CENTRALE
PARIS

ÉCOLE CENTRALE
PARIS

THESE DE DOCTORAT
Spécialité Génie Industriel

Présenté par
Walid BEN AHMED

SUJET

SAFE-NEXT: A SYSTEMIC APPROACH FOR KNOWLEDGE DISCOVERY IN DATABASES.
APPLICATION IN ACCIDENT SCENARIO DEVELOPMENT AND INTERPRETATION

Rapporteurs :

Daniel COUTELLIER– Professeur E.N.S.I.M.E. Valenciennes/Directeur Général PRIMECA
Rose DIENG – Chef du projet ACACIA, INRIA Sophia Antipolis
Christian TAHON – Maître de conférence, Université de Valenciennes et du Hainaut

Examinateurs :

Michel TOLLENAERE – Professeur INPG, Grenoble
Claudia ECKERT – Professeur, Engineering Design Center, University of Cambridge
Yves PAGE – Adjoint du directeur du LAB PSA Peugeot-Citroën Renault
Directeur de thèse : Mounib MEKHILEF – HDR à l'Université d'Orléans
Co-directeur de thèse : Michel BIGAND – Maître de Conférence, Ecole Centrale Lille

2004

Contents

1 ABSTRACT

Nowadays, given the automation of data collection, very large databases are constructed. The exploitation of these data in accidentology and several others fields (e.g. marketing, engineering, etc.) requires automatic techniques of Knowledge Discovery in Databases (KDD). Incorporating expert knowledge in the KDD process is fundamental to handle the complexity of data, domain and knowledge. This necessitates the development of approaches, methods and techniques intended to identify, represent and operationalize expert knowledge.

In this dissertation, we propose a new approach, SAFE-Next (Systemic Approach For Enhanced kNowledge EXTraction), which integrates the following four approaches: the first one, ASMEC (Approche Systémique de ModElisation des Connaissances), allows knowledge modeling according to multiple viewpoints and granularity levels. The second approach, AICEF (Approche d'Incorporation des Connaissances Expertes dans la Fouille de données), uses the ASMEC knowledge model to elaborate multi-view metadata. It then uses these metadata as a tool for incorporating expert knowledge into the KDD process. The third approach, ASAIC (Approche Systémique d'Analyse d'Impact de Changement), uses the ASMEC knowledge model to carry out multi-view change impact analysis. The fourth approach, ASEM (Approche Systémique d'Evaluation de Modèles), provides an assessment framework for knowledge models.

The epistemological and methodological foundations of our work are constructivism and systemic approach (or cybernetics). Based on these backgrounds, our research contributions concern several disciplines, ranging from Accidentology, Knowledge Engineering, Knowledge Discovery in Databases and Design. In accidentology, SAFE-Next provides experts with an efficient tool for knowledge management. It enables the elaboration of multi-view accident scenarios, which are a powerful tool for understanding accident mechanisms in order to develop safety counter-measures. Furthermore, SAFE-Next provides a knowledge capitalization tool. In knowledge engineering, SAFE-Next supplies, via ASMEC, a multi-view knowledge model and thereby allows the integration of different viewpoints stemming from different users. Furthermore, it provides a multi-granularity knowledge model and in that way addresses the difficulty of knowledge identification and formalization. At the same time, SAFE-Next permits, via ASEM, the evaluation of knowledge models, an issue rarely addressed in literature. In Knowledge Discovery in Database, SAFE-Next enables, via AICEF, the incorporation of domain knowledge in the data preprocessing step (i.e. the first step in a KDD process) and more specifically in the attribute selection task. Likewise, the multi-view metadata enables the incorporation of domain knowledge in the interpretation step (i.e. the last step in a KDD process). In design, SAFE-Next provides safety system developers with an efficient tool to construct the design space. Scenarios enable them to understand complex behaviors and thereby to define solutions and alternatives. SAFE-Next also provides, via ASAIC, an approach for multi-view change impact analysis. Moreover, it proposes an extension of the change impact analysis to the use process of a given product as well as the evaluation process instead of limiting it to the design process.

2 INDUSTRIAL CONTEXT AND MAIN ISSUES

While considerable improvements in road safety have been made, the number of accidents remains high. An estimated 86992 road accidents occurred in France in 2003. It is known that more than 5000 citizens die every year (5400 in 2003). From an economic perspective, a financial loss of 24 billion Euros per year is estimated.

The two French car manufacturers, PSA (Peugeot-Citroën) and Renault, created the LAB (Laboratory of Accidentology, Biomechanics and Human Behavior) in 1969. Its main objective is to provide carmakers with the required accidentology knowledge in order to develop and improve safety systems. The LAB carries out in-depth studies on the scene of the accident to construct accident databases. Experts in the LAB use these databases in order to develop the required knowledge.

To analyze an accident, several approaches were used. Some of these studies focus on the accident's causal aspect. Others focus on the accident's sequential aspect [Brenac, 1997; Fleury et Brenac, 2001; Kurucz et al., 1977], or on the human mechanisms of error production and of information processing in accidents [Fuller et Santos, 2002; Van Elslande et Alberton, 1997]. Some studies in cognitive psychology analyze the driver's behavior as a process of skill learning and automatization [Summala, 2000], or as a management process of risks and task difficulties [Fuller, 2000]. Thus, each of these studies focuses on a given aspect of the accident. However, when considering the complexity of the accident behavior, several approaches should be combined in order to handle this complexity.

Accident Scenario is a prototypical behavior of a group of accidents having similarities. Accidentologists assume that similar accident factors entail similar safety countermeasures. Based on this assumption, accidentologists recognize accident scenario as a powerful tool to provide safety system developers with the knowledge required. However, developing accident scenario is a complex task due to the following issues: accident scenario elaboration requires expert knowledge, the knowledge required is often implicit and comes in different forms, there is a multiplicity of viewpoints (psychology, mechanics, etc.), the fact that expertise is time-consuming, different granularity levels and ways of representing accident scenarios, etc. Furthermore, accident scenario elaboration requires the combination of *clinical approaches* (experts) and *data mining* approaches, more specifically clustering.

To handle these issues, we developed SAFE-Next, a Systemic Approach For Enhanced kNowledge EXTraction. Our approach integrates four main approaches:

- ASMEC: A Knowledge Engineering Approach for Multi-Experts Knowledge Modeling and Acquisition: it allows the elaboration of a multi-view representation of domain knowledge, a multi-granularity level of domain knowledge. Moreover, it can be used to guide a multi-view elicitation in a knowledge acquisition task;
- AICEF: an Approach For the Incorporation of Domain Knowledge in a KDD process. It uses ASMEC domain knowledge in order to incorporate experts knowledge in the KDD process;

- ASAIC: a Systemic Approach for Change Impact analysis: it uses ASMEC domain knowledge and allows a multi-view analysis of a given change in road safety;
- ASEM: a Generic Systemic Approach for Model Assessment.

Figure 1 presents the different approaches we developed in SAFE-Next and the relationship between these approaches. ASMEC allows the elaboration of a domain knowledge model according to a multi-view and multi-granularity framework. We use XML language for the formalization of ASMEC model. So, our knowledge model is expressed through a *multi-view metadata.* AICEF, the second approach in SAFE-Next, uses these metadata in order to incorporate domain knowledge in KDD process. ASAIC, the third approach in SAFE-Next use the metadata in order to carry out multi-view change impact analysis in road safety. ASEM, the third approach in SAFE-Next allows the assessment of the different models.

