

Zadání projektu z předmětu IPP 2020/2021

Zbyněk Křivka

E-mail: krivka@fit.vut.cz

1 Základní charakteristika projektu

Navrhněte, implementujte, dokumentujte a testujte sadu skriptů pro interpretaci nestrukturovaného imperativního jazyka IPPcode21. K implementaci vytvořte odpovídající stručnou programovou dokumentaci. Projekt se skládá ze dvou úloh a je individuální.

První úloha se skládá ze skriptu `parse.php` v jazyce PHP 7.4 (viz sekce 3) a dokumentace k tomuto skriptu (viz sekce 2.1). Druhá úloha se skládá ze skriptu `interpret.py` v jazyce Python 3.8 (viz sekce 4), testovacího skriptu `test.php` v jazyce PHP 7.4 (viz sekce 5) a dokumentace těchto dvou skriptů (viz sekce 2.1).

2 Požadavky a organizační informace

Kromě implementace skriptů a vytvoření dokumentace je třeba dodržet také řadu následujících formálních požadavků. Pokud některý nebude dodržen, může být projekt hodnocen nula body! Pro kontrolu alespoň některých formálních požadavků lze využít skript `is_it_ok.sh` dostupný v *Souborech* předmětu IPP.

Termíny:

připomínky¹ k zadání projektu do 14. února 2021;

fixace zadání projektu od 15. února 2021;

odevzdání první úlohy ve středu **17. března** 2021 do 23:59:59;

odevzdání druhé úlohy v úterý **20. dubna** 2021 do 23:59:59.

Dodatečné informace a konzultace k projektu z IPP:

- *Wiki* předmětu IPP v IS FIT včetně Často kladených otázek (*FAQ*).
- *Fórum* předmětu IPP pro ak. rok 2020/2021, témata IPP: **Projekt.***.
- Zbyněk Křivka (garant projektu): dle konzultačních slotů (viz webová stránka) nebo po dohodě e-mailem (uvádějte předmět začínající "IPP:"), viz <http://www.fit.vut.cz/person/krivka>. Další cvičící najdete na kartě předmětu.
- Dušan Kolář (garant předmětu; jen v závažných případech): po dohodě e-mailem (uvádějte předmět začínající "IPP:"), viz <http://www.fit.vut.cz/person/kolar>.

¹Naleznete-li v zadání nějakou chybu či nejasnost, dejte prosím vědět na *Fóru* předmětu nebo e-mailem na krivka@fit.vut.cz. Validní připomínky a upozornění na chyby budou oceněny i bonusovými body.

Pokud máte jakékoliv dotazy, problémy či nejasnosti ohledně tohoto projektu, tak po přečtení *FAQ* a *Fóra* předmětu (využívejte i možnosti hledání na *Fóru*), neváhejte napsat na *Fórum* (u obecného problému, jenž se potenciálně týká i vašich kolegů) či kontaktovat garanta projektu, cvičícího (v případě individuálního problému), nebo nouzově garanta předmětu, kdy do předmětu vždy uveďte na začátek řetězec "IPP:". Na problémy zjištěné v řádu hodin až jednotek dní před termínem odevzdání některé části projektu nebude brán zřetel. Začněte proto projekt řešit s dostatečným předstihem.

Forma a způsob odevzdání: Každá úloha se odevzdává individuálně prostřednictvím IS FIT v předmětu IPP (odevzdání e-mailem není možné nebo je bodově postihováno) a odevzdáním stvrzujete výhradní autorství skriptů i dokumentace.

Do termínu „Projekt - 1. úloha v PHP 7.4“ odevzdáte archiv pro první úlohu v jazyce PHP 7.4 (skript `parse.php`; včetně dokumentace tohoto skriptu). Po termínu odevzdání 1. úlohy bude otevřen termín „Projekt - 2. úloha v Pythonu 3.8 a testovací skript v PHP 7.4“ pro odevzdání archivu pro druhou úlohu (skripty `interpret.py` a `test.php`; včetně dokumentace pro obsažené skripty). Součástí archivu první úlohy mohou být i nedokončené skripty a dokumentace odevzdávané k hodnocení až v druhé úloze a naopak v archivu druhé úlohy se smí vyskytovat skript z první úlohy a adresář s vašimi testy. V případě odevzdání již plně funkčního skriptu `test.php` v rámci archivu 1. úlohy lze získat 1 bonusový bod.

Každá úloha bude odevzdána ve zvláštním archivu, kde budou soubory k dané úloze zkomprimovány programem ZIP, TAR+GZIP či TAR+BZIP do jediného archivu pojmenovaného `xlogin99.zip`, `xlogin99.tgz`, nebo `xlogin99.tbz`, kde `xlogin99` je váš login. Velikost každého archivu bude omezena informačním systémem (pravděpodobně na 1 MB). Archiv nesmí obsahovat speciální či binární spustitelné soubory. Názvy všech souborů mohou obsahovat pouze písmena anglické abecedy, číslice, tečku, pomlčku a podtržítka. **Skripty budou umístěny v kořenovém adresáři odevzdaného archivu.** Po rozbalení archivu na serveru *Merlin* bude možné skript(y) spustit. Archiv smí obsahovat v rozumné míře pomocné adresáře (typicky pro vaše vlastní testy, vaše knihovny a pomocné skripty nebo povolené knihovny, které nejsou nainstalovány na serveru *Merlin*).

Hodnocení: Výše základního bodového hodnocení projektu v předmětu IPP je **maximálně 20 bodů**. Navíc lze získat maximálně 5 bonusových bodů za kvalitní a podařené řešení některého z rozšíření nebo kvalitativně nadprůměrnou účast na *Fóru* projektu apod.

Hodnocení jednotlivých skriptů: `parse.php` až 6 bodů; `interpret.py` až 8 bodů; `test.php` až 3 body. Tj. v součtu až **17 bodů**. Dokumentace a úprava zdrojových textů skriptů bude dohromady hodnocena až **3 body** (až 1 bod za popis skriptu `parse.php` a až 2 body za popis skriptů `interpret.py` a `test.php`), avšak maximálně 30 % ze sumy hodnocení skriptů dané úlohy (tedy v případě neodevzdání žádného funkčního skriptu v dané úloze bude samotná dokumentace hodnocena 0 body). Body za každou úlohu včetně její dokumentace se před vložením do IS FIT zaokrouhlují na desetiny bodu.

Skripty budou spouštěny na serveru *Merlin* příkazem: `interpret skript parametry`, kde `interpret` bude `php7.4` nebo `python3.8`, `skript` a `parametry` závisí na dané úloze a skriptu. Hodnocení většiny funkčnosti bude zajišťovat automatizovaný nástroj. Kvalitu dokumentace, komentářů a strukturu zdrojového kódu budou hodnotit cvičící.

Vaše skripty budou hodnoceny nezávisle na sobě, takže je možné odevzdat například jen skript `interpret.py` nebo jen `test.php` do 2. úlohy bez odevzdání 1. úlohy.

Podmínky pro opakující studenty ohledně případného uznání hodnocení loňského projektu najdete v *FAQ* na *Wiki* předmětu.

