

複数の述語間の関係を考慮した End-to-End 日本語述語項構造解析

松林優一郎[◆] 乾健太郎[◆]

[◆]東北大学 [◇]理化学研究所AIP センター
 {y-matsu, inui}@ecei.tohoku.ac.jp

1 はじめに

述語項構造は、文章内の述語とその項との関係を規定する構造である。例えば下の図1の文では、「尋ねる」や「答える」という表現が述語であり、「記者団」や「首相」という表現が述語「尋ねる」の項である。このように、文章中の要素を述語との関係によって構造的に整理する事で、複雑な構造を持った文章において「誰が、何を、どうした」のような文章理解において重要な情報を抽出することができる。

記者 [団] が 尋ねる と [首相] が 答えた。

図1 述語項構造解析の例：下線は説明のために現在注目している述語。鉤括弧は述語の項の主辞。下付き文字は格関係。破線は統語係り受け関係。

日本語の述語項構造解析における現状の主要な課題は、述語と直接的な係り受け関係にない項 (*Zero* 事例と呼ぶ) の推定精度の向上である。日本語の述語項関係においては、図1における「尋ねる」と「首相」の関係のように、項の省略や並列構造、機能動詞構文などを伴って述語と項が間接的な統語関係となる事例が全体の25%存在するが^{*1}、これらの解析精度は、項が同一文中にある場合に対象を限っても F_1 値で53.6%程度と低い水準に留まっている [7]。

この問題の解決にあたって、これまでに大別して二つのアプローチが取られてきた。一つは、述語項間の統語関係とは相補的な情報として、大規模コーパスから述語と項、格関係の共起情報を獲得し、これを推定に用いる方法である [10, 11]。もう一方は、述語間の項の共有関係をモデル化することで、比較的推定精度のよい述語項関係の情報を他の述語項関係の推定に利用し、全体の推定精度を向上させる手法である [5, 8, 9]。

本稿では、日本語述語項構造解析の精度向上に向けて後者のアプローチを取る。我々は、複数の述語に関する項の推定情報を統合する新しい項構造解析モデルを提案する。このモデルは、多層双方向 RNN を用いた end-to-end の項構造解析モデルをベースとし (図2)、入力層と、最終層の直前に複数の述語情報を考慮するための機構を新たに導入する (図3, 4)。実験では、我々の提案するモデルが、標準的なベンチマークコーパスにおいて、特に述語と項が間接的な統語関係となる事例の精度を改善し、事例全体の F_1 値においてもこれまでの最高精度を達成したことを示す。

2 タスク設定

本研究では、日本語述語項構造解析のベンチマークコーパスの一つである NAIST Text Corpus (NTC) [4] 1.5 の仕様、および既存研究のタスク設定に習い問題の入出力を定める。

入力として、単語区切りが与えられた文 $w = w_1, \dots, w_n$ と、述語の位置 $p = p_1, \dots, p_q$ が与えられる。解析器は、与えられ

^{*1} NAIST Text Corpus 1.5 における統計。

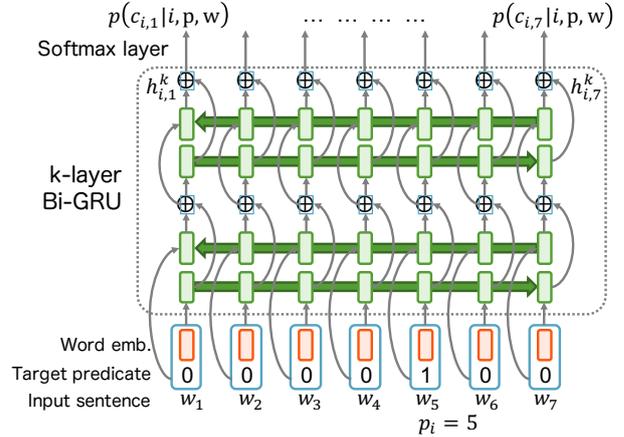


図2 ベースモデルのネットワーク図

たそれぞれの述語に対し、文中からその述語のガ、ヲ、ニ格に対応する項をそれぞれ高々一つ出力する。項は、その範囲として適切な文字列の最右の語で表現する。本稿では既存研究 [7, 8] に習い、述語項関係のうち特に述語と項が同一文内にある事例のみを解析対象とする。また、実験での比較対象である Ouchi et al. (2017) [9] に習い、項が述語と同一文節内にある場合は評価対象から除外する。

