



Introduction to ontologies

Nicola Guarino
Laboratorio di Ontologia Applicata (LOA)
Istituto di Scienze e Tecnologie della Cognizione (ISTC-CNR)
Trento, Italy

1. Why ontologies: the need to focus on content

The focus of ontological analysis: from form to *CONTENT*

- The key problems
 - content-based information access (*semantic matching*)
 - content-based information integration (*semantic integration*)
- To approach them, content must be studied, understood, analyzed *as such*, independently of the way it is represented.
- Computer technologies are not really good for that (focus is usually on representation and reasoning)
- ...and users of computer systems are often *confused by technology*

Ontologies: a *magic solution*?



No ontology without *ontological analysis!*

The problem: subtle distinctions in meaning

“Trying to engage with oo many partners too fast is one of the main reasons that ***so many online market makers have foundered***.

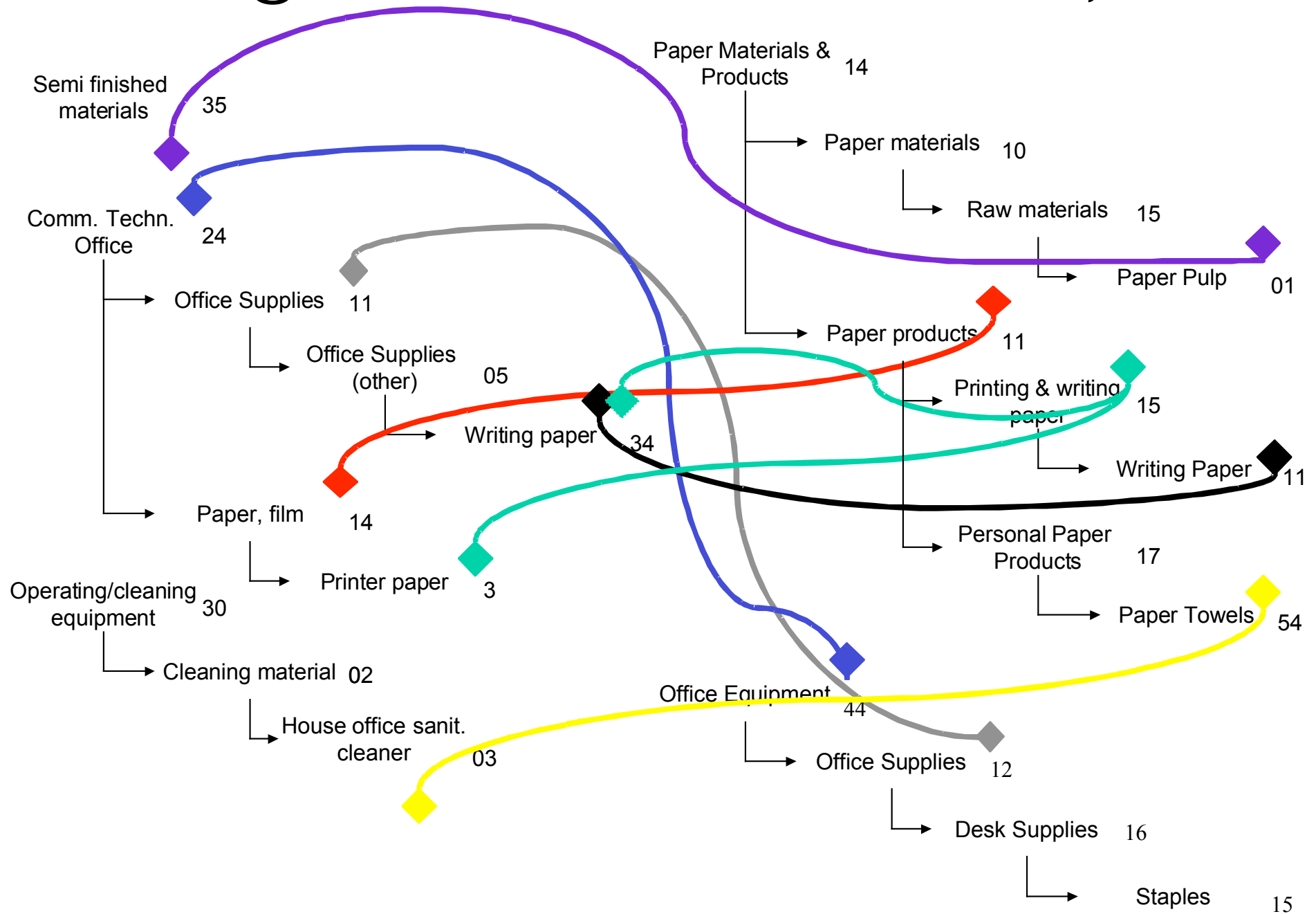
The transactions they had viewed as simple and routine actually involved many ***subtle distinctions in terminology and meaning***”

Harvard Business Review, October 2001



Ecl@ss

UNSPSC, UCEC



Subtle distinctions in meaning...

- What is an *application* to a public administration?
- What is a *working place*? an *address*?
- What is an *unemployed person*?
- What is a *customer*?
- What is a *passenger*?
- What is an *organization*?
- What is a *document*?
- What is a *contract*?
- What is a *lake*, a *river*, a *valley*?
- What is a *population*? a *species*?
- What is a *service*?



