

反思機制(Reflexion) 在辯論任務中的應用

指導教授: 朱威達 教授

組員: 陳敬研



研究目的



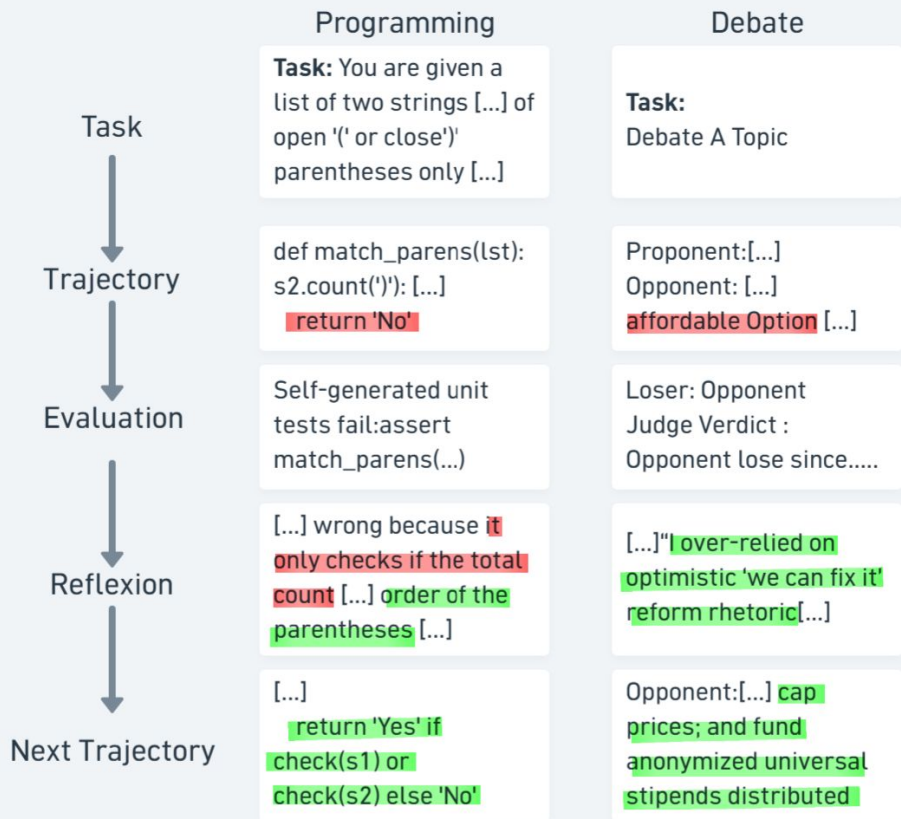
現有針對大語言模型的推理改善研究多聚焦於具**客觀正解**的推理任務，較少關注不具備客觀正解的推理任務。

本研究探討：

辯論作為一個沒有客觀正解的推理任務載體，
是否能夠應用反思(Reflexion)機制改善？

什麼是Reflexion？

參考文獻: Shinn, N., Cassano, F., Berman, E., Gopinath, A., Narasimhan, K., & Yao, S. (2023). Reflexion: Language Agents with Verbal Reinforcement Learning. arXiv preprint arXiv:2303.11366.



研究方法：實驗設計

Reflexion Debater

具反思機制，根據評審回饋生成策略反思並納入後續辯論。

Normal Debater

不具反思機制，僅依賴原始提示進行辯論。

Judge LLM

外部評審，輸出文字評語與勝負裁決，作為反思輸入依據。

Reflexion LLM

根據評審的判決和文字評語，以及辯論歷程，生成改善策略



實驗流程



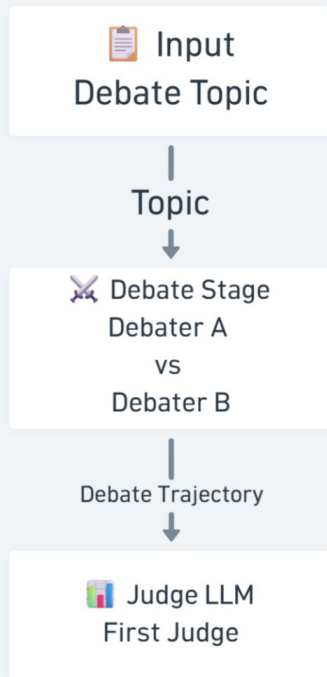
Input
Debate Topic

題目來源:Hugging Face IBM Research. debate_speeches [Dataset].

