COMPUTER SCIENCE NEA

Quantum Computation Simulator



Tom Piercy (5514)

City of Norwich School (18133)

***Contents***

[***Analysis*** 4](#_Toc154049808)

[Background to the problem 4](#_Toc154049809)

[Current solution 4](#_Toc154049810)

[Client interview 5](#_Toc154049811)

[Intended User and Prerequisite knowledge: 6](#_Toc154049812)

[User needs and project limitations: 6](#_Toc154049813)

[Data sources and destinations 7](#_Toc154049814)

[Data dictionary 8](#_Toc154049815)

[Data flow diagram 9](#_Toc154049816)

[Entity relationship diagram 12](#_Toc154049817)

[Object orientation plan 13](#_Toc154049818)

[Project goals, objectives and KPIs 13](#_Toc154049819)

[Potential solutions 13](#_Toc154049820)

[Chosen solution 14](#_Toc154049821)

[***Design*** 17](#_Toc154049822)

[Database design 17](#_Toc154049823)

[**Initial solution** 17](#_Toc154049824)

[**Final solution** 19](#_Toc154049825)

[IPSO chart 20](#_Toc154049826)

[Sample of planned SQL queries 20](#_Toc154049827)

[Validating user input 21](#_Toc154049828)

[**Regular expressions** 21](#_Toc154049829)

[**SQL injection and system security** 23](#_Toc154049830)

[**Error handling** 23](#_Toc154049831)

[Algorithm research 24](#_Toc154049832)

[**Quantum algorithms** 24](#_Toc154049833)

[**Classical algorithms** 25](#_Toc154049834)

[Algorithm Design 27](#_Toc154049835)

[Class definitions 29](#_Toc154049836)

[Project hierarchy chart 29](#_Toc154049837)

[UI/UX design 29](#_Toc154049838)

[Project directory layout 33](#_Toc154049839)

[**Initial tree** 33](#_Toc154049840)

[**Final tree** 33](#_Toc154049841)

[***Technical solution*** 34](#_Toc154049842)

[Database implementation 34](#_Toc154049843)

[UI implementation 34](#_Toc154049844)

[Simulator implementation 34](#_Toc154049845)

[Runtime screenshots 34](#_Toc154049846)

[***Testing*** 35](#_Toc154049847)

[Trace tables 35](#_Toc154049848)

[Test tables 35](#_Toc154049849)

[Code iterations 38](#_Toc154049850)

[UX testing 41](#_Toc154049851)

[***Evaluation*** 42](#_Toc154049852)

[Achieved solution versus project objectives 42](#_Toc154049853)

[Client feedback 42](#_Toc154049854)

[Independent feedback 42](#_Toc154049855)

[Scope for further development 42](#_Toc154049856)

[***Appendix*** 43](#_Toc154049857)

# ***Analysis***

## **Background to the problem**

Mark, a first-year undergraduate student, is studying the optional module “Introduction to quantum computation” but is struggling to understand some of the key principles he is coming across. He would like a sandbox learning resource which balances teaching content with practical exercises so that he can get some experience working with the concepts he is learning. Additionally, he doesn’t understand when and where some of the knowledge he is gathering is used in real life and would like some exposure to the different applications of quantum computation.

Some definitions are provided to aid understanding of the project:

* A simulator is a piece of software that aims to model a physical system.
* Quantum computation refers to calculations performed using quantum states as a basis.
* A quantum computer is a computer that uses quantum states as opposed to standard classical bits to perform calculations.

Quantum computer simulators will naturally not perfectly reflect an actual quantum computer. Whilst this is fine for implementing most algorithms and explaining principles, it does mean that these simulators are often noiseless, stable and run over a large memory space.

## **Current solution**

Currently, Mark only has two places he can look for help: he can ask his lecturers, who are always very busy and inaccessible, or he can look on the internet. When Mark searches for the questions he wants answers to, he is greeted by a screen full of symbols and academic articles far too advanced for him to understand. Good introductory texts and videos can be found but they are spread across multiple sites and are hard to compile. Additionally, none of these resources are interactive so Mark has no real world understanding of what he is reading which is quite a mental block for him. An example of one such pre-existing resource is <https://quantum-computing.ibm.com/>



## **Client interview**

Before I began the project, I had the opportunity to talk to Mark about how he felt about his studies at the moment. The transcript is given below:

**Tom: “What issues are you facing at the moment with finding online resources?”**

**Mark:** “I think it's really hard to actually find stuff online as it is aimed towards people who have a deeper understanding than me. I really like videos because I understand them the best, but they don't really give me a chance to practice or consolidate anything. Loads of the websites I've found are from the early 90s and are running with the UI of a potato. I don't like this because I don't want to stay on the websites if they look bad.”

**Tom: “Are there any good parts to this process?”**

**Mark:** “There is loads of content online which is good in theory, but I am really overwhelmed by it, and I don't want to navigate it myself. There are probably some really good resources online but not for exactly what I need. I really want just one resource that can do everything.”

**Tom: “If you could improve any part of the current process of gathering resources, what would it be?”**

**Mark:** “As I mentioned earlier, there is stuff online but just way too much and I don't want to gather them, I just want them all available to me.”

**Learning points of the interview:**

1. Mark needs some form of resource collation as well as referring links so that if he finds a topic that particularly interests him, he can research it further.
2. Interactivity is a big thing for my client, he needs to be able to apply the concepts that he learns otherwise they won't stick in his head.
3. My client would like a teaching resource that can expose him to lots of different topics and what he needs to look for to explore them further.
4. The program solution should look nice since he will be on it for extended periods of time.
5. As an extension to 4, the program should be simple to learn to use and quick to pick up.
6. The program needs to be fun, otherwise my client won’t be inclined to learn from it.
7. Finally, the program needs to provide incentive to continue learning but not pressure the user in any way.

## **Intended User and Prerequisite knowledge:**

Whilst the project is being designed with Mark in mind (the primary client), the new system will be available to anyone who wants to increase their knowledge in the area of quantum computation. The program will not be bespoke or personalized to Mark in any way and will allow for multiple users on the same device to work through the lessons and play on the simulator independently of each other.

In terms of prerequisites, the only thing that will be required to use the software is a willingness to learn. The lesson design assumes no prior knowledge in the field. Some proficiency with technology would be helpful so as to not limit the software’s functionality; however, it is not required since the system will be designed to be as fluid and intuitive as possible. I will implement assertion statements and exception blocks to elegantly handle errors so that the user doesn’t get overwhelmed and can see clearly what error happened, why it happened and how the user should (if they can) fix it.

