Formal Ontology, Patterns and Anti-Patterns for Next-Generation Conceptual Modeling

Giancarlo Guizzardi

Ontology and Conceptual Modeling Research Group (NEMO) Federal University of Espírito Santo, Brazil

<u>also at:</u> Laboratory for Applied Ontology, ISTC/CNR, Trento, Italy





# Acknowledgements

I am grateful to Gerd Wagner, Nicola Guarino, Ricardo Falbo, Renata Guizzardi, João Paulo Almeida, John Mylopoulos and the members of NEMO for many years of fruitful collaboration.

"Conceptual Modeling is the activity of describing aspects of the physical and social world for the purpose of understanding and communication... the adequacy of a conceptual modeling notation rests in its ability to promote understanding about that world among its human users"

(John Mylopoulos, Conceptual Modeling and Telos, 1992)

#### Why now?



The **Taxonomy** of Animals in *The Celestial Emporium of Benevolent Knowledge (Borges)* 

- Those that belong to the emperor
- Those that resemble flies from a distance
- Those that have just broken a flower vase
- Embalmed ones
- Fabulous ones

#### "Those that resemble flies from a distance" is a logically possible way to group objects, but it's not how we naturally make sense of the world. No real language would have a noun for such a category...Real nouns capture something deep; they refer to **kinds** of things that are thought to share deep properties..."

(Paul Bloom, How Pleasure Works, 2010)

#### "...As the evolutionary theorist Stephen Jay Gould put it, our classifications don't just exist to avoid chaos, they are "theories about the basis of natural order."

(Paul Bloom, How Pleasure Works, 2010)

# Carving reality at its joints [Plato]:





### "Carving up Reality"

#### We need to guarantee

#### Intra-worldview Consistency

and

#### Inter-worldview Interoperability

#### Why now?



The increasing need to answer Cross-Silo Questions









Tweet < 0

F Like < 0

Timekeeper

Having lost one legal case to insurers, the towers' leaseholder wins a second

Dec 9th 2004 | NEW YORK | From the print edition

SEVEN months ago, a jury in lower Manhattan ruled that under the forms covering insurance of the World Trade Centre, the striking of the twin towers by two aeroplanes constituted only one "occurrence". Consequently, Larry Silverstein, who had recently leased the Trade Centre complex, was entitled to one payment, not two—a difference of \$3.5 billion. On December 6th, in the same courtroom with the same judge presiding, another jury decided that under the documents used by nine other insurers the attacks were two events, thus qualifying for two payments. The verdict will provide Mr Silverstein with as much as \$1.1 billion extra for rebuilding the Trade Centre. It will also ensure that he remains in control of the project.

Why, after two weeks of deliberation, did the second jury come to a different conclusion from the first? The main reason lay in the preliminary paperwork signed by the underwriters. Because the Trade Centre had been leased to Mr Silverstein only weeks before the attack, the final insurance contracts had yet to be signed. The insurers in the first trial had signed a form with a much tighter definition of an "occurrence" than in the form signed by the nine insurers in the second trial. In addition, the insurance companies' claim that they always define FaceTime e" precisely may have been undermined by testimony that they had been flexible in other cases—for example, involving sequences of

#### Advertisement



#### Follow The Economist



#### Why now?



The Criticality of the Domains Involved and Complexity of the Concepts at hand

# "Carving up Reality"

- There is not doubt about the brute reality. The issue is interpreting that part of reality according to a certain system of categories
- These categorization operations are in a sense a prioristic

# Ontology

- For that we need a *a prioristic* system of categories and their ties addressing issues of Identity, Unity (Parts and Wholes), Individuation, Change, Classification and Taxonomic Structures, Dependence (Existential, Historical, Relational, Notional), Causality, Essential and Accidental Characterization
- We need Formal Ontology and Ontological Analysis

#### Ontology-Driven Conceptual Modeling

A discipline aiming at developing ontology-based methodologies, computational tools and **modeling languages** for the area of Conceptual Modeling

### UFO

#### (Unified Foundational Ontology)

- Over the years, we have built a Philosophically and Cognitively well-founded Ontology to contribute to the general goal of serving as a Foundation for Conceptual Modeling
- This Ontology has been used to as a theory for addressing may classical conceptual modeling constructs such as Object Types and Taxonomic Structures (CAISE 2004, CAISE 2007, CAISE 2012), Part-Whole Relations (CAISE 2007, CAISE 2009, FOIS2010, CAISE 2011), Intrinsic and Relational Properties (ER 2006, ER 2008, ER 2011, CAISE 2015), Weak Entities, Attributes and Datatypes (ER 2006), Events (ER 2013), Services (EDOC 2013), Capabilities (EDOC 2013), Goals, Communities, Organizational Structures, etc...













