CS441/CS241 Automata Theory and Formal Languages

ربيع 2024

د. عدنان محمود عبدالله الشريف
adnan.sherif@uot.edu.ly

1

التعابير المنتظمة و اللغات المنتظمة
Regular Expressions and Languages

- ذكرنا سابقا ان اللغة هي مجموعة السلاسل المختارة من المجموعة Σ^* والمتكونة من الابجدية Σ ونرمز لها بالرمز (L) بحيث $L \subseteq \Sigma^*$ و يتم وصف اللغة بتحديد شروط تكوين السلاسل التي تنتمي الى اللغة.
- كما درسنا بعض العمليات على اللغات وهي:
  - اتحاد لغتين (Union): اتحاد لغتين L_1 و L_2 ويمثل بـ $L_1 \cup L_2$ هي اللغة التي تحتوي على السلاسل الموجودة في L_1 او L_2 او كلاهما:
  $$L_1 \cup L_2 = \{x | x \in L_1 \vee x \in L_2\}$$
  - تعاقب لغتين (Concatenation): تعاقب لغتين L_1 و L_2 ويمثل بـ $L_1.L_2$ هي اللغة التي تحتوي على السلاسل المشكلة من تعاقب السلاسل في L_1 بجميع السلاسل في L_2 :
  $$L_1 \cdot L_2 = \{x \cdot y | x \in L_1 \wedge y \in L_2\}$$
  - إغلاق لغة (Closure): إغلاق اللغة L المعرفة على الابجدية Σ هي اللغة الناتجة عن تعاقب كل سلاسل اللغة L ونرمز لإغلاق اللغة L بـ L^*:
  $$L^* = L^0 \cup L^1 \cup L^2 \cup L^3 \ldots$$

2

التعابير المنتظمة و اللغات المنتظمة
Regular Expressions and Languages

- تعريف اللغة المنتظمة: يمكن تعريف اللغة المنتظمة على ابجدية Σ بالتكرار كالتالي:
  - $\{\varepsilon\}$ هي لغة منتظمة على الابجدية Σ.
  - إذا كان **a** رمز من رموز الابجدية Σ عندئذ تكون $\{a\}$ لغة منتظمة على الابجدية Σ.
  - إذا كانت L لغة منتظمة على الابجدية Σ فإن L^* و L^n لغات منتظمة على الابجدية Σ.
  - إذا كانت L_1 و L_2 لغات منتظمة فإن $L_1 \cup L_2$ و $L_1 \cap L_2$ و $L_1 \cdot L_2$ (L_1L_2) لغات منتظمة على الابجدية Σ.
  - لا توجد لغات منتظمة أخرى.
- نلاحظ ان كل اللغات المنتظمة يمكن تكوين اوتومات منتهية لتمثيلها و التعرف على السلاسل التي تتبع اللغة التي تمثلها.

التعابير المنتظمة و اللغات المنتظمة
Regular Expressions and Languages

- التعابير المنتظمة هي طريقة جبرية للتعبير عن اللغات المنتظمة.
- التالي تعريف التعابير المنتظمة بقواعد جبرية حيث الدالة L ترجع اللغة التي تمثلها التعبير المنتظم:
  - $\emptyset$ هو التعبير المنتظم الذي يحدد اللغة الفارغة
    $$L(\emptyset) = \emptyset = \{\}$$
  - ε هو التعبير المنتظم الذي يحدد اللغة التي تحتوي السلسلة الفارغة
    $$L(\varepsilon) = \{\varepsilon\}$$
  - لأي رمز a من الابجدية Σ فإن a هو التعبير المنتظم الذي يحدد اللغة
    $$L(a) = \{a\}$$
  - إذا كان E و F تعابير منتظمة فإن كل من التالي تعابير منتظمة

$$L(E+F) = L(E) \cup L(F)$$
$$L(EF) = L(E) \cdot L(F)$$
$$L(E^*) = \big(L(E)\big)^*$$
$$L\big((E)\big) = L(E)$$

الأولوية

التعبير المنتظم التالي:

$$01^* + 1 \equiv (0(1)^*) + 1$$

التعابير المنتظمة واللغات المنتظمة
Regular Expressions and Languages

مثال 1: اكتب التعبير المنتظم للغة التي تمثل الاعداد الثنائية بطول 3 ثم بين فئة السلاسل التي تنتمي الى هذه اللغة.

