BSPro - A First Bachelor Semester Project in BiCS-land

Monday 29th November, 2021 - 12:11

Stefan Capitanescu
University of Luxembourg
Email: stefan.capitanescu.001@student.uni.lu

This report has been produced under the supervision of:
Nicolas Guelfi
University of Luxembourg
Email: nicolas.Guelfi@uni.lu

Abstract

This document is a template for the scientific and technical (S&T for short) report that is to be delivered by any BiCS student at the end of each Bachelor Semester Project (BSP). The Latex source files are available at: ***https://github.com/nicolasguelfi/lu.uni.course.bics.global***

This template is to be used using the Latex document preparation system or using any document preparation system. The whole document should be in between 6000 to 8000 words [1] *(excluding the annexes) and the proportions must be preserved. The other documents to be delivered (summaries, ...) should have their format adapted from this template.*

A tutor (or any person having contributed to the BSP work) is not a co-author per se for a student's work. It is possible to exploit a BSP report to produce a scientific and technical publication. In this case, the authors list has to be discussed and agreed with the concerned parties.

1. Plagiarism statement

This 350 words section without this first paragraph must be included in the submitted report and placed after the conclusion. This section is not counting in the total words quantity.

I declare that I am aware of the following facts:

- As a student at the University of Luxembourg I must respect the rules of intellectual honesty, in particular not to resort to plagiarism, fraud or any other method that is illegal or contrary to scientific integrity.
- My report will be checked for plagiarism and if the plagiarism check is positive, an internal procedure will be started by my tutor. I am advised to request a pre-check by my tutor to avoid any issue.
- As declared in the assessment procedure of the University of Luxembourg, plagiarism is committed whenever the source of information used in an assignment, research report, paper or otherwise published/circulated piece of work is not properly acknowledged. In other words, plagiarism is the passing off as one's own the words, ideas or work of another person, without attribution to the author. The omission of such proper acknowledgement amounts to claiming authorship for the work of another person. Plagiarism is committed regardless of the language of the original work used. Plagiarism can be deliberate or accidental. Instances of plagiarism include, but are not limited to:

1) Not putting quotation marks around a quote from another person's work
2) Pretending to paraphrase while in fact quoting
3) Citing incorrectly or incompletely
4) Failing to cite the source of a quoted or paraphrased work
5) Copying/reproducing sections of another person's work without acknowledging the source
6) Paraphrasing another person's work without acknowledging the source
7) Having another person write/author a work for oneself and submitting/publishing it (with permission, with or without compensation) in one's own name ('ghost-writing')
8) Using another person's unpublished work without attribution and permission ('stealing')
9) Presenting a piece of work as one's own that contains a high proportion of quoted/copied or paraphrased text (images, graphs, etc.), even if adequately referenced

Auto- or self-plagiarism, that is the reproduction of (portions of a) text previously written by the author without citing that text, i.e. passing previously authored text as new, may be regarded as fraud if deemed sufficiently severe.

1. i.e. approximately 12 to 16 pages double columns excluding the Plagiarism Statement

2. Introduction ($\pm$ 5% of total words)

This paper presents the bachelor semester project made by Motivated Student together with Motivated Tutor as his motivated tutor. It presents the scientific and technical dimensions of the work done. All the words written here have been newly created by the authors and if some sequence of words or any graphic information created by others are included then it is explicitly indicated the original reference to the work reused.

This report separates explicitly the scientific work from the technical one. In deed each BSP must cover those two dimensions with a constrained balance (cf. [BiCS(2021)]). Thus it is up to the Motivated Tutor and Motivated Student to ensure that the deliverables belonging to each dimension are clearly stated. As an example, a project whose title would be "PLAYTOUCH - A multi-user game for multi-touch devices" could define the following deliverables:

- Possible scientific [Armstrong and Green(2017)] deliverables:
  - What are concurrency models and how are they implemented?
  - How is measured ergonomics in human-computer interaction?
  - How to model the concurrency of a multi-touch devices?
  - Can PLAYTOUCH enter in a blocking state?
  - How to model the design of PLAYTOUCH?
- Possible Technical deliverables:
  - PLAYTOUCH Implementation
  - PLAYTOUCH Tests implementation
  - Hardware end system configuration for PLAYTOUCH

The length of the report should be from 6000 to 8000 words excluding images and annexes. The sections presenting the technical and scientific deliverables represent $\pm$ 80% of total words of the report.

