



**Maria do Carmo Correia Marques**

Mestre

## **Risk-based Decision Support System for Life Cycle Management of Industrials Plants**

Dissertação para obtenção do Grau de Doutor em  
Engenharia Electrotécnica e de Computadores

Orientador: Prof. Doutor Rui Neves-Silva,  
Professor Auxiliar,  
Faculdade de Ciências e Tecnologia  
Universidade Nova de Lisboa

Júri:

Presidente: Prof. Doutor Paulo da Costa Luís da Fonseca Pinto  
Arguente(s): Prof. Doutora Teresa Maria de Gouveia Torres Feio Mendonça  
Prof. Doutor José António Barata de Oliveira

Vogais: Prof. Doutor Paulo Miguel Garcia e Costa O'Conner Shirley  
Doutora Ana Rita Gamito Bentes de Campos  
Doutora Gabriella Caporaletti



**Junho, 2011**



## **Copyright**

A Faculdade de Ciências e Tecnologia e a Universidade Nova de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objectivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.



*To Sandrine,  
your light will always guide me*

## **Acknowledgements**

Over the last years I had the opportunity to meet some of the most determining people to the success of this journey. Some of them represented a step further in my professional and personal development. Taking the risk of forgetting someone I would like to start by thanking to all the participants in the InLife project (FP6-2005-NMP2-CT-517018), partially funded by the 6<sup>th</sup> Framework Program of the European Union. The research developed within this project provided the foundations for the work presented in this thesis and some of the most important ideas were born during the several fruitful discussions between the partners.

My gratitude is also to the motivator and mentor to my work, Dr. Rui Neves-Silva. Along the years that I have been working with him he has always demonstrated his support even in situations where doubts seem to be taking advantage. His calm and focused orientation helped me to concentrate in the essential and his valuable critics and suggestions contributed for the final result of this journey. I also would like to thank him for his friendship which made the difference in so many stressful times.

I would like to thank to all my colleagues at UNINOVA, especially to Ana Rita Campos, Filipa Ferrada and Ana Inês Oliveira whose support and friendship have been invaluable during these years. A special thank goes for the students that I taught for they taught me back too, providing me one of the most rewarding experiences.

I must not forget my family since this journey would not be possible without their incommensurable love and understanding. To my parents for the lessons that encouraged me to never stop learning and seeking self development. To my sister for the supporting words that are always an antidote for any crisis and to my nephews for their joy and energy. To my grandmother for the example of strength and perseverance that guided me, through life, as role model.

Last, but definitely not least, I would like to thank Carlos for standing by me for better, for worse, for richer, for poorer, in sickness and in health.



## Abstract

The objective of this thesis is to contribute for a better understanding of the decision making process in industrial plants specifically in situations with impact in the long term performance of the plant.

The way decisions are made, and especially the motivations that lead to the selection of a specific course of action, are sometimes unclear and lack on justification. This is particularly critical in cases where inappropriate decisions drive to an increase on the production costs. Industrial plants are part of these cases, specifically the ones that are still lacking enhanced monitoring technologies and associated decision support systems.

Maintenance has been identified as one of the critical areas regarding impact on performance. This is due to the fact that maintenance costs still represent a considerable slice of the production costs. Thus, understanding the way maintenance procedures are executed, and more important, the methods used to decide when maintenance should be developed and how, have been a concern of decision makers in industrial plants.

This thesis proposes a methodology to efficiently transform the existing information on the plant behaviour into knowledge that may be used to support the decision process in maintenance activities. The development of an appropriate knowledge model relating the core aspects of the process enables the extraction of new knowledge based on the past experience. This thesis proposes also a methodology to calculate the risk associated to each maintenance situation and, based on the possible actions and on the impacts they may have in the plant costs performance, suggests the most appropriate course. The suggestion is made aiming the minimization of the life cycle costs.

Results have been validated in test cases performed both at simulation and real industrial environments. The results obtained at the tests cases demonstrated the feasibility of the developed methodology as well as its adequateness and applicability in the domain of interest.

**Keywords:** Industrial Knowledge Models, Risk Analysis, Intelligent Decision Support Systems, Life cycle management





## Resumo

O objectivo desta tese é contribuir para uma melhor compreensão do processo de tomada de decisão em ambientes industriais, com especial foco nas situações com impacto no comportamento de longo prazo da instalação.

O modo como as decisões são tomadas e, especialmente as motivações que conduzem à selecção de curso de acção específico, são muitas vezes pouco claras e carecem de justificação. Isto é particularmente crítico nos casos em que uma decisão inapropriada pode conduzir a uma escalada dos custos de produção. Este é muitas vezes o caso de instalações industriais que não possuem tecnologias de monitoração associadas a sistemas de suporte à decisão.

A area da manutenção foi identificada com uma das áreas críticas no que diz respeito ao impacto no comportamento da instalação. Isto deve-se ao facto de os custos de manutenção continuarem a representar uma parte importante dos custos de produção. Assim, compreender o modo como a manutenção é executada e, mais importante, os métodos que são usados para decidir quando e como se deve fazer manutenção, tem sido uma preocupação dos agentes de decisão em ambiente industrial.

Esta tese propõe uma metodologia para transformar a informação existente acerca do comportamento da instalação em conhecimento que pode ser utilizado para suportar o processo de tomada de decisão em acções de manutenção. O desenvolvimento de um modelo de conhecimento apropriado, que relaciona os aspectos centrais do processo, permite a extracção de novo conhecimento baseado na experiência passada. Esta tese propõe ainda uma metodologia para calcular o risco associado a cada situação de manutenção e, com base nas possíveis acções a implementar e nos diferentes impactos associados a cada uma delas em termos de custos, sugere o curso de acção mais apropriado. A sugestão é feita com o objectivo de minimizar os custos de ciclo de vida.

Os resultados foram validados em casos de teste que foram efectuados tanto em ambiente de simulação como em ambiente industrial. Os resultados obtidos demonstraram a exequibilidade da metodologia desenvolvida, bem como a sua adequação e aplicabilidade ao domínio de interesse.

**Palavras-chave:** Modelos de Conhecimento Industrial; Análise de Risco; Sistemas Inteligentes de Suporte à Decisão; Gestão do Ciclo de Vida



# Table of Contents

<b>1. INTRODUCTION .....</b>	<b>1</b>
1.1. General Context .....	1
1.2. Motivations for decision support in industry.....	2
1.2.1. Maintenance impact in the life cycle of industrial plants .....	4
1.3. Research problem and scientific contributions .....	7
1.4. Outline .....	9
<b>2. MAINTENANCE FOR LIFE CYCLE MANAGEMENT .....</b>	<b>11</b>
2.1. Classic maintenance methodologies .....	12
2.1.1. Run-to-failure .....	12
2.1.2. Preventive maintenance.....	13
2.1.3. Predictive maintenance and Condition-based maintenance .....	14
2.2. Maintenance improvement methodologies .....	17
2.2.1. Total productive maintenance .....	17
2.2.2. Reliability-centered maintenance .....	18
2.2.3. Risk based inspection / Risk based maintenance .....	19
2.2.4. Life cycle maintenance.....	20
2.3. Conclusions .....	22
<b>3. RISK ANALYSIS AND DECISION SUPPORT SYSTEMS .....</b>	<b>25</b>
3.1. Risk analysis.....	25
3.1.1. Historical overview .....	25
3.1.2. Probabilistic risk assessment .....	26
3.1.3. Decision under risk.....	29
3.2. Decision support systems.....	32
3.2.1. Reaching a decision: overview of the decision-making process .....	34
3.2.2. Brief history of decision support systems .....	40
3.2.3. Intelligent decision support systems .....	42
3.2.4. Decision support systems for life cycle management .....	50
3.3. Conclusions .....	53
<b>4. PROPOSED DECISION SUPPORT METHODOLOGY BASED ON RISK ANALYSIS.....</b>	<b>55</b>
4.1. Problem framework .....	56
4.2. Concept for risk-based decision support system.....	57
4.2.1. Information collection .....	60
4.2.2. Knowledge model.....	61
4.2.3. Information correlation and aggregation .....	62
4.2.4. Dealing with uncertainty .....	64
4.2.5. Risk analysis .....	65
4.2.6. Decision model.....	67
4.2.7. Proposed decision algorithm.....	69
4.2.8. Numerical example .....	72
4.3. Conclusions .....	75
<b>5. VALIDATION THROUGH SIMULATION .....</b>	<b>77</b>
5.1. Development of the simulator.....	77
5.1.1. Industrial plant simulation.....	80
5.1.2. Decision support system .....	84
5.2. Simulation results .....	91
5.2.1. Learning capabilities of the algorithm.....	91

5.2.2. Comparison with standard maintenance methodologies.....	95
5.3. Conclusions .....	102
<b>6. VALIDATION IN INDUSTRIAL APPLICATIONS.....</b>	<b>105</b>
6.1. Validation methodology.....	106
6.1.1. Development of the life cycle management system prototype .....	106
6.1.2. Gathering and correlating information .....	111
6.1.3. Assessment methodology .....	113
6.2. Scenario I: condition-based maintenance of a manufacturing plant .....	114
6.2.1. Description and objectives .....	114
6.2.2. Testing scenario .....	115
6.2.3. Test results.....	119
6.3. Scenario II: condition-based maintenance from the product perspective .....	121
6.3.1. Description and objectives .....	121
6.3.2. Testing scenario .....	122
6.3.3. Test results.....	127
6.4. Conclusions .....	129
<b>7. CONCLUSIONS AND FUTURE WORK .....</b>	<b>131</b>

## Table of Figures

Figure 1.1. The total cost visibility .....	5
Figure 2.1. Bathtub Curve .....	13
Figure 2.2. Typical cost of CBM installation and operation .....	15
Figure 2.3. Typical potential savings produced by CBM .....	16
Figure 2.4. The P-F curve .....	19
Figure 2.5. Maintenance evolution paradigm .....	23
Figure 3.1. Example of an Event Tree.....	29
Figure 3.2. Examples of utility functions.....	30
Figure 3.3. Example of value function .....	31
Figure 3.4. Weighting function.....	32
Figure 3.5. The decision table .....	38
Figure 3.6. The Case-based Reasoning cycle .....	45
Figure 3.7. Bayesian approach .....	47
Figure 3.8. Deriving decisions in Bayesian approach .....	48
Figure 3.9. Example of a Bayesian network.....	49
Figure 3.10. Life cycle of industrial plants .....	51
Figure 4.1. Focusing the problem.....	56
Figure 4.2. Concept for Risk-based Decision Support System .....	59
Figure 4.3. Developed model for the Knowledge Repository .....	61
Figure 4.4. Expected behaviour of state variable $x_1$ considering different maintenance strategies .....	62
Figure 4.5. Behaviour of $x_1$ according with the performed actions .....	63
Figure 4.6. Correlation between $x_1$ and $\Delta T_f$ .....	64
Figure 4.7. Risk analysis based on probabilities .....	67
Figure 4.8. Decision model.....	68
Figure 4.9. Decision Model for the current example.....	75
Figure 5.1. Conceptual architecture of an Intelligent Decision Support System .....	78
Figure 5.2. System workflow .....	79
Figure 5.3. General state function .....	80
Figure 5.4. Uncertainty on the result of an action using an approximated state function.....	80
Figure 5.5. Calculation of the threshold accordingly with the action performed .....	81
Figure 5.6. Probability of Developing a Failure vs the Logit of risk .....	82
Figure 5.7. State function and its relation with threshold establishment .....	83
Figure 5.8. Discrete state function.....	84
Figure 5.9. Plant workflow .....	84
Figure 5.10. Workflow for building the decision model.....	86
Figure 5.11. Risk Analysis workflow .....	87
Figure 5.12. Workflow to provide suggestion .....	88
Figure 5.13. Workflow for updating the decision model .....	89
Figure 5.14. Risk based Decision Support System: Results Interface .....	90

Figure 5.15. Testing accuracy with 50 instances .....	92
Figure 5.16. Testing accuracy with 300 instances .....	93
Figure 5.17. Learning curve with no previous knowledge at set-up .....	94
Figure 5.18. Learning curve with 50 problems introduced at set-up .....	95
Figure 5.19. Accumulated costs for test case 1 .....	97
Figure 5.20. Accumulated costs for test case 2 .....	98
Figure 5.21. Accumulated costs for test case 2 with different maintenance plans.....	99
Figure 5.22. Accumulated costs for test case 3 .....	100
Figure 5.23. Accumulated costs for test case 3 with minimum threshold on 0.5.....	101
Figure 5.24. Accumulated costs for test case 2 with minimum threshold on 0.1.....	102
Figure 5.25. Comparison of the three test cases .....	103
Figure 6.1. InLife rationale.....	106
Figure 6.2. Life Cycle Management System concept.....	107
Figure 6.3. Flowchart for automatic behaviour of LCMS .....	108
Figure 6.4. LCMS Condition-based Maintenance service home. ....	109
Figure 6.5. Simplified model.....	110
Figure 6.6. InLife set-up tool.....	113
Figure 6.7. Functional scheme and life cycle parameters in scenario II data model.....	124

## Table of Tables

Table 3.1. Expectancy Theory table .....	31
Table 3.2. Major components of a case .....	44
Table 4.1. Stratification of $\Delta t$ .....	65
Table 4.2 Specifications for numerical example: Symptom, Causes, Actions and Costs .....	72
Table 4.3. Specifications for numerical example: intervals for maintenance actions .....	73
Table 4.4. Cases caused by $C_1$ and associated sub-group.....	74
Table 4.5. Cases caused by $C_2$ and associated sub-group.....	74
Table 5.1. Software tools.....	79
Table 5.2. Example of conditional probability for action $A_i$ .....	81
Table 5.3. Characteristics specified for actions.....	96
Table 5.4. Outcomes probabilities for specified actions .....	96
Table 6.1. Common information sources .....	112
Table 6.2. Data model for the manufacturing plant of scenario I.....	115
Table 6.3. Data model for life cycle parameters of scenario I .....	117
Table 6.4. List of causes and associated actions for scenario I .....	117
Table 6.5. List of causes and associated consequences for scenario I .....	117
Table 6.6. Scenario for testing Condition-based Maintenance.....	118
Table 6.7. Testing results for scenario I .....	120
Table 6.8. Data model for GAAS acclimatization units' family .....	122
Table 6.9. Data model for GAAS acclimatization unit .....	123
Table 6.10. Data model for life cycle parameters of scenario II .....	125
Table 6.11. List of causes and associated actions for scenario II .....	125
Table 6.12. List of causes and associated consequences for scenario II .....	125
Table 6.13. Scenario for testing Condition-based Maintenance.....	126
Table 6.14. Testing results for scenario II .....	129





## List of Acronyms and Abbreviations

<b>AFR</b>	Average Failure Rate
<b>AHP</b>	Analytic Hierarchy Process
<b>Aml</b>	Ambient Intelligence
<b>APL</b>	A Programming Language
<b>ASP</b>	Application Service Provider
<b>BN</b>	Bayesian Networks
<b>BI</b>	Breakdown Intensity
<b>CBM</b>	Condition-Based Maintenance
<b>CBR</b>	Case-Based Reasoning
<b>CMMS</b>	Computerized Maintenance Management System
<b>DSS</b>	Decision Support Systems
<b>EIS</b>	Executive Information Systems
<b>ETA</b>	Event Tree Analysis
<b>EUT</b>	Expected Utility Theory
<b>EV</b>	Expected Value
<b>FMEA</b>	Failure Modes and Effects Analysis
<b>FR</b>	Failure Rate
<b>GDM</b>	Group Decision Making
<b>GDSS</b>	Group Decision Support Systems
<b>IA</b>	Inherent Availability
<b>IDE</b>	Integrated Development Environment
<b>IDSS</b>	Intelligent Decision Support Systems
<b>InLife</b>	Integrated Ambient Intelligence and Knowledge-Based Services for Optimal Life Cycle Impact of Complex Manufacturing and Assembly Lines
<b>KBS</b>	Knowledge-Based Systems
<b>LCC</b>	Life Cycle Costs
<b>LCM</b>	Life Cycle Management
<b>LCP</b>	Life Cycle Parameters

<b>LCMS</b>	Life Cycle Management System
<b>MAUT</b>	Multi-Attribute Utility Theory
<b>MTBF</b>	Mean Time Between Failure
<b>MTTF</b>	Mean-Time-To-Failure
<b>NEE</b>	Net Equipment Effectiveness
<b>OA</b>	Operational Availability
<b>OCE</b>	Overall Craft Effectiveness
<b>ODSS</b>	Organizational Decision Support Systems
<b>OEE</b>	Overall Equipment Effectiveness
<b>OLAP</b>	On-Line Analytical Processing
<b>PdM</b>	Predictive Maintenance
<b>P-F</b>	Potential - Functional
<b>PLM</b>	Product Lifecycle Management
<b>PM</b>	Preventive Maintenance
<b>PT</b>	Prospect Theory
<b>RbDSS</b>	Risk-based Decision Support System
<b>RBI</b>	Risk Based Inspection
<b>RBM</b>	Risk Based Maintenance
<b>RBS</b>	Rule-Based Systems
<b>RCM</b>	Reliability-Centered Maintenance
<b>RTF</b>	Run-To-Failure
<b>SDM</b>	Supra Decision Maker
<b>SI</b>	Spontaneity Intensity
<b>SME</b>	Small and Medium Enterprise
<b>SUDM</b>	Sensory-Updated Degradation-based Maintenance
<b>TBM</b>	Time-Based Maintenance
<b>TPM</b>	Total Productive Maintenance
<b>TQM</b>	Total Quality Management
<b>UBM</b>	Usage-Based Maintenance

## Glossary

**Life cycle management** is a management approach that can be used by all types of organizations to improve their processes and thus the sustainability performance of the companies and associated value chains.

