

# Logique et sémantique

Damien Nouvel



# Plan

1. Graphes pour la logique
2. Graphes conceptuels
3. Logiques de description
4. Web sémantique

# Logique et représentations

- ▶ Représentations logiques par **diagrammes**
  - **Euler** (~1750) : ensembles
  - **Venn** (~1880) : négation par coloration
  - **Caroll** (~1880) : négation par dichotomie
- ▶ **Sémiotique de C. S. Peirce** (~1910)
  - **Triade** : representamen / interprétant / référent
  - ⇒ **Triangle sémiotique**
  - **Signe** (representamen)
    - Indice : trace laissée par le référent
    - Icône (image, diagramme, métaphore) : représente le référent
    - Symbole : lien arbitraire avec le référent (règle)
  - **Coupure sémiotique** : distance entre signe et référent

# Graphes existentiels (Peirce)

- ▶ Représentation graphique
  - **Alpha** : logique propositionnelle
  - **Beta** : logique du premier ordre
  - Gamma : logique modale

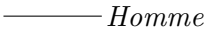
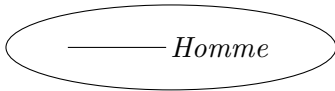


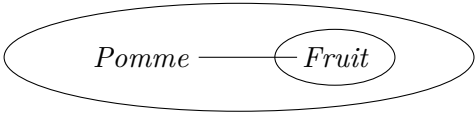
⇒ Utilisation de **symboles** et de **lignes**

- Feuille blanche : assertions vraies
- Ligne continue fermée : négation (ou coupure)
- Ligne d'identité : quantification existentielle
- Ligne entre symbole : relation (prédicat)

## Graphes existentiels alpha

Formule	Graphe
$A \wedge B$	$A$ $B$
$A \wedge \neg B$	$A$ $(B)$
$A \vee B \equiv \neg(\neg A \wedge \neg B)$	$(A)$ $(B)$
$A \rightarrow B \equiv \neg(A \wedge \neg B)$	$A$ $(B)$

## Graphes existentiels beta

Formule	Graphe
$\exists x, \text{Homme}(x)$	
$\neg \exists x, \text{Homme}(x)$	
$\exists x, \neg \text{Homme}(x)$	
$\exists x, \exists y, \text{Homme}(x) \wedge \text{Pomme}(y) \wedge \text{Mange}(x, y)$	
$\forall x \text{Pomme}(x) \rightarrow \text{Fruit}(x)$	

# Réseaux sémantiques

- ▶ Taxinomies de **Quillian** et **Collins** (~1965)
  - **Nœuds** : concepts (termes)
  - **Arcs** typés
    - **is-a** : relation d'agrégation (subsumption)
    - **has-a** : relation de composition (méronymie)
    - **kind-of** : relation d'instanciation
  - Wordnet (1985, anglais)
    - **Synsets** : synonymes
    - **Noms** : hyperonymes / hyponymes, meronymes / holonymes
    - **Verbes** : hyperonymes, troponymes, implication (entailment)
- ⇒ En 2012 : ~ 150 000 mots, ~ 120 000 synsets
- ⇒ Sémantique des arcs / liens / relations
- ⇒ Nombre limité de relations
- ⇒ Pas de disjonction explicite
- ⇒ Quantification par relation

# Plan

1. Graphes pour la logique
2. Graphes conceptuels
3. Logiques de description
4. Web sémantique



# Graphes conceptuels

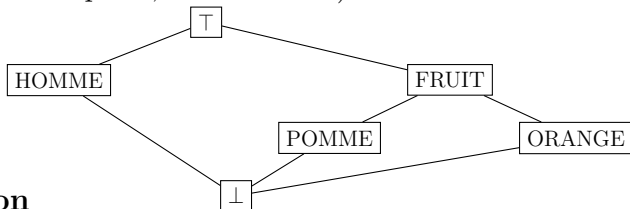
- ▶ J. F. Sowa (1984) : représentation de **connaissances**
- ⇒ Rapprochement entre **langage naturel** et **sémantique**
- ⇒ Notion de **scène** et **procès**
- ▶ Graphe inspiré des réseaux sémantiques
  - **Noeuds**
    - **Concepts** (objets du monde modélisé)
    - **Relations** (entre les objets)
  - **Arcs** orientés non typés
- ⇒ **Bipartite** (concepts / relations), connecté et fini
- ⇒ Les relations sont des noeuds



