Multidimensional Models Meet the Semantic Web: Defining and Reasoning on OWL-DL Ontologies for OLAP

Nicolas PRAT, ESSEC Business School

Imen MEGDICHE, CNAM

Jacky AKOKA, CNAM & Institut Mines Télécom

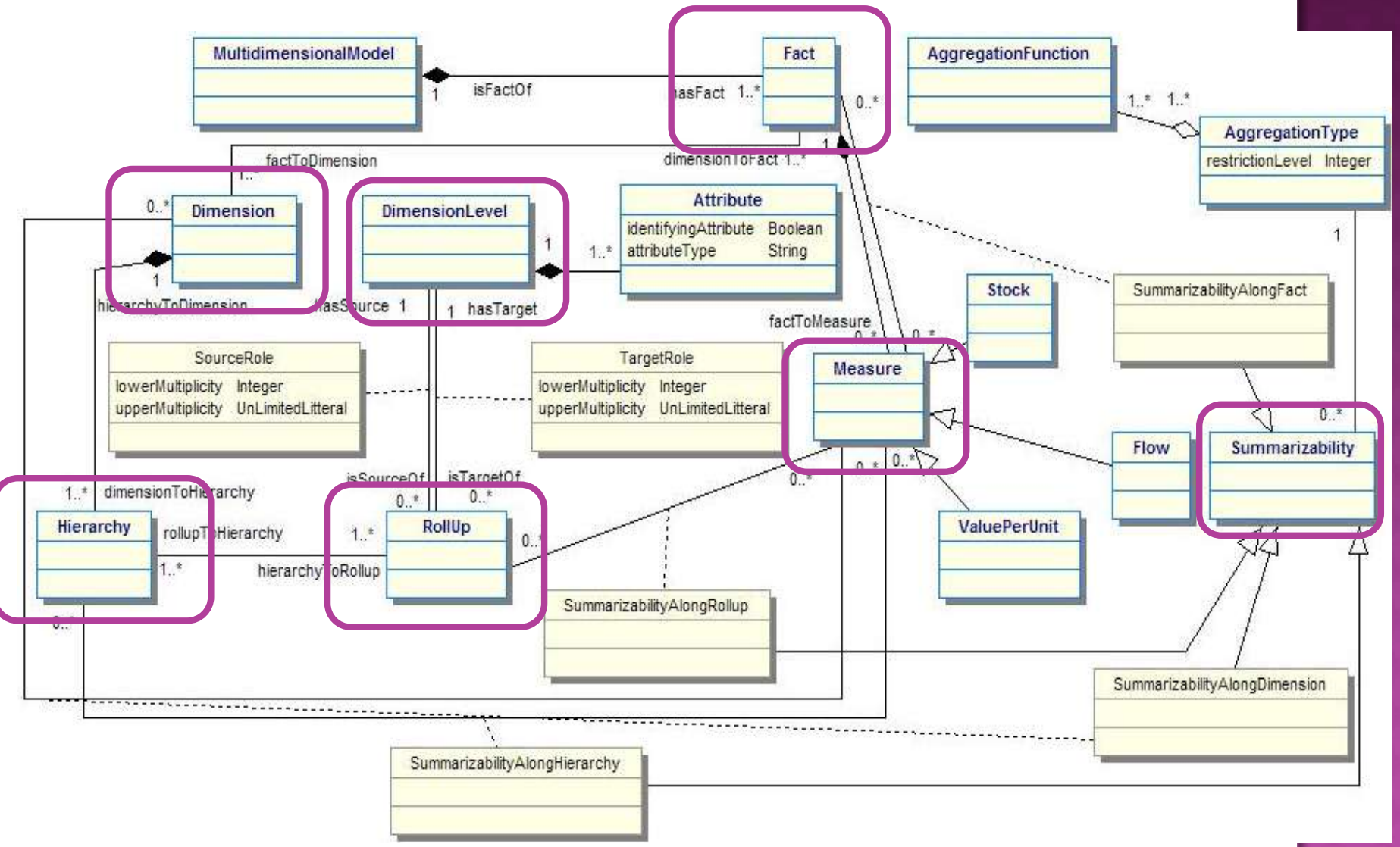
CONTEXT OF THE RESEARCH

- Purpose:
  - Check multidimensional models, in particular summarizability, to ensure correct OLAP analysis.
- Idea:
  - Check models by reasoning on these models.
  - Use the OWL-DL language=> represent multidimensional models as OWL-DL ontologies.
- Previous research:
  - Summarizability (additivity) extensively studied in the literature.
  - *However*, complete and specific mapping rules for representing multidimensional models as OWL-DL ontologies missing from the literature.

