

# **Ontology-Based Lexicon for Creation of Domain Ontologies**

**Kiril Simov**

Linguistic Modelling Laboratory  
Bulgarian Academy of Sciences

GL2009

19th September, 2009, Pisa, Italy



# Plan of the Talk

- Creation of domain ontology
- Definition of *ontology-to-text* relation
- Problems with the model
- Middle layer ontology and lexicon



# Role of the Ontology

Ontology has to support

- Semantic annotation of text documents
- Semantic search
- Representation of application data
- Interaction with the users



# Creation of Domain Ontology

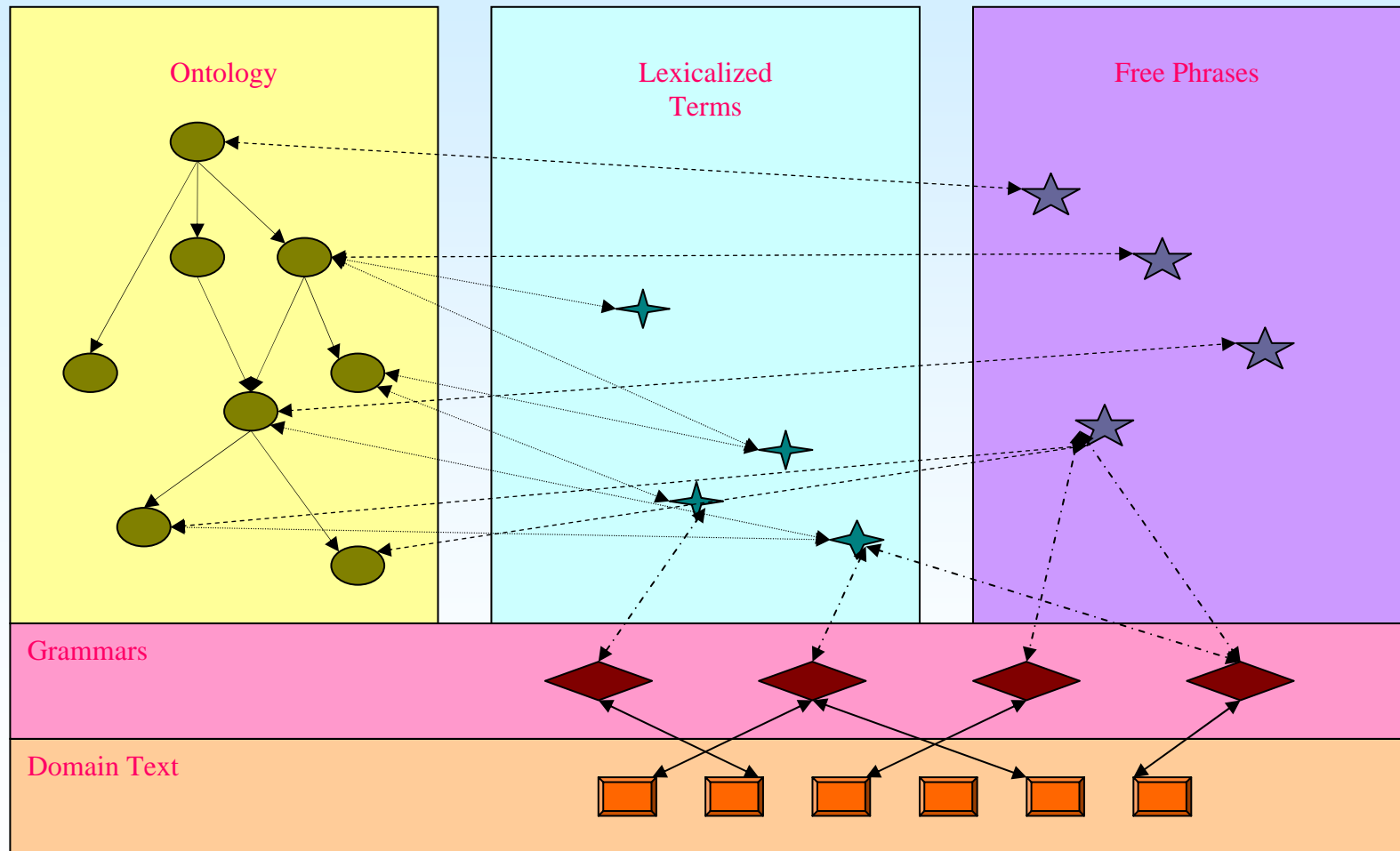
The main steps are:

- Linguistic processing of domain texts (industrial standards, text books, vocabularies, etc)
- Formalization of the terms
- Link to an upper ontology (in our case DOLCE via OntoWordNet)
- Evaluation and Documentation
- Lexicons and concept annotation grammar creation

# Ontology-to-Text Relation

- Ontology is the repository for word senses
  - **Polysemy and metonymy are encoded as interrelated concepts**
- Lexicon represents the relation between word sense (concept, relation, instance in the ontology) and other lexical knowledge – morpho-syntactic features, etc
  - **Human oriented features**
- Grammar represents the relation between lexical items in the lexicon and their realization in the text

# Ontology-to-Text Relation (2)



# Problems with the Model

- Domain semantic annotation is sparse
  - 8 concepts within 100 tokens
  - 14.8 tokens per sentence
  - 1.19 concepts per sentence
- Upper ontology is hard to be presented to users
- Solutions
  - Better annotation grammar – Co-reference, etc
  - Interaction with general lexica – WordNet
  - Middle layer ontology

# Middle Layer Ontology

- Concepts between the upper ontology and domain ontologies
- Good abstraction over the domains
- More comprehensible as a representation to the end users
- Better granularity for semantic annotation
- Event sub-hierarchy

# Middle Layer Ontology (2)

- Concepts have to be defined precisely with respect to the general world knowledge (relations and axioms)
- Concepts have to meet the requirements of an ontology creation methodology (e.g. OntoClean)
- Is it possible to define middle layer concepts that are "neutral" with respect to an upper ontology?
- Is it possible to be language independent?

# MLO Ontology-based Lexicon

- General lexica aligned to the MLO via two relations – *equality* and *subsumption*
- It has to support the following tasks
  - Semantic annotation on the level of the upper and middle layer ontology
  - User understanding of the ontology
  - Mapping from a domain ontology to the upper ontology
- Domain lexicons are extension of MLO lexicon

