Перечисление взвешенных плоских деревьев

А. К. ЗВОНКИН
Университет Бордо, Франция
e-mail: zvonkin@labri.fr

УДК 519.175.3

Ключевые слова: взвешенные деревья, перечисление.

Аннотация

У взвешенных деревьев рёбра снабжены положительными целочисленными весами. Мы перечисляем двукрашенные плоские взвешенные деревья в соответствии с их весом и количеством рёбер.

Abstract

A. K. Zvonkin, Enumeration of weighted plane trees, Fundamentalnaya i prikladnaya matematika, vol. 18 (2013), no. 6, pp. 135—144.

In weighted trees, all edges are endowed with positive integral weights. We enumerate weighted bicolored plane trees according to their weight and number of edges.

1. Введение

Взвешенные деревья представляют интерес как самостоятельный и естественный комбинаторный объект, а также играют важную роль в некоторых вопросах теории чисел (см., например, [1, 6, 7]).

Определение 1 (взвешенное дерево). *Взвешенное двукрашенное плоское дерево*, или *взвешенное дерево*, или просто *дерево* для краткости — это плоское дерево, вершины которого выкрашены в чёрный и белый цвет двудольным образом, а рёбрам которого приписаны положительные целочисленные *веса*. Вес дерева — это сумма весов всех его рёбер.

Степень вершины определяется как сумма весов всех прилегающих к ней рёбер. Очевидно, что сумма степеней чёрных вершин, а также сумма степеней белых вершин равны полному весу дерева. Пусть вес дерева равен n и оно имеет p чёрных вершин степеней $\alpha_1, \ldots, \alpha_p$ и q белых вершин степеней $\beta_1, \ldots, \beta_q$. В этом случае пара разбиений (α, β) числа n называется *паспортом* дерева.

Если веса всех рёбер дерева равны 1, то такое дерево называется *обычным*. Обычное дерево, которое получается из взвешенного дерева отбрасыванием всех весов, называется *топологическим деревом*.

Фундаментальная и прикладная математика, 2013, том 18, № 6, с. 135—144.
© 2013 *Центр новых информационных технологий МГУ,*
Издательский дом «Открытые системы»

Прилагательное *плоский* в этом определении означает, что мы рассматриваем наши деревья не как графы, а как карты на плоскости. Более точно, это означает, что циклический порядок рёбер в каждой вершине фиксирован и изменение этого порядка, вообще говоря, меняет дерево. Все деревья, рассматриваемые в этой заметке, являются плоскими, и в дальнейшем мы этот факт более не оговариваем.

Пример 2 (взвешенное дерево). На рис. 1 изображено взвешенное дерево. Его вес равен $n = 18$, а паспорт имеет вид $(\alpha, \beta) = (5^2 2^3 1^2, 7^1 6^1 4^1 1^1)$.

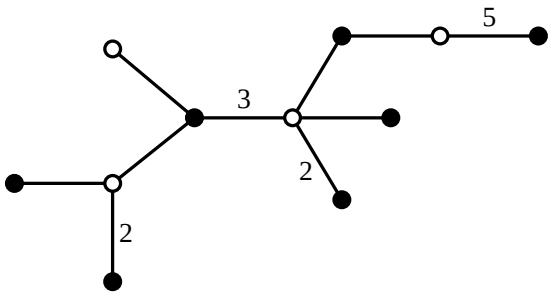

Рис. 1. Взвешенное дерево. Веса, не указанные явно, равны 1

Определение 3 (корневое дерево). Дерево с выделенным ребром называется *корневым*, а само выделенное ребро — его *корнем*. Мы считаем, что корень ориентирован в направлении от чёрной вершины к белой.

Цель этой статьи — перечисление корневых (двукрашенных плоских) деревьев.

