

# Case-Based Reasoning Applied to Geometric Production Measurements

Erik Olsson · Peter Funk · Alf Andersson

Received: date / Accepted: date

**Abstract** Measurements from products are continuously collected to allow adjustments in the production line to certify a feasible product quality. Case-Based Reasoning is a promising methodology for this type of quality assurance. It allows product measurements and its related adjustments to the production line to be stored as cases in a Case-Based Reasoning system. The idea is to describe an event of adjustments based on deviations in geometric measurement points on a product and connect these measurements to their correlated adjustments done to the production line. Experience will implicitly be stored in each case in the form of uniquely weighted measurement points according to their positive influence on adjustments. Methods have been developed in order to find these positive correlations between measurements and adjustments by analysing a set of historical product measurements and their following adjustments. Each case saved in the case base will be "quality assured" according to this methods and

only cases containing strong positive correlations will be used by the system. The correlations will be used to supply each case with its own set of individual weights.

**Keywords** Decision Support Systems · Experience Reuse · Case-Based Reasoning · Quality Improvement

## 1 Introduction

Production of cars is a very complex process which involves many process steps. Each step contributes to the variation of the quality of the final product. Therefore, it is essential to have good control over each process as well as the interaction between the ingoing processes. A car body consists of over 350 individual parts which have their own variations. These parts are assembled in several steps into the final product: the complete car body. The car body have it's specification to follow, in order to fulfil the demands for mounting of all other equipment in the car. If the car body is out of specification, problems will arise e.g.: bad fitting of interior parts, leakage of sound and water, closing problems with door, hood and other hang on parts. These are just examples of problems which arise due to that the geometry of the car body is out of specifications. The list of problems can be made much longer. Therefore, it is important to understand how each of the ingoing parts and processes affects and interacts with each other. In this study, performed at Volvo Car Corporation, the assembly process of one sub assembly cell (the gore line) for Volvo S80 was analysed. The approach was to investigate a set of selected measurements of the assembled part geometry and compare that to the process adjustments which have been done. We believe that efficient decision support is helpful in this kind of

---

Erik Olsson  
School of Innovation, Design, and Engineering Mälardalen  
University, P.O. Box 883, SE- 721 23 Västerås, Sweden  
Tel.: +4621107335  
Fax: +4621101460  
E-mail: erik.m.olsson@mdh.se

Peter Funk  
School of Innovation, Design, and Engineering Mälardalen  
University, P.O. Box 883, SE- 721 23 Västerås, Sweden  
Tel.: +46703339150  
Fax: +4621101460  
E-mail: peter.funk@mdh.se

Alf Andersson  
Volvo Car Corporation Manufacturing Engineering Dept.  
81280, Geo. Loc. 26G01 SE-293 80 Olofström  
and Product Development Chalmers University SE-412 96  
Gothenburg, Sweden  
Tel.: +46454265280  
E-mail: paander58@volvocars.com

situation and in this paper we present a method to connect measurements of car body parts to their resulting adjustments done in the production line. Our approach is to investigate a set of selected production measurements and their resulting adjustments to the welding frame. We are also investigating the outcome of the adjustments. The methodology of Case-Based Reasoning (CBR) is applied to enable a connection between measurements, adjustments and their outcome and enable improved performance with increased number of collected cases. This approach enables events of measurements, adjustments and their outcome to be connected and to be saved as cases in a CBR system for future use and re-use as previous similar cases. A case library of such cases is assembled and it will be made readily available to provide real-time decision support in any situation to technicians. Figure 1 depicts a situation of a case solving a problem occurring late in the production line that may result in a defective part if not corrected in time. The case connects measurement features from off-target parts (problems) to action taken previously to adjust production line (solution) to bring production back to target.

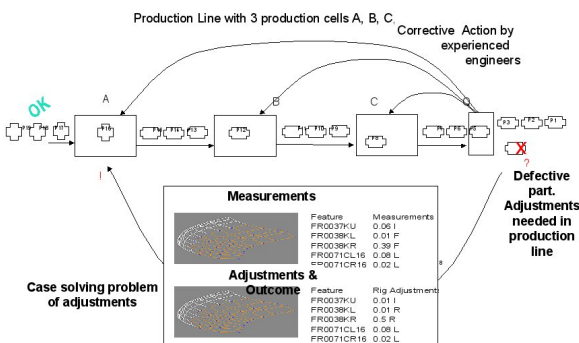


Fig. 1 Case solving a problem occurring late in the production line

In this kind of situation it is vital that the correct cases are retrieved to assure that the correct adjustments are done. We have developed algorithms especially suited for this context to provide efficient case matching, case retrieval and case adaptation to provide technicians with adequate data from previously stored cases of adjustments and their corresponding outcome. The paper is organized as follows. Section 2 gives a brief introduction to CBR. Section 3 describes our approach of CBR applied to geometric production measurements. Section 4 describes a prototype system involving a web interface, a server and a database developed for evaluation of our approach in a live production environment and finally section 5 concludes this paper with summary and conclusions.

## 2 Case-Based Reasoning

Case Based Reasoning is an Artificial Intelligence method that is inspired by the way humans reuse past experience to solve new problems. It offers an alternative to implement intelligent diagnosis systems for real-world applications [Watson, 1997]. Motivated by the doctrine that similar situations lead to similar outcomes. CBR is derived from instance-based learning which is a machine learning method [Mitchell, 1997] used in the artificial intelligence discipline. The technique of CBR had its theoretical origins in the mid 1970s and originally came from research in cognitive science [Watson, 1997]. It a feasible model of the reasoning process performed by our brain e.g. when we are subjected to stereotypical situations such as going to a restaurant or visiting a hairdresser. If a similar situation is encountered a second time, memories of these situations are already recorded in our brains and stored as scripts that inform us what to expect and how to behave. The original work in CBR was performed by Schank and Abelson in 1977. In 1983 Janet Kolodner developed the first CBR system designated CYRUS [Kolodner, 1993]. Cyrus was an implementation of Schanks dynamic memory model and contained knowledge, as cases, about the travels and meetings of a former U.S. Secretary of state. CBR has been known outside the research community since about 1990 when Lockheed began to use a CBR system named CLAVIER [Mark, 1989] for the baking of composite parts in an industrial oven.

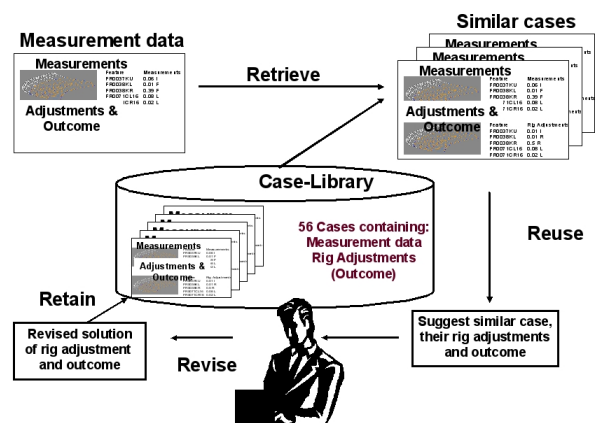


Fig. 2 CBR applied to geometric production measurements

CBR fits well to classify new measurements or sensor signals based on experiences of past categorizations [Nilsson et al., 2003]. The main strength lies in the fact that it enables directly reusing concrete examples in history and consequently eases the knowledge acquisition bottleneck. The CBR methodology is used to solve

new problems based on learning from similar cases (i.e. existing experience) stored in a case library that is obtained by storing previous similar situations. It bases its learning from past cases and it has the ability to build up experience, improve its performance and adapt to a changing environment. The CBR cycle is based on four main steps: Retrieve, Reuse, Revise and Retain. The quality of the retrieval process is measured in how well it identifies similar problems that may be reused. Adaptation may also be implemented in the reuse step. Adaptation enables one or more solutions to be adapted to solve a similar problem. A proposed solution may be revised to better fit the new problem before it is retained in the case base. A CBR system generally follow this cycle [Aamodt and Plaza, 1994] as depicted in figure 2.

The approach is especially suitable where simulation and modelling is too complex and adjustments cannot be calculated in real time. This is often the case in a real production environment where the relationship between the result and adjustment is so complex that it cannot be predicted, to many factors influence the outcome. Even so skilled technicians learn over time how to adjust in order to get the desired result, experience that may have been acquired through costly mistakes. CBR enables harvesting this experience and also transferring it between experts and to less experienced experts [Olsson and Funk, 2009].

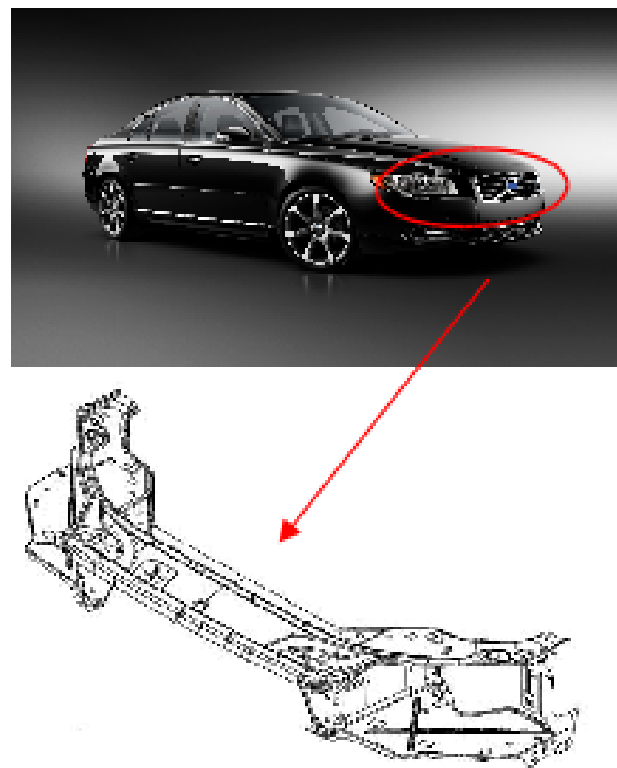
### 3 CBR Applied to Geometric Production Measurements

When a novel case is presented to the CBR system a list with similar cases will be presented. This is done using individual weights stored in each of the presented cases. These weights will amplify or attenuate the closeness of the stored case in relation to the novel case. The result will be a list of the nearest neighbouring cases containing candidate solutions according to each set of individual weights stored in each of the cases. Each individual case in this list has a global ranking in the list. The list can be used as decision support for technicians. Cases can easily be browsed using an intuitive web interface and the technician can decide which case or combination of cases to use to guide which adjustments to be done. In the above manner a possible scenario may contain a number of deviating measurements that may need a combination of stored solutions in order to make a complete adjustment. The technician may use a number of top-ranked cases from the list to perform the necessary adjustments. E.g. case#2 in the list may be similar to deviations in measurement points #1,2,5 presented in the novel case and case#4 in

the list may contain solutions to adjustments of deviating measurement point #6 presented in the novel case. The technician can now make a compound adjustment by combining information about adjustment done on case#2 and case#4 and if these adjustments prove to be successful this new compound case may be injected using the web interface into the case base to be used to guide future adjustments.

#### 3.1 Case Representation and Case Base

The position of the gore in the car can be seen in figure 3.



**Fig. 3** Position of the gore in the car

The assembly cell can be seen in figure 4.

The gore consists of nine ingoing parts which each one contributes with its own variation. The position of the ingoing parts is fixed in an assembly fixture. The internal position of the parts can be adjusted by adding shims at the position of the reference- and support points. Adjustments are used to compensate for the variation of the ingoing parts in order to reach the specification demands on the final sub assembly, the gore. The geometry of the parts are controlled by measurement of predefined points, spread over the geometry to



**Fig. 4** Assembly cell

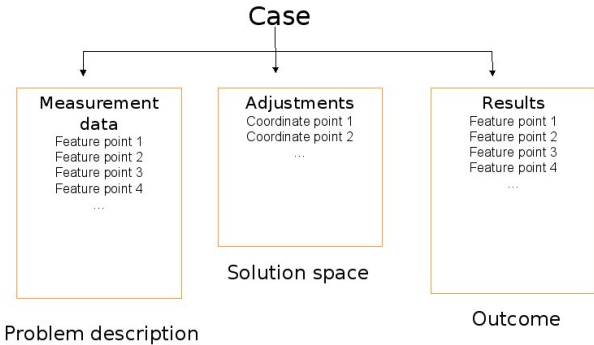
reflect the process stability. These points are measured with a Coordinate Measurement Machine (CMM) and indicates if the process are stable or not. If the results indicate that the part is out of specification, problem will arise when mounting the gore into the car and will generate problems such as lamp fixation etc. The parts are continuously measured and stored in a database. Trends can be followed over time since parts are measured during the entire lifetime of the part. Since the part is very sensitive for variations, many adjustments are required and the experience of the operator, who is responsible, contributes significantly to the quality of the assembled part. A tool for relating how adjustments in the assembly fixture affects the part geometry of the gore would be a valuable tool for the operator in order to take the correct decisions of adjustments of the assembly fixture.

Our first task was to populate the case-base with cases containing connections of measurements  $\rightarrow$  adjustments  $\rightarrow$  resulting adjustments. All measurements were stored in a database but adjustments done to the assembly fixture were manually recorded in a log binder. Our first task was to extract all useful information from the log binder about adjustments done. Secondly we needed to sync these adjustments to the measurements stored in the database. We used time stamps from the adjustment logs to find relevant measurements done just before and just after an adjustment had been made. The case base consists of such cases containing measurements, adjustments and outcome (results). Each case is represented with a unique case id. There are 56 cases in total including 290 measurement points, 30 adjustment coordinates and 290 resulting measurement points. Our intention is to use the measurement points as features representing a discrete event of adjustments. Consequently each case in

**Table 1** Adjustment point 11

Case2	Case4	Case5	Case8	Case13	Case15
0.2	0.2	0.4	-0.4	-0.3	0.3

the case-base represent a discrete event of adjustment made to the assembly fixture. Figure 5 depicts the case representation.



**Fig. 5** Case representation of geometric production measurements

### 3.2 Individual Weighting

To achieve an accurate and realistic matching of cases we intend to find significant measurement points and weight them accordingly. Individual weights are calculated according to relationships between measurements and adjustments using correlation analysis based on the information in the case base. Our intention is to find possible relations between one or more measurement points and one or more adjustment points. Also one or more adjustment points can relate to one or more resulting measurement points. Our main approach is to find all the cases with similar adjustment coordinates (e.g. all cases containing information about the alteration of a particular adjustment point) and in these cases we investigate whether a relation exists between its measurement points and its resulting measurement points. Example: There are six cases: 2, 4, 5, 8, 13 and 15 where adjustment point 11 is altered as depicted in table 1. We name this table  $Ad_{11}$ .

Measurement data from all the above cases is collected as depicted in table 2. We name this table  $Mea_{11}$ .

And result measurement data from the above cases is collected as depicted in table 3. We name this table  $Rslt_{11}$ .

In the above data we now look if a relation between measurement points ( $Mea_{11}$ ) and resulting mea-

**Table 2** The measurement data

	Case2	Case4	Case5	Case8	Case13	Case15
Mp1	-0.17	-0.17	-0.24	-0.10	-0.13	-0.22
...	...	...	...	...	...	...
Mp290	0.92	1.00	0.90	0.53	0.40	0.62

**Table 3** Resulting measurements

	Case2	Case4	Case5	Case8	Case13	Case15
Mp1	-0.05	-0.10	-0.01	-0.10	-0.13	-0.22
...	...	...	...	...	...	...
Mp290	0.92	1.00	0.70	0.13	0.40	0.20

surement points  $Rslt_{11}$  can be found. It means that the alteration of the particular adjustment coordinate 11 influence a change in affected measurement points and this change is visible when comparing measurement points in cases: 2, 4, 5, 8, 13 and 15 with resulting measurement points in the same cases. We name this differences matrix  $D_{11}$ . First, we make this comparison by calculating the difference between the measurements and the resulting measurement points ( 1).

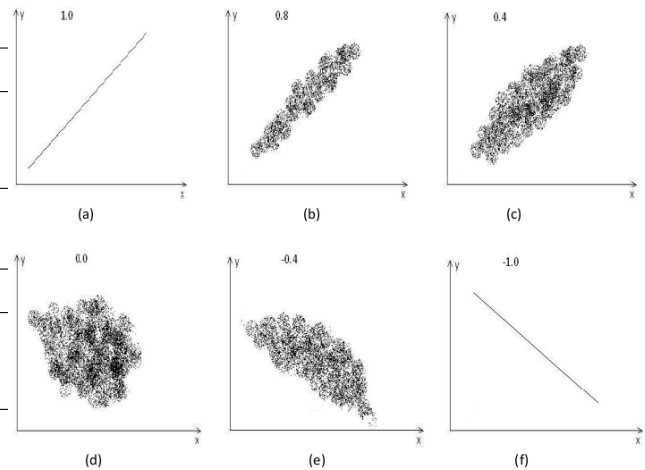
$$D_{11} = Mea_{11} - Rslt_{11} \quad (1)$$

Secondly, we use correlation analysis to find the relationship between this difference and the adjustments. In general statistical usage, saying that correlation exists implies that there is a relationship between two mathematical variables or measured data values. There are several correlation coefficients. The Pearson correlation coefficient is special among these coefficients as it is sensitive to a linear relationship between two variables. The definition of the Pearson correlation coefficient R is ( 2) [Rodgers and Nicewander, 1988]:

$$R = \frac{1}{n-1} \sum_{i=1}^n \frac{x_i - \varphi_x}{\sigma_x} * \frac{y_i - \varphi_y}{\sigma_y} \quad (2)$$

where  $\varphi_x$  and  $\sigma_x$  denote the sample mean and the sample standard deviation respectively for the variable x and  $\varphi_y$  and  $\sigma_y$  denote the sample mean and the sample standard deviation respectively for the variable y. R is in the range from -1 to 1. The case R=1 as shown in Figure 6 (a), represents the maximum possible linear positive connection between two variables. If R=-1, a maximum possible linear negative connection exists. The magnitude of R indicates the strength of the association, while the sign implies the direction of association.

With this information we can select candidate measurement points, which may be affected if a certain adjustment coordinate (e.g. adjustment point 11) alters [Gao, 2010]. If a strong enough relationships exist, the

**Fig. 6** Several sets of (x, y) points, with the correlation coefficient of x and y for each set

affected measurement point can be considered as important as it is shown to be influenced by adjustment point 11. We use this information to calculate weights accordingly. A criterion is set to indicate the strength of a relationship. The criterion has three grades of relationships; strong, medium and non-existent. Figure 7 depicts colour-coded measurements points (lines) according to their relationship with adjustment points (columns). The criterion is set as 0.1. A p-value lower than 0.1 is considered as a strong relationship and marked red while the yellow marked values higher than 0.75 are defined as non-existent relationships. The p-values of the white marked values are in the range of 0.1 to 0.75 and their relationship criterion is defined as medium. As an example from the below chart in figure 7, measurement point FR0179CL15/Y (no 6) is strongly affected by adjustment point 11 and 12, medially affected by adjustment point 27 and weakly (not) affected by adjustment point 31.

Feature	Char	NO	11	12	27	31
FR0037KU	Y	1	0,84	0,93	0,28	0,60
FR0038KL	X	2	0,59	0,43	0,52	0,02
FR0038KR	X	3	0,07	0,09	0,74	0,16
FR0071CL16	Z	4	0,76	0,78	0,98	0,42
FR0071CR16	Z	5	0,55	0,81	0,89	0,77
FR0179CL15	Y	6	0,01	0,01	0,55	0,83
FR0179CL17	Z	7	0,71	0,64	0,15	0,20
FR0179CR15	Y	8	0,58	0,44	0,19	0,32
FR0179CR17	Z	9	0,16	0,07	0,63	0,07
FR0244AL	NOR	10	0,96	0,87	0,11	0,19
FR0244AR	NOR	11	0,74	0,75	0,21	0,34
FR0245AL	NOR	12	0,21	0,24	0,93	0,23
FR2009AL11	NOR	13	0,24	0,35	0,58	0,76
FR2009AR11	NOR	14	0,19	0,11	0,25	0,11
FR2010AL08	NOR	15	0,66	0,68	0,60	0,66
FR2010AR08	NOR	16	0,18	0,10	0,64	0,39
FR2010BL09	NOR	17	0,52	0,35	0,39	0,23
FR2010BR09	NOR	18	0,23	0,37	0,07	0,11

**Fig. 7** Colour-coded measurements points (lines) according to their relationship with adjustment points (columns)

### 3.3 Case-Based Classification using Weighted Measurements

As mentioned before, the features of the problem part of a case consist of sample measurements from a finished body part and the solution part of the case consists of the adjustments done accordingly. Our approach is to match and classify a novel case based on comparing its problem part (ie its sample measurements) with the problem parts of the cases saved in the case base. When matching a novel case all measurements are vital but only cases representing partial (or full) similarities in deviating measurements will be retrieved and these cases will only represent events of successful adjustments with strong correlations between adjustments and measurements as described in the previous section. This is done using individual weights stored in each of the retrieved cases. These weights will amplify or attenuate the similarity of the stored cases in relation to the novel case. Generally the problem part of the novel case can be described as a feature vector of  $n$  measurement features and is formulated as ( 3).

$$FV = \{m_1, m_2, \dots, m_i, \dots, m_n\} \quad (3)$$

where each  $m$  is a geometric measurement. In comparison, each individually weighted case retrieved from the case base is formulated as ( 4):

$$FV = \{m_1w_1, m_2w_2, \dots, m_iw_i, \dots, m_nw_n\} \quad (4)$$

where each  $m$  is a geometric measurement and each  $w$  is a weight indicating the strength of the influence  $m$  as shown to have on resulting adjustments. We then perform case-based reasoning to achieve a list of the nearest neighbouring cases containing candidate solutions according to each set of individual weights stored in each of the retrieved cases. The retrieval process consists of the following steps:

1. Compare the feature vector with the known cases in the library by means of similarity calculation.
2. Select the  $k$  nearest cases exhibiting the highest similarity degrees.

Given a feature vector:

$FV = \{m_1w_1, m_2w_2, \dots, m_iw_i, \dots, m_nw_n\}$  its similarity degree with case  $C$  in the case library is defined as ( 5):

$$similarity(FV, C) = \sqrt{\sum_{i=1}^n w_i \times (FV_i - C_i)^2} \quad (5)$$

## 4 Prototype System

We have developed a prototype system involving a web interface, a server and a database for classification of geometric production measurements [Dahlgren and Castillo,

2010]. This has been done in close collaboration with technicians at Volvo. The system is developed in order to aid technicians to adjust production equipment according to geometric product measurements. The prototype is able to import production data and store in a database accessible from a web interface . The system has been made easy to maintain to allow further development as our research progresses. It has a layered design and by modularising the different functions the software is easier to evaluate. The use case depicted in figure 8 presents the main tasks of the system. The user can upload a measurement, enter adjustments in the adjustment log and finally upload the outcome measurement. A Case in the system consists of three parts:

1. Measurement values defined as out of tolerance.
2. Adjustments done accordingly and an associated log which shows what adjustments made, how much and why.
3. The outcome measurement values showing the resulting measurements after necessary adjustments.

A Case is defined as "unfinished" if it lacks information about adjustments and/or outcome measurements and it is defined as "finished" if it includes all the above different parts. The system retrieves the nearest case using data from the uploaded measurement and displays it to the user. The system also finds out which adjustment points affects the selected feature points and present them to the user. When the initial measurement and outcome measurement are uploaded and the adjustment data is entered the Case is considering finalised and the system saves it to the database. With data from the new case a new relationship matrix is calculated and the data in the database is updated.

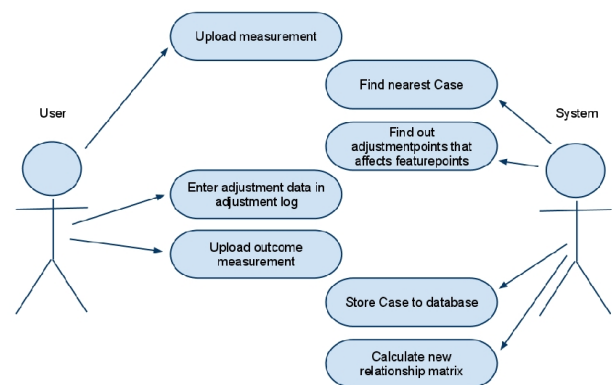


Fig. 8 System use case

The user interface is designed to be easy for the typical user; an engineer with knowledge about the production. All buttons are placed at the top of the application, this makes the interface consistent as the user

finds the controls at the same location. The division of the tasks: "create new case" and "finish case" is a natural dividing since the tasks are separated and they may not be done at the same time.

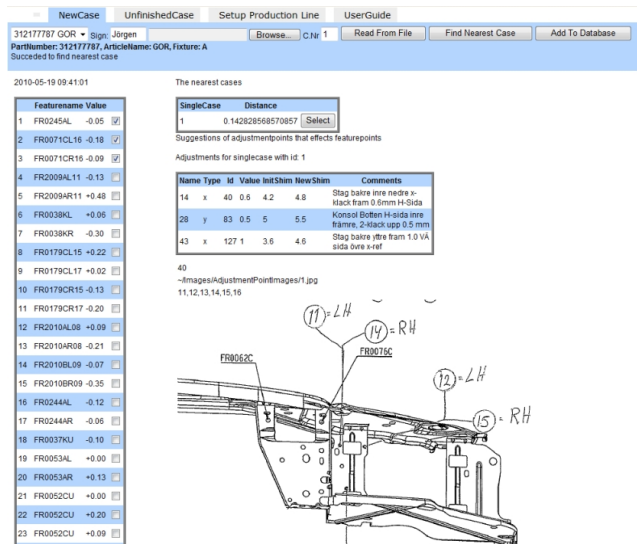


Fig. 9 Screenshot of the prototype system

Figure 9 displays a screenshot of the prototype system. It displays a nearest neighbour retrieval of the top-ranked case from the case base. Initial measurement values from a novel case can be seen in the column to the left of the screen. In the column; measurement values defined as out of tolerance can be selected to be used for case retrieval by the system. In the middle of the screen information about suggested adjustments from the top-ranked case is shown. Further information about relevant adjustment points are depicted in the illustration in the bottom of the screen. An evaluation of the prototype will be done in order to validate its performance as decision support and its on-site functionality at the factory. Hopefully it will aid technicians in their daily work at the production line.

## 5 Conclusions

We believe that our approach of applying case-based reasoning to geometric production measurements can provide decision support that helps technicians to make fast and correct adjustments to the production line and therefore make the production of car body parts more efficient. An efficient reuse of past cases of adjustments act as powerful decision support for technicians on all experience levels. The case library stores measurement → adjustment → outcome triplets capturing past experience. These cases can be reused and adapted to the

current situation and the system will over time collect considerable experience as the case library grows.

**Acknowledgements** Our thanks to the sponsors of eMaintenance2010 for their intellectual and financial support and to Volvo Car Corporation and ProViking SSF who funded this case study.

## References

- A. Aamodt and E. Plaza. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *Artificial Intelligence Com*, 7:39–59, 1994.
- E. Dahlgren and P. C. Castillo. Development of web interface, server and database for intelligent process surveillance with focus on geometry variations. BSc Thesis, Volvo Car Corporation, Uppsala and Mlardalen University, Sweden, 2010.
- C. Gao. Connect measurements of car body parts to adjustments in production line. MSc Thesis, Volvo Car Corporation, Uppsala and Mlardalen University, Sweden, 2010.
- J. Kolodner. *Case-Based Reasoning*. Morgan Kaufmann, 1993.
- B. Mark. Case-based reasoning for autoclave management. *Proceedings of the Case-Based Reasoning Workshop*, 1989.
- T. M. Mitchell. *Machine Learning*. McGraw-Hill, 1997.
- M. Nilsson, P. Funk, and M. Sollenborn. Complex measurement in medical applications using a case-based approach. *Workshop Proceedings of the International Conference on Case-Based Reasoning*, pages 63–73, 2003.
- E. Olsson and P. Funk. Agent-based monitoring using case-based reasoning for experience reuse and improved quality. *Journal of Quality in Maintenance Engineering*, 15, July 2009.
- J. L. Rodgers and W. A. Nicewander. *Thirteen ways to look at the correlation coefficient*, volume 42. The American Statistician, 1988.
- I. Watson. *Applying Case-Based Reasoning: Techniques for Enterprise Systems*. Morgan Kaufmann, 1997.