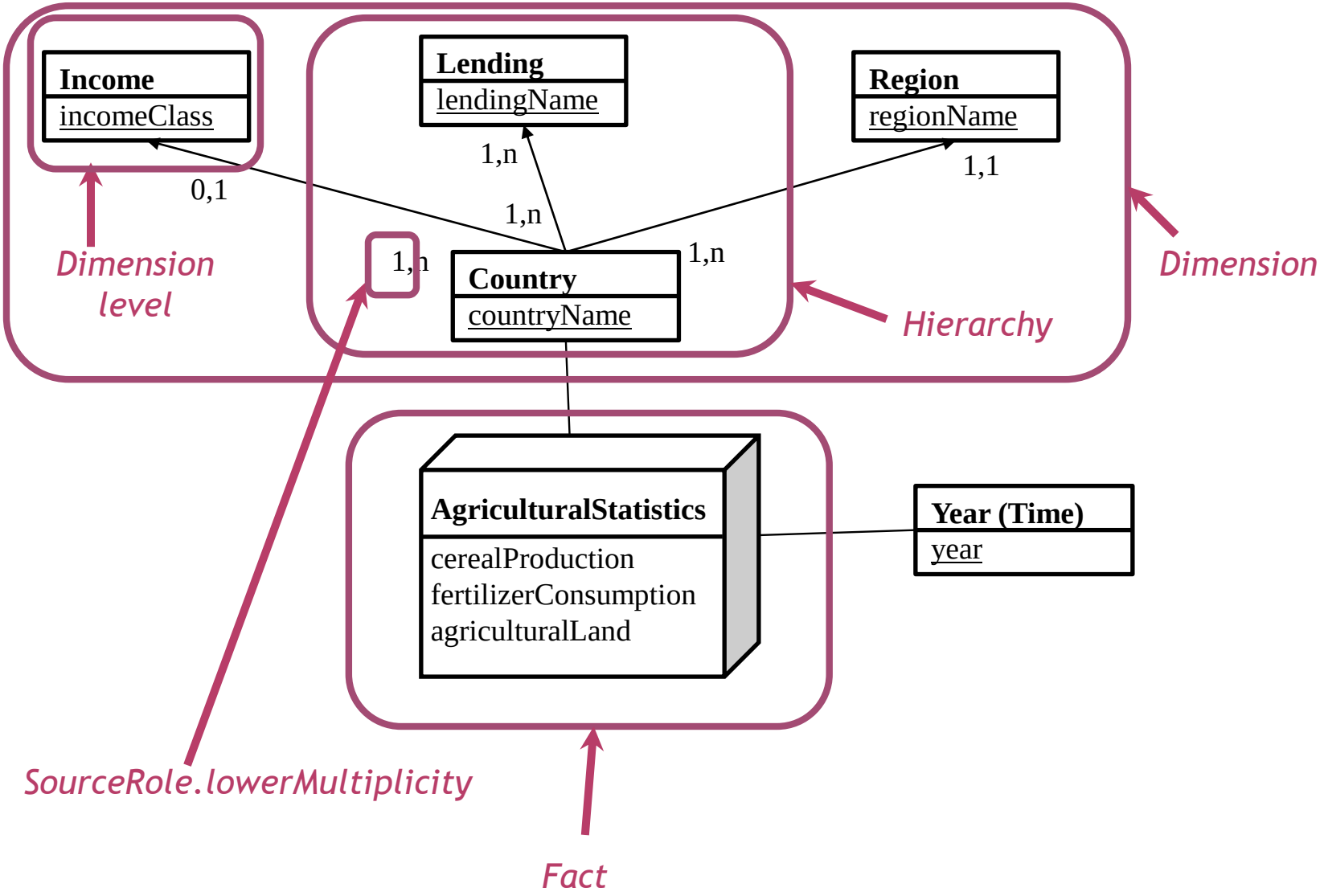
CONTRIBUTION & OUTLINE

- Our contribution:
  - Mapping transformations to represent multidimensional models as OWL-DL ontologies (extension and improvement of transformations defined in RCIS 2012 paper).
  - Reasoning on OWL-DL multidimensional ontologies to check the multidimensional models (incl. summarizability).
  - Connection with the RDF Data Cube Vocabulary for implementing our approach in the semantic Web.
- Outline:
  - Multidimensional metamodel
  - OWL-DL
  - Mapping transformations
  - Reasoning on multidimensional models
  - Application scenario
  - Complementarities with the RDF Data Cube Vocabulary
  - Conclusion and perspectives.

