

言語表現された推論知識による帰結文生成の技術的課題

大西 良明[†] 乾 健太郎[†] 松本 裕治[†]

[†] 奈良先端科学技術大学院大学 情報科学研究科

E-mail: †{yoshiaki-o,inui,matsu}@is.naist.jp

あらまし 計算機で高度な言語処理を行うためには、言語解析だけでなく世界知識に基づいた高度な推論が必要となる。本稿では、自然言語で表現された推論規則、より具体的には述語項構造で表現された事態対によって記述された推論規則を使って推論を行う場合の技術的課題について論じる。実際に、国語辞典の語釈文から推論規則を抽出し、これを入力文に適用する予備的実験を行った。誤り分析の結果、(1) 岩波国語辞典の語義の粒度で述語の多義性を解消したとしても、さらに粒度の細かい多義性によって推論誤りが生じる場合が少なくないこと、(2) 言語モデルによる帰結文のフィルタリングによって推論誤りをある程度回避できること、(3) 推論規則のカバレージを確保するには格交替を柔軟に扱える機構が必要なこと、(4) 項に「A の B」の構造が内在する場合に「B」の顕在化・潜在化を柔軟に扱う仕組みが必要なこと、などが明らかになった。

キーワード 推論制御, 述語項構造, 含意関係認識, プラン認識, 語釈文

Issues in Generating Entailed Sentences Using Inference Rules Represented by Natural Language

Yoshiaki ONISHI[†], Kentaro INUI[†], and Yuji MATSUMOTO[†]

[†] Graduate School of Information Science, Nara Institute of Science and Technology

E-mail: †{yoshiaki-o,inui,matsu}@is.naist.jp

Abstract In order to handle the language on computers, it is important not only to analyze language but to make inference based on world-knowledge. In this paper, we extract the pairs of the predicate argument structures as a rule of inference from dictionary. And we break down the problems and discuss technical challenges we must resolve.

Key words inference rules, predicate argument structure, textual entailment recognition, plan recognition, gloss

1. はじめに

人間のような言語処理能力を計算機で実現するには、頑健な言語解析技術だけでなく大量の世界知識による高度な推論を実現する必要がある。

我々が実現したい推論は大きく2種類ある。一つは、テキスト間の含意関係を認識する推論である。PASCAL RTE Challenge [3] に代表されるように含意関係の認識は近年研究者の注目を集めている。

- (1) a. 夏目漱石は明治39年の春に『坊ちゃん』を雑誌「ホトトギス」に発表した。
b. 『坊ちゃん』の著者は夏目漱石である。

(1) の例文において (1a) が真であれば文 (1b) も真である。このような関係を含意関係という。含意関係を認識する技術は、質問応答や情報抽出、機械翻訳など多くの言語処理アプリケー

ションで重要な役割を果たすことが期待される。

もう一つの推論は人間の行動やプランに関する推論である。例えば「洗濯物を干したとたんに雨が降ってきた」という文が与えられたときに以下のような推論を行いたい。

- (2) a. 「洗濯物を干した」目的が「洗濯物を乾かす」で、
b. 「雨」によって「洗濯物を乾かす」が達成されなかった。

このような推論はプラン認識と呼ばれ、対話プランニングや人の行動支援などに有用である。プラン認識を実現するためには人の行動に対する前提条件や手段、目的など多岐に渡る知識が必要と考えられる。

言うまでもなく、こうした推論を実現するには、モノだけでなくコト、すなわち事態に関する上位下位関係や部分全体関係、因果関係など様々な知識が必要である。本稿では事態の記述表現として述語項構造を使うことを考える。事態に関する知識を述語項構造、すなわち自然言語で表現するアプローチは、次節

で述べるように、知識表現としての頑健性や知識獲得の容易性の面から形式的な知識表現に比べて有利である。ただし、そうした自然言語で表現された上述のような推論を行うには、自然言語がもつ多義性や同義性を頑健に処理する推論機構を開発する必要がある。本稿では、推論の基本パタンの一つとしてモダスポネンス、すなわち (3a) のように「P ならば Q」の形式で表現される推論規則 (以下、単に推論規則) と (3b) のような入力文 P から (3c) のような帰結文 Q を推論 (生成) する問題に焦点をあて、P や Q の事態を述語項構造で表現する場合に生じる技術的課題を論じる。