3. Project description ($\pm$ 10% of total words)

3.1. Domains

3.1.1. Scientific . Provide a short description of the scientific domain(s) in which the project is being made.

3.1.2. Technical. Provide a short description of the technical domain(s) in which the project is being made.

3.2. Targeted Deliverables

3.2.1. Scientific deliverables. Provide a synthetic and abstract description of the scientific deliverables that have been produced. Each BSP must contain some work done according to the principles of the scientific method. It basically means that you should define at least one question related to the knowledge domain of your BSP and follow part of the scientific method process to answer this question. The description of the work done to answer this question is a scientific deliverable.

Other examples of question could be:

- Is Python an adequate language for concurrent programs?
- How can we measure the ergonomics of a graphical user interface?
- How can we ensure that a program will not fail?

An answer to such question should be the result of applying partly or totally the scientific method according to its standard definition which can be found in the literature.

As you can see in this template, the scientific deliverable is entirely separated from the technical deliverable. It the default case it addresses a question closely related to the technical deliverable.

3.2.2. Technical deliverables. For this bachelor semester project, the target to reach is to create a software that could export an audio file of none, one or multiple songs which are defined by different criteria.

The program was planned to be written in python, a world-known programming language open-source which is widely used for its versatility and various libraries and modules. For this project, the musical module "pydub" had to be used to import and export songs with only the path of the audio file searched. Even if this module has more interesting features they normally will not be used for the code. A graphical user interface was expected to be done for the BSP with Tkinter, an python blinding easy to use and with a great documentation all over the internet. Provide a synthetic and abstract description of the technical deliverables that were targeted to be produced. A technical deliverable in this report is the description of a product build by the student using software or hardware technologies.

4. Pre-requisites ($[5\%..10\%]$ of total words)

Describe in these sections the main scientific and technical knowledge that is required to be known by you before starting the project. Do not describe in details this knowledge but only abstractly. All the content of this section shall not be used, even partially, in the deliverable sections. It is important not to include in this section all the knowledge you have been obliged to acquire in order to produce the deliverable. It should only state the knowledge the student possessed before starting the project and that was mandatory to possess to be capable to produce the deliverables. It explicitly defines the technical and scientific pre-condition for the project. It is also useful to avoid project failures due to over or under complex subjects.

4.1. Scientific pre-requisites

4.2. Technical pre-requisites

5. A Scientific Deliverable 1

For each scientific deliverable targeted in section 3
2 provide a full section with all the subsections described

below.

5.1. Requirements (± 15% of section's words)

This scientific deliverable will cover different questions but all related to the main question which is "How can we evaluate the similarities between two instrumentals tracks". All the work in the production section will divided into parts of different questions we will answer to and they are all related one to each other to have an overview of the main scientific question.

5.2. Design (± 30% of section's words)

Provide the necessary and most useful explanations on how those deliverables have been produced.

5.3. Production (± 40% of section's words)

Since the starting of the millennium, engineers and scientists have been working for the past two decades on the music recognition. Constantly evolving, the different project became more and more efficient and quicker. The most known software for music recognition, Shazam, develloped by Shazam Entertainment Limited, is using a method to recognize songs called the "fingerprint algorithm" with the help of considerable databases relying on. This will be the base of our research to answer the question previously quoted, "How can we evaluate the similarities between two instrumentals tracks" ? But in a first place, let's answer the question : "How Shazam and other recognition software work" ? In a way to analyse sound and get results, the software needs an input. Fortunately, nearly all the computers and mobile devices have a microphone input that can be used. The method of recognition a song by only a degraded sample is called the "audio fingerprint" and it consists grossly of analysing from a point in time, witch graphically represent the frequency of a musical note or from a rhythm in time. Those points are replaced on a time-frequency graphic that can be visualized. (insert picture from main doc). Those points are replaced and analysed in a way to see if it correspond to a song or if it has a lot of same points. From one random point the algorithm analyses the This methods is called the "pattern matching" method. This is possible due to enormous databases

5.4. Assessment (± 15% of section's words)

Provide any objective elements to assess that your deliverables do or do not satisfy the requirements described above.