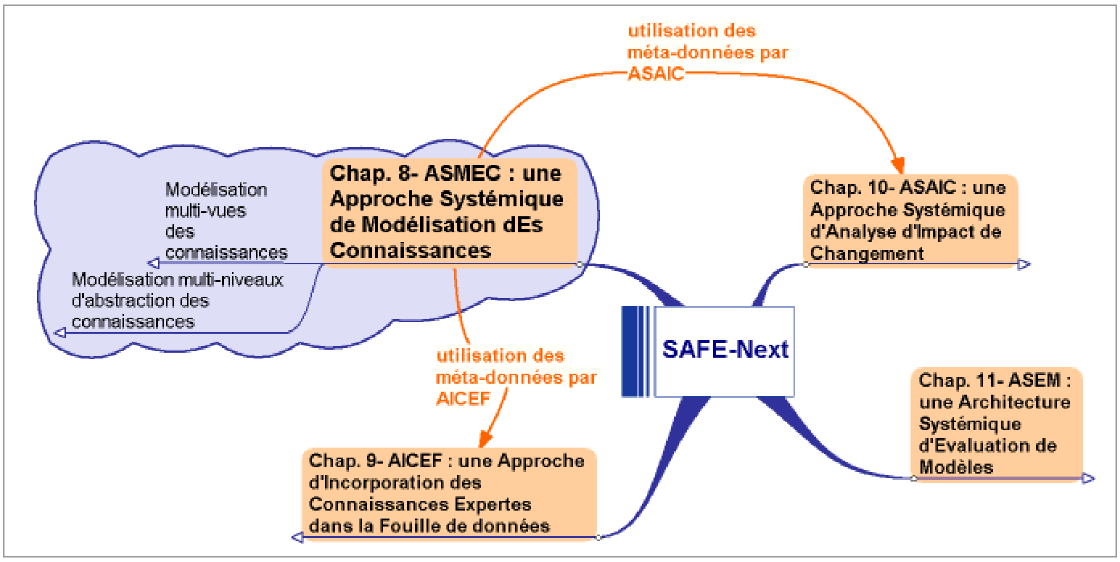


Figure 1- ASMEC : Systemic Approach for Multi-view Knowledge Modeling (the first approach in SAFE-Next)

[Brenac, 1997] T. Brenac. *L'analyse séquentielle de l'accident de la route: comment la mettre en pratique dans les diagnostics de sécurité routière, Outil et méthode,* Rapport de recherche n°3, INRETS, 1997.

[Fleury et Brenac, 2001] D. Fleury et T. Brenac. Accident prototypical scenarios, a tool for road safety research and diagnosis studies. *Accident Analysis & Prevention*, vol. 33, p. 267-276, 2001.

[Fuller, 2000] R. Fuller. The Task-Capability Interface Model Of The Driving Process. *RTS, Recherche Transports Sécurité*, vol. 66, Tome 1, p. 35-45, 2000.

[Fuller et Santos, 2002] R. Fuller et J. A. Santos. Human Factors For High-way Engineers, Elsevier, Pergamon, 2002.

[Kurucz *et al.*, 1977] Kurucz, Morrow, Fogarty, Janicek et Klapper. The Effectiveness of ABS in Real Life Accidents. *14th international technical conference on Enhancing Safety Vehicles*, Munich, Allemagne, 1977.

[Summala, 2000] H. Summala. Automatization, automation, and modeling of driver's behavior. *RTS, Recherche Trans-ports Sécurité*, vol. 66, Tome 1, p. 35-45, 2000.

[Van Elslande et Alberton, 1997] P. Van Elslande et L. Alberton. *Scénarios-types de production de l'erreur humaine dans l'accident de la route, problématique et analyse qualitative,* Rapport de recherche N°218, INRETS, 1997.

3 THEORETICAL FOUNDATIONS OF OUR RESEARCH: CYBERNETICS AND GENERAL SYSTEM THEORY

The epistemological foundation of our work is constructivism, which assumes that knowledge depends on how the individual "constructs" meaning from his or her experience. A system, in a constructivist perspective, is recognized as a representation of reality seen by some people in a given context. Indeed, we assume that each expert in accidentology has an individual perception of the same behavior (i.e. road accidents in our case). Our study is intended to identify these different viewpoints in order to construct a multi-view domain model. Constructivism is opposed to positivism, which assumes that science must be concerned only with the things of which we have a direct experience. According to positivism, science must limit itself to what is observable and measurable ("empiricism"). Systems, in a positivist perspective, are seen as existing and as real entities, with universally valid features, which can be identified and studied as such. This is sometimes called a "hard" perspective.

3.1 Cybernetics & General System theory

The methodological foundation of our research is cybernetic [Von Foerster, 1995; Wiener, 1948], which is a modeling approach stemming from the constructivist epistemology. Cybernetics is the science that studies the abstract principles of organization in complex systems. It is concerned not so much with what systems consist of, but how they function. Cybernetics perceives *system* (observed object) as an agent interacting with another *agent* (observer) in a given *context*. Observer and observed cannot be separated, and the result of observations (i.e. models, representation, analysis, etc.) will depend on their interaction [Heylighen et Joslyn, 2001].

Hence, there are usually three distinctions to make in order to model a system in a cybernetic perspective:

(a) The **observed object**: this viewpoint is related to the following aspects: the structure, behavior, operation, or other characteristics of a real-world process, concept, or system. In our case, the object observed is Driver-Vehicle-Environment (DVE) system;

(b) The **observer**: the person who performs the observation and perception of the studied system in order to model and/or to analyze it. The perception may be different according to different observers. In our case, the observers are experts in accidentology (called henceforth accidentologists). Each of these experts has, according to his perception (viewpoint), at least one model of the observed object;

(c) The **context of observation**: it consists of the context in which the observer and the observed interact. The context is characterized by the aim of the study and the environment (different agents implicated, their relationship etc.) in which the study is carried out etc. In our case, the environment includes especially safety system designers since the aim of our study is to develop a shared representation of accidents between accidentologists and designers. Designers use these representations to define new safety counter-measures and/or to assess safety systems.

General Systems Theory (GST) is the second methodological background on which we based our work. GST was founded at about the same time as cybernetic by Ludwig von Bertalanffy [Bertalanffy, 1969]. Its history dates back to the 1940's and 1950's. Some of the founding fathers of the system theory are von Bertalanffy, Wiener and Ashby. In 1950s, the two schools (Cybernetics and GST) combine to build a unified science called *systemic approach*.

The systemic approach is then the transdisciplinary study of the abstract organization of phenomena, independently of their substance, type, spatial or temporal scale of existence [Le Moigne, 1999]. It touches virtually all the traditional disciplines, from mathematics, technology, and biology to philosophy and social sciences.

According to the systemic approach, systems can be classified in two categories: complicated systems and complex systems [Le Moigne, 1999; Morin et Le Moigne, 1999]. Complicated systems are characterized by a behavior that can be predicted by analyzing the interactions between components. Complicated systems are deterministic systems (e.g. a computer is a complicated system). Complex systems are systems for which the behavior cannot be predicted by such an analysis. Complex systems are non-deterministic.

3.2 The application of the systemic approach on road accident modeling

Behavior in road accidents is complex, within a complex system that consists of the triptych Driver-Vehicle-Environment. It is not the number of components involved in the accident occurrence, neither the number of variables interacting during the accident, nor the number of their interactions, that make the accident complex. But, most of all, it is the impossibility to predict the DVE system behavior that entails this complexity. This unpredictability is notably due to the fact that human actions are strongly involved in accident causation, and that human behavior is unpredictable. Furthermore, during the road accident, the DVE system performs some functions (i.e. perception, interpretation, anticipation, decision, action), which generate transformations (i.e. new situation, new interpretation, new purpose, new requirement, etc.), which in turn generate new functions and behaviors, etc. According to Morin and Le Moigne (1999), one can describe complex behavior through feedback and recursive loops. Feedback loops can be either negative, that is moving towards stability, or positive, that is moving towards amplification. Recursive loops are responsible for the auto-production and auto-organization of the system [Morin et Le Moigne, 1999]. Due to these feedback and recursive loops, the DVE behavior cannot be reproduced, and it is also impossible to identify with exhaustiveness and certainty all the failures and dysfunction mechanisms at stake.

According to Miller's definition of a living system [Miller, 1995], the DVE is an open and living system as much as each component (i.e. driver, vehicle, infrastructure, traffic, etc.) is constantly interacting with its environment by means of information and matter-energy exchanges.

