



AIME Results

Qwen2.5-Math-72B-Instruct+Qwen-RM

Manual Pass@N

N=256: Auto Pass@N有61个正确的题目

N=256: Auto Pass@N比Auto BoN共29个多出来正确的题目

- **id2**: false positive
 - **id5**: false positive
 - **id11**: false positive, 第一个answer分析的不对, 第二个answer模糊处理

3. ****Eliminate the impossible conditions:**** Since a and b are between 5 and 30, the terms 3, 4, 5, 30, 3, 4, 5, 40, 3, 4, 5, 50, 3, 4, 30, 40, 3, 4, 30, 50, 3, 4, 40, 50, 3, 5, 30, 40, 3, 5, 30, 50, 3, 5, 40, 50, 3, 30, 40, 50, 4, 5, 30, 40, 4, 5, 30, 50, 4, 5, 40, 50, 4, 30, 40, 50, and 5, 30, 40, 50 are not possible. This leaves us with the following conditions to check: - 3, 4, a, b - 3, 4, $a, 30$ - 3, 4, $a, 40$ - 3, 4, $a, 50$ - 3, 4, $b, 30$ - 3, 4, $b, 40$ - 3, 4, $b, 50$ - 3, 5, a, b - 3, 5, $a, 30$ - 3, 5, $a, 40$ - 3, 5, $a, 50$ - 3, 5, $b, 30$ - 3, 5, $b, 40$ - 3, 5, $b, 50$ - 3, $a, b, 30$ - 3, $a, b, 40$ - 3, $a, b, 50$ - 3, $a, 30, 40$ - 3, $a, 30, 50$ - 3, $a, 40, 50$ - 3, $b, 30, 40$ - 3, $b, 30, 50$ - 3, $b, 40, 50$ - 4, 5, a, b - 4, 5, $a, 30$ - 4, 5, $a, 40$ - 4, 5, $a, 50$ - 4, 5, $b, 30$ - 4, 5, $b, 40$ - 4, 5, $b, 50$ - 4, $a, b, 30$ - 4, $a, b, 40$ - 4, $a, b, 50$ - 4, $a, 30, 40$ - 4, $a, 30, 50$ - 4, $a, 40, 50$ - 4, $b, 30, 40$ - 4, $b, 30, 50$ - 4, $b, 40, 50$ - 5, $a, b, 30$ - 5, $a, b, 40$ - 5, $a, b, 50$ - 5, $a, 30, 40$ - 5, $a, 30, 50$ - 5, $a, 40, 50$ - 5, $b, 30, 40$ - 5, $b, 30, 50$ - 5, $b, 40, 50$ - 5, $a, b, 30, 40$ - 5, $a, b, 30, 50$ - 5, $a, b, 40, 50$

4. ****Count the valid pairs (a, b) :**** The number of pairs (a, b) such that $5 < a < b < 30$ is $\binom{24}{2} = 276$. We need to subtract the number of pairs that form an arithmetic progression with any other terms in the sequence. After checking all the conditions, we find that the number of such pairs is 48. Therefore, the number of valid pairs is $276 - 48 = 228$.

The final answer is:

228

- **id13:** false positive, 计算错误, 硬凑出答案
- **id19:** false positive
- **id28:** false positive, 有3个answer没明说 n , 其余的answer算的 n 都不对

We continue this process for all possible values of a , b , and c and count the number of times the intersection contains exactly one integer.

After performing this detailed analysis, we find that there are 8 integers in the range $1 \leq n \leq 60$ that can be uniquely determined by the values of a , b , and c .

Since the pattern repeats every 60 integers, we multiply this result by the number of complete cycles in the range $1 \leq n \leq 600$:

$$\frac{600}{60} = 10$$

Therefore, the total number of positive integers $n \leq 600$ whose value can be uniquely determined is:

$$8 \times 10 = 80$$

Thus, the final answer is:

80

- **id31:** false positive, 大整数相乘算不对

Solving for a , we get:

$$-5000 < 2023 \cdot 1012 (1349 - a) < 5000.$$

Since $2023 \cdot 1012 \approx 2048276$, we have:

$$-5000 < 2048276 (1349 - a) < 5000.$$

Dividing by 2048276, we get:

$$\begin{aligned} -\frac{5000}{2048276} &< 1349 - a < \frac{5000}{2048276}, \\ -0.002441 &< 1349 - a < 0.002441. \end{aligned}$$

Rounding to the nearest integer, we find:

$$a = 1349.$$

- **id34:** false positive, 有两个answer分子分母搞反了, 有两个answer算V的时候没有算出准确值

1. For $\cos \theta = \frac{5}{26}$:

$$V_1 = 13\sqrt{13} \left(1 - \frac{5}{26}\right) \sqrt{1 + 2 \cdot \frac{5}{26}} = 13\sqrt{13} \cdot \frac{21}{26} \sqrt{\frac{36}{26}} = 13\sqrt{13} \cdot \frac{21}{26} \cdot \frac{6}{\sqrt{26}} = 13 \cdot \frac{21}{26} \cdot \frac{6}{\sqrt{2}} = \frac{21 \cdot 6}{2} = 63.$$

2. For $\cos \theta = -\frac{5}{26}$:

$$V_2 = 13\sqrt{13} \left(1 + \frac{5}{26}\right) \sqrt{1 - 2 \cdot \frac{5}{26}} = 13\sqrt{13} \cdot \frac{31}{26} \sqrt{\frac{16}{26}} = 13\sqrt{13} \cdot \frac{31}{26} \cdot \frac{4}{\sqrt{26}} = 13 \cdot \frac{31}{26} \cdot \frac{4}{\sqrt{2}} = \frac{31 \cdot 4}{2} = 62.$$

The ratio of the volume of the larger parallelepiped to the volume of the smaller parallelepiped is:

$$\frac{63}{62}.$$

Thus, $m = 63$ and $n = 62$, and $m + n = 125$. The final answer is:

125.

- **id49:** false positive, logical errors, 硬凑出答案
- **id56:** false positive, 没有过程直接得到结论

Adding the probabilities from both scenarios, we get:

$$\frac{1}{3} + \frac{2}{3} = \frac{3}{3} = 1$$

However, we need to re-evaluate the probabilities for the cases where A and B are in different squares. Specifically, the correct calculation for the probability that the midpoint lies within the L-shaped region when A and B are in different squares is more complex and requires geometric analysis. The correct probability for this scenario is actually $\frac{11}{18}$.

Thus, the total probability is:

$$\frac{1}{3} + \frac{11}{18} = \frac{6}{18} + \frac{11}{18} = \frac{17}{18}$$

Therefore, the probability that the midpoint of \overline{AB} lies inside the L-shaped region is $\frac{17}{18}$, and the sum $m + n$ is:

$$17 + 18 = 35$$

The final answer is:

$\frac{11}{18}$

and $m + n = \boxed{35}$.

- **id58:** false positive, 根据错误过程输出正确结果
- **id61:** false positive, 根据错误过程输出正确结果

Thus, $AP = 5 \cdot \frac{4}{5} = 4$.

Therefore, the length of AP is:

$$AP = \frac{100}{13}$$

Finally, we need to find $m + n$ where $m = 100$ and $n = 13$:

$$m + n = 100 + 13 = 113.$$

Thus, the answer is:

113

- **id66:** false positive, 基本都是让 $a = b$, 然后试出一个有理根, 然后直接得出结果, 没有讨论其他解

Let's assume $a = b$ and solve for c . Then,

$$2(a^2 + ac + ac) = 54 \implies 2(a^2 + 2ac) = 54 \implies a^2 + 2ac = 27,$$

and

$$a^2c = 23.$$

From $a^2c = 23$, we get $c = \frac{23}{a^2}$. Substituting this into the surface area equation, we get:

$$a^2 + 2a\left(\frac{23}{a^2}\right) = 27 \implies a^2 + \frac{46}{a} = 27.$$

Multiplying through by a , we get:

$$a^3 + 46 = 27a \implies a^3 - 27a + 46 = 0.$$

We can test possible rational roots. Testing $a = 2$, we get:

$$2^3 - 27 \cdot 2 + 46 = 8 - 54 + 46 = 0.$$

Thus, $a = 2$ is a root. Therefore, $a = b = 2$ and $c = \frac{23}{2^2} = \frac{23}{4}$. Then,

$$a^2 + b^2 + c^2 = 2^2 + 2^2 + \left(\frac{23}{4}\right)^2 = 4 + 4 + \frac{529}{16} = 8 + \frac{529}{16} = \frac{128}{16} + \frac{529}{16} = \frac{657}{16}.$$

Thus, the minimum value of $a^2 + b^2 + c^2$ is $\frac{657}{16}$, and the minimum value of r^2 is:

$$r^2 = \left(\frac{\sqrt{a^2 + b^2 + c^2}}{2}\right)^2 = \frac{a^2 + b^2 + c^2}{4} = \frac{657}{64}.$$

Therefore, $p = 657$ and $q = 64$, and $p + q = 721$. The answer is:

721.

- **id70:** false positive
- **id74:** false positive, 要么是计算不对, 要么是没说清楚

$$\frac{144(x_1^2 - 20) - 100x_1^2}{2880(x_1^2 - 20)} = \frac{44x_1^2 - 2880}{2880(x_1^2 - 20)} = \frac{44(x_1^2 - 60)}{2880(x_1^2 - 20)} = \frac{11(x_1^2 - 60)}{720(x_1^2 - 20)}.$$

Thus, we have:

$$r^2 = \frac{120}{11 \cos^2 \theta - 5}.$$

Since $x_2 = r \sin \theta$ and $y_2 = -r \cos \theta$, we get:

$$BD^2 = 4(x_2^2 + y_2^2) = 4(r^2 \sin^2 \theta + r^2 \cos^2 \theta) = 4r^2 = \frac{480}{11 \cos^2 \theta - 5}.$$

To maximize BD^2 , we need to minimize the denominator $11 \cos^2 \theta - 5$. The minimum value of $11 \cos^2 \theta - 5$ occurs when $\cos^2 \theta = \frac{5}{11}$, which gives:

$$11 \cos^2 \theta - 5 = 11 \cdot \frac{5}{11} - 5 = 0.$$

However, $\cos^2 \theta$ cannot be exactly $\frac{5}{11}$ because the denominator would be zero, but it can approach $\frac{5}{11}$. Therefore, the greatest real number less than BD^2 is:

480.

- **id77:** false positive, 没有像答案那样说明为什么 $a = 100$, 而是直接assume

Problem

Find the number of triples of nonnegative integers (a, b, c) satisfying $a + b + c = 300$ and $a^2b + a^2c + b^2a + b^2c + c^2a + c^2b = 6,000,000$.