6. A Technical Deliverable 1

For each technical deliverable targeted in section 3
2 provide a full section with all the subsections described below. The cumulative volume of all deliverable sections represents 75% of the paper's volume in words. Volumes below are indicated relative the the section.

6.1. Requirements (± 15% of section's words)

This BSP will be mainly rely on python programming language. Python is currently on of the most known programming language for it simplicity for the multiple extensions possible and widely supported on different operating systems. The technical part is partly based on audio manipulation and the library pydub is going to be used for this project. This library is not already implement in modern IDE because it is an external library. This library will be used to import song from their path, concatenate songs, one after the others and finally to export the final result by giving a name to the file we want to export, the audio file extension and the bitrate. I will add more if there is any more tool the project will be needed

6.2. Design (± 30% of section's words)

Before starting to code the actual program, a mock up was made using Miro, a free website and application whose main functionality is to let the user create a mock up with tools integrated. This platform is also used for meetings, brainstorming with other people and to manage and plan for a work team. Here is the realised mock up for this project :
(insert the mockup)
This platform was advised by the tutor of this project for the intuitive interface, the variety of tools proposed, and the convincing result a user can achieve without using it before. The realised mock up used only a few functionalities that were proposed because it was not an important part of this project and it has been used to show only how the program would work and ,if achieved, a design for the software produced. After that, this mock up is used as a template to remind the main function that the program offers and how it operates.

The actual program runs in python, an open source programming language that offers a diversity of very useful libraries when it comes to manipulates different type of files or creating a program that is not only based on the basic functionalities of python. For manipulating audio files there are several libraries that can be used, but for the purpose of this Bachelor Semester Project the simplicity and capability of the Pydub library corresponded to the program needs. Published firstly in 2011, Pydub created by James Robert has been upgraded and is still upgraded nowadays, for the purpose of this project the functionalities of importing, exporting and concatenating songs. Using this library has multiples advantages. Firstly, it is one of the most used libraries for audio file manipulations, it has a decent amount of documentation all over the internet and the commends are really intuitive and does corporate well with the python programming language.

To insert the data about the audio files containing tracks, we will use a simple text file. with on each line having the different information about the songs separated by a comma. This file is after interpreted when the user launches the program. This technique is a very good idea, instead having every description of every songs in the real python program and make it less readable, to have an external text file has

several advantages and the open built-in function from python is clear and is an old functionality. Even thought the actual code is not complex and does not have a great length, it remains better to keep data in another file to not disturb the main program.

The program uses class to define criteria the project will rely on classes.

6.3. Production (± 40% of section's words)

cf. section 6 applied to the technical deliverable

6.4. Assessment (± 15% of section's words)

cf. section 6 applied to the technical deliverable

Acknowledgment

The authors would like to thank the BiCS management and education team for the amazing work done.

7. Conclusion

The conclusion goes here.

References

[BiCS(2021)] BiCS Bachelor Semester Project Report Template. https://github.com/nicolasguelfi/lu.uni.course.bics.global University of Luxembourg, BiCS - Bachelor in Computer Science (2021).

[BiCS(2021)] Bachelor in Computer Science: BiCS Semester Projects Reference Document. Technical report, University of Luxembourg (2021)

[Armstrong and Green(2017)] J Scott Armstrong and Kesten C Green. Guidelines for science: Evidence and checklists. *Scholarly Commons*, pages 1–24, 2017. https://repository.upenn.edu/marketing_papers/181/

8. Appendix

All images and additional material go there.