

組合せ特徴量を用いた述語項構造解析

平 博順 藤田 早苗 永田 昌明

NTT コミュニケーション科学基礎研究所

{taira.hirotooshi, fujita.sanae, nagata.masaaki}@lab.ntt.co.jp

1 はじめに

意味解析の一つである述語項構造解析は、述語と項（日本語では述語と格関係にある名詞句）との関係を同定する技術である。表1に述語項構造のコーパスの一つであるNAISTテキストコーパスの定義に基づく述語項構造の例を示す。

表1: 述語項構造

元テキスト: 花子が太郎から与えられた本			
述語	ガ格(主格) 項	ヲ格(対格) 項	ニ格(与格) 項
与える	太郎	本	花子

NAIST テキストコーパスでは、テキスト中の述語の基本形に対して意味的にガ格(主格)、ヲ格(対格、対象格)、ニ格(与格、目的格)となる名詞句を正解の項としてタグ付けがされている¹。「花子が太郎から与えられた本」というテキストがあった場合、述語「与えられた」の基本形「与える」に対してガ格、ヲ格、ニ格となる項はそれぞれ「太郎」、「本」、「花子」となる。このように、テキストから述語と名詞句の間の意味的な関係を自動抽出する技術は、さまざまな言語処理タスクにおいて有用であると考えられ、情報抽出[2]、質問応答[7]、自動要約[5]といったタスクにおいて、その利用が試みられてきている。

NAIST テキストコーパスを対象とした述語項構造解析では、これまで項の組合せ制約をSVMで求め決定リストで項を特定する方法[8]、最大エントロピー法を用いた方法[4]、変換ベース学習を用いた方法[9]、Markov Logicを用いた方法[10]などが提案されている。

述語項構造解析、特に項の同定に関しては、単語、品詞、係り受け状態などの基本特徴量だけでなく、それらを組合せた組合せ特徴量が有効であると考えられる。しかしながら、積極的に組合せ特徴量を用いる方

法は、文献[8]以外ではほとんど提案されていない。また、自動的に基本特徴量を組み合わせる方法としては、カーネル法が考えられるが、非線形カーネルの計算コストは高いため、実際にはほとんど線形カーネルしか用いられていない。

そこで本稿では、線形SVMを用いて述語項構造解析を行う手法をベースラインとして、そこに有効な組合せ特徴量を追加して、精度向上が図れる可能性があることを示す。有効な組合せ特徴量を効率的に見つけ出すことは一般に困難であるが、本稿では、変換ベース学習の枠組みを使用した組み合わせ特徴量選択を行うことで、効率的に有効な組合せ特徴量を抽出することを試みた。

本稿の構成は次の通りである。2節で本稿でのベースラインとなるSVMを用いた述語項構造解析手法、3節で提案手法である変換ベース学習を用いた分類に有効な組合せ特徴量抽出方法について述べる。4節で実験結果について述べ、最後に5節で結論を述べる。

2 SVMを用いた述語項構造解析

我々は、SVMを用いた述語項構造解析をベースライン手法として訓練フェーズは、次のような手順で行った。1) 訓練用テキストに対して形態素解析、係り受け解析を行う²。2) ある単語が述語であるか否かの分類学習をSVMで行う。3) 各正解述語に対して、同一文中の各単語が項となるか否かについての分類学習をSVMで行う。4) 各項がどの項タイプ(本稿ではガ格、ヲ格、ニ格の3タイプ)であるかについての分類学習を各項タイプ毎にSVMで行う。

一方、テストフェーズは、次のような手順で行った。

1) テスト用テキストに対して形態素解析、係り受け解析を行う 2) 訓練フェーズ2) で学習されたモデルを用いて各単語が述語か否かを特定する。3) 2) で予測さ

¹事態性名詞に対する項にもタグ付けされているが本稿では述語に対する項のみ扱う。

²本稿では、形態素解析器としてChasen、係り受け解析器としてCabochaを用いた。

れた述語に対し、同一文中の単語が項となるか否かの予測を訓練フェーズ 3) で学習されたモデルを用いて予測する。4) 3) で項と予測された単語について、訓練フェーズ 4) で学習されたモデルを用いて各単語がどの項タイプとなるかを予測する。ただしこのとき、予測値の高い順に項のスロットを埋める。

ここで、訓練フェーズの 2) と 4) は必ずしも分ける必要が無く、項か否かの分類を行わず、項タイプ分類のみを行うことも考えられるが、予備実験を行ったところ、2) の手順を行わないと精度が下がることが確認されたため、この手順で行っている。これは、Pradhan らが、英語の Semantic Role Labeling (SRL) タスクで、SVM を用いて項同定を行っている研究 [6] でも見られたように、項か否か分類と、項タイプの分類では識別に有効な特徴量がかなり異なっているためであると考えられる。

