Fermi-MCTS：フェルミ推定のための LLM 推論フレームワーク

丸田敦貴 [1]　加藤誠 [1,2]
[1] 筑波大学大学院　[2] 筑波大学 図書館情報メディア系　[3] 国立情報学研究所
s1711576@klis.tsukuba.ac.jp　mpkato@acm.org

概要

本研究では，フェルミ推定タスクにおいて，大規模言語モデル（LLM）の内部知識と数学的推論能力を統合的に扱うための推論フレームワーク Fermi-MCTS を提案する．本手法は，モンテカルロ木探索を基盤とし，LLM のキャリブレーションに基づく中間数値の信頼性と，推論全体の妥当性を報酬として組み込むことで，信頼性が高く妥当な推論経路を探索する．実験の結果，提案手法は CoT や ToT，RAG 手法と比較して，高い推定精度を示した．

1　はじめに

近年，大規模言語モデル（LLM）の発展により，常識的推論や数学的推論など幅広いタスクで顕著な進歩が見られている [1, 2, 3]．これらの成果は LLM が知識の保持に加え，論理構造の理解や数学的推論を行えることを示唆している．一方で，これらの研究の多くは，推論に必要な数値情報や前提が与えられている状況を想定している．現実世界の多くの課題は，推論に必要な数値や前提条件を自ら補完し，それらを組み合わせて推論することが求められる．

そのような課題の代表例として，フェルミ推定タスク [4, 5, 6] が挙げられる．このタスクは，直接測定が困難な数値を複数の中間数値に分解し，それらの数値を組み合わせて推定するタスクである．たとえば「NLP2026 で飲まれるコーヒーは何杯？」という質問に対して，参加人数や 1 人あたりの消費量などの中間数値を仮定して計算を行う．このようなタスクを LLM が解けるようになれば，人間が行う概算の見積もりや，十分なデータが得られていない状況における意思決定を支援することが期待される．

フェルミ推定は中間数値の仮定と計算を伴うため，数学的推論能力の向上が重要となる．LLM の数学的推論能力を向上させる手法として，段階的に推論を行う Chain-of-Thought（CoT）[1] や，複数の推論を探索する Tree-of-Thought（ToT）[2]，推論を強化学習によって探索するモンテカルロ木探索（MCTS）[3] などの手法が提案されている．しかし，これらの手法をそのままフェルミ推定に適用すると，次の二つの課題が生じる．

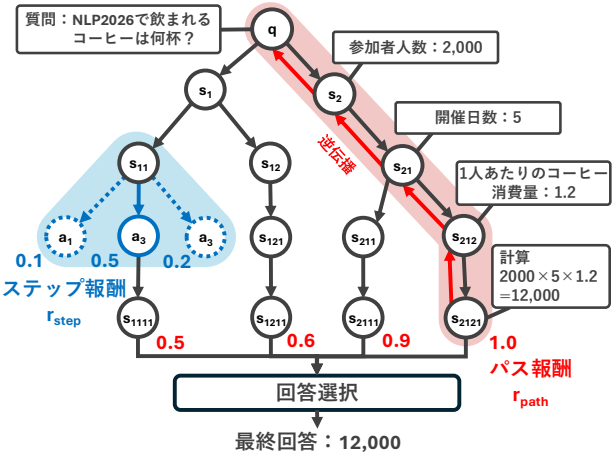

図 1　提案手法 Fermi-MCTS の概要図．

課題 1：中間数値の誤り　フェルミ推定では複数の中間数値を仮定して計算を行うが，LLM の生成する数値には誤差が含まれている可能性があり，誤差の大きい中間数値を用いると，最終的な推定値の精度が大きく損なわれる．

