Mathématiques et sciences humaines

Marcel Cori

L’interrogation : ébauche d’une représentation formelle par des équations

Mathématiques et sciences humaines, tome 77 (1982), p. 127-154
<http://www.numdam.org/item?id=MSH_1982__77__127_0>

Math. Sci. hum. (*20*^{e} année, n°77, *1982*, p.127-154)

L'INTERROGATION : EBAUCHE D'UNE REPRESENTATION FORMELLE PAR DES EQUATIONS

MARCEL CORI (*)

Le problème de l'interrogation se pose aux informaticiens dans au moins deux domaines d'application : les systèmes "questions-réponses" et l'interrogation de bases de données. Il s'agit dans les deux cas de données liées par des relations, et l'on veut obtenir la restitution d'une partie de ces données et des relations qui les lient. La seule différence est dans le volume de données que l'on manipule, beaucoup plus important dans le deuxième cas.

Si l'on veut que l'interrogation se fasse en utilisant un langage "quasi-naturel" (au sens d'un langage artificiel semblable, pour l'utilisateur, à la langue qu'il emploie), la démarche qui s'impose est la suivante :

1°) utiliser un système de représentation métalinguistique et des procédures qui associent à chaque énoncé une représentation dans le système choisi ;

2°) avoir une représentation des questions compatible avec la représentation des énoncés enregistrés ;

3°) traiter l'interrogation proprement dite au moyen d'une procédure qui permette une restitution (éventuellement partielle) des données en fonction de la question.

(*) Université Paris VII, Département d'Informatique Générale, Tour 13, 1er étage, 75251 Paris Cédex 05.

Ce travail a été effectué dans le cadre de l'ERA 642 du CNRS (PITFALL Projet Interdisciplinaire de Traitement Formel et Automatique des Langues et du Langage) pour l'ATP "Informatique et Sciences Humaines".

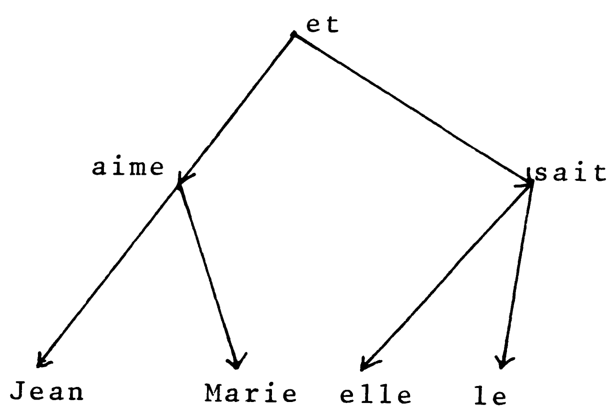
La tentative de résoudre ces problèmes nous a amené à étudier des structures définies à partir des graphes sans circuit.

En effet, la plupart des systèmes de représentation métalinguistique qui existent, que ce soient les grammaires syntagmatiques [CHO,57] , les grammaires transformationnelles [CHO,57] [HAR,65] ou les grammaires de dépendance [TES,59] [HAY,62] aboutissent à la construction d'arborescences. Ce qui ne permet pas, entre autres, de rendre compte des phénomènes anaphoriques. Il est par là illusoire de traiter efficacement l'interrogation, même dans des cas simples.

Donnons un exemple. Prenons l'énoncé :

(1) <u>Jean aime Marie et elle le sait</u>

On le représenterait à l'aide d'un "arbre de dépendance" de la façon suivante :



Ce type de représentation ne permet pas de répondre à des questions aussi élémentaires que :

<u>Que sait Marie ?</u>

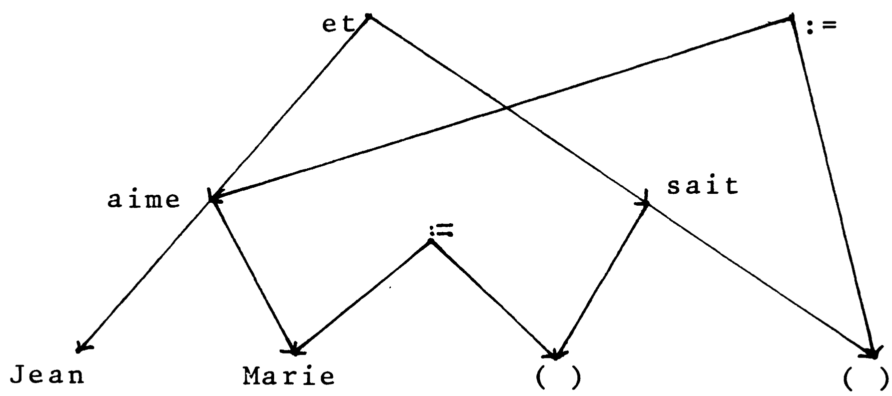
<u>Qui sait que Jean aime Marie ?</u>

C'est pourquoi nous nous intéressons particulièrement aux représentations formelles issues de la théorie des opérations prédicatives et énonciatives développées dans le cadre du programme PITFALL (voir notamment [CLI/DES, 1975]). L'anaphore est représentée par une relation d'"identification" notée ":=". Cette relation est constituée par un opérateur d'identification unaire qui identifie une place d'objet (représentée par "( )") insérée dans une relation prédicative avec un autre objet déjà déterminé ; il en résulte une relation d'identification entre cette place d'objet qui est instanciée et l'objet déterminé. Ainsi, dans l'énoncé (1), le marqueur elle est la trace linguistique de l'identification de la place vide, dans la forme propositionnelle

( ) le sait

et du terme Marie. L'identification a un statut qui est défini précisément dans [DES, 80].

La relation prédicative sous-jacente à l'énoncé (1) est représentée par :



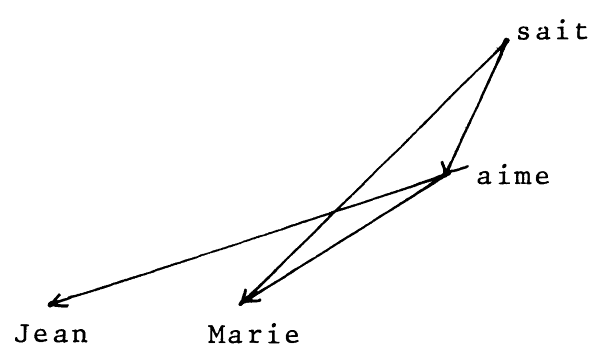
(Nous n'indiquons pas, sur cette représentation, les sortes des objets identifiés, sortes qu'il faut prendre en compte).

Cette représentation ne tient pas compte des conditions d'énonciation responsables notamment des marqueurs de temps et de personne. On a obtenu une approximation (justifiée par des objectifs finalisés) d'une représentation plus complexe où seraient inscrites les conditions d'énonciation.

