Teoria dell'Informazione e della Trasmissione — Appello del 12.9.2016

Esercizio A. Dire se ciascuna delle seguenti asserzioni è vera o falsa giustificando la risposta.

1. $H(X) + H(Y) \leq H(X, Y)$
2. Se X, Y sono variabili casuali indipendenti, allora $H(X - Y) \geq H(X)$.

Esercizio B. Sia $\langle \mathcal{X}, p \rangle$ tale che $\mathcal{X} = \{1, 2, 3\}$ con $p(1) = \frac{1}{2}$, $p(2) = \frac{1}{4}$ e $p(3) = \frac{1}{4}$. Si consideri il codice $C(1) = 0$, $C(2) = 01$ e $C(3) = 11$. Rispondere alle domande seguenti giustificando la risposta.

1. C è non singolare?
2. C è istantaneo?
3. Si calcoli il codice di Shannon binario per $\langle \mathcal{X}, p \rangle$.

Esercizio C. Calcolare la lunghezza media di un codice di Huffman binario per la sorgente i cui simboli hanno probabilità pari alle frequenze dei caratteri (inclusi gli spazi) che compaiono nella frase

```
she sells sea shells
```

Esercizio D. Siano X, Y_1 e Y_2 tre variabili casuali tali che per ogni x, y_1, y_2 vale

$$p(y_1 \mid x) = p(y_2 \mid x) \qquad \text{e} \qquad p(y_1, y_2 \mid x) = p(y_1 \mid x) p(y_2 \mid x) \; .$$

Si dimostri che $I\big(X, (Y_1, Y_2)\big) = 2I(X, Y_1) - I(Y_1, Y_2)$.
Nota: si osservi che $p(y_1 \mid x) = p(y_2 \mid x)$ per ogni x, y_1, y_2 implica $p(y_1) = p(y_2)$.

Soluzione esercizio A.

- Falso. Infatti, $H(X) + H(Y) \geq H(X, Y)$ dato che $I(X, Y) = H(X) + H(Y) - H(X, Y) \geq 0$.
- Vero. Infatti, $H(X - Y) \geq H(X - Y \mid Y) = H(X)$ dove l'ultima disuguaglianza vale perché $f(X) = X - Y$ è una funzione biunivoca una volta che Y è noto e quindi l'entropia di $f(X)$ è pari all'entropia di X nello spazio condizionato su Y.

Soluzione esercizio B.

- Sì, infatti $C(x)$ è unico per ogni $x \in \mathcal{X}$.
- No, infatti $C(1)$ è un prefisso di $C(2)$.
- Dato che $\log_2 \frac{1}{p(1)} = 1$ e $\log_2 \frac{1}{p(2)} = \log_2 \frac{1}{p(3)} = 2$ esiste un codice istantaneo con lunghezze $\{1, 2, 2\}$. Per esempio, $C'(1) = 0$, $C'(2) = 10$ e $C'(3) = 11$.

Soluzione esercizio C. Le frequenze sono

$$\begin{pmatrix} A & H & \text{sp.} & S & L & E \\ \frac{1}{20} & \frac{2}{20} & \frac{3}{20} & \frac{6}{20} & \frac{4}{20} & \frac{4}{20} \end{pmatrix}$$

Le lunghezze della parole di un codice di Huffman C sono pertanto

$$\begin{pmatrix} 2 & 2 & 3 & 3 & 3 & 3 \\ \frac{4}{20} & \frac{4}{20} & \frac{3}{20} & \frac{2}{20} & \frac{2}{20} & \frac{2}{20} \end{pmatrix}$$

e la lunghezza media è pari a $\frac{49}{20}$.

Soluzione esercizio D. Usando $H(Y_1, Y_2 \mid X) = H(Y_1 \mid X) + H(Y_2 \mid X)$ (per l'indipendenza di Y_1 e Y_2 nello spazio condizionato su X) e le uguaglianze $H(Y_1) = H(Y_2)$, $H(Y_1 \mid X) = H(Y_2 \mid X)$ otteniamo

$$\begin{aligned} I\big(X, (Y_1, Y_2)\big) &= H(Y_1, Y_2) - H(Y_1, Y_2 \mid X) \\ &= H(Y_1) + H(Y_2) - I(Y_1, Y_2) - H(Y_1 \mid X) - H(Y_2 \mid X) \\ &= 2H(Y_1) - I(Y_1, Y_2) - 2H(Y_1 \mid X) \\ &= 2I(X, Y_1) - I(Y_1, Y_2)\ . \end{aligned}$$