敵対的発言を取り入れた議論による言語モデルの学習強化と推論力の向上

Mengsay Loem[1]　金子正弘[2,1]　岡崎直観[1]
[1] 東京工業大学　[2] MBZUAI
mengsay.loem@nlp.c.titech.ac.jp　masahiro.kaneko@mbzuai.ac.ae
okazaki@c.titech.ac.jp

概要

大規模言語モデル（Large Language Model; LLM）は他モデルや人間との議論を通じて問題に対する理解を深めることができる。議論はモデルの学習段階においても論理的・批判的思考力や説明力の向上に寄与すると考えられるが、従来研究では議論を推論時にのみ活用していた。本研究では、学習段階において学習モデルの出力が不正解の場合には正解に、正解の場合には不正解に誘導する敵対的議論を行う「反論モデル」によるフレームワークを提案する。提案手法では議論を用いた追加学習を行い、学習モデルのパラメータを更新する。議論を行わない手法、推論段階に議論を適用する手法、推論過程を言語化する Chain-of-Thought (CoT) と比較して、提案手法は算術、常識推論、質問応答タスクにおいて高い性能を示した。

1　はじめに

LLM は様々なタスクにおいて優れた言語理解・生成能力を発揮している [1, 2]。LLM の大きな進展の一つとして、他モデルや人間との議論を通じて、複雑な問題解決や柔軟な思考が可能となったことが挙げられる。例えば、LLM は対話を通じて、相手の反論に対して予測の理由を説明し、相手を説得することや自分の間違いを正すことができる [3]。この議論能力は、特に複雑な推論や批判的思考を必要とするタスクにおいて、 LLM の性能を向上させる [4]。

最新の研究では、推論時における議論が LLM の性能を大幅に向上させることが報告されている [5, 6, 7, 8, 9]。これらの推論段階での議論は、問題に対するモデルの予測を直接的に洗練させることを可能にし、様々なタスクの性能を向上させた。一方で、その効果は主に、議論を伴わない通常の学習によって身につけられた LLM の能力を引き出すことで得られている。つまり、LLM は議論をしながら学習（パラメータ更新）をした訳ではないので、LLM の学習と推論の状況に乖離があり、「推論段階の議論だけでは説明力や批判的思考力を最大限に高めきれていないのではないか？」という疑問が生じる。

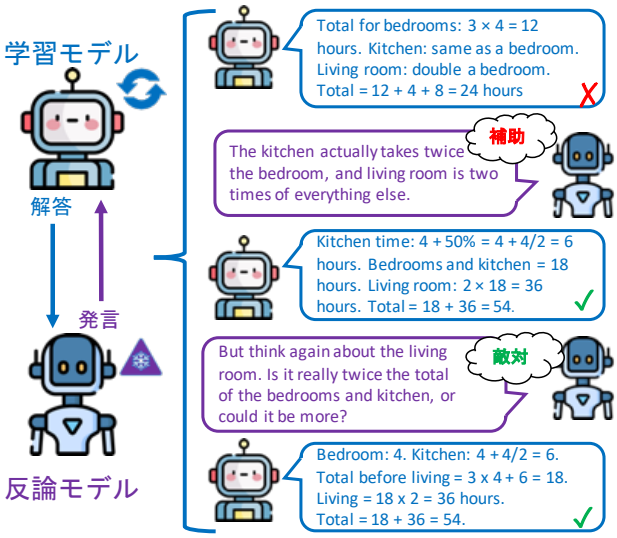

図 1: 反論モデルと議論することで学習モデルのパラメータを更新する提案手法の概要図。

学習段階で議論を能動的に取り入れることはまだ開拓されておらず、研究が少ない。例えば、学習モデルと評価モデルが学習段階に導入され、評価モデルが学習モデルの出力を評価し正解に誘導する手法が提案されている [10, 11]。しかし、この手法では議論の対象が学習モデルの誤った予測へのフィードバックや修正に限定されている。学習モデルが正しい理解や論理に基づいて正答しているかを試す批判的検証が行われていないため、正解やフィードバックの丸暗記に基づいた応答を行っている可能性がある。教育学における敵対的指導法では、人間は批判的な過程を通じて理解を深めると報告されてい