**Decision-making process** is a mental cognitive process that results in the selection of a course of action among several alternative scenarios. Every decision making process produces a final choice.

**Decision support systems** are computer-based information systems that support business or organizational decision-making activities.

**Intelligent Decision Support Systems** are a type of decision support systems that make extensive use of artificial intelligence techniques.

**Risk analysis** is a technique to identify and assess factors that may jeopardize the success of a project or achieving a goal.

**Decision under Risk** is the process of making a decision without knowing the exact outcomes.

**Maintenance methodology** is the method followed by an industrial company to organise their maintenance activities.

**Life cycle maintenance** is an approach to maintenance procedures in order to include them as part of a life cycle management perspective.



# *1. Introduction*

---

This thesis intends to contribute for a better understanding of the decision making process in industrial plants, specially, on the long term performance impact of short term decisions.

The way decisions are made in industrial environments, and especially the motivations that lead to the selection of a specific course of action, are sometimes unclear and may drive to an increase on the production costs. This is particular critical in industrial plants still lacking enhanced monitoring technologies and associated decision support systems.

Maintenance has been identified as one of the critical areas regarding long term impact on performance. This is due to the fact that maintenance costs still represent a considerable slice of the overall production costs and strongly affects the lifetime of the plant under acceptable performance. Thus, understanding the way maintenance procedures are executed and, more important, the methods used to decide when maintenance should be developed and how, have been a concern of decision makers in industrial plants.

## **1.1. General Context**

If something positive has born from the, sometimes not so clear, basis of the globalized economy was the recognition of knowledge as a major asset in people's life. Knowledge is, and will remain being, the main factor for competition both at personal and organizational levels, conducting tendencies at global markets.

This is reflected in the way more traditional markets try to compete with the emergent ones. The use of their knowledge, acquired through a long time experience, is being used to battle for a place in the world wide economy. This is achieved, most of times, by a combination of high quality raw materials together with a growing encouragement of applied scientific research.

In fact, in the latest years we have been observing that the companies that successfully survived the impact of globalisation were the ones that added to their traditional way of making business the technological development component. This puts in evidence that even if knowledge exists it

will not be any helpful if the company does not take the most out of it, i.e. the way knowledge is used may represent the difference between success and failure. Finding the best ways of using that knowledge, and transform it into a primary asset, is then a key issue for any company that embraces the challenge of competing in the global market.

Nonetheless, and although many companies are aware of the need for changing, they are sometimes frightened by the investments needed to perform that transformation. Understand the risks of investment and the potential benefits may be an obstacle for many companies, but the decision making will be highly facilitated if the company has at least a rough idea of the possible outcomes of its decisions. Thus the need for a complete and assertive risk analysis method, able of supporting the decisions to be made, is fundamental for the success of the company strategy.

Indeed, the concept of using risk analysis as a support for decision making is not new, and has proved its efficacy in several domains from business to finance, from management to market sales (Grünig & Kühn, 2005), (Lapin, 1987), (Lindley, 1991).

However, despite the importance that making a decision may have a decision is not valid *per se*. Its value is only reflected on the main objective for which that decision will contribute. Companies that try to find new and improved ways to foster their performance, normally reflect this position in the refinement of their decision strategy. The analysis of the impact of the decisions that are made is used in the refinement process evidencing that the decision strategy adopted affects the life cycle of the company and, eventually, determines its success. The integration of these impacts in a more holistic view contribute for the development of a Life Cycle Management (LCM) approach aiming to contribute for improving the overall success of the plant on the long term perspective.

## **1.2. Motivations for decision support in industry**

The capacity of reaching a decision is an important task in the daily life of both companies and people. In fact the decision making process is present in all choices that a person, or a group, has to perform. This process can be made intuitively, when little knowledge about the situation exists or it can follow some kind of protocol using existing knowledge to reach a consistent decision.

In some cases there is the need to collect knowledge and points of view from different specialists in different areas. In these cases the establishment of the correct set of criteria is a crucial point. Additionally it may also be needed to add a sort of voting system that enables reaching a conclusion at the end of the process. This situation is very common in many companies where decisions are no longer made by one person but are part of an integrated process where the different parts involved have a word to say in their field of experience. These decisions are normally oriented for reaching a subsequent higher level objective that is part of the overall strategy of the company.

Together with the result of the decision itself, the way decisions are made may have a strong impact in company success. The adoption of a decision support system can help to streamline the

decision making process. This adoption is normally stimulated by different aspects such as the ones pointed out by Holsapple and Whinston (Holsapple & Whinston, 2000):

- **Economic limitations:** Humans are expensive. Although the lack of politeness in this sentence the fact is that increasing the number of humans to compensate cognitive limitations can become economically prohibitive;
- **Temporal limitations:** A decision may be needed in a time frame that puts severe pressure on a decision maker. This may affect the decision maker and contribute negatively to the resultant decision;
- **Cognitive limitations:** Human memory is limited when handling high amounts of information. This affects the correlation capacity of human beings compromising their ability of reaching a conclusion;
- **Competitive strategy implications:** the existence of decision support systems are thought of having a positive impact in the company strategy since, in principle, the time, money and other resources spent on decision making process will tend to decrease.

Although some time has passed since this identification was made, its validity remains intact. Note that almost all aspects highlighted in the previous list are based on limitations that are associated with human nature. In any case, humans remain an important part of the process since they are the ones who, eventually, will have to validate the final decision.

Nonetheless, and as already mentioned, more important than reaching a decision is the result of the decision itself, especially in the long term perspective. Thus, as important as quickly reach a decision to solve a problem is the knowledge of the impact that decision will have in the future of the entire company.

The same conclusion is valid when the focus is on the decisions that are made, on a daily basis, in an industrial plant. When focusing on the industrial companies the production plant is a critical part of the company's performance. Its capacity of giving answer, in time, to the required demands, maintaining production quality is crucial for the company success. This highlights the importance of the plant as a central part of any decision that affects production patterns.

Thus, the existence of structured decision mechanisms to support the decision making has been recognised as a key issue in several aspects of industrial plants: from location (Badri, 2007) to process (Aldrich, Schmitz, & Gouws, 2000), from logistics (Tan & Kumar, 2006) to maintenance (Bouza-Fernandez, Gonzalez-Filgueira, de las Heras Jimenez, & Vazquez-Gonzalez, 2010). This aspect assumes an additional dimension when the long term perspective is considered. In fact the way the plant is managed is vital to increase its production and extend its lifetime, thus making it more efficient from long term perspective.

### 1.2.1. Maintenance impact in the life cycle of industrial plants

Life cycle management of industrial plants strongly relies in taking decisions on a daily basis: what configuration is best, what production pattern, when to replace a part, who should be involved in a maintenance process, etc.

From a general point of view, maintenance activities are seen as a necessary trouble, and the general feeling is that nothing can be done to reduce maintenance costs. This might be true when considering maintenance in its traditional scope in which the activities were limited to the operation phase. For many years maintenance was seen as a pure tactical task since it was concerned with quickly restoring failures to an operating condition (Barringer, 2001).

However, nowadays, as the paradigm of industry shifts towards realizing a sustainable society, there is a recognizable change in the role of maintenance. The goal of industry is no longer solely based on efficiency, but also on the provision of functions needed by society while minimizing material and energy consumption (Takata, et al., 2004). And, to achieve this goal, the need for a holistic view, which results in the use of a life cycle approach to maintenance, is crucial.

Taking this context into consideration, the role of maintenance must be redefined as an essential means for life cycle management. In this view the decision of performing a maintenance action each day will definitely have an impact on the life cycle of the plant.

Thus, it becomes clear that the life cycle concept relies heavily on maintenance issues. This idea has been pursued by studies performed in industrial environments in which the use of historical maintenance data is seen as a good prediction mechanism for future failure behaviour ( (Gitzel, 2010), (Sondalini, 2008)). Example of this work is the study carried out by *ABB Full Service* and *ABB Corporate Research Life Cycle Science group* which goal was to see whether a statistical analysis of the data provided by a Computerized Maintenance Management System (CMMS) could be used as decision support for maintenance and investments (Gitzel, 2010).

Unfortunately, in Gitzel's study concluded that the benefits of using the stored data were not so obvious due to the complexity of the simulations needed to reach some usable result. In fact if one spends a long time in analysing and simulating data then it is likely to conclude that no important impact will come from that approach. And, especially when we are dealing with industry, time and resource consuming approaches are normally signs of growing costs. Thus, the development of methods to treat industrial plants complexity, without increasing it, should be sought.

The problem of cost is one of the key issues in maintenance, since its costs are a major part of the total costs of any industrial plant and can represent 15% to 60% of the costs of the goods produced, depending on the type of industrial sector (Mobley, 2002).

In reality, many manufacturers do not know exactly the true costs of their products, since they cannot quantify the after-sales cost that are incurred and often the total cost of the machinery is not visible. Truth is that visible costs of any purchase only represent a small proportion of the total



cost of ownership. Figure 1.1 gives a graphic representation using an iceberg analogy that highlights the dangers of poor financial management if only the apparent costs are considered.



Figure 1.1. The total cost visibility<sup>1</sup>

Studies performed on maintenance management effectiveness indicated that one-third of all maintenance costs are wasted as the result of unnecessary or improperly carried out maintenance (Ellis, 2003). Moreover, ineffective maintenance management has a significant impact in the production capacity of the plant, especially when companies have to compete in a globalised economy where there is little room for the non-evolving ones.

The main cause for this ineffective management is the lack of data that would enable to assess the need for repair or maintain plant machinery, equipment, etc. For this reason, in many cases, maintenance is still performed based on simple statistical trend data or on the actual failure of the plant. On the other hand the current access to an entire world of sensors and microprocessors can be used to monitor the operating condition of the plants, and help reducing unnecessary repairs, preventing machine failures, and reducing the impact of maintenance on plant operation. This is already being implemented by some industrial plants contributing for their performance improvement.

Thus, the implementation of maintenance strategies for improving system reliability, preventing the occurrence of unexpected system failures, and reducing maintenance costs is a key issue (Kaiser & Gebraeel, 2009). This leads us to a double based approach where both aspects, prevent failure and reduce costs would have an important impact in the overall life cycle management of the plant.

Unfortunately aggregating both aspects can be a quite challenging task since, from a general point of view they are inversely proportional, which means that when we reduce maintenance costs equipment will probably register more failures and vice-versa.

---

<sup>1</sup> Adapted from *Life Cycle Costing* by Public Competition and Purchasing Unit, 1992

Often these impacts are not obvious or immediate, in fact many of them are hidden or indirect, and they only appear when a more holistic view is taken. Thus the central problem is on how to combine the existing data so that the impact of a specific decision can be foreseen. To solve this problem there is the need to understand plant behaviour and for this the implementation of monitoring technologies is necessary.

Nonetheless, even in cases where technology exists, the data available on the plant behaviour is not structured and it is scattered along the plant making extremely difficult to correlate it and reach a consistent conclusion, i.e. a decision.

Additionally, an industrial plant can produce, through the available instrumentation, a huge amount of data from where decisions could be supported if well interpreted (Marques & Neves-Silva, 2009). The problem is to identify what data should be collected and how it should be processed in order to be effectively useful. This may not be trivial to overcome since each industrial plant represents a unique problem and the solution requires the participation of experts with great knowledge on the specificities of the plant.

It is therefore vital to develop new methods and tools that contribute for the mitigation of these problems, facilitating the collection of data about the plant behaviour and enabling the establishment of cause-effect relations transforming the information into knowledge. These relations allow the development of a decision support system which, by processing the present situation together with the historical information, is able of calculating the risk involved in the current situation and propose the most adequate course of action (i.e. suggest the maintenance procedure that should be followed). The continuous collection of information permits the adaptation of the decision process to new demands assuring its effectiveness.

This thesis proposes a methodology to support decision process in maintenance aiming to contribute for improving industrial plant life cycle based on the risk analysis of each course of action. The methodology proposed aims at accommodating the following aspects:

- **Good level of problem insight:** In general, plants can be very complex systems which, to be treated, must be fully understood, especially in their constituents parts so that influence can be analyzed and conclusions can be extracted;
- **Effortless usability:** Although recognizing the high level of complexity of a plant, the solution cannot be more complicated than the problem;
- **Straightforward understanding:** One of the most important parts of any decision support method is its capacity of being explained, i.e. understand how the decision was reached and which were, in terms of parameters, the main contributors.

Such an approach is thought of contributing for reducing the negative impact of maintenance actions on the overall production and the consequent enhancement of the life cycle of the industrial plant. Hence the decision to focus this PhD thesis on the definition of appropriate data to be used for decision support within maintenance, together with the development of a decision

support methodology based on risk analysis covers an important part of the aimed life cycle management approach.

### **1.3. Research problem and scientific contributions**

Taking into consideration the driving reasons mentioned above it is possible to conclude that having complete knowledge about the impact of a specific decision in the life cycle of an industrial plant could influence greatly the way decisions are made and executed.

Thus, the research question addressed by this work can be summarised as:

#### **Research question**

To what extent can a short term decision influence the long term perspective of an industrial plant and how to find the best option to optimise its life cycle impact?

The adopted work hypothesis to address the research question is defined below:

#### **Research Hypothesis**

The use of simulation together with data models, instantiated with specific information of the industrial plant, can contribute to provide insight knowledge about existing options and the impact of each of them in the life cycle of the plant.

Note that the use of the word *simulation* in the above research hypothesis is deeply related with the establishment of “what-if” scenarios and does not represent the possibility of developing temporal simulations, which can be obtained by using the simulator developed (see chapter 5).

The main scientific contributions that have been developed under the framework of this thesis are:

### **Main Contributions**

1. Development of a **plant generic model** based on the analysis performed. This model describes the important parts of the plant and is used to store the accumulated knowledge which is the basis for the decision process. For an appropriate result in terms of decision the model needs to be filled with plant specific data, especially in what concerns the main cause-effect relations. If little knowledge on plant behaviour exists then the system will be able to learn from experience, filling the model and refining the decision result along time.
2. Definition of a **risk analysis strategy** based on probabilistic risk assessment that uses conditional probabilities associated to the different parts of the plant model. These probabilities are established both by previous knowledge on the behaviour of the plant as well as by the continuous updating during operating period. The probabilities are refined along time contributing for an improvement of the methodology results.
3. Development of an **algorithm that combines the risk analysis strategy with a decision tree model** to incorporate the effect that different actions performed over the plant may have in its long term behaviour. The effect is established by comparing the current situation with situations occurred in the past. The actions performed in those past situations, and their outcomes, are used to establish the costs of each course of action. The final result is the selection of the most adequate course of action, as a combination of cause of the situation and action to be executed. This selection is based on the course of action presenting lower costs.
4. Development of a **simulator together with a collection of industrial test cases** to be used as test base for the extensive tests needed to assess the quality of the approach. The **test cases** were developed in real industrial environments and were focused in testing specific parts of the methodology, namely, the generic plant model and the risk analysis strategy. The **simulator** was developed to overcome the difficulty in achieving life cycle results in a suitable time frame. The algorithm combining risk analysis with decision trees was extensively tested in simulation environment providing insight about its capabilities.

## 1.4. Outline

This thesis is organised in seven chapters: ***Introduction, Maintenance for Life Cycle Management, Risk Analysis and Decision Support Systems, Proposed Decision Support Methodology based on Risk Analysis, Validation through Simulation, Validation in Industrial Applications***, and ***Conclusions and Future Work***.

On chapter 1 ***Introduction*** the problematic of decision is presented highlighting its presence in the daily life of both companies and people. Additionally the scope of the work is introduced together with the associated motivations that served as baseline for the research presented. The chapter also states the research problem, the research hypothesis and the main original contributions of the work. In the end the outline of the thesis is presented.

Chapter 2, ***Maintenance for Life Cycle Management***, starts by describing several maintenance methods that are currently used in industry together with different maintenance methods and programs that may be used to reach a specific objective. The underlying concept of life cycle maintenance is presented highlighting its potential impact on the company performance towards a more sustainable strategy.

Chapter 3, ***Risk Analysis and Decision Support Systems***, congregates the aspects of risk analysis with the decision support issue. It starts by describing the historical perspective of risk problematic, showing the importance of analysing risks when a decision is to be made. Additionally some aspects on risk calculation are also addressed. With respect to decision support systems, the main concepts related with decision making process are here introduced to clarify the importance of developing systems to help that process. Then an historical overview of the decision support systems is presented together with a special reference to Intelligent Decision Support Systems. In the end the link between the utilization of decision support systems to improve life cycle management is established.

In chapter 4, ***Proposed Decision Support Methodology based on Risk Analysis***, the ideas that supported the research developed are presented. The chapter starts by presenting the problem and propose the concept for the risk-based decision support system. Then, the approaches used to achieve the aimed results at each part of the work are detailed namely in what concerns the way that valuable information on plant status is treated, the developed knowledge model, the strategy used for risk analysis and how it was combined with a decision tree model. The formalisation of the developed decision algorithm is presented at the end of the chapter.

Chapter 5, ***Validation through Simulation***, presents the simulator developed to support the intensive testing of the methodology, especially in what concerns the decision algorithm combining risk analysis with decision trees. The need for this simulation is due to the difficulty in achieving long term conclusions when dealing with real industrial environments. Thus simulation enables the extraction of life cycle results that can be generalised for a real industrial plant. Additionally several tests on learning and accuracy capabilities of the system are also presented

together with comparisons between the utilization of the developed methodology against other maintenance methods.

