О сложности сборки графов

Д. В. Зайцев

Введение

В работе изучается сложность схем построения графов при помощи двух операций склейки вершин, которые заключаются в отождествлении пары вершин с удалением петель и параллельных ребер. Первая операция применяется к паре вершин одного графа, вторая — к паре вершин двух графов, не имеющих общих элементов. В качестве простейшего берется граф, состоящий из одного ребра. Любой промежуточный результат, то есть построенный на каком-то этапе граф, разрешено использовать неоднократно. Такие схемы могут представлять интерес в связи со сжатием информации. Под сложностью схемы понимается число применений операций над графами. Получены оценки функции Шеннона для сложности таких схем, а также оценки сложности схемного задания некоторых конкретных графов. Схемный подход к заданию графов уже рассматривался ранее, в частности, в работе [1], однако в ней не изучался вопрос о сложности схемы построения графов, и использовались другие операции.

1. Основные определения и результаты

Пусть G — неориентированный связный граф, в котором нет петель и параллельных ребер. Определим две операции склейки вершин. Одноместная операция склейки — отождествление двух заданных вершин одного графа. Двуместная операция склейки — отождествление двух заданных вершин пары графов, не имеющих общих элементов. Всегда производится удаление петель и параллельных ребер.

Сборкой графа G назовем конечную последовательность графов $G_1, G_2, \ldots, G_T$ такую, что последний граф последовательности G_T изоморфен G, первый граф последовательности G_1 изоморфен ребру, каждый граф последовательности G_s, где $1 < s \leqslant T$, получен в результате применения одной операции склейки следующим образом: одноместная операция применяется к графу изоморфному одному из $G_1, ..., G_{s-1}$, двуместная операция склейки применяется к паре графов, каждый из которых изоморфен одному из $G_1, ..., G_{s-1}$. То есть каждый новый граф склеен из изоморфных предыдущим. Потребуем также, чтобы каждый граф использовался, то есть участвовал в получении хотя бы одного графа с большим номером в данной последовательности.

Пусть длина последовательности сборки G — количество операций склейки, затраченных для построения последовательности.

Сложностью $L(G)$ сборки G назовем минимальную длину последовательности сборки G.

Введем функции шенноновского типа, характеризующие сложность класса графов с точки зрения сборки. Пусть $L(n) = \max\limits_{|G|=n} L(G)$ — максимальная сложность в классе графов с n вершинами. Обозначим через D дерево с корнем. Пусть $D(n) = \max\limits_{|D|=n} L(D)$ — максимальная сложность в классе корневых деревьев с n вершинами.

Далее для удобства часто будем говорить не о склейках вершин, а о склейках графов, указывая множество пар склеиваемых вершин. Шагом сборки, или просто шагом, будем называть некоторое количество склеек графов, достраивающих сборку от одного указанного графа до другого.

Обозначим через $K_{p,q}$ полный двудольный граф с количеством вершин p и q в соответствующих долях. Обозначим через K_n полный граф с n вершинами. Получены следующие результаты.

Теорема 1. *Справедливы положения:*
а) при $q \geqslant p > 0$ для $K_{p,q}$ выполнено $L(K_{p,q}) \leqslant 3q + 5\log_2 q$;
б) $L(K_n) < \frac{7}{2}n + 8\log_2 n$.

Теорема 2. *Имеет место $\frac{n^2}{\log_2 n} \lesssim L(n) \lesssim \frac{n^2}{\sqrt{\log_2 n}}$ при $n \to +\infty$.*

Теорема 3. *Имеет место* $D(n) \asymp \frac{n}{\log_2 n}$ *при* $n \to +\infty$.

2. Доказательство теоремы 1

Собираем полный двудольный граф $K_{]\frac{n}{2}[,]\frac{n}{2}[}$. Строим последовательность графов, где на каждом шаге i граф K_{m_i,m_i} получается из $K_{]\frac{m_i}{2}[,]\frac{m_i}{2}[}$, склеенного на предыдущем шаге.

