On the Relations between Structural Case-Based Reasoning and Ontology-based Knowledge Management

Ralph Bergmann & Martin Schaaf

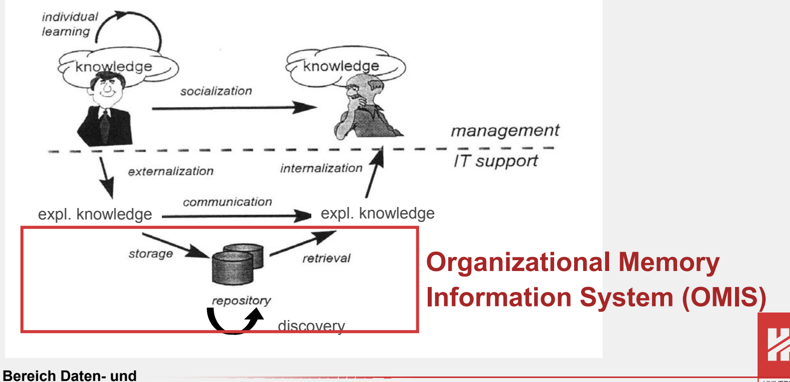
University of Hildesheim
Data- and Knowledge Management Group

www.dwm.uni-hildesheim.de

Information Technology for KM

- Technology to support KM activities:

  Capture, Store, Select (Retrieval), Share, Apply, Discover



OMIS Principles

- Acquiring and maintaining OMIS entries is an investment:
  - Only administer knowledge which is productive
  - Use existing sources of knowledge like manuals, faq's, etc., For building up and maintaining the knowledge base
- It is not necessary to model the domain up to 100%
- Supporting people in routine tasks through:
  - access to more information
  - higher degree of decisions quality
- "Soft Technology": support and not replacement

S-CBR and OBKM

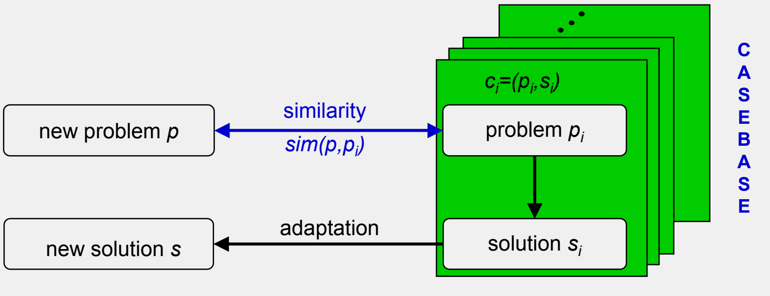
- Different approaches are applied for implementing OMIS:
  - Structural CBR (S-CBR)
    - Methodological: (Extended) CBR Cycle viewed also as organizational structure
    - Technology: Using similarity-based reasoning, e.g. tool orenge
  - Ontology-based Knowledge Management (OBKM)
    - Methodological: Organizing knowledge according to an ontology
    - Technology: Using deductive inference, e.g. tool Ontobroker
- Thesis:

  S-CBR and OBKM
  - have a different origin but base on similar principles
  - technologies have complementary strengths
- Conclusion:
  - cross-fertilization between both approaches desirable

Overview

1. Introduction
2. Structural CBR applied to Knowledge Management
3. Ontology-based Knowledge Management
4. On the Relations
5. Conclusion

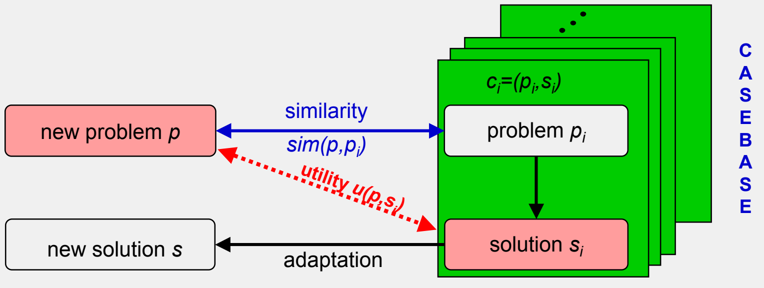
2. Structural CBR applied to Knowledge Management



- Basic CBR assumption:

Similar problems have similar solutions!

Similarity and Utility



- Basic CBR assumption re-phrased:

Similarity between problems
approximates utility of the solution!