Une application fondée sur ce modèle a été développée par M.J. Blosseville et R. Bouilloux [BLO/BOU, 80] . Leur objectif était : 1- la représentation d'énoncés empruntés à un corpus juridique, 2- la construction d'une base de données à partir de ce corpus, 3- l'interrogation de cette base de données dans un langage quasi-naturel. Ils ont, pour ce faire, restreint le système de représentation métalinguistique à un sous-système plus simple, en spécifiant les approximations retenues et donc les limites de ce sous-système.

Répondre à une question se ramène à la recherche d'une sous-structure dans le graphe représentatif d'un énoncé, immédiatement ou après transformation de ce graphe. Cela est effectué au moyen d'algorithmes généraux décrits dans [COR, 80]. Mais la détermination d'une réponse peut donner lieu à une combinatoire importante, croissant très rapidement avec la longueur de l'énoncé pris en compte.

Nous proposons donc ici une représentation légèrement différente : les sommets liés par une identification sont fusionnés afin d'obtenir des graphes sans circuit plus "compacts". Ainsi oublie-t-on l'ordre des unités lexicales qui apparaissent dans un énoncé. Cela donne, sur notre exemple :

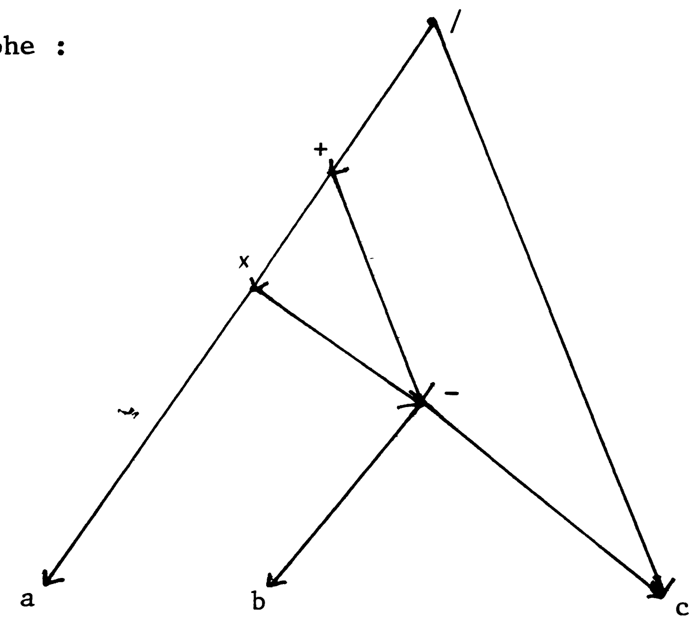


Plusieurs occurrences d'un même syntagme étant confondues (par exemple Marie ou Jean aime Marie), il en résulte une diminution de la combinatoire dans la recherche de sous-graphes. D'où des avantages qui compensent, du point de vue du traitement automatique, les approximations introduites.

Notre approche s'inspire notamment d'une technique informatique particulière utilisée en compilation (c'est-à-dire dans la traduction de programmes écrits dans des langages de haut niveau en suites d'instructions exprimées dans un code immédiatement compréhensible par la machine). L'analyse syntaxique du programme conduit à la construction d'arbres ; mais si l'on veut optimiser la génération de code, il est nécessaire de rendre compte de l'"anaphore", c'est-à-dire du fait qu'un même nom fait référence à un même objet, ou que deux occurrences de sous-arbres représentent les mêmes calculs. Pour cela on transforme les arbres en graphes sans circuit, ou "DAGs" (Directed Acyclic Graphs) [AHO/ULL, 1977]. Donnons un exemple très simple. L'expression arithmétique

$$[a \times (b - c) + (b - c)] / c$$

est représentée par le graphe :



Ainsi, lors de l'exécution d'un programme contenant cette expression, on effectue le calcul de (b - c) une seule fois.

Mais les DAGs n'ont pas reçu une formalisation précise. Nous avons donc défini les structures de base que nous utilisons : les "grillages" [COR, 80]. Ce sont des graphes sans circuit auxquels peuvent être associés canoniquement des opérateurs. On constate une nouvelle fois que certains outils mathématiques servent à la fois en linguistique et en informatique théoriques.

Dans le présent article : 1- nous précisons les définitions des structures qui nous intéressent, 2- nous définissons des morphismes qui permettent de formaliser l'interrogation, aboutissant à la notion d'"équation" définie sur ces structures, 3- nous illustrons notre propos par des exemples de représentations d'énoncés ainsi que de questions et nous donnons une idée des procédures d'interrogation sur ces exemples.

0. NOTATIONS

0.1 S^* désigne le monoïde libre sur l'ensemble S.

0.2 A\B est l'ensemble des éléments de l'ensemble A qui ne sont pas dans l'ensemble B.

0.3 La composée de 2 applications :

$$f : X \to Y \quad \text{et} \quad g : Y \to Z \quad \text{est notée :} \quad gf : X \to Z$$

0.4. Les notations sur les graphes sont pour la plupart tirées de [BER, 58]. Un graphe est donné par un couple $G = \langle X, \gamma \rangle$ où X est un ensemble (de sommets) et γ une application de X dans l'ensemble de ses parties.

Pour un élément x de X, $\gamma(x)$ désigne l'ensemble de ses descendants immédiats.

La fermeture transitive de γ, notée $\hat{\gamma}$, définit les descendants de tout sommet. Enfin, si Y est un sous-ensemble de X, on note par extension :

$$\hat{\gamma}(Y) = \bigcup_{x \in Y} \hat{\gamma}(x)$$

0.5 Si $G = \langle X, \gamma \rangle$ est un graphe, l'ensemble des sommets terminaux de G, c'est-à-dire des sommets qui n'ont pas de descendant immédiat, est noté Tèr (G).

1 - LES STRUCTURES

1. 1 Grillages

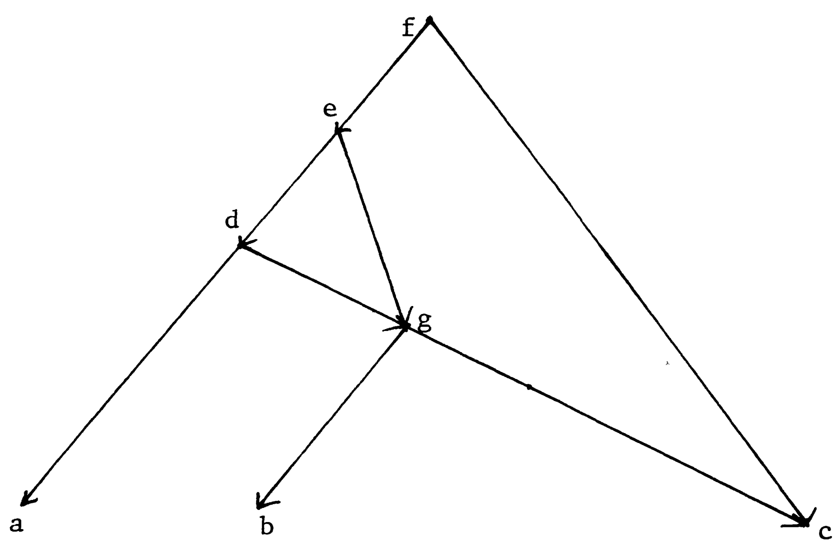
Définition : Un grillage est un triplet $G = \langle X, \theta, \gamma \rangle$ où X est un ensemble fini, θ une relation d'ordre total sur X et $< X, \gamma>$un graphe sans circuit.

