

Compilers Techniques

Lecture 5

المحاضرة الخامسة

Lexical Analyser exercises

تمارين خاصة بالتحليل المعجمي

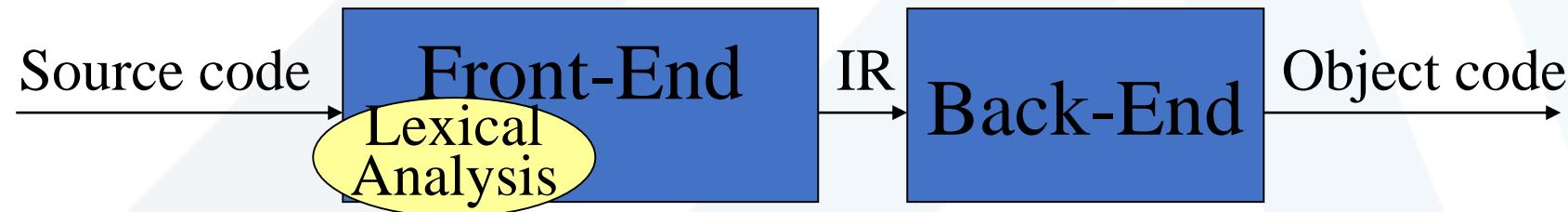
Reference: Aho2 Sections 3.6-3.7; Aho1 pp. 113-125; Grune 2.1.6.1-2.1.6.6 (different style); Hunter 3.3 (very condensed); Cooper1 2.4-2.4.3

السنة الرابعة – المستوى السابع - الهندسة المعلوماتية

تمارين على التحليل المmorphي

Lectures Topics

تمارين عن التعبير المنتظمة - NFA – DFA - الاختصار -



تمرین رقم 1 استجابة المحلل المعجمي لسلسل الدخل

لديك سلسل الدخل التالي ووضح كيفية استجابة المحلل المعجمي (الماسح) له أو وضح خرج المحلل المعجمي!

```
For(int i=0;i<=n;i++)  
    x=x+10;
```

الحل:

```
<FOR><PAR,><INT><ID,i><ASSIGN,=><NUM,0><SEMI,;><ID,i>  
<ASSIGN,<=><ID,n><SEMI,;><ID,i><DPLUS,++><PAR,)><ID,x><ASIGN,=>  
<ID,x><PLUS,+><NUM,10><SEMI,;>
```

ماذا عن العبارة

if (X==10) then if (Y==-1) then z=1; else z=2;

ولد سلاسل أو تطابقات صحيحة وفق التعبير المنتظم RE التالي

تمرين رقم 2

التعبير المنتظم :RE

`\d{4}[-,]? \d{4}[-,]? \d{4}[-,]? \d{4}`

علمًاً أن `d` هي تعبير منتظم فرعي يدل على الأرقام من 0 وحتى 9. $d: [0-9]$.

الجزء `\d{4}` يدل على أربع أرقام بغض النظر عن ماهية هذه الأرقام (أي رقم في المجال من 0 إلى 9).

الجزء `? [-]` يدل على وجود جزء بعد الأرقام الأربعية وهو إما إشارة - أو الفاصلة أو فراغ أو لا شيء أبداً.

يتم بعدها تكرار الجزأين السابقين متتاليتين ثم يتم الانتهاء بالجزء `\d{4}` يدل التعبير المنتظم السابق على 16 حرف (أربع مجموعات يفصل بين كل مجموعتين منها إما - أو فاصلة أو فراغ أو تكون غير مفصولة).

أعط مثال على سلسلة تطابق هذا التعبير المنتظم:

يطلب توليد تعبير منظم RE يمكن من خلاله كتابة سلاسل تعبر عن ايميل أو بريد الكتروني.

علمًا أن البريد الإلكتروني يبدأ بحرف صغير أو كبير يليه حرف أو أكثر ويمكن أن يتخلل الحروف أرقام أو رموز مثل إشارة + أو النقطة أو إشارة _ أو إشارة -

يلي جزء الاسم إشارة @ يلي إشارة @ الجزء الثاني من الايميل وهو اسم المجال مثلاً gmail أو yahoo وكذلك هو اسم يبدأ بحرف صغير أو كبير ويليه حرف أو أكثر يمكن أن يتخلله نقطة . أو إشارة - فقط يلي اسم المجال رمز النقطة .

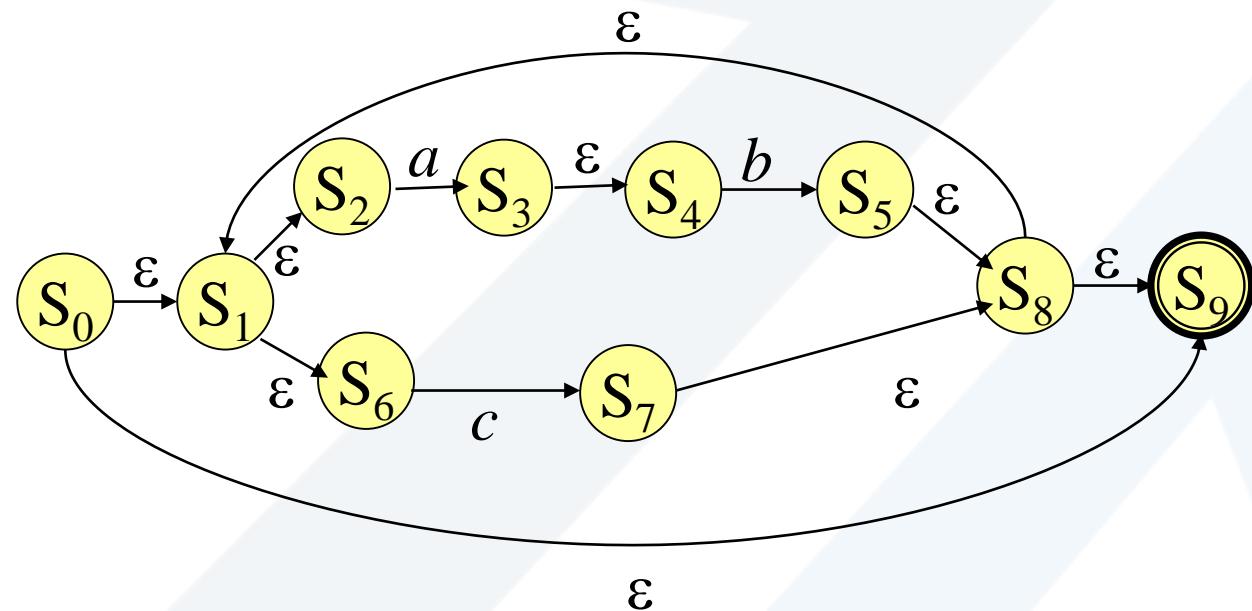
وبعدها اللاحقة التي تعبر عن نوع أو بلد الموقع مثل .com أو .sy وهذا يحدد عدد الأحرف بمجال بين حرفين إلى 5 أحرف.

`[a-zA-Z0-9.+_-]@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,5}`

ابحث في بناء تعبير منظم يعبر عن وسم HTML الخاص بتكوين مقطع أو paragraph
`<p>....</p>`

ولد مخطط NFA للتعبير المنظم

$$((ab)|c)^*$$



يمكنك رسم مخططات NFA أخرى غير الموصوف أعلاه بشرط أن يطابق ذات التعبير المنظم المطلوب!

سؤال شامل

صمم تعبيراً منتظماً لكشف السلسلة الأطول في الأبجدية {a-z} مستعيناً بالخصائص التالية:

- 1) تبدأ السلسلة بحرف a وتنتهي بحرف z.
- 2) بعد الحرف a، يمكن أن تتضمن السلسلة أحد خيارين اثنين. الأول يبدأ بحرف b متبوع بصفر نسخة أو أكثر من 's' وينتهي بحرف 'e' أما الثاني فيبدأ بحرف c لكنه لا يتضمن أي حرف 's' وينتهي بحرف 't'.
- 3) في أي نقطة من التسلسلين السابقين، وعند ورود حرف x يتم اعتبار التسلسل متوقفاً أو متنياً وبمعنى آخر إن الرمز x يعتبر كأنه الرمز 'e' في التسلسل الأول وكالرمز 't' في الثاني أي الرمز الذي تنتهي به السلسلة.
- 4) كلا التسلسلين غير متداخلين لكن يمكن تكرارهما بأي ترتيب.

مثلاً التسلسل **abssseccaaabbba~~f~~ghtbz_ez** صحيح
التعبير **abssxcfftz** أيضاً صحيح.

الحل:

$$RE = a \cdot [(b \cdot (s^*) \cdot (x|e)) \mid (c \cdot (\text{notSTX}^*) \cdot (x | t))]^+ \cdot z$$

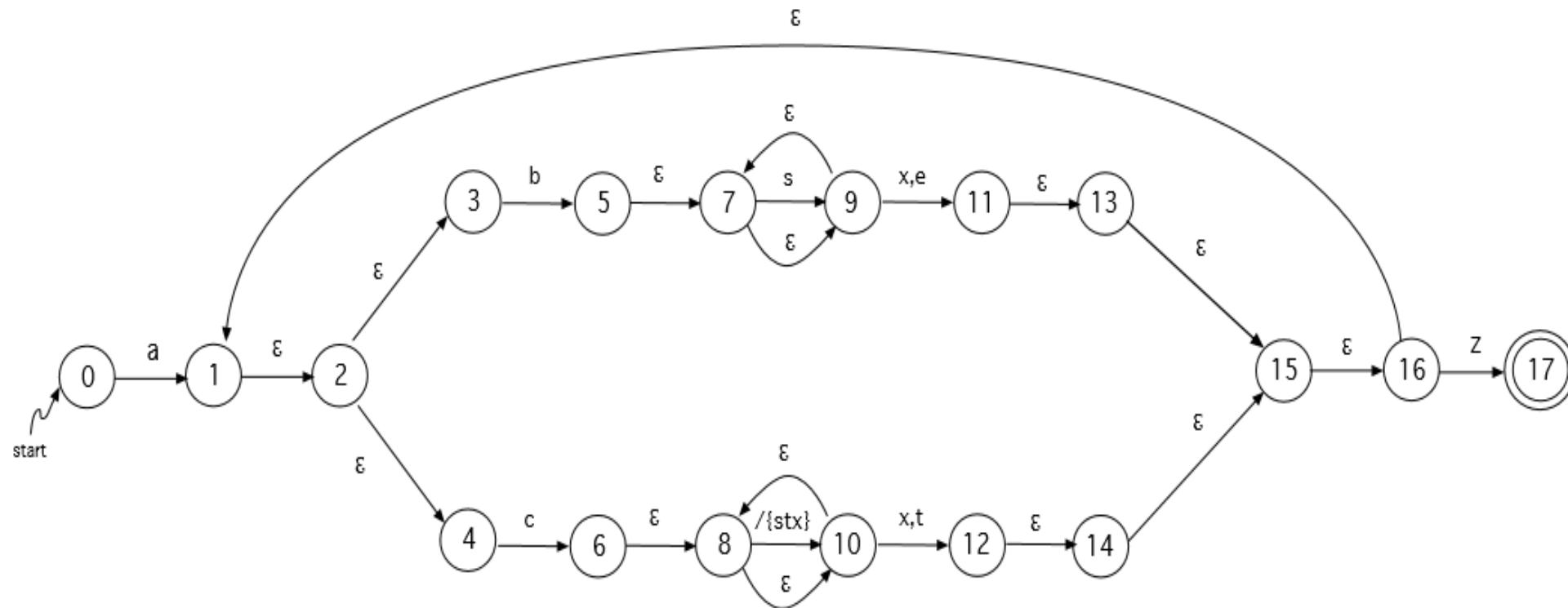
حيث:

$$\text{notSTX} = \{a-r, u, v, w, y, z\}$$

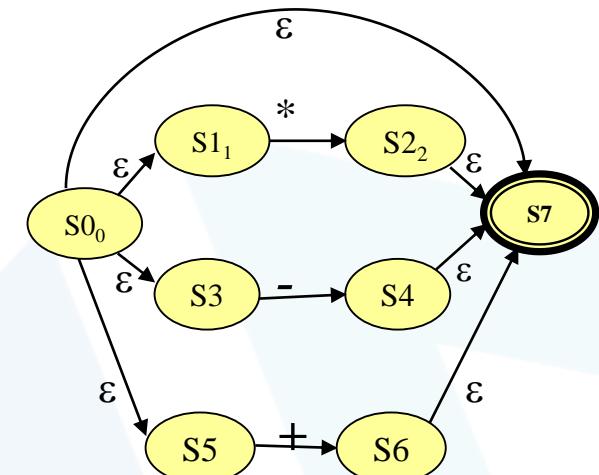
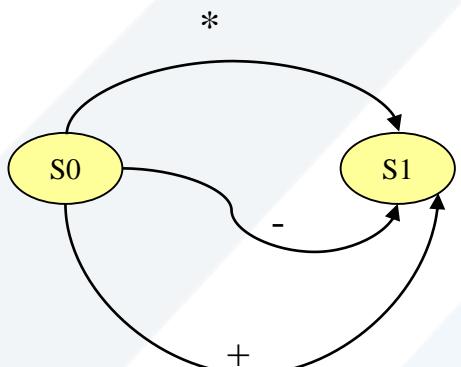
سؤال شامل

تمرين رقم 5

باستخدام التعبير المنظم الذي قمت بتأليده في الطلب السابق، قم باشتقاق مخطط NFA الموافق له مستخدماً طريقة طومسون



	ϵ -closure(mov e($s_i, -$))	ϵ -closure(mov e($s_i, +$))	ϵ -closure(mov e($s_i, *$))
$A = \{0, 1, 3, 5, 7\}$	ϵ -closure(4) = {7} B	ϵ -closure(6) = {7} B	ϵ -closure(2) = {7} B
$B = \{7\}$	-	-	-



تمرين رقم 5

سؤال شامل

باستخدام طريقة بناء المجموعات الجزئية حول مخطط NFA السابق إلى مخطط DFA الموافق.

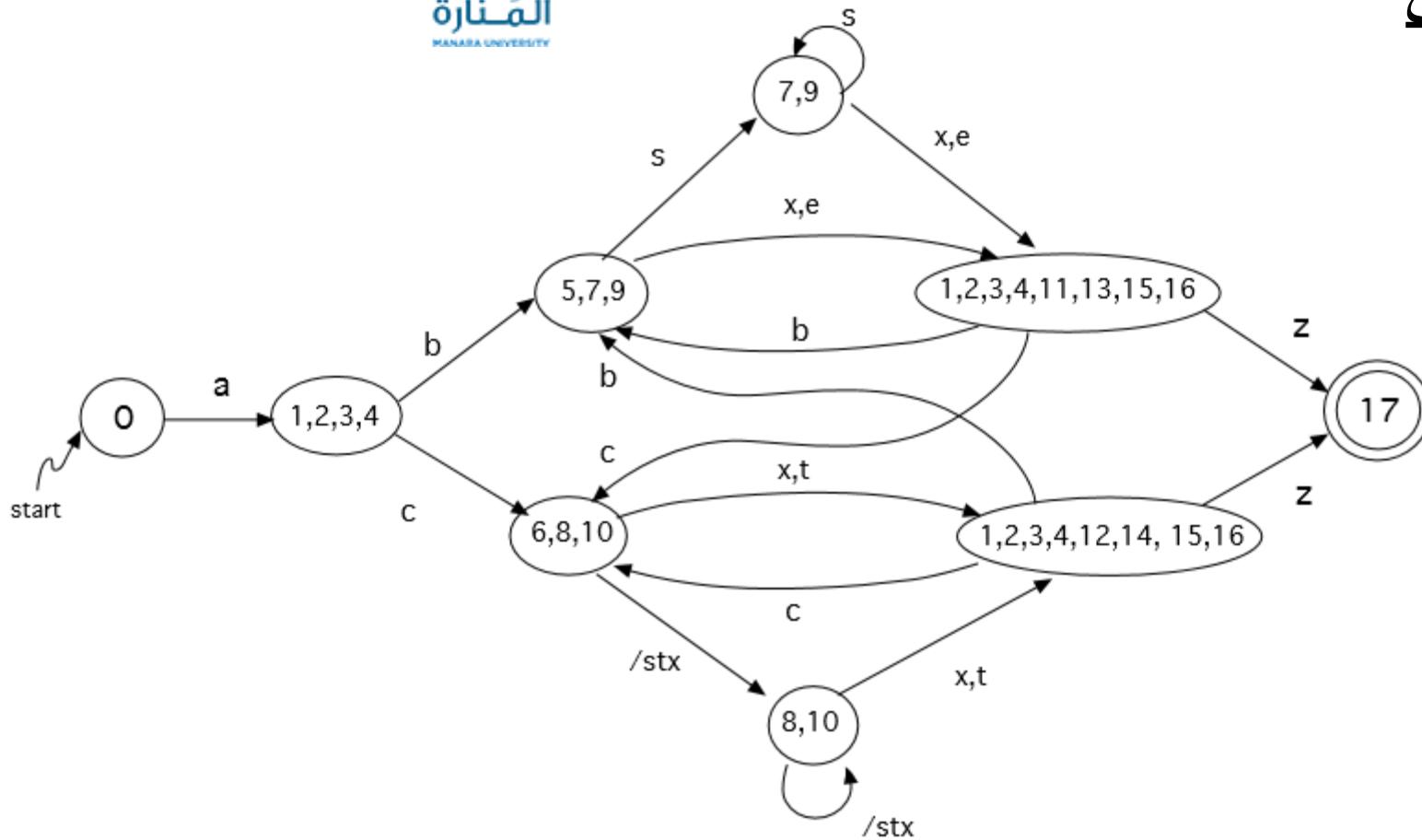
	ϵ -closure(move (s_i, a))	ϵ -closure(move (s_i, b))	ϵ -closure(move (s_i, c))	ϵ -closure(move ($s_i, x/e$))	ϵ -closure(move ($s_i, /stx$))	ϵ -closure(move ($s_i, x/t$))	ϵ -closure(move (s_i, s))	ϵ -closure(move (s_i, z))
A={0}	ϵ -closure(1)={1,2,3,4} B	-	-	-	-	-	-	-
B={1,2,3,4}	-	ϵ -closure(5)={5,7,9} C	ϵ -closure(6)={6,8,10} D	-	-	-	-	-
C={5,7,9}	-	-	-	ϵ -closure(11)={1,2,3,4,11,13,15,16} E	-	-	ϵ -closure(9)={7,9} F	-
D={6,8,10}	-	-	-	-	ϵ -closure(10)={8,10} G	ϵ -closure(12)={1,2,3,4,12,14,15,16} H	-	-
E={1,2,3,4,11,13,15,16}	-	C	D	-	-	-	-	ϵ -closure(17)={17} I

	ϵ -closure(move(s_i, a))	ϵ -closure(move(s_i, b))	ϵ -closure(move(s_i, c))	ϵ -closure(move($s_i, x/e$))	ϵ -closure(move($s_i, /stx$))	ϵ -closure(move($s_i, x/t$))	ϵ -closure(move(s_i, s))	ϵ -closure(move(s_i, z))
$F=\{7,9\}$	-	-	-	ϵ -closure(11) <u>E</u>	-	-	ϵ -closure(9)={7,9} <u>F</u>	-
$G=\{8,10\}$	-	-	-	-	- ϵ -closure(10) <u>G</u>	ϵ -closure(12) <u>H</u>	-	-
$H=\{1,2,3,4,12,14,15,16\}$	-	ϵ -closure(5) <u>C</u>	ϵ -closure(6) <u>D</u>	-	-	-	-	ϵ -closure(17)={17} <u>I</u>
$I=\{17\}$	-	-	-	-	-	-	-	-

مخطط DFA الموافق



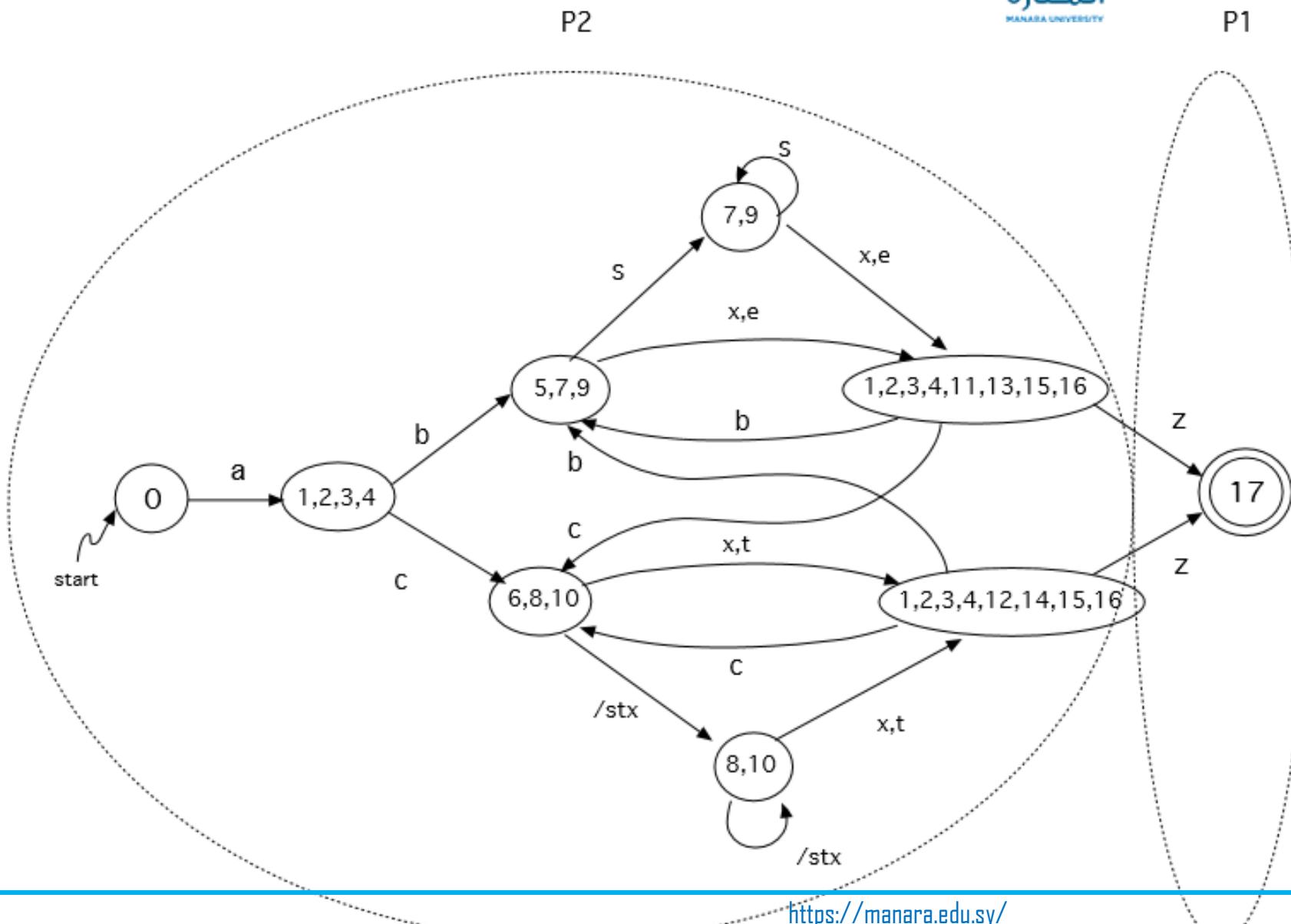
جامعة
المنارة
MANARA UNIVERSITY

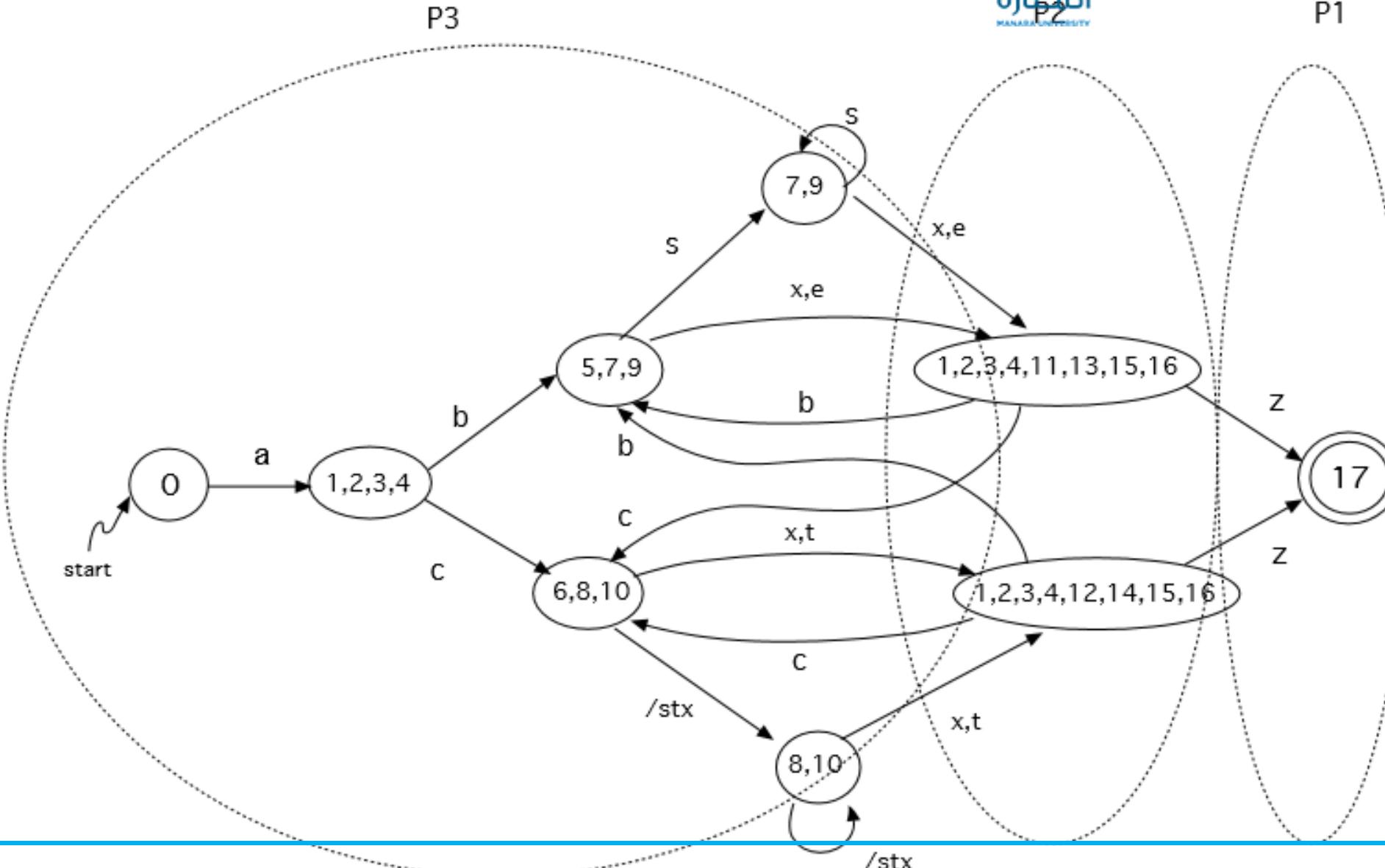


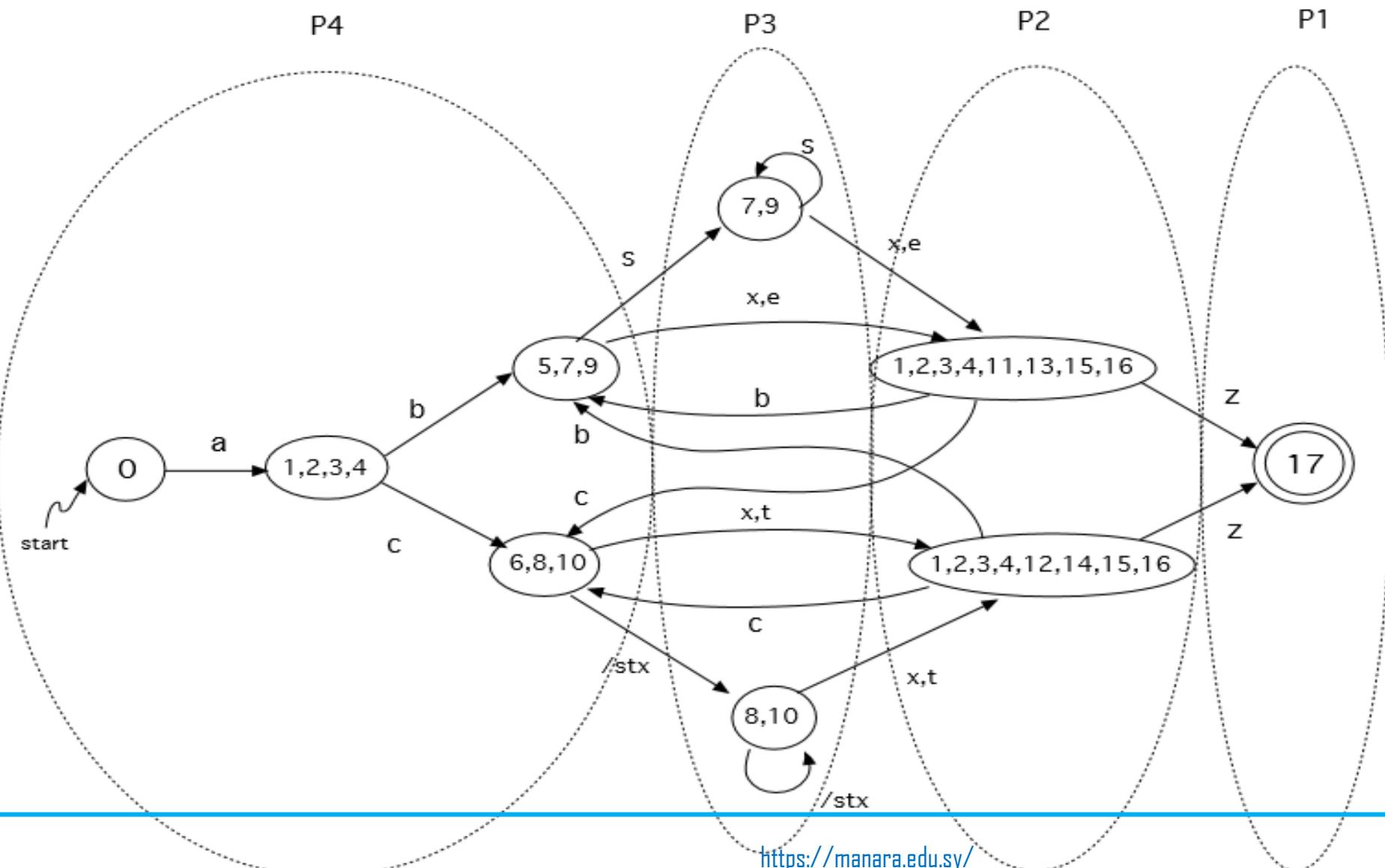


P2

P1

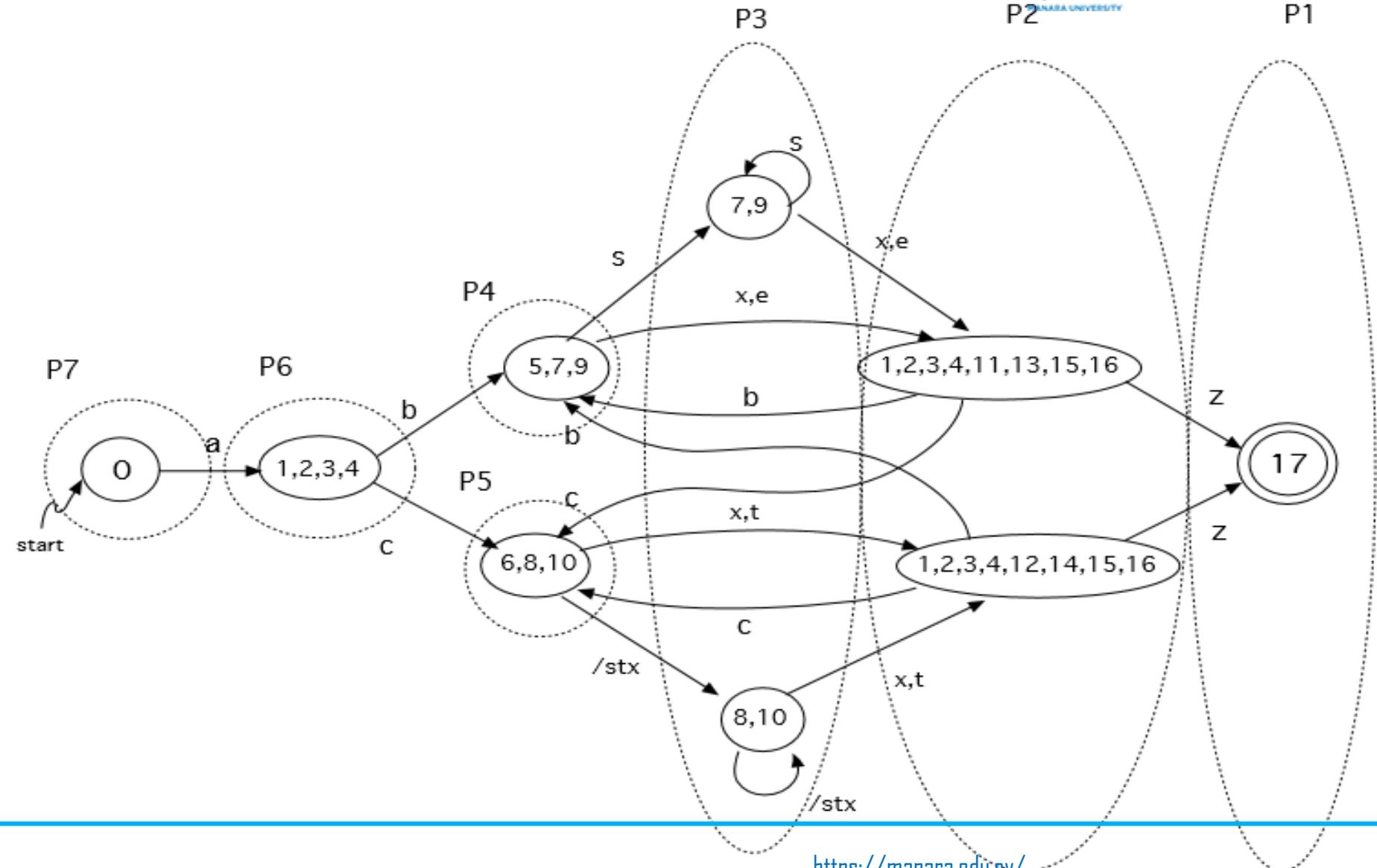
اختصار مخطط DFA

اختصار مخطط DFA

 اختصار مخطط DFA

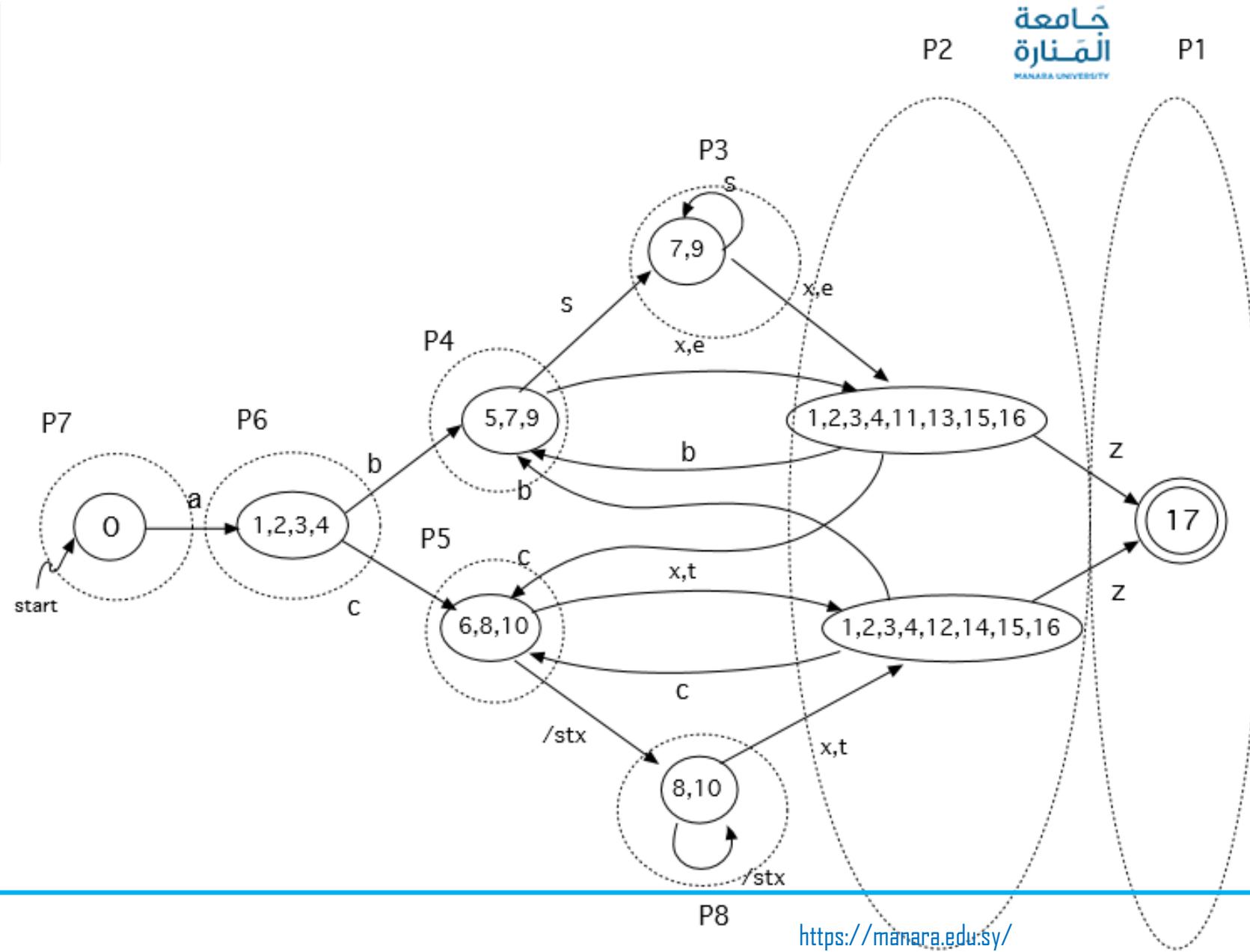


اختصار مخطط DFA





اختصار مخطط DFA



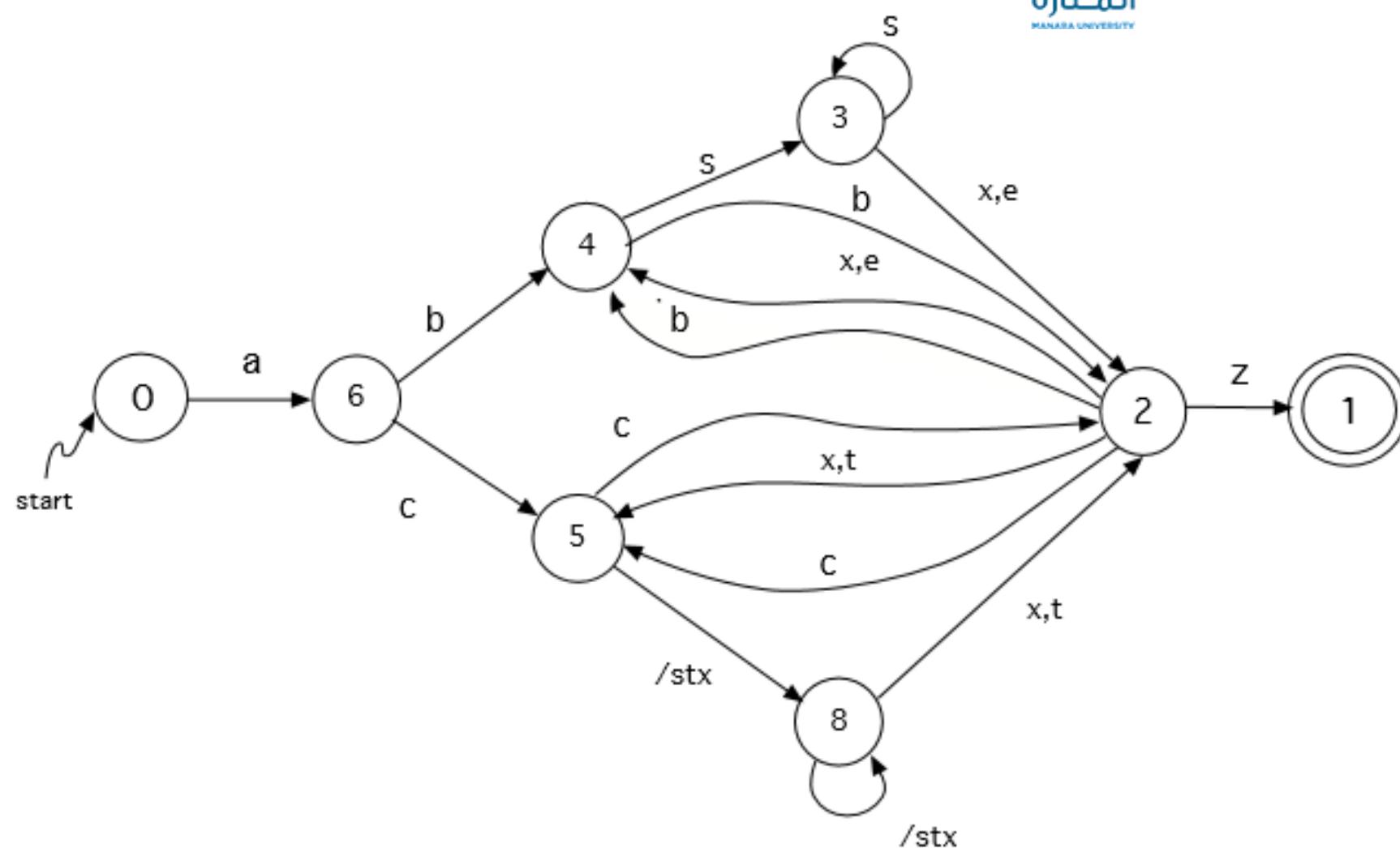
DFA مخطط المختصر النهائي

سؤال شامل



باستخدام طريقة بناء المجموعات الجزئية حول مخطط NFA السابق إلى مخطط DFA الموافق.

جامعة
المنارة
MANARA UNIVERSITY



Some Rules for RE

x	the character x
"x"	the character x, even if it is a special character
\x	the character x, even if it is a special character
x\$	the character x at the end of a line
^x	the character x at the beginning of a line
x?	Zero or one occurrence of x
x+	One or more occurrences of x
x*	Zero or more occurrences of x
xy	the character x followed by the character y
x y	the character x or the character y
[az]	the character a or the character z
[a-z]	from character a to character z
[^a-z]	Any character except from a to z
x{ n }	n occurrences of x
x{ m,n }	between m and n occurrences of x
x/y	x if followed by y (only x is part of the pattern)
.	Any character except \n
(x)	same as x, parentheses change operator priority
<<EOF>>	end of file