

クラウドソーシングとAIエージェントによる 推論ネットワーク構築最適化

西田 亮祐[†] 伊藤 寛祥^{††} 森嶋 厚行^{††}

[†] 筑波大学情報学群情報メディア創成学類

〒 305-0821 茨城県つくば市春日 1 丁目 2

^{††} 筑波大学 図書館情報メディア系 〒 305-8550 茨城県つくば市春日 1 丁目 2

E-mail: [†]Ryosuke.Nishida.2025b@mlab.info, ^{††}{ito,mori}@slis.tsukuba.ac.jp

あらまし 近年ファクトチェックなどの問題に対して、木田らはクラウドソーシングを用いて確信度付き推論ネットワークを構築し、最適化する手法を提案した。しかし木田らの手法では構築と最適化が分離されており、構造の動的更新ができないという問題があった。本研究では、ネットワーク構築の段階から強化学習を導入し構築過程そのものを最適化する手法を提案する。具体的には、構築中のネットワークを状態、次の構築操作を行動とするマルコフ決定過程でモデル化し、Deep Q-Network (DQN) に基づく AI エージェントによって最適化する。本研究における AI エージェントとは、行動タイプ選択・仮説選択・ノード選択の 3 つの出力ヘッドを持つ DQN を用いた強化学習エージェントであり、各状態における次の構築操作とその操作先を決定する。本論文では「木田らの手法で得られた推論ネットワークと同等の確信度差を達成する推論ネットワークを、構築段階から強化学習を導入することでどの程度低コストで構築できるのか」というリサーチクエスチョンに取り組んだ。シミュレーション実験の結果、提案手法は木田らの手法と比較して約 60% のコスト削減を達成し、効率よく確信度の高い推論ネットワーク構造を得られることが示唆された。

キーワード クラウドソーシング, ヒューマンコンピテーション, Human-in-the-loop システム, 強化学習, 意見集約

1 はじめに

近年、ファクトチェックなどの複雑な問題に対して多くの人々から情報を収集し問題解決に役立てる方法として、クラウドソーシングは注目を集めている [2], [4], [8]。しかし、従来の手法では結果のみを集計し、主に多数決といった単純な手法で意見を統合するため、人々の意見や根拠は反映されず、なぜその結果に至ったのかを説明することが困難であるという問題が存在する。

これに対し、木田らは MYCIN [10] を代表とする推論モデルに着想を得て、ワーカからの意見を「論証」という形で収集し、確信度付き推論ネットワークを構築する手法を提案した。この手法により、結果だけでなくその結果に対する論理的な説明を得ることが可能となった。また木田らは、推論ネットワーク構築後に新たな根拠の追加先を強化学習によって最適化することで、効率よく確信度を向上させられることを示した。

しかし、木田らの手法では、推論ネットワークをあらかじめ構築し、その後の根拠の追加先のみを対象に強化学習による最適化を行っている。すなわち、ネットワーク構築と最適化が分離されており、構築済みのネットワーク構造そのものには手を加えない前提で設計されている。しかし現実には、最適化を進める過程でネットワーク構造の不備が判明する場合もあり、構造自体が動的に更新されるべきである。また、最適化の対象を構築過程そのものへと拡張することで、より効率的に確信度の

高いネットワークを構築できると考えられる。そこで本研究のリサーチクエスチョンは以下の通りである。

木田らの手法で得られた推論ネットワークと同等の確信度差を達成する推論ネットワークを、構築段階から強化学習を導入することでどの程度低コストで構築できるのか？

本研究では、確信度付き推論ネットワーク構築の段階から強化学習を導入し、構築過程そのものを最適化する手法を提案する。具体的には、Deep Q-Network (DQN) に基づく AI エージェントを用いて、構築操作の選択を最適化する。本研究における AI エージェントとは、行動タイプ選択 (論証追加・命題統合・根拠追加の 3 種類から選択)、仮説選択 (論証追加時の対象仮説を選択)、ノード選択 (根拠追加時の対象ノードを選択) の 3 つの出力ヘッドを持つ DQN を用いた強化学習エージェントである。この AI エージェントが各状態において最適な構築操作とその操作対象を決定することで、効率的なネットワーク構築を実現する。シミュレーション実験の結果、提案手法は木田らの手法と比較して効率よく確信度の高い推論ネットワーク構造を得られることが明らかとなった。

図 1 に、本研究で提案する確信度付き推論ネットワークの全体像を示す。

本研究の貢献は以下の通りである。

1. 確信度付き推論ネットワーク構築過程の最適化枠組みの提案：従来手法では構築後の根拠追加のみを最適化対象とし

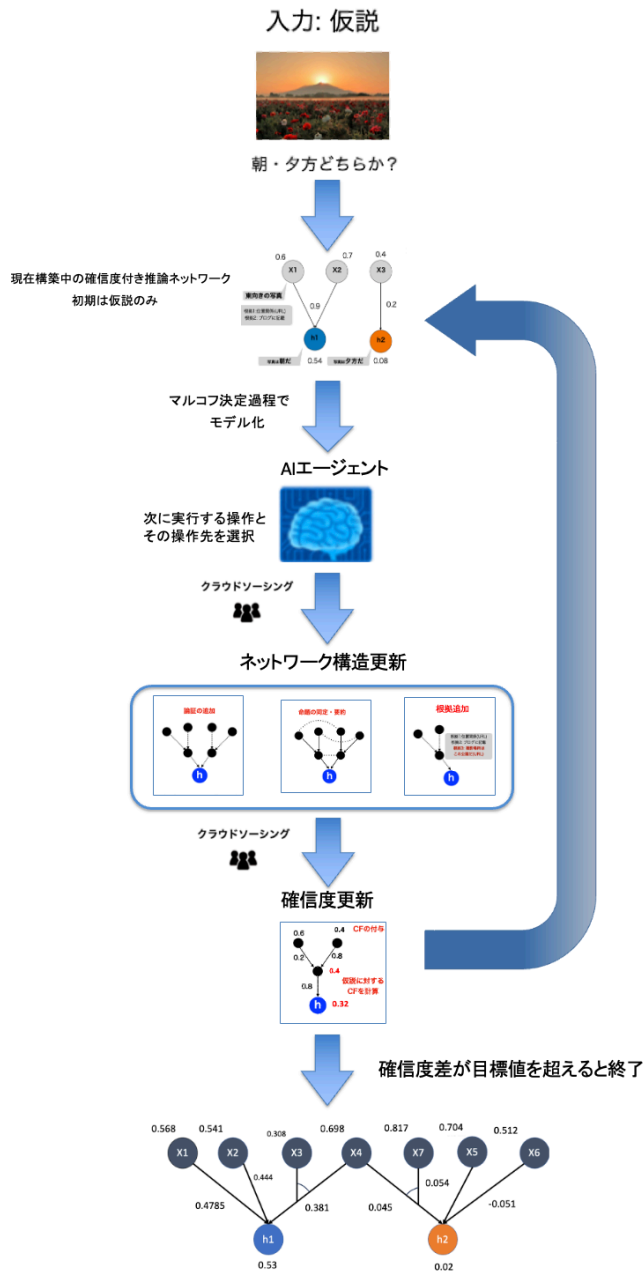


図1 提案手法の全体像

- ていたのに対し、本研究ではネットワーク構築の段階から強化学習を導入する枠組みを提案した。具体的には、構築中のネットワークを状態、構築操作（論証追加・命題統合・根拠追加）およびその対象を行動とする階層的マルコフ決定過程として定式化し、3つの出力ヘッドを持つAIエージェントによって構築過程を最適化する枠組みを考案した。
2. シミュレーション実験による有効性の検証：提案手法を実装し、木田らが構築した推論ネットワークデータを用いてシミュレーション実験を行った。その結果、木田らの手法と同等の確信度差を達成する推論ネットワークを約60%低いコストで構築できることを実証した。

