

# 含意関係認識による金融ドキュメントチェックへの取り組み

馬目 華奈<sup>1,a</sup> 外園 康智<sup>2</sup> 長谷川 貴博<sup>3</sup> 小西 優祐<sup>3</sup> 渡邊 知樹<sup>4</sup>  
 築 有紀子<sup>1</sup> 谷中 瞳<sup>4</sup> 田中 リベカ<sup>1</sup> 峯島 宏次<sup>1</sup> 戸次 大介<sup>1</sup>

<sup>1</sup>お茶の水女子大学 <sup>2</sup>野村総合研究所 <sup>3</sup>オメガ・パートナーズ <sup>4</sup>東京大学

{<sup>a</sup>g1420542, yana.yukiko, tanaka.ribeka, bekki}@is.ocha.ac.jp  
 {takahiro.hasegawa, yusuke.konishi}@omegapartners.jp  
 {hitomiyataka, watanabe-kazuki163}@g.ecc.u-tokyo.ac.jp  
 y-hokazono@nri.co.jp minesima.koji@ocha.ac.jp

## 1 はじめに

金融業界において、金融商品の説明書や、当局への届出書など多くのドキュメントには、投資判断を誤らせないためのガイドラインやコンプライアンスルールが設定されている。ドキュメントは商品や契約の細かな差異により、単語や表現のバリエーションは多いが、ルールは、適用範囲を広くするために抽象化されている。そのため、金融ドキュメントの正確なチェックには、金融の専門知識に基づく文の意味内容の深い理解と推論が要求される。この専門性と正確性が要求されるチェックを自動で行うシステムを構築することは、未だに困難な問題である [1]。

含意関係認識 (Recognizing Textual Entailment, RTE) は、意味処理のための基本タスクであり、与えられたテキスト  $T$  から仮説  $H$  が推論されるかどうかを判定する課題である。  $T$  と  $H$  間で一致している単語数など表層的な情報に着目する手法だけでは、複雑な意味的関係を伴う推論を正確に扱うことが困難であることから、近年、頑健な構文・意味解析と論理的手法を組み合わせた手法が提案されている [2, 3]。論理ベースの手法では、テキストは論理式に変換され、定理証明に基づいて含意関係の判定が行われる。

こうした論理ベースの意味解析・推論システムは、テンプレートや独自辞書に基づくカスタマイズ・拡張の容易さ、アウトプットとなる意味表現・推論結果の解釈性の高さという利点があり、また大規模な教師データを必要としないという点で、ドメインを絞り込んだ応用が期待される。しかし、現状では論理ベースの RTE システムを実社会の課題に応用する試みは十分に検討されていない。そのため、理論と応用の間にギャップがあるのが現状である。

この状況をふまえて、本研究では金融ドキュメントのチェックという課題に論理ベースの RTE の手法を適用することを試みる。チェック対象としたのは、企業の上場申請書類とコーポレートアクション情報である。意味解析・推論システムとしては、cgg2lambda [2, 4] を用いる。具体的には、実際の金融テキストに基づく RTE テストデータを構築し、それに基づいて意味解

析・推論システムの改良を行う。また、論理推論において必要な公理を補完するアブダクション [5] の仕組みを、日本語の推論システムに対応するように拡張する。

以下では、提案システムの全体像を示した上で (§2)、RTE データの構築方法 (§3)、意味解析・推論システムの改良方法 (§4) について説明する。その上で、改良版のシステムによる実験結果について報告し、今後の展望・残された問題点について述べる (§5)。

## 2 提案システムの概要

提案手法の全体像を図 1 に示す。解析の対象となる文書セット (企業の上場申請書類やコーポレートアクション情報) に対しては、チェック項目のリストが質問文の形で用意されている。まず、文書全体から各チェック項目に関連する前提群  $T_1, \dots, T_n$  を選択し、チェック項目の質問を対応する仮説文  $H$  へと変換する。これにより、文書セットがチェック項目を満たすか否かは、 $T_1, \dots, T_n$  が  $H$  を含意するか否か、という含意関係認識の問題として解くことができる。

提案手法では、前提文  $T_1, \dots, T_n$  と仮説文  $H$  は、CCG パーザによる構文解析、cgg2lambda による意味解析を経て論理式に変換され、必要なドメイン知識を表現した知識ベースのもとで、 $T'_1 \wedge \dots \wedge T'_n \rightarrow H'$  が証明可能であるかどうかの判定を行う。

