Presentation PI2 Project: DicMath

Brief description:

Our project aims to design a software that allows blind students to dictate mathematical equations and receive vocal feedback to solve them. The software works on computers and all commands can be executed through keyboard or voice input, eliminating the need for a mouse. The main users of the software are blind students (up to high school level), so it only provides basic mathematical operations such as addition, subtraction, multiplication, division, power, and functions. Our software implements both input and output of equations - the user dictates the equation, and the software verbally returns the understood equation. Given the uncertainty of the input equation, the software will emit a sound signal when the equation is completed, repeating the equation. Given the available resources in the library, we implemented dictation in as many languages as possible, including English and French.

Team Number: n°101

Project Partner: Microsoft France

Project Number: n°001

Team Member	Project Role (if defined)
Annabel MERCERON	Referent
Amine MOUSSA	Member
Benjamin TOUBIANA	Member
Yuwen CAI	Member

The Project Closure Report contains key descriptive information about the project.

As the last document written on the project, it analyzes the outcome of the project and the process by which that outcome was produced.

Its purpose is twofold: to ensure that the closure activities are carried out properly and to facilitate the transfer of experience (or the transfer of experience if the project is not completed or if a follow-up to the project is envisaged) to the customer's organization.

Table des matières

1 Objectives Set versus Results Achieved

1.1 Project Initial Objectives

Doing **math** is fun, but it has always been a tricky task that requires concentration. In this context, **visually impaired students** cannot maximize that focus if they spend their time finding effective ways to do math around their disability. This is where we come in to solve this problem.

At the kickoff meeting with our Microsoft partner Mr. Philip Trottin, our goal was to **develop an AI-powered software that helps visually impaired students do math using their voices**. The software should allow students to dictate math equations (by voice) and navigate through those equations to modify them. So it should be **accessible to our clients and easy to use** to avoid misleading them.

From day 1, **we devided the tasks/modules** based on our qualities and weaknesses but planned that everybody must be involved in all the tasks.

Team Repartition

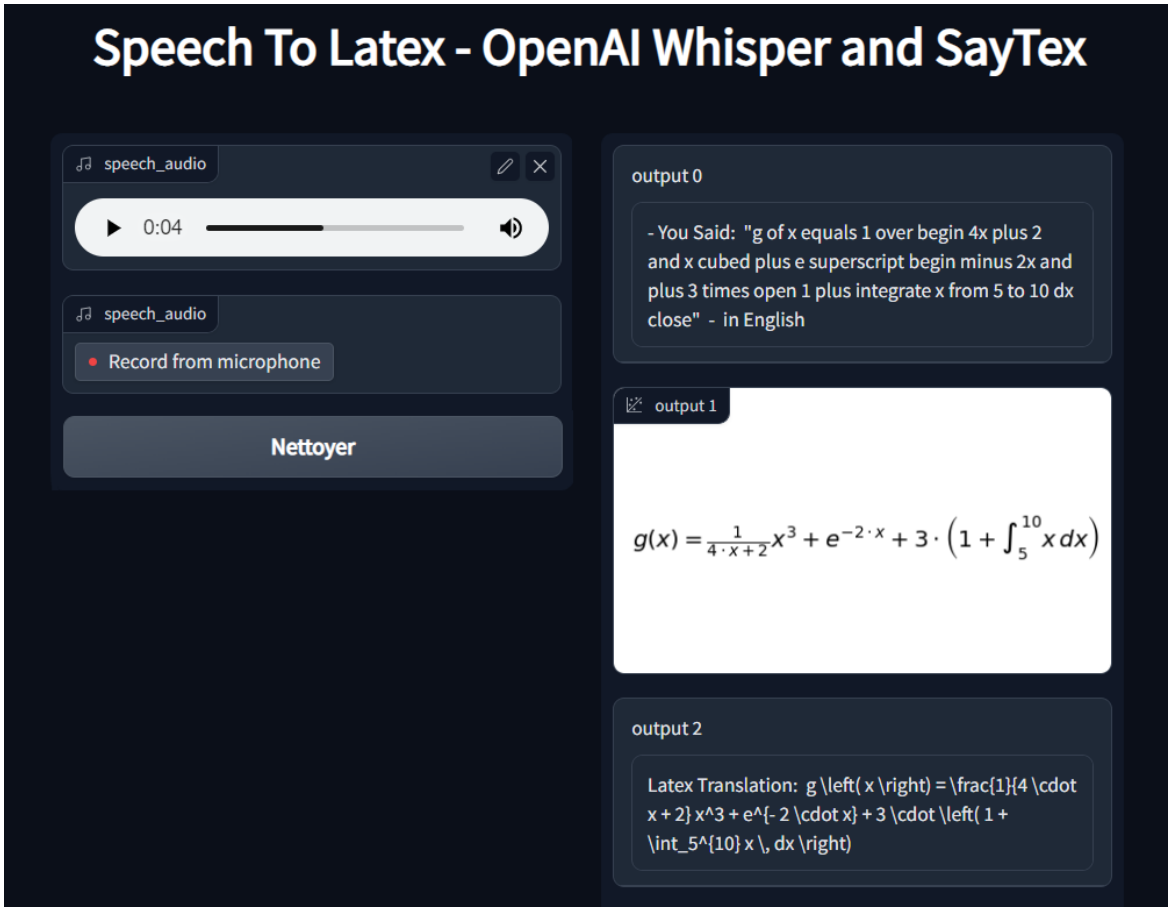
Team Member	Task
Amine MOUSSA	- Speech-To-Text (OpenAI Whisper, Azure Cognitive Services) - Text-To-Latex - Graphical User Interface (GUI)
Annabel MERCERON	- Text-To-Speech (Azure Cognitive Services) - Graphical User Interface (GUI)
Benjamin TOUBIANA	Equations' Navigation
Yuwen CAI	Documentation