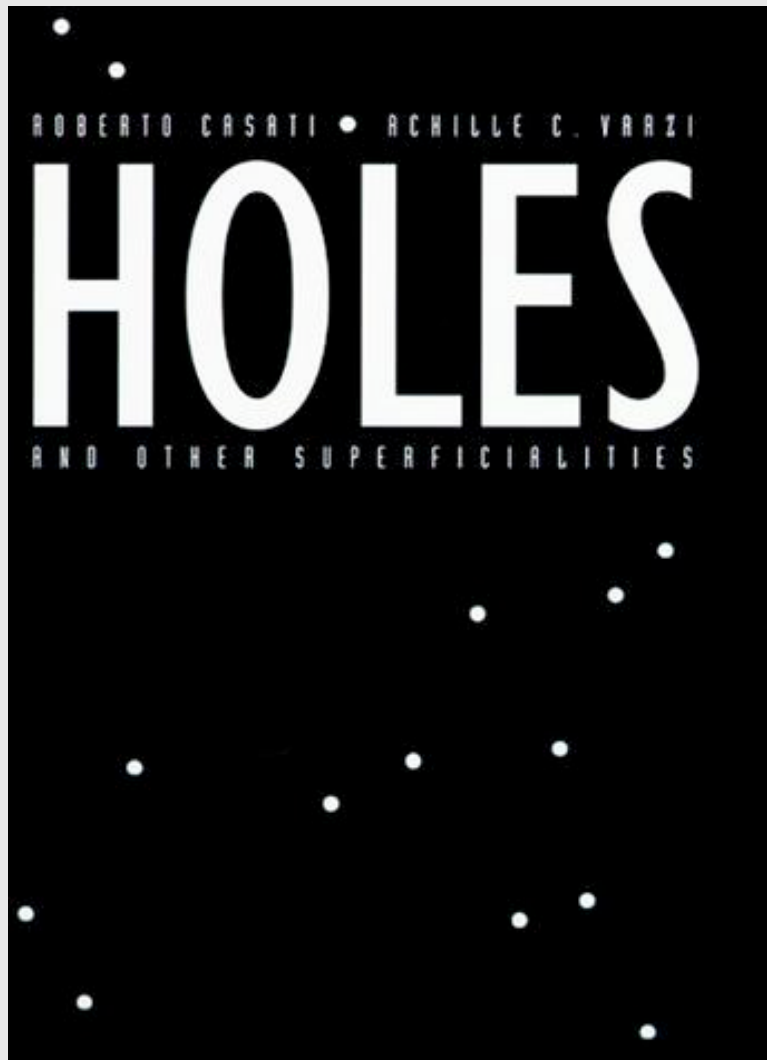
Where subtle distinctions are important

- 2000 US Presidential elections: how many *holes*?
- 2001 twin towers catastrophe: how many *events*?

...only *ontological analysis* solves these problems!!



The ontology of holes



Books by Roberto Casati and Achille C. Varzi (MIT Press):

- Holes and other superficialities
- Parts and places



A common alphabet is not enough...

- “XML is only the first step to ensuring that computers can communicate freely. ***XML is an alphabet for computers*** and as everyone who travels in Europe knows, knowing the alphabet doesn't mean you can speak Italian or French”

Business Week, March 18, 2002



Standard glossaries can help, but...

- Defining standard vocabularies is *difficult and time-consuming*
- Once defined, standards *don't adapt well*
- Heterogeneous domains need a *broad-coverage vocabulary*
- People don't implement standards correctly anyway
- Vocabulary definitions are often *ambiguous or circular*



Do we know what to REpresent?

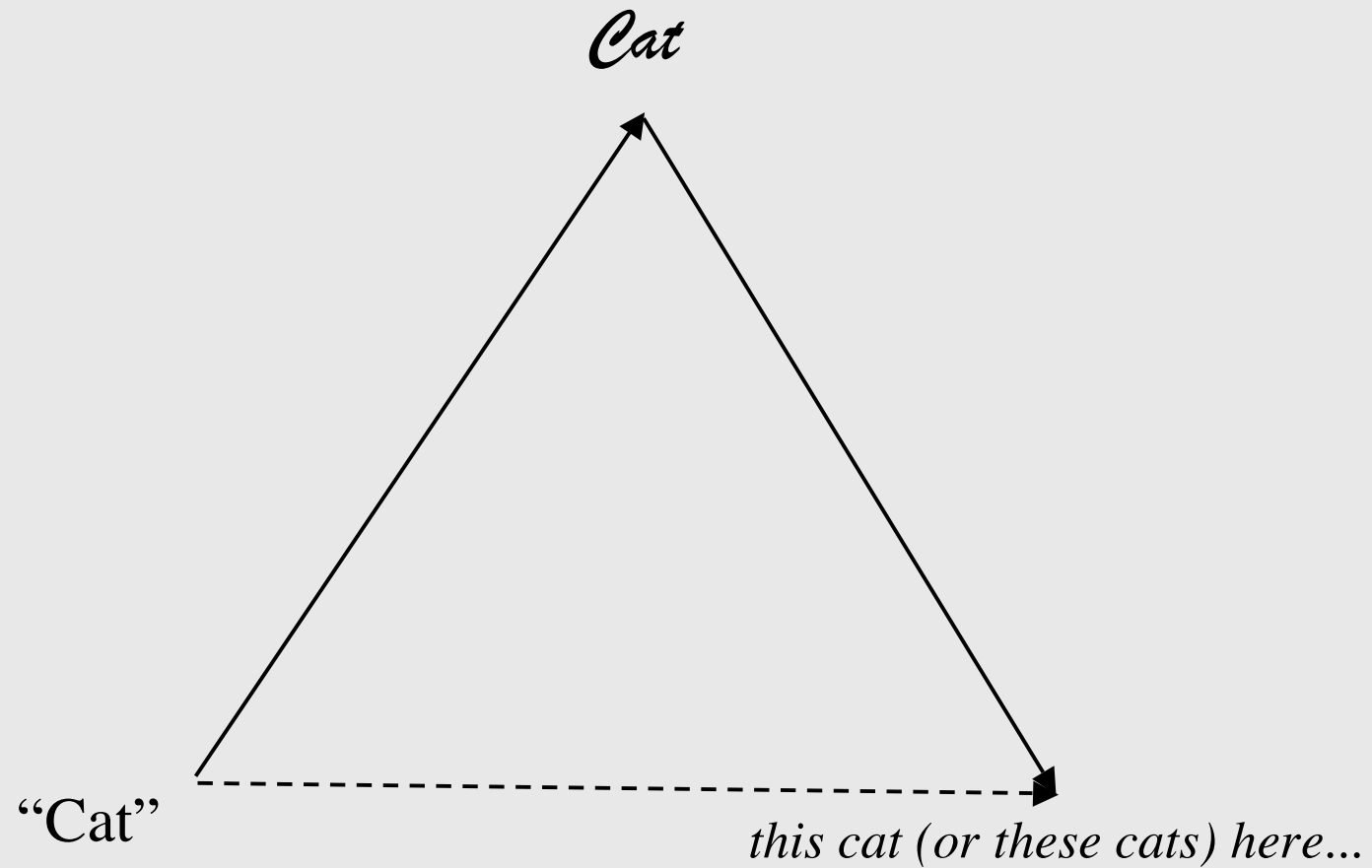
- ***First*** ontological analysis,
- ***THEN*** knowledge representation...