今回はシステム全体のうち、主に意味解析・推論部の開発を行った。より具体的な文の解析プロセスを以下、順に説明する。

### 2.1 CCG による構文・意味解析

まず、前提文・仮説文について、組合せ範疇文法 (Combinatory Categorical Grammar, CCG) [6, 7] に基づく構文解析を行う。CCG では、各語に付与された統語範疇に基づいて、少数の組合せ規則によって文の統語構造と意味合成の計算を同時に行う。CCG パーザは日本語 CCGBank [8] に基づく Jigg [9] と depccg [10] を用いる。形態素解析には、kuromoji<sup>1</sup> を用いる。

<sup>1</sup><https://www.atilika.com/ja/kuromoji/>

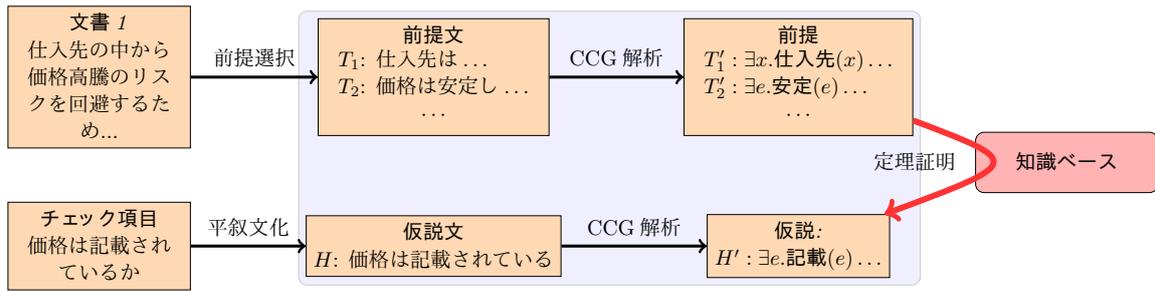


図 1: 提案手法の全体像

構文解析で得られた CCG の導出木をラムダ計算に基づいて論理式へと変換する。各語に対する意味割り当てには、統語範疇単位で意味を指定するテンプレートと、論理語・機能語に対して意味を指定する語彙項目の二種類を使用する。文の骨格となる動詞句に対しては Neo-Davidsonian Event Semantics [11] のアプローチに基づいて意味表示を与える。Event Semantics は、動詞句がイベントを導入するという考え方に基づいており、イベントを中心とした述語項構造に加えて、様々な粒度の論理表現・機能表現の意味関係を表現することができる。cgg2lambda では、意味辞書 (テンプレート・語彙項目) が YAML<sup>2</sup>ファイルで定義されており、対象となるドメインに適合した意味論を開発することが可能となっている。

## 2.2 定理証明とアブダクション

推論は証明支援系である Coq を用いて行う。Coq は一階述語論理の部分系に対する自動推論を含む自動証明機能も備えており、これを高階の公理と組み合わせることで、自然言語の自動推論が可能となる。推論に必要な公理は、ドメインの知識を記述した知識ベースを拡張する形で必要に応じて追加することが可能である。cgg2lambda では、外部知識から前提と仮説の述語間の意味関係を公理として補完するアブダクションの機能が追加 [5] されており、後述するように本研究ではこの機能の日本語への拡張を行った。

## 3 RTE テストデータの構築

金融テキストを対象とする含意関係認識システムの実現性をテストするために、企業の上場申請書類・コーポレートアクションを元にして RTE テストセットを構築した。この作業では、書類申請の際に実際に用いられるチェックリストを参考にした。具体例として、「反社会的勢力の排除に向けた具体的な取り組み状況」に関するテキストのチェック項目には、

1. 反社会的勢力排除に関する要領を定めているか?
2. 取引先について、反社会的勢力との関わりがないことを事前に調査しているか?

などのチェック項目があり、それぞれ項目に対応して、

<sup>2</sup><http://www.yaml.org/spec/>

	パターン
(Ax1)	$P$ ことを定める $\Rightarrow P$
(Ax2)	$X$ は $Y$ をいう $\Rightarrow X = Y$
(Ax3)	$P$ するために $Q$ する $\Rightarrow P$ する