- (3) a. 推論規則: X ガ Y Ⅱ 取り組む → X ガ Y Ⅲ 解決する
 b. 入力文: 政府が少子化に取り組む
 c. 帰結文: 政府が少子化を解決する

以下、まず 2 節で自然言語による知識表現について述べる。3 節では我々が行った国語辞典語釈文からの推論規則について述べ、4 節では作成した推論知識を用いて実際に行った推論で発生した問題を分析する。

2. 知識をどのように表現するか

知識を表現する上で、どのような表現を用いるかは重要な問題である。我々は自然言語表現で知識を記述する可能性を考える。

事態を表す表現は「雨」や「運動会」のような語、「コンビニで本を買う」のような述語項構造、「毛布がなくて寒い」のような述語項構造を組み合わせたもの、さらにそれらを組み合わせたものなど多岐に渡る。これらのうち事態を表す単位として考えられるのは「雨」のような事態性名詞、「コンビニで本を買う」のような述語項構造がある。さらに「雨」が表す事態は「雨が降る」のような述語項構造で表すことができる。それゆえに我々は事態を表す単位として述語項構造を用いることを考える。

述語項構造で事態を表現するという事は、自然言語に用意された動詞や名詞などのオープンクラスな語彙を基本的にそのまま利用するという事を意味する。このように自然言語の語彙を知識表現に用いる方法は、個々の事態概念やモノ概念に記号を与えて人工的に形式的なオントロジーを用意するアプローチに比べ、形式性が確保できないという欠点はあるものの、オントロジーの開発コスト、また新しい概念に対する頑健性の面で優れて有利である。また人間にとって非常に扱いやすい表現なのでコーパスなどの言語資源からの獲得が容易に行え、述語項構造ベースの資源の構築にかかるコストが少ない。

また、述語項構造での表現では、様々な事態を表す統語構造間に生じる差異を吸収し、オープンクラスの語彙を扱いながらも標準的な表現が可能となる。例えば、「夏目漱石によって発表された『坊ちゃん』」や「夏目漱石による『坊ちゃん』の発表」や「夏目漱石は『坊ちゃん』を発表した」は統語構造が大きく異なるが、これらを述語項構造による表現にすれば [[夏目漱石ガ][『坊ちゃん』ヲ] 発表する] に統一される。このように述語項構造で扱うことは、名詞句や動詞句などの統語構造を気にす

ることなく記述することが可能となる。

さらに述語項構造解析技術は、英語では PropBank や NomBank, また日本語でも京都テキストコーパス第 4.0 版や GDA コーパス, NAIST テキストコーパスなどの資源の蓄積とともに近年急速に発展してきており [7], [8], [10], 述語項構造を単位として含意関係計算にアプローチすることは技術的にも現実性が出てきた。

3. 国語辞典語釈文からの推論規則の抽出

岩波国語辞典を用いて推論規則の抽出を手で行った。国語辞典の事態性を持った見出し語を「見出し語事態」、語釈文中に含まれる事態を「定義語事態」とする。また見出し語事態に述語項の情報を付与したものを「見出し語項構造」、定義語事態に付与したものを「定義語項構造」とする。

我々は事態性のある見出し語の各語義に対し、見出し語項構造を作成し、その語釈文から定義語項構造を作成し、これらの述語項構造間に意味関係を付与した。具体的な作業手順を以下に述べる。

3.1 項の対応関係

見出し語の語義に対して必須と考える項と、項のスロットに適当な変数を与えたものを見出し語に加え、見出し語項構造とする。

- 畳む - X ガ Y Ⅲ 畳む
- 挿入する - X ガ Y Ⅲ Z Ⅱ 挿入する

語釈文から事態性がある語を全て抜き出し、それぞれに対し見出し語と同様に必須と思われる項を加える。同時に見出し語の項のスロットに代入されている変数の対応をつけ、項の対応関係をつける。

- (4) a. 畳む (31542-0-0-1-0)
 b. 折り返して重ねる。広がっているものを折って小さくする。
 c. X ガ Y Ⅲ 畳む → X ガ Y Ⅲ 折り返す
 d. X ガ Y Ⅲ 畳む → X ガ Y Ⅲ 重ねる
 e. X ガ Y Ⅲ 畳む → Y ガ Ⅲ 広がっている
 f. X ガ Y Ⅲ 畳む → X ガ Y Ⅲ 折る
 g. X ガ Y Ⅲ 畳む → X ガ Y Ⅲ 小さくする

