

Analyse sémantique (automatique) du langage naturel

Pascal Amsili¹ Pascal Denis²

¹Laboratoire de Linguistique Formelle
Université Paris Diderot - Paris 7

²Laboratoire ALPAGE
INRIA & Université Paris Diderot - Paris 7

Master de Linguistique Informatique, année 2009-2010

- 1 Introduction
 - Analyse sémantique
 - Organisation du cours
- 2 Approches compositionnelles
 - Principes
- 3 Construction compositionnelle de représentations
 - Rappels
 - Grammaire de Montague
 - Quantification et portée

Plan

- 1 Introduction
 - Analyse sémantique
 - Organisation du cours
- 2 Approches compositionnelles
 - Principes
- 3 Construction compositionnelle de représentations
 - Rappels
 - Grammaire de Montague
 - Quantification et portée

Analyse sémantique (du LN)

- Définition linguistique : relation des signes à la « réalité »
- Définition en TAL : ??? quel TAL n'est pas sémantique ?

Définition empirique : **niveau de représentation**

Tâches :

- Compréhension automatique : extraction d'information
- Calcul de référence, relations de discours...

Ressources :

- lexique, réseau sémantique
- corpus étiquetés de référence

cf. cours Analyse syntaxique (du LN)

Contenu du cours

- ① Approches compositionnelles
 - ① Construction compositionnelle de représentation logique
 - ② Evaluation des conditions de vérité
- ② Sémantique lexicale
 - ① Désambiguation des sens en contexte (WSD) (avec Wordnet)
 - ② Ressource : Thésaurus distributionnel
- ③ Discours et (co-)Référence
 - ① Co-référence (définition et mesures pour évaluation)
 - ② Résolution anaphorique

Organisation pratique

- Responsables : Pascal Amsili (6 séances)
Pascal Denis (6 séances)
- Alternance CM / TP
- Contrôle des connaissances : travaux rendus à l'issue de la séance ou quelques temps après (détails à déterminer)
- 12 séances sur 13 semaines : pas de cours le 3 nov.

Page web :

<http://www.linguist.univ-paris-diderot.fr/~amsili/Ens/AnaSem.php>

Plan

- 1 Introduction
 - Analyse sémantique
 - Organisation du cours
- 2 Approches compositionnelles
 - Principes
- 3 Construction compositionnelle de représentations
 - Rappels
 - Grammaire de Montague
 - Quantification et portée

Approches compositionnelles

Sémantique vériconditionnelle

Comprendre une phrase c'est être capable d'énoncer ses conditions de vérité

- (1) A cat is on the mat

Approches compositionnelles

Sémantique vériconditionnelle

Comprendre une phrase c'est être capable d'énoncer ses conditions de vérité

(1) A cat is on the mat



Approches compositionnelles

Sémantique vériconditionnelle

Comprendre une phrase c'est être capable d'énoncer ses conditions de vérité

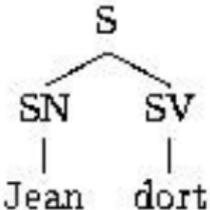
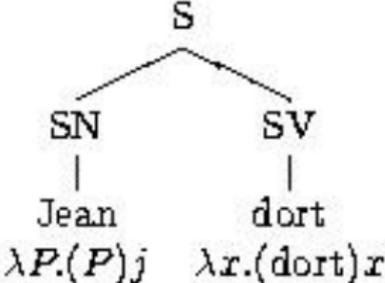
(1) A cat is on the mat



Approches compositionnelles

Principe de compositionnalité du sens

Le sens d'une phrase est fonction du sens de ses parties

| | | |
|---|-------------|--|
| Jean dort | → | (dort) <i>j</i> |
| yacc ↓ | | λ-calcul |
|  | attributs → |  |

Plan

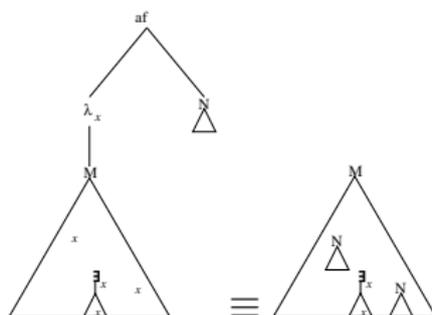
- 1 Introduction
 - Analyse sémantique
 - Organisation du cours
- 2 Approches compositionnelles
 - Principes
- 3 Construction compositionnelle de représentations
 - Rappels
 - Grammaire de Montague
 - Quantification et portée