課題 2：計算全体の妥当性の欠如　個々の中間数値が正確でも，それらを組み合わせる数式全体が妥当でなければ，正しい推定は困難である．

そこで本研究では，フェルミ推定タスクにおいて LLM の内部知識と数学的推論を統合的に扱うための推論フレームワーク Fermi-MCTS を提案する（図 1）．本フレームワークでは，モンテカルロ木探索を用いて推論経路を探索し，LLM 内部の数値情報の信頼性と推論過程の妥当性に基づいて推論経路を誘導することで，信頼性が高く妥当な数値推定を行う．

工夫 1：LLM 内部知識の信頼性の評価　LLM のキャリブレーション研究 [7, 8, 9] の知見を応用し，推論過程で生成される中間数値に対し，確

This work is published without peer review and is licensed by the author(s) under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

信度に基づく**ステップ報酬**を与えることで，信頼性の高い数値知識を用いる推論に誘導する．

工夫２：推論の妥当性の評価　回答に達した段階で，推論過程全体の妥当性を評価する**パス報酬**を与えることで，局所的には正しくても全体として非現実的な推論を抑制する．

工夫３：推論の妥当性に基づく回答選択　探索で得られた複数の推論経路の中から，推論過程の妥当性に基づいて最終推定値を選択することで，より信頼性の高い最終回答を導出する．

実験では，フェルミ推定データセットを用いて評価を行った結果，提案手法は CoT や ToT，RAG 手法を上回る性能を示し，本フレームワークによる LLM のフェルミ推定能力の向上が確認された．

論文の貢献は以下の 3 点である．

1. 数学的推論と一般常識を必要とするフェルミ推定タスクに対し，LLM の内部知識を活用した MCTS フレームワークを提案した．
2. LLM 内部の数値情報の信頼性と推論過程の妥当性を定量化するための報酬関数を提案した．
3. 実験を通じて，本手法がフェルミ推定タスクにおいて既存手法よりも高い推定精度を示した．

2　Fermi-MCTS

本研究では，LLM の内部知識と数学的推論能力を統合し，フェルミ推定を行うための推論フレームワーク Fermi-MCTS を提案する（図 1）．提案手法は，MCTS [10] を基盤として，フェルミ推定特有の中間数値の誤りと計算全体の妥当性の欠如に対処する報酬設計および回答選択手法を導入する．具体的には，（1）**ステップ報酬**，（2）**パス報酬**，（3）**回答選択**の 3 要素から構成される．これら 3 つの仕組みにより，局所的な中間数値の誤差を抑制しつつ，妥当性のあるフェルミ推定を実現する．

2.1　MCTS による推論アルゴリズム

本研究では，フェルミ推定タスクを以下のように定式化する．与えられた質問 q に対し，推論は一連の推論ステップ $s_1, s_2, \ldots, s_T$ から構成される．各推論ステップ s_t は，中間数値の仮定や計算など，推論過程における 1 つの操作を表し，推論の状態 τ_t は，それまでの推論軌跡 $\tau_t = (q, s_1, s_2, \ldots, s_t)$ と定義する．推論生成モデル f_{gen} は現在の状態 τ_t を入力として，次に追加可能な推論ステップ候補 $\{a_{t+1}^{(1)}, \ldots, a_{t+1}^{(K)}\}$ を生成する．各候補 $a_{t+1}^{(k)}$ は，選択されることで新たな推論ステップ s_{t+1} となり，推論状態は $\tau_{t+1} = \tau_t \oplus s_{t+1}$ へ遷移する．この探索を通じて最終出力 y を求める．