Throughout the project, **we didn't** really **vary our main objectives**. We just simplified the "Equation Navigation" module to be able to implement a better GUI (graphical user interface) with more efficient AI services.

1.2 Results Achieved

A few days after the kick-off, we were lucky since the company OpenAI released a new powerful **Speech-To-Text tool named Whisper**. So a team member started implementing it as well as other state-of-the-art services.

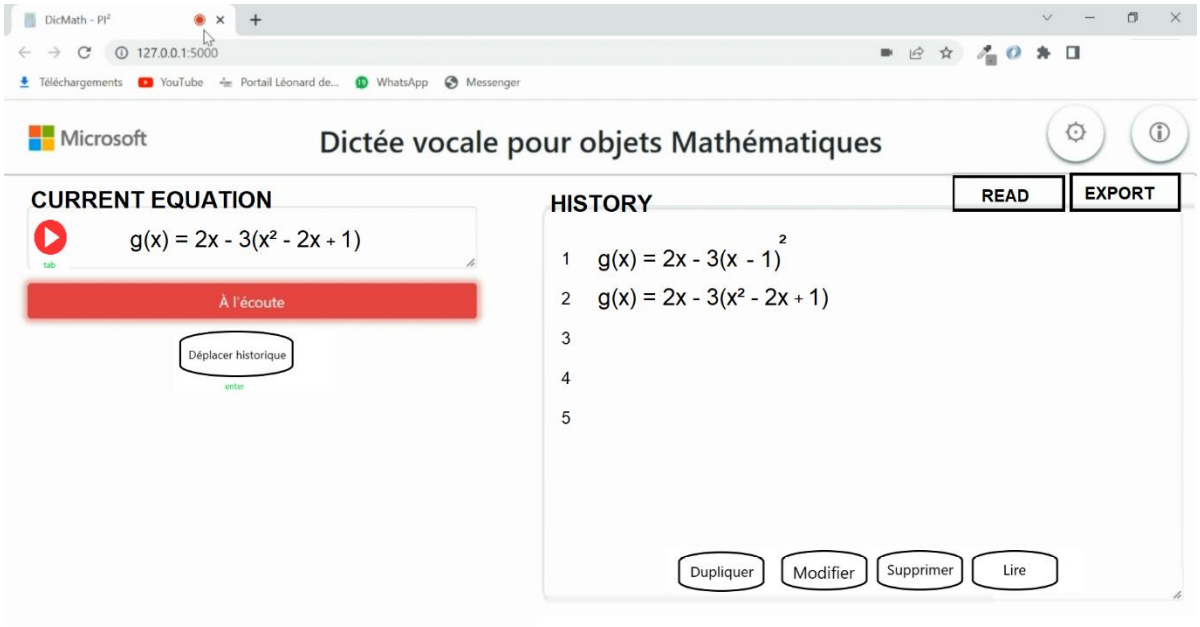
→ The first result was a python-based GUI that allows the user to dictate a mathematical equation through the microphone, and using AI, the GUI displays the speech transcribed, the equation in humanly readable format, and the equation in Latex format.