الحل: من المعطيات نستنتج التالي:

1. الابجدية لتمثيل الأرقام الثنائية $\Sigma = \{0,1\}$
2. اللغة التي تمثل الاعداد الثنائية بطول 3 هي

$$L = \{000,001,010,011,100,101,110,111\}$$

نلاحظ ان اللغة المطلوب تكوين تعبير منتظم لها تتكون من سلاسل تحتوي على 3 خانات وكل خانة يمكن ان تكون 0 او 1 وعليه التعبير المنتظم الذي يعبر عن هذه اللغة هو

$$(0+1)^3$$

او

$$(0+1)(0+1)(0+1)$$

التعابير المنتظمة واللغات المنتظمة
Regular Expressions and Languages

مثال 1(تابع): يمكن إيجاد الفئة المطلوبة بتطبيق الدالة L على التعبير المنتظم باستخدام القواعد الجبرية كالتالي:

$$L\big((0+1)(0+1)(0+1)\big) \equiv$$
$$L\big((0+1)\big)\cdot L\big((0+1)(0+1)\big) \equiv$$
$$L(0+1)\cdot L\big((0+1)(0+1)\big) \equiv$$
$$\big(L(0)\cup L(1)\big)\cdot L\big((0+1)(0+1)\big) \equiv$$
$$(\{0\}\cup\{1\})\cdot L\big((0+1)(0+1)\big) \equiv$$
$$\{0,1\}\cdot L\big((0+1)(0+1)\big) \equiv$$
$$\{0,1\}\cdot L\big((0+1)\big)\cdot L\big((0+1)\big) \equiv$$
$$\{0,1\}\cdot\{0,1\}\cdot\{0,1\} \equiv$$
$$\{0,1\}\cdot\{00,01,10,11\} \equiv$$
$$\{000,001,010,011,100,101,110,111\}\ \#$$

التعابير المنتظمة و اللغات المنتظمة
Regular Expressions and Languages

مثال 2: اكتب التعبير المنتظم للغة على الابجدية $\Sigma = \{0,1\}$ بحيث كل سلسلة في اللغة عبارة عن 0 يتبعه 1 او العكس 1 يتبعه 0.

الحل: من المعطيات نستنتج التالي:

1. الابجدية $\Sigma = \{0,1\}$
2. السلسلة التي تنتمي لهذه اللغة يمكن ان تبدأ 0 وتنتهي بـ 1 او0 كما يمكن ان تبدأ بـ1 وتنتهي بـ 0 او 1

يمكن ان نكون التعبير المنتظم للحالة الأساسية 01 او 10 كالتالي

$$(01) + (10)$$

ثم يمكن إضافة معامل الاغلاق لتكرار السلسلة بحيث يصبح التعبير المنتظم

$$(01)^* + (10)^*$$

لكن ممكن ان تبدأ بـ 0 وتنتهي بـ 0 او تبدأ بـ 1 وتنتهي بـ1 ويمكن تعديل التعبير المنتظم كالتالي

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

التعابير المنتظمة و اللغات المنتظمة
Regular Expressions and Languages

مثال 2(تابع): يمكن إيجاد الفئة المطلوبة بتطبيق الدالة L على التعبير المنتظم باستخدام القواعد الجبرية كالتالي:

$$L\big((01)^* + (10)^* + 1(01)^* + 0(10)^*\big) \equiv$$

$$L((01)^*) \cup L\big((10)^*\big) \cup L(1(01)^*) \cup L(0(10)^*) \equiv$$

$$\{01\}^* \cup \{10\}^* \cup L(1) \cdot L((01)^*) \cup L(0) \cdot L((10)^*) \equiv$$

$$\{01\}^* \cup \{10\}^* \cup \{1\} \cdot \{01\}^* \cup \{0\} \cdot \{10\}^* \equiv$$

$$\{\varepsilon, 01,0101,010101, \dots\} \cup \{\varepsilon, 10,1010,101010, \dots\} \cup \{1\} \cdot \{01\}^* \cup \{0\} \cdot \{10\}^* \equiv$$

$$\{\varepsilon, 01,0101,010101, \dots\} \cup \{\varepsilon, 10,1010,101010, \dots\} \cup \{1,101,10101,1010101 \dots\} \cup \{0,010,01010,0101010 \dots\} \equiv$$

$$\{\varepsilon, 0,1,01,10,101,010,0101,1010,10101,01010 \dots\}\#$$

التعابير المنتظمة و اللغات المنتظمة
Regular Expressions and Languages

مثال 3: بين اللغة التي تمثلها التعابير المنتظمة التالية:

1. لغة على الابجدية $\Sigma = \{0,1\}$ بحيث $(0+10)^*1^*$
2. لغة على الابجدية $\Sigma = \{0,1\}$ بحيث $(0+1)^*(1+0)^*$
3. لغة على الابجدية $\Sigma = \{0,1\}$ بحيث $(0^*1^*)^*000(0+1)^*$

الحل: (1) باستخدام الدالة L

$$L\big((0+10)^*1^*\big) \equiv L\big((0+10)^*\big) \cdot L(1^*) \equiv \big(L(0+10)\big)^* \cdot L(1)^*$$
$$\equiv \big(L(0) \cup L(10)\big)^* \cdot \{1\}^* \equiv \Big(\{0\} \cup \big(L(1) \cdot L(0)\big)\Big)^* \cdot \{\varepsilon, 1,11,111, \dots\}$$
$$\equiv \big(\{0\} \cup (\{1\} \cdot \{0\})\big)^* \cdot \{\varepsilon, 1,11,111, \dots\} \equiv (\{0\} \cup \{10\})^* \cdot \{\varepsilon, 1,11,111, \dots\}$$
$$\equiv \{0,10\}^* \cdot \{\varepsilon, 1,11,111, \dots\}$$
$$\equiv \{\varepsilon, 0,10,00,010,100,1010,000,0100,1000,10100,0010,0010,01010,10010,101010 \dots\}$$
$$\cdot \{\varepsilon, 1,11,111, \dots\}$$
$$\equiv \{\varepsilon,0,01,011,0111,\dots,10,101,1011,10111, \dots, 00,001,0011,00111, \dots\}$$

التعابير المنتظمة و اللغات المنتظمة
Regular Expressions and Languages

مثال 3 تابع: (2) باستخدام الدالة L

$$L\big((0+1)^*(1+0)^*\big) \equiv L\big((0+1)^*\big) \cdot L\big((1+0)^*\big)$$
$$\equiv L\big((0+1)\big)^* \cdot L\big((1+0)\big)^* \equiv \big(L(0+1)\big)^* \cdot \big(L(1+0)\big)^*$$
$$\equiv \big(L(0) \cup L(1)\big)^* \cdot \big(L(1) \cup L(0)\big)^* \equiv (\{0\} \cup \{1\})^* \cdot (\{1\} \cup \{0\})^*$$
$$\equiv (\{0,1\})^* \cdot (\{0,1\})^*$$
$$\equiv \{\varepsilon, 0,1,00,01,10,11,000,001,010,011,100,101,110,111, \dots\}$$
$$\cdot \{\varepsilon, 0,1,00,01,10,11,000,001,010,011,100,101,110,111, \dots\}$$
$$\equiv \{\varepsilon, 0,1,00,01,10,11,000,001,010,011,100,101,110,111, \dots\}$$

نلاحظ ان هذه اللغة هي جميع الاعداد الثنائية.

فقرة 3 من المثال تمرين في المحاضرة لديكم 7 دقائق

التعابير المنتظمة و اللغات المنتظمة
Regular Expressions and Languages

• قواعد جبرية للتعابير المنتظمة حيث R,P و Q تعابير منتظمة

القاعدة	رقم	القاعدة	رقم
$R(P+Q) = RP + RQ$	9	$\emptyset + R = R$	1
$(R+P)Q = RQ + PQ$	10	$\emptyset R = R\emptyset = \emptyset$	2
$R^*R^* = R^*$	11	$\varepsilon R = R\varepsilon = R$	3
$R^*R = RR^* = R^+$ حيث $R^* = (\varepsilon + R^+)$	12	$\varepsilon^* = \varepsilon = \emptyset^*$	4
$(R^*)^* = R^*$	13	$R + R = R$	5
$\varepsilon + R^* = R^*$	14	$R + Q = Q + R$	6
$(RP)^*R = R(PR)^*$	15	$R + (P+Q) = (R+P) + Q$	7
$(R+P)^* = (R^*P^*)^* = (R^* + P^*)^*$	16	$R(PQ) = (RP)Q$	8

