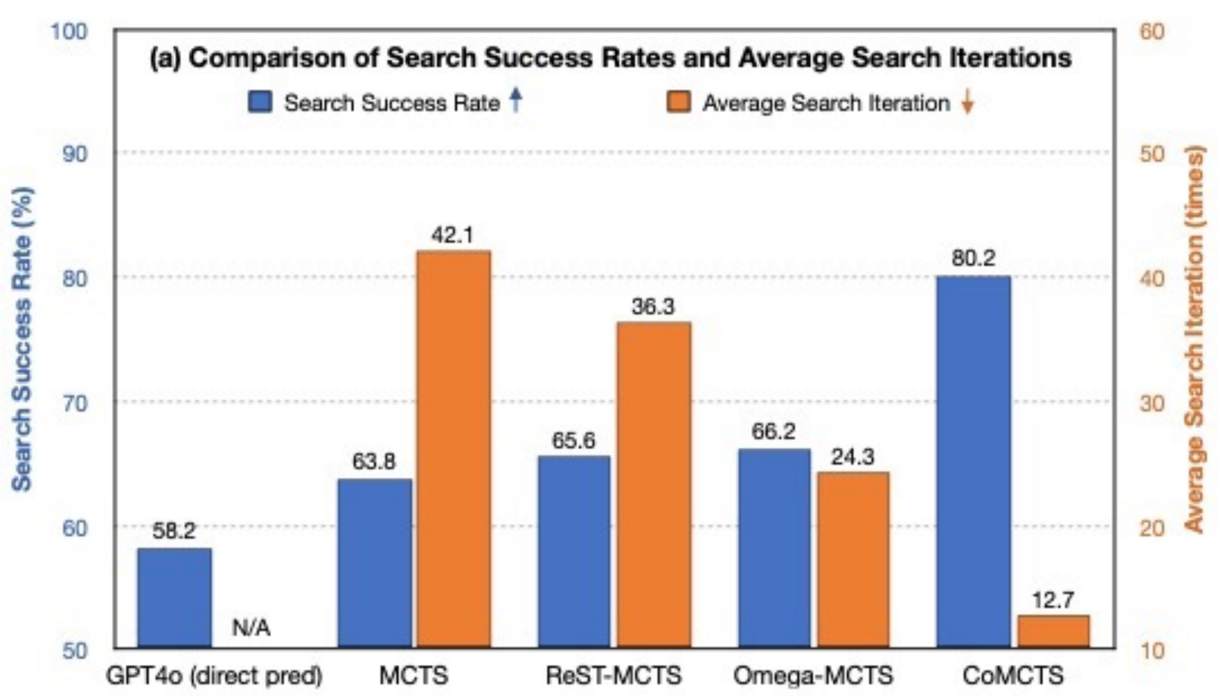
Mulberry: Empowering MLLM with o1-like Reasoning and Reflection via Collective Monte Carlo Tree Search

