

# 構文解析

工学博士 奥村 学 監修

博士(工学) 鶴岡 慶雅 共著

博士(情報理工学) 宮尾 祐介

コロナ社

## 刊行のことば

人間の思考，コミュニケーションにおいて不可欠なものである言語を計算機上で扱う自然言語処理という研究分野は，すでに半世紀の歴史を経るに至り，技術的にはかなり成熟するとともに，分野が細かく細分化され，また，処理対象となるものも，新聞以外に論文，特許，WWW上のテキストなど多岐にわたり，さらに，応用システムもさまざまなものが生まれつつある。そして，自然言語処理は，現在では，WWWの普及とともに，ネットワーク社会の基盤を支える重要な情報技術の一つとなっているといえる。

これまでの自然言語処理に関する専門書は，自然言語処理全般を広く浅く扱う教科書（入門書）以外には，情報検索，テキスト要約などを扱う，わずかの書籍が出版されているだけという状況であった。この現状を鑑みるに，読者は，「実際にいま役に立つ本」，「いまの話題に即した本」を求めているのではないかと推測される。そこで，これまでの自然言語処理に関する専門書では扱われておらず，なおかつ，「いま重要と考えられ，今後もその重要さが変わらない」と考えられるテーマを扱った書籍の出版を企画することになった。

このような背景の下生まれた「自然言語処理」シリーズの構成を以下に示す。

1. 自然言語処理で利用される，統計的手法，機械学習手法などを広く扱う  
近年の自然言語処理は，コーパスに基づき，統計的手法あるいは機械学習手法を用いて，規則なり知識を自動獲得し，それを用いた処理を行うという手法を採用することが一般的になってきている。現状多くの研究者は，他の先端的な研究者の論文などを参考に，それらの統計的手法，機械学習手法に関する知識を得ており，体系的な知識を得る手がかりに欠けている。そこで，そのような統計的，機械学習手法に関する体系的知識を与える専門書が必要と感じている。
2. 情報検索，テキスト要約などと並ぶ，自然言語処理の応用を扱う  
自然言語処理分野も歴史を重ね，技術もある程度成熟し，実際に使えるシステム，技術として世の中に少しずつ流通するようになってきている

## ii 刊 行 の こ と ば

ものも出てきている。そのようなシステム、技術として、検索エンジン、要約システムなどがあり、それらに関する書籍も出版されるようになってきている。これらと同様に、近年実用化され、また、注目を集めている技術として、情報抽出、対話システムなどがあり、これらの技術に関する書籍の必要性を感じている。

### 3. 処理対象が新しい自然言語処理を扱う

自然言語処理の対象とするテキストは、近年多様化し始めており、その中でも、注目を集めているコンテンツに、特許（知的財産）、WWW上のテキストが挙げられる。これらを対象とした自然言語処理は、その処理結果により有用な情報が得られる可能性が高いことから、研究者が加速度的に増加し始めている。しかし、これらのテキストを対象とした自然言語処理は、これまでの自然言語処理と異なる点が多く、これまでの書籍で扱われていない内容が多い。

### 4. 自然言語処理の要素技術を扱う

形態素解析、構文解析、意味解析、談話解析など、自然言語処理の要素技術については、教科書の中で取り上げられることは多いが、技術が成熟しつつあるにもかかわらず、これまで技術の現状を詳細に説明する専門書が書かれることは少なかった。これらの技術を学びたいと思う研究者は、実際の論文を頼らざるを得なかったというのが現状ではないかと考える。

本シリーズの構成を述べてきたが、この構成は現在の仮のものであることを最後に付記しておきたい。今後これらの候補も含め、新たな書籍が本シリーズに加わり、本シリーズがさらに充実したものとなることを祈っている。

本シリーズは、その分野の第一人者の方々に各書籍の執筆をご快諾願えたことで、成功への最初の一步を踏み出せたのではないかと考えている。シリーズの書籍が、読者がその分野での研究を始める上で役に立ち、また、実際のシステム開発の上で参考になるとしたら、この企画を始めたものとして望外の幸せである。最後に、このような画期的な企画にご賛同下さり、実現に向けた労をとって下さったコロナ社の各氏に感謝したい。

