

多言語に対応した汎用的な構成的意味論の 計算フレームワークの構築

峯島宏次（慶應義塾大学） 窪田悠介（国立国語研究所）
宮尾祐介（東京大学） 窪田愛（東京大学） 船蔵颯（京都大学）

要旨 自然言語の文を高階論理の意味表示へ構成的に対応付ける新しい計算フレームワークを作るプロジェクトを導入し、その現状と今後の展望について報告する。具体的には、まず組合せ範疇文法のパーザを利用して、その出力である導出木に意味タグの情報を付与することで、できるだけ個別言語に依存しない形で適切な意味表示への対応付けを行う構成的意味論の計算システムを実装する。その上で、英語の統語論のテキストから抽出された例文に対して意味表示を対応づける意味表示のテストセットを作成することで、現状のシステムによりどの程度の範囲の言語現象を扱うことができるのかの検証を行う。また、人間の言語処理・言語獲得の計算モデリングへの応用、自然言語処理への応用を中心に「計算可能な構成的意味論」の意義について考察する。

1. フレームワークの概要

文をその統語構造に即して何らかの意味表現に対応づける構成的意味論は、現代の形式意味論の核にある考え方である。これを計算可能なシステムとして実装し、様々な言語の実テキストを対象として解析を行う研究は、組合せ範疇文法 (CCG) (Steedman 2000) など構成的意味論と相性のよい文法理論に基づくコーパスの出現 (Hockenmaier and Steedman 2007, Uematsu et al. 2015) を契機として始まったが、まだその端緒についたばかりの段階にある。既存研究の一つ `ccg2lambda` (Mineshima et al. 2016, Martínez-Gómez et al. 2016) では、英語と日本語に対してそれぞれCCGに基づく意味計算システムを実装している。本研究では、そこから共通のコンポーネントを取り出し、できるかぎり個別言語に依存しないフレームワークを構築する。具体的には、ここで提案するフレームワークは、以下の三つからなる。

1. CCGから高階論理(HOL)の意味表示を導出する構成的意味論の計算システム
2. そのために必要な意味辞書（意味タグ）
3. 各言語で文と意味表示を対応付けるコーパス

現段階では、英語と日本語を対象としている。英語CCGパーザ (Lewis and Steedman 2014, Yoshikawa et al. 2017) の出力を利用することで、たとえば、(1a)の文に対しては(1b)のHOLの意味表示が対応付けられる。システムが実際に出力したCCG導出木を図1に示す。

- (1) a. Every child quietly read the book.
b. ((every_{UNIV} child_{NN}) (quietly_{ADV} (read_{TV} (the_{DT} book_{NN}))))

この例が示すように、HOLの意味表示とは、標準的な形式意味論のLFに限りなく近いものである。HOLは、語順の情報など個別言語の多様性を捨象した意味に関する共通言語を目指して設計されており、このレベルの意味表示を適切に設計することで、理論言語学を計算論的手法で検証する研究が飛躍的に進む可能性がある。

また、HOLの意味表示は、1階述語論理 (FOL) や談話表示構造 (DRS) など、目的に応じて多様な意味表示に展開可能である。本研究では、FOLとDRSの意味表示としては、イベント意味

論 (Parsons 1990)に基づくものを採用している。図1の導出木の各ノードには、HOLの表現とそれに対応するイベント意味論の表現が付与されている。(1a)の文に対してシステムから自動的に導出されたFOLとDRSの意味表示を図2に示す。

