

日本語 CCG パーザに基づく意味解析・推論システムの提案

田中 リベカ¹ 峯島 宏次^{1,4} Pascual Martínez-Gómez^{1,2,4} 宮尾 祐介^{2,3} 戸次 大介^{1,2,4}

¹ お茶の水女子大学 ² 国立情報学研究所 ³ 総合研究大学院大学 ⁴ JST, CREST

{tanaka.ribeka, bekki}@is.ocha.ac.jp, koji.mineshima@ocha.ac.jp,
{pascual, yusuke}@nii.ac.jp

1 はじめに

本稿では、日本語 CCG パーザの導出木をラムダ計算に基づいて論理式へと変換し、高階論理に基づく自動推論を行う統合的システムを提案する。

論理に基づく意味表現は、係り受け解析や述語項構造解析などの従来の解析を超える深い意味解析、またそれに基づく推論システムを実現する上できわめて重要である。英語を対象とする研究では、組合せ範疇文法 (CCG) のような現代的な統語論に基づくパーザの出力結果を述語論理の論理式に変換し、推論を行う手法が提案されており (Bos, 2008), 質問応答や含意関係認識などのタスクに広く用いられている。

しかし、日本語に対する本格的な実装はまだ存在していない。とりわけ、日本語に特有の統語構造を論理式へと変換するプロセスは自明なものではなく、適切な意味表現を得るためには、日本語 CCG の導出木に対する意味合成プロセスの詳細な分析が不可欠である。

本稿では、Mineshima et al. (2015) で提案された、英語 CCG パーザの導出木を高階論理の論理式へと変換し自動推論を行うシステムを拡張し、日本語 CCG パーザ (能地ら, 2015) に基づく意味解析と推論の統合的システムを提案する。またこのシステムを含意関係認識に適用し、量化や否定などの論理推論、モダリティや内包的述語などを伴う複雑な意味推論が正確かつ効率的に実現可能であることを示す。システムの評価実験には、日本語意味論テストセット JSeM (Kawazoe et al. 2015) を用いる。

2 システムの概要と関連研究

この節では、含意関係認識タスクを念頭に置いて、システムの概要を示す。含意関係認識は、質問応答・情報検索・テキスト要約・機械翻訳などの幅広い応用が期待される意味処理のためのタスクであり、典型的には、文 T から仮説 H が推論されるかどうかを判定する問題である。たとえば以下の例は、 T と H は矛盾するため、含意関係が成立しない例である。

T : 太郎は7時の電車に乗り損なった。

H : 太郎は7時の電車に乗った。

T と H において一致している単語数などの文の表層的な情報に着目する手法では、量化や否定、テンス・アスペクト・モダリティなど複雑な意味情報を伴う推論を体系的に行うことが難しいことから、英語圏では様々な論理ベースの手法が提案されてきた。論理的アプローチでは、文 T と H の意味は論理式によって記述され、文の間の含意関係の成立は $T \rightarrow H$ が定理として証明可能であるか否かによって判定される。

論理的手法に基づく代表的なシステムとしては、談話表示理論 (DRT) による意味表示と一階述語論理 (FOL) による自動定理証明を組み合わせた Boxer/Nutcracker (Bos, 2008) がある。FOL に基づくシステムは、FOL の汎用的な定理証明器及びモデル構成器を利用できるという利点がある一方で、一般化量子子や様々な内包的表現など、自然言語に特徴的な複雑な意味的知識を記述することが難しいという問題がある。そのため、自然言語の形式意味論では高階論理 (HOL) を用いて意味記述を行うことが標準的となっている。また、近年の Natural Logic の研究が明らかにしているように (van Benthem, 2008), 自然言語推論には一般化量子子など FOL を超える部分がある一方で、量子子の入れ子構造に制限があるなど、FOL の定理証明を用いなくとも処理可能な部分が大半を占めると予想される。高階論理には、FOL のような汎用的な定理証明器が存在しないものの、自然言語推論に特化した自動推論を実装する余地は残されている。Mineshima et al. (2015) は、高階論理に基づく推論システムが、FraCaS データセット (Cooper et al, 1994) による評価実験において、FOL に基づく Boxer/Nutcracker よりも優れた正答率・速度を示すことを明らかにしている。また、Abzianidze (2015) は、Natural Logic に基づくタブロー系による高階論理推論システムを実装し、SICK データセット (Marelli et al, 2014) による評価実験において、トップの RTE システムに匹敵する正答率を達成している。

本稿で提案するシステムの構成を図1に示す。統語解析は、日本語 CCG Bank (Uematsu et al., 2015) に基づいて実装された日本語 CCG パーザ (能地ら,



図 1: CCG に基づく意味解析・推論システムの構成.

$$\frac{X/Y : f \quad Y : a}{X : fa} > \quad \frac{Y : a \quad X \setminus Y : f}{X : fa} <$$

図 2: CCG の組合せ規則の例 (関数適用規則).

