

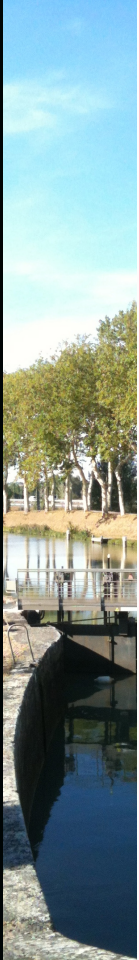
**Syntaxe, sémantique et lexique
dans un même cadre logique**

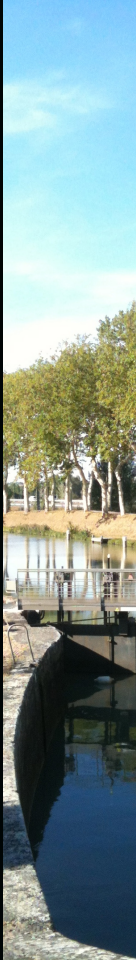
Richard Moot and Christian Retoré
CNRS et Université de Bordeaux

LIRMM, 3 juillet 2013

PLAN

- A Thématique : analyse automatique du langage naturel par des méthodes logiques**
- B D'une phrase à une formule logique, sur un exemple**
- C Quelques questions actuelles**
- D Applications : logiciel d'analyse sémantique et étude de la cognition**





A Thématique : analyse automatique du langage naturel par des méthodes logiques

A.1. Analyse syntaxique et sémantique profonde

Texte / parole retranscrite \longrightarrow formule logique

Quelqu'un repeint une classe.

$\exists x \exists c \text{ repeint}(x, c)$

ou plus fin :

$\exists x : \text{humain} \exists c : \text{gr_humain} \text{ repeint}(x, (f(c))^{\phi})$

avec $\text{repeint} : \text{humain} \rightarrow \phi \rightarrow \mathbf{t}$

A.2. Méthodes logiques en syntaxe

Règles universelles

Lexique : mot \rightarrow catégorie syntaxique

Dérivation grammaticale = preuve formelle
(dans une logique linéaire (ll) et non commutative (nc))

- (1) * Tout le monde connaît Mathieu Mathieu. (ll)
- (2) Mathieu monde tout connaît le. (nc)

A.3. Méthodes logiques en sémantique

Catégorie syntaxique \longrightarrow type sémantique

groupe nominaux \longrightarrow individus

groupes verbaux \longrightarrow prédicats

Analyse syntaxique \longrightarrow assemblage de formules incomplètes (= λ -termes, preuves formelles)

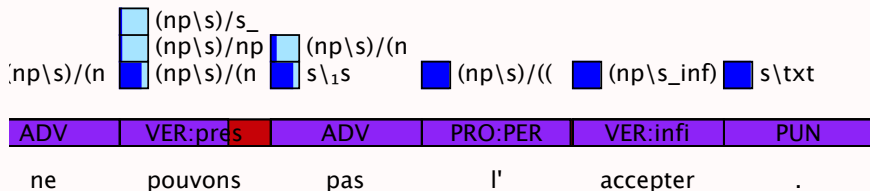
Lexique : mot \longrightarrow formules incomplètes = lambda termes

assemblage, réduction
(en pratique : λ -DRT ; dans cet exposé : λ -calcul à la Montague)

