Computers in Industry 59 (2008) 936–948

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.elsevier.com/locate/compind

Knowledge-based collaborative engineering of pipe networks in the upstream and downstream petroleum industry

Leonid Sheremetov *, Ildar Batyrshin, Manuel Chi, Armando Rosas

Mexican Petroleum Institute (IMP), Eje Central Lazaro Cardenas 152, San Bartolo Atepehuacan, Distrito Federal, Mexico

ARTICLE INFO

Article history:
Available online 30 August 2008

Keywords:
Collaborative design
Knowledge-based engineering
Interoperability framework
Agents

ABSTRACT

This paper describes an approach to the development of a framework for the automation of the detailed engineering cycle based on the integration of concurrent and collaborative engineering models with knowledge-based engineering. The framework is illustrated by the example of pipe networks design and stress and flexibility analysis in the upstream or downstream petroleum industry. Engineering tools are integrated with a Multi-Agent System for helping stress analysis engineers to identify whether a given pipe arrangement can cope with weight, thermal and pressure stress at safe operation levels and suggesting available solutions on pipeline layout and support distribution based on historical cases. Overall, the system architecture is discussed and implementation details are provided.

1. Introduction

In today's world, engineering companies must be able to reduce their time to market and adapt to the changing environments. Decisions on design alternatives must be made quickly and they must be done right the first time out. Concurrent and collaborative engineering (CCE) is a business strategy which replaces the traditional product development process with one in which tasks are done in parallel by collaborating engineering teams. This strategy enables an early consideration for every aspect of a product's development process increasing product quality while reducing work time. The Department of Project Engineering of the Mexican Petroleum Institute (IMP) has focused its activity to CCE business processes in the constant search for efficiency and productivity of engineering projects. The key enabler of these innovative practices is the advanced information technology.

The usage of engineering software is a daily engineering practice. The software tools, like CAD/CAE packages and simulators, facilitate the routine work. Nevertheless, there are several challenges to be addressed. First, current tools often tend to isolate information at tool boundaries, creating islands of automation. Second, concurrent engineering models are usually organization-specific, which require adaptation of 3D CAD systems employing feature-based and parametric modeling capabilities to the company's engineering practices. Finally, a decision-making process is usually grounded in the experience of the designers: no reasoning is delivered from simulators about what to do if input data does not assure a required solution. The engineering software does not come-up with an acceptable solution; engineers have to manipulate variables and parameters to achieve a convenient output that fulfils the mean requirements.

The solution here lies in the integration of product lifecycle management, CCE, and knowledge-based engineering (KBE) paradigms. For more than two decades, the knowledge-based systems (KBSs) have successfully provided intelligent support to humans during the process of analysis and design [10]. At the beginning, the interest was namely focused in expert systems. In particular, in the area of pipeline networks design, the identification of piping design rules and how these rules can be incorporated into an expert system using a common subset of LISP language was reported by [7]. Another expert system named FRAES was designed to assist the technical personnel of petroleum companies in the analysis, design and diagnosis of flexible pipes (flowlines or risers) for offshore applications [4]. Numerical algorithms, databases and expert knowledge were explored by the inference mechanism of the system.

Recently the interest in KBS has been renewed by the challenge of dealing with complex dynamics in business process re-engineering as well as in concurrent engineering in an automated fashion. Such systems look for integrated engineering environments also known as KBE or intelligent CAD systems in order to reduce the amount of design man-hours and human errors. In [8], an expert system shell and a geometric modeling kernel were

* Corresponding author. Tel.: +52 55 9175 7546; fax: +52 55 9175 7463.
E-mail addresses: sher@imp.mx (L. Sheremetov), batyr@imp.mx (I. Batyrshin), machi@imp.mx (M. Chi), aelguera@imp.mx (A. Rosas).

0166-3615/$ – see front matter
doi:10.1016/j.compind.2008.07.004

Table 1
Software tools cited in the paper and their description

Software tool	Manufacturer	Description
AutoCAD	AutoDesk	CAD software application for 2D and 3D design and drafting
CADDS 5	PTC	Specialized CAD/CAM software for discrete manufacturers that design massive, complex assemblies
CADDS 5 ISSM	PTC	Integrated Surface and Solid Modeler is a suite of object-oriented software libraries providing 3D surface and solid modeling capabilities as a customer programming option for CADDS
CADWorx Plant	COADE Engineering Software	AutoCAD-based design and drafting system for creating orthographic and isometric 2D and 3D piping drawings
CAEPIPE	SST Systems	Software for stress and deformation analysis, eigenfrequency and dynamic analysis of a piping system in accordance with international standards
CAESAR II	COADE Engineering Software	Software for pipe stress analysis evaluating structural response of piping systems and their code-defined stresses caused by a wide variety of loads in accordance with international standards
CAPNET Knowledge Acquisition Tool and Expert System Shell	Mexican Petroleum Institute	Generic tools for capturing and processing of knowledge bases from any application domain with inference machines based on fuzzy algebras originally developed for Component Agent Platform over .NET (CAPNET)
LISP	Massachusetts Institute of Technology	List Processing Language is a family of computer programming languages for artificial intelligence applications
MS Biztalk Server	Microsoft	Business process management server permitting to automate and integrate business processes
Nexpert Object	Neuron Data	Expert System Shell using knowledge representation based on objects
OntoCAPE	AVT	A large-scale ontology for the domain of CAPE
PDMS	AVEVA	Plant Design Management System for 3D process plant design
PDS	Intergraph	Plant Design System is a CAD/CAE application for plant design, construction, and operations
SmartPlant 3D	Intergraph	3D plant design environment
WizWhy	WizSoft	Data Mining tool for issuing predictions, summarizing data, and if-then rules generation

integrated for design process automation. The CADDS 5[1] was used as the overall CAD environment, the Nexpert Object was used as the expert system shell, and the CADDS 5 ISSM was used to build the user interface through which geometric models of pipes are created and modified. In [9], a general description of the process used to create a KBE for CCE was proposed. As shown, linking them together provides a wide variety of synergistic effects. In [17], a Web-based KBE architecture was proposed for collaborative product development as a result of the analysis of several existing implementations principally focused on the design. Information sharing aspect of concurrent engineering also has been focus of research work since the paradigm was proposed [11].

Recently, engineering resources integration and engineering services orchestration (provided by these resources) have been identified as crucial issues to construct collaborative design environments. For such environments, an intelligent software agent has advantages over existing engineering frameworks in sharing engineering information, data, and knowledge among the engineers [5,18]. In this paper an integrated framework for automation of the detailed engineering cycle based on the integration of CCE model with KBE is described. This paper extends the framework first proposed in [21], by going deep into details of knowledge-based components and its interoperability architecture. These components are represented by fuzzy expert systems for (i) critical lines identification and ranking, (ii) historical layout similarity check, and (iii) pipeline thickness analysis. The first two components are described in the paper. The main contributions of the proposed approach are (i) the extension of the standard pipe design domain representation model with fuzzy reasoning capabilities and (ii) the interoperability architecture enabling interaction of intelligent agents and engineering tools at the knowledge level in order to automate the whole process of stress analysis using agent-enabled Service Oriented Architecture (SOA) and orchestration by a business process management system. Thus, the framework addresses the above mentioned challenges in several aspects: (i) enables information sharing between engineering tools using intelligent software agent technology; (ii) intelligent agents wrap the reasoning capabilities of knowledge-based components thus helping both experienced and inexperienced engineers to come-up with acceptable solutions; (iii) bridges the gaps between the concurrent engineering model and 3D CAD systems enabling engineering services orchestration by the workflow of pipe stress analysis (PSA) process.

The framework is developed for the design process of pipe networks and stress and flexibility analysis in the upstream or downstream petroleum industry. Two cases are presented in the paper to illustrate the proposed methodology. In the first case study, the ability of the system to classify critical lines using different methods (rule base and heuristic algorithm) was tested and compared with that of an expert. The obtained results are discussed and some of the problems with existing procedures are highlighted. In the second case study the system was used to implement the whole workflow and record the behavior of the framework in a real engineering project scenario.

The rest of the paper is organized as follows. In the next section, decision-making in PSA is briefly described and a general solution scheme is developed. In Section 3 the methods implemented in knowledge-based components of the proposed framework are explained. Section 4 describes the interoperability architecture for CCE and knowledge management capabilities of the infrastructure. In Section 5 implementation details of the proposed architecture are presented. Section 6 describes the results of the experiments followed by conclusions.

2. Motivation for research: pipe detailed engineering

The design process of pipe networks can cover up to 40% of the whole process of detailed engineering of facilities both in the Upstream (offshore oil platforms) and Downstream Petroleum

[1] Please find a brief description of the cited software tools in Table 1 at the end of the paper.