La suite constituée par tous les sommets de X ordonnés selon θ sera notée X_θ . De même pour tout x de X on notera $\gamma_\theta(x)$ la suite ordonnée de ses descendants immédiats.

Exemple : $G = <X, \theta, \gamma>$ avec :

$X_\theta = <a, d, b, e, g, f, c>$

x	a	b	c	d	e	f	g
$\gamma(x)$	ϕ	ϕ	ϕ	{a,g}	{d,g}	{c,e}	{b,c}



1. 2 S-grillages

On se donne un ensemble fini non vide S qui est un vocabulaire de "sortes". Ces sortes permettent de représenter, en informatique, des notions caractérisant les objets manipulés, telles que : "entier", "réel", "booléen", ou, en linguistique, les catégories primitives "nom" et "proposition".

Définition : Un S-grillage est un 4-uple $G = <X, \theta, \gamma, \xi>$ où $<X, \theta, \gamma>$ est un grillage à chaque sommet duquel est attribuée une sorte par l'application ξ de X dans S.

Exemple : On prend S = {m, n, r, z} et on définit l'application ξ sur les sommets du grillage de l'exemple précédent par :

x	a	b	c	d	e	f	g
$\xi(x)$	n	n	m	z	z	r	z

1. 3 Σ-grillages

1. 3. 1 Définitions ([DES, 76], [DES, 80]) :

Un S-schéma (d'opérateurs) est un couple $\Sigma = \langle V, \tau \rangle$ où V est un ensemble fini non vide de "vocables" et τ une application :

$$\tau : V \to S \times S^{*}$$

$$v \to \langle s, \mu \rangle$$

Un élément $\sigma = \langle v, s, \mu \rangle$ de Σ est un opérateur simple, v en est le vocable, $\tau(\sigma) = \langle s, \mu \rangle$ en est le type (fonctionnel). Plus précisément, on note :

$\tau_1(\sigma) = s$ (type source)

$\tau_2(\sigma) = \mu$ (type but)

L'opérateur est dit "n-aire" si n est la longueur de μ. En particulier, si $\tau_2(\sigma) = \lambda$ (mot vide), l'opérateur est 0-aire.

1. 3. 2 Définition : Un Σ-grillage est un 5-uple $H = \langle X, \theta, \gamma, \xi, \omega \rangle$ où $\langle X, \theta, \gamma, \xi \rangle$ est un S-grillage et ω une application associant un opérateur à chaque sommet non-terminal et, éventuellement, à certains sommets terminaux du grillage :

$$\omega : X^{\circ} \to \Sigma$$

(X° est un sous-ensemble de X qui comprend tous les sommets non-terminaux du graphe).

ω doit vérifier la condition suivante : pour tout x de X°, si

$\gamma_\theta(x) = \langle x_1, \ldots, x_n \rangle$, alors : $\tau\omega(x) = \langle \xi(x), \xi(x_1) \ldots \xi(x_n) \rangle$

<u>Remarque</u> : Si x est terminal, l'opérateur associé $\omega(x)$ est 0-aire.

<u>Définition</u> : le Σ-grillage H est <u>complet</u> si et seulement si X° = X.

<u>Exemple</u> : On se donne un S-schéma

$\Sigma = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4\}$ avec :

$\sigma_1 = \langle /, \underline{r}, \underline{z}\,\underline{m} \rangle$

$\sigma_2 = \langle +, \underline{z}, \underline{z}\,\underline{z} \rangle$

$\sigma_3 = \langle \times, \underline{z}, \underline{n}\,\underline{z} \rangle$

$\sigma_4 = \langle -, \underline{z}, \underline{n}\,\underline{m} \rangle$

On reprend le S-grillage de l'exemple précédent et on définit l'application ω par :

$\omega : \{d, e, f, g\} \rightarrow \Sigma$

x	d	e	f	g
$\omega(x)$	σ_3	σ_2	σ_1	σ_4

On voit que ω vérifie bien la condition pour que l'on obtienne un Σ-grillage.

Si l'on associe à chaque sorte de S un ensemble (par exemple à $\underline{z}$ l'ensemble des entiers relatifs) et à chaque opérateur simple de Σ une opération (par exemple à σ_2 l'addition entre entiers) ce Σ-grillage pourra servir à représenter l'expression arithmétique citée en introduction. On voit, sur cet exemple, pourquoi il n'est pas nécessaire que des opérateurs soient associés aux sommets terminaux du graphe.

Remarque : Etant donné un S-schéma Σ, il a été défini récursivement dans [DES, 80] l'ensemble $\mathbb{T}[\Sigma]$ de tous les multiopérateurs sur Σ. $\mathbb{T}[\Sigma]$ contient notamment les opérateurs simples de Σ et est fermé pour 2 opérations appelées "greffe" et "intrication". Si l'on munit Σ d'une certaine "sémantique", à tout multiopérateur σ correspond une application $\sigma^{\bullet}$; de même à tout Σ-grillage H est associée une application f_H.

Dans [COR, 80] ont été décrites des procédures qui à tout multi-opérateur σ associent un Σ-grillage H (et réciproquement) de telle manière que les applications $\sigma^{\bullet}$ et f_H soient identiques.

C'est afin d'assurer une telle liaison entre opérateurs et graphes sans circuit qu'a été introduit l'ordre sur les sommets des graphes. En premier lieu, il est nécessaire que l'ensemble des descendants immédiats d'un même sommet soit totalement ordonné pour que puisse être associé naturellement un opérateur à chaque sommet non-terminal ; et si l'on veut associer un multi-opérateur au Σ- grillage, il est indispensable que les différents ordres soient compatibles.

Ceci induit un ordre partiel sur les sommets. S'il a été introduit un ordre total contenant cet ordre partiel, c'est afin de donner une plus grande commodité à la définition de structures : sans cela il faudrait vérifier pour toute structure si l'ordre partiel a été défini correctement, cette vérification étant non-triviale.

2 - MORPHISMES ET SOUS-STRUCTURES

2. 1 Q-morphismes de grillages

2. 1. 1 Définition : Soient $G = \langle X, \theta, \gamma \rangle$ et $G' = \langle X', \theta', \gamma' \rangle$ 2 grillages, $f : X \to X'$ une application. f est un Q-morphisme de G dans G' si et seulement si :

$$\forall x \in X \setminus \mathrm{Ter}(G) \qquad f\,\gamma_\theta(x) = \gamma'_{\theta'}\, f(x)$$

2. 1. 2 Propriété : La composée de 2 Q-morphismes est aussi un Q-morphisme.

Preuve : Soient les grillages $G = \langle X, \theta, \gamma\rangle$, $G' = \langle X', \theta', \gamma'\rangle$

et $G'' = \langle X'', \theta'', \gamma''\rangle$ et soient $f : X \to X'$ et $g : X' \to X''$ 2 Q-morphismes.

