ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное
учреждение высшего профессионального образования
Национальный исследовательский университет
«Высшая школа экономики»

Московский институт электроники и математики им. А.Н. Тихонова
Департамент прикладной математики
Кафедра «Компьютерная безопасность»

КУРСОВАЯ РАБОТА
по дисциплине «Теория информации»

«Оценка числа компонент графа итерации случайного отображения»

Выполнил:
студент группы СКБ-151
Михалицын П.К.
Проверил:
старший преподаватель кафедры
«Компьютерная безопасность»
Миронкин В.О.

МОСКВА 2018

Оглавление

Введение **3**

1 Оценка числа компонент графа отображения **4**

1.1 Случай любого случайного отображения 4

1.2 Случай биективного случайного отображения 8

1.3 Анализ полученных формул 9

2 Оценка числа компонент графа итерации случайного отображения **11**

2.1 Основная часть 11

2.2 Оценка числа максимального числа компонент при итерации случайного отображения 15

2.3 Оценка числа компонент итерации биективного случайного отображения 16

2.4 Анализ полученных формул 17

2.4.1 Анализ основной формулы 17

2.4.2 Анализ второго раздела 18

2.4.3 Анализ третьего раздела 19

Вывод **21**

Литература **21**

Введение

В данной работе мы выведем формулы среднего числа компонент, получаемых при итерации случайного отображения. Отдельно будут рассмотрены случаи биективного отображения и оценка среднего максимального числа компонент при итерации. В вводной части, будет рассмотрен случай, когда итерации не производятся – для лучшего освоения материала и вывода основных вспомогательных формул, которые мы рассмотрим в уже основной части данной работы.

Эта работа будет полезна в алгоритмах шифрования и выработки ключа, особенно в получении оценки максимального количества итераций, при которых можно использовать данный ключ при взятии случайного отображения, для уменьшения вероятности вхождения в цикл. Таким образом можно оценивать количество непересекающихся множеств ключей, которые будут создавать те самые компоненты графа.

Существует возможность расширения области применения данной работы до приложений в сферах, которые были указаны выше. Всего этого можно добиться, без труда выведя небольшое количество формул. т.к. основные формулы и концепции их получения будут рассмотрены в данной работе.

Оценка числа компонент графа отображения

1.1 Случай любого случайного отображения

Прежде чем начать поиск числа компонент графа случайного отображения, определим как выглядит граф этого отображения.

В первую очередь рассмотрим пример графа случайного отображения, чтобы понять с чем мы будем иметь дело и что считать.

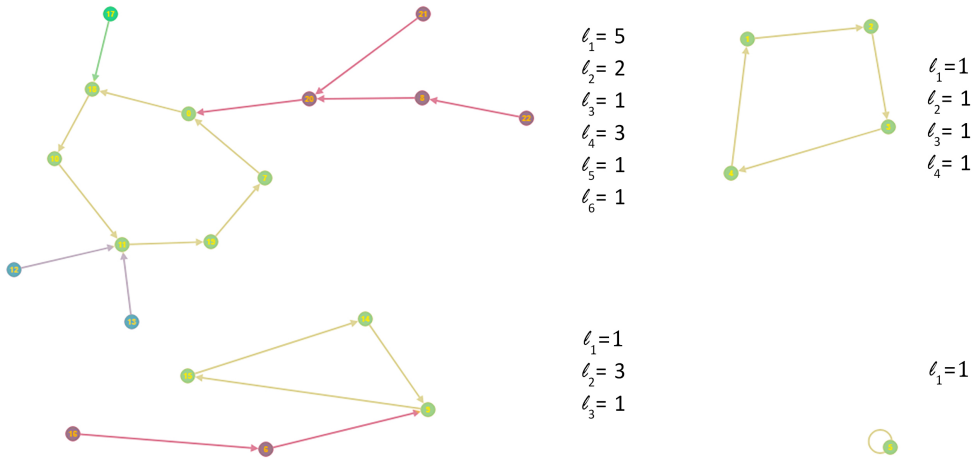

Рис. 1.1: Пример графа случайного отображения

Очевидно, что будет образовываться некоторое подобие леса. Причем каждая компонента этого леса будет ориентированным графом ровно с одним циклом, более того все ребра будут направлены в сторону этого цикла.

Теперь давайте приступим к подсчету количества таких графов. Для начала начнем с одной компоненты. Введем следующие параметры, которые могут определить данную компоненту и определим число компонент с данными параметрами.

1. n – кол-во вершин в компоненте
2. $l_1, l_2, ..., l_k$ – число вершин, находящихся в "побочных путях", ведущих до циклической части графа, (каждая из них больше 0 и при этом в сумме они дают n). Стоит заметить, что под побочными путями имеются в виду ориентированные деревья, каждая из вершин которых имеет однозначный путь до цикла (в силу того, что это деревья)
3. k – число вершин в цикле

Для наглядности на рисунке 1.1 желтым цветом определен цикл, а другими цветами определены побочные компоненты и указаны соответствующие параметры графа.

Приступим к подсчетам.

Для первого "побочного пути" нужно выбрать какие вершины будут в нем состоять ($\frac{n!}{l_1!(n-l_1)!}$ способов) После этого нужно будет разместить их на дереве, которое будет соединено впоследствии с циклом. Согласно теореме Келли о числе деревьев : $l_1^{l_1-2}$ – количество различных деревьев, которые можно построить на l_1 вершинах. После этого на собранном дереве нужно будет выбрать вершину, которая окажется в цикле (l_1 способов). Очевидно, что выбирая на дереве вершину, которая будет в цикле мы всегда однозначным образом определяем направление, в котором будут идти стрелки от каждой другой вершины до выбранной(т.к. в дереве, кратчайший путь от одной вершины к другой всегда однозначен и в нем отсутствуют циклы). Итого получим кол-во способов выбора такого "побочного пути":

$$\frac{n!}{l_1!(n-l_1)!}l_1^{l_1-1}$$

Проделав то же самое для второго побочного пути, получим кол-во возможных комбинаций:

$$\frac{(n-l_1)!}{l_2!(n-l_1-l_2)!}l_2^{l_2-1}$$

и т.д. для kой компоненты получим кол-во способов:

$$\frac{(n-l_1-...-l_{k-1})!}{l_k!(n-l_1-l_2-...-l_k)!}l_k^{l_k-1} = \frac{(n-l_1-...-l_{k-1})!}{l_k!}l_k^{l_k-1}$$

Осталось определить как будут расположены вершины на цикле ($(k-1)!$ способов. Мы убрали оттуда k т.к. любую из этих вершин можно считать начальной)

Перемножая полученные комбинаторные значения, получаем итоговую формулу:

$$(k-1)!\frac{n!}{l_1!(n-l_1)!}l_1^{l_1-1} * ... * \frac{(n-l_1-...-l_{k-1})!}{l_k!}l_k^{l_k-1} = (k-1)!C_n^{l_1,l_2,...,l_k}l_1^{l_1-1}l_2^{l_2-1}...l_k^{l_k-1}$$

, где $C_n^{l_1,l_2,...,l_k}$ – мультиномиальный коэффициент,

Теперь нужно просуммировать данные значения при различных $l_1, l_2, ..., l_k$ и учитывая возможное возникновение повторений при подсчете, получим итоговую формулу для подсчета количества числа таких компонент.

$$\sum_{k=1}^{n}\frac{1}{k}\sum_{l_1+l_2+...+l_k=n}C_n^{l_1,l_2,...,l_k}l_1^{l_1-1}l_2^{l_2-1}...l_k^{l_k-1} =$$

Полученные числа, показывают кол-во определенных выше компонент графа (один цикл и все побочные компоненты ведут к нему), которые можно построить на n вершинах. Обозначим их $M(n)$.