表 1: 追加した公理のパターンの抜粋

- H1. 反社会的勢力排除に関する要領を定めている。
- H2. 取引先や仕入れ先について、反社会的勢力チェックを行っている。

という仮説文を立てることができる。この例のようにチェックリストの多くは疑問文で記述されているため、今回は、過去の類似記載文も参考に、対応する平叙文を手で作成した。構築したテストセットでは、元となる金融ドキュメントから 50 文を抽出し、RTE の前提文とした。それに基づいてチェックリストを参考に、131 問の推論問題を作成した。各問題には yes (含意する)、no (矛盾する)、unknown (どちらでもない) のうちいずれかを正解ラベルとして付与した。yes は「テキストにチェック項目が記載されている」、no 及び unknown は「テキストにチェック項目が記載されていない」ということを意味する。加えて、no の場合は「テキストにチェック項目と矛盾する内容が記載されている」という情報も含んでいる。yes/no/unknown の件数の内訳は、56/37/38 問である。

## 4 cgg2lambda の改良

### 4.1 テンプレートの改良と高階論理の公理

構築した RTE テストセットの一部に基づいて Event Semantics のテンプレート及び高階論理の公理の改良を行った。テンプレートには新たに「利益剰余金」「資本剰余金」「反社会的」などテストセットに登場した重要語 13 語に対して意味割り当てを追加した。これは主に複合表現が表現する述語を一つにまとめるための措置である。

推論に必要な公理を追加する際には、必要な公理を一般的なパターンごとに管理することが重要となる。今回追加した公理のパターンの一部を表 1 に示す。(Ax1) は、いわゆる implicative predicate [12] に分類される述語に関する公理、(Ax2) は定義を述べる述語に関する公理、(Ax3) は目的節に関する公理である。これらのパターンは、高階論理と Event Semantics

を組み合わせることで比較的シンプルに記述することができる。例えば、公理 (Ax1) は、

$$(Ax1) \forall e \forall P (\text{定める}(e) \wedge \text{Accl}(e, P) \rightarrow P)$$

のように記述される。ここで、 $e$  はイベント変数、 $P$  は命題変数であり、Accl は述語「定める」が導入するイベントと「こと」や「と」によって導入される補文の内容  $P$  を結びつける意味役割である。

## 4.2 アブダクションによる公理補完

本節では、アブダクションによって推論に必要な公理を補完する方法について説明する。アブダクションによる公理補完は、チェック時に、担当者の持つ金融知識や一般常識を使うことに対応する。知識ベースとしては、日本語 WordNet [13] を使用した。

具体例として、次の入力文  $\mathbf{T}$  と  $\mathbf{H}$  は、意味解析によって、以下の  $\mathbf{T}'$ ,  $\mathbf{H}'$  に変換される。

$\mathbf{T}$ : ある株式会社が国債を購入する。

$\mathbf{H}$ : ある組織が証券を購入する。

$\mathbf{T}'$ :  $\exists e \exists x \exists y (\text{購入}(e) \wedge \text{株式会社}(x) \wedge \text{国債}(y) \wedge (\text{subj}(e) = x) \wedge (\text{obj}(e) = y))$

$\mathbf{H}'$ :  $\exists e \exists x \exists y (\text{購入}(e) \wedge \text{組織}(x) \wedge \text{証券}(y) \wedge (\text{subj}(e) = x) \wedge (\text{obj}(e) = y))$

ここで、ccg2lambda の推論システムは、 $\mathbf{T}' \rightarrow \mathbf{H}'$  の証明を実行し、どのような単語間知識が不足しているのかを証明過程を追跡することで検出する [5]。この場合、(株式会社( $x$ ), 組織( $x$ )) と (国債( $y$ ), 証券( $y$ )) という述語ペアが証明のギャップとして認識され、日本語 WordNet の単語間の語彙関係<sup>3</sup>から、証明のギャップを埋めるために必要な二つの公理が生成される。

(補完公理 1)  $\forall x (\text{国債}(x) \rightarrow \text{証券}(x))$

(補完公理 2)  $\forall x (\text{株式会社}(x) \rightarrow \text{組織}(x))$

外部知識を使わない場合はシステムの出力は **unknown** であるのに対して、補完公理を加えて証明を再実行するときの出力は **yes** となる。

次の例の場合、(組織( $x$ ), 株式会社( $x$ )) が証明のギャップとして認識されるが、組織を下位語、株式会社を上位語とする語彙関係が WordNet に存在しないため、出力は正しく **unknown** となる。

