



From Natural Language Requirements to a Conceptual Model

Christian Kop, Günther Fliedl, Heinrich C. Mayr
Alpen-Adria-Universität Klagenfurt



Introduction

- Talk focuses on the transformation from requirements statements (nat. lang. sentences) to a conceptual model
- It will show the possibilities and the problems during this transformation
- It will introduce an interlingua for communication with the end users



Content

- Linguistic Processing and Interpretation
- Problems of Interpretation
- The Interlingua
- Defect Detection
- Mapping



Linguistic Processing Techniques

- Parsing or
- Tagging
- Chunking
- Usage of linguistic resources (lexicons) and or ontologies for syntactic (“semantic”) features of words



Interpretation

- Analyses the structure provided by the linguistic processing step;
- provides proposals for the target (e.g. conceptual) model

Classes

- Individual nouns
- Noun phrase (adjective + noun) → specialized class

Instance

- Proper noun



Relationship types

- Verbs with their nouns
- Prepositions (with, of ..)

Attributes

- Derived by „specials rules“ i.e., nouns which occur in a certain context.
- Example.: Chen's Rule 7: „*The X of Y is Z* and Z not a proper noun“ (e.g. adjective) → X is attribute of Y and has value Z



Other rules

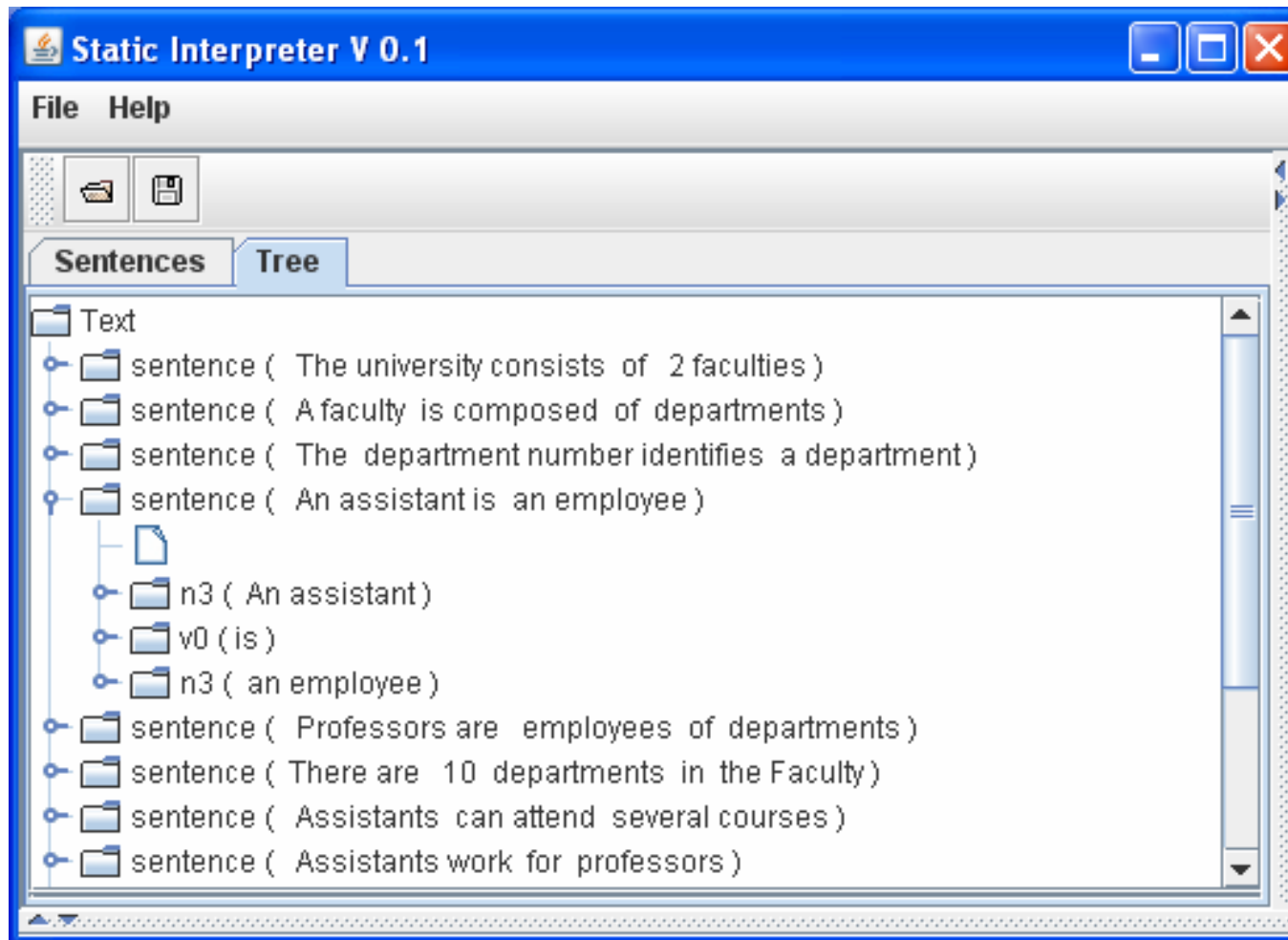
- cardinalities
 - Quantification description (many, 1, 5) before the syntactic object
- Specialization (is-a)
 - verb „is“ in combination with article “a/an“.
 - Adjective – Noun combination
 - Even prepositions



