Formal and Relational Concept Analysis approaches in Software Engineering: an overview and an application to learn model transformation patterns in examples

Xavier Dolques, Marianne Huchard, Clémentine Nebut, Hajer Saada

To cite this version:

HAL Id: lirmm-00616272
https://hal-lirmm.ccsd.cnrs.fr/lirmm-00616272v2
Submitted on 2 Dec 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Formal and Relational Concept Analysis approaches in Software Engineering:
an overview and an application to learn model transformation patterns in examples

Xavier Dolques, Marianne Huchard, Clémentine Nebut, Hajer Saada

1 INRIA, Centre Inria Rennes - Bretagne Atlantique
2 LIRMM - Université Montpellier 2, CNRS

First ICESE Virtual Workshop on Software Engineering and Artificial Intelligence, may 2011
Motivations

Software contains plenty of data analysis problems involved in:
- the forward engineering process
- the re-engineering tasks
- various analyses

Focusing on Formal Concept Analysis
- an exploratory data analysis / data mining method
- an unsupervised machine learning approach
- produces clusters, classification and implication rules
Motivations

Highlight main characteristics of FCA
- defining FCA
- main applications of FCA in SE
- multi-relational data analysis with RCA
- young applications of RCA in SE

Learning model transformation (MT) patterns
- on examples of MT
- building of MT examples using ontology alignment
- learning MT patterns with RCA
Formal Concept Analysis (FCA)

What is FCA?
- a formalization of the philosophical notion of concept
- an approach for data analysis and knowledge processing
- many existing experiences and projects
- algorithms, graphical representations, tools
- an active research community (3 conf. ICFCA, CLA, ICCS)

source http://people.aifb.kit.edu/jvo/fca4sw/
What is a concept?
The concept *bird*

- a set of objects (concept’s extent):

- a set of attributes / characteristics (concept’s intent):

  *feathers, with a bill, etc.*
The concept *flamingo* is a subconcept of the concept *bird*

- inclusion of concept’s extents:
  the set of flamingos is included in the set of birds

- inclusion of the concept’s intents:
  the attributes of birds are included in the attributes of flamingos
The formal context

Things that are known about the world

<table>
<thead>
<tr>
<th></th>
<th>flying (fl)</th>
<th>nocturnal (n)</th>
<th>feathered (fe)</th>
<th>migratory (m)</th>
<th>duck-billed (db)</th>
<th>web (w)</th>
</tr>
</thead>
<tbody>
<tr>
<td>flying squirrel (S)</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>bat (B)</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ostrich (O)</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>flamingo (F)</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sea-gull (G)</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Concepts can be derived from a formal context

A formal concept is a pair \((X, Y)\) where

- \(Y\) is the set of attributes common to the objects of \(X\)
- \(X\) is the set of objects having all attributes of \(Y\)
The concept lattice: specialization order

**Introduction**
- FCA
- RCA
- By Example
- Example
- Matching
- Transformation Generation
- Conclusion

more general concepts

more specific concepts

Concepts and labels:
- Concept_1: flying
- Concept_2: nocturnal bat
- Concept_3: flying squirrel
- Concept_4: feathered ostrich
- Concept_5: web
- Concept_6: migratory flamingo
- Concept_7: duck-billed sea-gull
The concept lattice: clusters

- **Flying Animals**
  - Concept 2: nocturnal bat
  - Concept 3

- **Birds**
  - Concept 1: flying
  - Concept 5: feathered ostrich
  - Concept 4

- **Flying birds**
  - Concept 6: migratory flamingo
  - Concept 7: duck-billed sea-gull
  - Concept 8: web flying squirrel
The concept lattice: implication rules

- Concept_0
- Concept_1
  - flying
- Concept_5
  - feathered
  - ostrich
- Concept_2
  - nocturnal
  - bat
- Concept_8
  - web
  - flying squirrel
- Concept_4
- Concept_6
  - migratory
  - flamingo
- Concept_7
  - duck-billed
  - sea-gull
- Concept_3

