Министерство образования и науки Российской Федерации
Санкт-Петербургский политехнический университет Петра Великого
Институт компьютерных наук и технологий
Высшая школа киберфизических систем и управления

Работа допущена к защите

Директор ВШ КФСУ

_____В.П. Шкодырев

«___»_____2018 г.

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА МАГИСТРА

ИЗВЛЕЧЕНИЕ ЗНАНИЙ И УПРАВЛЕНИЕ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ
KNOWLEDGE EXTRACTION AND CONTROL OF AN UNCERTAIN MODEL
по направлению 09.04.01 – Информатика и вычислительная техника
по образовательной программе
09.04.01_17 Интеллектуальные системы (международная образовательная
программа на иностранном языке)

Выполнил
студент гр. 23541/8   К.В. Блэйк

Руководитель
директор ВШ КФСУ ИКНТ,   В.П. Шкодырев
д.т.н., профессор

Консультант по   Л.А. Киселева
нормоконтролю

Санкт-Петербург

2018

Министерство образования и науки Российской Федерации
Санкт-Петербургский политехнический университет Петра Великого
Институт компьютерных наук и технологий
Высшая школа киберфизических систем и управления

УТВЕРЖДАЮ

Директор ВШ КФСУ

_____В.П. Шкодырев

«__»_____ 2018 г.

ЗАДАНИЕ
по выполнению выпускной квалификационной работы
студенту Кристофер Виллиам Влэйк

1. The Topic of the Master's thesis:

   Knowledge Extraction and Automatic Control

2. Date of Master's submission: 17.09.2018

3. The initial data for the Master's thesis:

   a. Systems and products are developed on a daily basis, all of which require a control process. The development of such a control process often requires extensive analysis and requires domain-specific knowledge. The task is to develop an automatic or semi-automatic process for developing control systems to enable faster-to-market, more capable, and better understood products. Such a system should use as little prior information about the black box system as possible.

4. The content of Master's thesis (list of analyzed questions):

   a. What are the unique data experienced by the black box?

   b. What are the repeating structures within the data?

   c. What are the primitive control functions or mechanisms of the black box?

5. The list of graphical materials (with the exact names of illustrations):

   3.1   Black Box Model
   3.2   Split range

6. Master's thesis advisor: Вячеслав Петрович Шкодырев

7. Date of task issue: 05.02.2018

Master's thesis supervisor   Вячеслав   Петрович Шкодырев

The task is accepted «___» _____20___г.

Student   Кристофер   Виллиам Блэйк

*Print on both sides of one paper sheet

РЕФЕРАТ
На 21 с., 2 рисунков, 5 таблицы, 12 приложениий

ЧЁРНЫЙ ЯЩИК, ИЗВЛЕЧЕНИЕ ЗНАНИЙ, СИСТЕМА УПРАВЛЕНИЯ, МАШИННОЕ ОБУ-ЧЕНИЕ, ДЕРЕВО РЕШЕНИЙ

Был создан адаптивный шаблон для полуавтоматического форматирования магистерской диссертации. Этот шаблон LaTeX предназначен для создания двух версий: одной для Университета Лейбница в Ганновере и одной для Санкт-Петербургского политехнического университета Петра Великого. Пользователь должен только создавать страницы содержимого и определять структуру в главном файле, и это же содержимое будет использоваться в обоих тезисах.

(Автоматический перевод с помощью Google Translate.)

ABSTRACT
21 pages, 2 figures, 5 tables, 12 algorithms

BLACK BOX, KNOWLEDGE EXTRACTION, CONTROL SYSTEM, MACHINE LEARNING, DECISION TREE

An adaptive template has been created for semi-automating the formatting of a master thesis. This LaTeX template serves to create two versions, one for Leibniz University Hannover and one for Peter the Great Saint Petersburg Polytechnic University. A user must only produce the content pages and define the structure in a main file and that same content will be used in both thesis reports.

CONTENT

INTRODUCTION

This document is intended to serve as a template for the dual master's degree program between Leibniz University Hannover and Peter the Great Saint Petersburg Polytechnic University. It includes two "main.tex" files that already include the required formatting rules, which automates most of the work. It also includes examples of tables, equations and figures for the difficult to automate parts.

