Fakulta riadenia a informatiky

Informatika

Algoritmy a údajové štruktúry  
Semestrálna práca 2   
Aplikácia na spracovanie výsledkov parlamentných volieb - Dokumentácia

Andrej Držík, 5ZY022

Obsah

[Návrh aplikácie 3](#_Toc40627534)

[Popis zvolených štruktúr 4](#_Toc40627535)

[Table 4](#_Toc40627536)

[SequenceTable 5](#_Toc40627537)

[UnsortedSequenceTable 6](#_Toc40627538)

[HashTable 7](#_Toc40627539)

[Popis výpočtu každého vyhľadávcacieho, resp triediaceho kritéria 8](#_Toc40627540)

[Kritérium 8](#_Toc40627541)

[Kritéria pre územné jednotky: 8](#_Toc40627542)

[Kritéria pre politické strany: 8](#_Toc40627543)

[Kritéria pre kandidátov: 8](#_Toc40627544)

[Filter 9](#_Toc40627545)

[Výberové kritérium 9](#_Toc40627546)

[Zložitosti operácií 3 – 10 10](#_Toc40627547)

[Načítavanie údajov do tabuľky 10](#_Toc40627548)

[Napĺňanie údajov prvkom v tabuľke 10](#_Toc40627549)

[Úloha 1 – 2 10](#_Toc40627550)

[Úloha 3 (Vypísanie informácií o územných jednotkách po prefiltrovaní) 10](#_Toc40627551)

[Úloha 4 (Zoradenie územných jednotiek spĺňajúcich daný filter) 11](#_Toc40627552)

[Úloha 5 (Vypísanie výsledkov zvolených politických subjektov vo zvolených územných jednotkách) 11](#_Toc40627553)

[Úloha 6 (Zoradenie politických strán podľa výsledkov vo zvolených územných jednotkách) 12](#_Toc40627554)

[Úloha 7 (Vypísanie výberových kritérií pre politické subjekty vo zvolených územných jednotkách) 12](#_Toc40627555)

[Úloha 8 (Zoradenie územných jednotiek podľa úspešnosti zvoleného politického subjektu) 13](#_Toc40627556)

[Úloha 9 (Vypísanie výsledkov kandidáta za politický subjekt v územnej jednotke) 13](#_Toc40627557)

[Úloha 10 (Zoradenie výsledkov kandidátov za politické subjekty v územnej jednotke) 14](#_Toc40627558)

[Diagram tried 15](#_Toc40627559)

[Všeobecne 15](#_Toc40627560)

[Kritérium 16](#_Toc40627561)

[Filter 17](#_Toc40627562)

[Vyberove kritérium 18](#_Toc40627563)

[Comparator 19](#_Toc40627564)

[Sorter 20](#_Toc40627565)

[Používateľská príručka 21](#_Toc40627566)

# Návrh aplikácie

Aplikácia ma za úlohu načítať dáta z priložených CSV súborov a následne tieto dáta filtrovať a triediť podľa užívateľom zadaných parametrov.

Dáta o územných jednotkách, politických stranách a kandidátoch načítaných z CSV súborov uchovávam v HashTable. Zvolil som tak kvôli jej rýchlosti (operácie vlož, zruš, nájdi prvok majú zložitosť O(1)).

Ako kľúč v tabuľkách som zvolil kód obce/číslo politickej strany/číslo kandidáta, keďže každé z týchto čísiel je unikátne a teda aj každý kľúč bude unikátny.

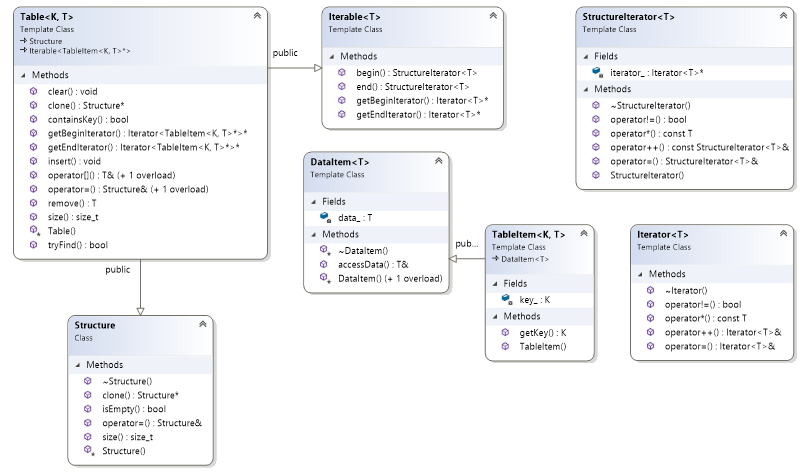
Pri filtrovaní tabuľky dáta ukladám do UnsortedSequenceTable, aby mohli byť ďalej triedené podľa zadaného kritéria. (operácie vlož, zruš majú zložitosť O(1), nájdi prvok O(N)).

Pre triedenie UnsortedSequenceTable používam metódu triedenia QuickSort. Výkon tohto algoritmu je v rozmedzí O(N\*Log2N) až O(N2). Tento však závisí od voľby pivota, ktorého hodnotu neviem vopred určiť a teda neviem určiť ani jeho skutočný výkon.

Pre filtrovanie dát som v práci použil kritéria a filtre zo zadania, ktoré su neskôr v dokumente popísané. Kritéria využívam aj v Comparator triedach využitých pri sortovaní.