Стоит заметить, что данные числа встречаются в работе Дональда Кнута "Искусство программирования". В онлайн энциклопедию целочисленных последовательностей они занесены под номером A001865 и обозначают количество функций, у которых соответствующий им граф отображения имеет одну компоненту связанности (Number of connected functions on n labeled nodes) [1].

В этой же энциклопедии имеется более простая формула для вычисления этих последовательностей, а именно имеет место теорема, которую мы оставим без доказательства

Теорема 1.1. *Суммарное число компонент графов случайных эндоморфизмов над алфавитом мощности n можно подсчитать по формуле*

$$\sum_{k=1}^{n} \frac{n! n^{n-k-1}}{(n-k!)} \tag{1.1}$$

Причем кое слагаемое внутри суммы – это количество искомых графов, в которых содержится цикл длины k.

Зная данный факт, давайте попробуем найти экспоненциальную производящую функцию для данных чисел, т.к. после этого вычисление данных чисел будет занимать всего одну строчку кода в пакете Wolfram Mathematica и к тому же этот подсчет будет происходить гораздо быстрее.

Теорема 1.2. *Экспоненциальной производящей функцией для количества отображений, которые имеют одну компоненту связанности, является*

$$-\ln(1+W(-x))$$

, где $W(x)$ *–W функции Ламберта [2]*

Доказательство.

$$W(x) = \sum_{n=0}^{\infty} (-1)^{n-1} \frac{n^{n-1} x^n}{n!}$$

Зная замечательное свойство ряда Тейлора для степени W-функции Ламберта:

$$W^k(x) = -k \sum_{n=k}^{\infty} \frac{(-n)^{n-k-1}}{(n-k)!} x^n$$

увидим, что данная формула очень напоминает слагаемые той, которую мы нашли в онлайн энциклопедии для наших чисел. А именно, если мы рассмотрим такую функцию:

$$(-1)^k \frac{W^k(-x)}{k} = \sum_{n=k}^{\infty} \frac{n^{n-k-1}}{(n-k)!} x^n \tag{1.2}$$

, то заметим, что именно она будет являться экспоненциальной производящей функцией для количества односвязанных искомых графов с циклом длины k

Теперь, если мы рассмотрим ряд Тейлора от функции:

$$-W(-x) + \frac{W^2(-x)}{2} - \frac{W^3(-x)}{3} + \frac{W^4(-x)}{4} - ... =$$

, то коэффициент при x^n будет в точности как из формулы (1.1). Обнаружим, что такой ряд имеет именно функция

$$-\ln(1+W(-x))$$

■

Осталось только определить число графов, состоящих из таких компонент, построенных на n вершинах. Если в каждой компоненте будет n_i вершин, а всего компонент k, то искомое число таких графов

$$\frac{1}{k!} \sum_{n_1+n_2+...+n_k=n} C_n^{n_1,n_2,...,n_k} M(n_1) M(n_2) ... M(n_k) \tag{1.3}$$

Предложение 1.1. Экспоненциальная производящая функция для этих чисел будет

$$(-1)^k \frac{1}{k!} \ln^k(1 + W(-x)) \tag{1.4}$$

Действительно, если мы будем возводить наш ряд в степень k, то перед x^n будет находиться число, стоящее в формуле в соответствии со свойствами возведения в степень экспоненциальных производящих функций

Данная особенность получения чисел с помощью экспоненциальной производящей функции позволит нам упростить расчеты, сделать их более эффективными в плане вычислений. Так на компьютере автора поиск искомых чисел с помощью формулы (1.3) позволял работать только с алфавитом мощности 20-30, в то время как работа с помощью формулы (1.4) позволила добиться увеличения производительности в десятки раз. В анализе полученных результатов мы получим явные формулы для данных чисел, которые повысят производительность вычислений уже в сотни раз.

Теперь приступим к подсчету чисел, ради которых и была написана данная глава этой работы–подсчету суммарного числа компонент всех графов отображений. Зная суммарное число компонент можно без особых трудностей получить среднее поделив суммарное число компонент на количество всех функций (n^n).

Теорема 1.3. *Суммарное число компонент графов всех эндоморфизмов можно получить с помощью экспоненциальной производящей функции*

$$\frac{-\ln(1 + W(-x))}{(1 + W(-x))}$$

Доказательство. Поскольку мы знаем, как получить число наших графов с n вершинами обладающими k компонентами (воспользовавшись тем фактом, что это nый коэффициент в ряду Тейлора у функции (1.4)), попробуем узнать вид экспоненциальной производящей функции, коэффициенты ряда Тейлора которого – суммарное число компонент всех графов соответствующие всем эндоморфизмам над множеством мощности n.

Если мы знаем кол-во графов с одной компонентой, двумя, тремя и т.д., то суммарное число компонент, это сумма произведений числа графов с k компонентами умноженными на k, где $k = 1, ..., n$.

т.к. эти числа лежат в коэффициентах функций вида (1.4), то проделаем несколько манипуляций с этими функциями:

$$\sum_{k=1}^{n} k(-1)^k \frac{1}{k!} \ln^k(1 + W(-x)) = \sum_{k=1}^{n} \frac{1}{(k-1)!}(-\ln(1 + W(-x)))^k$$

Заметим, что свободный коэффициент в (1.4) всегда равен 0 (число отображений в графе с $n = 0$ вершинами на $k \neq 0$ компонентами равно 0). Поэтому с увеличением мощности алфавита и при произведении расчетов для эндоморфизмов над алфавитом мощности n, первые $k < n$ компонент будут такими же как и для эндоморфизмов над алфавитом мощности k т.к. в нашей формуле появится только $-\ln(1+W(-x))$ в степени n, а значит первые $k < n$ коэффициентов не изменятся.

Таким образом мы можем сделать переход до $n = \inf$ и взять оттуда нужные нам коэффициенты.

$$\sum_{k=1}^{\infty} \frac{1}{(k-1)!} p^k = e^p * p$$

Делая обратно подстановку $p = -\ln(1 + W(-x))$ Получим:

$$\frac{-\ln(1 + W(-x))}{(1 + W(-x))}$$

■

Предложение 1.2. Для того, чтобы узнать среднее количество компонент графов функций над алфавитом мощности n нужно взять nый член ряда Тейлора полученной функции и поделить его на общее число функций (n^n). Как уже отмечалось выше, сам nый коэффициент означает суммарное число компонент графа от всех отображений над алфавитом мощности n.

