

Lehký příklad: Vyřazování písmen (10 bodů)

Napište program `remove.py`, který z vět vyřadí zadanou množinu písmen. Program na výstupu vytiskne 1) věty, které obsahovaly nejvíce vyřazených písmen, 2) věty, kterým po vyřazení písmen zbylo nejvíce písmen a 3) věty s nejvyšším poměrem zbylých písmen oproti vyřazeným písmenům.

Vstup Dva argumenty z příkazové řádky:

- První argument je jméno **souboru**. Soubor obsahuje na každé řádce jednu větu. Věta obsahuje malá a velká písmena anglické abecedy, mezery a interpunkční znaménka.
- Druhý argument je **řetězec** malých písmen anglické abecedy.

Výstup Výstup obsahuje:

- Větu s vyřazenými zadanými písmeny, ze které bylo odstraněno nejvíce písmen. Pokud by bylo těchto vět se stejným počtem vyřazených písmen více, pak je vytiskněte postupně — na každou řádku jednu větu v pořadí, v jakém byly uvedeny ve vstupním souboru.
- Větu s vyřazenými zadanými písmeny, která po vyřazení písmen obsahuje nejvíce zbylých písmen. Pokud by bylo těchto vět s vyřazenými písmeny více, pak je vytiskněte postupně — na každou řádku jednu větu v pořadí, v jakém byly uvedeny ve vstupním souboru.
- Větu s vyřazenými zadanými písmeny, která má nejvyšší poměr zbylých písmen ku vyřazeným písmenům. Pokud nebylo vyřazeno žádné písmeno je poměr $+\infty$ (tudíž největší možný poměr). Pokud by bylo těchto vět více, pak je vytiskněte postupně — na každou řádku jednu větu v pořadí, v jakém byly uvedeny ve vstupním souboru.

Odevzdání Program nahrajte do Brute jako soubor `remove.py` – úloha **QT5**.

Poznámky:

- Předpokládejte, že vstup je zadán korektně, tzn. vstupní soubor existuje a obsahuje alespoň jednu větu a zadaná množina písmen k vyřazení obsahuje alespoň jedno písmeno.
- Zadaná písmena k vyřazení jsou vždy malá písmena, ale z věty je nutné vyřadit i jejich velké varianty. Například pokud se má vyřadit písmeno `t` z věty `Tata jel do Tabora.`, bude výsledek `aa jel do abora.`
- Poměr zbylých písmen ku vyřazeným lze počítat jako podíl dvou čísel pouze pokud je počet vyřazených písmen větší než 0. Pokud je počet vyřazených písmen 0, pak je tento poměr větší než všechny ostatní poměry.
- Speciální znaky jako mezera a znaky `.`, `,`, `?`, `!` nejsou písmena.

Bodování

Při řešení můžete použít libovolné funkce jazyka Python, včetně standardních knihoven dostupných v systému Brute.

Testovací data

Obsah souborů test1.txt, test2.txt, test3.txt můžete stáhnout z adresy <https://cw.fel.cvut.cz/wiki/courses/b3b33alp/cviceni>.

Příklad I Volání `python3 remove.py test1.txt aeiouy`

Obsah souboru test1.txt:

```
Bylo, nebylo.  
Proc ja sem lez.  
Vzdyt je to lehke.  
Kazdy program ma chybu.
```

Výstup

```
Kzd prgrm m chb.  
Kzd prgrm m chb.  
Prc j sm lz.
```

Komentář Z první věty zbyde B1, nbl. a bylo z ní odstraněno 5 samohlásek a ve větě zbylo 5 písmen.

Z druhé věty zbyde Prc j sm lz. a byly z ní odstraněny 4 samohlásky a ve větě zbylo 8 písmen.

Z třetí věty zbyde Vzdt j t lhk. a bylo z ní odstraněno 5 samohlásek a ve větě zbylo 9 písmen.

Ze čtvrté věty zbyde Kzd prgrm m chb. a bylo z ní odstraněno 7 samohlásek a ve větě zbylo 12 písmen.

Tedy nejvíce písmen bylo odstraněno z poslední věty, v této větě zbylo také nejvíce písmen (první a druhý řádek výstupu) a největší poměr zbylých ku vyřazeným je pro druhou větu Prc j sm lz. a to $\frac{8}{4} = 2$ (ostatní věty mají $\frac{5}{5} = 1$, $\frac{9}{5} = 1.8$, $\frac{12}{7} \approx 1.7$).

Příklad II Volání `python3 remove.py test2.txt abcde`

Obsah souboru test2.txt:

```
Lorem ipsum dolor sit amet.  
Sed elit dui, pellentesque.  
Nullam rhoncus aliquam metus.  
Fusce tellus odio, dapibus.  
Integer pellentesque quam.
```

Výstup

```
S lit ui, pllntsqu.  
Nullm rhonus liqum mtus.  
Lorm ipsum olor sit mt.
```

Komentář Z vět bylo postupně odstraněno 4, 8, 5, 7 a 7 písmen a ve větách zbylo 18, 14, 20, 15 a 16 písmen. Tedy nejvíce písmen bylo odstraněno z druhé věty, nejvíce písmen zbylo v třetí větě a největší poměr zbylých ku vyřazeným je v první větě a to $\frac{18}{4} = 4.5$

Příklad III Volání `python3 remove.py test3.txt aiu`

Popis části	Počet testů	Timeout	Max. bodů	Bodování
Náhodná data	30	1 s/test	6	0.2b za správný výsledek
Existují věty neobsahující vybraná písmena	20	1 s/test	4	0.2b za správný výsledek

Obsah souboru test3.txt:

```
Ano.  
Ne.  
Mozna  
Nikdy  
Uvidime  
Az naprsi?  
Nasnezi!
```

Výstup

```
vdme  
z nprs?  
z nprs?  
Nsnez!  
Ne.
```

Komentář Z vět bylo postupně odstraněno 1, 0, 1, 1, 3, 3, 2 písmen a ve větách zbylo 2, 2, 4, 4, 4, 5, 5 písmen. Tedy nejvíce písmen bylo odstraněno z páté a šesté věty, nejvíce písmen zbylo v šesté a sedmé větě a největší poměr zbylých ku vyřazeným je v druhé větě a to $\frac{2}{0} = \infty$