

Dokumentace úlohy **CST: C Stats v Python 3** do IPP 2014/15

Jméno a příjmení: **Václav Bayer**

Login: **xbayer05**

Úvod

Při tvorbě projektu nebylo využito objektového návrhu. Skript by se dal dělit do několika následujících částí: zpracování parametrů ze vstupu, ošetření získaných parametrů-jsou-li korektní vzhledem k zadání, prohledávání a vytisknutí výsledků.

Implementace

Popis postupu implementace zadané úlohy.

Zpracování parametrů

Z počátku byla snaha využít `argparse` ale kvůli určitým nedostatkům bylo použito zpracování pomocí regulárních výrazů. Cyklem `for` je brán argument po argumentu a porovnáván s povolenými možnostmi. Pokud se zde vyskytuje jakýkoli nepovolený argument, program je ukončen s chybou. Získané argumenty a jejich hodnoty jsou uloženy jako globální proměnné.

Ošetření parametrů

Funkce `check_args()` slouží k ošetření argumentů dle zadání. Probíhá zde kontrola, zda nedošlo k nepovolené kombinaci argumentů, zda se má tisknout nápověda, zda se některé argumenty neopakují, nebo naopak nechybí, atd. Další důležitou částí této funkce je ošetření parametru `input`, jeho hodnoty, kontrola vůči parametru `nosubdir` a následné vyhledání souborů, adresářů a popřípadě podadresářů. Výstupem funkce je pole souborů, které je předáno funkci `open_file()`.

Zahájení hledání

Funkce `open_file()` ze získaného pole bere soubor po souboru, se kterým dále pracuje. Nejprve se pokouší otevřít soubor s následným čtením obsahu, v případě neúspěchu je skript ukončen s chybou. Po úspěšném otevření, proběhne kontrola, zda je zadán argument `p` a podle toho se do slovníku přidá výsledná cesta k souboru. Je-li obsah souboru prázdný, hledání se spouští vůbec nebude a automaticky je k danému souboru ve slovníku přiřazena hodnota nula, v opačném případě se dle zadaného jednoduchého argumentu provede požadovaný typ hledání.

Ve všech případech požadovaného hledání, až na hledání (`search_words()`) daného vzoru (argument `w` s hodnotou) je z funkce `open_file()` volána funkce `lex_parser()`, které je předán jako parametr obsah daného souboru. Tato funkce bere znak po znaku a pomocí rozpoznávání stavů vypustí z textu veškeré komentáře, makra a řetězce a filtrovaný text vrátí k dalšímu hledání. V průběhu dochází i k počítání znaků komentářů a znaků komentářů v makrech (rozšíření `COM`), výsledek je uložen v globální proměnné `count`. Funkce sice není úplně nejprůhledněji implementována, ale pro implementaci rozšíření stačilo poupravit pouze jeden řádek kódu. Pro větší přehlednost a lepší porozumění byla snaha kód dostatečně okomentovat. Po vykonání funkce je přefiltrovaný text navrácen funkci `open_file()` pro další zpracování.

Má-li skript vypsat počet znaků komentáře, je použita hodnota `count` z předešlého filtrování a další hledání se již neprovádí. V případě vypsání počtu klíčových slov (`search_keys()`), identifikátorů (`search_iden()`) nebo operátorů (`search_operators()`), je brán filtrovaný text a předán jako parametr funkci, která odpovídá požadavku.

Počet nalezených výskytů je přiřazen danému souboru ve slovníku a po analýze všech souborů je ukončena funkce `open_file()`, která vrací daný slovník.

Vytisknutí výsledků

Posledním krokem je funkce `print_result()`, která má za úkol vytisknout výsledky analýzy. Funkce získá jako parametr slovník s našimi výsledky, který lexikograficky seřadí a sečte všechny počty výskytů. Dále se provede výpočet délek správného odsazení a výsledek se vytiskne na standartní výstup nebo do souboru (ovlivněno přepínačem `output`). Má-li být proveden tisk do souboru a nepodaří se, skončí skript s chybou. Jestliže proběhne vše v pořádku, skript je ukončen s hodnotou nula.