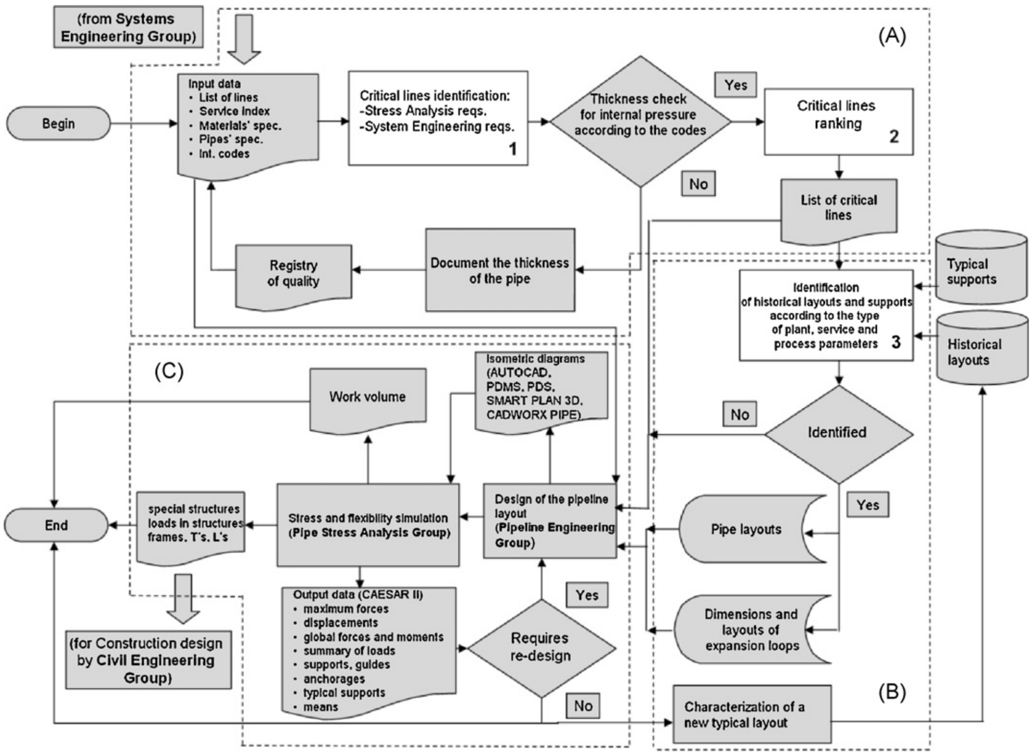

Fig. 1. Pipe stress analysis workflow in pipe detailed engineering.

Industry (refineries and chemical plants). PSA is a part of pipe design. It is used to identify whether a pipe arrangement will cope with weight, thermal and pressure stress at acceptable levels under engineering design standards for safe operation. Decision-making includes: identification of critical pipelines, top-down ranking of critical pipelines, flexibility of the piping system as for relaxing overstressed points, and proper and balanced supports distribution.

In the conventional practice, PSA is carried out at the final stages of the pipe system design process (block C from Fig. 1). It is usually done using simulation software such as CAEPIPE or CAESAR II. A pipe simulation model is constructed from piping general layout, piping isometric drawings, piping and valve specifications. Pipe design is usually carried out in different engineering platforms like PDS, PDMS, SmartPlant 3D, CADWorx Plant, etc. depending upon the client's requirements. If design data input to simulator fulfils specified constraints and there are no overstressed points, then the simulator's output is OK. Otherwise, the system is "stiff" and conditions violation is indicated. Whenever a formal PSA of the pipe design indicates an overstress point, decision-making over possible solutions imply routing changes to make the systems more flexible, supports distribution or both. Such back and forth design iterations between piping and stress departments continue until a suitable layout and support scheme are found.

It is obvious that the sooner (at least preliminary) PSA is done, less time is spent on the re-design of the pipe network. A 'design and analyze' process developed in this paper for existing or to-be designed pipe networks is illustrated in Fig. 1. Compared to the traditional practice, it includes two additional blocks. At the first stage (block A), the critical lines are preliminary identified according to the design data provided by systems engineering group, e.g., pipeline lists, service indexes and specification for pipeline materials. In traditional practice, this identification is performed by the PSA engineers considering the type of the facility (offshore platform, refinery, petrochemical plant, etc.) and pipes' operational conditions. A process to check thickness of pipelines can be also performed at this point based on ASME standard, feeding back the designers if necessary for appropriate reviewing. The list of critical lines is ranked based on the lines' parameters from most critical to less critical ones.

At the second stage (block B), in order to provide the pipe designers with guidelines on critical lines design, similar historical cases are analyzed and suggested to the piping designers, based on the operational parameters such as the type of connected equipment, temperature, pressure and diameter. At the same time, for the header pipes, the dimensions of expansion loops are calculated and the isometric drawings of the suggested solutions are generated. Finally, once the detailed Pipe Layout Design is complete, the isometric drawings of the most critical lines along with the additional information stored within engineering databases are returned to the PSA group for the detailed stress and flexibility analysis using simulation software (like CAESAR II). These activities are similar to that of the traditional process (block C). The only difference consists in identifying and characterization of new typical layouts to be stored in the DB of historical cases along with the results of stress and flexibility tests. The results also allow for developing specifications for final building structures used in construction design by Civil Engineering Group and calculations of final work volume. In the following section it is explained in details how the main activities of this process are implemented using KBE techniques (white boxes in Fig. 1).

3. KBE components of the framework

KBE components of the framework implementing the functionality described in Fig. 1, are specialized on critical lines identification, applying both system and stress design criteria (box A-1 in Fig. 1), critical lines ranking (box A-2), and identification of similar historical cases both for pipe layout and supports (box B-3).

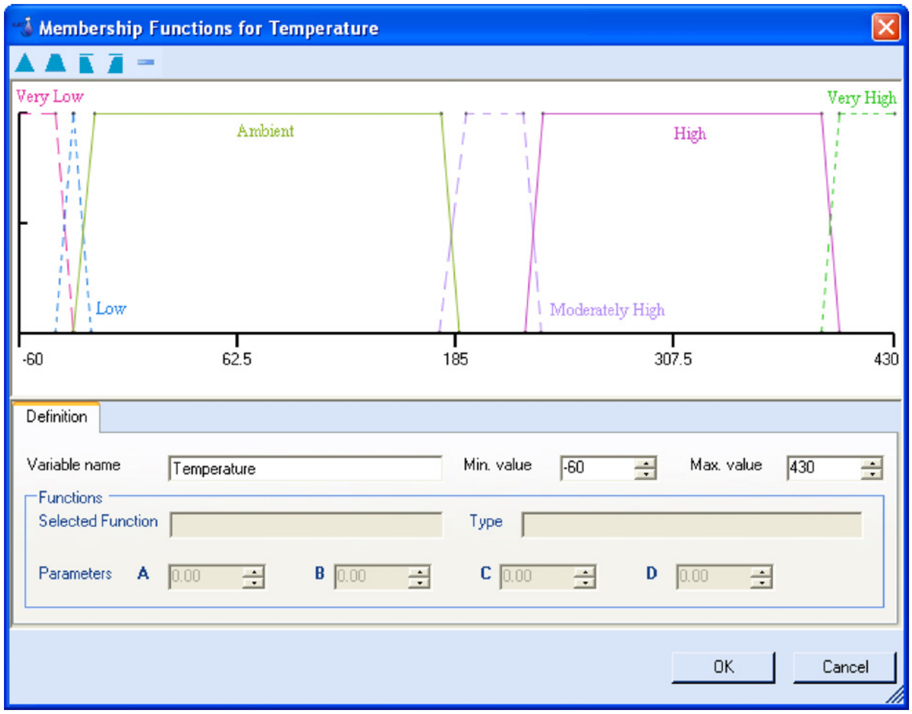

Fig. 2. Definition of "Temperature" as a linguistic variable.

3.1. Critical lines identification and ranking

In spatial design of piping systems including pipe routing, component placement and placement of pipe supports, the typical projects contain 500–800 pipes and up to 70% of them are usually considered as potentially critical subject to a wide variety of factors that take into account weight, pressure, thermal, seismic and other static and dynamic conditions [1]. Identification of possible critical lines at the project's early stages is based on the analysis of those parameters that are available at this stage from the systems engineering team:

- Operational conditions (temperature and internal pressure).
- Equipment to which it connects.
- Fluid that transports.
- Material of the pipe and accessories (valves, bridles, elbows, etc.).
- Diameter of the pipe.
- Fluid type: toxics, explosives or that is operated in two phases.
- Systems of pipes, whose fault is critical for the technological process like vents or transfer lines.
- Processes whose licensing requires special additions (expansion joints, pulsation bottles) or layouts without loss of pressure.

The KBE components for critical lines identification and ranking implement two models: one based on the rule-based model of the domain and the second one using a heuristic method defined by the PSA experts.

