

Формула оценки: **0.25 дз + 0.3 контест + 0.15 кр + 0.3 экзамен + Бонус**

Домашние задания: сдавать устно(раз в неделю асистенты устраивают доп пару, запись онлайн) или latex, дедлайн 10-21 день.

Контесты: Длинные(код реview), короткие(раз в 2 недели), неточные, бонусные(идет к бонусу). Штрафов нет.

Контрольные работы: раз в модуль, тестовые вопросы.

Бонусы: бонусные контесты, ACM, работа на семинаре.

Материалы:

- Кормен
- en.wikipedia
- викиконспекты
- e-maxx
- Корте-Фанен Комбинаторная оптимизация

Теория вероятности. $(\Omega, 2^\Omega, P)$ - вероятностная пространство.

$A \subset \Omega$, $P(A) = \sum_{w \in A} P(w)$.

Def: A, B - события, $P(B) > 0$. $P(A|B)$ - вероятность события A , если наступило событие B . Тогда
$$P(A|B) = \frac{\sum_{w \in A \cap B} P(w)}{\sum_{w \in B} P(w)} = \frac{P(A \cap B)}{P(B)}$$
.

Def: A и B независимые, если $P(A|B) = P(A)$.

Тогда, если A и B независимые, то $P(A \cap B) = P(A) \cdot P(B)$.

$\xi : \Omega \rightarrow \mathbb{R}$ - случайная величина. $\xi(w)$ - значение, $w \in \Omega$.

Пример: Есть 5 марок автомобиля, их стоимости и их количества. A - 1000 - 100; B - 2000 - 5; C - 3000 - 5; D - 2000 - 20; E - 1500 - 30; Тогда нас интересуют $P(\xi = 1000) = \frac{100}{160}$, $P(\xi = 1500) = \frac{30}{160}$, $P(\xi = 2000) = \frac{25}{160}$, $P(\xi = 3000) = \frac{5}{160}$.

Матожидание $E(\xi) = \sum_{w \in \Omega} \xi(w) \cdot P(w) = \sum_x x \cdot P(\xi = x)$.

Индикаторная случайная величина: $I_A = \begin{cases} 1, & w \in A \\ 0, & w \notin A \end{cases}$ Тогда $E(I_A) = P(A)$.

Пусть есть 2 случайной величины ξ_1 и ξ_2 . Тогда $E(\alpha\xi_1 + \beta\xi_2) = \alpha E(\xi_1) + \beta E(\xi_2)$.
 $E(\alpha\xi_1 + \beta\xi_2) = \sum_{w \in \Omega} (\alpha\xi_1(w) \cdot P(w) + \beta\xi_2(w) \cdot P(w)) = \alpha \sum_{w \in \Omega} \xi_1(w) \cdot P(w) + \beta \sum_{w \in \Omega} \xi_2(w) \cdot P(w) = \alpha E(\xi_1) + \beta E(\xi_2)$

Две случайные величины называются независимые, если $\forall x, y : P(\xi_1 = x \text{ и } \xi_2 = y) = P(\xi_1 = x) \cdot P(\xi_2 = y)$.
 n случайных величин называются попарно независимыми, если любые 2 величины независимы.
независимы в совокупности - см семинар

ξ_1 и ξ_2 - случайные независимые величины. Тогда $E(\xi_1\xi_2) = E(\xi_1)E(\xi_2)$.
 $E(\xi_1\xi_2) = \sum_{w \in \Omega} \xi_1(w)\xi_2(w)P(w) = \sum_x x \cdot P(\xi_1\xi_2 = x) = \sum_{(u,v)} uv \cdot P(\xi_1 = u \text{ и } \xi_2 = v) = [\xi_1 \text{ и } \xi_2 \text{ независимы}] = \sum_{(u,v)} uv \cdot P(\xi_1 = u) \cdot P(\xi_2 = v) = (\sum_u u \cdot P(\xi_1 = u)) \cdot (\sum_v v \cdot P(\xi_2 = v)) = E(\xi_1)E(\xi_2)$.

Задача о назначениях. Есть n работников и n работ. Есть таблица, где a_{ij} - сколько i -ый работник берет за j -ую работу. Нужно распределить работников по работам так, чтобы суммарная плата за все работы была минимальна. Оценим матожидание затрат при случайном решении. A_{ij} - событие, когда i -ый работник делает j -ую работу. $\xi = \sum_{(i,j)} I_{A_{ij}} \cdot a_{ij}$. Тогда $E(\xi) = \sum_{(i,j)} E(I_{A_{ij}}) = \sum_{(i,j)} a_{ij}P(A_{ij}) = \sum_{(i,j)} a_{ij} \cdot \frac{1}{n}$.

Найти максимальный разрез в неориентированном невзвешанном графе.

Будем строить случайный разрез(каждую вершину либо в A , либо в \bar{A}). Тогда ξ - величина нашего разреза. $\xi = \sum_{e \in E(G)} I_{B_e}$, где B_e - событие, когда e лежит в разрезе. $P(e \in \text{разрез}) = \frac{1}{2}$. Тогда $E(\xi) = E(\sum_{e \in E(G)} I_{B_e}) = \sum E(I) = \frac{1}{2} |E(G)|$.

Есть перестановка $p_1 \dots p_n$. Алгоритм жадно набирает возрастающую подпоследовательность. Какое матожидание длины этой подпоследовательности?
Событие A_i - алгоритм возьмет p_i . $E(\xi) = E(\sum I_{A_i}) = \sum P(A_i)$. $P(A_i) = P(\forall j < i : p_j < p_i) = \frac{1}{i}$. Тогда $E(\xi) = \sum_{i=1}^n \frac{1}{i} = \Theta(\log n)$.

Дисперсия $D(\xi) = \sum_{w \in \Omega} P(w)(\xi(w) - E(\xi))^2$.

Свойства:

- $D(\xi_1 + \xi_2) = D(\xi_1) + D(\xi_2)$, ξ_1 и ξ_2 независимы

- $D(\lambda\xi_1) = \lambda^2 D(\xi_1)$

Неравенство Маркова. $\xi : \Omega \rightarrow \mathbb{R}_+$. $P(\xi(w) \geq E(\xi) \cdot k) \leq \frac{1}{k}$.

Неравенство Чебышева. $\xi : \Omega \rightarrow \mathbb{R}$. $P(|\xi - E(\xi)| \leq \alpha) \leq \frac{D(\xi)}{\alpha^2}$.

Модели

RAM-модель (Random Access Machine) Вопросы, возникающие при создании модели

1. адресация
2. какие инструкции
3. рекурсия
4. где лежат инструкции
5. размер данных
6. кол-во памяти
7. случайность

Адресация Есть ячейки, в которых можно хранить целые числа (ограничения на $MAXC$ разумные, и на них введена неявная адресация)

Замечание. Явная адресация — при создании элемента получаем адрес и можем пользоваться только этим адресом. Неявно — можем получать адреса каким-то своим образом, к примеру, $ptr + 20$.

Кол-во памяти Неявное соглашение RAM — время работы не меньше памяти. По дефолту считаем, что мы его инициализируем мусором

Где инструкции Хранить инструкции можно в памяти и где-то снаружи. Мы будем хранить снаружи (внутри — RASP-модель). Иначе говоря, инструкции и данные отделены.

Какие инструкции В нашей модели есть инструкции следующих типов:

- работа с памятью
- ветвление
- передача управления ($=goto$),
- арифметика (at least $a + b, a - b, \frac{a}{b}, \cdot, mod, \lfloor \frac{a}{b} \rfloor$)
- сравнения (at least $a < b, a > b, a \leq b, a \geq b, a = b, a \neq b$)
- логические (at least $\wedge, \vee, \oplus, \neg$)
- битовые операции ($>>, <<, \&, |, \sim, \oplus$)
- математические функции (опять-таки, в рамках разумного)
- rand

Все инструкции работают от конечного разумного числа operandов (не умеем в векторные операции)

Размер данных $\exists C, k : C \cdot A^k \cdot n^k$ — верхнее ограничение на величины промежуточных вычислений.

Рекурсия Рекурсия всегда линейна по памяти относительно глубины.

Случайность Мы считаем, что у нас есть абсолютно рандомная функция. Будем полагать, что у нас есть источник энтропии, выдающий случайности в промежутке $[0, 1]$.

Время работы.

- наихудшее — $t = \max_{input, random} t(input, random)$
- наилучшее — $t = \max_{input} \min_{random} t(input, random)$
- ожидаемое — $E t = \max_{input} Average_{random} t(input, random)$
- на случайных данных — $t = Average_{input} Average_{random} t(input, random)$

Алгоритмы

Методы доказательства корректности алгоритма.

1. индукция
2. инвариант
3. от противного

Способы оценки времени работы:

- Прямой учет
- Рекурсивная оценка
- Амортизационный анализ

Прямой учет Время работы строчки — произведение верхних оценок по всем строчкам-предкам нашей.

К примеру

```
while (!is_sorted()) { // O(# inversions) = O(n^2)
    for (int i = 0; i + 1 < n; i++) { // O(n) * O(parent) = O(n^3)
        if (a[i] > a[i + 1]) {
            swap(a[i], a[i + 1]); // O(n^3)
        }
    }
}
```

Рекурсивная оценка Пример — сортировка слиянием

Нас интересует две вещи: инвариант и переход. Для оценки времени используем рекурренту вида $T(n) = O(f(n)) + \sum_{n' \in \text{calls}} T(n')$

При этом если мы доказываем время работы, то показываем $T(n) \leq c \cdot f(n)$, зная, что для n' $\exists c : T(n') \leq c \cdot f(n)$

Важно, что c глобальное и не должно увеличиваться в ходе доказательства

stable sort Делает сортировку, не меняя порядок равных элементов относительно исходной последовательности. Merge-sort стабилен.

inplace-algorithm Не требует дополнительной памяти и делает все прямо на данной памяти (у нас есть $\log n$ памяти на рекурсию). Quick-sort inplace.

Время работы qsort

$$T(n) = \max_{\text{input}} \text{average}_{rand} t(\text{input}, rand) = \max_{|\text{input}|=n} Et(\text{input})$$

$$\begin{aligned} T(n) &\leq \Theta(n) + \frac{1}{n} \sum_{k=0}^{n-1} (T(k+1) + T(n-k)) \leq \Theta(n) + \frac{2}{n} \cdot \sum_{k=1}^n T(k-1) \leq a \cdot n + \frac{2}{n} \sum_{k=1}^n c \cdot (k-1) \cdot \log(k-1) \leq \\ &\leq a \cdot n + \frac{2}{n} \sum_{k=1}^{\frac{n}{2}} c \cdot (k-1) \cdot \log n - \frac{2}{n} \sum_{k=1}^{\frac{n}{2}} c \cdot (k-1) + \frac{2}{n} \sum_{k=\frac{n}{2}+1}^n c \cdot (k-1) \cdot \log n \leq \\ &\leq a \cdot n + \frac{2}{n} \cdot n^2 \cdot \log n - \frac{2c}{n} \cdot \frac{(n-2)^2}{4} \leq a \cdot n + cn \log n - \frac{c(n-2)}{4} \leq cn \log n \end{aligned}$$

$$\frac{c(n-2)}{4} \geq a \cdot n$$

$$c \cdot n - 2c \geq 4 \cdot a \cdot n$$

$$c \geq \frac{4 \cdot a \cdot n}{(n-2)}$$

, что верно для достаточно больших n .

Ограничение на число сравнений в сортировке Бинарные сравнения на меньше.

Рассмотрим дерево переходов. Для перестановки есть хотя бы один лист — листьев хотя бы $n!$

$$L(T) \leq 2^x, d(T) \leq x$$

, если x — ответ

$$L(T) \geq n!$$

$$d(T) \geq \log n! \geq \log \frac{n^{\frac{n}{2}}}{2} = \log 2^{(\log \frac{n}{2}) \cdot \frac{n}{2}} = \frac{n}{2} \cdot \log \frac{n}{2} = \Omega(n \log n)$$

1 Сортировки основанные на внутреннем виде данных

Имеем n чисел $[0, U - 1]$, $U = 2^w$, числа укладываются в RAM-модель

Сортировка подсчетом Заводим массив $cnt[U]$, $cnt[x] = |\{i : a_i = x\}|$. Дальше переводим $count \rightarrow pref$, $pref[x] = pref[x - 1] + count[x]$
 $O(n + U)$

Поразрядная сортировка b_{ij} — j -й бит i -го числа. (Сортируем бинарные строки длины w)

Поочередно сортируем строки, разбивая их на классы эквивалентности по $iter$ последним символам. После чего мы стабильно сортируем по $(iter + 1)$ -му символу.

$O(n \log U)$

Bucket sort Разбиваем множество на корзины, каждой корзине соответствует отрезок. В каждой корзине запускаемся рекурсивно.

$O(n \log U)$

В продакшне используют первую пару итераций, чтобы сильно снизить размерность на реальных данных.

Пусть мы хотим отсортировать равновероятные числа из $[0, 1]$. В каждом бакете отсортируем за квадрат. Получим $O(n)$.

$$\begin{aligned} t(n) &\leq \sum_{i=1}^n c \cdot (1 + E(cnt_i)^2) \leq c \cdot n + c \cdot \sum_{i=1}^n E(cnt_i^2) = \\ &= c \cdot n + c \cdot \sum_{i=1}^n \sum_{j=1}^n EI_{A_{ij}} = c \cdot n + c \cdot n^2 \cdot \frac{1}{n} \leq 2 \cdot c \cdot n \end{aligned}$$

2 Иерархия памяти

Нас интересует задержка (latency), пропускная способность (throughput). Подгрузка x данных занимает $l + \frac{x}{t}$

От долгой к быстрой

1. external machine / internet
2. HDD
3. SSD
4. RAM
5. L3
6. L2
7. L1
8. registers

3 Алгоритмы во внешней памяти

n — размер задачи

M — размер RAM

B — блок данных

$B \ll M \ll n$

$\log n \ll B$

$B < \sqrt{M}$ (но это неявно и не факт)

Mergesort во внешней памяти Обычный mergesort, но три типа событий в *merge*:

1. Кончился первый буфер — подгружаем новый
2. Кончился второй — аналогично
3. Кончился буфер для слияния — выписываем обратно в RAM и сбрасываем

$$O\left(\frac{n}{B} \log n\right) \rightarrow O\left(\frac{n}{B} \log \frac{n}{B}\right) \rightarrow O\left(\frac{n}{B} \log_{\frac{M}{B}} \frac{n}{B}\right)$$

+2 идеи:

1. Дошли до размера M — явно посортируем в RAM
2. Можем сливать сразу $\frac{M}{B}$ массивов

1 Простые структуры данных

Требования к структуре данных:

- От СД мы хотим обработку каких-то наших запросов.
- online-offline. Бывает, что мы знаем все запросы, бывает, что мы узнаем запрос только тогда, когда отвечаем на предыдущий
- При обсуждении времени работы отделяется время на препроцессинг и на последующие ответы на запросы (query).

1.1 Data structure / interface

Структура данных — это какой-то математический объект, который умеет отвечать на наши запросы конкретным способом. Красно-черное дерево — это структура данных.

Интерфейс — это объект, с которым может взаимодействовать пользователь, который каким-то образом умеет отвечать на наши запросы (пользователю все равно, как, его волнует только то, что интерфейс реализует, и за какое время(память) он это делает). `std :: set` — это интерфейс.

Итератор — это специальный объект, отвечающий непосредственно за ячейку в структуре данных. Для того, чтобы удалить элемент, мы должны иметь итератор на этот элемент. Т.е. удаление по ключу работает за $O(\text{erase})$, а удаление по значению за $O(\text{find}) + O(\text{erase})$

list Списки бывают двусвязными, односвязными, циклическими. У каждого элемента есть ссылка на следующий (и иногда на предыдущий), а также есть отдельный глобальный указатель на начало списка.

stack Стек — структура данных, которая умеет делать добавление в конец, удаление из конца, взятие последнего элемента, за $O(1)$.

queue Очередь — структура данных, которая умеет делать добавление в конец, удаление из начала, взятие первого элемента, за $O(1)$.

deque Двусторонняя очередь — структура данных, которая умеет делать добавление, удаление и взятие элемента с любого конца последовательности, за $O(1)$.

priority_queue Очередь с приоритетами aka куча — структура данных, которая умеет делать добавление, удаление, и быстрые операции с минимумом, представляющая из себя дерево с условием $\text{parent}(u) = v \rightarrow \text{value}(v) \leq \text{value}(u)$.

Все эти структуры реализуются как на массиве (храним последовательную память и указатели на начало/конец), так и на списках (что на самом деле является тем же самым, что и на массиве, просто ссылка вперед эквивалентна $a_i \rightarrow a_{i+1}$ в терминологии массивов, *прим. автора*).

Структура данных на массиве кратко быстрее аналогичной на ссылках, потому что массив проходится по кэшу и не требует дополнительной памяти.

data structure	add	delete	pop	find	top	build	min	get by index
stack	$O(1)$	-	$O(1)$	-	$O(1)$	$O(n)$	$O(n)$	-
dynamic array	$O(1)$	$O(n)$	$O(1)$	$O(n)$	$O(1)$	$O(n)$	$O(n)$	$O(1)$
queue	$O(1)$	-	$O(1)$	-	$O(1)$	$O(n)$	$O(n)$	-
deque	$O(1)$	-	$O(1)$	-	$O(1)$	$O(n)$	$O(n)$	$O(1)$
linked list	$O(1)$	$O(1)$	$O(1)$	$O(n)$	$O(1)$	$O(n)$	$O(n)$	$O(n)$
sorted array	$O(n)$	$O(n)$	$O(1)$	$O(\log n)$	$O(1)$	$O(n \log n)$	$O(1)$	$O(1)$
priority_queue	$O(\log n)$	$O(\log n)$	$O(\log n)$	$O(1)$	$O(n)$	$O(1)$	-	

1.2 Двоичная куча

Реализация двоичной кучи на массиве — создаем массив размера sz , и создаем ребра $i \rightarrow 2 \cdot i$, $i \rightarrow 2 \cdot i + 1$.

От такой кучи мы хотим:

- insert(x)
- get_min()
- extract_min()
- erase
- change = {decrease_key, increase_key}

Для такой кучи мы реализуем $sift_up(x)$, $sift_down(x)$ — просеивание вниз и вверх. Процедура должна устраниить конфликты с элементом x . Остальные операции умеют реализовываться через нее.

$$insert = add_leaf + sift_{up}$$

$$extract_min = swap(root, last) + last - 1 + sift_{down}(root)$$

$$erase = decrease_key(-\infty) + extract_min$$

Отдельно отметим построение кучи за $O(n)$ — $sift_down$ поочередно для всех элементов $n, n - 1, \dots, 1$.

```

void sift_up(int v) { // v >> 1 <=> v / 2
    if (key[v] < key[v >> 1]) {
        swap(key[v], key[v >> 1]);
        sift_up(v >> 1);
    }
}

void sift_down(int v, int size) { // indexes [1, size]
    if (2 * v > size) {
        return;
    }
    int left = 2 * v;
    int right = 2 * v + 1;
    int argmin = v;

```

```
if (key[v] > key[left]) {
    argmin = left;
}
if (right <= size && key[right] < key[argmin]) {
    argmin = right;
}
if (argmin == v) {
    return;
}
swap(key[v], key[argmin]);
sift_down(argmin, size);
}
```

Какое-то сегодняшнее дополнение про бинарную кучу есть в прошлом конспекте

К-чная куча Куча на полном К-чном дереве. Этую кучу можно так же хранить в массиве, с 0-индексацией.

- *sift_up* такой же, $O(\log_k n)$
- *sift_down* ищет минимум среди k элементов на каждом шаге, поэтому работает за $O(k \cdot \log_k n)$.

	$2 - \text{heap}$	$k - \text{heap}$
<i>insert</i>	$O(\log n)$	$O(\log_k n)$
<i>extract_min</i>	$O(\log n)$	$O(k \log_k n)$
<i>decreasekey</i>	$O(\log n)$	$O(\log_k n)$
<i>increasekey</i>	$O(\log n)$	$O(k \log_k n)$

Дейкстра на K-ной куче Алгоритм Дейкстры достает минимум n раз, и улучшает ключ m раз

$$a \cdot \log_k n = b \cdot k \cdot \log_k n, a = m, b = n$$

Отсюда $k = \frac{a}{b} = \frac{m}{n}$ в случае Дейкстры. Еще отметим, что $k \geq 2$.

В случае когда $a = b^q$, $q > 1$, то k -куча структура работает за $O(1)$ (вроде бы этот факт мы докажем в домашке), причем с хорошей константой, поэтому применимо на практике (привет, фибоначчиева куча!).

Амортизационный анализ Идея в том, что мы хотим оценить суммарное число операций, а не на каждом шаге работы. То есть вполне может быть итерация алгоритма за $O(n)$, но нам важно, что суммарное число $O(n \log n)$

Так что есть $t_{\text{real}} = t$, $t_{\text{amortized}} = \tilde{t}$.

