Implementačná Dokumentácia k 1. úlohe z IPP 2020/2021

Meno a Priezvisko : Jakub Drobena
Login : xdrobe01

Postup práce:

Vstupné dáta som načítal pomocou STDIN a následne som ich pomocou metódy „`explode`“ načítaný string rozdelil na dve časti a to programovú a komentáre. Keďže komentárová časť sa neskôr nepoužíva, som do premennej riadku „`line`“ priradil len hodnotu operačného kódu. Ako ďalšie som pomocou metódy „`str_replace`“ v danom stringu nahradil problematické znaky. Načítaný riadok som ešte rozdelil na jednotlivé tokeny pomocou metódy „`explode`“ do array. Jednotlivé tokeny som následne predával mnou definovanej metóde „`addElement`“ ktorá sa stará o prevod jazyka IPP2022 na formát XML. Pre pracovanie s XML formátom som použil triedu „`SimpleXMLElement`“, ktorou som do premennej xml inicializoval xml strom. Finálny spracovaný XML strom som nakoniec pomocou „`dom`“ -Document Object Model sformátoval a vypísal na štandardný výstup.

Metóda addElement:

Metóda pomocou switchu zistí z prvej pozície v array „words“ (kde sa nachádzajú jednotlivé tokeny) typ inštrukcie ktorý sa nachádza na prvej pozicii a následne sa vykoná syntaktická a lexikálna analýza. Ako prvé sa pomocou metódy „count“ zistí v array „words“ počet argumentov ktoré daná inštrukcia vyžaduje. Následne pomocou metód : „`isString(), isSymbol(), isBool(), isInt() a isLabel()`“ je prevádzaná lexikálna analýza kde sa zisťuje, či dané tokeny spĺňajú požiadavky ich typu a či vykonávajú syntaktickú analýzu argumentov. Ako posledné, keď sú všetky podmienky splnené, sa pridávajú jednotlivé elementy do XML stromu. Môže sa stať, že je potrebné oddeliť typ a hodnotu v tokenoch a to vo všetkých prípadoch okrem premennej a label. Typ a obsah sa získajú pomocou metódy `getValue` a `getType` ktorá vracajú hodnotu a typ oddelene. Následne už len vytvorím v XML strome jednotlivé elementy pomocou získaných hodnôt.

Výber niektorých regexov na určovanie typu a správnosti:

`\A(LF|GF|TF)@[a-zA-Z0-9_\-$;&%*!?]*\z/i` <- Regex pre premenné

`\A(int)@(\+|\-|)([0-9]+|(0x[0-9A-Fa-f]+)|(0X[0-9A-Fa-f]+)|`

`(0o[0-7]+)|(0O[0-7]+))\z` <- regex pre typ int vo formatoch hex,dec,octa(rozdelené do 2 riadkov)