2. Формулировка основной теоремы

Теорема 4 (перечисление взвешенных деревьев). *Пусть a_n обозначает число корневых двукрашенных плоских деревьев веса n. Тогда производящая функция*

$$f(t) = \sum_{n \geqslant 0} a_n t^n$$

равна

$$f(t) = \frac{1 - t - \sqrt{1 - 6t + 5t^2}}{2t} = \\ = 1 + t + 3t^2 + 10t^3 + 36t^4 + 137t^5 + 543t^6 + 2219t^7 + 9285t^8 + \dots . \quad (1)$$

Числа a_n удовлетворяют рекуррентному соотношению

$$a_0 = 1, \quad a_1 = 1, \quad a_{n+1} = a_n + \sum_{k=0}^{n} a_k a_{n-k} \quad \text{при} \quad n \geqslant 1. \tag{2}$$

Асимптотическое поведение чисел a_n при $n \to \infty$ выражается формулой

$$a_n \sim \frac{1}{2}\sqrt{\frac{5}{\pi}} \cdot 5^n n^{-3/2}. \tag{3}$$

Пусть $b_{m,n}$ обозначает число корневых взвешенных двуцветных плоских деревьев веса n с m рёбрами. Тогда производящая функция

$$h(s,t) = \sum_{m,n \geqslant 0} b_{m,n} s^m t^n$$

равна

$$h(s,t) = \frac{1 - t - \sqrt{1 - (2+4s)t + (1+4s)t^2}}{2st} =$$
$$= 1 + st + (s + 2s^2)t^2 + (s + 4s^2 + 5s^3)t^3 + (s + 6s^2 + 15s^3 + 14s^4)t^4 + \dots . \tag{4}$$

Явная формула для чисел $b_{m,n}$ выглядит следующим образом:

$$b_{m,n} = \binom{n-1}{m-1} \cdot \mathrm{Cat}_m = \binom{n-1}{m-1} \cdot \frac{1}{m+1}\binom{2m}{m}, \tag{5}$$

где Cat_m — это m-е число Каталана.

Обозначим через $|\operatorname{Aut}(T)|$ порядок группы автоморфизмов дерева T. Пусть c_n обозначает число неизоморфных некорневых деревьев T веса n, причём вклад дерева T в общую сумму равен $1/|\operatorname{Aut}(T)|$. Тогда

$$c_n = \sum_T \frac{1}{|\operatorname{Aut}(T)|} = \sum_{m=1}^{n} \frac{b_{m,n}}{m}, \tag{6}$$

где первая сумма берётся по всем неизоморфным некорневым деревьям T веса n.

Последовательность a_n присутствует в «Энциклопедии числовых последовательностей» [5] под индексом A002212. Она имеет множество различных интерпретаций, причём некоторые из них идут из химии. Среди этих интерпретаций со ссылкой на Р. Баше (2005 г.) упоминаются «мультидеревья»: они соответствуют нашим взвешенным деревьям. Р. Баше сообщил автору, что мультидеревья появились у него в качестве побочного продукта его топологических исследований. Некоторые из приведённых выше формул также могут быть найдены в указанной энциклопедии.

Пример 5 (деревья веса 4). На рис. 2 изображены деревья веса 4. Мы видим десять деревьев, но в действительности имеется 16 неизоморфных некоренных деревьев веса 4. В самом деле, при перемене чёрного и белого цветов четыре дерева остаются изоморфными самим себе, в то время как для шести

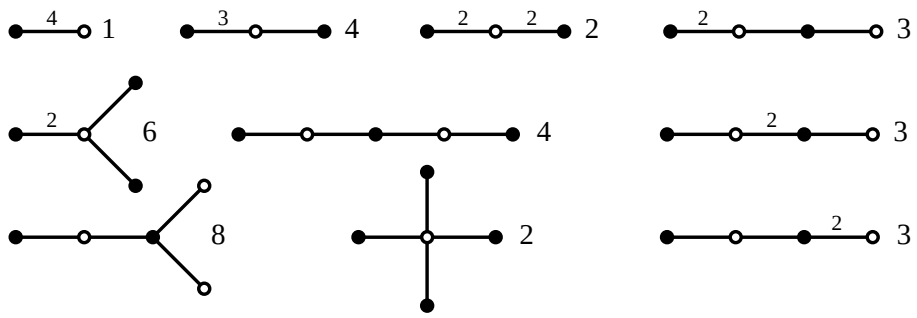

Рис. 2. Деревья веса 4. Рядом с каждым деревом указано количество возможных выборов корня (с учётом перекрашивания). Число корневых деревьев равно 36

оставшихся деревьев это свойство неверно. Таким образом, к нашим рисункам следовало бы добавить ещё шесть недостающих деревьев.

