

# 系列ラベリングを統合した遷移システムによる 漸進的意味役割付与

大内 啓樹<sup>1,2</sup> 吉野 幸一郎<sup>1,3</sup> 中村 哲<sup>1,2</sup> 松本 裕治<sup>1,2</sup>

<sup>1</sup> 奈良先端科学技術大学院大学 情報科学研究科

<sup>2</sup> 理化学研究所 革新知能統合研究センター AIP

<sup>3</sup> 科学技術振興機構 さきがけ

{ouchi.hiroki.nt6,koichiro,s-nakamura,matsu}@is.naist.jp

## 1 はじめに

意味役割付与は、「いつ、どこで、誰が、何を、誰に、どうした」といった述語と項の関係を同定するタスクである。図1は意味役割付与の例を示している。例えば、述語 *makes* の動作主 (A0) は *She* であり、被動作主 (A1) は *computers* である。こうした解析における先行研究の多くは、完結した文を入力されることを前提としていた。これに対し、文末を待たずに、逐次的な入力に対して意味役割付与を漸進的に行う技術が注目を集めている [7]。

漸進的意味役割付与研究の動機の一つとして、音声対話・同時翻訳などの音声言語システムへの応用 [2, 9] が挙げられる。こうしたシステムではリアルタイム性が重要な要素であり、音声認識からの逐次的な単語入力に対し、各入力時点に応じた意味役割情報が必要となる。もう一つの動機として、人間の言語理解過程における漸進性のモデル化が挙げられる。意味役割認識過程における漸進性は心理言語学などの分野で認識されており [11]、漸進的な意味役割付与モデルを構築することは人間の言語理解過程を明らかにする上でも重要である。

漸進的意味役割付与では、時刻  $t$  までに入力された単語列  $w_{1:t} = w_1, \dots, w_t$  に対して決定可能な意味役割を予測する。図2は、時刻4で  $w_4 = \text{repairs}$  が入力された際の例を表している。ここで、述語 *makes* と *repairs* が共に *She* を動作主 (A0) の項として持つことが示されている。一方、被動作主 (A1) となる単語はまだ入力されておらず、解析が保留となっている。既存研究 [7] では、木接合文法に基づいてこうした解析を行なっている。しかし、木接合文法の学習にはコストがかかり、かつ、構文解析誤りが意味役割付与の予測性能を大きく損なう問題がある。

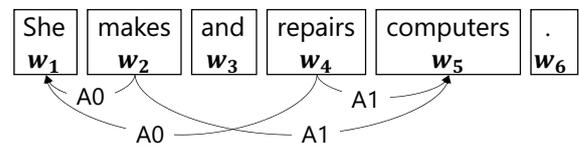


図1: 意味役割付与の例。

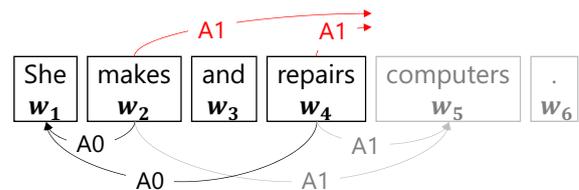


図2: 逐次的入力に対する評価対象の意味役割 (単語  $w_4$  入力時) の例。

本研究では、構文情報を用いず単語情報のみから高精度な漸進的意味役割付与を可能とする手法を提案する。提案手法は、構文解析分野で使用される遷移システム (Transition-Based System) を意味役割付与のために拡張し、系列ラベリング技術を統合することによって、各時刻における意味役割付与を行う。

## 2 提案手法

遷移型依存構造解析アルゴリズムを拡張し、漸進的意味役割付与に特化した手法を提案する。このアルゴリズムでは、入力された単語列をバッファに、述語と判断された単語をスタックに保持する。また状態遷移操作に応じ、スタックに保持された述語とバッファに保持された単語間の意味役割を予測する。本章では提案アルゴリズムについて詳述する。

タイプ	操作	現在の状態	次の状態
Shift	PRED	$[S], [w_{1:t} B], [A]$	$[S w_t], [w_{1:t} B], [A]$
	NO-PRED	$[S], [w_{1:t} B], [A]$	$[S], [w_{1:t} B], [A]$
Label	$\text{LABEL}(\{p_i\}_1^L)$	$[S p_{1:L}], [w_{1:t} B], [A]$	$[S p_{1:L}], [w_{1:t} B], [A \cup \{y_{1:t}^i\}_1^L]$

表 1: 提案手法における状態遷移操作.

