Title: Innovative CAPTCHA test

Team Member's: Andrew Wong, Gurpreet Singh

1. Introduction

1.1 Problems

Every day there are millions of people creating a new account on apps, websites, and financial platforms. Oftentimes when going through the process of making said accounts you've run into a CAPTCHA test. These tests even appear when you are trying to download a program, or even when you've forgotten your password and are constantly trying to login. CAPTCHA stands for Completely Automated Public Turing test to tell Computers and Humans Apart. It is definitely a mouth full and that is why we keep it short with CAPTCHA. In short, CAPTCHAs were made to create a simple test that humans can solve but automated bots can't.

This was a huge breakthrough in the security industry because in the year 2000 yahoo email, the biggest email company had no way of authenticating users. This created a big issue because the public was able to create free emails infinitely. Especially when malicious users made bots / scripts to generate millions of spam accounts. These accounts were used to spread viruses, sell fake products, and steal bank information. This is where CAPTCHA came into play and aided in the mass production of spam accounts. This process will filter out the bots and only allow authentic users to access a specific page or site. To the majority of humans this test is supposed to be simple and quick. However, slowly these tests have become more difficult and time consuming. The main reason being that bots are being programmed to better fill out these CAPTCHA tests. The very first captcha test created was having a simple image with a mix of numbers and characters. The user would have to type in the correct code in order to proceed. Sooner or later people created bots to solve these captcha tests and to counteract this, CAPTCHAS became distorted in one way or another for it to be harder for bots to read the image. Meanwhile, humans just have to take a second longer to figure out the proper characters displayed in the image. This eventually led to humans becoming frustrated and slower at solving these CAPTCHA tests. On the other hand, as bots gather more tests and data they are getting better at solving these tests. The first CAPTCHA was invented by Luis von Ahn and the test was meant to be simple. That is no longer the case and CAPTCHAs have since been upgraded and improved.

1.2 Brief approach and result

Our approach to this issue was to create our own CAPTCHA test that hits the following requirements: free, quick, simple. We decided to combine some of the strategies from previous CAPTCHA tests that decide whether the user is a bot or a human. Basically, the program will be a time based test with a skill check on human reaction time. Our design is to have a chemistry test tube that will have a red line placed randomly. The objective of the game is to press the spacebar key to stop the water just at the red line. The red line will be thick so it is nearly impossible to miss the line. However, this combination of a time task

and image won't allow bots to beat the game. The user must complete the task under a specific amount of time which counteracts bruteforce attempts. The animation of the water flowing also helps make it harder for the bot to know the exact time to stop the water. There is no way to test the results except we have a functioning CAPTCHA test on a website. I say this because I'm not sure where to source a CAPTCHA bot to test our CAPTCHA against. As stated in our project proposal, there exist CAPTCHA farms that constantly brute force CAPTCHA tests to continue creating these spam accounts. They can collect libraries of distorted images to give the best guess to what character the CAPTCHA test is asking for. Given the time restraint for this project it was also already hard enough to learn new technologies and implement it in creating the CAPTCHA game. There is no way of really testing our test against CAPTCHA farms.

2. Background

The technologies used to create this CAPTCHA is the pygame library, in the python coding language, and photoshop. You will need to download pygame on your windows or linux system to work on this CAPTCHA test. Background knowledge in photoshop is needed for the animation of the rising water created in the game. From scratch I drew each frame of water rising in the tube. Looping through an array of the images allowed us to generate this animation. That being said, in coding the user must understand the fundamentals of classes, loops, data types, arrays, conditional statements and functions.

There are many types of CAPTCHA tests. There exist simple math problems, word problems, social media sign-ins, time-based, picture identification, and noCaptcha Recaptcha. Most users are used to picture identification and noCaptcha Recaptcha on a day to day basis. OCR (Optical Character Recognition) is a very labor intensive process that bots use to answer these CAPTCHAS. We want to create a test that will make it even harder for the bot to solve. Not only will the animation aspect of our program make it hard but the time constraint to our program. This doesn't allow the bot to bruteforce past our CAPTCHA test.

3. Solutions and Methods

The way we created this CAPTCHA test is through pygame. Pygame is a library that you can download and import into our python code. We start off with the import and we want to set up the basics of the program. The most basic skeleton being initializing pygame, creating a screen size for the game (in our case 800 x 800 px screen) and a while loop that keeps the program open until the user clicks the quit program button on the top of the window. We use an event listener for that. Now for all functionality and logic in the CAPTCHA must be in this loop. Everything else aesthetic wise can be outside the loop. For example, in the bottom of the code we have our aesthetics, and frames per second for our animation. We set the background color to white, put the group of sprites on the screen,

give it a speed to display on the screen, and a clock to control the frame rate. On the screen we write the instructions that the user must click to start the game and press spacebar to top the water at the red line. The red line is placed randomly on the test tube each time the game is played. This red line is purposely made to be thick so the user has an easier time stopping at the right time. They can stop the water within 3 frames of the animation and still pass the test.

