О весе функций, заданных бесповторными И/ИЛИ формулами

Еременко А.Р., Яшунский А.Д.

Рассматривается множество функций, заданных бесповторными формулами с бинарными операциями конъюнкции (логического И) и дизъюнкции (логического ИЛИ). Для функций, заданных формулами с фиксированным числом операций, исследуются значения весов — числа наборов, на которых функция принимает значение 1. Найдены асимптотические оценки для числа бесповторных формул, задающих функции с весом из определенных диапазонов, в частности, — числа формул, задающих функции, у которых доля единиц среди значений не превышает четверти.

Ключевые слова: булева функция, бесповторная формула, вес функции, асимптотическая оценка.

1. Введение

Рассмотрим плоское корневое бинарное дерево с n внутренними вершинами. Будем считать, что все листья дерева помечены символами переменных $x_1, \dots, x_{n+1}$ или их отрицаний, причем любым двум различным листам соответствуют различные переменные, а каждая внутренняя вершина помечена либо символом & (конъюнкция, логическое И), либо символом $\vee$ (дизъюнкция, логическое ИЛИ). Тогда каждому дереву можно сопоставить булеву функцию от переменных $x_1, \dots, x_{n+1}$. Такая функция будет выражаться *бесповторной* формулой, т. е. каждый символ переменной будет встречаться в ней ровно один раз. Подробнее о таких функциях, часто также называемых бесповторными, см. [1]. Напомним, что *весом* булевой функции называется число наборов, на которых она принимает значение 1. Очевидно, что вес функции от $n+1$ переменной — целое число от 0 до 2^{n+1}.

Пусть вместо всех переменных бесповторной функции подставлены независимые случайные величины, равные 0 и 1 с вероятностями $\frac{1}{2}$.

Несложно видеть, что в этом случае вероятность обращения функции в единицу будет в точности равна ее весу, разделенному на 2^{n+1}. Кроме того, легко проверить, что эта вероятность (равно как и вес функции) не зависят от расстановки отрицаний на символах переменных, а зависит только от структуры дерева и пометок на его внутренних вершинах.

Множество значений вероятностей (а вместе с ними и возможных весов бесповторных функций) можно описать следующей формальной конструкцией. Определим *вероятностные формулы* по индукции. Будем считать, что x — вероятностная формула сложности $L(x) = 0$ со значением $\pi(x) = \frac{1}{2}$. Если Φ_1 и Φ_2 — вероятностные формулы, то выражения $(\Phi_1 \& \Phi_2)$ и $(\Phi_1 \vee \Phi_2)$ также являются вероятностными формулами сложности $L(\Phi_1) + L(\Phi_2) + 1$, а их значения, соответственно равны:

$$\pi((\Phi_1 \& \Phi_2)) = \pi(\Phi_1)\pi(\Phi_2), \tag{1}$$

$$\pi((\Phi_1 \vee \Phi_2)) = \pi(\Phi_1) + \pi(\Phi_2) - \pi(\Phi_1)\pi(\Phi_2). \tag{2}$$

Общее число вероятностных формул сложности n в точности равно числу деревьев с n внутренними вершинами, помеченными символами & и $\vee$. Оно выражается как $2^n c_n$, где c_n — числа Каталана, которые можно задать соотношениями $c_0 = 1$ и $c_{n+1} = \sum\limits_{k=0}^{n} c_k c_{n-k}$.

Из результатов Р. Л. Схиртладзе [2] известно, что значением вероятностных формул могут быть любые двоично-рациональные числа. При этом значениями формул сложности n являются всевозможные дроби с нечетным числителем и знаменателем 2^{n+1}, и только они. С учетом этого факта, положим:

$$f(l,n) = \left| \left\{ \Phi \colon \pi(\Phi) = \frac{2l-1}{2^{n+1}} \right\} \right|$$

для $l = 1, \ldots, 2^n$.

Отметим, что среди бесповторных формул, заданных И/ИЛИ деревьями с фиксированным порядком переменных, ровно $f(l,n)$ выражают монотонную функцию веса l от $n+1$ переменной, и ровно $2^{n+1} f(l,n)$ выражают функцию веса l от $n+1$ переменной. При этом, например, количество бесповторных монотонных функций веса l от $n+1$ переменной можно оценить сверху величиной $(n+1)! f(l,n)$. Далее в тексте работы мы будем оперировать исключительно с величинами $f(l,n)$ не обращаясь более к их интерпретации через веса соответствующих булевых функций.

Как несложно видеть, при всех n имеет место равенство $\sum\limits_{l=1}^{2^n} f(l,n) = 2^n c_n$. Также по индукции несложно доказать, что у *двойственных* вероятностных формул, т. е. таких, которые получаются друг из друга заменой символов & на $\vee$, а $\vee$ на & соответственно, значения в сумме равны 1. Отсюда вытекает следующее важное соотношение:

$$f(l,n) = f(2^n - l + 1, n). \tag{3}$$

Таким образом, за исключением случая $n = 0$, все вероятностные формулы могут быть разбиты на пары, в которых сумма значений равна 1, следствием чего оказывается выполнение при $n \geqslant 1$ равенства:

$$\sum_{l=1}^{\frac{1}{2}2^n} f(l,n) = \frac{1}{2} 2^n c_n. \tag{4}$$

Отсюда же вытекает, что среднее значение $\pi(\Phi)$ по формулам фиксированной сложности равно $\frac{1}{2}$. Это частный случай более общего свойства вероятностных формул, доказанного в [3].

