

Toulmin モデルのベイジアンネットワーク表現を用いた論証推敲支援システム

宇都 雅輝^{†a)} 鈴木 宏昭^{††} 植野 真臣[†]

Toulmin Model based Argument Elaboration Support System using Bayesian Network Representation

Masaki UTO^{†a)}, Hiroaki SUZUKI^{††}, and Maomi UENO[†]

あらまし 本論では、アカデミックライティングにおける論証の推敲を支援するシステムを開発する。従来の論証推敲支援システムでは、論証の規範モデルとして知られる Toulmin モデルにユーザの論証を当てはめ可視化する支援を行っていることが多い。しかし、論証の主目的である「主張」の正当化のためには、Toulmin モデルへの当てはまりの良さよりも、文章間の因果の強さ、すなわち「論証の強さ」を重視した論証の推敲が重要である。論証の推敲では、論証構成が複雑になったとき、以下の問題が生じると考えられる。1. 「論証の強さ」を全ての文章間について評価することが困難である。2. 論証中の各文章がどの程度正当化できているかの推定が難しい。3. 「主張」の正当化に対して各文章がどのように影響しているかを把握することが困難である。これらの問題を解決するために、本論では、Toulmin モデルのベイジアンネットワーク表現を用いて、1. 論証の強さ、2. 文章の正当性、3. 主張への影響度、という 3 つの指標を算出し、その値に応じて論証改訂のためのアドバイスをフィードバックする論証推敲支援システムを開発する。

キーワード Toulmin モデル、ベイジアンネットワーク、アカデミックライティング、論証推敲支援システム

1. はじめに

情報の効果的な伝達を目指すアカデミックライティングでは [1]、説得的な「論証」の構築が重要課題となる [2]。論証とは「自らの主張を説得的に行うための言語行為 [3]」と定義される。説得的な論証を構築するためには、全ての文章間の因果関係を把握し、客観的かつ反復的に論証を推敲する必要がある。しかし、特に初心者にとって、このような作業は容易ではないことが知られている [4]。

そのため、アカデミックライティングにおける論証の推敲を支援するシステムがこれまでに多数開発されてきた [5]。このようなシステムの多くでは、Toulmin

モデル [6] と呼ばれる論証スキーマに論証を当てはめることを支援する（例えば [4], [7] ~ [10]）。Toulmin モデルとは、論証を「主張」、「データ」、「論拠」、「裏付け」、「限定」、「論駁」の 6 つの小単位に分類し、それぞれの関係を有向グラフとして形式化したものであり [11]、論証の基準として多様な分野で利用され、その妥当性が示されている（例えば [12] ~ [15]）。

Toulmin モデルに基づく論証推敲支援システムとして、館野ら [7] は、文章中の Toulmin モデルの構成要素に対応する箇所をハイライトでき、更にそれを他のユーザと共有して相互評価できるシステム「カラコメ！」を開発している。Brudvik ら [8]、Azilawati ら [9] は、文章に Toulmin モデルの各構成要素に対応するタグ付けを行うことで、それらを階層化して表示するシステム「Voice of Reason」を開発している。また、Chris ら [10] は、論証を Toulmin モデルに当てはめ、ダイアグラムとして可視化できるシステム「Araucaria」を開発している。

これらのシステムでは、論証を Toulmin モデルに表層的に当てはめ可視化する支援に留まっている。し

[†] 電気通信大学大学院 情報システム学研究所, 東京都
Graduate School of Information Systems, The University
of Electro-Communications, 1-5-1 Chofugaoka, Chofu-shi,
Tokyo 182-8585, Japan

^{††} 青山学院大学 教育人間科学部, 東京都
Department of Education, Aoyama Gakuin University, 4-4-
25 Shibuya, Shibuya-ku, Tokyo 150-8366, Japan

a) E-mail: uto_masaki@ai.is.uec.ac.jp

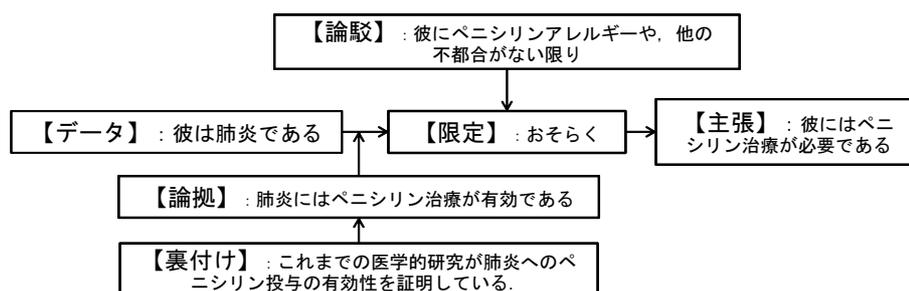


図 1 Toulmin モデル
Fig.1 Toulmin model

かし、論証の主目的である「主張」の正当化のためには、論証の Toulmin モデルへの当てはまりの良さよりも、文章間の因果の強さ、すなわち「論証の強さ」を重視した論証の推敲が重要である。一般の論証では、複数の複数の「論拠」や「データ」、「裏付け」から主張を導くことが多く、論証の構成が複雑化すると、以下の問題が生じると考えられる。1. 「論証の強さ」を全ての文章間について評価することが困難である。2. 論証中の各文章がどの程度正当化できているかの推定が難しい。3. 「主張」の正当化に対して各文章がどのように影響しているかを把握することが困難である。

以上の問題を解決するために、本論では、Toulmin モデルのページアンネットワーク表現を用いた論証推敲支援システムを開発する。ここでは、論証中の文章を確率変数、文章間の因果を条件付確率で表すことで、Toulmin モデルに当てはめて構築された論証をページアンネットワークとして定式化する。確率は主観確率とし、確率値に対応付けられた 6 段階のカテゴリを用いてユーザが評価する。本システムでは、このページアンネットワーク表現を用いて、1. 論証の強さ、2. 文章の正当性、3. 主張への影響度、という論証の特性を表す 3 つの指標（論証特性指標と呼ぶ）を算出し、それらの推定結果に応じて論証改訂のためのアドバイスをフィードバックする。本システムの特徴は以下の通りである。

(1) ページアンネットワークの特殊形のひとつである Noisy or モデル [16] を利用することで、ユーザが評価する主観確率の個数を減少させた上で、全ての論証の強さを推定できる。

(2) 論証中の各文章がどの程度正当化されているかを数値化でき、更に、この値が低い場合、その文章を正当化するための論証改訂方法をアドバイスできる。

(3) Sensitivity Analysis [17], [18] を利用した論証の分析により、「主張」の正当化に対する各文章の影響を定式化できる。

更に、本論では、被験者実験により本システムの有効性を示す。

2. Toulmin モデル

本論では、論証の規範モデルとして Toulmin モデルを採用する。

Toulmin モデルとは、論証を「主張 (Claim)」、「データ (Data)」、「論拠 (Warrant)」、「裏付け (Backing)」、「限定 (Qualifier)」、「論駁 (Rebuttal)」の 6 つの小単位に分類し、それぞれの関係を有向グラフとして形式化したものであり [11]、一般に図 1 のようなダイアグラムで表現される。ここで「主張」とは、正当化されるべき命題であり、「データ」とは、結論を基礎づけるために参照される事実である。「論拠」とは、データと結論を架橋し、その移行を権威づける一般的な仮言的弁明であり、「裏付け」は、論拠に権威と信用を与える言明である。「限定」は、データから結論への移行を正当化する強さを表し、「論駁」は、論拠の一般的正当性が退けられなければならない状況を意味する [19]。

Toulmin モデルの構成要素には論証中の文章が対応付けられ、例えば、図 1 の Toulmin モデルは以下の論証 [20] を表現している。

彼にはペニシリン治療が必要である (主張)。
なぜなら、彼は肺炎であり (データ)、肺炎にはペニシリン治療が有効である (論拠) からだ。これまでの医学的研究が肺炎へのペニシリン治療の有効性を証明している (裏付け)。従って、彼がペニシリンアレルギーであったり、他の不都合な状況にない限り (論駁)、お

