

Complexité et algorithmique - Miniprojet 1

Basil dalié

Novembre 2018

Chapitre 1

Introduction

Dans le cadre du module Complexité et algorithmique inscrit dans le programme pédagogique de quatrième année en Sciences Informatiques à l'école Polytech Nice-Sophia, il nous a été demandé de réaliser quatre algorithmes pour résoudre le problème de la recherche du $n^{\text{ème}}$ bit sur une machine de turing. Ce problème consiste à retourner, pour une donnée de longueur N constitué d'un mot m sur l'alphabet $\{A, B, C, D\}$ suivi de la lettre X puis d'un nombre n , la $n^{\text{ème}}$ lettre du mot m ou alors X si le nombre n est supérieur à la longueur de m (e.g. si m est AABA, le résultat doit être B si n est égal à 3 et X si n est égal à 5). Chacun de ces algorithmes devaient respecter des contraintes distinctes : pour le premier, la machine de turing dispose de deux bandes, contre une bande pour les trois suivants. Le second devait avoir une complexité de l'ordre de $O(N^2)$ et les deux suivants de $O(N \log N)$. Pour le troisième algorithme la donnée n devait être donnée en binaire, alors qu'elle devait être donnée en unaire pour les trois autres. Au cours de ces pages, vous pourrez trouver pour chacun des algorithmes réalisés, sa description, ainsi que la preuve et l'analyse de sa complexité

Chapitre 2

Notation

Dans la suite de ce document nous utiliserons les notations suivantes :

- p et $|p|$ désignent respectivement le préfixe du mot donné en entrée qui se termine par le dernier symbole du mot m avant le séparateur X et sa longueur.
- P et $|P|$ désignent respectivement le préfixe du mot donné en entrée qui se termine par le séparateur X (inclus) et sa longueur.
- s et $|s|$ désignent respectivement le suffixe du mot donné en entrée qui commence par le premier symbole du nombre n après le séparateur X et sa longueur.
- S et $|S|$ désignent respectivement le suffixe du mot donné en entrée qui commence par le séparateur X (inclus) et sa longueur.

Chapitre 3

Première solution

3.1 Description

La première solution pour la recherche du $n^{\text{ème}}$ bit est une machine de Turing à deux bandes, le nombre n est donné en unaire. la complexité de l'algorithme est de l'ordre de $O(N)$.

3.1.1 Alphabet

Pour cette solution, l'ensemble des symboles T est égal à l'union de l'ensemble des symboles de données I avec l'ensemble $\{-\}$ où le symbole $-$ représente le symbole blanc. Ses éléments sont : $\{A, B, C, D, X, 1, -\}$, le symbole 1 représente l'unique chiffre pour représenter la donnée n en unaire.

3.1.2 Ensemble d'état

La machine de turing comporte 11 états en comptant l'état final q_f .

3.1.3 Fonction de transition

Voir table de transition en annexe A.

3.1.4 Description de l'algorithme

L'algorithme se déroule de la manière suivante : au démarrage, la tête de lecture/écriture de la première bande se trouve sur le premier caractère du mot m et celle de la seconde bande se trouve au même niveau. A l'état initial (q_0), tant qu'un symbole A, B, C, ou D est lû sous la tête de lecture/écriture de la première bande, celle-ci se déplace vers la droite et la machine reste dans la même état.

Lorsque la tête de lecture/écriture de la première bande atteint le symbole X, cette dernière est déplacée vers la droite et on passe dans l'état q_1 (Copie du nombre n sur la deuxième bande).

A partir de ce moment, tant qu'un symbole 1 est lu sous la tête de lecture/écriture de la première bande, ce symbole est effacé de la première bande et recopié sur la seconde bande et les deux têtes de lecture/écritures se déplacent vers la droite. Une fois que tous les symboles 1 ont été consommés, la machine passe dans l'état q_2 (Déplacement de la tête de lecture/écriture de la première bande sur le caractère X) dans lequel la tête de lecture/écriture de la première bande se déplace vers la gauche tant qu'un symbole blanc est lu.

Lorsque le symbole X est lu par la tête de lecture/écriture de la première bande, celle-ci se déplace vers la gauche et la machine passe dans l'état q_3 (Déplacement de la tête de lecture/écriture de la première bande sur le premier symbole du mot m).

Dans cet état, tant qu'un symbole A, B, C, ou D est lu par la tête de lecture/écriture de la première bande, celle-ci se déplace vers la gauche, et lorsque le symbole blanc ($-$) est lu, elle se déplace vers la droite et la machine passe dans l'état q_4 (Déplacement de la tête de lecture/écriture de la seconde bande en dessous du premier symbole du mot m).

