Laboratory for Applied Ontology

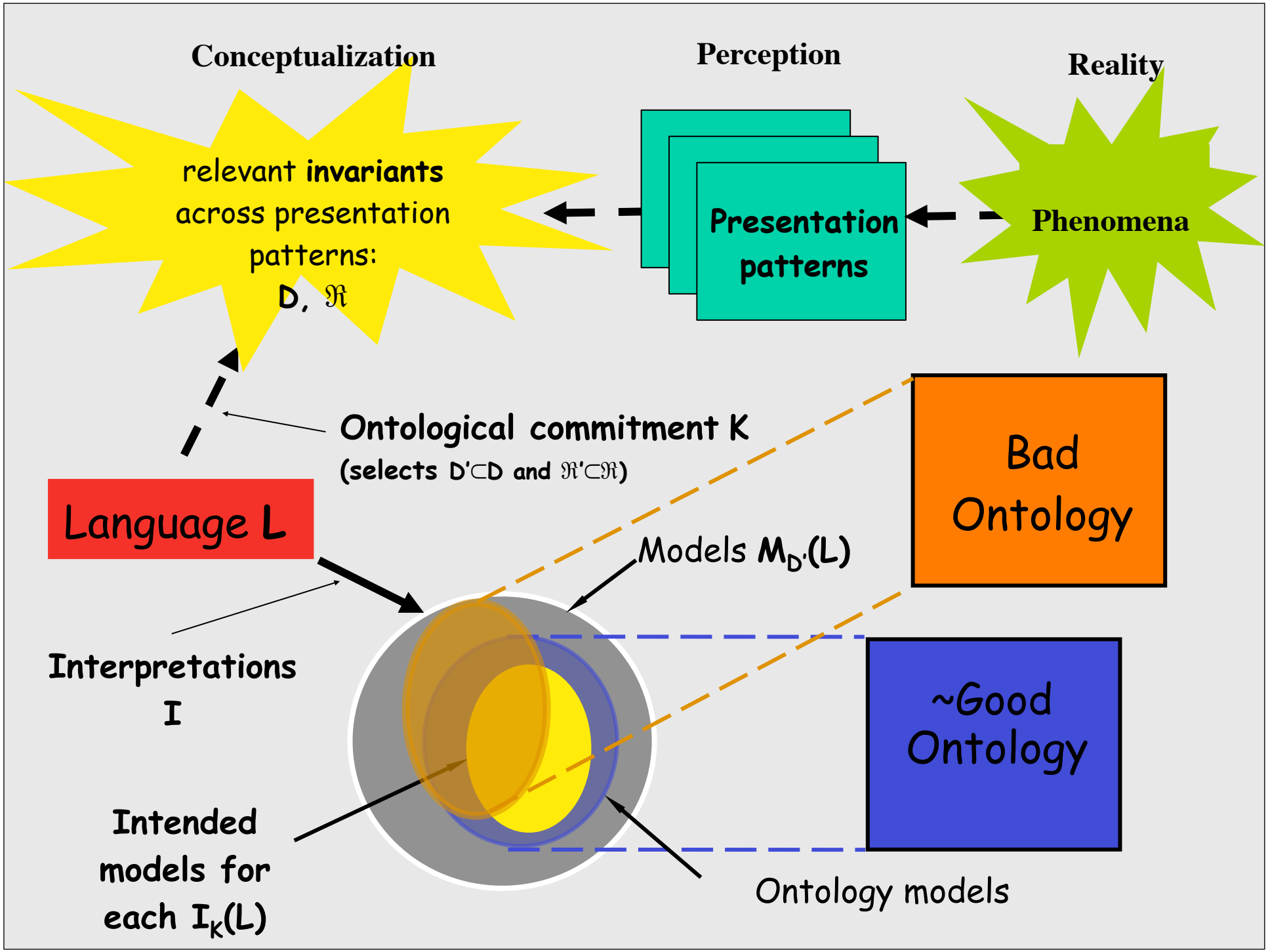
Institute of Cognitive Science and Technology
Italian National Research Council

