LL 문법으로의 커버링 변환의 단순화
(Simplification of Covering Transformation into LL Grammars)

이 경 옥 †
(Gyung-Ok Lee)

요 약 확장된 PLR 문법은 LR 문법의 부분 클래스로서 LL 문법에 대한 커버링 성질을 만족하는 변환이 존재하는 현재까지의 가장 큰 문법 클래스이다. 본 논문에서는 확장된 PLR 문법에 대한 기존 커버링 문법 변환을 단순화시키는 커버링 문법 변환을 제시한다. 본 논문에서 제시하는 커버링 문법 변환은 기존의 4개의 규칙 유형을 3개의 규칙 유형으로 축소시킨다.
키워드 : *LR 문법, LL 문법, 확장된 PLR 문법, 커버링 성질*

Abstract Extended PLR grammars are currently the largest subclass of LR grammars whose grammars are transformed into LL grammars satisfying covering property. This paper suggests a simplified covering transformation of the original covering transformation for extended PLR grammars. The proposed covering transformation reduces the original four rule types to the three rule types.
Key words : LR grammar, LL grammar, extended PLR grammar, covering property

1. 서 론

LR 파싱[1]은 대표적인 상향식 구문 분석 방법이며 이에 상응하는 LL 파싱[2]은 대표적인 하향식 방법이다. 이들 파싱을 적용할 수 있는 문법 클래스인 LR 문법은 길이 k의 미리보기(lookahead) 스트링을 이용한 결정적 파싱이 가능한 가장 큰 범위의 문법 클래스이며, LL 문법은 LR 문법의 부분 클래스이지만 개념적으로 단순하며 구현하기 쉽고 직관적인 LL 파싱이 가능한 문법 클래스이다. 이에 LR 문법에 대한 LL 문법으로의 변환 방법에 관한 연구가 진행되었으며, 특히 커버링 성질을 갖는 변환[3-5]은 원래의 문법 구조를 복구할 수 있는 장점이 있다.

확장된 PLR 문법[5]은 LR 문법의 부분 클래스로서 LL 문법에 대한 커버링 변환이 존재하는 현재까지 알려진 가장 큰 문법 클래스이다. 기존에 제시된 확장된 PLR 문법에 적용되었던 커버링 변환은 4개 규칙 유형으로 표현되었다. 본 논문에서는 기존 변환의 규칙 유형의 성질을 분석하여서 기존 4개 유형에서 3개 유형으로 줄일 수 있음을 보인다. 이로부터 기존 방법에 비해서 단순한 형태의 커버링 성질을 갖는 변환 문법을 생성할 수가 있다.

2장에서는 기본 정의와 표기법에 관해 언급하고, 3장에서는 기존 변환 문법의 규칙 유형의 성질을 분석한다. 4장에서는 단순화된 커버링 변환을 제시하고, 5장에서 결론을 맺는다.

2. 기본 정의와 표기법

본 논문의 기본적인 정의와 표기법은 관련 연구[5-7]를 따른다. 문법 G = (N, Σ, P, S)로 정의되며, 여기서 N은 비단말 심볼들의 집합, Σ은 단말 심볼들의 집합, P은 문법 규칙들의 집합, S는 시작 심볼을 나타낸다. k는 임의의 양수를 표기한다. 본 논문은 문법 G에 규칙 $S' \to S\$^k$을 추가한 문법을 대상으로 하며, G가 LR(k) 문법임을 가정한다. $A(\in N)$에 대해 G^A는 A를 시작 심볼로 하는 축소된(reduced) 문법[7]을 표기한다.

$\alpha \in V^*(V = N \cup \Sigma)$에 대해서 $k{:}\alpha$, $\alpha{:}k$는 각각 α의 길이 k의 전위 스트링(prefix)과 후위 스트링(suffix)을 표기한다. $L^G(\alpha) = \{x \mid \alpha \Rightarrow^* x,\ x \in \Sigma^*\}$, $L(G) = L^G(S)$, $FIRST_k^G(\alpha) = \{k{:}x \mid \alpha \Rightarrow^* x,\ x \in \Sigma^*\}$, $FOLLOW_k^G(\alpha) = \{k{:}x \mid S \Rightarrow^* \beta\alpha x,\ x \in \Sigma^*\}$, $RC_k^G(\alpha)$

---

† 종신회원 : 한신대학교 정보통신학과 교수
golee@hs.ac.kr
논문접수 : 2010년 6월 3일
심사완료 : 2010년 10월 20일

$= \{k{:}xz \mid S \Rightarrow^*_{rm} \beta Bz \Rightarrow_{rm} \beta\gamma\delta z \Rightarrow^*_{rm} \beta\gamma xz,\ \alpha = \beta\gamma,\ xz \in \Sigma^*\}$이다. 이들의 표기 방법에서 G가 문맥상 명확할 경우는 이를 생략한다.

$A \in N$, $R \subseteq FOLLOW_k(A)$라고 하자. 제한적 문법 유도를 나타내는 $\Rightarrow_{A,R}$은 다음과 같이 정의된다: $Ar \Rightarrow_{rm}^* \gamma Bzr \Rightarrow_{rm}^p \gamma X\delta zr (r \in R)$이 존재한다고 하자. $\gamma = \varepsilon$, $X = A$인 경우에는 $FIRST_k(\delta zr) \cap R = \emptyset$의 조건을 만족 시에 $\gamma Bzr \Rightarrow_{A,R}^p \gamma X\delta zr$이고, $\gamma = \varepsilon$, $X = A$가 아닌 경우에는 앞서의 조건에 무관하게 $\gamma Bzr \Rightarrow_{A,R}^p \gamma X\delta zr$이다. $RC_k^{A,R}(\alpha) = \{k{:}yzr \mid Ar \Rightarrow_{A,R}^* \beta Bzr \Rightarrow_{A,R} \beta\gamma\delta zr \Rightarrow_{A,R}^* \beta\gamma yzr,\ \alpha = \beta\gamma,\ r \in R,\ yz \in \Sigma^*\}$로 정의한다.

정의 2.1 문법 $G_1 = (N_1, \Sigma_1, P_1, S_1)$, $G_2 = (N_2, \Sigma_2, P_2, S_2)$에 대해서 $L(G_1) = L(G_2)$이며 h는 P_2에서 P_1^*로의 준동형사상(homomorphism)이라 하자. 다음의 조건을 만족하면 G_2는 G_1을 h의 관점에서 *좌에서 우로 커버(cover)*한다고 한다.

(i) G_2상에 $S_2 \Rightarrow_{lm}^{\pi} w$가 존재하면 G_1상에 $S_1 \Rightarrow_{rm}^{h(\pi)} w$가 성립한다.

(ii) G_1상에 $S_1 \Rightarrow_{rm}^{\pi'} w$가 존재하면 G_2상에 $S_2 \Rightarrow_{lm}^{\pi} w$, $h(\pi) = \pi'$가 성립한다. □

정의 2.1 (i),(ii)에서 rm이 lm으로 대치되면 G_2는 G_1을 h의 관점에서 *좌에서 좌로 커버*한다고 한다.

정의 2.2 [5] G가 확장된 PLR 문법이기 위한 필요충분 조건으로는 다음의 (문장 1)이 성립한다.

(문장 1) G의 생존 가능 스트링(viable prefix)인 $\alpha(|\alpha| > n)$, $L^G(\alpha)$의 원소인 x, $RC_k^G(\alpha)$의 원소인 v에 대해서 아래의 (조건 1)을 만족하는 B, W, $\gamma(\gamma \neq \varepsilon, |\gamma| \leq n, \alpha = \beta\gamma)$가 존재하도록 하는 상수 n이 존재한다.

(조건 1) $S' \Rightarrow^{\pi}_{rm} \beta\gamma z \Rightarrow_{rm}^* xz$, $k{:}z=v$가 G상에 존재할 때 마다 $S' \Rightarrow_{rm}^{\pi_1} \beta Bz''$, $Bw \Rightarrow^{\pi_2}_{B,W} \gamma z'w$, $w = k{:}z''(w \in W)$, $z'z'' = z$인 π_1와 π_2가 G상에 존재한다. (이때 n은 G의 <u>*확장된 PLR(k) 인덱스*</u>이다.) □

3. 기존 문법 변환의 분석

본 절에서는 독자의 편의를 위해서 기존 연구[5]에서의 변환 방법과 관련된 정의들이 주어지고, 단순화된 커버링 변환을 위한 기존 변환 문법의 규칙 유형의 성질을 분석한다.

$A \in N$, $r, z \in \Sigma^k$, $X \in N \cup \{\varepsilon\}$, $\alpha \in V^*$에 대해 $(A, r)\ d^{\alpha}\ (X, z)$는 $A \rightarrow \alpha X\beta \in P$, $r \in FOLLOW_k(A)$, $z \in FIRST_k(\beta r)$과 동치이다. 정의된 d 관계를 표기하는 d-그래프의 경로는 $(A_0, r_0)\ d^{\alpha_1}\ (A_1, r_1)\ d^{\alpha_2}\ (A_2, r_2) \ldots (A_{n-1}, r_{n-1})\ d^{\alpha_n}\ (A_n, r_n)$의 유한 순서이다. 이때 $A_n \in \Sigma$이거나 $A_n = \varepsilon$이면 *단말(terminal) 경로*라 한다. 경로 $h = (A_0, r_0)\ d^{\alpha_1}\ (A_1, r_1)\ d^{\alpha_2}\ (A_2, r_2) \cdots (A_{n-1}, r_{n-1})\ d^{\alpha_n}\ (A_n, r_n)$의 *$|\beta|$-세그먼트*는 $\alpha_1\alpha_2 \cdots \alpha_m = \beta$, $\alpha_{m+p+1} \cdots \alpha_n = \gamma$, $\beta\gamma = \alpha$, $\alpha_m \neq \varepsilon(m \neq 0)$, $\alpha_{m+p+1} \neq \varepsilon\ (m+p \neq n)$을 만족하는 h의 부분 경로인 $(A_m, r_m)\ d^{\varepsilon}\ (A_{m+1}, r_{m+1}) \cdots (A_{m+p-1}, r_{m+p-1})\ d^{\varepsilon}\ (A_{m+p}, r_{m+p})$이다. 단말 경로 $h = (A_0, r_0)\ d^{\alpha_1}\ (A_1, r_1)\ d^{\alpha_2}\ (A_2, r_2) \ldots (A_{n-1}, r_{n-1})\ d^{\alpha_n}\ (A_n, r_n)$, $A_0 = A$, $r_0 = r$, $\alpha = \alpha_1\alpha_2 \ldots \alpha_n$, $A_n r_n = u$에 대해서 0-세그먼트 내에 첫번째 원소를 제외하고는 (A, r) 원소가 포함되지 않는 경우에 h를 *⟨A, r, α, u⟩-경로*라 한다.

정의 3.1(Π 관계)

$(A, R, \beta\gamma)\ \Pi_u\ (B, W, \gamma)$가 성립하기 위한 필요 충분 조건은 각 $\langle A, r, \beta\gamma, u\rangle$-경로 h, $r \in R$에 대해서 $(A_0, r_0)\ d^{\varepsilon}\ (A_1, r_1) \cdots (A_{n-1}, r_{n-1})\ d^{\varepsilon}\ (A_n, r_n)$을 $|\beta|$-세그먼트라고 하면

(i) (조건 1)을 만족하는 $m(1 \leq m \leq n)$이 존재한다.

(조건 1) $A_m = B$이고 $\cup_h \{r_i \mid A_i = B, 0 \leq i \leq m\} \cap \cup_h \{r_i \mid A_i = B, m+1 \leq i \leq n\} = \emptyset$이다.

(ii) m_s을 (조건 1)을 만족하는 가장 작은 m이라고 하면 $W = \cup_h \{r_{m_s}\}$이다.

(여기서 $\cup_h$의 표기는 모든 $\langle A, r, \beta\gamma, u\rangle$-경로의 포함을 의미한다) □

정의 3.2 (Φ 함수)

A, R, α, u가 주어졌다고 하자. $\Phi(A, R, \alpha, u)$ 함수값은 다음과 같이 정의된다:

(i) $(A, R, \alpha)\ \Pi_u\ (B_i, W_i, \gamma_i)$, $i = 0, \cdots, n$가 성립하는 모든 다른 $(B_0, W_0, \gamma_0), (B_1, W_1, \gamma_1), \cdots, (B_n, W_n, \gamma_n)$에 대해서 $(B_i, W_i, \gamma_i)\ \Pi_u\ (B_{i-1}, W_{i-1}, \gamma_{i-1})$, $i=1, 2, .., n$가 성립한다고 하자. 이때 $\Phi(A, R, \alpha, u) = (B_0, W_0, \gamma_0)$이다.

(ii) 그 밖의 경우에는 $\Phi(A, R, \alpha, u) = \text{undefined}$이다. □

다음은 기존 연구[5]에서 제시되었던 문법 변환을 정리하여 나타낸다.

변환 1 (기존 문법 변환)

G를 확장된 PLR(k) 문법, n을 G의 확장된 PLR(k) 인덱스라고 하자. G의 변환된 문법은 $T(G) = (N_T, \Sigma, P_T, S_T)$이다. 여기서 $S_T = [S, \{\$^k\}, \varepsilon, FIRST_k(S\$^k)]$는 시작 심볼, N_T은 P_T에서의 넌터미널들의 최소 집합, P_T는 다음에서 정의되는 문법 규칙의 최소 집합이다:

$A \in N$, $R(\neq \emptyset) \subseteq FOLLOW_k(A)$, $\alpha(|\alpha| \leq n)$는 G^A의 생존 가능 스트링, $U(\neq \emptyset) \subseteq RC_k^{A,R}(\alpha)$라고 하자. 또한 $Z = \{u \in U \mid \Phi(A, R, \alpha, u) \neq \text{undefined}\}$이다.

<u>*유형1.*</u> $[A, R, \alpha, U] \rightarrow a[A, R, \alpha a, V]$

여기서 V는 공집합이 아닌 집합 $\{v\ (=zb) \in RC_k^{A,R}(\alpha a) \mid az \in U - Z,\ (A, r)\ d^{*\alpha}\ (a, z),\ r \in R\}$이다.

유형2. $[A, R, \alpha, U] \rightarrow [A, R, \beta B, V]$

여기서 $\alpha = \beta\gamma$, V는 공집합이 아닌 집합 $\{v \mid (A, r)\ d^{*\beta}\ (B, v), B \rightarrow \gamma \in P, r \in R, v \in U - Z\}$이다.

유형3. $[A, R, \alpha, U] \rightarrow \varepsilon$

여기서 $R \cap U \neq \emptyset$, $\alpha = A$이다.

유형4. $[A, R, \alpha, U] \rightarrow [B, W, \gamma, V][A, R, \beta B, W]$

여기서 $\alpha = \beta\gamma$, V는 공집합이 아닌 집합 $\{v \mid \Phi(A, R, \alpha, v) = (B, W, \gamma), v \in U\}$이다. □

다음 보조정리들은 변환 문법 T(G)의 성질을 보인다.

보조정리 3.1 $[A, R, \alpha, U] \in N_T$, $A \rightarrow \alpha \in P$라고 하자. 각 $r \in R \cap U$에 대해서 $\Phi(A, R, \alpha, r)$ = undefined이다.

(증명) $A \rightarrow \alpha \in P$이기에 $Ar \Rightarrow_{A,R}^{\pi} \alpha r$이 G상에 존재한다. 여기서의 π에 대해서 $Ar \Rightarrow_{A,R}^{\pi^1} \beta Czr$, $Cw \Rightarrow_{C,Z}^{\pi^2} \zeta yw$, $w = k{:}zr(w \in Z)$, $yz = \varepsilon$, $\beta\zeta = \alpha$, $\pi = \pi^1\pi^2$인 $\pi^1(\neq \varepsilon)$, $\pi^2(\neq \varepsilon)$이 존재하지 않는다. 따라서 $(A, R, \alpha)\ \Pi_r\ (C, Z, \zeta)$인 C, Z, ζ가 존재하지 않으므로 $\Phi(A, R, \alpha, r)$ = undefined이다. □

보조정리 3.2 문법 규칙 $[A, R, \alpha, U] \rightarrow [A, R, \beta B, V] \in P_T$가 존재한다면 (a) $\beta = \varepsilon, B = A$ (b) $V = R \cap U$이다.

(증명)

(a) $\beta \neq \varepsilon$이거나 $B \neq A$라고 하자. $v \in V$를 임의로 선택한다. 이때 T(G)의 생성 규칙에 의해서 $Ar \Rightarrow_{A,R}^{+} \beta Bzr \Rightarrow_{A,R} \beta\gamma zr$, $k{:}zr = v$, $\beta\gamma = \alpha$, $r \in R$이 G상에 존재한다. 이로부터 $(A, R, \alpha)\ \Pi_v\ (B, W, \gamma)$가 성립하기에 $\Phi(A, R, \alpha, v) \neq$ undefined이다. 이것은 $v \in V$에 대한 모순이다.

(b) $V \neq R \cap U$라고 하자. 이때 다음의 두 경우가 존재한다.

b-1. $v \notin R \cap U$인 $v \in V$가 존재한다.

집합 V의 정의에 의해서 $V \subseteq U$이기에 $v \in U$, $v \notin R$이다. 한편 보조정리 3.2(a)에 의해서 $Ar \Rightarrow_{A,R}^{+} Azr \Rightarrow_{A,R} \alpha zr$, $k{:}zr = v$, $r \in R$이 G상에 존재한다. 이때 $\Phi(A, R, \alpha, v) \neq$ undefined이기에 $v \in V$는 모순이다.

b-2. $v \notin V$인 $v \in R \cap U$가 존재한다.

$v \in U$이고 $v \notin V$이기에 $v \in \{u \in U \mid \Phi(A, R, \alpha, u) \neq \text{undefined}\}$이다. 따라서 $\Phi(A, R, \alpha, v) \neq$ undefined이다. 이것은 보조정리 3.1에 모순이다.□

보조정리 3.3 $[A, R, \alpha, U] \Rightarrow_{lm} [A, R, \beta B, V] \Rightarrow_{lm}^{\pi} x$가 T(G)상에 존재한다면 $\pi = [A, R, A, V] \rightarrow \varepsilon$이다.

(증명) $[A, R, \alpha, U] \rightarrow [A, R, \beta B, V] \in P_T$이기에 보조정리 3.2에 의해 $\beta = \varepsilon$, $B = A$, $V = R \cap U$이다. 따라서 $[A, R, A, V]$을 좌측으로 갖는 규칙은 유일하게 $[A, R, A, V] \rightarrow \varepsilon$이다. □

보조정리 3.3으로부터 다음의 정리가 성립한다.

정리 3.1 T(G)의 문법 유도 과정에서 유형2 규칙의 다음으로는 항상 유형3 규칙이 따라온다. □

정리 3.1로부터 유형2 규칙 $[A, R, \alpha, U] \rightarrow [A, R, A, V]$와 유형3 규칙 $[A, R, A, V] \rightarrow \varepsilon$은 항상 같이 유도됨을 알 수가 있고, 이들은 $[A, R, \alpha, U] \rightarrow \varepsilon$으로 결합 가능함을 알 수가 있다.

4. 단순화된 문법 변환

단순화된 문법 변환의 이해를 돕기 위해서 다음에서 T(G)의 유형2 규칙의 성질이 주어진다.

성질 4.1 $[A, R, \alpha, U] \in N_T$, $A \rightarrow \alpha \in P$, $R \cap U \neq \emptyset$이면 $[A, R, \alpha, U] \rightarrow [A, R, A, V]$, $V = R \cap U$가 T(G)상에 존재한다.

(증명) $R \cap U \neq \emptyset$이기에 $r \in R \cap U$인 r이 존재한다. 따라서 $[A, R, \alpha, U] \rightarrow [A, R, A, V]$, $r \in V$가 T(G)상에 존재한다. 한편 보조정리 3.2에 의해서 $\beta = \varepsilon$, $B = A$, $V = R \cap U$이다. 따라서 본 성질이 성립한다. □

변환 2 (**단순화된 커버링 변환**) G를 확장된 PLR(k) 문법, n을 G의 확장된 PLR(k) 인덱스라고 하자. G의 변환된 문법은 $T(G) = (N, \Sigma, P, S)$이다. 여기서 $S = [S, \{\$^k\}, \varepsilon, FIRST_k(S\$^k)]$은 시작 심볼, N은 P에서의 넌터미널들의 최소 집합, P는 다음에서 정의되는 문법 규칙의 최소 집합이다:

$A \in N$, $R(\neq \emptyset) \subseteq FOLLOW_k(A)$, $\alpha(|\alpha| \leq n)$는 G^A의 생존 가능 스트링, $U(\neq \emptyset) \subseteq RC_k^{A,R}(\alpha)$라고 하자. 또한 $Z = \{u \in U \mid \Phi(A, R, \alpha, u) \neq \text{undefined}\}$이다.

유형A. $[A, R, \alpha, U] \rightarrow a[A, R, \alpha a, V]$

여기서 V는 공집합이 아닌 집합 $\{v\ (=zb) \in RC_k^{A,R}(\alpha a) \mid az \in U - Z, (A, r)\ d^{*\alpha}\ (a, z), r \in R\}$이다.

유형B. $[A, R, \alpha, U] \rightarrow \varepsilon$

여기서 $A \rightarrow \alpha \in P$, $R \cap U \neq \emptyset$이다.

유형C. $[A, R, \alpha, U] \rightarrow [B, W, \gamma, V][A, R, \beta B, W]$

여기서 $\alpha = \beta\gamma$, V는 공집합이 아닌 집합 $\{v \mid \Phi(A, R, \alpha, v) = (B, W, \gamma), v \in U\}$이다. □

정의 4.2 g는 P에서 P^*로의 준동형사상으로 다음과 같이 정의된다: $p \in P$라고 하자.

(i) p가 $[A, R, \alpha, U] \rightarrow \varepsilon$이면 $g(p) = [A, R, \alpha, U] \rightarrow [A, R, A, V] \cdot [A, R, A, V] \rightarrow \varepsilon (V = R \cap U)$이다.

(ii) 그 밖의 경우에는 $g(p) = p$이다. □

다음 두 보조정리는 T(G)와 T(G)간의 관계를 보이며, 문법 유도 길이에 관한 인덕션을 사용하여 쉽게 증명 가능하다.

보조정리 4.2 $[A, R, \alpha, U] \in N_T$(여기서 α = A인

경우에는 $R \cap U = \emptyset$이다)일 때, $[A, R, \alpha, U] \Rightarrow_{lm}^{\pi} x$가 T(G)상에 존재하면 $[A, R, \alpha, U] \Rightarrow_{lm}^{\pi'} x$가 T(G)상에 존재하며 $g(\pi') = \pi$가 성립한다. □

보조정리 4.3 $[A, R, \alpha, U] \Rightarrow_{lm}^{\pi'} x$가 T(G)상에 존재한다면 $[A, R, \alpha, U] \Rightarrow_{lm}^{\pi} x$가 T(G)상에 존재하며 $g(\pi') = \pi$가 성립한다. □

보조정리 4.2, 4.3로부터 다음 정리를 얻을 수가 있다.

정리 4.1 T(G)은 T(G)를 g의 관점에서 좌에서 좌로 커버한다. □

정의 4.3 h는 P_T에서 $P \cup \{\varepsilon\}$로의 준동형사상으로 다음과 같이 정의된다: $p_T \in P_T$라고 하자.

(i) $p_T = [A, R, \beta\gamma, U] \rightarrow [A, R, \beta B, V]$이면 $h(p_T) = B \rightarrow \gamma$이다.

(ii) 그 밖의 경우에는 $h(p_T) = \varepsilon$이다. □

정리 4.2 f를 $h \circ g$로 정의하면 T(G)는 G를 f의 관점에서 좌에서 우로 커버한다. (증명) T(G)은 T(G)를 g의 관점에서 좌에서 좌로 커버하고 T(G)는 G를 h의 관점에서 좌에서 우로 커버한다[5]. 따라서 본 정리는 성립한다. □

정리 4.3 T(G)는 LL(k)이다. □

정리 4.3은 기존 연구[3, 정리6.6]와 매우 유사하게 증명될 수가 있기에 증명 과정을 생략한다.

다음의 성질들은 T(G)가 갖는 장점에 관한 이해를 돕기 위해서 보인다.

성질 4.2 T(G)의 문법 규칙의 수는 T(G)의 문법 규칙의 수에 비해서 $O(|P| x 2^{|\Sigma^k|})$에 비례하여 감소된다.

(증명) T(G)의 문법 규칙에서의 유형A, 유형C는 각각 T(G)의 문법 규칙의 유형1, 유형4에 대응되며, 이들 간의 개수는 동일하다. 한편 T(G)의 유형B규칙은 유형2의 개수와 동일하다. 따라서 T(G)의 규칙의 개수는 T(G)의 유형4의 규칙의 개수만큼 감소되며 유형4의 규칙의 개수는 $O(|P| x 2^{|\Sigma^k|})$에 비례한다. □

성질 4.3 $x \in L(G)$에 대해서, $S' \Rightarrow^{\pi}_{rm} x$, $|\pi| = n$이 G상에 존재하며, T(G)상에 $S_T \Rightarrow_{lm}^{\pi'} x$, $|\pi'| = m$이 존재한다고 하자. 그러면 T(G)상에는 $S \Rightarrow_{lm}^{\pi''} x$, $|\pi''| = m-n$이 존재한다.

(증명) 보조정리 4.2에서 보듯이 T(G)상의 스트링 유도 과정와 T(G)상의 유도 과정은 g함수로 표현될 수 있다. g함수의 정의로부터 T(G)상의 유형2 규칙과 유형3 규칙이 T(G)상의 유형B 규칙으로 대치되며, 이 길이는 n의 길이와 일치한다. 따라서 본 성질이 성립한다. □

성질 4.3으로부터 다음의 성질을 얻을 수가 있다. 문법에서의 스트링 유도 과정의 길이는 유도 트리의 내부 노드의 수와 밀접하게 관련됨을 상기하기 바란다.

성질 4.4 $x \in L(G)$에 대해서 G에 대한 유도 트리 노드 수가 $n+|x|$이며, T(G)에 대한 유도 트리 노드 수가 m이라고 하자. 이때 T(G)상의 유도 트리 노드 수는 $m-n$개이다. □

5. 결 론

본 논문에서는 확장된 PLR 문법에 대한 기존 연구[5]에서 제안한 변환 문법의 성질을 분석하여 기존 변환을 단순화한 문법 변환을 제시하였다. 제안된 변환 문법에서 기존 변환 문법으로의 준동형사상이 존재함과 제안된 변환 문법이 원래의 문법을 좌에서 우로 커버링함을 보였다. 본 논문의 결과로서 LR 문법에서 LL 문법으로의 커버링 성질을 갖는 문법 변환이 기존의 4개의 규칙 유형에서 3개의 규칙 유형으로 축소 가능함을 알 수가 있다. 또한 축소된 문법은 기존 문법에 비해서 문법 규칙의 수, 스트링 유도 과정의 길이, 유도 트리의 노드 수를 감소시킴을 알 수 있다. 차후의 연구 과제로는 실용적인 문법에 대한 변환 문법의 구현을 생각해 볼 수가 있다.

참 고 문 헌

[1] D. E. Knuth, "On the translation of languages from left to right," *Information and Control*, vol.8, pp.607-623, 1965.

[2] D. E. Knuth, "Top down Syntax," *Acta Informatica*, vol.1, pp.79-110, 1971.

[3] M. Hammer, "A new grammatical transformation into LL(k) form," *Proc. of Sixth Annual ACM Symposium on Theory of Computing: ACM*, pp. 266-275, 1974.

[4] E. Soisalon-Soininen and E. Ukkonen, "A method for transforming grammars into LL(k) form," *Acta Informatica*, vol.12, pp.338-369, 1979.

[5] G.-O. Lee and K.-M. Choe, "A powerful LL(k) covering transformation," *SIAM J. Computing*, vol.35, no.2, pp.359-377, 2006.

[6] A.V. Aho and J.D. Ullman, *The Theory of Parsing, Translation and Compiling*, vols. 1 & 2, Prentice-Hall, Englewood Cliffs, NJ, 1972, 1973.

[7] S. Sippu and E. Soisalon-Soininen. *Parsing Theory*, vols. I & II, Springer, Berlin, 1990.

이 경 옥

1990년 2월 서강대학교 전자계산학과 졸업(학사). 1992년 8월 한국과학기술원 전산학과 졸업(석사). 2000년 2월 한국과학기술원 전산학과 졸업(박사). 2000년 8월부터 한신대학교 정보통신학과 근무. 현재 한신대학교 정보통신학과 부교수. 관심분야는 프로그래밍 언어와 컴파일러 등