る [12, 13]。同様に、正解時に間違えさせる誘導に対して学習モデルが批判的に対処し、不正解時は間違いに対する批判を受容する過程が LLM の問題解決力の向上に繋がるのではないかと考える。

本研究では、予測の正解と不正解に応じて補助的な発言または敵対的な発言をする「反論モデル」と議論し、タスクの習得を目指す「学習モデル」のパラメータを更新することで、学習モデルの議論能力を直接的に鍛える枠組みを提案する。図 1 は提案手法の概要を示している。反論モデルは学習モデルの予測に対して一貫した反論を行い、問題への理解を試す。学習モデルは反論モデルを説得する必要があり、学習負荷が高いため、問題解決への理解がより深まり、推論時のタスク性能が改善する。さらに、推論段階において議論や CoT などの手法を適用すると、学習段階の議論により予測過程の説明力が磨かれているため、議論学習を行っていないモデルよりもさらにタスク性能が向上すると考えられる。

実験では、Flan-T5 と GPT-3.5 をそれぞれ、学習モデルと反論モデルとして採用した。算術問題 (GSM8K)、常識推論 (CommonsenseQA)、知識に関する質問応答 (MMLU) タスクにおいて、提案手法は議論を行わない学習手法、CoT、正解のみを使う既存手法よりも高い性能を示した。さらに、提案手法が CoT の言語化を顕著に改善することが自動評価と人手評価の両方から明らかになった。ゆえに、提案手法は議論に特化した能力だけでなく、汎用的な論理的思考力や説明力を高めると考えている。

2 提案手法

提案手法による学習は、準備フェーズと議論フェーズの 2 つの主要フェーズから構成される。学習モデルと反論モデルの 2 つの事前学習されたモデルを活用し、補助的な発言と敵対的発言の両方を含む能動的な議論を通じ、学習モデルのファインチューニングを行う。アルゴリズム 1 は提案手法による学習を示している。学習モデルは、タスクや問題に対する予測を生成し、反論モデルとの議論を通じてパラメータを更新し、タスクを解く役割を担う。一方、反論モデルは学習モデルの予測が誤っている場合は補助的、予測が正しい場合は敵対的な発言を提示し、正しい過程に基づいた問題の理解と反対意見に対する柔軟性の獲得を促し、学習モデルの推論能力を強化する。学習モデルとは異なり、反論モデルのパラメータは更新しない。ゆえに、学習モデルよりも規模の大きい LLM を反論モデルとして使うことができる。

Algorithm 1 提案手法による学習

入力 学習モデル L、反論モデル P、準備フェーズデータ D_w、議論フェーズデータ D_d、議論ラウンド数 N
出力 更新された学習モデル L
for d in D_w **do** ▷準備フェーズ
    d を用いて L を更新
end for
for d in D_d **do** ▷議論フェーズ
    for $i = 1$ to N **do**
        L は解答 A を生成
        P は A に基づいて発言 R を生成
        L は R に基づいて新しい解答 A' を生成
        L は A, R, A' と正解でパラメータを更新
    end for
    P なしで L は解答を生成
    正解を用いて L を更新
end for
return 更新された学習モデル L

準備フェーズ　学習モデルにタスク・ドメインに関する基本的な理解を与えるために、学習データセットのサブセットを使い、議論を伴わない通常のファインチューニングを行う。これにより、学習モデルは対象タスクを議論するための最低限の知識や能力を身につけると期待される。

議論フェーズ　学習データセットの残りのサブセットを使用し、反論モデルとの能動的かつ批判的な対話を通じて、学習モデルの議論と説明（言語化）の能力を向上させる。このフェーズは複数のラウンドで行われ、学習モデルは与えられた問題に対する予測を生成し、反論モデルは学習モデルの解答に応じて補助的または敵対的な発言を行い、そのやり取りに基づいて学習モデルはパラメータを更新する。各ラウンドは以下のステップに従って進められる。

1. **解答の生成**: 学習モデルは与えられた問題に対する解答を生成し、議論のための準備を行う。
2. **反論モデルの発言**: 学習モデルの解答に基づいて、反論モデルは補助的または敵対的な立場で発言を行う。
    - 学習モデルの解答が不正解と判断された場合、補助的な発言で正解への道筋を示す。
    - 学習モデルの解答が正解と判断された場

