

Programovanie (1) v C/C++ 2025/26

Cvičenia 9, miniprojekt

Šifra

Tento **miniprojekt** je väčšia úloha, na ktorú máte viac času ako na bežné príklady z cvičení, ale máte za ňu aj viac bodov. Začnite na nej pracovať v predstihu, nenechávajte si ju na poslednú chvíľu.

Na miniprojekte pracujte **samostatne**, neukazujte svoj program spolužiakom a nepoužívajte nástroje umelej inteligencie. V prípade otázok kontaktujte vyučujúcich predmetu. Študenti, ktorí majú z testu pre pokročilých uznané cvičenia 9, budú mať uznaný aj tento miniprojekt a nemusia ho odovzdávať.

Body môžu získať iba programy, ktoré majú na testovači celkový status OK. Pri bodovaní ale budeme brať do úvahy aj to, či vaše riešenie spĺňa pokyny v zadaní. Dbajte aj na pekný programátorský štýl (zrozumiteľné mená premenných a funkcií v angličtine alebo slovenčine, odsadzovanie, komentáre). Namiesto opakovania podobného kódu vytvorte vhodné funkcie, napríklad na spočítanie počtu písmen v súbore.

Na prednáške 16 sme videli ako súbor zašifrovať jednoduchou Caesarovou šifrou. V tejto úlohe napíšete program, ktorý takto zašifrovaný súbor automaticky odšifruje, pričom uhádne správny posun K , o ktorý má posunúť všetky písmená v abecede. Základnou myšlienkou je nájsť taký posun, v ktorom sa jednotlivé písmená textu budú vyskytovať podobne často, ako sa vyskytujú v poskytnutej ukážke textu napísanej v tom istom jazyku ako zašifrovaný text.

Váš program bude načítavať dva súbory: `sifra.txt` a `korpus.txt`. V súbore `sifra.txt` je zašifrovaný text, pričom Cézarovou šifrou boli spracované iba malé písmená anglickej abecedy. Veľké písmená sa v texte nenachádzajú a ostatné znaky (medzery, čísla, interpunkcia a podobne) sú ponechané v pôvodnej forme. V súbore `korpus.txt` je ukážka textu v tom istom jazyku ako neznámy zašifrovaný text. V tomto texte sa tiež nenachádzajú veľké písmená.

Obidva vstupné súbory prečítajte a uložte si do tabuľky počty výskytov jednotlivých písmen **a až z**. Napríklad v prvom ukážkovom vstupe `korpus.txt` je $21\,337$ písmen, z toho $2\,618 \times a$, $475 \times b$, $721 \times c$, ..., $728 \times z$. Tieto počty výskytov prepočítajte na percentá tak, že každý počet vynásobíte číslom 100 a celočíselne vydelite celkovým počtom písmen **a až z** (dostaneme tak percentá zaokrúhlené nadol na celé číslo). Napríklad písmeno **a** tvorí z korpusu približne $2618 \cdot 100 / 21337 = 12\%$. Počty jednotlivých písmen aj tabuľky percentuálneho zastúpenia pre `korpus.txt` a `sifra.txt` vypíšte na konzolu vo formáte ako v príklade nižšie.

Potom skúšajte všetky posuny od 0 po 25 a pre každý posun spočítajte, ako by vyzerala tabuľka pre nezašifrovanú verziu súboru `sifra.txt`, pokiaľ sa pri šifrovaní použil daný posun. Na výstup vypíšte riadok, v ktorom je najskôr hodnota posunu, potom táto tabuľka a nakoniec číslo vyjadrujúce, ako veľmi sa tabuľka percent odlišuje od tabuľky pre súbor `korpus.txt`. Odlišnosť dvoch tabuliek **a** a **b** spočítajte ako súčet absolútnych hodnôt rozdielov ich zodpovedajúcich prvkov $|a[i] - b[i]|$. Absolútnu hodnotu počíta funkcia `abs` z knižnice `cstdlib`. Pri výpise použite formát ako v príklade.

Medzi vypísanými posunmi nájdite ten, pri ktorom je spočítaná odlišnosť najmenšia. Tento posun a jeho odlišnosť vypíšte na konzolu a použite ho na odšifrovanie súboru `sifra.txt`. Odšifrovaný text vypíšte do súboru `vystup.txt`. Pri odšifrovaní meňte len malé písmená, ostatné znaky vypíšte bez zmeny.

Súbory načítavajte znak po znaku, napríklad príkazom `getc`, a neukladajte celé súbory alebo ich riadky do pamäte. Súbor `sifra.txt` odporúčame čítať dvakrát: raz na výpočet tabuľky a raz na odšifrovanie.

