

# 係り元の末尾表現に着目した Hierarchical Tag Context Tree を利用した日本語意味役割付与システムの構築

石原 靖弘<sup>1,a)</sup> 竹内 孔一<sup>1,b)</sup>

受付日 2015年10月19日, 採録日 2016年4月5日

**概要:** 近年, 言語処理において, 述語とその係り元との関係をタイプ分けする意味役割付与の研究が英語圏を中心に進められている. 英語の意味役割付与では統計的学習モデルに対して構文解析情報を特徴量として利用している. 一方, 日本語では表層格を関係ラベルとする述語項構造解析において有効な特徴量が提案されており, 意味役割付与においても同様に有効であると考えられるが, より多様な関係ラベルを扱う意味役割付与において他に有効な特徴量が存在するかについては検討の余地があると考えられる. しかしながら, 日本語では意味役割付与データが少なく, そのような検討についてはこれまでほとんどなされていない. そこで, 最近構築された日本語意味役割付与データに対して, 統計的学習モデルによる意味役割付与システムを構築し, どのような特徴量が精度向上に寄与するか明らかにする. 本論文では日本語の機能表現が意味役割の決定に関与することに注目し, 階層的な可変長 n-gram をコーパスから獲得する HTCT (階層的タグ文脈木) を特徴量抽出器として利用する. 意味役割付与実験において HTCT から得た特徴量を使用した場合に付与精度が向上することを示す. また機能語辞書や長単位解析器を利用した場合と比較して HTCT を利用して抽出した特徴量がより効果的であることを示す. さらに, コーパスから直接獲得した固定長 n-gram の特徴量と組み合わせることでより良い精度が得られることを示す. また HTCT を含む提案する特徴量は述語項構造解析の特徴量とともに使用した場合においても有効に働くことを示す.

キーワード: 意味役割付与, 述語項構造解析, 階層的タグ文脈木

## Construction of Japanese Semantic Role Labeling System Using Hierarchical Tag Context Trees Extracted from Tail Expressions of Dependency Elements

YASUHIRO ISHIHARA<sup>1,a)</sup> KOICHI TAKEUCHI<sup>1,b)</sup>

Received: October 19, 2015, Accepted: April 5, 2016

**Abstract:** In recent natural language processing, semantic role labeling (SRL) which determines semantic relations between predicates and their arguments have been studied especially for English language. In English, syntactic features are effective for statistical learning model-based semantic role labelers. In Japanese, on the other hand, effective features have been being revealed in Japanese case marker-based predicate argument structure analysis, and they should also be effective for Japanese SRL because of the similarity of the tasks, however, just the features might not be enough for Japanese SRL due to the diversity of the labels it considers. Unfortunately, there had been few language resources in Japanese that can be used for SRL, and therefore, researches on effective features for SRL as well. This paper reveals effective features for SRL systems based on statistical learning approach with a Japanese semantic role labeled corpus released recently. In the preliminary study we found that functional multi-word expressions in the arguments have a great influence on determining their semantic role labels. Thus we exploit hierarchical tag context trees (HTCTs), which can obtain variable length n-grams to extract generalized functional multi-word expressions as features for SRL systems. The experimental results show that the SRL systems added the features obtained by HTCTs outperform the system with the features extracted using a dictionary on multiword expressions. Additionally, the systems using both the variable length features with HTCTs and the fixed length features directly extracted from the corpus show a better accuracy. The experimental comparison with the features developed in Japanese predicate-argument analysis shows that our proposed features are still effective when used with them.

**Keywords:** semantic role labeling, predicate-argument structure analysis, hierarchical tag context trees

## 1. はじめに

文の係り受け関係解析をさらに発展させた意味役割付与 (Semantic Role Labeling) が提案され、機械翻訳や作文支援、質問応答などの応用タスクにおいて利用されその有効性が示されている [1], [2], [3]. 意味役割付与とは文中の述語に対する係り元となる構成要素 (項と呼ぶ) に対し、述語との意味的關係を表すラベル (意味役割と呼ぶ) を付与するタスクであるが、本研究では項は係り元の構成要素だけでなく、下記のような場合も項として意味役割付与の対象とする。

1. 【対象】 詰め将棋の本を 【場所】 秋葉原で 買った
2. 【対象】 親を 慕う 【修飾】 姿
3. 【原因】 左側から衝突されたようになってるので、骨折しました
4. 【対象】 開いちゃったのを 熱湯に入れて、【様態】 少し かき回し してみる

それぞれの例文で太字が述語であり、[] 内が項を示し、【】内が意味役割である。1. では「詰め将棋の本を」という項は「買う」に対して「対象」という意味的關係にあり、「秋葉原で」という項は「場所」という意味的關係にあることを示している。2. は被連体修飾である「姿」に意味役割が付与される例である。3. は項が単純な名詞句ではなく、述語を含む文になっている例である。4. では並列構造により直接の係り元ではない要素に項として意味役割を付与する例である。

このような述語項間の意味的關係を表す意味役割には様々な種類のもものが提案されている。たとえば言語学の分析 [4], [5] をもとに Agent, Theme, Recipient といった関係概念を固定するものや、言語処理を重視して Arg1, Arg2 といった番号ベースにするもの [6] が英語を中心に提案されており、前者は VerbNet [7] や FrameNet [8], 後者は PropBank として意味役割が文に付与された事例データが利用可能である。

英語の意味役割付与では構文解析情報が重要であることが先行研究から示されている [9]. 一方、日本語では、表層格を述語項間の関係ラベルとして利用する述語項構造解析において、ゼロ照応を含む項に対して格関係の解析が研究されている。たとえば文献 [10] では NAIST テキストコーパスにおいて述語項構造解析で有効な特徴量が整理されて調べられている。これらの特徴量は意味役割付与においても期待できるが、ガ格、ヲ格、ニ格以外の関係も取り入れる意味役割付与において、提案されているもの以外にも有効な特徴量が存在するか検討する余地があると考えられる。

<sup>1</sup> 岡山大学大学院自然科学研究科  
Graduate School of Natural Science and Technology  
Okayama University, Okayama 700-8530, Japan

a) ishihara@cl.cs.okayama-u.ac.jp

b) koichi@cl.cs.okayama-u.ac.jp

そこで本論文では日本語の意味役割付与に対する効果的な特徴量を実験的に明らかにすることを目標とする。実験データとしては、近年、日本語で分野のバランスをとった均衡コーパス (BCCWJ) 上に意味役割が付与されたデータが構築されつつある [11]. こうした実験データに対して統計的手法に基づく意味役割付与システムを構築することで有効な特徴量を検討する。

日本語文法の分析 [12] によると助詞と同様に機能的な働きをする要素として複合助詞の存在が議論されている。たとえば「花子へ手紙を送った」は複合助詞「に対して」を用いて「花子に対して手紙を送った」と言い換えても同様の意味であり、「花子」は「送る」に対して送り先という意味的關係は変わらない。つまり、意味役割としては同じものを付与することになる。よって複合助詞は意味役割付与における特徴量として効果的と考えられる。日本語の形態素解析器 MeCab (IPADIC) では「に対して」など格助詞の連語として登録されているが、たとえば「ナイフでもってチーズを切る」[12] における「でもって」は連語として解析されない<sup>\*1</sup>など登録されていない連語も存在する<sup>\*2</sup>。また機能語を整理した日本語機能表現辞書「つつじ」[13] では上記の「でもって」なども登録されているが、複合助詞「といっしょに」<sup>\*3</sup>など意味役割に関連する表現が集められているわけではない。

助詞や複合助詞を意味役割の同定における特徴量として利用しようとする、現段階では網羅された辞書や言語ツールが存在しないため、学習データから獲得する必要がある。その際、文字列だけでなく読みや品詞などに抽象化した構造も利用したい。たとえば上記の「といっしょに」は漢字で「と一緒に」と表現されることもあるが読みでとらえれば片方が学習データに出現しなかったとしても、同じ表現としてとらえることができる。また「によらず」「にかかわらず」[12] などは「に-動詞-ず」または「助詞-動詞-助動詞」のように品詞レベルを利用すれば、知られていない複合助詞を広くとらえることができる。より拡張して意味役割付与の特徴量としてとらえ直すと、項の末尾表現に対して、表層や基本形、読み、品詞、といった単位の階層レベルを混在した連続列からなる特徴量が意味役割付与の手がかりとして有効である可能性がある。

そこで本研究では階層性のある連続列から有効な可変長 n-gram を取り出す Hierarchical Tag Context Tree (以降、HTCT と記述) [14] を利用する。項の末尾表現に HTCT を

<sup>\*1</sup> IPADIC Version 2.7.0-20070801 では連語として登録されていなかった。しかしながらマニュアル (<http://chasen.naist.jp/snapshot/ipadic/ipadic/doc/ipadic-ja.pdf>) には記載されており、一貫性やメンテナンスの問題と見受けられる。

<sup>\*2</sup> この文では「ナイフ」は「切る」に対する手段としての意味的關係である [12].

<sup>\*3</sup> 「おじいちゃんが子供といっしょに散歩する」では「子供」は「散歩する」に対して、動作主体である「おじいちゃん」と同様、共同動作を行う相手となる [12].

