

Jméno a příjmení: Milan Gardáš

Login: xgarda04

Postup řešení

Zpracování parametrů

Zpracování probíhá ihned po spuštění skriptu za pomoci jednoduchého cyklu a proměnné `sys.argv`, ze které zjišťuji, které parametry byly zadány, a následně nastavuji pomocné proměnné, jež reprezentují jednotlivé parametry. Pokud je zadána nepovolená kombinace parametrů nebo je zadán neznámý parametr, běh skriptu je okamžitě ukončen s odpovídajícím chybovým kódem. Parametry `--input`, `--output`, `--pretty-xml` a `--max-par` jsou rozděleny na pomoci indexu, za předpokladu svého zadání se správnou syntaxí.

Zpracování souboru

Vstupem může být nejen soubor, ale i složka. V případě zadání pouze souboru je situace velmi jednoduchá, protože ho stačí pouze otevřít a případně ověřit to zda se opravdu otevřel korektně. Pokud je zadána složka, je situace o něco komplikovanější, z důvodu jejího prohledání a zanořování do podadresářů. Na tento úkol využívám funkci `os.walk`. Soubory, které takto naleznou, postupně uložím do seznamu. Soubory otevírám v cyklu, ve kterém je zároveň předám funkcím, které obsah souboru zpracovávají a následně vypisují výsledek.

Detekce hlaviček funkcí a získání jejich informací

Tato úloha ke správné funkčnosti vyžaduje ze vstupního souboru odstranění řetězců, maker, komentářů a těla funkcí. Tyto odstranění realizují jednoduchými konečnými automaty. Vzhledem k tomu, že soubor ještě obsahuje nejen hlavičky funkcí, ale také může obsahovat deklaraci či definici nějaké proměnné například `int pokus = 5`, proto je nutné detekovat hlavičky funkcí a následně je zkopírovat do seznamu, který již bude obsahovat pouze hlavičky funkcí. Tuto detekci provádím pomocí regulárního výrazu, protože syntaxe hlavičky funkce v jazyce C je pevně dána. Z každé deklarace funkce je zapotřebí zjistit její název, návratový typ, typ parametrů a případně zda se jedná o funkci s proměnným počtem parametrů. Název, návratový typ a typ parametrů zjišťuji jednoduchými regulárními výrazy. Na zjišťování toho zda funkce má či nemá proměnný počet parametrů, slouží konečný automat, který spočítá čárky v parametrech funkce a hledá tři tečky, které značí právě proměnný počet parametrů a v případě že je nenalezne tak se k počtu parametrů přičte jednička a u funkce se poznačí, že nemá proměnný počet parametrů.

Zpracování vybraných parametrů příkazové řádky

V případě zadání parametru `--no-inline` je regulárním výrazem kontrolována zda aktuální funkce tento specifikátor neobsahuje a pokud ano tak je její zpracovávání zastaveno.

Pro správnou funkci parametru `--max-par` využívám již spočtený počet parametrů a tuto hodnotu porovnávám s hodnotu zadanou u parametru a nevyhovující funkce nevypisují.

V případě, že je zadán parametr `--no-duplicates`, je potřeba také kontrolovat to zda se některé funkce v našem seznamu nevyskytují vícekrát a případně ponechat pouze první výskyt těchto funkcí a ostatní zahodit, aby se ve finálním výpisu neobjevovali vícekrát. Toto kontroluji pomocí seznamu, do kterého postupně ukládám názvy funkcí jednoho souboru a kontroluji, zda se již název funkce v seznamu nenachází. Pokud se v něm nachází, tak výpis funkce přeskočím, jinak se uloží do seznamu. Po výpisu funkcí z jednoho souboru se tento seznam vymaže, aby byl připraven na jména funkcí z dalšího souboru.

V případě zadání parametru `--remove-whitespace`, je nutné odstraňovat přebytečné mezery, k čemuž využívám dva regulární výrazy. Nejprve hledám výskyt dvou a více bílých znaků za sebou a nahrazuji je pouze jednou mezerou a poté vyhledávám znak hvězdičky následovaný mezerou a nahrazuji ji pouze znakem hvězdičky.

Formátování výstupu

Formát výstupu je zadáním pevně daný a liší se pouze tím zda je zadán parametr **--pretty-xml** a případně jaká je jeho hodnota. Kontrola zadání parametru se provádí před každým výpisem. Pokud parametr není zadán je úloha velmi jednoduchá, protože stačí výsledky pouze vypisovat za sebou bez odřádkování ve správném pořadí. Jestliže je parametr zadán pak je nutné výsledky odřádkovat a také správně odsadit, což je řešeno vložением znaku nového řádku a pomocným cyklem, jehož počet opakování je roven hodnotě výše uvedeného parametru a jeho tělo obsahuje výpis jedné mezery.