5.1 Since I.C $\subseteq$ U.D,

minimizing over I.C. can do no better than minimizing over U.D.

$\Rightarrow \quad L_1 \geq L_2$

5.2 The Kraft inequality $\sum_{i=1}^{6} D^{-l_i} \leq 1$ must be satisfied.

Trying $D=2$ results in $\sum D^{-l_i} = \frac{7}{4} \not\leq 1$

Trying $D=3$ gives $\sum D^{-l_i} = \frac{26}{27} < 1$

So $D=3$ is the smallest possible D.

The title alludes to the possibility that Martians use base 3 because they may have 3 fingers.

5.3 Generate $z_1, z_2, z_3, \ldots$ iid with each z_ℓ uniform on the integers $0, 1, \ldots, D-1$, and stop generating the sequence at the smallest ℓ such that $(z_1, z_2, \ldots, z_\ell)$ is a codeword.

Codeword i of length l_i has probability D^{-l_i} of being generated.

By the union bound, the probability of stopping is bounded by

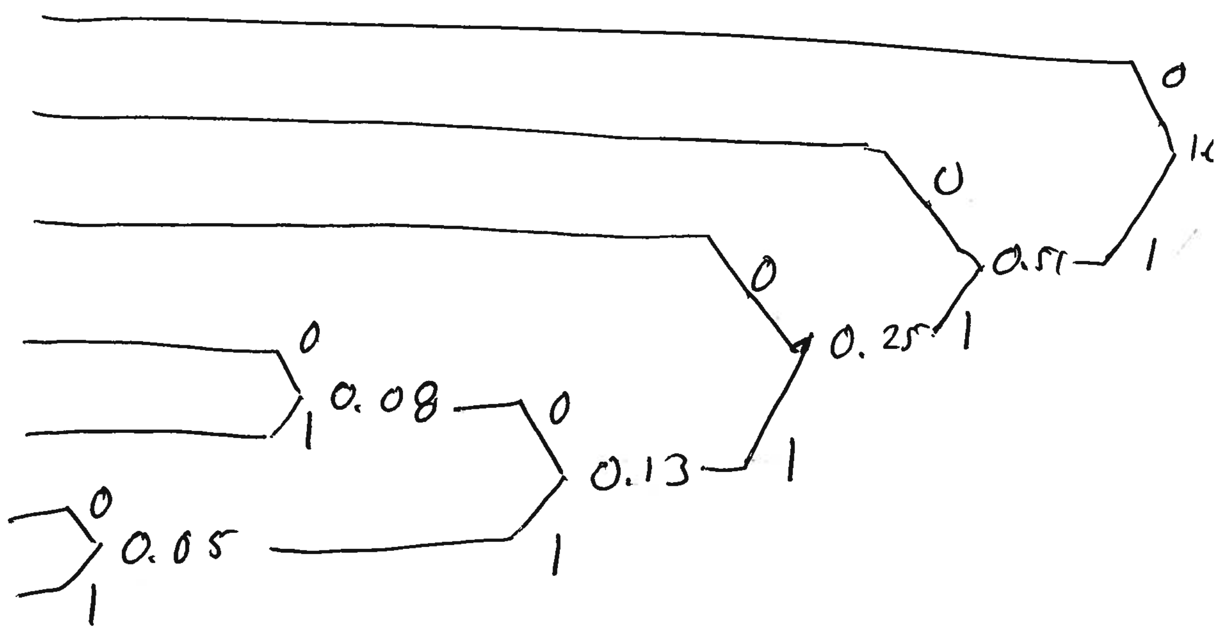
$$P[\text{Stop}] \le \sum_{i=1}^{m} D^{-l_i} < 1$$

$\Rightarrow$ $P[\text{never Stop}] > 0$

So there must be arbitrarily long sequences $z_1, z_2, \ldots$ such that $(z_1, \ldots, z_\ell)$ is not a codeword.

5.4

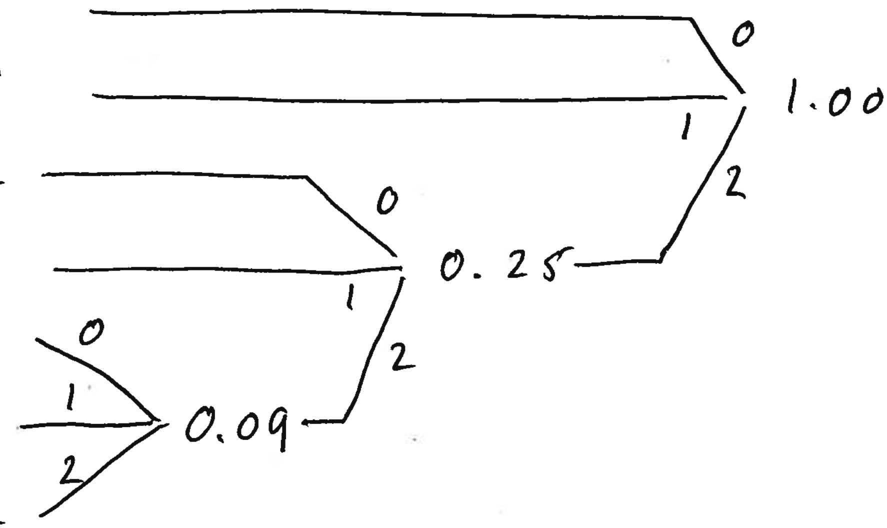
a)
0 0.49
10 0.26
110 0.12
11100 0.04
11101 0.04
11110 0.03
11111 0.02



b) $L = 1 \times 0.49 + 2 \times 0.26 + 3 \times 0.12 + 5 \times 0.04$
$+ 5 \times 0.04 + 5 \times 0.03 + 5 \times 0.02$
$= 2.02$

