

# Principi di Organizzazione Ontologica

Corso di Intelligenza Artificiale: Trattamento Automatico di Linguaggio Naturale

Veronica Tomatis  
Pisa, 10 giugno 2002

Dipartimento di Informatica - Università di Pisa

# Overview

- Definizione del concetto di ontologia (e motivazioni)
  - Filosofica
  - Informatica
  - Un esempio concreto: Yahoo!
  - Relazioni con altre metodologie
- Ontological Engineering
  - Criteri di progettazione
  - Approcci di sviluppo
  - Problemi aperti
  - Top-level Ontology & Formal Ontology: Ontoclean
- Conclusioni e Bibliografia

Veronica Tomatis, 10 Giugno 2002

2

# Ontologia in Filosofia

*Ontos + logos*

“Disciplina filosofica relativa all’ordine e alla struttura dell’essere in generale”

κατηγορια aristotelica

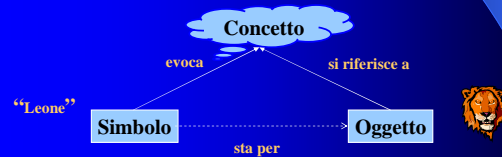
10 categorie per classificare ogni cosa che può essere detta o “predicata” su qualcosa

Veronica Tomatis, 10 Giugno 2002

3

# Ontologia in Filosofia (II)

- Domande aperte:
  - Che cos’è l’essere?
  - Quali sono le caratteristiche comuni a tutti gli esseri?



Veronica Tomatis, 10 Giugno 2002

4

## Ontologia in Informatica

“Catalogo delle tipologie che si assume esistano in un dominio di interesse **D** dalla prospettiva dell'individuo che usa un linguaggio **L** al fine di parlare di **D**”

“An ontology is a specification of a conceptualization” [Gruber '95]

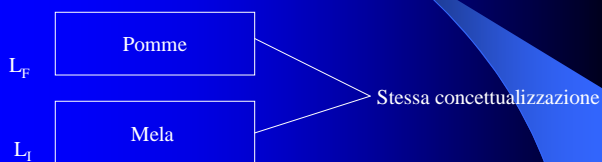
## Concettualizzazione

- La struttura formale della realtà come percepita e organizzata da un agente

*Indipendente da:*

- Vocabolario usato
- Specifica situazione

## Concettualizzazione: esempio



## Ontologia: AI

- Un artefatto ingegneristico costituito da uno specifico vocabolario usato per descrivere un certa realtà accompagnato da un set di assunzioni esplicite riguardo il significato delle parole del vocabolario
- “La dichiarazione di una teoria logica” [Gruber '95]
  - Logica del Primo Ordine è ontologicamente neutrale
    - No vincoli
    - No caratterizzazioni
  - Logica + Ontologia = linguaggio per esprimere relazioni tra entità del dominio di interesse

## Livelli di Ontologia

- **Lexicon:** Vocabolario con definizioni in NL
- **Tassonomia:** Ordinamento sistematico
- **Thesaurus:** Repertorio dei termini chiave utilizzati per il reperimento di informazioni in una banca dati (Tassonomia + termini relati)
- **Modelli Relazionali**
- **Teoria completamente assiomatica**

Veronica Tomatis, 10 Giugno 2002

9

## Motivazioni

- Le sorgenti di dati elettronici sono in continua espansione
- L'accesso, la ricerca e la sintesi di informazioni diventano sempre più complessi
- Problema: **Gap** fra

significato  
dell'informazione

vs

attuale  
memorizzazione delle  
informazioni

Veronica Tomatis, 10 Giugno 2002

10

## Scopi

- Fornire un linguaggio comune per condividere e riusare conoscenza su un fenomeno nel dominio di interesse  
*"People can't share knowledge if they don't speak a common language" [T.Davenport, Working Knowledge]*
- Campi di interesse: importante nelle attività in cui c'è bisogno di condividere conoscenza

Veronica Tomatis, 10 Giugno 2002

11

## Campi di Applicazione

- Natural language processing and understanding (WordNet)
- Information extraction
- Information integration
- Semantic Web
- E-Commerce (e.g. cataloghi di prodotti)
- Knowledge management
- Knowledge discovery
- Multi-agent systems

Veronica Tomatis, 10 Giugno 2002

12

## Compiti

Ad esempio nel Semantic Web (esempio di conoscenza eterogenea accumulata):

- Classificare
- Etichettare
- Definire
- Trovare
- Integrare
- Usare

Veronica Tomatis, 10 Giugno 2002

13

## Un Esempio di Base

Yahoo! : una gerarchia tassonomica di classi su Web

## Un esempio



- “Yet Another Hierarchical Officious Oracle” (indice gerarchico)

- David Filo and Jerry Yang, Ph.D. at Stanford University, February 1994
- 150.000 categorie

Veronica Tomatis, 10 Giugno 2002

15

## YAHOO! : Categorie

- Offrono un modo standard e universale di riferirsi e descrivere la natura di “cose”, attività e documenti
- Caratterizzano semanticamente il contenuto dei documenti
- Realizzato a mano!

<b>Arts &amp; Humanities</b> Literature, Photography...	<b>News &amp; Media</b> Full Coverage, Newspapers, TV...
<b>Business &amp; Economy</b> B2B, Finance, Shopping, Jobs...	<b>Recreation &amp; Sports</b> Sports, Travel, Autos, Outdoors...
<b>Computers &amp; Internet</b> Internet, WWW, Software, Games...	<b>Reference</b> Libraries, Dictionaries, Quotations...
<b>Education</b> College and University, K-12...	<b>Regional</b> Countries, Regions, US States...
<b>Entertainment</b> Ficks, Movies, Humor, Music...	<b>Science</b> Animals, Astronomy, Engineering...
<b>Government</b> Elections, Military, Law, Taxes...	<b>Social Science</b> Archaeology, Economics, Languages...
<b>Health</b> Medicine, Diseases, Drugs, Fitness...	<b>Society &amp; Culture</b> People, Environment, Religion...

Veronica Tomatis, 10 Giugno 2002

16

## Relazioni con Altre Metodologie

- Ontologie vs Data Base
- Ontologie vs Knowledge Base

## Ontologia vs Schema Concettuale

### Ontologia

- Semantica formale
- Si specificano tutti i vincoli rilevanti per caratterizzare il significato

### Schema Concettuale

- Non sempre ha un a semantica formale
- Si specificano solo i vincoli rilevanti per l'aggiornamento del database

Veronica Tomatis, 10 Giugno 2002

18

## Ontologia vs Basi di Conoscenza

### Ontologia

- Sviluppata con lo scopo di stabilire consenso sui termini
- Fortemente espressiva

### Basi di Conoscenza

- Disegnata per risolvere un determinato problema
- Limitatamente espressiva

Veronica Tomatis, 10 Giugno 2002

19

## Ontological Engineering

- Criteri
- Approcci
- Problemi Aperti
- Un approccio: Formal Ontology
- Un esempio: Ontoclean

## Ontological Engineering

*Principi per progettare, modificare, applicare e valutare un' ontologia*

Le caratteristiche dell'ontologia devono risultare sufficientemente complete, corrette e concise per i potenziali utenti

### *Ontological Commitment:*

accordo tra le parti ad adottare una particolare ontologia per comunicare su un particolare dominio di interesse in modo coerente e **consistente** e beneficiare della conoscenza altrui (non garantisce la **completezza**)

## Criteri

1. Chiarezza
2. Coerenza
3. Estendibilità
4. Minimizzare l' Encoding Bias
5. Minimizzare l' Ontological Commitment

## Criteri (I)

### 1. Chiarezza

- Definizioni oggettive indipendenti dal contesto
- Adottare un formalismo
- Definizioni espresso con assiomi logici
- Definizioni complete (predicati definiti da condizioni necessarie e sufficienti)
- Definizioni documentate in linguaggio naturale

### 2. Coerenza

- Le sentenze inferite dagli assiomi non devono essere in contraddizione con le definizioni

## Criteri (II)

### 3. Estendibilità

- Disegnare l' ontologia per anticiparne gli usi futuri
- Estensione e specializzazione del vocabolario
- No revisione del vocabolario esistente

## Criteria (III)

### 4. Minimizzare l' Encoding Bias

- Specifica della concettualizzazione a livello di conoscenza senza dipendere da un particolare livello simbolico di encoding
- Scelte di rappresentazione non basate su convenienza di rappresentazione o implementazione (diversi sistemi e stili di rappresentazione di conoscenza)

### 5. Minimizzare l' Ontological Commitment

- Definire (concettualizzare) solo i termini essenziali per condividere conoscenza

## Approcci

*Usati nella fase di progettazione o modifica (feedback-valutazione)*

- I. Creazione (Inspirational)
- II. Induzione
- III. Deduzione
- IV. Sintesi
- V. Collaborativo
- .....Modelli ibridi

## I. Approccio Creativo

*Perché è necessaria un'ontologia?*

- Uso dell' immaginazione individuale, della creatività del punto di vista personale sul dominio di interesse
- Ontologia che risponda ai bisogni specifici individuati
- Pericolo: limitato Ontological Commitment

## II. Approccio Induttivo

- Osservare, esaminare e analizzare specifici casi nel dominio di interesse
- La caratterizzazione per uno specifico caso è applicata agli altri casi del dominio di interesse

### III. Approccio Deduttivo

- Adottare principi generali
- Applicare i principi in modo adattivo alla costruzione di specifiche ontologie
- "Filtrare" e "distillare" le nozioni generali per dedurre ontologie *ad hoc*

### IV. Approccio Sintetico

- Identificare un set di base di ontologie (no sussunzione)
- Ogni ontologia fornisce una caratterizzazione parziale del dominio di interesse
- "Sintetizzare" per sviluppare un' ontologia unificata
  - Integrare sistematicamente i concetti
  - Eliminare le caratterizzazione meno sviluppate
  - Uniformare le differenti terminologie

### V. Approccio Collaborativo

- Disegno e sviluppo sono frutto del lavoro di un team
- Esperienza, background e punto di vista delle persone che collaborano allo sviluppo
- Pericolo: troppe persone direttamente coinvolte



### Confronto

- |  |   |
|--|---|
| <ul style="list-style-type: none"><li>● <b>Creativo</b><ul style="list-style-type: none"><li>- Manca di base teoriche</li><li>+ Ontologie uniche ed innovative</li></ul></li><li>● <b>Induttivo</b><ul style="list-style-type: none"><li>- Non generalizzabile</li><li>+ Adatto a casi specifici</li></ul></li><li>● <b>Deduttivo</b><ul style="list-style-type: none"><li>- Presuppone l' esistenza di uno schema di caratterizzazioni generali</li></ul></li></ul> | <ul style="list-style-type: none"><li>● <b>Sintetico</b><ul style="list-style-type: none"><li>- Si affida alla capacità di sintesi degli sviluppatori</li></ul></li><li>● <b>Collaborativo</b><ul style="list-style-type: none"><li>- Si affida alla capacità individuali dei singoli sviluppatori</li><li>+ Possibilità di valutazione (qualità) del lavoro in corso d'opera</li></ul></li></ul> |
|--|---|



## Problemi Aperti e Soluzioni Proposte

## Problemi Aperti

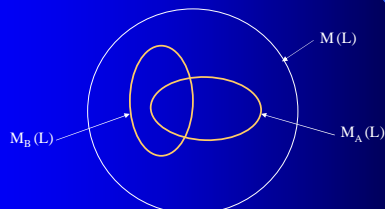
- Integrazione di ontologie
- Ontologie dipendenti dal dominio
- Valutazione delle scelte di progettazione

Veronica Tomatis, 10 Giugno 2002

34

## Integrazione di Ontologie (I)

Due sistemi A e B che usano lo stesso vocabolario possono "comunicare" solo se hanno la stessa concettualizzazione o se i modelli della loro concettualizzazione si "sovrappongono"



Veronica Tomatis, 10 Giugno 2002

35

## Integrazione di Ontologie (II)

L' approccio bottom-up all'integrazione di sistemi basati su ontologie locali non garantisce **consistenza:**

- Relazioni concettuali rilevanti per il contesto specifico
- Approssimazioni *ad hoc* per i modelli locali

**Alternative all' intersezione di differenti ontologie?**

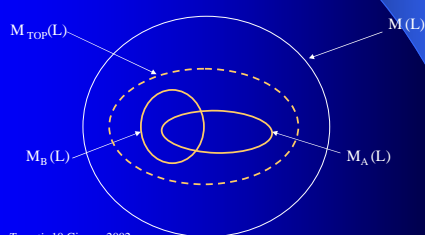
*Ontologia Top-Level*

Veronica Tomatis, 10 Giugno 2002

36

## Ontologia Top-Level

Descrive in modo molto generale concetti (eventi, azioni, oggetti ...) indipendenti dal dominio specifico

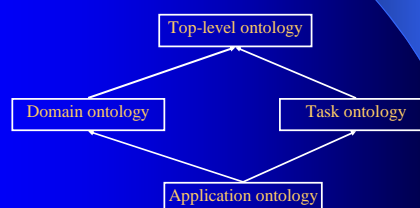


Veronica Tomatis, 10 Giugno 2002

37

## Ontologia: Livelli di Generalità (I)

Tipi di ontologie in rapporto al loro livello di generalità



Veronica Tomatis, 10 Giugno 2002

38

## Ontologia: Livelli di Generalità (II)

- *Top-level*: si descrivono concetti generali indipendenti dal dominio
- *Domain e Task*: si descrivono il vocabolario relativo al generico dominio (e.g. automobili) o compiti e attività generiche (e.g. vendere) specializzando i termini introdotti nell'ontologia Top-level
- *Application*: si descrive concetti dipendenti dal particolare dominio e compito (specilizzazione di Domain a Task Ont.)

Veronica Tomatis, 10 Giugno 2002

39

## Un Approccio Formale:

superare la dipendenza dal dominio e valutare le scelte di progettazione

- Formal Ontology
- Ontoclean

## Formal Ontology

Approccio metodologico per definire un framework  
**generale e rigoroso** per:  
*comprendere, paragonare e valutare* scelte ontologiche

Basato su:

- Logica
- Filosofia
- Linguistica

## Formal Ontology: Metodologia

Definire distinzioni e relazioni formali fra

- Entità del mondo, come percepite
- Categorie usate per parlare delle entità

**Goals:**

- Caratterizzare particolare e universale attraverso proprietà e relazioni formali indipendenti dal dominio
- Definire una metodologia di modellazione unificata per DB, KB e sistemi OO

## Un Modello: Ontoclean

- Metodologia basata su relazioni formali
- Si definiscono una serie di metaproprietà usate per caratterizzare aspetti rilevanti di classi e relazioni
- Le metaproprietà impongono vincoli alla struttura tassonomica delle ontologie

## Ontoclean: Nozioni Formali

- Essenza e Rigidità
- Identità e Unità

## Essenza e Rigidità (I)

- Una proprietà è *essenziale* per un' entità se necessaria a mantenerne l' identità

Es.:

- *Marco deve avere un cervello*
- *Marco deve essere una persona*

- Una proprietà è *rigida* se è essenziale per tutte le sue istanze

- Proprietà rigide: (*essere una persona*)
- Proprietà semi-rigide: (*essere rigido/duro*)
- Proprietà anti-rigide: (*essere uno studente*)

Veronica Tomatis, 10 Giugno 2002

45

## Essenza e Rigidità (II)

- Forniscono informazioni sul significato inteso delle proprietà
- Impongono vincoli sulla relazione di sussunzione (verifica della consistenza dei link tassonomici)

Una proprietà anti-rigida non può sussumere proprietà rigide



INCONSISTENZA

e.g. (*la classe studente non può sussumere la classe persona*)

Veronica Tomatis, 10 Giugno 2002

46

## Identità e Unità (I)

- **Identità:** problema di riconoscere entità individuali come uguali

e.g. Durata - Intervallo di tempo (sottoclasse)

- "1.00-2.00- di giovedì prossimo"
- "2.00-3.00- di mercoledì prossimo"

Stessa durata ma intervalli differenti!!!

- **Unità:** problema di riconoscere tutte le parti che compongono un'entità individuale

e.g. Acqua - Oceano (sottoclasse)

Veronica Tomatis, 10 Giugno 2002

47

## Identità e Unità (II)

Confusioni del linguaggio naturale in cui si può incorrere quando lo si usa per descrivere il "mondo"

"Tutti gli intervalli di tempo sono durate"

=

"Tutti gli intervalli di tempo hanno una durata"

&

"Gli oceani non sono tipi di acqua ma sono composti di acqua"

Veronica Tomatis, 10 Giugno 2002

48

## Ontoclean: Relazioni

- Sussunzione
- Istanziamento
- Parte di/Tutto
- Polisemia

## Sussunzione

*Sussunzione: relazione is-a, sottoclasse*

Formalmente: ogni istanza della sottoclasse è necessariamente istanza della superclasse

Estremamente utile per impartire una struttura all'ontologia ma spesso "inflazionata"

## Istanziamento

*È Umano una sottoclasse di Specie?*

- Specie: identificato dalla posizione assunta nella tassonomia biologica
- Umani: identificati dalla posizione spazio/temporale dei loro corpi (caso più semplice)

*Istanziamento non è sussunzione*

## Parte di/Tutto

Sottoclasse è analogo a sottoinsieme (= parte di)

*Macchina / Motore*

Macchina è sottoclasse di motore?

