|  |
| --- |
| В декартовом дереве из n узлов, приоритеты y которого являются [случайными величинами](https://neerc.ifmo.ru/wiki/index.php?title=%D0%94%D0%B8%D1%81%D0%BA%D1%80%D0%B5%D1%82%D0%BD%D0%B0%D1%8F_%D1%81%D0%BB%D1%83%D1%87%D0%B0%D0%B9%D0%BD%D0%B0%D1%8F_%D0%B2%D0%B5%D0%BB%D0%B8%D1%87%D0%B8%D0%BD%D0%B0" \o "Дискретная случайная величина) c равномерным распределением, средняя глубина вершины O(\log n).  **Доказательство:** |
| Будем считать, что все выбранные приоритеты y попарно различны.  Для начала введем несколько обозначений:   * x_k— вершина с k-ым по величине ключом; * индикаторная величина A_{i, j} = \left\{\begin{array}{lllc} 1 ,&& x_i\  \text{is ancestor of} \ x_j\\  0 ,&& \text{otherwise}\\ \end{array}\right. * d(v)— глубина вершины v;   В силу обозначений глубину вершины можно записать как количество предков:  d(x_k) = \sum\limits_{i = 1}^{n} A_{i,k}.  Теперь можно выразить [математическое ожидание](https://neerc.ifmo.ru/wiki/index.php?title=%D0%9C%D0%B0%D1%82%D0%B5%D0%BC%D0%B0%D1%82%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%BE%D0%B5_%D0%BE%D0%B6%D0%B8%D0%B4%D0%B0%D0%BD%D0%B8%D0%B5_%D1%81%D0%BB%D1%83%D1%87%D0%B0%D0%B9%D0%BD%D0%BE%D0%B9_%D0%B2%D0%B5%D0%BB%D0%B8%D1%87%D0%B8%D0%BD%D1%8B" \o "Математическое ожидание случайной величины) глубины конкретной вершины:  E(d(x_k)) = \sum\limits_{i = 1}^{n} Pr[A_{i,k} = 1]— здесь мы использовали линейность математического ожидания, и то что E(X) = Pr[X = 1]для индикаторной величины X(Pr[A] — вероятность события A).  Для подсчёта средней глубины вершин нам нужно сосчитать вероятность того, что вершина x_iявляется предком вершины x_k, то есть Pr[A_{i,k} = 1].  Введем новое обозначение:   * X_{i, k}— множество ключей \{x_i, \ldots, x_k\}или \{x_k, \ldots, x_i\}, в зависимости от i < k или i > k. X_{i, k} и X_{k, i} обозначают одно и тоже, их мощность равна |k - i| + 1.  |  | | --- | | **Лемма:**  Для любых i \ne k, x_iявляется предком x_kтогда и только тогда, когда x_iимеет наибольший приоритет среди X_{i, k}. | | **Доказательство леммы:** | | Если x_iявляется корнем, то оно является предком x_kи по определению имеет максимальный приоритет среди всех вершин, следовательно, и среди X_{i, k}.  С другой стороны, если x_k— корень, то x_i— не предок x_k, и x_kимеет максимальный приоритет в декартовом дереве; следовательно, x_iне имеет наибольший приоритет среди X_{i, k}.  Теперь предположим, что какая-то другая вершина x_m— корень. Тогда, если x_iи x_kлежат в разных поддеревьях, то i < m < kили i > m > k, следовательно, x_mсодержится в X_{i , k}. В этом случае x_i— не предок x_k, и наибольший приоритет среди X_{i, k}имеет вершина с номером m.  Наконец, если x_iи x_kлежат в одном поддереве, то доказательство применяется по индукции: пустое декартово дерево есть тривиальная база, а рассматриваемое поддерево является меньшим декартовым деревом. | |  |   **Вернемся к основному доказательству:**  Так как распределение приоритетов равномерное, каждая вершина среди X_{i, k}может иметь максимальный приоритет, мы немедленно приходим к следующему равенству:  Pr[A_{i, j} = 1] = \left\{\begin{array}{lllc} \dfrac{1}{k - i + 1} ,&& k \ > \ i\\  0 ,&& k\ =\ i\\ \dfrac{1}{i - k + 1} ,&& k \ < \ i\\ \end{array}\right.  Подставив последнее в нашу формулу с математическим ожиданием получим:  E(d(x_k)) = \sum\limits_{i = 1}^{n} Pr[A_{i,k} = 1] = \sum\limits_{i = 1}^{k - 1}\dfrac{1}{k - i + 1} + \sum\limits_{i = k + 1}^{n}\dfrac{1}{i - k + 1} \leqslant  \leqslant \ln(k) +  \ln(n - k)+2(здесь мы использовали неравенство \sum\limits_{i = 1}^{n}  \dfrac{1}{i} \leqslant \ln(n) + 1)  \log(n)отличается от \ln(n)в константу раз, поэтому \log(n) = O(\ln(n)).  В итоге мы получили что E(d(x_k)) = O(\log(n)). |
|  |

Таким образом, среднее время работы операций \mathrm{split} и \mathrm{merge} будет O(\log(n)).