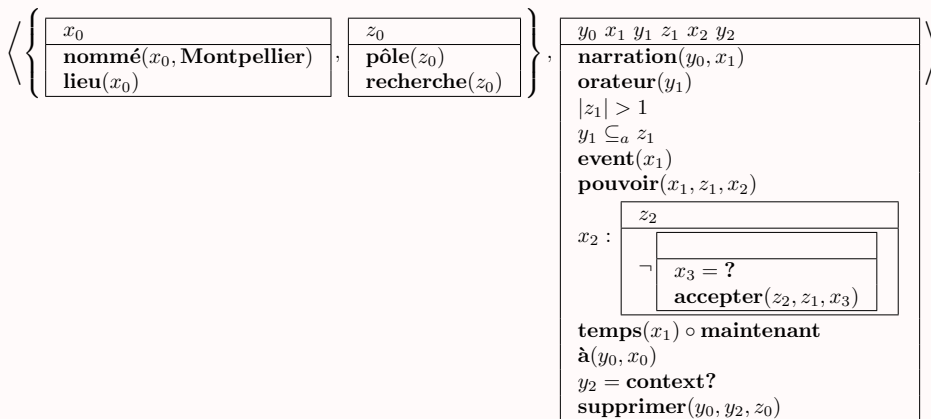
Sémantique lexicale : système multisorte, richement typé

A.4. Exemple — Google News de ce matin

1. Le pôle recherche est supprimé à Montpellier, nous ne pouvons pas l'accepter.



A.5. Exemple — Google News de ce matin



$$\equiv \exists x_0 \text{ nommé}(x_0, \text{Montpellier}) \wedge \exists z_0 \dots$$

A.6. Ressources nécessaires

Grammaires lexicalisées règles universelles
→ tout est dans le LEXIQUE

grammaire apprise sur corpus annoté
types sémantiques simples possiblement acquis
termes sémantiques ?
sens lexical et possibles adaptations au contexte ?





**B D'une phrase à une formule
logique, sur un exemple**

B.1. Syntaxe : Calcul de Lambek

$$\frac{A/B \quad B}{A} /E \qquad \frac{B \quad B \setminus A}{A} \setminus E$$
$$\dots \quad [B]^i \qquad [B]^i \quad \dots$$
$$\vdots \qquad \vdots$$
$$\frac{A}{A/B} /I_i \qquad \frac{A}{B \setminus A} \setminus I_i$$

B.2. Exemple

Quelqu'un repoint la classe
 $s/(np \setminus s)$ $(np \setminus s)/np$ $((s/np) \setminus s)/n$ n

B.3. Exemple

Quelqu'un repeint la classe

$$\frac{s/(np \setminus s) \quad (np \setminus s)/np \quad ((s/np) \setminus s)/n \quad n}{(s/np) \setminus s} / E$$

B.4. Exemple

Quelqu'un repeint

$$s/(np \setminus s) \quad (np \setminus s)/np \quad np \quad \frac{\text{la classe } ((s/np) \setminus s)/n \quad n}{(s/np) \setminus s} / E$$

B.5. Exemple

Quelqu'un repeint la classe
 $s/(np \setminus s)$ $\frac{(np \setminus s)/np}{np \setminus s} / E$ $\frac{((s/np) \setminus s)/n}{(s/np) \setminus s} / E$

B.6. Exemple

Quelqu'un repeint la classe
 $s/(np \setminus s)$ $\frac{(np \setminus s)/np}{np \setminus s} / E$ $\frac{((s/np) \setminus s)/n}{(s/np) \setminus s} / E$

B.7. Exemple

$$\frac{\frac{s/(np \setminus s)}{s/np} / l_1}{s} \frac{\frac{(np \setminus s)/np}{np \setminus s} \frac{[np]^1}{/E}}{\frac{((s/np) \setminus s)/n}{(s/np) \setminus s} \frac{n}{/E}} / E$$

B.8. Syntaxe et sémantique

(Type syntactique)* = Type sémantique		
s^*	=	t un phrase est une proposition
np^*	=	e un groupe nominal/pronom est un entité
n^*	=	$e \rightarrow t$ un nom est un ensemble d'entités
$(a \setminus b)^*$	=	$(b/a)^*$ = $a^* \rightarrow b^*$ extension de $(_)*$ à tout les formules

B.9. Dérivation syntaxique

$$\frac{\frac{s/(np \setminus s)}{s/np} / I_1 \quad \frac{\frac{(np \setminus s)/np \quad [np]^1}{np \setminus s} / E \quad \frac{((s/np) \setminus s)/n \quad n}{(s/np) \setminus s} / E}{s} / E$$