A partir de ce moment, tant qu'un symbole 1 est lu sous la tête de lecture/écriture de la seconde bande, celle-ci se déplace vers la gauche, et lorsque le symbole blanc est lu, elle se déplace vers la droite et la machine passe dans l'état q_5 (effaçage des symboles situés après le symbole d'indice n).

A ce moment de l'algorithme, la première bande contient le mot m suivi du symbole X , la seconde bande contient le nombre n écrit en unaire dont le premier symbole est situé juste en dessous du premier symbole du mot m . Les têtes de lecture/écriture des bandes 1 et 2 sont situés respectivement sur le premier symbole du mot m et sur le premier symbole du nombre n .

A l'état q_5 , tant que les têtes de lecture/écriture des bandes 1 et 2 lisent respectivement un symbole parmi $\{A, B, C, D\}$ et un symbole 1, les deux têtes de lecture/écriture se déplacent vers la droite. Si la tête de lecture/écriture de la première bande lit un symbole parmi $\{A, B, C, D, X\}$ et la tête de lecture/écriture de la seconde bande lit un symbole blanc (\cdot) on écrit le symbole blanc sous la tête de lecture/écriture de la première bande et les deux têtes de lecture/écriture se déplacent vers la droite. Dans le cas où le symbole X est lu par la tête de lecture/écriture de la première bande, si la tête de lecture/écriture de la seconde bande lit le symbole blanc ($n \leq \text{long}(m)$), les deux têtes de lecture/écriture se déplacent vers la gauche et la machine passe dans l'état q_6 (Positionnement sur le symbole d'indice m).

Sinon, toujours si un symbole X est lu sur la première bande, si un symbole 1 est lu sur la seconde bande ($n > \text{long}(m)$), les deux têtes de lecture/écriture se déplacent vers la gauche et la machine passe dans l'état q_7 (Effaçage des symboles situés à gauche).

Dans le cas où $n \leq \text{long}(m)$, la machine se trouve dans l'état q_6 , dans lequel les deux têtes de lecture/écriture se déplacent vers la gauche tant que le symbole blanc est lu sur les deux bandes. Lorsqu'un symbole parmi $\{A, B, C, D\}$ est lu sur la première bande et qu'un symbole 1 est lu sur la seconde bande (le symbole d'indice n est atteint), les deux têtes de lecture/écriture se déplacent vers la gauche et la machine passe dans l'état q_7 (Effaçage des symboles situés à gauche du mot d'indice n).

A ce moment, peu importe si $n \leq \text{long}(m)$ ou si $n > \text{long}(m)$, on se trouve dans l'état q_7 et les deux têtes de lecture/écriture sont situées au même niveau et de sorte que le symbole à renvoyer est situé à droite de la tête de lecture/écriture de la première bande. Dans cet état, tant que les têtes de lecture/écriture des bandes 1 et 2 lisent respectivement un symbole parmi $\{A, B, C, D\}$ et un symbole 1, un symbole blanc est écrit sur la première bande et les deux têtes de lecture/écriture se déplacent à gauche. Si deux symboles blancs sont lus sur les deux bandes, alors les deux têtes de lecture/écriture se déplacent vers la droite et la machine passe dans l'état q_8 (Effaçage de la seconde bande).

Dans l'état q_8 , tant qu'on lit un symbole 1 sur la seconde bande, un symbole blanc est écrit sous la tête de lecture/écriture de la seconde bande et celle-ci se déplace à droite. Quand un symbole blanc est lu sous la tête de lecture/écriture de la seconde bande, celle-ci se déplace vers la gauche et la machine passe dans l'état q_9 (Déplacement de la tête de lecture/écriture de la première bande sur le symbole à renvoyer).

A l'état q_9 , la première bande ne contient que le symbole à renvoyer des symboles blancs, et la seconde bande ne contient que des symboles blancs. Si $\text{long}(m) = 1$, la tête de lecture/écriture de la première bande est située sur le symbole à renvoyer, sinon elle est située à sa gauche. Ainsi, celle-ci se déplace à droite tant qu'un symbole blanc est lu sur la première bande. Quand un symbole parmi $\{A, B, C, D, X\}$ est lu sur la première bande (le $n^{\text{ème}}$ bit est atteint), la machine passe dans l'état final q_f .