Despite their being one primary client, the concept of points and achievements create incentive for Mark to share the program with his friends and compete with them in a friendly manner - thus it is important to have isolated user “accounts” so that anyone using the software can easily see their own personal stats as well as how they rank up to others.

## **User needs and project limitations:**

**Client requirements:**

1. The program is fun, simple and quick to pick up.
2. The program incentivises learning and competing against others.
3. The program is aesthetically pleasing.
4. The program is fully interactive.
5. The program is a springboard to further studies.

**Project limitations:**

1. Time: The project is time bounded to February 2024
2. Knowledge: My programming experience is primarily in Python so this is the language that I will be using for my project. Whilst there may be more suitable languages for my solution, using python allows me to easily implement various programming paradigms, such as OOP.
3. The Python language: Python comes built-in with lots of useful modules that I can use to easily connect various parts of my program (such as modules for managing GUIs and databases) as well as a large community presence that can provide support with module implementation. However, Standalone Python is not very good at making web or mobile applications, so this does raise some limitations for me.

## **Data sources and destinations**

**Current system**

|  |  |  |
| --- | --- | --- |
| **Description** | **Source** | **Destination** |
| Academic papers | Internet/web archive from user input search query | Hard copy printout or digital file |
| Videos | Video archives on streaming sites (such as “YouTube”) from user query | Playback in browser |

**Proposed system**

|  |  |  |
| --- | --- | --- |
| **Description** | **Source** | **Destination** |
| User login information | Input (creating, logging into an account) | Users DB table |
| Player score/high score/ achievements | User completing an action inside the program | Scores DB table |
| Saved results | Saved csv file | History DB table |

## **Data dictionary**

**Database**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Field name** | **Field purpose** | **Data type** | **Field size** | **Example** |
| Username | Uniquely identifies a user of the program (not a primary key though) | Unique string, varchar | <= 16 characters | “Markiplier” |
| Password | Hash of a user’s password that links to their username and grants access to features locked behind accounts | String - **[A-Z,a-z,0-9]{64}** | 64 characters | “2fd83795b2f270fc3be2d4bdf16674045ad5e2b1453998a020e305ba5c0c45ab” |
| UserID | Primary key of users table | Auto Integer primary key (>= 0) | 4 bytes | 1 |
| Highscore | Total score of user | Integer (>= 0) | 4 bytes | 10000 |
| Difficulty | Describes the difficulty of a level which is used to calculate points | integer (>= 0) | 0-99 (2-digit number) | 99 |
| ChallengeID | Uniquely identifies an achievement that the user can receive | Auto integer primary key (>=0) | 4 bytes | 3 |
| challengetext | Describes an achievement, the text that the user will see on completion | String, text | < 2^16 Characters | “Successfully entangle two qubits” |
| Reward | The number of points the user receives for completing a challenge | Integer (>=0) | < 6 characters | 555 |
| date | The current date as recorded when completing certain actions | String, datetime | 19 characters | “1000-01-01 00:00:00” |

**Program**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Location** | **Variable name** | **Data type** | **Maximum size/value** | **Start value** | **Description** |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

## **Data flow diagram**

Towards the beginning of my project, I decided to make some basic data flow diagrams. These are heavily abstracted as rather than detailing everything I decided to just use them to gain a basic understanding of what I need to do, and how I could decompose the problem. They do not show all connections or even what happens in the subroutines. They don’t show what inputs are being taken or what decisions are being made: they are purely designed to show where these routines may be run and where inputs may be taken. I will use these diagrams to make, test and improve my final system.

A diagram of a computer program

Description automatically generatedA final data-flow illustration has been provided to aid any other developers. It shows the path taken through the whole program on initialisation and is used to give an example of which functions should return to where in the call-stack. The hope is that it aids understanding of the overall program flow:

## **Entity relationship diagram**

If I decide that databases are the right fit for my project, then the entities in my program could have the following relationships:

Every user has to have exactly one unique username, although multiple users can have the same password. A user is required to have a password. Users can, if they so wish, have a collection in the saved projects database linking to their files but they don’t have to - however if they choose to do this then every project or collection of projects must be attributed to one user id only. The leaderboard is set up when the first user signs up (as the score is a required property of every user this means that even when no points have been earned the leaderboard will show 0). Additionally, the user can get a predetermined number of points for completing achievements although there is the possibility that the user never completes any of these so this is a weak entity.

To implement this database, I would most likely use SQLite3 as it is built into python and provides a vast array of supporting documentation as well as functioning exactly as any other dedicated SQL controller. What this means is that using SQLite gives me the freedom to implement commands in a language that I am already familiar in (the database would be controlled and operated in python) but when it comes to querying the database, I can use the extensive and powerful instruction set of SQL - giving me the best of both worlds. Finally, there are lots of GUIs for viewing SQLite data, giving me the ability to see the database intuitively. This would rapidly speed up development as I wouldn’t need to write a structured command to debug or test a small part of the database and would also make tracking errors easier if there were any. Since SQLite3 is the most popular SQL python module alongside the fact that it is built in means that there doesn’t seem to be any worry of it becoming deprecated any time soon- which means that my program will still be able to function in the future.

## **Object orientation plan**

## **Project goals, objectives and KPIs**

Using the learning points of the interview and marks requirements for the solution I have devised a table listing the goals of the project and how I aim to meet them.

|  |  |  |  |
| --- | --- | --- | --- |
| **Goal** | **Priority** | **Testing** | **Key performance indicator** |
| The program is intuitive, fluid and easy to use | High | Client and Independent UX testing during development | Positive feedback from testers |
| The program incentivises learning in a fun environment | High | Independent testing at end of project | Positive feedback from testers |
| The program should be aesthetic and easy to be on for extended periods of time | Moderate | Developer testing and client interviews | Positive feedback from interviews |
| The program should be fully interactive | High | Developer testing | User freedom |
| The program should be timeless and not quickly outdated | Moderate | Developer testing | N/A |
| The program should be feature rich | Moderate | Developer testing | Client feedback |
| The program should accurately reflect the system it is emulating | High | Developer testing | Client interviews |
| The program should contain examples and relevant links. | Low | Independent testing | Positive feedback from testers |
| The program should not be able to break, crash or be exploited | High | Developer testing | Positive feedback during testing period |

## **Potential solutions**

1. An interactive website learning resource that provides exercises, questions and additional links

- This solution is quite similar to the current solution, with the addition of extra interactive elements to benefit mark. This is potentially a good thing because it increases familiarity with the program format – by extension increasing fluidity.

