II Sınav

FORMAL DİLLER VE OTOMATLAR

DÜZGÜN DILLER VE DÜZGÜN GRAMERLER

Bir dilin düzgün olabilmesi için bu dili kabul eden bir otomat olması gerekir. Yani her düzgün dil bir DFA ya da NFA tarafından tanımlanabilir. Ancak bazı durumlarda düzgün dilleri ifade etmek için başka yöntemlere ihtiyaç duyarız (düzgün ifadeleri kullanırız).

Düzgün Dil: En basit dil olan boş dizgi veya 1 uzunluğundaki dilden başlayarak

-birleşim (union)

-ardarda ekleme (concatenation)

-Kleene işleci ya da yıldızı

işlemlerinden birinin veya birkaçının kullanılması ile yeni diller yaratılır ve bu şekilde elde edilen diller «düzgün diller» olarak adlandırılır.

- **Düzgün İfade:** ∑ alfabesi üzerinde bir düzgün ifade şöyle tanımlanır:

\- Ø, boş dile, Ø veya { } diline ait düzgün ifadedir.

\- λ, {λ} diline ait düzgün ifadedir.

\- ∀ a ∈ ∑ için a, {a} diline ait düzgün ifadedir.

\- ∑ üzerinde herhangi r ve s düzgün ifadeleri L_r ve L_s dillerine ait olmak üzere,

(rs) düzgün ifadesi L_rL_s diline

(r+s) düzgün ifadesi $L_r \cup L_s$ diline

(r*) düzgün ifadesi L^* diline aittir.

- $L \in \Sigma^*$ olmak üzere (∑ alfabesi üzerine oluşturulan bir L dili) düzgün ifadelerin biri veya birkaçı ile simgelenebiliyorsa bu dil düzgün dildir.

Regular Languages correspond to Regular Expressions

$L = \varnothing$	RE is $\varnothing$
$L = \{\lambda\}$	$RE = \lambda$
$L = \{a\}$	$RE = a$
$L = L_1 \cup L_2$	$RE = (r_1 + r_2)$
$L = L_1 L_2$	$RE = (r_1 r_2)$
$L = L_1^*$	$RE = (r_1^*)$

* = istenen sayıda λ = etkisiz
+ = ikisinden biri

DÜZGÜN IFADE ÖRNEKLERI

- 01* = {0, 01, 011, 0111, ….} 1 tane 0, istediğimiz kadar 1.
- 0* + 1 = {λ, 0, 1, 00, 000, 0000,…}
- (01*)(01) = {001, 0101, 01101, 011101,…}
- (0+1)* = {λ, 0, 1, 00, 01, 10, 11,…} (0 ve 1'lerin tüm dizgileri}
- (0+1)*00(0+1)* = {00, 100, 1001, 000, 0001, …}

(İçinde mutlaka 00 olan, 0 ve1'lerin tüm dizgileri)

- (0+1)*00(0+1)* + (0+1)*11(0+1)* → içinde 00 **ya da** 11 olan 0 ve 1'lerin tüm dizgileri
- (1+10)* = {λ} veya 1 ile başlayan ve 00 içermeyen tüm dizgilerin kümesi
- (0+1)*011 → 011 ile biten tüm dizgiler
- 0*1* = {λ, 0, 00, 000, …, 1, 11, 111,…, 01,001, 011…}

1'den sonra 0 gelmeyen tüm dizgiler

- 2'den fazla a içermeyen tüm katarlar:

  $(b + c)^*(\lambda + a)(b + c)^*(\lambda + a)(b + c)^*$

- 00^*11^* → en azından bir tane 0 ve 1 içeren ve 1'den sonra 0 gelmeyen tüm dizgiler

  → ya 0 ya λ akar.

- $(0+ \lambda)1^* = \{01^n : n \geq 0\} + \{1^n : n \geq 0\}$

  01^* $\lambda 1^*$

- {001} diline ait düzgün ifade 001

- {0, 1} diline ait düzgün ifade (0 + 1)

- $\{1, 1010, 110110\}^*$ diline ait düzgün ifade

  $(1 + 1010 + 110110)^*$

  → virgül

  "," yerine +

- $\{10, \lambda\}\{10\}$ diline ait düzgün ifade $((10+\lambda)10)$

  {1010, λ10} = {1010, 10}

DÜZGÜN IFADELERIN SADELEŞTIRILMESI

r bir düzgün ifade ise

$r^+ \Rightarrow$ en az 1.r demek.

- (rr) yerine r^2
- ((rr)r) yerine r^3
- (r^+) yerine ((r^*)r) yazılabilir

Düzgün ifadeler gerektiğinde basitleştirilebilmelidir.

1*(1+λ) düzgün ifadesi yerine 1*

1*1* yerine 1*

0*+1* yerine 1*+0* $\Rightarrow$ istenildiği kadar olduğu için, yer değiştirilebilir.

istenildiği kadar a veya b ve sonu kesinlikle a veya bb olacak.

Örnek: ∑={a,b} için r=(a+b)*(a+bb) düzgün ifadesi.

Bu ifadenin dili L(r)={a, bb, aa, abb, ba, bbb,…}

a ile ya da bb ile biten tüm dizgilerin kümesi

istenen sayıda 2 tane a

Örnek: r=(aa)*(bb)*b ifadesi 0 ya da çift sayıda a'ları takip eden tek sayıda b'lerden oluşan dizgiler kümesi.

$L(r)=\{a^{2n}b^{2m+1} : n \geq 0, m \geq 0\}$

Örnek: ∑={0,1} için dil

L(r) = {w∈∑* : w en azından bir çift ardarda gelen 0'lara sahiptir}

L(r) diline ait düzgün ifade r = (0+1)*00(0+1)*

$(0+1)^* =$ rastgele

bu iki kısım rastgele.

Örnek: L={w ϵ {0,1}* : w içinde ardarda gelen 0 bulunmasın}

Dizgi içinde her 0 görüldüğünde arkasından hemen 1 gelmeli

$(1^*011^*)^*$

Ancak 0 ile biten ya da sadece 1'lerden oluşan dizgiler de olabilir

$r = (1^*011^*)^*(0+\lambda) + 1^*(0+\lambda)$

$r = (1+01)^*(0+\lambda)$ ✓

Her ikisi de doğrudur. Bir dil için birden fazla düzgün ifade olabilir.

Örnek: $\Sigma=\{a,b,c\}$ olsun. En fazla iki tane a içeren tüm dizgiler.

$(b+c)^*(a+\lambda)(b+c)^*(a+\lambda)(b+c)^*$

Düzgün ifade yazmak için ipuçları:

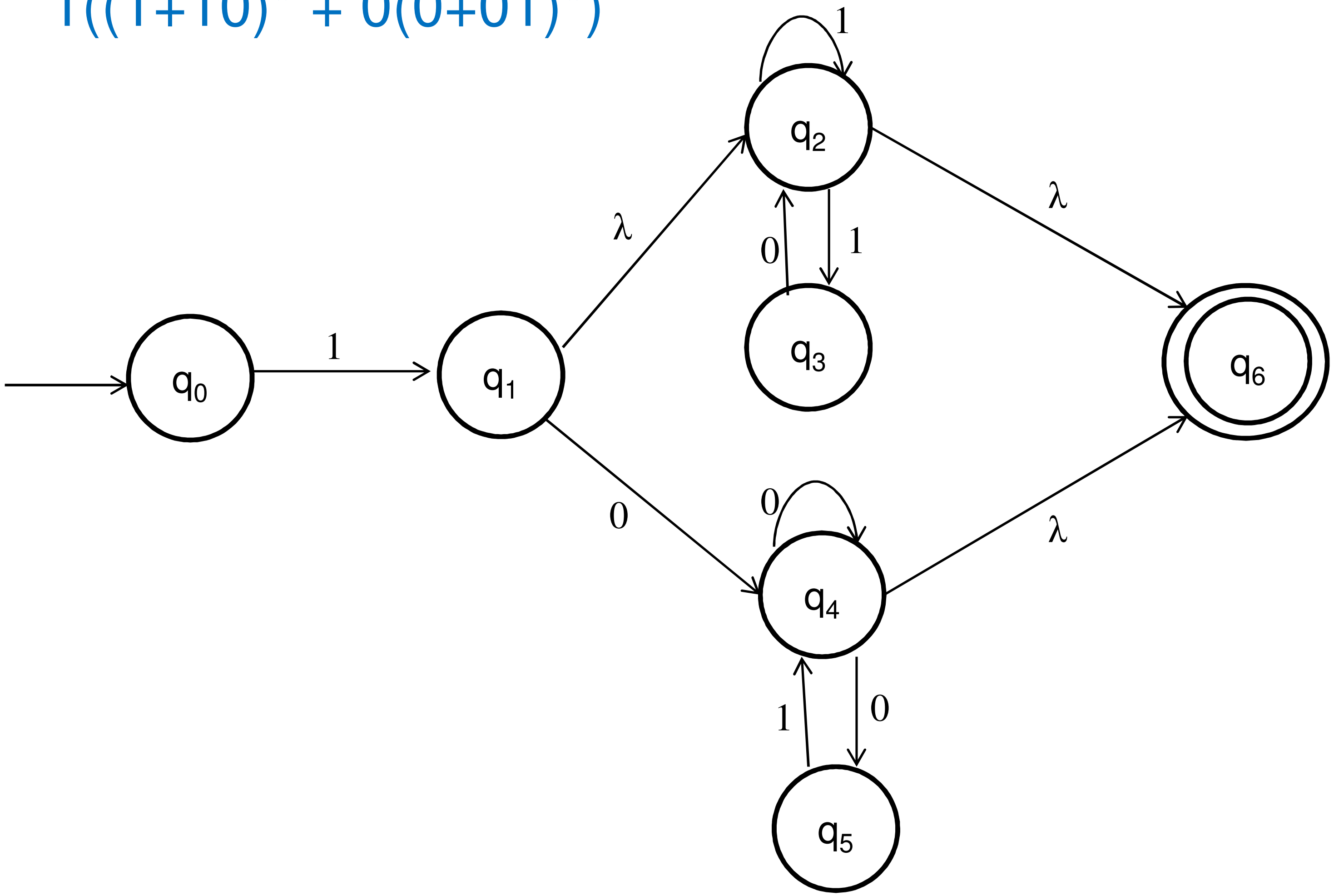
$\Sigma=\{a,b,c\}$ olsun

- Sıfır ya da daha fazla a : a^*
- Bir ya da daha fazla a : aa^* ; a^+
- Tüm dizgiler : $(a+b+c)^*$
- Boş dizgi dışında tüm dizgiler: $(a+b+c)(a+b+c)^*$
- a içermeyen tüm dizgiler : $(b+c)^*$
- Sadece 1 tane a içeren tüm dizgiler: $(b+c)^*a(b+c)^*$

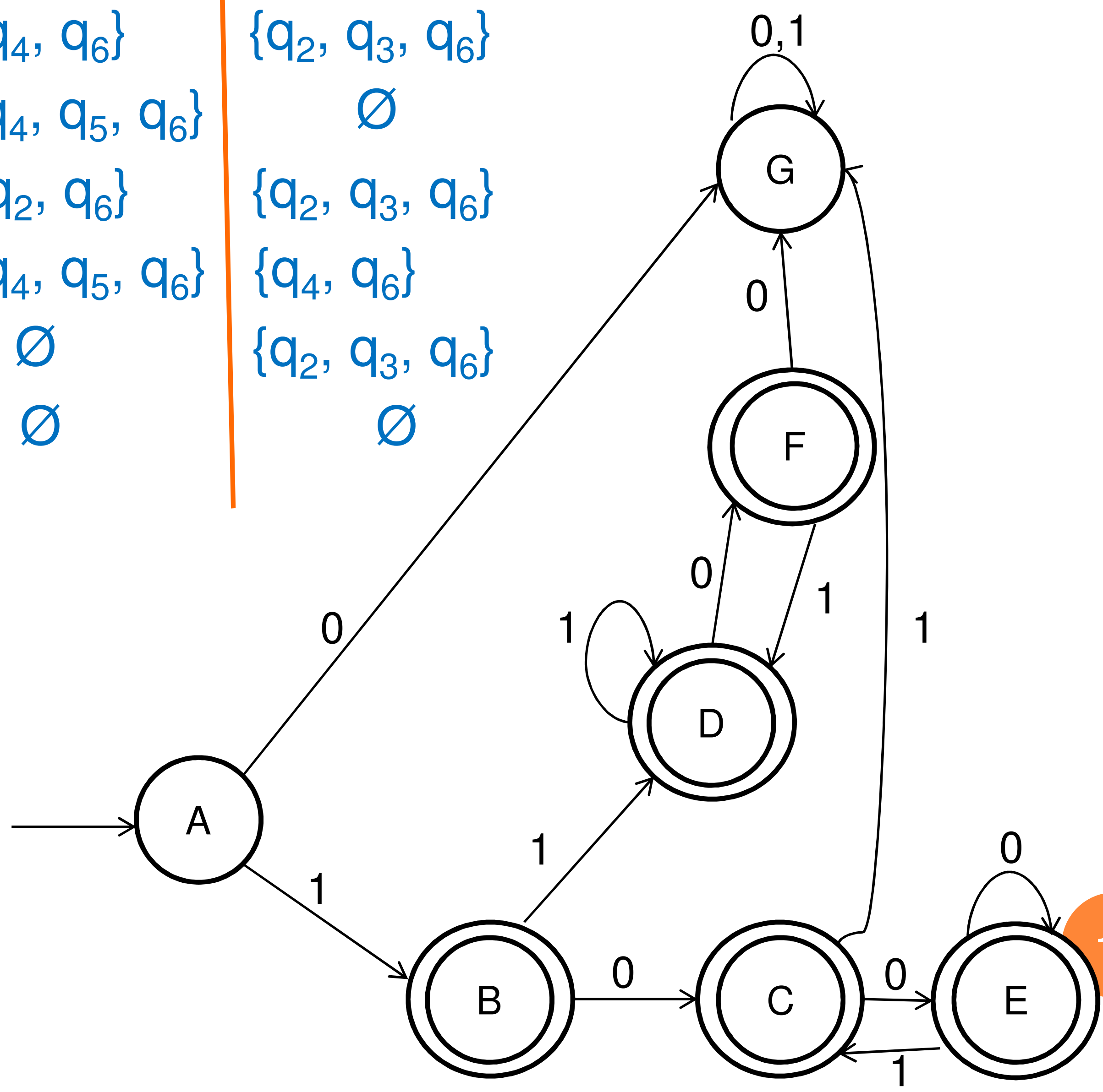
- Örnek: 1(1+10)* + 10(0+01)* dilini kabul eden DFA?

(Öncelikle NFA elde edilir, daha sonra DFA'ya dönüştürülür)

1((1+10)* + 0(0+01)*)



	0	1
$A=\{q_0\}$	$\varnothing$	$\{q_1, q_2, q_6\}$
$B= \{q_1, q_2, q_6\}$	$\{q_4, q_6\}$	$\{q_2, q_3, q_6\}$
$C= \{q_4, q_6\}$	$\{q_4, q_5, q_6\}$	$\varnothing$
$D= \{q_2, q_3, q_6\}$	$\{q_2, q_6\}$	$\{q_2, q_3, q_6\}$
$E=\{q_4, q_5, q_6\}$	$\{q_4, q_5, q_6\}$	$\{q_4, q_6\}$
$F=\{q_2, q_6\}$	$\varnothing$	$\{q_2, q_3, q_6\}$
$G= \varnothing$	$\varnothing$	$\varnothing$



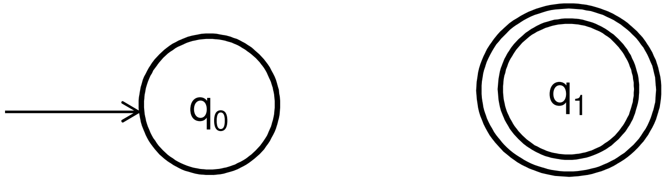
Düzgün İfadeler ve Düzgün Diller Arasındaki Bağlantı

Her iki kavram da temelde aynıdır. Her düzgün dil için bir düzgün ifade vardır ve her düzgün ifade için bir düzgün dil vardır.

