

# ディベート人工知能における影響関係認識のためのテキスト内の論理構造に関する考察

## Analysis about Logical Structures for Effect Relation Recognition in Debating AI

佐藤 美沙<sup>\*1</sup>  
Misa Sato

柳井 孝介<sup>\*1</sup>  
Kohsuke Yanai

柳瀬 利彦<sup>\*1</sup>  
Toshihiko Yanase

三好 利昇<sup>\*1</sup>  
Toshinori Miyoshi

丹羽 芳樹<sup>\*1</sup>  
Yoshiki Niwa

<sup>\*1</sup> 日立製作所 研究開発グループ  
Research & Development Group, Hitachi Ltd.

This paper describes a technique to recognize effect relations in a given text sentence. We did a preliminary experiment of effect recognition by using a simple rule based on syntax structures. Although all of the found relations are correct, most of the effect relations weren't found. Therefore, we analyzed the false negative errors and discuss what are important to recognize effect relations.

### 1. はじめに

筆者らの研究グループでは、電子化されたテキストデータを意思決定支援に活用することを目的として、ディベートをする人工知能を開発している。本研究では、新聞記事などから抽出したテキストが、ある意見を補強するような根拠であるか否かを自動で推定する方法について述べる。たとえば、

The casinos increased employees in Mississippi.  
というテキストは、casino に肯定的な立場である、

Casinos promote employment.  
という意見を補強できる。このような、意見を補強する具体的な事実を多数集めることで、より客観的な意思決定が可能となると考えられる。ディベート人工知能では、議論の対象となっているものが、人が価値を持つと考えるものにどのような作用を与えるかが意思決定における重要な観点になると考え[柳井 15]、この作用の情報をテキストから読みとることを主要な機能の一つとしている。以降、casino にあたるものを議論対象、employment にあたるものを価値と呼ぶ。その他の価値には、たとえば health, crime などがある。

本研究では、新聞記事などから抽出したテキストが、指定の対象から指定の価値への作用関係を含意するか否かを推定する問題について扱う。本稿ではこの問題を作用関係認識と呼ぶ。表1に問題の例を示す。作用関係認識では、議論対象 TGT が価値 VAL に与える作用について、与えられたテキストに記述があるか否かを推定する。先の例では、テキストは、TGT:casino が VAL:employment に与える作用関係を含意している。

事前に作用関係を表す述語を整備しておくことにより、作用関係認識が可能と考えられる。たとえば、「カジノ建設により犯罪が増加した」という文からは、「増加」という述語を手掛かりに、議論対象:カジノによる 価値:犯罪 への促進作用を認識できる。つまり、この文では作用述語の主語が議論対象、目的語が価値となっており、このパターンに当てはまるならば作用関係があると推定するルールによって認識できる。こうした述語の有効性は質問応答や情報抽出などの応用においても報告されている[Hashimoto 12]。

本稿では、作用関係を認識するのに有効なパターンにどのようなものがあるかを検討する。まず作用述語を用いた基本パ

ーンにより作用関係の認識を試みる。次に、再現率を上げるため、基本パターンでは認識できなかった作用関係の文について、どのようなパターンが含まれるかの分析を行う。そこから、作用関係認識において必要となるパターンや知識資源はどのようなものであるかを議論する。

表1. 作用関係認識の例

	項目	例
入力	文	The casinos increased employees in Mississippi.
	議論対象 TGT 価値 VAL	casino employment
出力	作用関係の有無	有り

### 2. 部分構造パターンによる作用関係認識手法

以下の手順により作用関係を認識する。

- (1) 入力文に対して構文解析を行い、句構造構文木を得る。
- (2) 構文木から、(主語, 述語, 目的語)の三つ組を獲得する。
- (3) 入力文がパターンに当てはまるか否かを判定し、当てはまった場合には作用関係があると推定する。

手順(2)では、作用関係を表す動詞を収集した作用述語辞書と、それらの動詞に対する主語と目的語を特定する構文ルール辞書を用いる。作用述語辞書には 416 個の動詞あるいは動詞句が含まれる。たとえば、increase, promote, decrease, suppress が含まれる。

手順(3)では基本パターンを用いる。基本パターンとは、(i)入力文に作用述語が含まれ、(ii)作用述語の主語である名詞句に議論対象を表す語句が含まれ、(iii)作用述語の目的語である名詞句に価値を表す語句が含まれる、の三条件をすべて満たすというものである。後述 4.1 節で説明する記法に従うと、基本パターンは TGT affect VAL である。このパターンに当てはまる文を作用関係ありと認識する。

### 3. 実験

前述の基本パターンによる作用関係認識の精度を調べた。文分割および構文解析には Stanford Core NLP [Manning 14] を用いた。

連絡先: 佐藤美沙, 株式会社日立製作所 研究開発グループ,  
〒185-8601 東京都国分寺市東恋ヶ窪一丁目 280 番地,  
042-323-1111, misa.sato.mw@hitachi.com

