自 然 言 語 処 理 101－9
（1 9 9 4. 5. 27）

Earley アルゴリズムの並列化手法

中村 貞吾 日高 達
(九州工業大学) (九州大学)

文脈自由文法の並列構文解析の手法に関しては，これまでに CYK 法に基づくものや，LR 解析法を一般の文脈自由文法に対して拡張した一般化 LR 法に基づいたアルゴリズム等が提案されている．本稿では，最も効率的な逐次型構文解析アルゴリズムである Earley 法に基づく一般の文脈自由文法に対する並列的構文解析アルゴリズムを与える．Earley 法は，入力に対して left-to-right に入力に同期した解析を行なうため，shift 操作は reduce 操作の終了まで待たなくてはならない．本アルゴリズムでは，Earley 法のアイテムに "gap" を導入して reduce 操作の連鎖を分解することによって，$O(n)$ 時間，$O(n^2)$ 空間，$O(n^2)$ プロセッサの能率を実現している．

Parallel Technique for Earley's Algorithm

Teigo NAKAMURA Toru HITAKA
(Kyushu Institute of Technology) (Kyushu University)

Parallel parsing methods for context-free grammars are developed based on CYK algorithm and generalized LR algorithm so far. This paper presents a parallel parsing algorithm based on Earley's algorithm, that is the most efficient sequencial algorithm, for arbitrary context-free grammars. This algorithm uses "item with gap" to break up sequences of reduction operations. It works in $O(n)$ time using $O(n^2)$ processors and $O(n^2)$ spaces.

1 はじめに

文脈自由文法の構文解析法としては，CYK 法 [1] や Earley 法 [2]，LR 解析法を一般の文脈自由文法に対して拡張した一般化 LR 法 [4] などが良く知られている．これらは，いずれも逐次処理のアルゴリズムであり，Earley 法では入力文字列の長さを n として $O(n^2)$ 空間，$O(n^3)$ 時間 (文法が曖昧でない場合は $O(n^2)$ 時間) で構文解析を行なうことができる．一方，文脈自由文法の並列構文解析の手法に関しては，これまでに CYK 法に基づくもの [3] や一般化 LR 法に基づいたアルゴリズム [6][7] 等が提案されている．

本稿では，最も効率的な逐次型構文解析アルゴリズムである Earley 法に基づく一般の文脈自由文法に対する並列的構文解析アルゴリズムを与える．Earley 法は入力に対して left-to-right に入力に同期した解析を行なうため，自明な並列化を行なった場合，解析時間は $O(n^2)$ となる．これを解決するために，本アルゴリズムでは Earley 法のアイテムに "gap" を導入して reduce 操作の連鎖を分解することによって，$O(n)$ 時間，$O(n^2)$ 空間，$O(n^2)$ プロセッサの能率を実現している．

2 準備

【定義 1】 文脈自由文法を $G = \langle N, \Sigma, P, S \rangle$ とする．特に断わらない限り，記号 $X, Y, Z, \ldots$ や $\alpha, \beta, \gamma, \ldots$ などをそれぞれ $X, Y, Z \in N$, $\alpha, \beta, \gamma \in (N \cup \Sigma)^*$ の意味で使用する．また空記号列を ϵ と記す．以下の議論では，構文解析の入力文字列を $w = a_1 a_2 \ldots a_n$ に固定する．そして $[i, j]$ で入力文字列の部分列 $a_{i+1} a_{i+2} \ldots a_j$ を表わすものとする．

並列アルゴリズムの記述は，PASCAL 的構文の PRAM(Parallel Randam Access Machine) プログラムによって行なう．
構文 **parallel** x : $P(x)$ **do** $operation(x)$ は，$P(x)$ が真となる各 x に対して $operation(x)$ を並列に実行することを意味するものとする． ■

3 並列 Earley アルゴリズム

高速な並列構文解析アルゴリズムを実現するためには，解析表作成の際の処理操作の連鎖を互いに独立な処理操作に分解することが必要になる．峯 [6] は，一般化 LR 法を基にして入力に対して非同期に解析を進めることによって並列性を高め，$O(n)$ 時間のアルゴリズムを与えた．また，Rytter[3] は CYK 法で作成する解析表中に解析が未完な部分を表わす "gap" を導入することによって高度の並列化を行ない，$O(\log n)$ 時間の認識アルゴリズムを実現している．Earley 法は，入力に対して left-to-right に同期を取る解析法であるため，その並列化を行なうにあたっては，[3] と同様に Earley 法のアイテムに gap を導入する．

3.1 gap 付きアイテム

Earley 法におけるアイテムおよび gap を導入したアイテムを図 1，図 2 に示す．Earley 法は開始記号からの top-down 予測により最左導出に基づいてアイテムを構成していくため，gap については終端記号列の直後に現れるもののみを考慮する．