1.2 Случай биективного случайного отображения

Формула, полученная выше, решает задачу для случая любой функции (не обязательно биективной). В связи с тем, что автор находит весьма красивым формулу для суммарного числа компонент графов биективных отображений, мы рассмотрим этот случай отдельно от функций общего вида.

В первую очередь заметим, что если мы будем рассматривать случай биективной функции, то графы, которые будут образовываться представят собой циклы.

Выведем формулу для подсчета кол-ва способов разбиения множества k на различные непересекающиеся циклы. Данная задача уже будет выглядеть довольно простой в сравнении с предыдущей задачей, связанной с любой функцией.

Рассмотрим случай, когда мы разбили наше множество на k циклов длины $l_1, l_2, ..., l_k$.

Предложение 1.3. Можно заметить что данные числа являются числами Стирлинга первого рода без знака и для них существует простая рекуррентная формула для подсчета.

Эта "особенность"не является случайной. Если внимательно посмотреть на определение чисел Стирлинга первого рода, то можно понять, что это именно те числа, которые мы и хотели получить изначально. А именно это количество перестановок из n элементов с k циклами.

Теорема 1.4. *Для вычисления чисел Стирлинга первого рода существует рекуррентное соотношение*

$$S(n,k) = S(n-1,k-1) + (n-1)S(n-1,k)$$
$$S(0,0) = 1$$
$$S(n,0) = 0, \ \textit{при } n > 0$$
$$S(0,k) = 0, \ \textit{при } k > 0$$

Доказательство. Смысл этого рекуррентного соотношения следующий. Для разбиения множества из n элементов на k циклов, можно взять уже готовое разбиение множества с $n-1$ элементами на $k-1$ циклов и добавить новый элемент в качестве нового цикл или же можно было взять разбиение множества с $n-1$ элементами на k циклов и добавить новый элемент в один из готовых циклов ($n-1$ способ). ■

Теперь приступим к подсчету среднего количества компонент графа. Обозначим через $A(n)$ это количество для языка мощности n при биективном отображении

$$A(n) = \frac{1}{n!}(1 * S(n,1) + 2 * S(n,2) + ... + n * S(n,n)) =$$

$$\frac{1}{n!}(1*(S(n-1,1-1)+(n-1)*S(n-1,1))+2*((S(n-1,2-1)+(n-1)*S(n-1,2)))+...+n*((S(n-1,n-1)+(n-1)*S(n-1,n)))) =$$

$$\frac{1}{n!}(1 * ((n-1) * S(n-1,1)) + 2 * ((S(n-1,2-1) + (n-1) * S(n-1,2))) + ... + n * ((S(n-1,n-1)))) =$$

$$\frac{1}{n!}((n-1)*(1*S(n-1,1)+2*S(n-1,2)+...+(n-1)*S(n-1,n-1))+2*S(n-1,1)+3*S(n-1,2)+...+n*S(n-1,n-1)) =$$

$$\frac{1}{n!}(n*(1*S(n-1,1)+2*S(n-1,2)+...+(n-1)*S(n-1,n-1))+S(n-1,1)+S(n-1,2)+...+S(n-1,n-1))=$$

$$\frac{1}{n!}(n*A(n-1)+(n-1)!)=$$

Очевидно, что $A(1)=1$. Теперь от рекурсивной функции перейдем к непосредственной

$$A(n)=\frac{1}{n!}(n*A(n-1)+(n-1)!)=\frac{1}{n!}(n*((n-1)A(n-2)+(n-2)!)+(n-1)!)=$$
$$\frac{1}{n!}(n*((n-1)((n-2)A(n-3)+(n-3)!)+(n-2)!)+(n-1)!)=$$
$$\frac{1}{n!}(n*((n-1)((n-2)A(n-3)+(n-3)!)+(n-2)!)+(n-1)!)=$$
$$\frac{1}{n!}(n*((n-1)((n-2)(...3(2*1+1!)+2!)...+(n-3)!)+(n-2)!)+(n-1)!)=$$
$$\frac{1}{n!}(n!+\frac{n!}{2}+\frac{n!}{3}+...+\frac{n!}{n-1}+\frac{n!}{n})=$$
$$1+\frac{1}{2}+\frac{1}{3}+...+\frac{1}{n-1}+\frac{1}{n}=\sum_{k=1}^{n}\frac{1}{k}$$

Теорема 1.5. *Среднее количество компонент графа, относящееся к биективному отображению над алфавитом мощности n на себя, является nым гармоническим числом.*

Предложение 1.4. Этот факт показывает например, что с увеличением мощности алфавита число компонент стремится к бесконечности (из свойств гармонических чисел).

1.3 Анализ полученных формул

Для того, чтобы начать анализ полученных формул давайте найдем непосредственно формулу для количества компонент графа случайного отображения. Для этого рассмотрим экспоненциальную функцию, которую мы получили.

$$\frac{-\ln(1+W(-x))}{1+W(-x)}$$

посмотрим, какой коэффициент будет при $x^n/n!$

$$\frac{-\ln(1+W(-x))}{1+W(-x)}=(-W(-x)+\frac{W^2(-x)}{2}-\frac{W^3(-x)}{3}+\frac{W^4(-x)}{4}-...)(1-W(-x)+W^2(-x)-W^3(-x)+W^4(-x)-...)=$$

$$\sum_{k=1}^{\infty}H(k)(-W(-x))^k$$

, где $H(n)$ – nое гармоническое число.

Т.к. мы знаем краткую формулу ряда Тейлора для $(-W(-x))^k$

$$\sum_{k=1}^{\infty}H(k)\sum_{n=k}^{\infty}\frac{kn^{n-k-1}n!}{(n-k)!}\frac{x^n}{n!}$$

и коэффициент при x^n появляется только если $k = 1, ..., n$, то получим итоговую формулу для коэффициента при $x^n/n!$

$$\sum_{k=1}^{n} H(k) \frac{kn^{n-k-1}n!}{(n-k)!}$$

Отсюда получаем

Теорема 1.6. *Среднее число компонент в графе случайного отображения будет задаваться формулой*

$$\sum_{k=1}^{n} H(k) \frac{kn^{-k-1}n!}{(n-k)!}$$

Данная формула еще больше повышает производительность вычислений.