- An advantage of having a web implementation is that no programs need to be installed natively, freeing memory and potentially increasing program reach and access by having a very vast distribution network. This would mean that it would be very easy for other people to find and use the program.  
- A disadvantage of web distribution is that using the system is reliant on a stable internet connection. This potentially limits the amount of people that could use the system and the times that they could use it. This may be inconvenient and put people off using the system  
- A central server would have to be introduced which would both require both more complex code and considerations towards balancing usage and load. The server would need to go through strength and resilience testing which would take up more time and be more labour intensive.

2. A cross-device mobile game that teaches the basics of quantum computing through play.

- This solution would require considerations about distribution, as it would be important to know how the product would be sent out to potential users. Both Apple and Google have their own apk distribution platforms with large reach and additional ease of accessibility for less technical users, however the drawback to these is getting approval from the companies to list the application on their servers.  
- An advantage of this system would be that once distribution was complete, the program would be very easy to navigate through and intuitive for end users. This would be due to the familiarity and fluency of users for mobile applications.

- The system would be installed locally, so there would be no need for complex server code, whilst still maintaining an effective and efficient way to push updates to the program through the app store.

- A disadvantage of this solution would be the complexity of the code required to produce it. Tools such as Android App Development Studio and mitLab exist to speed up the process although both of these still rely on mobile optimised languages such as Kotlin, which I am unfamiliar with. Another issue would be that to avoid alienated part of the user base, the app would ideally need to be compatible with IOS, for which most apps are coding in a language such as Swift as well as having different coding requirements to be accepted onto the apple distribution system.

3. A sandbox desktop application that teaches quantum computing by encouraging exploration and providing visual feedback

- This solution gives a good compromise with the problem of distribution: Getting the program out to users is slightly more difficult and updates would be infrequent or non-existent, but this comes at the benefit of not needing to code to a FAANGs company specifications but rather being free to take the project in my own direction.

- To address the issue of program repetitively and interest stagnation - on a desktop we have the resources at hand to effectively design a less restrictive, more sandbox environment allowing the user to decide for themselves what they want to do in any given session. Additionally, coding in for desktop environment allows for a more advanced I/O system. We have more screen space to play with, better input control and increased flow for an extended range of characters that can be represented with a physical keyboard.  
- The code complexity is also a good compromise. The code required for the solution wouldn’t be as complex as the other ones as well as being in a language I am familiar with, whilst retaining a certain degree of program intricacy.  
- Code distribution could be done through free application hosting sites on the internet, such as sourceforge or majorgeeks and cross-OS compatibility – whilst still an issue – will be much easier to address as well as affecting significantly less people than with the mobile application due to the widespread use of windows in professional environments.

## **Chosen solution**

After talking with my client and explaining all of the different solutions, we have decided to go ahead with solution 3. Together we feel that solution 3 strikes a good balance between being a complex and complete answer to the problem whilst maintaining achievability. We have decided not to go with solution 1 due to my inexperience with server and network coding as well as the limited time frame of the project. We have decided not to go with solution 2 due to issues with learning the required development language as well as issues with authorship and the distribution of the system. After a detailed discussion of solution 3, mark is aware of the limitations of the project and how we can mitigate them. A table is provided below:

|  |  |  |  |
| --- | --- | --- | --- |
| **Problem** | **Impact** | **Significance/risk** | **Mitigation** |
| Declining interest due to rigid, repetitive or reused lesson structures | Users will stop using the program quickly, opting for other solutions | 3/10 – Low harm | Freeing the user from restrictive lesson structures by creating an open sandbox experience |
| Difficulty distributing the system to users | Users won’t be able to find the program | 2/10 – Easily mitigated | Using free distribution systems that already exist on the web |
| Difficulty porting the solution to other platforms | Mac and Linux users won’t be able to user the program | 5/10 – high probability, Low harm | Linux users can build from source. Mac users will be able to run from the python interpreter until development can be completed. Community porting is also an option |
| Program modification by 3rd parties | Due to the python interpreter code could be modified for malicious purposes | 6/10 – moderate risk, low impact | Coding best practices and secure development limiting access. Secure hashing on sensitive data. Installed locally with no network connection so no risk of large data leaks. |
| Device compatibility issues | Not all devices or displays will be compatible with display libraries used | 4/10 – minimal impact | Program can be shipped as executable with the python interpreter and all necessary libraries to minimise user setup. Python handles screen resolution natively |
| Accessibility for non-native English users | Users might struggle to use program. Developers might struggle to read code | 2/10 – Low harm | Translation resources exist on the web as well as coding for a desktop environment providing a greater range of printable Unicode characters and keyboard setups. Automatic OS translation |
| System misuse | Database dumping, SQL injecting and other program exploitation | 6/10 – moderate risk, low harm | Local install nature of program limits impact of database attacking. Saving sensitive information securely limits attack vector. Program corruption through SQL injection can be reset through a program reinstall with little impact on user. |

# ***Design***

## **Database design**

### **Initial solution**

**0-NF**

The following table is its non-normalised initial state. This is because in it’s current state it is unclear which field should be the primary key, and the context of the rest of the data depends on this. For example, picking a primary key of “Challenge ID” means repetition in the userID field, creating issues with retrieving, updating and removing data.

**1-NF**

The following data is in 1st normal form. We have moved data into more appropriate locations as well as unifying styling across all fields. The primary key has been specified to be UserID. The problem with using this form of the database is that there are still pieces of unrelated or dependant data being stored in the same location. We have also checked that all data is atomic.

**2-NF**

The data has now been put into 2nd normal form, separating data into its appropriate linking tables. However, in the process of putting the database into 2nd normal form we have had to reintroduce some redundancy in our table which we will fix in 3rd normal form. “ChallengeID” has been chosen as the primary key for the new table Challenges.

**3-NF**

The database has now been put into 3rd normal form, which is the final form of the data. It was decided that difficult to handle pieces of data (in particular the date achieved section being reliant on the users device to provide information: having the potential to break the database) would be handled by a JSON file linked to the other tables via the use of foreign keys. This also helps us remove our redundant userID field in the challenges table and perform a successful inference test on the rest of the data. Scoring has also been placed into a new table to eliminate redundancy.



