|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  |  |  | | --- | --- | --- | | Meno: | Emma Macháčová | Hodnotenie projektu:  (max 10 (TS) bodov) | | Cvičenie: | pondelok 11:00 |  | | Dátum: | 22.04.2021 | |

**Projekt TZIV LS 2020/21 – TS**

|  |  |
| --- | --- |
| Zadanie: | **Počty rôznych slov**  Na vstupe je niekoľko slov oddelených oddelovačom $, reťazce sú z množiny znakov: písmena anglickej abecedy. Navrhnite Turingov stroj - riešenie, ktoré vypíše počet rôznych slov (case sensitive) v unárnej sústave.  Zadanie riešte deterministicky a nedeštruktívne (vstup spracovávajte po jednotlivých znakoch; vstup nepremazávajte, môžete použiť označenie - napr. prepísať malé na veľké písmená; výstup umiestnite vedľa na pásku; ak je to potrebné, použite textové oddeľovače). Na páske máte vstup, napravo za oddelovač riešite výstup zadania, všetky pomocné výpočty, či konštanty si môžete uviesť vľavo od vstupu, tiež za oddelovač. Zjednodušenie: množinu vstupných znakov si môžete zvoliť, ale v minimálnej mohutnosti 7 znakov. |
| Vstup: | **Akceptované vstupy** (písmená anglickej abecedy a-m; A-M)**:**  jama  ma$ja$e$ja$ja  ja$a$jaaa$jama$ja$j$a  abeceda$Abeceda$ABEc$EDA$jedla$DeDa  **Neakceptované vstupy** (čísla, diakritika, špeciálne symboly, n-z; N-Z)**:**  jana1$Dan2iela$dan45iela$pet5er$zuz25ana$zana$zuzana$jana  $Daniela$daniela$peter$zuzana$zana$zuzana$jana$daniela$  hocijaký$vstup$ktorý$má$diakritiku |

|  |  |
| --- | --- |
| Neformálne riešenie: | **Program v princípe funguje nasledovne:**   1. prečíta sa prvé písmeno slova z pásky a označí sa 2. hlava sa presúva doprava kým nenarazí na oddelovač slov 3. prvé neoznačené písmeno za oddelovačom sa porovná s načítaným písmenom v kroku č. 1 4. ak sa zhodujú písmeno sa označí, ak sa nezhodujú zmení sa oddelovač, aby sa ďalšie písmená tohoto slova už s tým ktoré aktuálne hľadáme neporovnávali (emma != eemmma; emma != ema) 5. krok č. 2 až 4 sa opakujú až po koniec vstupu, potom sa hlava vráti na začiatok pásky 6. načíta sa ďalšie neoznačené písmeno slova (kroky 2 – 6 sa opakujú dovtedy, kým sa neprečíta oddelovač slov alebo koniec pásky) 7. program skontroluje, či slová, ktoré sú označené sú označené celé (emma != emmaa) 8. zapíše sa jednotka za oddelovač, duplikované slová sa označia ako už nájdené, tie ktoré sa ignorovali sa odznačia 9. za už porovnávané slovo sad á oddelovač, aby sme sa k nemu už nevracali 10. takto sa páska postupne z ľava do prava spracuje kým nezostanú žiadne slová na spracovanie 11. páska sa vráti do pôvodného stavu   **Okrem tohoto, program dbá na nasledovné:**   * kontrola veľkých / malých písmen (ahoj != Ahoj) * nestačí, že slovo obsahuje rovnaké písmená, musia ísť v rovnakom poradí a slovo musí mať rovnakú dĺžku (aahhoojj != ahoj, ahoj != ah) * ak je vstup len jedno slovo započíta sa korektne |
| Zložitosť riešenia: | Program má jeden hlavný cyklus, v rámci ktorého sú využité ďalšie tri hlavnejšie, navzájom nezávislé vnorené cykly. V najlepšom prípade keď sú slová rovnaké porovnávanie prebehne raz a zložitosť je lineárna, v najhoršom keď je každé slovo iné je kvadratická.   |  |  |  | | --- | --- | --- | | **Dĺžka vstupu:** | **Počet krokov spolu:** | **Priemer krokov na písmeno:** | | 1 slovo (dĺžka 1) | 17 | 17 | | 1 slovo (dĺžka 2) | 29 | 15 | | 1 slovo (dĺžka 3) | 48 | 16 | | 2 slová (dĺžka 1) | 36 | 18 | | 2 slová (dĺžka 2) | 59 | 15 | | 5 slov (dĺžka 3) | 331 | 22 |   = **cca 17 krokov na písmeno** |
| Simulátor: | turingmachine.io |
| **(Kód priložený ako samostatný súbor,** github gist ID: 9e08eebbf10b3ec715fe0421d5bc3004 **)** | |