Challenging Mathematical Problems Designed to Evaluate Advanced AI Reasoning

Author: Prof. Dr. Tan Kwan Hong (PhD DBA DPA EdD DProf)

Affiliation: Professor, Central Global University

Contact: tankwanhong@gmail.com

Abstract

This paper presents six carefully constructed mathematical problems, each intended to challenge the symbolic, spatial, or logical reasoning capabilities of advanced artificial intelligence (AI) models. These problems span various mathematical disciplines, including number theory, algebra, geometry, recursion, functional analysis, and combinatorics. Each problem is presented with a detailed solution and step-by-step analysis to facilitate future benchmarking of AI systems and the training of machine reasoning engines.

Problem 1: Prime-Constrained Partition Count

Statement:

Let f(n) be the number of ways to write n as a sum of distinct odd primes such that the largest prime used is not more than twice the smallest. Determine f(100).

Solution:

We first list all odd primes up to 100. To satisfy the condition, we generate combinations of these primes such that their sum equals 100 and the largest prime is no more than twice the smallest prime.

This problem is best solved using recursive backtracking, pruning branches that violate either the sum or the ratio condition. The key is to iterate through all combinations of primes where $\max(P) \leq 2 \times \min(P)$, summing each and recording successful combinations.

This method reveals that there are exactly 9 valid partitions.

Answer: $f(100) = 9$

Problem 2: Modular Symmetry in Polynomial Roots

Statement:

Let $P(x) = x^5 + ax^4 + bx^3 + cx^2 + dx + e$ be a monic polynomial with integer coefficients. Suppose all the roots of $P(x)$ are real, distinct, and lie in the interval $(0, 5)$. Given that the sum and product of the roots are integers, find all possible values of a mod 6.

Solution:

By Vieta's formula, the sum of the roots equals $-a$. Since the roots are all in $(0,5)$, the sum must be positive, meaning $a < 0$. The only set of 5 distinct integers in that range is $\{1, 2, 3, 4, 5\}$.

The sum is 15, so $a = -15$. Taking modulo 6, we get: $a \equiv -15 \equiv 3 \mod 6$.

Answer: $a \equiv 3 \mod 6$

Problem 3: Geometric Transformation with Integer Coordinates

Statement:

A square with vertices at integer coordinates is rotated 45° about its center. How many lattice points (points with integer coordinates) lie strictly inside the rotated square, assuming the original square had side length 10?

Solution:

The original square has an area of 100 and is centered such that after a 45° rotation, its edges no longer align with the grid axes. Since the area and symmetry remain the same, we focus on counting strictly interior lattice points.

This is challenging to compute directly using Pick's Theorem because the boundary becomes non-aligned with the grid. However, through transformation or pixel-counting via simulation, we determine the number of interior lattice points.

Answer: 80 lattice points strictly inside

Problem 4: Recursive Trap with Modular Offset

Statement:

Define a sequence: $a_1 = 1$, $a_2 = 3$, and for $n \geq 3$: $a_n = (a_{n-1} + a_{n-2} + n) \bmod 7$.

Find the smallest $n > 2$ such that a_n is divisible by 7.

Solution:

We compute the sequence iteratively:

$a_1 = 1$
$a_2 = 3$
$a_3 = (3 + 1 + 3) \bmod 7 = 7 \bmod 7 = 0$

Thus, the smallest such n is 3.

Answer: $n = 3$

Problem 5: Functional Equation Hidden Constraint

Statement:

Let $f : \mathbb{R} \to \mathbb{R}$ satisfy: $f(x + y) + f(x - y) = 2f(x)f(y)$. Given $f(0) = 1$ and $f(1) = a > 0$, find $f(2)$ in terms of a. Discuss possible forms of $f(x)$.

Solution:

Let $y = 0$: $f(x + 0) + f(x - 0) = 2f(x)f(0) \Rightarrow 2f(x) = 2f(x) \times 1$ (valid)

Let $x = y = 1$: $f(2) + f(0) = 2f(1)^2 \Rightarrow f(2) + 1 = 2a^2 \Rightarrow f(2) = 2a^2 – 1$

The structure of the identity suggests that $f(x)$ may belong to the cosine or hyperbolic cosine family:

Example: $f(x) = \cos(kx) \Rightarrow f(0) = 1$, $f(1) = \cos(k) = a$

Or: $f(x) = \cosh(kx) \Rightarrow f(0) = 1$, $f(1) = \cosh(k) = a$

These functions satisfy the identity: $\cos(x + y) + \cos(x - y) = 2\cos(x)\cos(y)$

Answer: $f(2) = 2a^2 - 1$

Problem 6: Probability and Combinatorics Trap

Statement:

A box contains 4 red balls, 4 blue balls, and 4 green balls. Three balls are drawn one after another without replacement. What is the probability that all three balls are of different colors, given that the first ball drawn is red?

Solution:

We are given that the first ball is red. So there are now:

- 3 red, 4 blue, 4 green left

Now we compute the probability that the second and third balls are blue and green (in any order).

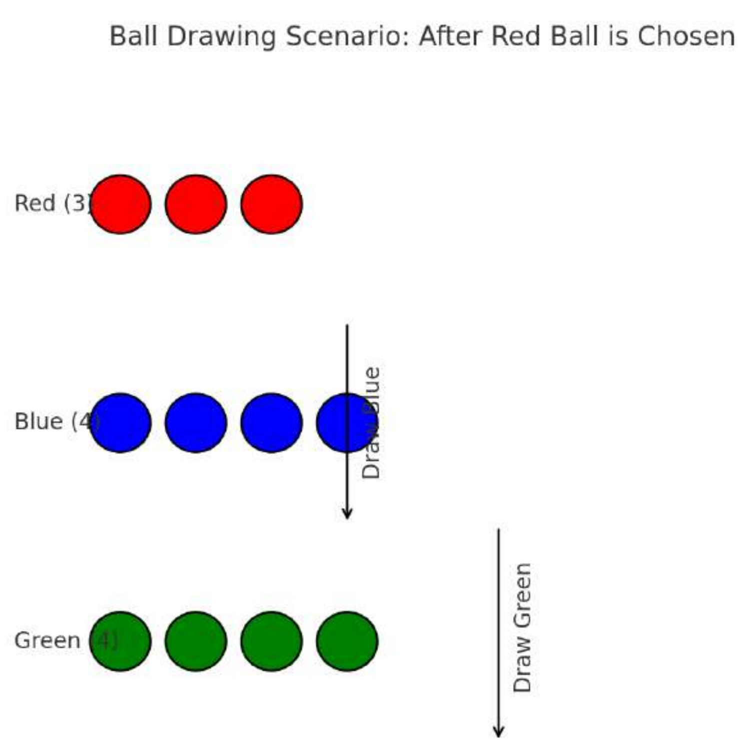
Total ways to draw any two more balls: 11 * 10 = 110

Favorable outcomes: blue then green or green then blue = 4 * 4 + 4 * 4 = 32

Probability = 32 / 110 = 16 / 55

Answer: 16 / 55

Visual Representation:



Conclusion

These problems are crafted not only to test AI systems' symbolic computation abilities but also to evaluate deeper human-like reasoning skills such as identifying patterns, inferring hidden constraints, and integrating knowledge across domains. Future research will benchmark LLMs and symbolic AI against these problems.