MULTIDIMENSIONAL METAMODEL



EXAMPLE OF MULTIDIMENSIONAL MODEL (ADAPTED FROM WORLD BANK)



OWL-DL

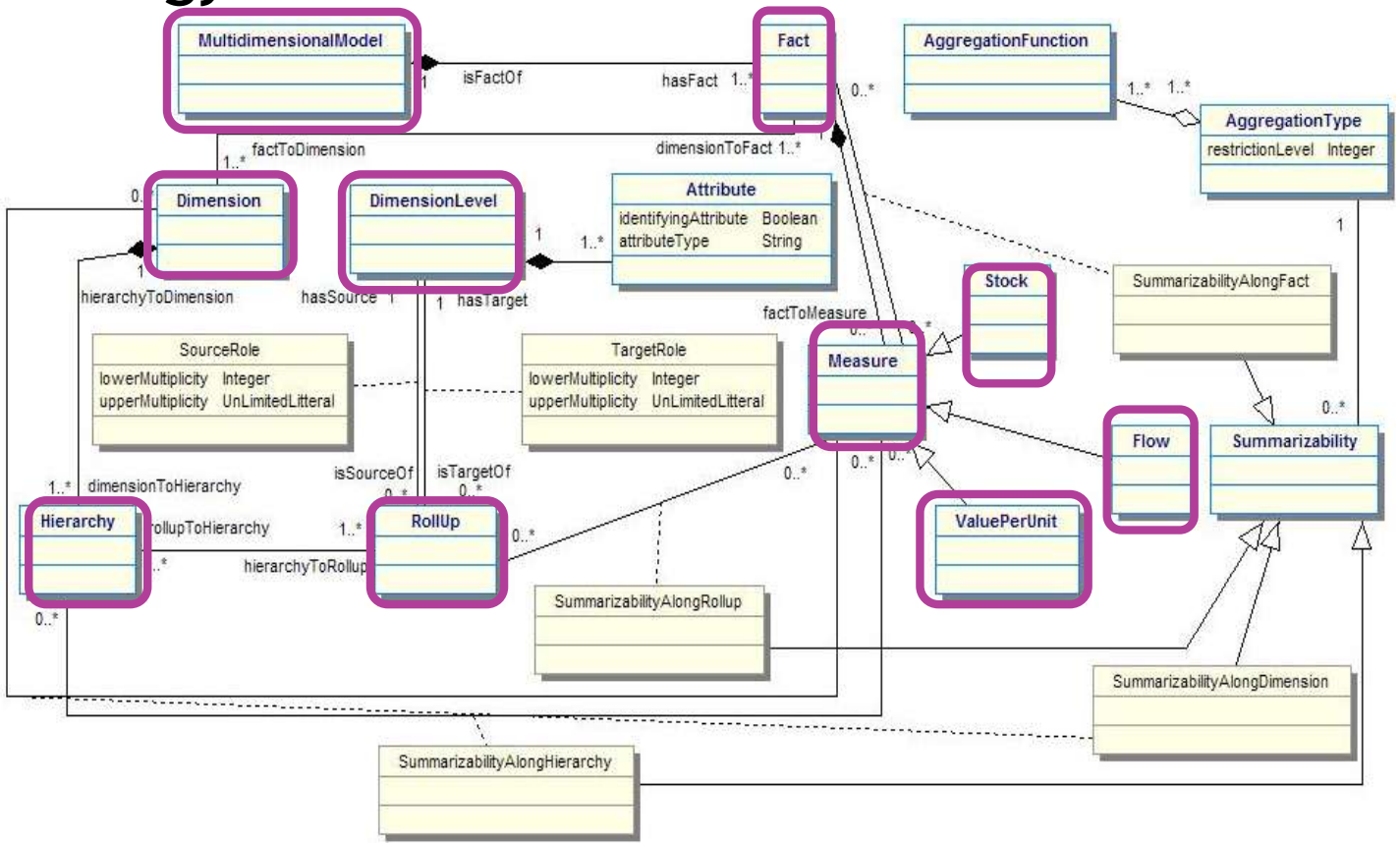
- OWL: standard for representing ontologies in the semantic Web.
- 3 variants:
  - OWL Full
  - OWL Lite
  - **OWL-DL (Description Logic) ⇨ Reasoning.**
- We use OWL v2. Includes property chains:
  - $P1 \circ P2 \subseteq P3$ (If P1(x,y) and P2(y,z), then P3(x,z))

MAPPING TRANSFORMATIONS

- Source:
  - Multidimensional model (instance of the multidimensional metamodel).
- Target:
  - OWL-DL ontology. May be implemented into ontology tool (e.g. Protégé), coupled with reasoner (e.g. Pellet).
- Two levels of transformations:
  - Metamodel-level (model-independent) transformations.
  - Model-level transformations.

