A review of the accuracy and user-response assessment of the Art-Related Turing Test (ARTT) application

Ahan Sabharwal

BACKGROUND

A CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) is a program that protects websites against bots by generating and grading tests that humans can pass but current computer programs cannot.[1] Usually, this is done by offering users an image of 'distorted' text that humans can identify easily, but is nearly impossible for a bot to decipher.

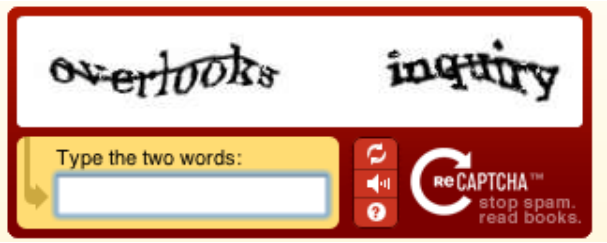


Figure 1: A CAPTCHA example

With the progress in artificial intelligence (AI) technologies, computer programs are getting better at such challenges, with CAPTCHAs scrambling to stay ahead of the state-of-the-art.

During my participation in the Applied Sciences and Engineering (ASE) track of the 2016 Yale Young Global Scholars (YYGS) summer program, my capstone project team[2] was tasked with proposing a solution to a research problem related to cryptography. Our team reviewed current CAPTCHA efforts and decided to create a better solution, which would

[1] CAPTCHA: Telling Humans and Computers Apart Automatically (http://www.captcha.net/)
[2] My YYGS Capstone team included the following students: Andrew Gagliostro; Sonal Gupta; and Sasha Valone

leverage the human capacity for creative thought and expression to identify the type of user input.

ARTT uses artificial intelligence to differentiate between drawings made by humans and those by bots. A detailed presentation of ARTT can be found at ahan.io/ARTT, and the prototype at ahan.io/ARTT-prototype. The prototype has been tested over 1000 times, and ARTT's artificial intelligence has used the tests to learn and become smarter.

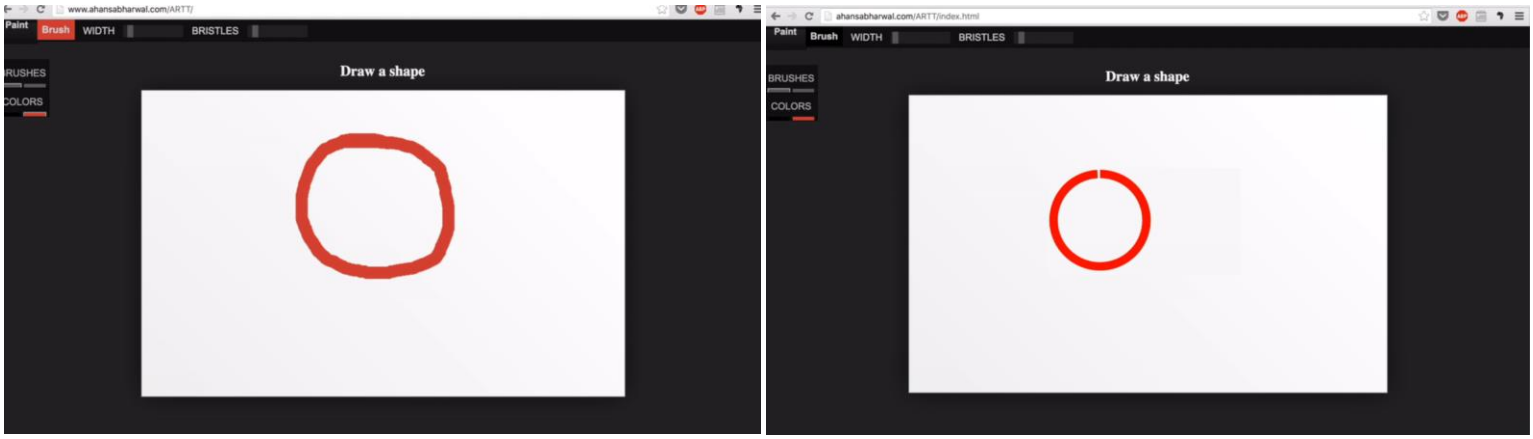


Figure 2 - Human's drawing (left) vs bot's drawing (right);
ARTT correctly identified the user in this case

Objective of this research

This research study tested whether ARTT would be able to successfully accept and identify all user-input accurately as human or machine. The fact is that each human thinks differently, resulting in creativity, progress, and development. Each person is born into a unique set of circumstances – time, place, family, natural affinities, intrinsic motivations, attraction to objects, activities and people – which shape his/her thought processes. All these factors make a human unique, which leads to different interpretations of a text-based or drawing-based prompt.

