

Sem vložte zadání Vaší práce.

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
KATEDRA TEORETICKÉ INFORMATIKY



Diplomová práce

Comma-shell, interaktivní debugger shellu

Bc. Tomáš Nesrovnal

Vedoucí práce: Ing. Jan Baier

22. dubna 2017

Poděkování

Doplňte, máte-li komu a za co děkovat. V opačném případě úplně odstráňte tento příkaz.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval(a) samostatně a že jsem uvedl(a) veškeré použité informační zdroje v souladu s Metodickým pokynem o etické přípravě vysokoškolských závěrečných prací.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona, ve znění pozdějších předpisů. V souladu s ust. § 46 odst. 6 tohoto zákona tímto uděluji nevýhradní oprávnění (licenci) k užití této mojí práce, a to včetně všech počítačových programů, jež jsou její součástí či přílohou, a veškeré jejich dokumentace (dále souhrnně jen „Dílo“), a to všem osobám, které si přejí Dílo užít. Tyto osoby jsou oprávněny Dílo užít jakýmkoli způsobem, který nesnižuje hodnotu Díla, a za jakýmkoli účelem (včetně užití k výdělečným účelům). Toto oprávnění je časově, teritoriálně i množstevně neomezené. Každá osoba, která využije výše uvedenou licenci, se však zavazuje udělit ke každému dílu, které vznikne (byť jen zčásti) na základě Díla, úpravou Díla, spojením Díla s jiným dílem, zařazením Díla do díla souborného či zpracováním Díla (včetně překladu), licenci alespoň ve výše uvedeném rozsahu a zároveň zpřístupnit zdrojový kód takového díla alespoň srovnatelným způsobem a ve srovnatelném rozsahu, jako je zpřístupněn zdrojový kód Díla.

V Praze dne 22. dubna 2017

.....

České vysoké učení technické v Praze

Fakulta informačních technologií

© 2017 Tomáš Nesrovnal. Všechna práva vyhrazena.

Tato práce vznikla jako školní dílo na Českém vysokém učení technickém v Praze, Fakultě informačních technologií. Práce je chráněna právními předpisy a mezinárodními úmluvami o právu autorském a právech souvisejících s právem autorským. K jejímu užití, s výjimkou bezúplatných zákonných licencí, je nezbytný souhlas autora.

Odkaz na tuto práci

Nesrovnal, Tomáš. *Comma-shell, interaktivní debugger shellu*. Diplomová práce. Praha: České vysoké učení technické v Praze, Fakulta informačních technologií, 2017. Dostupný také z WWW: (<https://github.com/nesro/nesro-tom-dip-2016>).

Abstrakt

V několika větách shrňte obsah a přínos této práce v češtině. Po přečtení abstraktu by se čtenář měl mít čtenář dost informací pro rozhodnutí, zda chce Vaši práci číst.

Klíčová slova Nahradte seznamem klíčových slov v češtině oddělených čárkou.

Abstract

Sem doplňte ekvivalent abstraktu Vaší práce v angličtině.

Keywords Nahradte seznamem klíčových slov v angličtině oddělených čárkou.

Obsah

Úvod	1
Motivace	1
1 Definice a pojmy	3
2 Historie shellu a dnešní využití	5
2.1 Windows	5
2.2 UNIX	5
2.3 Linux	7
2.4 Shell	7
2.5 cmd.exe	7
2.6 Power Shell	8
2.7 v9 - Thomson Shell	8
2.8 Sh - Bourne Shell	8
2.9 Dash - Debian Almquist Shell	9
2.10 Csh - C Shell	9
2.11 Ksh - Korn Shell	9
2.12 Bash - Bourne Again Shell	9
2.13 Zsh - Z Shell	9
2.14 Práce v příkazové řádce	10
2.15 Nastavení shellu	11
2.16 Terminál	11
2.17 Existující řešení	11
2.18 Bash frameworky pro psaní skriptů	12
2.19 Bash frameworky pro správu doplňků	12
2.20 Bash frameworky pro úpravu promptu	12
3 Analýza a návrh	15
3.1 Cíl práce	15
3.2 Fungování shellu	16

3.3	Debugování shellu	19
3.4	Možnosti debugování v interaktivním shellu	20
3.5	Statická analýza skriptů	21
4	Realizace	23
4.1	Nespouštění příkazů	23
4.2	Implementace debuggeru	23
4.3	Hooks	23
4.4	Bezpečný mód	24
4.5	Historie	25
4.6	Instalace	25
5	Testování	27
5.1	Automatizované spouštění příkazů	27
	Závěr	29
	Literatura	31
A	Seznam použitých zkratk	33
B	Obsah příloženého CD	35

Seznam obrázků

2.1	Historie unixu	6
2.2	Historie Unixového Shellu	7
3.1	Architektura shellu	16
3.2	Bash diagram	19

Úvod

Grafické uživatelské rozhraní (GUI) se jednoduše ovládá, ale ne vždy je k dispozici. To platí zejména při ovládání serverů.

Rozhraní příkazové řádky (CLI) je základní textové prostředí pro komunikaci s operačním systémem. Umožňuje spouštění programů, vkládat vstupní data a sledovat výstupní data v terminálu.

Jedním ze základních bodů UNIXové filosofie je mít jednoduché programy, které dělají pouze jednu věc, ale dělají ji dobře. To platí zejména pro základní příkazy ze sady GNU coreutils, tedy příkazy pro základní manipulaci se soubory, shellem a textem.

Tyto základní příkazy je možné řetězit a tím vytvářet užitečné jednořádkové skripty.

TODO: napsat o tom, že pro začátečníka to může být neintuitivní, musí si pamatovat spoustu příkazů. o návratových kódech, o historii příkazu a o logování

Motivace

Příkazová řádka a základní práce s ní by měly být jedna z prvních věcí, které by se uživatel měl naučit, pokud chce umět používat svůj systém na vyšší úrovni. Stejně jako v programování platí, že počítač provádí dokonale a přesně to, co mu řekneme. Ale i při sebemenší chybě v našem příkazu selže celá sekvence akcí, které se měly provést. Není těžké udělat i nějakou fatální chybu, po které není snadné uvést vše do původního stavu.

Náš projekt, Comma-shell, si dává za úkol vytvořit z příkazové řádky místo, ve kterém je těžší nějaké chyby udělat, případně před chybou varovat, nebo vysvětlit, jak danou věc udělat lépe. Důležitý cíl je také umožnit nastavit toto prostředí tak, aby bylo snadno použitelné, nebylo potřeba externích terminálů, nebo grafických aplikací a aby bylo možné snadno omezit repetitivní výstupy programu. Comma-shell se dělí na tři části, každá pomáhá řešit jiný problém.

