

# 潜在多様体に基づく中間推論のステアリング

風間健太郎<sup>1</sup> 白藤大幹<sup>1</sup> 斉藤辰彦<sup>1</sup>

<sup>1</sup> 三菱電機株式会社

{Kazama.Kentaro@cw, Shirafuji.Daiki@ay, Saito.Tatsuhiko@db}  
.MitsubishiElectric.co.jp

## 概要

Chain-of-Thought (CoT) は、中間推論を明示的に出力することで、大規模言語モデル (LLM) の性能を向上させることができる。しかしながら、LLM は最終的な回答が正解であっても、論理的に不整合な推論ステップを生成し、中間推論の信頼性を損なうことが指摘されている。本論文では、中間推論の品質を向上させるための潜在多様体に基づく手法を提案する。提案手法は、高品質な CoT の軌跡が形成する潜在空間を VAE により学習する。GSM8k データセットに対する評価実験の結果、提案手法は正答率を最大 2.6 ポイント、推論品質を最大 5.3 ポイント向上させた。本結果は、提案手法が LLM の中間推論品質の改善に有効であることを示唆する。

## 1 序論

大規模言語モデル (Large Language Model; LLM) の推論能力は急速に進展している [1]。Chain-of-Thought (CoT) [2]、Self-Consistency [3]、Tree-of-Thought [4] などの明示的に中間推論ステップを生成する手法は、LLM の下流タスクにおける精度向上を可能とする。LLM の推論過程を出力させることは、モデルがどのような思考で結論に至ったのかを分析することを容易にし、その推論挙動の解釈可能性を高める。

しかしながら、LLM は最終的な回答が正解であるにも関わらず、誤った推論過程を生成する場合があります [5, 6]、これは推論過程の論理的整合性の評価が必要であることを示す。本課題を解決するために、LLM の振る舞いを推論時に制御する手法である Activation Engineering [7] による推論過程のステアリングが活用されている。Kirtania は、証明生成における中間推論の整合性向上に Activation Steering を用いた [8]。一方で、既存手法はユークリッド空間上の線形演算による隠れ状態の操作に依存してお

り、隠れ状態が多様体として表現されることを考慮していない。

本研究では、潜在多様体上で LLM の隠れ状態をステアリングする手法を提案する。提案手法の概要を図 1 に示す。本手法では、gpt-oss-20b [9] により推論ステップごとの品質スコアを付与したデータセットで VAE を学習し、対象 LLM の隠れ状態を潜在空間上のより高品質な領域にステアリングする。VAE が隠れ状態を近似的にガウス分布に正規化された空間に写像することで、元の隠れ状態空間にみられる異方性が緩和されるため、潜在空間上での線形演算が安定すると考えられる。Llama3.2 [10] と Qwen3 [11] の 6 種類の LLM を対象に、GSM8k データセット [12] で提案手法を評価する。評価実験の結果、提案手法は最終回答の精度を向上するのみでなく、生成される推論過程の論理的整合性、忠実性およびステップ単位の一貫性のスコアを向上させることが示された。

本研究の主要な貢献は以下の通りである。

- 潜在空間上で隠れ状態を操作し、より論理的整合性の取れた推論過程を生成する手法の提案
- 提案手法が複数のモデルにおいて、CoT 推論の品質を向上させることの実証

## 2 関連研究

隠れ状態を操作して LLM の振る舞いを変化させる研究として、Activation Engineering が提案されている [7]。Turner らは、対照的な例から得たベクトルを中間層に加算することで、出力の属性を変化させた [13]。Gu らは、隠れ状態を部分空間に分解して LLM の振る舞いを制御した [14]。これらの先行研究は、推論時に介入するという点で、RLHF [15] や DPO [16] などのパラメータ更新を伴う手法と異なる。しかしながら、既存のステアリング手法の多くは、活性化空間における線形な操作を前提として