The systemic approach is opposed to reductionism and assumes that real systems are interacting with their environments and that they can acquire qualitatively new characteristics through emergence and continual evolution [Bertalanffy, 1969]. Rather than reducing an entity (e.g. the human body) to the features of its parts or elements (e.g. organs or cells), systemic focuses on the relationship between the parts, which connect them into a whole (holism). In other words, the systemic approach assumes that to handle a complex behavior, it is fundamental to make the junction between the following viewpoints:

- The ontological viewpoint (i.e. what is the system?): it allows a structure-oriented and contextual analysis of the system. In other words, it represents the sub-systems (the driver, infrastructure, traffic, ambient conditions, vehicle, etc.), their taxonomic groups, their contexts (the driver's professional status, family status, etc.), their structures, as well as the various interactions between these sub-systems and their components;

- The functional viewpoint (i.e. what does the system do?): it allows a function-oriented analysis of the system. It represents the global process of the DVE functioning during the road accident, which combines several procedures (perception, diagnostic, prognostic, decision and action) (Van Elslande et al., 1997);

- The transformational (or evolutionary) viewpoint (i.e. how does the system evolve? What does it become?): it allows a transformation-oriented analysis of the system. The DVE system behavior can be described as an evolution that goes through several states. The transformational viewpoint integrates the accident's sequential and causal models developed by the INRETS and described in the next section (Brenac, 1997; Fleury et al., 2001);

- The teleological (or intentional) viewpoint (i.e. what is the goal or intention of the system?): it allows a goal-directed analysis of the accident. In other words, it assumes that each of the DVE system components or functions has to serve a purpose in an active context in order to ensure the safety of the DVE system.

[Bertalanffy, 1969] L. v. Bertalanffy. General system theory: foundations, development, applications, George Braziller, New York, 1969.

[Heylighen et Joslyn, 2001] F. Heylighen et C. Joslyn. Cybernetics and Second Order Cybernetics. in *R.A. Meyers (ed.), Encyclopedia of Physical Science & Technology (3rd ed.)*, p. 155-177, Academic Press, New York, 2001.

[Le Moigne, 1999] J.-L. Le Moigne. La modélisation des systèmes complexes, Dunod, 1999.

[Miller, 1995] J. G. Miller. Living Systems, University Press of Colorado, 1995.

[Morin et Le Moigne, 1999] E. Morin et J. L. Le Moigne. L'intelligence de la complexité, L'Harmattan, Paris, 1999.

[Von Foerster, 1995] H. Von Foerster. The Cybernetics of Cybernetics (2nd edition), Future Systems Inc., Minneapolis, 1995.

[Wiener, 1948] N. Wiener. Cybernetics or control and communication in the animal and the machine, 2nd Edition, MIT Press, Cambridge, MA, 1948.

4 ASMEC: A KNOWLEDGE ENGINEERING APPROACH FOR MULTI-EXPERTS KNOWLEDGE MODELING AND ACQUISITION

Knowledge engineering is a field within artificial intelligence that develops Knowledge-Based Systems (KBS). Such systems are computer programs that contain large amounts of knowledge, rules and reasoning mechanisms to provide solutions to real-world problems.

4.1 General Issue:

Knowledge Acquisition and Representation are the bases of KBS building. However, knowledge engineers have to address many issues when they are performing these tasks. These issues are related to: the implicit nature of knowledge, different types of knowledge, the multidisciplinarity of domains, the multiplicity of viewpoints, the fact that expertise is time-consuming, different granularity levels and choices possible when representing knowledge, etc.

4.2 Research Objective:

The research aims to develop a knowledge engineering approach that allows:

- A multi-view representation of domain knowledge;
- A multi-granularity level of domain knowledge;
- A guiding of the elicitation step in a knowledge acquisition task.

4.3 Research Method & Result

The epistemological foundation of our research is constructivism. The methodology we apply is cybernetics and general system theory, which are complex system modeling approaches. Based on these backgrounds, we developed ASMEC (Approche Systémique de ModElisation des Connaissances) or Systemic Approach for Knowledge Modeling. ASMEC provides knowledge engineers with a Top-Down approach for knowledge acquisition and representation according to different generic viewpoints and different granularity levels.

4.3.1 Meta-model level

The first granularity level in ASMEC, ***the meta-model***, consists of two layers: the fist consists of three generic cybernetic viewpoints: the *observed object*, the *observer* and the *context of observation*. In the second layer, each of the first layer's components is composed of four generic systemic axes: *ontological*, *functional*, *transformational* and *teleological* (or *goal oriented*) viewpoints (see Figure 2 for an UML representation of the meta-model level).

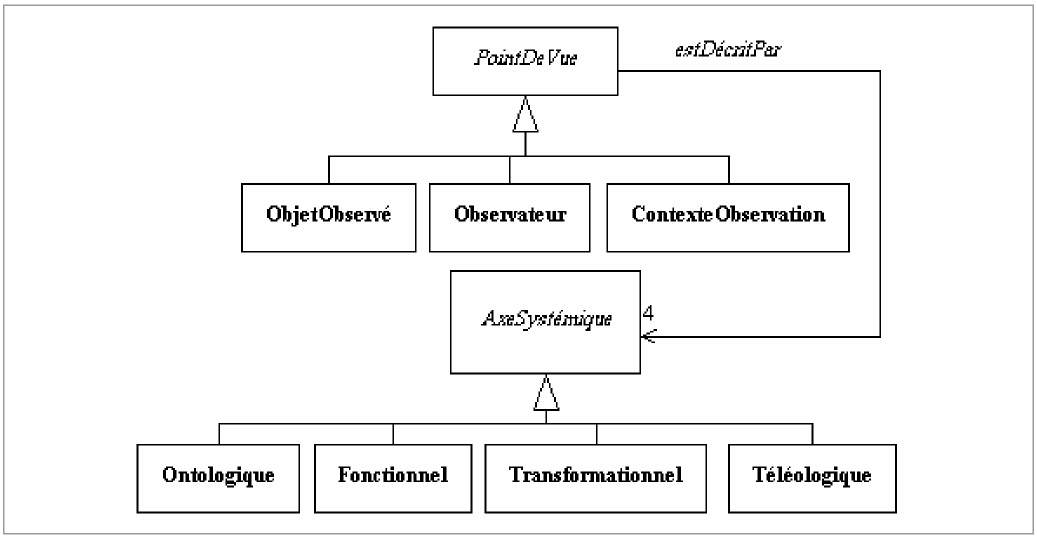


Figure 2- UML model of the meta-model level in ASMEC: four generic systemic views instrument each cybernetic viewpoint.

4.3.2 Model level

The second granularity level in ASMEC consists of ***domain models***. Each systemic view is composed of domain models. Let us take as an example the "observed object" view, which represents in our case the Driver-Vehicle-Environment system. We used the "*DVE model*", the "safety system model" and the "vehicle safety oriented model" to characterize the ontological view. We used the "sequential model" to characterize the transformational view. We used the "information processing model" to characterize the functional view and we used the "driving task model" to characterize the teleological view (see Figure 3)

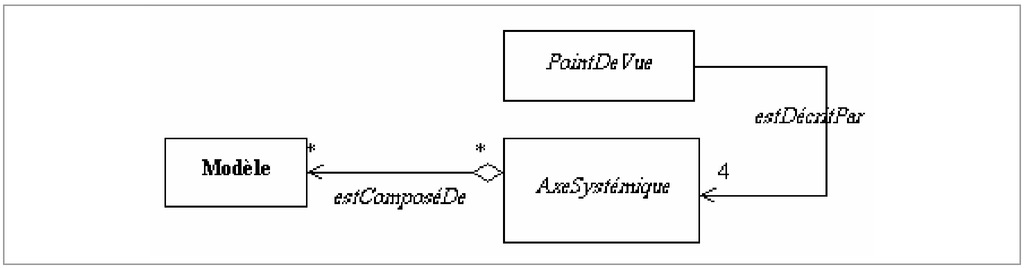


Figure 4- Using domain models for the instrumentation of ASMEC meta-model (i.e. 1st granularity level in ASMEC).