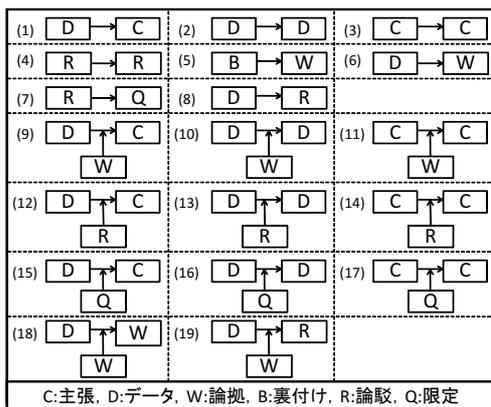


図 2 Toulmin モデルの構成規則
Fig. 2 Formation rules of Toulmin model

そらく (限定) 彼にはペニシリン治療を施すべきであろう。

Toulmin モデルは、日常的な文脈の論理に幅広く適用可能であり、法廷での論証過程の分析 [12] やクリティカルリーディング [13]、合意形成 [14]、意思決定 [15] など多様な分野で利用され、その妥当性が示されている。アカデミックライティングにおける論証支援においても、論証構築のための具体的な基準として利用されている [21]。そこで、本論でも Toulmin モデルを論証の規範として利用し、Toulmin モデルへの論証の当てはめを支援する。このために、ここでは Toulmin モデルの構造を定義する。

Toulmin モデル T は、図 1 に示した 6 つの構成要素を元とする集合 N と構成要素間の支持関係を表す有向辺集合 E により $T = \langle N, E \rangle$ で定義される。ここで構成要素間への有向辺の付与は任意には行えず、一定の規則 (以下、構成規則) に従う必要がある。本論では、この構成規則を定義するために、Toulmin モデルの構成規則に言及した先行研究 [10], [22] ~ [24] から共通する規則を全て抽出した。

抽出した構成規則を図 2 に示す。図 2 の構成規則の組み合わせにより、例えば、図 3 のような Toulmin モデルを表現できる。図 3 の論証は、「主張」に対して、二つの「データ」を提示し、それらの間の論証を支持するために「論拠」や「限定」を追加している。更に、「データ (2)」を補足するために「データ (3)」と「論拠 (3)」が付与されている。この論証は、図 2 の構成規則 (1),(2),(9),(10),(12),(15) により表現できる。

この例のように、Toulmin モデルでは、「主張」に対

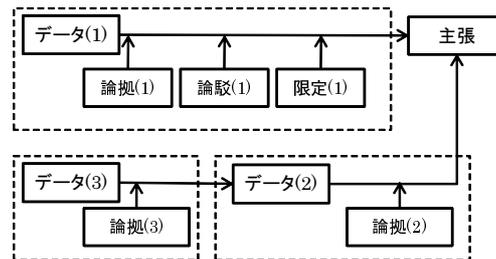


図 3 構成規則に従う Toulmin モデルの例
Fig. 3 An example of Toulmin model

して「データ」や「論拠」などを複数提示することで複雑な構造の論証も表現できる。本論では、図 3 のように「データ」が複数存在するとき、図中の点線で囲んだような局所的な論証を「局所論証」と呼び、以下のように定義する。ある「主張」(または「データ」) S_i に対し R_i 個の「データ」 $S_{d_i}^i = \{S_{d_1}^i, \dots, S_{d_r}^i, \dots, S_{d_{R_i}}^i\}$ が提示されるとき、 $S_{d_r}^i$ から S_i への有向辺を $S_{d_r}^i \rightarrow S_i$ と表す。更に、 $S_{d_r}^i \rightarrow S_i$ に接続する有向辺を持つ構成要素集合を Ψ_{id_r} と表記し、 $S_{d_r}^i \cup \Psi_{id_r}$ を A_{ir} と表す。「局所論証」は構成要素集合 A_{ir} から S_i への論証を表す。

3. Toulmin モデルのベイジアンネットワーク表現

本論では、Toulmin モデルに当てはめて構築された論証をベイジアンネットワークとして表現する。

3.1 Toulmin モデルのベイジアンネットワーク表現

一般に、論証内の文章の正しさは真偽の 2 値で断定できない場合が多く、「客観的に妥当か (以下、正当性)」といった蓋然的な基準で評価される [25]。そこで、本論では、論証中の文章を確率変数 $S_i (i = 1, \dots, N)$ とみなし、文章の正当性を主観確率 $p(S_i = k)$ として表す。ここで、 $k \in \{0, \dots, r\}$ (本論では $r = 1$) は確率変数 S_i の状態値を表し、真偽 (それぞれ 1, 0) の 2 値を取るものとする。更に、文章集合 Π_{S_i} が与えられたときの文章 S_i の正当性は条件付確率 $p(S_i = k | \Pi_{S_i} = j)$ で表す。ここで、 $\Pi_{S_i} = j (j = 1, \dots, q_i)$ は Π_{S_i} が j 番目のパターンを取ることを表す。主観確率は、不確実性を含む主観的な信念や信頼の測度として知られており [26]、正当性のような蓋然的かつ主観的な事象の数量化に適した測度であると考えられる。

他方、このような確率変数間の依存関係を簡潔に表

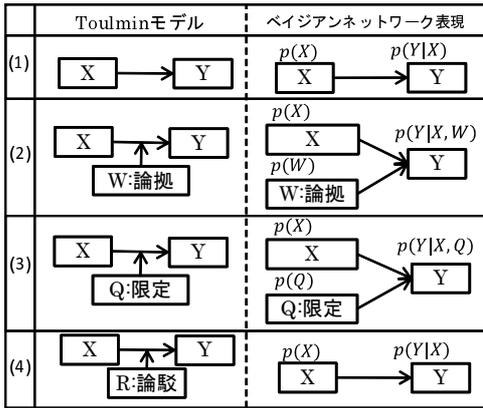


図4 Toulminモデルのベイジアンネットワーク表現
Fig. 4 Bayesian network representation of Toulmin model

現するモデルとして、ベイジアンネットワークが知られている。ベイジアンネットワークとは、次の定義を満たすモデルである。

[定義1] N 個の変数集合 $X = \{x_1, x_2, \dots, x_N\}$ を持つベイジアンネットワークは、以下の $\langle G, p \rangle$ で表現される。

- $G = \langle V, E \rangle$ は、 X に対応する頂点集合 V と、その変数間の確率的依存関係を表現する有向辺集合 E で構成される非循環有向グラフであり、ネットワーク構造と呼ばれる。
- p は、 X についての同時確率分布を表し、各変数 $x_i (i = 1, \dots, N)$ に対応する条件付確率分布 $p(x_i | \Pi_i, G)$ の積として、 $p(x_1, x_2, \dots, x_N) = \prod_{i=1}^N p(x_i | \Pi_i, G)$ で表される。ただし、 Π_i は変数 x_i の親変数集合を示す。

ベイズ統計学では、確率の解釈として、ユーザの知識と客観データを統合的に取り扱える主観確率を用いる。近年のベイジアンネットワークの研究では、データからネットワーク構造 G やパラメータ p を推定する統計的学習アプローチが盛んであるが(例えば[27])、そもそもは専門家の因果知識と主観確率を計測し、エキスパートシステムを構築することから始まっている[28],[29]。

ベイジアンネットワークは、1. 確率変数間の因果関係を表すモデルと解釈でき、2. 主観確率を扱うことができ、3. 任意の変数の確率を求める確率推論を効率よく実行できるといった特徴を持っており[30]、本論の目的に適した手法であると考えられる。

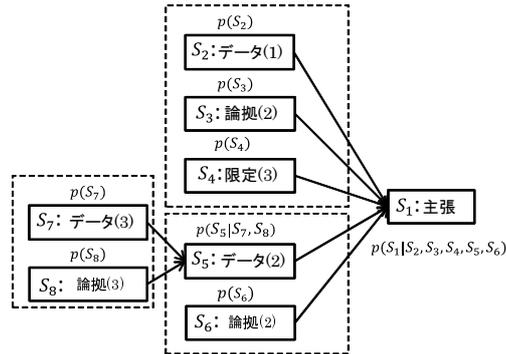


図5 ベイジアンネットワーク表現の例
Fig. 5 An example of transformation into Bayesian network representation

