

Otomata Teorisi

Ders-2

DÜZENLİ İFADELER (REGULAR EXPRESSIONS)

Tokat Gaziosmanpaşa Üniversitesi
Bilgisayar Mühendisliği Bölümü

Düzenli İfadeler

➤ ALFABE(Alphabet)



Kelimelerin üretiminde kullanılan birimlerin sonlu kümesine alfabe denir. Alfabenin elemanları genelde "sembol" veya "harf" olarak adlandırılmaktadır. Alfabe genelde Σ sembolü ile gösterilmektedir

➤ KELİME(Word)



Kelime, Σ alfabesinden seçilen sonlu sayıdaki sembolün bir araya gelmesinden oluşmaktadır.

➤ DİL(Language)



Bir alfabe üzerinden üretilmiş olan kelimelerin kümesi dildir.
Dile ait olan kelimelerin nasıl üretileceğinin kuralları tanımlanmış ise söz konusu dil biçimsel (formal) dil olarak adlandırılmaktadır

Düzenli İfadeler

➤ ALFABE(Alphabet)

Kelimelerin üretiminde kullanılan birimlerin sonlu kümesine alfabe denir. Alfabenin elemanları genelde "sembol" veya "harf" olarak adlandırılmaktadır. Alfabe genelde Σ sembolü ile gösterilmektedir

➤ KELİME(Word)

Kelime, Σ alfabesinden seçilen sonlu sayıdaki sembolün bir araya gelmesinden oluşmaktadır.

➤ DİL(Language)

Bir alfabe üzerinden üretilmiş olan kelimelerin kümesi dildir.

Dile ait olan kelimelerin nasıl üretilceğinin kuralları tanımlanmış ise söz konusu dil biçimsel (formal) dil olarak adlandırılmaktadır

➤ Düzenli İfadeler

Düzenli ifade (regular expression), dil tanımlayıcı ifadedir. Düzenli ifadeler ile tanımlanan diller düzenli diller (regular languages) olarak adlandırılmaktadır. Temel kavramlar bölümünde verilen örneklerde, diller küme biçiminde tanımlanmaktaydı. Örneğin, $\Sigma=\{x\}$ olmak üzere, $L=\{x^n, n=1,3,5,7,\dots\}$ kümesi tek sayıda x sembolünden oluşan kelimelerin dilini göstermektedir. Bu gösterim her ne kadar doğru olsa da, daha biçimsel (formal) bir gösterim de ortaya konmalıdır. Böylece, daha karmaşık dillerin de standart ve kolay bir biçimde tanımlanması sağlanmış olacaktır. Düzenli ifadeler, bu biçimsel gösterimi sağlamaktadır.

Düzenli İfadeler

Örnek -1

$\Sigma=\{x\}$ olmak üzere, $L=\{\Lambda, x, xx, xxx, xxxx, \dots\}$ kümesi ile tanımlanan dil, boş kelime de dahil olmak üzere içerisinde herhangi bir sayıda x sembolü olan tüm kelimelerin dilidir. L dilini tanımlayan düzenli ifade, x^* ifadesidir. Diğer deyişle, x^* ifadesi 0, 1, 2, n , ..., sonsuza kadar giden sayıda x sembolünden oluşan kelimeleri ifade eder.

Düzenli İfadeler

Örnek -2

$\Sigma=\{a,b\}$ olmak üzere,

$L=\{a, ab, abb, abbb, abbbb, \dots\}$ kümesi ile tanımlanan dil, **ab^*** düzenli ifadesi ile tanımlanabilecektir.

$L=\{\Lambda, ab, abab, ababab, \dots\}$ biçimindeki bir dil ise, **$(ab)^*$** düzenli ifadesi ile tanımlanabilecektir.

Düzenli İfadeler

Örnek -2

$\Sigma=\{a,b\}$ olmak üzere,

$L=\{a, ab, abb, abbb, abbbb, \dots\}$ kümesi ile tanımlanan dil, **ab^*** düzenli ifadesi ile tanımlanabilecektir.