The choice of the different model takes into account the other two cybernetic viewpoints. For instance, the instrumentation of the *observed object* (i.e. DVE system), takes into account the *observer* (accidentologist in our case) viewpoint and *the context of observation* (designer in our case) viewpoint. Figure 5 illustrates this aspect:

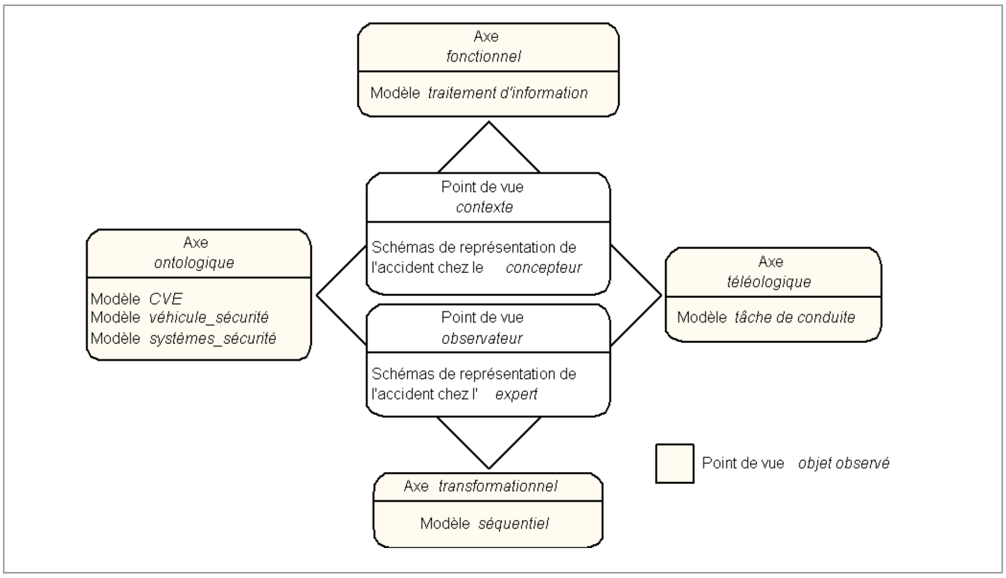


Figure 5- The development of the « observed object » viewpoint takes into account the other two cybernetic viewpoints (i.e. « observer » and « context of observation ».

4.3.3 Domain concepts

The third granularity level in ASMEC consists of ***domain concepts***. Domain concepts characterize each model in the previous level.

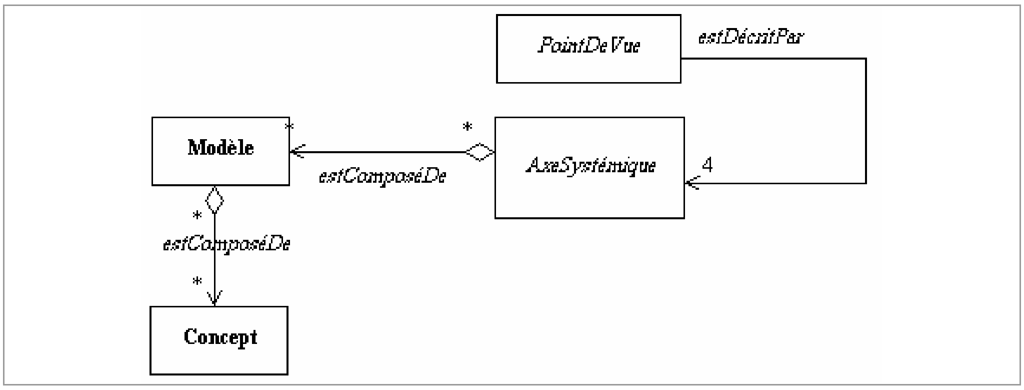


Figure 6- Using domain concepts for the characterization of domain models (i.e. 2nd granularity level in ASMEC).

Let us take an example: the following concepts: normal driving situation, accident situation, emergency situation and crash situation characterize the sequential model. “Perception”, “diagnosis”, “prognosis”, “decision” and “action” characterize the “information-processing model”.

4.3.4 Attribute level

The fourth granularity level in ASMEC consists of ***attributes***. A set of attributes characterizes each domain concept.

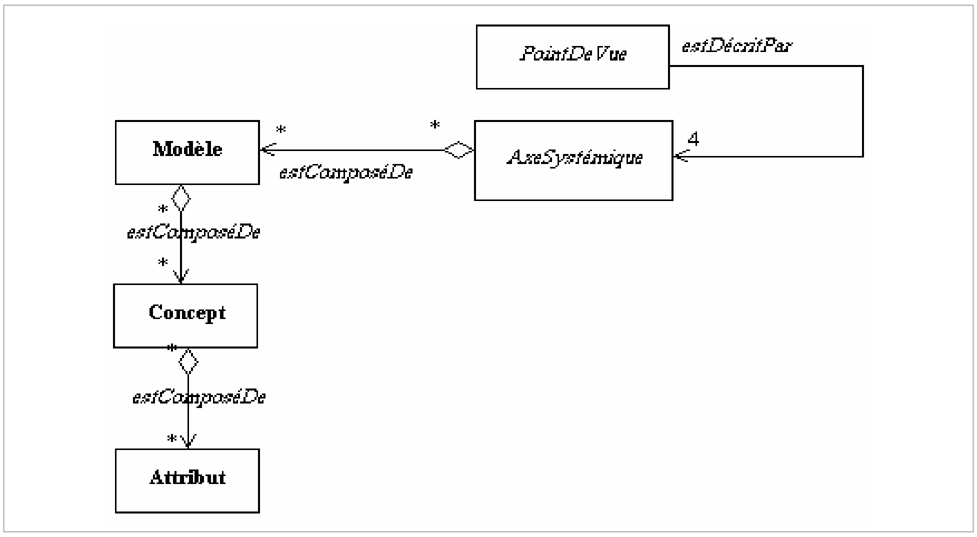


Figure 7- Using attributes for the characterization of domain concepts (i.e. 3rd granularity level in ASMEC).

Experts carry out the classification of attributes according to concepts and models.

4.4 Industrial Application

A knowledge base in accidentology is crucial to develop and assess safety countermeasures. However, there are many issues that have to be dealt with: the complexity of the road accident, the implicit nature of knowledge, the multi-disciplinary nature of accidentology (psychology, biomechanics, mechanics, ergonomics, design etc.), the multiplicity of expert viewpoints (functional, structural etc.) etc. To overcome these issues, we applied ASMEC. It allowed us to carry out a multi-view knowledge acquisition and thereby to elaborate a knowledge model in accidentology shared by vehicle and infrastructure experts, psychologists and safety system designers.

Two expert teams carried out the implementation of ASMEC in accidentology. Each team is composed of a psychologist, a vehicle expert and an infrastructure expert. The implementation consisted of the classification of attributes characterizing an accident (947 attributes are used in the LAB) according to the different viewpoints of ASMEC. Each attribute is assigned to one or more concept, each concept is assigned to one or more model, and each model is assigned to one or more systemic axis, which is related to the cybernetic viewpoints. To a certain extent, the implementation of ASMEC can be perceived as the elaboration of a ***metadata*** since the classification of these attributes according to different viewpoints and granularity levels is an elaboration of "*data about data*". We used XML to represent these metadata.

This study is funded by:

- The LAB (Laboratory of Accidentology, Biomechanics and Human Behavior), it is a common laboratory of the two French car manufacturers PSA Peugeot-Citroën and Renault
- The French Research Ministry.

5 AICEF: AN EFFICIENT APPROACH FOR DOMAIN KNOWLEDGE INCORPORATION IN KDD PROCESS

Nowadays, given the fast growth of large databases in business management, government administration, engineering and many other fields, Knowledge Discovery in Databases is a crucial issue. Over the last 20 years, several methods and techniques have been developed in order to automate knowledge extraction from these large databases. The KDD process is defined in [Fayyad et al., 1996] as *the nontrivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in data*.

5.1 General Issue

One can represent a KDD process as a sequence of three main steps: *data preparation* (or *pre-processing*), *data mining* and *result interpretation*. Expert knowledge incorporation in the first and the last steps is crucial to extract valid, useful, and understandable patterns in data.

Domain (or expert) knowledge is fundamental in addressing the following issues in the data preparation step: integrate data from different sources, complete missing data, handle redundancies, reduce the number of attributes, select relevant attributes according to different viewpoints, construct new attributes etc.

Domain knowledge is also essential in the interpretation step in order to make the extracted patterns useful, especially in a multi-user context. Two different experts may give a different interpretation of the same result stemming from a given data mining technique (tables, rules, features, graphs, etc.). Moreover, without expert knowledge, data mining results may be useless even if we use an efficient data mining technique. Therefore, incorporating domain knowledge is fundamental.