$\mathbf{T}$ : ABC コープは組織である。

$\mathbf{H}$ : ABC コープは株式会社である。

## 5 性能評価

構築した金融ドキュメント RTE テストデータ 131 件について、提案システムで含意関係の判定を行った。前提文  $T_1, \dots, T_n$  と仮説  $H$  について、 $T_1 \wedge \dots \wedge T_n \rightarrow H'$  が証明可能なら **yes** を、仮説を否定した  $T_1 \wedge \dots \wedge T_n \rightarrow \neg H'$  が証明可能なら **no** を出力し、いずれにも失敗した場合、あるいは 100 秒でタイムアウトした場合には

<sup>3</sup>上位・下位語、反義語、類似関係など 7 種を用いた。

	正答率		適合率		再現率	
	2 値	3 値	2 値	3 値	2 値	3 値
改良前	79.39	61.07	93.94	80.88	55.36	60.92
改良後	<b>91.60</b>	77.86	<b>95.92</b>	86.02	<b>83.93</b>	77.86

表 2: 評価実験結果

**unknown** を出力した。Jigg と depccg それぞれの構文解析によって得られた論理式について、Coq による定理証明を独立に試み、証明に成功した方をシステムの予測として採用した。

## 5.1 結果

表 2 に改良前の ccg2lambda のテンプレート・公理を用いたシステムと今回の改良後のシステムの結果を示す。金融ドキュメントのチェックでは、テキストが仮説 (チェック項目) を含意するか否かが主な問題となるため、**yes** を正例とした 2 値分類のスコアによって評価を行う。参考のため、**yes**, **no**, **unknown** の 3 値分類のスコア (マイクロ平均) も示した。一定の改良によって、改良前の適合率 (93.94%) を損なうことなく、高い正答率 (91.60%) と再現率 (83.93%) を達成することができたといえる。

表 3 に改良後のシステムで含意関係が正しく判定できた例の一部を示す。

## 5.2 考察

改良後のシステムで含意関係の判定に失敗した 11 件のうち、形態素解析・構文解析の誤りによるものは 6 件、意味解析 (テンプレート及び公理) の問題に起因するものは 5 件であった。

意味解析に起因する問題の一つに、推論における型の区別がある。現在のテンプレートでは、Event Semantics の標準的な分析に則り、Entity 型と Event 型を区別した意味割り当てを行っている。型 Entity の項は主に名詞に由来し、型 Event の項は主に動詞に由来する。たとえば、以下の文には「上場」という文字列が 2 度出現する。

- (1) 上場株式には金融商品取引所に上場する新株予約権が含まれる。

この文の論理式では、最初の「上場」は Entity 型の項をとる述語であり、サ変動詞に由来する二番目の「上場」は型 Event の項をとる述語となる。提案システムでは、このように同じ語から変換された述語であっても、型が異なる場合には異なる述語として区別しており、証明時の型エラーの発生を防いでいる。

Entity や Event といった型の区別を設けることにより、証明時の探索空間が狭まり、効率的な推論が可能となるという利点がある。しかし一方で、型の区別によって逆に推論が困難になるケースも存在する。

- (2) a. 今後のますますの発展を期待する。  
b. 今後ますます発展することを期待する。

T-H ペア	正解	予測
T: 反社会的勢力とは、関わりを持たないことを定めます。 H: 反社会的勢力とは、関わりを持たない。	yes	yes
T: 警察や弁護士、全国暴力追放推進センター等の外部専門機関と、綿密に連携しています。 H: 暴力追放推進センターと、連携している。	yes	yes
T: 株主、販売先、購買先についての反社会的勢力チェック業務の主管部署は、経営管理部です。 H: 反社会的勢力チェック業務の主管部署がある。	yes	yes
T: 契約書もしくは覚書で、反社会的勢力との絶縁を前提として取引を行う旨を明記しております。 H: 反社会的勢力との絶縁を前提として取引を行う。	yes	yes
T: 平成 21 年 1 月 1 日以後に支払を受けるべき上場株式等の配当所得については、申告分離課税を選択することができる。 H: ある配当所得については、申告分離課税を選択することができる。	yes	yes
T: 株式会社は、株主に対し、利益剰余金と資本剰余金の双方を原資として配当をすることができる。 H: 株式会社は、株主に対し、資本剰余金を原資として配当をすることができる。	yes	yes
T: 上場株式等には、日本の金融商品取引所に上場する新株予約権が含まれる。 H: 上場株式は、新株予約権を含む。	yes	yes
T: 「企業行動規範」において「反社会的勢力に対する姿勢」を明文化し、全職員の行動指針とする。 H: 全職員の行動指針を明文化する。	unk	unk
T: 株主、販売先、購買先についての反社会的勢力チェック業務の主管部署は、経営管理部です。 H: 販売先についての反社会的勢力チェック業務の副管部署は、総務部です。	unk	unk