Interpretation summary

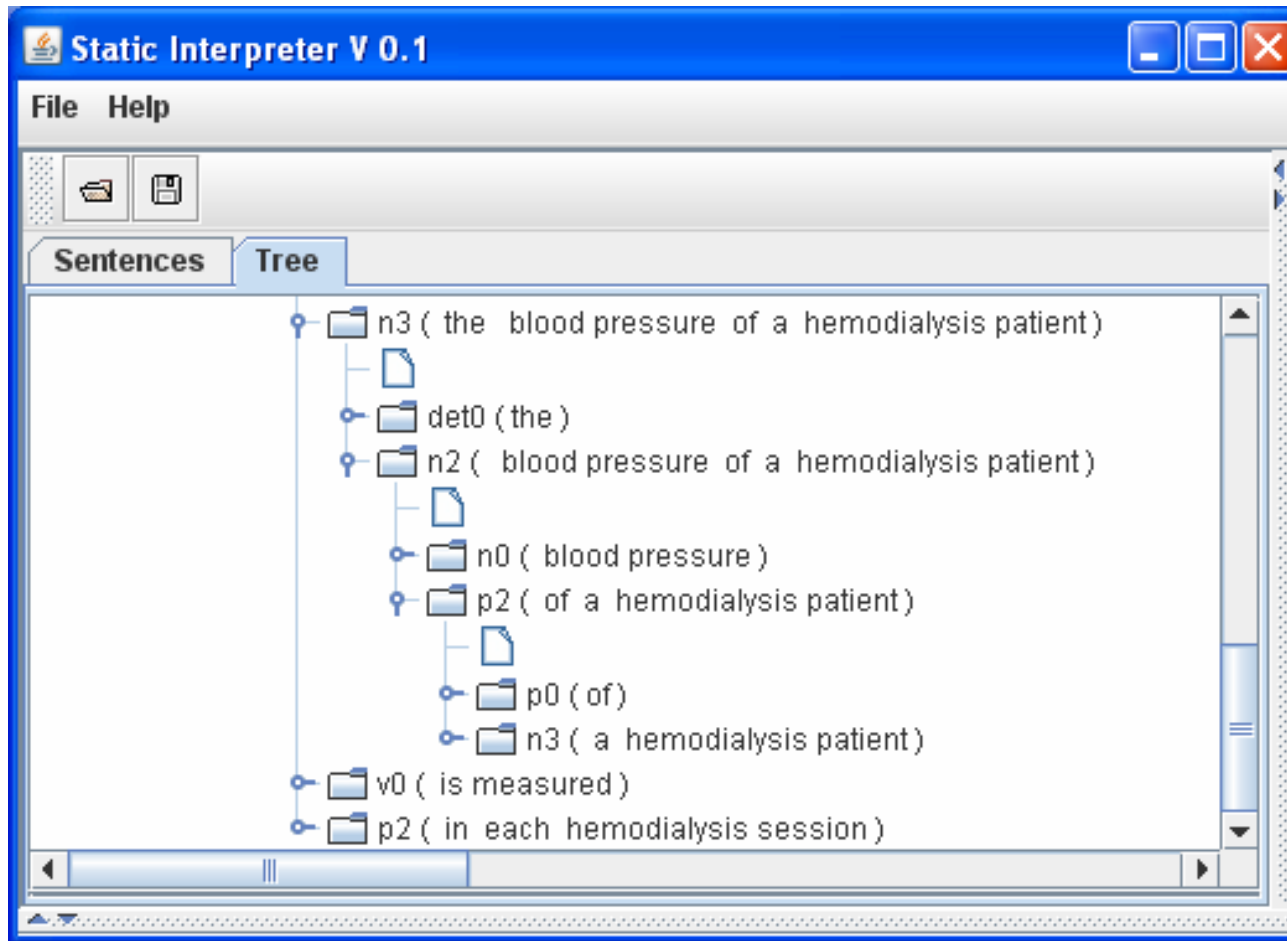
- The grammar of the natural language is the basis
- Grammatical pattern (nouns, verbs etc) are mapped to elements of the target language (e.g., classes associations etc.)
- Mapping is purely executed on syntactic criteria and patterns in both „languages“
- The specific way of how to do this depends on the concrete context of the source and target language

Source Language – natural language / NTMS





Source language – natural language / NTMS





Problems / 1

- Different interpretations for one and the same natural language pattern (e.g., adjective + noun)
- Rules to distinguish attributes from classes cannot be applied in general.
- Do I get all needed information during interpretation?