ARTT is required to successfully identify human-drawn responses versus machine-drawn responses, the accuracy of which is assessed in this research paper.

INTRODUCTION

The Art-Related Turing Test (ARTT) platform provides a clean and elegant drawing space, which allows users to choose different colors, brushes, bristle sizes, etc. After the user responds to the drawing prompt, the image drawn (by the user) is uploaded to the server. At the server end, the image is processed and analyzed by the Clarifai API, which has specifically been adapted for drawings rather than for photographs.

The AI-based Clarifai API contextualizes and views different parts of the image separately, and then visualizes the full drawing, as a whole. Figure 2 shows how ARTT analyzes separate parts of the image. The graphics in red/pink below indicate the analysis taking place.

Figure 3: ARTT analyzing parts of image

ARTT then generates tags appropriate to the image. These dynamically-formulated tags are equated to the prompt presented to the user. If they match, ARTT concludes that the user is

human. Figure 3 shows ARTT's 'input-compute-predict' cycle, which is the basis of how ARTT receives, understands and responds to the user's drawing.

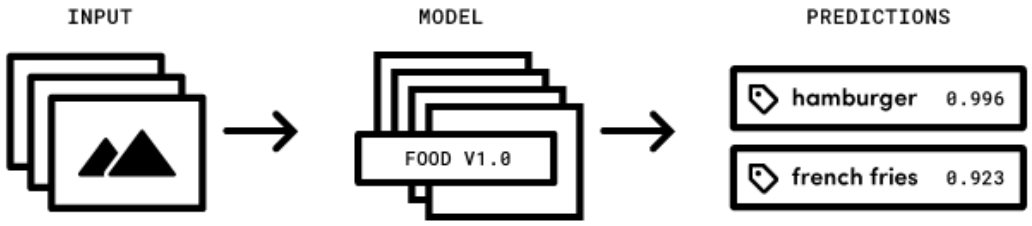


Figure 4 - input-compute-predict cycle for ARTT

The following is an extract from ARTT's code to check if one of the returned tags equals the required answer:

```
var responseTags;
          var isHuman = false;
          // get tags with an array of images
          var URL = 'http://ahan.io/ARTT-prototype/' + filename; // input
          function getTags() { // compute
            console.log (URL);

            Clarifai.getTagsByUrl(URL).then(
              function(response){
                responseTags = response.results[0].result.tag.classes;
                for(var i = 0; i<responseTags.length; i++){
                  console.log(responseTags[i]);
          if(responseTags[i] == "shape" || responseTags[i] == "figure"){
                    isHuman = true;
                  }
                }
                if(isHuman == true){
                  alert("You're a human"); // PREDICT
                }
                if(isHuman == false){
                  alert("You're a bot."); // PREDICT
                }
              },
              function(error){
                console.log("SOMETHING WENT WRONG");
              });
            }
```

METHODS

To test whether the ARTT system was plausible, participants were asked to read and respond to the prompts shown on ARTT. The prompts considered for this experiment are listed below. Although two example prompts are shown here, it is to be noted that a wide range of prompts could be set up for users in the future.

- Experiment 1 (exp. 1): “Draw the device that tells time”
- Experiment 2 (exp. 2): “Draw a (any) shape”

The users were then required to choose a color from the color palette, as shown below:

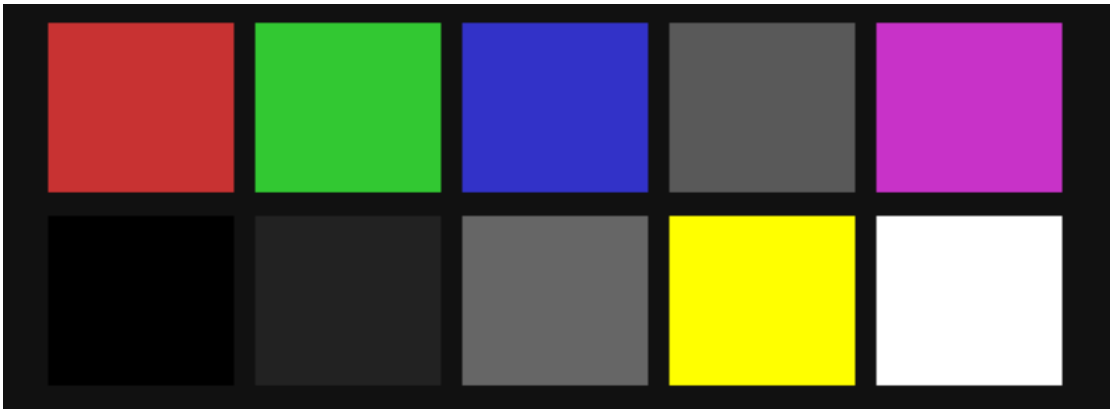
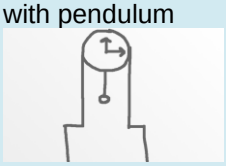


Figure 4 - ARTT color palette

After choosing a color (figure 4), users drew an image (on the computer) that they thought was an appropriate response to the prompt. The participants’ choice of color greatly affects the accuracy of ARTT, since ARTT’s artificial intelligence is coded to identify and correctly determine and understand color. ARTT then analyzed the user-response and concluded whether the drawing was created by a bot or a human. An external observer timed this entire process.

RESULTS

Table 1: Results of Exp.1 ("Draw the instrument that tells time")

Participant	Color chosen	User Response (image)	Time (seconds)	Recognized as Human?
1	Black	Wristwatch 	25	Yes
2	Red	Wristwatch 	23	Yes
3	Green	Only analog face 	21	No
4	Black	Only analog face 	26	Yes
5	Blue	Only hands of clock 	17	No
6	Grey	Grandfather clock with pendulum 	39	Yes
7	Yellow	Only analog face	29	Yes

Discussion:[3]

In experiment 1, the most commonly returned tags were: `watch`, `clock`, `hands`, `figure`, `wrist`, `wrist watch`, `timer`, `stopwatch`, `time piece`, `analogue`.

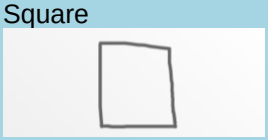
These results go to show that ARTT is not perfect: it failed to correctly recognize humans in two situations.

1) Participant 3 - It is believed that this case failed because of the color selected (green). ARTT returned tags such as `greenery` in this case.
2) Participant 5 - This case is believed to have failed due to the ambiguity of the drawing (user response).

Overall, ARTT did correctly recognize and accept the different interpretations of the prompt, including analog face, grandfather clock, and wristwatch.

[3] It cannot definitively be said why ARTT failed in some cases and did not in others due to the independent nature of Artificial Intelligence.

Table 2: Results of Exp.2 (“Draw a (any) shape”)

Participant	Color chosen	User Response (image)	Time (seconds)	Recognized as Human?
1	White	Circle	11	Yes
2	Green	Circle 	9	Yes
3	Grey	Square 	16	Yes
4	Pink	Triangle 	14	Yes
5	Black	Circle 	10	Yes
6	Red	Pentagon 	20	Yes
7	Red	Triangle and Circle 	18	Yes

Discussion:[4]

In experiment 2, the most common tags returned were: `shape`, `symbol`, `line`, `illustration`.

In this experiment, ARTT succeeded in correctly identifying human users in each of the participant trials. This is probably due to the comparatively easier prompt to "draw a shape" – the AI can comprehend and analyze simpler user drawings better than larger and more complicated ones. Therefore, from these findings, it is clear that ARTT should ask users to draw simpler objects, which are still vague and difficult enough to hinder bots.

From the experiment, it can be seen that ARTT adapted correctly to different colors, number of shapes and types of shapes. This experiment proves that ARTT's AI understands user input dynamically.

[4] It cannot definitively be said why ARTT failed in some cases and did not in others due to the independent nature of Artificial Intelligence.

CONCLUSION

The Art-Related Turing Test (ARTT) works quite effectively as a CAPTCHA, identifying human input accurately in most cases. The challenge for bots in this case is to interpret drawing instructions and to respond with a relevant drawing, with human-like imperfections in the output. A more detailed study needs to be conducted to compare the accuracy of ARTT with that of current CAPTCHAs.