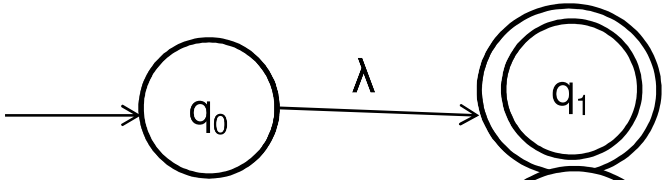
Düzgün Dilleri Gösteren Düzgün İfadeler:

r bir düzgün ifade ise L(r) de bu ifadeye ait olan düzgün bir dildir. Bir dilin düzgün olabilmesi için de herhangi bir DFA tarafından kabul edilmesi gerekmektedir. NFA ve DFA'ların denkliğinden, aynı şeyi NFA için de söyleyebiliriz.

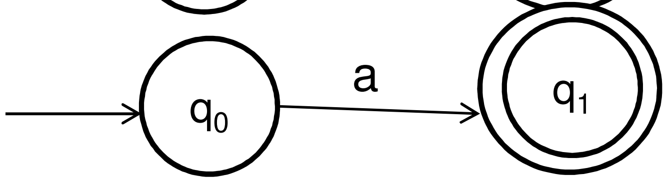
- Ø dilini kabul eden NFA



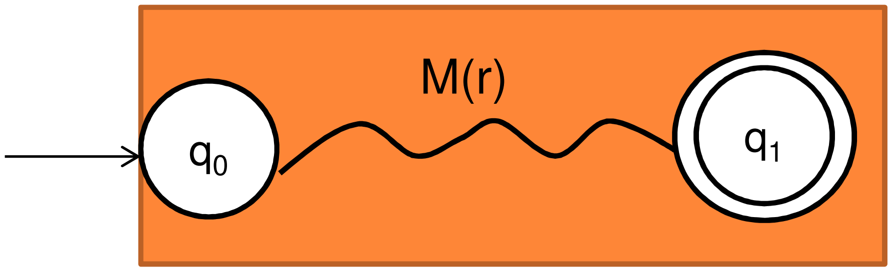
- {λ} dilini kabul eden NFA



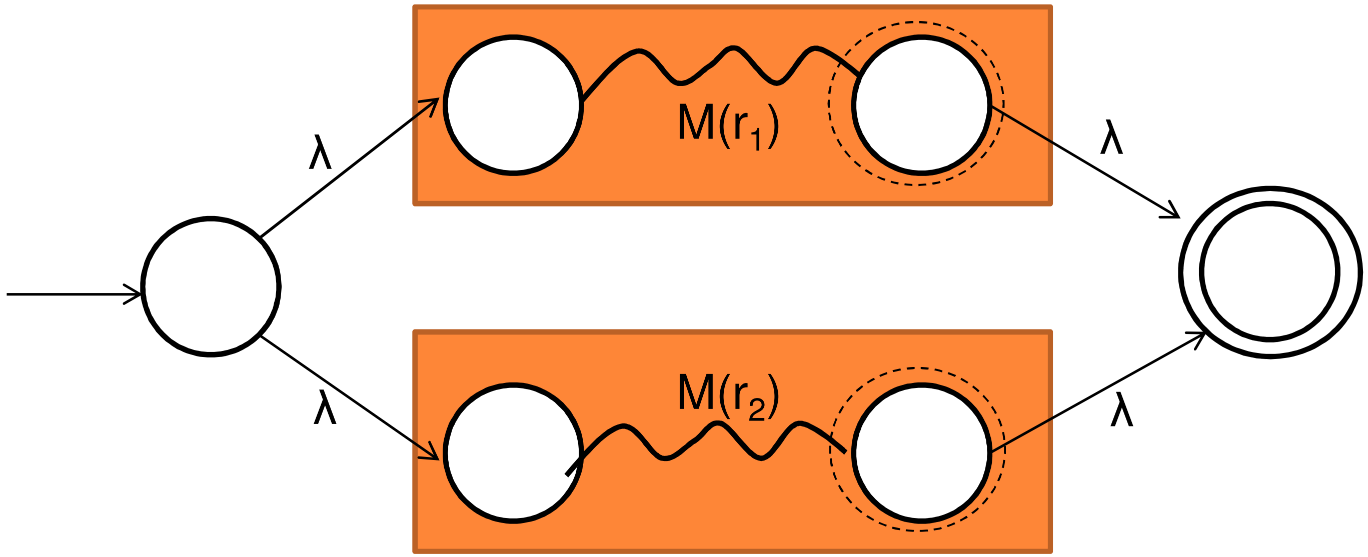
- {a} dilini kabul eden NFA



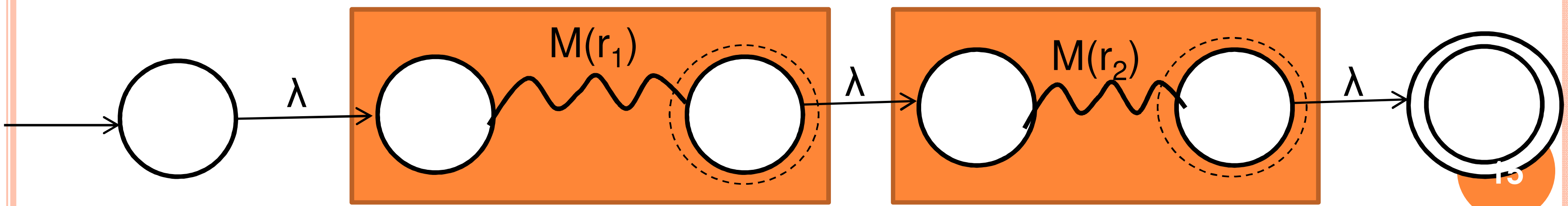
- L(r) dilini kabul eden NFA



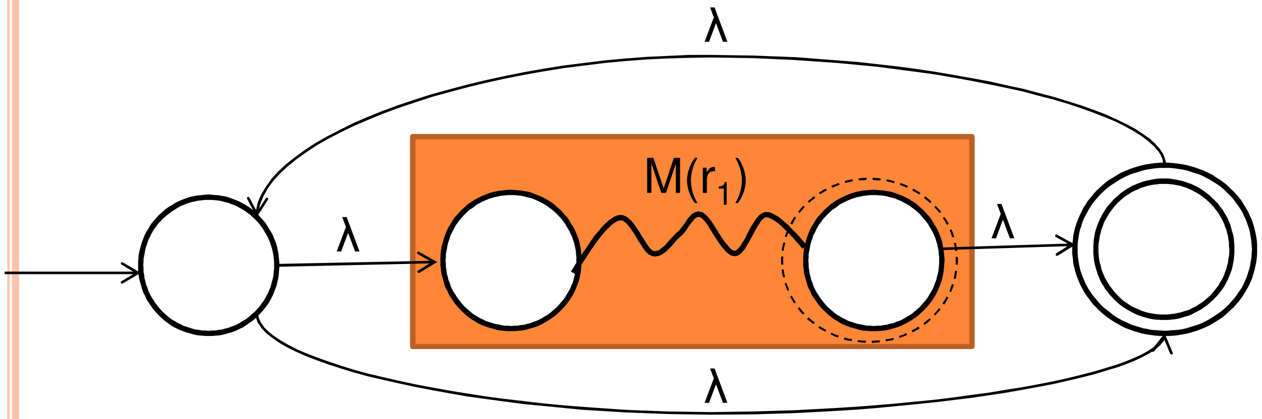
- $L(r_1 + r_2)$ dilini kabul eden NFA



