Non-Trivial Questions

Spyridon Manolidis

January & February 2025

1 Introduction

I have recently been looking for research opportunities as an aspiring physicist and mathematician. I found Non-Trivial, a nonprofit that specialises in helping young people make a difference and produce unprecedented research.

Unfortunately, I got rejected from the fellowship, but I feel I learned a lot from the application process as all the questions were quite perplexing but also incredibly interesting. I thought I'd share my thought process for each question for anyone interested.

IMPORTANT: Many of the quoted sections below are taken directly from the Non-Trivial Fellowship Application for 2025 - I did not come up with these problems myself, all credit should go to them.

2 Question 1 - A Modern Turing Test

I was prompted to 'write up to 140 characters that our human reviewer will recognize as human but our AI reviewer will classify as AI-generated.'

I wanted to tackle this problem using cryptography. I thought of creating a poem or something similar which would be made by AI, so it would pass the AI test, but the AI would be given very specific instructions to create a poem which actually contained a code.

The plan was to hide the message 'IAMHUMAN' within the poem by having each line represent one letter. The way this would be done would be by taking the first letter of each word on each line, and assigning it a dot if a vowel, and a dash if a consonant. The poem could then be recreated with each line being a sequence of dots and dashes, which were then corresponding letters in morse code.

The AI generated a poem with this within the first few attempts, but the problem was that the poem would be over 140 characters, thus not a valid response.

To overcome this problem, I decided to only hide the message 'HUMAN' within the poem, and I instructed the AI with the following:
'Can you please create me a poem where...
The first line has four words each beginning with a vowel
The second line has two words beginning with a vowel, then one with a consonant
The third line has two words beginning with a consonant
The fourth line has one word beginning with a vowel, and then one beginning with a consonant
The fifth and last line has a word beginning with a consonant, and then one

beginning with a vowel.
Bonus points if you can chuck the word morse in there.'

The AI then subsequently generated the poem:
'Here is a poem with the requested line structures:

Eager eyes, ancient oak
Icy, eerie night
Moonlit shadows
A whisper
Lingers, echoing'

Using the method above, converting the first letter of each word on each line to a dot if a vowel or to a dash if a consonant gives '..../..-/--/.-/-.' which, using a morse code table, translates to 'HUMAN.'

Then, as a final test I put the above into many AI detectors to ensure that it AI could detect that the poem was written by an AI. Many detectors required more than 140 characters, so I simply duplicated the poem as many times as I needed to and then most detectors recognised it as AI.

Thus, the above was my final solution.

3 Question 2 - Racing to the brink

In my opinion question 2 was by far the most interesting. All applicants were to be put in a game: 'You and four others submit a number. We subtract the sum of all five numbers from 50. This represents the possibility of harm from exploration. Then, we multiply the resulting number by your submission – that's your score. Your goal is to maximize your score.'

So in other words, you must pick a number, X, and your score is calculated using the following formula:

$$\text{Score} = X \cdot (50 - \sum_{i=1}^{4}(X_i) - X) \tag{1}$$

Where X_1, X_2, X_3 and X_4 are the number picked by the other players.

I have already published my solutions for this problem, but I'll reiterate...

3.1 The No-Risk Solution

Consider the following scenarios:

1. If $\sum_{i=1}^{4}(X_i)$, the sum of everyone else's numbers, is a very large positive number then $(50 - \sum_{i=1}^{4}(X_i) - X)$ becomes a very large negative number. This means that we could pick a negative number which would allow us to get a high score. However, choosing too large a number risks making $(50 - \sum_{i=1}^{4}(X_i) - X)$ positive, thus giving us a negative score.

2. Similarly, If $\sum_{i=1}^{4}(X_i)$ is a very large negative number, $(50 - \sum_{i=1}^{4}(X_i) - X)$ becomes a very large positive number. Thus, we could pick a positive number to get a high score, but if our number is too big, $(50 - \sum_{i=1}^{4}(X_i) - X)$ turns negative and we risk having a negative score yet again.

Using graphical methods will show you the same thing, that is: either you will always get a negative score or it will peak at a point and then collapse to negativity. Therefore, because the probability of having a negative number is greater than the probability of having a positive number, the best answer using this logic should be **0**, as this ensures that we never get a negative number.

3.2 The Altruistic Approach

Considering that this problem was for an application for an altruistic company, then we should all demonstrate our altruism by taking a risk and trying to maximise all of our scores.
If we all pick the same number, X, then our scores become $X \cdot (50 - 5X)$, which is a simple quadratic $(-5X^2 + 50X)$. This has an easily calculable maximum at $(5, 125)$, which means that we should all pick **5** to give us all a score of 125.

3.3 Considering All Possibilities Using Sums

Let's start by defining a function $f(x, y)$ which represents our score. Here, x represents the number we chose and $y = \sum_{i=1}^{4}(X_i)$, the sum of everyone else's numbers. Thus, using the information given to us in the question, we can deduce:

$$f(x, y) = x \cdot (50 - x - y) \tag{2}$$

We can set a constant y and see how varying x affects our score. For example, setting $y = 5$ will give us $f(x, 5)$, which is $x \cdot (50 - x - 5)$. No matter what we change y to, we should always get a graph that is a negative parabola.
But we want to consider all the numbers everyone else chooses - people are random and can have some pretty wild choices!
The way we would do this is by summing up all of the functions at different values of y (e.g. $(f(x, -2) + f(x, -1) + f(x, 0) + f(x, 1) + f(x, 2) + ...)$). This would give us a function $g(x)$ whose maximum (if it has one) would be the optimal solution, as it would take into account all the functions.

Let's say we want g for now to only consider the functions where y is an integer and $-10 \leq y \leq 10$. We can define g as:

$$g(x) = \sum_{i=-10}^{10} f(x,i) = \sum_{i=-10}^{10} (x \cdot (50 - x - i)) \tag{3}$$

Notice that the lower and upper bounds of y are the lower and upper bounds of the summation. Thus, we can deduce that for $y : y \in \mathbf{Z}, -u \leq y \leq u$, g is:

$$g(x) = \sum_{i=-u}^{u} f(x,i) = \sum_{i=-u}^{u} (x \cdot (50 - x - i)) \tag{4}$$

Putting this into a professional tool such as Wolfram Alpha, Equation 4 becomes:

$$g(x) = -x \cdot ((x - 50) \cdot (2u + 1)) \tag{5}$$

We see that for any (positive) u, there is a maximum at $x = 25$, even as u approaches infinity ($\lim_{u \to \infty}$). Thus, we can at last conclude that the optimal solution must be **25**.[1]

3.4 My Judgment

I originally ended up picking **5** because I feel it was a real way to demonstrate my altruism. However, I changed my mind because I felt the solution in Section 3.3 demonstrated creative & analytical mathematical ability, so I changed my answer to **25**.

4 Question 3 - Digital Trolley Problem

The original question is quite long, so I'll summarise it here.

You are given some background information first. You are told that it is 2045, and that scientists have achieved Whole Brain Emulation for humans. This is a working copy of a brain in a computer and the emulated brains show near-identical responses to their biological originals when given the same stimuli. In addition, the emulation can replicate all neurons and their connections, as well as brain chemistry.

You are then given the following scenario: 'A runaway trolley is hurtling down a track toward a human worker. You could flip a switch to divert it toward a

[1] An obvious flaw in the logic here is that we are only considering discrete steps of y as u is limited to integer values. It is left as an exercise to the reader to confirm that this wouldn't make a difference - as the function f is continuous and quadratic (for any $\mathbf{R}$ substitution of y), so any decimal values would yield results between the closest two integers.

data center housing X human brain emulations. No back-ups of the emulations are available and the people whose brains were scanned died in the process.'

You are asked 'Is there a number X where you would choose to sacrifice the human worker to save the emulations?' and 'what's the smallest such X where you would make this choice?'

To this, I gave the following response:
'If we value the emulations the same as a human life, the human and their descendants' exponential reproduction (2^x) will always exceed the number of emulations lost, given enough time. Whilst some emulated brains, such as the brain of a great scientist, might seem more valuable, there could be an even greater mind a few generations away. Without the human, this entire bloodline could not exist. Plus, while supercomputers might mimic thought and brain chemistry might emulate emotion, two defining factors of humanity, they lack free will in thought. What I mean by this is that their thoughts can be filtered and controlled. This, in my eyes, removes a critical aspect of the creativity of humanity, and declassifies them as humans. Also, natural death protects us from the indefinite reigns of dictators and autocrats, so perhaps it would be best to destroy all brain emulations to shield humanity. Lastly, resources are finite and I believe the living have priority, the dead have used theirs.'

Well that's it! I really enjoyed these questions so I'm very much looking forward to finding more brainteasers like these. If you have any, please don't hesitate to contact me and as always, thanks for reading.