Soit x un sommet non terminal de G. f(x) est un sommet non terminal de G'. On a donc :

$$g\, \gamma'_{\theta'}\, f(x) = \gamma''_{\theta''}\, g\, f(x)$$

$$\text{or} \quad f\, \gamma_{\theta}(x) = \gamma'_{\theta'}\, f(x)$$

donc $g\, f\, \gamma(x) = \gamma''_{\theta''}\, gf(x)$ ce qui démontre la propriété.

2. 1. 3 Proposition : Soient $G = \langle X, \theta, \gamma\rangle$ et $G' = \langle X', \theta', \gamma'\rangle$ 2 grillages, S un ensemble fini d'au moins 2 éléments et $f : X \to X'$ une application. f est un Q-morphisme de G dans G' si et seulement si quels que soient le S-schéma Σ et le Σ-grillage $H' = \langle X', \theta', \gamma', \xi', \omega'\rangle$ il existe un Σ-grillage $H = \langle X, \theta, \gamma, \xi, \omega\rangle$ tel que le diagramme suivant commute :

$$\begin{tikzcd} & \Sigma & \\ X \setminus Ter(G) \arrow[ur, "\omega"] \arrow[rr, "f"'] & & X' \setminus Ter(G') \arrow[ul, "\omega'"'] \end{tikzcd}$$

Preuve : Si f est un Q-morphisme, si Σ, ξ' et ω' sont donnés, on définit ξ par :

$$\forall\, x \in X \qquad \xi(x) = \xi'\, f(x)$$

et ω, de manière à ce que le diagramme commute, par :

$\forall\ x \in X \setminus Ter(G) \qquad \omega(x) = \omega'\ f(x).$

Pour tout x, si $\gamma_\theta\ (x) = \langle x_1, \ldots, x_n\rangle$, on a :

$$\gamma'_{\theta'}\ f(x) = \langle f(x_1), \ldots, f(x_n)\rangle$$

Donc $\quad \tau\ \omega'\ f(x) = \langle \xi'\ f(x),\ \xi'\ f(x_1) \ldots \xi'\ f(x_n)\rangle$

$\qquad\quad \tau\ \omega(x) \quad = \langle \xi(x),\ \xi(x_1) \ldots \xi(x_n)\rangle$

H est bien un Σ-grillage (ω n'étant pas définie pour ses sommets terminaux).

Réciproquement, supposons que Σ, G, G' et f soient donnés de manière à ce que le diagramme commute. On a :

$$\forall\ x \in X \setminus Ter(G) \qquad \tau\ \omega'\ f(x) = \tau\ \omega(x)$$

ce qui implique, si $\gamma'_{\theta'}\ f(x) = \langle x'_1, \ldots, x'_n\rangle$ et $\gamma_\theta\ (x) = \langle x_1, \ldots x_n\rangle$:

$$\forall\ x \in X \setminus Ter(G) \quad \langle \xi'\ f(x),\ \xi'(x'_1) \ldots \xi'(x'_n)\rangle = \langle \xi(x),\ \xi(x_1) \ldots \xi(x_n)\rangle$$

Si f n'était pas un Q-morphisme, il existerait x et x_i tels que $x'_i \neq f(x_i)$. Soit $H'' = \langle X', \theta', \gamma', \xi'', \omega''\rangle$ un Σ-grillage tel que $\xi''(x'_i) \neq \xi'(x'_i)$ (un tel Σ-grillage existe puisque S a au moins 2 éléments).

Il ne peut alors exister de Σ-grillage H où l'on aurait $\tau\omega''\ f(x) = \tau\omega(x)$.

2. 2 Sous-Grillages

2. 2. 1 <u>Définition</u> : Soient $G = \langle X, \theta, \gamma\rangle$ et $G' = \langle X', \theta', \gamma'\rangle$ 2 grillages tels que $X' \subseteq X$. G' est un <u>sous-grillage</u> de G si et seulement si θ' est la restriction de θ à X' et

$$\forall\ x \in X' \qquad \gamma'(x) \neq \phi \quad \Rightarrow \quad \gamma'(x) = \gamma(x)$$

2. 2. 2 Proposition : Soient $G = \langle X, \theta, \gamma\rangle$ et $G' = \langle X', \theta', \gamma'\rangle$ 2 grillages tels que $\langle X', \gamma'\rangle$ soit un sous-graphe partiel de $\langle X,\gamma\rangle$ et θ' soit la restriction de θ à X'. L'injection canonique de X' dans X détermine un Q-morphisme de G' dans G si et seulement si G' est un sous-grillage de G.

Preuve : Si G' est un sous-grillage de G, on a immédiatement

$$\forall\ x \in X' \setminus \mathrm{Ter}(G') \qquad \gamma'_{\theta'}(x) = \gamma_{\theta}(x)$$

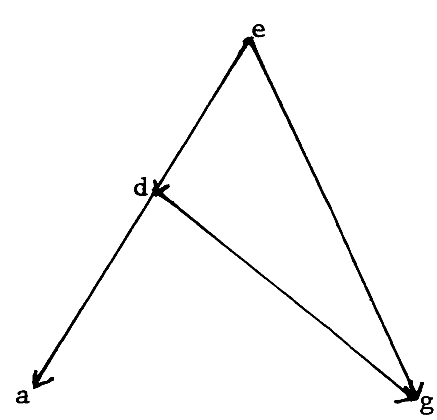
Réciproquement, s'il existe un sommet x non-terminal dans G' tel que $\gamma'(x) \neq \gamma(x)$, l'injection canonique ne peut être un Q-morphisme.

Remarque : Cette proposition, interprétée à la lumière de la proposition 2. 1. 3, indique que les sous-grillages sont les seules sous-structures des grillages qui conservent le rapport aux opérateurs : à chaque sommet non-terminal d'un sous-grillage d'un grillage donné on pourra associer un même opérateur, que l'on se réfère au grillage ou au sous-grillage.

Exemple : Nous reprenons le grillage du premier exemple et nous considérons le grillage $G' = \langle X', \theta', \gamma'\rangle$ tel que

$$X'_{\theta'} = \langle a, d, e, g\rangle$$

x	a	d	e	g
$\gamma'(x)$	ϕ	$\{a,g\}$	$\{d,g\}$	ϕ



Si l'on associe à e et à d les opérateurs σ_2 et σ_3 respectivement, on pourra représenter par ce graphe une expression arithmétique telle que :

$$(x \times y) + y$$

2. 3 Sous-grillages d'ensembles initiaux

2. 3. 1 Lemme : Soit $<X, \gamma>$ un graphe et Y un sous-ensemble de X. L'ensemble $X' = \hat{\gamma}(Y)$ vérifie :

$$\forall\ x \in X' \qquad \gamma(x) \subseteq X'$$

Preuve : Immédiate.

