

КУРСОВИ ПРОЕКТИ ПО СТРУКТУРИ ОТ ДАННИ

Факултет по математика и информатика на СУ „Св. Климент Охридски“,
Специалности „Информатика“ и „Софтуерно инженерство“, 2 курс

Зимен семестър 2016/2017

Архивиране на файлове

В рамките на този проект трябва да се разработи приложение, което сканира дадена директория и архивира нейното съдържание. Програмата трябва да създаде един архивен файл, който пази в себе си оригиналното съдържание на директорията. Съдържанието на файла трябва да се компресира, като за кодирането трябва да се използва алгоритъмът на Хъфман, а моделирането можете да изберете сами как да реализирате.

Относно разликата между моделиране и кодиране: Нека имаме следния стринг:

"abcd abab cdcd cdab"

Най-грубо казано, моделирането се опитва да анализира входа и да определи кои са отделните единици, които след това трябва да се представят по оптимален начин (кодиране). Например една възможност е моделирането да заключи, че в текста има четири думи, които подлежат на кодиране:

"abcd ", "abab ", "cdcd ", "cdab"

Ако моделирането им съпостави следните числа:

1:"abcd ", 2:"abab ", 3:"cdcd ", 4:"cdab" Тогава

кодираният текст ще изглежда така:

1234.

Възможно е обаче моделирането да реши, че имаме следните четири срички:

"ab", "ab ", "cd", "cd "

Нека отново предположим, че кодирането ще ги номерира от едно до четири:

1:"ab", 2:"ab ", 3:"cd", 4:"cd " Тогава

кодираният текст ще изглежда така:

14123431

Когато компресирате стринга, не можете да предадете просто кодираното представяне, защото зад него може да стоят безброй неща. Трябва да предадете и използвания речник. Така за последния пример, ако искаме да предадем на някого кодирания текст, можем да предадем например:

1:"ab", 2:"ab ", 3:"cd", 4:"cd ", 14123431

Както се вижда от примерите, при една и съща схема на кодиране, може да се получат различни резултати, в зависимост от резултатите от моделирането. Както вече написахме, в проектите трябва да се използва Хъфманово кодиране, а сами можете да изберете как да моделирате входа. Най-лесно, но не особено ефективно ще бъде да четете входа байт по байт и да кодирате отделните байтове.

При оценяването на проектите ви ще се взема предвид и как точно моделирате входа си.

Като функционалност приложението ви трябва да може да архивира или разархивира съдържанието на дадена директория. При стартирането му потребителят трябва да може или да укаже директория, която да се архивира или да укаже файл, който трябва да се разархивира.

Входът на програмата трябва да се получава от командния ред. Първият параметър указва каква операция да се изпълни. Възможните операции са както следва:

№	Команда	Описание
1.	-Pack	Архивиране. В този случай след командата се подава път до директория, която да се обработи и името на архивния файл, който трябва да се създаде. Например: > Archive.exe -Pack C:\Test C:\Temp\myfile.arc
2.	-Unpack	Разархивиране. В този случай след командата се подава път до съществуващ архив и път до директория, в която той да се разархивира. Например: > Archive.exe -Unpack C:\Temp\myfile.arc C:\AnotherDirectory\
3.	-List	Показване на съдържанието на архив. След тази команда се подава път до съществуващ архив. Програмата ви трябва да отвори архива и да изведе на екрана имената на всички файлове и директории, които се съдържат в него. Например: > Archive.exe -List C:\Temp\myfile.arc

Докато работи, приложението трябва да извежда някаква информация, за да може потребителят да знае дали то работи или е „увиснало“. Например може да се извежда броят на обработените файлове, колко процента от работата е свършена и т.н. Един сравнително прост вариант и за трите команди, е да извеждате името на всеки обработен файл.

Обърнете внимание, че във входната директория може да има не само файлове, но и други директории. Тъй като се иска да архивирате цялата директория, това означава, че трябва да можете рекурсивно да влезете в тези поддиректории и да запазите и тяхното съдържание. Също така, възможно е някои от поддиректориите да са празни (да не съдържат файлове) или да имате файлове с размер нула байта (празни файлове). Приложението ви не трябва да ги пропуска и след разархивиране, те трябва да се възстановят така, както са били в оригиналната директория.