Magistère d'Informatique et Modélisation

Devoir de Logique

À remettre le 5 octobre 1998

L'objectif du devoir est de définir trois concepts fondamentaux de la logique et de l'informatique, à savoir les mots, les arbres étiquetés et les termes.

Partie I : (Mots)

Un *alphabet* est un ensemble, dont les éléments sont appelés «lettres». Un *mot* sur A est un suite finie de lettres. Si $A = \{a, b, c\}$, $aaba$, a, acb sont des mots. La suite vide est un mot particulier, appelé aussi le *mot vide* et noté ε. Si $\alpha = a_1 ... a_m$ et $\beta = b_1 ... b_n$ sont deux mots, on peut créer un nouveau mot par concaténation, on le note alors $\alpha\beta$ (ou $\alpha \cdot \beta$) ; c'est le mot (la suite) $a_1 ... a_m b_1 ... b_n$. L'ensemble des mots sur A est noté A^*.

1. Quelle est la structure de $(A^*, \cdot, \varepsilon)$?
2. Comment définiriez-vous un morphisme de monoïde ? Comparez cette définition avec celle d'un morphisme de groupes qui sont pourtant des cas particuliers de monoïde.
3. Montrez que si $(M, *, 1)$ est un monoïde quelconque et si il existe une application $f : A \to M$, alors il existe un unique morphisme de monoïde $\xi_M : A^* \to M$ qui prolonge f (c-à-d tel que pour tout $a \in A$, $\xi_M(a) = f(a)$). Pour cette raison on dit que $(A^*, \cdot, \varepsilon)$ est un (et par abus de langage «le») monoïde libre sur A. Pourquoi l'abus de langage assimilant «un» à «le» est-il tolérable ?
4. Un sous-ensemble de A^* est appelé un langage sur A. On étend la concaténation aux langages de A^*. Si L et L' sont des langages sur A,

   $$L \; L' = \{\alpha\beta \mid \alpha \in L \text{ et } \beta \in L'\}.$$

   On définit l'opération L^* sur les langages par

   $$L^* = \bigcup_{n=0}^{\infty} L^n$$

   où $L^0 = \{\varepsilon\}$ et $L^{n+1} = L \; L^n$. Montrez que L^* est le plus petit langage de A^* qui satisfait l'équation $L^* = \{\varepsilon\} \cup L \; L^*$. Un langage L est dit *régulier* si il peut être décrit à partir de langages à un élément à l'aide de la concaténation de l'union et de l'opération *.

5. Décrivez à l'aide de la concaténation, de l'étoile et des langages finis, le langage N de $\{a,b\}^*$ des mots qui contiennent au moins une fois la séquence ab. Décrivez son complémentaire dans $\{a,b\}^*$.
6. Soit $\beta = b_1 \ldots b_n$. α est un préfixe de β ce qui s'écrit $\alpha \,\underline{\mathsf{prefix}}\, \beta$ s'il existe m, $m \leq n$ tel que $\alpha = b_1 \ldots b_m$. Montrez que pour tout mot α, $\varepsilon \,\underline{\mathsf{prefix}}\, \alpha$. Donnez quelques propriétés de la relation $\underline{\mathsf{prefix}}$. Si p n'est pas préfixe de q et si q n'est pas préfixe de p, p et q sont dits *étrangers*. Montrez que p est étranger à q si et seulement si pour tout r, p n'est pas préfixe de q.
7. On peut montrer qu'un langage L est régulier si $L = \xi_M^{-1}(M')$ où M, M' et ξ_M sont définis ainsi : $(M, \star, 1)$ est un monoïde fini, pour chaque $a \in A$, a_M est l'image dans M de l'élément a de A, M' est un sous-ensemble fini de M et ξ_M est l'unique morphisme de monoïde de $(A^*, \cdot, \varepsilon)$ vers $(M, \star, 1)$ tel que $\xi_M(a) = a_M$.

   Soit $A = \{a, b\}$ et si $M = \{1, \mathsf{a}, \mathsf{b}, \mathsf{ba}, \omega\}$ avec $a_M = \mathsf{a}$, $b_M = \mathsf{b}$ et

$\star$	1	a	b	ba	ω
1	1	a	b	ba	ω
a	a	a	ω	ω	ω
b	b	ba	b	ba	ω
ba	ba	ba	ω	ω	ω
ω	ω	ω	ω	ω	ω

   Admettant que $(M, \star, 1)$ est un monoïde, quel est le langage régulier défini en prenant pour M' le sous-ensemble $\{1, \mathsf{a}, \mathsf{b}, \mathsf{ba}\}$? Même question en prenant pour M' le sous-ensemble $\{\omega\}$?

1 Partie II (Arbres étiquetés)

$\mathbb{N}$ est l'ensemble des naturels et $\mathbb{N}^+$ l'ensemble des naturels privé de zéro. $(\mathbb{N}^+)^*$ est l'ensemble des mots sur $\mathbb{N}^+$.

Le *domaine* d'un arbre étiqueté est un sous-ensemble fini non vide D de $(\mathbb{N}^+)^*$ tel que si $\alpha \in D$ et $\beta \,\underline{\mathsf{prefix}}\, \alpha$ alors $\beta \in D$ et si $\alpha\ (n+1) \in D$ alors $\alpha\ n \in D$. Un arbre étiqueté par A est un application $T : D_T \to A$ où D_T est un domaine. Dans la suite, «arbre» signifiera «arbre étiqueté». D_T est le domaine de T ; on le note généralement $Dom(T)$. Si $p \in Dom(T)$, $T(p)$ s'appelle l'étiquette à la position p.

1. Montrez que tout domaine contient le mot vide. $T(\varepsilon)$ s'appelle la *racine* de T.
2. TT est la représentation d'un arbre sous forme d'un dessin (voir figure 1). $A = \{+, x, y, z, t\}$ et $Dom(TT) = \{\varepsilon, 1, 2, 21, 22, 221, 222\}$. On a $TT(\varepsilon) = +$, $TT(1) = x$, $TT(21) = y$, $TT(22) = +$. Quelles sont les valeurs que vous donneriez à $TT(2)$, $TT(221)$ et $TT(222)$?
3. $T_{|p}$ est le *sous-arbre à la position p* défini par $T_{|p}(q) = T(pq)$ où $Dom(T_{|p}) = \{q \mid pq \in Dom(T)\}$. Décrivez $Dom(TT_{|22})$ et $TT_{|22}$.

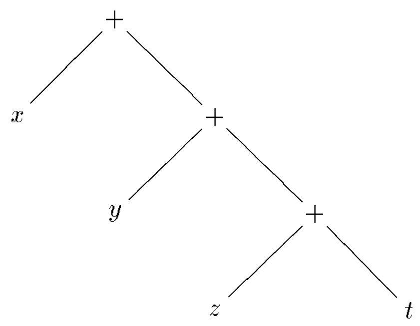

FIG. 1: Le dessin de l'arbre TT

4. Montrez que $(T_{|p})_{|q} = T_{|pq}$.
5. $T[U]_p$ est le *remplacement* du sous-arbre de T à la position p par l'arbre U, on le définit par

$$Dom(T[U]_p) = \{q \in Dom(T) \mid \neg(p\ \underline{\mathsf{prefix}}\ q)\} \cup \{pp' \mid p' \in Dom(U)\}$$

et

$$T[U]_p(q) = \begin{cases} T(q) \text{ si } \neg(p\ \underline{\mathsf{prefix}}\ q) \\ U(p') \text{ si } q = pp' \text{ avec } p' \in Dom(U) \end{cases}$$

Soit $UU = TT[TT]_{21}$. Que vaut $UU(212)$. Dessinez UU.
6. Montrez que $T[U[V]_q]_p = T[U]_p[V]_{pq}$.
7. Sous quelle condition a-t-on $T[U]_p[V]_q = T[V]_q[U]_p$?
8. Montrez que si p et q sont étrangers, alors $(T[U]_p)_{|q} = T_{|q}$.
9. Montrez que $(T[U]_p)_{|pp'} = U_{|p}$.
10. Quelles sont les valeurs r dans $(T[U]_p)_{|r}$ non atteintes par les deux questions précédentes. À la manière des deux questions précédentes quel est le membre droit d'une identité où vous auriez comme membre gauche $(T[U]_p)_{|r}$? Prouvez l'identité ainsi trouvée.
11. Soient $a \in A$ et une suite d'arbres $T_1, ..., T_n$. L'arbre $a(T_1, \ldots, T_n)$ est défini par

$$Dom(a(T_1, \ldots, T_n)) = \{\varepsilon\} \cup 1\, Dom(T_1) \cup \ldots \cup n\, Dom(T_n)\}$$

et $a(T_1, \ldots, T_n)(\varepsilon) = a$ et si $p \in Dom(T_i)$, $a(T_1, \ldots, T_n)(ip) = T_i(p)$. Définissez $T_{|p}$ et $T[U]_p$ en vous sevrant de cette nouvelle opération.

Part III (Termes)

Un *alphabet gradué* (ou *signature*) est la donnée d'un alphabet et d'une fonction $ar : A \to \mathbb{N}$, dite *arité*. En mettant de restrictions adéquates sur le arbres, définissez ce que serait, pour vous, un terme. Montrez qu'alors les opérations définies dans la partie II préservent la structure de terme.