

## 第26回計算言語学会年次大会報告

小暮 潔†                  野垣内 出†                  前田 広幸‡

†ATR自動翻訳電話研究所

‡シャープ(株)コンピュータシステム研究所

本年6月7日から10日の間、ニューヨーク州立大学バッファロー分校で開催された第26回国計算言語学会年次大会ACL-88について、その概要を報告する。この大会では、35件の一般講演、1件の招待講演、および、4件のチュートリアルが行われた。

The Report on  
the 26th Annual Meeting of the Association for Computational Linguistics

Kiyoshi KOGURE†, Izuru NOGAI†, and Hiroyuki MAEDA‡

†ATR Interpreting Telephony Research Laboratories

Twin 21 Bldg. MID Tower, 2-1-61 Shiromi, Higashi-ku, Osaka 540

‡Sharp Corporation, Computer Systems Laboratories

2613-1 Ichinomoto-cho, Tenri-shi, Nara 632

The 26th Annual Meeting of the Association for Computational Linguistics was held at the State University of New York at Buffalo, June 7 - 10 1988. During the conference, 35 talks were presented, as well as an invited talk, after 4 tutorials.

## 1. はじめに

今年6月7日から6月10日の間、ニューヨーク州バッファロー市の郊外にあるニューヨーク州立大学バッファロー分校(State University of New York at Buffalo)において、第26回国計算言語学会年次総会(26th Annual Meeting of the Association for Computational Linguistics)が開催された。

会議では、6月7日に4件のチュートリアル(表1)が行われ、その翌日から3日間、35件の一般講演(表2)と1件の招待講演が行われた。招待講演としては、Dedre Gentner が“Analogy and the Interpretation of Metaphor”という題で講演を行い、心理学における類推、メタファに関する研究が紹介された。

## 2. 講演の概要

以下では、講演の中から、報告者らが興味を持ったものを中心に報告する。

### Sentence fragments regular structures

標準的な言語処理システムの最小の拡張で、telegraphic fragmentsを誤りとしてではなく、正規の形で扱う方法について述べた。Unisysの自然言語処理システムPUNDITにインプリメントされている。英語におけるFragmentsの統語的、意味的、語用論的な解析について考察している。

### Multi-level plurals and distributivity

複数名詞に関する考察を行った。統語的には、同じ構造を持つ文が、複数名詞の存在と、ある種の動詞により、異なる意味構造を持つことは、よく知られている。例えば、“the boys walk”と“the boys gather”は、“ $\forall x \in \text{BOYS: WALK}[X]$ ”、“GATHER[BOYS]”のようになる。そして、“walk”と“gather”のような動詞をそれぞれ、distributive verb、collective verbという。これらの区別は、複数名詞の意味的取扱いに重要である。この発表は、著者のSchaのこれまでの研究を

曖昧性の問題、複数の持つ意味の曖昧性について拡張した。複数の持つ意味のあいまい性の例として次の例文をあげている。“The children ate the pizzas.”について、“What did the children eat?”には適切な答え“The pizzas”が与えられるが、“Who ate the pizzas?”には、食べていない子供がいる場合、“The children”では、適切ではない。このシステムでは、意味解釈を2段階とし、あいまい性を含むEnglish-Oriented Formal Language (EFL)を得て、これを曖昧性のない言語World Model Language (WML)に変換して、処理を行う。もし、曖昧性があれば、複数のWMLを得る。動詞区分に従い、EFLからWMLへの変換ルールを記述しこれらの問題に対処している。

### The interpretation of relational nouns

この報告では、計算機上での関係名詞の意味の扱いについて考察している。関係名詞とは“sister”、“commander”などの名詞で、John's brotherのbrotherのように明示的、もしくは非明示的な引数を持つ。また、Haveなどの動詞の要素となるという特徴がある。データベース、エキスパートシステムの自然言語処理インタフェイスのような応用などでは、これらの意味解釈は重要である。この報告では、特に機能名詞、“speed”、“distance”、“rating”などについて詳しく考察しており、関係名詞の解析は、意味的な解析が必要であるとし、Multilevel semantic frameworkを用いて記述している。すなわち、前発表と同様、EFLとWMLを用いて処理している。例えば、“Frederick has a speed of 15 knots.”は、意味的に高いレベルの言語EFLに解釈される。