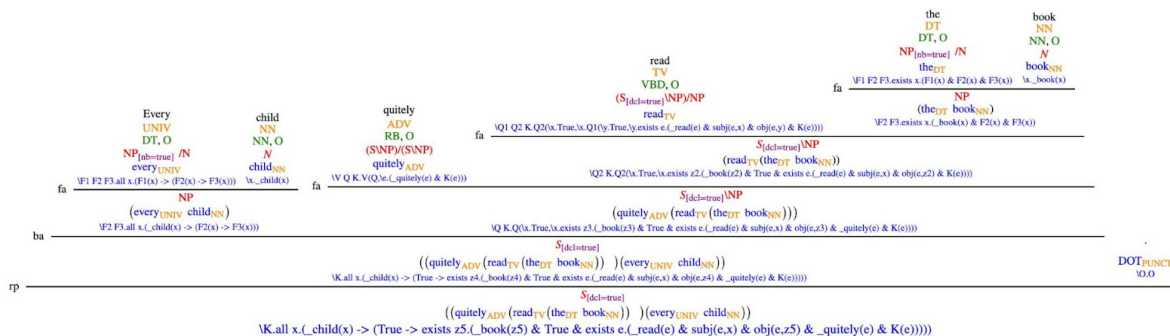


図1：Every child quietly read the bookのCCG導出木と意味合成

FOL: $\text{all } x1.(_child(x1) \rightarrow \text{exists } x2.(_book(x2) \& \text{exists } e3.(_read(e3) \& \text{subj}(e3,x1) \& \text{obj}(e3,x2) \& _quietly(e3))))$

DRS: $([[], [([x1], [child(x1)])] \rightarrow ([e3, x2], [book(x2), read(e3), subj(e3, x1), obj(e3, x2), quietly(e3)])])$

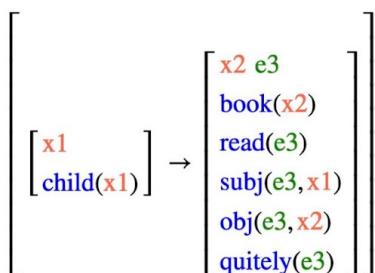


図2：FOLとDRSの2種類の意味表示

2. 意味タグの設計と意味表示テストセットの構築

意味合成や様々な意味表示の表現形式への展開をできるかぎり一般的な形で行うため、先行研究 (Abzianidze et al. 2017) に基づいて汎用的な意味タグを設計する。意味タグを導入することで、CCGの統語範疇や標準的な品詞タグよりも詳細な意味表示を単語に対して指定することが可能となる。たとえば、繰り上げ述語 *seem* とコントロール述語 *try* は、CCGbank (Hockenmaier and Steedman 2007)、及び、それに基づく既存のCCGパーザでは同じ統語範疇 (S\NP)/(S\NP) が割り当てられている。よって、(2a)と(3a)の文に対する導出木はまったく同じ形になり、そこから導かれる意味表示を区別することができない。そこで、*seem*と*try*にそれぞれ異なる意味タグ (RSとCONTR) を付与する。RSタグをもつ語には(2b)、CONTRタグをもつ語には(3b)に示す意味表示を割り当てる。これによって、(2c)と(3c)のように、それぞれの主語指向性の違いを考慮した適切な意味表示を導くことができる。

- (2) a. John seems to like pizza.
 b. $\lambda P \lambda x. \text{seem}_{RS}(P(x))$

- c. $\text{seem}_{RS}(\text{like}_{TV}(\text{pizza}_{NN})(\text{john}_{NNP}))$
 (3) a. John tries to like pizza.
 b. $\lambda P \lambda x.(\text{try}_{CONTR}(P))(x)$
 c. $(\text{try}_{CONTR}(\text{like}_{TV}(\text{pizza}_{NN}))) (\text{john}_{NNP})$

意味タグを導入するもう一つの利点は、意味合成をできるかぎり個別言語に依存しない形で行うことを可能にする点にある。日本語でコントロール述語に相当する語、例えば、「食べてみる」における「みる」の統語範疇は、 $(S\text{INP})/(S\text{INP})$ という形であり、英語での *try* の統語範疇 $(S\text{INP})/(S\text{INP})$ とはスラッシュの向きだけが異なる。これは $S\text{INP}$ という範疇の項をこの述語の左側に要求するか右側に要求するかの違いであり、意味計算の結果には影響しない。そこで共通の RS という意味タグを付与し、そのタグごとに意味割り当てを定義することにより、このような語順に関する違いを捨象することができる。既存研究(ccg2lambda)では英語と日本語で異なる意味割り当てが定義されていたのに対して、本研究の意味タグに基づく方法によって、言語間の共通部分を取り出し、より一般的な構成的意味論を実装することが可能となる。将来的には、多言語 CCGパーザ (Tran and Miyao 2022) と組み合わせることで、英語や日本語に限らず多言語の意味合成・推論システムが実現されることが期待される。