One of the challenges of this approach remains that it is quite 'involved', so many users may find it quite onerous for accessing simple applications/ websites. For more security-sensitive applications, e.g., banking, users may be more inclined to accepting such an approach. Additionally, it remains to be seen how quickly bot technology can overcome the challenge of interpreting and drawing a relevant response. As always, in the area of cyber-security and cryptography, the technology needs to stay ahead of that developed by the security hackers.

As AI becomes ever more sophisticated, the gap between what humans can do and what bots can achieve is becoming ever smaller. For now, ARTT uses the superior cognitive and creative skills of humans (including their imperfect execution/ drawing skills) to authenticate users. To ensure a level playing field, and more consistent results, the system can be set up to limit user drawings to 'simpler' output, such as by imposing an upper limit on the number of 'brush strokes' and colors. As demonstrated in this study, simpler drawings have a higher probability of accurate validation by the ARTT system. Additionally, the color palette may be modified to remove colors that may negatively influence test results for each user-prompt.

Further testing with bots may reveal some additional features of the system that need to be modified. The initial tests reveal that ARTT is a promising method of authenticating valid, human users and avoiding automated scripts / hacks.

REFERENCES

Ahn, L. V., Blum, M., Hopper, N. J., & Langford, J. (n.d.). CAPTCHA: Using hard AI problems for security. *CAPTCHA*. Retrieved from http://www.captcha.net/captcha_crypt.pdf

This source shows the similarities between AI and CAPTCHAs, as well as shows how the advancement of one can result in the evolution of the other.

Angre, A. R., Kapadia, M. D., & Ugale, M. (2015). PiCAPTion: Picture captchas for internet authentication. *International Journal of Computer Applications*, *114*(10), 6-9. Retrieved from http://research.ijcaonline.org/volume114/number10/pxc3901976.pdf

This explains the image CAPTCHAs and how they can be easily cracked but how they protect the security of the users.

Brown, B. (2013, November 7). Researchers dare AI experts to crack new 'gotcha' password scheme; Like captcha, gotcha's inkblot password system relies on humans' visual skills. *LexisNexis Academic*.

This source gave us important information about how AI was able to break CAPTCHAs.

Clarifai. (n.d.). Retrieved June 22, 2016, from Clarifai website: https://www.clarifai.com/#demo and https://www.clarifai.com/developer/guide/

This site is a great demo about an AI that can be easily manipulated to break CAPTCHAs and also contains images that explain the AI process.

Minor, J. (n.d.). AI startup develops captcha-cracking software. *PC Magazine*. Retrieved from ProQuest database.

This source shows how AI can crack CAPTCHAs and how both CAPTCHAs and AI must keep advancing to stay ahead of each other.

Every Brain Is Unique - Maren Schmidt. Link:marenschmidt.com/2010/05/every-brain-is-unique

Ollmann, G. (2008). The evolution of commercial malware development kits and colour-by-numbers custom malware. *Computer Fraud and Security*, *2008*(9), 4-7. Retrieved from http://www.sciencedirect.com/science/article/pii/S1361372308701350

This source informed us about the monetary gains that come from producing bots that hack CAPTCHAs and other advantages to having AI that can bypass CAPTCHAs.

Pope, C., & Kaur, K. (n.d.). Is it human or computer? Defending e-commerce with captchas. *IEEE*, 43-49. Retrieved from http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1425425

This source gives a very simple overview about CAPTCHAs and the overall impact of CAPTCHAs on society and the economy.

Thomas, V. A., & Kaur, K. (2013). Cursor CAPTCHA – captcha mechanism using mouse cursor. *International Journal of Computer Applications*, *67*(22), 13-17. Retrieved from http://research.ijcaonline.org/volume67/number22/pxc3887253.pdf

This source explains the new Google reCAPTCHA and how it monitors the cursor in how it moves toward the box and browses the browser history of the computer.

Zhu, B. B., Yan, J., Bao, G., Yang, M., & Xu, N. (2014). Captcha as graphical passwords—a new security primitive based on hard AI problems. *IEEE Transactions on Information Forensics and Security*, *9*(6), 891-904. Retrieved from http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6775249&tag=1

This source went into depth about graphical passwords which was necessary to our research, especially when making a new and improved CAPTCHA.