

TRANSMISIÓN DE DATOS 2006/07		
Práctica 1: Codificación Huffman		Grupo
		Puesto
Apellidos, nombre		Fecha
Apellidos, nombre		23/26 de Octubre de 2006

El objetivo de esta práctica es familiarizar al alumno con la codificación de fuentes (sin pérdidas) mediante el algoritmo Huffman.

En el fichero huffman.zip se encuentra la función de matlab `huffmanRecursivo` (fichero `huffmanRecursivo.m`) que permite calcular los códigos huffman recursivamente correspondientes a una lista de probabilidades de símbolos. Para llevar a cabo la práctica copie la función al directorio de realización de la práctica.

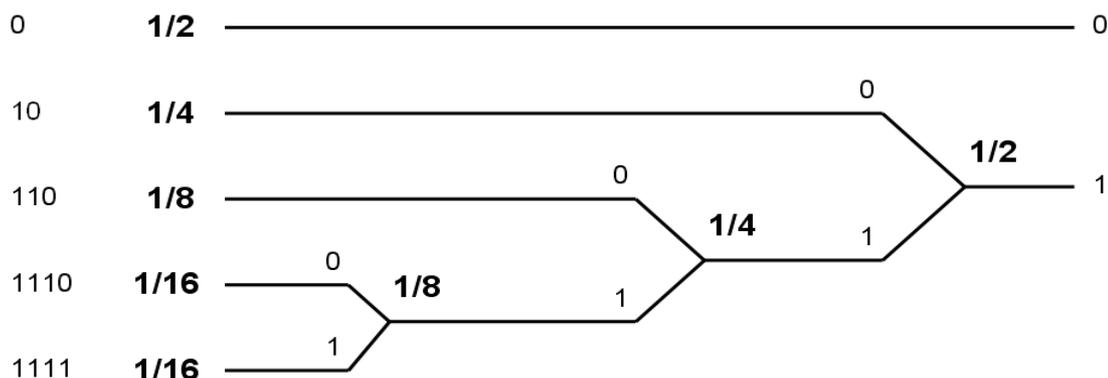
Justo antes de finalizar la práctica, comprima el proyecto en un fichero `TxDatosPIGXX.zip`, conéctese al sistema de entrega de prácticas de la Intranet y entréguelo en el grupo que corresponda (A o B). Guárdese adicionalmente una **copia personal**, para posibles futuras reutilización del código en prácticas posteriores.

1 Introducción

La codificación Huffman es uno de los métodos clásicos de codificación de fuente de la familia de los métodos estadísticos, esto es, que necesitan conocer la distribución probabilística de la fuente.

El algoritmo básico consiste en, tras ordenar los símbolos de mayor a menor en probabilidad, ir juntando parejas de menor probabilidad formando un árbol. Cuando solamente hay dos raíces, se asigna los símbolos 0 y 1 (o 1 y 0) a cada raíz, y se itera hacia atrás.

Dada, por ejemplo, una fuente F con 5 símbolos de probabilidades $\{1/2, 1/4, 1/8, 1/16, 1/16\}$ se podría construir el siguiente diagrama de árbol. En cada etapa los dos elementos con menor probabilidad se unen y se obtiene un elemento resultante cuya probabilidad es la suma de las dos anteriores. Cada vez que se realiza una unión de dos elementos se asigna a cada uno de ellos un '1' o un '0' (el resultado dependerá de que criterio de asignación se aplique). El proceso termina cuando únicamente quedan dos elementos y también a cada uno de ellos se le asigna un '1' o un '0'. Finalmente para conocer el código asociado a cada probabilidad se recorre el árbol en sentido inverso y se concatenan los unos o ceros por los que se va pasando hasta llegar al principio de cada ramificación.



2 Codificación Huffman

Haciendo uso de la función de matlab `huffmanRecursivo` proporcionada (fichero `huffmanRecursivo.m`) escriba un programa matlab `huffman` que:

- permita introducir una lista de probabilidades
- compruebe que su suma es correcta
- calcule la entropía de esa fuente
- calcule el código Huffman correspondiente (mediante la función `huffmanRecursivo`)
- calcule la longitud media del código generado

Nota: Para acceder correctamente a las celdas tipo string de la variable X de tipo cell que devuelve la función `huffmanRecursivo` utilice `X{i,2}`

2.1.1 Ejercicio 1: Fuente A

Utilizando el programa `huffman` calcular el código Huffman para la fuente A con 4 símbolos con probabilidades $\{0.6, 0.2, 0.1, 0.1\}$.

Calcule analíticamente la entropía, el código Huffman y la longitud media del código para la fuente A.

El código $\{0, 10, 110, 111\}$ también es un código Huffman válido para la fuente A. ¿En que se diferencia la generación del código para obtener este último. Modificar la función `huffmanRecursivo` para que genere este último código a partir de las probabilidades de la fuente A. Utilice a partir de ahora esta función en el programa `huffman`.

2.1.2 Ejercicio 2: Fuente B

Dada la fuente B con 4 símbolos con probabilidades $\{0.5, 0.3, 0.15, 0.05\}$.

Calcule analíticamente la entropía, el código Huffman y la longitud media del código para la fuente B.

2.1.3 Ejercicio 3: Fuente C

Dada la fuente C con 4 símbolos con probabilidades $\{0.5, 0.25, 0.15, 0.1\}$.

Calcule analíticamente la entropía, el código Huffman y la longitud media del código para la fuente C.

A la vista de los resultados para las fuentes A, B y C (analíticos y del programa) comente el algoritmo implementado por la función `matlab` que crea el código Huffman.

2.1.4 Ejercicio 4: Fuente D

Dada la fuente D con 8 símbolos con probabilidades $\{0.3, 0.15, 0.25, 0.2, 0.05, 0.025, 0.015, 0.01\}$.

Calcule analíticamente la entropía, el código Huffman y la longitud media del código para la fuente D.

Comente los resultados obtenidos analíticamente y mediante el programa.

3 Codificación Huffman con extensión de fuente

Modifique el programa `matlab huffman` para que adicionalmente a los pasos que realiza:

- solicite el orden de la extensión de fuente que se quiere calcular
- calcule el código Huffman correspondiente a la extensión
- calcule la longitud media del código generado para la extensión

3.1.1 Ejercicio 5: Fuente E

Dada la fuente E con 4 símbolos con probabilidades $\{0.4, 0.3, 0.2, 0.1\}$.

Calcule analíticamente la entropía, el código Huffman y la longitud media del código para la fuente E.

3.1.2 Ejercicio 6: Fuente F (E^2)

Calcule la tabla de probabilidades para la fuente F que sea el resultado de codificar la extensión de fuente de orden 2 de la fuente E.

Calcule, utilizando el programa implementado, la entropía, el código Huffman y la longitud media del código para la fuente F.

Comente los resultados frente a los de la fuente E.

3.1.3 Ejercicio 7: Fuente G (E^3)

Calcule la tabla de probabilidades para la fuente G que sea el resultado de codificar la extensión de fuente de orden 3 de la fuente E.

Calcule, utilizando el programa implementado, la entropía, el código Huffman y la longitud media del código para la fuente G.

Comente los resultados de la entropía y la longitud media del código para la fuente G.

Comente los resultados frente a los de las fuentes E y F.