Дроби с числителем 1 (соответственно $2^n - 1$) могут быть получены только из формулы, содержащей исключительно символы & (соответственно $\vee$), поэтому выполнены равенства $f(1,n) = f(2^n - 1, n) = c_n$.

Как следует из равенств (1) и (2) значение $\frac{2l-1}{2^{n+2}}$ может быть получено в том, и только в том случае, если значения $\pi(\Phi_1) = \frac{2l_1-1}{2^{n_1+1}}$ и $\pi(\Phi_2) = \frac{2l_2-1}{2^{n_2+1}}$ удовлетворяют одному из соотношений:

$$\left[\begin{array}{l} \frac{2l_1-1}{2^{n_1+1}} \cdot \frac{2l_2-1}{2^{n_2+1}} = \frac{2l-1}{2^{n+2}}, \\ \frac{2l_1-1}{2^{n_1+1}} + \frac{2l_2-1}{2^{n_2+1}} - \frac{2l_1-1}{2^{n_1+1}} \cdot \frac{2l_2-1}{2^{n_2+1}} = \frac{2l-1}{2^{n+2}}. \end{array}\right.$$

Второе из этих соотношений может быть переписано как

$$\frac{2(2^{n_1} - l_1 + 1) - 1}{2^{n_1+1}} \cdot \frac{2(2^{n_2} - l_2 + 1) - 1}{2^{n_2+1}} = \frac{2(2^{n+1} - l + 1) - 1}{2^{n+2}},$$

что, с учетом соображений двойственности (см. равенство (3)) приводит к следующей формуле для $f(l, n+1)$:

$$f(l, n+1) = \sum_{\frac{2l_1-1}{2^{n_1+1}} \cdot \frac{2l_2-1}{2^{n_2+1}} \in \left\{\frac{2l-1}{2^{n+2}}, \frac{2(2^{n+1}-l-1)+1}{2^{n+2}}\right\}} f(l_1, n_1) f(l_2, n_2). \tag{5}$$

Формула (5) позволяет для малых n непосредственно вычислить значения $f(l,n)$. Гистограммы значений $f(l,n)$, а также график нормированной нарастающей суммы $F(x,n) = \frac{1}{2^n c_n} \sum\limits_{1 \leqslant l \leqslant x 2^n} f(l,n)$, для случая $n = 10$ приведены на рис. 1. Пунктиром на графике $F(x,10)$ показана прямая $y = x$.

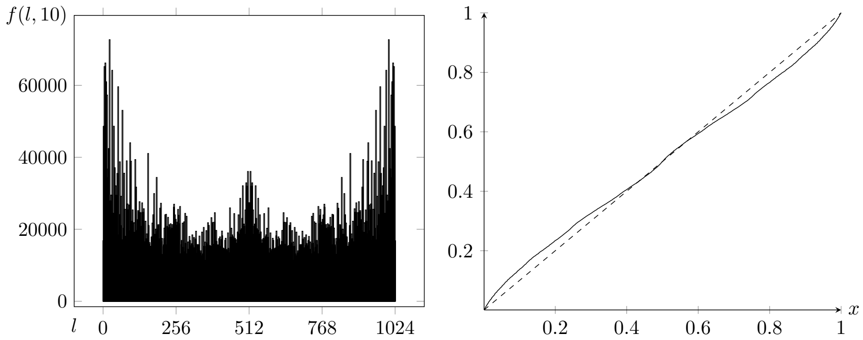

Рис. 1. Графики $f(l,10)$ и $F(x,10)$

Графики функций $F(x,n)$ с ростом n постепенно приближаются к прямой $y = x$, и уже для $n = 10$ график расположен достаточно близко от прямой. Это позволяет высказать гипотезу о том, что с ростом n значения $f(l,n)$ для $l = 1, \ldots, 2^n$ постепенно выравниваются. Однако, как будет показано далее, эта гипотеза не подтверждается.

2. Неравномерность $f(l,n)$

Рассмотрение в сумме из соотношения (5) лишь части слагаемых позволяет получать нижние оценки для $f(l,n)$.

Дробь $\frac{2l-1}{2^{n+1}}$ может быть представлена в виде $\frac{2l-1}{2^{k+1}} \cdot \frac{1}{2^{n-k-1+1}}$, при условии, что k удовлетворяет неравенству $k \geqslant \lceil \log_2 l \rceil$. Это позволяет выписать для $f(l,n+1)$ следующую оценку снизу:

$$f(l,n+1) \geqslant 2 \sum_{k=\lceil \log_2 l \rceil}^{n} f(l,k) f(1,n-k) = 2 \sum_{k=\lceil \log_2 l \rceil}^{n} f(l,k) c_{n-k}.$$

Положим $\alpha_0 = 1$ и определим $\alpha_{t+1} = 2 \sum\limits_{s=0}^{t} \alpha_t c_{t-s}$. Тогда индукцией по n несложно доказать следующее утверждение.