表 3: RTE の成功例

上記の 2 文はほぼ同一の意味をもち、一方が他方を含意する。しかし、(2a) の「発展」が名詞由来であるのに対し、(2b) の「発展」は動詞由来であるため、論理式への変換時には異なる述語として区別される。さらに、「今後」「ますます」といった表現についても、(2a) 中では連体詞、(2b) 中では副詞として出現しているため、変換後の述語の型は異なる。結果的に、(2a,b) は異なる述語から構成される論理式に変換され、その間の含意関係を捉えることは困難となる。このような場合には述語の型の差異を無視することで推論が可能になるが、その扱いは今後の課題としたい。

## 6 おわりに

本稿では、論理ベースの言語処理に基づいて金融ドキュメントのチェックを行うシステムの可能性を検討した。システム全体のうち、意味解析・推論部について、構築した金融 RTE テストデータに基づいて改良を行い、現時点での性能評価の結果を示した。

システムの改良プロセスは、具体的な推論課題をデータとして、チェック担当者のもつ金融に関する専門知識を明示化し、そのチェック能力（言語理解・推論能力）をできるだけ正確に再現する仕組みを形式化するプロセスに他ならない。前提選択やチェック項目から仮説文への変換は人手で行ったため、実験結果は最終的なシステムの現時点での上限を示すものである。また、システムの整合性を保ちつつ知識ベースを拡張するためには、外部知識として質の高い金融オントロジー<sup>4</sup>の構築・活用が不可欠である。今後、パーサのエラーを軽減するための *n*-best 解の利用や、論理の深い知識を要求しないテンプレート・公理の書き換えツール、ラッパーの整備を含めて、さらなるシステムの改良を進めていく予定である。

<sup>4</sup>英語ではそのようなオントロジーとして、Financial Industry Business Ontology (FIBO) がある。

## 参考文献

- [1] 長谷川貴博. AI化する銀行. 幻冬舎メディアコンサルティング, 2017.
- [2] Pascual Martínez-Gómez, Koji Mineshima, Yusuke Miyao, and Daisuke Bekki. ccg2lambda: a compositional semantics system. In *Proceedings of ACL 2016 System Demonstrations*, pp. 85–90, 2016.
- [3] Lasha Abzianidze. LangPro: Natural language theorem prover. In *Proceedings of EMNLP2017: System Demonstrations*, pp. 115–120, 2017.
- [4] 田中りべカ, 峯島宏次, Pascual Martínez-Gómez, 宮尾祐介, 戸次大介. 日本語 CCG パーザに基づく意味解析・推論システムの提案. 言語処理学会第 22 回年次大会発表論文集, 2016.
- [5] Pascual Martínez-Gómez, Koji Mineshima, Yusuke Miyao, and Daisuke Bekki. On-demand injection of lexical knowledge for recognising textual entailment. In *Proceedings of EACL2017*, pp. 710–720, 2017.
- [6] Mark Steedman. *Surface Structure and Interpretation*. The MIT Press, 1996.
- [7] 戸次大介. 日本語文法の形式理論: 活用体系・統語構造・意味合成. くろしお出版, 2010.
- [8] Sumire Uematsu, Takuya Matsuzaki, Hiroki Hanaoka, Yusuke Miyao, and Hideki Mima. Integrating multiple dependency corpora for inducing wide-coverage japanese CCG resources. In *Proceedings of ACL 2013*, pp. 1042–1051, 2013.
- [9] Hiroshi Noji and Yusuke Miyao. Jigg: a framework for an easy natural language processing pipeline. In *Proceedings of ACL2016 System Demonstrations*, pp. 103–108, 2016.
- [10] Masashi Yoshikawa, Hiroshi Noji, and Yuji Matsumoto. A\* CCG parsing with a supertag and dependency factored model. In *Proceedings of ACL2017*, pp. 277–287, 2017.
- [11] Terence Parsons. *Events in the Semantics of English*. MIT Press, 1990.
- [12] Lauri Karttunen. Implicative verbs. *Language*, Vol. 47, No. 2, pp. 340–358, 1971.
- [13] 栗林孝行, Francis Bond, 黒田航, 内元清貴, 井原均, 神崎享子, 鳥澤健太郎. 日本語ワードネット 1.0. 言語処理学会第 16 回年次大会発表論文集, 2010.