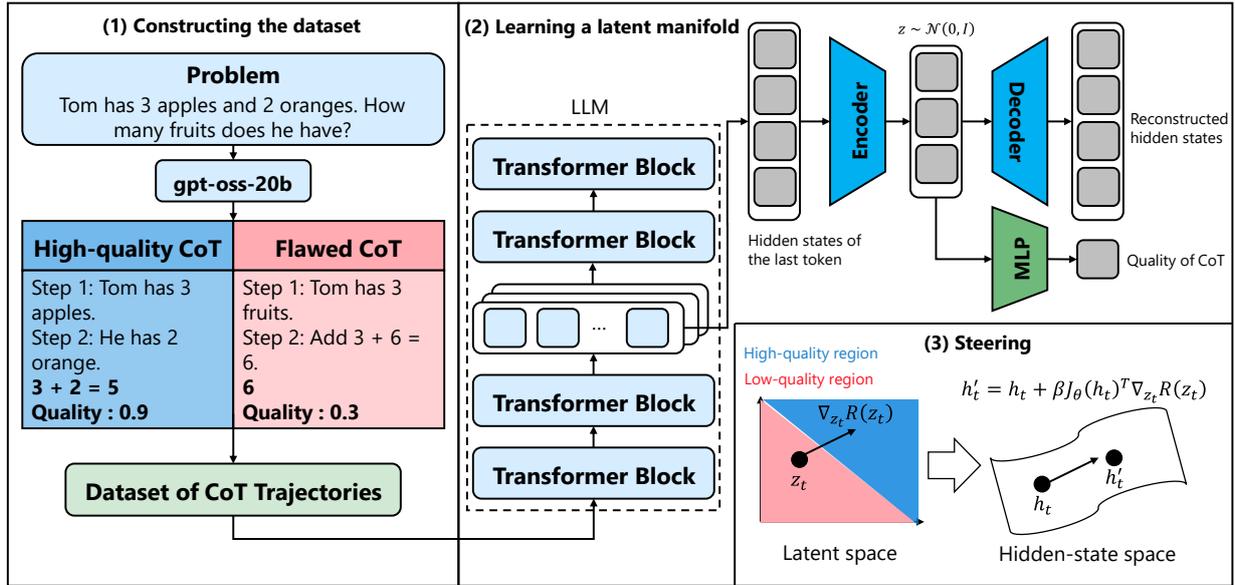


図 1 提案手法の概要図. (1) 各問題に対して複数の推論過程を生成し、各中間ステップに品質スコアを付与することでデータセットを構築し、(2) 高品質な推論の構造を捉える潜在多様体を得るために、VAE を学習し、(3) 推論時に、対象モデルの隠れ状態を、その潜在表現がより高品質な推論に関連付けられた領域に向けて更新する。

いる。

推論過程を改善するために、Process Reward Models (PRM) [17] や Tree-of-Thoughts [4] に基づく手法が提案されている。これらは中間推論の正当性を評価したり、正確な推論を選択したりすることができる。これらの既存研究の多くは生成されたトークン列を対象としており、生成中の内部表現そのものを動的に修正するアプローチは限定的である。

VAE [18] や  $\beta$ -VAE [19] は、複雑なデータの低次元構造を抽出する。LLM の表現も、意味的な抽象度に応じて、幾何学的なパターンや部分空間構造が存在することが示唆されている [20, 21]。本研究では、これらの潜在的な幾何学構造を利用することで、推論過程の制御を動的に可能とする。

### 3 手法

#### 3.1 データセット

$M_{teacher}$  を高品質な CoT を生成する LLM、 $P$  を問題の集合とする。各問題  $p \in P$  に対し、 $M_{teacher}$  は長さ  $T_i$  の CoT

$$C_p^i = [c_{0,0}^i, \dots, c_{T_i-1}^i]$$

を生成する。ここで  $i \in \{\text{pos}, \text{neg}\}$  は、その CoT が高品質であるか、推論過程に欠陥を含むかを表す。高品質な軌跡は、整合性が取れた推論過程を生成するように指示された際に、 $M_{teacher}$  により生成され