Лемма 1. *При* $n \geqslant \lceil \log_2 l \rceil$ *выполнено* $f(l,n) \geqslant \alpha_{n-\lceil \log_2 l \rceil} f(l, \lceil \log_2 l \rceil)$.

Выражение для элементов последовательности α_t может быть найдено явно. Обозначим $A(z) = \sum\limits_{t=0}^{\infty} \alpha_t z^t$ производящую функцию последовательности α_t, а через $C(z) = \sum\limits_{n=0}^{\infty} c_n z^n$ — производящую функцию чисел Каталана. Напомним, что для $C(z)$ имеет место равенство $C(z) = \frac{1}{2z}(1 - \sqrt{1-4z})$. Тогда из определения последовательности α_t вытекает, что $A(z)$ удовлетворяет уравнению

$$A(z) - 1 = 2zA(z)C(z),$$

из которого вытекает, что $A(z) = \frac{1}{1-2zC(z)}$. С учетом выражения для $C(z)$ получаем, что $A(z) = \frac{1}{\sqrt{1-4z}}$. Разложение функции $A(z)$ в ряд по степеням z дает явное выражение $\alpha_t = \binom{2t}{t}$. Учитывая известное соотношение $c_n = \frac{1}{n+1}\binom{2n}{n}$, получаем следующее утверждение.

Теорема 1. *При* $n \geqslant \lceil \log_2 l \rceil$ *выполнено неравенство*

$$f(l,n) \geqslant (n - \lceil \log_2 l \rceil + 1) f(l, \lceil \log_2 l \rceil) c_{n-\lceil \log_2 l \rceil}.$$

В частности, для $l = 2$ в качестве следствия получаем

$$f(2,n) \geqslant n f(2,1) c_{n-1} = n c_{n-1}.$$

Поскольку $\frac{nc_{n-1}}{c_n} \to \infty$ при $n \to \infty$, доказанное неравенство исключает возможность приближения значений $f(l,n)$ к c_n с ростом n при всех l, как это можно было бы предположить с учетом того, что среднее значение $\frac{1}{2^n} \sum\limits_{l=1}^{2^n} f(l,n)$ в точности равно c_n. Таким образом, если какой-то предельный закон для величин $f(l,n)$ и существует, его характер должен быть существенно сложнее.

3. Рекуррентные неравенства

Дальнейший анализ покажет, что гипотеза о какой бы то ни было равномерности в распределении значений формул не подтверждается даже в самом слабом варианте. Для того, чтобы это продемонстрировать, мы будем оценивать количество формул сложности n, значение которых не превышает $\frac{1}{4}$ — будем обозначать его $q(n)$, а также количество формул сложности n, значение которых лежит между $\frac{1}{4}$ и $\frac{1}{2}$ — его будем обозначать $r(n)$. Введем соответствующие определения.

Положим $q(0) = 0$, $q(1) = 1$, и далее для $n \geqslant 2$ определим

$$q(n) = \sum_{l=1}^{\frac{1}{4}2^n} f(l, n).$$

Оценим величину $q(n)$ снизу. Из соотношения (5) вытекает, что $q(n+1)$ в точности равно сумме произведений $f(l_1, n_1)f(l_2, n_2)$ по всевозможным значениям l_1, n_1, l_2, n_2, удовлетворяющих условиям

$$\left[\begin{array}{l} \frac{2l_1-1}{2^{n_1+1}} \cdot \frac{2l_2-1}{2^{n_2+1}} \leqslant \frac{1}{4}, \\ \frac{2l_1-1}{2^{n_1+1}} \cdot \frac{2l_2-1}{2^{n_2+1}} \geqslant \frac{3}{4}, \end{array}\right.$$

Оценка снизу для $q(n+1)$ получается рассмотрением только первого из этих двух условий. Заметим, что неравенство $\frac{2l_1-1}{2^{n_1+1}} \cdot \frac{2l_2-1}{2^{n_2+1}} \leqslant \frac{1}{4}$ заведомо выполняется, если какая-то из дробей, $\frac{2l_1-1}{2^{n_1+1}}$ или $\frac{2l_2-1}{2^{n_2+1}}$ не больше $\frac{1}{4}$, а также если обе дроби не превышают $\frac{1}{2}$. Объединяя эти два случая, учитывая их пересечение, получаем следующую оценку:

$$q(n+1) \geqslant \sum_{\substack{n_1+n_2=n, \\ \frac{2l_1-1}{2^{n_1+1}} \leqslant \frac{1}{2}, \frac{2l_2-1}{2^{n_2+1}} \leqslant \frac{1}{2}}} f(l_1, n_1)f(l_2, n_2) + 2\sum_{\substack{n_1+n_2=n, \\ \frac{2l_1-1}{2^{n_1+1}} > \frac{1}{2}, \frac{2l_2-1}{2^{n_2+1}} \leqslant \frac{1}{4}}} f(l_1, n_1)f(l_2, n_2). \tag{6}$$

