

## Implementačná dokumentácia k 1. úlohe do IPP 2019/2020

Meno a priezvisko: Tomáš Ďuriš

Login: xduris05

### Zadanie a popis skriptu

Úlohou bolo vytvoriť skript `parse.php` v jazyku PHP (konkrétne PHP7.4), ktorý načíta zdrojový kód v jazyku IPPcode2020 zo štandardného vstupu, skontroluje jeho lexikálnu a syntaktickú správnosť, a v prípade správneho vstupu následne vstup prevedie do formátu XML.

### Spôsob riešenia

Keďže som sa rozhodol riešiť aj rozšírenie, môj skript `parse.php` ako prvé načíta argumenty zo vstupu pomocou funkcie `getopt()`, skontroluje, či sa nachádzajú medzi povolenými argumentami a získa z nich potrebné informácie na uloženie. V prípade argumentu `--help` sa vypíše nápoveda a program sa ukončí. Následne program pokračuje načítaním vstupu zo štandardného vstupu. Ak prebehlo úspešne, pokračuje vygenerovaním hlavičky XML, rozdelením vstupu na riadky a odstránením prázdnych riadkov.

Následne prebehne kontrola po riadku, kde sa odstránia riadky s komentármi a uloží sa ich počet (prípadne časť riadku, ktorá obsahuje komentár). Odstránia sa aj redundatné biele znaky a riadky, ktoré obsahujú len tieto znaky. Skontroluje sa, či prvý riadok obsahuje správnu hlavičku a zvyšné riadky sa rozdelia na slová na základe použitého rozdelovača - medzera.

Skontroluje sa prvý parameter (operačný kód) a vráti sa očakávaný počet parametrov. Skontroluje sa, či sa zhoduje s počtom parametrov na riadku vstupu a ak áno, vygeneruje sa inštrukcia. Následne sa spracovávajú parametre. Spracovanie parametrov je riešené cez "switch", kde sú inštrukcie rozdelené do skupín so zhodnými očakávanými parametrami. Zistí sa, o akú inštrukciu ide, do ktorej skupiny patrí a prebehne kontrola lexikálnej správnosti parametra (či už sa jedná o konštantu, premennú, návěstie alebo typ). Ak ide o konštantu, a konkrétne o reťazec, tak prebehne kontrola správnosti escape sekvencií, ak sa nejaké v reťazci nachádzajú.

Pri kontrole lexikálnej správnosti parametrov som využíval regulárne výrazy, keďže podľa mňa išlo o najjednoduchšiu možnosť kontroly a mohol som využiť výhody, ktoré mi jazyk ponúka. V prípade splnenia všetkých formálnych požiadaviek na daný parameter následuje jeho zapísanie do XML a prechodu na ďalší parameter/inštrukciu. Na generáciu XML som použil `XMLWriter`, ktorý mi mimo iné uľahčil aj reprezentáciu problematických znakov ako (`<`, `>`, `&`) a umožnil jednoduché vytváranie XML elementov a pridávanie atribútov.

Ako posledné prebehne ukončenie generovania XML dokumentu a jeho vypísanie na štandardný výstup.

## Implementácia rozšírenia

V prípade zadania argumentov spadajúce pod rozšírenie **STATP** prebehne formálna kontrola a uloženie potrebných informácií, či už o výstupnom súbore alebo o očakávaných štatistikách na výstupe. Následne pri behu programu sa zaznamenávajú všetky štatistiky a po ukončení sa vypíšu tie, ktoré užívateľ zadal medzi argumenty pri vstupe, v takom poradí ako ich zadával a v takom množstve ako zadal vždy jedno číslo na jeden riadok.