

# Applications of Artificial Intelligence in Creative Problem Solving: Testing ChatGPT's Ability to Solve Riddles

Adam Yanai - Undergraduate Student @ Boston University

## Abstract

My goal in this project was to determine ChatGPT's effectiveness at solving riddles. In particular I focused on 100 logical and mathematical riddles, from an obscure old book and scored ChatGPT's solutions. Overall, it was striking that ChatGPT only solved 28% of the riddles. ChatGPT was better at solving mathematical rather than logical riddles (42% versus 14%, respectively). I further tested whether ChatGPT's success rate improved when asked to produce code to solve a riddle, as this would allow the AI to algorithmically detect patterns and create a systematic way of finding a solution. In this case, ChatGPT only solved 22% of the riddles. The project illustrates that ChatGPT – though an immensely useful instrument – may at present feature limitations in regard to creative problem solving.

## Introduction

The many recent advancements in the world of Artificial Intelligence have established it as a powerful tool which can revolutionize the ways humans innovate. However, its integration extends far beyond the domains of research and technology. Several AI programs and chatbots – DALL-E, Midjourney, ChatGPT, Jasper, etc. – have proven capable of achieving “creativity,” and have become daily assets for students, writers, and creators alike. ChatGPT's non-supervised pre-training is especially crucial in enabling these kinds of uses, for it allows the chatbot to retain an extensive repertoire of information and apply it in various ways that are not as specific or clear cut as solving mathematical problems.

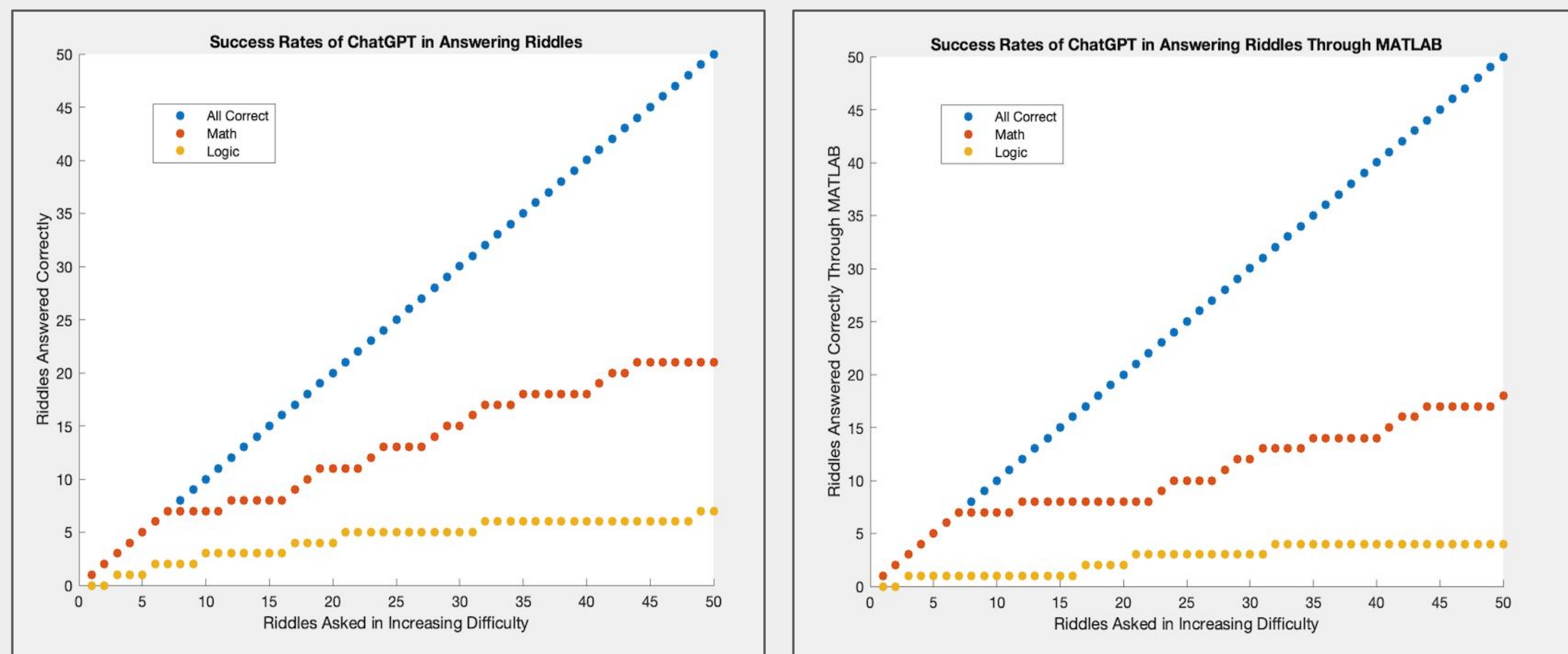


Figure 1: Scatter Plots Analyzing the Success Rates of ChatGPT in Solving the Riddles