Рассмотрим сначала первую сумму, выделив в ней слагаемые, в которых n_1 или n_2 — нулевое:

$$\begin{aligned} &\sum_{\substack{n_1+n_2=n, \\ \frac{2l_1-1}{2^{n_1+1}} \leqslant \frac{1}{2}, \frac{2l_2-1}{2^{n_2+1}} \leqslant \frac{1}{2}}} f(l_1, n_1)f(l_2, n_2) = \\ &\qquad = f(1,0)\left(\sum_{l_2 \leqslant \frac{1}{2}2^n} f(l_2, n)\right) + \left(\sum_{l_1 \leqslant \frac{1}{2}2^n} f(l_1, n)\right) f(1,0) + \\ &\qquad\qquad + \sum_{\substack{n_1+n_2=n, \\ n_1>0, n_2>0}} \left(\sum_{\frac{2l_1-1}{2^{n_1+1}} \leqslant \frac{1}{2}} f(l_1, n_1)\right)\left(\sum_{\frac{2l_2-1}{2^{n_2+1}} \leqslant \frac{1}{2}} f(l_2, n_2)\right). \end{aligned}$$

С учетом соотношения (4) рассматриваемая сумма принимает вид:

$$\sum_{\substack{n_1+n_2=n,\\ \frac{2l_1-1}{2^{n_1+1}}\leqslant\frac{1}{2},\frac{2l_2-1}{2^{n_2+1}}\leqslant\frac{1}{2}}} f(l_1,n_1)f(l_2,n_2) = 2\cdot\frac{1}{2}2^n c_n + \sum_{\substack{n_1+n_2=n,\\ n_1>0,n_2>0}} \frac{1}{2}2^{n_1}c_{n_1}\frac{1}{2}2^{n_2}c_{n_2} =$$

$$= 2^n c_n + \frac{1}{4}2^n\sum_{k=1}^{n-1} c_k c_{n-k} = 2^n c_n + \frac{1}{4}2^n(c_{n+1}-2c_n) = \frac{1}{2}2^n c_n + \frac{1}{8}2^{n+1}c_{n+1}.$$

Перейдем теперь к рассмотрению второй суммы из (6). Заметим, что если одно из значений n_1 или n_2 нулевое, то не найдется ни одной пары значений l_1, l_2, удовлетворяющих условиям $\frac{2l_1-1}{2^{n_1+1}} > \frac{1}{2}$ и $\frac{2l_2-1}{2^{n_2+1}} \leqslant \frac{1}{4}$. С учетом этого и равенства (4) получаем:

$$\sum_{\substack{n_1+n_2=n,\\ \frac{2l_1-1}{2^{n_1+1}}>\frac{1}{2},\frac{2l_2-1}{2^{n_2+1}}\leqslant\frac{1}{4}}} f(l_1,n_1)f(l_2,n_2) = \sum_{k=1}^{n-1}\left(\sum_{\frac{2l_1-1}{2^{k+1}}>\frac{1}{2}} f(l_1,k)\right)\left(\sum_{\frac{2l_2-1}{2^{n-k+1}}\leqslant\frac{1}{4}} f(l_2,n-k)\right) =$$

$$= \sum_{k=1}^{n-1}\frac{1}{2}2^k c_k q(n-k) = \frac{1}{2}\sum_{k=1}^{n-1}2^k c_k q(n-k).$$

Объединяя все выведенные выше неравенства, приходим к следующей рекуррентной оценке.

Лемма 2. *При $n \geqslant 1$ выполнено неравенство:*

$$q(n+1) \geqslant \frac{1}{8}2^{n+1}c_{n+1} + \frac{1}{2}2^n c_n + \sum_{k=1}^{n-1}2^k c_k q(n-k).$$

Определим теперь функцию $r(n)$, положив $r(0) = r(1) = 0$ и далее для $n \geqslant 2$:

$$r(n) = \sum_{\frac{1}{4}2^n<l<\frac{1}{2}2^n} f(l,n) = \sum_{\frac{1}{2}2^n<l<\frac{3}{4}2^n} f(l,n),$$

где равенство сумм очевидно имеет место в силу соотношения (3). Из данного определения получаем, что $r(2) = 2$. Отметим также, что при $n \geqslant 1$ выполнено $q(n) + r(n) = \frac{1}{2}2^n c_n$.

Представление (5) и симметрия $f(l,n)$, описываемая равенством (3), позволяют получить следующее соотношение для $r(n+1)$ при $n \geqslant 2$:

$$2r(n+1) = \sum_{\frac{1}{4}2^{n+1}<l<\frac{1}{2}2^{n+1}} f(l,n) + \sum_{\frac{1}{2}2^{n+1}<l<\frac{3}{4}2^{n+1}} f(l,n) =$$
$$= \sum_{\frac{1}{4}2^{n+1}<l<\frac{3}{4}2^{n+1}} \sum_{\frac{2l_1-1}{2^{n_1+1}} \cdot \frac{2l_2-1}{2^{n_2+1}} \in \left\{\frac{2l-1}{2^{n+2}}, \frac{2(2^{n+1}-l-1)+1}{2^{n+2}}\right\}} f(l_1,n_1)f(l_2,n_2) =$$
$$= 2\sum_{\substack{n_1+n_2=n,\\ \frac{1}{4}<\frac{2l_1-1}{2^{n_1+1}}\cdot\frac{2l_2-1}{2^{n_2+1}}<\frac{3}{4}}} f(l_1,n_1)f(l_2,n_2).$$

Из данного соотношения естественно вытекает равенство

$$r(n+1) = \sum_{\substack{n_1+n_2=n,\\ \frac{1}{4}<\frac{2l_1-1}{2^{n_1+1}}\cdot\frac{2l_2-1}{2^{n_2+1}}<\frac{3}{4}}} f(l_1,n_1)f(l_2,n_2), \tag{7}$$

которое будем использовать далее для получения рекуррентной нижней оценки величины $r(n+1)$.

