Student Number: <u>20053722</u>   Name: <u>Bryan Hoang</u>

5. **Answer:**

(a) *Proof.* Let X and Y by r.v.'s that take values in $\mathcal{X}$. The nonnegativity of $H(\cdot)$ implies that the sum to two conditional entropies is also nonnegative. Thus, we have that $\boxed{\Delta(X, Y) \geq 0}$, satisfying the nonnegativity of a metric.

By the commutative property of addition, we have

$$\begin{aligned} \Delta(X, Y) &= H(X|Y) + H(Y|X) \\ &= H(Y|X) + H(X|Y) \\ &= \Delta(Y, X) \end{aligned}$$

which implies that $\boxed{\Delta(X, Y) = \Delta(Y, X)}$ satisfies the symmetric property of a metric.

$\Delta(X, Y) = 0$ **if and only if** the value of X is completely determined by the value of Y and vice versa. i.e., $\exists f$ invertible such that $X = f(Y)$.

To prove that Δ satisfies the triangle inequality by contradiction, suppose otherwise. Then

$$\Delta(X, Y) + \Delta(Y, Z) < \Delta(X, Z)$$
$$H(X|Y) + H(Y|X) + H(Y|Z) + H(Z|Y) < H(X|Z) + H(Z|X)$$
$$\cancel{H(X|Y,Z)} + I(X; Z|Y) + H(Y|X, Z) + \cancel{I(Y;Z|X)} < \cancel{H(X|Y,Z)} + \cancel{I(X;Y|Z)}$$
$$+ H(Y|X, Z) + \cancel{I(X;Y|Z)} + \cancel{H(Z|X,Y)} + I(X; Z|Y) \quad + \cancel{H(Z|X,Y)} + \cancel{I(Y;Z|X)}$$
$$I(X; Z|Y) + H(Y|X, Z) + H(Y|X, Z) + I(X; Z|Y) < 0$$

which contradicts the nonnegativity of entropy and mutual information. Thus, the initial assumption was incorrect, and so Δ **does** satisfy the triangle inequality for being a metric. $\qquad\square$

(b) *Proof.* For the first inequality, suppose it is false. Then

$$\begin{aligned} H(X) - H(Y) &> \Delta(X, Y) \\ H(X) - H(Y) &> H(X|Y) + H(Y|X) \\ H(X) - \cancel{H(Y)} &> H(X, Y) - H(X) + H(X, Y) - \cancel{H(Y)} \qquad \text{by chain rule} \\ H(X) &> H(X, Y) \end{aligned}$$

which contradicts the fact that joint entropies are always greater than the marginal entropies. Therefore, $H(X) - H(Y) \leq \Delta(X, Y)$. WLOG, a similar argument can be made for $H(Y) - H(X) \leq \Delta(X, Y)$

For the second inequality, also suppose it is false. Then

$$\begin{aligned} H(X_1|X_2) - H(Y_1|Y_2) &> \Delta(X_1, Y_1) + \Delta(X_2, Y_2) \\ H(X_1|X_2) - H(Y_1|Y_2) &> H(X_1|Y_1) + H(Y_1|X_1) + H(X_2|Y_2) + H(Y_2|X_2) \\ H(X_1|X_2) &> \big(H(X_1|Y_1) + H(Y_2|X_2) + H(Y_1|Y_2)\big) + H(Y_1|X_1) + H(X_2|Y_2) \end{aligned}$$
$$(1)$$

But we also have that

$$\begin{aligned}
H(X_1|Y_1) &\geq H(X_1|Y_1,Y_2) \\
H(Y_2|X_2) &\geq I(X_1;Y_1;Y_2|X_2) \\
H(Y_1|Y_2) &\geq I(X_1;Y_1|X_2,Y_2) \\
H(X_1|X_2) &\leq H(X_1|Y_1,Y_2) + I(X_1;Y_1;Y_2|X_2) + I(X_1;Y_1|X_2,Y_2) \tag{2}
\end{aligned}$$

Then (1) and (2) contradict each other. Thus, the initial assumption about the second inequality was incorrect, and so $H(X_1|X_2) - H(Y_1|Y_2) \leq \Delta(X_1,Y_1) + \Delta(X_2,Y_2)$. WLOG, a similar argument can be made for $H(Y_1|Y_2) - H(X_1|X_2) \leq \Delta(X_1,Y_1) + \Delta(X_2,Y_2)$.　$\square$