5.2 Research Objective

The research aims to develop a knowledge engineering approach allowing the:

- Identification and representation of domain knowledge;
- Incorporation of domain knowledge to allow a multi-view selection of attributes (i.e. the 1st step of the KDD process) that will be used in the data mining step (i.e. the 2nd step in the KDD process);
- Incorporation of domain knowledge to allow a multi-view interpretation of data mining results (classes, rules etc.) (i.e. the last step in the KDD process).

5.3 Research Method & Result

AICEF (Approche d'Incoporation de Connaissances en Fouille de Données or Approach For the Incorporation of Domain Knowledge in a KDD process) is the second approach we developed in SAFE-Next (see Figure 8).

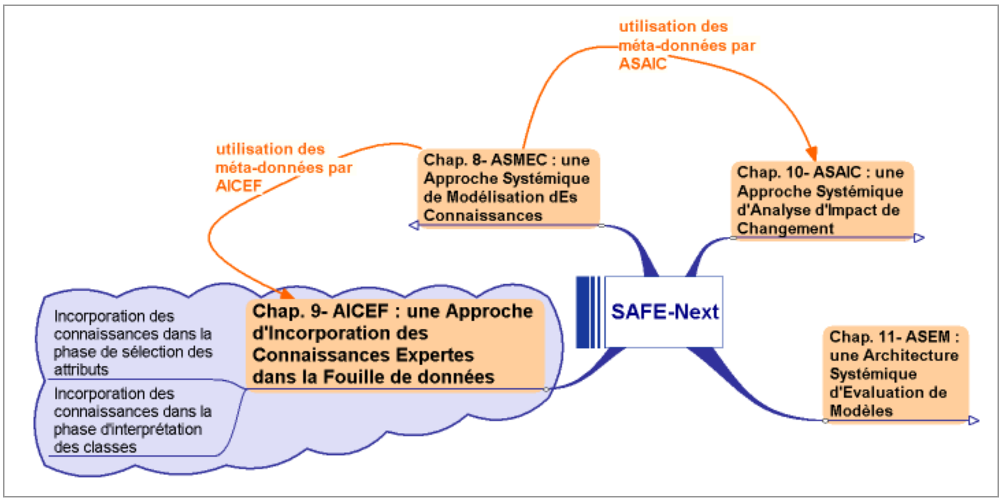


Figure 8- AICEF : an approach for the incorporation of domain knowledge in the KDD process (AICEF is the second approach in SAFE-Next, it uses the meta data stemmed from ASMEC)

AICEF uses the domain knowledge model (represented in metadata form) stemmed from ASMEC in order to incorporate domain knowledge in the KDD process. AICEF performs the link between ASMEC and accident database through the matching of *domain attributes* in ASMEC and columns in the studied database.

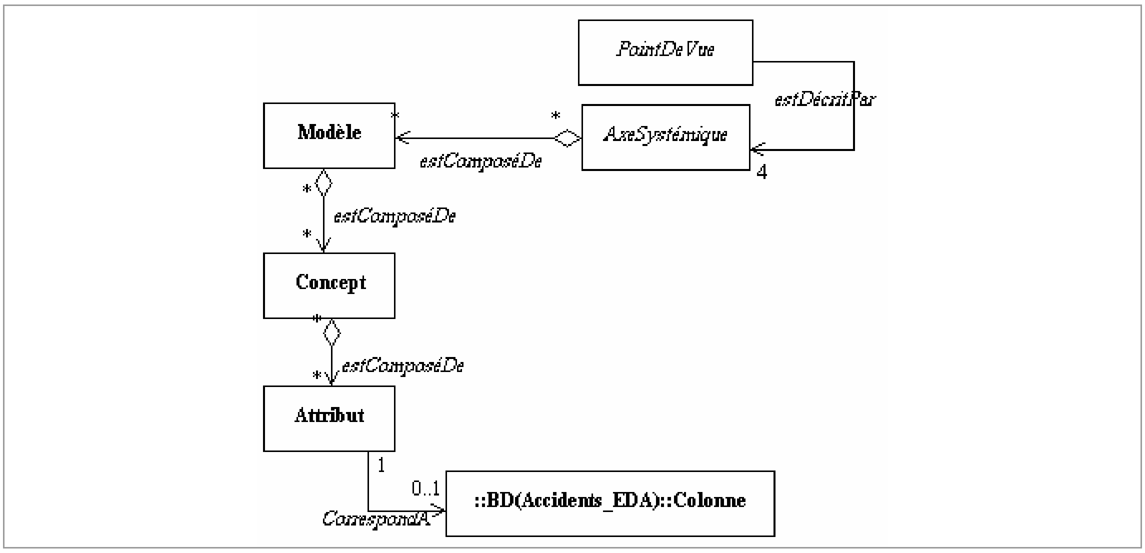


Figure 9- The link between ASMEC and the database: attributes in ASMEC correspond to columns in the studied database.

On the basis of the correspondence represented in Figure 9, AICEF enables the selection of attributes (1st step in a KDD process) according to the following viewpoints: *observed object*, *observer*, *context of observation*, *ontological*, *functional*, *transformational* and *teleological*. It also enables the selection of attributes according to the different domain models identified and formalized in ASMEC and according to the different concepts characterizing the different models. Hence, it allows a multi-view and a multi-granularity level selection of attributes. We use XML to represent the metadata (i.e. the classification of attributes according to different viewpoints in ASMEC). Then, we elaborate an interface, which allows the analyst to express his objective according to the different viewpoints and granularity levels. This allows him to select attributes for a given study according to the *metadata*.

We also use the same metadata in the interpretation step of the KDD process. In a clustering task, AICEF allows the representation of each cluster according to the different viewpoints implemented in the metadata. This multi-view representation is performed by the matching of domain attributes in ASMEC and attributes used in the clustering task (see Figure 10).