2 関連研究

2.1 クラウドソーシング

クラウドソーシングは、インターネットを介して多くの人々から情報を収集し、問題解決に役立てる方法として広く利用されている。クラウドソーシングを通じた問題解決の手法は、特にファクトチェックの分野で注目されている。

ただし、クラウドソーシングを用いた既存のファクトチェック研究では、ワーカが真偽の判断を行うのみにとどまり、判断の根拠となる情報を体系的に収集する取り組みは限定的である。こうした問題から、根拠情報の収集と評価を反復的に行う手法[1]や、マルチステップ検索とLLM（大規模言語モデル）を活用したフェイクニュースの検出などにも使われている。

2.2 推論ネットワーク

MYCIN[10]は、医療診断支援を目的とした代表的なエキスパートシステムである。専門家の知識や経験をルールベースで記述し、確信度という概念を導入することで、不確実性を伴う判断を支援する枠組みを提供した。MYCINにおける推論ネットワークは、ルール間の関係性を体系的に表現し、確信度を伴った結論の導出を可能にする。MYCIN以降、様々な研究がルールベース推論モデルの拡張や改良を行い、多様なドメインへの適用が試みられてきた。しかしながら、こうしたエキスパートシステムの枠組みにクラウドソーシングを統合するアプローチは、従来の研究ではほとんど見られない。

木田らは、MYCINの確信度付き推論モデルに着想を得て、クラウドソーシングを通じた知識獲得と評価のプロセスを組み込んだ確信度付き推論ネットワーク構築手法を提案した[7]。クラウドを活用した意思決定支援に関する先行研究[6]は存在するものの、これらは主に意見の集約に焦点を当てており、説明や根拠の明示的な収集には至っていない。また、こうした研究は最終的な意思決定の支援を主目的としており、説明性の向上を重視する観点とは異なる。木田らの手法は、クラウドソーシングを用いて推論ネットワークを構築することで、評価とその根拠を同時に得られる枠組みを実現した。

本研究では、木田らの確信度付き推論ネットワーク構築手法を基盤としつつ、構築過程そのものに強化学習を導入することで、より効率的なネットワーク構築を目指す。

2.3 強化学習

リソース配分の最適化は、クラウドソーシング研究における主要なトピックの一つとして位置づけられている。先行研究では、作業者の能力、タスクの難易度、および作業者の応答品質間の関係をベイズネットワークによりモデル化し、各タスクに対して部分観測マルコフ決定過程(POMDP)を定式化することで、最適なリソース配分戦略を導出する手法が提案されている[3]。

本研究では、木田らが用いたDeep Q-Network(DQN)[9]を同様に採用する。ただし、木田らの手法がネットワーク構築後の根拠追加を最適化対象としたのに対し、本研究ではネット

ワーク構築の段階から強化学習を導入し、構築過程におけるコスト配分を最適化する。

3 推論ネットワークの構築

本研究で扱う推論ネットワークは、木田ら [7] の論文に基づき、相反する仮説に対する論証を体系的に表現した構造である。図 2 に推論ネットワークの構成例を示す。ネットワークは命題（ノード）と含意（エッジ）から構成され、それぞれに確からしさを表す確信度（certainty factor）が付与される。各ノードおよびエッジの確信度から、後述する確信度付き推論のルールに従って、最終的に仮説に対する確信度が算出される。この枠組みは、MYCIN [10] で提案されたルールベース推論ネットワークに基づいている。

このような推論ネットワークを採用する理由は、限定された問題領域に対してコストを投下することで、クラウドワークから多様な知見を収集し、説明や根拠を段階的に蓄積することで、柔軟かつ評価可能な枠組みを実現できる点にある。

以下に本研究で用いる用語を定義する。これらの定義は木田ら [7] に準拠する。また、本研究で使用する記号とその説明を表 1 に示す。

- **命題 (Proposition)** : ネットワークの基本単位であり、各ノードとして表現される。
- **含意 (Implication)** : 命題間の論理的関係を表し、エッジとしてネットワーク内の接続を形成する。
- **論証 (Argumentation)** : 特定の命題間の関係を示し、含意エッジとともに構成される。
- **根拠 (Evidence)** : 各命題および含意を支持する情報であり、確信度算出の基盤となる。
- **確信度 (Certainty Factor)** : 各命題および含意の信頼性を定量的に表す指標であり、ネットワーク全体の推論の強度を評価するために用いられる。
- **仮説 (Hypotheses)** : 検証対象となる命題であり、ネットワークを通じてその妥当性が評価される。

表 1 記号一覧

記号	説明
$H = \{h_1, \dots, h_i\} (i \geq 1)$	検証対象となる仮説の集合
$X = \{x_1, \dots, x_k\}$	命題の集合
$R = \{r_1, \dots, r_k\}$	含意の集合
$E = \{e_1, \dots, e_k\}$	ワークから収集した根拠の集合
$A(x, h)$	「 x ならば h 」を示す論証
$S(x), S(r) = \{e_1, \dots, e_j\}$	X, R を支持する根拠の集合
$CF(x), CF(r)$	X, R の確信度

3.1 確信度付き推論の概要

本節では、木田ら [7] に基づき、確信度付き推論の定義と推論規則について述べる。

3.1.1 確信度の定義

確信度 (certainty factor: CF) は、ルールや事実に対する信頼性を数値化した尺度である。その値は-1 から 1 の範囲をとり、正の値は肯定的な信念を、負の値は否定的な信念を表す。絶対

値が 1 に近いほど信念が強く、0 に近いほど信念が弱いことを意味する。本研究では、薦田による確信度の解釈表 [12] に基づき、ワークの回答を-1 から 1 の範囲に変換して確信度として付与する。表 2 に各確信度の値とその解釈を示す。なお、ワークが「分からない」と回答した場合は、確信度として 0 を割り当てる。

表 2 確信度の値の解釈

確信度の値	解釈
-1	絶対に…でない (definitely not)
-0.8	ほぼ確実に…でない (almost certainly not)
-0.6	たぶん…でない (probably not)
-0.4	…でないかもしれない (maybe not)
-0.2 ~ 0.2	分からない (unknown)
0.4	…かもしれない (maybe)
0.6	たぶん…(probably)
0.8	ほぼ確実に…(almost certainly)
1	絶対に…(definitely)

3.2 推論規則

命題 X および含意 R の確信度に基づいて結論 Y の確信度を算出し、ネットワーク全体に確信度を伝播させることで、最終的な結論の確信度を導出する。以下に推論規則の主要な処理について述べる。

