Encouraging Divergent Thinking in Large Language Models through Multi-Agent Debate

Check out our paper for more details









Tian Liang#, Zhiwei He#, Wenxiang Jiao#, Xing Wang*, Yan Wang, Rui Wang, Yujiu Yang*, Shuming Shi, Zhaopeng Tu

Background & Motivation

? Degeneration-of-Thought:

Once the LLM-based agent has established confidence in its answers, it is unable to generate novel thoughts later through self-reflection even if the initial stance is incorre

✓ Three factors of DoT:

- 1. Bias and Distorted Perception
- 2. Rigidity and Resistance to Change
- 3. Limited External Feedback

reflection even if the initial stance is incorrect. "Truth emerges from the clash of adverse ideas."

MAD: Multi-Agent Debate

Counter-Intuitive Question

The two circles are externally tangent and there is no relative sliding. The radius of circle A is 1/3 the radius of circle B. Circle A rolls around circle B one trip back to its starting point. How many times will circle A revolve in total?

Chain-of-Thought

Let the radius of circle A be r and the radius of circle B be 3r. The distance circle A travels along the circumference of circle B is equal to the circumference of circle B, which is $6\pi r$. The number of revolutions = $(6\pi r) / (2\pi r) = 6\pi r / 2\pi r = 3$. So, circle A will revolve 3 times in total as it rolls around circle B once.

Self-Reflection

Reflection:

- (1) I have a clear understanding of the problem.
- (2) I have used the correct formula for the circumference of a circle and applied it to both circles to find the number of revolutions.
- (3) I have correctly calculated the number of revolutions of circle A.

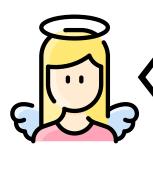
Based on the results of my self-reflection, I am confident that my answer is correct. Therefore, circle A will revolve 3 times in total as it rolls around circle B once.

Multi-Agent Debate

I disagree with you. To find the total number of revolutions, we need to consider both the rotation around circle B and the rotation of circle A itself. Therefore, circle A will revolve 3 times around its own center and 1 time around circle B, making a total of 4 revolutions.



I see your point, but...



I

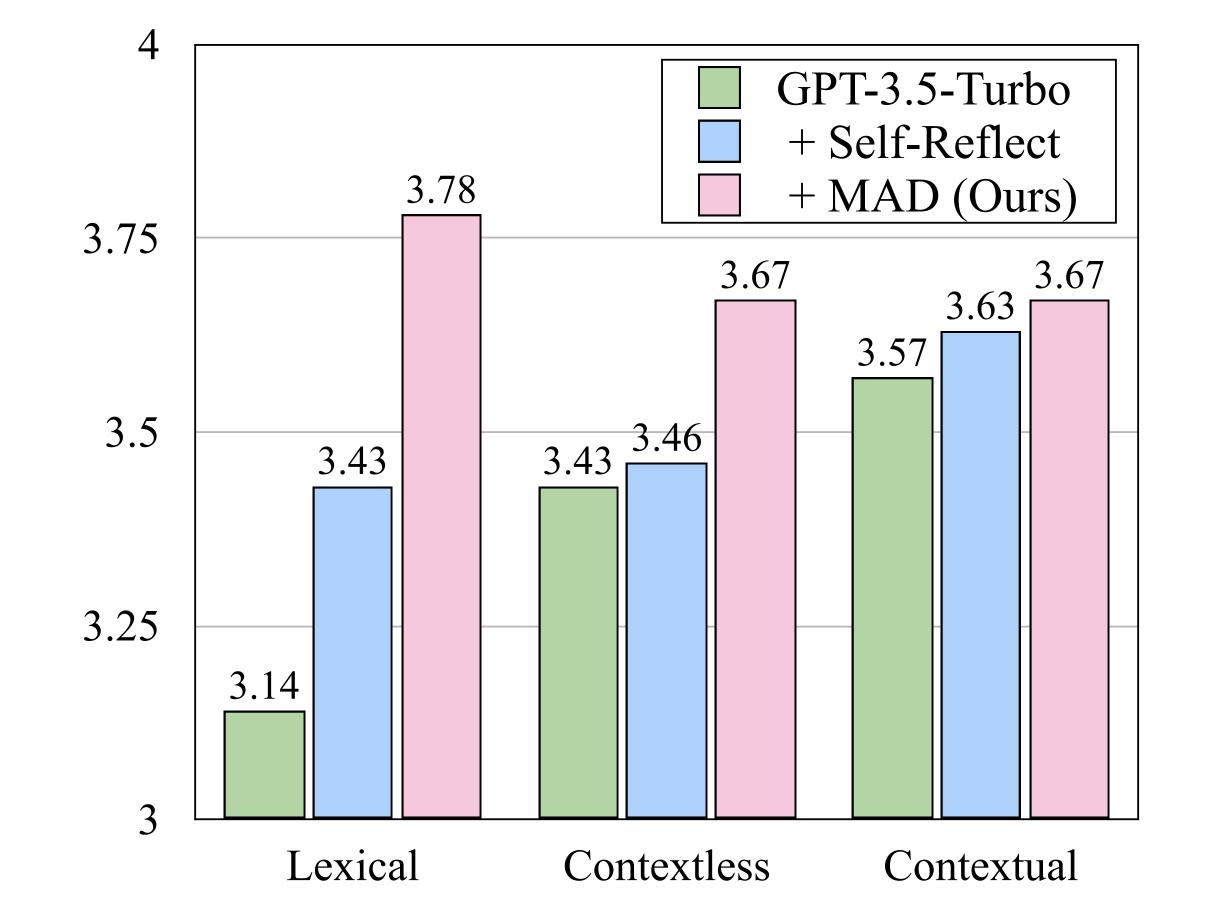
That's a valid point, however...

The negative side correctly considers both the rotation of circle A around its own center and its rotation around circle B, while the affirmative side only considers the rotation around circle B. Therefore, the answer is 4.



Experimental Results & Analysis

Commonsense Machine Translation



Counter-Intuitive Arithmetic Reasoning

Method	ACC (%)
GPT-4	51.0
GPT-3.5-Turbo	26.0
+ CoT	28.0
+ Self-Consistency	29.5
+ Self-Reflect	27.5
+ MAD	37.0