



Záverečný test

Zadanie



Ústav informatiky
Prírodovedecká fakulta
UPJŠ v Košiciach

Doplňujúce zdrojové kódy sú na stránke predmetu PAZ1b. Funkčnosť každého riešenia musí byť preukázaná spustením na testovacích vstupoch - nespustiteľné riešenia neumožňujú zisk príslušných bodov.

SpacePAZ

V súvislosti s vesmírom je teraz veľmi rušno: SpaceX pripravuje misiu na Mars. Anonymous zverejnili, že NASA ohlásí existenciu mimozemského života. Musk a Hawking vyhlásili, že pre ľudstvo je nevyhnutné začať kolonizovať a obývať aj vesmírne objekty mimo planétu Zem (Mars, Mesiac?). Svet sa mení a študenti UPJŠ musia byť pripravení na nové výzvy.

Misia Mars: Carbon dioxide processing (2+2 bodov, grafové algoritmy)

Skoro 96% atmosféry Marsu tvorí oxid uhličitý. No ľudia pre svoju existenciu potrebujú kyslík. Jedným z možných riešení je získavať ho z oxidu uhličitého podobnou cestou ako na Zemi, teda pomocou zelených rastlín. Aj keď na Zemi sú najväčšími spracovateľmi stromy, prostredie na Marse nie je pre ich rast vhodné. Preto je nutné začať, s jednoduchšími a odolnejšími rastlinami - machmi a lišajníkmi. Lenže aj machy a lišajníky potrebujú na svoj rast vlhkosť, kyslú pôdu a ich rastu pomáhajú aj niektoré druhy minerálov nachádzajúcich sa v pôde. Na založenie „martanskej farmy“ musíme mať všetko potrebné pre rastliny. Žiaľ to sa nemusí nachádzať na jednom mieste. Zo satelitných snímok vieme, kde sa nachádza voda (vlhkosť), vhodná pôda a ďalšie potrebné minerály. Navyše vieme, kde sú bezpečné prechody (bez skál a priepastí) a s dobrým povrchom medzi jednotlivými miestami, kde môžeme založiť farmu.



Úloha: Pre zadaný graf a 3 vrcholy, v ktorých sa spolu nachádza všetko potrebné pre farmu, nájdite najlepšiu polohu (vrchol), kde by sa mohla farma nachádzať.

(2b) Implementujte program, ktorý načíta neorientovaný graf z textového súboru a nájde najkratšiu cestu medzi každou dvojicou vrchov.

(2b) Rozšírte program tak, aby pre zadané tri vrcholy našiel v grafe taký vrchol, ktorého súčet vzdialenosti od týchto troch vrcholov je minimálny.

Misia Mars: Vylodenie (12 bodov, backtracking)

Jeden z tímov v SpaceX rieši presun pasažierov z materskej transportnej lode na planétu Mars - vylodenie. Na tento presun vytvorili znovu použiteľné pristávacie moduly. Každý modul je k-miestny a má maximálnu celkovú transportnú hmotnosť M . Keďže proces pristávania a vzletu je veľmi energeticky (palivovo) náročný, celkový počet pristátí musí byť minimalizovaný. Ako ale čo najoptimálnejšie prepraviť pasažierov na Mars?

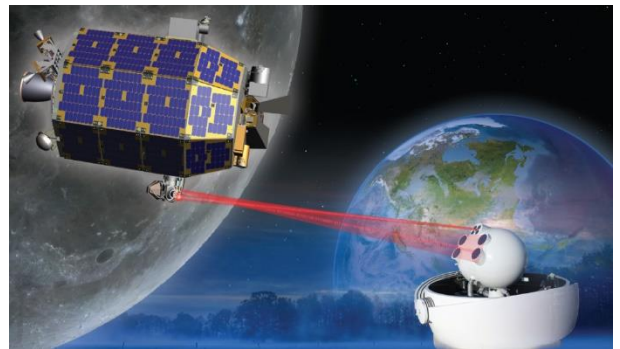


Pre zadaný zoznam hmotností pasažierov m_1, m_2, \dots, m_n nájdite prepravný plán spĺňajúce technické obmedzenia (max. k pasažierov s celkovou hmotnosťou M na jedno pristátie) s minimálnym počtom pristátí.

Hint: Je viacero možností, ako túto úlohu riešiť. Jedna z možností je preskúmať všetky spôsoby, ako by postupne zaradom mohli ľudia nastupovať do pristávacieho modulu.