Different CBR Approaches

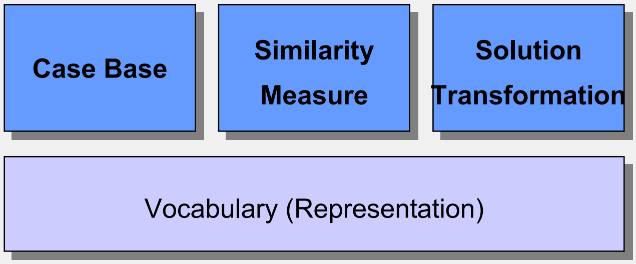
differentiated by the Case Representation Approach

- Textual CBR:
  - Based on free-text cases
  - Particular words or phrases are extracted as indexes
  - No common representation across the cases
- Conversational CBR
  - Cases are represented by a list of questions
  - No common representation across the cases
- Structural CBR
  - Common vocabulary is developed, e.g. attribute-value based or object-oriented used as a common data structure
  - All cases (and all other knowledge) are represented based on this vocabulary

Knowledge Containers for CBR

(Richter, 1995)

- Where Knowledge is contained -



- Knowledge can be distributed between the four containers according to application needs.
- Focus on knowledge in the case base.

Using SCBR for KM

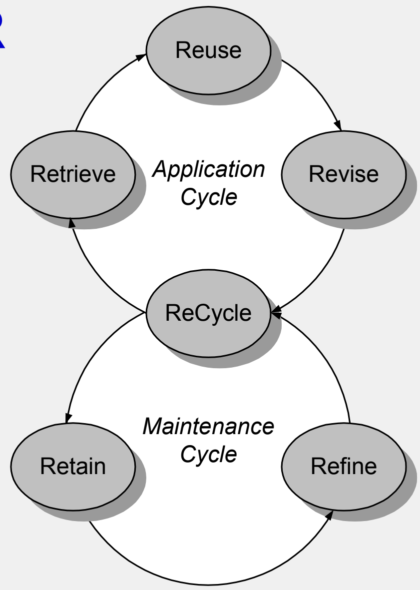
- Store knowledge items as cases ? No not really!

BUT:

- Keep knowledge items in their original knowledge form.
- *Characterize* each knowledge items by means of a particular vocabulary.
- Store the characterization of the knowledge item together with a link to the knowledge item as case in the case base:
  - Problem = Characterization of knowledge item.
  - Solution = Link to knowledge item in knowledge container.
- Develop task-specific similarity measure to assess the utility of a knowledge item only based on its characterization.
- Adaptation is usually not an issue.

Extended CBR Cycle

- Maintenance is becoming an issue



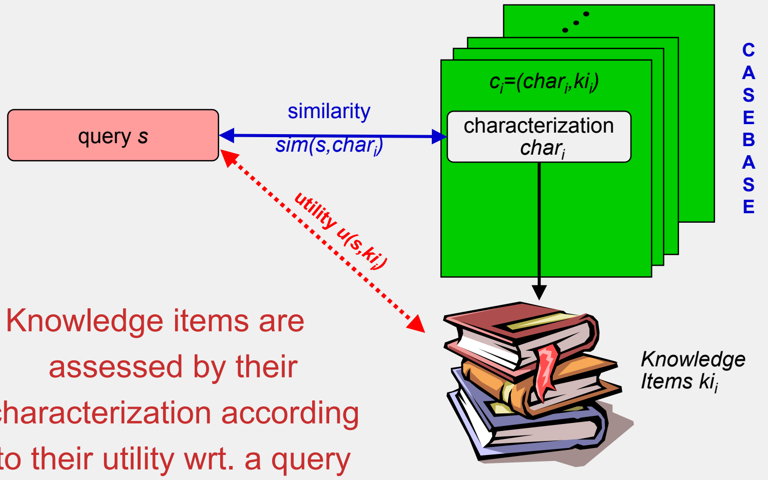
(Göker & Roth-Berghofer, 1999)

Example: HOMER System

(Göker & Roth-Berghofer, 1996-1999)

- HOMER = OMIS for Knowledge about problems with CAD/CAM Hard- and Software at DaimlerChrysler.
- Implemented using CBR Technology as part of the European project INRECA-II.
- KM Challenge:
  - High cost for supporting for CAD-Workstations at the car development.
  - In Sindelfingen in 2000: 1300 CAD-Clients, 16 Applications.
  - Frequent hard- and software updates.
- Main development steps:
  - Large vocabulary developed in cooperation with the hotline operators.
    - more than 200 concepts ...
  - Development of appropriate similarity measures on top of the vocabulary
  - Development of the knowledge base:
    - collected failure reports characterized manually.