そこで、本論では、Toulminモデルに当てはめて構築された論証をベイジアンネットワークにより定式化する。具体的には、ユーザが構築したToulminモデルをもとにベイジアンネットワークの構造 G を決定し、条件付確率は、主観確率値を対応付けた言葉を用いてユーザに評価させる。言葉を主観確率に変換する手法としては、[31]や[32]が知られている。増田ら[32]は、「十分~である」、「おそらく~である」、「どちらかという~である」、「どちらかという~でない」、「あまり~でない」、「全く~でない」などの言葉を主観確率に変換する手法を提案している。これは各ユーザの主観確率がベータ分布として求められる手法である。ここではこの結果から、言葉のベータ分布の期待値がほぼ等分されることを用いて、次のように主観確率を求めた。6. 十分妥当である(0.95), 5. おそらく妥当だ(0.77), 4. どちらかという妥当だ(0.59), 3. どちらかという妥当ではない(0.41), 2. あまり妥当ではない(0.23), 1. 全く妥当でない(0.05)。以降では、これらの言葉を評価カテゴリと呼ぶ。

3.2 Noisy or モデル

本論では、ベイジアンネットワークの構造を、Toulminモデルにおける各構成要素間の論理的関係を反映するように、ユーザが構築したToulminモデルから決定する。

まず「論拠」 W が任意の局所論証 $X \rightarrow Y$ に付与された、図2中(9),(10),(11),(18),(19)のようなToulminモデルを考える。これらのToulminモデルは、文章 Y の正当性が、文章 X と「論拠」 W に依存して決まることを意味する[23]。更に、文章 X と「論拠」 W は、一方が他方の正当性を高める関係になく独立

である．ここで、例として、図 1 の Toulmin モデルを考える．図 1 では、「主張：彼にはペニシリン治療が必要である」が Y 、「データ：彼は肺炎である」が X 、「論拠：肺炎にはペニシリン治療が有効である」が W に対応する．この例から、「主張」の正当性が「データ」と「論拠」の両方に依存して決まること、「データ」と「論拠」が独立な関係であることがわかる．以上より、 X, Y, W を確率変数とみなすと同時確率分布は $p(X, Y, W) = p(Y|X, W)p(X)p(W)$ と書け、図 4(2) 右図のようなベイジアンネットワークで表現できる．

また、図 2 中 (15),(16),(17) のように、局所論証 X Y に「限定」 Q が付与された Toulmin モデルについても、同様の説明が可能である．

一方、図 2 中 (12),(13),(14) のように、局所論証 X Y に「論駁」 R が付与された Toulmin モデルにおいて、「論駁」 R は文章 Y の正当性向上に直接には寄与しないため [22]、ベイジアンネットワークには反映しない．

以上より、Toulmin モデルからベイジアンネットワークを構築するための規則を図 4 にまとめた．例として、図 3 の Toulmin モデルのベイジアンネットワーク表現を図 5 に示す．

この変換では、論証が複雑化し、ある文章を支持する「データ」「論拠」「限定」が増加すると、ベイジアンネットワークのパラメータ数が指数的に増大し、ユーザが全ての条件付確率を評価することが困難になる．しかし、説明変数のどれか一つが成立したときに目的変数が成立するような場合には、ベイジアンネットワークの特殊形であり、計算量を線形オーダーに抑えることができる Noisy or モデルが利用できる．Noisy or モデルでは、各説明変数の不成立が変数間で独立に生じると仮定できるとき [16]、説明変数集合 Π_{S_i} の任意のパターン j に対する条件付確率を次式で計算できる．

$$p(S_i = 1 | \Pi_{S_i} = j) = 1 - \prod_{S_l \in I_j} p(S_l = 0 | S_l = 1) \quad (1)$$

ここで、 I_j は Π_{S_i} の j 番目のパターンにおいて 1 を取る変数の集合を表す．

他方、図 3 のように、ある文章に対して複数の局所論証が提示される場合、一つでも局所論証が成立すればその文章は正当化されると考えられる．したがって、その文章を目的変数、その文章を支持する各局所論証

を説明変数とみなすと、Noisy or モデルの仮定が満たされる．そこで、本論では、Noisy or モデルを以下のように用いる．

任意の「主張」(または「データ」) S_i に対し R_i 個の局所論証 $A_i = \{A_{i1}, \dots, A_{ir}, \dots, A_{iR_i}\}$ があるとき、 A_{ir} を所与とする条件付確率 $p(S_i = 1 | A_{ir})$ を $r = 1, \dots, R_i$ について、それぞれ独立にユーザに評価させ、最終的に条件付確率 $P(S_i = 1 | \Pi_{S_i})$ を式 (1) で決定する．これにより、ユーザに評価させる条件付確率の数を、 $2^{\sum_{r=1}^{R_i} |A_{ir}|}$ から $\sum_{r=1}^{R_i} 2^{|A_{ir}|}$ に軽減できる．

例えば、図 5 の場合、「 S_1 :主張」を支持する 2 つの局所論証を所与とした条件付確率 $p(S_1 = 1 | S_2, S_3, S_4)$ と $p(S_1 = 1 | S_5, S_6)$ をそれぞれユーザに評価させ、条件付確率 $p(S_1 = 1 | \Pi_{S_1})$ を式 (1) により決定する．これにより、 $p(S_1 = 1 | \Pi_{S_1})$ に関し、ユーザが評価する確率値の数は、 $2^5 = 32$ から $2^3 + 2^2 = 12$ に軽減される．

4. 論証特性指標

ここでは、Toulmin モデルのベイジアンネットワーク表現を用いて、1. 論証の強さ、2. 文章の正当性、3. 主張への影響度、という 3 つの論証特性指標を定義する．

4.1 論証の強さ

「論証の強さ」は文章間の因果の強さを表すため、変数間の依存関係の強さを表す条件付相互情報量を用いて定式化する．ここで、2 変数 S_i, S_j 間の依存関係を $S_i \rightarrow S_j$ とすると、論証の強さは次式で定義される．

$$I(S_i, S_j | \Psi_{S_{ij}}) = \sum_{\Psi_{S_{ij}}} p(\Psi_{S_{ij}}) \sum_{k=0}^r \sum_{k'=0}^r p(S_j = k | S_i = k', \Psi_{S_{ij}}) p(S_i = k') \log \frac{p(S_j = k | S_i = k', \Psi_{S_{ij}})}{p(S_j = k | \Psi_{S_{ij}})} \quad (2)$$

式 (2) の計算には、変数 S_i の親ノード集合 Π_{S_i} における任意の部分集合 $\beta \subseteq \Pi_{S_i}$ を所与とする条件付確率 $P(S_i = k | \beta)$ が必要である．これは条件付確率 $P(S_i = k | \Pi_{S_i})$ から β 以外の全ての変数 $\Pi_{S_i} \setminus \beta$ を周辺化し、 $\sum_{\Pi_{S_i} \setminus \beta} p(S_i = k | \Pi_{S_i}) p(\Pi_{S_i})$ により得られる．

条件付相互情報量は、0 を最小値とし変数間の依存関

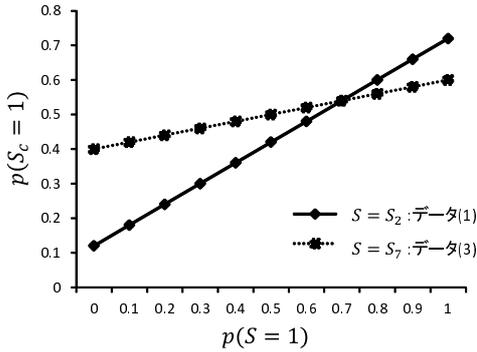


図 6 Sensitivity function
Fig. 6 Sensitivity functions

係が強いほど大きい値を取るため、 $I(S_i, S_j | \Psi_{S_{ij}}) < \epsilon_0$ (本論では $\epsilon_0 = 0.05$) のとき論証が弱いと判断する。

