|  |
| --- |
| Univerzita Pardubice |
| Datové struktury |
| Semestrální práce B |
|  |
| **Matěj Trakal** |
| **23.4.2012** |

|  |
| --- |
| Využití R-stromu pro bodové a intervalové vyhledávání. |

Vypracování semestrální práce

Pro semestrální práci jsem zvolil programovací jazyk C# a IDE Visual Studio 2010.

# Použité datové struktury

## R-strom

Jako vstupní strukturu pro vybudování R-stromu jsem použil strukturu pole s kapacitou (List<RVrchol>) kam se postupně vkládají vrcholy a nad nimi je následně po seřazení vybudován R-strom. Pro samotné udržení stromu pak využívám dynamické struktury RVrchol, kterou označím jako kořen a pod ní jsou uchovány pomocí odkazů všichni potomci.

Pro seřazení dat do správného pořadí pro využití R-stromu je využito Z-křivky (Mortonovi křivky).



Obrázek : ADT R-strom

### RVrchol

RVrchol je struktura zapouzdřující právě jeden vrchol nebo list ve struktuře stromu.

Třída vrcholů obsahuje opět datovou složku, klíč (nakonec nevyužit, ale připraven pro funkci) – vyhledávání dle klíče apod., obdélníkovou oblast, která obaluje nejmenší možnou oblast kolem cesty, pole potomků a ukazatel na rodiče pro případ, že by bylo třeba traverzovat zpět (v nějaké pokročilejší operaci, než kterých je využito v této práci).

Seznam potomků je realizován pomocí ADT List (pole s kapacitou), kde operace přidání má složitost O(1). Většina ostatních operací, kromě přístupu pomocí indexu O(1), složitost O(n). V tomto případě se ovšem využívá pouze budování stromu a následně jeho průchod, tedy složitost zůstává v obou případech O(1).



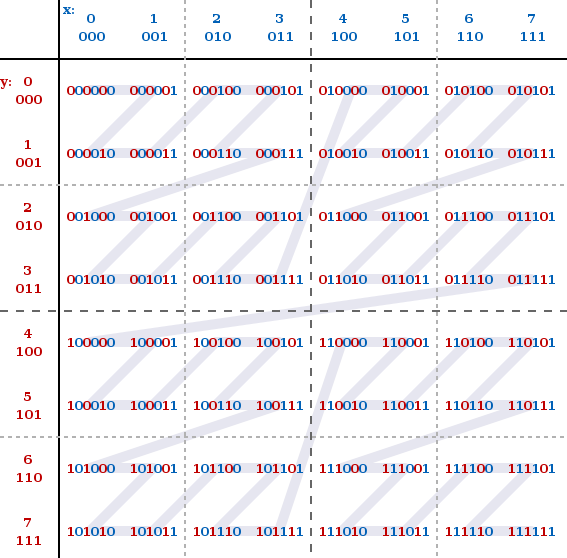
Obrázek : Struktura RVrchol

# Procházení prvků (Z-order)

Jelikož jsem nenalezl rozumný způsob, jak porovnávat v Z-průchodem datový typ Double ve kterém mám uchovány GPS souřadnice, zvolil jsem metodu, která by neměla výrazně ovlivnit výsledek a to, že jsem GPS souřadnice vynásobil konstantou a ty následně převedl na datový typ integer, který jsem zaokrouhlil.

Vzhledem k samotné podstatě R-stromu a oblastí, které se generují přes několik prvků, nemělo by docházet k velkým odchylkám, resp. neměli by být ovlivněny výsledky touto operací.

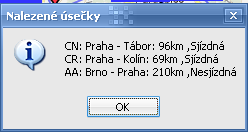
Datové typy integer jsou převedeny na bitové pole, kde se střídají po bitu y a x souřadnice, čímž se vytvoří posloupnost bitů, které jsou vhodné pro řazení Z-křivky.



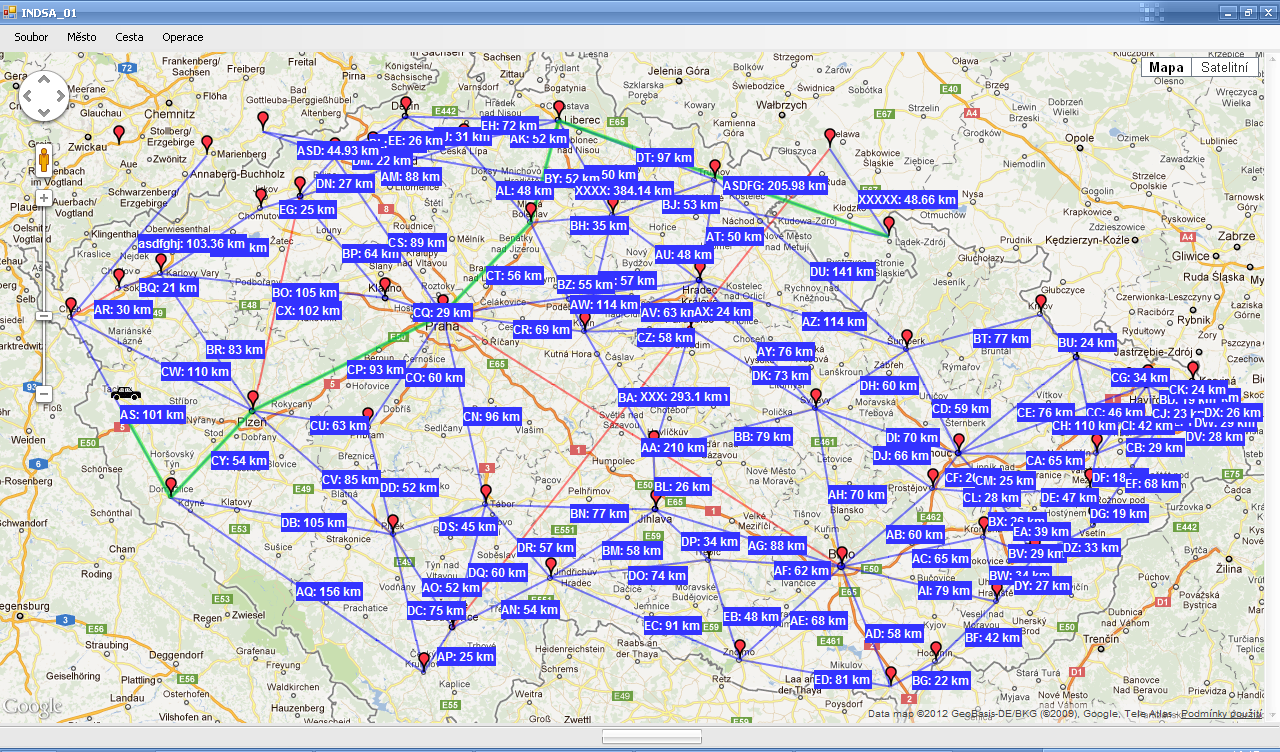
Obrázek : Průchod z-křivky [http://en.wikipedia.org/wiki/File:Z-curve.svg]

# GUI

Pro výstup jsem použil pouze slovní výpis vrcholů a informací o nich do vyskakovacího okna, jelikož pokud bych měl vykreslovat všechny obdélníky do podkladové mapy, nejspíše by v ní nebylo k nalezení již nic jiného a slovní popis, kde jsou ve výsledku vypsány body, které vyhledávací algoritmus považuje za vhodné výsledné kandidáty.



Obrázek : Zobrazené cesty zasahující do hledané oblasti.



Obrázek 5: Výsledná aplikace se zobrazenou nalezenou cestou.