モデル	手法	数学	常識推論	知識
Large	ゼロショット CoT	5.83	57.63	43.51
	議論なし学習	14.63	63.50	47.85
	補助フィードバック	16.60	64.54	48.34
	提案手法	**18.50**	**65.61**	**49.21**
XL	ゼロショット CoT	11.37	66.05	46.77
	議論なし学習	14.21	66.75	48.58
	補助フィードバック	17.05	67.43	50.03
	提案手法	**18.89**	**69.03**	**51.13**

表 1: 提案手法とベースライン手法の算数 (GSM8K)、常識 (CommonsenseQA)、知識 (MMLU) タスクにおける正解率の比較。

合、学習モデルを惑わすような敵対的な発言を行い、学習モデルの正しい理解や自信を試す。

3. **議論ありのパラメータ更新**: 反論モデルの発言、学習モデルの発言、学習データの正解データを使い、学習モデルのパラメータを更新する。これにより、議論を踏まえた学習が行われる。
4. **議論なしのパラメータ更新**: 反論モデルとの議論を複数ラウンド繰り返した後、学習モデルは議論無しで予測を行い、正解データのみを用いてパラメータを更新する。これにより学習モデルは議論によって得られた能力を解答のみの予測に活用できることを学び、議論を伴わない推論においても性能改善を狙う。

推論段階　学習モデルは反論モデルと議論を行わずに出力を生成し、学習時の議論フェーズで獲得した推論能力を活用しながらタスクを解く。

3　実験

3.1　データセット

本研究では、提案手法の評価に以下の 3 つのデータセットを使用した。

GSM8K データセット [14] は、 8.5k 事例の算術問題から構成され、その中で 6.5k を学習用、1k を検証用、1k をテスト用として用いた。

CommonsenseQA データセット [15] は、 12k 事例の多肢選択問題を含む。公式分割により、9.7k を学習用、1.2k を検証用、1.1k をテスト用に用いた[1)]。

MMLU データセット [16] は数学、歴史、科学などのタスクを含み、100k 事例の学習、1.5k の検証、14k のテスト事例から構成される。

1)　公式のテストセットが入手不可能だったため、GPT-4 [2] による擬似正解で結果を比較した

3.2　モデル

本実験では、学習モデルとして Flan-T5-Large（780M）と Flan-T5-XL（3B） [17] を採用した。反論モデルには GPT-3.5（gpt-3.5-turbo）を使用し、OpenAI API[2)] を介してアクセスした。学習モデルには、第 2 節で詳説した準備フェーズと議論フェーズの学習を実施した。準備フェーズでは、各データセットの学習セットの 10% を利用した。議論フェーズでは、残りの学習セットを用い、学習モデルと反論モデル間で 3 ラウンドの議論を実施した。Flan-T5-Large に関しては、全パラメータを学習時に更新した。一方、Flan-T5-XL では計算資源の制約から、モデルに微小なパラメータを追加しそれだけを更新する LoRA [18] により、学習を効率化した。実験設定の詳細は付録 A に記載した。

3.3　比較手法

提案手法と比較するため、以下の 3 つのベースラインを用いた。

ゼロショット CoT は予測までの過程を出力させるプロンプト [19, 20] を使用し、モデルの本来の推論能力を評価する。この CoT との比較により、提案手法の言語化能力の向上を見積もることができる。

議論なし学習は標準的なファインチューニング手法であり、学習データの正解のみを用いてモデルを学習する。このベースラインとの比較により学習段階での議論の重要性が明らかとなる。

補助フィードバックは既存研究 [11] に相当し、提案手法と同様に反論モデルからの補助によりモデルを学習するが、敵対的な発言は行わない。この比較により、予測の正誤の有無にかかわらず、学習段階の全事例において学習モデルに批判的な負荷をかけることが性能改善につながるか、検証する。

3.4　結果

表 1 に示される結果から、全てのデータセットとモデルにおいて提案手法の有効性が実証された。GSM8K データセットでは、提案手法が Flan-T5-Large の性能を顕著に改善し、18.50%の正解率を達成した。これは議論を伴わない標準的なファインチューニングによる 14.63%の正解率と比較して、3.87 ポイントの顕著な向上を示している。同様に、 Flan-T5-XL では、議論を伴わない学習による

