AUTOMATES, LANGAGES ET LOGIQUE

Yann PEQUIGNOT

Jacques DUPARC
Alessandro FACCHINI

Section de Mathématiques | Projet de semestre | Automne 2009

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Résumé

Cet exposé porte sur les langages reconnaissables par des automates finis. Le premier chapitre porte sur les différentes notions d'automates finis. Après avoir défini les automates sur les mots finis et avoir exposé leurs propriétés de base, nous étudions les automates sur les mots infinis en comparant l'expressivité des différentes conditions d'acceptation et les propriétés de clôture de l'ensemble des langages qu'elles sont capables de reconnaître.

Le deuxième chapitre est consacré à la démonstration du Théorème de McNaughton selon la méthode de détermination dite de Safra. Ce résultat permet d'achever le tableau commencé au chapitre premier en établissant l'équivalence expressive de toutes les notions d'automates excepté celle d'automate de Büchi déterministe dont l'expressivité est plus faible.

Le troisième chapitre est dévolu à la démonstration du Théorème de Büchi selon lequel la logique monadique du second ordre sur les mots possède la même expressivité que la notion d'automate fini. Nous terminons ce chapitre sur deux conséquences de ce résultat en logique.

L'exposition de chacun des sujets suit principalement [Perrin and Pin, 2004].

TABLE DES MATIÈRES

Résumé 1

Notations 5

I Les automates sur les mots 7
- I.1 Les automates sur les mots finis 7
- I.2 Les automates sur les mots infinis 10
- I.3 Automate de Büchi 11
  - I.3.1 Automate déterministe de Büchi 11
  - I.3.2 Automate non déterministe de Büchi 14
- I.4 Automate de Muller 16
  - I.4.1 Automate non déterministe de Muller 17
  - I.4.2 Automate déterministe de Muller 18
- I.5 Automates de Rabin et de Strett 20

II Le Théorème de McNaughton 25
- II.1 Le Théorème de McNaughton 25
- II.2 Les arbres 26
- II.3 La construction de Safra 27

III Automates et logique monadique du second ordre 33
- III.1 La logique monadique du second ordre sur les mots 33
  - III.1.1 La syntaxe 33
  - III.1.2 La sémantique 34
- III.2 MSO et les automates 36
  - III.2.1 Une formule pour chaque automate 37
  - III.2.2 Un automate pour chaque formule 38
- III.3 Quelques conséquences pour la logique 42

Bibliographie 45

Notations

Soit E est un ensemble fini. Nous utilisons les notations suivantes.

$\omega = \{0, 1, 2, \ldots\}$	dénote l'ensemble des nombres naturels.
$u_0 u_1 \ldots u_{n-1} \in E^n$	représente une suite finie $u : n = \{0, \ldots, n-1\} \to E$.
$E^* = \bigcup_{n \in \omega} E^n$	dénote l'ensemble des suites finies sur E avec la convention que $E^0 = \{\epsilon\}$, où ϵ est la suite vide.
$E^+ = E^* \backslash \{\epsilon\}$	dénote l'ensemble des suites non vides finies sur E.
$\|u\|$	dénote la longueur de la suite u, elle vaut n lorsque $u \in E^n$.
$\mathscr{P}(E)$	dénote l'ensemble des sous-ensembles de E.
E^ω	dénote l'ensemble des suites infinies $\alpha : \omega \to E$, notées $\alpha_0 \alpha_1 \cdots$
$\mathrm{Inf}(\alpha)$	dénote pour $\alpha \in E^\omega$ l'ensemble des éléments de E qui apparaissent infiniment souvent dans α, formellement : $$\mathrm{Inf}(\alpha) = \{e \in E \mid \forall n \in \omega . \exists i > n . \alpha(i) = e\}$$.
$\alpha[m, n]$	dénote pour $\alpha \in E^\omega$ et pour des entiers $m \leq n$ la restriction $\alpha_m \cdots \alpha_n$ de α à $\{m, m+1, \ldots, n\}$.
$\mathrm{Occ}(u)$	dénote pour une suite finie ou infinie sur E, l'ensemble des éléments de E qui apparaissent dans u, i.e. $$\mathrm{Occ}(u) = \{e \in E \mid \exists n \in \omega . u_n = e\}.$$

CHAPITRE I

Les automates sur les mots

Les automates sont en quelque sorte des machines dans le sens où ils peuvent, en principe du moins, être réalisés pratiquement. Étymologiquement, qui se meut de soi-même, les automates considérés ici sont plus exactement capables de lire d'eux-mêmes et d'aboutir à l'acceptation ou au refus d'un mot. Ainsi, ces automates sont potentiellement l'ensemble des mots qu'ils acceptent, leur langage. Bien que souvent infinis, les langages qui sont le propre d'un automate possèdent un caractère particulièrement tangible du fait qu'ils sont potentiellement contenus dans une machine finie.

I.1 Les automates sur les mots finis

Définition I.1.1. Un **automate non déterministe (fini)** (NFA) est la donnée d'un quintuple $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, Acc)$ où :

i) Q est un ensemble fini dont les éléments sont appelés les **états** ;

ii) Σ est un ensemble fini appelé l'**alphabet** et dont les éléments sont appelés des **lettres** ou des **symboles** ;

iii) $\Delta \subseteq Q \times \Sigma \times Q$ est une relation appelée la **relation de transition** ;

iv) $q_I \in Q$ est un état appelé l'**état initial** ;

v) Acc est la **composante d'acceptation**.

Un **automate déterministe (fini)** (DFA) est la donnée d'un automate non déterministe $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, Acc)$ dont la relation de transition Δ est une fonction $\Delta : Q \times \Sigma \to Q$, i.e. pour tout $q \in Q$ et tout $a \in \Sigma$ il existe un unique $p \in Q$ tel que $(q, a, p) \in \Delta$.

Par définition, les DFAs sont des cas particuliers de NFAs. Nous parlons simplement d'un automate pour désigner un NFA. La composante d'acceptation sera précisée de diverses manière dans la suite.

Un automate $\mathfrak{A}$ possède un ensemble Σ appelé alphabet. Dans ce cadre, une suite finie sur l'alphabet est appelée un **mot fini** et un ensemble de mots finis $L \subseteq \Sigma^*$ est appelé un **langage sur les mots finis**. Les automates lisent des mots finis sur leur alphabet de la façon suivante.

Définition I.1.2. Soient $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F, Acc)$ un NFA et $u = u_0 u_1 \dots u_{n-1} \in \Sigma^*$ un mot fini sur Σ. Un **déroulement** de $\mathfrak{A}$ pour u est une suite finie d'états $d = d_0 d_1 \dots d_n \in Q^*$ telle que :

i) $d_0 = q_I$;

ii) pour tout $1 \leq k \leq n$, $(d_{k-1}, u_{k-1}, d_k) \in \Delta$.

Nous notons $\mathscr{D}(\mathfrak{A}, u)$ l'ensemble des déroulements de l'automate $\mathfrak{A}$ pour le mot fini u. Dans le cas d'un NFA, $\mathscr{D}(\mathfrak{A}, u)$ peut être vide ou contenir plusieurs éléments. Remarquer par contre que dans le cas où $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F)$ est un automate déterministe, pour tout mot fini $u \in \Sigma^*$, il y a exactement un déroulement $d \in Q^*$ de $\mathfrak{A}$ pour u et donc $\mathscr{D}(\mathfrak{A}, u) = \{d\}$.

Un automate ne se contente pas de lire des mots, il en accepte certains. Nous précisons à présent en quoi consiste la composante d'acceptation dans le cas des mots finis.

Définition I.1.3. Un automate (non) déterministe $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F)$ où la composante d'acceptation est un sous-ensemble d'états $F \subseteq \mathscr{P}(Q)$ appelés **états finaux** est appelé un **automate (non) déterministe sur les mots finis** s'il **accepte** un mot fini $u = u_0 u_1 \dots u_{n-1} \in \Sigma^*$ sur Σ s'il existe $d = d_0 d_1 \dots d_n \in \mathscr{D}(\mathfrak{A}, u)$ vérifiant la **condition d'acceptation** $d_n \in F$.

Le **langage reconnu** par un automate sur les mots finis $\mathfrak{A}$ est défini comme

$$l(\mathfrak{A}) = \{u \in \Sigma^* \mid \mathfrak{A} \text{ accepte } u\}.$$

Fixons à présent un alphabet fini Σ contenant m lettres et considérons les automates sur cet alphabet. Pour chaque naturel $n \in \omega$, les automates sur Σ qui possèdent n états sont en nombre fini, à savoir $m \cdot n \cdot 2^n \cdot n \cdot 2^n$. Par conséquent, les automates sur Σ sont en nombre infini dénombrable. Or, l'ensemble Σ^* des mots finis sur Σ ayant lui aussi pour cardinal $\aleph_0$, l'ensemble $\mathscr{P}(\Sigma^*)$ des langages sur Σ possède la cardinalité du continu. Il existe donc forcément des langages qui ne sont reconnus par aucun automate. Nous formulons donc la définition suivante.

Définition I.1.4. Un langage $L \subseteq \Sigma^*$ est appelé **régulier** s'il existe un NFA sur les mots finis $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F)$ qui reconnaît L, i.e. tel que $l(\mathfrak{A}) = L$. L'ensemble des langages réguliers sur Σ est noté $\mathrm{REG}(\Sigma)$.

Remarque I.1.5. Pour un alphabet fini Σ, n'importe quel langage $L \subseteq \Sigma^*$ est reconnu par l'automate possédant un nombre infini dénombrable d'états donné par $(\Sigma^*, \Sigma, \Delta, \{\epsilon\}, L)$ avec la relation de transition qui représente la concaténation, i.e.

$$(u, a, v) \in \Delta \Leftrightarrow ua = v.$$

Deux automates $\mathfrak{A}$ et $\mathfrak{B}$ sur un même alphabet sont dits **équivalents** s'ils reconnaissent le même langage, c'est-à-dire si $l(\mathfrak{A}) = l(\mathfrak{B})$. Il est étonnant d'observer que les automates non déterministes sur les mots finis ne sont pas plus expressifs que leurs homologues déterministes dans la mesure où :

Théorème I.1.6. *Tout automate non déterministe sur les mots finis est équivalent à un automate déterministe sur les mots finis.*

Preuve. Soit $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F)$ un NFA. Définissons le DFA $\mathfrak{A}' = (Q', \Sigma, \Delta', q_I', F')$ par :

a) $Q' = \mathscr{P}(Q)$;

b) Pour tout $U \subseteq Q$ et tout $a \in \Sigma$:

$$\Delta'(U,a) = \{q \in Q \mid \exists p \in U.(p,a,q) \in \Delta\} \in Q';$$

c) $q'_I = \{q_I\}$;

d) $F' = \{R \subseteq Q \mid R \cap F \neq \emptyset\}$.

Nous montrons que $l(\mathfrak{A}) = l(\mathfrak{A}')$. Soit $u \in l(\mathfrak{A})$ et $d = d_0 \ldots d_n \in \mathscr{D}(\mathfrak{A}, u)$ un déroulement de $\mathfrak{A}$ pour u avec $d_n \in F$. Notons $D = D_0 \ldots D_n \in (Q')^*$ l'unique déroulement de $\mathfrak{A}'$ pour u. Alors par définition de $\mathfrak{A}'$, nous avons pour tout $0 \leq k \leq n$ que $d_k \in D_k$. En particulier, nous avons $d_n \in D_n \cap F \neq \emptyset$ et donc $D_n \in F'$. Ceci signifie que u est accepté par $\mathfrak{A}'$ et donc $u \in l(\mathfrak{A}')$. Par conséquent, $l(\mathfrak{A}) \subseteq l(\mathfrak{A}')$.

Pour montrer l'autre inclusion, soit $u = u_0 \ldots u_{n-1} \in l(\mathfrak{A}')$ arbitraire. Nous notons $D = D_0 \ldots D_n \in (Q')^*$ l'unique déroulement de $\mathfrak{A}'$ pour u. Puisque u est accepté par $\mathfrak{A}'$, $D_n \in F'$ et il existe donc $d_n \in D_n \cap F$. Par définition de la fonction de transition de $\mathfrak{A}'$, il existe $d_{n-1} \in D_{n-1}$ avec $(d_{n-1}, u_{n-1}, d_n) \in \Delta$. Par induction finie, nous obtenons un déroulement $d_0 \ldots d_n \in Q^*$ de $\mathfrak{A}$ pour u et puisque $d_n \in F$, $\mathfrak{A}$ accepte u. Ceci montre que $l(\mathfrak{A}) \supseteq l(\mathfrak{A}')$. □

En corollaire, nous avons qu'un langage $L \subseteq \Sigma^*$ est régulier si et seulement s'il existe un automate déterministe fini qui le reconnaît.

En outre, l'ensemble des langages réguliers sur un alphabet fini Σ possède une caractérisation équivalente. Considérons les opérations suivantes sur les langages sur Σ. Pour $A, B \subseteq \Sigma^*$,

$$\begin{array}{lrcl}
\text{(union)} & A + B & = & A \cup B \\
\text{(concaténation)} & AB & = & \{uv \mid u \in A \text{ et } v \in B\} \\
\text{(étoile)} & A^* & = & \{u_0 \cdots u_n \mid n \in \omega \text{ et } u_i \in A \text{ pour tout } i \in \{0, \ldots, n\}\} \cup \{\epsilon\}
\end{array}$$

L'ensemble des **langages rationnels** sur Σ sont alors défini comme la clôture minimale de l'ensemble :

$$R_0(\Sigma) = \{\{a\} \mid a \in \Sigma\} \cup \{\emptyset\} \cup \{\{\epsilon\}\},$$

pour les opérations union, concaténation, et étoile.

La preuve du résultat suivant peut être lue dans le chapitre 1 de [Sipser, 2006].

Théorème I.1.7. *L'ensemble* $\mathrm{REG}(\Sigma)$ *des langages réguliers sur* Σ *égale l'ensemble des langages rationnels sur* Σ.

En particulier, ce théorème établi que l'ensemble des langages réguliers est clos pour les opérations d'union, concaténation et étoile. L'ensemble des langages réguliers jouit aussi d'une autre propriété.

Proposition I.1.8. *Soit* $L \in \mathrm{REG}(\Sigma_1 \times \Sigma_2)$. *Notons* $\pi_1 : \Sigma_1^* \times \Sigma_2^* \to \Sigma_1^*$ *la projection sur la première composante. Le langage sur* Σ_1

$$\pi_1(L) = \{u_1 \in \Sigma_1^* \mid \exists u_2 \in \Sigma_2^* \, (u_1, u_2) \in L\}$$

appartient à $\mathrm{REG}(\Sigma_1)$.

Preuve. Remarquons tout d'abord que nous avons la bijection $\Sigma_1^* \times \Sigma_2^* = (\Sigma_1 \times \Sigma_2)^*$. Considérons un NFA sur les mots finis $\mathfrak{A} = (Q, \Sigma_1 \times \Sigma_2, \Delta, q_I, F)$ qui reconnaît L. Le langage $\pi_1(L)$ est reconnu par le NFA sur les mots finis $\mathfrak{A}_1 = (Q, \Sigma_1, \Delta_1, q_I, F)$ où

$$\Delta_1 = \{(p, a_1, q) \in Q \times \Sigma_1 \times Q \mid \exists a_2 \in \Sigma_2 \ \big(p, (a_1, a_2), q\big) \in \Delta\}.$$

En effet, si $u_1 = u_{1,0} u_{1,1} \cdots u_{1,n-1} \in l(\mathfrak{A}_1)$ alors il existe un déroulement $d = d_0 d_1 \cdots d_n \in Q^*$ de $\mathfrak{A}_1$ pour u_1 tel que $d_n \in F$. Choisissons alors une fonction $r : \Delta_1 \to \Sigma_2$ telle que pour tous $(p, a_1, q) \in \Delta_1$, nous avons $\big(p, (a_1, r(p, a_1, q)), q\big) \in \Delta$. Alors le mot $u_2 \in \Sigma_2$ défini par

$$u_{2,i} = r(d_i, u_{1,i}, d_{i+1}) \quad i = 0, \ldots, n-1,$$

est tel que d est un déroulement de $\mathfrak{A}$ pour le mot $(u_1, u_2) \in \Sigma_1^* \times \Sigma_2^*$ et donc $(u_1, u_2) \in l(\mathfrak{A})$. Par conséquent, $u_1 = \pi_1(u_1, u_2) \in \pi_1(l(\mathfrak{A}))$.

Réciproquement, si $u_1 \in \pi_1(l(\mathfrak{A}))$ alors il existe $u_2 \in \Sigma^*$ tel que $(u_1, u_2) \in L(\mathfrak{A})$. Un déroulement acceptant de $\mathfrak{A}$ pour (u_1, u_2) est un déroulement acceptant de $\mathfrak{A}_1$ pour u_1. Par conséquent, $u_1 \in l(\mathfrak{A}_1)$. □

I.2 Les automates sur les mots infinis

Étant donné un alphabet fini Σ, les suites infinies sur Σ sont appelées des **mots infinis**. Un automate lit un mot infini sur son alphabet de la manière suivante.

Définition I.2.1. Soient $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, Acc)$ un NFA et $\alpha = \alpha_0 \alpha_1 \ldots \in \Sigma^\omega$ un mot infini sur Σ. Un **déroulement** de $\mathfrak{A}$ pour α est une suite infinie d'états $\rho = \rho_0 \rho_1 \ldots \in Q^\omega$ telle que

i) $\rho_0 = q_I$;

ii) pour tout $i \in \omega$, $(\rho_i, \alpha_i, \rho_{i+1}) \in \Delta$.

De la même façon que pour la lecture des mots finis, nous notons $\mathscr{D}(\mathfrak{A}, \alpha) \subseteq Q^\omega$ l'ensemble possiblement vide des déroulements de $\mathfrak{A}$ pour le mot infini α. Remarquer que si $\mathfrak{A}$ est déterministe, alors pour tout mot infini α, il existe un unique déroulement $\rho \in Q^\omega$ de $\mathfrak{A}$ pour α et donc $\mathscr{D}(\mathfrak{A}, \alpha) = \{\rho\}$.

De façon similaire au cas des mots finis, les automates ne vont pas se contenter de lire des mots infinis mais ils vont en accepter certains. Nous allons étudier plusieurs manières de définir l'acceptation d'un mot infini par un automate.

Un **ω-langage** sur un alphabet fini Σ est ensemble de mots infinis sur Σ, i.e. un sous-ensemble de Σ^ω. Le **ω-langage reconnu par** $\mathfrak{A}$ est l'ensemble

$$L(\mathfrak{A}) = \{\alpha \in \Sigma^\omega \mid \mathfrak{A} \text{ accepte } \alpha\},$$

où le sens de « $\mathfrak{A}$ accepte α » sera précisé de différentes façon dans les sous-sections suivantes.

Comme pour les automates sur les mots finis, deux automates $\mathfrak{A}$ et $\mathfrak{B}$ sur un même alphabet, munis d'une composante d'acceptation pour les mots infinis, sont dits **équivalents** s'ils reconnaissent le même ω-langage, i.e. si $L(A) = L(B)$.

I.3 Automate de Büchi

Définition I.3.1. Un automate (non) déterministe $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F)$ où la composante d'acceptation est un sous-ensemble d'états $F \subseteq Q$ appelés **états acceptants** est appelé un **automate (non) déterministe de Büchi** s'il **accepte** un mot infini $\alpha \in \Sigma^\omega$ s'il existe $\rho \in \mathscr{D}(\mathfrak{A}, \alpha)$ vérifiant la **condition de Büchi**

$$\mathrm{Inf}(\rho) \cap F \neq \emptyset.$$

Considérons un automate $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F)$ où la composante d'acceptation est un sous-ensemble d'états $F \subseteq Q$. D'une part, l'automate $\mathfrak{A}$ est appelé un automate sur les mots finis s'il accepte un mot fini selon la condition d'acceptation définie à la section précédente et reconnaît alors le langage sur les mots finis $l(\mathfrak{A}) \subseteq \Sigma^*$. D'autre part, il est appelé un automate de Büchi s'il accepte un mot infini selon la condition de Büchi et reconnaît alors le langage sur les mots infinis $L(\mathfrak{A}) \subseteq \Sigma^\omega$. Par la suite, nous allons largement exploiter ces deux interprétations d'un tel automate et pour un automate de Büchi $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F)$, nous parlerons de son langage sur les mots finis $l(\mathfrak{A}) \subseteq \Sigma^*$ et de son ω-langage $L(\mathfrak{A}) \subseteq \Sigma^\omega$.