Рядом с каждым деревом указано число возможных выборов корня (с учётом возможного перекрашивания). Общее число деревьев равно 36, и это число — коэффициент a_4 перед t^4 в функции $f(t)$ (см. (1)). Среди этих 36 деревьев одно дерево состоит из единственного ребра, шесть деревьев имеют по два ребра, 15 деревьев имеют по три ребра и 14 деревьев имеют по четыре ребра. Эти числа являются коэффициентами многочлена $s + 6s^2 + 15s^3 + 14s^4$, стоящего перед t^4 в $h(s,t)$ (см. (4)).

Число c_4 в соответствии с формулой (6) равно

$$1 + \frac{6}{2} + \frac{15}{3} + \frac{14}{4} = 12\frac{1}{2}.$$

Действительно, среди 16 неизоморфных некорневых деревьев десять являются асимметричными, четыре имеют симметрию порядка 2 и два дерева имеют симметрию порядка 4, что даёт

$$10 + 4 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 12\frac{1}{2}.$$

Мы оставляем детали читателю.

Доказательство теоремы 4 приведено в разделе 4. Предварительная работа, необходимая для проведения доказательства, осуществляется в разделе 3.

3. Слова Дика и взвешенные слова Дика

Имеется стандартный способ кодирования корневых топологических (т. е. невзвешенных) плоских деревьев *словами Дика* и *путями Дика*. Мы совершаем обход вокруг дерева по часовой стрелке, начиная свой путь на левом берегу ребра-корня, и каждый раз записываем букву x, когда проходим вдоль

некоторого ребра впервые, и букву y, когда проходим вдоль того же ребра вторично с другой стороны. Так получается слово Дика. Путь Дика, соответствующий этому слову, стартует в начале координат. Каждой букве x ставится в соответствие шаг $(1,1)$ по координатной сетке, а каждой букве y ставится в соответствие шаг $(1,-1)$.

Эти объекты легко характеризуются. Рассмотрим слово w, являющееся конкатенацией трёх слов, $w = u_1u_2u_3$, причём каждое из слов u_i может быть пустым. Тогда u_1 называется *префиксом*, u_2 — *фактором*, а u_3 — *суффиксом* слова w.

Определение 6 (слова Дика и пути Дика). *Слово Дика* — это такое слово w в алфавите $\{x,y\}$, что $|w|_x = |w|_y$ (здесь $|w|_x$ и $|w|_y$ обозначают число вхождений букв x и y в w), в то время как для любого префикса u слова w мы имеем $|u|_x \geqslant |u|_y$. *Путь Дика* — это путь на плоскости, который стартует в начале координат, состоит из шагов $(1,1)$ и $(1,-1)$ и финиширует на оси абсцисс, всё время оставаясь при этом в верхней полуплоскости.

Рисунок 3 иллюстрирует введённые выше понятия. Корневое ребро выделено жирным, а начальная буква x показана более крупным шрифтом.

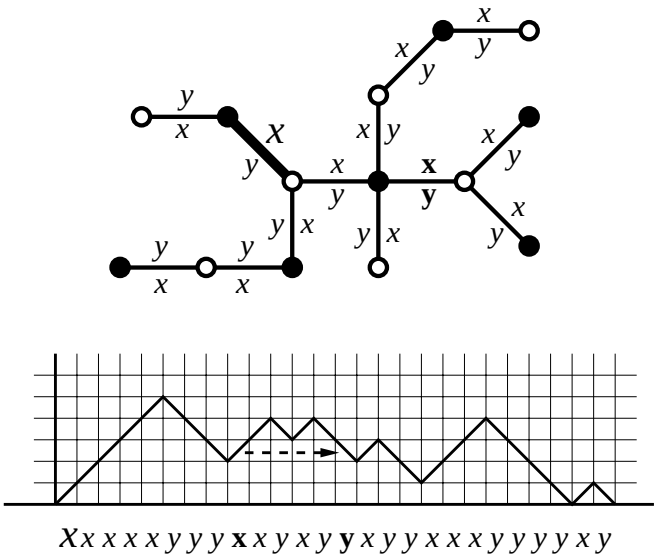

Рис. 3. Корневое двукрашенное плоское дерево и соответствующие ему слово Дика и путь Дика

Замечание 7 (пустое слово Дика). Слово Дика может быть пустым. В этом случае оно соответствует дереву, состоящему из единственной вершины. В виде

исключения мы не окрашиваем эту вершину ни в чёрный, ни в белый цвет. Тогда имеется только одно дерево без рёбер и одно пустое слово.

