# Expressions régulières

**Introduction informelle**

[A-F] = {A, B, C, D, E, F}

[0-99] = [0-9,9] = [0-9]

[0-24-6] = {0,1,2,4,5,6}

. = n’importe quel caractère

[.] = ça veut dire le point

. [.]. [..] = any. any any.

ε = mot vide

A\* = ε, A, AA, AAA, …

M+ = M\* - {ε} = M M\*

[1 + 2\*] = {1, +, 2, \*}

M ? = (0 ou 1 fois) ε, M

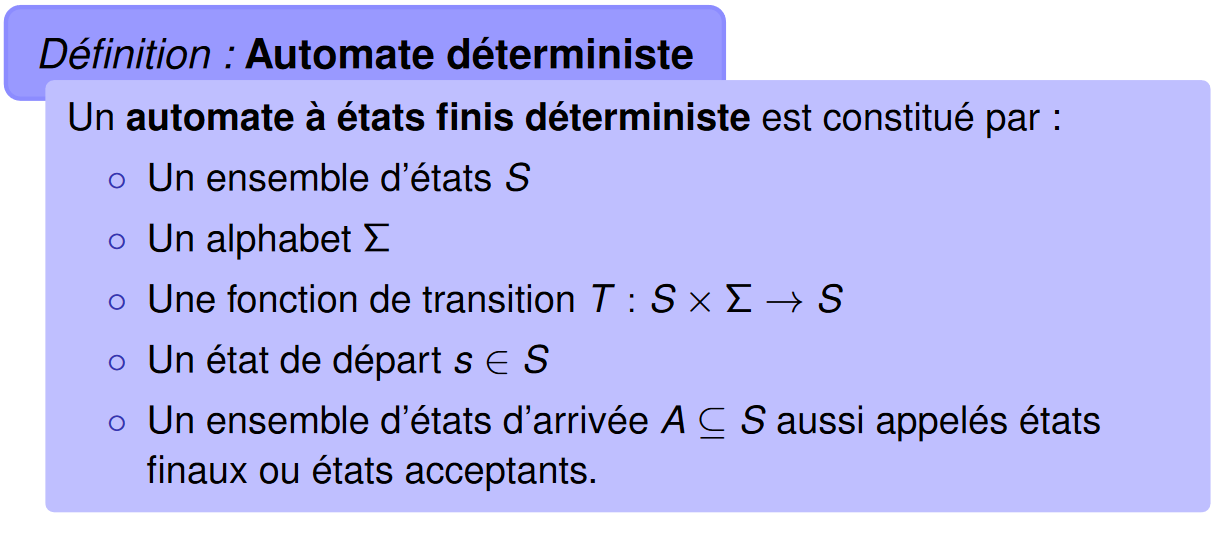
[1] [23] ? = {1, 12, 13}

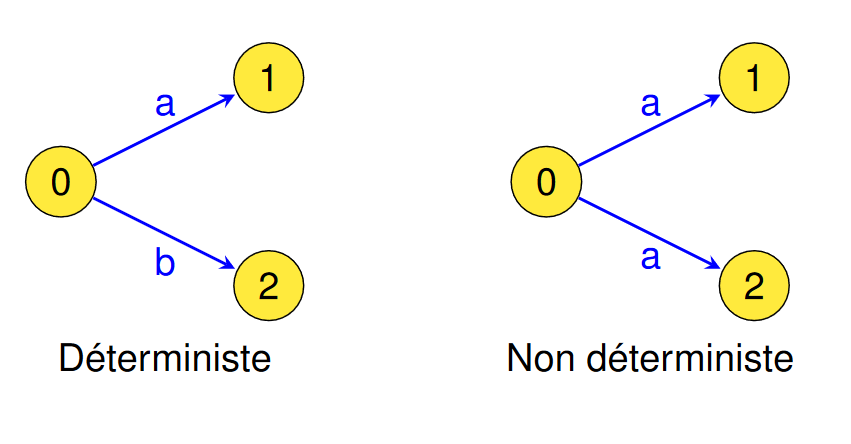
A.\* = A … (tous les mots qui commencent par A ou A tout seul)

[0-9]+|[a-z]+ = le truc de gauche ou le truc de droite

[^35] = any – {3-5} (n’importe quel caractère sauf 3 ou 5)