3.1.5 Preuve de l'algorithme

Dans le cas où $n \leq \text{long}(m)$, il est évident qu'après la recopie des symboles constituant le nombre n sur la seconde bande à partir du niveau du premier symbole du mot m , le dernier symbole 1 de la seconde bande est au même niveau que le $n^{\text{ème}}$ symbole du mot m (c'est à dire le symbole à renvoyer). Il suffit alors de synchroniser les deux têtes de lecture/écriture et de les placer au niveau du premier symbole du mot m , puis d'avancer les deux têtes de lecture/écriture vers la droite jusqu'à ce qu'un symbole blanc soit lu sur la seconde bande pour détecter l'emplacement du symbole à renvoyer, ensuite il reste seulement à effacer les symboles autour de celui-ci puis d'effacer la seconde bande. Le cas où $n > \text{long}(m)$ est détecté lors du déplacement des deux têtes de lecture/écriture vers la droite de manière synchronisée, si on lit le symbole X sur la première bande et le symbole 1 sur la seconde. A ce moment, les deux têtes de lecture/écriture se trouvent alors sur le symbole à renvoyer et comme tous les symboles 1 de la seconde bande ont été consommés, il reste seulement à effacer tous les symboles à gauche de X sur la première bande puis d'effacer la seconde bande.

3.1.6 Analyse de la complexité

Comme décrit dans la section 2.1.4, les séquences de transition entre deux états différents de la machine menant au résultat attendu sont $q_0 \rightarrow q_1 \rightarrow q_2 \rightarrow q_3 \rightarrow q_4 \rightarrow q_5(\rightarrow q_6)? \rightarrow q_7 \rightarrow q_8 \rightarrow q_9 \rightarrow q_f$. Pour obtenir la fonction $T(N)$ mesurant le nombre de transition à emprunter pour aboutir au résultat

en fonction de la longueur de la donnée N , il suffit de calculer le nombre de transition effectuées dans chacun de ces états en fonction de N puis d'en faire la somme.

- Pour q_0 le nombre de transition est égal à $N - |s|$ avec $|s| \leq N - 2$ donc strictement positif.
- Pour q_1 le nombre de transition est égal à $N - |P| + 1$ avec $|P| \leq N - 1$ donc strictement positif.
- Pour q_2 le nombre de transition est égal à $N - |p|$ avec $|P| \leq N - 2$ donc strictement positif.
- Pour q_3 le nombre de transition est égal à $N - |S|$ avec $|S| \leq N - 1$ donc strictement positif.
- Pour q_4 le nombre de transition est égal à $N - |P| + 1$ donc strictement positif
- Pour q_5 le nombre de transition est égal à $N - |s|$ qui est strictement positif
- Pour q_6 le nombre de transition est égal à $|p| - |s| + 1$ avec $|p| > |s|$ car $n \leq \text{long}(m)$ (autrement on n'entrerait pas dans cet état), ce qui revient à $N - 2 * |s|$ qui est strictement positif
- Pour q_7 , le nombre de transition est égal à $N - |P| + 1$ si $n \leq \text{long}(m)$ ou $N - |S| + 1$ si $n > \text{long}(m)$ donc dans tous les cas $N - k + 1$ avec $k < N - 1$ qui est strictement positif
- Pour q_8 , le nombre de transition est égal à $N - |P| + 1$ qui est strictement positif
- Pour q_9 , le nombre de transition est égal à $N - |P|$ si $n \leq \text{long}(m)$ ou $N - |S|$ si $n > \text{long}(m)$ donc dans tous les cas $N - l + 1$ avec $l < N - 1$ qui est strictement positif
- Pour q_f , le nombre de transition est égal à 0

On en déduit la fonction $T(N) = 9 * N - |p| - 3 * |P| - 2 * |s| - |S| + 5 - k - l + (N - 2 * |s|)$:

La définition de Big-O dit qu'une fonction $f(x)$ est égale à $O(g(x))$ si il existe un facteur constant c et un nombre réel x_0 tel que $|f(x)| \leq c * g(x)$.

Pour $x_0 = 0$ et $c = 1$, $|T(N)| \leq 9 * N + (N - 2 * |s|) + 5$ car $|p| < N$, $|P| < N$, $|s| < N$ et $|S| < N$
Donc $T(N) = O(9 * N + (N - 2 * |s|))$

Pour $x_0 = 0$ et $c = 1$, $|T(N)| \leq 10 * N + 5$ car $N - 2 * |s|$ est strictement positif
Donc $T(N) = O(10 * N + 5)$

Pour $x_0 = 0$ et $c = 5$, $|T(N)| \leq 10 * N$
Donc $T(N) = O(10 * N)$

Pour $x_0 = 0$ et $c = 10$, $|T(N)| \leq N$
Donc $T(N) = O(N)$

On en déduit que la complexité de l'algorithme est dans l'ordre de $O(N)$.

