Determinant factors of using data and knowledge in bridge engineering

Janusz Szelka ^{1,*}, Zbigniew Wrona ²

¹ University of Zielona Góra, Poland

² College of Management “Edukacja” in Wrocław, Poland

Abstract. Complexity and heterogeneity of information and decision processes in bridge construction necessitates the use of heterogeneous and complex information resources. However, for data collection and processing in this area, transactional database systems are mainly employed. They allow only the processing of detailed precise data as well as static ones. Other categories of data used in this area, including, for instance, fuzzy, temporal or active data, make it necessary to use additional specialized database systems. A similar remark can be made in relation to the various categories of acquired and processed knowledge in bridge construction projects (e.g. complete or incomplete knowledge in the reasoning process). Numerous independent and autonomous software applications used for this purpose do not offer comprehensive support for complex and multi-stage problems. It appears then that rational IT solutions in this area should be sought within the integrated systems of intelligent databases, supplemented with data fuzzing mechanisms and active data support, and, additionally, enabling the exchange of data with heuristic reasoning systems, in particular with expert systems that use the rule-based organization of knowledge, cooperating with case-based reasoning systems.

1 Introduction

Computer-aided support for information and decision projects in bridge construction requires the employment of increasingly extensive information resources. The projects in this area, at every stage of their execution (design, organization of construction, monitoring or prediction of the technical condition of the bridge), can be classified as tasks of high complexity. This complexity translates into both the amount of information resources that need to be collected and processed, as well as their significant diversity. The problem is not in the absence of software mechanisms or unfamiliarity with the methods (techniques) that should be employed to collect and process information and knowledge, but in such a large heterogeneity of mechanisms (methods, tools) which need to be applied that their complementary use becomes a major challenge.

In a situation where some of the information resources available are in the form of precise data, whereas others are expressed by linguistic variables, are characterized by vagueness (fuzziness), and sometimes are also incomplete, a question about rational methods and tools to be used for their collection and processing becomes highly significant. In consequence, the then emerging necessity of simultaneously making use of several different IT applications frequently means that one has to limit oneself to supporting only selected stages of the entire project or is made to significantly simplify the description of the

problem situation, which may considerably affect the quality (effectiveness) of the final effect.

The essence of the problem under consideration is presented in Fig. 1.

Fig. 1. Essence of the problem.
 Source: Own elaboration

* Corresponding author: j.szelka@interia.pl

2 The specificity of information and decision resources in bridge construction projects

A solution to a problem situation in bridge construction is provided as a result of the analysis of various factors, such as: needs regarding the use of the facility, stocks of local materials, or the size of potentials of various categories (human, hardware) that can be made use of in the construction/reconstruction of the facility. Not only should the influence of each factor on the area of acceptable solutions be taken into account, but also the interrelationship between individual factors. For instance, when seeking rational solutions in the case of designing a temporary bridge in a crisis situation (e.g. after a flood event), one should use a typology of such bridges as well as relevant solutions related to construction, technology and organization. A constant will only be the required bridge load capacity, while the variables in time will be the parameters of the water obstacle, and thus the different possibilities of detour. Then, along with the change of those parameters, the following will also undergo a change: the number and type of supports, and the number and breadth of bridge spans. A similar situation will be connected with the availability of the assembly yard (when using foldable structures), which, in turn, affects the organization of the assembly of the span structure and bridge supports. It will also require the selection of special construction equipment (e.g. on a tracked chassis) due to the difficulties related to accessing the place where bridge works are being carried out. Each of these factors affects the adoption of the proper (preferably rational) structural layout of the facility as well as its technology and organization of construction.