λ -calcul

$$(\lambda x. \alpha x \gamma) A \xrightarrow{\beta} \alpha A \gamma$$

$$(\lambda x. M) N \equiv [x := N] M$$

$$(\lambda x. \boxed{\text{ronfle}(x)}) \boxed{\text{Jean}} \equiv \boxed{\text{ronfle}(\boxed{\text{Jean}})}$$



Types

- e est un type
- t est un type
- Si a et b sont des types, alors $\langle a, b \rangle$ est un type

D_e = domaine des entités

D_t = domaine des valeurs de vérité

$D_{\langle a, b \rangle}$ = domaine des fonctions de D_a vers D_b

| | | |
|---|----------------------|--------------------------------------|
| e | entités | noms propres |
| $\langle e, t \rangle$ | prédicats | noms, adjectifs, verbes intransitifs |
| $\langle e, \langle e, t \rangle \rangle$ | prédicat binaires | verbes transitifs |
| $\langle \langle e, t \rangle, t \rangle$ | quantif. généralisés | NP |

Langage typé

- Pour chaque type a , les variables et les constantes de type a sont dans ME_a .
- Pour tous types a et b , si $\alpha \in ME_{\langle a,b \rangle}$ et $\beta \in ME_a$ alors $(\alpha)\beta \in ME_b$.
- Pour tous types a et b , si u est une variable de type a et $\alpha \in ME_b$, alors $\lambda u.\alpha$ est dans $ME_{\langle a,b \rangle}$.
- Si φ et ψ sont dans ME_t , alors les expressions suivantes sont aussi dans ME_t : $\neg\varphi$, $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$.
- Pour tout type a , si φ est dans ME_t et u est une variable de type a , alors $\forall u\varphi$ et $\exists u\varphi$ sont dans ME_t .

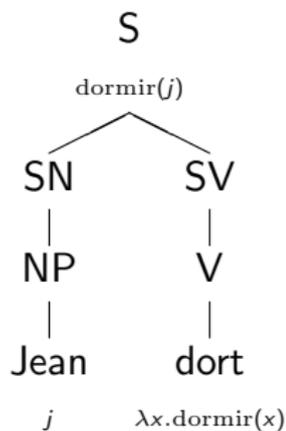
Exemples

 j $\lambda x.dort(x)$ $dort(Jean)$ $\lambda P.P$ $\lambda P\lambda Q.\exists t((P)t \wedge (Q)t)$ $\lambda Q.\exists x(homme(x) \wedge (Q)x)$ $\lambda P\lambda x.(P)\lambda y.((aime)x)y$ $\lambda x.\lambda y.aimer(x)(y)$ e $\langle e, t \rangle$ t $\langle a, a \rangle$ $\langle \langle e, t \rangle, \langle \langle e, t \rangle, t \rangle \rangle$ $\langle \langle e, t \rangle, t \rangle$ $\langle \langle \langle e, t \rangle, t \rangle, \langle e, t \rangle \rangle$ $\langle e, \langle e, t \rangle \rangle$

Méthode des fragments

- 1 définir une analyse syntaxique (une portion de grammaire)
- 2 définir un mode de composition des représentations (application fonctionnelle)
- 3 définir une représentation pour chaque lexème participant à la construction