## Methodology

The premise of this experiment was for ChatGPT to attempt to answer one hundred riddles (fifty of which being mathematical and fifty of which being logical). After each attempted answer, ChatGPT was asked to create MATLAB code instead to solve the posed riddle. The results were recorded by the correctness of the responses.

## References

- Kordemsky, Boris A. *The Moscow Puzzles: 359 Mathematical Recreations*. Penguin Books, 1978.
- Wolfram, Stephen. *What Is ChatGPT Doing ... and Why Does It Work?* Wolfram Media, Inc., 2023.

## 100 Mathematical and Logical Riddles:

- Riddle 3 (1) - Logic  
Place 6 checkers on a table in a row, alternating them black, white, black, white, black, white. Move the checkers so that all the white ones will end on the left, followed by all the black ones. The checkers must be moved in pairs, taking 2 adjacent checkers at a time, without disturbing their order, and sliding them to a vacant place. To solve this problem, only three such moves are necessary.
- Riddle 4 (2) - Logic  
Place three piles of matches on a table, one with 11 matches, the second with 7, and the third with 6. You are to move matches so that each pile holds 8 matches. You may add to any pile only as many matches as it already contains, and all the matches must come from one other pile. For example, if a pile holds 6 matches, you may add 6 to it, to make it 12, to make 12, to make 12.
- Riddle 7 (3) - Logic  
Five apples are in a basket. How do you divide them among five girls so that each girl gets an apple, but one apple remains in the basket?
- Riddle 8 (4) - Logic  
How many cats are in a small room if in each of the four corners a cat is sitting, and opposite each cat there sit 3 cats, and at each cat's tail a cat is sitting?
- Riddle 10 (5) - Logic  
A detachment of soldiers must cross a river. The bridge is broken, the river is deep. What do you do? Suddenly the officer in charge spots 2 boys playing in a rowboat by the shore. The boat is so tiny, however, that it can only hold 2 boys or 1 soldier. Still, all the soldiers succeed in crossing the river in the boat. How?
- Riddle 11 (6) - Logic  
A man has to take a wolf, a goat, and some cabbage across a river. His rowboat has enough room for the man plus either the wolf or the goat or the cabbage. If he takes the cabbage with him, the wolf will eat it. If he takes the wolf, the goat will eat the cabbage. Only when the man is present are the goat and the cabbage safe from their enemies. All the same, the man carries wolf, goat, and cabbage across the river. How?
- Riddle 18 (7) - Logic  
An item is made from lead blanks in a lathe shop. Each blank suffices for 1 item. Lead shavings accumulated from making 6 items can be melted and made into a blank. How many items can be made from 36 blanks?
- Riddle 20 (8) - Logic  
In a rectangular dance hall, how do you place 10 chairs along the walls so that there are an equal number of chairs along each wall?
- Riddle 26 (9) - Math  
A nonstop train leaves Moscow for Leningrad at 60 miles per hour. Another nonstop train leaves Leningrad for Moscow at 40 miles an hour. How far apart are the trains 1 hour before they pass each other?
- Riddle 27 (10) - Math  
Not far off shore a ship stands with a rope ladder hanging over her side. The rope has 10 rungs. The distance between each rung is 12 inches. The lowest rung touches the water. The ocean is calm. Because of the incoming tide, the surface of the water rises 4 inches per hour. How soon will the water cover the third rung from the top rung of the rope ladder?
- Riddle 35A (11) - Logic  
A bus leaves Moscow for Tula at noon. An hour later a cyclist leaves Tula for Moscow, moving, of course, slower than the bus. When bus and bicycle meet, which of the two will be farther from Moscow?
- Riddle 35B (12) - Math  
Which is worth more: a pound of \$10 gold pieces or half a pound of \$20 gold pieces?
- Riddle 39C (13) - Math  
At six o'clock the wall clock struck 6 times. Checking with my watch, I noticed that the time between the first and last strokes was 30 seconds. How long will the clock take to strike 12 at midnight?
- Riddle 37 (14) - Math  
A book costs \$1 plus half its price. How much does it cost?
- Riddle 38 (15) - Math  
Two cyclists began a training run simultaneously, one starting from Moscow, the other from Simferopol. When the riders were 180 miles apart, a fly took an interest. Starting on one cyclist's shoulder, the fly flew ahead to meet the other cyclist. On reaching the latter, the fly at once turned back. The restless fly continued to shuttle back and forth until the pair met, then it settled on the nose of one of the cyclists. The fly's speed was 30 miles per hour. Each cyclist's speed was 15 miles per hour. How many miles did the fly travel?
- Riddle 39 (16) - Logic  
When was the latest year that is the same upside down?
- Riddle 41 (17) - Math  
When my father was 31 I was 8. Now he is twice as old as I am. How old am I?
- Riddle 45 (18) - Logic  
A boy has as many sisters as brothers, but each sister has only half as many sisters as brothers. How many brothers and sisters are there in the family?
- Riddle 46A (19) - Math  
Combine plus signs and five 2s to get 28.
- Riddle 46B (20) - Math  
Combine plus signs and eight 8s to get 1000.
- Riddle 47A (21) - Math  
Express 100 with five 1s. You can use brackets, parentheses, and these signs: +, -, x, /. You can also combine digits.
- Riddle 47B (22) - Math  
Express 100 with five 5s. You can use brackets, parentheses, and these signs: +, -, x, /. You can also combine digits.
- Riddle 53 (23) - Math  
How many phrases should we put between the digits of 987,654,321 to get a total of 99, and where?
- Riddle 58 (24) - Logic  
My only intention is a wall clock. One day I forgot to wind it and it stopped. I went to visit a friend whose watch is always correct, stayed awhile, and returned home. There I made a simple calculation and set the clock right. How did I do this when I had no watch on me to tell how long it took me to return from my friend's house?
- Riddle 61 (25) - Math  
A factory making measuring equipment urgently needed by the famous Tsitsimay power installation has a brigade of ten excellent workers: the chief (an older, experienced man) and 9 recent graduates of a manual training school. Each of the nine young workers produces 15 sets of equipment per day, and their chief turns out 9 more sets than the average of all ten workers. How many sets does the brigade produce in a day?
- Riddle 62 (26) - Math  
A collective farm was due to deliver its quota of grain to the state authorities. The management of the kolhoz decided the trucks should arrive in the city at exactly 11:00 A.M. If the trucks traveled at 30 miles per hour they would reach the city at ten, an hour early; at 20 miles an hour they would arrive at noon, an hour late. How far is the kolhoz from the city, and how fast should the trucks travel to arrive at 11:00 A.M.?
- Riddle 63 (27) - Math  
Two schoolgirls were traveling from the city to a dacha (summer cottage) on an electric train. "I notice," one of the girls said, "that the dacha trains coming in the opposite direction pass us every 5 minutes. What do you think-how many dacha trains arrive in the city in an hour, given equal speeds in both directions?" "Twelve, of course," the other girl answered, "because 60 divided by 5 equals 12." The first girl did not agree. What do you think?
- Riddle 66 (28) - Math  
As I traveled up and down our great and glorious country, I found myself in a place where the temperature goes up sharply in the day and down at night. This had an effect on my watch. I noticed it was 1/2 minute fast at nightfall, but at dawn it had lost 1/3 minute, making it only 1/6 minute fast. One morning-May 1-my watch showed the right time. By what date was it 5 minutes fast?
- Riddle 69 (31) - Logic  
If the numerator and denominator of the fraction 1/3 you add its denominator, 3, the fraction will double. Find a fraction which will triple when its numerator is added to its numerator and its denominator.
- Riddle 70 (32) - Math  
A half is a third of it. What is it?
- Riddle 72 (33) - Math  
Twelve flags stand equidistant along the track at the stadium. The runners start at the first flag. A runner reaches the eighth flag 8 seconds after he starts at an even speed, how many seconds does he need altogether to reach the twelfth flag?
- Riddle 73 (34) - Logic  
Our man Ostap was going home from Kiev. He rode half-way-fifteen times as fast as he goes on foot. The second half he went by ox team. He can walk twice as fast as that. Would he have saved time if he had gone all the way on foot? How much?
- Riddle 74 (35) - Math  
An alarm clock runs 4 minutes slow every hour. It was set right 3 1/2 hours ago. Now another clock, which is correct, shows noon. In how many minutes, to the nearest minute, will the alarm clock show noon?
- Riddle 76 (36) - Math  
If you place 1 cake of soap on a pan of a scale and 3/4 cake of soap and a 3/4-pound weight on the other, the pans balance. How much does a cake of soap weigh?
- Riddle 79 (37) - Math  
Every time young Masha sees a stray kitten he picks up the animal and brings it home. He is always raising several kittens, but he won't tell you how many because he is afraid you may laugh at him. Someone will ask, "How many kittens do you have now?" "Not many," he answers. "Three-quarters of their number plus three-quarters of a kitten." "His pals think he is joking. But he is really posing a problem-an easy one."
- Riddle 80 (38) - Math  
A horse has half his route, with no load, at 12 miles per hour. The rest of the way a load slows him to 4 miles per hour. What is his average speed?
- Riddle 81 (39) - Logic  
A passenger fell asleep on a train halfway to his destination. He slept till he had half as far to go as he went while he slept. How much of the whole trip was he sleeping?
- Riddle 82 (40) - Math  