3.2.1 単純論証

単一の前提条件からなる単純論証「 $X \Rightarrow Y$ 」において、結論 Y の確信度 $CF(Y, \{X\})$ は以下の式で定義される。

$$CF(Y, \{X\}) = CF(X) \times CF(R) \quad (1)$$

3.2.2 結合論証

前提条件が論理積 ($X_1 \wedge X_2 \wedge \dots$) または論理和 ($X_1 \vee X_2 \vee \dots$) で構成される場合、結論 Y の確信度はそれぞれ以下の式により算出される。

$$CF(Y, \{X_1, X_2, \dots\}) = \min_i(CF(X_i)) \times CF(R) \quad (\text{論理積}) \quad (2)$$

$$CF(Y, \{X_1, X_2, \dots\}) = \max_i(CF(X_i)) \times CF(R) \quad (\text{論理和}) \quad (3)$$

論理積と論理和が混在する場合は、上記の方法を組み合わせて確信度を導出する。例えば、

$$(X_1 \vee X_2) \wedge (X_3 \vee X_4) \wedge X_5 \Rightarrow Y$$

の場合、 Y の確信度は以下のように算出される。

$$\min(\max(CF(X_1), CF(X_2)), \max(CF(X_3), CF(X_4)), CF(X_5)) \times CF(R)$$

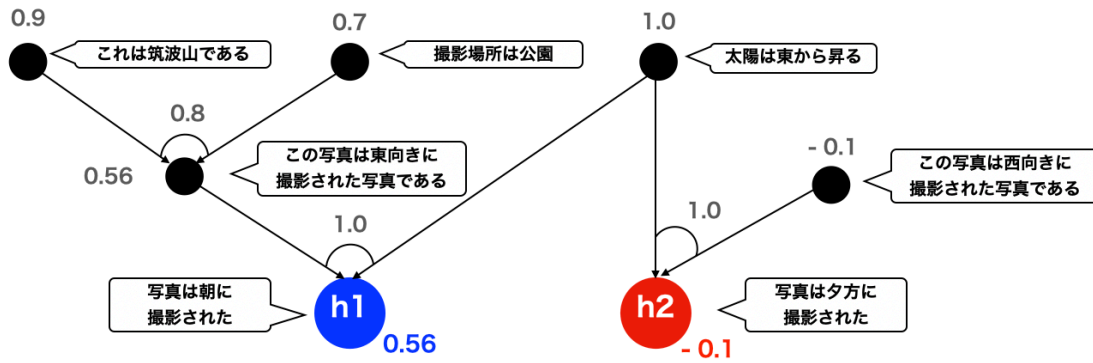


図2 推論ネットワークの基本構造 (木田ら [7] より引用)

3.2.3 合流論証

複数のルールから同一の結論が導出される場合、各ルールの確信度を合成し、総合的な確信度 CF_{COMB} を求める。例えば、ルール「 $X_1 \Rightarrow Y$ 」から導かれる確信度を CF_1 、ルール「 $X_2 \Rightarrow Y$ 」から導かれる確信度を CF_2 とすると、合成確信度は以下の式で算出される。

$$CF_{COMB} = \begin{cases} CF_1 + CF_2(1 - CF_1) & (CF_1, CF_2 > 0) \\ \frac{CF_1 + CF_2}{1 - \min(|CF_1|, |CF_2|)} & (CF_1 CF_2 < 0) \\ CF_1 + CF_2(1 + CF_1) & (CF_1, CF_2 < 0) \end{cases} \quad (4)$$

両ルールの確信度が同符号の場合、より高い確信度が強調される形で合成される。異符号の場合は、両者が相殺される形となる。3つ以上のルールが存在する場合は、逐次的に合成を適用し、最終的な確信度を算出する。

4 強化学習を用いた推論ネットワーク構築

本研究では、木田ら [7] が提案した EMV ワークフローを基盤とし、構築過程における各操作の選択を強化学習によって最適化する手法を提案する。木田らの手法では、推論ネットワーク構築後に根拠追加先のリソース配分を最適化していたのに対し、本研究では構築段階から強化学習を導入することで、より効率的に確信度の高い推論ネットワークを構築することを目指す。

具体的には、構築中のネットワーク（初期は仮説に対応するノードのみ）を状態、次の構築操作を行動とするマルコフ決定過程 (MDP) としてモデル化し、各状態における操作先を強化学習によって最適化する。行動として選択可能な構築操作は以下の3種類である。

- **論証の追加 (Expansion)** : 木田らの Expansion フェーズに相当し、仮説を支持する論証を収集してネットワークに追加する操作。クラウドソーシングを用いて、ワーカーに対して、指定された仮説への論証をインターネットなど使い収集してもらう。これらを既存の推論ネットワークに追加する。
- **命題の統合 (Merge)** : 木田らの Merge フェーズに相当し、

重複する命題を統合してネットワークの冗長性を排除する操作。

- **根拠の追加 (Support)** : 命題および含意を補強するための根拠 (インターネットを用いて収集した URL 等) を追加する操作。これは木田らが EMV ワークフロー構築後に最適化を行った操作に相当する。

各操作の実行後、確信度更新 (木田らの Verification フェーズに相当) を行い、推論規則に基づいて仮説の確信度を再計算する。

4.1 ワークフローの概要

本手法のワークフローは以下の手順で構成される。

1. **初期化** : 検証対象となる仮説に対応するノードのみで構成される初期ネットワークを生成する。
2. **状態観測** : 現在のネットワーク構造を状態として観測する。
3. **行動選択** : AI エージェントが、次に実行する操作 (Expansion・Merge・Support) およびその対象を選択する。
4. **操作実行** : 選択された操作を実行し、ネットワーク構造を更新する。
5. **確信度更新** : 更新されたネットワークに対して確信度を再計算する (Verification に相当)。
6. **終了判定** : 終了条件 (コスト上限到達、確信度収束など) を満たすまで手順 2~5 を繰り返す。

4.2 各操作の詳細

4.2.1 命題の追加 (Expansion)

木田らの Expansion フェーズに相当する操作である。仮説を支持する論証を収集し、推論ネットワークに追加する。クラウドワーカーに対して、AI エージェントから指定された仮説に対する論証の提供を依頼する。収集された論証は、命題ノードおよび含意エッジとして既存のネットワークに追加される。

4.2.2 命題の統合 (Merge)

木田らの Merge フェーズに相当する操作である。推論ネットワーク内の重複する命題を統合する。意味的に重複する命題を検出し、同定・要約を行うことで、ネットワーク内の冗長性を排除し、各命題が独立した意味を持つ構造を構築する。

4.2.3 根拠の追加 (Support)

命題 X および含意 R を補強するための根拠 (インターネットを用いて収集した URL 等) をクラウドワークに収集してもらい、追加する操作である。これは木田らが EMV ワークフローによるネットワーク構築後に、強化学習を用いて最適化を行った操作に相当する。本研究では、この操作をネットワーク構築段階から行動選択の対象として組み込むことで、構築過程全体を通じた最適化を実現する。

4.2.4 確信度更新 (Verification)