Comma-shell hooks, nebo-li části kódu, které lze vykonat před, nebo po vykonání spuštěného příkazu a dokonce zabránit jeho spuštění umožňují ochránit uživatele před vykonáním příkazu, který je v například nějakém smyslu nesprávný, špatný, nebo je v nepořádku. Pokud tedy napíše příkaz, který obsahuje nějaký překlep, nebo chybu, která by způsobila neočekávané chování, může být uživatel varován, než bude takový příkaz proveden. Nejde jen o ochranu uživatele před špatnými příkazy. Díky informacím o spouštěném příkazu a následně o jeho výsledku lze psát různé skripty, které by jinak bylo složité vytvořit.

Comma-shell debugger, tedy ladící program pro příkazy spouštěné z interaktivní příkazové řádky. Pokud uživatel píše složitější program, obsahující například smyčku, nebo roury a příkaz nefunguje tak, jak uživatel očekává, tak nejjednodušší je rozdělit si příkaz na podpříkazy a ujistit se, že každý příkaz dělá to, co od něj uživatel očekává. Comma-shell debugger by měl tento postup automatizovat.

Poslední částí jsou takzvané bezpečné příkazy. Pokud uživatel vykoná příkaz, jehož efekt byl jiný než očekával, chce spustit opačný příkaz, aby napravit svůj omyl, nebo má často mnohem větší problém, pokud šlo například o nesprávné použití programu pro mazání souborů. Cílem bezpečných příkazů je tedy zaprvé umožnit vrátit provedené změny, ale i informovat uživatele o tom, co se stane. Pokud uživatel používá ke specifikování souborů, které mají být cílem nějakého příkazu, je častý omyl vybrat nějaké soubory nechtěně. Nebo například provést příkaz ze špatného adresáře.

Definice a pojmy

Příkazová řádka, anglicky Command Line Interface, je uživatelské rozhraní ovládané příkazy. Příkazy spuštěné z příkazové řádky vypisují výstup také na příkazovou řádku a další interakce se spuštěným programem probíhá zase zadáváním příkazů. Na rozdíl od textového nebo grafického rozhraní uživatel nemůže nijak zasahovat do již vypsané či vykreslené části.

TODO: obrazek prikazove radky, tui a gui

Shellem se obecně myslí uživatelské prostředí pro využívání funkcí jádra operačního systému. V této práci budeme slovem shell označovat právě interpret příkazů v příkazové řádce. Tedy program, ovládaný příkazy, který umí spouštět ostatní programy a zajišťovat jejich vstup a zobrazovat jejich výstup.

Prompt je krátký řetězec znaků, který se objeví na začátku řádky, do které se bude vepisovat příkaz v příkazové řádce. Je používán k zobrazování důležitých informací, jako například, ve kterém adresáři se uživatel nachází či jaké je jméno počítače, na kterém se nachází.

GNU je projekt zabývající se tvorbou a propagací svobodného software. Cílem je vytvořit kompletní svobodný operační systém unixového typu. Protože jádro operačního systému od GNU s názvem Hurd nebylo tak úspěšné, začalo se používat jádro Linux. Kombinace operačního systému GNU a jádra Linux se označuje jako GNU/Linux. GNU vytvořilo i svůj shell, Bash. Sada základních příkazů, jejichž existence je předpokládána v nějaké formě na každém operačním systému, se jmenuje GNU coreutils. Patří do ní příkazy jako ls, cp, rm, atd. Je důležité si uvědomit, že příkazy jako cd, echo jsou vestavěné příkazy shellu Bash.

Terminál byl z historického hlediska elektronické zařízení k základní komunikaci s počítačem. Dnes se v operačních systémech spouští terminálové emulátory, ve kterých lze spustit různé programy. Nejčastěji je to právě shell.

Provádění příkazů můžeme rozdělit na dvě kategorie. Jednou z nich je interaktivní režim, kdy interpret spouští příkazy tak, jak je uživatel píše do příkazové řádky. V tomto režimu může uživatel také používat klávesové zkratky, aliasy, automatické doplňování a další funkce v interaktivním režimu. Druhou

1. DEFINICE A POJMY

kategorií je dávkový režim, kdy uživatel nejprve příkazy napíše do souboru. Takovému souboru obsahujícímu příkazy se říká skript. Takový soubor může být zavolán interpretem shellu, který postupně vykoná všechny příkazy. Tento postup se hodí pro složitější a opakující se příkazy.

Historie shellu a dnešní využití

V této kapitole stručně shrneme historii rodin nepoužívanějších operačních systému a shellů, které se v nich používají.