2015) を用いて行う. 意味表示は高階論理 (HOL) の表現を用い, Mineshima et al. (2015) に基づいて, 高階論理・型理論の証明支援系である Coq を用いて自動証明を行う. Coq は半自動の証明支援系として知られているが, FOL 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, 自然言語の効率的な自動推論が可能となると予想される. 高階論理の推論システム (公理及び証明探索法) は日本語・英語で共通であり, Mineshima et al. (2015) に準じたものを使用した.

3 CCG に基づく意味合成

3.1 日本語 CCG と合成的意味論

CCG (Steedman, 2000; Bekki, 2010) はいわゆる語彙化文法の一つで, 統語構造から意味表示 (論理式) を合成的に導出するのに適した文法体系として知られている. CCG の統語範疇は, 基底範疇 (S , NP など) と二項演算子 $/$, \setminus を用いて再帰的に定義される関数的範疇の二種類がある. 語の統語的な振る舞いは, 少数の組合せ規則として定式化され, 各組合せ規則は, 同時に語の意味合成の計算方法を指定する. 図 2 に組合せ規則の例を示す. たとえば, $>$ 規則により, X/Y という形の統語範疇及び意味 f をもつ語は, その右側にある Y という形の統語範疇及び意味 a をもつ語と結びつき, X という統語範疇及び意味 fa をもつ語が形成される. 語の意味はラムダ項によって表現される. 組合せ規則に基づいて得られたラムダ項を β 簡約することで最終的な意味表示 (論理式) が得られる.

各語に対する意味割り当てには, NP (名詞句) や $S \setminus NP$ (自動詞) などの統語範疇単位で意味を指定する図式的な意味割り当て (テンプレート) と, 量化や否定表現などの限られた範囲の論理語・機能語に対して単語単位で意味を指定する語彙項目の二種類を使用する. 主要な統語範疇に対する意味割り当てのテンプレートを表 1 に示す¹. 動詞句と名詞句に対する意味表示の詳細については以下で説明する.

¹ $\lambda X \lambda Y.Z$ を $\lambda XY.Z$, $(Px)y$ を $P(x, y)$ と表記している. Base は各語の基底形が現れる位置を表す.

品詞	統語範疇	意味表示
名詞	NP	$\lambda NF. \exists x (N(\text{Base}, x) \wedge F(x))$
自動詞	$S \setminus NP_{ga}$	$\lambda QK. Q(\lambda I. I, \lambda x. \exists v (K(\text{Base}, v) \wedge \text{Nom}(v) = x))$
他動詞	$S \setminus NP_{ga} \setminus NP_o$	$\lambda Q_2 Q_1 K. Q_1(\lambda I. I, \lambda x_1. Q_2(\lambda I. I, \lambda x_2. \exists v (K(\text{Base}, v) \wedge \text{Nom}(v) = x_1 \wedge \text{Acc}(v) = x_2)))$
副詞	S/S	$\lambda SK. S(\lambda I v. K(\lambda v'. (J(v') \wedge \text{Base}(v')), v))$
形容詞	NP/NP	$\lambda QNF. Q(\lambda Gx. N(\lambda y. (\text{Base}(y) \wedge G(y)), x), F)$

表 1: 統語範疇に対する意味割り当てのテンプレート.

$$NP^* = ((E \rightarrow \text{Prop}) \rightarrow E \rightarrow \text{Prop}) \rightarrow (E \rightarrow \text{Prop}) \rightarrow \text{Prop}$$

$$S^* = ((Ev \rightarrow \text{Prop}) \rightarrow Ev \rightarrow \text{Prop}) \rightarrow \text{Prop}$$

$$(C1/C2)^* = (C1 \setminus C2)^* = C2^* \rightarrow C1^*$$

図 3: 統語範疇と意味表示の型との対応関係.