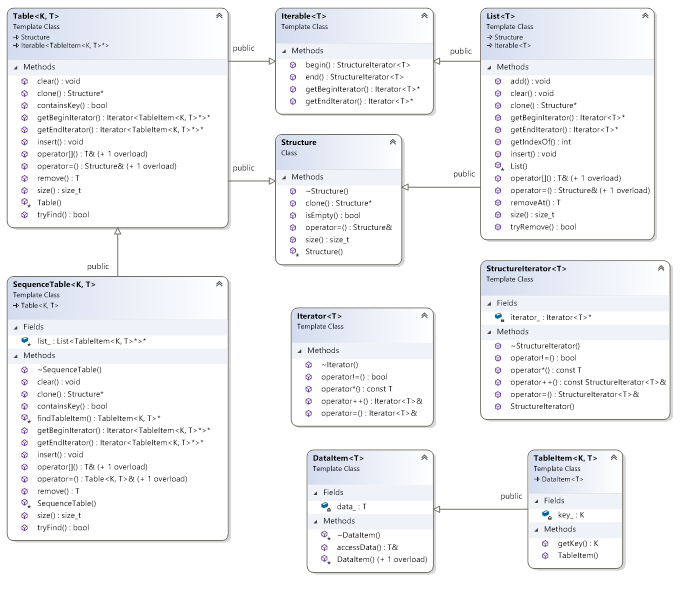
# Popis zvolených štruktúr

## Table



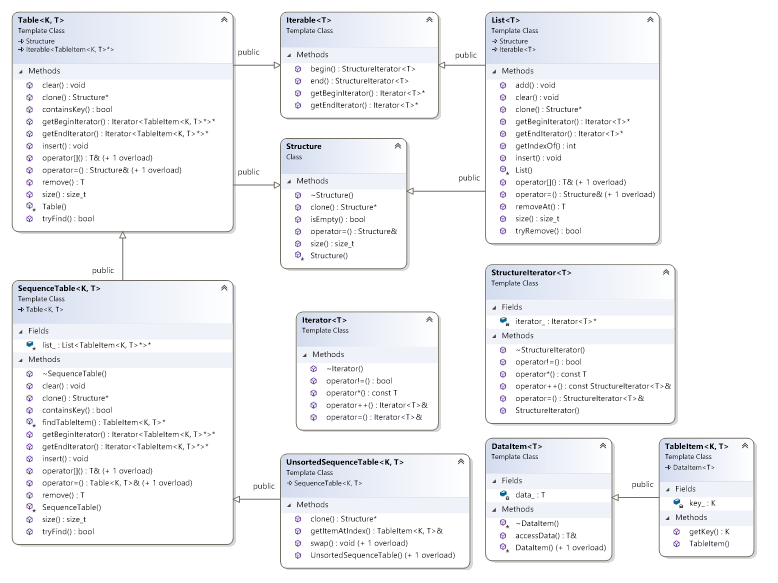
Table<K, T> je údajová štruktúra, kde každý prvok (TableItem<K, T>) je charakterizovaný jednodznačným kľúčom K a dátami T.

## SequenceTable



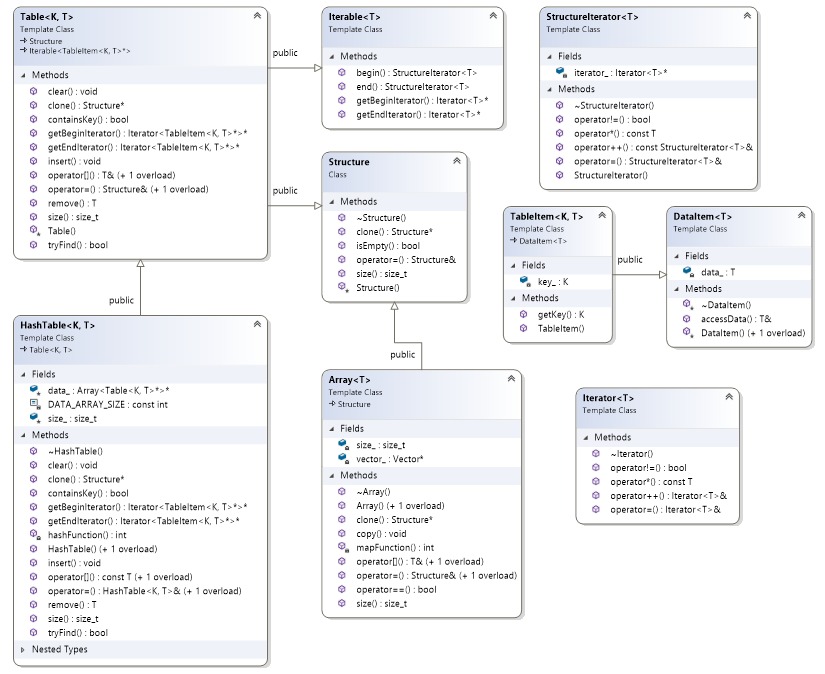
SequenceTable<K, T> dedí od Table <K, T>. Je to implementácia tabuľky ako sekvencia prvkov, ktorá spája dáta s ich kľúčom.

## UnsortedSequenceTable



UnsortedSequenceTable<K, T> dedí od SequenceTable<K, T>. Tabuľka je implementovaná implicitne implementovaným zoznamom a je neutriedená.

## HashTable



HashTable<K, T> dedí od Table<K, T>. Je to implementácia tabuľky pomocou poľa, kde pre získanie adresy každého prvku v tomto poli používam hashovaciu funkciu.

# Popis výpočtu každého vyhľadávcacieho, resp triediaceho kritéria

## Kritérium

Kritérium predstavuje funkciu, ktorá preberie vstupný objekt 𝜔𝜔 a vráti hodnotu niektorej jeho vlastnosti.