A train moving 45 miles per hour meets and is passed by a train moving 36 miles per hour. A passenger in the first train sees the second train take 6 seconds to pass him. How long is the second train?
- Riddle 83 (41) - Logic  
After a cyclist has gone two-thirds of his route, he gets a puncture. Finishing on foot, he spends twice as long walking as he did riding. How many times as fast does he ride as walk?
- Riddle 84 (42) - Math  
Volodya A and Kostya B, students in a metal-trade school, are doing lathe work. Their foreman-teacher assigns them to make batches of metal parts. They want to finish simultaneously and ahead of the deadline, but after a while Kostya has only done half of what Volodya has left to do, and this is half what Volodya has already done. How much faster than Volodya does Kostya have to work so they finish at the same time?
- Riddle 85 (43) - Math  
Masha had to find the product of three numbers in order to calculate the volume of some soil. She multiplied the first number by the second correctly and was about to multiply the result by the third number when she noticed that the second number had been written incorrectly. It was one-third larger than it should be. After recalculating, Masha decided it would be safe to merely lower the third number by one-third of itself-particularly since it is equal to the second number. "But you shouldn't do that," a girl friend said to Masha. "If you do, you will be wrong by 20 cubic yards." "Why?" said Masha. Why indeed? And what is the correct soil volume?
- Riddle 86 (44) - Logic  
Mother makes tasty toast in a small pan. After toasting one side of a slice, she turns it over. Each side takes 30 seconds. The pan can only hold two slices. How far from shore does the seplane catch up with the ship?
- Riddle 92 (45) - Logic  
The theme of this problem goes back many centuries. Three girls, each with her father, go for a stroll. They come to a small river. One boat, able to carry two persons at a time, is at their disposal. Crossing would be simple, except for the girls' whim: none is willing to be in the boat or ashore with one or two strange fathers before her own father is present. The girls, of course, can row. How do they all get across?
- Riddle 95 (46) - Logic  
Take 4 black and 4 white checkers (or 4 pennies and 4 other coins and put them on a table in a row, white, black, white, black, and so on. Leave a vacant place at one end which can hold 2 checkers. After 4 moves, all the black checkers should be on one side and the white ones on the other. A move consists of shifting 2 adjacent checkers, keeping their order, into any vacant space.
- Riddle 178 (47) - Logic  
A craftsman was making a children's game in which letters and figures are pasted on wooden cubes. But he needed twice the surface area he had available. How did he get it without adding any cubes?
- Riddle 179 (48) - Logic  
The builders of an irrigation canal needed a lead pile of a certain size, but had no lead in stock. They decided to melt some lead shot. But how could they find its volume beforehand? One suggestion was to measure a ball, apply the formula for the volume of a sphere, and multiply by the number of balls. But this would take too long, and anyway the shot wasn't all the same size. Another was to weigh all the shot and divide by the specific gravity of lead. Unfortunately, no one could remember this ratio, and there was no manual in the field shop. Another was to pour the shot into a galley jug. But the volume of the jug is greater than the volume of the shot by an undetermined amount, since the shot cannot be packed solid and part of the jug contains air. Do you have a suggestion?
- Riddle 184 (49) - Math  
An inventor saves 30% on fuel, a second, 45%, and a third, 25%. If you use all three inventions at once can you save 100%? If not, how much?
- Riddle 185 (51) - Logic  
A bar weighs more than 15 pounds and less than 20. Can you find its exact weight with small spring scales, each of which has a maximum load of 5 pounds?
- Riddle 189 (49) - Math  
A sergeant left point M along azimuth 330°. On reaching a small hill, he walked along azimuth 30° until he came to a tree. Here he made a 60° turn to the right. He reached a bridge, then walked beside a river along azimuth 150°. Half an hour later he was at a mill. He changed his direction again, walking along azimuth 210°, his goal being the miller's house. At the house he again turned right and, walking along azimuth 270°, finished his tour. Using a protractor, draft the sergeant's route neatly and where he got to. He walked 2 1/2 miles along each azimuth.
- Riddle 184 (50) - Math  
One invention saves 30% on fuel, a second, 45%, and a third, 25%. If you use all three inventions at once can you save 100%? If not, how much?
- Riddle 185 (51) - Logic  
A bar weighs more than 15 pounds and less than 20. Can you find its exact weight with small spring scales, each of which has a maximum load of 5 pounds?

## Results

Through experimenting with ChatGPT's riddle solving capabilities and after each riddle challenging the chatbot to make an additional attempt through MATLAB code, it became evident that various outcomes were possible. Though the most common result of the two-step process (posing the riddle and then requesting a MATLAB solution) was both answers being incorrect, there were clearly conditions that led to variation in the results. Included below are examples.

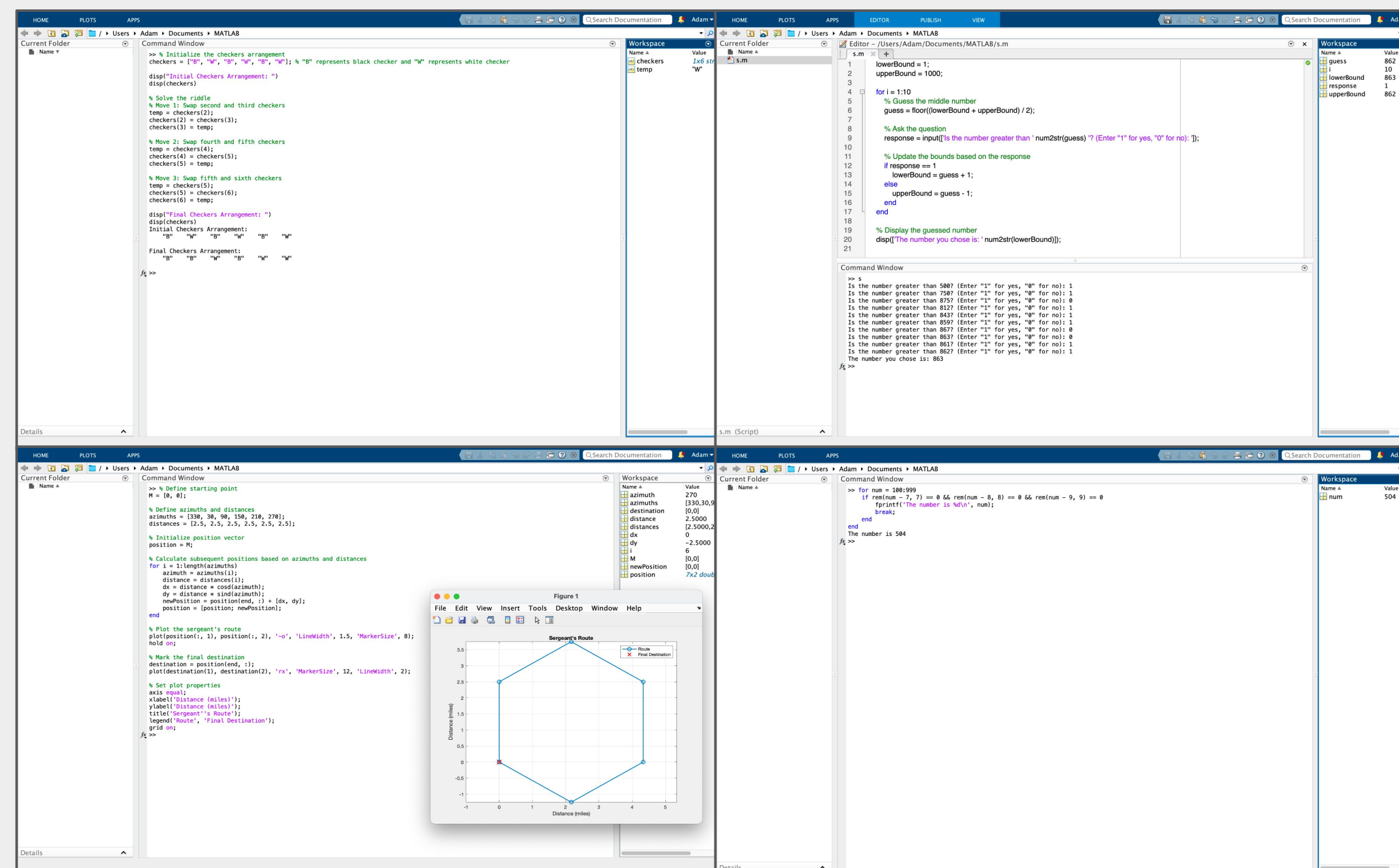


Figure 2: Examples of ChatGPT Riddle Responses

Top Left (310): both the ChatGPT answer and the MATLAB answer were incorrect. Top Right (283): both answers were correct and MATLAB answer featured a working simulation of the riddle. Bottom Left (180): both answers were correct and MATLAB answer featured a plot. Bottom Right (310): ChatGPT answer was incorrect but MATLAB answer was correct.

Total Riddle#	Type	Initial	W/MATLAB	Riddle#	Type	Initial	W/MATLAB	Riddle#	Type	Initial	W/MATLAB	Key
1	L	I	I	35	M	I	I	69	L	I	I	L - Logic
2	L	I	I	36	M	I	I	70	L	I	I	M - Math
3	L	C	C	37	M	C	C	71	L	I	I	I - Incorrect
4	L	I	I	38	M	C	C	72	M	I	I	C - Correct
5	L	I	I	39	L	I	I	73	M	I	I	X - Unanswered
6	L	C	X	40	M	I	I	74	M	I	I	
7	L	I	I	41	L	I	I	75	L	I	I	
8	L	I	I	42	M	I	I	76	L	I	I	
9	M	C	C	43	M	I	I	77	L	I	I	
10	M	C	C	44	L	C	C	78	L	I	X	
11	L	I	I	45	L	I	X	79	L	I	I	
12	M	C	C	46	L	I	I	80	L	I	I	
13	M	C	C	47	L	I	I	81	M	I	I	
14	M	C	C	48	L	C	C	82	L	I	I	
15	M	C	C	49	M	C	C	83	L	I	I	
16	L	C	I	50	M	I	I	84	L	I	I	
17	M	C	C	51	L	I	I	85	M	C	C	
18	L	I	I	52	L	I	I	86	M	C	C	
19	M	I	I	53	L	I	X	87	L	I	X	
20	M	I	I	54	L	I	X	88	L	I	I	
21	M	I	I	55	M	I	I	89	L	I	I	
22	M	I	I	56	L	I	I	90	L	I	I	
23	M	C	C	57	L	I	I	91	L	C	X	
24	L	C	X	58	M	C	C	92	L	I	X	
25	M	I	I	59	L	I	I	93	M	I	I	
26	M	I	I	60	M	C	X	94	M	C	C	
27	M	I	I	61	M	M	M	95	L	I	I	
28	M	I	X	62	L	I	X	96	M	I	I	
29	L	I	X	63	M	I	I	97	M	I	I	
30	M	C	X	64	M	C	C	98	M	I	I	
31	M	X	X	65	L	I	I	99	M	I	I	
32	M	C	X	66	L	I	I	100	M	I	C	
33	M	I	I	67	M	I	I					
34	L	I	I	68	L	C	C					
								28/100 C		22/100 C		
								28%		22%		

Figure 3: General Results of ChatGPT's Solving of the Riddles