数理解析研究所講究録
第 716 巻 1990 年 71-80



# 推論の並列化

九州大学理学部基礎情報学研究施設
宮　野　　悟
Satoru MIYANO

## 1　はじめに

推論を並列に行うことができれば，より速く結論を得られると考えるのは自然である．しかし，どの程度高速に結論を導き出すことが可能なのであろうか．次の推論規則を例にとろう．

$$
\begin{array}{lllllll}
A_1 & \leftarrow & & \qquad & A_9 & \leftarrow & \\
A_a & \leftarrow & & & A_b & \leftarrow & \\
A_c & \leftarrow & & & A_d & \leftarrow & \\
A_e & \leftarrow & & & A_f & \leftarrow & \\
A_2 & \leftarrow & A_1, A_9 & & A_3 & \leftarrow & A_2, A_a \\
A_4 & \leftarrow & A_3, A_b & & A_5 & \leftarrow & A_4, A_c \\
A_6 & \leftarrow & A_5, A_d & & A_7 & \leftarrow & A_6, A_e \\
A_8 & \leftarrow & A_7, A_f & & & &
\end{array}
$$

**例 1**

ここで，矢印の右側が空である推論規則の左のものを公理ということにする．ここでは，$A_1, A_9, A_a, A_b, A_c, A_d, A_e, A_f$が公理である．公理に推論規則を次々に適用して得られるものを定理というが，上の例では，すべての $A_t$ $(t \in \{1, ..., 9, a, ..., f\})$ が定理である．$A_8$が定理であることを示す証明木は図 1 のようになるが，$A_8$を導き出すために，7 段の推論が行われている．並列に推論が可能であったとしても，この例ではより短い段数で結論を出すことができないようにみえる．この様な例に対しては，並列化は役に立たないのであろうか．

ここでは，推論を並列に行うことができるとき，その推論の体系を変換することにより，より高速に結論を導き出すことが可能なことを示す．この推論規則の変換法の基本的アイディアは [5], [10], [11] によるが，この方法は，問題の中に潜在している自明でない並列性を抽出するのに極めて有用である．第 3 節では，いくつかの例によりその有用性を確かめる．

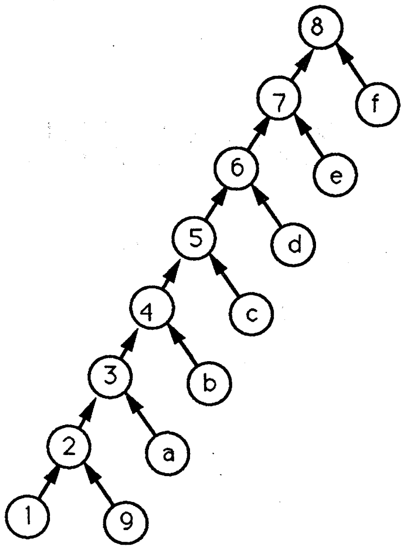


**図 1 :** $A_8$の証明木

# 2 バランス化による効率化

**推論システム**は組 $Q=(X,R)$ で与えられる．$X$は**文**の有限集合，$R$は**推論規則**の集合である．推論規則は$\alpha \leftarrow \beta_1,...,\beta_n$ の形をしている．ただし，$\alpha,\beta_1,...,\beta_n \in X$ $(n \geq 0)$ である．特に $\alpha \leftarrow$ の形のとき，文 $\alpha$ を**公理**とよび，公理の集合を $\mathbf{axiom}(Q)$ とかく．

$\mathbf{TH}(Q)$ は公理を含み，推論規則に関して閉じている最小の集合である．すなわち，次の条件を満たす $X$の最小の部分集合である．$\mathbf{TH}(Q)$ の要素を**定理**と呼ぶ．

1. $\mathbf{axiom}(Q) \subseteq \mathbf{TH}(Q)$.
2. $\beta_1,...,\beta_n \in \mathbf{TH}(Q)$ かつ $\alpha \leftarrow \beta_1,...,\beta_n \in R$ ならば，$\alpha \in \mathbf{TH}(Q)$ である．

$Q$ の**推論木**とは，$X$の要素を頂点のラベルとする有限木 $T$で次の条件を満たすものである：もしラベル $\alpha$ のついた頂点が，ラベル $\beta_1,...,\beta_n$ のついた $n$ 個 $(n \geq 0)$ の頂点を子として持つならば $\alpha \leftarrow \beta_1,...,\beta_n$ である．推論木 $T$ の頂点の個数および高さを，それぞれ $size(T)$ および $height(T)$ で表す．