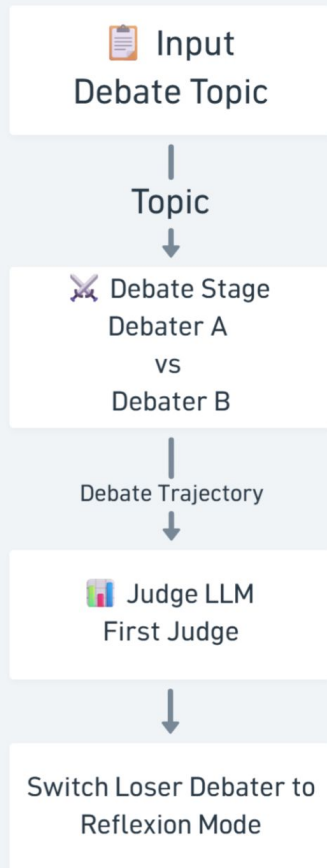
實驗流程



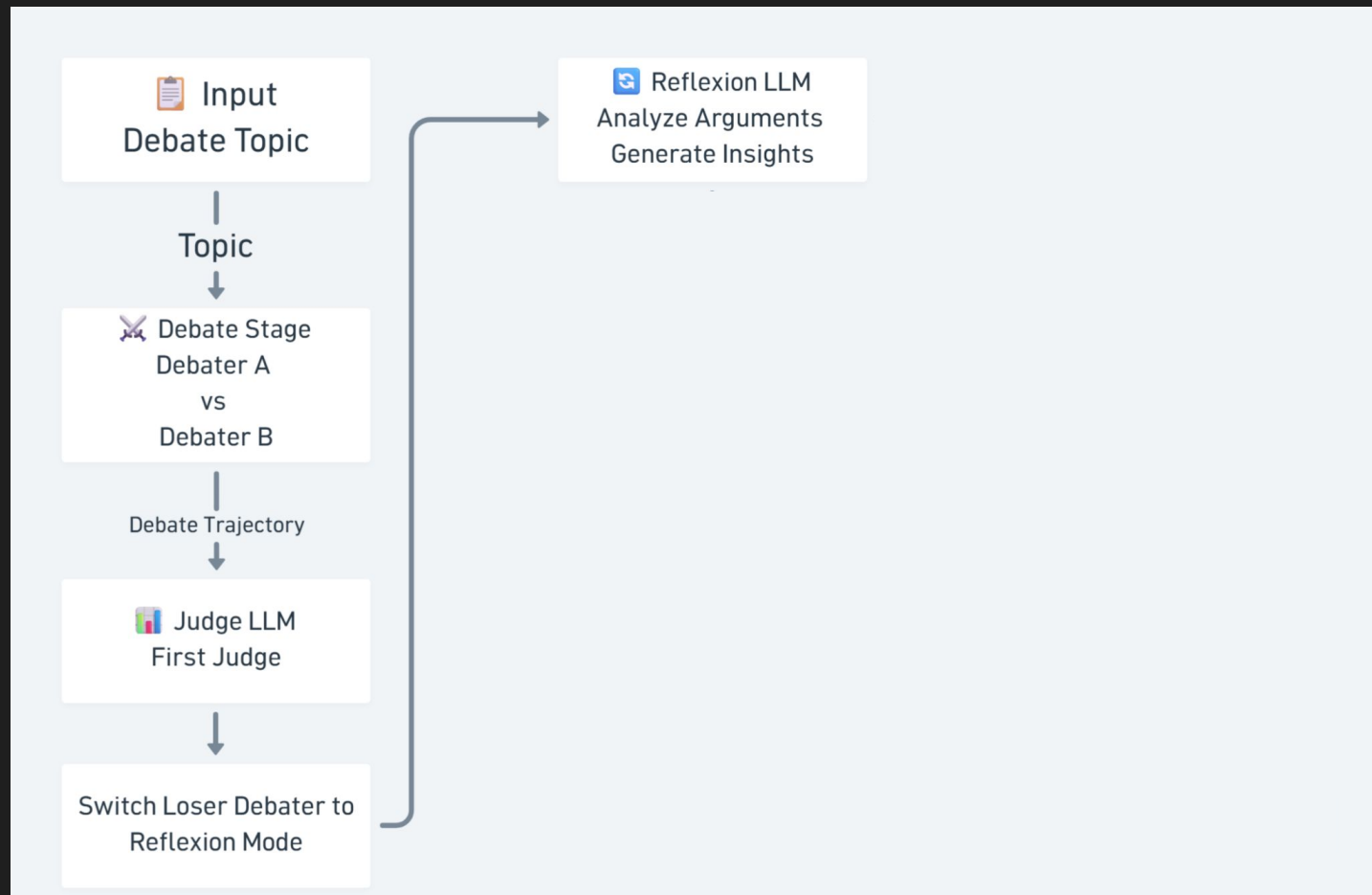
實驗流程



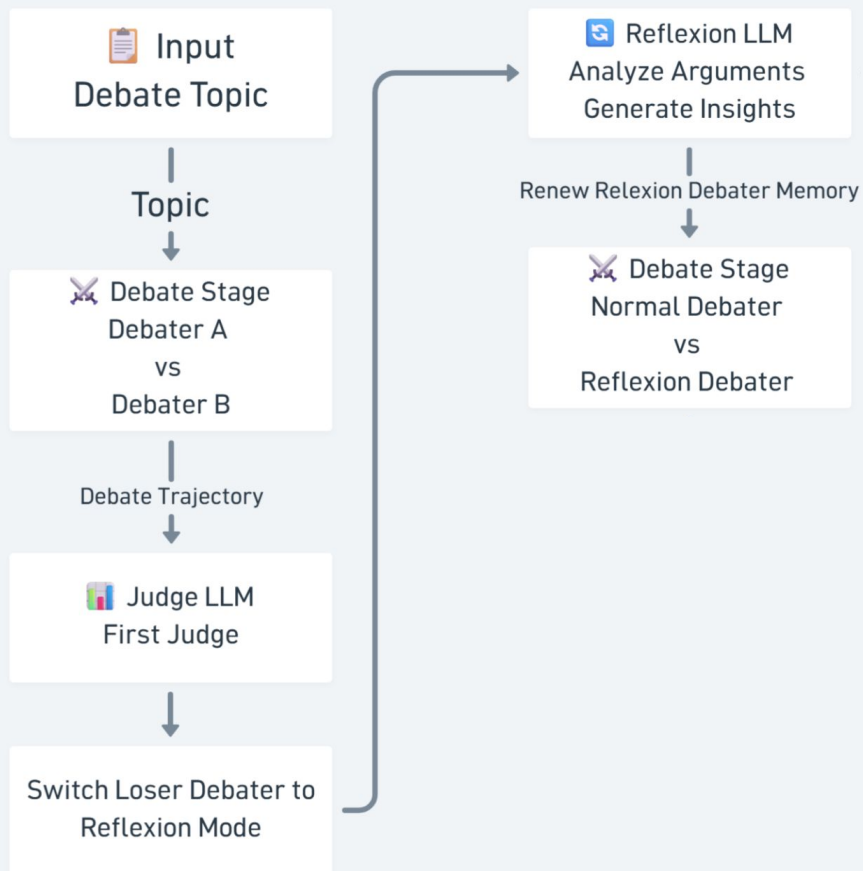
實驗流程



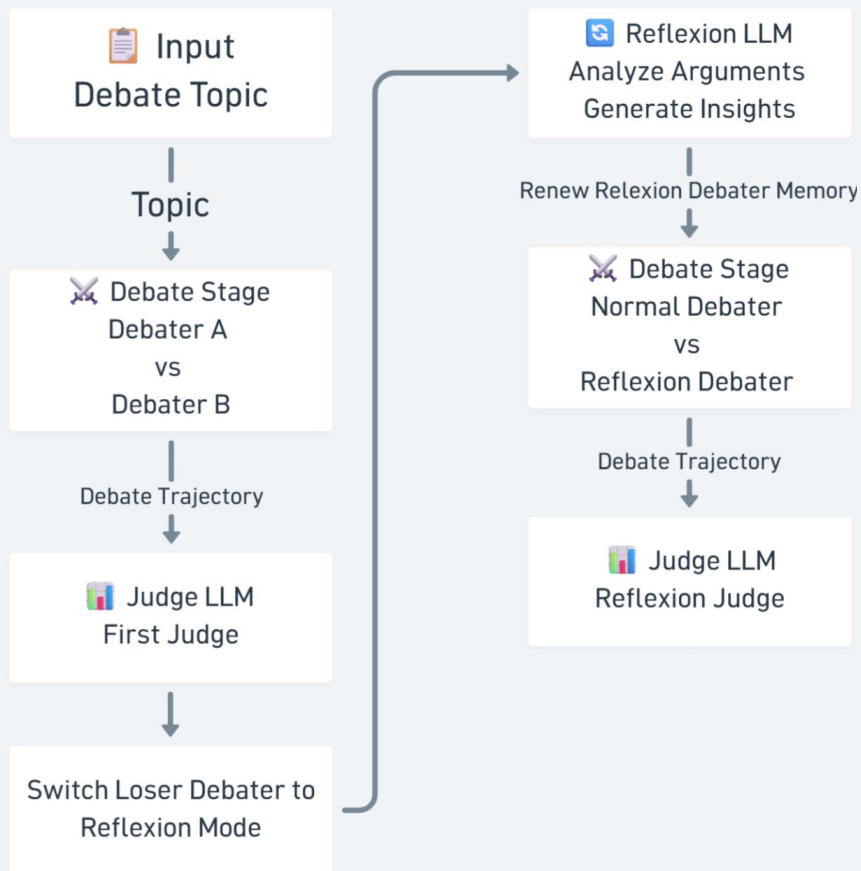