Na stránke nájdete štyri sady vstupných súborov. Uvádzame, ako má vyzeráť výstup na konzolu pre prvý z nich. Nezabudnite, že ešte treba vypísať aj odšifrovaný text do súboru `vystup.txt`. Ručne si môžete skontrolovať, či je odšifrovaný text skutočne rozpoznateľným textom v rovnakom jazyku ako vstupný korpus.

Príklad výstupu na konzolu:

```
Pocty korpus:
a:2618 b:475 c:721 d:952 e:1953 f:24 g:32 h:566 i:1419 j:407 k:770 l:877 m:686 n:1234 o:2015 p:559 q:0 r:939 s:1152 t:1172 u:770 v:770 w:1 x:0 y:497 z:728
Percenta korpus:
a:12 b:2 c:3 d:4 e:9 f:0 g:0 h:2 i:6 j:1 k:3 l:4 m:3 n:5 o:9 p:2 q:0 r:4 s:5 t:5 u:3 v:3 w:0 x:0 y:2 z:3
```

Pocty sifra:
a:376 b:332 c:247 d:0 e:9 f:174 g:117 h:671 i:65 j:173 k:305 l:466 m:45 n:52 o:126 p:411 q:156 r:237 s:166 t:228 u:433 v:572 w:239 x:0 y:332 z:372
Percenta sifra:
a:5 b:5 c:3 d:0 e:0 f:2 g:1 h:10 i:1 j:2 k:4 l:7 m:0 n:0 o:1 p:6 q:2 r:3 s:2 t:3 u:6 v:9 w:3 x:0 y:5 z:5

