Ontology-Driven Conceptual Modeling

Chris Welty
IBM Watson Research Center

Acknowledgements

People	Organizations
	Vassar College, USA
Nicola Guarino	LADSEB-CNR, Padova
Cladio Masolo	CNR Cognitive Science
Aldo Gangemi	Institute, Trento
Alessandro Oltramari	
Bill Andersen	OntologyWorks, Inc.

Outline

- Setting the record straight 
- Motivation
- Formal foundation
- "Upper Level" distinctions
- Common pitfalls

What is Ontology?

- A discipline of Philosophy
  - *Meta-physics* dates back to Artistotle
    - Meta (after) + physica (physical, real)
  - *Ontology* dates back to 17th century
    - Ontos (that which exists) + logos (knowledge of)
    - As in TorONTO, ONTario, ON TOp
  - The science of what *is* (in the universe)
  - "One universe, One ontology"
- Quine, 1969:
  "To exist is to be the value of a quantified variable"

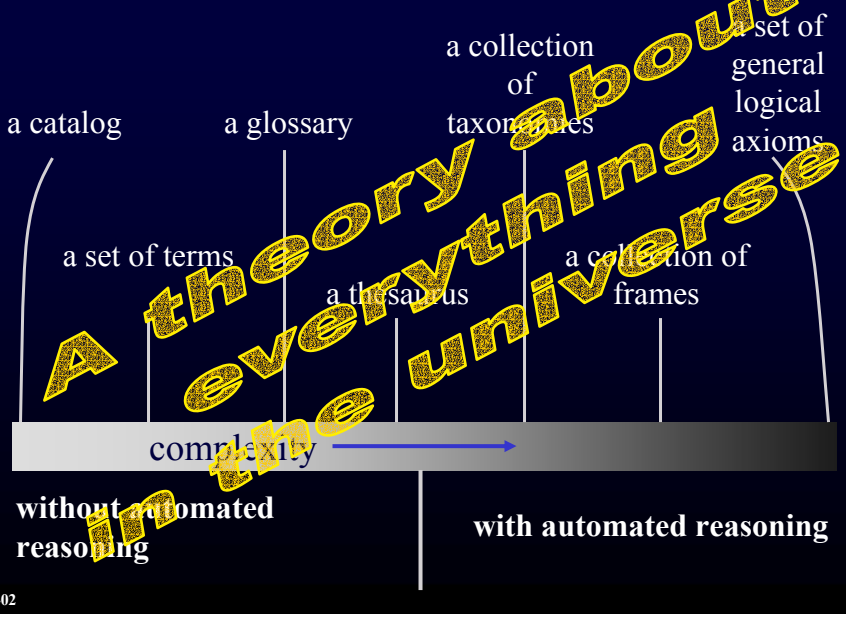
What is Ontology?

- Borrowed by AI community
  - McCarthy (1980) calls for "a list of things that exist"
  - Specify all the *kinds of things* that can be the values of variables
- Evolution of meaning in CS
  - Now refers to domain modeling, conceptual modeling, knowledge engineering, etc.
- Note: not a "new name for an old thing"

What is *an* Ontology?

- Poor definition:
  "Specification of a conceptualization" [Gruber, 1993]
- Better:
  "Description of the kinds of entities there are and how they are related."
- Good ontologies should provide:
  - **Meaning**
  - Organization
  - Taxonomy
  - Agreement
  - Common Understanding
  - Vocabulary
  - Connection to the "real world"

What is *an* Ontology?