2013年12月

監修者 奥村 学

## ま え が き

構文解析というのは不思議な研究分野である。日本語や英語といった自然言語に内在する（と考えられている）真の構文構造を見たことのある人間は誰もいない。それでも構文解析を研究する人達は、言語現象に対する観察や内省によって、それらが存在するものとし、その構造をうまく表現するためのモデルやそれを解析するための手法を日々精力的に研究している。

構文解析が多くの研究者を引き付ける理由はなんだろうか。一つの理由は、言語がもつ構造やそれを表現するための枠組みそのものに対する単純な好奇心のようにも思われる。人間の知的な活動を支える大きな柱の一つである言語が、本質的にどのような構造をもっていて、それを計算機でどのように表現できるのか、ということは多くの研究者にとって心躍る話題であろう。もう一つの大きな理由は、構文構造を解析することが、自然言語処理の多くのアプリケーションにおいて有用だと信じられているからである。実際、大量のテキストからさまざまな情報を抽出するためのシステムや、計算機が自然言語で質問に答えるシステムでは、構文解析が入力テキストに対する前処理として利用されることが多い。したがって、構文解析の技術が向上すれば、広範な自然言語処理アプリケーションの性能が向上することになる。

かくも魅力的な構文解析というテーマであるが、ひとたびその分野の研究論文を読もうとすると、初学者は多くの壁に突き当たることになる。構文解析に関する最近の論文の多くは、離散最適化や機械学習、言語理論といったそれぞれ比較的独立に発展してきた複数の分野の知識を前提としていて、ある一つの教科書を読めば必要な前提知識が一通りそろそろようにはなっていない。構文解析の研究者は、それらの多岐にわたる知識を論文を読みながら徐々に身に付けていくのだが、それはとても時間のかかるプロセスである。

本書では、理工系の学部3年生程度の知識をもった読者が、構文解析に関する最近の研究論文を読むために必要な知識を、ワンストップで得られるようなものを目指した。構文解析に必要な機械学習技術などについても、特段の前提知識を必要とすることなく理解できるよう、できるだけ丁寧に記述することを心がけた。著者の力量不足によりその目標がすべて理想的に実現できたわけではないが、構文解析技術に関するまとまった情報を類書では得られない形で提供できているのではないかと考えている。

本書を書くきっかけを与えていただき、また草稿の段階から丁寧に査読をしていただいた奥村学先生とコロナ社の方々に感謝する。能地宏さんには原稿の細かいところまでチェックしていただいた。最後に、週末の原稿書きを長きにわたって暖かく見守ってくれた家族に大いに感謝したい。

2017年6月

著 者

# 目 次

## 1. はじめに

1.1 構文解析とは .....	1
1.2 なぜ構文解析を学ぶのか .....	4
1.3 本書の構成 .....	5

## 2. 品詞解析と機械学習

2.1 品詞解析の役割 .....	6
2.2 機械学習による品詞タグ付け .....	8
2.3 隠れマルコフモデル .....	11
2.3.1 確率モデル .....	11
2.3.2 Viterbi アルゴリズム .....	16
2.4 最大エントロピーモデル .....	19
2.5 最大エントロピーマルコフモデル (MEMM) .....	26
2.6 条件付き確率場 (CRF) .....	28
2.7 構造化パーセプトロン .....	35
2.8 ビーム探索 .....	39
2.9 生コーパスを利用した学習 .....	43
2.9.1 自己学習 .....	43
2.9.2 クラスタリング .....	44
2.10 浅い構文解析 .....	45

