Mathématiques de l'informatique

(durée 6 heures)

Ce problème a pour objet l'étude de quelques aspects de la théorie des traces qui est un des modèles les plus étudiés du parallélisme.

Dans tout le problème, A désigne un alphabet fini. Un mot u sur A est une suite finie $(a_1, \ldots, a_n)$. Ce mot sera noté simplement $u = a_1 \ldots a_n$ si aucune confusion n'est possible. L'entier n est la longueur du mot u, notée $|u|$. Le mot vide, noté 1, est l'unique mot de longueur nulle. Pour toute lettre a de A, le nombre d'occurrences de la lettre a dans le mot u est noté $|u|_a$. Le sous-ensemble de A constitué des lettres ayant au moins une occurrence dans le mot u est appelé alphabet du mot u et est noté $\mathrm{alph}(u)$. L'ensemble de tous les mots sur un alphabet A muni de l'opération binaire de concaténation (i.e. de "mise bout à bout") a une structure de monoïde dont l'élément neutre est le mot vide. C'est le monoïde libre engendré par A, noté A^*. La concaténation de $u, v \in A^*$ est notée simplement uv.

On munit A d'une relation $I \subseteq A \times A$ anti-réflexive et symétrique, dite relation d'indépendance. Le complémentaire de I est la relation de dépendance $D = (A \times A) \setminus I$. Intuitivement deux actions indépendantes peuvent être exécutées simultanément par des processeurs distincts. Deux sous-ensembles B et C de A sont dits indépendants si $B \times C \subseteq I$. Deux mots $u, v \in A^*$ sont dits indépendants si leurs alphabets $\mathrm{alph}(u)$ et $\mathrm{alph}(v)$ le sont. Par abus de notation, on note alors simplement $(u, v) \in I$.

La relation I induit une relation binaire R sur A^* définie par :

$$R = \{(uabv, ubav) \mid u, v \in A^*, (a, b) \in I\}$$

La clôture réflexive et transitive de la relation R est notée $\sim_I$ ou simplement $\sim$. Deux mots u, v de A^* vérifiant $u \sim_I v$ sont dits $\sim_I$-équivalents ou simplement équivalents.

Exemple 1 : Soit $A = \{a, b, c, d\}$ et $I = \{(a, b), (b, a), (b, c), (c, b)\}$. Nous avons alors $aabcbcd \sim_I baaccbd \sim_I aabbccd$. Les mots $acaa$ et bb sont indépendants, par contre les mots $acaa$ et bd ne le sont pas.

On note $\mathbb{M}$(A,I) le quotient du monoïde libre A^* par la relation d'équivalence $\sim_I$. Un élément de $\mathbb{M}$(A,I) est appelé trace et $\mathbb{M}$(A,I) est le monoïde des traces engendré par A et I.

1 Combinatoire

1. Soient u, v deux mots équivalents. Montrer que $|u|_a = |v|_a$ pour toute lettre a de A et que $\mathrm{alph}(u) = \mathrm{alph}(v)$ et $|u| = |v|$.

2. Si $B \subseteq A$, on note Π_B la projection de A^* sur B^* effaçant toutes les lettres n'appartenant pas à B. Montrer que deux mots sont équivalents si et seulement si $\Pi_{\{a,b\}}(u) = \Pi_{\{a,b\}}(v)$ pour tout couple $(a, b) \in D$.

3. En déduire un algorithme linéaire en la longueur du mot pour décider si deux mots sont équivalents.

4. Montrer que si u, v, x et y sont des mots vérifiant $xuy \sim_I xvy$ alors $u \sim_I v$.

5. Soit u un mot. Une lettre a est dite minimale dans u, on note alors $a \in \min(u)$, si $u = u'au''$ avec $(a, u') \in I$. Montrer que $a \in \min(u)$ si et seulement si :

$$\forall b \in A, (a, b) \in D \Longrightarrow \Pi_{\{a,b\}}(u) \text{ commence par la lettre } a$$

6. Déterminer un algorithme qui énumère tous les mots équivalents à un mot donné. Quelle est la complexité de cet algorithme dans le pire des cas?

2 Formes normales

On munit maintenant l'alphabet A d'une relation d'ordre total $<$. Cette relation définit une relation d'ordre total sur A^* (notée encore $<$ par abus de notation), dit ordre lexicographique, par $u < v$ si ou bien $v = uw$ avec $w \neq 1$ ou bien $u = xay, v = xbz$ avec $a, b \in A$ et $a < b$. Un mot u est dit sous forme normale lexicographique (fnl en abrégé) s'il est plus petit que tout autre mot équivalent :

$$\forall v \in A^*, u \sim_I v \Longrightarrow u = v \text{ ou } u < v$$

1. Déterminer un algorithme linéaire en la longueur du mot calculant l'unique mot sous fnl équivalent à un mot donné.

Un mot non vide u est dit sous forme normale de Foata (fnf en abrégé) si $u = u_1 \dots u_n$ avec, pour tout i dans $\{1, \dots, n\}$:

F0 $\forall a \in \mathrm{alph}(u_i), |u_i|_a = 1$

F1 $\forall a, b \in \mathrm{alph}(u_i), a = b$ ou $(a, b) \in I$

F2 u_i est un mot sous fnl

F3 si $i \geq 2$, $\forall b \in \mathrm{alph}(u_i)$, $\exists a \in \mathrm{alph}(u_{i-1})$ tel que $(a, b) \in D$

2. Un mot sous fnl est-il toujours sous fnf? Et réciproquement?

3. Montrer qu'il existe un unique mot sous fnf équivalent à un mot donné.

4. Déterminer un algorithme linéaire en la longueur du mot calculant l'unique mot sous fnf équivalent à un mot donné.

3 Reconnaissables

Soit M un monoïde d'élément neutre 1_M.

Un sous-ensemble $X \subseteq M$ est dit reconnaissable s'il existe un morphisme η de M dans un monoïde fini S et un sous-ensemble $F \subseteq S$ tels que $X = \eta^{-1}(F)$. On dira aussi que le morphisme η reconnaît X. L'ensemble des sous-ensembles reconnaissables de M est noté $\mathrm{Rec}(M)$.

La concaténation de $K, L \subseteq M$ est définie par :

$$KL = \{kl \mid k \in K, l \in L\}$$

et l'itération de $K \subseteq M$ par :

$$K^* = \bigcup_{n \geq 0} K^n \text{ où } K^0 = \{1_M\} \text{ et } K^{n+1} = K^n K \text{ pour tout } n \geq 0$$

L'ensemble des parties rationnelles de M, noté $\mathrm{Rat}(M)$, est la plus petite famille de parties contenant les parties finies et close par union, concaténation et itération.

Le théorème de Kleene assure que dans le monoïde libre, les langages rationnels et les langages reconnaissables coïncident : $\mathrm{Rec}(A^*) = \mathrm{Rat}(A^*)$.

Cette partie du problème et la suivante ont pour objet d'établir un équivalent de ce théorème pour les monoïdes de traces.

D'après la question 1.1 on peut définir sans ambiguité l'alphabet d'une trace et son nombre d'occurrences d'une lettre donnée, on étend alors les notations alph( ), $|\ |$ et $|\ |_a$ aux éléments de $\mathbb{M}$(A,I). Si $u, v \in \mathbb{M}(A, I)$ sont deux traces, on notera encore $(u, v) \in I$ si $\mathrm{alph}(u)$ et $\mathrm{alph}(v)$ sont indépendants.

Une trace $t \in \mathbb{M}(A, I)$ est dite connexe si la relation de dépendance D restreinte à $\mathrm{alph}(t)$, c'est à dire le graphe $(\mathrm{alph}(t), D \cap (\mathrm{alph}(t) \times \mathrm{alph}(t)))$, est un graphe connexe. La notion de mot connexe est définie de manière analogue. Soit φ la surjection canonique de A^* sur $\mathbb{M}$(A,I). Si l'on reprend l'exemple 1 ci-dessus, les traces $\varphi(abccbd)$ et $\varphi(bdd)$ sont connexes alors que $\varphi(abccb)$ et $\varphi(cbb)$ ne le sont pas.

Une composante connexe d'une trace t est une trace connexe non vide u telle que $t = uv$ avec $(u, v) \in I$. Si $T \subseteq \mathbb{M}(A, I)$ est un langage de traces, on note $\mathrm{c}(T)$ l'ensemble de toutes les composantes connexes des éléments de T.

Sur $\mathbb{M}(A, I)$, l'itération concurrente c-$*$ est alors définie pour tout langage de traces $T \subseteq \mathbb{M}(A, I)$, par $T^{c-*} = (c(T))^*$. L'ensemble des langages $c-$rationnels de $\mathbb{M}$(A,I), noté $\mathrm{c-Rat}(\mathbb{M}(A, I))$, est la plus petite famille de langages de $\mathbb{M}$(A,I) contenant les langages finis et close par les opérations d'union, concaténation et itération concurrente.

1. Soit $T \subseteq \mathbb{M}(A, I)$ un langage de traces. Si $u \in \mathbb{M}(A, I)$ est une trace, on note $u^{-1}T$ l'ensemble $\{v \in \mathbb{M}(A, I) \mid uv \in T\}$. Montrer que les trois assertions suivantes sont équivalentes.

   (i) Le langage T est reconnaissable.

   (ii) L'ensemble $\varphi^{-1}(T)$ est un langage reconnaissable de A^*.

   (iii) La famille $\{u^{-1}T \mid u \in \mathbb{M}(A, I)\}$ est finie.

2. Un morphisme η défini sur $\mathbb{M}$(A,I) est dit alphabétique si $\text{alph}(u) = \text{alph}(v)$ dès que $\eta(u) = \eta(v)$. Montrer qu'un langage de traces est reconnaissable si et seulement si il est reconnu par un morphisme alphabétique.

3. Montrer que si T est un langage reconnaissable, le langage $\text{c}(T)$ est encore reconnaissable.

4. On note FNL l'ensemble de tous les mots de A^* sous forme normale lexicographique (cf partie 2). Montrer que FNL est un langage reconnaissable de A^*.

5. Soit w un mot de FNL tel que w^2 soit aussi dans FNL. Montrer que w est connexe.

6. Montrer que $\text{Rec}(\mathbb{M}(A, I)) \subseteq \text{c}-\text{Rat}(\mathbb{M}(A, I))$.

4 Rang

On veut maintenant démontrer réciproquement que tout langage de $\text{c}-\text{Rat}(\mathbb{M}(A, I))$ est dans $\text{Rec}(\mathbb{M}(A, I))$.

1. Soient $x, y, u_0, \ldots, u_n$ $n + 3$ mots de A^*. Montrer que les assertions suivantes sont équivalentes.

   (i) $xy \sim u_0 \ldots u_n$.

   (ii) Il existe des mots $x_0, \ldots, x_n$, $y_0, \ldots, y_n$ tels que :

   - $x \sim x_0 \ldots x_n$, $y \sim y_0 \ldots y_n$
   - $u_i \sim x_i y_i$ pour tout $0 \leq i \leq n$
   - $(y_i, x_{i+1} \ldots x_n) \in I$ pour tout $0 \leq i < n$

Soient $x, y \in A^*$ et L un langage de A^*. D'après la question précédente, le produit xy est équivalent à un élément de L s'il existe des mots $x_0, y_0, \ldots, x_n, y_n$ de A^* tels que $x \sim_I x_0 \ldots x_n$, $y \sim_I y_0 \ldots y_n$, $x_0 y_0 x_1 y_1 \ldots x_n y_n \in L$ et $(y_i, x_{i+1} \ldots x_n) \in I$ pour tout $0 \leq i < n$. On définit alors $\text{rang}(x, y, L)$ comme étant le plus petit entier n pour lequel une telle factorisation est possible. Le rang (fini ou infini) d'un langage $L \subseteq A^*$ est alors

$$\text{rang}(L) = \max\{\text{rang}(x, y, L) \mid x, y \in A^*, xy \in \varphi^{-1}(\varphi(L))\}$$

Un langage $L \subseteq A^*$ est dit I-clos (ou simplement clos) si tout mot équivalent à un mot de L est encore dans L.

2. On considère l'alphabet $A = \{a, b\}$ et la relation d'indépendance $I = \{(a, b), (b, a)\}$. Calculer $\text{rang}(\{ab\}^*)$ et $\text{rang}(\{a\}^*\{b\}^*)$.

3. Soient $L_1, L_2 \subseteq A^*$ deux langages I-clos. Montrer que $\text{rang}(L_1L_2) \leq 1$.

4. Soit $L \subseteq A^*$ un langage I-clos connexe. Montrer que $\text{rang}(L^*) < 2|A|$.

5. Soit $L \subseteq A^*$ un langage reconnaissable de rang fini. Montrer que $\varphi(L)$ est un langage reconnaissable de traces.

6. Que dire de la réciproque de la question précédente?

7. Montrer que $\text{c}-\text{Rat}(\mathbb{M}(A, I)) \subseteq \text{Rec}(\mathbb{M}(A, I))$.

5 Graphes de dépendance

On note ω le cardinal de l'ensemble $\mathbb{N}$ des entiers naturels. L'addition usuelle dans $\mathbb{N}$ est étendue en posant $\omega + n = n + \omega = \omega$ pour tout $n \in \mathbb{N} \cup \{\omega\}$. Pour $a \in A$ et $n \in \mathbb{N} \cup \{\omega\}$, on définit l'ensemble des sommets $V_a(n) = \{(a, i) \mid i < n\}$.

Un graphe de dépendance sur (A, I) est une paire $g = ((n_a)_{a \in A}, E)$ où $n_a \in \mathbb{N} \cup \{\omega\}$ pour tout $a \in A$ et $E \subseteq V \times V$ avec $V = \bigcup_{a \in A} V_a(n_a)$ vérifiant les trois conditions :

C1 E est sans cycle : $\forall v \in V, (v, v) \notin E^+$ (où E^+ désigne la clôture transitive de E)

C2 $\forall a, b \in A, \forall x \in V_a(n_a), \forall y \in V_b(n_b)$:

$$(a, b) \in D \iff ((x, y) \in E \text{ ou } (y, x) \in E \text{ ou } x = y)$$

C3 $\forall a \in A$, $\forall (a,i),(a,j) \in V_a(n_a)$:

$$((a,i),(a,j)) \in E \Longleftrightarrow i < j$$

L'ensemble des graphes de dépendance sur (A,I) est noté $\mathbb{G}$(A,I). Le sous-ensemble de $\mathbb{G}$(A,I) constitué des graphes ayant un nombre fini de sommets est noté $\mathbb{F}$(A,I). Pour $g = ((n_a)_{a\in A}, E_g) \in \mathbb{G}(A,I)$, on définit $\mathrm{alph}(g) = \{a \in A \mid n_a \neq 0\}$ et $\mathrm{alphinf}(g) = \{a \in A \mid n_a = \omega\}$.

On définit une opération de concaténation partielle dans $\mathbb{G}$(A,I) de la manière suivante. Soient $g = ((n_a)_{a\in A}, E_g) \in \mathbb{G}(A,I)$ et $h = ((m_a)_{a\in A}, E_h) \in \mathbb{G}(A,I)$ avec $\mathrm{alphinf}(g)\times\mathrm{alph}(h) \subseteq I$ (cette condition est trivialement vérifiée si $g \in \mathbb{F}(A,I)$), la concaténation de g et h, notée $g \cdot h$ est le graphe de dépendance $f = ((n_a + m_a)_{a\in A}, E_f)$ avec

$$\begin{aligned} E_f = E_g \ & \bigcup \{((a, n_a + i), (b, n_b + j)) \mid ((a,i),(b,j)) \in E_h\} \\ & \bigcup \{((a,i),(b, n_b + j)) \mid (a,b) \in D, i < n_a, j < m_b\} \end{aligned}$$

1. Montrer que $\mathbb{F}$(A,I) muni de la concaténation $(\cdot)$ est isomorphe à $\mathbb{M}$(A,I).

Soient $g, f \in \mathbb{G}(A,I)$, g est un préfixe de f s'il existe $h \in \mathbb{G}(A,I)$ tel que $f = g \cdot h$. La relation "être préfixe de" est une relation d'ordre partiel sur $\mathbb{G}$(A,I) notée $\leq$.

