

論文 / 著書情報
Article / Book Information

題目(和文)	先読みを持つ文法とその微分に関する研究
Title(English)	Studies on Grammars with Lookahead and Their Derivatives
著者(和文)	宮寄貴之
Author(English)	Takayuki Miyazaki
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第12512号, 授与年月日:2023年9月22日, 学位の種別:課程博士, 審査員:南出 靖彦,田中 圭介,増原 英彦,鹿島 亮,脇田 建
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第12512号, Conferred date:2023/9/22, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

(博士課程)

論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名	宮 寄 貴 之		
論文審査 審査員		氏 名	職 名		氏 名	職 名
	主査	南出靖彦	教授		脇田 建	准教授
	審査員	田中 圭介	教授	審査員		
		増原 英彦	教授			
		鹿島 亮	准教授			

論文審査の要旨（2000 字程度）

本論文は「Studies on Grammars with Lookahead and Their Derivatives(先読みを持つ文法とその微分に関する研究)」と題し、英文で全 6 章から構成されている。

正規表現は正規表現ライブラリを通じて広くプログラミングに用いられており、文脈自由文法はプログラミング言語の構文解析の基礎理論となっている。しかし、正規表現ライブラリや構文解析では、理論的な扱いが確立されていない様々な拡張が用いられている。そのような拡張の一つが肯定先読み、否定先読みと呼ばれる拡張で、文字列を消費しない照合を可能にする。本論文は、先読み付き言語という概念を導入し、先読み付き正規表現と文脈自由文法の意味論を定式化している。また、正規表現の微分を先読み付き正規表現と先読み付き文脈自由文法に拡張することで、前者の決定性オートマトンへの変換、後者の所属判定アルゴリズムを得ている。先読み付き文脈自由文法の意味論の定義においては、否定先読みが先読み付き言語に対して単調増加ではないことが問題となったが、極限による意味論と区間による意味論を与えている。区間による意味論は、接続を 2 個の変数に制限した標準形への変換を可能にし、 $O(n^3)$ の所属判定アルゴリズムを与えている。

第 1 章「Introduction」では、論文の主題となる正規表現と構文解析における先読みについて概観するとともに、主要な技術的な道具となる形式言語理論における微分について述べている。さらに、本論文の貢献についてその概要を述べている。

第 2 章「Regular Expressions with Lookahead」では、先読み付き正規表現の意味論と微分による決定性オートマトンへの変換について述べている。先読み付き正規表現に、形式言語理論的な意味を与えるために、まず、先読み付き言語（文字列の組の集合）とその演算を定義し、その代数的な構造がテスト付きクリーネ代数になっていることを示している。次に、この先読み付き言語とその演算を用いて、先読み付き正規表現の意味論を定式化している。また、正規表現の微分が先読み付き正規表現に自然に拡張できることを示し、決定性オートマトンへの変換を与えている。

第 3 章「Context-Free Grammars with Lookahead and Limit Semantics」では、先読み付き文脈自由文法を導入し、極限による意味論を与え、閉包性などの性質を議論している。先読み付き文脈自由文法の意味論の定義においては、否定先読みが先読み付き言語に対して単調増加ではないことが問題となる。本章では、Boolean Grammar に関する研究を参考にし、空集合からの極限を用いて意味論を定式化している。この意味論において、先読み付き文脈自由文法が文脈自由文法と解析表現文法（Parsing Expression Grammar）の拡張となっていることを示している。さらに、先読み付き文脈自由文法によって定義される言語の族が、ブール演算、マーク付きの接続、マーク付きクリーネ閉包について閉じていることを示している。

第 4 章「Least Interval Semantics of Context-Free Grammars with Lookahead」では、先読み付き文脈自由文法に対して、区間を用いた意味論を与えている。極限による意味論には代入や置換が意味を保存しないという問題があることを示し、標準形への変換等で障害となることを述べている。この問題を解決するために、先読み付き言語の区間を用いた意味論を導入している。先読み付き言語の区間に対する演算では否定先読みも単調増加になり、最小不動点として意味論を定義することが可能になっている。また、接続を 2 個の変数に制限した標準形への変換を与えている。

第 5 章「Derivatives of Context-Free Grammars with Lookahead」では、微分を先読み付き文脈自由文法に拡張し、 $O(n^3)$ 時間の所属判定アルゴリズムを与えている。Might らによる文脈自由文法の微分のアイデアを先読み付き文脈自由文法に応用することで、その微分を得ている。所属判定は、文字列による微分と空文字列の所属判定に帰着できる。区間による意味論では、空文字列に対する所属判定が文法の大きさに対して線形時間で可能であることを用いて、微分による所属判定が文字列の長さ n に対して $O(n^3)$ になることを示している。

第6章「Conclusions」では本論文の主要な結果を総括するとともに、未解決の問題について議論している。

以上のように、本論文では実用的な正規表現、構文解析で用いられる先読みに対して、その実装に囚われず、形式言語理論やプログラム意味論の枠組みに基づき意味を与え、閉包性などの結果を示している。また、微分を用いて先読み付き文脈自由文法に対して所属判定アルゴリズムを与え、その時間計算量が $O(n^3)$ であることを示している。これら本論文で与えられた知見は、実用的な正規表現・構文解析技術と形式言語理論の橋渡しとなるものであり形式言語理論分野への新たな価値を創造したという観点からも理学的貢献すること大である。よって、本論文は博士(理学)の学位論文として十分価値があるものと認める。

注意:「論文審査の要旨及び審査員」は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。