2. 3. 2 Définition : Soient $G = <X, \theta, \gamma>$ un grillage et Y un sous-ensemble de X. $G' = <X', \theta', \gamma'>$ est le sous-grillage d'ensemble initial Y de G si et seulement si :

(i) $X' = \hat{\gamma}(Y)$

(ii) $\forall\ x,y \in X' \qquad x\ \theta' y \Leftrightarrow x\ \theta\ y$

(iii) $\forall\ x \in X' \qquad \gamma'(x) = \gamma(x)$

Il est évident que G' est un sous-grillage de G.

2. 3. 3 Proposition : Soit $G = <X, \theta, \gamma>$ un grillage et $G' = <X', \theta'\ \gamma'>$ un sous-grillage de G. Pour qu'il existe $Y \subseteq X$ tel que G' soit le sous-grillage d'ensemble initial Y, il faut et il suffit que :

$$\forall\ x \in X' \qquad x \in Ter(G') \Rightarrow x \in Ter(G)$$

Preuve : Si G' est un sous-grillage d'ensemble initial, la condition 2.3.2 (ii) entraîne :

$$\gamma'(x) = \phi \quad \Rightarrow \gamma(x) = \phi$$

Réciproquement, si les sommets terminaux du sous-grillage G' de G sont aussi terminaux dans G, cela entraîne :

$$\forall \ x \in X' \qquad \gamma'(x) = \gamma(x)$$

Comme G' est un grillage, on a, pour tout élément x de X'

$$\gamma(x) = \gamma'(x) \subseteq X'$$

d'où on tire $\hat{\gamma}\ (X') \subseteq X'$

et donc $\hat{\gamma}\ (X') = X'$

G'est par conséquent le sous-grillage d'ensemble initial X'.

Remarque : On déduit de cette proposition que si un Σ-grillage sert à représenter une expression arithmétique, un sous-grillage servant à représenter une sous-expression sera nécessairement un sous-grillage d'ensemble initial.

Exemple : Nous reprenons ici encore le grillage du premier exemple et nous en définissons le sous-grillage $G'' = < X'', \theta'', \gamma'' >$ par :

$$X''_{\theta''} = <a, d, b, e, g, c>$$

x	a	b	c	d	e	g
$\gamma''(x)$	ϕ	ϕ	ϕ	$\{a,g\}$	$\{d,g\}$	$\{b,c\}$

Ce sous-grillage peut servir à représenter l'expression :

$$a \times (b - c) + (b - c)$$

2. 3. 4 Proposition : Soit $H = < X, \theta, \gamma, \xi, \omega>$ un Σ-grillage complet, $G' = <X', \theta', \gamma'>$ un sous-grillage de $G = <X, \theta, \gamma>$ et f l'injection canonique de X' dans X. Il existe $H' = < X', \theta', \gamma', \xi', \omega'>$

Σ-grillage complet tel que le diagramme suivant commute

$$\begin{tikzcd} & \Sigma & \\ X' \arrow[ur, "\omega'"] \arrow[rr, "f"'] & & X \arrow[ul, "\omega"'] \end{tikzcd}$$

si et seulement si G' est un sous-grillage d'ensemble initial de G.

Preuve : G' étant un sous-grillage de G, f est un Q-morphisme et le diagramme ci-dessus commute pour les sommets non terminaux de X' si l'on définit ξ' par

$$\forall\ x \in X' \qquad \xi'(x) = \xi\ (x)$$

et si l'on prend :

$$\forall\ x \in X' \setminus \mathrm{Ter}(G') \qquad \omega'(x) = \omega(x)$$

Si G' est un sous-grillage d'ensemble initial de G, tout x de Ter(G') appartient aussi à Ter (G). On peut donc construire H' en définissant ω' par :

$$\forall\ x \in X' \qquad \omega'(x) = \omega\ (x)$$

et donc faire commuter le diagramme ci-dessus.

Réciproquement, si le diagramme commute, on a, pout tout sommet terminal x de G'

$$\omega'(x) = \omega(x)$$

$\omega(x)$ est un opérateur 0-aire et donc x est aussi terminal dans G. Ce qui prouve que G' est un sous-grillage d'ensemble initial de G.

2. 4 Equations définies sur des Σ-grillages

2. 4. 1 Pour nous, une équation sera donnée par un Σ-grillage complet

$H = < X, \theta, \gamma, \xi, \omega>$, un Σ-grillage $H_1 = < X_1, \theta_1, \gamma_1, \xi_1, \omega_1>$, H et H_1 étant des Σ-grillages constants, un Σ-grillage $H'_1 = <X'_1, \theta'_1, \gamma'_1, \xi'_1, \omega'_1>$ et un Q-morphisme f de $G_1 = <X_1, \theta_1, \gamma_1>$ dans $G'_1 = <X'_1, \theta'_1, \gamma'_1>$ inconnus tels que G'_1 soit un sous-grillage de $G = <X, \theta, \gamma>$ et que le diagramme suivant commute :

$$\begin{tikzcd}
& \Sigma & \\
X_1^\circ \arrow[ur, "\omega_1"] \arrow[r, "f"'] & X'_1 \arrow[u, "\omega'_1"'] \arrow[r, "g"'] & X \arrow[ul, "\omega"']
\end{tikzcd}$$

(g est l'injection canonique de X'_1 dans X).

Résoudre l'équation reviendra à trouver le Σ-grillage H'_1 et le Q-morphisme f tels que le diagramme commute.

H'_1 sera alors une solution de l'équation.

Remarque : La donnée des Σ-grillages constants suffisant à déterminer une équation, nous noterons en général l'équation : $< H_1, H >$.

2. 4. 2 Propriété : Si H'_1 est une solution de l'équation, G'_1 est un sous-grillage d'ensemble initial de G.

Preuve : Immédiate, d'après la commutativité de la partie droite du diagramme.

2. 4. 3 Proposition : Pour que l'équation $<H_1, H>$ admette une solution, il faut et il suffit qu'il existe un Q-morphisme h de G_1 dans G tel que le diagramme suivant commute :

$$\begin{tikzcd}
& \Sigma & \\
X_1^\circ \arrow[ur, "\omega_1"] \arrow[rr, "h"'] & & X \arrow[ul, "\omega"']
\end{tikzcd}$$

Preuve : Si l'équation admet une solution, soit h = gf la composée des Q-morphismes f et g : c'est aussi un Q-morphisme. Et le diagramme ci-dessus commute.

Réciproquement, si h est un Q-morphisme, soit $G'_1 = \langle X'_1, \theta'_1, \gamma'_1 \rangle$ le sous-grillage de G d'ensemble initial $h(X_1)$, soit ω'_1 la restriction de ω à X'_1 et soit g l'injection canonique de X'_1 dans X. D'après la proposition 2.3.4, le diagramme ci-dessous commute :

$$\begin{tikzcd} \Sigma & \\ X'_1 \arrow[u, "\omega'_1"] \arrow[r, "g"'] & X \arrow[ul, "\omega"'] \end{tikzcd}$$

On voit que h décrit un Q-morphisme de G_1 dans G'_1 et que ω_1 qui vérifie

$$\forall x \in X_1^\circ \qquad \omega_1(x) = \omega\, h(x)$$

vérifie aussi :

$$\forall x \in X_1^\circ \qquad \omega_1(x) = \omega'_1\, h(x)$$

On peut donc bien obtenir un diagramme qui commute définissant une solution de l'équation donnée.

