

Semestrálna práca 1 – analýza výkonu údajových štruktúr

Úlohou prvej semestrálnej práce je experimentálne porovnať výkon rozdielnych implementácií vybraných ADT. Výstupom semestrálnej práce bude:

1. Knižnica Vami testovaných údajových štruktúr (podľa zvolenej úrovne semestrálnej práce).
2. Aplikácia schopná otestovať štruktúry a vyprodukovať výsledné CSV súbory.
3. CSV súbory s výstupmi z testov.
4. Dokumentácia, ktorá bude obsahovať
 - o popis zvolených údajových štruktúr vrátane UML diagramu,
 - o návrh testov a aplikácie,
 - o popis štruktúry CSV súborov,
 - o metodiku spracovania výsledkov testov (CSV súborov), a
 - o prezentáciu výsledkov a záverov vyplývajúcich z testovania.

Údajové štruktúry musia byť Vami naprogramované, správne objektovo navrhnuté, univerzálne, a **efektívne** implementované z pohľadu výpočtovej a pamäťovej zložitosti. **V kóde používajte namiesto číselných konštánt symbolické.** Údajové štruktúry a testovacia aplikácia musia byť naprogramované v **jazyku s manuálnou správou pamäte (t. j. bez garbage collector-u)**.

Testovacia aplikácia bude slúžiť na testovanie štruktúr a vyprodukovanie (a prípadné spracovanie) výstupu. Musí obsahovať intuitívne používateľské rozhranie (konzola alebo grafické rozhranie). Aplikácia umožní zvoliť testovanú štruktúru a spustiť sadu testov. Testy vyprodukujú CSV súbor s ich výsledkami. Po ukončení behu aplikácie musí byť pamäť **preukázateľne čistá** (nevznikli „memory leak-y“). Testy musia byť naprogramované univerzálne. Vstup do testov bude ADT štruktúry (napr. `List<T>`, nie `ArrayList<T>` ani `LinkedList<T>`)!

CSV (comma-separated values) súbor je textový súbor obsahujúci dáta oddelené čiarkami. Takéto súbory je možné ďalej jednoducho spracovávať tabuľkovými editormi (napr. MS Excel). Dáta v CSV súboroch sú jednoducho prenositeľné medzi aplikáciami. Štruktúra súboru nie je nijako pevne daná (teda koľko musí mať riadkov, koľko dát je v jednotlivých riadkoch, atď.) Význam jednotlivým riadkom/stĺpcom je potrebné určiť a zdokumentovať v dokumentácii semestrálnej práce. Viac o CSV súboroch sa dočítate napr. na <https://datahub.io/docs/data-packages/csv>.

Vyprodukované CSV súbory je potrebné spracovať (akokoľvek - ručne, pomocou tabuľkového editora, pomocou ďalšej alebo tej istej aplikácie). Metodické popísanie spôsobu spracovania dát je súčasťou dokumentácie semestrálnej práce. Pomocou popísaného postupu musí byť možné nové dáta spracovať rovnako, ako dáta predtým. Výsledky spracovania (grafy, priemery, odhady, atď.) je potrebné zdokumentovať a interpretovať (vyvodiť závery). Identifikácia relevantných veličín je súčasťou semestrálnej práce.

(5b) ADT zoznam

Otestujte výkon rôznych implementácií ADT zoznam. Zoznam musí podporovať operácie definované v Tab. 1.

Tab. 1 Operácie podporované ADT zoznam

Operácia	Parametre	Návratová hodnota
Vytvor		prázdny zoznam
Zruš		
Vlož prvý	prvok	
Vlož posledný	prvok	
Vlož na index	prvok, index	
Zruš prvý		prvok
Zruš posledný		prvok
Zruš na indexe	index	prvok
Sprístupni	index	prvok
Nastav	index, prvok	
Index prvku	prvok	index

V troch scenároch definovaných v Tab. 2 porovnajte výkon implementácie:

- poľom (`ArrayList`),
- zretäzenou pamäťou (`LinkedList`).

V každom scenári vykonajte spolu 1 000 000 operácií. Jednotlivé operácie sú v jednotlivých scenároch volané náhodne tak, aby na konci súhlasil podiel jednotlivých operácií (neplatí, že najskôr sa zavolajú operácie Vlož, potom operácie Zruš, potom operácie Sprístupni resp. Nastav a nakoniec operácia Index prvku). Ak je možné vybrať z viacerých operácií, operácie sa vyberú s rovnakou pravdepodobnosťou. Parametre do operácií sú taktiež náhodné; ako index je možné zvoliť akýkoľvek aktuálne platný index.