項の対応関係が間接照応になっているものも存在する。それらについても同様に以下のように項の対応関係をとる。

- (5) a. 挿入する (29620-0-0-0-0)
 b. 中にさし込む。はさみ込む。
 c. X ガ Y Ⅲ Z Ⅱ 挿入する → X ガ Y Ⅲ Z の中 Ⅱ 差し込む

3.2 意味関係の定義

我々はテキスト間の含意関係認識や人のプラン認識をするための基礎知識を獲得することが目的である。そこで、前節で獲得した述語項構造間につけるべき関係を (a) 述語項構造間の時間的關係 (b) 含意関係が必然的であるか、蓋然的であるか (c) 述語項構造が行動であるか出来事であるか (意思性があるかないか) の観点で分類を行った。そして我々は事態間関係として、上位・同義、結果 (状態)、不可分、前提条件、付帯状況、手段、

目的、反義語の8つを定義した。しかし、これらはあくまで上記の分類基準に従って分類したものであり、推論に必要な事象間関係を全て記述できるとは限らない。事象間関係にどのような分類を行うのが適切であるかは、実際の推論における正確性と頑健性、さらに知識獲得の現実性に基づいて今後検討していく必要がある。以下にいくつかの関係について例示する。より詳しい関係の定義などは文献[1]を参照されたい。

【上位・同義】同時、必然的、行為/出来事→行為/出来事

- XガYヲ使いこなす→XガYヲ使う
- Xガ漂流する→Xガ流れる

【付帯状況】同時、蓋然的、出来事→行為/出来事

- Xガ流行る→Xガ次々と伝わる
- XガYヲ防ぐ→XガYヲさえぎる

【手段】同時、蓋然的、行為→行為

- XガYヲ炙る→XガYヲ火にあてる
- Xガ稼ぐ→Xガ働く

【結果(状態)】定義語が後、必然的、行為/出来事→出来事

- XガYヲ傾ける→Yガ傾く
- XガYヲ起こす→Yガ立つ

【目的】定義語が後、蓋然的、行為→行為/出来事

- XガYヲメモする→XガYヲ忘れない
- XガYヲ味見する→XガYの味加減ヲ知る

【前提条件】定義語が前、必然的、行為/出来事→出来事

- Xガ言い逃れる→Xガ問い詰められる
- XガYヲ迎え撃つ→Yガ攻める

3.3 作業経過

「岩波国語辞典コーパス 2004」^(注1)として入手可能な岩波国語辞典第4版を用いて行った。岩波国語辞典は一つの見出し語に対して、大分類、中分類、小分類と階層的に語義が付けられており、最も細かい小分類された語義を今回は使用した。さらに、このコーパスには語釈文に使用されている動詞、名詞の語義が特定されている。以下は実際に今回の作業で抽出した推論規則の例である。

- (6) a. XガYヲ寄せ付ける (53474-0-0-0-0)→Yガ近寄る (32591-0-0-1-0)【結果(状態)】
- (7) a. XガYヲ盗掘する (53474-0-0-0-0)→XガYをとる (37713-0-0-2-4)【同義・上位】

このように、今回の知識抽出の作業は岩波国語辞典の語義を基準としており、その意味では語義の曖昧性が解消されたデータを使用していることになる。ただし岩波国語辞典の語義が推論を行う上で全ての曖昧性を解消しているわけではない。このことについては、詳しくは後述する。

現在、登録されている事象性を含んだ語義 24078 個に対して作業を行った。これは RWC コーパス中における事象性を含む語義の出現回数の合計の 99.17 %を網羅している。抽出した関係数は 42478 個で内訳は表 1 のとおりである。

表 1 国語辞典から抽出した関係と個数

関係	抽出個数
同義・上位	21852
結果(状態)	1435
前提条件	2083
付帯状況	5963
手段	1222
目的	610
反義語	424
不可分	154