# Concepts et types

## ▸ Types

- **Ordre partiel** avec  $\top$  et  $\perp$
- ⇒ **Treillis** (subsumption, relation **is-a**)



## ▸ Quantification

- Par défaut : **existentielle**
- Référent comme **individu** ou **ensemble**
  - Homme : #58
  - Homme : Damien
  - Homme : Pierre, Paul, Jacques
  - Homme :  $\{*\}@4$
  - Homme :  $\forall$

# Rôles thématiques

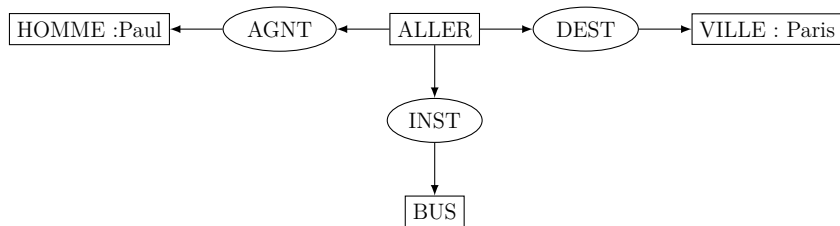
- ⇒ Importance des relations normalisées
- ⇒ Relations aussi partiellement ordonnées
  - ▶ **Relations** usuelles
    - **Agent** : acteur de l'action
    - **Patient** : subit l'action (siège)
    - **Instrument** : moyen pour réaliser l'action
    - Mais aussi : destination, thème, objet, expérient, cause, résultat, source, etc.
- ⇒ Lien entre la **syntaxe** et les rôles sémantiques
- ⇒ Étiquetage en rôles sémantiques
- ⇒ Encore beaucoup de difficultés
  - Voix passive
  - Relation temporelles
  - Adverbes et circonstants
  - ...

# Extensions des graphes conceptuels

- ▶ Possibilité d'imbriquer des graphes
  - Propositions relatives / conditionnelles
  - Difficultés pour les raisonnements
- ▶ Conversion en formules logiques
  - Fonction qui transforme un graphe en **forme linéaire**
  - Formats et syntaxe : CGIF, KIF, prédicats
  - Principe
    - Sommets non instanciés : variables
    - Sommets instanciés (marqueurs individuels) : constantes
    - Concepts : prédicats unaires
    - Relations : prédicats n-aires
  - ⇒ Quantification existentielle et conjonction
  - ⇒ Ensembles de **triplets**

