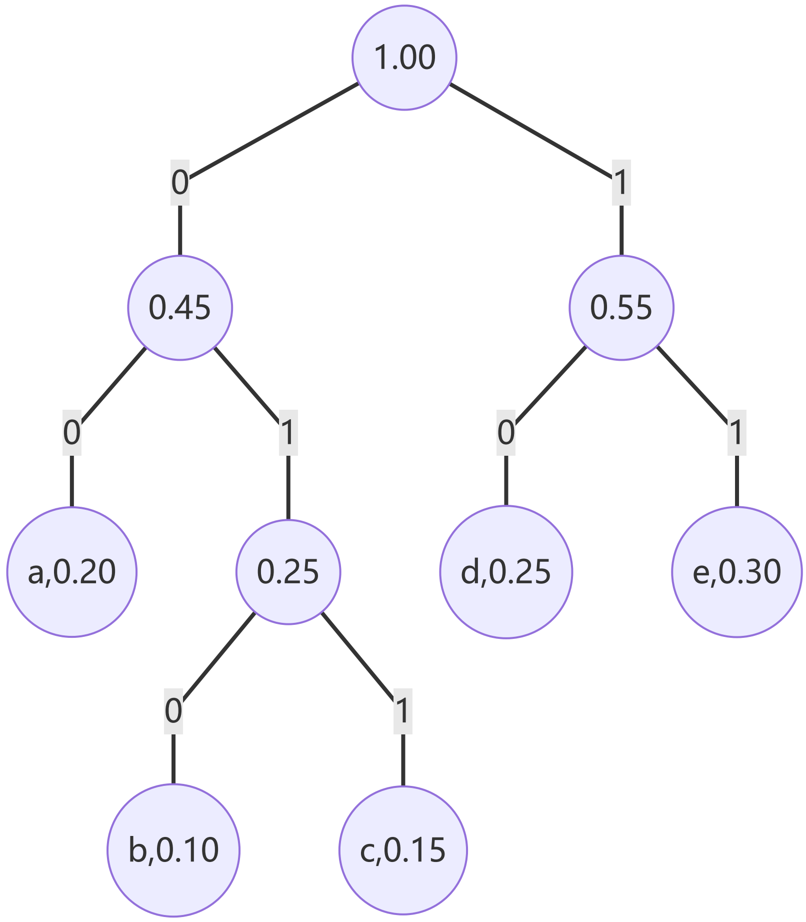
二.19

- 记T的分支点有s个，则$\nu = s + t = 2s + 1$
- 则$s = t - 1$，$\varepsilon = \nu - 1 = s + t - 1 = 2t - 2$

二.21

- 

- $APL = \frac{1.00+0.45+0.25+0.55}{1} = 2.25$

二.24

证明: 概率最低的两个消息符号有相同的概率

- 记$s_i = |\{j|p_j = 2^{-s_i}, j = 1, 2, \cdots, n\}|$，$i = 0, 1, 2, \cdots, n, \cdots$
- 记m为满足$s_i \neq 0$的最大的i
- 则$1 = \sum_{j=1}^{n} p_j = \sum_{i=0}^{m} \frac{s_i}{2^i}$，且显然有$s_0 \leq 1$
- 对于任一满足$s_i > 1$的i，可令$s'_i = s_i - 2$，$s'_{i-1} = s_{i-1} + 1$，$s'_k = s_k$（$k \neq i, i-1$）

- 则有$1=\sum_{i=0}^{m}\frac{s_i}{2^i}=\sum_{i=0}^{m}\frac{s'_i}{2^i}$
- 在重复上述操作足够多次后，必然可得到序列$s'_0,s'_1,\cdots,s'_m$满足$\forall i\in[0,m]$都有$s'_i\le 1$且$1=\sum_{i=0}^{m}\frac{s'_i}{2^i}$
- 若$s'_0=1$，则$s'_1=s'_2=\cdots=s'_m=0$
- 若$s'_0=0$，则$1=\sum_{i=0}^{m}\frac{s'_i}{2^i}\le\sum_{i=1}^{m}\frac{1}{2^i}=1-\frac{1}{2^m}$不成立
- 故必有$s'_m=0$
- 由于$s_m>0$，故一定是对$i=m$进行过不少于一次操作使得s_m减小至s'_m，即有$s_m>1$
- 综上，概率最低的两个消息符号有相同的概率$\frac{1}{2^m}$

证明: 在此概率分布下,Huffman编码的平均编码长度为$-\sum p_i\log_2 p_i$

- 进行Huffman编码，且每次合并概率最低的两个消息符号或合并所得节点
- 则对于每个消息符号，其祖先路径上的每一个祖先都对应着一次合并，节点相应的概率也会相应地翻倍，直至到达根节点，则该消息符号的深度为$-\log_2 p_i$
- $APL=\frac{\sum_{i=1}^{n}p_i(-\log_2 p_i)}{\sum_{i=1}^{n}p_i}=-\sum_{i=1}^{n}p_i\log_2 p_i$