2.11	日本語形態素解析	46
2.12	英語以外の言語	47
2.13	近年の動向	48

### 3. 句構造解析

3.1	句構造	50
3.2	文脈自由文法	52
3.2.1	CKY アルゴリズム	54
3.2.2	Earley 法	57
3.3	確率文脈自由文法	61
3.3.1	PCFG を用いた CKY アルゴリズム	64
3.3.2	内側外側 (Inside-Outside) アルゴリズム	65
3.4	確率文脈自由文法の拡張	67
3.4.1	語彙化	67
3.4.2	非終端記号の詳細化	72
3.5	識別モデルによる再順位付け	73
3.6	遷移型モデルによる句構造解析	75
3.7	評価法	81
3.8	近年の動向	83

### 4. 依存構造解析

4.1	依存構造とは	84
4.2	CKY 法	88
4.3	Eisner 法	90
4.4	MST 法	94

4.5 遷移型依存構造解析 .....	98
4.6 日本語構文解析 .....	102
4.7 評価法 .....	104
4.8 近年の動向 .....	104
4.8.1 高次の関係を用いた解析 .....	104
4.8.2 疑似非交差解析 .....	106

## 5. 文法理論, 深い構文解析

5.1 組合せ範疇文法 (CCG) .....	109
5.2 主辞駆動句構造文法 (HPSG) .....	113
5.3 深い構文解析 .....	118
5.4 Supertagging .....	120
5.5 文法獲得 .....	121
5.6 意味役割付与, 格解析, 述語項構造解析 .....	124
5.7 Semantic Parsing .....	127
5.8 評価法 .....	131
5.9 近年の動向 .....	132

## 6. 構文解析の応用

6.1 情報抽出 .....	134
6.2 単語クラスタリング .....	135
6.3 評判解析 .....	136
6.4 機械翻訳 .....	137
6.5 テキスト間含意関係認識 .....	139



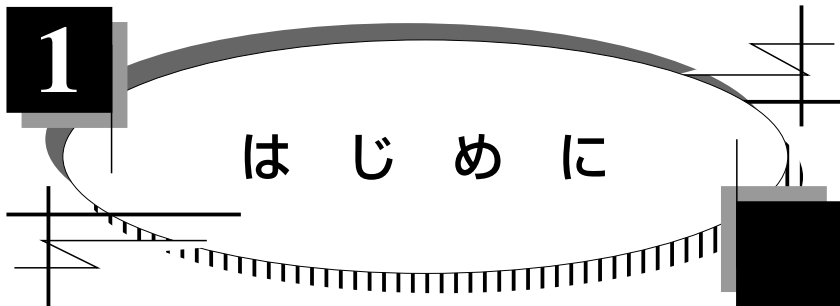
## 7. 構文解析ツール

7.1	Stanford Parser	144
7.2	Berkeley Parser	145
7.3	BLLIP Parser	145
7.4	ZPar	146
7.5	SyntaxNet	146
7.6	Mate	147
7.7	TurboParser	147
7.8	MaltParser	147
7.9	MSTParser	148
7.10	C&C Parser	148
7.11	EasyCCG	148
7.12	OpenCCG	149
7.13	Enju	149
7.14	KNP	149
7.15	CaboCha	150
7.16	KyTea	151
7.17	Jigg	151

## 8. ツリーバンク

8.1	Penn Treebank	152
8.2	Prague Dependency Treebank	154
8.3	Universal Dependencies	154
8.4	京都大学テキストコーパス	155

8.5	NAIST テキストコーパス .....	156
8.6	BCCWJ .....	157
8.7	PropBank .....	157
引用・参考文献 .....		158
索	引 .....	170



## 1.1 構文解析とは

コンピュータで文章の意味を理解するという事は、自然言語処理の究極の目標の一つであるが、文の意味を正しくとらえるためには、文中の単語の並びを表面的に処理するだけでは不十分である。例えば、さまざまなものやことの性質や、それらの間に成り立つ関係に関する情報を英文テキストから抜き出すために、「名詞+動詞+名詞」という単語の並びに着目すると

Everything in Italy tastes good.