Chapter 6, ***Validation in Industrial Applications***, presents the system developed and used as validation prototype in real industrial environments. The chapter includes the aspects related to the implementation of the prototype, based on the algorithms proposed. The validation at this stage was mainly focused on the methods to collect and correlate information, the knowledge model building and the risk analysis strategy. The test cases developed for validation were deployed in real industrial environments using real data to feed the knowledge model.

Chapter 7, ***Conclusions and Future Work***, discusses the main conclusions of this work, detailing the problems found along the way and pointing out potential solutions for them. Additionally, the main results and contributions are summarised, and some possible directions for further research are mentioned.

## *2. Maintenance for life cycle management*

---

As previously acknowledged maintenance has always contributed with a significant part for the production costs on industrial companies. In the past, maintenance was simply regarded as repair work: machines were operated until they broke down, and there was no way to predict failures. And, although maintenance concepts and methodologies have advanced significantly over the past several decades, maintenance still has a negative image since some companies still look at it as a direct measure against troubles. In this context the maintenance departments are many times seen as cost-centres, which do not create profits.

With the advent of mass production, maintenance changed from fixing things to replacing them. And, for many years the general idea was that it was not worth to fix machines and would be better to buy new.

The growing worries about environment and sustainability turn this view obsolete in many areas. Industrial managers are currently especially aware of these new demands especially due to the increasing costs that may represent decommissioning and waste treatment. This new picture represents an opportunity for the development of a new vision for maintenance positioning inside the industrial processes.

In fact if one looks at the role of maintenance from the perspective of life cycle management, it is easy to see that this is the new picture industry must seek. Based on the main purpose philosophy of life cycle management, the goal should be focus in the control of the conditions of production so that new functionalities, required by customers or by society, may be provided while keeping the environmental load at a minimum and maintaining appropriate corporate profits (Takata, et al., 2004).

## **2.1. Classic maintenance methodologies**

To understand the need for a new view on the maintenance management perspective this section goes deep into the traditional approaches for performing maintenance. Industrial plants may employ different types of maintenance management methods: from run-to-failure to condition-based maintenance associated with prediction techniques. Their main features are described in the following text.

### **2.1.1. Run-to-failure**

The logic of run-to-failure is straightforward: each time a machine breaks down, fix it. This method implies that if the machine is not broken then nothing should be done and, in the early industrial age, it seemed quite a good approach since no money was spent in maintenance until a serious failure occurs (Mora, 2010).

Run-to-failure is then a reactive management technique that is only triggered by the failure of equipment. This characteristic leads some critics to say that this is “no maintenance” management approach at all (Mobley, 2002) since when something is seriously damaged the odds of putting it back to work are very little and normally the equipment must be replaced.

This strategy (or lack of it) leads in general also to an increase in costs due to the break/replace approach, since equipment is not repaired until it fails to operate. According to Mobley (Mobley, 2002) the major costs associated to this type of maintenance are:

- **High spare parts inventory costs:** maintenance department is forced to maintain extensive spare parts (including spare machines or at least major components) for all critical equipment of the plant or rely on equipment vendors for immediate delivery;
- **High overtime labour costs:** maintenance personnel must be able to react immediately to all failures to minimise the impact on production;
- **High machine downtime:** the time to repair is dependent on the existence of spare parts and personnel availability. If something is missing then the machine downtime increases; if nothing is missing then the costs increase due to need of having everything available;
- **Low production availability:** with higher machine downtime, production availability decreases. Once again the alternative is to guarantee full availability of spares and personnel... and pay for them.

For these reasons few plants use a true run-to-failure philosophy, especially nowadays. In several cases maintenance departments perform some basic preventing tasks (e.g. lubrication and minor adjustments) waiting for total failure to occur.



### 2.1.2. Preventive maintenance

With the development of reliability engineering in the 1950s, the concept of Preventive Maintenance (PM) was advocated by a group of Japanese engineers (Mora, 2011), and notions such as Time-Based Maintenance (TBM) and Usage-Based Maintenance (UBM) were introduced.

There are many definitions of preventive maintenance, but all of them rely of a time-driven approach in which maintenance tasks are performed based on elapsed operation time. Thus, in preventive maintenance repairs are scheduled based on measures of the Mean-Time-To-Failure (MTTF)<sup>2</sup>.

In this scope, it is widely accepted that, in what concerns equipment failure, there is a strong connection between the age of the equipment and its failure rate. In fact, it has long been known that most groups of similar machines will exhibit failure rates that are somewhat predictable if averaged over a long time. This idea drives to the so-called "Bathtub Curve" (named from the cross-sectional shape of a bathtub) which relates failure rate to operating time (see Figure 2.1). The bathtub curve has been widely used in several reliability related work (e.g. (Amstadter, 1977), (Barlow & Proschan, 1975), (Henley & Kumamoto, 1981)).

The curve is built taking into consideration three aspects:

- An initial high, but decreasing, failure rate – resulting from early "Infant Mortality" failure;
- An intermediate and constant failure rate – resulting from constant (random) failures; and
- A final increasing failure rate: resulting from wear out failures.

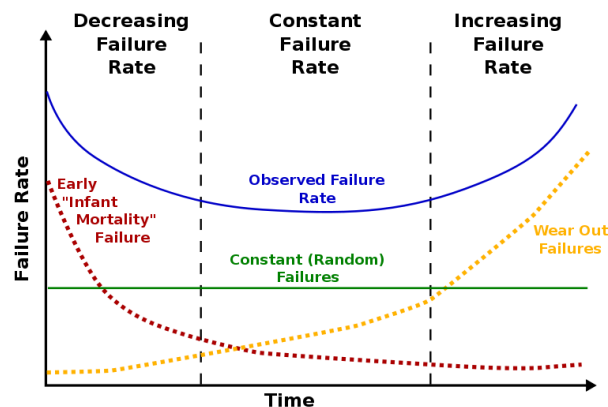


Figure 2.1. Bathtub Curve<sup>3</sup>

The "Bathtub Curve" is then generated by mapping the rate of early "infant mortality" failures when first introduced, the rate of random failures with a constant failure rate during its "useful life", and finally the rate of "wear out" failures as the product exceeds its design lifetime.

<sup>2</sup> MTTF is calculated from the total accumulated operating time divided by the number of failures during the same period

<sup>3</sup> Image taken from [http://commons.wikimedia.org/wiki/File:Bathtub\\_curve.jpg](http://commons.wikimedia.org/wiki/File:Bathtub_curve.jpg). As a work of the U.S. federal government, the image is in the public domain.

One could be tempted to think that the direct application of the Bathtub Curve would solve the problem of finding an appropriate maintenance strategy but, unfortunately this is not the case. In fact some studies concluded that it cannot be directly applied and, although it might be used successfully in some cases, it drives into an inefficient use of resources for most machines. Moreover, some critics have been made to the realism of the curve, especially in what concerns the foundations of the failure rates (Klutke, Kiessler, & Wortman, 2003).

Thus, instead of looking for pre-defined recipes as the Bathtub Curve, one should prefer to analyze failure-data and determine whether or not the assumed benefits of this technique are realized. In fact, if this analysis is not performed carefully, the normal result of using this technique is either unnecessary repairs or catastrophic failures. The reason is simple and it is related with the operational differences of each case, i.e. each plant has its own specific variables which affect the normal operating life of equipment.

The analysis of the specific variables is then useful for a complete understanding of the impact that maintenance costs have in the overall production costs.

### **2.1.3. Predictive maintenance and Condition-based maintenance**

The premise of Predictive Maintenance (PdM) is that regular monitoring of the actual mechanical conditions, operating efficiency, and other indicators of the operating conditions and process systems will provide the data required to ensure the maximum interval between repairs as well as minimize the number and cost of unscheduled outages created by failures (Mobley, 2002). However this concept can still be extended since a comprehensive predictive maintenance management program is a mean of improving productivity, product quality, and overall effectiveness of industrial plants.

Predictive maintenance is a condition-driven preventive maintenance program. Instead of relying on plant average statistics (like the above mentioned Bathtub Curve) it uses direct monitoring to determine mean-time-to-failure or loss of efficiency for plant equipment. This data provides the maintenance personnel with information for an appropriate scheduling of the maintenance activities. This way the complete monitoring of the equipment contributes for an early detection of problems which can then be minimized and major repairs can usually be prevented.

Despite its obvious advantages this method has also a couple of drawbacks namely the cost of installation of a complete monitoring system. In fact the initial cost of implementing a predictive maintenance program can be very high and a complete analysis on the need for the investment must be made at the beginning phase. Another drawback results from the first one and is related with the difficulty of identifying the most important parts of the plant to focus the monitoring investment in there. Once again a complete study of the plant is the only way to solve this important question.

### 2.1.3.1. Condition-based maintenance

As soon as the limitations of preventive maintenance were recognized, the concept of condition-based maintenance (CBM) started to gain visibility. Although it was firstly introduced in the late 1940s in Rio Grande Railway Company it was during the 1970 decade that the technology emerged.

CBM is based on the development of machine diagnostic techniques (Takata, et al., 2004) and preventive actions are taken when symptoms of failures are recognized through monitoring or diagnosis. Therefore, CBM enables taking the proper actions at the right timing to prevent failures, if there is a proper diagnostic technique.

In general, condition-based maintenance is a technique which assesses the overall health of equipment by regularly measuring and analyzing the data gathered during operation. With robust, regular and consistent data, maintenance personnel can be able to schedule maintenance actions based on the knowledge of the condition of the equipment rather than on its the running hours (UBM) or time alone (TBM) (Marques & Neves-Silva, 2008).

Condition-based maintenance methods assume that the condition of a machine is monitored and maintenance is only undertaken if conditions warrant it.

It uses real-time data to prioritize and optimize maintenance resources which allows determining the equipment's health and acting only when maintenance is actually necessary. This method equally applies to manufacturing processes where the settings of some machines or components may need to be altered based on the monitored condition of the process.

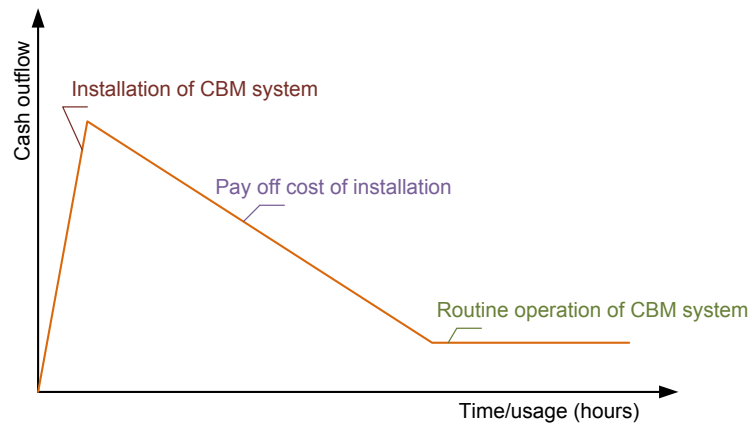


Figure 2.2. Typical cost of CBM installation and operation<sup>4</sup>

However, pure CBM is not always the best method of maintenance, especially from the perspective of cost effectiveness. In fact and as already mentioned for predictive maintenance, the initial cost of CBM can also be extremely high depending on the type of equipment we want to monitor and the level of certainty we demand. Therefore, it is important to decide the importance

<sup>4</sup> Adapted from Mobley (Mobley 2002)

of the investment before adding CBM to all equipment. Figure 2.2 presents the typical cost graphic for CBM installation.

Nonetheless, as instrumentation and information systems tend to become cheaper and more reliable, CBM may become an important tool for running a plant as supported by Figure 2.3 which presents the typical curve of potential saving for installation of a CBM system. More optimal operations will lead to lower production cost and lower use of resources. In turn lower use of resources may be one of the most important differentiators in a future where environmental issues become more important by the day.

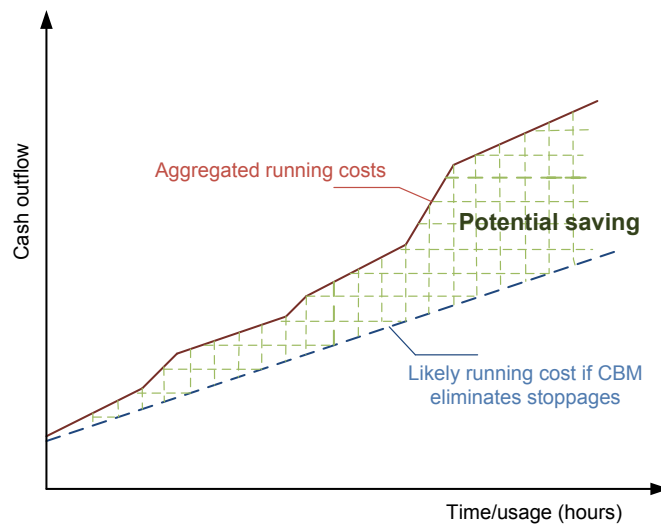


Figure 2.3. Typical potential savings produced by CBM<sup>5</sup>

The use of prediction techniques can help complementing CBM adding to it the capacity of, based on the trends of the monitored signals, predict the occurrence of failures. From a practical point of view the prediction is made based on sensory information collected by appropriate monitoring techniques. These sensorial signals often exhibit characteristic patterns known as degradation signals which can be used to predict a system's remaining lifetime (Nelson, 1990). In fact a conditioned based maintenance strategy using predictive technologies can be the key to extending equipment life, reducing maintenance costs and increasing asset utilization.

An example of the work being produced using a combination of these two techniques is the one developed by Kaiser (Kaiser & Gebraeel, 2009) presenting a sensory-updated degradation-based maintenance (SUDM) policy, which extends conventional CBM by combining population-based degradation characteristics with real-time monitoring information to predict the remaining lifetime.

The results of the methods developed so far are quite promising for achieving a more structured contribution for a life cycle approach for maintenance.

<sup>5</sup> Adapted from Mobley, 2002

## **2.2. Maintenance improvement methodologies**

The management methods described above are seen as the most used in industrial environments and, in many of them, preventive maintenance is still the one that prevails. Nevertheless, the use of some methods that combine aspects of maintenance with reliability has also gained importance in the recent years. These methods normally aggregate a maintenance method together with some strategy defined by the company in terms of the goals to be achieved. In the next sections some of the most well known methods in this area will be described.

### **2.2.1. Total productive maintenance**

This concept has its roots on the so-called total quality management (TQM). TQM evolved as a direct result of Dr. W. Edwards Deming's influence on Japanese industry. As a statistician, Dr. Deming initially began to show the Japanese how to use statistical analysis in industry and how to use the resulting data to control quality during the industrial process. The initial statistical procedures and the resulting quality control concepts fuelled by the Japanese work ethic soon became a paradigm for Japanese industry (X-Stream LEAN, 2007) and total productive maintenance (TPM) appeared as a new maintenance philosophy.

TPM is not a simple maintenance management program, but more like a partnership that involves production, maintenance, engineering and technical personnel to improve overall equipment effectiveness (OEE). Its goals are aggregated in terms of capacity, product production and total production cost, and it is aimed at improving the following six aspects:

- **Equipment breakdowns:** Results in equipment downtime for repairs. Costs can include downtime (and lost production opportunity or yields), labour, and spare parts.
- **Setup and adjustment slowdowns:** Results in lost production opportunity (yields) that occurs during product changeovers, shift change or other changes in operating conditions.
- **Idling and short-term stoppages:** Results in frequent production downtime, from zero to ten minutes in length, which are difficult to record manually. As a result, these losses are usually hidden from efficiency reports and are built into machine capabilities but can cause substantial equipment downtime and lost production opportunity.
- **Reduced capacity:** Results in productivity losses when equipment must be slowed down to prevent quality defects or minor stoppages. In most cases, this loss is not recorded because the equipment continues to operate.
- **Quality related losses:** Results in off-spec production and defects due to equipment malfunction or poor performance, leading to output which must be reworked or scrapped as waste.
- **Start-up/restart losses:** Results in wear and tear on equipment that reduces its durability and productive life span, leading to more frequent capital investment in replacement equipment.

To cope with these six losses TPM developed a model based on five pillars:

1. **Improve equipment effectiveness:** find out which are the causes for equipment inefficiency;
2. **Involve operators in daily maintenance:** not necessarily in the maintenance actions but planning, scheduling, etc.
3. **Improve maintenance efficiency and effectiveness:** the involvement of several departments results in an improved overall process;
4. **Educate and train personnel:** operators and maintenance personnel cooperate and learn how to better operate and maintain the equipment;
5. **Design and manage equipment for maintenance prevention:** suggestions from operators and maintenance personnel are introduced in the design process of equipment.

Overall equipment effectiveness (OEE) is then the benchmark used for TPM programs and may be expressed by the formula:

$$OEE = Availability \times Performance\ rate \times Quality\ rate$$

The evaluation of operating and maintaining costs enables an improved knowledge about equipment behaviour throughout its life cycle. In a long term perspective it is expected that these costs are minimized which will drive into a maximized OEE.

### 2.2.2. Reliability-centered maintenance

The term reliability-centered maintenance (RCM) was firstly introduced in the late 1970s by a group of American engineers from United Airlines to describe a process used to determine the optimum maintenance requirements for aircraft (Cotaina, et al., 2000).

The initial basic premise of RCM is that machines have a finite useful life and they all will eventually fail. Starting from this belief RCM is based on the typical Potential – Functional (P-F) curve (see Figure 2.4) and uses Failure Modes and Effects Analysis (FMEA<sup>6</sup>) and Weibull<sup>7</sup> distribution analysis to anticipate when these failures will occur. Both analyses assume proper design, installation, operation and maintenance of a plant to define their probability functions.