Soit $(Z, \leq)$ un ensemble partiellement ordonné. Un sous-ensemble $Y \subseteq Z$ est cohérent si pour tout $x, y \in Y$ il existe $z \in Z$ tel que $x \leq z$ et $y \leq z$. Un sous-ensemble non vide $Y \subseteq Z$ est dirigé si pour tout $x, y \in Y$ il existe $z \in Y$ tel que $x \leq z$ et $y \leq z$. Un élément $x \in Z$ est premier (resp. compact) si pour tout sous-ensemble (resp. pour tout sous ensemble dirigé) $Y \subseteq Z$ admettant une borne supérieure $\sqcup Y$, $x \leq \sqcup Y$ entraine l'existence d'un élément $y \in Y$ tel que $x \leq y$.

2. Montrer qu'un ensemble $K \subseteq \mathbb{G}(A,I)$ est cohérent si et seulement si il admet une borne supérieure.

3. Caractériser les éléments premiers de $\mathbb{G}(A,I)$ et montrer que tout graphe de dépendance est la borne supérieure de l'ensemble des éléments premiers qui lui sont inférieurs ou égaux.

4. Caractériser les éléments compacts de $\mathbb{G}(A,I)$ et montrer que tout graphe de dépendance est la borne supérieure de l'ensemble des éléments compacts qui lui sont inférieurs ou égaux.

5. Montrer que tout élément de $\mathbb{G}(A, I)$ est la borne supérieure d'une suite croissante d'éléments de $\mathbb{F}(A, I)$.

6 Equations

1. Soit M un monoïde d'élément neutre 1_M. Une longueur est un morphisme γ de M dans $\mathbb{N}$ telle que $\gamma^{-1}(0) = \{1_M\}$. Un monoïde M est finiment engendré s'il existe K fini inclus dans M tel que tout élément de M est produit d'éléments de K. Montrer qu'un monoïde finiment engendré M muni d'une longueur est isomorphe à un monoïde de traces si et seulement si, pour tout $x, y, z, t \in M$, l'égalité $xy = zt$ entraîne l'existence de $p, q, r, s \in M$ tels que $x = pq$, $y = rs$, $z = pr$, $t = qs$ et $qr = rq$.

2. Soit M un monoïde. Deux éléments u et v de M sont dits conjugués s'il existe w dans M tel que $uw = wv$. Les éléments u et v sont dits transposés s'il existe x, y dans M tels que $u = xy$ et $v = yx$.

   (a) Montrer que les notions de conjugaison et de transposition coïncident si M est un monoïde libre.

   (b) Donner un exemple de monoïde de traces où ces deux notions sont distinctes.

   (c) Montrer que dans un monoïde de traces, la conjugaison est la clôture transitive de la relation de transposition.

Soit $X = \{x, y\}$ avec $x \neq y$ un alphabet de 2 lettres. Une équation $e = f$ sur X est une paire de mots distincts $e, f \in X^*$. Une solution de $e = f$ dans $\mathbb{M}(A, I)$ est un morphisme $\theta : X^* \longrightarrow \mathbb{M}(A, I)$ telle que $\theta(e) = \theta(f)$. Si une solution θ est définie par $\theta(x) = u$ et $\theta(y) = v$, on écrit simplement que (u, v) est une solution de $e = f$ (dans $\mathbb{M}(A, I)$). Une solution (u, v) est dite cyclique s'il existe une trace t telle que $u, v \in t^*$ (i.e. sont des puissances de t).

3. Soit (u, v) une solution de $e = f$ dans $\mathbb{M}(A, I)$. Montrer que (u, v) est cyclique dès que l'une des deux conditions suivantes est satisfaite :

   (i) $\mathbb{M}$(A,I) est libre c'est à dire $I = \emptyset$.

(ii) Il existe des nombres entiers $\alpha, \beta > 0$ tels que $\alpha|u|_a = \beta|v|_a$ pour toute lettre a de A.

4. Soit (u, v) une solution de l'équation $e = f$. Soit u' (resp. v') une composante connexe de u (resp. de v). Montrer que ou bien $\mathrm{alph}(u') = \mathrm{alph}(v')$ ou bien $(u', v') \in I$.

5. Soit (u, v) une solution de l'équation $e = f$. Montrer qu'il existe des traces deux à deux indépendantes et connexes $t_1, \ldots, t_k$ telles que $u, v \in t_1^* \ldots t_k^*$.