### 3.1 データセットの作成

TGTとVALの4通りの組合せ (smoking, income), (smoking, disease), (violent video game, crime), (violent video game, moral) に対して、データセットを作成した。たとえば、smoking (喫煙を禁止すべきか) に対して、income (収入), disease (病気) を価値としたことを意味している。

これらの組み合わせに対し、以下の手順で実験データセットを作成した。表2に作成したデータセット内の文の数を示す。

(1) 英文新聞記事を集めた Annotated English Gigaword [Napoles 12] に対して、TGT, VAL およびそれらの類義語をクエリとして全文検索を行うことで、上位 200 記事を選択した。

(2) 選択された記事内から、TGTとVALの両方が前後 2 文内に 1 回以上出現するような文を抽出した。

(3) 抽出した文に対して、著者ら以外の人の手により、作用関係の有無のラベルを付与した。

表2. データセット

作用関係あり	作用関係なし	合計
270 文	838 文	1108 文

### 3.2 結果

結果を表3に示す。データセット中の 23 文が作用関係ありと認識され、すべて実際に作用関係のある文であったため識別率は 100%であった。一方、再現率は 8.5%であった。再現率を上げる必要がある。

表3. 基本パターンによる作用関係認識結果

識別率	再現率
100%	8.5%

## 4. 文例の分析

再現率向上のため、基本パターン以外のパターンを実装する必要がある。そこで、実際は作用関係があるにもかかわらず認識結果では作用関係がないとされた False Negative の誤り例に対して分析を行った。本章ではその分析結果を述べる。特に、ディベート人工知能において根拠として抽出したい例を中心に見ていく。

まず本章で用いるテキスト内部分構造の記法を説明する。

議論対象は TGT, 価値語は VAL で表す。作用述語は affect で表す。また、述語との相対的な位置により、主語あるいは目的語であることを表す。基本パターン TGT affect VAL は、affect クラスの述語がテキスト内に含まれる場合において、述語の主語に述語クラス名の左側に記された TGT が含まれ、かつ述語の目的語に述語クラス名の右側に記された TGT が含まれることを示す。

作用述語以外に以下の述語に着目する。具体例は各文例に譲る。

- prefer: 主語の態度を表す。
- be: 等価を表す。
- significant: 主語の良し悪しを表す。観点を目的語とする。
- contrast: 比較を表す。比較対象を目的語とする。「contrast:VAL」は、述語が価値を表す語の比較級であることを示す。

知識情報は[]で表す。以下の知識情報がある。

- [authority]: 権威を表す。政府や研究結果などの、信頼がおけると通常考えられるものを表す。
- [alternative TGT]: 議論対象の代替手段を表す。たとえば、TGT が generic drug であるとき、branded drug が代替手段に含まれる。

一文には複数の部分構造が含まれる。複数の部分構造を記号^でつなげる。

以下、発見された部分構造パターンと該当する文例を挙げる。例文中の太字は TGT あるいは VAL を、下線を引いた文字は着目する述語を示し、知識やそれ以外の着目点をイタリックで示す。一部の例は前節で説明したデータセットには含まれないが、データセット同様 Annotated English Gigaword [Napoles 12] 内の文である。各例文の後ろに出典を記す。

段落題は、「抽出された部分構造→正解の作用関係」を表す。矢印の右側に、入力文に含まれる正解の作用関係を、対応する基本パターンの形式によって記した。この矢印はルールとして見ることができ、これらのルールを2章の手順(3)で用いることを想定している。

### 4.1 簡単な推論が有効となる例

#### (1) TGT ^ affect- VAL- → TGT affect- VAL-

Announcing a proposed four-year **medical cannabis** trial, New South Wales Premier Bob Carr said his government had an obligation to explore all avenues available to minimize pain among the sick. (©1994-2010 Agence France Presse)

価値 pain は作用述語 minimize の目的語である一方、議論対象 cannabis は minimize の主語に含まれてはいないが他の部分に出現している。議論対象と価値の両方もが作用述語と直接繋がってはいなくとも、このようなパターンの組合せを認めることで、作用関係を認識できる。

#### (2) TGT be+ X ^ X affect- VAL- → TGT affect- VAL-

Mrs. Bradley, who uses a wheelchair, said for many people, **cannabis is** the only drug that helps with the pain. (©1994-2010 The Associated Press)

議論対象 cannabis が句 X (the only drug) と等価関係にあり、その句 X を主語とする作用述語 helps with の目的語が価値 pain となっている。

TGT と句 X が等価関係であることと、X affect VAL であることから、TGT affect VAL の基本パターンと類似の意味を持つと考えることでこの作用関係は認識できる。このような簡単な推論が有効と考えられる。この文例は(1)のパターンもあてはまるが、このパターンの方がより詳細であり、より高い識別率が期待される。

#### (3) VAL+ affect- TGT → TGT affect- VAL+

Agencies also said the nature of their work and data **security concerns** made it difficult for them to support telework programs. (©1994-1998, 2003-2009 Los Angeles Times-Washington Post News Service, Inc.)