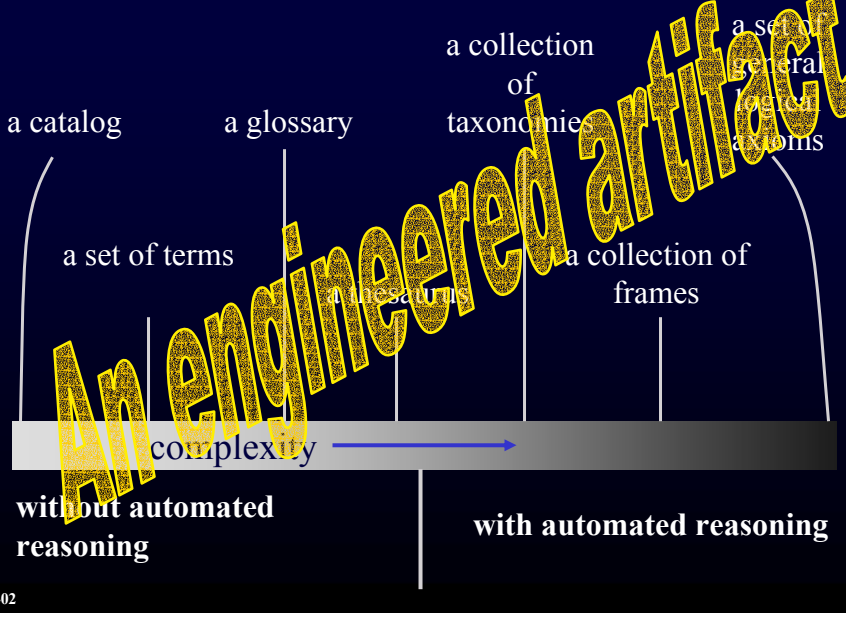


A theory about everything in the universe

without automated reasoning

with automated reasoning

What is *an* Ontology?



An engineered artifact

without automated reasoning

with automated reasoning

Key Challenges

- Must build/design, analyze/evaluate, maintain/extend, and integrate/reconcile ontologies
- Little guidance on how to do this
  - In spite of the pursuit of many *syntactic* standards
  - Where do we *start* when building an ontology?
  - What criteria do we use to *evaluate* ontologies?
  - How are ontologies *extended*?
  - How are different ontological choices *reconciled*?
- Ontological Modeling and Analysis
  - Does your model mean what you intend?
  - Will it produce the right results?

Outline

- Setting the record straight
- Motivation 
- Formal foundation
- "Upper Level" distinctions
- Common pitfalls

Motivation

Provide a sound basis for analyzing ontological decisions

"If you can give me a way to shorten the length of the arguments I have with these doctors, you have made a significant contribution…"

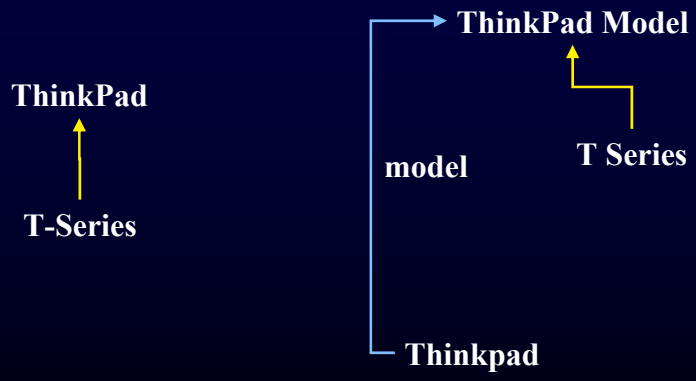
-Alan Rector

Most ontology efforts *fail*

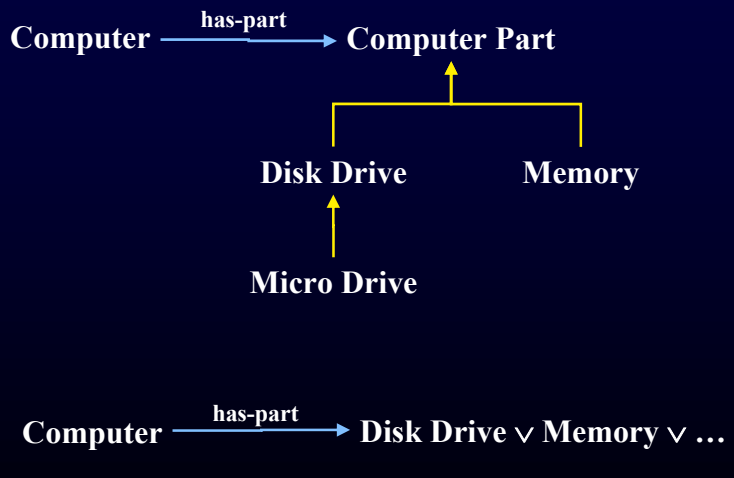
- Why?
  - The quality of the ontology dictates its impact
  - Poor ontology, poor results
  - Ontologies are built by people

…The average IQ is 100

Which one is better?



