



Title	自然言語理解における不完全テキストの復元
Author(s)	唐沢, 博
Citation	大阪大学, 1986, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/35169
rights	
Note	著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、 ＜a href="https://www.library.osaka-u.ac.jp/thesis/#closed">https://www.library.osaka-u.ac.jp/thesis/#closed >大阪大学の博士論文について <a> をご参照ください。

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

氏名・（本籍）	から 唐	さわ 沢	ひろし 博
学位の種類	工	学	博 士
学位記番号	第	7 3 0 1	号
学位授与の日付	昭和 61 年 3 月 25 日		
学位授与の要件	基礎工学研究科 物理系専攻 学位規則第 5 条第 1 項該当		
学位論文題目	自然言語理解における不完全テキストの復元		
論文審査委員	(主査) 教 授 豊田 順一 (副査) 教 授 藤澤 俊男 教 授 高島 堅助 教 授 嵩 忠雄 教 授 都倉 信樹 教 授 鳥居 宏次		

論 文 内 容 の 要 旨

本論文は内容が二部から構成されている。第一部「不完全なテキストを知識を用いて復元する英文補完システム」は不完全な英語テキストを補完手法を用いてより完全なテキストに復元することを目的とし、第二部「省略を含む日本語テキストの復元システム」は日本語にみられる省略表現において省略されている概念を復元することを目的とする。

第一部では、①語、句、文の集まりからなるテキストを解析する手法、および②より完全なテキストを形成するために不足している語句を補う方法を提案する。①の目的のために、チャンキングおよび単一化と呼ぶプロセスを実現するアルゴリズムを考案し、これをもとに従来機械処理の対象とするのが困難であった ill-formed な文の解析について一つの方法を示す。②の目的のために、文脈・連想・推論・常識といった知識を用いて必要な語句を補う補完法を体系化し、これをもとにより完全なテキストを生成する。以上の考えに基づいたシステムについて補完能力を評価したところ、より完全なテキストを生成するのに連想と推論が寄与することを確認した。この結果からシステムをチューニングする技法を導く。

第二部では、日本語文の省略表現から、省略されている概念を復元する手法を述べる。省略表現を、③助詞の省略、④前文とのアナロジーに基づく省略、⑤予測文脈に基づく省略に大別する。オブジェクト指向モデルに基づいた解析手法を用いて係り受け関係をメッセージ交換の許容性として扱うことで③の省略を復元する。穴あけ構文のような④の省略に対しては、前文の解析結果とのアナログカルなマッチングを用いて復元する。文脈知識に依存して省略されたような⑤の省略は MOPs と呼ぶ階層的概念記憶構造のもつ予測能力を用いて復元を行う。省略表現を含む文章はその解釈に付随する曖昧性が増大

するが、この問題に対処するために多世界モデルを適用する。一つの解釈は一つの世界観の内で行われ、その妥当性は証明可能性として扱われる。

論文の審査結果の要旨

近年、電子計算機システムの高度化にともない対話インタフェースの重要性がはっきりと認識され、さまざまな研究が行われている。いわゆる自然言語理解の研究がこの範疇に入る。

本論文の第一部では、文法的に不完全な英文テキストを入力し、そのテキストを理解して、より文法構造の整ったテキストを生成する自然言語理解システムについて論じている。すなわち、英文断片の羅列を解析し、Grammatical terms Inventory と呼ぶ深層構造表現に変換した後、これに対して文法的知識、文脈知識あるいは一般的常識などを用いて書かれていなかった情報を補う補完法という手法を体系化している。さらに、補完能力の評価を行うため、補完システムの利用者が頭に描いた内容とこのシステムが生成したテキストとの比較、評価を行う実験心理学手法を導入し、この評価により補完体系の性質を明らかにすると共に、補完システムに対する tuning 技法としても有効なものであることを示している。

第二部では、第一部で得られた知見をもとに省略表現のされた日本語テキストを対象として、省略以前のテキストを復元する技法について述べたものである。省略現象を助詞の省略、前文とのアナロジーに基づく省略および予測文脈に基づく省略の3種類に制限した上で、省略・復元モデルを考え、構文・意味・談話の知識を用いて入力テキストを解析している。解析の際生ずる多義性に対しては、多世界モデルと解釈可能性に対する評価基準の導入で対処している。

自然言語処理研究において一般化された省略現象の研究はほとんど例がなく、本論文で述べられた研究成果はこの種の問題に対する新たな知見を与えているものと考えられる。よって、博士論文として価値あるものと認める。