価値 security からの議論対象 telework への負の作用が記されている。基本パターンとは主語と目的語が反対でありながら、telework が価値 security に対して負の作用を持つことを含意している。こうした主語目的語が反対の例を扱う作用関係認識ルールを書くことができる。

#### (4) [authority] prefer+ VAL+ TGT → TGT affect+ VAL+

Canada and Holland allow medicinal cannabis use and nine US states have passed similar measures although Washington

remains opposed to any relaxation in drug laws. (©1994-2010 Agence France Presse)

権威者が議論対象について肯定的/否定的態度を示すときは、何らかの根拠に基づいている可能性がある。そのときに、価値への言及があった場合、議論対象が価値へ作用していることを根拠としていると考えることで、作用関係が認識できる。

## 4.2 代替手段の知識が有効となる例

- (1) affect- VAL- ^ contrast+ [alternative TGT]  
→ TGT affect- VAL-

Donors now expected recipients in Africa to scale back expenditures by procuring cheaper alternatives to branded drug products, the consultancy added. (©1995-2010 Xinhua News Agency)

議論対象 generic drug は文中に明確には記述されていない。代わりに TGT の代替手段となりうる branded drug products が比較対象として示されている。価値 expenditure は作用述語 scale back の目的語であることから、4.1 (1) と類似のルールに従うことで [alternative TGT] affect- VAL- と推定できる。

代替手段についての否定的な根拠は、議論対象についての肯定的な根拠となりうる。[alternative TGT] affect- VAL- → TGT affect+ VAL- という考えにより議論対象についての作用関係を認めることができる。

- (2) [alternative TGT] contrast:VAL- TGT  
^ [alternative TGT] significant-  
^ prefer- [alternative TGT]  
→ TGT affect- VAL-

Others are refusing to buy new drugs that are more expensive but not much more effective than generic versions. (©1994-2010 New York Times)

代替手段 new drugs が議論対象と価値 expensive の観点で比較されていることから、作用関係が推測される。加えて、代替手段について、否定的態度 refusing to buy を示されていること、否定的特徴 not much more effective が記されていることにより、作用関係である可能性が高いと予想される。このように抽出されるパターンの数が多いほど、作用関係を含んでいる可能性が高いとみなすことが考えられる。

## 4.3 その他の特徴が有効となる例

- (1) [authority] prefer+ TGT ^ VAL+ ^ (シンプルさ)  
→ TGT affect+ VAL+

Australia's most populous state on Tuesday unveiled plans to allow cannabis use for medical purposes. (©1994-2010 Agence France Presse)

権威 The commission が議論対象 cannabis に否定的態度を示していること、および、価値 medical への言及があること、により作用関係を認識できる。

4.2 (2) と比べて、significant を含まない点で作用関係である可能性が弱いと考えられるが、文全体としてのシンプルさでは優っている。シンプルな文ほど、複数パターンが抽出される場合の作用関係らしさは高まると考えられる。したがって、句構造構文木上のノード数などの、文のシンプルさを表す指標は手がかかりとなりうる。

## 5. おわりに

本稿では、新聞記事などのテキストから議論の根拠となるような文を抽出するために、作用関係認識という問題を扱い、(主語、

述語、目的語)の三つ組に基づいた基本パターンにより認識する方法について検討した。基本パターン方式では、識別率は100%であるものの再現率が14%と低いことが確認された。次に文例の分析を行った。分析の結果から、簡単な推論方法の実装や、代替手段などの知識の整備が望ましいと思われる。今後はこうした観点に基づいて作用関係認識手法の設計に取り組む。

## 参考文献

- [Hashimoto 12] Hashimoto, C., Torisawa, K., De Saeger, S., Oh, J.H., and Kazama, J.: Excitatory or Inhibitory: A New Semantic Orientation Extracts Contradiction and Causality from the Web, in Proc. of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, Association for Computational Linguistics, 2012
- [Manning 14] Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., and McClosky, D.: The Stanford CoreNLP Natural Language Processing Toolkit, in Proc. of the ACL 2014: System Demonstrations, Association for Computational Linguistics, 2014
- [Napoles 12] Napoles, , Courtney, , Gormley, M., and Durme, B. V.: Annotated English Gigaword LDC2012T21, Philadelphia: Linguistic Data Consortium, 2012
- [柳井 15] 柳井 孝介, 三好 利昇, 柳瀬 利彦, 佐藤 美沙, 丹羽 芳樹, Reisert, P., 乾 健太郎: ディベート人工知能における意見生成, 人工知能学会全国大会論文集, 第 29 巻, pp. 1-3, 人工知能学会, 2015