

Bonusová úloha č. 4

Naimplementujte algoritmus, který počítá Levenshteinovu vzdálenost pro slova v češtině podle následujících pokynů:

- nerozlišujte velká a malá písmena, například tedy 'n' a 'N' považujte za shodné znaky
- znaky lišící se diakritikou považujte za různé
- 'ch' berte jako dva znaky ('c' a 'h')

Pracujte se souborem `jmena.txt`, který obsahuje všechna kalendářní jména, a využijte algoritmus k vyřešení následujících úkolů.

1. Nalezněte taková dvě jména, jejichž Levenshteinova vzdálenost je maximální (v případě existence více takovýchto dvojic stačí uvést jednu z nich).

2. Uživatel vyplňuje do textového pole v elektronickém formuláři křestní jméno. Nežadá ho ale správně, vloží 5-ti znakový řetězec `Vada1`. Simulujte chování spellcheckeru a navrhněte pro uživatele smysluplné opravy, tj. korektní kalendářní jména, vybraná podle vámi navrženého kritéria. Seznam navržených oprav uveďte do odpovědi.

3. Uvažujme následující logickou hádanku. Máme danu množinu slov \mathcal{S} a v ní vybraná dvě různá slova u a z . Úkolem je přepsat počáteční slovo u na koncové z . Přepisy provádíme v krocích. V rámci jednoho kroku je povoleno přepsat aktuální slovo v jiným slovem $v' \in \mathcal{S}$, musí být ale splněna podmínka, že Levenshteinova vzdálenost mezi v a v' je maximálně 2.

Pokud například \mathcal{S} odpovídá našim kalendářním jménům, $u = \text{Diana}$ a $z = \text{Iveta}$, pak řešením je přepsat `Diana` na `Ivana` (vzdálenost je 2) a poté `Ivana` na `Iveta` (vzdálenost je opět 2). Potřebovali jsme tedy dva přepisy. Vzhledem k tomu, že vzdálenost mezi `Diana` a `Iveta` je 4, řešení sestávající ze dvou přepisů je nejkratší možné.

Nalezněte takové zadání pro množinu kalendářních jmen, které má řešení a které zároveň mezi všemi řešitelnými zadáními maximalizuje délku nejkratšího řešení. Jinými slovy, sestavte zadání, které řešitele donutí použít co největší množství přepisů, i když bude postupovat optimálně. Do odpovědi uveďte kromě počátečního a koncového slova také příslušnou nejkratší posloupnost přepisů.

Poznámka: Do odevzdávacího systému uploadněte `.zip` soubor, který bude obsahovat pdf zprávu s odpověďmi a také všechny zdrojové kódy.