A full workings of LaTeX will not be described here. However for reference to newer users, the main advantage of LaTeX, is the automation of several tedious items like figure numbering, table numbering, figure placement, tracking references, creating table of contents, and creating the bibliography. Make sure to learn these commands are look through examples to save time!

1 Background

1.1 Problem Statement

Students at the end of a master's degree program must produce a document outlining and describing the full details of a project. The university requires documentation for such projects to relatively similar and thus follow the same formatting rules. Additionally all such papers are expected to have certain content. A template with the correct requirements will aid students in this goal, producing professional quality documents and saving time.

1.2 Objective

This work plans to describe a process for using a template for the documentation of a single master thesis, producing two different versions for two different universities at the same time. It will provide templates as well as examples, outlining the major differences and clarifying the requirements.

- Create a template for both universities.

- Show required sections.

- Provide examples of figures, tables and equations.

1.3 History and Overview

A background of why the thesis topic is relevant and can make a difference.

1.4 Applications

Below are three example usage scenarios for using latex templates. However, only the first will be explored during the development of this work.

1. **Multiple Document Formats** – Sometimes documents required several different versions. Rather than maintain multiple copies fo the same content, content can be seperated from formatting. This enables applications like an abbridged version, normal version, commented version or large-font version.

2. **Parrallel Work** - By seperating the content and formatting into several files, multiple people can contribute independently and at the same time. A master editor can focus on layout while different writers focus on different chapters such as theory or experimental results.

3. **Versioning** – Sections of a book can be quickly and easily be replaced, without effecting the rest of the document. Randomization functions can also be used for selecting random content as well. Examples could be randomly sampling employee reports or randomly selecting 10 questions to create a new unique exam.

1.5 Scope and Limitations

This document is only relevant for the dual master's degree program between Leibniz University Hannover and Peter the Great Saint Petersburg Polytechnic University. Any other usage may have unintended consequences. Also, this template should be verified by your professor at both locations. It provides only a getting started point with no guarantees of formatting requirements.

2 Related Work

A major requirement of all research work is to discuss existing work and the discovered gaps in knowledge or capability. After establishing the prior work, this naturally leads into discussing what the new work (that you produced) is contributing.

3 Guidelines

The two "main.tex" files of this template includes 6 areas. As a user, it is not recommended to modify "Formatting Rules" or "Generic Setup", as these define the rules of the template and make it work properly.

The content of the thesis is specified in the "Content" section at the end of the "main.tex" file. The front matter and back matter should not need modified and will create the table of contents, table of figures, and bibliography. The main matter will be the most modified section and will specify the chapters, sections and subsections.

1. Name and Title - Specifies the title page and authors/professors.

2. Formatting Rules - Defines rules for the university.

3. Generic Setup - Defines rules for general latex functioning. Necessary packages are included here.

4. Packages - Allows installing of additional packages for the thesis.

5. New Commands - Allows specifying shortcuts for the thesis.

6. Content - Defines the structure and thesis sections.

LUH Requirements

Having spoken to a few professors, there does not appear to be a university standard. Some departments are different and have there own. However, having looked at several example thesis document, the developed template seems to match several of those.

SPbPU Requirements

A copy of the original formatting rules document [1] in russian and english (google translated) is available in the reference folder of this repository.

3.1 SPbPU Paper Format

Table 3.1 – SPbPU General Paper Format

Name	Requirement
Paper Size	A4
Font	Times New Roman or Arial
Font Size	14pt
Line Spacing Interval	1.5
Margins	left:30mm, right:10mm, top:20mm, bottom:20mm
Printing	One-sided
Page Numbering	Arabic numerals
Page Number Location	top-right, 10mm from right edge

3.2 SPbPU Structure

The title page and task-list page are provided from the department as .doc files. Simply open them up, edit them, and save them as PDFs with the same name. The template will automatically load them into the correeect locations.

1. Title Page - Provided by SPbPU as a seperate .doc file.

2. Task Page - Provided by SPbPU as a seperate .doc file. If necessary, print on both sides.

3. Abstract - in English and Russian. An example is available on page 43 of the official document. [1]

4. Content - The table of contents. This <u>must</u> physically be page 4, but not have a page number. An example is available on pages 44/45 of the official document. [1]

5. Introduction

6. Main Content of Thesis

7. Conclusion

8. References

9. Appendices

4 Examples

4.1 Figures

Figures should automatically be adjusted by the template. Table 4.1 shows the important rules and Fig. 4.1 provides an example.