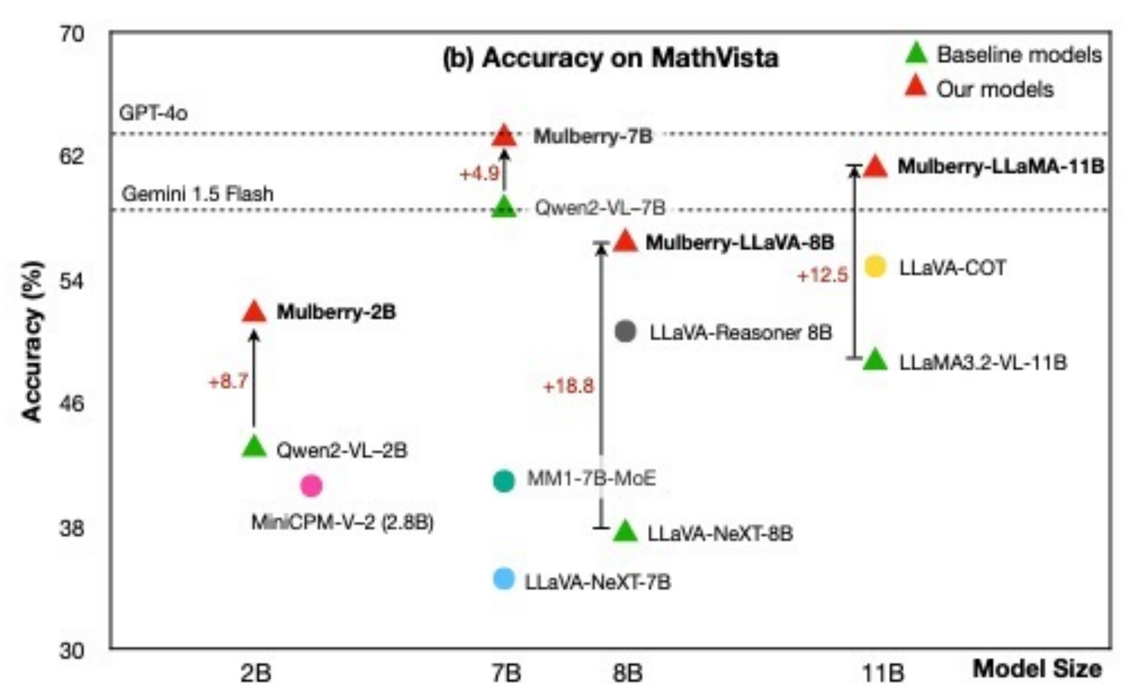
Huanjin Yao[2,3], Jiaxing Huang[1], Wenhao Wu[3], Jingyi Zhang[1], Yibo Wang[2], Shunyu Liu[1], Yingjie Wang[1], Yuxin Song[3], Haocheng Feng[3], Li Shen[4], Dacheng Tao[1]

Paper GitHub WeChat

Background (o1-like & MCTS)