本フレームワークは MCTS の 4 段階（選択，展開，シミュレーション，逆伝播）を繰り返す探索アルゴリズムに基づいており，探索木の各ノードは 1 つの推論ステップ s を表し，ノードまでのパスは推論状態 τ に対応する．まず**選択**では，これまでに構築された探索木の中から，次に探索を行うべき推論状態を決定する．具体的には，各ノード s に蓄積された Q 値と訪問回数に基づき，Upper Confidence bound applied to Trees [10] を用いて展開するノードを選択する．次に，**展開**では，選択された推論状態 τ に対して，推論生成モデル f_{gen} を用い，推論ステップ候補 $\{a_1 \ldots a_K\}$ を生成する．各候補に対してステップ報酬を計算し，最も高い報酬を持つ候補を選択することで，新たな推論ステップ s を探索木に追加する．そして**シミュレーション**では，回答に到達するか，深さ制限 L に達するまで推論を進め，得られた推論軌跡 τ に対してパス報酬を計算する．**逆伝播**では，パス報酬を推論軌跡上の各ノードに逆伝播させ，Q 値と訪問回数を更新し，次回以降の選択に反映させる．なお，アルゴリズムの詳細は付録 A に記す．

2.2　報酬設計

Fermi-MCTS では，中間的な推定数値の信頼性と計算方法の妥当性を担保するため，ステップ報酬関数 f_{step} とパス報酬関数 f_{path} の 2 つを導入する．ステップ報酬は各中間数値の信頼性を評価し，パス報酬は最終的な推論経路の妥当性を評価する．これら二つの報酬は MCTS における展開と逆伝播で利用され，局所的に信頼性が高く，かつ全体的に妥当性のある推論経路を選択する役割を果たす．

2.2.1　数値信頼性に基づくステップ報酬

各推論ステップ a_k において，LLM が生成する中間数値の一貫性に基づいて，その信頼性を定量化する．本研究では，出力の一貫性を信頼性の指標とするキャリブレーション手法 [7, 8, 9] を数値出力に応用し，同一推論ステップに対する複数生成結果の分布からステップ報酬 r_{step} を計算する．具体的には，推論ステップ a_k の中の指示文（例：日本の総人口を推定する）に対し LLM を m 回生成し，数値候補を得る：$y_i = f_{\text{gen}}(a_k)(i = 1, 2, \cdots, m)$．各候補値に常

This work is published without peer review and is licensed by the author(s) under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

用対数変換を行い，$x_i = \log_{10}(y_i)$ を得る．次に，対数変換後の数値候補集合 $\{x_i\}$ に対してカーネル密度推定を適用し，LLM が出力した数値の確率分布 $\hat{p}(x)$ を得る：

$$\hat{p}(x) = \frac{1}{Z} \sum_{i=1}^{m} K_h(x - x_i). \tag{1}$$

ここで，$K_h(\cdot)$ はカーネル関数を表し，本研究では次式で定義されるガウスカーネルを用いる：

$$K_h(x - x_i) = \frac{1}{h} \exp\left(-\frac{(x - x_i)^2}{2h^2}\right). \tag{2}$$

h はバンド幅，Z は正規化定数である．得られた確率分布 $\hat{p}(y)$ に対して，そのエントロピー

$$H = -\int \hat{p}(x) \log \hat{p}(x)\, dy \tag{3}$$

を計算する．エントロピーが小さいほど，数値出力が特定の値に集中しており，LLM による出力が一貫していると考えられるため，ステップ報酬を次式のように定義する：

$$r_{\text{step}} = f_{\text{step}}(a_k) = \frac{1}{1 + H}. \tag{4}$$

2.2.2 推論の妥当性に基づくパス報酬

推論軌跡の終端状態 τ' において，推論過程の妥当性に基づいてパス報酬を $r_{\text{path}} = f_{\text{path}}(\tau')$ 算出する．f_{path} は、推論経路の妥当性を評価する関数であり，具体的には，LLM を用いて推論過程を評価し，計算手順の正当性や整合性に基づいて報酬を出力する．また，パス報酬のプロンプトを付録 B に示す．

2.3 回答選択手法

探索によって得られた終端ノードと対応するパス報酬の集合 $Y = \{(y_i, r_{\text{path},i}) \mid i = 1, 2, \ldots, M\}$ から，最終的な数値回答 $\hat{y}$ を決定する．本節では，節 2.2.1 で導入した対数空間におけるカーネル密度推定と同様の手法を用いて，最終候補値の分布を集約する．