る。これらは通常、正しい回答に至る。一方で、欠陥を含む CoT は、自然な言語構造を保ちつつも、計算ミスや論理的な誤りを含む。

CoT の  $t < T_i$  までの部分列

$$C_{p,t}^i = [c_{0,0}^i, \dots, c_{t-1}^i]$$

に対して品質スコア  $q_{p,t}^i \in [0, 1]$  を付与する。このスコアは、ステップ  $t$  までの中間推論が、最終的な正解を導くまでに十分な情報を含んでいるか否かを評価する。スコアは、付録 A.1 に示すプロンプトを  $M_{teacher}$  に与えることで生成する。

以上より、データセットは式 (1) で表される。

$$\mathcal{D} := \left\{ (p, C_p^i, Q_p^i) \mid p \in P, i \in \{\text{pos}, \text{neg}\} \right\} \quad (1)$$

ここで、 $Q_p^i = [q_{p,0}^i, \dots, q_{p,T_i-1}^i]$  は  $C_p^i$  に対応する品質スコアの集合である。データセットは、潜在多様体の学習および品質推定に用いられる。

#### 3.2 潜在多様体に基づくステアリング

隠れ状態の潜在多様体：ステアリング対象のモデルを  $M_{student}$  とし、各部分列  $c_{p,t}^i \subset C_p^i$  の最終トークンにおける隠れ状態を

$$H_p^i = [h_{p,0}^i, \dots, h_{p,T_i-1}^i], h_{p,t}^i \in \mathbb{R}^d$$

と計算する。ここで  $d$  は  $M_{student}$  の隠れ状態次元である。各要素は、ステップ  $t$  を生成した後の  $M_{student}$  による内部状態を表す。

潜在表現を得るために、エンコーダ

$$f_\theta : \mathbb{R}^d \rightarrow \mathbb{R}^k$$

と、デコーダ

$$g_\phi : \mathbb{R}^k \rightarrow \mathbb{R}^d$$

を持つ VAE を学習する。ここで  $k$  は潜在空間次元である。VAE は目的関数として

$\mathcal{L}_{VAE} = \mathbb{E}_{q_\theta(z|h)} [\|g_\phi(z) - h\|^2] + \text{KL}(q_\theta(z|h) \| p(z))$  を最小化する。ここで  $q_\theta(z|h)$  は事後分布、 $p(z)$  は事前分布、 $\text{KL}(\cdot)$  は Kullback-Leibler 情報量である。VAE を学習することにより、幾何学的に滑らかで、かつ元の隠れ状態空間の構造に忠実な潜在多様体が得られる。VAE の学習後、各隠れ状態は潜在ベクトルに射影され、潜在空間  $\mathcal{Z}$  上の軌跡として

$$Z_p^i = [z_{p,0}^i, \dots, z_{p,T_i-1}^i], z_{p,t}^i = f_\theta(h_{p,t}^i) \in \mathcal{Z}$$

が得られる。

**潜在品質関数**：VAE の学習と同時に、潜在ベクトルから品質スコアを予測するスコアリング関数

$$R_\psi : \mathcal{Z} \rightarrow \mathbb{R}$$

を学習する。スコアリング関数は、潜在ベクトルと品質スコアのペア  $(z_{p,t}^i, q_{p,t}^i)$  が与えられたとき、目的関数

$$\mathcal{L}_{score} = \frac{1}{|\mathcal{D}|} \sum_{p,i,t} \left( R_\psi(z_{p,t}^i) - q_{p,t}^i \right)^2$$

を最小化することで学習される。このとき、勾配  $\nabla_z R_\psi(z)$  は、潜在空間において中間推論の品質を向上させる方向を表す。

**隠れ状態のステアリング**：推論時、 $M_{student}$  が生成した隠れ状態  $h_t$  について、対応する潜在表現

$$z_t = f_\theta(h_t)$$

を用いて、勾配

$$\nabla_{h_t} R_\psi(z_t) = J_\theta(h_t)^\top \nabla_{z_t} R_\psi(z_t)$$

を計算する。ここで  $J_\theta(h_t)$  は  $f_\theta$  の  $h_t$  におけるヤコビアンである。これらを用いて、ステアリングされた隠れ状態  $h'_t$  は

$$h'_t = h_t + \beta \frac{\nabla_{h_t} R_\psi(z_t)}{\|\nabla_{h_t} R_\psi(z_t)\|}$$

と計算される。ここで  $\beta$  はステアリングの強度を制御するハイパーパラメータである。更新された状態  $h'_t$  は次のトークンの生成に使用される。すべての  $t$  について  $h'_t$  を計算することで、対象モデルの言語的振る舞いを維持しつつ、潜在多様体に基づくより高品質な領域へ誘導された推論が生成される。

## 4 実験

### 4.1 実験設定

**モデル**：対象 LLM として、Llama-3.2 モデル [10] (1B, 3B) と Qwen3 モデル [11] (0.6B, 1.7B, 4B, 8B) を使用する。評価対象は、ステアリングを実施しない **Baseline** と、提案手法である潜在空間上の更新を通して隠れ状態を修正する **Steered** である。データセット  $\mathcal{D}$  の作成に用いる  $M_{teacher}$  は、gpt-oss-20b[9] を使用する。

**学習データセット**：算術推論のベンチマークである GSM8k[12] を用いて学習データセットを作成する。各データは、自然言語による質問と、段階的な推論を通じて導出可能な回答のペアから構成される。 $M_{teacher}$  は、各事例に対して以下の 2 種類の CoT を生成する。

1. 高品質な CoT：正しく整合性の取れた推論ステップにしたがうもの。
2. 低品質な CoT：誤った、不完全な、あるいは矛盾した推論を含むもの。

生成された CoT は隠れ状態  $\{h_t\}$  に変換され、VAE によって潜在表現  $\{z_t\}$  が学習される。

**評価データセット**：GSM8k のテストセットを用いて、4-shot にて評価を行う。4 つのサンプルは GSM8k のテストセットからランダムに選択されたものである。few-shot プロンプトは付録 A.3 に示す。

**評価指標**：最終回答の正解率と推論品質を独立して評価するため、以下の 2 つの指標を採用する。

1. Exact Match (EM)：予測された最終回答が正解 (ground truth) と完全に一致するか測定する。
2. ペアワイズ勝率 (Win Rate)：GPT-4o により、Baseline と Steered によって生成されたそれぞれの CoT を、明瞭性、整合性、推論品質の観点で比較する。勝率は、Steered が選択された事例の割合である。ペアワイズ勝率の判定に用いたプロンプトは付録に記載する。

ステアリング強度は、付録 A.4 の通り変動させ、各評価指標のスコアが最良となる値を採用する。のように変動させ、各評価指標のスコアが最良となる値を採用する。

表 1 Llama-3.2, Qwen3 モデルにおける Exact Match (%) とペアワイズ勝率 (%). 太字は Baseline と比べて高い性能を発揮したものを表し, \*は McNemar 検定において  $p < 0.05$  で有意差のあった値を表す.

	Llama-3.2				Qwen3							
	1B		3B		0.6B		1.7B		4B		8B	
	Baseline	Steered	Baseline	Steered	Baseline	Steered	Baseline	Steered	Baseline	Steered	Baseline	Steered
EM (%)	42.7	<b>43.4</b> ( $\beta = 1.5$ )	73.0	<b>73.8</b> ( $\beta = 1$ )	60.0	60.0	82.3	<b>84.9*</b> ( $\beta = 150$ )	90.6	88.9	90.7	<b>91.4</b> ( $\beta = 200$ )
Win Rate (%)	20.55	<b>31.61*</b> ( $\beta = 5$ )	4.25	<b>7.66*</b> ( $\beta = 3.5$ )	10.01	<b>15.31</b> ( $\beta = 10$ )	2.81*	<b>6.82*</b> ( $\beta = 50$ )	0.68	<b>2.81*</b> ( $\beta = 10$ )	1.52	<b>8.04*</b> ( $\beta = 150$ )

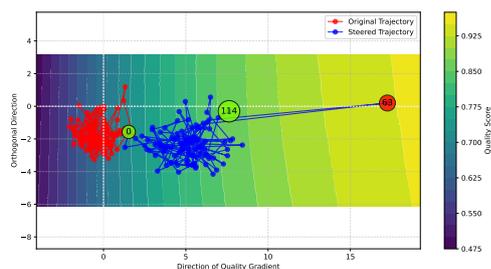


図 2 潜在ベクトルの軌跡を 2 次元平面に描写したものの. x 軸は品質スコアの勾配方向を表し, y 軸はその直交軸を表す. 赤線がベースラインを, 青線が提案手法を表す.

## 4.2 実験結果

表 1 に, 各モデルにおける EM とペアワイズ勝率 (Win Rate) の値を示す. Qwen3-1.7B は,  $\beta = 150$  で EM が 82.3% から 84.9% に向上した. Llama-3.2 では, EM の改善は限定的である. ペアワイズ勝率は, すべてのモデルで改善が見られ, 特に Llama-3.2-1B と Qwen3-0.6B において改善幅が大きい.

最適なステアリング強度  $\beta$  の値は, モデルサイズによって異なる. Llama-3.2 では,  $\beta \leq 5$  が最適であり, Qwen3-8B では  $\beta \in [150, 200]$  で最大の改善幅を観測した.

## 5 考察

提案手法が推論過程をどのように変化させるか検証するため, CoT を 2 次元平面上の軌跡として描写し, 実験を行った. 図 2 は, テストセット内の 1 問題に対する, x 軸を品質スコアの勾配方向としたときの潜在ベクトルの軌跡である. ベースラインの軌跡変化は, 推論全体を通して軽微である. 一方で, ステアリングされた軌跡は, あるトークンの付近で高品質方向への遷移を示している. 本例では, 文法的構造や表現は同一であるにも関わらず, 特定のトークンを発端として内部変化が生じる. すなわち, ステアリングが出力トークンだけではなく内部

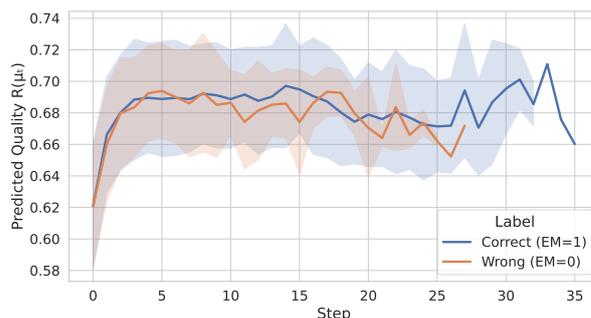


図 3 推論ステップごとの品質スコアの変化

表現自体を変容させ, 推論の構造的な一貫性を向上させる可能性がある.

図 3 は, 推論ステップが進むにつれて, 品質スコアの値がどのように変化したか描写したものである. 最終的な回答が正しいか否かにかかわらず, 品質スコアの値に大きな変化は見られないが, 最終的な回答が不正解の場合, 推論ステップ数が短くなる. 本洞察は, LLM は回答に至るまでの推論の長さ, 最終的な回答の正答率に関連がある可能性を示唆している.

## 6 結論

本研究では, LLM の推論過程を調整するための手法を提案した. 実験の結果, 提案手法は, Llama-3.2 および Qwen3 のモデルにおいて, 最終的な回答性能を損なわずに CoT の品質を向上させた. その結果, 提案手法は, 潜在空間上で品質スコアの勾配を計算することで, 高品質な CoT を生成することが示された. 考察を通して, 推論の品質と最終回答に関連性はあるものの, 等価ではないことを明らかにし, 推論過程そのものの品質を評価することの重要性を主張した.

今後の展望として, 大域的な論理的一貫性・冗長性などの性質を反映した品質スコアの定義, 評価指標の設計や, 人手による評価が挙げられる.