Problems / 2

- No guarantee that sentence information is relevant for the target schema.
- No guarantee to get a complete schema elements out of a sentences.
- No guarantee to get complete schema from all sentences.
- No guarantee to get a consistent schema from all sentences



Problems / 2

- Mapping rules from natural language sentences to a conceptual schema can only rely on syntactic features
- What was the intention of the requirements holder?



Solutions

- Let stakeholders (requirements holder) decide
- Introduce an Interlingua

Decision support

The screenshot shows the 'Static Interpreter V 0.1' application. The main window contains the following elements:

- Menu Bar:** File, Help
- Top-Left Pane:** Contains the sentence: "With regard to the monitoring of blood pressure n".
- Top-Right Pane:** Contains a table with the following data:

Relation	Concept 1	Concept 2
<monitoring>-of-[?..?]<blood pressure measurement>	monitoring	blood press...
<regard>-to-[?..?]<monitoring>	regard	monitoring
<hemodialysis patient>-has-[?..?]<blood pressure>	hemodialysi...	blood press...
<date>-is measured -[?..?]<blood pressure>	date	blood press...
- Bottom Pane:** Contains a table with the following data:

Connectiontype	Concept 1	Concept 2
<hemodialysis patient>-has-[?..?]<blood press...	hemodialysis patient	blood pressure

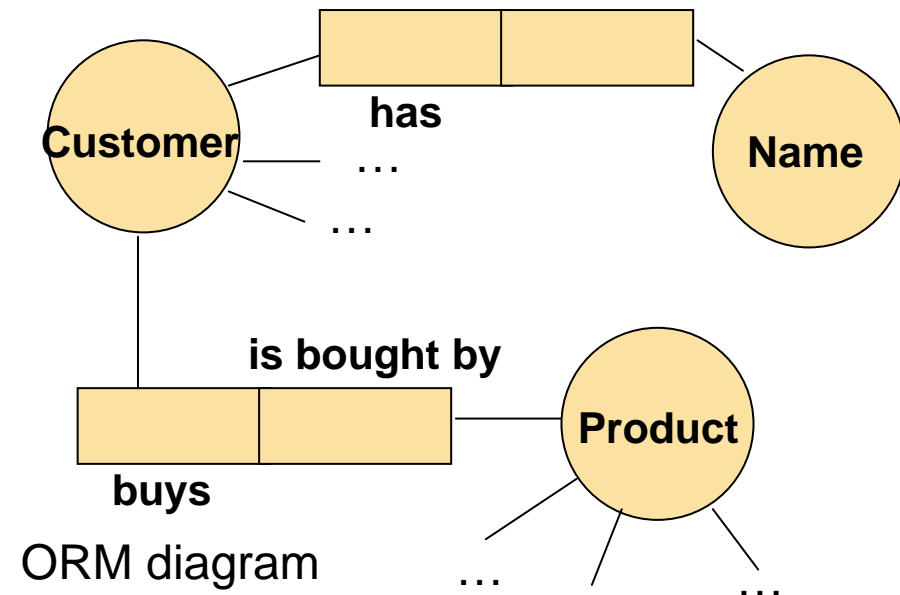
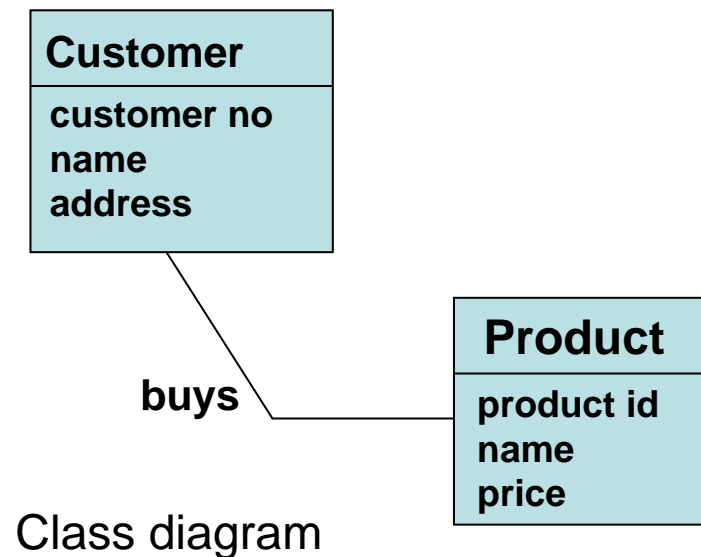


Goals of the Interlingua

- Original goals
 - Support to distinguish between classes and attributes
 - Mapping to different ER- and UML dialects
 - Understandable for different kinds of stakeholders
- Possibility to support detection of incompleteness in a schema

