

# Specifikace zápočtového programu (PRG030)

## KOMPRESI TEXTU

Petr Koupý, INF 1/X/32

Úlohou programu bude komprimovat a dekomprimovat textové soubory. Program bude ovládán výhradně pomocí parametrů načítaných z příkazové řádky, aby se dal používat v dávkových souborech. Parametrem bude název vstupního a výstupního souboru, stanovení kompresního nebo dekompresního režimu, případně také nějaký příznak debug režimu, kdy program bude vypisovat informace o průběhu zpracování do log souboru nebo na obrazovku. Podrobný návod pro ovládání a zápis parametrů bude obsažen v dokumentaci.

Jako kompresní algoritmus bude použito Huffmanovo kódování. Na začátku celého procesu se projde textový dokument a vytvoří se tabulka frekvenčního výskytu jednotlivých znaků. Z ní se následně vygeneruje binární strom. Potom se text druhým průchodem za pomoci stromu zkomprimuje. Do výstupního souboru budou uložena zkomprimovaná data a nezbytné informace pro rekonstrukci stromu. Dekompresní proces napřed rekonstruuje strom a následně jedním průchodem vytvoří původní soubor.

Program bude umět načíst a zpracovat libovolný znak ASCII tabulky. Ve výstupních řetězcích bude potřeba zachytit pouze 2 stavy – zda v binárním stromě zvolit levého nebo pravého syna stávajícího uzlu. Proto bude kvůli větší úspoře místa vhodné výsledný soubor ukládat v binárním režimu.