$L=\{\Lambda, ab, abab, ababab, \dots\}$ biçimindeki bir dil ise, **$(ab)^*$** düzenli ifadesi ile tanımlanabilecektir.

ÖNEMLİ!!!

$ab^* \neq (ab)^*$

Düzenli İfadeler

Örnek -3

$\Sigma=\{a,b,c\}$ alfabeti üzerinde tanımlı $L_4=\{a, c, ab, cb, abb, cbb, abbb, cbbs, \dots\}$ dili,

$(a+c)b^*$ düzenli ifadesi ile tanımlanabilecektir.

Bu örnekteki $(a+c)$ ifadesi, parantez içerisindeki sembollerden birisinin seçileceği (a veya c) anlamına gelmektedir.

Seçilen a veya c sembolünü, $n=0,1,2,3,\dots$ olmak üzere n tane b sembolü izleyebilmektedir.

Düzenli İfadeler

Örnek -4

$\Sigma=\{a,b\}$ olmak üzere, a ve b sembollerinden oluşan **üç** uzunluktaki tüm kelimeleri içeren dil L ise, bu dil **$(a+b)(a+b)(a+b)$** düzenli ifadesi ile tanımlanabilecektir.

Soru:

L dilinde kaç elman bulunmaktadır?

Düzenli İfadeler

Örnek -4

$\Sigma=\{a,b\}$ olmak üzere, a ve b sembollerinden oluşan **üç** uzunluktaki tüm kelimeleri içeren dil L ise, bu dil **$(a+b)(a+b)(a+b)$** düzenli ifadesi ile tanımlanabilecektir.

Soru:

L dilinde kaç elman bulunmaktadır?

L dilinde yalnızca sekiz eleman bulunmaktadır. Bunlar;

aaa,
aab,
aba,
abb,
baa,
bab,
bba,
bbb

Düzenli İfadeler

Örnek -5

$\Sigma=\{a,b\}$ olmak üzere $\mathbf{a^*b^*}$ düzenli ifadesi ile tanımlanan dil, boş kelime ve eğer varsa a sembollerinin b sembollerinden önce geldiği kelimelerin dilidir.

Bu dil, $L=\{\Lambda, a, b, aa, ab, bb, aaa, aab, abb, bbb, aaaa, aaabb, \dots\}$ biçiminde de tanımlanabilecektir.

Düzenli İfadeler

Örnek -6

$\Sigma=\{x\}$ olmak üzere, tek sayıda x sembolü içeren kelimelerin dili

$L=x(xx)^*$ veya $L=(xx)^*x$ olarak da tanımlanabilecektir. Bu örnekte, değişme özelliği görülmektedir.

Diğer taraftan, x^*xx^* düzenli ifadesi, belirtilen dili tanımlamamaktadır. Bunun nedeni, x^*xx^* ifadesi, faktörlere ayrılmış biçimiyle $(xx)(x)(x)$ yani $xxxx$ kelimesinin dilin elemanı olmasına izin vermektedir. Bu da tek sayıda x sembolü içermeye kuralına aykırıdır. Bir düzenli ifade, bir dile ait olmayan hiçbir kelimenin tanımına izin vermemelidir.

Düzenli İfadeler

Örnek -7

$\Sigma=\{a,b\}$ olmak üzere, a sembolü ile başlayan ve b sembolü ile biten tüm kelimelerin dilini tanımlayın.

Düzenli İfadeler

Örnek -7

$\Sigma=\{a,b\}$ olmak üzere, a sembolü ile başlayan ve b sembolü ile biten tüm kelimelerin dilini tanımlayın.

Bu dil **$L = a(a+b)^*b$** düzenli ifadesi ile tanımlanabilecektir.

