BACKGROUND KNOWLEDGE IN DATA MINING

Tomáš Klieger, Vojtěch Svátek, Milan Šimůnek, Daniel Šťastný and Andrej Hazucha
BACKGROUND KNOWLEDGE

- Background Knowledge Exchange Format (BKEF) XML Schema
- Semantic abstraction over BKEF – the Background Knowledge Ontology (BKOn)
DESIGN OBJECTIVES

Disadvantages of semantic format as primary standard

• verbosity

• poor readability due to structural complexity

• the need for specialized, not widely available APIs

Proposed solution in form of XML Schema (BKEF)
DESIGN OBJECTIVES

Integration with Other Specifications

- PMML
- FML
BASIC CONCEPTS

• Hierarchically structured metaattributes
• Meta-field - combination of metaattribute and its format
• Meta-field Value
• Patterns - capturing relations between metaattributes
BKEF

stores mining models

### BKEF Schema Overview

- Meta-Attributes [1..*]
  - Meta-Attribute [1..*]
    - Annotation [0..*]
    - Format [0..*]
- Patterns [0..1]
  - Association Rules
  - Mutual Influences [0..1]
  - Mutual Influence [0..*]
  - Background A. Rules [0..1]
  - Background A. Rule [0..*]
- Has Child [0..*]
- Transformed to [0..*]

### Meta-Attributes [example]

- Blood Pressure (Group Meta-A.)
  - Diastolic blood pressure
    - Formats: Format: mmHg
    - Format: kPa
  - Systolic blood pressure
    - Variability: stable
    - Formats: ...

### Format [example]

- mmHg
  - Author: MUDr. Plesny
  - Data Type: Float
  - Allowed Range: <50;300>
  - Collation: numerical/ascending
  - Preprocessing Hints
    - Discretization Hint: patient without Diabetes
      - Interval Enumeration
      - Interval Bin Name: <50;90>
        - normal
      - Interval Bin Name: <90;140>
        - increased
    - Child Meta Attribute: Smoking
    - Child Meta Attribute: Weight
• semantic abstraction over BKEF

• adds typed associations between attributes
  • Semantic KB utilizes these relations for reasoning

• automatic transformation of BKEF XML into instances of the ontology concepts
BKEF, BKOn and Data Mining

- Data mining process initialization (Collecting knowledge from experts)
  - web based editor to build BKEF XML documents
- Linking Background Knowledge with Mined Data
BKEF, BKOn and Data Mining

- Background Knowledge for Localizing Search
- Background Knowledge for Result Pruning
- Background Knowledge for Postprocessing
CONCLUSIONS

• Standardized format for BK interchange in DM
• Ontology for BK modeling
• Future development: Ontology design patterns