Annexe A

Table de transition - machine 1

| état | l1 | l2 | e1 | e2 | d1 | d2 | nouvel état | remarques |
|-------|----|----|----|----|----|----|-------------|--------------------------------|
| q_0 | A | - | A | - | R | S | q_0 | MOVE_TO_X_ON_TAPE_ONE |
| | B | - | B | - | R | S | q_0 | |
| | C | - | C | - | R | S | q_0 | |
| | D | - | D | - | R | S | q_0 | |
| | X | - | X | - | R | S | q_1 | |
| q_1 | 1 | - | - | 1 | R | R | q_1 | WRITE_N_ON_TAPE_TWO |
| | - | - | - | - | L | L | q_2 | |
| q_2 | - | 1 | - | 1 | L | S | q_2 | MOVE_TO_WORD_RIGHT_ON_TAPE_ONE |
| | X | 1 | X | 1 | L | S | q_3 | |
| q_3 | A | 1 | A | 1 | L | S | q_3 | MOVE_TO_WORD_LEFT_ON_TAPE_ONE |
| | B | 1 | B | 1 | L | S | q_3 | |
| | C | 1 | C | 1 | L | S | q_3 | |
| | D | 1 | D | 1 | L | S | q_3 | |
| | - | 1 | - | 1 | R | S | q_4 | |
| q_4 | A | 1 | A | 1 | S | L | q_4 | MOVE_TO_WORD_LEFT_ON_TAPE_TWO |
| | B | 1 | B | 1 | S | L | q_4 | |
| | C | 1 | C | 1 | S | L | q_4 | |
| | D | 1 | D | 1 | S | L | q_4 | |
| | A | - | A | - | S | R | q_5 | |
| | B | - | B | - | S | R | q_5 | |
| | C | - | C | - | S | R | q_5 | |
| | D | - | D | - | S | R | q_5 | |
| q_5 | A | 1 | A | 1 | R | R | q_5 | ERASE_RIGHT |
| | B | 1 | B | 1 | R | R | q_5 | |
| | C | 1 | C | 1 | R | R | q_5 | |
| | D | 1 | D | 1 | R | R | q_5 | |
| | A | - | - | - | R | R | q_5 | |
| | B | - | - | - | R | R | q_5 | |
| | C | - | - | - | R | R | q_5 | |
| | D | - | - | - | R | R | q_5 | |
| | X | - | - | - | L | L | q_6 | |
| | X | 1 | X | 1 | L | L | q_7 | |
| q_6 | - | - | - | - | L | L | q_6 | MOVE_TO_NTH |
| | A | 1 | A | 1 | L | L | q_7 | |
| | B | 1 | B | 1 | L | L | q_7 | |
| | C | 1 | C | 1 | L | L | q_7 | |
| q_7 | D | 1 | D | 1 | L | L | q_7 | ERASE_LEFT |
| | A | 1 | - | 1 | L | L | q_7 | |
| | B | 1 | - | 1 | L | L | q_7 | |
| | C | 1 | - | 1 | L | L | q_7 | |
| | D | 1 | - | 1 | L | L | q_7 | |
| q_8 | - | - | - | - | R | R | q_8 | ERASE_TAPE_TWO |
| | A | 1 | A | - | S | R | q_8 | |
| | B | 1 | B | - | S | R | q_8 | |
| | C | 1 | C | - | S | R | q_8 | |
| | D | 1 | D | - | S | R | q_8 | |
| | - | 1 | - | - | S | R | q_8 | |
| | A | - | A | - | S | L | q_9 | |
| | B | - | B | - | S | L | q_9 | |
| | C | - | C | - | S | L | q_9 | |
| | D | - | D | - | S | L | q_9 | |
| | - | - | - | - | S | L | q_9 | |
| q_9 | - | - | - | - | R | S | q_9 | MOVE_TO_NTH_FINAL |
| | A | - | A | - | S | S | q_f | |
| | B | - | B | - | S | S | q_f | |
| | C | - | C | - | S | S | q_f | |
| | D | - | D | - | S | S | q_f | |
| | X | - | X | - | S | S | q_f | |

TABLE A.1 – Table de transition