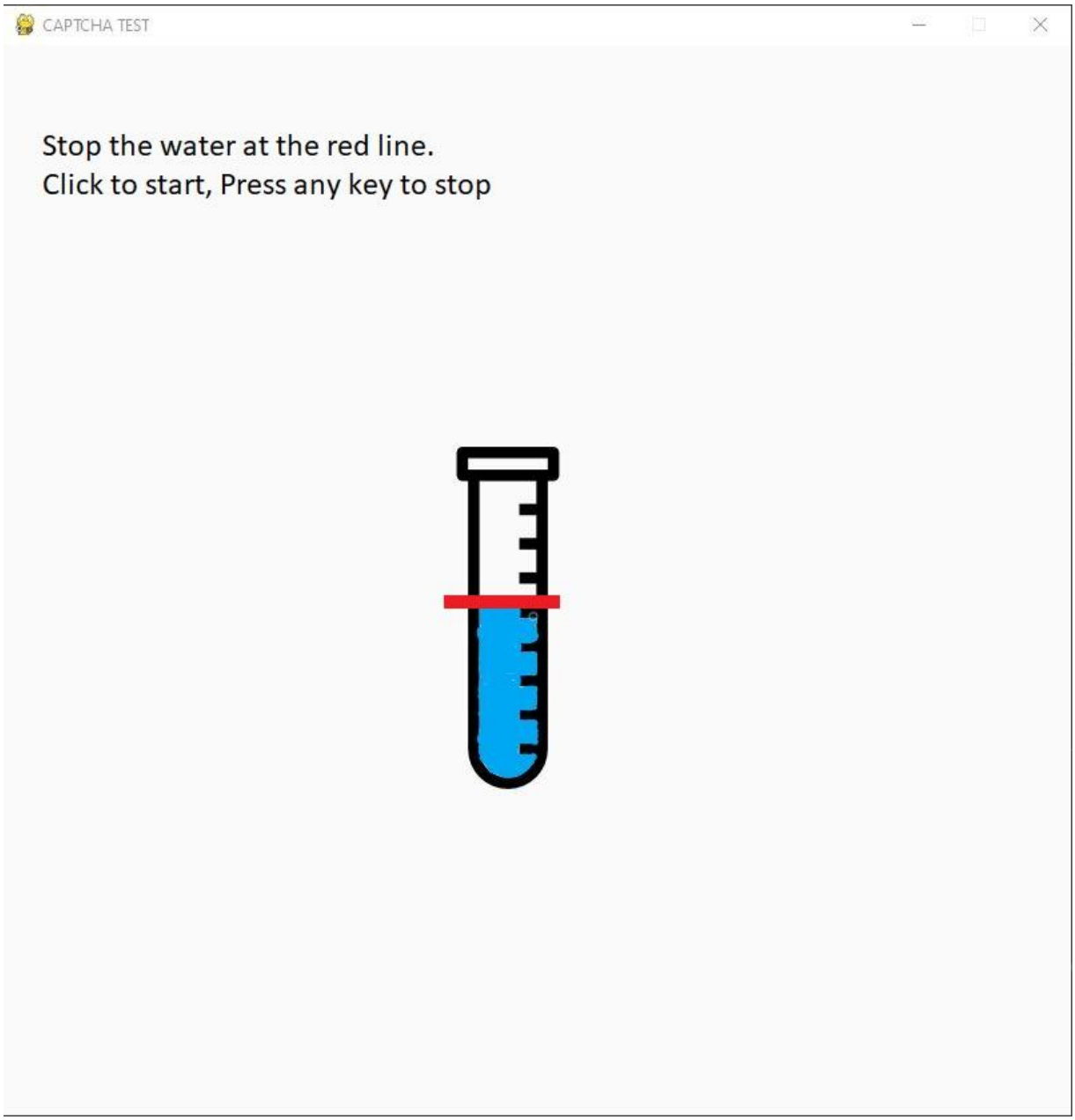

Fig. 1 A draft image of our CAPTCHA test.

4. Experiments and Analysis of Results

After thorough research online, we've found that the average time it takes to solve a CAPTCHA test is 10 seconds. Our goal was to come up with an alternative test that is faster than the average time of the typical CAPTCHA we are all used to. That being said, our CAPTCHA test is at max 2 seconds long. We can confirm this by the amount of time it takes for our "test tube" animation to fill up with water. That means any time a user passes our test it should take less than 2 seconds. We also took human reaction time into factor for our CAPTCHA test. The average human reaction time is 250 milliseconds which is about .25 seconds and for this exact reason we have our frame rate go at this speed. This allows for the user to get a better experience and actually complete our CAPTCHA game with ease. Now we also have to consider people who are below average with a slower reaction time. To solve this issue we decide to have some leniency by allowing the user to be within 2 frames of the actual answer. The thicker the red line allows multiple frames to be the correct answer plus-minus one from the actual correct answer. When creating our CAPTCHA test we wanted to accommodate all factors that relate to the time and difficulty of the test. Taking everything into consideration, we believe the design we came up with is a better alternative than the traditional free CAPTCHA tests on many web pages.

5. Related (prior) works

To understand the backbone and prior works of CAPTCHA we must look at its history. Mentioned previously, the beginning of CAPTCHA was originally created in the early 2000's. However over the years technology has also grown drastically. This meant that CAPTCHA had to get updated and maintain a balance of difficulty for humans to solve and making it harder for bots to solve. The ideal solution was to distort the characters in the image making it hard for a bot to distinguish what is on the screen. Leaving it an easy task for humans to recognize the text. However, in 2009 google took over CAPTCHA as the original inventors gave up on trademarking it.

Now that Google had complete control and rights over CAPTCHA they went ahead and revamped it by instituting reCAPTCHA v2. This new version of CAPTCHA is similar to the first one except with tiles of images. Each image either has an object that the human must identify and select or a random object that is not related to the topic at all. The test is simply to select all image tiles that have the said object in it. Not only does this require user input on each tile but it's not easy for the bot to distinguish images. Over time these tests have also become harder because although it is more accurate humans are constantly second guessing their answers. Making it not only frustrating but time consuming. Google then released reCAPTCHA v3. This third version is a CAPTCHA that tracks your navigation and pathing on a website. Based on how you interact on the web page it will grade you with a score between zero and one. Depending on your score you will be kicked off the website. The main issue with this version is that the predetermined score is based on the website owner's discretion. Meaning, the owner or whoever made the website gets to set anyone who scores below X score will be considered a bot and disconnected from the website. This

isn’t a 100% proof method because there is still the chance of a potential customer or user to get booted off the website. In my opinion it is better to allow some bots in than to ruin a users experience by kicking them off the website. The upside to this CAPTCHA is that it is a seamless test that requires no time for anyone to do anything.

Lastly, the latest version of CAPTCHA technology is reCAPTCHA enterprise. The main issue is, reCAPTCHA enterprise costs money. Consumers may not want to pay for the latest technology providing more protection than what is needed. There are other alternatives than Google's CAPTCHA which include simple games. For example, one being to click an arrow that rotates an image until it is upright. Once the image is upright under a specific time, the user passes the test. Another example is a simple slider where the user will slide an object into the proper spot within a time limit. The most effective method according to my research is using a premade account as a login for new apps / websites. Creating an account with an already existing social media account or email. This will allow most users to instantly create an account without having to solve any puzzles.

Our proposed idea is different from other CAPTCHA tests in the aspect of simplicity and time it takes to solve the test. The user should be able to pass the test in two clicks. Most importantly, our CAPTCHA will be a free alternative to Google's CAPTCHA tests. Mentioned in our intro the main goal and purpose was to fulfil three requirements: fast , simple ,and free. I believe that our CAPTCHA test satisfies these conditions.

6. Conclusion (Free page)

The way we implemented our solution with a new CAPTCHA game took in many different factors. From the planning phase to actually creating our test we tried our best to make something special to make websites just a bit more secure. With the time restraints of this semester there are some things that we didn’t get to implement that we wanted. Also we didn’t get a chance to polish and create a nice looking CAPTHCA test. Our prototype is functioning and can do its purpose however it isn’t a final / finished product. We could improve greatly on the aesthetics / art of the game. We could create something more pleasing to the eye along with instructions to give the user an idea on what to do.

In addition to the time restraint we haven’t had time to test our game against OCR technology. There are bot farms that exist to farm / complete CAPTCHA tests. Which is a very important part of security, constantly trying to break the security to make it stronger. Kind of like what happened to the original CAPTCHA they had to keep making it harder and harder for bots that it eventually became too hard for humans. In the future we would definitely want to test our CAPTCHA against bots. We would also like to polish the CAPTCHA game into a complete test that is appealing enough to use on the web.

Reference:

Why captchas are getting harder - youtube. (n.d.). Retrieved December 3, 2021, from https://www.youtube.com/watch?v=lUTvB1O8eEg.

Ackerman, E. (2021, June 24). *Upgrade to superhuman reflexes without feeling like a robot*. IEEE Spectrum. Retrieved December 3, 2021, from https://spectrum.ieee.org/enabling-superhuman-reflexes-without-feeling-like-a-robot.

The recaptcha project. The reCAPTCHA Project - Carnegie Mellon University CyLab. (n.d.). Retrieved December 3, 2021, from https://web.archive.org/web/20171027203659/https://www.cylab.cmu.edu/partners/success-stories/recaptcha.html.