## 参考文献

- [1] OpenAI. Gpt-4 technical report. **arXiv preprint arXiv:2303.08774**, 2023.
- [2] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed H. Chi, Quoc V. Le, and Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models. **arXiv:2201.11903**, 2022.
- [3] Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed H. Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. Self-consistency improves chain of thought reasoning in language models. **arXiv:2203.11171**, 2022.
- [4] Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L. Griffiths, Yuan Cao, and Karthik Narasimhan. Tree of thoughts: Deliberate problem solving with large language models. In **ArXiv preprint arXiv:2305.10601**, 2023.
- [5] Chujie Zheng, Zhenru Zhang, Beichen Zhang, Runji Lin, Keming Lu, Bowen Yu, Dayiheng Liu, Jingren Zhou, and Junyang Lin. Processbench: Identifying process errors in mathematical reasoning. **arXiv preprint arXiv:2412.06559**, 2024.
- [6] Evelyn Yee, Alice Li, Chenyu Tang, Yeon Ho Jung, Ramamohan Paturi, and Leon Bergen. Faithful and unfaithful error recovery in chain-of-thought. In **First Conference on Language Modeling (COLM)**, 2024.
- [7] Andy Zou, et al. Representation engineering: A top-down approach to ai transparency. **arXiv:2310.01405**, 2023.
- [8] Shashank Kirtania. Activation steering in neural theorem provers. **arXiv preprint arXiv:2502.15507**, 2025. v5 (latest version).
- [9] OpenAI. gpt-oss-120b gpt-oss-20b model card. **arXiv preprint arXiv:2508.10925**, 2025.
- [10] Meta AI. The llama 3 herd of models. **arXiv preprint arXiv:2407.21783**, 2024.
- [11] Qwen Team. Qwen3 technical report. **arXiv preprint arXiv:2409.12186**, 2024.
- [12] Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Mark Chen, Heewoo Jun, Lukasz Kaiser, Matthias Plappert, Jerry Tworek, Jacob Hilton, Reiichiro Nakano, et al. Training verifiers to solve math word problems. **arXiv preprint arXiv:2110.14168**, 2021.
- [13] Alexander Matt Turner, Lisa Thiergart, Gavin Leech, David Udell, Juan J. Vazquez, Ulisse Mini, and Monte MacDiarmid. Steering language models with activation engineering. **arXiv preprint arXiv:2308.10248**, 2023.
- [14] Andy Zou, Long Phan, Sarah Chen, James Campbell, Phillip Guo, Richard Ren, Alexander Pan, Xuwang Yin, Mantas Mazeika, Ann-Kathrin Dombrowski, Shashwat Goel, Nathaniel Li, Michael J. Byun, Zifan Wang, Alex Mallen, Steven Basart, Sanmi Koyejo, Dawn Song, Matt Fredrikson, J. Zico Kolter, and Dan Hendrycks. Representation engineering: A top-down approach to ai transparency. **arXiv preprint arXiv:2310.01405**, 2023.
- [15] Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. Training language models to follow instructions with human feedback. In **Advances in Neural Information Processing Systems**, 2022.
- [16] Rafael Rafailov, Archit Sharma, Eric Mitchell, Stefano Ermon, Christopher D. Manning, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. In **Advances in Neural Information Processing Systems**, 2023.
- [17] Zhaohui Yang, Chenghua He, Xiaowen Shi, Shihong Deng, Daxin Jiang, et al. Beyond the first error: Process reward models for reflective mathematical reasoning. **arXiv preprint arXiv:2503.21295**, 2025.
- [18] Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. **International Conference on Learning Representations**, 2014.
- [19] Irina Higgins, Loic Matthey, Arka Pal, Christopher Burgess, Xavier Glorot, Matthew Botvinick, Shakir Mohamed, and Alexander Lerchner.  $\beta$ -VAE: Learning basic visual concepts with a constrained variational framework. In **International Conference on Learning Representations**, 2017.
- [20] Kawin Ethayarajh. How contextual are contextualized word representations? comparing the geometry of BERT, ELMo, and GPT-2 embeddings. In **Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing**, 2019.
- [21] John Hewitt and Christopher D. Manning. A structural probe for finding syntax in word representations. In **Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics**, 2019.

