논·문 15-1-4

s-extended PLR(k) 문법
(s-extended PLR(k) grammars)

이경옥
한신대학교 컴퓨터정보통신학부 정보통신학전공
golee@hucc.hanshin.ac.kr

요 약

본 논문에서는 PLR(k) 문법과 비좌순환(non-left recursive) k-transformable 문법을 포함하는 s-extended PLR(k) 문법을 정의하고, 이 문법에 대한 LL 커버링 문법 변환을 제안한다.

1. 서론

Extended PLR(k) 문법 [4]은 k-transformable 문법 [2,3]과 PLR(k) 문법 [6]을 포함하는 현재까지 가장 큰 LL(k) 커버링 문법이 제안된 문법 클래스이다. 한편 extended PLR(k) 문법과 이 문법 클래스에 대해 제안된 LL(k) 커버링 문법 변환을 설명하기 위해서는 보편적인 LR 이론을 수정한 새로운 정형식들에 대한 이해가 필요하다. 이에 본 논문에서는 보편적 LR 이론을 그대로 이용한 LL(k) 커버링 문법 변환을 제안한다. 이 변환에 적용가능한 문법 클래스를 *s-extended PLR*(k) 문법으로 정의한다. 이 문법 클래스에 대해서 [4]와 유사하게 문법 유도 과정에 근거하여 특징화가 보여진다. 이 클래스는 PLR(k) 문법 클래스와 비좌순환(non-left-recursive) k-transformable 문법 클래스를 포함한다.

2. 기본 정의

표기법과 기본 정의들은 [1,5]을 따르며, 독자가 이에 친숙함을 가정한다.

심볼 G는 임의의 고정된

$G=(N,\Sigma, P,S)$ ($V=N\cup\Sigma$)를 나타낸다. 심볼 k는 임의의 고정된 양수를 나타낸다. N에 속하지 않는 S'에 대해서 $G'=(N\cup\{S'\},\Sigma,P\cup\{S'\to S\$^k\},S')$는 G의 **확장된(augmented) 문법**이다. G는 S'에 의해 확장된 문법임을 가정하고 또한 **리듀스된 (reduced) 문법** [5]임을 가정한다. 그리고 G는 어떤 $k(k\geq1)$에 대한 LR(k) 문법임을 가정한다.

생존가능 스트링(viable prefix) α에 대한 **k-우문맥(right context)**은 $RC_k(\alpha)$ $=\{k\,yz \mid S' \Rightarrow^*_{rm} \beta Bz \Rightarrow_{rm} \beta\gamma\delta z \Rightarrow^*_{rm} \beta\gamma yz, \beta\gamma=\alpha\}$이다. G^A는 A를 시작 심볼로 하는 리듀스된 G의 축소 문법을 나타낸다. α가 G^A의 생존가능 스트링인 경우에 $RC_k^{G^A}(\alpha)=\{k\,yzr \mid Ar \Rightarrow^*_{rm} \beta Bzr \Rightarrow_{rm} \beta\gamma zr \Rightarrow^*_{rm} \beta\gamma yzr, \beta\gamma=\alpha, r\in FOLLOW_k(A)\}$로 정의된다. 간략하게 $RC_k^{G^A}$은 RC_k^A으로 표현한다.

3. 변환

이 장에선 LL(k)-유사 성질을 찾기 위한 Π 관계를 정의한 후에 이에 바탕을 둔 LL(k) 커버링 변환 문법을 제안한다.

정의(Π 관계)

$(A,\beta\gamma)\neq(B,\gamma)$일 때 $(A,\beta\gamma)\ \Pi\ (B,\gamma)$가 성립하기 위해서는 다음이 성립한다. 즉, 다음의 (유도과정 1) $Ar \Rightarrow^+_{rm} \beta\gamma wr$,

$$r\in FOLLOW_k(A),\ k\,wr\in RC_k^B(\gamma)$$

이 존재할 때마다 (유도과정 1)은 $Ar \Rightarrow^+_{rm} \beta Bzr \Rightarrow^*_{rm} \beta\gamma yzr,\ yz=w$ 의 형태이다. □

위 관계는 A로부터 $\beta\gamma$가 유도되고 $RC_k^B(\gamma)$에 속하는 스트링이 생성될 경우에 γ가 B로부터 생성됨을 항상 예상할 수 있음을 의미한다.

다음은 Π 관계의 부분 관계 Π_{LL}을 정의한다. 이는 제안될 커버링 문법 변환에서의 LL 충돌을 막기 위함이다.