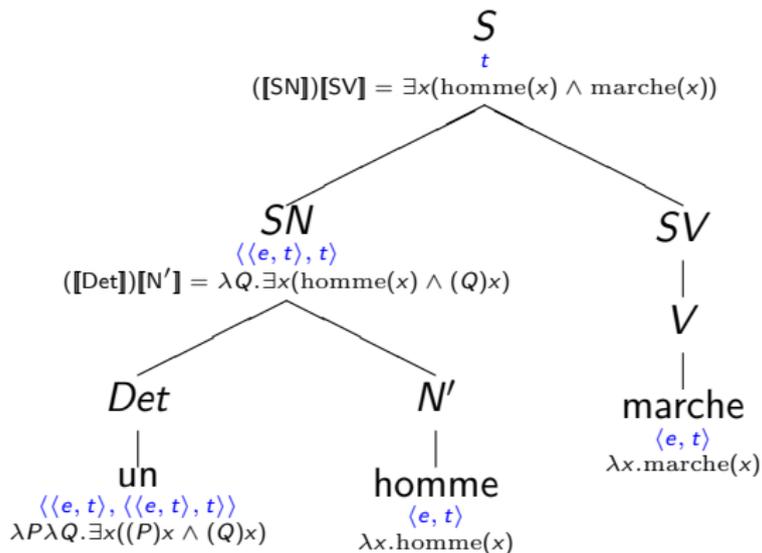
Exemple



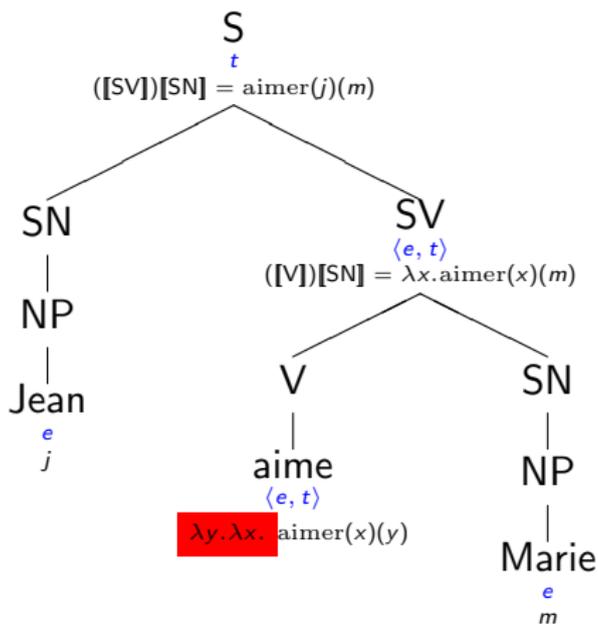
| | | | |
|------|---------------|-------|------|
| S | \rightarrow | SN | SV |
| 0 | \leftarrow | (2) | 1 |
| SN | \rightarrow | NP | |
| 0 | \leftarrow | 1 | |
| SV | \rightarrow | V | |
| 0 | \leftarrow | 1 | |

| | | |
|------|---------------|-----------------------|
| NP | \rightarrow | Jean |
| 0 | \leftarrow | j |
| V | \rightarrow | dort |
| 0 | \leftarrow | $\lambda x.dormir(x)$ |