Registrovaná rozšíření: V případě implementace některých registrovaných rozšíření za bonusové body bude odevzdaný archiv obsahovat soubor **rozsireni**, ve kterém uvedete na každém řádku identifikátor jednoho implementovaného rozšíření² (řádky jsou ukončeny unixovým koncem řádku, tj. znak s dekadickou ASCII hodnotou 10). V průběhu řešení mohou být zaregistrována nová rozšíření úlohy za bonusové body (viz *Fórum* předmětu IPP). Nejpozději do termínu pokusného odevzdání dané úlohy můžete na *Fórum* zasílat návrhy na nová netriviální rozšíření, která byste chtěli navíc implementovat. Cvičící rozhodne o přijetí/nepřijetí rozšíření a hodnocení rozšíření dle jeho náročnosti včetně přiřazení unikátního identifikátoru. Implementovaná rozšíření neidentifikovaná v souboru **rozsireni** nebudou hodnocena.

Pokusné odevzdání: Pro zvýšení motivace studentů pro včasné vypracování úloh nabízíme koncept nepovinného pokusného odevzdání. Výměnou za pokusné odevzdání do uvedeného termínu (cca týden před finálním termínem) dostanete zpětnou vazbu v podobě zařazení do některého z pěti rozmezí hodnocení (0–10 %, 11–30 %, 31–50 %, 51–80 %, 81–100 %). Bude-li vaše pokusné odevzdání v prvním rozmezí hodnocení, máte možnost osobně konzultovat důvod, pokud jej neodhalíte sami. U ostatních rozmezí nebudou detailnější informace poskytovány.

Pokusné odevzdání bude relativně rychle vyhodnoceno automatickými testy a studentům zaslána orientační informace o správnosti pokusně odevzdané úlohy z hlediska části automatických testů (tj. nebude se jednat o finální hodnocení; proto nebudou sdělovány ani body či přesnější procentuální hodnocení). Využití pokusného termínu není povinné, ale jeho nevyužití může být negativně vzato v úvahu v případě reklamace hodnocení projektu.

Formální požadavky na pokusné odevzdání jsou totožné s požadavky na finální termín a odevzdání se bude provádět do speciálních termínů „Projekt - Pokusné odevzdání 1. úlohy“ do **10. března 2021** a „Projekt - Pokusné odevzdání 2. úlohy“ do **13. dubna 2021**. Není nutné zahrnout dokumentaci, která spolu s rozšířeními pokusně vyhodnocena nebude. U skriptu **test.php** bude pokusné hodnocení omezeno jen na velmi základní automatizovatelné testy (bez manuální kontroly skutečného obsahu).

2.1 Dokumentace

Implementační dokumentace (dále jen dokumentace) musí být stručným a uceleným průvodcem **vašeho způsobu řešení** skriptů 1., resp. 2. úlohy. Bude vytvořena ve formátu **PDF** nebo **Markdown** (viz [4]). Jakékoliv jiné formáty dokumentace než PDF či Markdown³ (přípona **md**) budou ignorovány, což povede ke ztrátě bodů za dokumentaci. Dokumentaci je možné psát buď česky, slovensky (s diakritikou, formálně čistě), nebo anglicky (formálně čistě).

Dokumentace bude popisovat celkovou filozofii návrhu, interní reprezentaci, způsob a váš specifický postup řešení (např. řešení sporných případů nedostatečně upřesněných zadáním, konkrétní řešení rozšíření, případné využití návrhových vzorů, implementované/nedokončené vlastnosti). Dokumentace může být doplněna např. o UML diagram tříd, navržený konečný automat, pravidla vámi vytvořené gramatiky nebo popis jiných formalismů a algoritmů. Nicméně **nesmí obsahovat ani částečnou kopii zadání**.

Rozsah textu dokumentace pro každý skript bude přibližně 1 normostrana. Vysázená dokumentace 1. úlohy popisující skript **parse.php** by neměla přesáhnout 1 stranu A4 a u 2. úlohy (popisující skripty **interpret.py** a **test.php**) pak 2 strany A4. Doporučení pro sazbu: 10bodové písmo Times New Roman pro text a Courier pro identifikátory a skutečně krátké úryvky zajímavého kódu (či zpětné apostrofy v Markdown); nekládejte žádnou zvláštní úvodní stranu, obsah ani závěr. V ro-

²Identifikátory rozšíření jsou uvedeny u konkrétního rozšíření tučně.

³Bude-li přítomna dokumentace v Markdown i PDF, bude hodnocena verze v PDF, kde je jistější sazba.

zumné míře je vhodné používat nadpisy první a druhé úrovně (12bodové a 11bodové písmo Times New Roman či `##` a `###` v Markdown) pro vytvoření logické struktury dokumentace.

Nadpis a hlavička dokumentace⁴ bude na prvních třech řádcích obsahovat:

Implementační dokumentace k `%cislo%`. úloze do IPP 2020/2021

Jméno a příjmení: `%jmeno_prijmeni%`

Login: `%xlogin99%`

kde `%jmeno_prijmeni%` je vaše jméno a příjmení, `%xlogin99%` váš login a `%cislo%` je číslo dokumentované úlohy.

Dokumentace bude v kořenovém adresáři odevzdaného archívu a pojmenována `readme1.pdf` nebo `readme1.md` pro první úlohu a `readme2.pdf` nebo `readme2.md` pro druhou úlohu.

V rámci dokumentace bude hodnoceno i komentování zdrojových kódů (minimálně každá funkce a modul (třída) by měly mít komentář o svém účelu a parametrech; u složitějších funkcí okomentujte i omezení na vstupy či výstupy).

2.2 Programová část

Zadání projektu vyžaduje implementaci tří skriptů⁵, které mají parametry příkazové řádky a je definováno, jakým způsobem manipulují se vstupy a výstupy. Skript (vyjma `test.php`) nesmí spouštět žádné další procesy či příkazy operačního systému. Veškerá chybová hlášení, varování a ladicí výpisy směřujte pouze na standardní chybový výstup, jinak pravděpodobně nedodržíte zadání kvůli modifikaci definovaných výstupů (ať již do externích souborů, nebo do standardního výstupu). Jestliže proběhne činnost skriptu bez chyb, vrací se návratová hodnota 0 (nula). Jestliže dojde k nějaké chybě, vrací se chybová návratová hodnota větší jak nula. Chyby mají závazné chybové návratové hodnoty:

- 10 - chybějící parametr skriptu (je-li třeba) nebo použití zakázané kombinace parametrů;
- 11 - chyba při otevírání vstupních souborů (např. neexistence, nedostatečné oprávnění);
- 12 - chyba při otevření výstupních souborů pro zápis (např. nedostatečné oprávnění, chyba při zápisu);
- 20 – 69 - návratové kódy chyb specifických pro jednotlivé skripty;
- 99 - interní chyba (neovlivněná vstupními soubory či parametry příkazové řádky; např. chyba alokace paměti).

