伸展树时间复杂度的一个上界

王祝康

2022 年 6 月 2 日

定理 1. 对于一棵有 n 个节点的伸展树, 进行 m 次伸展操作，时间复杂度存在上界 $O((n+m)\log n)$.

证明. 采用势能分析方法.

定义势函数 $\Phi:\mathbb{D}\mapsto\mathbb{R}$, 其中 D_i 是伸展树经过 i 次操作后得到的数据状态, $i=1,2,\cdots,m$. $\Phi(D_i)$ 是该状态下的的势函数值. 不妨设 c_i 为第 i 次操作的实际代价, 则均摊代价 $\hat{c}_i$ 定义为

$$\hat{c}_i = c_i + \Phi(D_i) - \Phi(D_{i-1}) \tag{1}$$

其中 $i=1,2,\cdots,m$. 因此有

$$\Phi(D_m) + \sum_{i=0}^{m} c_i = \Phi(D_0) + \sum_{i=0}^{m} \hat{c}_i \tag{2}$$

若 $\exists\ \Phi:\mathbb{D}\mapsto\mathbb{R}$, 使得 $\forall\ i\in\mathbb{N}$, $\Phi(D_i)$ 都有界, 那么显然

$$\begin{aligned}\sum_{i=0}^{m} c_i &= \sum_{i=0}^{m} \hat{c}_i + \Phi(D_0) - \Phi(D_m)\\ &\le \sum_{i=0}^{m} \hat{c}_i + M\end{aligned} \tag{3}$$

其中 M 为常数. 这就说明, 均摊复杂度给出了实际复杂度的一个上界.

下面说明一些基本约定. 对伸展树进行 k 次伸展操作后, 定义势函数

$$\Phi(D_k) = \sum_{x\in D_k} \log(rank(x)) \tag{4}$$

其中 x 为任意节点, $rank(x)$ 为以 x 为根的子树的节点个数. 定义任意节点的势能

$$\phi(x) = \log(rank(x)) \tag{5}$$

定义 v,p,g 分别为当前节点、父节点、祖父节点, v',p',g' 分别为操作后的对应节点.

下面对伸展树节点的各类旋转操作进行对应的势能变化分析.

1. L/R 旋转操作, 以 L 为例:

$$
\begin{aligned}
\Delta\Phi &= \phi\left(v'\right) + \phi\left(p'\right) - \phi(v) - \phi(p) \\
&= \phi(p) + \phi\left(p'\right) - \phi(v) - \phi(p) \\
&= \phi\left(p'\right) - \phi(v) \\
&\leq \phi\left(v'\right) - \phi(v)
\end{aligned} \tag{6}
$$

2. LL/RR 旋转操作, 以 LL 为例:

$$
\begin{aligned}
\Delta\Phi &= \phi\left(v'\right) + \phi\left(p'\right) + \phi\left(g'\right) - \phi(v) - \phi(p) - \phi(g) \\
&= \phi(g) + \phi\left(p'\right) + \phi\left(g'\right) - \phi(v) - \phi(p) - \phi(g) \\
&= \phi\left(p'\right) + \phi\left(g'\right) - \phi(v) - \phi(p) \\
&\leq \phi\left(v'\right) + \phi\left(g'\right) - 2\phi(v)
\end{aligned} \tag{7}
$$

可以证明: $\phi\left(v'\right) + \phi\left(g'\right) - 2\phi(v) \leq 3\left(\phi\left(v'\right) - \phi(v)\right)$, 即 $\phi(v) + \phi(g') - 2\phi(v') \leq 0$. 事实上, 考虑 $rank(v) = a$, $rank(g') = b$, 则 $rank(v') \geq a + b$. 因此有

$$
\phi(v) + \phi\left(g'\right) - 2\phi\left(v'\right) \leq \log\frac{ab}{(a+b)^2} \leq \log\frac{1}{4} < 0 \tag{8}
$$

即

$$
\Delta\Phi \leq 3\left(\phi\left(v'\right) - \phi(v)\right) \tag{9}
$$

3. LR/RL 旋转操作, 以 LR 为例:

$$
\begin{aligned}
\Delta\Phi &= \phi\left(v'\right) + \phi\left(p'\right) + \phi\left(g'\right) - \phi(v) - \phi(p) - \phi(g) \\
&= \phi(g) + \phi\left(p'\right) + \phi\left(g'\right) - \phi(v) - \phi(p) - \phi(g) \\
&= \phi\left(p'\right) + \phi\left(g'\right) - \phi(v) - \phi(p) \\
&\leq 2\phi\left(v'\right) - 2\phi(v) \\
&= 2\left(\phi\left(v'\right) - \phi(v)\right)
\end{aligned} \tag{10}
$$

综合式 (6)(9)(10) 可知, 对于任意节点 x 的任何一次旋转操作, 势能变化量有上界

$$
\Delta\Phi \leq O(\phi(x') - \phi(x)) \tag{11}
$$

现假设第 k 次伸展操作中, 旋转的节点为 $x_1, x_2, \cdots, x_s$, 其中 $1 \leq s \leq n$. 则总的势能变化量为

$$
\Phi(D_k) - \Phi(D_{k-1}) \leq O\left(\sum_{i=1}^{s-1} \phi(x_{i+1}) - \phi(x_i)\right) = O\left(\phi(x_s) - \phi(x_1)\right) \leq O\left(\log n\right) \tag{12}
$$

由势能分析理论, 一次伸展操作中, 旋转的常数时间可以忽略, 因 $\hat{c_k}$ 与 $\Phi(D_k) - \Phi(D_{k-1})$ 有相同的复杂度, 即

$$\hat{c_k} \le O\left(\log n\right) \tag{13}$$

又由假设, 有 $\Phi(D_m) > 0$, $\Phi(D_0) < n \log n$. 结合式 (3) 可知

$$\sum_{k=0}^{m} c_k \le m\hat{c_k} + n \log n \le O\left(m \log n + n \log n\right) \tag{14}$$

证毕[1]. □

[1]可参考另一种风格的证明: Daniel D. Sleator, Robert Endre Tarjan, A data structure for dynamic trees