**flying and feathered**

**migratory**
A Survey of Formal Concept Analysis Support for Software Engineering Activities, Tilley et al., FCA 2005

... And many research work during the past 5 years

- Requirement Analysis: elaborating requirements [Andelfinger], reconciling stake-holders [Düwel et al.], linking use cases and classes [Böttger et al.]
- Component / Web service classification and retrieval [Lindig, Fisher, Aboud et al., Azmeh et al.]
- Exploring a formal specification [Tilley]
- Dynamic analysis: debug temporal specifications [Ammons et al.], test coverage [Ball], locating features [Eisenbarth et al., Bojic et al.], fault localization [Cellier et al.]
Analysis of legacy systems:
- Configuration structure [Snelting]
- Grouping fields in COBOL systems [Van Deursen et al., Kuipers et al.]
- Migrating COBOL towards Corba components [Canfora]
- Migrating from imperative to OO paradigm [Sahraoui et al., Siff et al., Tonella]
- Reengineering class hierarchies [Snelting et al., Schupp et al., Godin et al., Huchard et al.]
- Detecting patterns [Tonella and Antoniol, Arévalo et al.]
- Order for reading classes [Dekel]
- Bad smell correction [Bhatti et al.]
- Conceptual code exploration [Cole et al.]
- Aspect mining [Tonella and Ceccato, Tourwé and Mens]
- Access-guided client class extraction for Eiffel [Ardourel and Huchard]
- Mining Source Code for Structural Regularities [Lozano et al.]
1 Introduction
2 FCA
3 RCA
4 “Model Transformation By example” approaches
5 Illustrative Example
6 Models Matching
7 Transformation patterns/rules generation
8 Conclusion
Relational Concept Analysis (RCA)

- Extend the purpose of FCA for taking into account relations between objects
- The RCA process relies on the following main points:
  - a relational model based on the entity-relationship model
  - a conceptual scaling process allowing to represent relations between objects as relational attributes
  - an iterative process for designing a concept lattice where concept intents include non-relational and relational attributes.
- RCA provides relational structures that can be represented as ontology concepts within a knowledge representation formalism such as description logics (DLs).

Relational Concept Analysis (RCA)

A relational model based on the entity-relationship model ...

Pizza story
Objects and links
Relational Concept Analysis (RCA)

Pizza story

Pizza data

- four object/attribute contexts
  - $K_{People} \subseteq \text{People} \times \text{people names}$
  - $K_{Pizza} \subseteq \text{Pizza} \times \text{pizza names}$
  - $K_{Food} \subseteq \text{Food item} \times \text{food names}$
  - $K_{Country} \subseteq \text{Country} \times \text{country names}$

- four object/object contexts
  - eats $\subseteq \text{People} \times \text{Pizza}$
  - contains $\subseteq \text{Pizzas} \times \text{Food item}$
  - producedIn $\subseteq \text{Food item} \times \text{Country}$
  - hasForNational $\subseteq \text{Country} \times \text{People}$
### Formal contexts

**K\text{People}**

<table>
<thead>
<tr>
<th></th>
<th>Amedeo</th>
<th>Flavia</th>
<th>Dagrun</th>
<th>Lars</th>
<th>Uma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amedeo</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flavia</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dagrun</td>
<td></td>
<td>×</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lars</td>
<td></td>
<td></td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uma</td>
<td></td>
<td></td>
<td></td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

**K\text{Pizza}**

<table>
<thead>
<tr>
<th></th>
<th>Dronning</th>
<th>Kampanje</th>
<th>Margherita</th>
<th>Marina</th>
<th>Norwegian</th>
<th>Regina</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dronning</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kampanje</td>
<td></td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Margherita</td>
<td></td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marina</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norwegian</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td>Regina</td>
<td></td>
<td>×</td>
<td></td>
<td>×</td>
<td>×</td>
<td></td>
</tr>
</tbody>
</table>

**K\text{Ingredients}**

<table>
<thead>
<tr>
<th></th>
<th>basilic</th>
<th>chicken</th>
<th>cream</th>
<th>ham</th>
<th>mozzarella</th>
<th>mushroom</th>
<th>olive</th>
<th>tomato</th>
<th>salmon</th>
<th>swisscheese</th>
</tr>
</thead>
<tbody>
<tr>
<td>basilic</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chicken</td>
<td></td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cream</td>
<td></td>
<td></td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ham</td>
<td></td>
<td></td>
<td></td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mozzarella</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mushroom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>olive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>×</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tomato</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>salmon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>swisscheese</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>×</td>
</tr>
</tbody>
</table>

