TELECOM Nancy 1A — Mathématiques Appliquées pour l'Informatique

Théorie des langages : déterminisation des automates finis

Introduction

Soit un automate $\mathcal{A} = (A,\ Q,\ I,\ E,\ T)$ a priori non déterministe, l'objectif est de construire un automate déterministe $\mathcal{A}' = (A,\ Q',\ e_0,\ \delta',\ T')$ tel que $\mathcal{L}(\mathcal{A}') = \mathcal{L}(\mathcal{A})$, c'est-à-dire reconnaissant le même langage.

Remarque préliminaire

L'automate déterministe doit tenir compte de toutes les transitions possibles de l'automate indéterministe, plus précisément il doit mémoriser tous les états dans lesquels pourraient se trouver l'automate indéterministe. Une solution possible est qu'à chaque état de l'automate déterministe corresponde un ensemble d'états de l'automate indéterministe.

Les figures 1 et 2 montrent cette correspondance pour éliminer l'indéterminisme dû aux transitions (q_0, a, q_0) et (q_0, a, q_1) puis aux transitions (q_0, a, q_1) et (q_0, a, q_2).

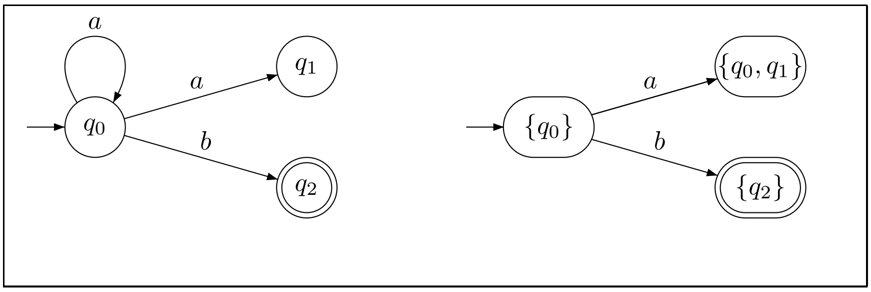


Figure 1 – Elimination de l'indéterminisme

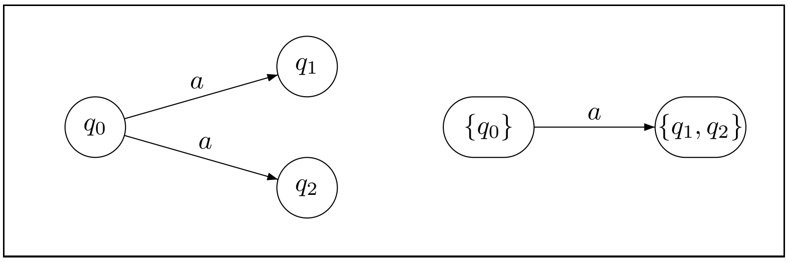


Figure 2 – Elimination de l'indéterminisme

Formalisation de la construction

Définition

Soient $\mathcal{A} = (A,\ Q,\ I,\ E,\ T)$ un automate et $p \in Q$ un état, $V(p) = \{q,\ q \in Q \text{ et } p \xrightarrow{\varepsilon} q\}$ est l'ensemble des états qui sont les extrémités de calculs dont l'origine est p et l'étiquette ε, c'est-à-dire les états pouvant être atteints à partir de p par des suites de ε-transitions.

Nous définissons maintenant les différents éléments de l'automate $\mathcal{A}' = (A,\ Q',\ e_0,\ \delta',\ T')$.

1. Q', l'ensemble des états de $\mathcal{A}'$ est $Q' = \mathcal{P}(Q)$ l'ensemble des sous-ensembles de Q.
2. e_0 l'état initial de $\mathcal{A}'$ est $e_0 = \bigcup_{r \in I} V(r)$, e_0 est l'ensemble composé des états de $\mathcal{A}$ que l'on peut atteindre par des ε-transitions à partir des états initiaux de $\mathcal{A}$. On a évidemment $I \subseteq e_0$.

3. Pour $e \in Q'$ et $a \in A$, on définit $\delta'(e,a) = \bigcup_{p \in e} \{q,\ q \in Q\ et\ p \xrightarrow{a} q\}$ où $p \xrightarrow{a} q$ signifie que l'on a un calcul d'origine p, d'extrémité q et d'étiquette a. En posant $\phi(p,a) = \{q,\ q \in Q\ et\ p \xrightarrow{a} q\}$, on obtient

$$\delta'(e,a) = \bigcup_{p \in e} \phi(p,a) \tag{1}$$

.

4. Un état de $\mathcal{A}'$ est un état final s'il contient un état final de $\mathcal{A}$, donc $T' = \{e,\ e \in Q'\ et\ e \cap T \neq \emptyset\}$.

Remarques

Le calcul de la fonction de transition δ' pose deux problèmes :

Il existe parmi les éléments de Q' des états inaccessibles, il est donc inutile de définir δ' pour ces états. On peut pallier cet inconvénient en générant seulement les états accessibles de e_0.

La définition de δ' (formule 1) fait intervenir une réunion d'ensembles et aussi des calculs pouvant comporter des ε-transitions. Les ε-transitions sont une source d'erreur lorsque l'on veut générer les états, pour cela il est préférable de définir une table intermédiaire, calculant ϕ, qui prend en compte séparément les ε-transitions.

L'algorithme suivant [1] met en œuvre ces deux idées.