#### Kinds



# Anti-Rigid Sortals (**Roles** and **Phases**)



# Anti-Rigid Sortals (**Roles** and **Phases**)



## **Rigid Mixins**



## **Anti-Rigid Mixins**



## **Anti-Rigid Mixins**









(these ontological distinctions with the meaning they have in OntoUML were first presented at Guizzardi, Wagner, Guarino & van Sinderen in CAISE, 2004 and are strongly based on the ontological distinctions underlying the OntoClean methodology, see Guarino & Welty, 2002, 2004)


Ontologically well-defined, formally characterized and cognitively sound systems of types

2



#### Problem (1)

- 1. Characterize the difference between the following types:
  - Person, Apple, Car, Dog, Organization
  - Student, Singer, President, Employee
  - Adult, Puppy, Metropolis
  - Crime Weapon, Insurable Item, Sharp Object, Rational Agent, Cultural Heritage Item





#### Solution

- 1. Characterizing the difference between:
  - NATURAL TYPE/KIND (e.g., Person, Apple, Car) = RIGID
    SORTAL
  - ROLE (e.g., Student, President, Employee) = ANTI-RIGID
    + RELATIONALLY DEPENDENT SORTAL
  - PHASE (e.g., Living Person, President, Employee) = ANTI-RIGID + RELATIONALLY INDEPENDENT SORTAL
  - MIXIN (e.g., Crime Weapon, Insurable Item, Sharp Object, Rational Agent, Cultural Heritage Item)? = MIXIN



Incorporation of ontological constraints in the language metamodel to guarantee *ontological consistency by design* 

#### Role

- All instances of a given ROLE are of the same KIND (e.g., all Students are Person)
- All instances of a ROLE instantiate that type only contingently (e.g., no Student is necessarily a Student)
- Instances of a KIND instantiate that ROLE when participating in a certain RELATIONAL CONTEXT (e.g., instances of Person instantiate the Role Student when enrolled in na Educational Institution)
- A ROLE cannot be a supertype of a Rigid Type





#### WORLD W



#### WORLD W



#### WORLD W'



#### WORLD W'







We run into a logical contradiction!

#### Role

- All instances of a given ROLE are of the same KIND (e.g., all Students are Person)
- All instances of a ROLE instantiate that type only contingently (e.g., no Student is necessarily a Student)
- Instances of a KIND instantiate that ROLE when participating in a certain RELATIONAL CONTEXT (e.g., instances of Person instantiate the Role Student when enrolled in na Educational Institution)
- A ROLE cannot be a supertype of a Rigid Type





#### The Emerging Role Pattern



#### The Emerging Phase Pattern



Problem (2)

 Suppose that I want to represent that the ROLE Customer can be played by entities of different KINDS, namely, People and Organizations. How to relate the ROLE and its *allowed types* using subtyping relations?

#### A Classic Problem





#### A Possible Alternative?



«roleMixin» **Customer** 





## The emerging **RoleMixin** Pattern



#### Which one is better?



by Chris Welty

#### The RoleMixin Pattern in ORM



by Terry Halpin



Precise methodological guidelines for mapping into different implementation environments







#### Precise modal semantics with implications for validation





Valid state of affairs according to the representation

Intended state of affairs according to the Conceptualization




**Over-constraining** 



False Agreement











## Conceptual Model = Structure + Axiomatization (Ontological Commitment)





### Conceptual Model = Structure + Domain-Independent Axioms + Domain-Specific Axioms















## Data Modeling Guide (DMG) For An Enterprise Logical Data Model (ELDM)

Version 2.3

March 15, 2011

The U.S. Government has rights in this document in accordance with DFAR 252.227-7013 Rights in Technical Data – Noncommercial Item (Nov 1995) under contract number H98230-09-C-1180.