適用して、項の意味役割識別に有効な n-gram 特徴量を抽出する。意味役割付与システムとして Conditional Random Field (CRF) を利用し、HTCT で獲得した n-gram を特徴量として利用し、意味役割を付与する。意味役割付与実験の結果から、HTCT の特徴量を加えたモデルの方が、HTCT の特徴量を加えない場合や、機能語辞書「つつじ」を利用した特徴量を加えた場合よりも精度が向上したことを報告する。また日本語の述語項構造解析で明らかになった有効な特徴量との比較を行い、そのような特徴量とともに用いた場合においても、HTCT の特徴量を含む項末尾の特徴量は有効に働くことを示す。

以下では、2 章で HTCT を利用した特徴量の抽出法について述べ、3 章で意味役割付与実験について述べる。4 章は関連研究について述べ、5 章でまとめと今後の展望を述べる。

## 2. HTCT による項末尾の特徴量の抽出

本研究で提案する HTCT は一般的な確率的接尾辞木の予測対象を意味役割に変更し、さらに文脈構成要素に対する階層性を取り入れたモデルである。まず、下記では意味役割を予測する HTCT について説明した後、項の末尾表現から関連のある特徴量を取り出す方法についての記述する。提案する HTCT の構築法自体は特定のコーパスの構造に依存するものではないが、対象とする学習コーパスは 3.2 節で述べるように 1 文は長単位\*4を最小単位として表層や読みが付与されており、項と意味役割は長単位列に対して付与されている。よってここでは長単位をベースに確率的接尾辞の構築を説明する。

### 2.1 意味役割を予測する HTCT

本研究では HTCT を利用して意味役割付与に有効な長単位連続列を特徴量として取り出したい。そこで、通常の確率的接尾辞木の予測対象を意味役割に変えたモデルを導入する。まず、階層性を考慮しない長単位列に対する確率的接尾辞木について記述した後、階層性を考慮した HTCT について述べる。図 1 に 3 長単位 (a, b, c) の場合に意味役割  $l$  を予測する確率的接尾辞木の例を示す。接尾辞木は根ノードから各ノードまでの枝が文脈となる長単位列を表す。通常の確率的接尾辞木では各ノードではそれまでの文脈長単位列に対して、次の長単位が出現する確率を定義するが、ここでは意味役割を予測する。

たとえば根ノードは空の長単位  $\epsilon$  を表し、文脈がない場合の意味役割  $l$  の出現確率は式で表すと  $p(l|\epsilon)$  となる。また、ノード  $ab$  では、文脈長単位列  $ab$  のときの意味役割の出現確率  $p(l|ab)$  を保持している。よってこの確率的接尾辞木を作成することで、可変長の長単位に対する意味役割

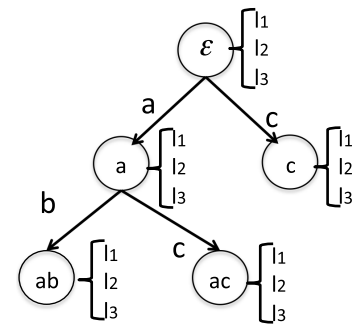


図 1 確率的接尾辞木の例

Fig. 1 Example of probabilistic suffix trees.

の出現確率を計算することができる。

次に確率的接尾辞木を構築する方法について述べる。今、文脈を  $c$  としたときにある意味役割  $l$  が出る確率を  $P(l|c)$  とすると、頻度  $n(l|c)$  を用いて下記のように定義する。

$$P(l|c) = \frac{n(l|c)}{\sum_{x \in L} n(x|c)} \quad (1)$$

ここで  $L$  は候補となる意味役割すべての集合を表す。長単位列と意味役割の確率的接尾辞木は既存の文脈に対して、新たな長単位を加えることを再帰的に行うことで構築する。構築のポイントは既存の文脈を持つノードに対して新たな長単位を加えて、より長い文脈のノードを加えるかどうか判定する部分である。この評価式として先行研究 [14], [15] で提案されている情報量基準に基づく手法を採用し、文脈  $c$  に対して新たな長単位  $t$  を加えた場合の評価値  $\Delta(ct)$  を下記のように定義する。

$$\begin{aligned} \Delta(ct) &= \sum_{l \in L} n(l|ct) \log \frac{P(l|ct)}{P(l|c)} \\ &= n(ct) \sum_{l \in L} P(l|ct) \log \frac{P(l|ct)}{P(l|c)} \\ &= n(ct) D_{KL}(P(\cdot|ct), P(\cdot|c)) \end{aligned} \quad (2)$$

ここで、 $n(ct)$  は文脈  $c$  に続いて  $t$  が出現した回数を表し、 $D_{KL}$  はカルバックライブラ情報量を表す。この式は新たな長い文脈  $ct$  での確率分布と既存の文脈  $c$  での確率分布との差を評価した値で、 $\Delta(ct)$  がしきい値を超えた場合、新たなノードを追加する。よって確率的接尾辞木は根ノード  $\epsilon$  から順に式 (2) を適用し、しきい値を超えたノードを再帰的に追加する山登り法で構築する。

次に確率的接尾辞木に階層性を導入する。1 つの長単位には表層文字、基本形、読み、品詞の 4 種類の長単位情報 (以降、タグまたは文脈タグと記述) があることを仮定し、図 2 に示すように上位ほど抽象的なタグとなる階層構造を設定する。これにより上位のタグほど、表層の異なりを吸収した適応範囲の広い文脈を指定することができる。確率的接尾辞木の枝はこの 4 階層のうちのいずれかのタグを選択しながら新たなノードを追加する。長単位に対する 4 階層のタグの集合を  $T$  とし、既存の文脈タグ列  $c$  に対して

\*4 [http://pj.ninjal.ac.jp/corpus\\_center/bccwj/morphology.html](http://pj.ninjal.ac.jp/corpus_center/bccwj/morphology.html)

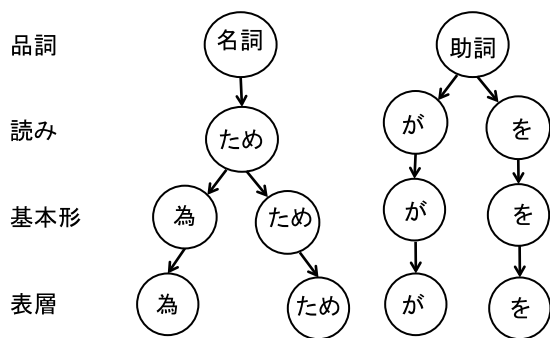


図 2 長単位の 4 つの階層構造

Fig. 2 Hierarchical structure of a long-unit-word.

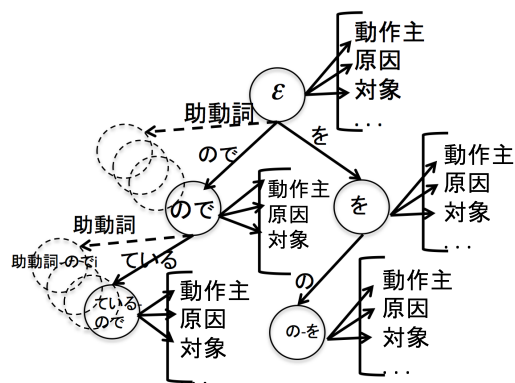


図 3 階層性を考慮した確率的接尾辞木

Fig. 3 Probabilistic suffix tree applied with a long-unit-word hierarchy.

新たな文脈タグ  $t$  を追加するかどうかの評価は式 (2) をもとに式 (3) のように定義する.

$$\max_{t \in T} \Delta(ct) \tag{3}$$

つまり、4 階層の各層それぞれに対して式 (2) を適用して、しきい値を超える中で最も値が大きい文脈タグのノードを追加する。もし、4 階層ともしきい値を超えなければ、新たな文脈タグ  $t$  は追加しない。接尾辞木全体を構築するには、階層がない場合と同様に根ノードから有効な文脈タグ列を式 (3) によって順次追加していく山登り法を採用する。

### 2.2 項に対する HTCT の構築

前節で定義した HTCT を構築し、HTCT から有効な文脈タグ  $n$ -gram を取り出し、CRF ベースの意味役割付与システム (詳細は 3.3 節参照) の特徴量として利用する。以下では、項に対する HTCT の構築法について具体的に記述した後、HTCT から文脈タグ  $n$ -gram を抽出する手法について述べる。

意味役割の決定に有効な特徴のある表現は項の末尾に現れるがその長さは分からない。そこで、項の末尾から項の先頭に向かって HTCT を構築する。具体例を図 3 に示す。1 章の例にあげた「ている-ので」や「の-を」(ここでハイ

フンは長単位の区切りを表す) の長単位列に対して、それぞれ末尾から接尾辞木を構築している様子を示している。

各ノードは、文脈タグ列における意味役割 (図では「動作主」「原因」「対象」など) に対する確率が付与されている。各枝の部分文脈タグを示しており、たとえば、長単位「ので」の枝では「助動詞」など他の階層の可能性 (図 3 中の破線で示した部分) の中から表層の「ので」が選択されている様子を示している。また続く長単位「ている」の枝では品詞階層である場合の文脈タグ列「助動詞-ので」から表層の場合の「ている-ので」の選択候補の中から、表層長単位列の「ている-ので」が文脈タグ列として選択されたことを示している。前節で述べたように HTCT の構築は山登り法により根ノードから順に構築する。そのため直前の文脈タグが 1 つ決定されればそれ以降のタグでは、別の候補のタグは検討されない。具体的には、図 3 ではノード「ている」の直前のタグは表層である「ので」に決定されており、品詞「助動詞」や長単位「ので」といった文脈タグ列は計算に考慮されない。単純な構築手法であるが、こうして獲得された可変長の文脈タグ列が意味役割付与に有効であることを 3 章で実験的に示す。

### 2.3 HTCT を利用した特徴量の抽出

構築された HTCT から意味役割付与システムに入力する特徴量を抽出する。HTCT には各意味役割に対してどのような文脈タグ列において高い確率で生成されるかが記録されている一方で、意味役割の決定には、項の末尾表現のほかに述語の動詞や項に含まれる名詞など複数の特徴量が必要であるため、文脈タグ列は 1 特徴量として扱う方が好ましい。つまり、HTCT でも項の長単位から意味役割を予測できるが、最終的な意味役割の判定は多数の特徴量を取り扱うことができる CRF に任せる。一方で、意味役割付与システムの中での HTCT の位置付けとしては、入力項をなす長単位列に対して特徴的な文脈タグ列を取り出す特徴量出力関数となる。HTCT は式 (1) にあるように接尾辞木の各ノードに意味役割に対する確率分布  $P(l|c)$  を持っているため、意味役割が与えられればその意味役割に対して最も高い確率を与える文脈タグ列を得ることができるが、特徴量を取り出す段階ではどの意味役割かは不明である。そこでここでは単純に、入力されたある項  $a$  に対して意味役割を無視して、HTCT の中で最長の文脈タグ列を特徴量として抽出する。今、項  $a$  を末尾側から HTCT に適用する際、上記 4 つの長単位の階層の組を長さ  $n$  の文脈タグ列  $nct$  を生成するとすると、HTCT を利用して得られる文脈タグ列  $\hat{nct}$  は下記のように求める。

$$\hat{nct} = \operatorname{argmax}_{l, nct} \operatorname{Length}(nct) \tag{4}$$

$$\text{s.t. } nct \in \operatorname{feat}(a) \text{ かつ } P(l|nct) > 0$$

ここで  $\operatorname{Length}(nct)$  は文脈タグ列  $nct$  の長さを測る関数で、

$feat(a)$  は入力項  $a$  から長単位に分割して、4つの階層情報を取り出し長さ1から  $n$  までの文脈タグ列  $nct$  を生成する関数である。また  $l$  は式(1)の意味役割を表しており、 $l$  と  $nct$  の組合せで最長の  $\hat{nct}$  を求めるが、最終的には  $l$  は利用しない。取り出した文脈タグ列  $\hat{nct}$  は特徴量としてCRFに登録する。具体的な特徴量登録の事例は3.4.3項で記述する。