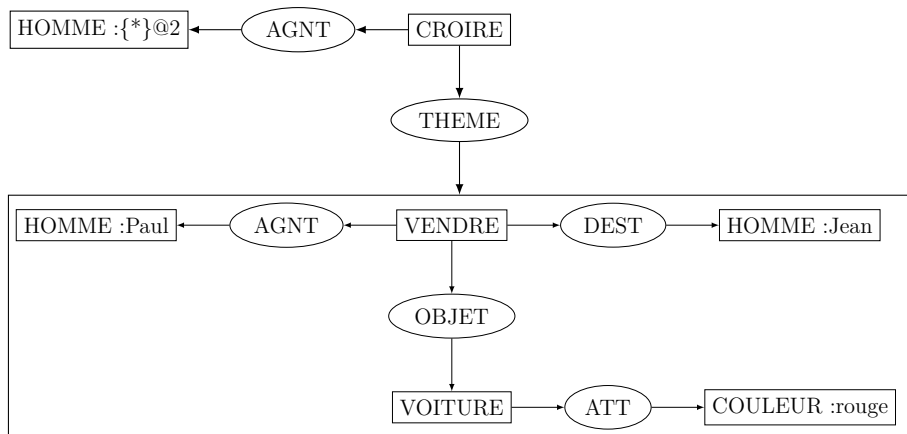
# Exemple

- ▶ Paul va en bus à Paris



## Exemple

- Les deux types croient que Paul vend sa voiture rouge à Jean



# Opérations sur les graphes

- ▶ **Dérivation** de graphes (spécialisation)
  - **Simplification** : fusion de relations identiques
  - **Jointure interne** : fusion de concepts identiques
  - **Restriction de type** : changement de type
  - **Restriction de référence** : ajout d'un individu
- ▶ **Composition** de graphes
  - **Somme** : juxtaposition de graphes
  - **Jointure** : copie de graphes avec fusion d'un concept
  - ⇒ Compatibilité des individus avec les types
  - **Projection** : recherche d'un graphe dans un autre
  - ⇒ Sous-graphe spécialisé
- ▶ Graphes de **définition** : expansion, contraction

# Algèbre des graphes

- ▶ Opérations de spécialisation
  - **Sous-graphe** : sélection de nœuds
  - **Sous-types** : identique excepté les types (restriction)
  - **Individu** : identique excepté les marqueurs (restriction)

⇒ Relation de spécialisation : **ordre partiel** sur les graphes ( $\leq$ )

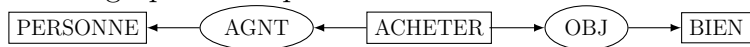
- ▶ **Canon**
  - Hiérarchie de types
  - Ensemble de marqueurs individuels
  - Relation entre marqueurs et types
  - Ensemble de graphes

⇒ Dérivations à partir de la **base de connaissances**



# Exercice

- Soit le graphe conceptuel :



- Indiquez les opérations et graphes résultants
- La personne qui achète est Jean
  - Le bien acheté est un livre
  - Quelqu'un achète un bien à Marie
  - C'est Jean qui achète un livre à Marie
  - Marie a lu un livre qui lui a été acheté
  - Qui a lu le livre acheté par Jean ?

# Plan

1. Graphes pour la logique
2. Graphes conceptuels
3. Logiques de description
4. Web sémantique

# Logique et terminologie

- ▶ Extension des **réseaux sémantiques**
- ▶ Représentation de connaissances **terminologique** :  $\langle T, A \rangle$ 
  - $T$  : T-Box, système Terminologique (notions, concepts)
  - ⇒ Définition des concepts
  - ⇒ Règles, formules
  - $A$  : A-Box, Assertions sur les individus (instances)
  - ⇒ Types de concepts :  $C(a)$  ou  $a : C$
  - ⇒ Relations (rôles) entre concepts :  $R(a, b)$  ou  $(a, b) : R$
  - ⇒ Faits, description de situations
- ⇒ Formalisme simplifié et **opérationnel**
  - ▶ Éléments (disjoints deux-à-deux)
    - **Concepts**
    - **Rôles**
    - **Individus**

# Concepts et rôles terminologiques

## ▸ T-Box

### • Concepts

- Ensembles d'individus
- **Inclusion** (conditions nécessaires) :  $\sqsubseteq$
- **Équivalence** (conditions nécessaires et suffisantes) :  $\equiv$   
⇒ Majuscules

### • Rôles

- Relations entre individus (ensemble, produit cartésien)  
⇒ Toujours binaires  
⇒ Minuscules

## ▸ A-Box

- Prédication unaires (concepts) ou binaires (rôles)
- Uniquement avec des individus

# Opérateurs des logiques de description

## ► Symboles

- Concepts : universel  $\top$  et absurde  $\perp$
  - Définition de concepts :  $\sqsubseteq$  et  $\equiv$
  - Négation :  $\neg$
  - Conjonction :  $\sqcap$
  - Quantification :  $\forall$  ou  $\exists$
  - Dénombrément :  $=$ ,  $\leq$  ou  $\geq$
- ⇒ Pas de variables
- ⇒ Pas de disjonctions