Which one is better?



Contributions

- Methodology to help analyze & build consistent ontologies
  - Formal foundation of ontological analysis
  - Meta-properties for analysis
  - "Upper Level" distinctions
    - Standard set of upper-level concepts
    - Standardizing semantics of ontological relations
- Common ontological modeling pitfalls
  - Misuse of intended semantics
- Specific recent work focused on clarifying the subsumption (is-a, subclass) relation

Outline

- Setting the record straight
- Motivation
- Formal foundation 
- "Upper Level" distinctions
- Common pitfalls

Approach

- Draw ***fundamental notions*** from Philosophy
- Establish a set of useful ***meta-properties***, based on behavior wrt above notions
- Explore the way these meta-properties combine to form relevant ***property kinds***
- Explore the ***constraints*** imposed by these property kinds.

Basic Philosophical Notions

(taken from *Formal Ontology*)

- Identity
  - How are instances of a class distinguished from each other
- Unity
  - How are all the parts of an instance isolated
- Essence
  - Can a property change over time
- Dependence
  - Can an entity exist without some others

Essence and Rigidity

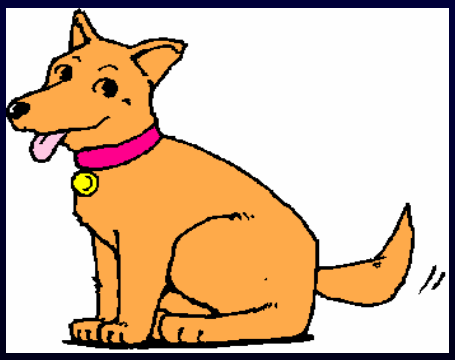
- Certain entities have ***essential*** properties.
  - Hammers must be hard.
  - John must be a person.
- Certain properties are essential to ***all*** their instances (compare *being a person* with *being hard*).
- These properties are ***rigid*** - if an entity is ever an instance of a rigid property, it must always be.

Formal Rigidity

- ϕ is rigid (+R): $\forall x \blacklozenge\phi(x) \rightarrow \Box \phi(x)$
  - e.g. Person, Apple
- ϕ is non-rigid (-R): $\exists x \ \phi(x) \wedge \neg \Box \phi(x)$
  - e.g. Red, Male
- ϕ is anti-rigid (~R): $\forall x \blacklozenge\phi(x) \rightarrow \neg \Box \phi(x)$
  - e.g. Student, Agent

Identity and Unity

- Identity: is this my dog?
- Unity: is the collar part of my dog?



Identity criteria

- ***Classical formulation:***

$$\phi(x) \wedge \phi(y) \rightarrow (\rho(x,y) \leftrightarrow x = y)$$

- ***Generalization:***

$$\phi(x,t) \wedge \phi(y,t') \rightarrow (\Gamma(x,y,t,t') \leftrightarrow x = y)$$

(synchronic: $t = t'$; diachronic: $t \neq t'$)

- In most cases, Γ is based on the ***sameness*** of certain ***characteristic features***:

$$\Gamma(x,y,\ t,t') = \forall z\ (\chi(x,z,t) \wedge \chi(y,z,t'))$$

A Stronger Notion: Global ICs

- ***Local*** IC:

  $\phi(x,t) \wedge \phi(y,t') \rightarrow (\Gamma(x,y,t,t') \leftrightarrow x = y)$

- ***Global*** IC ***(rigid properties only)***:

  $\phi(x,t) \rightarrow (\phi(y,t') \wedge \Gamma(x,y,t,t') \leftrightarrow x = y)$