정의 (Π_{LL} 관계) Π_{LL} 관계는 Π의 부분 관계로서 다음의 조건을 만족한다. 즉 $(A,\alpha)\ \Pi_{LL}\ (B,\gamma)$, $(A,\alpha)\Pi_{LL}\ (C,\delta)$이면 $RC_k^B(\gamma)\cap RC_k^C(\delta)=\varnothing$이다. □

다음에선 Π_{LL}을 이용한 변환을 제시한다.

알고리즘1($T(G,\Pi_{LL})=(N_T,\Sigma,P_T,S_T)$의 생성)

입력: G, Π_{LL}

출력: 이 알고리즘이 성공적으로 끝나면 $T(G,\Pi_{LL})$를 생성한다.

방법:

$S_T=[S',\varepsilon], N_T=\{S_T\}, P_T=\varnothing$으로 초기화한다.

repeat

(a) 각 $[A,\alpha](\alpha\neq\varepsilon)\in N_T$에 대해서 $Z=\{z\in RC_k^B(\gamma) \mid (A,\alpha)\ \Pi_{LL}\ (B,\gamma)\}$로 정의하며 아래의 과정을 수행한다.

(i) *Type1.* $P_T=P_T\cup\{[A,\alpha]\to a[A,\alpha a]\}$ 이때 $V=\{k\,zwr \mid Ar \Rightarrow^*_{rm} \beta Bwr \Rightarrow_{rm} \beta\gamma a\delta wr \Rightarrow^*_{rm} \beta\gamma azwr,\ \beta\gamma=\alpha,$ $r\in FOLLOW_k(A),\ k\,azwr\in Z\}$일 때 $V\neq\varnothing$이다.

(ii) $Type2.$ $P_T = P_T \cup \{[A,\alpha] \to [A,\beta B]\}$ 이 때 $V = \{k:zwr \mid Ar \Rightarrow^*_{rm} \beta Bwr \Rightarrow_{rm} \beta\gamma wr,\ \beta\gamma = \alpha,\ r \in FOLLOW_k(A),\ k:zwr \in Z\}$일 때 $V \neq \varnothing$이다.

(iii) $Type3.$

$P_T = P_T \cup \{[A,\alpha] \to [B,\gamma][A,\beta B]\}$

이때 (A,α) Π_{LL} (B,γ)이 성립해야 한다.

(b) P_T내에 새로이 형성된 넌터미널은 N_T에 더해진다.

until (N_T 집합이 변하지 않는다.)

$Type4.$ $P_T = P_T \cup \{[A,A] \to \varepsilon\}$

이때 $[A,A] \in N_T$이다.

□

주어진 문법 G에 대해서 어떤 Π_{LL}를 선택하는가에 따라서 위 알고리즘은 성공적으로 끝나서 $T(G,\Pi_{LL})$ 문법을 생성하거나, 무한 루프에 빠지거나 한다. 전자의 경우에 G를 **s-extended PLR(*k*) 문법**이라 정의한다.

h를 P_T에서 F로의 함수로 다음과 같이 정의하자. $r = [A,\beta\gamma] \to [A,\beta B]$인 경우에 $h(r) = B \to \gamma$이고 그 밖의 경우엔 $h(r) = \varepsilon$이다.

다음의 두 보조 정리는 G상의 유도 과정과 $T(G,\Pi_{LL})$상의 유도 과정 사이의 관계를 보인다.

보조정리 1. $[A,\alpha] \in N_T$라 하자. $A \Rightarrow^{\pi}_{rm} \alpha x$인 π가 G상에 존재하면 $h(\pi_T) = \pi^R$이고 $[A,\alpha] \Rightarrow^{\pi_T}_{lm} x$인 π_T가 $T(G,\Pi_{LL})$상에 존재한다.

(증명) $|\pi|$에 대한 인덕션에 의해 증명이 가능하며, 자세한 과정은 지면상의 제약으로 생략한다. □

보조정리 2. $[A,\alpha] \Rightarrow^{\pi_T}_{lm} x$인 π_T가 $T(G,\Pi_{LL})$상에 존재하면 $A \Rightarrow^{\pi}_{rm} \alpha x$인 π가 G상에 존재한다.

(증명) $|\pi_T|$의 길이에 대한 인덕션으로 증명이 가능하며, 자세한 과정은 생략한다. □

보조정리 1과 보조정리 2로부터 다음의 정리 1을 얻을 수 있다.

정리 1. $T(G,\Pi_{LL})$는 h의 관점에서 G를 좌측에서 우측으로 커버(left-to-right cover)한다. □

한편 G의 LR(k) 성질과 $T(G,\Pi_{LL})$의 프러덕션 생성 조건에 의해서 $T(G,\Pi_{LL})$는 LL(k) 문법임을 보일 수 있다. 따라서 $T(G,\Pi_{LL})$는 G의 LL(k) 커버링 문법이다.

4. s-extended PLR(*k*) 문법

s-extended PLR(k) 문법 클래스는 다음과 같이 문법 유도를 이용한 특징화가 가능하다.

정리 2. G가 s-extended PLR(k) 문법이기 위한 필요 충분한 조건으론 G에 의해 결정되는 상수 n이 존재하여 다음 (조건 1)을 만족한다. (조건 1): $\alpha, |\alpha| \geq n$가 G의 존속가능 스트링이고 $v \in RC_k(\alpha)$일 때 B,γ $(\alpha = \beta\gamma,\ |\gamma| \leq n)$가 존재하여 $S' \Rightarrow^*_{rm} \beta\gamma z$, $k:z \in RC^B_k(\gamma)$일때마다 $S' \Rightarrow^*_{rm} \beta B z'' \Rightarrow^*_{rm} \beta\gamma z' z''$ $(z'z'' = z)$이 존재한다. □

정리 3. G가 PLR(k) 문법이면 G는 s-extended PLR(k) 문법이다.

(증명) α를 G의 존속가능 스트링이고 $v \in$

$RC_k^G(\alpha)$이라 하자. 우선 $\alpha \neq \varepsilon$인 경우를 생각한다. 이때 유도과정

$S' \Rightarrow^*_{rm} \beta Bx \Rightarrow_{rm} \beta X\delta\zeta x \Rightarrow^*_{rm} \beta X\delta y_2 x \Rightarrow^*_{rm} \beta X y_1 y_2 x$ ($\beta X\delta = \alpha$, $X\delta \neq \varepsilon$, $k\,y_2x = v$)이 존재한다. 한편 임의의 유도과정 $S' \Rightarrow^*_{rm} \beta' Ax' \Rightarrow_{rm} \beta'\beta'' X\delta'\zeta' x' \Rightarrow^*_{rm} \beta'\beta'' X\delta' y'_2 x' \Rightarrow^*_{rm} \beta'\beta'' X y'_1 y'_2 x'$ ($\beta'\beta'' = \beta$, $k\,y'_1 y'_2 x' = k\,y_1 y_2 x$)이 존재한다면 PLR($k$) 문법의 정의에 의해서 $A = B, \beta'' = \varepsilon$이다. (조건 1)을 만족하는 n이 존재함을 보이기 위해서 n을 G의 프러덕션 우변의 최대 길이를 나타내고 $\gamma = X\delta$라고 하자. 이때 앞에서의 B와 γ에 대해서 $S' \Rightarrow^*_{rm} \beta\gamma z$, $k\,z \in RC_k^B(\gamma)$가 존재할 때마다, $S' \Rightarrow^*_{rm} \beta B z_2 \Rightarrow^*_{rm} \beta\gamma z_1 z_2$ ($z_1 z_2 = z$)인 유도과정이 존재함을 알 수 있다. 또한 이때 $|\gamma| \leq n$, $v \in RC_k^B(\gamma)$을 만족시킨다.

$\alpha = \varepsilon$인 경우에도 B와 γ를 S'와 ε으로 각각 할당할 경우에 앞에서의 n에 대해서 $|\gamma| \leq n$이고 (조건 1)이 성립함이 명백하다. 따라서 G는 s-extended PLR(k) 문법이다. □

정리 3과 유사한 방법으로 다음의 정리를 얻을 수 있다.

정리 4. G가 비좌순환 k-transformable 문법이면 G는 s-extended PLR(k) 문법이다. □

정리 3과 정리 4에 의해서 s-extended PLR(k) 문법은 PLR(k) 문법과 비좌순환 k-transformable 문법을 포함한다.

감사의 글

본 논문은 2001년도 한신대학교 학술연구비 지원을 받았습니다.

참고문헌

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

[2] Hammer, M.: A New Grammatical Transformation into deterministic top down form. MIT, Mass., Project MAC Technical Report TR-119, 1974

[3] Hammer, M.: A New Grammatical Transformation into LL(k) form. Proc. of Sixth Annual ACM Symposium on Theory of Computing, 266-275 (1974)

[4] Lee, Gyung-Ok: On LL-to-LR Covering Grammars, Ph.D. Thesis, KAIST (2000)

[5] Sippu, S. and Soisalon-Soininen, E.: Parsing Theory, vol I, II, Springer-Verlag Berlin Heidelberg, 1990

[6] Soisalon-Soininen, E. and Ukkonen, E.: A Method for Transforming Grammars into LL(k) Form. Acta Informatica 12, 338-369 (1979)

이경옥

한국정보과학회 프로그래밍 언어 논문지 제 14권 2호 참조