Следующее предложение является тривиальным следствием приведённой выше конструкции.

Предложение 8 (деревья и слова Дика). *Существует взаимно-однозначное соответствие между корневыми двукрашенными плоскими деревьями и словами Дика.*

Для описания взвешенных деревьев осталось сделать совсем немного. Имеется естественное понятие спаривания букв в слове Дика: две буквы x и y называются *парными* (или образуют *пару*), если они стоят на разных сторонах одного и того же ребра. Парные буквы в слове или в пути Дика легко распознаются. Выберем в слове Дика w две буквы x и y, такие что x идёт раньше y, и рассмотрим фактор xuy. Тогда (x, y) образуют пару в том и только в том случае, если слово u, стоящее между x и y, само является словом Дика. В пути Дика выберем восходящий шаг, соответствующий букве x, и будем двигаться от него по горизонтали до тех пор, пока не наткнёмся на нисходящий шаг: этот шаг как раз и соответствует букве y, образующей пару с выбранной ранее буквой x. На рис. 3 одна из пар выделена полужирным шрифтом (и на дереве, и в слове Дика), а пунктирная стрелка показывает, как найти «противоположный» шаг пути Дика. Пара (x, y), в которой буква x является самой первой буквой слова Дика, отвечает корневому ребру.

Вернёмся к взвешенным деревьям и проделаем следующую операцию: для каждого ребра заменим соответствующую ему пару (x, y) на пару (x_i, y_i), где i — это вес данного ребра.

Определение 9 (взвешенные слова Дика). *Взвешенное слово Дика* — это слово в бесконечном алфавите $\{x_i, y_i\}_{i \geqslant 1}$, представляющее собой слово Дика, в котором каждая пара букв (x, y) заменена на некоторую пару (x_i, y_i). Мы говорим, что пара (x_i, y_i) имеет вес i, а вес всего слова — это сумма весов составляющих его пар.

Предложение 10 (взвешенные слова и деревья). *Имеется взаимно-однозначное соответствие между взвешенными деревьями и взвешенными словами Дика.*

4. Доказательство основной теоремы

Каждое непустое слово Дика w однозначно представляется в виде конкатенации $w = xuyv$, где u и v сами являются словами Дика (возможно, пустыми). Здесь, очевидно, x — это первая буква слова w, а y образует с ней пару. Нисходящий шаг пути Дика, соответствующий этой букве y, отвечает моменту первого возвращения на горизонтальную ось.

Аналогично любое непустое взвешенное слово Дика w однозначно представляется в виде конкатенации $x_i u y_i v$ для некоторого $i \geqslant 1$; здесь u и v являются взвешенными словами Дика.

Обозначим через $\mathcal{D}$ формальную сумму всех взвешенных слов Дика, т. е. формальный степенной ряд с некоммутирующими переменными x_i, y_i, $i = 1, 2, \ldots$ (через ε обозначается пустое слово):

$$\mathcal{D} = \varepsilon + x_1y_1 + x_2y_2 + x_1x_1y_1y_1 + x_1y_1x_1y_1 + x_3y_3 + x_1x_2y_2y_1 + \\ + x_1y_1x_2y_2 + x_2x_1y_1y_2 + x_2y_2x_1y_1 + x_1x_1x_1y_1y_1y_1 + \ldots . \quad (7)$$