Paradigms

- ER und OO approaches (e.g.: class diagram)
- Fact oriented (e.g. ORM)





Pro and cons of these paradigms

- Class diagram
 - Compact look and feel (+)
 - Changes necessary (class \leftrightarrow attribute) (-)
- ORM
 - No changes necessary (class / attribute) (+)
 - More space necessary to draw it (-)



Requirements for an Interlingua (KCPM)

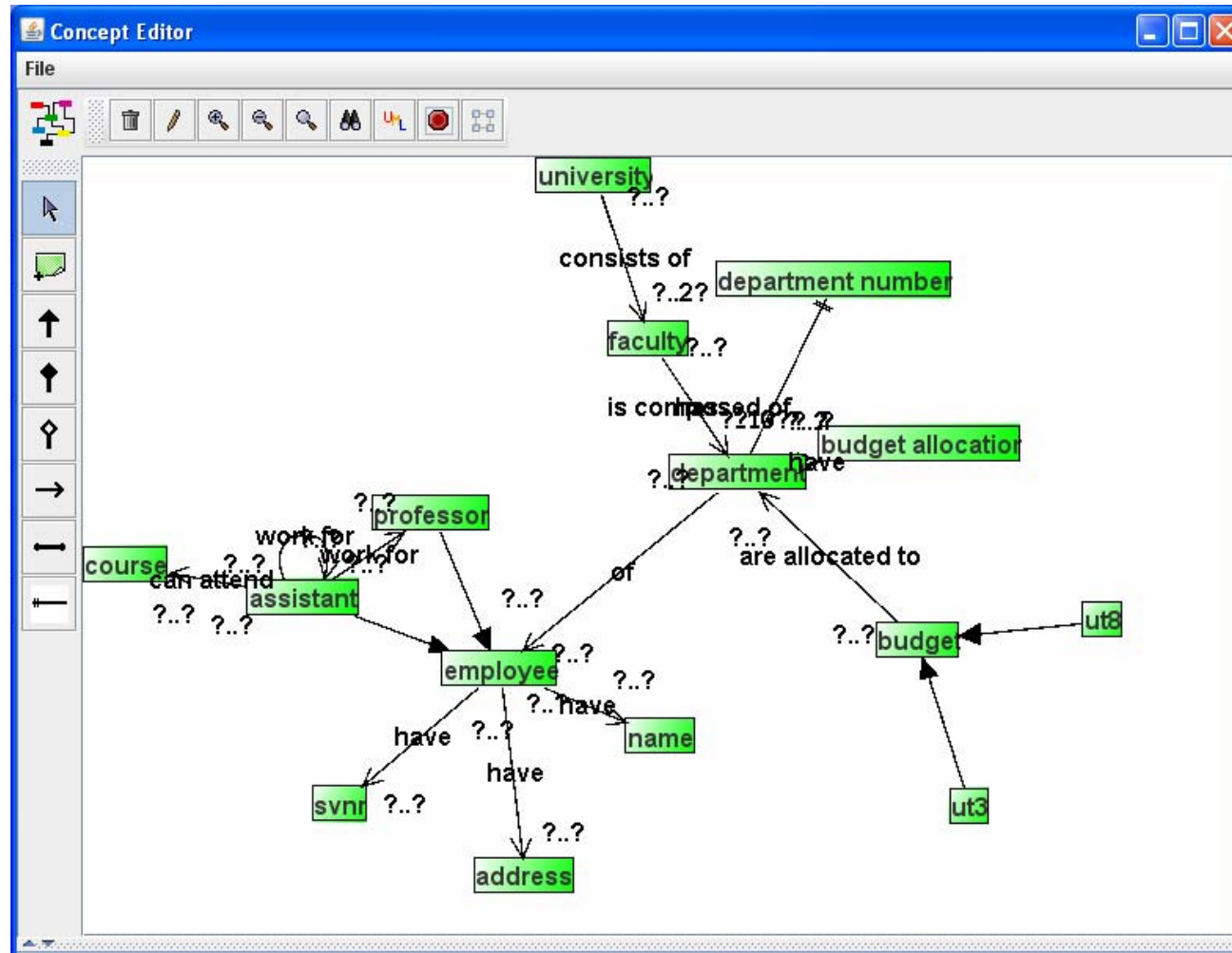
- Simpler extraction of concepts → therefore fact oriented paradigm
- Easy transformation to a class diagram
- Not too technical
- Introduction of modeling notions which are necessary in the early phase.