## 2.1 遷移システム

遷移システムの状態  $c$  は,  $(S, B, A)$  の三つ組で定義される.  $S$  はスタック,  $B$  はバッファ,  $A$  はラベル集合である. スタックには, 「述語である」と判断された単語が積まれる. バッファには, 各時刻の入力単語が入る. ラベル集合には, 予測された意味役割ラベルの集合が逐次的に追加される.

遷移システムの操作は, 表 1 が示す 2 つの Shift 系操作と 1 つの Label 系操作として定義される. まず, 時刻  $t$  に新しい単語  $w_t$  がバッファの末尾へ追加され, この単語に対して Shift 系操作がなされる. Shift 系操作の PRED は, バッファの末尾の単語  $w_t$  をスタックに積む. これは, 単語  $w_t$  が「述語である」と判断された場合に行われる. 一方, もう 1 つの Shift 系操作である NO-PRED は, 単語  $w_t$  が「述語ではない」と判断された場合に選ばれ, 状態に変化を及ぼさず次のアクションである Label 系操作へと移行する. Label 系操作である  $\text{LABEL}(\{p_i\}_1^L)$  は, スタックに積まれた述語列  $p_{1:L}$  の各述語  $p_i$  に対して, バッファ内にある単語全ての意味役割を予測する. ここで  $L$  は, スタックに積まれた述語数である.

これらの操作は一定の順序で必ず行われる. つまり, 時刻  $t$  において単語  $w_t$  がバッファに入力された際に, Shift 系操作 1 つ (PRED または NO-PRED) と Label 系操作 1 つ (LABEL) が順に行われる. この点が, 一般的な遷移型構文解析アルゴリズムと大きく異なる.

## 2.2 操作に利用するモデル

### Shift 系操作に利用するモデル

各時刻で行う Shift 系操作として, PRED と NO-PRED のどちらかを選ぶ二値分類問題を解く. まず, バッファに保持された単語列  $w_{1:t}$  から特徴量ベクトル  $\mathbf{h}$  を計算する. この計算には双方向型リカレントニューラルネットワーク (Bi-RNN) [4] を用いる.

$$\mathbf{h} = \text{Bi-RNN}(w_{1:t})$$

次に, シフト系操作として行うべき操作が PRED である確率を, ベクトル  $\mathbf{h}$  を用いて計算する.

$$\text{Pr}(\text{PRED}) = \text{sigmoid}(\mathbf{w} \mathbf{h})$$

ここで,  $\mathbf{w}$  は  $\mathbf{h}$  と同次元のパラメータベクトルであり, sigmoid はシグモイド関数である. この確率が 0.5 以上である場合は PRED が選ばれ, 0.5 未満の場合は NO-PRED が選ばれることとなる. パラメータの学習はバイナリ交差エントロピーに基づいて行う.

### Label 系操作に利用するモデル

LABEL 操作では, スタックの各述語  $p$  とバッファに保持された単語列  $w_{1:t}$  を引数とし, 系列ラベリング問題として意味役割を予測する. まず, Bi-RNN を用いて単語列  $w_{1:t}$  に対応する特徴量ベクトル  $\mathbf{h}_{1:t}$  を計算する.

$$\mathbf{h}_{1:t} = \text{Bi-RNN}(w_{1:t}, p)$$

次に, ベクトル  $\mathbf{h}_{1:t}$  を用いて意味役割ラベルの確率を計算する.

$$\mathbf{y}_{1:t} = \text{softmax}(\mathbf{h}_{1:t})$$

ここで,  $\mathbf{y}_{1:t}$  は単語列  $w_{1:t}$  に対応するラベル確率ベクトル列である. つまり, 各ベクトル  $\mathbf{y}_i$  の各次元  $l$  はラベル  $l$  の確率値である. これにより, 各ベクトルにおける確率最大のラベルを意味役割として出力する. パラメータの学習は交差エントロピーに基づいて行う.