V práci som využil kritéria dané zadaním práce:

### Kritéria pre územné jednotky:

**KriteriumUJNazov** – ako vstupný parameter nepreberá nič, vracia názov danej územnej jednotky

**KriteriumUJTyp** – ako vstupný parameter nepreberá nič, vracia typ danej územnej jednotky

**KriteriumUJPrislusnost** – ako vstupný parameter preberá inú územnú jednotku, vracia hodnotu boolean, či patrí pod túto územnú jednotku alebo nie

**KriteriumUJUcast** – ako vstupný parameter nepreberá nič, vracia účasť (pomer zúčastnení voliči / voliči \* 100) ako hodnotu double v intervale <0, 100>

**KriteriumUJVolici** – ako vstupný parameter nepreberá nič, vracia počet voličov ako hodnotu integer v intervale <0, ∞>

**KriteriumUJZucastneniVolici** – ako vstupný parameter nepreberá nič, vracia počet zúčastnených voličov ako hodnotu integer v intervale <0, ∞>

**KriteriumUJOdovzdaneObalky** – ako vstupný parameter preberá spôsob odovzdania, vracia počet odovzdaných obálok daným spôsobom ako hodnotu integer v intervale <0, ∞>

**KriteriumUJPodielPlatnychHlasov** – ako vstupný parameter nepreberá nič, vracia podiel platných hlasov (platné hlasy/odovzdané obálky(spolu) \* 100) ako hodnotu double v intervale <0, 100>

**KriteriumUJPocetPlatnychHlasov** – ako vstupný parameter nepreberá nič, vracia počet platných hlasov ako hodnotu integer v intervale <0, ∞>

### Kritéria pre politické strany:

**KriteriumPSNazov** – ako parameter nepreberá nič, vracia názov politickej strany

**KriteriumPSHlasyZiskane** – ako parameter preberá územnú jednotku, vracia počet hlasov získaných v danej územnej jednotke

**KriteriumPSUspesnost** – ako parameter preberá územnú jednotku, vracia úspešnosť v danej územnej jednotke

### Kritéria pre kandidátov:

**KriteriumKMeno** – ako parameter nepreberá nič, vracia meno kandidáta v tvare Meno Priezvisko

**KriteriumKPrislusnost** – ako parameter preberá politickú stranu, vracia hodnotu boolean, či kandiduje za danú politickú stranu alebo nie

**KriteriumKPrednostneHlasy** – ako paramter preberá územnú jednotku, vracia počet získaných prednostných hlasov v tejto územenej jednotke

## Filter

Filter spolupracuje s nejakým kritériom. Filter preberie vstup a vráti taký výstup, kde všetky objekty vstupu spĺňajú toto kritérium ako UnsortedSequenceTable.

V práci som využil filtre dané zadaním práce:

**FilterUJNazov** – ako paramter preberá názov, spolupracuje s kritériom KriteriumUJNazov

**FilterUJTyp** – ako parameter preberá typ, spolupracuje s kritériom KriteriumUJTyp

**FilterUJPrislusnost** – ako parameter preberá hodnotu boolean a územnú jednotku, spolupracuje s kritériom KriteriumUJPrislusnost

**FilterUJUcast** – ako parameter preberá dve hodnoty double ako min a max, spolupracuje s kritériom KriteriumUJUcast

**FilterUJVolici** – ako parameter preberá dve hodnoty int ako min a max, spolupracuje s kritériom KriteriumUjVolici

**FilterPSNazov** – ako parameter preberá názov, spolupracuje s kritériom KriteriumPSNazov

**FilterKMeno** – ako parameter preberá meno, spolupracuje s kritériom KriteriumKMeno

**FilterKPrislusnost** – ako parameter preberá politickú stranu, spolupracuje s kritériom KriteriumKPrislusnost

## Výberové kritérium

Výberové kritérium je špeciálny druh filtra, ktorý vráti najviac jeden prvok z množiny, na ktorú je aplikované.

V práci som využil filtre dané zadaním práce:

**VKriteriumMaxPSHlasy** – ako parameter preberá územnú jednotku, spolupracuje s kritériom KriteriumPSHlasyZiskane

**VKriteirumMinPSHlasy** – ako parameter preberá územnú jednotku, spolupracuje s kritériom KriteriumPSHlasyZiskane

# Zložitosti operácií 3 – 10

## Načítavanie údajov do tabuľky

Dáta načítavané do akejkoľvek tabuľky (územné jednotky, politické strany, kandidáti) sú získané z csv súboru, kde jeden riadok zodpovedá jednému prvku a súbor má N riadkov. Dáta sú načítavané do HashTable, ktorá má zložitosť operácie vlož O(1). Vkladanie dát do tabuľky má teda zložitosť O(N), kde N je počet prvkov vkladaných do tabuľky.

## Napĺňanie údajov prvkom v tabuľke

Pri napĺňaní údajov do jednotlivých prvkov v akejkoľvek tabuľke sa musí spristúpniť daný prvok, čo má zložitosť O(1). Napĺňanie údajov má teda zložitosť k\*O(N), kde N je počet prvkov tabuľky, k počet vkladaných údajov (k\*O(1)).

## Úloha 1 – 2

Úlohy 1 až 2 som sa rozhodol nezahrnúť do práce z dôvodu použitia funkcionality týchto úloh v nasledujúcich úlohách.

## Úloha 3 (Vypísanie informácií o územných jednotkách po prefiltrovaní)

Užívateľ zadá parametre 2 filtrom určeným zo zadania a následne si vyberie dodatočný filter. Zložitosť prvého filtrovania je O(UJ), kde UJ je počet prvkov v tabuľke územných jednotiek. Výsledok prvého filtrovania sa znovu filtruje cez druhý filter, čo má zložitosť O(UJ1), kde UJ1 je počet prvkov po prvom filtrovaní. Výsledok sa znovu filtruje so zložitosťou O(UJ2), kde UJ2 je počet prvkov po druhom filtrovaní. Výsledná zložitosť filtrovania je **O(UJ + UJ1 + UJ2)**.

Následne sa vypíšu údaje o každom prvku vo výslednej tabuľke. Toto má zložitosť **k\*O(UJ3),** kde UJ3 je počet prvkov po filtrovaní posledným filtrom a k je počet údajov (k\*O(1)).