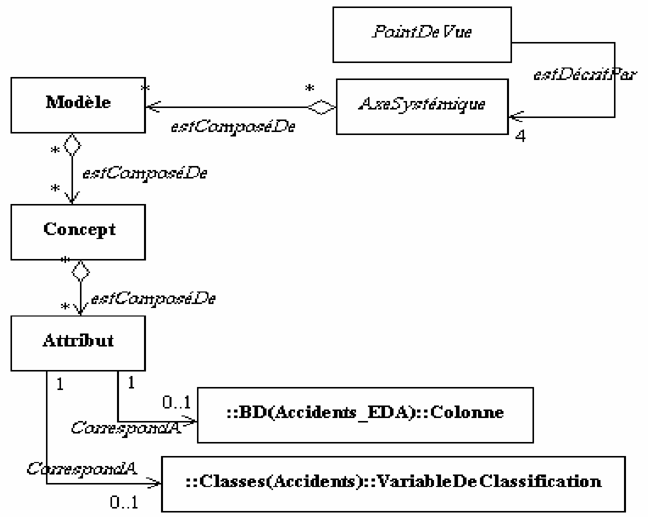
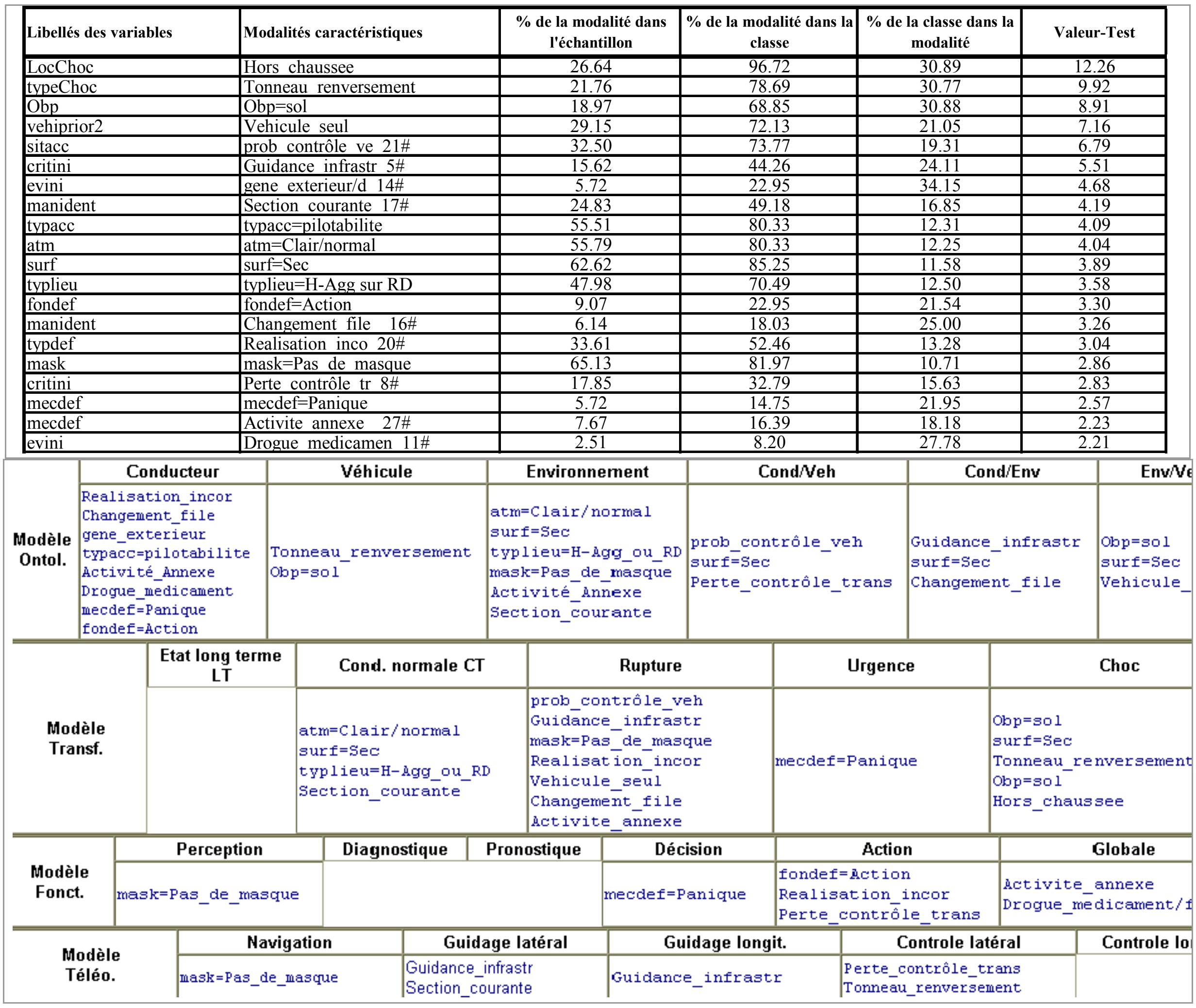


Figure 10- The link between ASMEC and the clustering results: attributes in ASMEC correspond to attributes used in the clustering task.

5.4 Industrial Application

We applied AICEF to carry out a multi-view selection of attributes in order to elaborate accident clusters. AICEF allowed us to perform a multi-view selection of attributes used in the clustering task. Moreover, AICEF allows a reduction in the number of attributes and the time required for the selection task. We compared results with previous studies carried out without using AICEF. We pointed out the improvement of the "quality" of selection in the sense that different viewpoints were taken into account.

We also applied AICEF in the interpretation step. This allowed us to interpret each accident cluster according to the different viewpoints, models and concepts formalized in ASMEC. These models and concepts may be related to different fields (design, psychology, mechanics etc.) and users. AICEF thereby makes the clusters usable by different agents stemming from different disciplines. The same object (i.e. cluster in this case) can be analyzed and interpreted from different viewpoints. Figure 11 illustrates the multi-view presentation of each cluster: the first table contains attributes characterizing one cluster. The second table contains the projections of the same cluster according to different models.



Libellés des variables	Modalités caractéristiques	% de la modalité dans l'échantillon	% de la modalité dans la classe	% de la classe dans la modalité	Valeur-Test
LocChoc	Hors chaussee	26.64	96.72	30.89	12.26
typeChoc	Tonneau renversement	21.76	78.69	30.77	9.92
Obp	Obp=sol	18.97	68.85	30.88	8.91
vehiprior2	Vehicule seul	29.15	72.13	21.05	7.16
sitacc	prob contrôle ve 21#	32.50	73.77	19.31	6.79
critini	Guidance infrastr 5#	15.62	44.26	24.11	5.51
evini	gene exterieur/d 14#	5.72	22.95	34.15	4.68
manident	Section courante 17#	24.83	49.18	16.85	4.19
typacc	typacc=pilotabilite	55.51	80.33	12.31	4.09
atm	atm=Clair/normal	55.79	80.33	12.25	4.04
surf	surf=Sec	62.62	85.25	11.58	3.89
typlieu	typlieu=H-Agg sur RD	47.98	70.49	12.50	3.58
fondef	fondef=Action	9.07	22.95	21.54	3.30
manident	Changement file 16#	6.14	18.03	25.00	3.26
typdef	Realisation inco 20#	33.61	52.46	13.28	3.04
mask	mask=Pas de masque	65.13	81.97	10.71	2.86
critini	Perte contrôle tr 8#	17.85	32.79	15.63	2.83
mecdef	mecdef=Panique	5.72	14.75	21.95	2.57
mecdef	Activite annexe 27#	7.67	16.39	18.18	2.23
evini	Drogue medicamen 11#	2.51	8.20	27.78	2.21

	Conducteur	Véhicule	Environnement	Cond/Veh	Cond/Env	Env/Ve
Modèle Ontol.	Realisation_incor Changement_file gene_exterieur typacc=pilotabilite Activité_Annexe Drogue_medicament mecdef=Panique fondef=Action	Tonneau_renversement Obp=sol	atm=Clair/normal surf=Sec typlieu=H-Agg_ou_RD mask=Pas_de_masque Activité_Annexe Section_courante	prob_contrôle_veh surf=Sec Perte_contrôle_trans	Guidance_infrastr surf=Sec Changement_file	Obp=sol surf=Sec Vehicule_

	Etat long terme LT	Cond. normale CT	Rupture	Urgence	Choc
Modèle Transf.		atm=Clair/normal surf=Sec typlieu=H-Agg_ou_RD Section_courante	prob_contrôle_veh Guidance_infrastr mask=Pas_de_masque Realisation_incor Vehicule_seul Changement_file Activite_annexe	mecdef=Panique	Obp=sol surf=Sec Tonneau_renversement Obp=sol Hors_chaussee

	Perception	Diagnostique	Pronostique	Décision	Action	Globale
Modèle Fonct.	mask=Pas_de_masque			mecdef=Panique	fondef=Action Realisation_incor Perte_contrôle_trans	Activite_annexe Drogue_medicament/f

	Navigation	Guidage latéral	Guidage longit.	Controle latéral	Controle lo
Modèle Téléo.	mask=Pas_de_masque	Guidance_infrastr Section_courante	Guidance_infrastr	Perte_contrôle_trans Tonneau_renversement	

Figure 11- A multi-view projection of clusters.

Figure 11 makes accident clusters more useful for both accidentologists and safety system designers. In other words, these multi-view representations of clusters (or scenarios) provide accidentologists and designers with a shared language. This enhances their collaboration in the definition of new safety systems.

This study is funded by:

- The LAB (Laboratory of Accidentology, Biomechanics and Human Behavior), it is a common laboratory of the two French car manufacturers PSA Peugeot-Citroën and Renault
- The French Research Ministry.

[Fayyad *et al.*, 1996] U. Fayyad, G. Piatetsky-Shapiro et P. Smyth. The KDD Process for Extracting Useful Knowledge from Volumes of Data. *Communications Of The ACM*, vol. 39(11), 1996.

6 ASAIC: AN APPROACH FOR A MULTI-VIEW CHANGE IMPACT ANALYSIS

Due to the increasing complexity of the modern industrial context in an evolutionary environment, the Engineering Change Impact Analysis (ECIA) approach has emerged as the key to enable organizations (e.g. companies, production systems, etc.) to anticipate their environmental evolution and therefore to survive. In this complex environment, changes can occur not only in high tech engineering environments, but also in any environment affected by technological advances, new technology emergence, human errors, fluctuations in the availability or the cost of specific components and raw materials and changing consumer demand. Given this situation, it is crucial for a company to have ECIA tools to respond quickly and effectively.

