## Otomata Teorisi Bölüm 4

Bağlamdan Bağımsız Dilbilgisi –Tür 2

## 4.BAĞLAMDAN-BAĞIMSIZ DİLBİLGİSİ

- Programlama dilleri, derleyiciler, yorumlayıcılar, sözdizim çözümleyiciler, aritmetik deyimler, ..vb. birçok yazılım bileşeninin bünyesinde yer aldığı için bağlamdan-bağımsız dilbilgisi ve diller ile bu dilleri tanıyan makine modeli bilgisayar bilimleri ve mühendisliği açısından önem taşır.
- Bağlamdan bağımsız dilbilgisi bir dörtlüdür.

V<sub>N</sub>= Sözdizim değişkenleri kümesi(sonlu bir küme)

V<sub>T</sub>=Uç simgeler kümesi (sonlu bir küme)

 $V_N$  ve  $V_T$  ayrık kümelerdir.  $V_N \Pi V_T = \emptyset$ 

S: Başlangıç değişkeni S ε V<sub>N</sub>

P: Yeniden yazma ya da türetme kuralları

A=>b  $A \in V_N$   $b \in V^*$ 

```
G_{4,1} = \langle V_N, V_T, P, S \rangle
     V_N = \{ S \}
     V_{\tau} = \{+, -, *, /, (, ), v, c \}
     P: S=> S+S | S-S | S*S | S/S | (S) | v | c
G<sub>4 1</sub> tarafından türetilen tümcelerden biri bulunmak istendiğinde:
S=> S*S => S*(S) => S*(S-S) => v*(S-S) => v*(v-S) => v*(v-c)
S=>S*S=>(S)*S=>(S)*(S)=>(S+S)*(S)=>(S+S)*(S-S)=>(S+S)*(S-S/S)
=>(v+S)*(S-S/S) => (v+c)*(S-S/S) => (v+c)*(v-S/S) => (v+c)*(v-v/S)
=> (v+c)*(v-v/c)
S=> S+S => S*S +S => (S)*S+S => (S/S)*S+S => (S/(S))*S+S
\Rightarrow (S/(S-S))*S+S => (v/(S-S))*S+S => (v/(v-S))*S+S=> (v/(v-c))*S+S
\Rightarrow => (v/(v-c))*v+S => (v/(v-c))*v+c
```

```
G<sub>4.2</sub>=<V<sub>N</sub>, V<sub>T</sub>, P, S>
V<sub>N</sub>= { S, A, B }
V<sub>T</sub>={a, b, c, d}
P: S=> aAdd
A=> aAd | Ad | bBcc
B=>bBc | Bc | λ

G<sub>4.2</sub> tarafından türetilen tümcelerden biri bulunmak istendiğinde:
S=> aAdd =>abBccdd =>abccdd
S=> aAdd =>aaAddd => aabBccddd =>aabbccdddd
S=>aAdd =>aaAddd => aaAdddd=> aaabBcccddddd=> aaabBcccdddddd => aaabBccccdddddd => aaabBccccdddddd
```

Yukarıdaki örneklerden bu bağlamdan bağımsız dilbilgisi tarafından  $L(G_{4.2})$  dili aşağıdaki gibi tanımlanabilir:

```
L(G_{4.2}) = \{a^n b^m c^k d^p \mid n, m, p, k \ge 1, p > n, k > m\}
```

```
G_{4.3} = \langle V_N, V_T, P, S \rangle
V_N = \{ S, X, Y, Z \}
V_T = \{ a, b, c \}
P: S = \rangle XY
X = \rangle aXbb | aZbb | abb
Y = \rangle cY | c
Z = \rangle Zb | Xb
```

G<sub>4.3</sub> tarafından türetilen tümcelerden biri bulunmak istendiğinde:

S=> XY => abbY => abbcY =>abbcc

Yukarıdaki örneklerden bu bağlamdan bağımsız dilbilgisi tarafından  $L(G_{4.2})$  dili aşağıdaki gibi tanımlanabilir:

$$L(G_{4.3}) = \{a^n b^m c^k \mid n \ge 1, k \ge 1 \ m \ge 2n\}$$

```
G_{4.4} = \langle V_N, V_T, P, S \rangle
V_N = \{ S, A \}
V_T = \{ 0, 1 \}
P: S = \rangle 0S0 \mid 1S1 \mid 0A0 \mid 1A1
A = \rangle 0A1 \mid 1A0 \mid 01 \mid 10
```

G<sub>4.4</sub> tarafından türetilen tümcelerden biri bulunmak istendiğinde:

```
S=> 0S0 => 01A10 => 010A110 => 01010110

S=> 0S0 => 01S10 => 011S110 => 0110A1110 => 01100A11110 => 011001011110

S=> 0S0 => 00S00 => 000S000 => 0001A1000 => 00010A11000 => 000100A111000 => 0001000A1111000 => 000100A1111000 => 0001000A1111000  => 0001000A1111000 => 0001000A1111000 => 000100A1111000 => 0001000A1111000 => 000100A1111000 => 0001000A1111000 => 0001000A111000 => 0001000A11000 => 0001000A11000 => 000100A11000 > 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 000100A11000 => 0001000A11000 => 0001000A11000 => 0001000A11000 => 0001000A11000 => 0001000A11000 => 0001000A1000 => 0001000A1000 => 00010000A1000 => 00010000 => 00010000 => 00010000 => 00010000 => 000100000
```

Yukarıdaki örneklerden bu bağlamdan bağımsız dilbilgisi tarafından  $L(G_{4.2})$  dili aşağıdaki gibi tanımlanabilir:

```
L(G_{4.3}) = \{uv(v')^R u^R \mid u, v \in (0+1)^*, |u| \ge 1, |v| \ge 1\}
u^R \quad u'nun tersi(reverse)
(v')^R \quad V'nin tümlerinin tersi
```

# 4.2 Türetme Ağacı

 Bir dilbilgisi tarafından türetilen tümcesel yapı ve tümceler aşağıdaki gibi gösterilebilir:

$$S = > \beta_1 = > \beta_2 = > \beta_3 = > \beta_{n-1} = > \beta_n = > w \beta_i \epsilon V^*$$
,  $w \epsilon V_T^*$ 

 Bağlamdan-bağımsız dillerin her tümcesel yapısına ve her tümcesine bir türetme ya da ayrıştırma ağacı (derivation or parsing tree) karşı gelir.

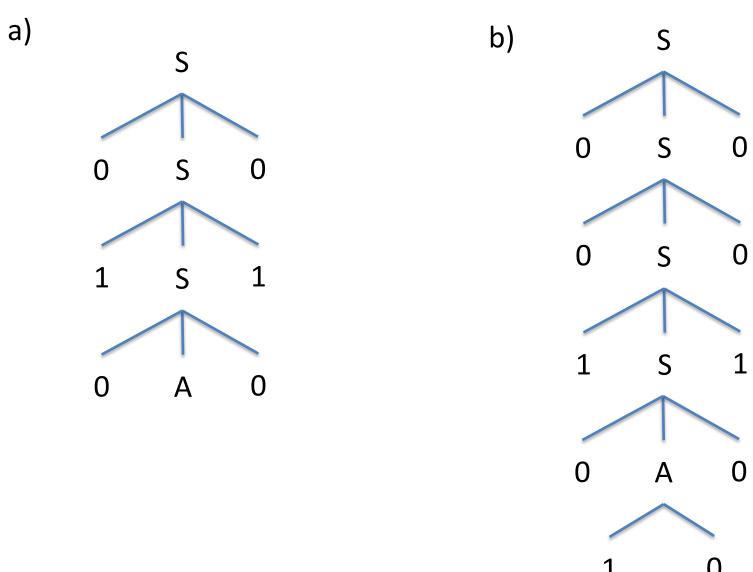
Örneğin **G**<sub>4,4</sub> tarafından türetilen:

 $\beta$  = 010A010 tümcesel yapısı ile

**w** = 0010100100 tümcesine

karşı gelen türetme ağaçları Çizim 4.1'de görülmektedir.