Algorithme

1. Calcul de l'état initial $e_0 = \bigcup_{r \in I} V(r)$.
2. Construction d'une table intermédiaire $\begin{array}{cccc} \phi & : Q \times A & \to & \mathcal{P}(Q) = Q' \\ & (p,\ a) & \mapsto & \phi(p,\ a) \end{array}$

   Cette étape consiste à compléter la table suivante :

ϕ	a_0	$\dots$	a_i	$\dots$	a_n
q_0	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
q_i	$\dots$	$\dots$	$\phi(q_i,\ a_i)$	$\dots$	$\dots$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
q_n	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$

   où $A = \{a_0, \dots, a_n\}$ et $Q = \{q_0, \dots, q_n\}$.
3. Construction d'une table définissant δ'

δ'	a_0	$\dots$	a_i	$\dots$	a_n
e_0	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$

   Le remplissage de la table se fait ligne par ligne en utilisant la table précédente et la formule 1. Au départ seul e_0 est présent, on remplit alors la ligne de e_0 et on reporte dans la première colonne les nouveaux états générés, on remplit ensuite les lignes de ces états. Le processus s'arrête lorsque l'on ne génére plus de nouveau état et que toutes les lignes sont remplies.
4. Détermination de l'ensemble des états finaux $T' = \{e,\ e \in Q'\ et\ e \cap T \neq \emptyset\}$.

Cette construction permet d'énoncer la proposition suivante :

Proposition

Tout automate $\mathcal{A}$ est équivalent à un automate déterministe complet $\mathcal{A}'$. Si $\mathcal{A}$ a n états, on peut construire $\mathcal{A}'$ avec au plus 2^n états.

Un automate déterministe $(A,\ Q,\ q_0,\ \delta,\ T)$ est dit complet si sa fonction de transition δ est définie pour tout élément de $Q \times A$.

1. Pierre Marchand, Mathématiques Discrètes, Automates, langages, logique et décidabilité. Dunod

Exercice 1

Soit l'automate $\mathcal{A}$ de la figure 3.

1. Donner toutes les raisons pour lesquelles $\mathcal{A}$ est un automate indéterministe.
2. Déterminiser $\mathcal{A}$ en utilisant l'algorithme précédemment défini.
3. Dessiner le diagramme sagittal de l'automate déterministe obtenu.

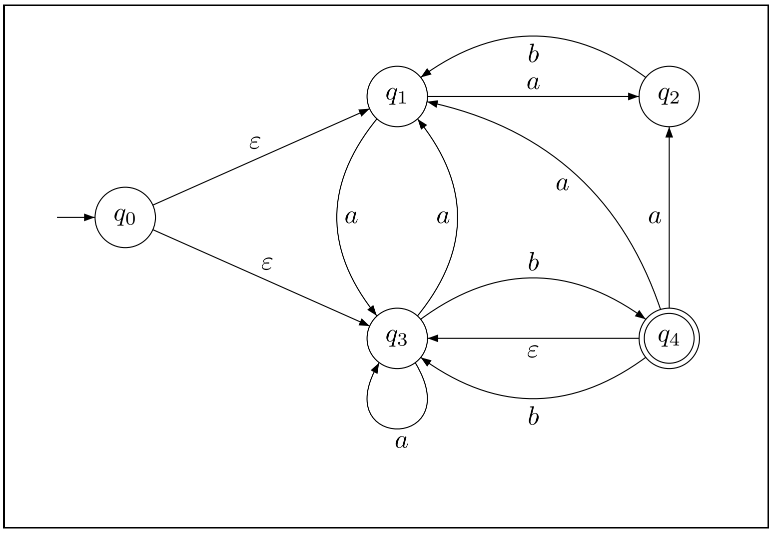


FIGURE 3 – Automate à déterminiser

Exercice 2

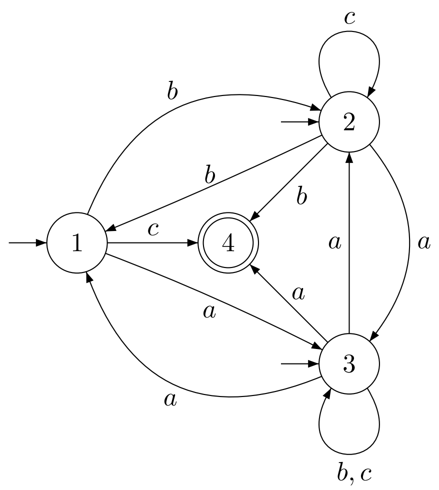
Le problème de la recherche d'un mot dans un texte apparaît sous des formes diverses dans des domaines variés (moteur de recherche sur le "Web", remplacement et recherche de chaînes de caractères dans les éditeurs de texte, recherche de mots dans les correcteurs orthographique ...)

Ces problèmes de localisation d'occurrences de mots appartenant à un ensemble fini X dans un texte écrit sur un alphabet A, se ramènent à reconnaître le langage rationnel A^*X et donc à construire l'automate correspondant.

1. Soit l'alphabet $A = \{a,\ c,\ o\}$, construire un automate reconnaissant le langage A^*cacao et donner son diagramme sagittal.
2. Déterminiser l'automate précédemment construit.
3. Dessiner le diagramme sagittal de l'automate déterministe et exécuter le mot *ccaccacocacao* sur cet automate.

Exercice 3

Soit l'automate indéterminsite suivant obtenu dans "le probléme du barmann aveugle avec des gants de boxe" (td précédent).



1. Déterminiser cet automate.
2. Trouver une stratégie gagnante pour le barmann.