6.1 General Issue

The first issue we are addressing here is related to the fact that, in complex domains such as accidentology, several viewpoints have to be taken into account in order to perform a change impact analysis task.

The second issue is related to the fact that most research in literature has been focused on either product models [Murdock et al., 1997], process models [Blessing, 1996] or product/process/resource integrated models [Ma et al., 2003]. Nevertheless, these different models seem insufficient to handle an ECIA task especially in a complex context. They do not take into account all the life cycle steps especially the ***use process*** and the ***evaluation process*** of the developed system.

6.2 Research Objective

The research aims to elaborate a Knowledge-Based System (KBS) allowing a multi-view analysis of a given change impact. This KBS can be used as a communicating interface between designers and domain experts to handle an Engineering Change Impact Analysis task. We also intend to propose a generic framework for extending change impact analysis to the use and evaluation processes instead of limiting it to the design process.

6.3 Research Method & Result

Analyzing a change impact on complex system requires a domain knowledge base. Approaches such as ASMEC can provide such a knowledge base. We developed ASAIC (Approche Systémique d'Analyse d'Impact de Changement or Systemic Approach For A Multi-view Change Impact Analysis), which uses the domain knowledge model (formalized in metadata) stemmed from ASMEC in order to carry out a multi-view ECIA task. ASAIC is the third approach we developed in SAFE-Next (see Figure 12).

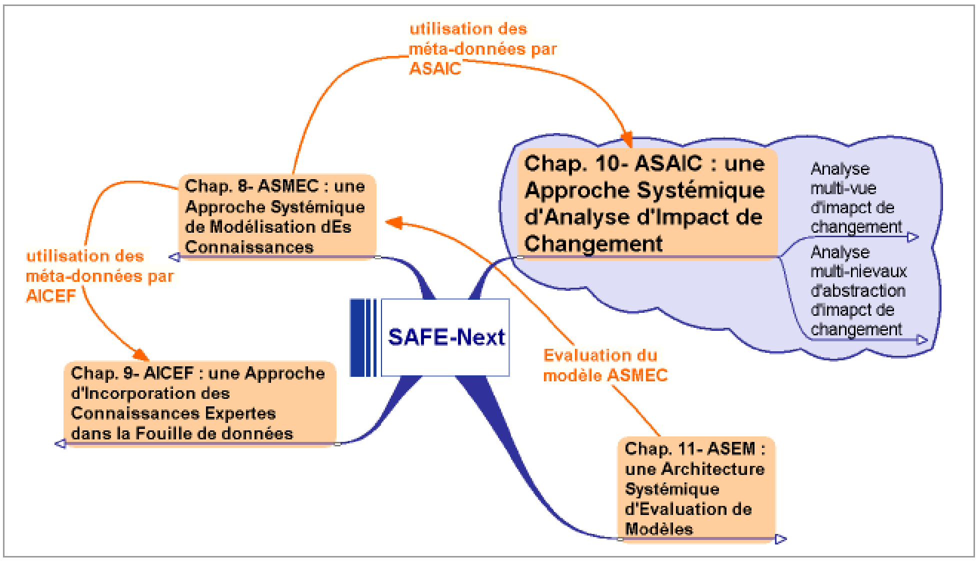


Figure 12- ASAIC : the use of the multi-view meta-data stemmed from ASMEC in order to carry out ECIA task.

Based on ASAIC, we elaborated a knowledge-Based System for a multi-view change impact analysis (see next paragraph).

In addition, we proposed a systemic representation of both the use and evaluation processes of a given product in order to analyze change impact according to these two viewpoints.

6.4 Industrial Application

Road safety is a complex domain influenced by technological, economical, political and social change. Any engineering change in the safety system design (e.g. new technology, new system etc.) can dramatically affect the Driver-Vehicle-Environment (DVE) system behavior (e.g. driver takes more risks) that in turn may have an impact on the design requirement (e.g. new safety requirements).

Developing safety systems without previously analyzing the impact of the changes introduced by these technologies on driver behavior may result in undesirable effects. For instance, a driver with a vehicle equipped with an ABS (Antilock Braking System) or ESP (Electronic Stability Program) system can take more risks than with a non-equipped vehicle. The type of accident may even change and this can imply new safety requirements or constraints that safety systems developers have to consider. On the other hand, a change in user behavior (the Driver-Vehicle-Environment system) due to a new law or a social change can imply new safety requirements or constraints that safety systems developers must take into account.

However, analyzing a change impact on the DVE's behavior requires an accidentology knowledge. Based on ASAIC, we developed a Knowledge-Based System (KBS) for Change Impact Analysis in accidentology. We also proposed a systemic framework intended to allow change impact analysis according to four viewpoints: ontological, functional, behavioral and teleological. This framework provides designers and accidentologists with comprehensive support for change analysis, information sharing and therefore an efficient handling of the change impact in a complex context. Concretely, the KBS provides analysts with three granularity levels: The first analysis level consists of identifying which systemic aspect (i.e. ontological, functional, transformational or teleological) the change affects. The second analysis level consists of identifying which element (i.e. structural component, functional step, transformation situation or teleological level) of each of the four systemic viewpoints is affected by the change. The third analysis level is based on attributes characterizing a road accident and it consists of identifying which feature of the impacted element the change concerns. Safety system developers must take into consideration the affected features of the DVE system in order to avoid undesirable effects.

The identification of the features that a new safety system may affect is crucial for safety system designers. However, this identification requires expert knowledge in accidentology. Hence, the interface between designers and accidentologists that our KBS provides is fundamental.

This study is funded by:

- The LAB (Laboratory of Accidentology, Biomechanics and Human Behavior), it is a common laboratory of the two French car manufacturers PSA Peugeot-Citroën and Renault;
- The French Research Ministry.

[Blessing, 1996] L. T. M. Blessing. Design Process Capture and Sup-port. *Proceedings of the 2nd Workshop on Product Structuring*, Delft, 1996.

[Ma *et al.*, 2003] S. Ma, B. Song, W. Feng Lu et C. F. Zhu. A Knowledge-Supported System For Engineering Change Impact Analysis. *Proceedings of DETC'03 ASME 2003, Design Engineering Technical Conferences*, Chicago, Illinois, USA, 2003.

[Murdock *et al.*, 1997] J. W. Murdock, S. Szykman et R. D. Sriram. An Information Modeling Framework to Support Design Databases and Repositories. *Proceedings of the 1997 ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, Sacramento, California, USA, 1997.

7 ASEM: A GENERIC FRAMEWORK FOR MODEL ASSESSMENT

The cybernetic epistemology understands knowledge as a model that a subject or a group constructs according to a given perception. Models are the interface between the subject and the real world in constructing knowledge and are therefore crucial to handle growing environmental complexity (e.g. industrial, social, economic contexts).

7.1 General Issue

Since models are crucial for knowledge construction, the assessment of these models is a fundamental task to see whether the constructed models fit the modeled system reality as well as intended. Furthermore, as noticed in recent research, assessment can influence perceptions about the modeled system and thereby the result model, which in turn influence the constructed knowledge. However, cybernetic and general system theory states the context-dependence of models and their continuous evolution. In spite of this variability and subjectivity, our research argues that a generic framework for model assessment can be developed.

7.2 Research Objective:

The research aims to develop a generic framework for model assessment. This framework is intended to provide users with generic assessment criteria to tackle model assessment. These criteria may be used as:

- Modeling constraints during a modeling process;
- A posteriori assessment criteria after a modeling process;
- Selection criteria to compare existing models and to select relevant models for a given objective.

7.3 Research Method & Result

The theoretical foundation of our paper is derived from cybernetics and general systems theory [Bertalanffy, 1969; Von Foerster, 1995] which are two interdisciplinary academic domains derived from constructivist epistemology. Studies in cybernetics and general systems theory are generally transdisciplinary and related to organization, communication, control and modeling. They concern several disciplines ranging from mathematics, philosophy, technology, biology and social sciences. They are more specifically related to "sciences of complexity", including artificial intelligence, neural networks, dynamical systems, chaos and complex adaptive systems. Cybernetics and general systems theory date back to the 1940's and 1950's. Some of the founding fathers are von Bertalanffy, Wiener, Ashby and von Foerster.

