Homework 7 Solutions

December 4, 2019

5.4

Show that if an undirected graph with n vertices has k connected components, then it has at least $n - k$ edges.

Solution

Each of the k components must contain at least one node. The remaining $n - k$ nodes must be connected to some of the first k nodes via at least one edge each. Thus, the total number of edges is at least $n - k$.

5.10

Let T be an MST of graph G. Given a connected subgraph H of G, show that $T \cap H$ is contained in some MST of H.

Solution

Proof. (By constructing a solution)

We have to show that all edges in $T \cap H$ are included in some MST T_H of H. We construct T_H as follows.

By the cut property, each edge e in $T \cap H$ must be the lightest edge across the two partitions in some cut$(S, V - S)$ of G. Let V_H be the node set of H. This is also true in the cut$(S \cap V_H, V_H - S \cap V_H)$, unless some other edge e' has an equal weight with e. In that case, we can always choose to include e into T_H. In this way, we can add every edge in $T \cap H$ into T_H. Then we add additional edges if necessary to complete T_H. Thus, we must have $T \cap H \subseteq T_H$. □

5.16

Prove the following two properties of the Huffman encoding scheme.

(a) If some character occurs with frequency more than 2/5, then there is guaranteed to be a codeword of length 1.

Solution

Proof. (By contradiction). We assume no codeword of length 1 if some character occurs with frequency $> 2/5$. Using the Huffman encoding scheme, one can create a binary tree T where there is a node $a \in T$ which $Freq(a) > 2/5$. Let T_0 be subtree of T where $a \in T_0$, we have nodes $b, c, d \in T_0$.

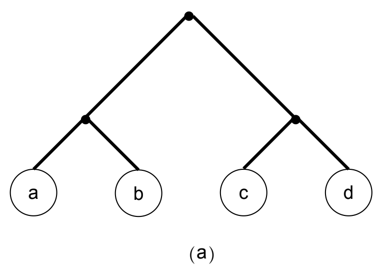

Figure 1: Subtree T_0 of the Huffman tree T.

We use a, b, c, d to represent $Freq(a), Freq(b), Freq(c), Freq(d)$

We have:

$$\begin{cases} a > 2/5 \\ a + b + c + d = 1 \\ a \leq c + d \\ b \leq c + d \\ b, c, d > 0 \end{cases}$$

By $a > 2/5$ and $a \leq c + d$, we get $c + d > 2/5$.

By $c + d > 2/5$, $a > 2/5$ and $a + b + c + d = 1$, we get $b < 1/5$

Then, we need to determine if nodes a, b picked before than c, d to build tree. If a, b picked first, we have $c > a > 2/5$, $d > a > 2/5$, thus will have $a + b + c + d > 6/5 > 1$. So, we determine nodes c, d were picked before a, b, thus get $c < b$ and $d < b$.

Now we have $b + b > c + d > 2/5$ and get $b > 1/5$, which contradicts $b < 1/5$ we got before.

So, the assumption is wrong and the original statement is correct. □

(b) If all characters occur with frequency less than 1/3, then there is guaranteed to be no codeword of length 1.

Solution

Proof. (By contradiction) We assume there exists a codeword of length 1 when all characters occur with frequency $< 1/3$.

Similar to (a), we have a binary tree T and subtree $T_0 \subset T$. Nodes $a, b, c \in T_0$. Let a be a codeword of length 1.

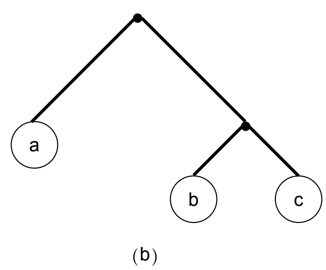

Figure 2: Subtree T_0 of the Huffman tree T.

Use a, b, c to represent $Freq(a), Freq(b), Freq(c)$, we have:

$$\begin{cases} a+b+c=1 \\ a<1/3 \\ b \leq a \\ c \leq a \end{cases}$$

By $a < 1/3$ and $b \leq a$, get $b < 1/3$.

By $a < 1/3$ and $c \leq a$, get $c < 1/3$.

We get $a+b+c < 1/3+1/3+1/3 = 1$ which contradicts $a+b+c = 1$. Thus the assumption is wrong, original statement is correct. □