



Combining Case-Based Reasoning and Process Mining to improve collaborative decision-making in products design

Fatima-Zahra Berriche, Besma Zeddini, Hubert Kadima, Alain Riviere

► To cite this version:

Fatima-Zahra Berriche, Besma Zeddini, Hubert Kadima, Alain Riviere. Combining Case-Based Reasoning and Process Mining to improve collaborative decision-making in products design. AICCSA 2015, Nov 2015, Marrakech, Morocco. pp.1 - 7, 10.1109/AICCSA.2015.7507199 . hal-01396533

HAL Id: hal-01396533

<https://hal.science/hal-01396533>

Submitted on 15 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Combining Case-Based Reasoning and Process Mining to improve Collaborative Decision-Making in Products Design

Fatima-Zahra BERRICHE, Besma ZEDDINI, Hubert
KADIMA

Laboratoire QUARTZ/EISTI
Avenue du Parc
95000 Cergy Pontoise, France
[fbe,bzi,hk}@eisti.eu](mailto:{fbe,bzi,hk}@eisti.eu)

Alain RIVIERE

Laboratoire QUARTZ/SUPMECA
3, rue Fernand Hainaut
93407 Saint Ouen, France
alain.riviere@supmeca.fr

Abstract— *In this paper, we present a research in progress that expose an integral collaborative decision making process combining Case-Based Reasoning approach and the Process Mining techniques (CBR-Mining) to improve designing of manufacturing products. In collaborative decision-making participating actors have different objectives, constraints, knowledge, and viewpoints. The purpose of this paper is to illustrate via a use case study how process mining techniques may be integrated into Case-base-Reasoning.*

Keywords— *Product development process; knowledge management; decision support; case based reasoning; process mining; Industrial design*

I. INTRODUCTION

Rapid innovation and costs reduction during the collaborative design of industrial products is a strategic issue for industrial companies. In this context, several actors are involved in a collaborative environment with different views during different phases of the product lifecycle, while exchanging knowledge expressed in heterogeneous formats. This knowledge is related to different concepts such as product structure, usage, project history and design activities, parameters and constraints. Increasingly, enterprises are realizing how knowledge is accumulated in recent years, in order to emphasize the "know-what and Know-how". The exploration and management of these knowledge allow for the discovery of the inflow of resources that can be used to optimize product innovation and to make knowledge-driven decisions.

In collaborative decision-making, participating actors formalize the system, according to different viewpoints during different phases of the product life cycle. Collaborative Decision-making is a common decision process for two or more operational processes. It depends on the information, outputs of operational processes that are continually changing

and thus introduce the unpredictability in the decision-making process.

Our research approach involves the construction of a knowledge mining model to facilitate support decision in product design. Firstly, we try to develop a generic model which covers all information on past design processes including the applied solutions. Secondly, since the design process is the process that is based on the experience of the designers, it oriented us to a new combination of process mining and case-based reasoning. In our case, the case based reasoning is used to browse the previous design cases to provide good design practices inspired by previous experiences. Process mining on the other hand aims at discovering a process model that merges all previous similar solutions.

This paper is organized as follows. Section II introduces case based reasoning and process mining, we also present some contributions for their integration. Section III describes some major features needed for the capitalization and reuse of knowledge in this context. Section IV presents the standard design model product. The Section V describes our approach, which is used to support decisions in design. Section VI illustrates the first results. Section VII concludes the paper and discusses some future improvements.

II. CASE BASED REASONING AND PROCESS MINING

The CBR solves problems by finding similar cases in its knowledge base and adapting to the particular case. This solving is made of a number of phases: case representation, indexing, similarity comparison, retrieval and adaptation. In the literature, several contributions to the combination of data mining techniques with the CBR process were proposed. Arshadi and Jurisica have combined CBR with data mining to improve case-based classification [1], Kim and Han proposed a new method of indexation and classification based cases,

based on the use of artificial neural networks competitive [2], Beddoe used genetic algorithms to facilitate selecting and weighting features to personnel rostering [3].