Identity meta-properties

- ***Supplying*** (global) identity (+**O**)
  - Having some "own" IC that doesn't hold for a subsuming property
- ***Carrying*** (global) identity (+**I**)
  - Having an IC (either own or inherited)
- ***Not carrying*** (global) identity (**-I**)

Unity Criteria

- An object x ***is a whole*** under ω iff ω is an equivalence relation that binds together all the parts of x, such that

$$P(y,x) \rightarrow (P(z,x) \leftrightarrow \omega(y,z))$$

but not

$$\omega(y,z) \leftrightarrow \exists x(P(y,x) \wedge P(z,x))$$

- P is the ***part-of*** relation
- ω can be seen as a ***generalized indirect connection***

Unity Meta-Properties

- If all instances of a property ϕ are wholes under ***the same*** relation, ϕ carries unity (**+U**)
- When at least one instance of ϕ is not a whole, or when two instances of ϕ are wholes under ***different*** relations, ϕ does not carry unity (**-U**)
- When no instance of ϕ is a whole, ϕ carries anti-unity (**~U**)

Property Dependence

- Does a property holding for *x* depend on something else besides *x*? (***property dependence***)
  - $P(x) \rightarrow \exists y\ Q(y)$
  - *y* should not be a part of *x*
- Example: Student/Teacher, customer/vendor

Outline

- Setting the record straight
- Motivation
- Formal foundation
- "Upper Level" distinctions 
- Common pitfalls

"Upper Level" Ontology

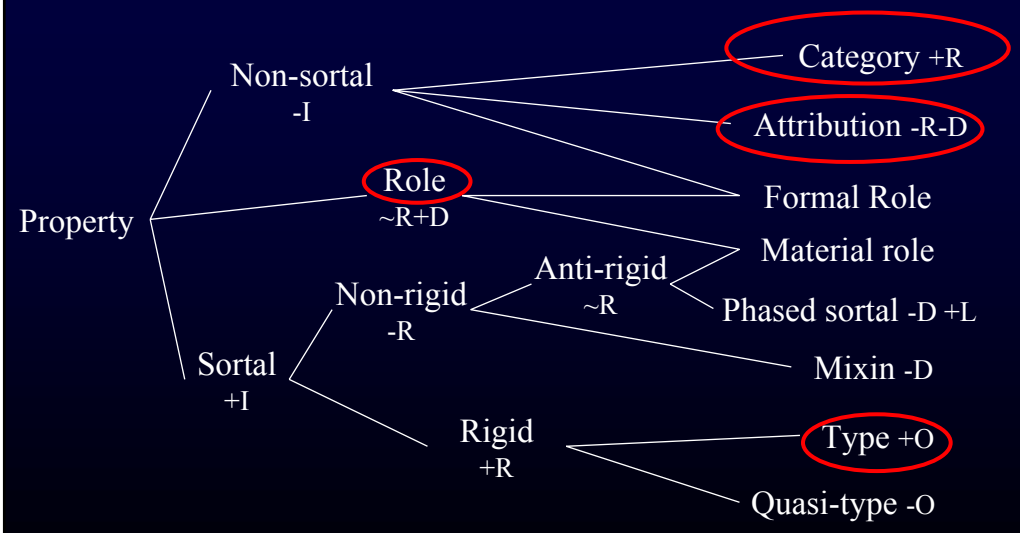
- The "media independent" knowledge
  - Fundamental truths of the universe
  - Non contextual (aka *formal*)
- Is there only one?
- Upper level ≠ Large
- Proven value
  - A place to start
  - Semantic integration

Upper Level

Where do I start?

