Information Theory with Applications

Homework Set 2

p. 115, ex. 3.4

Let $\mathbb{A}$ have size K^2, and assign for $a_n \in \mathbb{A}$, $n \in \{1, 2, \dots K^2\}$ the code

$$\phi(a_n) = \begin{cases} n, & n \leq K \\ \{m, l\}, & n = mK + l \\ & \text{with } 1 \leq m \leq K-1, \\ & 0 \leq l \leq K-1 \end{cases}$$

Choose $\mathbb{Q}$ to be uniform prob measure on $\mathbb{A}$, so

$$H_K(\mathbb{Q}) = \log_K |\mathbb{A}| = \log_K K^2 = 2.$$

However,

$$\mathbb{E}[\,l(X)\,] = \mathbb{P}(X \leq K) + 2\,\mathbb{P}(X > K)$$

$$< 2,$$

p. 117 ex 3.10

We prove $H(Q) = \sum_{j=1}^{\#\text{nodes}} q_j H(Q_j)$

by induction over nodes in tree.

$n=1$. Only root and leaves,

$$H(Q) = -q_{i_1} \log q_{i_1} - q_{i_2} \log q_{i_2}$$
$$\cdots - q_{i_k} \log q_{i_k},$$

nothing to show.

$n \to n+1$. Consider tree with $n+1$ nodes.

Pick one, say $n+1$ having only leaves as children. Define a new tree with n nodes by making $(n+1)$th node a leaf with probability $q'_{n+1} = \sum_{l=1}^{k} q_{i_l}$.

For this tree, and corresp. measure Q' we know by assumption

$$H(Q') = \sum_{j=1}^{n} q'_j H(Q'_j).$$

Assume n-th node has former $(n+1)$th as child.

Then full tree has entropy

$$H(Q) = \sum_{j=1}^{n} q_j' \, H(Q_j')$$

$$+ \underbrace{\cancel{q_n'} \, \frac{q_{n+1}}{\cancel{q_n'}} \log \frac{q_{n+1}}{q_n'} - \sum_{i=1}^{\ell} q_{i\ell} \log q_{i\ell}}_{q_{n+1} \underbrace{\sum_{i=1}^{\ell} \frac{q_{i\ell}}{q_{n+1}} \log \frac{q_{i\ell}}{q_{n+1}}}_{H(Q_{n+1})}}$$

$$= \sum_{j=1}^{n+1} q_j \, H(Q_j) \, ;$$

p. 117 ex 3.11

WLOG we assume $l_i > l_j$, otherwise rename, or if $l_i = l_j$, nothing to prove.

Want to show $l_i - l_j \le 1 + \frac{q_j}{q_i}$.

By optimality, we know $q_i \le q_j$, otherwise we could switch and make code better.

Let k-th node be ancestor of i-th and at level $l_k = l_j$. We show

(*) $q_k \ge (l_i - \underbrace{l_k}_{l_j}) q_i$

by induction over $l_i - l_k$.

Let $l_i - l_k = 1$, then $q_k \ge q_i$ by optimality.

Given that (*) holds for $l_i - l_k \le n$, and $l_i - l_k = n+1$, pick two children of k-th node c_1 and c_2, with c_1 being ancestor of i-th node. By assumption (*)

$$q_{c_1} \ge (l_i - l_k - 1) q_i$$

and by optimality

$$q_{c_2} \ge q_k \ge q_{i_2}$$

so

$$q_k \ge (l_i - l_k - 1) q_i + q_i = (l_i - l_k) q_i.$$

Now we conclude ~~for~~ $l_i - l_j = 1 \leq 1 + \frac{q_j}{q_i}$ is clear by optimality, and if $l_i - l_j \geq 2$ there exists a k-th node at level $l_j + 1$ which is ancestor of i-th node and thus

$$q_j \geq q_k \geq (l_i - l_k)\, q_i$$

$$= (l_i - l_j - 1)\, q_i$$

$$\Rightarrow \frac{q_j}{q_i} + 1 \geq l_i - l_j.$$

p. 124 ex. 3.28 Given $q_j = Q(X=j)$, and $q_m \geq \sum_{k=m+1}^{\infty} q_k$ for infinitely many m, construct a Huffman tree.

Take sequence $\{m_n\}$ and create a Huffman tree T_1 for a modified prob. vector Q_1 on alphabet which collapses outcomes $\{X > m_1\}$ into one symbol. By assumption, this symbol can be chosen to have the longest code length. Now construct a Huffman tree T_2 for $\{m_1+1, m_1+2, \ldots m_2, m_2+1\}$ where m_2+1 has prob. $\sum_{k=m_2+1}^{\infty} q_k$. Attach this tree by its root to the symbol for $\{X > m_1\}$ in T_1. The concatenation will still be Huffman because the sum of all probabilities of leaves in T_2 is smaller than those in T_1, without the leaf for $\{X > m_1\}$. Proceed iteratively to build the full tree. By dominated (or monotone) convergence, $\mathbb{E}[L_j]$ for concatenation of $T_1 \circ T_2 \circ \dots \circ T_n$ converges to $\mathbb{E}[L]$, and resulting tree T cannot be further improved by swapping leaves b/c any two leaves eventually belong to a tree $T_1 \circ T_2 \circ \dots \circ T_n$ for some n.