

Le langage, logique !

Maxime Amblard

15 Octobre 2015

Les jeudis des sciences
Colloquium général
Leçon 217
semestre XXIX

LORIA - UMR 7503
Université de Lorraine

MSH-Lorraine - USR 3261
Inria Nancy Grand Est

CNRS

Plan

- 1 Langage, logique, etc.
- 2 Pour une histoire croisée : linguistique, informatique et logique
- 3 Sémantique(s) des langues naturelles

Plan

- 1 Langage, logique, etc.
- 2 Pour une histoire croisée : linguistique, informatique et logique
- 3 Sémantique(s) des langues naturelles

C'est quoi le langage ?

- capacité d'exprimer une pensée au moyen de signes avec une sémantique. *ie* un moyen de communication
- langage naturel : langue parlée par un être humain

Modéliser la langue = gérer les ambiguïtés

- phonologie
- morphologie

- syntaxe

- sémantique

- pragmatique

Modéliser la langue = gérer les ambiguïtés

Ginette Garcin

- phonologie
- morphologie

- syntaxe

- sémantique

- pragmatique

Modéliser la langue = gérer les ambiguïtés

- phonologie
- morphologie

autossa = dans l'auto
autossaniko = dans mon auto

- syntaxe
- sémantique
- pragmatique

Modéliser la langue = gérer les ambiguïtés

- phonologie
- morphologie

la : pronom, nom ou déterminant ?

- syntaxe
- sémantique
- pragmatique

Modéliser la langue = gérer les ambiguïtés

- phonologie
- morphologie

- syntaxe

Jean regarde (la fille avec un télescope)
Jean regarde (la fille) (avec un télescope)

- sémantique

- pragmatique

Modéliser la langue = gérer les ambiguïtés

- phonologie
- morphologie

- syntaxe

Esther pense que Bertille croit que Hélène dit qu'elle conduira la voiture

- sémantique
- pragmatique

Modéliser la langue = gérer les ambiguïtés

- phonologie
- morphologie

- syntaxe

- sémantique

- pragmatique

La belle ferme la porte
étoile : célébrité ou astre ?

Modéliser la langue = gérer les ambiguïtés

- phonologie
- morphologie

- syntaxe

- sémantique

- pragmatique

La pièce est dans le porte-monnaie
Le porte-monnaie est dans la pièce

Modéliser la langue = gérer les ambiguïtés

- phonologie
- morphologie

- syntaxe

- sémantique

- pragmatique

Pour faire quoi ?

Pour faire quoi ?



Real Humans, saison 2

Où et pour faire quoi ?

- mails, SMS, Twitter, Facebook, commentaires, blogs, articles, pages web, ...

Où et pour faire quoi ?

- mails, SMS, Twitter, Facebook, commentaires, blogs, articles, pages web, ...

- interface homme-machine en langue naturelle
- dialogues automatiques (réservations d'hôtel, Siri)
- génération automatique de textes (résumé, bulletin météo)
- Watson !
- reconnaissance d'écriture, identification vocale

Où et pour faire quoi ?

- synthèse vocale
- traduction automatique, correction automatique
- recherche d'information, moteurs de recherche
- jeux en ligne (syntaxe, sémantique)
- détection d'opinion, de leader, suivi de marques sur internet
- génération de sous-titres
- avatar (LSF)

Langage de la logique

Mais difficulté de la langue naturelle à exprimer clairement les relations logiques

Langage de la logique

Mais difficulté de la langue naturelle à exprimer clairement les relations logiques

- “le produit de la somme de deux nombres par leur différence est identique à la différence des carrés de ces nombres”

Langage de la logique

Mais difficulté de la langue naturelle à exprimer clairement les relations logiques

- “le produit de la somme de deux nombres par leur différence est identique à la différence des carrés de ces nombres”
- $(n + m)(n - m) = n^2 - m^2$

⇒ émergence de langues spécifiques, comme les mathématiques, la langue de la logique : $\forall \exists \diamond \square \rightarrow \wedge \vee$

C'est quoi le langage ?