4.2 文章の正当性

「文章の正当性」は、ある文章が、論証によりどの程度正当化されているかを表す指標である。ここでは、これを文章 S_i が真である周辺確率として式 (3) で定義する。

$$p(S_i = 1) = \sum_{j=1}^{q_i} p(S_i = 1 | \Pi_{S_i} = j) p(\Pi_{S_i} = j) \quad (3)$$

この計算は Variable Elimination アルゴリズム [33] により、根ノードから順に変数消去することで厳密かつ容易に実行できる。

この値が低い場合、すなわち、 $p(S_i) < \epsilon_1$ (本論では $\epsilon_1 = 0.7$) のとき、その文章は十分に正当化されていないと解釈する。

4.3 主張への影響度

「主張への影響度」は、ある文章 S_i が「主張」 S_c の正当化にどの程度影響を与えるかを表す指標である。より具体的には、文章 S_i の正当性が向上するように論証を改訂したとき、どの程度「主張」の正当性が向上するかを表す。ベイジアンネットワークにおいてこのような分析を行う手法として Sensitivity Analysis [17] [18] が有用である。ここで、ある変数 S_i が真である確率 $p(S_i = 1)$ が変化するとき、「主張」 S_c が真である確率 $p(S_c = 1)$ がどのように変化するかを調べたいとする。Sensitivity Analysis では、 $p(S_c = 1)$ を $p(S_i = 1)$ の関数 (Sensitivity function と呼ばれる) として以下のように表現する。

$$p(S_c = 1) = \mu_i p(S_i = 1) + \nu_i$$

$$\begin{cases} \mu_i = \frac{p^0(S_c=1, S_i=1)}{p^0(S_i=1)} - \frac{p^0(S_c=1, S_i=0)}{p^0(S_i=0)} \\ \nu_i = \frac{p^0(S_c=1, S_i=0)}{p^0(S_i=0)} \end{cases} \quad (4)$$

ここで、 p^0 は $p(S_i = 1)$ の値を変化させる前の確率分布を表す。

例として、図 5 のベイジアンネットワークにおける、「 S_2 :データ (1)」及び「 S_5 :データ (3)」が真である確率 $p(S_2 = 1)$ と $p(S_5 = 1)$ を変化させた場合の Sensitivity function を図 6 に示す。図 6 より、「 S_2 :データ (1)」の正当性 $p(S_2 = 1)$ が向上するように論証を改訂した方が、より効率よく「主張」の正当性 $p(S_1 = 1)$ を向上できることがわかる。

このように Sensitivity function を比較することで、論証中の各文章の正当性向上が「主張」の正当化にどの程度影響するかを判断できる。

4.4 論証特性指標の妥当性評価

本節では、本論で提案した 3 つの論証特性指標の妥当性を評価する。ここでは、専門家による「論証の強さ」「文章の正当性」「主張への影響度」の評点と、本論で提案した各論証特性指標の推定値との相関を実験により算出する。

実験手順は以下の通りである。はじめに、筆者らが用意した 2 つの論証を専門家に配布した。ここでは、文献 [3] (pp.171-174) [34] (pp.197-205) で悪文として例示された論証を用いた。これらの論証の内容と、論証を Toulmin モデルに当てはめたときの構成要素数、総文字数、ネットワークの平均深さと平均幅を表 1 にまとめた。ここで、ネットワークの平均幅は、Toulmin モデルのベイジアンネットワーク表現における、各ノードの親ノード数の平均を表し、ネットワークの平均深さは、根ノードから各葉ノードまでの各パスに含まれるノード数の平均を表す。それぞれ括弧内は分散を表している。なお、本実験で用いた論証には、「弱い論証」「正当性の低い文章」が各一つ以上出現していた。専門家には、原文とその論証の Toulmin モデルを記載した紙資料を配布した。

専門家には、まず、2 つの論証中の全ての「論証の強さ」と「文章の正当性」「主張への影響度」を 3 段階 (3: 高, 2: 中, 1: 低) で評価させ、Toulmin モデルが記載された紙面上に評点を記入させた。次に、これらの Toulmin モデルのベイジアンネットワーク表

表 1 実験に利用した論証

Table 1 Argumentations used in the experiments

	内容	構成要素数	総文字数	ネットワークの平均深さ	ネットワークの平均幅
論証 1	現行の中学校の歴史教科書が教室での利用に不適切であることを主張する論証文	20	1238	3.16(0.75)	2.67(1.92)
論証 2	高等動物を用いた現行の動物実験が倫理的に不当であることを主張する論証文	16	1687	3.80(1.10)	2.80(1.64)

表 2 論証特性指標の妥当性評価結果

Table 2 Evaluation results of argumentation characteristic indexes

	論証の強さ	文章の正当性	主張への影響度
相関係数	.87	.79	.67
p	< .01	< .01	< .01

現に必要な条件付確率値を、筆者らが作成した条件付確率表に、3.1 で挙げた 6 段階の評価カテゴリで入力させた。

ここでは、入力された評価カテゴリを条件付確率値に変換し、これを用いて算出した各論証特性指標の推定値と、専門家による評点とのスピアマンの順位相関係数を計算した。ただし、「主張への影響度」は関数で定義されるため、専門家による評点との相関係数を直接求めることができない。「主張」への影響は、式 (4) で表される Sensitivity function の傾きが大きいほど強いと解釈できるため、本実験ではこの値を用いて相関係数を算出した。結果を表 2 に示す。

表 2 より、全ての論証特性において、有意水準 1% で専門家の評点と論証特性指標の推定値に正の相関が認められたことがわかる。このことから、提案した各論証特性指標の推定値は専門家による評点と類似した傾向となることが示された。

4.5 ユーザによる主観評価の妥当性評価

本システムでは、論証特性指標の推定値やフィードバックの内容が、ユーザによる条件付確率の主観評価の結果に依存して決まるため、これらの主観評価が適切に行われていることが重要となる。本論では、ユーザは、あらかじめ論証と文章の妥当性評価に関する指導や説明を受け、適切に主観評価を行うことができると仮定する。

ここでは、まず、これらの指導を受けたユーザが、主観評価を適切に行うことができるかを評価する。被験者として論文執筆経験のない理系大学生 10 名ずつを集めた。被験者には、実験内容を説明した後、Toulmin モデルと各論証特性指標に関する説明、および、文章の正当性に対応する条件付確率の評価例を例文を用い

て 20 分程度で解説した。更に理解を深めさせるために簡単な例題を用いた演習問題と解説を 10 分程度実施した。その後、被験者には、原文とその論証の Toulmin モデルを記載した紙資料、更に、Toulmin モデルのベイジアンネットワーク表現に必要なすべての条件付確率表について、確率値の箇所を空欄にしたものを記載した紙資料を配布した。被験者には、これらを用いて、各条件付き確率表の空欄に 6 段階の評価カテゴリを記入させた。

これらのデータを用いて、ここではまず、4.4 の実験で専門家が入力した条件付確率に対応する評価カテゴリと、被験者が入力した評価カテゴリについて、スピアマンの順位相関係数を算出した。結果として、論証 1 では $r = 0.82 (< .01)$ 、論証 2 では $r = 0.71 (< .01)$ となり、有意水準 1% で、被験者と専門家の主観評価結果に正の相関が認められた。このことから、本実験の被験者が、条件付確率を適切に評価できたことがわかる。

次に、被験者による条件付確率の主観評価値を用いて算出される各論証特性指標値の妥当性を確認する。ここでは、被験者と専門家の主観評価結果をそれぞれ用いて各論証特性指標の値を算出し、これらの相関を求めた。ここではピアソンの累積相関係数を用いた。また、4.4 の実験同様「主張への影響度」に関する相関を求めるために Sensitivity function の傾きを利用した。結果を表 3 に示す。

表 3 より、全ての論証特性指標について、専門家の主観評価結果を用いたときの推定値と、被験者の主観評価結果を用いたときの推定値に、有意な正の相関があったことが確認できる。このことから、本実験で被験者にフィードバックされた各論証特性指標値は、専門家の主観評価結果から推定した値と同傾向であったことが示された。