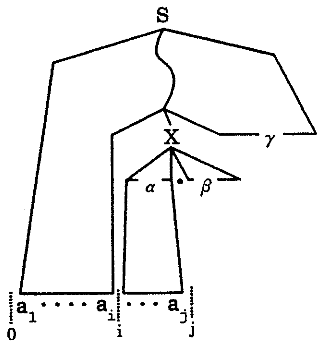

図 1: Earley 法のアイテム $\langle X \rightarrow \alpha \cdot \beta,\ i,\ j \rangle$

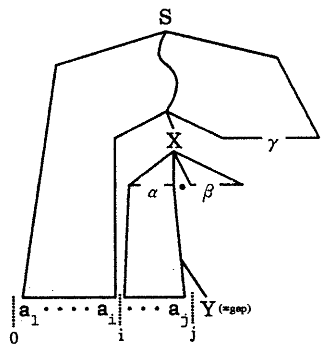

図 2: gap 付きアイテム $\langle X \rightarrow \alpha \cdot \beta,\ i,\ j,\ Y \rangle$

3.2 解析表作成アルゴリズム

【定義 2】 解析表 PL は次の性質を満たすアイテムの集合である．

$$\langle X \to \alpha\cdot\beta, i, j, Y\rangle \in PL$$
$$\Updownarrow$$
$$\begin{cases} S \stackrel{*}{\Longrightarrow} [0,i]X\gamma,\ \alpha \stackrel{*}{\Longrightarrow} [i,j]Y, \\ 0 \le i \le j \le n,\ X \to \alpha\beta \in P \end{cases}$$

また，$\langle *, *, j, * \rangle$ の形 (* は任意) のアイテム集合を PL_j で表わす． ■

解析表作成アルゴリズムを図 3 に示す．記号 ➡ は共有メモリ上に確保された解析表に対するアイテムの登録操作を表わす．アルゴリズム中で用いられている $Y \stackrel{*}{\Longrightarrow} X\eta$ および $Z \stackrel{*}{\Longrightarrow} Y$ の判定に関しては，文法が与えられた時点で予め計算してテーブル化しておくことにより，アルゴリズム実行中では定数ステップで判定を行なうことができる．

本アルゴリズムは ϵ-free の文脈自由文法に対して入力文字列の解析表を作成する．またアルゴリズム記述の便宜上，ダミーアイテムとして $\langle \phi \to \cdot S, 0, 0, \epsilon\rangle$ を初期登録する．ここで，$\phi \notin (N \cup \Sigma)$ である．解析表に登録されるアイテムのうち $\langle X \to \alpha\cdot\beta, i, j, \epsilon\rangle$ の形のアイテムは，逐次型の Earley 法で登録されるアイテム $\langle X \to \alpha\cdot\beta, i, j\rangle$ に対応している．作成された解析表中に $\langle \phi \to S\cdot, 0, n, \epsilon\rangle$ が登録されていれば，入力文字列は受理される．

3.3 構文構造の抽出

アルゴリズム A1 によって入力が受理された場合，PL を検索することによって入力文字列の構文構造を出力することになる．図 4 に構文構造の抽出アルゴリズムを与える．

PL 作成アルゴリズム A1 および構文構造抽出アルゴリズム A2 に対して次の定理が成り立つ．

【定理 1】 PL 作成アルゴリズム A1 は $O(n^2)$ プロセッサ，$O(n^2)$ 記憶域を用いて $O(n)$ 時間で PL を作成する．また，構文構造抽出アルゴリズム A2 は $O(n)$ プロセッサを用いて PL から $O(n)$ 時間で構文構造を 1 つ抽出する． ■

上記アルゴリズムの正当性に関しては，付録 A に示す．

$\langle \phi \to \cdot S, 0, 0, \epsilon\rangle$ ➡ PL_0 /*ダミーアイテムの登録*/
for $j := 0$ **to** n **do**
  begin
    [step-1] **parallel** $i : 0 \le i \le j-1$ **do**
      $\langle X \to \alpha'\cdot a_j\beta, i, j-1, \epsilon\rangle \in PL_{j-1}$ に対して
      $\langle X \to \alpha' a_j\cdot\beta, i, j, \epsilon\rangle$ ➡ PL_j
    [step-2] **parallel** $(i,k) : 0 \le i \le k \le j-1$ **do**
      $\langle X \to \alpha\cdot\beta, i, k, W\rangle \in PL_k$, $\langle W \to \gamma\cdot, k, j, \epsilon\rangle \in PL_j$ に対して
      $\langle X \to \alpha\cdot\beta, i, j, \epsilon\rangle$ ➡ PL_j
    [step-3] **parallel** $i : 0 \le i \le j$ **do**
      $\langle Z \to \gamma\cdot Y\delta, i, j, \epsilon\rangle \in PL_j$, $Y \stackrel{*}{\Longrightarrow} X\eta$, $X \to \beta \in P$ に対して
      $\langle X \to \cdot\beta, j, j, \epsilon\rangle$ ➡ PL_j
    [step-4] **parallel** $i : 0 \le i \le j$ **do**
      $\langle X \to \alpha'\cdot Z\beta, i, j, \epsilon\rangle \in PL_j$, $Z \stackrel{*}{\Longrightarrow} Y$ $(Y \in N)$ に対して
      $\langle X \to \alpha' Z\cdot\beta, i, j, Y\rangle$ ➡ PL_j
    [step-5] **parallel** $(i,k) : 0 \le i \le k \le j-1$ **do**
      $\langle X \to \alpha\cdot\beta, i, k, W\rangle \in PL_k$, $\langle W \to \gamma\cdot, k, j, Y\rangle \in PL_j$ $(Y \in N)$
      に対して $\langle X \to \alpha\cdot\beta, i, j, Y\rangle$ ➡ PL_j
end