$\exists c \{c \in \text{SPEEDS} \mid \text{OF}(c, \text{amount}(15, \text{KNOTS}))\}$ :  
HAVE (FREDERICK, c) [EFL]

これに、EFL-to-WML変換規則を適用して

F-SPEED(FREDERICK) = amount(15, KNOTS)  
[WML]を得る。このWMLは曖昧性のない言語である。前発表とともに名詞に関する興味深い研究である。

### The integrated framework for semantic and pragmatic interpretation

意味的、語用論的解釈のための枠組として、条件付き解釈(Conditional Interpretation)と名付けられた考えを導入している。条件付き解釈は、宣言的な解釈ルールの集合に従い、構成的に構築されるものである。この方法は、広い範囲の語用論的な現象を扱うことができる。また、文脈依存な解釈とそうでないものを分離するなどの特徴が

表1 チュートリアルの講師と題名

●Peter Sells <i>Contemporary Syntactic Theories</i>
●Martha Palmer, Lynette Hirschman, & Deborah Dahl <i>Text Processing Systems</i>
●David McDonald <i>Natural Language Generation</i>
●Masaru Tomita <i>Efficient Parsing Algorithms</i>

表2 一般講演

- Roy J. Byrd and Everyn Tzoukermann (IBM Thomas J. Watson Research Center)  
*Adapting an English Morphological Analyzer for French*
- Maricia C. Linebarger, Deborah A. Dahl, Lynette Hirschman, and Rebecca J. Passonneau (Paoli Research Center, Unisys Corporation)  
*Sentence Fragments Regular structures*
- Remko Scha and David Stallard (BBN Laboratories Inc.)  
*Multi-Level Plural and Distributivity*
- Jos de Bruin and Remko Scha (BBN Laboratories Inc.)  
*The Interpretation of Relational Nouns*
- Douglas B. Moran (Artificial Intelligence Center, SRI International)  
*Quantifier Scoping in the SRI Core Language Engine*
- Bruce W. Ballard (AT&T Bell Laboratories)  
*A General Computational Treatment of Comparatives for Natural Language Question Answering*
- Manny Rayner (SICS) and Amelie Banks (UPMAIL)  
*Parsing and Interpreting Comparatives*
- Robin F. Karlin (Department of Computer and Information Science, University of Pennsylvania)  
*Defining the Semantics of Verbal Modifiers in the Domain of Cooking Tasks*
- Mary Dalrymple (Artificial Intelligence Center, SRI International)  
*The Interpretation of Tense and Aspect in English*
- Martha E. Pollack (Artificial Intelligence Center, SRI International) and Fernando C. N. Pereira (Cambridge Computer Science Research Center, SRI International)  
*An Integrated Framework for Semantic and Pragmatic Interpretation*
- Eugene Charniak and Robert Goldman (Department of Computer Science, Brown University)  
*A Logic for Semantic Interpretation*
- Jerry R. Hobbs, Mark Stickel, Paul Martin, and Douglas Edwards (Artificial Intelligence Center, SRI International)  
*Interpretation as Abduction*
- Robin Haigh, Geoffrey Sampson, and Eric Atwell (Center for Computer Analysis of Language and Speech, University of Leeds)  
*Project APRIL: A Progress Report*
- Bonnie Lynn Webber (Department of Computer and Information Science, University of Pennsylvania)  
*Discourse deixis: Reference to Discourse Segments*
- Steve Whittaker and Phil Stenson (Hewlett-Packard Laboratories)  
*Clues and Control in Expert-Client Dialogues*
- Janyce M. Weibe and William J. Rapaport (Department of Computer Science, State University of New York at Buffalo)  
*A Computational Theory of Perspective and Reference in Narrative*
