и s_2 (возможно, пустых) однозначно восстанавливается $s_1(s_2)$. При этом s_1 и s_2 независимы – между ними не требуется никаких соотношений, кроме того, что суммарно в них должно быть n пар скобок. Обозначаем количество пар скобок в s_2 через k, перебираем все возможные варианты k, а для каждого k – все пары из двух правильных скобочных строк и получаем рекуррентное соотношение для чисел Каталана:

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}. \qquad (2)$$

Для доказательства равенства (1) попробуем похожим образом разбить на две независимые части лес (более общо – мультимножество структур).

Возьмем лес из n вершин и каким-нибудь образом расположим все его деревья слева направо, сгруппировав одинаковые. Отметим какую-нибудь одну вершину в этом графе (таким образом выбрать лес и отметить в нем вершину можно nb_n способами). Выделим блок из нескольких деревьев, а именно дерево с отмеченной вершиной и все такие же деревья *справа* от него (рис. 3). Таким образом мы разбили лес на две части: остаточную невыделенную, представляющую собой лес (возможно, пустой), и выделенную, представляющую со-

Рис. 3

[1] Наше доказательство равенства (1) – точнее, его специального случая, равенства (3) – также независимо нашли М.Тихомиров и Д.Гринберг.

бой блок из некоторого ненулевого количества одинаковых деревьев. Количество вершин в каждом дереве выделенного блока обозначим d, а общее количество вершин в блоке – k; так, на рисунке 3 имеем $d = 5$, $k = 15$, $n = 57$.

Заметим, что даже если забыть про то, как мы расположили деревья в лесу и какая именно вершина в блоке был изначально выделена, то из остаточного леса и блока одинаковых деревьев можно однозначно собрать исходный лес, причем остаточный лес и блок одинаковых деревьев могут быть любыми, единственное ограничение – общее количество вершин должно равняться n (прямо как со скобочными строками!). В то же время любая пара из остаточного леса и блока одинаковых деревьев, получаемая из данного исходного леса, может быть получена ровно d способами – по числу вершин в самом левом дереве из блока. Поэтому если мы сосчитаем всевозможные пары из леса и блока одинаковых деревьев с суммарным числом вершин k, учтя каждую пару с весом d (определяется блоком), то мы получим в точности количество способов составить лес на n вершинах и выделить в нем одну вершину, оно равняется nb_n. Учитывая, что количество лесов на $n-k$ вершинах равно b_{n-k}, количество деревьев на d вершинах равно a_d и k должно делиться на d, получаем равенство (1).

Подсчет деревьев с корнем

Вернемся теперь к деревьям с корнем. Обозначим c_n количество деревьев с корнем на n вершинах. В них замечательно то, что каждая «растущая» из корня (рис. 4,*а*) «ветвь» дерева сама по себе может считаться деревом с корнем, поскольку она, несомненно, является деревом, а ее корнем естественно назначить вершину, ближайшую к корню исходного, «большого» дерева. Отсюда вытекает взаимно однозначное соответствие между деревьями с корнем из $n+1$ вершин и лесами из деревьев с корнем, имеющими в общей сложности n вершин: имея дерево с корнем, убираем из него корень, назначаем во всех образовавшихся деревьях корнями те вершины, что были связаны с прежним корнем, получаем лес; обратно, имея лес, добавляем к нему новую вершину, соединяем ее с корнями всех деревьев леса, назначаем ее новым корнем, получаем дерево с корнем. Это значит, что c_n является последовательностью, которая переходит сама в себя со сдвигом индекса на 1 при преобразовании Эйлера!

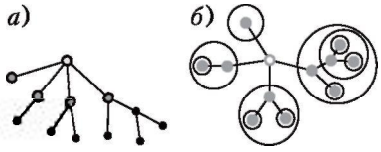

Рис. 4

Равенство (1) превращается в рекуррентное соотношение для c_{n+1} при $n > 0$, в которой не используются никакие дополнительные последовательности:

$$c_{n+1} = \frac{1}{n}\sum_{k=1}^{n} c_{n-k+1}\sum_{d|k} dc_d. \qquad (3)$$

При этом будем считать, что $c_0 = 0$, поскольку в дереве с корнем все-таки должна быть хоть одна вершина – корень, и $c_1 = 1$.

Последовательность 0, 1, 1, 2, 4, 9, 20, 48, ..., полученная таким образом, имеется в OEIS под номером A000081. Ее можно связать и с количеством деревьев без корня.