Ontology quality, ontology design patterns, and competency questions

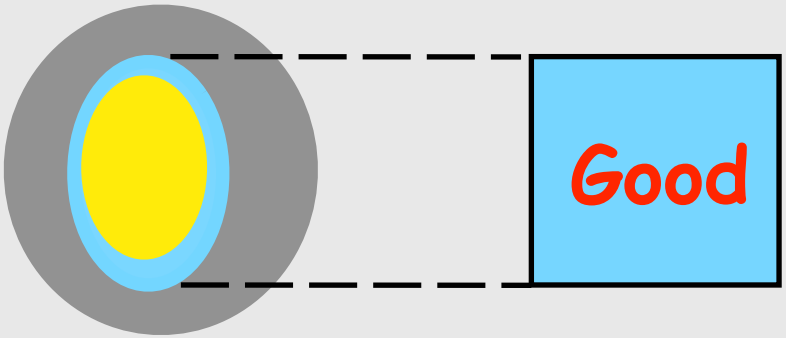
Nicola Guarino

Italian National Research Council
Institute of Cognitive Sciences and Technology (ISTC-CNR)
Laboratory for Applied Ontology, Trento

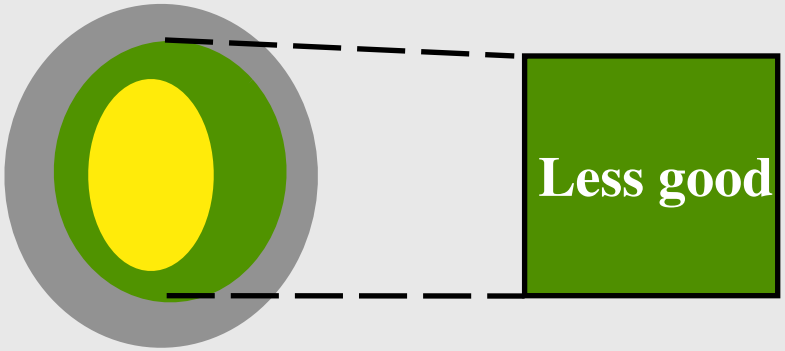
Good ontologies and bad ontologies


Conceptualization
Perception
Reality
relevant invariants across presentation patterns: D, ℜ
Presentation patterns
Phenomena
Ontological commitment K
(selects D'⊂D and ℜ'⊂ℜ)
Language L
Models $M_{D'}(L)$
Bad Ontology
Interpretations I
~Good Ontology
Intended models for each $I_K(L)$
Ontology models

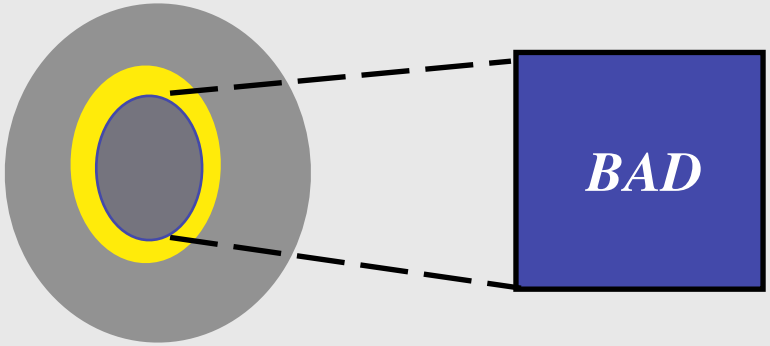
Ontology Quality: Precision and Correctness



