

How class axioms affect instance data

Viktoria Pammer

Graz University of Technology & Know-Center, Austria



Example

ABox: Facts

Crete is an island which is located in greece and in the mediterranean sea.

$Island(crete)$
 $locatedIn(crete, greece)$
 $locatedIn(crete, mediterranean)$

Existing TBox: General Knowledge

Two types of natural areas are described: land and water areas. Islands are land areas. There are also nations, which are geopolitical areas. All these are geographic areas.

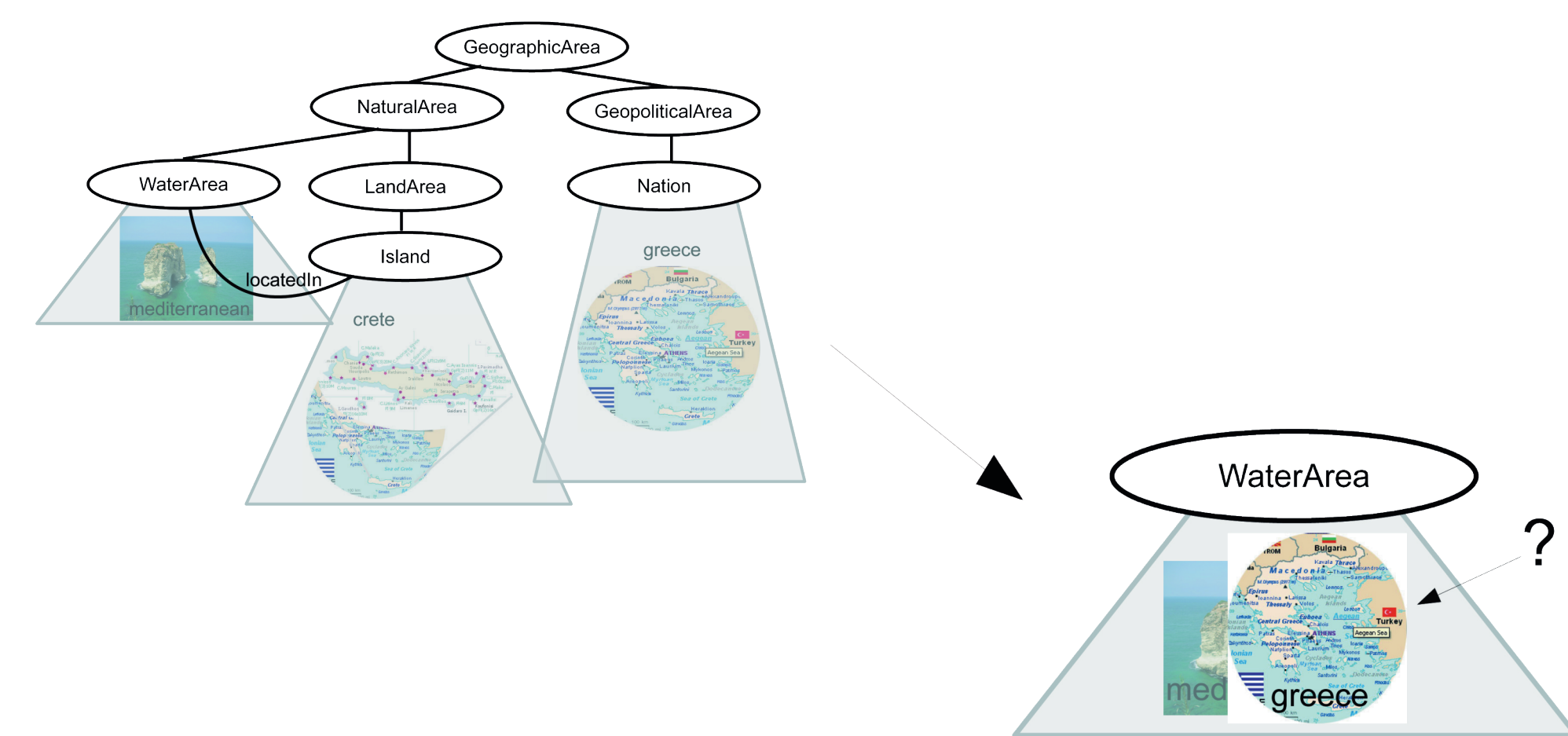
$GeopoliticalArea \sqsubseteq GeographicArea$
 $Nation \sqsubseteq GeopoliticalArea$
 $NaturalArea \sqsubseteq GeographicArea$
 $LandArea \sqsubseteq NaturalArea$
 $Island \sqsubseteq LandArea$
 $WaterArea \sqsubseteq NaturalArea$

New axioms: Additional details about Island

Every island is located in a water area. Everything in which an island is located must be a water area.