В сумме из соотношения (7) выделим слагаемые, в которых n_1 или n_2 меньше 2:

$$r(n+1) = 2\sum_{\frac{1}{4}<\frac{2l_2-1}{2^{n+1}}<\frac{3}{4}} f(1,0)f(l_2,n) + 2\sum_{\frac{1}{4}<\frac{3}{4}\cdot\frac{2l_2-1}{2^{n-1+1}}<\frac{3}{4}} f(2,1)f(l_2,n-1) + \sum_{\substack{n_1+n_2=n,\, n_1>1,\, n_2>1,\\ \frac{1}{4}<\frac{2l_1-1}{2^{n_1+1}}\cdot\frac{2l_2-1}{2^{n_2+1}}<\frac{3}{4}}} f(l_1,n_1)f(l_2,n_2).$$

В представлении выше удвоение первых двух сумм обусловлено тем, что значение 0 (соответственно 1) может принимать как n_1, так и n_2. Во второй сумме выбор $n_1 = 1$ автоматически влечет $l_1 = 2$, так как при $l_1 = 1$, мы получали бы для l_2 заведомо невыполнимое условие $\frac{1}{4} < \frac{1}{4} \cdot \frac{2l_2-1}{2^{n_2+1}}$.

Оценим каждую из сумм в приведенном выше соотношении (напомним, что оно рассматривается для $n \geqslant 2$). Для первой суммы имеет место точное равенство:

$$2\sum_{\frac{1}{4}<\frac{2l_2-1}{2^{n+1}}<\frac{3}{4}} f(1,0)f(l_2,n) = 2\sum_{\frac{1}{4}<\frac{2l-1}{2^{n+1}}<\frac{3}{4}} f(l,n) = 4r(n).$$

Вторая сумма может быть оценена снизу, поскольку значения l_2, удовлетворяющие неравенствам $\frac{1}{4} < \frac{3}{4} \cdot \frac{2l_2-1}{2^{n-1+1}} < \frac{3}{4}$, эквивалентным неравенствам $\frac{1}{3} < \frac{2l_2-1}{2^n} < 1$, заведомо включают множество значений l_2, удовлетворяющих $\frac{1}{2} < \frac{2l_2-1}{2^n}$:

$$2\sum_{\frac{1}{4}<\frac{3}{4}\cdot\frac{2l_2-1}{2^{n-1+1}}<\frac{3}{4}} f(2,1)f(l_2,n-1) = 2\sum_{\frac{1}{4}<\frac{3}{4}\cdot\frac{2l-1}{2^n}<\frac{3}{4}} f(l_2,n-1) \geqslant$$

$$\geqslant 2\sum_{\frac{2l-1}{2^n}>\frac{1}{2}} f(l,n-1) = 2\cdot\frac{1}{2}2^{n-1}c_{n-1} = 2^{n-1}c_{n-1}.$$

Наконец, рассмотрим третью сумму (отметим, что она содержит какие-либо слагаемые только если $n \geqslant 4$). Множество значений l_1, n_1 и l_2, n_2, удовлетворяющих неравенствам $\frac{1}{4} < \frac{2l_1-1}{2^{n_1+1}} \cdot \frac{2l_2-1}{2^{n_2+1}} < \frac{3}{4}$ заведомо включает наборы, удовлетворяющие соотношениям:

$$\left[\begin{array}{l} \frac{1}{2} < \frac{2l_1-1}{2^{n_1+1}} < \frac{3}{4}, \frac{3}{4} < \frac{2l_2-1}{2^{n_2+1}} < 1; \\ \frac{1}{2} < \frac{2l_1-1}{2^{n_1+1}} < \frac{3}{4}, \frac{1}{2} < \frac{2l_2-1}{2^{n_2+1}} < \frac{3}{4}; \\ \frac{3}{4} < \frac{2l_1-1}{2^{n_1+1}} < 1, \frac{1}{2} < \frac{2l_2-1}{2^{n_2+1}} < \frac{3}{4}. \end{array}\right.$$