The character of the product design process oriented us towards the use of process mining, because it is a new method that has been recently used in various domains, such as healthcare [4], time prediction [5] and in various other areas for analyzing resource behavior [6]. But it has not yet been applied in the domain of product design. In our work, we decide to adopt an approach that combines case based reasoning with process mining in order to improve the design products process and to support decision. We will discuss in more detail this approach in the next section.

A. Case based reasoning

The Case-Based Reasoning (CBR) was introduced in 1975 by Minsky. In 1994, Kolodner formalized the foundation of CBR in his book [7] and in the same time Aamodt and Plaza formalized the CBR cycle [8]. The CBR is a process that involves the reuse of past experiences, and was used in expert systems and cognitive science. In this approach, the user trying to solve a new problem recognizes similarities with previously solved problems called cases. A case is a commonly specific problem that has been identified, resolved, stored and indexed in a memory with the solution, and optionally the process for obtaining it [9][10]. The CBR systems are applied in many fields such as: medicine, trade, industrial diagnosis, monitoring and financial analysis.

1) The lifecycle of CBR

The lifecycle of a CBR is composed of four steps: Retrieve, Reuse, Revise, and Retain.

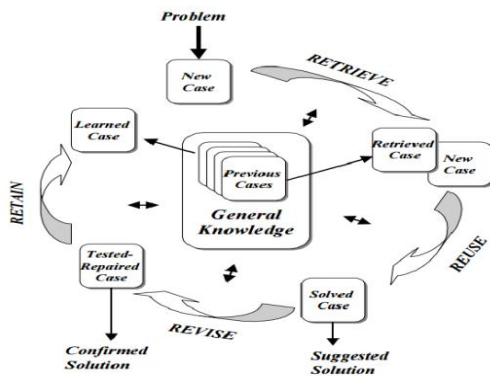


Fig. 1. The lifecycle of CBR [9]

a) Retrieve:

Given a target problem, retrieve from memory cases that are relevant to solving it. A case consists of a problem, its solution, and, typically, annotations about how the solution was derived.

b) Reuse:

Map the solution from the previous case to the target problem. This may involve adapting the solution as needed to fit the new situation.

c) Revise:

Having mapped the previous solution to the target situation, test the new solution in the real world (or a simulation) and, if necessary, revise.

d) Retain:

After the solution has been successfully adapted to the target problem, store the resulting experience as a new case in memory.

B. Process mining

Process mining aims at discovering, monitoring and improving real processes by extracting knowledge from event logs available in today's information systems. Log files may contain data recording three perspectives: behavioral (tasks and their time of execution), informational (data used and produced by tasks) and organizational (actors which perform tasks and their relationships).

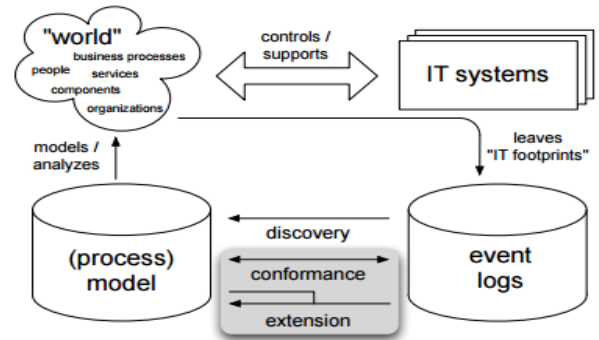


Fig. 2. Three types of process mining: **Discovery**, **Conformance**, and **Extension (Enhancement)** [11]

a) Discovery:

Produces a process model from event logs without using any *a priori* information. The discovered models can be represented by formal languages such as Petri nets or notations such as Business Process Model and Notation (BPMN).

b) Conformance:

Check if reality conforms to the process model: compare an existing process model with an event log of the same process. It is also used to detect, locate and explain deviations, and to measure the severity of these deviations.

c) Enhancement:

Extend or improve an existing process model using information about the actual process recorded in some event log.

