Student Number: [redacted] Name: Bryan Hoang

4. (10 points)

(a) **Answer:**

Since the prefix code is first-order, $n = 1$.

Part 1. (p_1)

Calculating $\overline{R}(\mathcal{C}, p_1)$ yields

$$\begin{aligned}
\overline{R}(\mathcal{C}, p_1) &= \sum_{x \in \mathcal{X}} p_1(x) l(f(x)) \\
&= 3 \cdot \frac{1}{4} \cdot 1 + 4 \cdot \frac{1}{16} \cdot 2 \\
&\boxed{= 1.25 \, \frac{\text{quaternary units}}{\text{source symbol}}}.
\end{aligned}$$

The theoretical limit of $\overline{R}(\mathcal{C}, p_1)$, the entropy rate $H(\mathcal{X})$ of the DMS, is

$$\begin{aligned}
H_4(\mathcal{X}) &= H_4(X) \\
&= -\sum_{x \in \mathcal{X}} p_1(x) \log_4 p_1(x) \\
&= -\left(3 \cdot \frac{1}{4} \log_4 \frac{1}{4} + 4 \cdot \frac{1}{16} \log_4 \frac{1}{16}\right) \\
&\boxed{= 1.25 \, \text{quaternary units}}.
\end{aligned}$$

Thus, we have that $\boxed{\overline{R}(\mathcal{C}, p_1) = H_4(\mathcal{X})}$.

Part 2. (p_2)

Calculating $\overline{R}(\mathcal{C}, p_2)$ yields

$$\begin{aligned}
\overline{R}(\mathcal{C}, p_2) &= \sum_{x \in \mathcal{X}} p_2(x) l(f(x)) \\
&= \frac{1}{4} \cdot 1 + 2 \cdot \frac{1}{8} \cdot 1 + 4 \cdot \frac{1}{8} \cdot 2 \\
&\boxed{= 1.5 \, \frac{\text{quaternary units}}{\text{source symbol}}}.
\end{aligned}$$

The theoretical limit of $\overline{R}(\mathcal{C}, p_2)$, the entropy rate $H(\mathcal{X})$ of the DMS, is

$$\begin{aligned}
H_4(\mathcal{X}) &= H_4(X) \\
&= -\sum_{x \in \mathcal{X}} p_2(x) \log_4 p_2(x) \\
&= -\left(\frac{1}{4} \log_4 \frac{1}{4} + 6 \cdot \frac{1}{8} \log_4 \frac{1}{8}\right) \\
&\boxed{= 1.375 \, \text{quaternary units}}.
\end{aligned}$$

Student Number: [redacted] Name: Bryan Hoang

Thus, we have that $\boxed{H_4(\mathcal{X}) \leq \overline{R}(\mathcal{C}, p_2) < H_4(\mathcal{X}) + 1}$.

(b) **Answer:**

Qualitatively, the source distribution p_1 is 4-adic which is a nice property to have since certain codes (e.g., Shannon code) will be optimal with this distribution.

In terms of compression efficiency, the corresponding average code rate for p_2 is slightly greater than its entropy rate by part (a). But $\boldsymbol{p_1}$ **is the preferable distribution** for the source since its corresponding average code rate is the same as its entropy rate and that $\overline{R}(\mathcal{C}, p_1) < \overline{R}(\mathcal{C}, p_2)$, making it more efficient.