2.1 Windows

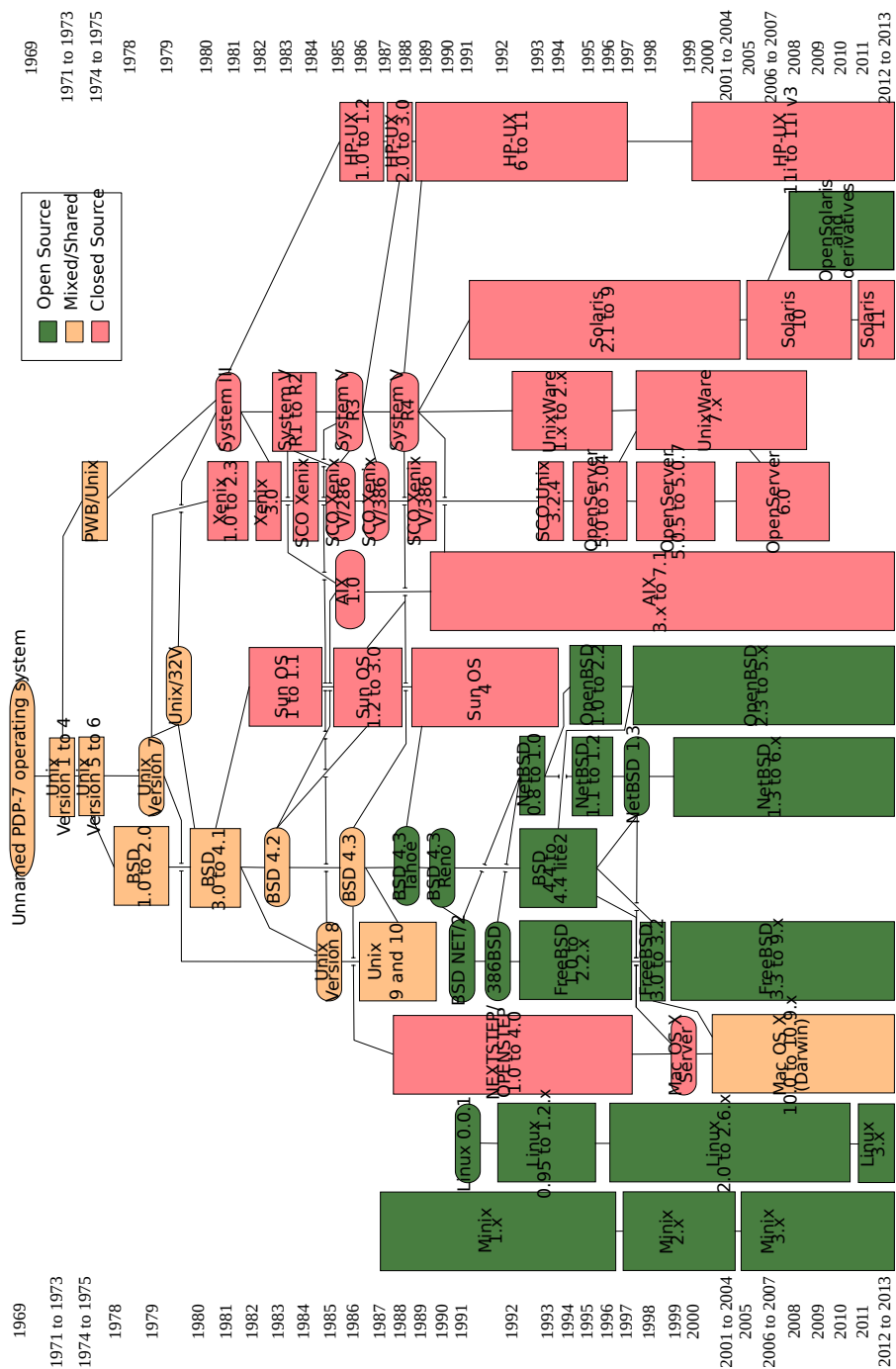
Operační systém Windows je obecně uživatelsky přívětivější, protože skoro vše je možné nastavit přes grafické rozhraní. Ovšem ne vždy je grafické prostředí dostupné a příkazová řádka je jediná možnost, jak ovládat počítač.

2.2 UNIX

Unix je rodina operačních systémů, které zvládnou spustit více úkonů najednou a do kterých se v jeden okamžik může připojit více uživatelů. Historie těchto systémů sahá až na konec sedmdesátých let, kdy v Bell laboratořích firmy AT&T byl v jazyce C vytvořen systém UNIX, jehož název byl zaregistrován jako ochranná známka. Časovou osu ukazuje obrázek 2.1. Specifikace "Single UNIX Specification" je souhrn standardu, které operační systém musí splňovat a dodržovat, aby se mohl označovat za UNIX.

Filosofie Unixu je sada doporučení a pravidel, která vznikla postupem času dle zkušeností tvůrců Unixu. Filosofie Unixu zdůrazňuje tvoření jednoduchého, přehledného a hlavně snadno rozšiřitelného a znovupoužitelného softwaru. Váží si programátorův čas více, než čas práce počítače. Nejznámější pravidlo je tvořit programy, které dělají jen jednu věc a dělají ji správně, rychle a bez chyb. Takové programy lze použít jako filtry a skládat je za sebe. To se často používá v příkazové řádce, kdy se pomocí rour spojují výstupy programů se vstupy jiných programů. Zdánlivě složitý úkol tak lze udělat rychle a efektivně bez nutnosti programování v nějakém jazyce.

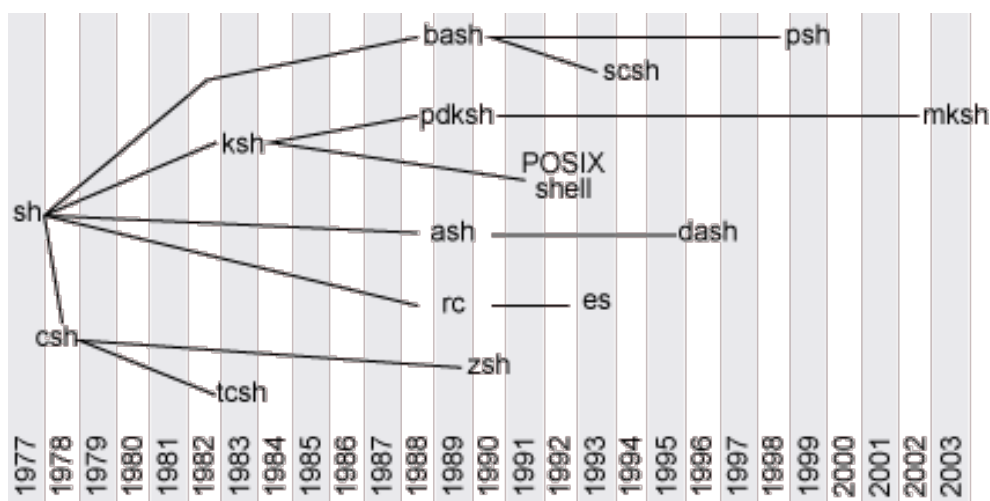
2. HISTORIE SHELLU A DNEŠNÍ VYUŽITÍ



2.3 Linux

Linux je otevřený a svobodný Unix-like operační systém. Linuxem se obecně myslí GNU operační systém s jádrem Linux. Unix-like znamená, že přestože je takový operační systém velice podobný systému Unix, nemá certifikaci "Single UNIX Specification". Linuxová distribuce je GNU/Linuxový systém, který většinou obsahuje balíčkovací systém, kterým je snadné doinstalovat ostatní programy. Linuxové distribuce jsou i předpřipravené, s grafickým prostředím. Například v této práci ukazujeme návod na vytvoření takového připraveného systému, GNU/Linuxové distribuci Debian s předpřipraveným grafickým prostředím LXDE a aplikacemi z rodiny LXDE.

2.4 Shell



Obrázek 2.2: Historie Unixového Shellu

2.5 cmd.exe

Shellem v operačním systému MS-DOS a PC-DOS byl COMMAND.COM, 16-bitový intepretr příkazů. V novějších systémech Windows byl tento shell nahrazen CMD.EXE, 32-bitovým interpretem příkazů. V CMD.EXE byla dostupná emulace COMMAND.COM, pro zpětnou kompatibilitu.

