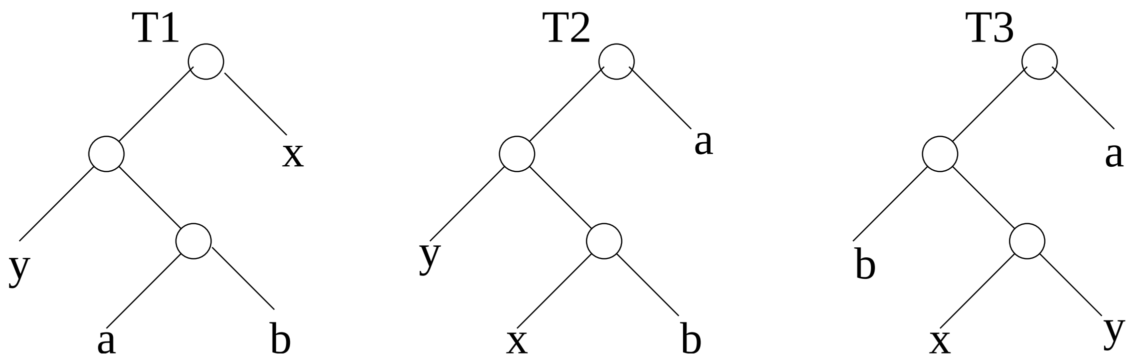
Correctness



- After the first change the difference in cost

$$
\begin{aligned}
B(T1) - B(T2) &= \sum_{c \in C} f(c)\, d_{T1}(c) - \sum_{c \in C} f(c)\, d_{T2}(c) \\
&= f(x)\, d_{T1}(x) + f(a)\, d_{T1}(a) - f(x)\, d_{T2}(x) - f(a)\, d_{T2}(a) \\
&= f(x)\, d_{T1}(x) + f(a)\, d_{T1}(a) - f(x)\, d_{T1}(a) - f(a)\, d_{T1}(x) \\
&= (f(a) - f(x))(d_{T1}(a) - d_{T1}(x)) \\
&\geq 0
\end{aligned}
$$

Optimal Substrucure Property

Let T represent an optimal prefix code over C. Consider any two characters x and y that appear as sibling leaves in T, and let z be their parent. Then considering z as a character with frequency $f(z) = f(x) + f(y)$, the tree $t' = T - \{x, y\}$ represents and optimal prefix code for $C' = C - \{x, y\} \cup \{z\}$.

- For $\forall c \neq x, y \in C$, we have $c \in C'$, and

$$f(c) \cdot d_T(c) = f(c) \cdot d_{T'}(c)$$

- Since, $d_T(x) = d_T(y) = d_{T'}(z) + 1$, we have

$$\begin{aligned} f(x)\, d_T(x) + f(y)\, d_T(y) &= (f(x) + f(y))(d_{T'}(z) + 1) \\ &= f(z) d_{T'}(z) + (f(x) + f(y)) \end{aligned}$$