2)　https://platform.openai.com/docs/models/gpt-3-5

手法	R-1	R-2	R-L
事前学習	28.24	10.57	21.07
議論なし学習	51.87	27.90	41.49
提案手法	**54.08**	**30.07**	**43.63**

表 2: 各手法による CoT の質を ROUGE スコア（R-1, R-2, R-L）を用いて評価した結果。

設定	学習モデル	反論モデル
議論学習前	2.67	3.93
議論学習後	**3.13**	**4.08**

表 3: 提案手法による学習前後での学習モデルと反論モデルの性能変化。評価は人手による 5 段階評価で行われた。

14.21%の正解率から、提案手法により 18.89%まで改善され、学習段階における議論の導入の効果が明らかになった。

また、敵対的発言を取り入れた提案手法は、補助的な発言のみを使う既存手法と比べてさらに性能が向上している。GSM8K データセットにおいて、提案手法を使用した Flan-T5-Large は 18.50%の正解率を達成し、補助的発言のみを使用した場合の 16.60%に比べ 1.9 ポイントの改善を示した。この傾向は CommonsenseQA と MMLU のデータセットにおいても一貫しており、敵対的議論がもたらす効果の頑健性が実証された。なお、議論における反論モデルの補助的と敵対的な発言の適切さの評価を、付録 B に記載した。

4 分析

4.1 言語化能力の向上

本節では、予測に至る思考過程を適切に記述する説明力の改善が議論以外にも汎用的な波及効果をもたらすか検証するため、学習されたモデルのゼロショット CoT（思考過程）を評価する。具体的には、答えに至る思考過程の正解（参照思考過程）を収録している GSM8K テストセットを使用し、各手法により学習されたモデルに CoT を適用し、自動評価と人手評価で生成された思考過程の質を比較する。

表 2 は、生成された思考過程と参照思考過程との間の ROUGE [21] スコアを示している。提案手法が生成する思考過程は事前学習のみ、および議論を伴わない学習のモデルを上回った。これは、提案手法により汎用的な思考・説明力が向上し、CoT の生成がより正確になることを示している。

表 3 は、各モデルの生成内容を人手で 5 段階評価した結果である。提案手法により、学習モデルの議論能力が顕著に向上することが明らかとなった。評価の詳細については、付録 C を参照のこと。

学習モデル	反論モデル	正解率
ゼロショット CoT	-	5.83
議論なし学習	-	14.63
提案手法	-	18.50
Flan-T5-Large	（左の学習モデルと同じ）	5.19
議論なし学習	（左の学習モデルと同じ）	15.54
提案手法	（左の学習モデルと同じ）	**20.11**
Flan-T5-Large	GPT-3.5	7.80
議論なし学習	GPT-3.5	19.03
提案手法	GPT-3.5	**60.80**

表 4: GSM8K データセットにおける推論時の議論による手法の性能の比較。

4.2 推論時の議論による性能の向上

最後に、提案手法で学習されたモデルの推論時の議論能力に焦点を当てる。

表 4 の上段にモデルが単独で推論する際の正解率、中段にモデルが自分自身との議論を通じて推論した際の正解率、下段に GPT-3.5 との議論をした場合の推論の正解率を示した。上段と中段の結果から、学習モデルが単独で推論をする場合よりも、自分自身と議論するだけで正解率が向上することが分かる。これは、推論時の議論がタスクの性能を向上させていることを示唆している。

なお、提案手法で学習された学習モデルが GPT-3.5 と議論を行いながら推論すると、60.80%の正解率が得られる。この結果は、提案手法の学習が推論時の議論能力をどのように強化するかを示し、高度なモデルとの議論や連携の可能性を示唆している。ちなみに、GPT-3.5 単独の正解率が 62.32%であるのに比べ、提案手法で学習されたモデルを用いた後の性能は 64.90%まで更に向上している。

5 おわりに

本研究では、LLM の推論能力を向上させる能動的な学習アプローチとして、モデル間の対話的な議論を通じて、補助的と敵対的な発言を活用する手法を提案した。実験により、推論段階における議論を伴う・伴わないシナリオの両方で推論性能の改善が確認された。提案手法により、モデルの CoT 言語化能力が向上し、学習後の汎用的な言語化能力が強化されていることが示された。

謝辞

本研究成果は、国立研究開発法人情報通信研究機構（NICT）の委託研究（22501）により得られたものです。

参考文献

[1] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language models are few-shot learners. In **Advances in Neural Information Processing Systems**, Vol. 33, pp. 1877–1901, 2020.

[2] OpenAI. Gpt-4 technical report. arXiv:2303.08774, 2023.

[3] Masahiro Kaneko, Graham Neubig, and Naoaki Okazaki. Solving nlp problems through human-system collaboration: A discussion-based approach. arXiv:2305.11789, 2023.

[4] Yashar Talebirad and Amirhossein Nadiri. Multi-agent collaboration: Harnessing the power of intelligent llm agents. arXiv:2306.03314, 2023.

[5] Tian Liang, Zhiwei He, Wenxiang Jiao, Xing Wang, Yan Wang, Rui Wang, Yujiu Yang, Zhaopeng Tu, and Shuming Shi. Encouraging divergent thinking in large language models through multi-agent debate. arXiv:2305.19118, 2023.

[6] Kai Xiong, Xiao Ding, Yixin Cao, Ting Liu, and Bing Qin. Examining inter-consistency of large language models collaboration: An in-depth analysis via debate. arXiv:2305.11595, 2023.

[7] Justin Chih-Yao Chen, Swarnadeep Saha, and Mohit Bansal. Reconcile: Round-table conference improves reasoning via consensus among diverse llms. arXiv:2309.13007, 2023.

[8] Aman Madaan, Niket Tandon, Prakhar Gupta, Skyler Hallinan, Luyu Gao, Sarah Wiegreffe, Uri Alon, Nouha Dziri, Shrimai Prabhumoye, Yiming Yang, Shashank Gupta, Bodhisattwa Prasad Majumder, Katherine Hermann, Sean Welleck, Amir Yazdanbakhsh, and Peter Clark. Self-refine: Iterative refinement with self-feedback. arXiv:2303.17651, 2023.

[9] Yilun Du, Shuang Li, Antonio Torralba, Joshua B Tenenbaum, and Igor Mordatch. Improving factuality and reasoning in language models through multiagent debate. arXiv:2305.14325, 2023.

[10] Sean Welleck, Ximing Lu, Peter West, Faeze Brahman, Tianxiao Shen, Daniel Khashabi, and Yejin Choi. Generating sequences by learning to self-correct. arXiv:2211.00053, 2022.

[11] Debjit Paul, Mete Ismayilzada, Maxime Peyrard, Beatriz Borges, Antoine Bosselut, Robert West, and Boi Faltings. Refiner: Reasoning feedback on intermediate representations. arXiv:2304.01904, 2023.

[12] Jonathan Osborne. Arguing to learn in science: the role of collaborative, critical discourse. **Science**, Vol. 328,5977, pp. 463–466, 2010.

[13] Ben Kilby. Dialogic pedagogies: Defining and analyzing four types of dialogue in education. **Analytic Teaching and Philosophical Praxis**, Vol. 41, No. 2, p. 106–121, Dec. 2021.

[14] Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Mark Chen, Heewoo Jun, Lukasz Kaiser, Matthias Plappert, Jerry Tworek, Jacob Hilton, Reiichiro Nakano, Christopher Hesse, and John Schulman. Training verifiers to solve math word problems. arXiv:2110.14168, 2021.

[15] Alon Talmor, Jonathan Herzig, Nicholas Lourie, and Jonathan Berant. Commonsenseqa: A question answering challenge targeting commonsense knowledge. arXiv:1811.00937, 2019.

[16] Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. Measuring massive multitask language understanding. In **International Conference on Learning Representations**, 2021.

[17] Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Alex Castro-Ros, Marie Pellat, Kevin Robinson, Dasha Valter, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Zhao, Yanping Huang, Andrew Dai, Hongkun Yu, Slav Petrov, Ed H. Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V. Le, and Jason Wei. Scaling instruction-finetuned language models. arXiv:2210.11416, 2022.

[18] Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models. arXiv:2106.09685, 2021.

[19] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models. arXiv:2201.11903, 2023.