Pokud zadání níže nestanoví jinak, veškeré vstupy a výstupy jsou v kódování UTF-8. Pro účely projektu z IPP musí být na serveru *Merlin* ponecháno implicitní nastavení `locale`⁶, tj. `LC_ALL=cs_CZ.UTF-8`.

Jména hlavních skriptů jsou dána zadáním. Pomocné skripty nebo knihovny budou mít příponu dle zvyklostí v daném programovacím jazyce (`.php` pro PHP 7 a `.py` pro Python 3). Vyhodnocení

⁴Anglické znění nadpisu a hlavičky dokumentace najdete v *FAQ* na *Wiki* předmětu.

⁵Tyto skripty jsou aplikace příkazové řádky neboli konzolové aplikace.

⁶Správné nastavení prostředí je nezbytné, aby bylo možné používat a správně zpracovávat parametry příkazové řádky v UTF-8. Pro správnou funkčnost je třeba mít na UTF-8 nastaveno i kódování znakové sady konzole (např. u programu PuTTY v kategorii *Window.Translation* nastavíte *Remote character set* na UTF-8). Pro změnu ovlivňující aktuální sezení lze využít unixový příkaz `export LC_ALL=cs_CZ.UTF-8`.

skriptů bude prováděno na serveru *Merlin* s aktuálními verzemi interpretů (dne 1. 2. 2021 bylo na tomto serveru nainstalováno **php7.4**⁷ verze 7.4.14 a **python3.8**⁸ verze 3.8.6).

K řešení lze využít standardně předinstalované knihovny obou jazykových prostředí na serveru *Merlin*. Případné využití jiné knihovny kromě knihovny podporující načítání/ukládání formátu XML, zpracování parametrů příkazové řádky a zpracování regulárních výrazů je třeba konzultovat s garantem projektu (především z důvodu, aby se řešení projektu použitím vhodné knihovny nestalo zcela triviálním). Seznam povolených a zakázaných knihoven bude udržován aktuální na *Wiki* předmětu. Ve skriptech v jazyce PHP jsou některé funkce z bezpečnostních důvodů zakázány (např. **header**, **mail**, **popen**, **curl_exec**, **socket_***; úplný seznam je u povolených/zakázaných knihoven na *Wiki*).

Každý skript bude pracovat s jedním společným parametrem:

- **--help** vypíše na standardní výstup nápovědu skriptu (nenačítá žádný vstup), kterou lze převzít ze zadání (lze odstranit diakritiku, případně přeložit do angličtiny dle zvoleného jazyka dokumentace), a vrací návratovou hodnotu 0. Tento parametr nelze kombinovat s žádným dalším parametrem, jinak skript ukončete s chybou 10.

Kombinovatelné parametry skriptů jsou odděleny alespoň jedním bílým znakem a mohou být uváděny v libovolném pořadí, pokud nebude řečeno jinak. U skriptů je možné implementovat i vaše vlastní nekolizní parametry (doporučujeme konzultaci na *Fóru* nebo u garanta projektu).

Není-li řečeno jinak, tak dle konvencí unixových systémů lze uvažovat zástupné zkrácené (s jednou pomlčkou) i dlouhé parametry (se dvěma pomlčkami), které lze se zachováním sémantiky zaměňovat (tzv. alias parametry), ale testovány budou vždy dlouhé verze.

Je-li součástí parametru i soubor (např. **--source=***file* nebo **--source=**"*file*") či cesta, může být tento soubor/cesta zadán/a relativní cestou⁹ nebo absolutní cestou; výskyt znaku uvozovek a rovnítko ve *file* neuvažujte. Cesta/jméno souboru mohou obsahovat i Unicode znaky v UTF-8.

3 Analyzátor kódu v IPPcode21 (**parse.php**)

Skript typu filtr (**parse.php** v jazyce PHP 7.4) načte ze standardního vstupu zdrojový kód v IPPcode21 (viz sekce 6), zkontroluje lexikální a syntaktickou správnost kódu a vypíše na standardní výstup XML reprezentaci programu dle specifikace v sekci 3.1.

Tento skript bude pracovat s těmito parametry:

- **--help** viz společný parametr všech skriptů v sekci 2.2.

Chybové návratové kódy specifické pro analyzátor:

- 21 - chybná nebo chybějící hlavička ve zdrojovém kódu zapsaném v IPPcode21;
- 22 - neznámý nebo chybný operační kód ve zdrojovém kódu zapsaném v IPPcode21;
- 23 - jiná lexikální nebo syntaktická chyba zdrojového kódu zapsaného v IPPcode21.

⁷Upozornění: Na serveru *Merlin* je třeba dodržet testování příkazem **php7.4**, protože pouhým **php** se spouští verze, která nemá přístup k souborovému systému!

⁸Upozornění: Na serveru *Merlin* je třeba dodržet testování příkazem **python3.8**, protože pouhým **python** se spouští stará nekompatibilní verze! Python 3.x není zpětně kompatibilní s verzí 2.x!

⁹Relativní cesta nebude obsahovat zástupný symbol ~ (vlňka).

3.1 Popis výstupního XML formátu

Za povinnou XML hlavičkou¹⁰ následuje kořenový element `program` (s povinným textovým atributem `language` s hodnotou `IPPCODE21`), který obsahuje pro instrukce elementy `instruction`. Každý element `instruction` obsahuje povinný atribut `order` s pořadím instrukce (počítáno od 1, souvislá posloupnost) a povinný atribut `opcode` (hodnota operačního kódu je ve výstupním XML vždy velkými písmeny) a elementy pro odpovídající počet operandů/argumentů: `arg1` pro případný první argument instrukce, `arg2` pro případný druhý argument a `arg3` pro případný třetí argument instrukce. Element pro argument má povinný atribut `type` s možnými hodnotami `int`, `bool`, `string`, `nil`, `label`, `type`, `var` podle toho, zda se jedná o literál, návěští, typ nebo proměnnou, a obsahuje textový element.

Tento textový element potom nese buď hodnotu literálu (již bez určení typu a bez znaku `@`), nebo jméno návěští, nebo typ, nebo identifikátor proměnné (včetně určení rámce a `@`). U proměnných vypisujte označení rámce vždy velkými písmeny, jak by mělo být již na vstupu. Velikosti písmen samotného jména proměnné ponechejte beze změny. Aby zde nebylo třeba definovat přesně výstupní formát čísel, vypisujte na výstup číselné literály přesně ve formátu ze zdrojového kódu (např. zůstanou kladná znaménka čísel nebo počáteční přebytkové nuly) a není třeba kontrolovat jejich lexikální správnost (na rozdíl od řetězových literálů). U literálů typu `string` při zápisu do XML nepřevádějte původní escape sekvence, ale pouze pro problematické znaky v XML (např. `<`, `>`, `&`) využijte odpovídající XML entity (např. `<`, `>`, `&`). Podobně převádějte problematické znaky vyskytující se v identifikátorech proměnných. Literály typu `bool` vždy zapisujte malými písmeny jako `false` nebo `true`.