- Hiroyuki Maeda, Susumu Kato, Kiyoshi Kogure, and Hitoshi Iida (ATR Interpreting Telephony Research Laboratories)  
*Parsing Japanese Honorifics in Unification-Based Grammar*
- John Bateman (USC/Information Science Institute)  
*Aspects of Clause Politeness in Japanese: An Extended Inquiry Semantics Treatment*
- Seiji Miike, Koichi Hasebe, Harold Somers, and Shin-ya Amano (Research and Development Center, Toshiba Corporation)  
*Experiences with an On-Line Translating Dialogue System*
- Eduard H. Hovy (USC/Information Science Institute)  
*Planning Coherent Multisentential Text*
- Douglas Appelt and Kurt Konolige (Artificial Intelligence Center and Center for Study of Language and Information, SRI International)  
*A Practical Nonmonotonic Theory for Reasoning about Speech Acts*
- Eduard H. Hovy (USC/Information Science Institute)  
*Two Types of Planning in Language Generation*
- James Raymond Davis (The Media Laboratory, MIT) and Julia Hirschberg (AT&T Bell Laboratories)  
*Assigning Intonational Features in Synthesized Spoken Directions*
- Megumi Kameyama (MCC)  
*Atomization in Grammar Sharing*
- Gerard Salton (Department of Computer Science, Cornell University)  
*Syntactic Approaches to Automatic Book Indexing*
- Andrew David Beale (Unit for Computer Research on the English Language, University of Lancaster)  
*Lexicon and Grammar in Probabilistic Tagging of Written English*
- Thomas Ahlswede and Martha Evens (Computer Science Department, Illinois Institute of Technology)
- Naoki Abe (Department of Computer and Information Science, University of Pennsylvania)  
*Polynomial Learnability and Locality of Formal Grammars*
- Robert T. Kasper (USC/Information Science Institute)  
*Conditional Descriptions in Functional Unification Grammar*
- Mark Johnson (Brain and Cognitive Sciences, MIT)  
*Deductive Parsing with Multiple Levels of Representation*
- Masaru Tomita (Center for Machine Translation and Computer Science Department, Carnegie-Mellon University)  
*Graph-Structured Stack and Natural Language Parsing*
- Yves Schabes and Aravind K. Joshi (Department of Computer and Information Science, University of Pennsylvania)  
*An Earley-Type Parsing Algorithm for Tree Adjoining Grammars*
- Remo Pareschi (Department of Computer and Information Science, University of Pennsylvania)  
*A Definite Clause Version of Categorical Grammar*
- David J. Weir and Aravind K. Joshi (Department of Computer and Information Science, University of Pennsylvania)  
*Combinatory Categorical Grammars: Generative Power and Relationship to Linear Context-Free Rewriting Systems*
- Andreas Eisele and Jochen Dörre (Institute of Maschinelle Sprachverarbeitung, Universität Stuttgart)  
*Unification of Disjunctive Descriptions*

ある。ここでの解釈とは、センスと仮定の集合からなる。句のセンスは、意味的な情報のみからなるのではなく、句の言語学的文脈からの情報を含む。仮定は、このとき最適な解釈が句に与えられるような制約である。この仮定は、解釈に新しいパラメタを導入し、文脈中の対応物へこのパラメタを束縛する束縛仮定、解釈に新しいパラメタを導入せずに、既存のパラメタで束縛されているものに制限を与える制限仮定の2種である。条件付き解釈では、“the jets failed.”の解釈を

<fail(x),{bind(x,def,jet)}>

としている。fail(x)はセンスを、{bind(x,def,jet)}は仮定の集合を示し、xが定指示で束縛できるjetのソートのどれかを示しているという情報を表している。