3.1.1. Rule-based component

For the development of the first model, rule mining techniques are used for developing the rule set and representing it in CAPNET Knowledge Representation Format (KRF). The CAPNET KRF is an extension of the FIPA-RDF and represents elements from the world by using Objects [2]. These are identifiable entities from application domain with a unique name and a list of Properties defining their state. If values for some properties are restricted to a list of possibilities, then it is said that the values have Constraints. Objects are grouped in a wider concept known as Resource. Pipe design domain representation model is based on OntoCAPE [14]. OntoCAPE as of version 1.0, contains about 600 chemical engineering concepts and 400 relations.

One of the characteristic features of the model is the possibility to handle fuzzy variables by defining fuzzy sets and membership functions. As an example from the PSA domain, let us consider fluid temperatures, which could be *very low*, *low*, *ambient*, *moderately high*, *high* and *very high* (Fig. 2).[2] Fig. 2 presents a screenshot of a module used for definition of membership functions (MFs). In each specific case they are defined by experts or experienced users. Trapezoidal MF with overlapping curves is the most popular MF used in fuzzy modeling [6]. From Fig. 2 for instance, a temperature of 399 °C lies between the two latter ranges with an equal membership value of 0.5 in [0,1] membership scale. These membership values converted to the values of plausibility *pv* in {0,1,2,3,4,5,6} scale are used in the rules. So, once interpreted in terms of a linguistic variable, two facts with *average possibility* (plausibility value equals to 3 in plausibility scale) are generated:

Fact 1: Temperature is high <*pv* = 3>
Fact 2: Temperature is very high <*pv* = 3>

Rules define relationships between known facts and information that can be concluded. Antecedents and consequents are propositions that relate properties from resources to some values. Available pipes lists with expert identifications of the critical lines were used for rules generation applying WizWhy 3 Data Mining tool [12]. WizWhy identifies associations between the parameters of the data sets and represents them in a form of IF-THEN association rules. Once the rule set is generated, it has to be converted into CAPNET KRF rule format. As explained in [19], CAPNET KRF model assigns plausibility values to the rules in order to manage uncertainties. The following method to convert the WizWhy rules into CAPNET rules was used: all the positive supporting rules (with the conclusion: *line is critical*) were given

[2] MSSP-69 Design Codes were used as a reference [14].

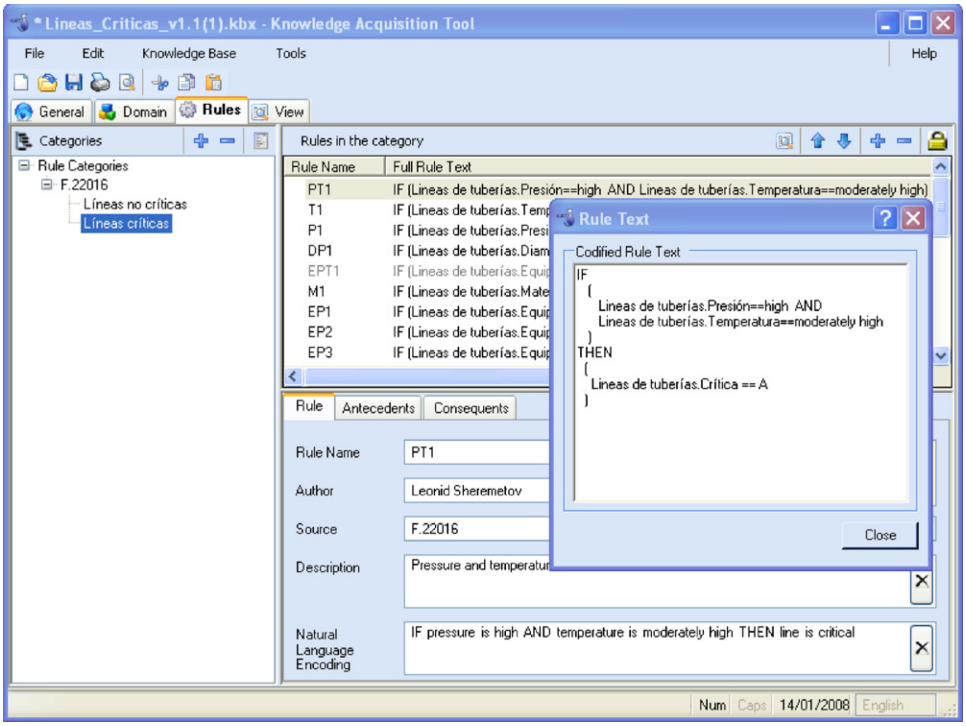

Fig. 3. A screenshot of the KAT v2.0: rule edition mode.

plausibility values from 6 to 4 in {0,1,2,3,4,5,6} scale depending upon the rule's probability[3] in the following way: probabilities from 70 to 80 were given value *possible* (4), from 80 to 90 – *highly possible* (5), and from 90 to 100 – *sure* (6). Similarly, the negation supporting rules (with the conclusion: *line is not critical*) were given plausibility values from 4 to 6. Let us illustrate the above by an example. Suppose the following rule was generated by WizWhy from the training sample:

Rule 1: If *pressure* is 19.00 ... 158.40 (average = 61.80)

and *temperature* is 190.40 ... 395.00 (average = 197.90)

Then

Critical line is Yes

Rule's probability: 0.914

Comparing the temperature values with those from Fig. 2, it can be seen that the rule interval covers several linguistic values for the *temperature* variable (the same happens with the *pressure* value). So, using fuzzy variables component this rule will be converted to the CAPNET KRF rule:

Rule 1: IF (*pressure* == high OR *pressure* == very high)

AND (*temperature* == moderately high OR *temperature* == high OR *temperature* == very high)

THEN line == critical, <weight = 6>. (rule's *pv* = 6)

The described KRF was implemented in CAPNET Knowledge Acquisition Tool (KAT) and used for agents' KB development. Fig. 3 illustrates a fragment of the rule development using CAPNET KAT. At the left-hand side of the screenshot a structure of rule categories and problem domain resources appear. A structure of a rule (for the first part of Rule 1 described above) appears in the window Rule Text.[4] Antecedents and consequents of the rules are propositions that relate properties from resources to some value by one of the following operators: equal to, greater than, greater than or equal to, lower than or equal to and not equal. The rules are stored both in internal (proprietary kbx format used by the CAPNET KAT) and external (XML) format used for file sharing purposes.

CAPNET Expert System Shell implements both conjunctive and disjunctive multi-set based algebras to manipulate uncertainties in CAPNET KRF rules [19]. The Shell has its own interface and can be used for debugging purposes during expert system development. On the other hand, the Shell provides a set of API functions incorporated in the class *CAPNET_Engine* that serve for loading a domain KB and initial facts, invocation of inference mechanisms (forward and backward chaining), retrieving the list of inferred facts with explanations, etc. This API is used by knowledge-based agents. The details on the handling of the possibility of the facts and rules (*pv* in the case of facts and *weight* in the above rule example) by means of plausibility values are introduced in [19]. To handle this and other types of uncertainties, two fuzzy inference engines, forward chaining and backward chaining, working with linguistic scales were developed. These inference engines combine traditional forward chaining and backward chaining inference mechanisms with multi-set based algebras of fuzzy inference described in [19].

In order to rank the critical lines based on the number of fired rules and their plausibility values, the generated rules have to be converted into the accumulative weights of the critical lines. The following conventions are adopted: positive supporting rules obtain positive factors from 1 to 3 for plausibility values from 4 to 6, respectively. Similarly, negative supporting rules obtain negative values from -1 to -3. It is supposed that the final knowledge base does not have rules with weights less than 4.

3.1.2. Knowledge component based on experts' heuristics

Traditionally experts marked the critical lines in the project based on their experience considering most of the factors mentioned above. To order all lines by criticality level it was

[3] The term "probability" designates what other data mining tools call "Confidence Level".

[4] Instead of using OR connector actually Rule 1 is divided into several simple rules.

proposed an approach taking into account the importance of criticality factors. The total evaluation of the criticality of the lines $C(L)$ was calculated as a weighted sum of partial evaluations C_J, $J \in \{T, E, P, D, M, F\}$ for temperature (T), equipment (E), pressure (P), diameter (D), material (M) and fluid (F):

$$C(L) = w_T * C_T + w_E * C_E + w_P * C_P + w_D * C_D + w_M * C_M + w_F * C_F, \quad (1)$$

where the partial evaluations C_J are calculated by the method specific for each factor. For temperature, pressure and diameter they are calculated as respective values of considered characteristics in comparison with their maximum values for all lines in a project. Partial evaluations for equipment, fluid and material are given by experts and depend on the type of considered factors describing the line.