- Particulars
  - Concrete
    - Location, event, object, substance, …
  - Abstract
    - information, story, collection, …
- Universals
  - Property (Class)
  - Relation
    - Subsumption (subclass), instantiation, constitution, composition (part)

A formal ontology of properties



Sortals, categories, and other properties

- Sortals (*horse, triangle, amount of matter, person, student...)*
  - Carry identity
  - Usually correspond to ***nouns***
  - High organizational utility
  - Main subclasses: ***types*** and ***roles***
- Categories *(universal, particular, event, substance...)*
  - No identity
  - Useful generalizations for sortals
  - Characterized by a set of (only necessary) formal properties
  - Good organizational utility
- Other non-sortals *(red, big, decomposable, eatable, dependent, singular...)*
  - No identity
  - Correspond to ***adjectives***
  - Span across different sortals
  - Limited organizational utility (but high semantic value)

Formal Ontology of Relations

- Subsumption
- Instantiation
- Part/Whole
- Constitution
- Spatial (Cohn)
- Temporal (Allen)

Subsumption

- The most pervasive relationship in ontologies
  - Influence of taxonomies and OO
- AKA: Is-a, a-kind-of, specialization-of, subclass (Brachman, 1983)
  - "horse is a mammal"
- Capitalizes on general knowledge
  - Helps deal with complexity, structure
  - Reduces requirement to acquire and represent redundant specifics
- What does it *mean*?

$$\Box\ \forall x\ \phi(x) \rightarrow \rho(x)$$

Every instance of the subclass is necessarily an instance of the superclass

The *Backbone Taxonomy*

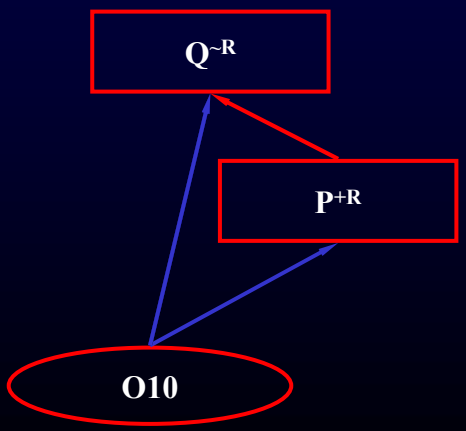
Assumption: ***no entity without identity***
Quine, 1969

- Since identity is supplied by types, every entity must instantiate a type
- The taxonomy of types spans the whole domain
- Together with categories, types form the ***backbone taxonomy,*** which represents the ***invariant structure*** of a domain (rigid properties spanning the whole domain)

Rigidity Constraint

$+R \not\subset \sim R$

- Why?
  - $\forall x\, P(x) \rightarrow Q(x)$



Identity Conditions along Taxonomies

- Adding ICs:
  - Polygon: same edges, same angles
    - Triangle: two edges, one angle
      - **Equilateral triangle: one edge**
- Just inheriting ICs:
  - Person
    - Student

Identity Disjointness Constraint

Besides being used for recognizing sortals, ICs impose ***constraints*** on them, making their ontological nature explicit:

Properties with incompatible ICs are ***disjoint***

Examples:
- sets vs. ordered sets
- amounts of matter vs. assemblies

Unity Disjointness Constraint

Properties with incompatible UCs are ***disjoint***
+U ⊄ ~U

Taxonomic Constraints

- +R ⊄ ~R
- -I ⊄ +I
- -U ⊄ +U
- +U ⊄ ~U
- -D ⊄ +D

- Incompatible IC's are disjoint
- Incompatible UC's are disjoint

- Categories subsume everything
- Roles can't subsume types

Outline

- Setting the record straight
- Motivation
- Formal foundation
- "Upper Level" distinctions
- Common pitfalls 

