

Lexique Génératif à la Montague

Bruno Mery

2 Mai 2007

Représentations et formalisations sémantiques

Sous la direction de MM. Bassac, Retoré.

LaBRI, équipes Méthodes Formelles
INRIA Futurs – Projet Signes

Plan

- Linguistique computationnelle et traitement automatique des langues
- Lexiques et sémantiques
- Le lexique génératif et ses formalisations structurelles et logiques
- Notre modèle semi-Montagovien

Linguistique computationnelle

- Méthodes et objet
 1. Années 50 : "calcul" des langues
 2. Tentatives et désillusions
 3. Succès des cas "spécialisés"
 4. Refonte des méthodes
- Niveaux d'analyse
 1. Phonologie
 2. Syntaxe et morphologie
 3. Sémantique
 4. Pragmatique et discours

Traitement automatique des langues

- Aspect pratique et éminemment lucratif
- Bases de données, recherches. . .
- Images et sons
- Représentation des connaissances
- Analyses statistiques pures
- Analyses basées sur la syntaxe
- . . . et sur la sémantique ?

Représentations du lexique

Un ensemble de mots élémentaires. . .

- Membres d'un ensemble "Alphabet" : $L \subseteq A^*$
- Représentants de catégories syntaxiques : $L \subset (N + V + N)^*$
- Porteurs d'un certain sens couplé à la syntaxe

$$\text{dort} : \lambda P \lambda x. P(x) \wedge \text{dort}(x)$$

- Porteurs de plusieurs significations

$$\text{avoir} : \lambda xy. \text{possede}(x, y)$$

$$\text{avoir} : \lambda xy. \text{priopriete}(y, x)$$

- Servant de base à des mécanismes supérieurs

Le principe d'un lexique vecteur de sens

Approches :

- Lexique référentiel
- Lexique à sens multiples

Problèmes :

- Mots finis, sens non bornés
- Polymorphie(s), polysémie(s)
Tente, tu, canon / Ligne, bien, faible

Le sens est-il donc *compositionnel* ?

Exemples compositionnels. . .

- Les adjectifs ont-ils un *sens* ?
Rapide, bon, rouge. . .
- Ambiguïté des verbes
Commencer, vouloir. . .
- Polysémies discursives
Auteur, compositeur, interprète, public. . .
- Figures de style
Voile, lame, bécane. . .

Le Lexique Génératif

- Une base aristotélicienne
- Structures d'arguments et d'événements
- Principes forts de co-composition et de coercion
- Sous-spécification
- Liage sélectif