Предложение 1.5. В вышеуказанной онлайн энциклопедии имеется формула, показывающая к какой функции асимптотически сходится данная сумма [3], а именно

$$\frac{(\ln(2n) + \gamma)}{2}$$

, где γ – константа Эйлера

Данный результат хорошо приближает вычисления. Так, например, на графике ниже представлена зависимость значения среднего числа компонент от мощности языка,

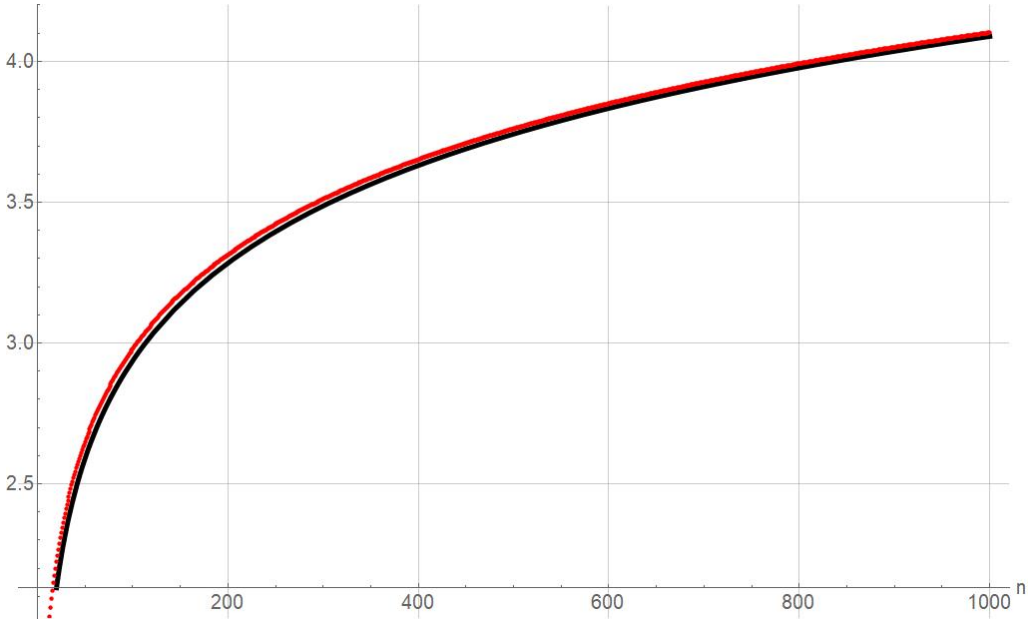

Рис. 1.2: График зависимости среднего числа компонент графа отображения от мощности языка

Как видно из этого графика, данная функция действительно хорошо приближает значение среднего числа компонент, так для мощности языка 10000 погрешность в вычислениях меньше 10^{-2}

Про асимптотику числа компонент нет смысла говорить в связи с тем, что гармонический ряд аппроксимируется данной функцией:

$$\ln n + \gamma$$

и это является общеизвестным фактом.

Оценка числа компонент графа итерации случайного отображения

2.1 Основная часть

Перейдем теперь к сути исходной задачи, которая стояла перед нами, а именно как будет изменяться суммарное число компонент графа при итерации случайного отображения.

В первую очередь заметим, что при итерации отображения мы будем перепрыгивать через несколько ребер нашего графа, причем в результате данных "прыжков"циклы могут распадаться и собираться вместе. Из этого свойства следует один из важных фактов, который позволит вычислить при какой степени суммарное число компонент графа будет максимальным. Давайте рассмотрим при каком условии наши компоненты содержащие циклы различной длины распадутся на компоненты с циклами длиной 1. Такое возможно если в цикле любой длины мы будем возвращаться обратно в свою же вершину. Для этого нужно, чтобы степень отображения делилась на длину любого цикла в компонентах, то есть была наименьшим общим кратным для всех чисел, начиная с единицы и заканчивая длиной самого максимального цикла (то есть мощности алфавита). То есть, например, для 5 это число будет $\mathrm{lcm}(1,2,3,4,5) = 60$.

Запишем этот факт в качестве предложения

Предложение 2.1. Суммарное чило образовавшихся компонент графа в результате итерации случайного отображения над алфавитом мощности n максимально в том случае, когда итерация происходит $i\,\mathrm{lcm}(1,2,...,n)$, где $i = 1,2,...$

Данный факт мы оставим на рассмотрение в следующем разделе, а теперь перейдем к основной задаче. Начнем с нахождения количества компонент,на которые распадется одна компонента графа при возведении случайного отображения в степень r. Имеет место теорема

Теорема 2.1. *Если компонента с циклом длины k, то число компонент, на которое распадется данная компонента при возведении отображения в степень r будет* $\gcd(r,k)$.

Доказательство. Пусть n – минимальное число вершин, которые мы пройдем до попадания в исходную, которая стоит на позиции p. Тогда, должно выполниться равенство

$$p = p + rn \mod k$$

Поскольку n является минимальным, то оно должно равняться $k/\gcd(r,k)$. Действительно, пусть z – это минимальное число, при котором выполняется данное равенство. Заметим, что наше условие эквивалентно условию

$rz = 0 \mod k$, а значит в силу свойств сравнения по модулю должно выполняться и то, что

$\alpha_1 z = 0 \mod \alpha_2$, где $\alpha_1 = \frac{r}{\gcd(r,k)}$, а $\alpha_2 = \frac{k}{\gcd(r,k)}$.Потому как α_1 и α_2 взаимопросты, то данное равенство выполняется только при $z = \alpha_2$, что мы и хотели показать. А следовательно и число циклов, на которое распадется компонента будет $\gcd(r,k)$. Т.к. при прохождении всех вершин, начиная с какой-либо, мы разбиваем наш цикл на непересекающиеся циклы длины $\frac{k}{\gcd(r,k)}$, скольких у нас $\gcd(r,k)$ ■

На "побочные пути"до цикла можно не обращать внимания т.к. они не смогут создать новый цикл (в силу их ацикличности) и соединятся с одним из образовавшихся циклов. Так же стоит заметить, что в результате итерации отображения, две компоненты слиться не смогут в силу того, что одна компонента не имеет общих вершин с другой, из вершины лежащей на одной компоненте, за любое число шагов не возможен переход до вершины другой компоненты.

Теперь, когда мы знаем на сколько компонент распадется наша,обладающая циклом длины k, мы можем найти производящую функцию для суммарного числа компонент при возведении отображений в степень, причем отображений, имеющих одну компоненту.

Предложение 2.2. Как мы помним из прошлой главы, функция (1.2) является экспоненциальной производящей для числа связанных графов, которые обладают циклом длины k. Значит экспоненциальная производящая функция для суммарного числа компонент, которые образуются при возведении отображения в степень, граф которых обладает циклом длины k, будет:

$$\gcd(r,k)(-1)^k\frac{W^k(-x)}{k}$$

Докажем свойство наибольшего общего делителя, которое понадобится нам в дальнейшем.