以上のようにして出来上がった見出し語項構造と定義語項構造の対を以下では推論規則と呼び、推論規則の集合を知識データベースと呼ぶ。

4. 推論規則の適用による帰結文生成

次に、3 節で述べた推論規則を「PならばQ」の形式の推論規則と見なし、それらを実際の文に適用する実験を行った。

まず、実験に用いる推論規則の数を制限するため、前件の動詞語義の出現頻度が高いものから順に 500 個の推論規則を選択した。ただし、語義の出現頻度は新聞記事に語義タグが付与されている SenseVal コーパス(白井[12])を使って推計した。これらからランダムに 338 件を抜き出し、それに語義 ID 情報が付与された述語項構造の入力文を人手で作成し適用した。例文は当該語義の用法と思われるものであり、この際に必須格は全て埋めるように指示した。作業は上述の推論規則について何の知識も持たない作業者が行った。338 件の入力文に適用された推論規則の数は 1444 件である。1444 件のうち、推論が行われたものは 1374 件である。この中で正しく推論が行われたものは 952 件で推論が行われた 1374 件の 69.3 %にあたる。推論で生成された帰結文の一部を表 2 に示す。

間違いは大きく 2 種類に分類された。一つは語義の粒度の問題であり、もう一つは推論規則における問題である。

4.1 語義の粒度の問題

不適切と判断された帰結文 422 件のうち、290 件(68.7 %)は語義の粒度に起因する誤りであった。

語義の問題ははじめから予想された問題であり、それ以外の問題に焦点を当てるために、我々は語義情報が付いた入力文を用いた。にも関わらず語義の粒度に起因する誤りが生じたのは岩波国語辞典による語義の分類では粒度が適切でない場合があり、それにより誤った帰結文が多数生成されたためである。

- (8) a. 出来上がる (34973-0-0-0-0)
- b. すっかりできる。完成する。また俗に、酒を飲んで、よい機嫌に酔った状態になる。
- c. Xガ出来上がる→Xガすっかりできる
- d. Xガ出来上がる→Xガ完成する
- e. Xガ出来上がる→Xガ酒を飲む
- f. Xガ出来上がる→Xの機嫌がよい
- g. Xガ出来上がる→Xガ良い機嫌に酔う

「出来上がる」のこの語義に対して「X=料理」であるならば(c)(d)は成立するが、(e)(f)(g)は不適切となる。逆に「X=部

(注1) : <http://www.gsk.or.jp/catalog.html>

表 2 推論規則により生成された帰結文の例

ラベル	入力文	帰結文	関係
正解	首相が委員会を設ける	首相が委員会を構える	同義・上位
正解	彼が字を面用紙に書く	彼が字を面用紙に記す	同義・上位
正解	JR が満足する	JR が作られる	付帯状況
正解	タイガースがカーブに負ける	タイガースが戦う	前提条件
正解	国民が与党を支持する	国民が与党に賛成する	付帯状況
正解	世界が平和を求める	世界が平和を得る	目的
正解	母が演歌を歌う	母が節を母の声につける	手段
正解	物価上昇が現状に留まる	物価上昇が現状から進まない	結果 (状態)
正解	曖昧さが誤解を招く	曖昧さが誤解を引き起こす	同義・上位
正解	店が利益を見込む	店が利益を予想する	前提条件
不正解	判定がルールに沿う	判定が進む	同義・上位
不正解	罰金が 5 万円で済む	罰金が解決する	同義・上位
不正解	交通事故が生ずる	交通事故が生える	同義・上位
不正解	罪が消える	罪が聞こえない	結果 (状態)
不正解	彼が友人を訪ねる	彼が友人に行く	付帯状況
不正解	アメリカが兵隊を展開する	アメリカが兵隊を開く	付帯状況
不正解	母親が母乳を乳児に与える	母親が母乳を乳児に渡す	手段
不正解	息子が許可を待つ	息子が早く来るのを望む	上位・同義
不正解	学力が分かる	十分な理解力が学力にある	付帯状況
不正解	死相が顔に表れる	死相の様子が顔に出る	上位・同義
未生成	上司が亡くなる	—	—
未生成	彼が気持ちを踏み切る	—	—
未生成	日銀が金利を 2 パーセントに引き上げる	—	—

長」ならば (c)(d) は不適切だが, (e)(f)(g) は成立する。つまりこれらは推論規則としては別々に対処されるべきである。このような問題は今回の実験では語釈文から推論規則の抽出を行ったが, 複数の語釈文からなる場合に多く見られた。さらに以下のような例もある。