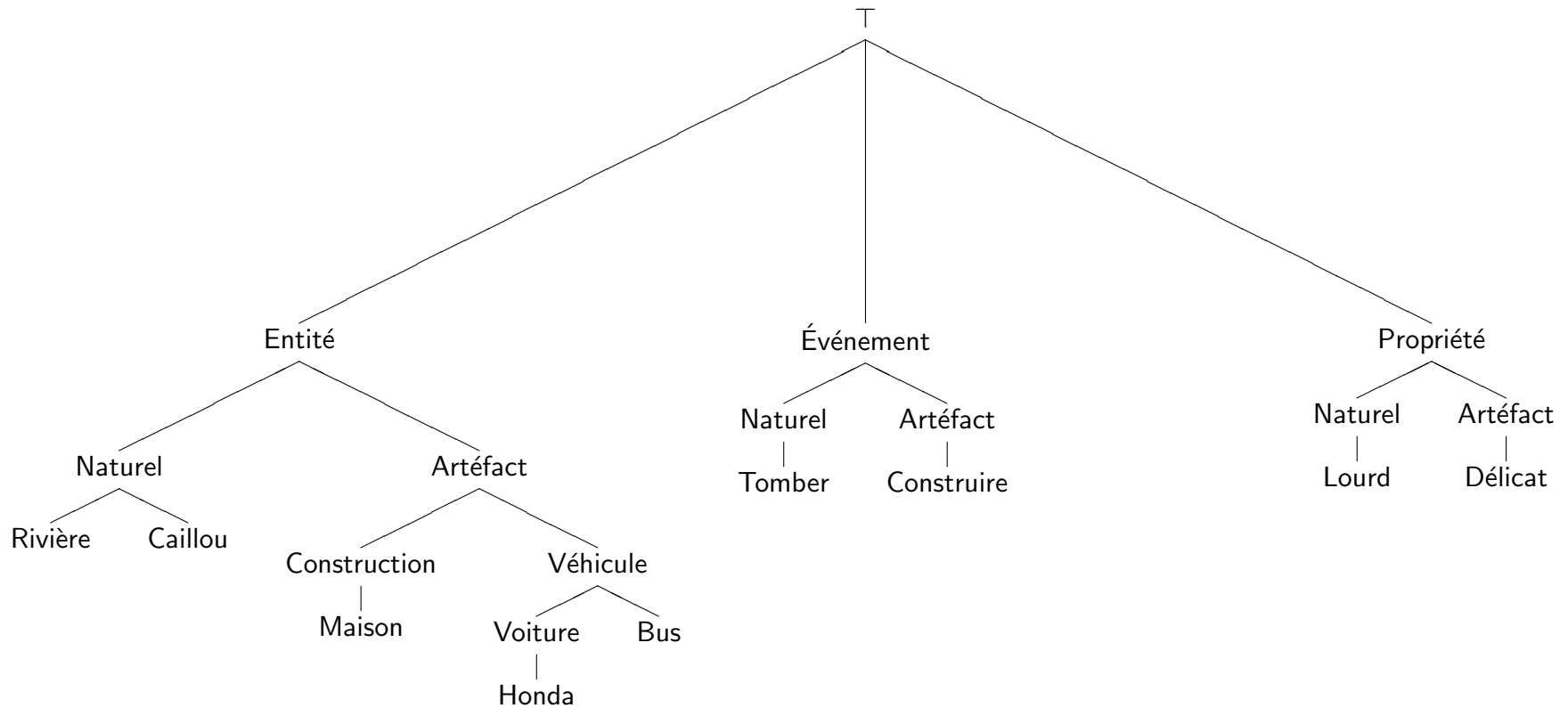
Structure de types riche

Maison	
Argument	: $[x, \text{Maison}(x)]$
Event	: $\left[\begin{array}{l} e_1 : \text{Process} = \text{Construction}(x, y) \\ e_2 : \text{Result} = \text{Achievement}(x) \\ e_1 \leq e_2 \end{array} \right]$
Qualia	: $\left[\begin{array}{l} \text{Formal} : x, \text{Maison}(x) \wedge \text{Spacieuse}(x) \dots \\ \text{Const} : \varepsilon_1, \text{Briques}(\varepsilon_1), \dots \\ \text{Agentive} : y, \text{Ouvriers}(y), \dots \\ \text{Telic} : \exists z : \text{People}/\text{VivreDans}(x, z) \end{array} \right]$
Inheritance	: Construction

Typage et hiérarchie

- Unité de type
- Entités, événements, propriétés
- Types naturels ou artificiels
- Hiérarchisation et héritage
- Spécification croissante
- Accommodation de types

Essai de taxonomie



Un formalisme objet

- Types et classes
- Structures et champs
- Exploitation et méthodes
- Coercition et *casting*
- La représentation de Kalyan Moy Gupta

Extraits de structure (nominaux)

```
EMPLOYEE
term:<employee>
behaviors:<
  beh:<EMPLOY_ACT(EMPLOYER, this)>>
creative events:<
  creat.:<HIRE_ACT(EMPLOYER, PERSON)>>
relatives:<
  rel.:<EMPLOYER>>
*role players:<
  player:<PERSON>>
```

Formules logiques (1)

- Une dérivation du sens fondée sur la syntaxe : prédicats et arguments
- Sémantique vériconditionnelle, par Montague

$$\overbrace{\lambda x^e . \underbrace{\text{dort}}_{e \rightarrow t}(x)}^t [c^e]$$

$$\wedge^{t \rightarrow t \rightarrow t}, \Rightarrow^{t \rightarrow t \rightarrow t}, \forall, \exists \dots$$

Formules logiques (2)

- Types différenciés : $x^e, \text{chat}(x) \rightarrow x^{\text{chat}}$
- Prédicats typés : $\lambda x^{\text{ouvrier}} y^{\text{construction}} . \text{construire}(x, y)$
- Accommodation par héritage

$$\frac{\lambda x^{\text{vehicule}} . \text{conduire}(x) \quad [y^{\text{Honda}}], \quad \text{Honda} \sqsubseteq \text{Voiture} \sqsubseteq \text{Vehicule}}{\lambda x^{\text{vehicule}} . \text{conduire}(x) \quad [y^{\text{vehicule}}]}$$

- Coercitions de types

$$\frac{\lambda x^{\text{personne}} . \text{penser}(x) \quad [y^{\text{article}}], \quad \overbrace{\text{Agentive}(\text{article})}^{\text{personne}}}{\lambda x^{\text{personne}} . \text{penser}(x) \quad [y^{\text{personne}}]}$$

Le cas des *dot objects*

- Quand un type ne suffit pas. . . Entités multi-aspectuelles
- Cas de "livre" : *J'ai lu tous les livres de cette étagère*
Aspect physique, aspect informationnel, relation contenant-contenu. . .
- Cas de "déjeuner" : *Ce déjeuner était tardif, mais très bon*
Aspect événementiel, aspect physique, structure complexe
- Etc. . .
Réunion, cours, concert, fenêtre, verre de vin. . .