1) Process mining tools

In the literature, there are many tools using process mining, we present some of these products in the following table :

TABLE I. EXAMPLES OF PROCESS MINING PRODUCTS[7]

Product name	Type	Organization
ARIS Process Performance Manager	Commercial tools	Software AG
Entreprise Visualization Suite	Commercial tools	Businesscape
Disco	Commercial tools	Fluxicon
Genet/Petrify	Academic tools	Universitat Politècnica de Catalunya
Intersage BPME	Commercial tools	Fujitsu
OKT Process mining suite	Open-source tools	Exeura s.r.l
Process Discovery Focus	Commercial tools	Iontas
ProcessAnalyser	Commercial tools	QPR
Rbminer/Dbminer	Academic tools	Universitat Politècnica de Catalunya
ProM	Open-source tools	Process mining group
Reflect	Commercial tools	Futura Process Intelligence

We use for our work, ProM tool [11] because it is an Open Source framework and it supports all of the process mining techniques presented in figure 2. For example, each of the discovery algorithms (genetic mining, heuristic mining, fuzzy mining, alpha algorithm, etc.) are integrated in the ProM tool.

III. MODELING KNOWLEDGE FOR PRODUCT DESIGN PROCESS

The innovation and performance of design activities for industrial products become a major issue for business success. To be placed high on the globalized market, it must not only combine customer satisfaction, productivity and competitiveness, but also to deal with the growth of technology and the increasing volume of available and accessible information. The management of their own information has always been one of the issues for companies. This information is currently processed and managed taking into account the meaning and semantics, which means that it manages instead of knowledge. In order to help engineers from various disciplines work together to identify the choices that will guide and justify decision making in design product, we decide to build our knowledge base capturing the knowledge generated by the practice of product development process and to represent them by the principle proposed by MOKA methodology (Methodology and tools Oriented to Knowledge based engineering Applications)[12][13]. This principle consist of collecting "ICARE" forms defining knowledge elements as follows:

- Illustrations representing comments, past experiences, specific cases and complex explanations
- Constraints describing the product's or its component's limitations
- Activities to describe problems resolution stages
- Rules to describe knowledge that directs the choices in the activities
- Entities to represent knowledge elements that describe the product, its components, its assemblies, parts and features. An entity can be structural or functional (see figure below).

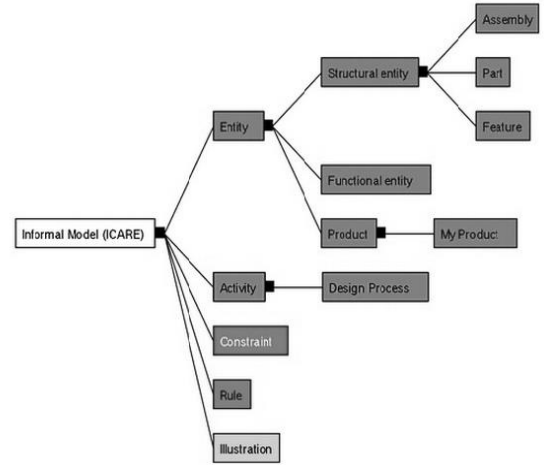


Fig. 3. Informal model of MOKA [12]

The ICARE form translates the meta product models and processes. In our case, we choose to propose a generic product design model by integrating the different common concepts of the product appearance and process. This model allows driving the design activity from a design situation previously formalized in order to evaluate support design. In this way, we will propose our generic product design model.

IV. GENERIC PRODUCT DESIGN MODEL

In this section, we propose a generic product design model to provide a unifying view of all the concepts involved in a design activities, our work is based on the meta model of the MOKA principle(see figure 4). This model is divided into four packages: design activities description, product, process, resources:

- The design activities description: it's our base case, contains all past design activities as problem description, and all the information about the design activity itself: start date, end date, structure, rules of calculations, dimensions, resources, etc.

- Product : covers five essential models(structure, function, behavior, constraint)
 - Structure: contains the all components of the product as the various geometries to achieve.
 - Function: identify different rules for the calculation and design that allows to sequence and control the execution of design tasks
 - Behavior: is the study of the behavior of products in its use phase
 - Constraint: define limitations Product
- Process : activities that represent a breakdown of the design process or a set of elementary operations
- Resources : describe the technology associated at each product