#### Data Modeling Guide (DMG) For An Enterprise Logical Data Model, V2.3; 15 March 2011

### Preface

The success of this Data Modeling Guide for an Enterprise Logical Data Model could not have been possible without the inspired and exhaustive research of Giancarlo Guizzardi, notably his "Ontological Foundations for Structural Conceptual Models," published in 2005 in association with the Centre for Telemetrics and Information Technology, which provided the theoretical foundation for the methodologies describe within, and from which real world, practical implementations have already ensued.

At the core of Guizzardi's modeling paradigm are the principles of Rigidity, Uniform Identity and Existential Dependence. From those foundational tenets he extrapolates the concepts of SortalUniversal (Unified Principle of Identity), MixinUniversal (Disparate Set of Concepts), and finally the constructs of SubstanceSortal (Kind, Quantity, and Collective), Subkind, Phase, Role, Category, RoleMixin and Mixin. In short, the total package offered to us by Guizzardi contained a complete and fully integrated set of concepts and constructs that left us wanting for nothing.

# OntoUML Model Benchmark

- Model benchmark with 56 models
- Models in domains such as Provenance in Scientific Workflow, Public Cloud Vulnerability, Software Configuration Management, Emergency Management, Services, IT Governance, Organizational Structures, Software Requirements, Heart Electrophisiology, Amazonian Biodiversity Management, Human Genome, Optical Transport Networks, Federal Government Organizational Structures, Normative Acts, and Ground Transportation Regulation

## The Emerging Anti-Pattern: Relation Between Overlapping Types (**RelOver**)



## The Emerging Anti-Pattern: Relation Specialization (**RelSpec**)





(b)





# Anti-Pattern Catalogue

- Association Cycle
- Binary Relation Between Over. Types
- Deceiving Intersection
- Free Role Specialization
- Imprecise Abstraction
- Multiple Relational Dependency
- Part Composing Over. Roles
- Whole Composed by Over. Parts
- Relator Mediating Over. Types
- Relation Composition
- Relator Mediating Rigid Types
- Relation Specialization
- Repeatable Relator Instances

- Relationally Dependent Phase
- Generalization Set With Mixed Rigidity
- Heterogeneous Collective
- Homogeneous Functional Complex
- Mixin With Same Identity
- Mixin With Same Rigidity
- Undefined Formal Association
- Undefined Phase Partition

Anti-Patterns (AP)	AP Occurrences	Relevant Model Construct (RMC)	RMC /AP Ratio	% of Qualified Models with AP Occurrence	
RelSpec	817	Association	4.92	48.15%	
ImpAbs	758	Association	5.30	72.22%	
AssCyc	1809	Association	2.22	92.59%	
RelOver	149	Relator	8.08	25%	
RepRel	319	Relator	3.77	64.58%	
BinOver	224	Association	17.93	48.15%	

Anti-Patterns (AP)	AP Occurrences	Relevant Model Construct (RMC)	RMC /AP Ratio	% of Qualified Models with AP Occurrence	
RelSpec	817	Association	4.92	48.15%	
ImpAbs	758	Association	5.30	72.22%	
AssCyc	1809	Association	2.22	92.59%	
RelOver	149	Relator	8.08	25%	
RepRel	319	Relator	3.77	64.58%	
BinOver	224	Association	17.93	48.15%	

Anti-Patterns (AP)	AP Occurrences	Relevant Model Construct (RMC)	RMC /AP Ratio	% of Qualified Models with AP Occurrence	
RelSpec	817	Association	4.92	48.15%	
ImpAbs	758	Association	5.30	72.22%	
AssCyc	1809	Association	2.22	92.59%	
RelOver	149	Relator	8.08	25%	
RepRel	319	Relator	3.77	64.58%	
BinOver	224	Association	17.93	48.15%	

Anti-Patterns (AP)	AP Occurrences	Relevant Model Construct (RMC)	RMC /AP Ratio	% of Qualified Models with AP Occurrence	
RelSpec	817	Association	4.92	48.15%	
ImpAbs	758	Association	5.30	72.22%	
AssCyc	1809	Association	2.22	92.59%	
RelOver	149	Relator	8.08	25%	
RepRel	319	Relator	3.77	64.58%	
BinOver	224	Association	17.93	48.15%	