3 ベースモデル

我々の提案モデルは、多層双方向 RNN を用いた end-to-end の項構造解析モデル [13] を拡張し、複数の述語情報を考慮する機構を導入したものである。ただし、拡張前のモデル (ベースモデル) についても既存の end-to-end 項構造解析における知見を踏まえて我々が新たに設計したものとなっている。

我々のベースモデルのネットワークを図2に示す。形式的に、ベースモデルは入力として単語列 $w = w_1, \dots, w_n$ と解析対象となる述語の位置 p_i を取り、入力された単語列に対応するラベルの確率 $p(c_{i,1}|i,p,w), \dots, p(c_{i,n}|i,p,w)$ を出力する。ここで $c_{i,t} \in \{ガ, ヲ, ニ, NONE\}$ は単語 w_t が述語 w_{p_i} の特定の格になるかを表すものである。入力層では、He et al. (2017) [3] の報告を参考に、入力の各単語を埋め込んだベクトルと、解析対象の述語の位置をバイナリ値として表現したものを結合する。次に、これらのベクトル列を多層双方向 RNN に入力する。一般に、項構造解析の分野では奇数層を順方向、偶数層を逆方向とし、一方の RNN の出力を逆方向の RNN の入力に用いる形式の双方向 RNN が用いられてきたが [3, 9, 13]、この形式の双方向 RNN では層が深くなるため学習上の工夫を必要とした [3, 9]。そこで、我々はこの方法の代わりに、Marcheggiani & Titov (2017) [6] と同様に、一層ごとに順方向と逆方向の RNN の出力を結合して次の層に入力する、より一般的な多層双方向 RNN の形式を利用することで深層化を軽減する方法を取る (図2中央)。RNN には GRU [2] を採用する。その後、 k 層双方向 RNN の各時刻 t の出力 $h_{i,t}^k$

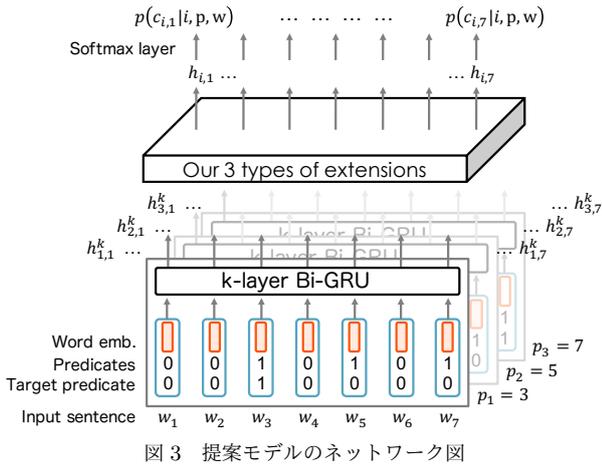


図3 提案モデルのネットワーク図

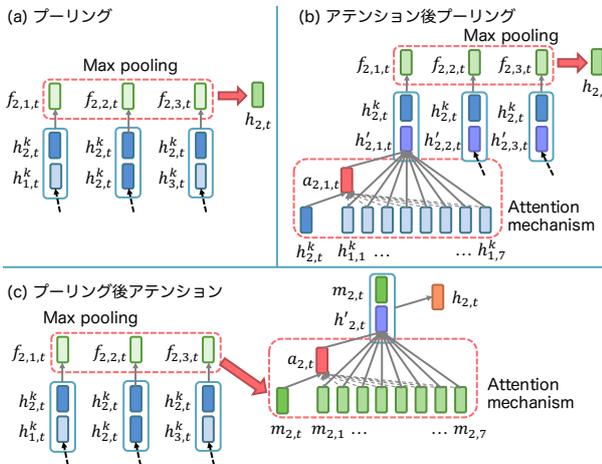


図4 提案モデルにおける3種類の拡張

を softmax 層で確率 $p(c_{i,t}|i, p, w)$ を表す 4 次元ベクトルに変換する。最終的に、ガ、ヲ、ニの各ラベルについて、ラベルの確率が出力の閾値 θ_c を超えるものがあれば、文中で最大の確率を持つ w_t をその格の項として一つ選び出力する。

