日本語 Textual Entailment のデータ構築と自動獲得した類義表現に基づく推論関係の認識

小谷 通隆 †　　柴田 知秀 †　　中田 貴之 ††　　黒橋 禎夫 †

† 京都大学大学院 情報学研究科　†† 東京大学大学院 情報理工学系研究科
odani@nlp.kuee.kyoto-u.ac.jp, shibata@nlp.kuee.kyoto-u.ac.jp,
nakata@koseki.t.u-tokyo.ac.jp, kuro@i.kyoto-u.ac.jp

1　はじめに

近年 Textual Entailment の認識（Recognizing Textual Entailment,RTE）が注目を集めている．RTE とは，以下に示すような **text**(以下 **t** と略記) と **hypothesis**(以下 **h** と略記) を与え，**t** から **h** が推論されればYES，そうでなければ NO とする判断をシステムで行うタスクである．

t: About two weeks before the trial started, I was in Shapiro's office in Century City.

h: Shapiro works in Century City.

推論判定: YES

この認識は，質問応答 (QA) や情報検索 (IR) などテキストの内容の理解が必要な高次タスクにおいて重要な問題となる．

RTE の実現を目指す試みとして，共通に大規模な正解データを作成して参加者がそれぞれのシステムの解析結果を提出し，結果を比較検討する評価型ワークショップが活発に行われている [1]．2007 年に行われた PASCAL-RTE3 では，上記のような **t** と **h** のペアに推論判定の正解を人手で付与した英語の Textual Entailment データ (TE データ) が 1600 個作られ，成績トップのシステムの精度は 80 %であった．このとき作られた TE データは，QA，IR などの評価データやシステムの出力などをもとに作られており，このデータを正しく推論できるエンジンが出来れば直接 QA，IR などの精度向上につながる．

しかし，このようにして作られたデータは複数の要因が複合的に作用して推論関係が導かれる事例が多く，問題点を議論しにくい．そこで本稿では，1 つ，多くとも 2 つの要因によって判断できる事例のみを集めた日本語の TE データを構築する．さらに類義表現に基づく推論関係を認識する手法を提案し，構築したデータを用いて検証する．

2　日本語 Textual Entailment のデータ構築

2.1　推論関係の分類

まず，推論の要因を「包含」「語彙 (体言)」「語彙 (用言)」「構文」「推論」の 5 つに分類し，さらにそれぞれの分類に下位分類を設けた．そして，各分類に当てはまる事例を作成した．表 1 に作成した TE データの分類とそこに属するデータ数を示す．分類間にデータの偏りがないようにして約 2700 セットのデータを構築している[1]．以下に分類の説明と例を示す．

包含　**t** に **h** がほぼそのまま含まれているようなデータである．例として下位分類「補文」に含まれるデータを示す．

(1)　ビルが選ばれたのは不幸だ．→ ビルが選ばれた．

語彙 (体言)　**t** にある名詞の意味や性質から **h** の真偽の情報が与えられるようなデータである．例として下位分類「定義的」に含まれるデータを示す．

(2)　埼玉は内陸に位置する．
→ 埼玉は海に面していない．

語彙 (用言)　**t** にある用言の意味や性質から **h** の真偽の情報が与えられるようなデータである．例として下位分類「言い換え」に含まれるデータを示す．