Düzenli İfadeler

Örnek -8

$\Sigma=\{a,b\}$ olmak üzere, herhangi bir yerinde en az bir a sembolü bulunan kelimelerin dilini tanımlayın.

Düzenli İfadeler

Örnek -8

$\Sigma=\{a,b\}$ olmak üzere, herhangi bir yerinde en az bir a sembolü bulunan kelimelerin dilini tanımlayın.

Bu dil **$(a+b)^*a(a+b)^*$** düzenli ifadesi ile tanımlanabilmektedir. Boş kelime ve yalnızca b sembollerinden oluşan kelimeler budilin elemanı değildir.

Düzenli İfadeler

Örnek -9

$\Sigma=\{a,b\}$ olmak üzere, en az iki a sembolü içeren kelimelerin dilini tanımlayınız.

Düzenli İfadeler

Örnek -9

$\Sigma=\{a,b\}$ olmak üzere, en az iki a sembolü içeren kelimelerin dilini tanımlayınız.

Bu dil $(a+b)^*a(a+b)^*a(a+b)^*$ düzenli ifadesi ile tanımlanabilmektedir. Diğer taraftan, $b^*ab^*a(a+b)^*$ düzenli ifadesi de en az iki a içeren kelimelerin dilini tanımlıyor olarak nitelendirilebilecektir.

Düzenli İfadeler

Örnek -10

$\Sigma=\{a,b\}$ olmak üzere, yalnızca iki a sembolü içeren kelimelerin dili tanımlayınız.

Düzenli İfadeler

Örnek -10

$\Sigma=\{a,b\}$ olmak üzere, yalnızca iki a sembolü içeren kelimelerin dili tanımlayınız.

Bu dil **$b^*ab^*ab^*$** düzenli ifadesi ile tanımlanabilecektir.

Düzenli İfadeler

Örnek -11

- $\Sigma=\{a,b\}$ olmak üzere, en az bir a sembolü ve en az bir b sembolü içeren kelimelerin dili tanımlayınız.

Düzenli İfadeler

Örnek -11

- $\Sigma=\{a,b\}$ olmak üzere, en az bir a sembolü ve en az bir b sembolü içeren kelimelerin dili tanımlayınız.

Bu dil aşağıdaki düzenli ifade ile tanımlanabilecektir.

$[(a+b)^*a(a+b)^*b(a+b)^*]+[(a+b)^*b(a+b)^*a(a+b)^*]$

Düzenli İfadeler

Örnek -12

$\Sigma=\{a,b\}$ olmak üzere, yalnızca b sembollerinden oluşan veya bir a sembolünü izleyen b sembollerinden oluşan dil, $L= b^*+ab^*$ düzenli ifadesi ile tanımlanabilecektir.

Diğer bir alternatif nedir?

Düzenli İfadeler

Örnek -12

$\Sigma=\{a,b\}$ olmak üzere, yalnızca b sembollerinden oluşan veya bir a sembolünü izleyen b sembollerinden oluşan dil, $L= b^*+ab^*$ düzenli ifadesi ile tanımlanabilecektir.

Diğer bir alternatif nedir?

$(\Lambda+a)b^*$ şeklinde tanımlanabilir.

Düzenli İfadeler

Örnek -13

$\Sigma=\{0,1\}$ olmak üzere, çift sayıda uzunluktaki kelimelerin dili nedir?

Düzenli İfadeler

Örnek -13

$\Sigma=\{0,1\}$ olmak üzere, çift sayıda uzunluktaki kelimelerin dili nedir?

$L=(00+01+10+11)^*$ düzenli ifadesi ile çift sayıda uzunluktaki kelimelerin dili tanımlanmaktadır.

Düzenli İfadeler

Örnek -14

- $\Sigma=\{0,1\}$ olmak üzere, tek sayıda uzunluktaki kelimelerin dilini tanımlayabilecek düzenli ifadeleri yazınız.