Výsledná zložitosť operácií v úlohe 3 (Vypísanie informácií o územných jednotkách po prefiltrovaní) je pri použití všetkých filtrov **O(UJ + UJ1 + UJ2 + k\*UJ3).**

## Úloha 4 (Zoradenie územných jednotiek spĺňajúcich daný filter)

Užívateľ zadá parametre 4 filtrom určeným zo zadania. Zložitosť filtrovania pri využití všetkých 4 filtrov je **O(UJ + UJ1 + UJ2 + UJ3)**, kde UJ je počet prvkov v tabuľke územné jednotky, UJ1 je počet prvkov po prvom filtrovaní, UJ2 počet prvkov po druhom filtrovaní, UJ3 počet prvkov po treťom filtrovaní.

Výsledok filtrovania následne zoradíme pomocou QuickSort podľa užívateľom zadaného kritéria. Skutočnú zložitosť algoritmu QuickSort nemôžeme určiť, preto budem ďaľej v dokumentácií pracovať s jeho hornou zložitosťou O(N2). Zložitosť triedenia po filtrovaní je teda **O(UJ42),** kde UJ4 je počet prvkov po štvrtom filtrovaní.

Následne sa vypíšu údaje o každom prvku vo výslednej tabuľke. Toto má zložitosť **k\*O(UJ4)**, kde UJ4 je počet prvkov po filtrovaní posledným filtrom a k je počet údajov (k\*O(1)).

Výsledná zložitosť operácií v úlohe 4(Zoradenie územných jednotiek spĺňajúcich daný filter) je pri použití všetkých filtrov je **O(UJ + UJ1 + UJ2 + UJ3 + UJ42 + k\*UJ4).**

## Úloha 5 (Vypísanie výsledkov zvolených politických subjektov vo zvolených územných jednotkách)

Užívateľ zadá parametre 4 filtrom určeným zo zadania. Zložitosť filtrovania pri využití všetkých 4 filtrov je **O(UJ + UJ1 + UJ2 + UJ3)**, kde UJ je počet prvkov v tabuľke územné jednotky, UJ1 je počet prvkov po prvom filtrovaní, UJ2 počet prvkov po druhom filtrovaní, UJ3 počet prvkov po treťom filtrovaní.

Ďalej užívateľ zadá parameter filtra pre politickú stranu (názov). Zložitosť filtrovania je **O(PS)**, kde PS je počet prvkov v tabuľke politické strany.

Následne sa vypíšu údaje kritéria (získané hlasy politickej strany v danej územnej jednotke) každého prvku vo výslednej tabuľke filtrovania politických strán a spolu s tým údaje o každom prvku vo výslednej tabuľke územných jednotiek. Toto má zložitosť **O(PS1\*k\*UJ4)**, kde PS1 je počet prvkov po filtrovaní politických strán, k je počet údajov(k\*O(1)), ktoré sa vypisujú a UJ4 je počet prvkov po štvrtom filtrovaní územných jednotiek.

Výsledná zložitosť operácií v úlohe 5 (Vypísanie výsledkov zvolených politických strán vo zvolených územných jednotkách) je pri použití všetkých filtrov **O(UJ + UJ1 + UJ2 + UJ3 + PS + PS1\*k\*UJ4).**

## Úloha 6 (Zoradenie politických strán podľa výsledkov vo zvolených územných jednotkách)

Užívateľ zadá parametre 4 filtrom určeným zo zadania. Zložitosť filtrovania pri využití všetkých 4 filtrov je **O(UJ + UJ1 + UJ2 + UJ3)**, kde UJ je počet prvkov v tabuľke územné jednotky, UJ1 je počet prvkov po prvom filtrovaní, UJ2 počet prvkov po druhom filtrovaní, UJ3 počet prvkov po treťom filtrovaní.

Uložím si politické strany do pomocnej tabuľky UnsortedSequenceTable, to má zložitosť **O(PS)**, kde PS je počet politických strán.

Každému prvku v tejto pomocnej tabuľke politických strán nastavím pomocný súčet ako súčet hlasov získaných v územných jednotkách po filtrovaní. Toto má zložitosť **O(PS\*UJ4)**, kde PS je počet politických strán a UJ4 počet územných jednotiek po 4. filtrovaní.

Pomocnú tabuľku politických strán následne zoradím pomocou QuickSort podľa hodnoty pomocného súčtu získaného v prechádzajúcom bode. Zložitosť triedenia po filtrovaní je teda **O(PS2).**

Následne sa vypíše súčet počtu hlasov získaných vo výsledných územných jednotkách pre každú politickú stranu a spolu s ním počet získaných hlasov v jednotlivých územných jednotkách pre každú politickú stranu. Toto má zložitosť **O(PS\*k\*UJ4)**, kde PS je počet politických strán, k je počet údajov(k\*O(1)), ktoré sa vypisujú a UJ4 je počet prvkov po štvrtom filtrovaní územných jednotiek.

Výsledná zložitosť operácií v úlohe 6 (Zoradenie politických strán podľa výsledkov vo zvolených územných jednotkách) je pri použití všetkých filtrov **O(UJ + UJ1 + UJ2 + UJ3 + PS + PS\*UJ4 + PS2 + PS1\*k\*UJ4).**

## Úloha 7 (Vypísanie výberových kritérií pre politické subjekty vo zvolených územných jednotkách)

Užívateľ zadá parametre 4 filtrom určeným zo zadania. Zložitosť filtrovania pri využití všetkých 4 filtrov je **O(UJ + UJ1 + UJ2 + UJ3)**, kde UJ je počet prvkov v tabuľke územné jednotky, UJ1 je počet prvkov po prvom filtrovaní, UJ2 počet prvkov po druhom filtrovaní, UJ3 počet prvkov po treťom filtrovaní.