という文から “Italy tastes good” という誤った情報を抜き出してしまふ。

構文解析を行うと、文を単なる単語の並びとしてとらえるのではなく、その背後にある文法的な構造を明らかにすることができる。例えば、3章で解説する句構造に基づいて上の例文を構文解析すると、図 1.1 のような階層構造が得られる。ここでは、“Everything in Italy” は名詞句 (noun phrase, NP) というまとまりであることが示されていて、それが動詞句 (verb phrase, VP) とともに文を構成していることがわかる。このことから、名詞句の一部である “Italy” を “tastes” の主語として抽出することが誤りであることがすぐにわかる。このように、構文解析は自然言語文の内部構造を明らかにすることで、文が表す意味を正しくとらえる手助けをする。ちなみに、この句構造の中で、各単語の直上にあるノード、例えば、“Everything” に対する Noun や “in” に対

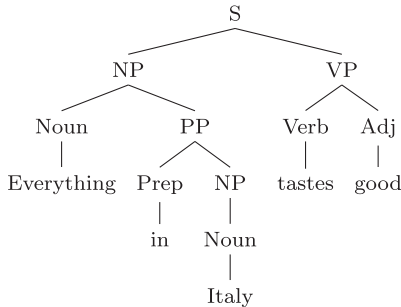


図 1.1 句構造の例

する Prep は、各単語の品詞を表している。このように、句構造解析では、品詞解析の処理をその中にも含むこともあるが、品詞解析を構文解析とは独立して行う場合、その処理を品詞タグ付けと呼ぶ。上記の文に対して品詞タグ付けを行った結果は

Everything/Noun in/Prep Italy/Noun tastes/Verb good/Adj.

ということになる。

句構造と並んで構文解析のための構造としてよく用いられているのが、単語間の係り受け関係を表す依存構造である。上記の文の依存構造は、図 1.2 のように表すことができる。動詞“tastes”の係り元を見ると、“Everything”と“good”となっており、このことから“Italy”が“tastes”の主語ではないことがわかる。依存構造解析については 4 章で詳しく解説する。

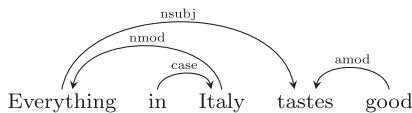


図 1.2 依存構造の例

構文解析に関する研究の歴史は、自然言語処理の研究と同じくらい長いですが、実用に耐える解析精度と速度をもつ構文解析器が誰でも自由に使えるようになったのは比較的最近のことである。詳しくは 7 章で紹介するが、現在では、英語や日本語にかぎらず、多くの言語で高精度な構文解析のためのツールが利用可能になっている。また、6 章で紹介するように、機械翻訳をはじめ、情報抽出、

質問応答など、さまざまな自然言語処理アプリケーションに構文解析技術が使われている。

近年では、句構造や係り受け構造といった構文構造に加えて、述語項構造や意味役割といった、「深い」構文構造を解析することのできる構文解析器が使えるようになってきている。例えば、以下の三つの文

The kid ate the cake.

The cake was eaten by the kid.

I miss the cake eaten by the kid.

はそれぞれ異なった構文構造をもっていて、動詞“eat”（の変化形）と他の名詞“kid”や“cake”の間の係り受け関係も異なっている。しかしながら、どの文においても、“eat”の意味的な主語は“kid”であり、意味的な目的語は“cake”である。詳しくは5章で紹介するが、意味役割付与や述語項構造解析、格解析といった深い構文解析を行うことで、句構造や依存構造には直接には表れない意味的關係を得ることができる。これからの自然言語処理研究が、文章の意味の理解に向かうことは疑いないが、このような深い構文解析技術は、文の意味を正しく解析する上で重要な基盤技術になるであろう。

近年の実用的な構文解析技術の発展の最も大きな原動力となったのは、2章で紹介する機械学習技術である。（教師あり）機械学習とは、与えられた学習モデルと大量の入出力のペア（学習データ）に対して、それらの入出力関係を実現するようなモデルパラメータを計算するための技術的フレームワークである。入力を文、出力をその文の構文構造とすれば、その対応関係の学習はまさに構文解析の問題に他ならない。機械学習に基づく構文解析技術が発展した背景には、ツリーバンクと呼ばれる、各文に構文構造が付与されたテキストコーパスの存在が不可欠であった。これらの各種コーパスについては、8章で述べる。

