AUTOMATED ANALYSIS AND EVALUATION OF TECHNICAL SURVEYING PROCESSES WITH KNOWLEDGE-BASED SYSTEMS

Andreas EICHHORN[1], Klaus CHMELINA[2]

[1]*Institute of Geodesy and Geophysics, Engineering Geodesy*
Vienna University of Technology, Vienna, Austria
[2]*GEODATA Ziviltechnikerges. mbH, Vienna, Austria*

Abstract: The widely automated analysis and evaluation of measurement processes is a still unsolved task in technical surveying. Complex measurement systems, increasing time pressure and more and more operators with non-academic background require the development of new strategies for a comprehensive and user-friendly presentation and explanation of quality information as part of the process. To solve these tasks, a possible way is the application of knowledge-based systems. The paper describes the development of a knowledge-based system for the automated evaluation of the quality of free stationing processes. One special focus is set on the complex knowledge acquisition and evaluation phase. A software prototype is presented that processes quality relevant information from hybrid sources and automati-cally reports the evaluation and explanation to the user.

1. AUTOMATION OF TECHNICAL SURVEYING PROCESSES - A DILEMMA

When we look at the typical surveying processes taking place in many areas of technical surveying, we see reached a quite high degree of automation in data acquisition. But when we consider the whole process, including project preparation and data processing and evaluation, real life technical surveying processes are still far from being automated. They require human interactions and decisions here and there. So we feel that, if we want to achieve visible further progress in process automation, it is necessary to think of the implementation of other, maybe even more visionary, concepts. These concepts must address the problematic human interac-tion issue. To develop them, we first analyse the current situation and discover two dilemmas.

1.1. The dilemma in the field

Surveying sensors like tachymeters, lasertrackers, optical trackers, laserscanners, GNSS-sensors, digital cameras, etc. become more and more motorized, miniaturized, modularized and computer-controllable. Manufacturers try to pack and integrate several of these sensors into one instrument body or try to realize at least the idea of interchangeable sensor modules to achieve maximum flexibility and versatility.

For specific problems complex automatic surveying systems are developed and installed that consist of and control a number of different geodetic (and other) sensors spread somewhere in the field. These systems or instrumentations can perform a series of data acquisition operations simultaneously and/or subsequently (e.g. tachymeter networks or GNSS-sensor networks for deformation monitoring).

Other surveying systems integrate sensors on static or mobile multi-sensor platforms from which monitoring and mapping tasks can be carried out automatically and even kinematically (e.g. mobile mapping systems, machine guidance systems).

Autonomously operating monitoring and mapping robots, logically and for sure, will be commercially available in future. They are under development already.

But all modern surveying systems today support us only to a limited extent when regarding our real life field work problems. Of course, field work can be performed quicker and more comfortable than ever before and automatic measuring modes improve data acquisition fundamentally. But however, there is a dilemma.

Our surveying systems succeed in producing more and more data automatically, kinematically, continuously, etc. but, apart from these technical superlatives, keep silly and stupid as ever. They neither know what they measure, what would be the best method nor how to react on a change of situation, etc. They have no on-board intelligence at all. They just can dumbly execute predefined mission configurations. In doing so they only solve the most simple part of what is really needed but – what a glory – do this better and quicker than ever. But why don't they tell us our mistakes, what to do next, what to be aware of or how to avoid troubles? Why don't they efficiently assist the many many unskilled and unexperienced field operators that, by far, build the majority in practice?

1.2. The dilemma in the office

From advanced automatic measuring systems we can receive raw and pre-processed data in real time or near real time, in some cases even continuously. But can current data evaluation software check data quality reliably or further evaluate measuring data to meaningful high quality information at same speed, maybe even in real time? Is there data evaluation software existing that can really prepare decisions or suggest actions trustworthily? In practice, real time data interpretation and analysis is either missing at all or done at very low level meaning that, for example, data is checked naively against predefined thresholds what we can find in most alarming systems.

In standard surveying processes, data evaluation is commonly separated from data acquisition, meaning that it is done later in the office after field work. But also there a comprehensive data quality check is missing in the evaluation software packages we use in practice. Data quality check is an annoying and unattractive manual work, especially for large data amounts. As a consequence it is done on a random basis and superficially rather than continuously and thoroughly - if it is done at all. Software components that allow for a competent and reliable automated data check that has in mind the specific requirements of the surveying purpose and the application field do not exist.

Problem-oriented data analysis and interpretation as a final step of data evaluation suffers from the same lack of tools. It is manual work to be done by so called experts that have little assistance. Data interpretation resists automation successfully up to now.

To summarize, we perceive a fundamental lack of methods and tools to overcome these dilemmas. And for some reason and time we think that methods from artificial intelligence have the potential to deliver us above mentioned visionary concepts. In the following we tried out one such method and describe its development and our results achieved.

2. CREATION OF A KNOWLEDGE-BASED SYSTEM FOR FREE STATIONING

2.1. Motivation and major requirements

Geodata ZT GmbH is an Austrian SME that provides technical surveying services worldwide. One main business field is construction surveying and about 50 employees daily work on construction sites, mostly tunnel sites. Following the current state of the art, they in most cases perform free stationing for all their typical surveying tasks such as setting out, network measuring and displacement monitoring. Let us assume each of these 50 surveyors to perform 5 free stationings per day (a quite realistic assumption). Then we obtain 250 free stationing data sets daily that – to avoid problems – should be checked before further processing. Practically, this is done by quickly viewing a calculation protocol that lists adjustment statistics and further figures. But as we have to face the fact that normally only academics have the knowledge to fully understand adjustment statistics and, further, time pressure is usual, data quality check is a real life problem for Geodata.

This specific problem is one example of the above mentioned automation dilemmas. To provide a solution we now present a knowledge-based system that is designed to automate free stationing data quality check on construction sites.

The major requirements are to automatically load and process free stationing data sets, to represent and use geodetic knowledge to interprete data quality and to output results that are understandable even for non-academics.

2.2. Strategy for prototype development

The development of a knowledge-based system prototype for automated free stationing data quality check is realised by a cooperation between Geodata ZT GmbH and the research group Engineering Geodesy, TU Vienna (see Gattringer, 2006 ; Chmelina and Eichhorn, 2008). From the possible implementation forms for a knowledge-based system, it is decided to create a rule-based system (see Jackson, 1999 ; Leondes, 2000) which is particularly suitable for this task. The reason for this is that a high percentage of free stationing knowledge aims for geometrical or environmental conditions which can be best represented by rules.

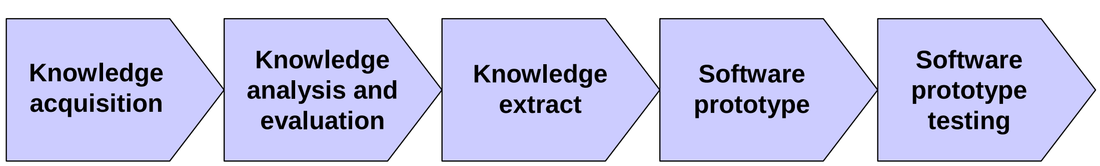

Figure 1 - Phases for the development of the software prototype

Principally, the development process can be divided into five subsequent phases (see Figure 1). The ‘knowledge acquisition phase’ contains the methodical collection of free stationing knowledge from human experts and other sources like literature. The collected knowledge must be analysed and evaluated concerning certain quality aspects like correctness or accuracy, which is performed in the 'knowledge analysis and evaluation phase'. The implementation of the evaluated knowledge into a knowledge-based systems requires the 'creation of the knowledge extract', which in the present work means the knowledge description as a systematic tabular summary ready for being represented by the chosen knowledge-based technique. This is an essential precondition for the 'creation of the software prototype', which follows in the next phase and generates the tool for the practical application in surveying processes. In the last phase, the 'software prototype testing' completes the development. In the following a more detailed description is presented.

3. KNOWLEDGE ENGINEERING: ACQUISITION AND PREPARATION

3.1. Expert interviews and collection of metaknowledge

One appropriate way for the acquisition of free stationing knowledge is the interview with experts who are dealing with this measuring process as practitioners on construction sites (like tunneling) or are teaching the more theoretical background in lectures at university. To address these people special questionnaires are created and distributed at TU Vienna and in several offices (e.g. Geodata). The developed structure of the questionnaires enables on one side to assign the large variety of personal knowledge derived from practical experiences to a manageable amount of main criteria but on the other side also gives space for intuitive answers. An example for more 'specialised' (a) and more 'intuitive' (b) questions concerning the evaluation of free stationing quality is shown in Figure 2.

(a)

Questionnaire for the Free Stationing process

1. Which theoretical and practical criteria must be considered with the choice of the station ?

1.1 Try to evaluate when a criterium is good – middle – or bad fulfilled.

•
•
•

3. Which parameters / quantities / criteria influence the accuracy of the results (station and targets) ?

3.1 Try to sequence the practical and theoretical relevance of the criteria (e.g. important – less important – not important)

•
•
•

(b)

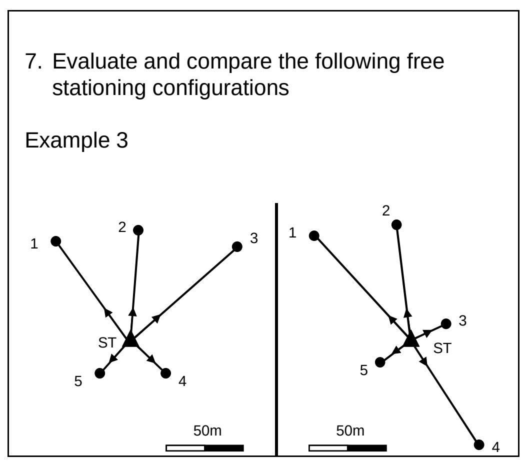

Figure 2 - Examples for 'specialised' (a) and 'intuitive' (b) questions in the questionnaire

During the following evaluation process one big problem for the knowledge engineer is the separation and classification of independent information. This means the identification of redundant statements in subjectively verbalised answers. Another challenge is discovering uncertain or incorrect knowledge. This requires a detailed knowledge analysis which - in this project - is primarily performed by a classification of the given answers by the knowledge engineer himself. The classification process is supported by the personal knowledge of the engineer and an additional literature study for completion of possible knowledge lacks. It can be stated that knowledge acquisition and an accurate analysis are a very challenging and time-consuming (nearly 6 months) part within the workflow for the prototype development.

3.2. Knowledge analysis and evaluation

The knowledge analysis and evaluation contains the examination of the questionnaires regarding content and quality of the collected knowledge. The main goals are to

- eliminate incorrect and incomprehensible answers,
- to extract and evaluate all possible hidden knowledge,
- to aggregate answer classes by the assignment of the extracted knowledge to more generalized generic terms,
- to define rankings for free stationing quality (e.g. the list of important – less important – or not important criteria for station choice (see also Figure 2a)),
- and to develop the preconditions for the final knowledge extract (see Section 3.3).

The extraction of the knowledge is performed by the knowledge engineer for each single question. The quality of the answers is statistically evaluated regarding correctness and accuracy (containing unambigousness and comprehensibility). For each given answer the quality evaluation is realised by comparison with the other related answers from the questionnaires, additional knowledge from the knowledge engineer himself and 'reference knowledge' derived from literature.

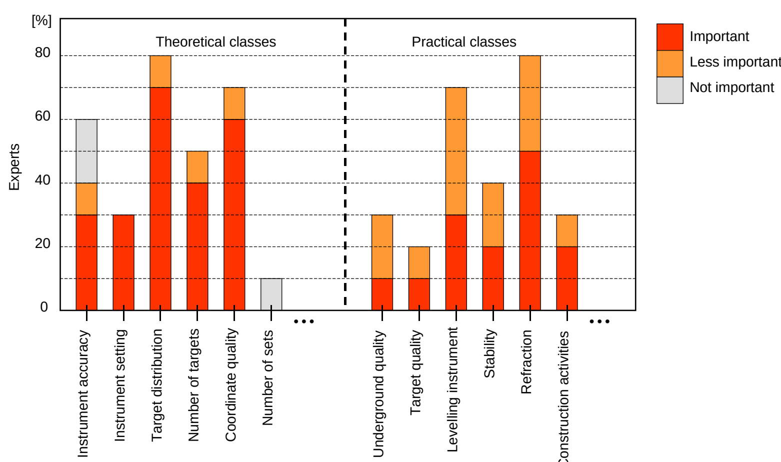

Figure 3 - Empirical determination of answer classes and rankings (cutout from question 3)

In this project the classification of the extracted knowledge to independent answer classes is also performed by the engineer. Additional knowledge derived from literature is used to fill up possible knowledge lacks (see also Section 3.1) and to create a kind of 'pseudo total knowledge'. The aggregation process to answer classes is a very difficult step, because highly influenced by the personal view of the operator. In the present case it is tried to minimise the personal influence by introducing a kind of expert round (consisting of three experts) discussing the decisions of the knowledge engineer.

Some results from knowledge analysis are shown in the histogram in Figure 3. The cutout of the analysed answers from question 3 (see Figure 2a) shows that is it possible to aggregate theoretical answer classes like ‘instrument accuracy’ and ‘geometrical target distribution’ and also more practical classes (dependent on the environmental conditions on the construction site) like ‘underground quality’ and ‘refraction’. In total 25 answer classes are aggegrated for this question. The absolute frequency of the answer classes and the given rankings in question 3.1 (again Figure 2a) also enable to define a superior ranking for the influence of the different classes to free stationing quality. In Figure 3 it is shown that e.g. ‘geometrical target distribution’, ‘coordinate quality’ and ‘refraction’ have an much higher ranking than the ‘number of measured sets’.

3.3. Final knowledge extract as input for the software prototype

One example for a final knowledge extract suitable for implementation in the software prototype is shown in Table 1.

Criterion	Evaluation		
	Good	Sufficient	Bad
Distribution of target points	Symmetrical in all directions	Partly symmetrical	None symmetrical, bad intersections
Number of target points	> 5	4 - 5	< 4
Pre-info coordinate quality	Accuracy info from precise, higher-ranking network	Accuracy info available but no reliable source	No accuracy info available
Time of coord. determination	Within last 10 years	Older (approx. 10-40 years)	Older than 40 years or unknown
⋮	⋮	⋮	⋮
Distance to station	Close targets: 15m < s < 50m Remote targets: 100m < s < 300m	Close targets: s = 10-15m or 50-100m Remote targets: s = 70-100m or 300-600m	Close targets: s < 10m or > 100m Remote targets: s < 70m or > 600m
Visibility	Target is well visible	Only temporarily visible	Permanently non visible
Contrast	Clear contrast to background	Weak contrast to background	Not clear observable / non visible
Accessibility	Easy accessible	Accessible	Poor / non accessible
⋮	⋮	⋮	⋮

Table 1 - Knowledge extract for target points as rough input for the rule-based system

The example shows a cutout of the systematic tabular summary of all collected knowledge regarding the influence of target points to the quality of the free stationing process. The table shows a first classification of possible scenarios (criteria) in three evaluation classes (marks): ‘good’, ‘sufficient’ and ‘bad’. In the prototype itself (see Sections 4 and 5), the top class ‘good’ is further differentiated in ‘very good’ and ‘good’. In addition, the weights of the different criteria for the final evaluation of the free stationing quality are derived from the empirically determined rankings shown in Figure 3.

Principally, the knowledge extract represented in Table 1 can now be implemented in a rule-based system. As an example, possible rules derived from the second criterion (number of target points) are:

1. If the number of target points is > 5, then the criterion gets a 'good'.
2. If the number of target points is 4-5, then the criterion gets a 'sufficient'.
3. Else the number of target points is < 4, and the criterion gets a 'bad'.

4. SOFTWARE PROTOTYPE

4.1. Development environment CLIPS

The software prototype of the rule-based system for automated evaluation of free stationing quality is developed by the representation of the knowledge extract with the expert system development environment CLIPS (= C Language Integrated Production System ; see CLIPS, 2008). CLIPS was developed in 1985 at the Johnson Space Center (NASA) and is now free available for public use. It is considered suitable as it has already been successfully used in related applications (e.g. Chmelina, 2003). This means that already existing parts of CLIPS-code can also be applied to the current problem. It is a very flexible tool and enables the representation of a wide spektrum of knowledge using several knowledge representation techniques like facts and rules, objects and procedures.

4.2. Creation of the prototype

Basic elements of the CLIPS prototype are the fact-list (contains all collected facts like sensor accuracy, determined thresholds, etc.), the rule-base (contains all collected rules, e.g. derived from Table 1) and the forward-chaining inference mechanism (organises the execution of the rules, e.g. the required sequence). The principle of the inference mechanism is shown in Figure 4.

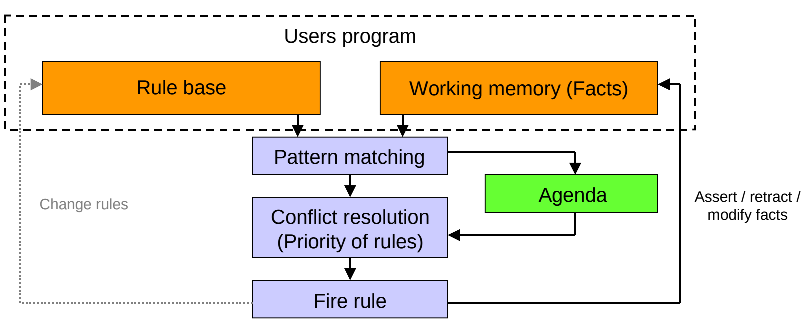

Figure 4 - Inference mechanism implemented in the prototype (Accord. to WXCLIPS, 2008)

At the beginning of the inference process the facts and rules are uploaded into the 'working memory'. The preconditions for a rule execution (so called 'left hand side' of the rule) are defined as a combination of facts and /or objects (defining a so called 'pattern'), which must be present in the current working memory. The compliance with the preconditions is checked by 'pattern matching'. Prior to their execution all pattern matched rules are stored in the 'agenda' and 'sorted' according to their priority. This sorting process is called 'conflict resolution', for which CLIPS offers seven different strategies (depth, breadth, LEX, MEA, complexity, simplicity, and random). In our case the depth strategy is used simply meaning that the 'agenda' is dynamically reordered from highest to lowest priority every time a new rule enters the 'agenda'.

The execution of rules leads to actions, which are defined as 'right hand side' of the rule. The results from executed rules can be e.g. 'asserted, retracted or modified facts', which leads to a dynamic update of the working memory. A 'change of rules' is principally also possible, but is not implemented in the present prototype.

In detail, the prototype contains the following forms of extracted facts and rules:

- Primary facts: these facts represent free stationing input data, e.g. station and target coordinates, standard deviations, type of measurements performed to the targets, environmental conditions like underground, weather, etc.
- Secondary facts: these facts represent heuristic knowledge of boundary conditions like limits, thresholds, ranges and also the rankings and marks as extracted in the knowledge acquisition phase.
- Heuristic rules: the rules evaluate the primary facts against the secondary facts and create specific (related to certain criteria like the distribution of remote targets, close targets or the underground quality) and aggregated ratings (like the total distribution of close and remote targets) of the quality of the current free stationing situation. At the end one overall quality evaluation is created.

The prototype creates an output file (in the form of a detailed report), which presents the evaluation results to the user. If necessary, the report also presents an explanation for the conclusion made.

5. TESTING THE PROTOTYPE

In the current phase of the project, the prototype testing is realised with simulated data only. The integration in a real construction process will be the task of future work. In addition, not all extracted knowledge is implemented in the prototype. At the moment, it is restricted to the representation of main quality influences like geometry and quality of the target points.

The simulation is performed by loading the relevant information of a free stationing scenario in the knowledge-based system. The scenario is derived from an example given in the questionnaire. The extracted expert evaluations from the scenario are used for the verification of the generated report.

The geometrical aspects of the free stationing scenario are shown in Figure 5a. The measuring configuration consists of three remote targets (1 to 3) with distances of approximately 200m to the station (ST) and two close targets (4 and 5) with distances between 20 and 30m. The remote targets are well distributed around the station, which is located in their center. The close targets show a more one-sided configuration. They are situated comparatively close to each other. Further conditions like the standard deviations of the target coordinates and their visibility are also available and fed into the system.

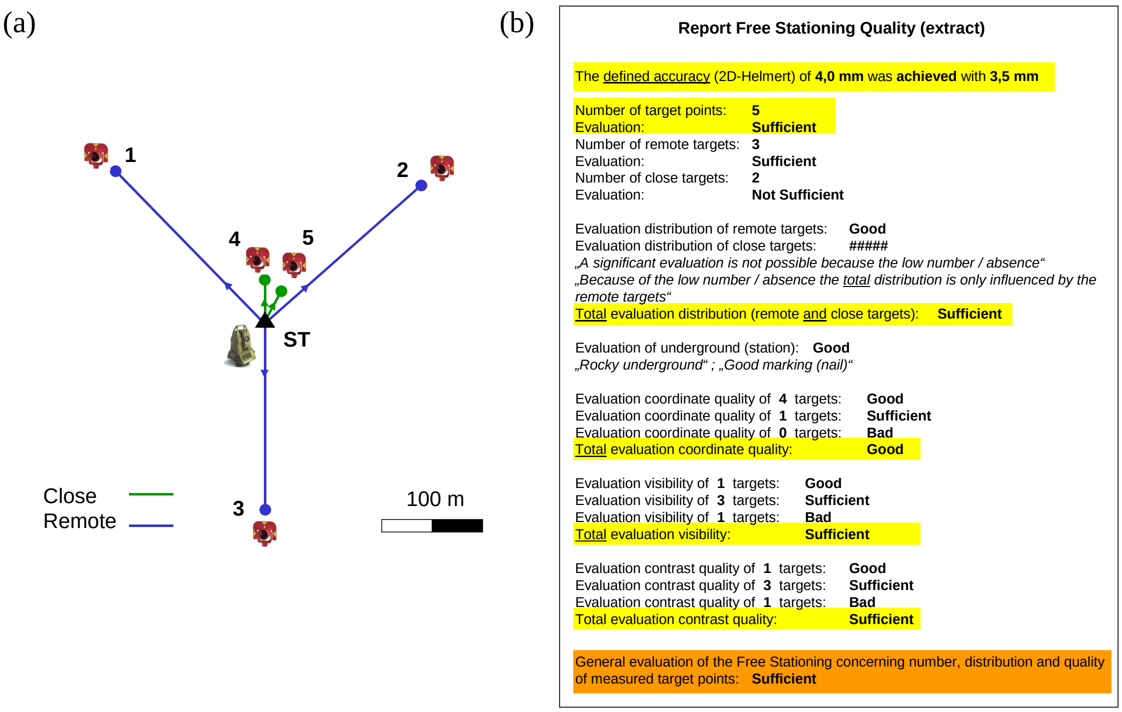

Figure 5 - (a) Simulation of a free stationing scenario and (b) extract of generated report

Based on the available information, the prototype generates a report, which is presented in Figure 5b. It begins with the numerical indication of the achieved accuracy (2D-Helmert error from the adjustment of the free stationing). The achieved accuracy (**3,5mm**) is very close to the pre-defined admissible boundary (**4,0mm**). Possible explanations for this effect are given in the further progress of the report.

The total number of target points (**5**) is evaluated only with **Sufficient**. Looking to the distribution of the targets, only the three remote targets contribute to the evaluation. They are well distributed (**Good**) but only a small number, so in total the system again gives only a **Sufficient**. Station underground and target coordinate quality are evaluated with **Good**, but the visibility and contrast of the targets are in total only **Sufficient**.

The general evaluation of the knowledge-based system (of course strongly influenced from the results for number and distribution of the target points) is a **Sufficient**. This shows that the configuration and environmental conditions of this free stationing example are only moderate and the obtained accuracy can be expected.

These results are representing well the intuitive evaluations given by the experts in the knowledge acquisition phase. The testing of further scenarios obtains similar good results (see Gattringer, 2006). It can be stated that the system is on a good way for a real automated evaluation of free stationing quality.

6. CONCLUSIONS AND OUTLOOK

For routinely and frequently performed technical surveying processes (e.g. free stationing) we see a high potential for the successful use of methods of Artificial Intelligence and knowledge-based systems. Especially for the automation of not-too-sophisticated data analysis and interpretation tasks (e.g. data quality check) they already offer the possibilities to store and represent geodetic knowledge (e.g. as facts and rules) and to draw conclusions (e.g. by forward-chaining inferencing) efficiently.

The prototype described in this paper, indeed, seems somehow primitive and the outcome poor when comparing all the development effort made with the final problem solving competence achieved. But there is possibility for technical improvements in many ways and we have a lot of ideas all motivating us to further research the topic.

We want to close with a spunky vision. In this vision geodetic systems have become capable to perform data analysis and interpretation by applying advanced methods of Artificial Intelligence. They support us geodesists routinely by playing the smart and friendly role of competent tutors and advisors. They even talk to us and we trust their voices as we already do today when hearing 'now turn right' from our car navigation system. So easily we can imagine a future total station telling us warmly in the field: 'Sorry, your free stationing has 87 quality problems! Do you want me to tell you details or a suggestion of how to proceed?'. We don't have to fear such an instrument as long as it reacts correctly when we answer: 'Shut up'.

REFERENCES

Chmelina, K. (2003). *Wissensbasierte Analyse von Verschiebungsmessungen im Tunnelbau*. Geowissenschaftliche Mitteilungen, TU Vienna, Institute for Geodesy and Geophysics, No. 62. Vienna.

Chmelina, K. and Eichhorn, A. (2008). Applications of Knowledge-Based Systems in Technical Surveying. *Journal of Applied Geodesy*. Berlin: de Gruyter, in print.

CLIPS (2008). *CLIPS – A Tool for Building Expert Systems*. URL: http://clipsrules.Source-forge.net/, last access 02/08.

Gattringer, B. (2006). *Entwicklung eines Wissensbasierten Systems zur Qualitätskontrolle tachymetrischer Auswertungen*. Diploma thesis, Institute of Geodesy and Geophysics, Engineering Geodesy, TU Vienna. Vienna: not published.

Jackson, P. (1999). *Introduction to Expert Systems*. 3rd ed. Addison Wesley.

Leondes, C.T. (ed.) (2000). *Knowledge-Based Systems*. Vol. 4. Academic Press.

WXCLIPS (2008). *Hompage of wxCLIPS*. URL: http://www.anthemion.co.uk/wxclips/, last access 02/08

Corresponding author contacts

Andreas EICHHORN
andreas.eichhorn@tuwien.ac.at
Institute of Geodesy and Geophysics, Engineering Geodesy
Vienna University of Technology
Austria