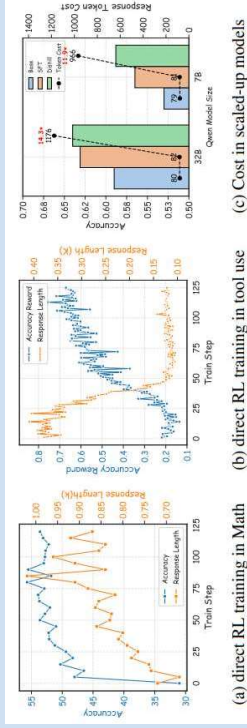




Problem & Motivation

The Challenge of Tool-Use Scaling

- Integrating Large Language Models with tools is critical for General Artificial Intelligence.
- Problem A: Reasoning Collapse.** Direct RL training often fails as reasoning trajectories collapse into short, insufficient paths for complex tasks.
- Problem B: Token Inefficiency.** Scaled-up models (e.g., via distillation) tend to overthink simple problems, wasting computational resources.

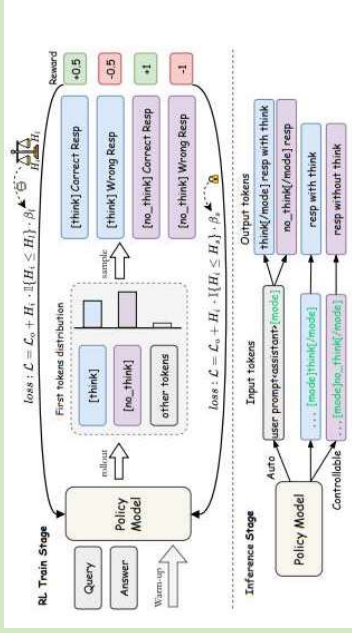


Key Pre-study Insights

- Data Distribution:** Analysis shows that while data difficulty affects convergence speed, it is not the primary cause of reasoning collapse.
- Entropy Correlation:** Low information entropy, which measures a model's exploration capability, is strongly correlated with reasoning collapse.

Methods

Decoupled Adaptive Entropy Constraints



Sample-level policy loss:

$$\beta_t = \beta_s \cdot m_t + \beta_l \cdot (1 - m_t) \cdot \mathbb{I}\{H_t \leq H_l\}$$

$$\mathcal{L}_p = \frac{1}{N} \sum_{i=1}^N \left[-\min(\rho_i \hat{A}_i, \text{clip}(\rho_i, 1 - \epsilon, 1 + \epsilon) \cdot \hat{A}_i) - \beta_t H_t \right]$$

Adaptive Entropy Coefficient Loss:

$$\mathcal{L}_\beta = \frac{1}{\sum_{j=1}^N (1 - m_j)} \sum_{i=1}^N (1 - m_i) \cdot \beta_i \cdot (H_i - H_l)$$

Two-Stage Training Framework:

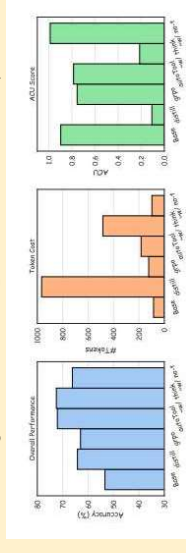
- Warm-up SFT: mixed long and short reasoning trajectories to perceive problem difficulty
- RL with Decoupled Constraints: differentiates between reasoning modes

Experimental Results

- Overall performance:** AutoTool-7B model achieving a +11.95% accuracy boost over SFT-based baselines.
- Complex Tasks:** Achieves significant gains in Multi-turn scenarios (+28.5% compared to PubTool-SFT).
- Dynamic Adaptation:** The model automatically adjusts thinking rates: **45%** for complex scenarios and **0%** for simple ones.

Type	Model	Non-Live	Live	Multi-Turn	Overall Acc
◆Bise	LlaMA-3.1-8B-Instruct	84.21	61.08	9.62	50.87
	Qwen2.5-7B-Instruct	86.46	67.44	7.62	55.69
	Qwen2.5-32B-Instruct	85.81	74.23	17.75	59.67
▼Frontier	GPT-4o-2024-11-20	87.67	79.88	43.00	70.42
	o1-2025-04-16	81.42	73.43	56.12	70.32
	Gemini-1.5-Pro	89.54	76.83	30.62	65.48
◆SFT	Hammer2.1-7B(Liu et al., 2024)	88.65	75.11	23.50	61.83
	ToolACE-8B (Liu et al., 2024)	87.54	78.59	7.75	58.42
	xAAM-7B(Zhang et al., 2024a)	81.06	75.22	10.00	54.75
	PubTool-SFT ¹	88.98	77.28	9.68	58.17
	PubTool-Distilled ¹	87.73	78.64	15.65	60.30
◆RLVR	DeepSeek-R1-ZS28	75.20	77.30	38.88	63.79
	OwO-38B(Team, 2025b)	88.81	78.54	33	66.34
	OwO-32B(Team, 2025b)	87.33	75.61	14.50	58.30
	Tool-N1-7B (Zhang et al., 2025b)	89.25	80.38	-	-
	ToolRL-7B(Qian et al., 2025)	82.21	74.90	18.12	58.38
◆Ours	Thinkless(Fang et al., 2025)	86.92	77.62	24.64	63.06
	AdaptRL(Huang et al., 2025)	86.36	73.12	15.63	58.37
	PubTool-GRPO ¹	88.87	78.93	10.77	60.13
	AutoTool-7B ¹	89.76%	80.22%	38.18%	70.12%
	+ think	89.86%	80.43%	39.28%	70.71%
+ no-think	87.36%	78.60%	27.63%	63.34	

Token Savings: Reduces computational overhead by **81%**.



Conclusion: AutoTool provides a practical auto-scaling solution that balances high-level reasoning performance with extreme inference efficiency.