- (9) a. 要る (2963-0-0-3-0)
 b. それが, なくてはならないものになる。必要だ。かかる。
 c. X が 要る → X が なくてはならない
 d. X が 要る → X が 必要だ
 e. X が 要る → X が かかる

この場合, 「X=時間」であるならば (c)(d)(e) は正しい帰結文を生成する。しかし「X=車」の場合 (c)(d) は正しい帰結文を生成するが (e) が生成する「車がかかる」は言語表現として不適切である。

以上のように同一の語義として定義されてはいるものの, 項に入力される変数によっては関係が成立しないものが多く存在した。この問題を解決するには適切な帰結文を生成する語義の粒度に分解する必要がある。しかし, 粒度を推論の適格性が保証されるように最適化をするということは困難な問題であることが予想される。さらに「要る」の例が示唆するように最適な粒度が格要素に依存するとならば, 格要素で独立になるような粒度に語義を分解すること自体が不可能である。さらに粒度を細かくすれば多義性解消にかかる負荷も大きくなる。

我々が無駄なコストを生じることなくこのような粒度の問題を解決するためには, 適切な粒度を求めるのではなく, 多少の曖昧性を許容しつつ推論制御を行っていくことである。そのためにまず有効であると考えられるのは帰結文中の動詞と格要素の組み合わせの適格性を言語モデルで判断する方法が考

えられる。語義の誤りと思われる間違い 286 個中 134 個 (46.9%) が統語的に正しいながらも適格性が低いと思われるものだった。実際に動詞と格要素の組み合わせの適格性を用いて推論を制御するものとして Pantel ら [11] がある。Pantel らは推論規則に代入される名詞のペアを同時選択制限と独立選択制限の 2 つのモデルで推論を制御することを試みている。また SemEval-2007 の結果からも動詞と格要素の適格性を用いることの有効性が示唆されるところだが (Giuliano et al. [5]), 今後大規模な実験によって検証していく必要がある。

しかしこのような言語モデルをもってしてもまだ正しい推論が行えるわけではない。

- (10) a. 伸びる (40174-0-0-0-1)
 b. 育ったり引っ張られたりして, 長くなる, また, 真っすぐになる。
 c. X が 伸びる → X が 育つ
 d. X が 伸びる → X が 引っ張られる
 e. X が 伸びる → X が 長くなる
 f. X が 伸びる → X が 真っすぐになる

この場合「X=髪」であるならば, (e) のみが正しい帰結文を生成する。(c) に関しては先の言語モデルを用いて解決可能であるが, (d)(f) に関しては適格性を考慮しても解決できない。

4.2 前件と入力文の照合時の問題

(格交替による誤り)

格交替が原因で作成した推論規則と照合しない入力文が 338 件中 11 件存在した。以下に一致しなかった推論規則の例を挙げる。

- (11) a. X が Y に 怒る → X が Y に 腹を立てる
 b. 入力文: 父が息子を怒る

- (12) a. W ガ X ヲ Y カラ Z へ 移す → W ガ X ヲ Y カラ Z へ 動かす
b. 入力文：銀行が視点を駅前に移す
- (13) a. X ガ Y ト 代わる → X ガ Y ト 交替する
b. 入力文：祖父が父に代わる

以上のように推論規則が規定する項と実際に入力される文の項が異なる場合がある。これはその動詞に格交替があることに起因する。例えば「怒る」の推論規則では二格が動作の対象となっているが入力文ではヲ格が動作の対象になっている。

このような推論規則と実際の入力文間における表面的な項構造の異なりは、何らかの形で吸収する必要がある。

これには2つの方法が考えられる。一つは格交替のパターンをコーパスから自動獲得するという方法である。もう一つは竹内ら[14]の語意概念構造(LCS)に基づく事態オントロジーの構築で行っている動詞の意味分類を利用して同一のカテゴリに属していると思われる動詞に同じ格パターンを与え、全体としての記述を簡易化する方法が考えられる。

(AのBが不適切な表現となる誤り)

帰結文が不適切となったもの492件中28件(5.7%)が(14)(15)で示されるような間違いであった。

- (14) a. 消える(10304-0-0-0-3)
b. X ガ 消える → X の形ガ 見えない
c. 姿が消える → *姿の形が見えない
- (15) a. 越える(16933-0-0-1-0)
b. X ガ Y ヲ 越える → X ガ Y の上ヲ 過ぎる
c. 旅人が峠を越える → *旅人が峠の上を過ぎる