التعابير المنتظمة و اللغات المنتظمة
Regular Expressions and Languages

• نظرية اردن (Ardins Theorem) : في حال ان P و Q تعابير منتظمة على نفس الابجدية Σ بحيث P لا تحتوي على السلسلة الفارغة ε فإن التعبير المنتظم على الصورة

$$R = Q + RP$$

يمكن استبداله بالتعبير المنتظم

$$R = QP^*$$

الاثبات: من المعادلة الاصلية والتعويض عن R في الطرف الأيمن نجد ان

$$R = Q + RP$$

$$R = Q + (Q + RP)P = Q + QP + RP^2$$

$$R = Q + QP + (Q + RP)P^2 = Q + QP + QP^2 + RP^3$$

بعد التعويض n من المرات نحصل على

$$R = Q + QP + QP^2 + QP^3 + \cdots + QP^n + RP^{n+1}$$

نلاحظ ان هذه المتوالية تستمر الى ما لانهاية بنفس الشكل وعليه بأخذ Q عامل مشترك نحصل على

$$R = Q(\varepsilon + P + P^2 + P^3 + \cdots + P^n + P^{n+1} + \cdots)$$

$$R = QP^* \quad \#$$

التعابير المنتظمة و اللغات المنتظمة
Regular Expressions and Languages

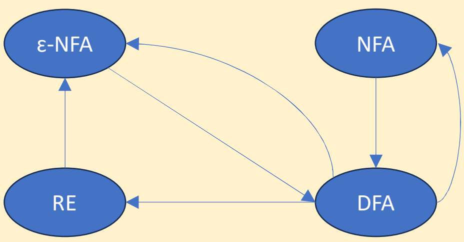
- مثال 4: استخدم القواعد الجبرية لاختصار التعابير المنتظمة في المثال رقم 3

الحل:

القاعدة المستخدمة	التعبير المنتظم	رقم
باستخدام القاعدة رقم 16	$(0 + 10)^*1^*$ $(0^*(10)^*)^*1^*$	1
باستخدام القاعدة رقم 6 باستخدام القاعدة رقم 11	$(0 + 1)^*(1 + 0)^*$ $(0 + 1)^*(0 + 1)^*$ $(0 + 1)^*$	2
باستخدام القاعدة رقم 16	$(0^*1^*)^*000(0 + 1)^*$ $(0 + 1)^*000(0 + 1)^*$	3

الاوتومات المنتهية و التعابير المنتظمة
Finite Automata and Regular expressions

- كل لغة يمكن وصفها باستخدام احدى الاوتومات المنتهية يمكن تكوين تعبير منتظم لها. وكل تعبير منتظم له اوتومات يمثله.

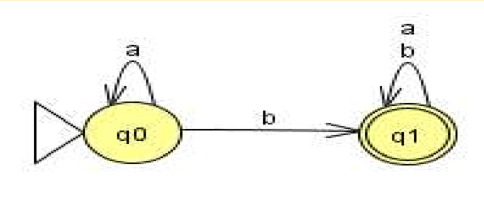


الاوتومات المنتهية و التعابير المنتظمة
Finite Automata and Regular expressions

- خطوات إيجاد التعبير المنتظم لاوتوماتا منتهية حتمية (DFA) :
  1. قم بإنشاء معادلة لكل حالة بالشروط التالية:
     1. حالة البداية يتم إضافة ε
     2. لكل سهم داخل الى الحالة يتم إضافة رمز الحالة القادم منها السهم تم الرمز (الحرف) على السهم

مثال

$$q_0 = \varepsilon + q_0 a$$
$$q_1 = q_0 b + q_1 a + q_1 b$$



15

الاوتومات المنتهية و التعابير المنتظمة
Finite Automata and Regular expressions

- خطوات إيجاد التعبير المنتظم لاوتومات منتهية حتمية (DFA) تابع :
  2. يتم اختصار المعادلات باستخدام القواعد الجبرية حتى نتحصل على معادلات لكل حالة بدون رموز الحالات في الطرف الأيمن

$$q_0 = \varepsilon + q_0 a$$
$$= \varepsilon a^*$$
$$= a^*$$
$$q_1 = q_0 b + q_1 a + q_1 b$$
$$= a^* b + q_1 a + q_1 b$$
$$= a^* b + q_1 (a + b)$$
$$= a^* b\ (a + b)^*$$