ただし、表 3 より、論証 2 の「主張への影響度」は、他の指標と比べ、相関係数の値が小さいことがわかる。このことは、論証 2 が弱い論証を繰り返して主張を導

表 3 被験者へのフィードバックの妥当性評価
Table 3 Evaluation results of system feedbacks

		論証の強さ	文章の正当性	主張への影響度
論証 1	相関係数	.43	.65	.89
	p	< .01	< .01	< .01
論証 2	相関係数	.67	.76	.27
	p	< .01	< .01	< .05

いている悪文であり、ネットワークが深く、主観評価値の設定が難しいことに原因していると考えられる。

「主張への影響度」は、ネットワークが深い場合、個別の局所論証に対する主観評価値の誤差が累積し、推定値の信頼性が低下することがある。論証 1 に比べ、論証 2 はネットワークが深く、かつ、専門家と被験者間の主観評価値の誤差が若干大きかったことから、論証 2 の「主張への影響度」では相関係数が小さく算出されたと考えられる。

更に、一般的な論証でどのようなネットワークが構築されるかを調べるために、鈴木ら [21] の実験で用いられたテーマ「最寄駅から大学までの最も良い経路について論じよ」で、7 人の被験者にレポートを執筆させた。完成した論証から Toulmin モデルを構築したところ、平均ノード数は 6.29(1.89)、ネットワークの平均深さは 2.22(0.37)、平均幅は 2.61(1.01)であった。更に、2012 年の電子情報通信学会論文誌からランダムに選択した 5 件の情報システム論文の序論の論証を分析し、Toulmin モデルを構築したところ、平均ノード数は 8.80(2.48)、ネットワークの平均深さは 2.33(0.34)、平均幅は 3.40(0.55)であった。

以上より、論証 2 のように弱い論証が繰り返される悪文の場合には、初心者の構成する「主張への影響度」の信頼性は低くなる傾向があるが、それでも有意水準 5% で専門家の判断とも相関があり有用である。文章を推敲して修正することにより、「主張への影響度」も専門家の判断に近づいていくと考えられる。

5. 論証推敲支援システム

5.1 論証特性指標に基づく論証改訂方法のフィードバック

ここでは、論証特性指標の推定結果に応じて論証改訂のためのアドバイスをフィードバックする手法を提案する。

論証中の任意の文章を正当化するためには、以下の論証改訂方法のパターンが考えられる。

(A) その文章を支持する証拠を新たに追加する。

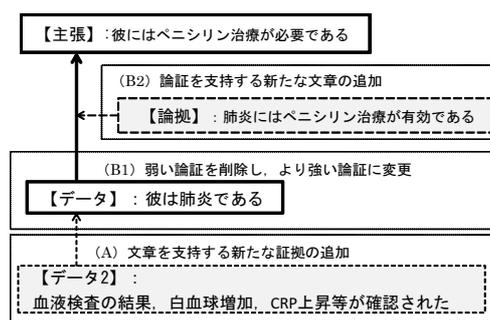


図 7 論証改訂方法のパターン
Fig. 7 Patterns of revising an argument

(B) その文章を支持する既存の局所論証を次のいずれかの方法で改訂する。

(B1) その論証を削除し、より強い論証に変更する。

(B2) その論証を支持する文章を新たに追加する。

これらについての具体例を図 7 に示す。図 7 において、文章「彼は肺炎である (データ)」の正当性を高めたい場合には (A) この文章を支持する証拠 (この例では「血液検査の結果、白血球増加および CRP 上昇が確認された (データ)」) を追加すればよい。一方、文章「彼にはペニシリン治療が必要である (主張)」の正当性を高めたい場合には、以下の改訂方法が考えられる (A) この文章を支持する証拠を新たに追加する。(B) 既存の論証「彼は肺炎である (データ)」。よって、彼にはペニシリン治療が必要である (主張) を次のいずれかの方法で改訂する (B1) 「彼は肺炎である (データ)」を削除し、より強い「データ」を提示する (B2) この論証を支持する文章 (例では「なぜなら、肺炎にはペニシリン治療が有効であるからだ (論拠)」) を追加する。

本システムでは、これらの論証改訂方法を用いて、論証特性指標の推定結果に応じた次のようなアドバイスのフィードバックを行う。

まず、任意の文章 S_i の正当性が低い、すなわち、 $p(S_i = 1) < \epsilon_1$ である場合、その文章の正当性を高めるために、上記の (A) または (B) の方法で改訂を行うようにアドバイスを与える。ただし、改訂方法 (B) に対応したアドバイス文「この文章を支持する既存の局所論証を改訂しましょう」は、その文章を支持する局所論証が少なくとも一つ存在するときのみ提示される。

更に、その文章 S_i を正当化したとき (具体的には $p(S_i = 1) = \epsilon_1$ となったとき)、どの程度「主張」の

表 4 アドバイス文の生成規則
Table 4 Rules for advice generation

論証特性指標の推定結果	アドバイス文
文章の正当性が低い $p(S_i = 1) < \epsilon_1$	この文章は正当性が低いようです。 この文章を正当化すると、 $\Delta p(S_c = 1)$ 程度の「主張」の正当性向上が見込めます。 この文章を正当化するためには、この文章の正当性を $\epsilon_1 - p(S_i = 1)$ 以上高める必要があります。 以下のような方法で、この文章の正当性を高めましょう。 ・この文章を支持する以下のような文章を追加しましょう。 *この文章を支持する「データ」を追加しましょう *この文章に対する「裏付け」を追加しましょう ・この文章を支持している既存の局所論証を改訂しましょう。 **
論証が弱い $I(S_i, S_{d_r}^i \Psi_{S_{d_r}}) < \epsilon_0$	この論証は弱いようです。以下のような方法で論証を改訂しましょう。 ・この論証を削除し、より強い論証に変更しましょう。 ・この論証を支持する以下のような文章を追加し、論証を強化しましょう。 *この論証に対する「論拠」を追加しましょう *この論証に対する「限定」を追加しましょう

*: Toulmin モデルの構成規則に従い、その文章/論証に追加可能な構成要素のみ提示

** : その文章を支持する局所論証が、少なくとも一つ存在するときに提示

正当性向上が見込めるかを、「主張への影響度」の式 (4) を用いて次式で計算し、フィードバックする。

$$\Delta p(S_c = 1) = (\mu_i \epsilon_1 + \nu_i) - p(S_c = 1) \quad (5)$$

また、その文章を正当化するためにはどの程度の正当性向上が必要か、すなわち、 $\epsilon_1 - p(S_i = 1)$ もフィードバックする。

一方、任意の論証について、論証が弱い、すなわち $I(S_i, S_{d_r}^i | \Psi_{S_{d_r}}) < \epsilon_0$ である場合には (B1) または (B2) の方法で論証を改訂するようにアドバイス文を提示する。

なお、改訂方法 (A) と (B2) に対応するアドバイスを提示する際には、Toulmin モデルの構成規則 (図 2) を用いて、着目している文章/論証に追加可能な Toulmin モデルの構成要素を決定し、追加を促すアドバイス文 (例えば「この論証に対する「論拠」を追加しましょう」) を提示する。

表 4 に、各論証特性指標の推定結果に応じて生成されるアドバイス文をまとめる。表 4 では、任意の文章/論証に関する各論証特性指標の推定結果が、各行の 1 列目の条件を満たすとき、その行の 2 列目に列挙されたアドバイス文が提示されることを表す。例えば、ある文章 S_i に着目したとき、「文章の正当性が低い $p(S_i = 1) < \epsilon_1$ 」ならば、表 4 の 1 行目に記述されたアドバイス文が全て提示される。

5.2 論証推敲支援システム

本節では、ここまで述べた手法を実装した論証推敲支援システムについて詳述する。

本システムは図 8 右のようなインターフェイスを持つ。ユーザは、任意のワードプロセッサ (図 8 左) を

用いて文章執筆を行いながら、本システムを用いて以下の手順で Toulmin モデルを構築し、論証の推敲を行う。

(1) システムを起動すると、システム下部にアドバイス文「主張を入力してください」が表示される。ユーザはワードプロセッサに文章を執筆し、「主張」に対応する文章をシステム左下のテキストボックスにコピーしてアドバイス文をクリックする。

(2) 入力された文章が Toulmin モデルのノードとしてシステム上部に可視化され、同時に、その文章の正当性を自己評価するように促す図 9 のようなウィンドウが表示される。