- $L(r_1 r_2)$ dilini kabul eden NFA

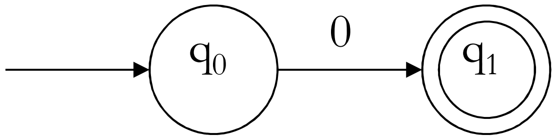


- L(r_1^*) dilini kabul eden NFA

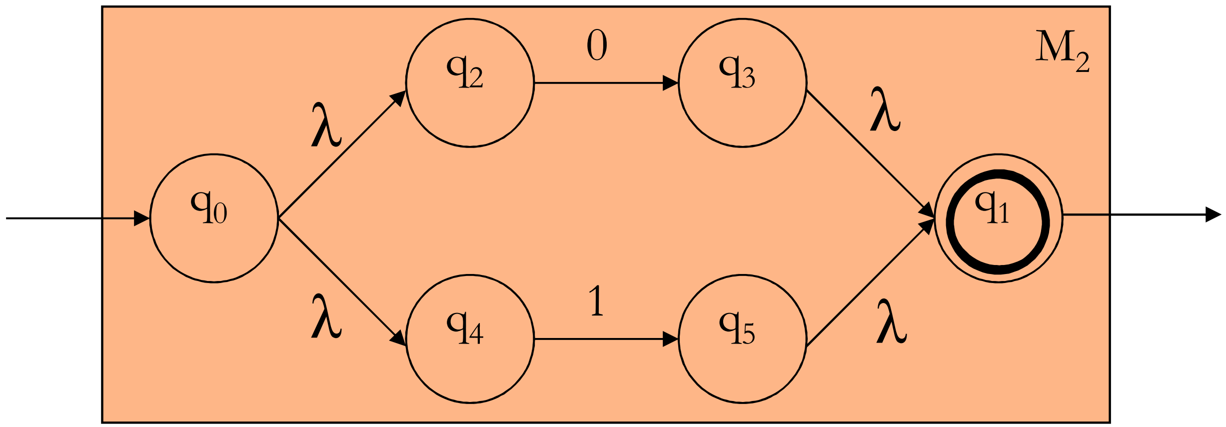


Örnek: Düzgün ifade → λNFA

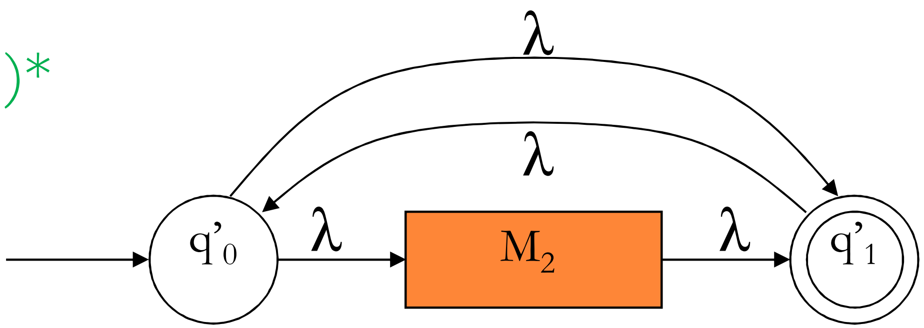
- $R_1 = 0$



- $R_2 = 0 + 1$



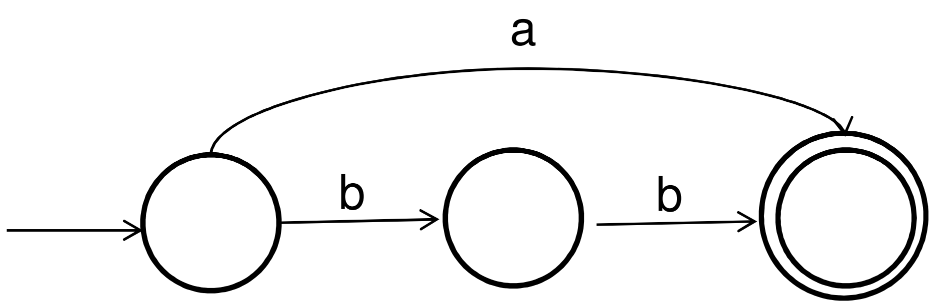
- $R_3 = (0 + 1)^*$



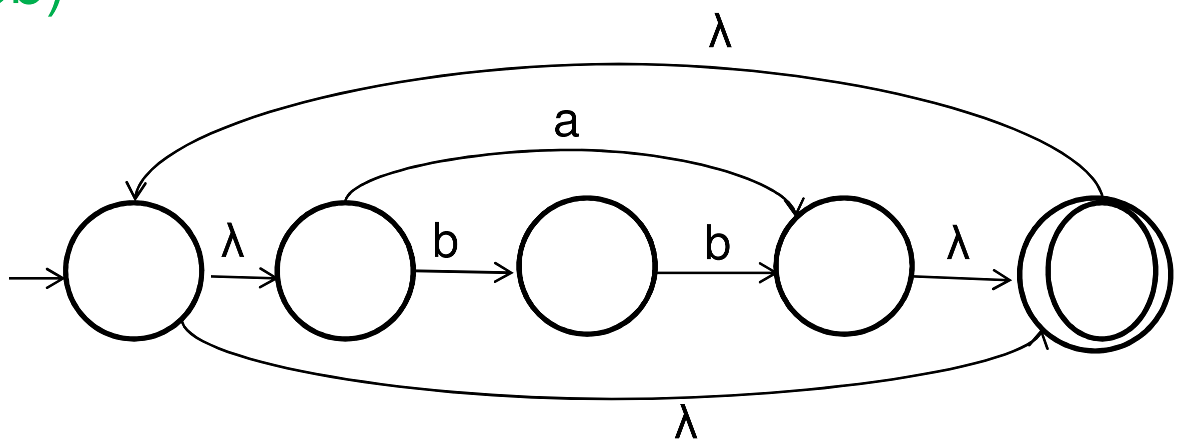
Örnek: $r=(a+bb)^*(ba^*+\lambda)$ olmak üzere L(r) dilini kabul eden NFA bulun.

L(a+bb) dilini kabul eden otomat M_1 olsun.

M_1 :

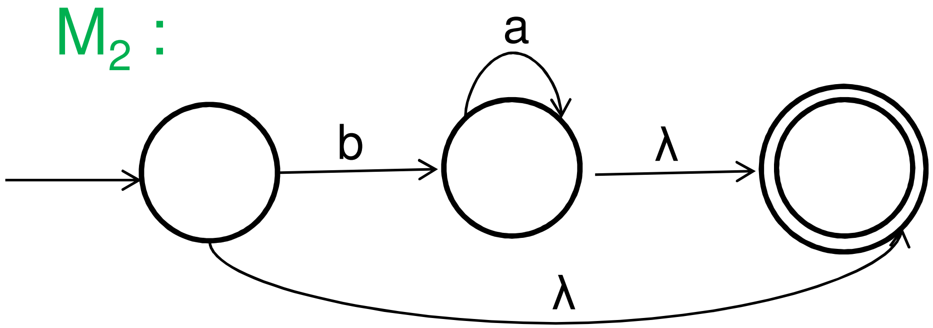


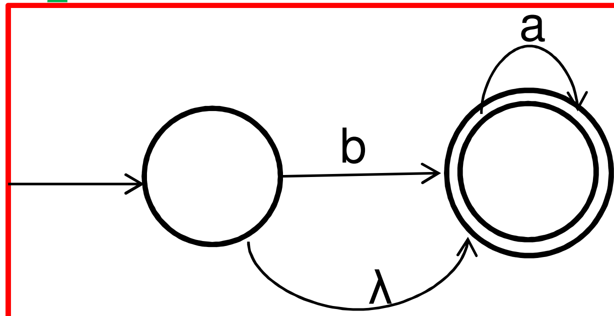
$L(a+bb)^*$



- L(ba*+λ) dilini kabul eden otomat M_2 olsun.

