

Comparison between automatically generated task templates and manually designed templates

Experiment Setup

Aspect	Details
Dataset	We randomly sampled 500 problems from the MATH dataset.
Manual Template Preparation	We manually designed task templates for each of these 500 problems.
Automatic Template Generation	For the BoT method, we did not provide any pre-existing thought templates. Instead, we ran BoT twice on these problems to accumulate thought templates.
Evaluation Metrics	We compared the performance of manually designed templates and automatically generated templates using accuracy.

Method	Manually designed template + GPT 3.5	BoT + GPT 4.0 (Accumulated high-level thought template)	BoT + GPT 3.5 (After accumulation)
MATH-500	52.8%	73.4%	78.4%

The results demonstrate our automatically generated task templates are of better quality. Here we also use an example to show the superiority of our thought template:

Question: A truncated cone has horizontal bases with radii 18 and 2. A sphere is tangent to the top, bottom, and lateral surface of the truncated cone. What is the radius of the sphere?

Manually dsigned template	Automatically generated template
Step 1: Identify Geometric Elements	Step 1: Identify the radii of the top and bottom bases of the truncated cone (R_1 and R_2)
Step 2: Establish Relationships	Step 2: Define Key Points and Distances
Step 3: Apply Geometric Formulas	Step 3: Calculate the distance between the bases (BC)
Step 4: Form Equations	Step 4: Use the Pythagorean theorem to find the height of the trapezoid formed by the points of tangency: $CH = \sqrt{BC^2 - (R_1 - R_2)^2}$
Step 5: Solve Equations System	Step 5: The radius of the sphere is half the height CH : Radius of the sphere = $\frac{1}{2} \sqrt{BC^2 - (R_1 - R_2)^2}$
Step 6: Verify Solution	