(3) ユーザは 3.1 で挙げた 6 段階の評価カテゴリを用いて自己評価を行う。システムは評価結果を主観確率に変換した後、論証中の全ての文章と局所論証について論証特性指標を計算する。

(4) システム上に可視化されている Toulmin モデルのノードまたは有向辺をクリックすると、システム下部に各論証特性指標の算出結果が表示される。更に、これらの値が指標ごとに定めた閾値 ϵ 未満である場合、表 4 の規則に従い、論証改訂のためのアドバイス文を表示する。

(5) これらの論証特性指標及びアドバイスをもとに、ユーザはワードプロセッサ上の文章を改訂または加筆する。このとき、改訂方法 (A) または (B2) のアドバイスに従い、Toulmin モデルの構成要素に対応する文章を執筆した場合には、システム左下のテキストボックスにその文章をコピーしアドバイス文をクリックする。

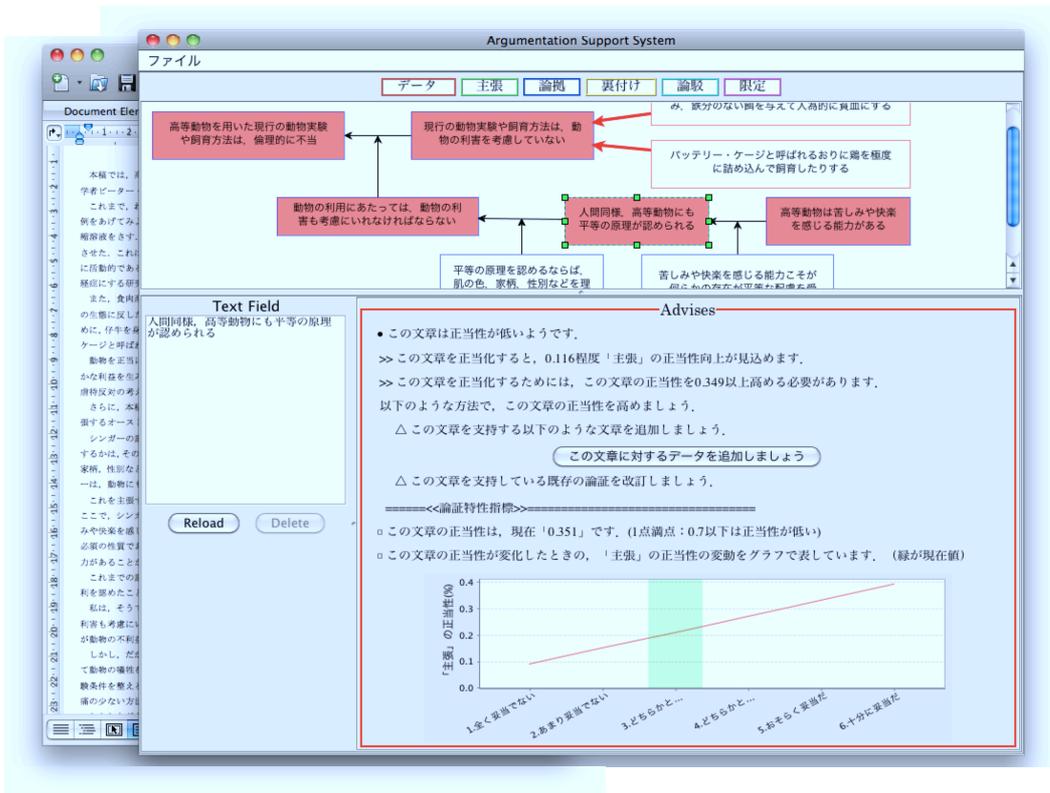


図 8 システムインターフェイス
Fig. 8 System interface

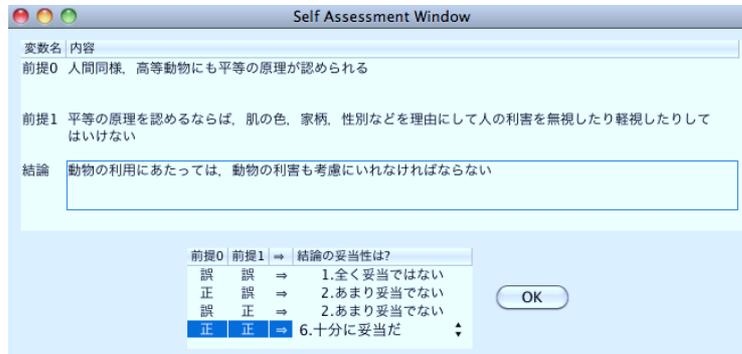


図 9 自己評価ウィンドウ
Fig. 9 Self assessment window

(6) Toulmin モデルにノードと有向辺が追加され、システム内部では Toulmin モデルがベイジアンネットワーク表現に変換される。同時に、追加された文章に関する自己評価を促すウィンドウが表示される。例えば、「主張」 S_C に対して「データ」 S_D を

追加した場合、 $p(S_D = 1)$ 及び $p(S_C = 1|S_D = 1)$, $p(S_C = 1|S_D = 0)$ の主観確率を問う質問文「 D が真である確率」、「 D が真のとき C が真となる確率」、「 D が偽のとき C が真となる確率」が図 9 のように提示される。

(7) 以降, (3)~(7) を繰り返すことで, 論証を Toulmin モデルに当てはめて可視化しながら, 「主張」が十分に正当化されるまで文章と論証の推敲を繰り返す.

なお, システム上に可視化される Toulmin モデルのノードは, 枠線が構成要素ごとに異なる色でハイライトされている. 更に, 文章の正当性が低い箇所と論証の弱い箇所は, 対応するノード背景と有向辺が赤色でハイライトされる.

本システムは Java (Swing) を用いて実装した.

本システムが支援対象とする文章は, ある問題に対する自らの「主張」を, それを支える「データ」や「論拠」を用いて論理的に展開する論証である. 論文のように段落単位の構成が求められる文章であっても, Toulmin モデルの各構成要素として, 段落全体あるいは各段落のトピックセンテンスを対応付けることで, Toulmin モデルを構築できるため, 本システムを利用できる. また, 論証が複雑な場合でも, 画面右にスクロールバーが出現し, ネットワークとアドバイスの提示が可能となる.

6. 評価実験

ここでは, 被験者実験により, 本システムの有効性を評価する. 具体的には, 論証が複雑なとき, 1. 論証の強さ, 2. 文章の正当性, 3. 主張への影響度, を考慮した論証の推敲を本システムが支援できるかを評価する.

6.1 実験方法

上記の評価を行うために, 本実験では, 筆者らが用意した比較的複雑な論証を, 本システム及び比較用システムを用いて被験者に推敲させ, 被験者が行った改訂の内容と改訂後の論証の質を評価する. 具体的な実験手順は以下の通りである.

(1) 被験者を, 本システムを利用する A 群と, 論証を Toulmin モデル [6] に当てはめ可視化する機能のみを持つシステムを利用する B 群にランダムに振り分けた. 各グループの被験者は論文執筆経験のない理系大学生 10 名ずつとした.

(2) 実験内容を説明した後, Toulmin モデルと各論証特性指標, システムの利用方法に関する説明を 20 分程度行った. また, 文章の正当性に対応する条件付確率の評価例を例文を用いて解説した. 更に, 実験内容とシステムの操作について理解を深めさせるために簡単な例題を用いた演習と解説を 10 分程度実施した.

(3) 4.4 の実験で利用した 2 種類の論証をユーザに配布した. 被験者には, 原文とその論証の Toulmin モデルを記載した紙資料を配布した.

(4) 被験者には, まず, 資料と同じ構造の Toulmin モデルをシステム上に構築させた後, システムを自由に操作させながら, 次の作業を行わせた.

(a) 論証の改訂が必要だと思う箇所 (以下, 指摘箇所) を Toulmin モデルに記載された紙面上に全て指摘する.

(b) 指摘した箇所について, 改訂が必要だった理由 (以下, 指摘理由) と, どのように改訂すべきだと思うか (以下, 改訂方法) を自由記述する.