## 2.3 解析例

表 2 は提案手法での解析例を示している. 時刻  $t$  は  $t$  番目の単語がバッファに入力された時点を示している. 時刻 1 でバッファに She が入力され, この単語は述語ではないため, スタックには積まれない. 時刻 2 でバッファに makes が入力され, この単語は述語であるため, PRED によってスタックに積まれる. また, LABEL(makes) によって, スタックに積まれた述語  $p_1 = \text{makes}$  に対するバッファの単語  $w_{1:2}$  の意味役割ラベル予測が行われる. 予測されたラベル列  $y_{1:2}^i$  は, 時刻 2 における述語  $p_1 = \text{makes}$  の意味役割として保存される. 以降も同様に, 文末まで解析が行われる.

## 3 評価実験

CoNLL-2009 Shared Task [5] の意味役割付与データセットを用いて, 提案手法の有効性を検証する. 評価には先行研究で定義された指標を用い, また漸進的意

Time	Action	Stack	Buffer	A
0		[]	[]	
1	NO-PRED LABEL( $\phi$ )	[]	[She]	$A \cup \phi$
2	PRED LABEL({makes})	[makes]	[She makes]	$A \cup y_{1:2}^1$
3	NO-PRED LABEL({makes})	[makes]	[She makes and]	$A \cup y_{1:3}^1$
4	PRED LABEL({makes, repairs})	[makes repairs]	[She makes and repairs]	$A \cup \{y_{1:4}^i\}_1^2$
5	NO-PRED LABEL({makes, repairs})	[makes repairs]	[She makes and repairs computers]	$A \cup \{y_{1:5}^i\}_1^2$
6	NO-PRED LABEL({makes, repairs})	[makes repairs]	[She makes and repairs computers.]	$A \cup \{y_{1:6}^i\}_1^2$

表 2: 提案手法での解析例.

味役割付与における先行研究との比較を行った. 提案手法は遷移システムにおける操作の決定と意味役割付与の系列ラベリングを含むため, 遷移操作に正解が与えられた場合 (Shift-Gold) と, 遷移操作もモデル予測した場合 (Shift-Auto) の 2 種類を調査した.

### 3.1 既存手法

設定が異なるため厳密な比較はできないが, [7] で提案された以下の手法を参考として記載する.

- Konstas+ 14 (LE-Gold)
- Konstas+ 14 (LE-Auto)
- MAJORITYBASELINE
- MALTBASELINE

1 つ目の手法の Konstas+ 14 (LE-Gold) は, 木接合文法に基づいて構文解析する際に, 正解の Lexical Entry (LE-Gold) を使用する. したがって, 品詞などはすべて正解のものであり, 解析対象の述語も正解のものを容易に選ぶことができる. 2 つ目の手法の Konstas+ 14 (LE-Auto) は, 予測した Lexical Entry (LE-Auto) を使用しているため, より現実的な設定となっている. 3 つ目の手法の MAJORITYBASELINE は, 述語項間に正解の係り受けラベルが与えられた上で最も頻度の高い意味役割を選択する. 4 つ目の手法の MALTBASELINE は MAJORITYBASELINE と同様, 係り受けラベルに対して最も頻度の高い意味役割を選択するが, 使用する係り受けラベルは MALTPARSER<sup>1</sup> で予測したものである.

