

論理推論による症例検索に向けた日本語症例テキストの複合語解析の試案

石田 真捺 (指導教員：戸次大介)

1 はじめに

医療ドメインのテキストを用いた自然言語処理の研究は英語圏を中心に活発に進められている [8, 4, 1]. 一方で、日本語の医療ドメインのテキストを用いた研究 [12, 3, 10, 7] も活発に進められているが、固有表現抽出などの基礎解析が中心であり、症例テキストの意味解析や推論の研究は依然として発展途上である。

症例テキストを有効活用する方法の一つとして、時系列を考慮した症例検索が考えられるが、これを実現するためには、高度な意味解析が求められる。そこで本研究では、高度な意味解析システムである ccg2lambda [2] を用いた時系列症例検索の実現を目指す。ccg2lambda とは、組合せ範疇文法 (Combinatory Categorical Grammar, CCG) [5, 11] に基づく高精度な構文解析と、高階論理に基づく自動推論システムを組み合わせた高度な意味解析・推論システムである。

日本語の症例テキストの意味解析を行う上で重要な問題の一つとして、複合語の扱いがある。以下に例を示す。

- (1) 出産から2週間後に発熱し内服薬を処方された例
 $\exists e_1 e_2 e_3 (\text{出産}(e_1) \wedge \text{発熱}(e_2) \wedge \text{処方}(e_3) \wedge \text{obj}(e_3) = \text{内服薬} \wedge \text{before}(e_1, e_2, 2\text{weeks}) \wedge \text{during}(e_3, e_2))$
- (2) 出産2週間後より発熱、発疹が出現し、近医で感冒等と診断され内服薬処方を受けたが…
 $\exists e_1 e_2 e_3 (\text{出産2週間後}(e_1) \wedge \text{発熱}(e_2) \wedge \text{受ける}(e_3) \wedge \text{obj}(e_3) = \text{内服薬処方} \wedge \text{before}(e_1, e_2, 2\text{weeks}) \wedge \text{during}(e_3, e_2)) \dots$

(1) と (2) の意味表示を比較してみる。どちらも「出産の2週間後に発熱」「内服薬を処方される」という内容を含んでいるため、この部分は同じ意味内容を表すはずである。しかし、現在の実装では複合語にあたる下線部分の意味が正しく解析できていないため、意味の同一性が得られない。そこで本研究では、ccg2lambda に複合語解析モジュールを追加することで、症例報告のテキストも正しい意味表示を導出可能にする。

2 症例テキストの意味解析

症例テキストの意味解析システムの全体像を図1に示す。まず、入力文に対して日本語形態素解析器 Janome¹ を用いてトークナイズを行う。次に、CCG 構文解析器 depccg [9] を用いて、CCG 構文木を導出する。さらに、構文解析の出力結果から複合語を取り出し、それを複合語解析モデルに渡し、複合語内の意味関係を予測する。これには、BiLSTM の系列ラベリングモデル [6] を用いる。そして、複合語解析モデルの出力結果に基づいて、複合語の箇所の部分木を修正し、全体の CCG 構文木から修正前の部分木と修正後の部分木を入れ替える。最後に、修正した CCG 構文木と意味テンプレートに基づいて、意味表示 (論理式) を合成的に導出する。

¹<https://mocabeta.github.io/janome/>

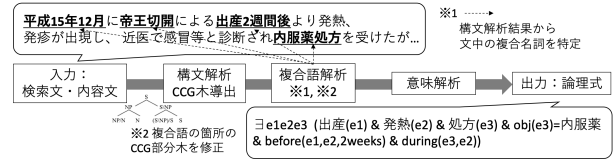


図 1: ccg2lambda を用いた症例テキスト意味解析

3 症例テキストの複合語解析

3.1 アノテーション

複合語解析モデルを構築するにあたり、訓練に用いるアノテーションデータが必要となる。意味現象タグは、表1に示す15種類を用意した。本研究で構築した意味現象タグデータでは、タグ付けに brat² を使用した。図2にタグ付けの例、表2に意味現象タグの分布を示す。