Лемма 1. $\gcd(m,n)=\gcd(m+i*n,n)$ *при любых целых* i

Доказательство. Пусть $d=\gcd(m,n)$, тогда$da_1=m, da_2=n$, используя факт из теории чисел, заключающийся в том, что существуют такие числа b_1, b_2, что

$$m*b_1+n*b_2=d$$

причем d – минимальное положительное значение функции стоящей слева при целых b_1 и b_2

представим b_2 как ib_1+b_2', тогда:

$$b_1*(m+in)+b_2'*n=d$$

, в силу того, что меньше значение у этой функции получиться не может (иначе бы получилось меньше значение у функции выше), то это и есть наибольший общий делитель чисел $m+in$ и n

■

Зная данный факт, мы можем сказать, что число компонент, на которое распадется компонента с циклом длины k, совпадает с числом компонент, на которое распадется компонента с циклом длины $k+ir$. Теперь, можно приступить к подсчету суммарного числа компонент графов, которые имеют одну компоненту, в результате его распада при возведении в степень.

Т.к. число компонент, которые имеют циклы длины k равно $(-1)^k\frac{W^k(-x)}{k}$, то ряд экспоненциальной производящей функции для тех чисел, которые мы хотим найти:

$$-\gcd(r,1)W(-x)+\gcd(r,2)\frac{W^2(-x)}{2}-\gcd(3,k)\frac{W^3(-x)}{3}+...+$$
$$+\gcd(r,r-1)(-1)^{r-1}\frac{W^{r-1}(-x)}{r-1}+\gcd(r,r)(-1)^r\frac{W^r(-x)}{r}+\gcd(r,1)(-1)^{r+1}\frac{W^{r+1}(-x)}{r+1}+...+$$
$$+\gcd(r,r-1)(-1)^{2r-1}\frac{W^{2r-1}(-x)}{2r-1}+\gcd(r,r)(-1)^{2r}\frac{W^{2r}(-x)}{2r}+\gcd(r,1)(-1)^{2r+1}\frac{W^{2r+1}(-x)}{2r+1}+...+$$
$$...$$
$$=$$

$$-\gcd(r,1)(W(-x)+(-1)^r\frac{W^{r+1}(-x)}{r+1}+(-1)^{2r}\frac{W^{2r+1}(-x)}{2r+1}+...)+$$
$$+\gcd(r,2)(\frac{W^2(-x)}{2}+(-1)^r\frac{W^{r+2}(-x)}{r+2}+(-1)^{2r+2}\frac{W^{2r+2}(-x)}{2r+2}+...)+...+$$
$$+(-1)^r\gcd(r,r)(\frac{W^r(-x)}{r}+(-1)^r\frac{W^{2r}(-x)}{2r}+(-1)^{3r}\frac{W^{3r}(-x)}{3r}+...)$$

Посмотрим, чему равен ряд находящийся внутри скобок. Он имеет вид:

$$\sum_{n=0}^{\infty}(-1)^{nr}\frac{x^{nr+k}}{nr+k}$$

Для того, чтобы узнать, чему равен этот ряд рассмотрим гипергеометрическую функцию ${}_2F_1(a,b,c,x)$ [4], ряд которой имеет вид :

$${}_2F_1(a,b,c,x)=\sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{(c)_n}\frac{x^n}{n!}$$

, где $(x)_n$ символ Похгаммера

$$(x)_n=\prod_{k=1}^{n}(x+k-1)=x(x+1)(x+2)...(x+n-1)$$

Лемма 2. *Рассмотрим ряд функции* ${}_2F_1(1,\frac{k}{r},\frac{k}{r}+1,(-x)^r)\frac{x^k}{k}$. *Ее ряд будет иметь вид*

$$\sum_{n=0}^{\infty}(-1)^{nr}\frac{x^{nr+k}}{nr+k}$$

Доказательство.

$${}_2F_1(1,\frac{k}{r},\frac{k+r}{r},(-x)^r)=\sum_{n=0}^{\infty}\frac{n!(\frac{k}{r})(\frac{k}{r}+1)...(\frac{k}{r}+n-1)}{(\frac{k}{r}+1)...(\frac{k}{r}+n)}\frac{(-1)^{nr}x^{nr}}{n!}=\sum_{n=0}^{\infty}\frac{\frac{k}{r}}{\frac{k}{r}+n}(-1)^{nr}x^{nr}=\sum_{n=0}^{\infty}\frac{k}{nr+k}(-1)^{nr}x^{nr}$$

Таким образом, домножив получившийся ряд на $\frac{x^k}{k}$, мы получим искомый ряд. ■

Теорема 2.2. *Экспоненциальная производящая функция для суммарного числа компонент, образующихся в результате итерации отображения, обладающих одной компонентой, будет*

$$\sum_{k=1}^{r}(-W(-x))^k\gcd(r,k)\frac{{}_2F_1(1,\frac{k}{r},\frac{k}{r}+1,(-(W(-x)))^r)}{k} \tag{2.1}$$

Обозначим данную функцию как $F(r,x)$.

Нетрудно заметить, что формула, полученная ниже, для одной степени $r=1$ является частным случаем полученной формулы.

Теперь осталось узнать производящую функцию суммарного числа компонент образующихся в результате возведения всех отображений в степень r. Отображение состоит из нескольких компонент (обозначим это число как a) и каждая компонента состоящая из k вершин образует в результате число компонент равное kому коэффициенту $F(r,x)$ (обозначим это число как b). Тогда число раз, в которое увеличилось число компонент равняется $\frac{b}{a}$

Теперь, пусть отображение состоит из компонент мощности $n_1, n_2, ..., n_k$ (некоторые из них могут быть равны 0) и существует всего a_{n_i} компонент над алфавитом мощности n_i, где $i = 1, 2..., k$. Тогда первая компонента увеличится в $\frac{b_{n_1}}{a_{n_1}}$ раз, вторая в $\frac{b_{n_2}}{a_{n_2}}$ раз и т.д. kая компонента увеличится в $\frac{b_{n_k}}{a_{n_k}}$ раз. Т.к. число всех таких отображений будет $C_n^{n_1,n_2,...,n_k} a_{n_1} a_{n_2} ... a_{n_k}$, то суммарное число компонент, которое образуется в результате определенных $a_{n_1}, a_{n_2}, ..., a_{n_k}$:

$$C_n^{n_1,n_2,...,n_k}(\frac{b_{n_1}}{a_{n_1}} + \frac{b_{n_2}}{a_{n_2}} + ... + \frac{b_{n_k}}{a_{n_k}})a_{n_1} a_{n_2} ... a_{n_k}$$

, что равносильно

$$C_n^{n_1,n_2,...,n_k} \sum_{i=1}^{k} b_{n_i} a_{n_1} a_{n_2} ... a_{n_{i-1}} a_{n_{i+1}} ... a_{n_k}$$

Просуммируем по всевозможным $n_1, n_2, ..., n_k$ и поделим на число повторений $k!$

$$\frac{1}{k!} \sum_{n_1+...+n_k=n} C_n^{n_1,n_2,...,n_k} \sum_{i=1}^{k} b_{n_i} a_{n_1} a_{n_2} ... a_{n_{i-1}} a_{n_{i+1}} ... a_{n_k}$$