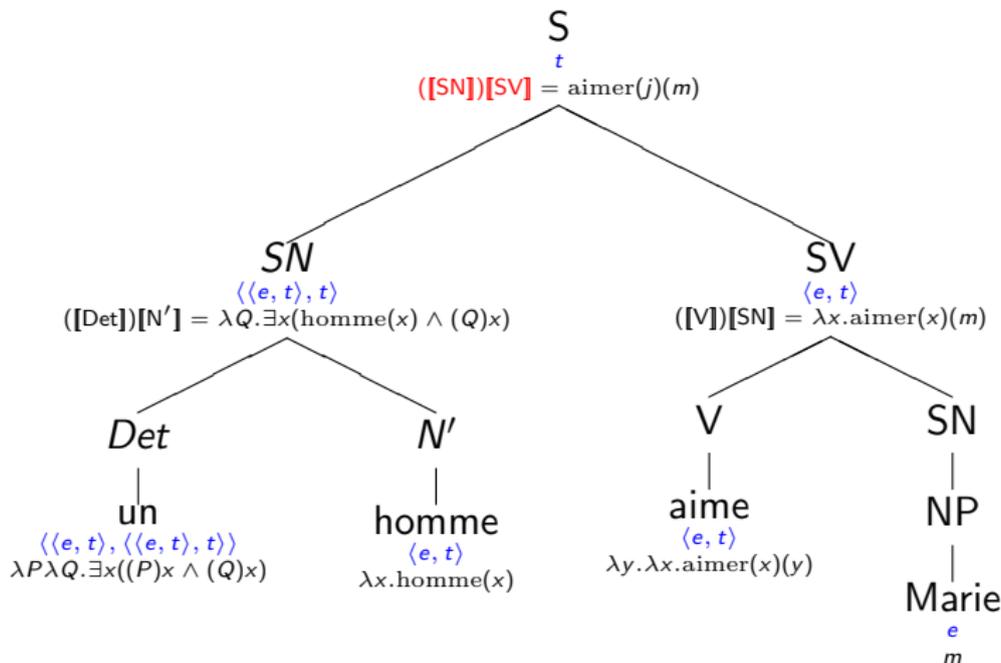
Quantificateurs



Verbes transitifs

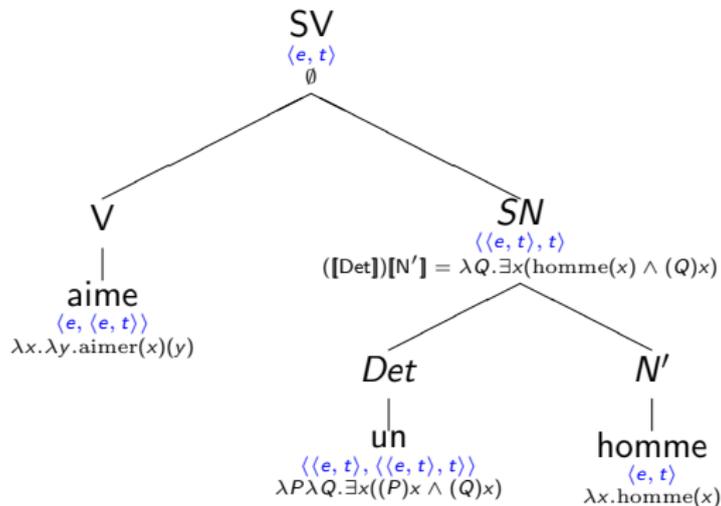
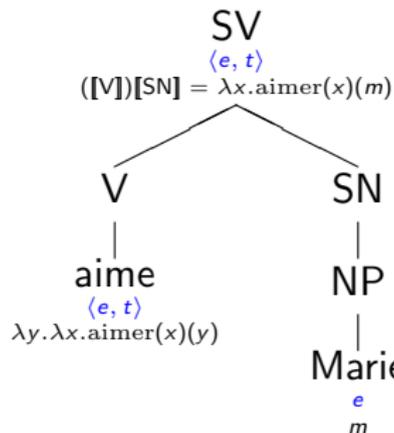


Verbes transitifs (suite)