### **Final solution**

I had the opportunity to speak to mark about the final database design and it was agreed that we would go ahead with the 3rd normal form design. In order of priority of implementation- mark commented that, whilst to best meet the project objectives we need the inclusion of every table, the highscores table was less important to him than the core functionality that the Users table (and associated login system) provides. We settled on the priority order: *Users, Challenges, JSON, Highscores*.

## **IPSO chart**

|  |  |
| --- | --- |
| INPUT   * Username * Password * Single line commands | PROCESS   * Verify user login details/register user * Store/Load progress * Simulate command * Recognise and award score and achievements * Update leaderboard * Draw diagram |
| STORAGE   * Leaderboard DB table * User credential DB table * User project DB table | OUTPUT   * Text/graphics * Achievement text/notifications * Leaderboard * Open project menu * Draw diagram |

## **Sample of planned SQL queries**

For testing and debugging purposes (as well as perhaps in the final program - see [“SQL injection and system security”](#_wfv8mqmrep4u)), all SQL queries that make changes to the database will be prefixed with the line “begin transaction” so that any changes can be rolled-back if the query was written incorrectly or made false changes to the database. Once the query has been verified to behave appropriately the changes can be committed.

\*Note that in SQL the “;” at the end of the line isn’t required so I will be omitting it for brevity

**Create**

**INSERT INTO** users (username, hash) **VALUES** (?, ?)

**INSERT INTO** leaderboard (userid,username,score) **VALUES** (?,?,?)

**Get**

**SELECT** userid **FROM** users **WHERE** (username = *“xyz”* and hash = “*xyz”)*

**Update**

**UPDATE** leaderboard **SET** (score) **WHERE** userid = “*xyz”*

**Delete**

**DELETE** **FROM** leaderboard **WHERE** (userid = *“xyz”)*

## **Validating user input**

### **Regular expressions**

A useful method of input validation is pattern matching. This is the process of pulling out patterns and general forms from an input string then comparing them to accepted and rejected states to determine if the input is valid. This is better than individually checking the string against others because it allows for easy adaptation as well as being a deterministic process (such that an accepted string will always remain an accepted string unless the underlying pattern matcher changes). For this reason, I decided to use the regex pattern matcher in my program (built-in: the “***re”*** module in python). Below are some regex tables containing the pattern match expressions that I could implement into my program as a method of validating user input.

|  |  |  |  |
| --- | --- | --- | --- |
| **Regex expression** | **Breakdown** | **Test strings** | **Pass/fail** |
| ^([A-Za-z0-9]|[A-Za-z0-9](([a-zA-Z0-9,=\.!\-#|\$%\^&\*\+/\?\_\{\}~]+)\*)[a-zA-Z0-9,=!\-#|\$%\^&\\*\+/\?\_\{\}~])@(?:[0-9a-zA-Z-]+\.)+[a-zA-Z]{2,9}$  (taken from [regex - How can I validate an email address using a regular expression? - Stack Overflow](https://stackoverflow.com/questions/201323/how-can-i-validate-an-email-address-using-a-regular-expression)) | **EMAIL VALIDATION**  Whilst the complete breakdown of this regex is far to complicated to fit into this table or indeed this project, the creator helpfully included a FSM diagram to illustrate how it works. You can find it [here](https://i.stack.imgur.com/YI6KR.png). | test  test@  test@test  test.com  test@test.com  test@test.co.uk  1test@test.t  test1@test.t  Test1@test.t  test!@test.t  \_test@test.t  !!!!@test.t  test@@test.com  @test@test.com | No (Pass)  No (Pass)  No (Pass)  No (Pass)  Yes (Pass)  Yes (Pass)  Yes (Pass)  Yes (Pass)  Yes (Pass)  Yes (Pass)  Yes (Pass)  Yes (Fail)  No (Pass)  Yes (Fail) |
| (?=.\*\d.\*)(?=.\*[\p{P}\p{S}].\*)(?=.\*[a-zA-Z].\*).{8,} | **PASSWORD VALIDATION**  **Positive Lookahead**  **(?=.\*\d.\*) -** Assert that the Regex below matches  **.** -matches any character (except for line terminators).  **\*** - matches the previous token between zero and unlimited times, as many times as possible.  **\d** - matches a digit.  **Positive Lookahead**  **(?=.\*[a-zA-Z].\*) -** Assert that the Regex below matches.  **[a-zA-Z] -** Match a single character present in the list  **{8,}** matches the previous token between 8 and unlimited times, as many times as possible.  **Global pattern flags**  **g** modifier: **g**lobal. All matches (don't return after first match)  **m** modifier: **m**ulti line. Causes ^ and $ to match the begin/end of each line (not only begin/end of string) | test  testtest  test1234  test123!  Test  Test1!  !test123  Test123!  test!!!! | No (Pass)  No (Pass)  No (Pass)  Yes (Pass)  No (Pass)  No (Pass)  Yes (Pass)  Yes (Pass)  No (Pass) |
| ^([a-zA-Z]+:)|(:[a-zA-Z]+:) | **SUPPLIMENT SYNTAX**  **^** - asserts position at start of a line  **:** - matches the character “:”  **+** - matches the previous token between one and unlimited times | test  test:  :test:  :  ::  test:  :test: | No (Pass)  Yes (Pass) Yes (Pass) No (Pass) No (Pass) No (Pass) Yes (Pass) |
| [a-zA-Z\_]\w\*\([a-zA-Z\_]\w\*\) | **OBJECT SYNTAX**  **\w** - matches a word character. | test  test()  test(a)  test(a )  test (a)  test(a)  test(\_)  test(%) test(a a) | No (Pass) No (Fail) Yes (Pass) No (Pass) No (Pass) Yes (Pass)  Yes (Pass)  No (Pass) No (Pass) |
| [0-9]+(\.[0-9]+)? | **DIGIT SYNTAX  ?  -** matches the previous token between zero and one times | 0  10  10.1  10.  10.0.10 | Yes (Pass) Yes (Pass)  Yes (Pass) Yes (Pass) Yes (Fail) |

### **SQL injection and system security**

In my login procedure, I will have to be particularly careful about sanitizing user input because it will be fed into a SQL query that gets sent to and executed on the users DB. This means that the user could theoretically abuse the login process to escape the query and execute their own query - this wouldn’t be especially difficult, due to the user actually being able to view exactly what query gets sent to the database since Python is an interpreted language. One safeguard around this is converting the program to an executable which I probably will anyway for user convenience: this doesn’t however fix the overall problem.