В силу симметрии первого и третьего варианта в совокупности выше, получаем следующую нижнюю оценку:

$$\sum_{\substack{n_1+n_2=n,\\ n_1>1,\, n_2>1,\\ \frac{1}{4}<\frac{2l_1-1}{2^{n_1+1}}\cdot\frac{2l_2-1}{2^{n_2+1}}<\frac{3}{4}}} f(l_1,n_1)f(l_2,n_2) \geqslant \quad 2\sum_{\substack{n_1+n_2=n,\\ n_1>1,\, n_2>1,\\ \frac{1}{2}<\frac{2l_1-1}{2^{n_1+1}}<\frac{3}{4},\, \frac{2l_2-1}{2^{n_2+1}}>\frac{3}{4}}} f(l_1,n_1)f(l_2,n_2) +$$

$$+\sum_{\substack{n_1+n_2=n,\\ n_1>1,\, n_2>1,\\ \frac{1}{2}<\frac{2l_1-1}{2^{n_1+1}}<\frac{3}{4},\, \frac{1}{2}<\frac{2l_2-1}{2^{n_2+1}}<\frac{3}{4}}} f(l_1,n_1)f(l_2,n_2) =$$

$$= 2\sum_{\substack{n_1+n_2=n,\\ n_1>1,\, n_2>1}} \left(\sum_{\frac{1}{2}<\frac{2l_1-1}{2^{n_1+1}}<\frac{3}{4}} f(l_1,n_1)\right)\left(\sum_{\frac{2l_2-1}{2^{n_2+1}}>\frac{3}{4}} f(l_2,n_2)\right) +$$

$$+\sum_{\substack{n_1+n_2=n,\\ n_1>1,\, n_2>1}} \left(\sum_{\frac{1}{2}<\frac{2l_1-1}{2^{n_1+1}}<\frac{3}{4}} f(l_1,n_1)\right)\left(\sum_{\frac{1}{2}<\frac{2l_2-1}{2^{n_2+1}}<\frac{3}{4}} f(l_2,n_2)\right) =$$

$$= \sum_{\substack{n_1+n_2=n,\\ n_1>1,\, n_2>1}} (2r(n_1)q(n_2)+r(n_1)r(n_2)) = \sum_{k=2}^{n-2} r(k)\left(q(n-k)+\frac{1}{2}2^{n-k}c_{n-k}\right).$$

Объединяя все полученные оценки, получаем следующее утверждение.

Лемма 3. *При $n \geqslant 2$ выполнено неравенство:*

$$r(n+1) \geqslant 4r(n) + 2^{n-1}c_{n-1} + \frac{1}{2}\sum_{k=2}^{n-2} r(k)2^{n-k}c_{n-k} + \sum_{k=2}^{n-2} r(k)q(n-k).$$

4. Асимптотические оценки

Для получения дальнейших результатов воспользуемся аппаратом производящих функций и подходом, систематизированным Ф. Флажоле и Р. Седжевиком в монографии *Аналитическая комбинаторика* [4]. Этот подход заключается в рассмотрении производящих функций как комплексных функций комплексного переменного. При этом асимптотическое поведение коэффициентов производящей функции оказывается связанным с поведением функции в особых точках, находящихся на границе круга сходимости степенного ряда функции, представляющего ее разложение в точке $z = 0$. Эта связь, в частности, формализуется следующей теоремой, непосредственно используемой в дальнейших рассуждениях.

Теорема 2 (Флажоле—Одлыжко [5])**.** *Пусть функция $f(z) = \sum\limits_{n=0}^{\infty} f_n z^n$ аналитическая в круге радиуса $R > 1$ с центром в точке $z = 0$, рассеченном вдоль действительной оси правее точки $z_0 = 1$. Пусть при этом в окрестности точки $z_0 = 1$ выполнено равенство $f(z) = O((1-z)^{-\alpha})$ где $\alpha \in \mathbb{R}$. Тогда при $n \to \infty$ имеет место равенство $f_n = O(n^{\alpha-1})$.*

Теорема 2 применима в случае, когда $z_0 = 1$ — единственная особая точка на границе круга сходимости функции $f(z)$, однако, как несложно понять, замена переменной в функции $f(z)$ позволяет обобщить теорему на случай, когда единственная особая точка на границе круга сходимости есть $\omega \in \mathbb{R}^+$. В этом случае при выполнении в окрестности $z_0 = \omega$ равенства $f(z) = O((1-\omega z)^{-\alpha})$ коэффициенты f_n будут удовлетворять асимптотическому равенству $f_n = O(\omega^{-n} n^{\alpha-1})$.

Перейдем к получению асимптотических оценок величин $q(n)$ и $r(n)$. Положим $d_n = 2^n c_n$, тогда производящая функция этой последовательности $D(z) = \sum\limits_{n=0}^{\infty} d_n z^n$ удовлетворяет уравнению $D(z) = 1 + 2zD^2(z)$ и может быть выражена следующим образом: $D(z) = \frac{1}{4z}(1 - \sqrt{1-8z})$, где квадратный корень понимается как одна из ветвей соответствующей многозначной аналитической функции (та, которая, в соответствии с комбинаторным смыслом функции $D(z)$, обеспечивает, что коэффициенты ее разложения в степенной ряд по степеням z действительные и неотрицательные).

Определим последовательность κ_n, положив $\kappa_0 = 0$, $\kappa_1 = 1$ и далее для $n \geqslant 1$:

$$\kappa_{n+1} = \frac{1}{8}d_{n+1} + \frac{1}{2}d_n + \sum_{k=1}^{n-1} d_k \kappa_{n-k}, \tag{8}$$

где при $n = 1$ сумма в соотношении выше считается нулевой. Тогда из леммы 2 и определения κ_n индукцией по n легко выводится соотношение $q(n) \geqslant \kappa_n$.