Tato příkazová řádka, shell, již umí přesměrovávat vstupy a výstupy a to jak do souboru tak mezi procesy. Také je schopná vykonávat příkazy ze souborů, které se nazývají dávkové a zpravidla mívají příponu bat.

2.6 Power Shell

Protože je základní příkazová řádka Windows omezená co se funkcionality týče, existují další shelly pro operační systémy Windows, často i s otevřeným zdrojovým kódem. Firma Microsoft chtěla vytvořit i svůj shell, který bude umět více věcí. S jeho vývojem začala v roce 2002 a po 3 letech jej zveřejnila jako Microsoft Command Shell (MSH) s krycím jménem Monad, který byl později přejmenován na Power Shell.

Power Shell je integrován s platformou .NET od Microsoftu. Je tedy možné využívat již vytvořené funkce a nástroje, které usnadní a urychlí vývoj.

Oproti cmd.exe umí Power Shell mnoho věcí. Power Shell umožňuje spouštět příkazy v intervalech, nebo v nějaký přesný čas. Přibyly konstrukce pro větvení. Příkazy je možné spouštět na pozadí, nebo vzdáleně.

2.7 v9 - Thomson Shell

Pro první shell pro systém UNIX napsal Ken Thomson v roce 1971 a byl nazván V9 shell. Jednalo se již o uživatelský program, tedy spuštěný mimo jádro operačního systému. Shell byl velice jednoduchý a některá základní funkcionality byla umožněna dalšími programy. Například expanze parametrů byla zajišťována příkazem glob. Základní větvení, if, bylo také jako samostatný program.

V9 shell přinesl syntaxi pro přesměrování a pipelining. Uměl spouštět více oddělených příkazů a dokázal příkazy spouštět na pozadí. Naopak chyběla třeba podpora pro spouštění skriptů.

Díky jednoduchosti a oddělenosti ostatní funkcionality měly zdrojové kódy pouze 900 řádek v jazyce C.

Originální zdrojové kódy jsou stále dostupné, <https://github.com/yvesnrb/Thompson-Shell>. Projekt Osh, <https://v6shell.org/>, obsahuje jak port tohoto původního shellu, tak i jeho vylepšenou verzi.

2.8 Sh - Bourne Shell

Bourne shell byl vytvořen pro UNIX V7 v roce 1977 Stephenem Bournem v laboratořích Bell a AT&T. Na rozdíl od Thomsonova v9 shellu umí Bourne shell spouštět příkazy ze souborů, skriptů. Šlo tedy psát znovupoužitelné sady příkazů, které zjednodušili a urychlili vývoj různých aplikací. Interaktivní režim je samozřejmě podporován také.

Bourne shell přinesl také podporu smyček, proměnných, signálů, subshelly a HERE dokumenty. Skripty v shellu mohly sloužit i jako filtry. Z tohoto shellu poté začaly vznikat další, jak je vidět na obrázku 2.2.

Dnes se Bourne shell stále používá. Je totiž menší, rychlejší a stabilnější než ostatní shelly. Má také minimum závislostí na externí knihovny.

2.9 Dash - Debian Almquist Shell

Dash je moderní implementace Bourne Shellu. Snaží se být stále malý a rychlý, ale přesto třeba i bezpečný a stabilní při chybách systému. Důležitý je fakt, že tento shell splňuje standardy POSIX 1003.2 a 1003.2a, tedy standardy o regulárních výrazech, respektive jak se má chovat shell a jeho základní nástroje.

Tento shell byl nastaven jako shell, který spouští příkazy při bootování i skripty, které mají jako shebang `/bin/sh` na operačních systémech Ubuntu a Debian.

2.10 Csh - C Shell

C Shell byl vytvořen Bill Joyem v sedmdesátých letech. Hlavní myšlenkou byla syntaxe více podobná jazyku C. Dnes se používá vylepšená verze `tcsh`.

Přestože není tento shell moc oblíbený v dnešní době, při svém vzniku představil některé nové vlastnosti, které pak začali adaptovat i ostatní shelly. Byla to například historie příkazů, nebo podpora aliasů.

2.11 Ksh - Korn Shell

Korn Shell je Unixový shell, který vytvořil David Korn a jeho kolegové z Bell laboratoří na začátku osmdesátých let. Korn Shell byl vyvíjen z Bourne Shellu a zachovává zpětnou kompatibilitu.

Korn Shell umožňuje upravovat psaný příkaz, jako v editorech Vi, nebo Emacs. Podporu spravování spuštěných programů, historie a aliasy příkazů byly převzaty z C Shellu. Korn Shell dovoluje vytvářet asociativní pole a umí pracovat s desetinnými čísly.

2.12 Bash - Bourne Again Shell

Bash je open source shell projektu GNU. Bash se snaží být zpětně kompatibilní s Bourne shellem, tedy skripty napsané v Bourne shellu pravděpodobně půjdou spustit i v Bashi. Naopak to je problém, který se řeší program `checkbashisms`, o kterém bude řeč dále.

todo

2.13 Zsh - Z Shell

Z Shell je Unixový shell, pojmenovaný po přihlašovacím jménu profesorovi z Yale, Zhong Shao. Byl v roce 1990 napsán Paulem Falstadem. Ze standardních shellů.

Z Shell nejenom že nabízí nejvíce funkcí mezi používanými Shelly pro Unixové systémy, ale také je kolem něj velká aktivní komunita spravující mnoho pluginů, v čele s Oh My Zsh, o kterém bude řeč dále.

Funkcí navíc oproti ostatním shellům je hodně. Hned po instalaci Z Shellu se uživateli nabídne možnost nastavení, jako například nastavení ukládání historie, nebo automatické doplňování. Z Shell má na rozdíl od Bashe nativní podporu automatického doplňování, která je nejen rychlejší, je více způsobů, jak s ní iteragovat, ale je i snadnější na implementaci do uživatelských skriptů.

Z Shell umožňuje zobrazit hodnotu promptu PS1 i v pravo. Tím se otevírají další možnosti pro zobrazování informací.

Ukládání historie je automaticky sdíleno mezi jednotlivými instancemi Z Shellu. Takové chování je možné nastavit i například v Bashi, ale je to pro to potřeba ukládat každý příkaz do souboru na disk.

Další z věcí ovlivňující pohodlnost práce v Z Shellu je takzvaný globbing, který je lepší než v ostatních shellech. Jedná se o popsání jednoho, nebo několik řetězců, zpravidla názvu souborů, kratším zápisem. Například *.txt se zamění za všechny soubory ve složce s příponou txt. V Z Shellu je ale možné popsat cestu k souborům pomocí zkratk a nechat ji doplnit. Například /h/u/t/s se změní na /home/user/test/script.sh.