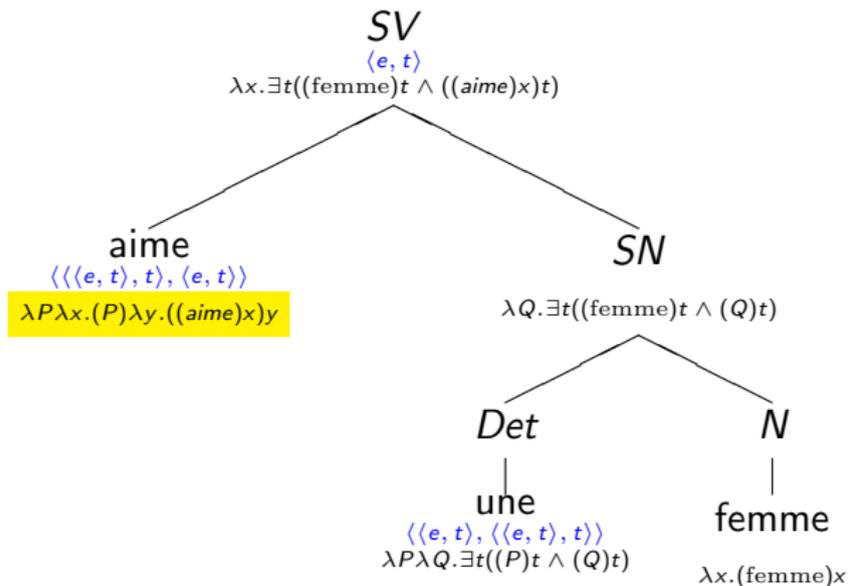
Verbes transitifs

Problème



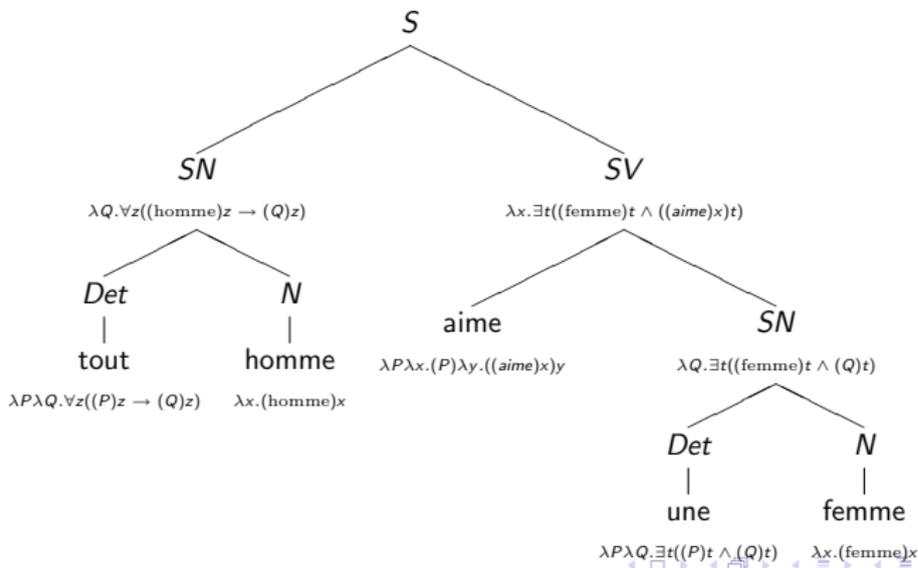
Verbes transitifs

Solution



Fragment : récapitulatif

- (2)
- Tout homme aime une femme
 - $\forall x \exists y ((\text{homme}(x) \wedge \text{femme}(y)) \rightarrow \text{aime}(x, y))$
 - $\forall z ((\text{homme})z \rightarrow \exists t ((\text{femme})t \wedge ((\text{aime})z)t))$



Type shifting : 1er exemple

- Hypothèse 1 :

$$\begin{aligned}V &\rightarrow \text{aime} \\ \llbracket V \rrbracket &= \lambda P \lambda x. (P) \lambda y. ((\text{aime})x)y\end{aligned}$$

- Hypothèse 2 :

$$\begin{aligned}V &\rightarrow \text{aime} \\ \llbracket V \rrbracket &= \lambda x. \lambda y. ((\text{aime})x)y \\ V' &\rightarrow V \\ \llbracket V' \rrbracket &= (\Phi) \llbracket V \rrbracket\end{aligned}$$

$$\Phi = ?$$

Type shifting : 1er exemple (suite)

$$\Phi = ?$$

$$(\Phi)\lambda x.\lambda y.((\textit{aime})x)y = \lambda P\lambda x.(P)\lambda y.((\textit{aime})x)y$$

$$\Phi = \lambda Q.\lambda P.\lambda x_1.(P)(Q)x_1$$

Type de Φ ?

Type shifting : 1er exemple (suite)

$$\Phi = ?$$

$$(\Phi)\lambda x.\lambda y.((\textit{aime})x)y = \lambda P\lambda x.(P)\lambda y.((\textit{aime})x)y$$

$$\Phi = \lambda Q.\lambda P.\lambda x_1.(P)(Q)x_1$$

Type de Φ ?

$$\langle\langle e, \langle e, t \rangle \rangle, \langle\langle\langle e, t \rangle, t \rangle, \langle e, t \rangle \rangle\rangle$$

Type shifting : 1er exemple (suite)

$$\Phi = ?$$

$$(\Phi)\lambda x.\lambda y.((\text{aime})x)y = \lambda P\lambda x.(P)\lambda y.((\text{aime})x)y$$

$$\Phi = \lambda Q.\lambda P.\lambda x_1.(P)(Q)x_1$$

Type de Φ ?

$$\langle\langle e, \langle e, t \rangle \rangle, \langle\langle\langle e, t \rangle, t \rangle, \langle e, t \rangle \rangle\rangle$$

- Vérification :

$$(\lambda Q.\lambda P.\lambda x_1.(P)(Q)x_1)$$

Type shifting : 1er exemple (suite)

$$\Phi = ?$$

$$(\Phi)\lambda x.\lambda y.((\text{aime})x)y = \lambda P\lambda x.(P)\lambda y.((\text{aime})x)y$$

$$\Phi = \lambda Q.\lambda P.\lambda x_1.(P)(Q)x_1$$

Type de Φ ?

$$\langle\langle e, \langle e, t \rangle \rangle, \langle\langle\langle e, t \rangle, t \rangle, \langle e, t \rangle \rangle\rangle$$

- Vérification :

$$(\lambda Q.\lambda P.\lambda x_1.(P)(Q)x_1) \lambda x.\lambda y.((\text{aime})x)y$$

Type shifting : 1er exemple (suite)

$$\Phi = ?$$

$$(\Phi)\lambda x.\lambda y.((\textit{aime})x)y = \lambda P\lambda x.(P)\lambda y.((\textit{aime})x)y$$

$$\Phi = \lambda Q.\lambda P.\lambda x_1.(P)(Q)x_1$$

Type de Φ ?

$$\langle\langle e, \langle e, t \rangle \rangle, \langle\langle\langle e, t \rangle, t \rangle, \langle e, t \rangle \rangle\rangle$$

- Vérification :

$$\begin{aligned} & (\lambda Q.\lambda P.\lambda x_1.(P)(Q)x_1) \lambda x.\lambda y.((\textit{aime})x)y \\ = & \lambda P.\lambda x_1.(P)(\lambda x.\lambda y.((\textit{aime})x)y)x_1 \end{aligned}$$