具体的には，各候補値 y_i に対して常用対数変換を施し，得られた集合 $\{x_i\}$ に対して，パス報酬 $r_{\text{path},i}$ を重みとするカーネル密度推定を行う：

$$\hat{p}(x) = \frac{1}{Z} \sum_{i=1}^{M} r_{\text{path},i} K_h(x - x_i). \tag{5}$$

ここで，カーネル関数 $K_h(\cdot)$ は式 2 で定義されるガウスカーネルである．最後に，各候補点 x_i における密度 $\hat{p}(x_i)$ を評価し，最大値を与える点に対応する元の数値 y_i を最終回答として選択する：

$$\hat{y} = y_{j^*}, \quad j^* = \arg\max_i \hat{p}(x_i). \tag{6}$$

3 実験

3.1 実験設定

データセット　本研究では，Kalyan らが提案した 1,557 件のフェルミ推定データセット [6] を評価に利用した．このデータセットには，質問文に加えて推論を補助するための補足情報が含まれているが，本研究では LLM の内部知識と推論能力をどの程度活用できるかを検証するため，質問文のみを入力とする設定を採用した．

評価指標　数値推定の精度を評価するために，既存のフェルミ推定の評価に用いられる fp score[6] をより一般化した指標 fp score@k（fp@k）を用いる．モデルの予測値 A' と正解値 A の許容誤差を指定するパラメータ k を導入した fp@k を以下に定義する：

$$\text{fp@}k = \max\left(0, \frac{1}{k}\left|\log_{10}\frac{A'}{A}\right|\right) \tag{7}$$

利用モデル・パラメータ　推論生成モデル f_{gen} および報酬関数には Llama3.1-8B-Instruction を用い，tempreture は 0.6 とした．また，探索パラメータ c は 2，ロールアウト回数 N_{iter} は 32, 深さ制限 L は 8，推論ステップ生成数 K は 8，ステップ報酬のサンプル数 m は 16 とした．

ベースライン　本研究では，提案手法の有用性を検証するために，既存の RAG 手法と LLM の内部知識に基づく手法をベースラインとして設定した．

まず，RAG ベースラインについて説明する．**MCR** [11] は，Vicuna-13B を用いた RAG フレームワークであり，多段階推論が必要な質問をサブ質問に分解し，関連する文書を反復的に検索する手法である．**SPRING** [12] は，Llama2-7B 用いた RAG のための追加トークンを導入した手法である．本研究では，検索モデルに BM25 [13]，外部知識に Wikipedia[1] を使用した．

次に，内部知識ベースラインについて説明する．**直接推定** は，中間的な推論過程を伴わず，直接数値を生成する手法である．**CoT** [1] は，質問に対して段階的に推論を行う手法である．**CoT+SC@32** [14] は，同一プロンプトに対して 32 本の CoT をサンプリングし，多数決によって最終回答を選択する手法である．**ToT@32** [2] は，推論過程を木構造として表

1) https://huggingface.co/datasets/Tevatron/wikipedia-nq-corpus

This work is published without peer review and is licensed by the author(s) under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

表 1　主な実験結果．既存のベースライン手法と提案手法の比較．

手法	fp@3	fp@5	fp@10
RAG ベースライン			
MCR [11]	0.234	–	–
SPRING [12]	0.307	0.413	0.571
内部知識ベースライン			
直接推定	0.146	0.196	0.266
CoT [1]	0.463	0.569	0.700
CoT+SC@32 [14]	0.466	0.560	0.681
ToT@32 [2]	0.434	0.555	0.708
提案手法			
Fermi MCTS	**0.489**	**0.599**	**0.733**

表 2　提案手法における報酬のアブレーションスタディ．

報酬設定	fp@3	fp@5	fp@10
Fermi MCTS	**0.489**	**0.599**	0.733
– **報酬なし**	0.434	0.553	0.705
– **パス報酬なし**	0.437	0.553	0.701
– **ステップ報酬なし**	0.484	0.597	**0.736**