Unfortunately, this is not the current practice...

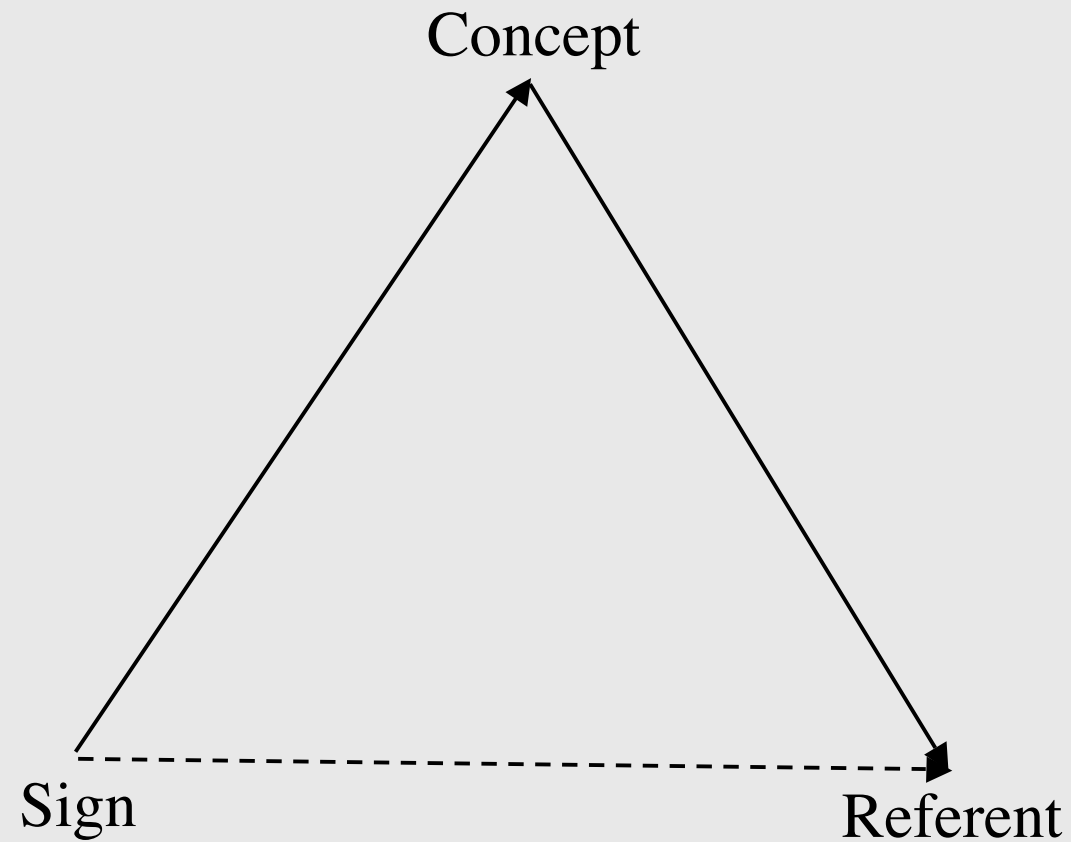


2. Meanings and signs

The triangle of meaning - 1



The triangle of meaning - 2



Intension ed extension

- Intension (concept): part of meaning corresponding to general principles, rules to be used to determine reference (typically, abstractions from experience)
- Extension (object): part of meaning corresponding to the effective reference
- Only by means of the **concept** associated to the **sign** “cat” we can correctly **interpret** this sign in various **situations**
- The sign’s referent is the result of this interpretation
- Such interpretation is a **situated intentional act**



An example: the concept of *red*

a b



→ {a}



→ {b}



→ {a,b}



→ {}



Concepts, properties, and relations

- By means of concepts, we ***ascribe*** properties and relations to things. We can say that concepts *describe* properties or relations.
- Concepts describing relations are also called ***conceptual relations***:
 - *friend-of, father-of...*
- Conceptual relations are NOT sets of tuples! Their *extension* is a set of tuples.



Relations vs. Conceptual Relations

ordinary relations are defined on a *domain* D :

$$r_n \in 2^{D^n}$$

conceptual relations are defined on a *domain space* $\langle D, W \rangle$

$$\rho_n : W \rightarrow 2^{D^n} \quad (\text{Montague's intensional logic})$$



3. What is an ontology

Ontology, lexicon, semantics

- Distinctions among contents: *Ontology*
- Reference to content: *Lexicon*, via *Semantics*
- Every organization, every computer system
 - Makes (implicit) ontologic assumptions
 - Adopt a certain lexicon, to which an *intended semantics* is ascribed.



Ontology and Ontologies

- **Ontology**: the philosophical discipline
 - Study of **what there is**
(content *qua* content, even independently of its existence...)
 - Study of the **nature** and **structure** of “reality”
- **ontologies**:

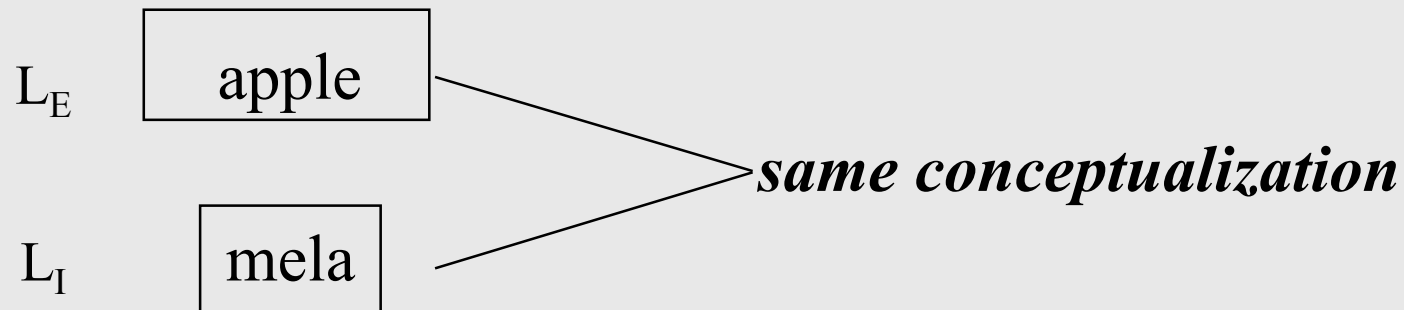
Specific (theoretical or computational) artifacts expressing the **intended meaning** of a **vocabulary** in terms of **primitive** categories and relations describing the **nature** and **structure** of a **domain of discourse**

Gruber: “Explicit and formal specifications of a **conceptualization**”



What is a conceptualization

- Formal structure of (a piece of) reality *as perceived and organized by an agent, independently of*:
 - the **vocabulary** used
 - the actual occurrence of a specific **situation**
- Different situations involving same objects, described by different vocabularies, may share the same conceptualization.