# Türetme Ağacı: Çizim 4.1



## Türetme ya da Ayrıştırma Ağacının Tanımı

- a) Ağacın kökünün etiketi S'dir.
- b) Kök dışındaki ara düğümlerin etiketleri sözdizim değişkenleridir  $(A \epsilon V_N)$ .
- c) Eğer ağaç bir tümcesel yapıya karşı geliyorsa, yaprakların etiketleri sözdizim değişkenleri ya da uç simgeler olabilir (X $\epsilon$ V). Eğer ağaç bir tümceye karşı geliyorsa yaprakların etiketleri yalnız uç simgeler (a $\epsilon$ V $_{\rm T}$ ) olabilir.
- d) Eğer ara bir düğümün etiketi A, bu ara düğümün hemen altındaki düğümlerin etiketleri ise soldan sağa  $X_1$ ,  $X_2$ ,  $X_3$ ,...,  $X_k$  ise, dilbilgisinin yeniden yazma kuralları arasında

$$A = X_1 X_2 X_3 \dots X_k$$

 $X_1, X_2, X_3, ..., X_k \in V$  kuralı yer almalıdır.

e) Eğer bir düğümün etiketi λ ise, bu düğüm bir uç düğüm(yaprak) olmalı ve bu düğümün kardeşi bulunmamalıdır.

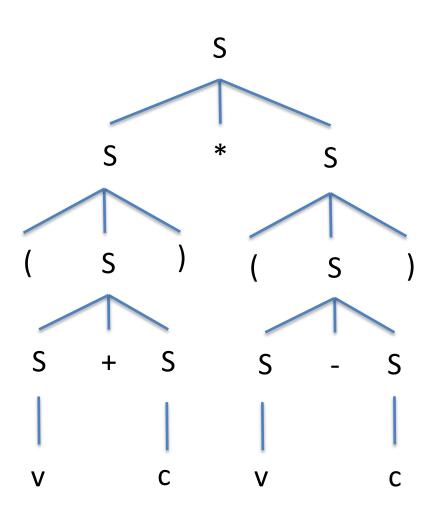
## Soldan ve Sağdan Türetme

- Yapraklarının etiketleri uç simgeler olan her türetme ağacına dilin bir tümcesi karşı gelir. Eğer dil belirgin (unambiguous) bir dil ise de, dilin her tümcesine bir türetme ağacı karşı gelir. Eğer dilin bir tümcesine birden çok türetme ağacı karşı geliyorsa, dil belirgin olmayan (ambiguous) bir dildir. Belirginlik dilbilgisinin ve dilin en temel özelliklerinden biridir.
- Tümceleri birden çok anlam taşıyan belirgin olmayan dillerin uygulamada hiçbir değeri yoktur. Bu nedenle, kullanılan dillerin tümünün belirginlik özelliğini sağladığı varsayılacaktır. Buna göre dilin her tümcesine bir türetme ağacı, yapraklarının etiketleri uç simgeler olan her türetme ağacına da bir tümce karşı geldiğini söyleyebiliriz.
- Dilin her tümcesine bir türetme ağacı karşı geldiği gibi, bir soldan türetme, bir de sağdan türetme karşı gelir.

# Soldan veya Sağdan Türetme

- Eğer bir tümce türetilirken, her adımda en soldaki değişkene bir türetme uygulanıyorsa, yapılan türetmeye soldan türetme (leftmost derivation) denir. Soldan türetmenin her ara adımında, eğer tümcesel yapı birden çok değişken içeriyorsa, öncelik en soldaki değişkene verilir.
- Eğer bir tümce türetilirken, her adımda en sağdaki değişkene bir türetme uygulanıyorsa, yapılan türetmeye sağdan türetme (rightmost derivation) denir. Sağdan türetmenin her ara adımında, eğer tümcesel yapı birden çok değişken içeriyorsa, öncelik en sağdaki değişkene verilir.