表 1: 意味現象タグの一覧

現象	タグ	例
複合語	MWE	... と診断され、 内服薬処方を受け...
後ろの形態素の一部	PP	当院耳鼻科
エンティティ	EN	GIS 所見
イベント	EV	回腸切除術施行
ガ格	GA	アミロイド沈着
ヲ格	WO	内服薬処方
ニ格	NI	当院耳鼻科入院
ノ格	POS	当院耳鼻科
エンティティの修飾	M_EN	GIS 所見
部位	PA	骨盤内リンパ節
数詞	NUM	10月24日永眠
助数詞(月、日)	CL	10月24日永眠
助数詞(ヶ月、週間)	CLP	出産2週間後
時間関係	TE	出産2週間後
術後	JTE	術後3ヶ月目

平成15年6月に左手蜂窩織炎のため当院整形外科入院中に

血便あり 当科再診となった。

図 2: 意味現象タグのアノテーション例

3.2 モデル

複合語解析では、複合語を入力として、複合語の意味現象タグを系列として出力するように、モデルを学習させる。本論文では複合語解析モデルに標準的な系列ラベリングモデルの一つである BiLSTM の系列ラベリングモデル [6] を用いた。

3.3 評価実験

実験は、Google Colaboratory (GPU を使用) で行った。3.1 節のアノテーションデータを用いて複合語解析モデルを学習し、10 分割交差検証でモデルの評価を行った。今回はデータを段階的に増やし、実験を行った。1 回目の実験では、対象文に含まれる複合語の数は全部で 131 件だが、1 つの文ペアに同じ複合語が重複して出現していた場合、その重複を取り除いたため、残った 103 件を学習データとしている。2 回目の実験は、1 回目の対象文を含む 301 文を対象とした。これ

²<http://brat.nlplab.org/>

表 2: 意味現象タグの分布

タグ	1 回目の出現数		2 回目の出現数	
	出現数	割合 (%)	出現数	割合 (%)
EV	101	24.75	212	25.95
PP	59	14.46	146	17.87
TE	39	9.56	83	10.16
GA	36	8.82	56	6.85
NUM	36	8.82	50	6.12
EN	30	7.35	59	7.22
WO	30	7.35	81	9.91
CL	26	6.37	37	4.53
PA	16	3.92	17	2.08
M_EN	12	2.94	21	2.57
CLP	10	2.45	10	1.22
NI	8	1.96	29	3.55
POS	4	0.98	8	0.98
JTE	1	0.25	8	0.98
合計	408	100	817	100

表 3: 意味現象タグの予測精度

	学習データ数	精度の平均	標準偏差
1 回目	103 件	0.6638	0.1364
2 回目	209 件	0.7750	0.0330

に含まれる複合語は 247 件で、重複を除いた学習データは、209 件である。

表 3 は、意味現象タグの予測精度である。1 回目の実験では精度が安定していないのに対し、2 回目の精度は安定しており、1 回目と比べると向上が見られた。

3.4 エラー分析

表 4 は、タグの予測結果である。下線部分のタグが、間違っ て予測されてしまったものである。1 回目と 2 回目の予測結果を比較してみると、データを増やした 2 回目の方が、間違っ て予測してしまっ たタグが減っ ている。以上のことから、データを増やしたことで精度が向上したことがわかる。

表 4: タグの予測結果。下線部分は誤っている箇所。

入力 (複合語)	正解タグ	予測結果 (1 回目)	予測結果 (2 回目)
“LA 分類 gradeD”	['PP', 'POS', 'EN']	['PP', 'POS', 'EN']	['PP', 'POS', 'EN']
“手術 所見”	['M_EN', 'EN']	['M_EN', 'EN']	['M_EN', 'EN']
“回腸 切除術 施行”	['WO', 'EV', 'WO', 'EV']	['WO', 'EV', 'TE', 'PAD']	['WO', 'EV', 'EN', 'EV']
“肝 左葉 切除術”	['PA', 'PP', 'WO', 'EV', 'EN']	['PA', 'PP', 'WO', 'EV', 'WO']	['PA', 'PP', 'WO', 'EV', 'EN']
“術後 3 ヶ月 目”	['JTE', 'NUM', 'CLP', 'TE']	['PP', 'NUM', 'NUM', 'CLP']	['JTE', 'NUM', 'CLP', 'TE']