⇒ langues non naturelles : $a^n b^n c^n$

usage en informatique

- définir des programmes (des outils concrets)
- penser des algorithmes (des outils abstraits de modélisation)

C'est quoi le langage ?

⇒ langues non naturelles : $a^n b^n c^n$

usage en informatique

- définir des programmes (des outils concrets)
- penser des algorithmes (des outils abstraits de modélisation)
qui sont eux-mêmes définis par un langage simple spécifique

C'est quoi le langage ?

⇒ langues non naturelles : $a^n b^n c^n$

usage en informatique

- définir des programmes (des outils concrets)
- penser des algorithmes (des outils abstraits de modélisation)
qui sont eux-mêmes définis par un langage simple spécifique

- si on comprend un langage, on pourrait le définir pour qu'une machine le comprenne (?)

Qu'est-ce que la langue naturelle ?

- difficile de dire précisément si une phrase est grammaticale/non-grammaticale
- pas d'ensemble de règles grammaticales officiel

Qu'est-ce que la langue naturelle ?

- difficile de dire précisément si une phrase est grammaticale/non-grammaticale
 - pas d'ensemble de règles grammaticales officiel
- ⇒ les modèles de la langue ne sont que des approximations cherchant à définir des processus cognitifs
- ⇒ on cherche à définir les propriétés formelles :
- suisse allemand $S_1 S_2 S_3 V_1 V_2 V_3 O_1 O_2 O_3$.

Différence entre un linguiste et un informaticien ?

- linguiste :

- description de la langue naturelle
- intérêt dans la compréhension du processus cognitif
(comment la grammaticalité est caractérisée ?)

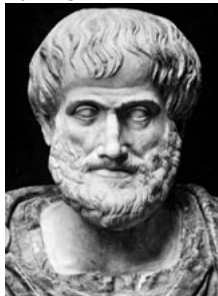
- informaticien :

- simuler une tâche humaine
- réussir une tâche

(le résultat doit pouvoir être réutilisé !)

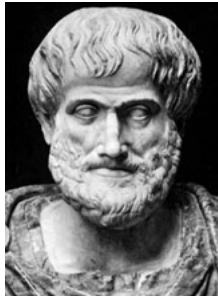
Est-ce qu'il y a de la logique dans la langue ?

- syllogisme d'Aristote



Est-ce qu'il y a de la logique dans la langue ?

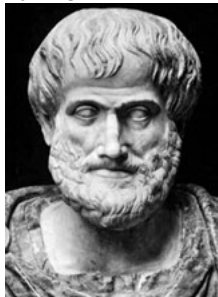
- syllogisme d'Aristote



Tous les hommes sont mortels

Est-ce qu'il y a de la logique dans la langue ?

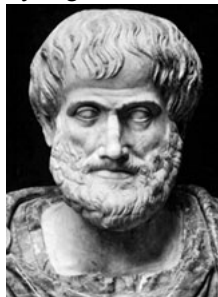
- syllogisme d'Aristote



Tous les hommes sont mortels
Socrate est un homme

Est-ce qu'il y a de la logique dans la langue ?

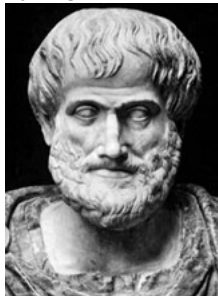
- syllogisme d'Aristote



Tous les hommes sont mortels
Socrate est un homme
⇒ Socrate est mortel

Est-ce qu'il y a de la logique dans la langue ?

- syllogisme d'Aristote

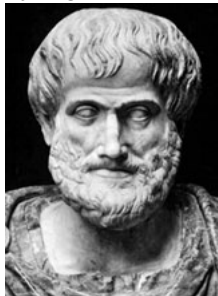


Tous les hommes sont mortels
Socrate est un homme
⇒ Socrate est mortel

Toutes les créatures à dents sont kleptomanes
Les poules ont des dents
⇒ Les poules sont kleptomanes

Est-ce qu'il y a de la logique dans la langue ?

