SID: Multi-LLM Debate Driven by Self Signals

Yuan Sui, Bryan Hooi

26 Mar 2026 (modified: 26 Mar 2026) | OpenReview Conference | Keywords: | Decisions: Withdrawn | 22 KB ×42

Keywords: Multi-agent debate, LLM reasoning

Abstract: Large language Models (LLMs) have demonstrated impressive capabilities across diverse application domains. Recent work has explored Multi-LLM Agent Debate (MAD) as a way to enhance performance by leveraging diversity among models. These studies have shown that heterogeneous model collaboration can outperform single-model inference, suggesting that different LLMs contribute complementary reasoning capabilities. This raises the question of how such collaboration might be further improved. We observe that most MAD frameworks rely on uniform interaction patterns regardless of model identity or task difficulty, which limits the potential benefits of collaboration. In this paper, we propose Self-Identified Debate (SID), a multi-LLM debate framework that introduces lightweight inter-model communication. Models are prompted to jointly predict answer correctness and confidence alongside their reasoning, enabl ing selective attention to more reliable peers. We also propose uncertainty-aware evaluation, which weighs agents by their self-reported confidence rather than treating all agents equally. Across 6 benchmarks spanning arithmetic, knowledge, and reasoning tasks, we show that our method consistently outperforms vanilla MAD and ablations that remove the communication or uncertainty-weighting components, indicating that both components contribute to the observed gains. We also find that the benefit of debate varies across model compositions and task types, with the largest improvements on challenging tasks where at least one model in the pair is weak. We release our code at https://github.com/sandi-gt/sid_debate.

Primary Area: Foundations or Frontiers in Models, including LLMs

Submission Number: 22 (Track: Main Conference Track) | Submitted: 26 Mar 2026 | Last Modified: 26 Mar 2026

Latex Template used: Yes | Submission Type: Regular Paper | Submission Length: Regular paper (7 - 9 pages) | Anonymization: This submission does not use double-blind review.

Acknowledgments: All Yes

Reciprocal Reviewing: I understand that if my submission is accepted, I will be a reviewer for the Main Track of the Conference.

Supplementary Material: Code, Data, Videos, Demos, Extended Results

Code Of Conduct: Yes

Submission Number: 22

Filter by title, author, ... | Sort by activity | Search keywords... | Apply Selected Filter | ⊘ | ⊞ | ☰ | ⊟ | ⊞ | ⊞ | ⊞