Домашняя работа

Котов Артем, МОиАД2020

6 декабря 2020 г.

Содержание

Is it time to lakonichnyi solutions? — Well…it‘s not

Task 1: Бесконечно несловарное слово

Решение.

■

Task 2: Поиск с k ошибками

Решение.

Работает через хеши: посчитаем для префиксов t и s хеши, это стоило нам $\mathcal{O}(|s| + |t|)$.

Возьмем окошко длины $|s|$. Пройдемся им по тексту, бинпоиском находим максимальный общий префикс через сравнение хешей, это стоит нам $\mathcal{O}(\log |s|)$. Если наша строка не совпала с окошком, то значит следующий элемент после префикса точно отличается, значит надо проверить, что в оставшейся (правой) части образца отличий не больше, чем $k - 1$. То есть мы свели задачу к задаче поменьше в смысле k. Так мы можем сделать максимум k раз, так как иначе заведомо больше отличий (можно об этом думать, как о том, что мы поочередно откусывает различные символы). Если мы так дошли до конца окошка не больше, чем за k шагов, то есть совпадение. В итоге мы потратили $\mathcal{O}(|s| + |t| + |t|k \log |s|) = \mathcal{O}(|t|k \log |s|)$, в последнем переходе я посчитал, что $|t| \geq |s|$. ■

Task 3: T9

Решение.

Заводим бор, при этом в терминалах храним еще дополнительную информацию — частоту, и еще из них будут ссылки на 5 самых частых. Посмотрим на первый запрос: спускаемся по бору, находим слово, увеличиваем его частоту. Для каждой вершины на обратном пути (даже не терминальных!) к корню после этого надо проверить, надо ли обновить ссылки на самых частых. Это можно сделать через сравнение минимальной частоты среди текущих 5, и если надо, то обновляем ссылку на только что “увеличенную” вершину, каждая такая операция стоит нам $\mathcal{O}(1)$, кажется, что это дорого, но у нас даже суммарная длина слов не больше 10^6.

Замечание. *Глубже спускаться смысла нет, так как они явно не суффиксы слова, у которого мы увеличили частоту.*

Теперь, когда приходит второй запрос, мы спускаем по бору, находим вершину, в которой надо остановиться (это может быть не терминальной вершиной), смотрим на 5 самых встречаемых (у нас есть вот такие вспомогательные ссылки). Чтобы вывести эти 5 мы переходим по этим ссылочкам, которые указывают в терминальные вершины, поднимаемся от них к корню, это будет как бы слово наоборот (в смысле обычное слово получается у нас спуском из корня в терминал, а если мы поднимаемся из терминала к копню, то мы получаем слово, записанное наоборот), следовательно, надо просто развернуть прочитанное и выдать пользователю. Такое, вообще говоря, будет нам стоить O-большое от суммарной длины этих самых 5 частых слов.

Замечание. *Ясно, что быстрее нельзя, так как хотя бы сам вывод будет столько стоить.*

■

Task 4: Семейства функций

Решение.

Важное предположение, я считаю в дальнейшем, что функции переводят каждый элемент их левого множества, в какие-то (необязательно различные, но единственные) элементы правого множества, вроде бы, это называется инъекцией.

1) Представим это себе как две доли графа, слева у нас 2^n вершин, справа — 2^m. Тогда нам комбинаторно надо посчитать кол-во различных ребер из левой доли в правую, т.е. по отношению каждой вершине справа задать вопрос, кто в нее приходит, тогда кол-во инъективных отображению 2^{m2^n}.

2) По определению, посчитаем вероятность, что $\forall x_1 \neq x_2 : h(x_1) = h(x_2)$. Такая вероятность равна $P = \frac{\#\{h(x_1)=h(x_2)\}}{|H|} = 2^{m(2^n-1)}/2^{m2^n} = \frac{1}{2^m} = \frac{1}{|\mathbb{F}_2^m|}$. То есть по определению для любых n и m. Так же показывается и 2-независимость $\frac{\#\{h(x_1)=y_1, h(x_2)=y_2\}}{|H|} = \frac{2^{m(2^n-2)}}{2^{m2^n}} = \frac{1}{(2^m)^2} = \frac{1}{|\mathbb{F}_2^m|^2}$

3) Хм, стоит отметить, что бесконечности бывают разные, остановимся на случае счетного множества. Тогда, кажется, что сломается условие на произвольные x, но что-то не получается формально показать.

■