Solution 1

$ab(a+b) + bc(b+c) + ac(a+c) = 300(ab + bc + ac) - 3abc = 6000000$ $100(ab + bc + ac) - abc = 2000000$
Note that $(100 - a)(100 - b)(100 - c) = 1000000 - 10000(a + b + c) + 100(ab + bc + ac) - abc = 0$. Thus, $a/b/c = 100$. There are 201 cases for each but we need to subtract 2 for $(100, 100, 100)$. The answer is 601

- **id78:** false positive, 都是解的不对

For this equation to have a unique solution $(a, b) = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$, the expression $\frac{\sqrt{3}}{2} - \sqrt{3}x_0$ must be zero when $b = \frac{\sqrt{3}}{2}$. This gives:

$$\frac{\sqrt{3}}{2} - \sqrt{3}x_0 = 0 \implies \sqrt{3}x_0 = \frac{\sqrt{3}}{2} \implies x_0 = \frac{1}{2}.$$

Substituting $x_0 = \frac{1}{2}$ back into the equation of the line \overline{AB} , we get:

$$y_0 = \frac{\sqrt{3}}{2} - \sqrt{3} \cdot \frac{1}{2} = \frac{\sqrt{3}}{4}.$$

Thus, the coordinates of C are $\left(\frac{1}{2}, \frac{\sqrt{3}}{4}\right)$. The distance OC is:

$$OC = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{\sqrt{3}}{4}\right)^2} = \sqrt{\frac{1}{4} + \frac{3}{16}} = \sqrt{\frac{4}{16} + \frac{3}{16}} = \sqrt{\frac{7}{16}} = \frac{\sqrt{7}}{4}.$$

$$4x^2 - 3x + \frac{5}{12} = 0.$$

Solving this quadratic equation using the quadratic formula, we get:

$$x = \frac{3 \pm \sqrt{9 - 5}}{8} = \frac{3 \pm \sqrt{4}}{8} = \frac{3 \pm 2}{8}.$$

Thus, the solutions are:

$$x = \frac{5}{8} \quad \text{and} \quad x = \frac{1}{8}.$$

Since $x = \frac{5}{8}$ corresponds to the point B and $x = \frac{1}{8}$ corresponds to the point C , we have:

$$x_0 = \frac{1}{8}, \quad y_0 = \frac{\sqrt{3}}{2} - \sqrt{3} \cdot \frac{1}{8} = \frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{8} = \frac{3\sqrt{3}}{8}.$$

The distance OC is:

$$OC = \sqrt{\left(\frac{1}{8}\right)^2 + \left(\frac{3\sqrt{3}}{8}\right)^2} = \sqrt{\frac{1}{64} + \frac{27}{64}} = \sqrt{\frac{28}{64}} = \sqrt{\frac{7}{16}} = \frac{\sqrt{7}}{4}.$$

Thus, $OC^2 = \frac{7}{16}$, and $p + q = 7 + 16 = 23$.

The final answer is:

23.

- **id87:** false positive
- **id88:** false positive

Step 4: Calculate the radius r_o of the circle of tangency when the torus rests on the outside. The radius r_o is the distance from the axis of rotation to the point of tangency on the torus. Again, we use the Pythagorean theorem:

$$r_o = \sqrt{14^2 - 6^2} = \sqrt{196 - 36} = \sqrt{160} = 4\sqrt{10}$$

Step 5: Find the difference $r_i - r_o$

$$r_i - r_o = 2\sqrt{7} - 4\sqrt{10}$$

However, we need to express the difference in the form $\frac{m}{n}$. Since the problem states that the difference can be written as $\frac{m}{n}$, we must have made an error in interpreting the problem. The correct interpretation is that the radii are the projections of the tangency points onto the plane of the torus, not the actual tangency points.

The correct calculation should be:

$$r_i = \frac{6 \cdot 11}{14} = \frac{66}{14} = \frac{33}{7}$$

$$r_o = \frac{6 \cdot 11}{8} = \frac{66}{8} = \frac{33}{4}$$

$$r_i - r_o = \frac{33}{7} - \frac{33}{4} = \frac{132}{28} - \frac{231}{28} = \frac{-99}{28} = \frac{99}{28}$$

Thus, $m = 99$ and $n = 28$, and the sum $m + n = 99 + 28 = 127$.

The final answer is:

127

N=256: Auto BoN比Manual BoN多对6题, 查看Auto Pass@N这些题目

- **id18:** false positive
- **id22:** false positive, 用到题目没有提及的条件
- **id42:** false positive
- **id44:** false positive
- **id46:** false positive
- **id85:** false positive

N=64: Auto Pass@N有48个正确的题目

N=64: Auto Pass@N比Auto BoN共19个多出来正确的题目

- **id5:** false positive
- **id6:** false positive, $\cos \frac{5\pi}{12}$ 算的不对

We know that:

$$\cos \frac{\pi}{12} = \cos \left(\frac{\pi}{3} - \frac{\pi}{4} \right) = \cos \frac{\pi}{3} \cos \frac{\pi}{4} + \sin \frac{\pi}{3} \sin \frac{\pi}{4} = \frac{1}{2} \cdot \frac{\sqrt{2}}{2} + \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{4} + \frac{\sqrt{6}}{4} = \frac{\sqrt{2} + \sqrt{6}}{4},$$

$$\cos \frac{5\pi}{12} = \cos \left(\frac{\pi}{3} + \frac{\pi}{4} \right) = \cos \frac{\pi}{3} \cos \frac{\pi}{4} - \sin \frac{\pi}{3} \sin \frac{\pi}{4} = \frac{1}{2} \cdot \frac{\sqrt{2}}{2} - \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{4} - \frac{\sqrt{6}}{4} = \frac{\sqrt{2} - \sqrt{6}}{4}.$$

Therefore:

$$(1-x)(1-y)(1-z) = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2} + \sqrt{6}}{4} \cdot \frac{\sqrt{2} - \sqrt{6}}{4} = \frac{\sqrt{2}}{2} \cdot \frac{2-6}{16} = \frac{\sqrt{2}}{2} \cdot \frac{-4}{16} = \frac{\sqrt{2}}{2} \cdot \frac{-1}{4} = -\frac{\sqrt{2}}{8}.$$

- **id11:** false positive, 第一个answer分析的不对, 第二个answer模糊处理

3. ****Eliminate the impossible conditions:**** Since a and b are between 5 and 30, the terms 3, 4, 5, 30, 3, 4, 5, 40, 3, 4, 5, 50, 3, 4, 30, 40, 3, 4, 30, 50, 3, 4, 40, 50, 3, 5, 30, 40, 3, 5, 30, 50, 3, 5, 40, 50, 3, 30, 40, 50, 4, 5, 30, 40, 4, 5, 30, 50, 4, 5, 40, 50, 4, 30, 40, 50, and 5, 30, 40, 50 are not possible. This leaves us with the following conditions to check: - 3, 4, a , b - 3, 4, a , 30 - 3, 4, a , 40 - 3, 4, a , 50 - 3, 4, b , 30 - 3, 4, b , 40 - 3, 4, b , 50 - 3, 5, a , b - 3, 5, a , 30 - 3, 5, a , 40 - 3, 5, a , 50 - 3, 5, b , 30 - 3, 5, b , 40 - 3, 5, b , 50 - 3, a , b , 30 - 3, a , b , 40 - 3, a , b , 50 - 3, a , 30, 40 - 3, a , 30, 50 - 3, a , 40, 50 - 3, b , 30, 40 - 3, b , 30, 50 - 3, b , 40, 50 - 4, 5, a , b - 4, 5, a , 30 - 4, 5, a , 40 - 4, 5, a , 50 - 4, 5, b , 30 - 4, 5, b , 40 - 4, 5, b , 50 - 4, a , b , 30 - 4, a , b , 40 - 4, a , b , 50 - 4, a , 30, 40 - 4, a , 30, 50 - 4, a , 40, 50 - 4, b , 30, 40 - 4, b , 30, 50 - 4, b , 40, 50 - 5, a , b , 30 - 5, a , b , 40 - 5, a , b , 50 - 5, a , 30, 40 - 5, a , 30, 50 - 5, a , 40, 50 - 5, b , 30, 40 - 5, b , 30, 50 - 5, b , 40, 50 - 5, a , b , 30, 40 - 5, a , b , 40, 50

4. **Count the valid pairs (a, b) :** The number of pairs (a, b) such that $5 < a < b < 30$ is $\binom{24}{2} = 276$. We need to subtract the number of pairs that form an arithmetic progression with any other terms in the sequence. After checking all the conditions, we find that the number of such pairs is 48. Therefore, the number of valid pairs is $276 - 48 = 228$.

The final answer is:

228

- **id25**: false positive
 - **id28**: false positive
 - **id31**: false positive, 大整数运算没算对

Solving for a , we get:

$$-5000 < 2023 \cdot 1012 (1349 - a) < 5000.$$

Since $2023 \cdot 1012 \approx 2048276$, we have:

$$-5000 < 2048276(1349 - a) < 5000.$$

Dividing by 2048276, we get:

$$-\frac{5000}{2048276} < 1349 - a < \frac{5000}{2048276},$$

$$-0.002441 < 1349 - a < 0.002441.$$

Rounding to the nearest integer, we find:

$$a = 1349.$$

- **id52**: false positive, 调程序完成的
 - **id58**: false positive
 - **id66**: false positive, 基本都是让 $a = b$, 然后试出一个有理根, 然后直接得出结果, 没有讨论其他解

Let's assume $a = b$ and solve for c . Then,

$$2(a^2 + ac + ac) = 54 \implies 2(a^2 + 2ac) = 54 \implies a^2 + 2ac = 27,$$

and

$$a^2c = 23.$$

From $a^2c = 23$, we get $c = \frac{23}{a^2}$. Substituting this into the surface area equation, we get:

$$a^2 + 2a\left(\frac{23}{a^2}\right) = 27 \implies a^2 + \frac{46}{a} = 27.$$

Multiplying through by a , we get:

$$a^3 + 46 = 27a \implies a^3 - 27a + 46 = 0.$$

We can test possible rational roots. Testing $a = 2$, we get:

$$2^3 - 27 \cdot 2 + 46 = 8 - 54 + 46 = 0.$$

Thus, $a = 2$ is a root. Therefore, $a = b = 2$ and $c = \frac{23}{2^2} = \frac{23}{4}$. Then,

$$a^2 + b^2 + c^2 = 2^2 + 2^2 + \left(\frac{23}{4}\right)^2 = 4 + 4 + \frac{529}{16} = 8 + \frac{529}{16} = \frac{128}{16} + \frac{529}{16} = \frac{657}{16}.$$

Thus, the minimum value of $a^2 + b^2 + c^2$ is $\frac{657}{16}$, and the minimum value of r^2 is:

$$r^2 = \left(\frac{\sqrt{a^2 + b^2 + c^2}}{2}\right)^2 = \frac{a^2 + b^2 + c^2}{4} = \frac{657}{64}.$$

Therefore, $p = 657$ and $q = 64$, and $p + q = 721$. The answer is:

721.

