

Ontologie

Olivier Corby

Ontologie

Etude de ce qui *est*
terme emprunté à la philosophie

Pour nous : *formalisation d'une conceptualisation*

Ontologie

Se mettre d'accord sur le *sens* des termes employés dans une organisation, une communauté, un métier

Faire en sorte que les personnes et les logiciels se *comprennent*

Utile pour des applications distribuées telles que le Web

3

Ambiguïté

‘chambre’ :

Chambre d'hôtel ?

Chambre d'écho ?

Chambre des députés ?

Chambre d'enregistrement ?

Chambre noire ?

Chambre funéraire ?

4

Ontologie

- Identifier, modéliser les *concepts* d'un domaine, pertinents pour une/des applications
- Se mettre d'accord, au sein d'une communauté, sur les *termes* employés pour se référer à ces concepts

5

terme

Mot de la langue naturelle qui désigne un (des) concept(s) :

cat, chat, greffier, matou

termes qui désignent le concept de *chat*

→ synonymie : plusieurs termes dénotent le même concept

→ ambiguïté : plusieurs concepts dénotés par le même terme

6

Ontologie

- Identifier/modéliser les concepts et les *relations* conceptuelles
- formaliser la conceptualisation, et le vocabulaire correspondant
- Formalisation pour lever toute ambiguïté

7

Pour quoi faire ?

- une base de connaissances à objets
- un système d'annotation
- un système d'indexation documentaire, de recherche d'information
- Commerce électronique

8

Ontologie

- Différentes acceptions du mot ontologie :
- Vocabulaire technique,
- Référentiel métier,
- Terminologie/thesaurus,
- Système de classes d'une représentation par objet : UML ?
- Base de connaissances terminologique ⁹

Conception

- Identifier/modéliser les concepts et les termes pertinents
- Identifier les relations pertinentes : subClassOf, isa, partOf, hasPart, closeTo, over, under, contain, connected, etc.
- Règles pour combiner les concepts et les relations : partOf est transitive

Modéliser

- Une ontologie est une *modélisation* des connaissances
- Au sens où une théorie physique est un modèle de la réalité
- Il peut y avoir plusieurs modèles d'une même réalité :
 - expérience de la chute des corps
 - théorie de la gravitation de Newton
 - relativité générale d'Enstein

11

Connaissance

- Continuum : donnée, information, connaissance
 - donnée : ... --- ...
 - information : SOS
 - connaissance : **en cas d'alerte,
déclencher les secours**
- La connaissance permet de produire de nouvelles données, informations, connaissances : inférence

12

Types de connaissances

- Connaissances de résolution de problème :
 - conception, diagnostic, évaluation, planification
 - tâches, inférences
- Connaissances du domaine
 - e.g. électronique, mécanique, médecine, etc.
 - ontologie : *réutilisable*
 - modèles de domaine : fonctionnel, causal,

13

Ontologie

concept Object

concept Engine < Object

concept Vehicle < Object

relation partOf :

Object -> Object

14

Modèles de domaine

Structurel :

Engine e1 partOf Vehicle v2

causal :

Engine breakdown => Vehicle stop

15

Modèles de tâche

Conception :

- specify, design, assess

Diagnostic :

- identifier des symptômes, choisir un modèle causal, émettre des hypothèses, imaginer des conséquences, les tester, etc.

16

Ontologie

- L'accent est mis sur les *concepts* et les *relations* du domaine
- Orienté modélisation, spécification, réutilisation, partage, standardisation.
- Relations spécifiques du domaine considéré

17

Objets

- L'accent est mis sur le côté *opérationnel*.
- Opérations rendues possibles par le modèle objet : classification, évolution, simulation, calcul.
- Orienté résolution de problème

18

Terminologie/Thesaurus

- L'accent est mis sur les *termes* utilisés pour dénoter les concepts du domaine (linguistique, langue naturelle).
- Les relations entre termes sont *linguistiques*.
- Orienté indexation.

19

Thesaurus (2)

Recueil documentaire alphabétique de termes servant de descripteur pour :

- analyser un corpus
- indexer des documents
- Relations prédéfinies standardisées pour les thésaurus

20

Relations linguistiques

- BT : Broader Term
- NT : Narrower Term
- TT : Top Term
- RT : Related Term (other than BT, NT, TT, etc.)

21

Relations (2)

- USE : Preferred Term
- UF : Use for, non preferred synonym, quasi synonym
- SN : Scope Note, Une note pour expliquer un terme

22

Ontologie ?

- **Composant réutilisable**
- **Réutilisation** : généralité, abstraction (reuse)
- **Partage** : consensus, standardisation (sharing)
- Accord sur conceptualisation partagée : engagement ontologique (commitment)

23

Caractéristiques

- Formalisation : lever les ambiguïtés
- Indépendance (relative) par rapport à une tâche, ou un problème précis
- Utilisable pour différentes tâches : conception, diagnostic, maintenance, recherche d'information

24

Caractéristiques (2)

- Indépendante d'une implémentation : limiter le biais dû à un formalisme de représentation : niveau conceptuel
- Des inférences sont possibles (relations transitives, axiomes, etc.)

25

Types d'Ontologies

Ontologies générales, abstraites, de haut niveau :

- Ontologie des catégories conceptuelles :
- objet, événement, état, processus, action, temps, espace

Ontologies théoriques :

- physique, mathématique, cinématique

26

Types (2)

Ontologie applicative :

- Médecine, automobile, patrimoine culturel, organisation, etc.

27

Principes de conception

- Clarté : communiquer le sens des termes définis
- Cohérence
- Extensibilité
- Indépendance vs implémentation
- Modularité

28

Outils/Langages

- KIF : Knowledge Interchange Format
- Ontolingua
- Logiques de description (orientée classification)
- Graphe conceptuel (support)
- RDF Schema
- DAML OIL

29

Exemples

- Ontolingua : www-ksl.stanford.edu/sns.html
- WordNet : Ontologie de la langue naturelle
- Enterprise Ontology
- Kactus : ingénierie
- UMLS : Unified Medical Language System
- Cyc

30