N° d'ordre : 1176

UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

THÈSE

Présentée à

L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

pour obtenir le titre de

DOCTEUR DE 3ème CYCLE

(Informatique)

par

Erick TIMMERMAN

FEUILLAGES D'ARBRES INFINIS.

Thèse soutenue le 22 juin 1984 devant la commission d'examen

Membres du Jury :

Président :	G. COMYN
Rapporteur :	M. DAUCHET
Examinateur :	A. ARNOLD
	C. CARREZ
	B. COURCELLE
	M. LATTEUX

P R O F E S S E U R S CLASSE EXCEPTIONNELLE

M. CONSTANT Eugène	I.E.E.A.
M. FOURET René	Physique
M. GABILLARD Robert	I.E.E.A.
M. MONTREUIL Jean	Biologie
M. PARREAU Michel	Mathématiques
M. TRIDOT Gabriel	Chimie
M. VIVIER Emile	Biologie
M. WERTHEIMER Raymond	Physique

P R O F E S S E U R S 1ère classe

M. BACCHUS Pierre	Mathématiques
M. BEAUFILS Jean-Pierre (dét.)	Chimie
M. BIAYS Pierre	G.A.S.
M. BILLARD Jean (dét.)	Physique
M. BOILLY Bénoni	Biologie
M. BOIS Pierre	Mathématiques
M. BONNELLE Jean-Pierre	Chimie
M. BOUGHON Pierre	Mathématiques
M. BOURIQUET Robert	Biologie
M. BREZINSKI Claude	I.E.E.A.
M. CELET Paul	Sciences de la Terre
M. CHAMLEY Hervé	Biologie
M. COEURE Gérard	Mathématiques
M. CORDONNIER Vincent	I.E.E.A.
M. DEBOURSE Jean-Pierre	S.E.S.
M. DYMENT Arthur	Mathématiques

PROFESSEURS 1ère classe (suite)

M.	ESCAIG Bertrand	Physique
M.	FAURE Robert	Mathématiques
M.	FOCT Jacques	Chimie
M.	GRANELLE Jean-Jacques	S.E.S.
M.	GRUSON Laurent	Mathématiques
M.	GUILLAUME Jean	Biologie
M.	HECTOR Joseph	Mathématiques
M.	LABLACHE COMBIER Alain	Chimie
M.	LACOSTE Louis	Biologie
M.	LAVEINE Jean Pierre	Sciences de la Terre
M.	LEHMANN Daniel	Mathématiques
Mme	LENOBLE Jacqueline	Physique
M.	LHOMME Jean	Chimie
M.	LOMBARD Jacques	S.E.S.
M.	LOUCHEUX Claude	Chimie
M.	LUCQUIN Michel	Chimie
M.	MIGEON Michel Recteur à Grenoble	E.U.D.I.L.
M.	MIGNOT Fulbert (dét.)	Mathématiques
M.	PAQUET Jacques	Sciences de la Terre
M.	PROUVOST Jean	Sciences de la Terre
M.	ROUSSEAU Jean-Paul	Biologie
M.	SALMER Georges	I.E.E.A.
M.	SEGUIER Guy	I.E.E.A.
M.	SIMON Michel	S.E.S.
M.	STANKIEWICZ François	S.E.S.
M.	TILLIEU Jacques	Physique
M.	VIDAL Pierre	I.E.E.A.
M.	ZEYTOUNIAN Radyadour	Mathématiques

P R O F E S S E U R S 2ème classe

M. ANTOINE Philippe	Mathématiques (Calais)
M. BART André	Biologie
Mme BATTIAU Yvonne	Géographie
M. BEGUIN Paul	Mathématiques
M. BELLET Jean	Physique
M. BERZIN Robert	Mathématiques
M. BKOUCHE Rudolphe	Mathématiques
M. BODARD Marcel	Biologie
M. BOSQ Denis	Mathématiques
M. BRASSELET Jean-Paul	Mathématiques
M. BRUYELLE Pierre	Géographie
M. CAPURON Alfred	Biologie
M. CARREZ Christian	I.E.E.A.
M. CAYATTE Jean-Louis	S.E.S.
M. CHAPOTON Alain	C.U.E.E.P.
M. COQUERY Jean-Marie	Biologie
Mme CORSIN Paule	Sciences de la Terre
M. CORTOIS Jean	Physique
M. COUTURIER Daniel	Chimie
M. CROSNIER Yves	I.E.E.A.
M. CURGY Jean-Jacques	Biologie
Mle DACHARRY Monique	Géographie
M. DAUCHET Max	I.E.E.A.
M. DEBRABANT Pierre	E.U.D.I.L.
M. DEGAUQUE Pierre	I.E.E.A.
M. DELORME Pierre	Biologie
M. DELORME Robert	S.E.S.
M. DE MASSON D'AUTUME Antoine	S.E.S.
M. DEMUNTER Paul	C.U.E.E.P.

PROFESSEURS 2ème classe (Suite 1)

M. DENEL Jacques	I.E.E.A.
M. DE PARIS Jean-Claude	Mathématiques (Calais)
Mle DESSAUX Odile	Chimie
M. DEVRAINNE Pierre	Chimie
M. DHAINAUT André	Biologie
Mme DHAINAUT Nicole	Biologie
M. DORMARD Serge	S.E.S.
M. DOUKHAN Jean-Claude	E.U.D.I.L.
M. DUBOIS Henri	Physique
M. DUBRULLE Alain	Physique (Calais)
M. DUBUS Jean-Paul	I.E.E.A.
M. FAKIR Sabah	Mathématiques
M. FONTAINE Hubert	Physique
M. FOUQUART Yves	Physique
M. FRONTIER Serge	Biologie
M. GAMBLIN André	G.A.S.
M. GLORIEUX Pierre	Physique
M. GOBLOT Rémi	Mathématiques
M. GOSSELIN Gabriel (dét.)	S.E.S.
M. GOUDMAND Pierre	Chimie
M. GREGORY Pierre	I.P.A.
M. GREMY Jean-Paul	S.E.S.
M. GREVET Patrice	S.E.S.
M. GUILBAULT Pierre	Biologie
M. HENRY Jean-Pierre	E.U.D.I.L.
M. HERMAN Maurice	Physique
M. JACOB Gérard	I.E.E.A.
M. JACOB Pierre	Mathématiques
M. JEAN Raymond	Biologie
M. JOFFRE Patrick	I.P.A.

PROFESSEURS 2ème classe (suite 2)

M. JOURNEL Gérard	E.U.D.I.L.
M. KREMBEL Jean	Biologie
M. LANGRAND Claude	Mathématiques
M. LATTEUX Michel	I.E.E.A.
Mme LECLERCQ Ginette	Chimie
M. LEFEVRE Christian	Sciences de la Terre
Mle LEGRAND Denise	Mathématiques
Mle LEGRAND Solange	Mathématiques (Calais)
Mme LEHMANN Josiane	Mathématiques
M. LEMAIRE Jean	Physique
M. LHENAFF René	Géographie
M. LOCQUENEUX Robert	Physique
M. LOSFELD Joseph	C.U.E.E.P.
M. LOUAGE Francis(dét.)	E.U.D.I.L.
M. MACKE Bruno	Physique
M. MAIZIERES Christian	I.E.E.A.
M. MESSELYN Jean	Physique
M. MESSERLIN Patrick	S.E.S.
M. MONTEL Marc	Physique
Mme MOUNIER Yvonne	Biologie
M. PARSY Fernand	Mathématiques
Mle PAUPARDIN Colette	Biologie
M. PERROT Pierre	Chimie
M. PERTUZON Emile	Biologie
M. PONSOLLE Louis	Chimie
M. PORCHET Maurice	Biologie
M. POVY Lucien	E.U.D.I.L.
M. RACZY Ladislas	I.E.E.A.
M. RAOULT Jean François	Sciences de la Terre
M. RICHARD Alain	Biologie

PROFESSEURS 2ème Classe (suite 3)

M.	RIETSCH François	E.U.D.I.L.
M.	ROBINET Jean-Claude	E.U.D.I.L.
M.	ROGALSKI Marc	Mathématiques
M.	ROY Jean-Claude	Biologie
M.	SCHAMPS Joël	Physique
Mme	SCHWARZBACH Yvette	Mathématiques
M.	SLIWA Henri	Chimie
M.	SOMME Jean	G.A.S.
Mle	SPIK Geneviève	Biologie
M.	STAROSWIECKI Marcel	E.U.D.I.L.
M.	STERBOUL François	E.U.D.I.L.
M.	TAILLIEZ Roger	Institut Agricole
Mme	TJOTTA Jacqueline (dét.)	Mathématiques
M.	TOULOTTE Jean-Marc	I.E.E.A.
M.	TURRELL Georges	Chimie
M.	VANDORPE Bernard	E.U.D.I.L.
M.	VAST Pierre	Chimie
M.	VERBERT André	Biologie
M.	VERNET Philippe	Biologie
M.	WALLART Francis	Chimie
M.	WARTEL Michel	Chimie
M.	WATERLOT Michel	Sciences de la Terre
Mme	ZINN JUSTIN Nicole	Mathématiques

CHARGES DE COURS

M. ADAM Michel	S.E.S.

CHARGES DE CONFERENCES

M. BAFCOP Joël	I.P.A.
M. DUVEAU Jacques	S.E.S.
M. HOFLACK Jean	I.P.A.
M. LATOUCHE Serge	S.E.S.
M. MALAUSSENA DE PERNO Jean-Louis	S.E.S.
M. NAVARRE Christian	I.P.A.
M. OPIGEZ Philippe	S.E.S.

Je tiens à exprimer ma profonde gratitude à tous ceux qui m'ont apportés aide et soutien au cours de l'élaboration de cette thèse.

Je remercie particulièrement Max Dauchet qui fut à l'origine de ce travail, en a assuré l'orientation et le suivi, et qui a toujours su répondre à mes nombreuses sollicitations.

Je remercie également les membres du Jury : André Arnold, Christian Carrez, Gérard Comyn, Bruno Courcelle et Michel Latteux, pour l'intérêt qu'ils ont porté à ce travail et, chacun pour leur part, pour les critiques, remarques, suggestions et conseils, tant lors de l'élaboration que lors de la rédaction.

Enfin je remercie Bénédicte ainsi que Monsieur Glane qui ont réalisé avec compétence la dactylographie et le tirage de cette thèse.

E. Tison

TABLE DES MATIERES

Pages

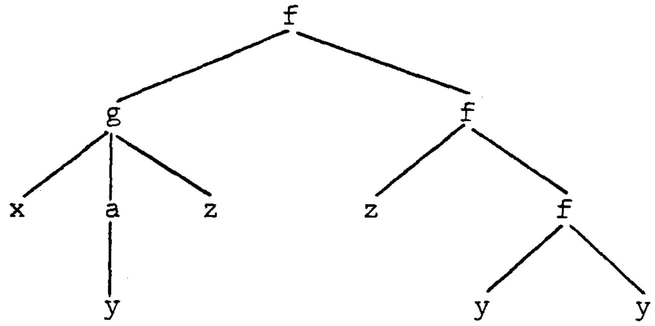
INTRODUCTION

L'opération de feuillage, bien définie pour les arbres finis, établit une liaison importante entre arbres et mots, forêts et langages. Nous étendons cette opération aux arbres infinis à partir de spécifications, par des constructions standards. Les "mots" infinis que nous définissons lors de cette construction sont de nature plus générale que les mots infinis usuels et que les mots bi-infinis ; ils se rapprochent plutôt des "arrangements" dénombrables introduits par Courcelle et Heilbrunner. Plus précisément, l'opération de feuillage obtenue s'avère moins fine que celle de "frontière" d'arbres infinis définie à l'aide des arrangements.

Le résultat important, qui est établi alors, est la décidabilité de l'égalité des feuillages d'arbres infinis réguliers. Nous nous intéressons ensuite à la structure de l'ensemble de mots construit et montrons que cet ensemble est un 2/3 de SFP mais pas un SFP, ces deux classes de C.P.O. ayant été introduites en sémantique par Plotkin. Cette propriété permet de métriser convenablement cet ensemble.

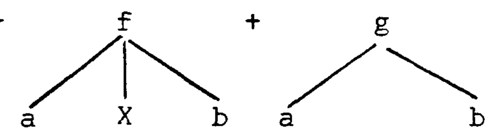
Les arbres sont des objets qui apparaissent partout en informatique ; qu'il s'agisse d'arbres programmatiques, d'arbres comme structures de données, d'arbres syntaxiques, d'arbres sémantiques, ils ont été abondamment étudiés et leur rôle fondamental n'est plus à démontrer.

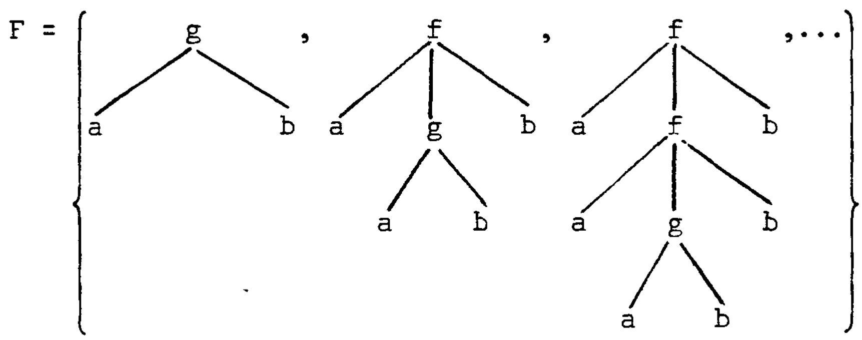
Le feuillage d'un arbre est un mot représentant les feuilles de l'arbre, dans un ordre donné. Considérons par exemple l'arbre suivant :



son feuillage est le mot x y z z y y

Son intérêt est évident dès lors qu'on manipule des arbres de production, de dérivation. Il est bien connu que le feuillage d'une forêt reconnaissable est un langage algébrique et réciproquement tout langage algébrique est feuillage d'une forêt reconnaissable. Par exemple la forêt reconnaissable F définie par la grammaire régulière d'arbre : $X \rightarrow$



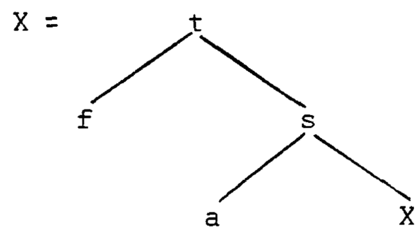


a pour feuillage le langage algébrique : $\{a^n\, b^n / n \in N_+\}$.

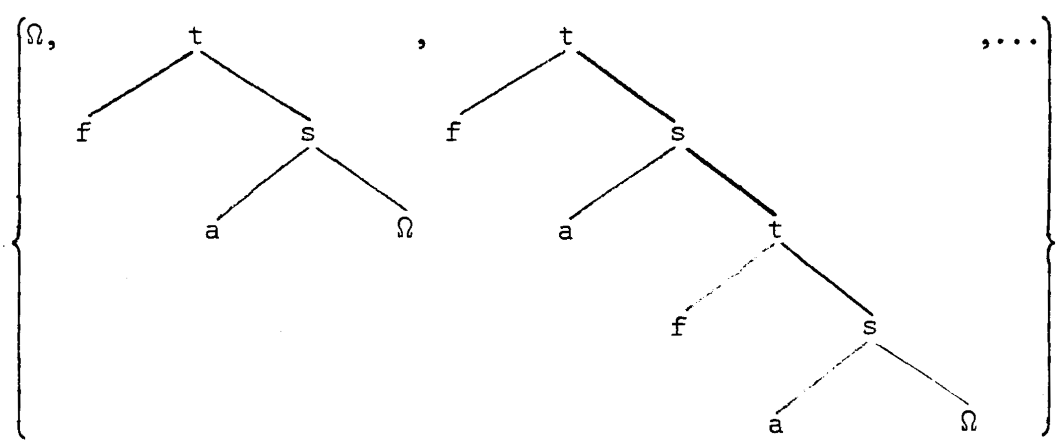
Des propriétés des langages algébriques peuvent alors être obtenues à partir de propriétés plus simples des forêts reconnaissables par application du feuillage ; un exemple classique est le lemme d'itération de la double étoile des langages algébriques. Cette idée de se plonger dans une structure plus riche, celle d'arbre, plutôt qu'opérer lourdement sur la structure apparente, plus pauvre, celle de mot, est utilisée en compilation par attribut (Knuth, Courcelle, Lohro,...) où l'on travaille sur un arbre ayant pour feuillage le programme (un mot) considéré.

Quant aux arbres infinis, ils apparaissent naturellement dans l'étude de la sémantique des langages de programmation. Ils s'obtiennent par exemple en développant à l'infini les itérations d'un programme (Cousineau, Jacob,...) ou en développant, à l'infini également, des schémas de programmes itératifs ou récursifs (Nivat, Courcelle, Guessarian, Kott,...). Les arbres obtenus de cette façon sont algébriques, et avec l'ordre syntaxique ("moins défini que") sur l'ensemble des arbres finis et infinis, sont à la base de la sémantique algébrique définie par Nivat ; la sémantique d'un programme est alors définie par l'arbre associé (limite des arbres finis d'approximation représentant la suite des calculs finis).

Un cas particulier important est celui des arbres réguliers ; ils ont un nombre fini de sous-arbres distincts ; ce sont ceux que l'on obtient en développant les itérations d'un programme, ou en déployant les schémas de programme itératif. Les arbres réguliers sont également les solutions de certains systèmes d'équations d'arbres (Cousineau, Courcelle,...), dits systèmes réguliers. Prenons par exemple le système, réduit à une équation :

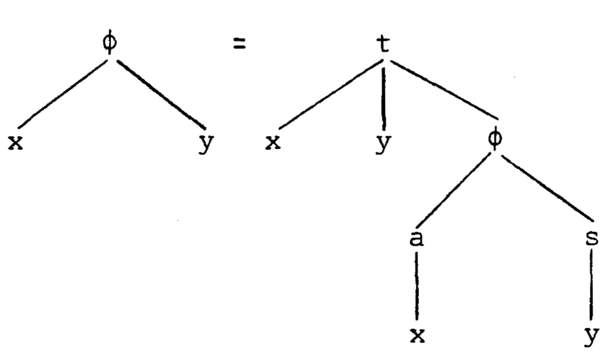


en réécrivant itérativement la variable X dans le second membre de l'équation, et en remplaçant cette variable par le symbole "indéfini" Ω on obtient la suite croissante (au sens de l'ordre syntaxique) d'arbres :

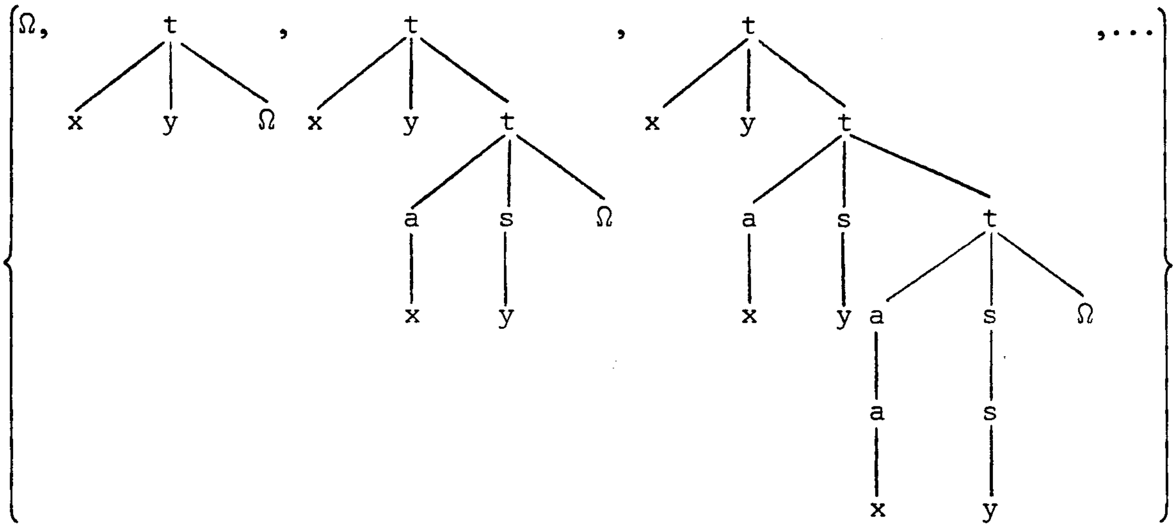


Le supremum (sup) de cette suite est un arbre infini régulier.

Plus généralement, les arbres algébriques sont les solutions de systèmes algébriques d'arbres, dont un exemple simple est le suivant :



En procédant comme précédemment par réécritures et en remplaçant les sous-arbres du type ϕ par le symbole Ω, on obtient la suite croissante d'arbres :



qui a pour sup un arbre infini algébrique.

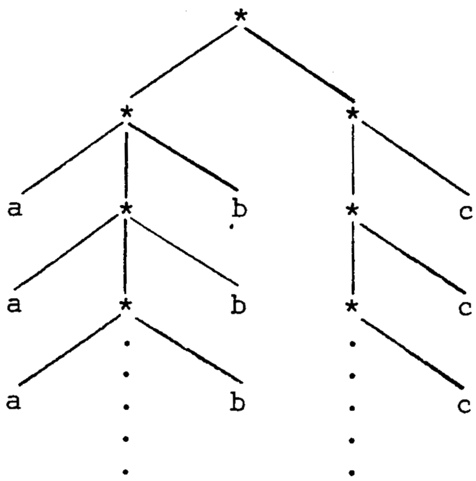
Les mots infinis, souvent considérés comme cas particulier des arbres infinis, et les langages infinitaires permettent de décrire les comportements infinis de processus (Büchi, Mac-Naughton,...) et de formaliser les problèmes de synchronisation et du parallélisme de processus (Arnold, Nivat,...).

Les mots "bi-infinis" (Nivat, Perrin) sont une extension naturelle des mots infinis ; alors que les mots infinis ont par définition un début, les mots bi-infinis n'ont ni début, ni fin ; ils représentent des suites de lettres indexées par l'ensemble des entiers relatifs.

Ces types de mots sont évidemment insuffisants pour pouvoir étendre aux arbres infinis l'opération de feuillage ; on ne pourrait, à priori, décrire convenablement que les feuillages d'arbres ayant une, voire deux branches infinies.

Les "arrangements" dénombrables, généralisation plus forte des mots infinis, ont permis de définir la "frontière" d'arbres infinis (Courcelle), un arrangement étant une suite de lettres indexées par un ensemble, quelconque, totalement ordonné. La frontière d'un arbre est alors l'arrangement défini par les noeuds terminaux (feuilles) de l'arbre. Considérons par exemple l'arbre suivant :

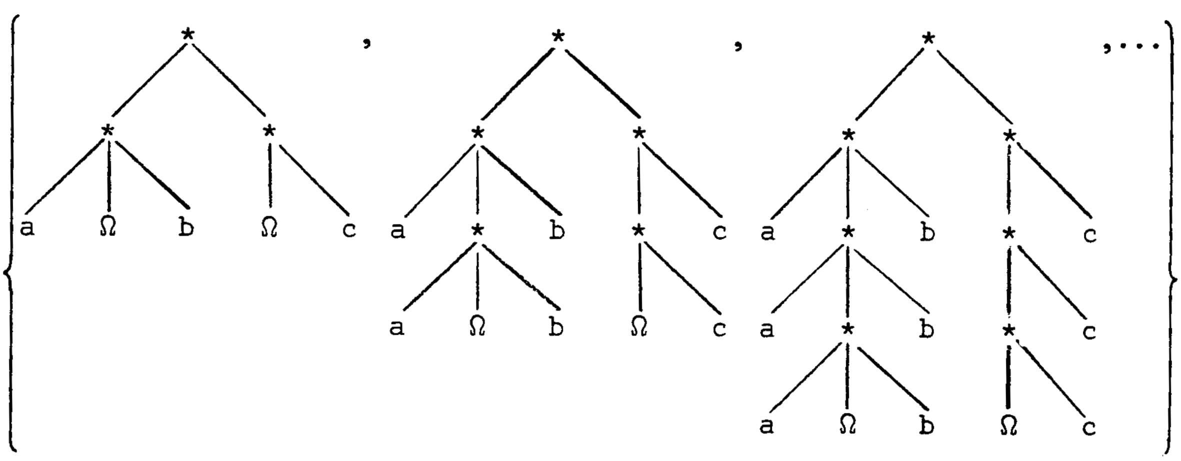
t =



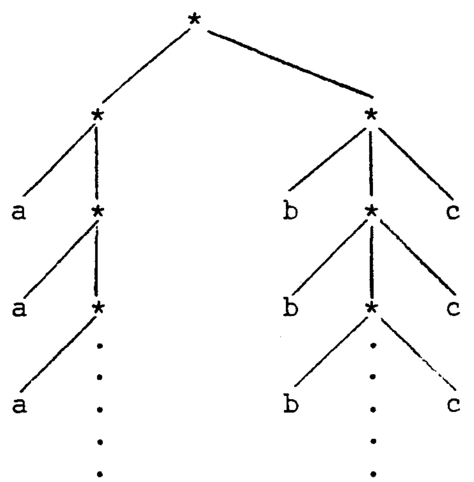
La frontière de cet arbre est l'arrangement dénombrable, noté $a^{\omega} \bullet b^{-\omega} \bullet c^{-\omega}$, ($b^{-\omega}$ indiquant une infinité de lettres 'b' obtenues en "remontant" dans l'arbre de l'infini vers la racine) que l'on peut schématiser de la façon suivante :

$$
\begin{array}{ccccccccc}
1\ 2\ldots\ldots n & \ldots +\infty & \ldots +\infty & \ldots\ldots p & \ldots 2\ 1 & \ldots +\infty & \ldots\ldots q & \ldots 2 & 1 \\
\downarrow \quad \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
a\ a\ a\ldots\ldots a\ a & \ldots a & \ldots b & \ldots\ldots b\ b\ b & \ldots b\ b & \ldots c & \ldots\ldots c\ c\ c & \ldots c & c
\end{array}
$$

La frontière est effectivement une extension du feuillage d'arbres finis, mais ne possède cependant pas la "bonne" propriété d'être continue ; on ne peut donc pas identifier la frontière d'un arbre infini à une "limite" d'approximations finies. Reprenons l'exemple précédent, la suite d'arbres :



est croissante et a pour sup l'arbre t considéré. L'ensemble des frontières de ces arbres finis est le langage : $\{a^n \Omega b^n \Omega c^n / n \in N_+\}$. Si l'on considère maintenant l'arbre : t' =



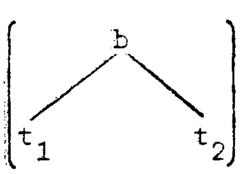
sa frontière est $a^{\omega} \cdot b^{\omega} \cdot c^{-\omega}$

elle est différente de celle de t, alors qu'en considérant, comme précédemment, les approximations de cet arbre on obtient les mêmes frontières finies.

Les frontières d'arbres réguliers ont été étudiées (Courcelle, Heilbrunner) et si l'on peut les exprimer à l'aide d'expressions, la décidabilité de l'équivalence reste un problème ouvert.

Que pourrait alors être une généralisation plus satisfaisante de la notion de feuillage ?

La première contrainte est de pouvoir concaténer les mots infinis (ce qui est le cas pour les arrangements) pour avoir la propriété intrinsèque du feuillage, noté ϕ :

ϕ 

$= \phi(t_1) \cdot \phi(t_2)$

La deuxième contrainte est celle de continuité afin de pouvoir considérer des approximations finies représentant les mots infinis. Les arbres, munis de l'ordre syntaxique, ayant une structure de CPO, il paraît donc raisonnable de s'imposer que le feuillage soit une application continue d'un CPO dans un autre CPO, c'est-à-dire vérifie :

$$\Phi(\sup\{t_i \mid i \in N\}) = \sup\{\Phi(t_i) \mid i \in N\}$$

si $\{t_i \mid i \in N\}$ est une suite croissante d'arbres.

Ces contraintes induisent certaines propriétés que doit vérifier l'ensemble des mots, concernant la relation d'ordre, la concaténation et la compatibilité entre les deux.

La solution la plus générale répondant à ces spécifications (objet initial d'une catégorie) est obtenue par des constructions standards (complétion d'ensemble ordonné, extensions par continuité). L'ordre obtenu sur les mots est plus général que l'ordre préfixe usuel ou l'ordre "facteur de" ; ainsi les mots se précisent en tout point où l'on rencontre le symbole indéfini Ω. Par exemple, le mot $a\Omega bb\Omega c\Omega$ est plus petit que le mot $ac\Omega abbbac$: la première occurrence de la lettre Ω se précise en $c\Omega a$, la deuxième en ba, et la troisième en ε, le mot vide.

Il est intéressant de noter qu'avec cet ordre, si deux arbres fonctionnels sont en relation par l'ordre syntaxique, les termes canoniquement associés (des mots) sont en relation également.

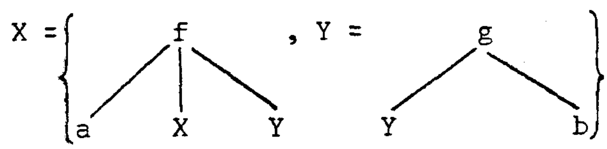
Les mots infinis et bi-infinis sont des classes particulières des "mots" infinis que nous construisons.

La comparaison entre frontière et feuillage montre que l'opération de frontière est plus fine ; plus précisément, pour les arbres localement finis, des arbres ayant même frontière ont même feuillage, la réciproque étant fausse.

On peut donc définir une surjection de l'ensemble des arrangements dénombrables dans une partie, très large, de l'ensemble de (nos) mots infinis. Reprenant les arbres t et t' de l'exemple précédent, ils ont même feuillage qui est égal à $\mathrm{Sup}\{a^n\ \Omega\ b^n\ \Omega\ c^n \mid n \in N_+\} = \mathrm{Sup}(a^*\ \Omega\ b^*\ \Omega\ c^*)$ et à l'image par cette surjection des arrangements : $a^{\omega} \cdot b^{-\omega} \cdot c^{-\omega}$ et $a^{\omega} \cdot b^{\omega} \cdot c^{-\omega}$.

La liaison entre systèmes d'équations régulières d'arbres et systèmes d'équations algébriques de mots, par l'opération de feuillage, est immédiate. Il n'est donc pas surprenant d'obtenir que les feuillages d'arbres réguliers sont les (plus petites) solutions des systèmes algébriques de mots.

La démarche utilisée est exactement celle de Nivat en sémantique algébrique, et consiste à opérer par substitutions (réécritures des non-terminaux) et à remplacer les non-terminaux (variables) par Ω de façon à ordonner les approximations obtenues. Par exemple, du système régulier d'arbres :



on déduit le système algébrique de mots :

$$\{X = aXY \ ; \ Y = Yb\}$$

les approximations de Y seront de la forme Yb^n, $n \in N$ ce qui donnera la suite croissante $\{\Omega b^n \mid n \in N\} = \Omega b^*$ dont le sup est le feuillage de l'arbre solution de la deuxième équation. Pour X on obtiendra par réécritures, en X uniquement, des mots de la forme $a^n X Y^n$, $n \in N$ puis $\{a^n\Omega\ (\Omega b^p)^n \mid n,\ p \in N\}$ dont le sup est égal au feuillage de l'arbre infini régulier solution (en X) du système, est encore égal à $\sup(a^*\Omega\ (\Omega b^*)^*)$.

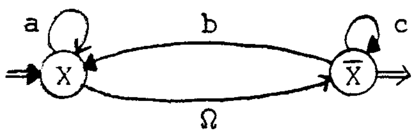
Sur cet exemple, simple, on voit que l'on a pu exprimer la solution (la plus petite en fait) du système algébrique, et donc le feuillage des deux arbres réguliers considérés, à l'aide de langages rationnels. De façon générale on montre que c'est toujours possible, c'est-à-dire que pour tout système algébrique de mots chaque composante de la plus petite solution est le sup d'un langage rationnel, dont on peut calculer une expression. Nous en déduisons alors la décidabilité de l'égalité des feuillages d'arbres infinis réguliers.

Il est à noter que des arbres d'allures très différentes peuvent avoir même feuillage et que l'ensemble des approximations de la solution d'un système algébrique, obtenues selon le procédé décrit précédemment, est un langage algébrique non trivial en général. Il suffit, pour s'en convaincre, de considérer la simple équation :

$X = aXbXc$ résolue par le langage rationnel $a^*(\Omega c^* ba^*)^* \Omega c^*$ dont le sup est également la première composante de la solution du système :

$$\{X = X_1 X_2 X_3 \,;\, X_1 = aX_1 \,;\, X_2 = X_2 X_3 b X_1 \,;\, X_3 = X_3 c\}.$$

L'obtention d'un rationnel peut être faite en construisant un automate associé au système d'équation, ou par transformation de l'expression représentant la frontière de l'arbre de dérivation, expression calculée par l'algorithme de Heilbrunner de résolution des systèmes algébriques. Par exemple pour l'équation $X = aXbXc$, l'automate associé est le suivant :



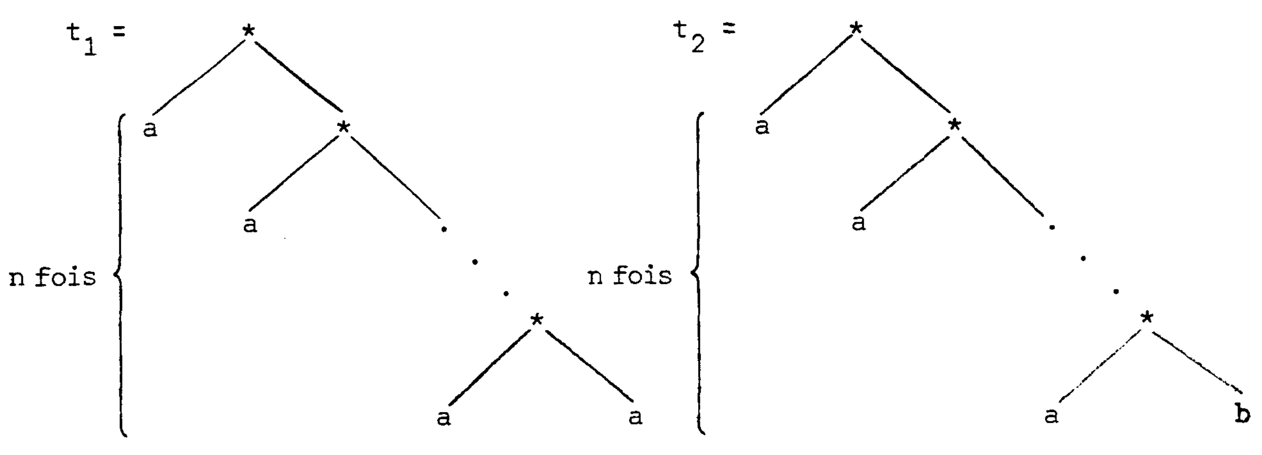
et en utilisant l'algorithme d'Heilbrunner on obtient l'expression rationnelle : $a^*\Omega(\Omega c^* ba^*\Omega)^*\Omega c^*$ représentant un langage équivalent (au sens de l'ordre) à celui reconnu par l'automate.

Par un codage simple, nous montrons ensuite que le problème de décidabilité de l'égalité des arbres algébriques est réductible à celui de l'égalité des feuillages d'arbres algébriques. Utilisant cette démarche pour les arbres réguliers on retrouve alors la décidabilité de l'égalité des arbres réguliers.

Le problème de faire de l'ensemble des mots un espace métrique s'inscrit dans le cadre général de la métrisation des C.P.O. D'un point de vue général cet aspect est important ; ainsi dans la sémantique du non-déterminisme (Arnold, Nivat, Boudol,...) c'est l'aspect espace métrique complet des arbres qui est fondamental, la suite des approximations n'étant plus dirigée du fait des différents choix possibles induits par le non-déterminisme.

Ce problème est résolu pour certaines classes de C.P.O ; pour les C.P.O. ω-algébriques (pour lesquels on peut "séparer" les objets infinis des objets finis (finitaires)), Comyn et Dauchet ont défini une ultramétrique associée à la topologie dite de Lawson (ou de Cantor). Cet espace métrique n'est cependant pas complet en général, mais pour certaines classes plus fines de C.P.O, les SFP et 2/3 de SFP, introduites en sémantique par Plotkin, cet espace est complet et il y a donc isomorphie entre le complété métrique et le complété au sens de l'ordre. Il fallait donc situer notre C.P.O. parmi ces classes ; nous montrons qu'il s'agit d'un 2/3 de SFP qui n'est pas un SFP. Intuitivement le fait que ce ne soit pas un SFP indique que l'ordre sur les mots n'est pas "simple" ; la propriété d'être 2/3 de SFP équivaut à pouvoir "bien" métriser.

La métrique obtenue s'avère être de nature quelque peut différente de celle sur les arbres ou sur les mots usuels : en effet pour ces derniers la distance entre deux éléments est fonction de la plus grande "longueur" d'information commune, il y a donc prédominance de l'information positive, alors que dans notre cas elle dépend avant tout du plus petit sous-mot qui les sépare (i.e. qui apparaît dans l'un et pas dans l'autre) et privilégie l'information négative. Par exemple, les deux arbres suivant :



ont une distance de l'ordre de $\frac{1}{n}$, alors que leurs feuillages (i.e. les mots a^{n+1} et $a^n b$) ont pour nous une distance de l'ordre de 1 du fait de l'occurrence de la lettre b dans l'un et pas dans l'autre.

Dans le premier chapitre nous rappelons les définitions et propriétés essentielles des objets que nous manipulons, à savoir : les mots, les arbres, les C.P.O. et l'opération de feuillage d'arbres finis.

Dans le second, après avoir spécifié le problème de généralisation du feuillage, nous construisons ce feuillage et donc l'ensemble des mots résultants. Nous comparons alors ces objets avec les mots infinis et bi-infinis, et par l'intermédiaire de la frontière, avec les arrangements dénombrables.

Au cours du troisième chapitre, nous nous intéressons aux systèmes d'équations. Le but est de donner un cadre formel à la notion de construction par approximations successives de la (plus petite) solution. Nous rappelons tout d'abord pour cela les résultats concernant les systèmes d'équations d'arbres. Le résultat est obtenu par la même démarche que pour ces systèmes. Le lien entre systèmes réguliers d'arbres et systèmes algébriques de mots, et entre leur solution est ensuite établi.

Le quatrième chapitre est consacré au problème de décision, évoqué précédemment, et donc à la résolution, par deux méthodes différentes, des systèmes algébriques de mots, puis au codage des arbres algébriques.

Enfin dans le cinquième chapitre, nous ne faisons qu'effleurer le problème de métrisation du C.P.O. en le situant parmi les classes connues, après avoir brièvement rappelé les résultats dont nous avons besoin.

I. RAPPEL DES NOTIONS DE BASE

1. Monoïdes, Mots

Un **monoïde** est un ensemble (E) muni d'une opération (application de $E \times E$ dans E) associative et qui possède un élément neutre.

Soit (E, o, ε) un monoïde (E = l'ensemble, o désigne l'opération et ε l'élément neutre).

Une relation d'équivalence R sur E est une **congruence** si pour tout $x, y, z, t \in E$ on a :

$x \; R \; y$ et $z \; R \; t$ implique $(x \circ z) \; R \; (y \circ t)$.

Une application ϕ d'un monoïde $(E, \circ, \varepsilon_E)$ dans un monoïde $(F, \square, \varepsilon_F)$ est un **morphisme de monoïde** si elle vérifie :

- $\phi(\varepsilon_E) = \varepsilon_F$
- $\forall x, y \in E : \phi(x \circ y) = \phi(x) \square \phi(y)$

Notons N_+ l'ensemble des entiers naturels positifs et $[n] = \{1, 2, \ldots, n\}$, $n \in N_+$.

Soit X un alphabet non vide ;

- un mot fini (u) sur X est une application d'une partie initiale de N_+ dans X

i.e. $u : [n] \to X$, $n \in N_+$; n est la longueur du mot u, notée $|u|$

ou $u : \phi$ (l'ensemble vide) $\to X$

le mot vide, noté ε, de longueur nulle.

- un mot infini sur X est une application de N_+ dans X

X^* désigne l'ensemble des mots finis sur X

X^ω l'ensemble des mots infinis sur X

et $X^\infty = X^* \cup X^\omega$.

Concaténation

$\forall u, v \in X^\infty$, le concaténé de u et v est le mot w, noté uv, défini par :

- si $u \in X^\omega$ (u est infini), $w = u$

- si $u \in X^*$, $u : [n] \to X$

$w(i) = u(i) \quad \forall i \in [n]$
$w(j) = v(j-n) \quad \forall j > n$

Exemple :

$X = \{a, b\}$

Soit $u = a^5 = aaaaa$

(i.e. $u : [5] \to X$, $u(i) = a \quad \forall i \in [5]$).

Soit v le mot infini : $v : N_+ \to X$, $v(i) = b \quad \forall i \in N_+$.

Notation :

$v = b^\omega$

$w = uv = a^5 b^\omega$ est défini par :

$w(i) = a \quad \forall i \in [5]$
$w(i) = b \quad \forall i > 5$

et $w' = vu = v = b^\omega$.

X^∞ muni de la concaténation est un monoïde, d'élément neutre ε le mot vide.

Relations d'ordre sur les mots :

- Ordre préfixe ou "facteur gauche".

$\forall u, v \in X^\infty$, u est préfixe de v, noté $u < v$, si il existe $w \in X^\infty$ tel que $uw = v$.

Le mot vide, ε, est le plus petit élément de X^∞ pour cette relation d'ordre, et tout mot infini est maximal, i.e. $\forall u \in X^\omega$, $u < v \Rightarrow u = v$.

- Ordre lexicographique : $\leq_\ell$

Soit $r : X \to N_+$ une numérotation des lettres de X.

$\forall u, v \in X^\infty$

$$u \leq_\ell v \quad \text{ssi} \quad \begin{cases} - \ u < v \\ \text{ou} \\ - \text{ si } i \text{ est le plus petit entier tel que} \\ u(i) \neq v(i) \text{ alors } r(u(i)) < r(v(i)) \end{cases}$$

2. Les arbres

Soit Σ un alphabet gradué ; i étant un entier, soit Σ_i l'ensemble des éléments de Σ de arité, ou degré, i.

Domaine d'arbre :

Une partie A de $(N_+)^*$ (l'ensemble des mots finis sur l'alphabet N_+) est un domaine d'arbre, si elle vérifie les deux conditions :

1) - A est close par préfixe et n'est pas vide
i.e. $\forall \alpha, \beta \in (N_+)^*$

$\alpha\beta \in A$ implique $\alpha \in A$

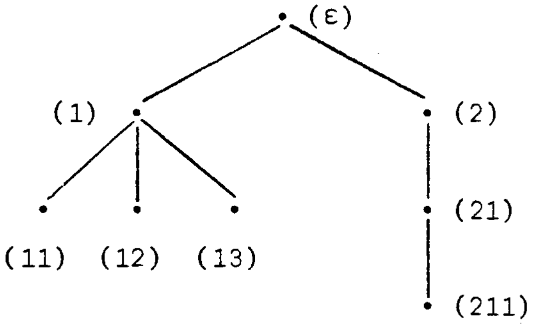
2) - $\forall \alpha \in (N_+)^*$, $\forall i \in N_+$:

$\alpha i \in A$ implique $\alpha j \in A$ pour tout entier j inférieur à i.

Exemples :

- $A = \{\varepsilon, 1, 2, 11, 12, 13, 21, 211\}$

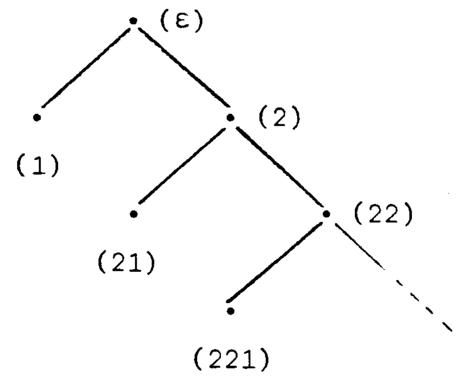
Ce domaine d'arbre correspond au "schéma arborescent" suivant :



- $B = 2^* + 2^*1$

i.e. $B = \{2^n \mid n \in N\} \cup \{2^p 1 \mid p \in N\}$.

B est un domaine d'arbre infini, représenté schématiquement par



<u>Arbre</u> :

Un arbre, sur l'alphabet Σ, est une application partielle $t : (N_+)^* \to \Sigma$ telle que son domaine, noté dom(t), soit un domaine d'arbre et qui vérifie :

$\forall \alpha \in \mathrm{dom}(t)$, si $t(\alpha) \in \Sigma_i$ (de arité i) $\alpha j \in \mathrm{dom}(t)$ ssi $j \leq i$ pour tout entier j.

Un arbre est fini si son domaine est fini, infini sinon.

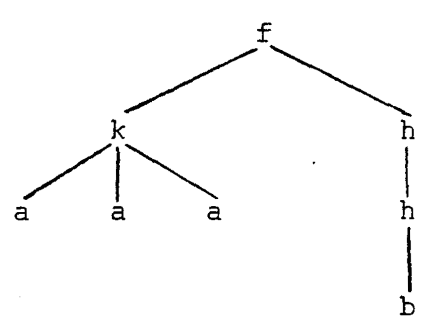
<u>Exemples</u> :

Soit $\Sigma = \Sigma_0 \cup \Sigma_1 \cup \Sigma_2 \cup \Sigma_3$ avec $\Sigma_0 = \{a, b\}$, $\Sigma_1 = \{h\}$, $\Sigma_2 = \{f, g\}$, $\Sigma_3 = \{k\}$.

Soit l'arbre t défini par :

$t(\varepsilon) = f$, $t(1) = k$, $t(11) = t(12) = t(13) = a$, $t(2) = t(21) = h$, $t(211) = b$.

t est représenté par la figure suivante :

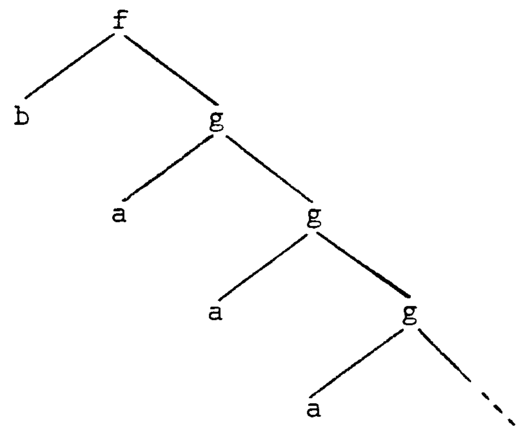


ou par l'expression f(k(a, a, a), h(h(b))).

Soit maintenant l'arbre t' défini par :

- $t'(2^n) = g$, $\forall n > 0$
- $t'(2^n1) = a$, $\forall n > 0$
- $t'(\varepsilon) = f$
- $t'(1) = b$.

t' est un arbre infini schématisé par :



On notera : $T(\Sigma)$ l'ensemble des arbres finis et $T^{\infty}(\Sigma)$ l'ensemble des arbres finis ou infinis sur l'alphabet Σ.

Remarque :

$T(\Sigma)$ peut également être défini comme étant le plus petit ensemble vérifiant :

- $\Sigma_0 \subset T(\Sigma)$

- $\forall n \in N_+$, $\forall f \in \Sigma_n$, $\forall t_1, \ldots, t_n \in T(\Sigma)$

$$f(t_1, \ldots, t_n) \in T(\Sigma).$$

Sous-arbres :

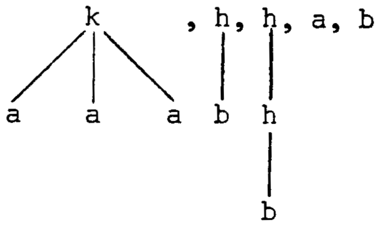
Soient t, t' deux arbres de $T^{\infty}(\Sigma)$.

t est un sous-arbre de t' s'il existe un mot α de $(N_+)^*$ tel que :

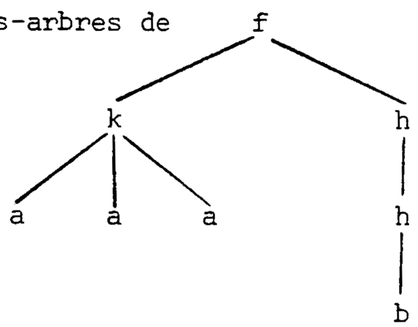
$\forall w \in (N_+)^*$, $w \in dom(t)$ ssi $\alpha w \in dom(t')$ et $t(w) = t'(\alpha w)$.

Exemples :

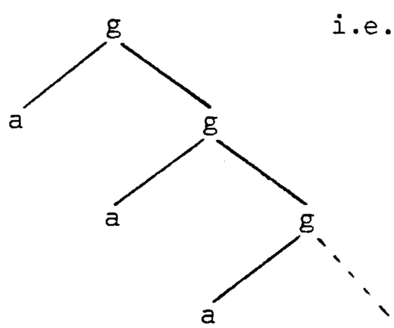
Les arbres :



sont des sous-arbres de



l'arbre t =



i.e. $t(2^n) = g \quad \forall n \geq 0$

$t(2^n 1) = a \quad \forall n \geq 0$

est un sous-arbre de l'arbre infini t' défini dans un exemple précédent.

- Ordre syntaxique : $<$

Soit Ω un symbole n'appartenant pas à Σ, de arité nulle. (Ω = "l'indéfini").

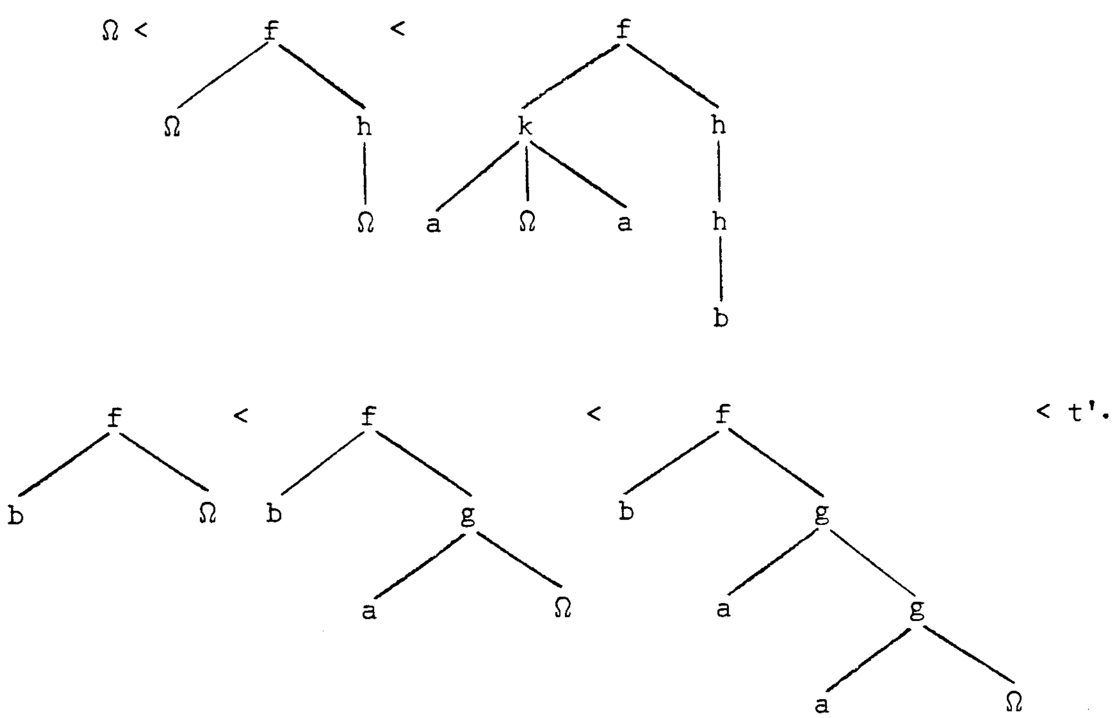
Notons, de façon usuelle, $T_\Omega(\Sigma)$ pour $T(\Sigma \cup \{\Omega\})$ et $T^\infty_\Omega(\Sigma)$ pour $T^\infty(\Sigma \cup \{\Omega\})$.

L'ordre syntaxique sur les arbres est défini par : $\forall t, t' \in T^\infty_\Omega(\Sigma)$,
$t < t'$ ssi $dom(t) \subset dom(t')$ et $\forall \alpha \in dom(t)$, $t(\alpha) \neq \Omega$ implique $t'(\alpha) = t(\alpha)$.

Sur $T_\Omega(\Sigma)$, (arbres finis), l'ordre syntaxique peut être défini récursivement par :

$$t < t' \text{ si } \left| \begin{array}{l} t = \Omega \\ \text{ou} \\ t = f(t_1,\ldots, t_k) \text{ et } t' = f(t'_1,\ldots, t'_k) \\ \text{avec } t_i < t'_i \quad \forall i = 1,\ldots, k \\ f \in \Sigma_k \end{array} \right.$$

<u>Exemples</u> :



L'arbre Ω (de domaine $\{\varepsilon\}$, image de $\varepsilon = \Omega$) est le plus petit élément de $T^\infty_\Omega(\Sigma)$ pour l'ordre syntaxique. $T^\infty(\Sigma)$ est l'ensemble des arbres maximaux de $T^\infty_\Omega(\Sigma)$ relativement à cet ordre.

<u>Tronqué d'un arbre</u> :

Etant donné un arbre t de $T^\infty_\Omega(\Sigma)$, le "tronqué" de t à la profondeur n, n entier, est l'arbre fini noté $H_n(t)$ et défini par :

- $\mathrm{dom}(H_n(t)) = \{\alpha \in \mathrm{dom}(t) \ / \ |\alpha| \leq n\}$

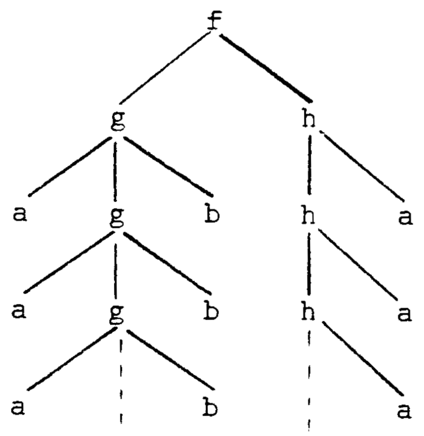
  ($|\alpha|$ = longueur du mot α)

- $\forall\alpha \in \mathrm{dom}(H_n(t))$

  $H_n(t)(\alpha) = t(\alpha)$ si $|\alpha| < n$
  $= \Omega$ si $|\alpha| = n$.

Exemple :

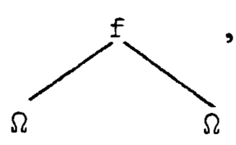
Soit t =



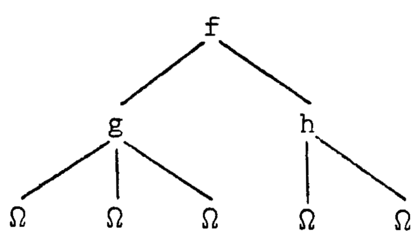
i.e. $t(\varepsilon) = f \quad \forall n \in N$
$t(12^n1) = a$
$t(12^n) = g$
$t(12^n3) = b$
$t(21^n) = h$
$t(21^n2) = a$

$H_0(t) = \Omega$

$H_1(t) =$



, $H_2(t) =$



, etc...

Bien sur : $\forall n \in N$, $H_n(t) < t$ et si t est un arbre fini il existe un entier p tel que $\forall n \geq p$: $H_n(t) = t$.

Arbres localement finis :

Un arbre t de $T^\infty(\Sigma)$ est dit localement fini si pour tout mot w de dom(t) il existe un mot v de $(N_+)^*$ tel que : $wv \in \mathrm{dom}(t)$ et $t(wv) \in \Sigma_0$.

Autrement dit, de tout noeud de l'arbre il existe un chemin (descendant) fini vers une feuille.

On notera $T^{Loc}(\Sigma)$ l'ensemble des arbres localement finis sur l'alphabet Σ.

On a : $T(\Sigma) \subsetneq T^{Loc}(\Sigma) \subsetneq T^{\infty}(\Sigma)$.

Exemple :

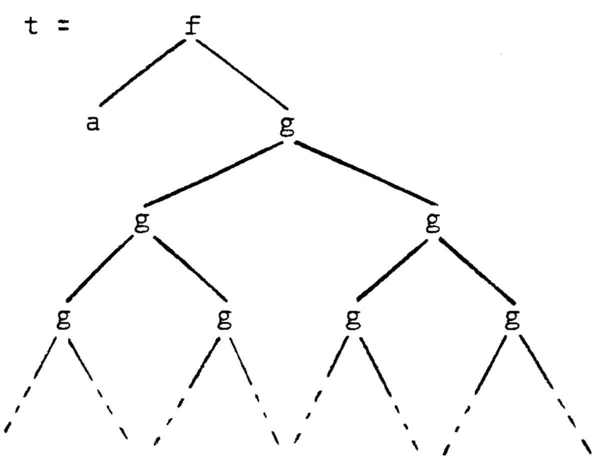
Les arbres infinis des exemples précédents sont localemnt finis ; par contre l'arbre suivant ne l'est pas :

$t(\varepsilon) = f$

$t(1) = a$

et $t(w) = g \quad \forall w \in 2\cdot(1+2)^*$ (i.e. tout mot formé avec des 1 et 2 commençant par un 2).

t =



3. Les C.P.O. ω-algébriques

Soit un ensemble E muni d'une relation d'ordre (partiel), notée $<$.

Une partie D de E est dite **dirigée** ssi toute partie finie de D est majorée dans D ; autrement dit :

$\forall x, y \in D, \exists z \in D$ tel que $x < z$ et $y < z$.

Définition :

Un ensemble ordonné E est un **C.P.O.** (Complete Partial Order) ssi :

- E possède un plus petit élément (souvent noté $\perp$)

et

- toute partie dirigée de E admet un sup (plus petit majorant).

Exemples :

- X^{∞}, l'ensemble des mots finis et infinis sur l'alphabet X, muni de l'ordre préfixe est un C.P.O. Le plus petit élément est ε le mot vide ; toute partie dirigée de X^{∞}, pour cet ordre, est une suite croissante.

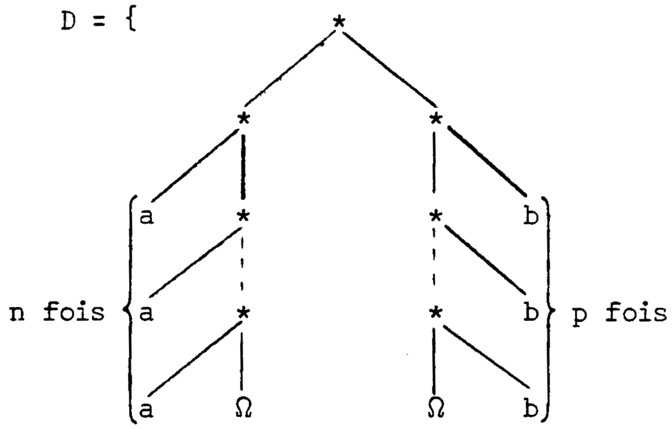
- X^*, l'ensemble des mots finis sur X, muni de l'ordre préfixe n'est pas un C.P.O.

En effet toute suite infinie croissante n'a pas de sup dans X^*.

- $T^{\infty}_{\Omega}(\Sigma)$, l'ensemble des arbres sur $\Sigma \cup \{\Omega\}$, muni de l'ordre syntaxique est un C.P.O.

Exemple de partie dirigée :

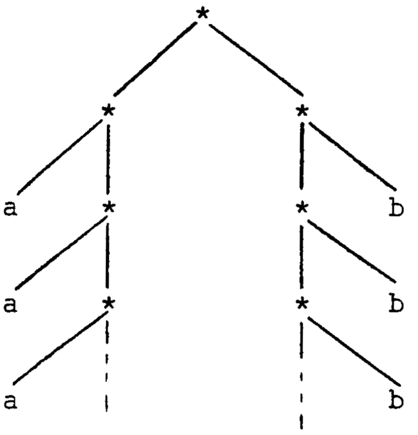
D = {



, n et p $\in$ N}.

D n'est pas une suite croissante ; son sup est l'arbre infini $t \in T^{\infty}(\Sigma)$:

t =



<u>Base d'un C.P.O.</u> :

Un sous-ensemble B d'un C.P.O. E est une base de E ssi tout élément de E est le sup d'une partie dirigée de B ; i.e.

$\forall x \in E, \exists D \subset B$, D dirigée telle que $x = \sup(D)$.

Par exemple : - X^* est une base du C.P.O. X^∞, tout mot infini est le sup d'une suite croissante de mots finis.

- $T_\Omega(\Sigma)$ est une base de $T^\infty_\Omega(\Sigma)$, tout arbre infini est le sup d'une partie dirigée d'arbres finis.

<u>Complétion</u> :

Tout ensemble ordonné, B, possédant un plus petit élément peut être "complété" en un C.P.O. noté B^∞ tel que B soit isomorphe à une partie de B^∞ et constitue une base de B^∞.

Cette complétion peut être faite par classes d'équivalence de parties dirigées de B ; deux parties dirigées D et D' étant équivalentes si tout élément de D est majoré par un élément de D' et réciproquement.

Exemples :

- N l'ensemble des entiers naturels, muni de l'ordre usuel, sera complété en $N \cup \{+\infty\}$.

- X^{∞} (respectivement $T_{\Omega}^{\infty}(\Sigma)$) est le complété de X^{*} (respectivement $T_{\Omega}(\Sigma)$).

Elément finitaire :

Un élément b d'un C.P.O. E est dit "finitaire" ssi pour toute partie dirigée D de E on a :

$(b < \sup(E) \Rightarrow \exists c \in D \text{ et } b < c)$.

Remarque :

Dans X^{∞} comme dans $T_{\Omega}^{\infty}(\Sigma)$ la notion d'élément finitaire coïncide avec celle d'élément fini (les mots et arbres finis sont finitaires ; les mots et arbres infinis ne sont pas finitaires) ; ceci n'est pas vrai de façon générale.

Considérons par exemple l'ensemble des représentations décimales des réels positifs, muni de l'ordre classique sur R. C'est un C.P.O., de plus petit élément 0, ayant pour base l'ensemble des représentations décimales finies. Soit la suite croissante $(x_i)_{i \in N}$ avec $x_i = 1,\underbrace{99\ldots9}_{i \text{ fois}}$.

Cette suite admet comme sup le réel 2 d'où $2 \leq \sup\{x_i \mid i \in N\}$ mais il n'existe aucun élément de cette suite majorant 2 ; 2 n'est donc pas finitaire bien que fini.

Base finitaire :

Une base B d'un C.P.O. E est dite finitaire si elle est constituée d'éléments finitaires.

Si un C.P.O. admet une base finitaire dénombrable alors elle est unique et tout élément finitaire du C.P.O. appartient à cette base.

C.P.O. ω-algébrique :

· Un C.P.O. est dit "ω-algébrique" s'il admet une base finitaire dénombrable.

Soit E un C.P.O. ω-algébrique de base finitaire B alors $E = B^{\infty} = B \cup B^{\omega}$, $B \cap B^{\omega} = \phi$ où B^{ω} est l'ensemble des éléments non finitaires ou "purement infinitaires", c'est-à-dire l'ensemble des sup de suites infinies strictement croissantes.

Exemples :

- X^{∞} (muni de l'ordre préfixe) est un C.P.O. ω-algébrique de base finitaire X^{*}.

- $T^{\infty}_{\Omega}(\Sigma)$ (muni de l'ordre syntaxique) est un C.P.O. ω-algébrique de base finitaire $T_{\Omega}(\Sigma)$.

- L'ensemble des représentations décimales des réels n'est pas un C.P.O. ω-algébrique.

4. Feuillage d'arbres finis

Soit Σ un alphabet gradué ; Σ_0 est l'ensemble des symboles de Σ de arité nulle (constantes).

L'opération feuillage est une application $fg : T(\Sigma) \to \Sigma_0^{*}$ qui peut être définie récursivement de la façon suivante :

$\forall t \in T(\Sigma)$,

- si $t \in \Sigma_0$, $fg(t) = t$

- si $t = f(t_1, \ldots, t_p)$, $f \in \Sigma_p$, $t_1, \ldots, t_p \in T(\Sigma)$

$$fg(t) = fg(t_1)\ fg(t_2) \ldots\ fg(t_p).$$

Soit $t \in T(\Sigma)$, $t : dom(t) \to \Sigma$.

Soit $fr(t) = \{w \in dom(t) \,/\, t(w) \in \Sigma_0\}$.

i.e. $fr(t)$ est l'ensemble des mots du domaine correspondant aux feuilles de l'arbre t.

Alors si $u_1, u_2, \ldots, u_k$ est la suite ordonnée, suivant l'ordre lexicographique, des mots de $fr(t)$ on a :

$$fg(t) = t(u_1)\, t(u_2) \ldots t(u_k)$$

Exemple :

Soit l'arbre t défini par :

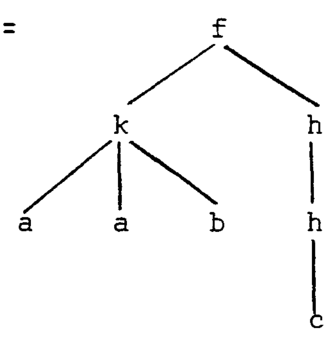
$dom(t) = \{\varepsilon, 1, 2, 11, 12, 13, 21, 211\}$

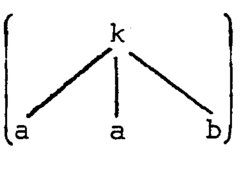
et $t(\varepsilon) = f$, $t(1) = k$

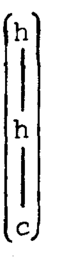
$t(11) = t(12) = a$, $t(13) = b$, $t(2) = t(21) = h$

$t(211) = c$.

i.e. $t =$



$fg(t) = fg$ 

$\cdot\, fg$ 

$= \ldots = aabc.$

$fr(t) = \{11, 12, 13, 211\}$.

$fg(t) = t(11) \cdot t(12) \cdot t(13) \cdot t(211) = aabc.$

II. FEUILLAGE D'ARBRES INFINIS

Nous voulons étendre aux arbres infinis l'opération de feuillage. L'extension ne peut se faire par continuité à fg : $T_\Omega^\infty(\Sigma) \to \Sigma^\infty$ du fait de la non compatibilité de l'ordre préfixe sur les mots avec la concaténation et avec l'ordre syntaxique sur les arbres. Nous allons donc établir le cahier des charges d'une opération de feuillage généralisée, puis construire l'objet libre répondant à ces spécifications. Nous comparons ensuite le résultat obtenu avec les notions existantes, en particulier avec la frontière d'arbres infinis introduite par Courcelle [6].

1. Spécifications

Nous voulons obtenir une application continue du C.P.O. des arbres dans un autre C.P.O, ensemble de mot qui permette la concaténation de mots infinis.

Il faut donc une application ϕ : $T_\Omega^\infty(\Sigma) \to M$, M un ensemble de mots, vérifiant :

- a) ϕ est stricte i.e. $\phi(\Omega) = \Omega$

- b) ϕ est croissante i.e. $\forall t, t' \in T_\Omega^\infty(\Sigma)$
  $t < t'$ implique $\phi(t)$ plus petit que $\phi(t')$

- c) ϕ est continue i.e. $\forall t \in T_\Omega^\infty(\Sigma)$
  si $t = \sup(D)$, D partie dirigée de $T_\Omega^\infty(\Sigma)$
  alors $\phi(t) = \sup(\phi(D))$
  où $\phi(D) = \{\phi(x) / x \in D\}$ est dirigée par b)

- d) ϕ vérifie la propriété de feuillage, c'est-à-dire :
  $\forall t \in T_\Omega^\infty(\Sigma)$, si $t = \begin{array}{c} f \\ \diagup \quad \diagdown \\ t_1 \; \cdots\cdots \; t_p \end{array}$

  $f \in \Sigma_p$; $t_1, \ldots, t_p \in T_\Omega^\infty(\Sigma)$
  alors $\phi(t) = \phi(t_1) \bullet \cdots \bullet \phi(t_p)$

Ces quatre propriétés impliquent certaines contraintes sur la nature et la structure de l'ensemble M d'arrivée :

- 1) M muni d'une relation d'ordre ($\leq$) est un C.P.O. de plus petit élément Ω (par a), b) et c).)

- 2) M muni d'une opération de concaténation ($\bullet$) est un monoïde (par d).)

- 3) La relation d'ordre est compatible avec la concaténation i.e. $\forall x, y, z, t \in M$, ($x \leq y$ et $z \leq t$) implique $x \bullet z \leq y \bullet t$ (par b), c) et d))

- 4) La concaténation est continue i.e. $\forall D, D'$ parties dirigées de M $\mathrm{Sup}(D) \bullet \mathrm{Sup}(D') = \mathrm{Sup}(D \bullet D')$ où $D \bullet D'$ est la partie dirigée (par 3).) $= \{x \bullet y \mid x \in D \;\&\; y \in D'\}$ (par c) et d).)

__Définition__ :

On appelle **"feuillage"** une application ϕ vérifiant a), b), c) et d), de $T^{\infty}_{\Omega}(\Sigma)$ dans un ensemble M vérifiant les conditions 1), 2), 3) et 4).

__Propriété immédiate__ :

Pour tout arbre fini t de $T_{\Omega}(\Sigma)$, si $u_1, \ldots, u_k$ est la suite ordonnée (lexicographiquement) des mots de fr(t) ($\subseteq$ dom(t)) alors on a : pour tout feuillage $\phi : T^{\infty}_{\Omega}(\Sigma) \to M$

$$\phi(t) = \phi(t(u_1)) \bullet \ldots \bullet \phi(t(u_k))$$

__Conséquence__ :

Tout feuillage est complètement défini par ses valeurs sur Σ_0.

2. Construction du feuillage initial

2.1. Construction du co-domaine

Soit un alphabet X ne contenant pas le symbole Ω. Considérons le monoïde libre $(X \cup \{\Omega\})^*$; les contraintes de réflexivité, compatibilité avec la concaténation et sur le rôle particulier de Ω conduisent à la définition de la relation suivante :

$\forall x, y \in (X \cup \{\Omega\})^*$

$$\underline{x \leq_\Omega y} \text{ ssi } \left| \begin{array}{l} - \; x \in X^* \;\&\; y = x \\ \text{ou} \\ - \; x = x_0 \Omega x_1 \ldots \Omega x_n, \text{ les } x_i \in X^* \\ \text{et } y = x_0\, y_1\, x_1 \ldots\, y_n\, x_n, \text{ les } y_j \in (X \cup \{\Omega\})^* \\ n \in N_+ \end{array} \right.$$

Il est immédiat que $\leq_\Omega$ est un pré-ordre sur $(X \cup \{\Omega\})^*$ vérifiant :

- a) $\forall x \in (X \cup \{\Omega\})^*$, $\Omega \leq_\Omega x$

- b) $\leq_\Omega$ est compatible avec la concaténation

<u>Propriété</u> :

$\leq_\Omega$ est le plus petit pré-ordre sur $(X \cup \{\Omega\})^*$ vérifiant les propriétés a) et b) définies ci-dessus.

<u>Preuve</u> :

Soit $<$ un pré-ordre vérifiant a) et b) ; $\forall x, y \in (X \cup \{\Omega\})^*$ tels que $x \leq_\Omega y$

- si $x \in X^*$ alors $x = y$ et $x < y$ (réflexivité)

- sinon $x = x_0\, \Omega\, x_1\, \Omega \ldots \Omega\, x_n$, $x_i \in X^*$
  $y = x_0\, y_1\, x_1 \ldots\, y_n\, x_n$, $y_j \in (X \cup \{\Omega\})^*$

On a $\forall i = 0, 1, \ldots, n \qquad x_i < x_i$

et ((a)) $\forall j = 1, \ldots, n \qquad \Omega < y_j$

d'où (par (b)), $x < y$.

Remarque :

$\leq_\Omega$ n'est pas antisymétrique et n'est donc pas un ordre.

Il suffit de noter que : $\Omega \leq_\Omega \Omega\Omega$ et $\Omega\Omega \leq_\Omega \Omega$ puisque $\Omega \leq_\Omega \varepsilon$ (le mot vide).

Passage à l'ordre :

Soit $\simeq$ la congruence sur $(X_\cup\{\Omega\})^*$ définie par $\Omega\Omega \simeq \Omega$; on a immédiatement :

$\forall x, y \in (X_\cup\{\Omega\})^*$, $x \simeq y$ ssi ($x \leq_\Omega y$ et $y \leq_\Omega x$).

Par conséquence, la relation $\leq_\Omega$ est un ordre (partiel) sur l'ensemble des classes de congruence : $(X_\cup\{\Omega\})^*/\simeq$, et Ω (plus précisément la classe de Ω : $\Omega\Omega^*$) est le plus petit élément.

Une classe de congruence est un langage de la forme suivante :

$x_0\ \Omega^+\ x_1\ \Omega^+ \ldots x_{n-1}\ \Omega^+\ x_n,\ n \in N$

où : x_0 et $x_n \in X^*$
$x_j \in X^+$ pour $j = 1, \ldots, n-1$.

Le représentant canonique d'une classe est celui des mots comportant le moins d'occurrences du symbole Ω i.e. $x_0\ \Omega\ x_1\ \Omega \ldots x_{n-1}\ \Omega\ x_n$.

La concaténation sur $(X_\cup\{\Omega\})^*/\simeq$ découle directement de la concaténation sur le monoïde libre et du point de vue des représentants canoniques tout se passe comme si on avait modifié la concaténation en posant $\Omega\Omega = \Omega$.

La classe du mot vide (ε) (réduite à lui-même) est l'élément neutre et $(X_\cup\{\Omega\})^*/\simeq$ un monoïde.

Le passage à l'ordre préserve la propriété de compatibilité de la relation $\leq_\Omega$ avec la concaténation.

Notations :

On notera W(X) l'ensemble des classes de congruence : $(X_\cup\{\Omega\})^*/\simeq$, et, sauf lorsqu'il sera nécessaire de faire la distinction, on confondra les mots et les classes auxquelles ils appartiennent.

Complétion :

Du P.O. : $\langle W(X), \leq_\Omega, \Omega\rangle$, par complétion, on obtient le C.P.O. : $\langle W^\infty(X), \leq_\Omega, \Omega\rangle$.

Par construction W(X) est une base du C.P.O. ; cette base est dénombrable (puisque $(X_\cup\{\Omega\})^*$ l'est).

Propriété :

$W^\infty(X)$ est un C.P.O. ω-algébrique.

Preuve :

Il suffit de montrer que tout élément de la base W(X) est finitaire.

Pour cela, il suffit de remarquer que tout mot fini (de W(X)) a un nombre fini de minorants, et donc ne peut être le sup d'une suite infinie strictement croissante.

Conséquence :

$W^\infty(X) = W(X) + W^\omega(X)$, $W(X) \cap W^\omega(X) = \phi$ où $W^\omega(X)$ est l'ensemble des éléments non finitaires appelés "termes infinis".

Extension de la concaténation :

Puisque $W^\infty(X)$ est un C.P.O. ω-algébrique, pour tout terme infini x il existe une partie dirigée D d'éléments de la base finitaire W(X) telle que $x = \mathrm{Sup}(D)$.

D'autre part, l'ordre étant compatible avec la concaténation sur W(X), pour toutes parties dirigées D, D' de W(X) l'ensemble $D \bullet D' = \{uv / u \in D \ \& \ v \in D'\}$ est une partie dirigée de W(X), d'où :

Définition :

$\forall x, y \in W^\infty(X)$,

$$xy = \mathrm{Sup}(D \bullet D') \quad \text{si } x = \mathrm{Sup}(D) \text{ et } y = \mathrm{Sup}(D').$$

Cette extension, par continuité, confère à $W^\infty(X)$ une structure de monoïde (élément neutre ε : le mot vide) et préserve la propriété de compatibilité de l'ordre avec la concaténation.

Exemples de termes infinis :

$w_1 = \mathrm{Sup}\{a^n \Omega \mid n \in N\}$;

$w_2 = \mathrm{Sup}\{\Omega a^n \mid n \in N\}$; $w_2 \neq w_1$

$w_3 = \mathrm{Sup}\{\Omega a^n \Omega \mid n \in N\}$; $w_3 \leq_\Omega w_1$ et $w_3 \leq_\Omega w_2$

$w_4 = \mathrm{Sup}\{a^n \Omega b^p \mid n, p \in N\}$; $w_1 \leq_\Omega w_4$

$w_5 = \mathrm{Sup}\{\Omega a^n b^p \Omega \mid n, p \in N\}$; $w_3 \leq_\Omega w_5$

$w_6 = \mathrm{Sup}\{(a^n \Omega b^p)^q \mid n, p, q \in N\}$; $w_1 \leq_\Omega w_6$

$w_7 = \mathrm{Sup}\{(a^n \Omega)^p \mid n, p \in N\} = w_1$.

Eléments maximaux de $W^{\infty}(X)$; **X fini, fixé :**

De par la définition de la relation d'ordre ($\leq_{\Omega}$), tout mot de X^* (i.e. ne comportant pas le symbole Ω) est maximal dans $W(X)$ et donc dans $W^{\infty}(X)$.

Soit $\rho : W(X) \to (X_{\cup}\{\Omega\})^*$ l'application qui à tout élément de $W(X)$ (une classe de congruence de mots) associe son représentant canonique dans $(X_{\cup}\{\Omega\})^*$.

Notons :

- $X^{\omega} = \{Sup(D) \mid D$ partie dirigée infinie
  $\& \; \forall w \in D,\ \rho(w) \in X^*\Omega\}$

(on retrouve les mots infinis "classiques", voir II.3.1.).

- $X^{-\omega} = \{Sup(D) \mid D$ partie dirigée infinie
  $\& \; \forall w \in D,\ \rho(w) \in \Omega X^*\}$.

- $M = \{w \in W(X) \mid \rho(w) \in \Omega X^* \Omega\}$
  M est une partie dirigée et a donc un Sup : $m = Sup(M)$.

Propriété :

$X^* \cup X^{\omega} \cdot \{m\} \cdot X^{-\omega}$ est l'ensemble des éléments maximaux de $W^{\infty}(X)$ où $X^{\omega} \cdot \{m\} \cdot X^{-\omega}$ dénote l'ensemble des termes infinis de la forme $u_1 \cdot m \cdot u_2$, $u_1 \in X^{\omega}$, $u_2 \in X^{-\omega}$.

Preuve :

Montrons tout d'abord que tout élément de $X^{\omega} \cdot \{m\} \cdot X^{-\omega}$ est maximal.

Soit $u = u_1 \cdot m \cdot u_2$, $u_1 \in X^{\omega}$, $u_2 \in X^{-\omega}$
$u_1 = Sup(D_1)$ avec $\rho(D_1) \subseteq X^*\Omega$
$u_2 = Sup(D_2)$ avec $\rho(D_2) \subseteq \Omega X^*$.

Soit $w \in W^{\infty}(X)$ tel que $u \leq_{\Omega} w$.

Il existe une partie dirigée $D \subseteq W(X)$ telle que $w = Sup(D)$ ($W^{\infty}(X)$ C.P.O. ω-algébrique).

Il faut montrer que $w \leq_{\Omega} u$ et pour cela que : $\forall x \in D, \exists y \in D_1 \cdot M \cdot D_2$ tel que $x \leq_{\Omega} y$.

On a $x \leq_{\Omega} y$ ssi $\rho(x) \leq_{\Omega} \rho(y)$ dans $(X_{\cup}\{\Omega\})^{*}$

$\rho(x)$ est de la forme suivante :

$$\rho(x) = x_0 \; \Omega \; x_1 \; \Omega \ldots \; x_{n-1} \; \Omega \; x_n$$

avec $x_0, x_n \in X^{*}$
$x_1, \ldots, x_{n-1} \in X^{+}$

et $n > 0$ ($\rho(x)$ contient au moins une occurrence du symbole Ω, sinon x serait maximal et D serait finie et donc $u \not\leq_{\Omega} w = Sup(D) = x$ puisque u est purement infinitaire).

Le fait que $u \leq_{\Omega} w$ implique :

- $\exists v_1 \in D_1$ tel que $x_0\Omega \leq_{\Omega} \rho(v_1)$

- $\exists v_2 \in D_2$ tel que $\Omega x_n \leq_{\Omega} \rho(v_2)$

(sinon contradiction avec l'hypothèse $u \leq_{\Omega} w$).

D'autre part : $\Omega x_1 \; \Omega \ldots \Omega \; x_{n-1} \leq_{\Omega} x_1 x_2 \ldots x_{n-1} \in X^{*}$ puisque $\Omega \leq_{\Omega} \varepsilon$.

Or $y = v_1 \cdot x_1 x_2 \ldots x_{n-1} \cdot v_n \in D_1 \cdot M \cdot D_2$ et $x \leq_{\Omega} y$ puisque l'ordre est compatible avec la concaténation ; d'où $w \leq_{\Omega} x$ et donc $w = x$ ($x \leq_{\Omega} w$ par hypothèse) ce qui prouve que x est maximal.

Montrons maintenant que tout élément maximal non finitaire ($\notin X^*$) appartient à $X^\omega \cdot \{m\} \cdot X^{-\omega}$.

Soit $w \in W^\omega(X)$ tel que w soit maximal i.e. $\forall u \in W^\infty(X)$ on ait ($w \leq_\Omega u \Rightarrow u = w$).

Remarquons tout d'abord que $w \in X^\omega \cdot W^\infty(X) \cdot X^{-\omega}$.

En effet, soit D partie dirigée telle que $w = \mathrm{Sup}(D)$.

Soit $D_1 = \{x\Omega \in W(X) \mid \exists y \in D \;\&\; \rho(y) = x\Omega y',\ x \in X^*\}$

On a D_1 est une suite croissante de $W(X)$

(sinon D n'est pas dirigée) et D_1 est infinie
(sinon w n'est pas maximal) et $\mathrm{Sup}(D_1) \in X^\omega$ par définition de D_1.

Soit $D_2 = \{\Omega x \in W(X) \mid \exists y \in D \;\&\; \rho(y) = y''\Omega x,\ x \in X^*\}$ de la même façon D_2 est une suite infinie croissante de $W(X)$ et $\mathrm{Sup}(D_2) \in X^{-\omega}$.

Ceci montre qu'il existe une partie dirigée $D \subseteq \Omega \cdot W(X) \cdot \Omega$ telle que $w = \mathrm{Sup}(D_1) \cdot \mathrm{Sup}(D) \cdot \mathrm{Sup}(D_2) = \mathrm{Sup}(D_1 \cdot D \cdot D_2)$.

D'autre part pour toute partie dirigée $D \subseteq \Omega \cdot W(X) \cdot \Omega$ on a $\mathrm{Sup}(D) \leq_\Omega m = \mathrm{Sup}(M)$

d'où (w est maximal) $w = \mathrm{Sup}(D_1) \cdot m \cdot \mathrm{Sup}(D_2) \in X^\omega \cdot \{m\} \cdot X^{-\omega}$.

2.2. Définition du feuillage

On considère l'ensemble des arbres : $T_\Omega^\infty(\Sigma)$, et l'ensemble des mots $W^\infty(\Sigma_0)$.

Soit $\phi : T_\Omega^\infty(\Sigma) \rightarrow W^\infty(\Sigma_0)$, l'application satisfaisant aux spécifications (strictitude, croissance, continuité, propriété de feuillage) définie par :

$\forall x \in \Sigma_0 : \phi(x) = x.$

<u>Propriété</u> :

ϕ existe et est bien définie.

<u>Preuve</u> :

Par la construction explicite de ϕ.

a) <u>Pour les arbres finis</u> :

Soit $\phi : T_\Omega(\Sigma) \rightarrow W(\Sigma_0)$ définie par :

- $\phi(\Omega) = \Omega$
- $\forall x \in \Sigma_0, \phi(x) = x$
- $\forall t \in T_\Omega(\Sigma)$, si $t = f(t_1,\dots, t_p)$

$f \in \Sigma_p$; $t_1,\dots, t_p \in T_\Omega(\Sigma)$; $p \in N_+$

$\phi(t) = \phi(t_1) \bullet\dots\bullet \phi(t_p).$

On a immédiatement :

- $\forall t \in T_\Omega(\Sigma), \phi(t) \simeq fg(t)$
- ϕ est croissante

i.e. $\forall t, t' \in T_\Omega(\Sigma), t < t' \Rightarrow \phi(t) \leq_\Omega \phi(t').$

Cette propriété découle directement des définitions de l'ordre syntaxique sur les arbres ($<$) et de celui sur les mots ($\leq_\Omega$).

<u>Conséquence</u> :

. Pour toute partie dirigée Δ d'arbres de $T_\Omega(\Sigma)$, $\phi(\Delta) = \{\phi(t) \mid t \in \Delta\}$ est une partie dirigée de $W(\Sigma_0)$.

b) <u>Extension aux arbres infinis</u> :

ϕ est étendu à $T_\Omega^\infty(\Sigma) \to W^\infty(\Sigma_0)$ par continuité (grâce à la remarque précédente) en posant :

$\forall t \in T_\Omega^\infty(\Sigma)$, $\phi(t) = \mathrm{Sup}(\phi(\Delta))$
si $t = \mathrm{Sup}(\Delta)$.

<u>Propriété de ϕ</u> :

- ϕ est stricte par définition

- ϕ est croissante (l'extension par continuité conserve cette propriété)

- ϕ est continue par définition

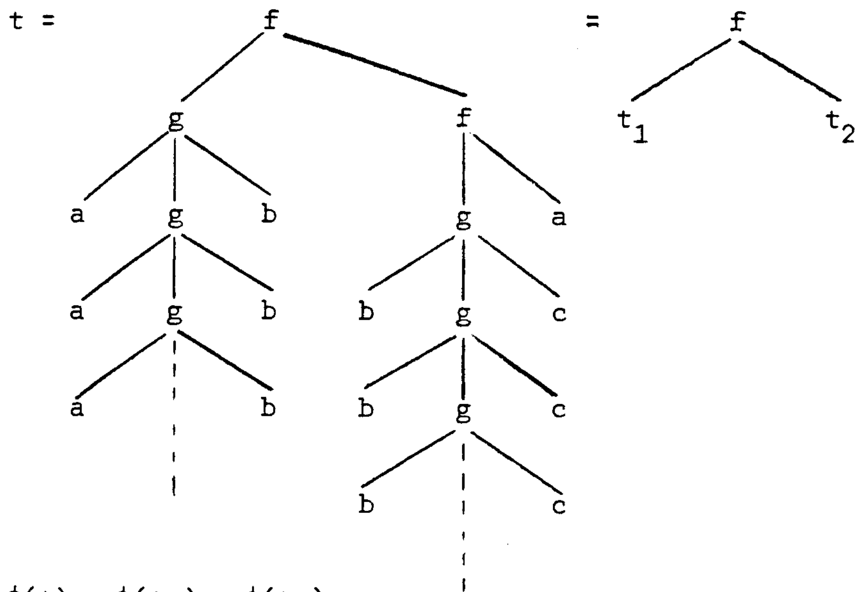
- $\phi(f(t_1,\ldots, t_p)) = \phi(t_1) \cdot \ldots \cdot \phi(t_p)$

  $\forall p \in N_+$, $\forall f \in \Sigma_p$, $\forall t_1,\ldots, t_p \in T_\Omega^\infty(\Sigma)$ (même remarque)

d'où : $\phi : T_\Omega^\infty(\Sigma) \to W^\infty(\Sigma_0)$ est un feuillage.

<u>Exemple</u> :

Soit l'arbre infini t schématisé par



$\phi(t) = \phi(t_1) \bullet \phi(t_2)$.

En considérant des approximations finies de t_1 et t_2, (par exemple les tronqués) on obtient :

$$\phi(t_1) = \sup\{a^n \,\Omega\, b^n \mid n \in N\} = \sup(a^* \,\Omega\, b^*)$$

$$\phi(t_2) = \sup\{b^p \,\Omega\, c^p\, a \mid p \in N\} = \sup(b^* \,\Omega\, c^*\, a)$$

et donc $\phi(t) = \sup(a^* \,\Omega\, b^*) \bullet \sup(b^* \,\Omega\, c^*\, a)$
$\qquad\qquad = \sup(a^* \,\Omega\, b^* \,\Omega\, c^*\, a)$.

<u>Théorème 1</u> :

Le feuillage $\phi : T^\infty_\Omega(\Sigma) \to W^\infty(\Sigma_0)$ est initial i.e. pour tout feuillage $\phi_M : T^\infty_\Omega(\Sigma) \to M$ il existe un et un seul morphisme de monoïde continu h tel que $\phi_M = h \circ \phi$.

<u>Preuve</u> :

Soit $\phi_M : T^\infty_\Omega(\Sigma) \to M$ un feuillage.

Notons $\bullet$: la concaténation dans M
λ : l'élément neutre de M
$<$: l'ordre sur M.

<u>Lemme</u> :

$\forall t, t' \in T_\Omega(\Sigma)$

$\phi(t) = \phi(t')$ implique $\phi_M(t) = \phi_M(t')$.

Il suffit de remarquer que dans M on a : $\Omega \bullet \Omega = \Omega$ (puisque $\Omega < \lambda$ et $<$ est compatible avec la concaténation $\bullet$) et que $\phi_M(\Omega) = \Omega$.

Soit $h : W(\Sigma_0) \to M$ défini par :

1) $h(\Omega) = \Omega$
2) $h(\varepsilon) = \lambda$
3) $\forall x \in \Sigma_0$, $h(x) = \phi_M(x)$
4) $\forall u, v \in W(X)$, $h(uv) = h(u) \bullet h(v)$

Ω étant le plus petit élément de M et l'ordre ($<$) étant compatible avec la concaténation ($\bullet$), il est facile de vérifier que h est croissante ce qui permet son extension par continuité à :

$h : W^\infty(X) \to M$ en posant (M est un C.P.O.)

$\forall w \in W^\infty(X)$, $h(w) = \mathrm{Sup}(h(\Delta))$ si $w = \mathrm{Sup}(\Delta)$.

L'application h, ainsi étendue, est un morphisme de monoïde continu (par définition) et on a :

$\forall t \in T^\infty_\Omega(\Sigma)$, $\phi_M(t) = h(\phi(t))$.

En effet pour tout arbre fini $t \in T_\Omega(\Sigma)$ si $u_1, \ldots, u_k$ est la suite ordonnée lexicographiquement des mots de fr(t) (mots des noeuds terminaux du domaine de t) on a : $\phi(t) = t(u_1) \bullet \ldots \bullet t(u_k)$
d'où $h(\phi(t)) = h(t(u_1)) \bullet \ldots \bullet h(t(u_k))$, puis les $t(u_i)$ appartenant à Σ_0 :

$h(\phi(t)) = \phi_M(t(u_i)) \bullet \ldots \bullet \phi_M(t(u_k)) = \phi_M(t)$.

D'autre part, pour tout arbre infini $t \in T^{\infty}_{\Omega}(\Sigma)$, t est le Sup d'une partie dirigée D d'arbres finis et $\Phi(t) = \mathrm{Sup}(\Phi(D))$ par définition de Φ $\Phi_M(t) = \mathrm{Sup}(\Phi_M(D))$ par propriété de Φ_M d'où $h(\Phi(t)) = \mathrm{Sup}(h(\Phi(D)))$ par définition de h.

D étant un ensemble d'arbres finis : $h(\Phi(D)) = \Phi_M(D)$ par le résultat précédent, ce qui montre :

$$h(\Phi(t)) = \mathrm{Sup}(\Phi_M(D)) = \Phi_M(t).$$

Quant à l'unicité de h, elle est évidente puisque h est entièrement définie par ses valeurs sur Σ_0.

Remarques :

a) $\Phi : T^{\infty}_{\Omega}(\Sigma) \to W^{\infty}(\Sigma_0)$ est une surjection i.e. $\forall w \in W^{\infty}(\Sigma_0)$, $\exists t \in T^{\infty}_{\Omega}(\Sigma)$ tel que $\Phi(t) = w$.

b) Considérant les arbres localement finis on note que $\{\Phi(t) \mid t \in T^{Loc}(\Sigma)\}$ est strictement inclus dans $W^{\infty}(\Sigma_0)$. Cet ensemble sera noté $W^{Loc}(\Sigma_0)$ et ses éléments seront appelés termes localement finis.

3. Comparaison avec les notions existantes

3.1. Les mots infinis

Soit un alphabet X ; on considère $X^{\infty} = X^{*} \cup X^{\omega}$ l'ensemble des mots sur X muni de l'ordre préfixe ("facteur gauche de") ; c'est un C.P.O. ω-algébrique de base finitaire X^{*} et de plus petit élément ε (le mot vide). X^{∞} peut être identifié à une partie de $W^{\infty}(X)$.

En effet, soit $b : X^{\infty} \to W^{\infty}(X)$ défini par :

- $\forall w \in X^{*}$, $b(w) = w$
- $\forall w \in X^{\omega}$, $b(w) = \mathrm{Sup}\{w[n]\,\Omega \mid n \in N\}$

où w[n] dénote le facteur gauche de longueur n de w.

Remarque :

Dans X^∞ avec l'ordre préfixe on a :

$$w = \mathrm{Sup}\{w[n] \mid n \in N\}$$

d'où $\{w[n]\ \Omega \mid u \in N\}$ est une suite croissante de $W(X)$.

On a immédiatement : b est injective et
$b(X^\infty) = X^* \cup \{w \in W^\infty(X) \mid \exists D \subseteq X^*\ \Omega \text{ et } w = \mathrm{Sup}(D)\}$.

Feuillage avec les mots infinis :

Afin d'obtenir les propriétés nécessaires à un feuillage il nous faut modifier quelque peu l'approche des mots infinis, reprenant pour cela celle de S. Tison [19].

Considérons les mots de $M = X^\infty \cup X^*\ \Omega$ avec l'ordre suivant : $\forall u, v \in M$.

$$u < v \text{ ssi } \left| \begin{array}{ll} - \ u \in X^\infty \text{ et } u = v & \\ \text{ou} & \\ - \ u = u_1\ \Omega & u_1 \in X^* \\ \quad v = u_1 v_1 & v_1 \in M \end{array} \right.$$

M muni de cet ordre est un C.P.O. ω-algébrique de base finitaire $X^* \cup X^*\ \Omega$ et de plus petit élément Ω.

Concaténation dans M : $(*)$ $\forall u, v \in M$

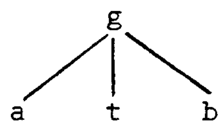
$$u * v = \left| \begin{array}{l} uv \text{ si } u \in X^\infty \text{ (concaténation usuelle)} \\ \text{ou} \\ u \text{ si } u \in X^*\ \Omega \end{array} \right.$$

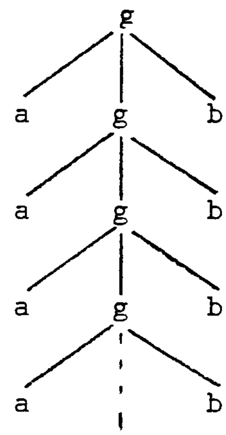
L'ordre sur M est compatible avec la concaténation et la concaténation est continue. Ceci permet de définir le feuillage $f : T^\infty_\Omega(\Sigma) \to M$ (avec $X = \Sigma_0$)
par - $f(\Omega) = \Omega$
- $\forall x \in \Sigma_0,\ f(x) = x$
et - $\forall t \in T^\infty_\Omega(\Sigma)$, si $t = g(t_1, \ldots, t_p)$
$f(t) = f(t_1) * \ldots * f(t_p)$.

Il est clair que par f on ne considèrera que les feuilles situées à gauche de la branche infinie la plus à gauche des arbres infinis.

Exemples :

1) Soit $t \in T^{\infty}(\Sigma)$ défini par :

t = 

allure de t : 

$t = \mathrm{Sup}\{t[n] \mid n \in N\}$

où t[n] est le tronqué à profondeur n de t

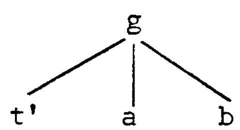
d'où $f(t) = \mathrm{Sup}\{f(t[n]) \mid u \in N\}$

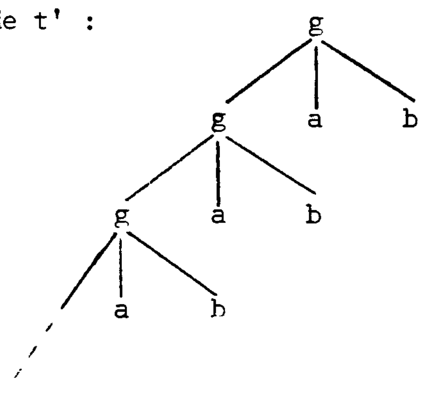
$= \mathrm{Sup}\{a^n * \Omega * b^n \mid n \in N\} = \mathrm{Sup}\{a^n\ \Omega \mid u \in N\}$

$= a^{\omega} \in X^{\infty}$

et $\Phi(t) = \mathrm{Sup}\{a^n\ \Omega\ b^n \mid n \in N\} \in W^{\infty}(X)$

2) Soit $t' \in T^{\infty}_{\Omega}(\Sigma)$ défini par :

t' = 

allure de t' : 

$t' = \mathrm{Sup}\{t'[n] \mid n \in N\}$

d'où $f(t') = \mathrm{Sup}\{f(t'[n]) \mid u \in N\}$
$= \mathrm{Sup}\{\Omega * (ab)^n \mid n \in N\}$
$= \mathrm{Sup}\{\Omega\} = \Omega$

et $\phi(t') = \mathrm{Sup}\{\Omega(ab)^n \mid n \in N\} \in W^\infty(\Sigma_0)$.

Soit $h : W^\infty(\Sigma_0) \to M$ le morphisme de monoïde continu défini par $h(\Omega) = \Omega$ et $\forall x \in \Sigma_0$, $h(x) = x$. On a bien $f(t) = h(\phi(t))$ et $f(t') = h(\phi(t'))$.

3.2. Les mots bi-infinis (Perrin - Nivat [15]) :

Rappel des définitions :

X dénotant toujours un alphabet, on considère les "bimots" ou couples de mots (u, v) du produit cartésien $X^\infty \times X^\infty$.

Sur cet ensemble est défini une relation d'équivalence notée $\sim$: $\forall (u, v), (u', v') \in X^\infty \times X^\infty$

$(u, v) \sim (u', v')$ ssi il existe un mot $w \in X^*$

tel que : $\begin{cases} v = wv' \text{ et } u' = w^R u \\ \text{ou} \\ v' = wv \text{ et } u = w^R u' \end{cases}$

On appelle mot "bilatère" la classe d'équivalence d'un bimot et on note ${}^\infty X^\infty$ l'ensemble des mots bilatères.

Un bimot (u, v) peut être :

- fini si u et $v \in X^*$
- infini à droite si $u \in X^*$ et $v \in X^{\omega}$
- infini à gauche si $u \in X^{\omega}$ et $v \in X^*$
- bi-infini si u et $v \in X^{\omega}$

Chacune de ces familles de bimots est saturée par l'équivalence $\sim$, ce qui permet la même classification des mots bilatères.

L'ensemble des mots bilatères finis et identifié à X^* par : $\forall (u, v) \in X^* \times X^*$, $(u, v) \to u^R v \in X^*$ (on a bien $(u, v) \sim (u', v')$ ssi $u'^R v' = u^R v$).

De la même façon, l'ensemble des mots bilatères infinis à droite est identifié à X^{ω} par :

$$\forall (u, v) \in X^* \times X^{\omega},\ (u, v) \to u^R v \in X^{\omega}.$$

L'ensemble des mots bilatères infinis à gauche est mis en bijection avec X^{ω} par :

$$\forall (u, v) \in X^{\omega} \times X^*,\ (u, v) \to v^R u \in X^{\omega}.$$

<u>Relation avec $W^{\infty}(X)$</u> :

L'ensemble des mots bilatères ${}^{\infty}X^{\infty}$ peut être identifié à une partie de $W^{\infty}(X)$. En effet :

Soit $b : X^{\infty} \times X^{\infty} \to W^{\infty}(X)$ définie par :

$\forall (u, v) \in X^{\infty} \times X^{\infty}$

- si u et $v \in X^*$, $b((u, v)) = u^R v \in X^* \subset W^{\infty}(X)$

- si $u \in X^*$ et $v \in X^{\omega}$, $(v = \mathrm{Sup}\{v[n] \mid n \in N\})$
  $b((u, v)) = \mathrm{Sup}\{u^R v[n]\, \Omega \mid n \in N\}$

- si $u \in X^{\omega}$ et $v \in X^*$, $(u = \mathrm{Sup}\{u[n] \mid n \in N\})$
  $b((u, v)) = \mathrm{Sup}\{\Omega (u[n])^R\, v \mid n \in N\}$

- si u et v $\in X^{\omega}$,
  $b((u, v)) = \mathrm{Sup}\{\Omega(u[n])^{R} \cdot v[n]\ \Omega \mid n \in N\}$.

On a immédiatement les propriétés suivantes :

a) $\forall (u, v), (w, t) \in X^{\infty} \times X^{\infty}$

$(u, v) \sim (w, t)$ ssi $b(u, v) = b(w, t)$

ce qui permet de définir l'application

$\overline{b} : {}^{\infty}X^{\infty} \to W^{\infty}(X)$ par $\forall \alpha \in {}^{\infty}X^{\infty}$

$\overline{b}(\alpha) = b((u, v))$ pour $(u, v) \in \alpha$

et d'obtenir :

b) $\overline{b} : {}^{\infty}X^{\infty} \to W^{\infty}(X)$ est injective

et $\overline{b}({}^{\infty}X^{\infty}) = X^{*} + \{w \in W^{\infty}(X) \,/\, \exists D \subseteq X^{*}\ \Omega + \Omega\ X^{*} + \Omega\ X^{*}\ \Omega\ \&\ w = \mathrm{Sup}(D)\}$.

3.3. Frontières d'arbres infinis et arrangements

La notion de "frontière" d'arbre infini, introduite par Courcelle [6] repose sur la définition des arrangements ou "mots infinis généralisés". Après en avoir rappelé les définitions et propriétés essentielles, nous verrons que feuillages et frontières ne coïncident pas.

3.3.1. Arrangements dénombrables

Soit X un alphabet fini.

Un arrangement u sur l'alphabet X est un triplet $<D, \Pi, h>$ où :

et
- D est un ensemble
- Π une relation d'ordre total sur D
- h une application de D dans X

- u est dit dénombrable ssi D est dénombrable.

- $A_\omega(X)$ dénote l'ensemble des arrangements dénombrables.

- Les mots de X^* sont identifiés avec les arrangements de la forme : $<[n], \leq, h>$ où $[n] = \{1,\ldots, n\}$, $\leq$ l'ordre sur les entiers.

Exemple :

$\forall w \in X^* \quad w = a_1 a_2 \ldots a_n$

$n \in N_+, \; a_i \in X \quad \forall i \in N$

w est identifié à l'arrangement $<[n], \leq, h>$ avec $h : [n] \to X$, $h(i) = a_i$.

Les arrangements de la forme $<N_+, \leq, h>$ correspondent aux mots infinis de X^ω.

Relation d'équivalence :

Soient $u = <D_u, \Pi_u, h_u>$ et $v = <D_v, \Pi_v, h_v>$ deux arrangements.

On dira que u est équivalent à v et on notera $u \equiv v$ ssi :

$\exists f : D_u \to D_v$ vérifiant

- f est bijective
- f préserve l'ordre i.e. $\forall x, y \in D_u$
  $f(x) \; \Pi_v \; f(y) \iff x \; \Pi_u \; y$

et - $\forall x \in D_u : h_v(f(x)) = h_u(x)$.

Concaténation (de manière simplifiée) :

Soient u et v définis comme précédemment avec $D_u \cap D_v = \phi$.

Alors $u \bullet v = <D, \Pi, h>$ avec

- $D = D_u \cup D_v$

- Π : • $\forall x \in D_u, \forall y \in D_v \quad x \Pi y$

  • Π restreint à $D_u = \Pi_u$

  • Π restreint à $D_v = \Pi_v$

- $h : D \to X$ tel que

  $\forall x \in D_u \quad h(x) = h_u(x)$

  $\forall y \in D_v \quad h(y) = h_v(y)$

Exemples et notations :

Soit $h : D \to X$ tel que $\forall i \in D \quad h(i) = a \in X$

- $\langle N, \leq, h\rangle = a^{\omega} \in X^{\omega}$
- $\langle N, \geq, h\rangle = a^{-\omega} \neq a^{\omega}$
- $\langle Z, \leq, h\rangle = a^{Z} = a^{-\omega} \cdot a^{\omega}$

3.3.2. Frontières d'arbres infinis

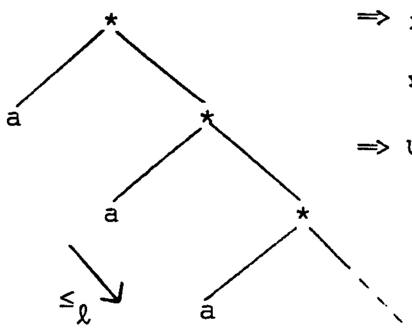
On considère $T^{\infty}(\Sigma)$ l'ensemble des arbres finis et infinis sur Σ avec $\Sigma_0 = X$.

Définition :

$\forall t \in T^{\infty}(\Sigma)$

frontière$(t) = \psi(t) = \langle fr(t), \leq_{\ell}, t\rangle$
où : $fr(t)$ = l'ensemble des mots du domaine de t ayant leur image dans Σ_0 ($\to$ feuilles de t) par $t : dom(t) \to \Sigma$ et $\leq_{\ell}$ l'ordre lexicographique sur les mots.

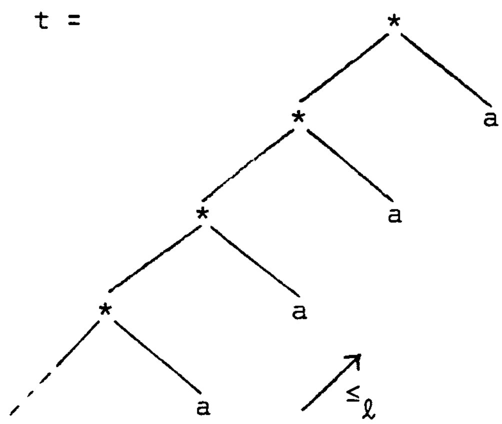
Exemples :

t = 

$\Rightarrow fr(t) = 1^*0$

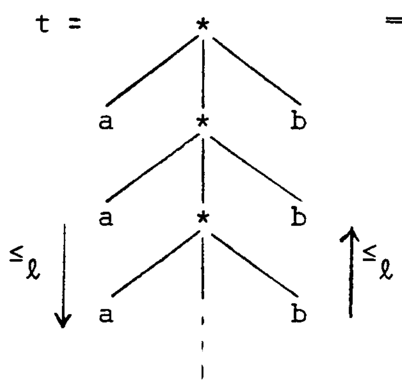
$\forall u \in fr(t) \quad t(u) = a$

$\Rightarrow \psi(t) = \langle 1^*0, \leq_\ell, t\rangle \equiv a^\omega$

t = 

$\Rightarrow fr(t) = 0^*1$

$\psi(t) = \langle 0^*1, \leq_\ell\, t\rangle \equiv a^{-\omega}$

t = 

$\Rightarrow fr(t) = 1^*0 + 1^*2$

$\psi(t) = \langle fr(t), \leq_\ell, t\rangle \equiv a^\omega \cdot b^{-\omega}$

Propriété :

· $\forall t \in T^{\infty}(\Sigma)$

si $t = f(t_1,\ldots, t_p)$, $f \in \Sigma_p$

$\psi(t) \equiv \psi(t_1) \bullet\ldots\bullet \psi(t_p)$.

Théorème (Courcelle [6]) :

$$A_\omega(\Sigma_0) = \{\psi(t) \mid t \in T^{\infty}(\Sigma)\}$$

$$= \{\varepsilon\} \cup \{\psi(t) \mid t \in T^{Loc}(\Sigma)\}.$$

3.3.3. Feuillage et frontière

Pour tout arbre fini t de $T(\Sigma)$ (pas de symbole Ω) on a :

$\psi(t) = fg(t) = \phi(t) \in \Sigma_0^*$

Pour tout arbre fini t de $T_\Omega(\Sigma)$:

$\psi(t) = fg(t) \in (\Sigma_0 \cup \{\Omega\})^*$.

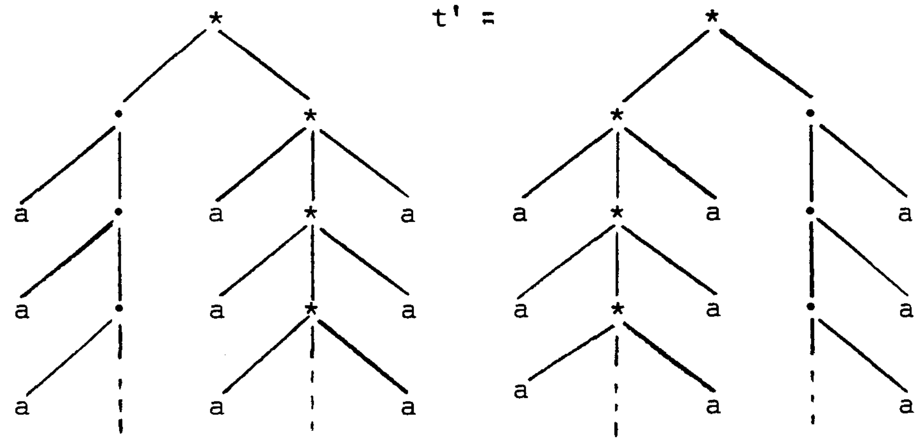
L'opération de frontière (ψ) n'est pas continue (sens C.P.O.) en effet :

Soient $\{t_i \mid i \in N\}$ et $\{t'_i \mid i \in N\}$ deux suites croissantes d'arbres finis de $T_\Omega(\Sigma)$ ayant leur Sup(t et t') dans $T^{\infty}(\Sigma)$, on a :

$(\forall i \in N, \psi(t_i) \equiv \psi(t'_i))$ n'implique pas $\psi(t) \equiv \psi(t')$.

Exemple (Courcelle [6]) :

$t =$ $\qquad$ $t' =$



$$\psi(t) \equiv a^{\omega} \bullet a^{\omega} \bullet a^{-\omega} \qquad \psi(t') = a^{\omega} \bullet a^{-\omega} \bullet a^{-\omega}$$

$$\psi(t) \not\equiv \psi(t')$$

et si on considère pour tout $n \in N$ les tronqués $t[n]$ et $t'[n]$ de t et t' à profondeur n on a :

$$\psi(t[n]) \equiv \psi(t'[n]) = a^n \, \Omega \, a^n \, \Omega \, a^n.$$

<u>Conséquence</u> :

$$\forall t, t' \in T^{\infty}(\Sigma) \text{ (et même } T^{Loc}(\Sigma))$$

$\phi(t) = \phi(t')$ n'implique pas $\psi(t) \equiv \psi(t')$.

Autrement dit deux arbres ayant même feuillage n'ont pas nécessairement même frontière.

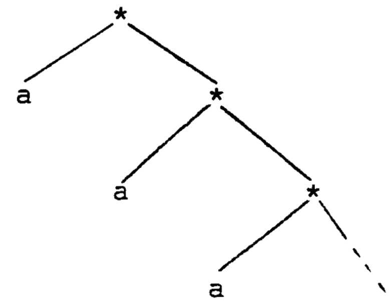
(Voir exemple précédent : $\phi(t) = \phi(t') = \mathrm{Sup}\{a^n \, \Omega \, a^n \, \Omega \, a^n \mid n \in N\}$.

Réciproquement dans le cas général on a :

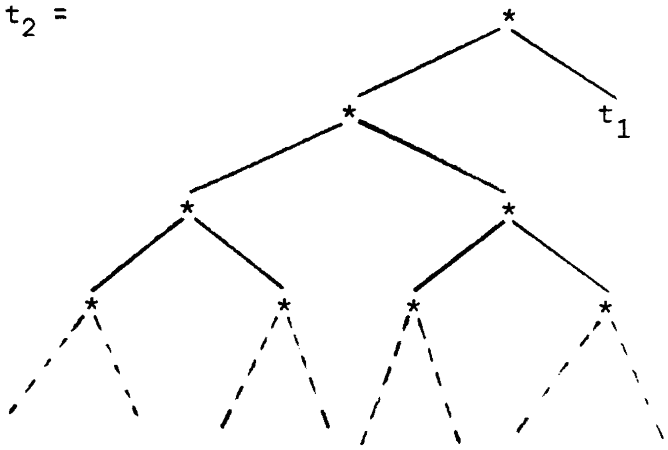
$$\forall t, t' \in T^{\infty}(\Sigma), \ \psi(t) \equiv \psi(t') \not\Rightarrow \phi(t) = \phi(t').$$

<u>Exemple</u> :

t_1 :



t_2 =



$\psi(t_1) \equiv a^\omega$ $\qquad\qquad$ $\psi(t_2) \equiv a^\omega$

et $\phi(t_1) = \mathrm{Sup}\{a^n\ \Omega \mid n \in N\} \neq \phi(t_2) = \mathrm{Sup}\{\Omega\ a^n\ \Omega \mid n \in N\}$.

Par contre, en se limitant aux arbres localement finis :

<u>Propriété</u> :

$\forall t, t' \in T^{Loc}(\Sigma)$

$\psi(t) \equiv \psi(t')$ implique $\phi(t) = \phi(t')$.

<u>Preuve</u> :

Soient t et u, deux arbres localement finis tels que $\psi(t) \equiv \psi(u)$ i.e. $<fr(t), \leq_\ell, t> \equiv <fr(u), \leq_\ell, u>$.

Par définition de l'équivalence des arrangements, il existe une bijection $h : fr(t) \to fr(u)$ vérifiant

- $\forall x, y \in fr(t) : h(x) \leq_\ell h(y)$ ssi $x \leq_\ell y$
- $\forall x \in fr(t) : u(h(x)) = t(x)$.

$T^{\infty}_{\Omega}(\Sigma)$ étant un C.P.O. de base finitaire $T_{\Omega}(\Sigma)$ il existe une suite croissante $\{t_i \mid i \in N\}$ de $T_{\Omega}(\Sigma)$ telle que $t = \mathrm{Sup}\{t_i \mid i \in N\}$ de même $u = \mathrm{Sup}\{u_j \mid j \in N\}$.

Par propriété du feuillage $\{\phi(t_i) \mid i \in N\}$ et $\{\phi(u_j \mid j \in N\}$ sont des suites croissantes de $W(\Sigma_0)$ et $\phi(t) = \mathrm{Sup}\{\phi(t_i) \mid i \in N\}$

$$\phi(u) = \mathrm{Sup}\{\phi(u_j) \mid j \in N\}.$$

Pour montrer que $\phi(t) = \phi(u)$, il suffit de montrer que ces deux suites sont équivalentes c'est-à-dire :

a) $\forall i \in N$, $\exists j \in N$ tel que $\phi(t_i) \leq_{\Omega} \phi(u_j)$ et réciproquement :

b) $\forall k \in N$, $\exists \ell \in N$ tel que $\phi(u_k) \leq_{\Omega} \phi(t_{\ell})$.

Pour tout arbre fini v de $T_{\Omega}(\Sigma)$ on considère :

- son domaine : Dom(v)
- ses noeuds terminaux (feuilles) : fr(v)

et - les feuilles "terminales"

$$ft(v) = \{x \in fr(v) \,/\, v(x) \neq \Omega\}$$

Il est clair que les feuilles terminales des approximants t_i de l'arbre t sont des feuilles de t i.e. $\forall i \in N$, $ft(t_i) \subset fr(t)$ et $\forall j > i$, $ft(t_i) \subset ft(t_j)$.

Ceci montre en particulier que $\forall i \in N$, $\forall x \in ft(t_i)$, $h(x)$ est défini et appartient à $fr(u)$.

Considérant que $T^{\infty}_{\Omega}(\Sigma)$ est un C.P.O. ω-algébrique et que $ft(t_i)$ est fini pour tout i et que la remarque précédente s'applique à l'arbre u, on obtient :

<u>**Lemme**</u> :

$\forall i \in N$, $\exists j \in N$ tel que

$h(ft(t_i)) \subseteq ft(u_j)$.

Il reste à montrer que pour tout i, $\phi(t_i) \leq_\Omega \phi(u_j)$ pour un j obtenu grâce au Lemme.

Le représentant canonique de $\phi(t_i)$ est un mot de la forme :

$$x_0 \,\Omega\, x_1 \,\Omega\, \dots\, x_{n-1} \,\Omega\, x_n,\ n \in N_+ \text{ avec } x_0, x_n \in \Sigma_0^*$$
$$x_1, \dots, x_{n-1} \in \Sigma_0^+$$

Il faut donc montrer que $\phi(u_j)$ est de la forme :

$$x_0\, y_1\, x_1 \dots x_{n-1}\, y_n\, x_n\ ;\ y_1, \dots, y_n \in (\Sigma_0 \cup \{\Omega\})^*.$$

<u>Remarque 1</u> :

Puisque $h(ft(t_i) \subset ft(u_j)$, h préserve l'ordre et $u(h(x)) = t(x)$ pour tout $x \in ft(t_i)$:

$\phi(u_j)$ contient et dans le même ordre toute la séquence de lettres terminales ($\in \Sigma_0$) de $\phi(t_i)$.

<u>Remarque 2</u> :

Puisque t et u sont localement finis et que h préserve l'ordre on a :

• si x est le plus petit élément (ordre lexicographique) de $fr(t_i)$ et $x \in ft(t_i)$ $(t(x) \neq \Omega)$ alors h(x) est le plus petit élément de $fr(u_j)$ et appartient à $ft(u_j)$ (cas où $x_0 \neq \varepsilon$).

• si x est le plus grand élément de $fr(t_i)$ et $x \in ft(t_i)$ alors h(x) est le plus grand élément de $fr(u_j)$ et $\in ft(u_j)$ (cas où $x_n \neq \varepsilon$).

• si y est le successeur de x (ordre lexicographique) dans $fr(t_i)$ et x et y appartiennent à $ft(t_i)$ alors h(y) est le successeur de h(x) dans $fr(u_j)$ et h(x) et h(y) appartiennent à $ft(u_j)$.

Ce dernier point montre que deux lettres terminales successives dans $\phi(t_i)$, le sont également dans $\phi(u_j)$ ce qui n'est pas forcément le cas si u n'est pas localement fini (il pourrait y avoir insertion de Ω).

Ces deux remarques montrent que $\phi(u_j)$ a bien la forme voulue i.e. $x_0\, y_1\, x_1 \ldots x_{n-1}\, y_n\, x_n$ et donc que $\phi(t_i) \leq_\Omega \phi(u_j)$.

Pour montrer que : $\forall k \in N$, $\exists \ell \in N$ tel que

$\phi(u_k) \leq_\Omega \phi(t_\rho)$

On procède exactement de la même manière en utilisant cette fois h^{-1} (h est bijective) ce qui prouve la propriété.

Conséquence de la propriété :

Il existe une application surjective s de l'ensemble des arrangements dénombrables dans l'ensemble des termes infinis "localement finis" (i.e. $s : A_\omega(\Sigma_0) \to W^{Loc}(\Sigma_0)$) vérifiant : $\forall t \in T^{Loc}(\Sigma)$, $\phi(t) = s(\psi(t))$.

Définition de s :

Soit u : un arrangement dénombrable sur un alphabet X.

Si u est fini (son domaine est fini), u est équivalent ($\equiv$) à un mot w de X^* et dans ce cas $s(u) = w$.

Soit u non fini, $u = \langle D, \Pi, f \rangle$.

Soit $\nu : N_+ \to D$ une énumération bijective de D.

Pour tout $n \in N_+$, considérons la suite croissante $i_1, \ldots, i_n$ des éléments de $[n] = \{1, 2, \ldots, n\}$ ordonnée selon l'ordre suivant : i plus petit que j ssi $\nu(i)\ \Pi\ \nu(j)$ dans D. Soit v_n le mot de W(X) défini de la façon suivante :

$v_n = \alpha_0\, f(\nu(i_1))\, \alpha_1 \ldots \alpha_{n-1}\, f(\nu(i_n))\, \alpha_n$ où les α_j peuvent valoir ε ou Ω et vérifient :

- $\alpha_0 = \varepsilon$ ssi $\nu(i_1)$ est le plus petit élément de D,

- $\alpha_n = \varepsilon$ ssi $\nu(i_n)$ est le plus grand élément de D,

- pour j de 1 à n-1, $\alpha_j = \varepsilon$ si et seulement si $\nu(i_{j+1})$ est le successeur de $\nu(i_j)$ selon l'ordre total Π sur D.

La suite $\{v_n \mid n \in N_+\}$ ainsi définie est une suite croissante de W(X), en effet :

Pour tout entier n, le passage de v_n à v_{n+1} se fait en remplaçant l'une des occurrences du symbole Ω, dans v_n, par le mot $f(\nu(n+1))$ ou $\Omega f(\nu(n+1))$ ou $f(\nu(n+1))\Omega$ ou $\Omega f(\nu(n+1))\Omega$. On a donc bien $v_n \leq_\Omega v_{n+1}$. D'autre part le Sup de cette suite appartient à $W^{Loc}(X)$ (par construction). On peut donc définir : $s(u) = \mathrm{Sup}\{v_n \mid n \in N\}$.

<u>Remarque</u> :

s, ainsi définie, ne dépend pas de l'énumération choisie.

<u>Propriété</u> :

$\forall u, u' \in A_\omega(X)$

$u \equiv u'$ implique $s(u) = s(u')$.

<u>Preuve</u> :

Soit $u = \langle D, \Pi, f\rangle$
$u' = \langle D', \Pi', f'\rangle$

et h la bijection $h : D \to D'$ préservant l'ordre et telle que $f' \circ h = f$.

Il suffit de choisir ν' énumération bijective de D' égale à $\nu \circ h$ pour obtenir $v_n = v'_n$ pour tout $n \in N_+$.

Propriété :

$\forall u, u' \in A_\omega(X)$

$s(u \bullet u') = s(u) \bullet s(u')$.

Preuve :

Soit $u = \langle D, \Pi, f \rangle$; $u' = \langle D', \Pi', f' \rangle$

ν énumération bijective de D ; $D \cap D' = \phi$
ν' énumération bijective de D' ;

Soit ρ l'énumération de $D \cup D'$ (domaine de $u \bullet u'$) définie par : $\forall n \in N_+$

$\rho(2n) = \nu'(n) \in D'$
$\rho(2n-1) = \nu(n) \in D$

En notant respectivement v_n, v'_n, w_n les approximants d'ordre n de s(u), s(u') et s(uu') avec ces énumérations, on a : $\forall n \in N_+$, $w_{2n} = v_n \bullet v'_n$ et donc $s(u) \bullet s(u') = s(uu')$.

Propriété :

$\forall t \in T^{Loc}(\Sigma) : \phi(t) = s(\Psi(t))$.

Preuve :

Considérons les tronqués t[n] à profondeur $n \in N_+$ de t, t étant un arbre infini localement fini.

Soit ν une énumération bijective de fr(t) vérifiant : $\forall x, y \in fr(t)$, $|x| < |y|$ implique $\nu^{-1}(x) < \nu^{-1}(y)$ (x est énuméré avant y).

En notant v_n l'approximant d'ordre n de $s(\Psi(t))$ par cette énumération, il est clair qu'il existe une suite d'entiers strictement croissante et infinie : $\{i_j \mid j \in N_+\}$ tel que pour tout $n \in N_+$ on ait :

$$v_{i_n} = \phi(t[n]).$$

Ce qui prouve la propriété du fait de la définition de s et de la continuité de ϕ.

III. SYSTEMES D'EQUATIONS

Nous nous intéressons ici aux systèmes d'équations algébriques de mots et à leurs solutions dans le C.P.O. $W^{\infty}(X)$. Des résultats analogues à ceux bien connus des systèmes d'équations d'arbres sont obtenus.

L'opération de feuillage permet d'établir simplement la liaison entre systèmes réguliers d'arbres et algébriques de mots d'une part et leur plus petite solution d'autre part.

Il nous faut tout d'abord rappeler les définitions des substitutions du 1^{er} et du 2^{nd} ordre dans les arbres et étendre aux termes infinis les homomorphismes.

Dans tout ce chapitre on notera :

Σ un alphabet gradué

V un ensemble de variables syntaxiques

$$V = \{v_1, \ldots, v_n, \ldots\} \text{ et } V_k = \{v_1, v_2, \ldots, v_k\}$$

et Z un ensemble fini de variables graduées

$$Z = \{z_1, \ldots, z_p\} \text{ disjoint de } \Sigma \text{ et } \forall i \in [p],\ a_i = \text{arité de } z_i.$$

1. Substitutions

1.1. Substitutions du premier ordre

Ce sont des opérations qui consistent à substituer un arbre aux feuilles ayant pour symbole une variable.

Elles peuvent être définies de façon inductive puis par continuité ou de façon directe par les domaines d'arbre.

Soit une application $\sigma : V \to T_{\Omega}^{\infty}(\Sigma)$.

- σ est étendue à $T_\Omega(\Sigma \cup V)$ par les règles suivantes :

•) $\sigma(\Omega) = \Omega$

•) $\sigma(f) = f$ si $f \in \Sigma_0$

•) $\sigma(f(t_1,\ldots, f_n)) = f(\sigma(t_1),\ldots, \sigma(t_n))$

si $f \in \Sigma_p$, $p > 0$, $t_1,\ldots, t_p \in T_\Omega(\Sigma \cup V)$

- σ est ensuite étendue par continuité à

$\sigma : T^\infty_\Omega(\Sigma \cup V) \to T^\infty_\Omega(\Sigma)$

En remarquant que : $\forall t, t' \in T_\Omega(\Sigma \cup V)$

$t < t'$ implique $\sigma(t) < \sigma(t')$.

On obtient si D est une partie dirigée de $T_\Omega(\Sigma \cup V)$ alors $\sigma(D) = \{\sigma(t) \mid t \in D\}$ est une partie dirigée de $T_\Omega(\Sigma)$.

Et si D et D' sont deux parties dirigées équivalentes (même sup), $\sigma(D)$ et $\sigma(D')$ sont équivalentes ; ceci permet de poser par définition :

$\forall t \in T^\infty_\Omega(\Sigma \cup V)$
$\sigma(t) = \sup(\sigma(D))$ si $t = \sup(D)$.

Définition :

$\sigma : T^\infty_\Omega(\Sigma \cup V) \to T^\infty_\Omega(\Sigma)$ ainsi définie est une substitution du premier ordre.

On remarque que le résultat de la substitution σ à un arbre t est l'arbre $\sigma(t)$ obtenu en remplaçant simultanément dans t toute occurrence de variable (en feuille de t) par l'arbre associé.

Une autre définition, équivalente à la précédente [7] peut être formulée de la façon suivante :

Soit $\sigma : V \to T_\Omega^\infty(\Sigma)$

$\forall t \in T_\Omega^\infty(\Sigma \cup V)$ $(t : \mathrm{dom}(t) \to \Sigma \cup V \cup \{\Omega\})$ $\sigma(t)$ est l'arbre de $T_\Omega^\infty(\Sigma)$ dont le domaine vérifie : $\forall\alpha \in N_+^*$

$$\alpha \in \mathrm{dom}(\sigma(t)) \text{ ssi : } \left| \begin{array}{l} - \ \alpha \in \mathrm{dom}(t) \text{ et } t(\alpha) \notin V \\ \quad \text{dans ce cas } \sigma(t)(\alpha) = t(\alpha) \\ \text{ou} \\ - \ \alpha \text{ est de la forme } \alpha = \beta\gamma \\ \quad \text{avec } \beta \in \mathrm{dom}(t) \text{ et } t(\beta) = v \in V \\ \quad \text{et } \gamma \in \mathrm{dom}(\sigma(v)) \\ \quad \text{dans ce cas } \sigma(t)(\alpha) = \sigma(v)(\gamma) \end{array} \right.$$

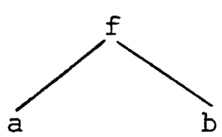
Notations :

On notera $t \cdot [t_1 \mid v_1, \ldots, t_n \mid v_n]$ le résultat de la substitution du premier ordre σ sur l'arbre t avec σ définie par

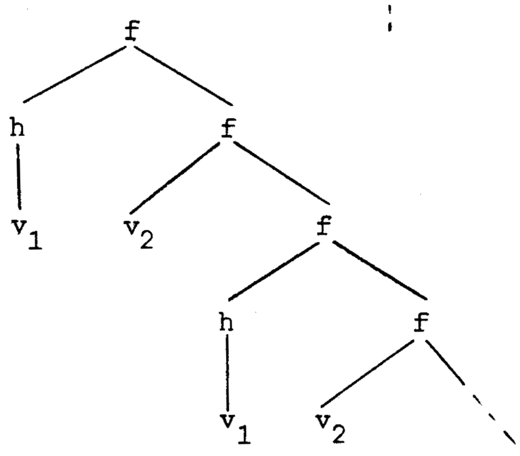
$\sigma(v_1) = t_1, \ldots, \sigma(v_n) = t_n$.

Exemple :

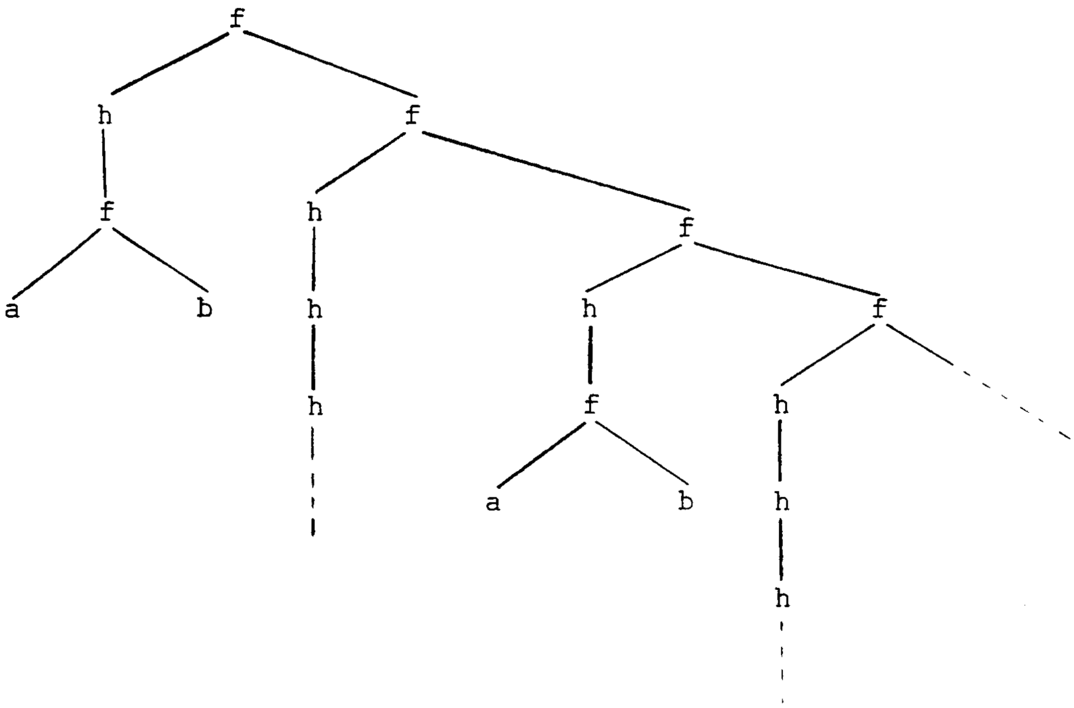
Soit σ définie par :

$\sigma(v_1) =$ 

et $\sigma(v_2) =$ 

soit l'arbre t = 

$\sigma(t)$ est l'arbre figuré de la façon suivante :



1.2. Substitutions du second ordre

Elles consistent à substituer un arbre à des sous-arbres ayant un symbole de variable fonctionnelle à la racine.

Soit θ une application de Z (variables graduées) dans $T^{\infty}(\Sigma \cup V)$ vérifiant :

$\forall i = 1,\ldots, p \quad \theta(z_i) \in T^{\infty}(\Sigma \cup V_{a_i})$

(a_i est la arité de z_i).

θ est étendue par induction à θ : $T(\Sigma \cup Z)$ à $T^{\infty}(\Sigma)$ par :

1) $\theta(f) = f$ si $f \in \Sigma_0$

2) $\theta(f(t_1,\ldots, t_n)) = f(\theta(t_1),\ldots, \theta(t_n))$

si $f \in \Sigma_n$, $n > 0$, $t_1,\ldots, t_n \in T(\Sigma \cup Z)$

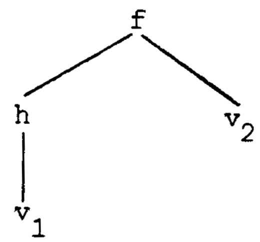
3) $\theta(z_i(t_1,\ldots, t_{a_i}))$
$= \theta(z_i) \cdot [\theta(t_1) \mid v_1,\ldots, \theta(t_{a_i}) \mid v_{a_i}]$

$\forall z_i \in Z,\ t_1, \ldots, t_{a_i} \in T(\Sigma \cup Z)$

θ est par définition une substitution du second **ordre**.

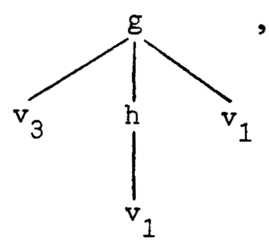
Exemple :

Soit $\theta(z_1) =$



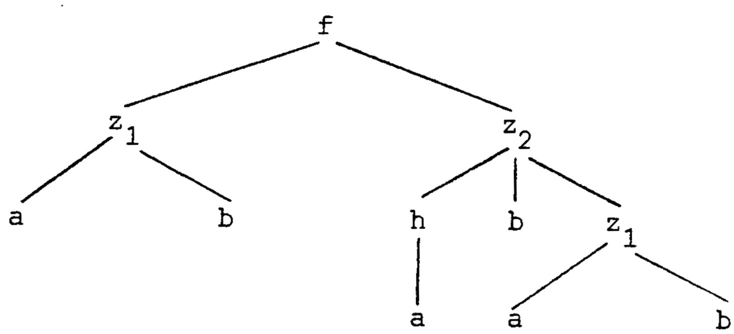
$, a_1 = 2$

$\theta(z_2) =$

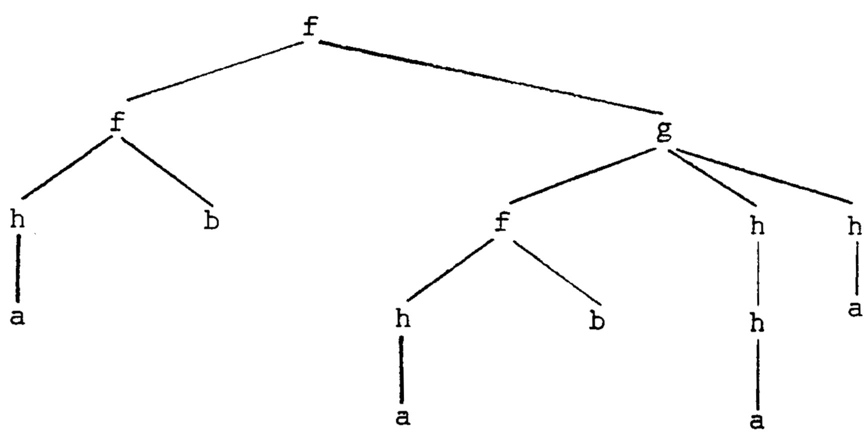


$, a_2 = 3$

$t =$



$\theta(t) =$



1.3. Homomorphismes étendus

Soient X et Y deux alphabets (non gradués).

Soit h un homomorphisme de $(X \cup \{\Omega\})^*$ dans $(Y \cup \{\Omega\})^*$ vérifiant : $h(\Omega) = \Omega$.

On remarque tout d'abord :

$\forall u, v \in (X \cup \{\Omega\})^*$

$u \simeq v$ implique $h(u) \simeq h(v)$

($\simeq$ est la congruence définie par $\Omega\Omega \simeq \Omega$).

Ceci nous permet de considérer h de W(X) dans W(Y) (classes de congruence) ; c'est évidemment un morphisme de monoïdes.

<u>$1^{ère}$ extension</u> :

Soit $h : W(X) \to W(Y)$ un morphisme vérifiant $h(\Omega) = \Omega$. Il est facile de vérifier que :

$\forall u, v \in W(X)$, $u \leq_\Omega v$ implique $h(u) \leq_\Omega h(v)$.

Ceci montre que pour toute partie dirigée D de W(X), $h(D) = \{h(u) \mid u \in D\}$ est une partie dirigée de W(Y) et que si D' est une partie dirigée équivalente à D, h(D) et h(D') sont équivalentes (même sup).

D'où l'extension par continuité de h à

$h : W^\infty(X) \to W^\infty(Y)$.

En posant $h(\sup(D)) = \sup(h(D))$, $\forall D$ partie dirigée.

Remarque :

. Si h est non effaçant ($h(x) \neq \varepsilon$, $\forall x \in X$) l'image d'un terme infini est un terme infini, ce qui n'est pas forcément vrai si h est effaçant ($\exists x \in X$ tel que $h(x) = \varepsilon$). L'image d'un mot fini est un mot fini dans tous les cas.

Seconde extension :

Cette deuxième extension consiste à considérer que l'image d'une lettre peut être un terme infini donc :

$$h : W^{\infty}(X) \to W^{\infty}(Y) \text{ avec } h(\Omega) = \Omega$$
$$\text{et } h(x) \in W^{\infty}(Y),\ \forall x \in X.$$

Tout d'abord, du fait de la continuité de la concaténation ($\mathrm{Sup}(D) \cdot \mathrm{Sup}(D') = \mathrm{Sup}(D \cdot D')$), l'image d'un mot fini est parfaitement définie par :

$$\forall w \in W(X),\ w = x_1\, x_2 \ldots\, x_n,\ x_i \in X \cup \{\Omega\}$$

$$n \in N : h(w) = h(x_1) \cdot h(x_2) \ldots\, h(x_n).$$

Si w est un terme infini, soit D une partie dirigée de W(X) ayant pour sup w.

Pour tout mot u de D, h(u) est défini et du fait de la compatibilité de l'ordre avec la concaténation, la propriété :

$$\forall u, v \in W(X),\ u \leq_{\Omega} v \text{ implique } h(u) \leq_{\Omega} h(v)$$

reste valide, ce qui prouve que h(D) est une partie dirigée de $W^{\infty}(Y)$ et permet de poser :

$$\forall w \in W^{\infty}(X),\ h(w) = \sup(h(D)) \text{ si } w = \sup(D).$$

Une telle application est par définition un morphisme de monoïde continu de $W^{\infty}(X)$ dans $W^{\infty}(Y)$.

Par la suite on appelera simplement "homomorphisme" toute application de ce type.

Propriété :

A toute substitution du premier ordre

$\sigma : T^{\infty}_{\Omega}(\Sigma \cup V) \to T^{\infty}_{\Omega}(\Sigma)$ on peut associer un et un seul homomorphisme $h : W^{\infty}(\Sigma_0 \cup V) \to W^{\infty}(\Sigma_0)$ vérifiant : $\forall t \in T^{\infty}_{\Omega}(\Sigma \cup V)$

$\phi(\sigma(t)) = h(\sigma(t))$.

Il suffit de définir h par :

$$h(\Omega) = \Omega$$
$$h(x) = x \quad \forall x \in \Sigma_0$$
$$h(v) = \phi(\sigma(v)) \quad \forall v \in V.$$

Le résultat est immédiat par les propriétés de σ, h et ϕ. L'homomorphisme obtenu est de plus non effaçant.

Substitutions "dirigée" :

Soit σ une substitution de $W(X) \to 2^{W(Y)}$, σ est dite dirigée ssi :
- $\sigma(\Omega) = \{\Omega\}$
- $\forall x \in X$, $\sigma(x)$ est une partie dirigée de $W(Y)$.

Propriétés :

σ étant une substitution dirigée,

1) $\forall w \in W(X)$, $\sigma(w)$ est une partie dirigée de $W(Y)$.

2) Pour toute partie dirigée D de $W(X)$.

$$\sigma(D) = \bigcup_{u \in D} \sigma(u) \text{ est une partie dirigée de } W(Y).$$

3) Il existe un et un seul homomorphisme h de $W^\infty(X)$ dans $W^\infty(Y)$ vérifiant :

$$h(\sup(D)) = \sup(\sigma(D)) \quad \forall D \text{ partie dirigée}$$

h est défini par : $h(\Omega) = \Omega$

$$h(x) = \sup(\sigma(x)) \quad \forall x \in X.$$

4) Réciproquement à tout homomorphisme h on peut associer une substitution dirigée σ vérifiant la propriété précédente.

2. Systèmes d'équations d'arbres

2.1. Le cas algébrique

Définition :

Un système d'équations algébriques d'arbres (ou système algébrique) est un système d'équations de la forme :

$$S = \begin{cases} z_i(v_1,\ldots, v_{a_i}) = u_i \in T(\Sigma \cup Z \cup V_{a_i}) \\ i = 1,\ldots, p \qquad p \in N_+ \end{cases}$$

où $Z = \{z_1,\ldots, z_p\}$: variables graduées ; (a_i = arité de z_i)

$V \supset V_{a_i} = \{v_1,\ldots, v_{a_i}\}$: variables syntaxiques

Σ = alphabet gradué terminal.

Un tel système est dit sous forme normale de Greibach si chaque arbre u_i a, à la racine, un symbole terminal ($\in \Sigma$) i.e. : $u_i(\varepsilon) \in \Sigma$.

Solutions :

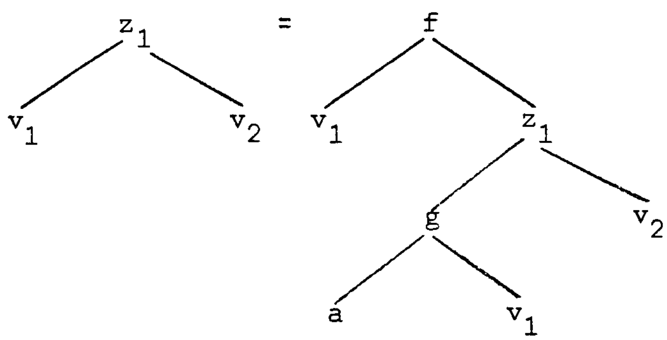
. Une solution d'un système algébrique S est un p-uple d'arbres $(t_1, \ldots, t_p)$ satisfaisant les équations c'est-à-dire vérifiant :

- $t_i \in T^{\infty}_{\Omega}(\Sigma \cup V_{a_i}) \quad \forall i = 1, \ldots, p$

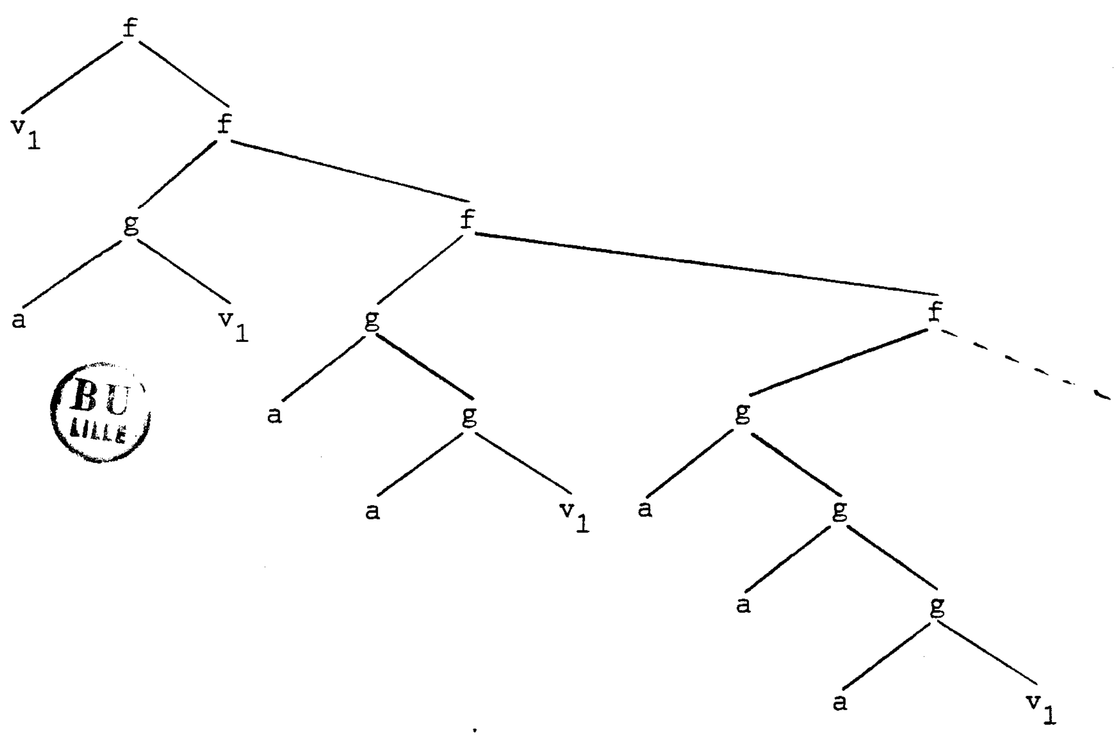
- $t_i = \Theta(u_i)$ avec Θ la substitution du second ordre définie par $\Theta(z_j) = t_j \quad \forall j = 1, \ldots, p.$

Exemple :

Soit S (réduit à une équation) :



L'arbre figuré ci-dessous est une solution de cette équation.



<u>**Théorème [7]**</u> :

Un système algébrique S a une plus petite solution $(t_1, \ldots, t_p)$ (i.e. toute autre solution $(t'_1, \ldots, t'_p)$ vérifie : $t_i < t'_i$, $i = 1, \ldots, p$). Si de plus S est sous forme normale de Greibach, il a une solution unique composée d'arbres maximaux ($t_i \in T^{\infty}(\Sigma \cup V_{a_i})$, $i = 1, \ldots, p$).

Un moyen simple de décrire la plus petite solution d'un système algébrique est de considérer celui-ci comme un système de réécriture :

$$S = \begin{cases} z_i(v_1, \ldots, v_{a_i}) \to u_i \\ i = 1, \ldots, p \end{cases}$$

Il s'agit alors d'une grammaire algébrique d'arbres et cette grammaire est déterministe dans le sens où il n'y a qu'une règle de production (dérivation) par variable.

Pour tout arbre t de $T(\Sigma \cup V \cup Z)$ notons $\hat{L}(S, t)$ l'ensemble des arbres qui se dérivent de t par S

i.e. $\hat{L}(S, t) = \{u \in T(\Sigma \cup V \cup Z) \,/\, t \xrightarrow[S]{*} u\}$.

Soit h la substitution (du second ordre) définie par :

$h(z_i) = \Omega \qquad \forall i = 1, \ldots, p$

Les autres lettres restant inchangées.

On a alors :

<u>**Propriété [7]**</u> :

Pour tout arbre t de $T(\Sigma \cup V \cup Z)$, $h(\hat{L}(S, t))$ est une partie dirigée de $T_{\Omega}(\Sigma \cup V)$.

Il suffit de remarquer que : $t \xrightarrow[S]{} t' \Rightarrow h(t) < h(t')$.

Théorème [7] :

Le p-uple d'arbres $(t_1,\dots, t_p)$ défini par

$$t_i = \mathrm{Sup}(h(\hat{L}(S, z_i(v_1,\dots, v_{a_i})))), \; i = 1,\dots, p$$

est la plus petite solution du système algébrique S.

2.2. Cas particulier : Systèmes réguliers

Définition :

Un système d'équations régulières d'arbres (ou système régulier) est un système algébrique pour lequel toutes les variables sont de arité nulle.

C'est donc un système de la forme :

$$S = \begin{cases} z_i = u_i \\ \\ i = 1,\dots, p \qquad p \in N_+ \end{cases}$$

avec les u_i appartenant à $T(\Sigma \cup Z)$ et arité $(z_i) = a_i = 0 \;\; \forall i = 1,\dots, p$.

On remarque qu'un système régulier est sous forme normale de Greibach (ou de Greibach) si aucun des u_i n'est réduit à une variable.

Une solution d'un système régulier S est un p-uple d'arbres de $T^{\infty}_{\Omega}(\Sigma)$: $(t_1,\dots, t_p)$ vérifiant les équations, c'est-à-dire dans ce cas :

$$t_i = u_i \cdot [t_1 \mid z_1,\dots, t_p \mid z_p] \qquad \text{(substitution du premier ordre)}$$

pour $i = 1,\dots, p$.

Les résultats précédents s'appliquent évidemment à ce cas particulier plus simple. La grammaire d'arbres associée à un système régulier est une grammaire régulière (réécriture en feuilles uniquement) et la substitution h $(h(z_i) = \Omega,\ z_i \in Z)$ est du premier ordre.

3. Systèmes d'équations de mots

Soient un alphabet terminal X (non gradué) et un ensemble de variables non graduées $Z = \{z_1, \dots, z_p\}$ disjoint de X.

Définition :

Un système d'équations algébriques de mots (ou système algébrique) est un système d'équations de la forme :

$$S = \begin{cases} z_i = u_i \\ i = 1, \dots, p \qquad p \in N_+ \end{cases}$$

où les u_i appartiennent à $(X \cup Z)^*$.

De tels systèmes ont été étudiés par B. Courcelle [6] et Heilbrunner [11] dans le cadre des arrangements et frontières d'arbres infinis ; nous en parlerons dans le chapitre suivant. Nous nous intéressons ici aux solutions, de systèmes algébriques, p-uples de termes infinis.

Solutions :

Une solution d'un système algébrique S est un p-uple de mots de $W^\infty(X)$: $(w_1, \dots, w_p)$ satisfaisant les équations, c'est-à-dire vérifiant :

- $\forall i = 1, \dots, p$:

$$w_i = \theta(u_i) \text{ avec } \theta \text{ l'homomorphisme défini par :}$$

$\theta(\Omega) = \Omega$

$\theta(x) = x \quad \forall x \in X$

$\theta(z_j) = w_j$ pour $j = 1, \dots, p$

Un résultat analogue à celui concernant les systèmes d'arbres peut être établi directement ; par l'opération de feuillage il en découle immédiatement.

Comme précédemment, considérons le système algébrique S comme un système de réécriture :

$$S = \begin{cases} z_i \to u_i \\ \\ i = 1, \dots, p \end{cases}$$

S est une grammaire algébrique déterministe (une règle de production par variable) de mots d'alphabet terminal X et de non-terminal Z.

Soit h le morphisme de monoïdes : $h : (X \cup Z)^* \to W(X)$ défini par :

$h(x) = x \quad \forall x \in X$

$h(z_i) = \Omega \quad \forall i = 1, \dots, p$

On remarque alors que : $\forall v, w \in (X \cup Z)^*$

$v \to w$ implique $h(v) \leq_\Omega h(w)$

d'où :

<u>**Propriété**</u> :

$\forall v \in (X \cup Z)^*$

$h(\hat{L}(S, v))$ est une partie dirigée de $W(X)$

avec $\hat{L}(S, v) = \{w \in (X \cup Z)^* \,/\, v \xrightarrow[S]{*} w\}$.

Théorème :

Tout système algébrique $S = \begin{cases} z_i = u_i \\ \\ i = 1,\ldots, p \end{cases}$ a une plus petite solution dans $(W^\infty(X))^p$.

Le p-uple $(w_1,\ldots, w_p)$ défini par :

$w_i = \mathrm{Sup}(h(\hat{L}(S, z_i)))$ pour $i = 1,\ldots, p$

est la plus petite solution du système S.

Preuve :

a) Montrons que $(w_1,\ldots, w_p)$ est solution de S et pour cela calculons $\theta(u_i)$ avec θ l'homomorphisme défini par :

- $\theta(z_i) = w_i \qquad i = 1,\ldots, p$
- $\theta(x) = x \qquad \forall x \in X$

Le mot $u_i \in (X \cup Z)^*$ est de la forme :

$u_i = \alpha_0\, z_{j_1}\, \alpha_1\, z_{j_2} \ldots \alpha_{n-1}\, z_{j_n}\, \alpha_n$, $n \in N$

avec $\alpha_0, \alpha_1,\ldots, \alpha_n \in X^*$

d'où $\theta(u_i) = \alpha_0\, w_{j_1}\, \alpha_1\, w_{j_2} \ldots \alpha_{n-1}\, w_{j_n}\, \alpha_n$

notons L_i le langage $\hat{L}(S, z_i)$.

On obtient $\theta(u_i) = \alpha_0 \cdot \sup(h(L_{j_1}))\, \alpha_1 \ldots \alpha_{n-1} \cdot \sup(h(L_{j_n}))\, \alpha_n$ par définition des w_i, et donc par propriété du C.P.O. $W^\infty(X)$:
$\theta(u_i) = \sup(\alpha_0 \cdot h(L_{j_1})\, \alpha_1 \ldots \alpha_{n-1} \cdot h(L_{j_n})\, \alpha_n)$ et puisque les α_i appartiennent à X^* et par propriété de h (morphisme) : $\theta(u_i) = \sup(h(\alpha_0\, L_{j_1}\, \alpha_1 \ldots \alpha_{n-1}\, L_{j_n}\, \alpha_n))$.

Or $\alpha_0 L_{j_1} \alpha_1 \ldots \alpha_{n-1} L_{j_n} \alpha_n$ est inclus dans L_i, par définition des langages L_i, et tout mot (sauf z_i) de L_i appartient à $\alpha_0 L_{j_1} \alpha_1 \ldots \alpha_{n-1} L_{j_n} \alpha_n$, donc $\alpha_0 L_{j_1} \alpha_1 \ldots \alpha_{n-1} L_{j_n} \alpha_n = L_i - \{z_i\}$et $\theta(u_i) = \sup(h(L_i)) = w_i$ ce qui prouve que $(w_1, \ldots, w_p)$ est solution de S.

b) Montrons maintenant que c'est la plus petite solution.

Soit $(v_1, \ldots, v_p) \in (W^{\infty}(X))^p$ une solution de S.

Par définition des solutions on a donc :

$\forall i = 1, \ldots, p : v_i = \theta(u_i)$ avec θ l'homomorphisme défini par :
$\theta(z_i) = v_i$, $i = 1, \ldots, p$
$\theta(x) = x \quad \forall x \in X$

Il faut montrer que $w_i \leq_\Omega v_i$ pour tout i et pour cela que :
$\forall y \in \hat{L}(S, z_i)$ (i.e. $z_i \xrightarrow[S]{*} y$)

$h(y) \leq_\Omega v_i$

Comme précédemment u_i est un mot de la forme

$u_i = \alpha_0 z_{j_1} \alpha_1 \ldots \alpha_{n-1} z_{j_n} \alpha_n$, $n \in N$, $\alpha_k \in X^*$

donc $v_i = \alpha_0 v_{j_1} \alpha_1 \ldots \alpha_{n-1} v_{j_n} \alpha_n$.

Par définition de l'ordre et l'expression de v_i on a $h(z_i) = \Omega \leq_\Omega v_i$ et $h(u_i) = \alpha_0 \Omega \alpha_1 \ldots \alpha_{n-1} \Omega \alpha_n \leq_\Omega v_i$.

Il est clair d'autre part qu'à toute dérivation dans S :
$z_i \to y_1 \to y_2 \to \ldots \to y_k = y$, on peut associer une transformation de l'expression de v_i de façon que si $y = \beta_0 z_{k_1} \beta_1 \ldots \beta_{\ell-1} z_{k_\ell} \beta_\ell$ avec $\beta_0, \ldots, \beta_\ell \in X^*$ on ait :
$v_i = \beta_0 v_{k_1} \beta_1 \ldots \beta_{\ell-1} v_{k_\ell} \beta_\ell$ en remplaçant itérativement dans l'expression de v_i l'un des v_j par son expression (puisque $(v_1, \ldots, v_p)$ est solution de S) et donc on obtient : $h(y) \leq_\Omega v_i$ pour tout $y \in \hat{L}(S, z_i)$ ce qui établit le résultat.

Remarques :

Les systèmes algébriques n'ont pas en général une solution unique. Il suffit pour s'en convaincre de considérer l'exemple très simple suivant :

$S = \{z = a\, z\, b\}$ avec $X = \{a, b\}$

la plus petite solution de S est $w = \mathrm{Sup}\{a^n \,\Omega\, b^n \,/\, n \in N\}$ une autre solution est par exemple $u = \mathrm{Sup}\{a^n \,\Omega\, (ab)^n \,\Omega\, b^n \,/\, n \in N\}$.

Pour les systèmes d'arbres, le fait d'être sous forme normale de Greibach garantit l'unicité de la solution. Cette notion de forme normale n'a évidemment aucun sens pour les systèmes de mots.

Les systèmes algébriques n'ont pas en général une plus grande solution.

Exemple :

$S = \{z = z\, ab\, z\}$ avec $X = \{a, b\}$

$w_1 = \mathrm{Sup}(a^* \,\Omega\, (a+b)^* \,\Omega\, b^*)$ est solution

$w_2 = \mathrm{Sup}(b^* \,\Omega\, (a+b)^* \,\Omega\, a^*)$ est solution

$w_1 \neq w_2$ et w_1, w_2 sont maximaux dans $W^\infty(X)$.

4. Liaison entre systèmes réguliers d'arbres et systèmes algébriques de mots

La liaison est évidemment établie par l'opération de feuillage.

Soit S un système régulier d'arbres :

$$S = \begin{cases} z_i = u_i \in T(\Sigma \cup Z) \\ \\ i = 1, \ldots, p \end{cases}$$

Soit S_m le système algébrique de mots dérivé de S :

$$S_m = \begin{cases} z_i = \phi(u_i) \in (\Sigma_0 \cup Z)^* \\ i = 1,\ldots, p \end{cases}$$

Propriété :

Si $(t_1,\ldots, t_p)$ est la plus petite solution de S alors $(\phi(t_1),\ldots, \phi(t_p))$ est la plus petite solution de S_m.

(Les $\phi(t_i)$ appartenant à $W^\infty(\Sigma_0)$)

c'est immédiat.

Réciproquement, à tout système algébrique de mots S_m

$$S_m = \begin{cases} z_i = w_i \in (X \cup Z)^* \\ i = 1,\ldots, p \end{cases}$$

on peut associer un système régulier d'arbre S

$$S = \begin{cases} z_i = u_i \in T(\Sigma \cup Z)^* \\ i = 1,\ldots, p \end{cases}$$

tel que la plus petite solution de S_m soit le feuillage de la plus petite solution de S (feuillage d'un p-uple d'arbres : p-uple des feuillages des arbres).

De plus S peut être choisi sous forme normale de Greibach.

Il suffit de choisir $\Sigma_0 = X$ et les arbres u_i, non réduits à une variable, vérifiant $\phi(u_i) = w_i$.

IV. DECISION DE L'EGALITE DES FEUILLAGES

. En montrant que le feuillage d'un arbre régulier peut être déterminé par un langage rationnel dont il est le sup, on en déduit que l'égalité des des feuillages d'arbres réguliers est décidable. Il est ensuite montré, à l'aide d'un codage simple, que le problème de décision de l'égalité des arbres algébriques est réductible en celui de l'égalité des feuillages d'arbres algébriques.

1. Feuillages d'arbres réguliers

Définition :

Un arbre est régulier si l'ensemble de ses sous-arbres est fini (i.e. s'il a un nombre fini de sous-arbres distincts).

Bien sûr, tout arbre fini est régulier.

Une importante caractérisation des arbres réguliers est la suivante :

Propriété :

Un arbre est régulier si et seulement s'il est une composante de la solution unique d'un système d'équations régulières d'arbres sous forme normale de Greibach.

Soit un arbre régulier $t \in T^{\infty}(\Sigma)$ donné par un système régulier S (sous-entendu sous forme normale de Greibach) :

$$S = \begin{cases} z_i = u_i \in T(\Sigma \cup Z) \\ i = 1,\ldots, p \quad p \in N_+ \end{cases}$$

On peut toujours supposer, quite à renuméroter les variables, que l'arbre t est la première composante de la solution du système S.

Grâce au chapitre précédent, nous savons que le feuillage de t est alors la première composante de la plus petite solution du système algébrique de mots

$$S_m = \begin{cases} z_i = \phi(u_i) \in (\Sigma_0 \cup Z)^* \\ i = 1,\ldots, p \end{cases}$$

et $\phi(t) = \sup(h(L_1))$

avec $L_1 = \{w \in (\Sigma_0 \cup Z)^* \mid z_1 \xrightarrow[S_m]{*} w\}$

et $h : (\Sigma_0 \cup Z)^* \to W(\Sigma_0)$ défini par
$h(z_i) = \Omega \quad \forall i = 1,\ldots, p$
$h(x) = x \quad \forall x \in \Sigma_0$

Réciproquement, toute composante de la plus petite solution d'un système algébrique de mots est le feuillage d'un arbre régulier.

Etant donné un autre arbre régulier t' donné par un système régulier S' (sous les mêmes hypothèses que t) d'où dérive le système S'_m, décider de l'égalité $\phi(t) = \phi(t')$ c'est donc décider de l'équivalence des parties dirigées $h(L_1)$ et $h(L'_1)$ de $W(\Sigma_0)$.

On sait que deux parties dirigées D et D' sont équivalentes

ssi tout élément de D est majoré par un élément de D' et réciproquement,
ssi elles ont même Idéal
i.e. $Id(D) = Id(D')$
avec $Id(L) = \{w \in W(\Sigma_0) \mid \exists v \in L \,\&\, w \leq_\Omega v\} \ \forall L \subset W^\infty(\Sigma_0)$,
ssi l'ensemble des représentants canoniques des éléments de Id(D) est égal à celui des éléments de Id(D').

Notation (Rappel) :

On notera ρ l'application de $W(X)$ dans $(X \cup \{\Omega\})^*$ (pour tout alphabet X) qui à tout élément de $W(X)$ associe son représentant canonique dans $(X \cup \{\Omega\})^*$ i.e. $\forall w \in W(X)$, $\rho(w)$ est le mot de la classe de congruence w contenant le moins d'occurrences de la lettre Ω.

D'après ce qui précède nous avons donc :

(I) $\phi(t) = \phi(t')$ ssi $\rho(Id(h(L_1))) = \rho(Id(h(L_1')))$

Théorème IV-1 :

Pour tout système algébrique de mots S_m

$$S_m = \begin{cases} z_i = u_i \in (X \cup Z)^* \\ i = 1,\dots, p \qquad p \in N_+ \end{cases}$$

Pour tout $i = 1,\dots, p$: le langage $\rho(Id(h(L_i)))$ de $(X \cup \{\Omega\})^*$ est rationnel et l'on peut construire une expression rationnelle le représentant.

Théorème VI-2 :

L'égalité des feuillages d'arbres réguliers est décidable.

C'est une conséquence directe du théorème précédent et de l'équivalence (I), puisque l'égalité de langages rationnels est décidable.

2. Preuve du Théorème IV-1

Nous allons tout d'abord établir un lemme simple qui permet de simplifier le problème. Nous donnerons ensuite quelques exemples montrant que pour certaines classes de systèmes l'obtention d'un langage rationnel satisfaisant est simple et qu'elle est moins évidente à priori pour un système quelconque. Deux preuves différentes seront alors données : l'une "directe" s'appuyant sur la construction d'un automate associé à un système algébrique, l'autre utilisant les résultats de Courcelle [6] et Heilbrunner [11] concernant la résolution de systèmes algébriques dans le cadre des arrangements et frontières d'arbres infinis.

2.1. Préliminaire

Soit un langage L de $(X \cup \{\Omega\})^*$.

A ce langage correspond une partie, notée également L, de W(X) : le saturé par la relation de congruence ($\simeq$) c'est-à-dire l'ensemble des classes de congruences de W(X) qui sont représentées dans le langage L.

ρ étant toujours défini comme précédemment :

Lemme :

Pour tout langage rationnel R de $(X \cup \{\Omega\})^*$, on peut construire le langage rationnel $\rho(Id(R))$.

Preuve :

Soit un automate A reconnaissant le langage R. Pour obtenir les mots "pré-plus petit" ($\leq_\Omega$) qu'un mot w de R, il faut pouvoir insérer des Ω dans w et remplacer toute lettre de w par Ω.

Ceci correspond à la transformation de l'automate A par l'ajout de transitions :

a) Pour tout état q, ajouter la transition $q \overset{\Omega}{\circlearrowleft}$
(i.e. passer de l'état q à lui-même en lisant Ω)

b) Pour toute transition $q \xrightarrow{x} q'$, $\forall x \in X$ ajouter la transition $q \xrightarrow[\Omega]{} q'$

Pour obtenir ensuite tous les représentants canoniques équivalents aux mots reconnus, il suffit d'effectuer sur l'automate la clôture transitive des Ω-transitions

i.e. si $q \xrightarrow{\Omega} q' \xrightarrow{\Omega} q''$

ajouter la transition $q \xrightarrow{\Omega} q''$

Le langage reconnu alors par cet automate modifié est égal à $\rho(Id(R)) \cup M$ avec M tel que : $M \cap \rho(Id(R)) = \phi$ et $\forall w \in M$, $\exists u \in \rho(Id(R))$ & $w \simeq u$.

Il suffit alors d'effectuer l'intersection rationnelle avec le langage $X^*(\Omega\ X^+)^*\ \ (\Omega + \varepsilon)$ pour éliminer les mots de M et donc reconnaître $\rho(Id(R))$.

Remarque :

La preuve que $\rho(Id(R))$ est rationnel peut également être faite par la définition d'une transduction rationnelle.

Ce lemme montre, en particulier, que pour la preuve du Théorème il suffit de construire un rationnel R_i de $(X \cup \{\Omega\})^*$ vérifiant :

- $R_i \subseteq W(X)$ est une partie dirigée.

- $Sup(R_i) = i^{\text{ème}}$ composante de la plus petite solution du système S_m.

2.2. Exemples

a) Système linéaire :

Un système est linéaire si chaque membre droit d'équation possède au plus une occurrence de variable.

Soit

$$S_1 = \begin{cases} z_1 = w_1\ z_2\ w_2 \\ z_2 = w_3\ z_3\ w_4 \\ z_3 = w_5\ z_2\ w_6 \end{cases} \qquad \begin{array}{l} Z = \{z_1,\ z_2,\ z_3\} \\ \text{les } w_i \in X^* \\ \ \end{array}$$

par simple substitution dans les équations on obtient le système équivalent :

$$S'_1 = \begin{cases} z_1 = w_1\ z_2\ w_2 \\ z_2 = w_3\ w_5\ z_2\ w_6\ w_4 \\ z_3 = w_5\ w_3\ z_3\ w_4\ w_6 \end{cases}$$

et $L_1 = w_1\ L_2\ w_2$

$$h(L_2) = \{(w_3\ w_5)^n\ \Omega\ (w_6\ w_4)^n \,/\, n \in N\}$$

$$h(L_3) = \{(w_5\ w_3)^n\ \Omega\ (w_4\ w_6)^n \,/\, n \in N\}$$

puis $h(L_1) = \{w_1(w_3\ w_5)^n\ \Omega\ (w_6\ w_4)^n\ w_2 \,/\, n \in N\}$.

Ces langages sont bien sur algébriques, mais il est facile de voir que les langages rationnels R_1, R_2, R_3 définis ci-dessous leur sont respectivement équivalents (même sup).

$$R_1 = w_1(w_3\ w_5)^*\ \Omega\ (w_6\ w_4)^*\ w_2$$

$$R_2 = (w_3\ w_5)^*\ \Omega\ (w_6\ w_4)^*$$

$$R_3 = (w_5\ w_3)^*\ \Omega\ (w_4\ w_6)^*$$

b) Système quasi-rationnel :

Un système est quasi-rationnel s'il ne possède pas de variables expansives c'est-à-dire s'il n'existe pas d'équations permettant des réécritures de la forme :

$$z_i \xrightarrow{*} x\ z_i\ y\ z_i\ z, \quad x, y, z \in (X \cup Z)^*$$

Autrement dit, pour tout i, pour tout mot w appartenant à L_i, le nombre d'occurrences de z_i dans w est au plus égal à 1.

Exemple :

$$S_2 = \begin{cases} z_1 = a\ z_2\ b\ z_1\ c\ z_2\ d \\ z_2 = u\ z_3\ v \\ z_3 = w\ z_2\ x \end{cases} \qquad \begin{matrix} \\ Z = \{z_1, z_2, z_3\} \\ a, b, c, d, u, v, w, x \in X^* \end{matrix}$$

Le sous-système composé des deux dernières équations est linéaire et peut être résolu comme précédemment :

$$R_2 = (u\ w)^*\ \Omega\ (x\ v)^*$$

$$R_3 = (w\ u)^*\ \Omega\ (v\ x)^*$$

La première équation est linéaire en la variable z_1 (z_2 considérée comme constante) et résolue par :

$$R_1 = (a\ z_2\ b)^*\ \Omega\ (c\ z_2\ d)^*$$

Il suffit alors de substituer R_2 à z_2 dans l'expression de R_1 pour obtenir le rationnel cherché :

$$R_1 = (a(u\ w)^*\ \Omega\ (x\ v)^*\ b)^*\ \Omega\ (c(u\ w)^*\ \Omega\ (x\ v)^*\ d)^*$$

Cette expression peut aussi être obtenue directement par l'expression du langage algébrique $h(L_1)$:

$$h(L_1) = \{(a(u\ w)^n\ \Omega\ (x\ v)^n\ b)^p\ \Omega\ (c(u\ w)^q\ \Omega\ (x\ v)^q\ d)^p \,/\, n, p, q \in N\}$$

On a bien $Sup(R_1) = Sup(h(L_1))$.

c) Système non quasi-rationnel

Exemple :

$$S_3 = \{z_1 = a\ z_1\ z_1\ b\},\ X = \{a, b\}$$

Le langage $h(L_1)$ est bien entendu algébrique, mais ne peut être exprimé simplement comme dans les exemples précédents. Ce système, par la construction donnée plus loin, est résolu par le rationnel :

$$R_1 = (a^* \ \Omega\ b^*)^*$$

Autre exemple :

$$\begin{aligned} S = \{z_1 &= a\ z_2\ b\ z_3\ c\ z_1\ d\ z_4\ e, \\ z_2 &= u\ z_2\ v\ z_3\ w, \\ z_3 &= x\ z_2\ y\ z_3\ t, \\ z_4 &= a\ z_4\} \end{aligned}$$

avec $Z = \{z_1, z_2, z_3, z_4\}$.

Cet exemple sera développé au cours de la preuve.

2.3. Preuve par construction d'automates

Le principe de cette preuve est de partitionner le système initial en sous-systèmes particuliers ; pour chacun de ces sous-systèmes construire un automate permettant d'obtenir des langages rationnels dont les sup forment la plus petite solution, puis par substitutions rationnelles sur ces langages obtenir les expressions rationnelle résolvant le système initial.

Soit un système algébrique S, d'alphabet terminal X et de variables $Z = \{z_1, \ldots, z_p\}$, $p \in N_+$

$$S = \begin{cases} z_i = u_i \in (X \cup Z)^* \\ \\ i = 1,\dots, p \end{cases}$$

Notons w_z la composante en la variable z de la plus petite solution de S, pour toute variable $z \in Z$.

Soit la relation d'équivalence $(\equiv_S)$ sur Z définie par :

$\forall i, j = 1,\dots, p$

$$z_i \equiv_S z_j \text{ ssi } \begin{cases} - \ z_i \xrightarrow[S]{*} \alpha \, z_j \, \alpha' & \alpha, \alpha' \in (X \cup Z)^* \\ \text{et} & \\ - \ z_j \xrightarrow[S]{*} \beta \, z_i \, \beta' & \beta, \beta' \in (X \cup Z)^* \end{cases}$$

Cette relation permet de partitionner Z en

$Z = Z_1 \cup \dots \cup Z_q$, $q \leq p$

et donc le système S en sous-systèmes disjoints $S_1, \dots, S_q$:
$S_j = \{z_i = u_i \,/\, z_i \in Z_j\}$

Chacun de ces sous-systèmes (S_j) peut être considéré comme un système algébrique d'alphabet terminal $X \cup (Z - Z_j)$ et de variables Z_j, dans lequel toutes les variables sont équivalentes. Toute composante de la plus petite solution de S_j est alors un élément de $W^\infty(X \cup (Z - Z_j))$.

Notons w_z^j la composante en la variable $z \in Z_j$ de la plus petite solution de S_j, $\forall z \in Z_j$ et pour tout $j = 1,\dots, q$.

L'ensemble $\{S_1, \ldots, S_q\}$ des sous-systèmes est naturellement ordonné par la relation :

$$S_i < S_j \text{ s'il existe } z_k \in S_i \text{ et } z_\ell \in S_j$$

tels que $z_k \xrightarrow[S_i]{*} \alpha\, z_\ell\, \beta$, $\alpha, \beta \in (X \cup Z)^*$

Les sous-systèmes maximaux, relativement à cette relation, sont en fait des systèmes sur l'alphabet terminal X et chaque composante de leur (plus petite) solution appartient à $W^\infty(X)$. Avec les notations précédentes on a donc :

$$\text{si } S_j \text{ est maximal, } \forall z \in Z_j : w_z^j = w_z.$$

Soit le système S', sous-système de S, formé de la réunion des sous-systèmes maximaux :

$$S' = \{z_i = u_i \,/\, z_i \in Z_j \text{ et } S_j \text{ maximal}\}$$

en notant Z' l'ensemble des variables de S'.

Soit un sous-système S_i vérifiant :

$$\forall z_k = u_k \in S_i,\ u_k \in (X \cup Z \cup Z_i)^*$$

c'est un système sur l'alphabet terminal $X \cup Z'$ et les w_z^i appartiennent à $W^\infty(X \cup Z')$ $\forall z \in Z_i$.

Soit θ l'homomorphisme défini par :

$$\theta(z) = w_z \qquad \forall z \in Z'$$
$$\theta(x) = x \qquad \forall x \in X$$

Il est clair alors, d'après les relations d'équivalence et d'ordre que l'on a :

$$\forall z \in Z_i : w_z = \theta(w_z^i)$$

Supposons maintenant résolus les systèmes S_i et S' c'est-à-dire que l'on ait construit les langages rationnels :

$R_z \subseteq (X \cup \{\Omega\})^*$ pour tout $z \in Z'$

tel que $\sup(R_z) = w_z$

et $R_z^i \subseteq (X \cup \{\Omega\} \cup Z')^*$ pour tout $z \in Z_i$

tel que $\sup(R_z^i) = w_z^i$.

La substitution rationnelle σ définie par :

- $\sigma(z) = R_z$, $\forall z \in Z'$
- $\sigma(x) = x$, $\forall x \in X$
- $\sigma(\Omega) = \Omega$

est dirigée et vérifie :

$\forall D$ dirigée de $W(X \cup Z')$
$\theta(\sup(D)) = \sup(\sigma(D))$

ce qui montre en particulier que :

$\forall z \in Z_i$, $w_z = \theta(w_z^i) = \theta(\sup(R_z^i))$

est égal à $\sup(\sigma(R_z^i))$

Le langage $\sigma(R_z^i)$ est un rationnel de $(X \cup \{\Omega\})^*$ dont l'expression est obtenue en remplaçant dans l'expression de R_z^i toute occurrence de variable de Z' par l'expression rationnelle associée.

Le système S_i + S' est alors résolu.

En réitérant cette opération un nombre fini de fois on obtiendra la solution (les rationnels) du système initial S, sous l'hypothèse, bien sur, que l'on ait résolu chacun des sous-systèmes. C'est ce dernier point qu'il reste à considérer.

Exemple (suite de l'exemple précédent) :

La relation d'équivalence conduit à considérer les trois sous-systèmes :

$S_1 = \{z_1 = a\ z_2\ b\ z_3\ c\ z_1\ d\ z_4\ e\} \qquad Z_1 = \{z_1\}$

$S_2 = \{z_2 = u\ z_2\ v\ z_3\ w,\ z_3 = x\ z_2\ y\ z_3\ t\} \qquad Z_2 = \{z_2, z_3\}$

$S_3 = \{z_4 = a\ z_4\} \qquad Z_3 = \{z_4\}$

on a $S_1 < S_2$, $S_1 < S_3$.

S_2 et S_3 sont maximaux et $S' = S_2 \cup S_3$.

S_1 et S_3 sont linéaires et peuvent être résolus facilement :

$R_1^1 = (a\ z_2\ b\ z_3\ c)^* \ \Omega\ (d\ z_4\ e)^*$

$R_4^3 = a^* \ \Omega = R_4$

Supposons S_2 résolu par les rationnels R_2 et R_3.

Par substitution dans R_1^1 de R_4 à z_4, R_2 à z_2 et R_3 à z_3 on obtient :

$R_1 = (a\ R_2\ b\ R_3\ c)^* \ \Omega\ (d\ a^* \ \Omega\ e)^*$

Résolution des sous-systèmes :

Soit un système $S = \begin{cases} z_i = u_i \in (X \cup Z)^* \\ i = 1, \ldots, p \end{cases}$ tel que toutes les variables soient équivalentes ($\equiv_S$) deux à deux.

On peut éliminer le cas où S est réduit à une équation terminale (i.e. le membre droit est un mot de X^*) pour lequel la solution est immédiate.

Dans tous les autres cas, chaque membre droit des équations contient au moins une occurrence de variable.

Soit l'automate A associé à S défini de la façon suivante :

- <u>ensemble d'états Q</u> :

$$Q = \{q_1, \ldots, q_p\} \cup \{\bar{q}_1, \ldots, \bar{q}_p\}$$

i.e. pour chaque variable z_i on définit deux états q_i et $\bar{q}_i$.

- <u>alphabet</u> = $X \cup \{\Omega\}$.

- <u>transitions</u> : elles sont de quatre types différents pour tout $i = 1, \ldots, p$ on aura :

a) $q_i \bullet \xrightarrow{\Omega} \bullet \bar{q}_i$

b) $q_i \bullet \xrightarrow{\alpha_i} \bullet q_j$ si on a $z_i = \alpha_i z_j w$, $\alpha_i \in X^*$

t) $\bar{q}_k \bullet \xrightarrow{\beta_i} \bullet \bar{q}_i$ si on a $z_i = w' z_k \beta_i$, $\beta_i \in X^*$

d) $\bar{q}_j \bullet \xrightarrow{\gamma_i} \bullet q_k$ si on a $z_i = w_1 z_j \gamma_i z_k w_2$, $\gamma_i \in X^*$.

- <u>état initial</u> = q_i pour un i fixé.

- <u>un seul état terminal</u> = $\bar{q}_i$ pour le même i.

<u>Remarques</u> :

- Tout mot reconnu par l'automate possède au moins une occurrence du symbole Ω.

- Les transitions sont définies par des mots de X^* (sauf celles de type a)) et non par des lettres. On pourra toujours, si besoin est, déduire de cet automate un automate sous forme usuelle reconnaissant le même langage.

- Etant donné un état q_i, il existe un et un seul état q_j tel qu'il existe une transition de q_i à q_j.

- Etant donné un état $\bar{q}_i$, il existe un et un seul état $\bar{q}_k$ tel qu'il existe une transition de $\bar{q}_k$ à $\bar{q}_i$.

Soit R_i le langage reconnu par cet automate ; pour terminer la preuve du théorème, il faut montrer que R_i (considéré dans W(X)) est une partie dirigée dont le sup est égal à la ième composante de la plus petite solution de S.

Pour cela il suffit de montrer :

1 Pour tout mot $w \in (X \cup Z)^*$ tel que $z_i \xrightarrow[S]{*} w$, le mot h(w) est reconnu par l'automate avec h l'homomorphisme de $(X \cup Z)^*$ dans $(X \cup \{\Omega\})^*$ défini par : $-h(z) = \Omega \quad \forall z \in Z$

$- h(x) = x \quad \forall x \in X$

2 Pour tout mot w reconnu par l'automate, il existe un mot $w' \in (X \cup Z)^*$ tel que $- z_i \xrightarrow[S]{*} w'$

et $- w \leq_\Omega h(w')$

Le premier point (1) est évident. Il suffit en effet de remarquer que pour toute équation du système S :

$z_j = u_j$ si u_j est de la forme

$u_j = \alpha_0 z_{j_1} \alpha_1 z_{j_2} \dots \alpha_{n-1} z_{j_n} \alpha_n$, les $\alpha_k \in X^*$

On a dans l'automate la suite de transitions suivante :

$$q_j \xrightarrow{\alpha_0} q_{j_1} \xrightarrow{\Omega} \bar{q}_{j_1} \xrightarrow{\alpha_1} q_{j_2} \xrightarrow{\Omega} \bar{q}_{j_2} \to \dots \to \bar{q}_{j_{n-1}} \xrightarrow{\alpha_{n-1}} q_{j_n} \xrightarrow{\Omega} \bar{q}_{j_n} \xrightarrow{\alpha_n} \bar{q}_j$$

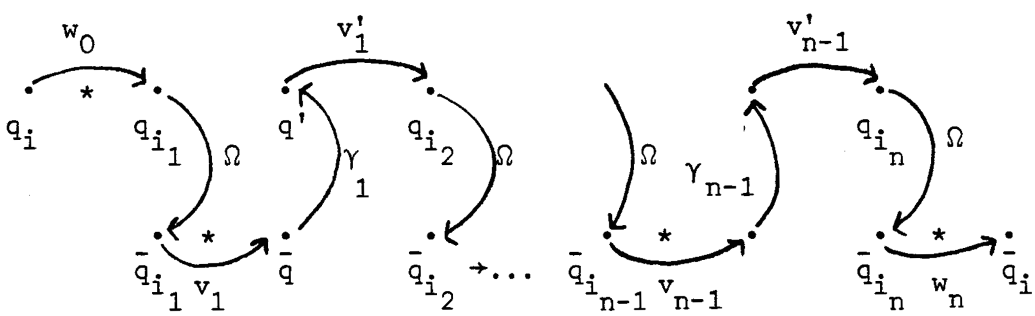
Preuve du (2)

Soit $w \in R_i$. w est un mot de la forme $w = w_0\ \Omega\ w_1\ \Omega \ldots \Omega\ w_{n-1}\ \Omega\ w_n$, $n \in N_+$ et les $w_i \in X^*$.

D'après la définition de l'automate, il existe alors une suite d'états $q_{i_1}, q_{i_2}, \ldots, q_{i_n}$ telle que l'on ait :

- En lisant w_0, on passe de l'état initial q_i à l'état q_{i_1}, sans rencontrer un état barré.
- En lisant w_n, on passe de l'état $\bar{q}_{i_n}$ à l'état terminal $\bar{q}_i$, sans rencontrer un état non barré.
- Pour $j = 1$ à $n-1$, en lisant w_j on passe de l'état $\bar{q}_{i_j}$ à l'état $q_{i_{j+1}}$ avec une et une seule transition d'un état barré à un état non barré. w_j se décompose donc en $v_j\ \gamma_j\ v'_j$ avec γ_j assurant cette transition particulière.

De façon schématique :



Lemme 1 :

Pour toute suite de transitions de la forme :

$$q_{j_1} \xrightarrow{\beta_1} q_{j_2} \xrightarrow{\beta_2} q_{j_3} \cdots q_{j_n} \xrightarrow{\beta_n} q_{j_{n+1}}, \quad n \in N_+$$

On a dans le système S :

$$z_{j_1} \xrightarrow[S]{n} \alpha_1\, \alpha_2 \ldots \alpha_n\, z_{j_{n+1}}\, w' \qquad w' \in (X \cup Z)^*$$

C'est immédiat d'après la définition de ce type de transitions.

Lemme 2 :

Pour toute suite de transitions de la forme :

$$\bar{q}_{j_1} \xrightarrow{\beta_1} \bar{q}_{j_2} \xrightarrow{\beta_2} \bar{q}_{j_3} \ldots \bar{q}_{j_n} \xrightarrow{\beta_n} \bar{q}_{j_{n+1}}, \quad n \in N_+$$

On a dans S :

$$z_{j_{n+1}} \xrightarrow[S]{n} w'\, z_{j_1}\, \beta_1\, \beta_2 \ldots \beta_n$$

Même remarque.

Lemme 3 :

Pour toute suite de transitions de la forme :

$$\bar{q}_{j_1} \xrightarrow{\beta_1} \bar{q}_{j_2} \xrightarrow{\beta_2} \ldots \bar{q}_{j_n} \xrightarrow{\beta_n} \bar{q}_k \xrightarrow{\gamma} q_{k_1} \xrightarrow{\alpha_1} q_{k_2} \xrightarrow{\alpha_2} \ldots q_{k_\ell} \xrightarrow{\alpha_\ell}$$

$n, \ell \in N_+$

On a dans S:

$\exists z \in Z$ tel que : $z \xrightarrow[S]{n+\ell+1} w'\, z_{j_1}\, \beta_1\, \beta_2 \ldots \beta_n\, \gamma\, \alpha_1\, \alpha_2 \ldots \alpha_\ell\, z_m\, w''$

puisque l'on a : $\exists z \in Z$ tel que $z \xrightarrow[S]{} u_1\, z_k\, \gamma\, z_{k_1}\, u_2$

et par le lemme 1 $z_k \xrightarrow{n} v_1\, z_{j_1}\, \beta_1\, \beta_2 \ldots \beta_n$

et par le lemme 2 $z_{k_1} \xrightarrow{\ell} \alpha_1\, \alpha_2 \ldots \alpha_\ell\, z_m\, v_2$.

Grâce au lemme 1 on obtient donc par des dérivations gauches

$$z_i \xrightarrow[S]{*} w_0\, z_{i_1}\, w'$$

avec w' de la forme : $w''\, z\, \beta$, $z \in Z$, $\beta \in X^*$

ou bien $w' \in X^*$, $w' = \beta$

Deux cas peuvent alors se produire :

- β est un facteur droit de w_n

i.e. $w_n = \beta'\, \beta \qquad \beta' \in X^*$

ou

- w_n est un facteur droit de β

dans ce cas on a donc :

$$z_i \xrightarrow[S]{*} w_0\, v\, w_n, \text{ v contenant au moins une variable.}$$

Dans le cas précédent, par le lemme 2, en appliquant des dérivations à droite on aura

$$z_i \xrightarrow[S]{*} w_0\, z_{i_1}\, x' \xrightarrow[S]{*} w_0\, v\, w_n$$

<u>Remarque</u> :

Le fait que l'on ne puisse rencontrer que ces deux possibilités provient de l'unicité des transitions :

$$q_i \xrightarrow{\alpha_i} q_j \qquad \text{pour } i \text{ donné}$$

et

$$\bar{q}_k \xrightarrow{\alpha_\ell} \bar{q}_\ell \qquad \text{pour } \ell \text{ donné}$$

On a donc obtenu : $z_i \xrightarrow[S]{*} w_0\ v\ w_n$.

v contient au moins une variable ; par le lemme 3 on sait qu'il existe une variable z telle que :

$z \xrightarrow[S]{*} v_1\ w_1\ v_1'$, v_1' contenant une variable au moins.

Or toutes les variables sont équivalentes, ce qui implique que :

$$z_i \xrightarrow[S]{*} w_0\ v'\ z_j\ v''\ w_n \xrightarrow{*} w_0\ v'\ u\ z\ u'\ v''\ w_n$$

et donc $z_i \xrightarrow[S]{*} w_0\ v\ w_1\ x\ w_n$, $v, x \in (X \cup V)^*$

et x contenant une variable.

Par le même raisonnement à partir d'une variable dans x on obtiendra w_2 et ainsi de suite jusqu'à w_n et l'on aura :

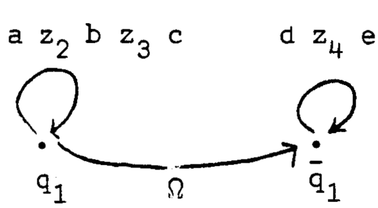
$$z_i \rightarrow w_0\ v_0\ w_1\ v_1\ w_2\ v_2 \ldots w_{n-1}\ v_{n-1}\ w_n = w'$$

où les $v_i \in (X \cup V)^*$

et il est évident par définition que $w \leq_\Omega h(w')$ ce qui établit le résultat.

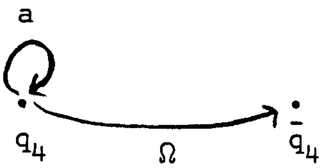
<u>Exemple (suite)</u> : Automates associés à S_1, S_2 et S_3.

Automate A_1 :



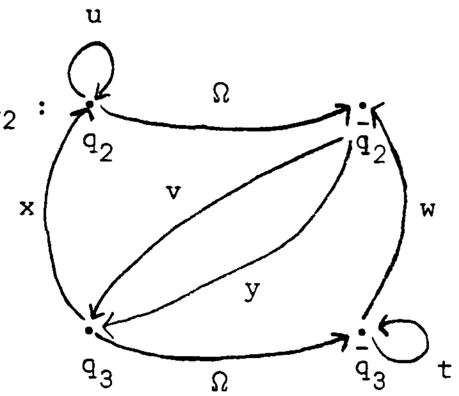
$S_1 = \{z_1 = a\ z_2\ b\ z_3\ c\ z_1\ d\ z_4\ e\}$

On retrouve le rationnel $R_1^1 = (a\ z_2\ b\ z_3\ c)^*\ \Omega\ (d\ z_4\ e)^*$.

Automate A_3 : 

$S_3 = \{z_4 = a\ z_4\}$

A_3 reconnaît $R_4 = a^* \Omega$

Automate A_2 : 

$$S_2 = \begin{cases} z_2 = u\ z_2\ v\ z_3\ w \\ z_3 = x\ z_2\ y\ z_3\ t \end{cases}$$

Avec q_2 comme état initial (et donc $\bar{q}_2$ comme état terminal) on obtient alors l'expression rationnelle :

$$R_2 = u^* \Omega\ ((v+y)\ (x\ u^* \Omega + \Omega\ t^* w))^*.$$

Avec l'état initial q_3 on obtient :

$$R_3 = ((\Omega\ t^* w + x\ u^* \Omega)\ (v+y))^* \Omega\ t^*$$

2.4. Deuxième preuve

Les systèmes algébriques de mots ont été étudiés par Courcelle [6] et Heilbrunner [11] dans le cadre des arrangements dénombrables et frontières d'arbres infinis. Une solution d'un système étant alors un p-uple d'arrangements dénombrables. Ces systèmes sont résolus à l'aide d'expressions régulières représentant des arrangements ; chaque composante de la solution obtenue étant égale à la frontière de l'arbre de dérivation maximal correspondant.

Expressions régulières :

Pour résoudre les systèmes quasi-rationnels, Courcelle définit les expressions régulières à partir des opérations de concaténation, puissance ω et puissance $-\omega$. A ces opérations Heilbrunner ajoute celle de "shuffle" permettant la résolution de systèmes algébriques quelconques.

Sans rappeler les définitions précises, voyons les propriétés caractéristiques de ces opérations.

- **Puissance ω :**

Si a est un arrangement.

a^{ω} est l'arrangement solution de l'équation $x = a\,x$

d'où $a^{\omega} = \langle N, <, a\rangle$ $\quad a : N \to \{a\}$ application constante

a^{ω} = aaaa....... aaa....... une infinité de fois.

- **Puissance $-\omega$:**

$a^{-\omega}$ est la solution de l'équation $x = x\,a$

d'où $a^{-\omega} = \langle N, >, a\rangle$

$a^{-\omega}$ = ... aa....... aaa

- **Shuffle :**

Si $a_1, a_2, \ldots, a_n$ sont des arrangements.

Soit $A = \{a_1, \ldots, a_n\}$ $\qquad n \in N_+$

Le shuffle de A noté A^{η} est la solution de l'équation :
$x = x\,a_1\,x\,a_2\,x \ldots x\,a_n\,x$.

Si $A^{\eta} = \langle D, \Pi, f\rangle$ avec $f : D \to A$ on a la propriété :

$\forall x, y \in D$, $x \neq y$ et $x\,\Pi\,y$

$\forall a \in A$, $\exists z \in D$ tel que :

$x\,\Pi\,z$ et $z\,\Pi\,x$

et $f(z) = a$.

Les expressions régulières sont alors définies de la façon suivante :

X est un alphabet.

1) Toute lettre $x \in A$ est r.e. (expression régulière)

ε le mot vide est r.e.

2) Si R et R' sont r.e. alors

RR' est r.e.

3) $(R)^{\omega}$ et $(R^{-\omega})$ sont r.e. si R l'est.

4) Si $T = \{R_1, \dots, R_n\}$ où les R_i sont r.e.

alors T^{η} est r.e.

Heilbrunner [11] donne un algorithme permettant de résoudre tout système algébrique. Cet algorithme construit des expressions régulières représentant la solution du système égale au p-uple des frontières des arbres de dérivations maximaux du système ayant à la racine une variable.

La démarche de l'algorithme est similaire à celle que nous avons vue précédemment, elle consiste à décomposer le système en sous-systèmes, à résoudre séparément les sous-systèmes et par substitutions à obtenir la solution générale.

Nous allons donc utiliser ce résultat et transformer les expressions régulières représentant des arrangements en expressions rationnelles représentant des langages rationnels de $(X \cup \{\Omega\})^*$ et montrer que ces rationnels vérifient bien la propriété que nous voulons établir.

Soit un alphabet X ; considérons $A_\omega(X)$ l'ensemble des arrangements dénombrables sur X, et $W^\infty(X)$.

Soit g la transformation, qui à toute expression régulière représentant un arrangement de $A_\omega(X)$ associe un langage rationnel de $(X \cup \{\Omega\})^*$, définie récursivement par :

- $g(\varepsilon) = \varepsilon$
- $g(x) = x \quad \forall x \in X$
- $g(E \cdot E') = g(E) \cdot g(E')$, où E et E' sont r.e.
- $g(E^\omega) = (g(E))^* \, \Omega$
- $g(E^{-\omega}) = \Omega(g(E))^*$
- si $E = \{E_1, \ldots, E_k\}$ ensemble de r.e.
  $g(E^\eta) = (\Omega(g(E_1) + g(E_2) + \ldots + g(E_k)))^* \, \Omega$

Soit s l'application : $s : A_\omega(X) \to W^{Loc}(X)$ définie en II.3.3.

$\forall$ arbre t localement fini : $\phi(t) = s(\psi(t))$.

<u>Propriété</u> :

Soit E une expression régulière représentant un arrangement u de $A_\omega(X)$, alors :

g(E) est une partie dirigée de W(X)

et s(u) = sup(g(E)).

<u>Remarque</u> :

Cette propriété suffit pour établir le théorème dans le cas restreint des systèmes algébriques vérifiant : pour toute variable z du système S, l'arbre maximal de dérivation associé à z est localement fini ce qui est équivalent à : la plus petite solution de S ne contient pas de composante égale à Ω.

En effet :

$$S = \begin{cases} z_i = u_i \in (X \cup V)^* \\ i = 1,\ldots, p \end{cases}$$

Soit $E_1,\ldots, E_p$ les expressions régulières obtenues par l'algorithme de Heilbrunner, alors par la propriété :

$\forall i = 1,\ldots, p$

$g(E_i)$ est une partie dirigée de $W(X)$

$g(E_i)$ est une expression rationnelle

et si t_i est l'arbre de dérivation maximal associé à z_i :

$s(\psi(t_i)) = \phi(t_i)$

et $s(\psi(t_i)) = \mathrm{Sup}(g(E_i))$.

Or $\phi(t_i)$ est la $i^{\text{ème}}$ composante de la plus petite solution de S et donc $g(E_i)$ est bien un rationnel satisfaisant.

<u>Preuve de la propriété</u> :

- Si $E = \varepsilon$, u est le mot vide, $s(u) = \varepsilon$

  $g(E) = \varepsilon$ et $\sup(g(E)) = \varepsilon$.

- Si $E = x$, $x \in X$

  $u = x$, $s(u) = x = g(E)$, $\sup(g(E)) = x$.

- Si $E = E_1 E_2$ avec E_1 et E_2 vérifiant la propriété

  E_1 représentant l'arrangement u_1, $s(u_1) = \sup(g(E_1))$

  E_2 représentant l'arrangement u_2, $s(u_2) = \sup(g(E_2))$

On a $u = u_1 \; u_2$

$s(u) = s(u_1) \bullet s(u_2)$ par propriété de s

d'où $s(u) = \sup(g(E_1)) \bullet \sup(g(E_2))$

$g(E) = g(E_1) \bullet g(E_2)$ par définition

$g(E_1)$ et $g(E_2)$ étant des parties dirigées, $g(E)$ l'est également et donc

$\mathrm{Sup}(g(E)) = \sup(g(E_1) \bullet g(E_2))$

$= \sup(g(E_1) \bullet \sup(g(E_2))$ par propriété de C.P.O.

$= s(u)$.

Ces trois premiers points montrent en particulier que la propriété est vraie pour les arrangements finis (mots de X^*).

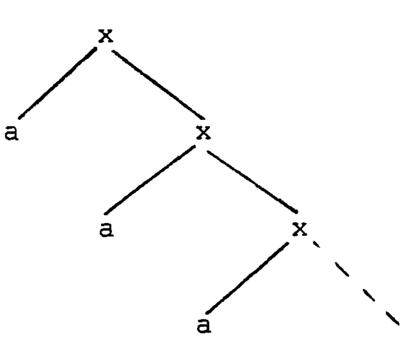
• Soit a une lettre de l'alphabet x

et $E = a^\omega$ représentant u

E est l'expression régulière solution de l'équation $x = a\,x$.

Soit t l'arbre maximal de dérivation associé à cette équation

t =



t est localement fini et donc $\Phi(t) = s(\Psi(t))$.

. $g(E) = a^{*}\,\Omega$ par définition.

On a immédiatement $\Phi(t) = \sup(g(E))$ et donc puisque $u \equiv \Psi(t)$, $s(u) = \sup(g(E))$.

• De la même manière, en utilisant l'équation $x = x\,a$, on obtient le résultat pour l'expression $E = a^{-\omega}$.

• Soient $a_1, \ldots, a_n$ des lettres de X

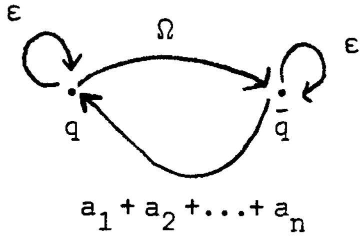
et $E = \{a_1, \ldots, a_n\}^{\eta}$ représentant u.

E est solution de l'équation $x = x\,a_1\,x\,a_2\,x \ldots x\,a_n\,x$

et $g(E) = (\Omega(a_1 + a_2 + \ldots + a_n))^{*}\,\Omega$, c'est évidemment une partie dirigée de W(X).

Comme précédemment soit t l'arbre de dérivation associé à x, t est localement fini d'où $\Phi(t) = s(\Psi(t)) = s(u)$.

Il reste à montrer que $\Phi(t) = \sup(g(E))$, ce qui peut être fait directement en considérant des approximations de t et $\Phi(t)$, ou en utilisant la preuve précédente et en construisant l'automate associé à l'équation :



qui donne le résultat.

• Il ne reste, pour établir la propriété, qu'à montrer qu'elle demeure valide lorsqu'on substitue une expression, la vérifiant, à une lettre dans une expression la vérifiant également.

Soit x une lettre de l'alphabet X.

Soient E et E_x des expressions régulières représentant respectivement les arrangements u et u_x de $A_\omega(X)$ et vérifiant la propriété.

i.e. $s(u) = \sup(g(E))$

$s(u_x) = \sup(g(E_x))$

Soit F l'expression régulière obtenue en substituant dans E l'expression E_x à chaque occurrence de la lettre x. Notation $F = E \bullet [x \mid E_x]$.

F représente l'arrangement $w \equiv u \bullet [x \mid ux]$.

(u auquel on substitue u_x à x).

• $s(w) = s(u \bullet [x \mid u_x]) = s(u) \bullet [x \mid s(u_x)]$ où $[x \mid s(u_x)]$ est l'homomorphisme qui a la lettre x associe le mot $s(u_x)$ et laisse les autres lettres inchangées.

• $g(F) = g(E \bullet [x \mid E_x]) = g(E) \bullet [x \mid g(E_x)]$ par propriété de g où $[x \mid g(E_x)]$ est une substitution dirigée associée à l'homomorphisme $[x \mid s(u_x)]$ puisque par hypothèse $s(u_x) = \sup(g(E_x))$, d'où :

g(F) est bien une partie dirigée de W(X),
et $\mathrm{Sup}(g(F)) = \sup \bullet (g(E) \bullet [x \mid g(E_x)])$

$= \sup \bullet (g(E)) \bullet [x \mid \sup(g(E_x))]$

$= s(u) \bullet [x \mid s(u_x)] = s(w)$

Remarque :

Dans le cas où le système algébrique n'est pas conforme à la restriction qu'il a fallu imposer pour la résolution par cette méthode (i.e. il existe une composante de la solution égale à Ω), cette méthode peut quand même être utilisée. Il suffit de modifier quelque peu l'algorithme de Heilbrunner en posant : si l'algorithme fournit l'expression ε pour une composante alors prendre l'expression Ω.

Exemple fin :

L'algorithme de Heilbrunner, donne les expressions :

pour l'équation $z_4 = a\ z_4$: $E_4 = a^{\omega}$

pour $z_2 = u\ z_2\ v\ z_3\ w$ $\qquad E_2 = u^{\omega} \cdot T^{\eta} \cdot t^{-\omega}\ w$

et $z_3 = x\ z_2\ y\ z_3\ t$ $\qquad E_3 = x\ u^{\omega} \cdot T^{\eta} \cdot t^{-\omega}$

avec $T = \{t^{-\omega}\ w\ v\ x\ u^{\omega},\ t^{-\omega}\ w\ y\ x\ u^{\omega}\}$

puis $E_1 = (a\ E_2\ b\ E_3\ c)^{\omega} \cdot (d\ E_4\ e)^{-\omega}$

Par la transformation g on obtient :

$R_4 = a^* \ \Omega$

$R_2 = u^* \ \Omega\ (t^*\ w\ (v+y)\ x\ u^*\ \Omega)^*\ t^*\ w$

$R_3 = x\ u^*\ \Omega\ (t^*\ w\ (v+y)\ x\ u^*\ \Omega)^* \cdot t^*$ après quelques simples transformations puis

$R_1 = (a\ R_2\ b\ R_3\ c)^*\ \Omega\ (d\ a^*\ \Omega\ e)^*$

3. Arbres algébriques

Définition :

Un arbre est algébrique si et seulement si il est une composante de la solution unique d'un système d'équations algébriques d'arbres sous forme normale de Greibach.

Soit Σ un alphabet gradué : $\Sigma = \Sigma_0 + \Sigma_1 + \ldots + \Sigma_n$, $n \in N_+$.

Soit un ensemble de variables graduées $Z = \{z_1, \ldots, z_p\}$.

Soit V un ensemble de variables syntaxiques.

Soit S un système algébrique d'arbres

$$S = \begin{cases} z_i(v_1, \ldots, v_{a_i}) = u_i \in T(\Sigma \cup Z \cup V_{a_i}) \\ i = 1, \ldots, p \end{cases}$$

où a_i est la arité de z_i

et $V_{a_i} = \{v_1, \ldots, v_{a_i}\} \subset V$.