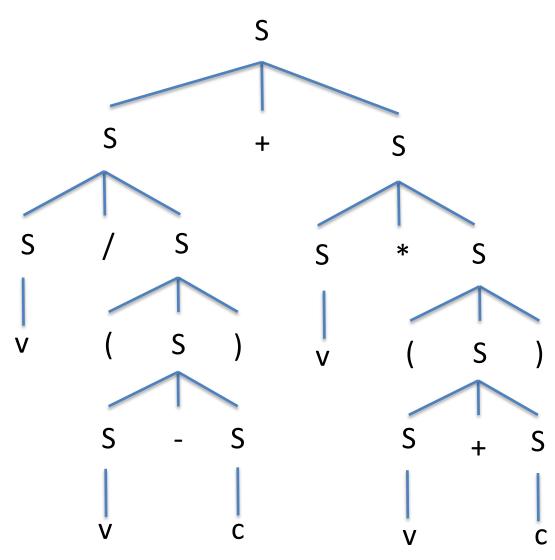
# Türetme ağacı: $w_1=(v+c)*(v-c)$



# $w_1$ =(v+c)\*(v-c) soldan ve sağdan türetme

b) 
$$w_1$$
'in sağdan  
türetilmesi:  
 $S=>S*S$   
 $=>S*(S)$   
 $=>S*(S-S)$   
 $\Rightarrow S*(S-C)$   
 $\Rightarrow S*(v-C)$   
 $\Rightarrow (S)*(v-C)$   
 $\Rightarrow (S+S)*(v-C)$   
 $=> (S+C)*(v-C) => (v+C)(v-C)$ 

# Türetme ağacı: $w_2=v/(v-c)+v*(v+c)$



# w<sub>2</sub>=v/(v-c)+v\*(v+c) soldan ve sağdan türetme

```
w<sub>2</sub>'in soldan türetilmesi:
                                                         b) w<sub>2</sub>'in sağdan türetilmesi:
S=>S+S
                                                      S => S + S
=>S/S+S
                                                      \RightarrowS+S*S
=>v/S*S
                                                      \Rightarrow S+S*(S)
=> v/(S)+S
                                                      \RightarrowS+S*(S+S)
=>v/(S-S)+S
                                                      \Rightarrow S+S*(S+c)
=>v/(v-S)+S
                                                      \Rightarrow S+S*(v+c)
=> v/(v-c)+S
                                                      \Rightarrow S+v*(v+c)
=> v/(v-c)+S*S
                                                      \Rightarrow S/S + v*(v+c)
=>v/(v-c)+v*S
                                                      \Rightarrow S/(S-S) + v*(v+c)
=>v/(v-c)+v*(S)
                                                      \Rightarrow S/(S-c) + v*(v+c)
=>v/(v-c)+v*(S+S)
                                                      \Rightarrow S/(v-c) + v*(v+c)
=> v/(v-c)+v*(v+S)
                                                      \Rightarrow v/(v-c) + v*(v+c)
\Rightarrow v/(v-c)+v*(v+c)
```

# 4.3 Dilbilgisinin Yalınlaştırılması

- Bağlamdan bağımsız bir dilbilgisi verildiğinde, bu dilbilgisinin daha az sayıda değişken ve kural içeren ve belirli biçimdeki kuralları içermeyen bir dilbilgisine dönüştürülmesine dilbilgisinin yalınlaştırılması denilir.
- 4.3.4. Yararsız Değişken, Simge ve Kurallar
- Eğer bir dilbilgininin uç simgelerinden biri, bu dilbilgisi tarafından türetilen tümcelerden hiçbirinde yer almıyorsa, bu uç simgeye **yararsız simge** (useless symbol) denilir.
- Eğer bir dilbilgisinin değişkenlerinden biri (A), en az bir tümce türetilirken elde edilen tümcesel yapıların en az birinde yer alıyorsa
- S=>  $\alpha_1$  A $\alpha$ 2 => w
- Bu değişken yararlı (usefull) bir değişkendir. Bu özelliği taşımayan değişkenlere ise yararsız değişken (useless variable) denilir. Yararsız değişken hiçbir tümcenin türetilmesinde kullanılmayan; kullanıldığı türetmelerin hiçbirinin sonunda ise bir tümce elde edilemeyen değişkendir.