(3)　何をやっても眠気がとれない．→ 何をしても眠い．

構文　**t** と **h** が構文の形を変えただけの関係になっているようなデータである．例として下位分類「主語の変換」に含まれるデータを示す．

(4)　化粧は女を化かす．女は化粧で化ける．

[1] 本研究においてはメタファーやメトニミー，慣用表現に関するものは扱わない．

表 1: 作成した TE データの分類と属するデータ数

分類	下位分類	◎	○	△	×	計	合計
包含	節	43	5	1	43	92	263
	並列	34	1	1	34	70	
	補文	25	2	4	25	56	
	名詞句	21	3	0	21	45	
語彙 (体言)	定義的	92	60	19	46	217	557
	同義語	57	18	5	45	125	
	下位 → 上位	35	3	3	20	61	
	上位 → 下位	9	8	4	25	46	
	対義語	18	5	1	14	38	
	名詞の格関係	18	5	0	13	36	
	性質	7	9	6	12	34	
語彙 (用言)	言い換え	126	87	10	74	297	767
	前提的	66	61	16	22	165	
	含意	45	67	19	10	141	
	副詞	33	21	2	21	77	
	内包	26	9	5	7	47	
	対義語	14	4	3	19	40	
構文	主語の変換	60	7	7	35	109	219
	複文の変換	13	24	4	15	56	
	強調構文	19	3	1	11	34	
	名詞句	10	0	0	10	20	
推論	結果 → 原因	48	93	67	28	236	868
	原因 → 結果	43	95	51	41	230	
	省略の類推	65	51	6	19	141	
	副助詞+対応	23	27	18	28	96	
	副助詞+一般化	17	13	7	10	47	
	時間軸 · 数量	20	5	1	10	36	
	順接 · 逆接+一般化	10	12	2	10	34	
	順接 · 逆接+具体化	9	5	4	10	28	
	間接発話行為	3	8	8	1	20	
		1009	711	275	679	2674	

一般的な推論　上記の分類には当てはまらない類の推論形式のデータである．例として下位分類「結果 → 原因」に含まれるデータを示す．

(5)　霜柱がたった．→ 気温が低くなった．

2.2　推論判定の基準

PASCAL-RTE3 で作られた TE データでの推論判定は YES か NO の 2 値である．しかし，実際の推論判定においては 2 値で分類すると判断が困難である場合も多いので，以下の 4 種類を用意した．

◎: t が真であったとき **h** が必ず真であるといえる場合

(6)　彼は人間である．→ 彼は哺乳類である．

○: t が真であったとき **h** が正しいと常識的には考えられる場合

(7)　彼はオムレツを作った．→ 彼は卵を使った．

△: t が真であったとき **h** が真である可能性がある程度考えられる場合

(8)　信号は赤ではなかった．→ 信号は青だった．

×: t が真であったとき **h** が全くの誤りだとわかる場合

(9)　彼は読書が好きだ．→ 彼は読書が嫌いだ．

3　自動獲得した類義表現に基づく推論関係の認識

本稿では，まず表 1 の分類のうち「語彙 (体言)」に含まれる語や句の同義・上位下位関係に基づく推論関係の認識を行う．

(10)　太郎は人間だ．→ 太郎は哺乳類だ．

(11)　哺乳類は肺呼吸だ．→ 人間は肺呼吸だ．

上記の推論ではともに「人間」の上位語が「哺乳類」であるという知識が使われている．また，同じ知識を用いながらもそこから導かれる推論関係は多様で，(10) では **t** の「人間」が **h** では上位語「哺乳類」に，(11) では **t** の「哺乳類」が **h** では下位語「人間」に置き換わっている．すなわち，語や句の関係を知識として獲得し，さらにそれによって導かれる推論パターンを認識しておく必要がある．3.1 節に語や句の関係の獲得，3.2 節では推論パターンを用いた自動推論システムの構築について述べる．

3.1　語や句の同義・上位下位関係の自動獲得

3.1.1　国語辞典とウェブテキストからの自動獲得

我々は，国語辞典とウェブテキストを利用して類義表現を自動的に獲得している [3]．まず，国語辞典から語の定義文の文末パターンを認識することで一般的な語の基本的な同義・上位下位関係を獲得できる．

夕食: 夕方の食事．夕飯．
⇒夕食 → 食事 (定義文 1 文目主辞)
夕食＝夕方の食事＝夕飯 (定義文が短い)

しかし，国語辞典から獲得できる関係は基本的な語彙に関するものに限定されるため，ウェブテキストから「対称な括弧関係」に着目して「ケータイ＝携帯電話」，「放射性同位体＝ RI」といった語彙の同義表現を獲得している．以上の手法で 90 %を越える精度で語や句の同義・上位下位関係を獲得している．

