Ejercicios Tema 4

Algoritmia para problemas difíciles

Departamento de Informática e Ingeniería de Sistemas
Escuela de Ingeniería y Arquitectura – Universidad de Zaragoza

4 de diciembre de 2023

Ejercicios sobre algoritmos de compresión

Notas: Cualquiera de los ejercicios de implementación debe acompañarse de dicha implementación y una memoria explicativa de las prestaciones del algoritmo y las pruebas realizadas.
Se utiliza el término "compresor real" para referirse a implementaciones terminadas y muy utilizadas de algoritmos de compresión que a menudo forman parte de sistemas operativos o bibliotecas.

Ejercicio 1 Encuentra una cota superior y una cota inferior para el número de frases del algoritmo LZ78 que sean válidas para cualquier entrada y lo más ajustadas posibles. Busca ejemplos en los que se vea la utilidad de dichas cotas. Nota: el número de símbolos del diccionario se supone fijo y conocido, no se puede variar.

Ejercicio 2 Modificar el algoritmo de compresión de Lempel Ziv 78 (LZ78) para que el tamaño máximo del diccionario sea un parámetro M dado por el usuario (es decir, la entrada es ahora una cadena w a comprimir y un entero M) y dicho diccionario contenga las primeras M frases del LZ78 que son todas las que se pueden usar en la compresión, una vez lleno el diccionario es fijo.
Encontrar los casos mejor y mejor (en términos de ratio de compresión) para el algoritmo anterior y utilizarlos para razonar cotas superior e inferior al tamaño de la compresión alcanzada. Nota: El número de símbolos del diccionario se supone fijo y conocido, no se puede variar. Interesan los casos en que la longitud de la entrada es mucho más grande que el tamaño del diccionario.

Ejercicio 3 Compara experimentalmente el nivel de compresión de LZ78 y Huffman. Para ello puedes realizar tu propia implementación de los mismos y/o utilizar los compresores "reales" que sean lo más parecidos posible a ellos. Nota: En caso de implementarlos, el lenguaje elegido tiene que permitir el uso eficiente de bits.

Ejercicio 4 Implementa en tu lenguaje favorito de programación el algoritmo LZ78 estrictamente como se publicó inicialmente (que coincide con la explicación vista en clase). Genera strings binarios aleatorios largos y comprueba su grado de compresión con LZ78 y con con compresor "real" como `gzip` o `bzip2`. Nota: El lenguaje elegido tiene que permitir el uso eficiente de bits.

Ejercicio 5 Implementa en tu lenguaje favorito de programación el algoritmo LZ78 estrictamente como se publicó inicialmente (que coincide con la explicación vista en clase). Comprueba el nivel de compresión que alcanza sobre alguno (o varios) de los libros del proyecto Gutenberg (o algún repositorio similar). Compara con el nivel de compresión que se alcanza con compresor "real" como `gzip` o `bzip2`. Nota: El lenguaje elegido tiene que permitir el uso eficiente de bits.

Ejercicio 6 Averigua exactamente cómo está implementado el compresor `zpaq`. ¿Qué algoritmo/s utiliza y qué implementación exacta? ¿Qué formato devuelve? ¿Usa algún diccionario o ventana y de qué tamaño? ¿Se trata de una única implementación o bien hay variantes? ¿Forma parte de algún sistema operativo o librería y en ese caso depende de la versión del mismo/a?

Ejercicio 7 Averigua exactamente cómo está implementado el compresor `compress` (forma parte del

sistema `Unix` y está formado por dos programas, `compress` y `uncompress`). ¿Qué algoritmo/s utiliza y qué implementación exacta? ¿Qué formato devuelve? ¿Usa algún diccionario o ventana y de qué tamaño? ¿Se trata de una única implementación o bien hay variantes? ¿Forma parte de algún sistema operativo o librería y en ese caso depende de la versión del mismo/a?

Ejercicio 8 Averigua exactamente cómo está implementado el compresor `Deflate`. ¿Qué algoritmo/s utiliza y qué implementación exacta? ¿Qué formato devuelve? ¿Usa algún diccionario o ventana y de qué tamaño? ¿Se trata de una única implementación o bien hay variantes? ¿Forma parte de algún sistema operativo o librería y en ese caso depende de la versión del mismo/a?

Ejercicio 9 Averigua exactamente cómo está implementado el compresor `gzip`. ¿Qué algoritmo/s utiliza y qué implementación exacta? ¿Qué formato devuelve? ¿Usa algún diccionario o ventana y de qué tamaño? ¿Se trata de una única implementación o bien hay variantes? ¿Forma parte de algún sistema operativo o librería y en ese caso depende de la versión del mismo/a?

Ejercicio 10 Averigua exactamente cómo está implementado el compresor `bzip2`. ¿Qué algoritmo/s utiliza y qué implementación exacta? ¿Qué formato devuelve? ¿Usa algún diccionario o ventana y de qué tamaño? ¿Se trata de una única implementación o bien hay variantes? ¿Forma parte de algún sistema operativo o librería y en ese caso depende de la versión del mismo/a?

En caso de entregar alguno de estos ejercicios, la fecha límite es el lunes 11 de diciembre.

Antes de realizar cualquiera de estos ejercicios el alumno debe seleccionarlo en moodle.

Cualquier fuente utilizada en la resolución de estos ejercicios debe ser indicada claramente en la solución.