各操作の実行後に行われる処理であり、木田らの Verification フェーズに相当する。命題 X および含意 R について、それぞれの根拠 $S(X), S(R)$ を踏まえた確信度スコアをクラウドソーシングを通じて算出し、推論規則に基づき仮説の確信度を更新する。

5 マルコフ決定過程 (MDP) によるモデル化

本節では、推論ネットワーク構築過程を MDP としてモデル化する。木田らの手法 [7] では、構築済みネットワークに対する根拠追加のみを行動として定義していたのに対し、本研究では構築操作 (論証追加・命題統合・根拠追加) を含む行動空間を定義し、構築過程全体を最適化の対象とする。

5.1 状態空間

状態空間は、構築中の推論ネットワークの構造を表現する。具体的には、グローバル情報と各ノードの情報から構成される。

5.1.1 グローバル情報

推論ネットワーク全体の状態を表す以下の情報を含む。

- 各仮説の確信度 ($CF_{h_1}, CF_{h_2}, \dots$)
- 各仮説に属するノード数

5.1.2 ノード情報

各ノード (命題および含意) に対して、以下の情報を保持する。

- 確信度の平均値 \bar{x}_i
- 確信度の分散 $\text{Var}(x_i)$
- 根拠数 $|S(x_i)|$
- ノードタイプ (命題または含意)
- アクティブ状態 (ネットワークに追加済みか否か)

これらをまとめて、時刻 t における状態 \mathbf{s}_t を以下のように定義する。

$$\mathbf{s}_t = \left\{ \begin{array}{l} \text{グローバル} : (CF_{h_1}(t), CF_{h_2}(t), N_{\text{active}}^{h_1}(t), N_{\text{active}}^{h_2}(t)), \\ \text{ノード} : (\bar{x}_i(t), \text{Var}(x_i(t)), |S(x_i)|, \text{type}_i, \text{active}_i) \\ \forall i \in \{1, \dots, N_{\text{max}}\} \end{array} \right\}$$

ここで、 N_{max} はネットワークが保持可能な最大ノード数である。

5.2 行動空間

本研究では、推論ネットワークの構築操作を行動空間として定義する。行動選択は 2 段階 (Stage1 および Stage2) で構成される。

5.2.1 Stage1: 行動タイプ選択

Stage1 では、以下の 3 種類の行動タイプから 1 つを選択する。

- **論証追加 (Expansion)** : 仮説を支持する新たな論証をネットワークに追加する。
- **命題統合 (Merge)** : ネットワーク内の重複する命題を統合する。
- **根拠追加 (Support)** : 既存の命題または含意に対して根拠を追加する。

Stage1 の行動空間は以下のように定義される。

$$A_{\text{Stage1}} = \{\text{Expansion, Merge, Support}\}, \quad |A_{\text{Stage1}}| = 3$$

5.2.2 Stage2: ターゲット選択

Stage2 では、Stage1 で選択した行動タイプに応じて、具体的な操作対象先を選択する。

- **論証追加時**: 論証を追加する対象の仮説を選択する。

$$A_{\text{Stage2}}^{\text{Expansion}} = \{h_1, h_2, \dots, h_k\}, \quad |A_{\text{Stage2}}^{\text{Expansion}}| = k$$

- **命題統合時**: ターゲット選択は不要であり、ネットワーク内の全ての重複命題を自動的に統合する。
- **根拠追加時**: 根拠を追加する対象のノード (命題または含意) を選択する。

$$A_{\text{Stage2}}^{\text{Support}} = \{x_1, x_2, \dots, x_n \mid \text{active}(x_i) = 1\}$$

このように、行動 a は Stage1 の行動タイプ $a_1 \in A_{\text{Stage1}}$ と Stage2 のターゲット $a_2 \in A_{\text{Stage2}}^{a_1}$ の組として表現される。

$$a = (a_1, a_2)$$

5.3 報酬関数

本研究で扱う問題は、相反する複数の仮説に対する確信度を適切に導出することである。理想的には、正しい仮説の確信度は +1 (絶対に正しい) に、誤った仮説の確信度は -1 (絶対に誤り) に収束すべきである。

木田らの報酬設計 [7] に基づき、相反する 2 つの仮説を扱う場合、それぞれの確信度の差を報酬として定義する。仮説 h_1 を正しい仮説、仮説 h_2 を誤った仮説とすると、報酬 r は以下のように定義される。

$$r = CF(h_1) - CF(h_2)$$

この報酬設計により、正しい仮説の確信度を高め、誤った仮説の確信度を低下させる行動が促進される。

5.4 状態遷移

行動を実行することで、推論ネットワークの構造および確信度が更新され、次状態へと遷移する。状態遷移は選択された行動タイプに応じて異なる。

5.4.1 論証追加時の状態遷移

選択された仮説に対して新たな論証（命題ノードおよび含意エッジ）がネットワークに追加され、対応するノードのアクティブ状態が更新される。

5.4.2 命題統合時の状態遷移

重複する命題が統合され、ネットワーク構造が簡素化される。統合された命題の根拠は集約され、確信度が再計算される。

5.4.3 根拠追加時の状態遷移

選択されたノードに対して根拠が追加され、木田ら [7] が提案したベイズ更新に基づいて確信度が更新される。以下にその手順を示す。

1. 証拠数による減衰

既存の証拠数を N 、証拠の基本強度を S とし、減衰した証拠の強さ \tilde{S} を以下の式で定義する。

$$\tilde{S} = \frac{S}{(1+N)^\gamma}$$

ここで、 γ は減衰の度合いを制御するパラメータである。

2. 目標値と更新方向の設定

現在の確信度 \bar{x} に基づき、目標値 T と更新方向 D を以下のように定める。

$$T = \begin{cases} +1, & \text{if } \bar{x} > 0.2, \\ -1, & \text{if } \bar{x} < -0.2, \\ 0, & \text{if } -0.2 \leq \bar{x} \leq 0.2. \end{cases}$$

$$D = -\text{sign}(\bar{x} - T)$$

3. ベイズ更新による確信度の更新

尤度平均 $\mu_L = \bar{x} + \tilde{S} \cdot D$ を計算し、事前分布と尤度分布を用いてベイズ更新を行う。事後分布の平均 μ_{post} を新たな確信度として、範囲を $[-1, 1]$ にクリップする。

$$\bar{x} \leftarrow \text{clip}(\mu_{\text{post}}, -1, +1)$$

5.5 終了条件

以下のいずれかの条件を満たした場合、エピソードを終了する。

- 最大ステップ数に到達
- 確信度が収束（変化量が閾値以下）
- コスト上限に到達

6 AI エージェントによるネットワーク構築最適化実装の設計

本節では、本研究で提案する AI エージェントの設計について述べる。本研究における AI エージェントとは、Deep Q-Network (DQN) を基盤とした強化学習エージェントであり、行動タイプ

選択・仮説選択・ノード選択の 3 つの出力ヘッドを持つニューラルネットワークによって構成される。この AI エージェントは、推論ネットワークの状態を観測し、次に実行すべき構築操作とその対象を自律的に決定する。