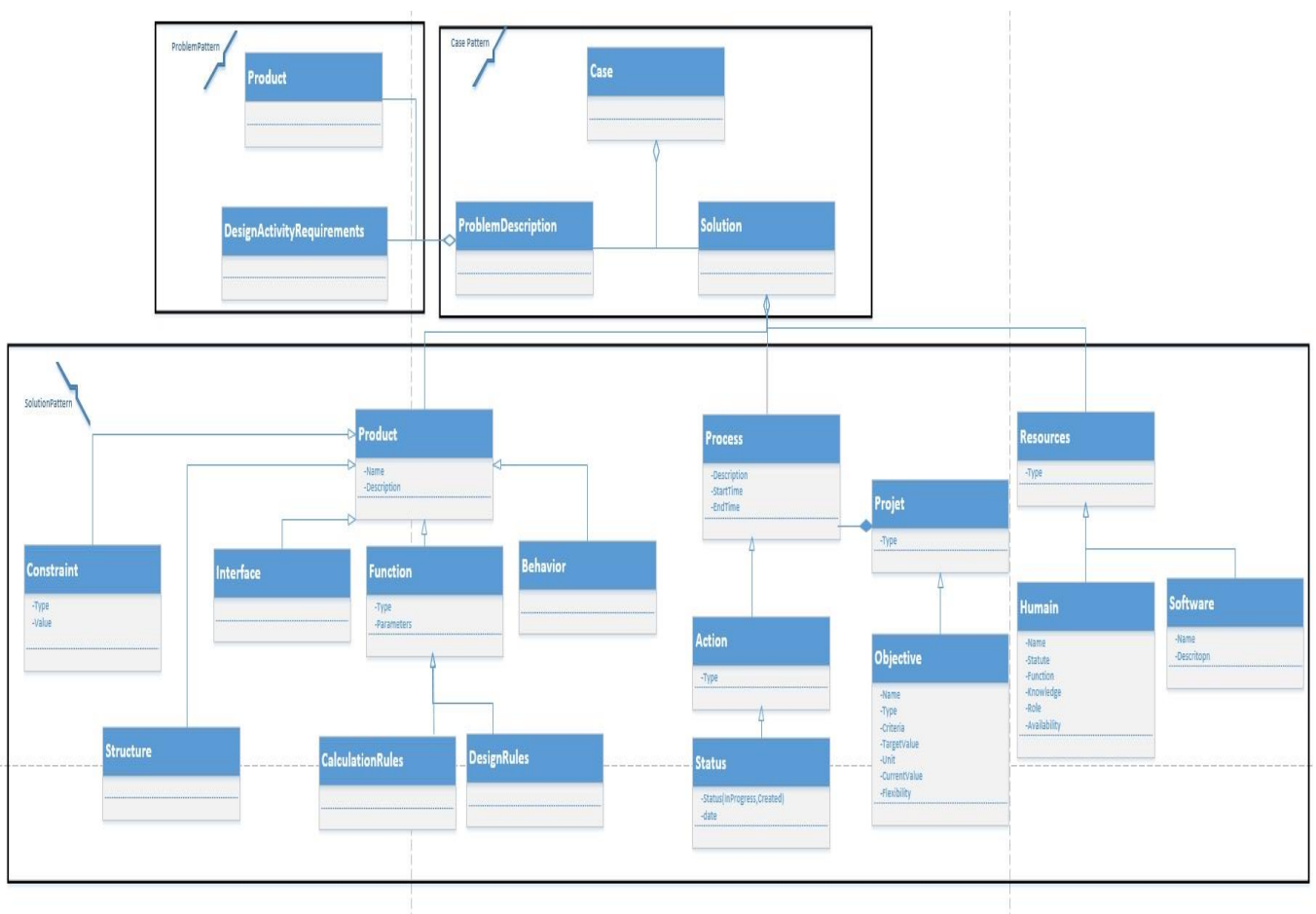


Fig. 4. Design meta-model

V. CBR-MINING :COMBINING CASE BASED REASONING AND PROCESS MINING

In this section, we present our approach of coupling CBR and process mining to provide a methodological framework to support design. The idea is to provide a model process design activity responding to a current design activity. In fact, this model process is found by analyzing all related design activities. For each current design activity, we calculate the degree of similarity compared to the process stored in the case base. We use the process mining to extract these similar processes and generate a model process. Moreover, to use CBR and to adapt process mining to our problem, we assume that:

- The knowledge base used is relevant, it covers a large number of process design and is filled automatically.
- Only effective solutions judged and considered as good practice, are stored.