The complexity and heterogeneity of information and decision processes in this area necessitates the use of heterogeneous and complex information resources. This heterogeneity concerns the data types used, but, above all, it is visible in the diversification of the categories of data and knowledge resources. A data type specifies the range of values (the domain of the attribute) that can be used to identify the state of a variable. Information and decision projects in bridge construction make use of all the data types that are currently recognized and used programmatically (in different programming languages), in particular: textual data, numerical data, dates and times, and, more and more frequently, multimedia data, including images, sounds and video sequences. In the context of the specificity of the tasks under discussion, a little more attention should be given to the term: data category. This is not an unambiguous term and can refer to both data type (detailed, aggregated, precise, fuzzy) as well as data specificity, customarily expressed through additional tools for their description (active data, temporal data). The most widely used data category in bridge construction is precise detailed data of a static nature. Most numerical data and some textual data related to, for instance, water obstacle parameters or assembly type (depth = 5m, bottom soil type = 'stony', type of intermediate support = 'P264') can be included in this category. However, it is enough to refer

the description of the problem situation to the crisis situation (flood conditions), and it will transpire that in the same example (specification of water obstacle parameters) the data will not be static (e.g. the variable will be the width of the water obstacle), and, in addition, in analytical and decision processes, the "history of variability" of this parameter will become significant. In this situation, the data representing the width of the water obstacle can therefore be included in the temporal data category. For this data category, in addition to the attribute value itself, timestamps should be used, specifying both the time of measurement and the time when it stops being current. It is worth noting, however, that a significant number of attributes depicting the problem situation can be represented with the so-called linguistic variables, for which certain variable categories are specified (e.g. bottom soil type: 'stony', 'cohesive/cohesive and very', etc.). In the majority of situations, typical divalent logic (bottom soil type: stony: YES/NO) is sufficient to express the values of variables, but in some cases clear-cut boundaries between several categories of linguistic values, assigned to a variable, are too far-reaching a simplification. In such cases, fuzzy data (employing multivalued logic) should be used, with which the actual situation can be reflected more accurately [1]. This means that by using appropriate fuzzing functions, bottom soil type can, for instance, be classified as stony with a probability of 0.7 and as coarse with a probability of 0.3.

A significant variability of the situation one may deal with in bridge construction projects, translating into the variability of data values, may require an automated (not human-initiated) data update in real time. A bridge monitoring system may be an example of such mechanism. One of the significant risk factors is the fatigue factor. If a temporal variable is used to store the value read from a sensor, it will be enough in the so-called active rule, associated with this variable (this rule is referred to as the trigger), to write, for instance, that if a triple reading from the sensor in the bridge monitoring system exceeds the limit value, the alarm for the system should be activated. The data provided with mechanisms for their automated updating is referred to as active data.

The specificity of the selected categories of information used in bridge construction is presented in Fig. 2.

Information resources used in information and decision processes in bridge construction are not limited only to the categories of primary or processed data (information). A significant part of these processes require resources in the form of knowledge, understood as a set of reliable and interrelated information [2]. Knowledge can be represented (structured) in various forms (including: decision rules, case description, decision trees, or neural networks). The most widely used form of knowledge representation is the rule-based representation, reflected in the form of decision rules.

Fig. 2. Information categories (in the notation: entity-attribute-value) characteristic for bridge construction.
 Source: Own elaboration

The elementary portions of rule-based knowledge are recorded with the following structure: **if...then...**. For instance, one of the rules defining the way of assembling a temporary bridge may take the following form: **IF** assembly_compartments **AND** (number_cranes = "none" **OR** number_cranes = "one") **THEN** assembly = "no assembly" [3].

The effectiveness of rule-based reasoning is significantly limited by the requirement that the knowledge accumulated in a rule base and a fact base must be complete. In situations where knowledge is incomplete (and this may often be the case, for instance, in crisis situations), other methods and tools of heuristic reasoning should be used to work out conclusions. Case-Based Reasoning (CBR) systems are an effective and relatively easily scalable tool. A case base in such a system collects information on problem situations that have occurred (e.g. regarding the selection of span structures of temporary bridges). Reasoning, consisting in the necessity of solving the current problem situation, is performed by searching the case base for the situation that is the most similar to the one under consideration, and if the similarity is sufficient, the conclusion is adapted to the current situation, based on the case that is similar enough. Information resources include here a formalized way of the textual description of cases, and the mechanism of case comparison uses selected similarity functions (e.g. the weighted average of individual attributes depicting the compared cases).

Each of the above-mentioned categories of information resources, including portions of data or knowledge, appears to be significant for the quality of information and decision processes, but it should also be noted that the possibility of using each of them requires the application of specific and dedicated IT tools that ensure their efficient collection and processing.