現し，複数の候補となる推論分岐を探索する手法である．本研究では，幅優先探索を適用し，32 本の推論経路が生成された時点で探索を終了した．

3.2　実験結果

表 1 は提案手法とベースライン手法との比較結果を示す．まず，RAG ベースラインの MCR，SPRING と比較すると，提案手法はすべての指標で大きく上回った．RAG ベースラインは外部知識を明示的に参照しているにもかかわらず，提案手法は内部知識を活用した推論のみ高い推定精度を示した．これは，LLM の内部知識を適切に活用しつつ，探索を通じて推論の妥当性を高める本手法の有効性を示唆している．次に，内部知識ベースラインとの比較では，提案手法はすべての手法と比べて高い推定精度を達成した．これは，単純な推論よりも，報酬に基づいて多様な推論を探索する手法がより効果的であることを示唆している．

表 2 に，提案手法の報酬を除去した場合の性能変化を示す．報酬をすべて除去した場合とパス報酬を除いた場合に大きなスコア低下が見られた．このことから，パス報酬のような推論経路の妥当性を評価する仕組みがより効果的であることが示唆される．

表 3　回答選択手法の比較．

回答選択手法	fp@3	fp@5	fp@10
パス報酬最大値	0.408	0.527	0.683
重み付き多数決	0.428	0.545	0.694
提案手法	**0.489**	**0.599**	**0.733**

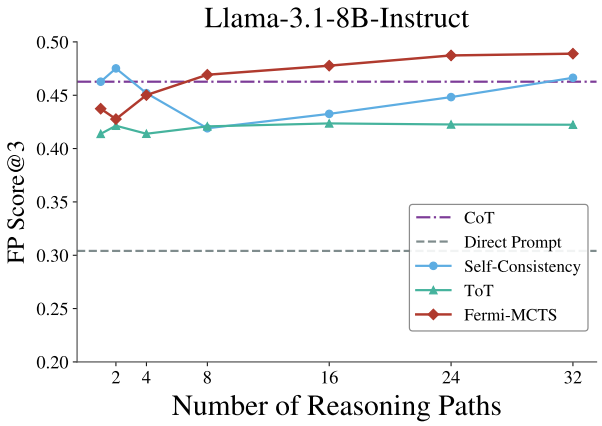

図 2　推論経路数による性能の変化．

表 3 に，最終回答の選択方法について様々な手法を適用したときの性能比較を示す．パス報酬の最大値を最終回答とする手法やパス報酬を重みして多数決で回答を決定する手法は，低い予測精度を示した．一方で，確率密度分布に基づいて回答を選択する提案手法は大幅な性能改善を示した．この結果は，単純な多数決ではなく，回答の分布を解析して代表値を選択する方法が，フェルミ推定において有効であることを示している．

図 2 は推論経路数による性能の変化を図示している．提案手法は推論経路数の増加に伴って精度が改善する傾向が見られる．一方，ToT は推論回数を増やしても性能が向上せず，CoT+SC は一度性能が低下する傾向を示した．これは，単純に推論経路を増やすだけでは，全体の精度が向上しないことを示唆している．

4　おわりに

本研究はフェルミ推定タスクのために，LLM 内部の数値情報の信頼性および推論過程の妥当性を報酬として付与し，MCTS によって推論を探索するフレームワーク Fermi-MCTS を提案した．実験の結果，提案手法 CoT や ToT，RAG 手法と比較して，高い推定精度を示した．今後の発展として，MCTS に外部知識の検索を組み込むことが挙げられる．

This work is published without peer review and is licensed by the author(s) under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

謝辞

本研究は JSPS 科研費 JP24K03048 及び JST 次世代研究者挑戦的研究プログラム JPMJSP2124 の支援を受けたものです．また，産総研及び AIST Solutions が提供する ABCI 3.0 を利用しました．

参考文献

[1] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. Chain-of-thought prompting elicits reasoning in large language models. **Advances in neural information processing systems**, Vol. 35, pp. 24824–24837, 2022.