# Syntaxe des logiques de description

- ▶ Concepts ensemblistes
  - $\neg C_1$
  - $C_1 \sqcap C_2$
- ▶ Restrictions
  - $\forall r_1 C_1$  ou  $\exists r_1 C_1$
  - $= x r_1 C_1, \leq x r_1 C_1$  ou  $\geq x r_1 C_1$
- ▶ Axiomes
  - $C_1 \sqsubseteq C_2$
  - $r_1 \sqsubseteq r_2$
  - $C_1 \equiv C_2$
  - $r_1 \equiv r_2$

# Exemple

## ▸ T-Box

- Un homme est une personne
- $Homme \sqsubseteq Personne$
- Une femme est une personne
- $Femme \sqsubseteq Personne$
- On ne peut être homme et femme
- $Homme \sqcap Femme \sqsubseteq \perp$
- Une équipe est un ensemble d'au moins deux personnes
- $Equipe \equiv Ensemble \sqcap \forall membre. Personne \sqcap \geq 2membre$
- Une petite équipe contient moins de 5 membres
- $PetiteEquipe \equiv Equipe \sqcap \leq 5membre$
- Une équipe structurée contient un(e) responsable
- $Responsable \sqsubseteq Personne$
- $EquipeStructuree \equiv Equipe \sqcap \exists membre. Responsable$

# Interprétation

## ► Interprétation $I$

- Domaine  $\Delta^I$  : individus
- Fonction d'interprétation ensembliste
  - $\top^I = \Delta^I$
  - $\perp^I = \emptyset$
  - $(\neg A)^I = \Delta^I \setminus A^I$
  - $(C_1 \sqcap C_2)^I = C_1^I \cap C_2^I$
  - $(\forall r_1 C_1)^I = \{a \in \Delta^I \mid \forall b, (a, b) \in R_1 \rightarrow b \in C_1\}$
  - $(\exists r_1 C_1)^I = \{a \in \Delta^I \mid \exists b, (a, b) \in R_1 \wedge b \in C_1\}$



# Propriétés des logiques de description

## ▶ Satisfiabilité

- Un concept  $C$  est satisfiable s'il existe une interprétation  $I$  telle que  $C^I \neq \emptyset$

## ▶ Incohérence

- Un concept  $C$  est incohérent si pour toute interprétation  $I$  telle que  $C^I = \emptyset$

## ▶ Subsumption

- Un concept  $C_1$  est subsumé par un concept  $C_2$  si pour toute interprétation  $I$ ,  $C_1^I \subset C_2^I$

## ▶ Incompatibilité

- Deux concepts  $C_1$  et  $C_2$  sont incompatibles ssi pour toute interprétation  $I$ ,  $C_1^I \cap C_2^I = \emptyset$

# Extensions

- ▶ Logique **minimale**  $\mathcal{AL} = \{\top, \perp, \neg A, C_1 \sqcap C_2, \forall r_1 C_1, \exists r_1\}$
- ▶ Certaines parties sont des **extensions**
  - Négation de concepts complexes  
 $\mathcal{ALC} = \mathcal{AL} \cup \{\neg C\}$
  - Disjonction de concepts  
 $\mathcal{ALU} = \mathcal{AL} \cup \{C_1 \sqcup C_2\}$
  - Quantification existentielle typée  
 $\mathcal{AL\epsilon} = \mathcal{AL} \cup \{\exists r_1 C_1\}$
  - Cardinalités de rôles  
 $\mathcal{ALN} = \mathcal{AL} \cup \{= x r_1 C_1, \leq x r_1 C_1, \geq x r_1 C_1\}$
  - Conjonction de rôles  
 $\mathcal{ALR} = \mathcal{AL} \cup \{r_1 \sqcap r_2\}$
  - Rôles inverses  $\mathcal{AL\mathcal{I}}$ , composition  $\mathcal{AL\mathcal{H}}$ , range  $\mathcal{AL\mathcal{Q}}$ , ...