4 提案モデル

ベースモデルは文中の各述語に対して独立に項を推定するモデルである。しかし実際の文では、例えば図1の文における「首相」のように、「尋ねる」「答える」という二つの述語の意味関係から「尋ねる」の二格と「答える」のガ格が同一であると推測される例が存在する。このような複数の述語間での項の依存関係をとらえるため、提案モデルではベースモデルの入力層と最終層の直前に複数の述語情報を考慮するための変更を加える。図3、4にモデルのネットワーク図を示す。

End-to-endの項構造解析において、述語間の項の依存関係を考慮する既存のモデルとして Grid RNN を用いた Ouchi et al. (2017) [9] の手法があるが、このモデルでは推定情報の統合に RNN を用いるため、情報の統合が語順や語の距離の影響を受け、結果として語の表層的な距離によって推定情報の統合に損失が現れる可能性がある。我々の手法では、単語間の統語関係を表現する部分は下層の双方向 RNN に委ね、拡張部である上層のレイヤーでは、推定に必要な語の情報同士を、プーリングやアテンションといった機構で語順や語の表層的な距離とは無関係に直接的に結びつける。

入力層では、解析対象の述語位置に加えて、文中の全ての述語の位置を表現したバイナリ値の次元を新たに与える。この拡張を行ったモデルを MP モデルと呼ぶ。後の実験で、この極めて単純な変更が解析精度に正の効果を与えることを示す。

また、ベースモデルの softmax 層の直前に、図4に示した次の3種類の機構を追加し、効果を比較する。

■(a) プーリング: MP-POOL モデル 図1の例では、「首相」は「尋ねる」の二格であるが、「答える」のガ格でもある。一般に、文中である述語の主格となった項は他の述語の項にもなりやすい。また、並列関係にある述語間では主語や目的語を共有しやすい。このような述語間の項の共有関係をモデル化するため、述語と項候補の語のペア (w_{p_i}, w_t) に関するラベルの確率を推定する際に、この語 w_t がどの程度他の述語の項になりそうかという情報を利用したい。

そこで、各単語 w_t に関して、図4(a)のように、ある述語 w_{p_i} についての k 層双方向 RNN の出力 $h_{i,t}^k$ と、文中の任意の述語 w_{p_j} ($i = j$ の場合を含む) についての出力 $h_{j,t}^k$ を結合し、非線形変換を行ったものを max pooling にかけることで、 w_t についての他の述語に関する推定情報を統合する。

$$h_{i,t} = \text{maxpool}_j(f_{i,j,t}) \quad (1)$$

$$f_{i,j,t} = \text{ReLU}(W_f[h_{i,t}^k, h_{j,t}^k] + b_f) \quad (2)$$

ここで、 $\text{maxpool}_j(f_{i,j,t})$ は $f_{i,j,t}$ の各次元について、 $\{f_{i,1,t}, \dots, f_{i,q,t}\}$ の中での最大値を取り出す操作である。このようにして得られた $h_{i,t}$ を単語 w_t の述語 w_{p_i} に関する新たな表現とし、softmax 層に入力する。

■(b) アテンション後プーリング: MP-ATT-POOL モデル 前段落では、項の共有関係のモデル化を考えたが、他にも、項となる二つの異なる語の間関係もモデル化したい。例えば、「取る」という述語は、目的格が「パン」であれば主格は人である可能性が高いが、目的格が「新入社員」であれば主格は会社である可能性が高い。また、「息子」が「食べる」お弁当は「父」や「母」が「作る」可能性が高いといった関係のように、複数の述語の項をまたぐ知識関係も存在する*2。これをとらえるため、ある述語 w_{p_i} の語 w_t に関する推定を行う際に、文中の任意の述語 w_{p_j} の各語 $w_{t'}$ に関する推定情報を統合したい。この際、解析対象のペア (w_{p_i}, w_t) と関係の深い語の情報を中心に活用するために、アテンション機構 [1] を利用して情報の重み付けを行う。