B.10. Dérivation sémantique

$$\frac{\frac{(e \rightarrow t) \rightarrow t \quad \frac{e \rightarrow (e \rightarrow t) \quad [e]^1}{e \rightarrow t} \rightarrow E}{\frac{t}{e \rightarrow t} \rightarrow I_1} \rightarrow E \quad \frac{(e \rightarrow t) \rightarrow ((e \rightarrow t) \rightarrow t) \quad e \rightarrow t}{(e \rightarrow t) \rightarrow t} \rightarrow E}{t} \rightarrow E$$

B.11. Correspondance Curry-Howard

$$\frac{q^{(e \rightarrow t) \rightarrow t} \quad \frac{r^{e \rightarrow (e \rightarrow t)} \quad [x^e]^1}{(r \ x)^{e \rightarrow t}} \rightarrow E}{(q \ (r \ x))^t \rightarrow E} \rightarrow E \quad \frac{l^{(e \rightarrow t) \rightarrow ((e \rightarrow t) \rightarrow t)} \quad c^{e \rightarrow t}}{(l \ c)^{(e \rightarrow t) \rightarrow t}} \rightarrow E$$

$$\frac{\frac{(q \ (r \ x))^t \rightarrow E}{(\lambda x. (q \ (r \ x)))^{e \rightarrow t}} \rightarrow I_1 \quad \frac{l^{(e \rightarrow t) \rightarrow ((e \rightarrow t) \rightarrow t)} \quad c^{e \rightarrow t}}{(l \ c)^{(e \rightarrow t) \rightarrow t}} \rightarrow E}{((l \ c) (\lambda x. (q \ (r \ x))))^t} \rightarrow E$$

B.12. Lexique sémantique

Mot	type sémantique u^* terme sémantique : λ - terme de type u^* x^v la variable ou la constante x est de type v
<i>Quelqu'un</i>	$(e \rightarrow t) \rightarrow t$ $\lambda P^{e \rightarrow t} \exists (e \rightarrow t) \rightarrow t (\lambda x^e (P x))$
<i>repeint</i>	$e \rightarrow (e \rightarrow t)$ $\lambda y^e \lambda x^e ((\text{repeindre}^{e \rightarrow (e \rightarrow t)} x) y)$
<i>la</i>	$(e \rightarrow t) \rightarrow ((e \rightarrow t) \rightarrow t)$ $\lambda P^{e \rightarrow t} \lambda Q^{e \rightarrow t} (\exists (e \rightarrow t) \rightarrow t (\lambda x^e (\wedge^{t \rightarrow (t \rightarrow t)} (P x)(Q x))))$
<i>classe</i>	$e \rightarrow t$ $\lambda x^e (\text{classe}^{e \rightarrow t} x)$

B.13. Lexique sémantique simplifié

Mot	type sémantique u^* terme sémantique : λ -terme de type u^* x^v la variable ou la constante x est de type v
<i>Quelqu'un</i>	$(e \rightarrow t) \rightarrow t$ $\lambda P^{e \rightarrow t} \exists x^e (P x)$
<i>repeint</i>	$e \rightarrow (e \rightarrow t)$ $\lambda y^e \lambda x^e \text{repeindre}(x, y)$
<i>la</i>	$(e \rightarrow t) \rightarrow ((e \rightarrow t) \rightarrow t)$ $\lambda P^{e \rightarrow t} \lambda Q^{e \rightarrow t} \exists x^e (P x) \wedge (Q x)$
<i>classe</i>	$e \rightarrow t$ $\lambda x^e \text{classe}(x)$