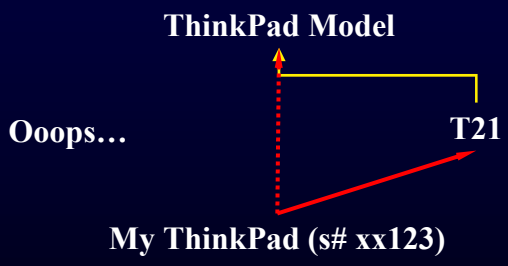
Overloading Subsumption

Common modeling pitfalls

- Instantiation
- Constitution
- Composition
- Disjunction
- Polysemy

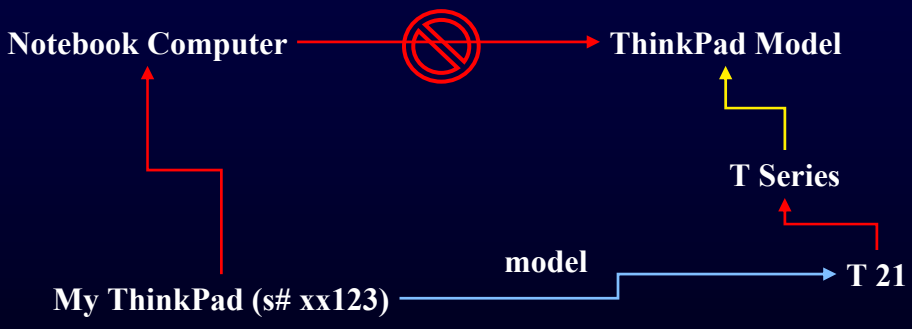
Instantiation (1)

Does this ontology mean that **My ThinkPad** *is a* **ThinkPad Model**?

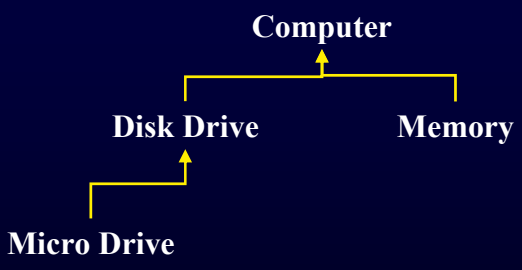


Question: What ThinkPad models do you sell?
Answer should NOT include My ThinkPad -- nor yours.

Instantiation (2)

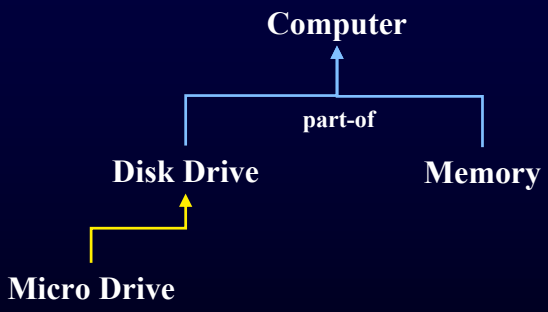


Composition (1)

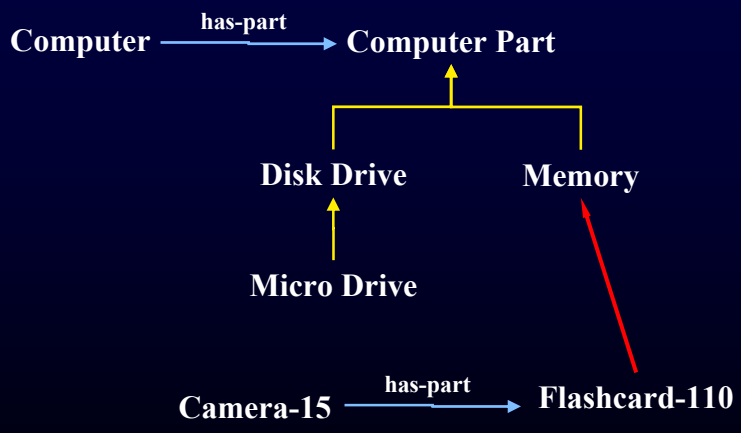


Question: What Computers do you sell?
Answer should NOT include Disk Drives or Memory.

Composition (2)



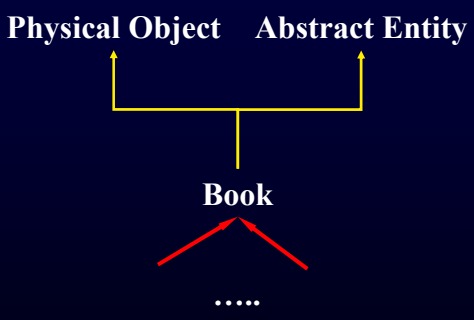
Disjunction (1)



Unintended model: flashcard-110 is a computer-part

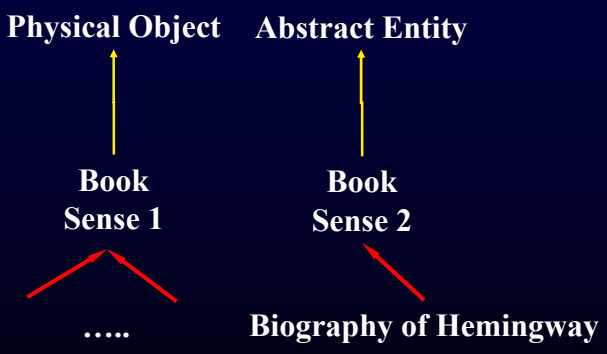
Disjunction (2)

Polysemy (1)
(Mikrokosmos)

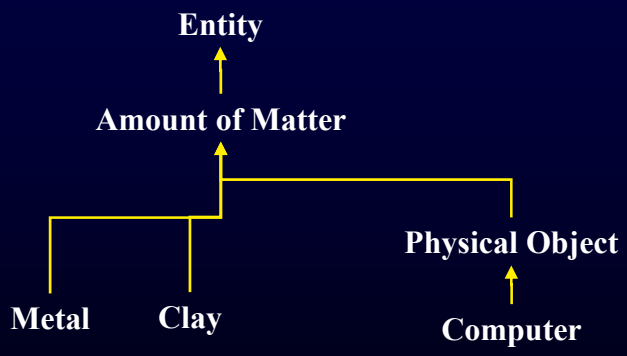


Question: How many books do you have on Hemingway?
Answer: 5,000

Polysemy (2)
(WordNet)

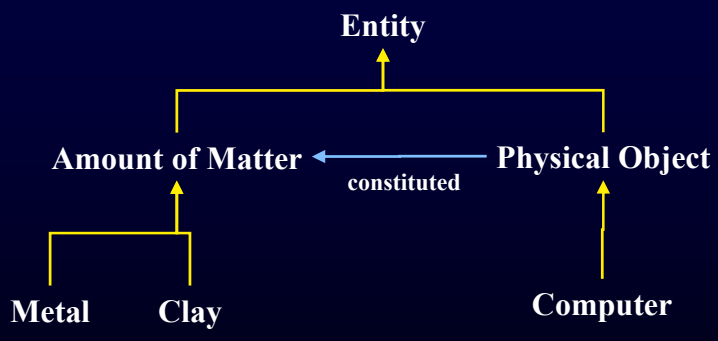


Constitution (1)
(WordNet)



Question: What types of matter will conduct electricity?
Answer should NOT include computers.

Constitution (2)



Technical Conclusions

- Subsumption is an overloaded relation
  - Influence of OO
  - Force fit of simple taxonomic structures
  - Leads to misuse of *is-a* semantics
- Ontological Analysis
  - A collection of well-defined knowledge structuring relations
  - Methodology for their consistent application
    - Meta-Properties for ontological relations
    - Provide basis for disciplined ontological analysis

Applications of Methodology

- *Ontologyworks*
- IBM
- *Ontoweb*
- *TICCA, WedODE, Galen, …*
- Strong interest from and participation in
  - Semantic web (w3c)
  - IEEE SUO
  - Wordnet
  - Lexical resources

References

- Guarino, Nicola and Chris Welty. 2002. *CACM.* **45**(2):61-65.
- Smith, Barry and Chris Welty. 2001. Ontology: Towards a new synthesis. In *Formal Ontology in Information Systems*. ACM Press.
- Welty, Chris and Nicola Guarino. 2001. In *J. Data and Knowledge Engineering.* **39**(1):51-74. October, 2001.
- Guarino, Nicola and Chris Welty. 2000. In *Proceedings of ER-2000: The 19th International Conference on Conceptual Modeling.*
- Guarino, Nicola and Chris Welty. 2000. In *Proceedings of EKAW-2000*
- Guarino, Nicola and Chris Welty. 2000. In *Proceedings of ECAI-2000: The European Conference on Artificial Intelligence.*
- Upcoming special issue of *AI Magazine* on Ontologies.