Т.к. a_i это iый коэффициент у $F(1, x)$, а b_i это i коэффициент у $F(r, x)$, где $i = 1, 2, ..., k$, то нетрудно убедиться, что производящая функция для отображения с k числом компонент:

$$\frac{1}{(k-1)!} F(r, x) F^{k-1}(1, x)$$

Так будет получаться в силу свойств перемножения рядов и при коэффициенте x^n будет именно нужный нам коэффициент. Действительно:

$$\frac{1}{(k-1)!} F(r, x) F^{k-1}(1, x) = \frac{1}{(k-1)!} \sum_{n_1=0}^{\infty} \frac{b_{n_1}}{n_1!} x^{n_1} \sum_{n-n_1=0}^{\infty} \left(\sum_{n_2+...+n_k=n-n_1} C_{n-n_1}^{n_2,...,n_k} a_{n_2} ... a_{n_k} \right) \frac{x^{n-n_1}}{(n-n_1)!} =$$

$$\frac{1}{(k-1)!} \sum_{n=0}^{\infty} \left(\sum_{n_1+n_2+...+n_k=n} C_n^{n_1,n_2,...,n_k} b_{n_1} a_{n_2} ... a_{n_k} \right) \frac{x^n}{n!} =$$

$$\sum_{n=0}^{\infty} \left(\frac{1}{k!} \sum_{n_1+n_2+...+n_k=n} C_n^{n_1,n_2,...,n_k} \sum_{i=1}^{k} b_{n_i} a_{n_1} a_{n_2} ... a_{n_{i-1}} a_{n_{i+1}} ... a_{n_k} \right) \frac{x^n}{n!}$$

Что и требовалось показать.

На финальном шаге нам осталось просуммировать это число при всевозможных k, так же как и в случае с $r = 1$. Производя такое суммирование, получим:

$$\sum_{k=1}^{\infty} \frac{1}{(k-1)!} F(r, x) F^{k-1}(1, x) = F(r, x) e^{F(1,x)}$$

На этом этапе вспоминаем, что

$$F(1, x) = -\ln(1 + W(-x))$$

И таким образом получаем итоговую формулу:

Теорема 2.3. *Экспоненциальная производящая функция суммарного числа компонент, образующаяся при итерации всевозможных случайных отображений из алфавита мощности n на себя:*

$$\frac{F(r,x)}{1+W(-x)}$$

, где F(r,x) – это функция (2.1)

Предложение 2.3. Данная экспоненциальная производящая функция позволяет легко определить суммарное количество компонент, которое образуется в результате итерации случайного отображения. Если же мы хотим определить наиболее вероятное число компонент, которое мы получим в результате применения случайного отображения, то надо поделить это число на число отображений (n^n).

Как бы странно это не звучало, формула, полученная в первой главе, является всего лишь частным случаем полученной в этой главе формулы. Более того, найдя экспоненциальную производящую функцию для наших чисел, мы опять-таки облегчили программирование нашей формулы (в пакете Wolfram Mathematica расчеты производятся в 3 строчки. При желании (и потери читабельности кода) данную формулу можно записать и в одну строчку), увеличили производительность при расчете чисел. Так, для сравнения, при поиске данных чисел "в лоб"(т.е. перебором всех отображений) поиск числа компонент затягивался на несколько минут для $r = 2$ и мощности алфавита $n = 6$, в то время как при использовании производящей функции мы могли работать с алфавитом мощности $n = 600$ и степенью $r = 2$, при этом на все расчеты уходило примерно 45 секунд!

2.2 Оценка числа максимального числа компонент при итерации случайного отображения

В данном разделе давайте попробуем найти суммарное максимальное число компонент, которое будет получаться при итерации случайного отображения, то есть когда степень является $\mathrm{lcm}(1, 2, ..., n)$ для отображений над алфавитом мощности n.

Предложение 2.4. Насколько мы знаем из начала первого раздела данной главы, число компонент на которое распадется наша компонента с циклом длины k будет $\gcd(r, k)$(см. Теорема 2.1), а значит максимально возможное число будет k. Основываясь на этом факте и зная экспоненциальную производящую функцию для числа компонент с циклом длины k (то есть функции (1.2)), мы можем найти производящую функцию для суммарного максимально возможного числа компонент, на которое распадется компонента графа построенного на n вершинах. А именно:

$$-W(-x) + W^2(-x) - W^3(-x) + ... = \frac{1}{1+W(-x)} - 1 = \frac{-W(-x)}{1+W(-x)}$$

Зная данный факт мы можем с легкостью найти суммарное максимальное число компонент графа отображения, которые образуются в результате итерации отображения, граф которого имеет k компонент, нужное количество раз. А именно, по аналогии с прошлыми разделами:

Теорема 2.4. *Экспоненциальная производящая функция суммарного максимально возможного числа компонент, образующихся в результате итерации случайного эндоморфизма, граф которого обладает k компонентами, над алфавитом мощности n будет иметь вид:*

$$\frac{1}{(k-1)!}\frac{-W(-x)}{1+W(-x)}(-\ln(1+W(-x)))^{k-1}$$

В целях экономии места в данной работе оставим доказательства этого факта в качестве упражнения на усвоение материала читателю

Теперь так же, как и в остальных случаях просуммируем при всевозможных k данные функции. Получим:

Теорема 2.5. *Экспоненциальная производящая функция суммарного максимально возможного числа компонент, образующихся в результате итерации случайного отображения над алфавитом мощности n на себя будет иметь вид:*

$$\sum_{k=1}^{\infty} \frac{1}{(k-1)!} \frac{-W(-x)}{1+W(-x)} (-\ln(1+W(-x)))^{k-1} = \frac{-W(-x)}{(1+W(-x))^2}$$

Данный результат мы более подробно рассмотрим в последующих разделах, когда будем находить асимптотическую функцию коэффициентов данной экспоненциальной производящей функции и явную формулу данных чисел.

2.3 Оценка числа компонент итерации биективного случайного отображения

Начнем поиск экспоненциальной производящей функции для данной задачи начиная с тех же соображений, которые мы использовали в начале первого раздела (т.е. мы будем отталкиваться от того, что мы знаем, на сколько компонент распадется компонента с циклом длины k).

