

ЛКШ, ЛКШ.2018.Август В' В', конспект лекции

Собрано 5 августа 2018 г. в 10:43

Содержание

| | |
|--|----------|
| 1. Правильные скобочные последовательности | 1 |
| 1.1. Определение | 1 |
| 1.2. Проверка СП на правильность | 1 |
| 1.3. Проверка на правильность СП из одного типа скобок | 1 |
| 1.4. Количество ПСП заданной длины | 2 |
| 1.5. Количество ПСП заданной длины, способ 2 | 2 |
| 1.6. Формула для количества ПСП | 2 |
| 1.7. Вычисление первых n чисел Каталана | 3 |

Тема #1: Правильные скобочные последовательности

4 августа

1.1. Определение

Формальное определение ПСП:

1. Пустая строка является ПСП.
2. Если строки S_1 и S_2 являются ПСП, то строка S_1S_2 тоже является ПСП.
3. Если строка S является ПСП, то строка (S) тоже является ПСП (вставить сюда другие виды скобок при необходимости).

Если написать корректное арифмитическое выражение и стереть всё, кроме скобок, то у нас останется ПСП.

1.2. Проверка СП на правильность

Пусть у нас есть скобочная последовательность из нескольких типов скобок. Мы хотим проверить, является ли она правильной. Есть простой алгоритм, делающий это с помощью структуры данных «стек».

Стек — структура данных, позволяющая добавлять элемент в конец и доставать элемент из конца. Это умеют плюсовый `vector` и массив в Питоне.

Как проверить СП на правильность? Каждой закрывающейся скобке должна соответствовать открывающаяся того же типа и наоборот.

Давайте будем делать что-то вроде сканлайна по последовательности. События — скобки. Если мы встречаем открывающуюся скобку, добавим ее в стек. Встречаем закрывающуюся — посмотрим на верх стека. Если там лежит открывающаяся скобка того же типа, то удаляем ее из стека и продолжаем сканлайн. Иначе последовательность неправильная.

Не забудем проверить, что в конце стек должен оказаться пустым.

При обработке закрывающей скобки не забудьте, что стек мог оказаться пустым.

1.3. Проверка на правильность СП из одного типа скобок

Введём понятие баланса.

Балансом скобочной последовательности на отрезке назовем разность между количеством открывающихся и закрывающихся скобок на этом отрезке.

Заметим, что теперь нам не нужен стек. По сути, баланс на префиксе — размер стека при проходе предыдущим алгоритмом на том же префиксе.

Вот такой алгоритм получился:

```
1 balance = 0
2 for (bracket : brackets) {
3     if (bracket == '(')
4         balance++
5     else if (--balance < 0)
6         :-(
7 }
8 if (balance == 0) :-)
9 else :-(
```

1.4. Количество ПСП заданной длины

Давайте посчитаем количество ПСП из n открывающих и закрывающих скобок.

Для этого воспользуемся динамическим программированием.

Пусть dp_i — количество ПСП из i открывающихся и закрывающихся скобок.

База: $dp_0 = 1$

Переход: как мы можем получить ПСП из n открывающихся и n закрывающихся скобок?

Посмотри на любую ПСП требуемой длины. Посмотрим на позицию парной к первой скобке.

Пусть это $2i + 1$. На отрезке $[1; 2i]$ лежит ПСП из i открывающихся и закрывающихся скобок.

Зафиксируем какое-то i . Тогда внутри первых скобок мы получим dp_i вариантов расставить скобки, а справа dp_{n-i-1} вариантов. По правилам комбинаторики их надо перемножить.

Просуммируем по всем i . Получим рекурренту

$$dp_0 = 1; dp_n = \sum_{i=0}^{n-1} dp_i \cdot dp_{n-i-1}$$

Этот способ требует $\mathcal{O}(n^2)$ времени и $\mathcal{O}(n)$ памяти.

1.5. Количество ПСП заданной длины, способ 2

Будем считать другую динамику. $dp_{pos, bal}$ — сколько существуют скобочных последовательностей из pos скобок таких, что их баланс равен bal и bal не бывает меньше нуля.

База: $dp_{0, ??} = 0$; $dp_{0, 0} = 1$, $dp_{?, -1} = 0$

Переход: Как мы можем получить последовательность из pos скобок с балансом bal ? Взять последовательность из $pos - 1$ скобок с балансом $bal - 1$ и приписать к ней открывающую скобку; взять последовательность из $pos - 1$ скобок с балансом $bal + 1$ и приписать к ней закрывающую скобку.

Получаем динамику $dp_{pos, bal} = dp_{pos-1, bal-1} + dp_{pos-1, bal+1}$.

Этот способ требует $\mathcal{O}(n^2)$ времени и $\mathcal{O}(n^2)$ памяти (хотя его можно улучшить до линии памяти). Такой подход пригодится нам в будущем.

1.6. Формула для количества ПСП

Количество ПСП из n открывающих и n закрывающих скобок равно n -му числу Каталана (C_n).

Для чисел Каталана верна формула

$$C_n = \frac{C_{2n}^n}{n+1}$$

Давайте ее докажем.

