

## I. IDENTIFIKAČNÍ ÚDAJE

## POSUDEK OPONENTA ZÁVĚREČNÉ PRÁČE

Název práce:	An Environmnet for Evaluation of Robotic Experiments
Jméno autora:	Tomáš Verner
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Oponent práce:	Ing. Vojtěch Vonásek, Ph.D.
Pracoviště oponenta práce:	Katedra kybernetiky

## II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

<b>Zadání</b>	<b>lehčí</b>
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Práce spočívá ve vytvoření skriptu pro načtení výsledků experimentů (tj. textové soubory obsahující jednotlivá měření) a vytvoření tabulky (ve formátu Latex/OpTeX) s výsledky (shrnutím, průměry atd.). Zadání předpokládá tvorbu tabulek pro konkrétní robotické experimenty, výsledek práce nelze použít na generování obecných tabulek (a pokud ano, tak to není demonstrováno).	
<b>Splnění zadání</b>	<b>splněno</b>
<i>Posudte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
Zadání neobsahuje konkrétní kritéria, které má výsledná implementace splňovat, stačí pouze generovat nějaké tabulky. Zadání je splněno.	
<b>Zvolený postup řešení</b>	<b>správný</b>
<i>Posudte, zda student zvolil správný postup nebo metody řešení.</i>	
Student zpracovává textové soubory s výsledky, vytvoří z nich SQL databázi (zde je použit engine SQLite), SQL je použita pro tvorbu agregovaných dat, výpočet statistik, a z té následně tvoří LaTeXovské tabulky dle zadání. Postup je v pořádku.	
<b>Odborná úroveň</b>	<b>D-uspokojivě</b>
<i>Posudte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Výstupem bakalářské práce je jednak vlastní text, dále přiložený datový soubor (ten obsahuje vytvořené skripty v jazyce Julia) a krátký soubor README, který pouze popisuje parametry hlavního programu. Na konci práce je uveden odkaz na neveřejný gitlab repozitář (kam oponent nemá přístup). Za jedinou dostupnou dokumentaci vytvořeného programu lze tedy brát text bakalářské práce. Ta je však psána chaoticky a podle popisu není moc jasné, jak program používat, jaký je „workflow“ (návaznost výpočtů a skriptů), jak program případně upravit pro generování jiných tabulek.	
V práci chybí popis (a vztah) tabulek v SQL databázi, jejich provázání přes primární klíče, datové typy sloupců atd. Nejsou uvedeny dotazy pro tvorbu agregovaných dat. Není popsán způsob práce s neúplnými daty (víme jen, že jsou nahrazeny hodnotou NULL), ale není popsáno, jak to dále ovlivňuje výpočet výsledků a statistik (např. hodnot PDM a PDB).	
Student nesprávně používá termín „parsování“ ve smyslu načítání dat, dále jsou často zaměňovány pojmy „table“ (ve smyslu SQL tabulka a ve smyslu výsledná LaTeXovská tabulka): „During the creation of the database, many tables with different purposes are generated“ - zde jsou asi míněny SQL Tabulky? Nebo naopak program rovnou při načítání vstupních dat tvoří i výsledné LaTeXovské tabulky? Tyto nedostatky zhoršují porozumění textu.	
Není jasné, jestli je možné LaTeXovské tabulky generovat z již existující SQL databáze, nebo se musí před vygenerováním LaTeX tabulek znova vytvářet celá SQL tabulka včetně načítání původních dat?	

Pojem „dictionary“ není jasný: „The files are parsed into a dictionary..“, myslí se tím datová struktura „hash table“ v jazyce Julia, nebo se jedná o „adresář“ v souborovém systému? Každopádně, informace o „dictionary“ je pro uživatele bezvýznamná, neboť důležitým výsledkem načtení dat je uložení do SQL databáze, která jediná slouží v dalším výpočtu, agregacím, atd., a tvorbě tabulek.

Závažným problémem je zaokrouhlování čísel: „The Precision parameter specifies how numeric values should be rounded, if at all, when being stored into the database“.