This coupling aims to improve the performance of CBR. In our case, we focus only on the retrieving step to extract previously designed activities and on the reuse step in which we used process mining to discover a new process, merging past similar solutions. The figure 5 illustrates our approach

A. Similar Cases Retrieval:

In the CBR process, when a new design task comes, it is represented as a new case. We apply an algorithm to extract the best similar previous cases, based on a highest similarity degree factor. We used the nearest neighbor algorithm [14] which is a simple approach that calculates the similarity between previous stored cases based on features selected by the user.

B. Process mining:

In order to use the process mining, the processes selected will replace the event log and they must respect all the important parts of the event log structure. After we apply the alpha algorithm [15][16] of process mining which aims to discover a workflow from a set of execution traces of the same workflow. The alpha-algorithm is based on the workflow-net (WF-net) and the workflow-log(WF-log) to represent the workflow.

- A (WF-net) is a Petri net responding to structural constraints.
- A workflow-log (WF-log) is a sequence of events produced by the execution of the workflow. Each event is characterized by a type and each type is associated with a transition from (WF-net) that produced.

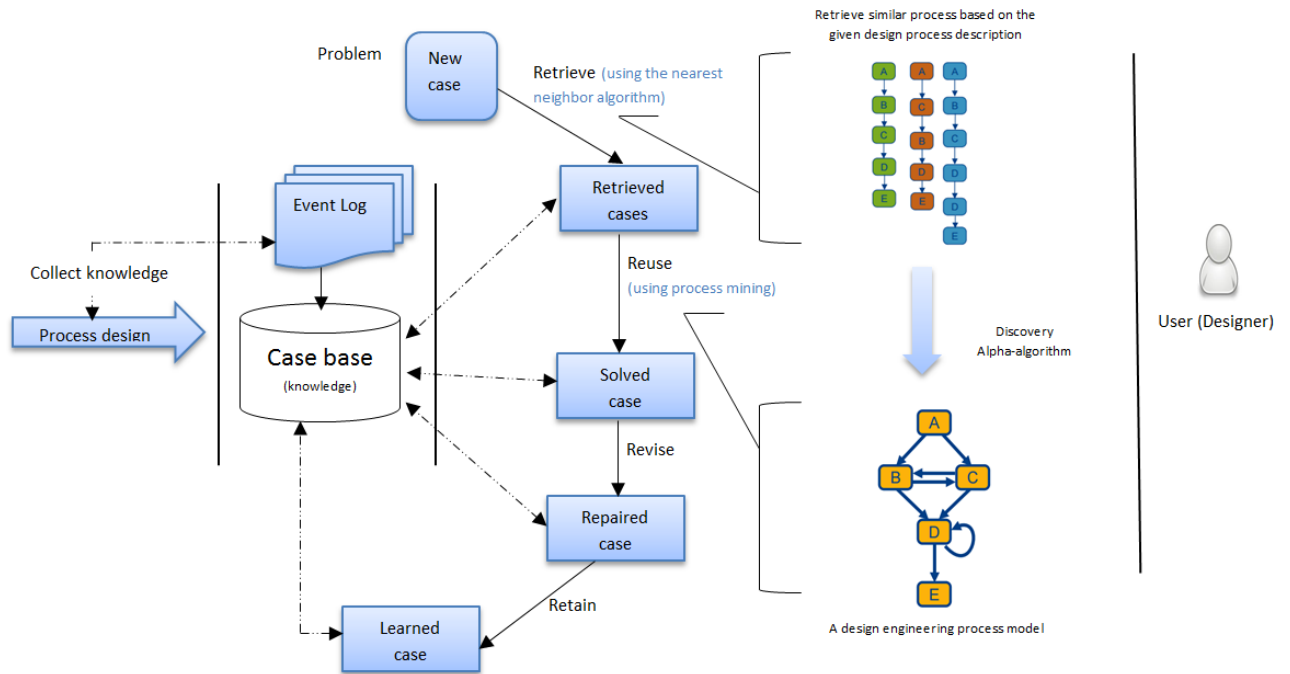


Fig. 5. CBR-Mining approach

VI. A CASE STUDY

In this section, a case study is discussed to illustrate the analyses that are possible and the process mining algorithms that are available in order to validate our approach. As we have already mentioned, we assume that our knowledge base covers a considerable number of design activities and it filled automatically. We choose the design activity of a “Carrosserie” process as use case. The aim of our approach is for the user to be able to recover all similar processes to his target case (Design activity) stored in his knowledge base using the nearest neighbor algorithm. The recovery of these processes can be expressed into two groups of factors: product type and process feature factor (design task process requirement). As can be seen in Figure 5, we recover all process similar to the activity design “carrosserie”

Case ID	Event ID	Actor	Product	dd-MM-yy	Role	Activity	Costs
Case 1	35654423	Fati	Carrosserie	30-12-201	Design	Create part par-choc arrier	5i
Case 1	35654424	Emmanuel	Carrosserie	31-12-201	Design	create part par-choc avant	40i
Case 1	35654425	Guillaume	Carrosserie	05-01-201	Design	create part porte	10i
Case 1	35654426	Remy	Carrosserie	06-01-201	assembly	assemblies of parts	20i
Case 1	35654427	Remy	Carrosserie	07-01-201	assembly	calculation assembly constraints	20i