DQN は Q 学習をニューラルネットワークで近似する深層強化学習手法であり、Mnih ら [9] によって提案された。従来の Q 学習では、状態空間や行動空間の次元が増加するにつれて Q 値テーブルの管理が困難となる課題があった。DQN を用いることで、高次元の状態空間においても Q 関数を効率的に近似することが可能となる。さらに、Double DQN [5] や Dueling DQN [11] といった拡張手法への発展も容易であり、今後の性能向上が期待できるため、本研究では DQN を採用した。

本研究では、前節で定式化した MDP に基づき、AI エージェントが報酬を最大化する方策を学習する。木田らの手法 [7] では単一の行動空間（根拠追加先の選択）を対象としていたのに対し、本研究では複数の行動空間（行動タイプの選択およびターゲットの選択）を扱うため、複数の出力ヘッドを持つアーキテクチャを採用する。

6.1 環境設計

• 観測空間：

観測は、推論ネットワーク全体の状態を表すグローバル情報と、各ノードの情報を結合した 1 次元特徴量ベクトルとして表現される。グローバル情報には各仮説の確信度およびアクティブノード数が含まれる。各ノードの情報は、確信度平均、確信度分散、根拠数、ノードタイプ、アクティブ状態から構成される。

• 行動空間：

行動は 2 段階の階層的構造として定義される。Stage1 では行動タイプ（論証追加・命題統合・根拠追加）を選択し、Stage2 では選択された行動タイプに応じたターゲット（仮説またはノード）を選択する。この階層的設計により、AI エージェントは構築操作の種類と対象を段階的に決定する。

• 報酬関数：

報酬は、相反する 2 つの仮説の確信度差として定義される。具体的には、正しい仮説の確信度から誤った仮説の確信度を減じた値を報酬とする。この報酬設計により、AI エージェントは正しい仮説の確信度を正方向へ、誤った仮説の確信度を負方向へ誘導する行動方策を学習する。

• 状態遷移：

選択された行動タイプに応じて状態遷移が実行される。論証追加時には新たなノードおよびエッジがネットワークに追加され、命題統合時には重複ノードが統合される。根拠追加時には、木田ら [7] が提案したベイズ更新手法に基づいて確信度が更新される。各操作の実行後、推論規則に従って仮説の確信度が再計算される。

• エピソード終了条件：

規定ステップ数への到達、または確信度の収束をエピソード終了条件とする。なお、無効な行動はステップ数としてカウントせず、有効な行動のみがステップ数を消費

する。

6.2 AI エージェントのアーキテクチャ

本研究では、階層的行動空間に対応するため、共通の特徴抽出層と複数の出力ヘッドから構成される AI エージェントのアーキテクチャを採用する。

- **共通特徴抽出層：**

入力層には観測ベクトルを入力し、複数層の全結合層 (ReLU 活性化関数付き) を通じて状態の特徴表現を抽出する。この特徴表現は後続の各出力ヘッドで共有される。過学習を防止するため、各層には Dropout を適用する。

- **Stage1: 行動タイプヘッド：**

共通特徴抽出層の出力を入力とし、3 種類の行動タイプ (論証追加・命題統合・根拠追加) に対応する Q 値を出力する。

- **Stage2: ターゲット選択ヘッド：**

行動タイプに応じて 2 種類のヘッドを使い分ける。論証追加時には仮説選択ヘッドを用いて対象仮説の Q 値を出力し、根拠追加時にはノード選択ヘッドを用いて対象ノードの Q 値を出力する。命題統合時にはターゲット選択を行わない。

6.3 AI エージェントの学習

AI エージェントは、経験再生バッファを活用して過去の経験データを再利用しながら、状態から行動価値 (Q 値) を予測するニューラルネットワークを学習する。学習の安定性を確保するため、ターゲットネットワークをはじめとする DQN 固有の安定化手法を導入する。

- **ニューラルネットワークによる Q 関数の近似：**

入力層には観測ベクトルを受け取り、複数層の全結合層 (非線形活性化関数付き) を経由して特徴を抽出する。出力層には、各出力ヘッドに対応するニューロンを配置し、それぞれが特定の行動に対する Q 値を表現する。

- **方策 (Policy)：**

ϵ -greedy 法に基づいて Q 値から行動を選択する。学習初期には探索率 ϵ を高く設定してランダムな探索を促進し、学習の進行に伴って探索率を減衰させることで、より高い Q 値を持つ行動を優先的に選択するよう誘導する。

- **ターゲットネットワークと経験再生バッファ：**

ターゲットネットワークを導入することで、ポリシーネットワークの学習を安定化させる。一定エピソードごとにポリシーネットワークの重みをターゲットネットワークにコピーし、学習対象の急激な変動を抑制する。経験再生バッファには、過去の状態・行動・報酬・次状態を蓄積し、ランダムにサンプリングしたミニバッチを用いて学習を行う。これにより、サンプル間の時間的相関を低減し、学習の安定性を向上させる。

以上の設計により、MDP として定式化された推論ネットワーク構築過程に対して、報酬最大化方策を学習する枠組みを構築する。これにより、構築操作の選択が AI エージェントによっ

て最適化され、効率的に確信度の高い推論ネットワークを構築することが可能となる。

7 実験

本章では、木田ら [7] が構築した「時間帯判定タスク」の推論ネットワークデータを用いて、提案手法のシミュレーション実験を行った結果を示す。本実験はシミュレーション環境での検証であり、今後は提案手法を用いて実際にクラウドソーシングプラットフォーム上で実験を行う予定である。

7.1 実験の目的

本実験の目的は、リサーチクエスト「木田らの手法で得られた推論ネットワークと同等の確信度差を達成する推論ネットワークを、構築段階から強化学習を導入することでどの程度低コストで構築できるのか」に回答することである。具体的には、木田らの手法が達成した確信度差 0.51 を目標値として設定し、提案手法がこの目標値に到達するまでに要するコストを計測する。両手法のコストを比較することで、構築段階から強化学習を導入することによるコスト削減効果を定量的に評価する。

7.2 タスクの概要

本研究では、木田ら [7] が使用した「時間帯判定タスク」を対象として実験を行った。このタスクは、朝焼けの写真 (図 3) に対し、「写真は朝に撮影されたものである」「写真は夕方に撮影されたものである」のいずれかを判定するものである。木田らの予備実験では、「写真は夕方に撮影されたものである」という誤った回答をしたワーカが 67% (100 人中 67 人) を占めており、事前情報がない場合にワーカの回答が割れる問題設定となっている。



図 3 朝焼けの写真

7.3 シミュレーション環境

本実験では、木田ら [7] がクラウドソーシングを通じて構築した推論ネットワークのデータを基に、提案手法の有効性を検証するシミュレーション環境を構築した。結果の信頼性を確保するため、本シミュレーションは 100 回の独立した試行を行い、その平均値および分布を評価指標として用いた。シミュレーション環境の設定を表 3 に示す。

表3 シミュレーション環境の設定

項目	設定値
最大ノード数	80 (朝 40、夕方 40)
最大ステップ数	100
シミュレーション回数	100 回
目標	朝 CF 値 - 夕方 CF 値 ≥ 0.51

