Case-Based Reasoning for the Design of Start-Stop Logic of Hydroelectric Power Stations

Fabrízio Nicolai Mancini¹*, Alexandre Rasi Aoki ¹²
¹ Programa de Pós-Graduação em Desenvolvimento de Tecnologia - Prodetec, Curitiba, Brazil; ² Institutos Lactec, Curitiba, Brazil.

ABSTRACT

The present article focuses on the construction of an Intelligent System for the Project Start-Stop Logic of Hydropower Plants, work carried out as master dissertation, allowing the retention of knowledge of previous cases which set a valuable intellectual capital, organizing knowledge generated by specialists and consultants of the area. The tool aims at saving time as well as standardizing solutions, allowing future developments for the construction of other necessary logic (such as those related to asset management and tools to aid the operation and maintenance of hydroelectric power plants). The MyCBR, software that proved to be of easy programming, but with some limitations, is used for such an attempt of a shell of current use. The recovery of the cases based on calculations of global and local similarity allows the selection of solutions already used in hydroelectric plants with analogous technological solutions. The results were satisfactory in view of the periodicity of the use of the tool and the possibility of increasing the case base as the logics are implemented in other projects, besides allowing the inclusion of other logics and documentation pertinent to the Hydropower Plants.

Keywords: Hydropower plant. Start and stop logic. Intelligent Systems - Case Based Reasoning. MyCBR.

* Author for correspondence: fabrizionicolaimancini@gmail.com
INTRODUCCION

Electricity is fundamental to modern life, it brings comfort, safety, and guarantees health and other benefits for the population. Brazil is a country with one of the cleanest energy matrixes, according to the International Energy Agency (IEA) [1], and the major source for electricity production is in hydroelectric plants. Despite the importance of the electric sector, its digitization was late and the automation of the hydroelectric plants did not have an adequate implementation according to Jardini; Mendes [2]. Usually, the same logic of conventional systems (based on electromechanical relays) has been implemented, showing an underutilization of the new technology, simply replicating the knowledge of specialists in the area. The computational advance and the digitization of the construction of these logics (start-stop of generating units of hydroelectric plants) is still carried out on a primary basis. Based on the growth of the field of artificial intelligence and its wide application in the various sectors of the economy, it is possible to use these techniques to improve productivity and evolution of the model itself.

OBJECTIVE

The general objective of this article is to contribute to the promotion of the automation of hydroelectric power plants, specifically through the modeling of an intelligent system that allows the construction of the hydroelectric power plant logic needed for the start and stop of the generating units, indicating the equipment usually used and which information is made available to the Digital Control and Supervision System (DCSS): Generator, Excitation System Turbine, Speed Regulation Systems, Lift Transformer, Shielded Bus, Auxiliary Services Transformer; using the Case-Based Reasoning (CBR) analysis technique and its peculiarities (Case Construction, Similarity Measures, Case Selection, Reuse, Case Adaptation and Case Base Feedback); evaluating and adapting to the knowledge a shell for CBR implementation; and finally discussing results and validating them.

METHODS

The method used to carry out the research is firstly based on a bibliographical survey on the control of hydroelectric generation and the intelligent system used, CBR. Subsequently we analyzed the documents of a company of the sector, Intertechne Consultores S.A., to build a model of knowledge. A shell was selected for the use of the technique. Finally, the CBR model was implemented, tested and validated. The structure of the problem for CBR is the fundamental step of the construction of the intelligent system, considering the need to organize the knowledge so that the selected technique can be applied. The logics for starting and stopping a hydroelectric plant have their solution provided in the logic diagrams, which contain individual logics. Due to these divisions of the solution, a survey was made of what is common in all logics, the identification of objects based on the logical knots presented by Standard 61850-7-410 [3] and the guide provided by 61850-5 -710 [4], as well as the present systems, equipment and services.
General Flowchart

For the construction of the CBR, in order to generate start-stop logic for hydroelectric plants, some activities were achieved systematically, which happened in two phases: data selection/treatment and CBR modeling [5].

The selection and processing of data aim to prepare the documentation to be used in the shell, considering eventual inconsistencies due to the multiple forms of representation of the data and the absence of a standardization as already presented, which can be seen in 1.

![Chart 1 – Data Selection and Processing](image)

Source: Authors based on [6]

The modeling of CBR involved several steps, according to Chart 2: case modeling, involving the determination of the type of knowledge representation, possible types and values; definition of similarity measures, both global and individual, including their specific formulations; insertion of the treated data to characterize each case in the case basis; case-study tests to validate case basis; availability of shell for users.
The knowledge representation allows the segregation of the knowledge produced in the construction of a particular Hydroelectric Plant so that it can be object of CBR, structuring the knowledge for its computability, especially for the calculation of similarity.

The choice of data types and possible values address the need to categorize knowledge, structuring it according to its nature: textual, symbolic or numerical. In addition to this, it allows the selection of possible values, being able to limit entries and visualize the similarities or nonexistence among all possibilities.

The definition of the similarity’s measure allows the approximation of the knowledge, identifying if a solution can partially or totally meet the query, stressing that local and global similarity measures must be generated.

The insertion of the data in the database seeks the adequacy of the data to the knowledge model, establishing the model or allowing its revision in order to fit all existing data.

After the cases are entered, simple tests validate the case basis to verify if the data have been correctly posted or if the data conform to the existing design and documentation.

**Data Selection and Treatment**

The selection of data begins with the survey of the available data, firstly examining the data provided by the adopted solution in previous cases, the logical diagrams. This choice stems from the need to assess whether the logic diagrams can provide all the data needed to build the system. The diversity of constructions for logics, sometimes in the same enterprise, demonstrates the need to adapt the data to the construction of a standardization of knowledge, allowing the comparison of projects, their systems and equipment, through global and local similarity measures.
The treatment of these data consists of the division of each of the pertinent information of every logic in order to identify the objects and the services provided by each of the systems of the hydroelectric power station. With the purpose of storing preferably the data that are pertinent to the desired solution and determinant for the consultation, the delimitation of the knowledge representation was applied. To assist in solving the discrepancies displayed in the logic diagrams and to identify the logical information used for purposes not explained in the logic diagrams (for example, points that serve only to prepare alarms for the operators, but do not influence on the digital process except in order to avoid that certain problems converge to trips, causing the stop of a generating unit) points lists (rich material, with relative ease for manipulation and data treatment) were used. In an ancillary way the workstatement, design criteria, single-line diagrams and three-way diagrams made it possible to indicate the project's specific choices, such as philosophies and architecture, allowing the transcription of important information in a reduced time.

Knowledge Representation

The representation of knowledge has as a limitation in the form of the possible representation in the used shell, that is, textual. Thus, the construction of the case basis will have as a source the construction of concepts and attributes, being possible to use concepts as attributes of other concepts, enabling a hierarchical construction of knowledge, following the flowchart of Figure 3: Logical diagram analysis; identification of systems; identification of equipment; transcription of the logic referring to the system, transcription of the point list necessary for the logic.

An example is that the protection system can point to some equipment: the 86M and 86E blocking relays and the unit protection relay. Logic can predict the need for all of them not to be actuated. The point list identifies the equipment with the actuated protections. Thus we have:

SYSTEM: Protection
PROTECTION EQUIPMENT: 86M Block Relay, 86E Block Relay, Generating Unit Protection Relay.
LOGIC TRANSCRIPTION:
LDN Blocking Relay 86M ACTUATED
ANDN Blocking Relay 86E ACTUATED
ANDN Generating Unit Protection Relay ACTUATED

The point list, preliminary to this logic, will have:
Blocking Relay 86M ACTUATED
Blocking Relay 86E ACTUATED
Generating Unit Protection Relay ACTUATED

Regarding this system, by compiling other documents, corroborated by a logical diagram, it can be established that the protection system of a plant has a protection system with only one-unit protection relay (with no alternating or secondary protection). Therefore, it is possible to establish another attribute of the protection, its configuration, which should be marked as a simple protection.

Types of Data and Values of Cases

The data types present in the used shell, MyCBR are: boolean, concept, data, double, real, integer, range, character, and symbol. The usual one is the symbolic type, since it allows the registration of data of the hydroelectric plants in a restrictive way, allowing only the knowledgeable engineer or the person in charge of the maintenance of the system to insert new data (such as systems, their configurations, equipment, point list among others). Another type of data used is character, because it allows, through strings, the construction of texts, which can be compared, but do not present a similarity measure. In order to perform the standardization, a large part of the equipment data and the system configurations were treated as a symbol, allowing the data to be restricted, while the required standardization was performed. For each attribute there was an option to include other information, including a specific attribute for this information.

Similarity Measures

The measures of similarities are composed by the identification of characteristics of the HPP, elements of the systems, their configuration and existing equipment, promoting the construction of local similarities and ways to identify more interesting alternatives when thinking about the reuse of the cases, since in some enterprises the system could be similar and fully utilized, creating an adaptation to take advantage of local similarity.

The different types of similarity measures are derived from the existing data type, for example, it is possible to restrict the protection systems in view of the power of the generator, which will identify, at the moment the logic is worked, those minimum protections for each of the operating ranges of the generating units.

RESULTS

The construction of the Intelligent System for the Start-Stop Logic of Hydroelectric Power Stations had as a starting point the survey and presentation of a theoretical framework related to the control of hydroelectric generation and the technique of artificial intelligence used, the reasoning based on cases. The control of hydroelectric generation and the identification of the components of the hydroelectric power stations used to construct the logics of starting and stopping allowed the restriction of the research object, causing an economy in the processing of the information made available in the CBR case basis. Another important factor was the organization of knowledge model based on the control logics identified by the bibliography, which, although restricted, corroborated with the documents used to construct the case basis.
The shell that was used to prove the case-based reasoning technique, called MyCBR [7], was chosen because it was developed by the Competence Center Case-Based Reasoning (CCCBR) of the Deutschen Forschungszentrum für Künstliche Intelligenz (DFKI) and by the School of Informatics and Technology of the University of West London (UWL) of the United Kingdom, studied and used by academics of the area of artificial intelligence, including in Brazil. The data and their treatments are treated in relation to the shell. Although they provide multiple representations, they also limit them by allowing only the use of texts and numbers, even in the symbolic categories.

**DISCUSSION**

The analysis of the results and their discussion allow the verification of the adequacy of the method used in the existing problem, making clear what was correct, what was not and its correction, as well as exposing the limitations of the technique and/or the tool used.

**Selection and Preliminary Treatment of the Data**

In order to select the data, we assessed the existing documents in a company that has carried out projects in the area for almost 30 years; these documents being organized in two management softwares called EPDM and SGI (legacy), which made the search faster and more punctual, allowing the location of at least 20 distinct hydroelectric plants, 134 logical diagrams, 40 point lists and 98 different documents, including architectures, single-line diagrams and design criteria. For the complete tests, due to the complexity and the time available for the completion of the work, four projects were chosen with six hydroelectric plants - some of them had small plants used as a sanitary flow and/or auxiliary service. It was determined 260 variables that could be stored for each case, restricting to the similarity calculation to a smaller universe of only 120 elements, that allowed the characterization of the hydroelectric power station with sufficient elements to search for the solution through the applied technique.

**Representation of Knowledge**

For the construction of the model, a multilevel separation was sought, identifying elements of the hydroelectric power plants based on objects and systems, which would allow the data to be split in a more orderly way, identifying concepts and their attributes as attributes of a larger concept. For example: the hydroelectric power plant would be the largest concept and within this concept there is an attribute with the specific configurations of each of the bearings, such as: Generator Anchor - Lubrication - Configuration, allowing the indication of the technological solution based on the models of the Hydraulic Power Plant or the inexistence of the system. The method was adequate, but with adaptations, considering the need to meet the peculiarities of a hydroelectric power station and of the shell used, since for each concept used as attribute there was a need to build instances (particular cases) which should not be linked to the cases (previous solutions), except when used to narrow the search spaces. The obstacle to this technique arises from the need to feed the information in the MyCBR itself, for it is not possible to establish instances with different nomenclature from the numerical one by means of text or table (as the shell allows), even with the export and the import of the case basis, incongruence also occurs, duplicating cases and attributes.
Similarity Measures
For each data a specific weight was established based on the statistic of the case basis and/or importance of the element (even if in the statistics it was not clear the existence of these distinctions, as it is the case of the turbines) or those in which it was statistically witnessed a technological diversity that makes it possible to limit the researched cases.

Case Basis
In order to use the instances in a given case basis, it is necessary to link it to the case basis. If there is no linking of the instances to the case basis at the time of recovery of cases, the result is null, due to the lack of elements to be searched.

Recovery of Cases
The recovery of cases is carried out through a search tool present in the shell that enables the consultation, the linking and the selection of the most appropriate solution based on the insertion of the data by the user within the limitations foreseen in the knowledge model and with focus on the criteria chosen by the constructor of the search system.

Result With a Complete Search
Tests were carried out with the completion of all elements listed for the calculation of similarity, obtaining answers as in the case 0706-MA, with the search linked to Fall, flow and other attributes, totalling 69 (one of them is the one containing the data of the operations, including the logics, point lists and others), a similarity of 0.63 was obtained, as compared to 0.47 of the others (due to the identity of solutions and little divergence in relation to the main criteria such as type of turbine, type of generator, among others). The selected case can be exported to a CSV spreadsheet, allowing the use of its data for the construction of the start-stop logic through the most similar cases and the knowledge of which systems have a similarity, being always exported the 4 cases with the greater similarity.

REUSE Cases
Reuse is effectively decided by the user, who may or may not use the solutions thoroughly or, based on local similarities, select the most appropriate solution, based on the systems that identify the knowledge fraction. In order to implement this step, 3 variables were created for each System / Equipment required for the start-stop logic: Combinational Logics, Sequential Logics and Point List. However, it is important to point out that they are in a different place, but at the disposal of the user, rather than in the General Data of the Hydroelectric Power Plant.

Solution Review
The review of the case should also be carried out by the user identifying incorrect solutions, excluding from the base imperfections and matching the search criteria and weights to make the search more consistent with each use. Two distinct moments can be used to review the cases: through a simulation to prove the effectiveness of the solution, made possible through the use of another tool containing a mathematical
model of a generating unit with the identified peculiarities; and with the construction of the as built of the logical diagram, presenting the consolidated version of the logic that in theory served (or will serve) as a basis for programming the logic needed by the hydroelectric plant.

Solution Retention

The retention of new cases should only be used once it has already been approved by the client (even before an as built), since it will have been reviewed by at least four different instances: verifier, approver and client itself, depending on the flow of documents set in the project by an owner’s engineering and integration phase. Its consolidation through the as built is desirable as it will bring greater credibility to the selected solution.

CONCLUSION

The technique is functional, because, from the characteristics of a hydroelectric power station informed by the shell, it identifies in the case basis a possible solution, indicating the logics to be used. It was also identified that, although the sample is representative, from a company with decades of experience in the sector, most cases are related to a specific type of turbine, a Francis Turbine, but other elements are common and for having similar technological solutions, which allows a significant use of logic, even with different turbines. It was also identified that queries generate inconsistencies in the case basis, specifically when using multiple-valued symbols, showing a limiting factor for shell utilization, which were bypassed by creating copies of that case basis, so that it can be imported with each new use (which is a plausible situation in view of the periodicity of the use of the tool and the necessity/possibility of including new cases). The Case-Based Reasoning technique is efficient for the construction of a tool that increases the productivity of projects in the electrical sector, allowing the consolidation of expert knowledge in a database that can evolve according to the construction of new knowledge of the sector, bringing safety, reliability and agility in the design of the start-stop logic of generating units in hydroelectric power stations. In addition, the work opens new fronts with the possibility of increasing the technique to other areas of the stations involving logic, as well as the increment in the case basis of specific solutions or fragments of solutions to attend specific logics for Pelton and Bulb type turbines, which are not present in the case basis, including important elements such as the logics for Pelton turbine injectors and deflectors as well as the peculiarities and similarities of Bulb and Kaplan turbines.

REFERENCES


Received: December 21, 2017;
Accepted: August 01, 2018