Doporučení: Všimněte si, že analýza `IPPCODE21` je tzv. kontextově závislá (viz přednášky), kdy například můžete mít klíčové slovo použito jako návěští a z kontextu je třeba rozpoznat, zda jde o návěští, nebo ne. Při tvorbě analyzátoru doporučujeme kombinovat konečně-stavové řízení a regulární výrazy a pro generování výstupního XML využít vhodnou knihovnu.

Výstupní XML bude porovnáváno s referenčními výsledky pomocí nástroje `A7Soft JExamXML`¹¹, viz [2]. Pozor, v PHP příkaz `return` neslouží pro návrat chybového kódu, použijte funkci `exit`. Pro výpis varování na standardní chybový výstup uveďte na začátku skriptu příkaz

```
ini_set('display_errors', 'stderr');
```

3.2 Bonusová rozšíření

STATP Sbírání statistik zpracovaného zdrojového kódu v `IPPCODE21`. Skript bude podporovat parametr `--stats=file` pro zadání souboru *file*, kam se budou vypisovat statistiky v parametrech umístěných za tímto `--stats`. Statistiky se do souboru vypisují po řádcích dle pořadí v parametrech s možností jejich opakování; na každý řádek nevypisujte nic kromě požadovaného číselného výstupu a odřádkování; případně existující soubor je přepsán. Pro sběr skupin statistik do různých souborů se použije další výskyt parametru `--stats` s jiným jménem souboru a následovaný dalšími parametry nové skupiny statistik. Parametr `--loc` vypíše do statistik počet řádků s instrukcemi (nepočítají se prázdné řádky ani řádky obsahující pouze komentář ani úvodní řádek). Parametr `--comments` vypíše do statistik počet řádků, na kterých se vyskytoval komentář. Parametr `--labels` vypíše do statistik počet definovaných návěští (tj. unikátních možných cílů skoku). Parametr `--jumps` vypíše do statistik počet všech instrukcí

¹⁰Tradiční XML hlavička včetně verze a kódování je `<?xml version="1.0" encoding="UTF-8"?>`

¹¹Nastavení `A7Soft JExamXML` pro porovnávání XML (soubor `options`) je v *Souborech* k projektu v IS FIT.

návratů z volání a instrukcí pro *skoky* (souhrnně podmíněné/nepodmíněné skoky a volání), `--fwjumps` počet dopředných skoků, `--backjumps` zpětných skoků a `--badjumps` počet skoků na neexistující návěští. Je-li uveden pouze parametr `--stats` bez upřesnění statistik k výpisu, bude výstupem prázdný soubor. Chybí-li při zadání parametrů statistik `--loc`, `--comments` a dalších před nimi parametr `--stats`, jedná se o chybu 10. Pokus o zápis více skupin statistik do stejného souboru během jednoho spuštění skriptu vede na chybu 12. [1,5 b]

4 Interpret XML reprezentace kódu (`interpret.py`)

Skript (`interpret.py` v jazyce Python 3.8) načte XML reprezentaci programu a tento program s využitím vstupu dle parametrů příkazové řádky interpretuje a generuje výstup. Vstupní XML reprezentace je např. generována skriptem `parse.php` (ale ne nutně, takže může obsahovat chyby) ze zdrojového kódu v IPPcode21. Interpret navíc oproti sekci 3.1 podporuje existenci volitelných dokumentačních textových atributů `name` a `description` v kořenovém elementu `program`. Sémantika jednotlivých instrukcí IPPcode21 je popsána v sekci 6. Interpretace instrukcí probíhá dle atributu `order` vzestupně (sekvence nemusí být souvislá na rozdíl od sekce 3.1).

Tento skript bude pracovat s těmito parametry:

- `--help` viz společný parametr všech skriptů v sekci 2.2;
- `--source=file` vstupní soubor s XML reprezentací zdrojového kódu dle definice ze sekce 3.1;
- `--input=file` soubor se vstupy¹² pro samotnou interpretaci zadaného zdrojového kódu.

Alespoň jeden z parametrů (`--source` nebo `--input`) musí být vždy zadán. Pokud jeden z nich chybí, jsou chybějící data načítána ze standardního vstupu.

Chybové návratové kódy specifické pro interpret:

- 31 - chybný XML formát ve vstupním souboru (soubor není tzv. dobře formátovaný, angl. *well-formed*, viz [1]);
- 32 - neočekávaná struktura XML (např. element pro argument mimo element pro instrukci, instrukce s duplicitním pořadím nebo záporným pořadím) či lexikální nebo syntaktická chyba textových elementů a atributů ve vstupním XML souboru (např. chybný lexém pro číselný nebo řetězcový literál, neznámý operační kód apod.).

Chybové návratové kódy interpretu v případě chyby během interpretace jsou uvedeny v popisu jazyka IPPcode21 (viz sekce 6.1).

Doporučení: Doporučujeme použít knihovnu pro načítání XML. V případě nekompletní implementace se zaměřte na funkčnost globálního rámce, práce s proměnnými typu `int`, instrukce `WRITE` a instrukce pro řízení toku programu.

¹²Vstup/vstupní soubor může být prázdný; např. neinterpretuje-li se žádná instrukce `READ`.

4.1 Bonusová rozšíření

FLOAT Podpora typu `float` v IPPcode21 (**hexadecimální zápis** v analyzátoru i v interpretu včetně načítání ze standardního vstupu; např. `float@0x1.2000000000000p+0` reprezentuje 1,125; viz funkce `float.fromhex()` a `float.hex()` v Python 3). Podporujte instrukce pro práci s tímto typem: `INT2FLOAT`, `FLOAT2INT`, aritmetické instrukce (včetně `DIV`), atd. (viz [3]). Podpora v `parse.php` je možná, ale nebude testována. [1 b]

STACK Podpora zásobníkových variant instrukcí (přípona `S`; viz [3]): `CLEARs`, `ADDs/SUBs`, `MULs/IDIVs`, `LTS/GTS/EQS`, `ANDs/ORS/NOTs`, `INT2CHARs/STR2INTs` a `JUMPIFEQs/JUMPIFNEQs`. Zásobníkové verze instrukcí z datového zásobníku vybírají operandy se vstupními hodnotami dle popisu tříadresné instrukce od konce (tj. typicky nejprve $\langle symb_2 \rangle$ a poté $\langle symb_1 \rangle$). Podpora v `parse.php` je možná, ale nebude testována. [1 b]

STATI Sbírání statistik interpretace kódu. Skript bude podporovat parametr `--stats=file` pro zadání souboru *file*, kam se agregované statistiky budou vypisovat (po řádcích dle pořadí v dalších (případně opakovaných) parametrech; na každý řádek nevypisujte nic kromě požadovaného číselného výstupu a odřádkování). Podpora parametru `--insts` pro výpis počtu tzv. vykonaných instrukcí (tj. kromě ladicích instrukcí a speciální instrukce `LABEL`) během interpretace do statistik. Parametr `--hot` vypíše do statistik hodnotu atributu `order` u vykonané instrukce, která byla provedena nejvícekrát a má hodnotu atributu `order` nejmenší. Podpora parametru `--vars` pro výpis maximálního počtu inicializovaných proměnných přítomných v jeden okamžik ve všech platných rámcích během interpretace zadaného programu do statistik. Chybí-li při zadání `--insts`, `--hot` či `--vars` parametr `--stats`, jedná se o chybu 10. [1 b]

5 Testovací rámec (`test.php`)

Skript (`test.php` v jazyce PHP 7.4) bude sloužit pro automatické testování (postupné) aplikace `parse.php` a `interpret.py`¹³. Skript projde zadaný adresář s testy a využije je pro automatické otestování správné funkčnosti jednoho či obou předchozích skriptů včetně vygenerování přehledného souhrnu v HTML 5 na standardní výstup. Pro hodnocení `test.php` budou dodány referenční implementace `parse.php` i `interpret.py`. Testovací skript nemusí u `parse.php` ani u `interpret.py` testovat jejich dodatečnou funkčnost aktivovanou parametry příkazové řádky (s výjimkou potřeby parametru `--source` a/nebo `--input` u `interpret.py`).