7.4 コスト設定

木田ら [7] の実験結果と比較可能とするため、各操作に対するコストを表 4 のように設定した。

表4 操作別コスト設定

操作	コスト	備考
論証追加 (Expansion)	100 円	1 回あたり
命題統合 (Merge)	0 円	-
根拠追加 (Support)	50 円	1 回あたり

また、各操作後の確信度更新 (Verification) に要するコストは、直前の操作に応じて表 5 のように設定した。

表5 Verification コスト設定

直前の操作	コスト	備考
論証追加後	250 円	25 人による投票
命題統合後	$100 \times N_{active}$ 円	N_{active} : アクティブノード数
根拠追加後	250 円	25 人による投票

そして木田らが実施したクラウドソーシングタスクの金額の詳細を表 6 に示す。なお根拠追加 (Support) は本研究の設定と同じ、一回あたり 50 円とする。

フェーズ	タスク	プラットフォーム	人数	料金	タスク数
Expansion	時間帯	Lancers	30	100 円	1
Verification	時間帯	Yahoo ¹	150	10 円	11

表6 フェーズ別タスク料金表

7.5 各操作の詳細

本シミュレーションでは、木田ら [7] が実際のクラウドソーシングで収集したデータを基に、以下の操作をシミュレートした。

7.5.1 論証追加 (Expansion)

本操作は、木田らの Expansion フェーズに相当し、仮説を支持する新たな論証をネットワークに追加する。シミュレーションでは、木田らが Lancers を通じて収集した論証データ (命題および含意) をサンプルデータとして使用した。

- **対象選択:** 朝の仮説または夕方の仮説を選択する。
- **データ追加:** サンプルデータから命題 (X) と対応する含意 (R) をランダムに選択し、観測空間に追加する。
- **初期値設定:** 追加された命題および含意には、サンプルデータに基づく初期確信度、分散、根拠数が設定される。

7.5.2 命題統合 (Merge)

本操作は、木田らの Merge フェーズに相当し、推論ネットワーク内の重複する命題を統合する。統合処理は木田ら [7] と同様の手順で実施した。

1. **重複検出:** 同一コンテンツ (content) を持つアクティブな命題および含意を検索する。
2. **グループ化:** コンテンツごとにノードをグループ化する。
3. **統計計算:** グループ内のノードに対して、以下の統計量を計算する。
 - 確信度の平均値: $\bar{CF} = \frac{1}{n} \sum_{i=1}^n CF_i$
 - 分散の平均値: $\bar{Var} = \frac{1}{n} \sum_{i=1}^n Var_i$
 - 根拠数の合計: $N_{total} = \sum_{i=1}^n N_i$
4. **非アクティブ化:** 元の命題を非アクティブ状態に変更する。
5. **新規作成:** 統合された統計量を持つ新たな命題を生成する。

7.5.3 根拠追加 (Support)

本操作は、木田らが EMV ワークフロー構築後に最適化を行った操作に相当し、既存の命題または含意に対して根拠を追加する。根拠追加後の確信度更新は、木田ら [7] が提案したベイズ更新手法に基づいて実施した。

1. **対象選択:** アクティブなノード (命題または含意) から根拠追加先を選択する。
2. **根拠数更新:** 選択されたノードの根拠数を 1 増加させる。
3. **ベイズ更新:** 以下の手順で確信度を更新する。
 - a. 証拠数による減衰: $\tilde{S} = \frac{S}{(1+N)^\gamma}$
 - b. 目標値と更新方向の設定
 - c. 尤度平均の計算: $\mu_L = \bar{x} + \tilde{S} \cdot D$
 - d. 事後分布の計算とクリッピング: $\bar{x} \leftarrow \text{clip}(\mu_{\text{post}}, -1, +1)$

ベイズ更新に使用するパラメータを表 7 に示す。これらのパラメータ (S, γ, σ_L) は、木田ら [7] が Nelder-Mead 法を用いて最適化したものである。木田らの研究では、時間帯判定タスクと国判定タスクの両方のデータを用いてパラメータ最適化を行った結果、両タスクで近似した値が得られたことが報告されている。この結果は、異なるトピックのタスクに対しても同様のパラメータが適用可能である可能性を示唆している。

本シミュレーションでは、木田らが最適化したパラメータをそのまま使用している。ただし、これらのパラメータは木田らの手法 (EMV ワークフロー構築後の根拠追加最適化) に対して最適化されたものであり、本研究の提案手法 (構築段階からの強化学習導入) に対しても最適であるかは検証されていない。提案手法に適したパラメータの探索については、今後の研究課題である。

表7 ベイズ更新パラメータ

パラメータ	値	説明
S (evidence_strength)	0.0865	根拠の基本強度
γ (gamma)	3.8829	減衰パラメータ
σ_L (likelihood_std)	0.0187	尤度の標準偏差

7.5.4 確信度更新 (Verification)

各操作の実行後、推論規則に基づいて仮説ノードの確信度を

1: Yahoo!クラウドソーシング

再計算する。これは木田らの Verification フェーズに相当する。シミュレーションでは、各命題および含意の確信度から、単純論証・結合論証・合流論証のルールに従って仮説の確信度を算出する。

7.6 評価指標

提案手法の評価には、以下の指標を用いた。

- **確信度差 (CF 差)**: 朝の仮説の確信度から夕方の仮説の確信度を減じた値。正しい仮説 (朝) の確信度が高く、誤った仮説 (夕方) の確信度が低いほど大きな値となる。
- **達成コスト**: 目標 (CF 差 ≥ 0.51) を達成するまでに要した総コスト。
- **平均ステップ数**: 目標達成までに要した平均行動数。

8 結果

本章では、「時間帯判定タスク」において、提案手法である AI エージェントを用いたシミュレーション実験の結果を示す。木田ら [7] の手法をベースラインとして、コストおよび確信度の観点から比較を行う。

8.1 学習後の AI エージェントの行動選択

学習後の AI エージェントがどのような行動を選択し、推論ネットワークを構築していくかを分析した。図 4 に、ステップごとの行動選択比率の推移と確信度の推移を示す。

8.1.1 行動選択の傾向

図 4 の上段に示す行動比率の推移から、以下の傾向が確認された。

- **初期段階 (0~5 ステップ目)**: 論証追加 (Expansion) が非常に高い確率で選択されている。この段階では、推論ネットワークの基盤となる命題および含意を積極的に追加し、ネットワーク構造を構築していることがわかる。
- **中盤以降 (6 ステップ目~)**: 論証追加の選択確率が急激に低下し、代わりに根拠追加 (Support) の選択確率が上昇している。これは、ネットワーク構造がある程度構築された後は、既存のノードに対する根拠追加によって確信度を向上させる戦略に切り替えていることを示している。
- **命題統合 (Merge)**: 全ステップを通じて一定の低い確率で選択されている。これは、重複命題が発生した場合のみ統合操作が有効となるためと考えられる。

8.1.2 確信度の推移

図 4 の中段および下段に示す確信度の推移から、以下の傾向が確認された。

- **初期段階 (0~5 ステップ目)**: 確信度差 (朝 CF - 夕方 CF) が急激に上昇している。この段階は論証追加が主に選択されており、ネットワーク構造の構築によって確信度が大きく変化していることがわかる。
- **中盤以降 (6 ステップ目~)**: 確信度差は緩やかな上昇傾向を示している。この段階では根拠追加が主に選択されており、既存ノードの確信度を段階的に向上させていることがわかる。