### 3. 意味役割付与実験

HTCTを利用して取り出した特徴量が意味役割付与にどの程度貢献するか長単位列に対して意味役割を付与する実験を行うことで確認する。以下では、まず意味役割付与タスク、使用するコーパス、評価枠組み、意味役割付与システムについて定義した後、本研究で行った2種類の実験について記述する。1つはHTCTを利用した特徴量を含む項の末尾表現に関する特徴量間の比較実験であり、もう1つは既存の日本語述語項構造解析システムで用いられている特徴量とHTCTを利用した特徴量との比較実験である。

#### 3.1 意味役割付与タスク

本研究では複数の長単位のまとまり(項候補と呼ぶ)からなる系列に対して意味役割を付与する。これを品詞のタグ付け[16]や固有表現抽出[17]タスクと同様に系列ラベリング問題として扱う。つまり入力文が項候補に分解されていると仮定して項に対しては意味役割を付与し、項でないものにはOを付与する。たとえば1章の事例は図4のように意味役割を付与し、正解評価としては意味役割で評価する\*5。こうした系列による学習を行うことで、「対象」の後に「場所」が出現するといったラベルの連続性の特徴を学習に活かすことができる。

意味役割付与システムの評価では、2分割交差検定で評

項候補	ラベル
詰め将棋の-本-を	対象
秋葉原-で	場所
買っ-た	O
親-を	対象
慕う	O
姿	修飾
左側-から-衝突さ-れ-た-よう-に-なっ-て-い-る-の-で-	原因
骨折し-まし-た	O
開い-ちゃっ-た-の-を	対象
熱湯-に-入-れ-て	O
少し	様態
かき回し-てみる	O

図4 実験コーパスのラベル例

Fig. 4 Examples of labels for evaluation corpus.

\*5 単位は3.2節で利用するコーパスが採用する長単位で示している。そのため「詰め将棋」は1つの単位となる。

価する。評価尺度には適合率、再現率、およびF値を利用する。適合率および再現率は項候補単位でかつ意味役割の数(Oラベルを除く)で評価する。これにより意味役割付与システムが意味役割をどの程度認識できたかが評価できる。

$$\text{適合率} = \frac{\text{システムが正しく意味役割を出力した数}}{\text{システムが出力した意味役割数}} \quad (5)$$

$$\text{再現率} = \frac{\text{システムが正しく意味役割を出力した数}}{\text{正解の意味役割数}} \quad (6)$$

$$F \text{ 値} = \frac{2 \times (\text{適合率}) \times (\text{再現率})}{(\text{適合率}) + (\text{再現率})} \quad (7)$$

#### 3.2 使用するコーパス

本研究で使用する実験用の意味役割付与コーパスは現代日本語書き言葉均衡コーパス(以降、BCCWJと記述)\*6に意味役割を付与した述語項構造シソーラスコーパス[11]\*7(以降、PTCPと記述)に対して項でない長単位連続列にOラベルを付与して作成する\*8。PTCPでは1章で示したように項は述語の係り元である場合だけでなく、被連体修飾や並列化などにより直接係り関係にない長単位連続列にも項として人手で判定し、意味役割が付与されている。実験コーパス作成においては同じ意味役割が付与された長単位連続列を項候補として1つにまとめる。また意味役割が付与されていない長単位についてはOラベルを付与し意味役割と同様に項候補としてまとめる。

1章の事例を実験コーパスに変換した例を図4に示す。

図4では各行が項候補を表しており、ハイフンで区切られている単位が長単位である。太字の部分の述語に対する項に意味役割が付与され、そうでないものにはOラベルが付与されている。

各長単位にはBCCWJで構築された読み、基本形、品詞が付与されており、これらの特徴量として利用できる。PTCPでは文内で述語として現れる動詞(サ変名詞に「する」が付与されている場合を含む\*9)に対して、意味役割を付与している。述語項構造シソーラスに掲載されている動詞に対する意味役割付与事例を増やす目的で構築されているため、各動詞に対して20例文以下で意味役割付与例が登録されている。PTCPでは同じ文に対して複数の動詞に対する意味役割が付与されている場合がある。この場合、実験コーパスでは、1つの動詞に対する意味役割付与事例を1例文として、動詞ごとに異なる意味役割付与事例として登録する。

こうして整理した実験コーパスは3,800例文(付与対象

\*6 [http://pj.ninjal.ac.jp/corpus\\_center/bccwj/](http://pj.ninjal.ac.jp/corpus_center/bccwj/)

\*7 <http://pth.cl.cs.okayama-u.ac.jp/>

\*8 PTCPで付与されている語義概念は利用しないので削除する。

\*9 BCCWJの長単位における品詞体系では動詞に分類されている。

表 1 使用する実験コーパス内で頻出する上位 10 件の意味役割

Table 1 Top 10 semantic roles in the corpus.

意味役割	頻度
対象	1,750
動作主	661
様態	346
修飾	304
経験者	290
時間 (点)	278
着点	221
手段	204
原因	202
副詞相当	199

の動詞は延べ 3,800, 318 種類), 13,133 カ所の項候補のうち, 6,232 カ所の項に 64 種類の意味役割が付与されている。各項に対する上位 10 種類の意味役割の事例を表 1 に示す。また O ラベルの頻度は 6,901 件である。表 1 で最も多いのは「対象」ラベルであり, 次に動作主である。これは PropBank における Arg1, Arg0 に対応しており, 英語の意味役割と同様に頻度が高い。

### 3.3 意味役割付与システム

意味役割ラベルを付与する枠組みとして, 統計的学習モデルとして系列ラベリング問題で高い精度を示す CRF [18] を利用する<sup>\*10,\*11</sup>。CRF は入力長単位列  $\mathbf{x} = x_1, \dots, x_N$  に対して出力ラベル列  $\mathbf{y} = y_1, \dots, y_N$  を下記の確率に基づいて出力する。

$$P(\mathbf{y}|\mathbf{x}) = \frac{1}{Z_{\mathbf{x}}} \exp \left( \sum_{i=1}^N \sum_k \lambda_k f_k(y_{i-1}, y_i, \mathbf{x}) \right) \quad (8)$$

$$\hat{\mathbf{y}} = \operatorname{argmax}_{\mathbf{y} \in \mathcal{Y}(\mathbf{x})} P(\mathbf{y}|\mathbf{x}) \quad (9)$$

ここで,  $f_k(\cdot)$  は特徴量の関数であり, 入力項候補列に対してどの部分を特徴として取り上げるかを指定する関数で,  $\lambda_k$  はその関数に対する重みを示し, 学習により求める。  $Z_{\mathbf{x}}$  は入力の項候補列がとりうる全ラベル列に対して和をとったもので

$$Z_{\mathbf{x}} = \sum_{\mathbf{y} \in \mathcal{Y}(\mathbf{x})} \exp \left( \sum_{i=1}^N \sum_k \lambda_k f_k(y_{i-1}, y_i, \mathbf{x}) \right)$$

式 (8) を正規化する働きを持つ。

CRF を利用した意味役割付与システムの全体象を図 5 に示す。図 5 はベースとなるシステムであり, 特徴量とし

<sup>\*10</sup> CRF の実装として CRF++ (<http://taku910.github.io/crfpp/>) を利用する。

<sup>\*11</sup> CRF の特徴としてラベル決定の際に多数の特徴量を利用できる点がある。linear-chain モデルを適用するので, たとえば「場所」の次に「O」ラベルが来ることや, 「動作主」の次に「対象」が来るといったラベル間の状態遷移の性質を取り込むことができる。しかしながら, 日本語のようなかき混ぜ規則や項の省略がある言語において, linear-chain が有効かどうかは分かっておらず今後の課題としたい。

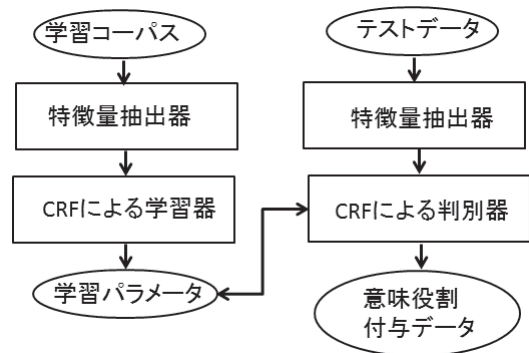


図 5 CRF を利用した意味役割付与システム

Fig. 5 CRF-based semantic role labeling system.

て実験コーパスに付与されている長単位情報 (表層, 読み, 基本形, 品詞), 述語情報, および述語項構造解析 [10] で提案された特徴量を利用する。

### 3.4 項の末尾表現に関する特徴量の比較実験

3.3 節のベースシステムに対して, 提案手法である HTCT による特徴量を導入したモデルのほか, 項候補の末尾表現を異なる方法で特徴化した比較用システムを構築する。1 つは項候補末尾の固定長 n-gram を特徴量として導入したモデル, もう 1 つは機能語辞書「つつじ」を利用したモデル, そして長単位解析器 Comainu を利用したモデルである。以下, 各モデルについて説明する。

#### 3.4.1 基本特徴量

表 2 に基本特徴量を示す。基本特徴量は 3 つの部分からなり, 1 つ目は各項候補内における長単位単独の特徴量 (b1 から b5), 2 つ目は前後の文脈に関する項候補の特徴量 (b6 から b21), 3 つ目は組合せの特徴量 (b22 から b25) である。

まず 1 つ目の長単位単独の特徴量では項候補内の最終長単位の名詞に関する情報, ならびに項候補の末尾にある助詞や述語を取り込んでいる。最終長単位の名詞とは対象とする項候補の意味的な主辞である。項候補内の主辞である名詞と述語との組が意味役割の決定に大きく寄与することから特徴量として利用する。具体例を示すと 3.1 節の例では, 「詰め将棋の本を」の項候補に対する特徴量として b1 から b3 の最終名詞は「本」になり b4 は「を」, b5 は「買う」となる (表 4 を参照)。

2 つ目の文脈の特徴量では予備実験結果<sup>\*12</sup>で良かった 2 つ前から 2 つ後の項候補内の長単位の情報を利用する。

3.1 節の例で「詰め将棋の本を」に対する特徴量としては b14 から b17 の最終名詞は「秋葉原」, b17 の助詞は「で」となり b6 から b13 と b18 から b21 は該当がないためダミータグ (null) を挿入する。

<sup>\*12</sup> 実験としては 1 つ前および 1 つ後, ならびに, 2 つ前および 2 つ後の場合を行い, 比較して文脈を 2 にした方が適していたので採用した。3 つ目の文脈については検討しなかった。

表 2 基本特徴量

Table 2 Basic feature set.