→ In the other hand, a second member used the Microsoft Azure Cognitive Services to implement a **human like Text-To-Speech bot**. The code translates the newly generated Latex equations to a text that'll be read for the future user (students).

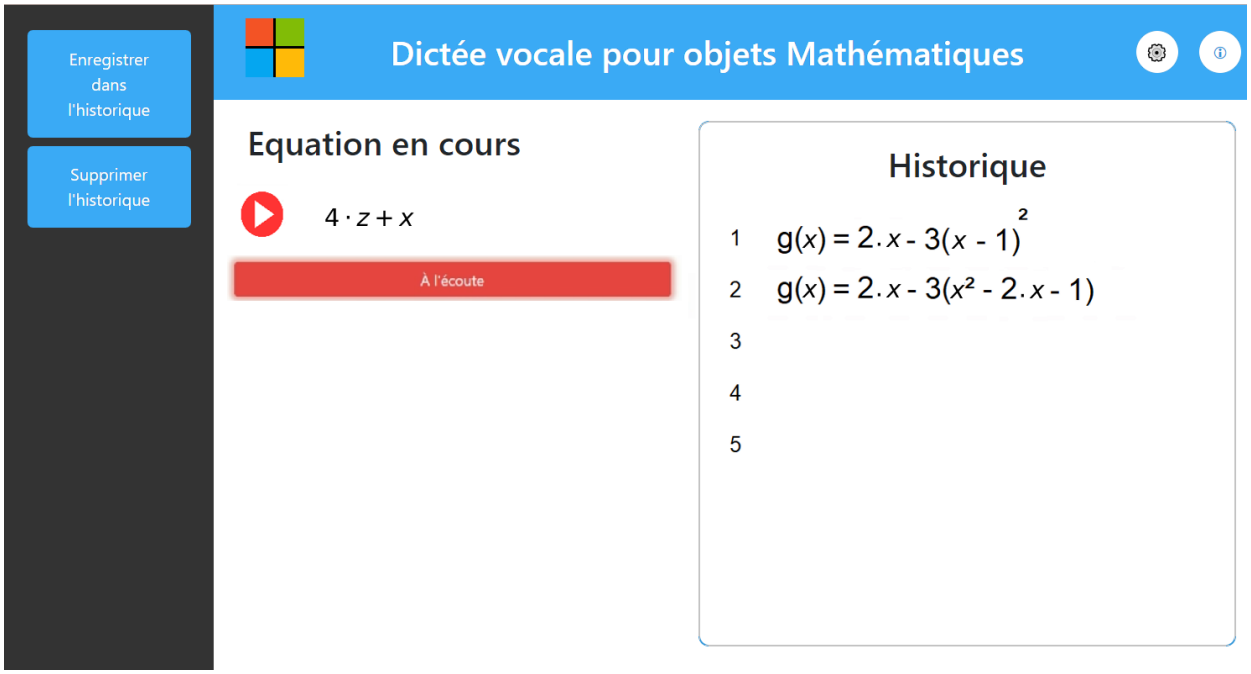
→ Some of the remaining members started looking for efficient ways to develop **an accessible web page** (HTML, CSS, JS) for the software which must suit visually impaired students.

At some point where almost **all the modules were ready, we merged them into 1 project**. The web page was linked to the Python codes that use Artificial Intelligence services.



The **feedback** of the Microsoft partner was always **positive and encouraging**. He helped us reach our objectives with quality advice.

→ After discussing our prototype with him, he was thrilled about the efficiency of the product. He gave us suggestions concerning the GUI to make it more accessible. His expertise in the field of Accessibility was a strong support to this project (especially concerning the client endpoint). **So here is the final product's main page.**



The only problem we encountered was the difficulty finding **common free time** to work together sometimes, especially when we needed **multi-disciplinary profiles**.

1.3 List of Deliverables

In this project, there are 3 main deliverables (all accepted by the partner):

- ✓ **The GitHub Repository:** Python, HTML, CSS, JS codes which interoperate in the final GUI.
- ✓ **A Documentation**: complete tutorial of the use of the product (also included in GUI)
- ✓ **A flyer:** A1 flyer to describe our project's objective and functionalities.