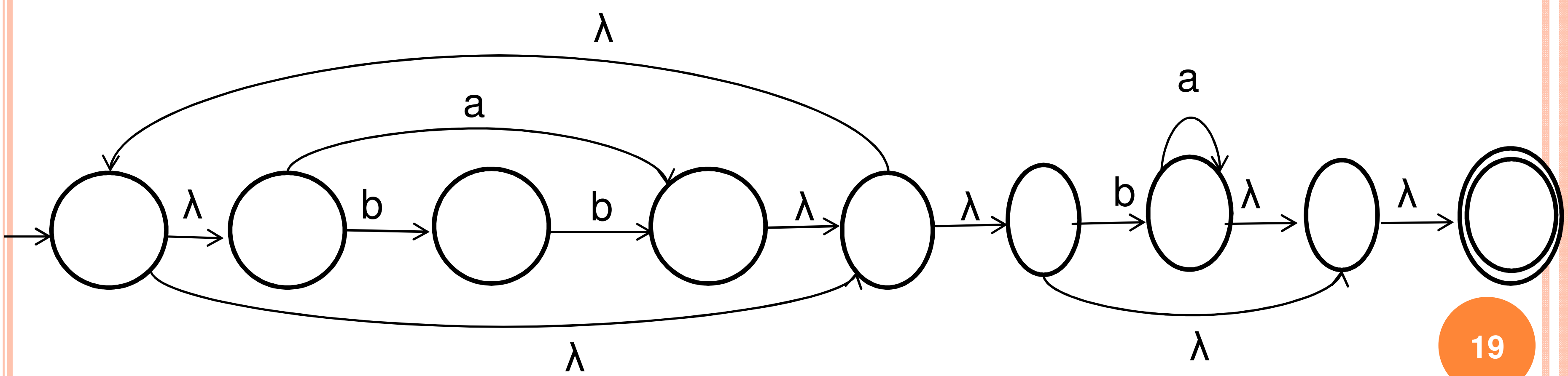
M_2 :



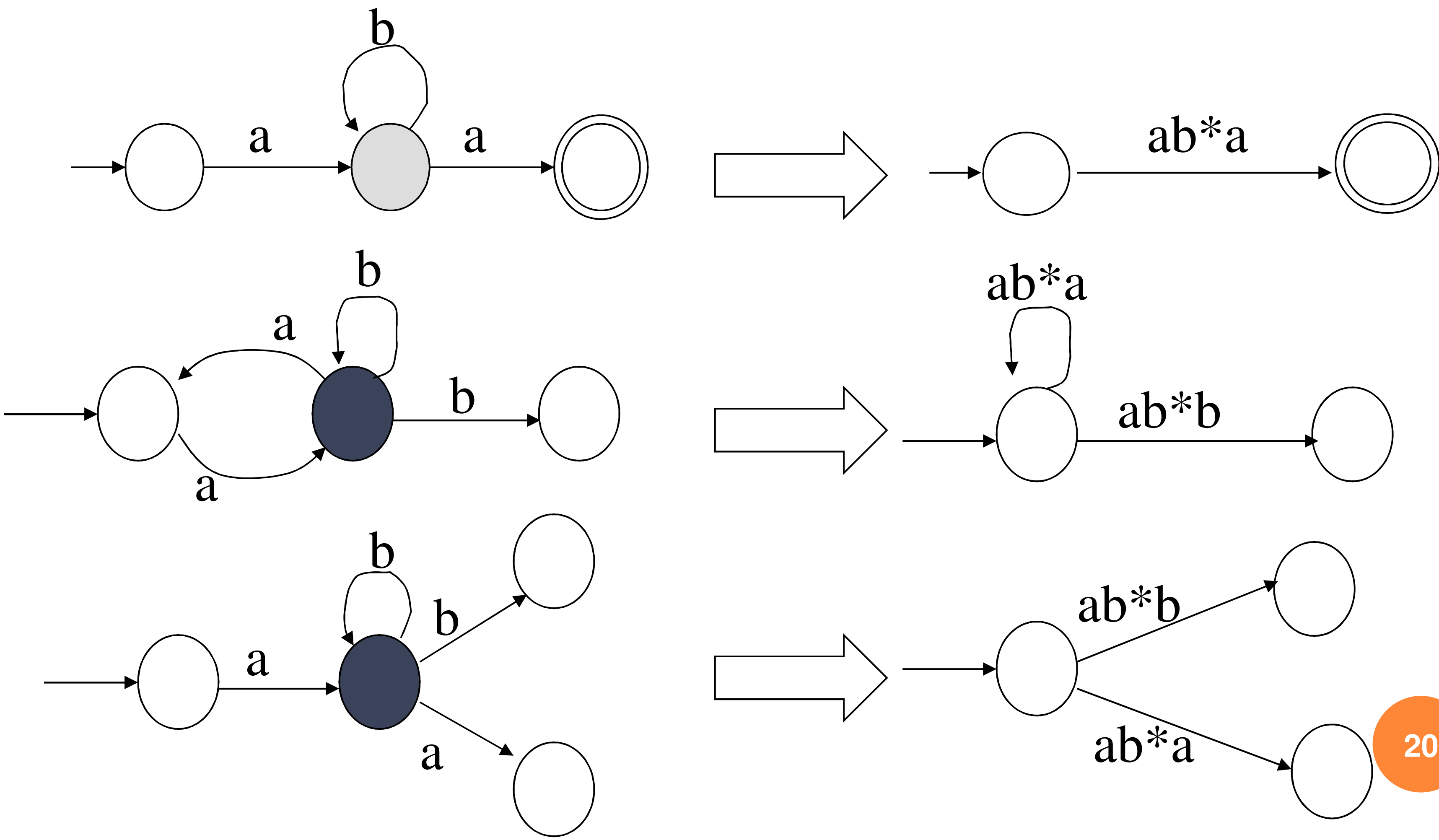


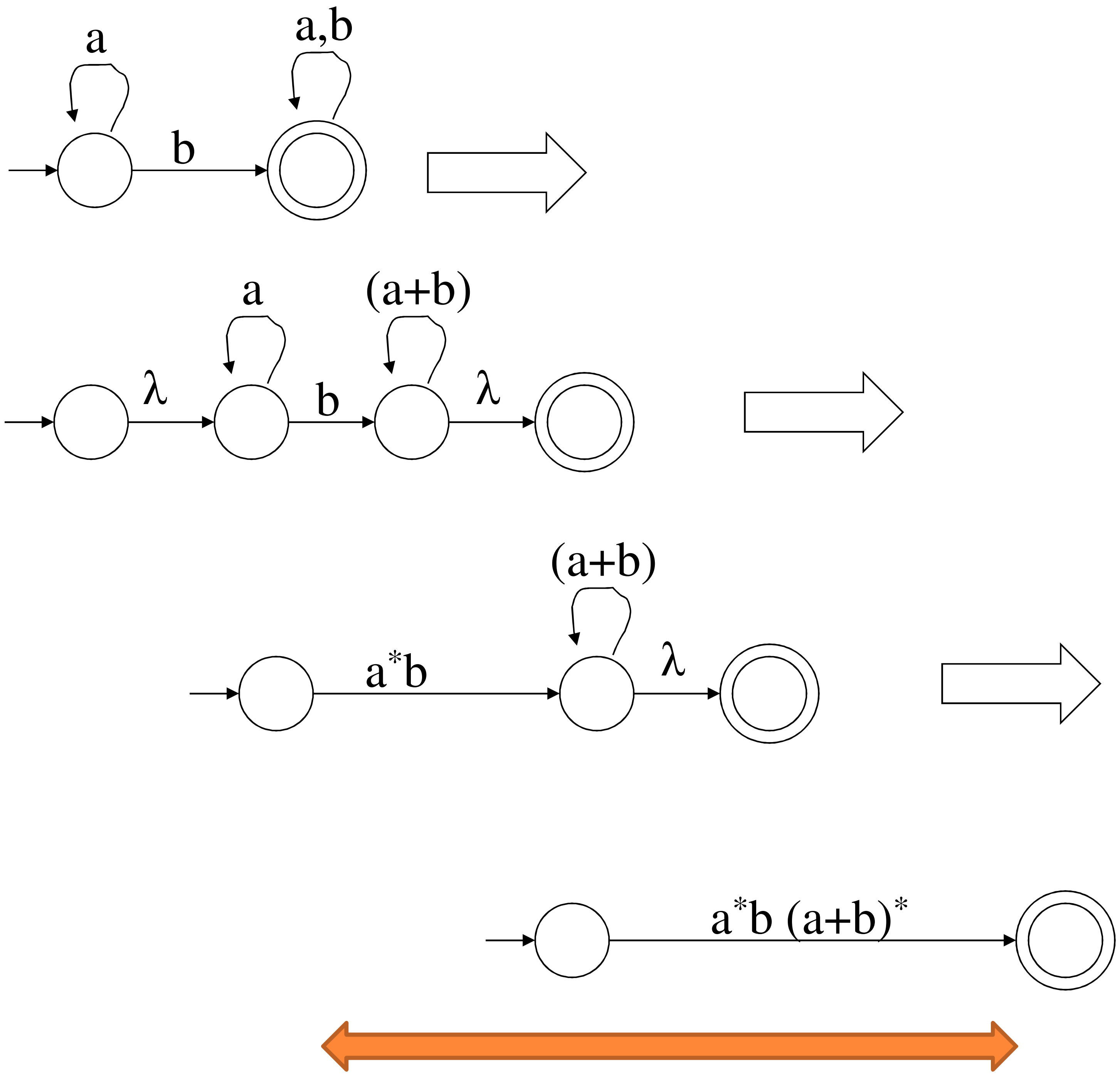
a*'ı da kabul edeceği için bu otomat olmaz!!