There are two main ways that I can attempt to fix this issue. The first is input sanitisation, which would mean controlling the users input to prevent them from injecting commands and also prevent users from accidentally setting their username or password to a banned string. This is probably the easiest of the two methods however, it depends entirely on having appropriate sanitisation which can be hard to do. Covering every single possible attack string as well as differentiating them from legitimate user credentials could be nearly impossible.

The second method would be much harder to implement but would function better (in fact, a combination of both methods would provide the best security but as I am limited by time, I'm not sure if I would have the resources to implement this). It revolves around using “Rollback” and “Commit” transaction commands in the SQL query itself. The program could read the database changes after executing the given command and compare these with the expected changes to decide whether or not to commit the command or rollback to before it was issue: for example, a select query should not produce any changes in the database, so if the program reads any updates when it supposedly sent a SELECT request then something has gone wrong.

### **Error handling**

My program takes full advantage of python’s call-stack based error handling by forcing certain interactions at the interpreter level – meaning that the user never has to fear of an unexpected crash or warning. The program makes full use of python’s built-in statements such ***try… except…*** for their optimised performance and ease of integration into large programs or projects that need to scale.

Another way of cleaning up console output for the user is by forced function calls. The program regularly uses statements such as ***assert***, ***raise*** and ***exit.*** These built-in functions can be used in conjunction with clever program design to produce clean program closing and human readable error messages when the user needs to respond to a problem with the interpreter. This solution is the most elegant because it includes functions that have existed in python for a long time and have become optimised to the interpreter rather than wasting time writing new code that wouldn’t function as well.

## **Algorithm research**

Algorithm research was vital to producing a solution that fit mark as well as being the most effective tool for learning. During my research, I drew on knowledge of algorithms that I had studied in education as well as those that go beyond the curriculum to help me create an efficient project. I explored both quantum and classical algorithms to set my project apart.

### **Quantum algorithms**

To fully understand the intricacies of a quantum computer, I had to effectively research various quantum algorithms. This provided me a platform to make sure that my program worked as accurately as possible. It also helped me gain the necessary base knowledge to build my system appropriately. In order to research this topic, I used resources I found online as well as a few series of lecture notes – I then collated these into a large binder for ease of research. This provided me an extensive array of resources that I could turn to when I needed guidance with my project.

**Grover's algorithm:**

One of the algorithms that can be implemented into my system is Grover’s algorithm. This is a searching algorithm with order as opposed to its classical problem counterpart which cannot be solved in fewer than steps for an unsorted array. It does this by solving multiple oracles with the intent of raising the probability of finding the correct answer and lowering the probability of measuring an incorrect answer. The probability states are then measured to resolve which solution is most likely to be correct.

A black background with a black square

Description automatically generated with medium confidence

**Superposition:**

Not an algorithm but a really important element of quantum computers and one that fits well to talk about here. Superposition is the name given to the state of probability that an entity exists in because of the non-deterministic nature of quantum mechanics. It is a very important concept in quantum computing as it is the mechanism that allows exponential speed up of subroutines by solving multiple oracles at the same time. It is akin to parallel processing or threading, but it doesn’t take up additional resources in the same way because it is an inherent property of nature as opposed to a human construction. Superposition is a far more complicated topic than has just been described however the hope is that this small paragraph sufficiently aids understanding for the other algorithms discussed below.

**Shor’s algorithm:**

A black background with a black square

Description automatically generated with medium confidenceShor’s algorithm is a complex quantum algorithm for factorising large numbers. It was developed in 1994 by Peter Shor. On a quantum computer it runs in polynomial time: and has important implications for cybersecurity: undermining the current security standard of the intractability of factoring large numbers. At this moment in time, creating a stable system with enough qubits is impossible.

**Entanglement:**

Entanglement is the process of equally splitting the probability of finding an entity in any given state with no leaning. This means that the object has no preference to be in one state versus any other: it is achieved with the Hadamard gate and is very important in quantum computing for setting/resetting qubits as it is symmetrical (it is its own inverse), cyclical and deterministic. This means that applying the algorithm a set number of times from a set state will always result in the same result.

### **Classical algorithms**

**Diffusion process:**

To make my program more visual and appealing to users, I wanted to include a diffusion model for a particle so that the user could “see” the effects their commands were having on the simulation and to hopefully help them retain that knowledge.

In doing this I went through many different methods before settling on the final one. Initially I tried using a combination of a SoftMax algorithm to generate proportions and then a normalisation function to get it into an appropriate range. The thing that was difficult about this method was that there were lots of variables that needed to be considered and lots of temporary lists and changes that needed to be saved. Not only was this harmful due to the excessive memory use but also because of how python is interpreted as a language – having a large base of variables before proper testing of function application and scope means that data can get overwritten or stored unnecessarily throughout the program’s runtime.

****

A screenshot of a computer program

Description automatically generated

In the end I decided to use Gaussian noise to create an element of randomness proportional to the step of the loop. This meant that my random noise function could smooth out over time to be more representative of a diffusion function. It is also much more efficient, as no diffusion calculations need to take place: we can mimic them with a large amount of accuracy without having to spend the computational cost of full simulation.

A blue and red line graph

Description automatically generated

**Tree traversal algorithms:**

For parsing data and commands taken via user input, the standard method is to use the regular development pipeline of ***“lexer -> abstract syntax tree -> action tree -> interpretation”.*** Whilst not strictly necessary when you are not building an entire language from the ground up, it is always beneficial to adopt standard or conventional methods as this makes development significantly easy as well as increasing the options for adaption. Since AST’s and AT’s both make use of the tree data structure, they rely on traversal algorithms to function.



As of such, I had to research various different traversal algorithms including the most common in/pre/post order methods.

## **Algorithm Design**

In order to properly plan my project’s python implementation, I designed a set of simple algorithms in pseudocode that I can use to prepare for application in my program.