## 1.2 なぜ構文解析を学ぶのか

構文解析の研究が大きく進展したことにより、現在は高精度かつ高速な構文解析ソフトウェアを簡単に使うことができる。これらのソフトウェアを使うだけなら、構文解析技術の中で用いられる基礎技術についてわざわざ学ぶ必要はないかもしれない。実際、構文解析の中ではさまざまな技術が複合的に使われており、一見、たいへん複雑に見える。避けて通れるなら避けたほうがよいと思われるかもしれない。

それでも、自然言語処理にたずさわる研究者やエンジニアは、構文解析の基礎について一通り学ぶべきであると考えます。これは、構文解析技術は、機械学習、動的計画法、探索法、文法理論といった、自然言語処理のさまざまな場面（構文解析にかぎらない）において必要となる多くの技術の上に成り立っているからである。つまり、構文解析の基礎技術について学ぶことで、自然言語処理で駆使されるさまざまな先人の知恵を学習することができるのである。例えば、機械翻訳は異なる言語の間の変換と考えることができ、一見構文解析とは関係ないように見えるが、ここでも構文解析のアイデアがふんだんに使われている。

極論すれば、自然言語処理に関するプログラムは、構文解析技術を理解していればほとんど理解できる（あるいは自分で書ける）といってよい。逆に、構文解析を知らなければ、自然言語処理のプログラムの半分以上は理解できないかもしれない。つまり、テキストデータを扱うためのさまざまなアイデアは、構文解析の一連の基礎技術の中に凝縮されているのである。構文解析について学ぶことで、これらの技術の有機的なつながりを理解することができ、それは構文解析以外の多くの問題においても有用なのである。

## 1.3 本書の構成

2章では、品詞タグ付けのための手法と、機械学習の基礎的な事項について説明する。品詞タグ付けのための手法として、隠れマルコフモデル、最大エントロピーモデル、構造化パーセプトロンについて述べる。今日の構文解析システムのほとんどは機械学習技術を利用しているが、本章では、品詞タグ付けの問題を例にとり、機械学習の基本的な考え方やアルゴリズムを説明する。

3章では、句構造解析について述べる。句構造解析のための文法として、文脈自由文法とそれを用いた構文解析アルゴリズムについて説明する。構文構造の曖昧性を解消する手法として、確率文脈自由文法に関して述べる。また、確率文脈自由文法を拡張し、実用的なレベルまで精度を向上させるいくつかの手法について述べる。

4章では、依存構造解析について述べる。依存構造を解析するためのアルゴリズムとして、CKY法、Eisner法、最大全域木に基づく手法、遷移型解析アルゴリズムを紹介する。

5章では、文法理論、深い構文解析について述べる。本章では、句構造や依存構造では明示されない文法構造を計算することのできる文法として、組合せ範疇文法 (CCG)、主辞駆動句構造文法 (HPSG) を紹介し、それらを用いた構文解析アルゴリズムおよび文法獲得の手法について説明する。また、句構造解析や依存構造解析の結果に意味的關係を付け加える手法として、意味役割付与、格解析、述語項構造解析について述べる。

6章では、構文解析の応用例を紹介する。本章では、数ある構文解析の応用例から、特に、情報抽出、単語クラスタリング、評判解析、機械翻訳、テキスト間含意関係認識について述べる。