KM with SCBR



Knowledge items are assessed by their characterization according to their utility wrt. a query

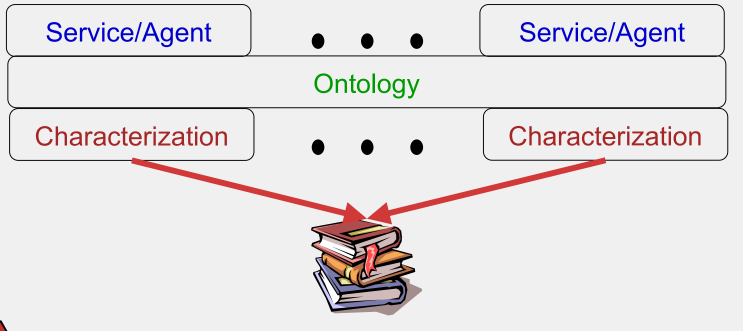
Vocabulary Development Issues

- Vocabulary used ...
  - for characterizing knowledge and formulating queries.
  - by different people in the organization.
  - for a particular purpose, e.g. diagnostic problem solving.
- Hence, the vocabulary must fulfill the following properties:
  - *formal*: must allow similarity-based retrieval.
  - *shared*: common understanding of its usage among the people using it.
  - *utility-distinguishability*: complete as to enable to distinguish the utility of knowledge items for a particular situation.
- Additional helpful properties:
  - functional independence aimed at, if possible.
    - simplifies similarity/utility definition.

This closely matches the aim of ontologies

3. Principles of OBKM

- OMIS: Creation, accumulation, sharing, reuse, and further development of knowledge in an organization within the context of explicitly defined conceptual models (ontologies)
- Meta-Data characterizations associated to knowledge items



What is an Ontology

“An ontology is a formal, explicit specification of a shared conceptualization.”
Tom Gruber.

- An ontology is:
  - *formal*: is machine understandable.
  - *explicit specification*: contains explicitly defined concepts, properties, relations, functions, constraints, axioms, ... about the domain of discourse.
  - *shared*: represents consensual knowledge of a community.
  - *conceptualization*: abstract model of some phenomenon in the world.
- can range from a simple taxonomic hierarchy of classes to a logic program utilizing first-order predicate logic, modal logic, or even higher order logics with probabilities.

Ontological Engineering (1)

Very early origins:

- CYC project: addressed the issue of reusability and modularity of large knowledge bases
- KL-ONE: first logical formalization of a frame-based semantic network
  - inspired an entire new discipline in logical frame-based languages called terminological logics or description logics
  - T-Box: Axioms of the knowledge base
  - A-Box: Instance level knowledge (facts)

- Today:
  - Systems like FLORID/FloXML (U. of Freiburg) or OntoBroker (Ontoprise) based on F-Logic, a descendant of KL-ONE

F-Logic Example

publication[
  title => *string*;
  authors =>> *author*;
  location => *URI*;
  submittedTo => *event*;
  topics =>> *topic*;
].

workshop : *event*.
event[
  title => *string*;
  cfp => *URI*;
  appDay@(*string*) => *date*;
  topics =>> *topic*;
].

$$\forall_{X,T,E}\ X : publication[topics->>T] \leftarrow X[submittedTo->E] \wedge E[topics->>T] \quad (*)$$

conceptualizes

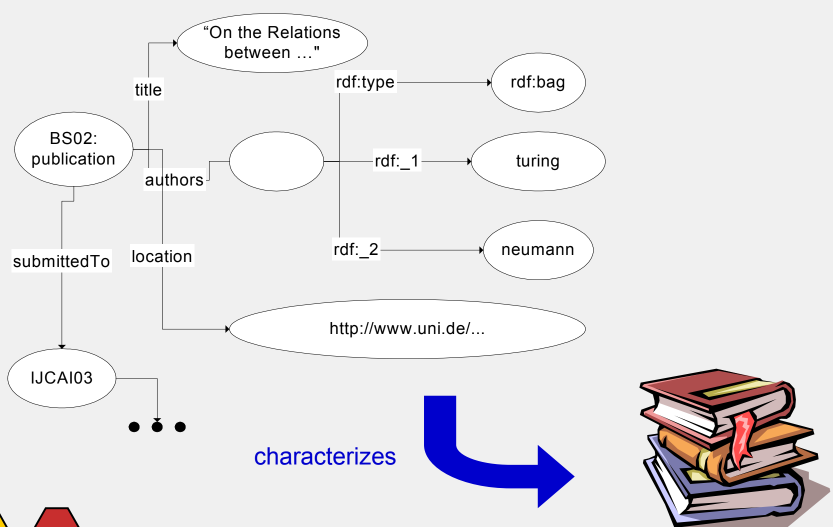
Formal Characterization (Meta-Data Annotation) of Knowledge Items (Instance-level Knowledge, A-Box)