Vesmírne dátové prenosy (10 bodov, ???)

Osídľovanie vesmíru znamená aj požiadavku na vybudovanie vesmírneho internetu. Problém však je, že rádiová technológia je obmedzená rýchlosťou svetla. Pri uvážení vzdialeností medzi vesmírnymi kolóniami to vedie ku komunikačným spojeniam s vysokou latenciou. Na strane druhej každá kolónia by mala mať prístup k digitalizovaným znalostiam ľudstva. SpaceX sa rozhodol, že najoptimálnejšie je mať v každej kolónii lokálne dátové centrá. Tie budú dátami naplnené tak, že dáta zo Zeme sa uložia na prenosné read-only média s extrémne vysokou dátovou hustotou a odolnosťou (vyvinuté na UPJŠ na báze nanotechnológií a LAO techniky) a vesmírnymi loďami sa prepravujú do lokálneho dátového centra v kolónii. Bude to síce chvíľu trvať, ale preniesie sa takto extrémne množstvo dát. Keď takéto exabajtové médium dorazí do dátového centra v kolónii, treba ho spracovať. Na začiatku média je súbor s popisom obsahu. Niektoré súbory môžu byť z pohľadu dátového centra nové, iné sú len potenciálne aktualizácie súborov už uložených v dátovom centre. Aby bol v súboroch poriadok, každý súbor je jedinečne identifikovaný UUID, čo je: „A **universally unique identifier (UUID)** is a 128-bit number used to identify information in computer systems.“ UUID je zapísaný vo formáte xxxxxxxx-xxxx-Nxxx-Nxxx-xxxxxxxxxxxx. Príklad UUID: 123e4567-e89b-12d3-a456-426655440000.



Tie budú dátami naplnené tak, že dáta zo Zeme sa uložia na prenosné read-only média s extrémne vysokou dátovou hustotou a odolnosťou (vyvinuté na UPJŠ na báze nanotechnológií a LAO techniky) a vesmírnymi loďami sa prepravujú do lokálneho dátového centra v kolónii. Bude to síce chvíľu trvať, ale preniesie sa takto extrémne množstvo dát. Keď takéto exabajtové médium dorazí do dátového centra v kolónii, treba ho spracovať. Na začiatku média je súbor s popisom obsahu. Niektoré súbory môžu byť z pohľadu dátového centra nové, iné sú len potenciálne aktualizácie súborov už uložených v dátovom centre. Aby bol v súboroch poriadok, každý súbor je jedinečne identifikovaný UUID, čo je: „A **universally unique identifier (UUID)** is a 128-bit number used to identify information in computer systems.“ UUID je zapísaný vo formáte xxxxxxxx-xxxx-Nxxx-Nxxx-xxxxxxxxxxxx. Príklad UUID: 123e4567-e89b-12d3-a456-426655440000.

Úloha: Máme 2 textové súbory potenciálne veľkosti niekoľkých desiatok gigabajtov, t.j. nevieme ich načítať do RAM a môžeme ich len sekvenčne čítať po riadkoch. Súbory v každom riadku obsahujú práve jedno UUID. UUID záznamy sú v súboroch (lexikograficky) usporiadané. UUID v i -tom riadku je stále (lexikograficky) menšie ako UUID v riadku $i+1$. Vytvorte program, ktorý zistí, koľko je takých UUID záznamov v prvom súbore (UUID súborov na médiu), ktoré sa nachádzajú aj v druhom súbore (UUID súborov v dátovom centre).

V jednote je sila (12 bodov, stromy)

V minulosti každá krajina a každá spoločnosť, ktorá realizovala lety do vesmíru, používala vlastnú techniku, vlastný software a vlastný personál. V ére spoločných projektov je potreba zjednotiť toho čo najviac. Ako prvý sa zjednotil vzduchový uzáver, ktorý na vstupe do ISS používajú všetky zúčastnené krajiny a firmy. Podobné zjednotenie dátových štruktúr je teraz tvoja úloha. Ved' NASA (USA), ESA (EU), FKA (Rusko), CNSA (China), ISRO (India), Space-X, Orbital ATK a mnohé ďalšie používajú vlastné systémy a aj dáta. Pred zjednotením je nutné najprv zistiť, či vôbec ide o tie isté údaje. Ved' aj jednoduchá štruktúra ako strom môže mať veľa rôznych spôsobov implementácie. Preto je nutné porovnať stromy uložené v rôznych implementáciách, či reprezentujú to isté.