**K\text{Country}**

<table>
<thead>
<tr>
<th></th>
<th>Italy</th>
<th>Norway</th>
<th>Switzerland</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>×</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Introduction**

- FCA
- RCA
- By Example
- Example
- Matching
- Transformation
- Generation
- Conclusion
Relational context: $R_{eats}$

<table>
<thead>
<tr>
<th></th>
<th>Dronning</th>
<th>Kampanje</th>
<th>Margherita</th>
<th>Marina</th>
<th>Norwegian</th>
<th>Regina</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amedeo</td>
<td></td>
<td></td>
<td>×</td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td>Flavia</td>
<td></td>
<td></td>
<td></td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dagrun</td>
<td></td>
<td></td>
<td>×</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lars</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>Uma</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>×</td>
</tr>
</tbody>
</table>
Relational context: $R_{\text{contains}}$

<table>
<thead>
<tr>
<th></th>
<th>basilic</th>
<th>chicken</th>
<th>cream</th>
<th>ham</th>
<th>mozzarella</th>
<th>mushroom</th>
<th>olive</th>
<th>tomato</th>
<th>salmon</th>
<th>swisscheese</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dronning</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kampanje</td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Margherita</td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Marina</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norwegian</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Regina</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Relational context: $R_{\text{producedIn}}$

<table>
<thead>
<tr>
<th></th>
<th>Italy</th>
<th>Norway</th>
<th>Switzerland</th>
</tr>
</thead>
<tbody>
<tr>
<td>basilic</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chicken</td>
<td></td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>cream</td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td>ham</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mozzarella</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mushroom</td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td>olive</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tomato</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>salmon</td>
<td></td>
<td></td>
<td>×</td>
</tr>
<tr>
<td>swisscheese</td>
<td></td>
<td></td>
<td>×</td>
</tr>
</tbody>
</table>
Relational context: $R_{\text{hasForNational}}$

<table>
<thead>
<tr>
<th></th>
<th>Amedeo</th>
<th>Flavia</th>
<th>Dagrun</th>
<th>Lars</th>
<th>Uma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Italy</td>
<td>$\times$</td>
<td>$\times$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td></td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td></td>
<td></td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
</tbody>
</table>
A RCF $\mathcal{F}$ is a pair $(K, R)$ with:

- $K$ is a set of formal contexts $K_i = (O_i, A_i, I_i)$
- $R$ is a set of relational contexts $R_j = (O_k, O_l, I_j)$,

**Pizza RCF**

$K = K_{People}, K_{Pizza}, K_{Food}, K_{Country}$

$R = R_{eats}, R_{contains}, R_{producedIn}, R_{hasForNational}$
An iterative approach (RCA)

Learned concepts are used in a next step to learn more
RCA - Step 0 - Initial Lattices

Introduction FCA RCA By Example Example Matching Transformation Generation Conclusion
Scaling relations

Integrating concepts in the relational contexts

Amedeo eats Margherita ; Margherita ∈ extent(Concept_11)
→ ∃p ∈ Concept_11, s.t. Amedeo eats p
→ (Amedeo, eats:Concept_11)
→ (Amedeo, Concept_11) belongs to the existentially scaled relation eat*, (Amedeo, ∃ eat :Concept_11) stands

<table>
<thead>
<tr>
<th></th>
<th>Amedeo</th>
<th>Flavia</th>
<th>Dagrun</th>
<th>Lars</th>
<th>Uma</th>
</tr>
</thead>
<tbody>
<tr>
<td>eats</td>
<td>Concept_7</td>
<td></td>
<td>Concept_8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>eats</td>
<td>Concept_11</td>
<td></td>
<td>Concept_12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>eats</td>
<td>Concept_13</td>
<td></td>
<td>Concept_14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amedeo</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flavia</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dagrun</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lars</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Uma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
RCA - Lattices at step 1

Introduction FCA RCA By Example Example Matching Transformation Generation Conclusion
RCA - Lattices at step 2
An excerpt of the iteration
Applications