4 症例テキストの意味解析

4.1 構文解析

複合語解析モデルが予測した複合語の意味関係に基づいて、複合語の箇所の CCG 構文木を修正するため、意味現象タグに基づいて正しい構文木を出力する CFG パーザを構築した。設計した CFG 規則の一部を表 5 に示す。

表 5: CFG 規則 (一部)

MWE	→	EN.1 EV.1 EV.2 TE.1 PT.1
EV.1	→	“EV” PP.1 “EV” GA.1 EV.2 WO.1 EV.2 NI.1 EV.2 PT.1 EV.2 SP.1 EV.2
EV_wo	→	WO.2 “EV”
EN.1	→	“EN” PP.1 “EN” M_EN EN.2 POS “EN” GA.2 “EN” EV.1 “EN” EV.2 “EN” PA “EN” PT.1 EN.2 TE.1 EN.2
SP.1	→	“NUM” “CLP” PP.1 SP.2
PT.1	→	“NUM” “CL” PT.3 PT.3 PP.1 PT.2 PP.1 PT.3
TE.1	→	“TE” “JTE” EN.2 TE.2 EV.1 TE.2 SP.1 TE.2
M_EN	→	“M_EN” PP.1 “M_EN”
WO.1	→	“WO” PP.1 “WO” EV_wo WO.2 WO.2 WO.2 “NUM” WO.2
NI.1	→	“NI” PP.1 “NI” POS “NI” PA NI.2 M_EN NI.2
GA.1	→	“GA” PP.1 “GA” PA GA.2 M_EN GA.2
POS.1	→	“POS” PP.1 “POS”
PA	→	“PA” PP.1 “PA”
PP.1	→	“PP” PP.2 “PP”

4.2 意味解析

表 6 に意味テンプレートの一部を示す。統語範疇をキーとして意味表示を割り当てるため、統語範疇のみ

表 6: 意味テンプレート (一部)

タグ	統語範疇	意味表示
EV	S[ev]	$\lambda C1. \lambda C2. \lambda C3. \lambda K. \exists x.(K(\text{Surf}, e))$
GA	S[ga]/S[X]	$\lambda S. \lambda C1. \lambda C2. \lambda C3. \lambda K. \exists x(\text{Surf}(x) \& S(C1, C2, C3, \lambda J. \lambda E.K(J, E) \& C1(x, E, \text{Nom})))$
WO	S[wo]/S[X]	$\lambda S. \lambda C1. \lambda C2. \lambda C3. \lambda K. \exists x(\text{Surf}(x) \& S(C1, C2, C3, \lambda J. \lambda E.K(J, E) \& C2(x, E, \text{Acc})))$
NI	S[ni]/S[X]	$\lambda S. \lambda C1. \lambda C2. \lambda C3. \lambda K. \exists x(\text{Surf}(x) \& S(C1, C2, C3, \lambda J. \lambda E.K(J, E) \& C3(x, E, \text{Dat})))$
EN	NP[en]	$\lambda N. \lambda F. \exists x.(N(\text{Surf}, x) \& F(x))$
POS	NP[pos]/NP[X]	$\lambda M. \lambda N. \lambda F. \exists y.(N(\text{Surf}, y) \& M(N, \lambda x.\text{Of}(x, y) \& F(x)))$
M_EN	NP[en]/NP[X]	$\lambda M. \lambda N. \lambda F. \exists y.(N(\text{Surf}, y) \& M(N, \lambda x.\text{Rel}(x, y) \& F(x)))$