3.2. Pipe layout similarity identification

The project uncovered a substantial volume of historic information on detailed engineering projects developed by the IMP Engineering Department for different plants of the Mexican oil company *Petroleos Mexicanos* (PEMEX), which were carefully analyzed by the experts in order to identify the most typical pipe layouts. Unfortunately it was impossible to use project database (DB) directly as a DB of historical cases and most of pipe layout solutions had to be re-captured. The need to re-create the isometrics was caused by several reasons. Firstly, historical traces were stored in different incompatible formats, so the decision was made to use AutoCAD *dwf* as a portable format. Secondly, most of these layouts corresponded to relatively complex lines and were represented in several isometric drawings (usually from 3 to 7). So, in order to obtain a complete circuit they had to be combined in only one drawing. Another reason was to have so called "intelligent drawings" including links to referenced elements. These elements are of two types: typical and standard supports stored in the DB and results of the simulation tests. The latter were included in order not to store just the most frequently used pipe layouts connecting different equipment types under certain values of parameters (temperature, pressure, etc.), but really typical layouts understanding the ranges of parameters they could cover. To get this, additional simulations of the parametrical limits of these layouts in terms of flexibility and stress were carried out.

The knowledge-based component for the similarity identification of the layout and supports of pipes is based on the heuristic model, developed by the PSA experts, that allows determining the similarity of historical layouts from the typical layouts repository with respect to the current project's pipes. This model allows measuring that similarity quantitatively, since it provides a numerical value whose interpretation will allow determining to what extent a historical layout is applicable to the requirements of a new pipe.

The model of calculation of similarity distance $d(L_c,L_t)$ is based on the following function:

$$d(L_c, L_t) = d_E(\lambda_T d_T + \lambda_D d_D + \lambda_P d_P + \lambda_M d_M), \quad (2)$$

where

- L_c is a pipe identified like critical during the process of identification and ranking,
- L_t is a typical pipe layout contained in the historical layouts repository,
- E, T, D, P, M are parameters defining pipe's critical factors: type of equipment, temperature, diameter, pressure and material respectively,
- λ_X are the weights of each component X of similarity determined during the identification of critical lines,
- d_X is the distance or similarity between the critical line and the typical pipe layout with respect to the parameter X.

For the calculation of similarity by temperature, pressure and diameter, the fuzzy set theory is used to define the membership functions of trapezoidal type for these parameters in a similar way that was described above in Section 3.1.1. For example, in the case

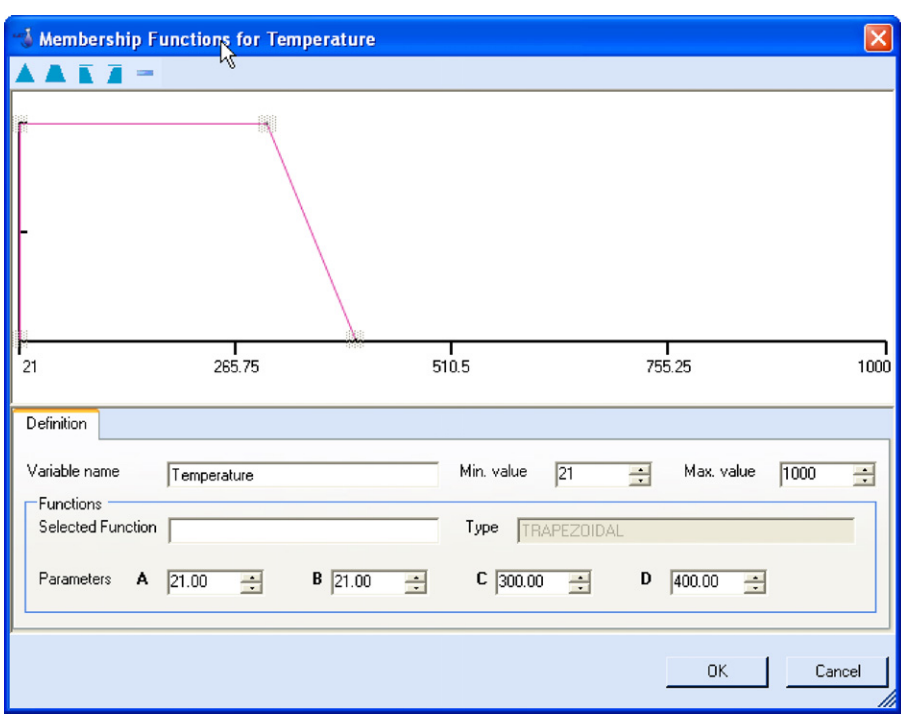

Fig. 4. Trapezoidal membership function for the temperature parameter.

of the positive temperature, the trapezoidal function, visualized graphically with fuzzy values definition software component, is exemplified in Fig. 4.

In Fig. 4, $A = B = 21$ (the room temperature), C = maximum temperature supported by the typical pipe, determined by a flexibility analysis using CAESAR II simulator and D = an approximate value determined by the PSA experts.

In order to calculate the value of the similarity considering the material of the pipes, the maximum allowed stress values according to the ASME B31.1 design code are used [15]. The procedure of this calculation is the following. First, the maximum values of the allowed stresses s_1 and s_2 are determined for the material of the critical and typical lines respectively at the required temperature. Then the distance is calculated as follows:

$$d_M = \left\{ \begin{array}{ll} 0 & \text{if } s_1 = 0 \text{ or } s_2 = 0 \\ 1 & \text{if } s_2 \geq s_1 \\ \dfrac{s_2}{s_1} & \text{if } s_2 < s_1 \end{array} \right\} \quad (3)$$

4. Technological infrastructure for knowledge-based collaborative engineering

In this section, the interoperability architecture developed to integrate knowledge-based components of the proposed framework with engineering applications during the project's lifecycle is described. This technological infrastructure allows teams from different Engineering disciplines (Processes, Systems, Instruments, Pipes, Civil, Mechanical, PSA) to work, on a shared project, in joint fashion and simultaneously. This way the information, produced by each design group, can be shared within the project context. This architecture allows interchanging data and integrating different engineering applications, simulators and tools (Fig. 5).

In order to enable homogeneous and consistent access to the information within different design processes, the approach shares the philosophy of SOAs, recent developments of W3C and international standards. The central part of this architecture consists on Project's Lifecycle Management (PLM) component and a set of Web services (WS) accessed through a façade called Project's Information Integration System (PIIS). PLM orchestrates the whole design process, also managing the repository for documents generated by different design teams.

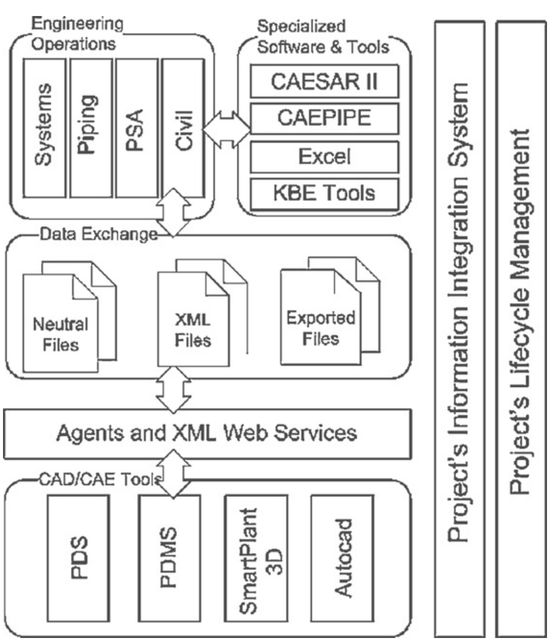

Fig. 5. Interoperability architecture.

Each one of the PIIS WS performs a specific transform task for a given pipe network design platform (PDS, PDMS, SmartPlant 3D), by producing output and input exchange files wherever is possible. Exchange data formats of this architecture include neutral files which have been developed by industry, built-in-house XML files based on schemas following ISO 15926 guidelines, and finally files resulting from direct import/export operations of these design platforms.

KBE tasks accomplished by agents are seamlessly integrated within the framework using agent and Web services interoperability [20]. The first two agents (responsible for critical line identification and ranking) are enabled with CAPNET Expert System Shell component as embedded fuzzy inference engine. They also interact with personal assistants of design teams in order to interchange information on critical events for the CCE design process as shown below. Agents share a domain knowledge representation model providing a common and consistent symbolic representation of application (offshore platforms or petrochemical plants) and problem (detailed engineering) domains. The methods of rule-based knowledge representation used by agents are described in Section 3.1.1.

PLM and PIIS façade is accessible to the users through a web portal. By using this portal, the users get the whole picture of project lifecycle current status, and collaborate through the access and update of document repository. By using the façade facilities, the users obtain suitable files in order to get their specialization tasks finished. In this way, for example, PSA engineers can get files with suitable CAESAR II format for flexibility analysis, or civil engineers can apply ad hoc tools to calculate work volume from 3D design files.

Communication between the components takes advantage of agent technology and CAPNET content language enabling communication at knowledge level. Many portions of a design space are of interest only to one agent, while other portions must be common to many agents. So, the main elements of the interaction scheme are: the domain model based on shared ontology and a communication language enabling knowledge transfer among agents.