Table 4.1 – Figure format

Item	LUH	SPbPU
Title Font	Normal	Italic
Title Location	Below	Below
Title Alignment	Centered	Centered
Number	7.3	7.3
Seperator	:	-

$$\text{Input Space} \Rightarrow \boxed{\text{Black Box}} \rightarrow \text{Output Space}$$

Input Space
$X \in I^d$

Black Box

Output Space
$Y \in O^k$

Function Space
$$F = \{f_1(X), f_2(X), \dots f_l(X)\}$$
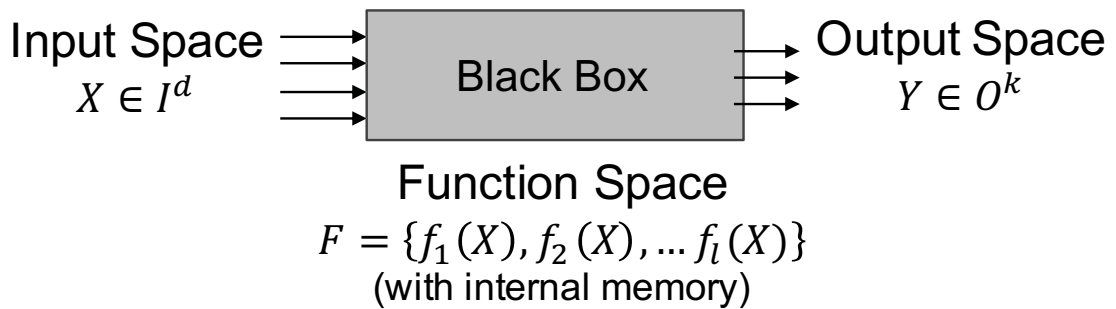(with internal memory)

Figure 4.1 – Example figure

4.2 Tables

Tables should automatically be adjusted by the template. Table 4.2 shows the important rules and Table 4.3 provides an example.

Table 4.2 – Figure format

Item	LUH	SPbPU
Title Font	Normal	Italic
Title Location	Above	Above
Title Alignment	Centered	Left
Number	7.3	7.3
Seperator	:	-

Table 4.3 – Example instances of knowledge with descriptions

Instance	Knowledge Content	Description
c_1	$(x_1 = 0.0)$	Input x_1 has a value 0.0.
c_2	$(x_1 = 1.0)$	Input x_1 has a value 1.0.
c_3	$(x_2 = 0.0)$	Input x_2 has a value 0.0.
c_4	$(x_2 = 1.0)$	Input x_2 has a value 1.0.
c_5	$(y_1 = 0.0)$	Output y_1 has a value 0.0
c_6	$(y_1 = 1.0)$	Output y_1 has a value 1.0
c_7	(c_1, c_3)	$x_1 = 0$ and x_2=0 simultaneously.
c_8	(c_1, c_4)	$x_1 = 0$ and x_2=1 simultaneously.
c_9	(c_2, c_3)	$x_1 = 1$ and x_2=0 simultaneously.
c_{10}	(c_2, c_4)	$x_1 = 1$ and x_2=1 simultaneously.
c_{15}	$(c_{14}; 3c_2; c_{13})$	x_1 switched from 0 to 1, held for $3t$ then switched to 0.

4.3 Equations

Equations are partly automatic. The number scheme is automatic, but the commas and period must be manually added/removed for the two versions. Table 4.4 shows the important rules and the below equations provide an example.

Table 4.4 – Figure format

Item	LUH	SPbPU
Suffix	None	commas ending with period
Eq. Location	Centered	Centered
Number Location	Right	Right
Number Style	(7.3)	(7-3)

$$N = \sum 1_t \tag{4-1}$$

$$Sum = \sum v_t \tag{4-2}$$

$$SqSum = \sum v_t{}^2 \tag{4-3}$$

$$\mu = Sum/N \tag{4-4}$$

$$\sigma = SqSum - 2N\mu + N\mu^2 \tag{4-5}$$

$$N = \sum 1_t, \tag{4-6}$$

$$Sum = \sum v_t, \tag{4-7}$$

$$SqSum = \sum v_t{}^2, \tag{4-8}$$

$$\mu = Sum/N, \tag{4-9}$$

$$\sigma = SqSum - 2N\mu + N\mu^2. \tag{4-10}$$

4.4 Algorithms

Algorithms are not automatic. LUH prefers the algorithmic style while SPbPU requires figures. Below are examples of both.