Пока объясняется принцип сборки, считаем, что число n является степенью двойки. Заметим, что здесь $n \geqslant 2$. При $n = 2$ имеем полный двудольный граф, состоящий из одного ребра. Берем его в качестве первого графа последовательности сборки. Шаг состоит из трех склеек графов (см. рис. 1). Каждая склейка графов в данном случае состоит из $\frac{m_i}{2}$ склеек вершин. Берутся два двудольных графа (с долями A, B и C, D, соответственно), каждый из которых изоморфен $K_{\frac{m_i}{2},\frac{m_i}{2}}$. Долю A склеим с C. Пары склеиваемых вершин произвольные. Получаем полный двудольный граф $K_{\frac{m_i}{2},m_i}$. Берем новый двудольный граф, изоморфный $K_{\frac{m_i}{2},m_i}$. Это возможно, так как $K_{\frac{m_i}{2},m_i}$ теперь добавлен в последовательность сборки. В большей доле нового графа выделяем подмножества вершин H и F по $\frac{m_i}{2}$ вершин в каждом. D склеим с F, B с H, так получим K_{m_i,m_i}. Таким образом, в последовательности сборки графа $K_{\frac{n}{2},\frac{n}{2}}$ содержатся все K_{m_i,m_i}, где $m_i = 2^i$, $i = 1, \ldots, \log_2 \frac{n}{2}$.

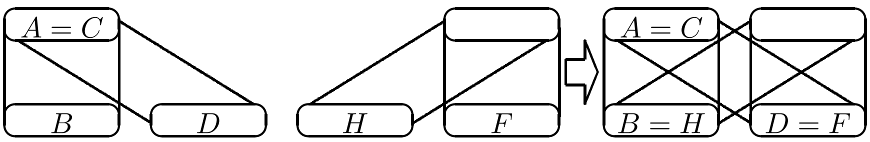

Рис. 1. Шаг сборки полного двудольного графа.

Воспользуемся полученными при сборке $K_{\frac{n}{2},\frac{n}{2}}$ двудольными графами для сборки полного графа K_n. Строим последовательность графов, где на каждом i шаге граф K_{m_i} получается из двух графов $K_{\frac{m_i}{2}}$, полученных на $i-1$ шаге ($K_{\frac{m_i}{2}} \cong K_{m_{i-1}}$), и одного $K_{\frac{m_i}{2},\frac{m_i}{2}}$. Берем граф K_2, состоящий из одного ребра, в качестве исходного. Шаг состоит из двух склеек графов. Каждая склейка графов состоит из $\frac{m_i}{2}$ склеек вершин (см. рис. 2).

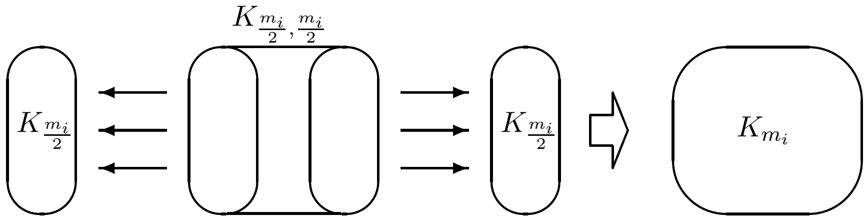

Рис. 2. Шаг сборки полного графа.

Склеиваем $K_{\frac{m_i}{2}}$ с одной из долей $K_{\frac{m_i}{2},\frac{m_i}{2}}$. Полученный граф склеиваем снова с графом, изоморфным $K_{\frac{m_i}{2}}$. Получаем окончательно K_{m_i}. Итак, мы можем построить $K_{\frac{n}{2},\frac{n}{2}}$ и K_n в случае, если n является степенью двойки.