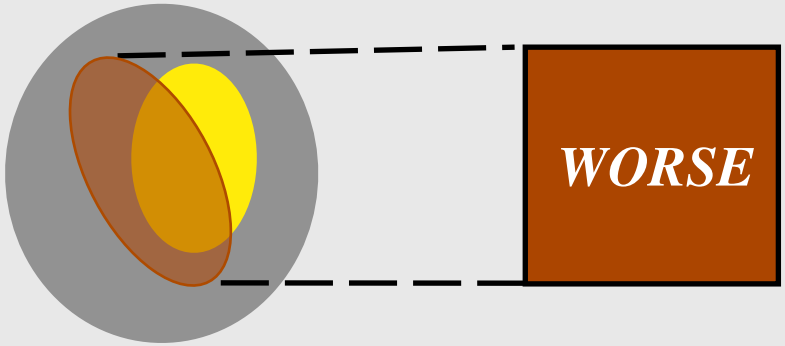
High precision, max correctness



Low precision, max correctness



Max precision, low correctness



Low precision, low correctness

When precision is not enough

Only one binary predicate in the language: on

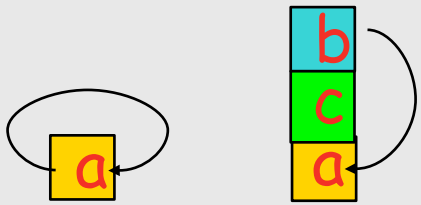
Only three blocks in the domain: a, b, c.

Axioms (for all x,y,z):

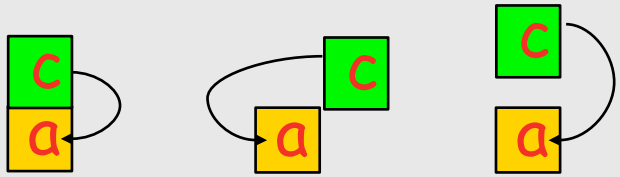
$on(x,y) \rightarrow \neg on(y,x)$

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

All non-intended models (for the given domain) are excluded, but still some counterexamples can't be excluded



Excluded counterexamples



Indistinguishable examples

A third dimension for ontology quality: *accuracy*

- In general, a single intended ***model*** may not discriminate between positive and negative ***examples*** because of a ***mismatch*** between*:*
  - Cognitive domain and domain of discourse: lack of ***entities***
  - Conceptual relations and ontology relations: lack of ***primitives***

- Capturing all intended models is not sufficient for a “perfect” ontology

  Precision: non-intended *models* are excluded

  Accuracy: negative *examples* are excluded

Correctness, precision, and accuracy

- Correctness: no ontology constraint is wrong
- Precision: the ontology discriminates between wrong and correct ***statements***
- Accuracy: the ontology discriminates between wrong and correct ***examples***

Ontology design patterns and ontology quality

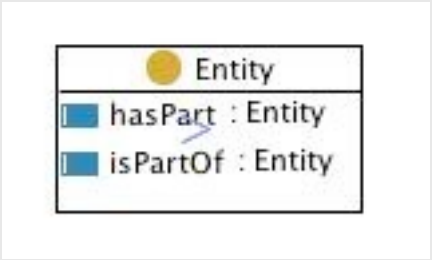
"ontology design patterns play an important role in obtaining higher quality ontologies"
(EKAW 2010 Ontology Quality Workshop CFP)

A critical tradeoff: reusability vs. interoperabilty

- *ontology patterns* are supposed to be highly ***reusable***...

- Are they ***interoperable***?
- Differently from generic software, interoperability is the *reason d'être* of ontologies...

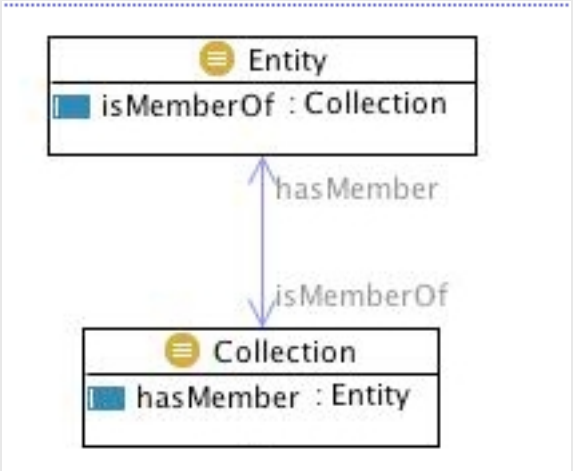
A *content pattern*[1]: Part-of



- Name: part of
- Intent: To represent entities and their parts
- Competency questions:
  - *What is this entity part of?*
  - *What are the parts of this entity?*