Case 1	35654483	Remy	Carrosserie	30-12-201	assembly	verify assembly constraints	5i
Case 1	35654485	Mike	Carrosserie	30-12-201	Simulation	verify assembly	10i
Case 1	35654487	Mike	Carrosserie	30-12-201	Simulation	simulated	40i
Case 1	35654488	Samuel	Carrosserie	05-01-201	Manager	Validate	20i
Case 2	35654489	Fati	Carrosserie	08-01-201	Design	Create part par-choc arrier	20i
Case 2	35654521	Emmanuel	Carrosserie	30-12-201	Design	create part par-choc avant	5i
Case 2	35654522	Guillaume	Carrosserie	30-12-201	Design	create part porte	40i
Case 2	35654524	Remy	Carrosserie	30-12-201	assembly	assemblies of parts	10i
Case 2	35654525	Remy	Carrosserie	06-01-201	assembly	calculation assembly constraints	20i
Case 2	35654526	Remy	Carrosserie	06-01-201	assembly	verify assembly constraints	20i
Case 2	35654527	Mike	Carrosserie	06-01-201	Simulation	verify assembly	40i
Case 2	35654530	Mike	Carrosserie	08-01-201	Simulation	simulated	10i
Case 2	35654531	Samuel	Carrosserie	09-01-201	Manager	Notify modification design	20i
Case 3	35654533	Fati	Carrosserie	15-01-201	Design	Create part par-choc arrier	20i
Case 3	35654641	Emmanuel	Carrosserie	06-01-201	Design	create part par-choc avant	5i
Case 3	35654643	Guillaume	Carrosserie	07-01-201	Design	create part porte	10i
Case 3	35654644	Remy	Carrosserie	08-01-201	assembly	assemblies of parts	40i
Case 3	35654645	Remy	Carrosserie	09-01-201	assembly	calculation assembly constraints	20i
Case 3	35654647	Remy	Carrosserie	12-01-201	assembly	verify assembly constraints	20i
Case 3	35654711	Mike	Carrosserie	06-01-201	Simulation	verify assembly	5i
Case 3	35654712	Mike	Carrosserie	07-01-201	Simulation	simulated	40i
Case 3	35654714	Samuel	Carrosserie	08-01-201	Manager	Validate	10i
Case 4	35654715	Fati	Carrosserie	10-01-201	Design	Create part par-choc arrier	20i
Case 4	35654716	Emmanuel	Carrosserie	11-01-201	Design	create part par-choc avant	20i
Case 4	35654718	Guillaume	Carrosserie	14-01-201	Design	create part porte	10i
Case 4	35654719	Remy	Carrosserie	16-01-201	assembly	assemblies of parts	40i
Case 4	35654720	Remy	Carrosserie	19-01-201	assembly	calculation assembly constraints	20i
Case 4	35654721	Remy	Carrosserie	20-01-201	assembly	verify assembly constraints	20i

Fig. 6. An event log

In order to extract information from this event log file and to test the applicability of our approach, we choose to use ProM framework[11] which is an open source framework for process mining. Initially the data are in form of Excel file but using ProM Framework, the data was converted into MXML[15], a format that is commonly used in ProM which will be used as the input to the process mining algorithm. Therefore, in our use study we apply the mining algorithm alpha algorithm that we have explained above.