**IN-ORDER TRAVERSAL:**

**SUB** in\_order(root)

**IF** root **exists** **THEN**

in\_order(root.left)

**OUTPUT** root.value

in\_order(root.right)

**ENDIF**

**POST-ORDER TRAVERSAL:**

**SUB** post\_order(root)

A diagram of a network

Description automatically generated with medium confidence **IF** root **exists** **THEN**

post\_order(root.right)

post\_order(root.left)

**OUTPUT** root.value

**ENDIF**

**PRE-ORDER TRAVERSAL:**

**SUB** pre\_order(root)

**IF** root **exists** **THEN**

**OUTPUT** root.value

pre\_order(root.left)

pre\_order(root.right)

**ENDIF**

**BREADTH FIRST SEARCH:**

**SUB** BFS(step)

        visited **->** [**False**] **\*** (**MAX**(self.graph) **+** 1)

        queue **->** **LIST**()

**ENQUEUE** step, queue

        visited[s] **->** **True**

**WHILE** queue

            step **->** **POP** 0, queue

**OUTPUT** s

**FOR** I **IN** self.graph[step]

**IF** visited[i] **=** **False** **THEN**

**ENQUEUE** i, queue

                    visited[i] **->** **True**

**ENDIF**

**ENDFOR**

**ENDWHILE**

**DEPTH FIRST SEARCH:**

visited -> **SET**()

**SUB** DFS(vertex, visited)

**ADD** vertex, visited

**OUTPUT** vertex

**FOR** neighbour **IN** self.graph[v]

**IF** neighbour **NOT IN** visited **THEN**

                DFS(neighbour, visited)

**ENDIF**

**ENDFOR**

**GROVERS ALGORITHM:**

**SUB** oracle(x, target)  
 **IF** x = target **THEN**

**RETURN TRUE**

**ELSE RETURN FALSE**

**ENDIF**

N -> **LENGTH** list

**FOR** qubit **IN** list

**HADAMARD** qubit

**ENDFOR**

**FOR** k = 1 **TO** **SQRT** N

Oracle(list[k],target)

**FOR** qubit **IN** list

**HADAMARD** qubit  
 **X** qubit  
 **IF** **FOR** **EVERY** element **IN** list, element -> |1> **THEN**

**Z** qubit  
 **ENDIF  
 HADAMARD** qubit

**ENDFOR**

**ENDFOR**

result -> **MEASURE** list

OUTPUT result

## **Class definitions**

#THIS SECTION CONTAINS CLASS PSEUDOCODE

## **Project hierarchy chart**

#THIS SECTION CONTAINS A DIAGRAM OF THE PROGRAMS HIERACHY, AS ONE PROBLEM SPLIT INTO SUBPROBLEMS AS WELL AS SHOWING INHERITENCE

## **UI/UX design**

**Title design:**

The first part of the program that the user interacts with once will be a console interface for the login system (which could also be implemented as a GUI for more accessibility later down the development cycle for the program). As of such, I wanted this to look interesting to leave a positive impression on the user: this is entirely cosmetic, it has no functionality other than making the program appear more friendly to new users or those less familiar with a CLI environment. I went through a few designs for the splash screen which I have included below:

A screenshot of a computer screen

Description automatically generated

In the end I settled on the third design as I felt it suited the rest of my design much better as well as being easiest to read and most compatible with different machines (as it is comprised entirely of 7-bit ascii symbols)

**GUI design:**

Since a large portion of my project requires visual feedback to the user it was important to design a GUI that looks and feels satisfying. Additionally, the objective of making the program as aesthetic and simple as possible fed into my design process. I ended up doing some digital drafts of the generic layout I was looking for in my program before implementing these in tkinter and tweaking as required.

A black rectangular object with white text

Description automatically generatedA drawing of a whiteboard

Description automatically generated

A screen shot of a computer

Description automatically generatedA drawing of a grid

Description automatically generated

**A drawing of a square object

Description automatically generated**A screenshot of a computer screen

Description automatically generated

A computer screen with a white screen

Description automatically generatedA screenshot of a computer

Description automatically generatedAfter this I went into figma to properly plan how the user might react with the program as well as to curate an effective and aesthetic program flow. The designs I produced are shown below.

A screenshot of a computer

Description automatically generated

Using figma helped me properly plan the interactions that the user would have with the system so that I could effectively program an implementation in tkinter.

**CLI design:**

An integral part of my project was the CLI that the user would interact with throughout the runtime of the program. As of such, it was vital that the CLI was designed to be fluid, intuitive and visually appealing. I went back into figma to sketch some drafts of what I wanted the user to interact with and plot the general flow of how they would perform these interactions. This was useful because it allowed me to have a physical blueprint to work to when programming.

A screenshot of a computer

Description automatically generatedAt first, I was unsure of how I wanted the user to perform entry in my program. I had a wide selection to choose from and I didn’t want any key binds to interfere with each other due to the growing number of tabs in my program. I found a good module to use called getch which solved some of these problems. Getch works by directly scanning the standard output buffer for new input characters and sends them directly to the program. This has the effect of being able to take input in python without having the user have to use the enter key – reducing collisions with keybinds in matplotlib and tkinter. Since it is written in python it can also be called directly as a function whenever you as the developer want to check the buffer. This provides much greater control, especially in programs that may have multiple input points perhaps even simultaneously.

A number and number in black squares

Description automatically generated with medium confidence

As for drawing a quantum circuit. It as been a part of my project from the beginning, and I have always wanted to implement it into my solution. After further research however, designing my own ascii printer seems to be beyond the scope of my project: as of such I have decided to investigate using pre-existing tools[[1]](#footnote-2) as I imagine it will be easier to parse the users input into the format that the tool requires than it would be to rewrite the tool from scratch for my project. I will still consider writing my own, however it would be as a later iteration of the code, and I won’t worry about implementing it until I have finished everything else.

## **Project directory layout**

### **Initial tree**

**>nea**

**---->pycache**

---->cbit.py

---->interface.py

---->login.py

---->main.py

---->qbit.py

---->vector.py

>users.db

### **Final tree**

# ***Technical solution***

## **Database implementation**

## **UI implementation**

## **Simulator implementation**

## **Runtime screenshots**

# ***Testing***

## **Trace tables**

## **Test tables**

My code was fully tested at two main points in my project: roughly halfway through as well as at the end. This was to ensure that no errors slipped through the smaller testing that went on during the development process. This was important to do because it gave me the opportunity to step back from my project for a bit and just focus on fixing any bugs or issues that might have come up before continuing to improve and adapt the project.