Un ω-langage $L \subseteq \Sigma^\omega$ est **Büchi reconnaissable** s'il existe un automate non déterministe de Büchi $\mathfrak{A}$ tel que $L(\mathfrak{A}) = L$. Un ω-langage Büchi reconnaissable est **déterministe Büchi reconnaissable** s'il est reconnu par un automate déterministe de Büchi.

I.3.1 Automate déterministe de Büchi

Nous commençons par décrire le langage reconnu par un automate déterministe de Büchi. Dans ce but nous définissons l'opération suivante qui associe à chaque langage fini un ω-langage. Si $L \subseteq \Sigma^*$ est un langage fini, nous définissons le ω-langage noté $\lim L$ par :

$$\lim L = \{\alpha \in \Sigma^* \mid \forall n \in \omega . \exists i > n . \alpha[0, i] \in L\}.$$

Nous avons le résultat suivant.

Théorème I.3.2. *Soit $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F)$ un automate déterministe de Büchi. Le ω-langage reconnu par $\mathfrak{A}$ est* $\lim l(\mathfrak{A})$.

Preuve. Montrons que $L(\mathfrak{A}) \subseteq \lim l(\mathfrak{A})$. Soit $\alpha \in L(\mathfrak{A})$. Notons $\rho \in Q^\omega$ l'unique déroulement de $\mathfrak{A}$ pour α. Puisque $\mathfrak{A}$ accepte α, pour tout $n \in \omega$ il existe $i > n$ tel que $\rho_i \in F$ et donc $\alpha[0, i-1] \in l(\mathfrak{A})$. Par conséquent, $\alpha \in \lim l(\mathfrak{A})$.

Montrons maintenant que $\lim l(\mathfrak{A}) \subseteq L(\mathfrak{A})$. Soient $\alpha \in \lim l(\mathfrak{A})$ et $\rho \in Q^\omega$ l'unique déroulement de $\mathfrak{A}$ pour α. Pour tout $n \in \omega$, il existe $i > n$ tel que $\alpha[0, i] \in l(\mathfrak{A})$. Puisque $\mathfrak{A}$ est déterministe, $\rho_0 \cdots \rho_{i+1}$ est l'unique déroulement de $\mathfrak{A}$ (en tant qu'automate sur les mots finis) pour $\alpha[0, i]$ et donc $\rho_{i+1} \in F$. Puisque F est fini, il s'ensuit que $\mathrm{Inf}(\rho) \cap F \neq \emptyset$ et par conséquent $\alpha \in L(\mathfrak{A})$. $\square$

Remarquer que si $\mathfrak{A}$ est un automate de Büchi non déterministe, alors le théorème précédent n'est pas vrai. Cependant, si nous regardons $\mathfrak{A}$ comme automate sur les mots finis, le Théorème I.1.6 nous assure qu'il y a un automate déterministe qui reconnaît le même langage fini L. Si nous regardons cet automate déterministe comme un automate de Büchi, alors son langage est $\lim L$ par le théorème précédent. Il reste vrai que $L(\mathfrak{A}) \subseteq \lim L$, cependant l'autre inclusion n'est en général pas vérifiée.

Nous notons $\mathrm{REG}^\omega_{\text{Dét}}(\Sigma)$ les ω-langages déterministe Büchi reconnaissables.

Corollaire I.3.3. *L'ensemble des ω-langages déterministe Büchi reconnaissables sur un alphabet fini Σ est :*

$$\mathrm{REG}^{\omega}_{Dét}(\Sigma) = \{\lim L \mid L \in \mathrm{REG}(\Sigma)\}.$$

Les ω-langages Büchi déterministe reconnaissables possèdent la propriété de clôture suivante.

Proposition I.3.4. *L'ensemble* $\mathrm{REG}^{\omega}_{Dét}(\Sigma)$ *des ω-langages déterministe Büchi reconnaissables est clos par union finie et intersection finie.*

Preuve. Pour l'union, le résultat découle de la formule

$$\lim L_1 \cup \lim L_2 = \lim(L_1 \cup L_2),$$

pour des ensembles $L_1, L_2 \subseteq \Sigma^*$, du fait que $\mathrm{REG}(\Sigma)$ est clos par union finie et du Théorème I.3.2.

Pour l'intersection, nous avons besoin d'une construction spécifique. Considérons deux automates déterministes de Büchi $\mathfrak{A}_i = (Q_i, \Sigma, \Delta_i, I_i, F_i)$ avec $i = 1, 2$. Nous construisons un automate déterministe de Büchi reconnaissant le ω-langage $L(\mathfrak{A}_1) \cap L(\mathfrak{A}_2)$. Posons

$$Q = Q_1 \times Q_2 \times \{0, 1\},$$

et définissons pour tout $(p_1, p_2, s) \in Q$ et tout $a \in \Sigma$,

$$\Delta\big((p_1, p_2, s), a\big) = \big(\Delta_1(p_1, a), \Delta_2(p_2, a), \delta((p_1, p_2, s), a)\big),$$

où la fonction $\delta : Q \times \Sigma \to \{0, 1\}$ est définie par :

$$\delta\big((p_1, p_2, s), a\big) = \begin{cases} 0 & \text{si } (p_2 \in F_2 \wedge \Delta_1(p_1, a) \notin F_1) \vee \big(p_2 \notin F_2 \wedge s = 0 \wedge \Delta_1(p_1, a) \notin F_1\big); \\ 1 & \text{si } (p_2 \in F_2 \wedge \Delta_1(p_1, a) \in F_1) \vee \big(p_2 \notin F_2 \wedge (s = 1 \vee \Delta_1(p_1, a) \in F_1)\big). \end{cases}$$

Nous posons également

$$F = Q_1 \times F_2 \times \{1\}.$$

Nous considérons l'automate déterministe de Büchi $\mathfrak{A} = (Q, \Sigma, \Delta, (I_1, I_2, 1), F)$. Intuitivement, lors d'un déroulement de $\mathfrak{A}$, la composante booléenne d'un état mémorise le fait qu'un état de F_1 a été visité depuis la dernière visite d'un état de F_2. Plus précisément, la composante booléenne est initialisée à 0 directement à la suite d'une visite d'un état de F_2 et demeure nulle jusqu'à ce qu'un état de F_1 soit visité. Si un état de F_1 est alors visité, la composante booléenne prend la valeur 1 à l'état suivant et demeure égale à 1 jusqu'à la visite d'un nouvel état de F_2, moment auquel elle vaut encore 1 avant d'être réinitialisé à 0 à l'état suivant si ce n'est pas un état de F_1, et ainsi de suite.

Nous affirmons que l'automate $\mathfrak{A}$ reconnaît le ω-langage $L(\mathfrak{A}_1) \cap L(\mathfrak{A}_2)$.

Supposons que $\alpha \in L(\mathfrak{A})$ et notons $\rho = \rho_0\rho_1 \cdots \in Q^{\omega}$ l'unique déroulement de $\mathfrak{A}$ pour α. Pour $i = 1, 2$, en projetant sur la $i^{\text{ème}}$ composante des états de ρ, nous obtenons l'unique déroulement $\rho_i = \rho_{i,0}\rho_{i,1}\rho_{i,2} \cdots$ de $\mathfrak{A}_i$ pour α. Il découle du fait que $\mathrm{Inf}(\rho) \cap F \neq \emptyset$ qu'il existe une suite strictement croissante de naturels $(n_k)_{k \in \omega}$ telle que, pour tout $k \in \omega$, l'état $\rho_{n_k} = (\rho_{1,n_k}, \rho_{2,n_k}, s_{n_k})$ possède la propriété que $\rho_{2,n_k} \in F_2$ et $s_{n_k} = 1$. Il s'ensuit directement que $\mathrm{Inf}(\rho_2) \cap F_2 \neq \emptyset$ et par conséquent $\alpha \in L(\mathfrak{A}_2)$. Par ailleurs, pour chaque $k \geq 1$, le fait que $s_{n_k} = 1$ et $\rho_{2,n_{k-1}}, \rho_{2,n_k} \in F_2$ indique qu'il existe $n_{k-1} + 1 \leq m_k \leq n_k$ tel que $\rho_{1,m_k} \in F_1$. Par conséquent $\rho_1 \cap F_1 \neq \emptyset$ et $\alpha \in L(\mathfrak{A}_1)$. Nous concluons que $\alpha \in L(\mathfrak{A}_1) \cap L(\mathfrak{A}_2)$.

Réciproquement, supposons que $\alpha \in L(\mathfrak{A}_1) \cap L(\mathfrak{A}_2)$. Notons $\rho_1 \in Q_1^\omega$, respectivement $\rho_2 \in Q_2^\omega$, l'unique déroulement de $\mathfrak{A}_1$, respectivement $\mathfrak{A}_2$, pour α. Nous avons $\mathrm{Inf}(\rho_1) \cap F_1 \neq \emptyset$ et $\mathrm{Inf}(\rho_2) \cap F_2 \neq \emptyset$. Notons

$$\rho = (\rho_{1,0}, \rho_{2,0}, s_0)(\rho_{1,1}, \rho_{2,1}, s_1) \ldots \in Q^\omega$$

l'unique déroulement de $\mathfrak{A}$ pour α. Maintenant, soit $n \in \omega$ arbitraire. Il existe $n_2 \geq n$ tel que $\rho_{2,n_2} \in F_2$. Pour ce n_2, il existe $n_1 \geq n_2$ tel que $\rho_{1,n_1} \in F_1$. Pour ce naturel n_1, notons alors N_2 le plus petit entier plus grand ou égal à n_1 tel que $\rho_{2,N_2} \in F_2$. Nous avons par définition de δ que $s_{N_2} = 1$ et donc $\rho_{N_2} = (\rho_{1,N_2}, \rho_{2,N_2}, 1) \in F$. Cela montre que $\mathrm{Inf}(\rho) \cap F \neq \emptyset$ et donc que $\alpha \in L(\mathfrak{A})$. □

Nous terminons cette sous-section sur les automates déterministes de Büchi en considérant la complexité topologique des ω-langages qu'ils reconnaissent. Rappelons que Σ est muni de la topologie discrète et Σ^ω de la topologie produite dont les ouverts sont les ensembles de la forme $A\Sigma^\omega$ avec $A \subseteq \Sigma^*$. Nous rappelons de plus que la classe de Borel $\mathbf{\Pi}_2^0$ des intersections dénombrables d'ouverts de Σ^ω est formée exactement des ensembles de la forme $\lim A$ pour $A \subseteq \Sigma^*$. Par souci de complétude, nous exposons la démonstration de ce résultat.

Proposition I.3.5. *Soit $X \subseteq \Sigma^\omega$. Alors,*

$$X \in \mathbf{\Pi}_2^0(\Sigma^\omega) \Longleftrightarrow \exists A \subseteq \Sigma^* \textit{ tel que } X = \lim A.$$

Preuve. ($\Leftarrow$) : Soit $A \subseteq \Sigma^*$ et posons $X = \lim A$. Définissons pour chaque $n \in \omega$,

$$A_n = \{u \in A \mid |u| \geq n\}.$$

Les ensembles $A_n\Sigma^\omega$ sont ouverts, par conséquent leur intersection est dans $\mathbf{\Pi}_2^0(\Sigma^\omega)$. Montrons que $X = \bigcap_{n\in\omega} A_n\Sigma^\omega$. Premièrement, si $x \in \lim A$, alors par définition, pour tout $n \in \omega$, il existe $i > n$ tel que $x[0,i] \in A$ et donc $x[0,i] \in A_n$ et $x \in A_n\Sigma^\omega$. Par conséquent, $x \in \bigcap_{n\in\omega} A_n\Sigma^\omega$. Réciproquement, soit $x \in \bigcap_{n\in\mathbb{N}} A_n\Sigma^\omega$ et supposons qu'il existe $N \in \omega$ tel que pour tout $i > N$ nous avons $x[0,i] \notin A$. Alors, par définition des A_n, il s'ensuit que $x \notin A_{N+1}\Sigma^\omega$ et par conséquent $x \notin \bigcap_{n\in\omega} A_n\Sigma^\omega$. C'est une contradiction et donc nécessairement, nous devons avoir $x \in \lim A$.

($\Rightarrow$) : Soit $X = \bigcap_{n\in\mathbb{N}} U_n$ où chaque U_n est un ouvert de Σ^ω. Posons $V_n = \bigcap_{0\leq i\leq n} U_i$. Chaque V_n est ouvert en tant qu'intersection finie d'ouverts, et s'écrit donc $B_n\Sigma^\omega$ pour un certain $B_n \subseteq \Sigma^*$. Observer de plus que $B_n\Sigma^\omega = (B_n \backslash B_n\Sigma^+)\Sigma^\omega$ et notons $A_n = B_n \backslash B_n\Sigma^+$. De plus, $X = \bigcap_{n\in\omega} V_n$ avec $V_{n+1} \subseteq V_n$ pour chaque $n \in \omega$. Puisque $A_n\Sigma^\omega = A_n\Sigma^n\Sigma^\omega$, nous avons également $X = \bigcap_{n\in\omega} A_n\Sigma^n\Sigma^\omega$. Nous affirmons que $X = \lim A$, où $A = \bigcup_{n\in\omega} A_n\Sigma^n$.

En effet, supposons premièrement que $x \in X$ et qu'il existe $N \in \omega$ tel que pour tout $i > N$ nous avons $x[0,i] \notin A$. Puisque $x \in X = \bigcap_{n\in\mathbb{N}} A_n\Sigma^n\Sigma^\omega$, alors en particulier $x \in A_{N+1}\Sigma^{N+1}\Sigma^\omega$ et donc il existe $i > N$ tel que $x[0,i] \in A_{N+1}\Sigma^{N+1} \subseteq A$. Cela contredit notre hypothèse et donc nécessairement, $x \in \lim A$.

Réciproquement, si $x \in \lim A$, alors il existe une suite strictement croissantes d'entiers $(n_k)_{k\in\omega}$ telle que $x[0,n_k] \in A$ pour tout $k \in \omega$. Il s'ensuit que pour chaque $k \in \omega$ il existe $m_k \in \omega$ tel que $x[0,n_k] \in A_{m_k}\Sigma^{m_k}$. Et par définition des A_n, si $k \neq l$, alors $m_k \neq m_l$. Par conséquent, $x \in A_n\Sigma^n\Sigma^\omega = V_n$ pour une infinité de $n \in \omega$ et puisque la suite des V_n est décroissante nous avons $x \in \bigcap_{n\in\mathbb{N}} V_n = X$. □

Par le Corollaire I.3.3, nous obtenons que les ω-langages déterministes Büchi reconnaissables sur un alphabet fini Σ sont contenus dans $\mathbf{\Pi}^0_2(\Sigma^\omega)$ et le niveau de ces ensembles dans la hiérarchie de Borel est donc borné.

Corollaire I.3.6. *Soit Σ un alphabet fini. Les ω-langages sur Σ reconnus par un automate déterministe de Büchi sont contenus dans la classe de Borel $\mathbf{\Pi}^0_2(\Sigma)$, i.e.*

$$\mathrm{REG}^\omega_{Dét}(\Sigma) \subseteq \mathbf{\Pi}^0_2(\Sigma^\omega).$$

I.3.2 Automate non déterministe de Büchi

À présent, nous caractérisons en termes des langages réguliers les ω-langages qui sont Büchi reconnaissables.

Théorème I.3.7. *Un ω-langage $L \subseteq \Sigma^\omega$ est Büchi reconnaissable si et seulement s'il est de la forme :*

$$L = \bigcup_{i=1}^{n} U_i V_i^\omega \text{ avec } U_1, V_1, \ldots, U_n, V_n \in \mathrm{REG}(\Sigma).$$

Preuve. 1. Montrons la première implication. Soit $L \subseteq \Sigma^\omega$ un ω-langage reconnu par un automate non déterministe $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F)$. Notons pour tout couple d'états $(p, q) \in Q \times Q$ par $\mathfrak{A}_{pq}$ l'automate $(Q, \Sigma, \Delta, p, \{q\})$. Vu comme un automate sur les mots finis, $\mathfrak{A}_{pq}$ reconnaît le langage $L_{pq} \subseteq \Sigma^*$. Nous affirmons que

$$L = \bigcup_{q \in F} L_{q_I q} L^\omega_{qq}.$$

En effet, supposons que $\alpha \in L$ et notons $\rho \in Q^\omega$ un déroulement de $\mathfrak{A}$ pour α tel que $\mathrm{Inf}(\rho) \cap F \neq \emptyset$. Il existe alors un état acceptant $q \in F$ et $(n_k)_{k\in\omega}$ une suite strictement croissante d'entiers tels que $\rho_{n_k} = q$ pour tous $k \in \omega$. Il s'ensuit d'une part que $\alpha[0, n_0] \in L_{q_I q}$ et d'autre part que $\alpha[n_{k-1} + 1, n_k] \in L_{q,q}$ pour tout $k \geq 1$. Par conséquent, $\alpha \in \bigcup_{q\in F} L_{q_I q} L^\omega_{qq}$.

Supposons maintenant que $\alpha \in L_{q_I q} L^\omega_{qq}$ pour un $q \in F$ et soit $(n_k)_{k\in\omega}$ une suite strictement croissante d'entiers telle que $\alpha[0, n_0] \in L_{q_I q}$ et $\alpha[n_{k-1} + 1, n_k] \in L_{qq}$ pour tout $k \geq 1$. Alors il existe un déroulement ρ^0 de $\mathfrak{A}_{q_I q}$ pour $\alpha[0, n_0]$ qui se termine en q et pour tout $k \geq 1$ il existe un déroulement ρ^k de $\mathfrak{A}_{qq}$ pour $\alpha[n_{k-1} + 1, n_k]$ qui se termine en l'état q. Nous formons donc la suite d'états $\rho = \rho^1 \rho^2 \rho^3 \cdots$ qui est un déroulement de $\mathfrak{A}$ pour α tel que $q \in \mathrm{Inf}(\rho) \cap F$. Par conséquent $\alpha \in L$.

2.. La deuxième implication découle des trois résultats suivants.

- $2a$. $L \in \mathrm{REG}(\Sigma) \Rightarrow L^\omega$ Büchi reconnnaissable.
- $2b$. $L \in \mathrm{REG}(\Sigma)$ et $K \subseteq \Sigma^\omega$ Büchi reconnnaissable $\Rightarrow LK$ Büchi reconnnaissable.
- $2c$. $L_1, L_2 \subseteq \Sigma^\omega$ Büchi reconnnaissables $\Rightarrow L_1 \cup L_2$ Büchi reconnnaissable.

Nous démontrons ces trois affirmations.

2a. Soient $L \in \mathrm{REG}(\Sigma)$ et $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F)$ un DFA sur les mots finis reconnaissant L. Nous définissons l'automate non déterministe de Büchi $\mathfrak{B} = (Q \cup \{q_0\}, \Sigma, \Delta', q_0, \{q_0\})$, où $q_0 \notin Q$, en stipulant sa relation de transition comme étant :

$$\Delta' = \Delta \cup \{(q_0, a, q) \mid (q_I, a, q) \in \Delta\} \cup \{(q, a, q_0) \mid \exists p \in F.(q, a, p) \in \Delta\}.$$

Noter en premier lieu que $\mathfrak{B}$ en tant qu'automate sur les mots finis reconnaît le langage $L \cup \{\epsilon\}$.

Pour voir que $L(\mathfrak{B}) = L^\omega$, supposons en premier lieu que $\alpha \in L(\mathfrak{B})$ et soit $\rho \in (Q \cup \{q_0\})^\omega$ un déroulement de $\mathfrak{B}$ pour α tel que $q_0 \in \mathrm{Inf}(\rho)$. Notons $(n_k)_{k\in\omega}$ la suite strictement croissante des entiers $n \in \omega$ tel que $\rho_n = q_0$. Alors, par définition de l'automate $\mathfrak{B}$, $\alpha[0, n_1] \in L$ et $\alpha[n_k + 1, n_{k+1}] \in L$ pour tout $k \geq 1$. Par conséquent, $\alpha \in L^\omega$.

Supposons maintenant que $\alpha = \alpha^0\alpha^1 \cdots \in L^\omega$ avec $\alpha^i \in L$ pour tout $i \in \omega$. Alors par définition de $\mathfrak{B}$, pour chaque $i \in \omega$ il existe un déroulement fini ρ^i de $\mathfrak{B}$ pour α^i se terminant en q_0. La suite infinie d'états de $\mathfrak{B}$ donnée par $\rho = \rho^1\rho^2 \cdots$ est alors un déroulement de $\mathfrak{B}$ pour α et clairement $q_0 \in \mathrm{Inf}(\rho)$. Par conséquent, $\alpha \in L(\mathfrak{B})$.

2b. Soient $L \in \mathrm{REG}(\Sigma)$ reconnu par un DFA sur les mots finis $\mathfrak{A}_1 = (Q_1, \Sigma, \Delta_1, q_1, F_1)$ et $K \subseteq \Sigma^\omega$ un ω-langage reconnu par un automate non déterministe de Büchi $\mathfrak{A}_2 = (Q_2, \Sigma, \Delta_2, q_2, F_2)$ avec $Q_1 \cap Q_2 = \emptyset$. L'automate non déterministe de Büchi $\mathfrak{B} = (Q_1 \cup Q_2, \Sigma, \Delta_\mathfrak{B}, q_1, F_2)$ avec la relation de transition :

$$\Delta_\mathfrak{B} = \begin{cases} \Delta_1 \cup \Delta_2 \cup \{(p, a, q_2) \mid \exists q \in F_1.(p, a, q) \in \Delta_1\} & \text{si } q_1 \notin F_1, \\ \Delta_1 \cup \Delta_2 \cup \{(p, a, q_2) \mid \exists q \in F_1.(p, a, q) \in \Delta_1\} \cup \{(q_1, a, q) \mid (q_2, a, q) \in \Delta_2\} & \text{si } q_1 \in F_1, \end{cases}$$

reconnaît le langage LK.

2c. Soient $L_1, L_2 \subseteq \Sigma^\omega$ des ω-langages reconnus respectivement par des automates non déterministes de Büchi $\mathfrak{A}_1 = (Q_1, \Sigma, \Delta_1, q_1, F_1)$ et $\mathfrak{A}_1 = (Q_2, \Sigma, \Delta_2, q_2, F_2)$ avec $Q_1 \cap Q_2 = \emptyset$. Définissons l'automate non déterministe de Büchi $\mathfrak{B} = (Q_1 \cup Q_2, \Sigma, \Delta_\mathfrak{B}, q_I, F_1 \cup F_2)$ avec $q_I \notin Q_1$ et $q_I \notin Q_2$ et avec relation de transition :

$$\Delta_\mathfrak{B} = \Delta_1 \cup \Delta_2 \cup \{(q_I, a, p) \mid (q_1, a, p) \in \Delta_1\} \cup \{(q_I, a, q) \mid (q_2, a, q) \in \Delta_2\}.$$

L'automate non déterministe de Büchi $\mathfrak{B}$ reconnaît le langage $L_1 \cap L_2$. □

Nous notons $\mathrm{REG}^\omega(\Sigma)$ l'ensemble des ω-langages Büchi reconnaissables sur l'alphabet fini Σ. Cet ensemble est aussi appelé la ω-clôture de Kleene de $\mathrm{REG}(\Sigma)$. Le dernier théorème nous dit que

$$\mathrm{REG}^\omega(\Sigma) = \left\{ \bigcup_{i=1}^n L_i K_i^\omega \;\middle|\; n \in \omega \text{ et } L_1, K_1, \ldots, L_n, K_n \in \mathrm{REG}(\Sigma) \right\}.$$

Les ω-langages Büchi reconnaissables jouissent également de la propriété suivante qui est l'analogue pour les ω-langages de la Proposition I.1.8.

Proposition I.3.8. *Soit $L \in \mathrm{REG}^\omega(\Sigma_1 \times \Sigma_2)$. Notons $\pi_1 : \Sigma_1^\omega \times \Sigma_2^\omega \to \Sigma_1^\omega$ la projection sur la première composante. Le ω-langage sur Σ_1*

$$\pi_1(L) = \{\alpha_1 \in \Sigma_1^\omega \mid \exists \alpha_2 \in \Sigma_2^\omega\, (\alpha_1, \alpha_2) \in L\}$$

appartient à $\mathrm{REG}^\omega(\Sigma_1)$.

Preuve. Remarquons tout d'abord que nous avons la bijection $\Sigma_1^\omega \times \Sigma_2^\omega = (\Sigma_1 \times \Sigma_2)^\omega$. Considérons un automate non déterministe de Büchi $\mathfrak{A} = (Q, \Sigma_1 \times \Sigma_2, \Delta, q_I, F)$ qui reconnaît L. Le langage $\pi_1(L)$ est reconnu par l'automate non déterministe de Büchi $\mathfrak{A}_1 = (Q, \Sigma_1, \Delta_1, q_I, F)$ où

$$\Delta_1 = \{(p, a_1, q) \in Q \times \Sigma_1 \times Q \mid \exists a_2 \in \Sigma_2\, \big(p, (a_1, a_2), q\big) \in \Delta\}.$$

En effet, si $\alpha_1 = \alpha_{1,0}\alpha_{1,1} \cdots \in L(\mathfrak{A}_1)$ alors il existe un déroulement $\rho = \rho_0\rho_1 \cdots \in Q^\omega$ de $\mathfrak{A}_1$ pour α_1 tel que $\mathrm{Inf}(\rho) \cap F \neq \emptyset$. Choisissons alors une fonction $r : \Delta_1 \to \Sigma_2$ telle que pour tous

$(p, a_1, q) \in \Delta_1$, nous avons $\big(p, (a_1, r(p, a_1, q)), q\big) \in \Delta$. Alors le mot infini $\alpha_2 \in \Sigma_2$ défini par pour tout $i \in \omega$ par

$$u_{2,i} = r(d_i, u_{1,i}, d_{i+1}),$$

est tel que ρ est un déroulement de $\mathfrak{A}$ pour le mot $(\alpha_1, \alpha_2) \in \Sigma_1^\omega \times \Sigma_2^\omega$ et donc $(\alpha_1, \alpha_2) \in L(\mathfrak{A})$. Par conséquent, $\alpha_1 = \pi_1(\alpha_1, \alpha_2) \in \pi_1(L(\mathfrak{A}))$.

Réciproquement, si $\alpha_1 \in \pi_1(L(\mathfrak{A}))$ alors il existe $\alpha_2 \in \Sigma^*$ tel que $(\alpha_1, \alpha_2) \in L(\mathfrak{A})$. Un déroulement acceptant de $\mathfrak{A}$ pour (α_1, α_2) est un déroulement acceptant de $\mathfrak{A}_1$ pour α_1. Par conséquent, $\alpha_1 \in L(\mathfrak{A}_1)$. □

Nous observons maintenant que si l'alphabet contient au moins deux symboles, alors les automates déterministes de Büchi ont une moins grande capacité expressive que les automates non déterministes de Büchi.

Théorème I.3.9. *Soit Σ un alphabet fini de cardinalité plus grande ou égale à 2. Il existe des ω-langages qui sont Büchi reconnaissables mais qui ne sont pas déterministe Büchi reconnaissables.*

Preuve. Soient $a, b \in \Sigma$ distincts. Par le Théorème I.3.7, le ω-langage $\{a,b\}^*\{a\}^\omega$ est Büchi reconnaissable. Supposons par l'absurde qu'il est déterministe Büchi reconnaissable, alors par le Théorème I.3.2 il existe $L \in \mathrm{REG}(\Sigma)$ tel que $\{a,b\}^*\{a\}^\omega = \lim L$. Puisque $ba^\omega \in \{a,b\}^*\{a\}^\omega$, il s'ensuit qu'il existe $n_0 \in \omega$ tel que $ba^{n_0} \in L$. Par induction, nous définissons pour tout $k \geq 0$ un entier $n_{k+1} \in \omega$ tel que $ba^{n_0}ba^{n_1}\cdots ba^{n_k}ba^{n_{k+1}} \in L$. Cela est possible puisque par hypothèse

$$ba^{n_0}\cdots ba^{n_k}ba^\omega \in \lim L = \{a,b\}^*\{a\}^\omega.$$

Par conséquent, le mot infini $ba^{n_0}ba^{n_1}\cdots \in \lim L$ mais puisqu'il contient une infinité de b, il n'appartient pas $\{a,b\}^*\{a\}^\omega$. Nous avons donc une contradiction et par conséquent $\{a,b\}^*\{b\}^\omega$ n'est pas déterministe Büchi reconnaissable. □

Le dernier théorème établi que si $\mathrm{Card}(\Sigma) \geq 2$, alors $\mathrm{REG}^\omega_{\text{Dét}}(\Sigma) \subsetneq \mathrm{REG}^\omega(\Sigma)$. Par conséquent, en se référant au Corollaire I.3.3 et à la Proposition I.3.5, lorsque $\mathrm{Card}(\Sigma) \geq 2$, il existe des ω-langages Büchi reconnaissables sur Σ qui n'appartiennent pas à la classe de Borel $\mathbf{\Pi}^0_2(\Sigma^\omega)$.

I.4 Automate de Muller

Nous nous intéressons à une autre condition d'acceptation de mots infinis par un automate.

Définition I.4.1. Un automate (non) déterministe $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, \mathscr{F})$ où la composante d'acceptation est un ensemble de sous-ensembles d'états $\mathscr{F} \subseteq \mathscr{P}(Q)$ est appelé un **automate (non) déterministe de Muller** s'il **accepte** un mot infini $\alpha \in \Sigma^\omega$ s'il existe $\rho \in \mathscr{D}(\mathfrak{A}, \alpha)$ vérifiant la **condition de Muller**

$$\mathrm{Inf}(\rho) \in \mathscr{F}.$$

Remarquons immédiatement qu'étant donné un automate de Büchi, il est aisé de donner un automate de Muller qui lui soit équivalent. En outre, cette transformation fournit un automate déterministe de Muller dans le cas où l'automate de Büchi est déterministe.

Théorème I.4.2. *Soit $L \subseteq \Sigma^\omega$ un ω-langage reconnu par un automate (non) déterministe de Büchi $\mathfrak{B} = (Q, \Sigma, \Delta, q_I, F)$. L'automate (non) déterministe de Muller $\mathfrak{M} = (Q, \Sigma, \Delta, q_I, \mathscr{F})$ où $\mathscr{F} = \{G \in \mathscr{P}(Q) \mid G \cap F\}$ reconnaît L.* □

I.4.1 Automate non déterministe de Muller

Nous montrons à présent que si $L \subseteq \Sigma^\omega$ est un ω-langage reconnu par un automate non déterministe de Muller, alors il est reconnu par un automate non déterministe de Büchi.

Théorème I.4.3. *Soit $L \subseteq \Sigma^\omega$ un ω-langage reconnu par un automate non déterministe de Muller $\mathfrak{M} = (Q, \Sigma, \Delta, q_I, \mathscr{F})$. Il existe un automate non déterministe de Büchi qui reconnaît L.*

Preuve. Nous définissons un automate non déterministe de Büchi $\mathfrak{B}$ sur l'alphabet Σ dont l'ensemble des états est

$$Q' = Q \cup \bigcup_{G \in \mathscr{F}} \{G\} \times G \times \mathscr{P}(G),$$

la relation de transition est

$$\begin{aligned}\Delta' = \Delta \cup &\{(p, a, (G, q, \emptyset)) \mid G \in \mathscr{F} \text{ et } q \in G \text{ et } (p, a, q) \in \Delta\} \\ \cup \bigcup_{G \in \mathscr{F}} &\{((G, p, R), a, (G, q, R \cup \{q\})) \mid p, q \in G \text{ et } R \subsetneq G \text{ et } (p, a, q) \in \Delta\} \\ \cup \bigcup_{G \in \mathscr{F}} &\{((G, p, G), a, (G, q, \emptyset)) \mid p, q \in G \text{ et } (p, a, q) \in \Delta\},\end{aligned}$$

l'état initial est $q_I \in Q'$ et l'ensemble des états acceptants est

$$F' = \{(G, q, \emptyset) \mid G \in \mathscr{F} \text{ et } q \in G\}.$$

Montrons que l'automate non déterministe de Büchi $\mathfrak{B} = (Q', \Sigma, \Delta', q_I, F')$ reconnaît le ω-langage L de l'automate de Müller $\mathfrak{M}$.

Supposons que $\alpha \in L(\mathfrak{B})$ et notons $\rho \in (Q')^\omega$ un déroulement de $\mathfrak{B}$ pour α tel que $\mathrm{Inf}(\rho) \cap F' \neq \emptyset$. Il existe alors $G \in \mathscr{F}$ et $q \in G$ tel que $(G, q, \emptyset) \in \mathrm{Inf}(\rho)$. Notons $N \in \omega$ le plus grand entier tel que $\rho_N \in Q$ et remarquons que par définition de la relation de transition de $\mathfrak{B}$, nous avons $\rho_n \in \{G\} \times G \times \mathscr{P}(G)$ pour tout $n \geq N+1$. Notons alors $\pi : \{G\} \times G \times \mathscr{P}(G) \to G$ la projection sur la deuxième composante et observons que $\tilde{\rho} = \rho_0 \cdots \rho_N \pi(\rho_{N+1}) \pi(\rho_{N+2}) \cdots$ est un déroulement de $\mathfrak{M}$ pour α. Puisque $\pi(\rho_n) \in G$ pour tout $n \geq N+1$, nous avons $\mathrm{Inf}(\tilde{\rho}) \subseteq G$. De plus,par définition de $\mathfrak{B}$ à partir $\rho_{N+1} = (G, p, \emptyset)$ la troisième composante des états est un sous-ensemble de G qui garde en mémoire les états de G visités, et remet le compteur à $\emptyset$ lorsque tous les états de G ont été visités. Par conséquent, le fait que $(G, q, \emptyset) \in \mathrm{Inf}(\rho)$ implique que $\mathrm{Inf}(\tilde{\rho}) = G$ et donc $\alpha \in L$.

Réciproquement, supposons que $\alpha \in L = L(\mathfrak{M})$ et soit $\rho \in Q^\omega$ un déroulement de $\mathfrak{M}$ pour α tel que $\mathrm{Inf}(\rho) = G$ pour un $G \in \mathscr{F}$. Il existe alors $N \in \omega$ tel que $\rho_n \in G$ pour tout $n \geq N+1$ et nous prenons N minimal. Nous définissons une suite infinie d'états de $\mathfrak{B}$ par :

$$\begin{aligned}\tilde{\rho}_n &= \rho_n \text{ pour tout } 0 \leq n \leq N, \\ \tilde{\rho}_{N+1} &= (G, \rho_{N+1}, \emptyset)\end{aligned}$$

puis par induction pour $n \geq N+2$, si $\tilde{\rho_n} = (G, \rho_n, R)$ alors :

$$\tilde{\rho}_{n+1} = \begin{cases} (G, \rho_{n+1}, R \cup \rho_{n+1}) & \text{si } R \subsetneq G, \\ (G, \rho_{n+1}, \emptyset) & \text{si } R = G. \end{cases}$$

Alors, $\tilde{\rho}$ est un déroulement $\mathfrak{B}$ pour α et puisque $\mathrm{Inf}(\rho) = G$ et G est fini, nécessairement, il existe $q \in G$ tel que $(G, q, \emptyset) \in \mathrm{Inf}(\tilde{\rho})$. Par conséquent, $\alpha \in L(\mathfrak{B})$. □

Il est important de remarquer que dans le cas où l'automate de Muller est déterministe, alors l'automate de Büchi construit dans la preuve du théorème précédent est non déterministe. Nous verrons qu'il n'est pas possible en général de construire un automate déterministe de Büchi qui soit équivalent à un automate déterministe de Muller donné.

En combinant le théorème précédent avec la version non déterministe du Théorème I.4.2, nous obtenons :

Théorème I.4.4. *Un ω-langage $L \subseteq \Sigma^\omega$ est reconnaissable par un automate non déterministe de Muller si et seulement si $L \in \mathrm{REG}^\omega(\Sigma)$.* □

I.4.2 Automate déterministe de Muller

Nous commençons par montrer que l'ensemble des ω-langages reconnaissables par un automate déterministe de Muller est clos par complémentation.

Proposition I.4.5. *Soit $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, \mathscr{F})$ un automate déterministe de Muller. L'automate déterministe de Muller $\neg\mathfrak{A} = (Q, \Sigma, \Delta, q_I, \neg\mathscr{F})$ où $\neg\mathscr{F} = \{F \subseteq Q \mid F \notin \mathscr{F}\}$ reconnaît le ω-langage $\Sigma^\omega \setminus L(\mathfrak{A})$.*

Preuve. Remarquer que puisque $\mathfrak{A}$ et $\neg\mathfrak{A}$ sont des automates déterministes, pour tout $\alpha \in \Sigma^\omega$ il existe un unique déroulement ρ de $\mathfrak{A}$ pour α et que ce déroulement est aussi l'unique déroulement de $\neg\mathfrak{A}$ pour α. De plus, par définition de $\neg\mathscr{F}$,

$$\mathrm{Inf}(\rho) \in \neg\mathscr{F} \Leftrightarrow \mathrm{Inf}(\rho) \notin \mathscr{F},$$

et donc $\alpha \in L(\neg\mathfrak{A})$ si et seulement si $\alpha \notin L(\mathfrak{A})$. □

Les ω-langages reconnaissables par un automate déterministe de Muller sont en outre clos par intersection. Avec le résultat précédent, nous obtenons que les ω-langages reconnaissables par un automate déterministe de Muller forment une algèbre de Boole, i.e. ils sont clos par union finie, intersection finie et complémentation.

Proposition I.4.6. *Soient $\mathfrak{A}_i = (Q_i, \Sigma, \Delta_i, I_i, \mathscr{F}_i)$ avec $i = 1, 2$ deux automates déterministes de Muller. L'automate déterministe de Muller $\mathfrak{A} = (Q_1 \times Q_2, \Sigma, \Delta, (I_1, I_2), \mathscr{F})$ où*

$$\Delta\big((q_1, q_2), a\big) = \big(\Delta_1(q_1, a), \Delta_2(q_2, a)\big) \quad \textit{pour tout } (q_1, q_2) \in Q_1 \times Q_2 \textit{ et pour tout } a \in \Sigma,$$
$$\mathscr{F} = \{F_1 \times F_2 \mid F_1 \in \mathscr{F}_1 \textit{ et } F_2 \in \mathscr{F}_2\},$$

reconnaît $L(\mathfrak{A}_1) \cap L(\mathfrak{A}_2)$.

Preuve. Notons $\pi_i : Q_1 \times Q_2 \to Q_i$ pour $i = 1, 2$ les projections sur la première et la deuxième composante. Notons également $\pi_i : (Q_1 \times Q_2)^\omega \to Q_i^\omega$ les fonctions induites par ces projections sur les suites infinies d'états. Soit $\alpha \in \Sigma^\omega$ un mot infini sur Σ. Notons

$$\rho = (\rho_{1,0}, \rho_{2,0})(\rho_{1,1}, \rho_{2,1}) \cdots \in (Q_1 \times Q_2)^\omega$$

l'unique déroulement de $\mathfrak{A}$ pour α. Remarquons que l'unique déroulement de $\mathfrak{A}_1$, respectivement $\mathfrak{A}_2$, pour α est donné par $\pi_1(\rho) = \rho_{1,0}\rho_{1,1}\rho_{1,2}\cdots \in Q_1^\omega$, respectivement $\pi_2(\rho) = \rho_{2,0}\rho_{2,1}\rho_{2,2} \in Q_2^\omega$. Nous affirmons que

$$\pi_i(\mathrm{Inf}(\rho)) = \mathrm{Inf}(\pi_i(\rho)) = \text{ pour } i = 1,2. \tag{I.1}$$

En effet, pour $i = 1$, supposons que $q_1 \in \pi_1(\mathrm{Inf}(\rho))$. Il existe alors $q_2 \in Q_2$ tel que $(q_1, q_2) \in \mathrm{Inf}(\rho)$. Il s'ensuit que $q_1 \in \mathrm{Inf}(\pi_1(\rho))$. Réciproquement, supposons que $q_1 \in \mathrm{Inf}(\pi_1(\rho))$ et notons $(n_k)_{k\in\omega}$ la suite strictement croissante des naturels pour lesquels $\rho_{1,n_k} = \pi_1(\rho_{n_k}) = q_1$. La suite $r = (\rho_{2,n_k})_{k\in\omega}$ d'états de Q_2 associée vérifie $\mathrm{Inf}(r) \neq \emptyset$ car Q_2 est fini. Alors en choisissant un $q_2 \in \mathrm{Inf}(r)$ nous obtenons que $(q_1, q_2) \in \mathrm{Inf}(\rho)$. Par conséquent, $q_1 \in \pi_1(\mathrm{Inf}(\rho))$. L'égalité pour $i = 2$ se montre de façon identique.

Supposons maintenant que $\alpha \in L(\mathfrak{A})$, i.e. $\mathrm{Inf}(\rho) \in \mathscr{F}$. Par (I.1), l'unique déroulement $\pi_1(\rho)$ de $\mathfrak{A}_1$ pour α vérifie $\mathrm{Inf}(\pi_1(\rho)) = \pi_1(\mathrm{Inf}(\rho)) \in \mathscr{F}_1$ par définition de $\mathscr{F}$. Par conséquent, $\alpha \in L(\mathfrak{A}_1)$. La démonstration du fait que $\alpha \in L(\mathfrak{A}_2)$ est, *mutatis mutandis*, identique. Ainsi, $\alpha \in L(\mathfrak{A}_1) \cap L(\mathfrak{A}_1)$.

Réciproquement, supposons que $\alpha \in L(\mathfrak{A}_1) \cap L(\mathfrak{A}_2)$, i.e. $\mathrm{Inf}(\pi_1(\rho)) \in \mathscr{F}_1$ et $\mathrm{Inf}(\pi_2(\rho)) \in \mathscr{F}_2$. Par (I.1) et par la définition de $\mathscr{F}$, nous avons

$$\mathrm{Inf}(\rho) = \pi_1(\mathrm{Inf}(\rho)) \times \pi_2(\mathrm{Inf}(\rho)) = \mathrm{Inf}(\pi_1(\rho)) \times \mathrm{Inf}(\pi_2(\rho)) \in \mathscr{F}.$$

Donc $\alpha \in L(\mathfrak{A})$.

Nous avons ainsi obtenu $L(\mathfrak{A}) = L(\mathfrak{A}_1) \cap L(\mathfrak{A}_2)$. □

Nous donnons à présent une caractérisation des ω-langages qui sont reconnus par un automate déterministe de Muller en terme des ω-langages Büchi déterministes reconnaissables.

Théorème I.4.7. *Soit $L \subseteq \Sigma^\omega$. Les assertions suivantes sont équivalentes.*

1) *L est reconnu par un automate déterministe de Muller ;*
2) *L est de la forme*
$$L = \bigcup_{i=1}^{n} (U_i \cap V_i^c),$$
où les U_i et les V_i sont des ω-langages déterministe Büchi reconnaissables.
3) *L est une combinaison booléenne de ω-langages déterministe Büchi reconnaissables.*

Preuve. 1)⇒ 2). Soit $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, \mathscr{F})$ un automate déterministe de Muller. Pour chaque état $q \in Q$, nous notons $L(q)$ le ω-langage reconnu par l'automate déterministe de Büchi $(Q, \Sigma, \Delta, q_I, \{q\})$. Nous avons la formule suivante :

$$L(\mathfrak{A}) = \bigcup_{F\in\mathscr{F}} \left(\bigcap_{q\in F} L(q) \cap \Big(\bigcup_{p\in Q\setminus F} L(p) \Big)^c \right).$$

En outre, par la Proposition I.3.4, les ω-langages $\bigcap_{q\in F} L(q)$ et $\bigcup_{p\in Q\setminus F} L(p)$ sont déterministe Büchi reconnaissables.

2)⇒3). Trivial.

3)⇒ 1). Cette implication découle des deux assertions suivantes que nous avons déjà démontré. D'une part, un ω-langage reconnaissable par un automate déterministe de Büchi est reconnaissable par un automate déterministe de Muller par le Théorème I.4.2. Et d'autre part, les ω-langages reconnaissables par un automate déterministe de Muller forment une algèbre par les propositions I.4.5 et I.4.6. □

À l'aide du Corollaire I.3.6, nous obtenons le corollaire suivant.

Corollaire I.4.8. *Soit Σ un alphabet fini. Les ω-langages reconnus par un automate déterministe de Muller sont contenus dans la classe de Borel $\mathbf{\Delta}_3^0(\Sigma^\omega)$.*

Preuve. Par le Théorème I.4.7, les ω-langages reconnaissables par un automate déterministe de Muller sont exactement les combinaisons booléennes de ω-langages reconnaissables par un automate déterministe de Büchi. Par le Corollaire I.3.6, les ω-langages reconnaissables par un automate déterministe de Büchi sont contenus dans la classe de Borel $\mathbf{\Pi}_2^0(\Sigma^\omega)$. Or, de façon générale, la classe $\mathbf{\Pi}_2^0$ est contenue dans la classe $\mathbf{\Delta}_3^0$ qui est une algèbre. Il en découle que les ω-langages reconnaissables par un automate déterministe de Muller sont contenus dans $\mathbf{\Delta}_3^0(\Sigma^\omega)$. □

I.5 Automates de Rabin et de Strett

Définition I.5.1. Un automate (non) déterministe $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, \Omega)$ où la composante d'acceptation est un ensemble fini de couples de sous-ensembles d'états $\Omega = \{(E_1, F_1), \ldots, (E_n, F_n)\}$ avec $F_1, E_1, \ldots, E_n, F_n \subseteq Q$ est appelé :

1) un **automate (non) déterministe de Rabin** lorsqu'il **accepte** un mot infini $\alpha \in \Sigma^\omega$ s'il existe $\rho \in \mathscr{D}(\mathfrak{A}, \alpha)$ vérifiant la **condition de Rabin**

$$\exists (E, F) \in \Omega \quad (\mathrm{Inf}(\rho) \cap E = \emptyset) \wedge (\mathrm{Inf}(\rho) \cap F \neq \emptyset). \tag{R}$$

2) un **automate (non) déterministe de Strett** lorsqu'il **accepte** un mot infini $\alpha \in \Sigma^\omega$ s'il existe $\rho \in \mathscr{D}(\mathfrak{A}, \alpha)$ vérifiant la **condition de Strett**

$$\forall (E, F) \in \Omega \quad (\mathrm{Inf}(\rho) \cap E \neq \emptyset) \vee (\mathrm{Inf}(\rho) \cap F = \emptyset), \tag{S}$$

ou de façon équivalente,

$$\forall (E, F) \in \Omega \quad (\mathrm{Inf}(\rho) \cap F \neq \emptyset) \Rightarrow (\mathrm{Inf}(\rho) \cap E \neq \emptyset).$$

Remarquer que la condition de Strett (S) est équivalente à la négation de la condition de Rabin (R).

Observons tout de suite qu'il est facile de donner un automate de Muller qui soit équivalent à un automate de Rabin ou de Strett. En outre, cette transformation fournit un automate déterministe de Rabin ou de Strett dans le cas où l'automate de Muller est déterministe.

Théorème I.5.2. *Soit $L \subseteq \Sigma^\omega$ un ω-langage reconnu par un automate (non) déterministe de Rabin, respectivement par un automate (non) déterministe de Strett. Le ω-langage L est reconnu par l'automate (non) déterministe de Muller $\mathfrak{M} = (Q, \Sigma, \Delta, q_I, \mathscr{F})$ où*

$$\mathscr{F} = \{G \subseteq Q \mid \exists (E, F) \in \Omega \quad (G \cap E = \emptyset) \wedge (G \cap F \neq \emptyset)\},$$

respectivement

$$\mathscr{F} = \{G \subseteq Q \mid \forall (E, F) \in \Omega \quad (G \cap E \neq \emptyset) \vee (G \cap F = \emptyset)\}.$$ □

De plus, étant donné un automate de Büchi, il est facile de donner un automate de Rabin ou de Strett qui reconnaisse le même langage.

Théorème I.5.3. *Soit $L \subseteq \Sigma^\omega$ un ω-langage reconnu par un automate (non) déterministe de Büchi $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F)$. L'automate (non) déterministe de Rabin $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, \{(\emptyset, F)\})$ et l'automate (non) déterministe de Strett $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, \{(F, Q)\})$ reconnaissent L.* □

En mettant en commun les deux derniers théorèmes pour le cas non déterministe et en se référant au Théorème I.4.4, nous obtenons comme résumé que toutes les notions d'automate non déterministes que nous avons définies reconnaissent le même ensemble de ω-langages à savoir $\mathrm{REG}^\omega(\Sigma)$.

Théorème I.5.4. *Soit $L \subseteq \Sigma^\omega$ un ω-langage. Les assertions suivantes sont équivalentes.*
a) L est reconnaissable par un automate non déterministe de Büchi.
b) L est reconnaissable par un automate non déterministe de Muller.
c) L est reconnaissable par un automate non déterministe de Rabin.
d) L est reconnaissable par un automate non déterministe de Strett.
e) $L \in \mathrm{REG}^\omega(\Sigma)$. □

Remarquer cependant que, comme nous l'avons déjà mentionné, la construction de la preuve du Théorème I.4.3 mène à un automate non déterministe de Büchi, même dans le cas d'un automate déterministe de Muller. Il n'est par conséquent pas possible de déduire des théorèmes I.5.2 et I.5.3 que les automates déterministes de Rabin, de Strett et de Muller reconnaissent le même ensemble de ω-langages. Nous montrons cela à présent.

Théorème I.5.5. *Soit $L \subseteq \Sigma^\omega$ un ω-langage reconnu par un automate déterministe de Muller $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, \mathscr{F})$. Il existe un automate déterministe de Rabin qui reconnaît L.*

Preuve. Supposons sans perte de généralité que $Q = \{1, \ldots, k\}$ pour un entier k et que $q_I = 1$. Nous définissons un automate déterministe de Rabin $\mathfrak{R} = (Q', \Sigma, \Delta', q'_I, \Omega)$ dont les composantes sont les suivantes. L'ensemble de ses états est :

$$Q' = \Big\{(q_1, \ldots, q_k, i) \in Q^k \times \{1, \ldots, k\} \mid \forall i, j \in \{1, \ldots, k\}.(i \neq j \Rightarrow \bar{q}_i \neq \bar{q}_j\Big)\}.$$

Son état initial est $q'_I = (1, \ldots, k, k)$. Intuitivement, lors d'un déroulement de l'automate $\mathfrak{M}$, un état $\bar{q} = (q_1, \ldots, q_k, i) \in Q'$ représente les états de Q dans l'ordre de leur dernière occurence dans le déroulement correspondant de $\mathfrak{A}$. L'état q_1 étant l'état actuel, q_2 le dernier état avoir été visité avant q_1, q_3 le dernier état avoir été visité avant q_1 et q_2, et ainsi de suite. Pour tout $q \in Q$, notons $r_q : Q' \to \{1, \ldots, k\}$ la fonction qui indique le rang de l'état q, i.e. $r_q\big((q_1, \ldots, q_k, i)\big) = l$ si $q_l = q$.

La fonction de transition de $\mathfrak{M}$ réarrange l'ordre selon la fonction de transition Δ de $\mathfrak{A}$. Elle est donnée pour tout $\bar{q} = (q_1, \ldots, q_k, i) \in Q'$ et tout $a \in \Sigma$ par

$$\Delta'\big((q_1, \ldots, q_k, i), a\big) = (\Delta(q_1, a), q_1, \ldots, q_{l-1}, q_{l+1}, \ldots, q_k, l) \in Q', \quad \text{où } l = r_{\Delta(q_1, a)}(\bar{q}).$$

En notant $\pi : Q' \to \{1, \ldots, k\}$ la projection sur la dernière composante, remarquer que nous avons pour tout $\bar{q} \in Q'$ et tout $a \in \Sigma$

$$\pi(\Delta'(\bar{q}, a)) = r_{\Delta(\bar{q}_1, a)}(\bar{q}). \qquad (\star)$$

En d'autres termes, si $\Delta'(\bar{q}, a) = \bar{q}'$, alors $\bar{q}'_1 = \Delta(\bar{q}_1, a)$ tandis que $\pi(\bar{q}') = r_{\bar{q}'_1}(\bar{q})$.

La composante d'acceptation de $\mathfrak{M}$ est $\Omega = \{(E_1, F_1), \dots, (E_k, F_k)\}$ avec pour $i = 1, \dots, k$:

$$\begin{aligned} E_i &= \{\bar{q} \in Q' \mid \pi(\bar{q}) > i\} \\ F_i &= \{\bar{q} \in Q' \mid \pi(\bar{q}) > i\} \cup \{\bar{q} \in Q' \mid \pi(\bar{q}) = i \wedge \{\bar{q}_1, \dots, \bar{q}_i\} \in \mathscr{F}\}. \end{aligned}$$

Nous montrons que $\mathfrak{M}$ reconnaît L. Soit $\alpha \in \Sigma^\omega$. Notons $\rho \in Q^\omega$ l'unique déroulement de $\mathfrak{A}$ pour α et $\bar{\rho} \in (Q')^\omega$ l'unique déroulement de $\mathfrak{M}$ pour α. Observer que pour tout $n \in \omega$

$$\rho_n = (\bar{\rho}_n)_1 \text{ et } (\bar{\rho}_{n+1})_1 = \Delta(\rho_n, \alpha_n).$$

Nous montrons tout d'abord l'assertion suivante. Si $\rho \in Q^\omega$ est l'unique développement de $\mathfrak{A}$ pour un mot infini $\alpha \in \Sigma^\omega$, alors l'unique déroulement $\bar{\rho} \in (Q')^\omega$ de $\mathfrak{R}$ pour α vérifie

$$\exists N \in \omega. \forall n \geq N. \pi(\bar{\rho}_n) \leq \mathrm{Card}(\mathrm{Inf}(\rho)).$$

En effet, il existe $M \in \omega$ tel que pour tout $n \geq M$, nous avons $\rho_n \in \mathrm{Inf}(\rho)$. Pour ce naturel M, il existe $N \geq M$ tel que $\mathrm{Occ}(\rho[M, N-1]) = \mathrm{Inf})(\rho)$. Intuitivement,en notant $\mathrm{Card}(\mathrm{Inf}(\rho)) = i$, pour tout $n \geq N-1$, les i première composantes de tout état $\bar{\rho}_n$ sont les j états de $\mathrm{Inf}(\rho)$. Formellement, pour tout $n \geq N$ nous avons $r_p(\bar{\rho}_n) \leq i$ pour tout $p \in \mathrm{Inf}(\rho)$. Il découle alors de ($\star$) que $\pi(\tilde{\rho}_n) \leq i$ pour tout $n \geq N$. L'assertion est démontrée.

Supposons maintenant que $\alpha \in L$. Alors, il existe $J \in \mathscr{F}$ tel que $\mathrm{Inf}(\rho) = J$ et notons $j = \mathrm{Card}(J) \in \{1, \cdots, k\}$. Par l'assertion, il existe $N \in \omega$ tel que $\pi(\bar{\rho}_n) \leq j$ pour tout $n \geq N$. Par conséquent, $\mathrm{Inf}(\bar{\rho}) \cap E_j = \emptyset$. Montrons que de plus $\mathrm{Inf}(\bar{\rho}) \cap F_j \neq \emptyset$. Pour voir ceci, notons p_0 l'état de J tel que $r_{p_0}(\bar{\rho}_N) = j$. Puisque $\mathrm{Inf}(\rho) = J$, il existe un $n_0 > N$ minimal tel que $\rho_{n_0} = p_0$ et nous avons alors $\pi(\bar{\rho}_{n_0}) = j$. Notons alors p_1 l'état de J tel que $r_{p_1}(\bar{\rho}_{n_0}) = j$, à nouveau il existe $n_1 > n_0$ minimal tel que $\rho_{n_1} = p_1$ et nous avons alors $\pi(\bar{\rho}_{n_1}) = j$. Par induction, nous construisons alors une suite strictement croissante de naturels $(n_k)_{k \in \omega}$ qui vérifie $\bar{\rho_{n_k}} \in F_j$ pour tout $k \in \omega$. Il s'ensuit que $\mathrm{Inf}(\bar{\rho}) \cap F_j \neq \emptyset$. Par conséquent, $\alpha \in L(\mathbb{R})$.

Supposons réciproquement que $\alpha \in L(\mathfrak{M})$. Il existe alors $j \in \{1, \dots, k\}$ tel que $\mathrm{Inf}(\bar{\rho}) \cap E_j = \emptyset$ et $\mathrm{Inf}(\bar{\rho}) \cap F_j \neq \emptyset$. Puisque Q est fini, $\mathscr{F} \subseteq \mathscr{P}(Q)$ est fini et il existe $J \in \mathscr{F}$ tel que

$$\mathrm{Inf}(\bar{\rho}) \cap \{\bar{q} \in Q' \mid \pi(\bar{q}) = j \wedge \{\bar{q}_j, \dots, \bar{q}_j\} = J\}. \tag{I.2}$$

De plus, il existe $N \in \omega$ tel que pour tout $n \geq N$ nous avons $\pi(\bar{\rho_n}) \leq j$. Ainsi, en prenant $M > N$ tel que

$$\bar{\rho}_M \in \{\bar{q} \in Q' \mid \pi(\bar{q}) = j \wedge \{\bar{q}_1, \dots, \bar{q}_j\} = J\},$$

puisque par ($\star$), nous avons

$$r_{\rho_{M+1}}(\bar{\rho}_M) = \pi(\bar{\rho}_{M+1}) \leq j,$$

nous obtenons

$$\rho_{M+1} = (\bar{\rho}_{M+1})_1 = \Delta(\rho_M, \alpha_M) \in J.$$

Par induction, nous obtenons pour tout $n \geq M$

$$\rho_n = (\bar{\rho}_n)_1 \in J.$$

Nous avons donc obtenu $\mathrm{Inf}(\rho) \subseteq J$. Pour voir que $\mathrm{Inf}(\rho) = J$, supposons par l'absurde que $\mathrm{Card}(\mathrm{Inf}(\rho)) = k < j = \mathrm{Card}(J)$. Par l'assertion, cela implique qu'il existe un naturel N tel que $\pi(\bar{\rho}_n) \leq k < j$ pour tout $n \geq N$. Cela contredit (I.2). Par conséquent, $\mathrm{Inf}(\rho) = J$ et donc $\alpha \in L$. $\square$

Le dernier théorème nous donne un automate déterministe de Rabin équivalent à un automate déterministe Muller donné. De plus, le Théorème I.5.2 nous fournit notamment un automate déterministe de Muller équivalent à un automate déterministe de Strett donné. Ainsi, nous avons également un automate déterministe de Strett équivalent à un automate déterministe de Rabin donné. En résumé :

Théorème I.5.6. *Soit $L \subseteq \Sigma^\omega$ ω-langage. Les assertions suivantes sont équivalentes.*

a) L est reconnaissable par un automate déterministe de Muller.

b) L est reconnaissable par un automate déterministe de Rabin.

c) L est reconnaissable par un automate déterministe de Strett.

CHAPITRE II

Le Théorème de McNaughton

Le premier chapitre aboutit à trois ensembles a priori distincts de langages sur les mots infinis reconnaissables par un automate. Le premier formé par ceux reconnaissables par un automate déterministe de Büchi qui sont strictement inclus dans le deuxième formé par ceux reconnaissables par un automate non déterministe de Büchi, de Muller, de Rabin ou de Strett. Le troisième est la clôture booléenne du premier, il est formé des langages qui sont reconnaissables par un automate déterministe de Muller, de Rabin ou de Strett. Ce chapitre est dévolu à la démonstration du Théorème de McNaughton qui établit que le premier et le deuxième sont égaux.

Nous fixons pour ce chapitre un alphabet fini Σ contenant au moins deux symboles distincts. Nous sous-entendons, sauf mention contraire, que les automates que nous considérons ont tous pour alphabet Σ.

II.1 Le Théorème de McNaughton

Au chapitre précédent, nous avons montré que les automates non déterministes de Büchi, de Muller, de Rabin et de Strett reconnaissent tous le même ensemble de ω-langages que nous notons $\mathrm{REG}^{\omega}(\Sigma)$. Nous avons également vu que les automates déterministes de Büchi reconnaissent un ensemble $\mathrm{REG}^{\omega}_{\mathrm{Dét}}(\Sigma)$ de ω-langages qui est strictement inclus dans $\mathrm{REG}^{\omega}(\Sigma)$. Nous avons aussi montré que les automates déterministes de Muller, de Rabin et de Strett reconnaissent tous le même ensemble de ω-langages à savoir la clôture booléenne de $\mathrm{REG}^{\omega}_{\mathrm{Dét}}(\Sigma)$.

Le but de ce chapitre est de démontrer le Théorème de McNaughton qui établit que les automates non déterministes de Büchi ont la même expressivité que les automates déterministes de Muller, de Rabin et de Strett.

Théorème II.1.1 (Théorème de McNaughton)**.** *Soit $L \subseteq \Sigma^{\omega}$ un ω-langage reconnu par un automate non déterministe de Büchi. Il existe un automate déterministe de Rabin qui reconnaît L.*

Avec ce résultat, nous obtenons que toutes les notions d'automates du chapitre précédent, sauf celle d'automate déterministe de Büchi, reconnaissent le même ensemble de ω-

langages que nous notons $\mathrm{REG}^{\omega}(\Sigma)$. Par souci de simplicité, un ω-langage $L \subseteq \Sigma^{\omega}$ sera appelé **reconnaissable** s'il appartient à $\mathrm{REG}^{\omega}(\Sigma)$. Nous avons donc le corollaire.

Corollaire II.1.2. *Soit $L \subseteq \Sigma^{\omega}$ un ω-langage. Les assertions suivantes sont équivalentes.*

1) *L est reconnaissable, i.e. L est reconnaissable par un automate non déterministe de Büchi.*
2) *L est reconnaissable par un automate non déterministe de Muller.*
3) *L est reconnaissable par un automate déterministe de Muller.*
4) *L est reconnaissable par un automate non déterministe de Rabin ou de Strett.*
5) *L est reconnaissable par un automate déterministe de Rabin ou de Strett.*

En vertu du Théorème I.4.7, nous savons que l'ensemble des ω-langages reconnaissables par un automate déterministe de Muller est la clôture booléenne de l'ensemble $\mathrm{REG}^{\omega}_{\text{Dét}}(\Sigma)$ formé par les ω-langages reconnaissables par un automate déterministe de Büchi. Or par le Collaire II.1.2, les langages reconnaissables sont exactement ceux qui sont reconnaissables par un automate déterministe de Muller. Ainsi, d'un point de vue topologique, il découle alors du Corollaire I.4.8, que les ω-langages reconnaissables sont contenus dans la classe de Borel $\boldsymbol{\Delta}^0_3(\Sigma^{\omega})$ et que le niveau des ces ensembles dans la hiérarchie de Borel est borné.

Nous consacrons la fin de ce chapitre à la démonstration du Théorème de McNaughton. Nous utilisons la construction dite de Safra pour construire un automate déterministe de Muller équivalent à un automate non déterministe de Büchi donné. Les états de l'automate déterministe de Muller serons des arbres. C'est pourquoi nous exposons quelques notions de base sur les arbres dans la section suivante.

II.2 Les arbres

Définition II.2.1. Un arbre est un triplet $A = (N, r, p)$ où

i) N est un ensemble dont les éléments sont appelés les **nœuds** ;
ii) r est un élément de N appelé la **racine** ;
iii) $p : N \backslash \{r\} \to N$ est une fonction appelée la **fonction parent** ;

tel que la propriété suivante soit vérifiée :

$$\forall n \in N \quad \exists k \in \omega \quad p^k(n) = r.$$

Un arbre $A = (N, r, p)$ est dit **fini** si N est fini, sinon il est dit **infini**.

Pour un arbre $A = (N, r, p)$, il est possible de définir la **fonction enfant** qui associe à chaque nœud l'ensemble des nœuds dont il est le parent. Formellement, la fonction enfant de A est la fonction $c_A : N \to \mathscr{P}(N)$ définie par $c_A(n) = \{m \in N \mid p(m) = n\}$ pour tout $n \in N$.

Un nœud $n \in N$ est un **ancêtre** d'un nœud m s'il existe $k \in \omega$ tel que $p^k(m) = n$, le nœud m est alors appelé un **descendant** de n.

Soit n est un nœud d'un arbre $A = (N, r, p)$, dénotons par N_n l'ensemble des descendants de n. Le **sous-arbre** de A de racine n est l'arbre $A_n = (N_n, n, p_n)$ où $p_n : N_n \backslash \{n\} \to N_n$ est la restriction de p.

L'**arité** d'un nœud n d'un arbre $A = (N, r, p)$ avec fonction enfant c_A est la cardinalité de $c_A(n)$. L'arité d'un arbre est le supremum des arités de chacun de ces nœuds.

Un **chemin** de longueur $n \in \omega$ dans un arbre $A = (N, r, p)$ est une suite $s \in N^n$ telle que pour tout $i \in n-1$, nous avons $p(s_{i+1}) = s_i$. Un **chemin infini** dans un arbre $A = (N, r, p)$ est une suite infinie $s \in N^\omega$ telle que pour tout $i \in \omega$, nous avons $p(s_{i+1}) = s_i$.

Nous aurons besoin du lemme suivant.

Lemme II.2.2 (Lemme de König)**.** *Soit $A = (N, r, p)$ un arbre infini d'arité finie. Il existe un chemin infini dans A.*

Preuve. Nous montrons par induction qu'il existe une suite infinie de nœuds $s = s_0 s_1 \cdots \in N^\omega$ tel que $s_0 = r$ et pour tout $k \in \omega$ le sous-arbre de A de racine n_k est infini. Cette suite est un chemin infini dans A. Le fait que A est infini nous assure la base de la récurrence. Supposons maintenant pour $k \in \omega$ que s_k est nœud tel que le sous-arbre de racine s_k est infini. Puisque $\mathscr{A}$ est d'arité finie, s_k ne possède qu'un nombre fini d'enfants. Par conséquent, il existe un enfant m de n_k tel que le sous-arbre de racine m est infini. Nous choisissons un tel enfant m pour s_{k+1}. □

Nous aurons besoin dans la suite d'arbre dont les enfants de chaque nœud sont ordonnés. Un tel arbre est appelé un **arbre orienté** et peut être défini par un triplet $A = (N, r, c)$ où N est l'ensemble de nœud, $r \in N$ est la racine, et $c : N \to N^*$ est une fonction où pour chaque $n \in N$ la suite de nœuds $c(n) = c_0 \cdots c_k$ représente, dans l'ordre, les enfants du nœud n de l'aîné au cadet. Nous étendons les ordres sur les enfants de chaque nœud à un ordre partiel sur l'ensemble des nœuds d'un arbre orienté. Étant donné deux nœuds $v, w \in N$ qui ne sont pas ancêtres l'un de l'autre, notons a leur plus proche ancêtre commun et a_v, respectivement a_w, l'enfant de a qui est l'ancêtre de v, respectivement w. Nous disons v est **à la gauche** de w si a_v est l'aîné de a_w.

Nous utiliserons des arbres orientés dont chaque nœud possède une étiquette. Un tel arbre est appelé un **arbre orienté étiqueté**, il est décrit par un quadruplet $A = (N, r, f, e)$ où (N, r, f) est un arbre orienté et $e : N \to E$ est une fonction qui associe à chaque nœud un élément d'un ensemble d'**étiquettes** E.

II.3 La construction de Safra

Nous nous fixons un automate non déterministe de Büchi $\mathfrak{A} = (Q, \Sigma, \Delta, 1, F)$ dont les états sont $Q = \{1, \ldots, n\}$. Nous supposons de plus sans perte de généralité que l'état initial de $\mathfrak{A}$ n'est pas un état acceptant, i.e. $1 \notin F$. Nous allons petit à petit définir un automate déterministe de Rabin $\mathfrak{R}$ qui reconnaît le même langage que $\mathfrak{A}$. L'ensemble de ces états est formé par les arbres de Safra sur Q que nous définissons à présent.

Définition II.3.1. Un **arbre de Safra** sur $Q = \{1, \ldots, n\}$ est quadruplet $A = (N, f, e, M)$ où

i) $N \subseteq V = \{1, \ldots, 2n\}$ est l'ensemble des nœuds de A ;

ii) $f : N \to N^*$ est la fonction qui associe à chaque nœud de A la suite ordonnée de ces enfants de l'aîné au cadet ;

iii) $e : N \to \mathscr{P}(Q) \backslash \{\emptyset\}$ est la fonction qui étiquette chaque nœud avec un sous-ensemble non vide de Q ;

iv) $M \subseteq N$ est l'ensemble des nœuds **marqués** de A ;

tel que les propriétés suivantes sont vérifiées :

a) $(N, 1, f, e)$ est un arbre orienté étiqueté et nous notons $c : N \to \mathscr{P}(N)$ sa fonction enfant ;

b) Un nœud marqué n'a pas d'enfants, i.e. $c(m) = \emptyset$ pour tout $m \in M$.

c) L'union des étiquettes des enfants d'un nœud est un sous-ensemble strict de l'étiquette de ce nœud, i.e. pour tout $v \in N$, nous avons :

$$\bigcup_{n \in c(v)} e(n) \subsetneq e(v);$$

d) Les étiquettes de deux nœuds qui ne sont pas ancêtres l'un de l'autres sont disjointes, i.e. pour tous $v, w \in N$, si v n'est pas un ancêtre de w et w n'est pas un ancêtre de v, alors

$$e(v) \cap e(w) = \emptyset.$$

L'ensembles des arbres de Safra sur $Q = \{1, \ldots, n\}$ est noté $\mathscr{S}_n$.

L'ensemble $\mathscr{S}_n$ des arbres de Safra sur $\{1, \ldots, n\}$ est fini, nous avons même le résultat suivant :

Lemme II.3.2. *Soit $A = (N, f, e, M) \in \mathscr{S}_n$ un arbre de Safra sur $\{1, \ldots, n\}$. Le nombre de nœuds de A est inférieur ou égal à n.*

Preuve. Notons $c : N \to \mathscr{P}(N)$ la fonction enfant de A et définissons pour tout nœud $v \in N$:

$$r(v) = e(v) \setminus \bigcup_{n \in c(v)} e(n) \subseteq e(v) \subseteq Q = \{1, \ldots, n\}.$$

Par la propriété c) d'un arbre de Safra, $r(v) \neq \emptyset$ pour tout $v \in N$. De plus, pour deux nœuds distincts $n_1, n_2 \in N$, nous avons $r(n_1) \cap r(n_2) = \emptyset$. En effet, si les nœuds n_1, n_2 ne sont pas ancêtres l'un de l'autre, alors par la propriété d) d'un arbre de Safra, nous avons que $r(n_1)$ et $r(n_2)$ sont disjoints. Tandis que si le nœud n_1 est un ancêtre de n_2, ou inversement, cela découle de la propriété c) d'un arbre de Safra et de la définition de r.

Finalement, nous obtenons :

$$\mathrm{Card}(N) = \sum_{v \in N} 1 \leq \sum_{v \in N} \mathrm{Card}(r(v)) \leq \mathrm{Card}(Q) = n. \qquad \square$$

L'ensemble des états de notre automate déterministe de Rabin est $\mathscr{S}_n$. Nous nous attelons à présent à décrire se fonction de transition Δ'. Pour tout $A = (N, f, e, M) \in \mathscr{S}_n$ et tout $a \in \Sigma$, l'arbre de Safra $\Delta'(A, a) \in \mathscr{S}_n$ est obtenu de A en effectuant les étapes suivantes.

Étape 1) Nous formons l'arbre $A_1 = (N, f, e_1, \emptyset)$ en fixant l'ensemble des nœuds marqués à l'ensemble vide et en appliquant la relation de transition de $\mathfrak{A}$ à chaque étiquette en définissant pour tout $v \in N$,

$$e_1(v) = \{q \in Q \mid \exists p \in e(v).(p, a, q) \in \Delta\} = \Delta(e(v), a).$$

Étape 2) Nous formons l'arbre $A_2 = (N_2, f_2, e_2, M_2)$ à patir de A_1 en ajoutant à chaque nœud dont l'étiquette contient un état acceptant un nouvel enfant (qui devient le cadet) avec pour étiquette les états acceptants de l'étiquette de son parent et nous marquons ce nœud. Formellement, nous notons

$$\tilde{N} = \{v \in N \mid e_1(v) \cap F \neq \emptyset\},$$

et nous fixons une injection $j : \tilde{N} \to V \backslash N$. L'existence d'une telle injection nous est assurée par le Lemme II.3.2 selon lequel $\mathrm{Card}(\tilde{N}) \leq \mathrm{Card}(N) \leq n$. Nous posons alors $N_2 = N \cup \mathrm{Im}(j)$ et définissons

$$f_2(v) = \begin{cases} f(v) & \text{si } v \in N \backslash \tilde{N}; \\ f(v)j(v) & \text{si } v \in \tilde{N}; \\ \epsilon & \text{si } v \in \mathrm{Im}(j); \end{cases} \qquad e_2(v) = \begin{cases} e_1(v) & \text{si } v \in N \\ e_1(j^{-1}(v)) \cap F & \text{si } v \in \mathrm{Im}(j); \end{cases}$$

Nous marquons finalement les nœuds ajoutés en posant $M_2 = \mathrm{Im}(j)$.

Étape 3) Nous formons l'arbre $A_3 = (N_2, f_2, e_3, M_2)$ à partir de A_2 en supprimant de l'étiquette de chaque des nœud v les états qui apparaissent dans les étiquettes des nœuds qui se trouvent à la gauche de celui-ci. Ainsi, l'arbre A_3 vérifie la propriété d) d'un arbre de Safra. Formellement, nous définissons e_3 pour tout $v \in N_2$ par

$$e_3(v) = e_2(v) \backslash \bigcup_{\substack{w \in N_2 \text{ tel que} \\ w \text{ est à la gauche de } v}} e_2(w).$$

Étape 4) Nous formons l'arbre $A_4 = (N_4, f_4, e_4, M_4)$ à partir de A_3 en supprimant les nœuds dont l'étiquette est vide et en adaptant en conséquence les fonctions f_2 et e_3 ainsi que l'ensemble des nœuds marqués. Ainsi, l'arbre A_4 ne contient plus de nœuds avec une étiquette vide. Formellement, nous fixons

$$N_4 = \{v \in N_2 \mid e_3(v) \neq \emptyset\},$$

nous posons $M_4 = N_3 \cap N_4$ et nous définissons e_4 comme la restriction de e_3 à N_4. Pour chaque nœud $v \in N_4$, la suite finie $f_4(v)$ est obtenue de $f_2(v)$ en effaçant les éléments de $N_2 \backslash N_4$.

Étape 5) L'arbre $A_5 = (N_5, f_5, e_5, M_5)$ est obtenu de A_4 en ajoutant à l'ensemble des nœuds marqués chaque nœud $v \in N_4$ dont l'étiquette égale l'union des étiquettes de ces enfants, i.e. tel que

$$e_4(v) = \bigcup_{w \in N_4 \text{ enfant de } v} e_4(w),$$

et en supprimant tous les descendants de ces nœuds. Les fonctions f_4 et e_4 sont adaptées en conséquence pour obtenir f_5 et e_5. L'arbre ainsi obtenu vérifie la propriété d) d'un arbre de Safra.

Pour un symbole $a \in \Sigma$, en appliquant ces 5 étapes à un état $A \in \mathscr{S}_n$, nous obtenons un arbre de Safra $A_5 = \Delta'(A, a) \in \mathscr{S}_n$.

Nous fixons l'état initial de notre automate déterministe de Rabin comme étant l'abre de Safra réduit à un unique nœud non marqué avec comme étiquette le singleton $\{1\}$ où 1 est l'état initial de $\mathfrak{A}$. Formellement, nous définissons l'état initial A_I de $\mathfrak{R}$ comme l'arbre de Safra

$$A_I = (\{1\}, \{(1, \epsilon)\}, \{(1, \{1\})\}, \emptyset).$$

Nous définissons finalement la composante d'acceptation de Rabin Ω comme

$$\Omega = \{(E_v, F_v) \mid v \in V\},$$

avec pour tout $v \in V = \{1, \ldots, 2n\}$,

$$\begin{aligned} E_v &= \{A \in \mathscr{S}_n \mid v \text{ n'est pas un nœud de } A\}, \\ F_v &= \{A \in \mathscr{S}_n \mid v \text{ est un nœud marqué de } A\}. \end{aligned}$$

Observons directement que par le Théorème I.5.2, cette composante d'acceptation est équivalente à la composante de Muller définie par :

$$F \in \mathscr{F} \Leftrightarrow \exists v \in V \begin{cases} \forall A \in F \; v \text{ est un nœud de } A, \text{ et} \\ \exists A \in F \text{ tel que } v \text{ est marqué dans } A. \end{cases}$$

Nous allons étudier plus précisément le comportement de l'automate déterministe de Rabin $\mathfrak{R} = (\mathscr{S}_n, \Sigma, \Delta', A_I, \Omega)$ que nous avons défini. Afin de simplifier la suite de l'exposition, nous faisons la définition suivante. Soit $\mathfrak{B} = (Q_{\mathfrak{B}}, \Sigma, \Delta_{\mathfrak{B}}, q_I, Acc)$ un automate non déterministe. Un **chemin** dans $\mathfrak{B}$ pour un mot fini $u = u_0 u_1 \cdots u_{m-1} \in \Sigma^*$ est une suite $q_0 q_1 \cdots q_m \in Q^*_{\mathfrak{B}}$ d'états de $\mathfrak{B}$ telle que pour tout $i = 0, \ldots, m-1$, $(q_i, u_i, q_{i+1}) \in \Delta_{\mathfrak{B}}$. Si $\mathfrak{B}$ est déterministe, pour tout état $q_0 \in Q_{\mathfrak{B}}$ et tout mot fini $u = u_0 u_1 \cdots u_{m-1} \in \Sigma^*$, il existe un unique chemin $q_0 q_1 \cdots q_m$ dans $\mathfrak{B}$ pour u commençant en q_0.

Lemme II.3.3. *Soient $v \in V = \{1, \ldots, 2n\}$, $u = u_0 u_1 \cdots u_{m-1} \in \Sigma^*$ un mot fini, et $A_0 A_1 \cdots A_m \in \mathscr{S}_n^*$ un chemin dans $\mathfrak{R}$ pour u tel que, en notant $A_i = (N_i, f_i, e_i, M_i)$ pour tout $i = 1, \ldots, m$, nous ayons les propriétés suivantes*

i) pour tout $i = 1, \ldots, m$, v est un nœud de A_i, i.e. $v \in N_i$;

ii) le nœud v est marqué seulement dans A_0 et dans A_m, i.e. $v \in M_i$ si et seulement si $i = 0$ ou $i = m$.

Pour tout $p \in e_m(v)$, il existe un chemin $q = q_0 q_1 \cdots q_m$ dans $\mathfrak{A}$ pour u vérifiant

1) pour tout $i = 1, \ldots, m$, $q_i \in e_i(v)$;

2) le chemin q se termine en p, i.e. $q_m = p$;

3) le chemin q visite au moins un état acceptant après son état de départ q_0, i.e.

$$\mathrm{Occ}(q_1 \cdots q_m) \cap F \neq \emptyset.$$

Preuve. Commençons par remarquer que de façon générale, pour $i = 1, \ldots, m_1$,

$$e_{i+1} \subseteq \Delta(e_i(v), u_i).$$

En effet, lors de l'obtention de l'arbre $A_{i+1} = \Delta'(A_i, u_i)$ par les cinq étapes, à la première étape nous avons provisoirement $e_{i+1} = \Delta(e_i(v), u_i)$, à l'étape 3, certains éléments de $e_{i+1}(v)$ sont possiblement supprimés, et remarquons que v n'est pas supprimé à l'étape 4 ou à l'étape 5 puisque $v \in N_{i+1}$.

Notons pour chaque $i = 0, \ldots, m-1$, c_i la fonction enfant de A_i. Nous montrons à présent l'assertion suivante par induction sur $i = 0, \ldots, m-1$.

Assertion. Si $p \in \bigcup_{w \in c_i(v)} e_i(w)$, alors il existe un chemin $q_0 q_1 \cdots q_i \in Q^*$ dans $\mathfrak{A}$ pour $u_0 u_1 \cdots u_{i-1}$ tel que

1) pour tout $j = 0, \ldots, i-1$, $q_j \in e_j(v)$;
2) le chemin $q_0 q_1 \cdots q_i$ se termine en p, i.e. $q_i = p$;
3) le chemin $q_0 q_1 \cdots q_i$ visite F après son état de départ, i.e. $\mathrm{Occ}(q_1 \cdots q_i) \cap F \neq \emptyset$.

La base de l'induction est vérifiée car v est marqué dans A_0 et n'a par conséquent pas d'enfants. Supposons donc l'assertion vérifiée pour $0 \leq i \leq m-2$ et montrons que cela implique qu'elle est vérifiée pour $i+1$. Soit $p \in \bigcup_{w \in c_{i+1}(v)} \subsetneq e_{i+1}(v)$. Il existe par notre première remarque $q_i \in e_i(v)$ tel que $(q_i, u_i, p) \in \Delta)$. Nous distinguons alors deux cas.

Cas 1. L'état q_i appartient à l'étiquette d'un enfant de v dans A_i, i.e. $q_i \in \bigcup_{w \in c_i(v)} e_i(w)$. Nous pouvons alors conclure par l'hypothèse d'induction.

Cas 2. L'état q_i n'appartient pas à l'étiquette d'un enfant de v dans A_i. Ce cas peut se produire notamment si v n'a pas d'enfants dans A_i. Dans ce cas, p appartient à l'étiquette d'un nœud créé à l'étape 2 et donc $p \in e_{i+1}(v) \cap F$. Il existe un chemin $q_0 q_1 \ldots q_{i+1}$ dans $\mathfrak{A}$ pour $u_0 u_1 \cdots u_i$ tel que $q_{i+1} = p \in F$. Ce chemin remplit donc les conditions de l'assertion.

Remarquons maintenant que puisque $v \in M_m$ et que $v \in N_{m-1}$, il reçoit sa marque à l'étape 5. Ceci signifie qu'après avoir appliqué les étapes 1 à 4 à A_{m-1}, l'union des étiquettes des enfants de v égalait l'étiquette de v. Il s'ensuit que chaque état de $e_m(v)$ appartenait à l'étiquette d'un enfant de v après l'application des étapes 1 à 3 à A_{m-1}. Ainsi, si $p \in e_m(v)$, deux cas se présentent à nous.

Cas 1. L'état p appartient à l'étiquette d'un enfant de v après l'application de l'étape 1 à $A - m - 1$. Par conséquent il existe $q_{m-1} \in \bigcup_{w \in c_{m-1}(v)} e_{m-1}(w)$ tel que $(q_{m-1}, u_{m-1}, p) \in \Delta$. Par l'assertion, il existe un chemin de $q_0 q_1 \cdots q_{m-1}$ dans $\mathfrak{A}$ pour $u_0 u_1 \cdots u_{m-1}$ qui visite un état acceptant après son état de départ. Nous le prolongeons en un chemin $q_0 q_1 \cdots q_{m-1} p$ dans $\mathfrak{A}$ pour u qui satisfait l'énoncé du lemme.

Cas 2. L'état p appartient à l'étiquette d'un enfant de v seulement après l'application de l'étape 3. Par conséquent, $p \in \Delta(e_{m-1}(v), u_{m-1}) \cap F$ et il existe donc un chemin dans $\mathfrak{A}$ pour u qui commence en un état de $e_0(v)$ se termine en $p \in F$. Ce chemin vérifie l'énoncé du lemme. $\square$

Nous montrons à présent au travers de deux lemmes que l'automate déterministe de Rabin $\mathfrak{R} = (\mathcal{S}_n, \Sigma, \Delta', A_I, \Omega)$ que nous avons définit reconnaît le même langage que notre automate non déterministe de Muller $\mathfrak{A} = (Q, \Sigma, \Delta, 1, F)$.

Lemme II.3.4. *Nous avons $L(\mathfrak{R}) \subseteq L(\mathfrak{A})$.*

Preuve. Soit $\alpha \in L(\mathfrak{R})$ un mot infini accepté par l'automate déterministe de Rabin $\mathfrak{R}$. Notons $\rho = A_0 A_1 \cdots \in \mathcal{S}_n$ l'unique déroulement de $\mathfrak{R}$ pour α et notons v un élément de V pour lequel $\mathrm{Inf}(\rho) \cap E_v = \emptyset$ et $\mathrm{Inf}(\rho) \cap F_v \neq \emptyset$. Notons $A_i = (N_i, f_i, e_i, M_i)$ pour chaque $i \in \omega$.

Il existe $M \in \omega$ tel que pour tout $n \geq M$, v est un nœud de A_n, i.e. $v \in N_n$. Notons de plus, $(m_k)_{k \in \omega}$ la suite strictement croissante des naturels $n \geq M$ vérifiant $v \in M_n$. Pour tout $k \in \omega$, la suite $A_{m_k} A_{m_k+1} \cdots A_{m_{k+1}}$ est un chemin dans $\mathfrak{R}$ pour le mot $\alpha_{m_k} \cdots \alpha_{m_{k+1}-1}$ qui vérifie les hypothèses de Lemme II.3.3 pour v. Nous pouvons donc choisir pour tout entier $k \geq 1$ et tout $q_k \in e_{m_k}(v)$ un état $p(q_k) \in e_{m_{k-1}}(v)$ et une suite $\tilde{\rho}(q_k)$ d'états de $\mathfrak{A}$ telle que la suite $p(q_k)\tilde{\rho}(q_k)$ est un chemin dans $\mathfrak{A}$ pour $\alpha_{m_k} \cdots \alpha_{m_{k+1}-1}$ avec la propriété que $\mathrm{Occ}(\tilde{\rho}(q_k)) \cap F \neq \emptyset$. Afin d'appliquer le Lemme de König (II.2.2), nous définissons un arbre (N, r, p) dont l'ensemble des nœuds est

$$N = \{r\} \cap \{(k, q) \mid k \in \omega \text{ et } q \in e_{m_k}(v)\},$$

et la fonction parent est définie pour tout $(k,q)\in N\backslash\{r\}$ par

$$p(k,q)=\big(k-1,p(q_k)\big).$$

Cet arbre (N,r,p) est infini et d'arité finie, il existe donc dans (N,r,p) un chemin infini

$$r(0,q_0)(1,q_1)(2,q_2)\cdots\in N^*.$$

Choisissons alors un chemin fini $\tilde{\rho}_0\tilde{\rho}_1\cdots\tilde{\rho}_{m_0}$ dans $\mathfrak{A}$ pour $\alpha[0,m_0]$ commençant en l'état initial de $\mathfrak{A}$ et se terminant en q_0, i.e. tel que $\tilde{\rho}_0=1$ et $\tilde{\rho}_{m_0}=q_0$. Un déroulement de $\mathfrak{A}$ pour α est alors donné par concaténation :

$$\tilde{\rho}=\tilde{\rho}_0\tilde{\rho}_1\cdots\tilde{\rho}_{m_0}\tilde{\rho}(q_1)\tilde{\rho}(q_2)\cdots$$

et puisque pour tout $k\in\omega$, $\mathrm{Occ}(\tilde{\rho}(q_k))\cap F\neq\emptyset$, ce déroulement vérifie $\mathrm{Inf}(\tilde{\rho})\cap F\neq\emptyset$. Par conséquent, $\alpha\in L(\mathfrak{A})$. □

Lemme II.3.5. *Nous avons* $L(\mathfrak{A})\subseteq L(\mathfrak{R})$.

Preuve. Soient $\alpha\in L(\mathfrak{A})$ et $\tilde{\rho}=q_0q_1\cdots\in Q^\omega$ un déroulement de $\mathfrak{A}$ pour α tel que $\mathrm{Inf}(\tilde{\rho})\cap F\neq\emptyset$. Notons $\rho=A_0A_1\cdots\in\mathscr{S}_n^\omega$ l'unique déroulement de $\mathfrak{R}$ pour α et pour chaque $i\in\omega$ notons $A_i=(N_i,f_i,e_i,M_i)$. Nous commençons par montrer l'assertion suivante.

Assertion. Soit $v\in V$. Supposons qu'il existe $n_v\in\omega$ tel que $q_n\in e_n(v)$ pour tout $n\geq n_v$, et donc en particulier que $\mathrm{Inf}(\rho)\cap E_v=\emptyset$. Il y a deux possibilités qui s'excluent mutuellement.

1) Pour tout $n\in\omega$ il existe $i>n$ tel que $v\in M_i$, i.e. $\mathrm{Inf}(\rho)\cap F_v\neq\emptyset$.

2) Il existe $w\in V$ et $n_w\in\omega$ tel que pour tout $n\geq n_w$:

$$w \text{ est un enfant de } v \text{ dans } A_n \text{ et } q_n\in e_n(w).$$

En effet, supposons que **1)** ne soit pas vérifiée et notons $n'\in\omega$ le plus petit entier tel que pour tout $i>n'$, nous avons $v\notin M_i$. Notons alors m le plus petit naturel tel que $m>n'$ et $q_m\in\mathrm{Inf}(\tilde{\rho})\cap F$. Puisque q_m est un état acceptant, il apparaît dans l'étiquette d'un enfant w_1 de v dans A_m. Par conséquent, q_{m+1} appartient à l'étiquette de w_1 dans A_{m+1} à moins qu'il appartiennent à l'étiquette d'un nœud se trouvant à gauche de w_1, ce qui serait le résultat de l'application de l'étape 3. En fait, q_{m+1} ne saurait se trouver dans l'étiquette d'un nœud se trouvant à gauche de w_1 autre qu'un enfant w_2 de v se trouvant à gauche de w_1, car $q_{m+1}\in e_{m+1}(v)$ par hypothèse. Cependant, un tel déplacement vers un enfant de v plus âgé ne peut se produire qu'un nombre fini de fois. De plus, comme v n'est plus marqué à partir de n', cela implique que jamais ensuite tous les enfants ne seront supprimés à l'étape 5. Par conséquent, 2) est vérifiée.

Conclusion. Nous montrons qu'il existe $v\in V$ tel que $\mathrm{Inf}(\rho)\cap E_v=\emptyset$ et $\mathrm{Inf}(\rho)\cap F_v\neq\emptyset$ et donc que $\alpha\in L(\mathfrak{R})$. Commençons par remarquer que pour tout $i\in\omega$, l'état q_i appartient à l'étiquette de la racine de A_i, i.e. $q_i\in e_i(1)$. L'élément $1\in V$ satisfait les hypothèses de l'assertion. Par conséquent, ou bien 1 est marqué infiniment souvent et donc $\alpha\in L(\mathfrak{R})$, ou bien il existe $v_1\in V$ et $N_{v_1}\in\omega$ tel que pour tout $n\geq N_{v_1}$, d'une part $q_n\in e_n(v_1)$ et d'autre part v_1 est un enfant de la racine 1. En d'autres termes, ou bien 1 est le nœud cherché, ou bien nous avons un $v_1\in V$ qui à partir d'un certain temps est un enfant de 1 et qui satisfait les hypothèses de l'assertion. Notre nouveau candidat est v_1, et on répète l'argument pour v_1. Nous procédons ainsi de suite, jusqu'à trouver un $v\in V$ avec la propriété cherchée. Or V est fini, ou en d'autres termes la hauteur des arbres est finie, et donc forcément notre recherche va aboutir. □

CHAPITRE III

Automates et logique monadique du second ordre

> Nous avons vu que les automates permettent de définir des langages. Or, il existe une autre façon de décrire un ensemble de mots. Plus naturelle, dans un certain sens. Elle consiste à énoncer une propriété partagée par tous les mots de cet ensemble et seulement par ceux-ci. C'est ce que nous faisons par exemple sur l'alphabet binaire lorsque nous disons « l'ensemble des mots finis qui se terminent par un 1 » ou encore « l'ensemble des mots infinis pour lesquels un 1 survient toujours après un nombre pair de 0 ». La question se pose alors de savoir quel fragment de notre langage possède une expressivité semblable à celle des automates. Ce chapitre est consacré à la réponse à cette question.

Nous nous fixons un alphabet fini Σ.

III.1 La logique monadique du second ordre sur les mots

Le fragment de notre langage qui possède une expressivité identique aux automates est la logique monadique du second ordre.

III.1.1 La syntaxe

Les variables de la logique monadique du second ordre (MSO) sont de deux types. Les **variables du premier ordre** dont nous fixons l'ensemble dénombrable à :

$$\mathscr{V} = \{x_1, x_2, x_3, \ldots\},$$

et les **variables du deuxième ordre** ou **de prédicat** dont nous fixons l'ensemble dénombrable à :

$$\mathscr{P} = \{X_1, X_2, X_3, \ldots\}.$$

Nous allons considérer la signature

$$\mathscr{S} = \{<\} \cup \{Q_a \mid a \in \Sigma\}$$

dans laquelle la symbole primitif $<$ est un symbole de relation binaire et chaque symbole primitif Q_a est symbole de prédicat. La logique que nous décrivons est égalitaire, le symbole d'égalité est donc implicitement présent et son interprétation sera la relation d'égalité.

L'ensemble des formules atomiques de la logique monadique du second ordre de signature $\mathscr{S}$ est alors donné par

$$\mathrm{At(MSO)} = \{x = y, x < y, Q_a(x), X(x) \mid x, y \in \mathscr{V}, X \in \mathscr{P} \text{ et } a \in \Sigma\}.$$

Afin de définir les formules, nous définissons par induction F_n pour tout $n \in \omega$ par :

$$\begin{aligned} F_0 =& \mathrm{At(MSO)} \\ F_{n+1} =& \{\neg\phi, \phi \vee \psi \mid \phi, \psi \in F_n\} \\ & \cup \{\exists x\phi \mid x \in \mathscr{V} \text{ et } \phi \in F_n\} \\ & \cup \{\exists X\phi \mid X \in \mathscr{P} \text{ et } \phi \in F_n\} \end{aligned}$$

et posons l'ensemble des formules de MSO comme

$$\mathscr{F}(\mathrm{MSO}) = \bigcup_{n\in\omega} F_n.$$

Nous faisons les définitions habituelles :

$$\begin{aligned} \forall x\phi &\stackrel{\text{déf}}{=} \neg\exists x\neg\phi & \phi \wedge \psi &\stackrel{\text{déf}}{=} \neg(\neg\phi \vee \neg\psi) \\ \forall X\phi &\stackrel{\text{déf}}{=} \neg\exists X\neg\phi & \phi \to \psi &\stackrel{\text{déf}}{=} \neg\phi \vee \psi \end{aligned}$$

Pour une formule $\phi \in \mathscr{F}(MSO)$, nous notons $FV_1(\phi)$ l'ensemble des variables libres du premier ordre de ϕ et $FV_2(\phi)$ l'ensemble des variables libres du second ordre de ϕ. Une formule ϕ est close si $FV_1(\phi) = \emptyset = FV_2(\phi)$.

III.1.2 La sémantique

Nous allons décrire deux sémantiques distinctes pour le langage de MSO. La première interprétera les formules de MSO, dans les structures de mots finis, tandis que la deuxième les interprétera dans les structures de mots infinis.

La **structure d'un mot fini**(non vide) $u = u_0u_1\cdots u_{n-1} \in \Sigma^*\backslash\{\epsilon\}$ est définie comme le triplet

$$\underline{u} = (\mathrm{dom}(u), <^u, (Q_a^u)_{a\in\Sigma}),$$

où $\mathrm{dom}(u) = \{0, 1, \ldots, n-1\} = n$, $<^u$ est la relation d'ordre sur n et pour tout $a \in \Sigma$,

$$Q_a^u = \{i \in \mathrm{dom}(u) \mid u_i = a\}.$$

Nous définissons de façon analogue, la **structure d'un mot infini** $\alpha = \alpha_0\alpha_1\cdots\alpha_{n-1} \in \Sigma^\omega$ est définie comme le triplet

$$\underline{\alpha} = (\mathrm{dom}(\alpha), <^\alpha, (Q_a^\alpha)_{a\in\Sigma}),$$

où $\mathrm{dom}(\alpha) = \omega$, $<^\alpha$ est la relation d'ordre sur ω et pour tout $a \in \Sigma$,

$$Q_a^\alpha = \{i \in \omega \mid \alpha_i = a\}.$$

En particulier, les structures des mots finis ou infinis sont des structures au sens habituel pour la signature $\mathscr{S}$ dans le sens où ce sont des ensembles muni d'une interprétation de chaque symbole primitif. La notion de fonction d'interprétation et la relation de satisfaction entre structures et formules que nous utilisons sont tout à fait standard. Une **fonction d'interprétation** pour une structure de mot fini $\underline{u}$ est la donnée $\delta = (\delta_1, \delta_2)$ de deux fonctions

$$\begin{aligned}\delta_1 &: \mathscr{V} \to \mathrm{dom}(u)\\ \delta_2 &: \mathscr{P} \to \mathscr{P}(\mathrm{dom}(u)).\end{aligned}$$

Soit $\delta = (\delta_1, \delta_2)$ une fonction d'interprétation pour $\underline{u}$. Pour $k \in \mathrm{dom}(u)$ et pour une variable du premier ordre $x \in \mathscr{V}$, nous notons $\delta^{x\to k} = (\delta_1^{x\to k}, \delta_2^{x\to k})$ la fonction d'interprétation pour $\underline{u}$ définie par $\delta_2^{x\to k} = \delta_2$ et pour tout $x_i \in \mathscr{V}$

$$\delta_1^{x\to k}(x_i) = \begin{cases} \delta_1(x_i) & \text{si } x_i \neq x;\\ k & \text{si } x_i = x.\end{cases}$$

Pour $K \subseteq \mathrm{dom}(u)$ et pour une variable du deuxième ordre $X \in \mathscr{P}$ nous notons $\delta^{X\to K} = (\delta_1^{X\to K}, \delta_2^{X\to K})$ la fonction d'interprétation pour $\underline{u}$ définie par $\delta_1^{X\to K} = \delta_1$ et pour tout $X_i \in \mathscr{P}$

$$\delta_2^{X\to K}(X_i) = \begin{cases} \delta_2(X_i) & \text{si } X_i \neq X;\\ K & \text{si } X_i = X.\end{cases}$$

De façon similaire, une **fonction d'interprétation** pour une structure de mot infini $\underline{\alpha}$ est la donnée $\delta = (\delta_1, \delta_2)$ de deux fonctions

$$\begin{aligned}\delta_1 &: \mathscr{V} \to \omega\\ \delta_2 &: \mathscr{P} \to \mathscr{P}(\omega).\end{aligned}$$

Comme dans le cas des structures de mots finis, nous définissons les fonctions d'interprétation $\delta^{x\to k}$ et $\delta^{X\to K}$ pour toute fonction d'interprétation δ pour $\underline{\alpha}$ et tous $k \in \omega$, $x \in \mathscr{V}$, $K \subseteq \omega$, et $X \in \mathscr{P}$.

Le relation de satisfaction $\models$ entre formules, structures de mots finis ou infinis et fonctions d'interprétation est définie de façon usuelle, par induction sur la hauteur des formules. Soit $\underline{w} = (\mathrm{dom}(w), <^w, (Q_a^w)_{a\in\Sigma})$ la structure d'un mot fini ou infini w et $\delta = (\delta_1, \delta_2)$ une fonction d'interprétation pour $\underline{w}$. Pour les formules atomiques, nous posons :

$$\begin{array}{lll} \underline{w}, \delta \models x = y & \text{si et seulement si} & \delta_1(x) = \delta_1(y)\\ \underline{w}, \delta \models x < y & \text{si et seulement si} & \delta_1(x) <^w \delta_1(y)\\ \underline{w}, \delta \models X(x) & \text{si et seulement si} & \delta_1(x) \in \delta_2(X)\\ \underline{w}, \delta \models Q_a(x) & \text{si et seulement si} & \delta_1(x) \in Q_a^w \end{array}$$

tandis que la marche d'induction est définie par :

$$\begin{array}{lll} \underline{w}, \delta \models \neg\phi & \text{si et seulement si} & \text{non } \underline{w}, \delta \models \phi\\ \underline{w}, \delta \models \phi \vee \psi & \text{si et seulement si} & \underline{w}, \delta \models \phi \text{ ou } \underline{w}, \delta \models \psi\\ \underline{w}, \delta \models \exists x\phi & \text{si et seulement si} & \text{il existe } k \in \mathrm{dom}(w) \text{ tel que } \underline{w}, \delta^{x\to k} \models \phi\\ \underline{w}, \delta \models \exists X\phi & \text{si et seulement si} & \text{il existe } K \subseteq \mathrm{dom}(w) \text{ tel que } \underline{w}, \delta^{X\to K} \models \phi. \end{array}$$

Nous allons également nous intéresser à la sémantique dans laquelle les variables du second ordre ne portent que sur les sous-ensembles finis du domaine. Nous appelons cette logique **MSO faible** (*Weak* MSO), elle diffère de MSO seulement par la définition de sa relation de satisfaction. La relation de satisfaction $\models_f$ de MSO faible est définie similairement à celle de MSO, sauf pour la dernière ligne ci-dessus qui est remplacée par la ligne suivante

$$\underline{w}, \delta \models_f \exists X\phi \quad \text{si et seulement si} \quad \begin{array}{l}\text{il existe un sous-ensemble fini } K \subseteq \mathrm{dom}(w)\\ \text{tel que } \underline{w}, \delta^{X\to K} \models \phi.\end{array}$$

Remarquons que pour la structure $\underline{w}$ d'un mot fini ou infini w et $\phi \in \mathscr{F}(MSO)$ fixés, la validité de $\underline{w}, \delta \models \phi$ ne dépend que des valeurs de la fonction d'interprétation δ sur les variables libres du premier et du second ordre de ϕ. Nous exploiterons amplement cette remarque dans la section suivante.

Pour une formule close ϕ et un mot fini ou infini w, la validité de $\underline{w}, \delta \models \phi$ ne dépend pas de la fonction d'interprétation δ et nous notons dans ce cas $\underline{w} \models \phi$.

Le **langage sur les mots finis reconnu par une formule close** $\phi \in \mathscr{F}(MSO)$ est défini par

$$l(\phi) = \{u \in \Sigma^* \backslash \{\epsilon\} \mid \underline{u} \models \phi\}.$$

Remarquons que MSO et MSO faible sont identiques sur les mots finis puisque les domaines des structures de mots finis sont finis et donc

$$\{u \in \Sigma^* \backslash \{\epsilon\} \mid \underline{u} \models_f \phi\} = l(\phi)$$

De façon analogue, nous définissons le **langage sur les mots infinis reconnu par une formule close** $\phi \in \mathscr{F}$(MSO) par

$$L(\phi) = \{\alpha \in \Sigma^\omega \mid \underline{\alpha} \models \phi\}.$$

et nous distinguons cette fois le **langage sur les mots infinis reconnu faiblement par une formule close** $\phi \in \mathscr{F}$(MSO) qui est défini par

$$L_f(\phi) = \{\alpha \in \Sigma^\omega \mid \underline{\alpha} \models_f \phi\}.$$

Nous définissons la relation S de successeur sur les naturels comme

$$S(x_i, x_i) \stackrel{\text{déf}}{=} (x_i < x_j) \wedge \forall y \big(x_i < y \rightarrow (x_j = y \vee x_j < y)\big).$$

Remarquons aussi que si nous avions choisi le symbole de relation S de successeur parmi nos symboles primitifs à la place de la relation d'ordre $<$, et si nous avions muni les structures de mots de la relation de successeur sur leur domaine plutôt que de la relation d'ordre, nous aurions pu définir la relation d'ordre comme :

$$x_i < x_j \stackrel{\text{déf}}{=} \exists X \Big[\forall x \forall y \big((x \in X \wedge S(x,y)) \rightarrow y \in X\big)\Big] \wedge x_i \notin X \wedge x_j \in X.$$

Cette formule signifie intuitivement qu'il existe un intervalle X de la forme $[k, +\infty[$ contenant x_j mais pas x_i. Le quantificateur existentiel doit donc porter également sur les sous-ensembles infinis du domaine. Toutefois, pour exprimer la relation d'ordre à partir de la relation de successeur dans MSO faible, il est possible d'utiliser la formule

$$x_i < x_j \stackrel{\text{déf}}{=} \exists X [\forall x \forall y (x \in X \wedge S(x,y)) \rightarrow (y \in X \vee y = x_j)] \wedge x_i \in X.$$

III.2 MSO et les automates

Le but de cette section est de montrer que les automates ont la même capacité expressive que la logique MSO tant sur les mots finis que sur les mots infinis.

III.2.1 Une formule pour chaque automate

Théorème III.2.1. *Soit $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F)$ un automate non déterministe où la composante d'acceptation est un sous-ensemble $F \subseteq Q$. Il existe deux formules closes de la logique monadique du second ordre $\phi^*, \phi^\omega \in \mathscr{F}(MSO)$ telle que*

$$l(\phi^*) = l(\mathfrak{A}) \backslash \{\epsilon\} \quad \text{et} \quad L(\phi^\omega) = L(\mathfrak{A}).$$

Preuve. Nous supposons que $Q = \{1, \dots, n\}$ et que l'état initial de $\mathfrak{A}$ est l'état 1. Nous écrivons une formule ψ dont la clôture existentielle signifie, pour une structure de mot fini ou infini $\underline{w}$, l'existence d'un déroulement de l'automate $\mathfrak{A}$ pour w. Nous allons exprimer que pour chaque $k \in Q$, il existe un sous-ensemble du domaine X_k dont les éléments sont les indices auxquels le déroulement visite l'état k. La formule dit que les X_k sont disjoints deux à deux et que si $x \in X_k$, $x + 1 \in X_{k'}$ et le $x^{\text{ème}}$ symbole est $a \in \Sigma$, alors $(k, a, k') \in \Delta$. Formellement,

$$\psi = \left(\bigwedge_{\substack{k,k' \in Q \\ \text{tel que } k \neq k'}} \neg \exists x \big(X_k(x) \wedge X_{k'}(x)\big) \right) \wedge \left(\forall x \Big(\big(\forall y \neg S(y,x)\big) \to X_1(x) \Big) \right)$$
$$\wedge \left(\forall x \forall y \left(S(x,y) \to \bigvee_{(k,a,k') \in \Delta} \big(X_k(x) \wedge Q_a(x) \wedge X_{k'}(y)\big) \right) \right).$$

$l(\phi^*) = l(\mathfrak{A})$. Nous écrivons maintenant dans le cas des mots finis une formule qui exprime le fait qu'un déroulement est acceptant. Il suffit de stipuler que pour une transition $(k, a, f) \in \Delta$ avec $f \in F$, l'élément maximal du domaine, i.e. l'indice de l'avant dernier état du déroulement, est dans X_k et le dernier symbole du mot est a. Ceci nous donne dans MSO :

$$\phi^* = \exists X_1 \dots \exists X_n \left[\psi \wedge \forall x \left((\forall y \neg S(x,y)) \to \Big(\bigvee_{\substack{(k,a,f) \in \Delta \\ \text{tel que } f \in F}} X_k(x) \wedge Q_a(x) \Big) \right) \right].$$

Cette formule de MSO reconnaît bien le langage sur les mots finis de $\mathfrak{A}$. En effet, si $u = u_0 u_1 \cdots u_{m-1} \in l(\mathfrak{A})$, alors il existe un déroulement $d = d_0 d_1 \cdots d_m \in Q^*$ de $\mathfrak{A}$ pour u tel que $d_m \in F$. Pour voir que $\underline{u} \models \phi^*$, il suffit de choisir pour chaque $i = 1, \dots, n$:

$$X_i = \{ j \in \text{dom}(u) \mid d_j = i \}.$$

Réciproquement, si $u \in l(\phi^*)$ donnons nous une partition $X_1, \dots, X_n \subseteq \text{dom}(u)$ satisfaisant ϕ^*, alors un déroulement acceptant de $\mathfrak{A}$ pour u est donné par $d_j = i$ où i est tel que $j \in X_i$, pour $j = 1, \dots, m-1$ et $d_m = f$ tel que $(d_{m-1}, u_{m-1}, f) \in \Delta$.

$L(\phi^\omega) = L(\mathfrak{A})$. Nous exprimons maintenant pour les mots infinis une formule exprimant le fait qu'un déroulement est acceptant. Il suffit d'écrire dans MSO qu'une infinité d'éléments du domaine ω se trouvent dans un X_f avec $f \in F$. Formellement,

$$\phi^\omega = \exists X_1 \dots \exists X_n \left[\psi \wedge \Big(\bigvee_{f \in F} \forall x \exists y (x < y \wedge X_f(y)) \Big) \right].$$

Cette formule de MSO reconnaît bien le langage sur les mots infinis de $\mathfrak{A}$. En effet, si $\alpha = \alpha_0 \alpha_1 \cdots \in L(\mathfrak{A})$, alors il existe un déroulement $\rho = \rho_0 \rho_1 \cdots \in Q^\omega$ de $\mathfrak{A}$ pour α tel que $\text{Inf}(\rho) \cap F \neq \emptyset$. Pour voir que $\underline{\alpha} \models \phi^\omega$, il suffit de choisir pour chaque $i = 1, \dots, n$

$$X_i = \{ j \in \omega \mid \rho_j = i \}.$$

Réciproquement, si $\alpha \in L(\phi^\omega)$, il existe une partition $X_1, \dots, X_n$ de ω satisfaisant ϕ_ω. Un déroulement acceptant de $\mathfrak{A}$ pour α est alors donné pour tout $j \in \omega$ par $\rho_j = i$ où i est tel que $j \in X_i$. $\square$

III.2.2 Un automate pour chaque formule

Nous allons montrer à présent la réciproque du théorème de la sous-section précédente, dont voici l'énoncé.

Théorème III.2.2. *Soit $\phi \in \mathscr{F}(MSO)$ une formule close de MSO. Il est possible de construire effectivement un automate sur les mots finis $\mathfrak{A}$ et un automate de Büchi $\mathfrak{B}$ tels que*

$$l(\phi) = l(\mathfrak{A}) \quad et \quad L(\phi) = L(\mathfrak{B}).$$

Nous allons procéder par induction sur la hauteur des formules dans MSO. Pour cela, nous allons devoir considérer le langage reconnu par une formule contenant des variables libres. Cette approche peut sembler un détour fastidieux du fait que notre véritable intérêt est dans les formules closes, alors que les formules contenant des variables libres semblent plutôt être un artifice de construction, dont la présence a posteriori est plutôt indésirable. Toutefois, c'est un moyen clair d'atteindre notre but compte tenu de la définition que nous avons des formules et de la relation de satisfaction.

Afin de définir la satisfaction d'une formule ϕ contenant des variables libres, il nous faut donner une interprétation des variables de ϕ. Comme nous l'avons déjà remarqué, la satisfaction d'une formule ϕ par une structure $\underline{w}$ et une fonction d'interprétation δ pour $\underline{w}$ ne dépend pas des valeurs prisent par δ sur les variables qui ne sont pas libres dans ϕ. Seules les valeurs prisent par δ sur les variables libres de ϕ jouent un rôle. Pour toute formule $\phi \in \mathscr{F}(MSO)$, nous adoptons la notation suivante :

$$p(\phi) = \max\{i \in \omega \mid x_i \text{ est libre dans } \phi\} \quad \text{et} \quad q(\phi) = \max\{i \in \omega \mid X_i \text{ est libre dans } \phi\},$$

avec la convention que $p(\phi) = 0$ s'il n'y a aucune occurrence libre d'une variable du premier ordre dans ϕ et $q(\phi) = 0$ s'il n'y aucune occurrence libre d'une variable du second ordre dans ϕ.

Pour parler de la satisfaction d'un formule ϕ, il est suffisant d'interpréter les variables du premier ordre $x_1, \dots, x_{p(\phi)}$ et les variables du second ordre $X_1, \dots, X_{q(\phi)}$.

Afin d'interpréter les variables libres, nous allons considérer un alphabet modifié. Pour tous $p, q \in \omega$ nous définissons l'alphabet

$$\Sigma_{p,q} = \Sigma \times \{0,1\}^p \times \{0,1\}^q,$$

et nous allons voir les mots finis sur un tel alphabet à travers la bijection

$$\Sigma_{p,q}^* = \Sigma^* \times \underbrace{\{0,1\}^* \times \cdots \{0,1\}^*}_{p+q \text{ fois}}$$

et les mots infinis sur un tel alphabet seront vus à travers la bijection

$$\Sigma_{p,q}^\omega = \Sigma^\omega \times \underbrace{\{0,1\}^\omega \times \cdots \{0,1\}^\omega}_{p+q \text{ fois}}.$$

Pour tout $p, q \in \omega$, nous allons considérer le langage de mots fini sur $\Sigma_{p,q}$ défini par

$$K^*_{p,q} = \{(u_0, u_1, \ldots, u_p, u_{p+1}, \ldots, u_{p+q}) \in \Sigma^*_{p,q} \mid \text{ pour tout } i = 1, \ldots, p \ |u_i|_1 = 1\}$$

ainsi que le langage sur les mots infinis défini par

$$K^\omega_{p,q} = \{(u_0, u_1, \ldots, u_p, u_{p+1}, \ldots, u_{p+q}) \in \Sigma^\omega_{p,q} \mid \text{ pour tout } i = 1, \ldots, p \ |u_i|_1 = 1\}.$$

Indiquons déjà qu'étant donné un mot $(u_0, u_1, \ldots, u_{p+q}) \in K^*_{p,q}$, les p mots $u_1, \ldots, u_p$ sur $\{0,1\}$ codent l'interprétation des variables du premier ordre $x_1, \ldots, x_p$, tandis que les mots $u_{p+1}, \ldots, u_{p+q}$ sur $\{0,1\}$ codent l'interprétation des variables du second ordre $X_1, \ldots, X_q$. Pour les mots de $K^\omega_{p,q}$, l'idée est similaire et nous allons préciser ceci par la suite. Nous commençons par montrer que les langages $K^*_{p,q}$ et $K^\omega_{p,q}$ sont reconnus par des automates.

Lemme III.2.3. *Pour tous $p, q \in \omega$, nous avons $K^*_{p,q} \in \mathrm{REG}(\Sigma_{p,q})$ et $K^\omega_{p,q} \in \mathrm{REG}^\omega(\Sigma_{p,q})$.*

Preuve. Soient $p, q \in \omega$. Considérons l'automate non déterministe $\mathfrak{A} = (Q, \Sigma_{p,q}, \Delta, q_I, F)$ dont l'ensemble des états est $Q = \{0,1\}^p$, dont l'état initial est $q_I = (0, \ldots, 0) \in Q$, dont la relation de transition est donnée par

$$\Delta = \Big\{ \big((v_1, \ldots, v_p), (a_0, a_1, \ldots, a_{p+q}), (w_1, \ldots, w_p)\big) \in Q \times \Sigma_{p,q} \times Q \\ \Big| \forall i \in \{1, \ldots, p\} \big((v_i = 0 \vee a_i = 0) \wedge (w_i = |a_i - v_i|)\big) \Big\},$$

et dont la composante d'acceptation est le singleton $\{(1, \ldots, 1)\} \subseteq Q$. Intuitivement, lors de la lecture d'un mot $(w_0, w_1, \ldots, w_{p+q})$, l'automate $\mathfrak{A}$ débute à l'état $(0, \ldots, 0)$, il se déplace en $(0, \ldots, 0, 1, 0, \ldots, 0)$ avec un 1 à la $i^{\text{ème}}$ position lorsqu'il lit un 1 dans w_i, il s'arrête s'il lit à nouveau un 1 dans w_i. Il arrive en $(1, \ldots, 1)$ lorsqu'il a lu un 1 et un seul dans chaque w_i pour $i = 1, \ldots, p$ et stationne dès lors dans cet état, à moins qu'il lise un nouveau 1 dans w_i avec $i = 1, \ldots, p$ auquel cas il s'arrête. Ainsi, en tant qu'automate sur les mots finis $\mathfrak{A}$ reconnaît $K^*_{p,q}$, i.e. $l(\mathfrak{A}) = K^*_{p,q}$. Tandis qu'en tant qu'automate de Büchi, il reconnaît $K^\omega_{p,q}$, i.e. $L(\mathfrak{A}) = K^\omega_{p,q}$. □

Nous formalisons maintenant l'idée selon laquelle un mot dans $K^*_{p,q}$ ou $K^\omega_{p,q}$ est un mot sur Σ muni d'une interprétation des p premières variables du premier ordre et des q premières variables du second ordre. Pour tout $u = (u_0, u_1, \ldots, u_{p+q}) \in K^*_{p,q}$, nous définissons la **fonction d'interprétation partielle** de u comme le couple $\delta^u = (\delta^u_1, \delta^u_2)$ où

$$\delta^u_1 : \{x_1, \ldots, x_p\} \to \mathrm{dom}(u_0)$$
$$\delta^u_2 : \{X_1, \ldots, X_q\} \to \mathscr{P}(\mathrm{dom}(u_0)),$$

sont définies par $\delta^u_1(x_i) = n_i$ où $n_i \in \mathrm{dom}(u_0)$ est l'index de l'unique 1 du mot u_i, et

$$\delta^u_2(X_i) = \{j \in \mathrm{dom}(u_0) \mid u_{p+i,j} = 1\}.$$

De façon analogue, pour tout mot $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{p+q}) \in K^\omega_{p,q}$, nous définissons la fonction d'interprétation partielle de α comme le couple $\delta^\alpha = (\delta^\alpha_1, \delta^\alpha_2)$ où

$$\delta^\alpha_1 : \{x_1, \ldots, x_p\} \to \omega$$
$$\delta^\alpha_2 : \{X_1, \ldots, X_q\} \to \mathscr{P}(\omega),$$

sont définies par $\delta_1^\alpha(x_i) = n_j$ où $n_j \in \omega$ est l'index de l'unique 1 du mot α_i, et

$$\delta_2^\alpha(X_i) = \{j \in \omega \mid \alpha_{p+i,j} = 1\}.$$

Soit maintenant $\phi \in \mathscr{F}(\mathrm{MSO})$ une formule de MSO contenant possiblement des variables libres et un couple (p,q) de naturels tel que $p \geq p(\phi)$ et $q \geq q(\phi)$. Nous disons qu'un mot fini $u = (u_0, u_1, \ldots, u_{p+q}) \in K^*_{p,q}$ **satisfait** ϕ si le couple $(\underline{u_0}, \delta_u)$ formé de la structure $\underline{u_0} = (\mathrm{dom}(u_0), <^{u_0}, (Q_a^{u_0})_{a\in\Sigma})$ du mot u_0 et de la fonction d'interprétation partielle δ^u de u vérifie $\underline{u_0}, \delta^u \models \phi$ au sens de la section précédente. De la même façon, nous disons qu'un mot infini $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{p+q}) \in K^\omega_{p,q}$ **satisfait** ϕ si $\underline{\alpha_0}, \delta^\alpha \models \phi$.