### Interpretation as abduction

文解釈の過程は、文が真であることの最良の説明を与えることと見なすことができる。この過程を発見的推論(abductive inference)を用いて行っている。

### A logic for semantic interpretation

意味と語用論的扱いを同一形式上で扱えるWimp2について述べ、この方法が語義・格・照応の曖昧性の除去に有効であるとしている。Wimp2は推論エンジンとしてATMSを使用したFrailを用いる。意味解釈と語用論的処理のインタフェイスは、これらの曖昧性の除去を行う場合、語用論的な推論が必要なため、問題をはらんでいた。しかし、ここでのインタフェイスは、意味解釈と語用論的処理を同じ語彙、推論エンジンで行うため、簡潔となっている。また、統語の意味的扱いは、解釈の副作用として行われる。ここでは、語の曖昧性、格の曖昧性の解消、および語用論的推論による照応(名詞・代名詞)などの解析について詳しく説明している。照応の解析は定名詞句の解析であるが、将来は、不定名詞句の解析まで取り込む予定であると述べている。

### Quantifier scoping in the SRI core language engine

英語の“a”、“every”、“each”、“some”などを含む限量子の作用域について、述べている。限量子の作用域は、他の限量子、様相、否定、ある種の統語的構造などに関連して、決定される。このための方法は、アルゴリズムに埋めこまれたものではなく、ルールによってまとめられている。これまでに、指摘されたルールを多く含むが、こ

れらの調整を行い、良い結果を得ている。実験によると、人間の決める作用域は、断片的である。これらに対応するため、また修正や拡張がしやすいように、ルールによってまとめられている。多くのルールの例とこれらを遂行するためのアルゴリズムについて詳しく考察が加えられており、これまでのものと比べ効率と拡張性に優れるとしている。多くの限量子の作用域の例が具体的に説明されている。

第2日目の最初のセッションでは、談話に関する発表が3件行われた。

### Discourse deixis: Reference to discourse segments

Bonnie Lynn Webber は、談話理解に対する計算理論的アプローチの目標を、①理解プロセスの解明を要する問題は何かを同定し、② そのプロセスとそこに関与するデータ構造を明らかにすることであるとした上で、指示の問題に関して、これまでは主に代名詞(he, she, it, etc)や定表現(the ... )による談話内対象への指示プロセスを定式化することにのみ注意が払われてきたが、直指表現(this, that, etc)による、テキストそのものに対する指示(“... That's all I know about it. ...”, etc)に関しては、これまでの研究で扱われてきたのとは別種の指示プロセスを想定する必要があることを豊富な例を挙げながら議論した。さらに、そのような直指表現の使用により指示可能なテキストは、discourse segments 等と呼ばれる単位を形成するものに限られることを示した上で、今後は、その中に現れる個々の名詞表現についてだけではなく、discourse segments そのものの性質と表現に関して研究を進めることが重要であると結んだ。

### Cues and control in expert-client dialogues

ソフトウェアに関する問題についてExpertとClientとの間で交わされた、電話会話の録音から書き起こしたデータ(450ターン)を基に、発話タイプ(Assertion、Command、Question、Promptの四つ)と、発話の主導権(Control)に関する分析を行った結果、発話タイプと主導権の移動(Control Shift)との間には密接な関連が認められたのに対して、近年の談話理論で主張されるほど、cue words(“so”、“but”、“anyway”など)と主導権の移動との間に密接な関連は認められず、(a) control phase, (b) topic, (c) global organizationの3レベルからなる談話構造を同定するにあたっては、発話の主導権が重要な役割を演じるものと報

告している。彼ら自身も述べているように、より幅広いデータによる検証が求められる。

### **A computational theory of perspective and reference in narrative**

物語中、ある登場人物の視点から記述された主観的文脈(Subjective Context)では、普通名詞や固有名詞、人称代名詞などの指示表現は、その人物の信念を反映した形で用いられる。本報告では、心理動詞(例えば、“think”)などの表層上の言語的  
手掛かりを基に現在の視点(Current Perspective)を認識し、物語中どの部分が主観的文脈であるかを同定する手法、および、その結果に基づき、指示表現の解釈を行う手法が提案されている。有効な言語的  
手掛かりなしに客観的文脈と主観的文脈間の移行が行われた場合や、手掛かりが後から示された場合(“Jody was rich and famous. Why wasn't she happy? Bill wondered.”)の扱い、あるいは、時制表現や直指表現、登場人物の目標構造や談話における注意の焦点(focus of attention)、議論構造(argument structure)等に関する情報の利用は今後の課題であると結んでいる。