Пусть теперь n — произвольное число вершин (не обязательно является степенью двойки). Получим верхние оценки сложности сборки полного двудольного графа и полного графа. Для этого рассмотрим худший случай, когда дополнительные склейки требуются на каждом шаге. Это возможно, когда m_i нечетно на каждом шаге, то есть при делении m_i на два с избытком получается нечетное число вершин. Числа, относящиеся к этому случаю, запишем в виде $m_i = 2^i + 1$. Действительно, $\left]\frac{m_i}{2}\right[= \left]2^{i-1} + \frac{1}{2}\right[= 2^{i-1} + 1$ — нечетное число того же вида. Для получения полного двудольного графа $K_{]\frac{m_i}{2}[,]\frac{m_i}{2}[}$ из $K_{]\frac{m_i}{4}[,]\frac{m_i}{4}[}$ сначала склеиваем доли, как было описано выше (см. рис 1). Так получаем граф $K_{]\frac{m_i}{2}[+1,]\frac{m_i}{2}[+1}$ за $3 \cdot \left]\frac{m_i}{4}\right[$ склеек. В этом графе убираем две лишние вершины одной склейкой в каждой доле. Для получения полного графа сначала склеиваем два графа $K_{]\frac{m_i}{2}[}$ при помощи графа $K_{]\frac{m_i}{2}[,]\frac{m_i}{2}[}$ (см. рис. 2), затем убираем лишнюю вершину. Требуется одна дополнительная склейка. В описанном случае верно, что
$L\left(K_{]\frac{n}{2}[,]\frac{n}{2}[}\right) \leqslant L\left(K_{]\frac{n}{4}[,]\frac{n}{4}[}\right) + 3 \cdot \left]\frac{n}{4}\right[+ 2 = L\left(K_{]\frac{n}{8}[,]\frac{n}{8}[}\right) + 3\left]\frac{n}{8}\right[+ 2 + {} + 3\left]\frac{n}{4}\right[+ 2 < \frac{3}{2}n + 5 \cdot [\log_2 n] \leqslant \frac{3}{2}n + 5\log_2 n$ и
$L(K_n) - L\left(K_{]\frac{n}{2}[,]\frac{n}{2}[}\right) \leqslant L\left(K_{]\frac{n}{2}[}\right) + 2 \cdot \left]\frac{n}{2}\right[+ 1 < 2n + 3[\log_2 n] \leqslant 2n + {} + 3\log_2 n$. Значит, $L(K_n) < \frac{7}{2}n + 8\log_2 n$.

До сих пор мы рассматривали полные двудольные графы, у которых доли имеют одинаковое количество вершин. Обобщим результат для полных двудольных графов с произвольным количеством вершин в каждой доле. Рассмотрим полный двудольный граф $K_{p,q}$. Пусть $p \leqslant q$ для определенности. Рассмотрим снова худший случай, когда $p = 2^k + 1$, $q = 2^l + 1$, при некоторых k, l. На каждом шаге надо получать граф с нечетным числом вершин в каждой доле, поэтому требуются дополнительные склейки. Достаточно показать, что в этом случае $L(K_{p,q}) \leqslant L(K_{q,q})$.

Пусть построен граф $K_{p,p}$ таким способом, который был описан раньше. Доля с p вершинами получена. Дальнейшие шаги будем проводить так, как показано на рис. 3.

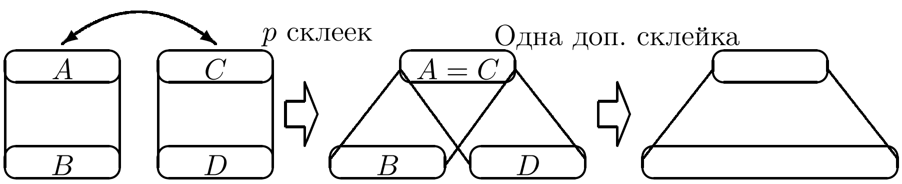

Рис. 3. Шаг сборки полного двудольного графа $K_{p,q}$ на втором этапе.

Долю A склеим с долей C за p склеек вершин. В большей доле нового графа удалим лишнюю вершину одной дополнительной склейкой. Через $l-k$ шагов будет получен граф $K_{p,q}$. Видно, что начиная с шага, на котором получена меньшая доля с p вершинами, количество склеек в шаге не меняется и равна $p+1$. Однако при сборке графа $K_{q,q}$ количество склеек на каждом шаге продолжает увеличиваться. Это доказывает справедливость последнего неравенства.

Получаем, что

$$L(K_{p,q}) \leqslant L\left(K_{q,q}\right) \leqslant L\left(K_{\left]\frac{q}{2}\right[,\left]\frac{q}{2}\right[}\right) + 3 \cdot \left]\frac{q}{2}\right[+ 2 < 3q + 5\log_2 q.$$

Теорема доказана.

3. Доказательство теоремы 2

Любой граф G можно склеить следующим тривиальным способом. Если G состоит из одной вершины, то мы получаем его из ребра одной одноместной склейкой. Если G изоморфен ребру, то сборка уже есть по определению. Пусть G содержит несколько ребер. Допустим, уже построен некоторый граф G_s, который изоморфен подграфу графа G. Всегда одной или двумя склейками можно склеить G_s с ребром так, что полученный в результате склейки граф G_{s+1} будет иметь на одно ребро больше и снова будет изоморфен подграфу графа G. Поскольку ребер конечное число, на некотором шаге будет получен граф, который изоморфен графу G.

Самым сложным случаем для тривиальной сборки является полный граф. В этом случае имеем сложность $2 \cdot C_n^2 = 2\frac{n(n-1)}{2} = n^2 - n$ из расчета, что каждое ребро присоединяется двумя склейками.