この結果から、学習後の AI エージェントは「初期段階でネットワーク構造を構築し、その後は根拠追加によって確信度を向上させる」という効率的な構築戦略を獲得していることが確認された。

8.2 コストの比較

提案手法と木田ら [7] の手法について、時間帯判定タスクにおける推論ネットワーク構築にかかったコストを比較した。図 5 にコストの分布を示し、表 8 にコストの比較結果を示す。

表 8 推論ネットワーク構築コストの比較 (時間帯判定タスク)

手法	平均コスト	削減率
木田らの手法 [7]	21,600 円	-
提案手法 (AI エージェント)	8,572 円	約 60% 削減

表 8 に示すように、提案手法を用いた場合の平均コストは **8,572 円** であり、木田らの手法における平均コスト **21,600 円** と比較して、**約 60% のコスト削減** を達成した。

このコスト削減は、提案手法が構築過程において各操作を適切に選択し、効率的にネットワークを構築できることに起因する。具体的には、以下の要因が考えられる。

- **論証追加の最適化**: 初期段階で集中的に論証追加を行い、必要十分なネットワーク構造を効率的に構築している。
- **根拠追加への切り替え**: ネットワーク構造が構築された後は、コストの低い根拠追加 (50 円/回) を主に選択することで、総コストを抑制している。
- **不要な操作の回避**: 強化学習によって、確信度向上に寄与しない操作を回避する戦略を学習している。

時間帯判定タスクの推論ネットワーク

表 9 に最終的に得られた命題、図 6 に提案手法で作られた推論ネットワークの一例と図 7 に木田らの手法で根拠を全て提示した上での最終的な推論ネットワークを提示する。推論ネットワークにおいて仮説 h1: 「写真は朝に撮影された」、仮説 h2: 「写真は夕方に撮影された」を意味する。最終的に得られた 7 つの命題のうち、X1 から X4 の命題が朝を導く論証であり、X4 から X の 7 命題が夕方を導く論証である。

表 9 推論ネットワークを構成する命題 (時間帯判定タスク)

Proposition_ID	命題
X1	筑波山の背後に低い太陽とポピー花畑がある
X2	朝霧が立ち込めている
X3	撮影場所は小貝川ふれあい公園である
X4	筑波山が写っている
X5	空がオレンジ色で地平線が赤みを帯びている
X6	花が咲いている
X7	撮影場所はフラワーパークである