その次のセッションでは、ATRからの発表も含め、日本語関係の発表が3件続いた。

### **Parsing Japanese honorifics in unification-based grammar**

PATR-IIのような素性構造の単一化に基づく文法の枠組みにおいて、「ご覧になる」や「ございます」といった固定表現を一般的に記述しようとすると、循環的な素性構造の単一化を扱うことが必要となる。ATRでは、Wroblewskiのグラフ構造の単一化アルゴリズムを拡張することにより、このような機能を必要に応じて用いることを可能としたCFGパーサーを開発している。その上で検証された次の2点について報告を行った。

- (i) 様々な情報の統一的記述に適した単一化文法の枠組みを利用して、発話中の敬語表現の使用に関する運用論的制約を得ることができる。すなわち、話題の人物や聞き手といった談話関与者に対する話し手の各種待遇態度に関する制約を各敬語表現構成素の語彙記述に含めることにより、発話場面で満足されるべき制約が計算され、これにより人称代名詞の解釈過程などにおいて重要な情報を得ることができる。また、
- (ii) 各種敬語表現構成要素が担う下位範疇化情報として、尊敬語接尾辞(「～なさる」等)、謙譲語接尾辞(「～申し上げます」等)、丁寧語接尾辞(「～ます」等)はそれぞれ自動詞句、他動詞句、文に下位

範疇化される補助成分であると記述することにより、動詞述語部分における敬語の可能な承接(「お誘い申し上げます」「\*お誘いなさり申し上げます」等)が構文論的に規制される。

なお一つめの点についてはさらに、談話表示理論の枠組みを用いて、談話内照応の「その」を例に、敬語文脈における代名詞の解釈過程の定式化を試みた。

### **Aspects of clause politeness in Japanese: An extended inquiry semantics treatment**

昨年まで京都大学に滞在し、現在は南カリフォルニア大学のISIで研究を行っているJ.Batemanによる、日本語の敬語を扱った文生成に関する講演である。システミック文法に基づく従来のテキスト生成システムでは、値選択可能な文法特徴を節点とする、ネットワーク構造により表現された文法に従い、各節点に対応付けられたChooserが知識ベース(Environment)に質問を發し、必要な決定を順次行うことにより文生成が行われた。ISIのNigelプロジェクトにおけるテキスト生成システムは、この方式に基づき幅広い範囲の英文を生成するのに成功していたが、日本語の敬語表現に端的に反映される、意味の対人関係機能をも考慮した生成を行おうとすると、文法ネットワークにおける節点とChooserを一対一に対応させることは困難であり、文法ネットワークにおける節点とChooserとの間の柔軟な対応付けを許すことにより、両レベルの記述に統一性と首尾一貫性を与えることが可能になると議論している。そして、以上のような考え方に立って実現された日本語文生成システムが生成する、敬語を含む文の実例が示された。文法特徴が適切に抽出されていれば文法ネットワークにおける節点とChooserとを厳密に対応させることにより文生成システムが構築できるとする当初の思想とこの提案は相入れないのであるかという前田からの質問に対しては、選択に際してのモジュラ性は維持されているし、またこの拡張により、文法的な選択と具体的場面での使用に関する選択との橋渡しが可能になるという主旨の答えが返ってきた。

### **Experiences with an on-line translating dialogue system**

機械翻訳装置を介して、スイスと日本との間でキーボード入力による異言語話者間(英語↔日本語)の対話実験を行った結果と、それに対する考察が報告された。対話はUNIXのtalkと類似した形で進められ、3分割された画面の上段に自分の入力