What is a *conceptualization*?

A cognitive approach

- Humans isolate **relevant invariances** from physical reality (quality distributions) on the basis of:
 - Perception (as resulting from evolution)
 - Cognition and cultural experience
 - Language
- A set of **atomic stimuli** (input pattern) is associated to each world's moment
- Synchronic level: **spatial invariants**
 - Unity properties are **ascribed** to input patterns: topological and morphological wholes (**percepts**) emerge
- Diachronic level: **temporal invariants**
 - **Objects**: equivalence relationships among **percepts** belonging to different moments
 - **Events**: unity properties are ascribed to **percept sequences** belonging to different moments



A *conceptualization* for D is a tuple $\mathbf{C} = \langle D, W, \mathfrak{R} \rangle$, where \mathfrak{R} is a set of conceptual relations on $\langle D, W \rangle$

A *model* for a language L with vocabulary V is a structure

$\langle \mathbf{S}, I \rangle$, where $\mathbf{S} = \langle D, \mathbf{R} \rangle$ is a *world structure* and $I: V \rightarrow D \cup \mathbf{R}$ is the usual interpretation function.

A model fixes a particular extensional interpretation of the language. Analogously, we can fix an *intensional* interpretation by means of a structure

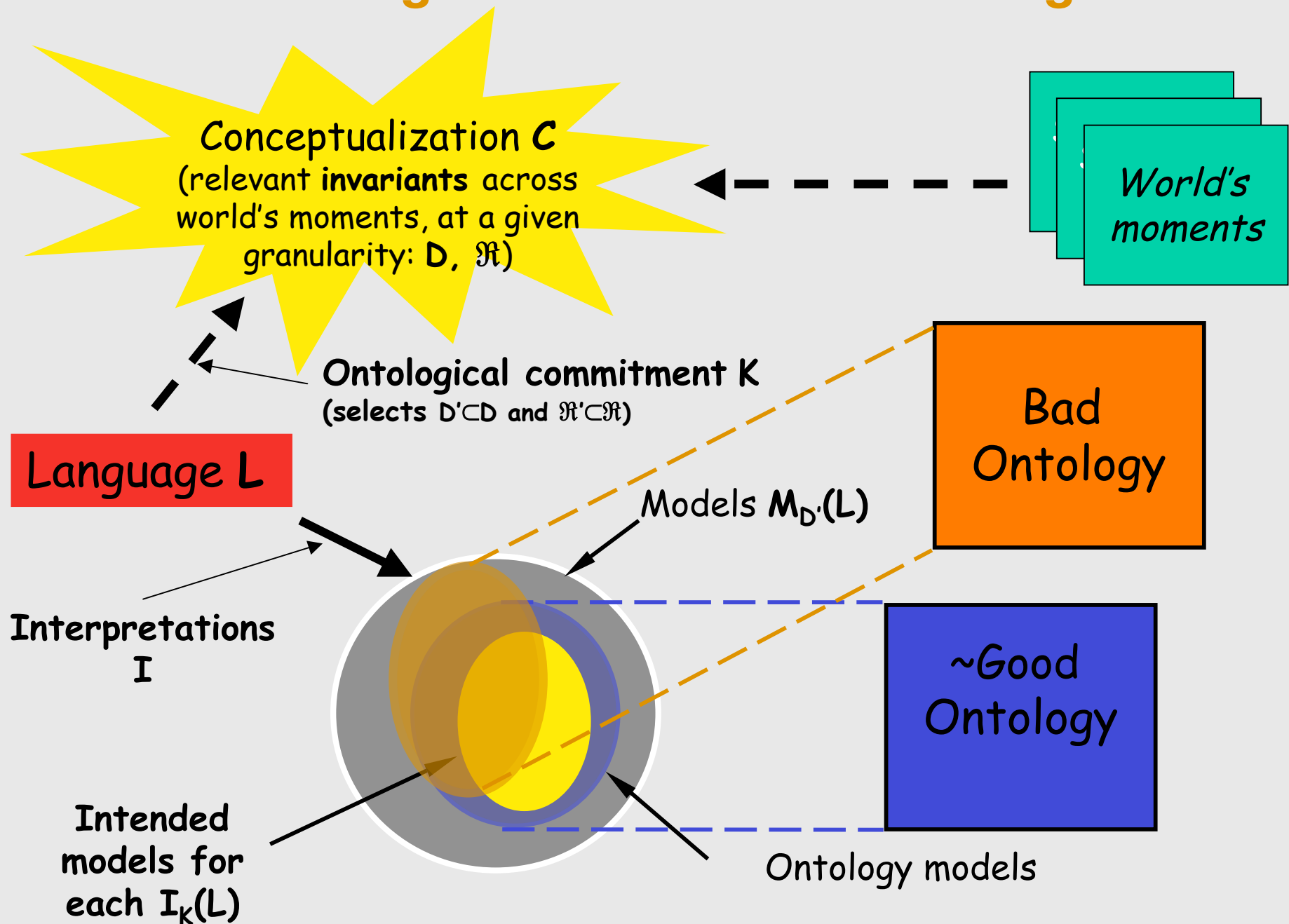
$\langle \mathbf{C}, \mathfrak{I} \rangle$, where $\mathbf{C} = \langle D, W, \mathfrak{R} \rangle$ is a conceptualization and $\mathfrak{I}: V \rightarrow D \cup \mathfrak{R}$ is an *intensional interpretation function*.

We call such a structure $\mathbf{K} = \langle \mathbf{C}, \mathfrak{I} \rangle$ an *ontological commitment* for L .

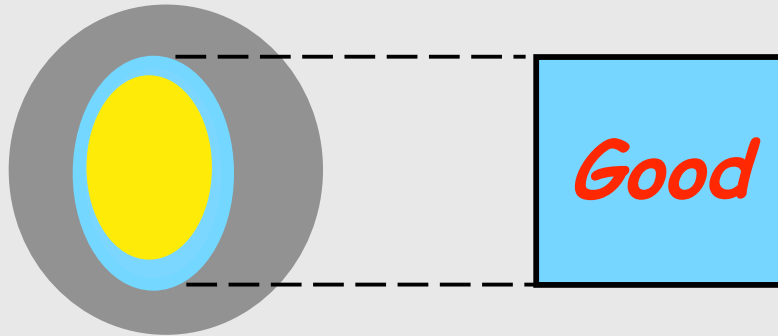
L *commits* to \mathbf{C} by means of \mathbf{K} .

\mathbf{C} is the *underlying conceptualization* of \mathbf{K} .

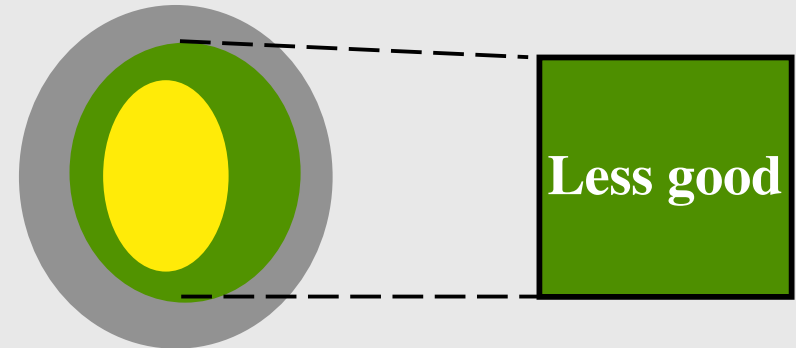
Ontologies and intended meaning



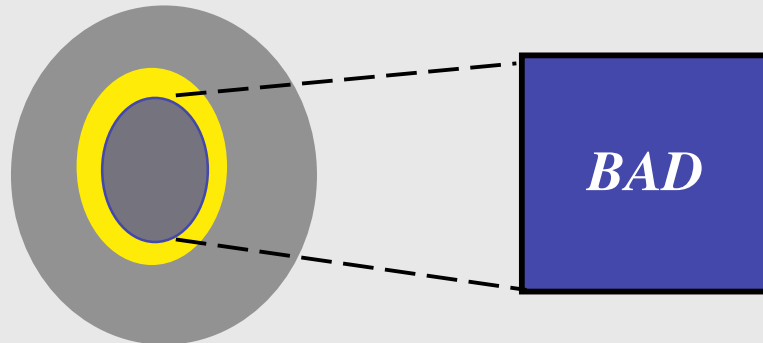
Ontology Quality: Precision and Coverage



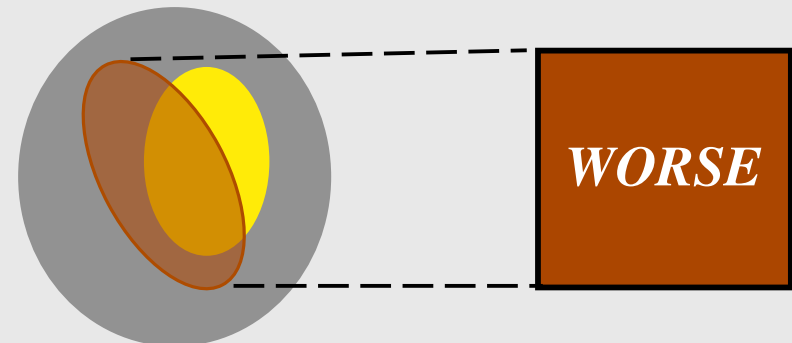
High precision, max coverage



Low precision, max coverage

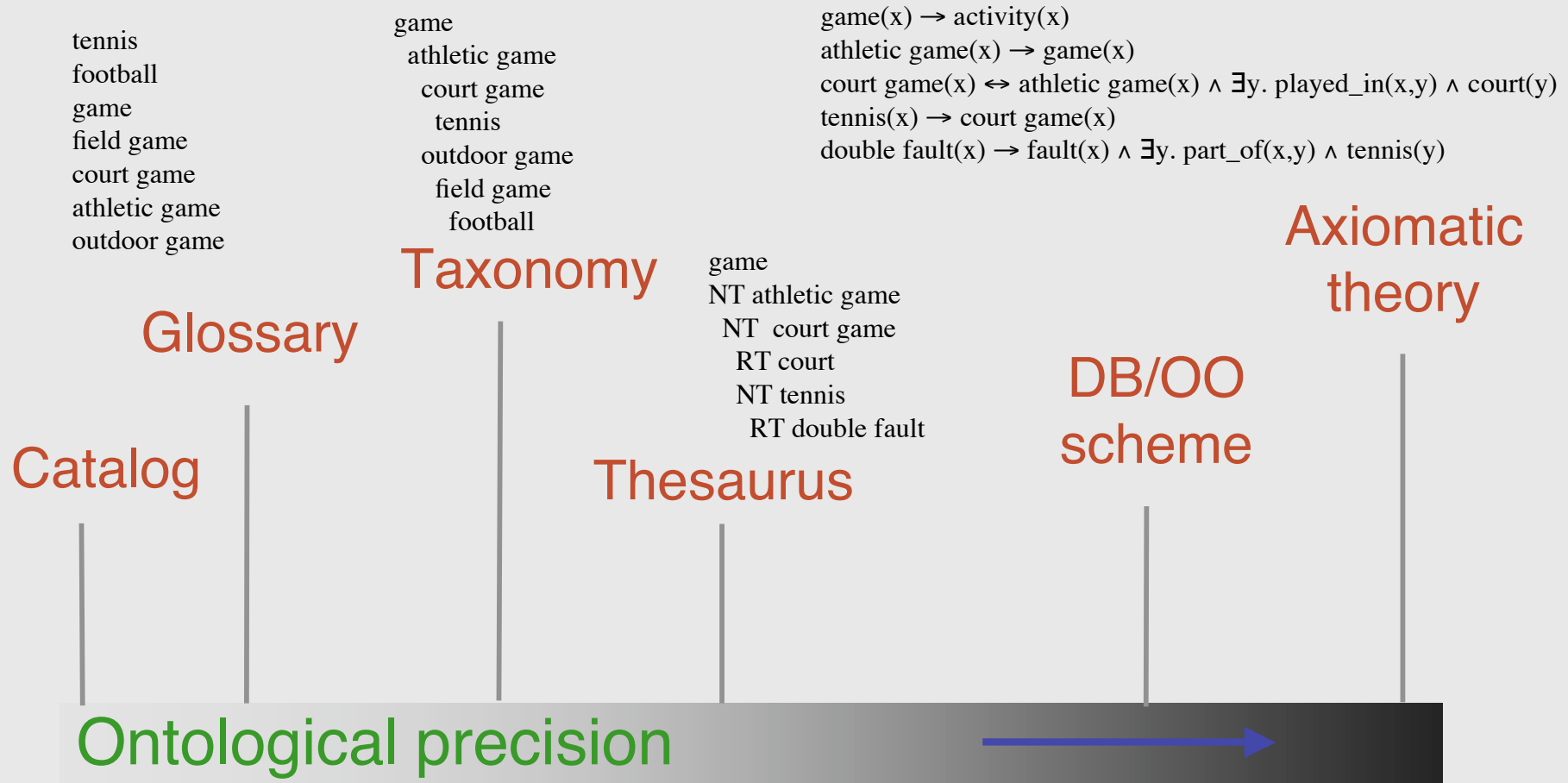


Max precision, limited coverage

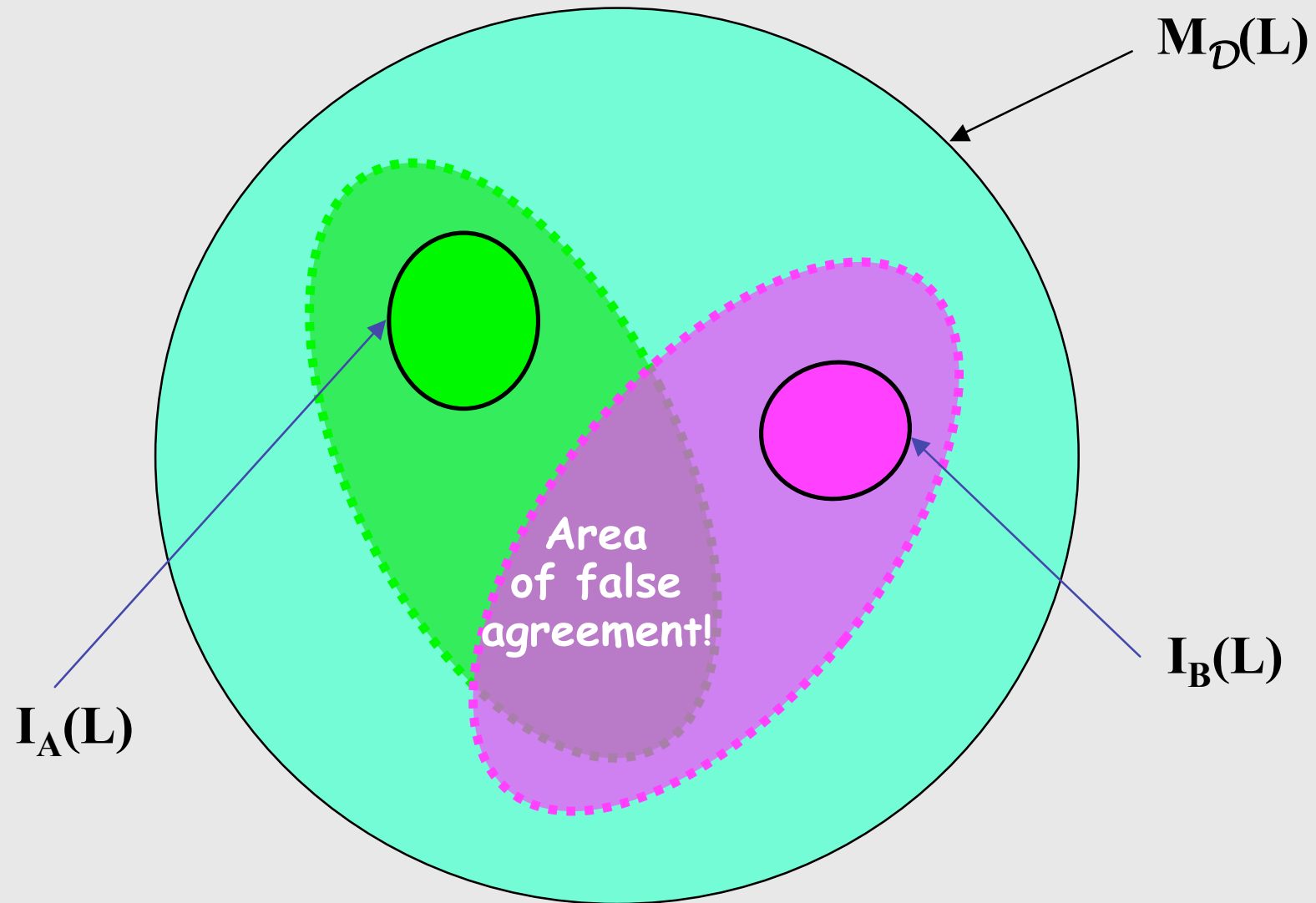


Low precision, limited coverage

Levels of Ontological Precision



Why precision is important



When precision is not enough

Only one binary predicate in the language: **on**

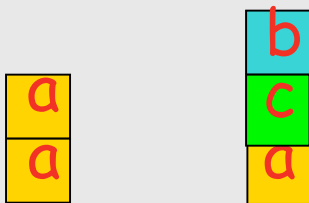
Only three blocks in the domain: **a**, **b**, **c**.

Axioms (for all x, y, z):

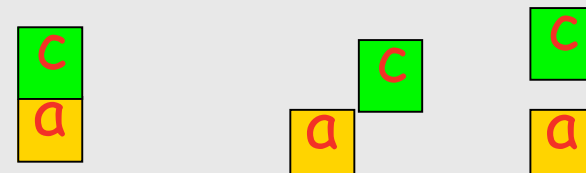
$$\text{on}(x, y) \rightarrow \neg \text{on}(y, x)$$

$$\text{on}(x, y) \rightarrow \neg \exists z (\text{on}(x, z) \wedge \text{on}(z, y))$$

Non-intended *models* are excluded, but the rules for the competent usage of **on** in different *situations* are not captured.