Ontology/Characterization Representation Languages

- Knowledge exchange and reuse require standardized representation languages with well-founded semantics
- Current advancements: Semantic Web

Ontology Representation Languages	Languages for structured characterizations
Resource Description Framework Schema (RDFS)	Resource Description Framework(RDF)
XML Ontology Exchange Language (XOL)	
Ontology Inference Layer (OIL)	
DARPA Agent Markup Language (DAML)	
Ontology Web Language (OWL, based on DAML/OIL)	
XML Topic Maps (XMT)	

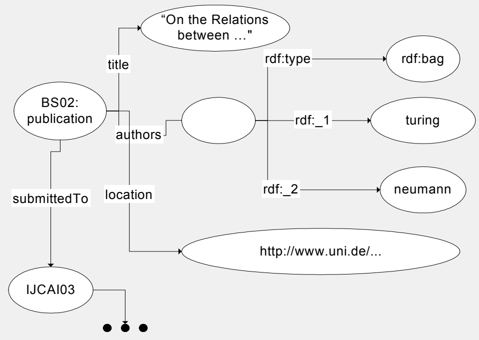
Characterization Example (RDF)



Inference Example

$$\forall x \leftarrow X : publication[authors->> \{turing\}]$$

should generate a binding
for x



But only with a proper set/subset semantic

Ontological Engineering (2)

- Systematic development of ontologies in a reusable and modular fashion and their maintenance
  - Standardized Ontology and Characterization Representation Languages
  - Widely accepted well-founded semantic
  - Ontologies are standard conceptualizations of the domain of discourse
    - committed by all agents/services that do reasoning

4. On the Relations ... (1)

- Vocabularies in SCBR and ontologies are pretty much the same.
  - Both are formal models for restricting possible interpretations of knowledge items.
  - Enable semantic-based access to knowledge items.
  - Rely on similar representational paradigms.
    - SCBR: object-oriented representations.
    - OBKM: frame-logic (frame-based semantic networks) derivates.

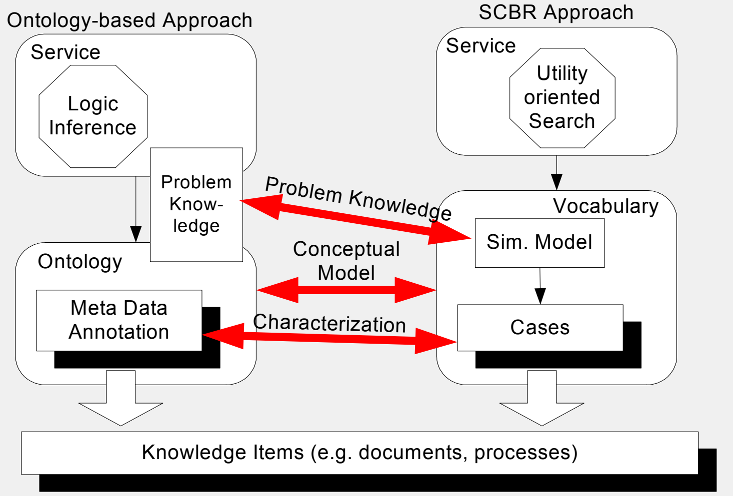
4. On the Relations ... (2)

- Both rely on characterizations of knowledge items requiring some formalization in advance.
  - SCBR: Problem part of cases.
  - OBKM: metadata annotation.
  .
- Ontologies/Vocabularies, characterizations, and knowledge items can by physically dispersed.
  - Different databases or content management systems.
  - Semantic Web.

But there are also differences ...

SCBR	OBKM
mostly isolated; not developed with respect to cooperation with other systems	open; cooperation among agents within an Ontology-based OMIS very important
standardization of representation languages not a big issue	W3C standardizations for the semantic web
systems mostly rely on proprietary representations, although XML-based, no standardized semantic	ontologies claim to provide a standardized conceptualization of the domain of discourse
Vocabularies don‘t conceptualize the domain of discourse per se, but on a task-specific manner	ontologies should be „problem free“ (nearly impossible)
Utility-based inference Suitable for many real world appl. Suggestions if no exactly matching solutions can be found	**Logic-based inference** Correct and provable results Required by computer agents for further processing

Ontology vs. SCBR Knowledge Containers



5. Conclusions

- SCBR for KM could benefit from semantic web developments
  - use of standardized ontology and characterization languages
  - interoperability with other agents
- Ontology standardization could benefit form considering results on similarity modeling
- Integration of utility and inference-based reasoning is desirable:
  - Toolbox Integration: different reasoning on same representation
  - Seamless integration: combined reasoning principle