Modeling elements of the Interlingua

- Concept (Thing type)
 - Example.: *customer, customer name, contract, idea, customer ID*
 - subordinated elements: *example, synonym, quantity description, definition, value constraint*
- Connection type
 - Example.: *customer <- buys -> product*
 - Subordinated elements: *perspective, cardinality*
- Source (sentences)

Example: University

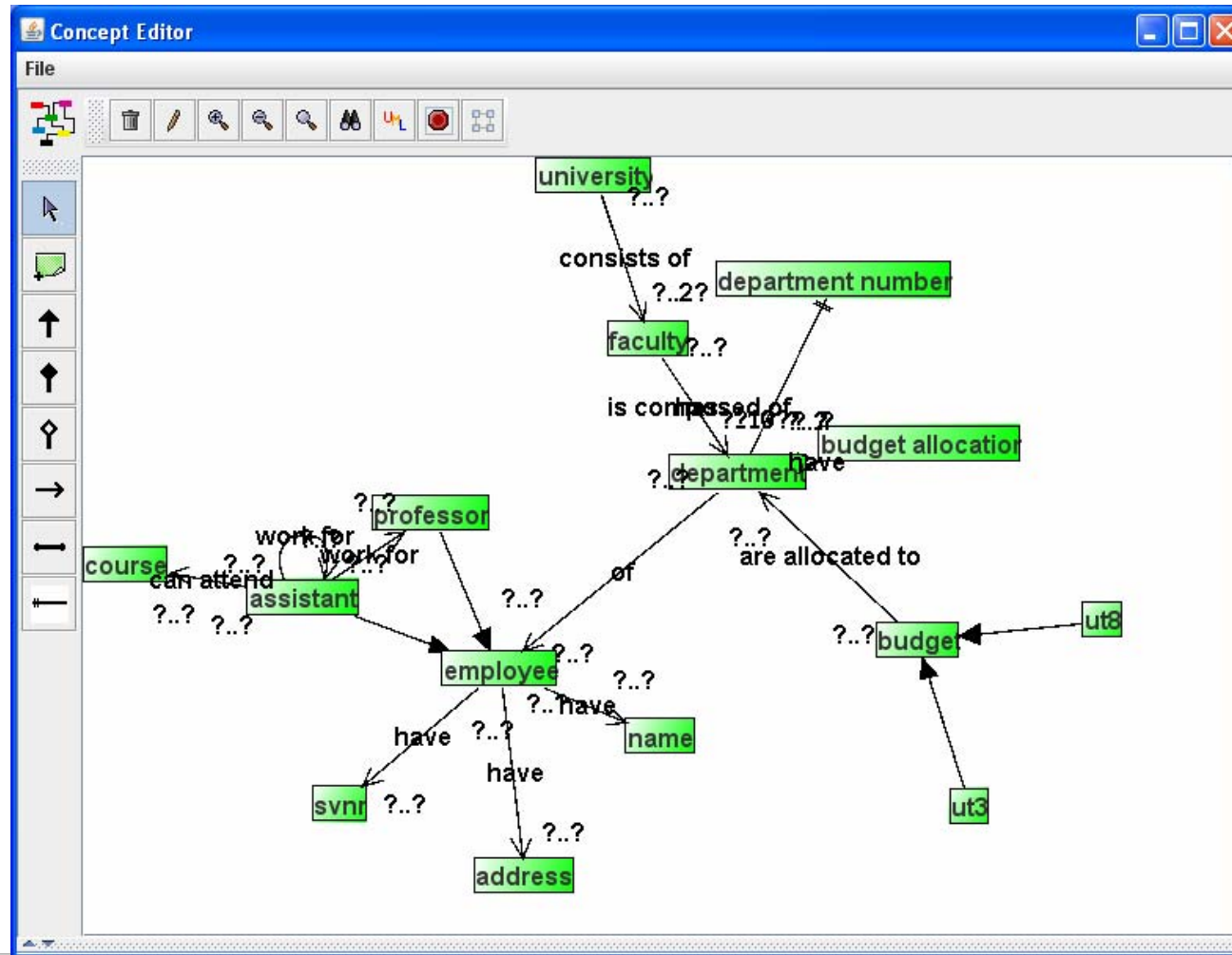




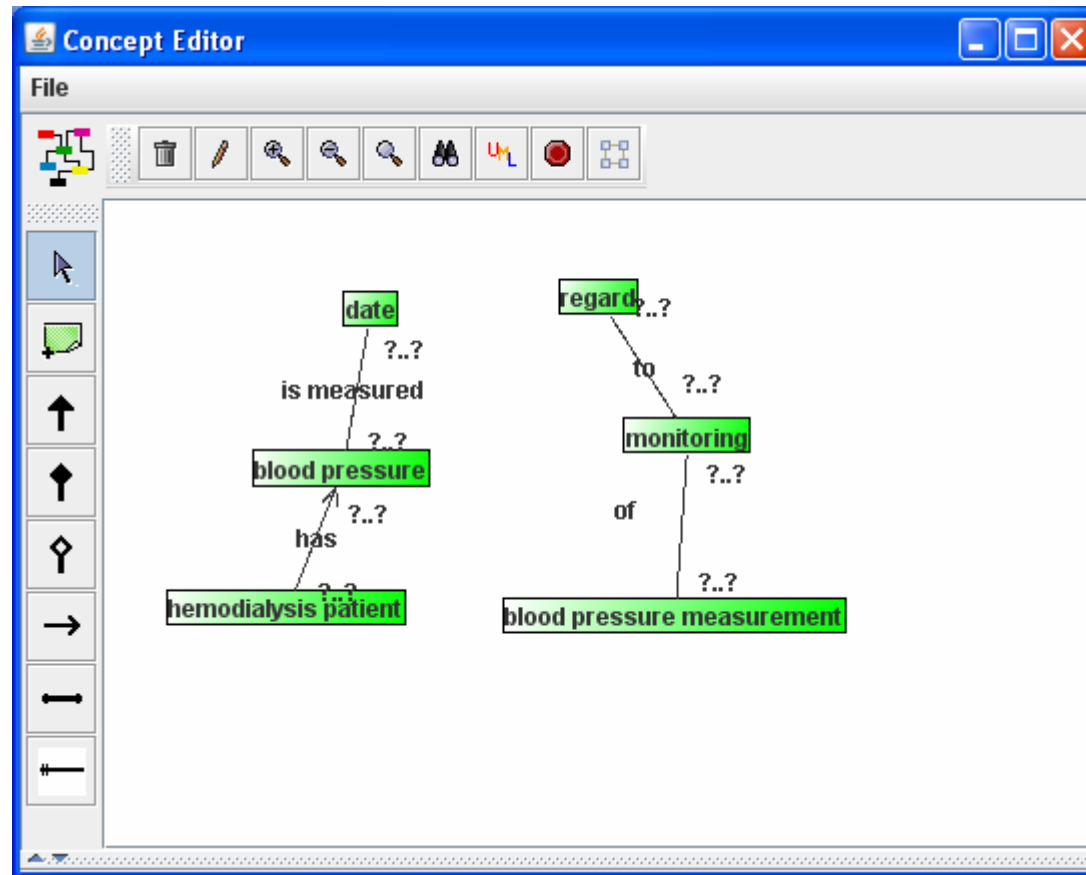
Detection of defects

- Multiplicities
- Class without attributes
- Central Concepts

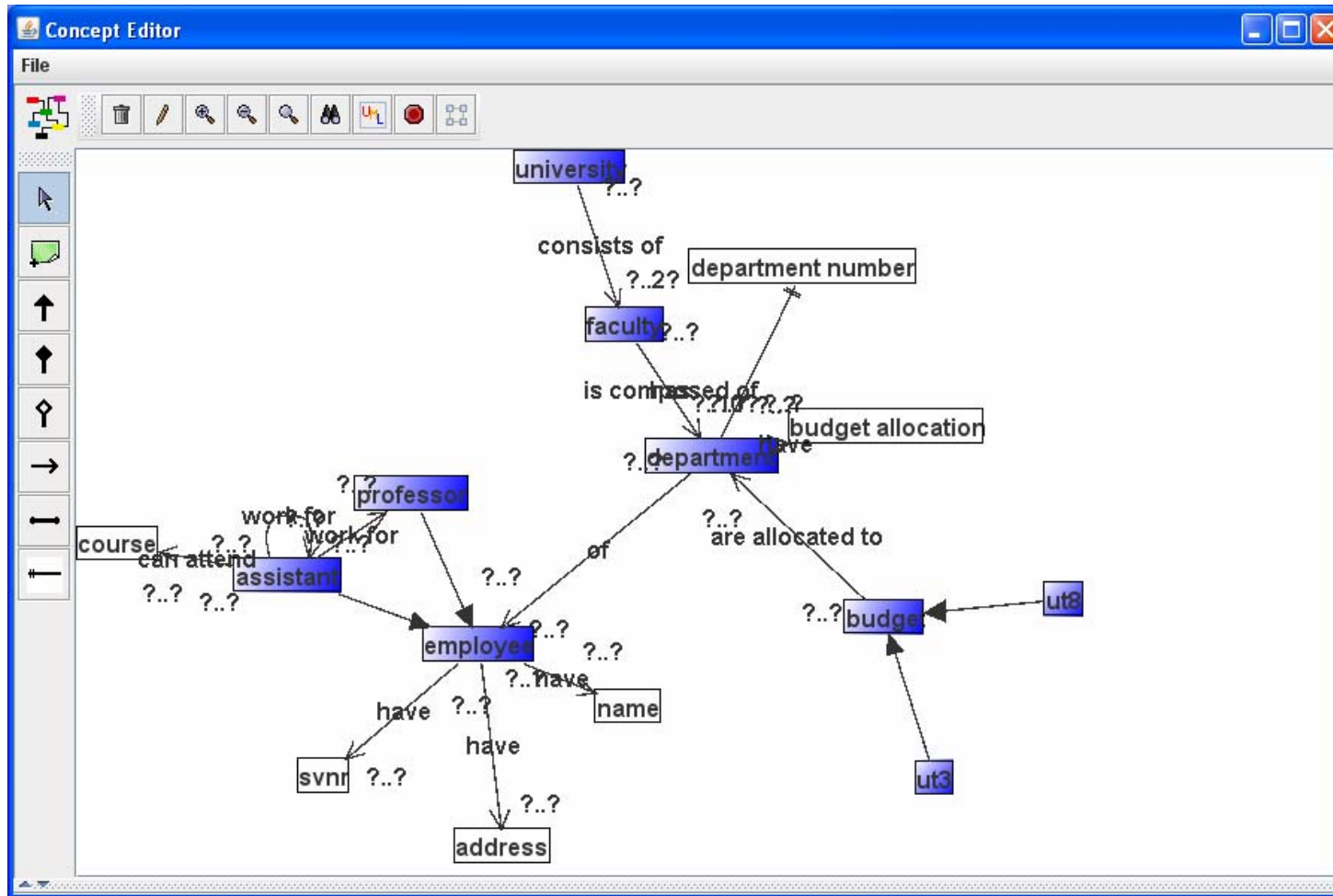
Missing Multiplicities



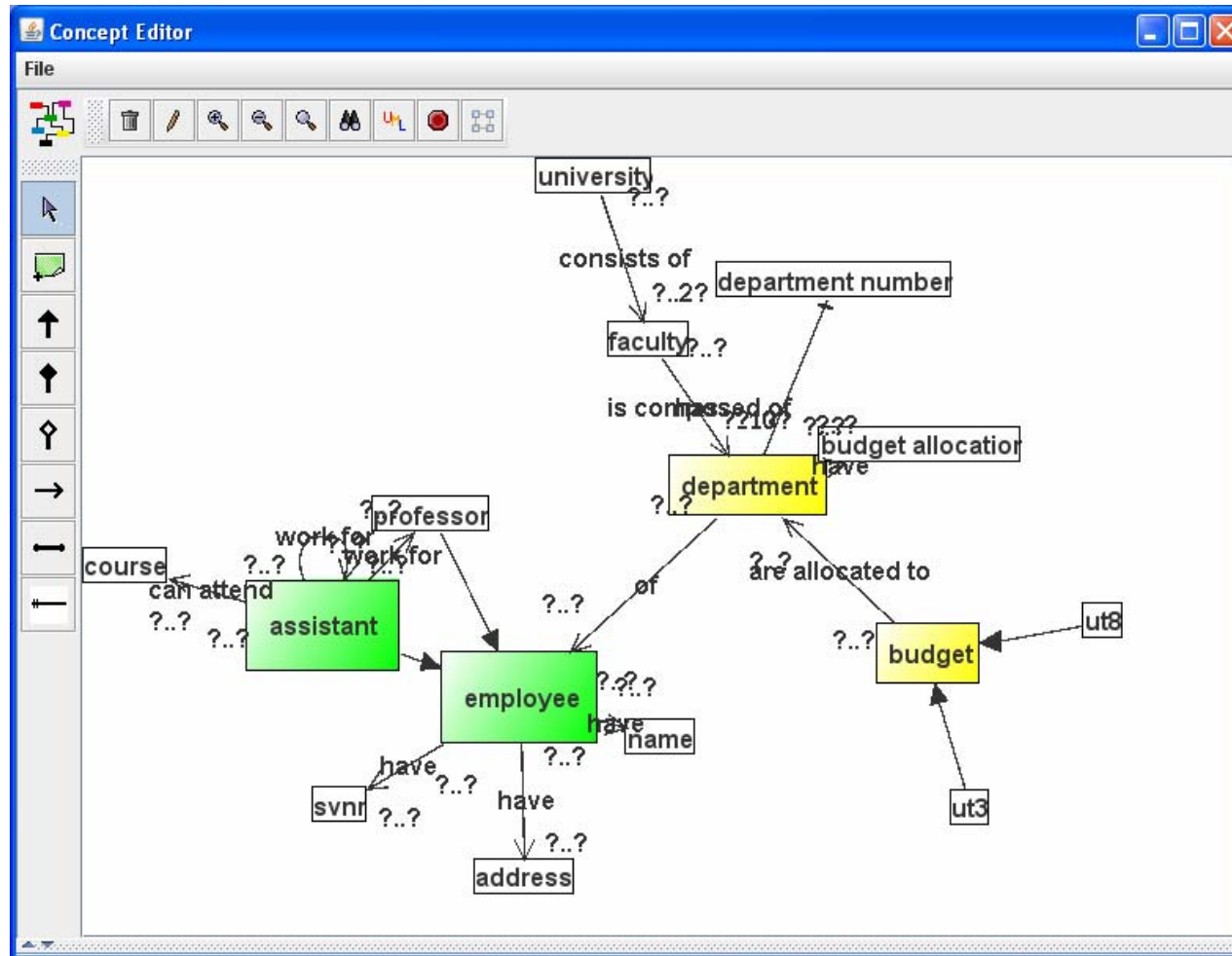
Multiplicities, more than one graph



Classes without attributes