7章では、構文解析ツールとして自然言語処理技術の研究や開発によく利用されるソフトウェアを紹介する。

8章では、構文解析の研究において、構文解析モデルやアルゴリズムの学習用・評価用データとして利用されているツリーバンクを紹介する。

# 索引

<p><b>【あ】</b></p> <p>浅い構文解析 45</p> <p><b>【い】</b></p> <p>依存構造 2, 84</p> <p>依存構造解析 84</p> <p>意味表現 110</p> <p>意味役割付与 125</p> <p><b>【う】</b></p> <p>後ろ向きアルゴリズム 34</p> <p>内側確率 65</p> <p>内側外側アルゴリズム 65</p> <p><b>【え】</b></p> <p>エンコーダ・デコーダモデル 83</p> <p><b>【お】</b></p> <p>重みパラメータ 22</p> <p><b>【か】</b></p> <p>開発データ 10</p> <p>過学習 10</p> <p>係り受け構造 84</p> <p>格解析 102, 127</p> <p>学習データ 10</p> <p>確率的勾配降下法 26</p> <p>確率文脈自由文法 61</p> <p>隠れマルコフモデル 11</p> <p>カテゴリ 109</p> <p>完全部分木 92</p>	<p><b>【き】</b></p> <p>機械学習 3, 9</p> <p>木構造制約 85</p> <p>疑似曖昧性 90</p> <p>疑似非交差解析 106</p> <p>機能タグ 121</p> <p>基本カテゴリ 110</p> <p>キュー 78</p> <p>京都大学テキストコーパス155</p> <p>京都テキスト解析ツール</p> <p>キット 151</p> <p>局所性 29</p> <p><b>【く】</b></p> <p>句 45</p> <p>句構造 1, 50</p> <p>組合せ規則 110</p> <p>組合せ範疇文法 109</p> <p>クラスタリング 44</p> <p><b>【け】</b></p> <p>形態素解析 6</p> <p>現代日本語書き言葉均衡コーパス 157</p> <p><b>【こ】</b></p> <p>子 85</p> <p>語彙化 69</p> <p>語彙項目 113</p> <p>語彙ルール 53</p> <p>交差検定法 24</p> <p>高次モデル 105</p> <p>構造化パーセプトロン 35</p> <p>構造子測 6</p>	<p>勾配法 24</p> <p>構文解析 1</p> <p>構文ルール 52</p> <p><b>【さ】</b></p> <p>最急降下法 26</p> <p>再現率 81</p> <p>最大エントロピーマルコフモデル 26</p> <p>最大エントロピーモデル 20</p> <p>最大全域木法 94</p> <p><b>【し】</b></p> <p>自己学習 43</p> <p>弱教師あり学習 129</p> <p>修飾句 122</p> <p>主辞 69, 85</p> <p>主辞駆動句構造文法 113</p> <p>主辞後置法 138</p> <p>述語項構造 3</p> <p>述語項構造解析 125, 127</p> <p>準ニュートン法 26</p> <p>条件付き確率場 28</p> <p>条件付き尤度 21</p> <p>情報抽出 134</p> <p><b>【す】</b></p> <p>スタック 78</p> <p><b>【せ】</b></p> <p>生成ルール 52</p> <p>正則化 23</p> <p><b>【そ】</b></p> <p>素性 20</p>
---	---	--



素性関数	22	データスパースネス	13	付加部	72
素性構造	113			不完全部分木	92
外側確率	66	<b>【と】</b>		文脈自由文法	52
ソフトマックス回帰	20	動詞句	1	<b>【へ】</b>	
<b>【た】</b>		導出	109	平均化パーセプトロン	38
対数線形モデル	20	動的計画法	17	<b>【ま】</b>	
多クラスロジスティック		特徴量	20	前向きアルゴリズム	32
回帰	20	<b>【に】</b>		前向き・後ろ向きアルゴリ	
多項ロジスティック回帰	20	ニューラル機械翻訳	139	ズム	35
単一化	115	<b>【は】</b>		マルコフ化	71
単語クラスタリング	135	バックポインタ	18	<b>【め】</b>	
談話表示構造	148	半教師あり学習	43	名詞句	1
<b>【ち】</b>		<b>【ひ】</b>		<b>【ら】</b>	
チャート	58	非交差制約	86	ラベル付き係り受け精度	104
チャンキング	45	必須項	122	ラベルなし係り受け精度	104
チョムスキー標準形	54	ビーム探索	40	ラムダ式	110
<b>【つ】</b>		評価データ	10	<b>【り】</b>	
ツリーバンク	152	評判解析	136	リランキング	74, 106
<b>【て】</b>		品詞	6	<b>【る】</b>	
適合率	81	品詞タガー	6	ルールベースの手法	8
テキスト間含意関係認識	140	品詞タグ付け	2, 6		
デコーディング	16	<b>【ふ】</b>			
テストデータ	10	深い構文解析	108, 118		