The figure 7 shows some results regarding the process design till the validation of the prototype. These results have been used for answering the need of the user (designer). In order to design products in a short time, the user analyzes the event log including all the solutions of similar past experiences to the activity design “carrosserie” and we deduce a model process presented in the following figure. This model presents the tasks that are executed in the process:

- *Design activities:* presents all the workflows Elements and the Event Log used in the design of our product.
- *Assembly activities:* Presents all the workflows Elements and the Event Log used in the assembly of the pieces.
- *Simulation activities:* presents all the workflows Elements and the Event Log used in the simulation of the virtual prototype before proceeding to manufacture.
- *Decide:* allows the project manager to validate project or to notify Modification before transmitting it to client.

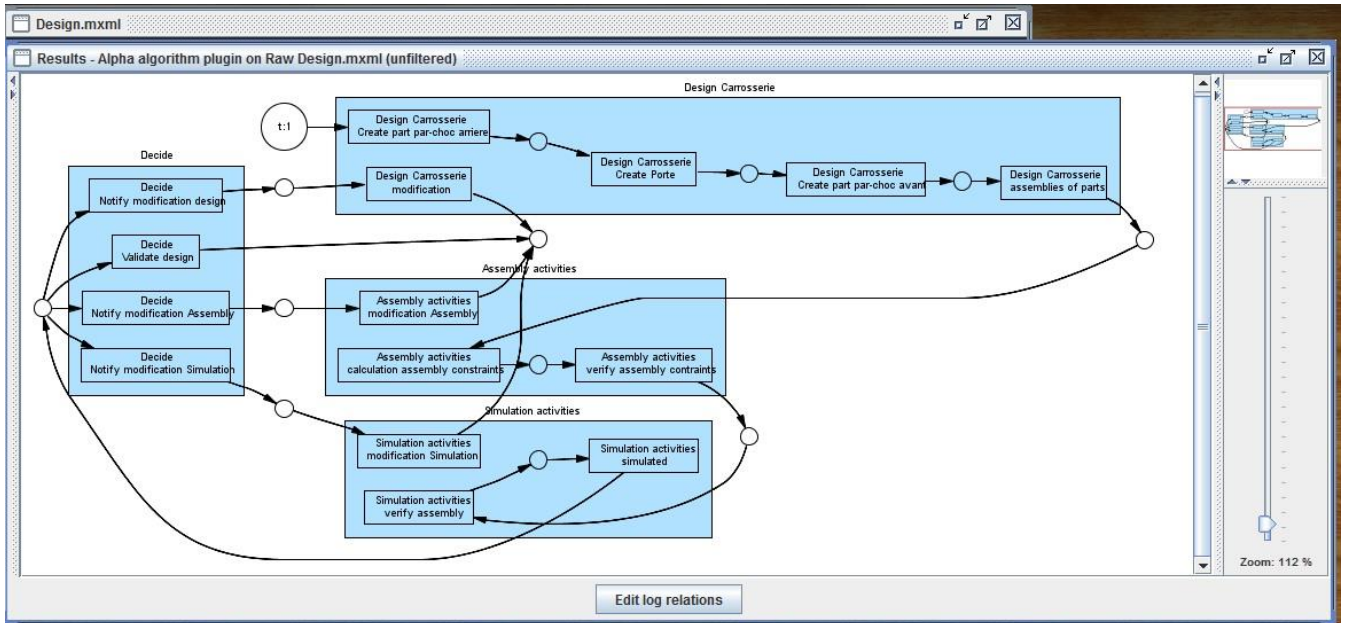


Fig. 7. Describing the design process using the process mining

In this paper, we presented CBR-Mining approach which aims to facilitate the technical design of industrial products. The proposed approach provides a domain case base with knowledge reuse capabilities. The objective is to support, optimize and improve engineered products realization and enhance software applications. The originality lies in the discovery of a process model based on past similar processes that are stored in a knowledge base. To validate our proposal, we have chosen the extraction of an automobile design process as a case study.