However, with the advent of predictive maintenance technologies this premise disappeared. The ability to detect minor deviations and react accordingly to avoid deterioration has contributed to effectively prevent degradation and resulting failures. Today RCM is defined by the technical standard SAE JA1011 (Netherton, 1998) which sets the minimum criteria that any process should meet before it can be called RCM.

---

<sup>6</sup> Failure Modes and Effects Analysis (FMEA) is a theoretical method based on probabilities to analyze the potential failure modes and classify them severity and likelihood.

<sup>7</sup> Weibull distribution is a continuous probability distribution which, if appropriately parameterized, gives a distribution for which the failure rate is proportional to a power of time.

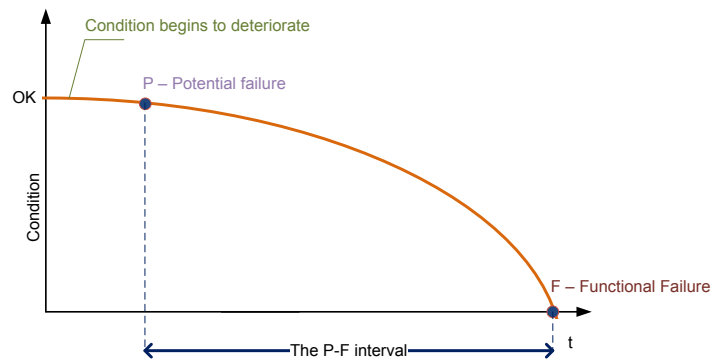


Figure 2.4. The P-F curve

The process starts with the following 7 questions, which should be worked out in the order they are listed:

1. What is the item supposed to do and its associated performance standards?
2. In what ways can it fail to provide the required functions?
3. What are the events that cause each failure?
4. What happens when each failure occurs?
5. In what way does each failure matter?
6. What systematic task can be performed proactively to prevent, or to diminish to a satisfactory degree, the consequences of the failure?
7. What must be done if a suitable preventive task cannot be found?

In this new context, RCM enables the definition of a complete maintenance regime, looking at maintenance as a mean to maintain the functions a user may require of machinery in a defined operating context. Reliability-Centered Maintenance can be used to create a cost-effective maintenance strategy to address dominant causes of equipment failure. It is maintained by means of constant review and update of equipment parameters based on the experience gained over time and, therefore, is a systematic approach for defining a routine maintenance program composed of cost-effective tasks that preserve important functions. RCM focuses on the, so many times, scarce economic resources on those items that would cause the most disruption if they were to fail.

### 2.2.3. Risk based inspection / Risk based maintenance

From the latter half of the 1980s, the importance of selecting proper maintenance strategies has been acknowledged in various areas. Together with the previously described RCM, Risk Based Inspection (RBI), or Risk Based Maintenance (RBM), (ASME, 1994), are the most well known methodologies for this purpose.

The approach is based on risk to prioritize and plan inspection used in engineering industries, and predominant in the oil and gas industries. This type of inspection planning analyses the probability (or likelihood) and consequence of failure of an asset to calculate its risk of failure. The level of

risk is then used to develop a prioritized inspection plan outlining the type and frequency of inspection for the asset.

Items with high probability and high consequence (i.e. high risk) are given a higher priority for inspection than items that are high probability but for which failure has low consequences. Thus, this strategy allows for a rational investment of inspection resources.

The idea is to assist a company in the selection of a cost effective and appropriate maintenance and inspection tasks and techniques, as well as to optimize such efforts while shifting from a reactive to a proactive maintenance regime. Also the approach aims at producing an auditable system, to give an agreed “operating window”, and to implement a risk management strategy. According to McCalley, Voorhis and Jiang (McCalley, Voorhis, & Jiang, 2003), the approach has the following specific attributes:

- The condition information is used to estimate equipment failure probability.
- Failure consequences are estimated and utilized in the prioritization of the maintenance tasks.
- Equipment failure probability and consequence at any particularly time are combined into a single metric called “risk”.
- Equipment risk may be accumulated over a time interval (e.g., a year or several years) on an hour-by-hour basis to provide a cumulative risk associated with each piece of equipment.
- The prioritization (and thus selection) of maintenance tasks is based on the amount of reduction in cumulative risk that is achieved by each task.
- Scheduling and selection of maintenance tasks is performed at the same time (using optimization algorithms) since the amount of reduction in cumulative risk depends on the time a maintenance task is implemented.

The reasons or drivers to adopt a risk based approach to the management of a plant can be varied. However, it is generally agreed that one of the main drivers is to optimize the costs of complying with statutory obligations. When properly carried out, this approach is able to manage the likelihood and consequences of plant failure at an acceptable level, and thereby avoid unreasonable risks.

#### **2.2.4. Life cycle maintenance**

The new challenges of industrial companies have changed the way maintenance was traditionally viewed. Currently maintenance is seen as a complementary way of preserving the condition of the plant so as to fulfil its required functions throughout its life cycle. Maintenance is then an important part of a life cycle management approach, whose main purpose is to enhance the sustainability of the plant life cycle. The term "life cycle maintenance" tries to capture the degree of importance that maintenance should have when plants are seen from a life cycle perspective (Rockwell Automation, 2004).



There are two reasons why it is necessary to control the conditions of the plant: changes in the condition of the plant due to deterioration, and changes in the production needs due to new demands from the market. Both changes generate gaps between the required function and the realized function and, if correctly planned, maintenance is executed to compensate these gaps by means of treatment or upgrading. According to Takata et. al (Takata, et al., 2004) to achieve this, maintenance should involve the following activities:

- **Maintainability design:** improving design based on evaluating maintainability in the plant development phase and providing the design data for maintenance strategy planning and maintenance task control.
- **Maintenance strategy planning:** Selecting a maintenance strategy appropriate to each part of the plant.
- **Maintenance task control:** Planning and executing the maintenance tasks based on the selected strategy.
- **Evaluation of maintenance results:** Evaluating the results of maintenance to determine whether the maintenance strategy planning and maintenance task control are appropriate.
- **Improvement of maintenance and plant:** improving maintenance task control, maintenance strategy planning, and even plant design based on the evaluation of maintenance results.
- **Dismantling planning and execution:** Planning and execution of dismantling at the end of the plant life cycle.

In life cycle maintenance, the list of activities presented above must be managed in an effective way throughout the life cycle of the plant aiming to achieve the following objectives (Takata, et al., 2004):

- **Adaptation to various changes during life cycle:** During the plant life cycle, there could be various changes in the required functions, in the operating environment, in the operating conditions and in the plant itself. Maintenance management should be flexible enough to adapt to these changes, because maintenance methods depend on these factors;
- **Continuous improvement of the plant:** In practice, it is impossible to design a perfect plant. Therefore, maintenance should include a mechanism for continuous improvement of the plant based on experience and knowledge acquired during its life cycle (and the life cycle of similar plants). This mechanism is also effective for functional upgrade of the plant to cope with shortening the plant life cycle due to rapid changes in market demands and technology development;
- **Integration of maintenance information:** For effective maintenance management, all information associated with maintenance should be integrated in such a way that it is available from any phase of the life cycle. In the development phase, for example, it is

essential to know the real operating situations and the problems encountered during past operations. On the other hand, it is necessary to have exact design data for maintenance strategy planning and maintenance task control.

Note that one of the critical aspects in life cycle maintenance is the assessment of the impact of the decisions made, since decisions are not free of risk. In fact, if one does not have an idea of the result a specific action in the overall plant behaviour or what costs are involved on it, becomes very hard to choose between different courses of action.

Thus, to overcome this issue the development of decision support systems aggregated to life cycle maintenance strategies enable to understand to what extent an action performed today influences the behaviour of the plant in the future.

### **2.3. Conclusions**

Over time companies began to realize that when equipment breaks down it always costs more, and take longer, to fix it than if it was maintained on a regular basis. Thus, companies started questioning this maintenance policy, understanding that is more cost effective shut down equipment for shorter periods that live with major breakdowns.

This view led companies to establish preventive maintenance strategies. These may be based on time elapsed between tasks or include some kind of data collection on equipment state allowing condition prediction.

The implementation of predictive and statistical techniques for monitoring equipment is a step forward on the companies' maintenance strategy helping to prevent a wide range of failures. From vibration analysis to temperature monitoring the information produced can, if properly used, reduce failures significantly.

At this point companies are able to discover problems before they evolve to a major breakdown which is already a great achievement. Freed of this concern it is possible to concentrate on refining the process by searching for what more can be improved. For example, the development of equipment failures analysis can be used to prevent future or repetitive problems. Figure 2.5 illustrates this evolution of maintenance towards a point where efforts should emphasize elimination of failures that require maintenance.

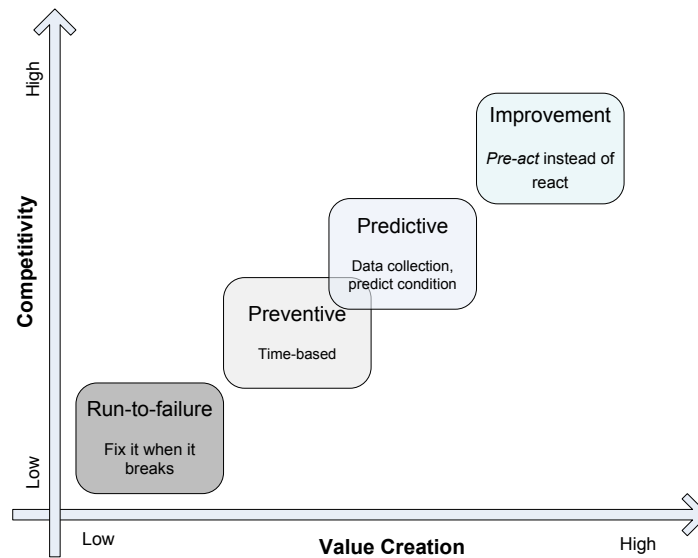


Figure 2.5. Maintenance evolution paradigm<sup>8</sup>

Nowadays companies are evolving for a stage of maturity where they recognise that producing high quality products for the lowest price, considering sustainability issues, is their main goal. From the product design phase to final production, competitiveness demands the use of the most advanced techniques and processes.

The author of this thesis defends that this stage of maturity should also congregate the aspects related with equipment maintenance recognising this as a primary issue on a company life cycle.

<sup>8</sup> Adapted from Mobley (Mobley 2002)



## 3. *Risk analysis and Decision Support Systems*

---

### 3.1. *Risk analysis*

*“Risk analysis is the science of evaluating health, environmental, and engineering risks resulting from past, current, or anticipated future activities. The use of these evaluations include providing information for determining regulatory actions to limit risk, presenting scientific evidence in legal settings, evaluating products and potential liabilities [...]. Risk analysis is an interdisciplinary science that relies on epidemiology and laboratory studies, collection of exposure and other field data, computer modelling, and related social and economic and communication considerations. In addition, social dimensions of risk are also addressed by social scientists [...]”*<sup>9</sup>

From the above definition it is possible to conclude that risk analysis is performed in various contexts since its application is as wide as any field where decision and uncertainty are present. In the scope of industrial plants, which are normally complex engineered technological entities, risk analysis is performed by means of probabilistic risk assessment which is seen as a systematic and comprehensive methodology to evaluate this specific type of risk.

#### 3.1.1. *Historical overview*

To make good choices, companies must be able to calculate and manage the attendant risks. Until a few hundred years ago risk management consisted of faith and hope, since until that time humankind’s understanding of numbers was too under developed (Bernstein, 1998).

As explained by Buchanan (Buchanan & O’Connell, 2006) during the Renaissance period, scientists and mathematicians such as Girolamo Cardano were obsessed with probability and

---

<sup>9</sup> Definition from the International Journal of Risk Analysis, an official publication of the Society for Risk Analysis.

games of chance. Around 1630, Blaise Pascal and Pierre de Fermat developed a way to determine the likelihood of each possible result of a simple game but it was only in the following century, with Daniel Bernoulli (Bernoulli, 1738 (1954)) taking up the study of random events, that the scientific basis for risk management took shape.

Bernoulli focused not on events themselves but on the human beings who desire or fear certain outcomes to a greater or lesser degree. His intent was to create mathematical tools that would allow anyone to “estimate his prospects from any risky undertaking in light of specific financial circumstances”. In other words, given the chance of a particular outcome, how much are you willing to bet?

In the nineteenth century, other scientific disciplines started to contribute for risk analysis such as the geodesic and astronomical research brought by Carl Friedrich Gauss (Gauss, 1809) as well as the concept of regression from Francis Galton (Galton, 1886). But it wasn't until after World War I that risk gained a fundamental position in economic analysis. In 1921, Frank Knight (Knight, 1921) distinguished between risk, when the probability of an outcome is possible to calculate, and uncertainty, when the probability of an outcome is not possible to determine. This called the attention insurance companies.

Around two decades later, John von Neumann and Oskar Morgenstern (Neumann & Morgenstern, 1947) laid out the fundamentals of game theory, which deals in situations where people's decisions are influenced by the unknowable decisions of “live variables”. Nowadays, companies try to know as much as is humanly and technologically possible, deploying modern techniques as derivatives, scenario planning, business forecasting, and real options. Nevertheless, a great amount of uncertainty remains present and choosing is in many situations a real hard task.

### **3.1.2. Probabilistic risk assessment**

The risk process begins with two fundamental elements: a need on the part of an individual or a group, and a vision held by the person or group that will implement the stochastic solution (Koller, 1999). Various reasons can be the trigger for companies to perform risk assessment, namely:

- Need to consider a range of possibilities rather than a single-value answer;
- Uncertainty associated to various options.

Considering the above mentioned reasons it has become obvious that risk assessment is also a powerful tool for decision making and portfolio management.

Risk is then a concept that denotes a potentially negative impact to an asset or some characteristic of value that may arise from present process or future events. In everyday usage, "risk" is often a synonym of "probability" of a loss or threat. Nevertheless, risk should combine the probability of an event occurring with the impact that the event would have and with its different circumstances, ( (Holton, 2004), (Stamatelatos, 2000)).

Thus, risk  $R$  is defined as the result of the product between the probability of occurrence,  $p$ , of a specific incident  $E$  and the impact  $I$  of that incident (in money or injuries), i.e.:

$$R = p(E) \times I(E) \quad (3.1)$$

When there is a change on normal status, the information provided by the plant allows the probability of an incident occurrence to be estimated. Also, the estimated possible consequences of the incident allow estimating the level of risk associated. The functions that describe both incident probability and losses are supposed to be known with some level of uncertainty. This information is collected during set-up phase and during operation using workers knowledge in order to compute the risk.

This probabilistic risk approach tries to give answer to the following basic questions:

1. What can go wrong with the plant under observation, or what are the initiators or initiating events (undesirable starting events) that lead to adverse consequence(s)?
2. What and how severe are the potential detriments, or the adverse consequences that the plant may suffer as a result of the occurrence of the initiator?
3. How likely to occur are these undesirable consequences, or what are their probabilities or frequencies?

The establishment of a classification system for the consequences that may be involved is the first step of the process. Additionally some thresholds levels may also be defined considering the type of application since normally are not universally adequate. This way, a common classification of levels of risk in a qualitative scale, based on the consequences of the incident is here presented (U.S. DEPARTMENT OF TRANSPORTATION, 1998):

- **Hazardous:** reduces the capability of the system or the operator ability to cope with adverse conditions to the extent that there would be - large reduction in safety margin or functional capability;
- **Major:** reduces the capability of the system or the operators to cope with adverse operating conditions to the extent that there would be - significant reduction in functional capability;
- **Minor:** does not significantly reduce system safety. Actions required by operators are well within their capabilities - Slight reduction in functional capability.

Also qualitatively, the probability of occurrence  $p$ , also called likelihood of occurrence, may be classified in the following manner where  $p_0$ ,  $p_1$  and  $p_2$  are to be defined during system customization and  $p_0 > p_1 > p_2$ :

- **Probable:**
  - Qualitative: Anticipated to occur one or more times during the entire system/operational life of an item;

- Quantitative: Probability of occurrence given an operational time is greater than  $p_0$ :

$$p(E|\Delta t) > p_0 \quad (3.2)$$

- **Remote:**

- Qualitative: Unlikely to occur to each item during its total life. May occur several times in the life of an entire system;
- Quantitative: Probability of occurrence given an operational time is less than  $p_0$ , but greater than  $p_1$ :

$$p_0 > p(E|\Delta t) > p_1 \quad (3.3)$$

- **Extremely Remote:**

- Qualitative: Not anticipated to occur to each item during its total life. May occur a few times in the life of an entire system;
- Quantitative: Probability of occurrence given an operational time is less than  $p_1$  but greater than  $p_2$ :

$$p_1 > p(E|\Delta t) > p_2 \quad (3.4)$$

- **Extremely Improbable:**

- Qualitative: So unlikely that it is not anticipated to occur during the entire operational life of an entire system;
- Quantitative: Probability of occurrence given an operational time is less than  $p_2$ :

$$p(E|\Delta t) < p_2 \quad (3.5)$$

The use of Event Tree Analysis (ETA), for assessing the probability of occurrence is one of the most recognized methods to perform risk analysis of technological systems and identify improvements in protection systems and other safety functions (Rausand & Høyland, 2004).

ETA is an inductive procedure that shows all possible outcomes resulting from an accidental (initiating) event, taking into account safety barriers (whether installed, functioning or not) and additional events and factors. By studying all relevant accidental events, ETA can be used to identify all potential accident scenarios and sequences in a complex system. Design and procedural weaknesses can be identified, and probabilities of the various outcomes from an accidental event can be therefore determined.