Предложенный ниже способ позволяет получить сборку произвольного графа G с n вершинами со сложностью по порядку меньше n^2 при $n \to +\infty$.

Лемма 1. *При $n \to +\infty$ имеет место $L(n) < C_n \frac{n^2}{\sqrt{\log_2 n}}$, где $C_n \to \frac{3}{\sqrt{2}} + 0$.*

Доказательство. Разобьем множество вершин графа G на подмножества мощности не больше R из $\mathbb{N}$. Получим, что множество вершин G разбито на $k = \left]\frac{n}{R}\right[$ частей. Пусть это будет $l = k - 1 = \left[\frac{n}{R}\right]$ целых частей по R вершин и остаточная часть, где число вершин меньше R. Остаточная часть участвует в получении верхней оценки наравне с остальными. Для сборки графа G поступаем так.

1) Сначала соберем тривиальным способом k графов, в которых число вершин не превосходит R, так, что каждый из них изоморфен соответствующему подграфу графа G, индуцированному множеством вершин одной из частей разбиения.

2) Соберем также все двудольные графы с количеством вершин в каждой доле не большем R.

3) После соберем G. Для этого заполним l целых частей и остаточную часть соответствующими графами из множества, полученного на

первом этапе. Соединим части двудольными графами из множества всевозможных двудольных графов, полученных на втором этапе.

Оценим сложность такой сборки. Рассмотрим следующие верхние оценки. Используем неравенство $\left]\frac{n}{R}\right[< \frac{n}{R} + 1$.
$C_0(R) = R(R-1)$ — верхняя оценка сложности сборки тривиальным способом одного произвольного графа, у которого число вершин не превосходит R.
$C_1(R) = \left]\frac{n}{R}\right[< \frac{n}{R}+1$ — верхняя оценка количества графов, в каждом из которых не более R вершин, соответствующих $k = \left]\frac{n}{R}\right[$ частям.
$C_2(R) = 2R^2$ — верхняя оценка сложности сборки произвольного двудольного графа, у которого в каждой доле не больше R вершин.
$C_3(R) = 2^{R^2}$ — верхняя оценка количества двудольных графов, у которых число вершин в каждой доле не больше R.
$C_4(R) = \frac{1}{2}\cdot\left]\frac{n}{R}\right[\left(\left]\frac{n}{R}\right[-1\right) < \frac{1}{2}\left(\frac{n}{R}+1\right)\frac{n}{R}$ — верхняя оценка количества пар частей разбиения G.
$C_5(R) = 2R$ — верхняя оценка сложности склейки одной пары графов, в каждом из которых не более R вершин, с помощью двудольного графа.

Для верхней оценки сложности сборки произвольного графа G с n вершинами имеем

$$L(n) < C_0 \cdot C_1 + C_2 \cdot C_3 + C_4 \cdot C_5 = \frac{n^2}{R} + 2R^2 \cdot 2^{R^2} + nR + R^2 - R.$$

С одной стороны, увеличение R (количества вершин в одной части разбиения) замедляет рост сложности с увеличением n; с другой — видно, что если взять $R = n$ (то есть максимально возможную величину R), то рост сложности становится экспоненциальным. Найдем по возможности большее $R(n)$ такое, чтобы значение $2^{R^2(n)} \cdot 2R^2(n)$ стремилось к $\frac{n^2}{R(n)}$ слева при $n \to +\infty$.

Итак, необходимо, чтобы при достаточно больших n выполнялось неравенство

$$\frac{n^2}{R} \geqslant 2^{R^2} \cdot 2R^2. \qquad (*)$$

Будем искать $R(n)$ в виде $\left[\sqrt{\log_2 n^{m(n)}}\right]$, $m(n) \geqslant 0$ таким образом.

