

Implementační dokumentace k 1. úloze do IPP 2020/2021

Jméno a příjmení: Vanessa Jóriová

Login: xjorio00

1. Analýza zdrojového kódu IPPcode21

Script na lexikálnu a syntaktickú analýzu jazyka IPPcode21 je implementovaný v súbore *parse.php*. Vstup program po riadkoch načítava zo štandardného vstupu a na výstup generuje XML reprezentáciu kódu. Najprv sa kontroluje korektnosť hlavičky kódu (akékoľvek biele znaky či nadbytočné odriadkovanie sa zanedbáva, rovnako ako aj veľkosť písmen samotných). Následne sa kontrolujú inštrukcie samotné. Rovnako ako v prípade hlavičky, aj tu sú zanedbané prípadné nadbytočné riadky či medzery medzi jednotlivými operandmi, rovnako ako sú ignorované aj komentáre.

Jednotlivé inštrukcie sú utriedené do skupín (casov switchu) podľa počtu a typu ich operandov. Počet operandov analyzovanej inštrukcie a korektnosť ich typu je skontrolovaná voči očakávanému počtu a typu operandov danej inštrukcie. Lexikálna korektnosť daného operandu je zisťovaná tým, či zodpovedá regulárnym výrazom opisujúcim očakávaný typ.

Zároveň s lexikálno-syntaktickou analýzou dochádza aj ku generácii xml stromu. Ten je po úspešnom ukončení lexikálno-syntaktickej analýzy vytlačený na štandardný výstup.

2. Spracovanie argumentov

Program implementuje aj rozšírenie STATP, ku ktorému sa viaže väčšina argumentov. Základné zadanie podporuje jediný argument *--help*, ktorý vypíše stručnú nápovedu k programu. Tento argument nemôže byť kombinovaný so žiadnym iným a negeneruje výstup.

Zvyšné argumenty podporujú rozšírenie STATP. Podľa zadaných argumentov program vypisuje štatistiky o zdrojovom kóde (napr. počet riadkov, počet komentárov, počet skokov...) do súboru zadaného argumentom *--stats=file*. Tento argument sa musí nachádzať pred argumentmi inicializujúcimi výpis danej štatistiky do súboru – jeho neprítomnosť vedie na chybu. Rovnako je chybné aj zadanie rovnakého súboru v novom argumente *--stats* (program kontroluje, či názov daného súboru nebol použitý už predtým), no zadanie iného súboru vedie na zapisovanie ďalších štatistík nasledujúcich za argumentom do spomínaného súboru.

3. Implementácia rozšírenia

V projekte je implementované rozšírenie STATP, ktoré zbiera štatistiky (počet riadkov, počet komentárov atď.). S týmto rozšírením sa spája aj nutnosť parsovania väčšieho množstva argumentov (popísané vyššie). Štatistiky sú vypisované do súborov zadaných argumentmi po riadkoch, pričom každý typ štatistiky je písaný na samostatný riadok a v prípade otvorenia neprázdného súboru je pôvodný obsah prepisovaný.

Za účelom rozšírenia v programe pribudlo niekoľko premenných monitorujúcich typ danej štatistiky. Pri správnom spracovaní riadku inštrukcie sa inkrementuje premenná zaznamenávajúca počet riadkov kódu, pri nájdení komentára sa zase tento fakt zaznamená do premennej počítajúcej komentáre. Nájdenie špecifickej inštrukcie LABEL zase zvýši počet návští pri zadaní parametru *--label*, *--jumps* zasa počítá inštrukcie pre podmienené a nepodmienené skoky a inštrukcie CALL a RETURN. Štatistiky o počte dopredných skokov (*--fwjumps*), skokov dozadu (*--backjumps*) a skokov na neexistujúce návestia (*--badjumps*) sa vyhodnocujú týmto spôsobom: ukladajú sa už nájdené názvy návští LABEL. Pokiaľ je dané návstie v čase skoku už nájdené, jedná sa o skok dozadu (*backjump*). Na konci behu programu sa každému návstiu, ktoré nebolo vyhodnotené ako skok dozadu, hľadá v poli zaznamenaných názvov návští príslušný cieľ skoku – pokiaľ sa nájde, jedná sa o dopredný skok (v čase analýzy skokovej inštrukcie nebol daný LABEL nájdený, musel byť teda súčasťou až nasledujúcich inštrukcií), pokiaľ sa nenájde, je skok vyhodnotený ako skok na neexistujúce návstie.