An Event Tree starts from an undesired initiator (loss of critical supply, component failure, etc.) and follows possible further system events through a series of final consequences (see Figure 3.1). As each new event is considered, a new node on the tree is added with a split of probabilities taking either branch. In the end the probabilities of a range of 'top events' arising from the initial event can then be seen (Russel & Norvig, 2003). These events correspond to the severity of the occurrence (Hazardous, Major or Minor) and by this the risk level is established.



The method includes the consideration of the following:

- Identification and delineation of the combinations of events that, if they occur, could lead to an accident (or other undesired event);
- Estimation of the chance of occurrence for each combination; and
- Estimation of the consequences associated with each combination.

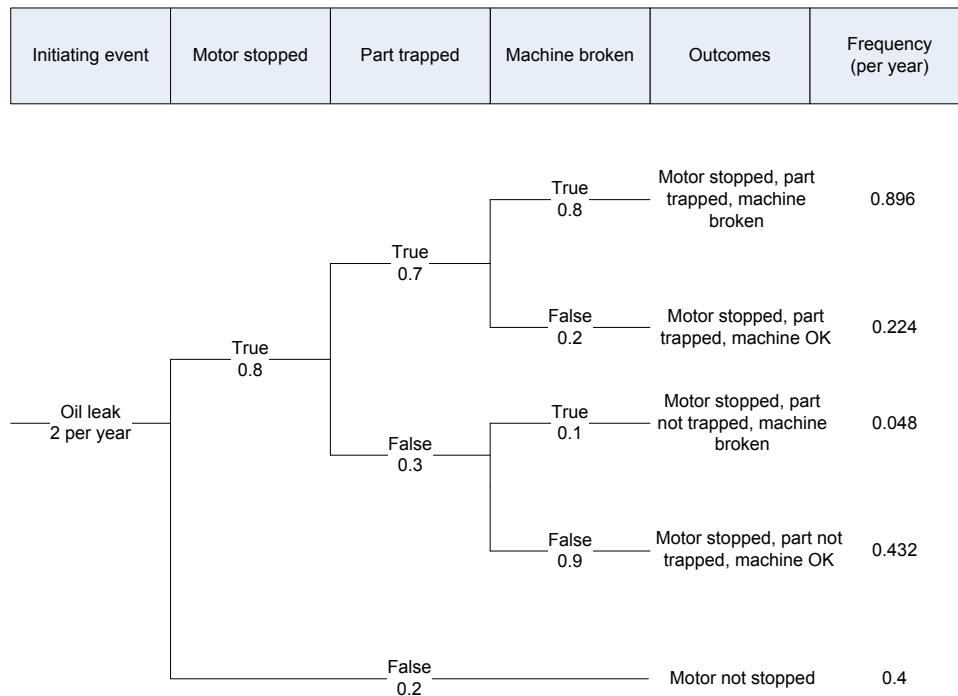


Figure 3.1. Example of an Event Tree

From what was exposed it is possible to conclude that probabilistic risk assessment is one of the most powerful tools, not only to assess risk, but also to support decision making when risk is an important aspect to be considered. The problem with these models is that they do not include any procedure to deal with human option, i.e. to incorporate the effects that even though risk analysis suggested an option, the final decision can be different one. Indeed, when dealing with industrial plant, these cases are not as scarce as one could be tempted to think, especially in cases where no specific protocol procedures exist.

### 3.1.3. Decision under risk

When a decision has to be made and the exact outcome of that decision is not known then decision maker faces a decision under risk problem. In this case, the different outcomes are normally associated to probabilities of occurrence, which are established by real observations or simulations of the events. These outcomes represent the expected value that a specific variable may take accordingly with the decision made. When using the Expected Value (EV) criterion, the expected value of that variable is calculated by means of a weighted average of all possible values that a random variable can take.

This would work perfectly if humans were not involved in the process. However, decisions are most of the times influenced by people's expectations of what may be the result of them. To incorporate this idea of decision connected to expected results the Expected Utility Theory (EUT) was originally proposed by Bernoulli in 1738 (Bernoulli, 1738 (1954)) and was revisited by John von Neumann and Oskar Morgenstern in 1947 (Neumann & Morgenstern, 1947).

EUT's main concept relies on the idea that a person's preference regarding uncertain outcomes may be represented by a function of payouts. The theory states that if an individual always chooses his/her most preferred alternative available, then the individual will choose one gamble over another if, and only if, there is a utility function such that the expected utility of one exceeds that of the other. The expected utility of any gamble may be expressed as a linear combination of the utilities of the outcomes with the weights being the respective probabilities.

Utility functions are normally continuous functions and are referred to as "von Neumann–Morgenstern utility functions". The attitude to risk is directly related to the curvature of the utility function i.e. risk neutral individuals have linear utility functions, while risk seeking individuals have convex utility functions, and risk averse individuals have concave utility functions. Therefore, the degree of risk aversion can be measured by the curvature of the utility function (see Figure 3.2).

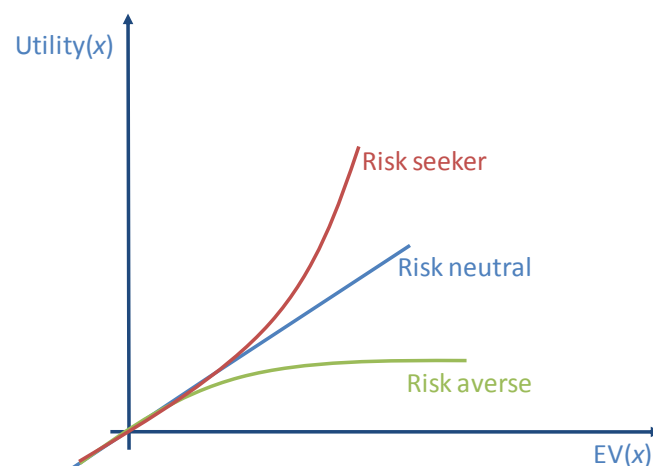


Figure 3.2. Examples of utility functions

Once again these would work just fine if humans were not many times influenced by external factors that condition their choices (e.g. previous experience, third part experience, problem presentation, etc.). In fact it became clear over some time that human beings tend to weight in a different way gains and losses using a specific reference values. In cases where gains are expected they tend to prefer lower risks for moderate to high probabilities and higher risks when the probabilities are low. On the other hand when losses are expected these preferences are inverted (see Table 3.1).

Table 3.1. Expectancy Theory table

	Low probabilities	Moderate – high probabilities
Gains	Risk preference	Risk aversion
Losses	Risk aversion	Risk preference

Based on this theory a set of studies were performed which constituted the basis to develop the Prospect Theory (PT) (Kahneman & Tversky, 1979). These studies investigated apparent anomalies and contradictions in human behaviour. The conclusions were divided into two sub-groups:

- **Gains:**
  - 84% of individuals prefer to go for a certain gain of 500 than to bet in a 1000 gain with 0.5 probability: this reflects the risk aversion for moderate probabilities;
  - 72% of individuals prefer to accept the challenge of winning 5000 with 0.001 probability against certain 5: this reflects the risk preference when in a low probabilities scenario;
- **Losses:**
  - 69% of individuals accept a 0.5 probability when referring to a loss of 1000 against a certain loss of 500: this reflects the risk preference for moderate probabilities;
  - 83% of individuals prefer to lose 5 against a 0.001 probability of losing 5000: this reflects the risk aversion for low probabilities.

Kahneman and Tversky found empirically that humans tend to underweight outcomes that are merely probable in comparison with outcomes that are obtained with certainty. Additionally they also concluded that people generally discard components that are shared by all prospects under consideration. Under prospect theory, value is assigned to gains and losses rather than to final assets; also probabilities are replaced by decision weights.

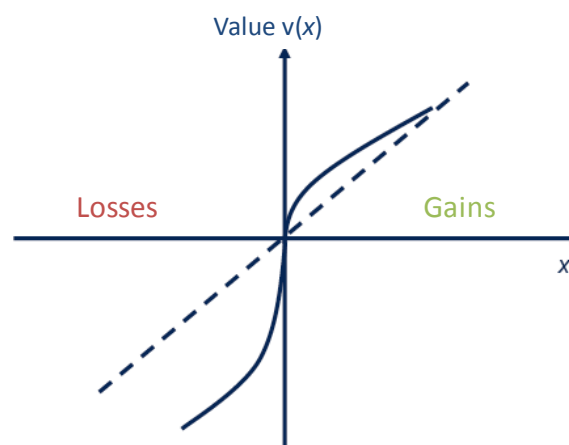


Figure 3.3. Example of value function

Thus while in expected utility theory the utility function is necessarily linear in the probabilities, in the prospect theory value function is not. Additionally, whereas utility is dependent on final payout, value is defined in terms of gains and losses. The value function is defined on deviations from a reference point and is normally concave for gains (implying risk aversion), commonly convex for losses (risk seeking) and is generally steeper for losses than for gains (loss aversion) (see Figure 3.3). The gradient for the value function for  $x < 0$  is superior than the one verified for  $x > 0$ , evidencing the aversion to losses.

The decision weights result from people's tendency to weight differently low and high probabilities. In fact for low probabilities people tend to overweight probability, believing that winning the lottery is easier than it actually is. On the other hand for high probabilities the opposite behaviour is observed, driving to underweighting the probability (see Figure 3.4).

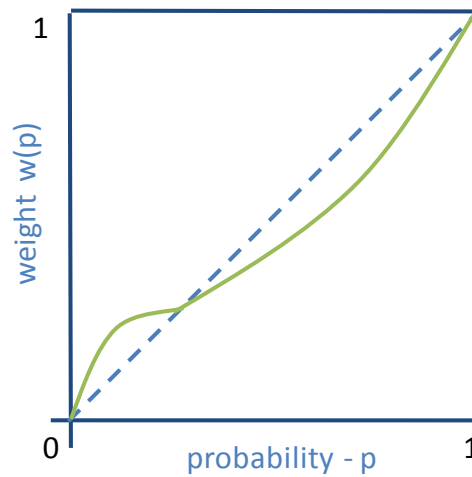


Figure 3.4. Weighting function

The value is calculated using:

$$V = \sum_{i=1}^n w(p_i) \times v(x_i) \quad (3.6)$$

Thus, and from a general point of view, EUT predicts that the better alternative will always be chosen, independently from the decision making scenario. On the other hand PT recognizes the information processing limits of human decision makers and their tendency toward satisfying decision behaviour. As such, it suggests that decision makers are often not consistent in their preferences and are subject to influence by the way alternatives are presented to them.

### 3.2. Decision support systems

In terms of decision theory, reaching a conclusion is choosing among a set of alternatives. Decision making is the study of the identification and choice among alternatives, based on the values and preferences of the decision maker. According to Harris (Harris, 2009) making a decision implies that there are alternative choices to be considered, and in such case the

objective is not only to identify as many of these alternatives as possible but to choose the one that best fits with the predefined goals, objectives, desires, values, and so on.

Some research using naturalistic methods show that, in situations with higher time pressure, higher stakes, or increased ambiguities, decision-makers tend to follow recognition primed decision approach, to fit a set of indicators into the expert's experience and immediately arrive at a satisfactory course of action without weighing alternatives (Nja & Rake, 2009). When trying to reach a decision, the main variation in the decision making process is related with criticality. This means that the process is affected by the urgency and by the risk associated to the situation. In fact, if the situation is critical it is highly unlikely to expect that the decision maker will follow the standard decision making process, step-by-step. In many situations decision makers choose to follow straightforward procedures, most of them intensively tested to ensure repeatability.

Due to the large number of considerations involved in many decisions, computer-based decision support systems have been developed to assist decision makers in considering the implications of various courses of thinking. These systems can help reduce the risk of human errors due to cognitive and temporal limitations ( (Diasio & Agell, 2009), (Holsapple & Whinston, 2000)).

All these aspects become clear when the decision must be made in a noisy environment or in a distressing situation, which are very common conditions in industrial plants. Thus, the use of systems that structure and correlate information are seen, nowadays, as being of key importance to the decision making process since they can help reducing the risk of human errors.

Decision Support Systems (DSS) are a class of knowledge based systems that, in very different manners, support decision making activities. As a very simple definition a DSS is a system for making decisions when decision is a choice between alternatives based on estimates of the values of those alternatives. In other words, supporting a decision means supporting a choice based on estimation, evaluation and/or comparison of alternatives. In practice, speaking of DSS mean speaking of computer applications that perform such a supporting role.

The concept of decision support has evolved from two main areas of research: the theoretical studies of decision making and the technical work on interactive computer systems.

Questions about who makes decisions and how they are made have shaped our world, especially when defining systems of government, justice, and social order. Additionally the panoply of fields that may be involved, from mathematics to sociology, from psychology to economics, demonstrates the level of complexity that decision making might reach. For this reason the development of Decision Support Systems has been a topic with great investment from research community especially with the advent of computer systems.

In the next sections aspects such as historical positioning and associated concepts will be described and analysed.

### 3.2.1. Reaching a decision: overview of the decision-making process

The history of decision-making strategies is not one of pure progress toward perfect rationalism. In fact human have always faced constraints when it comes to make optimal choices, namely: complex circumstances, limited time, inadequate mental computational power, etc. Some researchers defend that people would make economically rational decisions if only they could gather enough information (Simon, 1991), while others are able to identify factors that cause people to decide against their economic interest even when they know better (Tversky & Kahneman, 1981).

Damasio (Damasio, 1994) worked with brain-damaged patients to demonstrate that in the absence of emotion it is impossible to make any decisions at all. Despite its results some critics say that his object of study was not appropriate for the kind of conclusions he achieved.

From “humble decision making” (Etzioni, 1989) to “fast and frugal” heuristics (Gigerenzer, 1995) theorists have searched ways to achieve, if not optimal outcomes, at least acceptable ones. In a more rational perspective, several methods and approaches can be used to reach a decision. The following sections provide some insight about the problematic of decision making.

#### 3.2.1.1. *Elements on the decision process*

As explained before decision making is the study of identifying and choosing alternatives based on the values and preferences of the decision maker.

Structured rational decision making is an important part of all science-based professions, where specialists apply their knowledge in a given area for making informed decisions. For example, medical decision making often involves making a diagnosis and selecting an appropriate treatment. However the decision method is directly affected by the criticality, impact and importance of that decision. Concerning the resulting actions more intuitive decision prevails in highly critical cases, while more structured approaches are normally followed in less stressful ones.

Before explaining the complete process, and for a better understanding of all the vocabulary, a brief list of the most common expressions used in the decision domain is presented:

- **Alternative:** One of a number of possibilities from which one must be chosen. For decision making purposes this is used synonymously with possible or potential solution, option or choice.
- **Criterion (also referred to as “attribute”):** Is a rule or a principle for evaluating, testing, or discriminating among alternatives, and must be based on the goals. It is defined as objective measures of the goals to determine how well each alternative achieves the goals.
- **Decision:** The act or process of deciding; the result of the decision making process; the selection of one (and only one) alternative.

- **Decision Method:** The techniques used to support and justify the decision making process.
- **Goal:** A broad statement of intent. A goal goes beyond the minimum essential *must haves* to *wants* and *desires*. In mathematical form, a goal is an objective.
- **Problem:** A concise and unambiguous statement, agreed by all decision makers and stakeholders, identifying root causes, limiting assumptions, system and organizational boundaries and interfaces, and any stakeholder issues, together with the initial conditions and the desired conditions.
- **Requirement:** A condition that any acceptable solution to the problem must meet. A requirement spells out what the solution to the problem must do. In mathematical form, the requirement is a constraint describing the feasible (admissible) solution of the decision problem.

### 3.2.1.2. *Decision making process*

The process of reaching a decision is called decision making process and it begins when decision makers need to do something but they do not know what. Decision making is a reasoning process which can be rational or irrational, and can be based on explicit assumptions or tacit assumptions. Decision making is said to be a psychological construct. This means that although a decision cannot be seen, it can be inferred, from observable behaviour, that a decision has been made. It is a construction that imputes commitment to action. Thus, based on observable actions, one can assume that people have made a commitment to affect the action.

According to Baker (Baker, Bridges, Hunter, Johnson, Krupa, & Sorenson, 2002), decision making should start with the identification of the decision maker(s) and stakeholder(s) in the decision, reducing the possible disagreement about problem definition, requirements, goals and criteria. Then, a general decision making process can be divided into the following steps:

- **Step 1 - Define the problem:** This process must identify root causes, limiting assumptions, system and organizational boundaries and interfaces, as well as any stakeholder issues. The goal is to express the issue in a clear, "one-sentence" problem statement that describes both the initial and the desired conditions.
- **Step 2 - Determine requirements:** It is very important that even if subjective or judgmental evaluations may occur in the following steps, the requirements must be stated in exact quantitative form, i.e. for any possible solution it has to be decided unambiguously whether it meets the requirements or not.
- **Step 3 - Establish goals:** The goals may be conflicting but this is natural in practical decision situations.
- **Step 4 - Identify alternatives:** Any alternative must meet the requirements. If the number of the eligible alternatives is finite, it is possible to check one by one if it meets the requirements (the infeasible ones must be disregarded in further consideration). If the

number of the possible alternatives is infinite, the set of alternatives is considered as the set of the solutions fulfilling the constraints in the mathematical form of the requirements.

- **Step 5 - Define criteria:** Since the goals will be measured in the form of criteria, every goal must generate at least one criterion but complex goals may be represented only by several criteria. It can be helpful to group together criteria into a series of sets that relate to separate and distinguishable components of the overall objective for the decision.
- **Step 6 - Select a decision making tool:** There are several tools for solving a decision problem. The selection of an appropriate tool is not an easy task and depends on the concrete decision problem, as well as on the objectives of the decision makers.
- **Step 7 - Evaluate alternatives against criteria:** Depending on the criterion, the assessment may be objective (factual), with respect to some commonly shared and understood scale of measurement (e.g. money), or can be judgmental, reflecting the subjective assessment of the evaluator. After the evaluations, the selected decision making tool can be applied to rank the alternatives or to choose a subset of the most promising alternatives.
- **Step 8 - Validate solutions against problem statement:** The alternatives selected by the applied decision making tools have always to be validated against the requirements and goals of the decision problem.