[2] Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Tom Griffiths, Yuan Cao, and Karthik Narasimhan. Tree of thoughts: Deliberate problem solving with large language models. **Advances in neural information processing systems**, Vol. 36, pp. 11809–11822, 2023.

[3] Shibo Hao, Yi Gu, Haodi Ma, Joshua Hong, Zhen Wang, Daisy Wang, and Zhiting Hu. Reasoning with language model is planning with world model. In **Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing**, pp. 8154–8173, 2023.

[4] Germaine L. Taggart, Paul E. Adams, Ervin Eltze, John Heinrichs, James Hohman, and Karen Hickman. Fermi questions. **Mathematics Teaching in the Middle School**, Vol. 13, No. 3, pp. 164–167, 2007.

[5] Costas J Efthimiou and Ralph A Llewellyn. Cinema, fermi problems and general education. **Physics education**, Vol. 42, No. 3, p. 253, 2007.

[6] Ashwin Kalyan, Abhinav Kumar, Arjun Chandrasekaran, Ashish Sabharwal, and Peter Clark. How much coffee was consumed during emnlp 2019? fermi problems: A new reasoning challenge for ai. In **Conference on Empirical Methods in Natural Language Processing (EMNLP 2021)**, pp. 7318–7328. Association for Computational Linguistics, 2021.

[7] Stephanie Lin, Jacob Hilton, and Owain Evans. Teaching models to express their uncertainty in words. **Transactions on Machine Learning Research**, 2022.

[8] Katherine Tian, Eric Mitchell, Allan Zhou, Archit Sharma, Rafael Rafailov, Huaxiu Yao, Chelsea Finn, and Christopher D Manning. Just ask for calibration: Strategies for eliciting calibrated confidence scores from language models fine-tuned with human feedback. In **Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing**, pp. 5433–5442, 2023.

[9] Potsawee Manakul, Adian Liusie, and Mark Gales. Selfcheckgpt: Zero-resource black-box hallucination detection for generative large language models. In **Proceedings of the 2023 conference on empirical methods in natural language processing**, pp. 9004–9017, 2023.

[10] Levente Kocsis and Csaba Szepesvári. Bandit based monte-carlo planning. In **European conference on machine learning**, pp. 282–293. Springer, 2006.

[11] Ori Yoran, Tomer Wolfson, Ben Bogin, Uri Katz, Daniel Deutch, and Jonathan Berant. Answering questions by meta-reasoning over multiple chains of thought. In **Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing**, pp. 5942–5966, 2023.

[12] Yutao Zhu, Zhaoheng Huang, Zhicheng Dou, and Ji-Rong Wen. One token can help! learning scalable and pluggable virtual tokens for retrieval-augmented large language models. In **Proceedings of the AAAI Conference on Artificial Intelligence**, Vol. 39, pp. 26166–26174, 2025.

[13] Stephen E Robertson, Steve Walker, Susan Jones, Micheline M Hancock-Beaulieu, Mike Gatford, et al. **Okapi at TREC-3**. British Library Research and Development Department, 1995.

[14] Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc V. Le, Ed H. Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. Self-consistency improves chain of thought reasoning in language models. In **The Eleventh International Conference on Learning Representations**, 2023.

This work is published without peer review and is licensed by the author(s) under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).