Velmi užitečnou funkcí je automatická oprava překlepů. V Bashi je velmi limitovaná oprava příkazů, ale v Z Shellu se opravují nejen názvy příkazů, ale také argumentů. Pokud je argumentem cesta k souboru, který neexistuje, Z Shell navrhne cestu k souboru, který existuje a má podobnou cestu.

Přestože Z Shell se zdá být ve všech směrech lepší než Bash, tak není základním shellem ve většině Linuxových distribucích.

<http://zsh.sourceforge.net/FAQ/>

2.14 Práce v příkazové řádce

Práce v příkazové řádce může být mnohem efektivnější na rozdíl od práce v grafickém prostředí. Při normální práci není tolik nutné používat myš. Složitější, nebo méně používané funkce a programy je zpravidla rychlejší napsat, než vyhledat v nějakém kontextovém menu. To samozřejmě platí až tehdy, kdy uživatel ví, jakým příkazem a s jakými parametry se příkazy spouští.

Nový uživatel, který ještě není s prací v příkazové řádce zběhlý, bude ze začátku méně efektivní, ale časem si základní příkazy zapamatuje a jeho efektivita bude růst.

Další výhodou je možnost používání příkazů ve formě filtrů. I složitou věc je často možné vyřešit pomocí spuštění několika příkazů v rouři za sebou. Příkazy v mezikrocích mohou být také uživatelské skripty.

Velice snadné je také automatizace. Pokud uživatel často potřebuje vykonat nějakou sekvenci příkazů, může složené příkazy vyhledat v historii a třeba z nich vytvořit skript.

Nevýhodou na rozdíl od práce v grafickém prostředí je často fakt, že zpravidla ne všechny potřebné informace jsou stále na očích. Tento problém lze vyřešit rozšířeným a vylepšeným promptem, nebo emulátorem terminálů, které mohou vyhradit nějakou část terminálu pro zobrazování informací. I přes to se stává, že uživatel musí často spouštět nějaké příkazy, které vypíší nějaké informace.

S rychlejší prací přichází také větší zodpovědnost za spuštěné příkazy. Překlep v příkaze, který provádí nějaké změny může způsobit komplikace. Některé programy se uživatele dotazují, zda-li chce operaci opravdu provést, ale není to pravidlem. Nejlepší prevencí je si tedy příkaz před spuštěním přezkontrolovat.

2.15 Nastavení shellu

GNU/Linuxové distribuce, které nabízejí předpřipravené prostředí, mají pro výchozí shell bash připravené takzvané rc soubory, které nějakým způsobem upravují běžící interaktivní shell. Jednou z nejviditelnějších změn je nastavení proměnné PS1, o které bude řeč v pozdější kapitole, která určuje styl promptu. Ve výchozím nastavení bashe, tedy po spuštění bashe bez načtení rc souborů (`bash -norc`), se v PS1 zobrazuje pouze název shellu, jeho verze a zdali je uživatel root. Po načtení výchozích rc souborů se v PS1 ukazují informace jako je například jméno uživatele, jméno počítače a co je velmi důležité, jméno aktuálního adresáře.

Tyto informace jdou snadno získat použitím některých základních příkazů, například `whoami`, `hostname`, `pwd`. Protože tyto informace jsou důležité a potřebujeme je vědět pořád, chceme je mít pořád na očích.

2.16 Terminál

Protože je terminál elektronické zařízení, ukážeme zde pouze emulátory takových terminálů.

Prvním druhem jsou programy, které vytvoří nové okno v desktopovém prostředí operačního systému. Na operačních systémech Linux je to zpravidla X Window System.

Druhým druhem jsou programy, které se spustí už v nějakém terminálu a emulují další terminál. To se hodí například při vzdálené práci. Takový emulátor běží na straně serveru a pokud se přeruší spojení, terminál zůstane neporušen.

2.17 Existující řešení

Mezi seznamem různých zajímavých řešení awesome (github.com/sindresorhus/awesome) existuje i podsekcce awesome-shell (github.com/alebcay/awesome-shell), ve

které lze nalézt spoustu užitečných nástrojů pro práci s příkazovou řádkou, nebo psaním skriptů v bashi.

Spousta těchto nástrojů je nad rámec této práce.

2.18 Bash frameworky pro psaní skriptů

Existuje mnoho projektů, jejichž cílem je vytvořit framework v bashi, který má nějakým způsobem zjednodušit vytváření, především větších, skriptů v bashi. Spousta dnešních vývojářů je zvyklá na objektově orientovaný přístup k programování a je tedy pro těžké vytvořit větší program, který by zůstal přehledný.

2.18.1 Bash OO framework

Bash OO framework (github.com/niieani/bash-oo-framework) je framework napasny v bashi, který umožňuje vytvářet třídy, výjimky a testy. Jeho cílem je vytvořit prostředí pro psaní skriptu, kde se bude snadněji psát čitelný kód, bez částí, které se opakují.

2.19 Bash frameworky pro správu doplňků

Existuje spousta věcí, které uživatel příkazové řádky potřebuje občas řešit. Řešení pro daný problém je víc. Buď si uživatel vyřeší problém sám, nebo bude hledat řešení na internetu. Shellové frameworky, které umějí i spravovat doplňky, mohou mít řešení pro daný problém. Instalace a použití pak bude velmi jednoduché, intuitivní a bude zde nějaká záruka o funkčnosti a kvalitě.

Přestože je tato práce převážně o shellu bash, zsh je v tomto mnohem rozšířenější.

2.19.1 Oh My Zsh

Oh My Zsh ([ohmyz.sh](https://github.com/ohmyzsh)) je framework pro Zsh.

2.19.2 Bash-it

Bash-it (<http://github.com/Bash-it/bash-it>, <http://itsfoss.com/bash-it-terminal-tool/>) je framework pro Bash.

2.20 Bash frameworky pro úpravu promptu

2.20.1 Liquid Prompt

Projekt liquidprompt github.com/nojhan/liquidprompt je adaptivní prompt v interaktivním bashi.

2.20.2 Sexy Bash Prompt

Projekt sexy-bash-prompt <https://github.com/twolfson/sexy-bash-prompt> je další používaný prompt v interaktivním bashi.

Analýza a návrh

3.1 Cíl práce

Prvotním cílem práce bylo navrhnout a vytvořit shell, který bude asistovat studentům na začátku jejich studia shellu. Takový shell by je měl upozorňovat na chyby a navrhopvat lepší řešení. Přitom mělo být maximálně využito existujících řešení.