意味表示のラムダ項の型は, Entity (E), Event (Ev), Proposition (Prop) の三つの基本型と関数型からなり, 以下のように定義される.

$$T ::= E \mid Ev \mid \text{Prop} \mid T_1 \rightarrow T_2$$

日本語 CCG パーザは S と NP の二つの基底範疇をもち, 統語範疇から型への写像 $(\cdot)^*$ を図 3 のように再帰的に定義できる. 統語範疇と型の間に対応関係を保持した意味割り当てを行うことで, CCG の導出木の葉ノードに対する意味割り当てから合成的に well-formed な意味表示が得られることが保証される.

3.2 日本語動詞句の意味合成

文の骨格となる動詞句に対しては, いわゆる Neo-Davidsonian Event Semantics (Parsons, 1990) のアプローチに基づいて意味表示を与える. イベント意味論は, 動詞句がイベントないし状態を個体として導入するという考え方に基づく意味論であり, 特に副詞的な修飾表現をイベントに対する述語とみなすことで, 時間表現を含めた動詞にかかわる推論を簡潔に扱うことができる点に特徴がある. たとえば, (1) が (2), (3) を含意し, (2), (3) がそれぞれ (4) を含意するという事実は, イベント意味論の各意味表示によって捉えることができる (v はイベントを表す変数).

- (1) 太郎が駅までゆつくり歩いた.
 $\exists v (\text{歩く}(v) \wedge (\text{Nom}(v) = \text{太郎}) \wedge \exists x (\text{駅}(x) \wedge \text{Goal}(v, x)) \wedge \text{ゆつくり}(v) \wedge \text{Past}(v))$
- (2) 太郎が駅まで歩いた.
 $\exists v (\text{歩く}(v) \wedge (\text{Nom}(v) = \text{太郎}) \wedge \exists x (\text{駅}(x) \wedge \text{Goal}(v, x)) \wedge \text{Past}(v))$
- (3) 太郎がゆつくり歩いた.
 $\exists v (\text{歩く}(v) \wedge (\text{Nom}(v) = \text{太郎}) \wedge \text{ゆつくり}(v) \wedge \text{Past}(v))$
- (4) 太郎が歩いた.
 $\exists v (\text{歩く}(v) \wedge (\text{Nom}(v) = \text{太郎}) \wedge \text{Past}(v))$

Neo-Davidsonian Event Semantics では、動詞はイベントを項にとる 1 項述語として分析され、項 (必須格) と副詞表現・時間表現を含む付加詞 (非必須格) はともに、イベントを項にとる述語として分解される。この分析は、項と付加詞の区別が必ずしも明確でなく、また項の省略が頻繁に起こる日本語文の意味記述に適していると考えられる。

イベント意味論に基づく意味表現を合成的に与えるにあたってまず問題となるのは、(1)–(4) の各論理式の冒頭部に現れるイベントの存在量化をどのタイミングで導入するかという点である。従来標準的であったのは、文が完結する S ノードにおいて存在量化の操作 (存在閉包) を適用するという考え方であったが、量化名詞句や否定を含む文では、一般的にイベントの存在量化よりも量化名詞句や否定が広いスコープをとることが知られており、問題になる。ここでは、Champollion (2015) の分析に従って、動詞自体がイベントの量化を導入すると考える。また、副詞等の修飾表現は、動詞によって導入された存在量化子のスコープ内に出現しなくてはならない。そこで、動詞の型を繰り上げ、動詞が修飾表現を項にとると考える。表 1 のテンプレートに従って合成された、文「学生がゆっくり走った」の意味表示の導出木を図 4 に示す。

また、動詞を修飾するオペレータの中には、イベントの成立を含意しないものがあり、含意関係認識では重要なクラスをなす。たとえば、「電車に乗り損なった」は「電車に乗った」を含意せず、「論文がほぼ完成した」は「論文が完成した」を含意しない。この場合、たとえば「～損なう」という単語単位で、次のような種類の意味割り当てを行う。

統語範疇: $S \setminus S$
 意味表示: $\lambda QNF.Q(\lambda Gx.N(\text{fail}(G, x), F))$
 基底形: 損なう

ここで fail は「乗る」のような述語に適用される高階のオペレータである。これにより、「太郎は 7 時の電車に乗り損なった」という文には、 $\exists v(\text{fail}(\text{乗る}, v) \wedge (\text{Nom}(v) = \text{太郎}) \wedge (\text{Dat}(v) = \text{7 時の電車}) \wedge \text{Past}(v))$ という意味表示を与えられ、この文から「太郎は 7 時の電車に乗った」への推論は正しくブロックされる。