ここで実装したシステムによりどの程度の範囲の言語現象をカバーできるのかをテストするため、英語の統語論のテキスト (Kubota and Levine, to appear) に現れる例文約200文を対象に、自動的に HOL の意味表示を付与し、人手でその正しさをチェックしたコーパス (意味表示テストセット) を作成した。各文には、図1と図2の例と同様に、CCG導出木と3種類の意味表示 (HOL , FOL , DRS の意味表示) が結びつけられている。CCG導出木を得るため、英語の3つのCCGパーザ、C&C (Clark and Curran 2007)、easyccg (Lewis and Steedman 2014)、depccg (Yoshikawa et al. 2017) を使用し、各文に対する出力からそれぞれ適切な導出木を選択した。その導出木に基づいて新たな意味タグを追加することにより、基本的な項構造 (受身、使役)、複雑述語 (コントロール述語、繰り上げ述語)、量化表現、否定、モーダルなど多様な現象を含む意味表示を構築できることを確認した。このテストセット構築のために定義した意味タグの数は34個である。現状のCCGパーザでは正しい出力が得られない現象としては、等位接続や省略が挙げられる。このテストセットは公開する予定である。

3. 人間の言語処理・言語獲得の計算モデリングへの応用

言語学研究への応用の最重要課題は、 HOL を用いた人間の言語処理・言語獲得の計算モデリングである。言語が複雑かつ抽象的な意味を担う離散的な記号システムであることは疑いの余地がないが、現在までの研究は、意味情報の参照が極めて断片的かつ表面的なものにとどまっており (Abend et al. 2017, Alishahi and Stevenson 2008, Mitchener and Becker 2010 など参照)、このことが言語研究におけるモデリング手法の活用の余地を大きく限定している可能性がある。たとえば、Mitchener and Becker (2010) では、英語の繰り上げ・コントロール述語の獲得を計算モデリングの手法で扱っているが、繰り上げ・コントロール述語の区別を、事実上、主語の有生性と動詞句のイベント性という二つの二値素性における傾向性の違いで近似することで機械実装可能な言語獲得のモデルを構築している。しかしながら、子供が文の意味を考慮に入れず、このような表面的な手がかりのみで動詞の意味クラスを習得しているという仮定は現実的ではない。本研究で構築するフレームワークを用いれば、たとえば、文の意味を理解することによって入手可能な情報 (含意関係に基づく推論など) がこのような文法事項の習得に影響を及ぼすか、という問題を具体的に検証可能な計算モデリングの課題に落とし込んで検証することが可能となる。このことにより、理論言語学的知見をより直接的に取り込んだ形でのモデリング研究が可能となり、人間の言語処理・言語習得の研究にブレイクスルーをもたらす可能性がある。

4. 自然言語処理への応用

自然言語処理に対しては、本研究の高階意味表示・意味合成フレームワークにより、高精度な多言語意味解析・意味推論技術が実現される可能性がある。高階意味表示に基づく先行研究（ccg2lambda）は、これまで自然言語処理の文脈で研究されており、論理推論 (Mineshima et al. 2016, Haruta et al. 2022)、意味的類似度の計算 (Yanaka et al. 2017)、アブダクションによる知識獲得 (Yoshikawa et al. 2019) などに応用されている。

現在の自然言語処理においては、含意関係認識といった意味解析・推論技術において大規模言語モデルを用いた手法が主流であり、実用レベルの高い精度を実現している (Bowman et al. 2015, Williams et al. 2018)。しかし、学習データと異なるドメインを対象とする場合や、複数の意味論的現象が関わる深い推論が必要な場合には、大規模言語モデルを用いた意味解析・推論技術は精度が著しく低下することが示されている (McCoy et al. 2019, Yanaka and Mineshima 2019, 2022)。大規模言語モデルを用いた高精度な意味解析・推論と、本研究が提案する高階意味表示による意味解析・推論を組み合わせることで、上記の問題を解決した意味解析・推論システムが実現されることが期待される。