が、中段に相手の入力 of 翻訳結果が、下段には原語のまま相手の入力が表示される。タイプミス、固有名詞や略語の使用、あるいは省略などのため生じる誤りを除いて、システムはほぼ了解可能な翻訳を行ったが、この対話実験において特に注目される現象として、彼らが **metadialogue** と呼ぶものがあつたという。すなわち、翻訳内容の意が通じない場合に誤りを訂正するため交わされる、前の発言やシステム自身に関する会話である。将来的には、ユーザーのタイプミスや文法的誤りを早期に発見して修正を求める機能、**metadialogue** を行う際のきまりの確立、あいづちなどの適切な翻訳を行うためのシステムのチューン、システムの限界とくせをユーザーへ伝えることが求められると結んでいる。

### Planning coherent multisentential text

#### Two types of planning in language generation

Yale大学のSchank教授のグループからISIへ移ったE. Hovyは、生成におけるプランニングに関する研究として、2件の講演を行っている。1件目の講演では、結束性を保った文章を生成するために、MannとThompsonによって提唱された修辞構造理論(Rhetorical Structure Theory)において仮説される、連続したテキスト単位間の関係(Purpose、Sequenceなど20種類)の一部を、Cohen、Levesque、Perraultらの信念に関するモデルを基に定式化する試みが報告され、その定式化を用いて実現されたテキスト・プランナが生成する文章例が示された。

また、2件目の講演では、プランニングの種類を、階層化されたゴール構造をトップダウンに展開する **prescriptive** なものと、現在のゴール状態により複数の可能性の中から選択を行う **restrictive** なものに整理している。そして、運用論的側面を考慮した自然な文体の文章を生成するためには、両者のプランニングを行うことが可能な、彼の提唱する限定関与プランニング(**limited commitment planning**)という一種の遅延方式が有効であると述べている。この方式は、**PAULINE** システムにおいて実現され、その構成と生成文章例が示されている。

### A practical nonmonotonic theory for reasoning about speech acts

PerraultのDefault論理を用いた発話行為の分析に関する問題点を指摘し、代案として、階層的自己認識論理(**Hierarchic Autoepistemic Logic**、**HAEL**)を用いて分析する方法を提案して

いる。Perraultの枠組みにおける問題は、発話者が根拠のない命題を宣言することにより、発話者自身がその命題を信じるという誤った結論を導き出すということである。この枠組みでは、一般的な行為と信念の翻意の統一的な枠組みが問題として残されている。信念の変更を取り扱う枠組みとしては、非単調論理の形式化の一つである自己認識論理がある。この論理の中で発話行為を取り扱うためには、2つの問題があるとしている。優先性の問題と、不動点の中に命題が含まれるかどうかの判定が半決定可能であるということである。それらへの解決策として階層性を導入した**HAEL**を提案している。

### Assigning intonational features in synthesized spoken directions

統語構造だけではなく、意味や談話構造を考慮した適切なイントネーションで文章の音声合成を行うことを目的とした研究の発表である。談話構造に関する理論としては、Grosz & Sidnerの理論を、イントネーションの意味に関する理論としては、Pierrehumbertの理論を基礎にしている。上位の目的を持つ談話セグメントに対しては、下位のものよりも大きなピッチ幅を与え、下位の目的を持つ談話セグメントが後接する場合には、連続性を表すため発話末における音調の下げを控える、などのような知識を利用して、各発話にイントネーションを付与する。このことにより、いかに合成音声の自然性が改善されるかが、道路案内システム**Direction Assistant**の生成部に変更を加えることにより確かめられている。

### Atomization in grammar sharing

素性構造の持つ記述の包含関係が束を構成するという性質を活用した多言語の文法の共有手法の提案を行っている。**CUG (Categorial Unification Grammar)**における語彙記述は、素性構造で表されるが、この素性構造の包含関係の束から語彙記述のテンプレートを構成しようという試みである。従来の枠組みでは困難であった語彙記述の階層性の学習に関する新しい可能性を示している。