**Vector.py initial**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **#** | **Location** | **Test** | **Intended termination behaviour** | **Pass/Fail** | **Comments** |
| 1 | \_\_init\_\_ | Incorrect size parameter type | Error Message + exit | Pass |  |
| 2 | setElement | Valid data | Return True | Pass |  |
| 3 | setElement | Index larger than list | Return False | Pass |  |
| 4 | setElement | Wrong index type | Return False | Pass |  |
| 5 | getElement | Valid query | Return correct data | Pass | Data returned is from the right place |
| 6 | getElement | Wrong index type | Return False | Pass |  |
| 7 | getElement | Index larger than list | Return False | Pass |  |
| 8 | scalarMul | Valid query | Return multiplied vector | Fail | Returns (0,0) always |
| 9 | scalarMul | Bad type for ‘num’ | Return False | Fail | No error checking |
| 10 | \_\_mul\_\_ | Valid query | Return multiplied vector | Fail | Fails because of scalarMul |
| 11 | setElements | Valid entry | Return True | Pass |  |
| 12 | setElements | Invalid entries | Return False | Fail | No error handling |
| 13 | setN | Valid float entry | Return True | Pass |  |
| 14 | setN | String entry | Return False | Fail | Allows string entry |
| 15 | allZeros | Zeroed vector | Return True | Pass |  |
| 16 | allZeros | Non-zero vector | Return False | Fail | Works when any element other than the 0th index is not 0 |
| 17 | magnitude | Valid vector object | Return correct magnitude | Pass |  |
| 18 | isUnit | Unit vector input | Return True | Pass |  |
| 19 | isUnit | Non-unit vector input | Return False | Pass |  |
| 20 | unit | Valid vector object | Return correct unit vector | Pass |  |
| 21 | tensor | 2 vector objects – same length | Return correct tensor | Pass |  |
| 22 | tensor | 2 vector objects – different length | Return correct tensor | Pass |  |
| 23 | tensor | 1 vector object – one other type | Return False | Fail | No exception handling |
| 24 | \_\_repr\_\_ | Valid vector object | Return correct string representation | Pass |  |

**Cbit.py initial**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **#** | **Location** | **Test** | **Intended termination behaviour** | **Pass/Fail** | **Comments** |
| 25 | \_\_init\_\_ | Incorrect dirac type | Stops execution | Pass | Needs better exception handling |
| 26 | \_\_init\_\_ | Bad sub value | Stops execution | Pass | Needs better exception handling |
| 27 | \_\_init\_\_ | Bad sub type | Stops execution | Pass | Needs better exception handling |
| 28 | \_\_init\_\_ | Large dirac – no sub | Return correct tensor | Pass |  |
| 29 | \_\_init\_\_ | Large dirac – suitable sub | Return correct tensor | Pass |  |
| 30 | setElement | Index larger than list | Return False | Pass |  |
| 31 | setElement | Bad index type | Return False | Pass |  |
| 32 | setElement | Bad value | Return False | Pass | Cbits can only have 1’s or 0’s as elements |
| 33 | setElement | Valid data | Return True | Pass |  |
| 34 | measure | Valid vector | Return correct measurement | Pass |  |
| 35 | measure | Vector larger than 2 | Return False | Pass | Use probcollapse instead |
| 36 | probcollapse | Standard cbit | Print correct probabilities | Pass | Can’t be used in conjunction with setElement |
| 37 | probcollapse | Larger Vector | Print correct probabilities | Pass |  |
| 38 | \_\_repr\_\_ | Valid cbit object | Print correct string | Pass |  |

**Qbit.py initial**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **#** | **Location** | **Test** | **Intended termination behaviour** | **Pass/Fail** | **Comments** |
| 39 | \_\_init\_\_ | Incorrect dirac type | Stops execution | Pass | Code repeated from Cbit.py |
| 40 | \_\_init\_\_ | Bad sub value | Stops execution | Pass | ‘’ |
| 41 | \_\_init\_\_ | Bad sub type | Stops execution | Pass | ‘’ |
| 42 | measure | Basic Cbit configuration | Return Cbit value | Pass |  |
| 43 | measure | Modified Cbit | Return collapsed odds cbit | Pass | Returns a value pulled from a list with magnitudes as weights |
| 44 | measure | Larger than expected Qbit | Return False | Pass | Use probcollapse instead |
| 45 | \_softmax | Valid input | Return probabilities | Pass | Returns correct values |
| 46 | \_softmax | Non-list argument | Return False | Fail | Needs exception handling |
| 47 | \_softmax | Non integer array | Return False | Fail | ‘’ |
| 48 | \_normalise | Valid input | Return modified list | Pass |  |
| 49 | \_normalise | Non-list parameter | Return False | Fail | Needs exception handling |
| 50 | \_normalise | Non-float maxprime | Return False | Fail | Iterates over string, All operations performed are considered valid between str and list types as they are both iter. |
| 51 | \_normalise | Smaller than expected maxprime | Return modified list | Pass | Function should work for any integers |
| 52 | \_normalise | Negative maxprime | Return modified list | Pass | ‘’ |
| 53 | \_normalise | 0 maxprime | Return modified list | Fail | ‘’ |
| 54 | \_applygauss2d |  |  |  |  |

## **Code iterations**

Over the course of my project, each function and class had its own micro development cycle. This meant that some functions were initially written with no regard to speed, efficiency or readability to produce an MVP that would solve the problem. This gave me lots of scope for development as I could produce a working section of code and then go back through it to refine and improve it. This meant that sometimes code was scrapped completely because there existed a much better solution that was found later in the project. Some examples of code and how they changed between the start of the project and the present are given below to highlight this. Each function and class followed its own small spiral development, which helped me add all of the features my program needed in a timely manner avoiding a feature crunch towards the end of the project, in which functionality is shoe-horned in at the last minute without a stable or strong supporting code base. This is obviously not ideal. A black and white screen with a circular pattern

Description automatically generated with medium confidence

**Initial login:**



A white background with text

Description automatically generated

**Final login:**

A screenshot of a computer program

Description automatically generated

A screenshot of a computer program

Description automatically generated

**Graph plotting and threads:**

The graph plotting code went through many different variations until I found the one that I wanted to settle with. This is because I had to run multiple main loops simultaneously and independently: which forced me to teach myself some basic asynchronous programming and threading in python. Ultimately, I decided to use the built-in function ***matshow*** from matplotlib – this was because it was just more optimised than anything I could have possible written by myself as well as being built with 2d arrays in mind, which was a base type that I am already very familiar with rather than having to learn how use NumPy arrays or another custom type.