で元のタグが同定できるように設計した。そのため、意味現象タグと統語範疇、意味表示は 1 対 1 で対応している。

4.3 意味解析の結果

実験に用いた文を簡易化した 73 文に対し意味解析を行ったところ、39 文について正しい意味表示を得ることができた。図 3 に成功例を示す。

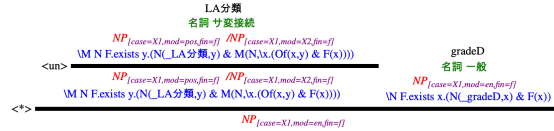


図 3: 複合語箇所の意味解析の結果

5 おわりに

本稿では、時系列を考慮した高度な症例検索に向けて、ccg2lambda を用いた症例テキストの複合語解析の手法を提案した。症例テキスト中の複合語の意味解析を行うために、複合語の意味関係をアノテーションしたデータセットを構築し、BiLSTM を用いた系列ラベリングモデルを実装した。実験の結果、学習データの増加に伴って、予測の精度向上がみられた。予測した意味現象タグに基づいた意味解析を行ったところ、73 件中 39 件のテキストについて正しい意味表示を得ることができた。今後は、アノテーションデータを増やすことで、複合語解析モジュールの精度向上を目指していく予定である。

謝辞：本研究の一部は、JST AIP-PRISM JP-MJCR18Y1 の助成を受けたものである。

参考文献

- [1] Bethard, S., Savova, G., Palmer, M. and Pustejovsky, J.: SemEval-2017 Task 12: Clinical TempEval, in *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, pp. 565–572, Vancouver, Canada (2017), Association for Computational Linguistics.
- [2] Martínez-Gómez, P., Mineshima, K., Miyao, Y. and Bekki, D.: ccg2lambda: a compositional semantics system, in *Proc. of ACL System Demonstrations*, pp. 85–90 (2016).
- [3] Mutinda, F. W., Nigo, S., Shibata, D., Yada, S., Wakamiya, S. and Aramaki, E.: Detecting Redundancy in Electronic Medical Records Using Clinical BERT, 言語処理学会第 26 回年次大会発表論文集 (2020).
- [4] Romanov, A. and Shivade, C.: Lessons from Natural Language Inference in the Clinical Domain.
- [5] Steedman, M.: *The Syntactic Process*, MIT Press (2000).
- [6] Wang, P., Qian, Y., Soong, F. K., He, L. and Zhao, H.: Part-of-speech tagging with bidirectional long short-term memory recurrent neural network, *arXiv preprint arXiv:1510.06168* (2015).
- [7] Yada, S., Joh, A., Tanaka, R., Cheng, F., Aramaki, E. and Kurohashi, S.: Towards a Versatile Medical-Annotation Guideline Feasible Without Heavy Medical Knowledge: Starting From Critical Lung Diseases, in *Proceedings of the 12th Language Resources and Evaluation Conference*, pp. 4565–4572, Marseille, France (2020), European Language Resources Association.
- [8] Yadav, P., Steinbach, M., Kumar, V. and Simon, G.: Mining Electronic Health Records (EHRs): A Survey, *ACM Comput. Surv.*, Vol. 50, No. 6 (2018).
- [9] Yoshikawa, M., Noji, H. and Matsumoto, Y.: A* CCG Parsing with a Supertag and Dependency Factored Model, in *Proc. of ACL*, pp. 277–287 (2017).
- [10] 山本英弥, 伊藤藤, 荒牧英治: 複合語の構成素情報を考慮した病名難易度の推定, 言語処理学会第 25 回年次大会発表論文集 (2019).
- [11] 戸次大介: 日本語文法の形式理論, くろしお出版, 東京 (2010).
- [12] 田川裕輝, 西埜徹, 谷口元樹, 谷口友紀, 大熊智子, 若宮翔子, 荒牧英治: 生成された読影所見の自動評価に向けた固有表現認識とモダリティ推定, 言語処理学会第 26 回年次大会発表論文集 (2020).