本稿ではこれに加え，国語辞典からの自動獲得の手法を用いて Wikipedia からも以下のような関係が獲得している．

アメリカン・ショートヘア:
…強力で大きな脚を持つ中型の猫．鼻・口の部分はほぼ四角い形をしており，…
⇒アメリカン・ショートヘア → 猫

3.1.2 名詞句の上位表現認識

一般に，名詞句においてその主辞が名詞句の上位表現となる．例えば,「チョコレート工場」の上位表現は「工場」であり,「海の生物」の上位表現は「生物」である．しかし，地名を表す語が連続する「中国遼寧省」や時間を表す語が連続する「12 月上旬」では，主辞以外の「中国」,「12 月」という語を名詞句の上位表現とみなす．

3.2 推論自動判定システムの構築

3.2.1 語や句の同義関係による推論パターン

t 中の表現を同義表現に置き換えたものは，**t** から推論可能なものである．つまり，ある表現 X の同義表現が X' であるときに，語や句の同義関係による推論パターンを記述すると例えば以下のようになる．

推論パターン: X を V する →X' を V する
例: 夕食 (X) を食べる → 夕飯 (X') を食べる

3.2.2 語や句の上位下位関係による推論パターン

語や句の上位下位関係が導く推論は単純ではなく，語が現れる文のタイプや語の文中での役割によって変化する．そこで，文を「行為を表す文」と「性質・状態を表す文」に分け，ある表現 S(Specific) の上位表現が G(General) であるとして推論パターンを記述する．

行為を表す文 **t** 中の表現を上位表現に置き換えたものは，**t** から推論可能なものである．

推論パターン: S に V する →G に V する
例: インコ (S) に餌をやる．→ 鳥 (G) に餌をやる．

性質・状態を表す文 **t** の述語以外の場所に現れる表現については **t** の表現を上位表現に置き換えたものが，**t** から推論可能なものである．

推論パターン: G は V する →S は V する
G は N である →S は N である
例: 鳥 (G) は翼を持つ．→ 燕 (S) は翼を持つ．

しかし，述語に現れる表現は「行為を表す文」と同様に上位表現に置き換えたものが，**t** から推論可能なものである．

推論パターン: N は S である →N は G である
例: カツオは海の生物 (S) である．
→ カツオは生物 (G) である．

語や句の上位下位関係による推論パターンを利用するためには文のタイプを判断する必要があるが．これは難しい問題である．そこで，本稿では性質・状態を表す文を以下の条件によって認識する．

時制 現在時制である
述語 判定詞文である，または述語が形容詞か動詞「ある」「持つ」である[2]，もしくは可能表現である

以上の条件を満たすものを「性質・状態を表す文」であるとし，それ以外を「動作を表す文」であるとする．

4 実験結果と考察

2 節で構築した TE データを用いて，構築したシステムの検証を行った.「語彙 (体言)」の「同義語」「下位 → 上位」「上位 → 下位」，計 189 事例を対象とした (複数の下位分類に属する 43 事例は除外した)．人手による判定は 4 段階の評価のうち◎と○を YES(Y)，△と×を NO(N) とした．結果を表 2 に示す．

人手での評価が YES である事例をシステムが誤って NO と判断した原因は，大半が語や句の関係が獲得できていないことが原因であった．例えば次の例は知識不足から推論関係を認識できない．

同義語
(12) お相撲さんを見た．→ 力士を見た．

また，同義語に含まれるデータにはその文脈においてのみ同義関係が成り立つ例も含まれている．

同義語
(13) 懐かしい顔が揃った．→ 懐かしい人が揃った．

一般的には「顔」と「人」は同義であるとはいえないため，このような事例は困難である．

同様に，上位下位関係にある表現が文脈によっては同義関係にあるとみなせる事例が存在する．

上位 → 下位
(14) ボタンを糸で縫いつけた．→ ボタンをボタン糸で縫いつけた．

(14) は「行為を表す文」であるので，**t** の「糸」を下位語「ボタン糸」に置き換えた **h** は **t** から推論することはできないはずである．しかしこの事例では，**t** の文脈での「糸」は「ボタン糸」であることは明らかであり,「糸」と「ボタン糸」が同義関係にあるとみなせる．