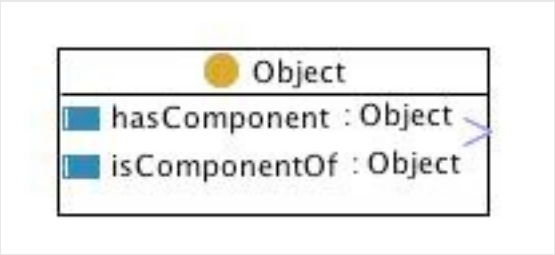
[1]from *www.ontologydesignpatterns.org*

Member-collection



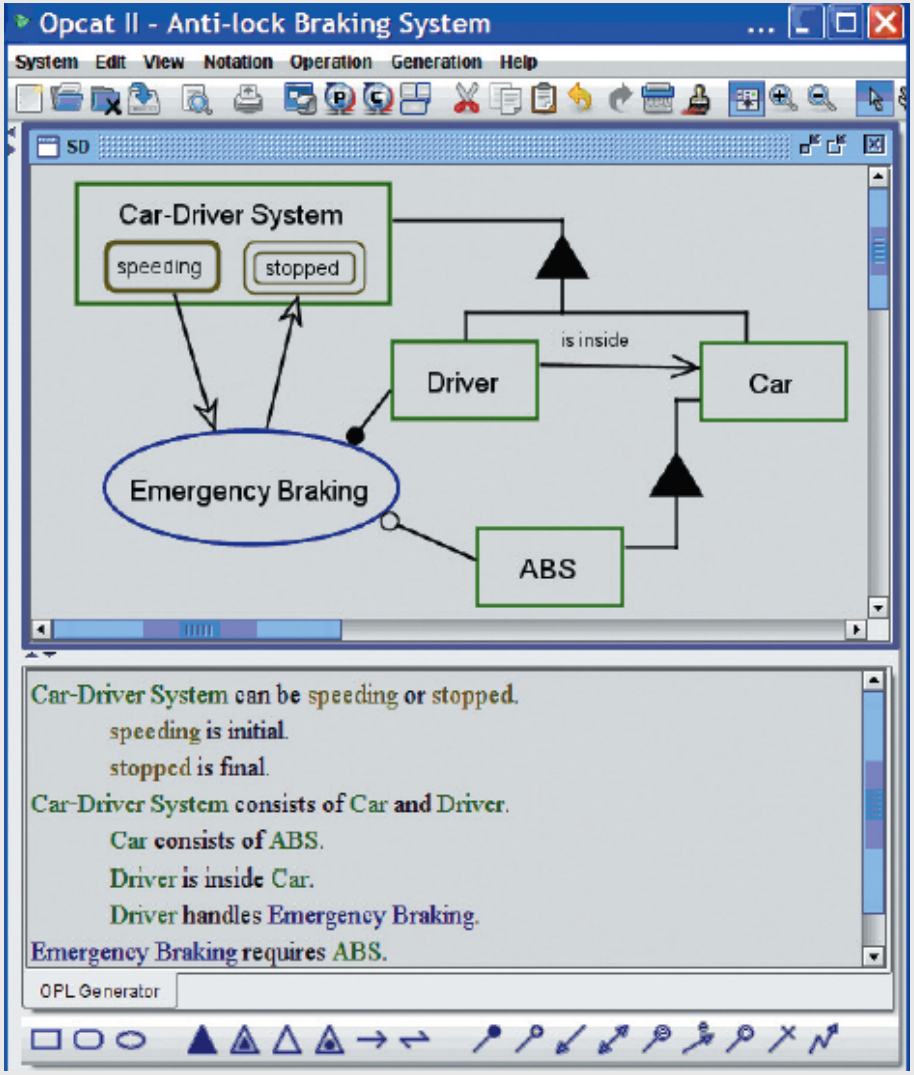
- Competency questions:
  - *What things are contained in this collection?*
  - *What collections this thing is member of?*