本研究の意味表示・合成手法ではCCG等の導出木と意味辞書（意味タグと意味表示の対応関係）が与えられれば意味解析・推論が実現できるため、本質的に学習データを必要としないものである。したがって、言語リソースが限られた多くの言語において意味解析・推論システムを実現する上でも重要な貢献となることが期待される。

謝辞

本研究はJST CREST、JP-MJCR2114 の支援を受けたものである。

参考文献

- Abend, O., Kwiatkowski, T., Smith, N. J., Goldwater, S., & Steedman, M. (2017). Bootstrapping language acquisition. *Cognition*, 164, 116–143.
- Abzianidze, L. and Bos, J. (2017). Towards universal semantic tagging. In *12th International Conference on Computational Semantics (IWCS 2017)*.
- Alishahi, A., & Stevenson, S. (2008). A computational model of early argument structure acquisition. *Cognitive Science*, 32, 789–834.
- Bowman, S., Angeli, G., Potts, C., & Manning, C. D. (2015). A large annotated corpus for learning natural language inference. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 632–642.
- Clark, S., & Curran, J. R. (2007). Wide-coverage efficient statistical parsing with CCG and log-linear models. *Computational Linguistics*, 33(4), 493–552.
- Haruta, I., Mineshima, K., & Bekki, D. (2022). Implementing natural language inference for comparatives. *Journal of Language Modelling*, Vol.10, No.1, 139–191.
- Kubota, Y. & Levine, R. (to appear). *Type-Logical Grammar: A Syntactic Introduction*.
- Lewis, M., & Steedman, M. (2014). A* CCG parsing with a supertag-factored model. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, 990–1000.

- Martínez-Gómez, P., Mineshima, K., Miyao, Y., & Bekki, D. (2016). ccg2lambda: a compositional semantics system. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, System Demonstrations*, 85–90.
- McCoy, R. T., Pavlick, E., & Linzen, T. (2020). Right for the wrong reasons: Diagnosing syntactic heuristics in natural language inference. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3428–3448.
- Mineshima, K., Tanaka, R., Martínez-Gómez, P., Miyao, Y., & Bekki, D. (2016). Building compositional semantics and higher-order inference system for a wide-coverage Japanese CCG parser. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 2236–2242.
- Mitchener, W. G., & Becker, M. (2010). Computational models of learning the raising-control distinction. *Research on Language and Computation*, 8, 169–207.
- Parsons, T. (1990). *Events in the Semantics of English: A Study in Subatomic Semantics*. MIT Press.
- Steedman, M. (2000). *The Syntactic Process*. MIT press.
- Tran, T.-A., & Miyao, Y. (2022). Development of a Multilingual CCG Treebank via Universal Dependencies Conversion. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, 5220–5233.
- Hockenmaier, J., & Steedman, M. (2007). CCGbank: a corpus of CCG derivations and dependency structures extracted from the Penn Treebank. *Computational Linguistics*, 33(3), 355–396.
- Uematsu, S., Matsuzaki, T., Hanaoka, H., Miyao, Y., & Mima, H. (2015). 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.
- Yoshikawa, M., Noji, H., & Matsumoto, Y. (2017). A* CCG parsing with a supertag and dependency factored model. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, 277–287.
- Yoshikawa, M., Mineshima, K., Noji, H., & Daisuke Bekki. (2019) Combining axiom injection and knowledge base completion for efficient natural language inference. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence*, 7410–7417.
- Yanaka, H., Mineshima, K., Martínez-Gómez, P., & Bekki, D. (2017). Determining semantic textual similarity using natural deduction proofs. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 681–691.
- Yanaka, H., & Mineshima, K. (2021). Assessing the generalization capacity of pre-trained language models through Japanese adversarial natural language inference. In *Proceedings of the Fourth BlackboxNLP Workshop on Analyzing and Interpreting Neural Networks for NLP*, 337–349.
- Yanaka, H., & Mineshima, K. (2022). Compositional Evaluation on Japanese Textual Entailment and Similarity. *Transactions of the Association for Computational Linguistics*, 10, 1266-1284.
- Williams, A., Nangia, N., & Bowman, S. R. (2018). A broad-coverage challenge corpus for sentence understanding through inference. In *2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1112–1122.