At the end of the process, if well applied, a decision should arise resulting from the most promising alternative validated against all the necessary constraints.

### **3.2.1.3. Nature of decision making**

There are several approaches for classifying decision making. However, besides the existing diversity, the one used in this dissertation is ultimately inherent to all the others. From a general perspective, decision making can be divided into two types:

- Rational, logical or quantitative; and
- Irrational, subjective or qualitative.

The first one is the objective and uses a logical sequential process to reach a decision or a quantitative metric. The second one is associated with non-measurable criteria and its result might to be not so easy to justify. Thus, and for the sake of validation, the goal should be to make any decision making as objective as possible. Nevertheless, cases exist where this approach might not be possible due to the lack of information and knowledge about the problem.

One way of solving the problem of subjectivity is to combine small subjective decisions to make an overall objective decision. The idea is to make subjective judgements on very small increments of the total decision. In fact, it is widely accepted that human ability to make small simple judgements is much higher than the ability to make large complex judgements. Once the subjective feelings and opinions are rated by putting numbers to them, it is possible to make a subjective decision as objective as possible.



After completing the identification of the alternatives, and conclude about the nature of the decision making there is the need to identify the criteria that will help us selecting the best option. At this point a new question must be made: is the decision problem ruled by only one criterion or by several?

This leads to another categorization of the decision making problem:

- single criterion decision making, and
- multi-criteria decision making.

In the first case, decision problems may have a single criterion or a measure that aggregates different values, e.g. cost. In these cases the decision problem may be solved simply by determining the alternative with the best value of the single criterion or aggregate measure. This approach results in having the classic form of an optimization problem where the objective function is the single criterion and the constraints are the requirements on the alternatives.

In the second case the problems are characterized by having a multiple but finite number of criteria. Here a new division needs to be made:

- **Cases that belong to the field of multiple criteria optimization:** cases involving decisions where the number or criteria is finite and the alternatives are infinite or given in implicit form. According to Steuer (Steuer, 1986) here the focus is on applying mathematical algorithms to identify alternatives that are optimal or efficient, under certain constraints, with respect to a few objectives that are expressed mathematically using decision variables. These are usually continuous therefore most problems have infinitely many alternatives which are defined by distinct combinations of values for decision variables. Additionally, techniques of multiple criteria optimization can also be used when the number of feasible alternatives is finite but they are given only in implicit form.
- **Cases that are part of the so called multi-criteria decision analysis:** cases involving decision making problems where the number of the criteria and alternatives is finite, and the alternatives are given explicitly (see Annex A for additional details about multi-criteria decision making methods). To solve these problems decision makers normally use a number of techniques to support them in the identification, comparison and evaluation of the existing alternatives. This process is performed taking into consideration the diversity of criteria, which are, in most of the cases, conflicting ones.

A standard feature of multi-criteria decision analysis methodology is the decision table, shown in Figure 3.5, where each row belongs to a criterion,  $C$ , and each column to the performance of an alternative,  $\mathcal{A}$ .

The score  $a_{ij}$  describes the performance of alternative  $\mathcal{A}_j$  against criterion  $C_i$ .

Additionally,  $w_1, \dots, w_m$  are weights assigned to the criteria, where  $w_i$  reflects the relative importance of criteria  $C_i$ . The weights of the criteria are usually determined on subjective basis, since they represent the opinion of a single decision maker or synthesize the opinions of a group of experts using a group decision technique as well.

Finally, the values  $x_1, \dots, x_n$  associated with the alternatives in the decision table are used in the methods based on the Multi-Attribute Utility Theory (MAUT) and are the final ranking values of the alternatives.

$$\begin{array}{cc}
 & \begin{array}{ccc} x_1 & \dots & x_n \\ A_1 & \dots & A_n \end{array} \\
 \begin{array}{cc} w_1 & C_1 \\ \vdots & \vdots \\ w_m & C_m \end{array} & \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}
 \end{array}$$

Figure 3.5. The decision table

As discussed by Roy (Roy, 1996), three fundamental problems can be applied to the assessment of a set of alternatives  $\mathcal{A} = \{\mathcal{A}_1, \mathcal{A}_2, \dots\}$ .

- **Choosing:** choose the best alternative  $\mathcal{A}_i$ .
- **Sorting:** sort the alternatives of  $\mathcal{A}$  into relatively homogeneous groups, which can then be arranged in preference order.
- **Ranking:** rank the alternatives of  $\mathcal{A}$  from best to worst.

The multiple criteria optimization is mainly focused on the first fundamental problem, i.e. choosing, while the multi-criteria decision analysis considers the three fundamental problems identified.

In the last decades, many multi-criteria decision analysis methods have been proposed for choosing and ranking. Examples include the MAUT elaborated by Keeney and Raiffa (Keeney & Raiffa, 1976), the Analytic Hierarchy Process (AHP) proposed by Saaty (Saaty T. L., 1980), and the Outranking methods proposed by Roy (Roy, 1968).

In recent years research efforts seem to be concentrated in finding new approaches for solving the sorting problem. As a decision analysis method, sorting is a prescriptive approach to assist individuals making wise classification decisions, involving the decision maker's preferences. Examples of the developments made are the results presented by Doumpos and Zopounidis ( (Doumpos & Zopounidis, 2001), (Doumpos & Zopounidis, 2002)) related with the efficiency of multi-criteria decision analysis classification approaches and the role of preference disaggregation when performing classification. Doumpos and Zopounidis ( (Doumpos & Zopounidis, 2002), (Doumpos & Zopounidis, 2002)) summarize many applications of sorting problems in multi-criteria decision analysis. Additionally, Chen (Chen, Kilgour, & W., 2006) proposed case-based distance methods to solve sorting problems, and Malakooti and Yang (Malakooti & Yang, 2004) proposed clustering of multiple criteria alternatives aiming to decrease the set of alternatives and the number of criteria and, by this, simplify sorting problem.

In any decision situation the decision making process is always driven by the existing knowledge about a specific problem. In fact, if no knowledge exists then, in most cases, decisions can only be made based in intuition. Nevertheless, the lack of knowledge does not necessarily mean the lack of objectivity. In fact, a problem may be highly subjective but, if the people involved in the

decision process are well aware of its nature, i.e. have past experience on dealing with similar situations, then it is likely that they reach a consistent decision. The key point here resides in the fact that knowledge is the base of the decision process and, normally, it exists in the mind of experts that are familiar with the processes requiring a decision.

Extensive work has been developed to facilitate the ways how this knowledge is acquired, processed and transformed into valuable information to be used in the decision making process, especially in cases where computational systems are used. Examples include the work developed by Choy et al. (Choy, Lee, Lau, & Choy, 2005), Vandaie (Vandaie, 2008) and Xu and Bernard (Xu & Bernard, 2011). Thus, if knowledge is the base for the decision making, and if this knowledge is somehow stored at peoples' mind, then it becomes clear who needs to be involved in the decision making process.

At this point it is important to make another categorization to distinguish:

- **decision making processes that are led by one person:** is normally the easiest one since the decision is concentrated in only one person so no conflicts will arise;
- **decision making processes that are led by a group of individuals:** decision depends on the agreement of a group of people, which constitutes the so-called Group Decision Making (GDM).

Due to its importance the study of groups and how they cooperate to reach a common goal, has been a field of great interest for a lot of researches. Its study gave its first steps around 1890, as part of the emerging field of social psychology. Follett (Follett, 1918) defended the value of conflict in achieving integrated solutions, but the breakthrough in understanding group dynamics occurred just after World War II when Lewin (Lewin, 1948), (Lewin, 1951)) proposed his influential "field theory", defending that actions are determined, in part, by social context and that even group members with very different perspectives will act together to achieve a common goal.

Over the next decades, knowledge about group dynamics and the care and feeding of teams evolved rapidly with the establishment of the circumstances under which group decision making is appropriate (Vroom & Yetton, 1973). The definition of the components required for successful teams (Belbin, 2004), the explanation of how groups exploit "external help" in the form of mediators and facilitators (Raiffa, 1982) and the suggestion that the most important decision may not be made by the team itself but rather by management about what kind of team to use (Drucker, 1986) are concepts that had arise after that.

From a practical point of view, a group decision situation involves multiple actors (decision makers), each with different skills, experience and knowledge relating to different aspects (criteria) of the problem. In a correct method for synthesizing group decisions, the competence of the different actors to the different professional fields has also to be taken into account. As stated by Kobashikawa (Kobashikawa, Hatakeyama, Dong, & Hirota, 2009) the essence of GDM is to find the alternative among the set of feasible alternatives, which best reflects the preferences of the group of individuals as a whole. Thus, identifying the preferences of the people involved in the

decision making process, i.e. the decision makers, is a key issue in GDM. Together with the identification process it is also essential to classify, somehow, those preferences aiming to confer them a value.

Nonetheless, as addressed by Nikolova (Nikolova, Shulus, Toneva, & Tenekedjiev, 2005) and by Tenekedjiev, Nikolova and Dimitrakiev (Tenekedjiev, Nikolova, & Dimitrakiev, 2004), is it widely accepted that humans have finite discriminating abilities which is reflected in their difficulty in provide singletons as preference values. Consequently GDM methods that only work with crisp preference values become hard to be used in real world, which is the reason for the amount of work developed using fuzzy algorithms to solve this problem.

GDM assumes that each actor considers the same sets of alternatives and criteria, as well as there is a special actor with authority for establishing consensus rules and determining voting powers to the group members on the different criteria. Keeney and Raiffa (Keeney & Raiffa, 1976) called this entity the Supra Decision Maker (SDM). The final decision is derived by aggregating (synthesizing) the opinions of the group members according to the rules and priorities defined by the SDM.

However it is important to stress that one of the most important aspects of group decision making is the diversity of positions and perspectives that should be envisaged when a team is selected. In fact, poor group decisions are often attributed to the failure to mix things up and question assumptions. Consensus is good, unless it is achieved too easily, in which case it becomes suspect driving into situations of so-called “groupthink” (Janis, 1972).

### **3.2.2. Brief history of decision support systems**

It is considered that the concept of Decision Support Systems (DSS) became an area of research of its own in the middle of the 1970s, before gaining intensity during the 1980s. During 1960s, some researchers tested the implementation of computerized quantitative models to assist in decision making and planning ( (Holt & Huber, 1969), (Raymond, 1966), (Urban, 1967)). Turban (Turban, 1967) made the first approach to mathematical models for decision making in industrial plants while Ferguson and Jones (Ferguson & Jones, 1969) reported the first experimental study using a computer aided decision system to investigate a production scheduling application running on an IBM 7094 computer.

Also during the 1960s, Scott-Morton's developments in building, implementing and testing an interactive, model-driven management decision system, showed how computers and analytical models could help managers making a recurring key business planning decision ( (Scott-Morton, 1967), (Scott-Morton & McCosh, 1968) (Scott-Morton & Stephens, 1968)).

These early researches were the base for the 1970s developments when the business journals started to publish articles on management decision systems, strategic planning systems and decision support systems. The term decision support system was firstly used by Gorry and Morton (Gorry & Scott-Morton, 1971) and in the following years the concept of DSS became an

area of research of its own. Little (Little, 1970) identified four criteria for designing models and systems to support management decision-making which included: robustness, ease of control, simplicity, and completeness of relevant detail. Forty years later these four criteria remain relevant when evaluating modern DSS. Later, Little (Little, 1975) expanded the frontiers of computer-supported modelling with the Brandaid system designed to support product, promotion, pricing and advertising decisions, and planted the seed for the development of the financial and marketing modelling language known as EXPRESS<sup>10</sup>.

The idea that man-computer interaction could enhance both the quality and efficiency of human problem solving, guided researchers in the following decades as the research in the DSS area gained intensity. In the middle and late 1980s DSS evolved to specialised systems focusing on specific needs:

- **Executive Information Systems (EIS):** devoted to facilitate and support the information and decision-making needs of senior executives;
- **Group Decision Support Systems (GDSS):** based in a collaboration technology designed to support meetings and group work; and
- **Organizational Decision Support Systems (ODSS):** evolved from the single user perspective into a set of users that interact with the same set of tools and make various, but still interrelated and autonomous, decisions.

When entering the 1990s, DSS was all about data warehousing and On-Line Analytical Processing (OLAP). According to Pendse (Pendse, 1997), OLAP had its origins in the APL<sup>11</sup> programming language and in systems like EXPRESS and System W<sup>12</sup>. By using a multidimensional data model, OLAP allowed complex analytical and ad-hoc queries with a rapid execution time, making it much more efficient to effectively use a DSS.

As the turn of the millennium approached, new Web-based analytical applications were introduced. Application Service Providers (ASPs) began hosting the application software and technical infrastructure for decision support capabilities. Additionally sophisticated "enterprise knowledge portals" were introduced by vendors that combined information portals, knowledge management, business intelligence, and communications-driven DSS in an integrated Web environment (Bhargava & Power, 2001).

Nowadays DSS research and development continues exploiting new technology developments benefiting from progress in very large data bases, artificial intelligence, human-computer interaction, simulation and optimization, software engineering, telecommunications, and from

---

<sup>10</sup> Currently EXPRESS is the standard modeling language for product data, and is part n°11 (ISO 10303-11) of the ISO Standard for the Product Data Representation and Exchange - STEP. Additional information can be found at <http://ng.tc184-sc4.org/>.

<sup>11</sup> APL (A Programming Language) is an interactive array-oriented language and integrated development environment based on a mathematical notation developed by Kenneth E. Iverson.

<sup>12</sup> System W, a multidimensional DSS developed by Comshare Inc., won the "Software Product of the Year" award in 1992 from Japan's Software Information Center Foundation.

more basic research on behavioural topics like organizational decision making, planning, behavioural decision theory and organizational behaviour.

### 3.2.3. Intelligent decision support systems

One of the main research lines on Decision Support System has evolved towards including intelligent abilities in those systems. These systems are based on artificial intelligence or intelligent agent technologies and are commonly called **Intelligent Decision Support Systems** (IDSS) (Holsapple & Whinston, 2000). Their main objective is to realize decision making functions by gathering and analyzing evidence, identifying and diagnosing problems, proposing possible courses of action and evaluating the proposed actions representing some human brain competences.

The approach used to structure the system learning ability can be used to classify them in two major groups<sup>13</sup>:

- **Based on behaviourism – “learning to”**: defends that learning engages the formation of associations between specific actions and specific events (*stimuli*) in the environment. These *stimuli* may either precede or follow the action (antecedents vs. consequences). Radical approaches (operant conditioning/behaviour modification/behaviour analysis) avoid any intervening variables and focuses on descriptions of relationships between behaviour and environment (“functional analysis”).
- **Based on cognition – “learning that”**: states that learning takes place in the mind, not in behaviour. It involves the formation of mental representations of the elements of a task and the discovery of how these elements are related. Behaviour is used to make inferences about mental states but is not of interest in itself (“methodological behaviourism”).

This classification is unanimous in terms of the basic concept that aggregates this kind of systems which is their ability to learn, represented by the introduction of the notion of intelligence<sup>14</sup> in their designation.

Reaching a decision is, in many aspects, similar to solving a problem (Chen Z. , 2000). Indeed, in both cases reasoning has shown to be a critical aspect since it forms the basis for evaluation and judging the received information. Additionally, perception and cognition have also been recognized as important aspects for effective decision making. Consequently, intelligent decision support systems incorporate these elements along with knowledge based systems that act as decision advisors. Thus, human reasoning and learning mechanisms have been the inspiration

---

<sup>13</sup> Based on the concepts followed by Douglas J. Navarick in his work on “Behavioral vs. Cognitive Views of Learning”: <http://psych.fullerton.edu/navarick/behavcog.ppt>.

<sup>14</sup> Herminio Duarte-Ramos proposed the use of the neologism *intellegence* instead of *intelligence* to distinguish between human and machine capacities. In his view these systems, independently from their complexity, do not present emotions nor possess any creativity, behaving in pre-programmed form. This way they realize *elective* actions and are not ready to approach general problems.

source for the development of IDSS, highly supported by the achievements made in the field of artificial intelligence<sup>15</sup>. Accordingly to this approach IDSS comprehends the following components:

- **Knowledge base:** it is responsible for storing all the knowledge available, which is provided by experts of the area of interest. To achieve its main objective the knowledge collected should be arranged in a specific format so that relations between problems and solutions could be derived (Benbya, 2008);
- **Acquisition mechanisms:** provides mechanisms for knowledge assimilation and management functions for the frame knowledge base (e.g. aggregate new knowledge, rearrange when knowledge is deleted, etc.), maintaining consistency and non-redundancy (Ishizuka & Matsuda, 1990);
- **Reasoning mechanisms:** this is where the “intelligence” of the system lays since it is responsible for manipulating the existing knowledge and create new one. Several reasoning methods can be used to implement these mechanisms, from the more formal ones based on mathematical expressions, to the more intuitive ones based on induction algorithms (Lucas, 2010).

The combination of existing knowledge, on a particular application domain, with an appropriate reasoning mechanism is the key for the system success, which is measured in terms of correctness of the proposed decisions. If well implemented, high levels of accuracy and consistency can be achieved, especially when dealing with deterministic problems. If this is not the case, performance can be compromised when a measure of uncertainty is not considered.

In the early days of knowledge based systems (KBS), rule-based systems (RBS)<sup>16</sup>, were popular. Early reasoning with uncertainty was, therefore, studied in the context of such rule-based systems with rules of the form  $A \rightarrow B_x$ , where  $x$  represents measure of uncertainty. The meaning of this rule is that when  $A$  is absolutely true then  $B$  is also true, but with certainty  $x$ . Note that this rule structure does not allow learning from past experiences, unless new rules are added to the system or existing rules are adapted to new knowledge. In both cases this is something that most systems do not perform by themselves and normally requires human intervention.

Rule-based systems emerged in the 1970 decade and they maintained their popularity until the mid 1990 (Osborne, 2009). By that time researchers started working on possible solutions for solving the lack of learning capacity of rule-based systems. Two solutions which demonstrated to be quite successful were Case-Based Reasoning (CBR) and Bayesian Networks (BN).

---

<sup>15</sup> A lot of discussion exists around the use of the term Intelligence regarding these systems, from John McCarthy' Artificial Intelligence (1954) to Zadeh's Soft Computing (1994). For the sake of simplicity, and since the author does not aim addressing the diversity of terminology in the area, the term Artificial Intelligence is used along the text.

<sup>16</sup> Rule based systems are systems that use rules to make deductions or choices. They can be looked at as systems that represent knowledge as rules of the form  $A \rightarrow B$  (Lucas 2010).

### 3.2.3.1. Case-based reasoning

The foundations for CBR systems were established based on the work of Roger Schank ( (Schank, 1982), (Paine, 1996)). Although having emerged in the laboratories of artificial intelligence (where research related with cognitive processes was developed) it has rapidly evolved for industrial and business applications. During the 1990s the interest on CBR promoted its fast grow with several systems being successfully developed (e.g.: SMART and CLAVIER were developed to support Compaq's customer service and composite part fabrication respectively).

Additionally combinations of RBS with CBR also became very popular during that period. The idea was to incorporate in rule-base systems the learning and adaptation capabilities demonstrated by CBR. Examples of this interest can be found in the work developed by Golding (Golding & Rosenbloom, 1991) and Babka (Babka & Whar, 1997).

The combination of CBR with other techniques still remains an active research field especially in the areas of diagnostics and decision ( (Berenji, Wang, Saxena, & A., 2005), (Bouchon-Meunier, 2009), (Campos, 2010), (Chang, Wang, Liu, & Qi., 2006), (Mi, Qian, Liu, & Chang, 2008), (Scully, 2006), (Thibault, Siadat, & Martin, 2006)).

Table 3.2. Major components of a case

Major Components	Description
Problem description	Goals to be achieved
	Constrains on the goals
	Features of the problem situation and relationship between parts
Solution	Solutions
	Reasoning steps
	Justifications for decisions
Outcome	The outcome itself
	Explanation of expected violation and/or failure
	Repair strategy
	Pointer to next attempt at solution

CBR systems are based on a starting set of cases that are structured in an appropriate format in order to constitute training examples. It is a problem solving approach that works by identifying commonalities between a retrieved case and the target problem. According to Kolodner (Kolodner, 1993) a case is "a contextualized piece of knowledge representing an experience that teaches a lesson fundamental to archiving the goal of the reasoned". Its main components are the ones presented in Table 3.2.



This approach proposes that a new problem is solved by recognizing its similarities to other specific known problems (i.e. problems that occurred in the past and from which it was acquired valid experience), and adapts the solution used to solve the past problems to the new one. For this reason, case-based methods are strongly influenced by cognitive science since they mime the human behaviour when trying to solve a problem or make a decision.

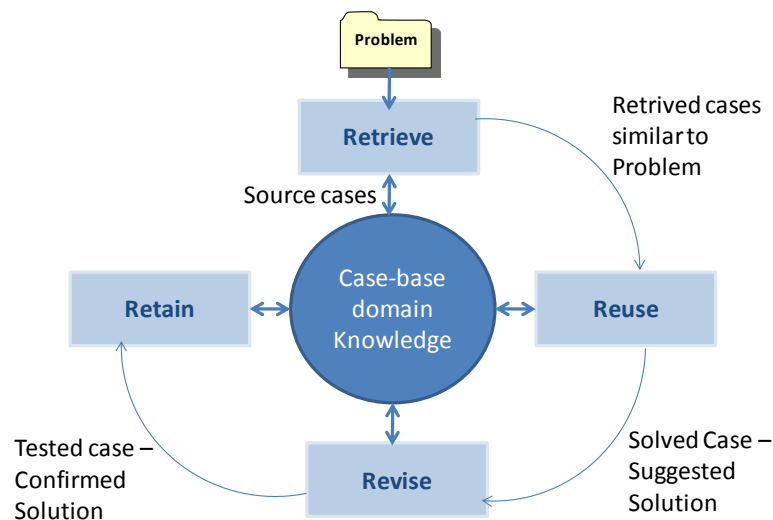


Figure 3.6. The Case-based Reasoning cycle

Reasoning is defined by the process of drawing inferences or conclusions, thus moving from what is known (fact) to what is unknown (inference). As illustrated in Figure 3.6 four phases comprise this process, namely:

- **Retrieving:** the output of retrieving phase is a set of the most similar cases;
- **Reusing:** the knowledge about the retrieved case(s) is reused to solve the actual problem;
- **Revising:** the solution adopted in the retrieved case(s) is then revised and adapted to the new problem solution, generating the suggested solution;
- **Retaining:** this solution should then be confirmed and the knowledge/experience gained with this process is retained to be used in the future, in the form of a case.

There are some characteristics that indicate that a specific domain is suitable for the use of CBR. Among these characteristics is the existence of records of previously solved problems. Additionally the idea that historical cases are an asset that must be preserved is also indicative of the possibility of using CBR. Nonetheless, when building a CBR system there is a set of requirements that must be fulfilled so that the system can work properly and provide *de facto* support in the domain of interest:

1. The first step is the definition of what is a *case* and how its representation will be made in order to capture its true meaning. A case must always have a description and a solution and, in complex domains, the description may have to be divided into several

characteristics so that a full explanation of the case can be achieved. Thus the case representation is strongly connected to the domain of application. It is not possible to define a set of parameters to be included in its representation that fit all of them, and for this reason, a careful identification of the representative parameters should be done by experts on each application domain;

2. Afterwards, cases should be indexed so that their retrieval becomes quicker and easier. An index is a computational data structure that can be stored in memory and searched quickly (Watson, 1995);
3. Then cases need to be retrieved through a retrieval algorithm that finds the most similar cases to the current problem. Case retrieval requires a combination of search and matching, and the major CBR applications use, in general, two retrieval techniques: nearest neighbour retrieval algorithm and inductive retrieval algorithm. The first computes the similarity between stored cases and new input case, based on weight features; while inductive retrieval algorithm is a technique that determines which features do the best job in discriminating cases, and generates a decision tree type structure to organize the cases in memory. Both techniques are widely applied in CBR applications and tools and both present strengths and weaknesses. Thus, the choice between nearest-neighbour retrieval and inductive retrieval requires experience and experimentation. Usually, it is a good choice using nearest-neighbour retrieval without any pre-indexing (Watson, 1997) but, if retrieval time becomes an important issue, inductive retrieval is preferable.
4. At the final stage, the adaptation of a solution that worked in the past to the new problem should be applied and the new case is added to the case-base.

#### **3.2.3.2. Bayesian networks**

By the middle of the 1980s Bayesian Networks made their appearance as a new option to rule-based system. The term Bayesian refers to Thomas Bayes<sup>17</sup>, who proved a special case of what is now called Bayes' theorem (Price, 1764). The theorem defends that the probability of an event A given an event B depends not only on the relationship between events A and B but also on the marginal probability of occurrence of each event.

In Bayesian statistics the probability of an event  $x$  is a person's *degree of belief* in that event, thus is a property of the person who assigns the probability (in opposition to classical probability which is a property of the physical world). The process of measuring a degree of belief is commonly referred to as a *probability assessment*. One problem with the probability assessment is that of precision. Nonetheless, in most cases, probabilities are used to make decisions, which are not sensitive to small variations in probabilities.

---

<sup>17</sup> Thomas Bayes (1702-1761) was an English mathematician and Presbyterian minister, known for having formulated a specific case of the theorem that bears his name. The theorem was only published in 1764.

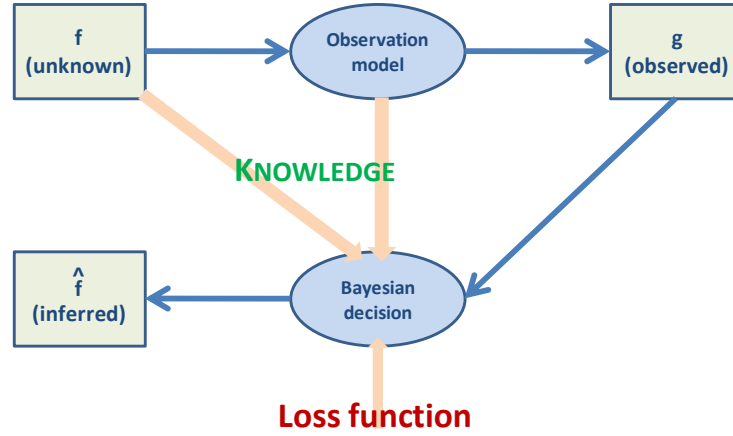


Figure 3.7. Bayesian approach<sup>18</sup>

Figure 3.7 expresses the basics of the Bayesian approach. The decision rules in Bayesian theory are derived accordingly with the applications of the following fundamental principles:

- Knowledge is expressed by means of probability functions;
- Inferences are conditioned to observations resultant from the application of the conditionality principle stating that “any inference must be based on observed data” (Birnbaum, 1962);
- The information contained in the observations can only be carried by the likelihood function as stated in the likelihood principle (Barnard, Jenkins, & Winsten, 1962);
- The loss function represents the loss associated with a bad estimation of specific variable measured in terms of the degree of imprecision.

Thus, knowledge about  $f$  once  $g$  is observed, is expressed by the posterior probability function,  $p(f|g)$ , which is also called Bayes law:

$$p(f|g) = \frac{p(g|f) \times p(f)}{p(g)} \quad (3.7)$$

The expected value,  $E$ , for the loss function,  $L(f, \hat{f})$ , is also defined considering the observations of  $g$ , which represents the so-called “posterior expected loss”:

$$E[L(f, \hat{f})|g] = \int L(f, \hat{f}) p(f|g) df \equiv \rho(p(f), \hat{f}|g) \quad (3.8)$$

The optimal decision rule is the one that corresponds to a minimization of  $\rho(p(f), \hat{f}|g)$ :

$$\widehat{f_{Bayes}} = \arg \min_{\hat{f}} \rho(p(f), \hat{f}|g) \quad (3.9)$$

<sup>18</sup> From Figueiredo (Figueiredo 1998)

Thus, it is possible to draw a more detailed diagram of the Bayesian approach, considering these specific aspects (see Figure 3.8).

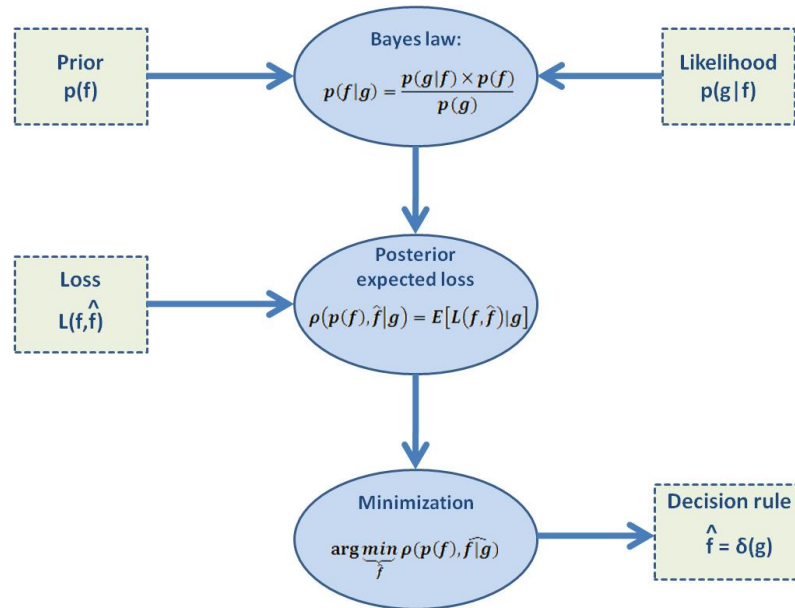


Figure 3.8. Deriving decisions in Bayesian approach<sup>19</sup>

Naturally these concepts were used to build the basis for the development of Bayesian Networks. The term was firstly introduced by Pearl (Pearl, 1985) and it rapidly attracted the interest of the research community due its capacity of handling incomplete data sets and allowing learning about causal relationships. BNs are very used in problems involving classification or regression, and even in cases where input variables are strongly anti-correlated and one or more inputs are not observable, it is possible to encode such dependencies. Additionally they are also useful when trying to gain understanding about a problem domain, and the knowledge of causal relationships allows making predictions in the presence of interventions. The research field remains very active in a variety of areas, e.g. pattern classification (Kim, Ha, & Lee, 2009), signal processing (Wang, Kuruoğlu, Yang, Xu, & Huang, 2010) and social systems (Sales, Schwaab, & Nassar, 2010).

In conjunction with Bayesian statistical techniques, BNs facilitate the combination of domain knowledge and data. Additionally they have a causal semantics that makes the encoding of causal prior knowledge particularly straightforward by using probabilities. Thus, prior knowledge and data can be combined with well studied techniques from Bayesian statistics.

Pragmatically a BN is a graphical model for probabilistic relationships among a set of variables (see Figure 3.9) enabling the representation of uncertain expert knowledge (Heckerman, Geiger, & Chickering, 1995). In a BN the entities of interest (e.g., decision criteria and sub-criteria, factors that influence them, etc.) are treated as random variables and represented as nodes in the network, connected by directed arcs indicating probabilistic dependencies between them. The

<sup>19</sup> From Figueiredo (Figueiredo 1998)

network structure, together with conditional probability tables associated with each node, provides a compact representation of the joint probability distribution of all variables (Wattayay & Peng, 2004).

Generally speaking the structure of a BN is a directed acyclic graph in whose nodes correspond to random variables of interest and the directed arcs represent direct causal or influential relation between nodes. The uncertainty of the interdependence of variables is represented locally by the conditional probability table  $p(x_i|\pi_i)$  associated with each node  $x_i$ , where  $\pi_i$  is the parent set of  $x_i$ .

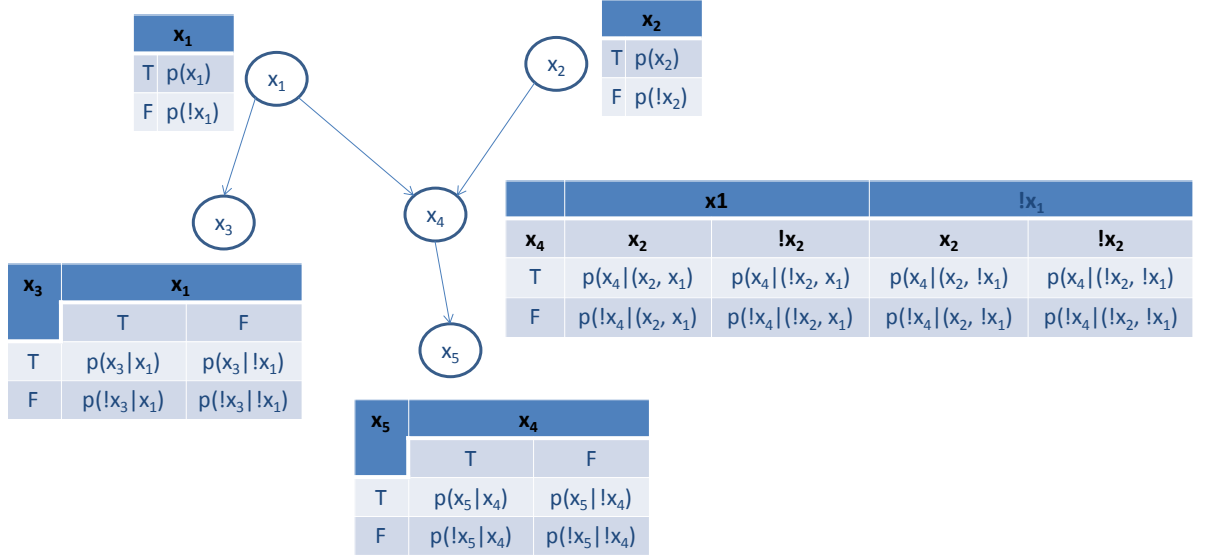


Figure 3.9. Example of a Bayesian network

An independence assumption is made in BN stating that  $x_i$ , given its parents  $\pi_i$ , is independent of any other variables except its descendants. The graphical structure of BN allows an unambiguous representation of interdependency between variables. Therefore, together with the independence assumption, this leads to one of the most important features of BN, i.e., the joint probability distribution of  $X = (x_1, \dots, x_n)$  can be factored out as a product of the conditional distributions in the network:

$$p(X = x) = \prod_{i=1}^n p(x_i, \pi_i) \quad (3.10)$$

With the joint probability distribution, Bayesian networks can support in theory any probabilistic inference in the joint space. Moreover, probabilistic inference algorithms have been developed by exploring the interdependency captured by the network structure. Most important among them are algorithms for computing posterior probabilities,  $p(x_i|e)$ , where  $e$  denotes *evidence*, i.e. the observed values for some variables. By this, it becomes straightforward to determine the effect of observations  $e$  (facts known with certainty) on the uncertainty of any set of random variables  $x_i$  by computing the conditional probability distribution  $p(x_i|e)$  for  $i = 1 \dots n$ .