---

		beam search	40	chart	58
<b>【A】</b>		Berkeley Parser	73, 145	Chomsky normal form	54
adjunct	72	BFGS 法	26	chunking	45
argument	122	Binary ルール	57	Chu-Liu/Edmonds 法	94
arg max	38	BLLIP Parser	145	CKY アルゴリズム	54
atomic category	110	BOS	15	CKY 法	54, 88
averaged perceptron	38	Brown クラスタリングアル		CNF	54
		ゴリズム	44	Cocke-Kasami-Younger	
<b>【B】</b>		<b>【C】</b>		algorithm	54
backward algorithm	34	CaboCha	150	Collins Parser	67
Baum-Welch アルゴリズム		category	109	combinatory categorial	
	44	CCG	109	grammar	109
Baum-Welch algorithm	44	CFG	52	combinatory rule	110
BCCWJ	157			complete span	92

conditional likelihood	21	forward-backward algo-		locality	29
conditional random field	28	rithm	35	log-linear model	20
CoNLL Shared Task	88	function tag	121	LP	81
context-free grammar	52	f-score	81	LR	81
CRF	28				
Cross Bracketing	82	<b>[G]</b>		<b>[M]</b>	
cross-validation	24	gradient descent	24	MaltParser	147
C&C Parser	148			Markovization	71
		<b>[H]</b>		Mate	147
<b>[D]</b>		head	85	maximum entropy model	20
data sparseness	13	headword	69	maximum spanning tree	94
decoding	16	head finalization	138	maxmium entropy Markov	
deep parsing	108	Head Percolation Table	70	model	26
dependency parsing	84	head-driven phrase struc-		Max Violation	42
dependency structure	84	ture grammar	113	MEMM	26
dependent	85	hidden Markov model	11	modifier	122
derivation	109	higher-order model	105	MST 法	94
discourse representation		HMM	11	MSTParser	148
structure	148	HPSG	113	multinomial logistic regres-	
DRS	148			sion	20
		<b>[I]</b>		multi-class logistic regres-	
<b>[E]</b>		incomplete span	92	sion	20
Earley 法	57	information extraction	134		
Early Update	42	inside probability	65	<b>[N]</b>	
EasyCCG	148	inside-outside algorithm	65	NAIST テキストコーパス	156
Edge-Factored モデル	105			neural machine translation	
Eisner 法	91	<b>[J]</b>			139
EM アルゴリズム	44	Jigg	151	noun phrase	1
EM algorithm	44			NP	1
encoder-decoder model	83	<b>[K]</b>			
Enju	149	KNP	149	<b>[O]</b>	
EOS	15	KyTea	151	OpenCCG	149
Evalb	82			outside probability	66
expectation-maximization		<b>[L]</b>			
algorithm	44	labeled attachment score		<b>[P]</b>	
			104	Parent Annotation	72
<b>[F]</b>		labeled precision	81	part-of-speech tagger	6
F スコア	81	labeled recall	81	part-of-speech tagging	6
feature	20	LAS	104	PCFG	61
feature function	22	lexicalization	69	Penn Treebank	7, 152
feature structure	113	Lexicalized PCFG	68	phrase	45
first-order model	105	lexical entry	113	phrase structure	50
forward algorithm	32	lexical rule	53	POS tagger	6