Pre každý prvok z výsledku vyfiltrovaných územných jednotiek sa vypíše hodnota výberového kritéria(politická strana najviac hlasov) a hodnota výberového kritéria(politická strana najmenej hlasov). Zložitosť takýchto výberových kritérií je O(PS). Toto má teda zložitosť **O(UJ4\*(2\*PS)).**

Výsledná zložitosť operácií v úlohe 7 (Vypísanie výberových kritérií pre politické strany vo zvolených územných jednotkách) po použití všetkých filtrov je **O(UJ + UJ1 + UJ2 + UJ3 + UJ4\*2\*PS).**

## Úloha 8 (Zoradenie územných jednotiek podľa úspešnosti zvoleného politického subjektu)

Užívateľ zadá parametre 4 filtrom určeným zo zadania. Zložitosť filtrovania pri využití všetkých 4 filtrov je **O(UJ + UJ1 + UJ2 + UJ3)**, kde UJ je počet prvkov v tabuľke územné jednotky, UJ1 je počet prvkov po prvom filtrovaní, UJ2 počet prvkov po druhom filtrovaní, UJ3 počet prvkov po treťom filtrovaní.

Ďalej užívateľ zadá parameter filtra pre politickú stranu (názov). Zložitosť filtrovania je **O(PS)**, kde PS je počet prvkov v tabuľke politické strany.

Vytvorím si pomocnú UnsortedSequenceTable, kde ukladám ako kľúč kód územnej jednotky a dáta ako úspešnosť daného politickej strany v danej územnej jednotke. Vkladanie údajov do takejto tabuľky má zložitosť **O(UJ4)**, kde UJ4 je počet prvkov po 4. filtrovaní územných jednotiek.

Pomocnú tabuľku následne zoradíme pomocou QuickSort podľa hodnoty double. Zložitosť triedenia po filtrovaní je teda **O(UJ42).**

Následne sa vypíšu údaje o každom prvku v tejto utriedenej pomocnej tabuľke. Toto má zložitosť **k\*O(UJ4),** kde UJ4 je počet prvkov po filtrovaní posledným filtrom a k je počet údajov (k\*O(1)).

Výsledná zložitosť operácií v úlohe 8 (Zoradenie územných jednotiek podľa úspešnosti zvoleného politického subjektu) je teda **O(UJ + UJ1 + UJ2 + UJ3 + PS + UJ4 + UJ42 + k\*UJ4)**

## Úloha 9 (Vypísanie výsledkov kandidáta za politický subjekt v územnej jednotke)

Užívateľ zadá parametre 4 filtrom určeným zo zadania. Zložitosť filtrovania pri využití všetkých 4 filtrov je **O(UJ + UJ1 + UJ2 + UJ3)**, kde UJ je počet prvkov v tabuľke územné jednotky, UJ1 je počet prvkov po prvom filtrovaní, UJ2 počet prvkov po druhom filtrovaní, UJ3 počet prvkov po treťom filtrovaní.

Ďalej užívateľ zadá parameter filtra pre kandidáta (názov). Zložitosť filtrovania je **O(K)**, kde K je počet prvkov v tabuľke politické strany.

Následne sa vypíšu údaje kritéria (získané hlasy kandidáta v danej územnej jednotke) každého prvku vo výslednej tabuľke filtrovania kandidáta a spolu s tým údaje o každom prvku vo výslednej tabuľke územných jednotiek. Toto má zložitosť **O(K1\*k\*UJ4)**, kde K1 je počet prvkov po filtrovaní politických strán, k je počet údajov(k\*O(1)), ktoré sa vypisujú a UJ4 je počet prvkov po štvrtom filtrovaní územných jednotiek.

Výsledná zložitosť operácií v úlohe 9 (Vypísanie výsledkov kandidáta za politický subjekt v územnej jednotke) je pri použití všetkých filtrov **O(UJ + UJ1 + UJ2 + UJ3 + K + K1\*k\*UJ4).**

## Úloha 10 (Zoradenie výsledkov kandidátov za politické subjekty v územnej jednotke)

Užívateľ zadá parametre 4 filtrom určeným zo zadania. Zložitosť filtrovania pri využití všetkých 4 filtrov je **O(UJ + UJ1 + UJ2 + UJ3)**, kde UJ je počet prvkov v tabuľke územné jednotky, UJ1 je počet prvkov po prvom filtrovaní, UJ2 počet prvkov po druhom filtrovaní, UJ3 počet prvkov po treťom filtrovaní.

Ďalej užívateľ zadá parameter filtra pre kandidáta (príslušnosť k politickej strane). Zložitosť filtrovania je **O(K)**, kde K je počet prvkov v tabuľke politické strany.

Uložím si politické strany do pomocnej tabuľky UnsortedSequenceTable, to má zložitosť **O(K1)**, kde K1 je počet politických strán.

Každému prvku v tejto pomocnej tabuľke kandidátov nastavím pomocný súčet ako súčet hlasov získaných v územných jednotkách po filtrovaní. Toto má zložitosť **O(K1\*UJ4)**, kde K1 je počet kandidátov po prefiltrovaní a UJ4 počet územných jednotiek po 4. filtrovaní.