3 Determinant factors of IT support for information and decision projects in bridge construction

For IT support of information and decision projects in bridge construction, most frequently classic database systems (using the so-called transactional databases) are most frequently used.

Transactional databases can be used to memorize values of different types: alphanumeric entries, graphics, sound material or video recording. Their implementations may include, for instance, libraries of technical and operational parameters of bridges, spans, etc., or directories of crossing equipment. This category may also incorporate data sets that include bridge and road structures along with their parameters, or the parameters of temporary crossings assembly.

Transactional databases have significant limitations both in terms of the form of data collected and the processes of their processing. The principal limitations include the following:

- detailedness of stored data;
- explicitness (sharpness) of attribute values;
- operations on static attribute values;
- use of the HADP (Human Active Database Passive) model.

These properties do not make them the right tool for processing data other than detailed, explicit and static. They also do not allow the collection and processing of data, the description of which would be supplemented with the time attribute (temporal data sequence). Despite these limitations, transactional databases are currently the most frequently used tools for IT support of information and analytical projects in bridge construction. Their widespread use is, however, mainly due to the high availability and popularity of relational database technologies, not their universality. Limiting oneself only to transactional databases entails having to give up all the categories of information that are not consistent with the data model accepted by transactional databases (e.g. fuzzy or temporal data). However, since such data can constitute a crucial part of the problem situation description that an engineer has at his disposal, one should expect that giving them up may have a significant impact on the quality of analytical and decision processes. An alternative solution is an attempt to apply additional, often innovative, database technologies.

An appropriate IT tool for gathering and processing vague data are the so-called fuzzy databases. Such systems ensure not only data fuzzification based on appropriately selected functions (e.g. a trapezoidal one), but also fuzzification of database queries. For instance, for a bridge monitoring system, the use of fuzzy queries enables the performance of such tasks:

1. Search for the elements of the structure for which the monitored parameters are changing *too fast*.
2. Search for structural elements for which the values of the monitored parameters *are approaching* the warning level.

The programming language used to create fuzzy data-

bases and to manage them is SQLf, a variation of the classic SQL (Structured Query Language) database language.

The traditional approach to the processing of data (both precise and fuzzy ones) assumes a static value of each attribute and is employed in most of the recording and analytical systems. In numerous cases in bridge construction, it is necessary to take account of the dynamics of the structure's parameters (behaviours), which necessitates the assignment of timestamps to data and the use of these stamps in data processing processes. Such challenges are addressed by temporal databases. A temporal database is a variation of a relational database, in which each entry has a timestamp, specifying the time in which its value remains true (the time of activating the data value and the time of its expiration). It also has built-in mechanisms that allow one to manipulate temporal data (retrieve data histories). The possibility of the temporary variation of data flowing into the system appears to be particularly significant for bridge construction systems management, in crisis situations. For instance, selection of a foldable bridge structure (under flood conditions) will be influenced by numerous factors describing the dynamically changing situation, e.g. in terms of water obstacle parameters, bank conditions, road passableness, etc. An ability to store a sequence of temporal attribute values, offered by temporal databases, is only one of the desired factors here. In addition, timestamps associated with each entry into such a database allow for the automation of many activities that in other cases have to be initiated by man. An example may include timestamps describing the expiration date of the parameter that determines the depth of a water obstacle. After the expiry of the validity of this attribute value, the database can automatically generate a query for its next value, ensuring the continuity of readings at a specified time interval. The sequences of registered entries (successive attribute values) make up the history of the structure (system) description in time and can be used to perform various analytical and decision tasks (e.g. prognostic analysis, optimization decisions related to the scheduling of temporary bridge construction tasks). The programming language that enables the creation of temporal databases is a variation of the SQL language, called TQL (Temporal Query Language).

The model of operation as indicated above, in which the activity associated with initiating data operations is transferred from man to database, is called HPDA (Human Passive Database Active) and requires the use of special solutions appropriate to the so-called active databases.

