



Софийски университет „Св. Климент Охридски“  
Факултет по математика и информатика

## Теми за проекти

курсове *Структури от данни;  
Структури от данни и програмиране*

специалности *Информатика, Информационни системи и  
Компютърни науки (1-ви поток)*

*зимен семестър 2021/22 г.*

### Тема 1: LZW компресиране

---

В този проект трябва да реализирате програма, която архивира и компресира файлове.

Програмата трябва да получава параметрите си от командния ред. Тя трябва да може да архивира и компресира множество файлове и директории в един файл, използвайки LZW алгоритъма (LZ78 или негов производен). Трябва да се поддържат следните функционалности:

1. ZIP: Създаване на архив по подаден списък от файлове и/или директории. Допустимо е използването на символите \* и ? в имената за по-лесно посочване на множество файлове. При работа с директории те се архивират в дълбочина (рекурсивно).
2. UNZIP: Разархивиране на един или всички файлове от архива по посочено име и директория в която да се записва резултата. Допустимо е използване на \* и ? в имената за по-лесно посочване на множество файлове. При разархивиране трябва да се запази структурата на директории.
3. INFO: Извеждане на списък на компресираните файлове с информация за нивото на компресия.
4. REFRESH: Обновяване на вече архивиран файл с друга негова версия. Изисква се тази операция да е по-бърза от създаването на целия архив наново.
5. EC: Проверка за повреден файл в архив – проверява дали не е възникнала промяна в архивния файл, която води до това, че един или повече от

архивираните файлове не могат да се възпроизведат коректно при разархивиране. За целта включете някакъв вид шумозащитно кодиране.

Уточнения:

- Операциите UNZIP и REFRESH трябва да позволяват да се работи с отделни файлове в архива. Не се считат за коректни решения, при които за разархивиране или промяна на един отделен файл, трябва да се обработят и всички останали в архива.
- При посочване на имената на файлове символ \* замества нула или повече произволни символи, а символ ? - точно един произволен символ.
- Възможно е в дървото, което сканирате, да срещнете празна директория (такава, в която няма никакви файлове). Трябва да намерите начин да запазите информация за нея в архива и след това да я възстановите.
- Когато архивирате дадена директория, нейното съдържание трябва да се запази с относителни пътища, които след това да се възпроизведат в директорията, в която тя се разархивира. Например:

Да допуснем, че архивираме директорията C:\Users\John Smith\Documents. В нея се съдържат следните:

```
C:\Users\John Smith\Documents\Homework.docx  
C:\Users\John Smith\Documents\music\song.mp3  
C:\Users\John Smith\Documents\video\video.mp4
```

Ако разархивираме получения файл в директория D:\Extracted, трябва да получим следните:

```
D:\Extracted\Homework.docx  
D:\Extracted\music\song.mp3  
D:\Extracted\video\video.mp4
```