Student zaokrouhluje vstupní čísla na předem zadaný počet desetinných míst (sekce 3.1) ještě předtím, než jsou uložena do databáze. To nedává smysl, paměťová náročnost SQL databáze se tím nezmenší (za předpokladu, že data jsou uložena v datovém typu REAL). Lepší postup by byl uložit čísla s nejvyšší přesností (tj. takovou, jaká je ve zdrojových datech), provádět výpočty na těchto „přesných“ číslech a zaokrouhlení provést až při výpisu výsledků do LaTeXovských tabulek.

### Formální a jazyková úroveň, rozsah práce

### C - dobře

*Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.*

Práce je psána anglicky, text obsahuje mnoho překlepů (např. již v poděkování je uvedeno „Intelligent and mobile Romotics Group“ namísto správného „Robotics“). Text je asi psán horkou jehlou, což se projevuje na množství překlepů a typografických prohřešků.

- Kapitola „Chapter“ by měla začínat na liché stránce (vpravo), což není vždy dodrženo.
- Student na všechny části textu odkazuje jako na „Chapter X.Y“ místo správného „Section X.Y“.
- Text obsahuje neexistující referenci: „After examining a few tables from theses about robotic experiments, path-finding algorithms in particular ??, the default value has been set to 2.“
- Různé zacházení s číslovkami: „However, each parameter has its default value and the user can therefore give 2 empty files ..“ vs. „The environment was then tested on two libraries..“ (obě věty jsou ze závěru).
- „For the first 4 options, the ones with a 1-line border, there is some resemblance to the Bolden parameter.“ Pojem „Bolden“ není nikde v textu vysvětlen.
- Nastavení pravidel (rovnice 1-5, str. 15, sekce 3.2.5) – text je nesprávně formátován ve formě rovnic.

### Výběr zdrojů, korektnost citací

### F - nedostatečně

*Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.*

Práce neobsahuje přehled existujících řešení. Citováno je pouze sedm zdrojů (hlavně odkazy na jazyk Julia, Tex, a SQLite). Reference [3] ze zadání není citována. Proč nebyl zdroj [3] využit?

V úvodu se dočteme, že „The main goal of the thesis is therefore not to outperform some already existing version..“, čímž je naznačeno, že nějaké řešení sice existuje, ale student se related-work nebude zabývat. Nástrojů, které umožňují vytvářet uživatelem definované tabulky (včetně výstupu do LaTeX tabulek) existuje více. Bylo by vhodné je popsat a argumentovat, proč je třeba psát vlastní systém. Existující nástroje jsou například balíčky Pandas, Koalas, Polars, atd.. viz [1-6]:

[1] <https://ctan.org/pkg/datatool>

[2] <https://www.ctan.org/pkg/pgfplotstable>

[3] <https://pandas.pydata.org/docs/index.html>

[4] <https://datatable.readthedocs.io/en/latest/start/quick-start.html>

[5] <https://pola-rs.github.io/polars-book/user-guide/index.html>

[6] <https://koalas.readthedocs.io/en/latest/index.html>

Pokud práce neobsahuje rozbor problematiky na pokyn vedoucího práce, bylo by dobré to do práce napsat, případně věnovat related-work alespoň několik kratších odstavců.

#### Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Vložte komentář (nepovinné hodnocení).

### III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Text práce je chaotický, není jasné jak programy fungují. Navržené programy jsou psány na míru jednomu konkrétnímu stylu práce s experimentálními daty a nejsou obecně použitelná. Vzhledem k neexistujícímu popisu metod a databázových schémat nelze říci, jestli je program možné generalizovat i pro generování jiných tabulek. Zaokrouhlování čísel není vhodně řešeno. Text obsahuje hodně překlepů, chyb a nevhodné formátování. Student se zřejmě zaměřil více na programování než na psaní textu.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **C - dobře**.

#### Otázky:

- Uveďte schéma SQL databáze.
- Jaký je rozdíl mezi tabulkou „valid“ a „results“? Proč jsou v tabulce „valid“ uložena u nevalidní data, zatímco v tabulce „results“ jen validní? Proč není použita jen jedna tabulka?
- Proč k zaokrouhlování čísel dochází již při ukládání do SQL a nikoliv až při generování výstupních tabulek?
- Proč nebyla provedena rešerže existujících řešení?

Datum: 6.6.2023

Podpis: V. Vonásek