5. Implementation of the framework

The proposed framework has been implemented in the application called SATD-AFE which stands for Decision Support System for Stress and Flexibility Preliminary Analysis (in Spanish). The business process model developed for the orchestration of the Web and agent services (AS) composing SATD-AFE follows the PSA process model described above (Fig. 1) according to the business process scheme shown in Fig. 6.

The first step in the process is critical lines identification and ranking within a particular project applying the methods described in Section 3.1. These lines are analyzed for their critical value. Agents applying system and stress criteria, interchange their results in order to produce an integrated list for ranking. This process involves the use of the *IMP_SATD-AFE_CL* Web service, which is actually an orchestration of several services provided by the above mentioned agents, specified in BPEL4WS and implemented using MS Biztalk Server [13]. At this point thickness reviewing agent is executing a task, whose results will define if the process continues or adjustments for pipeline specification are to be made, as shown in the workflow in Fig. 1. The KB for this agent includes the information items of application domain required to make calculations based on ASME code for thickness review. Once the critical lines ordering is complete, an agent for searching the

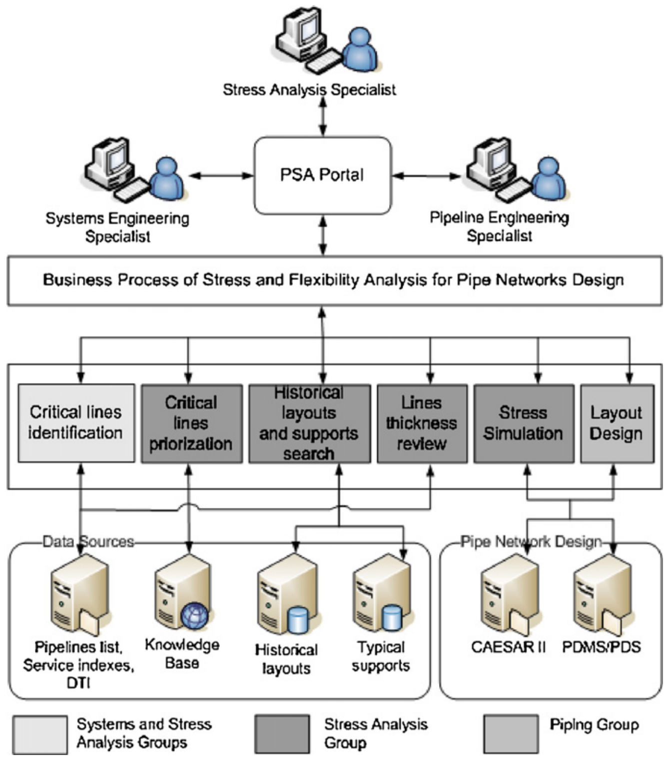

Fig. 6. Business process implementation scheme.

historical layouts uses its KB and a fuzzy calculations engine in order to get similarity measures for historical cases, as the mentioned workflow states.

All the involved agents expose their functionality as WS using the CAPNET AS-WS integration mechanism described in details in [20]. In the case study, several distributed information sources were modeled: pipelines list, service indexes and pipe and instrumentation diagrams (DTI). To get information from these sources a *GD_System_Service* WS was developed. This service is responsible for converting respective formats into CAPNET KRF objects.

As shown in Fig. 7, at the PSA portal the lines are ranked based on their criticality index (see column 2). For each pipe, its detailed ranking can be further decomposed in the absolute and relative fractions of the criticality index, defining the contribution of each of the critical factors (see P8445 service as an example). At the same time lines thickness analysis is carried out. As a result of the ranking procedure, the *inform* message is sent to the pipe designers assistant agent informing it that the first phase of preliminary stress analysis has finished and the Pipe Layout Design personal can access the list of layouts based on their similarity grades available for the most potentially critical pipes, which criticality index is above the threshold. For each pipe selected from the list, similar historical cases are searched in the DB of historical cases. The next screenshot (Fig. 8) displays the list of similar layouts found in the DB for the P8445 service. The table displays the similarity index (%) calculated according to the formula (2) for each found pipe calculated as described in Section 3.2. For each particular case the detailed levels of similarity can also be shown (historical layout with the isometric file named AE-586-300-301 C1168 in Fig. 8).

Once the similar historical cases are identified, the user can select the case to be used as a reference for the Pipe Layout Design. Obviously, the larger is the similarity factor, the greater is the probability that the historical case will fit the design requirements. At the moment, the DB contains typical layouts with associated metadata for eight successful projects of the IMP Department of Project Engineering. In the case the AE-586-300-311 C1168 pipe (95% similarity) is selected (the last row shown in Fig. 8), its isometric diagram is displayed as shown in Fig. 9. At the same time,

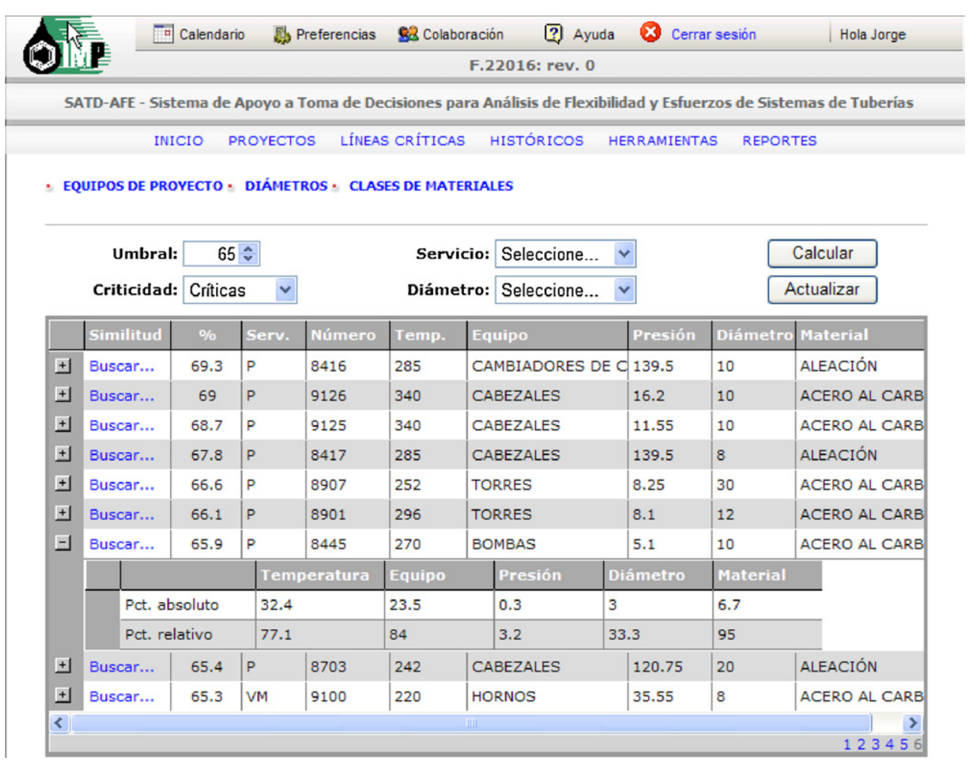

Fig. 7. SATD-AFE interface for critical lines ranking with the threshold equals to 65.

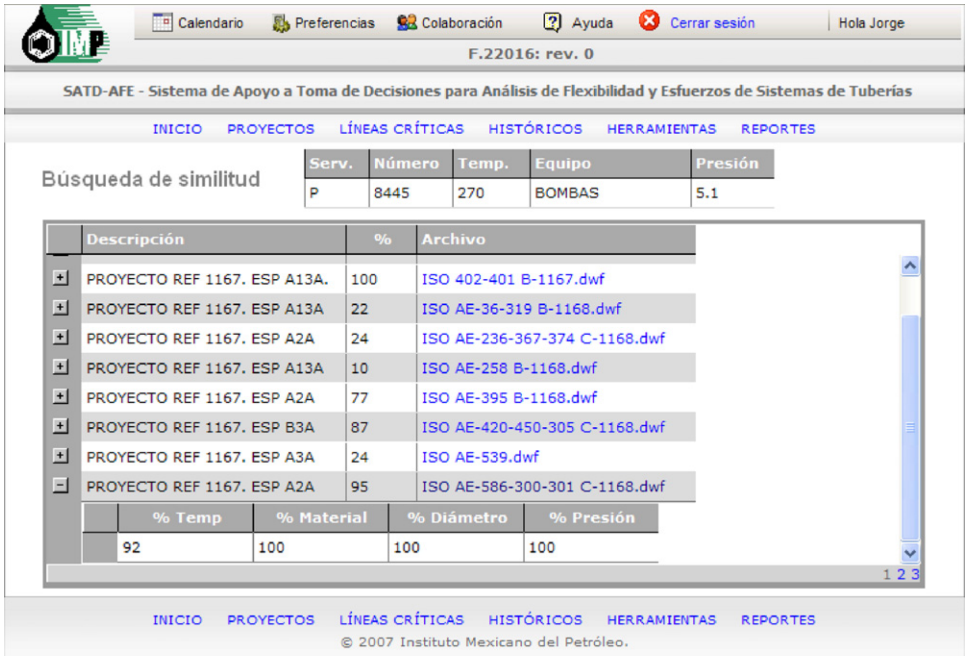

Fig. 8. SATD-AFE interface for similarity search from the historical cases DB.

the hangers and supports used for the corresponding historical layout are identified along with simulation results on flexibility and stress analysis. The flexibility and stress analysis report shows detailed information for loads applied on every node of the design which are the results of the simulations carried out for the typical layouts using CAESAR II® simulator. Overall of about 100 typical and standard supports are stored at the moment. An isometric diagram stored in the DB usually contains two types of links on these additional documents associated with each node: the first leads to the associated hangers and supports while the second displays the parameters of flexibility and stress checks (shown in separate windows in Fig. 9).

At the same time, the sum of loads is sent to the Civil Engineering personal assistant to ask for preliminary work load calculation. The particular agent interested in obtaining this information has to be previously subscribed by posing a persistent query within the vocabulary of the common ontology. When relevant information is published, it is forwarded on to the requesting agent as though told directly by the publisher. The *request* message for calculation action has the following coding:

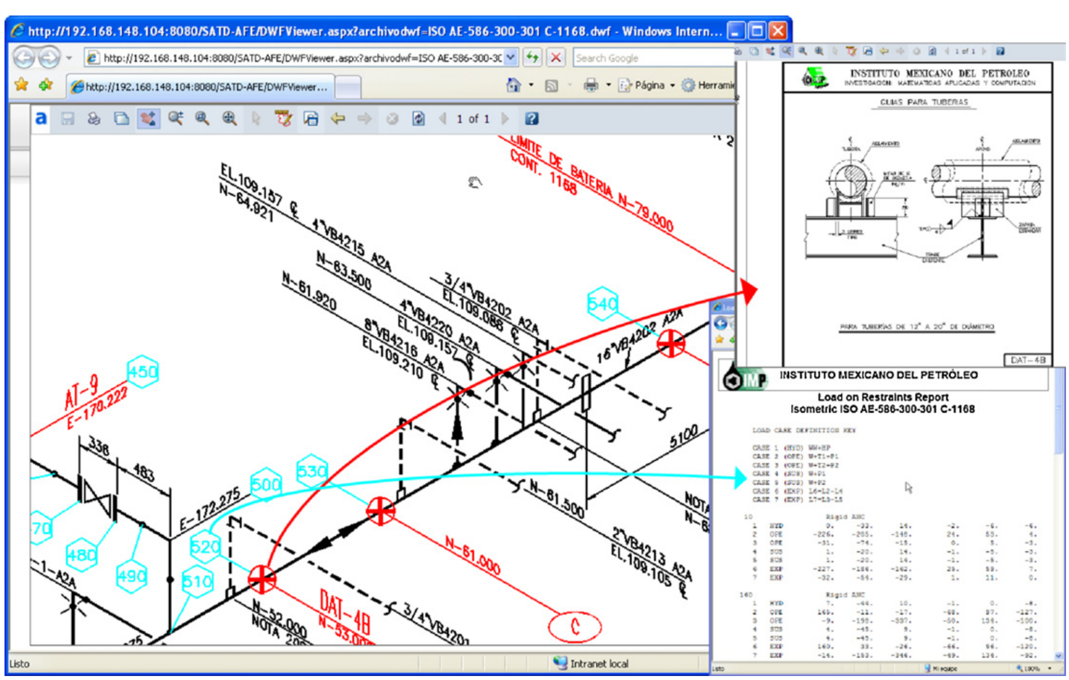

Fig. 9. A screenshot of the AutoCAD viewer for similar historical designs (file AE-586-300-311 C1168.dwf) with the associated typical support 4B corresponding to the node 520 (right-hand side upper window) and the parameters of flexibility and stress checks for this node as reported by CAESAR II® (right-hand side lower window).

```
<fipa-message act="request">
 <sender><agent-identifier><name id="Support@imp.mx" />
<addresses><href="http://192.168.142.63:4444/MTSURI"/>
    </addresses></agent-identifier>
 </sender>
 <receiver><agent-identifier><name id="civil@design" />
  <addresses><"href=http://192.168.142.63:4444/MTSURI"/>
      </addresses></agent-identifier>
  </receiver>
 <content>
  <request_content>"type=CAPNETCL-ACTION"><RDF0Action id= "id_civil">
<actor>transport1</actor><act>calculateSupportWorkLoad<act/>
      <ActionArguments>
     <property>
              <propertyname>support</propertyname>
              <propertytype>String</propertytype>
               <propertyvalue>…</propertyvalue>
            </property>
        </ActionArguments></RDF0Action>
   </request_content>
  </content>

 <language>capnet-cl</language>
 <content-language-encoding>fipa.acl.rep.xml.std</content-language-encoding>
 <ontology>ontoCAPE</ontology>
 <ontology>IMP_design</ontology>
</fipa-message>
```

Once the layout design is finished by the design group, stress analysis personal assistant agent is informed. In the case the historical layout was used directly for the new design, the hits counter for this layout is increased. On the other hand, if modifications to the layout were made (or completely new layout was created), stress simulation should be done. Two different WSs *GD_PDS_Service* and *GD_PDMS_Service* create input data files for CAESAR II® simulator depending upon the CAE/CAD tool used by pipe designers. Stress simulation is made by a stress designer. CAESAR II also has a neutral data file format for independent use in exchanging data with other programs. Piping input and output can be also directed to an ODBC database for data review and manipulation outside CAESAR II. Once the simulation is over and the layout fulfils specified constraints and there are no over-stressed points, the PSA personal assistant agent determines if the layout should be included in the DB by analyzing its similarity to the already stored typical layouts. For the implementation of the framework, from the server side the following tools were used:

- MS .Net Framework 3, the base platform for development, deployment and execution.
- MS SQL Server 2000, used as repository for historical layouts and data for support of the entire framework.
- MS Biztalk Server 2004 for orchestration and choreography of the business process.
- MS IIS Web server for final deployment of the web system.
- CAPNET Agent Platform with Expert System Shell for hosting of executing agents.
- MS Sharepoint Portal server, used as infrastructure for meeting, collaboration and information exchange of Department of Project Engineering specialists.

6. Analysis of the experimental results

At the first phase of the experiments, the ability of the system to classify and rank critical lines using the proposed methods was tested and compared with that of an expert. Later, the similarity identification method between critical and typical lines was tested. Due to the space limits, only the results of the former experiments with the heuristic method of lines ranking (Section 3.1.2) are discussed below.

The application of the approach to evaluating the criticality level of lines gave the possibility to order lines by criticality level and design first those lines that had the highest values of this level. The obtained results caused some discrepancy with direct expert indication of critical lines such that some lines evaluated by a specialist as critical have been obtained small level of criticality and, vise versa. Two reasons of such discrepancy were determined: (1) different levels of expertise of PSA specialists and as a consequence, errors in indication of critical lines in traditional approach; and (2) subjectivity in evaluating weights of importance of line characteristics in multi-criteria evaluation of criticality level of lines (1). Fig. 10 illustrates this error on one of the example projects. Distribution of critical and non-critical pipes in a project with 837 lines with respect to the calculated weights of line criticality: (a) distribution of critical (1) and non-critical (0) lines evaluated by a PSA specialist with respect to heuristic multi-criteria evaluation of criticality of lines, (b) number of correctly classified critical (529) and non-critical (308) lines depending upon the threshold (level of criticality), (c) the absolute classification error as a function of selected level of criticality: (minimum criticality level = 0.337, minimum error = 105 lines or 12.468%) and (d) the relative classification error as a function of selected