- syllogisme d'Aristote

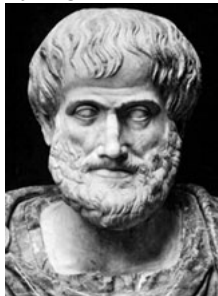


Tous les hommes sont mortels
Socrate est un homme
 \Rightarrow Socrate est mortel

Toutes les créatures à dents sont kleptomanes
Les poules ont des dents
 \Rightarrow Les poules sont kleptomanes
valide mais non vrai !

Est-ce qu'il y a de la logique dans la langue ?

- syllogisme d'Aristote



Tous les hommes sont mortels
 Socrate est un homme
 ⇒ Socrate est mortel

Toutes les créatures à dents sont kleptomanes
 Les poules ont des dents
 ⇒ Les poules sont kleptomanes
 valide mais non vrai !

- ça veut dire quoi sémantique ?

LANGUAGE < - - - - - > WORLD

Plan

1 Langage, logique, etc.

2 **Pour une histoire croisée : linguistique, informatique et logique**

3 Sémantique(s) des langues naturelles

Linguistique (et informatique)

- 500 av JC grammaire du Sanskrit de Panini
- 7^{ème} et 8^{ème} siècle : définition de l'arabe classique
(en parallèle de l'algorithmique)
- 19^{ème} siècle les frères Grimm et la description systématique de la correspondance des sons dans les langues indo-européennes
p-f : père-father ; pied-foot ; ...
- première partie du XX^{ème} : Ferdinand de Saussure
 - diachronie et synchronie
 - signifiant ('arbre') et signifié (le concept d'arbre)
 - Langue (système) et Parole
⇒ Structuralisme

Logique



Leibniz

“Le seul moyen de rectifier notre raisonnement est de le rendre **aussi tangible que celui des mathématiciens**, alors nous pouvons trouver nos erreurs, et quand il y a discussion, nous pouvons simplement dire : allons-y, calculons, sans grande difficulté, pour voir qui a raison.”

L'art de la découverte, 1685

Logique (et linguistique)

fin du 19^{ème} la vue idéale de Leibniz est partiellement réalisée :

Frege : la logique comme calcul, LPO



Peano : axiomatisation de l'arithmétique
(linguiste italien)



Russell : à la suite de Peano, appliquer la logique
aux fondements des mathématiques

Carnap : fonder les connaissances du monde sur la
logique

Adjukiewicz-Bar Hillel : grammaire catégorielle
(structure de la langue, catégories sémantiques)

Informatique (et logique)



Alonzo Church : équivalence entre ce qui est calculable et les fonctions récursives → d'où le λ -calcul. Premier lien fort entre description calculatoire et description mathématique



Turing : Machines de Turing (implémentation des propositions de Church) (puis cryptographie, IA, morphogénèse, ...)



Shannon and Weaver : approche probabiliste pour quantifier la quantité moyenne d'information (codage, texte, compression, cryptographie, ...)

Le tournant de la sémantique

Autour des années 1930, la vision syntaxique est étendue avec le développement de la théorie des modèles et le tournant sémantique



Tarski (logicien polonais) émigré aux états-unis en 1939 : définition célèbre de la **satisfaction** et introduction de la **theorie des modèles**

Une théorie est valide si on peut définir un univers dans laquelle elle est vraie

⇒ introduction de la **notion de vérité**
[tarski1944] [tarski1956]

Le tournant de la sémantique

Donc transformer une phrase S en langage naturelle en une formule ϕ fournit une interprétation du sens de S .



Objectifs

- représenter le sens des énoncés grâce à des formules logiques (proposition, premier ordre, classique, intuitioniste, ...)
- développer des algorithmes pour produire des représentations logiques (et utiliser ces représentations)
- techniques fondamentales pour construire de représentations sémantiques : λ -calcul

Plusieurs difficultés

- pourquoi utiliser des représentations ?
pourquoi ne pas utiliser directement la langue naturelle ?

Plusieurs difficultés

- pourquoi utiliser des représentations ?
pourquoi ne pas utiliser directement la langue naturelle ?
- peuvent-elles être utiles pour la pragmatique ?
prendre en compte que nous utilisons réellement le langage ?