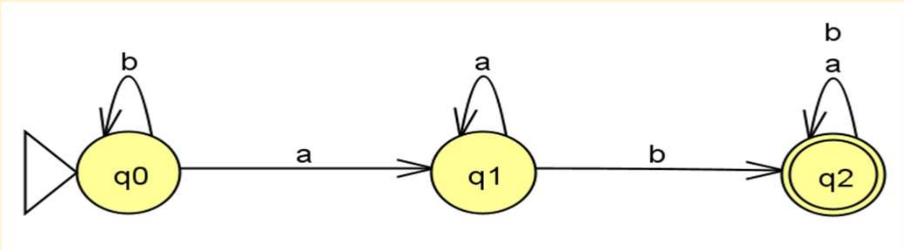
  3. التعبير المنتظم في حالة النهاية هو التعبير المنتظم المكافئ للاوتومات. في حال وجود أكثر من حالة نهاية يتم الدمج بينهم باستخدام +

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

16

الاوتومات المنتهية والتعابير المنتظمة
Finite Automata and Regular expressions

- مثال 5: اوجد التعبير المنتظم للاوتومات المنتهية الحتمية التالية:



الاوتومات المنتهية والتعابير المنتظمة
Finite Automata and Regular expressions

- مثال 5: الحل:

1. نكون المعادلات لكل حالة

$$q_0 = \varepsilon + q_0\mathrm{b}$$
$$q_1 = q_0a + q_1a$$
$$q_2 = q_1b + q_2a + q_2b$$

2. اختصار المعادلات للتخلص من q_0,q_1 و q_2 من الطرف الأيمن

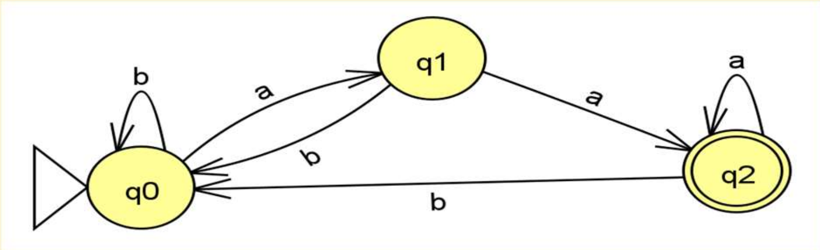
$$q_0 = \varepsilon + q_0\mathrm{b} = \varepsilon\mathrm{b}^* = \mathrm{b}^*$$
$$q_1 = q_0a + q_1a = b^*a + q_1a = b^*aa^*$$
$$q_2 = q_1b + q_2a + q_2b = b^*aa^*b + q_2a + q_2b = b^*aa^*b + q_2(a+b) = b^*aa^*b(a+b)^*$$

3. حالة النهاية هي q_2 عليه التعبير المنتظم الممثل لهذه الاوتوماتا هو

$$b^*aa^*b(a+b)^*$$

الاوتومات المنتهية والتعابير المنتظمة
Finite Automata and Regular expressions

- مثال 6: اوجد التعبير المنتظم للاوتومات المنتهية الحتمية التالية:



19

الاوتومات المنتهية والتعابير المنتظمة
Finite Automata and Regular expressions

- مثال 6: الحل:

1. نكون المعادلات لكل حالة

$$q_0 = \varepsilon + q_0\mathrm{b} + q_1\mathrm{b} + q_2\mathrm{b}$$
$$q_1 = q_0 a$$
$$q_2 = q_1 a + q_2 a$$

2. اختصار المعادلات للتخلص من q_0, q_1 و q_2 من الطرف الأيمن

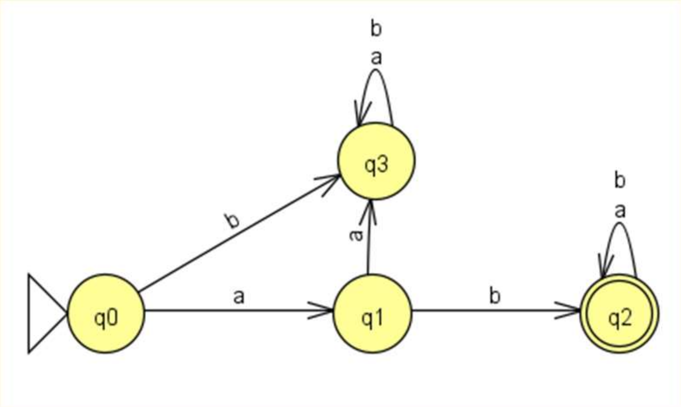
$$q_1 = q_0 a$$
$$q_2 = q_1 a + q_2 a = q_0 aa + q_2 a = q_0 aaa^*$$
$$q_0 = \varepsilon + q_0\mathrm{b} + q_1\mathrm{b} + q_2\mathrm{b}$$
$$q_0 = \varepsilon + q_0\mathrm{b} + q_0 a\mathrm{b} + q_0 aaa^*\mathrm{b}$$
$$q_0 = \varepsilon + q_0(\mathrm{b} + a\mathrm{b} + aaa^*\mathrm{b})$$
$$q_0 = \varepsilon(\mathrm{b} + a\mathrm{b} + aaa^*\mathrm{b})^* = (\mathrm{b} + a\mathrm{b} + aaa^*\mathrm{b})^*$$