番号	特徴量の説明
b1	項候補の最終名詞の表層
b2	項候補の最終名詞の基本形
b3	項候補の最終名詞の読み
b4	項候補の最終助詞の表層
b5	述語の基本形
b6	2つ前の項候補の最終名詞の表層
b7	2つ前の項候補の最終名詞の基本形
b8	2つ前の項候補の最終名詞の読み
b9	2つ前の項候補の最終助詞の表層
b10	1つ前の項候補の最終名詞の表層
b11	1つ前の項候補の最終名詞の基本形
b12	1つ前の項候補の最終名詞の読み
b13	1つ前の項候補の最終助詞の表層
b14	1つ後の項候補の最終名詞の表層
b15	1つ後の項候補の最終名詞の基本形
b16	1つ後の項候補の最終名詞の読み
b17	1つ後の項候補の最終助詞の表層
b18	2つ後の項候補の最終名詞の表層
b19	2つ後の項候補の最終名詞の基本形
b20	2つ後の項候補の最終名詞の読み
b21	2つ後の項候補の最終助詞の表層
b22	b1 と b5
b23	b2 と b5
b24	b3 と b5
b25	b4 と b5

最後に、3つ目の組合せの特徴量では項候補の最終名詞と述語との組合せを登録する。これはすでに記述したように意味的主辞 (b1 から b4) と述語の組が意味役割の決定に寄与すると考えられるため特徴量として登録して CRF に学習させる。

### 3.4.2 末尾固定長モデル

次に、上記の基本特徴量に各項候補の末尾の特徴量を加えた固定長モデルを提案する。このモデルは項候補の末尾から固定の長さで長単位のすべての情報を特徴量として取り込む。末尾からの長さは2個と3個の場合を仮定する。3個の長単位の場合の特徴量を表 3 に示す。たとえば3.1節の例文では「詰め将棋の本を」に対する特徴量として項候補末尾の3つの長単位は「の」「本」「を」であり、これらの各長単位が持つ4階層の情報を特徴量としてすべて登録する\*13。また、項候補内の長単位の数が設定より少ない場合は、ダミータグを挿入する。たとえば固定長3で、項候補内に2つの長単位しかない場合は、m9 から m12 はダミータグを入れる。

### 3.4.3 HTCT モデル

2.3節で取り出したHTCTの特徴量をCRFに登録する。項に対してHTCTを構築する際、予備実験結果において

\*13 これらの特徴量は意味役割のほか、Oラベルを判定する項候補でも同様に利用する。

表 3 各項候補の末尾長単位の特徴量 (固定長 3)

Table 3 Features of tail long-unit-words in each target (in fixed3 model).

番号	特徴量の説明
m1	項候補の最後の長単位の表層
m2	項候補の最後の長単位の基本形
m3	項候補の最後の長単位の読み
m4	項候補の最後の長単位の品詞
m5	項候補の最後から2番目の長単位の表層
m6	項候補の最後から2番目の長単位の基本形
m7	項候補の最後から2番目の長単位の読み
m8	項候補の最後から2番目の長単位の品詞
m9	項候補の最後から3番目の長単位の表層
m10	項候補の最後から3番目の長単位の基本形
m11	項候補の最後から3番目の長単位の読み
m12	項候補の最後から3番目の長単位の品詞

長さを制限せずしきい値だけで木を作成すると長さ9までの長い文脈タグ列が取り出されたため、HTCTの深さ(つまり文脈タグ列の長さ)を2から5まで制限したモデルを考える。またHTCT構築時のパラメータであるしきい値は0と設定する。

HTCTから取り出した文脈タグ列は意味役割を付与する項候補に対して、1つの特徴量として登録する。具体的に例を示して説明する。まずHTCT構築の際に木の深さを3に制限したモデルHTCT-3を学習データから構築する。次に特徴量を付与する対象である項候補、たとえば「詰め将棋の-本-を」に対してHTCT-3を適用して最長となる文脈タグ列を取り出す。この場合「/助詞-格助詞/名詞-普通名詞-一般/助詞-格助詞」という3つの長単位の文脈タグ列が得られたとする。ここで記号/は文脈タグの区切りを表し、項候補の末尾から木を作成しているため、文脈タグ列も末尾から構築されている。つまり今の場合、「を-本-の」に対する階層化された文脈タグ列が取り出されている。

次に文脈タグ列を対象とする項候補の特徴量として登録する。3.1節の例に対して基本特徴量とHTCT-3の特徴量を付与したCRF学習用のデータを表4に示す\*14。

表4では1行が1つの項候補を表している。文は3つの項候補から成り立っており、上から順に「詰め将棋の-本-を」「秋葉原-で」「買った」に対するb1からb5までの特徴量とHTCT-3の特徴量、意味役割を示している。「秋葉原-で」は2つの長単位で構成されているためこの項候補に対するHTCT-3の出力も長さ2の文脈タグ列になっている。またOラベルである「買った」の項候補に対しても同様にHTCT-3を適用する。ここで、HTCTを適用する学習データは意味役割付与システムのCRFが学習するデータと同じものとする。そしてHTCTを適用して特徴量を取り出すコーパスは学習データならびにテストデータであ

\*14 CRF++のデータフォーマットに合わせている。

表 4 HTCT-3 を利用した特徴量の付与事例  
Table 4 Example of features extracted with HTCT-3.

b1	b2	b3	b4	b5	HTCT-3 による特徴量	ラベル
本	本	ホン	を	買う	/助詞-格助詞/名詞-普通名詞-一般/助詞-格助詞	対象
秋葉原	秋葉原	アキハバラ	で	買う	/助詞-格助詞/名詞-固有名詞-一般	場所
null	null	null	null	買う	/助動詞/動詞-一般	O

る。よって意味役割付与システムがテストデータに対してラベルを付与する際、テストデータの知見はいっさい使用していない。

### 3.4.4 「つつじ」モデル

日本語機能表現辞書「つつじ」は日本語の項の末尾や文末に現れる機能的な語を収集した辞書であり、表記の異なり（たとえば「からすると」や「からしますと」など）や活用を吸収するために階層構造で記述されている。最下層の L9 ではすべての機能語の異表記や活用の異なりが登録されている。そこで L9 に登録されている 16,801 件の機能表現のうち表層の表記が重複したものを除いた 15,887 件を利用して（これを L9 リストと呼ぶ）、各項候補の末尾表現を同定して、CRF の特徴量として登録する。

たとえば、「理想-の-ために-に」という項候補に対して最終長単位部分から文字列で L9 リストとマッチングを行い、最長の機能語を同定し、その機能語を登録する。この例では「ために」がマッチするので「ために」を CRF の特徴量として登録する。ここで項候補末尾に句読点がある場合は句読点を飛ばしてマッチングを行う。また L9 リストに該当する文字列が存在しない場合はダミータグを特徴量として登録する。

### 3.4.5 Comainu モデル

長単位解析器 Comainu [19]<sup>\*15</sup> は入力文を BCCWJ の長単位と同様の要素に区切るシステムである。たとえば文が入力されると「に関して」など機能的な末尾表現などを一つにまとめて出力する。そこで対象とする項候補を文字列として Comainu に与え、分割された要素で末尾にある助詞を末尾表現の特徴量として登録するモデルを構築する。

### 3.4.6 実験結果

各モデルに基づく意味役割付与システムを作成して 2 分割交差検定により評価した結果を示す<sup>\*16</sup>。まず基本特徴量と固定長モデルの結果を表 5 に示す。各モデルは特徴量の違いだけであり、base は b1 から b25 までの基本特徴量みのモデルである。fixed は表 3 の項候補の末尾長単位の特徴量を付与したもので、fixed1 は base の特徴量に m1 から m4 の特徴量を加えたモデル、fixed2 は base と m1 から m7 まで、fixed3 は base と m1 から m12 まで、特徴量を加えたモデルである。fixed2c2 は fixed2 の特徴量に項候補最後の 2 つの長単位の特徴量の組合せ（16 要素）を加え

表 5 基本特徴量 (base) と固定長 (fixed) に基づく意味役割付与精度

Table 5 Results of semantic role labeling using the basic features and fixed-size features of chunks' tail.

モデル名	適合率 (%)	再現率 (%)	F 値
base	49.35	44.00	46.52
fixed1	52.02	46.90	49.33
fixed2	54.52	49.47	51.87
fixed3	54.37	49.74	51.95
fixed2c2	54.65	49.68	52.05
fixed3c2	54.53	49.98	52.16

表 6 HTCT ならびに「つつじ」または Comainu を利用したモデルの意味役割付与精度

Table 6 Results of semantic role labeling using HTCT, Tsutsuji or Comainu models.