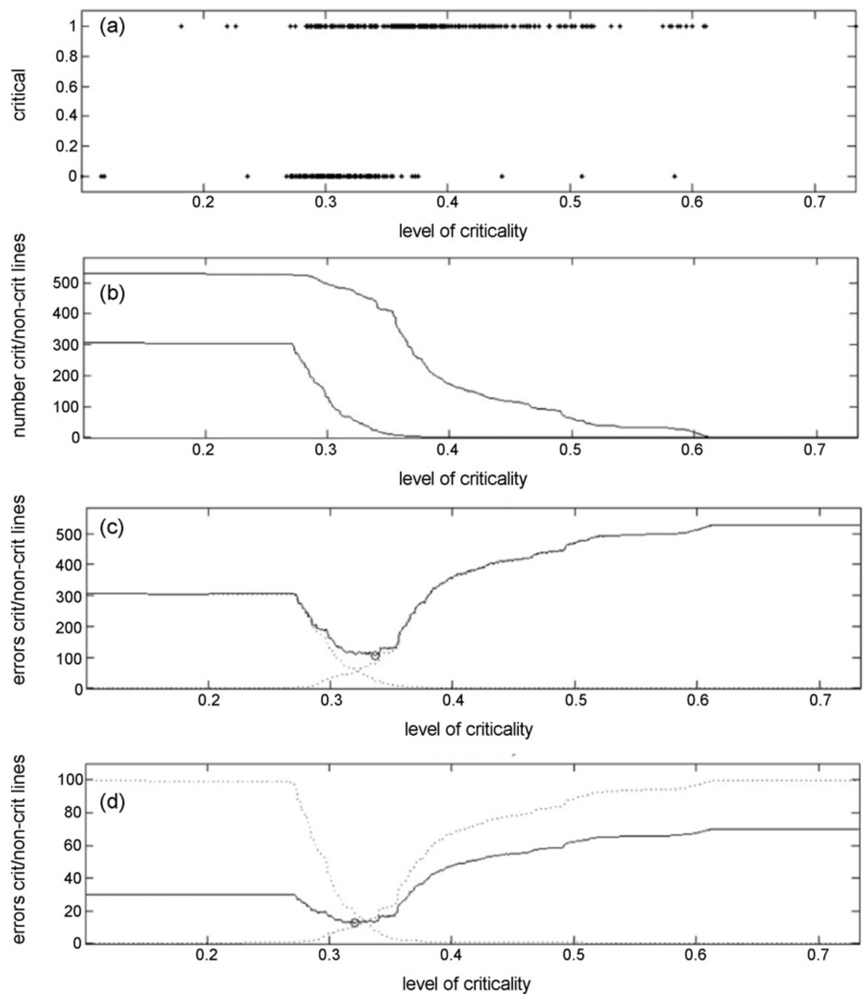

Fig. 10. Illustration of error in a sample project.

level of criticality: (minimum criticality level = 0.32, minimum error = 12.468%). For the optimal level of classification the general error can be less than 12%.

In order to overcome both limitations the basic algorithm was modified. First, cross-validation of expert indications of critical lines was done for several projects when experts also marked critical lines for each of the following factors: temperature, pressure and equipment. In this case a line was indicated as critical if it was critical at least by one of these factors. Cross-validation of expert indications showed that in some projects 33% of lines indicated initially by a specialist as non-critical after detailed separate analysis of most important factors were evaluated as critical. This surprising result can be explained by a well-known fact that it is difficult for people to make multi-criteria decisions in large number of experiments such as evaluation of 500–800 lines usually contained in one project. In the following, expert indications of critical lines based on separate analysis of most important characteristics were considered as more reliable.

Second, to minimize the error between expert indications of critical lines and multi-criteria evaluation of criticality level of lines the optimization procedure was applied which adjusted subjectively defined weights of importance of line factors in multi-criteria evaluation of criticality level of lines.

As it was found during the experiments, the more careful analysis of lines by an expert based on evaluation of different characteristics of lines usually increase the number of lines considered as critical. In current work, the evaluation of the level of criticality of lines is based on the following characteristics: temperature, pressure, fluid type, diameter and material of pipes, equipment used and expert opinion. Hence, there is a multi-criteria problem of automatic evaluation of the level of criticality of the lines. Several possible solutions of this problem based on discussions with experienced engineers were adopted. First, the critical values of characteristics based on MSSP and ASME B31 Process piping standard [15] were determined. Further, all factors were ordered and weighted with respect to their influence on the evaluation of a criticality of a line. For temperature, pressure and diameter they were calculated as respective values of considered characteristics in comparison with their maximum values for all lines in a project but this respective values are normalized in a different way depending whether the value of characteristics is greater or less than the given critical value. The set of factors has a hierarchical structure because the weight of some factors like equipment and material depends on their specific value for considered line. For example there are more than 30 types of equipment used in pipelines: ovens, compressors, turbines, caldrons, reactors, pumps, turbo generators, etc. material: steel, stainless steel, and alloy (see Fig. 7 for an example of the lines and their descriptive parameters belonging to one of the tested projects).

Further testing of the proposed approach to evaluation of criticality of lines on different projects shows good accordance of

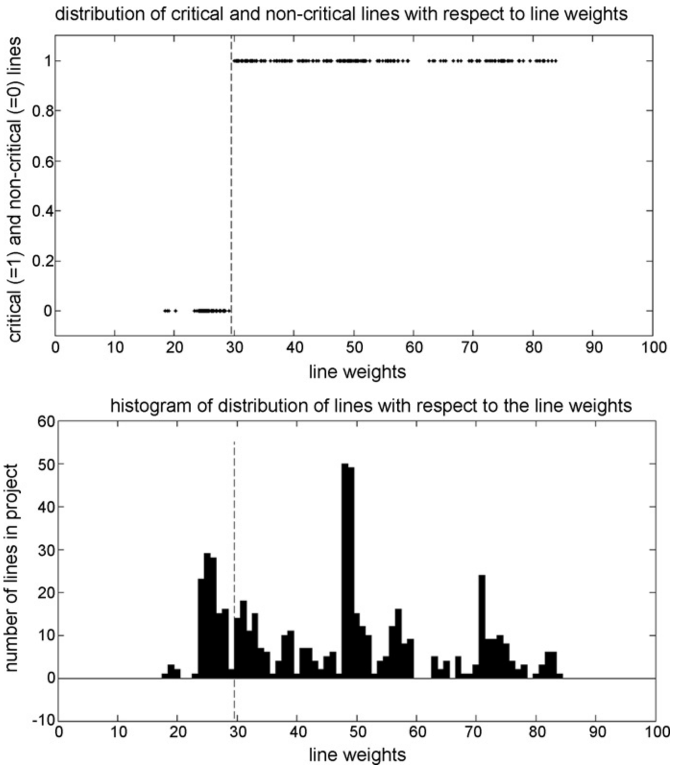

Fig. 11. Distribution of critical and non-critical pipes in a project with 548 lines with respect to the calculated weights of line criticality.

obtained evaluations with expert evaluations and experience. Fig. 11 shows an example of distribution of lines with respect to criticality evaluation for another engineering project containing 548 lines with 428 lines evaluated by expert as critical. As it can be seen, the evaluation of criticality index of the lines at a certain level can be split into critical and non-critical lines; the result which is in a perfect correspondence with expert evaluations.

At the current stage of experiments, the system has been used to implement the whole workflow and record the behavior of the framework in a real engineering project scenario. In this scenario, the interoperability architecture, response time, and user satisfaction were tested. The first project has been just finished and evaluations revealed a 20% time save for the PSA and pipe design teams.

7. Conclusions

The paper describes the framework for knowledge-based collaborative engineering, implemented for pipe design, stress and flexibility analysis tasks for engineering projects in petroleum industry. Rule mining techniques, fuzzy logic, expert systems and agent technologies were used as a technological base of the framework. The main idea behind the developed framework is the integration of different types of information and knowledge (numerical, graphical, linguistic, etc.) typical for engineering applications. Moreover, the developed models permit the integration of knowledge with different types of uncertainty such as fuzziness, possibility and probability. For the needs of sharing engineering information, data, and knowledge among the engineers, the proposed approach extends the existing engineering tools by intelligent software agent technology. Intelligent agents wrap the reasoning capabilities of knowledge-based components. Their interaction with engineering tools takes place at the knowledge level. To facilitate collaboration of teams of piping, stress and civil engineers in pipe networks design, agent-enabled service oriented architecture and orchestration by a business process management system are used. The interoperability architecture is grounded in the novel integration model for agent and Web services. The main advantages of the proposed framework are the following:

- The ‘‘downstream’’ aspects (final cost, manufacturability, safety, packaging, and recyclability) are now considered at design stage (thickness parameter).
- Integration of data produced by different applications and engineering platforms.
- Minimal errors in the data input between design teams, guaranteeing consistency and integrity of the data and technical information.
- Scalability offering the possibility of integrating new platforms and tools.

The proposed framework has been implemented in the application called SATD-AFE for the design of pipe networks and their stress and flexibility analysis and tested in real engineering project scenarios. Experimental results permitted to enhance the precision of the developed methods achieving a perfect match with the expert classifications. Further efforts are focused on extending the database of historical layouts which could allow increasing the similarity hits. The ongoing work is also centered at the generalization of the proposed framework to the whole process of detailed engineering of oil&gas facilities which could be the subject of the next paper.

According to the classification proposed in [16], the main industrial efforts nowadays are centered at level's-1 interoperability (agreed data representation formats, like WITSML and ISO 15926-2:2003), where the data itself can be translated from one application to another one and interoperability between applications is orchestrated by the companies themselves [3]. Nevertheless, for concurrent engineering applications, the upper levels, such as level 2, where application components can request to be notified when the contents of data objects encapsulated by servers change, are of primary interest. The paper shows how agent-enabled architecture permits to respond to this challenge. The level 3, which enables users' control of multiple applications through a common interface, is being developed within the IMP project D.00415. This functionality will take full advantage of using common ontologies. This way, combining agent and Web services in collaborative engineering processes it will be possible to solve complex operations and to create services with a greater added value.

Acknowledgements

Partial support for this research work has been provided by the IMP, within the project D.00413. The authors gratefully acknowledge the contributions of Ing. Lorena Velazquez, Ing. Javier Chavez, Ing. Jorge Martínez and Hugo Galicia. We also highly appreciate the support of the Director of the IMP Project Engineering Business Unit Ing. Arturo Rosales and the Coordinator of the Research Program on Applied Mathematics and Computing Dr. Jorge Velasco, as well as our technical peers Ing. Juan Antonio Rodriguez from the IMP and Ing. Jose Luis Gonzales and Ing. Sergio Melendez, both from PEMEX. The principal investigator of the project D.00415 is Jorge Perez. The authors gratefully appreciate the valuable comments of the reviewers that were very helpful for preparing the final version of the paper.

References

[1] M. Alvarado, M.A. Rodríguez-Toral, A. Rosas, S. Ayala, Decision-making on pipe stress analysis enabled by knowledge-based systems, Journal of Knowledge and Information Systems 12 (2) (2007) 255–278 (Springer Verlag).

[2] M. Contreras, E. Germán, M. Chi, L. Sheremetov, Design and implementation of a FIPA compliant Agent Platform in .NET, Journal of Object Technology 3 (9) (2004) 5–28 (ETH Zurich).

[3] daratechPLANT 2007. http://www.daratech.com/plant2007/.

[4] J.A.B. Geymayr, N.F.F. Ebecken, Integration of knowledge sources for flexible pipe evaluation and design, Expert Systems with Applications 10 (1) (1996) 29–36.

[5] Q. Hao, W. Shen, Z. Zhang, S.-W. Park, J.-K. Lee, Agent-based collaborative product design engineering: an industrial case study, Computers in Industry 57 (1) (2006) 26–38.

[6] J.-S.R. Jang, C.T. Sun, E. Mizutani, Neuro-Fuzzy and Soft Computing. A Computational Approach to Learning and Machine Intelligence, Prentice-Hall International, 1997.

[7] R.H. Jellesed, Developing an expert system to link the piping designer to computer-aided design, Pressure Vessels and Piping Division 177 (1989) 49–54.

[8] S.-S. Kang, S. Myung, S.-H. Han, Design expert system for auto-routing of ship pipes, Journal of Ship Production 15 (1) (1999) 1–9.

[9] J. Marra, Use of knowledge-based engineering in compressor rotor design, in: Proceedings of the International Gas Turbine and Aeroengine Congress and Exposition, June 5–8, Houston, TX, 1995, 95-GT-384.

[10] R.M. O'Keefe, A.D. Preece, The development, validation and implementation of knowledge-based systems, European Journal of Operational Research 92 (3) (1996) 458–473.

[11] J.G. McGuire, D.R. Kuokka, J.C. Weber, J.M. Tenenbaum, T.R. Gruber, G.R. Olsen, SHADE: technology for knowledge-based collaborative engineering, Journal of Concurrent Engineering: Applications and Research 1 (2) (1993).

[12] A. Meidan, Wizsoft's Wizwhy, in: Data Mining and Knowledge Discovery Handbook, Springer, US, 2005, pp. 1365–1369 (see also: www.wizsoft.com).

[13] Microsoft Biztalk. http://www.microsoft.com/biztalk/default.mspx.

[14] J. Morbach, A. Yang, W. Marquardt, OntoCAPE—a large-scale ontology for chemical process engineering, Process Systems Engineering 20 (2) (2007) 147–161.

[15] Piping. Codes and Standards of American Society of Mechanical Engineers. http://store.asme.org/category.asp?catalog%5Fname=Codes+and+Standards&category%5Fname=Piping&Page=1.

[16] J.F. Rainaud, A short history of the last 15 year's quest for IT Interoperability in the Petroleum E&P Industry, Oil and Gas Science and Technology 60 (4) (2005) 597–605.

[17] K. Rodriguez, A. Al-Ashaab, Knowledge web-based system architecture for collaborative product development, Computers in Industry 56 (January (1)) (2005) 125–140.

[18] W. Shen, D.H. Norrie, J.P. Barthes, Multi-Agent Systems for Concurrent Intelligent Design and Manufacturing, Taylor and Francis, London, UK, 2001.

[19] L. Sheremetov, I. Batyrshin, D. Filatov, J. Martínez-Muñoz, An Uncertainty Model for Diagnostic Expert System Based on Fuzzy Algebras of Strict Monotonic Operations, in: Lecture Notes in Artificial Intelligence, vol. 4293, Springer Verlag, 2006, pp. 165–175.

[20] L. Sheremetov, M. Contreras, Industrial application integration using Agent-Enabled SSOA: CAPNET Case Study, in: W. Shen (Ed.), Information Technologies for Balanced Manufacturing Systems, IFIP Series, Springer Verlag, 2006, pp. 109–118.

[21] L. Sheremetov, I. Batyrshin, M. Chi, E. Vergara, A. Rosales, Knowledge-based Collaborative Engineering of Pipe Networks in the Upstream and Downstream Petroleum Industry, in: W. Shen, Y. Yang, J. Yong, I. Hawryszkiewycz, Z. Lin, J.-P. Barthès, M.L. Maher, Q. Hao, M. Hong Tran (Eds.), Proceedings of the 11th International Conference CSCW in Design, Melbourne, Australia, April 26–28, IEEE Computer Society Press, 2007, pp. 458–463.

Leonid Borisovitch Sheremetov graduated (with honours) from St. Petersburg State Marine Technical University (former Leningrad Shipbuilding Institute), Russia in 1982. The same year he joined St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences were he obtained his Ph.D. in computer science in 1990. Currently he is a Principal Investigator of the Research Program in Applied Mathematics and Computer Science of the Mexican Petroleum Institute (IMP). He has led a number of projects for the energy industry with special focus on industrial exploitation of computational intelligence and agent technologies. His research interests include Multi-Agent Systems, Hybrid Intelligent Systems, Data Mining, and Semantic Service Oriented Architectures. He has published more than 200 refereed technical papers in reputable journals, conferences and book chapters. In 2007 he was awarded the Annual IMP Award in Research. He serves as Associate Editor of the IEEE Transactions on Systems, Man and Cybernetics, Part C and of the Int. Journal Computer Science and Systems. Served as a Guest Editor of several journals. He is member of the Technical Committee on Computer Supported Cooperative Work in Design of the IEEE Systems, Man and Cybernetics Society and member of the Technical Committee on Artificial Intelligence and Expert Systems of the IASTED.

Ildar Zakirzjanovich Batyrshin (M.S., Moscow Physical-Technical Institute, 1975; Ph.D., Moscow Power Engineering Institute, 1983; Dr.Sci., Highest Attestation Committee of Russia, 1996) joined the Department of Informatics and Applied Mathematics at the Kazan State Technological University, Tatarstan, Russia, in 1975, and served as its chairman and full professor from 1997 to 2003. Since 1999, he has also been with the Institute of Problems of Informatics of Academy of Sciences of Tatarstan, as a Leading Researcher. Since 2003 he is a Researcher of Research Program in Applied Mathematics and Computer Science in Mexican Petroleum Institute. His areas of research are Soft Computing, Data Mining, Decision-Making and Expert Systems. He is a co-author and co-editor of several books and special journal issues on Soft Computing and Data Mining published in Russian and English. He was awarded the State Scientific Scholarship, given by the Presidium of Russian Academy of Sciences in 1997 and 2000, the title Honored Scientist of the Republic of Tatarstan, Russia, in 2002. In 2007 he was awarded the Annual IMP Award in Research. He has held long-term visiting research positions at various institutions in Germany, Turkey, Hungary, Italy and Czech Republic. Currently he is a Secretary of the Society of Mathematics of Uncertainty, a member of Council of International Fuzzy Systems Association (IFSA) Society and Past-President of the Russian Association for Fuzzy Systems and Soft Computing.

Manuel Chi Chim graduated from Mathematics Faculty of Yucatan University (Mexico). Currently he is about to get a M.S. degree in Computer Science at the National Polytechnic Institute of Mexico. Currently he is a research assistant of the Research Program in Applied Mathematics and Computer Science of the Mexican Petroleum Institute. His research interests include Service Oriented Architectures, Distributed Systems, Software Security and Multi-Agent Systems.

Armando Rosas Elguera received the Diploma in Civil Engineering from National Polytechnic Institute of Mexico. He joined Mexican Petroleum Institute in 1979 and has about 30 years of experience as a specialist in flexibility and supports of critical piping systems for the process industry. During his career back in 1979 he was piping stress and flexibility specialist, later office head of piping flexibility, coordinator and representative of the IMP in the Laguna Verde project (a nuclear power plant in Mexico). He was also the pipe stress analysis department head from 1994–1998. Currently he is a researcher in the applications of pipe stress analysis. He has deep practical experience in pipe stress analysis for nuclear power projects, for process and power plants involving all the different phases of engineering projects, from engineering design to plants start up and operation.