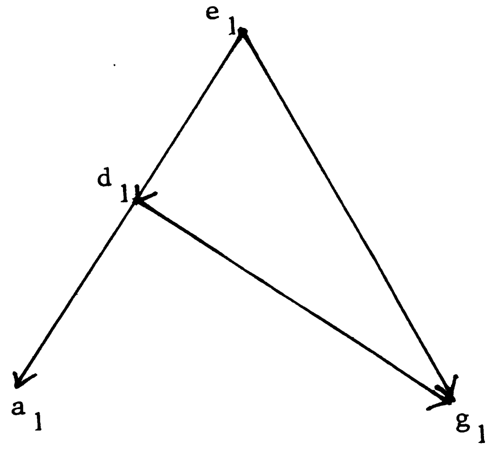
Exemple : Soit H le Σ-grillage de l'exemple 1.3.2. que l'on complète par :

x	a	b	c
$\omega(x)$	$\langle a, \underline{n}, \lambda \rangle$	$\langle b, \underline{n}, \lambda \rangle$	$\langle c, \underline{m}, \lambda \rangle$

et soit $H_1 = \langle X_1, \theta_1, \gamma_1, \xi_1, \omega_1 \rangle$ le Σ-grillage défini par :

$$X_{1\,\theta_1} = \langle a_1, d_1, e_1, g_1 \rangle$$

x	a_1	d_1	e_1	g_1
$\gamma_1(x)$	ϕ	$\{a_1, g_1\}$	$\{d_1, g_1\}$	ϕ
$\xi_1(x)$	$\underline{n}$	$\underline{z}$	$\underline{z}$	$\underline{z}$
$\omega_1(x)$	$<a, \underline{n}, \lambda>$	σ_3	σ_2	non défini



On vérifie que l'application h de X_1 dans X définie par :

x	a_1	d_1	e_1	g_1
$h(x)$	a	d	e	g

détermine un Q-morphisme de G_1 dans G, et on détermine une solution de l'équation, solution construite à partir du sous-grillage d'ensemble initial {a, d, e, g } de G qui n'est autre que le sous-grillage d'ensemble initial de l'exemple 2.3.3.

3 - APPLICATIONS

Dans cette partie, nous nous contentons d'illustrer notre propos, surtout par des exemples présentés intuitivement, sans donner de justifications linguistiques des représentations qui sont choisies.

3. 1 Représentation d'énoncés

Nous prenons $S = \{\underline{n}, \underline{p}\}$ où $\underline{n}$ représente la catégorie des noms ou des syntagmes nominaux et $\underline{p}$ la catégorie des propositions. Ces deux catégories primitives sont les mêmes que celles utilisées, par exemple, dans les grammaires catégorielles de Bar Hillel [BAR, 64] .

Un opérateur simple agit sur une suite finie d'objets de catégories déterminées. La suite de ces catégories forme le type but de l'opérateur, le type source étant la catégorie obtenue par l'action de l'opérateur.

Par exemple, l'opérateur aime fait correspondre à un couple d'éléments de catégorie $\underline{n}$ un élément de catégorie $\underline{p}$. On le note $\langle \text{aime}, \underline{p}, \underline{nn} \rangle$

Les noms élémentaires sont représentés par des opérateurs 0-aires. Par exemple : $\langle \text{Jean}, \underline{n}, \lambda \rangle$. Ainsi obtient-on un S-schéma Σ et un énoncé sera représenté par un Σ-grillage complet.

Exemple : Prenons l'énoncé :

(2) Jean aime Marie et elle le sait. Marie connaît la mère de Jean.

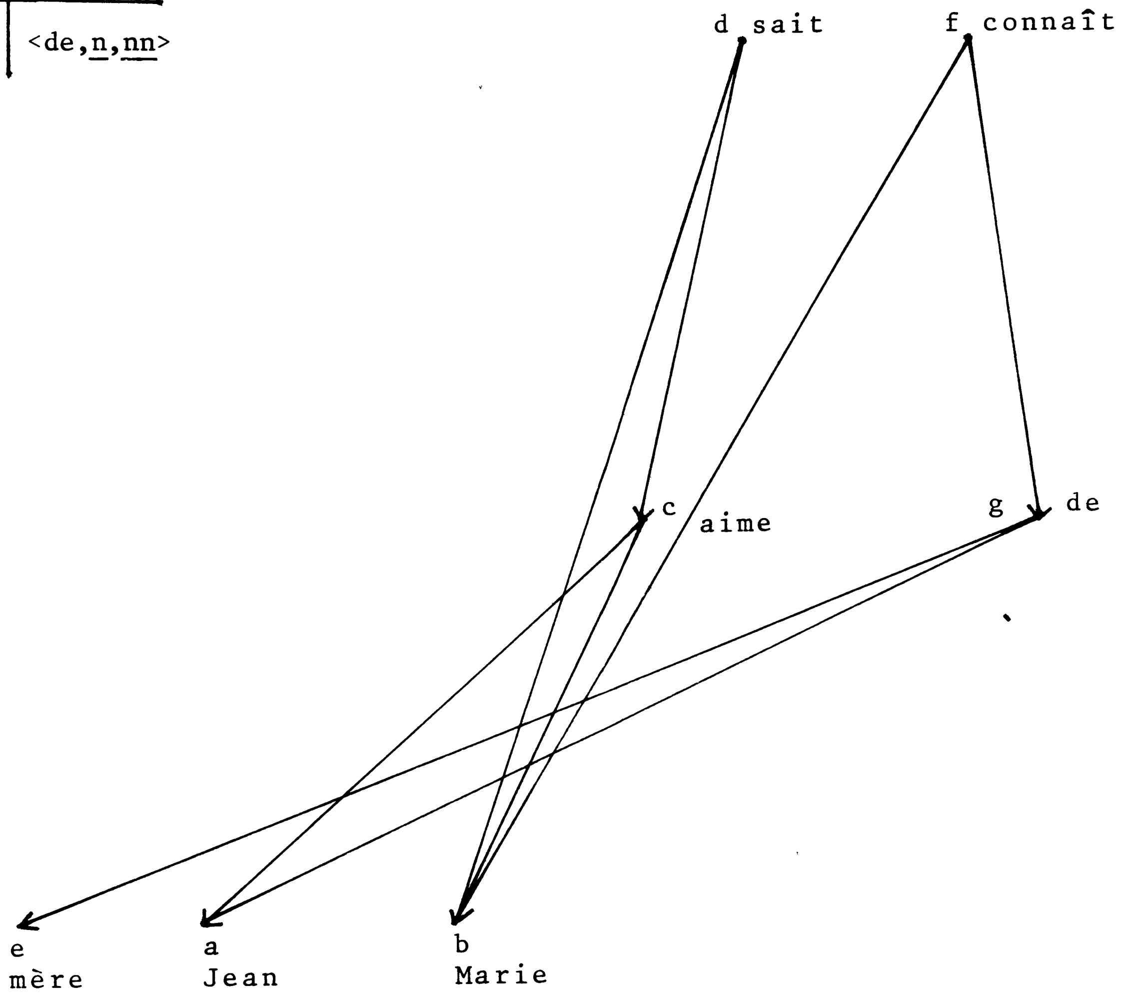
On pourra le représenter par le Σ-grillage

