# Methodological Assistance for Integrating Data Quality Evaluations into Case-Based Reasoning Systems

Annett Bierer

Chemnitz University of Technology
Faculty for economics and business administration
Chair of business process management and knowledge management
Reichenhainer Straße 39/611
D-09126 Chemnitz
Germany
Tel.: +49(0)371/531-33975; Fax: +49 (0) 371/531-26529
annett.bierer@wirtschaft.tu-chemnitz.de

**Abstract.** Case-based reasoning systems are used in more and more problem-solving domains supporting the long-term reusing and storing of experience. The performance of these systems essentially depends on the quality of the experience items in their knowledge base, represented as data. Defects in the quality of these data may interfere with the system's performance. By means of inspection and review the data quality is measured, evaluated, assured and improved. To support these activities in a case-based reasoning system, data quality criteria and control processes are required. Previous work in the field of data quality in case-based reasoning remains at a comparatively coarse-grained level. Existing approaches mostly do not provide sufficient methodological assistance in defining fine-grained quality criteria or designing and implementing control processes for the measurement and evaluation of the data quality. Therefore this paper proposes two approaches for methodological assistance in developing data quality inspections and data quality management for case-based reasoning systems.

**Keywords:** data quality, data quality management, closed loop control, goal-question-metrics-approach.

## 1 Introduction

Technologies for reusing and storing experience based on case-based reasoning (CBR)-methodology are now mature and case-based reasoning systems (CBR systems) are increasingly applied for long-term use in practice. The performance of CBR systems is essentially affected by the quality of the experience items, contained as data in their knowledge base (hereinafter referred to as data quality). If the current data quality level drops beyond an expected level, quality defects are indicated. They, in turn, are signs of errors. The notion of error includes all differences between a current data quality level and a required one,

R.O. Weber and M.M. Richter (Eds.): ICCBR 2007, LNAI 4626, pp. 254–268, 2007.

such as differences to the experience in the real world as well as entry errors and processing failures. Causes of insufficient data quality can be found in both the development and the use of a CBR system [14]. Causes of defects in system development may result from insufficient surveys and analyses of the relevant experience, from an inappropriate representation of the experience in the data and data structures, from implementation faults, or from testing and training. Causes of defects during system use, on which the paper focuses, may result from entry errors (e.g., entry of queries, acquisition of cases) and from changes in the relevant environment (e.g., incremental advances in the experience, radical or sudden technological and organisational changes).

To keep the data quality in the knowledge base at a constant level during the whole system's lifetime, the CBR system must be able to evaluate its current data quality level at any time. Developers and administrators are faced with the difficulty of several data structures for heterogeneous experience items, which will have different requirements for their measurement and evaluation. This is because:

- the data, representing heterogeneous experience items, show a varying sensitivity with respect to errors and failures,
- not necessarily every defect must directly result in triggering and executing maintenance operations, and
- no general criteria are applicable to the evaluation of the data quality of the knowledge base.

Numerous contributions in case-based reasoner maintenance provide learning algorithms for assuring the quality for several experience items. But there are only few criteria (e.g., problem-solution regularity, problem-distribution regularity, efficiency, competence) available which provide a sufficient granularity for the purposive execution of these algorithms. More analytical work is needed to identify potential sources for causes of defects in the knowledge base.

This paper presents initial steps towards an understanding of the importance of fine-grained evaluations of data quality in the knowledge base. After introducing a case study and the framework for the examination, an approach to define, analyse and interpret measures for the evaluation of data quality in CBR systems will be introduced. For the evaluation of data quality assurance and improvement, control processes based upon the principle of closed-loop control will be examined.

## 2 A CBR System for Quotation Processing

For a better understanding of main issues of the paper, the descriptions are illustrated by a simple case study. Its central components in the knowledge base will be presented in the following [19,22].

The knowledge base of a CBR system for assistance in quotation processing contains experience about the interrelation between technological features and manufacturing costs for rolls for the napping of fabrics. It consists of the following components:

- The cases in the case base are composed of the problem part that describes technological features in the form of attribute-value pairs, which are central for estimating manufacturing costs, and the solution part that describes among others cost accountings which will be dynamically generated in reference to a costing database. It also includes an explanation part for further information.
- The vocabulary contains characteristics of the technological features such as whether the feature is numeric or symbolic and referencing specifications for the data interchange with the costing database.
- The features are partly numeric (e.g., length of the roll, diameter) and partly symbolic (e.g., impact of fabric draft). Because of this, knowledge-intensive local similarity measures are used (for the notion of knowledge-intensive similarity measures see [27]).
- For the adaptation of a proposed solution to the current problem situation the solution transformation includes methods for similarity-based cost estimation as known from costing in early engineering stages [8,9].
- Additionally, membership functions are needed for the mapping of numeric features to linguistic terms in order to guarantee a numeric computation for vague feature values in queries (e.g., the length of the roll is given as "medium" to indicate that the length would probably be somewhere between 300 and 700 mm).