3 変換ベース学習を用いた有効な組合せ特徴量抽出

変換ベース学習 (Transformation based learning; TBL) は、これまで、形態素解析、構文解析、チャンキングタスクなどで高精度な解析に成功している、Brill によって考案された学習方法である [1]。変換ベース学習では、まず、ベースライン解析器と呼ばれるベースラインとなる解析器で訓練サンプルを解析し、訓練サンプルに正例か負例かの予測ラベルを付与する³。ここで、あらかじめ訓練コーパス中で付与された正解ラベルと予測ラベルとを比較し、予測が間違っているサンプルのみを抽出する。抽出されたサンプルから、特徴量を取り出し、その特徴量をサンプルが持ったときに、現在の予測ラベルから正解ラベルへと変換を行うというルールを生成する。次に、生成されたルールを全サンプルに対して適用した際のラベリング誤りの削減数を計算する。その中で最も誤り削減数が大きかったルールを選択し、それを全サンプルに適用して、予測ラベルを更新する。以降、予測が間違っているサンプルから特徴量を取り出し、ルールを生成し、誤り削減数が最大のルールを生成し、全サンプルへ適用する、という手続きを繰り返す。最大のエラー削減数が 0 になるまで、これを繰り返し終了する。テストフェーズでは、訓練フェーズで生成されたルールを決定リストとして、順に適用して予測ラベルを求める。

³TBL では一般には複数カテゴリーの分類が可能であるが、ここでは簡単のため正負例の 2 カテゴリー場合について説明する。

この手法の特徴としては、得られる分類ルールの可読性が高いこと、生成されるルール数が比較的少ないにも関わらず、高い分類性能が得られることが挙げられる。本稿では、ルール数が少ないが、高い分類性能が得られる点に注目し、有効な組合せ特徴量の抽出への適用を試みる。

以下に、本稿で用いた変換ベース学習を用いた有効な組合せ特徴量の抽出手順について述べる。まず、全訓練サンプルについて、予測ラベルの初期状態を「負例」にしておく。それから、上記で述べたように変換ベース学習を行う。変換ルールの条件には、あらかじめ全訓練サンプルから抽出しておいた基本特徴量の一つを用い、その基本特徴量を持つ場合のみに変換をかけるルールとする。実際に選択されたルールを用いてサンプルのラベルを変換する際には、各サンプルに持たせたりリスト構造の変換ヒストリに変換条件に使われた基本特徴量を Push 動作する。こうして変換ベース学習を終えた後、生成された各サンプルの持つ変換ヒストリに入った基本特徴量の異なりリストは、正例か負例かを決定付ける有効な特徴量組合せを近似的に反映していると考えられるため、これを有効な組合せ特徴量とし、SVM 学習の一つの特徴量 (feature) として使用する。

4 実験結果

4.1 実験データ

実験データとしては、NAIST テキストコーパス 1.4 β [3] 全体を訓練用データ (記事: 1 月 1 日 ~ 1 月 11 日、社説: 1 月 ~ 8 月)、開発用データ (記事: 1 月 12 日、1 月 13 日、社説: 3 月)、テスト用データ (記事: 1 月 14 日 ~ 1 月 17 日、社説: 10 月 ~ 12 月) の 3 つの集合に分割し、訓練用データを用いて学習を行い、テスト用データを用いて精度評価を行った。なお、この訓練データ集合にはのべ 4 万 9527 個の述語が含まれていた。表 2 に我々が用いたデータセットの内訳を示す。今回は、これらの中で係り受け解析器の判定で項と述語の間に係り受け関係のあったものを用いて評価を行った。

次に、SVM による分類学習に用いた基本特徴量を表 3 に示す。述語の分類では、述語の形態素解析で得られた特徴量に加えて、単語の意味カテゴリや用言分類なども用いている。項/非項の分類では、述語分類で用いた特徴量を述語候補、項候補それぞれから抽出し用いる他、項候補と述語候補間での係り受けのパスを

表 3: 分類に用いた基本特徴量

分類対象	値の例
述語/非述語	単語基本形、品詞 (レベル 1,2,3,4)、活用型、活用形、 固有表現、機能語、 「～ダ」、受身表現、使役表現、 単語意味カテゴリ、用言分類、 最終一文字、閉じ括弧 品詞レベル 2 + 機能語一文字目 単語表層 + 機能語一文字目
項/非項	述語候補に対する上記特徴量 項候補に対する上記特徴量 述語候補の活用型 + 項候補の単語基本形 述語候補の活用型 + 項候補の品詞レベル 1 述語候補の活用型 + 項候補の品詞レベル 2 述語候補の活用型 + 項候補の意味カテゴリ 項候補から述語候補への係り受けを辿ったときの path 文字列 述語候補から項候補への係り受けを辿ったときの path 文字列 述語候補と項候補が同一文節中にあるかどうか 受身形のときの項候補の単語 + 述語候補までの係り受け path 受身形のときの項候補の品詞 + 述語候補までの係り受け path 受身形のときの項候補の意味カテゴリ + 述語候補までの係り受け path 受身形のときの述語候補の単語 + 項候補までの係り受け path 受身形のときの述語候補の品詞 + 項候補までの係り受け path 受身形のときの述語候補の意味カテゴリ + 項候補までの係り受け path