Метод кредитов Элементам структуры сопоставляем сколько-то монет. Этими монетами элемент «расплачивается» за операции. Также мы накидываем сколько-то монет на операцию. Запрещаем отрицательное число монет. Начинаем с нулем везде.

Обозначим состояния структуры за S_0, S_1, \dots, S_n . Каждый переход стоил t_i , $t_i \geq |\text{operations}|$, где t_i — это сколько мы потратили. Также на i -м шаге мы вбрасываем в систему \tilde{t}_i монет. Тогда

$$\sum t_i \leq \sum \tilde{t}_i \leq A \rightarrow O(A)$$

Стек с минимумом

- *min_stack*
- *push*
- *pop*
- *get_min*

$m_i = \min(m_{i-1}, a_i)$ — поддерживаем минимумы. Операции тривиальны

Очередь с минимумом на двух стеках Храним два стека с минимумом, один из которых мысленно наращиваем в одну сторону, а другой в другую, при этом очередь выглядит как бы как склеенные стеки. То есть мы добавляем элемент в первый стек, а извлекать хотим из второго.

$$X \rightarrow a_n, a_{n-1}, \dots, a_1, |, b_1, b_2, \dots, b_n \rightarrow Y$$

Тогда единственная сложная операция — если мы хотим извлечь минимум, а второй стек пустой. Тогда мы все элементы из первого перекинем во второй по очереди с помощью «извлеки-добавь»

Почему это работает за $O(1)$ на операцию амортизированно? Представим каждому элементу при рождении 2 монеты, одну из которых мы потратим на добавление в первый стек, а вторую на удаление через второй.

set для бедных Хотим не делать *erase*, только *insert*, *find*, *get_min*. Храним $\log n$ массивов, $|a_i| = 2^i$, каждый из которых по инварианту будет отсортирован. Тогда *get_min* работает за $O(\log n)$ — просто берем минимум по всем массивам. Аналогично *find* делается бинпоисками за $O(\log^2 n)$

А как добавлять за $\tilde{O}(\log n)$? Каждый элемент при добавлении в структуру получает $\log n$ монет. Когда мы добавляем элемент, мы создаем новый массив ранга 0. Если было два массива ранга 0, сольем их в новый массив ранга 1 за суммарный размер (и заберем монетку у всех элементов во время слияния), и так далее, пока не создадим уникальный массив для текущего ранга.

Метод потенциалов $\Phi(S_i)$ — потенциал, который зависит только от состояния структуры (**не от последовательности действий**, которая к такому состоянию привела).

Опять вводим t_i , $\sum t_i = O(f(n))$. Определим амортизированное время работы:

$$\tilde{t}_i = t_i + (\Phi(S_{i+1}) - \Phi(S_i))$$

$$t_i = \tilde{t}_i + \Phi(S_i) - \Phi(S_{i+1})$$

Пусть мы показали $\tilde{t}_i \leq f(n)$. Тогда

$$\sum t_i \leq n \cdot f(n) + \Phi(0) - \Phi(n)$$

Нормальный потенциал — такой, что из неравенства выше все еще можно показать О-оценку на $\sum t_i = O(n \cdot f(n))$.

deque для богатых Хотим deque с поддержкой минимума.

Храним два стека как для обычной очереди. Все операции хорошо работают как на очереди, кроме перестройки структуры. В случае с очередью надо было переливать стеки только в одну сторону, а теперь иногда нужно туда-сюда.

Теперь мы будем перекидывать только половину элементов. Тогда нам понадобится 3 стека, один из которых будет вспомогательным для перестройки (там иначе стеки развернуты).

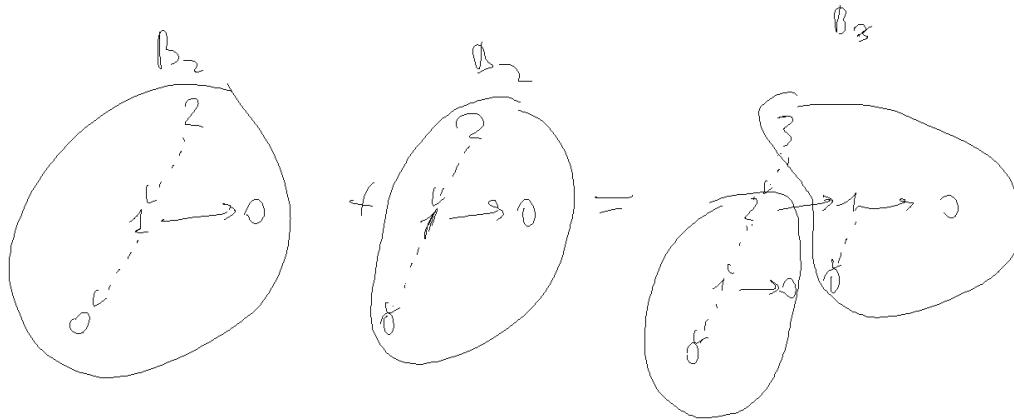
$$\Phi(S_i) = |Size_1 - Size_2|$$

Куча: insert, extract_min, decrease_min, decrease_key, increase_key, merge. Хотим добавить операцию merge - объединить 2 кучи. Считаем, что можем делать операции амортизировано (не разделяем кучи и кучи не персистентные). Меньшую кучу будем добавлять в большую ("переливать" в большую).

Хотелось бы раздать каждой вершине по $\log n$ монет и говорить, что, когда мы "переливаем" кучу, все элементы этой кучи платят по монете. Тогда монет хватит, так как при каждом переливании размер кучи, где находится вершина увеличивается вдвое, значит каждая вершина перельется не более $\log n$ раз. Но все ломается из-за того, что мы можем удалять вершины. Можно ввести потенциалы (подумать), а можно сказать, что у каждого элемента есть ранг ($r(i)$) и при каждом переливании все ранги меньшей кучи увеличиваются на 1. Тогда амортизировано merge будет работать за $O(\log^2 n)$.

Биномиальная куча:

Вершин в биномиальной кучи 2^n , на k -ом слое C_n^k вершин. Из каждой вершины храним ребро в старшего сына, в предка и в следующего брата. $B_k = \text{merge}(B_{k-1}, B_{k-1})$. Заметим, что merge деревьев одного ранга работает за $O(1)$.



Если у нас есть n чисел, то как мы их поместим в кучу размера 2^k ? $n = 2^{k_1} + 2^{k_2} + \dots + 2^{k_m}$. Тогда можем хранить все элементы как кучи рангов k_1, \dots, k_m . Тогда при merge нужно объединять два списка биномиальных куч (будем делать это 2 указателями по 2 массивам), будем объединять кучи одинаковых рангов

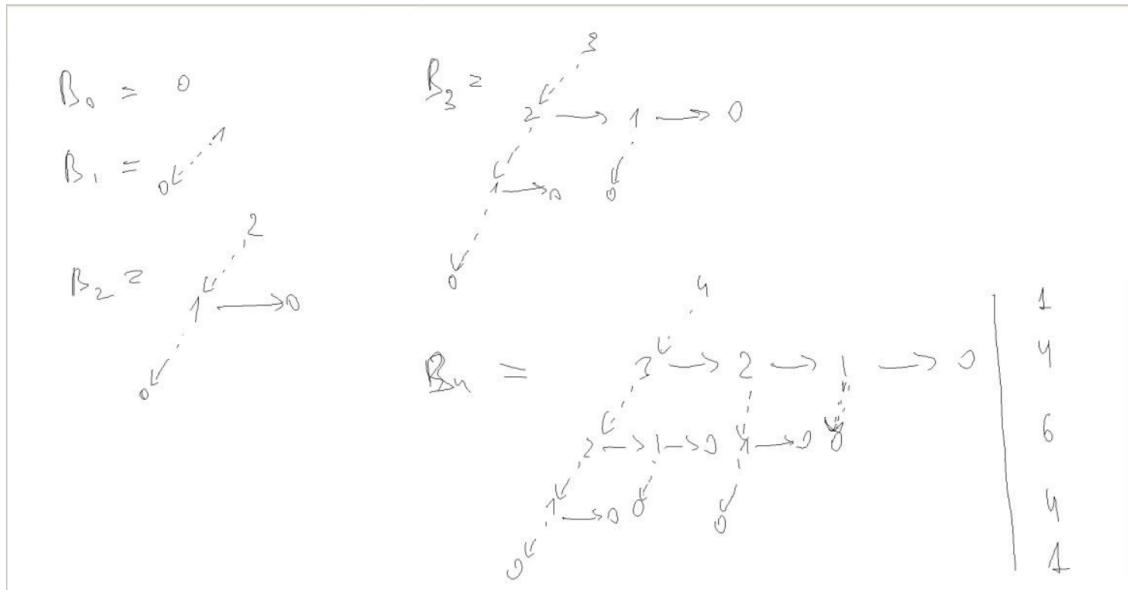
insert - создаем кучу ранга 0 с нашим элементом и делаем merge за $O(\log n)$.

Для поиска минимума будем просто поддерживать глобальный минимум, изменяя его за $O(\log n)$ (пробегаясь по всем кучам) при каждом запросе изменения.

decrease_key

extract_min

increase_key: decrease_key($v, -\infty$) \rightarrow extract_min \rightarrow insert(x). Работает за $O(\log n)$.



Фибоначчиева куча: (чет я устал тешать, чекайте у Кости)

Операции	binary heap	binomial heap	fibonacci heap
insert	$O(\log n)$	$O(\log n)$	$O(1)$
extract_min	$O(\log n)$	$O(\log n)$	$\tilde{O}(\log n)$
decrease_min	$O(\log n)$	$O(\log n)$	$\tilde{O}(1)$
increase_min	$O(\log n)$	$O(\log n)$	$\tilde{O}(\log n)$
merge	$\tilde{O}(\log^2 n)$	$O(\log n)$	$O(1)$
get_min	$O(1)$	$O(1)$	$O(1)$

Больше куч!

	<i>binaryheap</i>	<i>binomialheap</i>	<i>fibonacciheap</i>
<i>insert</i>	$O(\log n)$	$O(\log n)$	$O(1)$
<i>extract_min</i>	$O(\log n)$	$O(\log n)$	$\tilde{O}(\log n)$
<i>decrease_key</i>	$O(\log n)$	$O(\log n)$	$\tilde{O}(1)$
<i>increase_key</i>	$O(\log n)$	$O(\log n)$	$\tilde{O}(\log n)$
<i>merge</i>	$\tilde{O}(\log^2 n)$	$O(\log n)$	$O(1)$
<i>get_min</i>	$O(\log n)$ or $O(1)$	$O(1)$	

Биномиальная куча

Храним биномиальные деревья. Каждому дереву сопоставим ранг. Ранг дерева полностью определяет его структуру. Дерево ранга 0 — одна вершина. Дерево ранга 1 — одно ребро. В общем случае, дерево ранга n содержит корень и полное двоичное дерево размера 2^{n-1} . Дерево ранга $n+1$ — это два слитых вместе дерева ранга n — корень второго указывает на корень первого, а корень первого теперь будет указывать на оба бинарных дерева.

Биномиальная куча — это набор из логарифма биномиальных куч.

Слияние двух деревьев мы научились делать за $O(1)$ — меньшая вершина по ключу становится новым корнем, а дальше перекидываем указатели.

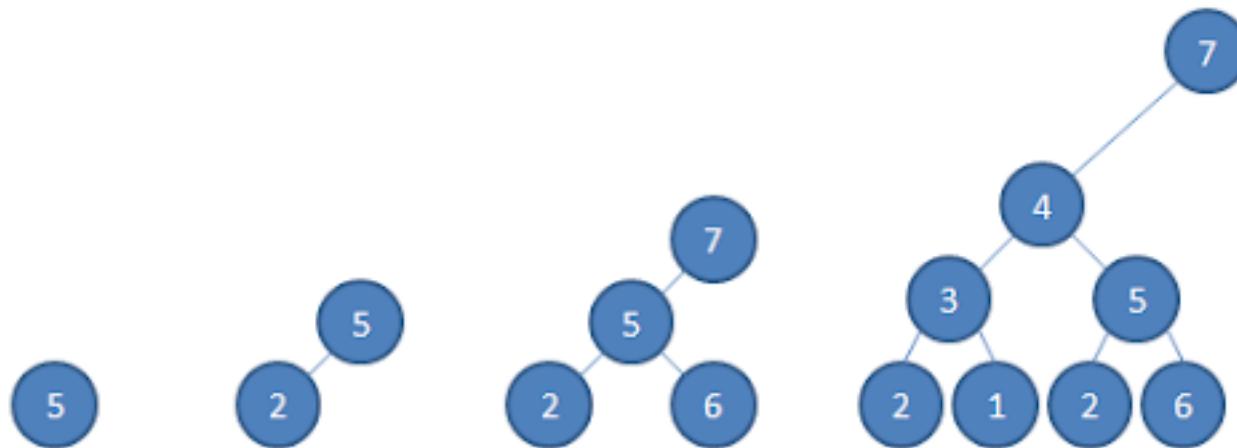
Слияние двух куч — это алгоритм сложения двоичных чисел — при слиянии двух деревьев ранга n мы «переносим» прибавление кучи ранга $n+1$

Как добавлять элемент? Создать кучу на 0 элементов и слить их вместе.

Уменьшение ключа — напишем *sift_up* на нашей новой куче

Удаление минимума — Заметим, что если в правом дереве пройти по правым детям и обозначим их за корни, а их левых детей за полные бинарные деревья, то мы получим набор деревьев рангов $0, 1, \dots, n-1$. Обозначим их за новую кучу, и сольем все вместе.

Увеличение ключа — удалим соответствующий элемент и добавим другой.



Фибоначчиева куча

Хотим сделать биномиальную кучу с послаблениями — делать операции в самый последний момент, менее четкую структуру, etc

Есть деревья, их корни храним в двусвязном закольцованным списке.

Всех детей для всех вершин храним в двусвязном закольцованным списке.

Новый ранг — это количество вершин в списке детей.

На каждом дереве выполнена куча, а также поддерживаем глобальный минимум.

Улучшение ключа делается так — удаляем вершину из своего списка, вместе с поддеревом. Добавляем в корневой список. Если мы удалили уже вторую вершину в поддереве родителя, то делаем каскадное вырезание — прыгаем по предкам с $mark = 1$, и вырезаем их в корневой список, причем все вырезания делаются по очереди.

Удаление делается так — мы приписываем всех детей к корневому списку, а потом вызываем *compact*, которая должна спасти наше дерево и навести порядок.

Псевдокод типовых операций:

```
struct Node{
    Node *child;
    Node *left;
    Node *right;
    Node *parent;
    int rank;
    bool mark;
    int value;
}

list<Node *> roots;

void insert(int x) {
    Node *node = new Node(x);
    roots.insert(node);
}

void merge(list<Node *> a, list<Node *> b) {
    merge(a, b); // O(1) haha super easy
}

int getmin() {
    return argmin->value;
}
```

compact

- Сбрасываем пометки корневого списка в 0
- Переводим дерево в состояние, где все ранги различны
- Храним ранги, мерджим одинаковые
- Как мерджим? Берем меньший корень, и записываем в его детей второй корень

Обозначим $R = \max rank$, $t(H) = root list size$, $m(H) = \sum_v mark(v)$
 $compact$ работает за $O(R + t(H))$.

Анализ времени работы $extract_min$ & $increase_key$ — $\tilde{O}(R)$.

$$\Phi(H) = t(H) + 2m(H) < 3n$$

Пусть каскадное вырезание сделало t_i действий. $m : 1 \rightarrow 0$, $t : +1$. Тогда $\Phi'(H) = \Phi(H) - 1$ за каждое вырезание. Тогда амортизированно вырезание работает за $O(1)$.

Compact:

$$t(H) \leq R, t'(H) = t(H) - R$$

Амортизированно работает за $2R + 1$

Хотим показать $R = O(\log n)$.

$$\forall v \in H \ sz(v) \geq A^{rank(v)}, A > 1$$

Тогда

$$r(v) \leq \log_A s(v) \leq c \log n$$

Возьмем

$$s(v) \geq \phi^{rank(v)-2}$$

<Оффтоп про числа Фибоначчи>

$$1. F_n = F_{n-1} + F_{n-2}, F_0 = F_1 = 1$$

$$2. F_n \geq \phi^{n-2}$$

$$3. \forall i \geq 2 : F_i = \sum_{j=0}^{i-2} F_j + 1$$

</Оффтоп про числа Фибоначчи>

Почему инвариант на размеры сохраняется? Возьмем вершину v с k детьми. Рассмотрим детей в том порядке, в котором их склеивал компакт. Тогда на момент добавления i -й вершины в структуру, ее ранг совпадал с рангом v . Тогда из условия на удаление не более чем одного сына ($mark$) следует, что $rank(i) \geq i - 1$. Тогда мы доказываем по индукции, что у нас все размеры — хотя бы числа фибоначчи,

соответствующие рангу. Тогда $s(v) = \sum_{u \in g(v)} s(u) + 1 \geq \sum_{u \in g(v)} F_{rank(u)} + 1 \geq \sum_{i=0}^{rank(v)-2} F_i + 1 \geq F_{rank(v)-2}$

Хэширование Есть задача сравнения объектов:

$$U = \{objects\}, u, v \in U : u \neq v?$$

Введем функцию $h : U \rightarrow \mathbb{Z}_m$, такую что $\forall u, v \in U : u \sim v \rightarrow h(v) = h(u)$. Обычно $m \ll |U|$.

Зачем юзать хэши, а не наивное сравнение?

- Бывает много сравнений, и мы не хотим дублировать вычисления
- Безопасность
- Для некоторых хешей верно, что $h(f(v, u)) = g(h(v), h(u))$. То есть можно вычислить хэш от некоего объекта, пользуясь уже посчитанными хэшами для других объектов.
- Иногда мы сравниваем объекты не на равенство, а на изоморфность.
- Сравнивать объекты бывает дорого

Требования к хэшу:

- вычислим за линейное время
- детерминирован
- семейство хэш-функций \mathbb{H} (и можем взять сколько угодно оттуда)
- равномерность $\forall_{v \neq u} p_{h \in \mathbb{H}}(h(u) = h(v)) \simeq \frac{1}{m}$
- масштабируемость
- необратимость
- (*optional*) лавинный эффект (при маленьком изменении объекта хэш меняется сильно)

Важно, что мы хотим брать случайную функцию из семейства \mathbb{H} на момент старта программы, потому что псевдослучайная функция на самом деле нам дает детерминированный алгоритм, который неверен.

Идеальная хэш-функция Так как объектов счетно, то отсортируем их, далее для каждого объекта запомним случайную величину от 1 до m (или даже можем запоминать ее лениво!).

Полиномиальный хэш Сводим объекты к строкам и хэшируем строки:

$$s = c_0c_1c_2 \dots c_{n-1}, c_i > 0$$

$$h_b(s) = \sum_{i=0}^{n-1} c_i \cdot b^i \pmod{m}$$

$$p_b(h_b(s_1) = h_b(s_2)) \leq \frac{n}{m} \text{ для фиксированного } m$$

Доказательство: Рассмотрим функцию как многочлен, теперь для равных функций смотрим на то, равна ли разность многочленов нулю для какого-то b . У многочлена от b степень равна n , а вероятность попасть в конкретный модуль $\frac{1}{m}$.

$$h(s_1 + s_2) = h(s_1) + h(s_2) \cdot b^{|s_1|}$$

Хэш-таблица

Key-value storage: три типа операций — $set(x, y)$, $get(x)$, $has(x)$. Хэш-таблица умеет выполнять такие запросы за $O(1)$.

Хотим делать индексацию не по ключу, а по хэшу от ключа $x_0 \rightarrow h(x_0)$.

К сожалению, бывает так, что в одну и ту же ячейку попало много элементов (матожидание числа коллизий порядка $\frac{n^2}{m}$, где m — размер хэш-таблицы). Мы будем называть это коллизиями.

Коллизии можно решать двумя способами, соответствующие хэш-таблицы имеют **открытую** или **закрытую** адресацию.

Закрытая адресация (или решение коллизии методом цепочек) — в каждой ячейке храним список, в который будем добавлять соответствующие элементы. Матожидание длины списка будет порядка $\frac{n}{m}$.

Хэш-таблица работает линейно от числа элементов, которые в ней когда-либо были, поэтому динамическая хэш-таблица работает за амортизированное время.

«stop-the-world»-концепция — если внутренние параметры системы бьют тревогу, сделаем глобальное изменение. В случае хэш-таблицы, если в таблице сейчас t элементов, создадим новую хэш-таблицу удвоенного размера, в которую перехэшируем оставшиеся элементы.

Замечание. Очень часто разработчик не хочет амортизированное время работы, потому что боится внезапного «stop-the-world», потому что это выключает систему на длительное время. Пример — финал Т18.

Открытая адресация

Делаем вид, будто коллизий не бывает (то есть, в каждой ячейке храним только одно значение). Кроме того, рядом с ячейкой храним ключ элемента (или -1 , если там пусто). Тогда если мы хотим найти элемент x , мы смотрим в ячейку $h(x)$, и идем от нее вправо до тех пор, пока не встретим x или -1 .