### 3.2 評価指標

評価指標として, Full-sentence SRL Score (FSS) と Combined Incremental SRL Score (CISS) [7] の 2 つを

<sup>1</sup><http://www.maltparser.org/>

用いる. FSS は, CoNLL-2009 Shared Task で用いられる標準的な指標である. つまり, 末尾まで文の入力を終えた段階での意味役割の F1 値を評価する. CISS は, 時刻  $t$  で推定可能な意味役割のみを評価する. 例えば, 図 2 の  $w_4$ =repairs が入力された時点で, 既知の単語から解くことができるのは  $A_0(\text{She, makes})$  および  $A_0(\text{She, repairs})$  のみである. そのため, これらの関係が解けたかどうかだけを F1 値で評価する. 未入力の単語との関係である  $A_1(\text{computers, makes})$  および  $A_1(\text{computers, repairs})$  は評価対象外である. 漸進的意味役割付与では, 各単語入力時点での解析性能が重要となるため, このような指標が用いられる.

### 3.3 実験結果

#### CISS による評価

図 3 は, CISS の F1 値を表している. 既存研究 [7] に従い, 単語数が 10, 20, 30, 40 語からなる文の各時刻における F1 値をプロットしている. 提案手法の Shift-Gold も Shift-Auto も同様の傾向が見られた. つまり, 文の長さに関わらず, 文のはじめの方の入力時点で F1 値のピークを迎えている. これは, 図 2 のように,  $A_0$  の意味役割が文のはじめの方に出現しやすく, かつ, その他の意味役割よりも比較的予測しやすいことに起因していると考えられる. これは, 既存研究 [7] も同様の傾向を示している. 一方で, 既存研究 [7] ではピークを迎えてから緩やかに精度が下がり続ける傾向が見られたが, 本提案手法はほぼ一定の精度を保ったまま文の末尾まで解析が実現できている.

#### FSS による評価

表 3 は, FSS の F1 値<sup>2</sup>を表している. 提案手法は, 構

<sup>2</sup>Konstas+ 14, MAJORITYBASELINE, MALTBASELINE の結果は [7] から引用.

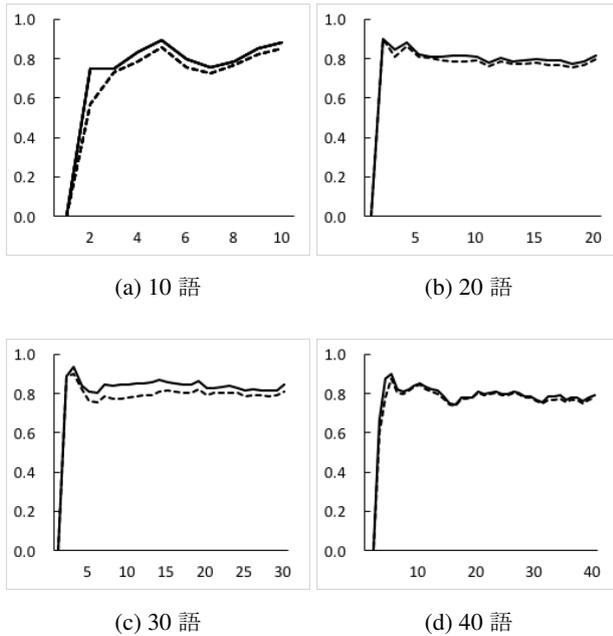


図 3: Combined Incremental SRL Score (CISS). 実線・点線は提案手法の Shift-Gold・Auto を表している。

System	P	R	F1
Proposed (Shift-Gold)	84.78	80.96	82.82
Proposed (Shift-Auto)	81.24	78.11	79.65
Konstas+ 14 (LE-Gold)	81.48	75.51	78.38
Konstas+ 14 (LE-Auto)	-	-	68.00
MAJORITYBASELINE	71.05	58.10	63.92
MALTBASELINE	60.90	46.14	52.50

表 3: Full-sentence SRL score (FSS).