(Для того чтобы записать такой ряд, следует выбрать какой-нибудь полный порядок на множестве слов. Конкретный вид этого порядка безразличен. В формуле (7) слова ранжированы по весу, для каждого веса — по длине, а при фиксированных весе и длине — по алфавиту.) Тогда упомянутое выше разложение слов ряда $\mathcal{D}$ в виде $x_i u y_i v$ приводит к следующему уравнению для $\mathcal{D}$:

$$\mathcal{D} = \varepsilon + x_1\mathcal{D}y_1\mathcal{D} + x_2\mathcal{D}y_2\mathcal{D} + \ldots = \varepsilon + \sum_{i=1}^{\infty} x_i\mathcal{D}y_i\mathcal{D}. \quad (8)$$

Теперь проделаем следующее:

— заменим ε, а также каждое вхождение буквы y_i в $\mathcal{D}$ множителем 1;
— заменим каждое вхождение буквы x_i в $\mathcal{D}$ на st^i;
— сделаем переменные s и t коммутирующими.

Тогда каждое слово w ряда $\mathcal{D}$ приобретает вид $s^m t^n$, где m — это число вхождений в w букв x_i, $i \geqslant 1$ (или, что то же самое, число рёбер во взвешенном дереве T_w, отвечающем слову w), а n — это вес слова w (или, что то же самое, вес дерева T_w). Приводя подобные члены, получаем производящую функцию

$$h(s,t) = \sum_{m,n \geqslant 0} b_{m,n} s^m t^t.$$

В то же время уравнение (8) преобразуется в следующее квадратное уравнение относительно $h(s,t)$:

$$h = 1 + s\left(\sum_{i=1}^{\infty} t^i\right)h^2 = 1 + \frac{st}{1-t} \cdot h^2. \quad (9)$$

Решая это уравнение и выбирая знак перед квадратным корнем таким образом, чтобы избежать полюса в нуле, получаем формулу (4). Подставляя в неё $s = 1$, получаем (1).

Асимптотическое выражение (3) для чисел a_n получается прямым применением к функции $f(t)$ готовых формул асимптотического анализа коэффициентов производящих функций (см., например, [2, гл. VI]). Надо лишь заметить, что подкоренное выражение в $f(t)$ равно

$$1 - 6t + 5t^2 = (1-t)(1-5t).$$

Для доказательства формулы (5) вспомним, что имеется Cat_m топологических корневых деревьев с m рёбрами. Будем обходить такое дерево по часовой стрелке, отправляясь от корня, и присваивать каждому впервые встретившемуся ребру какойл-ибо ненулевой вес. Имеется $\binom{n-1}{m-1}$ способов распределить полный вес n по m рёбрам. В самом деле, разложим в ряд n точек и расставим $m-1$ разделителей среди $n-1$ мест между точками. Эта процедура делит число n на m ненулевых частей.

Для доказательства рекуррентного соотношения (2) рассмотрим по отдельности деревья веса $n+1$, у которых корень имеет вес 1, и деревья, у которых вес корня равен $i \geqslant 2$. Деревья первого типа соответствуют словам вида $x_1 u y_1 v$, где u и v — взвешенные слова Дика суммарного веса n. Если вес u равен k, то вес v равен $n-k$. Суммируя по $k = 0, 1, \ldots, n$, получаем слагаемое $\sum\limits_{k=0}^{n} a_k a_{n-k}$ в правой части формулы (2). Деревья веса $n+1$ с корнем веса $i \geqslant 2$ получаются из деревьев веса n с корнем веса $i-1$ добавлением единицы к весу корня. Это даёт нам слагаемое a_n в правой части (2).

Наконец, чтобы получить формулу (6), заметим, что имеется m способов выбора корня в дереве с m рёбрами. Однако, если дерево имеет нетривиальные симметрии, некоторые из этих выборов приводят к изоморфным корневым деревьям. Количество неизоморфных корневых деревьев равно $m/|\operatorname{Aut}(T)|$. Деля на m, получаем присутствующийщий в (6) множитель $1/|\operatorname{Aut}(T)|$.

Теорема 4 доказана. □

5. Перечисление в соответствии с паспортом: возможна ли явная формула?