POS tagging	6	semantic representation	110		
Prague Dependency Tree-		semi-supervised learning	43	<b>[U]</b>	
bank	154	sentiment analysis	136	UAS	104
precision	81	SGD	26	Unary ルール	54
probabilistic context free		shallow parsing	45	Universal Dependencies	88, 144, 154
grammar	61	Shift 操作	79	unlabeled attachment	
projectivity	86	Shift-Reduce 法	78, 98	score	104
PropBank	157	softmax regression	20	unlabeled precision	81
Proposition Bank	157	spurious ambiguity	90	unlabeled recall	81
pseudo-projective parsing		stack	78	UP	81
	106	Stanford CoreNLP	144	UR	81
		Stanford Dependencies	144	<b>[V]</b>	
<b>[Q]</b>		Stanford Parser	144	verb phrase	1
quasi-Newton method	26	steepest descent	26	Viterbi アルゴリズム	16
queue	78	stochastic gradient descent	26	VP	1
			26	<b>[W]</b>	
<b>[R]</b>		structured perceptron	35	weakly supervised learning	129
Ratnaparkhi Parser	76	structured prediction	6	word clustering	135
recall	81	supertag	120	<b>[Z]</b>	
recognizing textual entail-		Supertagging	120	ZPar	146
ment	140	SyntaxNet	146	~~~~~	
Reduce 操作	79	syntax rule	52	<b>[数字]</b>	
regularization	23			1 次モデル	105
reranking	74	<b>[T]</b>			
RTE	140	treebank	152		
<b>[S]</b>		tree substitution grammar	83		
self-training	43		83		
Semantic Parsing	128	TSG	83		
		TurboParser	147		

奥村 学 (おくむら まなぶ)  
1984年 東京工業大学工学部情報工学科卒業  
1989年 東京工業大学大学院博士課程修了  
(情報工学専攻), 工学博士  
1989年 東京工業大学助手  
1992年 北陸先端科学技術大学院大学助教授  
2000年 東京工業大学助教授  
2007年 東京工業大学准教授  
2009年 東京工業大学教授  
現在に至る

鶴岡 慶雅 (つるおか よしまさ)  
1997年 東京大学工学部電気工学科卒業  
1999年 東京大学大学院修士課程修了 (電子情  
報工学専攻)  
2002年 博士 (工学) (東京大学)  
2009年 北陸先端科学技術大学院大学准教授  
2011年 東京大学准教授  
現在に至る

宮尾 祐介 (みやお ゆうすけ)  
1998年 東京大学理学部情報科学科卒業  
2001年 東京大学大学院博士課程中途退学 (情  
報科学専攻)  
2001年 東京大学助手  
2006年 博士 (情報理工学) (東京大学)  
2007年 東京大学助教  
2010年 国立情報学研究所准教授  
現在に至る

## 構 文 解 析

Syntactic Analysis

© Yoshimasa Tsuruoka, Yusuke Miyao 2017

2017年8月10日 初版第1刷発行

検印省略

監 修 者 奥 村 学  
著 者 鶴 岡 慶 雅  
宮 尾 祐 介  
発 行 者 株 式 会 社 コ ロ ナ 社  
代 表 者 牛 来 真 也  
印 刷 所 三 美 印 刷 株 式 会 社  
製 本 所 有 限 会 社 愛 千 製 本 所

112-0011 東京都文京区千石 4-46-10

発 行 所 株 式 会 社 コ ロ ナ 社  
CORONA PUBLISHING CO., LTD.  
Tokyo Japan

振替 00140-8-14844 ・ 電話 (03) 3941-3131 (代)

ホームページ <http://www.coronasha.co.jp>

ISBN 978-4-339-02759-4 C3355 Printed in Japan

(金)



**JCOPY** < 出版者著作権管理機構 委託出版物 >

本書の無断複製は著作権法上での例外を除き禁じられています。複製される場合は、そのつど事前に、出版者著作権管理機構 (電話 03-3513-6969, FAX 03-3513-6979, e-mail: info@jcopy.or.jp) の許諾を得てください。

本書のコピー、スキャン、デジタル化等の無断複製・転載は著作権法上での例外を除き禁じられています。購入者以外の第三者による本書の電子データ化及び電子書籍化は、いかなる場合も認めていません。落丁・乱丁はお取替えいたします。