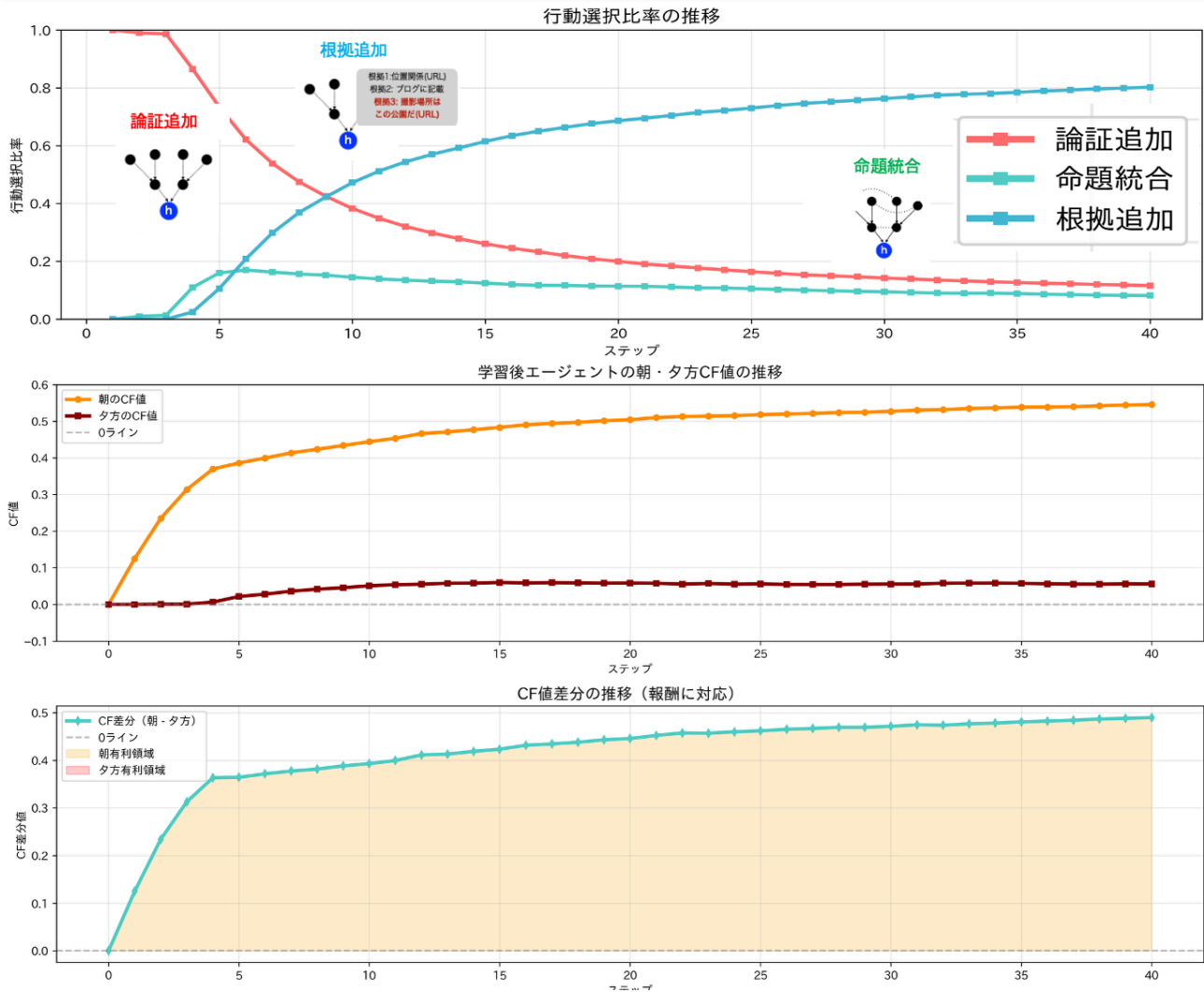


図4 学習後の AI エージェントが選択する行動比率の推移

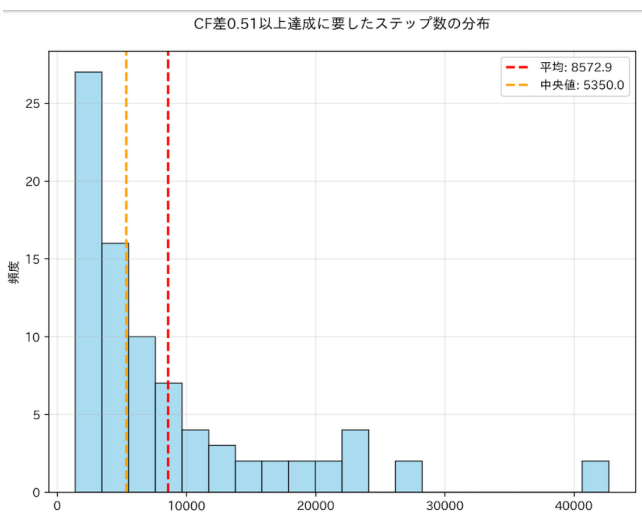


図5 提案手法におけるコストの分布

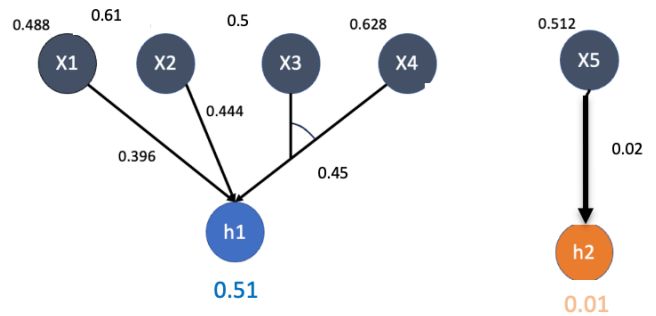


図6 提案手法で得られた推論ネットワークの一例

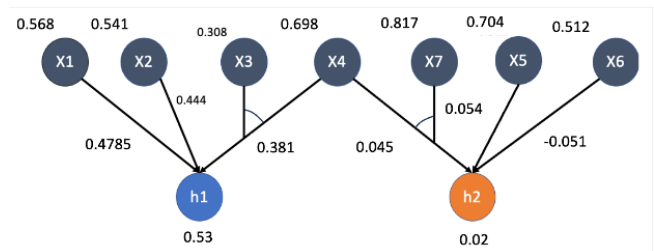


図7 木田らの手法で得られた推論ネットワーク

9 おわりに

本研究では、確信度付き推論ネットワークの構築過程そのものを強化学習により最適化する手法を提案した。木田ら [7] の手法では、推論ネットワーク構築後に根拠追加先のリソース配分を最適化していたのに対し、本研究ではネットワーク構築の段階から強化学習を導入することで、構築操作（論証追加・命題統合・根拠追加）の選択を最適化するアプローチを採用した。具体的には、構築中のネットワークを状態、構築操作およびその対象を行動とする階層的マルコフ決定過程（MDP）としてモデル化し、Deep Q-Network（DQN）に基づく AI エージェントを適用することで、効率的なネットワーク構築戦略の学習を実現した。シミュレーション実験の結果、提案手法は木田らの手法と比較して、効率よく確信度の高い推論ネットワーク構築を得られることが示唆された。

本研究の貢献は以下のとおりである。まず、確信度付き推論ネットワークの構築過程をマルコフ決定過程として定式化し、構築段階から強化学習を導入する枠組みを提案した点である。次に、複数の行動空間（行動タイプ選択およびターゲット選択）を扱う AI エージェントのアーキテクチャを設計し、構築操作の選択を最適化可能であることを示した点である。

しかしながら、本研究にはいくつかの課題も残されている。まず、本研究ではシミュレーション環境での実験にとどまっており、実際のクラウドソーシング環境での有効性は未検証である。シミュレーションでは理想化されたワーク応答を仮定しているが、実環境では応答の質のばらつきやワークの特性が結果に影響を与える可能性がある。また、推論ネットワークの構築において、命題の同定精度や情報の重要性の違いを考慮するなど、さらなる改善の余地がある。

今後の研究としては、以下の点に取り組む予定である。第一に、本研究で提案した手法を用いて、実際にクラウドソーシングプラットフォーム上で実験を行い、シミュレーション結果との比較を通じて実環境における有効性を検証する。これにより、シミュレーションと実環境の差異を明らかにし、実用的な手法への改良を図る。第二に、本手法を異なるドメインや問題設定に適用し、その汎用性を評価する。これにより、本手法がどのようなタスクにおいて有効であるかを分析し、適用範囲を明確化する。第三に、ワークの特性（信頼性、専門性など）を考慮した推論ネットワークの設計に取り組み、より高品質な推論結

果を得られる枠組みの構築を目指す。

謝 辞

本研究の一部は JSPS 科研費 (22H00508, 22H00011, 23K28095, 25K21807, 25K03415) と JST CREST(Grant Number JPMJCR22M2) の支援を受けたものである。ここに謝意を示す。

文 献

- [1] David La Barbera, Eddy Maddalena, Michael Soprano, Kevin Roitero, Gianluca Demartini, Davide Ceolin, Damiano Spina, and Stefano Mizzaro. 1. crowdsourced fact-checking: Does it actually work? *Information Processing and Management*, 2024.
- [2] Jay J. Van Bavel, Helena Gil-Buitrago, Irene Cisma, Rosamunde Cecilie Hendricks, and Daniela Lizarazo Villarreal. 4. scaling crowdsourcing interventions to combat misinformation, 2024.
- [3] Peng Dai, Christopher H. Lin, Mausam, and Daniel S. Weld. Pomdp-based control of workflows for crowdsourcing. *Artificial Intelligence*, Vol. 202, pp. 52–85, 2013.
- [4] William Godel, Zeve Sanderson, Kevin Aslett, Jonathan Nagler, Richard Bonneau, Nathaniel Persily, and Joshua A. Tucker. Moderating with the mob: Evaluating the efficacy of real-time crowdsourced fact-checking. *Journal of Online Trust and Safety*, Vol. 1, No. 1, Oct. 2021.
- [5] Hado van Hasselt, Arthur Guez, and David Silver. Deep reinforcement learning with double q-learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 30, pp. 2094–2100. AAAI Press, 2016.
- [6] Simo Hosio, Jorge Goncalves, and Theodoros Anagnostopoulos. Leveraging wisdom of the crowd for decision support. In *Proceedings of the EWIC/HCI 2016 Conference*, 2016.
- [7] Kai Kida, Hiroyoshi Ito, Nobutaka Suzuki, and Atsuyuki Morishima. 群衆による論証構築のための最適資源配分. 第 17 回データ工学と情報マネジメントに関するフォーラム (DEIM 2025), 2025.
- [8] Santos Espina Mairal, Florencia Bustos, Guillermo Solovey, and Joaquin Navajas. 2. interactive crowdsourcing to fact-check politicians. *Journal of Experimental Psychology: Applied*, 2023.
- [9] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Human-level control through deep reinforcement learning. *Nature*, Vol. 518, No. 7540, pp. 529–533, 2015.
- [10] Edward H. Shortliffe and Bruce G. Buchanan. A model of inexact reasoning in medicine. *Mathematical Biosciences*, Vol. 23, No. 3, pp. 351–379, 1975.
- [11] Ziyu Wang, Tom Schaul, Matteo Hessel, Hado van Hasselt, Marc Lanctot, and Nando de Freitas. Dueling network architectures for deep reinforcement learning. In *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, pp. 1995–2003. PMLR, 2016.
- [12] 安信千津子 薦田憲久. エキスパートシステムの設計と開発. 情報系教科書シリーズ. 昭晃堂, 1997.