# 4.3.4. Yararsız Değişken, Simge ve Kurallar

- Eğer bir dilbilgisinin yeniden yazma kurallarından biri (A=> β), en az bir tümcenin türetilmesi sırasında kullanılıyorsa, bu kural yararlı bir kuraldır(usefull rule). Hiçbir tümcenin türetilmesinde kullanılmayan kurallara ise yararsız kurallar (useless rules) adı verilir.
- Eğer bağlamdan-bağımsız bir dilbilgisi yararsız değişken, uç simge ya da kural içeriyorsa, bu dilbilgisine eşdeğer olan ve yararsız uç simge, değişken ve kural içermeyen bir dilbilgisi bulunabilir. Bunun için de öncelikle değişkenlerden hangilerinin yararlı olduğunun bulunması gerekir. Bir değişkenin (A) yararlı bir değişken olabilmesi için:
  - 1. Bu değişkenden başlanarak, en az bir uç simge dizisinin türetilebilmesi:

  - 2. Başlangıç değişkeninden bu değişkene ulaşmanın mümkün olması:

$$S=>\alpha_1 A \alpha 2$$

Bağlamdan-bağımsız bir dilbilgisi (G) verildiğinde, yararsız değişken ve kurallar atılarak eşdeğer bir dilbilgisinin bulunması iki adımda gerçekleştirilir.

## 4.3.4.1. Yararsız Değişken ve Kuralların Atılması

- Bağlamdan-bağımsız bir dilbilgisi (G) verildiğinde, yararsız değişken ve kurallar atılarak eşdeğer bir dilbilgisinin bulunması iki adımda gerçekleştirilir.
- Birinci adımda, algoritma 4.5 kullanılarak uç simge dizgisi türeten değişkenler kümesi bulunur. Bu kümede yer almayan değişkenler yararsız değişkenlerdir. Dilbilgisinden bu değişkenler ile bu değişkenlerin yer aldığı kurallar çıkarılarak eşdeğer bir dilbilgisi (G') bulunur.
- Algoritma 4.5. Uç simge dizgisi türeten değişkenler kümesinin bulunması:
- 1.  $T_D = \{A \mid (A => w) \in P \text{ ve } w \in V_T^* \};$
- 2. (T=T<sub>eski</sub>) oluncaya kadar aşağıdaki işlemleri tekrarla:
  - $2.1 T_{eski} = T_{D:}$
  - 2.2 Dilbilgisindeki her değişken (A) için aşağıdaki işlemleri tekrarla:
    - 2.2.1 Her (A=>β) kuralı için aşağıdaki işlemleri tekrarla:

eğer  $\beta \in (T_{eski}U V_T)^*$  ise A'yı  $T_D$ 'ye ekle

Son 2.2.1:

Son 2.2

Son 2.

# -4.3.4.1. Yararsız Değişken ve Kuralların Atılması

- İkinci adımda, algoritma 4.6 kullanılarak, birinci adımda bulunan dilbilgisindeki (G') değişkenlerden hangilerine başlangıç değişkeninden ulaşılabildiği bulunur. Başlangıç değişkeninden ulaşılamayan değişkenler de yararsız değişkenlerdir. Dilbilgisinden (G') bu değişkenler ile bu değişkenlerin yer aldığı kurallar çıkarılarak eşdeğer dilbilgisi (G'') bulunur.
- Algoritma 4.6.

