

# Examen de théorie des langages

La clarté et la précision de la rédaction seront prises en compte dans l'évaluation.

Le barème est indicatif.

Aucun document n'est autorisé.

Toutes les machines (ordinateurs, téléphones et calculatrices) sont interdites.

Toutes les réponses doivent être justifiées.

Durée : 1h 30min

L'acide ribonucléique (ARN) est une molécule biologique qui permet l'expression du code génétique. Il est composé de 4 bases nucléiques qui sont l'adénine, la cytosine, la guanine et l'uracile que l'on notera respectivement **a**, **c**, **g** et **u**.

Dans la suite, on considèrera qu'un brin d'ARN est un mot sur l'alphabet  $\{a, c, g, u\}$ .

## 1 Suite de bases (1+3+3)

Nous nous intéressons dans cette question aux suites de bases sur l'alphabet  $\Sigma_1 = \{a, u\}$  commençant par un **a**, finissant par un **u** et composées d'un nombre quelconque de **a** et de **u**.

1. Donner un Automate Fini Déterministe (AFD) qui reconnaît ces séquences.
2. Appliquer la mise en équation et la résolution via le lemme d'Arden afin d'en déduire le langage associé sous forme d'une expression régulière.
3. Appliquer la technique des quotients gauches sur l'expression régulière obtenue précédemment, afin de trouver l'AFD minimal permettant de reconnaître ces séquences.

## 2 C'est la fin de l'ARN! (0,5+1,5+1+2+3)

Nous nous intéressons dans cette question aux suites de bases de longueur quelconque sur l'alphabet  $\Sigma_2 = \{a, c, g, u\}$  terminées par un codons stop (**uaa**, **uag** ou **uga**). (On supposera que les codons stop seuls font partie du langage).

1. Donner le langage sous sa forme d'expression régulière.
2. Donner une grammaire de ce langage.
3. Donner le type du langage (avec explications).
4. Écrire un automate non déterministe (le plus simple possible) permettant de reconnaître ce langage.
5. Utiliser l'algorithme de détermination (avec explications) pour en déduire un automate déterministe.

## 3 Mélange équilibré? (5)

Nous nous intéressons dans cette question à des suites de bases sur l'alphabet  $\Sigma_3 = \{a, g, u\}$  de longueur quelconque qui contiennent exactement le même nombre de chacune des trois bases.

1. Justifier clairement la construction d'une machine abstraite qui reconnaît ce langage.



## Bonus

Quel code a été adopté internationalement il y a exactement 112 ans?