Type shifting : 1er exemple (suite)

$$\Phi = ?$$

$$(\Phi)\lambda x.\lambda y.((\textit{aime})x)y = \lambda P\lambda x.(P)\lambda y.((\textit{aime})x)y$$

$$\Phi = \lambda Q.\lambda P.\lambda x_1.(P)(Q)x_1$$

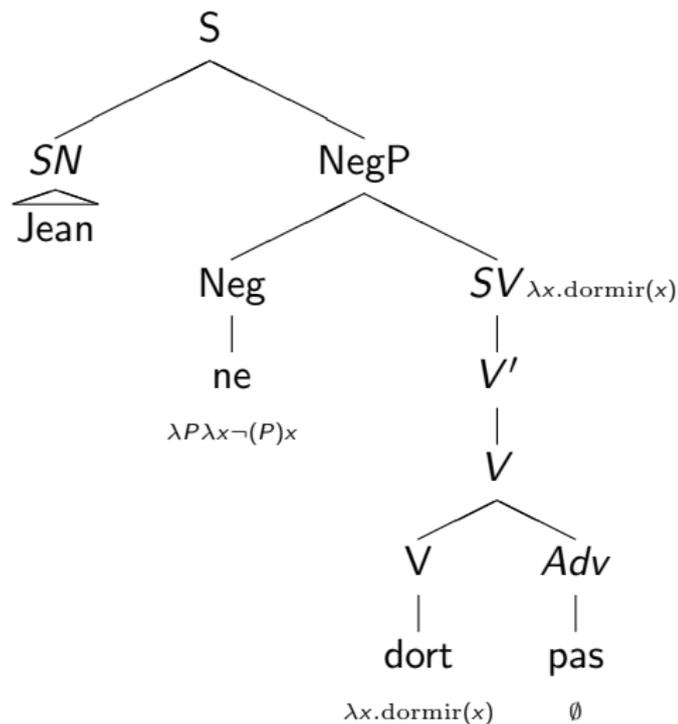
Type de Φ ?

$$\langle\langle e, \langle e, t \rangle \rangle, \langle\langle\langle e, t \rangle, t \rangle, \langle e, t \rangle \rangle\rangle$$

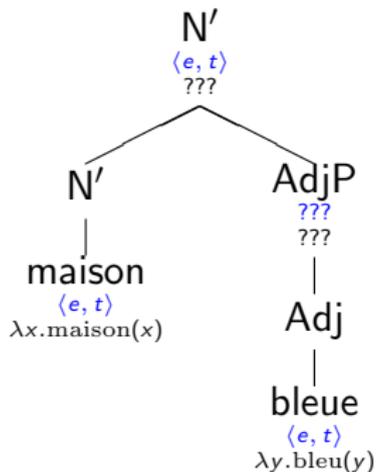
- Vérification :

$$\begin{aligned} & (\lambda Q.\lambda P.\lambda x_1.(P)(Q)x_1) \lambda x.\lambda y.((\textit{aime})x)y \\ = & \lambda P.\lambda x_1.(P)(\lambda x.\lambda y.((\textit{aime})x)y)x_1 \\ = & \lambda P.\lambda x_1.(P)\lambda y.((\textit{aime})x_1)y \end{aligned}$$

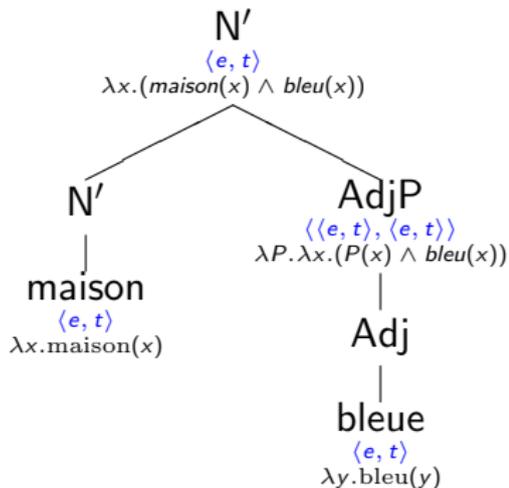
Négation



Modification (adjectivale)



Modification (adjectivale)