$Island \sqsubseteq \forall locatedIn. WaterArea$
 $Island \sqsubseteq \exists locatedIn. WaterArea$

A knowledge base consists of a TBox (terminological axioms) and an ABox (assertional axioms). Assume that the ABox and the existing TBox above are a small excerpt from an existing knowledge base KB . A knowledge engineer extends the TBox of the KB in order to more formally describe the concept $Island$.



Effects on assertional axioms

Example

- ▶ $WaterArea(mediterranean)$
- ▶ $WaterArea(greece)$
- ▶ $(\neg Island \sqcup \forall locatedIn. WaterArea)(mediterranean)$
- ▶ $(\forall locatedIn. WaterArea)(crete)$

Is $WaterArea(greece)$ intended?

Is $(\neg Island \sqcup \forall locatedIn. WaterArea)(mediterranean)$ informative?

Questions a knowledge engineer could ask

1. Does the extended KB entail any new facts about individuals (assertional axioms)?
2. Which new facts about individuals are entailed?

Computation

- ▶ Computing effects on assertional axioms is an open issue if complex concept assertions like $(\forall locatedIn. WaterArea)(crete)$ are considered. In \mathcal{ALC} already, the set of such types is infinite in general.

Meaningfulness

- ▶ The axiom $Island \sqsubseteq \forall locatedIn. WaterArea$ states that all islands are located in some water area. This entails that every instance is either not an island or is located in some water area. Which kind of complex concepts are meaningful / informative to a knowledge engineer?



Effects on models

Example 1

Assume the KB also contains the fact that kos is an $Island$.

- ▶ In all models of the extended KB, kos is located in something which is a $WaterArea$.

Example 2

Assume the KB is extended with $WaterArea \equiv (StagnantWater \sqcup FlowingWater)$.

- ▶ In all models of the extended KB, the $mediterranean$ sea is either a $StagnantWater$ or a $FlowingWater$ (or both).

In both cases, models of the original KB did not necessarily contain this knowledge.

Questions a knowledge engineer could ask

1. Do all models of the extended KB contain additional knowledge?
2. Which knowledge is added?



Computation

There is an infinite number of models for one KB in the general case.

- ▶ A procedure deciding whether there are effects could rely on minimal models.

No unique name assumption (UNA)

In the semantic web, the UNA should not be made: If Alice calls something $mediterranean$ and Bob calls something $mediterraneanSea$, these instance names could refer to the same entity in the models.

- ▶ Computing minimal models is highly complex if the UNA is not made.

Application

Useful in cases where ...

- ▶ A knowledge base is constructed or extended collaboratively.
- ▶ A large knowledge base is constructed or extended.

Formal modelling: Test suitability of TBox to a known kind of data

1. Collect typical facts of the domain that shall be held by the KB, and keep them as testdata
2. Design an ontology (terminological axioms)
3. Populate the ontology with testdata and compute effects of the TBox on the testdata.
4. Have domain experts validate the effects (inferred knowledge)

Knowledge base refinement/extension: Ascertain relevance/harmlessness w.r.t. given data

1. Refine knowledge base: Add terminological axioms to existing KB
2. Compute effects of an extended TBox on existing data in the KB
 - 2.1 If there are no effects: New TBox is potentially redundant/has no undesirable side-effects on existing data.
 - 2.2 If there are effects: Have domain experts validate the effects

Outlook

Effects on assertional axioms:

Which complex types should be considered?

Effects on models:

Efficient computation for effects on models.

Usability:

Group individuals according to types and compute effects only on exemplary individuals.

Some more details...

- i) There is a strong relation between what I call "effects on assertional axioms" and conservative extension in description logics: Conservativity is a precondition for new axioms to be "effective".
- ii) All effects on assertional axioms are also effects on models.
- iii) Computing minimal models without UNA: For 10 individuals, there are already 115975 possible combinations.

Acknowledgements

This work has been partially funded under grant 027023 in the IST work programme of the European Community. The Know-Center is funded within the Austrian COMET Program - Competence Centers for Excellent Technologies - under the auspices of the Austrian Ministry of Transport, Innovation and Technology, the Austrian Ministry of Economics and Labor and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.