Аналогично определим последовательность ρ_n, положив $\rho_0 = \rho_1 = 0$, $\rho_2 = 2$ и далее для $n \geqslant 2$:

$$\rho_{n+1} = 4\rho_n + d_{n-1} + \frac{1}{2}\sum_{k=2}^{n-2} \rho_k d_{n-k} + \sum_{k=2}^{n-2} \rho_k \kappa_{n-k}, \tag{9}$$

где суммы считаются нулевыми, если верхний предел меньше нижнего. Из леммы 3 и определения последовательности ρ_n также как и в случае выше вытекает, что $r(n) \geqslant \rho_n$.

С учетом значений последовательностей κ_n, ρ_n и d_n при $n = 0$ и $n = 1$ рекуррентные соотношения (8) и (9) могут быть преобразованы следующим образом:

$$\kappa_{n+1} = \frac{1}{8}d_{n+1} + \frac{1}{2}d_n + \sum_{k=0}^{n} d_k \kappa_{n-k} - \kappa_n,$$

$$\rho_{n+1} = 4\rho_n + d_{n-1} + \frac{1}{2}\left(\sum_{k=0}^{n} \rho_k d_{n-k} - 2\rho_{n-1} - \rho_n\right) + \sum_{k=0}^{n} \rho_k \kappa_{n-k} - \rho_{n-1},$$

что после дополнительных преобразований дает:

$$\kappa_{n+1} + \kappa_n = \frac{1}{8}d_{n+1} + \frac{1}{2}d_n + \sum_{k=0}^{n} d_k \kappa_{n-k},$$

$$\rho_{n+1} + 2\rho_{n-1} = \frac{7}{2}\rho_n + d_{n-1} + \frac{1}{2}\sum_{k=0}^{n} \rho_k d_{n-k} + \sum_{k=0}^{n} \rho_k \kappa_{n-k}.$$

Введем производящие функции последовательностей κ_n и ρ_n, положив $K(z) = \sum\limits_{n=0}^{\infty} \kappa_n z^n$ и $R(z) = \sum\limits_{n=0}^{\infty} \rho_n z^n$. Умножив соотношения выше на z^{n+1}

и просуммировав по n, приходим к равенствам:

$$K(z) - z + zK(z) = \frac{1}{8}(D(z) - 2z - 1) + \frac{1}{2}z(D(z) - 1) + zD(z)K(z),$$
$$R(z) - 2z^2 + 2z^2R(z) = \frac{7}{2}zR(z) + z^2(D(z) - 1) + \frac{1}{2}zR(z)D(z) + zR(z)K(z).$$

Эти равенства можно переписать в виде:

$$(1 + z)K(z) = \frac{1}{4}z - \frac{1}{8} + \frac{1}{8}D(z) + \frac{1}{2}zD(z) + zD(z)K(z),$$
$$(1 + 2z^2)R(z) = z^2 + \frac{7}{2}zR(z) + z^2D(z) + \frac{1}{2}zR(z)D(z) + zR(z)K(z).$$

Из полученных соотношений и явного представления для функции $D(z)$, приведенного ранее, могут быть выведены явные выражения для функций $K(z)$ и $R(z)$, из которых непосредственно вытекает, что у обеих функций, как и у функции $D(z)$ ближайшей к нулю особой точкой является $z_0 = \frac{1}{8}$, и других особых точек на границе круга сходимости нет. Более того, у каждой из этих функций точка $z_0 = \frac{1}{8}$ также как и у функции $D(z)$ является точкой ветвления второго порядка.

Согласно теореме Ньютона—Пюизё (подробнее см. [4]) для функций $D(z)$, $K(z)$ и $R(z)$ в окрестности точки $z_0 = \frac{1}{8}$ имеют место представления:

$$D(z) = \delta_0 + \delta_1\sqrt{1 - 8z} + o(\sqrt{1 - 8z}),$$
$$K(z) = k_0 + k_1\sqrt{1 - 8z} + o(\sqrt{1 - 8z}),$$
$$R(z) = r_0 + r_1\sqrt{1 - 8z} + o(\sqrt{1 - 8z}),$$

где $\delta_0 = D\left(\frac{1}{8}\right)$, $k_0 = K\left(\frac{1}{8}\right)$, $r_0 = R\left(\frac{1}{8}\right)$. Эти значения могут быть легко найдены с помощью имеющихся уравнений, путем подстановки в них значения $z = \frac{1}{8}$. В результате получим $\delta_0 = 2$, $k_0 = \frac{9}{28}$, $r_0 = \frac{7}{64}$.

Соотношения между δ_1, k_1 и r_1 могут быть получены подстановкой разложений функций в уравнения и приравниванием коэффициентов при $\sqrt{1-8z}$. Продемонстрируем это для функции $K(z)$:

$$\left(1+\frac{1}{8}-\frac{1}{8}(1-8z)\right)(k_0+k_1\sqrt{1-8z}+o(\sqrt{1-8z}))=$$
$$=\frac{1}{8}(\delta_0+\delta_1\sqrt{1-8z}+o(\sqrt{1-8z})-1)+$$
$$+\left(\frac{1}{8}-\frac{1}{8}(1-8z)\right)\left(\frac{1}{4}+\frac{1}{2}\delta_0+\frac{1}{2}\delta_1\sqrt{1-8z}+o(\sqrt{1-8z})+\right.$$
$$\left.+(\delta_0+\delta_1\sqrt{1-8z}+o(\sqrt{1-8z}))(k_0+k_1\sqrt{1-8z}+o(\sqrt{1-8z}))\right).$$

Тогда соотношение между коэффициентами при $\sqrt{1-8z}$ в правой и левой частях имеет вид:

$$\frac{9}{8}k_1=\frac{1}{8}\delta_1+\frac{1}{8}\left(\frac{1}{2}\delta_1+\delta_0k_1+\delta_1k_0\right),$$

откуда $7k_1=\left(\frac{3}{2}+\frac{9}{28}\right)\delta_1$, что влечет $\frac{k_1}{\delta_1}=\frac{51}{196}=\frac{51}{49}\cdot\frac{1}{4}\approx 0,26$.

Выполняя аналогичные преобразования в уравнении, связывающем $R(z)$, приходим к соотношению:

$$\left(1+2\cdot\frac{1}{8^2}\right)r_1=\frac{7}{2}\cdot\frac{1}{8}r_1+\frac{1}{8^2}\delta_1+\frac{1}{2}\cdot\frac{1}{8}(r_1\delta_0+r_0\delta_1)+\frac{1}{8}(r_1k_0+r_0k_1),$$

из которого получаем, что $\frac{r_1}{\delta_1}=\frac{373}{6144}=\frac{373}{384}\cdot\frac{1}{16}\approx 0,06$.

Согласно теореме Флажоле—Одлыжко асимптотика роста значений последовательностей d_n, κ_n и ρ_n совпадает с асимптотикой роста коэффициентов функции $\sqrt{1-8z}$ умноженной, соответственно, на δ_1, k_1 и r_1. Как следствие получаем, что выполнены асимптотические соотношения:

$$\frac{\kappa_n}{d_n}\sim\frac{k_1}{\delta_1}=\frac{51}{196},\quad \frac{\rho_n}{d_n}\sim\frac{r_1}{\delta_1}=\frac{373}{6144}.$$

Поскольку для последовательности d_n имеется явное выражение через числа Каталана, для величин $q(n)$ и $r(n)$ в силу определений последовательностей κ_n и ρ_n получаем следующие асимптотические (при $n\to\infty$) оценки:

$$q(n)\gtrsim\frac{51}{196}2^nc_n,\quad r(n)\gtrsim\frac{373}{6144}2^nc_n.$$

С учетом равенства $q(n)+r(n)=\frac{1}{2}2^nc_n$, можно дополнить нижние оценки верхними.

Теорема 3. *При $n \to \infty$ имеют место асимптотические неравенства:*

$$\frac{51}{196} \lesssim \frac{q(n)}{2^n c_n} \lesssim \frac{2699}{6144}, \quad \frac{373}{6144} \lesssim \frac{r(n)}{2^n c_n} \lesssim \frac{37}{196}.$$

Полученные оценки очевидно не являются окончательными, однако даже они позволяют сделать некоторые выводы о поведении величин $f(l,n)$. В частности, гипотеза о приближении с ростом n распределения значений $f(l,n)$ к какому-либо «равномерному», явно опровергается вытекающей из теоремы 3 оценкой $q(n) \gtrsim \frac{51}{196} > \frac{1}{4}$.

Дальнейшее уточнение оценок теоремы 3, по-видимому, возможно путем более детального рассмотрения способов представления двоичных дробей в виде произведений. Продвижение на этом пути, возможно, позволит применить аналогичные методы для оценки при растущем n сумм вида $\sum\limits_{\frac{t_1}{2^m}2^n < l < \frac{t_2}{2^m}2^n} f(l,n)$, где $m > 2$ — фиксированное число.

Список литературы

[1] Golumbic M. C., Gurvich V. A., *Boolean Functions: Theory, Algorithms and Applications*, eds. Crama Y., Hammer P.L., Cambridge University Press, Cambridge, 2011.

[2] Схиртладзе Р. Л., “О синтезе p-схемы из контактов со случайными дискретными состояниями”, *Сообщ. АН ГрузССР*, **26**:2 (1961), 181–186.

[3] Яшунский А. Д., “Об асимптотике вероятности значений случайных булевых выражений”, *Дискретн. анализ и исслед. опер., сер. 1*, **13**:2 (2006), 59–99.

[4] Flajolet Ph., Sedgewick R., *Analytic combinatorics*, Cambridge University Press, New York, NY, USA, 2009.

[5] Flajolet Ph., Odlyzko A. M., “Singularity analysis of generating functions”, *SIAM Journal on Algebraic and Discrete Methods*, **3**:2 (1990), 216–240.

On the weight of functions, defined by read-once AND/OR formulas

Eremenko A.R., Yashunsky A.D.

We consider the set of functions defined by read-once formulas with binary conjunction (logical AND) and disjunction (logical OR) operations. For the functions defined by formulas with a fixed number of operations we investigate the weights of functions, i.e. the number of tuples on which the function has value 1. We establish asymptotic bounds on the number of read-once formulas that define functions with specific weights, in particular on the number of formulas, that define functions with no more than a quarter of ones among their values.

Keywords: Boolean function, read-once formula, function weight, asymptotic bound.