Step-by-step reasoning enables LLMs to tackle complex tasks, with MCTS being a key technique. However, applying MCTS to MLLMs poses challenges in both search effectiveness and search efficiency.





Method (CoMCTS)

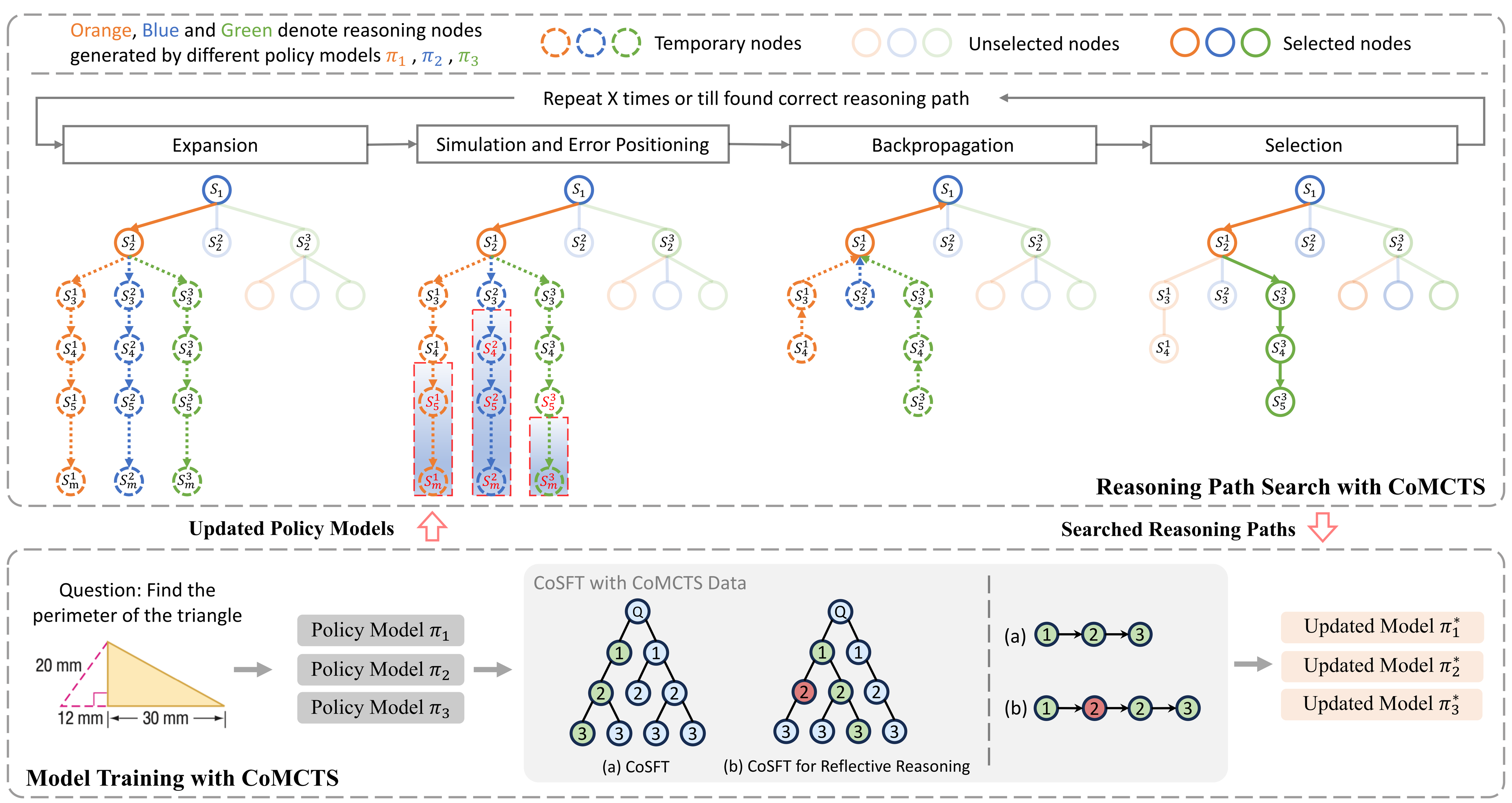
Introduce Collective Learning into MCTS for effective and efficient reasoning-path searching.