In a cybernetic perspective, we consider a model as an ontology (ideas, expressions, rules, patterns) open to, and interacting with, its environment through a given functioning. Via this interaction, it qualitatively acquires new properties resulting in continual evolution in order to fulfill a given teleology (goal/motivations of the subject for acquiring knowledge and then for developing a model and even for carrying out the modeling task).

In this research, we proposed an assessment framework of a complex system model, which is based on the four systemic viewpoints, namely its ontology, functioning, evolution and teleology. We developed some generic assessment criteria according to each of the systemic viewpoints. Many of these criteria are adapted from definitions stemmed from other fields (e.g. computer science, evaluation theory, performance assessment, decision making etc.).

7.4 Industrial Relevance

Modeling is a human process intrinsic to any human task [Le Moigne, 1999] and models are the basis of knowledge construction. In an industrial engineering context, models are used not only to analyze an existing system (industrial or natural) and therefore to understand and predict its behavior, but also to construct the industrial system itself. Models are considered by cybernetics and general systems theory [Bertalanffy, 1969; Von Foerster, 1995] as a perception of the real-world in a given context and thereby a tool a subject uses to construct knowledge. Theories are an example of the most known and shared models used to construct knowledge.

Hence, since models are the interface between a subject (e.g. a manger, a designer, a decision maker etc.) and the real world (e.g. a company, an organization, a natural system etc.) and therefore the basis of knowledge construction, the assessment of these models is crucial. Indeed, model assessment assures the quality of the constructed knowledge and makes these models efficient in terms of their ability and effectiveness to achieve the goal that motivated their creation.

We carried out this research with the collaboration of Mounib Mekhilef and Bernard Yannou.

[Bertalanffy, 1969] L. v. Bertalanffy. General system theory: foundations, development, applications, George Braziller, New York, 1969.

[Le Moigne, 1999] J.-L. Le Moigne. La modélisation des systèmes complexes, Dunod, 1999.

[Von Foerster, 1995] H. Von Foerster. The Cybernetics of Cybernetics (2nd edition), Future Systems Inc., Minneapolis, 1995.

8 NEW SYSTEM DEVELOPMENT

The challenge for a company is to bring to market a stream of new and improved, added value, products and services. We are focusing especially on issues related to the development of new products and new systems in general.

8.1 General Issue

Several approaches, methods and tools exist in the literature to help designers develop new systems and functions. Function Analysis and Query Functional Deployment (QFD) for example allow a designer to structure his design space (a design space is the set of design solutions and alternatives). However, these methods suppose that the main functions (functions related to the requirements) exist. Therefore, these methods only allow the deployment of *the main functions and the structuring of* ***the design space***. When one deals with new systems development, the primary need is for ***tools to construct the design space***. In other words, we need tools to define functions that will be realized in technical solutions. There is a lack of research in literature dealing with this issue.

8.2 Research Objective

The research aims at developing a scenario-based approach intended to help designers construct the design space. Let us take some examples:

- To develop new safety systems, designers need tools to understand accident behavior in order to determine adequate counter-measures;
- To develop new products, designers need tools to understand customer behavior in order to develop the adequate products.

A scenario is a *prototypical behavior of a group* of subjects or objects (customers, accidents, etc.) having similarities. It provides a knowledge base crucial in developing new systems.

8.3 Research Method & Result

We apply ASMEC (the first approach in our main approach SAFE-Next) in order to construct a multi-view domain knowledge model. Then, we apply clustering techniques in order to elaborate homogenous classes (of accidents, of customers, of behaviors etc.). Then, we apply AICEF, (the second approach in SAFE-Next) in order to carry out a multi-view attribute selection and a multi-view clusters interpretation.

We developed software which enables us to represent the same scenario according to different models specific to different fields, i.e. safety system design field and accidentology fields. Each scenario user can represent the scenario in his own representation space.

8.4 Industrial Application

We used a set of 717 accidents characterized by 947 attributes to elaborate accident scenarios. Accidentologists and designers use these scenarios in safety system development and/or assessment.

9 OUR CONTRIBUTIONS

Our research involved several disciplines, ranging from Knowledge Management, Accidentology, Knowledge Engineering, Knowledge Discovery in Databases and Design.

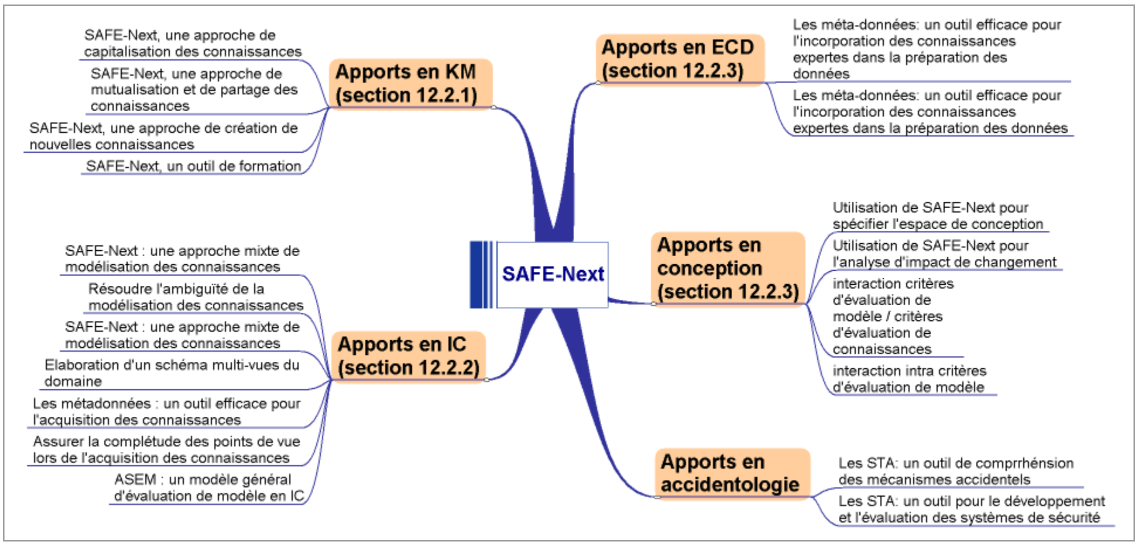


Figure 13- Our contributions in *Knowledge Engineering, Knowledge Discovery in Databases, Knowledge Management, Design and Accidentology*

In accidentology, SAFE-Next provides experts with an efficient tool for knowledge management. It enables the elaboration of multi-view accident scenarios, which are a powerful tool for understanding accident mechanisms in order to develop safety counter-measure. Furthermore, SAFE-Next provides a knowledge capitalization tool. In fact, the multi-view domain model stemming from ASMEC provides accidentologists with a synthetic representation of different types of domain knowledge. This knowledge can be shared between users, reused, updated, etc. The multi-view meta-data allows the formalization of implicit knowledge and thereby makes this knowledge more accessible and useful.

In knowledge engineering, SAFE-Next supplies, via ASMEC, a mutli-view knowledge model (observed object, observer, context of observation, functional view, ontological view, transformational view and teleological view) and thereby allows the integration of different viewpoints stemming from different users. Furthermore, ASMEC provides a multi-granularity knowledge model (meta-model, models, concepts and attributes) and in that way addresses the difficulty of knowledge identification and formalization. On another hand, SAFE-Next permit, via ASEM, the evaluation of knowledge models, an issue rarely addressed in literature.

In Knowledge Discovery in Database, SAFE-Next enables, via AICEF, the incorporation of domain knowledge in the data preprocessing step (i.e. the first step in a KDD process) and more specifically in the attribute selection task. Hence, a multi-view attribute selection can be performed according to the meta-data stemming from ASMEC. Likewise, the multi-view meta-data enables the incorporation of domain knowledge in the interpretation step (i.e. the last step in a KDD process). For instance, the same accident cluster can be represented according to several viewpoints.

In design, SAFE-Next provides safety system developers with an efficient tool to construct the design space. Scenarios enables designers to understand complex behaviors and thereby to define solutions and alternatives. SAFE-Next also provides, via ASAIC, an approach for multi-view change impact analysis. Moreover, SAFE-Next proposes to extend the change impact analysis to the *use process* of a given product as well as the *evaluation process* instead of limiting it to the design process.