Intelligent Decision Support Systems

(Case Study 2 – THE TEXTILE INDUSTRY CASE: THE MODSIMTex PROJECT)

Miquel Sànchez-Marrè, Beatriz Sevilla

Intelligent Data Science and Artificial Intelligence Research Research Centre (IDEAI-UPC)
Knowledge Engineering and Machine Learning Group (KEMLG-UPC)

Computer Science Dept.
Universitat Politècnica de Catalunya · BarcelonaTech

miquel@cs.upc.edu, bea.sevilla@gmail.com
http://www.cs.upc.edu/~miquel

Course 2020/2021

https://kemlg.upc.edu
CASE 4 - THE TEXTILE INDUSTRY CASE: THE MODSIMTex PROJECT
MODSIMTex Project

- The **objective** is to develop a system which will dramatically *reduce the cost* to develop *new technical textile products* by reducing the time, energy and raw material waste during the production machinery setup process.

Problem: Trial and error to set up the machinery ⇒ waste of time, effort, material and human resources
MODSIMTex Model

Raw materials
Variables & Values

Machinery settings
Variables & Values

End product
Variables & Values

Estimated values

Attribute without value

System
MODSIMTex Requirements

- For each particular problem $\Rightarrow$ a particular solution

- At each problem description (process) user inputs
  - Several unknown attributes
  - Different at each time.

- Incremental and adaptive
  - New experience along the time

- Non Intrusive system
  - Non user intervention
MODSIMTex A.I. System Requirements

- Use old configurations of textile machine settings relevant to the actual production batches to be produced
  - Raw Materials
    - Kind of yarn
    - Yarn features
  - End-product quality
    - End-product features: porosity, twist
  - Machine settings
    - Setting of main parameters (rotor speed, ...)
- Highly parameterized ⇒ Easily Configurable
- Possibility to be called several times, any time with different configuration ⇒ different results
- Independent on the different sector in MODSIMTex (spinning, knitting, weaving, non woven)
- Incremental & Adaptive ⇒ database grows over the time with the new experience
- Able to predict all the attributes in the same query.
CBR Characteristics

- Open enough to be parameterized \(\Rightarrow\) each phase can be customized
- Able to predict more than one attribute
- Lazy Learning
  - Not creates a model
    - Allow to change configuration each time that is called
    - Allow to be incremental introducing new experience in the case base
MODSIMTex Internal Workflow

External modules

- CBR module
- DB
- Eq solver
- Math module
- FEM module

Calculation manager

User data input

Data validation

Data output

User interface
How is the MODSIMTex Data

- Small number of examples’ sets
- Process = \langle attribute, value \rangle
- Large number of attributes (high dimensionality)
- Unknown dependencies between attributes
- Heterogeneous Data:
  - Continuous attributes
  - Discrete attributes
    - Non Ordered
    - Ordered
CASE-BASED REASONING

Using experience to solve problems

https://kemlg.upc.edu
Case-Based Reasoning

Case: Problem Description (characterization) + Solution
Query: Query at hand defines new case (problem description without solution)
Retrieve: New case is used to find the most similar case/s among the known (previous) cases (case base)
Reuse (Adapt): New and retrieved case/s are combined to a proposed case
Revise (Evaluate): Suggested case is applied and evaluated
Retain (Learn): Useful experiences from applying the case are retained by adapting the case base and the conceptual knowledge

[Mantaras’06; Aadmodt’94; kolodner’93; Riesbeck’89; Schank’77]
Case Structure

- Case
  - Description: characterization of the problem.
  - Solution
  - Optional to improve the CBR performance:
    - Evaluation: evaluates how good is the solution for this problem
    - Utility: measures how useful has been this case in the CBR system
Case Base Indexation

- Structure of how the cases are organized in the case base
- Indexation (Organization of the memory):
  - Flat (list):
    - (+) best matching
    - (-) Expensive retrieval time
  - Hierarchical:
    - (+) Efficient retrieval time
    - (-) Expensive the maintenance
    - (-) Not always optimal matching case
  - Self-Organizing Maps:
    - (+) Previously unknown relationships
    - (+) Cheap maintenance
    - (-) Not always optimal matching case
CBR - Retrieval

- Given a new case and a case base, search the most similar case/s with respect to this new case.
- Depends on:
  - Case Base Indexation
  - Similarity measure $\Rightarrow$ Distance metric
    - Depends on the case structure
    - Attribute weights
CBR - Retrieval

- Given a new case, the most similar ones are selected.
- The relevance of the parameters influence the similarity results.
- The solutions of the selected cases are adapted to the new case

\[
\bar{a} = a_s, c, \bar{w} = \{w_{shape}, w_{color}\}, w_{shape} > w_{color}
\]
CBR - Retrieval

- How many cases to retrieve?
  - Fix Number (m)
  - Distance threshold

- Which distance metric to use?
  - Depends on the case structure
    - Numerical attributes ⇒ Minkowsky’s metrics, Canberra ...
    - Heterogeneous data ⇒ L’Eixample, Gower, ...
    - Graph, images, ...