Excluded situations



Indistinguishable situations

Precision vs. Accuracy

- In general, a single intended *model* may not discriminate among relevant alternative *situations* because of
 - Lack of *primitives*
 - Lack of *entities*
- Capturing all intended models is not sufficient for a “perfect” ontology
 - Precision*: non-intended *models* are excluded
 - Accuracy*: non-intended *situations* are excluded



When is a precise (and well-founded) ontology useful?

1. When *subtle distinctions* are important
2. When *recognizing disagreement* is important
3. When *general abstractions* are important
4. When *careful explanation and justification* of ontological commitment is important
5. When *mutual understanding* is more important than interoperability.



4. Ontologies and...

Ontologies and classifications

- Classifications focus on:
 - ***access***, based on pre-determined criteria (encoded by ***syntactic keys***)
- Ontologies focus on:
 - ***Meaning*** of terms
 - ***Nature*** and ***structure*** of a domain



Ontologies and Database Schemas

- Database schemas:
 - Constraints focus on *data integrity*
 - Relationships and attribute values out of the DoD
- Ontologies:
 - Constraints focus on *intended meaning*
 - Relationships and attribute values first class citizens



Ontologies vs. Knowledge Bases

- Knowledge base
 - Assertional component
 - reflects *specific (epistemic) states of affairs*
 - designed for *problem-solving*
 - corresponding to *episodic memory*
 - Terminological component (*ontology*)
 - *independent* of particular *states of affairs*
 - Designed to support *terminological services*
 - corresponding to *semantic memory*

Ontological formulas are (assumed to be)
invariant, necessary information



5. The ontological level

A single, imperialistic ontology?

- An ontology is first of all *for understanding each other*
 - ...among people, first of all!
 - not necessarily for thinking in the same way
- A single ontology for multiple applications *is not necessary*
 - Different applications using different ontologies can co-exist and co-operate (not necessarily inter-operate)
 - ...if linked (and compared) together *by means of a general enough basic categories and relations (primitives)*.
- If basic assumptions are not made explicit, any imposed, common ontology risks to be
 - seriously mis-used or misunderstood
 - opaque with respect to other ontologies



Which primitives?

The role of ontological analysis

- Theory of Essence and Identity
- Theory of Parts (Mereology)
- Theory of Wholes
- Theory of Dependence
- Theory of Composition and Constitution
- Theory of Properties and Qualities

The basis for a common ontology vocabulary

*Idea of Chris Welty, IBM Watson Research
Centre, while visiting our lab in 2000*

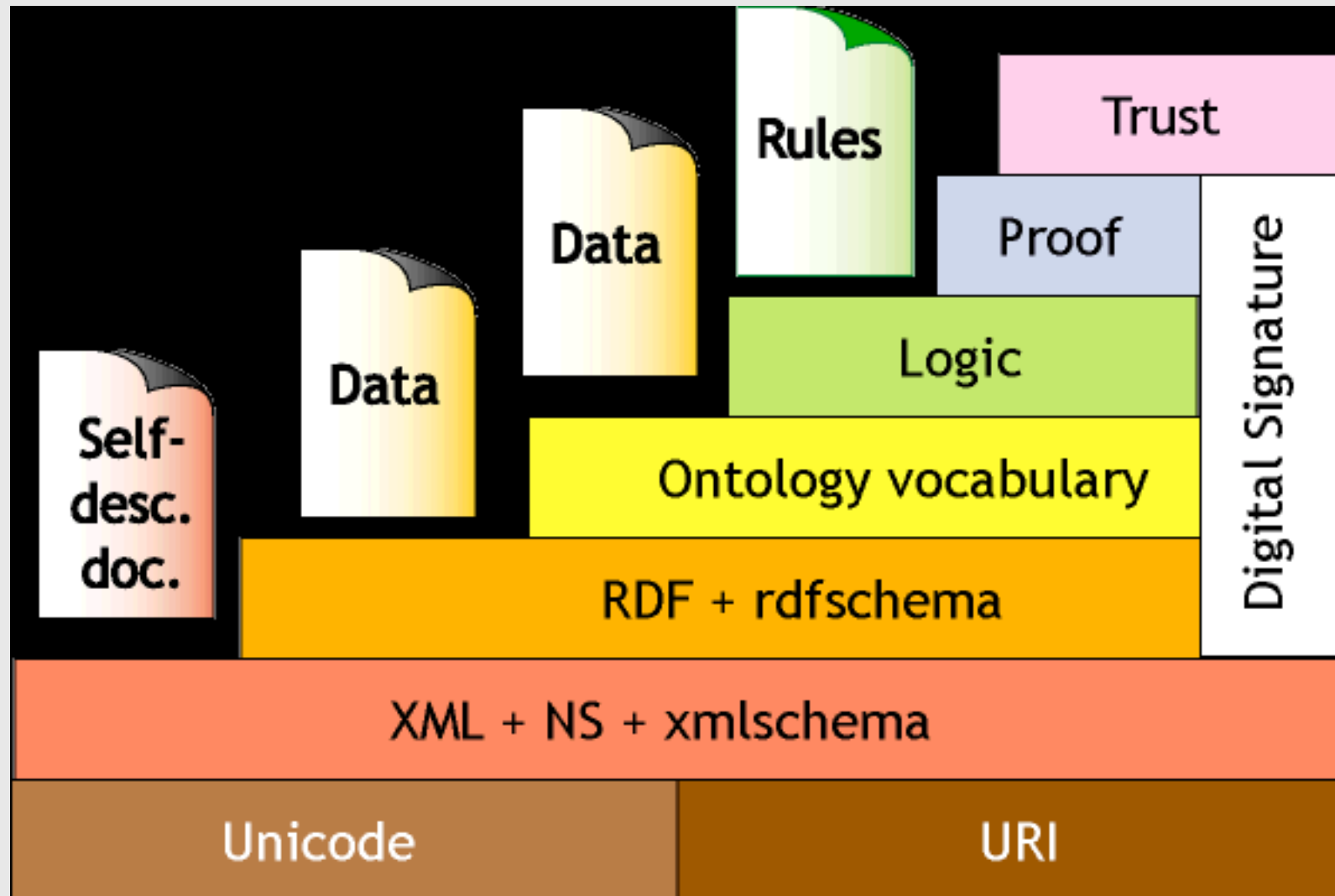


The Ontological Level

(Guarino 94)