第3日目は、統語解析・テキスト処理を中心とした講演が行われた。

### Syntactic approaches to automatic book indexing

自動索引付けシステムの場合、単なるキーワード抽出システムと異なり、複数の単語から構成される名詞句を抽出する必要があるが、この報告で

は、テキストの統語解析を行うことにより得られた情報から索引を生成する方法を提案している。統語解析によって名詞句を同定し、その名詞句内の被修飾語と修飾語から索引の候補を生成する。そして、出現頻度、出現場所の偏り、表題などでの使用などの要因を考慮したスコアを用い、そのスコアで上位の候補を索引としている。

### Lexicon and grammar in probabilistic tagging of written English

Lancaster 大学の the Unit for Computer Research on the English Language (UCREL) では、無制限、無修正のテキストを解析するためのソフトウェアを構築することを目標に、大規模な確率付き辞書、および、文法を作成するための基本データとなる統語情報付きのテキスト・データベース(Parsed databank)を構築している。この報告では、この統語情報付きのテキスト・データベース構築の手法について報告している。文法と統語情報付きのテキスト・データベースを並行して開発するために、約6000規則の文脈自由文法を基に、文法作成者がテキストを解析し、その過程で困難な場面に直面すると、文法規則を修正・追加するという方法を用いている。テキストとして、AP通信のニュース1年分を用いて、この方法により、約10000規則の文脈自由文法を開発している。この過程において、規則のマニュアルと、CLAWS (Constituent-Likelihood Automatic Word-tagging System) という、単語にクラスを表すタグを割当てするシステムを用いている。CLAWSは、単語のタグの候補の中から、コーパスを用いて学習された確率情報を用いて、候補の選択を行っている。CLAWSによって得られたタグ付きの単語列から、テキストDB作成者がエディタを用いて解析木を表す括弧付けを行うことにより、解析木が蓄積される。

### Parsing vs. text processing in the analysis of dictionary definitions

Webster's seventh New Collegiate Dictionary 中の語釈文からの意味関係の抽出を、Sager の LSP(Linguistic String Parser)による構文解析と、grep や awk のようなUNIX の汎用テキスト処理ユーティリティの組合せによる2種類の方法で行い、両者の比較検討を行っている。ここでは、意味関係として、単語<sub>1</sub> - 関係 - 単語<sub>2</sub> という3つ組みで表されるもの、例えば、funny - SYN(nonym) - amusing、build - TAX(nonym) - form のような関係を抽出することを考えてい

る。結論として、このような単純な構造の抽出だけを行うのであれば、LSPのような一般的な構文解析は、非効率であるという結論を導いている。

### Polynomial learnability and locality of formal grammars

Goldの研究に代表されるような「極限における同定」の概念に基づく文法の学習のように正確に文法を同定するのではなく、計算の複雑性の理論の概念に基づく、近似的な文法の学習可能性 polynomial learnability を取上げ、K-local 文脈自由言語が polynomially learnable であるということを示している。

### Conditional descriptions in functional unification grammar

Kasper は、ACL で86、87年と続けて、素性を用いた言語情報記述の定式化とその拡張、および、それらの実現手法を報告してきている。この報告では、システミック文法で用いている条件を含む記述(A→Bの形式の記述)を素性に基づくシステムの上で直接的に取り扱うために、彼の従来の単一化手法を拡張し、条件記述を取り扱うことができる素性構造単一化手法を提案している。素性構造記述の充足性と両立性を定義し、その上で否定的な情報の取り扱いについて、経路に関する同値関係の記述を含まない、制限された範囲で議論を展開している。これらに従い、前年提案した選言的記述を含む素性構造の単一化手法を提案している。

### Deductive parsing with multiple levels of representation

GB理論に基づく文法を解析するための制御方式の異なる4種類の解析機構を提案している。これらの解析機構は、「演繹としての解析」アプローチのもとで、GBのPF表現を人力とし、LF表現を出力する。ここでは、GB理論で用いる4種類の表現(PF表現、LF表現、S-構造、および、D-構造)を陽に用い、これらの間の原則を満足するようにして得られるLF表現を出力する。