## A 付録

### A.1 品質スコア判定プロンプト

gpt-oss-20b による品質スコアの判定には、以下のプロンプトを用いた。

```
You are an expert reasoning evaluator.
You will rate how logically consistent and correct the reasoning is
with respect to solving the following problem.
Problem:
{problem_text}
Reasoning prefix:
{prefix_text}
Give a single numeric score between 0.0 and 1.0 where:
- 1.0 = perfectly correct reasoning leading to a valid conclusion,
- 0.0 = completely wrong or illogical reasoning,
- values in between indicate partial correctness.
Respond only with the numeric score.
```

### A.2 ペアワイズ勝率判定プロンプト

GPT-4o によるペアワイズ勝率の判定には、以下のプロンプトを用いた。

```
You are an expert evaluator of reasoning quality and answer correctness.
You will be given two responses, Response A and Response B, to the
same problem.
Your job is to judge which one is better overall in terms of:
1. Correct final answer (most important)
2. Logical coherence and clarity of reasoning steps
3. Conciseness (less irrelevant detail)
4. Self-consistency (no arithmetic errors, no contradictions)
**Instructions:**
1. Extract the final answer from each response and compare it to the
ground truth. If one answer is correct and the other is incorrect, the
correct one *wins*.
2. If both answers are correct, compare reasoning: choose the one
that is more logically coherent, clearer, and more concise.
3. If both are incorrect, compare how close each reasoning is to a
correct reasoning path (e.g., minor arithmetic slip vs completely wrong
approach).
4. If both responses are equivalent in correctness and reasoning
quality (no clear winner), score as a **Tie**.
5. Provide exactly one of the following tokens as a following format:
'<<A>>', '<<B>>', or '<<Tie>>'.
6. After that, provide a short justification (1-3 sentences) for your
decision.
**Inputs:**
Problem: {problem}
Response A: {answer_a}
Response B: {answer_b}
Ground truth answer: {ground_truth}
Judge:
```

### A.3 few-shot プロンプト

Qwen3 系列における評価時に  $M_{student}$  に与える few-shot プロンプトは、以下の通りである<sup>1)</sup>。

```
<|im_start|>system
You are a concise math assistant. Use <think></think> for internal
reasoning,
but after </think> only output the final numeric answer in the for-
mat:
'The final answer is <number>.'
<|im_end|>
<|im_start|>user
Natalia sold clips to 48 of her friends in April, and then she sold half
as many clips in May. How many clips did Natalia sell altogether in
April and May?
<|im_end|>
<|im_start|>assistant
The final answer is 72.
<|im_end|>
<|im_start|>user
Weng earns $12 an hour for babysitting. Yesterday, she just did 50
minutes of babysitting. How much did she earn?
<|im_end|>
<|im_start|>assistant
The final answer is 10.
<|im_end|>
<|im_start|>user
Betty is saving money for a new wallet which costs $100. Betty has
only half of the money she needs. Her parents decided to give her $15
for that purpose, and her grandparents twice as much as her parents.
How much more money does Betty need to buy the wallet?
<|im_end|>
<|im_start|>assistant
The final answer is 5.
<|im_end|>
<|im_start|>user
Julie is reading a 120-page book. Yesterday, she was able to read 12
pages and today, she read twice as many pages as yesterday. If she
wants to read half of the remaining pages tomorrow, how many pages
should she read?
<|im_end|>
<|im_start|>assistant
The final answer is 42.
<|im_end|>
```

### A.4 ステアリング強度

#### Llama-3.2

$$\beta \in \{0.1, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 5, 10\}$$

#### Qwen3

$$\beta \in \{1, 10, 50, 100, 125, 150, 200, 300\}$$

1) Llama-3.2 シリーズでは、システムプロンプトから <think> タグに関する指示を削除したものをを用いる。