Théorie des Langages - EISTI - ING 1

Yannick Le Nir

Ecole Internationale des Sciences du Traitement de l'Information

yannick.lenir@eisti.fr

Automates

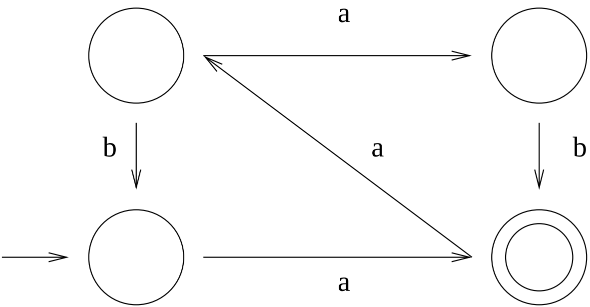
Concept

- Ensemble d'états (description du système à un instant donné)
- Ensemble d'évènements (communication vers l'extérieur)
- Ensemble de transition (triplets permettant l'évolution du système : etat-i, evt, etat-f)

Definition

Un automate d'états finis est un triplet $< Et, Ev, \Psi >$ où Et est un ensemble fini d'états, Ev un ensemble fini d'évènements et $\Psi :< Et, Ev > \longrightarrow Et$, une relation définissant les transitions.

Exemple d'automate



Chaînes reconnues : aaba , aaab, ...

Transitions

Modélisation

Les transitions entre états peuvent modéliser des problèmes autres que liés à l'axe du temps

Non-évènement

Une transition associée à un non-évènement, noté ε est appelée ε-transition (exemple : attente d'un message qui n'arrive pas).

Déterminisme

Un AFD est déterministe si $\forall (et_i, ev) \in < Et, Ev >$, il existe au plus un élément $et_f \in Et$ tel que $\Psi(et_i, ev) = et_f$. Un automate qui n'est pas déterministe est dit indéterministe.

Automate à reconnaissance de motifs

Ensemble de motifs

Soit G un ensemble et C une condition portant sur les éléments de cet ensemble. On définit $M_C(G) = \{g \in G / C(g) \text{ est vraie } \}$. Cet ensemble est appelé ensemble de motifs de G associés à C.

Définition

Un AFD à reconnaissance de motifs est un quintuplet $< Et, F, Ev, \Psi, q_0 >$ tel que :

- $< Et, Ev, \Psi >$ est un automate d'états finis
- $F \in Et$ est un ensemble non vide dit d'états finaux
- $q_O \in Et$ est l'état initial

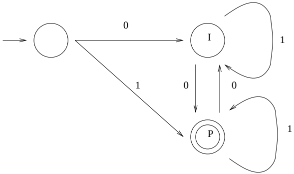
Reconnaissance de motifs

Exemple

Soit G l'ensemble des suites finies de nombres binaires (0 ou 1) et C la condition définie par $C(g)$: g contient un nombre pair de 0. L'automate à reconnaissance de motifs suivant reconnaît les nombres respectant la condition C :

- $Et = \{Debut, Pair, Impair\}$
- $F = \{Pair\}$
- $Ev = \{0, 1\}$
- Ψ est définie par les transitions

Evènement Etat	0	1
Debut	Impair	Pair
Impair	Pair	Impair
Pair	Impair	Pair



Déterminisation

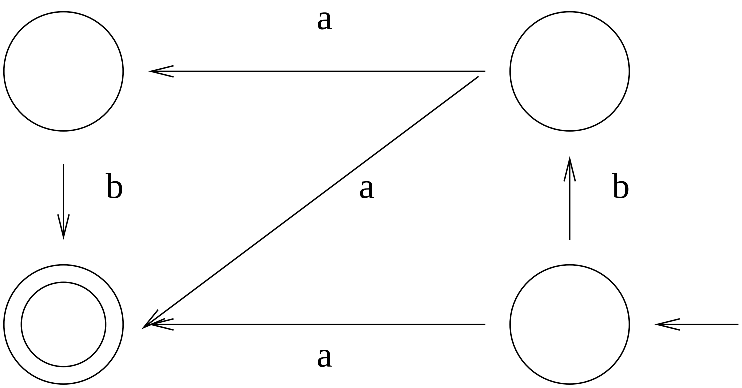
Théorème

Pour tout automate indéterministe à reconnaissance de motifs AIM $< Et, F, Ev, \Psi, q_0 >$, il existe un automate déterministe ADM $< Et', F', Ev', \Psi', q'_0 >$ à reconnaissance de motifs équivalent.

Preuve

- Construction de ADM par récurrence
  - Evénements identiques : $Ev' = Ev$
  - Et' sous ensemble de Et : $Et' \subseteq P(Et)$
- Base : $G_0 = \{\{q_0\}\}$, $E' = \emptyset$ et $q'_0 = \{q_0\}$
- Récurrence : init de G_n à $\emptyset$. $\forall et \in G_{n-1}$ et $\forall ev \in Ev$, $et' = \{et'_i \in E / \exists et_i \in E$ tel que $et'_i = \Psi(ev, et_i)\}$
  - Si $et' \neq \emptyset$ alors $E' = E' \cup \{et'\}$, $G_n = G_n \cup \{et'\}$
  - Si $et' \cap F \neq \emptyset$ alors $F' = F' \cup \{et'\}$
  - Ajout de la transition $et = \Psi'(et', ev)$

Exemple d'automate indéterministe



Motif reconnaissable

Definition

Soit $M_C(G)$ un ensemble de motifs et $A =< Et, F, Ev, \Psi, q_0 >$ un automate à reconnaissance de motifs. L'automate A permet de reconnaître les motifs de $M_C(G)$ ssi

1. $G \subseteq Ev^*$
2. $(u_1 \cdots u_n) \in M_C(G) \leftrightarrow$ Il existe au moins une configuration où l'automate après avoir été initialisé à l'état q_0 puis soumis aux événements $u_1, \cdots, u_n$ se trouve dans un état e tel que $e \in F$.

Langages réguliers et AFD

Théorème

Les mots d'un langage de type 3 (régulier ou rationnel) sont reconnaissables par un automate d'états finis à reconnaissance de motifs. Un automate d'états finis à reconnaissance de motifs définit de façon unique un langage de type 3.

Exemple d'application

Analyse lexicale de l'affectation numérique

Une grammaire $G =< T, N, mot, P >$ peut être définie par :

- $T = lettre \cup chiffre \cup \{+, -, *, /, =, .\}$
- $N = \{mot, nombre, operateur, identifiant\}$
- $P =$

$$\left\{ \begin{array}{ll} mot \rightarrow operateur, & operateur \rightarrow +|-|*|/|=, \\ mot \rightarrow nombre, & nombre \rightarrow (chiffre^{+})|(chiffre^{+}).(chiffre^{+}), \\ mot \rightarrow identifiant, & identifiant \rightarrow lettre\ |identifiant(lettre\ |chiffre) \end{array} \right\}$$

Langage reconnaissable par l'automate suivant :

Evènement Etat avant	a..z A..Z	0..9	+ - / *	=	.	◇
debut	identifiant	entier	opérateur	affect		debut
identifiant	identifiant	identifiant				identifiant
entier		entier			point	nombre
point		réel				point
réel		réel				nombre
opérateur						opérateur
affect						affect

Analyse lexicale pour l'affectation numérique