

### Потпуна и релативна путања

При раду са текстуалним фајловима за приступ позицији текстуалног фајла у директоријском систему можемо користити потпуну путању (C:\Temp\primer.txt) и релативну путању, у односу на директоријум из којег се апликација покреће (..\..\example.txt).

Ако се користи пуна путања, косе црте служе за одвајање фолдера.

У програмском језику C# то се може урадити на два начина: са дуплим бекслешом или са @ пре стринг литерала.

У примеру, путања до фајла "C:\Temp\radovi\test.txt" се може користити као стринг:

```
string fajl = "C:\Temp\radovi\test.txt";
string fajl = @"C:\Temp\radovi\test.txt";
```

Иако је употреба релативне путање компликованија пошто се током рада на пројекту може променити структура директоријума, и даље се препоручује избегавање коришћења потпуне путање.

Потпуна путања није добра због тога што је већа зависност апликације од окружења и сада је апликације нетрансферабилна.

### Кодовање фајлова

У меморији се све налази у форми бинарних података.

То значи да се текстуални фајлови треба да представе дигитално, да би се сместили у меморију. Овај процес се назива кодовање (encoding) или прецизније кодовање карактера из текстуалног фајла.

Процес кодовања се састоји од замене текст знакова са одређеним секвенцом бинарних вредности.

Кодовање карактера има правила за конверзију из текста у секвенцу бајтова и обрнуто.

Шема по којој се то ради је табела карактера са њиховим вредностима.

Два метода кодовања која се најчешће користе са ћириличним писмом у рачунарству су UTF-8 Windows-1251.

UTF-8 је универзална шема кодовања која подржава све језике.

Најчешће коришћени знаци (латинична слова, цифре) се кодују са 1B а ређе коришћени Уникод знаци (и ћирилични) се кодују са 2B, док сви остали се кодују са 3B или 4B.

Ова шема може да конвертује било који Уникод текст у бинаран облик и уназад.

Windows-1251 се обично користи за кодовање ћириличних текстова (поруке у мејловима).

Садржи 256 карактера, укључујући латинична слова, ћирилична и неке чешће коришћене симболе.

За рад са кодовањем унутар .NET Framework користи се класа System.Text.Encoding:

```
Encoding win1251 = Encoding.GetEncoding("Windows-1251");
```

### Читање ћириличног садржаја

Да би се могао читати ћирилични садржај, мора се користити одговарајуће кодовање које тачно тумачи ћириличне карактере.

За то се највише користи Windows-1251 кодовање и да би се могло користити треба га поставити као кодовање стрима:

```
Encoding win1251 = Encoding.GetEncoding("Windows-1251");
```

```
StreamReader citac = new StreamReader("test.txt", win1251);
```

Ако се експлицитно не постави тип кодовања за одређени фајл, .NET ће користити UTF-8 као дифолтни.

Уникод је индустријски сандард који омогућава рачунарима да приступе садржају и манипулишу текстом.

Препознаје преко 100000 карактера као и више шема кодовања.

Карактери у C# (типови char и string) су такође представљени у Уникоду.

Да би се могли читати карактери смештени у Уникоду, морају се користити подржане шеме кодовања за овај стандард.

Најпопуларнији стандард је UTF-8:

```
StreamReader citac = new StreamReader("test.txt",  
                                     Encoding.GetEncoding("UTF-8"));
```