(By comparison, $H(X) = 2.0128$ bits!)

c)
0 0.49
1 0.26
20 0.12
21 0.04
220 0.04
221 0.03
222 0.02



5.6

a) possible with $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{4}$ $p_3 = \frac{1}{4}$

b) This code could be shortened to
$\{00, 01, 10, 11\}$
Since it is not optimal, it cannot be a Huffman code.

c) The code can be shortened to $\{0, 1\}$ and remain instantaneous. So it is not Huffman

5.9 Consider a binary r.v. X such that

$$P[X=0] = \varepsilon$$

$$P[X=1] = 1-\varepsilon,$$

where $0 < \varepsilon$ is small. Then

$$H(X) = -(1-\varepsilon)\log(1-\varepsilon) - \varepsilon\log\varepsilon$$

is close to 0 since $H(X) \to 0$ as $\varepsilon \to 0$.

But we need 1 bit to encode X.

3.11 Consider reversing the order of the code.

For any received sequence, we work backwards from the end and look for reversed codewords.

Due to the suffix condition, the reversed codewords satisfy the prefix condition, and we can uniquely decode the reversed code.

Thus, to every suffix code, there is a prefix code of the same lengths, and vice versa. Therefore, we cannot achieve any lower average codeword lengths with a suffix than a prefix, and vice versa.

S.18

a) No since 0 is a prefix of 01

b) Yes. Find all "01"s, ~~which~~
and then parse remaining "0"s.

c) Yes because it is uniquely decodable.

5.20

a) We want to minimize $C = \sum p_i c_i l_i$

subject to $\sum 2^{-l_i} \le 1$ $(*)$

The optimal solution $l_1^*, \ldots, l_m^*$ will have equality in $(*)$ as otherwise we can improve upon the optimal solution by decrease one of the l_i^* until $(*)$ is satisfied with equality.

Following the proof of Thm 5.3.1, let

$$r_i = 2^{-l_i^*}$$

$$Q = \sum p_i c_i$$

$$q_i = p_i c_i / Q.$$

Then $\underline{r} = (r_1, \ldots, r_m)$ and $\underline{q} = (q_1, \ldots, q_m)$ are PMFs and

$$C = \sum p_i c_i l_i^*$$

$$= \sum p_i c_i \log \frac{1}{r_i}$$

$$= Q \sum q_i \log \frac{q_i}{r_i} - Q \sum q_i \log q_i$$

$$= Q\, D(\underline{q} \| \underline{r}) + Q\, H(\underline{q})$$

Since $H(\underline{q})$ doesn't depend on the l_i's, the optimal solution is for $D(\underline{q} \| \underline{r}) = 0$

$$\Rightarrow q_i = r_i$$

$$2^{-l_i^*} = \frac{p_i c_i}{Q}$$

$$l_i^* = -\log \frac{p_i c_i}{Q}$$

b) The same procedure, except replace p_i with $\frac{p_i c_i}{Q}$

c) Since C^* is the optimal soln of a) the ignores integer constraints, then

$$C^* \leq C_{Huffman}$$

Now, consider using codeword lengths

$$\lceil l_i^* \rceil = \lceil -\log q_i \rceil \leq l_i^* + 1$$

where l_i^* is optimal soln in a).

Then

$$\begin{aligned} C_{Huffman} &\leq \sum c_i p_i \lceil l_i^* \rceil \\ &\leq \sum c_i p_i (l_i^* + 1) \\ &= Q \sum q_i (l_i^* + 1) \\ &= \left(Q \sum q_i l_i^*\right) + Q \\ &= C^* + Q \end{aligned}$$

5.22 The maximal length of a codeword in an optimal prefix code is m since the Huffman tree cannot have more than m levels.

Therefore, there are a finite # K of codes to consider. Call them $C_1, \ldots, C_K$, and let the length of the codewords be

$$
\begin{array}{lll}
l_{1,1}, \ldots, l_{1,m} & \text{for} & C_1 \\
\vdots & \vdots & \vdots \\
l_{K,1}, \ldots, l_{K,m} & \text{for} & C_K
\end{array}
$$

The average length of code k with source distribution p is

$$L_k(p) = \sum_{i=1}^{m} l_{k,i}\, p_i,$$

and

$$L(p) = \min_{k=1,\ldots,K} L_k(p)$$

Since $L_k(p)$ is linear in p, $L_k(p)$ is continuous in p.
Since $L(p)$ is the min of K functions that are continuous in p, $L(p)$ is continuous in p.