Plusieurs difficultés

- pourquoi utiliser des représentations ?
pourquoi ne pas utiliser directement la langue naturelle ?
- peuvent-elles être utiles pour la pragmatique ?
prendre en compte que nous utilisons réellement le langage ?
- réalité philosophique, cognitive ou conceptuelle ? Jeu de symboles ou définition de ce que nous comprenons ?

Plusieurs difficultés

- pourquoi utiliser des représentations ?
pourquoi ne pas utiliser directement la langue naturelle ?
- peuvent-elles être utiles pour la pragmatique ?
prendre en compte que nous utilisons réellement le langage ?
- réalité philosophique, cognitive ou conceptuelle ? Jeu de symboles ou définition de ce que nous comprenons ?
- est-ce si pratique ? Le raisonnement logique est mathématiquement difficile

Première étape vers la construction sémantique



Principe de **compositionnalité** de **Frege** :

Le sens du tout est une fonction qui dépend du sens des parties.

Première étape vers la construction sémantique



Principe de **compositionnalité** de **Frege** :

Le sens du tout est une fonction qui dépend du sens des parties.

- mot = formule logique de base
- syntaxe exprime comment les composer

Première étape vers la construction sémantique



Principe de **compositionnalité** de **Frege** :

Le sens du tout est une fonction qui dépend du sens des parties.

- mot = formule logique de base
- syntaxe exprime comment les composer

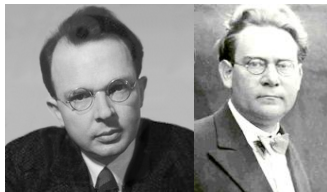


Historiquement c'est une idée assez récente

Long consensus pour penser que le lien était purement **analogique**, **Frege** et **Tarski** étaient très septiques quant à l'usage de la logique pour les langues naturelles

Historiquement c'est une idée assez récente

Long consensus pour penser que le lien était purement **analogique**, **Frege** et **Tarski** étaient très septiques quant à l'usage de la logique pour les langues naturelles

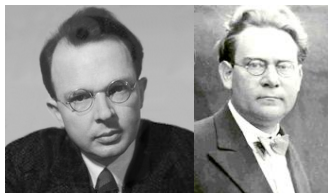


1930-1950 : certains logiciens ont pris ce lien au sérieux :

Carnap (modalités) et
Reichenbach (temps)

Historiquement c'est une idée assez récente

Long consensus pour penser que le lien était purement **analogique**, **Frege** et **Tarski** étaient très septiques quant à l'usage de la logique pour les langues naturelles



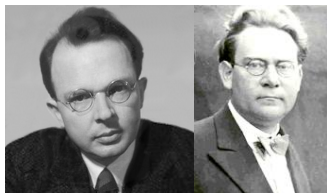
1930-1950 : certains logiciens ont pris ce lien au sérieux :

Carnap (modalités) et
Reichenbach (temps)

Mais **vision abstraite** et **non calculatoire** (algorithmique) de la logique

Historiquement c'est une idée assez récente

Long consensus pour penser que le lien était purement **analogique**, **Frege** et **Tarski** étaient très septiques quant à l'usage de la logique pour les langues naturelles



1930-1950 : certains logiciens ont pris ce lien au sérieux :

Carnap (modalités) et
Reichenbach (temps)

Mais **vision abstraite** et **non calculatoire** (algorithmique) de la logique

50 - 60 : nombreux philosophes se sont positionnés contre une vision partagée de la logique et de la langue naturelle

De Chomski à Montague

Dans ce contexte : apparition des travaux de **Chomsky** et **Montague**

De Chomski à Montague

Dans ce contexte : apparition des travaux de **Chomsky** et **Montague**



En 1957, **Chomski** propose une **formalisation de la syntaxe** des langues naturelles avec un versant cognitif. Pas de sémantique, mais tournant pour la **linguistique computationnelle**, et (via Montague) pour la sémantique computationnelle.

De Chomski à Montague

Dans ce contexte : apparition des travaux de **Chomsky** et **Montague**



En 1957, **Chomski** propose une **formalisation de la syntaxe** des langues naturelles avec un versant cognitif. Pas de sémantique, mais tournant pour la **linguistique computationnelle**, et (via Montague) pour la sémantique computationnelle.