(c) 論証中の全ての文章について, その文章が「主張」の正当化のためにどの程度重要だと思うか (以下, 影響度) を 3 段階 (1. 重要, 2. どちらとも言えない, 3. 重要でない) で評価する.

(d) 実際に文章と論証 (Toulmin モデル) を改訂する.

6.2 論証の推敲内容の妥当性評価

次に, 6.1 の実験手順 (4a),(4b),(4c) で被験者が行った「指摘理由」「改訂方法」「影響度」の評価が, 専門家 (筆者らの一人) が B 群と同じ条件で同様の活動を行った結果とどの程度一致するかを評価する.

そのために, まず, 上記の実験手順 (4b) で被験者が自由記述した「指摘理由」と「改訂方法」を表 5 のカテゴリに分類した. 表 5 には, 各群における各カテゴリの出現頻度の平均と標準偏差 (括弧内) も示した. 表より, どちらの群の被験者も「論証の強さ」「文章の正当性」の観点で論証を推敲していることがわかる.

表 5 のカテゴリを用いると, 実験手順 (4a),(4b),(4c) から得られたデータは, 被験者 $e \in \{1, \dots, n\}$ が各 Toulmin モデル中のノードと有向辺に付与した {指摘理由カテゴリ, 改訂方法カテゴリ, 影響度} の集合として表される.

ここでは, このデータを用いて, 複数の評価者による評定の一致率の指標として知られる κ 係数 [35] を, 専門家と各被験者間について算出した. κ 係数は, 被験者 e がカテゴリ i と評価し, 専門家がカテゴリ j と評価した頻度を $O_{ij}(e)$ とすると, 以下で定義される.

$$\kappa(e) = \frac{\sum_{i=1}^K O_{ii}(e)/N - \sum_{i=1}^K E_{ii}(e)/N}{1 - \sum_{i=1}^K E_{ii}(e)/N} \quad (6)$$

ここで, $E_{ii}(e) = \sum_{i=1}^K O_{ij}(e) \sum_{j=1}^K O_{ij}(e)/N$, $N = \sum_{i=1}^K \sum_{j=1}^K O_{ij}(e)$ を表す.

表 5 論証改定カテゴリ

Table 5 Categories of revising an argument

指摘理由カテゴリ	出現頻度		改訂方法カテゴリ	出現頻度	
	A 群	B 群		A 群	B 群
論証が弱い	4.65(2.84)	3.65(2.09)	その論証を削除する	3.80(2.34)	2.25(1.44)
			その論証を支持する文章を追加する	0.55(1.35)	1.35(1.63)
			その論証を支持する文章を変更する	0.30(0.45)	0.25(0.56)
文章の正当性が低い	1.30(0.81)	0.60(1.01)	その文章を削除する	0.05(0.23)	0.05(0.23)
			より正当性の高いものに書き換える	0.15(0.32)	0.07(0.27)
			その文章を支持する「データ」を追加する	1.10(0.88)	0.35(0.68)

表 6 実験結果

Table 6 System evaluation result

		κ 係数			専門家評価		追加要素数	削除要素数	総文章量	所要時間
		指摘理由	改訂方法	影響度	論証の強さ	文章の正当性				
論証 1	A 群	.39(.15)	.36(.18)	.38(.22)	3.70(.79)	3.70(.79)	2.50(2.14)	5.70(2.06)	936(294)	44.00(6.96)
	B 群	.21(.18)	.18(.17)	.15(.13)	2.10(1.10)	2.10(1.10)	1.40(1.77)	2.60(2.01)	1088(140)	39.30(4.14)
	<i>p</i>	< .05	< .05	< .01	< .01	< .01	< .10	< .01	< .10	< .05
論証 2	A 群	.31(.11)	.31(.12)	.41(.21)	3.00(1.15)	3.10(1.19)	1.30(1.33)	3.00(3.71)	1434(377)	32.60(8.86)
	B 群	.17(.18)	.17(.18)	.11(.12)	1.90(1.10)	1.80(1.03)	1.20(1.69)	1.30(1.06)	1641(96)	26.90(4.32)
	<i>p</i>	< .05	< .05	< .01	< .05	< .05	< .10	< .10	< .10	< .05

表 6 に、専門家と各群の被験者間について算出した κ 係数の平均と標準偏差 (括弧内) を示す。表中の、 $p < .05(.01)$ は対応のない *t* 検定の結果、有意水準 5% (1%) で有意差があったことを表す。表より、全ての項目において、A 群の κ 係数が B 群より有意に大きいことがわかる。このことから、本システムを利用することで、被験者は「論証の強さ」「文章の正当性」「主張への影響度」の観点で、より専門家に近い論証の推敲を行うことができたといえる。

6.3 改訂された論証の質の評価

次に、被験者が改訂した論証が、「論証の強さ」「文章の正当性」の観点で改善されたかを評価する。そのために、被験者が改訂した文章と論証を、以下の 2 項目について 4 段階 (1. 非常に多い, 2. やや多い, 3. あまりない, 4. ほとんどない) で専門家に採点させた。

(1) 論証が弱く、改訂が必要な個所がどの程度みられるか。

(2) 文章の正当性が低く、改訂が必要な個所がどの程度みられるか。

表 6 に、各項目における評価値の平均と標準偏差 (括弧内) を示す。表 6 中の専門家評価の列に記載した $p < .05(.01)$ は、Wilcoxon-Mann-Whitney 検定の結果、有意水準 5% (1%) で有意差があったことを表す。表より、全ての項目において、本システムを用いた A 群が有意に高い評点を得ていることがわかる。このことから、本システムの利用により、「論証の強さ」「文章の正当性」の観点で論証の改訂が促進されることが

示せた。

ここで、被験者による論証の改訂と専門家による評点との関係を分析するために、改訂時に追加・削除された Toulmin モデルの構成要素数と改訂後の総文章量、所要時間を計測した。結果を表 6 に示す。表より、被験者による改訂では、追加要素数に比べ削除要素数が多いことが確認できる。これは、本実験で利用した論証が、冗長な文章を多く含むこと、領域知識を要するテーマであり証拠を新たに創出することが困難であったことに起因するといえる。

一方、専門家による評点が高かった A 群の特徴として、追加・削除要素数ともに B 群に比べて多いことが挙げられる。ここで、追加・削除要素数と、専門家が採点した「論証の強さ」「文章の正当性」の平均点との相関係数を算出したところ、それぞれ $r = .38(p < .05)$ 、 $.54(p < .01)$ で有意な正の相関が確認された。このことから、A 群の被験者は、文章の追加・削除が必要な箇所を B 群の被験者より多く同定し効果的に改訂できたため、高い評点を得られたといえる。

更に、表 6 より、A 群の被験者の方が所要時間が有意に長かったことが確認できる。このことから、本システムが、被験者による注意深い論証の推敲を促進したと解釈できる。

以上の実験より、本システムが、1. 論証の強さ、2. 文章の正当性、3. 主張への影響度を考慮した論証の推敲を支援できることを示せた。

7. むすび

本論では、アカデミックライティングにおける論証の推敲支援として、論証が複雑なときにも「論証の強さ」を重視した論証の推敲を支援できるシステムを開発した。このために、本論では、Toulmin モデルに当てはめて構築された論証をベイジアンネットワークとして定式化する手法を提案した。更に、Toulmin モデルのベイジアンネットワーク表現を用いて、1. 論証の強さ、2. 文章の正当性、3. 主張への影響度という3つの論証特性指標を定義し、これらの指標を用いて論証改訂方法をアドバイスする手法を提案した。

また、被験者実験により、本システムが1. 論証の強さ、2. 文章の正当性、3. 主張への影響度、を考慮した論証の推敲を支援できることを示した。

なお、本論では、本システムを利用することで論証の推敲能力が向上することまでは主張できていない。今後はこの能力を伸ばすための利用方法についても研究する予定である。また、本システムを拡張し、ユーザ同士による協調的な論証推敲環境を構築していきたい。