Приступим сразу к подсчету числа компонент, которые образуются в результате итерации r раз биективного случайного отображения с одной компонентой в графе

$$\gcd(r,1)\frac{0!}{1!}x + \gcd(r,2)\frac{1!}{2!}x^2 + ... + \gcd(r,r-1)\frac{(r-2)!}{(r-1)!}x^{r-1} + \gcd(r,r)\frac{(r-1)!}{r!}x^r + \gcd(r,1)\frac{r!}{(r+1)!}x^{r+1} + ... =$$

$$\gcd(r,1)(\frac{0!}{1!}x + \frac{r!}{(r+1)!}x^{r+1} + ...) + ... + \gcd(r,r)(\frac{(r-1)!}{r!}x^r + \frac{(2r-1)!}{(2r)!}x^{2r} + ...) =$$

$$\gcd(r,1)(\frac{1}{1}x + \frac{1}{r+1}x^{r+1} + ...) + ... + \gcd(r,r)(\frac{1}{r}x^r + \frac{1}{2r}x^{2r} + ...) =$$

Если мы узнаем, какая функция имеет ряд вида

$$\sum_{n=0}^{\infty} \frac{x^{nr+k}}{nr+k}$$

, то это позволит нам записать нашу функцию более компактно. К счастью, данная функция существует, называется дзета-функцией Гурвица и обозначается как $\Phi(x,s,\alpha)$ [5]. Ее ряд имеет вид:

$$\Phi(x,s,\alpha) = \sum_{n=0}^{\infty} \frac{x^n}{(n+\alpha)^s}$$

Нетрудно проверить, что в нашем случае мы имеем дело с функцией $x^k\Phi(x^r,1,\frac{k}{r})/r$

Данный факт позволяет нам записать нашу экспоненциальную производящую функцию более компактно:

Теорема 2.6. *Экспоненциальная производящая функция суммарного числа компонент, образующихся в результате итерации случайного биективных отображений, граф которых имеет ровно один цикл, над алфавитом мощности n в себя будет иметь вид:*

$$\sum_{k=1}^{r} \gcd(r,k) \frac{x^k \Phi(x^r,1,\frac{k}{r})}{r}$$

Обозначим данную функцию как $F_b(r,x)$.

Т.к. производящая функция для числа биективных отображений с одной компонентой в графе это $-\ln(1-x)$ (проверьте это!), то производящая функция для числа компонент, которые будут образовываться в результате итерации отображения, граф которого имеет k компонент:

Теорема 2.7. *Экспоненциальная производящая функция суммарного числа компонент, образующихся в результате итерации случайных биективных отображений, граф которых имеет ровно* k *циклов, над алфавитом мощности* n *в себя будет иметь вид:*

$$\frac{1}{(k-1)!}F_b(z,x)(-\ln(1-x))^{k-1}$$

Проверьте данный факт, для закрепления материала!

Теперь опять-таки все суммируем по всевозможным k и получаем экспоненциальную производящую функцию для суммарного числа компонент:

Теорема 2.8. *Экспоненциальная производящая функция суммарного числа компонент, образующихся в результате итерации случайных биективных отображений, над алфавитом мощности* n *в себя будет иметь вид:*

$$\sum_{k=1}^{\infty}\frac{1}{(k-1)!}F_b(z,x)(-\ln(1-x))^{k-1}=\frac{F_b(z,x)}{1-x}$$

Мы опять-таки получили более общую формулу для подсчета числа компонент чем найденную в первой главе.

Предложение 2.5. Если мы хотим узнать вероятное число образовавшихся компонент в случае итерации случайно взятого отображения, то надо поделить число, которое мы получим, на общее число биективных отображений ($n!$)

Раздел, посвященный подсчету максимального числа компонент, которые могут образовываться в результате итерации случайного биективного отображения будет бессмысленным, т.к. число получившихся компонент очевидно – n, для алфавита мощности n. Связано это с тем, что все вершины находятся в циклах (в силу биективности) и все циклы распадаются на циклы длины 1.

2.4 Анализ полученных формул

2.4.1 Анализ основной формулы

Для начала попробуем оценить основную формулу этой главы. Для этого, так же как и в первой главе, начнем с вывода явной формулы для искомых чисел

$$\frac{F(r,x)}{1+W(-x)}=(-\gcd(r,1)W(-x)+\gcd(r,2)\frac{W^2(-x)}{2}-\gcd(3,k)\frac{W^3(-x)}{3}+...+$$
$$+\gcd(r,r-1)(-1)^{r-1}\frac{W^{r-1}(-x)}{r-1}+\gcd(r,r)(-1)^r\frac{W^r(-x)}{r}+\gcd(r,1)(-1)^{r+1}\frac{W^{r+1}(-x)}{r+1}+...+$$
$$+\gcd(r,r-1)(-1)^{2r-1}\frac{W^{2r-1}(-x)}{2r-1}+\gcd(r,r)(-1)^{2r}\frac{W^{2r}(-x)}{2r}+\gcd(r,1)(-1)^{2r+1}\frac{W^{2r+1}(-x)}{2r+1}+...+)$$
$$(1-W(-x)+W^2(-x)-W^3(-x)+...)=$$
$$=\sum_{k=1}^{\infty}(-W(-x))^k(gcd(r,1)+\frac{gcd(r,2)}{2}+\frac{gcd(r,3)}{3}+...+\frac{gcd(r,r)}{r}+\frac{gcd(r,1)}{r+1}+...+\frac{gcd(r,1+(k-1 \mod r))}{k})$$

Для простоты расчетов можно считать, что мы работаем с языком, мощность которого кратна r, т.к все равно средняя величина компоненты не будет сильно отличаться от этого значения. Это верно в силу того, что гармонический ряд очень медленно расходится при достаточно больших n

$$= \sum_{k=1}^{\infty}(-W(-x))^k(gcd(r,1) + \frac{gcd(r,2)}{2} + \frac{gcd(r,3)}{3} + ... + \frac{gcd(r,r)}{r} + \frac{gcd(r,1)}{r+1} + ... + \frac{gcd(r,r)}{k}) =$$

$$= \sum_{k=1}^{\infty}(-W(-x))^k(\sum_{p=1}^{r}\gcd(r,p)\sum_{z=0}^{k/r-1}\frac{1}{zr+p})$$

Осталось только разложить функцию $(-W(-x))^k$ в ряд Тейлора и собрать коэффициенты при $x^n/n!$

$$= \sum_{k=1}^{\infty}(\sum_{p=1}^{r}\gcd(r,p)\sum_{n=0}^{k/r-1}\frac{1}{nr+p})\sum_{n=k}^{\infty}\frac{kn^{n-k-1}n!}{(n-k)!}\frac{x^n}{n!}$$

Таким образом коэффициент при $x^n/n!$ будет

$$\sum_{k=1}^{\infty}(\sum_{p=1}^{r}\gcd(r,p)\sum_{z=0}^{k/r-1}\frac{1}{zr+p})\frac{kn^{n-k-1}n!}{(n-k)!}$$

Чтобы формулу можно было применить для любого значения n перепишем ее следующим образом

Теорема 2.9. *Суммарное число компонент графа, образующееся в результате итерации всех отображений алфавита мощности n на себя:*

$$\sum_{k=1}^{\infty}(\sum_{p=1}^{k}\frac{\gcd(r,p)}{p})\frac{kn^{n-k-1}n!}{(n-k)!}$$

Данная формула опять-таки повышает производительность в вычислениях искомых чисел. Заметим, что внутренняя сумма не превосходит гармонического числа, умноженного на r. Данная оценка позволяет дать асимптотическую функцию, которую не превосходит данная последовательность при $r > 1$:

Теорема 2.10. *Асимптотическая функция, которую не превосходит наша последовательность при $r > 1$ будет*

$$r\frac{(\ln(2n) + \gamma)}{2}$$

В дальнейшей работе мы попробуем улучшить эту оценку и попытаемся получить точную асимптотическую функцию для данной последовательности

