



Title	A Study on Some Context-Free Grammar Based Formalisms
Author(s)	中西, 隆一
Citation	大阪大学, 1994, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/38837
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	中 西 隆 一
博士の専攻分野の名称	博 士 (工 学)
学 位 記 番 号	第 11408 号
学 位 授 与 年 月 日	平成 6 年 3 月 25 日
学 位 授 与 の 要 件	学位規則第 4 条第 1 項該当 基礎工学研究科物理系専攻
学 位 論 文 名	A Study on Some Context-Free Grammar Based Formalisms (文脈自由文法に基づくいくつかの形式文法に関する研究)
論 文 審 査 委 員	(主査) 教 授 嵩 忠雄 (副査) 教 授 都倉 信樹 教 授 谷口 健一 教 授 菊野 亨 助教授 関 浩之

論 文 内 容 の 要 旨

自然言語を扱うシステムにおいて、様々な形式文法が自然言語の構文記述に用いられる。このような形式文法は (1) 自然言語を生成できる能力を持つこと, (2) 構文解析が妥当な時間 (理想的には入力系列長の多項式時間) で行えること, (3) 構文の記述が簡潔に行えること (高い記述能力を持つこと) 等の条件を満たすべきである。また自然言語の獲得過程をモデル化するため, そのような形式文法はある種の意味で学習可能であることが望ましい。そこで, 本研究では自然言語構文記述向き形式文法である多重文脈自由文法 (mcfg) 及び語彙機能文法 (lfg) について, それぞれ認識問題の計算量及び生成能力に関する考察を行った。また, mcfg 及び lfg の部分クラスを提案し, それらの生成能力についても考察した。さらに, 上述の理由で学習可能性を考えることは重要であるが, そのモデルの一つである質問からの厳密学習について, モデルの妥当性自体に問題があると考えた。そこで, 総当たりの検索を行うアルゴリズムがこのモデル上では学習アルゴリズムとなることをいくつかの場合について証明し, この結果に基づいて上記モデルの妥当性について論じた。

論 文 審 査 の 結 果 の 要 旨

文脈自由文法 (cfg) の拡張として, 多重文脈自由文法 (mcfg) や語彙機能文法 (lfg) が提案されている。本論文では, これらの形式文法, 及び計算論的学習に関し, 以下の研究結果がまとめられている。

2 章では mcfg の認識アルゴリズムが提案されている。まず, mcfg から派生する cfg が定義され, それを利用した以下の二つのアルゴリズムが示されている。

- (a) mcfg のある部分クラスに対する $O(n^2)$ 時間認識アルゴリズム (n は入力系列長)。この部分クラスは曖昧でない cfg をすべて含み, また $\{a^n b^n c^n d^n \mid n > 0\}$ や $\{wcw \mid w \in \{a, b\}^+\}$ を生成する mcfg を含む。
- (b) mcfg に対する P-RAM 上の並列認識アルゴリズム。その時間計算量は $O(\log^2 n)$ であり, プロセッサ数は n の多項式オーダーである。

与えられた mcfg が cfg であるとき, (b) のアルゴリズムの計算量は, Rytter らによる cfg 並列認識アルゴリズムとオーダー的に一致するという意味で, Rytter らのアルゴリズムの拡張となっている。

3章では lfg の生成能力について考察されている。先ず、一般の lfg の生成能力が 0 型文法と等しいことが示されている。効率的認識や構文解析を目的とした、lfg のいくつかの部分クラスが Kaplan と Bresnan, Gazdar と Pullum, 及び西野により提案されているが、本論文では、これらの部分クラスはいずれも NP 完全な言語を生成できることが示されている。従って、 $P \neq NP$ の仮定の下でこれらの部分クラスの多項式時間認識は不可能であると結論される。

さらに 4 章では、近年盛んに研究されている計算論的学習理論のモデルの一つである質問からの EXACT 学習モデルについて議論されている。具体的に、部分性質問と包含性質問の両方が利用可能である場合、多くの言語クラスの学習可能性が自明となること等を示すことにより、EXACT 学習モデルがどのような場合に妥当となるかについて考察されている。

以上のように、本論文は cfg の拡張である mcfg と lfg の生成能力及び認識問題の計算複雑さについて優れた結果を得ており、高く評価される。また、EXACT 学習モデルの妥当性に関して重要な知見を与えている。以上より、本論文は博士論文（工学）として価値のあるものと認められる。