Tento skript bude pracovat s těmito parametry:

- `--help` viz společný parametr všech skriptů v sekci 2.2;
- `--directory=path` testy bude hledat v zadaném adresáři (chybí-li tento parametr, tak skript prochází aktuální adresář);
- `--recursive` testy bude hledat nejen v zadaném adresáři, ale i rekurzivně ve všech jeho podadresářích;
- `--parse-script=file` soubor se skriptem v PHP 7.4 pro analýzu zdrojového kódu v IPPcode21 (chybí-li tento parametr, tak implicitní hodnotou je `parse.php` uložený v aktuálním adresáři);

¹³Za tímto účelem lze vytvářet dočasné soubory, které však nesmí přepsat žádný jiný existující soubor a potom musí být uklizeny. Případně můžete tento úklid potlačit podporou parametru `--debug` za účelem ladění rámce.

- `--int-script=file` soubor se skriptem v Python 3.8 pro interpret XML reprezentace kódu v IPPcode21 (chybí-li tento parametr, tak implicitní hodnotou je `interpret.py` uložený v aktuálním adresáři);
- `--parse-only` bude testován pouze skript pro analýzu zdrojového kódu v IPPcode21 (tento parametr se nesmí kombinovat s parametry `--int-only` a `--int-script`), výstup s referenčním výstupem (soubor s příponou `out`) porovnávejte nástrojem A7Soft JExamXML (viz [2]);
- `--int-only` bude testován pouze skript pro interpret XML reprezentace kódu v IPPcode21 (tento parametr se nesmí kombinovat s parametry `--parse-only` a `--parse-script`). Vstupní program reprezentován pomocí XML bude v souboru s příponou `src`.
- `--jexamxml=file` soubor s JAR balíčkem s nástrojem A7Soft JExamXML. Je-li parametr vynechán uvažuje se implicitní umístění `/pub/courses/ipp/jexamxml/jexamxml.jar` na serveru *Merlin*, kde bude `test.php` hodnocen.
- `--jexamcfg=file` soubor s konfigurací nástroje A7Soft JExamXML. Je-li parametr vynechán uvažuje se implicitní umístění `/pub/courses/ipp/jexamxml/options` na serveru *Merlin*, kde bude `test.php` hodnocen.

Chybový návratový kód specifický pro testovací rámec:

- 41 - zadaný adresář (*path* v parametru `--directory`) nebo zadaný soubor (*file* v parametrech `--parse-script`, `--int-script`, `--jexamxml` a `--jexamcfg`) neexistuje či není přístupný.

Každý test je tvořen až 4 soubory stejného jména s příponami `src`, `in`, `out` a `rc` (ve stejném adresáři). Hlavní soubor testu s příponou `src` obsahuje zdrojový kód v jazyce IPPcode21 (příp. jeho XML reprezentaci). Soubory s příponami `in`, `out` a `rc` obsahují vstup a očekávaný/referenční výstup a očekávaný první chybový návratový kód analýzy resp. interpretace nebo bezchybový návratový kód 0. Pokud soubor s příponou `in` nebo `out` chybí, automaticky se fyzicky dogeneruje prázdný soubor. V případě chybějícího souboru s příponou `rc` se vygeneruje soubor obsahující návratovou hodnotu 0. Při nenulovém návratovém kódu je provedení testu považováno za úspěšné, pokud je získán správný návratový kód v odpovídajícím skriptu¹⁴ (výstupy se neporovnávají). V případě nulového kódu je třeba provést i porovnání výstupu skriptu s referenčním výstupem. Při parametru `--parse-only` jsou výstupy porovnávány pomocí nástroje A7Soft JExamXML¹⁵, jinak je použit unixový nástroj `diff`.

Testy budou umístěny v adresáři včetně případných podadresářů pro lepší kategorizaci testů. Adresářová struktura může být libovolná (včetně zanoření). Není třeba uvažovat symbolické odkazy a skryté adresáře.

Požadavky na výstupní HTML verze 5: Přehledová stránka o úspěšnosti/neúspěšnosti jednotlivých testů a celých adresářů bude prohlédnuta ručně opravujícím, takže bude hodnocena její přehlednost a intuitivnost. Mělo by být na první pohled zřejmé, které testy uspěly a které nikoli (a kolik), a zda případně uspěly všechny testy (případně i po jednotlivých adresářích), takže je třeba zahrnout do výpisu všechny prováděné testy, ideálně v každém adresáři řazené abecedně. Výsledná stránka nesmí načítat externí zdroje¹⁶ a musí být možné ji zobrazit v běžném prohlížeči.

¹⁴Při postupné aplikaci obou filtrů závisí na návratové hodnotě `parse.php`, zda se následně spustí i `interpret.py`.

¹⁵Použití nástroje A7Soft JExamXML je popsáno na Wiki předmětu v článku *ProjectNotes*.

¹⁶Přímo do generované HTML stránky je možné vložit vlastní JavaScript nebo CSS kód.

5.1 Bonusová rozšíření

FILES Podporujte parametr `--testlist=file` sloužící pro explicitní zadání seznamu adresářů (zadaných relativními či absolutními cestami) a případně i souborů s testy (zadáva se soubor s příponou `.src`) formou externího souboru *file* místo načtení testů z aktuálního adresáře (nelze kombinovat s parametrem `--directory`). Dále podporujte parametr `--match=regex` pro výběr testů, jejichž jméno bez přípony (ne cesta) odpovídá zadanému regulárnímu výrazu *regex* dle PCRE syntaxe (tj. včetně oddělovačů; syntaktická chyba výrazu vede na chybu 11). [1 b]

V případě odevzdání již plně funkčního skriptu `test.php` v rámci archívu 1. úlohy lze získat 1 bonusový bod.