具体的には、図4(b)のように、まず、任意の述語 w_{p_j} に関する RNN の出力系列 $h_j^k = h_{j,1}^k, \dots, h_{j,n}^k$ に対して、解析対象のペア (w_{p_i}, w_t) についての出力 $h_{i,t}^k$ に応じた注目度 $a_{i,j,t}(\cdot)$ を計算し、重み付きの和を取った表現 $h'_{i,j,t}$ を計算する。

$$h'_{i,j,t} = \sum_{t'} a_{i,j,t}(t') \cdot h_{j,t'}^k \quad (3)$$

$$a_{i,j,t}(t') = \frac{\exp(W_a g_{i,j,t,t'} + b_a)}{\sum_{t''} \exp(W_a g_{i,j,t,t''} + b_a)} \quad (4)$$

$$g_{i,j,t,t'} = \tanh(W_g[h_{i,t}^k, h_{j,t'}^k] + b_g) \quad (5)$$

その後、これを再び解析対象ペアの出力 $h_{i,t}^k$ と結合し、非線形変換を行ったものを max pooling にかけることで、各述語

*2 ただし、このような特定の述語項間の関係知識の獲得は、NTC 規模の学習データでは高頻度の僅かな事例に限定されると推測される。

の各語に対する推定情報を (w_{p_i}, w_t) の推定情報と統合する。

$$h_{i,t} = \text{maxpool}_j(f_{i,j,t}) \quad (6)$$

$$f_{i,j,t} = \text{ReLU}(W_f[h_{i,t}^k, h'_{i,j,t}] + b_f) \quad (7)$$

■(c) プーリング後アテンション: MP-POOL-ATT モデル前段落の手法では、文の単語数 n と述語数 q に対して nq^2 個のアテンションを計算する必要があり、大きな計算コストがかかる。そこで図 4(c) のように、まず、任意の解析対象のペア (w_{p_i}, w_t) に関して第一の手法と同様のプーリングを事前に行うことで、単語 w_t の他の述語に関する推定情報を統合した表現 $m_{i,t}$ を得る。その後、ある述語 w_{p_i} に関する各語の表現系列 $m_{i,1}, \dots, m_{i,n}$ に対して、現在の解析対象の語 w_t に対応する表現 $m_{i,t}$ を用いて注目度 $a_{i,t}(\cdot)$ を計算し、重み付きの和を取った表現 $h'_{i,t}$ を得る。さらに、この $h'_{i,j,t}$ と $m_{i,t}$ を結合し、非線形変換を行うことで、解析対象のペアに関する最終的な表現 $h_{i,t}$ を得る。この方法では、計算が必要なアテンションの数を nq に削減できる。

$$h_{i,t} = \text{ReLU}(W_n[m_{i,t}, h'_{i,t}] + b_h) \quad (8)$$

$$h'_{i,t} = \sum_{t'} a_{i,t}(t') \cdot m_{i,t'} \quad (9)$$

$$a_{i,t}(t') = \frac{\exp(W_a g_{i,t,t'} + b_a)}{\sum_{t''} \exp(W_a g_{i,t,t''} + b_a)} \quad (10)$$

$$g_{i,t,t'} = \tanh(W_g[m_{i,t}, m_{i,t'}] + b_g) \quad (11)$$

$$m_{i,t} = \text{maxpool}_j(f_{i,j,t}) \quad (12)$$

$$f_{i,j,t} = \text{ReLU}(W_f[h_{i,t}^k, h'_{j,t}] + b_f) \quad (13)$$

5 実験

実験では、我々の提案する拡張の効果を確認する。

5.1 実験設定

■データセット・評価指標 実験には NTC 1.5 を用いる。データは、既存研究で広く用いられている Taira et al. (2008) [12] の分割に習い訓練・開発・評価セットに分割する。評価値として、システムが出力した項の位置とラベルが、NTC で正解の項を表現している共参照クラスタの中のいずれかの語のそれと一致すれば正解、しなければ不正解とし、適合率、再現率、 F_1 値を求める。実験では、各モデルについて同一の学習データとハイパーパラメータで 10 回独立な学習を行い、それらの評価値の平均と偏差を表に記載した。