- Macchina: capace di portare persone
- Motore: capace di generare una forza rotazionale

*"Parte di" non è sussunzione*

## Polisemia

Significati multipli di un unico termine

*Libro:* - volume rilegato con un peso, ecc.  
- entità astratta con autore, titolo ecc.

- Volume: identificato dalla posizione spazio/temporale
- Entità astratta: identificati da autore, titolo ecc.

Nessuna istanza può soddisfare entrambi i criteri di identità: si tratta di due diverse classi di identità

*Polisemia non è sussunzione*

## Vantaggi nell' Ontological Analysis

- Si acquisiscono importanti informazioni sulla struttura dell'ontologia e sulla natura di ciò che sta rappresentando
- Si identifica la spina dorsale della struttura tassonomica (proprietà e relazioni essenziali)
- Si scoprono gli usi inconsistenti della sussunzione nella tassonomia

## Conclusioni

- Definizione di Ontologia (filosofica, AI)
- Concetti di Ontological Engineering (criteri e approcci)

**Filosofia e Linguistica** = un contributo concreto in Informatica, AI per "ingegnerizzare" la conoscenza e rappresentare realtà di senso comune

- Nuovo punto di vista
- Idee per l'analisi e lo sviluppo
- Nuovi strumenti e soluzioni (Formal Ontology)

## Bibliografia (I)

1. Guarino, N. 1995. Formal Ontology, Conceptual Analysis and Knowledge Representation. International Journal of Human and Computer Studies, 43(5/6): 625-640.
2. Guarino, N. and Welty, C. 2002. Evaluating Ontological Decisions with OntoClean. Communications of the ACM, 45(2): 61-65.
3. Gruber, T.R. Towards principles for the design of ontologies used for knowledge sharing. In Guarino and Poli, R., Eds. International Workshop on Formal Ontology. N. Padova, Italy, 1993.
4. Castel, F. 2002. Ontological computing. Communications of the ACM, 45(2): 29-30.
5. Holsapple, C.W. and Joshi, K. D. 2002. A collaborative approach to ontology design. Communications of the ACM, 45(2): 42-47.
6. Henry, K. 2002. Predicting how ontologies for the semantic web will evolve. Communications of the ACM, 45(2): 48-54.
7. Everett, J.O. et al. 2002. Making ontologies work for resolving redundancies across documents. Communications of the ACM, 45(2): 61-65.
8. Jasper, R. and Uschold, M. A framework for understanding and classifying ontology applications. in ICAI-99 Ontology Workshop. Stockholm, Sweden, July, 1999.

## Bibliografia (II)

9. Guarino, N. 1998. Formal Ontology in Information Systems. In N. Guarino (ed.) Formal Ontology in Information Systems. Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998. IOS Press, Amsterdam: 3-15.
10. Labrou, Y. and Finin, T. Yahoo! as an ontology--Using Yahoo! categories to describe documents. In Proceedings of the 8th International Conference on Information and Knowledge Management. Kansas City, MO, Nov. 1999, pp. 180-87.
11. Guarino, N. 1994. The Ontological Level. In R. Casati, B. Smith and G. White (eds.), Philosophy and the Cognitive Science. Hölder-Pichler-Tempsky, Vienna: 443-456.
12. Guarino, N., Carrara, M., and Giaretta, P. 1994. An Ontology of Meta-Level Categories. In D. J., E. Sandewall and P. Torasso (eds.), Principles of Knowledge Representation and Reasoning: Proceedings of the Fourth International Conference (KR94). Morgan Kaufmann, San Mateo, CA: 270-280.
13. Gruber, T. R. (1993). A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition, 5(2), 199-220.

## Bibliografia (III)

14. Sowa, John F. "Building, Sharing, and Merging Ontologies"
15. Sowa, John F. (1996a) "Top-level ontological categories." International J. of Human-Computer Studies, 43:5/6, 1996, 669-686.