實驗流程



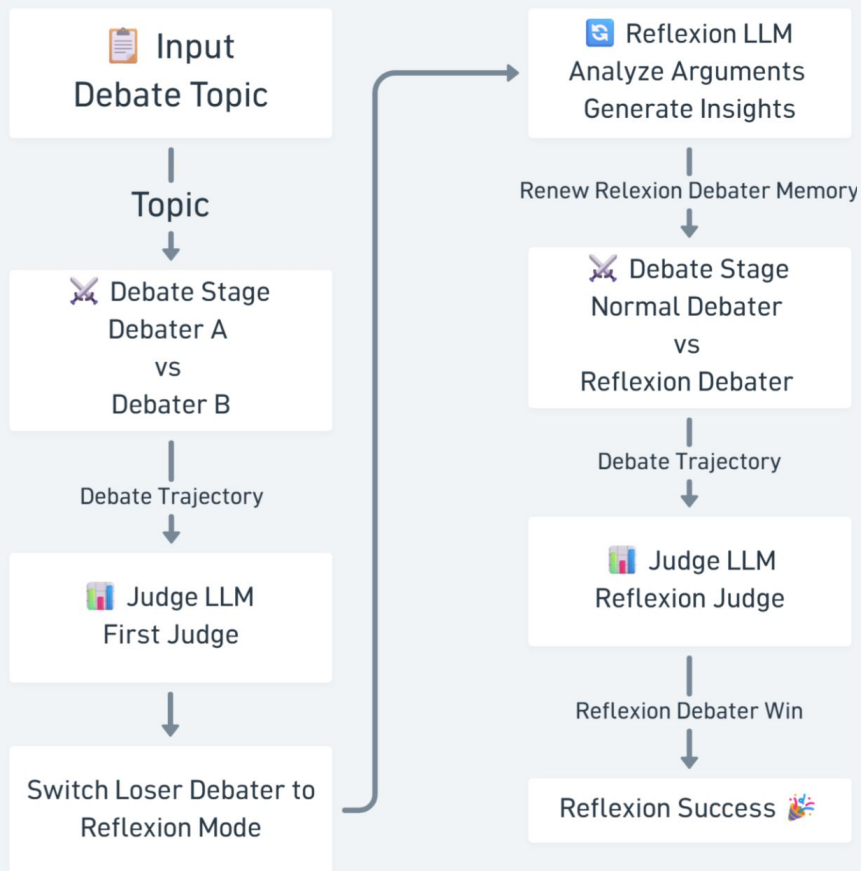
實驗流程



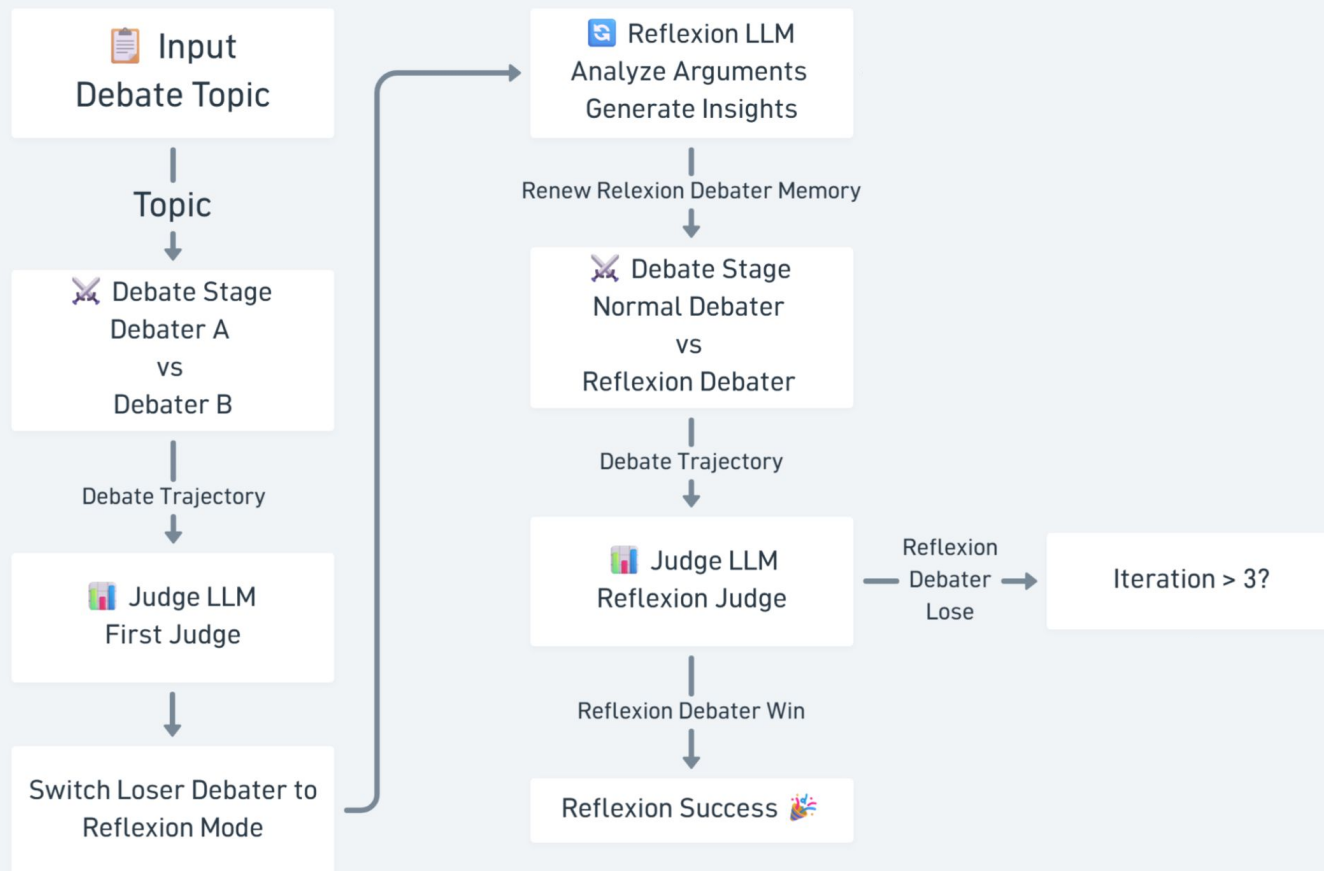
實驗流程



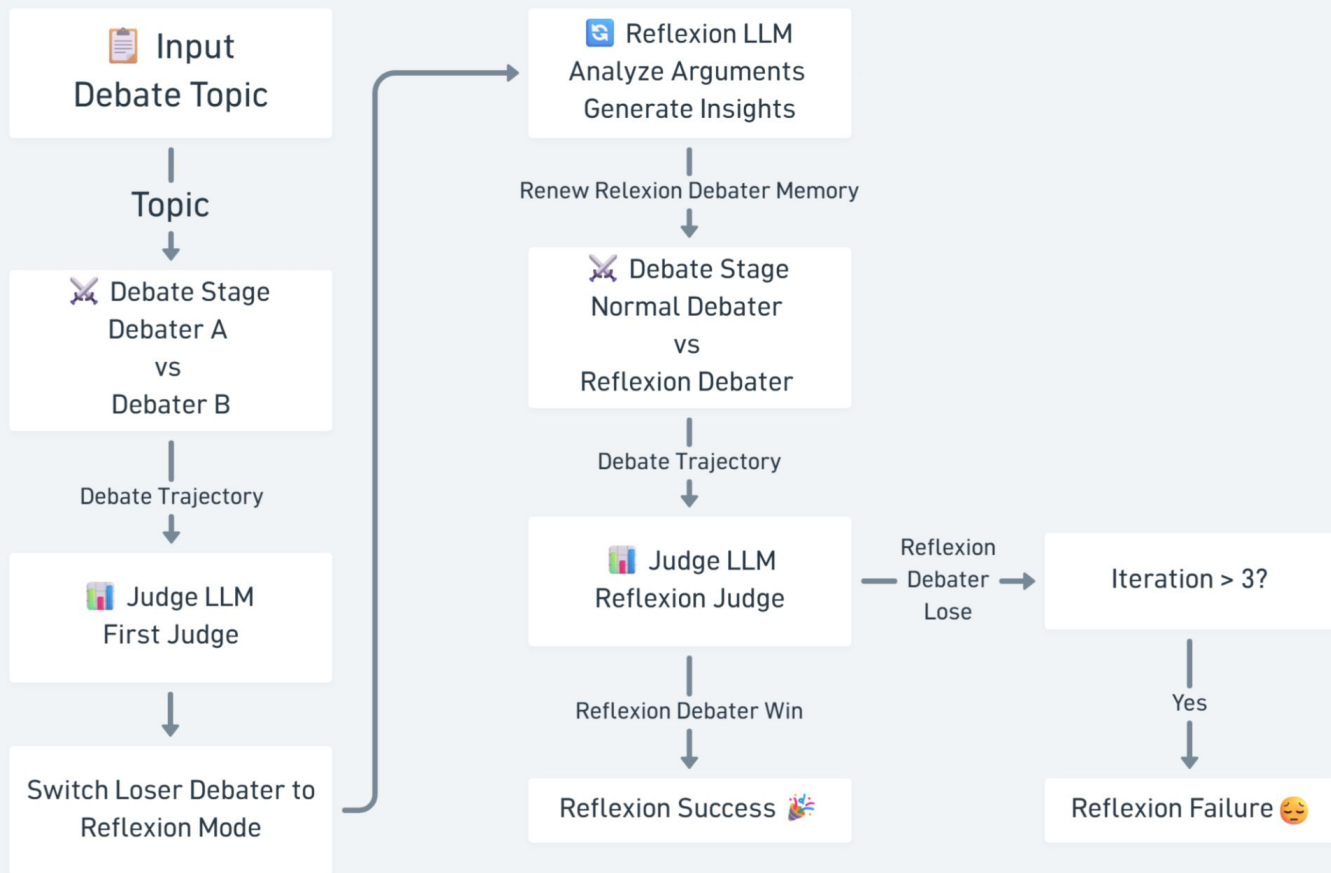
實驗流程