# Inférences

- ▶ Test de la subsomption
  - $C_1 \sqsubseteq C_2$  est vrai
  - $C_1 \sqcap \neg C_2$  est non satisfiable

⇒ Réfutation
- ▶ Méthode des **tableaux**
  - Décomposition des formules selon les connecteurs (arbres)
    - $a \wedge b$  : deux formules  $a$  et  $b$  sur la même branche
    - $a \vee b$  : chaque formule  $a$  et  $b$  sur une branche
  - Toutes les branches doivent être fermées (contradictions)
  - Branches non fermées : modèle pour l'ensemble de départ

# Avantages et implémentations

- ▶ **Avantages** des logiques de description
  - Par rapport au langage naturel
    - Sémantique formelle et définie
  - Par rapport à la logique
    - Décidabilité : entre propositions et du premier ordre

⇒ Utilisé pour le **web sémantique**

- ▶ Implémentations
  - Protégé <http://protege.stanford.edu>
  - Hermit <http://hermit-reasoner.com>
  - Pellet <http://clarkparsia.com/pellet>
  - Racer <http://www.racer-systems.com>
  - FaCT <http://www.cs.man.ac.uk/~horrocks/FaCT>

# Plan

1. Graphes pour la logique
2. Graphes conceptuels
3. Logiques de description
4. Web sémantique

# Ontologies

## ▶ Étymologie

- onto- : être
- -logie : science

⇒ Science de l'existence

⇒ Aristote, Thomas d'Aquin, Kant, Spinoza, Heidegger ...

## ▶ Informatique

- Modélisation des **connaissances** (explicatives) d'un domaine
- *A specification of a conceptualization.* (Gruber, 1992)
- *Une conceptualisation est une vue abstraite et simplifiée du monde que l'on veut représenter* (Gruber, 1992)

⇒ Visée applicative

- Partage de connaissances (interconnexion, web)
- Réutilisabilité (validité dans le temps)
- Vocabulaire consistant

⇒ Développement relativement récent

# RDF

- ▶ **Resource Description Framework** (W3C, 1999)
- ⇒ Bases de connaissances pour le web
  - Utilisation de **triplets** : (sujet, prédicat, objet)
  - ⇒ Base de connaissances
    - Description de faits
    - Intégrée au web (URLs, XML)
    - Requêtes de bases de données
  - ⇒ Mélange de standards
- ⇒ Partie A-Box
- ⇒ Pas vraiment de T-Box ...

# OWL

- ▶ **Ontology Web Language** (2004)
- ⇒ Extension de RDF / RDFS (schema)
- ⇒ Inspiré des logiques de descriptions
  - **Concepts**
  - **Rôles**
- ▶ Plusieurs versions
  - **OWL-Lite** : concepts et hiérarchies
  - **OWL-DL** : logique de description
  - **OWL-Full** : raisonnements (indécidables)
- ⇒ Partie T-Box



# Espaces de noms

- ▶ Restriction d'une ontologie à un **domaine**
  - Modéliser domaine par domaine
  - Importation d'ontologies de différents domaines
- ▶ Quelques ontologies populaires
  - **BabelNet** : équivalent WordNet multilingue
  - **FOAF** (Friend Of A Friend) : relation entre individus / sociétés
  - **Cyc** : univers du discours
  - **Dublin Core** : publication de documents
  - ...

# SKOS

- ▶ **Simple Knowledge Organization System** (2004)

⇒ Pour les terminologie / thésaurus

- ▶ Complète RDF et OWL
  - **Concepts** (URI)
  - Labels / étiquettes
  - Gestion du multilinguisme (XML **lang**)
- ▶ Relations sémantiques
  - Dans le vocabulaire
    - Généralisation : **broader**
    - Spécialisation : **narrower**
    - Associations : **related**
  - Entre vocabulaires : correspondances

# FOAF

▸ **Friend Of A Friend** (2000)

⇒ Pour modéliser les individus et leurs relations

- **Agent** : personnes, groupes, organisations
- **Document** : images, profils
- **Projets**

▸ Relation **membre** entre groupe et agents

▸ Propriétés

- Agents : date de naissance, organisation, email, site web ...
- Personnes : prénom, nom de famille, projet ...
- ...

⇒ Socle commun pour définir des organisations humaines

# Cas d'application : BNF

- ▶ **Bibliothèque Nationale de France**

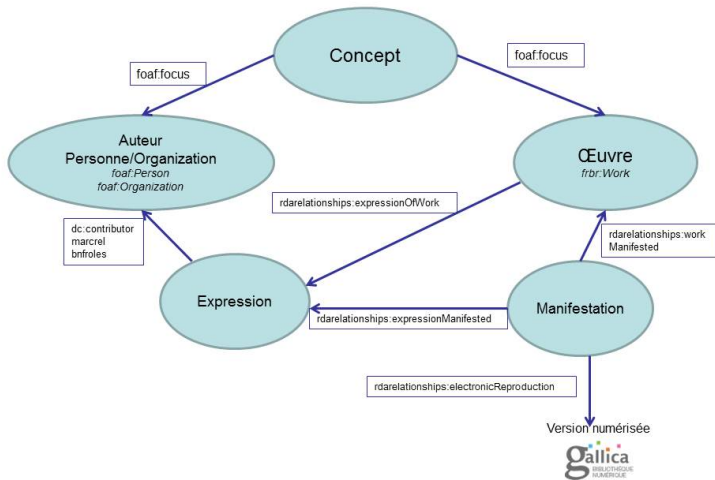
⇒ Catalogue d'ouvrages...d'auteurs, de pays, d'œuvres, etc.

- ▶ Utilisation de plusieurs standards du web sémantique, dont : Dublin Core, SKOS, FOAF, Geonames, IGN, INSEE, ISNI, OWL, RDFS ...

⇒ Volonté d'ouvrir l'accès aux données

- **Adresse (URI)** : <http://data.bnf.fr>
  - **Historique** : 2011 (35K pages) à aujourd'hui (900K auteurs)
  - Aspects **juridiques** : licence ouverte
- ▶ Possibilité de faire des requêtes SPARQL :  
<http://data.bnf.fr/sparql>

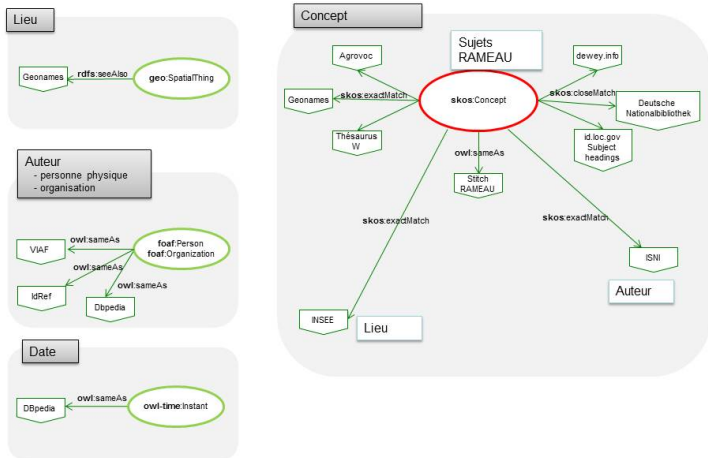
# Cas d'application : BNF (concepts)



## Cas d'application : BNF (structure)

data.bnf.fr

## Alignements vers des jeux de données externes



# Cas d'application : BNF (exercices)

- ▶ Faire les requêtes suivantes :
  - Quels auteurs sont morts pendant la 2ème guerre mondiale ?
  - Quels auteurs sont nés à Paris ?
  - Quels ouvrages sur Paris ont été écrit par des parisiens ?

# Exercice : habitation

- ▶ Ontologie des habitations
  - Concepts à hiérarchiser : bien, logement, espace (interieur ou exterieur), pièces possibles (cuisine, salle d'eau, douche, WC, salle de bains, salon, bureau, chambre), jardin, terrasse, etc.
  - Définir la propriété de bien qui contient des espaces
  - Définir par contraintes les logements suivants
    - Habitation salubre (avec salle d'eau)
    - Studio (chambre, salle d'eau)
    - Appartement (au moins 3 pièces)
    - Petite habitation (au plus 4 espaces intérieurs)
    - Grande habitation (au moins 4 pièces hors salles d'eau)
    - Bureau (pas de salle de bain)
    - Habitation prestigieuse (toutes prestations)
    - Villa (avec un jardin)
  
- ▶ Ajouter une description d'habitation comme individus...