文 献

- [1] N. Kando, "Text-level structure of research papers: Implications for text-based information processing systems," In Proceedings of the British Computer Society Annual Colloquium of Information Retrieval Research, pp.68–81, 1997.
- [2] 大井恭子, 石川直美, "「知識伝達モデル」から「知識変形モデル」への発展を志向するアカデミック・ライティング指導: 「問いかけ」とピア・レビューの重要性に着目して," 千葉大学教育学部研究紀要, vol.54, pp.105–117, 2006-02-28.
- [3] 戸田山和久, 論文の教室, 日本放送出版協会, 2002.
- [4] 館野泰一, 大浦弘樹, 望月俊男, 西森年寿, 山内祐平, 中原 淳, "アカデミックライティングにおける協同推敲活動環境の構築と評価," 日本教育工学会研究報告集, vol.8, no.3, pp.59–62, 2008-07-05.
- [5] O. Scheuer, F. Loll, N. Pinkwart, and B.M. McLaren, "Computer-supported argumentation: A review of the state of the art.," I. J. Computer-Supported Collaborative Learning, vol.5, no.1, pp.43–102, 2010.
- [6] S.E. Toulmin, The Use of Argument, Cambridge University Press, 1958.
- [7] 館野泰一, 大浦弘樹, 望月俊男, 西森年寿, 山内祐平, 中原 淳, "アカデミック・ライティングを支援する ict を活用した協同推敲の実践と評価," 日本教育工学会論文誌, vol.34, no.4, pp.417–428, 2011-03-30.
- [8] O.C. Brudvik, K.S. Hong, Y.S. Chee, and L. Guo, "Assessing the impact of a structured argumenta-
- tion board on the quality of students' argumentative writing skills," Proceeding of the 2006 conference on Learning by Effective Utilization of Technologies: Facilitating Intercultural Understanding, pp.141–148, IOS Press, Amsterdam, The Netherlands, The Netherlands, 2006.
- [9] J. Azilawati, C.M.L. Ho, and Y.S. Chee, "The impact of structured argumentation and enactive role play on students' argumentative writing skills," ICT: providing choices for learners and learning, 2007, pp.430–439, 2007.
- [10] C. Reed and G. Rowe, "Araucaria: Software for argument analysis, diagramming and representation," International Journal of AI Tools, vol.14, pp.961–980, 2004.
- [11] 堀田昌英, 榎戸輝揚, 岩橋伸卓, "多角的議論構造の可視化手法: 社会技術としての政策論議支援," 社会技術研究論文集, vol.1, pp.67–76, 2003.
- [12] 新田克己, 柴崎真人, 安村禎明, 長谷川隆三, 藤田 博, 越村三幸, 井上克巳, 白井康之, 小松 弘, "ダイアグラムに基づく法的論争支援システム," 人工知能学会論文誌, vol.17, pp.32–43, 2002-11-01.
- [13] 平柳行雄, "トゥールミン・モデルを利用したクリティカル・リーディング力向上," 日本英語コミュニケーション学会紀要, vol.16, no.1, pp.82–92, 2007-12.
- [14] C. Xiong, Y. Pan, and D. Li, "A discussion information-structuring model based on the toulmin formalism," Proceedings of the 1st international conference on Forensic applications and techniques in telecommunications, information, and multimedia and workshop, pp.44:1–44:6, e-Forensics '08, ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, Belgium, 2008.
- [15] 水山光春, "トゥールミンモデルを用いた意志決定過程を組み込んだ環境教育の授業設計," 京都教育大学環境教育研究年報, vol.4, pp.27–40, 1996-03-29.
- [16] F.V. Jensen and T.D. Nielsen, Bayesian Networks and Decision Graphs, 2nd edition, Springer Publishing Company, Incorporated, 2007.
- [17] H. Chan and A. Darwiche, "Sensitivity analysis in bayesian networks: from single to multiple parameters," Proceedings of the 20th conference on Uncertainty in artificial intelligence, pp.67–75, UAI '04, AUAI Press, Arlington, Virginia, United States, 2004.
- [18] L.C. van derGaag and S. Renooij, "Analysing sensitivity data from probabilistic networks," Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence, pp.530–537, UAI'01, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2001.
- [19] 氏川雅典, "トゥールミンの議論モデルの変容-批判から寛容へ," ソシオロギス, vol.31, pp.1–19, 2007.
- [20] 富田英司, 丸野俊一, "思考としてのアーギュメント研究の現在," 心理学評論, vol.47, no.2, pp.187–209, 2004.

- [21] 鈴木宏昭, 館野泰一, 杉谷祐美子, 長田尚子, 小田光宏, “Toulmin モデルに準拠したレポートライティングのための協調学習環境,” 京都大学高等教育研究, vol.13, pp.13-24, 2007-12-01 .
- [22] B. Verheij, “Evaluating arguments based on toulmin’s scheme,” Argumentation, vol.19, no.3, pp.347-371, 2005.
- [23] C. Reed and G. Rowe, “Translating toulmin diagrams: Theory neutrality in argument representation. argumentation,” Argumentation, vol.19, pp.267-286, 2005.
- [24] F.V. James, “Toulmin’s model and the solving of ill-structured problems,” Arguing on the Toulmin Model. New Essays in Argument Analysis and Evaluation, eds. by D.L. Hitchcock and B. Verheij, pp.303-311, Springer, 2006 .
- [25] 野矢茂樹, 論理トレーニング, 産業図書株式会社, 2006 .
- [26] 松原望, ベイズ統計学概説-フィッシャーからベイズへ, 培風館, 2010 .
- [27] 植野真臣, “ベイジアンネットワークの統計的学習,” 人工知能学会誌, vol.25, no.6, pp.803-810, 2010 .
- [28] J. Pearl, Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, 1 edition, Morgan Kaufmann, Sept. 1988.
- [29] D.E. Heckerman, Probabilistic similarity networks, Stanford University, Stanford, CA, USA, 1990.
- [30] 繁柘算男, 植野真臣, 本村陽一, ベイジアンネットワーク概説, 培風館, 2006 .
- [31] R. Beyth-Marom, “How probable is probable? a numerical translation of verbal probability expressions,” Journal of Forecasting, vol.1, no.3, pp.257-269, 1982.
- [32] 増田規彦, 繁柘算男, “言葉を尺度とした主観確率の測定,” 日本ファジィ学会誌, vol.8, no.6, pp.84-90, 1996-12-15 .
- [33] A. Darwiche, Modeling and Reasoning with Bayesian Networks, 1st edition, Cambridge University Press, New York, NY, USA, 2009.
- [34] 福澤一吉, 議論のレッスン, 生活人新書, 2002 .
- [35] J. Cohen, “A coefficient of agreement for nominal scales,” Educational and Psychological Measurement, vol.20, no.1, pp.37-46, 1960.

(平成 xx 年 xx 月 xx 日受付)



鈴木 宏昭



植野 真臣 (正員)

1988 年東京大学大学院教育学研究科博士課程単位取得退学。現在, 青山学院大学 教育人間科学部教育学科 教授。博士(教育学)。人間の学習, 思考, インターフェイスについての認知科学的研究に従事。著書に「知性の創発と起源(編著)」(オーム社)、「類似から見た心」(共立出版)、「コミュニケーションと思考(共著)」(岩波書店)などがある。日本認知科学会(編集委員), Cognitive Science Society, 日本心理学会, 教育工学会各会員。

1993 年神戸大学大学院教育学研究科修了, 1994 年東京工業大学大学院総合理工学研究科修了。現在, 電気通信大学大学院情報システム学研究科准教授。博士(工学)。



宇都 雅輝 (学生員)

2009 年電気通信大学人間コミュニケーション学科卒。2011 年電気通信大学大学院情報システム学研究科博士前期課程修了。同年, 同大学院博士後期課程入学, 現在に至る。

Abstract We propose an argument elaboration support system for academic writing. The traditional systems support construction of an argument as to fit Toulmin model. However, to elaborate an argument, focusing on “Strength of an argument”, which means strength of causal relationship between sentences, is more important. However, it is difficult to elaborate an argument when the argument scale grows up. To solve the problem, we propose a Bayesian network representation of Toulmin model. Using the Bayesian network, the proposed system estimates “Strength of an argument”, “Validity of a sentence” and “Influence of a sentence” and generates optimal advices.

Key words Toulmin model, Bayesian network, Academic writing, Argument elaboration support system