Anti-Pattern	#Occ.	#Error	#Error / #Occ.	#Refac. /#Error	
RelSpec	315	279	88.6%	97.1%	
RepRel	221	57	25.8%	84.2%	
RelOver	124	70	56.5%	77.1%	
BinOver	74	31	41.9%	74.2%	
AssCyc	20	14	70.0%	71.4%	
ImpAbs	125	11	8.8%	27.3%	
Total	879	462	52.56%	88.53%	

Anti-Pattern	#Occ.	#Error	#Error / #Occ.	#Refac. /#Error
RelSpec	315	279	88.6%	97.1%
RepRel	221	57	25.8%	84.2%
RelOver	124	70	56.5%	77.1%
BinOver	74	31	41.9%	74.2%
AssCyc	20	14	70.0%	71.4%
ImpAbs	125	11	8.8%	27.3%
Total	879	462	52.56%	88.53%

#### OntoUML Criminal Investigation +child «Role» parentOf 🕨 Child «Material» +parent 1..2 «Role» «Kind» «Role» «Role» «Material» Person Parent Captain Officer knows +detective +suspect conductedBy 🕨 investigates 🕨 1...\* «Role» «Role» «Mediation» «Mediation» 2...\* Detective Suspect 1 1...\* +investigation +investigation 1...\* +investigation +interrogator «Relator» Criminal Investigation 1...\* «Mediation» +investigation 1 relatedTo «Role» 1...\* Witness +investigative context +witness «Role» +leader responsibleFor 1 +witness «ComponentOf» Lead Detective «Mediation» +part of investigatic1...\* targetOf 🗩 1...\* «Mediation» +interrogation «Relator» «Mediation» Interrogation +interrogation conducts 📂





ext \_'Criminal Investigation'

				A				
	Relator Wi	th Overlapping Roles						
Relator: Criminal In	nvestigation							
Customizing Disjoin	ts Roles:			Add				
Lead Detective	Witness	Detective	Suspect					
<ul> <li>Exclusive</li> <li>Disjoint</li> <li>"Cr</li> </ul>	<ul> <li>Exclusive</li> <li>Exclusive from Table</li> <li>Disjoint</li> <li>Disjoint from Table</li> <li>"Criminal Investigation" Scope: 2 (at least)</li> <li>Execute With Analyzer</li> <li>OCL Solution</li> </ul>							
	1 🗸 🗸	۹   ۵	/   🛤 🤅	3 🖪 🗐				
1 context _'Cr 2 inv: self.wi 3 4 context _'Cr 5 inv: self.wi 6 7 context _'Dr 8 inv: not sel	riminal Invest itness.oclAsTy riminal Invest itness.oclAsTy etective' lf.oclIsTypeOf	igation' pe(Person)->as igation' pe(Person)->as (Suspect)	Set()->interse	ection(self.det	ective.oclAsTy	/pe(Person)->	>asSet())->	>size()=0 ize()=0
9								
🎹 Properties 🔥 V	Varnings 🔥 Erro	ors 📃 Output	OCL Editor					

"Few modelers, however,

have had the experience of subjecting their models to continual, automatic review.

Building a model incrementally with an analyzer,

**simulating** and checking as you go along, is a very different experience from using pencil and paper alone. The first reaction tends to be amazement: modeling is much more fun when you get instant, **visual feedback**. Then the sense of **humiliation** sets in, as you discover that **there's almost nothing you can do right**."

(Daniel Jackson, Software Abstractions : Logic, Language, and Analysis, 2006)

# The Humble Ontologist

[What] I have chosen to stress in this talk is the following. We shall do a much better ontology job in the future, provided that we approach the task with a full appreciation of its tremendous complexity,...,provided we respect the intrinsic limitations of the human mind and approach the task a Very Humble Ontologist

(paraphrasing Dijkstra's Humble Programmer, 1972)
## Demos, Tools and Model Repository