Protože nástroj na upozorňování chyb, konkrétně vynikající ShellCheck, již existuje, byla část cíle práce nastavena obecněji na analýzu a ovlivňování příkazů spouštěných z interaktivní příkazové řádky. Tím se v praxi myslí vytvořit nástroj, který dokáže spustit skripty před vykonáním příkazu a být je schopen ovlivnit, nebo zabránit jejich vykonávání.

Protože psaní špatných příkazů, ať je to jejich nesprávné použití, nebo použití nevhodných konstrukcí, může být i jen provedení nechtěné akce, dalším cílem práce je vytvořit sadu upravených základních programů z GNU coreutils, které uživatele nejprve upozorní na změny, které nastanou po spuštění takového příkazu a možnost vrátit stav změněných věcí do původního stavu.

Posledním cílem je vytvořit jednoduchý debugger příkazů spouštěných v interaktivním shellu. Na rozdíl od normálních debuggerů pro shell skripty, například bashdb - BASH Debugger, je cílem takového debuggeru maximální jednoduchost použití. Tím se například myslí spuštění debuggeru s předepsaným posledním příkazem pomocí spuštění jediného příkazu.

Nejdůležitějším cílem práce je vytvořit takový nástroj, který bude neinvazivní při normálním používání příkazové řádky. Nástroj, který bude začátečník intenzivně využívat, zatímco pokročilý uživatel ho bude mít na pozadí a čas od času ho upozorní na nějakou chybu, popřípadě mu pomůže urychlit práci na nějakém nesprávně fungujícím příkazu.

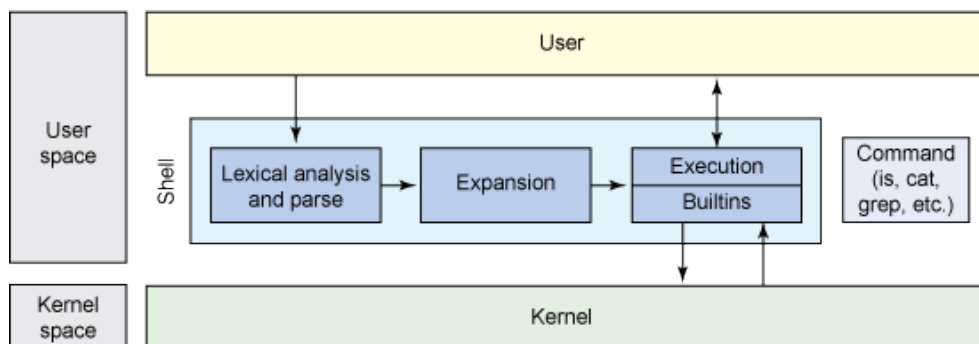
3.2 Fungování shellu

Abychom mohli implementovat jak hooky před a po příkazech, tak přepsat chování základních příkazů z GNU coreutils, je potřeba znát základní procesy, které se dějí při vykonávání příkazů.

Tyto postupy platí pro shell Bash, ale jsou velice podobné i v ostatních Unixových shellech.

3.2.1 Životní cyklus příkazu

V této kapitole vysvětlíme, co všechno všechno důležitého se stane mezi tím, kdy shell přijme příkaz a než je skutečně vykonán.



Obrázek 3.1: Architektura shellu

Prvním krokem je escapování, neboli převod znaků převod znaků, které mají ve svém kontextu speciální význam. V shellu lze tento postup provést pomocí znaku `\`, nebo tím, že řetězec znaků dáme do uvozovek. Příkladem může být znak `#`, který značí, že znaky za ním jsou pouze komentář. V ukázce 3.1 jsou čtyři příklady zavolání příkazu `echo`. V prvním případě ale znak `#` není vyescapován a Bash ho vyhodnotí jako začátek komentáře, proto se vypíše jen část, oproti ostatním voláním.

```
$ echo a # b
a
$ echo a \# b
a # b
$ echo "a # b"
a # b
$ echo 'a # b'
a # b
```

Ukázka 3.1: Escapování v shellu

Druhým krokem je odstranění komentářů. V Bashi lze nastavit, aby se v interaktivním shellu komentáře ignorovali. Slouží k tomu proměnná `interactive_comments`, v základním nastavení jsou komentáře zapnuté.

Třetím krokem je rozdělení celého vstupu na příkazy. K příkazům patří přeměňování vstupů a výstupů. Jednoduchý příkaz je buď přiřazení proměnné, nebo název příkazu. Co je příkaz bude řečeno dále. Příkazů v jednom vstupu může být více. Příkazy mohou být za sebou v rouře, odděleny logickým `and` nebo `or`. Příkazy mohou být také odděleny jednoduše pomocí středníku, nebo v případě spouštění ze skriptu i novou řádkou. Stejně může příkazy rozdělovat ampersand, který značí, že se příkaz spustí na pozadí. Některé konstrukty již z principu obsahují více příkazů, například podmínky nebo smyčky.

Čtvrtým krokem je expanze a substituce. TODO: tohle by mohla být samostatná kapitola?

Posledním krokem je samotné spuštění příkazu, popřípadě přiřazení do proměnné. O spuštění příkazů pojednává kapitola 3.2.3

3.2.2 Gramatika shellu

Soubor s gramatikou `bashe`, `parse.y`, má přes 6000 řádek. Chtěl bych zde napsat zjednodušenou gramatiku, která by se dala snadno pochopit (ono to zatím tak složité není).

v `man bash` je gramatika

Popsat jakým způsobem parsuje gramatiku BASH (`yacc`) a jakým to dělají parsery `bashlex` a `bashast`.

3.2.2.1 Projekty parsující shell

<https://github.com/bemeurer/beautysh/blob/master/beautysh/beautysh.py>

3.2.2.2 Bashlex

`Bashlex` je knihovna napsaná ve skriptovacím jazyce Python, která zjednodušeně imituje práci vnitřního parseru `Bashe`. Z velké části se jedná jen o přepsání zdrojových kódů C.

Existují rozdíly oproti parseru v `Bashi`. Za prvé se nespouští žádné příkazy, knihovna umí pouze parsovat. Na rozdíl od parseru `Yacc`, který používá `Bash`, je `Bashlex` reentrantní. Výstupem `Bashlexu` je kompletní abstraktní syntaktický strom.