■ハイパーパラメータ 各ハイパーパラメータは開発データでの F_1 値が最大となるよう値を定めた。結果として、単語埋め込みベクトルの次元数に 256 を、多層双方向 GRU の層の深さ k に $\{1, 2, 3, 4, 5\}$ から 4 を、GRU の隠れ状態の次元数 d_u に $\{32, 64, 128, 256\}$ から 128 を、GRU のドロップアウト率に $\{0.0, 0.1, 0.2\}$ から 0.1 を、プーリング時の非線形変換 f, g の出力次元数に $\{512, 768, 1024\}$ から 1024 を、MP-POOL-ATT モデルにおける $h_{i,t}$ の次元数に 1024 を選択した。Loss 関数には negative log likelihood、最適化手法には Adam を採用した。学習では、開発セットでのモデルの F_1 値が 4 回改善されなかった場合に学習係数を半分にし、 F_1 値最大時のパラメータから学習を再開した。この過程を繰り返し、新たに設定する学習係数が初期値の 16 分の 1 を下回る場合に学習を終了した。各モデルの学習係数の初期値には $\{0.0001, 0.0002, 0.0005, 0.001\}$ から最も適切な値を選択し

た。各ラベルを出力するための閾値 θ_c は学習データでの F_1 値が最大となるよう 0.01 刻みで調整した。

■前処理 単語埋め込みベクトルの初期値には Matsubayashi & Inui (2017) [7] で用いられた手法と同一の手法を適用し、日本語 Wikipedia 2016 年 9 月 1 日の dump データより得られたベクトルを利用した。これらの単語ベクトルは学習時に更新した。NTC 上の語が Wikipedia 単語ベクトルの語彙に対して未知語となる場合には、同手法で得られる品詞ベクトルを利用した。NTC の単語分割、および品詞推定には JUMAN 辞書を用いた CaboCha v0.68 の解析結果を利用した。

5.2 実験結果

表 1 に我々が提案する各モデルの解析精度、および既存研究における解析精度を示す。

■各拡張要素の効果 表 1 の上部では、我々の提案する各拡張要素の効果を比較する。入力層を拡張するモデルである MP モデルは、ベースモデルと比較して全体の F_1 値で 0.3 ポイント強の改善をもたらしており、入力層に文中の全ての述語の位置を指示する新たな次元を追加する単純な拡張が、効果的に解析精度を向上させることが分かった。

出力層付近への拡張を加えた MP-POOL、MP-ATT-POOL、MP-POOL-ATT の 3 モデルは、いずれも MP からの解析精度の向上が認められたが、中でも MP-POOL-ATT が最も高い解析精度を示した。また、我々の提案する拡張は特に *Zero* の事例における解析性能の向上に顕著な貢献をしていることが分かった。個別の結果を見ると、MP-POOL と MP-ATT-POOL は全体の F_1 値では同等の性能を示しているが、相対的には MP-POOL では *Zero* 事例、MP-ATT-POOL では述語と項が直接の係り受け関係を持つ事例である *Dep* に特徴的な性能向上がみられた。MP-POOL-ATT は *Dep*、*Zero* いずれの事例も MP-POOL、MP-ATT-POOL 双方の性能を上回る解析精度となり、総合的に優れた性能を持つモデルであることが示された。係り受けの距離別で見ると、MP-POOL-ATT モデルでは表中の全ての距離区分でベースモデルに比べて 1~1.5 ポイントの性能改善がみられた。

■既存研究との比較 表 1 の下部では既存研究との解析精度の比較を行う。Ouchi et al. (2017) [9] は Grid RNN を用いて、同一文中の複数の述語に関する項推定の依存関係を考慮する最先端の end-to-end 日本語述語項構造解析モデルである。我々の MP-POOL-ATT は、彼らのモデルと比較して全体の F_1 値で 2.2 ポイントの精度向上を示した。ただし、彼らの実験では RNN の隠れ状態の次元数が 32 次元と比較的低次元であったため、より対等な比較を目的として、我々のベースモデルの隠れ状態の次元数を 32 次元とした実験も行った。結果として、我々のベースモデルは Ouchi et al. (2017) の性能を全体の F_1 値で 0.4 ポイント上回ることが分かった。Ouchi et al. (2017) と我々のモデルでは拡張部分の設計上、ベースモデルの構造も本質的に異なってしまうため、拡張部分のみの効果を直接的に比較することは難しいが、彼らのモデルが、複数述語の項推定の依存関係を考慮しないベースのモデル (single-seq モデル) に対して全体の F_1 値で 0.27 ポイントの向上を示しているのに対して、MP-POOL-ATT はベースモデルと比べて 0.53 ポイントの向上が認められる。