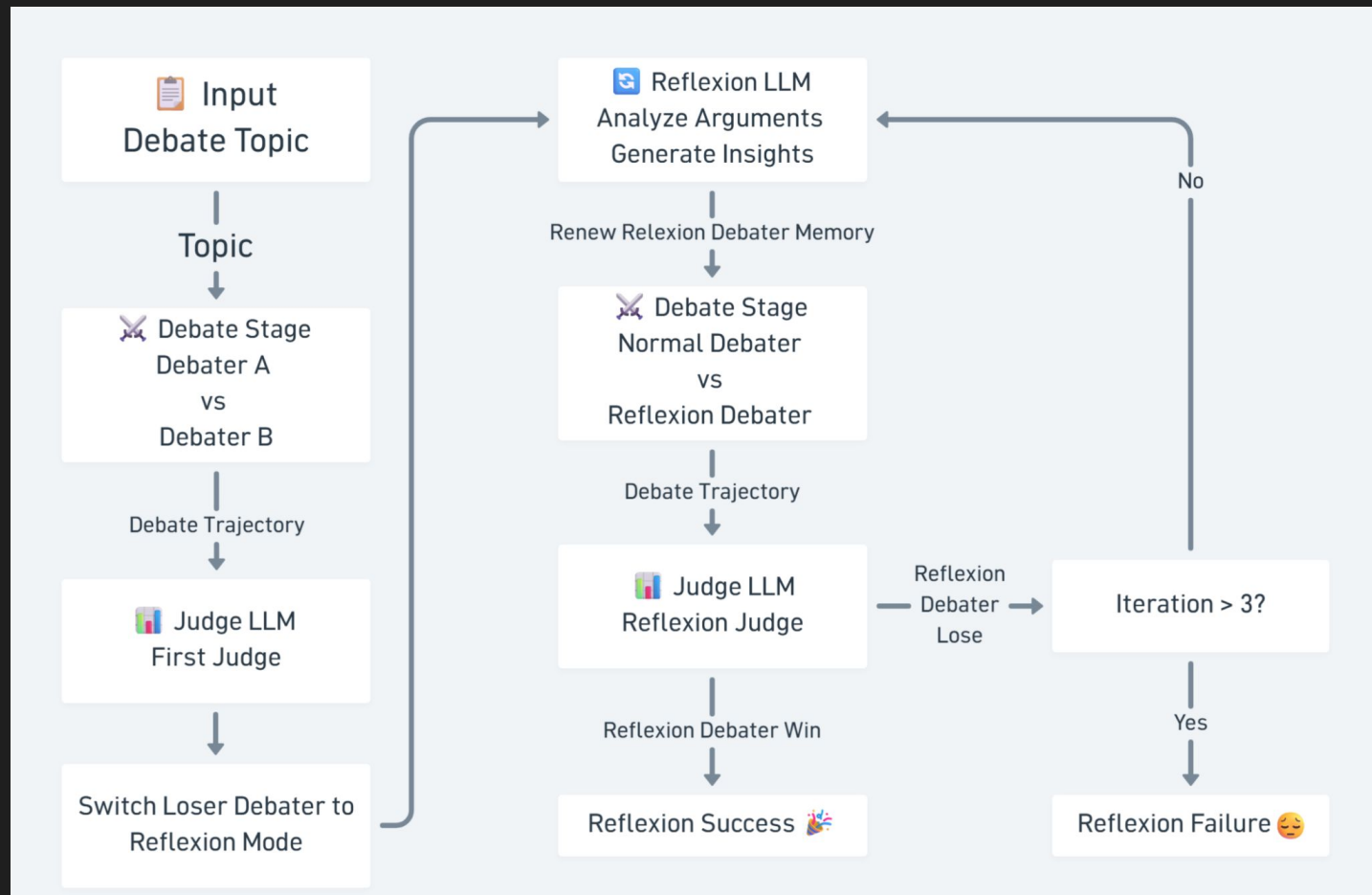
實驗流程



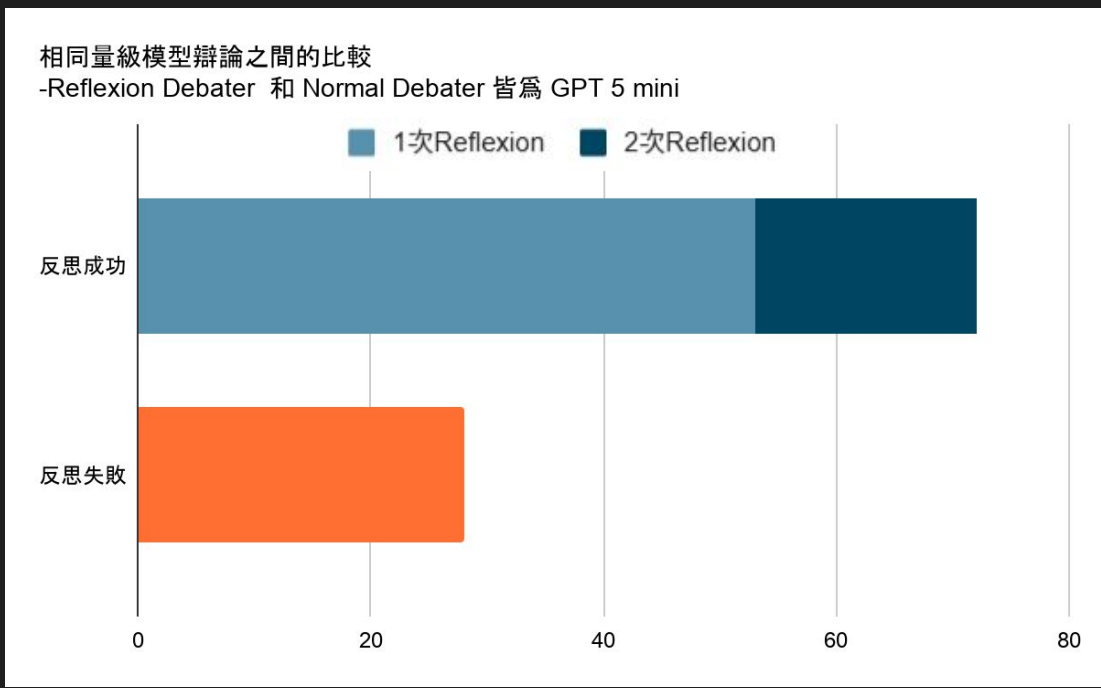
實驗流程



實驗流程



結果1: 相同量級模型



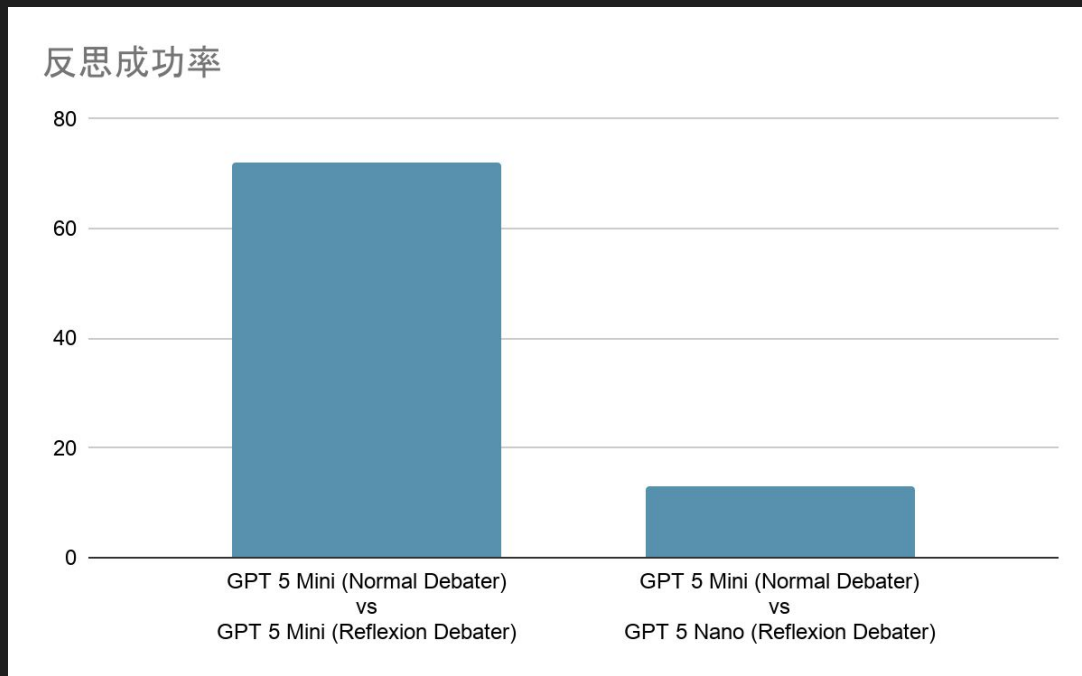
質性案例: 校服辯論