辿った場合の文字列などを加えている。項タイプ分類に関しては、項/非項分類で用いた基本特徴量と同じ基本特徴量を用いた。なお意味カテゴリに関しては、日本語語彙大系 [11] に基づき、単語毎に付与した。

表 2: 作成したコーパスの内訳

	train	dev	test
記事数	1,753	480	696
文数	24,263	4,834	9,287
文節数	227,191	46,167	87,412
述語数	67,187	13,589	25,584
項数	95,276	20,668	38,370
ガ格項	55,951	11,965	22,033
ヲ格項	26,963	5,553	10,372
ニ格項	12,362	3,150	5,965

4.2 比較手法

提案手法である変換ベース学習を用いた有効な特徴量抽出の効果を評価するために、以下の 5 つの特徴量セットで精度を比較した。正例のみから抽出した基本特徴量のみを使用 (basic_pos)、正負例すべての訓練サンプルから抽出した基本特徴量を使用 (basic)、変換ベース学習で得られた組合せ特徴量のみを使用 (tbl)、正例からの基本特徴量 + 変換ベース学習からの組合せ特徴量 (basic_pos + tbl)、正負例からの基本特徴量 + 変換ベース学習からの組合せ特徴量 (basic + tbl)。いずれの方法についても、述語分類、項/非項分類、項タイプ分類の 3 段階について線形カーネルを用いた SVM (実装は SVMlight) を用いた。

4.3 実験結果

最終的な項同定に対する F1 値の結果を表 4 に示す。なお、訓練データに元のデータの先頭 1000 記事を用いたものである。この実験結果から、変換ベース学習から抽出された組合せ特徴量単独では、基本特徴量を全て用いる場合よりも若干精度が下がるが、組合せ特徴量と基本特徴量のどちらも使用すると精度が向上することが分かる。

表 4: 実験結果 (係り受け関係がある場合)

手法	F1(%)
basic_pos	84.51
basic	86.58
tbl	85.01
basic_pos + tbl	85.25
basic + tbl (提案手法)	86.91

5 おわりに

本稿では日本語の述語項構造解析に関して、変換ベース学習を用いて有効な組合せ特徴量の抽出を行い、識別学習の特徴量に加えることで、解析精度が向上する可能性があることを示した。今後は、より詳細な手法に対する評価を進めるとともに、ゼロ代名詞解析などにも適用して、本手法の適用可能性について検討していきたいと考えている。

参考文献

- [1] Eric Brill. Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging. *Computational Linguistics*, Vol. 21, No. 4, pp. 543–565, 1995.
- [2] Lynette Hirschman, Patricia Robinson, Lisa Ferro, Nancy Chinchor, Erica Brown, Ralph Grishman, and Beth Sundheim. Hub-4 Event’99 general guidelines. 1999.
- [3] Ryu Iida, Mamoru Komachi, Kentaro Inui, and Yuji Matsumoto. Annotating a Japanese text corpus with predicate-argument and coreference relations. In *Proc. of ACL 2007 Workshop on Linguistic Annotation*, pp. 132–139, 2007.
- [4] Kenji Imamura, Kuniko Saito, and Tomoko Izumi. Discriminative approach to predicate-argument structure analysis with zero-anaphora resolution. In *Proc. of the ACL-IJCNLP 2009 Conference Short Papers*, pp. 85–88, 2009.
- [5] Gabor Melli, Yang Wang, Yudong Liu, Mehdi M. Kashani, Zhongmin Shi, Baohua Gu, Anoop Sarkar, and Fred Popowich. Description of SQUASH, the SFU question answering summary handler for the DUC-2005 summarization task. In *Proc. of DUC 2005*, 2005.
- [6] Sameer Pradhan, Wayne Ward, Kadri Hacioglu, James Martin, and Dan Jurafsky. Shallow semantic parsing using support vector machines. In *Proc. of the Human Language Technology Conference/North American Chapter of the Association of Computational Linguistics HLT/NAACL 2004*, 2004.
- [7] Dan Shen and Mirella Lapata. Using semantic roles to improve question answering. In *Proc. of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP/CoNLL)*, pp. 12–21, 2007.
- [8] Hirotoishi Taira, Sanae Fujita, and Masaaki Nagata. A japanese predicate argument structure analysis using decision lists. In *Proc. of the Conference on Empirical Methods in Natural Language Processing EMNLP 2008*, pp. 523–532, 2008.
- [9] Hirotoishi Taira, Sanae Fujita, and Masaaki Nagata. Predicate argument structure analysis using transformation based learning. In *Proc. of the Conference on ACL 2010*, 2010.
- [10] 吉川克正, 浅原正幸, 松本裕治. Markov logic による日本語述語項構造解析. 情報処理学会研究報告 (自然言語処理研究会) 2010-NL-199 No.5, 2010.
- [11] 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦. 日本語語彙大系. 岩波書店, 1997.