また、係り受け解析の情報を利用したモデルとして、Matsubayashi & Inui (2017) [7] のモデルとの比較を行った*3。こ

*3 ただし、Matsubayashi & Inui (2017) の解析精度は Ouchi et al. (2017) との厳密な比較のため、項が述語と同一文節内にある場合を評

Model	d_u	All		F_1 in different dependency distance							
		F_1 (%)	(σ)	Prec.	Rec.	Dep	Zero	2	3	4	≥ 5
Base model	128	83.09	(± 0.25)	85.64	80.70	89.76	53.26	59.95	51.6	43	32
MP	128	83.41	(± 0.15)	86.19	80.80	89.94	53.84	60.57	51.9	44	33
MP-POOL	128	83.56	(± 0.18)	86.20	81.08	90.01	54.46	61.35	52.5	45	32
MP-ATT-POOL	128	83.55	(± 0.15)	86.18	81.07	90.05	54.19	60.87	52.8	44	32
MP-POOL-ATT	128	83.62	(± 0.091)	86.09	81.29	90.07	54.53	61.22	53.1	44	33
Base model	32	81.80	(± 0.19)	84.65	79.13	88.83	50.24	56.78	48.6	41	29
Ouchi+ (2017) [8]	32	81.42				88.17	47.12				
M&I (2017) [7]		83.50	(± 0.17)	86.00	81.15	89.89	51.79	60.17	49.4	38	23
MP-POOL-ATT (ens. of 10 models)	128	85.25		88.31	82.39	91.11	57.59	64.12	56.8	47	34
M&I (2017) (ens. of 5 models)		84.07		86.09	82.15	90.24	53.66	61.94	51.8	40	24

表1 NTC 1.5 テストセットにおける各モデルの解析性能: d_u は RNN の隠れ状態の次元数。Dep は述語と項の文節係り受け距離が 1 の事例。Zero は距離が 2 以上の事例。

Model	Dep			Zero					
	ALL	ALL	NOM	ACC	DAT	ALL	NOM	ACC	DAT
MP-POOL-ATT	83.62	90.07	90.84	94.96	66.39	54.53	56.97	47.5	22
Ouchi+ (2017) [8]	81.42	88.17	88.75	93.68	64.38	47.12	50.65	32.4	7.5
M&I (2017) [7]	83.50	89.89	91.19	95.18	61.90	51.79	54.69	41.8	17
MP-POOL-ATT (ens. of 10 models)	85.25	91.11	91.86	95.55	69.54	57.59	59.84	51.6	24
M&I (2017) (ens. of 5 models)	84.07	90.24	91.59	95.29	62.61	53.66	56.47	44.7	16

表2 NTC 1.5 テストセットにおける格ごとの F_1 値 (%)

のモデルは NTC を Cabocha 0.68 で解析した結果を入力として用い、古典的な二値素性の他、述語と項候補の語の単語埋め込み、RNN に基づく係り受けパスの埋め込み等の特徴量として多層順伝播型ニューラルネットによりラベルの確率値を推定するモデルであり、NTC 1.5 において最高性能を示すモデルである。Matsubayashi & Inui (2017) は複数の述語に関する推定情報を統合せず、述語と項候補の語に対するローカルな推定を行うモデルではあるものの、我々の MP-POOL-ATT モデルは単一モデル間での比較において事例全体で 0.1 ポイントの改善、また、Dep、Zero のいずれの事例においてもこれを上回る解析性能を示し、NTC 1.5 における単一モデルでの最高精度を達成した。また、10 個の MP-POOL-ATT モデルの出力スコアの平均によって項を求めるアンサンブルモデルは、全体の F_1 値において単一モデルからさらに 1.6 ポイントの向上を示し、NTC 1.5 における最高精度を達成した。

表 2 には、格の種類ごとの結果を示した。単一モデルの比較において、我々の MP-POOL-ATT は Dep 事例のガ格、ヲ格では Matsubayashi & Inui (2017) に劣るものの、Zero 事例や二格の解析において解析精度を大きく向上させていることが分かる。アンサンブルモデルについてはいずれの格においても従来の解析精度を向上させる結果となった。