Posun 0: a:5 b:5 c:3 d:0 e:0 f:2 g:1 h:10 i:1 j:2 k:4 l:7 m:0 n:0 o:1 p:6 q:2 r:3 s:2 t:3 u:6 v:9 w:3 x:0 y:5 z:5 odlisnost 89
Posun 1: a:5 b:3 c:0 d:0 e:2 f:1 g:10 h:1 i:2 j:4 k:7 l:0 m:0 n:1 o:6 p:2 q:3 r:2 s:3 t:6 u:9 v:3 w:0 x:5 y:5 z:5 odlisnost 83
Posun 2: a:3 b:0 c:0 d:2 e:1 f:10 g:1 h:2 i:4 j:7 k:0 l:0 m:1 n:6 o:2 p:3 q:2 r:3 s:6 t:9 u:3 v:0 w:5 x:5 y:5 z:5 odlisnost 87
Posun 3: a:0 b:0 c:2 d:1 e:10 f:1 g:2 h:4 i:7 j:0 k:0 l:1 m:6 n:2 o:3 p:2 q:3 r:6 s:9 t:3 u:0 v:5 w:5 x:5 y:5 z:3 odlisnost 73
Posun 4: a:0 b:2 c:1 d:10 e:1 f:2 g:4 h:7 i:0 j:0 k:1 l:6 m:2 n:3 o:2 p:3 q:6 r:9 s:3 t:0 u:5 v:5 w:5 x:5 y:3 z:0 odlisnost 97
Posun 5: a:2 b:1 c:10 d:1 e:2 f:4 g:7 h:0 i:0 j:1 k:6 l:2 m:3 n:2 o:3 p:6 q:9 r:3 s:0 t:5 u:5 v:5 w:5 x:3 y:0 z:0 odlisnost 97
Posun 6: a:1 b:10 c:1 d:2 e:4 f:7 g:0 h:0 i:1 j:6 k:2 l:3 m:2 n:3 o:6 p:9 q:3 r:0 s:5 t:5 u:5 v:5 w:3 x:0 y:0 z:2 odlisnost 79
Posun 7: a:10 b:1 c:2 d:4 e:7 f:0 g:0 h:1 i:6 j:2 k:3 l:2 m:3 n:6 o:9 p:3 q:0 r:5 s:5 t:5 u:5 v:3 w:0 x:0 y:2 z:1 odlisnost 17
Posun 8: a:1 b:2 c:4 d:7 e:0 f:0 g:1 h:6 i:2 j:3 k:2 l:3 m:6 n:9 o:3 p:0 q:5 r:5 s:5 t:5 u:3 v:0 w:0 x:2 y:1 z:10 odlisnost 71
Posun 9: a:2 b:4 c:7 d:0 e:0 f:1 g:6 h:2 i:3 j:2 k:3 l:6 m:9 n:3 o:0 p:5 q:5 r:5 s:5 t:3 u:0 v:0 w:2 x:1 y:10 z:1 odlisnost 89
Posun 10: a:4 b:7 c:0 d:0 e:1 f:6 g:2 h:3 i:2 j:3 k:6 l:9 m:3 n:0 o:5 p:5 q:5 r:5 s:3 t:0 u:0 v:2 w:1 x:10 y:1 z:2 odlisnost 93
Posun 11: a:7 b:0 c:0 d:1 e:6 f:2 g:3 h:2 i:3 j:6 k:9 l:3 m:0 n:5 o:5 p:5 q:5 r:3 s:0 t:0 u:2 v:1 w:10 x:1 y:2 z:4 odlisnost 77
Posun 12: a:0 b:0 c:1 d:6 e:2 f:3 g:2 h:3 i:6 j:9 k:3 l:0 m:5 n:5 o:5 p:5 q:3 r:0 s:0 t:2 u:1 v:10 w:1 x:2 y:4 z:7 odlisnost 85
Posun 13: a:0 b:1 c:6 d:2 e:3 f:2 g:3 h:6 i:9 j:3 k:0 l:5 m:5 n:5 o:5 p:3 q:0 r:0 s:2 t:1 u:10 v:1 w:2 x:4 y:7 z:0 odlisnost 83
Posun 14: a:1 b:6 c:2 d:3 e:2 f:3 g:6 h:9 i:3 j:0 k:5 l:5 m:5 n:5 o:3 p:0 q:0 r:2 s:1 t:10 u:1 v:2 w:4 x:7 y:0 z:0 odlisnost 87
Posun 15: a:6 b:2 c:3 d:2 e:3 f:6 g:9 h:3 i:0 j:5 k:5 l:5 m:5 n:3 o:0 p:0 q:2 r:1 s:10 t:1 u:2 v:4 w:7 x:0 y:0 z:1 odlisnost 85
Posun 16: a:2 b:3 c:2 d:3 e:6 f:9 g:3 h:0 i:5 j:5 k:5 l:5 m:3 n:0 o:0 p:2 q:1 r:10 s:1 t:2 u:4 v:7 w:0 x:0 y:1 z:6 odlisnost 75
Posun 17: a:3 b:2 c:3 d:6 e:9 f:3 g:0 h:5 i:5 j:5 k:5 l:3 m:0 n:0 o:2 p:1 q:10 r:1 s:2 t:4 u:7 v:0 w:0 x:1 y:6 z:2 odlisnost 71
Posun 18: a:2 b:3 c:6 d:9 e:3 f:0 g:5 h:5 i:5 j:5 k:3 l:0 m:0 n:2 o:1 p:10 q:1 r:2 s:4 t:7 u:0 v:0 w:1 x:6 y:2 z:3 odlisnost 83
Posun 19: a:3 b:6 c:9 d:3 e:0 f:5 g:5 h:5 i:5 j:3 k:0 l:0 m:2 n:1 o:10 p:1 q:2 r:4 s:7 t:0 u:0 v:1 w:6 x:2 y:3 z:2 odlisnost 83
Posun 20: a:6 b:9 c:3 d:0 e:5 f:5 g:5 h:5 i:3 j:0 k:0 l:2 m:1 n:10 o:1 p:2 q:4 r:7 s:0 t:0 u:1 v:6 w:2 x:3 y:2 z:3 odlisnost 85
Posun 21: a:9 b:3 c:0 d:5 e:5 f:5 g:5 h:3 i:0 j:0 k:2 l:1 m:10 n:1 o:2 p:4 q:7 r:0 s:0 t:1 u:6 v:2 w:3 x:2 y:3 z:6 odlisnost 87
Posun 22: a:3 b:0 c:5 d:5 e:5 f:5 g:3 h:0 i:0 j:2 k:1 l:10 m:1 n:2 o:4 p:7 q:0 r:0 s:1 t:6 u:2 v:3 w:2 x:3 y:6 z:9 odlisnost 83
Posun 23: a:0 b:5 c:5 d:5 e:5 f:3 g:0 h:0 i:2 j:1 k:10 l:1 m:2 n:4 o:7 p:0 q:0 r:1 s:6 t:2 u:3 v:2 w:3 x:6 y:9 z:3 odlisnost 71
Posun 24: a:5 b:5 c:5 d:5 e:3 f:0 g:0 h:2 i:1 j:10 k:1 l:2 m:4 n:7 o:0 p:0 q:1 r:6 s:2 t:3 u:2 v:3 w:6 x:9 y:3 z:0 odlisnost 79
Posun 25: a:5 b:5 c:5 d:3 e:0 f:0 g:2 h:1 i:10 j:1 k:2 l:4 m:7 n:0 o:0 p:1 q:6 r:2 s:3 t:2 u:3 v:6 w:9 x:3 y:0 z:5 odlisnost 81
Najlepsi posun 7 odlisnost 17