これは「AのB」の表現において、Bが不要な場合に起こる。

我々は推論知識を獲得するうえで、見出し語項構造と定義語項構造の関係が一般的に成立するように獲得を行った。しかし、「B」の概念が「A」に含まれている場合は冗長な表現になってしまう。今後我々は一般的な名詞に関しては推論知識を用いて、「AのB」が不適切になるものには縮退させるような推論制御を行っていく必要がある。

また今回の実験では観察されなかったが以下のようなメトニミーにおける問題も考えられる。

- (16) a. 沸かす(55844-0-0-2-0)
b. X ガ Y ヲ 沸かす → X ガ Y ヲ 煮立たせる
c. 僕がポットを沸かす → *僕がポットを煮立たせる

このように、ある動詞においてはメトニミーとして成り立っていたものが、帰結文では成り立たないという状況が往々にしてあると考えられる。また項の一部が省略されている場合にも同様の問題が起こる可能性がある。このような一般的でないものが項に代入された述語項構造は、言語モデルによるフィルタリングで不適切な文と判断することはできるが、例えば「ポットの水を煮立たせる」という帰結文を今回の推論規則の枠組みで行うことはできない。

4.3 推論規則の抽出における問題

不適切な帰結文を導いた推論規則422件中58件(13.7%)は推論規則抽出自体が適切に行われていないものであった。

- (17) a. 望む(40112-0-1-1-0)
b. 自分の将来や他人・世間の事について、こうなればよい、そうしたいと思う。
c. X ガ Y ヲ 望む → Y ガ こうなる
d. X ガ Y ヲ 望む → X ガ そうする

このように、語釈文中の指示詞が指すものが文中に存在しない場合がある。これをそのまま適用してしまうと不適切になってしまう。今後、作業の対象としてこのような指示詞をどのように扱うのか検討する必要がある。

4.4 その他予測される誤り

(項の曖昧性問題)

今回我々は動詞に必須と思われる項のみに対し述語項構造を作成し、それを入力文とした。しかし実際に適用する場合には単一の項に複数の意味解釈を行えることが少なくない。以下のような例を考える。

- (18) a. 行く(52744-0-0-1-0)
b. X ガ Y ニ 行く → X ガ Y ニ 進む

この推論規則における二格は「父が京都に行く」のような移動の着点を表す二格である。しかし「父が釣りに行く」のように二格が動作の目的を表す場合も存在する。この場合、前者の例では推論規則は適切に扱われるが、後者の場合は推論規則を適用できない。このように語義だけでなく、項自体にも曖昧性が生じてくることが予想される。

(上位語に対する語釈文の性質)

今回我々は語釈文の主辞が見出し語の上位語にあたるという性質を利用して、同義・上位の関係を作成した。しかし語釈文中の主辞が上位語とはみなし難いものも存在する。

- (19) a. 見る(50038-0-0-1-0)
b. 視覚を動かして、ものの存在・形・様子・内容をとらえる。目で認める。
c. X ガ Y ヲ 見る → X ガ Y の存在・形・様子・内容ヲ とらえる
- (20) a. 知る(25357-0-0-0-1)
b. その存在を認め、またはその状況や内容、意味・価値を、こうだとつかむ。
c. X ガ Y ヲ 知る → X ガ Y の状況や内容・意味・価値ヲ つかむ

これらの語釈文は見出し語の動作を直接的に説明している。このような推論規則はそもそも語釈文の性質自体が他とは大きく異なり、見出し語項構造と定義語項構造を関係付けることは本来の目的から大きく外れる。しかし、以下のような関係を確認することはプラン認識などにおいて非常に重要である。

- (21) a. 彼が事件を見る → 彼が事件を知る(結果状態)

このような関係を得るためには、語彙概念レベルでの関係認識を行う必要があると考えられる。これに関しては竹内ら[14]が行っているLCSに基づく動詞分類を利用して語意概念間の関係を抽出できると考えている。

