Soluciones y comentarios sobre el parcial 3/10/2018

Enunciados en púrpura, comentarios en azul.

1. Sea $Y = XZ$, donde X toma valores en $(1, 2 \cdots n)$, con probabilidades $p_i = P(X = i)$, y Z (independiente de X) toma valores 0 y 1 , $P(Z = 1) = a$.

   a) Encontrar $H(Y)$, en función de p_i y a

   b) Si Y es la salida de un canal sin memoria, con entrada X, y a dado, ¿Cuál es la capacidad C?

   a) Notando que $H(Z|Y) = 0$ (saber el resultado me permite saber Z), podemos escribir $H(Z, Y) = H(Y) + H(Z|Y) = H(Z) + H(Y|Z)$ y entonces

$$H(Y) = H(Z) + H(Y|Z) = h(a) + aH(Y|Z = 1) + (1 - a)H(Y|Z = 0) \tag{1}$$

$$H(Y) = h(a) + aH(X) \tag{2}$$

   con $H(X) = -\sum p_i \log(p_i)$

   También es válido, aunque más engorroso y menos elegante, lo que hizo la mayoría: calcular $p(Y)$ y calcular la entropía:

$$H(Y) = -(1 - a)\log(1 - a) - \sum_i p_i a \log(p_i a) \tag{3}$$

   que, abriendo el logaritmo, se puede escribir como (2).

   b) Sabemos $I(X; Y) = H(Y) - H(Y|X)$. Pero $H(Y|X) = \sum p_i H(Y|X = i) = h(a)$

   Entonces, de (2) :

$$I(X; Y) = aH(X) \tag{4}$$

   Y $C = \max I(X; Y) = a \log(n)$ que se alcanza para distribución uniforme de X.

2. Una fuente X sin memoria emite símbolos de un alfabeto de 150 símbolos, equiprobables. Sea L_h la longitud media del código de Huffman binario (sin extensión). Encontrar cotas de L_h. Explicar qué sucede con la cota si se toman extensiones de la fuente.

   Sabemos que dsabemos todo código instantáneo cumple $L_h \geq H$ y sabemos (primer t. de Shannon) que existe un código instantáneo con $L_h < H + 1$. Como Huffman es instantáneo y óptimo, debe estar entre esas cotas. Como $H(X) = 7.23$

$$7.23 \leq L_h < 8.23 \tag{5}$$

   (Notar que la segunda desigualdad es estricta)

En este caso podemos afinar la cota superior, porque sabemos que existe un código trivial de longitud fija 8 que cubre 256 símbolos. Nuestro código de Huffman no puede andar peor que ese (problema 4.c), entonces

$$7.23 \leq L_h \leq 8 \tag{6}$$

Si tomamos extensiones, el mismo razonamiento que lleva a (5) se puede aplicar a los macrosímbolos ($n-$tuplas), con lo que para n grande la cota se acerca a $H(X)$:

$$7.23 \leq L_h^{(n)} < 7.23 + \frac{1}{n} \tag{7}$$

3. Demostrar la regla de la cadena de la información mutua

$$I(X, Y; Z) = I(X; Z) + I(Y; Z|X)$$

Sabemos que $I(X; Z) = H(Z) - H(Z|X)$ y $I(Y; Z|X) = H(Z|X) - H(Z|X, Y)$. Entonces:

$$I(X; Z) + I(Y; Z|X) = H(Z|X) - H(Z|X, Y) = I(X, Y; Z) \tag{8}$$

4. ¿Verdadero o falso? Desarrollar, señalar si hay algún error, justificar.

   a) $I(X; X) = H(X)$

   VERDADERO. Trivial. $I(X; X) = H(X) - H(X|X) = H(X)$

   b) Sean X, Y variables discretas, no necesariamente independientes, ambas con soporte en $\{1, 2, 3\}$. Entonces $H(X|Y) \leq H(X|Y = j)$ para cualquier $j \in \{1, 2, 3\}$

   FALSO. Como $H(X|Y) = \sum_j P(Y = j)H(X|Y = j)$, esto dice que $H(X|Y)$ es un promedio ponderado de los valores $H(X|Y = j)$ El promedio no puede ser menor a todos los valores promediados, a lo sumo será igual si todos los valores son iguales. En general, tendremos que $H(X|Y = j)$ a veces será mayor y a veces menor que $H(X|Y)$.

   Un contraejemplo, $Y = XZ$ con X, Z Bernoulli $(0, 1)$. Claramente $H(Y|X = 0) = 0$, pero $H(Y|X) > 0$.

   c) Si una fuente tiene 8 símbolos, la longitud promedio del código de Huffman binario nunca será mayor a 3.

   VERDADERO. El código de Hufmman es óptimo entre los instantáneos. Ahora bien, para 8 símbolos conocemos un código trivial instantáneo que tiene longitud 3, no importa la distribución de entrada. Entonces, no puede ser que nuestro código de Huffman tenga $L_h > 3$: no sería óptimo.

   (Muchos escribieron $H \leq L_h < H + 1$, pero por este lado no concluimos nada.)

   d) El código $\{01, 011, 11\}$ es unicamente decodificable.

   VERDADERO. No hay ambigüedad con las extensiones, porque los ceros delimitan los símbolos.

e) El código $\{10, 01, 11, 101\}$ no es instantáneo, pero es unicamente decodificable.

FALSO. No es instánteneo ni tampoco UD. Por ejemplo la secuencia 101101 puede decodidicarse de dos maneras diferentes.

f) Un código binario asigna a los símbolos (A, B, C, D, E) códigos de longitudes $(3, 1, 3, 3, 4)$. Entonces no puede ser un código de Huffman.

VERDADERO. Por su construcción (que es su definición) el código de Huffman asigna los dos símbolos menos probables al nivel más bajo del árbol, con lo cual resultan hojas "hermanas", con la misma longitud máxima. Aquí tenemos una sola hoja con longitud máxima (4). No puede ser un código de Huffman.

5. Calcular la capacidad de los canales dados por estas matrices de transición y especificar la distribución de entrada $p(X)$ correspondiente.

a. $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Llamando $p_i = P(X = i)$, $i = 1, 2, 3$

$I(X, Y) = H(Y) - H(Y|X) = H(p_1 + p_2, p_3) - 0$. Esto es maximizado la salida es uniforme (sobre los dos valores que tienen probabilidad no nula), lo cual occure cuando $p_3 = p_1 + p_2 = \frac{1}{2}$,y entonces $C = \log(2) = 1$ bit.

Esto también puede verse intuitivamente: las dos primeras entradas resultan indistinguibles, por lo que pueden considerarse una sola, con lo cual tenemos un canal binario sin ruido.

También es aceptable hacerlo de otras maneras, por ejemplo con $C = \log(2_1^C + 2_2^C) = \log(2^0 + 2^0)$

b $\begin{pmatrix} 1-\alpha & 0 & 0 & \alpha \\ 0 & 1-\alpha & 0 & \alpha \\ 0 & 0 & 1-\alpha & \alpha \end{pmatrix}$

Esto es un canal (no binario) de borrado. Y es esencialmente igual al del problema 1.

$$I(X;Y) = H(Y) - H(Y|X)$$
$$H(Y) = h(\alpha) + (1-\alpha)H(X)$$
$$H(Y|X) = h(\alpha)$$
$$I(X;Y) = (1-\alpha)H(X)$$
$$C = (1-\alpha)\log(3)$$

que se alcanza con entrada uniforme.

Otra manera (sin usar el problema 1):

$$I(X;Y) = H(X) - H(X|Y)$$

$$H(X|Y) = \sum_Y P(Y)H(X|Y) = P(Y=4)H(X|Y=4)$$

Pero $P(Y=4) = \alpha$ y $p(X|Y=4) = \frac{P(Y=4|X)p(X)}{P(Y=4)} = P(X)$, entonces $H(X|Y=4) = H(X)$ y

$$H(X|Y) = \alpha H(X)$$
$$I(X;Y) = (1-\alpha)H(X)$$

etc

Observaciones, errores comunes/notables/descalificadores:

- Varios confundieron código "únicamente decodificable" con código no singular.
- Muchos afirmaron que, como se cumple la desigualdad de Kraft (Mc Millan) el código es instantáneo (o unicamente decodificable). Eso es una barbaridad. Piensen en el código $\{A \to 0, B \to 01\}$ o, peor, en $\{A \to 0, B \to 00\}$
- Varios afirmaron que un código de Huffman tiene un arbol de tipo escalera, con hojas con todas las longitudes entre el minimo y el maximo. Falso.
- Varios afirmaron que un código de Huffman, por ser óptimo, tiene longitudes $l_i^* = -\log(p_i)$ (y por lo tanto $L_h = H$). Ambas son falsas.
- En el cálculo de las capacidades, varios obtuvieron valores absurdos. Una cosa es un error de cálculo, otra es el error conceptual. Por ejemplo, poner $C = 2$bits en un canal con tres entradas y tres salidas. También el resultado $C = 0$ debería resultar absurdo a simple vista en los canales del examen, debería ser claro que se puede trasmitir información. Lo mismo para los que en el $5b$ obtuvieron una capacidad que no depende de α.
- En el 1 algunos escribieron cosas como $H(X) = ap_i$ Eso no solo está mal, sino que no tiene el menor sentido. $H(X)$ es un numero, p_i es una función de variable discreta (índice i)!