

# Categorizing Case-Base Maintenance: Dimensions and Directions

David B. Leake and David C. Wilson

Presented by  
Ramakrishnan  
MSCS

# Introduction

Why the spotlight on case-base maintenance?

- Size of CBR system increasing -> case base size also increasing.
- Large case base raise concern about the problem of efficient and accurate case retrieval.
- Need to control the size of case-base.
- Even in smaller case-bases, accumulation of inconsistent and inaccurate cases reduces efficiency of the CBR system. eg. Case base having real world data.
- A framework for CBM would help in identifying current best practices, problems and opportunities for further study.

## Defining Case-Base Maintenance

- CBM implements policies for revising the organization or contents (representation, domain content, accounting information or implementation) of the case-base in order to facilitate future reasoning for a ***particular set of performance objectives***.
- Maintenance of case base may revise the case base at different levels:
  - Implementation level
  - Representational level
  - Knowledge level
  - Revise the maintenance policies themselves

# Categorization schemes for CBM

- Goal of categorization scheme for CMB is threefold:
  - Shed light on current practices -> increase understanding of current CBM approaches.
  - Finding gaps -> new research opportunities.
  - Cataloging approaches to identify which is best suited for a particular performance objective.
- The framework categorizes CBM approaches as CBM policies that determine when & how CBR systems perform CBM.
- Maintenance policies can be described in terms of:
  - Data Collection-> How they gather data relevant to maintenance.
  - Triggering-> When to trigger maintenance.
  - Operation Types -> Types of maintenance operations available.
  - Execution-> How selected maintenance operations are executed.

# Data Collection

- Data Collection :
    - Gathers information about individual cases.
    - About the case base in part or whole
    - Or about the overall processing behavior of the CBR system.
  - Types of data that can be collected:
    - **None** – no collection at all. -> **non-introspective**.
    - **Synchronic** - considers a snapshot of current case base in part or whole to formulate policy.
    - **Diachronic** – collect data over time and identify trends to formulate policies.
- } **Introspective**

## Data Collection cont..

**Timing** : When the data collection is performed..

- **Periodic** – collection happens at a set frequency w.r.t to the CBR cycle. E.g. After each problem solving cycle.
- **Conditional** – collection is performed in response to a well-defined but non-periodic condition. E.g. No. of cases in the library reaches a threshold.
- **Ad Hoc** – happens under ill-defined conditions determined externally to the CBR system. E.g. User-initiated.

## Data Collection cont..

Type of integration of data collection with the CBR system:

- On-line – during the course of an active reasoning episode.
- Off-line – during a pause in reasoning, such as waiting for user input or when idle between reasoning episodes.

Choice between on-line or off-line may affect the resources that can be devoted to the analysis problem -> important in time-constrained processing.

# Triggering

When to trigger maintenance?

- The *timing* and *integration* aspects of data collection give a guideline as to when to trigger.
- Strategy selection can be done periodically, conditionally, on an ad hoc basis and in on-line or off-line mode.
- Conditional triggering can be subdivided into:
  - Space-based. E.g. filling a limited amount of case storage.
  - Time-based. E.g. retrieval time exceeds a threshold.
  - Result-based. E.g. system fails to solve given problem or the wrong case is retrieved.

# Operation Types

Different maintenance policies revise different types of information (*target type*) at different levels (*revision level*)

- Target Type: Revision operations can focus on 4 types of target :
  - Indexing structure.
  - Domain contents
  - Accounting information.
  - Maintenance policies themselves.
- Revision level: Revision operations can make revisions at 3 levels:
  - Implementation level – list to D-tree.
  - Representational level – correcting case formats coming from different sources.
  - Knowledge level – correcting an erroneous feature value.

## Execution

Execution is characterized by :

1. The timing of maintenance operation.
2. Integration with other system processing.

### **Timing of execution:**

- **None** – e.g. a maintenance policy that simply informs that maintenance is required without making any changes.
- **Periodic** – execution on regular basis
- **Conditional** – changes may be accumulated for batch execution depending on a condition.
- **Ad hoc** – e.g. when an expert is available for doing maintenance.

### **Integration with other system processes:**

- **On-line** – execution during reasoning episode.
- **Off-line** – execution between reasoning episode.

## Scope of maintenance

- **Broad** – operations that affect a large subset or the entirety of the case base have broad scope.
- **Narrow** – operations that affect a single case or a small subset of the case-base.

This dimension is especially useful when characterizing resource-bound processing.

# Meta-Maintenance by Lazy CBM

- Update old cases in a 'lazy' manner – when an old case is retrieved to solve a new problem, CBM can simultaneously fix any defects present in the old case.
- Similar to case adaptation but its aim is to repair a problem in the old case rather than to fit that case to a specific new situation.
- Thus when a change must be applied to the entire case base, CBM can either :
  - Make that change to all cases simultaneously.
  - Or generate a maintenance rule to update each case that is retrieved, when its retrieved.
- This can be viewed as meta-maintenance – maintaining the system's maintenance knowledge.

## Meta-Maintenance cont..

- With a lazy updating scheme, cases that are obsolete must be distinguished from cases that have been updated.
  - A maintenance rule can be discarded when it has been applied to a sufficiently large proportion of retrieved cases – case base may have been sufficiently modified for the problems the system tends to encounter.
  - If application of maintenance rule is inexpensive, then keep the case-base unchanged, instead create new maintenance rules from old ones to obtain desired changes.

## Further exploration

- The role of usage trends of the maintenance rules in guiding maintenance.
- Example of trend that might be exploited – examination of patterns in the types of problems that are being solved – for a particular type of problem, a subset of the case base can be retrieved instead of entire case base.
- Integrate the design of maintenance rules with the design and performance objectives of the CBR system.



**Any Questions??**