Tab. 2 Testovacie scenáre pre ADT zoznam

Scenár	Podiel operácií			
	<i>Vlož prvý</i> <i>Vlož posledný</i> <i>Vlož na index</i>	<i>Zruš prvý</i> <i>Zruš posledný</i> <i>Zruš na indexe</i>	<i>Sprístupni</i> <i>Nastav</i>	<i>Index prvku</i>
A	15	15	60	10
B	30	30	30	10
C	40	40	5	5

(5b) ADT prioritný front

Otestujte výkon rôznych implementácií ADT prioritný front. Prioritný front musí podporovať operácie definované v Tab. 3.

V dvoch scenároch definovaných v Tab. 4 porovnajte výkon implementácie:

- zretäzeným zoznamom neutriedeným podľa priorít,
- ľavostrannou haldou.

V každom scenári vykonajte celkovo 1 000 000 operácií. Jednotlivé operácie sú v jednotlivých scenároch volané náhodne tak, aby na konci súhlasil podiel jednotlivých

operácií (neplatí, že najskôr sa zavolajú operácie *Vlož*, potom operácie *Vyber* nakoniec operácia *Ukáž*). Parametre do operácií sú taktiež náhodné; priorita je náhodné číslo z intervalu $\langle 0; 10\,000 \rangle$.

Tab. 3 Operácie podporované ADT prioritný front

Operácia	Parametre	Návratová hodnota
Vytvor		prázdny prioritný front
Zruš		
Vlož	prvok, priorita	
Vyber		prvok s najvyššou prioritou
Ukáž		prvok s najvyššou prioritou

Tab. 4 Testovacie scenáre pre ADT prioritný front

Scenár	Podiel operácií		
	<i>Vlož</i>	<i>Vyber</i>	<i>Ukáž</i>
A	45	45	10
B	50	25	25

(10b) ADT viacrozmerné pole – matica

Vytvorte testy na porovnanie rôznych implementácií matíc. Matica musí podporovať operácie definované v Tab. 5.

Tab. 5 Operácie podporované ADT viacrozmerné pole – matica

Operácia	Parametre	Návratová hodnota
Vytvor	počet riadkov, počet stĺpcov	matica, ktorej všetky položky sú inicializované na 0
Zruš		
Sprístupni	index riadku, index stĺpca	prvok
Nastav	index riadku, index stĺpca, prvok	

Vo vami navrhnutých dvoch testovacích scenároch porovnajte výkon implementácie:

- v súvislej pamäti,
- v nesúvislej pamäti (pole polí).

V rámci dokumentácie je nutné definovať počet operácií, podiely jednotlivých operácií a povahu generovaných dát.

(20b) Testovanie ďalších štruktúr

Otestujte nižšie uvedené implementácie vybraných ADT vo vami navrhnutých testovacích scenároch a zdokumentujte počet operácií, podiely jednotlivých operácií a povahu generovaných dát:

- **(5b) obojstranne zret'azený cyklický zoznam** (implementácia musí podporovať funkcionality ADT zoznam uvedenú v Tab. 1).
- **(5b) dvojzoznam ako implementácia prioritného frontu** – otestujte vplyv dĺžky kratšieho zoznamu na výkon štruktúry (implementácia musí podporovať funkcionality ADT prioritný front uvedenú v Tab. 3).
- **(10b) množina ako bitová mapa** – otestujte vplyv veľkosti bázeovej množiny na rýchlosť jednotlivých operácií. Množina musí podporovať funkcionality uvedené v Tab. 6. Vhodne navrhnete reprezentáciu bázeovej množiny prvkov.

Tab. 6 Operácie podporované ADT množina

Operácia	Parametre	Návratová hodnota
Vytvor	bázová množina	prázdna množina
Zruš		
Vlož	prvok	
Vyber	prvok	
Patrí	prvok	pravda/nepravda
Je rovná	množina	pravda/nepravda
Je podmnožinou	množina	pravda/nepravda
Zjednotenie	množina	nová množina
Prienik	množina	nová množina
Rozdiel	množina	nová množina

Bonus (5b)

Bonusovú úlohu je možné riešiť a odovzdať po splnení aspoň prvej úrovne semestrálnej práce (testovanie ADT zoznam). Navrhnete univerzálny spôsob zadávania testov ADT a následnej tvorby scenárov vo vašej testovacej aplikácii. Testy nebudú naprogramované vopred, ale používateľ ich bude môcť zostaviť. Umožnite, aby bolo možné takéto testovacie sady uložiť a opätovne načítať a, samozrejme, spustiť.

Bodovanie semestrálnej práce

Počet bodov za semestrálnu prácu	Úroveň	Požadovaná funkcionality
5	1	Zoznamy
10	2	Úroveň 1 + prioritné fronty
20	3	Úroveň 2 + matice
40	4	Úroveň 3 + ďalšie štruktúry
+5	Bonus	Aspoň úroveň 1 + používateľom zadávané testy