モデル名	適合率 (%)	再現率 (%)	F 値
htct2	52.54	47.30	49.78
htct3	51.66	46.42	48.90
htct4	51.14	45.86	48.35
htct5	51.16	45.96	48.42
tsutsuji	50.72	45.49	47.97
comainu	49.89	44.59	47.09

たモデルで、fixed3c2 は fixed3 の特徴量に項候補最後の 2 個の長単位の特徴量の組合せを加えたモデルである。

表 5 に示すように項候補の末尾長単位情報を利用しない base モデルに対して、最終長単位を 1 つだけ加えた fixed1 モデルが F 値において 2 ポイント以上上昇している。このことから項候補の末尾表現が意味役割付与の決定に寄与していることが分かる。そこで、項候補末尾の最終長単位の数を増やした場合 (fixed1 から 3) を比較すると最終の 3 長単位を利用する fixed3 が最も F 値が高くなった。さらに末尾長単位の表層や品詞といった特徴量の組を加えたモデルを作成した場合、fixed2 と fixed3 それぞれに対して F 値が向上した。よって F 値としては末尾長単位の組まで加えた fixed3c2 が最も高い値を示している。

次に HTCT, 「つつじ」, Comainu を利用した各モデルの意味役割付与精度を表 6 に示す。各モデルの特徴量は基本特徴量 (b1 から b25) に加えて、HTCT, 「つつじ」, Comainu を利用して取り出した 1 特徴量を加えたものである。表 6 の htct2 から htct5 は HTCT を学習データから構築する際に木の深さを制限したモデルで htct2 は深さ 2, htct5 は深さ 5 を示している。また tsutsuji は 3.4.4 項

<sup>\*15</sup> <http://comainu.org/>

<sup>\*16</sup> 各評価値の算出にはマイクロ平均を用いた。



表 7 固定長 (fixed) に HTCT (htct) の特徴量を加えたモデルの意味役割付与精度

Table 7 Results of semantic role labeling using fixed-size and HTCT-extracted features of chunks' tail.

モデル名	適合率 (%)	再現率 (%)	F 値
fixed2 + htct2*	54.94	49.94	52.32
fixed3 + htct2*	54.92	50.34	52.53
fixed2c2 + htct2	54.67	49.73	52.08
fixed3c2 + htct2	54.72	50.16	52.34

で述べた手法で「つつじ」から特徴量を取り出したモデル, comainu は 3.4.5 項の手法で Comainu からの特徴量を利用したモデルである.

まず htct での木の深さの違いによる精度の異なりを見ると, 深さ 2 の場合の htct2 の F 値が最も高く適合率, 再現率ともに他のモデルを上回っている. つまり最終長単位から 2 つ分の長単位情報が有効であることが分かる. また htct と tsutsuji, comainu の F 値を比較すると約 2 ポイント以上 htct2 の方が上回っている. tsutsuji と comainu も base に比べて F 値を最大で約 1 ポイント上回っているため, 人手で定義した表現をベースとした特徴量は意味役割の決定に有効ではあるものの, 意味役割の決定に有効な項候補末尾の特徴にとりこぼしがあると考えられる. この点, htct は学習により末尾の特徴量を取り出しているため, 有効な手法であると考えられる.

ただし htct2 の精度は fixed2 や fixed3c2 の精度と比較すると 2 ポイント以上下回っており接尾辞木で取り出した 2 長単位分の特徴量よりも 2 長単位分の表層や品詞といったすべての情報を利用したモデルの方が勝っていることが分かる. しかしながら htct モデルの特徴量は長単位タグ列であるため, fixed2 や fixed3 で利用した項候補末尾の長単位の情報を単に登録している特徴量とは性質が異なると考えられる. htct モデルでは接尾辞木から学習データでの有効な長単位タグ列を選択しているため, fixed モデルと htct モデルの特徴量を組み合わせることが有効であることが予測される. そこで, これらを複合させたモデルを構築する.

表 7 に固定長 (fixed) の特徴量に HTCT (htct) の特徴量を加えたモデルの意味役割付与精度を示す. このとき, 表 5 の元のモデルの出力と比較してマクネマー検定を行った. 表 7 のモデル名に\*印が付与されているモデルが有意水準 5% で精度が向上したモデルである. 結果として fixed2 および fixed3 モデルに対して, htct2 モデルの特徴量を加えた場合, 適合率, 再現率, F 値のすべてが向上し, 元のモデルに対して有意に精度が向上したことが明らかになった. fixed2 および fixed3 モデルはどちらも末尾の長単位尾の特徴量は単独で入っているが末尾長単位の組合せ情報は入っておらず, htct2 で新たに加えた効果が現れていると考えられる. 一方で, fixed2c2 や fixed3c2 に対しては htct2 を加えたことによる有意な差は認められなかつ

表 8 固定長 (fixed3) に HTCT (htct2) の特徴量を加えたときに正解に転じた意味役割の分布 (上位 10 件)

Table 8 Top 10 correctly labeled semantic roles of fixed3 + htct2 which fixed3 wrongly labeled.

意味役割	頻度	fixed3 で誤った意味役割 (頻度)
動作主	14	修飾 (7), O (3), 時間 (点) (2), 経験者 (1), 対象 (1)
対象	14	O (6), 動作主 (5), 着点 (1), 条件 (1), 修飾 (1)
時間 (点)	9	O (3), 対象 (2), 動作主 (2), 着点 (1), 修飾 (1)
程度	6	対象 (2), 時間 (点) (2), 動作主 (1), O (1)
場所	5	O (1), 時間 (1), 手段 (1)
着点	3	O (1), 対象 (1), 手段 (1)
経験者	2	動作主 (1), 修飾 (1)
原因	2	様態 (1), 起点 (1)
順接	2	原因 (2)
時間 (毎)	2	O (1), 時間 (点) (1)

た. これらの 2 つのモデルではすでに htct2 で取り込む長さ 2 の末尾長単位の組合せがすべて登録されているため, 有効に働かなかつたと考えられる.

### 3.4.7 HTCT により取り出された特徴量と結果の比較

HTCT による特徴量の追加でどのような意味役割の付与の変化があったか, ならびにどのような特徴量が追加されることで向上したか, 事例をあげて分析する. 対象としては, 最も F 値が高くなった fixed3 + htct2 モデルについて, 元のモデルである fixed3 からの差分を比較する. fixed3 では誤っていたが fixed3 + htct2 で意味役割が正しく付与された意味役割は 82 件であった. そのうちの上位 10 種類の意味役割の内訳を表 8 に示す.

表 8 において, htct2 の特徴量によって正しく付与されるようになった意味役割として「動作主」, 「対象」, 「時間 (点)」が多かったことが分かる. また表の右側の列では元々 fixed3 ではどのような意味役割として誤って識別されていたか頻度を付与して示している. 「動作主」の場合, fixed3 では「修飾」のほか, 「O」や「時間 (点)」と識別を誤っており, 「対象」の場合も, 「O」や「動作主」が多いため, 「動作主」, 「対象」, 「O」の間の識別誤りが多いことが分かる. これらのラベルが htct2 の特徴量を付与したことによって正しく付与されるようになった. 具体的に「動作主」の場合について正しく付与されるようになった項候補の特徴量を表 9 に示す.

表 9 は「小泉純一郎首相は 9 日.. 向かう途中..」という文の項候補「小泉純一郎首相は」の部分に対する fixed3 および fixed3 + htct2 の特徴量を示している. この項候補の長単位情報を表 10 に示す. また, 表中の特徴量の番号は htct2 を除いて, 表 2 と表 3 に示したものに对应してお

表 9 fixed3 と fixed3 + htct2 の特徴量と出力

Table 9 Features and output labels of fixed3 and fixed3 + htct2.

特徴量	内容	特徴量	内容
b1	小泉純一郎首相	m5	小泉純一郎首相
b2	小泉純一郎首相	m6	小泉純一郎首相
b3	コイズミジュンイチ ロウシュショウ	m7	コイズミジュンイチ ロウシュショウ
b4	は	m8	名詞-固有名詞-人名 -一般
b5	向かう	m9	null
m1	は	m10	null
m2	は	m11	null
m3	ハ	m12	null
m4	助詞-係助詞	htct2	/助詞-係助詞 /名詞-固有名詞-人名 -一般
出力意味役割 (fixed3)		O	
出力意味役割 (fixed3 + htct2)		動作主	

表 10 表 9 の項候補の長単位情報

Table 10 Long-unit-word information of the chunk in Table 9.

表層	読み	基本形	品詞
小泉 純一郎 首相 は	コイズミ ジュンイチロウ シュショウ ハ	小泉 純一郎 首相 は	名詞-固有名詞- 人名-一般 助詞-係助詞

り、特徴量 b6 から b25 は表 9 では省略しているが識別する際に特徴量として利用している。

この項候補の正しい意味役割は「動作主」であるが、出力された意味役割を比較すると fixed3 では「O」と誤って識別されて fixed3 + htct2 では「動作主」と正しく識別されている。特徴量の差を見ると 2 つのモデルの異なりは表 9 中の htct2 のみである。htct2 の特徴量「/助詞-係助詞/名詞-固有名詞-人名-一般」は項候補末尾から順に品詞列属列として「係助詞」の次に「人名」が来る長単位タグ列が特徴量として指定されている。CRF はこの情報をもとにデータから意味役割「動作主」との関係を学習したと考えられる。これは「(人名) + は」や「(人名) + も」といった表現が述語の「~した」という表現に対して「動作主」になりやすいという特性を取り出していると考えられる。実際、全データ内で特徴量「/助詞-係助詞/名詞-固有名詞-人名-一般」が付与されているデータを観測すると「動作主」が 24 件、「経験者」が 9 件（その他「O」、「使役」、相互（人）など）の意味役割にのみ付与されており、品詞レベルの連続であるが「動作主」を特徴付ける特徴量として有効に働いていると考えられる。

一方で、htct2 の特徴量を加えることで悪くなった事例が 35 件あった。これらの意味役割の上位 5 件を表 11 に示す。

表 11 固定長 (fixed3) に HTCT (htct2) の特徴量を加えたときに誤りに転じた意味役割の分布 (上位 5 件)

Table 11 Top 5 wrongly labeled semantic roles of fixed3 + htct2 which fixed3 correctly labeled.

意味役割	頻度	fixed3 + htct2 で誤った意味役割 (頻度)
動作主	5	対象 (2), 経験者 (1), O (1), 修飾 (1)
対象	5	O (3), 動作主 (1), 着点 (1)
原因	3	順接 (1), 対象 (1), 手段 (1)
修飾	3	動作主 (2), 対象 (1)
様態	3	O (1), 副詞 (1), 対象 (1)

表 12 fixed3 と fixed3 + htct2 の特徴量と出力

Table 12 Features and output labels of fixed3 and fixed3 + htct2.

特徴量	内容	特徴量	内容
b1	ダイエット商品	m5	など
b2	ダイエット商品	m6	など
b3	ダイエット ショウヒン	m7	ナド
b4	は	m8	助詞詞-副助詞
b5	入手し	m9	ダイエット商品
m1	は	m10	ダイエット商品
m2	は	m11	ダイエットショウヒン
m3	ハ	m12	名詞-普通名詞-一般
m4	助詞-係助詞	htct2	/助詞-係助詞 /助詞-副助詞
出力意味役割 (fixed3)		対象	
出力意味役割 (fixed3 + htct2)		動作主	

表 13 表 12 の項候補の長単位情報

Table 13 Long-unit-word information of the argument candidate in Table 12.

表層	読み	基本形	品詞
ダイエッ ト商品 など は	ダイエッ トショウヒン ナド ハ	ダイエッ ト商品 など は	名詞-普通名詞-一般 助詞-副助詞 助詞-係助詞

表 11 から「動作主」「対象」「時間 (点)」の誤りが上位を示しており、これは正解に転じた場合 (表 8) と同様で、頻度の多い意味役割で誤りが多く生じていることが分かる。これらの事例のうち、「動作主」について htct2 の特徴量を付与したことで誤るようになった特徴量の例を表 12 に示す。

表 12 は「全くダイエット商品などは..」という文の項候補「ダイエット商品などは」に対する fixed3 と fixed3 + htct2 の特徴量を示している (この項候補の長単位情報を表 13 に示す)。この項候補の正しい意味役割は「対象」であり、元々 fixed3 では正しく識別されていたものが fixed3 + htct2 では誤って「動作主」と識別されている。htct2 の特徴量「/助詞-係助詞/助詞-副助詞」は項候補末尾から「係助詞」の次に「副助詞」が来る長単位タグ列を指定してい

るがこの特徴量が付与された意味役割を全データで調べてみると、「対象」3件、「動作主」3件、「時間」1件と続き3種類の意味役割が該当する（全体では0も含めて8事例）。つまり意味役割の判別には有効に働かない特徴量であったことが考えられる。HTCTから獲得された特徴量のうち、さらに有効な特徴量をどのように絞り込むかについては今後の課題である。

### 3.5 述語項構造解析の特徴量との比較実験

述語項構造解析で提案されている松林ら [10] の特徴量と比較する。まず定性的な比較であるが、本論文で注目している項候補の末尾特性に関する特徴量は先行研究 [10] の (a6) 「項候補の文節において主辞以降に含まれる文字列」が該当する。本論文では末尾の表層だけでなく、読みや品詞、さらに HTCT など異なる特徴量を提案している。こうした違いが意味役割付与でどのように影響するか実験的に示す。松林らの特徴量は本研究で提案している特徴量と重複する部分と異なる部分がある。そこで松林らの特徴量と本研究で提案した基本特徴量（表 2）を合成して新たなベースラインとする。ここで特徴量の取り込みにおいて次の 2 点に注意した。

まず 1 つ目は実数値の特徴量についてである。意味役割付与モデルは CRF++ を利用しており、実数値を特徴量として利用することができない。そこで実数値の特徴量は、しきい値を設定し、しきい値を超えた場合にのみ発火させる。

2 つ目は外部資源についてである。項と述語の共起関係は意味役割の決定に関与することが考えられるため大規模格フレームデータなど有効に働くことが予測される。一方で本論文で提案する項の末尾情報を利用した固定長モデルや HTCT モデルはコーパス中にある情報を有効に利用する手法として提案している。外部資源データは学習データの量に関係なく働く一方で、提案手法の特徴量の効果は学習データの量に左右される。そこで評価の際に 2 分割交差検定だけでなく 10 分割交差検定でも評価を行う。

#### 3.5.1 先行研究からの特徴量の導入

以下では先行研究 [10] から取り入れる特徴量を 5 種類に分割して説明する。1 つ目は述語に関する特徴量 (mp1 から mp5)、2 つ目は項候補に関する特徴量 (mc1 から mc11)、3 つ目は述語と項候補の組合せの特徴量 (mj1 から mj5)、4 つ目は述語と項候補間の統語構造に関する特徴量 (ms1 から ms9)、5 つ目は大規模データによる特徴量 (ml1 から ml8) である。これらの特徴量と本論文が提案する b1 から b25 の特徴量とを融合して利用する。各特徴量では先行研究での特徴量との対応と、本論文でも提案している場合は対応付けがあることを各表で示す。以下、それぞれについて説明する。

まず 1 つ目の述語に関する特徴量を表 14 に示す。表内

表 14 述語に関する特徴量

Table 14 Feature set on predicates.

番号	先行	特徴量の説明
mp1	p2	述語の表層
mp2	p1	述語の基本形
mp3	p2	述語の読み
mp4	p2	述語の活用形 (MeCab)
mp5	p4	述語の末尾表現の種類

表 15 項候補に関する特徴量

Table 15 Feature set on argument candidates.

番号	本論	先行	特徴量の説明
mc1	b1	a2	項候補の最終名詞の表層
mc2	b2	a1	項候補の最終名詞の基本形
mc3	b3	a2	項候補の最終名詞の読み
mc4		a2	項候補の最終名詞の品詞 (第 2 階層まで)
mc5		a2	項候補の最終名詞の品詞 (全階層)
mc6		a2	項候補の最終名詞の固有名詞タグ (MeCab)
mc7		a4	項候補の先頭から最終名詞までの表層
mc8		a5	項候補の最終名詞の右隣の長単位の基本形
mc9		a5	項候補の最終名詞の右隣の長単位の品詞 (第 2 階層まで)
mc10		a5	項候補の最終名詞の右隣の長単位の品詞 (全階層)
mc11		a6	項候補の最終名詞から項候補末尾までの表層

の先行の列は文献 [10] での特徴量の表の行番号に対応している。mp4 の活用形は実験コーパスには活用形は付与されていないため、形態素解析器 MeCab を利用して文を解析し、対応する述語部分の活用形を利用する。自動解析であるため、「見逃し」など名詞と解析されて、活用形が獲得できない場合も含む<sup>\*17</sup>。また mp5 では「述語の接尾辞「れる」「せる」か補助動詞構文「である」であるか」を特徴量とした。

次に、各項候補における特徴量を表 15 に示す。文献 [10] では項候補は文節単位であるが、ここでは実験コーパスに付与されている長単位列からなる項候補に対して同様に付与する。mc6 の固有名詞タグは MeCab で解析した結果を利用する。その他は実験コーパスに付与されている長単位の情報を利用する。項候補の特徴量は実際に項であるかどうかには依存せず、同様に適用する。よって項候補に最終名詞が存在しない場合は、すべてダミータグを特徴量として与える。

項候補に関する特徴量の具体例を示すと表 4 の例では、「詰め将棋の本を」の項候補に対する最終名詞は「本」になり mc1 から mc6 までは「本」に関する情報が特徴量とな

<sup>\*17</sup> その場合は「\*」印とした。

表 16 述語と項候補の特徴量

Table 16 Feature set on predicates and argument candidates.

番号	本論	先行	特徴量の説明
mj1	b23	c1	述語の基本形と項候補の最終名詞の基本形の共起
mj2		c1	述語の基本形と項候補の先頭から最終名詞までの表層の共起
mj3		c2	述語の基本形と項候補の最終名詞から項候補末尾までの表層の共起
mj4		c3	述語の末尾表現の種類と項候補の最終名詞から項候補末尾までの表層の共起
mj5		c5	上記の特徴量と他の項候補の最終名詞から項候補末尾までの表層の辞書順連結の共起

表 17 述語と項候補間の統語構造に関する特徴量

Table 17 Syntactic feature set on predicate and chunk.

番号	先行	特徴量の説明
ms1	s1	項候補が述語より前か
ms2	s1	項候補が述語より後か
ms3	s10	述語の基本形と項候補が述語より前かの共起
ms4	s10	述語の基本形と項候補が述語より後かの共起
ms5	s11	項候補の最終名詞の基本形と項候補が述語より前かの共起
ms6	s11	項候補の最終名詞の基本形と項候補が述語より後かの共起
ms7	s12	述語の基本形と項候補の最終名詞の基本形と項候補が述語より前かの共起
ms8	s12	述語の基本形と項候補の最終名詞の基本形と項候補が述語より後かの共起
ms9	s14	述語の基本形と述語の活用形の共起 (MeCab)

る。mc7 は「詰め将棋の本」、mc8 から mc10 の最終名詞の右隣の長単位は「を」となり、mc9 と mc10 は「助詞-格助詞」となる。mc11 は「を」となる。

次に述語と項候補の特徴量を表 16 に示す。項候補に関する特徴量の具体例を示すと表 4 の例では、「詰め将棋の本を」の項候補と述語「買う」に対する特徴量で、項候補の最終名詞は「本」となる。mj1 は「買う-本」となり、mj2 は「買う-詰め将棋の本」となる。mj3 での項候補末尾までの表層は「詰め将棋の本を」となる。mj4 の「述語の末尾表現の種類」は mp5 の出力である。mj5 は対象とする項候補以外の項候補<sup>\*18</sup>に対する mc11 の出力である。たとえば、1 章の「[詰め将棋の本を][秋葉原で][買った]」の場合に、「買った」の項候補に注目したとき、その他の項候補「詰め将棋の本を」と「秋葉原で」に対する mc11 の出力「で」と「買った」を辞書順に並べる。

次に述語と項候補間の統語構造に関する特徴量を表 17 に示す。表 4 の例では、「詰め将棋の本を」の項候補と述語「買う」に対する特徴量では項候補が前であるため、ms1

<sup>\*18</sup> 名詞を含まない項候補など mc11 の出力がないものは除外する。

表 18 大規模データによる特徴量

Table 18 Feature set extracted from large scale data.