**A computer screen shot of a program

Description automatically generated**

## **UX testing**

# THIS SECTION SHOWS USER FEEDBACK ON UI FLOW AND EXPERIENCE

# ***Evaluation***

## **Achieved solution versus project objectives**

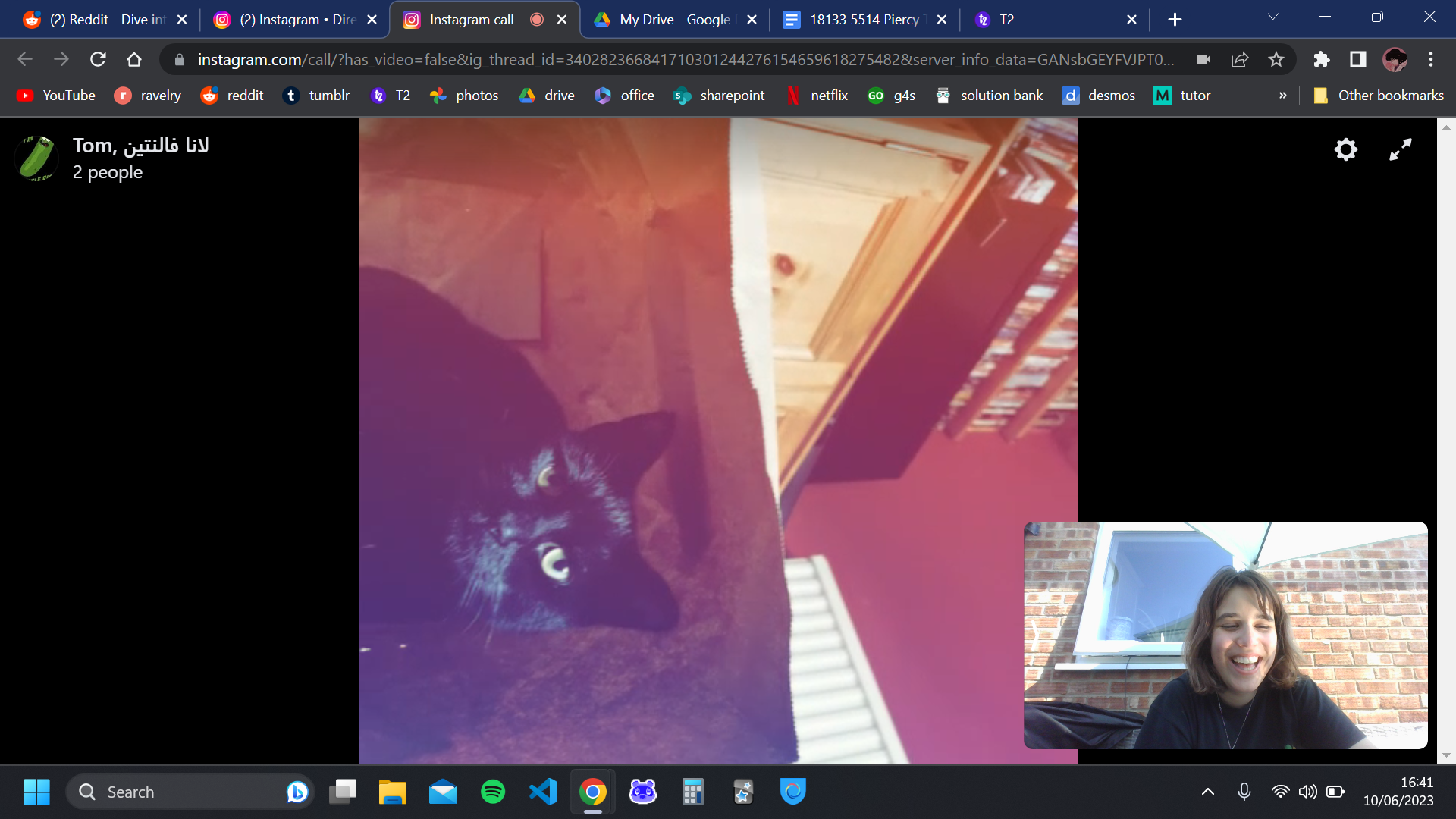
## **Client feedback**

#THIS SECTION CONTAINS MADE UP FEEDBACK FROM THE CLIENT

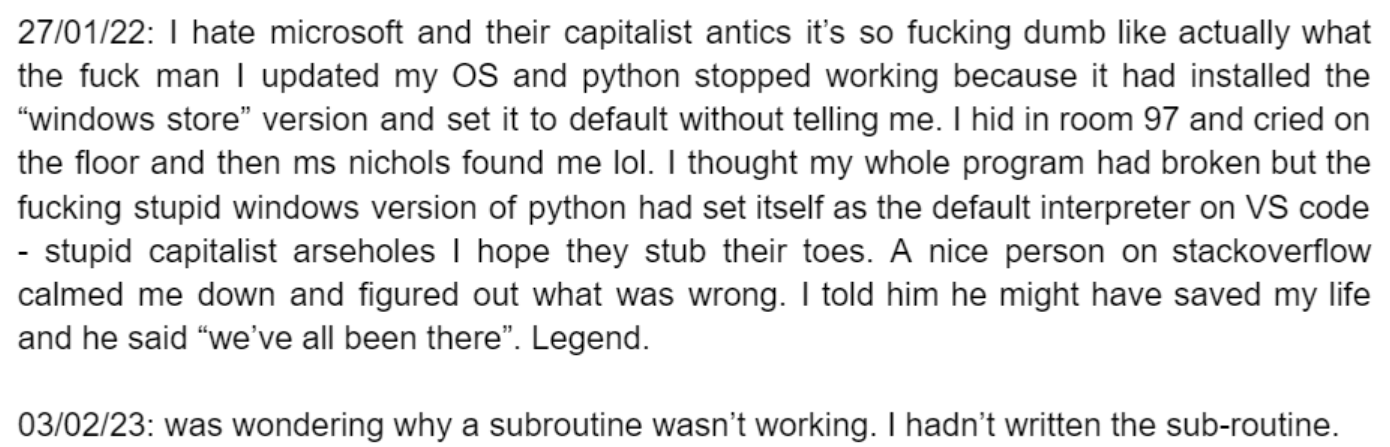
## **Independent feedback**

#THIS SECTION CONTAINS FEEDBACK FROM PEOPLE OTHER THAN THE INTENDED CLIENT

## **Scope for further development**



# ***Appendix***



1. <https://quantikz.krastanov.org/?circuit=> [↑](#footnote-ref-2)