6 Popis jazyka IPPcode21

Nestrukturovaný imperativní jazyk IPPcode21 vznikl úpravou jazyka IFJcode20 (jazyk pro mezikód překladače jazyka IFJ20, viz [3]), který zahrnuje instrukce tříadresné (typicky se třemi argumenty) a případně zásobníkové (typicky méně parametrů a pracující s hodnotami na datovém zásobníku). Každá instrukce se skládá z operačního kódu (klíčové slovo s názvem instrukce), u kterého nezáleží na velikosti písmen (tj. case insensitive). Zbytek instrukcí tvoří operandy, u kterých na velikosti písmen záleží (tzv. case sensitive). Operandy oddělujeme libovolným nenulovým počtem mezer či tabulátorů. Také před operačním kódem a za posledním operandem se může vyskytnout libovolný počet mezer či tabulátorů. Odřádkování slouží pro oddělení jednotlivých instrukcí, takže na každém řádku je maximálně jedna instrukce a není povoleno jednu instrukci zapisovat na více řádků. Každý operand je tvořen proměnnou, konstantou, typem nebo návěštím. V IPPcode21 jsou podporovány jednořádkové komentáře začínající mřížkou (`#`). Vyjma prázdných¹⁷ či zakomentovaných řádků obsahuje kód v jazyce IPPcode21 na začátku místo instrukce jen identifikátor jazyka (tečka následovaná jménem jazyka, kdy nezáleží na velikosti písmen):

.IPPcode21

6.1 Návrátové hodnoty interpretu

Proběhne-li interpretace bez chyb, vrací se návratová hodnota 0 (nula). Chybovým případům odpovídají následující návratové hodnoty:

- 52 - chyba při sémantických kontrolách vstupního kódu v IPPcode21 (např. použití nedefinovaného návěští, redefinice proměnné);
- 53 - běhová chyba interpretace – špatné typy operandů;
- 54 - běhová chyba interpretace – přístup k neexistující proměnné (rámec existuje);
- 55 - běhová chyba interpretace – rámec neexistuje (např. čtení z prázdného zásobníku rámců);
- 56 - běhová chyba interpretace – chybějící hodnota (v proměnné, na datovém zásobníku nebo v zásobníku volání);
- 57 - běhová chyba interpretace – špatná hodnota operandu (např. dělení nulou, špatná návratová hodnota instrukce EXIT);
- 58 - běhová chyba interpretace – chybná práce s řetězcem.

¹⁷Řádek obsahující pouze bílé znaky je považován za prázdný.

6.2 Paměťový model

Hodnoty během interpretace nejčastěji ukládáme do pojmenovaných proměnných, které jsou sdružovány do tzv. rámců, což jsou v podstatě slovníky proměnných s jejich hodnotami. IPPcode21 nabízí tři druhy rámců:

- globální, značíme GF (Global Frame), který je na začátku interpretace automaticky inicializován jako prázdný; slouží pro ukládání globálních proměnných;
- lokální, značíme LF (Local Frame), který je na začátku nedefinován a odkazuje na vrcholový/aktuální rámec na zásobníku rámců; slouží pro ukládání lokálních proměnných funkcí (zásobník rámců lze s výhodou využít při zanořeném či rekurzivním volání funkcí);
- dočasný, značíme TF (Temporary Frame), který slouží pro chystání nového nebo úklid starého rámce (např. při volání nebo dokončování funkce), jenž může být přesunut na zásobník rámců a stát se aktuálním lokálním rámcem. Na začátku interpretace je dočasný rámec nedefinovaný.

K překrytým (dříve vloženým) lokálním rámcům v zásobníku rámců nelze přistoupit dříve, než vyjmeme později přidané rámce.

Další možností pro ukládání nepojmenovaných hodnot je datový zásobník využívaný zásobníkovými instrukcemi.

6.3 Datové typy

IPPcode21 pracuje s typy operandů dynamicky, takže je typ proměnné (resp. paměťového místa) dán obsaženou hodnotou. Není-li řečeno jinak, jsou implicitní konverze zakázány. Interpret podporuje speciální hodnotu/typ `nil` a tři základní datové typy (`int`, `bool` a `string`), jejichž rozsahy i přesnosti jsou kompatibilní s jazykem Python 3.

Zápis každé konstanty v IPPcode21 se skládá ze dvou částí oddělených zavináčem (znak `@`; bez bílých znaků), označení typu konstanty (`int`, `bool`, `string`, `nil`) a samotné konstanty (číslo, literál, `nil`). Např. `bool@true`, `nil@nil` nebo `int@-5`.

Typ `int` reprezentuje celé číslo (přetečení/podtečení neřešte). Typ `bool` reprezentuje pravdivostní hodnotu (`false` nebo `true`). Literál pro typ `string` je v případě konstanty zapsán jako sekvence tisknutelných znaků v kódování UTF-8 (vyjma bílých znaků, mřížky (`#`) a zpětného lomítka (`\`)) a escape sekvencí, takže není ohraničen uvozovkami. Escape sekvence, která je nezbytná pro znaky s dekadickým kódem 000–032, 035 a 092, je tvaru `\xyz`, kde `xyz` je dekadické číslo v rozmezí 000–999 složené právě ze tří číslic¹⁸; např. konstanta

```
string@řetězec\032s\032lomítkem\032\092\032a\010novým\035řádkem
```

reprezentuje řetězec

```
řetězec s lomítkem \ a  
novým#řádkem
```

Pokus o práci s neexistující proměnnou (čtení nebo zápis) vede na chybu 54. Pokus o čtení hodnoty neinicializované proměnné vede na chybu 56. Pokus o interpretaci instrukce s operandy nevhodných typů dle popisu dané instrukce vede na chybu 53.

¹⁸Zápis znaků s kódem Unicode větším jak 126 pomocí těchto escape sekvencí nebudeme testovat.

6.4 Instrukční sada

U popisu instrukcí sázíme operační kód tučně a operandy zapisujeme pomocí neterminálních symbolů (případně číslovaných) v úhlových závorkách. Neterminál $\langle var \rangle$ značí proměnnou, $\langle symb \rangle$ konstantu nebo proměnnou, $\langle label \rangle$ značí návěští. Identifikátor proměnné se skládá ze dvou částí oddělených zavináčem (znak @; bez bílých znaků), označení rámce LF, TF nebo GF a samotného jména proměnné (sekvence libovolných alfanumerických a speciálních znaků bez bílých znaků začínající písmenem nebo speciálním znakem, kde speciální znaky jsou: $_$, $-$, $\$$, $\&$, $\%$, $*$, $!$, $?$). Např. GF@_x značí proměnnou _x uloženou v globálním rámci.

Na zápis návěští se vztahují stejná pravidla jako na jméno proměnné (tj. část identifikátoru za zavináčem).

Příklad jednoduchého programu v IPPcode21:

```
.IPPcode21
DEFVAR GF@counter
MOVE GF@counter string@ #Inicializace proměnné na prázdný řetězec
#Jednoduchá iterace, dokud nebude splněna zadaná podmínka
LABEL while
JUMPIFEQ end GF@counter string@aaa
WRITE string@Proměnná\032GF@counter\032obsahuje\032
WRITE GF@counter
WRITE string@\010
CONCAT GF@counter GF@counter string@a
JUMP while
LABEL end
```

Instrukční sada nabízí instrukce pro práci s proměnnými v rámci, různé skoky, operace s datovým zásobníkem, aritmetické, logické a relační operace, dále také konverzní, vstupně/výstupní a ladicí instrukce.