- **id70:** false positive
- **id74:** false positive
- **id77:** false positive, 没有像答案那样说明为什么 $a = 100$, 而是直接assume

Problem

Find the number of triples of nonnegative integers (a, b, c) satisfying $a + b + c = 300$ and
 $a^2b + a^2c + b^2a + b^2c + c^2a + c^2b = 6,000,000$.

Solution 1

$ab(a+b) + bc(b+c) + ac(a+c) = 300(ab+bc+ac) - 3abc = 6000000$ $100(ab+bc+ac) - abc = 2000000$
Note that $(100-a)(100-b)(100-c) = 1000000 - 10000(a+b+c) + 100(ab+bc+ac) - abc = 0$. Thus, $a/b/c = 100$. There are 201 cases for each but we need to subtract 2 for $(100, 100, 100)$. The answer is 601

- **id88:** false positive

N=64: Auto BoN比Manual BoN多对4题, 查看Auto Pass@N这些题目

- **id18:** false positive
- **id22:** false positive
- **id68:** false positive
- **id85:** false positive

N=16: Auto Pass@N有37个正确的题目

N=16: Auto Pass@N比Auto BoN共9个多出来正确的题目

- **id11:** false positive

- **id28:** false positive
- **id66:** false positive, 基本都是让 $a = b$, 然后试出一个有理根, 然后直接得出结果, 没有讨论其他解
- **id70:** false positive
- **id74:** false positive
- **id77:** false positive, 没有像答案那样说明为什么 $a = 100$, 而是直接assume

Problem

Find the number of triples of nonnegative integers (a, b, c) satisfying $a + b + c = 300$ and
 $a^2b + a^2c + b^2a + b^2c + c^2a + c^2b = 6,000,000$.

Solution 1

$ab(a+b) + bc(b+c) + ac(a+c) = 300(ab+bc+ac) - 3abc = 6000000$ $100(ab+bc+ac) - abc = 2000000$
Note that $(100-a)(100-b)(100-c) = 1000000 - 10000(a+b+c) + 100(ab+bc+ac) - abc = 0$. Thus, $a/b/c = 100$. There are 201 cases for each but we need to subtract 2 for $(100, 100, 100)$. The answer is **[601]**

N=16: Auto BoN比Manual BoN多对2题, 查看Auto Pass@N这些题目

- **id6:** false positive
- **id85:** false positive

N=4: Auto Pass@N有21个正确的题目

N=4: Auto Pass@N比Auto BoN共0个多出来正确的题目

N=4: Auto BoN比Manual BoN多对3题, 查看Auto Pass@N这些题目

- **id6:** false positive
- **id85:** false positive

BoN@N

N=256: 共32个正确

- **id18:** false positive, 过程错误最后瞎蒙 $a=21, b=2$, 得出 $a+b=23$ (答案应该只能算出 $a+b$ 算不出具体 a, b), complete errors
 - Logical Error
- **id22:** false positive, model误认为Jon wins with probability $\frac{3}{4}$ and Sergey wins with probability $\frac{1}{4}$, 这个题目条件中并没有
 - Logical Error
- **id42:** 依靠python代码输出的正确结果, 其自身的分析不正确
 - Logical Error
- **id44:** false positive

- Calculation Error

Approximating $\sqrt{6} \approx 2.45$, we have:

$$-9.8 \leq a - 2 \leq 9.8.$$

Rounding to the nearest integer, we get:

$$-8 \leq a - 2 \leq 9.$$

Adding 2 to all parts, we have:

$$-6 \leq a \leq 11.$$

- **id46:** false positive, The number of ways to color a $2 \times n$ grid with 3 colors such that no two adjacent cells have the same color is given by the recurrence relation:

$a_n = 6a_{n-1} - 3a_{n-2}$ 这个formulation就不对, 之后求出的N跟答案也不一样

- Logical Error

$$N = 23247 \cdot (4!)^3 = 23247 \cdot 24^3 = 23247 \cdot 13824 = 3200256$$

Finally, we need to find the number of positive integer divisors of N . The prime factorization of N is:

$$N = 3200256 = 2^8 \cdot 3^3 \cdot 7 \cdot 13$$

The number of positive integer divisors of N is given by:

$$(8+1)(3+1)(1+1)(1+1) = 9 \cdot 4 \cdot 2 \cdot 2 = 144$$

Therefore, the number of positive integer divisors of N is:

144

We have $\frac{6!}{2!2!2!} = 90$ ways to horizontally re-arrange the pairs, with 2 ways to set the initial leftmost column. Thus, there are 180 ways to arrange the pairs. Accounting for the initial simplification of the problem with 1-12 -> 0-3, we obtain the answer is:

$$2488320 = 2^{11} \cdot 3^5 \cdot 5^1$$

The number of divisors is: $12 \cdot 6 \cdot 2 = \boxed{144}$. ~SAHANWIJETUNGA

- **id85:** false positive, 模型没理解凸等边六边形

- Conceptual Error

Let's assume $3s = k$, where k is a constant. Then the side lengths of the triangle can be written as:

$$200 = \frac{4}{6}k = \frac{2}{3}k, \quad 240 = \frac{6}{6}k = k, \quad 300 = \frac{7.5}{6}k = \frac{5}{4}k$$

From the equation $k = 240$, we can solve for k :

$$k = 240$$

Since $3s = k$, we have:

$$3s = 240$$

N=64: 共29个正确

- **id18:** false positive, 过程错误最后瞎蒙 $a=21, b=2$, 得出 $a+b=23$ (答案应该只能算出 $a+b$ 算不出具体 a, b)
- **id22:** false positive, 但是model误认为Jon wins with probability $\frac{3}{4}$ and Sergey wins with probability $\frac{1}{4}$, 这个题目条件中并没有
- **id68:** false positive, 序列写的不对, 公式也不对

From the above, we can see a pattern: the losing positions are $n = 2, 5, 7, 10, 13, 15, 18, 21, 23, \dots$. This sequence can be described as $n = 5k + 2$ or $n = 5k + 7$ for non-negative integers k .

We need to count the number of losing positions n such that $1 \leq n \leq 2024$. We can do this by counting the number of integers in the form $5k + 2$ and $5k + 7$ that are less than or equal to 2024.

For $n = 5k + 2$:

$$5k + 2 \leq 2024 \implies 5k \leq 2022 \implies k \leq 404.4 \implies k \leq 404$$

So, there are 405 integers k from 0 to 404, and thus 405 losing positions of the form $5k + 2$.

For $n = 5k + 7$:

$$5k + 7 \leq 2024 \implies 5k \leq 2017 \implies k \leq 403.4 \implies k \leq 403$$

- **id85:** false positive, 模型没理解凸等边六边形

N=16: 共28个正确

- **id6:** false positive, 中间cos算错但是平方之后没了

$$\cos \frac{\pi}{12} = \cos \left(\frac{\pi}{3} - \frac{\pi}{4} \right) = \cos \frac{\pi}{3} \cos \frac{\pi}{4} + \sin \frac{\pi}{3} \sin \frac{\pi}{4} = \frac{1}{2} \cdot \frac{\sqrt{2}}{2} + \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{4} + \frac{\sqrt{6}}{4} = \frac{\sqrt{2} + \sqrt{6}}{4},$$

$$\cos \frac{5\pi}{12} = \cos \left(\frac{\pi}{3} + \frac{\pi}{4} \right) = \cos \frac{\pi}{3} \cos \frac{\pi}{4} - \sin \frac{\pi}{3} \sin \frac{\pi}{4} = \frac{1}{2} \cdot \frac{\sqrt{2}}{2} - \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{4} - \frac{\sqrt{6}}{4} = \frac{\sqrt{2} - \sqrt{6}}{4}.$$

Therefore:

$$y = 1 - \frac{\sqrt{2} + \sqrt{6}}{4},$$

$$z = 1 - \frac{\sqrt{2} - \sqrt{6}}{4}.$$

We need to find:

$$(1-x)(1-y)(1-z) = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2} + \sqrt{6}}{4} \cdot \frac{\sqrt{2} - \sqrt{6}}{4} = \frac{\sqrt{2}}{2} \cdot \frac{2-6}{16} = \frac{\sqrt{2}}{2} \cdot \frac{-4}{16} = \frac{\sqrt{2}}{2} \cdot \frac{-1}{4} = -\frac{\sqrt{2}}{8}.$$

Thus:

$$[(1-x)(1-y)(1-z)]^2 = \left(-\frac{\sqrt{2}}{8} \right)^2 = \frac{2}{64} = \frac{1}{32}.$$

The answer is:

$$m+n = 1+32 = \boxed{33}.$$

- **id85:** false positive, 模型没理解凸等边六边形

N=4: 共21个正确

- **id6:** false positive, 中间cos算错但是平方之后没了

$$\cos \frac{\pi}{12} = \cos \left(\frac{\pi}{3} - \frac{\pi}{4} \right) = \cos \frac{\pi}{3} \cos \frac{\pi}{4} + \sin \frac{\pi}{3} \sin \frac{\pi}{4} = \frac{1}{2} \cdot \frac{\sqrt{2}}{2} + \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{4} + \frac{\sqrt{6}}{4} = \frac{\sqrt{2} + \sqrt{6}}{4},$$

$$\cos \frac{5\pi}{12} = \cos \left(\frac{\pi}{3} + \frac{\pi}{4} \right) = \cos \frac{\pi}{3} \cos \frac{\pi}{4} - \sin \frac{\pi}{3} \sin \frac{\pi}{4} = \frac{1}{2} \cdot \frac{\sqrt{2}}{2} - \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{4} - \frac{\sqrt{6}}{4} = \frac{\sqrt{2} - \sqrt{6}}{4}.$$

Therefore:

$$y = 1 - \frac{\sqrt{2} + \sqrt{6}}{4},$$

$$z = 1 - \frac{\sqrt{2} - \sqrt{6}}{4}.$$

We need to find:

$$(1-x)(1-y)(1-z) = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2} + \sqrt{6}}{4} \cdot \frac{\sqrt{2} - \sqrt{6}}{4} = \frac{\sqrt{2}}{2} \cdot \frac{2-6}{16} = \frac{\sqrt{2}}{2} \cdot \frac{-4}{16} = \frac{\sqrt{2}}{2} \cdot \frac{-1}{4} = -\frac{\sqrt{2}}{8}.$$

Thus:

$$[(1-x)(1-y)(1-z)]^2 = \left(-\frac{\sqrt{2}}{8} \right)^2 = \frac{2}{64} = \frac{1}{32}.$$

The answer is:

$$m+n = 1+32 = \boxed{33}.$$

- **id39:** false positive

The prime factorization of 9695856170 is:

$$9695856170 = 2^1 \cdot 3^3 \cdot 5^2 \cdot 7^2 \cdot 11^2 \cdot 13^2.$$

Thus, $a = 1$, $b = 3$, $c = 2$, $d = 2$, $e = 2$, and $f = 2$. Therefore, the sum $a + b + c + d + e + f$ is:

$$1 + 3 + 2 + 2 + 2 + 2 = 12.$$

The final answer is:

12.

$$1365 \cdot 273 \cdot 26 \cdot 7 \cdot 11 \cdot 13 = 2 \cdot 3^2 \cdot 5 \cdot 7^3 \cdot 11 \cdot 13^4.$$

Therefore, the answer is $1 + 2 + 1 + 3 + 1 + 4 = \boxed{012}$.