2.4.2 Анализ второго раздела

Теперь давайте рассмотрим числа соответствующие максимально возможному числу компонент при должном числе итерации (то что рассматривали во втором разделе)

$$\frac{-W(-x)}{(1+W(-x))^2} = -W(-x)(1 - W(-x) + W^2(-x) - W^3(-x) + ...)(1 - W(-x) + W^2(-x) - W^3(-x) + ...) =$$

$$\sum_{k=0}^{\infty}(k+1)(-W(-x))^{k+1} = \sum_{k=0}^{\infty}(k+1)\sum_{n=k+1}^{\infty}\frac{(k+1)n^{n-k-2}n!}{(n-k-1)!}\frac{x^n}{n!}$$

Значит коэффициент при $x^n/n!$ будет

$$\sum_{k=0}^{n-1}(k+1)^2\frac{n^{n-k-2}n!}{(n-k-1)!}$$

Теорема 2.11. *Среднее максимальное число компонент образующихся в результате итерации случайного отображения над алфавитом мощности n можно подсчитать по формуле*

$$\sum_{k=0}^{n-1}(k+1)^2\frac{n^{-k-2}n!}{(n-k-1)!}$$

Найдем асимптотическую функцию, которая приближает данную последовательность. Для этого попробуем вынести из суммы $\sqrt{n}$ и покажем, что последовательность будет стремится к 5/4. Данное предположение было выдвинуто в связи с тем фактом, что график значений первых чисел очень напоминает график функции квадратного корня. На рисунке ниже показано его значение

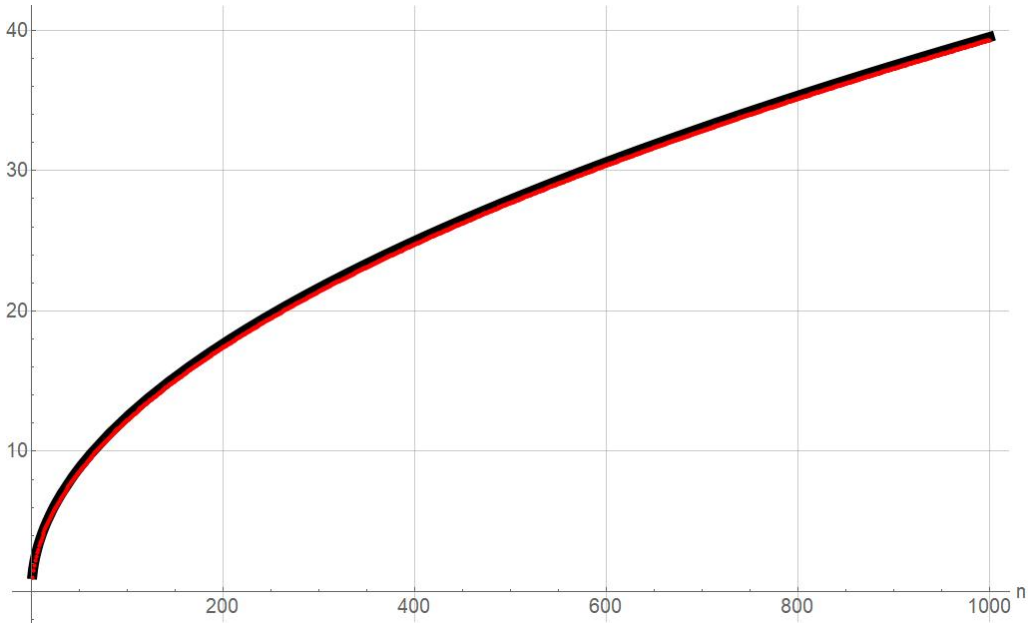

Рис. 2.1: График зависимости среднего максимального числа компонент графа отображения от мощности языка

Если вынести из под знака суммы $\sqrt{n}$, то можно заметить, что функция стремится к некоторой константе не превосходящей 1.25 (вывод делался при мощностях алфавита меньших 1000), подсчитав сумму при $n = 10000$, мы получим число $n = 1.2499$.Отсюда следует что данная сумма должна сходиться к данной константе. График 2.1 показывает это.

Теорема 2.12. *Таким образом функция, которая асимптотически приближается к данным числам, имеет вид*

$$\frac{5}{4}\sqrt{x}$$

2.4.3 Анализ третьего раздела

Как всегда начнем с вывода явной формулы чисел

$$\frac{F_b(z,x)}{1-x} = (\gcd(r,1)x + \gcd(r,2)\frac{x^2}{2} + ... + \gcd(r,r)\frac{x^r}{r} + \gcd(r,1)\frac{x^{r+1}}{r+1} + ...)(1 + x + x^2 + x^3 + ...) =$$

Для простоты расчетов можно считать, что мы работаем с языком мощности кратной r

$$\sum_{k=1}^{\infty}(\sum_{p=1}^{r}\gcd(r,p)\sum_{z=0}^{k/r-1}\frac{1}{zr+p})x^k$$

Таким образом коэффициент при $x^k/k!$ будет

$$n!\sum_{p=1}^{r}\gcd(r,p)\sum_{z=0}^{k/r-1}\frac{1}{zr+p}$$

Формулу можно записать проще и при этом она будет работать при любых k

Теорема 2.13. *Суммарное число компонент графа образующихся в результате итерации всех биективных отображений алфавита мощности n в себя*

$$k!\sum_{p=1}^{k}\frac{\gcd(r,p)}{p}$$

Из этой формулы можно сразу дать грубую верхнюю оценку для числа компонент, и т.к. $\gcd(r,p)\leq r$, то имеет место теорема

Теорема 2.14. *полученное число не превосходит значения функции:*

$$r(\ln(n)+\gamma)$$

, при достаточно больших n и $r\geq 2$

Вывод

В данной курсовой работе были получены основные формулы, которые помогут читателю в дальнейшем в создании более совершенных алгоритмов шифрования и выработки ключа.

В дальнейшем данная работа будет развиваться, можно получить более совершенные формулы аппроксимирования выведенных в 1 и 3 разделе второй главы числовых последовательностей, а также будет дано более математически корректное обоснование того, что последовательность из второго раздела второй главы асимптотически сходится к функции $5/4\sqrt{x}$.

Литература

[1] Sloane John. The on-line encyclopedia of integer sequences. Number of connected functions on n labeled nodes. https://oeis.org/A001865.

[2] Wikipedia. Lambert W function. https://en.wikipedia.org/wiki/Lambert_W_function.

[3] Sloane John. The on-line encyclopedia of integer sequences. The number of cycles in the digraph representation of all endofunctions on 1,2,...,n. http://oeis.org/A190314.

[4] Wikipedia. Hypergeometric function. https://en.wikipedia.org/wiki/Hypergeometric_function.

[5] Wikipedia. Lerch zeta function. https://en.wikipedia.org/wiki/Lerch_zeta_function.