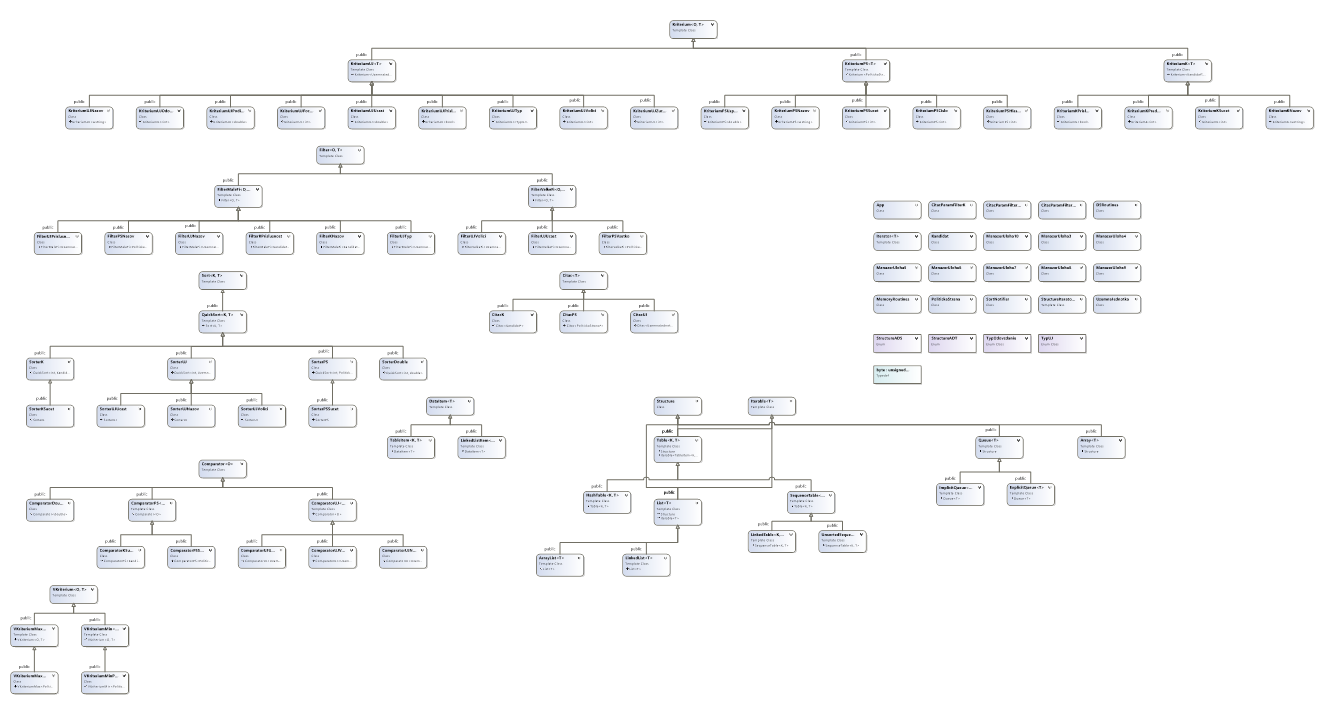
Pomocnú tabuľku kandidátov následne zoradím pomocou QuickSort podľa hodnoty pomocného súčtu získaného v prechádzajúcom bode. Zložitosť triedenia po filtrovaní je teda **O(K2).**

Následne sa vypíše súčet počtu hlasov získaných vo výsledných územných jednotkách pre každého kandidáta a spolu s ním počet získaných hlasov v jednotlivých územných jednotkách pre každého kandidáta. Toto má zložitosť **O(K1\*k\*UJ4)**, kde K1 je počet politických strán, k je počet údajov(k\*O(1)), ktoré sa vypisujú a UJ4 je počet prvkov po štvrtom filtrovaní územných jednotiek.

Výsledná zložitosť operácií v úlohe 10 (Zoradenie výsledkov kandidátov za politické subjekty v územnej jednotke) je pri použití všetkých filtrov **O(UJ + UJ1 + UJ2 + UJ3 + K + K1 + K1\*UJ4 + K2 + K1\*k\*UJ4).**