Пусть $\lambda \vdash n$, $\lambda = (\lambda_1, \lambda_2, \ldots \lambda_k)$ — разбиение числа n. Запишем λ в виде

$$\lambda = 1^{d_1} 2^{d_2} \ldots n^{d_n}, \quad \text{where} \quad \sum_{i=1}^{n} d_i = k, \quad \sum_{i=1}^{n} i \cdot d_i = n,$$

где d_i — это число частей разбиения λ, равных i. Обозначим

$$N(\lambda) = \frac{(k-1)!}{d_1!\, d_2! \ldots d_n!}. \tag{10}$$

Следующий результат был получен В. Т. Таттом в [8] (1964 г.) и впоследствии обобщён И. П. Гульденом и Д. М. Джексоном в [3] (1992 г.).

Теорема 11 (деревья с заданным паспортом). *Число обычных (т. е. не взвешенных) корневых двукрашенных плоских деревьев с паспоротом (α, β) равно*

$$nN(\alpha)N(\beta). \tag{11}$$

Соответственно, число неизоморфных обычных двукрашенных плоских деревьев с паспортом (α, β)*, подсчитываемых с весом* $1/|\operatorname{Aut}(T)|$ *каждое, равно*

$$\sum_T \frac{1}{|\operatorname{Aut}(T)|} = N(\alpha)N(\beta), \tag{12}$$

где сумма берётся по всем неизоморфным обычным двукрашенным плоским деревьям с паспортом (α, β).

Приведённая теорема является несравненно более сильной, чем наша теорема 4. Во-первых, она даёт *явную формулу*. Во-вторых, она перечисляет деревья не по одному или двум параметрам (таким, как вес и число рёбер), а по паспорту, содержащему гораздо более детальную информацию о дереве.

Возможна ли аналогичная формула для взвешенных деревьев?

Главным препятствием на пути к получению такой формулы является тот факт, что один и тот же паспорт может быть реализован как деревом, так и лесом: совсем простой пример этого явления показан на рис. 4. Поэтому прямолинейный подход (ни в коей мере не являющийся тривиальным) неизбежно приводит к процедуре включения-исключения, т. е. даёт не явную формулу, а алгоритм вычисления требуемого числа. Такой алгоритм получен в [4].

Рис. 4. Один и тот же паспорт $(5^1 3^1, 5^1 3^1)$ реализуется в виде леса и в виде дерева

Возникает до некоторой степени философский вопрос: что мы называем явной формулой? Ответ, разумеется, может быть лишь неформальным. Например, явная формула могла бы дать возможность «с одного взгляда» установить, что число взвешенных деревьев с данным паспортом не равно нулю. Чисто комбинаторное доказательство этого факта в [6] потребовало более двух страниц текста. Более тонкий вопрос формулируется так: в каком случае существует ровно одно взвешенное дерево с данным паспортом? Классификация таких деревьев, полученная в [6], явилась результатом длинных и запутанных рассуждений, растянувшихся более чем на 25 страниц. Наличие явной формулы могло бы кардинально упростить доказательство этого результата.

Литература

[1] Адрианов Н. М., Звонкин А. К. Взвешенные деревья с примитивными группами вращений рёбер // Фундамент. и прикл. мат. — 2013. — Т. 18, вып. 6. — С. 5—50.

[2] Flajolet P., Sedgewick R. Analytic Combinatorics. — Cambridge: Cambridge Univ. Press, 2009.

[3] Goulden I. P., Jackson D. M. The combinatorial relationship between trees, cacti and certain connection coefficients for the symmetric group // Eur. J. Combin. — 1992. — Vol. 13. — P. 357—365.

[4] Kochetkov Yu. Yu. Enumeration of one class of plane weighted trees. — 2013. — arXiv: 1310.6208v1.

[5] The On-Line Encyclopedia of Integer Sequences. — http://oeis.org/.

[6] Pakovich F., Zvonkin A. K. Minimum degree of the difference of two polynomials over $\mathbb{Q}$, and weighted plane trees // Selecta Math. (N. S.) — 2014. — P. 1—63. — DOI 10.1007/s00029-014-0151-0. — arXiv:1306.4141v1.

[7] Pakovich F., Zvonkin A. K. Minimum degree of the difference of two polynomials over $\mathbb{Q}$. Pt. II: Davenport—Zannier pairs. — http://www.labri.fr/perso/zvonkin/.

[8] Tutte W. T. The number of planted plane trees with a given partition // Am. Math. Mon. — 1964. — Vol. 71, no. 3. — P. 272—277.