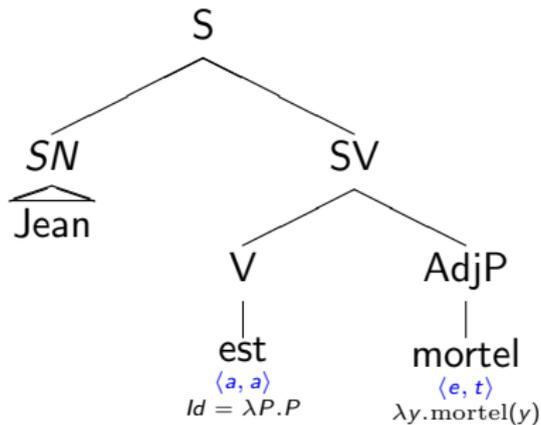
$$AdjP \rightarrow Adj$$

$$[[AdjP]] = \lambda Q \lambda P \lambda x. ((P)x \wedge ([Adj])x)$$

$$N' \rightarrow N' AdjP$$

$$[[N']] = ((\lambda Q \lambda P \lambda x. ((P)x \wedge (Q)x)) [[N']]) [[AdjP]]$$

Adjectif attribut



Interprétation des quantificateurs in situ ?

(3) Tous les acteurs du film aiment une femme.

Interprétation des quantificateurs

in situ ?

- (3) Tous les acteurs du film aiment une femme.
 $\forall x$ Mais ça n'est pas toujours la leur.

Interprétation des quantificateurs

in situ ?

(3) Tous les acteurs du film aiment une femme.

$\exists x$

Elle n'a pourtant pas le premier rôle

Interprétation des quantificateurs in situ ?

(3) Tous les acteurs du film aiment une femme.

Exemple bateau ?

Interprétation des quantificateurs

in situ ?

- (3) Tous les acteurs du film aiment une femme.
- (4)
- a. Tous les étudiants ont lu un article
 - b. Chaque nouveau venu doit passer un test
 - c. Un spécialiste relira chaque papier.
 - d. Un guide accompagnera chaque visiteur.
 - e. Il y a une étiquette à côté de chaque assiette.

Interprétation des quantificateurs

Deux problèmes

Ces situations posent un problème pour notre analyse compositionnelle. Deux types de problème en fait :

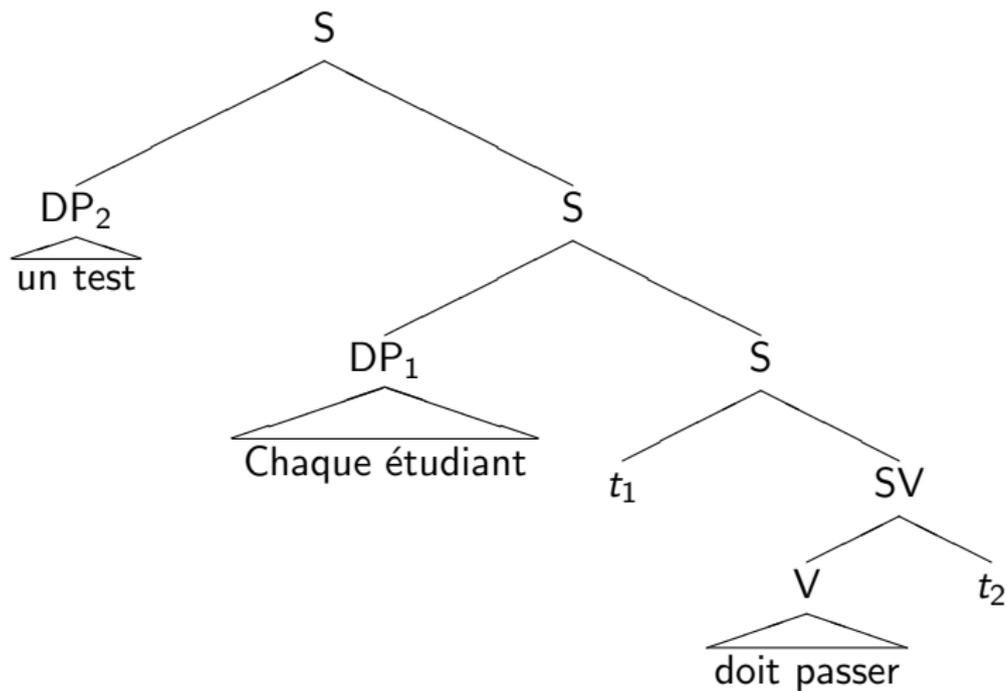
- Aucune ambiguïté n'est prévue dans notre système (ambiguïté sémantique)
- Même si on suppose qu'il n'y a pas ambiguïté, il faut être capable de dériver des formules logiques qui n'obéissent pas aux règles précédentes.

Interprétation des quantificateurs

Réponses possibles

- Mouvement (QR, May, 1987)
- Traitement par types : ambiguïté sémantique
- Traitement sémantique (stockage à *la Cooper*)
- Traitement par enrichissement de la logique : logique combinatoire sans variable (Jacobson), logique *independant-friendly* à la Hintikka...
- Traitement par sous-spécification

Mouvement



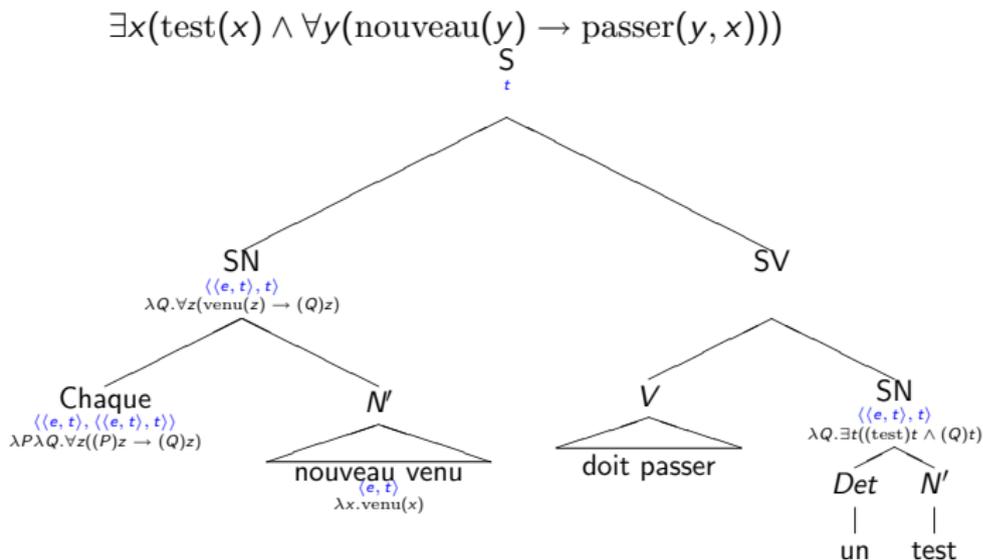
Ambiguïté sémantique

Problème :

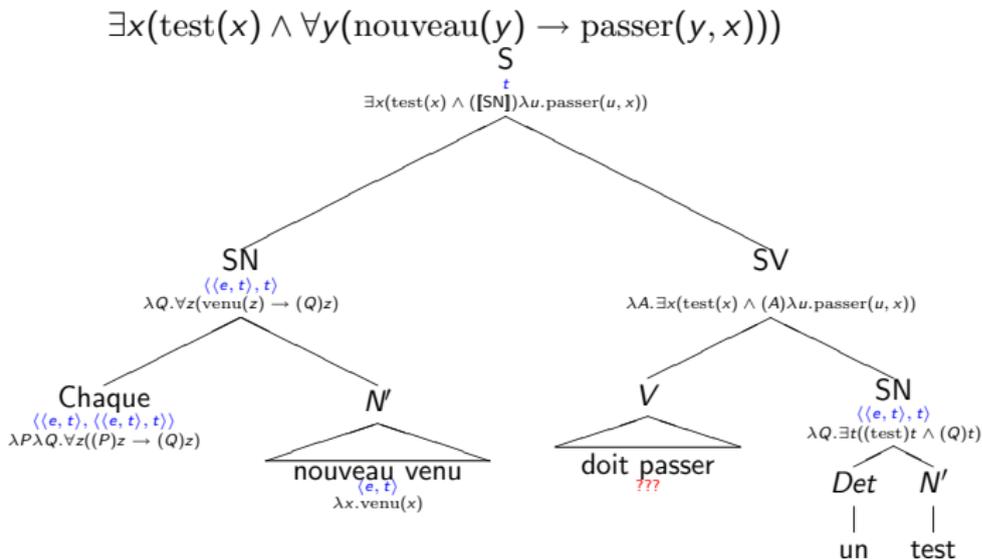
Comment définir la contribution de *un test* pour avoir la bonne lecture ?

- (5)
- a. Chaque nouveau venu doit passer un test
 - b. $\exists x(\text{test}(x) \wedge \forall y(\text{nouveau}(y) \rightarrow \text{passer}(y, x)))$

Ambiguïté sémantique



Ambiguïté sémantique



Cooper storage

- Représentation à deux niveaux
- Opérations supplémentaires : chargement/déchargement
- Ambiguïté rendue par la multiplicité de sites de “déchargement”

Enrichissement de la logique/Sous-spécification

$$\forall x \exists y (\text{homme}(x) \rightarrow (\text{femme}(y) \wedge \text{aime}(x, y)))$$