Úkol 2 – Šifrování a kryptoanalýza

Cílem je analyzovat a rozšifrovat předložené šifrové texty.

Soubor `ciphers.txt` obsahuje 3 zašifrované texty. Texty jsou v anglickém jazyce, malými písmeny a bez diakritiky. Vytvořte program, který bude provádět kryptoanalýzu a pokusí se zašifrované texty prolomit. Program bude očekávat jako jediný povinný parametr cestu k souboru `ciphers.txt`. Další parametry si můžete zvolit dle vašeho uvážení. Soubor obsahuje tři zašifrované texty, které jsou odděleny znakem konce řádky a znakem „#“.

Využijte frekvenční analýzu, případně metodu hrubou silou. Bude požadováno, aby výstupem vašeho programu byl rozšifrovaný text - tzn. programově rozhodnout, že se našel smysluplný text (např. porovnání s definovaným slovníkem). V případě, že váš program nedokáže šifru prolomit nebo by mu to trvalo neúměrně dlouhý čas, tak do výstupního souboru vypište řetězec: `“Cryptanalysis failed!”`

K úspěšnému odevzdání je potřeba, aby program vytvořil výstupní soubor `result.txt`, dle následujícího formátu:

```
text prvni sifry\n
#
text druhe sifry\n
#
text treti sifry\n
```

Korektní soubor `result.txt` může tedy vypadat např. takto:

```
meet me in the church\n
#
meet me in the cathedral\n
#
okay see you there\n
```

nebo takto:

```
meet me in the church\n
#
Cryptanalysis failed!\n
#
Cryptanalysis failed!\n
```

Prosím dodržte tento formát k usnadnění kontroly. Testovat se může i na jiné sadě zašifrovaných textů (stejné šifry), je nutné tedy program napsat dostatečně obecně. Vytvořte soubor `README.txt`, kde stručně popíšete a zhodnotíte vaše snažení a způsob řešení.

Dobré rady:

Je třeba mít na paměti, že u lámání složitějších šifer (např. metoda hrubé síly) je nemyslitelné ručně procházet a ověřovat velké množství možných výstupů. Proto je vhodné umět automaticky rozhodnout, že byla šifra prolomena (např. hledejte smysluplná slova či ngramy). Můžete si vytvořit vlastní slovník nejčastějších anglických slov či bigramů, trigramů.

Např. zde https://www3.nd.edu/~busiforc/handouts/cryptography/Letter%20Frequencies.html

Pokud program vykonává nějakou náročnou činnost, je vhodné zobrazovat informaci o progresu (progress bar nebo vypisovat counter, každých 100 nebo 1000 iterací cyklu).

Uvědomte si, že řada šifer má omezený klíč (např. u Caesarovy šifry číslo udávající posun abecedy apod.).

Hodnocení:

Max. počet bodů: **7** (1. a 2. šifra za 2 body, 3. šifra za 3 body)

Odevzdání:

Vytvořte a odevzdejte archiv `BIT_ukol2_<jméno_prijmeni>.zip` Součástí archivu bude soubor `README.txt`, a všechny vaše (i pomocné) skripty a programy.

Řádný termín odevzdání: 25.3.2021, 23:59:59

Mezní termín odevzdání: 1.4.2021, 23:59:59