- What are the Attribute relevance (weights)
  - Feature weighting
  - Expert knowledge
CBR – Adaptation (Reuse)

"Similar problems have similar solutions"[Lieber,07]

- Complex task ⇒ depends on the domain
- Methods:
  - Null ⇒ Copy the solution of the most similar case
  - Transformational ⇒ Transform the retrieved solutions to the new solution (Domain dependent)
    - Formula: formula pre-defined
  - Derivational ⇒ Reuse method/s or algorithm/S used to solve the retrieved solution/s (Domain dependent)
- Generic methods:
  - Null: copy the solution of best matching case
  - Mean/Mode: mean/mode of the selected retrieved cases
  - Weighted Mean/Mode: mean/mode weighted by the distance or an attribute, such as evaluation
CBR - Evaluation

- Evaluation of the proposed solution
- Opportunity to learn from the success or failure.
- None generic method in the literature
  - Experts’ feedback
CBR – Learning (Retain)

- Increase the problem solving capabilities by adding new experience (solved new case)
- Helps to maintain the case base
  - Strategies to clean the case base by deleting also redundant cases
- Depends of if there are only positive cases or also negative cases
CASE-BASED REASONING in MODSIMTex PROJECT

Using experience to make successful decisions

https://kemlg.upc.edu
MODSIMTex CBR

- Process = Input Material parameters + Machinery parameters + Output Material Parameters

- What is the query (new case)?
  - Given a process with some parameters which are specified (description) and others not. Goal: proposing all the values for this process parameters (solution).

- What is a case?
  - Process
    - Input material
    - Machinery
    - Output material

Description??
Solution??

Not a fixed description neither solution. At each new query can vary.
Evolution of MODSIMTex CBR

- Relevance of attributes depends on the specific problem to be solved \(\Rightarrow\) Weights can be changed in every new query
- Some attributes must be satisfied \(\Rightarrow\) Mark those as \textit{locked}
- Not easy to decide an specific value for the end users \(\Rightarrow\) value can be a \textit{range} for numerical attributes or \textit{list} for discrete ones.
- Some attributes with near values are considered as equal \(\Rightarrow\) A margin \textit{tolerance} is defined for numerical attributes:
  - Speed with tolerance 5\% \(\Rightarrow\) 300 = 310
- Relaxing “equality” of two cases \(\Rightarrow\) \textit{Range} & \textit{List}, \textit{Tolerance}, \textit{Maximum Equality Distance}

© Miquel Sànchez-Marrè & Beatriz Sevilla, KEMLG, 2013
MODSIMTex Query

● Description of each attribute:
  
  ▪ Numerical
    ◆ Value: Single value | Range
    ◆ Tolerance (%)
  
  ▪ Discrete
    ◆ Value: Single Value | Value List
  
  ▪ Locked: Restriction over the attribute. This is a “key” attribute that should be satisfied ⇒ distance = 0
  
  ▪ Weight: attribute relevance ⇒ influence in the distance assessment
**MODSIMTex Model**

<table>
<thead>
<tr>
<th>Raw materials</th>
<th>Machinery settings</th>
<th>End product</th>
</tr>
</thead>
<tbody>
<tr>
<td>MV1, MV2, MV3, MV4, MV5</td>
<td>EV1, EV2, EV3, EV4, EV5</td>
<td></td>
</tr>
</tbody>
</table>

- **Attribute without value**
  - Type = numerical
  - Tolerance = 2
  - Weight = 3
  - Locked = True | False
  - Value = value | range

- **Attribute with Value**
  - MV3 = [MV3 - 2%, MV3 + 2%]

- **Retrieved cases must have a value in MP3 with distance=0 to MV3**

- **Estimated values**

© Miquel Sànchez-Marrè & Beatriz Sevilla, KEMLG, 2013
MODSIMTex CBR - Retrieval

1. **If** cases with distance $< \text{maximum equality distance}$
2. **else** Retrieve cases with distance $< \text{maximum allowed distance}$

- Filter-based retrieval
  - For locked attributes
  - Filters cases having “equal” values in locked attributes
  - Depends on the attribute tolerance.

- Similarity-based retrieval
  - For non-locked attributes
  - Based on a distance metric
  - Depends on: *Distance metric, weights, tolerance, maximum Allowed Distance, number of retrieved cases (M)*
MODSIMTex CBR - Adaptation

- If 1 equal case (distance(new case, case)< equality distance):
  - Copy
- If >1 equal cases (distance(new case, case<sub>i</sub>)< equality distance)
  - Mean/Mode
- Else
  - Weighted Mean/Mode of the M retrieved cases where distance(new case, case<sub>i</sub>) < maximum allowed distance
MODSIMTex CBR - Evaluation & Retain

- Evaluation of the proposed solution
  - Confidence of each proposed value:
    - Depending on the variance of the used cases’ values and missing values.
    - $\text{Conficence}(At_i) = 1 - \left( \frac{\sum_{j=1}^{M} \text{distance}(case_{new}(At_i), case_j(At_i))}{M} - \frac{\bigcup_{j=1}^{M} \text{isMissing}(case_j(At_i))}{M} \right)$, where M is number of used cases
  - Possible extension
    - Possibility to compare the CBR solution against the real process parameters.
- Retain: only real values of process are stored.
MODSIMTexCBR CBR Diagram

Delta = maximum distance to determine if two cases are equal.
Delta’ = maximum distance allowed to take into account the compared case.

Yes -> Adapt
## CBR - Generic Tools

<table>
<thead>
<tr>
<th>Shell</th>
<th>Case</th>
<th>Case Base</th>
<th>Retrieve</th>
<th>Reuse</th>
<th>Revise</th>
<th>Retain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caspian</td>
<td>&lt;attribute&gt;</td>
<td>Flat</td>
<td>Search guided by user</td>
<td>Repair Rules by user</td>
<td>User</td>
<td>User</td>
</tr>
<tr>
<td></td>
<td>Solution:n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JColibri2</td>
<td>&lt;attribute&gt;</td>
<td>Flat</td>
<td>kNN</td>
<td>NULL</td>
<td>User</td>
<td>user</td>
</tr>
<tr>
<td></td>
<td>Solution:n</td>
<td></td>
<td>Distance: global &amp; Local</td>
<td>Numerical direct proportion</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Result &amp; Reason</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IUCBR</td>
<td>&lt;attribute&gt;</td>
<td>Flat</td>
<td>KNN</td>
<td>NULL</td>
<td>-</td>
<td>Custom Triggers. Delete old cases</td>
</tr>
<tr>
<td></td>
<td>Solution:n</td>
<td></td>
<td></td>
<td>Weighted Mean Weighted Majority</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>B-tree-backed</td>
<td></td>
<td></td>
<td>NULL</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(index by user)</td>
<td></td>
<td></td>
<td>Weighted Majority</td>
<td></td>
<td></td>
</tr>
<tr>
<td>myCBR</td>
<td>&lt;attribute&gt;</td>
<td>Flat</td>
<td>KNN</td>
<td>NULL</td>
<td>User</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Solution:1</td>
<td></td>
<td>Distance global &amp; local</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ReMind</td>
<td>&lt;attribute &gt;</td>
<td>ID3</td>
<td>ID3</td>
<td>Adjustment of difference (new case, retrieved cases)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Solution:1</td>
<td>Index by user. Automatic weighting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBR - Express</td>
<td>&lt;attribute&gt;</td>
<td>Flat</td>
<td>Trigrams Heuristics</td>
<td>-</td>
<td>-</td>
<td>User</td>
</tr>
</tbody>
</table>

© Miquel Sànchez-Marrè & Beatriz Sevilla, KEMLG, 2013
# CBR – Generic Tools

<table>
<thead>
<tr>
<th>Shell</th>
<th>Case</th>
<th>Case Base</th>
<th>Retrieve</th>
<th>Reuse</th>
<th>Revise</th>
<th>Retain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caspian</td>
<td><code>&lt;attribute&gt;</code> Solution:n</td>
<td>Flat</td>
<td>Search guided by user</td>
<td>Repair Rules by user</td>
<td>User</td>
<td>User</td>
</tr>
<tr>
<td>JColibri2</td>
<td><code>&lt;attribute&gt;</code> Solution:n Result &amp; Reason</td>
<td>Flat</td>
<td>kNN Distance: global &amp; Local</td>
<td>NULL Numerical direct proportion</td>
<td>User</td>
<td>user</td>
</tr>
<tr>
<td>IUCBRf</td>
<td><code>&lt;attribute&gt;</code> Solution:n</td>
<td>Flat B-tree-backed (index by user)</td>
<td>KNN</td>
<td>NULL Weighted Mean Weighted Majority</td>
<td>-</td>
<td>Custom Triggers. Delete old cases</td>
</tr>
<tr>
<td>myCBR</td>
<td><code>&lt;attribute&gt;</code> Solution:1</td>
<td>Flat</td>
<td>KNN Distance global &amp; local</td>
<td>NULL</td>
<td>User</td>
<td>-</td>
</tr>
<tr>
<td>ReMind</td>
<td><code>&lt;attribute&gt;</code> Solution:1</td>
<td>ID3 Index by user. Automatic weighting</td>
<td>ID3</td>
<td>Adjustment of difference (new case, retrieved cases)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CBR Gesconda</td>
<td><code>&lt;attribute&gt;</code> Solution:n (Evaluation) (Utility) (Source Cases)</td>
<td>Flat Hierarchy* SOM*</td>
<td>KNN Configurable global distance</td>
<td>NULL Mean/ Mode Weighted average Formula</td>
<td>User Simulation</td>
<td>User Learn Not Learn Conditions</td>
</tr>
</tbody>
</table>

© Miquel Sànchez-Marrè & Beatriz Sevilla, KEMLG, 2013