Деревья и конфигурации окружностей

Часто комбинаторные объекты и последовательности имеют несколько разных интерпретаций. Один пример мы видели только что – деревья с корнем из $n+1$ вершин и леса из деревьев с корнем из n вершин. Но деревья с корнем имеют и иную интерпретацию.

Возьмем дерево с корнем и для каждой вершины, кроме корня, нарисуем замкнутую кривую так, чтобы внутрь этой кривой попадала эта вершина и все другие вершины, находящиеся дальше ее от корня, вместе с соответствующими им кривыми. Ничто не мешает нам изобразить дерево на плоскости так, чтобы в качестве этих кривых можно было взять окружности (рис. 4,*б*). В результате плоскость оказалась разделена на области, в каждой из

которых находится ровно одна вершина, причем корень находится в единственной неограниченной, «внешней» области. С другой стороны, для любого расположения непересекающихся окружностей на плоскости мы можем нарисовать вершину внутри каждой окружности, еще одну вершину – корень – снаружи, соединить ребром каждую пару вершин, отделенных друг от друга только одной любой окружностью, и получить дерево с корнем. Если считать эквивалентными расположения окружностей, которые получаются друг из друга их движением и изменением размеров, но без пересечений, то таким образом оказывается установлено взаимно однозначное соответствие между расположением n окружностей на плоскости (или на диске) и деревом с корнем на $n+1$ вершинах.

Что же деревья без корня? Можно нарисовать на плоскости дерево и окружности так же, как в случае дерева с корнем, назначив любую вершину корнем. В этой конструкции, как и в случае дерева с корнем, две вершины соединены ребром тогда и только тогда, когда соответствующие им области разделены одной окружностью. В случае дерева с корнем снаружи всегда находится область, соответствующая вершине, которая ближе к корню, в случае же дерева без корня не должно быть разницы.

Существует операция, которая преобразует плоскость, разделенную окружностями на регионы, как раз так, что сохраняется набор областей и отношение между ними «граничит через окружность», но при этом может сделать любую область внешней. Эта операция – инверсия плоскости относительно окружности. Таким образом, каждое дерево без корня с $n+1$ вершинами соответствует конфигурации n не пересекающихся окружностей на плоскости с точностью до инверсий, перемещений и изменений размеров (без возникновения пересечений).

На рисунке 5,*а* показано действие одной такой инверсии. При инверсии меняются положение вершин и положение и размеры остальных окружностей, но поскольку конфигурация определена с точностью до

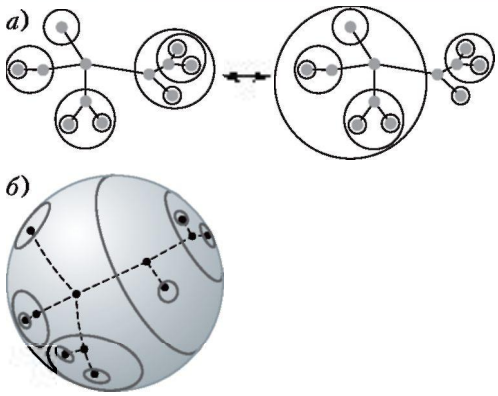

Рис. **5**

этих преобразований, то это несущественно. Важно лишь, что одна или несколько окружностей «выворачиваются наизнанку».

Или же можно дополнить плоскость бесконечно удаленной точкой (строго говоря, это необходимо, если мы рассматриваем все возможные инверсии) и через стереографическую проекцию свернуть плоскость в сферу, что вовсе устранит разницу между внутренностью и внешностью каждой окружности (рис. 5,*б*).

Соответственно, последовательности A000055 и A000081 подсчитывают не только деревья без корня и с корнем, но и конфигурации окружностей на плоскости и на сфере соответственно. Но лишь непересекающихся.

Конфигурации пересекающихся окружностей на плоскости – куда более сложная история, где в нашем распоряжении есть лишь компьютерный перебор; их подсчитывает последовательность A250001.

Список литературы

1. *Ф.Харари, Э.Палмер.* Перечисление графов. – М.: Мир, 1977.
2. *Р.Стенли.* Перечислительная комбинаторика, том 2. – М.: Мир, 2009.
3. *N.J.A.Sloane, S.Plouffe.* The Encyclopedia of Integer Sequences. – Academic Press, 1995.
4. *М.Гарднер.* Числа Каталана. – «Квант», 1978, №7.