

Implementační dokumentace k 1. úloze do IPP 2020/2021

Jméno a příjmení: Lukáš Baštýř

Login: xbasty00

Parser je implementován následovně:

Čte po řádcích standardní vstup. Pokud narazí na prázdnou linku, tak jí přeskočí. Narazí-li na komentáře, odstraní vše od # až do konce řádku. Následně se problémové znaky <, > a & nahradí za jejich *escape sekvence* pro XML. Nakonec se řádek rozdělí podle bílých znaků na jednotlivá slova (stringy).

Nyní se zkontroluje hlavička. Vezme se první (a jediné) slovo z řádku a porovná se (ve velkých písmenech) se slovem .IPPCODE21. Pokud jsou stejná, našli jsme hlavičku a může se pokračovat. V opačném případě je program ukončen s chybou 21.

Následně přichází veliký switch, který se používá na kontrolu operačního kódu. Zde jsou jednotlivé operační kódy rozděleny do skupin podle počtu a typu argumentů. Najdeme-li daný operační kód, je pak také potřeba zkontrolovat počet a typ operandů. Pokud operační kód nenajdeme, je program ukončen s chybou 22. Pokud dostane operační kód špatný počet operandů, program se ukončí s chybou 23

Operandy jsem rozdělil podle toho, jak se musí kontrolovat. Pro **proměnnou** se kontrolují pomocí regexu, kde musí být její rámec (GF@, LF@, TF@) následovaný povolenou kombinací znaků ze zadání. **Označení/label** se kontroluje stejně, akorát zde není rámec. **Typy** se kontrolují porovnáním s jejich názvy: **string**, **int** a **bool**.

Symboły (konsanty/proměnné) jsou zde asi nejzajímavější, protože stringová konstanta, může mít escape sekvence. Proměnné se kontrolují jak již bylo popsáno výše. Konstanty **int**, **bool** a **nil** se kontrolují pomocí regexu, kdy musí předcházet jejich typ (**string**, **int**, **bool**) + @ a povolená kombinace znaků. String je na tom ale dost jinak. Nejenom že musí předcházet jeho typ (**string@**), ale také musí podporovat escape sekvence (\ následováno 3mi čísly). To zde kontroluji tak, že porovnávám počet výskytů lomítek \ a lomítek za kterými následují aspoň 3 čísla \032. Pokud je jich stejně, jedná se o správný string s escape sekvencemi. Pokud je samotných lomítek více, chybí escape sekvence a string tedy není validní.

Pokud se narazí během kontroly operandů na nějaký problém, program se ukončí s chybou 23

Nakonec se daný operační kód vypíše na standardní výstup ve formátu XML. Toto se to opakuje pro každý řádek.

Všechna chybová hlášení se vypisují na standardní chybový výstup.