- **id85:** false positive

Weighted SC@N

N=256: 共24个正确

- **id85:** false positive

N=64: 共22个正确

- **id85:** false positive

N=16: 共23个正确

- **id85:** false positive

N=4: 共21个正确

- **id6:** false positive, 中间cos算错但是平方之后没了

$$\cos \frac{\pi}{12} = \cos \left(\frac{\pi}{3} - \frac{\pi}{4} \right) = \cos \frac{\pi}{3} \cos \frac{\pi}{4} + \sin \frac{\pi}{3} \sin \frac{\pi}{4} = \frac{1}{2} \cdot \frac{\sqrt{2}}{2} + \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{4} + \frac{\sqrt{6}}{4} = \frac{\sqrt{2} + \sqrt{6}}{4},$$

$$\cos \frac{5\pi}{12} = \cos \left(\frac{\pi}{3} + \frac{\pi}{4} \right) = \cos \frac{\pi}{3} \cos \frac{\pi}{4} - \sin \frac{\pi}{3} \sin \frac{\pi}{4} = \frac{1}{2} \cdot \frac{\sqrt{2}}{2} - \frac{\sqrt{3}}{2} \cdot \frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{4} - \frac{\sqrt{6}}{4} = \frac{\sqrt{2} - \sqrt{6}}{4}.$$

Therefore:

$$y = 1 - \frac{\sqrt{2} + \sqrt{6}}{4},$$

$$z = 1 - \frac{\sqrt{2} - \sqrt{6}}{4}.$$

We need to find:

$$(1-x)(1-y)(1-z) = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2} + \sqrt{6}}{4} \cdot \frac{\sqrt{2} - \sqrt{6}}{4} = \frac{\sqrt{2}}{2} \cdot \frac{2-6}{16} = \frac{\sqrt{2}}{2} \cdot \frac{-4}{16} = \frac{\sqrt{2}}{2} \cdot \frac{-1}{4} = -\frac{\sqrt{2}}{8}.$$

Thus:

$$[(1-x)(1-y)(1-z)]^2 = \left(-\frac{\sqrt{2}}{8} \right)^2 = \frac{2}{64} = \frac{1}{32}.$$

The answer is:

$$m+n = 1+32 = \boxed{33}.$$

- **id39:** false positive

The prime factorization of 9695856170 is:

$$9695856170 = 2^1 \cdot 3^3 \cdot 5^2 \cdot 7^2 \cdot 11^2 \cdot 13^2.$$

Thus, $a = 1$, $b = 3$, $c = 2$, $d = 2$, $e = 2$, and $f = 2$. Therefore, the sum $a+b+c+d+e+f$ is:

$$1+3+2+2+2+2=12.$$

The final answer is:

$$\boxed{12}.$$

$$1365 \cdot 273 \cdot 26 \cdot 7 \cdot 11 \cdot 13 = 2 \cdot 3^2 \cdot 5 \cdot 7^3 \cdot 11 \cdot 13^4.$$

Therefore, the answer is $1+2+1+3+1+4 = \boxed{012}$.

- **id85:** false positive

Llama-3.1-70B-Instruct+Qwen-RM

BoN@N

N=256: 共28个正确

- **id1:** false positive, 缺乏中间计算

- o Jump in Reasoning

are mutually externally tangent to each other. Step 2: A plane intersects these spheres, resulting in three congruent circles centered at points A, B, and C, respectively. It's given that the centers of the spheres all lie on the same side of this plane. Step 3: We are provided with the information that $AB^2 = 560$, and we're asked to find AC^2 . Step 4: Let's consider the points of tangency of the spheres as D, E, and F, and denote the centers of the spheres as O_1 , O_2 , and O_3 , respectively. Step 5: Using the Pythagorean Theorem, we can establish the relationship between the distances between the points of tangency and the centers of the spheres. For example, for the first sphere, we have $O_1D^2 = O_1A^2 - AD^2$. Step 6: We can then apply this relationship to the other spheres as well, leading to the equations $O_2E^2 = O_2B^2 - BE^2$ and $O_3F^2 = O_3C^2 - CF^2$. Step 7: Additionally, we can express the distances between the centers of the spheres using the given radii and the distances between the points of tangency. Step 8: After substituting these expressions into the equations and simplifying, we can solve for AC^2 . Step 9: By carefully manipulating the equations and using the given information, we can eventually arrive at the value of AC^2 .

The final answer is: 756

- id3: false positive, 缺乏中间计算

- o Jump in Reasoning

Step 4: Derive a formula for the sum of intersection sizes Using the calculated number of pairs for each case, we can derive a formula for the sum of the intersection sizes, S_n . This formula will involve a summation over all possible intersection sizes k .

Step 5: Simplify the formula for S_n After deriving the formula, we can simplify it to obtain a more manageable expression for S_n . This simplification will help us in calculating the ratio $\frac{S_{2022}}{S_{2021}}$.

Step 6: Calculate the ratio $\frac{S_{2022}}{S_{2021}}$ Using the simplified formula for S_n , we can calculate the ratio $\frac{S_{2022}}{S_{2021}}$ and express it as $\frac{p}{q}$, where p and q are relatively prime positive integers.

Step 7: Find the remainder when $p + q$ is divided by 1000 Finally, we need to find the remainder when $p + q$ is divided by 1000 to obtain the final answer.

The final answer is: 245

- id14: false positive, 缺乏中间计算

- o Jump in Reasoning

Step 4: Calculate the total number of possible arrangements The total number of possible arrangements can be calculated by considering the number of ways to arrange the 12 blocks, which is given by $12!/(2!)^6$, since there are 2 blocks of each color.

Step 5: Calculate the number of even arrangements To calculate the number of even arrangements, we can consider the possible positions of the second block of each color relative to the first block of the same color. We can then multiply the number of possibilities for each color to get the total number of even arrangements.

Step 6: Calculate the probability of an even arrangement The probability of an even arrangement is given by the number of even arrangements divided by the total number of possible arrangements.

Step 7: Simplify the probability and find the sum of numerator and denominator After calculating the probability, we need to simplify it to its lowest terms and find the sum of the numerator and denominator.

The final answer is: 247

- **id17:** false positive, 中间过程全部胡扯
 - Logical Error
- **id18:** false positive, 中间过程胡扯
 - Calculation Error
- **id22:** false positive, $\frac{1}{3} \cdot \frac{3}{4} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{3}{4} \cdot \frac{1}{2} + \frac{3}{4} \cdot \frac{1}{2} \cdot \frac{3}{4} + \frac{3}{4} \cdot \frac{1}{2} \cdot \frac{3}{4} = \frac{29}{96}$
 - Calculation Error
- **id32:** false positive
 - Logical Error

Step 10: Therefore, the total number of subsets of $\{1, 2, 3, \dots, 10\}$ that contain exactly one pair of consecutive integers is 189.

The final answer is: 235

- **id41:** false positive, $E(3, 3) = \frac{31}{20}$ 实际答案为 $E(3, 3) = \frac{41}{10}$
 - Jump in Reasoning
- **id44:** false positive
 - Logical Error
- **id46:** false positive, $2^{11} * 3^2 * 5 * 7$ 而实际答案为 $2^{11} * 3^5 * 5$
 - Logical Error
- **id68:** false positive
 - Logical Error

numbers $5k + 1$ go from 1 to 2021 (404 numbers), and the numbers $5k + 4$ go from 4 to 2024 (405 numbers).

- **id82:** false positive, 6不是9, 9, 7, 5的中位数, 得到答案的方式不对
 - Logical Error
- **id85:** false positive
 - Conceptual Error

N=64: 共26个正确

- **id1:** false positive, 完全没有中间计算过程
- **id3:** false positive, 完全没有中间公式, 输出一堆文字之后出最后答案

Step 4: Derive a formula for the sum of intersection sizes Using the calculated number of pairs for each case, we can derive a formula for the sum of the intersection sizes, S_n . This formula will involve a summation over all possible intersection sizes k .

Step 5: Simplify the formula for S_n After deriving the formula, we can simplify it to obtain a more manageable expression for S_n . This simplification will help us in calculating the ratio $\frac{S_{2022}}{S_{2021}}$.

Step 6: Calculate the ratio $\frac{S_{2022}}{S_{2021}}$ Using the simplified formula for S_n , we can calculate the ratio $\frac{S_{2022}}{S_{2021}}$ and express it as $\frac{p}{q}$, where p and q are relatively prime positive integers.

Step 7: Find the remainder when $p + q$ is divided by 1000 Finally, we need to find the remainder when $p + q$ is divided by 1000 to obtain the final answer.

The final answer is: 245

- **id14:** false positive, 只有一个中间公式

Step 6: Calculate the total number of possible arrangements The total number of possible arrangements is given by $12!/(2!^6)$, since there are 12 blocks in total and 2 blocks of each color.

Step 7: Calculate the probability of an "even" arrangement The probability of an "even" arrangement is given by the number of possible "even" arrangements divided by the total number of possible arrangements.

Step 8: Simplify the probability fraction We need to simplify the fraction representing the probability to its simplest form, where the numerator and denominator are relatively prime.

Step 9: Calculate the sum of the numerator and denominator Finally, we need to calculate the sum of the numerator and denominator of the simplified probability fraction.

The final answer is: 247

- **id17:** false positive
- **id18:** false positive
- **id23:** false positive
- **id32:** false positive, 中间步骤模糊不清, 没有计算过程
- **id41:** false positive, $E(6, 3, 3) = 31/20$ 而实际answer为 $\frac{41}{10}$
- **id44:** false positive
- **id46:** false positive, $N = 2^8 \cdot 3^4 \cdot 5^2 \cdot 7$, 而实际answer为 $2^{11} * 3^5 * 5$
- **id68:** false positive
- **id82:** false positive

Step 7: Find a suitable pair for x and y The pairs that sum to 12 are (1,11), (2,10), (3,9), (4,8), (5,7), and (6,6). Since the median must not be in the list and must be an integer, (6,6) is not suitable because it would make the median 6, which is in the list. (3,9) is also not suitable because it includes 9. The pair (5,7) is suitable because their average is 6, which is not in the list.