### Graph-structured stack and natural language parsing

富田の拡張LR解析で導入したグラフ構造スタックを、非決定的なスタック操作を取り扱うための一般的な装置としてとらえ、他の解析枠組みとの関係を述べている。グラフ構造スタックは、一般

的なスタックに、スタックの分割・結合・曖昧性の局所的なバッキングという3種類の操作を付け加えることにより、非決定的なスタック操作を許す装置である。この報告では、グラフ構造スタックの上で、ATN、LR解析、範疇文法(categorial grammar)の解析、原則に基づく文法(principle-based grammar)の解析などに見られる非決定性がどのように表されるかを示している。そして、グラフ構造スタックを用いることにより、これらの枠組みによる解析を多項式時間で行うことが可能であると述べている。また、チャート解析との関係として、スタックを用いたshift-reduce解析による全解探索が、左から右へのチャート解析であるとうことを主張している。

### An Earley-type parsing algorithm for tree adjoining grammars

TAG(Tree Adjoining Grammars)を解析するための、Earley型の解析アルゴリズムを提案している。TAGは、構文木に、別の構文木を特殊な形式で結び付けること(adjoining)により構文木を構成する文法である。そこで、部分的に解析されている構文木を表現するのに、ドット付き記号、ドット付き構文木などを導入し、これらに関する情報を含む解析状態、および、その状態の集合を定義した上で、これらの上の7種類の操作(Earleyの解析アルゴリズムにおけるscan、predict、completeに相当)を定義している。ある状態集合に対して1つの操作だけが適用でき、解析においては、これらの操作を状態集合に適用することにより進められる。この解析は、最悪の場合、 $O(n^6)$ で解析を行う。さらに、この解析では、一般の文脈自由文法に相当する構文木での置換操作を導入している。そして、このTAG解析アルゴリズムでは、置換操作を行う規則だけから構成される文法に従って解析する場合、Earleyのアルゴリズムと同様に、 $O(n^3)$ で解析を行うことができることを述べている。また、この報告では、TAGに素性に関する単一化方程式で表される制約を導入している。

### A definite clause version of categorial grammar

範疇文法における統語カテゴリを定節(definite clause)で表すことにより、1階述語で表す範疇文法を導入している。定節を用いることにより、複合的な型を表す論理式の一部で文字列中の位置を指定する整数の上の変数を共有させることにより、語順の制約を捕らえることができ、統語カテ

ゴリ中の結合方向の指定("","\")による指定)を不必要にしている。また、この方法で、位置指定の変数の間の不等式を用いることにより、関係節中の隣接していない構成素なども取り扱うことができることを述べている。

### Combinatorial categorial grammars: generative power and relationship to linear context-free rewriting systems

TAG、HGとLinear Indexed Grammarの3種類の独立に開発された文法形式によって生成される言語のクラスであるような言語のfamilarityが存在することが確立されているが、ここでは、Combinatorial Categorial Grammarも同一のクラスの言語を生成することを報告している。

### Unification of disjunctive feature descriptions

選言を含む素性構造の新しい表現方法を提案している。この方法では、値選言を含む記述を局所化することにより、素性構造の選言標準形への展開と、選言を含む素性構造の単一化における無矛盾性の確認のための単一化の繰り返しとを回避することができる。この方法は、KasperとRoundsの素性記述に関する論理を修正したものに基づき、素性記述を効率重視標準形(efficiency-oriented normal form、ENF)で表す。KasperとRoundsの論理において、経路に関する同値類を用いていたが、ここでは、非局所的な経路に関する等式を用いて表す。単一化手法においては、素性構造の非局所的な経路に関する等式を表すためにポインタを用い、このポインタに沿って非局所的な情報を移動させることにより処理を進める。ここで、素性に順序を付けることにより探索を効率化するとともに、経路に順序付けを行うことにより、情報移動の無限ループを回避している。その結果、循環構造を含む素性構造も単一化可能となっている。

### 参考文献

Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics