

LZSY005 - Linguistique computationnelle

Epreuve à mi-semestre

Travail individuel

Durée : 1h

1. [4 pts = 12']
Expliquez pourquoi on fait la différence, dans le domaine du TAL, entre **tâche** et **application**.

2. [6 pts = 18']
Soit l'alphabet $\Sigma = \{a, b, c, d\}$
- (a) Proposer un automate déterministe complet qui reconnaît le langage de tous les mots qui comprennent au moins 2 b ;
 - (b) Proposer un automate déterministe complet qui reconnaît le langage de tous les mots qui comprennent au plus 2 b ;
 - (c) Proposer un automate déterministe complet qui reconnaît le langage de tous les mots qui comprennent exactement 2 b .

3. [5 pts = 15']
- (a) Donner une expression rationnelle qui décrit le langage
 $L_3 = \{ab, bba, acba, babb, ccacc\}$.
 - (b) Soit l'expression rationnelle suivante :

$$a \left(c \left(a \left(b \mid \varepsilon \right) \mid \varepsilon \right) \mid \varepsilon \right) \mid \varepsilon$$

Donnez la liste des mots du langage ainsi décrit. Que peut-on dire de ce langage ?

- (c) Proposer une expression rationnelle pour tous les mots sur l'alphabet $\{a, b, c\}$ qui comprennent la suite abc .

4. [5 pts = 15']

Donner un automate qui reconnaît le langage décrit par les expressions rationnelles suivantes. On pourra s'inspirer des principes généraux de correspondance entre expressions rationnelles et automates, et utiliser par conséquent des ε -transitions. Cependant pour le troisième automate, on demande (aussi) un automate le plus simple possible sans ε -transition.

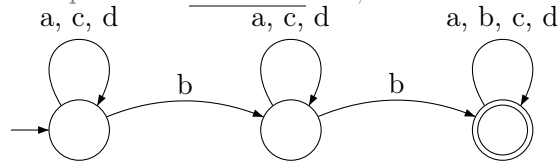
- $(a|c)(b|\varepsilon)d^*$
- $(ab|ba)^*$
- $(a|b)(a|b)^*$

1. Expliquez pourquoi on fait la différence, dans le domaine du TAL, entre **tâche** et **application**.

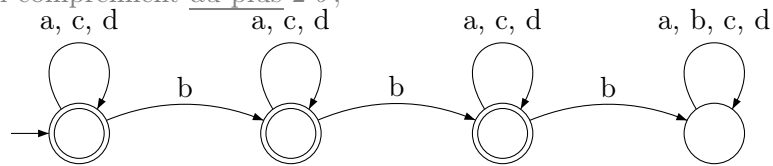
Question de cours non corrigée.

2. Soit l'alphabet $\Sigma = \{a, b, c, d\}$

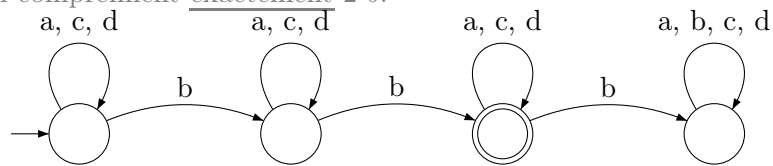
- (a) Proposer un automate déterministe complet qui reconnaît le langage de tous les mots qui comprennent au moins $2 b$;



- (b) Proposer un automate déterministe complet qui reconnaît le langage de tous les mots qui comprennent au plus $2 b$;



- (c) Proposer un automate déterministe complet qui reconnaît le langage de tous les mots qui comprennent exactement $2 b$.



3. (a) Donner une expression rationnelle qui décrit le langage

$$L_3 = \{ab, bba, acba, babb, ccacc\}.$$

Le choix le plus simple consiste à tout simplement considérer ce langage comme la disjonction entre tous les mots :

$$(ab \mid bba \mid acba \mid babb \mid ccacc)$$

On pouvait essayer aussi de « factoriser », mais le bénéfice n'est pas absolument évident :

$$(a(b|cba) \mid b(ba|abb) \mid ccacc)$$

- (b) Soit l'expression rationnelle suivante :

$$a \left(c \left(a \left(b \mid \varepsilon \right) \mid \varepsilon \right) \mid \varepsilon \right) \mid \varepsilon$$

Donnez la liste des mots du langage ainsi décrit. Que peut-on dire de ce langage ?

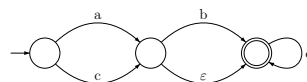
Le langage contient les mots $\{\varepsilon, a, ac, acb, acab\}$. On remarque qu'il s'agit de tous les préfixes de $acab$ — y compris les préfixes "triviaux" ε et le mot lui-même.

- (c) Proposer une expression rationnelle pour tous les mots sur l'alphabet $\{a, b, c\}$ qui comprennent la suite $babc$.

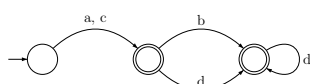
$$(a|b|c)^* babc (a|b|c)^*$$

4. Donner un automate qui reconnaît le langage décrit par les expressions rationnelles suivantes. On pourra s'inspirer des principes généraux de correspondance entre expressions rationnelles et automates, et utiliser par conséquent des ϵ -transitions. Cependant pour le troisième automate, on demande (aussi) un automate le plus simple possible sans ϵ -transition.

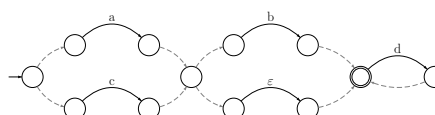
— $(a|c)(b|\epsilon)d^*$



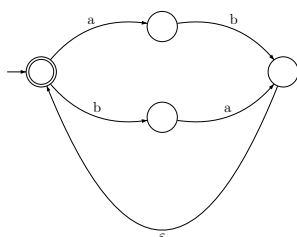
ou encore :



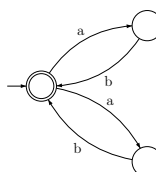
ou encore : (ϵ -transitions en pointillés)



— $(ab|ba)^*$



ou encore :



— $(a|b)(a|b)^*$