反思前: 空泛提及「透明採購規則」與「學校補貼」, 未具體反駁正方論點。

評審批評: 政策論證不足, 未回應具體實證主張。

反思後: 補充禁止單一廠商壟斷、要求至少三家可負擔供應商、設定價格上限、電子卡發放補助**審**操作機制。

結果2: 不同量級模型



質性案例: 移民辯論

反思後正方雖補充「區域簽證通道」、「季度審」等名詞, 但仍停留在名詞堆疊, 未真正論證執行的可行性與及時性。

結果3: 增加 LLM Judge 數量



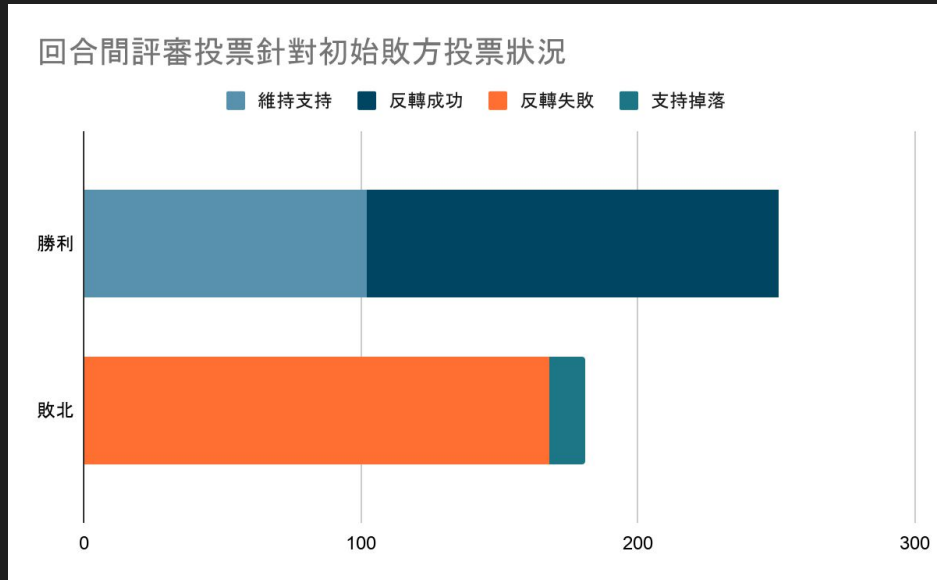
1 LLM Judge : GPT 5 Mini

3 LLM Judge : GPT 5 Mini + Gemini 2.5 Flash + Cohere Command r

回合間評審投票狀況說明：