文情報を用いずに高い F1 値を記録している。正解の Shift 操作 (Shift-Gold) の場合は F1 値が 82.82%, モデルで予測した Shift 操作 (Shift-Auto) の場合は 79.65% の F1 値を記録している。これらの結果から、文末まで入力を終えた時点でも本提案手法が有効に解析を行っていることが確認された。

## 4 関連研究

漸進的言語解析に関する研究は、これまでも多く行われてきた [3, 10, 6]。中でも、本研究に最も関連する研究として、Konstas ら [7] の漸進的意味役割付与が挙げられる。彼らは、漸進的意味役割付与の問題設定と評価指標の定式化を行なった。また、木接合文法に基づく漸進的な解析手法を提案し、ベースラインを上回る結果を報告している。本研究で提案する手法は、構文情報を用いずに、高精度な解析を行うことができる点が異なる。

類似した研究として、Choi ら [1] は遷移型システムを用いた意味役割付与手法を提案している。彼らの手法では、6 つの遷移操作によって遷移し、その過程で意味役割の予測が行われる。この手法は 1 文全体の入力を前提として設計されているため、単語の逐次入力に対応できない点が本研究と異なる。

## 5 おわりに

本研究では系列ラベリングを統合した遷移システムによる漸進的意味役割付与に取り組んだ。評価実験の結果から、提案手法は (i) 長い文でも頑健な漸進的解析が可能であり、(ii) 構文情報を使わずとも高精度に解析可能であることがわかった。本提案手法は多言語解析にも適用可能であるため、日本語を含めた他の言語への応用が可能である。また、本手法を対話・同時翻訳システムなどに組み込むことも予定している。

## 謝辞

本研究は JST さきがけ (課題番号:JPMJPR165B) の支援を受けて行った。

## 参考文献

- [1] Jinho D Choi and Martha Palmer. Transition-based semantic role labeling using predicate argument clustering. In *Proceedings of ACL Workshop on Relational Models of Semantics*, pages 37–45, 2011.
- [2] Bonaventura Coppola, Alessandro Moschitti, and Giuseppe Riccardi. Shallow semantic parsing for spoken language understanding. In *Proceedings of HLT-NAACL*, pages 85–88, 2009.
- [3] David DeVault, Kenji Sagae, and David Traum. Incremental interpretation and prediction of utterance meaning for interactive dialogue. *Dialogue & Discourse*, 2(1):143–170, 2011.
- [4] Alan Graves, Navdeep Jaitly, and Abdel-rahman Mohamed. Hybrid speech recognition with deep bidirectional lstm. In *Automatic Speech Recognition and Understanding (ASRU), 2013 IEEE Workshop*, 2013.
- [5] Jan Hajič et al. The CoNLL-2009 shared task: syntactic and semantic dependencies in multiple languages. In *Proceedings of CoNLL: Shared Task*, pages 1–18, 2009.
- [6] Arne Köhn and Wolfgang Menzel. Incremental predictive parsing with tur-boparser. In *Proceedings of ACL*, pages 803–808, 2014.
- [7] Ioannis Konstas, Frank Keller, Vera Demberg, and Mirella Lapata. Incremental semantic role labeling with tree adjoining grammar. In *Proceedings of EMNLP*, 2014.
- [8] Joakim Nivre. Incrementality in deterministic dependency parsing. In *Proceedings of the Workshop on Incremental Parsing: Bringing Engineering and Cognition Together*, pages 50–57, 2004.
- [9] Yusuke Oda, Graham Neubig, Sakriani Sakti, Tomoki Toda, and Satoshi Nakamura. Optimizing segmentation strategies for simultaneous speech translation. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 551–556, 2014.
- [10] Tomohiro Ohno, Shigeki Matsubara, Hideki Kashioka, Naoto Kato, and Yasuyoshi Inagaki. Incremental dependency parsing of japanese spoken monologue based on clause boundaries. In *Proceedings of Speech Communication and Technology*, pages 3449–3452, 2005.
- [11] Martin J Pickering, Matthew J Traxler, and Matthew W Crocker. Ambiguity resolution in sentence processing: Evidence against frequency-based accounts. *Journal of memory and language*, 43(3):447–475, 2000.