5. 関連研究

知識を辞書から抽出したものとしては、Mindnet [13] や鍛冶ら [6] がある。これらの研究はいずれも語釈文を自動解析し、全自動で知識を抽出することを試みている点では我々の試みに比べて有利である。ただし、Mindnet [13] では動詞間の関係はある程度分類できているものの項の対応関係についてはとらえられていない。また鍛冶ら [6] の手法については項の対応付けにアプローチしているが、扱っている関係が上位・言い換えに限定されている。

述語間の意味関係を人手で記述した代表的な資源に WordNet [4] と FrameNet [2] がある。WordNet には本稿で述べたような分類に近い意味関係が記述されている。FrameNet はフレーム間にも限定的ではあるが同様の意味関係が記述されている。

自然言語表現を用いた推論については、MacCartney et al. [9] がある。MacCartney et al. [9] では語の monotonicity に基づいて含意関係の認識を行っている。ただし、語の対応をとる際に edit distance を用いており統語的に近い文に対しては含意関係の認識が行えるが統語構造が異なったものには対応していない。

6. 終わりに

本稿では自然言語表現に基づく推論規則を国語辞典から抽出した。そしてその推論規則に入力文を適用した場合に起こる問題を、語義の粒度、推論規則における問題を中心に分類し、今後我々が推論を適切に行う為に必要な技術的な課題を論じた。

謝 辞

本研究は、文化省科研費基盤研究 (B) 「語彙意味論に基づく言い換え計算機構の工学的実現と言い換え知識獲得への応用」(17300047, 代表：乾健太郎) の支援を受けている。

文 献

- [1] 青山桜子, 阿部修也, 大西良明, 乾健太郎, 松本裕治. 事態間関係の獲得のための動詞語釈文の構造化. 言語処理学会 第 13 回年次大会, 2007.
- [2] Collin F. Baker, Charles J. Fillmore, and John B. Lowe. The berkeley framenet project. In *Proceedings of COLING-ACL*, pp. 86–90, 1998.
- [3] Ido Dagan, Oren Glickman, and Bernardo Magnini. The pascal recognising textual entailment challenge. In *Proceedings of Pascal Challenge Workshop on Recognising Textual Entailment*, 2005.
- [4] Christiane Fellbaum, editor. *WordNet-An Electronic Lexical Database*. The MIT press, 1998.
- [5] Claudio Giuliano, Alfio Glizzo, and Carlo Strapparava. Fbk-irst:lexical substitution task exploiting domain and syntagmatic coherence. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, pp. 145–148, 2007.
- [6] 鍛冶伸裕, 河原大輔, 黒橋禎夫. 国語辞典とコーパスを用いた用言の言い換え規則の学習. 言語処理学会 第 8 回年次大会, 2002.
- [7] P. Kingsbury and M. Palmer. From treebank to propbank. In *Proceedings of the Third International Conference on Language, Resources and Evaluation*, 2002.
- [8] 小町守, 飯田龍, 乾健太郎, 松本裕治. 述語項構造と共参照関係のアノテーション:naist テキストコーパス開発の経験から. 言語処理学会 第 13 回年次大会, 2007.
- [9] Bill MacCartney and Christopher D. Manning. Natural logic for textual inference. In *Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing*, pp. 193–200, Prague, June 2007. Association for Computational Linguistics.
- [10] Adam Meyers, Ruth Reeves, Catherine Macleod, Rachel Szekely, Veronika Zielinska, Brian Young, and Ralph Grishman. The nombank project:an interim report. In *Proceedings of HLT-NAACL Workshop on Frontiers in Corpus Annotation*, 2004.
- [11] Patrick Pantel, Rahul Bhagat, Bonaventura Coppola, Timothy Chklovski, and Eduard Hovy. ISP: Learning inferential selectional preferences. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Proceedings of the Main Conference*, pp. 564–571, Rochester, New York, April 2007. Association for Computational Linguistics.
- [12] 白井清昭. Senseval-2 日本語辞書タスク. 自然言語処理, 第 10 巻, pp. 3–24, 2003.
- [13] 鈴木久美, Gary Kacmarcik, Lucy Vanderwende, Arul Menezes. Mindnet/mnex:意味関係データベースの自動構築と解析のためのツール. 言語処理学会 第 11 回年次大会 発表論文集, 2005.
- [14] 竹内孔一, 乾健太郎, 藤田篤, 竹内奈央. 語彙概念構造に基づく事態上位オントロジーの構築. 言語処理学会 第 13 回年次大会, 2007.