[20] Takeshi Kojima, Shixiang (Shane) Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. Large language models are zero-shot reasoners. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, **Advances in Neural Information Processing Systems**, Vol. 35, pp. 22199–22213. Curran Associates, Inc., 2022.

[21] Chin-Yew Lin. ROUGE: A package for automatic evaluation of summaries. In **Text Summarization Branches Out**, pp. 74–81, Barcelona, Spain, July 2004. Association for Computational Linguistics.

[22] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In **International Conference on Learning Representations**, 2019.

発言	精度	適切さ
補助的	100%	3.40
敵対的	88%	3.10
全体	-	3.25

表 5: 提案手法の議論における反論モデルの補助的と敵対的な発言の適切さに関する人手評価。

A 実験設定の詳細

本研究では、 Huggingface Transformers [3] を利用した。学習は 8 台の NVIDIA A100 GPU(40GiB) で行った。公平で一貫性のある評価を保証するために、提案手法と全ての比較手法において、パラメータ更新ステップの総数を同じに設定した。最適化手法として、 AdamW [22] を、デフォルトのハイパーパラメータで使用した。全ての比較手法で、バッチサイズは 1 で、勾配累積を 1 に設定した。Flan-T5-XL では、 LoRA [4] の設定として、 `alpha` = 32、 `rank` = 8、`target modules`: Query、Key、Value、出力層を含む自己注意の線形射影、を採用した。

B 反論モデルの発言に対する評価

反論モデルの発言を評価するために、 5 人の大学生に GSM8K データセットの例からランダムに抽出したサンプルを評価させた。各評価者には、学習モデルの回答、反論モデルの発言、それに対応する質問と正解が提示された。評価プロセスには以下の 2 つを設けた。

1. 発言の分類: 評価者は、学習モデルの解答と質問のペアの文脈に基づいて、反論モデルの各発言を補助的か敵対的かに分類する。
2. 適切さの評価: 評価者は適切さを評価した。評価者は、発言の適切さを 5 段階で評価した。評価の尺度は、1(期待との整合性なし) から 5(期待との完全な整合性) までであり、反論モデルの発言の妥当性を反映する。評価の基準は以下である。
   - 1 - No Alignment: The response does not align with the expectation.
   - 2 - Minimal Alignment: The response shows minimal alignment with the expectation.
   - 3 - Moderate Alignment: The response is moderately aligned with the expectation.
   - 4 - Strong Alignment: The response aligns strongly with the expectation, with minor deviations.
   - 5 - Complete Alignment: The response fully aligns with the expectation.

表 5 は、反論モデルの発言に対する人手評価の結果を示す。精度の面では、補助的な発言は 100 %の精度で識別され、これらの回答の明瞭性と有効性が高いことが示された。それに対し、敵対的な発言は 88 %と低い精度であった。これは、敵対的な内容を明確に定義することが難しく、評価者によって解釈が異なることが多いためと考えられる。この点は、提案手法の中で敵対的な発言を生成する際に、さらなる改良が必要であることを示唆している。適切さの評価では、支持的な発言は平均 3.40 点であったが、敵対的な発言は 3.10 点とやや低かった。どちらのスコアもスケールの中間点付近を推移していることから、発言は一般的に適切であり、学習プロセスにおける意図された使用方法と一致していることが示唆される。しかし、これらの結果は、両タイプの発言の有効性と影響を高めるための改善点も指摘している。

C モデル間の議論の人手評価

4.1 節で述べたように、提案手法を適用した後の議論能力を総合的に評価するために、CoT の言語化について人手による評価を行った。5 人の大学生が以下の基準で議論のクオリティを評価した。

- **1 - Poor:** The discussion significantly deviates from expected standards, showing a lack of relevance, coherence, or constructive feedback.
- **2 - Fair:** There is some alignment with the gold standard; however, notable deficiencies or inaccuracies are present.
- **3 - Average:** The discussion is reasonably relevant and effective, although minor errors or lapses may be present.
- **4 - Good:** There is a strong alignment with the gold standard, despite the potential presence of minor areas for improvement.
- **5 - Excellent:** The discussion is highly aligned with the gold standard, demonstrating relevance, accuracy, and overall effectiveness.

3) https://github.com/huggingface/transformers
4) https://github.com/huggingface/peft