**Les automates**





**Expressions régulières**

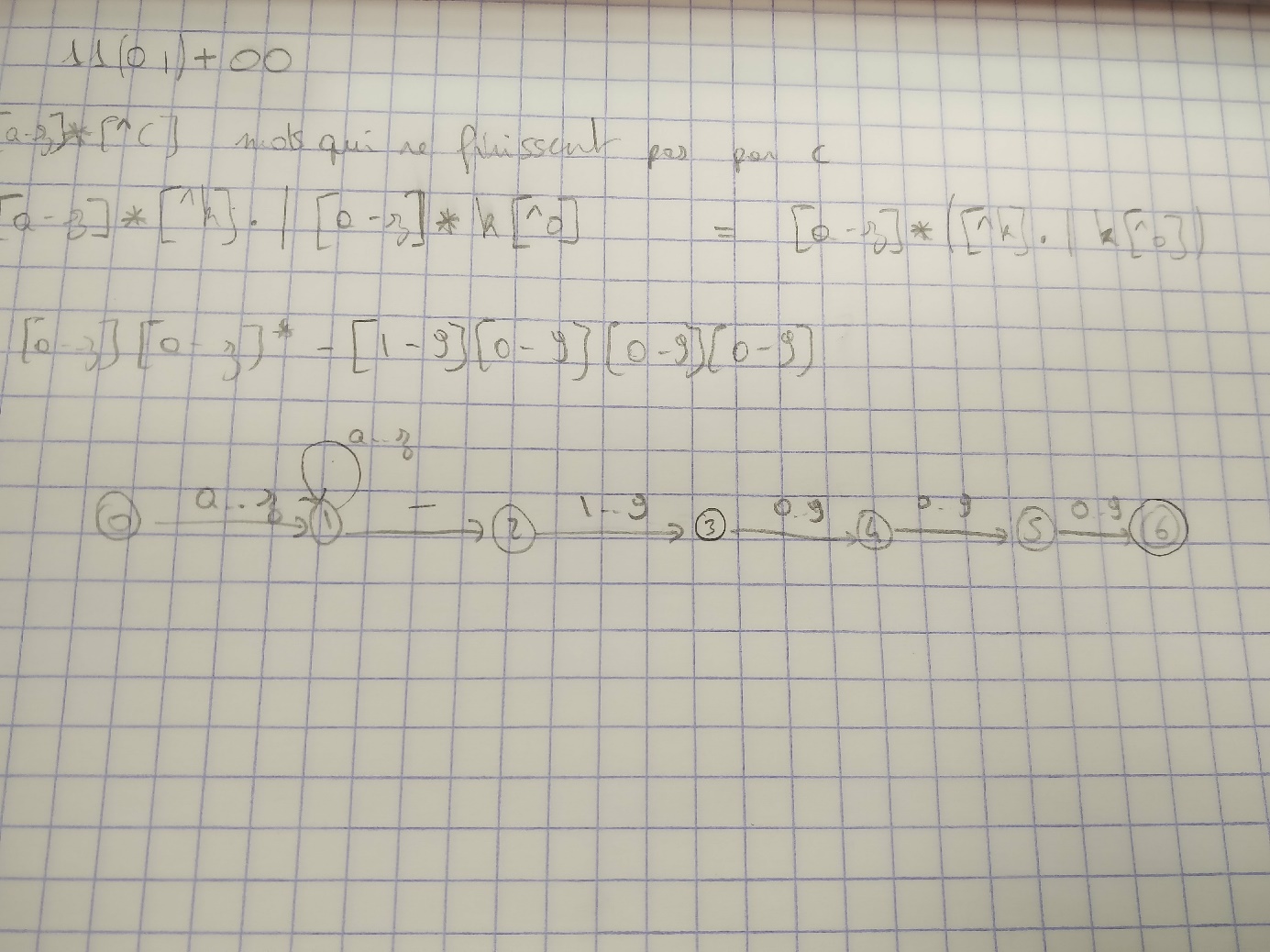
[a-z]\*[^c] = n’importe quel mot qui ne finit pas par un c

Ʃ = {a-z}

[a-z]\*[^k].| [a-z]\*k[^o] = [a-z]\*([^k].|k[^o]) = n’importe quel mot qui ne finit pas par ko

**Exercice**

[a−z][a−z]∗−[1−9][0−9][0−9][0−9]



**A retenir**

Notion d’expression régulière

Définition formelle (grammaire des regexps)

Équivalence avec les automates (deux sens)

**TD 1**

12092

A -> 192.168.2.9 port : 1215

B -> 192 .168.2.8 port : 80

<https://regex101.com/>

[21] correspond à la recherche de 2 et 1 séparément

[0-255] correspond à la recherche de 0, 2, 5

[0-9]{1,3} cherche tous les nombres compris entre 0 et 999

[0-9]+ cherche tous les nombres compris entre 0 et 999999…

Adresse IP -> [0-9]{1,3}[.][0-9]{1,3}[.][0-9]{1,3}[.][0-9]{1,3}

Port (22 ou 21) -> [0-9]{1,3}[.][0-9]{1,3}[.][0-9]{1,3}[.][0-9]{1,3} 2[12]

Cherche tous les nombres une seule fois -> grep -o -E « [0-9]+ » nom\_fichier|sort|uniq

Cherche tous les nombres une seule fois et trié -> grep -o -E « [0-9]+ » nom\_fichier|sort -n|uniq

Remplacer adresse mac par « ICI-MAC » -> sed -r « s/[0-9][a-f]{1-2}[ :]/ICI-MAC/g » extrait\_trace

Trouver url -> grep -o -E « https ?:\/\/[^’]» page\_microsoft.html |wc -l

Wc -l -> compte le nombre de ligne trouvé

Recherche commune dans le département 31 -> grep -o -E « 31[0-9]{3}\, » villes.txt |wc -l

Recherche des villes avec deux apostrophes -> grep -o -E « .\*[\’].\*[\’] » | wc -l