Son 3:

```
    T<sub>U</sub>={S};
    T<sub>eski</sub>=Φ;
    (T<sub>U</sub>=T<sub>Eski</sub>) oluncuya kadar aşağıdaki işlemleri tekrarla:
3.1. T<sub>yeni</sub>=T<sub>U</sub>-T<sub>eski</sub>;
```

3.2. T<sub>Eski</sub>=T<sub>U</sub>;
3.3 T<sub>yeni</sub>'deki her değişken için aşağıdaki işlemleri tekrarla:

3.3.1 Her (A=> $\beta$ ) yeniden yazma kuralı için aşağıdaki işlemleri tekrarla:  $\beta$ 'daki tüm değişkenleri  $T_{\cup}$ 'ya ekle;

```
Son 3.3.1;
Son 3.3.1;
```

```
G_{47} = \langle V_N, V_T, P, S \rangle
   V_N = \{ S, A, B, C, D, \}
E, F }
    V_{T} = \{a, b\}
    P: S=> B | AC
BS
        A=>aA\mid aF
        B=>CF | b
        C=> cC \mid D
        D=> C | aD
|BD
        E=> aA | BSA
        F=> bB | b
```

- (G<sub>4.7</sub>)'ye eşdeğer; yararsız değişken, simge ve kural içermeyen bir dilbilgisi bulmak için önce (G<sub>4.7</sub>)'deki değişkenlerden hangilerinin uç simge dizgisi türeten değişkenler olduğunu bulmak gerekir. Uç simge dizgisi türeten değişkenler kümesi algoritma 4.5 ile aşağıdaki gibi bulunur:
- T<sub>D</sub> = { S, A, B, E, F }
- Yukarıdaki kümede yer almayan C ve D değişkenleri yararsız değişkenlerdir.

$$G'_{4.7} = \langle V_N, V_T, P, S \rangle$$
 $V_N = \{ S, A, B, C, D, E, F \}$ 
 $V_T = \{ a, b \}$ 
 $P: S = \rangle B \mid BS$ 
 $A = \rangle aA \mid aF$ 
 $B = \rangle b$ 
 $E = \rangle aA \mid BSA$ 
 $F = \rangle bB \mid b$ 

$$G''_{4.7} = \langle V_N, V_T, P, S \rangle$$
 $V_N = \{ S, B \}$ 
 $V_T = \{ b \}$ 
 $P: S = \langle B | BS \rangle$ 
 $B = \langle b | B \rangle$ 

- Şimdi de (G'<sub>4.7</sub>)'deki değişkenlerden hangilerinin S'den ulaşılabilen değişkenler olduğunu bulmak gerekir. S'den ulaşılabilen değişkenler kümesi algoritma 4.6 kullanılarak aşağıdaki gibi bulunur:
- T<sub>U</sub> = { S, B }
- Yukarıdaki kümede yer almayan A, E ve F değişkenleri yararsız değişkenlerdir. Dilbilgisinden bu üç değişken ile bu değişkenlerin yer aldığı kurallar çıkarıldığında a simgesinin de yararsız bir simge olduğu görülür. Bu simge de çıkarıldığında (G<sub>4.7</sub>)'ye eşdeğer; yararsız simge, değişken ve kural içermeyen aşağıdaki dilbilgisi bulunur:

#### 4.5.2. Örnek 2

#### Ömek 2

Bağlamdan-bağımsız G<sub>S.4.2</sub> dilbilgisi aşağıdaki gibi tanımlanıyor.

```
G<sub>S.4.2</sub> = < V<sub>N</sub>, V<sub>T</sub>, P, S >
V<sub>N</sub> = { S, D, E }
V<sub>T</sub> = { a, b, c }
P: S ==> aaS | aaD |
D ==> bEb | ccD
E ==> bEb | c
```

- a) G<sub>S.4.2</sub> tarafından tanınan bağlamdan-bağımsız dilin (L<sub>S.4.2</sub>) küme tanımını oluşturunuz.
- b) G<sub>S.4.2</sub>'yi Chomsky normal biçimine (CNF) dönüştürünüz. Dönüştürme sırasında kullanacağınız yeni değişkenleri A, B, C, ve X<sub>1</sub>, X<sub>2</sub>, X<sub>3</sub>, ... diye adlandırınız. Oluşturduğunuz CNF dilbilgisinin tanımını eksiksiz ve düzenli biçimde yazınız.