Central concepts





Mapping information

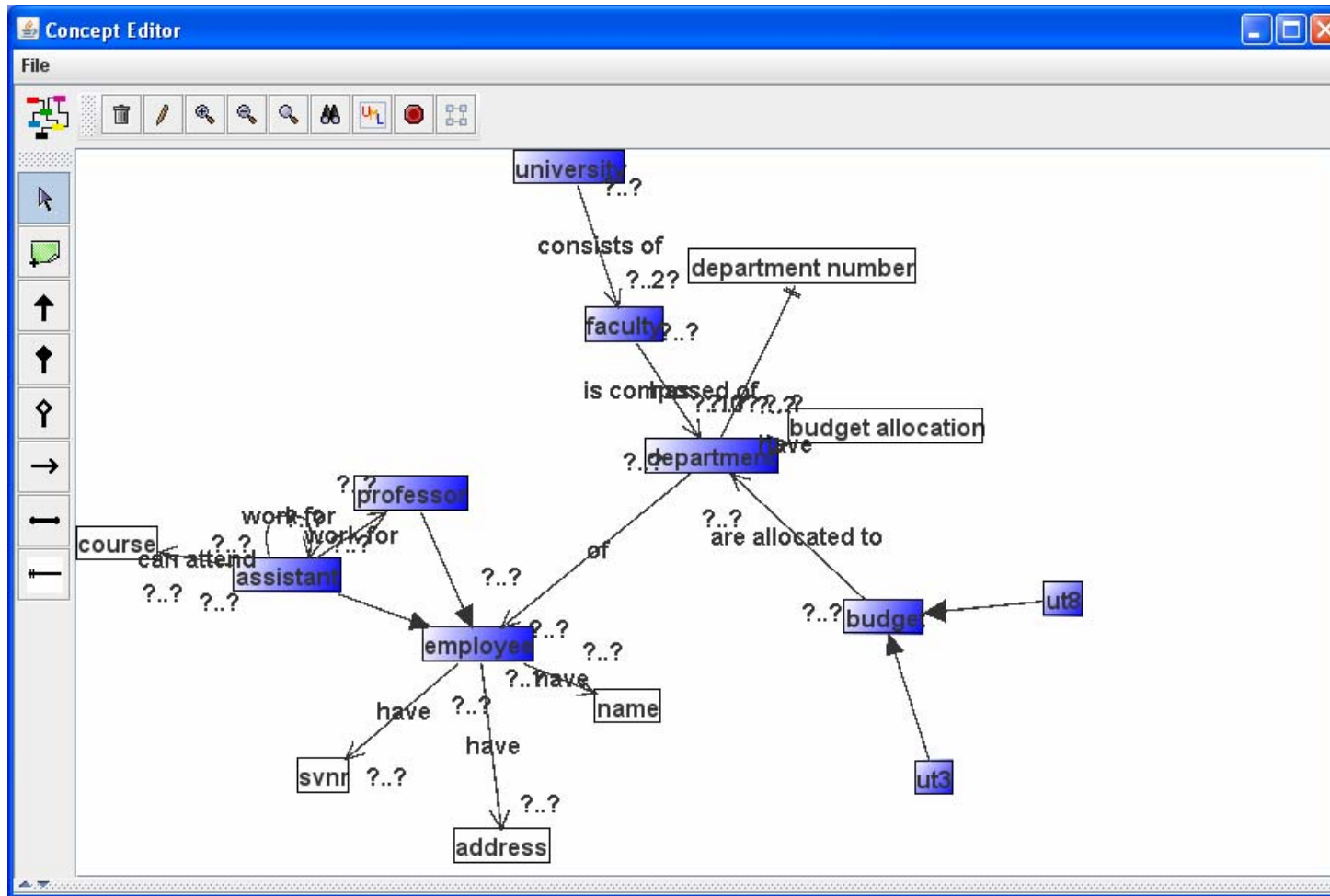
- Labels of thing types
- Labels of connection types
- Semantic connection types
- Value constraints, multiplicities
- Information are always considered in combination



Rule types

- Restrictiveness
 - law
 - proposal
- Order
 - Direct rules
 - Indirect rules
- Meta rules
 - Necessary for the interplay of rules
 - **Result:** mapping or contradiction

Mapping result





Summary

- Transformation from natural language needs interaction with end users
- Interaction can be at various stages (sentence selection; interlingua)
- An Interlingua can support the detection of defects before a conceptual model is finished