- **UML class diagram refactoring**

- **UML Use case diagram refactoring**
  * X. Dolques, M. Huchard, C. Nebut, and P. Reitz. Fixing generalization defects in UML use case diagrams. CLA 2010: 247-258

- **Blob design defect correction**

- **Extracting architectures in object-oriented software**
Applications

- **Learning model Transformation patterns in MDE**
  

- **Classification of web services**
  

- **Ontology construction**
  

- **Ontology pattern extraction**
  

- **Ontology restructuring**
  
A synthesis on RCA

- an iterative method to produce abstractions
- variations on scaling operators: $\exists, \forall, \forall\exists, \geq n r : c$, etc. (on relational contexts and on steps)
- object-attribute concept posets can be built instead of lattices to limit the complexity
- opportunities for enhancing application of FCA to software engineering domain

Tools
- Galicia: http://galicia.sourceforge.net/
- eRCA: http://code.google.com/p/erca/
Context and Motivations

Context
- Model driven development
- Development of a model transformation
- Source and target metamodels require domain experts

Motivations
- Ease and speed up the development process of model transformations
- Improve the integration of domain expert in the process
The “By Example” approach

Informal Example
The “By Example” approach
The “By Example” approach

Informal Example

Model 1

Metamodel 1

conforming to

Model 2

Metamodèle 2

modelised by

Introduction

FCA

RCA

By Example

Example

Matching

Transformation Generation

Conclusion
The “By Example” approach

Informal Example

modelised by

Model 1

conforming to

Metamodel 1

transformation

Model 2

conforming to

Métamodèle 2

modelised by

Introduction FCA RCA By Example Example Matching Transformation Generation Conclusion

lirmm-00616272, version 1 - 21 Aug 2011
The “By Example” approach
The “By Example” approach

UML
Person
name
Class
Property
ownedAttribute
Entity
Attribute

conforming to

(1)
(2)

instance(Class, x) ∧ ownedAttribute(x) ≠ ∅ ⇒ instance(Entity, t(x)) ∧ y ∈ ownedAttribute(x) → t(y) ∈ attribute(t(x))

instance(Property, x) ∧ association(x) = ∅ ⇒ instance(Attribute, t(x))
The “By Example” approach

(1) \( \text{instance}(\text{Class}, x) \land \text{ownedAttribute}(x) \neq \emptyset \Rightarrow \text{instance}(\text{Entity}, t(x)) \land (y \in \text{ownedAttribute}(x) \rightarrow t(y) \in \text{attribute}(t(x))) \)

(2) \( \text{instance}(\text{Property}, x) \land \text{association}(x) = \emptyset \Rightarrow \text{instance}(\text{Attribute}, t(x)) \)
“By Example” approaches

[Balogh et Varró(2009)]

- **Input matching**: set of typed couples of elements
- **Matching creation**: manually
- **Input specific development**: none
- **Learning principle**: Inductive Logic Programming
- **Output data**: transformation rules (VIATRA)
“By Example” approaches

[Wimmer et al. (2007)]

- **Input matching**: set of couples of elements
- **Matching creation**: manually
- **Input specific development**: explicit constraints of the transformation from concrete to abstract syntax
- **Learning principle**: \textit{ad hoc} method
- **Output data**: ATL code
“By Example” approaches

[Kessentini et al. (2008)]

- **Input matching**: set of block couples
- **Matching creation**: manually
- **Input specific development**: none
- **Learning principle**: metaheuristics
- **Output data**: a transformed model
Our approach

- **Input matching**: set of couples of elements
- **Matching creation**: matching assisted by tool
- **Input specific development**: none
- **Learning principle**: Relational Concept Analysis
- **Output data**: specification of transformation rules ordered in a lattice
General approach

Examples Models
Matching
Icons: http://www.openclipart.org (Public Domain)

Model 1
conforming to
Metamodel 1

Model 2
conforming to
Metamodel 2

modeled by

Transformation rules
1 Introduction
2 FCA
3 RCA
4 “Model Transformation By example” approaches
5 Illustrative Example
6 Models Matching
7 Transformation patterns/rules generation
8 Conclusion
Transformation Example

