

氏名	長尾真 なが お まこと
学位の種類	工学博士
学位記番号	論工博第118号
学位授与の日付	昭和41年11月24日
学位授与の要件	学位規則第5条第2項該当
学位論文題目	<b>STUDIES ON LANGUAGE ANALYSIS PROCEDURE AND CHARACTER RECOGNITION</b> (言語分析過程と文字認識の研究)
論文調査委員	(主査) 教授 坂井利之 教授 清野 武 教授 前田憲一

### 論 文 内 容 の 要 旨

この論文は、言語の分析過程と文字の自動認識に関するもので、言語分析過程と意味の解明の問題ならびに言語を形づくっている文字の自動認識とその認識機械の設計・実験を内容とし、2編からなっている。

第1編はオペレータ言語、コンテキスト・フリー言語の構文分析、意味を考慮した文章の電子計算機による生成を扱っていて5章からなっている。

第1章は序論で、電子計算機に用いられる形式言語、特にプログラミング言語のための基礎理論の少ないことから、この研究を行う意義を明らかにし、その位置づけを行っている。

第2章は句構造言語の中で、さらに制限の強いオペレータ言語の構造分析法を、2次元マトリックスに表現されたオペレータの優先順位のみをもとにして行えることを示している。すなわちオペレータに対して左右の優先順位を導入し、句構造法則からこの優先順位を導く方法である。

第3章はコンテキスト・フリーな句構造言語の一般構造分析法を扱っている。一般の句構造言語では言語構造にあいまいさが伴い、与えられた文章に対して複数の文章構造が考えられるので、あらゆる可能な構造を見出すアルゴリズムを句構造法則の標準形を使って記述している。ここに標準形とは句構造法則の右辺が2つ以下の要素からなるものである。

第4章では、自然言語のもつ意味を考慮して、英語の文章を電子計算機を用いて生成している。

最初に文章の中心となる概念を表わす単語（実際には主文の動詞）を選ぶ。次に展開法則を1つ選び、その法則の各要素に対して適当な単語を選ぶ際、既に選ばれた単語に対して意味的に矛盾のないよう行なう。このため約500の単語を意味分類し、各単語間の文法的位置を加味した意味的結びつきの表を作っている。生成された文章の適否により意味分類の良否を調べるのである。

さらにこのような意味情報を用いて日本語のわかち書きを電子計算機に行なわせる試みも示している。

第5章は研究の要約と将来の見透しについて述べている。

第2編は言語の要素である文字を工学的に自動的に読取る機械に関する研究およびこれら機械を電子計算機によってシミュレーションするためのプログラミング・システムについて述べたもので5章からなっている。

第1章は序論で、文字の自動認識における問題点をあげ、認識機械に対する考え方を示している。

第2章は非同期式文字読取装置の設計を論じたもので、著者の考案した文字読取の原理を3種類について詳述している。すなわち認識のパラメータとして採用している3つの論理条件に関して、その純粹の順序論理回路、禁止条件を適当に設ける方法、および定まった3つの特徴以外は禁止する方法である。

また入力装置、ベースライン検出回路等についても詳述し、判定に響く因子についても論じている。

第3章は、前章にのべた装置の改良、拡張、変形についてその優劣を論じ、2次元受光面をもつ文字読取装置が有望であることを示している。

第4章は、上記の認識装置を設計する際に、あらかじめ電子計算機でその働きをシミュレーションしている。このためのシミュレーション言語を開発し、同期式、非同期式論理回路にも適用しうるようにした点が特徴的で、実際にも極めて有効であったことを示している。

第5章は結論である。

## 論文審査の結果の要旨

言語の問題は情報処理工学の中でも最も重要であるが、また難しいテーマの一つである。

筆者は電子計算機の2つの重要な面であるソフトウェアの根本問題を第1編で扱い、ハードウェアの内現在最も展開を要望されている文字の自動読取による入力装置の問題を第2編で詳述している。

電子計算機を動作させるための形式言語と、人間の思考方法に基づく自然言語とのマッチングは情報工学における一つの大きい目標であって、筆者は自動プログラミングに用いられている言語の構造分析を解明してその手法に根拠を与え、また拡張を行っているほか、自然言語における意味分類を文章の自動生成によって把握しようとした研究が本論文の第1編である。

電子計算機の入力装置としての文字の光学的読取り、或いは文字パターンの認識、論理回路のシミュレーション・システムの開発とその実際回路への適用という研究が本論文の第2編である。

この論文の前者の内容の一部は、筆者と独立に外国の研究者によって同一結果に到達し、これは現在でもよく引用されている。後者の内容は内外に発表され、各国のこの方面の研究者のみでなくメーカーの注目を引いている。

本研究の主な成果をあげると次の通りである。

1) 電子計算機に用いられるプログラミング言語において、従来コンパイラーのための基礎理論がとぼしく、あいまい度を持たさないために経験的に約束していたことを、オペレータ言語の構造分析法としてマトリックス表現を用い、オペレータの優先順位を導入することによって、簡単にしかも明確に分析できること、隣接する要素をも調べられることを示した。

2) コンテキスト・フリーな句構造言語の一般分析法を提案した。すなわち文法を標準形にすることによって、文章の構文分析が任意の点までの情報であらゆる可能な分析ができることを示した。

これは最終的に文章にまで纏まらなると分析不可能というものでなくて一般的であり便利である。

3) 構文分析によって生ずるあいまいさは最終的には意味を持込む必要を生ずるが、筆者は文章の生成によって意味を形式で分類してその手法を展開した。これは初めての試みであり、計量言語の関係の国際会議で発表し注目を集めた。

4) タイプ活字でいろいろのホントのものに対して高速にしかも実用的に働く非同期式文字読取り方式を提案し、その論理設計・製作装置に対する実験を行ってその優秀性を確めた。この誤判定率や拒否率は発表されている同種のものに比して少く、機械に対する要求はゆるやかである。

5) 論理機械のシミュレーションのためのプログラミング・システムを開発し、これを試作した文字読取装置に適用し、製作前の改良を幾度か行ってその有効性を示した。また実験結果に基づき、試作装置の改良、新しい2次元入力判定方法を提案し、シミュレーションでその性能を調べた。

以上のように、この論文は言語分析過程と文字認識に関して、この分野の研究をかなり進展させ、学術上、工業上寄与するところが多い。よって本論文は工学博士の学位論文として価値あるものと認める。