Ожидаемое время — это $\frac{1}{1-\alpha}$, где $\alpha = \frac{n}{m}$. На практике все считают, что время — константа.

Удаление с открытой адресацией — нетривиальная задача, потому что удаление элемента рушит цепочки.

Удаление без «stop-the-world» — удаляем все элементы от нашего элемента до ближайшей -1 , но потом вернем их обратно с помощью добавлений

Удаление с «stop-the-world» — кроме пометки -1 делаем пометку «зарезервирована». Тогда при удалении элемента мы ставим в его ячейку пометку «зарезервирована». Тогда при линейном проходе мы делаем вид, что резерв — это настоящий элемент, а при добавлении мы можем записать в резерв. Резерв включается в параметр α , поэтому нам понадобится делать перестройки.

Кроме линейного сканирования можно делать что-то типа «прыжкам по хешу» ($h(x) + jh'(x)$), но на практике линейное сканирование — наш бро.

Совершенное хэширование

Нам изначально дано сколько-то ключей, мы хотим делать $get(x)$, $set(x, y)$ за гарантированные $O(1)$ с ожидаемым $O(n)$ на преподсчет.

Сделаем хэш-таблицу с закрытой адресацией. В каждой ячейке ожидаемое $O(1)$ элементов. Сделаем новую хэш-функцию, которая переносит все элементы из l_i в ячейки размера l_i^2 . Вероятность коллизии при таком хэшировании меньше $\frac{1}{2}$, поэтому мы будем просто рандомить хэш-функцию, пока не получим отсутствие коллизий, что займет у нас ожидаемое $O(1)$.

Фильтр Блума

insert, get

Запросы *get* работают необычным образом — No → No; Yes → Yes($1 - p$)/No(p). То есть, если структура говорит, что элемент в ней лежит, то он в ней точно не лежит. Далее минимизируется p .

Фильтр Блума является массивом из битов длины m . Фильтр выбирает k хэш-функций, элементу сопоставляется k значений хэша.

insert — в каждое из k значений ставим 1

get — проверяем, что во всех k значениях стоит 1.

Время работы. $k \sim \frac{m}{n}$. Рассмотрим вероятность ложноположительного срабатывания. Вероятность того, что клетка свободна — $(\frac{m-1}{m})^{kn}$. Тогда вероятность ложноположительного срабатывания это $(1 - (\frac{m-1}{m})^{kn})^k \sim (1 - e^{-\frac{kn}{m}})^k$.

Путем долгих вычислений получаем $k = \ln 2 \cdot \frac{m}{n}$.

Кроме того, ФБ поддерживает операции пересечения и объединения множеств.

Деревья поиска Храним структуру данных, которая должна уметь делать все то же самое, что можно делать с отсортированным массивом, но еще с добавлением-удалением.

Binaary search tree (BST) Корневое дерево. В вершине храним левого сына, правого сына, предка, ключ.

Обозначения:

- $l(v)$ — левый сын
- $r(v)$ — правый сын
- $p(v)$ — предок
- $key(v)$ — ключ
- $T(v)$ — поддерево
- $S(v) = |T(v)|$ — размер
- $A(v)$ — множество предков
- $seg(v) = [\min_{u \in T(v)} key(u); \max_{u \in T(v)} key(u)]$

Условие BST:

$$\forall u \in T(l(v)) : key(u) < key(v)$$

$$\forall u \in T(r(v)) : key(u) > key(v)$$

То есть, ключи лежат в порядке лево-правого обхода (в порядке «выписать левое поддерево - выписать вершину - выписать правое поддерево»)

$$v \in seg(u) \leftrightarrow v \in T(u)$$

Поиск в дереве — надо сделать спуск. То есть, если текущая вершина — не та, которая нам нужна, то можно понять, где лежит нужная нам, с помощью условия BST.

Нахождение следующего — либо спуск вправо, либо подъем по предкам до первого большего.

Основная проблема — операции вставки/удаления, которые должны делать дерево сбалансированным (таким, высота которого нас устраивает, то есть примерно $\log n$)

Декартово дерево У каждой вершины будем хранить не только ключ, но и какой-то приоритет y . Построим дерево так, что по y это куча, а по x это BST.

Тогда декартово дерево задается однозначно, если определить все приоритеты. Почему? Расставим точки на плоскости в соответствии с (x, y) , после чего найдем корень. У корня будет наименьший приоритет и поэтому он определяется единственным образом (будем считать, что приоритеты различны). Тогда к корню нужно присвоить слева и справа по дереву, которые рекурсивно строятся в левой и правых частях.

Утверждение Если взять случайные y , то матожидание глубины дерева $O(\log n)$

Доказательство: Нет.

Утверждение Если взять случайные y , то матожидание глубины каждой вершины $O(\log n)$

Доказательство. Матожидание высоты вершины — это число вершин, которые являются ее предками. Вершина i будет предком вершины j , если $y_i = \max\{y_i, y_{i+1}, \dots, y_j\}$. Матожидание суммы таких величин для i это $\sum_{j=0}^{i-1} \frac{1}{i-j} + \sum_{j=i+1}^{n-1} \frac{1}{j-i} \leq 2 \sum_{i=1}^n \frac{1}{i} = O(\log n)$

2-3 Дерево

Представляет собой структуру данных, которая является сбалансированным деревом поиска, удовлетворяющее двум условиям:

- Все листья будут на одной глубине, значения будут храниться в листьях
- У всех вершин число исходящих ребер $deg_v \in \{0, 2, 3\}$

Также мы в каждом поддереве будем хранить максимум, а детей будем хранить упорядоченными по величине максимума.

Операции на дереве: поиск Спуск будем делать рекурсивно. Поскольку мы хранили максимум, то мы среди детей находим первое число, большее x , спускаемся в соответствующего сына.

Операции на дереве: добавление Если степень предка после добавления стала равна 4, то создадим два сына размеров 2 и 2, и рекурсивно рассмотрим отца. Если дошли до корня, то создадим новый корень степени 2.

Операции на дереве: удаление Рассмотрим ситуации, которые могли возникать при удалении. Заметим, что у вершины есть отец, дедушка, дядя, брат (предок, другой сын прародителя, прародитель, другой сын предка).

Если у вершины было 2 брата, то после ее удаления ничего менять не надо.

Если у вершины был 1 брат, то нарушается инвариант на степени. Посмотрим на детей дяди. Если их было 3, то можно перераспределить 3 + 1 как 2 + 2, и инвариант не нарушится. Если же там было 2 ребенка, то склеимся в одну вершину степени 3, и уменьшим степень предка на 1. Рекурсивно запустим процесс балансировки от него.

Оптимальное дерево поиска Обозначим число детей за d . Тогда операции работают за $d \cdot \log_d n$. Найдем точку минимума: $(d \cdot \log_d n)' = \ln n \cdot \frac{\ln d - 1}{\ln^2 d}$. Нуевой корень производной при $d = e$. Таким образом, 2-3 дерево достаточно близко к оптимальному.

B+ дерево Зафиксируем константу T . Все значения опять храним в листьях. У вершин (кроме корня) степень от T до $2 \cdot T - 1$. $2 \leq deg_{root} \leq 2 \cdot T - 1$

Обычно T делают достаточно большим, чтобы дерево работало во внешней памяти.

Операции: вставка Если у вершины степень стала $2T$, то мы можем разделить ее на две вершины, подвесить их к предку, и рекурсивно запустить процесс у предка.

Операции: удаление Плохая ситуация при удалении — степень предка стала равна $T - 1$. Посмотрим на соседних братьев. Если один из них по степени больше T , то мы можем позаимствовать у него крайнего сына, чем починим свою степень. Если же у обоих братьев степень T , то мы можем смержить себя с братом, после чего рекурсивно продолжить процесс в предке, потому что у предка степень уменьшилась на 1.

Почему B+, а не B? В начале стоит сказать, что B-дерево обладает схожей структурой, но разрешает хранение значений не только в листьях. Его глубина не более чем на 1 меньше, чем у соответствующего ему B+-дерева. Но при этом элементы B+-дерева можно поддерживать в двусвязном списке, что удобно, а также можно итерироватьсь во внешней памяти.

Научимся в детермированный, так сказать, декартач.

2-3 дерево

В этом дереве есть вершины степени 2 и вершины степени 3. Элементы записываются только в листьях.

Ряд условий 2-3 дерева:

1) Все листья на одной и той же глубине. 2) Степень каждой внутренней вершины либо 2, либо 3.

В промежуточных вершинах храним ключи, которыми можно разделять вершины-сыновей.

У вершины с 3 детьми, два ключа - максимум поддеревьев двух первых детей, с 2 детьми - максимум в первом (а ещё храним максимум во всём поддереве).

Смотря на ключ, мы можем понять в какое из двух (или трёх) деревьев идти.

Вставка в 2-3 дерево.

Найдём место, в котором элемент должен быть и просто вставим его. Может случиться, что детей уже было 3, а теперь стало 4. Тогда разобьём нашу вершину (у которой стало 4 сына) и разобьём её на 2 и переподвесим их теперь к её матушке. Ну и продолжаем это процесс пока не дойдём до корня. Если вдруг мы расплели корень, то создадим новый корень и подвесим к нему наши две вершины. После того, как мы всё сделали, можно запуститься рекурсивно вверх от только что вставленной и идём вверх. Это работает быстро, потому что уровней логарифм.

Удаление из 2-3 дерева. Найдём элемент и тупо удалим его. Если у его предка было три ребёнка - всё хорошо. Предположим, что у вершины было 2 ребёнка, а остался один (трагично, не правда ли?). Теперь начинаются боль и мучения. В общем тут перебор случаев, который можно разобрать кроме того случая, когда у нас только один брат, у которого ровно два сына. Тогда мы переподвешиваем нашего сына к брату и переходим на уровень выше. Если мы упрёмся в корень, то просто возьмём корень и удалим его за ненадобностью.

Всё 2-3 дерево можно провязать двусвязными списками, связывающими вершины на одном уровне.

Почему бы вместо 2-3 дерева не сделать бы 5-11 дерево? Оценим это так. Пусть у всех вершин степень d . Тогда стоимость операции:

$$\begin{aligned} d \cdot \log_d n &= d \cdot \frac{\ln n}{\ln d} \Rightarrow \\ (d \cdot \log_d n)' &= (d \cdot \frac{\ln n}{\ln d})' \Rightarrow \\ (d \cdot \log_d n)' &= (d \cdot \frac{\ln d - 1}{(\ln d)^2})' \end{aligned}$$

0 производной в $d = e$. Значит, 2 и 3 - то хорошие приближения.

$B+$ - дерево.

Степень каждой вершины от t до $2t - 1$. Степень корня - от 2 до $2t - 1$.

Поиск вершины - также.

Такие деревья нужны для алгоритмов во внешней памяти.

Тогда асимптотика - $O(CPU \cdot (d \log_d n) + HDD \cdot \log_d n)$. Тогда, хочется брать достаточно большой d .

Вставка:

Ищем место для вставки и смотрим, куда вставить. Если в какой-то момент стало $2t$ детей, то сплитим на t и t и переподвешиваем.

Удаление: Пусть в вершине теперь стало $t - 1$ детей (иначе всё хорошо). Если у нас соседний брат - "Большой Брат" (хотя бы $t + 1$ ребёнок), то всё - мы нашли жертву, которую будем грабить (переподвешиваем одного из сыновей брата к нашей вершине). Если нет, то вдохновившись небезызвестным произведением Кена Кизи, подкидываем брату $t - 1$ кукушонка. Переходим на уровень выше и продолжаем процесс рекурсивно. Отдельно оговорим корень. Тут проблема только тогда, когда у него осталась 1 вершина, но тогда корень улетает в небытие, и в этом городе появляется новый корень (время для нового Жоко).

А сейчас будет нерекурсивная вставка в B -дерево. Ослабим ограничения на вершины теперь границы это $t - 1$ и $2t - 1$. Во время поиска вершины будут разбиваться в момент спуска. Как только вершина может переполниться (её степень $2t - 1$), тогда разобьём нашу вершину на $t - 1$ и t и пойдём дальше вниз.

Ещё одна отсечка, от которой плкается мозг - в каждой вершине нам надо хранить массив максимумов детей. Будем хранить эту вещь бинарным деревом поиска (ДД).

Задачи на отрезке. Введем какую-то произвольную операцию \oplus , и будем отвечать на запросы $get(l, r)$ на массиве a , ответом на которые будет $a_l \oplus a_{l+1} \oplus \dots \oplus a_r$. Также введем операцию изменения на отрезке $a_i := change(a_i, x)$.

Потребуем от \oplus ассоциативность — $a \oplus (b \oplus c) = (a \oplus b) \oplus c$.

Префиксные суммы. Если в задаче нет запросов изменения (и элементы образуют группу, то есть есть обратный элемент), то посчитаем $p_i = p_{i-1} \oplus a_i$. Тогда ответ на запрос — это $p_r \oplus p_{l-1}^{-1}$. Если операция некоммутативна, то нужно будет пострадать, но вроде бы можно просто сделать $p_{l-1}^{-1} \oplus p_r$. Построение за $O(n)$, запрос за $O(1)$.

Sparse table. Если элементы не образуют группу, но задача все еще статическая, то можно сохранить значения, соответствующие \oplus по всем отрезкам длины 2^k . Тогда, когда нужно ответить на запрос $get(l, r)$, можно взять перекрывающиеся отрезки $[l, l+2^i)$ и $(r-2^i, r]$. От операции требуется $a \oplus b = b \oplus a$ и $a \oplus a = a$. Есть модификация, позволяющая обойтись без второго свойства. Построение за $O(n \log n)$, запрос за $O(1)$.

Segment tree. Хотим к предыдущей задаче добавить обновление в точке. Хотим сохранить какое-то множество отрезков S , чтобы потом по нему восстанавливать ответ на произвольном отрезке $[l, r]$, склеивая не более чем $O(\log n)$ отрезков. Также должно быть не более $O(\log n)$ отрезков, содержащих какой-либо элемент.

Построим двоичное дерево над массивом, где вершина на глубине i будет отвечать за отрезок длины 2^{k-i} , где $n = 2^k$. Разбивать запросы на отрезки будем таким образом: рассмотрим все отрезки внутри запроса, и выкинем вложенные. На каждой глубине мы возьмем не более двух отрезков, поэтому суммарно запросы будут работать за $O(\log n)$.

								1: [0, 16)									
								2: [0, 8)									
4: [0, 4)				5: [4, 8)				6: [8, 12)				7: [12, 16)					
8: [0, 2)	9: [2, 4)	10: [4, 6)	11: [6, 8)	12: [8, 10)	13: [10, 12)	14: [12, 14)	15: [14, 16)	16: 0	17: 1	18: 2	19: 3	20: 4	21: 5	22: 6	23: 7	24: 8	25: 9
16: 0	17: 1	18: 2	19: 3	20: 4	21: 5	22: 6	23: 7	24: 8	25: 9	26: 10	27: 11	28: 12	29: 13	30: 14	31: 15		

Реализация будет такой — сделаем рекурсивную функцию get , которая хочет обойти дерево, зайти во все «ключевые» отрезки запроса, и посчитать итоговый ответ. Для $get(v, l, r)$ бывает три случая:

- v не отвечает ни за что из отрезка $[l, r]$, поэтому не делаем ничего и прекращаем работу.
- $[l, r]$ содержит отрезок вершины v , и тогда можно обработать этот отрезок и прекратить работу.
- Иначе надо спуститься в детей, и повторить процесс

Lazy propagation. Пусть у нас появился запрос $change$, модифицирующий отрезок. Будем хранить в вершине пометки вида «мы хотели сделать со всеми детьми такую-то операцию изменения», которые мы изначально будем проставлять с запросом изменения на все «ключевые отрезки». При этом во время запроса изменения мы по сути не сделаем изменений, только проставим пометки. Но теперь мы при каждом обращении к вершине проталкиваем модификатор (если есть) вниз, и меняем текущее

значение *value*. Таким образом, после проталкивания все величины остаются валидными (кроме тех, которые находятся где-то глубоко в дереве, но к которым мы не спустились).

Иначе говоря, мы считаем, что перед тем, как обратиться к какой-то вершине, мы обязаны обратиться ко всем вершинам-предкам, и тогда можем гарантировать, что в ней будут храниться актуальные значения параметров.

Требования к lazy propagation — дистрибутивность операции *change* относительно \oplus , а также модификаторы также должны являться полугруппой.

Запросы на деревьях. Будем обсуждать две задачи: **LCA** и **LA**.

Постановка задачи *LCA*: даются запросы на вычисления наименьшего общего предка двух вершин v, u . То есть, ответ на запрос — это вершина наибольшей глубины такой, что она является предком и v и u .

Постановка задачи *LA*: даются запросы на вычисления предка вершины v на глубине k .

Двоичные подъемы. Для каждой вершины преподсчитаем $p_{v,i}$, которая будет хранить информацию о том, какая вершина является 2^i -ым предком вершины v . Насчитать их можно с помощью обхода дерева за $O(n \log n)$.

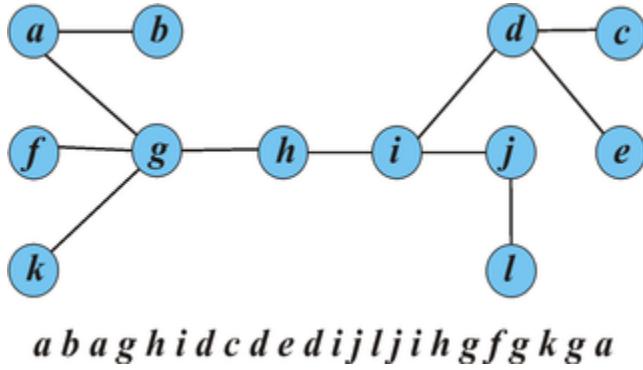
Решение задачи *LA* с помощью двоичных подъемов тогда будет таким: разложим число k на степени двойки, и сделаем соответствующие «прыжки»: $v \rightarrow p_{v,i_1} \rightarrow p_{p_{v,i_1}, i_2} \rightarrow \dots$

Решение задачи *LCA* будет таким: сначала мы выровняем вершины по глубине. Иначе говоря, решим задачу *LA*, чтобы после ее решения вершины запроса имели одинаковую глубину (при этом мы хотим двигаться только более глубокую вершину). Теперь, когда вершины на одной глубине (и если не совпали!), мы будем двигаться по степеням двойки от больших к меньшим таким образом: Если LA_{2^i} наших вершин не совпали, то тогда поднимем вершины запроса на 2^i . Такой процедурой мы найдем таких предков вершин u и v , которые не совпадают, находятся на одной глубине, и при этом имеют общего предка. Этот предок и будет ответом на задачу.

Решение с двоичными подъемами работает за $O(n \log n)$ преподсчета и $O(\log n)$ на запрос для обеих задач.

Offline. Решить *LA* в оффлайн можно обходом в глубину с поддержкой стека за $O(n+q)$. Решение *LCA* в оффлайн можно сделать с помощью алгоритма Тарьяна с СНМом (просто как факт) за $O((n + q)\alpha)$.

Эйлеров обход. Мысленно превратим каждое ребро в два ребра, одно из которых ориентировано вверх, а другое вниз. Тогда в таком графе можно сделать обход по типу Эйлерова — каждое ребро пройдем ровно один раз. Будем выписывать вершину v каждый раз, когда проходим по ребру из v .



Строго говоря, обход можно делать не из корня, а просто потом сделать циклический сдвиг, но обычно все запускают обход из корня и не парятся.

Важное свойство — подотрезок нашего обхода является путем. При этом путь между u и v в эйлеровом обходе содержит $LCA(u, v)$ (потому что путь из u в v точно проходит через $LCA(u, v)$), а еще не содержит предка $LCA(u, v)$ (потому что по ребру в него проход был дважды, и если мы посетим его после lca , то мы уже не могли спуститься обратно). То есть, *LCA* будет самой высокой вершиной на подотрезке обхода между u и v . Это является сведением к задаче *RMQ*. Тогда с помощью sparse table можно получить решение за $O(n \log n + q)$.

LA тоже можно решать с помощью эйлерова обхода, делая спуск по дереву отрезков с поиском первой вершины на высоте хотя бы k .

Метод четырех русских. Пушка, которая решит нам *LCA* за $O(n + q)$. Мы разобьем задачи на большие и маленькие. Нам не обязательно решать маленькие задачи при их появлении, если мы можем заранее решить все возможные маленькие задачи.

RMQ ± 1 . Заметим, что наше *RMQ* при поиске *LCA* обладала тем свойством, что $a_i = a_{i-1} \pm 1$, $a_i \neq a_{i-1}$. Разобьем массив на блоки размера k , в каждом блоке посчитаем максимум и получим массив $b_1, \dots, b_{\frac{n}{k}}$. Насчитаем на нем разреженные таблицы. Также в блоке насчитаем префиксные и суффиксные максимумы. Теперь мы умеем за $O(1)$ отвечать на все запросы, кроме тех, которые полностью лежат в одном блоке.

Для решения такой задачи мы посчитаем все возможные последовательности из ± 1 длины k , там возьмем все подотрезки, и для них за линию решим. То есть за $O(2^k \cdot k^3)$. Можно, наверное, и лучше, но нам пофиг.

Теперь положим $k = \lceil \frac{\log n}{2} \rceil$. Тогда маленькая задача решится за $O(\sqrt{n} \cdot \log^3 n)$. А для большой задачи преподсчет будет работать за $O(\frac{n}{k} \log n) = O(n)$. Таким образом, задача решена за $O(n + q)$.

Решение произвольного *RMQ* делается с помощью $RMQ \rightarrow LCA \rightarrow RMQ \pm 1$. Первый переход делается с помощью построения ДД на массиве с помощью стека. Тогда *RMQ* на отрезке это *LCA* для соответствующих вершин.

Ladder decomposition. Предложим другое решение задачи *LA*. Разобьем дерево на пути таким образом: возьмем самую высокую вершину, которая еще не покрыта путями, и возьмем из нее самый глубокий путь вниз. Также насчитаем двоичные подъемы. Такое можно решить за $O(n \log n)$. Каждый путь выпишем явно.

После этого мы мысленно удвоим все пути. Возьмем и выпишем еще столько же вершин вверх для каждого пути. Общая память все еще линейна.

Теперь пусть нам надо сделать подъем на k . Найдем наибольшее i такое, что $2^i \leq k$, сделаем такой прыжок. После чего мы оказываемся в вершине, самый глубокий путь из которой вниз был по длине не меньше, чем 2^i . Это значит, что ответ на задачу хранится в выписанном пути для этой вершины и его можно найти за $O(1)$.

Персистентность

Есть структура данных T и операции. Некоторые операции изменяют T , а некоторые не изменяют. Тогда можно говорить о версиях структуры T в разные моменты времени. *Персистентность* - это свойство структуры данных делать запросы к старым версиям.

Уровни персистентности:

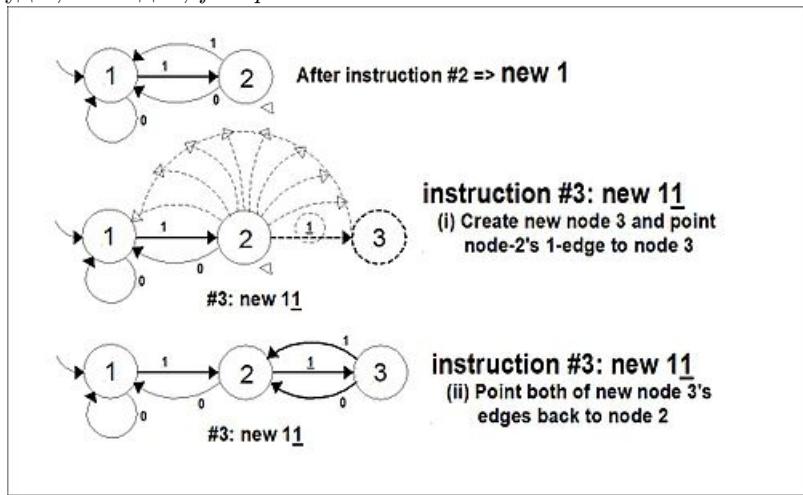
- *partial(weak)* - версии образуют цепочку, то есть каждая следующая версия - это измененная предыдущая (запросы изменения можно делать только к последней версии).
- *full(strong)* - версии образуют дерево (запросы изменения можно делать к любой версии).
- *functional(confluent)* - поддерживаются операции сразу для нескольких версий (допустим *merge* для версий декартова дерева).

Персистентный массив.

Для каждой клеточки храним вектор пар (t_x, v_x) - в момент t_x мы изменили этот элемент массива на v_x (храним в порядке возрастания t_x). Бинпоиском отвечаем на запрос *узнать значение элемента версии t* .

Такая персистентность *partial*, но не *full* и *functional*.

Pointer machine. Способ организовать структуру данных. Будем хранить все в узловой структуре, поддерживая связи между ними с помощью указателей. Тогда вместо изменения вершины мы просто создаем новую. Таким образом, чтобы изменить вершину x , нужно изменить суммарно $O(h)$ вершин. Такая структура данных будет, очевидно, *full persistent*.



Малополезная картинка

Full персистентный массив. Давайте создадим двоичное дерево над массивом размера n с высотой $O(\log n)$, реализовав его как pointer machine. Тогда теперь мы можем сделать изменение произвольного элемента в произвольной версии, получив $O(\log n)$ времени работы (и очевидную реализацию персистентного ДО в придачу).

Персистентное декартово дерево.

Заметим, что наше декартово дерево можно было бы реализовать через pointer machine, что даст нам что-то очень похожее на то, что мы хотим от такой структуры данных. К сожалению, такая структура данных не будет *functional persistent*, потому что есть конструктивный способ очень сильно расширить дерево в высоту (достаточно просто мерджить одну и ту же вершину с последней версией дерева, получая бамбук).

Проблема у нас возникла в тот момент, когда мы воспользовались старой идеей приоритетов. Теперь будем вычислять что-то типа приоритетов динамически. А именно, будем считать, что приоритет дерева L больше приоритета дерева R , если $\text{rand}() < \frac{S(L)}{S(L)+S(R)}$. Можно показать, что теперь высота ДД все еще $O(\log n)$.

Задачи оптимизации. Во всех предыдущих задачах на структуры данных мы работали, оптимизируя какую-либо очевидную задачу (например, сумму на отрезке). Теперь же мы будем решать задачи, которые непонятно как решать, кроме как полным перебором всех вариантов ответа.

В качестве примера возьмем задачу *subset sum*. В ней нам нужно найти подмножество заданного множества с фиксированной суммой S . Если перебрать все подмножества, и посчитать сумму по каждому подмножеству.

Перебором можно решить любую задачу, если перебрать все возможные ответы (множество вещественных чисел для нас тоже конечно!).

Подмножества хочется как-то пронумеровать. К сожалению, непонятно как закодировать 2^n чисел, если раньше мы разрешали в *RAM*-модели числа до $C^k \cdot n^k \cdot A$. Поэтому мы просто уточним *RAM*-модель, и разрешим $C^k \cdot t(n)^k \cdot A$ (Считая $t(n)$ таким временем работы, что внутри нет длинной арифметики).

Динамическое программирование. Иногда бывает полезно запоминать промежуточные величины перебора. Более того, часто перебор можно «ужать», если нам в переборе нужны не все величины (например, в задаче «*subset sum*» достаточно помнить только общую сумму, если перебирать элементы по очереди).

Сделаем $dp(i, x) \in \{0, 1\}$, которая будет говорить, можно ли набрать сумму x с помощью первых i элементов. Тогда $dp(i, x)$ можно пересчитать через $dp(i - 1, x)$ и $dp(i - 1, x - w_i)$.

Требования к нашей динамике:

- Граф вычислений ацикличен.
- «Состояния» динамики явно задают нам всю необходимую информацию.

Задача о рюкзаке. Пусть нам заданы n, S, w_i, c_i (то есть элементы с весами и стоимостями). Мы хотим выбрать некоторое подмножество с суммарным весом не более S и максимальной суммой стоимостей.

Мы можем сделать динамику $dp(i, w, c) \in \{0, 1\}$, которая решит нашу задачу. Но заметим, что наше решение монотонно по параметру c (то есть, для равных i и w стоит отдавать предпочтение ответу с максимальным c). Тогда c можно сделать *значением* динамики. То есть, пересчитывать динамику $dp(i, w) \in C$ как максимум из $dp(i - 1, w)$ и $dp(i - 1, w - w_i) + c_i$. Кстати, заметим, что тут задача монотонна по всем параметрам сразу.

Задача коммивояжера (TSP). Заданы точки на плоскости. Надо найти кратчайший кольцевой маршрут, проходящий по всем точкам хотя бы единожды.

Есть очевидное решение за $O((n-1)!)$. Воспользуемся ДП по подмножествам, основная идея которого — понять, что нам в состоянии важнее всего только то, в каком *множестве* вершин мы уже были, и в каких вершинах мы уже оказались. Закодировать множество мы можем с помощью двоичной маски. Решение с такой идеей отработает уже за $O(2^n n^2)$.

ДП по подстрокам. Отдельный трюк, когда подстроки пересчитываются через свои подотрезки. Важное отличие в том, что мы можем пересчитываться через несколько задач сразу ($dp(l, r) = dp(l, k) + dp(k, r)$), а за счет этого порядок пересчета на графике может быть неочевидным.

Meet-in-the-middle. Пусть нам при решении задачи динамического программирования нужно найти кратчайший путь из s в t на графе вариантов. Перебрать весь граф из вершины s может быть тяжело с точки зрения вычислений. Мы попробуем запустить поиск сразу из двух вершин — из s в прямую сторону, и из t в обратную. Тогда любая общая вершина для двух наших переборов задает путь из s в t .

Subset sum. Применим данную технику в задаче *subsetsum*. Поделим множество предметов на две равные группы. Тогда за $2^{\frac{n}{2}}$ мы можем найти все возможные суммы для каждой из двух групп элементов. Теперь за $O(2^{\frac{n}{2}})$ переберем все суммы x для первой группы, и с помощью хэш-таблицы проверить наличие $S - x$ для второй группы.

Рюкзак. Примерно то же самое можно сделать для задачи о рюкзаке, только теперь нам понадобится узнавать максимум на префикссе.

Максимальная клика в графе. Клика — связный подграф. Мы хотим найти максимальную клику в графе. Разобьем граф на две доли (левую и правую, размерами в пополам). Теперь посчитаем $dp_{submask}$ — максимальную подклику для подмножества вершин. Тогда, зная пересечение списков смежности по всем подмножествам вершин, мы можем поступать следующим образом: находить клику в левой доли, брать множество ее соседей в правой доле, и смотреть на соответствующее значение dp .

Компактность динамики. Будем считать динамику *компактной* по какому-то из измерений, если для пересчета значений динамики нужно помнить *не очень много* слоев по этому параметру. Например, динамика для НОП будет компактной по обоим параметрам (и там, и там достаточно помнить всего пару слоев).

Оптимизационные задачи. Часто человечество занимается тем, что берет перебор, и пытается сделать этот перебор оптимально возможным, чтобы работало ну очень-очень быстро. Для примера подобной задачи часто решают задачу коммивояжера (TSP), которая не имеет решения лучше, чем за экспоненту, но при этом современными методами решается для $n \sim 200$. Задача, между тем, актуальная, потому что транспортные компании, вот это все. Что должен сделать хороший перебор:

- Запоминать промежуточные решения, чтобы у нас была возможность экстренно остановить перебор.
- Поставить отсечения таким образом, чтобы не идти в те состояния, где **точно** не будет решений.

Задача о ферзях. Пусть мы хотим расставить n ферзей на доске $n \times n$. Как закодировать состояние? Можно, например, с помощью $2n$ координат. Заметим, что гораздо лучше зафиксировать перестановку, а потом делать ферзей (i, p_i) . Тогда такие ферзи гарантированно не бьют друг друга по вертикали или горизонтали. Остается только проверить диагонали. Тогда если перебор будет строить перестановку слева направо, то мы сможем «отрезать» некоторые состояния (например, не ставить ферзя в те клетки, которые еще не бились предыдущими ферзями).

Branch & bound method. Мы будем оптимизировать две вещи:

Границы. Нам хочется понимать, когда мы сможем отрезать ветку перебора, чтобы не упустить оптимальный ответ. Например, выходить из ветки, если текущий ответ превышает нынешний оптимальный. Идеально — ввести соответствующую функцию $f(p)$, и делать так:

```
if (cur + f(p) > best) {
    return;
}
```

Такая функция должна, в случае задачи TSP, возвращать такую длину пути, который нам **точно** понадобится пройти. Например, можно взять вес оставшегося дерева на оставшихся вершинах. Тогда $span \leq TSP \leq 2 \cdot span$. Хорошая функция оценки!

Кроме того, очень хорошо иметь нормальное приближение ответа $best$. То есть, изначально как-нибудь (отжигом, жадником, итд) найти неплохой ответ, чтобы перебор не шел в заранее ущербные шаги.

Ветви. Тут мы хотим сделать какую-то магию, чтобы перебирать ветки в правильном порядке. То есть, мы хотим запускать самые хорошие ветви в самом начале, особенно на первых слоях. Например, можно ввести оценочную функцию, и запускать перебор в порядке сортировки по этой оценочной функции. В задаче TSP, из физических соображений, можно запускать перебор сначала из ближайшей вершины к текущей. Также можно попытаться прыгнуть сразу на много уровней вниз, соптимизировав это какой-то простой динамикой ($dp_{mask, i, j}$ — наименьшая длина пути через всю маску, если мы начали в i и закончили в j) на маленьких подмножествах-клластерах. Тогда мы знали самый хороший способ взять какое-то подмножество, а тогда мы вместо 2^k ходов сделаем k ходов.

В ветвях есть много пространства для спекулятивных переборов. Например, можно идти в топ- k веток по оценочной функции.

Можно инициализировать решение для неспекулятивного перебора решением из спекулятивного перебора!

Решаем задачи на графе. Задачи о поиске максимальной клики, максимального независимого множества и минимального контролирующего множества сводятся друг к другу, поэтому если мы решили одну задачу (за какое-то $O(f(n))$), то и другие задачи решили. Будем решать задачу о поиске минимального вершинного покрытия.

Асимптотические оптимизации, которые не казались такими сначала. Задачу о вершинном покрытии можно решать за $O(2^n m)$: для каждой вершины решаем, берем мы ее или нет:

$$t(G, V) = t(G \setminus v, V \cup \{v\}) + t(G \setminus v, V)$$

Заметим, что если мы решили не брать какую-то вершину в ответ, то мы обязаны взять всех ее соседей. Тогда мы можем решать задачу на каком-то меньшем графе, а именно:

$$t(G, V) = t(G \setminus v, V \cup \{v\}) + t(G \setminus \{v \cup N(v)\}, V \cup N(v))$$

Рассмотрим максимальную степень вершины в графе $\max d$. Если, $\max d = 0$, то задача решается за $O(1)$. Тогда пусть $\max d \geq 1$. Таким образом, во втором случае число вершин уменьшается на два:

$$t_n = t_{n-1} + t_{n-2}$$

Получили решение за $O(\phi^n m)$, как оценка на числа Фибоначчи.

Давайте сделаем теперь $\max d \geq 2$ — когда у нас остались ребра, для каждого ребра возьмем одну вершину:

$$t_n = t_{n-1} + t_{n-3}$$

$$x^3 - x^2 - 1 = 0$$

Получили $O(1.47^n m)$.

Если $\max d = 2$, то граф разбивается на пути и циклы, в которых поиск вершинного покрытия тривиален. Аналогично получаем решение за $O(1.38^n m)$.

Перебор в антагонистических играх. Нам дано полное двоичное дерево четной глубины $2n$. Фиш-ка стоит в корне дерева. Каждый игрок на своем ходу перемещает фишку налево или направо. В каждом листе написано «0» или «1», причем мы не знаем значения в листах заранее, а можем только спрашивать у оракула (который не играет против нас, то есть неадаптивный).

Решение за 4^n запросов выглядит так: спросить про все листья, а потом посчитать на дереве динамику на выигрыш-проигрыш:

$$t_v = \begin{cases} 1 & \text{если } t_{v_l} = 0 \vee t_{v_r} = 0 \\ 0 & \text{иначе} \end{cases}$$

Это решение можно соптимизировать. Заметим, что если мы нашли переход из вершины в проигрышного сына, то второго сына можно не рассматривать, поэтому какие-то листья мы можем просто не посещать.

Введем t_n — матожидание количества посещенных листьев для глубины $2n$.

Для начала рассмотрим t_1 и все шесть конфигураций:

Листья	матожидание
[0, 0, 0, 0]	2
[0, 0, 0, 1]	2.5
[0, 1, 0, 1]	3
[0, 0, 1, 1]	2.5
[0, 1, 1, 1]	2.75
[1, 1, 1, 1]	3

Заметим, что за каждый спуск на два уровня вниз, мы рассматриваем в среднем не более трех детей. Тогда $t_n \leq 3 \cdot t_{n-1} \leq 3^n$

$\alpha\beta$ -отсечение. Проведем предыдущее рассуждение на минимаксной игре (это такая, где в листьях записаны числа, первый игрок хочет минимизировать итоговое число, а второй максимизировать). Введем параметры α и β — гарантии игроков. β — это минимальное число, которое первый игрок может себе гарантировать (то есть первый игрок знает, что не получит больше, чем β). Аналогично α — это максимальное число, которое гарантирует себе второй игрок). У нас должен соблюдаться инвариант $\alpha \leq \beta$, потому что остальные состояния неиграбельные. Как в них можно попасть по ходу перебора? Если в какой-то момент один из игроков сделает невыгодный для себя ход. Понятно, что при оптимальной игре обоих игроков они не будут делать невыгодные для себя ходы.

Как пересчитываются α и β ? В листе $\alpha = \beta = \text{get}(v)$. Если второй игрок может пойти в состояние со стоимостью x , то он может сделать $\alpha = \max(\alpha, x)$. Аналогично первый игрок будет уменьшать β . При этом эти гарантии будут переходить в сыновей вершины, но не в предков — в предки будут переходить только значение вершины (которое мы либо явно посчитали, либо вообще не считали, потому что состояние было неиграбельно).

На прошлой лекции мы показали, что паросочетание максимально \Leftrightarrow для него не существует дополняющей цепи.

Поиск максимального паросочетания

Ориентируем ребра парсоча справа-налево, а остальные - слева-направо. Тогда каждой удлиняющей цепи соответствует путь из левой свободной вершины в правую свободную. Путь можно искать за линейное время, получаем асимптотику поиска паросочетания за $O(nm)$ (поиск пути можно делать не больше n раз, так как в парсоче не больше n ребер).

Алгоритм Куна

Пробежимся по вершинам левой доли и попробуем найти удлиняющую цепь из текущей вершины.

Проблема: Если из какой-то вершины мы не смогли найти удлиняющую цепь, то потом можем найти.

- Пусть $r(A) = \max \text{ matching } (A, R)$, где $A \subset L$. $A_{i+1} = A_i \cup \{v_i\}$, $A_0 = \{v_0\}$. Докажем, что если $r(A_{i-1})$ найден алгоритмом Куна корректно, то и $r(A_i)$ будет найден корректно.

Понятно, что $r(A_{i-1}) \leq r(A_i) \leq r(A_{i-1}) + 1$.

- Если $r(A_{i-1}) = r(A_i)$, то $r(A_i)$ посчитается правильно

- Если $r(A_{i-1}) + 1 = r(A_i)$. Пусть M_i для $r(A_i)$ и M_{i-1} - парсоч для A_{i-1} . Тогда рассмотрим симметрическую разность M_i и M_{i-1} . Наш получившийся граф разился на пути и циклы, так как $|M_i| > |M_{i-1}|$, то есть удлиняющая цепь для M_{i-1} . Если ее конец не в v_i , то можем увеличить M_{i-1} - противоречие, значит есть удлиняющая цепь из v_i , а значит алгоритм Куна найдет ее и получит, что $r(A_i) = r(A_{i-1}) + 1$.

- Если какая-то вершина стала покрыта в течение алгоритма Куна, то она и останется покрытой.

Отсюда получим, что Куна находит лексикографически минимальный парсоч, то есть слева выбранные вершины будут лексикографически минимальными.

Теперь решим задачу: слева на вершинах есть неотрицательные веса, тогда хотим найти парсоч максимального веса (то есть максимизировать суммарный вес взятых вершин).

Переупорядочим вершины по весу, то есть туперь $w_1 \geq \dots \geq w_n \geq 0$. Теперь запустим Куна на левой доли. Тогда получим оптимальный парсоч. Пусть $p_1 < \dots < p_k$ - индексы 1 в этом числе, $p'_1 < \dots < p'_k$ - индексы взятых вершин в ответе, посчитанном Куном, p'_i - индексы взятых вершин в оптимальном ответе. Тогда $\exists i : p'_i < p_i$, но тогда $r(A_i)$ Куна посчитал неправильно.

Теперь добавим вершины и для правой доли. $a : L \rightarrow \mathbb{R}_+, b : R \rightarrow \mathbb{R}_+, w(uv) = a(u) + b(v)$.

Пусть M_1 - парсоч, построенный предыдущим алгоритмом для левой доли, M_2 - для правой доли. Рассмотрим $M_1 \cup M_2$ (в оригинале симметрическую разность, но в объединении просто будут циклы длины 2). Тогда граф разился на четные циклы и четные пути (иначе можно было увеличить парсоч). В циклах берем чередующиеся ребра (в том числе и циклах длины 2), в путях мы должны выбрать один из 2 концов, выбираем больший. Несложно понять, что полученный парсоч будем весом $w(M_1) + w(M_2)$.

Теорема Кёнига Размер максимального паросочетания в двудольном графе равен минимального вершинного покрытию.

Любое вершинное покрытие всегда не меньше любого паросочетания - на каждом ребре из парсоча должна быть хотя бы 1 вершина из вершинного покрытия.

Теперь предъявим какой-то парсоч равный какому-то вершинному покрытию. Ориентируем ребра парсоча справа-налево, а остальные - слева-направо. Запустим дфс из левой доли из вершин непокрытых паросочетанием ($L \setminus L(M)$, где $L(M)$ - множество покрытых вершин левой доли.) Тогда L^+ - множество вершин левой доли, до которых мы дошли, L^- - вершины левой доли, до которых не дошли, R^+ - вершины правой доли, до которых дошли, R^- - вершины правой доли, до которых не дошли. Тогда между L^+ и R^- нету ребер, так как там не может быть ребер не из парсоча (иначе можем дойти до R^- - ребра не из парсоча слева-направо), также не может быть ребер из парсоча, так как тогда конец этого ребра лежит в L^+ и покрыт, а значит мы туда попали, пройдя по этому ребру (так как у нас в каждую вершину левой доли может вести только ребро парсоча). (КЕК, УДАЧИ РЕБЯТА)

Тогда $L^- \cup R^+$ - вершинное покрытие. Все вершины L^- покрыты, так как все непокрытые сразу же в L^+ , все вершины R^+ покрыты, так как иначе есть удлиняющая цепь, а парсоч максимальный. Между L^- и R^+ не может быть ребер парсоча (справа налево), иначе можем дойти до L^- . Значит ребер парсоча хотя бы столько же, сколько

и вершин в вершинном покрытии, то есть мы нашли парсоч, мощность которого равна мощности вершинного покрытия. Отюда получаем, что $L^- \cup R^+$ - мин вершинное покрытие.
Отсюда $L^+ \cup R^-$ - максимальное независимое множество (как дополнение к мин вершинному покрытию).

Задача 1 Есть ДАГ (ориентированный ациклический граф). Надо покрыть все его вершины минимальным количеством вершинно-непересекающимися путями.

Рассмотрим выбранные ребра M , тогда путей будет $n - |M|$. Значит нам надо выбрать максимальное количество ребер так, чтобы исходящая степень каждой вершины была не больше 1, и входящая степень была не больше 1. Тогда построим двудольный граф: в левой и правой дали - по n вершин, тогда если в изначальном графе было ребро из i в j , то проведем ребро из i левой доли в j правой доли (неориентированное). Тогда парсоч и есть множество ребер M .

Задача 2 Теперь пути могут пересекаться. Рассмотрим транзитивное замыкание графа и решим предыдущую задачу. Тогда можно восстановить ответ из полученного (не успею немного дописать - там легко).

Теорема Дилвортса Размер максимальной антицепи равен количеству цепей в минимальном разбиении на цепи. (Понятно, что граф транзитивно замкнут по определению частично упорядоченного множества)

Мы уже умеем разбивать ЧУМ (частично упорядоченное множество) на минимальное количество цепей (*Задача 1*). Научимся строить отсюда максимальную антицепь. Рассмотрим паросочетание M из *Задачи 1*. Тогда $|VC| = |M|$, $|IS| = 2n - d$ в двудольном графе. Возьмем вершины, обе копии которых (из левой и правой доли) попали в $IS = L^+ \cup R^-$. Тогда получим антицепь (так как вершины лежат в IS). Теперь докажем, что полученная антицепь максимальна. Для этого докажем, что наши выбранные вершины (обе копии которых в IS) НЕ лежат в мин VC . Предположим обратное, тогда пусть обе копии какой-то вершины v попали в $\min VC$. По построению мин VC обе копии v покрыты парсочем, а концы соответствующих ребер парсоча (a и b) не лежат в $\min VC$. Но тогда $a \rightarrow v, v \rightarrow b$, но так как граф транзитивно замкнут, то $a \rightarrow b$, а значит это ребро не покрыто $\min VC$, противоречие.

DFS. Алгоритм обхода графа. Каждой вершине присваивается один из трех цветов — белый, серый или черный. Белые вершины мы еще не рассматривали, серые вершины мы рассматриваем сейчас, черные вершины мы больше не рассматриваем. Алгоритм примерно такой:

1. Покрасить текущую вершину в серый цвет.
2. Рекурсивно запуститься из всех белых вершин, соединенных с нашей.
3. Покрасить текущую вершину в красный цвет.

Что-то из важных утверждений про *dfs*:

- Серые вершины образуют путь.
- Ребро между двумя серыми несоседними вершинами существует тогда и только тогда, когда в графе есть цикл.
- Из черных вершин нет ребер в белые.

Топологическая сортировка. Пусть у нас есть ориентированный ациклический граф. Давайте выпишем такую перестановку вершин — p_v будет номером покраски вершины v в черный цвет. Тогда все наши ребра будут идти только справа налево.

Поиск компонент сильной связности и конденсация. Компонентой сильной связности называем такой класс эквивалентности, где v и u сильно связаны, если есть пути $u \rightarrow v$ и $v \rightarrow u$. Конденсацией мы назовем такой мета-граф, где все компоненты сильной связности «сжаты» в одну вершину, а ребра между новыми компонентами есть, если есть ребра между какой-то парой вершин из этих компонент. Такой граф уже точно будет ациклическим.

Сделаем процедуру, аналогичную топсорту, но теперь у нас ребра уже могут идти слева направо. Про самую последнюю вершину в новом порядке мы знаем, что она точно лежит в компоненте истока. Рассмотрим граф G' , построенный на обратных ребрах. Тогда если у нас есть ребро в G' справа налево $u \leftarrow v$, то это значит, что вершины u и v лежат в одной компоненте сильной связности. Тогда мы можем последовательно выделять КСС с помощью обхода по обратным ребрам.

2-sat. Автор опоздал на эту часть лекции, поэтому напишет ее сам позднее.

Мосты и точки сочленения. Назовем мостами такие ребра, при удалении которых граф теряет связность. Аналогичные вершины определим как точки сочленения. Как их найти? Сделаем dfs в неориентированном графе, и построим дерево dfs. Те ребра, по которым мы не переходили, мы назовем обратными. В дереве dfs эти ребра будут идти из вершины в ее какого-то предка.

Что тогда верно про мосты? Из поддерева ребра-моста нет ни одного обратного ребра, котороешло бы выше, чем наше ребро. А найти самое высокое ребро в поддереве можно обычной динамикой. Аналогично с точками сочленения.

Отношения вершинной и реберной двусвязности. Проще всего запомнить так — отношения вершинной двусвязности это отношение на ребрах, а отношение реберной — отношение на вершинах.

Вершины v и u находятся в одной компоненте реберной двусвязности, если существует реберно-простой цикл, содержащий u и v .

Отношение такой двусвязности транзитивно. Можно показать, что если $u \equiv v$ и $v \equiv w$ (при этом v и w соединены ребром), то и $u \equiv w$. Для этого надо склеить два цикла, и найти в них новый — из u в w .

Чтобы найти классы эквивалентности, можно удалить все мосты в графе.

Отношение вершинной двусвязности определяется аналогично, но на вершинах.

Задача о поиске кратчайшего пути. Пусть нам дан граф $G = (V, E)$. Каждому ребру задан какой-то вес, а весу пути соответствует суммарный вес всех ребер на пути.

Веса на ребрах могут быть естественными (неотрицательные), или не естественными (разрешаем отрицательный вес ребер). Также отдельный случай для нас — существование циклов отрицательного веса. Кроме того, мы иногда хотим искать любые пути, иногда реберно простые, иногда вершинно простые.

Bfs. Пусть мы хотим найти кратчайший путь, где вес каждого ребра равен 1. Идея такая — мы хотим построить слоистую декомпозицию, где уровню i соответствуют вершины на расстоянии i от стартовой. Алгоритм реализуется на очереди.

Кратчайшие расстояния в графе. В большинстве задач мы считаем, что циклов отрицательного веса нет, оптимальный путь простой, а на кратчайшие расстояния накладывается неравенство треугольника.

Алгоритм Форда-Беллмана. Задача — найти кратчайшие расстояния из s до всех вершин графа. Мы хотим на каждом шаге перебирать все ребра, и поддерживать инвариант — для шага i мы считаем, что алгоритм нашел кратчайшие расстояния среди всех путей длины i . Тогда на следующем шаге мы рассмотрим все ребра, и продолжим старые пути новыми ребрами, сделаем релаксации.

Циклы отрицательного веса. Заметим, что наш алгоритм не будет делать релаксации после шага n , потому что больше будет нечего релаксировать. Но в случае с циклами отрицательного веса это не так — мы знаем, что на каждом шаге после n -го мы будем делать релаксацию, проходя по улучшающему ответ циклу. Тогда, если мы для каждой релаксации запомнили, какая вершина улучшала наш текущий ответ, мы можем восстановить цикл (или оптимальный путь), просто проходясь по предкам.

Алгоритм Флойда. Задача — найти матрицу кратчайших расстояний. Пусть $f_k(u, v)$ — это кратчайшее расстояние между u и v , если в качестве промежуточных вершин разрешено использовать вершины $1, \dots, k$. Тогда при переходе $f_k(u, v) \rightarrow f_{k+1}(u, v)$ мы хотим сделать релаксацию пути как $u \rightarrow k + k \rightarrow v$.

Алгоритм Дейкстры. В реальных задачах редко присутствуют ребра отрицательного веса, и в таких ситуациях часто используется алгоритм Дейкстры. Инвариант алгоритма — известны кратчайшие расстояния от s до вершин множества A . Мы знаем, что $\max p(v) \leq \min p(u), v \in A, u \in V \setminus A$. Давайте для всех вершин поддерживать какое-то текущее $p(v)$. Понятно, что для вершин из A это будет итоговым ответом. Утверждается, что в A можно добавить вершину с минимальным текущим $p(v)$. Доказательство от противного — если текущее расстояние до v было не кратчайшим, то в v был путь, проходящий через вершину $u \in V \setminus A$, но тогда $p(u) < p(v)$, противоречие.

Тогда мы хотим добавлять вершины в A по очереди, по ходу обновляя значения $p(v)$ для всех соседей добавляемой вершиной. Тогда нам нужна структура данных, которая сделает n извлечений минимума и m уменьшений ключа. Оптимальная структура (асимптотически!) — фибоначчиева куча, с ней алгоритм работает за $O(n \log n + m)$.

Потенциалы. Создадим новый граф, где каждой вершине зададим потенциал c_v . Теперь за вес ребра примем $w_c(u, v) = w(u, v) + c_v - c_u$. То есть, вес ребра плюс разность потенциалов. Тогда длина пути $s \rightarrow t$ в новом графе определяется как длина в старом графе с дополнительным слагаемым $c_t - c_s = \text{const}$.

Задача о размещении потенциалов эквивалентна поиску отрицательного цикла. То есть, если цикл отрицательного веса есть, то не существует потенциалов, для которых выполняется неравенство треугольника (например, вес такого цикла в новом графе будет меньше нуля). Если же циклов нет, то сделаем $c_v = p(s, v)$. Тогда $w_p(u, v) = w(u, v) - p(v) + p(u)$, но $p(v) \leq p(u) + w(u, v)$, а тогда вес новых ребер положителен, а неравенство треугольника выполнено.

A-star. Пусть мы хотим решить задачу на плоскости, при этом хочется не посещать заведомо бесполезные состояния. Введем какую-то метрику $d(u, v)$, в качестве которой введем евклидово расстояние. Теперь мы вытаскиваем новые вершины по минимуму $p(u) + w(u, v) + d(v, t)$ (то есть минимальная сумма расстояния и оценки на метрику). Здесь t — это вершина, кратчайший путь в которую мы пытаемся найти. Поскольку на метрику введено неравенство треугольника, то наша новая Дейкстра все еще работает.

Двусторонняя Дейкстра. Аналогично обычной Дейкстры, запускаем алгоритм в две стороны, и обновляем ответ, если видим ребро $u \leftrightarrow v$, $v \in B$, $u \in A$.

Остовные деревья. Пусть нам дан взвешенный граф. Мы хотим оставить в нем подмножество ребер минимального суммарного веса так, чтобы граф оставался связным.

Лемма о безопасном ребре: Для подмножества вершин A есть ребро наименьшего веса, ведущее в дополнение к A . А именно, ведет из $u \in A$ в $v \notin A$. Существует остов, содержащий это ребро. Доказательство: От противного. Пусть мы не взяли ребро $u \leftrightarrow v$. Посмотрим на путь из u в v в дереве. В какой-то момент там нашлось ребро из A в $V \setminus A$. Тогда можно удалить это ребро и заменить его на $v \leftrightarrow u$ — стоимость не увеличится.

Алгоритм Прима. По аналогии с алгоритмом Дейкстры будем поэтапно строить множество A и добавлять в него минимальное ребро из леммы. Но тут важно, что лемма не гарантирует нам, что итоговый ответ хороший, потому что она говорила нам только о том, что ребро принадлежит какому-то ответу. Но есть усиленная версия леммы — если ребро меньше по весу, чем все остальные, то тогда оно обязательно будет в минимальном остове. Поэтому надо как-то неявно задать веса ребер, чтобы они были различны (это в реальной жизни не нужно, только для доказательства), после чего алгоритм Прима нам уже будет гарантировать минимальный ответ.

Система непересекающихся множеств. Структура данных, которая поддерживает следующие операции:

- $get(v)$ — узнать, в каком множестве находится элемент v
- $union(v, u)$ — объединить множество, содержащее v с множеством, содержащим u .
- $check(v, u)$ — проверить, находятся ли элементы в одном и том же множестве. Выражается через два get -а.

Алгоритм Краскала. В начале упорядочим ребра по весу. Будем поддерживать какое-то множество компонент связности. Если для очередного ребра компоненты вершин различны, то можно сделать $union$ в dsu . Работает за $O(sort(m) + union(n) \cdot n + check(n) \cdot m)$

Алгоритм Борувки. Выделим для каждой вершины минимальное ребро. Теперь добавим все эти ребра в остов, и сожмем граф. Дальше сделаем аналогичную операцию. Повторяем, пока не сойдется до одной компоненты. Каждый раз вершины соединяются в компоненты хотя бы по две вершины. Это значит, что итоговая оценка сложности будет $O(m \log n)$. Заметим, что на плотных графах Борувка работает за $O(m)$ — каждый раз число ребер уменьшается в два раза.

Чтобы объединить две асимптотики, получим:

$$\frac{n^2}{4^k} \leq m$$

$$4^k \geq \frac{n^2}{m}$$

$$k \approx \frac{1}{2} \log \frac{n^2}{m}$$

Значит, через k операций мы получим граф, который можем считать полным (а на нем мы работаем за линию!). Тогда асимптотика это $O(m \log \frac{n^2}{m})$.

Disjoint set union. Хотим реализовать следующий интерфейс:

- $get(x)$ — найти компоненту, в которой лежит x
- $unite(x, y)$ — объединить две компоненты
- $check(x, y)$ — проверить, что две вершины лежат в одной компоненте

$check(x, y)$ обычно выполняется как проверка $get(x) = get(y)$.

Под капотом мы будем хранить ориентированный лес. А именно, из каждой вершины будет идти ребро в предка. Компоненту будем идентифицировать номером корня в соответствующем дереве.

Как делать $get(x)$? Подниматься по ребрам $x \rightarrow parent_x \rightarrow \dots \rightarrow root$. Тогда $get(x) = root$

Для $unite(x, y)$ надо подвесить корень одной компоненты к какой-то вершине из другой компоненты.

Эвристики. Понятно, что в наивной реализации очень легко достигается время работы $O(n)$ на запрос. Но существует две легкие эвристики, которые значительно улучшают время работы. Формально, они являются не эвристиками, а просто оптимизациями алгоритма.

Эвристика размеров. При операции $unite$ будем подвешивать корень одного дерева к корню другого дерева (это значит, что нам понадобится в начале сделать две операции get). При этом мы будем подвешивать корень меньшего дерева к большему. Таким образом, любой подъем в get по ребру увеличивает размер компоненты в два раза. Мы получаем оценку $O(\log n)$ на запрос.

Эвристика рангов. Зададим каждой вершине ранг r_x . При подвешивании будем сравнивать корни не по размерам, а по рангам — к большему подвешивать меньший. Если два ранга совпали, то после подвешивания ранг корня увеличивается на 1. Тогда количество вершин с рангом x не больше, чем $\frac{n}{2^x}$. Это можно показать по индукции. Поскольку вершина с рангом $x+1$ получается как комбинация двух вершин с рангом x (а вершины с рангом x после этого становятся неактивными), то $A_{x+1} \leq \frac{A_x}{2}$, где A_x — число вершин ранга x . Считаем, что ранг одной вершины — 0.

Очевидное утверждение — ранг вдоль пути к предку возрастает. Поскольку рангов не больше, чем $O(\log n)$, то мы опять получаем оценку $O(\log n)$ на запрос.

Эвристика сжатия путей. Идея эвристики такая: если мы прошли по пути до корня, то мы для каждой вершины на этом пути узнали текущую компоненту. Это значит, что можно запомнить эту информацию и не подниматься в промежуточные вершины в дальнейшем — подниматься сразу в текущий корень компоненты. Утверждается, что эта эвристика без какой-либо эвристики на $unite$ работает амортизированно за $O(\log n)$ на запрос.

Как такую оценку показать? Разобьем ребра на легкие и тяжелые. Ребро называется *тяжелым*, если на нем висит хотя бы половина поддерева. Подняться по легкому ребру больше логарифма раз мы не можем. Утверждается, что суммарных проходов по тяжелым ребрам будет $O(n \log n)$.

Каждый раз, когда мы проходим по тяжелому ребру, оно становится легким (кроме, может быть, ребра в корень). Иногда легкие ребра опять становятся тяжелыми.

Будем рассматривать все вершины, кроме корня (все ребра, связанные с корнем, работают за $O(1)$, поэтому нам не важно). Для оставшихся вершин верно, что они только теряют тяжелых сыновей. При этом мы можем считать, что новые вершины не появляются — пририсовывается что-либо только к корню, а наша вершина уже никогда не будет корнем. Тогда к v присоединено не более $O(\log n)$ тяжелых ребер за все время ее существования. Это верно, потому что при удалении тяжелого ребра (переподвешивания его к корню), суммарный размер поддерева уменьшается в два раза. Таким образом, верна оценка $O(n \log n)$ на число тяжелых ребер, которые мы суммарно рассмотрим.

Объединение эвристик. Утверждается, что при объединении ранговой эвристики и эвристики сжатия путей достигается время $O(n\alpha)$. Это какое-то сложное доказательство, которое (возможно) будет потом. Мы покажем $O(n \log^* n)$.

Что такое $\log^* n$? Это функция, обратная к $f(a, b) = \begin{cases} 1 & , b = 0 \\ a^{f(a, b-1)} & \text{else} \end{cases}$. На всех тестовых данных, влезающих в современный компьютер, эта функция примерно равна четырем (кстати, обратная функция Аккермана тоже равна 4 на адекватных тестовых данных, — прим. автора).

Назовем ребро **крутым** (или **жестким**), если проход по ребру увеличивает ранг экспоненциально. А именно, $c^{r(v)} \leq r(\text{parent}_v)$ для какой-то c . Тогда, если мы пройдем по пути с a крутыми ребрами, то ранг будет не меньше, чем $f(c, a)$. А значит, количество крутых ребер $O(\log^* n)$.

Действия разбиваются на три типа:

1. Проходы по крутым ребрам — $O(\log^* n)$ на запрос амортизированно.
2. Проходы в корень — $O(1)$ на запрос.
3. Проходы по не крутым ребрам — сейчас докажем линейную амортизированную оценку.

Сколько раз нужно пройти по не крутым ребру, чтобы оно стало крутым? Каждый раз, когда мы переподвешиваем вершину v от предка parent_v к какой-то другой вершине, мы увеличиваем ранг предка. При этом ранг v зафиксирован. Это значит, что проходов из каждой вершины по не крутым ребру будет

не более чем $c^{r(v)}$. Тогда, просуммировав по всем вершинам, получаем $\sum_{x=0}^{\log n} A_x \cdot c^x \leq \sum_{x=0}^{\log n} \frac{nc^x}{2^x} \leq n \frac{1}{1 - \frac{c}{2}}$.

Это будет верно при таких c , которые меньше, чем 2 (чтобы мы получили убывающую геометрическую прогрессию). При этом $f(c, a)$ все еще должна уходить в бесконечность. $c = 1.9$ подойдет. Таким образом, число проходов по не крутым ребрам будет линейно.

Задача о двудольном максимальном паросочетании. Пусть нам дан двудольный граф. Это такой граф, в котором вершины делятся на два класса, а ребра проводятся только между вершинами разных классов. Паросочетанием называется такое множество ребер, что никакие два ребра из этого множества не смежны. Максимальным по включению паросочетанием называется такое паросочетание, которое нельзя дополнить никаким ребром (которое, аккуратно, не является максимальным). Двудольной кликой называется два подмножества вершин (левой и правой доли), чтобы между любой парой вершин разных подмножеств было ребро.

Удлиняющие цепочки. Удлиняющей цепочкой называется чередующийся путь из свободной вершины левой доли в свободную вершину правой доли. А именно, все нечетные ребра на пути не принадлежат паросочетанию, а четные, наоборот, принадлежат.

Очевидно, что если есть удлиняющая цепочка, то паросочетание можно увеличить — заменить статус всех ребер на противоположный. Количество ребер в паросочетании увеличится на единицу.

Оказывается, если удлиняющих цепочек нет, то паросочетание максимально. Рассмотрим такое M , а также M^* , для которого верно, что $|M^*| > M$. То есть, одно паросочетание не содержит удлиняющих цепочек, а другое больше по размеру.

Посмотрим на симметрическую разность этих множеств. То есть, на те ребра, которые принадлежат только одному из множеств. В таком графе степени вершин до двух. Тогда этот граф состоит из циклов четной длины и путей. В цикле четной длины одинаковое количество ребер из M и M^* . На путях четной длины количество ребер из M и M^* , аналогично, совпадает. Теперь посмотрим на нечетные пути. Ребер

одного из двух типов на нем будет больше. Значит, на каком-то из путей, ребер из M^* будет больше, чем из M . Но такой путь обязательно будет чередующейся цепочкой! Получаем противоречие.

Алгоритм Куна. Мы не обсудили, но если на следующей лекции его не будет, то я напишу

На прошлой лекции мы показали, что паросочетание максимально \Leftrightarrow для него не существует дополняющей цепи.

Поиск максимального паросочетания

Ориентируем ребра парсоча справа-налево, а остальные - слева-направо. Тогда каждой удлиняющей цепи соответствует путь из левой свободной вершины в правую свободную. Путь можно искать за линейное время, получаем асимптотику поиска паросочетания за $O(nm)$ (поиск пути можно делать не больше n раз, так как в парсоче не больше n ребер).

Алгоритм Куна

Пробежимся по вершинам левой доли и попробуем найти удлиняющую цепь из текущей вершины.

Проблема: Если из какой-то вершины мы не смогли найти удлиняющую цепь, то потом можем найти.

- Пусть $r(A) = \max \text{ matching } (A, R)$, где $A \subset L$. $A_{i+1} = A_i \cup \{v_i\}$, $A_0 = \{v_0\}$. Докажем, что если $r(A_{i-1})$ найден алгоритмом Куна корректно, то и $r(A_i)$ будет найден корректно.

Понятно, что $r(A_{i-1}) \leq r(A_i) \leq r(A_{i-1}) + 1$.

- Если $r(A_{i-1}) = r(A_i)$, то $r(A_i)$ посчитается правильно

- Если $r(A_{i-1}) + 1 = r(A_i)$. Пусть M_i для $r(A_i)$ и M_{i-1} - парсоч для A_{i-1} . Тогда рассмотрим симметрическую разность M_i и M_{i-1} . Наш получившийся граф разился на пути и циклы, так как $|M_i| > |M_{i-1}|$, то есть удлиняющая цепь для M_{i-1} . Если ее конец не в v_i , то можем увеличить M_{i-1} - противоречие, значит есть удлиняющая цепь из v_i , а значит алгоритм Куна найдет ее и получит, что $r(A_i) = r(A_{i-1}) + 1$.

- Если какая-то вершина стала покрыта в течение алгоритма Куна, то она и останется покрытой.

Отсюда получим, что Куна находит лексикографически минимальный парсоч, то есть слева выбранные вершины будут лексикографически минимальными.

Теперь решим задачу: слева на вершинах есть неотрицательные веса, тогда хотим найти парсоч максимального веса (то есть максимизировать суммарный вес взятых вершин).

Переупорядочим вершины по весу, то есть туперь $w_1 \geq \dots \geq w_n \geq 0$. Теперь запустим Куна на левой доли. Тогда получим оптимальный парсоч. Пусть $p_1 < \dots < p_k$ - индексы 1 в этом числе, $p'_1 < \dots < p'_k$ - индексы взятых вершин в ответе, посчитанном Куном, p'_i - индексы взятых вершин в оптимальном ответе. Тогда $\exists i : p'_i < p_i$, но тогда $r(A_i)$ Куна посчитал неправильно.

Теперь добавим вершины и для правой доли. $a : L \rightarrow \mathbb{R}_+, b : R \rightarrow \mathbb{R}_+, w(uv) = a(u) + b(v)$.

Пусть M_1 - парсоч, построенный предыдущим алгоритмом для левой доли, M_2 - для правой доли. Рассмотрим $M_1 \cup M_2$ (в оригинале симметрическую разность, но в объединении просто будут циклы длины 2). Тогда граф разился на четные циклы и четные пути (иначе можно было увеличить парсоч). В циклах берем чередующиеся ребра (в том числе и циклах длины 2), в путях мы должны выбрать один из 2 концов, выбираем больший. Несложно понять, что полученный парсоч будем весом $w(M_1) + w(M_2)$.

Теорема Кёнига Размер максимального паросочетания в двудольном графе равен минимального вершинного покрытию.

Любое вершинное покрытие всегда не меньше любого паросочетания - на каждом ребре из парсоча должна быть хотя бы 1 вершина из вершинного покрытия.

Теперь предъявим какой-то парсоч равный какому-то вершинному покрытию. Ориентируем ребра парсоча справа-налево, а остальные - слева-направо. Запустим дфс из левой доли из вершин непокрытых паросочетанием ($L \setminus L(M)$, где $L(M)$ - множество покрытых вершин левой доли.) Тогда L^+ - множество вершин левой доли, до которых мы дошли, L^- - вершины левой доли, до которых не дошли, R^+ - вершины правой доли, до которых дошли, R^- - вершины правой доли, до которых не дошли. Тогда между L^+ и R^- нету ребер, так как там не может быть ребер не из парсоча (иначе можем дойти до R^- - ребра не из парсоча слева-направо), также не может быть ребер из парсоча, так как тогда конец этого ребра лежит в L^+ и покрыт, а значит мы туда попали, пройдя по этому ребру (так как у нас в каждую вершину левой доли может вести только ребро парсоча). (КЕК, УДАЧИ РЕБЯТА)

Тогда $L^- \cup R^+$ - вершинное покрытие. Все вершины L^- покрыты, так как все непокрытые сразу же в L^+ , все вершины R^+ покрыты, так как иначе есть удлиняющая цепь, а парсоч максимальный. Между L^- и R^+ не может быть ребер парсоча (справа налево), иначе можем дойти до L^- . Значит ребер парсоча хотя бы столько же, сколько

и вершин в вершинном покрытии, то есть мы нашли парсоч, мощность которого равна мощности вершинного покрытия. Отсюда получаем, что $L^- \cup R^+$ - мин вершинное покрытие.
Отсюда $L^+ \cup R^-$ - максимальное независимое множество (как дополнение к мин вершинному покрытию).

Задача 1 Есть ДАГ (ориентированный ациклический граф). Надо покрыть все его вершины минимальным количеством вершинно-непересекающимися путями.

Рассмотрим выбранные ребра M , тогда путей будет $n - |M|$. Значит нам надо выбрать максимальное количество ребер так, чтобы исходящая степень каждой вершины была не больше 1, и входящая степень была не больше 1. Тогда построим двудольный граф: в левой и правой дали - по n вершин, тогда если в изначальном графе было ребро из i в j , то проведем ребро из i левой доли в j правой доли (неориентированное). Тогда парсоч и есть множество ребер M .

Задача 2 Теперь пути могут пересекаться. Рассмотрим транзитивное замыкание графа и решим предыдущую задачу. Тогда можно восстановить ответ из полученного (не успею немного дописать - там легко).

Теорема Дилвортса Размер максимальной антицепи равен количеству цепей в минимальном разбиении на цепи.
(Понятно, что граф транзитивно замкнут по определению частично упорядоченного множества)

Мы уже умеем разбивать ЧУМ (частично упорядоченное множество) на минимальное количество цепей (*Задача 1*). Научимся строить отсюда максимальную антицепь. Рассмотрим паросочетание M из *Задачи 1*. Тогда $|VC| = |M|$, $|IS| = 2n - d$ в двудольном графе. Возьмем вершины, обе копии которых (из левой и правой доли) попали в $IS = L^+ \cup R^-$. Тогда получим антицепь (так как вершины лежат в IS). Теперь докажем, что полученная антицепь максимальна. Для этого докажем, что наши выбранные вершины (обе копии которых в IS) НЕ лежат в мин VC . Предположим обратное, тогда пусть обе копии какой-то вершины v попали в мин VC . По построению мин VC обе копии v покрыты парсочем, а концы соответствующих ребер парсоча (a и b) не лежат в мин VC . Но тогда $a \rightarrow v, v \rightarrow b$, но так как граф транзитивно замкнут, то $a \rightarrow b$, а значит это ребро не покрыто мин VC , противоречие.

Алгоритмы на графах во внешней памяти Напомним, что мы хотим делать алгоритмы с менее, чем линейным числом обращения к памяти. Например, сложность $O(\frac{n \cdot \text{Poly}(\log n)}{B})$ нас устраивает, а сложность $O(n)$ нас категорически не устраивает. Всякие $O(\frac{\text{Poly}(n)}{\sqrt{B}})$ или $O(\frac{\text{Poly}(n)}{\log B})$ нам не нравятся, но иногда мы не можем лучше.

Dfs. Если мы смогли сохранить ребра так, что для каждой вершины ребра лежат в памяти последовательно, то мы можем доставать соседей v блоками. Но, к сожалению, это не делает сложность алгоритма хорошей, потому что у нас есть массив $used$, к которому нам придется обращаться. Обращения будут произвольными. Тогда, если $n < M$, то алгоритм имеет смысл, потому что мы можем сохранить массив в оперативной памяти, а иначе этот алгоритм работает за $O(n + \frac{m}{B})$.

Работа со списками. Имеем какой-то односвязный список. Он реализован как массив и указатели. То есть, в каждой ячейке массива есть указатель на следующую ячейку массива. Научимся с ним работать.

Групповые операции. Чтобы считать несколько значений из списка, мы можем либо спросить про каждый элемент запроса, либо про все элементы списка — $O(q)$ или $O(\frac{n}{B})$. Чтобы сделать групповые изменения, можно отсортировать все запросы заранее, после чего обрабатывать запросы группами за линейное время блоками. Сложность будет равняться $O(\text{sort}(q, M, B) + \frac{n+q}{B})$.

Задача list ranking. Мы хотим для каждого элемента узнать его порядковый номер в соответствующем списке. Это тривиально делается в оперативной памяти, потому что мы можем просто пройти по списку. Но для внешней памяти алгоритм не подойдет — прыжки по памяти произвольны.

Во-первых, мы хотим сделать список двусвязным. Это то же самое, что и набор запросов «Присвой следующему элементу мой индекс в поле предка». Групповой запрос на изменение мы уже научились обрабатывать.

Теперь сделаем что-то типа двоичных подъемов. На итерации k считаем, что мы правильно посчитали ранги для всех вершин, у которых ранг не более 2^k . Кроме этого, будем поддерживать для всех вершин указатель на вершину на расстоянии 2^k от нее (и вперед, и назад). Как задать новые ранги? Для тех вершин, у которых уже известны ранги, есть суперссылка в вершину без ранга. Но тогда ее ранг — это ранг старой вершины, увеличенный на 2^k . Осталось пересчитать суперссылки. Их можно пересчитывать двоичными подъемами — брать суперссылку от нашей суперссылки. Но у нас доступна в произвольную точку памяти. Поэтому мы сделаем так: пусть у нас есть вершина v , ее суперпредок u , и ее суперсын w . Тогда новый суперсын от u это w , а суперпредок от w это u . Это тоже операция группового присваивания.

Пусть у нас был оракул, который умел говорить, четный или нечетный ранг для элемента. Пусть мы выкинули из списка все нечетные элементы, а указатели на следующий элемент списка теперь указывают на элемент «через один» от нас. Например, список $1 \rightarrow 3 \rightarrow 4 \rightarrow 2$ перейдет в $1 \rightarrow 2$. Мы уменьшили длину списка в два раза. Пусть мы посчитали ранги в четном списке. Тогда мы на самом деле знаем ранги и для нечетного списка тоже — это просто ранг предка, увеличенный на единицу. Таким образом, мы получим сложность $t(n) = t(\frac{n}{2}) + \text{sort}(n, M, B)$.

Как создать такого оракула? Мы сделаем его недетерминированным. А именно, для каждого элемента случайно решим, выкинем мы его или нет. Но нам нужно, чтобы для элемента остался в новом списке либо он, либо его предок. Тогда мы выкидываем элемент, только если мы случайно решили, что хотим выкинуть его, и случайно решили, что мы не хотим выкинуть его предка. Так мы отбросим четверть элементов. Получаем сложность $t(n) = t(\frac{3n}{4}) + \text{sort}(n, M, B)$. Так мы можем проталкивать величину «чему равен номер моего списка».

Но поскольку у нас нумерация сместилась произвольным образом, ранги пока что пересчитываются не так просто. Раньше мы просто умножали ранг на 2 и прибавляли 1, но сейчас мы не выкидывали некоторые элементы.

Пусть рекурсия вернула наш новый список с рангами. Теперь мы для каждого нерассмотренного элемента знаем, чему должен быть равен ранг относительно предка. Но ранги предков сейчас поедут. Предварительно скажем, что ранг новых вершин — это ранг предка, увеличенный на 1. После чего мы получили два массива рангов — для старых вершин и для новых. Теперь нам надо сделать линейный проход, и при встрече равных рангов сначала брать старую вершину, а потом увеличивать все ранги на суффиксе на 1 (каким-нибудь счетчиком). По сути, мы просто избавились от равных чисел за сложность сортировки. Таким образом, мы решили list ranking за сложность $O(sort(n, M, B))$.

Поиск минимального остова во внешней памяти с помощью алгоритма Борувка. Отсортируем ребра по первой вершине. Теперь за $O(\frac{n}{B})$ мы находим минимальное ребро из каждой вершины. Теперь нам надо сделать самый сложный шаг — найти компоненты.

У нас есть какое-то множество ребер, которое задает лес. Мы хотим для каждой вершины узнать номер ее компоненты. Тогда мы потом сожмем компоненты, уменьшим размер в два раза, и тд (см. алгоритм Борувки). Разобъем каждое ребро на два ориентированных, и для каждой вершины зададим каждому ребру следующее, как в эйлеровом обходе. Теперь мы получаем списки ребер, и нам надо для каждого ребра просто понять номер его списка. Видим, что это и есть задача list ranking. Таким образом, мы сможем найти остов за $O(sort(n, M, B) \log n)$.

Потоки. Пусть нам дан некий ориентированный граф. У каждого ребра есть некоторая пропускная способность $c_{v,u}$ (можем полагать, что если ребра нет, то она равна нулю). Две вершины s, t обозначены как сток и исток.

Далее мы выделяем какое-то множество путей $P = \{p_1(s,t), p_2(s,t), \dots\}$. Каждому пути сопоставляем число $f_i \geq 0$. Требуется, чтобы для каждого ребра e сумма f_i по всем вхождениям ребра в пути не превосходила c_e . При таких условиях мы максимизируем $\sum f_i$.

Классический вид задачи. Обозначим потоком функцию от пары вершин $f : V \times V \rightarrow \mathbb{R}$, на которую накладываются следующие требования:

1. $f(v, u) \leq c(v, u)$
2. $f(v, u) = -f(u, v)$
3. $\forall v \in V \setminus \{s, t\} : \sum_u f(v, u) = 0$

Мы хотим максимизировать $|f| = \sum_u f(s, u) = \sum_u f(u, t)$.

Можно заметить, что можно определить сумму потоков $f_1 + f_2$, если выполняется пункт 1. Остальные свойства сохраняются.

Перевод из первой постановки задачи во вторую делается просто: вдоль каждого пути по прямым ребрам добавляем $+f_p$, а вдоль обратных ребер прибавляем $-f_p$. Тогда все три свойства выполняются, и все хорошо.

Определим $e_v = \sum_u f(v, u)$. Если мы выделим какое-то множество вершин, то поток, проходящий через множество равен сумме e_v . Можно заметить, что такое число это что-то из $\{0, |f|, -|f|\}$.

Ребро называется насыщенным если $f(v, u) = c(v, u)$. Остаточной пропускной способностью ребра $c^*(v, u)$ назовем $(c-f)(v, u)$. Остаточной сетью назовем граф на множестве ребер, для которых $c^*(v, u) \neq 0$.

Разрез. Это такое разбиение вершин на множества A и \bar{A} . Разрезы бывают ориентированные и неориентированные. В потоках мы почти всегда говорим про s/t -разрез — $s \in A, t \notin A$, стоимость разреза определяем как $\sum_{v \in A} \sum_{u \notin A} c(v, u)$.

Разрез называется насыщенным, если все его ребра насыщенные.

Декомпозиция потока. Поток можно представить в виде суммы $l \leq m$ элементарных потоков (то есть путей и циклов). Доказательство конструктивное: пока общий поток ненулевой, из s есть ребро, где $f > 0$. Проходя по нему, дальше тоже будет ребро с положительным потоком, итд. Получим путь $s \rightarrow t$, где можно уменьшить все f на минимальное из потоков по ребрам. Если встретили цикл, то уменьшаем цикл, и продолжаем. Поскольку каждый раз добавлялось одно ребро с $f = 0$, то всего мы выделим не более m элементарных потоков. В какой-то момент путей $s \rightarrow t$ не останется. Тогда в сети остались только циклы. Их можно выделить как элементарные потоки. Если нас интересует сведение второй постановки задачи к первой, то на циклы можно просто забить.

Размер декомпозиции можно оценить как $O(nm) - O(m)$ итераций, и $O(n)$ вершин в элементарном потоке. Найти за это время декомпозицию тоже можно. Для этого надо хранить указатель на текущее интересное исходящее ребро для каждой вершины. Тогда наш указатель либодвигается вправо (суммарно $O(m)$), либо мы находим какой-то новый элементарный поток. То есть, каждый шаг декомпозиции

работает за $O(n)$ переходов по ребру, и все сдвиги указателя суммарно работают за $O(m)$ по всем шагам декомпозиции, получаем сложность $O(nm)$.

Теорема Форда-Фалкерсона. Теорема состоит из нескольких утверждений:

1. Максимальный поток равен минимальному разрезу
2. Поток максимальен, если не существует пути $s \rightarrow t$ в остаточной сети

Давайте покажем первое утверждение. Для этого сначала скажем, что любой поток не больше любого разреза. Это верно, потому что размер разреза это сумма $\sum c_{v,u}$, а поток это $|f| = \sum f_{v,u}$ по $v \in A, u \notin A$. Теперь мы предъявим разрез, размер которого равен размеру какого-то потока — тогда мы одновременно найдем и минимальный разрез, и максимальный поток.

Пусть мы взяли какой-то существующий поток, в остаточной сети которого не было пути $s \rightarrow t$ (условие п.2). Тогда возьмем разрез, образованный насыщенными ребрами — то есть возьмем в множество A все вершины, достижимые из s в остаточной сети. Тогда для всех ребер разреза $c_{v,u} = f_{v,u}$, то есть вес разреза совпал с величиной потока. Значит, такой поток максимальен, а соответствующий разрез минимальный.

Алгоритм Форда-Фалкерсона. Самый простой алгоритм для поиска потока — найти произвольный путь в остаточной сети и пустить поток вдоль него. Работает данный алгоритм за $O(|f| \cdot E)$, что не очень хорошо. Стоит заметить, что алгоритм сохраняет дробность решений — то есть поток на целых чисел будет целым числом, для четных целых — четным целым (потому что происходят только операции взятия минимума и вычитания).

Алгоритм Эдмондса-Карпа. Идея алгоритма в том, чтобы дополнять поток вдоль кратчайшего по ребрам пути из s в t . Такой алгоритм имеет уже полиномиальную сложность. По сути, мы дополняем поток вдоль пути, который находится поиском в ширину. Для доказательства времени работы изобразим слоистую сеть, которую найдет bfs (то есть сгруппированные по расстояниям вершины). У нас есть ребра внутри одной компоненты, ребра в следующую, и ребра на сколько угодно назад. Утверждается, что при насыщении вдоль пути расстояние от s до v не уменьшается. Это верно, потому что при обновлении сети у нас удаляются какие-то ребра вдоль пути, и добавляются обратные. Следовательно, при старом разбиении на классы у нас не появляется ребер на два и более классов вправо, а значит расстояния могут только увеличиться.

На каждой итерации алгоритма насыщается какое-то ребро. Надо заметить, что оно может насыщаться еще раз, но для этого сначала надо пустить поток по обратному ребру. Сколько раз это можно сделать? $O(V)$. Потому что мы насыщаем вдоль ребра слева направо (из d в $d + 1$). Тогда, поскольку вершины двигаются только вправо, то не получится пустить по ребру $v \rightarrow u$ поток раньше, чем переместив v и u на слой вправо. А всего вершины могут быть удалены не более, чем на $O(V)$. Тогда мы делаем $O(VE)$ итераций, на каждой из которых вызываем bfs, получая оценку $O(VE^2)$.

Алгоритм Диница. Идея алгоритма Диница похожа на нашу предыдущую идею в декомпозиции потока. Если у нас уже есть зафиксированная слоистая сеть, то можно пускать поток вдоль пути, пока путь находится по ребрам между слоями. Тогда мы будем искать пути, пока слоистая сеть не потеряет связность, потом перестроим слоистую сеть, и так далее. Перестроек будет $O(V)$, потому что вершина t каждый раздвигается хотя бы на слой вправо. На каждой итерации нам надо найти все пути с насыщениями. Насыщений, как можно заметить, не более $O(E)$. Хочется много раз запустить dfs, который найдет все пути. При этом если каждый раз запускать dfs заново, то мы получим сложность $O(E^2)$. Но если запоминать указатель на первое нерассмотренное ребро, то сложность получится $O(VE)$. Почему? Потому что мы будем либо совершать неудачную попытку и двигать указатели к следующему

в списке элементу, либо найдем путь длины $O(V)$. Операций первого рода суммарно будет как раз $O(E)$ на все запуски.

Масштабирование. Давайте попытаемся улучшить асимптотику другой техникой — будем сначала искать большие потоки, а потом маленькие. А именно, давайте по очереди искать пути, по которым можно пропустить хотя бы 2^k единиц потока. Тогда на каждой итерации у нас будет не более $2^k \cdot E$ единиц потока, при этом мы будем находить пути с потоком хотя бы 2^{k-1} , что означает линейное от числа ребер число насыщений. Значит, Форд-Фалкерсон с масштабированием работает за $O(E^2 \log C)$.

Алгоритм Диница с масштабированием. Аналогично Форду-Фалкерсону с масштабированием, будем по очереди запускать алгоритм Диница, ставя ограничения на величину остаточных пропускных способностей, необходимых для того, чтобы ребро попадало в остаточную сеть. Общий вид алгоритма будет выглядеть как-то так:

```
int max_flow() {
    for (SCALE = 1 << 30; SCALE > 0; SCALE >>= 1) {
        while (bfs(s, t)) {
            while (path = dfs(s, t)) {
                relax(path);
            }
        }
    }
}
```

Какую мы получим выгоду от масштабирования? При переходе между итерациями мы знаем, что величина текущего минимального разреза не больше, чем $2 \cdot SCALE \cdot E$. То есть мы знаем, что теперь у нас будет суммарно не более $O(E)$ насыщений в рамках одной итерации.

Раньше мы оценивали Диница так: у нас был dfs , в котором мы считали число успешных и неуспешных итераций a_i и b_i . Мы получали, что $\sum a_i = O(VE)$, $\sum b_i = O(E)$. Сделаем новую оценку, используя тот факт, что насыщений всего $O(E)$. Здесь t_d — число путей длины d .

$$\sum_{d=0}^V \sum_{i=0}^{t_d} (a_{d,i} + b_{d,i}) \leq \sum_{d=0}^V (d \cdot t_d + E) \leq VE + V \cdot \sum_{d=0}^V t_d = O(VE)$$

В итоге на каждой итерации масштабирования мы сделаем $O(VE)$ шагов, получаем оценку на время работы $O(VE \log C)$.

Оценки Карзанова. Пусть в нашей сети единичные пропускные способности (например, в задаче о поиске максимального паросочетания). Тогда утверждается, что алгоритм Диница отработает за $O(E\sqrt{V})$.

Как это показать? Сначала покажем, что в рамках одной слойстой сети алгоритм делает $O(E)$ шагов. Для этого надо оценить сумму a_i , которая не превышает E , потому что после каждого успешного действия с ребром оно насыщается (читай — удаляется).

Отдельно покажем, что если определить $P = \sum p(v, u)$ (где $p(v, u)$ — потенциал вершины, сколько через нее максимально может пройти), то алгоритм Диница будет работать за $O(\sqrt{P}VE)$.

Определим вершинный разрез как такое множество вершин, через которое будет проходить любой путь из s в t .

Сделаем первые \sqrt{P} итераций Диница. Теперь в нашей сети будет не меньше, чем \sqrt{P} слоев. Поскольку эти слои не пересекались, то оставшийся поток в сети не больше минимума потока через вершинные разрезы, то есть он не больше, чем \sqrt{P} . А значит, алгоритм Диница найдет не более чем \sqrt{P} путей. Таким образом, мы получим сложность $O(\sqrt{P}VE)$. В применении к задаче о паросочетании $P = V$, а в рамках одной слойстой сети происходит не $O(VE)$, а $O(E)$ шагов, откуда получается сложность $O(E\sqrt{V})$.

Стоимостные потоки. Введем w — дополнительную стоимость потока вдоль ребра. Мы будем считать, что $w_{v,u} = -w_{u,v}$. Общей стоимостью назовем $\sum \frac{f_{v,u} \cdot w_{v,u}}{2}$ (у нас дважды учитывается поток вдоль одного ребра из-за обратных ребер). Обычно стоимость хотят минимизировать.

Выделяют несколько задач: min-cost-flow, min-cost-k-flow, min-cost-max-flow. В первой мы минимизируем вес, во второй мы ищем минимальный вес потока величины k , а в третьей найти максимальный поток, а среди таких потоков поток минимальной стоимости. Есть еще задача о циркуляции минимальной стоимости, к которой все эти задачи сводятся, если замкнуть $t \rightarrow s$ правильно заданным ребром.

Критерий оптимальности потока. Поток считается минимальным по стоимости потоком размера k , если и только если в его остаточной сети нет циклов отрицательного веса. Почему? Ну пусть мы нашли какой-то другой поток размера k , который имел меньшую стоимость. Если рассмотреть $f_1 - f_2$, то это будет циркуляцией. Если рассмотреть ее декомпозицию, в ней будут только циклы, при этом стоимость отрицательная. Значит, какой-то из циклов имеет отрицательную стоимость. Противоречие.

В доказательстве было узкое место: у нас могли нарушиться условия на пропускные способности, и поэтому не факт, что мы можем добавить этот новый цикл. Но на самом деле это правда, потому что f_1 и f_2 были корректными потоками. А именно, при добавлении у нас поток по ребру будет не больше чем максимум из потоков по ребру, что остается корректным потоком.

Алгоритм поиска min-cost-k-flow. Если у нас был поток f минимальной стоимости веса k , то поток минимальной стоимости веса $k + 1$ можно получить вдоль кратчайшего в плане стоимости пути в остаточной сети G_f . А это значит, что если мы сначала найдем минимальный поток веса 0 (брать отрицательные циклы, пока можем), а затем по очереди брать кратчайшие пути p из s в t .

Почему утверждение верно? Пусть был какой-то другой поток лучше. Тогда это значит? Что $w(f_1) + w(p) > w(f_2)$. Рассмотрим разность между этими потоками. В ней есть отрицательный цикл c . Рассмотрим $p + c$. Если его декомпозировать, то получим пути, каждый из которых не дешевле, чем p (потому что p был кратчайшим), и циклы, которые имели неотрицательную стоимость (иначе ими можно было бы уменьшить f_1). Тогда это значит, что поток $p + c$ имеет стоимость не меньше, чем p , противоречие.

Алгоритм поиска min-cost max-flow. В общем виде решение выглядит так — найти минимальный по весу путь, после чего вдоль него дополнять поток. Поскольку в сети бывают ребра отрицательного веса (и отрицательные циклы!), то искать кратчайший путь надо алгоритмом Форда-Беллмана. Тогда решить задачу можно за $O(|F| \cdot VE)$.

Алгоритм Джонсона. Во многих задачах сеть имеет более простой вид. Например, в задаче о назначениях в сети нет отрицательных циклов. Если отрицательных цикл нет, то можно воспользоваться идеей про потенциалы. Давайте посчитаем кратчайшие расстояния от s до всех остальных вершин. Тогда введем новый вес ребра, равный $w_{v,u} - \rho_u + \rho_v$ (их можно посчитать алгоритмом Форда-Фалкерсона). Таким образом, каждый вес ребра теперь не отрицательный. Поскольку вес любого пути $s \rightarrow t$ в новом графе — это вес в старом графе, взятый с какой-то константой, то кратчайший путь в новом графе — это кратчайший путь в старом. А такой кратчайший путь можно найти алгоритмом Дейкстры. После обновления вдоль пути у нас появятся обратные ребра, и потенциалы надо будет пересчитать. Как это делать? Заметим, что на кратчайшем пути цена всех ребер нулевая. А тогда цена обратных ребер тоже нулевая, и можно обновить кратчайшие расстояния через кратчайшие расстояния в новом графе. А именно, добавить к потенциалам кратчайшие расстояния в новом графе, которые мы тоже могли найти алгоритмом Дейкстры. В таком случае, сложность получается $O(|F| \cdot (E + V \log V) + VE)$, и в задаче о назначениях это дает асимптотику $O(V^3)$.

Строковые алгоритмы. Строкой мы будем называть конечную последовательность символов какого-то алфавита Σ . Если не сказано иного, то считать размер алфавита константой нельзя.

Задача о подстановке шаблона. Задача выглядит так: есть две строки s и t . Надо найти позиции, с которых в строке s можно прочитать строку t . При этом строка t состоит из обычных символов алфавита, а также два специальных символа Джокер (символ '?', который означает любой символ, а также астериск '*', который означает любую строку, возможно пустую). Соответственно, можно ли заменить вопросики на буквы, а звездочки на строки, чтобы получившаяся строка \bar{t} входила в s . Решать такую задачу можно с помощью динамического программирования за $O(|s| \cdot |t|)$, где состояние — это состояние вида «сколько символов из каждой строки мы прочитали».

Бор. Структура данных, которая позволяет делать добавление-удаление-поиск строк за их размер. Она реализуется как дерево, где на каждом ребре написана буква, а строке соответствует путь из корня до вершины. Соответственно, в каждой вершине надо хранить переходы по всем ребрам. Тогда добавление-поиск-удаление делаются просто спуском в дереве, с созданием вершин, если пока что переходов соответствующих не было. И надо еще хранить терминальные пометки, которые говорят о том, заканчивалась ли какая-то строка из множества в текущей вершине. Тривиальная реализация будет хранить $O(L|\Sigma|)$ памяти. Можно хранить список всех детей в списке, и тогда $O(L|\Sigma|)$ времени, но линейная память. Можно создать хеш-таблицу с переходами в вершине, тогда линейно по обоим параметрам. Если сжать бор (то есть сжать вершины степени 1), то тоже эффективнее будет.

π -функция. Префикс-функция от строки s определяется так: для каждой позиции i это такая максимальная длина l , что $s_{0,l-1} = s_{i-l+1,i}$. То есть, мы берем наибольший префикс строки s , совпадающий с суффиксом, заканчивающимся в позиции i . При этом мы искусственно запрещаем делать $l = i + 1$ (то есть запрещаем брать префикс, совпадающий с суффиксом).

Как эффективно вычислять эту функцию? Можно заметить, что $\pi_{i+1} \leq \pi_i + 1$. При этом нижней границы нет. То есть, если мы будем итеративно считать префикс-функцию, то значения у нас не могут увеличиваться больше чем на 1 на каждом шаге, поэтому суммарно уменьшаться на 1 они будут $O(n)$ раз.

Тогда префикс-функцию можно считать, например, так: сначала положить $\pi_i = \pi_{i-1} + 1$, а потом постепенно уменьшать, и делать проверку на равенство подстрок с помощью хэшей. Но стандартный алгоритм делает так: он по очереди пытается приписать символ к $\pi_{i-1}, \pi_{\pi_{i-1}-1}, \dots$. Это верно, потому что если строка максимальна, то ее надо было продлить относительно прошлой подстроки. Если так не получалось сделать (символы не совпадали), то надо было попробовать максимальную длину, которая подходит, после π_{i-1} . Но это ровно и есть значение префикс-функции от $s_{\pi_{i-1}}$, потому что $s_{0,l-1} = s_{i-l+1,i}$, и потому что мы запретили брать префикс, равный суффиксу. Тогда этот алгоритм так же работает за $O(|s|)$.

Чтобы найти подстроку t в строке s , можно запустить этот алгоритм на строке $t\$s$, после чего найти все точки, в которых $\pi_i = |t|$. Тут можно еще сэкономить память, и использовать только $O(|t|)$ памяти.

Автомат префикс-функции. Мы хотим, чтобы Кнут-Морис-Пратт (поиск подстроки в строке) работал неамортизированно. Для этого можно сначала насчитать все переходы, и потом делать переходы. А именно, мы хотим делать $go_{v,c}$ — переход из вершины v по символу c . Вершину v мы будем отождествлять с префиксом длины v . Тогда, если $s_{v+1} = c$, то $go_{v,c} = v + 1$, а иначе $go_{v,c} = go_{\pi_v,c}$.

Z-функция. В этот раз мы хотим найти для каждой позиции наибольший префикс $s_{0,n}$ и $s_{i,n}$. Тут функция монотонна, и ее можно считать наивно: пока можно увеличить значение, и оно корректно, надо увеличивать. Линейный алгоритм будет основан примерно на этом.

Будем строить z-функцию итеративно. Тогда каждый шаг алгоритма давал нам новую подстроку $s_{i,i+z_i}$. Мы их будем обозначать за l_i, r_i . Поддерживать мы из них будем всегда тот подотрезок, у которого правая граница как можно больше.

Тогда посмотрим на какое-то текущее i . Если $i \in [l, r]$, то $z_i \geq \min(r - i, z[i - l])$. Левое число нужно, чтобы мы не вышли за границу отрезка (потому что символы вне отрезка для нас неизвестные), а правое — это то, как мы используем текущий отрезок $[l, r]$. А именно, поскольку мы знаем, что наша строка сейчас такая же, как и $s_{0,r-l}$, то можно «подглядеть», чему было равно значение $z_{i'}$, где i' соответствует парному символу для i .

Дальше мы можем просто наивно увеличивать значение z-функции, и этого хватит для линейного времени. Время окажется линейным, потому что наивное увеличение надо использовать, только если $i + z_i = r$ (то есть, если мы хотим выйти за границы текущего отрезка). А это значит, что у нас сдвинется число r . Поскольку r может сдвинуться не более, чем n раз, то получаем оценку $O(|s|)$ времени и памяти.

Ахо-Корасик. Хотим решать следующую задачу: Дан набор из n строк s_i суммарного размера L на алфавите σ . Кроме этого, дан текст t , и задается вопрос насчет вхождений: сколько строк из набора входят в t как подстроки?

Первый шаг: построить бор на наборе строк. Теперь мы хотим сделать на этом боре детерминированный автомат. Автомат при чтении $t_0 \dots t_i$ должен переводить нас в вершину, которой соответствует самый длинный суффикс строки $t_0 \dots t_i$ из присутствующих в боре.

Для каждой вершины v создадим суффиксную ссылку $link$, которая будет вести в вершину, соответствующую самому длинному суффиксу $str(v)$, не совпадающему с $str(v)$, но присутствующему в боре. Кроме этого, обозначим переходы автомата за go_c . Тогда понятно, что $link_v$ можно найти, если посмотреть на суффиксную ссылку предка. А именно, если из суффиксной ссылки предка есть переход по символу, который написан на ребре (p_v, v) , то $link_v$ ведет именно туда. Если там перехода не было, то надо посмотреть на следующего предка в дереве суфсылок, и сделать аналогичную операцию. Но, если считать, что для предков посчитано go_c , то это то же самое, что взять $go_{c(p_v, v)}$ для $link_{p_v}$.

Для того, чтобы считать переходы в автомате, тоже можно воспользоваться предыдущими значениями. А именно, если из v есть ребро с символом c , то тогда go_c будет указывать в конец ребра. Иначе следует посмотреть на go_c для нашей суффиксной ссылки.

Обе динамики можно посчитать либо лениво, либо с помощью обхода в порядке возрастания глубин.

Оценка времени работы Можно оценить как $O(m|\sigma|)$, где m — число вершин в дереве, потому что динамика занимает ровно столько памяти (тривиально), и для нее верно $time = memory$. Также есть оценка $O(L)$, которую можно получить, если показать, что у нас всего спусков бывает не более чем $O(L)$, а подъемов суммарно не больше, чем спусков.

Суффиксное дерево. Суффиксным деревом мы будем называть сжатый бор, который будет хранить все суффиксы строки s , и только их. Сжатый бор — это бор, на ребрах которого написаны не буквы, а строки. Вместо строк мы будем хранить подотрезок $[l, l + len)$. Соответственно, если при добавлении строки оказывается так, что строка вдоль ребра не читается, то ребро нужно будет разделить на две части — общую и разную, и поставить между ними промежуточную вершину.

Добавлять в наш сжатый бор мы будем все суффиксы строки.

Квадратичный алгоритм. Будем считывать строку посимвольно. На ребрах будут написаны подотрезки. Поскольку мы не знаем число n , мы введем подотрезки $[l, \infty)$.

Алгоритм будет хранить все терминальные вершины в боре, и при добавлении нового символа будет продлевать все суффиксы строки на этот символ. Понятно, что все терминалы, в которые вели ребра $[l, \infty)$, можно никак не продлевать. Те суффиксы, которые лежали не на бесконечных ребрах, надо продлить на новый символ c . Если перехода нет, надо создать новую вершину, и сделать в нее ребро $[l, \infty)$.

Разделим суффиксы на определившиеся, неопределенные и неизвестные. Неизвестные — это суффикс, который еще не добавлен в бор. Определившийся суффикс — это такой суффикс, которому соответствует ребро $[l, \infty)$. Неопределенные — это такие суффиксы, которые являются префиксом определившегося суффикса. Будем называть их суффиксами 2, 1, и 0 типа (от определившихся к неизвестным).

Понятно, что суффиксы типа 2 не перестают быть типом 2 — это самый префикс такого типа. Тип 0 переходит в 1 или 2. 1 может переходить в 2. При этом важно понимать, что суффиксы всегда будут иметь вид 222221111110000. Почему? Потому что после какой-то единицы всегда будут идти единицы — если для позиции i суффикс s_i встречался раньше (на позиции $j < i$). то для s_{i+1} можно увидеть, что s_{j+1} соответствовал этому суффиксу.

Алгоритм Укконена. В терминологии 0-1-2 суффиксов мы хотим следующее: поддерживать в боре все 1 и 2-суффиксы, при этом каждый раз рассматривать только первый по счету 1-суффикс. При переходе по новому символу самый левый 1-суффикс может поменяться. Если он поменялся, то нам надо заменить подотрезок из 1-суффиксы на 2-суффиксы. Так мы получим амортизированную линейную сложность.

Тогда мы будем поддерживать указатель на j — первый 1-суффикс в дереве. Если мы видим переход, которого сейчас нет в боре, надо разделить ребро на два (если мы сейчас находимся в середине ребра). После этого из вершины делаем переход по новому символу. Теперь суффикс имеет тип 2. Сейчас осталось проделать аналогичную операцию для $j + 1$. Проблема в том, что непонятно, где находится соответствующий суффикс.

Введем суффиксные ссылки. Суффиксная ссылка из вершины v указывает на *позицию* в боре, которая соответствует суффиксу v , который на 1 короче текущего. То есть, для перехода мы сможем пользоваться суффиксными ссылками. Осталось понять, как их поддерживать.

Суффиксная ссылка из вершины всегда ведет в вершину. Это верно, потому что вершина соответствовала разветвлению, но тогда и для строки на 1 короче это тоже будет верно.

Пусть мы сейчас хотели сделать переход по суффиксной ссылке из вершины v . Посмотрим на $u = link(parent(v))$. Если мы прочитаем из вершины u ребро $parent(v) \rightarrow v$, то окажемся в позиции $link(v)$. Это либо будет вершиной, либо в ней сейчас произойдет ветвление. Причем чтение ребра надо реализовать за число вершин на пути, не за число символов.

За сколько это будет работать? Амортизированно за линию. Понятно, что нас интересует сейчас только суммарное время на чтение по ребрам во время поиска суффиксных ссылок. Тут можно ввести потенциал

$p(j)$ — число символов на ребре до предка-вершины j . Тогда, когда мы при чтении ребра проходим каждую вершину, мы уменьшаем свой потенциал.

Формальные грамматики и языки. Обозначим алфавит символов за σ , а множество конечных строк за $L \subset 2^{\sigma^*}$, где σ^* — это все последовательности конечной длины.

Регулярные выражения. Тривиальные регулярные выражения: (пустое множество) ϵ (пустая строка), $x \in \sigma$. Остальные регулярные выражения определяются рекурсивно относительно них. А именно, мы также разрешаем $A + B$ (последовательная запись), $A|B$ (выбор из двух регулярок), A^* (повторение строки из языка A 0, 1, 2, … раз). Звездочка еще называется замыканием *Kleene*.

Например, $\{0|1|2|3\}^* + \{0|1|2|3\}$ это множество всех непустых строк из символов $\sigma = \{0, 1, 2, 3\}$.

Способы задания языков. Регулярные выражения, ДКА, НКА, ϵ -НКА. Эти способы эквивалентны. Очевидно $L_{\text{ДКА}} \subset L_{\text{НКА}} \subset L_{\epsilon\text{-НКА}}$. Вложенность ϵ -НКА в ДКА можно показать, если рассмотреть алгоритм приведения одного автомата в другой. Например, можно выделить все подмножества вершин НКА как отдельные вершины ДКА (то есть, он будет иметь экспоненциальный размер). Тогда переход по ребру — это то же самое, что взять все вершины подмножества, и объединить переходы по ребрам из них. Еще стоит следить за переходами по пустому символу, потому что вершина сразу задает подмножество вершин, достижимых из нее по ϵ -ребрам.