Úloha: Implementujte program, ktorý načíta stromy v troch rôznych reprezentáciách a overí, či zodpovedajú identickému stromu. Navyše vieme, že každá hodnota v strome sa nachádza najviac raz. Rôzne reprezentácie:

1. Referencia na koreň stromu, ktorý je reprezentovaný triedou `Osoba` zo 4. prednášky.
2. Strom je reprezentovaný dvojicou polí, kde hodnoty na rovnakých indexoch spolu súvisia. Každý index reprezentuje jednu osobu. Prvé pole obsahuje meno osoby a druhé pole obsahuje meno rodiča danej osoby alebo `null`, ak je koreňom stromu (predpokladáme, že každá osoba má unikátne meno).
3. Pole reťazcov a matica susednosti orientovaného grafu, pričom orientovaný graf uložený v matici určuje potomkov a rodičov. Hrana ide od rodiča k potomkovi. Hodnota v poli na indexe i vyjadruje meno i -tej osoby.

Vesmírne vlaky (max. 15 bodov, dynamické programovanie)

Na diaľkové transportné lety SpaceX vytvoril vesmírne vlaky. Vesmírny vlak sa skladá z navzájom prepojených modulov - akoby vagónov. Počas letu sa niekedy stane, že vlak prechádza cez meteorický roj, ktorý nenávratne poškodí niektoré moduly. Vesmírny vlak je letu schopný, ak je v ňom poškodených nanajviš k modulov. Ak je poškodených viac ako k modulov, musia sa poškodené moduly odpojiť. Odpojiť je však možné len moduly z koncov vlaku. Takže pri tomto odpájaní modulov sa potenciálne odpoja aj funkčné moduly.



Úloha: Nech sa vesmírny vlak skladá z n modulov. Po prechode meteorickým rojom o každom module vieme, či je nenávratne poškodený (napr. v poli boolean-ov dĺžky n). Nájdite najdlhšiu súvislú podpostupnosť modulov, ktorá obsahuje nanajviš k poškodených modulov.

Hodnotenie:

- (1) 4 body za riešenie v čase $O(n^2)$ aj s výpisom riešenia
- (2) 12 bodov za riešenie v čase $O(n \cdot k)$ + 3 body za výpis riešenia

Hint: Označme si $R[i, j]$ dĺžku maximálnej súvislej podpostupnosti končiacu v i -tom module, ktorá obsahuje nanajviš j poškodených modulov.

Bipartitné grafy (12 bodov, grafy)

Galileo Galilei raz povedal: „*Velká kniha vesmíru je psaná jazykom matematiky*“. Preto možno čakať, že akákoľvek inteligentná mimozemská civilizácia bude tomuto jazyku rozumieť. Nuž a dôležitou časťou matematiky je teória grafov.

Zaujímavou skupinou grafov (aj z pohľadu praxe) sú takzvané bipartitné grafy. Aké to sú grafy? Graf G nazveme bipartitným, ak množinu jeho vrcholov $V(G)$ môžeme rozdeliť na dve disjunktné podmnožiny X a Y , t.j. $X \cap Y = \emptyset, X \cup Y = V(G)$, s takou vlastnosťou, že pre každú hranu grafu platí, že jeden jej koniec je v množine X a druhý v množine Y .

Úloha: Vytvorte program, ktorý pre zadaný neorientovaný (potenciálne aj nesúvislý) graf overí, či je tento graf bipartitný. Graf načítajte z textového súboru, formát si zvolte podľa vlastného uváženia. Očakáva sa riešenie v polynomiálnom čase od veľkosti grafu.

Rada: Nech G je súvislý bipartitný graf. Ak si zoberieme ľubovoľný vrchol $s \in X$, tak jeho susedia (ak nejakých má) sú z množiny Y . Toto pozorovanie ide rozšíriť. Ak si zoberieme ľubovoľný vrchol $s \in X$, potom vrcholy, ktorých vzdialenosť od s je párna, sú v množine X a vrcholy, ktorých vzdialenosť od s je nepárna, sú v množine Y . Čo ale s hocijakým súvislým grafom? Nuž ak by takýto graf mal byť bipartitný, potom po rozdelení vrcholov do množín X a Y podľa vzdialenosti od zvoleného vrcholu by malo platiť, že v grafe niet hrany, ktorá by mala oba konce buď v X alebo Y .