An Introduction to applied ontology

Thomas Bittner
bittner3@buffalo.edu
Department of Philosophy

Overview

1. Motivation for ontologies in information systems
2. What is an ontology?
3. Kinds of ontologies
4. Top level ontologies
5. Conclusions

Intro and motivation:
Why do we need ontologies?

Information Integration

Pieces of Information
- from different sources
- In different formats
- Of varying quality
- …

The ideal situation

Compatible data formats
Compatible semantics
Incompatible Data Formats

Incompatible Semantics

Semantic Heterogeneity

Different words similar meaning

Application 1
Lake Powell is a Body of water

Transform data

Application 2
Lake Powell is a Lake

How is a computer supposed to know that a lake is a body of water?

Application 1
Lake Powell is a Body of water

Transform data

Application 2
Lake Powell is a Lake

How is a computer supposed to know that some bodies of water are lakes?
Semantic Heterogeneity

**Same word different meaning**

‘tank’

Ontologies

help us (and computers) to deal with Semantic Heterogeneity

Comparing FMA and GALEN

**FMA**
- Foundational Model of Anatomy
- Uni-Washington
- Protégé – frame system

**GALEN**
- Anatomy, medical procedures, etc.
- Uni-Manchester
- GRAIL – description logic

Parthood relations

<table>
<thead>
<tr>
<th>FMA</th>
<th>GALEN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Transform data

How is a computer supposed to know that both applications mean completely different things when using the word ‘tank’?
First approach:
Let’s try linguistic analysis!!

We do linguistic analysis!!
Bodenreider at. al

We do linguistic analysis!!
Bodenreider at. al

- **Identify** anchor concepts by means of **linguistic analysis**
  - matching index: 0
    - GALEN[912]=polyhedral
    - FMA[69751]=Polyhedral
  - matching index: 9
    - GALEN[1204]=TemperatureValue
    - FMA[69888]=Temperature value
  - matching index: 13
    - GALEN[7242]=ReticuloendothelialSystem
    - FMA[57800]=Mononuclear phagocyte system

- **Verify** anchor concepts by means of **structural analysis**
  - Structural analysis: compare information about parthood relations that hold between the anchor concepts in FMA and in GALEN
  - Structural information provides **positive or negative** evidence for anchor concepts
    - Positive evidence: compatible part-of relations
    - Negative evidence: incompatible part-of relations
We do linguistic analysis!!
Bodenreider at. al

Results:
Anchors identified by lexical alignment
2,353 matching anchor concepts were identified lexically, accounting for about 4% of FMA concepts and 9% of GALEN concepts.

Why are there only so few matches?

The fundamental problem of linguistic approaches:

```
FMA
\downarrow\text{downgrade}
FMA
\text{as thesaurus}
```

```
\text{Mapping based on Linguistic and structural analysis}
\downarrow\text{downgrade}
GALEN
\text{as thesaurus}
```

Formal ontology which makes explicit the semantics of the terms that are used

```
\text{Ontology}
\downarrow
\text{FMA}
\text{Protégé, Frame system}
\downarrow
\text{GALEN}
\text{GRAIL, Description logic}
```

Formal ontology which makes explicit the semantics of the terms that are used

```
\text{Ontology}
\downarrow
\text{FMA}
\text{Protégé, Frame system}
\downarrow
\text{GALEN}
\text{GRAIL, Description logic}
```

Formal ontology which makes explicit the semantics of the terms that are used

```
\text{Ontology}
\downarrow
\text{FMA}
\text{Protégé, Frame system}
\downarrow
\text{GALEN}
\text{GRAIL, Description logic}
```

So what are Ontologies ???

- Definition (Guarino 1998):
  - A shared vocabulary plus a specification (characterization) of its intended meaning
Ontologies

- Guarino:
  - A shared **vocabulary** plus a specification (characterization) of its intended **meaning**

  Collection of symbols
  Meaning of the symbols

Ontologies

- Guarino:
  - A shared **vocabulary** plus a specification (characterization) of its intended **meaning**

  Collection of symbols:
  - "Tom-Bittner"
  - "tank"
  - "bald"
  - "part-of"

Symbols and their meaning

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Refers to</th>
<th>Referent</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Tom Bittner&quot;</td>
<td>Refers to</td>
<td>&quot;The presenter of the ontology class&quot;</td>
</tr>
<tr>
<td>&quot;bald&quot;</td>
<td>Refers to</td>
<td>&quot;...&quot;</td>
</tr>
</tbody>
</table>

Two symbols same referent
Symbols and their meaning

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Refers to</th>
<th>Referent</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;part-of&quot;</td>
<td>(Tom’s head, Tom), (Tom’s nose, Tom’s head), (Tom’s nose, Tom),</td>
<td></td>
</tr>
</tbody>
</table>

Ontologies

- Guarino:
  - A shared vocabulary plus a specification (characterization) of its intended meaning

Which symbol goes with which referent.

Ontologies

- Guarino:
  - A shared vocabulary plus a specification (characterization) of its intended meaning

Fish ontology:

```
“tank” Refers to
```

Ontologies

- Guarino:
  - A shared vocabulary plus a specification (characterization) of its intended meaning

Military ontology:

```
“tank” Refers to
```
Ontologies

- Guarino:
  - A shared vocabulary plus a specification (characterization) of its intended meaning

So how do ontologies specify semantics?

… it depends on the kind of ontology

Kinds of Ontologies

- Non-rigorous Ontologies: Meaning specified informally in natural language
- Rigorous Ontologies: meaning specified as a logical theory

In between a continuum of degree of rigor
Kinds of Ontologies

Terms

Thesauri

formal Taxonomies

Frames (OKBC)

Data Models (UML, STEP)

XML Schemas, & Data Models

Description Logics (DAML+OIL)

General Logic

Glossaries & Data Dictionaries

The GALEN ontology

The FMA – Foundational Model of Anatomy

Top-level ontologies

- DOLCE
- BFO

Description Logics (DAML+OIL)

3.2 Definitions

Mereological Definitions

(D14) \( PP(x, y) := PP(x, z) \land PP(z, y) \iff PP(x, y) \)

(D15) \( OP(x, y) := OP(x, z) \land OP(z, y) \iff OP(x, y) \)

(D16) \( AB(x, y) := AB(x, z) \land AB(z, y) \iff AB(x, y) \)

(D17) \( AF(x, y) := AF(x, z) \land AF(z, y) \iff AF(x, y) \)

(D18) \( x \in y \iff \exists z_{0} (z_{0} \in y) \land x \in z_{0} \land z_{0} \not \in y \)

(D19) \( x \in y \iff \exists z_{0} (z_{0} \in y) \land x \in z_{0} \land z_{0} \not \in y \)

(D20) \( TP(x, y) := TP(x, z) \land TP(z, y) \iff TP(x, y) \)

(D21) \( TO(x, y) := TO(x, z) \land TO(z, y) \iff TO(x, y) \)

(D22) \( TA(x, y) := TA(x, z) \land TA(z, y) \iff TA(x, y) \)

(Mereological Quality)

(D23) \( Q(x, y) := Q(x, z) \land Q(z, y) \iff Q(x, y) \)

(D24) \( Q(x, y) := Q(x, z) \land Q(z, y) \iff Q(x, y) \)

(D25) \( Q(x, y) := Q(x, z) \land Q(z, y) \iff Q(x, y) \)

(D26) \( Q(x, y) := Q(x, z) \land Q(z, y) \iff Q(x, y) \)

(D27) \( Q(x, y) := Q(x, z) \land Q(z, y) \iff Q(x, y) \)

Temporal and Spatial Qual conditions

(D28) \( Q(x, y) := Q(x, z) \land Q(z, y) \iff Q(x, y) \)

(D29) \( Q(x, y) := Q(x, z) \land Q(z, y) \iff Q(x, y) \)
**Domain ontologies** specify the semantics of the vocabulary of a particular domain:

- Biology:
  - ontology of biological organisms
  - ontology of ecosystems
  - ontology of biological processes

- Medicine:
  - ontology of diseases
  - ontology of human body parts (anatomy)
  - ontology of medical processes

Different domain ontologies use **different vocabularies** for structuring data

**Medicine:**
- Ontology of human body parts
  - heart, blood, pelvis, …
- Ontology of body processes
  - digestion, blood circulation, …
- Ontology of diseases:
  - Cancer, fracture, …

**Biology**
- Ontology of biological processes
  - Photosynthesis, …
- Ontology of biological organisms
  - Mammal, human being…
But all domain ontologies use the **same top level vocabulary** to define their domain-specific terms.

### Logical properties of the top-level vocabulary are specified in a Formal Ontology.

- Axiomatic logical theory
- Examples:
  - DOLCE developed by Nicola Guarino’s group in Italy
  - BFO developed at IFOMIS in Saarbruecken, National Center for Ontology Research (NCOR) in Buffalo

### Logical properties of top-level terms determine their semantics

- Specifying logical properties using axioms is the only way to specify semantics in a way that is accessible to a computer program
- Therefore critical for Interoperability and the Semantic Web

---

All domain ontologies use the **same top level notions**

![Diagram showing top-level ontology and domain ontology relationships]
All domain ontologies use the same top level notions

Facilitates interoperability

Kinds of Ontologies

Terms
Thesauri
Structured Thesauri
Proper Thesauri
Oxford Compendium
Principled Hierarchies
Principled Hierarchies (Yahoo)
"Ordinary" Hierarchies
Hierarchies
Glossaries
XML DTDs
XML Schemas
Data Models (UML, STEP)
Data Models (OKBC)
General Logic

We need the resources of full general logic to formalize top-level ontologies

Description logics may often be sufficient for domain ontologies

Three kinds of spatial entities

1. Independent entities (substances)
2. Dependents entities
3. Immaterial entities (Holes, flat boundaries)
4. Processes
5. Universals

Top-level ontologies

Basic categories
Substances

Do not depend (ontologically) on other entities

Complete bona fide boundary

Undetached (fiat) parts

- arms, legs, noses
- no natural completeness

Undetached (fiat) parts

- arms, legs, noses
- no natural completeness

Fiat boundary
Mount Blanc: Substance or non-detached part?

<table>
<thead>
<tr>
<th>Mountain</th>
</tr>
</thead>
<tbody>
<tr>
<td>bona fide upper boundaries</td>
</tr>
<tr>
<td>with a fiat base</td>
</tr>
</tbody>
</table>

| A nose has fiat boundary parts                |

<table>
<thead>
<tr>
<th>Aggregates of substances</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Aggregates:</td>
</tr>
<tr>
<td>- Your Family,</td>
</tr>
<tr>
<td>- the New York Philharmonic Orchestra</td>
</tr>
<tr>
<td>- No single connected boundary</td>
</tr>
<tr>
<td>- separated</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Three kinds of endurants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Substances</td>
</tr>
<tr>
<td>2. Dependents entities</td>
</tr>
<tr>
<td>3. Dependent spatial entities</td>
</tr>
</tbody>
</table>
Dependence

**a thought**

**Dependent endurant**

cannot exist without a thinker

**independent endurant**

(substance)

Dependent endurants:

- **one-place**:
  - your temperature, color, height
  - my knowledge of French
  - the whiteness of this cheese
  - the warmth of this stone
  - the fragility of this glass

Ontological Dependence

- **process → substance**
  - The erosion of the rock necessitates the existence of the rock

- **quality → substance**
  - The token *redness of the sand* necessitates the existence of the sand

- **state → substance**
  - the state of being pregnant necessitates the existence of someone being pregnant

Three kinds of endurants

1. Substances
2. Dependents entities
3. Imaterial spatial entities (Holes, flat boundaries, …)

Shadows are immaterial entities
What does *immaterial* mean? ???

- Do not ‘own’ or occupy the region at which they are located
- Can share their region
- Cannot have material parts

**Positive and negative parts**

- positive part (made of matter)
- negative part (not made of matter, immaterial)

**Holes are dependent entities**

- a hole requires a host

**Artifacts often have holes**

**Niches, environments are holes**
Places are holes

Holes of your body

Endurants vs. perdurants

Persistent entities

Endurants
- Persists through time in virtue of being **wholly present** at every time at which it exists at all.
  - I exist in full at this moment in time

Perdurants
- **Evolve** over time
- Do not exist in full at any given moment
  - This presentation does not exist in full right now

Endurants have spatial parts

Perdurants (processes) have spatio-temporal parts
Endurants do NOT have spatio-temporal parts

- I am an endurant
- The first 5-minute phase of my existence is not a spatio-temporal part of me
- It is a spatio-temporal part of my life
- My life is a process
Geographic perdurants

The movement of the water molecules which constitute the cold body of water is a perdurant (process)

Geographic entities

The weakening of the trade winds is a perdurant (process)

Universals and particulars

Two distinct kinds of entities

Universals		| particulars
---|---
types	| tokens
classes	| instances

The semantics of partonomic inclusion

\[ X \text{ PART-OF } Y \text{ iff } \]
For every instance \( x \) of \( X \) there exists an instance \( y \) of \( Y \) such that \( x \) part-of \( y \)

\[ \text{AND} \]
For every instance \( y \) of \( Y \) there exists an instance \( x \) of \( X \) such that \( x \) part-of \( y \)
Partonomic inclusion

Human-heads

Part-of

Part-of

Human-bodies

Partonomic inclusion

Human-heads

HasPart

HasPart

Human-bodies

Universals among dependent endurants

Is-a

quality

red

scarlet

R232, G54, B24

Inst-Of

this individual token of redness

(this token redness – here, now)

Barry Smith

Types and tokens

Of independent endurants

Of dependent endurants

Accidents

Species and instances

quality

color

red

scarlet

R232, G54, B24

this individual token of redness

(this token redness – here, now)

Barry Smith

Aristotle’s Ontological Square

<table>
<thead>
<tr>
<th>Substantial</th>
<th>Dependent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal</td>
<td></td>
</tr>
<tr>
<td>Second substance</td>
<td>Second accident</td>
</tr>
<tr>
<td>man</td>
<td>headache</td>
</tr>
<tr>
<td>cat</td>
<td>sun-tan</td>
</tr>
<tr>
<td>ox</td>
<td>red</td>
</tr>
<tr>
<td>Particular</td>
<td></td>
</tr>
<tr>
<td>First substance</td>
<td>First accident</td>
</tr>
<tr>
<td>this man</td>
<td>this headache</td>
</tr>
<tr>
<td>this cat</td>
<td>this sun-tan</td>
</tr>
<tr>
<td>this ox</td>
<td>this red spot</td>
</tr>
</tbody>
</table>

Barry Smith

Conclusions
The future: perfect information integration

Compatible data formats
Compatible semantics

All domain ontologies are built based on a sound top-level ontology

Kinds of Ontologies

Non-ontologies
Ontologies

Aristotle’s Ontological Square

<table>
<thead>
<tr>
<th></th>
<th>Substantial</th>
<th>Dependent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universal</td>
<td>Second substance</td>
<td>Second accident</td>
</tr>
<tr>
<td></td>
<td>man</td>
<td>headache</td>
</tr>
<tr>
<td></td>
<td>cat</td>
<td>sun-tan</td>
</tr>
<tr>
<td></td>
<td>ox</td>
<td>red</td>
</tr>
<tr>
<td>Particular</td>
<td>First substance</td>
<td>First accident</td>
</tr>
<tr>
<td></td>
<td>this man</td>
<td>this headache</td>
</tr>
<tr>
<td></td>
<td>this cat</td>
<td>this sun-tan</td>
</tr>
<tr>
<td></td>
<td>this ox</td>
<td>this red spot</td>
</tr>
</tbody>
</table>