A　推論アルゴリズムの詳細

Algorithm 1: MCTS を用いた LLM の推論アルゴリズム

Require: 質問 q，推論生成モデル f_{gen}，ステップ報酬関数 f_{step}，パス報酬関数 f_{path}，探索パラメータ c，探索回数 N_{iter}，最大推論深さ L，推論ステップ候補数 K，回答選択手法 AnswerSelector

Ensure: 最終回答 $\hat{y}$

Initialize $\tau_0 \leftarrow (q)$, $Q(s) \leftarrow 0$, $N(s) \leftarrow 0$ for all s, answer candidate set $Y \leftarrow \emptyset$
for $n \leftarrow 1$ to N_{iter} **do**
  $\tau \leftarrow \tau_0, t \leftarrow 0$
  while τ is expandable and $t < L$ **do**
    $s \leftarrow \arg\max_s \left[\frac{Q(s)}{N(s)} + c\sqrt{\frac{\ln N(\text{parent}(s))}{N(s)}} \right]$　選択
    $\tau \leftarrow \tau \oplus s, t \leftarrow t + 1$
  end while
  $\{a_1, \ldots, a_K\} \leftarrow f_{\text{gen}}(\tau)$　展開
  for $k \leftarrow 1$ to K **do**
    $r_{\text{step},k} \leftarrow f_{\text{step}}(a_k)$
  end for
  $a^* \leftarrow \arg\max_k r_{\text{step},k}$
  $s^* \leftarrow a^*$
  $\tau' \leftarrow \tau \oplus s^*$
  if τ' is terminal or $t = L$ **then**
    Extract answer y' from τ'
    $r_{\text{path}} \leftarrow f_{\text{path}}(\tau')$
    Append (y', r_{path}) to Y
    for each s along the path of τ' **do**
      $N(s) \leftarrow N(s) + 1, Q(s) \leftarrow Q(s) + r_{\text{path}}$
    end for
  end if
end for
$\hat{y} \leftarrow \text{AnswerSelector}(Y)$
return $\hat{y}$

B　パス報酬プロンプト

図 3 にパス報酬で用いた指示文を示す．

```
<Instruction>
You are a strict unit-consistency checker.
Your task is to evaluate whether given
    calculation steps produce a result that
     is dimensionally plausible for the
    stated quantity.
<Output Constraints>
You must return two elements in your
    answer:
1. A short textual rationale (1,2
    sentences) explaining your judgment.
2. A JSON object in the following form:
    {{"score": float number}}.
Do not output anything else beyond the
    rationale and the JSON object. After
    the JSON object, you MUST append the
    closing tag </answer> to indicate the
    end of the answer.
Scoring rubric:
1.0 = clearly consistent
0.5 = unclear but possibly plausible
0.0 = inconsistent
```

図 3　パス報酬のプロンプト

C　ステップ報酬に関する分析

表 4 に，ステップ報酬の設計を変化させた場合の性能を示す．キャリブレーションに関する既存研究では，言語的な自己申告に基づく言語ベース手法 [7, 8] と，出力の一貫性に基づく一貫性ベース手法 [9] の二つのアプローチが提案されており，これら二つの手法を数値推定タスクに適用し，比較した．

表 4　ステップ報酬を変更した場合の性能変化．

手法	fp@3	fp@5	fp@10
言語ベース			
確信度直接出力	0.469	0.582	0.726
誤差範囲出力	0.460	0.571	0.708
一貫性ベース			
厳密一致の割合	0.476	0.584	0.723
提案手法	**0.489**	**0.598**	**0.733**

まず，言語ベース手法では，確信度を直接出力させる方法と，出力数値の誤差範囲を出力させる方法を比較対象とした．いずれの方法も提案手法に比べて低い性能を示し，言語的な確信度が数値推定の信頼性を十分に反映していないことが示唆された．次に，一貫性ベース手法では，厳密一致の割合を用いる手法を比較対象とした．厳密一致の割合を用いる手法は言語ベースの手法よりも高い性能を示したが，提案手法と比較すると低い性能であった．以上の結果から，フェルミ推定タスクにおいて，一貫性ベース手法が信頼性の高いステップ報酬として機能することが確認され，特に数値分布から一貫性を計算する方法がより効果的であることが示唆された．

This work is published without peer review and is licensed by the author(s) under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/).