$G = \langle X, \theta, \gamma, \xi, \omega \rangle$ où :

$X_\theta = \langle e, a, b, d, c, f, g \rangle$

x	a	b	c	d	e	f
$\gamma(x)$	ϕ	ϕ	$\{a,b\}$	$\{b,c\}$	ϕ	$\{b,g\}$
$\xi(x)$	$\underline{n}$	$\underline{n}$	$\underline{p}$	$\underline{p}$	$\underline{n}$	$\underline{p}$
$\omega(x)$	$\langle \text{Jean}, \underline{n}, \lambda \rangle$	$\langle \text{Marie}, \underline{n}, \lambda \rangle$	$\langle \text{aime}, \underline{p}, \underline{\underline{nn}} \rangle$	$\langle \text{sait}, \underline{p}, \underline{\underline{np}} \rangle$	$\langle \text{mère}, \underline{n}, \lambda \rangle$	$\langle \text{connaît}, \underline{p}, \underline{\underline{nn}} \rangle$

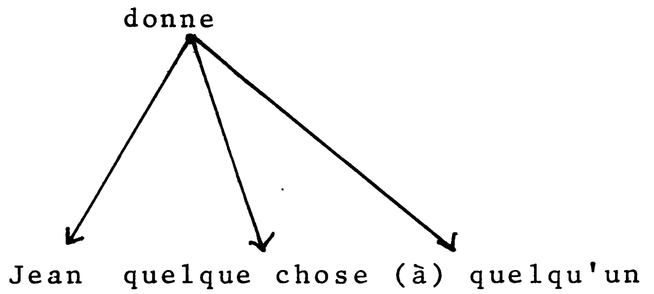
x	g
$\gamma(x)$	$\{e,a\}$
$\xi(x)$	$\underline{n}$
$\omega(x)$	$\langle \text{de}, \underline{n}, \underline{\underline{nn}} \rangle$



Remarques :

a. Nous ne posons pas ici le problème de la recherche du processus de représentation par un Σ-grillage d'un énoncé en langue naturelle. Ce processus requiert l'existence d'un lexique qui mette en correspondance opérateurs et unités lexicales.

Par exemple, à un verbe à 3 actants, tel que *donner* (*Jean donne quelque chose à quelqu'un*), correspondra un opérateur de type <$\underline{p}$, $\underline{n}$ $\underline{n}$ $\underline{n}$> et on aura une représentation telle que :



Cette représentation sera invariable, quel que soit l'ordre des termes dans l'énoncé, même si l'on a, par exemple :

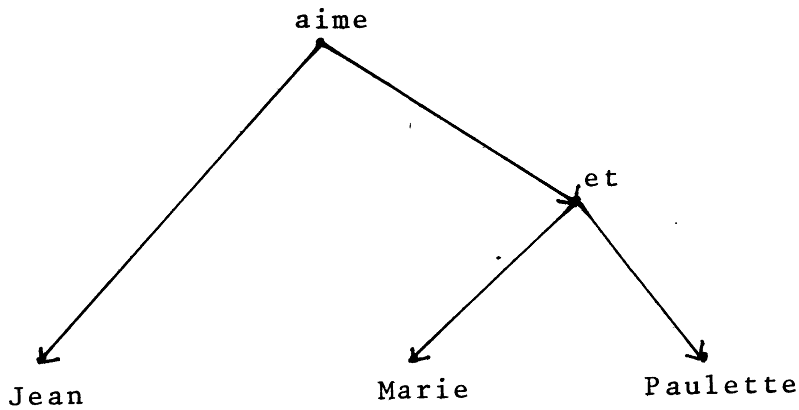
Jean donne à Pierre des nouvelles de Paul.

b. Il est clair que l'on essayera d'obtenir une représentation "standard", qui soit la même pour plusieurs énoncés considérés comme équivalents à certaines approximations près. Le choix de cette représentation se fera en fonction de la facilité potentielle d'interrogation.

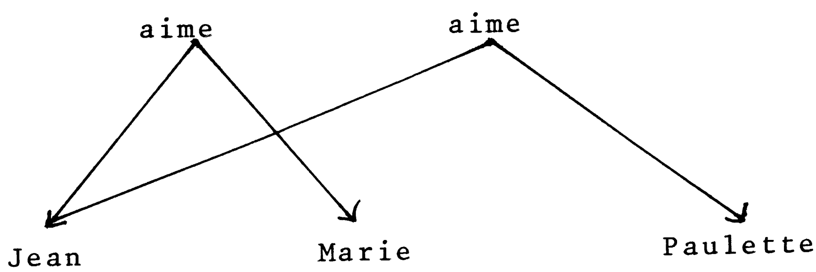
<u>Exemple</u> : L'énoncé

(3) <u>Jean aime Marie et Paulette</u>

ne sera pas représenté par :



mais par :



afin de traiter plus aisément une question comme :

<u>Qui aime Paulette</u> ?

Le passage à la deuxième représentation peut se faire, soit en paraphrasant l'énoncé originel en

(3') <u>Jean aime Marie et Jean aime Paulette</u>

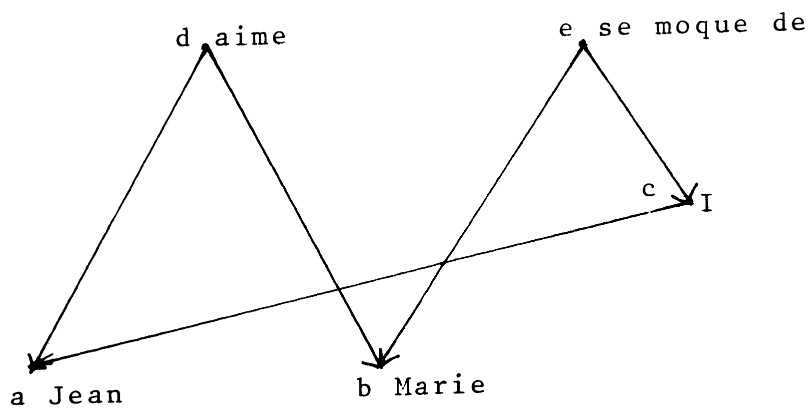
soit directement à partir de la première représentation. On pourra dans ce cas utiliser les transformations décrites en termes de grammaires de graphes dans [COR, 81].

c. Une difficulté provient de la nécessité que les ordres sur les descendants immédiats de mêmes sommets soient compatibles. En effet, si l'on veut par exemple représenter l'énoncé

(4) Jean aime Marie, mais Marie se moque de lui

l'opérateur aime aura Marie comme premier argument et Jean comme deuxième argument, tandis que pour l'opérateur se moque de cela sera l'inverse.