An active database system is characterized by the ability to change to the active state, even when no requests are explicitly directed to it. Changing the state of the database (from passive to active) may occur as a result of a specific external event, but also:

- completion of a specific set of transactions sent to the DBMS;
- passage of a given quantum of time;
- combination of the above cases.

An automated change of parameter value, e.g. consisting in the assignment of the variable: Exceeded level of stress value 'YES', as a result of the analysis of the value

of another variable (Stress level) coming from sensors, is made on the basis of active decision rules, the so-called triggers, based on the ECA (Event-Condition-Action) mechanism. A specific event in the system (retrieving a parameter value from a sensor) triggers an automatic check of the condition associated with this parameter, stored in an active rule, and then initiates the action (conditional upon the result of the check).

The issues in which this approach appears essential include the majority of tasks related to the monitoring of bridge construction systems (traffic analysis on bridges, diagnosis and prediction of the condition of bridge structures).

The possibilities of incorporating rule-based structures into data processing processes, offered by active databases, have become the premise for searching for universal solutions in the field of databases, which would enable both advanced data processing and the processing of rule-based knowledge. Such tools are referred to as intelligent databases. In addition to the traditional relational database module (the so-called extensional database), these systems contain a database of decision rules, and the attached software tools allow for the implementation of the reasoning process in which both the rules stored in the intensional database and the data stored in the extensional database can be used [4].

The programming tool that enables the implementation of deductive database systems is the language of logic, e.g. *Prolog* or *Datalog*, used in conjunction with the traditional language of creating and operating databases – *SQL*.

Neither active nor intelligent databases are able to completely replace heuristic reasoning systems. Generating a single action on data and developing rules on the basis of the data possessed is not equivalent to the implementation of a comprehensive reasoning process, which enables working out a conclusion for a specific category of problem situations. In this respect, there is still no alternative to expert systems that use a knowledge base. The process of working out conclusions for a decision problem in expert systems is implemented by the so-called reasoning mechanism, which, in addition to reasoning, allows explaining how to work out conclusions. The advantage of such systems (e.g. in combination with artificial neural network systems) is the method of knowledge representation: E-A-V (Entity-Attribute-Value), which is compatible with the basic method of data recording. This compatibility facilitates communication in the information resources layer between databases and knowledge processing systems. One example of an expert system that aids reasoning in bridge construction is the system for the selection of supports in the "SEDOP" road bridge [5]. The systems used in practice that perform reasoning based on expert knowledge relatively well deal with the uncertainty of information resources. The most widely used mechanism taking account of the level of the uncertainty of knowledge is the Certainty Factor (CF), which measures an expert's belief about the accuracy of the conclusion derived from given premises.

A much larger problem is encountered in the case of incomplete knowledge. As a result of unknown attribute values, an inability to determine the value of facts (e.g. in

the consultation process) or the lack of rules describing atypical situations, the reasoning process may reach the state of the so-called deadlock, in which the system is not in a position to activate any subsequent rule (the values of variables appearing in the premises of these rules are not known). This, in turn, means that it is impossible to work out a final conclusion.

To solve the problems for which it is not possible to obtain full information, CBR (Case-Based Reasoning) systems can be used [6]. In bridge construction, they can be used, for instance, to select the type of crossing (bridge, ferry, combined), or to select a span structure in flood conditions.

By compiling, based on the above description, lists of applicable IT tools supporting information and decision tasks in bridge construction and allowing the use of various categories of information resources, it is possible to designate the tools that support data collection and processing and the ones that can be used in heuristic reasoning processes (tools for knowledge collection and processing) – Fig. 3. Only few of the described IT solutions can qualify for both groups at the same time [7].

The categories of IT solutions shown in Fig. 3 can be used only as isolated, not cooperating with each other, or cooperating to a very limited extent.

Fig. 3. Categories of IT solutions supporting data and knowledge processing in bridge construction.
 Source: Own elaboration

In addition, their independent of each other use does not support complex tasks in which data and knowledge from different categories are included in the description of the problem situation. Therefore, it appears that when attempting to determine potential directions of the development of data and knowledge processing technologies in bridge construction, it is necessary to consider mainly the possibilities of integrating both the different categories of information resources and the modules that support these resources (Fig. 4).

Fig. 4. The idea of an integrated database system that works with reasoning systems in bridge construction.
 Source: Own elaboration

The fundamental component of such a system should include an integrated database system, containing a database (databases) storing data from a selected area of bridge construction, and an integrated database management system [8]. In addition to the basic operational database management module, it must contain mechanisms for handling active rules and a module for collecting fuzzy data and managing such data. The possibilities of the software integration of database mechanisms with knowledge gathering and processing tools are now very limited. However, it is possible to use heuristic reasoning systems as external systems that will transfer to the user (engineer) the conclusions developed for specified decision problems in the E-A-V format (Entity-Attribute-Value), via a database management system interface. The proposed idea of an integrated database system for bridge construction projects allows the use of various categories of data and knowledge in information and decision processes, but at this stage of information technology development it encounters significant difficulties related to the need to use software mechanisms appropriate for different languages of database handling (SQL, SQLf, TSQL) and the necessity of using software mechanisms for mapping reasoning results (e.g. CBR) to the level of data description. However, given, on the one hand, the needs of IT systems supporting information and decision processes in bridge construction, and, on the other hand, trends in information technologies to integrate data and processes, it can be assumed that integrated database systems cooperating with knowledge base systems will become an increasingly more significant software group in this area.

4. Conclusion

Computer-aiding for complex and heterogeneous analytical and decision projects in bridge construction requires the employment of various information technologies. In particular, a significant role is played by database systems that enable efficient search and processing of information resources. However, currently there are no such database technologies that would ensure efficient processing of both precise and fuzzy data, and at the same time would allow timestamping and automated initiation of data operations. An attempt to develop an integrated database that allows simultaneous storage and processing of detailed (static), fuzzy, temporal and active data, with a separate module for intensional and extensional data, appears to be the only rational solution, meeting the complexity of information and decision processes in bridge construction projects. An additional challenge is posed by the need to ensure cooperation between these systems and heuristic reasoning systems. Although at the current stage of database software development there is no homogenous and integrated language of databases supporting various data categories, given the general trends towards integrating IT systems with tools for their development (including programming languages integration), it should be assumed that IT solutions supporting information and decision processes in bridge construction will be heading in this very direction.

References

1. J. Konior, Przedsięwzięcia i procesy budowlane w kategorii zbiorów rozmytych. In O. Kapliński (Ed.), *Metody i modele badań w inżynierii przedsięwzięć budowlanych*, Warszawa, PAN (2007);
2. J. Szelka, Z. Wrona, Możliwości wykorzystania eksploracyjnej analizy danych w przedsięwzięciach inżynierskich, *Czasopismo Inż. Łąd., Środowiska i Architektury*, **63**, 1/I, pp. 213-220 (2016);
3. J. Szelka, Z. Wrona, Odkrywanie wiedzy z baz danych w obszarze przedsięwzięć budowlanych. In J. Sobczak-Piąska (Ed.), *Wybrane problem budownictwa*, J. Sobczak-Piąska (Ed.), Bydgoszcz, Wyd. Uczelniane Uniwersytetu Technologiczno-Przyrodniczego, pp. 171-178 (2015);
4. K. Jha, S. M. Ramjas, An Introduction To Deductive Database And Its Query Evalution, *International Journal Of Advanced Computer Technology*. V. **3**, No. 3: pp. 33-36 (2014);
5. J. Szelka, Obiektowy zapis wiedzy w systemach eksperckich wspomagających budowę mostów wojskowych. Warszawa, WAT (1999);
6. K. Myszkowski, P. Ociepka, J. Świder, Zastosowanie metody CBR do wspomagania procesu projektowania maszyn. *Modelowanie Inżynierskie*, V. **6**, No. 37, pp. 217-224 (2009).
7. J. Szelka, Z. Wrona, *Informatyczne wspomaganie procesów informacyjno-decyzyjnych w budownictwie komunikacyjnym*, Warszawa, PAN (2015);
8. J. Szelka, Z. Wrona, Wykorzystanie inteligentnych baz danych w budownictwie mostowym, *Acta Sci. Pol. Architectura* **16** (3), pp. 129-135 (2017).