Düzenli İfadeler

Örnek -14

- $\Sigma=\{0,1\}$ olmak üzere, tek sayıda uzunluktaki kelimelerin dilini tanımlayabilecek düzenli ifadeleri yazınız.

Örnek ifadeler:

$(0+1)(00+01+10+11)^*$

$(00+01+10+11)^*(0+1)$

$0+1+(00+01+10+11)(0+1)(00+01+10+11$

Düzenli İfadeler

Örnek -15

$\Sigma=\{0,1\}$ olmak üzere, $L=(a+b)^* (aa+bb) (a+b)^*$ düzenli ifadesi ile tanımlanan dili nasıl ifade edersiniz?

Düzenli İfadeler

Örnek -15

$\Sigma=\{0,1\}$ olmak üzere, $L=(a+b)^* (aa+bb) (a+b)^*$ düzenli ifadesi ile tanımlanan dili nasıl ifade edersiniz?

Cevap:

Herhangi bir yerinde aa veya bb gibi çift sembol içeren kelimelerin dilidir.

Düzenli İfadeler

Örnek -16

$\Sigma=\{0,1\}$ olmak üzere, çeşitli biçimlerde bir araya gelmiş 0 ve 1 sembollerini iki tane 1 sembolü ve bunu da eğer varsa 0 sembolünü mutlaka 1 sembol(lerinin) izlediği kelimelerin dili için düzenli ifade yazınız.

Düzenli İfadeler

Örnek -16

$\Sigma=\{0,1\}$ olmak üzere, çeşitli biçimlerde bir araya gelmiş 0 ve 1 sembollerini iki tane 1 sembolü ve bunu da eğer varsa 0 sembolünü mutlaka 1 sembol(lerinin) izlediği kelimelerin dili için düzenli ifade yazınız.

Cevap:

$(0+1)^*11(1+01)^*$

NOT: Bu tür örüntüler, görüntü işleme (image processing) veya örüntü tanıma (pattern recognition) konularında görülebilmektedir.

Düzenli İfadeler

Örnek -17

$\Sigma=\{0,1\}$ olmak üzere, $\text{EVEN-EVEN} = \{ \Lambda, aa, bb, aabb, abab, abba, baab, bbaa, aaaabb, \dots \}$ (Çift sayıda a ve b içeren kelimelerin dili). Bu dile ilişkin bir düzenli ifade yazınız.

Düzenli İfadeler

Örnek -17

$\Sigma=\{0,1\}$ olmak üzere, $\text{EVEN-EVEN} = \{ \Lambda, aa, bb, aabb, abab, abba, baab, bbaa, aaaabb, \dots \}$ (Çift sayıda a ve b içeren kelimelerin dili). Bu dile ilişkin bir düzenli ifade yazınız.

- $[aa + bb + (ab+ba)(aa+bb)^*(ab+ba)]^*$

Factor-1 Factor-2

Factor-3

Düzenli İfadeler

Örnek -18

Yanyana iki a (double a) içermeyen kelimelerin diline ilişkin bir düzenli ifade yazınız.

Düzenli İfadeler

Örnek -18

Yanyana iki a (double a) içermeyen kelimelerin diline ilişkin bir düzenli ifade yazınız.

- $b^*(abb^*)^*(\Lambda + a)$

Düzenli İfadeler

Örnek -19

$\Sigma = \{b, d, k, t\}$ olmak üzere

$L = k[(bd) + (bbdd) + (bbbddd)]^*t$ düzenli ifadesi veriliyor. Bu ifadeden üretilebilecek en kısa uzunluktaki kelimelerden 5 tanesini yazınız.

Düzenli İfadeler

Örnek -19

$\Sigma=\{b,d,k,t\}$ olmak üzere

$L=k[(bd)+(bbdd)+(bbbddd)]^*t$ düzenli ifadesi veriliyor. Bu ifadeden üretilebilecek en kısa uzunluktaki kelimelerden 5 tanesini yazınız.

- kt
- kbd t
- kbdbdt
- kbdbddt