TODO jak se pracuje s `bashlexem`, protože to používám

3. ANALÝZA A NÁVRH

```
$ python
>>> import bashlex
>>> parts = bashlex.parse('true && cat <(echo $(echo foo))')
>>> for ast in parts:
...     print ast.dump()
ListNode(pos=(0, 31), parts=[
  CommandNode(pos=(0, 4), parts=[
    WordNode(pos=(0, 4), word='true'),
  ]),
  OperatorNode(op='&&', pos=(5, 7)),
  CommandNode(pos=(8, 31), parts=[
    WordNode(pos=(8, 11), word='cat'),
    WordNode(pos=(12, 31), word='<(echo $(echo foo))', parts=[
      ProcesssubstitutionNode(command=
        CommandNode(pos=(14, 30), parts=[
          WordNode(pos=(14, 18), word='echo'),
          WordNode(pos=(19, 30), word='$(echo foo)', parts=[
            CommandsubstitutionNode(command=
              CommandNode(pos=(21, 29), parts=[
                WordNode(pos=(21, 25), word='echo'),
                WordNode(pos=(26, 29), word='foo'),
              ], pos=(19, 30)),
            ],
          ], pos=(12, 31)),
        ],
      ],
    ],
  ],
])
```

Ukázka 3.2: Výstup z knihovny Bashlex

3.2.2.3 Bashast

<https://github.com/neloe/libbash/blob/master/bashast/bashast.g>

3.2.3 Spouštění příkazů

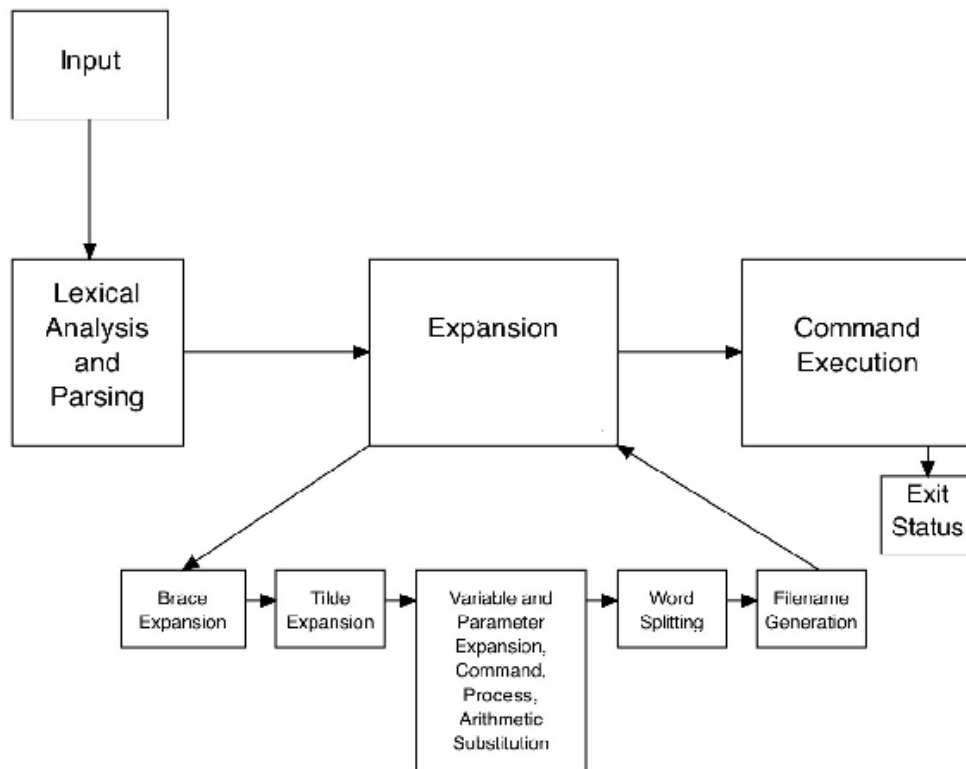
Popsat základní principy jak funguje shell. Popsat procesy v unixu, fork, exec, co všechno se musí stat, aby shell mohl spustit příkaz.

Důležité bude popsat, kdy se spouští debug trap a že můžeme příkazy spouštět přes exec.

3.2.4 Struktura BASHe

Tady je popsáno, jak bash funguje: http://www.aosabook.org/en/bash.html?cm_mc_uid=36339891185014923692137&cm_mc_sid_50200000=1492436376

Bash diagram 3.2.



Obrázek 3.2: Bash diagram

3.3 Debugování shellu

tohle je jeden bod zadání: "Proveďte řešení existujících nástrojů pro statickou analýzu, krokování a hledání chyb v BASH skriptech."

3.3.1 Debugování Bashe

todo: návod na debugování bashe přímo pomocí gdb. (tohle nakonec pro moji práci nebylo potřeba)

3.3.2 Interní nástroje

todo: popsat to, jak nástroje má v sobě bash zabudované v základu

3.3.2.1 Debugovací mód BASHe (jak funguje shopt s extdebug)

shopt s extdebug

3. ANALÝZA A NÁVRH

3.3.2.2 set x, u, v, e

příklady do skriptu

3.3.2.3 PS0, PS4

PS0 bude v novém bashi, my ji proto nebudeme používat, PS4 se vypisuje při debugování. todo: ps0 se nakonec nepovedlo

3.3.3 Externí nástroje

3.3.4 BASH Debugger

todo: popsat jak funguje, co všechno umí, nějaké příklady <http://bashdb.sourceforge.net/>

3.3.5 BashEclipse

BashEclipse je plugin do Eclipse, který umí krokovat v gui eclipse.

<http://unix.stackexchange.com/questions/131491/is-there-a-gui-debugger-for-shell-scripts> <https://sourceforge.net/projects/shelled/> <https://sourceforge.net/projects/basheclipse/>

3.4 Možnosti debugování v interaktivním shellu

Nebyl nalezen žádný nástroj pro debugování interaktivního shellu. todo

3.4.1 Definice debugování interaktivního shellu

todo: nema cenu debugovat jednoduché příkazy, ale složité ano. chceme umět rozkrokovat pipy a subshelly (tedy to co nam to dělá teď)

3.4.2 GNU Readline

GNU Readline umožňuje přemapovat enter tak, abysme mohli spustit příkaz v naší definované funkci. Problémem je, že takto upravený příkaz se uloží do historie. Dalším problémem jsou víceřádkové příkazy, tedy takové, pro jejichž napsání musíme několikrát zmáčknout enter. TODO: ukázka.

3.4.3 Napsání nového REPLu

Zprovoznění základní funkcionality by bylo snadné, vzhledem ke komplexnosti BASHe však téměř nemožné mít stejné chování jako v BASHi.