بالتعويض في q_2 بقيمة q_0 نحصل على التالي $q_2 = q_0 aaa^* = (\mathrm{b} + a\mathrm{b} + aaa^*\mathrm{b})^* aaa^*$

20

الاوتومات المنتهية و التعابير المنتظمة
Finite Automata and Regular expressions

- مثال 7: اوجد التعبير المنتظم للاوتومات المنتهية الحتمية التالية:



الاوتومات المنتهية و التعابير المنتظمة
Finite Automata and Regular expressions

- مثال 7: الحل:

1. نكون المعادلات لكل حالة

$$q_0 = \varepsilon$$
$$q_1 = q_0 a$$
$$q_2 = q_1 b + q_2 a + q_2 b$$

2. اختصار المعادلات للتخلص من q_0,q_1 و q_2 من الطرف الأيمن

$$q_0 = \varepsilon$$
$$q_1 = q_0 a = \varepsilon a = a$$
$$q_2 = q_1 b + q_2 a + q_2 b = ab + q_2 a + q_2 b = ab + q_2(a + b) = ab(a + b)^*$$

3. حالة النهاية هي q_2 عليه التعبير المنتظم الممثل لهذه الاوتوماتا هو

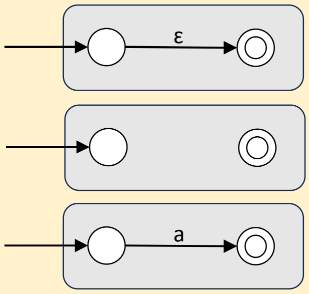
$$ab(a + b)^*$$

ملاحظة : عدم استخدام حالة الموت او الرفض q_3 . لما لم يتم استخدامها؟

الاوتومات المنتهية والتعابير المنتظمة
Finite Automata and Regular expressions

- قواعد الحصول على الاوتومات المنتهية اللاحتمية بحركة ε من تعبير منتظم

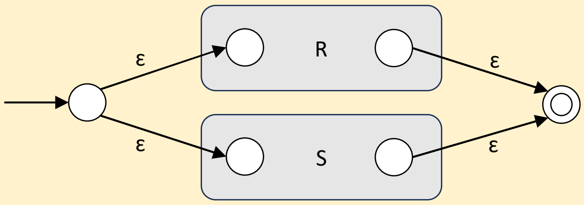
للوصول الى هذه الغاية يتم تكوين اوتوماتا لكل من القواعد الجبرية التي تم تعريفها للتعابير المنتظمة كالتالي



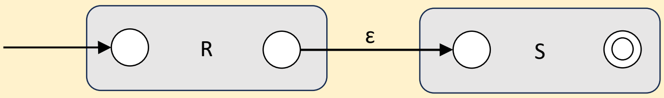
1. السلسلة الفارغة ε يتم تكوين اوتوماتا به حركة واحدة هي ε من حالة البداية الى حالة النهاية
2. $\emptyset$ وهذا يمثل اللغة الفارغة أي لا تحوي على سلاسل ويمكن تمثيلها بحالة البداية والنهاية فقط بدون حركة
3. a وهذا يمثل اللغة التي بها كلمة واحدة هي الرمز a ويمكن تمثيلها باوتوماتا به حركة واحدة من حالة البداية الى حالة النهاية بـa

الاوتومات المنتهية والتعابير المنتظمة
Finite Automata and Regular expressions

- قواعد الحصول على الاوتومات المنتهية اللاحتمية بحركة ε من تعبير منتظم (تابع)