Componency



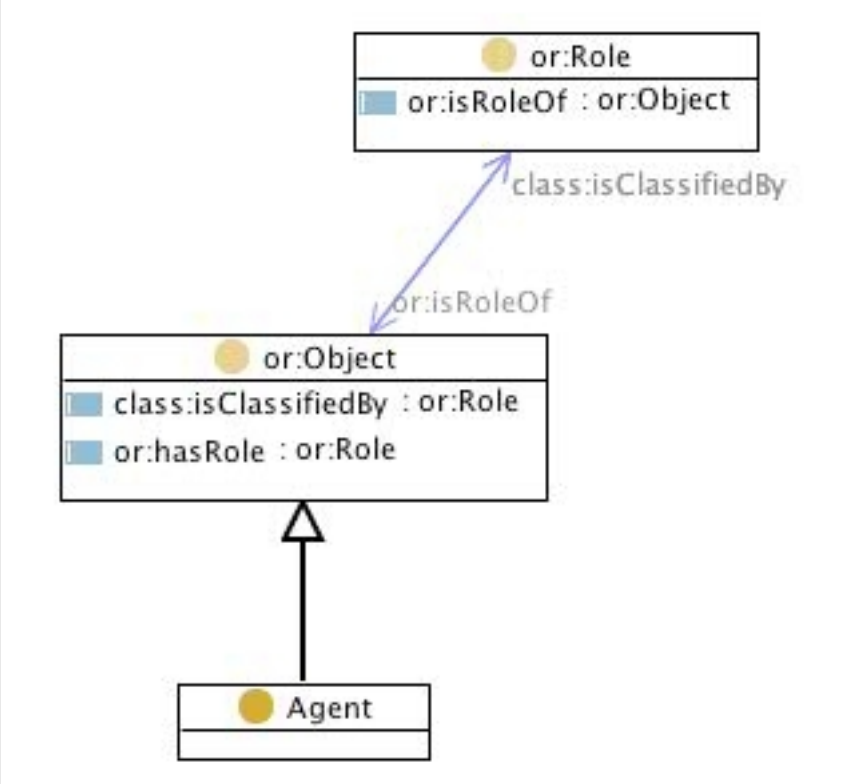
- Competency questions:
  - *What is this object component of?*
  - *What are the components of this object?*

An interpretation of “part-of”... or “component-of”?



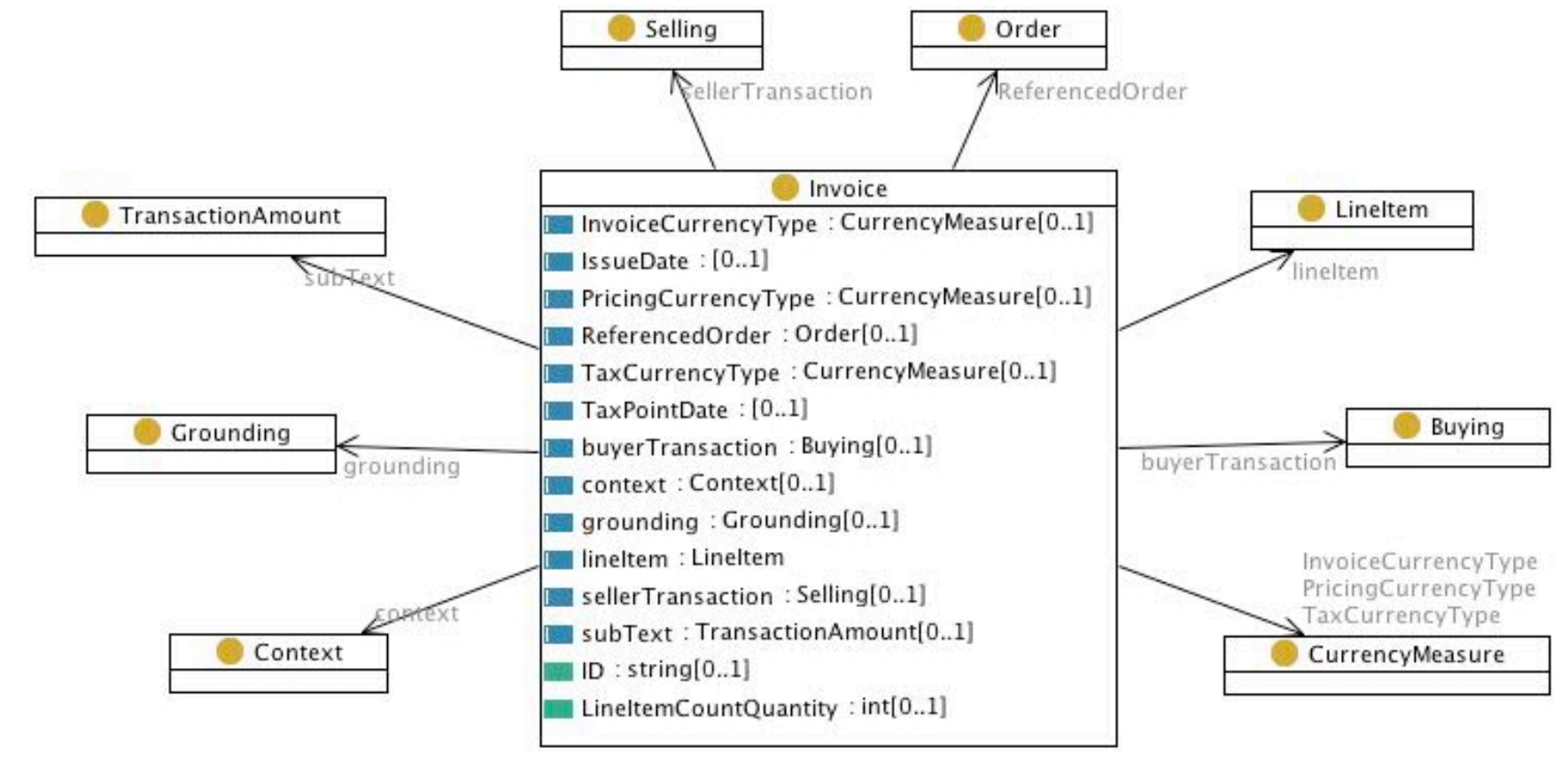
Dov Dory, Words from pictures for dual-channel processing, *Communications of the ACM* 51, 2008