## 3 Framework for Examination

The performance preservation of the CBR system is an issue of operational data quality management. Its main task is the continuous improvement of the data quality in the knowledge base.

### 3.1 Why Data Quality?

CBR systems as for quotation processing combine functionalities and properties of CBR and database and data warehouse techniques respectively. CBR provides a structuring for the problem-solving processes and the idea of storing concrete technical and costing episodes as cases. The content of the knowledge base is represented as data in heterogeneous databases (e.g., the solution description is contained in a costing database, the vocabulary is described in a technical database). Reasoning algorithms process the data and present it to the users. The human user interprets the data as information. Knowledge is what the user needs in order to perform the interpretation process, and what he gets learned from the new information [1]. The focus of the paper centers the underlying data which are the basis for executing the problem-solving processes and enabling the users to make more precise decisions.

### 3.2 Data Quality and Data Quality Management

Today's comprehension of quality comes from the Latin term "qualitas", which means characteristics, constitutions or conditions of objects and is not valued

[31]. Data quality can be described from several perspectives such as the view of objective characterisations of the data or the view of the users [12]. Analogous to the general notion of quality in the norms of the Deutsches Institut für Normung (DIN e.V.) and based upon the user-based approach data quality means: the totality of features of the data in the knowledge containers that fulfil stated or implied needs [7]. It is a multidimensional measure for the applicability of the data in order to fulfil the purpose of their acquisition/generation, whereas this applicability may change if the underlying needs are changing with time [21,31].

Data quality management refers to all activities within the framework of a quality management system that constitutes data relevant aspects of the quality policy, goals and responsibilities and their transformation into operational planning, controlling, assurance and improvement processes [7]. The management of data quality is in fact an executive function, whereby the quality-related tasks have to be integrated (e.g., specification of data quality strategies) at all management levels. The higher the management level the more data quality abstracts from the specific information system, here from the CBR system.

This paper only considers operational criteria and means for the design of quality inspection processes for evaluating data quality within a CBR system based on the user's needs [15,28].

### 3.3 Phases of Operational Data Quality Management

The activities of operational data quality management are implemented with reference to the contexts of the enterprise and the considered CBR system. Against the background of continuous quality improvement the process-related plan-do-check-act-cycle (PDCA-cycle) by [5] has been established for structuring the activities. The cycle represents the idea of a continuous quality improvement process through the cyclic sequence of the phases data quality planning (plan), control (do), assurance (check) and improvement (act) (Fig. 1).

**Data quality planning.** In data quality planning the expectations and needs of the users (e.g., in the domain of quotation processing) are acquired and gradually transferred into guidelines for the design of the data in the knowledge base. Goals and requirements for data quality are defined, metrics are derived, categorized, weighted and appropriate measuring points and methods within the CBR cycle are selected [13,28]. The outcome of the phase is a data quality plan, which includes the requirements, the needed processes, resources, means and methods for measuring, and required records for the verification of conformity of the provided data with the experience in the real world [14].

**Data quality control.** After the planning the current data quality status has to be checked and verified every now and then. The execution of checking and verifying is the function of data quality control. The aim is to hold the fixed quality specifications and to guarantee the mastery of the required processes [13]. For the achievement of the objectives, data quality control is responsible for the monitoring and controlling of data quality and the initiation of maintenance

