

TRANSMISIÓN DE DATOS 2006/07		
Práctica 2: Codificación Aritmética.		Grupo
		Puesto
Apellidos, nombre		Fecha
Apellidos, nombre		30 Octubre/ 2 Noviembre 2006

El objetivo de esta práctica es introducir al alumno en los fundamentos de la codificación aritmética y su estudio bajo determinadas circunstancias.

En el fichero codificacionAritmetica.zip se encuentran los archivos necesarios para la realización de esta práctica.

Justo antes del final de la práctica, comprima el proyecto en un fichero TxDatosP2GXX.zip (donde G indicará el grupo, A o B, y XX el puesto del laboratorio asignado), conéctese al sistema de entrega de prácticas de la Intranet y entréguelo en el grupo que corresponda (A o B). Guárdese adicionalmente una **copia personal**, para posibles futuras reutilización del código en prácticas posteriores.

1 Introducción

La **codificación aritmética** se fundamenta en el conocimiento a priori de las probabilidades de símbolo. Esto, lo convierte en un método estadístico.

Esta codificación se basa en la codificación de una secuencia de símbolos de una manera conjunta. Para ello partiremos de una colección de intervalos entre $[0,1)$ cada uno de los cuales estará asociado a cada uno de los posibles símbolos de fuente, teniendo una magnitud proporcional a la probabilidad de símbolo. Asociaremos cada uno de los posibles símbolos de fuente (S_i) con un intervalo (I_i), $\{S_i, I_i\}$ de tal forma que $\bigcup_{\forall i} I_i \equiv [0,1)$ y $\bigcap_{\forall i} I_i \equiv \Phi$. $I_i = [a_i, b_i)$. Definimos un intervalo inicial $[a_{actual}, b_{actual}) = [0,1)$

Para cada símbolo a codificar, modificaremos de la siguiente forma:

$$S_i \Rightarrow [a_{nuevo}, b_{nuevo}) = [a_{actual} + (b_{actual} - a_{actual}) \cdot a_i, a_{actual} + (b_{actual} - a_{actual}) \cdot b_i),$$

$$[a_{actual}, b_{actual}) = [a_{nuevo}, b_{nuevo})$$

El valor final de uno de los límites del intervalo $[a_{actual}, b_{actual})$ será suficiente para la recuperación de la secuencia de símbolos. En nuestro caso, tomaremos el límite inferior.

Por ejemplo:

$$\{S_1, S_2, S_3\}$$

$$p_1 = 0.5, p_2 = 0.2 \text{ y } p_3 = 0.3$$

$$I_1 = [0, 0.5), I_2 = [0.5, 0.7) \text{ y } I_3 = [0.7, 1)$$

$$[a_{\text{actual}}, b_{\text{actual}}) = [0, 1)$$

$$\text{De forma general: } I_i = [a_i, b_i)$$

Codificaremos la secuencia de símbolos:

$$S_1 S_2 S_1 S_3 S_2$$

$$S_1 \Rightarrow [a_{\text{nuevo}}, b_{\text{nuevo}}) = [0 + (1-0) \cdot 0, 0 + (1-0) \cdot 0.5) \Rightarrow [0, 0.5), [a_{\text{actual}}, b_{\text{actual}}) = [0, 0.5)$$

$$S_2 \Rightarrow [a_{\text{nuevo}}, b_{\text{nuevo}}) = [0 + (0.5-0) \cdot 0.5, 0 + (0.5-0) \cdot 0.7) \Rightarrow [0.25, 0.35), [a_{\text{actual}}, b_{\text{actual}}) = [0.25, 0.35)$$

$$S_1 \Rightarrow [a_{\text{nuevo}}, b_{\text{nuevo}}) = [0.25 + (0.35-0.25) \cdot 0, 0.25 + (0.35-0.25) \cdot 0.5) \Rightarrow [0.25, 0.3),$$

$$[a_{\text{actual}}, b_{\text{actual}}) = [0.25, 0.3)$$

$$S_3 \Rightarrow [a_{\text{nuevo}}, b_{\text{nuevo}}) = [0.25 + (0.3-0.25) \cdot 0.7, 0.25 + (0.3-0.25) \cdot 1) \Rightarrow [0.2850, 0.3),$$

$$[a_{\text{actual}}, b_{\text{actual}}) = [0.2850, 0.3)$$

$$S_2 \Rightarrow [a_{\text{nuevo}}, b_{\text{nuevo}}) = [0.2850 + (0.3-0.2850) \cdot 0.5, 0.2850 + (0.3-0.2850) \cdot 0.7) \Rightarrow [0.2925, 0.2955)$$

$$, [a_{\text{actual}}, b_{\text{actual}}) = [0.2925, 0.2955)$$

Luego, el valor a codificar será **0.2925**

La decodificación de una codificación se regirá por la siguiente regla, inversa de la de la codificación:

$$a_{\text{nuevo}} = a_{\text{actual}} + (b_{\text{actual}} - a_{\text{actual}}) a_i \Rightarrow a_i = \frac{a_{\text{actual}} - a_{\text{nuevo}}}{b_{\text{actual}} - a_{\text{actual}}}$$

Siendo $[a_{\text{actual}}, b_{\text{actual}})$ uno de los intervalos de los símbolos, el cual al que pertenece a_{actual} .

Para el ejemplo anterior, la el proceso de decodificación sería el siguiente:

$$a_0 = 0.2925 \Rightarrow a_{1actual} = 0.2925 \in I_1 \Rightarrow [a_{actual}, b_{actual}] = [0, 0.5) \Rightarrow S_1$$

$$a_1 = \frac{0.2925 - 0}{0.5 - 0} = 0.5850 \Rightarrow a_{2actual} = 0.5850 \in I_2 \Rightarrow [a_{actual}, b_{actual}] = [0.5, 0.7) \Rightarrow S_2$$

$$a_2 = \frac{0.5850 - 0.5}{0.7 - 0.5} = 0.4250 \Rightarrow a_{3actual} = 0.4250 \in I_2 \Rightarrow [a_{actual}, b_{actual}] = [0, 0.5) \Rightarrow S_1$$

$$a_3 = \frac{0.4250 - 0}{0.5 - 0} = 0.85 \Rightarrow a_{4actual} = 0.85 \in I_3 \Rightarrow [a_{actual}, b_{actual}] = [0.7, 1) \Rightarrow S_3$$

$$a_3 = \frac{0.85 - 0.7}{1 - 0.7} = 0.5 \Rightarrow a_{5actual} = 0.5 \in I_2 \Rightarrow [a_{actual}, b_{actual}] = [0.5, 0.7) \Rightarrow S_2$$

$$a_4 = \frac{0.5 - 0.5}{0.5 - 0} = 0 \Rightarrow FIN$$

2 Codificación Aritmética

Implementar la función matlab:

```
function [X]=codificacionAritmetica(P,Secuencia, d, n)
```

- P: Un vector fila con la probabilidad de los símbolos.
- Sec: Un vector fila con la secuencia de símbolos a codificar (el índice de cada uno de los símbolos correspondientes a las probabilidades de P).
- d: Número de decimales que se van a considerar para obtener el número entero para la codificación.
- n: El número de bits que la función podrá utilizar para codificar la secuencia.
- X: La secuencia codificada como un número binario de 'n' bits.

La función implementada deberá mostrar además por pantalla el valor decimal a codificar obtenido del cálculo de los intervalos y el valor entero obtenido a partir de dicho valor decimal como paso previo a la codificación binaria.

Nota: Para el paso de un valor decimal a binario se puede utilizar la función `dec2bin(d, n)`, donde 'd' es un número decimal cualquiera y 'n' el número mínimo de bits para codificar 'd'. Si no se especifica 'n' se realizará la codificación con el mínimo posible número de bits. Consultar la ayuda de `Matlab` para más información.

Notaz: La conocida problemática en cuanto a la precisión de representación de números en `Matlab` puede llevar a errores en la ejecución de esta función. Se recomienda aplicar la función 'redondea' proporcionada a cada elemento cuando se realicen comparaciones entre números.

2.1 Ejercicio 1

- 2.1.1 Calcular mediante la función implementada, `codificaciónAritmética`, la codificación correspondiente al ejemplo de la introducción. Utilice los valores 'd' y 'n' que considere oportunos. ¿Son coherentes los resultados? Comentar los resultados.

2.2 Ejercicio 2. Fuente A

Dada la fuente A con 4 símbolos $S \equiv \{S_1, S_2, S_3, S_4\}$ con probabilidades $P \equiv \{p_1, p_2, p_3, p_4\} \equiv \{0.1, 0.2, 0.4, 0.3\}$

- 2.2.1 Calcular analíticamente el valor a codificar, resultante de una codificación aritmética, para la secuencia S_1, S_2, S_4, S_2

- 2.2.2 Calcular mediante la función implementada, `codificaciónAritmética`, ese mismo valor. ¿Son coherentes los resultados? ¿Cuáles son los valores mínimos de 'd' y 'n' necesarios para poder codificar la secuencia? ¿Por qué?

2.3 Ejercicio 3. Fuente B

Dada la fuente B con 4 símbolos $S \equiv \{S_1, S_2, S_3, S_4\}$ con probabilidades $P \equiv \{p_1, p_2, p_3, p_4\} \equiv \{0.2, 0.1, 0.1, 0.6\}$

2.3.1 Calcular analíticamente el valor a codificar, resultante de una codificación aritmética, para la secuencia S_1, S_3, S_2

2.3.2 Calcular mediante la función implementada, `codificacionAritmetica`, ese mismo valor. ¿Son coherentes los resultados? ¿Cuáles son los valores mínimos de 'd' y 'n' necesarios para poder codificar la secuencia?

2.3.3 Calcular mediante la función implementada, `codificaciónAritmética` el valor codificado, resultante de una codificación aritmética, para la secuencia $S_1, S_3, S_2, S_1, S_2, S_3$. ¿Cuáles son los valores mínimos de 'd' y 'n' necesarios para poder codificar la secuencia?

2.4 *Comparativa*

2.4.1 Comente las ventajas e inconvenientes de aumentar o disminuir los valores 'd' y 'n' a la hora de realizar la codificación aritmética.

2.4.2 ¿Cree que existe alguna dependencia entre los valores 'd' y 'n' o por el contrario no dependen el uno del otro? ¿Qué otros factores podrían influir en dichos valores? Justifique la respuesta.

2.4.3 Compare los resultados obtenidos para las fuentes A y B y los distintos ejemplos ejecutados. ¿Qué diferencias y/o semejanzas se pueden encontrar? ¿A qué cree que se deben dichas semejanzas y/o diferencias?

3 Decodificación Aritmética

Implementar la función matlab:

```
function [V]=decodificacionAritmetica(B, P, d)
```

- B: Secuencia binaria correspondiente a un valor codificado mediante codificación aritmética.
- P: Un vector fila con la probabilidad de los símbolos codificados.
- d: Número de decimales que se han considerado para obtener el número entero para la codificación aritmética.
- V: Vector fila con la secuencia decodificada

La función implementada deberá mostrar además por pantalla el valor decimal obtenido a partir de la secuencia binaria como paso previo a la decodificación aritmética.

Nota: Para el paso de un valor binario a decimal se puede utilizar la función `bin2dec(b)`, donde 'b' es una secuencia binaria. Consultar la ayuda de `Matlab` para más información.

Nota2: La conocida problemática en cuanto a la precisión de representación de números en Matlab puede llevar a errores en la ejecución de esta función. Se recomienda aplicar la función 'redondea' proporcionada a cada elemento cuando se realicen comparaciones entre números.

Notas: La condición de parada en el bucle del algoritmo de decodificación aritmética puede no cumplirse nunca si no se han codificado los símbolos con la precisión suficiente. Para evitar esto se recomienda añadir un contador de iteraciones al bucle de forma que se detenga la ejecución cuando se supere un límite razonable (por ejemplo 50 iteraciones).

3.1 Ejercicio 1. Fuente A

Dada la misma fuente A del punto 2.2 con 4 símbolos $S \equiv \{S_1, S_2, S_3, S_4\}$ con probabilidades $P \equiv \{p_1, p_2, p_3, p_4\} \equiv \{0.1, 0.2, 0.4, 0.3\}$

- 3.1.1 Calcular mediante la función implementada, `codificacionAritmetica`, el valor codificado, C, resultante de una codificación aritmética, para la secuencia S_1, S_2, S_4, S_2 . Utilice los valores 'd' y 'n' que considere oportunos. A continuación decodifique la secuencia obtenida C utilizando la función implementada `decodificacionAritmetica` con el mismo valor 'd' utilizado para la codificación. ¿Qué valores 'd' y 'n' ha utilizado? Comente los resultados.

3.1.2 Repita el experimento anterior utilizando un valor 'd' tanto para codificación como para decodificación igual a 3. ¿Se reconstruye correctamente la secuencia original? ¿Por qué? Exponga los resultados obtenidos.

3.1.3 Repita el experimento anterior utilizando un valor 'd' tanto para codificación como para decodificación igual a 6. ¿Se reconstruye correctamente la secuencia original? ¿Por qué? Exponga los resultados obtenidos.

3.1.4 ¿Qué ocurre cuando los valores asignados a 'd' son diferentes en el proceso de codificación y en el de decodificación?