葉のラベルが全て公理である推論木を**証明木**とよぶ．定理$\gamma \in \mathbf{TH}(Q)$ に対して，$\gamma$を根のラベルとする証明木 $T(\gamma)$ が存在することは明らかであろう．このとき，

$$size(\gamma) = \min\{size(T(\gamma)) \mid T(\gamma) \text{ は } \gamma \text{ の証明木}\},$$
$$size(Q) = \max\{size(\gamma) \mid \gamma \in \mathbf{TH}(Q)\}$$

と定義する．

各推論規則 $\alpha \leftarrow \beta_1,...,\beta_n$ $(n \geq 3)$ に対して，新しい文$\gamma_1,...,\gamma_{n-2}$ を用いてこれを $(\gamma_1 \leftarrow \beta_1,\beta_2)$, $(\gamma_2 \leftarrow \gamma_1,\beta_3)$,..., $(\alpha \leftarrow \gamma_{n-2},\beta_n)$ に置き換える．この置き換えにより，推論規則の

個数は線形的にしか増加しない．これにより，以後は，$Q=(X,R)$ の推論規則の左側は高々2つの文しかないと仮定する．

与えられた推論システム $Q=(X,R)$ から次のような**バランス化推論システム** $Q^B=(X^B,R^B)$ を構成する．$X^B=X\cup(X\times X)\cup(X\times(X\times X))$ である．$R^B$ は次のものからなる．

1. $R$ の推論規則に対して，それぞれ $R^B$の推論規則を次のように定義する．

| $R$の推論規則 | $R^B$の推論規則 |
|---|---|
| $x\leftarrow$ | $x\leftarrow$ |
| $x\leftarrow y$ | $(y,x)\leftarrow$ |
| $x\leftarrow z,y$ | $(z,(y,x))\leftarrow$ |
| $x\leftarrow z,y$ | $(y,(z,x))\leftarrow$ |

2. $x\leftarrow y,(y,x)$
3. $(y,x)\leftarrow z,(z,(y,x))$
4. $(z,x)\leftarrow(z,y),(y,x)$

こうして構成される推論システム $Q^B$の文の個数および推論規則の個数は多項式的にしか増加していない．この $Q^B$に対して次の2つの命題が成立する．

**命題 1**　$\mathbf{TH}(Q)=\mathbf{TH}(Q^B)\cap X$.

**命題 2**　$\gamma\in\mathbf{TH}(Q)$ を $Q$ の定理とする．このとき，$Q$ における $\gamma$ の任意の証明木 $T(\gamma)$ に対して，バランス化推論システム $Q^B$ における $\gamma$ 証明木 $T'(\gamma)$ で

$$height(T'(\gamma))\leq 7\cdot\log(size(T(\gamma)))$$

を満たすものが存在する．

命題1は構成から明らかである．命題2の証明は次の補題を必要とする．

**補題 3**　$T$を$k$分木とする．このとき，

$$size(T')<k/(k+1)size(T)$$
$$size(T'')<k/(k+1)size(T)$$

を満たすような$T$の頂点$v$が存在する．ただし，$size(T)$ は葉の個数である．$T'$は$v$における$T$の部分木で，$T''$は$T$から$T'$を取り除いてできる部分木である．

**証明**　略．

**命題 2 の証明**　$x,y\in X$とする．$Q$ における**仮定** $y$**での** $x$ **の部分証明木**とは，$Q$ の推論木で，根のラベルが $x$，葉の中の1つがラベル $y$で他の葉のラベルは全て $Q$ の公理であるものをいう．このとき，次の (1), (2) は同値である．

(1) $(y,x)$ が $Q^B$の定理である．

(2) $Q$ における仮定 $y$ での $x$ の部分証明木が存在する．

正整数 $m$ に対して，$S_1(m), S_2(m)$ を次のように定義する．

(3) $S_1(m) = \{x \in X \mid$ サイズが高々 $m$ の $Q$ における $x$ の証明木がある $\}$.

(4) $S_2(m) = \{(y, x) \in X \times X \mid$ サイズが高々 $m$ の $Q$ における仮定 $y$ での $x$ の部分証明木がある $\}$.