(a) Expansion: Generate diverse, complementary subsequent reasoning nodes till the end.

(b) Simulation and Error Positioning: Simulate reasoning outcomes, position error candidate nodes and prune them along with their child nodes.

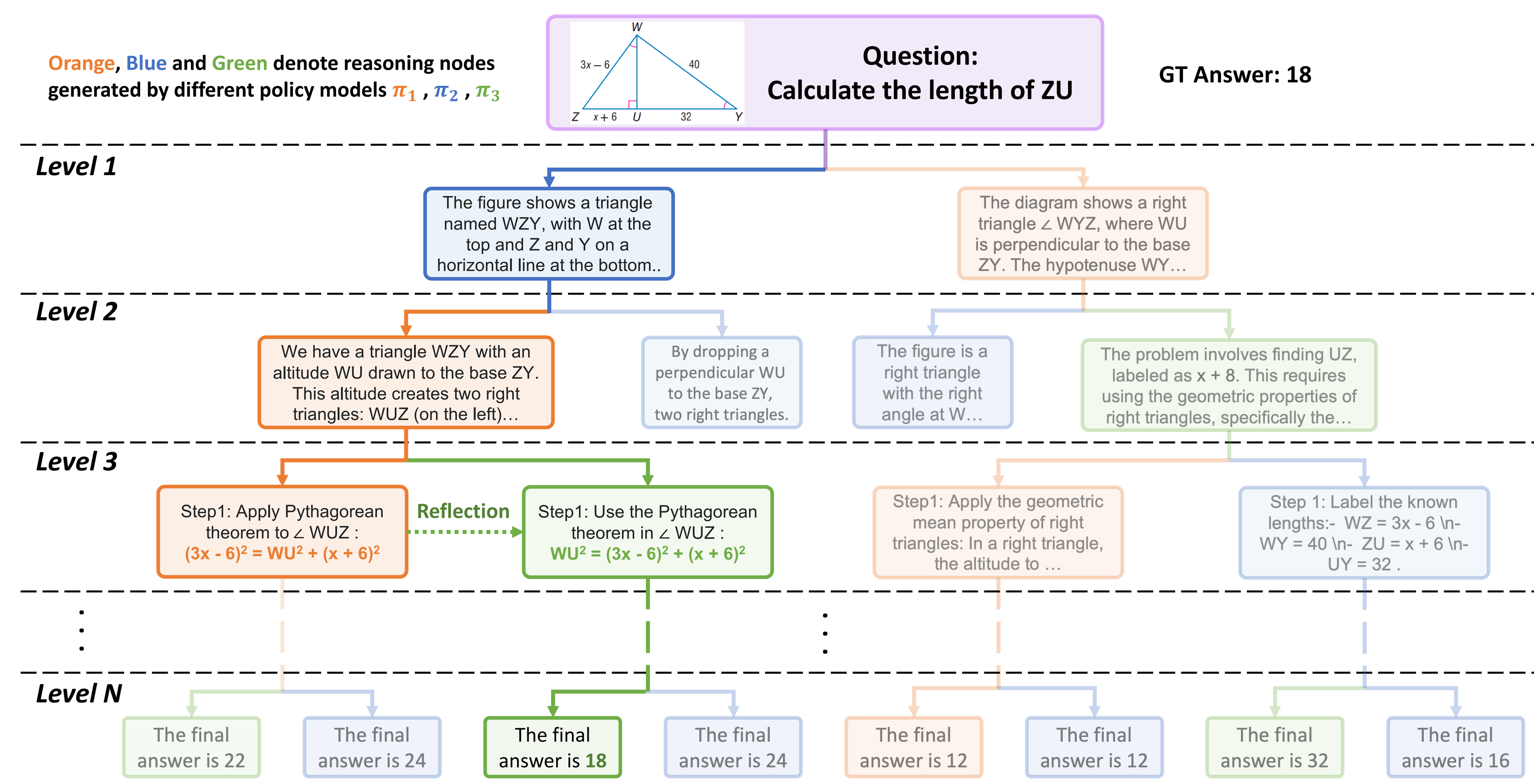
(c) Backpropagation: Update score and visit count of each node in a bottom-up manner.

(d) Selection: Select the leaf reasoning node with the highest UCB value as next start node.



Visualization of CoMCTS Data Construction

Using CoMCTS, we searched 260K reasoning and reflection step-by-step training data, named Mulberry-260K.



Question: Is Medium Orchid the minimum?
Answer yes or no.

Correct Answer: Yes

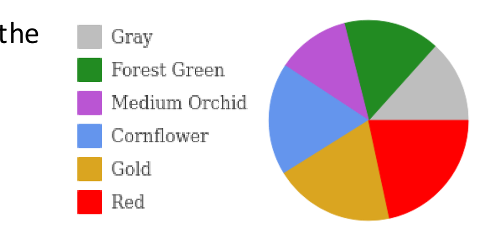


Image Description:
A pie chart divided into six segments representing different colors. Each segment has a label indicating its color. The colors include Gray, Forest Green, Medium Orchid, Cornflower, Gold, and Red.

Rationales:
- The pie chart visually represents proportions of each category (color).
- The size of each slice indicates how much of the whole each category occupies.
- Since the question asks if Medium Orchid is the minimum, we need to compare its slice size against others.

Let's think step by step.

Step 1:
Identify the sizes of all slices.
- Gray: Smallest slice
- Forest Green: Second smallest slice
- Medium Orchid: Third smallest slice
- Cornflower: Fourth largest slice
- Gold: Fifth largest slice
- Red: Largest slice

I think I have made a mistake. Let me rethink it.

Step 1:
Identify the slice labeled "Medium Orchid" on the pie chart.

Step 2:
Visually compare the size of the Medium Orchid slice to each of the other slices (Gray, Forest Green, Cornflower, Gold, Red).

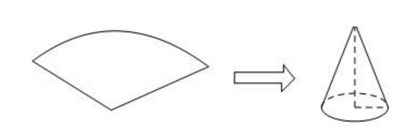
Step 3:
Determine if the Medium Orchid slice is the smallest among all the slices.