Nous adoptons les notations suivantes pour tout formule $\phi \in \mathscr{F}(\mathrm{MSO})$ et tout couple de naturel (p,q) tel que $p \geq p(\phi)$ et $q \geq q(\phi)$:

$$S^*_{p,q}(\phi) = \{u \in K^*_{p,q} \mid u \text{ satisfait } \phi\}$$
$$S^\omega_{p,q}(\phi) = \{\alpha \in K^\omega_{p,q} \mid \alpha \text{ satisfait } \phi\}.$$

Dans le cas d'une formule close ϕ de MSO, nous avons $p(\phi) = 0 = q(\phi)$ et nous retrouvons le langage reconnu par ϕ dans le sens où :

$$S^*_{0,0} = l(\phi) \quad \text{et} \quad S^\omega_{0,0} = L(\phi).$$

Nous allons montrer que pour toute formule $\phi \in \mathscr{F}(\mathrm{MSO})$ et tout couple de naturels (p,q) tel que $p \geq p(\phi)$ et $q \geq q(\phi)$, nous avons $S^*_{p,q}(\phi) \in \mathrm{REG}(\Sigma_{p,q})$ et $S^\omega_{p,q}(\phi) \in \mathrm{REG}^\omega(\Sigma_{p,q})$. Nous aurons en corollaire par la remarque précédente que pour tout formule close ϕ de MSO, $l(\phi) \in \mathrm{REG}(\Sigma)$ et $L(\phi) \in \mathrm{REG}^\omega(\Sigma)$.

La base de notre preuve par induction sur la hauteur des formules est assurée par le lemme suivant.

Lemme III.2.4. *Soient $\phi \in \mathrm{At}(MSO)$ une formule atomique de MSO et (p,q) un couple de naturels tels que $p \geq p(\phi)$ et $q \geq q(\phi)$. Le langage sur les mots finis $S^*_{p,q}(\phi)$ appartient à $\mathrm{REG}(\Sigma_{p,q})$ et le langage sur les mots infinis $S^\omega_{p,q}(\phi)$ appartient à $\mathrm{REG}^\omega(\Sigma_{p,q})$.*

Preuve. Nous montrons le lemme pour chacune des différentes formules atomiques de MSO. Nous allons dans chaque cas définir une relation de transition Δ_ϕ pour l'automate non déterministe $\mathfrak{A}_\phi = (Q, \Sigma_{p,q}, \Delta_\phi, q_I, F)$ dont l'ensemble des états est $Q = \{1,2\}$, dont l'état initial est $q_I = 1$, la composante d'acceptation est le singleton $F = \{2\} \subseteq Q$. Cet automate reconnaît le langage $l(\mathfrak{A}_\phi) \subseteq \Sigma^*_{p,q}$ en tant qu'automate sur les mots finis et le ω-langage $L(\mathfrak{A}_\phi) \subseteq \Sigma^\omega_{p,q}$ en tant qu'automate de Büchi. Dans chaque cas, la conclusion est obtenue du fait que Δ_ϕ est défini de telle sorte que

$$S^*_{p,q}(\phi) = K^*_{p,q} \cap l(\mathfrak{A}_\phi) \quad \text{et} \quad S^\omega_{p,q}(\phi) = K^\omega_{p,q} \cap L(\mathfrak{A}_\phi),$$

et du fait que l'ensemble des langages réguliers et l'ensemble des ω-langages reconnaissables sont clos par intersection.

$x_i = x_j$. Soient $p \geq \max\{i,j\}$ et $q \in \omega$. Nous définissons la relation de transition

$$\begin{aligned}\Delta_{x_i=x_j} = &\{(1, (a_0, a_1, \ldots, a_{p+q}), 1) \in Q \times \Sigma_{p,q} \times Q \mid a_i \neq 1 \wedge a_j \neq 1\}\\ &\cup \{(1, (a_0, a_1, \ldots, a_{p+q}), 2) \in Q \times \Sigma_{p,q} \times Q \mid a_i = 1 \wedge a_j = 1\}\\ &\cup \{2\} \times \Sigma_{p,q} \times \{2\}.\end{aligned}$$

$x_i < x_j$. Soient $p \geq \max\{i, j\}$ et $q \in \omega$. Nous définissons la relation de transition

$$\begin{aligned}\Delta_{x_i<x_j} = &\{(1,(a_0,a_1,\ldots,a_{p+q}),1) \in Q \times \Sigma_{p,q} \times Q \mid a_i \neq 1 \wedge a_j \neq 1\}\\ &\cup \{(1,(a_0,a_1,\ldots,a_{p+q}),2) \in Q \times \Sigma_{p,q} \times Q \mid a_i = 1 \wedge a_j \neq 1\}\\ &\cup \{2\} \times \Sigma_{p,q} \times \{2\}.\end{aligned}$$

$Q_a(x_i)$. Soient $p \geq i$ et $q \in \omega$. Nous définissons la relation de transition

$$\begin{aligned}\Delta_{Q_a(x_i)} = &\{(1,(a_0,a_1,\ldots,a_{p+q}),1) \in Q \times \Sigma_{p,q} \times Q \mid a_i \neq 1\}\\ &\cup \{(1,(a_0,a_1,\ldots,a_{p+q}),2) \in Q \times \Sigma_{p,q} \times Q \mid a_i = 1 \wedge a_0 = 1\}\\ &\cup \{2\} \times \Sigma_{p,q} \times \{2\}.\end{aligned}$$

$X_i(x_j)$. Soient $p \geq j$ et $q \geq i$. Nous définissons la relation de transition

$$\begin{aligned}\Delta_{X_i(x_j)} = &\{(1,(a_0,a_1,\ldots,a_{p+q}),1) \in Q \times \Sigma_{p,q} \times Q \mid a_j \neq 1\}\\ &\cup \{(1,(a_0,a_1,\ldots,a_{p+q}),2) \in Q \times \Sigma_{p,q} \times Q \mid a_j = 1 \wedge a_{p+i} = 1\}\\ &\cup \{2\} \times \Sigma_{p,q} \times \{2\}.\end{aligned}$$ □

L'étape d'induction est ensuite obtenue sur la base des deux prochains lemmes.

Lemme III.2.5. *Soient $\phi, \psi \in \mathscr{F}(MSO)$.*

1) Pour tout couple (p,q) de naturels tel que $p \geq \max\{p(\phi), p(\psi)\}$ et $q \geq \max\{q(\phi), q(\psi)\}$,

$$S^*_{p,q}(\phi) \in \mathrm{REG}(\Sigma_{p,q}) \text{ et } S^*_{p,q}(\psi) \in \mathrm{REG}(\Sigma_{p,q}) \Rightarrow S^*_{p,q}(\phi \vee \psi) \in \mathrm{REG}(\Sigma_{p,q}), \text{ et}$$
$$S^\omega_{p,q}(\phi) \in \mathrm{REG}^\omega(\Sigma_{p,q}) \text{ et } S^\omega_{p,q}(\psi) \in \mathrm{REG}^\omega(\Sigma_{p,q}) \Rightarrow S^\omega_{p,q}(\phi \vee \psi) \in \mathrm{REG}^\omega(\Sigma_{p,q}).$$

2) Pour tout couple (p,q) de naturels tel que $p \geq p(\phi)$,

$$S^*_{p,q}(\phi) \in \mathrm{REG}(\Sigma_{p,q}) \Rightarrow S^*_{p,q}(\neg\phi) \in \mathrm{REG}(\Sigma_{p,q}), \text{ et}$$
$$S^\omega_{p,q}(\phi) \in \mathrm{REG}^\omega(\Sigma_{p,q}) \Rightarrow S^\omega_{p,q}(\neg\phi) \in \mathrm{REG}^\omega(\Sigma_{p,q}).$$

Preuve. Le lemme est une conséquence des propriétés de clôture par intersection et complémentation de REG et REG^ω et des formules suivantes qui découlent des définitions. Nous avons d'une part

$$S^*_{p,q}(\phi \vee \psi) = S^*_{p,q}(\phi) \cup S^*_{p,q}(\psi) \quad \text{et} \quad S^\omega_{p,q}(\phi \vee \psi) = S^\omega_{p,q}(\phi) \cup S^\omega_{p,q}(\psi),$$

et d'autre part,

$$S^*_{p,q}(\neg\phi) = K^*_{p,q} \cap \left(S^*_{p,q}(\phi)\right)^c \quad \text{et} \quad S^\omega_{p,q}(\neg\phi) = K^\omega_{p,q} \cap \left(S^\omega_{p,q}(\phi)\right)^c,$$

où le complémentaire est pris dans $\Sigma^*_{p,q}$ et $\Sigma^\omega_{p,q}$ respectivement. □

Étant donné deux automates reconnaissant $S^\omega_{p,q}(\phi)$ et $S^\omega_{p,q}(\psi)$ respectivement, les résultats des Chapitres I et II nous fournissent une construction effective de deux automates reconnaissant $S^\omega_{p,q}(\phi \vee \psi)$ et $S^\omega_{p,q}(\neg\phi)$ respectivement. Les constructions effective de ces automates dans le cas des mots finis sont par exemple décrites dans [Sipser, 2006].

Nous regardons maintenant le cas des formules obtenue par quantification existentielle.

Lemme III.2.6. *Soient $\phi \in \mathscr{F}(MSO)$ et un couple (p,q) de naturels tels que $p \geq p(\phi)$ et $q \geq q(\phi)$. Notons*

$$\pi_p : \Sigma^*_{p,q} \to \Sigma^*_{p-1,q} \qquad \pi_q : \Sigma^*_{p,q} \to \Sigma^*_{p,q-1}$$
$$\pi_p : \Sigma^\omega_{p,q} \to \Sigma^\omega_{p-1,q} \qquad \pi_q : \Sigma^\omega_{p,q} \to \Sigma^\omega_{p,q-1}$$

les fonctions qui suppriment la $p^{ème}$ composante et la $p+q^{ème}$ composante respectivement. Nous avons

$$S^*_{p-1,q}(\exists x_p \phi) = \pi_p(S^*_{p,q}(\phi)) \qquad S^*_{p,q-1}(\exists X_q \phi) = \pi_q(S^*_{p,q}(\phi))$$
$$S^\omega_{p-1,q}(\exists x_p \phi) = \pi_p(S^\omega_{p,q}(\phi)) \qquad S^\omega_{p,q-1}(\exists X_q \phi) = \pi_q(S^\omega_{p,q}(\phi)).$$

Preuve. Cela découle directement des définitions. □

Remarquons aussi que les propositions I.1.8 et I.3.8 fournissent une construction effective d'un automate reconnaissant la projection d'un langage $\pi(L)$ à partir d'un automate reconnaissant le langage L. En particulier, les langages du lemmes précédents sont respectivement réguliers et reconnaissables.

Considérons maintenant une formule de la forme $\exists x_i \phi$ et des entiers p et q tels que $p \geq p(\exists x_i \phi)$ et $q \geq q(\exists x_i \phi)$. Nous pouvons substituer les occurrences de x_i dans ϕ par x_{p+1} pour obtenir la formule $\phi[x_{p+1}/x_i]$. La formule $\exists x_{p+1} \phi[x_{p+1}/x_i]$ est alors équivalente à la formule $\exists x_i \phi$ et nous avons

$$S_{p,q}(\exists x_i \phi) = S_{p,q}(\exists x_{p+1} \phi[x_{p+1}/x_i]) = \pi_{p+1,q}\big(S_{p+1,q}(\phi[x_{p+1}/x_i])\big),$$

tant sur les mots finis que sur les mots infinis, par le lemme précédent. Ceci termine la preuve du Théorème III.2.2.

III.3 Quelques conséquences pour la logique

Nous abordons comme conclusion deux conséquences de l'équivalence expressive entre automates sur les mots infinis et formules de la logique monadique du second ordre sur les mots. Commençons par regrouper les énoncés des théorèmes III.2.1 et III.2.2 pour les mots infinis. Ce résultat est connu sous le nom de Théorème de Büchi.

Théorème III.3.1. *Soit Σ un alphabet fini. Pour un ω-langage $L \subseteq \Sigma^\omega$ les deux assertions sont équivalentes :*

1) *L est reconnaissable.*
2) *L est définissable dans MSO.*

Nous montrons tout d'abord que la logique MSO sur les mots infinis d'un alphabet Σ est décidable. Nous entendons qu'il existe un algorithme qui décide étant donné une formule close de MSO si ϕ est une tautologie ou non, i.e. si $\underline{\alpha} \models \phi$ pour tout $\alpha \in \Sigma^\omega$, ou non.

Théorème III.3.2. *Soit Σ un alphabet fini. La logique MSO sur les mots infinis sur l'alphabet Σ est décidable.*

Preuve. Soit ϕ une formule close de MSO. Par le Théorème III.2.2, il est possible de construire effectivement un automate non déterministe de Büchi $\mathfrak{A}$ qui reconnaisse $L(\neg\phi)$, i.e. tel que pour tout mot infini $\alpha \in \Sigma^\omega$,

$$\alpha \in L(\mathfrak{A}) \quad \Leftrightarrow \quad \underline{\alpha} \models \neg\phi.$$

La question de savoir si ϕ est une tautologie est alors équivalente à la question de savoir si le langage de $\mathfrak{A}$ est vide. Or, le problème de la vacuité d'un automate de Büchi est décidable. En effet soit un automate de Büchi $\mathfrak{A} = (Q, \Sigma, \Delta, q_I, F)$, il suffit de calculer l'ensemble des états f de F qui sont atteignables depuis l'état initial q_I et d'ensuite regarder pour chacun de ces f si f est atteignable depuis f par un chemin non vide. L'existence d'une boucle non vide pour un état acceptant f atteignable depuis l'état initial étant équivalente à l'existence d'un mot infini α pour lequel il existe un déroulement ρ de $\mathfrak{A}$ avec $f \in \mathrm{Inf}(\rho)$. □

Nous montrons maintenant que la logique MSO faible est équivalente à la logique MSO sur les mots infinis d'un alphabet Σ. Nous entendons que pour chaque formule close $\phi \in \mathscr{F}(\mathrm{MSO})$, il existe d'une part une formule close $\psi \in \mathscr{F}(\mathrm{MSO})$ telle que $L(\phi) = L_f(\psi)$ et que d'autre part il existe une formule close $\psi \in \mathscr{F}(\mathrm{MSO})$ telle que $L_f(\phi) = L(\psi)$.

Théorème III.3.3. *Soit Σ un alphabet fini. La logique MSO sur Σ^ω et la logique MSO faible sur Σ^ω sont équivalentes.*

Nous décomposons la démonstration de ce résultat en plusieurs lemmes. Tout d'abord, montrons qu'il est facile d'associer à chaque formule de MSO faible une formule de MSO qui lui soit équivalente.

Lemme III.3.4. *Toute formule close de MSO faible est équivalente à une formule close de MSO.*

Preuve. Observons tout d'abord que nous pouvons écrire dans MSO une formule $\mathrm{Fini}(X)$ qui exprime que l'ensemble représenté par la variable du second ordre X est fini. En effet, posons

$$\mathrm{Fini}(X) = \exists y (\forall x X(x) \to x < y).$$

Ainsi étant donné une formule ϕ de MSO faible, remplaçons chaque sous-formule de ϕ de la forme $\exists X\psi$ par $\exists X(\mathrm{Fini}(X) \wedge \psi)$. La formule obtenue est une formule de MSO équivalente à ϕ. □

Pour montrer que pour chaque formule close ϕ, il existe une formule close ψ telle que $L(\phi) = L_f(\phi)$, il suffit par le Théorème de Büchi (III.3.1) de montrer que tout ω-langage reconnaissable est définissable dans MSO faible. Rappelons que par le Théorème de McNaughton, tout ω-langage reconnaissable est une combinaison booléenne de ω-langages de la forme $\lim L$ avec $L \subseteq \Sigma^*$ régulier. Par le Théorème III.2.1, il existe une formule ϕ de MSO telle que

$$\underline{u} \models \phi \quad \Leftrightarrow \quad u \in L.$$

Comme MSO et MSO faible se confondent sur les mots finis, on a de façon équivalente

$$\underline{u} \models_f \phi \quad \Leftrightarrow \quad u \in L.$$

Nous montrons alors le lemme suivant.

Lemme III.3.5. *Soit ϕ une formule close de MSO. Il existe une formule $\phi(x)$ avec seule variable libre la variable du premier ordre x avec la propriété que pour tout mot infini α et tout $n \in \omega$:*

$$\underline{w}, x \mapsto n \models_f \phi(x) \quad \Leftrightarrow \quad \underline{w[0,n]} \models_f \phi.$$

Preuve. Soit $\psi \in \mathscr{F}(\mathrm{MSO})$ une formule pas forcément close. Nous définissons $\psi(x)$ par induction sur la hauteur des formules. Si ψ est une formule atomique, posons $\psi(x) = \psi$. Ensuite, nous posons

$$\begin{aligned}
(\neg\psi)(x) &= \neg\psi(x) \\
(\phi \vee \psi)(x) &= \phi(x) \vee \psi(x) \\
(\exists z \psi)(x) &= \exists z \left((z < x) \wedge \psi(x)\right) \\
(\exists Z \psi)(x) &= \exists Z \Big(\left(\forall z (Z(z) \rightarrow z < x)\right) \wedge \psi(x)\Big)
\end{aligned}$$

Dans le cas d'une formule close ϕ, $\phi(x)$ possède les propriétés désirées. □

Ce lemme nous permet de conclure la preuve du Théorème III.3.3.

Lemme III.3.6. *Soit $\phi \in \mathscr{F}(MSO)$ une formule close. Il existe une formule close $\psi \in \mathscr{F}(MSO)$ telle que $L(\phi) = L_f(\psi)$.*

Preuve. Comme nous l'avons déjà remarqué, par le Théorème de Büchi, $L(\phi)$ est reconnaissable et c'est une conséquence du Théorème de McNaughton et du Théorème I.4.7 que $L(\phi)$ s'écrit comme une combinaison booléenne de langages de la forme $\lim L$ pour $L \subseteq \Sigma^*$ régulier. Il suffit donc de montrer que tout langage de la forme $\lim L$ pour $L \in \mathrm{REG}(\Sigma)$ est définissable dans MSO faible. Or, par le Théorème III.2.1, un langage régulier $L \subseteq \Sigma^*$ est définissable dans MSO faible par une formule φ. Il découle alors du Lemme III.3.5 qu'il existe une formule de MSO $\varphi(x)$ telle que pour tout mot infini $\alpha \in \Sigma^\omega$ et tout naturel $n \in \omega$

$$\underline{\alpha}, x \mapsto n \models_f \varphi(x) \quad \Leftrightarrow \quad \underline{\alpha[0,n]} \models_f \varphi.$$

Par conséquent, le langage $\lim L$ est reconnu faiblement par la formule

$$\psi = \forall y \Big(\exists x \left((x > y) \wedge \varphi(x)\right)\Big),$$

i.e. $L_f(\psi) = \lim L$. □

BIBLIOGRAPHIE

[Grädel et al., 2002] Grädel, E., Thomas, W., and Wilke, T. (2002). *Automata Logics, and Infinite Games*. Springer.

[Perrin and Pin, 2004] Perrin, D. and Pin, J.-E. (2004). *Infinite words*. Elsevier. 1

[Sipser, 2006] Sipser, M. (2006). *Introduction to the Theory of Computation*. THOMSON, Course Thechnology, second edition, international edition. 9, 41

[Thomas, 1996] Thomas, W. (1996). Languages, automata, and logic.

[Thomas, 2003] Thomas, W. (2003). Automata and reactive systems.