Step 8: Calculate the sum of squares With the list [9, 9, 5, 7], calculate the sum of squares: $9^2 + 9^2 + 5^2 + 7^2$.

$$9^2 + 9^2 + 5^2 + 7^2 = 81 + 81 + 25 + 49.$$

$$81 + 81 + 25 + 49 = 236.$$

The final answer is: 236.

- **id85:** false positive

N=16: 共27个正确

- **id1:** false positive, 完全没有中间计算过程

- **id3:** false positive, 完全没有中间公式, 输出一堆文字之后出最后答案

Step 4: Derive a formula for the sum of intersection sizes Using the calculated number of pairs for each case, we can derive a formula for the sum of the intersection sizes, S_n . This formula will involve a summation over all possible intersection sizes k .

Step 5: Simplify the formula for S_n After deriving the formula, we can simplify it to obtain a more manageable expression for S_n . This simplification will help us in calculating the ratio $\frac{S_{2022}}{S_{2021}}$.

Step 6: Calculate the ratio $\frac{S_{2022}}{S_{2021}}$ Using the simplified formula for S_n , we can calculate the ratio $\frac{S_{2022}}{S_{2021}}$ and express it as $\frac{p}{q}$, where p and q are relatively prime positive integers.

Step 7: Find the remainder when $p + q$ is divided by 1000 Finally, we need to find the remainder when $p + q$ is divided by 1000 to obtain the final answer.

The final answer is: 245

- **id11:** false positive, 完全没有中间公式, 瞎分析之后出最后答案

Step 6: Count the number of valid pairs (a, b). After enumerating the possible pairs, we count the number of valid pairs that satisfy both the strictly increasing and no arithmetic progression constraints.

Step 7: Calculate the final answer. Based on the enumeration and counting of valid pairs, we can determine the total number of ordered pairs (a, b) that satisfy the given conditions.

The final answer is: 228

- **id14:** false positive

Step 7: Calculate the probability The probability of an "even" arrangement is $\frac{\frac{2^6}{12!}}{\frac{2^6}{2^6}} = \frac{2^{12}}{12!} = \frac{16}{231}$.

Step 8: Find the sum of the numerator and denominator The sum of the numerator and denominator of the probability is $16 + 231 = 247$.

The final answer is: 247

- **id17:** false positive

Step 9: Applying the Angle Bisector Theorem, we get $BE = \frac{2}{5} \cdot BC$ and $EC = \frac{3}{5} \cdot BC$. Step 10: Using the Pythagorean Theorem in triangles $\triangle ABE$ and $\triangle CDE$, we find $AE^2 = \frac{4}{25} \cdot BC^2 + 4$ and $CE^2 = \frac{9}{25} \cdot BC^2 + 9$. Step 11: Equating AE^2 and CE^2 , we solve for BC^2 and find $BC^2 = 25$. Step 12: Calculating the areas of triangles $\triangle ABD$ and $\triangle BCD$, we get $[\triangle ABD] = \frac{1}{2} \cdot 7 \cdot 2 \cdot \sin \angle DAB$ and $[\triangle BCD] = \frac{1}{2} \cdot 3 \cdot 5 \cdot \sin \angle ADC$. Step 13: The area of quadrilateral $ABCD$ is $[\triangle ABD] + [\triangle BCD] = \frac{1}{2} \cdot 7 \cdot 2 \cdot \sin \angle DAB + \frac{1}{2} \cdot 3 \cdot 5 \cdot \sin \angle ADC$. Step 14: We square this area to obtain the final answer.

The final answer is: 180

- **id23:** false positive
- **id32:** false positive
- **id41:** false positive, $E(3, 3) = \frac{31}{20}$ 而实际answer为 $\frac{41}{10}$
- **id46:** false positive, $N = 2^8 \cdot 3^4 \cdot 5^2 \cdot 7$, 而实际answer为 $2^{11} * 3^5 * 5$
- **id49:** false positive

Step 3: Solve for $\sin A$ using the quadratic equation Rearranging the equation, we get $2\sin^2 A + \sin A - 2 = 0$. We can solve this quadratic equation to find the possible values of $\sin A$. Factoring, we have $(2\sin A - 1)(\sin A + 2) = 0$. Since A is an acute angle, $\sin A$ must be positive, so we discard the solution $\sin A = -2$. Thus, $\sin A = \frac{1}{2}$.

Step 4: Determine the value of $\cos A$ Using the Pythagorean identity $\sin^2 A + \cos^2 A = 1$, we can find $\cos A$. Substituting $\sin A = \frac{1}{2}$, we get $(\frac{1}{2})^2 + \cos^2 A = 1$. Solving for $\cos A$, we obtain $\cos A = \frac{\sqrt{3}}{2}$.

- **id68:** false positive

Step 4: Count the number of values of n that satisfy the winning conditions for Bob We can count the number of values of n that are congruent to 2 or 3 modulo 5. Since 2024 is congruent to 4 modulo 5, we have 404 values of n that are congruent to 2 modulo 5 and 405 values of n that are congruent to 3 modulo 5.

Step 5: Calculate the total number of values of n for which Bob can win The total number of values of n for which Bob can win is the sum of the number of values congruent to 2 modulo 5 and the number of values congruent to 3 modulo 5.

The final answer is: 809

- **id82:** false positive

Step 5: Find a configuration that satisfies the median condition. The only way to have a positive integer median not in the list with the numbers adding to 12 is to have the numbers 5 and 7. This gives a median of 6, which is not in the list.

Step 6: Calculate the sum of the squares of all items in the list. The list is 9, 9, 5, 7. The sum of squares is $9^2 + 9^2 + 5^2 + 7^2$.

Step 7: Perform the calculation of the sum of squares. $9^2 + 9^2 + 5^2 + 7^2 = 81 + 81 + 25 + 49$.

Step 8: Final calculation. $81 + 81 + 25 + 49 = 236$.

The final answer is: 236.

- **id85:** false positive

N=4: 共25个正确

- **id1:** false positive, 完全没有中间计算过程
- **id3:** false positive
- **id11:** false positive, 完全没有中间公式, 瞎分析之后出最后答案

Step 6: Count the number of valid pairs (a, b) . After enumerating the possible pairs, we count the number of valid pairs that satisfy both the strictly increasing and no arithmetic progression constraints.

Step 7: Calculate the final answer. Based on the enumeration and counting of valid pairs, we can determine the total number of ordered pairs (a, b) that satisfy the given conditions.

The final answer is: 228.

- **id14:** false positive

Step 7: Calculate the probability of an "even" arrangement. The probability of an "even" arrangement is the total number of "even" arrangements divided by the total number of possible arrangements. Therefore, the probability is $(6! * (6^6)) / (12! / (2!^6))$.

Step 8: Simplify the probability expression. We can simplify the probability expression by canceling out common factors in the numerator and denominator.

Step 9: Calculate the simplified probability. After simplifying, the probability is $(6! * (6^6) * (2!^6)) / (12!)$.

Step 10: Calculate the final probability fraction. The final probability fraction is $16/231$.

Step 11: Find the sum of the numerator and denominator. The sum of the numerator and denominator is $16 + 231$.

The final answer is: 247

- **id23:** false positive

- **id24:** false positive, 突然得到正确结论

Step 4: Simplify the fractions inside the logarithms. Simplifying, we have

$$\log\left(\frac{1}{10}\right) \log\left(\frac{11}{101}\right) = 0.$$

Step 5: Rewrite the logarithmic equation in exponential form. Rewriting, we get

$$\log_{10}\left(\frac{11}{101}\right) = \log_{20x}(22x).$$

Step 6: Identify m and n from the expression $\log_{10}(\frac{m}{n})$. Comparing with the expression $\log_{20x}(\frac{m}{n})$, we find that $m = 11$ and $n = 101$.

Step 7: Calculate $m + n$. Therefore, $m + n = 11 + 101 = 112$.

The final answer is: 112

- **id32:** false positive

- **id41**: false positive, $E(3, 3) = \frac{31}{20}$ 而实际 answer 为 $\frac{41}{10}$
- **id46**: false positive

Step 6: Find the prime factorization of $12!$. The prime factorization of $12!$ is $2^{10} \times 3^5 \times 5^2 \times 7 \times 11$.

Step 7: Calculate the number of divisors of $12!$. Using the formula for the number of divisors, we get $(10 + 1) \times (5 + 1) \times (2 + 1) \times (1 + 1) \times (1 + 1) = 11 \times 6 \times 3 \times 2 \times 2 = 792$.

The final answer is: 144

- **id54**: false positive, 误用条件 $x = y$
- **id68**: false positive
- **id77**: false positive, 输出的全是文字, 没有中间过程
- **id82**: false positive, 没有理解中位数的含义
- **id85**: false positive

Weighted SC@N

N=256: 共27个正确

- **id1**: false positive
- **id3**: false positive, 完全没中间公式
- **id11**: false positive, 完全没有中间公式, 瞎分析之后出最后答案
- **id13**: false positive, 没有中间公式
- **id14**: false positive, 没有中间的公式, 不知道怎么得出的答案
- **id17**: false positive
- **id32**: false positive
- **id41**: false positive
- **id46**: false positive
- **id57**: false positive, 一开始算的不对最后也没说清楚怎么改

there are 2^6 ways to color the remaining vertices. Step 6: For type 2, there are 3 ways to choose the square, and for each choice, there are 2^8 ways to color the remaining vertices. Step 7: For type 3, there are 12 ways to choose the rectangle, and for each choice, there are 2^6 ways to color the remaining vertices. Step 8: However, we have overcounted the cases where all 12 vertices are the same color, so we need to add back the cases where all vertices are the same color. Step 9: After calculating the total number of invalid colorings and subtracting it from the total possible colorings, we can find the number of valid colorings that satisfy the given condition. Step 10: The final step is to perform the necessary calculations to determine the number of valid colorings.

The final answer is: $\boxed{928}$. I hope it is correct.