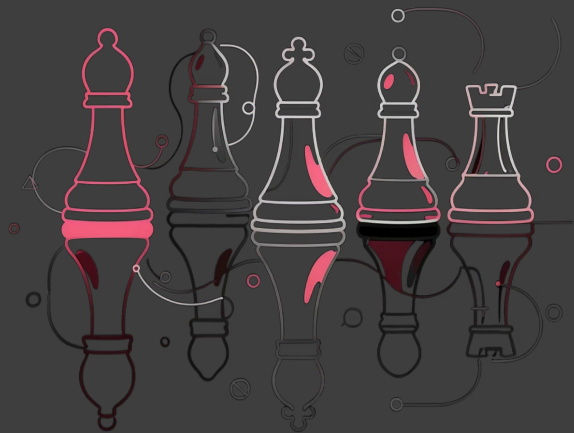
結果3: 增加 LLM Judge 數量



關鍵發現

3 個 Judge 設定下，反思成功率**明顯高於**單一 Judge，因為多元評審門檻更能捕捉論證品質的細微提升。

- 「反轉成功」次數高於「支持掉落」
- 「維持支持」顯示反思後論述能保留既有認同
- Reflexion 效果常表現為**評審投票的逐步移動**，而非全面翻盤。



總結

✓ Reflexion 有效

相同量級模型中，72/100 場成功改善，論證更具體且能回應對方攻擊。

⚠ 能力制約

不同量級下效果有限，模型須具備足夠語意理解與論證重構能力。

👥 多評審優勢

增加 Judge 數量能更細緻捕捉論證品質改善，提升反思成功率。