

Zhodnotenie zimného semestra pre ročníkový projekt

Téma

Témou môjho ročníkového projektu je kompresia dát pomocou techník spojených s dynamických bez-kontextových gramatík.

Ciele

Cieľ projektu sa dá rozdeliť na dve separátne časti:

- Teoretická časť – štúdium súčasných techník pre kompresiu dát pomocou dynamických bez-kontextových gramatík
- Praktická časť – prípadný návrh a implementácia algoritmu pre dosiahnutie efektívnejšej kompresie

Dosiahnuté výsledky

Počas zimného semestra som kládol dôraz najmä na prvú – teoretickú – časť zo spomenutých cieľov môjho ročníkového projektu štúdiom danej problematiky. Môj progres dosiahnutých míľnikov potrebných na porozumenie danej témy rozdeliť do nasledujúcich bodov:

- Úvod do problematiky
 - Prvým krokom bol úvod v ktorom som sa sústredil na základné piliere kompresie dát pomocou bez-kontextových gramatík ako sú:
 - Bez-kontextové gramatiky
 - Chomského normálny tvar
 - Straight-line programy
 - Orientované acyklické grafy
 - Problém najhlbšieho spoločného predka
- Teória v pozadí kompresie dát
 - Za každou efektívnou technikou stojí istá dávka teórie na ktorej sú postavené hlavné myšlienky. Pre kompresiu dát je potrebné porozumenie témam:
 - Teória informácie
 - Entropia
 - Problém najmenšej bez-kontextovej gramatiky
- Analýza existujúcich algoritmov
 - Keďže kompresia dát sa rozdeľuje do skupín offline (sú potrebné celé dáta na kompresiu) a online (kompresia počas čítania vstupu) algoritmov, tak som postupoval:
 - Porozumením jednoduchších offline algoritmov ako napr.
 - Lempel–Ziv–Welch (LZW)
 - Štúdium zložitejších online algoritmov
 - FOLCA plne online algoritmus (jeden z najefektívnejších dostupných)

- SOLCA vylepšenie FOLCA pre beh s konštantnou priestorovou zložitostou

Plány pre budúcu prácu

V priebehu nasledujúceho semestra mám v úmysle najmä hlbšie štúdium a analýzu FOLCA a SOLCA algoritmov s víziou možnej príležitosti na zefektívnenie kompresie. V prípade užitočného nápadu aj jeho následná implementácia.