UML model

conforming to

Entity-Relationship model

conforming to

UML metamodel

conforming to

Entity-Relationship metamodel
Metamodules involved in the transformation

**UML model**

**Entity-Relationship model**

**UML metamodel**

**Entity-Relationship metamodel**
Metamodels involved in the transformation

Excerpt of UML metamodel
Metamodels involved in the transformation

Entity-Relationship metamodel
Models involved in the transformation

UML model

conforming to

Entity-Relationship model

conforming to

UML metamodel

conforming to

Entity-Relationship metamodel
Models of our Example
A UML Model in concrete syntax
Models of our Example
A UML Model in abstract syntax
Models of our Example
A UML model (excerpt)
Models of our Example
An Entity-Relationship model in concrete syntax

Author
firstName lastName
writes foreword for
writes Text
year
Style
Novel Poetry
title
work
is a is a
has a

(firstName, lastName) (year) (title) (work) (is a, is a) (has a) (0, N)
Models of our Example
An Entity-Relationship model (excerpt)

- **Text**: `name="Text"`
  - **title**: `Attribute`
    - `name="title"`
  - **has a**: `Relationship`
    - `name="has a"`

- **Style**: `name="Style"`
  - **title**: `Attribute`
    - `name="title"`
  - **work**: `Role`
    - `name="work"`
  - **text**: `Role`
    - `name="text"`
  - **textC**: `Cardinality`
    - `min=0`
    - `max=-1`
  - **styleC**: `Cardinality`
    - `min=1`
    - `max=1`
Outline

1. Introduction
2. FCA
3. RCA
4. “Model Transformation By example” approaches
5. Illustrative Example
6. Models Matching
7. Transformation patterns/rules generation
8. Conclusion
Models Matching

Examples

Source MetaModel
Source Model
Target MetaModel
Target Model

Metamodel 1

Metamodel 2

Model 1

Model 2

Matching

Transformation generation

Transformation rules

Icons: http://www.openclipart.org (Public Domain)

conforming to
conformsTo
input/output

Transformation

modeled by

mandarine

bercamote

Introduction FCA RCA By Example Example Matching Transformation Generation Conclusion
the informal starting example contains named elements: those elements are to be found in the two models

source and target models structure are close enough
State of the Art

Constraints of the matching methods that can be applied
- can be applied on a graph structure (not only a tree)
- not strictly based on semantic analysis

Relevant matching approaches
- Similarity Flooding [Melnik et al., 2002]
- OLA [Euzenat et al., 2004]
- Anchor Prompt [Noy et Musen, 2001]

Advantage of anchorPrompt
- does not use similarity on relations names (relations names come from the metamodel level)
Description of process

Matching using Attributes values
- enumeration of values in the models
- matching of similar values
- matching of the elements containing those values

Matching Propagation
- using structure similarity assumption
Attribute Instances
In the UML source model
Attribute Instances
In the UML source model (excerpt)
Attribute Instances
In the Entity-Relationship target model (excerpt)
Value comparison

A source attribute instance and a target attribute instance match if all the following conditions are respected:

- they have the same value
- this value appears only once in the source model and the target model
Value matchings in attributes
A first Model Matching

Introduction

Matching
Propagation of the matching
The AnchorPrompt approach

$W(x, y) = 1 - \left| \frac{\text{index}(x)}{\text{length}(X) - 1} - \frac{\text{index}(y)}{\text{length}(Y) - 1} \right|$
Similarities propagation

Poetry: Entity
  name="Poetry"
  poetry: Role
    name="poetry"
  isA2: Relationship
    name="is a"
  text2: Role
    name="text"
  Text: Entity
    name="Text"
    text: Role
      name="text"
    has a: Relationship
      name="has a"

Poetry: Class
  name="Poetry"
  g2: Generalization
  Text: Class
    name="Text"
    text: Property
      name="text"
      lowerBound=0
      upperBound=-1
    has a: Association
      name="has a"
Similarities propagation

Text

Poetry: Entity

poetry: Role

isA2: Relationship

text2: Role

Text: Entity

has a: Relationship

Poetry: Class

g2: Generalization

Text: Class

property: Role

text: Property

has a: Association

Introduction FCA RCA By Example Example Matching Transformation Generation Conclusion
Empirical results