3.3 日本語名詞句の意味合成

名詞句の分析においても、量化構造を正しく捉えることがいわゆる単調推論 (monotonicity inference) などの基本的な含意関係の認識にとって重要である。英語の場合、 $a, the, every, some, no, most$ などの決定詞により名詞句の量化の種類を特定することができる。英語 CCG (Hockenmaier and Steedman, 2007) では、普通名詞の統語範疇を N 、決定詞が接続した形式の統

語範疇を NP と区別することで量化構造を捉えることが可能になっている。

一方、日本語の名詞句の特徴的な点は、決定詞を欠いた裸名詞句が頻出する点である。戸次 (2010) による日本語 CCG では、英語と同様に統語範疇 N と NP を用い、英語の冠詞の位置に空範疇の存在を仮定することで名詞句の量化の分析を与えている。日本語 CCG パーザは、この点で戸次 (2010) の文法を簡略化し、統語範疇 N を用いず、名詞句の統語範疇はすべて NP であるとしている。この簡略化された CCG に適合した合成的意味論を与えるには、動詞句の場合と同様、名詞句自体が量化 (デフォルトでは存在量化) を導入すると考えればよい。また名詞修飾表現の統語範疇が N/N ではなく NP/NP となるため、名詞句が導入した存在量化子のスコープの内側に修飾表現を導入する必要がある。ここでも、動詞句の分析と同様、型繰り上げにより名詞句がその修飾表現を項として要求すると考える (表 1 の NP の意味表示を参照)。これにより、たとえば全称文「すべての学生が来た」には $\forall x(\exists y(\text{学生}(y) \wedge (x = y) \rightarrow \exists v(\text{来る}(v) \wedge \text{Past}(v) \wedge (\text{Nom}(v) = x))))$ という意味表示が導出され、正しい推論を行うことが可能となる。

4 実験と評価

システムの評価を、日本語意味論テストセット JSeM (Kawazoe et al., 2015) を用いて行った²。JSeM は日本語の意味論的な現象に基づく含意関係データセットであり、含意関係認識システム及び意味理論の推論能力を評価する目的で設計されている。同様のテストセットとして英語では FraCaS データセットがあり、含意関係認識システムの評価に広く用いられている。JSeM は FraCaS の日本語訳に対応する問題と、独自に追加された問題から成る。各問題は 1 つ以上の前提と 1 つの結論からなり、答えは yes (前提が結論を含意する)、 no (前提が結論の否定を含意する)、 unknown (前提が結論も結論の否定も含意しない) の 3 パターンがある。問題の例を図 5 に示す。

JSeM では推論のタイプ (含意, 前提等) 及び推論に関与する意味現象が問題毎にラベルで明示されている。ここでは FraCaS で扱われている主要な現象のうち、量化を含む Generalized Quantifier と Plural, 内包的述語やモダリティを含む Adjective, Verb, Attitude の 5 つの現象を対象に、推論のタイプが含意 (entailment) である問題について評価を行った。

統語解析は日本語 CCG パーザ (能地ら, 2015)³ を用いて行い、形態素解析器には Kuromoji を使用した。

²システムは <https://github.com/myynlp/ccg2lambda> にて公開される。

³<https://github.com/myynlp/jigg>

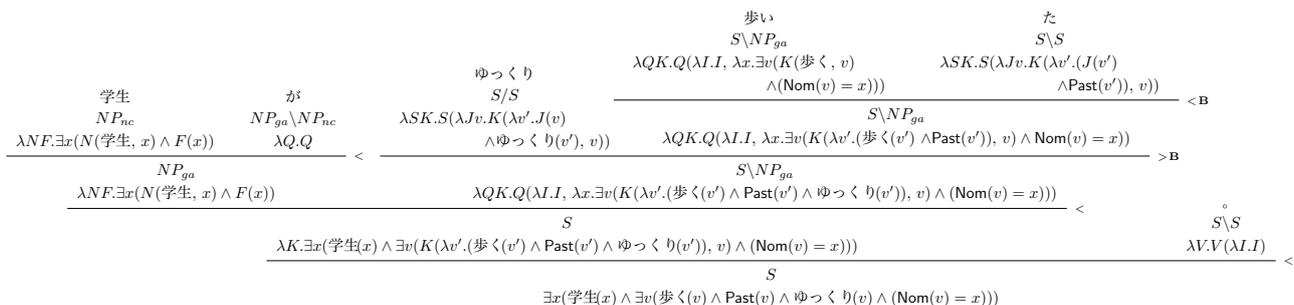


図 4: 「学生がゆつくり歩いた。」の導出木と意味表示.

jsem-327	現象: Generalized Quantifier, monotonicity 推論のタイプ: entailment
前提 1	全女性社員が書庫の鍵を持っている.
前提 2	すべての新入社員は女性社員である.
結論	全新入社員が書庫の鍵を持っている.
答え	Yes

図 5: JSeM のテストの例.