このとき，$H(m)$ を

$$H(m) = \max\{height_{Q^B}(\alpha) \mid \alpha \in S_1(m) \cup S_2(m)\}$$

と定義する．サイズが 1 の証明木をもつものは，$Q$ の公理だけであるので，$H(1) = 0$ である．また，サイズが 2 の証明木をもつものは $R$の推論規則 $x \leftarrow y$から得られるものであり，このとき $(y, x)$ は $Q^B$の公理であるので，$H(2) = 0$ である．そして，$H(3) = 1$ となる．$H(m)$ を評価するため次の 2 つの場合を考える．

1. $x \in S_1(m)$ としよう．このときサイズが高々$m$ の $Q$ における $x$ の証明木 $T$がある．補題 3 より，$T$の頂点 $y$で，$T_1$を $T$の $y$を根とする部分木とし，$T_2$を $T$から $T_1$の根を除いて $T_1$を取り除いてできる部分木とするとき，

$$size(T_1) < \frac{2}{3}m, \qquad size(T_2) < \frac{2}{3}m$$

となるものが存在する．よって，$x \in S_1((2/3)m), (y, x) \in S_2((2/3)m)$ である．このことから，$Q^B$の推論規則 $x \leftarrow y, (y, x)$ を使い $Q^B$における $x$ の証明木で高さが高々 $H((2/3)m) + 1$ のものがある．したがって

$$H(m) \leq H((2/3)m) + 1$$

となる．

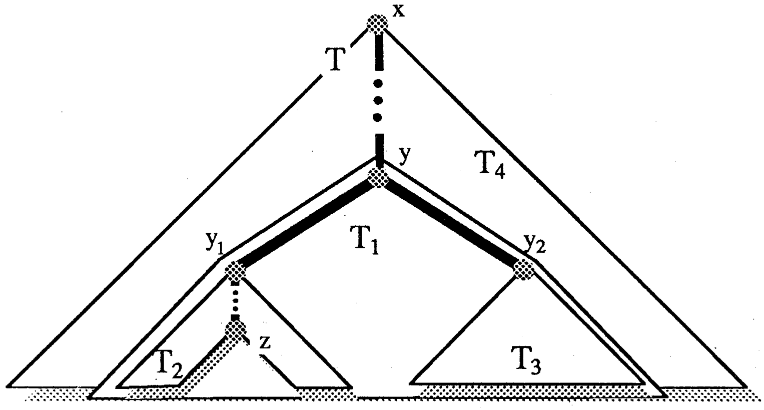


**図 2 :** $Q$ における仮定 $z$での $x$ の部分証明木

2. $(z,x) \in S_2(m)$ としよう．このときサイズが高々 $m$ の $Q$ における仮定 $z$ での $x$ の部分証明木 $T$ がある．そこで葉 $z$ から根 $x$ へいたる道を考える．この道を $z$ から $x$ へたどっていくとき $size(T_1) > size(T)/2$ となる初めての点を $y$ とする．ただし，$T_1$ は $y$ を根とする $T$ の部分木である．$y$ のとりかたより図 2 のようになり，$T_2$ を $y_1$ を根とする部分木から $z$ を根とする部分木を取り除いたもの，もし $y_2$ があれば $T_3$ を $y_2$ を根とする部分木，$T_4$ は $T$ から $T_1$ を取り除いた木であるとすると

$$size(T_4) \le \frac{1}{2}m, \qquad size(T_2) \le \frac{1}{2}m$$

である．また明らかに

$$size(T_3) \le m$$

である．したがって，$y_2$ については 1 で構成した高さが高々 $H((2/3)m) + 1$ の証明木を用い，図 3 のような $Q^B$ の証明木を構成すれば $(z,x)$ の証明木となる．そして，その高さは高々 $H((2/3)m) + 4$ である．よって

$$H(m) \le H((2/3)m) + 4$$

となる．

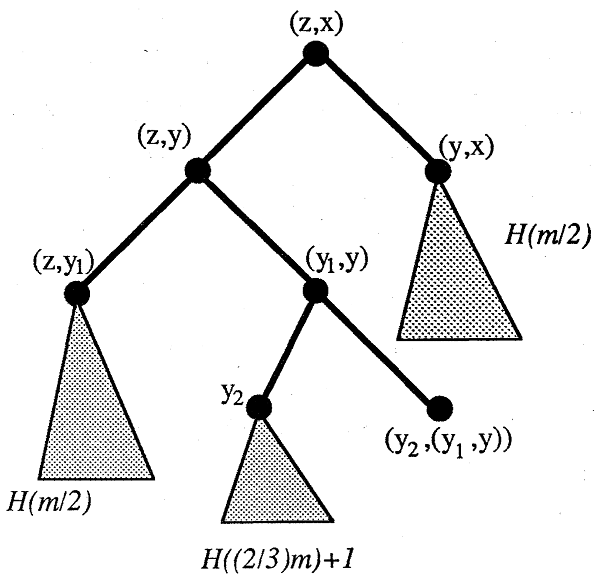


**図 3：$Q^B$ における $(z,x)$ の証明木**

以上のことから，$H(m) \le H((2/3)m)+4$ となり，この $H(m)$ の不等式を解くことにより

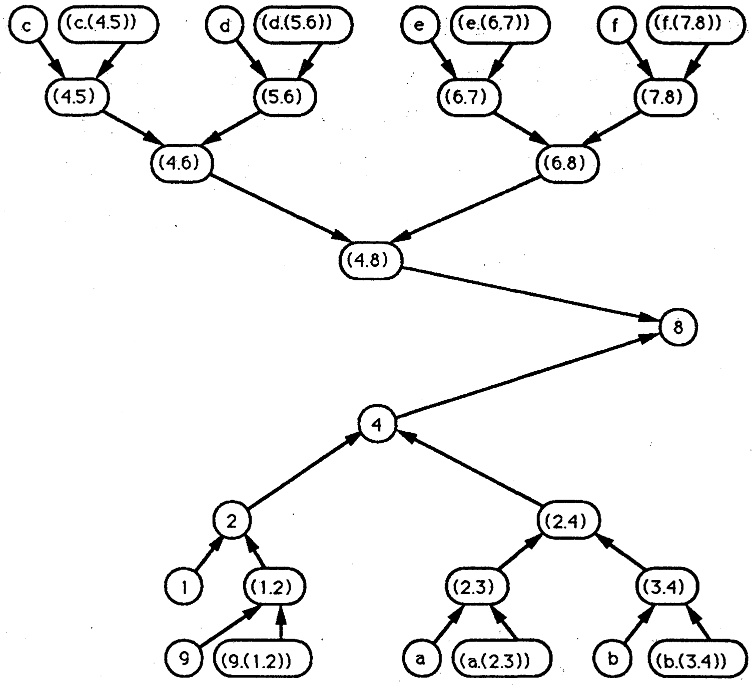


**図 4：バランス化推論システムの証明木**　例 1 の推論システムをバランス化したときの証明木である．推論の段数が減っている．

$$H(m) \leq 4\log_{\frac{3}{2}} m = \frac{4}{\log_2 3 - 1}\log m < 7\log m$$

を証明できる．□

図 4 は，例 1 の推論システムの図 1 の証明木に対応するバランス化推論システムの証明木の例である．推論の総数は 2 倍になっているが推論の段数は減少している．

命題 1, 2 より，CRCW PRAM 上で多項式個のプロセッサを使って，$\mathbf{TH}(Q^B)$ を時間 $O(\log(size(Q)))$ で計算できる (Algorithm 1).

**begin**
　**for** $k = 1$ **to** $7\cdot\log(size(Q))$
　　**pardo for each** $\alpha \leftarrow \beta_1, ..., \beta_n \in R^B$
　　　**if** $\{\beta_1, ..., \beta_n\} \subseteq \mathbf{TH}(Q^B)$
　　　　　　**then** $\mathbf{TH}(Q^B) \leftarrow \mathbf{TH}(Q^B) \cup \{\alpha\}$
　　**end pardo**
　**end for**
**end**

**Algorithm 1: $\mathbf{TH}(Q^B)$**

一般的には，$size(Q)$ は $|X|$ に関して指数的となるが，もし $size(Q)$ を多項式で押さえることができるような $Q$ にたいしては，$O(\log n)$ 時間の並列アルゴリズムとなる．

$\{Q_i = (X_i, R_i)\}_{i\geq 1}$ を推論システムの族とする．ある多項式 $p(n)$ が存在して，各定理 $\gamma \in \mathbf{TH}(Q_i)$ が $size(T(\gamma)) \leq p(|X_i|)$ となる $Q_i$ の証明木 $T(\gamma)$ をもつとき，この族は**多項式サイズ**であるという．

以上のことをまとめると次の定理になる．

**定理 4**　$\{Q_i\}_{i\geq 1}$ を多項式サイズの推論システムの族とする．このとき $\mathbf{TH}(Q_i)$ は，多項式個のプロセッサ数を用いて CRCW PRAM により $O(\log n)$ 時間で計算できる．ただし，$n = |X_i|$ とする．

## 3　逐次アルゴリズムから並列アルゴリズムへ

前節でのべた結果のおもしろい側面として，逐次アルゴリズムを並列アルゴリズムに変換できることがある．つまり，ある問題が，多項式サイズの推論システムの族に $\mathrm{NC}^1$還元可能ならば，その問題は $\mathrm{NC}^2$に入ることがいえる．このアイディアにより，いくつかの問題を並列化できる．以下にそのおもな例を述べよう．

1. 2-Satisfiability Problem (2SAT) [1], [4] は，多項式個のプロセッサの CRCW PRAM により時間 $O(\log n)$ で計算可能である．

   $S = \{\{\alpha_1, \beta_1\}, \{\alpha_2, \beta_2\}, ..., \{\alpha_m, \beta_m\}\}$ を 2SAT の instance とする．ただし，$\alpha_i, \beta_i \in \{x_1, \neg x_1, ..., x_n, \neg x_n\}$. $S$ の充足可能性を判定するには，次の unit resolution を行う多項式サイズの推論システム $Q_S = (X_S, R_S)$ を構成すればよい．

$$\begin{aligned} X_S &= \{\{\alpha, \beta\} \mid \alpha, \beta \in \{x_1, \neg x_1, ..., x_n, \neg x_n\}\} \\ &\quad \cup\{\Box\} \end{aligned}$$

   $Q_S$ の推論規則は次のものである．

$$\begin{array}{rcll} \{\alpha_i, \beta_i\} & \leftarrow & & (i = 1, ..., m) \\ \{\alpha, \gamma\} & \leftarrow & \{\alpha, \beta\}, \{\neg\beta, \gamma\} & \\ \{\gamma\} & \leftarrow & \{\beta\}, \{\neg\beta, \gamma\} & \\ \Box & \leftarrow & \{\beta\}, \{\neg\beta\} & \end{array}$$

   ここで $\alpha, \beta, \gamma$ は $\{x_1, \neg x_1, ..., x_n, \neg x_n\}$ のリテラルである．

   このとき，$S$ が充足可能でないことの必要十分条件は，リテラルの列 $\gamma_1, ..., \gamma_k$ が存在して，次の (1) と，(2) もしくは (3) のどちらか一方が成立することである [3]．そして $\Box \in \mathbf{TH}(Q_S) \Leftrightarrow S \notin$2SAT となる．$size(Q_S)$ は $O(m)$ である．

   (1) すべての $1 \leq i \leq k-1$ について，$\{\neg\gamma_i, \gamma_{i+1}\} \in S$.

   (2) $\{\gamma_1\} \in S$ かつ $\gamma_k = \neg\gamma_1$.

   (3) ある $1 < j < k$ に対して $\gamma_1 = \neg\gamma_j = \gamma_k$.

2. 文脈自由言語の認識は $NC^2$ で可能．

例で説明しよう．$G$ を次の生成規則より成る文脈自由文法とする．

$$S \to SS,\ S \to (S),\ S \to (),$$

ここで，“(” と “)” は終端記号である．記号列 $w = x_1 \cdots x_n$ に対して，推論システム $Q_w = (X_w, R_w)$ を次のように定義する．

$X_w = \{S[i,j] \mid j - i$ は偶数, $0 \leq i < j \leq n\}$.

$R_w$ は

(1) $x_{i+1}x_{i+2} = ()$ のとき，$S[i, i+2] \leftarrow$.

(2) $x_i = ($ かつ $x_{j+1} = )$ のとき，$S[i-1, j+1] \leftarrow S[i,j]$.

(3) $S[i,k] \leftarrow S[i,j], S[j,k]$.

よりなる．このとき，

$$S[i,j] \in \mathbf{TH}(Q_w) \iff S \Rightarrow^* x_{i+1} \cdots x_j$$

となり，族 $\{Q_w\}$ は定理 4 の条件を満たす．

3. 文脈自由言語のクラスとプッシュダウン・オートマトンによって受理される言語のクラスが一致することはよく知られている．このプッシュダウン・オートマトンをさらに一般化したものに非決定性 auxiliary プッシュダウン機械 [2] がある．これは，2 方向入力テープ（左右両方向に動ける読み取り専用のヘッドのついたテープ）をもつプッシュダウン・オートマトンに，読み書き可能な 1 本の作業用テープと書き込みのみ可能な出力テープが備わったものである．全ての入力に対して，出力テープに書かれる記号列が一意に決まるとき，関数を計算するという．クラス AuxPDA($n^{O(1)}, \log n$) は，多項式時間で走り，作業用テープの領域を高々$O(\log n)$ 使う（プッシュダウンスタックおよび出力テープに領域の制限はない）非決定性 auxiliary プッシュダウン機械によって計算可能な関数のクラスである．このクラスは，多項式サイズの推論システムを構成することで効率よく並列化可能である [7]．実際，AuxPDA($n^{O(1)}, \log n$)$\subseteq NC^2$ となることが知られている [8], [9]．このクラスは，ソーティング，パターンマッチング，パージング等の関数を含んでいるとても広いクラスである．

4. 辞書式順序で最初の極大な独立点集合問題 (LFMIS) は P 完全であるが，森 (forest) に制限すると，$NC^2$ にはいる [7]．

森 $G = (V, E)$，$V = \{1, ..., n\}$，に対して推論システムを次のように構成する：各 $i \in V$ に対して，$N(i) = \{j \mid \{i,j\} \in E,\ j < i\}$ とする．推論規則は次のように与える：

(1) 各 $j \in N(i)$ に対して，$\neg i \leftarrow j$.

(2) $i \leftarrow \neg j_1, ..., \neg j_t$. ただし $N(i) = \{j_1, ..., j_t\}$.

このとき，頂点 $i$ が辞書式順序で最初の極大な独立点集合に属していることと $i$ がこの推論システムの定理であることが同値となる．さらに，この推論システムは，多項式サイズである．

5. 完全グラフ $K_4$ と同相 (homeomorphic) な部分グラフをもたないグラフに対して，辞書式順序で最初の辺で生成される3サイクルなしの極大な部分グラフを求める問題は $NC^2$ にはいる．平面グラフで全ての頂点が1つの面上にあるものを外平面グラフ (outerplanar graph) という．外平面グラフであるための必要十分条件は，$K_4$ および $K_{2,3}$ に同相な部分グラフを含まないことが必要十分であるから，外平面グラフに対しても，同様のことがいえる．このことも多項式サイズの推論システムを構成することにより示すことができる [6]．

## 4 おわりに

本稿で述べた並列化技法は，問題の並列化に際し，一つの大域的な視点を与えるものと考えられる．しかし，並列化に必要なプロセッサの数は多項式とはいえ，あまり現実的とはいえないだろう．

## 参考文献


[1] Cook, S.A. and Luby, M.: A Simple Parallel Algorithm for Finding a Satisfying Truth Assignment to a 2-CNF Formula, *Inf. Process. Lett.*, Vol. 27 (1988), pp. 141-145.

[2] Hopcroft, J.E. and Ullman, J.D.: *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley Publishing Company, 1979.

[3] Jones, N.D., Lien, Y.E. and Laaser, W.T.: New Problems Complete for Nondeterministic Log Space, *Math. Systems Theory*, Vol. 10 (1976), pp. 1-17.

[4] Karp, R.M. and Wigderson, A.: A Fast Parallel Algorithm for the Maximal Independent Set Problem, *J. ACM*, Vol. 32 (1985), pp. 762-773.

[5] Miller, G. and Reif, J.: Parallel Tree Contraction and its Application, *Proc. 26th IEEE FOCS*, 1985, pp.478-489.

[6] Miyano, S.: A Parallelizable Lexicographically First Maximal Edge-Induced Subgraph Problem, *Inf. Process. Lett.*, Vol. 27 (1988), pp. 75-78.

[7] Miyano, S.: Parallel Complexity and P-Complete Problems, *Proc. Int. Conf. Fifth Generation Computer Systems 1988*, 1988, pp. 532-541.

[8] Ruzzo, W.L.: Tree-Size Bounded Alternation, *J. Comput. Syst. Sci.*, Vol. 21 (1980), pp. 218-235.

[9] Ruzzo, W.L.: On Uniform Circuit Complexity, *J. Comput. Syst. Sci.*, Vol. 22 (1981), pp. 365-383.

[10] Rytter, W.: The Complexity of Two-Way Pushdown Automata and Recursive Programs, *Proc. NATO Advanced Research Workshop on Combinatorial Algorithms on Words*, Springer-Verlag, 1985, pp. 341-356.

[11] Rytter, W.: Parallel Time $O(\log n)$ Recognition of Unambiguous CFL's, *Proc. Fundamentals of Computation Theory* (Lecture Notes in Computer Science, Vol. 199), 1985, pp. 380-389.