B.14. Substitution

$$(I\ c) \equiv (\lambda P^{e \rightarrow t} \lambda Q^{e \rightarrow t} \exists x^e (P\ x) \wedge (Q\ x)) \lambda y^e \text{classe}(y)$$

$$\lambda Q^{e \rightarrow t} \exists x^e ((\lambda y^e \text{classe}(y))\ x) \wedge (Q\ x)$$

$$\lambda Q^{e \rightarrow t} \exists x^e \text{classe}(x) \wedge (Q\ x)$$

B.15. Substitution

$$(q (r x)) \equiv (\lambda P^{e \rightarrow t} \exists y^e (P y)) ((\lambda z^e \lambda v^e \text{repeindre}(v, z)) x)$$

$$(\lambda P^{e \rightarrow t} \exists y^e (P y)) (\lambda v^e \text{repeindre}(v, x))$$

$$\exists y^e ((\lambda v^e \text{repeindre}(v, x)) y)$$

$$\exists y^e \text{repeindre}(y, x)$$

B.16. Substitution

$$\begin{aligned} & ((I\ c)(\lambda x.(q\ (r\ x)))) \\ & \equiv (\lambda Q^{e \rightarrow t} \exists x^e \text{classe}(x) \wedge (Q\ x)) \\ & \quad (\lambda y.\exists z^e \text{repeindre}(z, y)) \end{aligned}$$

$$\exists x^e \text{classe}(x) \wedge ((\lambda y.\exists z^e \text{repeindre}(z, y))\ x)$$

$$\exists x^e \text{classe}(x) \wedge \exists z^e \text{repeindre}(z, x)$$

B.17. Intégration de la sémantique lexicale

Typage riche et strict (restriction de sélection)

Quantifications sur les types pour factoriser les opérations

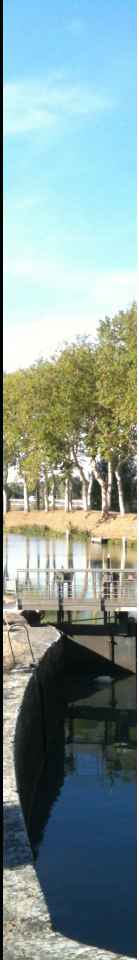
Le lexique associe aux mots des termes optionnels :
changement de type.

Gestion des coprédications (im)possibles.

Utilisés en cas d'argument de type différent de celui attendu.

- (3) C'est une bonne classe.
- (4) Un ouvrier a repeint une classe.
- (5) Un ouvrier a repeint une bonne classe.





C Quelques questions actuelles

C.1. Déterminants, quantificateurs et génériques

Très fréquents. Essentiels à la structure logique de la phrase.

Générique (un, le), quantificateur (un), sélectionneur (le), ... ambigus

- (6) **Un** parent d'élève de maternelle VIENT CHERCHER SON ENFANT en état d'ébriété, l'enseignant commet-il une faute en remettant l'enfant à ce parent ?
- (7) Il y avait **une panthère sortie de la cage**. **Elle** était attachée. **L'animal** a sauté sur moi.
- (8) Recueilli très jeune par les moines de l'abbaye de Reichenau, sur **l'île du lac de Constance**, en Allemagne, qui le prennent en charge totalement, Hermann étudie et devient l'un des savants les plus érudits du XI^{ème} siècle.

C.2. Modèles classiques, epsilon

Montague : quantificateur généralisé = fonction de deux prédicats P et Q
 $\exists x P(x) \& Q(x)$ ou $\forall x P(x) \Rightarrow Q(x)$.

- (9) différence syntaxe / sémantique
- elle écoutait une chanson de lassana hawa
 - SYNT. USUELLE :
(elle (écoutait (**une** (chanson (de lassana hawa))))))
 - SEM. & CG USUELLE :
(**(une** (chanson (de lassana hawa))) (λx elle écoutait x))
- (10) Le GN seul a une référence.
- Un luth, une mandore, une viole, que Michel-Ange [...].
- (11) Asymétrie entre thème et rhème :
- Certains politiciens sont des menteurs car ce qui les intéresse...
 - * Certains menteurs sont des politiciens car ce qui les intéresse...