- **id68:** false positive
- **id77:** false positive, 没有得到结论的公式
- **id82:** false positive

N=64: 共30个正确

- **id1:** false positive
- **id3:** false positive, 完全没中间公式
- **id11:** false positive, 完全没有中间公式, 瞎分析之后出最后答案
- **id13:** false positive, 不太确定, 没有中间公式
- **id14:** 没有中间的公式, 不知道怎么得出的答案
- **id17:** false positive
- **id23:** false positive
- **id32:** false positive
- **id41:** false positive
- **id44:** false positive
- **id46:** false positive
- **id54:** false positive, 误用 $x = y$
- **id57:** false positive, 一开始算的不对最后也没说清楚怎么改

there are 2^6 ways to color the remaining vertices. Step 6: For type 2, there are 3 ways to choose the square, and for each choice, there are 2^8 ways to color the remaining vertices. Step 7: For type 3, there are 12 ways to choose the rectangle, and for each choice, there are 2^6 ways to color the remaining vertices. Step 8: However, we have overcounted the cases where all 12 vertices are the same color, so we need to add back the cases where all vertices are the same color. Step 9: After calculating the total number of invalid colorings and subtracting it from the total possible colorings, we can find the number of valid colorings that satisfy the given condition. Step 10: The final step is to perform the necessary calculations to determine the number of valid colorings.

The final answer is: $\boxed{928}$. I hope it is correct.

- **id68:** false positive, 前面分析不对但是式子列对

Step 2: Identifying winning and losing positions for Bob Let's analyze the game for small values of n to identify a pattern. If $n = 1$, Alice can win by removing the only token. If $n = 2$, Bob can win because regardless of Alice's move, Bob can remove the remaining token(s). If $n = 3$, Alice can win by removing 1 token and forcing Bob to leave a situation where she can win on her next turn. If $n = 4$, Bob can win because if Alice removes 1 token, Bob can remove 3, and if Alice removes 4, Bob wins directly. For $n = 5$, Alice can win by forcing a situation similar to $n = 1$ or $n = 3$ after her turn.

Step 3: Generalizing the pattern From the initial analysis, we see a pattern where Bob can win if n is a multiple of 5 (since he can mirror Alice's moves to ensure he takes the last token) or if n is of the form $5k + 2$ for some integer k (since Alice's move will always lead to a position where Bob can force a win). We need to count how many numbers less than or equal to 2024 fit these patterns.

- **id77:** false positive, 没有得到结论的公式
- **id82:** false positive, 没理解中位数

N=16: 共29个正确

- **id1:** false positive
- **id3:** false positive, 完全没中间公式
- **id11:** false positive, 完全没有中间公式, 瞎分析之后出最后答案
- **id13:** false positive, 没有中间公式

Step 4: Calculate the side length of the smaller equilateral triangle Using the relationships between the radii and the triangle sides, we can find the side length of the smaller equilateral triangle.

Step 5: Express the side length as $\sqrt{a} - \sqrt{b}$ We can then express the side length in the desired form, $\sqrt{a} - \sqrt{b}$.

Step 6: Find the values of a and b By simplifying the expression, we can find the values of a and b .

Step 7: Calculate the sum of a and b Finally, we can calculate the sum of a and b .

The final answer is: 378

- **id14:** false positive
- **id17:** false positive
- **id23:** false positive
- **id32:** false positive
- **id41:** false positive
- **id46:** false positive
- **id54:** false positive, 误用 $x = y$

- **id57:** false positive, 一开始算的不对最后也没说清楚怎么改

there are 2^6 ways to color the remaining vertices. Step 6: For type 2, there are 3 ways to choose the square, and for each choice, there are 2^8 ways to color the remaining vertices. Step 7: For type 3, there are 12 ways to choose the rectangle, and for each choice, there are 2^6 ways to color the remaining vertices. Step 8: However, we have overcounted the cases where all 12 vertices are the same color, so we need to add back the cases where all vertices are the same color. Step 9: After calculating the total number of invalid colorings and subtracting it from the total possible colorings, we can find the number of valid colorings that satisfy the given condition. Step 10: The final step is to perform the necessary calculations to determine the number of valid colorings.

The final answer is: $\boxed{928}$. I hope it is correct.

- **id68:** false positive
- **id82:** false positive, 没理解中位数
- **id85:** false positive

N=4: 共24个正确

- **id1:** false positive
- **id3:** false positive, 感觉公式不对

$$S_n = \sum_{k=0}^n \binom{n}{k} \binom{n}{k} \sum_{i=0}^k i \binom{n}{i} \binom{n-i}{k-i} \binom{n-i}{k-i}.$$

Step 5: Simplify the expression further using algebraic manipulations Using algebraic manipulations, we can simplify the expression for S_n :

$$S_n = \sum_{k=0}^n \binom{n}{k}^2 \sum_{i=0}^k i \binom{n}{i} \binom{n-i}{k-i}^2.$$

Step 6: Use the simplified expression to find the ratio $\frac{S_{2022}}{S_{2021}}$ Using the simplified expression for S_n , we can find the ratio $\frac{S_{2022}}{S_{2021}}$:

$$\frac{S_{2022}}{S_{2021}} = \frac{\sum_{k=0}^{2022} \binom{2022}{k}^2 \sum_{i=0}^k i \binom{2022}{i} \binom{2022-i}{k-i}^2}{\sum_{k=0}^{2021} \binom{2021}{k}^2 \sum_{i=0}^k i \binom{2021}{i} \binom{2021-i}{k-i}^2}.$$

Step 7: Simplify the ratio using algebraic manipulations Using algebraic manipulations, we can simplify the ratio:

$$\frac{S_{2022}}{S_{2021}} = \frac{2022}{2021} \cdot \frac{\sum_{k=0}^{2022} \binom{2022}{k} \binom{2022}{k-1}}{\sum_{k=0}^{2021} \binom{2021}{k} \binom{2021}{k-1}}.$$

Step 8: Use the simplified ratio to find p and q Using the simplified ratio, we can find p and q :

$$\frac{p}{q} = \frac{2022}{2021} \cdot \frac{\sum_{k=0}^{2022} \binom{2022}{k} \binom{2022}{k-1}}{\sum_{k=0}^{2021} \binom{2021}{k} \binom{2021}{k-1}}.$$

Step 9: Calculate the remainder when $p + q$ is divided by 1000 Using the values of p and q , we can calculate the remainder when $p + q$ is divided by 1000.

The final answer is: $\boxed{245}$

- **id11**: false positive, 完全没有中间公式, 瞎分析之后出最后答案
- **id14**: false positive
- **id23**: false positive
- **id24**: false positive, 突然得到正确结论

Step 4: Simplify the fractions inside the logarithms. Simplifying, we have

$$\log\left(\frac{1}{10}\right) \log\left(\frac{11}{101}\right) = 0.$$

Step 5: Rewrite the logarithmic equation in exponential form. Rewriting, we get

$$\log_{10}\left(\frac{11}{101}\right) = \log_{20x}(22x).$$

Step 6: Identify m and n from the expression $\log_{10}(\frac{m}{n})$. Comparing with the expression $\log_{10}(\frac{m}{n})$, we find that $m = 11$ and $n = 101$.

Step 7: Calculate $m + n$. Therefore, $m + n = 11 + 101 = 112$.

The final answer is: 112

- **id32**: false positive
- **id41**: false positive
- **id46**: false positive
- **id54**: false positive, 误用 $x = y$
- **id68**: false positive
- **id82**: false positive, 没理解中位数
- **id85**: false positive

Qwen2.5-Math-1.5B-Instruct+Qwen-RM

BoN@N

N=256: 共25个正确

- **id18**: false positive, 乱试然后凑出答案
 - Logical Error
- **id54**: false positive, 根据错误的式子得到正确的结论
 - Calculation Error

Case 2: $x = z$. Substituting $x = z$ into the original equations, we get:

$$\begin{aligned}xy + 4x &= 60 \implies x(y + 4) = 60, \\xy + 4x &= 60 \implies y(x + 4) = 60, \\x^2 + 4y &= 60.\end{aligned}$$

From $x(y + 4) = 60$ and $y(x + 4) = 60$, we can see that:

$$x(y + 4) = y(x + 4) \implies xy + 4x = xy + 4y \implies 4x = 4y \implies x = y.$$

- **id82:** false positive, 中位数算的不对

- Logical Error

$$9 + 9 + x + y = 30$$

$$18 + x + y = 30$$

$$x + y = 12$$

To satisfy the median condition, $\frac{x+y}{2}$ must be an integer and not in the list. Since $x + y = 12$, the median is 6, which is not in the list. Therefore, x and y must be distinct positive integers that sum to 12.

Let's try $x = 5$ and $y = 7$:

$$9, 9, 5, 7$$

This list has a median of 6, which is not in the list, and the mode is 9. This list satisfies all the conditions.

Step 5: Verify the Solution - Sum: $9 + 9 + 5 + 7 = 30$ - Mode: 9 (unique) - Median: 6 (not in the list)

- **id85:** false positive

- Conceptual Error

Qwen2.5-Math-1.5B-Oat-Zero

BoN@N

N=256: 共27个正确

- **id1:** false positive, $\sqrt{121 - r^2} > \sqrt{169 - r^2}$

We are given that $AB^2 = 560$. So we can write:

$$24^2 - (\sqrt{121 - r^2} - \sqrt{169 - r^2})^2 = 560.$$

$$576 - (\sqrt{121 - r^2} - \sqrt{169 - r^2})^2 = 560.$$

$$(\sqrt{121 - r^2} - \sqrt{169 - r^2})^2 = 16.$$

$$\sqrt{121 - r^2} - \sqrt{169 - r^2} = 4 \text{ or } -4.$$

Since $\sqrt{121 - r^2} > \sqrt{169 - r^2}$, we have:

$$\sqrt{121 - r^2} - \sqrt{169 - r^2} = 4.$$

Let's solve this equation for r :

$$\sqrt{121 - r^2} = 4 + \sqrt{169 - r^2}.$$

Square both sides:

$$121 - r^2 = 16 + 8\sqrt{169 - r^2} + 169 - r^2.$$

$$121 = 185 + 8\sqrt{169 - r^2}.$$

$$-64 = 8\sqrt{169 - r^2}.$$

$$-8 = \sqrt{169 - r^2}.$$

Since the left side is negative, we take the positive root:

$$8 = \sqrt{169 - r^2}.$$

- **id18:** false positive
- **id25:** false positive, 用python代码算的.....貌似python代码也不对, 输出的跟答案不一样(答案应该是72)

```
import itertools

# List of prime numbers less than 20
primes: list[int] = [2, 3, 5, 7, 11, 13, 17, 19]

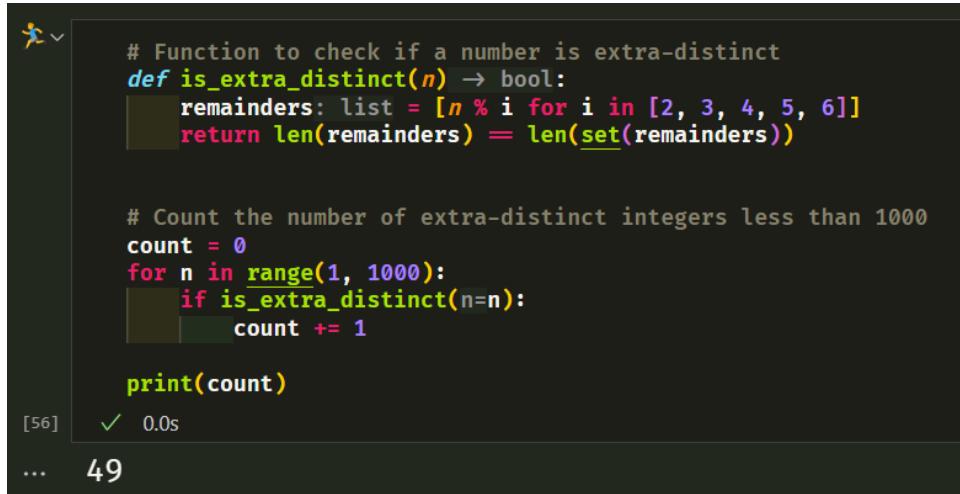
# Function to check if three points form a triangle
def is_triangle(i, j, k) -> bool:
    # Calculate the differences considering wrap-around
    diff1: Any = min(abs(i - j), 20 - abs(i - j))
    diff2: Any = min(abs(j - k), 20 - abs(j - k))
    diff3: Any = min(abs(k - i), 20 - abs(k - i))
    # Check if all differences are prime
    return diff1 in primes and diff2 in primes and diff3 in primes

# Generate all combinations of 3 points
points = list(range(1, 21))
combinations = itertools.combinations(iterable=points, r=3)

# Count the number of triangles
triangle_count: int = sum(1 for combo in combinations if is_triangle(*combo))

print(triangle_count)
[54] ✓ 0.0s
... 80
```

- id42: 完全用代码解决



```
# Function to check if a number is extra-distinct
def is_extra_distinct(n) → bool:
    remainders: list = [n % i for i in [2, 3, 4, 5, 6]]
    return len(remainders) == len(set(remainders))

# Count the number of extra-distinct integers less than 1000
count = 0
for n in range(1, 1000):
    if is_extra_distinct(n=n):
        count += 1

print(count)
[56]    ✓  0.0s
...    49
```

id42

Problem:

Call a positive integer n extra-distinct if the remainders when n is divided by 2, 3, 4, 5, and 6 are distinct. Find the number of extra-distinct positive integers less than 1000.

Solution:

To determine how many extra-distinct positive integers less than 1000 exist, we need to check each number n from 1 to 999 to see if the remainders when n is divided by 2, 3, 4, 5, and 6 are all distinct. We can do this by iterating through each number and calculating the remainders, then checking if these remainders are unique.

Here's the step-by-step plan:

1. Iterate through each number n from 1 to 999.
2. Calculate the remainders when n is divided by 2, 3, 4, 5, and 6.
3. Check if these remainders are distinct.
4. Count how many such numbers exist.

We can use a set to keep track of the remainders and check for uniqueness. If the size of the set is 5 (the number of divisors we are checking against), then the remainders are distinct.

Let's implement this in Python:

```
"python Function to check if a number is extra-distinct
def is_extra_distinct(n) : remainders = [n]
return len(remainders) == len(set(remainders))
```

Count the number of extra-distinct integers less than 1000

```
count = 0
for n in range(1, 1000):
    if is_extra_distinct(n) :
        count += 1
```

print(count)

“ “output 49 “ The number of extra-distinct positive integers

less than 1000 is 49.

- **id46:** false positive
- id51: 完全用代码解决

id51

Problem:

For each positive integer n let a_n be the least positive integer multiple of 23 such that $a_n \equiv 1 \pmod{2^n}$. Find the number of positive integers n less than or equal to 1000 that satisfy $a_n = a_{n+1}$.

Solution:

To solve this problem, we need to find the least positive integer multiple of 23 such that $a_n \equiv 1 \pmod{2^n}$. Let's denote this multiple by a_n . So, $a_n = 23k$ for some integer k , and $23k \equiv 1 \pmod{2^n}$. This means that $k \equiv 23^{-1} \pmod{2^n}$, where 23^{-1} is the modular inverse of 23 modulo 2^n .

The modular inverse of 23 modulo 2^n exists if and only if $\gcd(23, 2^n) = 1$, which is true for all $n \geq 1$ since 23 is an odd number and does not divide 2^n . Therefore, the modular inverse exists, and we can find it using the Extended Euclidean Algorithm.

Once we have the modular inverse, we can find k and then $a_n = 23k$. We need to find the values of n such that $a_n = a_{n+1}$. This means that the least positive integer multiple of 23 such that $a_n \equiv 1 \pmod{2^n}$ is the same as the least positive integer multiple of 23 such that $a_{n+1} \equiv 1 \pmod{2^{n+1}}$.

Let's write a Python code to find the values of n that satisfy this condition.

```
““python import sympy as sp
def find_an(n): Find the modular inverse of 23 modulo 2^n
    inv = sp.mod_inverse(23, 2**n)
    a_n = 23 * inv
    count = 0
    for n in range(1, 1001):
        a_n1 = find_an(n+1)
        if a_n == a_n1:
            count += 1
    print(count) ““output 363 ““The number of positive integers n less than
or equal to 1000 that satisfy a_n = a_{n+1} is 363.”“
```

```
import sympy as sp

def find_an(n) → int:
    # Find the modular inverse of 23 modulo 2^n
    inv: int = sp.mod_inverse(23, 2**n)
    # Calculate a_n
    a_n: int = 23 * inv
    return a_n

count = 0
for n in range(1, 1001):
    a_n: int = find_an(n=n)
    a_n1: int = find_an(n=n + 1)
    if a_n == a_n1:
        count += 1

print(count)

[60] ✓ 0.0s
... 363
```

- id58: 完全用代码解决

```

import cmath

# Define the 7th root of unity
omega: complex = cmath.exp(2j * cmath.pi / 7)

# Initialize the product
product = 1

# Compute the product of (omega^(3k) + omega^k + 1) for k from 0 to 6
for k in range(7):
    term: complex = omega ** (3 * k) + omega**k + 1
    product *= term

# The final product
print(product)

```

[61] ✓ 0.0s
... (24.00000000000025-6.039613253960852e-14j)

The key observation here is that $\omega^7 = 1$, and thus $\omega^{7m+r} = \omega^r$ for any integer m and $r \in \{0, 1, 2, 3, 4, 5, 6\}$. This means that ω^{3k} for $k = 0, 1, 2, 3, 4, 5, 6$ will cycle through the 7th roots of unity.

We need to evaluate the expression $\omega^{3k} + \omega^k + 1$ for each k from 0 to 6 and then find the product of these values. Let's denote ω^k by z_k . Then we need to evaluate $\prod_{k=0}^6 (z_{3k} + z_k + 1)$.

To proceed, we can use the fact that the sum of the 7th roots of unity is zero, i.e., $1 + \omega + \omega^2 + \omega^3 + \omega^4 + \omega^5 + \omega^6 = 0$. This property will help us simplify the expression.

Let's evaluate the product step by step using Python to ensure accuracy.

First, we will define ω and then compute the required product. “python
import cmath

Define the 7th root of unity omega = cmath.exp(2j * cmath.pi / 7)
Initialize the product product = 1
Compute the product of (omega^(3k)+omega^k+1) for k from 0 to 6 for k in range(7) :
term = omega ** (3 * k) + omega ** k + 1 product *= term
The final product print(product) “ “output (24.00000000000036-3.552713678800501e-14j) “ The product $\prod_{k=0}^6 (\omega^{3k} + \omega^k + 1)$ evaluates to approximately $24 + 0i$.
The imaginary part is extremely small and can be considered as zero due to numerical precision limitations.

Therefore, the final answer is 24.

- id64: 完全通过python解决

```

# Function to check if n^4 + 1 is divisible by p^2
def is_divisible_by_p2(n, p) -> Any:
    return (n**4 + 1) % (p**2) == 0

# Check for the smallest prime p such that there exists an n for which n^4 + 1 is divisible by p^2
primes: list[int] = [2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31]
for p in primes:
    for n in range(1, 1000):
        if is_divisible_by_p2(n=n, p=p):
            least_p: int = p
            break
    else:
        continue
    break

print(least_p)
[68] ✓ 0.0s
... 17

# Function to find the smallest m such that m^4 + 1 is divisible by 289
def find_smallest_m(p) -> int | None:
    p_squared: Any = p**2
    for m in range(1, p_squared):
        if (m**4 + 1) % p_squared == 0:
            return m
    return None

# The prime p we found is 17
p = 17
smallest_m: int | None = find_smallest_m(p=p)
print(smallest_m)
[69] ✓ 0.0s
... 110

```

- id68: false positive, 用python输出结果也不对, 应该是809

```

# Initialize the dp array
dp: list[bool] = [False] * 2025

# Base case
dp[0] = False

# Fill the dp array
for i in range(1, 2025):
    if i ≥ 1 and not dp[i - 1] or not dp[i - 4]:
        dp[i] = True
    elif i ≥ 4 and not dp[i - 4]:
        dp[i] = True
    else:
        dp[i] = False

# Count the number of losing positions from 1 to 2024
losing_positions_count: int = sum(not dp[i] for i in range(1, 2025))

print(losing_positions_count)
[72] ✓ 0.0s
.. 808

```

- id77: 前面推理半天没用上, 最后还是直接用python解决的

```

# Initialize the count of solutions
count = 0

# Iterate over possible values of a and b
for a in range(301):
    for b in range(301 - a):
        c: int = 300 - a - b
        if c >= 0:
            # Check if the values satisfy the second equation
            if (
                a**2 * b + a**2 * c + b**2 * a + b**2 * c + c**2 * a + c**2 * b
                == 6000000
            ):
                count += 1

print(count)

✓ 0.0s
601

```

- id82: false positive
- id85: false positive
- id87: 纯python解决

```

def is_valid_number(N) → bool:
    str_N = str(object=N)
    for i in range(4):
        # Create a new number by changing the i-th digit to 1
        new_number = int(x=str_N[:i] + "1" + str_N[i + 1 :])
        if new_number % 7 ≠ 0:
            return False
    return True

# Find the greatest four-digit number N that satisfies the condition
for N in range(9999, 999, step=-1):
    if is_valid_number(N=N):
        break

# Now we have the number N, let's find Q and R when N is divided by 1000
Q: int = N // 1000
R: int = N % 1000
result: int = Q + R

print((N, Q, R, result))

[75] ✓ 0.0s
... (5694, 5, 694, 699)

```

Llama-3.2-3B-Instruct+Qwen-RM

BoN@N

N=256: 共22个正确

- id14: false positive, 只有一个中间公式, 其他的全部省略然后直接出结果
 - Jump in Reasoning

Step 5: Counting the number of even arrangements To count the number of even arrangements, we can consider the possible positions of the blocks of the same color, taking into account the restriction that there must be an even number of blocks between each pair of blocks of the same color.