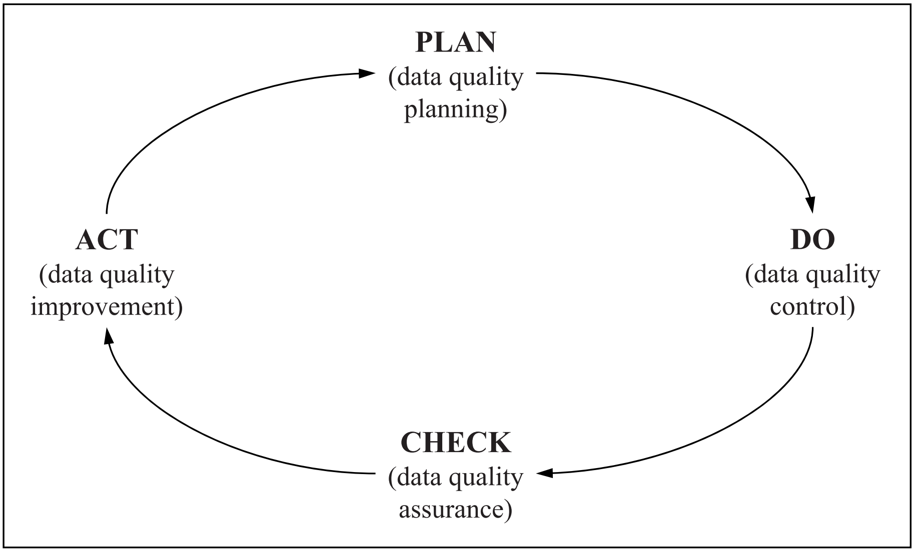


**Fig. 1.** Operational tasks in data quality management [24]

operations. Analysing to which extent the data quality and the requirements diverge is an issue of quality inspections [28].

**Data quality assurance and improvement.** Activities in data quality assurance and improvement are the initiation and execution of operations necessary to assure or to restore a required data quality level or even to increase it. Assurance, for instance, contains organisational arrangements like raising the user's awareness of data quality in the CBR system or automatically checking the inputs from users when entering queries or adding new cases. Restoring and improving are data administration tasks. These comprise maintenance operations, which in turn encompass case-based reasoner maintenance.

## 4 Metrics for Measuring and Evaluating Data Quality

The assessment of data quality presumes the setting of goals, metrics and measures respectively and their context-dependent interpretation derived from the goals. There are a variety of approaches and mechanisms to assist this top-down process. In this paper the Goal-Question-Metrics-Approach (GQM) will be used.

### 4.1 The Goal-Question-Metrics-Approach

For a CBR system to measure and to assess its data quality in a purposeful way, it requires (1) the specification of goals for the data quality in the knowledge base and the CBR system as whole, (2) their transfer to measuring data to operationalise the goals, and (3) the provision of a framework for domain-dependent interpretation of the measuring data to understand the goals [2]. First, the user requirements are acquired as accurately as possible in order to deliver quantifiable measures. These measures provide the basis of comparison for analysing the achievement of objectives with reference to the data quality level.

GQM is a systematic approach for defining and evaluating a set of operational measures, based on measurements. GQM can assist the adjustment and integration of goals, processes, components and component models into the CBR system. The approach assumes that measuring and evaluating the data quality requires [3]:

– the setting of goals for the quality of the data in the knowledge base (conceptual level),
– the refining of the goals into a set of quantifiable questions to characterize the assessment of specific goals (operational level), and
– the definition of measuring data associated with the questions in order to answer them in a quantitative way.

Every combination of these three levels builds up a hierarchical structure, which is called a GQM-model.

### 4.2 Deriving Data Quality Measures

Deriving data quality measures and evaluation models by means of the GQM-approach requires the existence of operationally defined data quality goals. In order to derive them the gap between the user's needs and their representation in goals has to be closed. In doing so, pyramids of needs may be built.

User requirements represent subjective expectations of the users regarding the performance of the CBR system. Usually, expectations are not directly transferable into data quality goals. They must be split up gradually until a suitable level for deriving questions is reached (Fig. 2) [16].

In a pyramid of needs each higher need (primary need) is decomposed step-by-step to partial needs (secondary, tertiary needs etc.) step-by-step until the transformation into product and process characteristics and the derivation of goals are feasible. For better understanding the figure below (Fig. 2) visualises the primary need "retrieval of experience-activating data for the efficient estimation of expected manufacturing cost", in short: *information*. Please note that this need has already been "translated". This means the voices of the users have already been transferred into a language understandable by the team of developers. The need information, for instance, can be split up into the three top-level-goals for CBR systems [25,26] *competence* (the range of cost estimation queries solved), *problem-solving efficiency* (e.g., the average retrieval time), and *solution quality* (e.g., the error level in the proposed solution). *Competence* may be split up further into *coverage* and *reachability* of the case base [25] or *problem-solution regularity* and *problem-distribution regularity* [18,30]. These tertiary needs have