Agent-Role



- which agent does play this role?
- what is the role that played by that agent?

Invoice



- What are the transactions involved in this invoice?
- What is the order this invoice is referring to?
- What is the line item for this invoice?
- What is the amount of the transactions involved in this invoice?
- What currency is applied to this invoice?

Some competency questions for the GoodRelations ontology

- CQ1: Which retrievable Web Resources describe an offer?
- CQ2: For which time frame is the offer valid?
- CQ3: Which types of customers are eligible?
- CQ4: Which are the eligible customer regions?
- CQ5: Which shipping / delivery methods are available?
- CQ6: Which methods of payment are accepted?
- CQ10: What is the mail address and which are the contact details of the offering business entity?

Competency questions, according to their inventors (1)

It is not a well-designed ontology if all competency questions have the form of simple lookup queries

Uschold & Gruninger 96

Competency questions, according to their inventors (2)

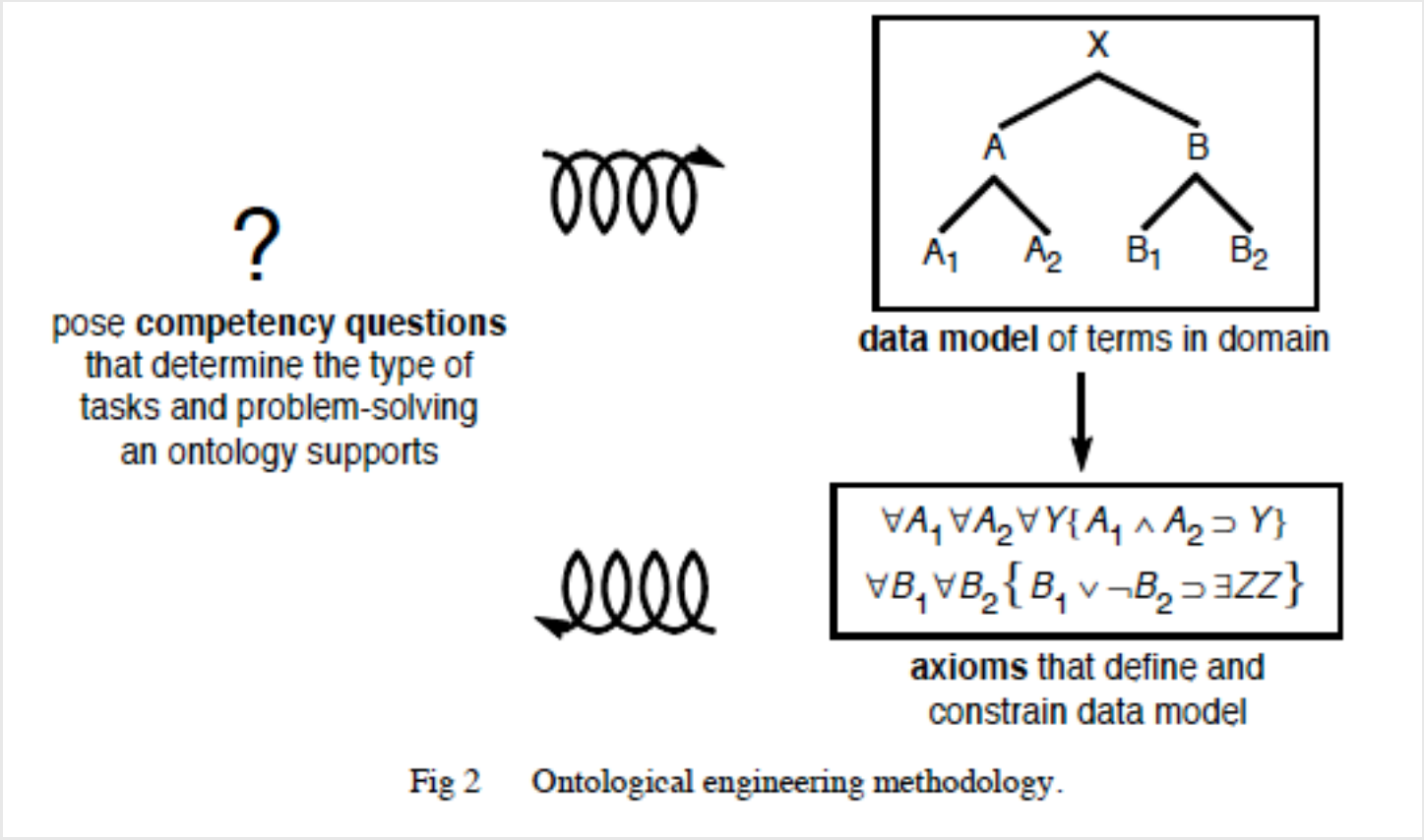


Fig 2 Ontological engineering methodology.

Kim, Fox & Gruninger 99

Original Gruninger's competency questions

- Planning and scheduling -- what sequence of activities must be completed to achieve some goal? At what times must these activities be initiated and terminated?
- Temporal projection -- Given a set of actions that occur at different points in the future, what are the properties of resources and activities at arbitrary points in time?