S est supposé sous forme normale de Greibach

i.e. $\forall u_i$ $\quad u_i$ est de la forme

$$\begin{array}{ccc} & f & \\ \swarrow & & \searrow \\ t_1 & \ldots\ldots & t_p \end{array}$$

pour un $f \in \Sigma$.

Nous allons montrer que le problème de décision de l'égalité des arbres algébriques est réductible à celui de l'égalité des feuillages d'arbres algébriques.

Pour cela nous définissons le codage suivant :

Soit $\overline{\Sigma}$ l'alphabet gradué défini par :

- $\forall\alpha \in \Sigma_0$, $\alpha \in \overline{\Sigma}_0$

- $\forall f \in \Sigma \backslash \Sigma_0$

  si arité de $f = i$ (≥ 1)

  les nouveaux symboles $f_0, f_1, \ldots, f_i \in \overline{\Sigma}_0$

  et le nouveau symbole $\overline{f} \in \overline{\Sigma}_{2i+1}$

Soit le codage $c : T(\Sigma \cup V \cup Z) \to T(\overline{\Sigma} \cup V \cup Z)$ défini récursivement de la façon suivante :

1) • $c(\alpha) = \alpha \qquad \forall\alpha \in \Sigma_0$

2) • $c(v) = v \qquad \forall v \in V$

3) • $$c\left(\begin{array}{ccc} & z_i & \\ \diagup & & \diagdown \\ t_1 & & t_{a_i} \end{array}\right) = \begin{array}{ccc} & z_i & \\ \diagup & & \diagdown \\ c(t_1) & & c(t_{a_i}) \end{array}$$

   pour toute variable z_i de z

   et $t_1, \ldots, t_{a_i}$ de $T(\Sigma \cup V \cup Z)$

4) • $$c\left(\begin{array}{ccc} & f & \\ \diagup & & \diagdown \\ t_1 & & t_k \end{array}\right) = \begin{array}{ccccccc} & & & \overline{f} & & & \\ f_0 & c(t_1) & f_1 & \ldots & f_{k-1} & c(t_k) & f_k \end{array}$$

   pour tout $f \in \Sigma \backslash \Sigma_0$, k étant la arité de f

   et $t_1, \ldots, t_k \in T(\Sigma \cup Z \cup V)$.

Remarque :

Ce codage est injectif.

i.e. $\forall t, t' \in T(\Sigma \cup Z \cup V)$

$t = t'$ ssi $c(t) = c(t')$.

En ajoutant la règle :

5) • $c(\Omega) = \Omega$

Le codage respecte l'ordre sur les arbres

i.e. $\forall t, t' \in T_{\Omega}(\Sigma \cup V)$

$t < t'$ ssi $c(t) < c(t')$

et peut donc être étendu par continuité à :

$c : T_{\Omega}^{\infty}(\Sigma \cup V) \to T_{\Omega}^{\infty}(\overline{\Sigma} \cup V)$

en posant $c(\sup(D)) = \sup(c(D))$ pour toute partie dirigée de $T_{\Omega}(\Sigma \cup V)$.

On a alors la propriété :

$\forall t, t' \in T_{\Omega}^{\infty}(\Sigma \cup V)$

$t = t'$ ssi $c(t) = c(t')$

Propriété :

$\forall t, t' \in T_{\Omega}(\Sigma \cup V)$

$t < t'$ ssi $\phi(c(t)) \leq_{\Omega} \phi(c(t'))$

<u>Preuve</u> :

Dans un sens c'est immédiat puisque

$t < t' \iff c(t) < c(t') \Rightarrow \phi(c(t)) \leq_{\Omega} \phi(c(t'))$

par propriété du feuillage.

Inversement il faut montrer que :

$\phi(c(t)) \leq_{\Omega} \phi(c(t')) \Rightarrow c(t) < c(t')$.

Récurrence sur la hauteur de l'arbre t.

Notons $h(t)$ la hauteur de t

- si $h(t) = 0$, t est réduit à une lettre de $\Sigma_0 \cup V \cup \{\Omega\}$ et $c(t) = t$

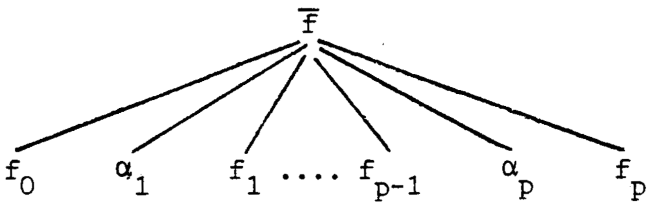
ou bien $t \neq \Omega$ et dans ce cas $t' = t$

ou bien $t = \Omega$ et bien entendu $t < t'$

- si $h(t) = 1$, t est alors de la forme :

$t = f(\alpha_1, \ldots, \alpha_p)$, les $\alpha_i \in \Sigma_0 \cup V \cup \{\Omega\}$

d'où $c(t) =$



et $\phi(c(t)) = f_0\, \alpha_1\, f_1\, \alpha_2 \ldots f_{p-1}\, \alpha_p\, f_p$

$\phi(c(t)) \leq_\Omega \phi(c(t'))$ implique que $\phi(c(t'))$ est de la forme :

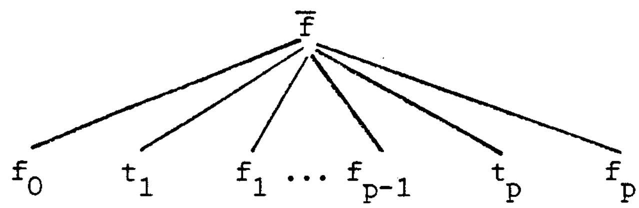
$$\phi(c(t')) = f_0\ \beta_1\ f_1\ \beta_2 \ldots\ f_{p-1}\ \beta_p\ f_p$$

avec $\beta_i = \alpha_i$ si $\alpha_i \neq \Omega$

et $\beta_i \in (\Sigma_0 \cup V \cup \{\Omega\})^*$ si $\alpha_i = \Omega$

et donc, d'après le codage, $c(t')$ est de la forme

$c(t') =$



avec $t_i = \alpha_i$ si $\alpha_i \neq \Omega$

et $t_i \in T_\Omega(\bar{\Sigma} \cup V)$ si $\alpha_i = \Omega$

ce qui prouve que $c(t) < c(t')$.

<u>**Hypothèse de récurrence**</u> :

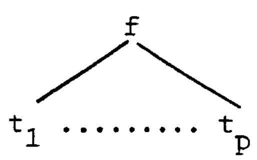
$\forall t \in T_\Omega(\Sigma \cup V)$ tel que $h(t) \leq n$, $n \in N_+$

$\forall t' \in T_\Omega(\Sigma \cup V)$

$\phi(c(t)) \leq_\Omega \phi(c(t')) \Rightarrow c(t) < c(t')$.

Soit un arbre t de hauteurs $n+1$

$t =$



, $f \in \Sigma_p$, $t_1, \ldots, t_p$ de hauteur inférieure ou égale à n.

$$c(t) = \begin{array}{c} \overline{f} \\ f_0 \qquad c(t_1) \qquad f_1 \;\ldots\; c(t_p) \qquad f_p \end{array}$$

$$\phi(c(t)) = f_0\ \phi(c(t_1)) f_1 \ldots\ f_{p-1}\ \phi(c(t_p))\ f_p$$

$\phi(c(t)) \leq_\Omega \phi(c(t'))$ implique, et par le codage que :

$$\phi(c(t')) = f_0\ \phi(c(t'_1))\ f_1 \ldots\ f_{p-1}\ \phi(c(t'_p))\ f_p$$

$$t'_1, \ldots, t'_p \in T_\Omega(\Sigma \cup V)$$

avec $\forall i = 1, \ldots, p \quad \phi(c(t_i)) \leq_\Omega \phi(c(t'_i))$

par hypothèse de récurrence on a donc :

$$\forall i = 1, \ldots, p \qquad c(t_i) < c(t'_i)$$

et donc $c(t) < c(t')$

puisque $t' = \begin{array}{c} f \\ t'_1 \quad \ldots \quad t'_p \end{array}$

ce qui établit la propriété.

<u>Conséquence</u> :

$$\forall t, t' \in T^\infty_\Omega(\Sigma \cup V)$$

$$t = t' \text{ ssi } \phi(c(t)) = \phi(c(t')).$$

Preuve :

· Dans un sens c'est évident puisque

$t = t'$ ssi $c(t) = c(t') \Rightarrow \phi(c(t)) = \phi(c(t'))$.

Soient t et $t' \in T^{\infty}_{\Omega}(\Sigma \cup V)$ tels que

$\phi(c(t)) = \phi(c(t'))$.

Soient D et D' des parties dirigées de $T_{\Omega}(\Sigma \cup V)$ ayant respectivement pour sup t et t' ; par l'extension de c on a :

$c(t) = \sup(c(D)) \qquad c(D) = \{c(t) \mid t \in D\}$

$c(t') = \sup(c(D'))$

d'où

$\phi(c(t)) = \phi(\sup(c(D))) = \sup(\phi(c(D))$

par propriété de ϕ

et $\phi(c(t')) = \sup(\phi(c(D)))$

donc : $\phi(c(t)) = \phi(c(t'))$ ssi $\forall$ arbre u de D

Il existe un arbre v de D' tel que $\phi(c(u)) < \phi(c(v))$

et réciproquement

ssi (par la propriété)

$\forall u \in D, \exists v \in D'$ tel que $u < v$

et réciproquement

ssi $t = t'$

Revenons au système algébrique S.

Soit le système

$$\overline{S} = \begin{cases} z_i(v_1,\ldots, v_{a_i}) = c(u_i) \\ i = 1,\ldots, p \end{cases}$$

associé à S. C'est évidemment un système algébrique sous forme normale de Greibach.

Soit $(t_1,\ldots, t_p)$ la solution de S, $t_i \in T^{\infty}(\Sigma \cup V_{a_i})$ il est immédiat que $(c(t_1),\ldots, c(t_p))$ est la solution de $\overline{S}$, les $c(t_i) \in T^{\infty}(\overline{\Sigma} \cup V_{a_i})$.

<u>Conséquence</u> :

Le problème de décidabilité de l'égalité des arbres algébriques est réductible à celui de l'égalité des feuillages d'arbres algébriques.

Puisque si t et t' sont algébriques, c(t) et c(t') le sont également, et $t = t'$ ssi $\phi(c(t)) = \phi(c(t'))$.

De façon plus générale, pour toute famille d'arbres close par le codage c, la décidabilité de l'égalité des feuillages implique la décidabilité de l'égalité des arbres de cette famille. C'est le cas en particulier pour les arbres réguliers.

V. QUELQUES PROPRIETES METRIQUES ET D'ORDRE DE $W^{\infty}(X)$

Nous savons déjà que le C.P.O. $W^{\infty}(X)$ est ω-algébrique c'est-à-dire qu'il possède une base finitaire et dénombrable. Cette propriété permet de définir une ultra-métrique (Comyn, Dauchet [5]) associée à la topologie de Lawson, sur le C.P.O. Certaines classes de C.P.O. ω-algébriques (infinitaire, SFP, 2/3 de SFP) possèdent de bonnes propriété concernant notamment l'espace métrique associé. Nous allons tout d'abord en rappeler les définitions et propriétés essentielles avant de situer notre C.P.O. parmi ces classes et de décrire sommairement la métrique obtenue.

1. Quelques rappels

1.1. C.P.O. des fonctions continues sur un C.P.O.

Soit P un C.P.O. ; notons $<$ l'ordre sur P et $\perp$ le plus petit élément.

Soit $[P \to P]$ = l'ensemble des fonctions continues de P dans lui-même, c'est-à-dire l'ensemble des fonctions f vérifiant :

$f(\sup(\Delta)) = \sup(f(\Delta))$ pour toute partie dirigée Δ.

Cet ensemble $[P \to P]$ est muni de l'ordre "par points" : $\forall f, g \in [P \to P]$

$f < g$ ssi $\forall x \in P$, $f(x) < g(x)$

Propriété :

$[P \to P]$ muni de l'ordre par points est un C.P.O. Son plus petit élément est la fonction $f_{\perp}$ définie par : $\forall x \in P \quad f_{\perp}(x) = \perp$

1.2. Topologie de Lawson, métrique associée

Soit B^{∞} un C.P.O. ω-algébrique de base finitaire B. La topologie de Lawson sur B^{∞} a comme sous-base les ensembles :

$P_b = \{x \in B^\infty / x > b\}$

et $N_b = \{x \in B^\infty / x \not> b\}$

pour tout élément $b \in B$.

P_b peut être interprété comme l'ensemble des éléments qui vérifient la "propriété b" ("information positive") et N_b comme l'ensemble des éléments ne vérifiant pas la "propriété b" ("information négative").

Cette topologie est séparée (de Hausdorff) c'est-à-dire que deux éléments quelconques distincts possèdent des voisinages disjoints.

Une suite $(x_n)_{n \in N}$ converge vers x pour cette topologie ssi :

pour tout élément y

- $y < x$ ssi $y < x_n$ pour tout n sauf un nombre fini

et

- $y \not< x$ ssi $y \not< x_n$ pour tout n sauf un nombre fini

Propriété :

Toute suite croissante $(x_n)_{n \in N}$ converge vers $\sup\{x_n \mid n \in N\}$.

Métrisation :

Notons $x \Delta y$ la différence symétrique entre Idéal(x) et Idéal(y), pour tout $x, y \in B^\infty$.

$(\text{Idéal}(z) = \{w \in B / w < z\})$

$x \Delta y = \{z \in B / z < x \text{ et } z \not< y$

$\text{ou } z \not< x \text{ et } z < y\}$

= l'ensemble des éléments de la base qui "séparent" x et y.

Soit $\nu : \mathbb{N} \to B$ une énumération de la base B.

Notons $\mu n[\nu(n) \in x \Delta y]$ pour le plus petit n tel que $\nu(n) \in x \Delta y$, en convenant qu'il est égal à $+\infty$ si $x \Delta y = \Phi$ (ssi x = y).

Soit $d_\nu : B^\infty \times B^\infty \to \mathbb{R}^+$ l'application définie par :

$\forall x, y \in B^\infty$

$$d_\nu(x, y) = \frac{1}{1 + \mu n[\nu(n) \in x \Delta y]}$$

<u>Propriété 1</u> :

d_ν est une ultra-métrique c'est-à-dire vérifie :

- $d_\nu(x, y) = 0$ ssi $x = y$
- $d_\nu(x, y) = d_\nu(y, x)$
- $d_\nu(x, z) \leq \mathrm{Max}\{d_\nu(x, y), d_\nu(y, z)\}$

pour tout x, y, z.

<u>Propriété 2</u> : (Comyn [4])

La topologie induite par d_ν ne dépend pas de l'énumération ν.

$\forall \nu, \nu'$ énumérations de B
d_ν et $d_{\nu'}$ sont uniformément équivalentes
($\Rightarrow$ toute suite de Cauchy pour d_ν est de Cauchy pour $d_{\nu'}$).

<u>Propriété 3</u> :

Pour tout C.P.O. ω-algébrique, d_ν définit une ultramétrique associée à la topologie de Lawson. (La topologie induite par d_ν et la topologie de Lawson sont identiques).

Cette propriété montre en particulier que toute suite croissante est convergente pour d_ν.

1.3. Les SFP

Soit P un C.P.O., $<$ l'ordre sur P.

Soit D une partie de P.

Notons U(D) l'ensemble des majorants minimaux de D c'est-à-dire l'ensemble des éléments minimaux parmi l'ensemble des majorants de D.

$$U(D) = \{x \in P \ / \ - \forall y \in D,\ y < x \quad \text{et} \ - \forall z \in P \text{ tel que } y < z \quad \forall y \in D \quad \text{on a } z < x \Rightarrow x = z\}.$$

U(D) est dit **complet** si tout majorant de D majore un élément de U(D), autrement dit :

$\forall z \in P$ tel que : $y < z,\ \forall y \in D$

$\exists x \in U(D)$ tel que $x < z$.

On notera $U^*(D)$ le plus petit ensemble contenant D et fermé par U c'est-à-dire vérifiant :

$\forall Y \subset U^*(D),\ U(Y) \subset U^*(D)$

Remarque :

En notant $U^o(D)$ pour D et $U^{n+1}(D) = U(U^n(D)) \quad \forall u \in N$

On a d'après la définition de U^* :

$$\bigcup_{i \geq 0} (U^i(D)) \subset U^*(D)$$

l'inclusion inverse n'étant pas vraie, en général.

Plotkin [16] définit les SFP (Sequences of Finite C.P.O.) comme limites de suites de C.P.O. finis et les caractérise de la façon suivante :

Caractérisation :

Un C.P.O. est un SFP si et seulement si il est ω-algébrique et vérifie :

- pour toute partie finie D de la base finitaire, U(D) est complet et $U^*(D)$ est fini.

Une autre caractérisation a été donnée par Smith [18] :

Propriété :

Un C.P.O. P est un SFP ssi P et $[P \to P]$ sont ω-algébriques.

1.4. Les 2/3 SFP

Définition :

Un C.P.O. est 2/3 SFP si il est ω-algébrique et vérifie :

- pour toute partie finie D de la base finitaire U(D) est fini et complet.

La différence avec les SFP est la suppression de la condition de finitude de $U^*(D)$. Tout SFP est donc évidemment 2/3 SFP.

Plotkin [16] donne une caractérisation topologique des 2/3 SFP.

Propriété :

Un C.P.O. ω-algébrique B^∞ est un 2/3 SFP si et seulement si la topologie de Lawson sur B^∞ est compacte (i.e. puisque c'est un espace métrique, toute partie infinie de B^∞ admet un point d'accumulation).

Conséquence :

Si B^∞ est 2/3 SFP, B^∞ muni de l'ultramétrique d_ν est complet (toute suite de Cauchy converge).

Il y a donc isomorphie entre B^{∞} le complété, au sens de l'ordre. de B et le complété métrique de B. i.e. Les objets obtenus par sup de parties dirigées de la base et ceux par limites de suites de Cauchy sont les mêmes.

2. Propriété de $W^{\infty}(X)$

Un C.P.O. est dit "conditionnellement complet" si toute partie majorée possède un sup.

Un C.P.O. est "infinitaire" si il est ω-algébrique et conditionnellement complet. Tout C.P.O. infinitaire est évidemment SFP.

Propriété :

$W^{\infty}(X)$ n'est pas un C.P.O. infinitaire il suffit de considérer l'exemple simple suivant :

$D = \{\Omega a \Omega,\ \Omega b \Omega\}$

D est majorée par $\Omega a \Omega b \Omega$ et $\Omega b \Omega a \Omega$ mais n'a pas de sup.

Théorème :

$W^{\infty}(X)$ est un 2/3 SFP, mais pas un SFP.

Preuve :

Montrons tout d'abord que $W^{\infty}(X)$ n'est pas un SFP.

Reprenons l'exemple précédent : $D = \{\Omega a \Omega,\ \Omega b \Omega\}$

On a facilement : $U(D) = \{\Omega a \Omega b \Omega,\ \Omega b \Omega a \Omega\}$

puis : $U^2(D) = \{\Omega a \Omega b \Omega a \Omega,\ \Omega b \Omega a \Omega b \Omega\}$

et de façon générale :

$$U^{2n}(D) = \{(\Omega a \Omega b)^n \ \Omega a \Omega, \ (\Omega b \Omega a)^n \ \Omega b \Omega\}$$

$$\text{et } U^{2n+1}(D) = \{(\Omega a \Omega b)^{n+1} \ \Omega, \ (\Omega b \Omega a)^{n+1} \Omega\}$$

pour tout $n \in N$.

Ce qui montre que $U^{*}(D)$ est infini, et donc que $W^{\infty}(X)$ n'est pas SFP.

Il est facile de voir que pour tout D fini, $D \subset W(X)$, $U(D)$ est fini. En supposant le contraîre $U(D)$ contiendrait des mots arbitrairement longs ce qui est impossible.

D'autre part si $U(D)$ n'est pas complet alors il existe une suite infinie décroissante, de majorants de D, formée de mots finis ; ceci est impossible puisque pour tout mot fini w, Idéal(w) = $\{z / z \leq_{\Omega} w\}$ est fini.

De façon plus générale on a :

<u>Propriété</u> :

Tout C.P.O. B^{∞} possédant une base dénombrable D et vérifiant : $\forall w \in B$, Idéal(w) fini est ω-algébrique et vérifie :

- $\forall D \subset B$, D fini : $U(D)$ est complet

c'est immédiat.

<u>Métrisation de $W^{\infty}(X)$</u> :

$W^{\infty}(X)$ étant un C.P.O. ω-algébrique, nous pouvons définir une ultramétrique, associée à la topologie de Lawson, à partir d'une énumération de la base finitaire $W(X)$.

Afin d'avoir une idée de ce que représentent les suites de Cauchy pour une métrique de ce type, nous allons choisir une énumération particulière de la base.

Définition :

P étant un ensemble ordonné dénombrable une énumération ν de P est dite "compatible avec l'ordre" si elle vérifie :

$$\forall x, y \in P,\ x < y \Rightarrow \mu n[\nu(n) = x] < \nu n[\nu(n) = y]$$

où $\mu n[\nu(n) = x]$ représente le plus petit entier n tel que $\nu(n) = x$.

Propriété : (Comyn [4])

Soit B^∞ un C.P.O. ω-algébrique de base finitaire B.

Il existe une énumération $\nu : \mathbb{N} \to B$ compatible avec l'ordre si et seulement si pour tout élément b de la base, Idéal(b) est fini,

où Idéal(b) = $\{x \in B \,/\, x < b\}$.

Cette propriété nous assure qu'une telle énumération existe pour W(X).

Remarque :

L'ordre $\leq_\Omega$ sur W(X) respecte, en quelque sorte, les "longueurs" de mot, longueur signifiant nombre d'occurrences de lettres de X ($\neq \Omega$) i.e. $\forall u, v \in W(X)$

$$u \leq_\Omega v \Rightarrow |u|_X \leq |v|_X$$

Pour obtenir alors une énumération de W(X) compatible avec l'ordre il faut énumérer par ordre croissant des longueurs (sur X) de mot et parmi tous les mots ayant même longueur (sur X) énumérer par ordre décroissant du nombre d'occurrence du symbole Ω (dans les représentants canoniques).

Soit donc $\nu : \mathbb{N} \to W(X)$ une énumération compatible avec l'ordre $\leq_\Omega$, et d_ν l'ultramétrique associée :

$$d_\nu : W^\infty(X) \times W^\infty(X) \to \mathbb{R}^+$$

$$\forall x, y \qquad d_\nu(x, y) = \frac{1}{1 + \mu n[\nu(n) \in x \Delta y]}$$

Notons n_0 pour $\mu n[\nu(n) \in x \Delta y]$.

Il est clair, par propriété de l'énumération ν, que le mot $\nu(n_0)$ a une longueur sur X ($|\nu(n_0)|_X$) minimale parmi ceux de $x \Delta y$, i.e.

$$\forall w \in x \Delta y, \ |\nu(n_0)|_X \leq |w|_X$$

La distance, d_ν, entre deux mots est donc directement liée à la longueur du plus petit mot qui les séparent.

Plus précisément, soit $d : W^\infty(X) \times W^\infty(X) \to \mathbb{R}^+$ définie par :
$\forall x, y \in W^\infty(X)$

$$d(x, y) = \frac{1}{\mathrm{Inf}\{|w|_X / w \in x \Delta y\}} \text{ si } x \neq y$$

$$d(x, y) = 0 \quad \text{si } x = y \text{ (ssi } x \Delta y = \Phi).$$

<u>Propriété</u> :

d est une ultramétrique sur $W^\infty(X)$ uniformément équivalente à d_ν.

<u>Preuve</u> :

- $d(x, y) = 0$ ssi $x = y$ par définition

- $d(x, y) = d(y, x)$ puisque $x \Delta y = y \Delta x$

- pour montrer que $d(x, z) \leq \mathrm{Max}\{d(x, y), d(y, z)\}$

. Il suffit de remarquer que :

$\forall w \in x \Delta z$, on a ($w \in x \Delta y$ ou $w \in y \Delta z$)

en notant n_{xy} pour $\mathrm{Inf}\{|w|_X / w \in x \Delta y\}$

on obtient donc $n_{xz} \geq \mathrm{Min}\{n_{xy}, x_{yz}\}$

et donc l'inégalité cherchée.

Quant à l'équivalence avec d_ν, elle est évidente d'après la définition de d et les remarques sur d_ν.

Les résultats concernant les 2/3 de SFP, montrent que $W^\infty(X)$ muni de la métrique d_ν (ou d) est complet (toute suite de Cauchy converge), et donc toute limite de suite de Cauchy est sup d'une partie dirigée de W(X).

Exemples :

Soit la suite : $\{a^n / n \in N\}$.

Cette suite est de Cauchy et converge vers un élément de $W^\infty(X)$:

- $d(a^n, a^{n+p}) = \frac{1}{n+1}$

On peut voir que sa limite est égale à $\sup(a^* \Omega a^*)$

Soit la suite $\{a^n b / n \in N\}$; elle est également de Cauchy et converge vers $\sup\{a^* \Omega a^* b\}$.

La suite $\{a^n b^n / n \in N\}$ converge vers $\sup\{a^* \Omega a^* b^* \Omega b^*\}$.

De façon générale, étant donné une suite de Cauchy, déterminer une partie dirigée (ou suite croissante) ayant même limite, ne paraît pas évident à priori.

Ce problème est résolu pour les C.P.O. infinitaires (Comyn [4]) et les SFP, mais pas pour les 2/3 SFP.

REFERENCES

[1] A. ARNOLD, M. NIVAT
"The metric space of infinite trees ; algebraical and topological properties".
Annales societatis mathematicae polonae, Série IV : Fundamenta Informaticae III-4, 1980, pp. 445-476.

[2] L. BOASSON, M. NIVAT
"Adherences of languages".
JCSS 20, 1980, pp. 285-309.

[3] G. BOUDOL
"Sémantique opérationnelle et algébrique des programmes récursifs non-déterministes".
Thèse d'Etat, LITP 80-28, 1980.

[4] G. COMYN
"Objets infinis calculables".
Thèse d'Etat, Lille I, 1982.

[5] G. COMYN, M. DAUCHET
"Approximations of infinitars objects".
9th Colloquium ICALP, Lecture Notes in Computer Science, Springer-Verlag, n° 140, pp. 116-127.

[6] B. COURCELLE
"Frontiers of infinite trees".
RAIRO Informatique Théorique, Vol. 12, pp. 319-337.

[7] B. COURCELLE
"Fundamental properties of infinite trees".
TCS 25, pp. 95-169.

[8] B. COURCELLE, M. NIVAT
"The algebraic semantics of recursive program schemes".
7th Symposium MFCS. Lecture Notes in Computer Science, n° 64, 1978, pp. 16-30.

[9] G. COUSINEAU
"Les arbres à feuilles indicées : un cadre algébrique pour l'étude des structures de contrôle".
Thèse d'Etat, 1977.

[10] S. EILENBERG
"Automata, languages and machines".
Vol. A, Academic Press, 1974.

[11] S. HEILBRUNNER
"An algorithm for the solutions of fixed point equations for infinite words".
RAIRO Informatique Théorique, Vol. 14, 1980, pp. 131-141.

[12] M. NIVAT
"Languages algébriques sur le magma libre et sémantique des schémas de programmes".
In Automata, Languages and Programming, 1st Colloquium, Springer-Verlag, 1973, pp. 293-307.

[13] M. NIVAT
"Sur les ensembles de mots infinis engendrés par une grammaire algébrique".
RAIRO Informatique Théorique, Vol. 12, 1978, pp. 259-278.

[14] M. NIVAT
"Infinite words, infinite trees, infinite computations".
In Mathematical Centre Tracts, Vol. 109, 1979, pp. 1-52.

[15] M. NIVAT, D. PERRIN
"Ensembles reconnaissables de mots bi-infinis".
LITP 81-52, 1981.

[16] G. PLOTKIN
"A powerdomain construction".
SIAM Journ. Comput. 5, 1976, pp. 452-487.

[17] G. PLOTKIN
"*I^{ω} as a universal domain*".
JCSS 17, 1978, pp. 209-236.

[18] M. SMITH
"*Powerdomains*".
JCSS 16, 1978, pp. 23-26.

[19] S. TISON
"*Mots infinis et processus. Objets infinitaires et topologie*".
Thèse de 3ème Cycle, Lille 1, 1983.

[20] S. TISON, M. DAUCHET, G. COMYN
"*Metrical and ordered properties of powerdomains*".
F.C.T. 1984.

RESUME :

L'opération de feuillage d'arbres est définie par des spécifications pour les arbres infinis. Le feuillage initial est obtenu par des constructions standards ; les "mots" infinis ainsi construits sont plus généraux que les mots infinis usuels et que les mots bi-infinis. En résolvant les systèmes d'équations algébriques de mots, on en déduit la décidabilité de l'égalité des feuillages d'arbres réguliers.

MOTS-CLES : Arbre, C.P.O., feuillage, frontière, infini, mot, terme.