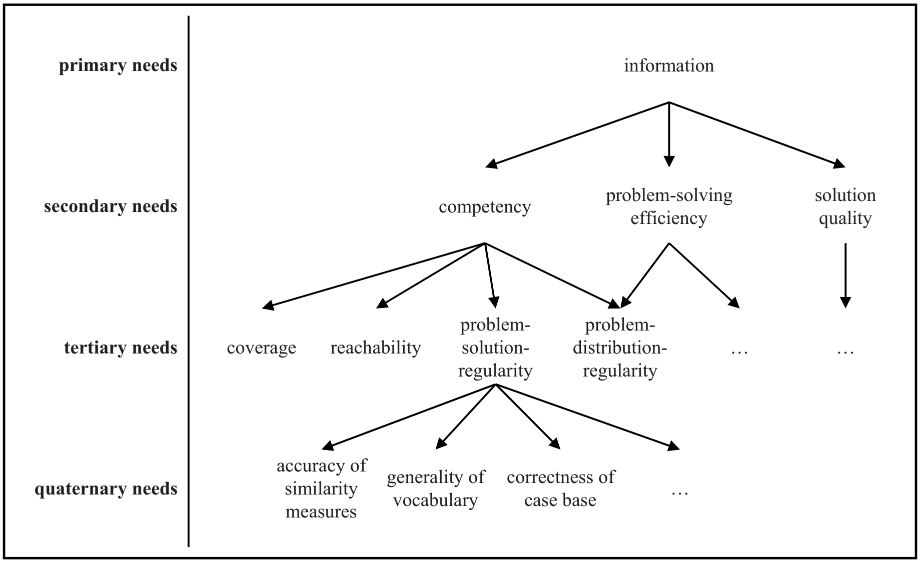


**Fig. 2.** Pyramid of needs for the example "information"

a relatively low granularity as a basis for assessing the data quality. Therefore an exclusive view on sinking *problem-solution regularity* in a CBR system is not sufficient for displaying directly whether there are defects in similarity measures, vocabulary or in the case base. Further analyses for identification of the knowledge base components causing the defects will be necessary. Therefore the tertiary need *problem-solution regularity* is here split up into quaternary needs like *accuracy of similarity measures*, *generality of vocabulary*, *correctness of the case base*.

The construction of pyramids of needs shows in a simple way the necessity for different goals and metrics for evaluating the data quality in a CBR system. Here, the quaternary needs are taken as a starting point for goal setting in GQM.

**Data Quality Goals.** Goals are defined for measuring objects (e.g., cases, retrieval, local similarity measures), for some purpose (e.g., evaluation, characterisation, improvement), with respect to various data quality criteria (e.g., minimality, consistency, speed), from various points of view (e.g., user, maintenance or experience engineer), and in relation to a relevant environment. Measuring objects are [3]:

- products: experience items, the data and data structures in the knowledge base, user queries or outcomes of the process steps retrieve, reuse and revise, which will be generated or processed during the system's lifetime;
- processes: activities associated with time like the processes of retrieve, reuse, revise itself;

– resources: employees as suppliers of experience, source systems etc., that are used in CBR for generating the outcomes.

The selection of an object must guarantee that it is possible to analyse and interpret its quality level directly with respect to a given data quality criterion, to expose the causes of its defects, and to repair the defect. As stated above a goal in GQM consists of the following components [2]:

**purpose:**
Analyse *(objects: products, processes, resources)*
for purpose of *(evaluation, decrease, improvement, ...)*
**perspective:**
with respect to *(quality criteria: timeliness, accuracy, validity, ...)*
from the point of view *(user, maintenance engineer, developer, ...)*
**environment:**
in the following context *(personal-, resource-, process-related factor, ...)*

Taking this scheme, an example for a concrete data quality goal is formulated as follows: "Analyse the *similarity assessment in the retrieval sets* for the purpose of *evaluation* of the *accuracy* from the point of view of an *experience engineer* depending on *advances in the experience of the human experts*".

**Questions and Metrics.** For assessing the achievement of goals they are refined into several questions. The questions must be qualified for characterizing the object of measuring with reference to the defined goal. The questions focused on the data quality goal break down the issues into major components for specific analyses. Please note that several GQM-models can have some questions in common.

After that each question is refined into a set of significant metrics. Several metrics may be used in order to answer different questions. For example the number of incorrect cases in a retrieval set for a given query may be a metric for both the generality of vocabulary and the accuracy of similarity measures.

Based on the metrics, skilled experience engineers and/or maintenance engineers as well as the CBR system, by itself, are able to evaluate the data quality of an object of measuring by analysing and interpreting its current values. In practice, several metrics may be aggregated to higher figures and key performance indicators respectively (e.g., the accuracy level of similarity measures and correctness level of the case base are aggregated for assessing the problem-solution regularity). When the domains of metrics differ, no aggregation by mathematical equations is feasible. But it is possible to combine and visualise their contribution to a higher-ranked figure by radar charts or other methods. The following table shows an example for questions and metrics using the GQM-goal defined above (Table 1).

The identification of questions and metrics is a nontrivial process, because both deriving questions and refining and relating appropriate metrics depend on various factors, among them (for the knowledge base of the case study):

**Table 1.** Example for questions and metrics for the data quality goal of evaluating the accuracy of similarity assessment

| | |
|---|---|
| Goal: | Analyse the similarity assessment in the retrieval sets for the purpose of evaluation of the accuracy from the point of view of an experience engineer depending on advances in the experience of the human experts. |
| Question: | What is the current accuracy of similarity assessment? |
| | Metric:<br>average maximum similarity of retrieved cases, average number of queries with no really "most similar" case in the retrieval set |
| Question: | Is the accuracy of similarity assessment improving with time? |
| | Metric: $\frac{\text{current average of accuracy}}{\text{baseline average of accuracy}} \times 100$<br>or a subjective rating by an experience engineer after consulting with the users |

- The understanding of correlations between technological features and the manufacturing cost for the development of suitable similarity measures with the developers, experience engineers and users of the CBR system.
- The existence of assessable and generalisable cost effects that will be needed for the definition of universally valid local similarity measures.
- The approach used for case representation (for an overview of case representation approaches see [4]) which for example determines what kind of similarity measure is needed or what has to be included in the vocabulary.
- The maturity of the measuring object similarity measures depending on the status of the system's life cycle. If the CBR system has reached a steady state, the metrics must allow the comparison with the real-world experience. If the CBR system is at the training stage, the metrics and their interpretation must allow learning and tuning the similarity measures up to an acceptable level for practical application.
- The learning process for refining and adapting the GQM-models. The defined metrics, must help in evaluating not only the measuring objects but also the reliability of the evaluation models.

The GQM-approach enables experience and maintenance engineers to define and to interpret operational and measurable knowledge bases for CBR systems. Because of numerous and various factors affecting the construction of GQM-models, the GQM-processes are usually nontrivial and highly contextual.

## 5 Processes for Measuring and Evaluating Data Quality

The defined goals and metrics may not be sufficient without appropriate measurement, evaluation and maintenance processes. Specific control processes for data quality inspections must be designed in order to meet a high-quality knowledge base. In conventional quality management the control process design is often based upon the cybernetic principle of closed-loop control [13,20].

### 5.1 Role and Types of Data Quality Inspections

For the assessment of a current data quality level and its comparison with fixed requirements quality inspections are useful instruments. There are several types of inspections [28]:

- static inspections used for dated off-line reviews of the experience-related data in the knowledge base in order to check their conformance with the defined goals (e.g., checking the case base for inconsistencies or redundancies);
- dynamic inspections used for on-line and off-line reviews by counts and tests associated with time in order to check the performance of the knowledge base over time (e.g., changes in accuracy of the similarity measures since start of the system's use); and
- defects and error analyses used for checking and revising errors, faults or failures documented error lists by the users during an active problem-solving cycle (e.g., user documents erroneous cases).

### 5.2 Closed-Loop Control as a Process Framework

Closed loop control is a process cycle. It is based on serial measuring of controlled variables. The variables are compared with some external reference value and are modified when the resulting differences are outside the limits [6]. For visualising the process of control control loops are constructed.

In the following the process of closed loop control [11] will be described in terms of the data quality goal of "analysing the feature weights for the purpose of evaluation of their accuracy from the point of view of the experience engineer and in the following context: the company has bought a new machine reducing the manufacturing costs for lathing the diameter of the rolls" (Fig. 3).

Closed loop control in the example aims at evaluating the data quality level of the feature weights in the global similarity measure of the CBR system for quotation processing (*controlled system*). The reason for checking the accuracy of the feature weights is to account for changes in the distribution of the manufacturing costs because of the new lathe (*disturbance variable*). First, the current values of the feature weights and the cost distribution are measured (*controlled variables*: weights, cost distribution). The current values are compared with the changed cost distributions (*reference variables*: distributions of costs because of the new machine). Then potential modifications are investigated by test retrievals and when the differences in the similarity assessments with the old and the new cost distribution are too high, maintenance operations are initiated (*controller*: comparison, test retrievals, initiating operations). Maintenance operations that could solve the inaccuracy of the feature weights have to be selected and executed (*manipulating variable*: selection, timing and realisation of operations automatically or through interaction with the maintenance engineer).

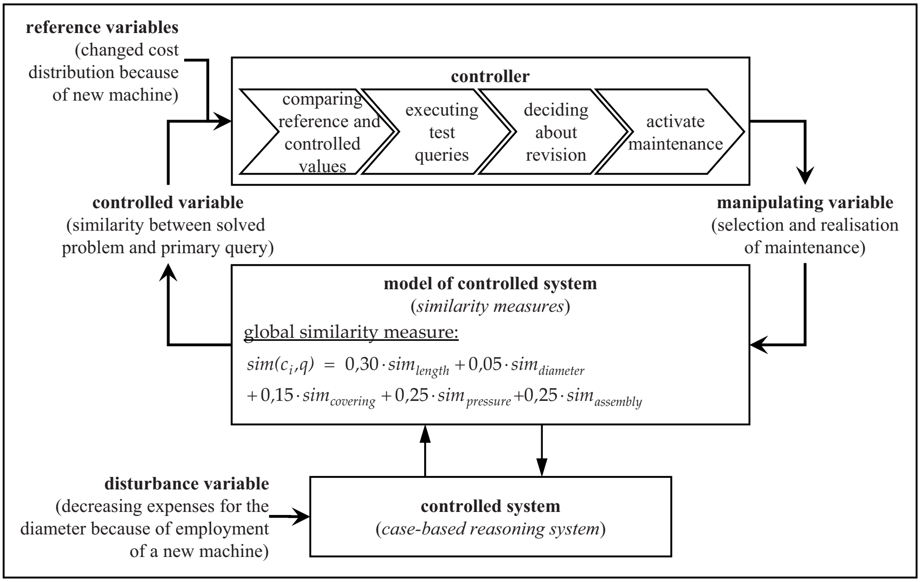


**Fig. 3.** Example of a control loop (following [11])

CBR systems bear some complexity in the knowledge base. Because of this, control processes do not operate directly on the controlled system but with models of the CBR system. The models contain only the components and relations of the knowledge base that are required for the respective process (*model of controlled system*: here the similarity measures and the case base is needed). The controlled system and its models interact by sensors (forwarding the disturbance and measuring data) and actors (reactivate the system after restoring the feature weights).

The example illustrated in Fig. 3 is only a simple instance of constructing and executing closed loop control. Usually, in real-world applications control loops are not as simple as that. In practice, hierarchically structured control loops are also needed. In this case, higher ranked loops will determine the reference variables and their values for subordinated stages and these in turn will be based upon the controlled variables of their subordinated control loops [13]. When assuming for the example above instead of changing the feature weights new features become necessary because of buying the new machine, changes are essential not only in the similarity measures but also in the vocabulary and the case base. The control loop for measuring, reviewing and restoring the vocabulary could provide disturbance variables and values for the controlled variables of the control loops for the evaluation of the similarity measures and the case base.

### 5.3 Integrating Control Loops in the Case-Based Reasoning Cycle

In addition to designing the basic sequences of closed loop control, the example above raises the question of when, where and how to specify and to integrate the three main tasks for controlling the data quality (measurements, evaluations and modifications) into the CBR system's processes efficiently. There are a variety of potential combinable strategies, which can be categorized as follows [17,29,30]:

- Strategies for the *scope* determine whether the main tasks affect only one component of the knowledge base (local), multiple components (multiple) or the whole knowledge base (global), or no inspections are made at all. Another differentiation would divide into operations affecting only a small data set (narrow) or a large data set or the whole knowledge base respectively (broad).
- Strategies for *triggering* determine the timing of the tasks. Triggering can be done at a set frequency (periodic), at every problem-solving cycle (continuous), in response to well-defined but nonperiodic conditions (conditional) or at externally given, nonperiodic and irregular conditions (ad hoc).
- Strategies for *integration* in the CBR cycle define whether one or all tasks are executed during an active problem-solving cycle (on-line) or during a pause of reasoning or in a separated maintenance cycle (off-line).
- Depending on the *integration of the users* into the control processes, especially the maintenance engineer, the tasks are executed by hand (manual), in interaction between the maintenance engineer and the system (interactive) or autonomously without human interaction (automatic).

By the combination of several strategies and their integration in the processes of control loops it will be possible to instantiate continuous data quality improvement processes. The postulation that the CBR system has to be able to evaluate the data quality of its knowledge base at any time requires at least the integration of measurement and evaluation into the CBR cycle.

For static quality inspections, measurement as well as the evaluation and modification are realised off-line in a separated maintenance cycle with the steps review and restore [23]. They are carried out off-line and may be periodic, interactive or automatic and local, multiple or global.

For dynamic quality inspections as well as defects and error analyses, collections of measuring data are carried out during an active problem-solving cycle (on-line, continuous, local, multiple, global). In order to enable the system to collect the measuring data the classical CBR steps retrieve, reuse, revise and retain will be enhanced by additional specific tasks (e.g., measuring data for number of queries) and methods (e.g., marking initially mismatches by an automatic counter). The continuous calculation of current averages of the metrics could be carried out during the active problem-solving cycle, too. The evaluation, in terms of comparisons of controlled variables and reference variables in the controller, is not integrated in the problem-solving cycle but in the maintenance cycle. The controller undertakes, autonomously or through interaction with the maintenance engineer, the task of reviewing the data quality level. When the measured and the required data quality level differ with reference to a particular quality goal, maintenance operations must be initialised. The selection and execution of these operations are issues of the restoring process and correspond with the tasks of the manipulating variables.

### 5.4 Combining Strategies for the Example in the Control Loop

For revising the feature weights to meet the changed distribution of manufacturing costs it is assumed that the revision can only be achieved by a manual adaptation executed by the maintenance engineer. At the time the revision has to be executed, no training data are available for the application of a learning algorithm for feature weighting. For this reason, a manual modification instead of an automatic learning is preferred here.

The control process is fully executed during a maintenance cycle (off-line). It is a static inspection because there are no dynamically collected measuring data that point out the technological change. The global similarity measure is the only component affected by the modifications (local). The process is triggered by the disturbing event of the new and more cost-effective machine (conditional). The activities in the controller, especially the comparison of weights before and after the reallocation and test retrievals for analysing the impacts of the changed weights, are triggered by the maintenance engineer in assistance with the experience engineer and the system (interactive). The maintenance operation consists of revising the feature weights in the global similarity measure and restoring the knowledge base for further application (manual).

After reactivation the case-based reasoning system data collection must be intensified for evaluating whether the accuracy of the modified weights meets the real world cost effects.

## 6 Conclusion

The growing and more long-term use of CBR systems requires fine-grained measurements and evaluations of the data quality in the whole knowledge base.

The presented methodologies are useful instruments for defining data quality goals and metrics as well as for designing and implementing control processes for continuous data quality improvement in the knowledge base. The integration of closed loop control processes into the CBR cycle as well as an enhanced maintenance cycle enable the measurement and evaluation of quality at a more fine-grained level. It is worthwhile to stress the fact out that both approaches are compatible with the strategies, frameworks, maintenance and learning algorithms developed for case-based reasoner maintenance.

The case study of the CBR system for quotation processing used to illustrate central aspects of the approaches is a relatively simple example. However, the idea of fine-grained quality evaluations and the approaches may be transferred to more complex CBR systems with different approaches for case representation and in several application domains.

## References

1. Aamodt, A., Nygård, M.: Different roles and mutual dependencies of data, information, and knowledge - an AI perspective on their integration. In: Data and Knowledge Engineering, vol. 16(3), pp. 191–222 (1995)

2. Basili, V.R.: Software modelling and measurement: The Goal Question Metric Paradigm. Technical Report CS-TR-2956, Department of Computer Science, University of Maryland (September 1992)
3. Basili, V.R., Caldiera, G., Rombach, H.D.: The Goal Question Metric Approach. In: Marciniak, J.J. (ed.) Encyclopedia of Software Engineering, vol. 1, pp. 528–532. John Wiley & Sons, Chichester (1994)
4. Bergmann, R., Althoff, K.-D., Breen, S., Göker, M., Manago, M., Traphöner, R., Wess, S. (eds.): Developing Industrial Case-Based Reasoning Applications. LNCS (LNAI), vol. 1612. Springer, Heidelberg (2003)
5. Deming, W.E.: Out of the Crisis. MIT Press, Cambridge (1986)
6. Deutsches Institut für Normung e.V (Hrsg.): DIN 19226-1:1994-02. Leittechnik; Regelungstechnik und Steuerungstechnik; Allgemeine Grundbegriffe. Beuth Verlag, Berlin Wien Zürich (1994)
7. Deutsches Institut für Normung e.V (Hrsg.): Qualitätsmanagement, Statistik, Zertifizierung: Begriffe aus DIN-Normen. Beuth Verlag, Berlin Wien Zürich (1995)
8. Ehrlenspiel, K., et al.: Konstruktionsbegleitende Kalkulation in der Produktentwicklung. In: Kostenrechnungspraxis, vol. 1, pp. 69–76 (1996)
9. Ehrlenspiel, K.: Kostengünstig Konstruieren: Kostenmanagement bei der integrierten Produktentwicklung. Springer, Heidelberg (2005)
10. Eppler, M.J.: Managing Information Quality. Increasing the Value of Information in Knowledge-intensive Products and Processes. Springer, Heidelberg (2003)
11. Ferstl, O.K, Sinz, E.J.: Grundlagen der Wirtschaftsinformatik. Band I. Verlag Oldenbourg, München Wien (2001)
12. Garvin, D.A.: What Does Product Quality' Really Mean? In: Sloan Management Review, vol. 26(1), pp. 25–43 (1984)
13. Helfert, M.: Planung und Messung der Datenqualität in Data-Warehouse-Systemen. Dissertation Universität St. Gallen, Nr. 2648. Difo-Druck GmbH, Bamberg (2002)
14. Hinrichs, H.: Datenqualitätsmanagement in Data Warehouse Systemen. Dissertation Universität Oldenburg, Logos, Berlin (2002)
15. Institute of Electrical and Electronics Engineers (eds.): IEEE Standard 729-1983: Glossary of Software Engineering Technology (1983)
16. Juran, J.M.: Handbuch der Qualitätsplanung. Verlag Moderne Industrie, Landsberg/Lech (1991)
17. Leake, D.B., Wilson, D.C.: Categorizing Case-Base Maintenance: Dimensions and Directions. In: Smyth, B., Cunningham, P. (eds.) EWCBR 1998. LNCS (LNAI), vol. 1488, pp. 196–207. Springer, Heidelberg (1998)
18. Leake, D.B., Wilson, D.C.: When Experience is Wrong: Examining CBR for Changing Tasks and Environments. In: Althoff, K.-D., Bergmann, R., Branting, L.K. (eds.) Case-Based Reasoning Research and Development. LNCS (LNAI), vol. 1650, pp. 218–232. Springer, Heidelberg (1999)
19. Meyer, S.: Verarbeitung unscharfer Informationen für die fallbasierte Kostenschätzung im Angebotsengineering, Dissertation Technische Universität Chemnitz. Verlag der Gesellschaft für Unternehmensrechnung und Controlling, Chemnitz (2001)
20. Pfeifer, T.: Qualitätsregelkreis. In: Zollonds, H.-D. (ed.) Lexikon Qualitätsmanagement: Handbuch des modernen Managements auf Basis des Qualitätsmanagements, pp. 998–1002. Hanser Verlag, München Wien (2001)

21. Quix, Chr.J.: Metadatenverwaltung zur qualitätsorientierten Informationslogistik in Data-Warehouse-Systemen. Dissertation an der Rheinisch-Westfälischen Technischen Hochschule Aachen (2003) (request date: 26.07.2006) Online-Resource, http://sylvester.bth.rwth-aachen.de/dissertationen/2003/263/03_263.pdf
22. Rösler, M.: Kontextsensitives Kosteninformationssystem zur Unterstützung frühzeitiger Produktkostenexpertisen im Angebotsengineering. Dissertation Technische Universität Chemnitz. Verlag der Gesellschaft für Unternehmensrechnung und Controlling, Chemnitz (2005)
23. Roth-Berghofer, Th.: Knowledge Maintenance of Case-Based Reasoning Systems. The SIAM Methodology. Dissertation der Universität Kaiserslautern. Dissertationen zur Künstlichen Intelligenz Nr. 262. Akademische Verlagsgesellschaft Aka GmbH, Berlin (2003)
24. Seghezzi, H.D.: Integriertes Qualitätsmanagement – das St. Galler Konzept. Hanser Verlag, München Wien (1996)
25. Smyth, B., McKenna, E.: Modelling the Competence of Case-Bases. In: Smyth, B., Cunningham, P. (eds.) EWCBR 1998. LNCS (LNAI), vol. 1488, pp. 208–220. Springer, Heidelberg (1998)
26. Smyth, B., McKenna, E.: Footprint-Based Retrieval. In: Althoff, K.-D., Bergmann, R., Branting, L.K. (eds.) ICCBR'99: Case-Based Reasoning Research and Development. LNCS (LNAI), vol. 1650, pp. 343–357. Springer, Heidelberg (1999)
27. Stahl, A.: Learning of Knowledge-intensive Similarity Measures in Case-Base Reasoning. Dissertation Universität Kaiserslautern. Verlag dissertation.de, Berlin (2004)
28. Wallmüller, E.: Software-Qualitätssicherung in der Praxis. Hanser Verlag, München Wien (1990)
29. Wilson, D.C.: Case Based-Maintenance: The husbandry of experience. Dissertation of the Indiana University (2001)
30. Wilson, D.C., Leake, D.B.: Maintaining case-based reasoners: dimensions and directions. In: Computational Intelligence, May 2001, vol. 17, pp. 196–213 (2001)
31. Würthele, V.: Datenqualitätsmetrik für Informationsprozesse. Dissertation Eidgenössische Technische Hochschule Zürich. Books on Demand GmbH, Norderstedt (2003)