1) Умножив обе части неравенства $(*)$ на $R(n)$, получим
$n^2 \geqslant 2^{\left[\sqrt{\log_2 n^m}\right]^2} \cdot 2\left[\sqrt{\log_2 n^m}\right]^3$.
2) Рассмотрим неравенства
$n^2 \geqslant 2^{\log_2 n^m} \cdot 2(\log_2 n^m)^{\frac{3}{2}} \geqslant 2^{\left[\sqrt{\log_2 n^m}\right]^2} \cdot 2\left[\sqrt{\log_2 n^m}\right]^3$.
Если верно первое неравенство, то верно и второе, поскольку $\sqrt{\log_2 n^m} \geqslant \left[\sqrt{\log_2 n^m}\right]$.
3) Обе части положительны, поэтому логарифмируя, получим
$2\log_2 n \geqslant m(n) \cdot \log_2 n + 1 + \frac{3}{2}\log_2(\log_2 n^{m(n)})$.
4) Раскрываем последнее слагаемое и умножаем на два, тогда
$4\log_2 n \geqslant 2 \cdot m(n)\log_2 n + 3 \cdot \log_2 m(n) + 3 \cdot \log_2(\log_2 n) + 2$.
5) Ищем $m(n)$ в виде $2 - \varepsilon(n)$, где $\varepsilon(n) \to 0+0$, тогда
$4\log_2 n \geqslant 4\log_2 n - 2\varepsilon(n)\log_2 n + 3\log_2(2-\varepsilon(n)) + 3\log_2(\log_2 n) + 2$,
и после сокращения
$0 \geqslant 3\log_2(2-\varepsilon(n)) + 2 + 3\log_2(\log_2 n) - 2\varepsilon(n)\log_2 n$.
6) Подберем $\varepsilon(n)$ так, чтобы сократились последние два слагаемых, и оставшиеся слагаемые стремились к нулю слева. Возьмем

$$\varepsilon(n) = \frac{3 \cdot \log_2(\log_2 n) + 5}{2 \cdot \log_2 n}, \quad \varepsilon(n) \to 0+0.$$

Проверим, что выбранное $\varepsilon(n)$ подходит. После подстановки $\varepsilon(n)$ и сокращений получаем

$$0 \geqslant 3\log_2\left(2 - \frac{3\log_2(\log_2 n) + 5}{2\log_2 n}\right) - 3.$$

Выражение в правой части стремится к нулю слева при $n \to +\infty$. Это доказывает, что

$$R(n) = \left[\sqrt{\log_2 n^{m(n)}}\right] = \left[\sqrt{\log_2 n^{2-\varepsilon(n)}}\right] = \left[\sqrt{\log_2 n^{2-\frac{3\log_2(\log_2 n)+5}{2\log_2 n}}}\right]$$

удовлетворяет нашим требованиям.

Вернемся к оценке сложности сборки произвольного графа G с n вершинами. При достаточно больших n верно

$$L(n) < \frac{n^2}{R} + 2R^2 \cdot 2^{R^2} + nR + R^2 - R < \frac{3n^2}{R}.$$

Напомним, что $\varepsilon(n) \to 0+0$ при $n \to +\infty$.

$$R(n)=\left[\sqrt{\log_2 n^{2-\varepsilon}}\right] > \sqrt{\log_2 n^{2-\varepsilon}}-1=\sqrt{2-\varepsilon}\cdot\sqrt{\log_2 n}-1=c_n\sqrt{\log_2 n},$$

здесь $c_n = \sqrt{2-\varepsilon_n} - \frac{1}{\sqrt{\log_2 n}} \to \sqrt{2}$ слева при $n \to +\infty$. Значит, окончательно имеем

$$L(n) < C_n\frac{n^2}{\sqrt{\log_2 n}}, \text{ где } C_n \to \frac{3}{\sqrt{2}}+0.$$

Лемма доказана.

Перейдем к рассмотрению нижней оценки сложности сборки произвольного графа G с n вершинами. Согласно определению, сборка графа G представляет собой последовательность графов

$$G_1 \xrightarrow{1} G_2 \xrightarrow{2} \ldots \xrightarrow{s-1} G_s \xrightarrow{s} \ldots \xrightarrow{T-1} G_T,$$

где G_1 — ребро, G_T изоморфен G, $1 \leqslant s \leqslant T$. По определению, каждый полученный в процессе сборки граф G_s используется для получения G. Везде дальше будем считать, что шаг сборки состоит из одной склейки.

Пусть B_s — множество вершин графа G_s. Оценим, в каких пределах может меняться количество вершин $|B_s|$ графа G_s в зависимости от номера шага s, и чем ограничена длина сборки T.