Our future work will focus on the development of a web application experience feedback to support the industrial design activities. We will integrate ProM tool and connect it to our knowledge base. The idea is that the ProM tool can import and operate directly log files stored in the knowledge base. Furthermore, to achieve semantic interoperability of data, we will transform MXML files imported by ProM to OWL format which is a language for knowledge representation.

This application design future work will focus on extending the already modelled domain knowledge and increase the capabilities of our Framework. An extension of the data types and their attributes in the underlying case-base will be performed. Further to that, the implementation of the knowledge enriched applications will be finalized and the framework will be deployed and tested in a real industrial pilot case under SOA (Service Oriented Architecture) in a Factory of the future equipments based on Internet of Things protocols connected to the Cloud.

ACKNOWLEDGMENT

This paper reflects only the authors' views. The work reported is carried out by the PLACIS project, which has received research funding from French Ministry for Higher Education and Research.

- [1] N.Arshadi and I.Jurisa, "Data Mining for Case-Based Reasoning in High-Dimensional Biological Domains". IEEE Transactions on Knowledge and Data Engineering, 1127-1137, 2005.
- [2] K.S. Kim,, and . I.G. Han "The cluster-indexing method for case-based reasoning using self-organizing maps and learning vector quantization for bond rating cases". Expert Systems with Applications, 21(3), 147-156. 2001.
- [3] G.R. Beddoe, S. Petrovic. "Selecting and weighting features using a genetic algorithm in a case-based reasoning approach to personnel rostering". European Journal of Operational Research 175649-671, 2006.
- [4] K. Anyanwu, A. Sheth, J. Cardoso, J. Miller, and K. Kochut. "Healthcare Enterprise Process Development and Integration". Journal of Research and Practice in Information Technology, 35(2):83-98, 2003.
- [5] P. Backus, M. Janakiram, S. Mowzoon, G.C. Runger, and A. Bhargava. "Factory cycle time prediction with a data-mining approach". IEEE Transactions on Semiconductor Manufacturing, vol. 19, no. 2, pp. 252-258, 2006.
- [6] W.M.P. van der Aalst, J. Nakatumba, A. Rozinat, and N. Russell , "Business Process Simulation: How to get it right?", in BPM Center Report BPM-08-07, BPM center. Org,2008.
- [7] E. Domeshek, and J. Kolodner., "Using the point of large cases". Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 7(2), 87-96, 1993.
- [8] A. Aamodt, "A knowledge representation system for integration of general and case-specific knowledge". In Proceedings from IEEE Tools with Artificial Intelligence (TAI), New Orleans, 5-12 November 1994, pp. 836-839, 1994.
- [9] I. Watson."Case-based reasoning is a methodology, not a technology".Knowledge based Systems,vol.12, p.303-308,1999.
- [10] G. Cortes Robles, "Management de l'innovation technologique et des connaissances : Synergie entre la théorie TRIZ et le Raisonnement à Partir de Cas, Application en génie des procédés et systèmes industriels". Thèse doctorale, Institut National Polytechnique de Toulouse,2006.
- [11] S. Mans, W. Aalst, M. P. Van Der, and H.M.W. Verbeek "Defining and Executing Process Mining Workflows with RapidProM'',
- [12] A. Bernard, "Déploiement d'une démarche de Knowledge Management dans le cadre d'un projet de spécification d'un outil KBE", June 2015
- [13] R .Klein. "Knowledge Modelling in Design – the MOKA framework, Proceedings of the International AI en Design", J.S.Gero (éd.), Kluwer, Worcester, MA, June 2000
- [14] J. Kolodner."Case-Based Reasoning", San Mateo, California:Morgan Kaufmann, 1993.
- [15] W.M.P. Van Der Aalst. Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer-Verlag, Berlin, 2011.
- [16] A. Medeiros, B. Van dongen, W. Van der aalst, A. Weijters. "Process mining : Extending the α -algorithm to mine short loops". Eindhoven University of Technology, Eindhoven,2004.