Il faut donc associer Jean (ou Marie) à 2 sommets distincts du graphe. On obtient alors



les sommets a, b, c étant dans cet ordre et l'opérateur associé à c étant I, opérateur identité de type $< \underline{n}, \underline{n} >$. I est un opérateur coprojectif (voir [DES, 1980]).

3. 2 Questions et interrogations

Une question sera représentée par un Σ-grillage G_i (en général non complet).

Une réponse à la question, en fonction du Σ-grillage G qui représente l'énoncé de référence, sera donnée par une solution de l'équation $<G_i, G>$, c'est-à-dire par un sous-grillage G'_i de G et un Q-morphisme f de G_i dans G'_i qui vérifient des propriétés énoncées plus haut.

<u>Exemples</u> :

a. La question

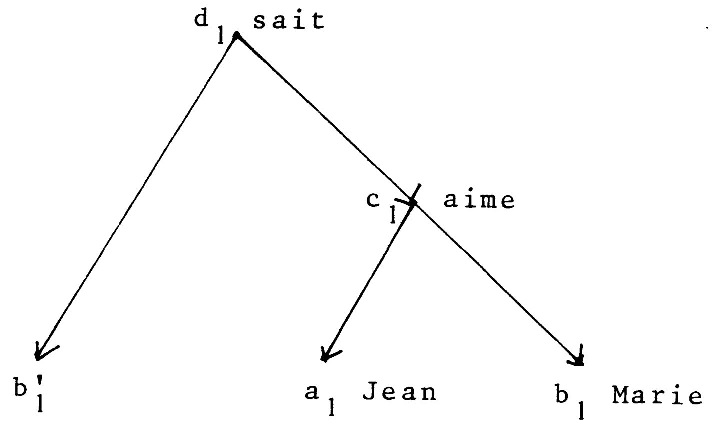
<u>Qui sait que Jean aime Marie</u> ?

sera représentée par le Σ-grillage

$G_1 = <X_1, \theta_1, \gamma_1, \xi_1, \omega_1>$ où :

$X_{1_{\theta 1}} = <b'_1, d_1, a_1, c_1, b_1>$

x	a_1	b_1	b'_1	c_1	d_1
$\gamma_1(x)$	ϕ	ϕ	ϕ	$\{a_1, b_1\}$	$\{b'_1, c_1\}$
$\xi_1(x)$	$\underline{n}$	$\underline{n}$	$\underline{n}$	$\underline{p}$	$\underline{p}$
$\omega_1(x)$	$<\text{Jean}, \underline{n}, \lambda>$	$<\text{Marie}, \underline{n}\ \lambda>$	non défini	$<\text{aime}, \underline{p}, \underline{nn}>$	$<\text{sait}, \underline{p}, \underline{np}>$



L'interrogation de la structure représentative G de l'énoncé (2) doit donner lieu à la construction d'un Q-morphisme f de G_i dans G :

x	a_1	b_1	b'_1	c_1	d_1
f(x)	a	b	b	c	d

La réponse à la question pourra être exprimée par l'énoncé :

Marie sait que Jean l'aime

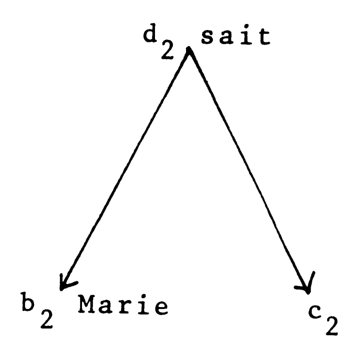
b. La question

Que sait Marie ?

sera représentée par le Σ-grillage $G_2 = \langle X_2, \theta_2, \gamma_2, \xi_2, \omega_2 \rangle$ avec :

$$X_{2\,\theta_2} = \langle b_2, d_2, c_2 \rangle$$

x	b_2	c_2	d_2
$\gamma_2(x)$	ϕ	ϕ	$\{b_2, c_2\}$
$\xi_2(x)$	$\underline{n}$	$\underline{p}$	$\underline{p}$
$\omega_2(x)$	$\langle$Marie,$\underline{n}$,$\lambda\rangle$	non défini	$\langle$sait,$\underline{p}$,$\underline{np}\rangle$



Une solution de l'équation $<G_2,G>$ sera donnée par le sous-grillage de G d'ensemble initial $\{b, c, d\}$.

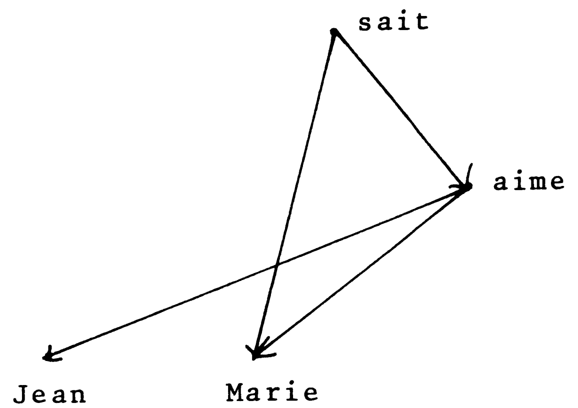
La réponse à cette question est la même qu'à la question précédente.

Remarque : Il est naturel qu'une réponse soit donnée par un sous-grillage d'ensemble initial, car les sous-grillages d'ensembles initiaux sont les seules sous-structures pouvant représenter des sous-énoncés (des syntagmes) d'un énoncé donné.

En effet, il résulte de la proposition 2.3.3 que tout autre sous-grillage conduirait à une structure où on ne connaîtrait pas tous les noms élémentaires sur lesquels agissent les opérateurs.

C. Marie sait-elle que Jean l'aime ?

Cette question sera représentée par un Σ-grillage complet tel que :



La réponse pourra encore une fois être exprimée par :

Marie sait que Jean l'aime.

REMERCIEMENTS :

Je remercie J.P. Desclès pour sa relecture attentive de cet article, ainsi que pour ses nombreux conseils.

BIBLIOGRAPHIE

[AHUL-77] AHO A.V., & ULLMAN J.D., Principles of compiler design, Addison Wesley, 1977.

[BAR-64] BAR-HILLEL Y., "Logical syntax and semantics", Jerusalem academic Press, 1964.

[BER-58] BERGE C., Théorie des graphes et ses applications, Dunod, Paris, 1958.

[CHO-57] CHOMSKY N., Syntactic structures, La Haye, Mouton, 1957.

[HAR-65] HARRIS Z.S., Transformation theory language, n°41, 1965.

[HAY-62] HAYS D.G., "Automatic language-data processing", Computer applications in the behavioral sciences, ed. H. Borko, Prentice Hall, 1962.

[TES-59] TESNIERE L., Eléments de syntaxe structurale, Paris, Klincksieck, 1959.

Les références aux publications du groupe PITFALL sont données p.84.