METAMODEL-LEVEL TRANSFORMATIONS

- T1.1: Each class of the multidimensional metamodel becomes a class in the OWL-DL ontology.



METAMODEL-LEVEL TRANSFORMATIONS

- T1.7: To represent summarizability along a fact, a dimension, a hierarchy, or a rollup, an object property is created in the OWL-DL ontology for each aggregation function. The domain of the object property is Measure, and its range is Fact, Dimension, Hierarchy or Rollup, respectively.

$summmableAlongFact$

$T \subseteq \forall summableAlongFact.Fact \cap$

$\forall summableAlongFact^{-}.Measure$

METAMODEL-LEVEL TRANSFORMATIONS

- T1.8: Aggregation types are mapped by defining (1) the object property summableAlongDimension as a subproperty of averageableAlongDimension [...]

  $summableAlongDimension \subseteq averageableAlongDimension$

  $averageableAlongDimension \subseteq countableAlongDimension$

  $minableAlongDimension \equiv averageableAlongDimension$

  $maxableAlongDimension \equiv averageableAlongDimension$

  …

MODEL-LEVEL TRANSFORMATIONS

- T2.1: Each dimension of the multidimensional model is defined as a subclass of the class Dimension in the OWL-DL ontology.

$Dim\_Country \subseteq Dimension$

$Dim\_Time \subseteq Dimension$

MODEL-LEVEL TRANSFORMATIONS

- T2.12: For each summarizability S in the multidimensional model relating a measure M with a dimension D: (i) if the restriction level of the aggregation type is 1, for the class corresponding to M, some values of the summableAlongDimension property are from the dimension D, (ii) if the restriction level of the aggregation type is 2, for the class corresponding to M, some values of the averageableAlongDimension property are from the dimension D and no values of the summableAlongDimension property are from the dimension D, and (iii), if the restriction level of the aggregation type is 3, for the class corresponding to M, some values of the countableAlongDimension property are from the dimension D and no values of the averagebleAlongDimension property are from the dimension D.

  Similarly, this transformation is applied to each summarizability relating a measure with a fact, a hierarchy or a rollup in the multidimensional model.

COMMON AXIOMS

- Axiom 2: Measures of type stock are not additive along temporal dimensions [Lenz & Shoshani, 1997].

$$Stock \sqsubseteq \neg(\exists summableAlongDimension.TemporalDimension)$$

COMMON AXIOMS

- Axiom 3: If a measure is summable (resp. averageable, minable, maxable, countable) along a fact, then it is summable (resp. averageable, minable, maxable, countable) along every dimension dimensioning this fact. […]

$$summableAlongFact \circ factToDimension \subseteq summableAlongDimension$$

…

DEFINED CLASSES

- Defined class: A drilldown complete rollup is a rollup for which the minimum cardinality between the source dimension level and the rollup is equal to one.
- Defined class: A drilldown complete hierarchy is a hierarchy made of drilldown complete rollups only.

VERIFYING MODELS & SUMMARIZABILITY

- Increasingly-restrictive levels of verification:

  CorrectModel⊆CompleteModel

  SummarizableModel⊆CorrectModel

  StrictlySummarizableModel⊆SummarizableModel

  LevelByLevelSummarizableModel⊆SummarizableModel

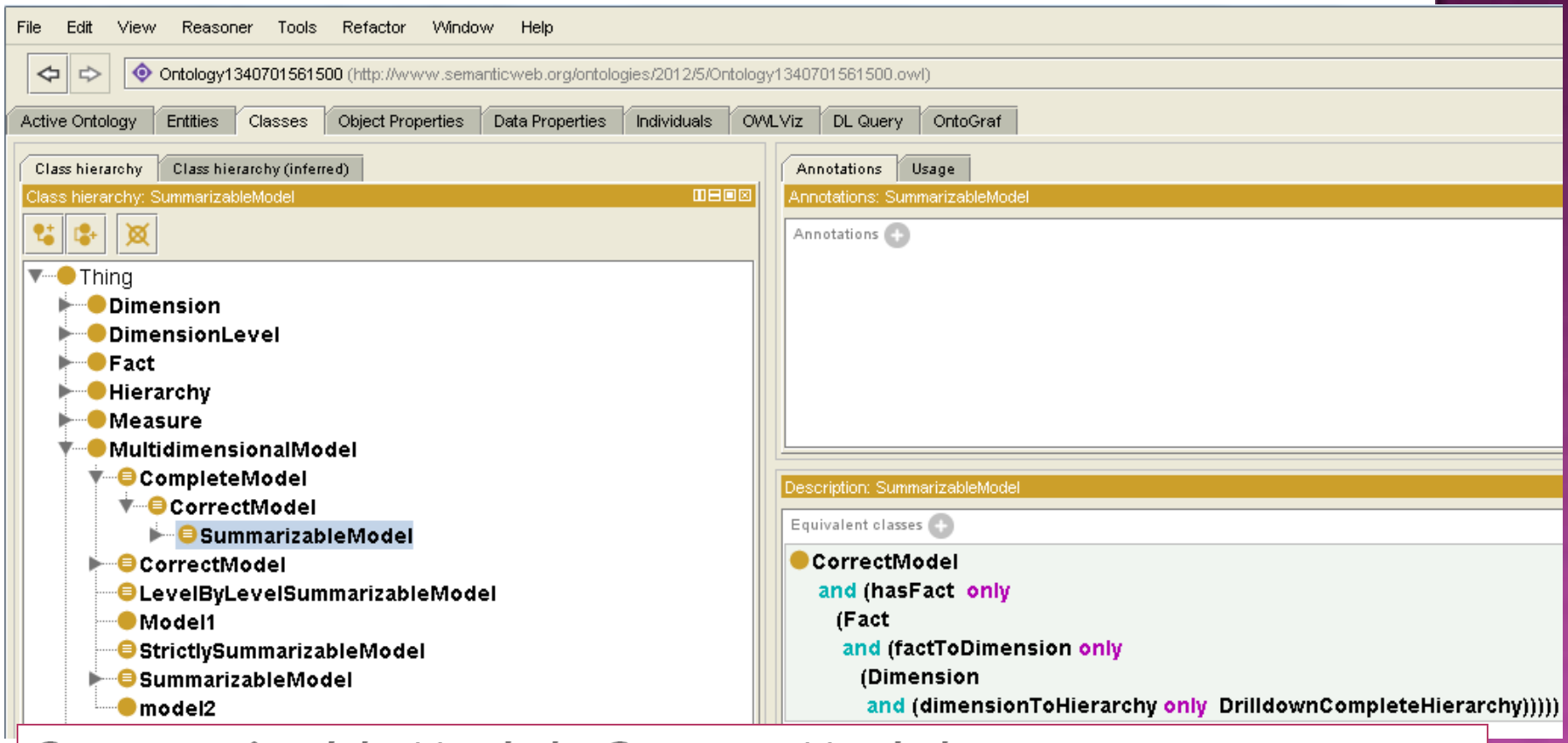
VERIFYING MODELS & SUMMARIZABILITY

- Rule 3: A summarizable model is a correct model in which all hierarchies are drilldown complete.

$$SummarizableModel \equiv CorrectModel \sqcap \forall hasFact.(Fact \sqcap \forall factToDimension.(Dimension \sqcap \forall dimensionToHierarchy.DrilldownCompleteHierarchy))$$

VERIFYING MODELS & SUMMARIZABILITY

- Rule 3 defined with Protégé:



$$SummarizableModel \equiv CorrectModel \cap \forall hasFact.(Fact \cap \forall factToDimension.(Dimension \cap \forall dimensionToHierarchy.DrilldownCompleteHierarchy))$$