番号	先行	特徴量の説明
ml1	w1	京大格フレーム内に (項候補の最終名詞の基本形, 項候補の最終助詞の表層, 述語の基本形) があるか
ml2	w2	京大格フレーム内に (項候補の先頭から最終名詞までの表層, 項候補の最終助詞の表層, 述語の基本形) があるか
ml3	w3	文脈類似語データベース (クラスタ数 500) における項候補の最終名詞のクラスタ ID (確信度が最も高いもの)
ml4	w4	文脈類似語データベース (クラスタ数 2,000) における項候補の最終名詞のクラスタ ID (確信度が最も高いもの)
ml5	w5	述語の基本形と項候補の最終名詞のクラスタ ID (クラスタ数 500) の共起
ml6	w5	述語の基本形と項候補の最終名詞のクラスタ ID (クラスタ数 2,000) の共起
ml7	w6	述語の基本形と項候補の最終助詞の基本形と項候補の最終名詞のクラスタ ID (クラスタ数 500) の共起
ml8	w6	述語の基本形と項候補の最終助詞の基本形と項候補の最終名詞のクラスタ ID (クラスタ数 2,000) の共起

が t(rue) となる。他は組合せであり、上記で説明している特徴量の組合せである。

最後に、大規模データによる特徴量を表 18 に示す。ml1 と ml2 は京都大学格フレーム (Ver 1.0)<sup>\*19</sup>を利用する。対象となる項と述語のペアがフレーム内に存在する場合に発火する特徴量である。項の中でどの長単位を選ぶかで ml1 と ml2 が異なっており、前者が最終名詞で後者がどれか項候補内でマッチすれば発火する。ml3 から ml8 は文脈類似語データベース (Ver 1.1.1)<sup>\*20</sup>を利用する。このデータベースでは類義語集合 ID が与えられるため、対象とする項の最終名詞と述語を類義語集合で抽象化して項と述語の共起関係を広くとりこむ特徴量である。

### 3.5.2 実験結果

3.5.1 項で整理した述語項構造解析の特徴量と表 2 の基本特徴量を融合した特徴量を新たにベースラインとして、固定長モデル fixed に htct を加えた特徴量の効果を確認する。まず表 19 に 2 分割交差検定の結果を示す。

ここで b はベースラインを表し、f2 や f3 は fixed モデル、h2 および h3 は htct モデル、c2 は組合せのモデルである。表 19 から明らかなように F 値において、f2 から f3c2h3 まですべてのモデルがベースラインを大きく上回っている。またマクネマー検定を b に対して行ったところ、すべての

<sup>\*19</sup> <http://www.gsk.or.jp/catalog/gsk2008-b/>

<sup>\*20</sup> [https://alaginrc.nict.go.jp/resources/nict-resource/li-info/li-outline.html#\#\\$A-1](https://alaginrc.nict.go.jp/resources/nict-resource/li-info/li-outline.html#\#$A-1)

表 19 述語項構造解析の特徴量を取り入れた意味役割付与精度 (2-fold)

Table 19 Results of semantic role labeling using the extended features developed from predicate-argument analysis (2-fold).

モデル名	適合率 (%)	再現率 (%)	F 値
b	56.26	50.43	53.19
f2	58.74	53.10	55.78
f2h2	58.82	53.11	55.82
f2h3	58.68	53.13	55.77
f3	<b>58.73</b>	53.56	<b>56.03</b>
f3h2	58.70	<b>53.58</b>	56.02
f3h3	58.60	53.43	55.90
f2c2	58.26	52.70	55.34
f2c2h2	58.51	52.89	55.56
f2c2h3	58.42	52.90	55.52
f3c2	58.40	53.32	55.75
f3c2h2	58.48	53.40	55.82
f3c2h3	58.56	53.40	55.86

表 20 述語項構造解析の特徴量を取り入れた意味役割付与精度 (10-fold)

Table 20 Results of semantic role labeling using the extended features developed from predicate-argument analysis (10-fold).

モデル名	適合率 (%)	再現率 (%)	F 値
b	60.64	54.99	57.68
f1h2	62.75	57.32	59.91
f1h3	62.67	57.30	59.87
f2	62.54	57.11	59.70
f2h2*(f2)	62.80	57.41	59.99
f2h3	62.71	57.45	59.96
f3	62.64	57.59	60.01
f3h2*(f3)	<b>63.17</b>	<b>58.10</b>	<b>60.53</b>
f3h3*(f3)	62.86	57.89	60.27
f2c2	62.64	57.19	59.79
f2c2h2	62.77	57.33	59.93
f2c2h3	62.69	57.37	59.91
f3c2	62.72	57.67	60.09
f3c2h2	62.72	57.69	60.10
f3c2h3	62.58	57.56	59.96

モデルで 5%有意水準で差異があることが確認できた。これにより提案している項の末尾特徴量は固定長や可変長でも、述語項構造で提案されている特徴量とともに使用した場合においても有効に働くことが実験的に示された。

しかしながら、表 19 では F 値が最も高いのは固定長モデルの f3 であり、htct モデルは精度が高くない。この原因として考えられるのは大規模データによる特徴量に対して学習データの量に影響される htct モデルが有効に働いていないのではないかと仮説を立てた。そこで、10 分割交差検定に切り替えて、同じ実験を行った結果を表 20 に示す。

表 20 では適合率、再現率、F 値の最も高いモデルは f3h2 (fixed3 + htct2) であった。また有意水準 5%でマクネマー検定を行ったところ f2 に対して f2h2, f3 に対して f3h2, および f3h3 で有意に差が生じたことが分かった。これより htct モデルは学習データが十分にある場合固定長で項候補の末尾情報を特徴量とするモデルに対してさらに精度を高めることができると考えられる。

#### 4. 関連研究

意味役割自動付与の研究は先駆的な研究である文献 [9] をはじめとして多くの研究がなされており、文献 [20], [21] にまとまった記述がある。意味役割付与コーパスは主に FrameNet と PropBank の 2 種類が利用されており、それぞれを利用した国際評価型ワークショップが開催されている。たとえば SemEval2007 [22] では FrameNet を利用しており、CoNLL では PropBank を利用しており 2004 年 [23], 2005 年 [24], 2008 年 [25], 2009 年 [26] に開催されている。付与対象も異なっており CoNLL では 2008 年から名詞の意味役割が付与されている\*21。また CoNLL-2008 までは自動付与対象の言語は英語であったが、CoNLL-2009 では多言語 (7 カ国) の意味役割付与コーパスを用意して同じシステムに対して評価を行っている。このように述語との意味的な関係を同定するタスクとして意味役割関連の研究が広く展開されておりすべてを取り上げることは困難であるため、特徴量に注目して英語の意味役割付与、英語以外の意味役割付与、近年提案されている構文解析を利用しないニューラルネットワークを利用した意味役割付与について取り上げる。

英語の意味役割付与では構文解析結果を利用した特徴量を利用してきた [21]。主要な構文解析まわりの特徴量は文献 [9], [21] で提案されており、句の種類 (Phrase Type), 支配カテゴリ (Governing Category), 構文木パス (Parse Tree Path), 述語との位置 (Position), 述語の態 (Voice), 主辞 (Head Word), 下位範疇化 (Subcategorization) があげられる。たとえば (S(NP (PRP He))(VP (VB ate)(NP (DT some)(NN fruits)))) という構文解析結果があり、動詞「ate」に対する意味役割を「He」に付与する場合、句の種類は「NP」であり、支配カテゴリは「S」、構文木パスは「VB ↑ VP ↑ S ↓ NP」、述語との位置は「before」、述語の態は「active」、主辞は「He」、下位範疇化は「(subject, object)」である\*22。

\*21 FrameNet では動詞だけでなく、形容詞、名詞に対しても Frame と FrameElement (意味役割に相当する) が述語と同様に設定されているため、FrameNet でのタスクは動詞以外の品詞の意味役割も考慮することになる。

\*22 このほか、様々な特徴量が利用されている。たとえば文献 [27] は 1 つの SRL システムではなく複数の SRL システムを利用して特徴量にするなど拡張が進んでいる。これも特徴量であるがここでは文の構造としてどのような特徴が意味役割に関係するかを議論するため割愛する。

これらの英語の特微量において、日本語の意味役割付与における「に関して」や「について」など機能的な末尾表現に対応する特微量は主辞である。たとえばこれらの英語表現は前置詞で「about」などで表現できるが、主辞は前置詞句の場合、前置詞を取り出すことを指定しており、文献 [9] ではこうした前置詞は case marker 相当の役割があることを示唆している。しかしながら、主辞で前置詞の表現がカバーできるかという英語でもたとえば「in terms of」のように複合語による機能的な表現が存在しており、これらを主辞としてどこを取り出すのかは明示的に文献 [9], [21] では記述されていなかった。また、動詞の意味役割に限定されている CoNLL-2004 と CoNLL-2005<sup>\*23</sup> で参加したシステムの論文を調べたが、複合語による機能的な表現を明示的に取り扱っている記述は見いだせなかった。複合語は文献 [28] にあるように英語でも困難な問題であることが指摘されており、英語での意味役割では取り扱っていないと考えられる。また、日本語と対応させて考えると、「に関して」は「about」であり、「～によって」は人（つまり意味役割で「動作主」）の場合、英語では「by」で表示され、原因の場合、「because」など異なる語が使われるため、英語では大きな問題になっていないのではないかと推察される。逆にとらえると、日本語の場合は主語や目的語の位置が入れ替わることができるため、項の助詞を含む末尾表現が重要であり、英語における意味役割付与研究では明らかにならなかった特微量であると考えられる。

一方、日本語では、表層格ベースの述語項構造解析において、ゼロ照応を含む項に対して格関係の解析が研究されており、トーナメントモデル [29] や構造化による手法 [30]、または特微量の観点から整理した研究 [10] が成されている。本論文での意味役割付与データは項の単位が文節ではないことや、付与ラベルの種類数が多いなど、述語項構造解析のタスクと異なる部分もある。しかし、述語項構造解析で明らかになった特微量は意味役割付与でも有効であることを今回、実験的に明らかにした。さらに項の末尾の特微量を加えることで意味役割付与精度をさらに向上させることができることを示した。

次に英語以外の言語での意味役割付与システムについて考察する。CoNLL-2009 ではカタルーニャ語、中国語、チェコ語、英語、ドイツ語、日本語、スペイン語について統一した形式での意味役割付与コーパスを構築して、13 チームの参加システムすべてが7カ国語の意味役割付与コーパス上で評価されている [26]。意味役割付与コーパスは係り受けや品詞、主辞情報が一貫した形式で構築されており、それらの中には複数語からなる末尾表現に対する処理の記述は見当たらなかった。参加する意味役割付与システムは言語に特化した処理は行わないので、コーパスに付与され