### Demos

#### https://www.youtube.com/channel/ UC9pKWYDVEevmDHaGQybAB1g

### Tools https://www.facebook.com/ ontoumleditor

<u>https://code.google.com/p/ontouml-</u> lightweight-editor/

http://www.menthor.net

### Model Repository

#### <u>http://www.menthor.net/model-</u> <u>repository.html</u>



**GIANCARLO GUIZZARDI** 

- SALLES, T.P., GUIZZARDI, G., Ontological Anti-Patterns: Empirically Uncovered Error-Prone Structures in Ontology-Driven Conceptual Models, Data & Knowledge Engineering (DKE) Journal, Volume 99, September 2015, Pages 72–104.
- RUY, F., REGINATO, C.C., SANTOS, V.A., FALBO, R., GUIZZARDI, G., Ontology Engineering by Combining Ontology Patterns, 35th International Conference on Conceptual Modeling (ER 2015) Stockholm, 2015.
- GUIZZARDI, G., Ontological Patterns, Anti-Patterns and Pattern Languages for Next-Generation Conceptual Modeling, invited companion paper to the Keynote Speech delivered at the 33rd International Conference on Conceptual Modeling (ER 2014), Atlanta, USA.

- NARDI, J., FALBO, R., ALMEIDA, J.P., GUIZZARDI, G., PIRES, L.F., VAN SINDEREN, M., GUARINO, N., FONSECA, C. M., A Commitment-Based Reference Ontology for Services, Information Systems, Oxford University Press, Volume 54, December 2015, Pages 263–288
- GUIZZARDI, G.; GRAÇAS, A.P., GUIZZARDI, R.S.S., Design Patterns and Inductive Modeling Rules to Support the Construction of Ontologically Well-Founded Conceptual Models in OntoUML, 3rd International Workshop on Ontology-Driven Information Systems (ODISE 2011), together with the 23rd International Conference on Advanced Information System Engineering (CAiSE'11), London, UK.
- BENEVIDES, A.B.; GUIZZARDI, G.; BRAGA, B.F.B.; ALMEIDA, J.P.A., Validating modal aspects of OntoUML conceptual models using automatically generated visual world structures, Journal of Universal Computer Science, Special Issue on Evolving Theories of Conceptual Modeling, Editors: Klaus-Dieter Schewe and Markus Kirchberg, 2010.

- GUIZZARDI, G., ZAMBORLINI, V., Using a Trope-Based Foundational Ontology for Bridging different areas of concern in Ontology-Driven Conceptual Modeling, Science of Computer Programming, Elsevier, 2014, DOI:10.1016/ j.scico.2014.02.022.
- ZAMBORLINI, V.; GUIZZARDI, G., On the representation of temporally changing information in OWL, IEEE 5th Joint International Workshop on Vocabularies, Ontologies and Rules for The Enterprise (VORTE) – Metamodels, Ontologies and Semantic Technologies (MOST), together with 15th International Enterprise Computing Conference (EDOC 2010), Vitória, Brazil, 2010.
- GUIZZARDI, G.; WAGNER, G.; GUARINO, N.; VAN SINDEREN, M., An Ontologically Well-Founded Profile for UML Conceptual Models, 16th International Conference on Advances in Information Systems Engineering (CAiSE), Latvia, 2004. Springer-Verlag, Berlin, Lecture Notes in Computer Science 3084, ISBN 3-540-22151-4

- GUARINO, N., The Ontological Level: Revisiting 30 Years of Knowledge Representation. Conceptual Modeling: Foundations and Applications 2009: 52-67
- GUARINO, N. WELTY, C., An Overview of OntoClean. Handbook on Ontologies 2004: 151-172
- GUARINO, N. WELTY, C., Evaluating ontological decisions with OntoClean. Commun. ACM 45(2): 61-65 (2002).
- Jackson, D.: Software Abstractions: Logic, Language, and Analysis. The MIT Press, Cambridge, Massachusetts (2012).

- Bauman, B. T., Prying Apart Semantics and Implementation: Generating XML Schemata directly from ontologically sound conceptual models, Balisage Markup Conference, 2009.
- U.S. Department of Defense, Data Modeling Guide (DMG) for an Enterprise Logical Data Model (ELDM), available online: http:// www.omgwiki.org/architecture- ecosystem/lib/exe/fetch.php? media=dmg\_for\_enterprise\_ldm\_v2\_3.pdf.
- Dijkstra, E.W., The Humble Programmer, Communications of the ACM, 15:10, Oct. 1972.
- Mylopoulos, J., Conceptual modeling and Telos. In P. Loucopoulos & R. Zicari (eds.), Conceptual Modeling, Databases, and CASE (Chapter 2, pp. 49–68). Wiley, 1992.



http://nemo.inf.ufes.br/ gguizzardi@inf.ufes.br