Precision calculation

![Precision Calculation Graph](chart.png)

**Substrings**
- precision substring
- precision anchorPrompt

**Keywords**
- delegation1
- Introducing
- Interface
- Introduce
- Delegate
- Redundant
- Inheritance
- Equivalence
- Attributes
- Associations
- Book2Publication
- Families2Persons
- uml.er1
- uml.er2
- UML2ClassDiagram2Table
- associations.persons
- KM32Problem
- emf2km3
- JavaSource2Table
- KM32EMF
To address the model matching problem

- Starting with 2 assumptions
  - the starting example contains named element
  - the models to match are structurally close

- We propose a process generating a model matching
  - using attributes values for a first matching with high confidence level
  - using an adaptation of AnchorPROMPT for extending the first matching
Transformation rules generation

Icons: http://www.openclipart.org (Public Domain)
Input Data

UML model

conforming to

Entity-Relationship model

conforming to

UML metamodel

conforming to

Entity-Relationship metamodel
Approach

- Consider the properties of the model elements:
  - their class
  - their relations with their neighbors
  - the properties of their neighbors
- Classify the different properties from the examples
- Classify the matching links considering the classification of the properties of their extremities

Formal Concept Analysis allows the classification of a set of objects considering their attributes
### Transformation of models in tables

- **Metamodel elements contexts**
  - Source metamodel context
  - Target metamodel context
- **Model elements contexts**
  - Source model context
  - Target model context
- **Matching links context**
- **Relations**
  - between model source elements and their class from the source metamodel
  - between model target elements and their class from the target metamodel
  - between model source elements: *e.g.* `ownedAttribute`
  - between model target elements: *e.g.* `attribute`
  - between matching links and their source from the source model
  - between matching links and their target from the target model

The management of relations requires RCA to iterate
Lattice of model elements

Source model (UML)

Concept 49

Introduction
FCA
RCA
By Example
Example
Matching
Transformation Generation
Conclusion
Lattice of matching links
Lattice interpretation

rules lattice (excerpt)
Rules examples
Hypothesis
- closed world
- needs a good cover of the transformation

Different kinds of characteristics
- needed characteristics
- allowed characteristics
- forbidden characteristics
Needed characteristics
Forbidden characteristics

rules_11 - Concept_81

Property

Class

Attribute

Entity

entity [0..*]

Property Classtype [0..0]

Attribute Entity

entity [0..*]

41 (S: 1)

A : Class
tA_10

Concept_42 (S: 1)

metaClassA : Property
elementA_2
elementA_3
elementA_5
elementA_14

Concept_49 (S: 1)
type : Concept_41
elementA_16

90

Introduction FCA RCA By Example Example Matching Transformation Generation Conclusion
## Empirical study