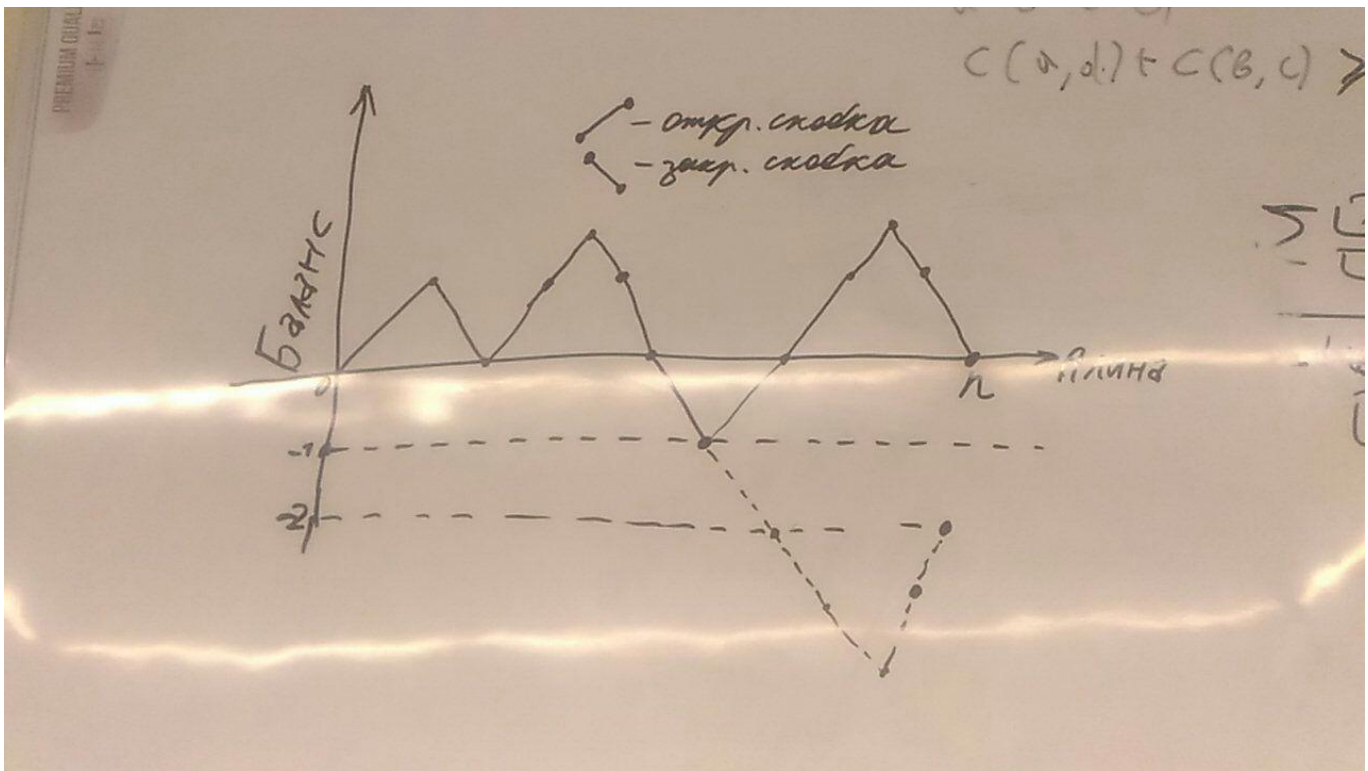
Сколько существует скобочных последовательностей с балансом 0 (балансы на префиксах могут быть отрицательны)? C_{2n}^n , так как мы по сути просто должны из $2n$ позиций для скобок выбрать n позиций, куда поставить открывающие, а на остальные поставим закрывающие.

Теперь мы хотим понять, сколько существует последовательностей с балансом 0, у которых есть отрицательный баланс на каком-то префиксе.

Для этого сделаем следующее: рассмотрим какую-то последовательность с балансом 0, с отрицательным балансом на каком-то префиксе. Сопоставим этой последовательности другую последовательность по следующему правилу:

1. Найдем первый момент, когда баланс стал равен -1 . Назовем эту позицию i .

2. Все скобки, начиная с $i + 1$ поменяем на противоположные ('(' на ')') и наоборот).



Заметим, что баланс новой последовательности будет -2 , так как баланс суффикса, который мы меняли, был 1 , а после изменения стал -1 .

Давайте докажем, что такое сопоставление — биекция. То есть, что любой последовательности с балансом 0 , уходящей на каком-то префиксе в минус, мы сопоставили какую-то ровно одну последовательность с балансом -2 и наоборот, любой последовательности с балансом -2 соответствует ровно одна последовательность с балансом 0 , где-то уходящая в минус.

Первое утверждение очевидно. Для доказательства второго найдем первый момент, когда баланс на префиксе стал -1 . Такой обязательно есть, так как мы как-то пришли в -2 . Снова инвертируем. Получили последовательность, которая приходит в ноль (и при этом на каком-то префиксе баланс отрицателен).

Таким образом, это биекция, значит, их количества равны. А количество последовательностей скобок, которые приходят в -2 посчитать легко. Из $2n$ позиций нужно выбрать $n - 1$ для открывающих скобок. Итого C_{2n}^{n-1} .

Значит, чтобы найти количество ПСП нужно из C_{2n}^n вычесть C_{2n}^{n-1} . Проверьте, что это равно $\frac{C_{2n}^n}{n+1}$.

1.7. Вычисление первых n чисел Каталана

Теперь научимся вычислять все первые n чисел Каталана за $\mathcal{O}(n)$.

Есть два способа. Первый — предподсчитать все факториалы до $2n$. Тогда ответ на запрос будет за $\mathcal{O}(1)$.

Второй — предподсчитать все C_{2k}^k для k от 1 до n . Это тоже можно сделать за линию:

База $C_2^1 = 2$, переход $C_{2k}^k \cdot \frac{(2k+1)(2k+2)}{(k+1)^2} = C_{2k+2}^{k+1}$.