## 4.5.2. Cevap 2

a) 
$$L_{S.4.2} = \{ (aa)^n (cc)^m b^k c b^k \mid n \ge 1, m \ge 0, k \ge 1 \}$$

### 4.5.5. Örnek 5

#### Ömek 5

Aşağıda tanımlanan bağlamdan-bağımsız dillerin her birini türeten bağlamdan bağımsız bir dilbilgisi tanımlayınız.

```
a)L<sub>4.5.1</sub> = { a^{2k} b^n c^{2n} d^k \mid k \ge 1, n \ge 1 }
```

b)L<sub>4.5.2</sub> = { 
$$a^{n+m} b^k c^{m+k} d^{2n} \mid n, m, k \ge 0$$
 }

## 4.5.5. Cevap 5

a) G 
$$_{4.5.1}$$
 = <  $V_N$ ,  $V_T$ ,  $P$ ,  $S$  >
$$V_N = \{ S, A \}$$

$$V_T = \{ a, b, c, d \}$$

$$P : S \Rightarrow aaSd \mid aaAd$$

$$A \Rightarrow bAcc \mid bcc$$

b) G 
$$_{4.5.2}$$
 = < V<sub>N</sub>, V<sub>T</sub>, P, S >   
V<sub>N</sub> = { S, A, B }   
V<sub>T</sub> = { a, b, c, d }   
P :S  $\Rightarrow$  aSdd | A   
A  $\Rightarrow$  aAc | B   
B  $\Rightarrow$  bBc |  $\lambda$ 

#### 4.5.6. Örnek 6

#### Ömek 6

Aşağıdaki bağlamdan-bağımsız her dilbilgisinin türettiği dilin küme tanımını veriniz.

## 4.5.6. Cevap 6

a) 
$$L_{s.4.6.1} = \{ a^n b a^m b a^k \mid n>1, m>0, k>1 \}$$

b) 
$$L_{S.4.6.2} = \{ a^{n+m} b c^{m+k} d e^{k+n} \mid n, m, k>0 \}$$

### 4.5.7. Örnek 7

#### Ömek 7

Aşağıda tanımlanan bağlamdan-bağımsız dillerin her birini türeten bağlamdan bağımsız bir dilbilgisi tanımlayınız.

a) 
$$L_{S.4.7.1} = \{ a^n b^m c^k \mid m > 0, k > 0, n > m + k \}$$

b) 
$$L_{S.4.7.2} = \{ a^n b^m c^k \mid n > 0, m > 0, k = 2n + m \}$$

## 4.5.7. Cevap 7

```
a) G<sub>S,4,7,1</sub> = < V<sub>N</sub>, V<sub>T</sub>, P, S >
V<sub>N</sub> = { S, A, B }
V<sub>T</sub> = { a, b, c }
P: S ==> aaAc
A ==> aAc | aA | aBb
B ==> aBb | aB | λ
```

### 4.5.8. Örnek 8

#### Ömek 8

Aşağıdaki bağlamdan-bağımsız her dilbilgisinin türettiği dilin küme tanımını veriniz.

## 4.5.8. Cevap 8

a) 
$$L_{s.4.8.1} = \{ a^n b^{2k+1} d^{2k+1} a^{2n} \mid n \ge 0, k \ge 0 \}$$

b) 
$$L_{S.4.8.2} = \{ (ab)^n (cd)^k (ba)^k (dc)^n \ n \ge 0, k \ge 0 \}$$