6 おわりに

本稿では、日本語述語項構造解析において複数の述語に関する項の推定情報を統合する新しい解析モデルを提案した。我々は RNN を用いた end-to-end の項構造解析モデルを拡張し、入力層と、最終層の直前に複数の述語情報を考慮するための機構を導入した。結果として、標準的なベンチマークデータである NTC 1.5 における最高精度を達成した。

一方で、実験で比較した Matsubayashi & Inui (2017) [7] は複数の述語に関する推定情報を統合しないシンプルなモデルであるが、単一モデルの解析精度は我々の MP-POOL-ATT モデルと僅差であり、ベースモデルとは F_1 値で 0.4 ポイント

優位である。このモデルは従来の項構造解析器と同様、文の係り受け構造を明示的に特徴量としており、今回の実験結果は、構文構造の有効な利用が我々のモデルにさらなる性能向上をもたらすことを示唆している。実際に、このような報告は英語の意味役割付与の研究においてもみられていることから [3, 6]、end-to-end モデルに対する構文構造の効果的な統合手法の開発は今後の我々の課題の一つである。

謝辞 東北大学の清野舜氏、渡邊研斗氏より本研究の遂行に有益なご意見をいただきました。また、本研究は JSPS 科研費 JP15K16045、および JP15H01702 の助成を受けたものです。

参考文献

- [1] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural Machine Translation by Jointly Learning to Align and Translate. In *ICLR*, pp. 1–15, 2015.
- [2] Kyunghyun Cho, Bart Van Merriënboer, Dzmitry Bahdanau, and Yoshua Bengio. On the Properties of Neural Machine Translation: Encoder-Decoder Approaches. In *SSST-8*, pp. 103–111, 2014.
- [3] Luheng He, Kenton Lee, Mike Lewis, Luke Zettlemoyer, and Paul G Allen. Deep Semantic Role Labeling: What Works and What’s Next. In *ACL*, pp. 473–483, 2017.
- [4] 飯田龍, 小町守, 井之上直也, 乾健太郎, 松本裕治. 述語項構造と照応関係のアンテーション: Naist テキストコーパス構築の経験から. 自然言語処理, Vol. 17, No. 2, pp. 25–50, 2010.
- [5] Ryu Iida, Kentaro Torisawa, Chikara Hashimoto, Jong-Hoon Oh, and Julien Kloetzer. Intra-sentential Zero Anaphora Resolution using Subject Sharing Recognition. In *EMNLP*, pp. 2179–2189, 2015.
- [6] Diego Marcheggiani and Ivan Titov. Encoding Sentences with Graph Convolutional Networks for Semantic Role Labeling. In *EMNLP*, pp. 1507–1516, 2017.
- [7] Yuichiro Matsubayashi and Kentaro Inui. Revisiting the Design Issues of Local Models for Japanese Predicate-Argument Structure Analysis. In *IJCNLP*, pp. 128–133, 2017.
- [8] Hiroki Ouchi, Hiroyuki Shindo, Kevin Duh, and Yuji Matsumoto. Joint Case Argument Identification for Japanese Predicate Argument Structure Analysis. In *ACL-IJCNLP*, pp. 961–970, 2015.
- [9] Hiroki Ouchi, Hiroyuki Shindo, and Yuji Matsumoto. Neural Modeling of Multi-Predicate Interactions for Japanese Predicate Argument Structure Analysis. In *ACL*, pp. 1591–1600, 2017.
- [10] Ryohei Sasano and Sadao Kurohashi. A Discriminative Approach to Japanese Zero Anaphora Resolution with Large-scale Lexicalized Case Frames. In *IJCNLP*, pp. 758–766, 2011.
- [11] Tomohide Shibata, Daisuke Kawahara, and Sadao Kurohashi. Neural Network-Based Model for Japanese Predicate Argument Structure Analysis. In *ACL*, pp. 1235–1244, 2016.
- [12] Hirotoshi Taira, Sanae Fujita, and Masaaki Nagata. A Japanese Predicate Argument Structure Analysis using Decision Lists. In *EMNLP*, pp. 523–532, 2008.
- [13] Jie Zhou and Wei Xu. End-to-end Learning of Semantic Role Labeling Using Recurrent Neural Networks. In *ACL*, pp. 1127–1137, 2015.

価対象から除外して再計算を行った。