Note: The caption of the below figure has been overridden to follow the SPbPU format.

Warning: This changes the list of figures/algorithms as well as the counts.

Algorithm 1 LUH Algorithm Example

1: //Global variable
2: $R \leftarrow rangesList$
3: **procedure** SPLITRANGE($r, value$)
4:     //Create new ranges
5:     $rLow \leftarrow range(r.Low, value)$
6:     $rHigh \leftarrow range(value, r.High)$
7:     //Remove old range and add new ranges
8:     $R.Remove(r)$
9:     $R.Add(rLow)$
10:     $R.Add(rHigh)$
11: **end procedure**

1: //Global variable
2: $R \leftarrow rangesList$
3: **procedure** SPLITRANGE($r, value$)
4:     //Create new ranges
5:     $rLow \leftarrow range(r.Low, value)$
6:     $rHigh \leftarrow range(value, r.High)$
7:     //Remove old range and add new ranges
8:     $R.Remove(r)$
9:     $R.Add(rLow)$
10:     $R.Add(rHigh)$
11: **end procedure**

Figure X.X - SPbPU Algorithm Example

5 Thesis Main Content

Below are example sections for defining the main content. They are not necessarily required.

5.1 Part 1

5.1.1 Content Subsection

5.1.2 Content Subsection

5.2 Part 2

5.2.1 Content Subsection

5.2.2 Content Subsection

6 Experiments Setup

Below are example sections for defining the experimental setup. They are not necessarily required.

6.1 Datasets

6.1.1 Datasets Subsection

6.1.2 Datasets Subsection

6.2 Equipment

6.2.1 Equipment Subsection

6.2.2 Equipment Subsection

6.3 Processes

6.3.1 Processes Subsection

6.3.2 Processes Subsection

7 Results

Below are example sections for defining the experimental results. They are not necessarily required.

7.1 Results Section

7.1.1 Results Subsection

7.1.2 results Subsection

7.2 Results Section

7.2.1 Results Subsection

7.2.2 results Subsection

8 Future Considerations

8.1 Topic Points

It is very likely that the new thesis will solve all problems and have no room for development or improvements. As such, these ideas should be documented and briefly introduced here to encourage others to build upon the new work described in the thesis. The formatting here is only for example.

Below is a list of topics that became apparent during development and testing of this work. They are not meant to be extensive, but rather to provide insight about any future developments. Entries in **bold** have proposed extensions in section 8.2.

1. **Algorithms** – SPbPU requires algorithms to be figures and generally other scientific writers prefer the algorithmic format. This is not part of the automatic formatting.

2. **Equations** – SPbPU requires equations to be the end of a sentence and be seperated by commas. LUH does not have this requirement. This is not part of the automatic formatting.

8.2 Proposed Extensions

1. **Algorithms** - SPbPU requires algorithms to be displayed as figures. Normally algorithms are displayed via the algorithm enviornment. A new environment could be created for changing the location of the figure name and number automatically as well as placing the content in a box.

2. **Equations** - SPbPU requires equations to be formatted as the end of a sentence with commas and a period. A new enviornment could be created with an option to turn these commas and period on/off. This would enable all this format to be adjust from the template, rather than in the content itself.

CONCLUSIONS

An adaptive template has been created for semi-automating the formatting of a master thesis. This LaTeX template serves to create two versions, one for Leibniz University Hannover and one for Peter the Great Saint Petersburg Polytechnic University.

A main.tex file is created for both universities, which defines the standard formatting rules for each university and then references various content files with minimal formatting. When compiled two seperated documents are produced from the same content.

To support further development, this work is available at `https://github.com/chriswblake/ThesisTemplate_LUH_SPbPU`

REVIEWERS

1   Peter the Great St. Petersburg Polytechnic University. Procedure for conducting state final certification. Technical report, 2019.