2 Methodological review

This project can be divided into two important parts, which will serve to distribute the tasks: the front-end part and the back-end part.

The front part includes the design of the human-computer interaction interface which means the visual model, through which the functional code of this last part is integrated and realized.

The back-end part includes all the programming code: dictating mathematical equations bidirectional which consists the text to speech and speech to text, setting up data structures for storing the equations, creating and improving artificial intelligence modules of voice synthesis sufficiently deterministic for the visually impaired. Professionals can use the software with complete peace of mind, navigation mode in the equation with various aids (shortcuts, voice feedback, etc.), multilingual translation of the application.

After completing the work of these two parts, realize the assembly to achieve a complete program with the required functions. And draft a manual for potential users or the possible further developer of the project.

On the first step of the project, the student SB is responsible for the design of the human-computer interface, while others focus on the problem of code programming. And finally, all the group participated in discussing the final design of the visual interface in order to get a better vision, and all focused on the python code to build the program.

Whoever it is, we do our best and use many resources to learn and program our bots without a lot of advanced coding knowledge.

In order to keep up with our partners' needs, we schedule several meetings to discuss our progress and write a report at the end of each weekly group meeting to have a clear picture of: work done or to be done. The problem we encountered and the solution we plan to apply to correct it.

3 Risk Management

This project is a complete project that traces the steps needed to make an app from start to finish. As a result, difficult spots that may slow down the progress of the project may be discovered.

3.1 Navigation

While some libraries allow a better understanding of certain parts, there is no guarantee of being able to accurately detect the different words required for comfortable navigation. It will then be necessary to review the operation of the interface to establish more keyboard shortcuts, in order to keep the application usable by the majority of the visually impaired.

3.2 Translation and faithful representation in LaTex

Now we use the libraries that use Speech-to-Text in order to make our task easier when we want to generate Mathematics in Latex. Unfortunately, these libraries are almost exclusively in English, which adds a difficulty for direct translation of a student's voice in English text to use as input, while retaining the exact mathematical meaning of what he wanted to express.

In this case, we have not found the solution yet, we review the rest of the application so that it works in English, and to add supports in other languages as additional features in its final phase.

3.3 Data storage for navigation

While it is possible to record and display equations quite faithfully to the voice input, it is more complicated to know how to easily access them once the recording is done. So, the idea is to simplify the problem as much as possible by using an "en bloc" separation of each equation, itself placed in a data structure which is compatible with the notions of 'previous and following' equations. If it turns out that these solutions are too complicated to implement, it will be necessary to rely on list and index manipulation.

3.4 Facilitation of vocabulary to use

Finally, as students without disabilities, we only have experience with software like Siri and Cortana to give effective voice commands, but we have no insight into how equation solving works.

mathematics for visually impaired students, it is then complicated to put oneself in their shoes and understand their needs without reference. To overcome this problem, it is planned to make an appointment with the INJA (national institute for young blind people) to better understand their need and know precisely which voice commands will be given priority in our application.

3.5 Design of Human-Computer Interface

Considering the intention to meet the needs of the visually impaired, and there is no opportunity to make an appointment with INJA to understand the user's needs initially, we need to understand the user's needs by empathizing with the user's vision problems, taking into account the use of a simple but colorful interface, and all the functions need to be represented by relevant text and buttons, and all functions need to be realized through keyboard input.

Due to the lack of direct contact and deep communication with possible users' group, we achieved the purpose of side-by-side understanding of requirements through meetings with collaborators and got relevant suggestions. The final programming implementation of the human-computer interaction interface will be implemented by IIM. After we get in touch with them, we will share the preliminary design drawings to facilitate their implementation.

4 Constraints

4.1 Code

What we can do is to use python on our own computer, which means that the programming environment is preserved, and all libraries have been imported. For subsequent use, the import process needs to be preserved to avoid code problems caused by environmental problems.

4.2 Aesthetics and Ergonomics

The visual effect design of the human-computer interaction interface achieves aesthetics, and all designs must conform to ergonomics. The convenience for the visually impaired is an important factor that we need to consider additionally.