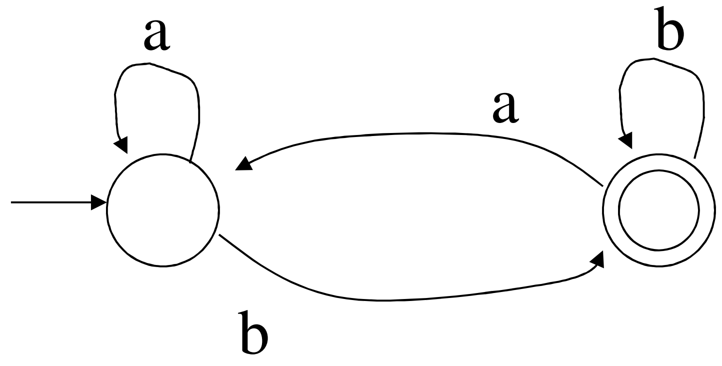
M_1 ve M_2'yi birleştirerek otomatı çizelim:

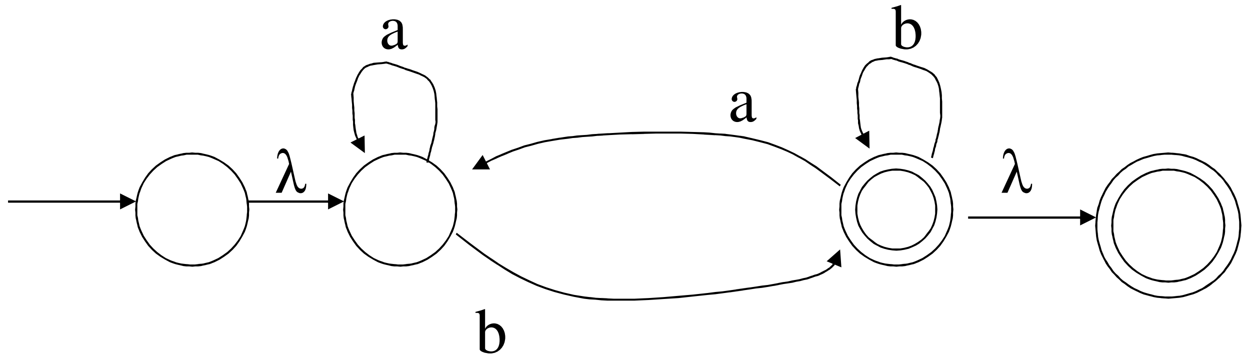


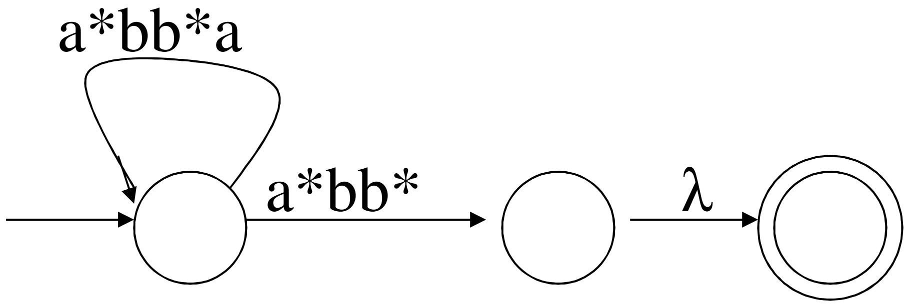
DFA’yı eşdeğer düzgün ifadeye dönüştürme:

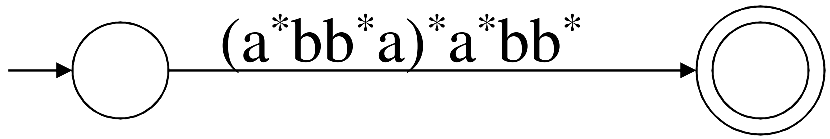


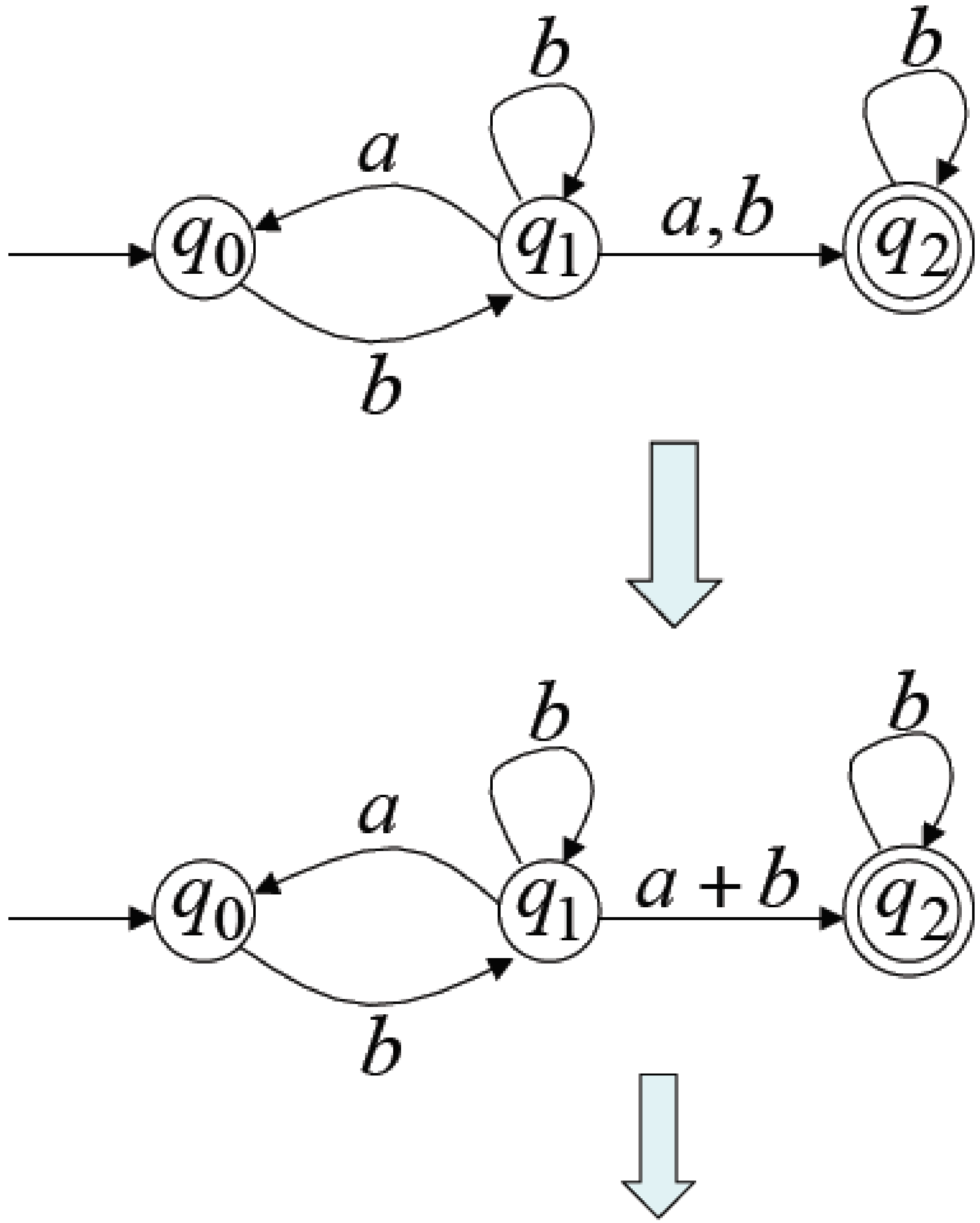

a
a,b
b
a
(a+b)
λ
b
λ
(a+b)
a*b
λ
a*b (a+b)*


a
b
a
b

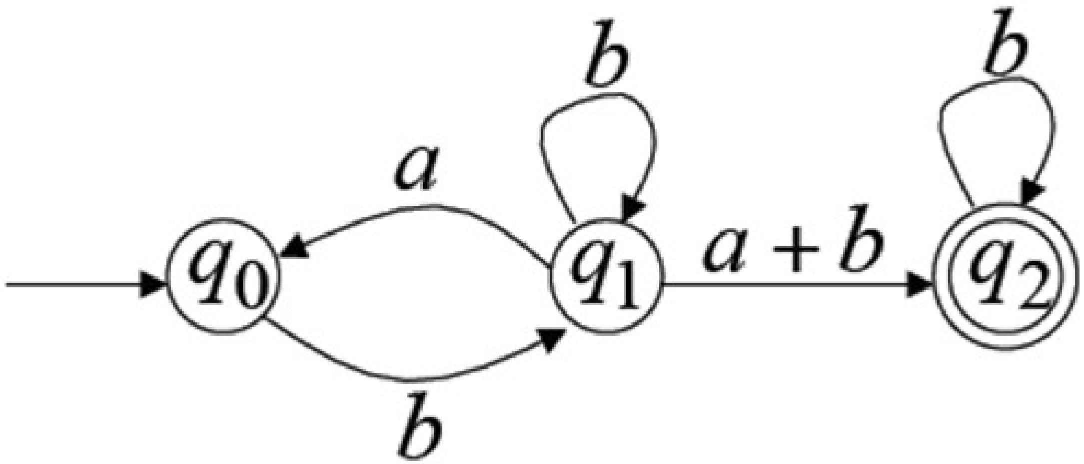

a
b
a
λ
λ
b

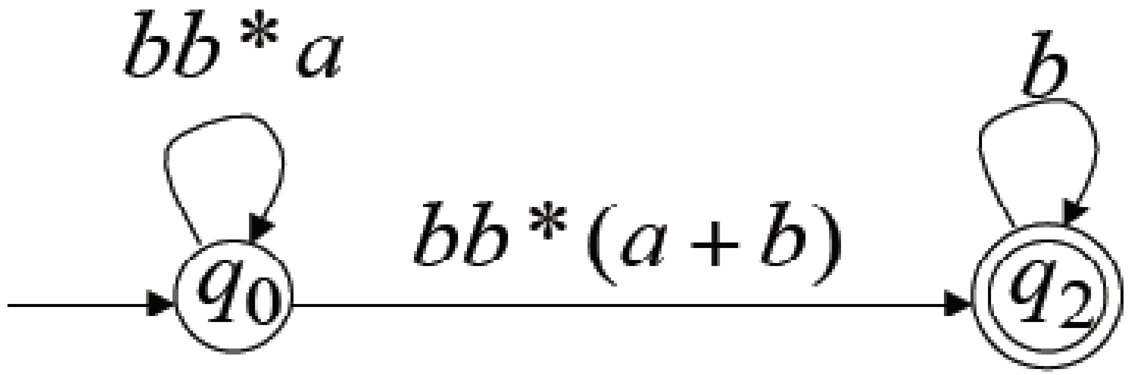

a*bb*a
a*bb*
λ


(a*bb*a)*a*bb*



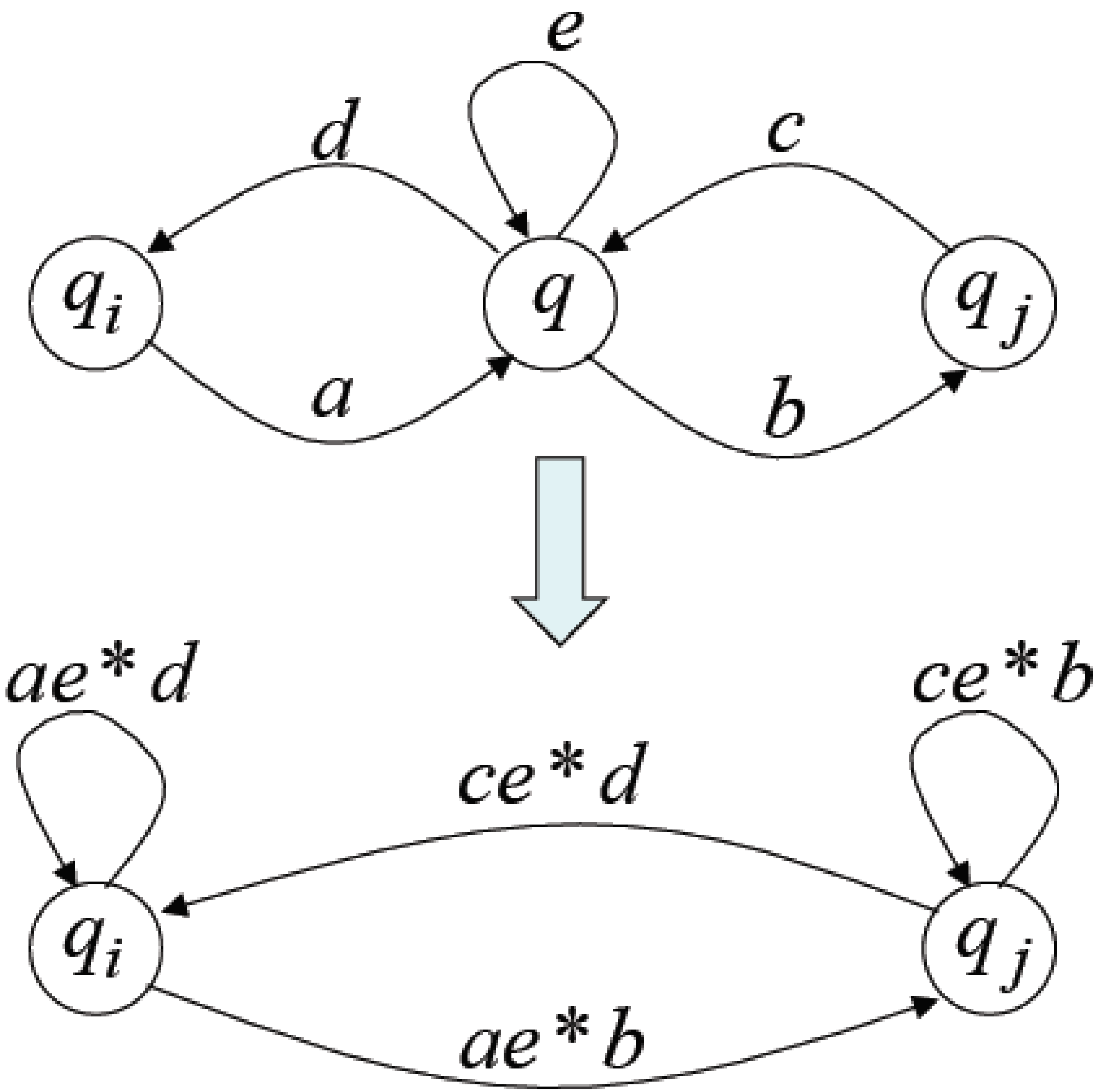
q_0
q_1
q_2
a
b
b
b
a,b
q_0
q_1
q_2
a
b
b
b
$a+b$



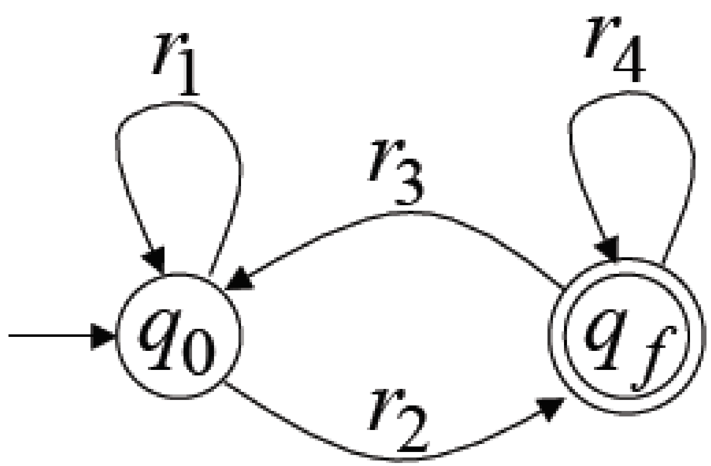


$$r = (bb^*a)^*bb^*(a+b)b^*$$

In General



The final transition graph:



The resulting regular expression:

$$r = r_1{}^* r_2 (r_4 + r_3 r_1{}^* r_2)^*$$

$$L(r) = L(M) = L$$

DÜZGÜN GRAMERLER

Düzgün dilleri tanımlamanın diğer bir yolu da basit gramerler kullanmaktadır. İlk olarak düzgün diller üreten gramerlere bakalım:

Sağdan ve Soldan Lineer Gramerler:

Bir G=(V, T, S, P) grameri A, B $\in$ V ve x $\in$ T* olmak üzere tüm türetim kuralları A→ xB | x şeklinde ise sağdan lineerdir.