図 3: 解析表 PL の作成アルゴリズム A1

$\langle \phi \to S\cdot,\ 0,\ n,\ \epsilon \rangle \in PL_n$ならば，以下の手続き $extract(\langle \phi \to S\cdot,\ 0,\ n,\ \epsilon \rangle)$ を実行する．

◆ $extract(\langle X \to \alpha \cdot \beta,\ i,\ j,\ \epsilon \rangle)$ の定義

if $\alpha = \alpha' a_j$ **then**

$extract(\langle X \to \alpha' \cdot a_j\beta,\ i,\ j-1,\ \epsilon \rangle)$

else if $\alpha = \alpha' Y\ (Y \in N)$ **then**

begin

parallel $k : i \le k \le j$ **do**

$\langle Y \to \gamma\cdot,\ k,\ j,\ \epsilon \rangle \in PL_j,\quad \langle X \to \alpha' \cdot Y\beta,\ i,\ k,\ \epsilon \rangle \in PL_k$

の組を検索し，共有メモリに書き込む．

共有メモリ上の 2 つのアイテムを読み込む．

$extract(\langle Y \to \gamma\cdot,\ k,\ j,\ \epsilon \rangle)$;

$extract(\langle X \to \alpha' \cdot Y\beta,\ i,\ k,\ \epsilon \rangle)$;

$Y \to \gamma$ を出力．

end

図 4:　構文構造抽出アルゴリズム A2

3.4　空規則に対する拡張

解析表作成アルゴリズム A1 を，ϵ 規則を含む一般の文脈自由文法に対して拡張する．逐次型の Earley アルゴリズムでは，ϵ 規則が含まれる文法に対しては終端記号のシフト処理を待ち合せた上で，解析表に新しいアイテムが登録されなくなるまで reduce 操作を繰り返す必要があった．ここでは，動的に ϵ 規則による reduce を行なうのではなく，各非終端記号 $X \in N$ に対して $X \stackrel{*}{\Longrightarrow} \epsilon$ かどうかを判定するテーブルを予め作成しておくことによって，シフト操作時に ϵ 規則による reduce 操作と同様な処理を行なう．そのアルゴリズムを図 5 に示す．

アルゴリズム A1 の **[step-1]**, **[step-2]** および **[step-3]** の各ステップの直後にそれぞれ以下のステップを挿入する．

[step-6]　**parallel** $i : 0 \le i \le j$ **do**

$\langle X \to \alpha \cdot \beta'\beta'',\ i,\ j,\ \epsilon \rangle \in PL_j,\quad \beta' \stackrel{*}{\Longrightarrow} \epsilon$ に対して

$\langle X \to \alpha\beta' \cdot \beta'',\ i,\ j,\ \epsilon \rangle \Rightarrow PL_j$

図 5:　ϵ 規則を含む文法に対する解析表作成アルゴリズム A3

4　解析実験

上記アルゴリズム A3 をワークステーション (SUN SparcStation 10) 上に Common Lisp を用いて実現し，並列動作のシミュレーションを行なった．実験では，9 つの文法と各文法に対する 20 ～ 30 個程度の入力文を用いて解析表の作成を行ない，本アルゴリズムによって作成されるアイテム数を調査した．全体の結果を表 1 に示し，文法 G_1 に対して $a + a * a$ を入力として与えた場合の解析表を以下に示す (ここで，ダミーアイテムは除く)．gap を持つアイテムの個数は gap を持たないアイテムの個数と同程度であることから，本アルゴリズムによって作成される解析表は通常の Earley 法の解析表の 2 倍程度になることがわかる．本実験は逐次マシン上でのシミュレーションであることから，実行時間に関してはここでは特に示していないが，並列動作の 1 ステップを $O(n^2)$ で模倣するため $O(n^3)$ 程度になることが確認できた．

PL_0

$\langle E \to \cdot T,\ 0,\ 0,\ \epsilon \rangle$
$\langle E \to \cdot T + E,\ 0,\ 0,\ \epsilon \rangle$
$\langle T \to \cdot F,\ 0,\ 0,\ \epsilon \rangle$
$\langle T \to \cdot F * T,\ 0,\ 0,\ \epsilon \rangle$
$\langle F \to \cdot [E],\ 0,\ 0,\ \epsilon \rangle$
$\langle F \to \cdot a,\ 0,\ 0,\ \epsilon \rangle$
$\langle T \to F \cdot * T,\ 0,\ 0,\ F \rangle$
$\langle T \to F \cdot,\ 0,\ 0,\ F \rangle$
$\langle E \to T \cdot + E,\ 0,\ 0,\ T \rangle$
$\langle E \to T \cdot + E,\ 0,\ 0,\ F \rangle$
$\langle E \to T \cdot,\ 0,\ 0,\ T \rangle$
$\langle E \to T \cdot,\ 0,\ 0,\ F \rangle$