### Table: Empirical metrics

<table>
<thead>
<tr>
<th>name</th>
<th>size of MapLinks lattice</th>
<th>number of extracted rules</th>
<th>premise size min</th>
<th>premise size max</th>
<th>premise size avg.</th>
<th>premise size std. dev.</th>
<th>premise depth min</th>
<th>premise depth max</th>
<th>premise depth avg.</th>
<th>premise depth std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>associations-persons</td>
<td>9</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>1.53</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>0.82</td>
</tr>
<tr>
<td>Book2Publication</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>1.21</td>
</tr>
<tr>
<td>delegation1</td>
<td>47</td>
<td>12</td>
<td>15</td>
<td>33</td>
<td>18</td>
<td>5.89</td>
<td>4</td>
<td>8</td>
<td>5</td>
<td>1.22</td>
</tr>
<tr>
<td>delegation2</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>17</td>
<td>9</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>0.87</td>
</tr>
<tr>
<td>Disaggregation</td>
<td>22</td>
<td>10</td>
<td>6</td>
<td>10</td>
<td>6</td>
<td>1.26</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>0.8</td>
</tr>
<tr>
<td>Ecore2Class</td>
<td>14</td>
<td>8</td>
<td>7</td>
<td>15</td>
<td>8</td>
<td>3.16</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>0.95</td>
</tr>
<tr>
<td>EliminateRedundantInheritance</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>11</td>
<td>8</td>
<td>1.63</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>1.27</td>
</tr>
<tr>
<td>emf2km3</td>
<td>26</td>
<td>19</td>
<td>14</td>
<td>52</td>
<td>20</td>
<td>9.48</td>
<td>5</td>
<td>7</td>
<td>4</td>
<td>0.77</td>
</tr>
<tr>
<td>EquivalenceAttributesAssociations</td>
<td>47</td>
<td>18</td>
<td>11</td>
<td>36</td>
<td>15</td>
<td>7.96</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>0.77</td>
</tr>
<tr>
<td>extractClass</td>
<td>10</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>1.15</td>
</tr>
<tr>
<td>Families2Persons_final</td>
<td>11</td>
<td>8</td>
<td>1</td>
<td>7</td>
<td>5</td>
<td>1.8</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>1.22</td>
</tr>
<tr>
<td>hideDelegate</td>
<td>53</td>
<td>22</td>
<td>17</td>
<td>117</td>
<td>35</td>
<td>28.7</td>
<td>4</td>
<td>7</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>IntroducePrimaryKey</td>
<td>31</td>
<td>17</td>
<td>11</td>
<td>18</td>
<td>12</td>
<td>2.5</td>
<td>4</td>
<td>7</td>
<td>2</td>
<td>0.91</td>
</tr>
<tr>
<td>IntroducingInterface</td>
<td>38</td>
<td>11</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>3.02</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>1.61</td>
</tr>
<tr>
<td>JavaSource2Table</td>
<td>17</td>
<td>10</td>
<td>5</td>
<td>9</td>
<td>7</td>
<td>1.3</td>
<td>3</td>
<td>7</td>
<td>2</td>
<td>0.87</td>
</tr>
<tr>
<td>KM32EMF</td>
<td>18</td>
<td>15</td>
<td>11</td>
<td>20</td>
<td>14</td>
<td>3.74</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td>1.35</td>
</tr>
<tr>
<td>KM32Problem</td>
<td>37</td>
<td>28</td>
<td>28</td>
<td>378</td>
<td>52</td>
<td>65.36</td>
<td>7</td>
<td>9</td>
<td>8</td>
<td>0.65</td>
</tr>
<tr>
<td>uml-er</td>
<td>35</td>
<td>24</td>
<td>2</td>
<td>11</td>
<td>6</td>
<td>2.75</td>
<td>2</td>
<td>9</td>
<td>5</td>
<td>1.15</td>
</tr>
<tr>
<td>uml-er-iccs</td>
<td>21</td>
<td>12</td>
<td>3</td>
<td>9</td>
<td>6</td>
<td>1.78</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>1.54</td>
</tr>
<tr>
<td>uml-er2</td>
<td>25</td>
<td>15</td>
<td>2</td>
<td>9</td>
<td>4</td>
<td>2.19</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>1.41</td>
</tr>
<tr>
<td>uml-er3</td>
<td>40</td>
<td>19</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>0.83</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>0.97</td>
</tr>
<tr>
<td>UML2ClassDiagram_to_KM3</td>
<td>200</td>
<td>40</td>
<td>28</td>
<td>187</td>
<td>58</td>
<td>40.51</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>2.72</td>
</tr>
</tbody>
</table>
## Empirical study