6.4.1 Práce s rámci, volání funkcí

MOVE $\langle var \rangle$ $\langle symb \rangle$

Přiřazení hodnoty do proměnné

Zkopíruje hodnotu $\langle symb \rangle$ do $\langle var \rangle$. Např. MOVE LF@par GF@var provede zkopírování hodnoty proměnné var v globálním rámci do proměnné par v lokálním rámci.

CREATEFRAME

Vytvoř nový dočasný rámec

Vytvoří nový dočasný rámec a zahodí případný obsah původního dočasného rámce.

PUSHFRAME

Přesun dočasného rámce na zásobník rámců

Přesuň TF na zásobník rámců. Rámec bude k dispozici přes LF a překryje původní rámec na zásobníku rámců. TF bude po provedení instrukce nedefinován a je třeba jej před dalším použitím vytvořit pomocí CREATEFRAME. Pokus o přístup k nedefinovanému rámci vede na chybu 55.

POPFRAME

Přesun aktuálního rámce do dočasného

Přesuň vrcholový rámec LF ze zásobníku rámců do TF. Pokud žádný rámec v LF není k dispozici, dojde k chybě 55.

DEFVAR $\langle var \rangle$

Definuj novou proměnnou v rámci

Definuje proměnnou v určeném rámci dle $\langle var \rangle$. Tato proměnná je zatím neinicializovaná a bez určení typu, který bude určen až přiřazením nějaké hodnoty. Opakovaná definice proměnné již existující v daném rámci vede na chybu 52.

CALL $\langle label \rangle$

Skok na návěští s podporou návratu

Uloží inkrementovanou aktuální pozici z interního čítače instrukcí do zásobníku volání a provede skok na zadané návěští (případnou přípravu rámce musí zajistit jiné instrukce).

RETURN

Návrat na pozici uloženou instrukcí CALL

Vyjme pozici ze zásobníku volání a skočí na tuto pozici nastavením interního čítače instrukcí (úklid lokálních rámců musí zajistit jiné instrukce). Provedení instrukce při prázdném zásobníku volání vede na chybu 56.

6.4.2 Práce s datovým zásobníkem

Operační kód zásobníkových instrukcí je zakončen písmenem „S“. Zásobníkové instrukce případně načítají chybějící operandy z datového zásobníku a výslednou hodnotu operace případně ukládají zpět na datový zásobník.

PUSHS $\langle symb \rangle$

Vlož hodnotu na vrchol datového zásobníku

Uloží hodnotu $\langle symb \rangle$ na datový zásobník.

POPS $\langle var \rangle$

Vyjmi hodnotu z vrcholu datového zásobníku

Není-li zásobník prázdný, vyjme z něj hodnotu a uloží ji do proměnné $\langle var \rangle$, jinak dojde k chybě 56.

6.4.3 Aritmetické, relační, booleovské a konverzní instrukce

V této sekci jsou popsány tříadresné instrukce pro klasické operace pro výpočet výrazu. Přetečení nebo podtečení číselného výsledku neřešte.

ADD $\langle var \rangle \langle symb_1 \rangle \langle symb_2 \rangle$

Součet dvou číselných hodnot

Sečte $\langle symb_1 \rangle$ a $\langle symb_2 \rangle$ (musí být typu int) a výslednou hodnotu téhož typu uloží do proměnné $\langle var \rangle$.

SUB $\langle var \rangle \langle symb_1 \rangle \langle symb_2 \rangle$

Odečítání dvou číselných hodnot

Odečte $\langle symb_2 \rangle$ od $\langle symb_1 \rangle$ (musí být typu int) a výslednou hodnotu téhož typu uloží do proměnné $\langle var \rangle$.

MUL $\langle var \rangle \langle symb_1 \rangle \langle symb_2 \rangle$

Násobení dvou číselných hodnot

Vynásobí $\langle symb_1 \rangle$ a $\langle symb_2 \rangle$ (musí být typu int) a výslednou hodnotu téhož typu uloží do proměnné $\langle var \rangle$.

IDIV $\langle var \rangle \langle symb_1 \rangle \langle symb_2 \rangle$

Dělení dvou celočíselných hodnot

Celočíselně podělí celočíselnou hodnotu ze $\langle symb_1 \rangle$ druhou celočíselnou hodnotou ze $\langle symb_2 \rangle$ (musí být oba typu int) a výsledek typu int přiřadí do proměnné $\langle var \rangle$. Dělení nulou způsobí chybu 57.

LT/GT/EQ $\langle var \rangle \langle symb_1 \rangle \langle symb_2 \rangle$

Relační operátory menší, větší, rovno

Instrukce vyhodnotí relační operátor mezi $\langle symb_1 \rangle$ a $\langle symb_2 \rangle$ (stejného typu; int, bool nebo string) a do $\langle var \rangle$ zapíše výsledek typu bool (**false** při neplatnosti nebo **true** v případě platnosti odpovídající relace). Řetězce jsou porovnávány lexikograficky a **false** je menší než **true**. Pro výpočet neostrých nerovností lze použít AND/OR/NOT. S operandem typu nil (další zdrojový operand je libovolného typu) lze porovnávat pouze instrukcí EQ, jinak chyba 53.

AND/OR/NOT $\langle var \rangle \langle symb_1 \rangle \langle symb_2 \rangle$

Základní booleovské operátory

Aplikuje konjunkci (logické A)/disjunkci (logické NEBO) na operandy typu bool $\langle symb_1 \rangle$ a $\langle symb_2 \rangle$ nebo negaci na $\langle symb_1 \rangle$ (NOT má pouze 2 operandy) a výsledek typu bool zapíše do $\langle var \rangle$.

INT2CHAR $\langle var \rangle \langle symb \rangle$

Převod celého čísla na znak

Číselná hodnota $\langle symb \rangle$ je dle Unicode převedena na znak, který tvoří jednoznakový řetězec přiřazený do $\langle var \rangle$. Není-li $\langle symb \rangle$ validní ordinální hodnota znaku v Unicode (viz funkce chr v Python 3), dojde k chybě 58.

STR2INT $\langle var \rangle$ $\langle symb_1 \rangle$ $\langle symb_2 \rangle$	Ordinální hodnota znaku
Do $\langle var \rangle$ uloží ordinální hodnotu znaku (dle Unicode) v řetězci $\langle symb_1 \rangle$ na pozici $\langle symb_2 \rangle$ (indexováno od nuly). Indexace mimo daný řetězec vede na chybu 58. Viz funkce <code>ord</code> v Python 3.	

6.4.4 Vstupně-výstupní instrukce

READ $\langle var \rangle$ $\langle type \rangle$	Načtení hodnoty ze standardního vstupu
Načte jednu hodnotu dle zadaného typu $\langle type \rangle \in \{\text{int}, \text{string}, \text{bool}\}$ a uloží tuto hodnotu do proměnné $\langle var \rangle$. Načtení provedte vestavěnou funkcí <code>input()</code> (či analogickou) jazyka Python 3, pak proveďte konverzi na specifikovaný typ $\langle type \rangle$. Při převodu vstupu na typ <code>bool</code> nezáleží na velikosti písmen a řetězec „true“ se převádí na <code>bool@true</code> , vše ostatní na <code>bool@false</code> . V případě chybného nebo chybějícího vstupu bude do proměnné $\langle var \rangle$ uložena hodnota <code>nil@nil</code> .	