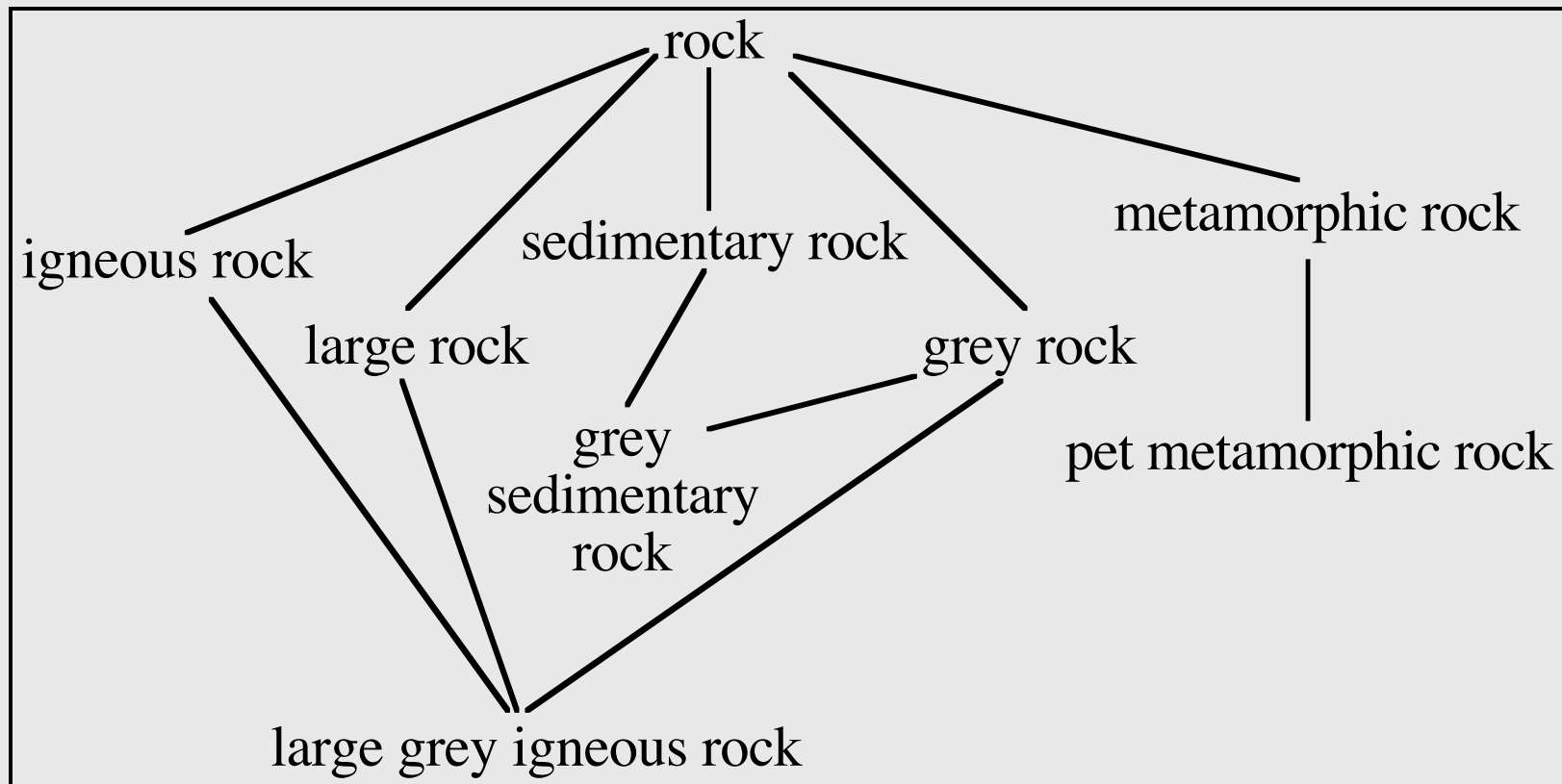
<i>Level</i>	<i>Primitives</i>	<i>Interpretation</i>	<i>Main feature</i>
Logical	Predicates, functions	Arbitrary	Formalization
Epistemological	Structuring relations	Arbitrary	Structure
Ontological	Ontological relations	Constrained (meaning postulate s)	Meaning
Conceptual	Conceptual relations	Subjective	Conceptualization
Linguistic	Linguistic terms	Subjective	Language dependence

The semantic web architecture [Tim Berners Lee 2000]



Kinds, roles, attributions

How many rock kinds are there?



[From Brachman, R., R. Fikes, et al. 1983. "Krypton: A Functional Approach to Knowledge Representation", *IEEE Computer*]



The answer

- According to Brachman & Fikes 83:
 - It's a dangerous question, only “safe” queries about analytical relationships between terms should be asked
- In a previous paper by Brachman and Levesque on *terminological competence in knowledge representation* [AAAI 82]:
 - “an *enhancement mode transistor* (which is a *kind* of transistor) should be understood as different from a *pass transistor* (which is a *role* a transistor plays in a larger circuit)”
- These issues have been simply *given up* while striving for logical simplification and computational tractability
- The OntoClean methodology, based on formal ontological analysis, allows us to conclude: *there are 3 kinds of rocks* (appearing in the figure)



A new journal: *Applied Ontology*



Editors in chief:

Nicola Guarino
ISTC-CNR

Mark Musen
Stanford University

IOS Press

Amsterdam, Berlin,
Washington, Tokyo, Beijing

www.applied-ontology-org

FOIS-2006

International Conference on Formal Ontology in Information Systems

<http://www.formalontology.org/>



**November 9-11, 2006
Baltimore, Maryland (USA)**



Conclusions

- Not all properties are the same
- Not all relations are the same
- Ontological distinctions *do matter*, and require to be represented *at the suitable level*
- A humble interdisciplinary approach is essential
- “...*But this is hard!!*”

why should it be EASY?!