3.4.4 DEBUG trap

Současné řešení. Při zapnutém extdebug je možné příkazy nepustit a jen evalovat poslední příkaz z historie. TODO: je potřeba popsat základní chování historie (např mezera na začátku příkazu, atd.)

3.5 Statická analýza skriptů

bash ma validaci syntaxe: bash -n, ale to nezjistí skoro žádné chyby.

3.5.1 Check Bashisms

<http://checkbashisms.sourceforge.net/> Některé skripty cheme kvůli rychlosti spouštět v Bourne-Shellu. Protože ten je pouze podmnožina bashe, tento nástroj najde výskyty syntaxe, která je podporovaná pouze v bashi a v Bourne-Shellu nemusí fungovat správně.

3.5.2 Explain Shell

Tato analýza nehledá chyby, ale vyparsuje vložený kód a vyhledá příslušnou dokumentaci v manuálových stránkách. Využívá knihovny bashlex, kterou využijeme v našem debuggeru. <http://www.explainshell.com/>

3.5.3 ShellCheck

todo: tady bych se mohl rozepsat víc. o tom, že to je napsané v haskellu, o tom že existuje databáze věcí, které se ve skriptech hledají a že ke každým je wiki..

Realizace

v zadání je: Navrhněte a implementujte nástroj, který umožní psát uživatelské skripty pro analýzu příkazů a ovlivňování jejich spouštění a vykonávání. = to jsou hooky

Nástroj musí umožňovat krokovat složitější skripty po jednotlivých příkazech. = to je debugger

Pro analýzu spouštěných skriptů využijte vhodný nástroj z řešeršní části. = to je shellcheck

4.1 Nespouštění příkazů

Pro zabránění spouštění používáme DEBUG trap. todo: popsat problémy a řešení

4.2 Implementace debuggeru

todo: napsat o tom, že část je napsána v pythonu a že se celá věc spouští v pre-hooku, takže to je přesně ten čas, kdy se může příkaz debuggovat a originální příkaz se nespustí.

popsat parsování v bashlex a jak se celá věc předává do bashu

4.3 Hooks

Aby byl kód přehledný, kromě bezpečných příkazů je funkcionality rozdělena do hooků, nebo-li modulů, které obsahují kód, který je spuštěn před, nebo i po vykonání příkazu. Kód vykonaný před příkazem může rozhodnout, zda-li má dojít k zabránění vykonání příkazu.

4.3.1 Implementace hooků

todo: napsat o tom, že se hooky pouští automaticky před i po a že je snadné vytvořit vlastní hook

4.3.2 Hooky před vykonáním příkazu

todo: napsat co všechno můžu dělat

4.3.3 Hooky po vykonání příkazu

todo: napsat co všechno můžu dělat

4.3.4 Předvytvořené hooky

todo: napsat o tom, že

4.3.5 Shellcheck hook

todo: tohle je nejlepší věc, kterou celé commash umí. tady by to možná chtělo ukázat případy, kdy se to hodí. taky popsat to, že některé věci jdou snadno zakázat

4.3.6 Bashlex hook

todo: tohle je bashovská část debuggeru

4.3.7 Explain RC hook

4.3.8 Notfound hook

4.4 Bezpečný mód

Bezpečný mód umožňuje dvě základní věci. Tou jednodušší je pouze vypnutí efektu příkazu, který má nějaké destruktivní následky. Složitější varianta dovoluje vrácení do stavu před vykonáním příkazu.

todo: napsat o tom, že po zkoušení jsem dosel k závěru, že je lepší uživateli nejprve ukázat co se děje, co se stane a jaké má možnosti. častokrát už

4.4.1 Bezpečné rm

todo: napsat o tom, jak to je vyřešené a proč. napsat, že `rm -i` i `rm -I` je k ničemu a proč je naše verze lepší popsat, jak funguje freedesktop.org trash a že ukládáme seznam souborů, které byly smazány jedním `rm`, aby se celý spuštěný příkaz `rm` vrátil

4.5 Historie

Popsat jak jsem vyresil ukladani historie prikazu, ukladani vystupu, navratove kody, jak vracet nasledky prikazu do puvodniho stavu. todo: historie taky neni v zadani a moc jsem to nedoresil. teoreticky by na to sel napsat pre a post hook, ktery zaznamena vsechno potrebne a zapise to nekam do souboru. todo: slo by snadno udelat: podobna historie co je ted, jen se ulozi odkud se prikaz spustil

4.6 Instalace

todo: odkaz na github, vypsati sem to je i tam, tedy: tvorba virtualniho stroje a instalace popsati, ze se predpoklada instalace do `/.commash`

Testování

5.0.1 Testování shell skriptů

Popsat jak fungují nektě testovací frameworky: shunit2, roundup

5.1 Automatizované spouštění příkazů

Popsat jak se dají automaticky spouštět příkazy. At už lokálně, nebo vzdalene. Popsat jak rekonstrukci z typescriptu, tak třeba Tcl, Expect.

5.1.1 Expect

todo: jak se pracuje s expectem

Pro testování byl zvolen skriptovací jazyk Expect, který je rozšířením jazyka Tcl. Lze tedy velice snadno otestovat chování Comma-shellu a jeho částí pro různé vstupy.

Testy se skládají ze vstupů a z řetězců, které se očekávají na výstupu. Přes Expect se potom vytvoří nová instance bashe, ve které se automaticky načte Comma-shell a postupně se posílají vstupy a sledují se výstupy pro očekávané řetězce.

Testy jsou rozdělené na menší části, kde každá testuje jednu konkrétní věc. Přes skript lze pustit sérii všech testů.

5.1.2 Testování Debuggeru

5.1.3 Testování Hooks

5.1.4 Testování Bezpečných příkazů

5.1.5 Měření paměťových nároků

Závěr

todo: urcite bych napsal. ze nikdo nic podobneho neudelal a ohlasy byly celkem pozitivni. pokud se toho chytne alespon par lidi a budou hlasit chyby, co je potreba vylepsit, nebo dokonce i prispivat, mohl by z toho byt jednou opravdu vyborny pomocnik

sem napište závěr Vaší práce

Literatura

Seznam použitých zkratek

GUI Graphical user interface

XML Extensible markup language

Obsah přiloženého CD

	readme.txt.....	stručný popis obsahu CD
	exe.....	adresář se spustitelnou formou implementace
	src	
	impl.....	zdrojové kódy implementace
	thesis.....	zdrojová forma práce ve formátu L ^A T _E X
	text	text práce
	thesis.pdf.....	text práce ve formátu PDF
	thesis.ps	text práce ve formátu PS