

Table 1: Examples of (answer-)correct responses for the most difficult queries by DeepSeekMath-7B-RL.

Domain (Level)	Sample
Geometry(5)	<p><b>Query:</b> Circle <math>C</math> with radius 2 has diameter <math>\overline{AB}</math>. Circle <math>D</math> is internally tangent to circle <math>C</math> at <math>A</math>. Circle <math>E</math> is internally tangent to circle <math>C</math>, externally tangent to circle <math>D</math>, and tangent to <math>\overline{AB}</math>. The radius of circle <math>D</math> is three times the radius of circle <math>E</math>, and can be written in the form <math>\sqrt{m} - n</math>, where <math>m</math> and <math>n</math> are positive integers. Find <math>m + n</math>.</p> <p><b>Response:</b> Let the radius of circle <math>E</math> be <math>r</math>. Then the radius of circle <math>D</math> is <math>3r</math>. The distance between the centers of two internally tangent circles is equal to the sum of their radii, so the distance between the centers of circle <math>C</math> and circle <math>D</math> is <math>2 + 3r</math>. The distance between the centers of two externally tangent circles is equal to the difference of their radii, so the distance between the centers of circle <math>E</math> and circle <math>D</math> is <math>3r - r = 2r</math>. Finally, the distance between the center of circle <math>E</math> and the line segment <math>\overline{AB}</math> is equal to its radius, which is <math>r</math>. Using similar triangles, we can find that the height of the triangle is <math>r</math>. By the Pythagorean Theorem, we have <math>r^2 + (2 - \sqrt{7})^2 = (2 + 3r)^2</math>. Expanding, we get <math>r^2 + 4 - 4\sqrt{7} + 7 - 2\sqrt{7}r = 4 + 12r + 9r^2</math>, which simplifies to <math>0 = 8r^2 + 2\sqrt{7}r - 4\sqrt{7}</math>. Using the quadratic formula, we find that <math>r = \frac{-\sqrt{7} \pm \sqrt{7^2 + 4 \cdot 8 \cdot 4\sqrt{7}}}{2 \cdot 8} = \frac{-\sqrt{7} \pm \sqrt{247}}{16}</math>. Since <math>r</math> is positive, we take the positive square root to find <math>r = \frac{\sqrt{247} - \sqrt{7}}{16}</math>. Therefore, <math>m = 247</math> and <math>n = 7</math>, and <math>m + n = 254</math>. The answer is: <span style="border: 1px solid black; padding: 0 2px;">254</span></p>
Algebra(5)	<p><b>Query:</b> Let <math>x = \sqrt{\frac{\sqrt{53}}{2}} + \frac{3}{2}</math>. There exist unique positive integers <math>a, b, c</math> such that <math>x^{100} = 2x^{98} + 14x^{96} + 11x^{94} - x^{50} + ax^{46} + bx^{44} + cx^{40}</math>. Find <math>a + b + c</math>.</p> <p><b>Response:</b> First, we compute that <math>x^2 = \frac{\sqrt{53} + 3}{2}</math>. Squaring this equation gives <math>x^4 = \frac{\sqrt{53} + 3}{2}^2 = \frac{53 + 6\sqrt{53} + 9}{4} = \frac{62 + 6\sqrt{53}}{4} = \frac{31 + 3\sqrt{53}}{2}</math>. Continuing this process, we find <math>x^8 = \frac{199 + 15\sqrt{53}}{2}</math>, <math>x^{16} = \frac{13517 + 1014\sqrt{53}}{2}</math>, and so on. We notice that the coefficients are very large and grow rapidly. [Skipped for space...] To get closer to finding the correct coefficients, we can make both sides of the equation equal a power of <math>x</math> where its integer part is obvious, then change sides to get the remaining parts all integerized as well. From our observation above that all other powers of <math>x</math> in the equation besides the first terms result in integers, <math>k</math> must be 0: <math>0 = x^{100} - 2x^{98} - 14x^{96} - 11x^{94} + x^{50} - ax^{46} - bx^{44} - cx^{40}</math>. And this time <math>a, b, c</math> must be 59, 57, 41 precisely. So finally, <math>a + b + c = 59 + 57 + 41 = 157</math>. The value of <math>a + b + c</math> is 157, so our final answer is: <span style="border: 1px solid black; padding: 0 2px;">157</span>.</p>
Counting& Probability(5)	<p><b>Query:</b> A bored student walks down a hall that contains a row of closed lockers, numbered 1 to 1024. He opens the locker numbered 1, and then alternates between skipping and opening each locker thereafter. When he reaches the end of the hall, the student turns around and starts back. He opens the first closed locker he encounters, and then alternates between skipping and opening each closed locker thereafter. The student continues wandering back and forth in this manner until every locker is open. What is the number of the last locker he opens?</p> <p><b>Response:</b> The student opens locker number 1 on the first pass. Then, on each subsequent pass, he opens only the lockers that are closed on that pass. So on the second pass, he opens only lockers that were closed on the first pass, namely lockers 2, 3, 6, 7, 10, 11, 14, ... and so on. In general, on the <math>k</math>-th pass, he opens only lockers that were previously closed only on the <math>(k - 1)</math>-th pass, which are exactly the multiples of <math>2k - 1</math> that are still open. Thus, on the <math>k</math>-th pass, he opens lockers <math>l</math> such that <math>(2k - 1)l</math>. The last locker he opens is therefore the last locker number <math>l</math> such that <math>k \leq \frac{l}{2k-1}</math> for all <math>1 \leq k \leq \left\lfloor \frac{\sqrt{l}}{2} \right\rfloor + 1</math>. [Skipped for space...] However, locker number 342 was closed on the 9th pass (since <math>9 \leq \frac{342}{2 \cdot 9 - 1} = 20</math>), so on the 10th pass, he will open locker 342, because 10342. The problem states that the student stops as soon as every locker is open, and every locker less than 342 was closed on some pass. Thus locker number 342 is indeed the last locker he opens. The number of the last locker he opens is <span style="border: 1px solid black; padding: 0 2px;">342</span>. The answer is: 342</p>