[2]「性質/効果がある」「性質/能力を持つ」などを想定している．

表 2: 実験結果

語彙 (体言)「同義語」

		正解		計
		Y	N	
システム	Y	10	0	10
	N	55	42	97
計		65	42	107

語彙 (体言)「上位 → 下位」

		正解		計
		Y	N	
システム	Y	3	0	3
	N	10	24	34
計		13	24	37

語彙 (体言)「下位 → 上位」

		正解		計
		Y	N	
システム	Y	8	2	10
	N	20	15	35
計		28	17	45

人手での評価が NO である事例をシステムが誤って YES と判断した原因は，名詞句の上位表現認識に問題があった．本稿の手法では「もみじ狩り」の上位表現が「狩り」,「小京都」の上位表現が「京都」になってしまい，このような比喩的な表現について考慮する必要がある．

次に文のタイプの判断の妥当性について考察する．本稿では現在時制であることと述語を調べることで「性質・状態を表す文」であることを認識し，データベースにおいては判断を誤る例がなかった．しかし，(15) のように判断を誤る例が存在する．

(15) 鳥は空を飛ぶ．→ ツバメは空を飛ぶ．

このような「性質・状態を表す文」であると認識できる事例を集め，認識方法を検証する必要がある．

また，今回構築した TE データには例が含まれていないが，特定の語に修飾されることで，**t** の語を上位語ではなく下位語に置き換えないといけない例も存在する．

上位→下位

(16) 全ての学生が避難した．
→ 全ての高校生が避難した．

このように 3.2 節で記述した推論パターンは全てを網羅したものではなく，把握されていない事例の収集が今後の課題である．

5 関連研究

MacCartney らは Natural Logic という理論を RTE システムに導入している [2]．Natural Logic とは **t** の表現を，指す内容が変わらない (編集後の文が **t** に包含される) ように少しづつ編集を繰返し，それによって **h** がつくり出せるかどうかを調べることで Textual Entailment を判断している．

日本語における RTE の実現については乾らによって研究がすすめられている [4]．乾らは，推論とは述語項構造を単位とした事態間の関係を認識することであるとし，事態に関する知識の総体を広く事態オントロジーと呼び，その構築を進めている．その構築するために，(a) 国語辞典の動詞語釈文の構造化，(b) 語彙概念構造に基づく事態上位オントロジーの開発，(c) コーパスからの知識獲得の 3 方向からアプローチしている．

6 おわりに

本稿では日本語 TE データの構築し，類義表現に基づく推論関係を自動認識する手法を提案した．語や句の関係についての知識の自動獲得と推論パターンを記述することによってを実現している．

今後の課題としては，語や句の関係をさらに幅広く獲得し，また推論パターンを整理することでシステムの精度をあげることである．また，表 1 の他の分類に属する事例を検討し，その扱いを模索していく．

謝辞

推論データの作成をしていただいた石川真奈見氏，二階堂奈月氏に感謝いたします．また，本研究の一部は情報通信研究機構の補助によって行なわれました．ここに記して感謝の意を表します．

参考文献

[1] Danilo Giampiccolo, Bernardo Magnini, Ido Dagan, and Bill Dolan. The Third PASCAL Recognising Textual Entailment Challenge. In *Proceeding of ACL-PASCAL Workshop on Textual Entailment and Paraphrasing*, pp. 1–9, 2007.

[2] Bill MacCartney and Christopher D.Manning. Natural logic for textual inference. In *Proc. WTEP 2007*, 2007.

[3] Tomohide Shibata, Michitaka Odani, Jun Harashima, Takashi Onishi, and Sadao Kurohashi. SYNGRAPH: A flexible matching method based on synonymous expression extraction from an ordinary dictionary and web corpus. In *Proc. of IJCNLP2008*, 2008.

[4] 乾健太郎. 事態オントロジー:言語に基づく推論のためのコトに関する基本知識. 言語処理学会第 13 回年次大会ワークショップ「言語的オントロジーの構築・連携・利用」発表論文集, pp. 27–30, 2007.