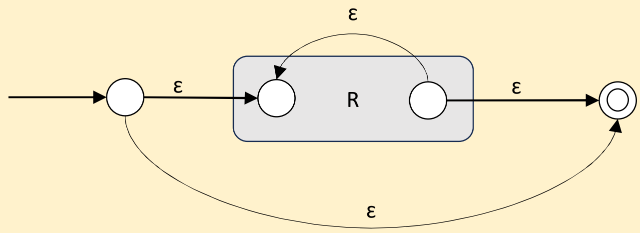
4. R+S



5. RS

الاوتومات المنتهية والتعابير المنتظمة
Finite Automata and Regular expressions

- قواعد الحصول على الاوتومات المنتهية اللاحتمية بحركة ε من تعبير منتظم (تابع)



6. R^*

الاوتومات المنتهية والتعابير المنتظمة
Finite Automata and Regular expressions

- مثال 8: كون الاوتومات المنتهية اللاحتمية بحركة ε للتعابير منتظم التالية:

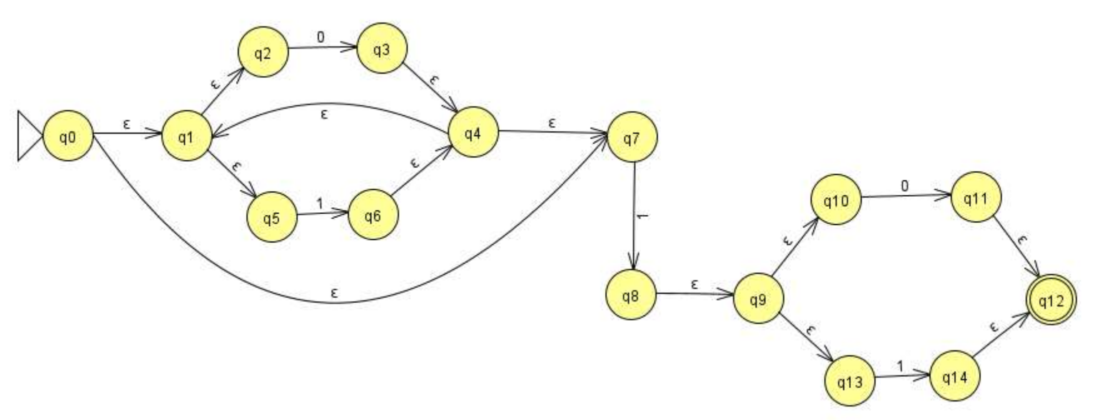
1. $(0+1)^*1(0+1)$

2. $(01)^*1$

3. $(0^*+1^*)^*$

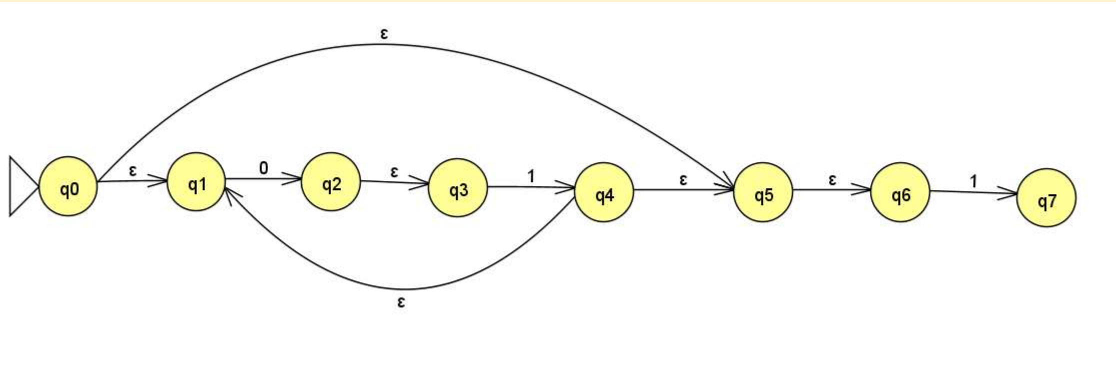
الاوتومات المنتهية والتعابير المنتظمة
Finite Automata and Regular expressions

- مثال 8 (1) الحل:



الاوتومات المنتهية والتعابير المنتظمة
Finite Automata and Regular expressions

- مثال 8 (2) الحل:



الاوتومات المنتهية والتعابير المنتظمة
Finite Automata and Regular expressions

• مثال 8 (3) الحل: