Wolfgang Schwegmann: An Embedded Expert System for the Automation of VLBI Data Analysis: Concept, Implementation and Results, IVS 2004 General Meeting Proceedings, p.287–291
http://ivscc.gsfc.nasa.gov/publications/gm2004/schwegmann2

An Embedded Expert System for the Automation of VLBI Data Analysis: Concept, Implementation and Results

Wolfgang Schwegmann

Bundesamt für Kartographie und Geodäsie (BKG)

e-mail: wolfgang.schwegmann@bkg.bund.de

Abstract

The delay between the observation and the availability of results is an important disadvantage of geodetic VLBI compared to other geodetic space techniques. One reason is the complicated and highly interactive analysis procedure, requiring well-founded expertise. While advances in e-VLBI technology promise to considerably cut the latency, automatic VLBI analysis tools have still to be developed to obtain VLBI results in near real-time. In this paper the concept, implementation and first results of an embedded expert system for the automation of the entire VLBI analysis procedure are presented.

1. Introduction

Geodetic VLBI data analysis is a very complex process and needs a lot of manual interactions. It is very time consuming because most tasks require a comprehensive knowledge of the whole procedure of the data analysis and the analyst needs a lot of experience and knowledge for solving the complex problems within the data analysis. Thus, an automation of the whole VLBI data analysis procedure would be very useful. It allows one to close the gap between the time of observation and the availability of results, i.e. a quicker turnaround can be achieved. The costs of operations will decrease, because the analysts will be relieved from their routine burdens. Thus, they can spent more time to improve the modelling or to investigate instrumental errors. The management of the data analysis will be much easier, because it is based on an explicitly modelled standard procedure that is performed by automatic programs. Last but not least, the results of the data analysis are much more consistent and reliable, because the same analysis strategy is going to be used and all steps of the analysis procedure are perfomed in the same way.

To achieve the automation, knowledge-based methods from the field of research of artificial intelligence are used. An *Intelligent Assistant for Data Analysis in VLBI* (IADA) is developed as an embedded expert system. Expert systems are software systems that are designed to reproduce the special knowledge and the expert's capability of reasoning within a narrow application area. By applying the explicitly modelled expert's knowledge such systems are able to solve complex problems. The term "embedded" is of particular interest, because embedding expert systems in the existing data processing environment is critical for the success of such a system. Embedding IADA in the Mark IV data analysis software CALC/SOLVE by building a powerful interface guarantees the automation of the whole VLBI analysis procedure.

Two consecutive steps are necessary to develop an expert system for the automation of the VLBI data analysis procedure: A so-called *Structured Model*, describing the tasks to be solved during the data analysis and the knowledge necessary to do that, has to be built. This model has to be implemented, i.e. the knowledge has to be formally represented within the expert system and an interface has to be developed to use it in conjunction with the existing analysis software.

2. Concept of IADA

The concept of IADA is based on the so-called *Structured Model of the VLBI Data Analysis Procedure*. This model describes the course of the VLBI data analysis and represents the knowledge necessary within the single steps of the analysis. For a successful automation it is important to model all the required knowledge in a structured and formal way. Thereto, the processing strategy of the data analysis has to be characterized by several steps that are performed one by one. Each step represents a major task within the data analysis and is subdivided into several substeps to describe the current task in more detail (cf. figure 1). Each substep has to be characterized by the following items:

- *Tasks*: The tasks of the current substep of the analysis procedure.
- *Parameterization*: Every task demands a special parameterization that depends on the results of previous steps of the analysis.
- *Models*: The models to be used.
- *Approach*: The approach to perform the tasks.
- *Evaluation criteria* (quality control): to evaluate the results of the current step of the data analysis.
- *Error sources*: To list all possible faults that can cause unsatisfactory results.
- *Trouble-shooting*: Methods to correct for faults; they are assigned to an error source.

Moreover, the *Structured Model* has to define the generic tasks to be solved during the automatic data analysis. Some tasks can be processed by the existing software while other tasks have to be handled by the expert system. The process flow between IADA and SOLVE and the several tasks that are performed by IADA are shown in figure 2.

The user interface for the usage of IADA is realized in SOLVE (cf. section 3). Thus, the analyst is able to call IADA for an automatic data analysis from within his usual software environment. The user interface allows the analyst to select a specific step and substep of the analysis procedure. For this substep IADA can be called in order to perform it automatically, to check the current parameterization, to evaluate the results of the analysis, to recover possible problems and to find solutions to solve the problems. From the user interface the analyst invokes the interface that handles the data and information flow between SOLVE and IADA.

For step 1, 2 and 3 of figure 1 the interface calls external programs (GETDB, XCALC, XLOG). These programs are independent modules within the Mark IV data analysis system and can be run in automatic or in interactive mode. They can be used independently from IADA as well as in conjunction with the expert system. In the latter case IADA is responsible for the control of the process flow, i.e. to call the programs when necessary and to check the results. GETDB, XCALC and XLOG are described in more detail in [2].

For the steps 4 to 8 (cf. figure 1) IADA is going to be used and according to figure 2 the following tasks are passed through: During the *Preparation* IADA specifies all information necessary to perform the current substep. The respective data are collected by the interface and included into the knowledge base of the expert system during the *Initialization*. Based on this information conclusions are drawn that are important for the next steps of the analysis (*Deduction*). Then the parameterization is being reviewed during the *Control*. Corrections to remove possible errors in the

Step	Substep
1 DownloadDB	1.1 X-Band | S-Band | X+S Band
2 Calc	2.1 X-Band | S-Band | X+S Band
3 LogFiles	3.1 Xlog
4 DataLoading	4.1 X-Band | S-Band | X+S Band
5 InitialSolution	5.1 InitialSettings 5.2 AprioriClock 5.3 AmbiguitySolution 5.4 EvaluateResults
6 IntermediarySolution	6.1 InitialSettings 6.2 OutlierElimination 6.3 EvaluateResults
7 FinalSolution	7.1 InitialSettings 7.2 OutlierElimination 7.3 ExamineSolution 7.4 FinalOutlierElimination 7.5 CheckCableCal 7.6 FinalParameterization 7.7 EvaluateResults
8 DatabaseUpdate	8.1 WebDoku 8.2 DBUpdate

Figure 1. Steps and substeps of the VLBI analysis procedure.

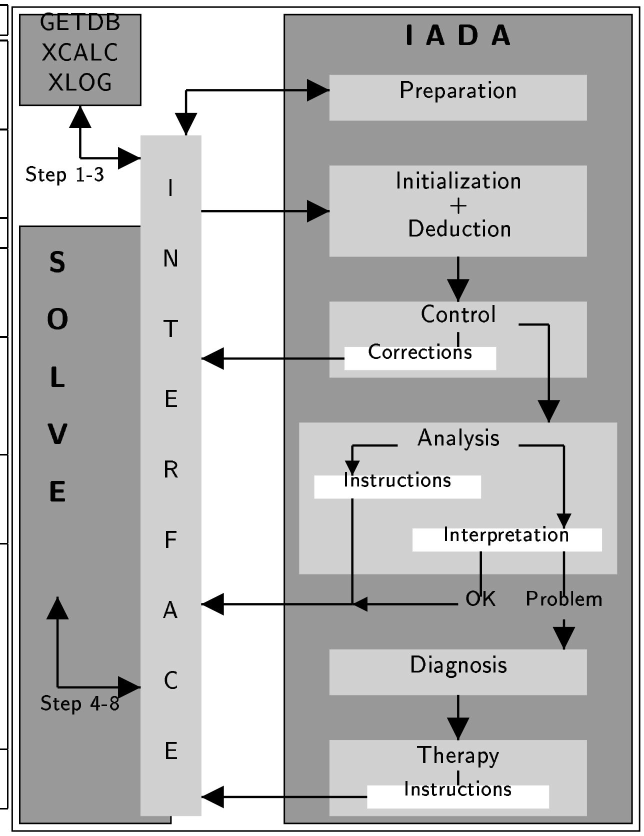


Figure 2. Process flow within IADA and SOLVE for each single substep.

parameterization are passed back to SOLVE and the analysis is continued with the *Initialization*. The actual tasks of the substep are performed during the *Analysis*, i.e. the instructions to be realized in SOLVE are going to be determined with respect to the current analysis situation or the current situation is going to be evaluated. If the result of this interpretation does not recover any problems the analysis is going to be continued with the next step or substep. Otherwise the *Diagnosis* is going to be started in order to find possible error sources for the actual problem. If symptoms of a well known problem can be identified methods to solve the problem are derived and the corresponding instructions are passed back to SOLVE (*Therapy*). Afterwards the current substep is going to be analyzed in IADA again.

3. Implementation of IADA

The Implementation of IADA requires two steps: During the so-called *Knowledge Representation* the *Structured Model of the VLBI Data Analysis Procedure* has to be transformed and stored in the knowledge base of the expert system. The problem of how to represent the knowledge of the VLBI data analysis procedure has already been described in [1]. In addition, a user interface has to be developed as well as an interface to exchange data and information between the existing software and the expert system. Both interfaces have been implemented within the existing analysis software, i.e. within the Mark IV data analysis system. The interface between SOLVE and IADA allows use of the existing software in conjunction with the expert system by applying the appropriate program for the different tasks of the VLBI data analysis. In principle IADA is responsible for the control of the data analysis process, for the evaluation of analysis results and for the detection and solution of problems (cf. figure 2). On the other hand the number crunching, the setting of parameters, the plotting of residuals, etc. is done further on in SOLVE.

A schematic representation of the embedded expert system for the automation of the VLBI data analysis is shown in figure 3. The main components are the existing SOLVE software package and the expert system (IADA). Some modules within SOLVE have been modified in order to allow an automatic data processing (dashed boxes). The boxes with a grey background are new automatic modules: GAMBSOL allows resolution of group delay ambiguities, CBREAK finds clock breaks, OUTLIER is an algorithm for outlier detection, IADAO is the user-interface for IADA and INTF is the interface to exchange data, information and commands between SOLVE and IADA.

IADAO allows the analyst to specify all information necessary for the automatic data analysis, e.g. database name, current step and substep of the analysis, application of IADA in interactive or automatic mode, and so on (cf. figure 4). This includes several configuration settings to control the automatic analysis (cf. figure 5). It can be specified, for example, whether to generate a comprehensive report in the form of web pages including several plots and protocol files. Beneath IADA the interface INTF is the most important component of the system. It consists of two main modules: `Solve->Iada` extracts on request of IADA all necessary information from the *work files* for further processing in IADA; `Iada->Solve` reads the results and instructions from IADA, makes appropriate changes in the *work files* and calls the appropriate programs to perform the instructions given by IADA.

4. Results

The embedded expert system has been applied to analyze 50 INTENSIVE sessions and 30 24 h experiments. All INTENSIVE sessions could be processed automatically and the results as well as an example of comprehensive web report are given in [2]. In the case of the 24 h experiments 70% of all experiments could be processed automatically. Most of the problems of the remaining 30% arose due to problems in solving the group delay ambiguities. However, in some cases the data could not be processed correctly because of a lack of knowledge in the expert system. This proves that expert systems can never replace a human expert because it is not possible to completely implement the expert's knowledge and problem solving strategies in an expert system. Nevertheless, IADA has shown that an embedded expert system can be used to automate the entire VLBI data analysis procedure in about 70% of the investigated 24 h experiments.

References

[1] Schwegmann, W., Automation of the VLBI Data Analysis, In: Measurement of Vertical Crustal Motion in Europe, TMR Network FMRX-CT96-0071 Scientific Report 1996-2001, J. Campbell, A. Nothnagel and R. Haas (eds.), ISBN 92-894-0763-8, 140–145, 2002.

[2] Schwegmann, W., G. Engelhardt, V. Thorandt, Towards an Operational Automatic VLBI Data Analysis Tool for INTENSIVE sessions, In: *2004 General Meeting Proceedings*, edited by N.R. Vandenberg and K.D. Baver, NASA/CP-2004-212255, this volume, 2004.

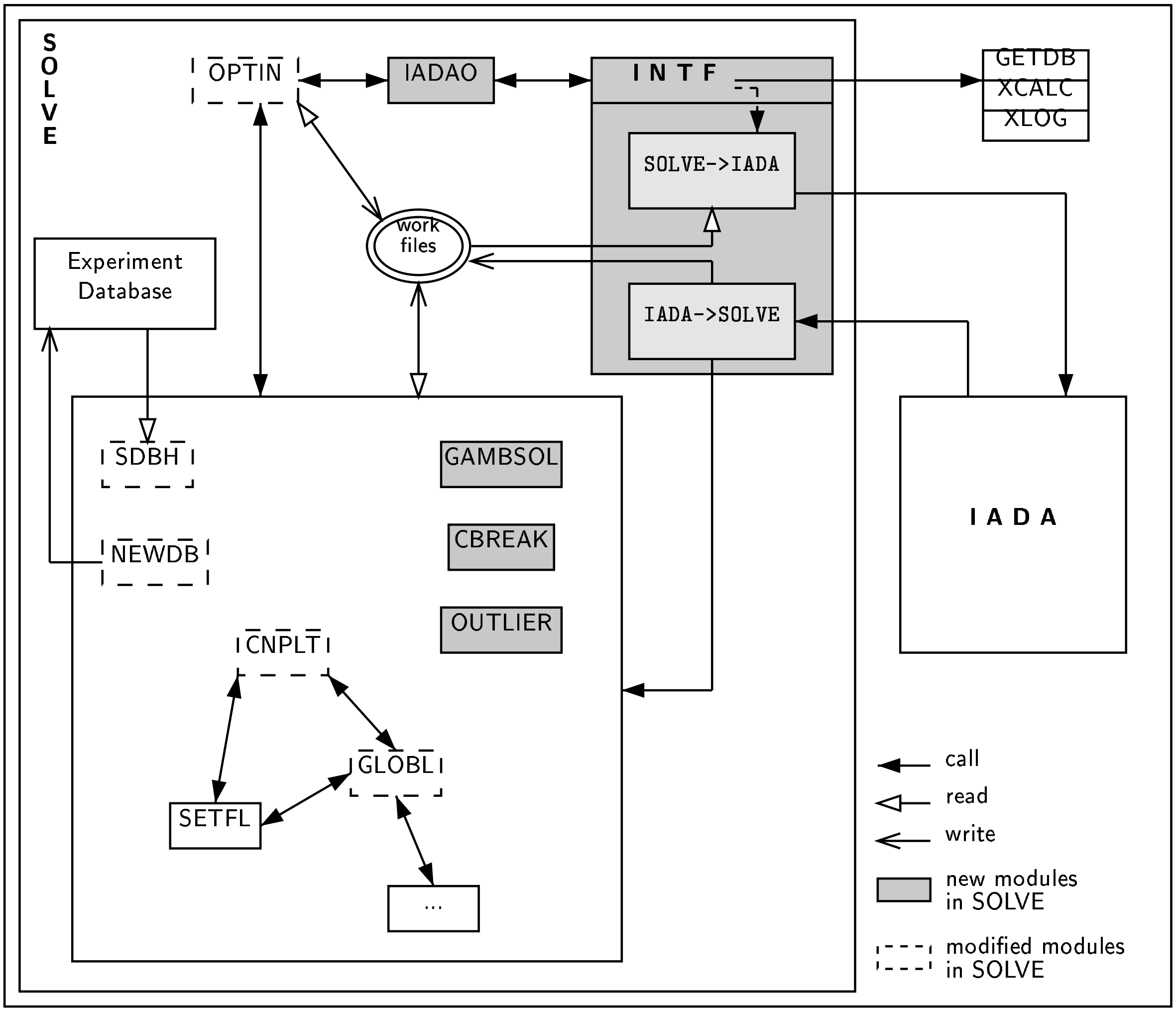
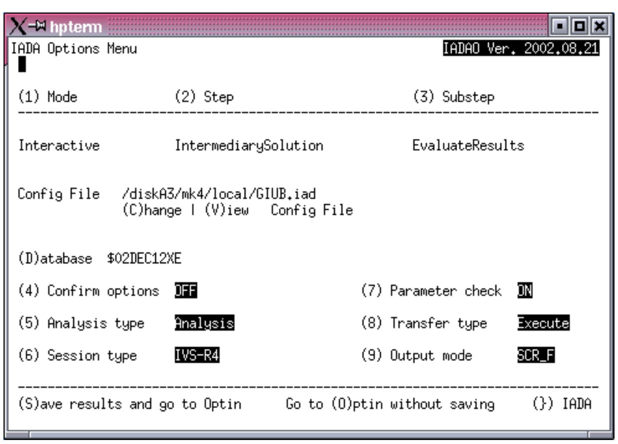


Figure 3. Schematic diagram of the embedded expert system for the automation of the VLBI data analysis.



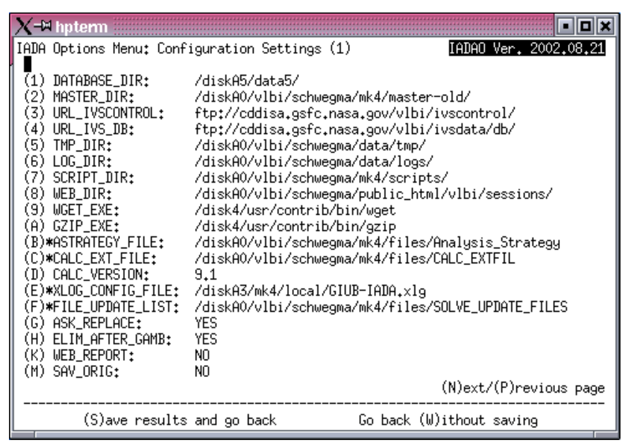


Figure 4. IadaO: The user-interface of Iada.

Figure 5. IadaO: Configuration settings.