ている特微量のみを利用している。よって、CoNLL-2009 でも末尾表現の特微量は議論されていない。また中国語の意味役割付与について文献 [21] で英語にはない特微量について記載されているが、中国語特有の構文処理、中国語用の動詞クラスの必要性和単語の切り出しが重要であることが指摘されている。本論文で注目している末尾表現に相当する特微量についての議論は見受けられなかった。

上記の先行研究は構文解析または係り受け解析を重視した手法であるが、近年、こうした解析をせずにニューラルネットワークにより直接単語列から意味役割を付与する手法が提案されてきている [31]。問題点としてとりあげているのは意味役割を付与する前処理の構文解析の誤りでありこれらを避けるために、言語的な知見は Wikipedia の文をニューラル言語モデル [32] で表しておき、テキストの解析は LSTM (Long Short Term Memory) [33] という RNN (Recurrent Neural Network) を拡張したものと融合して分析して中間表現を構築する。この中間表現を CRF の特微量として、意味役割をテキストに対して IOB2 タグで付与する。CoNLL-2005, CoNLL-2012 で評価した結果既存のシステムを超える結果が報告されている [31]。先行研究で議論してきた特微量をまったく使わない手法で興味深い。ただし英語は構造的に日本語とは異なる性質を持つため、日本語でこうした手法が将来適用された際、本研究成果は日本語意味役割付与精度の一事例として価値のある結果であると考えられる。

## 5. おわりに

本論文では日本語の意味役割付与において、項の末尾表現に着目した特微量を利用することで、統計的学習モデルを利用した意味役割付与システムの精度向上を試みた。末尾表現には機能語が多く現れるが、表現の揺れを吸収するために HTCT (階層的タグ文脈木) を利用して特徴化を行い、これに項に関する特微量を加えて CRF による意味役割付与システムを構築する手法を提案した。日本語に対する意味役割付与実験の結果、HTCT の特微量を利用しない場合に対して精度の向上が有意に見られ、HTCT の特微量が有効であることを実験的に示した。また、既存の日本語機能語辞書や、規則ベースで機能語を取り出すシステムを利用した場合より精度が向上した。また述語項構造解析で示された特微量に対して HTCT の特微量を加えた場合でも有効に働くことを実験的に示した。意味役割付与システムの研究は英語を中心に行われてきており、構文情報が有効に働く英語では本論文が提案している特微量は陽に取り上げられていないが、本論文が提案している特微量は日本語の意味役割付与に必要な特微量であると考えられる。

謝辞 数多くの有益なコメントをくださった査読者の皆様に感謝を申し上げる。

<sup>\*23</sup> <http://www.cs.upc.edu/~srlconll/>

## 参考文献

- [1] Shen, D. and Lapata, M.: Using Semantic Roles to Improve Question Answering, *Proc. 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pp.12–21 (2007).
- [2] Liu, D. and Gildea, D.: Semantic Role Features for Machine Translation, *Proc. 23rd International Conference on Computational Linguistics* (2010).
- [3] Liu, X., Han, B., Li, K., Stiller, S.H. and Zhou, M.: SRL-Based Verb Selection for ESL, *Proc. 2010 Conference on Empirical Methods in Natural Language Processing*, pp.1068–1076 (2010).
- [4] チャールズ J. フィルモア (著), 田中春美, 船城道雄 (訳): 格文法の原理, 三省堂 (1975).
- [5] Fillmore, C.J.: *Topics in Lexical Semantics*, pp.76–138, Indiana University Press (1977).
- [6] Palmer, M., Gildea, D. and Kingsbury, P.: The Proposition Bank: An Annotated Corpus of Semantic Roles, *Computational Linguistics*, Vol.31, No.1, pp.71–105 (2005).
- [7] Kipper, K., Dang, H.T. and Palmer, M.: Class-Based Construction of a Verb Lexicon, *Proc. 17th National Conference on Artificial Intelligence and 12th Conference on Innovative Applications of Artificial Intelligence*, pp.691–696 (2000).
- [8] Baker, C.F., Fillmore, C.J. and Lowe, J.B.: The Berkeley FrameNet project, *Proc. 36th Annual Meeting of the Association for Computational Linguistics*, pp.86–90 (1998).
- [9] Gildea, D. and Jurafsky, D.: Automatic Labeling of Semantic Roles, *Computational Linguistics*, Vol.28, No.3, pp.1–45 (2002).
- [10] 松林優一郎, 乾健太郎: 統計的日本語述語項構造解析のための素性設計再考, 言語処理学会第 20 回年次大会, pp.360–363 (2014).
- [11] Takeuchi, K., Ueno, M. and Takeuchi, N.: Annotating Semantic Role Information to Japanese Balanced Corpus, *Proc. MAPLEX 2015* (2015).
- [12] 日本語記述文法研究会: 現代日本語文法 2 第 3 部格と構文第 4 部ヴォイス, くろしお出版 (2009).
- [13] 松吉 俊, 佐藤理史, 宇津呂武仁: 階層構造による日本語機能表現の分類, 言語処理学会第 12 回年次大会 (2006).
- [14] Haruno, M. and Matsumoto, Y.: Mistake-Driven Mixture of Hierarchical Tag Context Trees, *Proc. 35th Annual Meeting of the Association for Computational Linguistics*, pp.230–237 (1997).
- [15] Weinberger, M.J., Rissanen, J.J. and Feder, M.: A Universal Finite Memory Source, *IEEE Trans. Information Theory* (1995).
- [16] Lafferty, J., McCallum, A. and Pereira, F.: Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data, *Proc. 18th International Conf. on Machine Learning*, pp.282–289 (2001).
- [17] Takeuchi, K. and Collier, N.: Bio-medical Entity Extraction Using Support Vector Machines, *Journal of Artificial Intelligence in Medicine*, Vol.33, No.2, pp.125–137 (2005).
- [18] Kudo, T., Yamamoto, K. and Matsumoto, Y.: Applying Conditional Random Fields to Japanese Morphological Analysis, *Proc. 2004 Conference on Empirical Methods in Natural Language Processing*, pp.230–237 (2004).
- [19] 小澤俊介, 内元清貴, 伝 康晴: BCCWJ に基づく中・長単位解析ツール Comainu, 言語処理学会第 20 回年次大会予稿集, pp.582–585 (2014).
- [20] Màrquez, L., Carreras, X., Litkowski, K.C. and Stevenson, S.: Semantic Role Labeling: An Introduction to the Special Issue, *Computational Linguistics*, Vol.34, No.2, pp.145–159 (2008).
- [21] Palmer, M., Gildea, D. and Xue, N.: *Semantic Role Labeling*, Morgan & Claypool Publishers (2010).
- [22] Baker, C., Ellsworth, M. and Erk, K.: SemEval' 07 Task 19: Frame Semantic Structure Extraction, *Proc. 4th International Workshop on Semantic Evaluations (SemEval-2007)*, pp.99–104 (2007).
- [23] Carreras, X. and Màrquez, L.: Introduction to the CoNLL-2004 Shared Task: Semantic Role Labeling, *Proc. CoNLL-2004 Shared Task* (2004).
- [24] Carreras, X. and Màrquez, L.: Introduction to the CoNLL-2005 Shared Task: Semantic Role Labeling, *Proc. CoNLL-2005 Shared Task* (2005).
- [25] Surdeanu, M., Johansson, R., Meyers, A., Màrquez, L. and Nivre, J.: The CoNLL-2008 Shared Task on Joint Parsing of Syntactic and Semantic Dependencies, *Proc. 12th Conference on Computational Natural Language Learning*, pp.159–177 (2008).
- [26] Hajič, J., Ciaramita, M., Johansson, R., Kawahara, D., Martì, M.A., Màrquez, L., Meyers, A., Nivre, J., Padò, S., Štěpánek, J., Straňák, P., Surdeanu, M., Xue, N. and Zhang, Y.: The CoNLL-2009 Shared Task: Syntactic and Semantic Dependencies in Multiple Languages, *Proc. 13th Conference on Computational Natural Language Learning (CoNLL): Shared Task*, pp.1–18 (2009).
- [27] Surdeanu, M., Màrquez, L., Carreras, X. and Comas, P.R.: Combination Strategies for Semantic Role Labeling, *Journal of Artificial Intelligence Research*, Vol.29, No.47, pp.105–151 (2007).
- [28] Sag, I.A., Baldwin, T., Bond, F., Copestake, A.A. and Flickinger, D.: Multiword Expressions: A Pain in the Neck for NLP, *Proc. 3rd International Conference on Computational Linguistics and Intelligent Text Processing*, pp.1–15 (2002).
- [29] 林部祐太, 小町 守, 松本裕治: 述語と項の位置関係ごとの候補比較による日本語述語項構造解析, 自然言語処理, Vol.21, No.1, pp.3–25 (2014).
- [30] 平 博順, 永田昌明: 構造学習を用いた述語項構造解析, 言語処理学会第 14 回年次大会, pp.556–559 (2008).
- [31] Zhou, J. and Xu, W.: End-to-end Learning of Semantic Role Labeling Using Recurrent Neural Networks, *Proc. 50th Annual Meeting of the Association for Computational Linguistics*, pp.1127–1137 (2015).
- [32] Bengio, Y., Ducharme, R., Vincent, P. and Janvin, C.: A Neural Probabilistic Language Model, *Journal of Machine Learning Research*, Vol.3, pp.1137–1155 (2003).
- [33] Hochreiter, S. and Schmidhuber, J.J.: Long Short-Term Memory, *Neural Computation*, Vol.9, No.8, pp.1735–1780 (1997).



石原 靖弘

2011 年岡山大学工学部情報工学科卒業。2013 年岡山大学大学院自然科学研究科博士前期課程修了。現在、岡山大学大学院自然科学研究科博士後期課程在学中。専門は自然言語処理。



竹内 孔一 (正会員)

1991年姫路工業大学工学部電子工学科卒業。1998年奈良先端科学技術大学院大学博士後期課程修了。博士(工学)。同年学術情報センター助手。2000年国立情報学研究所助手。2003年岡山大学工学部情報工学科講師。2005年

同大学大学院講師，現在に至る。主に，専門用語研究，語彙概念構造の研究，項構造の言語資源構築と解析に従事。言語処理学会，人工知能学会，電子情報通信学会，医療情報学会，ACM 各会員。