### Table of Empirical Study

<table>
<thead>
<tr>
<th>nom</th>
<th>min</th>
<th>max</th>
<th>avg.</th>
<th>std. dev.</th>
<th>min</th>
<th>max</th>
<th>avg.</th>
<th>std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>associations-persons</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>0.91</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>0.58</td>
</tr>
<tr>
<td>Book2Publication</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>delegation1</td>
<td>7</td>
<td>17</td>
<td>7</td>
<td>2.89</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>1.04</td>
</tr>
<tr>
<td>delegation2</td>
<td>15</td>
<td>26</td>
<td>20</td>
<td>4.96</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>0.63</td>
</tr>
<tr>
<td>Disaggregation</td>
<td>7</td>
<td>11</td>
<td>7</td>
<td>1.26</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>0.89</td>
</tr>
<tr>
<td>Ecore2Class</td>
<td>1</td>
<td>8</td>
<td>4</td>
<td>2.26</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>EliminateRedundantInheritance</td>
<td>7</td>
<td>11</td>
<td>8</td>
<td>1.63</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>0.82</td>
</tr>
<tr>
<td>emf2km3</td>
<td>11</td>
<td>20</td>
<td>14</td>
<td>3.59</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>1.17</td>
</tr>
<tr>
<td>EquivalenceAttributesAssociations</td>
<td>9</td>
<td>38</td>
<td>14</td>
<td>8.37</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>1.2</td>
</tr>
<tr>
<td>extractClass</td>
<td>16</td>
<td>20</td>
<td>18</td>
<td>2</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>0.87</td>
</tr>
<tr>
<td>Families2Persons_final</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>0.79</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>hideDelegate</td>
<td>17</td>
<td>186</td>
<td>39</td>
<td>37.75</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>0.88</td>
</tr>
<tr>
<td>IntroducePrimaryKey</td>
<td>9</td>
<td>14</td>
<td>9</td>
<td>1.21</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>1.03</td>
</tr>
<tr>
<td>IntroducingInterface</td>
<td>2</td>
<td>20</td>
<td>9</td>
<td>4.53</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>1.48</td>
</tr>
<tr>
<td>JavaSource2Table</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0.63</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0.63</td>
</tr>
<tr>
<td>KM32EMF</td>
<td>14</td>
<td>52</td>
<td>21</td>
<td>10.23</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>1.1</td>
</tr>
<tr>
<td>KM32Problem</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>uml-er</td>
<td>2</td>
<td>15</td>
<td>9</td>
<td>3.61</td>
<td>2</td>
<td>8</td>
<td>5</td>
<td>1.88</td>
</tr>
<tr>
<td>uml-er-iccs</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>0.29</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>0.76</td>
</tr>
<tr>
<td>uml-er2</td>
<td>2</td>
<td>15</td>
<td>9</td>
<td>3.71</td>
<td>2</td>
<td>8</td>
<td>5</td>
<td>1.88</td>
</tr>
<tr>
<td>uml-er3</td>
<td>2</td>
<td>15</td>
<td>9</td>
<td>3.47</td>
<td>2</td>
<td>8</td>
<td>5</td>
<td>1.84</td>
</tr>
<tr>
<td>UML2ClassDiagram_to_KM3</td>
<td>18</td>
<td>45</td>
<td>22</td>
<td>6.06</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>0.74</td>
</tr>
</tbody>
</table>

---

**Introduction**

- FCA
- RCA
- By Example
- Example
- Matching

**Transformation Generation**

**Conclusion**
The generation of transformation patterns is the result of the following steps:

- the transformation of the examples into relational contexts
- the application of the process RCA
- the interpretation of the final lattices
Outline

1. Introduction
2. FCA
3. RCA
4. “Model Transformation By example” approaches
5. Illustrative Example
6. Models Matching
7. Transformation patterns/rules generation
8. Conclusion
Contribution

An approach with tool for assisting the development of model transformations

Icons: http://www.openclipart.org (Public Domain)
Contribution

- Implementation independent with transformation metamodels
- Model matching assisted
- Result usable thanks to lattices
- Result genericity
Perspectives

Model matching

- Improving precision on the AnchorPROMPT adaptation
- Adapt other matching approaches for model matching

Transformation rules generation

- Building relations in final model for not simple case
- Improving interaction with user
- Insisting on approach validation

Evolutions

- Adapt the approach to other problems
- Investigating complementarity with other approaches
Thank you for your attention
Metamodel matching

Lopes et al. (2005)
Generating transformation definition from mapping specification: Application to web service platform.
*In CAiSE’05, LNCS 3520, pages 309–325, 2005.*

Del Fabro et Valduriez (2007)
Semi-automatic model integration using matching transformation and weaving models.

Falleri et al. (2008)
Meta-model Matching for Automatic Model Transformation Generation.
Balogh et Varró (2009)
Model transformation by example using inductive logic programming.


Kessentini et al. (2008)
Model Transformation as an Optimization Problem.

Wimmer et al. (2007)
Towards model transformation generation by-example.