4.3 Sustainability and usability

Since no additional components are required, the code can be implemented on any computer, so through subsequent in-depth research and improvement of the existing functions of the project, as well as the development and construction of other functions, this project is sustainable And it can be improved through user feedback after release.

4.4 Marketability

The target users of the program in this project are the visually impaired, and people who need to dictate mathematical equations at or below the bac level, so this software can be recommended to relevant institutions or individuals.

4.5 Cost

Even if a partner is willing to fund the account we need to use, we want to keep a reasonable budget and try to limit useless expenses.

5 Technical Review

In terms of resources, most of, if not all the effort that was put into the project ended up being finished in time for the result. However, some of it did not make it to the final cut.

For example, the case of navigation within an equation was supposed to work through a system of blocks and subblocks.
In short, an equation was an ensemble of members (left, right), separated by terms and identifiable thanks to additive operations. Cutting an equation this way would have allowed for a user to change small parts of the equation through voice after its initial inception.
The code to do this exists and is functional. It was made and took a significative amount of time to implement.
But because of the shrinking deadline, and the undecidedness on how to handle the case of small equation changes through the user interface, it was deemed an optional feature to add once the whole project is finished.

In the end, it remains the one thing that had effort put into that did not make it to the final version of DicMath.
It could be argued that if we hadn't worked on this in the first place, workload would have been re-allocated to improve on the features that did stay.

Otherwise, because we imagined our product as an interface first that would host multiple modules/features, all input work was transferred to immediate output results, with little to no hiccup.

6 Engineering standards

Your project does not require any specific certifications.

7 “Post-project” tasks

There are still certain actions and key deliverables to be carried out to guarantee the effective transmission of the project if it is not completed or if a follow-up is confirmed by the initiator.

- First of all, extending the solution to more languages is crucial to ensure that the final product is accessible to a wider audience. This involves translating all text and messages used in the user interface, as well as creating speech recognition models for additional languages.
- We noticed that for speech-to-text, Open AI Whisper was better in English while Azure Cognitives Services was better in French. If we had had more time, a new objective would be to ensure that these two technologies are used in parallel depending on the language used.
- Then, adding settings like voice feedback voice and voice dictation speed will allow the user to customize the experience according to their preference. This includes the ability to choose from different voice options, as well as the ability to adjust the reading speed of equations.
- It is also important to integrate the functionality of editing equations to allow the user to customize the calculations according to their specific needs. This functionality was originally planned but could not be integrated due to time constraints. It would therefore be necessary to spend more time and resources to develop this functionality for future integration.
- Finally, the integration of the DicMath solution with existing software, such as Microsoft Word or Powerpoint, will allow the user to benefit from the computing power of DicMath in other work contexts. This could include the ability to insert equations directly into documents or presentations, or the ability to synchronize calculations between different software.

In sum, to ensure an effective project transition, it is important to ensure that the key actions mentioned above are completed or accounted for for future follow-up.

8 Project Completion Recommendations

The Closing Report identified several key points that should be considered to identify key recommendations:

- Good practices to be generalized, such as effective development methods or testing techniques that have been used successfully during the project.

For example, communication was quite successful throughout the project, whether between team members or with our partner Philippe TROTIN. Indeed, each time a doubt remained, we made an appointment to clarify the expectations, to ensure that we were all in agreement on certain key points of the project. To illustrate this point, for example, at the start of the project, we spent long meetings talking about how we would handle the project, how to successfully navigate an equation, how to make the interface is inclusive, etc.

- Improvements to be made to methodological aspects, such as project management processes or communication protocols that could be reviewed to improve efficiency in the future.

To pick up on the previous point, good communication within the team was unfortunately our weak point: the discussions to bring the team to an agreement being very important to us, we don't hesitate to stop the project to Agree on the strategies to adopt, the methods to apply, the technologies to use, etc. This led us to slow down the project, in favor of greater cohesion and synergy.

- Peripheral actions to be taken: the corrective actions that have been observed during the project and which must be implemented to guarantee the correct functioning of the solution in the long term. In short, it is important to generalize best practices, revise methodological aspects and take corrective actions to ensure the future success of the project.

It was during the project that we realized this handicap that over-communication brought us. Corrective actions could be to divide the group into sub-groups, to only need the approval of 2 people (instead of 5) to advance a module of the project.

9 Annexes to the closure report.

Please see the code on the github: https://github.com/Annabel64/PI2_101_Microsoft_DicMath