Le epsilon de Hilbert répond à ces objections

$F(\varepsilon_x.F(x)) \equiv \exists x.F(x)$ $F(\varepsilon_x.\neg F(x)) \equiv \forall x.F(x)$

Un chat dort (sous ta voiture) $dort(\varepsilon_x \text{ chat}(x))$

C.3. Questions logiques

Epsilon fascinant !

Beaucoup de résultat faux (Asser, Leisenring,.. pas Hilbert !)
des travaux récents (Mints, Baaz)

Logique pour la syntaxe

GP (Blache, Guénot, Prost) 1er ordre, contraintes sur des termes

CG (Lambek) propositionnel, linéaire,

Lien ?

Lien théorie des types et calcul des prédicats :

$t : \alpha$ (irréfutable) $\tilde{\alpha}(t)$ (réfutable)

quel est le domaine des prédicats ?

quel est le domaines des opérateurs de Hilbert :

types ou prédicats/ensembles

C.4. Degrés et préférences

- (12) il a été habiter à côté de chez Rosalie que Rosalie elle savait pas
- (13) a. Il a quitté le labo il y a cinq minutes.
b. Il a quitté le labo il y a cinq ans.
- (14) Le lion a tendance à s'échapper.

Les GP peuvent quantifier la force des contraintes, pas les CG.

Comment utiliser JeuxDeMots ou de la sémantique distributionnelle pour préférer telle ou telle lecture.



**D Applications :
logiciel d'analyse sémantique
et étude de la cognition**

D.1. Analyseur sémantique du Français

Grail grammaires acquise

analyse catégorielle (MMCG)

analyse sémantique (lambda DRT plutôt que lambda calcul)

analyse sémantique avec sémantique lexicale
(quand on a le lexique : acquisition ?)

D.2. Corpus de récits de voyages du XIXe dans les Pyrénées

Reconstruction d'itinéraires.

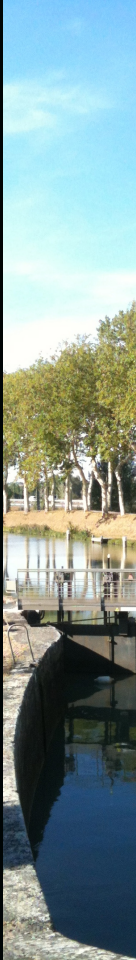
Typage riche, lexique sémantique fin pour

Temps

Espace

Coercions lexicales (fictive motion).

- (15) Nous coupons ici un sentier qui vient du port de Barroude (...)
- (16) La route suit le gave qui vient de Gavarnie.
- (17) Plus loin, de nobles hêtres montent sur le versant(...)
- (18) (...) cette route qui monte sans cesse pendant deux lieues
- (19) Le chemin pavé de calcaire et de pierres luisantes (...) serpente à travers fourrés de buis et de noisetiers
- (20) Puis, cinq minutes nous conduisent à un petit pont (...) qui nous porte sur la rive droite.



D.3. Modèles de la cognition humaine

PEPS CoLan Complexité et langage : une étude formelle et expérimentale des mécanismes de compréhension

Divers paramètres des modèles formels sont pressentis pour avoir une influence sur la difficulté à comprendre.

Que comprend-on ? Quelles mesures formelles s'accordent avec les expériences ?

Tests, suivi oculaire, temps de réaction (EEG)

- (21) Les Français vont en vacances à la mer "et ce sera encore le cas très longtemps" confirme J.-M. Mocquet,
- (22) a. Une secrétaire présentera chaque étudiant à un professeur.
b. À ce qu'il paraît, Marie conduit habituellement une Twingo
c. Un représentant de chaque labo connaît un conseiller de chaque commune.
- (23) J'ai fini mon livre. (lire, écrire, imprimer, autre ?)



Des questions ?