現段階での意味辞書は、統語範疇単位の意味割り当てのテンプレート 71 個、単語単位の意味割り当ての語彙項目 67 個から成る。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 30 秒でタイムアウトした場合には unknown を出力した。結果を表 2 に示す。Plain はパーザの出力結果のうち最もスコアの高い木を使用した場合の結果、Gold は不正解だった問題について人手で木を選択した場合の結果である。証明に要した平均時間は Plain が 3.7 秒/問、Gold が 3.6 秒/問であった。比較対象として、教師あり学習による分類器⁴ による結果を示す。我々のシステムは学習データを必要としないため、直接の比較はできないが、同等の正答率を達成することが示された。また、参考のため英語 FraCaS データセットの対応問題に対する Mineshima et al. (2015) の実験結果も示す。Plain と Gold の結果の比較から、構造的曖昧性の解消によりさらなる精度の向上が見込まれることが示唆された。

5 おわりに

本稿では、日本語 CCG パーザの出力を高階論理の論理式へ変換し、高階論理の推論器を用いて含意関係の判定を行うシステムを提案した。実テキストを論理式に変換し、論理推論を行うシステムの実装は、著者らの知る限り日本語では初の試みであり、今後様々な意味処理タスクへの応用が期待される。

⁴NTCIR RITE のベースラインツールを使用した (<http://www.cl.ecei.tohoku.ac.jp/rite2/doku.php?id=wiki:tools>).

現象	問題数	Plain	Gold	SL	M15
Quantifier	338	.80	.87	.90	.77 (74)
Plural	41	.56	.68	.61	.67 (33)
Adjective	65	.60	.62	.46	.68 (22)
Verb	36	.64	.64	.67	.62 (8)
Attitude	44	.57	.59	.70	.77 (13)
合計	524	.72	.78	.79	.73 (150)

表 2: JSeM での正答率。SL は教師あり学習の分類器による正答率、M15 は Mineshima et al. (2015) での英語 CCG パーザの Plain tree に基づく FraCaS での対応セクションの正答率、括弧内は FraCaS の問題数を表す。

参考文献

- Lasha Abzianidze. A tableau prover for natural logic and language. In *Proceedings of EMNLP2015*, pages 2492–2502, 2015.
- Johan Bos. Wide-coverage semantic analysis with Boxer. In *Proceedings of STEP '08*, pages 277–286, 2008.
- Lucas Champollion. The interaction of compositional semantics and event semantics. *Linguistics and Philosophy*, 38(1): 31–66, 2015.
- Robin Cooper et al. FraCaS — a framework for computational semantics. *Deliverable D6*, 1994.
- Julia Hockenmaier and Mark Steedman. CCGbank: a corpus of CCG derivations and dependency structures extracted from the Penn Treebank. *Computational Linguistics*, 33(3): 355–396, 2007.
- Ai Kawazoe, Ribeka Tanaka, Koji Mineshima, and Daisuke Bekki. An inference problem set for evaluating semantic theories and semantic processing systems for Japanese. In *Proceedings of LNL12*, pages 67–73, 2015.
- Marco Marelli et al. A sick cure for the evaluation of compositional distributional semantic models. In *Proceedings of LREC*, pages 216–223, 2014.
- Koji Mineshima, Pascual Martínez-Gómez, Yusuke Miyao, and Daisuke Bekki. Higher-order logical inference with compositional semantics. In *Proceedings of EMNLP2015*, pages 2055–2061, 2015.
- Terence Parsons. *Events in the Semantics of English*. MIT Press, 1990.
- Sumire Uematsu, Takuya Matsuzaki, Hiroki Hanaoka, Yusuke Miyao, and Hideki Mima. Integrating multiple dependency corpora for inducing wide-coverage japanese CCG resources. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 14(1):1–24, 2015.
- Johan van Benthem. A brief history of natural logic. In *Logic, Navya-Nyāya & Applications: Homage to Bimal Krishna Matilal*. College Publications, 2008.
- 能地 宏, 榊原 隆文, 宮尾 祐介. 日本語パイプライン処理のための簡易フレームワークの提案. 言語処理学会第 21 回年次大会発表論文集, 2015.
- 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.