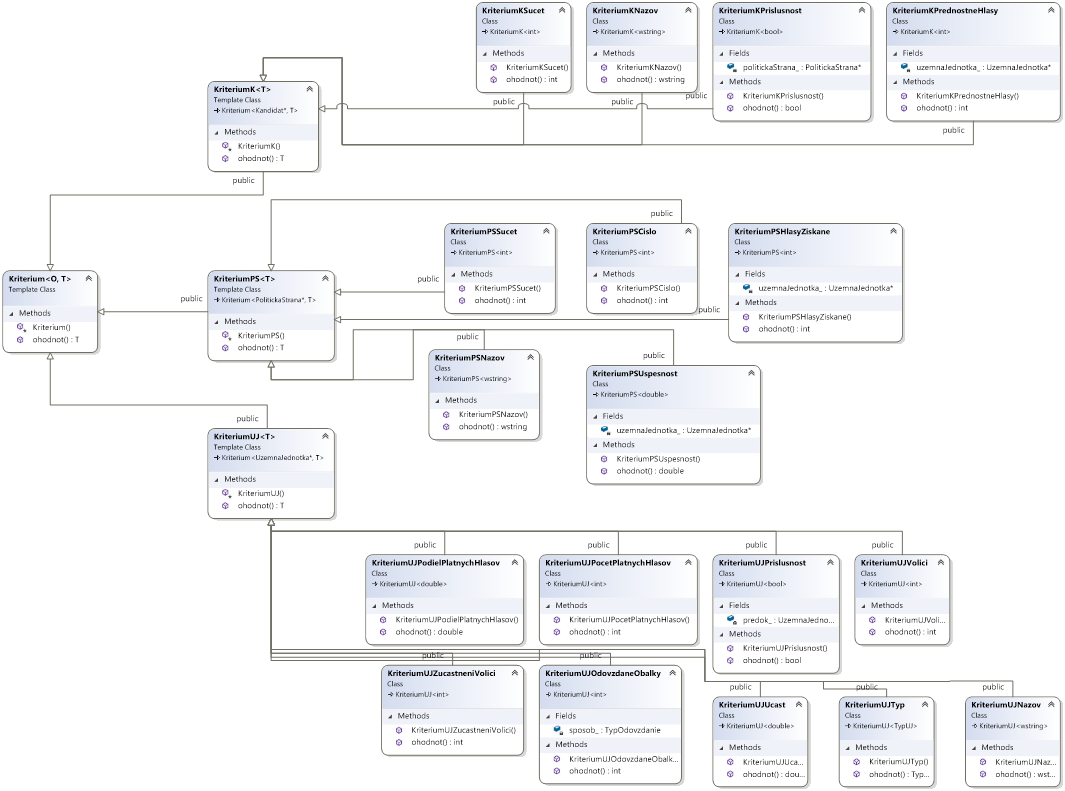
# Diagram tried

## Všeobecne

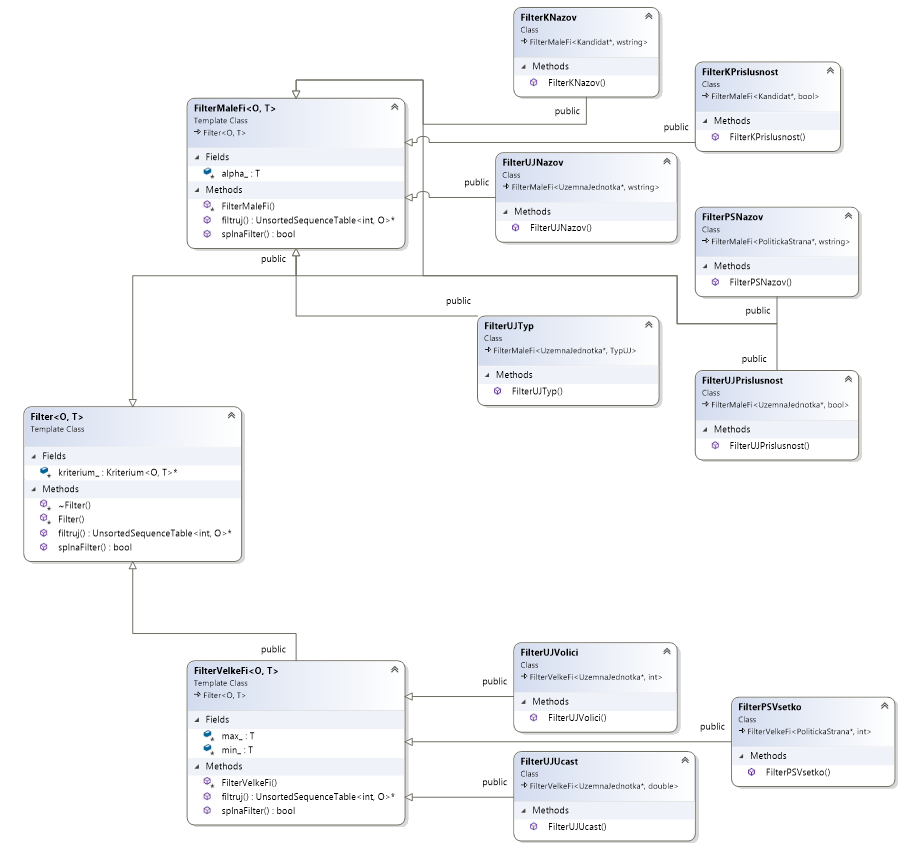


Celý diagram : <Diagram.cd>

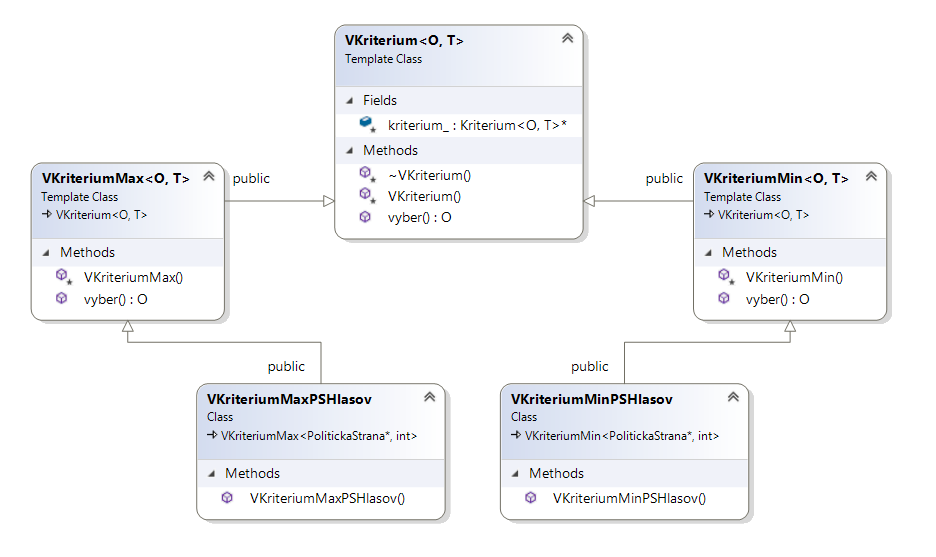
## Kritérium



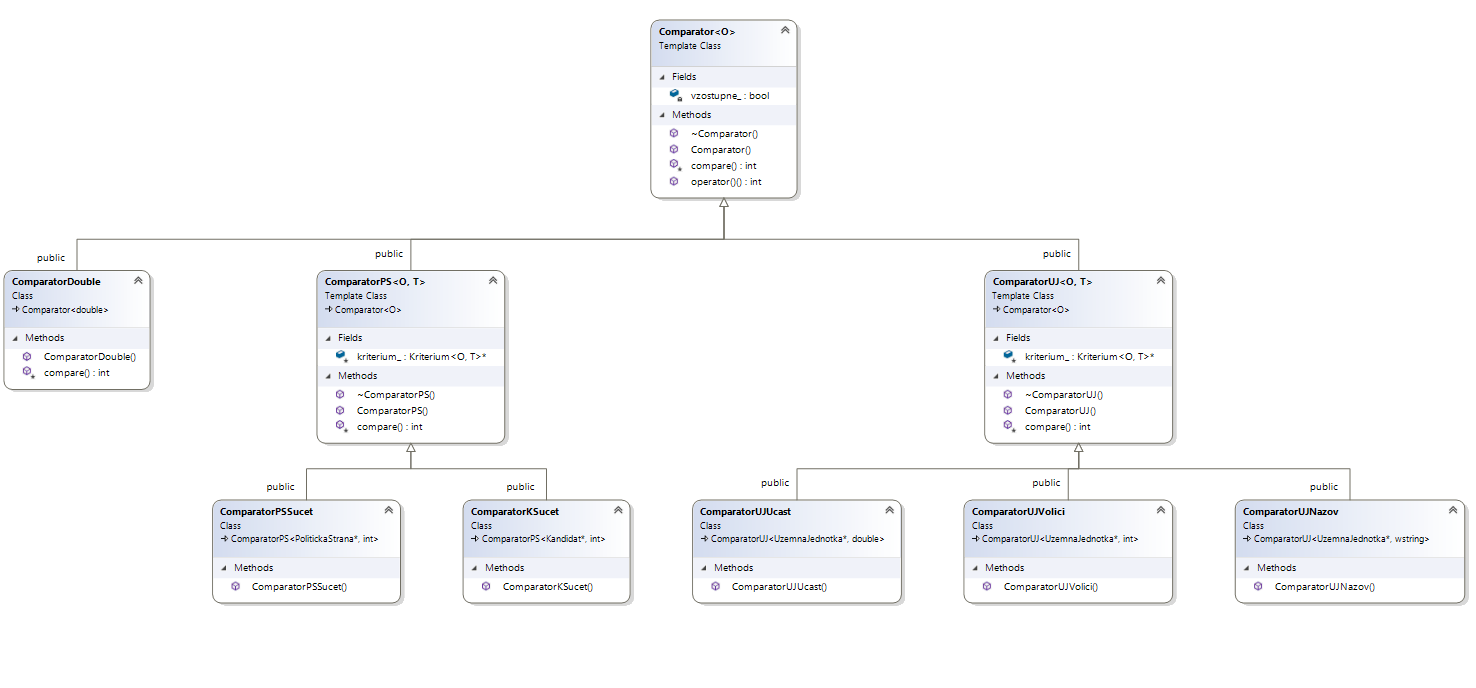
## Filter



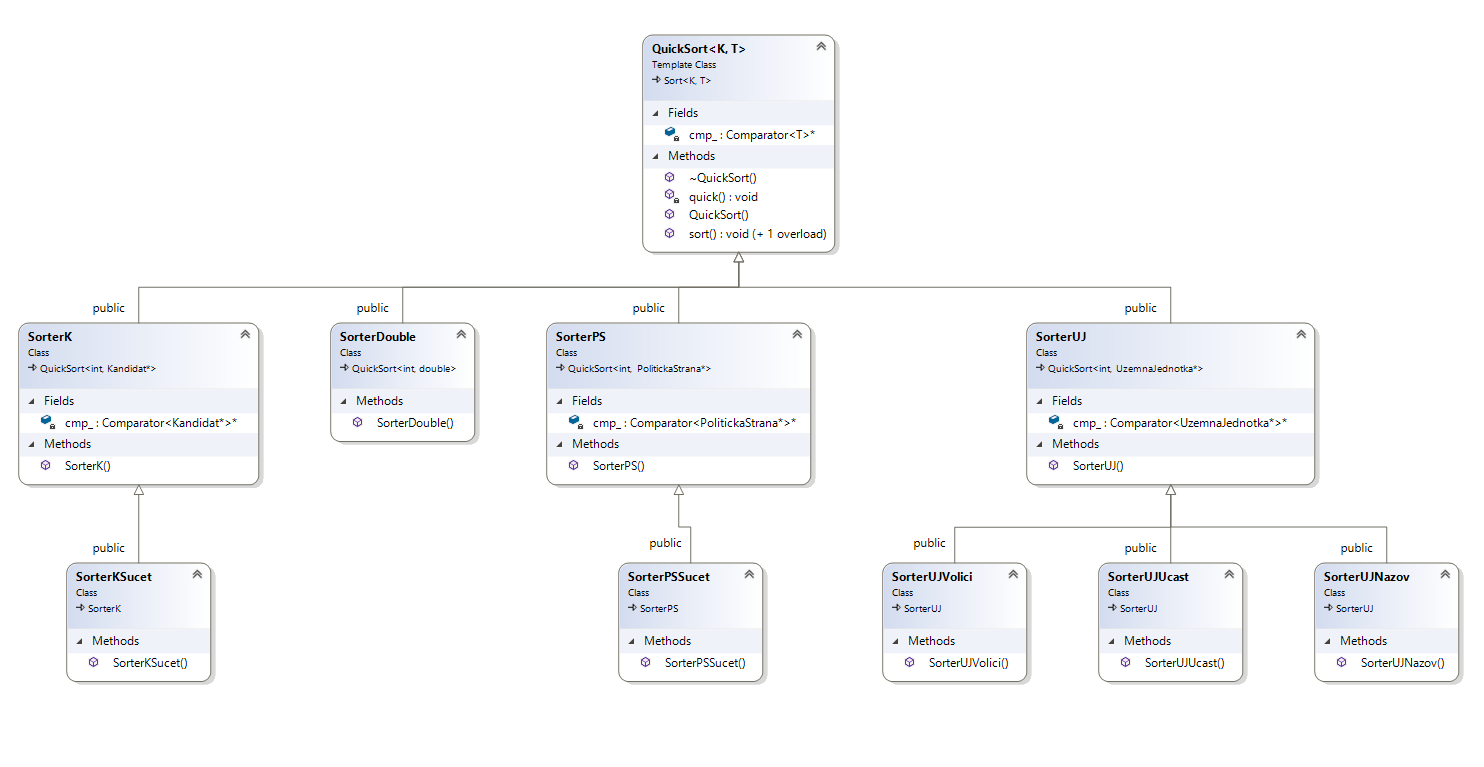
## Vyberove kritérium



## Comparator



## Sorter



# Používateľská príručka

Pre ovládanie aplikácie som sa rozhodol vytvoriť jednoduché konzolové rozhranie, kde má užívateľ po spustení aplikácie na výber, ktorú úlohu chce testovať. Po zvolení úlohy je vyzvaný k zadaniu parametra, taktiež má na výber, či daný filter chce použiť, alebo nie. Ak sa rozhodne daný filter nepoužiť, zadá „x” ako paramter. Po skončení konkrétneho testu sa znovu otvorí hlavné menu a užívateľ má na výber, či chce testovať inú úlohu , alebo zadá ako parameter “0” a tým sa aplikácia vypne.