Лемма 2. *1) Для любой сборки длины T произвольного графа G с n вершинами и для любого шага сборки с номером s $(1 \leqslant s \leqslant T)$, количество вершин $|B_s|$ графа G_s, полученного на шаге s, удовлетворяет системе неравенств*

$$\begin{cases} |B_s| \leqslant 2^{s-1}+1 \\ |B_s| \leqslant -s+n^2 \\ |B_s| \geqslant 1. \end{cases} \qquad (**)$$

2) Длина T любой сборки произвольного графа G с n вершинами удовлетворяет неравенствам

$$\log_2(n-1)+1 \leqslant T \leqslant n^2-n. \qquad (***)$$

Доказательство. Справедливость первого неравенства в системе $(**)$ следует из определения сборки. Пусть мы на каждом шаге выбираем среди полученных графов тот, у которого самое большое число вершин. Применение двуместной операции склейки к паре графов, изоморфных выбранному, позволяет получить граф, у которого не более чем в два раза больше вершин. Одна вершина исчезает, поскольку при склейке отождествляется пара вершин. Значит мощности множеств вершин получаемых графов при самом быстром росте образуют последовательность $2, 3, 5, 9, \ldots, 2^{s-1} + 1, \ldots$

Из сказанного также следует, что верно первое неравенство в системе $(***)$. Действительно, окончание сборки означает, что $|B_s| = n$, при $s = T$. Значит сборка может быть окончена, только когда T удовлетворяет неравенству

$$2^{T-1} + 1 \geqslant n, \quad \text{то есть} \quad T \geqslant \log_2(n-1) + 1.$$

Второе неравенство в $(* * *)$ является грубой верхней оценкой сложности сборки произвольного графа G с n вершинами. Лемма 1 дает более точную верхнюю оценку, но нам будет достаточно тривиальной оценки, поскольку это не повлияет на порядок получаемой далее нижней оценки сложности сборки.

Справедливость второго неравенства в системе $(**)$ также следует из определения сборки. Уменьшать число вершин в построенном графе можно только одноместной операцией склейки. За один шаг можно удалить одну вершину. Отсюда следует, что зависимость количества вершин от номера шага s в этом случае надо искать в виде $-s + C$, где C константа. Для получения очередного графа производится одна из операций склейки над графами, которые изоморфны полученным ранее. По определению, все полученные в процессе сборки графы используются, то есть используются все их вершины. Имеется в виду, что копии этих вершин, которые появляются при взятии изоморфных графов, войдут в очередной полученный граф, либо будут использованы при склейке вершин. Это говорит о том, что число вершин нельзя резко уменьшить. Пусть на шаге s получен граф G_s, и пусть все последующие шаги направлены на уменьшение числа вершин в получаемых графах. Значит число вершин зависит от номера шага следующим образом: $|B_s| = -s + C$. Окончание сборки означает, что $|B_s| = n$, при $s = T$. Получаем, что $-T + C = n$. Согласно грубой оценке $T \leqslant n^2 - n$, значит $-n^2 + n + C \leqslant n$. Отсюда находим $C \leqslant n^2$. Выходит, что $|B_s| \leqslant -s + n^2$, иначе, используя граф G_s, нельзя получить граф G, имеющий n вершин.

Третье неравенство в системе $(**)$ верно потому, что по определению графа множество его вершин не пусто, и операции склейки сохраняют это свойство.
Лемма доказана.

Лемма 3. *Число вершин любого графа G_s, участвующего в сборке графа G с n вершинами, меньше n^2.*

Доказательство. Достаточно доказать, что n^2 является верхней оценкой максимума функции $f(s) = \min(2^{s-1} + 1, -s + n^2)$. Легко заметить, что на отрезке $s \in [0, n^2 - n]$ функция $-s + n^2$ имеет максимум в нуле равный n^2. В свою очередь, $n^2 > \max\limits_{s \in [0, n^2 - n]} f(s)$ поскольку $-s + n^2$ монотонно убывает.
Лемма доказана.

Обозначим через $N(n, T)$ множество всевозможных сборок графа G с n вершинами длины T.

Лемма 4. *Имеет место $|N(n, T)| < n^{4T} \cdot T^{2T} \cdot 2^{1-T}$.*

Доказательство. Пусть $V(s)$ — количество способов продолжить сборку на шаге s. Покажем, что

$$V(s) = \frac{s(s-1)}{2} n^4 + s\frac{n^2(n^2+1)}{2} + s\frac{n^2(n^2-1)}{2}.$$

Рассмотрим каждое слагаемое подробнее.
$\frac{s(s-1)}{2}$ — количество способов образовать пару графов для склейки на шаге s. Здесь учтены только пары различных графов.
$n^4 = (n^2)^2$ — число способов склеить два различных графа из предыдущего пункта. n^2 — оценка количества вершин в каждом графе, по лемме 3.

s — количество способов образовать пару из двух изоморфных графов. На шаге сборки s мы имеем не более s различных графов. Выбираем граф, один из s, и клеим со своей копией.

$\frac{n^2(n^2+1)}{2} = C_{n^2}^2 + n^2$ — число пар вершин двух изоморфных графов. Прибавляется n^2, поскольку любая вершина может быть склеена также со своей копией из второго графа.

s из последнего слагаемого — количество графов, которые мы можем выбрать на шаге s, если производится склейка двух вершин, принадлежащих одному графу.

$\frac{n^2(n^2-1)}{2} = C_{n^2}^2$ — число пар вершин одного графа с n^2 вершинами.

После упрощения получаем $V(s) = n^4 \frac{s(s+1)}{2}$.

Наконец, посчитаем верхнюю оценку количества сборок $|N(n,T)|$. Для этого заметим, что

$$|N(n,T)| = \prod_{s=1}^{T-1} V(s) = \prod_{s=1}^{T-1} n^4 \cdot \prod_{s=1}^{T-1} \frac{s(s+1)}{2} = \\ = n^{4(T-1)} \cdot \frac{(T-1)!T!}{2^{T-1}} < n^{4T} \frac{T^{2T}}{2^{T-1}}.$$

Лемма доказана.

Обозначим через S_n множество графов с n вершинами. Нижняя оценка мощности множества S_n такова

$$\frac{2^{\frac{n(n-1)}{2}}}{n^{2n}} < |S_n| \, .$$

Лемма 5. *Имеет место* $\frac{1}{16} \cdot \frac{n^2}{\log_2 n} < L_(n)$.

Доказательство. Нижняя оценка доказывается из мощностных соображений. Найдем максимальную длину (сложность) T такую, что верно неравенство

$$\frac{|S(n)|}{|N(n,T)|} > 1,$$

где, как и выше, S — множество графов с n вершинами, N — множество сборок n-вершинных графов длины T.

Справедливость неравенства означает, что всех сборок длины T недостаточно, чтобы построить все графы из S, и длину необходимо увеличивать. Воспользуемся оценками для $|S(n)|$ и $|N(n,T)|$, тогда получим

$$\frac{2^{\frac{n(n-1)}{2}} \cdot n^{-2n}}{n^{4T} \cdot T^{2T} \cdot 2^{1-T}} > 1,$$
$$n^2 > n + 4n\log_2 n + 8T\log_2 n + 4T\log_2 T + 2 - 2T.$$

Будем искать T в виде $T = \frac{n^2}{c\log_2 n}$, тогда
$n^2 > 8\frac{n^2}{c\log_2 n}\log_2 n + 4\frac{n^2}{c\log_2 n}\log_2 n^2 - 4\frac{n^2}{c\log_2 n}\log_2(c\log_2 n) + n + 4n\log_2 n + 2 - 2\frac{n^2}{c\log_2 n}$,
$n^2 > \frac{16}{c}n^2 - 4n^2\frac{\log_2(c\log_2 n)}{c\log_2 n} - 2\frac{n^2}{c\log_2 n} + n + 4n\log_2 n + 2$,
$n^2 > \frac{16}{c}n^2 - \frac{n^2}{c\log_2 n}(4\log_2(c\log_2 n) + 2) + n(4\log_2 n + 1) + 2$.
Для выполнения неравенства, начиная с некоторого n, достаточно взять $c = 16$.
Лемма доказана.

Утверждение теоремы 2 следует из леммы 1 и леммы 5.

4. Доказательство теоремы 3

Количество вершин дерева n связано с количеством ребер m как $n = m+1$. Тривиальным способом подклейки по одному ребру любое дерево с n вершинами можно собрать со сложностью $n-1$.

Лемма 6. *Имеет место* $D(n) < 6 \cdot \frac{n}{\log_2 n}$.

Доказательство. Разобьем дерево D на поддеревья с не более R вершинами. Количество поддеревьев равно $\left]\frac{n}{R}\right[$. На сборку D из частей потребуется по одной склейке на каждую часть, всего $\left]\frac{n}{R}\right[$. Оценим сверху количество деревьев с R вершинами. 2^{2R} — известная верхняя оценка. Сборка каждого R-вершинного поддерева тривиальным способом потребует не более R склеек. Имеем $D(n) < 1 + \frac{n}{R} + R \cdot 2^{2R}$.

Будем искать R в виде $R=\frac{1}{c}\log_2 n$. Добьемся, чтобы $1+\frac{n}{R}\geqslant R\cdot 2^{2R}$ было как можно ближе к равенству, тогда оценка D будет иметь вид

$$D(n) < \frac{2n}{R} = \frac{2n}{\frac{1}{c}\log_2 n} = 2c\cdot\frac{n}{\log_2 n}.$$

Подберем константу c так:

$$n \geqslant R^2\cdot 2^{2R} - R,$$
$$n \geqslant \frac{1}{c^2}\log_2^2(n)\cdot n^{\frac{2}{c}} - \frac{1}{c}\log_2 n.$$

Достаточно взять константу $c > 2$. Пусть $c = 3$, тогда получаем оценку из условия леммы

$$D(n) < 6\cdot\frac{n}{\log_2 n}.$$

Лемма доказана.

Получим нижнюю оценку из мощностных соображений. Для этого сначала найдем верхнюю оценку мощности множества $N_D(n,T)$ сборок корневых деревьев с n вершинами длины T и нижнюю оценку мощности множества корневых деревьев $S_D(n)$.

Лемма 7. *Имеет место* $|N_D(n,T)| < 3^{T-1}\cdot T^{2T}\cdot n^{2T-2}$.

Доказательство. Найдем число $V_D(s)$ способов продолжить сборку дерева на шаге s. Покажем, что

$$V_D(s) = \frac{s^2-s}{2}\cdot 4n^2 + s\cdot\frac{4n^2+2n}{2} + s\cdot\frac{4n^2-2n}{2}.$$

Число способов образовать пару из различных деревьев последовательности длины s равно $\frac{s(s-1)}{2}$. Максимальное число вершин, которое может быть у дерева, входящего в сборку, не превосходит $2n$. Оценка получена способом, который подробно описан в леммах 2 и 3. $4n^2$ — число способов склеить два различных дерева. s из второго слагаемого — количество способов образовать пару из двух изоморфных деревьев. $\frac{4n^2+2n}{2}$ — число пар вершин двух изоморфных деревьев. s из последнего слагаемого — количество деревьев, которые мы можем выбрать на шаге s, если производится склейка двух вершин, принадлежащих одному дереву. $\frac{4n^2-2n}{2}$ — число пар вершин одного дерева с $2n$ вершинами. После упрощения получаем $V_D(s) = 2s^2n^2 + 2sn^2 < < 3(s+1)^2n^2$. Отсюда следует, что

$$|N_D(n,T)| = \prod_{s=1}^{T-1} V_D(s) < 3^{T-1} \cdot T^{2T} \cdot n^{2T-2}.$$

Лемма доказана.

Лемма 8. *Имеет место* $\frac{1}{4} \cdot \frac{n}{\log_2 n} < D(n)$.

Доказательство. Из мощностных соображений найдем максимальное T, при котором верно неравенство

$$\frac{|S_D(n)|}{|N_D(n,T)|} > \frac{2^n}{3^{T-1} \cdot T^{2T} \cdot n^{2T-2}} > 1,$$
$$n > (T-1)\log_2 3 + 2T \log_2 T + (2T-2)\log_2 n.$$

Ищем T в виде $\frac{n}{c \log_2 n}$. Имеем
$n > 2\frac{n}{c\log_2 n}\log_2(\frac{n}{c\log_2 n}) + 2\frac{n}{c\log_2 n}\log_2 n + \frac{n}{c\log_2 n}\log_2 3 - 2\log_2 n - \log_2 3$,
$n > \frac{2n}{c} - 2\frac{n}{c\log_2 n}\log_2(c\log_2 n) + \frac{2n}{c} + \frac{n}{c\log_2 n}\log_2 3 - 2\log_2 n - \log_2 3$,
$n > \frac{4n}{c} + \frac{n}{c\log_2 n}\log_2 3 - 2\frac{n}{c\log_2 n}\log_2(c\log_2 n) - 2\log_2 n - \log_2 3$.

Для выполнения неравенства, начиная с некоторого n, достаточно взять $c = 4$. Лемма доказана.

Утверждение теоремы 3 следует из лемм 6 и 8.

В заключение автор выражает благодарность профессору А. С. Подколзину, под руководством которого выполнена эта работа.

Список литературы

[1] Яблонский С. В. Введение в дискретную математику. — М.: Высшая школа, 2002.

[2] Харари Ф., Палмер Э. Перечисление графов. — М.: Мир, 1977.