PL_1

$\langle F \to a \cdot,\ 0,\ 1,\ \epsilon \rangle$
$\langle E \to T \cdot,\ 0,\ 1,\ \epsilon \rangle$
$\langle E \to T \cdot + E,\ 0,\ 1,\ \epsilon \rangle$
$\langle T \to F \cdot,\ 0,\ 1,\ \epsilon \rangle$
$\langle T \to F \cdot * T,\ 0,\ 1,\ \epsilon \rangle$

PL_2

$\langle E \to T + \cdot E,\ 0,\ 2,\ \epsilon \rangle$
$\langle E \to \cdot T,\ 2,\ 2,\ \epsilon \rangle$
$\langle E \to \cdot T + E,\ 2,\ 2,\ \epsilon \rangle$
$\langle F \to \cdot [E],\ 2,\ 2,\ \epsilon \rangle$
$\langle F \to \cdot a,\ 2,\ 2,\ \epsilon \rangle$
$\langle T \to \cdot F,\ 2,\ 2,\ \epsilon \rangle$
$\langle T \to \cdot F * T,\ 2,\ 2,\ \epsilon \rangle$
$\langle E \to T + E \cdot,\ 0,\ 2,\ E \rangle$
$\langle E \to T + E \cdot,\ 0,\ 2,\ F \rangle$
$\langle E \to T + E \cdot,\ 0,\ 2,\ T \rangle$
$\langle T \to F \cdot * T,\ 2,\ 2,\ F \rangle$
$\langle T \to F \cdot,\ 2,\ 2,\ F \rangle$
$\langle E \to T \cdot + E,\ 2,\ 2,\ T \rangle$
$\langle E \to T \cdot + E,\ 2,\ 2,\ F \rangle$
$\langle E \to T \cdot,\ 2,\ 2,\ T \rangle$
$\langle E \to T \cdot,\ 2,\ 2,\ F \rangle$

PL_3

$\langle F \to a \cdot,\ 2,\ 3,\ \epsilon \rangle$
$\langle E \to T + E \cdot,\ 0,\ 3,\ \epsilon \rangle$
$\langle E \to T \cdot,\ 2,\ 3,\ \epsilon \rangle$
$\langle E \to T \cdot + E,\ 2,\ 3,\ \epsilon \rangle$
$\langle T \to F \cdot,\ 2,\ 3,\ \epsilon \rangle$
$\langle T \to F \cdot * T,\ 2,\ 3,\ \epsilon \rangle$

PL_4

$\langle T \to F * \cdot T,\ 2,\ 4,\ \epsilon \rangle$
$\langle T \to \cdot F,\ 4,\ 4,\ \epsilon \rangle$
$\langle T \to \cdot F * T,\ 4,\ 4,\ \epsilon \rangle$
$\langle F \to \cdot [E],\ 4,\ 4,\ \epsilon \rangle$
$\langle F \to \cdot a,\ 4,\ 4,\ \epsilon \rangle$
$\langle T \to F * T \cdot,\ 2,\ 4,\ T \rangle$
$\langle T \to F * T \cdot,\ 2,\ 4,\ F \rangle$
$\langle T \to F \cdot * T,\ 4,\ 4,\ F \rangle$
$\langle T \to F \cdot,\ 4,\ 4,\ F \rangle$
$\langle E \to T + E \cdot,\ 0,\ 4,\ F \rangle$
$\langle E \to T + E \cdot,\ 0,\ 4,\ T \rangle$
$\langle E \to T \cdot,\ 2,\ 4,\ F \rangle$
$\langle E \to T \cdot + E,\ 2,\ 4,\ F \rangle$
$\langle E \to T \cdot,\ 2,\ 4,\ T \rangle$
$\langle E \to T \cdot + E,\ 2,\ 4,\ T \rangle$

PL_5

$\langle F \to a \cdot,\ 4,\ 5,\ \epsilon \rangle$
$\langle E \to T + E \cdot,\ 0,\ 5,\ \epsilon \rangle$
$\langle E \to T \cdot + E,\ 2,\ 5,\ \epsilon \rangle$
$\langle E \to T \cdot,\ 2,\ 5,\ \epsilon \rangle$
$\langle T \to F * T \cdot,\ 2,\ 5,\ \epsilon \rangle$
$\langle T \to F \cdot,\ 4,\ 5,\ \epsilon \rangle$
$\langle T \to F \cdot * T,\ 4,\ 5,\ \epsilon \rangle$

文法 G, 入力文字列 w, $\lvert w \rvert = n$	gap ありアイテム数	gap なしアイテム数
G_1: $E \to T+E \mid T,$ $T \to F*T \mid F,\quad F \to a \mid [E]$ w: $a(*[a+a])^m$ $(n = 6m+1)$	$m^2+28m+6$	$m^2/2+75m/2+11$
G_2: $S \to Ab,\quad A \to Ab \mid a$ w: ab^m $(n = m+1)$	2	$4n+2$
G_3: $S \to aB,\quad B \to aB \mid b$ w: $a^m b$ $(n = m+1)$	$n^2/2 - n/2$	$4n-2$
G_4: $S \to AB,\quad A \to Ab \mid a,$ $B \to bB \mid Bd \mid bc$ w: $ab^m cd$ $(n = m+3)$	n^2-3n+4	$13n-29$
G_5: $S \to SS \mid a$ w: a^m $(n = m)$	n^2+2n+1	n^2+3n+2
G_6: $S \to SSSSS \mid a$ w: a^m $(n = m)$	$5n^2/8+5n/4+5/8$	$5n^2/8+3n+2$
G_7: $S \to SS \mid aSb \mid \epsilon$ w: $(aabb)^m$ $(n = 4m)$	$m^2+14m+2$	$m^2+30m+5$
G_8: $S \to AB \mid DC,\quad A \to aA \mid \epsilon,$ $B \to bBc \mid \epsilon,\quad C \to cC \mid \epsilon,$ $D \to aDb \mid \epsilon$ w: $a^m b^m c^m$ $(n = 3m)$	m^2+6m+5	$m^2+24m+17$
G_9: $S \to NP\ VP \mid S\ PP,$ $NP \to NP\ PP \mid det\ n \mid n,$ $VP \to v\ NP,\quad PP \to p\ NP$ w: $n\ v\ det\ n(p\ det\ n)^m$ $(n = 3m+4)$	$3m^2+14m+18$	$3m^2/2+27m/2+21$

表 1: 解析実験結果

5 おわりに

Earley 法に基づく一般の文脈自由文法に対する並列構文解析法を与えた．このアルゴリズムは入力に対して left-to-right に入力に同期して動作し，$O(n^2)$ 個のプロセッサを用いて $O(n)$ 時間で解析表を作成する．これは，実行時間に関しては峯 [6] の手法と同等な能率である．また使用するプロセッサ数に関しては，常に $O(n^2)$ のプロセッサを必要とする点において峯 [6] の手法よりやや劣るが，本手法は入力に対して同期した解析を行なうことが可能であるため，確率文法などのように文法規則に優先度を持つ文法や，ヒューリスティクスを利用した解析などに対して有効に働くものと考えられる．今後，このアルゴリズムを実際の並列処理マシン上に実現する必要がある．

参考文献

[1] D. H. Younger : "Recognition and Parsing of Context-Free Languages in Time n^3", Information and Control, Vol. 14, 1967.

[2] J. Earley : "An Efficient Context-free Parsing Algorithm", Communications of the ACM, Vol. 13, No. 2, pp. 94–102, 1970.

[3] W. Rytter : "Parallel Time $O(\log n)$ Recognition of Unambiguous Context-free Languages", Information and Computation 73, pp.75–86, 1987.

[4] M. Tomita : "An Efficient Augmented-Context-Free Parsing Algorithm", Computational Linguistics, Vol.13, No. 1-2, pp. 31–46, 1987.

[5] 中村, 日高 : "文脈自由文法の並列構文解析法", 情報処理学会第 39 回全国大会, 4F-1, 1989.

[6] 峯, 谷口, 雨宮 : "一般の文脈自由文法に対する効率的な並列構文解析", 情報処理学会論文誌, Vol. 32, No. 10, pp. 1225–1237, 1991.

[7] 沼崎, 田中 : "Incomplete Stack を用いた並列一般化 LR パーザ PGLR について", 情報処理学会論文誌, Vol. 33, No. 1, pp. 18–27, 1992.

付録 A

アルゴリズム A1の正当性示す定理を以下に証明する．

【記法】 通常用いられている導出の記法を拡張して，導出において現在着目している記号位置を明示するために，着目している記号の直前に $*$ を挿入した文形式を考える．そして，生成規則は $*$ の直後の非終端記号にのみ適用可能であるとする．さらに，$*$ を右にシフトする規則 sh を新たに導入する．

$$X_1X_2\cdots X_k * X_{k+1}X_{k+2}\cdots X_m \stackrel{sh}{\Longrightarrow} X_1X_2\cdots X_kX_{k+1} * X_{k+2}\cdots X_m$$

また，導出 $*S \stackrel{*}{\Longrightarrow} [0,i] * X\delta \Longrightarrow [0,i] * \alpha\beta\delta \stackrel{*}{\Longrightarrow} [0,j]Y * \beta\delta$ を，$d = (\,X,\ \alpha,\ \beta,\ \delta,\ Y,\ i,\ j\,)$ と表わす． ■

【定理 2】

$$\langle\, X \to \alpha\cdot\beta,\ i,\ j,\ Y\,\rangle \in PL \iff \exists\delta,\ *S \stackrel{*}{\Longrightarrow} [0,i] * X\delta \Longrightarrow [0,i] * \alpha\beta\delta \stackrel{*}{\Longrightarrow} [0,j]Y * \beta\delta$$

■

(I) ($\Longrightarrow$ の証明)

PL にこれまでに登録されているアイテム数に関する帰納法を用いる．

<u>basis:</u> $\langle\, \phi \to \cdot S,\ 0,\ 0,\ \epsilon\,\rangle$ に対しては $*S \stackrel{*}{\Longrightarrow} [0,0] * \phi\delta \Longrightarrow *S \stackrel{*}{\Longrightarrow} *S$ なので成立する．

<u>inductive step:</u> PL_j に登録された $\langle\, X \to \alpha\cdot\beta,\ i,\ j,\ Y\,\rangle$ の形式のアイテムに対して，

(1) step-1 で登録されたアイテム $\langle\, X \to \alpha' a_j \cdot\beta,\ i,\ j,\ \epsilon\,\rangle$ に対して $\langle\, X \to \alpha'\cdot a_j\beta,\ i,\ j-1,\ \epsilon\,\rangle$ なるアイテムが既に PL_{j-1} に登録されている．よって，以下の導出が存在するので成立する．

$$\overbrace{*S \stackrel{*}{\Longrightarrow} [0,i]*X\delta \Longrightarrow [0,i]*\alpha' a_j\beta\delta \stackrel{*}{\Longrightarrow} [0,j-1]*a_j\beta\delta}^{(X,\ \alpha',\ a_j\beta,\ \delta,\ \epsilon,\ i,\ j-1)} \underbrace{\stackrel{sh}{\Longrightarrow} [0,j]*\beta\delta}$$

$$\underbrace{*S \stackrel{*}{\Longrightarrow} [0,i]*X\delta \Longrightarrow [0,i]*\alpha' a_j\beta\delta \stackrel{*}{\Longrightarrow} [0,j-1]*a_j\beta\delta \stackrel{sh}{\Longrightarrow} [0,j]*\beta\delta}_{(X,\ \alpha' a_j,\ \beta,\ \delta,\ \epsilon,\ i,\ j)}$$

(2) **step-2** で登録されたアイテム $\langle X \to \alpha\cdot\beta,\ i,\ j,\ \epsilon \rangle$ に対して,

$$\langle X \to \alpha\cdot\beta,\ i,\ k,\ W \rangle \in PL_k,\quad \langle W \to \gamma\cdot,\ k,\ j,\ \epsilon \rangle \in PL_j$$

なるアイテムが存在する. ここで, $i \le k < j$, $\alpha \ne \epsilon$ である. したがって,

$$*S \stackrel{*}{\Longrightarrow} [0,i]*X\delta \Longrightarrow [0,i]*\alpha\beta\delta \stackrel{*}{\Longrightarrow} [0,i][i,k]*W\beta\delta \stackrel{sh}{\Longrightarrow} [0,i][i,k]W*\beta\delta$$
$$*S \stackrel{*}{\Longrightarrow} [0,k]*W\eta \Longrightarrow [0,k]*\gamma\eta \stackrel{*}{\Longrightarrow} [0,j]*\eta$$

なる導出が存在する. よって, 導出 $(X,\ \alpha,\ \beta,\ \delta,\ \epsilon,\ i,\ j)$, すなわち

$$*S \stackrel{*}{\Longrightarrow} [0,i]*X\delta \Longrightarrow [0,i]*\alpha\beta\delta \stackrel{*}{\Longrightarrow} [0,k]*W\beta\delta \Longrightarrow [0,k]*\gamma\beta\delta \stackrel{*}{\Longrightarrow} [0,k][k,j]*\beta\delta$$

が存在する.

(3) **step-3** で登録されたアイテム $\langle X \to \cdot\beta,\ j,\ j,\ \epsilon \rangle$ に対して,

$$\langle Z \to \gamma\cdot Y\delta,\ k,\ j,\ \epsilon \rangle \in PL_j,\quad Y \stackrel{*}{\Longrightarrow} X\eta,\quad X \to \beta \in p$$

なるアイテムが存在し, したがって,

$$\underbrace{*S \stackrel{*}{\Longrightarrow} [0,k]*Z\xi \Longrightarrow [0,k]*\gamma Y\delta\xi \stackrel{*}{\Longrightarrow} [0,k][k,j]*Y\delta\xi}_{(Z,\ \gamma,\ Y\delta,\ \xi,\ \epsilon,\ k,\ j)} \stackrel{*}{\Longrightarrow} [0,j]*X\eta\delta\xi \Longrightarrow [0,j]*\beta\eta\delta\xi$$

(4) **step-4** で登録されたアイテム $\langle X \to \alpha' Z\cdot\beta,\ i,\ j,\ Y \rangle$ に対して,

$$\langle X \to \alpha'\cdot Z\beta,\ i,\ j,\ \epsilon \rangle \in PL_j,\quad Z \stackrel{*}{\Longrightarrow} Y$$

なるアイテム (ここで, $i<j$) が存在し,

$$*S \stackrel{*}{\Longrightarrow} [0,i]*X\delta \Longrightarrow [0,i]\alpha' Z\beta\delta \stackrel{*}{\Longrightarrow} [0,i][i,j]*Z\beta\delta \stackrel{*}{\Longrightarrow} [0,j]*Y\beta\delta \stackrel{sh}{\Longrightarrow} [0,j]Y*\beta\delta$$

(5) **step-5** で登録したアイテム $\langle X \to \alpha\cdot\beta,\ i,\ j,\ Y \rangle$ について

$$\langle X \to \alpha\cdot\beta,\ i,\ k,\ w \rangle \in PL_k,\quad \langle w \to \gamma\cdot,\ k,\ j,\ Y \rangle \in PL_j$$

が存在する. ここで, $i \le k < j$ である. したがって,

$$*S \stackrel{*}{\Longrightarrow} [0,i]*X\delta \Longrightarrow [0,i]*\alpha\beta\delta \stackrel{*}{\Longrightarrow} [0,i][i,k]*w\beta\delta \stackrel{sh}{\Longrightarrow} [0,i][i,k]w*\beta\delta$$
$$*S \stackrel{*}{\Longrightarrow} [0,k]*w\eta \Longrightarrow [0,k]*\gamma\eta \stackrel{*}{\Longrightarrow} [0,j]Y*\eta$$

なる導出が存在する. よって,

$$*S \stackrel{*}{\Longrightarrow} [0,i]*X\delta \Longrightarrow [0,i]*\alpha\beta\delta \stackrel{*}{\Longrightarrow} [0,k]*w\beta\delta \Longrightarrow [0,k]*\gamma\beta\delta \stackrel{*}{\Longrightarrow} [0,k][k,j]Y*\beta\delta$$

が存在する. 以上より (I) は証明された.

(II) ($\Longleftarrow$ の証明)

導出 $d = (X,\ \alpha,\ \beta,\ \delta,\ Y,\ i,\ j)$ のランク $r(d)$ を以下のように定め, $r(d)$ に関する帰納法を用いて証明する.

$$\begin{cases} r(d) = \langle r_1(d),\ r_2(d)\rangle,\quad r_1(d) = D_1(d) + D_2(d),\quad r_2(d) = D_2(d) \\ D_1(d):\ \text{導出}\ *S \stackrel{*}{\Longrightarrow} [0,i]*X\delta\ \text{の長さ},\quad D_2(d):\ \text{導出}\ *X \Longrightarrow *\alpha\beta \stackrel{*}{\Longrightarrow} [i,j]*\beta\ \text{の長さ} \\ r(d) < r(d') \iff r_1(d) < r_1(d')\ \text{または}\ r_1(d) = r_1(d'),\ r_2(d) < r_2(d') \end{cases}$$

<u>basis</u>: $r(d)$ が最小の導出 $d = (\phi,\ \epsilon,\ S,\ \epsilon,\ \epsilon,\ 0,\ 0)$ については $\langle \phi \to \cdot S,\ 0,\ 0,\ \epsilon \rangle \in PL_0$ であるので成立.

<u>inductive step</u>: 導出 $d = (X,\ \alpha,\ \beta,\ \delta,\ Y,\ i,\ j)$ に対して,

(1) $\alpha = \epsilon$ の場合 ($Y = \epsilon,\ i = j$)

$$d:\quad \underline{*S \overset{*}{\Longrightarrow} [0,j]*X\delta} \Longrightarrow [0,j]*\beta\delta$$

であるが，d の前半 (下線部) が

$$d':\quad *S \overset{*}{\Longrightarrow} [0,k]*Z\eta \Longrightarrow [0,k]*\gamma Y\xi\eta \overset{*}{\Longrightarrow} [0,k][k,j]*Y\xi\eta \overset{*}{\Longrightarrow} [0,j]*X\overbrace{\zeta\xi\eta}^{\delta}$$

となっている導出 $d'=(Z,\ \gamma,\ Y\xi,\ \eta,\ \epsilon,\ k,\ j)$ が存在する．$r(d')<r(d)$ であるので，d' に対しては，$\langle Z\to\gamma\cdot Y\xi,\ k,\ j,\ \epsilon\rangle\in PL_j$ なるアイテムが既に登録されており，また $Y\overset{*}{\Longrightarrow}X\zeta$ であるので，**step-3** によって，規則 $X\to\beta\in P$ に対して $\langle X\to\beta,\ j,\ j,\ \epsilon\rangle$ が PL_j に登録される．

(2)　$\alpha\neq\epsilon$ の場合

(i)　$\alpha=\alpha' a_j$ のとき　$(Y=\epsilon)$

導出 $d=(X,\ \alpha' a_j,\ \beta,\ \delta,\ \epsilon,\ i,\ j)$ は

$$\underline{*S \overset{*}{\Longrightarrow} [0,i]*X\delta \Longrightarrow [0,i]*\alpha' a_j\beta\delta \overset{*}{\Longrightarrow} [0,i][i,j-1]*a_j\beta\delta} \overset{sh}{\Longrightarrow} [0,i][i,j]*\beta\delta$$

であるので，前半 (下線部) に対応する導出 d' は $d'=(X,\ \alpha',\ a_j\beta,\ \delta,\ \epsilon,\ i,\ j-1)$ となる．
$r(d')<r(d)$ であるので，d' に対しては，$\langle X\to\alpha'\cdot a_j\beta,\ i,\ j-1,\ \epsilon\rangle\in PL_{j-1}$ であり，**step-1** によって $\langle X\to\alpha' a_j\cdot\beta,\ i,\ j,\ \epsilon\rangle$ が PL_j に登録される．

(ii)　$\alpha=\alpha' Z\ (Z\in N)$ のとき

(a)　$Y=\epsilon$ のとき．　導出 $d=(X,\ \alpha' Z,\ \beta,\ \delta,\ \epsilon,\ i,\ j)$ は

$$d:\quad *S \overset{*}{\Longrightarrow} [0,i]*X\delta \Longrightarrow [0,i]*\alpha' Z\beta\delta \overset{*}{\Longrightarrow} [0,i][i,j]*\beta\delta$$

また，$\alpha' Z\overset{*}{\Longrightarrow}[i,j]$ であり ϵ-free であるので $i<j$．したがって，$*\alpha' Z\overset{*}{\Longrightarrow}[i,j]*$ の導出において

$$*\alpha' Z \overset{*}{\Longrightarrow} [i,k]*W \Longrightarrow [i,k]*\underbrace{\gamma' a_j}_{\gamma} \overset{*}{\Longrightarrow} [i,k][k,j-1]*a_j \overset{sh}{\Longrightarrow} [i,j-1]a_j*$$

となる導出が存在する．ゆえに，

$$d':\quad *S \overset{*}{\Longrightarrow} [0,i]*X\delta \Longrightarrow [0,i]*\alpha' Z\beta\delta \overset{*}{\Longrightarrow} [0,i][i,k]*W\beta\delta \overset{sh}{\Longrightarrow} [0,i][i,k]W*\beta\delta$$
$$d'':\quad *S \overset{*}{\Longrightarrow} [0,k]*W\beta\delta \Longrightarrow [0,k]*\gamma\beta\delta \overset{*}{\Longrightarrow} [0,k][k,j]*\beta\delta$$

$r(d')<r(d),\ r(d'')<r(d)$ なので，$\langle X\to\alpha' Z\cdot\beta,\ i,\ k,\ W\rangle\in PL_k,\quad \langle W\to\gamma\cdot,\ k,\ j,\ \epsilon\rangle\in PL_j$．
step-2 によって $\langle X\to\alpha' Z\cdot\beta,\ i,\ j,\ \epsilon\rangle$ が PL_j に登録される．

(b)　$Y\in N$ のとき．　$d=(X,\ \alpha' Z,\ \beta,\ \delta,\ Y,\ i,\ j)$ であるので，導出は次のようになる．

$$d:\quad *S \overset{*}{\Longrightarrow} [0,i]*X\delta \Longrightarrow [0,i]*\alpha' Z\beta\delta \overset{*}{\Longrightarrow} [0,i][i,j]*Y\beta\delta \overset{sh}{\Longrightarrow} [0,i][i,j]Y*\beta\delta$$

ここで，$*\alpha'\overset{*}{\Longrightarrow}[i,k]*$ において (イ) $k<j$，(ロ) $k=j$ の場合について考える．

(イ)　$k=j$ のときは，$*\alpha' Z\overset{*}{\Longrightarrow}[i,j]*Z\overset{*}{\Longrightarrow}[i,j]*Y$ となっている (すなわち，$*Z\overset{*}{\Longrightarrow}*Y$)．ゆえに，

$$d:\quad \underbrace{*S \overset{*}{\Longrightarrow} [0,i]*X\delta \Longrightarrow [0,i]*\alpha' Z\beta\delta \overset{*}{\Longrightarrow} [0,i][i,j]*Z\beta\delta}_{d'=(X,\ \alpha',\ Z\beta,\ \delta,\ \epsilon,\ i,\ j)} \overset{*}{\Longrightarrow} [0,i][i,j]*Y\beta\delta \overset{sh}{\Longrightarrow} [0,j]Y*\beta\delta$$

$r(d')<r(d)$ なので $\langle X\to\alpha'\cdot Z\beta,\ i,\ j,\ \epsilon\rangle\in PL_j$．よって，**step-4** において $Z\overset{*}{\Longrightarrow}Y$ なる Y に対して $\langle X\to\alpha' Z\cdot\beta,\ i,\ j,\ Y\rangle$ が PL_j に登録される．

(ロ)　$k<j$ のときは，$*\alpha' Z\overset{*}{\Longrightarrow}[i,k]*Z\overset{*}{\Longrightarrow}[i,k]*W\Longrightarrow[i,k]*\gamma\overset{*}{\Longrightarrow}[i,k][k,j]*Y$ となる W が存在する．ゆえに，

$$d':\quad *S \overset{*}{\Longrightarrow} [0,i]*X\delta \Longrightarrow [0,i]*\alpha' Z\beta\delta \overset{*}{\Longrightarrow} [0,i][i,k]*Z\beta\delta \overset{*}{\Longrightarrow} [0,k]*W\beta\delta \overset{sh}{\Longrightarrow} [0,k]W*\beta\delta$$
$$d'':\quad *S \overset{*}{\Longrightarrow} [0,k]*W\beta\delta \Longrightarrow [0,k]*\gamma\beta\delta \overset{*}{\Longrightarrow} [0,k][k,j]*Y\beta\delta \overset{sh}{\Longrightarrow} [0,j]Y*\beta\delta$$

$r(d')<r(d),\ r(d'')<r(d)$ なので，$\langle X\to\alpha' Z\cdot\beta,\ i,\ k,\ W\rangle\in PL_k,\quad \langle W\to\gamma\cdot,\ k,\ j,\ Y\rangle\in PL_j$．
step-5 によって $\langle X\to\alpha' Z\cdot\beta,\ i,\ j,\ Y\rangle$ が PL_j に登録される．
以上より (II) が証明された．