- Execution monitoring and external events -- What are the effects on the enterprise model of the occurrence of external and unexpected events (such as machine breakdown or the unavailability of resources)?
- Hypothetical reasoning -- what will happen if we move one task ahead of schedule and another task behind schedule? What are the effects on orders if we buy another machine?

Competency questions revisited

- Epistemological:
  - what is this entity part of?
  - what are the parts of this entity?

- Ontological:
  - what does it mean to be a part of something?
  - can something be part of itself?
  - can something have only one (proper) part?
  - are two entities the same if they have the same parts?
  - does parthood imply contact?
  - what's the difference between parthood and spatial inclusion?
  - what's the difference between parts and components?
  - how are they related?

A simple methodology towards ontology quality

1. Isolate a ***target community***
2. For each term to be used in the ontology, check its possible ambiguities within the target community (collecting ***examples*** and ***counterexamples***)
3. Leveraging on axioms, and on the proper choice of domain and primitives
    1. Account for the differences among different senses
    2. Account for the relationships among different senses
4. ***Stop*** when all the terms used are unambiguous ***for the target community***

Conclusions: the risks of (current) ODPs in the light of ontology quality

- ***Underspecification***: simplicity encourages reusability but risks to decrease interoperability
- ***Isolation***: focusing on an isolated pattern risks to overlook important ***structural connections***