Осталось показать, что регулярные выражения эквивалентны автоматам. По регулярным выражениям можно достаточно легко строить ϵ -НКА, если отдельно рассмотреть все возможные способы задания языка, и дальше рекурсивно. То есть, $A + B$ выражается, если последовательно записать автоматы A и B , и перегнать терминалы A в стартовую вершину для B . Остальные каким-то похожим образом.

Вот теперь совсем последний шаг — надо показать, что любой ДКА можно задать регулярным выражением. Введем $L_i(v, u)$ — регулярное выражение, которое задает все строки, по которым можно пройти из v в u в ДКА, используя первые i ребер. Изначально это какие-то пустые регулярки. Потом добавляем ребро (x, y, c) . Тогда $L_{i+1}(v, u) = L_i(v, u) \cup \{L_i(v, x) + c + \{L_i(y, x) + c\}^* + L_i(y, u)\}$.

Лемма о накачке. Для любого регулярного языка L существует такое n , что для всех строк $|s| \geq n$ верно, $s = uvw$, $|uv| \leq n$, $\forall k \geq 0 s' = uv^k w \in L$. Доказательство такое: посмотрим на ДКА и возьмем n , равное числу вершин. Тогда для длинных строк их путь будет содержать цикл. Этот цикл как раз и соответствует строке v . Это значит, что любая достаточно длинная строка имеет циклическую структуру внутри.

Минимизация ДКА. Два автомата изоморфны, если они изоморфны как графы с буквами на ребрах и с учетом терминальных вершин. Мы хотим найти автомат минимального размера, который будет принимать тот же язык, что и заданный автомат.

Выделим классы эквивалентности по вершинам автомата. А именно, каждой вершине в соответствие ставим множество строк, которые из нее читаются. Мы покажем, что автомат, где вершинами будут классы эквивалентности, минимален.

Замечание. Классы будут либо терминальными, либо нет, в зависимости от того, принадлежит ли классу пустая строка.

Почему такой построить можно? Потому что если склеить вершины с одинаковым множеством читаемых строк, то множество строк, принимаемых автоматом, не изменится. Поэтому алгоритм построения такого автомата будет тривиальным — нам надо просто объединять вершины с одинаковым множеством строк.

Пусть есть автомат меньшего размера, принимающий такой же язык. Для представителя каждого класса эквивалентности выберем строку, которая в него ведет. Тогда какая-то пара строк будет вести в одну и ту же вершину минимального автомата. Язык, соответствующий этой вершине, не совпадает хотя бы с одним из языков классов эквивалентности.

Если автомат был ациклическим, то нам надо взять вершины в порядке топологической сортировки, после чего надо вершине ставить в соответствии множество $\text{terminal, class}(go_{c_1}), \text{class}(go_{c_2}), \dots, \text{class}(go_{c_n})$. Тогда достаточно просто смотреть, был ли класс с таким множеством раньше. Для этого можно воспользоваться хэш-таблицей, и получить алгоритм сложностью $O(m)$.

Алгоритм Хопкрофта. Теперь наш автомат может иметь циклы. Мы будем наращивать классы постепенно. А именно, если вершины относятся к разным промежуточным классам, то они уже точно разные, а вершины внутри одного класса еще могут быть одинаковыми. Изначально классы определяются терминальностью вершины. Потом можно брать и определять новый класс с помощью хэша от классов вершин-потомков. Тогда надо делать операцию, пока классы меняются. Почему этого хватает? Потому что если две вершины были разными, и отличались строкой L , и $|L|$ было минимальным из возможных, то тогда на каждом шаге $1, 2, \dots, |L|$ одна вершина будет менять свой класс. Работать алгоритм будет за $O(nm\sigma)$, но это наш медленный алгоритм.

Теперь, чтобы улучшить алгоритм, воспользуемся идеей о классах-сплиттерах. Если нашим промежуточным классом мы разбили другой класс по символу c (то есть, если для какого-то A верно, что из него есть переход по c в класс B , и вне класса B), то тогда для его непересекающихся подклассов B_1, B_2 верно, что в дальнейшем можно будет разбивать только по одному из них. Доказательство из теории множеств — если мы уже разбили по двум классам, то по информации $x \in B, x \in B_1$ однозначно достраивается $x \in B_2$

Кроме того, число разбиений линейно — если взять сумму $|C| - 1$ по всем классам, то каждое разбиение уменьшает эту сумму на 1. Поскольку эта величина неотрицательна и изначально $O(m)$, то и количество ненулевых разбиений будет линейно. А количество нулевых разбиений не более чем в два раза превышает количество ненулевых.

Ну тогда достаточно сохранить обратные переходы в автоматае, рассматривать классы в порядке очереди, а при разбиении класса на два класса поменьше можно добавить в очередь только меньший из A_1, A_2 . Тогда рассматривать конкретный обратный переход в автоматае мы будем не более, чем $O(\log m)$ раз, а значит суммарное время работы будет $O(m\Sigma \log m)$. Тут, правда, важно, что каждое разбиение должно быть проведено за $O(|B| + \sum |A_i|)$, где B — это класс, по которому мы разбиваем, а A_1 — это меньший из двух непустых классов разбиения (тут важно, что сумма A_1 оценивается как $O(m\Sigma \log m)$ только если $|A_1| < |A_2|$), но это не очень сложно сделать, если поддерживать всякие счетчики, и при

перенаправлении указателей для разбиения делать новый класс меньшим из двух.

МашинаТьюринга. МТ — это абстрактная модель вычислений, которая состоит из каретки и бесконечной ленты, по которой каретка будет двигаться. Программа для МТ — это автомат, который по состоянию и букве на позиции каретки говорит, что должна сделать каретка и в какое состояние должен перейти алгоритм. То есть, написать что-то в текущую клетку, сдвинуть каретку на позицию вбок. Будем считать, что P — это такой класс задач, для которых ответ бинарен, и для любого ввода алгоритм для МТ работает за полиномиальное время от размера ввода. Соответственно, для них автомат тоже имеет полиномиальный размер.

NP-полнота. Введем несколько задач с бинарным ответом (то есть да/нет), которые мы будем хотеть решать:

- Lin solver — решение системы линейных уравнений (Полиномиальна — алгоритм Гаусса)
- SAT — Задача о булевой разрешимости
- CNF-SAT — SAT, который задается набором конъюнктов, каждый конъюнктов внутри состоит из скольки-то дизъюнктов.
- Subset-sum — задача о рюкзаке
- k-Clique — задача о поиске клики размера k .
- Ham — задача о поиске Гамильтонова цикла/пути.
- Euler — Задача о поиске эйлерова цикла/пути (опять-таки, полиномиальна)

Эти задачи интересны тем, что полиномиального алгоритма решения для них пока что нет (ну кроме тех, про которые я написал обратное), но зато есть полиномиальный способ проверить, верен ли положительный ответ (сертификат). Для P -задач алгоритм проверки совпадает с решением — само условие задачи уже является своим сертификатом.

Будем считать, что задачи NP (non-deterministic polynomial) — это такие задачи, которые разрешимы на недетерминированной машине Тьюринга. В недетерминированной машине Тьюринга надо делать одну из операций, записанных в ячейке. Раньше в ячейке была только одна команда, а теперь может быть несколько — мы сами можем выбрать, какое нас интересует. Задача разрешима на такой машине, если все пути в дереве разбора имеют полиномиальный размер. Например, в задаче Subset-sum можно сделать автомат вида «полное бинарное дерево», глубина которого будет полиномиальна, а листов будет 2^n . Тогда сертификатом будет просто путь в этом дереве.

Если у задачи есть возможность полиномиально проверить отрицательный ответ, то она принадлежит классу $co - NP$.

Задача называется NP -трудной, если она хотя бы так же трудна, как и любая другая задача из NP .

Что такое «так же трудна»? Задача A сводится к задаче B , мы должны предъявить полиномиальный алгоритм, который преобразует задачу A в задачу B (переводит вход в вход B , решение B в решение A). То есть, задача называется NP -трудной, если ее полиномиальное решение автоматически решит все задачи из NP .

Задача называется NP -полной, если она и NP и NP -трудная.

Теорема Кука. Любая задача из NP может быть записана в виде CNF-SAT. По любой задаче из NP мы знаем ее дерево решения для недетерминированной машины Тьюринга. Кроме того, у нас есть детерминированная МТ, которая умеет проверять сертификат. Рассмотрим эту МТ. Мы можем взять все состояния во все моменты времени как переменные, все положения каретки во все моменты времени как переменные, все символы на ленте во все моменты времени как переменные. Тогда в каждый момент времени для нас эти параметры задают булевый вектор.

- В булевом векторе для состояния, положения и символов одной позиции должно быть ровно по одной единице.
- Значения для всех символов в момент времени 0 должны совпадать со стартовыми значениями, кроме тех позиций, которые соответствуют сертификату — они могут быть любыми.
- Поскольку у нас везде по одному переходу, то нам надо построить граф импликаций.
- Кроме того, надо сделать условие, что символ меняется только под кареткой. Это тоже какие-то импликации

Схема будет иметь полиномиальный размер. Решение схемы явно задаст стартовые переменные, которые выдадут сертификат-решение, поэтому решение схемы решит и задачу.

В целом, подробное доказательство гораздо сложнее, но тут уже можно додумать, а громоздких записей нет.

Быстрое возведение числа в степень. $a^n = a^{2^{p_1} + 2^{p_2} + \dots + 2^{p_k}}$. $(a^n)^2 = a^{2^{p_1+1} + 2^{p_2+1} + \dots + 2^{p_k+1}}$. Тогда чтобы получить значение a^n можно разложить n на степени двойки, а дальше работать с ним как с вектором степеней: мы можем либо приписать в конец 0 (то есть умножить на a) или прибавить ко всем числам 1 (то есть возвести в квадрат). Можно заметить, что такими операциями получится получить любую степень за логарифмическое время.

Диофантовы уравнения. Дано уравнение $ax + by = c$. Сначала разделим все на $\gcd(a, b)$, получим $a'x + b'y = c'$, где $\gcd(a', b') = 1$. Тогда существует решение уравнения $a'x + b'y = 1$, которое находится расширенным алгоритмом Евклида, а дальше корни можно получить умножением на c' .

Алгоритм Евклида для поиска \gcd использует то, что если $a > b$ $\gcd(a, b) = d$, то $a = dx$, $b = dy$, $a - b = d(x - y)$, $\gcd(a, b) = \gcd(b, a - b)$. Ну тогда можно сжать много операций вычитания подряд одним взятием остатка. Брать остаток можно не более $O(\log C)$ раз — если $a > 2b$, то b явно уменьшилось в два раза, а если $a < 2b$, то $a - b < \frac{a}{2} < b$.

Теперь рассмотрим расширенный алгоритм Евклида. Он будет возвращать решение уравнения $ax + by = 1$. Мы будем пользоваться тем, что точка останова для нашего алгоритма это $a = 1$, $b = 0$, тогда $x = 1$, $y = 0$. Тогда можно явно предъявить пересчет рекурсии:

$$bx + (a \bmod b)y = 1$$

$$bx + (a - \lfloor \frac{a}{b} \rfloor b)y = 1$$

$$ay + b(x - \lfloor \frac{a}{b} \rfloor y) = 1$$

Тогда $x' = y$, $y' = x - \lfloor \frac{a}{b} \rfloor y$.

КТО. Если у нас есть набор взаимно простых чисел, которые используются как модули, то остаток по каждому из этих модулей задает число. Иначе говоря, если взять число, меньшее чем НОК модулей, то для него существует единственное представление через вектор остатков. Понятно, что размеры множества векторов и чисел меньших НОК одинаковый, для любого числа есть представление в виде вектора. Значит нам важно только то, что для любого вектора остатков существует число, у которого именно такое представление.

Построим итеративно это число. $x \equiv r_i \pmod{m_i}$. Тогда $m_1 \cdot X - m_2 \cdot Y = r_2 - r_1$. Если разность остатков равна нулю, то это просто задает нам один общий остаток по произведению модулей, а иначе мы можем найти решение, которое, опять же, задает нам остаток по произведению двух модулей.

Факторизация числа. Число можно представить в виде произведения простых чисел, каждое из которых задано в какой-то степени.

Факторизовать одно число можно за $O(\sqrt{n})$, потому что не бывает двух простых делителей числа, больших чем \sqrt{n} (иначе их произведение больше чем n). Тогда можно пройтись по первым $O(\sqrt{n})$ чисел, и делить на них основное, чтобы узнать показатель степени. Это работает за $O(\sqrt{n} + \log n) = O(\sqrt{n})$.

Решето Эратосфена. Часто нам хочется сделать предподсчет, чтобы потом быстро факторизовать числа от 1 до n . Для этого хочется запомнить для всех чисел их минимальный простой делитель. Тогда факторизовать можно будет за $O(\log n)$.

Можно, например, перебрать делители, перебрать второй множитель от 1 до $\frac{n}{x}$, и проставить соответствующие пометки в клетках, которые получатся как произведение. Это будет работать за $O(n \sum_{i=1}^n \frac{1}{i}) = O(n \log n)$ (очень грубая оценка, можно оценить как сумму обратных простых как $O(n \log \log n)$)

Давайте сделаем решето за линейное время. Мы хотим теперь ставить пометки для каждого числа $m = x \cdot d$ только один раз — когда x это его минимальный простой делитель. Тогда, если мы для текущего числа d поставим пометку для всех чисел m , которые получаются домножением на $x \leq \min_divisor(d)$, то x будет минимальным простым делителем m , поэтому мы каждое m рассмотрим всего 1 раз.

Ро-метод Полларда.

$$\begin{aligned}\mathbb{Z}_n &= \{0, 1, \dots, n - 1\} \\ \mathbb{Z}_n^+ &= \{1, 2, \dots, n - 1\} \\ \mathbb{Z}_n^* &= \{z \in \mathbb{Z}_n^+ : \gcd(z, n) = 1\}\end{aligned}$$

Заметим, что для каждого числа количество чисел, не взаимно простых с составным n , хотя бы $O(\sqrt{n})$. Тогда можно было бы выбрать много случайных чисел и проверить, что $\gcd(z, n) > 1$. Для каждого такого числа вероятность найти делитель будет $\frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}}$. Нам не хватает такой точности.

Построим функциональный граф для какой-то псевдослучайной функции g (чаще всего $g(x) = x^2 + 1$ по модулю n) на остатках. Также навесим дополнительное требование на g , чтобы она сохранила остатки (то есть $g(x) \bmod a = g(x \bmod a) \bmod a$, наша функция этому удовлетворяет). Заметим, что в нем произвольный бесконечный путь будет выглядеть как буква ρ — сначала какой-то период, а потом цикл. С учетом случайности функции и парадокса дней рождения в этой букве ρ будет \sqrt{n} вершин.

Пока что мы еще ничего не выиграли, но осталось совсем немного. Возьмем и мысленно сделаем функциональные графы по всем остаткам меньше n (назовем их a). При этом мы явно будем генерировать только путь в функциональном графе для числа n . Поскольку у составного n был делитель меньше, чем $2\sqrt{n}$, то в каком-то функциональном графе мы зациклимся за $O(\sqrt[4]{n})$ шагов. Если мы возьмем все пары (x_i, x_{2i}) , то тогда они за линейное время относительно размера буквы ρ будут указывать на одинаковую вершину. А это будет значить, что $|x_i - x_{2i}| \equiv 0 \pmod{a}$. Тогда если a было делителем n , то $\gcd(|x_i - x_{2i}|, n) > 1$. Тогда мы нашли какой-то делитель и можно раскладывать рекурсивно.

Алгоритм Миллера-Рабина. Создадим тест-проверку на $a^{n-1} \equiv 1 \pmod{n}$. Если это было верно для всех a от 1 до $n - 1$, то число n было простым, потому что тогда оно было взаимно простым со всеми $a < n$. Мы хотим брать случайные числа a , при этом сделать немного итераций. Это называется тестом Ферма.

Просто брать случайные a нельзя — есть числа Кармайкла, которые работают в качестве контртеста. Их какое-то полиномиальное количество, хоть и не очень много.

Сделаем новый тест: $a^2 \equiv 1 \pmod{n}$, $a \neq 1, a \neq -1$, таких чисел не бывает для простых n (потому что это будет означать что $(a - 1)(a + 1) \equiv 0 \pmod{n}$)

Рассмотрим $n - 1 = 2^s \cdot k$, $k \equiv 1 \pmod{2}$.

Сделаем $A(x) = \{x^k, x^{2k}, x^{4k}, \dots, x^{n-1}\}$. Тогда если у нас есть пара соседей $(d, 1)$, $d \neq 1, d \neq -1$, то тогда наш второй тест провалился — n не простое. Если последним элементом последовательности было число $d \neq 1$, то провалился тест Ферма — число составное. Назовем свидетелями такие x , для которых один из этих тестов выполнился, остальных назовем лжецами. Мы хотим показать, что при случайном выборе x -ов, вероятность получить лжеца будет не выше $\frac{1}{2}$.

Тогда $x \in \mathbb{Z}_n^+$, $\mathbb{Z}_n^+ = W \cup L$. А еще запомним, что \mathbb{Z}_n^* — группа.

Пусть наше число не было числом Кармайкла. Тогда $\exists x \in \mathbb{Z}_n^* : x^{n-1} \not\equiv 1 \pmod{n}$. Тогда можно взять подгруппу $B = \{z \in \mathbb{Z}_n^* : z^{n-1} \equiv 1\}$ (она содержит единицу,)

План: Хотим показать, что $L \subseteq B \subseteq \mathbb{Z}_n^*$

Автор сломался понять доказательство, возможно затухает его позже.

Задача шифрования. В общем задача выглядит так: есть Алиса и Боб, которые хотят передавать друг другу информацию таким образом, чтобы их сообщение можно было перехватить, но перехвативший человек не понял бы, что в нем написано. Дальше будем считать, что канал между Алисой и Бобом прослушивается Евой. В принципе, Евой может быть кто угодно, потому что каналы публичные. Причем, даже если Ева прочитает и расшифрует сообщение, Алиса с Бобом даже не узнают про это.

Схема RSA. Пусть у Алисы и Боба было по два объекта: публичный и приватный ключ. Ключ — это просто какой-то шифровальный объект. При этом ключи взаимо обратимы, но для каждого из них по отдельности найти обратную функцию сложно, то есть $M = \text{private}(\text{public}(M)) = \text{public}(\text{private}(M))$. Если все участники знают публичные ключи, то обмен можно провести таким образом: Алиса производит $\text{public}_B(M)$, тогда Боб делает $\text{private}_B(\text{public}_B(M)) = M$.

Факты для RSA:

- $\forall a, n : a^{\phi(n)} \equiv 1 \pmod{n}$
- $(a, b) = 1 \rightarrow \phi(ab) = \phi(a)\phi(b)$
- Можно проверить число на простоту за $O(\log^k n)$
- Нельзя посчитать факторизацию n за $O(\log^k n)$. Этот факт не доказан человечеством.

Зафиксируем два простых ключа p_1, p_2 . Их произведение равно n . Мы знаем, чему равно $\phi(n)$ — это $(p_1 - 1)(p_2 - 1)$. При этому быстро посчитать $\phi(n)$ нельзя по факту 4.

Нашим публичным ключом будет пара (e, n) , приватным ключом будет пара (d, n) . e и d — это такие числа, что $ed \equiv 1 \pmod{\phi(n)}$. Тогда $M^{ed} = M$. Тогда Алиса сначала фиксирует какое-то e , взаимно простое с $\phi(n)$. Тогда Алиса может решить диофантово уравнение $ed + \phi(n)m = 1$, получая d .

Теперь пусть все знают e . Тогда Алиса посылает M^e , Боб делает $M^e d = M^{\phi(n)k+1} = M$.

Man in the middle и authority. На самом деле, описанный выше не работает, если у нам есть Мэллори, которая читает и модифицирует канал связи между Алисой и Бобом. Мэллори может выдать им свои публичные ключи в качестве публичного ключа друг друга. Тогда Мэллори будет явно читать весь канал между Алисой и Бобом, причем поскольку она знает публичные ключи Алисы и Боба, она может отправлять им что угодно.

Таким образом, мы хотим решить следующую задачу: получить публичный ключ, которому можно доверять.

Мы предположим, что у нас есть добрый Трент, которому все стороны доверяют, и публичный ключ которого известен всем. Тогда все могут сообщить Тренту свои публичные ключи и спросить у Трента чужие публичные ключи.

Цифровая подпись. Для того, чтобы все точно верили, что сообщение отправляет именно Трент, он может отправить пару $(M, \text{private}(M))$. Это называется цифровой подписью Трента. Теперь наши участники запрашивают сертификат у Трента, а Трент в качестве сертификата выдает Алисе свою цифровую подпись для этого ключа. То есть теперь Алиса может отправлять свой публичный ключ, представляясь Алисой, которую одобрил Трент.