Chomski est influent en linguistique par la théorie générative, mais également en informatique en théorie des langages.

Richard Montague (1930–1971)



Dans trois articles publiés à la fin des années 60, **Montague** ouvre la sémantique des langues naturelles modernes :

- English as a Formal Language
- The Proper Treatment of Quantification in Ordinary English
- Universal Grammar

Le travail de Montague

- logique intentionnelle
- quantificateurs généralisés (most, few, three, ...)
- premier modèle pour l'ambiguïté de portée des quantificateurs
- définition d'une interface syntaxe-sémantique rigoureuse

⇒ Il remplace l'analogie par l'algorithmique

De l'autre côté du diagramme



De l'autre côté du diagramme



Extension des systèmes :

- time, tense and aspect : **Reichenbach**, **van Benthem**, logique du temps à la **Allen**

De l'autre côté du diagramme



Extension des systèmes :

- time, tense and aspect : **Reichenbach**, **van Benthem**, logique du temps à la **Allen**
- événements : réification à la **Davidson**
- pluriel
- modalités
- ...



Limites des approches montageovienne

- donkey sentences

Every farmer who owns a donkey beats it

$$(\exists x \exists y. (\text{farmer } x \wedge \text{donkey } y \wedge \text{own } x y)) \rightarrow \text{beat } x y$$

- résolution d'anaphores inter-phrastiques

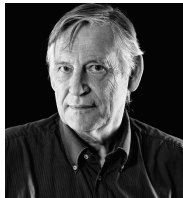
A man walks in the park. He whistle.

$$\exists x. (\text{man } x \wedge \text{walk_in_the_park } x) \wedge (\text{whistle } x)$$

Sémantique dynamique



- **Context Change Potential (CCP)** [Heim1983]
Interpretation en **fonction du contexte** et
contexte **modifié par l'interprétation**



- **Discourse Representation Theory (DRT)**
[Kamp1981]/**File Change Semantics (FCS)**
[Heim1982]
**niveaux intermédiaires entre représentation et
valeurs de vérités**

Type Theoretic Dynamic Logic



- **Dynamic Predicate Logic (DPL)**
[Groenendijk1991]
syntaxe standard mais sémantique
non-standard

Type Theoretic Dynamic Logic



- **Dynamic Predicate Logic (DPL)**
[Groenendijk1991]
syntaxe standard mais sémantique non-standard



- **Type Theoretic Dynamic Logic (TTDL)**
[deGroote2006] : purement montagovien pour le discours avec dynamicité par les continuations dans le λ -calcul

Type Theoretic Dynamic Logic

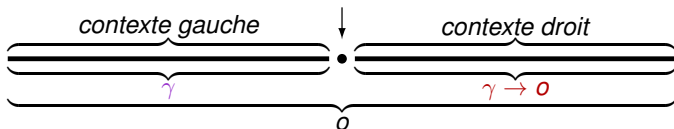


- **Dynamic Predicate Logic (DPL)**
[Groenendijk1991]

syntaxe standard mais sémantique
non-standard



- **Type Theoretic Dynamic Logic (TTDL)**
[deGroot2006] : purement montagovien pour le discours avec dynamicité par les continuations dans le λ -calcul



Retour sur le TAL : linguistique et informatique

- seconde guerre mondiale influence de la traduction automatique et de la cryptographie - grand essor jusqu'en 64 Alpac
the spirit is willing but the flesh is weak
The vodka is strong but the meat is rotten
- 70 - TAUM Météo Systran
- 80's approches symboliques premières grandes grammaires linguistique + logique manque de robustesse
- 90's premiers grands corpus, approches statistiques, premier PC domestique

Plan

- 1 Langage, logique, etc.
- 2 Pour une histoire croisée : linguistique, informatique et logique
- 3 Sémantique(s) des langues naturelles**

Des sémantiques

- sémantique lexicale (relations synonymie, antonymie, etc.)
- sémantique compositionnelle
- phénomènes sémantiques