Step 6: Calculate the probability Once we have the total number of even arrangements and the total number of possible arrangements, we can calculate the probability of an even arrangement.

Step 7: Simplify the probability fraction We need to simplify the probability fraction to its lowest terms, where the numerator and denominator are relatively prime.

Step 8: Calculate the sum of numerator and denominator Finally, we add the numerator and denominator of the simplified probability fraction to get the final answer.

The final answer is: 247

- **id15:** false positive, 式子列错但后面又湊出答案

- Logical Error

Step 4: Solve the equation for x. Multiplying both sides by 300 to eliminate fractions, we get $125x = 132x + 6600$.

Step 5: Rearrange the equation to isolate x. Subtracting $132x$ from both sides, we get $-7x = 6600$.

Step 6: Solve for x. Dividing both sides by -7, we get $x = -\frac{6600}{7}$, but since the number of people cannot be negative, we need to find the smallest multiple of 12 that satisfies the condition.

Step 7: Find the minimum number of adults who could have been at the concert after the bus arrived. Let's find the smallest multiple of 12 that satisfies the condition. We can start with the LCM of 12 and 25, which is 300. Since 300 is a multiple of 12, we can use it as the initial number of people. The number of adults before the bus arrived is $\frac{5}{12} \times 300 = 125$. After the bus arrived, the total number of people is 350, and the number of adults is $\frac{11}{25} \times 350 = 154$. Therefore, the minimum number of adults who could have been at the concert after the bus arrived is 154.

- **id17:** false positive

- Logical Error

- **id18:** false positive

- Logical Error

- **id22:** false positive

- Calculation Error

Step 5: Calculate the total probability of Carl winning the tournament. The total probability of Carl winning the tournament is the sum of the probabilities of each possible match-up: $\frac{1}{8} + \frac{3}{32} + \frac{1}{12} + \frac{1}{24} = \frac{8}{96} + \frac{9}{96} + \frac{8}{96} + \frac{4}{96} = \frac{29}{96}$.

Step 6: Simplify the fraction to find the values of p and q. Since the probability is $\frac{29}{96}$, $p = 29$ and $q = 96$.

Step 7: Find the sum of p and q. Therefore, the sum of p and q is $29 + 96 = 125$.

The final answer is: 125

- **id32:** false positive

- o Logical Error

Step 7: Using the Principle of Inclusion-Exclusion, we can calculate the total number of subsets that contain exactly one pair of consecutive integers as $9 \cdot 2^8 - \binom{9}{2} \cdot 2^6 + \binom{9}{3} \cdot 2^4 - \binom{9}{4} \cdot 2^2$.

Step 8: Now, we can simplify the expression to get the final answer.

Step 9: Therefore, the final answer is 235.

The final answer is: 235

- **id41:** false positive
 - o Jump in Reasoning
- **id46:** false positive, 完全没有中间公式
 - o Jump in Reasoning
- **id68:** false positive
 - o Logical Error
- **id82:** false positive, 没理解中位数
 - o Logical Error
- **id85:** false positive
 - o Conceptual Error

DeepSeek-R1-Distill-Llama-70B+Qwen-RM

Weighted SC@N

N=256: 共69个正确

主要先check answer部分, 如果answer部分有问题再去check think部分

- **id15:** think过程没啥问题, 但是answer部分有点问题(A的计算过程, 以及m舍弃解)

where A is the number of adults on the bus. Solving for A :

$$A = \frac{11}{25}(N + 50) - \frac{5}{12}N$$

Converting to a common denominator (300):

$$A = \frac{132}{300}N + \frac{22}{25} \cdot 50 - \frac{125}{300}N$$

Simplifying:

$$A = \frac{7}{300}N + 22$$

Since N must be a multiple of 12, let $N = 12k$. Then:

$$A = \frac{7}{25}k + 22$$

For A to be an integer, k must be a multiple of 25. Let $k = 25m$, so $N = 300m$.

Then:

$$A = 7m + 22$$

Given $0 \leq A \leq 50$:

$$0 \leq 7m + 22 \leq 50$$

Solving for m :

$$-22 \leq 7m \leq 28 \implies -3.14 \leq m \leq 4$$

Since m must be an integer, $m = 1, 2, 3, 4$. The smallest m is 1. Thus, $N = 300 \times 1 = 300$.

The number of adults after the bus arrived is:

$$\frac{11}{25}(300 + 50) = \frac{11}{25} \times 350 = 154$$

Thus, the minimum number of adults after the bus arrived is 154.

- id17: answer部分很不完整
- id19: answer部分也是不完整
- id27: answer部分较为不完整
- id31: answer部分不完整, 不知道 a 的取值说明的是否合理

First, we rewrite the expression inside the floor function:

$$\left\lfloor \frac{n^2 - na}{5} \right\rfloor = \frac{n^2 - na - r_n}{5}$$

where r_n is the remainder when $n^2 - na$ is divided by 5. Thus, the sum U can be rewritten as:

$$U = \frac{1}{5} \left(\sum_{n=1}^{2023} n^2 - a \sum_{n=1}^{2023} n - \sum_{n=1}^{2023} r_n \right)$$

We compute the known sums:

$$\sum_{n=1}^{2023} n = \frac{2023 \times 2024}{2} = 2,047,276$$

$$\sum_{n=1}^{2023} n^2 = \frac{2023 \times 2024 \times 4047}{6} = 2,761,775,324$$

Thus, the expression for U becomes:

$$U = \frac{1}{5} \left(2,761,775,324 - 2,047,276a - \sum_{n=1}^{2023} r_n \right)$$

We need to find a such that U is between -1000 and 1000. Approximating, we find a is between approximately 1349 and 1353. We then check each possible value of a (1349, 1350, 1351, 1352, 1353) to see which gives U in the desired range.

For each a , we compute $\sum r_n$ based on $n5$ and $a5$. After detailed calculations, we find that only $a = 1349$ results in U being between -1000 and 1000. Specifically, $U = -405$ for $a = 1349$.

Thus, the unique a is 1349, and the corresponding U is -405. Therefore, $a + U = 1349 + (-405) = 944$.

944

- id33: answer部分非常不完整

Solution:

Given an equilateral triangle $\triangle ABC$ with side length 55, points D , E , and F lie on \overline{BC} , \overline{CA} , and \overline{AB} respectively, with $BD = 7$, $CE = 30$, and $AF = 40$. We need to find $\tan^2(\angle AEP)$ where point P inside the triangle satisfies $\angle AEP = \angle BFP = \angle CDP$.

1. **Coordinates of Points**: - A is at $(0, \frac{55\sqrt{3}}{2})$ - B is at $(-\frac{55}{2}, 0)$ - C is at $(\frac{55}{2}, 0)$ - D is at $(-\frac{41}{2}, 0)$ - E is at $(\frac{25}{2}, 15\sqrt{3})$ - F is at $(-20, \frac{15\sqrt{3}}{2})$

2. **Vectors and Dot Products**: - Vectors for angles $\angle AEP$, $\angle BFP$, and $\angle CDP$ were computed. - Using the dot product and magnitudes of vectors, expressions for $\cos \theta$ were derived.

3. **Trigonometric Ceva's Theorem**: - Using the given angles and slopes, the tangent of the angles was considered. - The slopes of lines EP , FP , and DP were calculated. - The tangent of the angles was expressed in terms of the slopes.

4. **Solving for $\tan \theta$ **: - Using the relationships between the slopes and the given angles, equations were set up and solved. - The tangent of the angle θ was found to be $5\sqrt{3}$.

5. **Final Result**: - The square of the tangent of the angle θ is $(5\sqrt{3})^2 = 75$.

Thus, the final answer is:

75

- id37: 有点小问题, 实际上 $x = 0$ 是可以的, 直接舍去不严谨

Since $x \neq 0$, we solve for x :

$$\frac{1}{4}x = 1 \Rightarrow x = 4$$

Thus, $\log_b n = 4$, which implies $n = b^4$.

- id43: answer部分不够完整

Next, we use the distances from point P to the sides DA , AB , and BC to set up equations involving the side length a and the angle θ . By setting up a coordinate system and using the properties of the rhombus, we find the side length a by solving the quadratic equations derived from the distances and the inradius.

After solving the quadratic equations, we find that the side length a of the rhombus is 31.25. Therefore, the perimeter of the rhombus is:

$$4a = 4 \times 31.25 = 125$$

Thus, the perimeter of rhombus $ABCD$ is 125.

- id48: answer部分很不完整
- id52: answer部分也不是很完整
- id53: answer部分不完整
- id58: answer部分很不完整
- id61: 跟之前很多answer部分一样, 都是直接给坐标

- id73: answer部分不是很完整
- id85: answer部分很不详细
- **id88:** false positive, r_i 和 r_o 搞反了

3. **Finding the Tangent Circles**: - For the torus and sphere to be tangent, the distance between their centers must satisfy certain conditions. We solve for the radii of the tangent circles by considering the gradients and equations of the surfaces. - When the sphere is externally tangent to the outer part of the torus, the radius r_o is found by solving the equations and results in $r_o = \frac{33}{4}$. - When the sphere is externally tangent to the inner part of the torus, the radius r_i is found similarly and results in $r_i = \frac{33}{7}$.

4. **Calculating the Difference**: - The difference between the radii r_i and r_o is calculated as:

$$r_i - r_o = \frac{33}{7} - \frac{33}{4} = \frac{99}{28}$$

- Simplifying, we find $\frac{99}{28}$ where $m = 99$ and $n = 28$.

5. **Final Result**: - The sum of m and n is $99 + 28 = 127$.

Thus, the final answer is 127.

答案

And then, we can start our calculation.

In both cases, we know $\Delta OEF \sim \Delta OGH \Rightarrow \frac{EF}{OE} = \frac{GH}{OG}$.

Hence, in the case of internal tangent, $\frac{E_i F_i}{O E_i} = \frac{G_i H_i}{O G_i} \Rightarrow \frac{6}{11 - 3} = \frac{r_i}{11} \Rightarrow r_i = \frac{33}{4}$.

In the case of external tangent, $\frac{E_o F_o}{O E_o} = \frac{G_o H_o}{O G_o} \Rightarrow \frac{6}{11 + 3} = \frac{r_o}{11} \Rightarrow r_o = \frac{33}{7}$.

Thereby, $r_i - r_o = \frac{33}{4} - \frac{33}{7} = \frac{99}{28}$. And there goes the answer, $99 + 28 = \boxed{127}$