Step 4:
Validate the comparison by checking if any other slice is smaller than the Medium Orchid slice.

The final answer is:
Yes.

Results (Qualitative & Quantitative)

We conduct extensive experiments with four powerful baseline models, and comprehensively benchmark our Mulberry with various state-of-the-arts, including general and reasoning-based MLLMs.



Question: Use a sector paper sheet with a central angle of 120.0 and a radius of 6.0 to roll into a conical bottomless paper cap (as shown in the picture), then the bottom perimeter of the

Choices:
(A) 2 π cm; (B) 3 π cm; (C) 4 π cm; (D) 5 π cm

Correct Answer: (C)

Mulberry:
Image Description:
A diagram shows a circular sector being transformed into a cone without its base. The sector has a central angle labeled as **120°** and a given radius length of **6**.

Rationales:
To find the circumference or "bottom perimeter" of the resulting cone from rolling up this sector, we need to understand that when you form a cone using such a sector, the arc length becomes the circumference of the base circle of the cone.

Let's think step by step.

Step 1:
Calculate the arc length of the sector which forms the circumference of the cone's base. Arc Length = central angle/360 × 2πr. Here, r=6 cm and central angle = 120°.

Step 2:
Substitute values into the formula:
Arc Length = (120/360) × 2π × 6.
Simplify it further to get the actual arc length.

Step 3:
Perform calculations: Arc Length = 1/3 × 12π = 4π. This means the circumference of the cone's base equals **4π**.

The final answer is:
C

Qualitative Results

Method	MathVista	MMStar	MMMU	ChartQA	DynaMath	HallBench	MM-Math	MME_{sum}	AVG
Closed-Source Model									
GPT-4o [37]	63.8	63.9	69.1	85.7	63.7	55.0	31.8	2329	64.5
Claude-3.5 Sonnet [38]	67.7	62.2	68.3	90.8	64.8	55.0	-	1920	-
Open-Source Model									
MM-1.5-7B [39]	47.6	-	41.8	78.6	-	-	-	1861	-
Idefics3-LLaMA3-8B [40]	58.4	55.9	46.6	74.8	-	-	-	1937	-
InternVL2-8B [41]	58.3	**61.5**	51.8	83.3	39.7	-	-	2210	-
MiniCPM-V-2.6-8B [42]	60.6	57.5	49.8	-	-	48.1	-	2348	-
DeepSeek-VL2-MOE-4.5B [43]	62.8	61.3	51.1	86.0	-	-	-	2253	-
Reasoning Model									
LLaVA-CoT-11B [4]	54.8	57.6	-	-	-	47.8	-	-	-
LLaVA-Reasoner-8B [3]	50.6	54.0	40.0	83.0	-	-	-	-	-
Insight-V-8B [44]	49.8	57.4	42.0	77.4	-	-	-	2069	-
LLaVA-NeXT-8B [45]	37.5	42.1	41.7	69.5	22.7	33.4	0.6	1957	39.7
Mulberry-LLaVA-8B	56.3	54.5	43.0	79.5	34.1	47.5	18.9	2021	50.7 $^{11\uparrow}$
Llama-3.2-11B-V-Ins. [46]	48.6	49.8	41.7	83.4	34.3	40.3	4.1	1787	45.8
Mulberry-Llama-11B	61.1	58.5	45.6	83.5	37.2	48.9	18.7	2035	53.3 $^{7.5\uparrow}$
Qwen2-VL-2B [2]	43.0	48.0	41.1	73.5	24.9	41.7	1.0	1872	42.5
Mulberry-2B	51.7	51.3	42.0	77.7	30.0	44.9	13.9	2013	47.9 $^{5.4\uparrow}$
Qwen2-VL-7B [2]	58.2	60.7	54.1	83.0	42.1	50.6	5.9	2327	54.7
Mulberry-7B	**63.1**	61.3	**55.0**	**83.9**	**45.1**	**54.1**	**23.7**	**2396**	**58.9** $^{4.2\uparrow}$

Quantitative Results