Emmon Bach : Natural Languages Metaphysics

Des sémantiques

- sémantique lexicale (relations synonymie, antonymie, etc.)
explorer les relations structurelles entre les mots
- sémantique compositionnelle
- phénomènes sémantiques
Emmon Bach : Natural Languages Metaphysics

Des sémantiques

- sémantique lexicale (relations synonymie, antonymie, etc.)
explorer les relations structurelles entre les mots
- sémantique compositionnelle
comment construire les représentations ?
- phénomènes sémantiques
Emmon Bach : Natural Languages Metaphysics

Des sémantiques

- sémantique lexicale (relations synonymie, antonymie, etc.)
explorer les relations structurelles entre les mots
- sémantique compositionnelle
comment construire les représentations ?
- phénomènes sémantiques
Emmon Bach : Natural Languages Metaphysics
décrire la langue naturelle du point de vue de la réalité du monde

Des sémantiques

- sémantique lexicale (relations synonymie, antonymie, etc.)
explorer les relations structurelles entre les mots
- sémantique compositionnelle
comment construire les représentations ?
- phénomènes sémantiques
Emmon Bach : Natural Languages Metaphysics
décrire la langue naturelle du point de vue de la réalité du monde

Problèmes sémantiques

Problèmes sémantiques

- événements

Davidson

Juliette a embrassé Roméo

Problèmes sémantiques

- événements

Davidson

Juliette a embrassé Roméo

Juliette a embrassé Roméo rapidement

Problèmes sémantiques

- événements

Davidson

Juliette a embrassé Roméo

Juliette a embrassé Roméo rapidement

Juliette a embrassé Roméo rapidement à minuit

Problèmes sémantiques

- événements

Davidson

Juliette a embrassé Roméo

Juliette a embrassé Roméo rapidement

Juliette a embrassé Roméo rapidement à minuit

- souhaits

J'espère que vous viendrez

Problèmes sémantiques

- événements

Davidson

Juliette a embrassé Roméo

Juliette a embrassé Roméo rapidement

Juliette a embrassé Roméo rapidement à minuit

- souhaits

J'espère que vous viendrez

- modalité

Il est possible/probable/nécessaire que...

Problèmes sémantiques

- événements

Davidson

Juliette a embrassé Roméo

Juliette a embrassé Roméo rapidement

Juliette a embrassé Roméo rapidement à minuit

- souhaits

J'espère que vous viendrez

- modalité

Il est possible/probable/nécessaire que...

- temps

Aristote n'est vrai qu'à un moment de l'histoire

Problèmes sémantique et discours

- concepts flous

Paul est jeune, Beaucoup pensent que...

- présupposition

Pierre est allé chercher sa femme à l'aéroport

- résolution d'anaphore

Jean a une voiture. Elle est rouge.

- négation

Pierre n'est pas allé chercher sa femme à l'aéroport

Jean n'a pas une voiture. Elle est rouge.

Que faisons-nous ?

- développement de ressources pour la sémantique

framenet

grammaire sémantique du français et de l'anglais (ACG)

Que faisons-nous ?

- développement de ressources pour la sémantique

framenet

grammaire sémantique du français et de l'anglais (ACG)

- augmenter la pertinence de la modélisation

Que faisons-nous ?

- développement de ressources pour la sémantique

framenet

grammaire sémantique du français et de l'anglais (ACG)

- augmenter la pertinence de la modélisation
couverture - composition - calculabilité

Que faisons-nous ?

- développement de ressources pour la sémantique

framenet

grammaire sémantique du français et de l'anglais (ACG)

- augmenter la pertinence de la modélisation
couverture - composition - calculabilité
- réalité cognitive ?

Schizophrénie et Langage : Analyse et Modélisation

Que faisons-nous ?

- développement de ressources pour la sémantique

framenet

grammaire sémantique du français et de l'anglais (ACG)

- augmenter la pertinence de la modélisation
couverture - composition - calculabilité
- réalité cognitive ?

Schizophrénie et Langage : Analyse et Modélisation

- question importante pour la recherche : l'éthique

blog éthique et TAL <http://ethique-et-tal.org>