WRITE $\langle symb \rangle$	Výpis hodnoty na standardní výstup
Vypíše hodnotu $\langle symb \rangle$ na standardní výstup. Až na typ <code>bool</code> a hodnotu <code>nil@nil</code> je formát výpisu kompatibilní s příkazem <code>print</code> jazyka Python 3 s doplňujícím parametrem <code>end=''</code> (zamezí dodatečnému odřádkování). Pravdivostní hodnota se vypíše jako <code>true</code> a nepravda jako <code>false</code> . Hodnota <code>nil@nil</code> se vypíše jako prázdný řetězec.	

6.4.5 Práce s řetězci

CONCAT $\langle var \rangle$ $\langle symb_1 \rangle$ $\langle symb_2 \rangle$	Konkatenace dvou řetězců
Do proměnné $\langle var \rangle$ uloží řetězec vzniklý konkatenací dvou řetězcových operandů $\langle symb_1 \rangle$ a $\langle symb_2 \rangle$ (jiné typy nejsou povoleny).	

STRLEN $\langle var \rangle$ $\langle symb \rangle$	Zjistí délku řetězce
Zjistí počet znaků (délku) řetězce v $\langle symb \rangle$ a tato délka je uložena jako celé číslo do $\langle var \rangle$.	

GETCHAR $\langle var \rangle$ $\langle symb_1 \rangle$ $\langle symb_2 \rangle$	Vrať znak řetězce
Do $\langle var \rangle$ uloží řetězec z jednoho znaku v řetězci $\langle symb_1 \rangle$ na pozici $\langle symb_2 \rangle$ (indexováno celým číslem od nuly). Indexace mimo daný řetězec vede na chybu 58.	

SETCHAR $\langle var \rangle$ $\langle symb_1 \rangle$ $\langle symb_2 \rangle$	Změň znak řetězce
Zmodifikuje znak řetězce uloženého v proměnné $\langle var \rangle$ na pozici $\langle symb_1 \rangle$ (indexováno celočíselně od nuly) na znak v řetězci $\langle symb_2 \rangle$ (první znak, pokud obsahuje $\langle symb_2 \rangle$ více znaků). Výsledný řetězec je opět uložen do $\langle var \rangle$. Při indexaci mimo řetězec $\langle var \rangle$ nebo v případě prázdného řetězce v $\langle symb_2 \rangle$ dojde k chybě 58.	

6.4.6 Práce s typy

TYPE $\langle var \rangle$ $\langle symb \rangle$	Zjistí typ daného symbolu
Dynamicky zjistí typ symbolu $\langle symb \rangle$ a do $\langle var \rangle$ zapíše řetězec značící tento typ (<code>int</code> , <code>bool</code> , <code>string</code> nebo <code>nil</code>). Je-li $\langle symb \rangle$ neinicizovaná proměnná, označí její typ prázdným řetězcem.	

6.4.7 Instrukce pro řízení toku programu

Neterminál $\langle label \rangle$ označuje návěští, které slouží pro označení pozice v kódu IPPcode21. V případě skoku na neexistující návěští dojde k chybě 52.

LABEL $\langle label \rangle$	Definice návěští
Speciální instrukce označující pomocí návěští $\langle label \rangle$ důležitou pozici v kódu jako potenciální cíl libovolné skokové instrukce. Pokus o vytvoření dvou stejně pojmenovaných návěští na různých místech programu je chybou 52.	

JUMP $\langle label \rangle$	Nepodmíněný skok na návěští
Provede nepodmíněný skok na zadané návěští $\langle label \rangle$.	

JUMPIFEQ $\langle label \rangle$ $\langle symb_1 \rangle$ $\langle symb_2 \rangle$ Podmíněný skok na návěští při rovnosti
Pokud jsou $\langle symb_1 \rangle$ a $\langle symb_2 \rangle$ stejného typu nebo je některý operand nil (jinak chyba 53) a zároveň se jejich hodnoty rovnají, tak provede skok na návěští $\langle label \rangle$.

JUMPIFNEQ $\langle label \rangle$ $\langle symb_1 \rangle$ $\langle symb_2 \rangle$ Podmíněný skok na návěští při nerovnosti
Jsou-li $\langle symb_1 \rangle$ a $\langle symb_2 \rangle$ stejného typu nebo je některý operand nil (jinak chyba 53), tak v případě různých hodnot provede skok na návěští $\langle label \rangle$.

EXIT $\langle symb \rangle$ Ukončení interpretace s návratovým kódem
Ukončí vykonávání programu, případně vypíše statistiky a ukončí interpret s návratovým kódem $\langle symb \rangle$, kde $\langle symb \rangle$ je celé číslo v intervalu 0 až 49 (včetně). Nevalidní celočíselná hodnota $\langle symb \rangle$ vede na chybu 57.

6.4.8 Ladicí instrukce

Následující ladicí instrukce (DPRINT a BREAK) nesmí ovlivňovat standardní výstup. Jejich skutečnou funkcionalitu nebudeme testovat, ale mohou se v testech objevit.

DPRINT $\langle symb \rangle$ Vypis hodnoty na **stderr**
Předpokládá se, že vypíše zadanou hodnotu $\langle symb \rangle$ na standardní chybový výstup (stderr).

BREAK Vypis stavu interpretu na **stderr**
Předpokládá se, že na standardní chybový výstup (stderr) vypíše stav interpretu (např. pozice v kódu, obsah rámců, počet vykonaných instrukcí) v danou chvíli (tj. během vykonávání této instrukce).

Reference

- [1] Extensible Markup Language (XML) 1.0. W3C. World Wide Web Consortium [online]. 5. vydání. 26. 11. 2008 [cit. 2020-02-03]. Dostupné z: <https://www.w3.org/TR/xml/>
- [2] A7Soft JExamXML is a java based command line XML diff tool for comparing and merging XML documents. c2018 [cit. 2020-02-03]. Dostupné z: <https://www.a7soft.com/jexamxml.html>
- [3] Křivka, Z., a kol.: Zadání projektu z předmětu IFJ a IAL. c2020 [cit. 2021-01-30]. Dostupné z: <https://www.fit.vutbr.cz/study/courses/IFJ/private/projekt/ifj2020.pdf>
- [4] The text/markdown Media Type. Internet Engineering Task Force (IETF). 2016 [cit. 2020-02-03]. Dostupné z: <https://tools.ietf.org/html/rfc7763>

Revize zadání:

2021-02-10: Opravy překlepů, pravopisných a stylistických chyb.

2021-02-15: Některé statistiky ve STATP nebudou zahrnovat návraty z volání. Přeformulovány dvě věty v sekci 3.1.