

1 Zpracování vstupního souboru

Zdrojový kód je načítán ze standardního vstupu a následně jsou prováděny úpravy, jako je odstranění prázdných řádků a řádků obsahujících komentáře. Následuje kontrola vstupní hlavičky a syntaktická kontrola všech následujících operačních kódů a instrukcí. Nakonec je vytvořena XML reprezentace pomocí knihovny **xml.dom.minidom**.

Kontrola operačních kódů probíhá pomocí předem vytvořených seznamů obsahujících operační kódy. Tyto seznamy definují operace, které mají určitý počet argumentů. Tato kontrola je klíčová pro zajištění správnosti a bezpečnosti prováděných operací.

Dále následuje kontrola samotných argumentů. Tato kontrola je prováděna pomocí tří funkcí: **checkArgumentOne**, **checkArgumentTwo** a **checkArgumentThree**. Tyto funkce jsou volány před vytvořením XML reprezentace a slouží k ověření správnosti argumentů jednotlivých instrukcí.

Tvorba XML reprezentace se provádí ve funkcích **createXML_forN**, kde N označuje počet argumentů u operačních kódů. Kromě počtu argumentů některé z těchto funkcí mají možnost rozlišit konkrétní typy argumentů, jako jsou **symb**, **var** a **label**. Po každé kontrole se provádí vytvoření samotné XML reprezentace pro danou instrukci.

2 Implementace rozšíření

2.1 STATP

Toto rozšíření umožňuje skriptu sbírat statistiku o vstupním souboru. V mé implementaci sbírání statistik probíhá v souboru **stats.py**. Vytvoření statistik probíhá pomocí příkazů zadávaných do příkazové řádky. Zde je seznam všech možných statistik: **--stats=file** – Tiskne statistiky do zadaného souboru, **--print=string** – Tiskne řetězec do souboru, **--eol** – tiskne zalomení řádku do souboru, **--comments** – tiskne počet řádků, na kterých byl komentář, do souboru, **--jumps** – tiskne počet skoků do souboru, **--labels** – tiskne počet značek do souboru, **--fwjumps** – tiskne počet přímých skoků do souboru, **--backjumps** – tiskne počet zpětných skoků do souboru, **--badjumps** – tiskne počet neplatných skoků do souboru, **--frequent** – tiskne seznam názvů kódů operací, které jsou nejčastější podle počtu statických výskytů do souboru, **--loc** – tiskne počet řádek s instrukcemi do souboru.

3 Zpracování parametrů a spuštění skriptu

Pro spuštění skriptu stačí zadat příkaz **python3 parse.py < "vstupní soubor"**. Pro výpis nápovědy se používá parametr **--help**, díky kterému skript vypíše krátkou nápovědu a ukončí program. Pro výpis statistiky se používá parametr **--stats=filename** a další parametry pro statistiku. Pozor, parametry **--help** a **--stats=filename** nesmí být spolu v jednom příkazu.