Formalismes actuels (1)

- Types “Pointés” : $livre \subseteq P \bullet I$
- Coercition simple : $lire^{I \rightarrow t}$, $etagere^{P \rightarrow t}$
- Problème : perte de l’un des sens en co-prédication
- *Type Pumping*

$$\frac{\lambda x^P . etagere(x) \quad [y^{livre}]}{\lambda x^P . etagere(x) \quad [y^{\pi_2(livre)}]}$$

- Problème : différence des aspects
- Considérations philosophiques sur ce qu’est une composante de types. . .

Formalismes actuels (2)

- Asher : *logique de types compositionnelle avec types complexes*
- Principe : à chaque prédication sur un type complexe, on introduit une variable correspondant à l'aspect désiré.
- Exemple de règles : ●-Exploitation, ●-Introduction en coordination

$$\frac{\{\lambda P\phi(P(x)), c(P : (\alpha \bullet \beta)\gamma)\}[\psi, c'(\psi : \begin{bmatrix} \alpha' \\ \beta' \end{bmatrix} \gamma)], head(\psi)}{\{\lambda P\phi \left[\frac{\exists v(\Delta(\phi, x) \left[\frac{v}{x} \right] \wedge O - Elab(x, v))}{\Delta(\phi, x)} \right], c*(x : \begin{bmatrix} \alpha \Pi \alpha' \\ \beta \Pi \beta' \end{bmatrix}, v : \alpha \bullet \beta)\}[\psi, c']}$$

$$\frac{\{\lambda z\lambda x\lambda y(\phi[z][x][y] \wedge \psi[z][x][y]), c(\phi : \alpha \rightarrow \gamma, \psi : \beta \rightarrow \gamma, x : \delta, y, z : \theta)\}, \diamond x : \alpha \bullet \beta}{\lambda z\lambda x\lambda y(\exists v(\phi[z][v][y] \wedge O - Elab(v, x)) \wedge \exists u(\psi[z][u][y] \wedge O - Elab(u, x))), c*(x : \alpha \bullet \beta, v : \alpha, u : \beta)}$$

Principe de Réalité. . .

- Le formalisme n'est pas satisfaisant pour un logicien
- Problème de principe : multiplication des règles due au transfert de variable
- Problème d'efficacité : $\Delta(\phi, x)$, appui sur une inférence supérieure
- Problème de mise en œuvre. . .
- Plus simple : produit / projection

$$\frac{\lambda x^P . \textit{etagere}(x) \quad [y^{\textit{livre}}]}{\lambda x^P . \textit{etagere}(x) \quad \overbrace{[\pi_P(y)]}^P}$$

Retour à Montague

- Les objections “philosophiques” persistent.
- Plus important : la relation entre l’objet *livre* et son contenu est sensiblement différente de celle entre l’événement “déjeuner” et ce qui y est servi.
- Proposition : placer la polysémie au niveau des *termes* plutôt que de concentrer l’information dans leurs *types*. . .
- Termes optionnels : *changement de type*, relation explicite

<i>item</i>	<i>Main λ-term</i>	<i>Optional terms</i>
<i>livre</i>	$x : LivrePhys$	$g_1 : LivrePhys \rightarrow LivreInfo ;$ $f_T : LivrePhys \rightarrow Event ; \dots$
<i>livre</i>	$y : LivreInfo$	$g_2 : LivreInfo \rightarrow Entity ;$ $f_A : LivreInfo \rightarrow Agent ; \dots$

Un mécanisme intégré

- λ -calcul simple à la Montague, sans règles supplémentaires
- Chaque entrée lexicale peut, en sus de la structure argumentale donnant le λ -terme classique, apporter des termes supplémentaires qui peuvent (ou non) être utilisés. . .
- . . . soit par *auto-adaptation* : qualia, héritage, opérations destructrices ;
- . . . soit par *projection après sélection*, pour le cas des types complexes.
- Les opérateurs permettent un changement de type et explicitent la relation présente
- Les opérateurs peuvent aussi permettre de changer les termes : en général, $x \neq f(x)$

Exemples de fonctionnement

- Auto-adaptation (e.g., cuisson f_1 Animal $A \rightarrow$ Nourriture F) :

Le saumon était délicieux

$\lambda x^F . \text{delicieux}(x)[y^A]$

$\lambda x^F . \text{delicieux}(x)[f_1(y)]$

- Sélection puis projection (e.g., type *LivrePhys* / *LivreInfo*) :

J'ai lu ce livre, puis je l'ai rapporté

$x : \text{LivrePhys}, (\lambda y^\top . \text{lire}(\Pi_{\text{Info}}(y)))[x] \wedge (\lambda z^\top . \text{rapporter}(\Pi_{\text{Phys}}(z)))[x]$

$x : \text{LivrePhys}, \text{lire}(g_1(x)) \wedge \text{rapporter}(x)$

Au delà du Lexique

- En extension : interprétations multiples, choix et score

Bordeaux souhaite un pont, mais le maire s'y oppose

- Une intégration possible de la pragmatique et du contexte discursif ?

Je suis dans ce village depuis deux jours, ma chambre est très spacieuse

/

Je suis dans ce village depuis deux jours, ma yourte est très spacieuse