G gramerinin soldan lineer olması için tüm türetim kuralları A→Bx | x şeklinde olmalıdır.

- Düzgün bir gramer ya sağdan ya da soldan lineerdir. Lineer olması için türetim kurallarının sağ tarafında en fazla bir tane değişken bulunur ve bu değişken sağ tarafın ya en sağındaki ya da en solundaki sembol olmalıdır.

- Örnek: $G_1 = (\{S\}, \{a,b\}, S, P_1)$

$S \rightarrow abS \mid a$ Sağdan lineerdir

$G_2 = (\{S, S_1, S_2\}, \{a,b\}, S, P_2)$

$$\left.\begin{array}{l} S \rightarrow S_1ab \\ S_1 \rightarrow S_1ab \mid S_2 \\ S_2 \rightarrow a \end{array}\right\}$$ Soldan lineerdir.

Her iki gramer de düzgün gramerlerdir.

G_1 grameri için, $S \Rightarrow abS \Rightarrow ababS \Rightarrow ababa$

$L(G_1)$ dili $r = (ab)^*a$ düzgün ifadesi ile gösterilir.

$L(G_2)$ dili de $r = a(ab)^*ab = a(ab)^+$

- Örnek: G = ({S, A, B}, {a, b}, S, P)

$S \rightarrow A$

$A \rightarrow aB \mid \lambda$

$B \rightarrow Ab$

=>

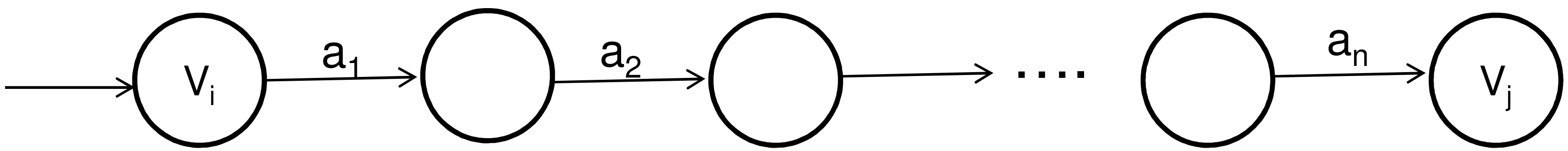
Bu gramer düzgün gramer değildir. Çünkü ne sağdan ne de soldan lineerdir.

Bu gramer ‘lineer gramer’e örnektir. Lineer gramer türetim kuralının sağ tarafında en fazla bir değişken olan gramerdir. Bu değişkenin yeri önemli değildir. Bu durumda düzgün bir gramer her zaman lineerdir, ancak bütün lineer gramerler düzgün değildir.

Sağdan lineer bir gramer tarafından türetilen dil her zaman düzgündür. Örneğin ab…cD => ab…cdE gibi bir türetimi D→dE türetim kuralı ile elde ettiğimizi varsayalım. Buna ait NFA'da gösterim D durumunda iken d sembolünü alarak E'ye gider. Bu durum aşağıdaki teoremin temelidir.

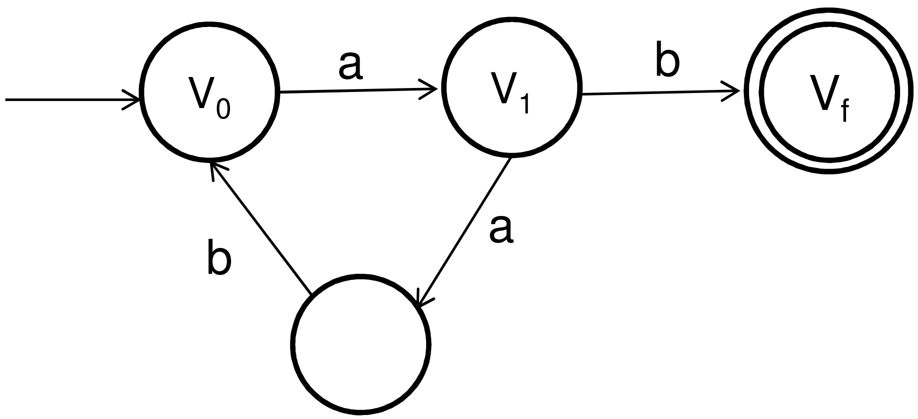
Teorem: G = (V, T, S, P) sağdan lineer bir gramer olsun. Bu durumda L(G) düzgün bir dildir. (Soldan lineer gramer için de geçerli)

$V_i \rightarrow a_1\ a_2 \ldots\ldots a_n V_j$



- Örnek: $V_0 \rightarrow aV_1$

  $V_1 \rightarrow abV_0 \mid b$ grameri ile türetilen dili kabul eden sonlu otomatı gösterin.



Bu gramer tarafından türetilen ve otomat tarafından kabul edilen dil L ((aab)*ab)

Teorem: $\sum$ alfabesi üzerinde tanımlı bir L dili düzgün ise L = L(G)'yi sağlayan sağdan lineer bir G = (V, $\sum$, S, P) grameri vardır. (Soldan lineer gramer için de geçerli)

Örnek: L(aab*a) için sağdan lineer bir gramer oluşturun.

$\delta(q_0, a) = \{q_1\}$	$q_0 \rightarrow aq_1$	$S \rightarrow aA$
$\delta(q_1, a) = \{q_2\}$	$q_1 \rightarrow aq_2$	$A \rightarrow aB$
$\delta(q_2, b) = \{q_2\}$	$q_2 \rightarrow bq_2$	$B \rightarrow bB$
$\delta(q_2, a) = \{q_f\}$	$q_2 \rightarrow aq_f$	$B \rightarrow aC$
$q_f \in F$	$q_f \rightarrow \lambda$	$C \rightarrow \lambda$

Örnek: $\sum = \{a, b\}$ alfabesi üzerinde en fazla üç a içeren tüm dizgileri kapsayan dili oluşturan düzgün bir gramer bulun.

$S \rightarrow bS \mid aA \mid \lambda$

$A \rightarrow bA \mid aB \mid \lambda$

$B \rightarrow bB \mid aC \mid \lambda$

$C \rightarrow bC \mid \lambda$

Örnek: Çift uzunluktaki dizgileri içeren dili oluşturan düzgün grameri bulun.

$S \rightarrow aaS \mid bbS \mid abS \mid baS \mid \lambda$

- Verilen bir gramerden bu gramerin hangi dili ürettiğini söyleyebilmeliyiz:

Örnek: $S \rightarrow aaSB \mid \lambda$

$B \rightarrow bB \mid b$

S => aaSB => aaB => aab

S => aaSB => aaB => aabB => aabb

S => aaSB => aaB => aabB => aabbB => aabbb

S=>aaSB=>aaB=>aabB=>aabbB=>aabbbB=>aabbbb

S=>aaSB=>aaaaSBB=>aaaaBB=>aaaaBb=>aaaabb

S=>aaSB=>aaaaSBB=>aaaaBB=>aaaaBbB=>

aaaaBbb=>aaaabbb

$L = \{ (aa)^n \, b^n \, b^* \} \; n \geq 0$

- Verilen bir dilden, bu dili üreten gramerin bulunması:

Örnek: {a, b, c} üzerinde, a ile başlayan, sadece iki b içeren ve cc ile biten tüm dizgileri içeren dile ait düzgün (sağdan lineer) bir gramer bulun.

S → aA

A → bB | aA | cA

B → bC | aB | cB

C → aC | cC | cD

D → c