SCENARIO

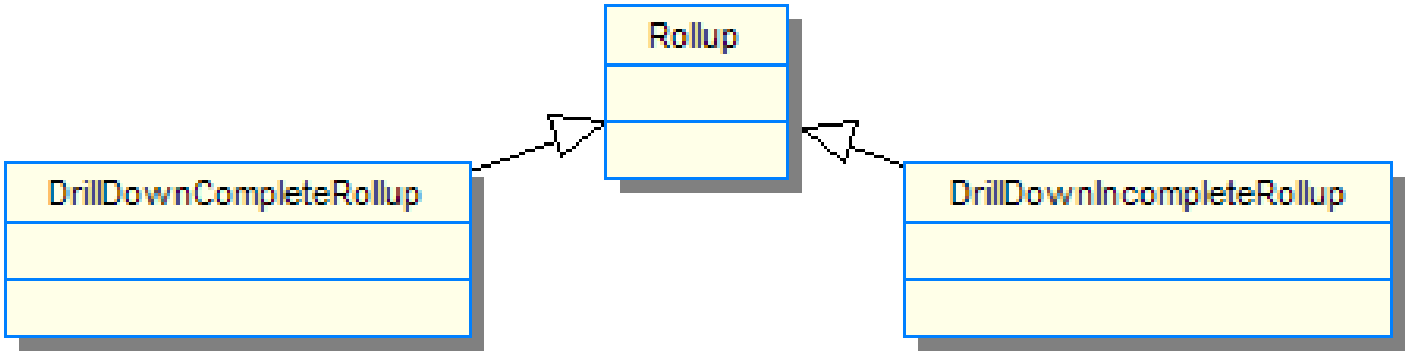
```
Axiom: Model1 subClassOf CompleteModel

Explanation(s):
1  cerealProduction subClassOf measureToFact only AgriculturalStatistics
2  CompleteModel equivalentTo MultidimensionalModel
      and hasFact only Fact
           and factToRollup only Rollup
              and isCountableAlongRollup some Measure
                 and measureToFact only Fact
                     or isSummableAlongRollup some Measure
                        and measureToFact only Fact
                            or isAverageableAlongRollup some Measure

3  Model1 subClassOf MultidimensionalModel
4  AgriculturalStatistics subClassOf Fact
5  RollupCountryLending subClassOf isSummableAlongRollup some cerealProduction
6  AgriculturalStatistics subClassOf factToRollup only RollupCountryLending
                                                        or RollupCountryRegion
7  RollupCountryRegion subClassOf isCountableAlongRollup some cerealProduction
8  Model1 subClassOf hasFact only AgriculturalStatistics
9  RollupCountryLending subClassOf Rollup
10 RollupCountryRegion subClassOf Rollup
11 cerealProduction subClassOf Measure
  ...
```

MULTIDIMENSIONAL METAMODEL VS RDF DATA CUBE VOCABULARY

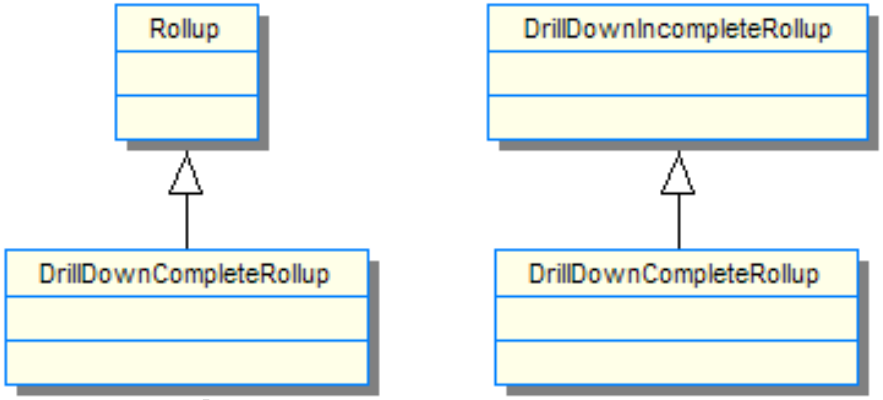
Concept (from ontology of multidimensional structures)	Multidimensional metamodel	RDF Data Cube vocabulary
DimensionElement	DimensionLevel	
Cube	Fact	DataSet
Dimension	Dimension	dimension
Measure	Measure	measure
Attribute	Attribute	
HierarchyLevel	Attribute (identifyingAttribute)	
Fact		Observation
Aggregation	AggregationFunction, AggregationType, Summarizability	
DrillDown	RollUp (hasSource)	
RollUp	RollUp (hasTarget)	
SliceAndDice		Slice

SOME ISSUES

- The logic of OWL-DL challenges our intuition.
  - Example:



  - vs:



- Incomplete reasoning on property chains at the type level (Pellet).

CONCLUSION

- Contribution of the paper:
  - Extension/improvement of mapping transformations to represent multidimensional models as OWL-DL ontologies.
  - Reasoning on OWL-DL multidimensional ontologies to check the multidimensional models (incl. summarizability).
  - Connection with the RDF Data Cube Vocabulary for implementation in the semantic Web.
- Future work:
  - Application to more extensive examples.