Examples of these algorithms can be found in “belief propagation” (Pearl, 1988), the “Junction tree” (Madson & Jensen, 1998), and more recently in algorithms that include various statistical sampling techniques, such as Markov Chain Monte Carlo sampling (Goel & Salganik, 2009).

### **3.2.4. Decision support systems for life cycle management**

Life Cycle Management is a business management approach that can be used by all types of business and organizations in order to improve their sustainability performance. The approach can be used equally by both large and small firms, and its purpose is to ensure more sustainable value chain management. LCM is used to target, organize, analyze and manage production-related information and activities (Remmen, Jensen, & Frydendal, 2007) towards continuous improvement along the production life cycle.

By learning how to more effectively manage this cycle, a company or an organisation can reveal a wealth of business, environmental and social value and, this way, make the choice to engage in more sustainable activities and production patterns. From this point of view, LCM is a framework for business planning and management that helps business managers to:

- Analyse and understand the life cycle stages of the business, product or service;
- Identify the potential economic, social, or environmental risks and opportunities at each stage; and
- Establish proactive systems to pursue the opportunities and manage or minimise the risks.

Thus, LCM is directly engaged with decision making, since the aim is to make more informed business decisions, acknowledging the risks and defining strategies to cope with them.

Normally, and even without a specific LCM strategy implemented, chances are that life cycle considerations are already influencing the decisions made in the daily operation of an industrial plant. The goal of using a structured LCM approach consists in helping making these decisions in a more deliberate and systematic way, contributing for a more sustainable production and consumption, and clearly defining and measuring the business value gained by doing so (UNEP/SETAC, Life Cycle Initiative, 2009). Note that the measuring of the success of this business approach goes beyond short-term success and aims at long-term value creation.

Life cycle models are not just a phenomenon of the life sciences. Industrial plants experience a similar cycle of life. Just as a person is born, grows, matures, and eventually experiences decline and ultimately death, so do industrial plants. The stages are the same for all plants but the length of each stage may vary from plant to plant and, even within the same plant, different parts may be at different life cycle stages. The strategies to cope with these differences must be integrated and dependents on the stage of the life cycle.

As in business, life cycle management of industrial plants strongly relies in taking decisions on a daily basis. From the early phases of plant installation until its decommissioning the implementation of decision support mechanisms are thought of contributing for the plant success.

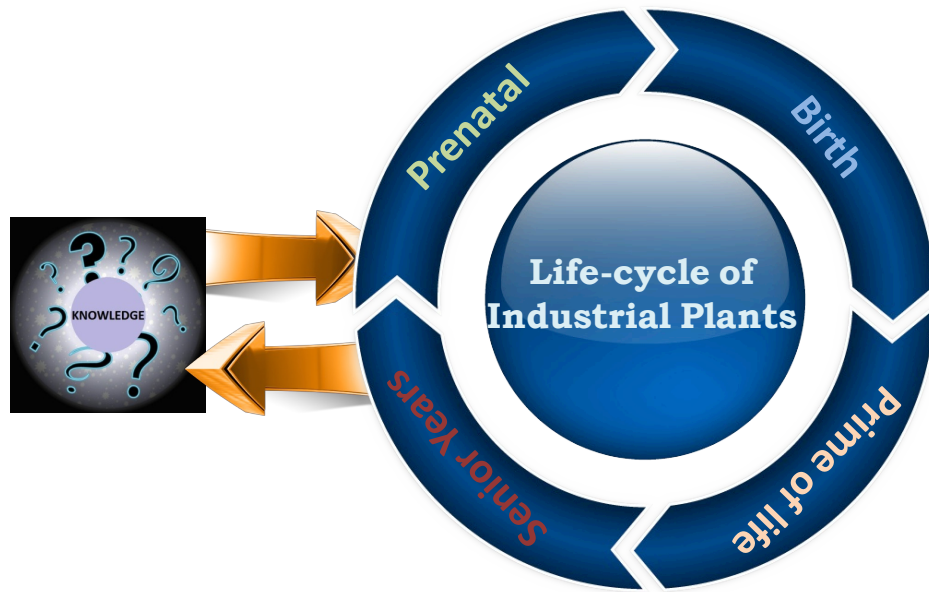


Figure 3.10. Life cycle of industrial plants

Figure 3.10 illustrates the life cycle of industrial plants, divided into four phases:

- **Prenatal:** this is the planning phase, in which the selection of the most suitable technological solutions is made. Additionally, at this phase the installation of appropriate equipment (e.g. sensors) to support plant status monitoring should be evaluated. If available, knowledge coming from already installed plants can be used to solve common problems;
- **Birth:** at this stage the plant is effectively installed following the recommendations and plans developed at *Prenatal* phase. The collection of problems occurred at *Birth* phase can contribute for smoothing future installations;
- **Prime of life:** this is the phase that concentrates the wider time horizon in terms of life cycle. The two previous phases should be planned and executed in order to maximise, and optimise, this one. Additionally the use of decision support strategies to support daily operation can also contribute to fully exploit the plant capabilities promoting the extension of this phase;
- **Senior years:** at this stage it is possible that the plant presents some ageing problems that cannot be solved with regular maintenance. The knowledge collected along the other phases can contribute for an improved management of this phase. Additionally this knowledge can also contribute for the identification of critical parts which replacement and/or improvement may have a positive impact in this phase.

The installation of appropriate equipment is crucial for an adequate development of a decision support strategy. In fact, an industrial plant can produce, through the available instrumentation, a

huge amount of information from where decisions can be supported if well interpreted (Marques & Neves-Silva, 2009). Thus, the first step is to guarantee the access to that information.

Subsequent success of the decision strategy, and consequently of the life cycle management approach, consists in the correct identification of the impacts that specific decisions may have in the long term behaviour of the plant. This idea relies on the fact that every action made during production (and even before) has an impact. Often these impacts are not obvious or immediate, in fact many of them are hidden or indirect, and they only appear when one takes a more holistic view. Thus the central problem is on how to combine the existing information so that the impact of a specific decision can be foreseen.

This task may be really hard to execute in industrial environments since many times information is not structured. Instead, it is scattered along the plant, making extremely difficult to correlate and reach a consistent conclusion, i.e. a decision. Thus the problem is to identify which information should be collected and how it should be processed in order to be effectively useful.

Every industrial plant considered must perform this information identification and collection, since each plant represents a unique problem and the solution requires the participation of experts with great knowledge on the plant specificities. Nonetheless, industrial plants producing the same (or similar) products, using the same (or similar) technology and the same (or similar) processes can re-use the knowledge acquired with the experience of others.

After this phase the knowledge collected can be used to derive life cycle parameters (LCP), which provide insight about the plant status. Additional details on LCP calculation can be found in Annex B. From a general point of view, these parameters can be grouped into the following categories:

- **Life cycle costs (LCC):** The term LCC is the total cost of a system during its life cycle from concept to scrap. These include:
  - Acquisition cost ( $A_c$ ): the initial cost which could be easily calculated during the conception phase;
  - Operation ( $O_c$ ): the total cost which is consumed for operating the system during operational phase;
  - Maintenance ( $M_c$ ): the cost for maintaining the system during operational phase. This is the total maintenance cost carried out only for the maintenance activities;
  - Decommissioning costs ( $D_c$ ): the cost for decommissioning the system during disposal phase.

The total LCC are given by:

$$LCC = A_c + O_c + M_c + D_c \quad (3.11)$$

- **Reliability:** The probability that a system will perform its intended function for a specified interval under stated condition. It is extrapolated from Average Failure Rate (AFR) and Mean Time Between Failure (MTBF);



- **Availability:** The ability of an item to be in a state to perform a required function under given conditions, at a given instant of time, or over a given time interval, assuming that the required external resources are provided. The concept relies on Inherent Availability (IA), which reflects the percentage of time a product would be available if no delays due to maintenance, supply, etc., were encountered, and Operational Availability (OA) which includes the effects of maintenance delays and other non-design factors;
- **Maintainability:** The relative ease and economy of time and resources with which an item can be retained in, or restored to, a specified condition when maintenance is performed by personnel having specified skill levels, and using prescribed procedures and resources, at each level of maintenance and repair. In this context, maintainability is a function of design;
- **Safety** The freedom from those conditions that can cause death, injury, occupational illness, damage to or loss of equipment or property, or damage to the environment
- **Other parameters:** (extrapolated from the ones mentioned above)
  - **Overall Equipment Effectiveness (OEE) and Net Equipment Effectiveness (NEE):** measures the combination of three elements for the physical asset; equipment asset availability, performance and quality output;
  - **Overall Craft Effectiveness (OCE):** focuses upon craft labour productivity and measuring/improving the value added contribution that people assets make. Three elements must be considered: effectiveness, efficiency and quality;
  - **Failure rate (FR):** Number of failure per unit of gross operating period in terms of time, events, cycles or number of parts;
  - **Spontaneity intensity (SI):** the ratio between total unplanned maintenance time to the total maintenance time;
  - **Breakdown intensity (BI):** the ratio between the number of breakdown occurred for a period of total production time;
- **Physical variables:** which can be measured directly, using sensors (e.g. tool wear).

The continuous analysis of these parameters is the basis for the development of an effective decision support strategy aiming to contribute for the life cycle optimisation of the industrial plant.

Normally decisions made over one of these parameters will eventually affect the behaviour of the others, since all of them are tightly connected. This notion of impact is important especially from the long term perspective of the plant in which the goal is determined by the maintenance of a specified performance level through finding equilibrium between the defined parameters.

### 3.3. Conclusions

This chapter aimed at clarifying the connection that exists between risk and decision. Since all decisions involving uncertainty present a degree of risk, it is important to be aware of it when selecting the appropriate course of action, which is, in principle, the one that minimises risk. In some cases, the characteristics of the outcomes may influence judgement. This is particularly true

for gambling where the prizes can sometimes be attractive enough to convince people playing an almost impossible game.

On the other hand, risk is a probabilistic game and the more information is available the more certainty is reached. Therefore, it is important to understand which variables influence risk in order to understand how their behaviour will affect the impact of the selected course of action. Using risk analysis strategies based on probabilities is then seen as a suitable way of dealing with the uncertainty problem, especially due to their capacity of providing insight about the impacts on the specific outcomes.

Nonetheless, in some cases, establishing relations between variables might not be so simple and, it is possible that, decision makers get lost over the huge amount of information being generated by some applications. Thus, to help solving this problem, selecting and organizing information are key aspects, for decision support systems. Their use helps structuring the existing connections between variables and, in some cases, may contribute for the establishment of relations between events that could seem to be uncorrelated.

Moreover, the addition of intelligence to decision support systems enables them to learn from past experience and adapt their answer to the new cases based on that learning process. This is actually what people do in their normal life, and the concept was adapted to the world of artificial intelligence. As in many other fields, the notion of intelligence is particularly important in industrial environments where most problems are recurrent and the way they are solved can make the difference in terms of plant performance which is the main goal for applying intelligent decision support systems to industrial plants.

When speaking of performance the impact will normally be reflected in the production costs, especially from a long term perspective. Thus, knowing exactly the parameters that should be considered to minimise production costs (without compromising performance goals) provides insight about the impact they have in the life cycle of the plant.

To conclude, the analysis of the plant is essential for the success of any strategy envisaging life cycle optimisation. For this reason, the knowledge acquired both by previous experiences and through plant experts, is vital and should not be overlooked.

## *4. Proposed Decision Support Methodology based on Risk Analysis*

---

Decision support in industrial environments can be performed using diverse information coming from several parts of the plant. Based on the concepts introduced in the previous chapters the approach presented in this thesis is focused on supporting decisions that must be made on a short term basis by providing information on what might be the impact of that decision in the long term perspective.

Industrial environments are, due to their complex nature, difficult to represent in a straightforward mathematical model or expression. This characteristic compromises the possibility of having them fully described by means of a single function. In fact, if one can say that their overall constitution can be compared with a regular system, i.e. with inputs and outputs, the nature of their inner components make it almost impossible to define a single relation between those inputs and outputs. Thus, if one wants to deal with the information coming from an industrial plant and, in return, after some appropriate treatment, feedback the plant with new information then some level of structure to relate the tasks performed in the plant with its outcome must be found.

Due to known impact of maintenance activities in the plant performance, the maintenance area was selected as the focus of this dissertation. In fact maintenance is a daily operation of industrial plants and, if correctly performed, may contribute positively for their life-expectancy. Thus, minimization of the impact of the maintenance activities in the plant operation, together with the associated costs, is the core of the methodology proposed in this section. As already mentioned, maintenance costs contribute with a substantial slice for the calculation of the life cycle cost of an industrial plant. Additionally the use of different maintenance strategies also has considerable impact in the entire life cycle of the plant.

If it is true that generally companies want to reduce their direct maintenance costs by reducing maintenance activities, it is also true that they do not want to reduce the life expectancy of their machines due to inappropriate maintenance strategies. Thus, to find the balance between these two aspects is the key issue of the work here proposed.

The methodology proposed is then based on the assumption that the collection of information about similar maintenance processes in similar industrial plants can be used for their optimisation, and by this contribute for improving its life cycle management.

Minimize maintenance costs and find the best maintenance strategy constitutes a two step decision problem that derives from the uncertainty in several parts of the system. In fact it is the existence of uncertainty that leads to the need for a decision. This uncertainty is due to several reasons, namely the impossibility of having two machines exactly equal, being operated in the same way, in the same conditions, etc., and the effect these aspects might have to achieve the most appropriate result are not quantifiable. In the next sections the detailed methodology will be explained.

#### 4.1. Problem framework

The work here presented aims at developing a system that uses data coming from different sources along the industrial plant to support decisions for life cycle management.

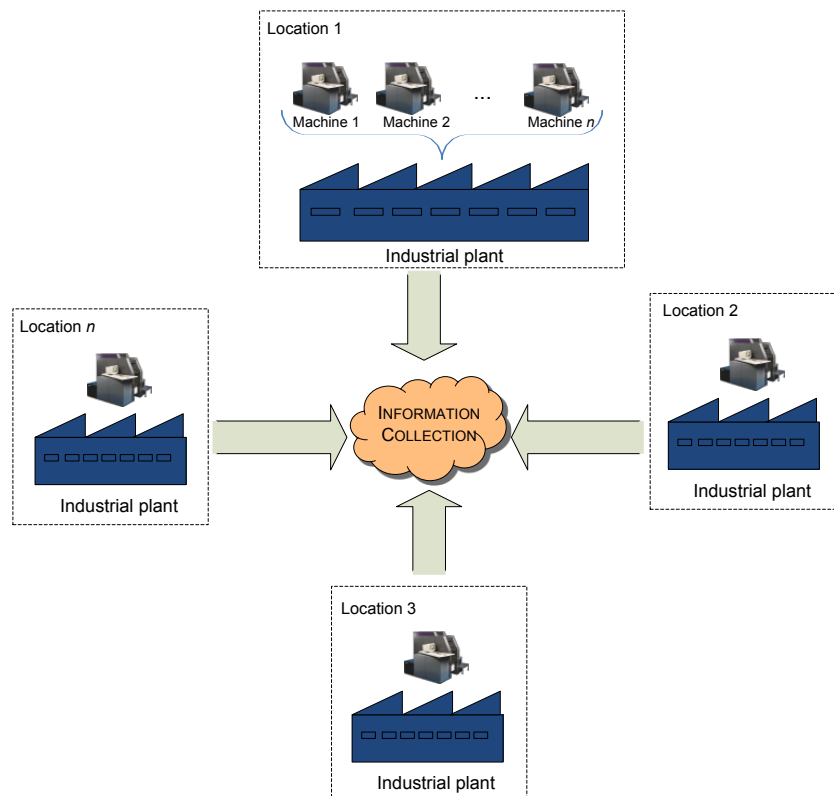


Figure 4.1. Focusing the problem

When speaking about industrial plants there is the need to clearly identify which kind of plants are being considered and how they are characterized. From a general point of view the focus of the proposed methodology is on industrial plants constituted by similar machines to realize the same function. Those machines may, or may not, be located in the same physical location. For example a company may have several plants equipped with similar machines but geographically distributed. What is essential is that all of them contribute with valuable information (see Figure 4.1). The value of the information is derived from the level of similarity found.

The information collected is used to establish cause-effect relations in order to assess the impact that a specific action might have in the behaviour of a plant. This assessment is made based on the historic background of the machines and the inputs from experts on machine behaviour. Thus, it is vital to settle an appropriate knowledge model to be used to interpret that information, and as a major assumption of this methodology, ***information on the behaviour of similar machines must be accessible in order to build the knowledge model.***

This knowledge model is developed taking into consideration the difficulty in finding one single expression able to characterize the plant environment. To solve this issue this thesis proposes a **case-based structure** capable of organising the available information in cases that can be used for further comparison. The case-base to be built must comprise information on several aspects about plant behaviour, namely:

- Important variables on plant status;
- Common problems that are related with the identified variables;
- Common causes for those problems;
- Maintenance actions to treat the problems.

The problem framework focuses exclusively on regular maintenance problems, thus it excludes situations where safety of operators and/or plant may be at stake. Having this in mind it is considered that all the problems occurring in the plant are able of being solved since and this solution is only a matter of cost.

## ***4.2. Concept for risk-based decision support system***

Based on the above assumptions the objective is to develop a system that supports the life cycle management of industrial plants, by helping the user when some decision must be made, based on the analysis of the received information.

To assess the overall status of the industrial plant a set of state variables, that describe its condition, are used. These variables characterize the status of the observable part of the system. The evolution of the state variables registered on the scope of abnormal situations is the base for establishing correlations that are used for building the decision model.

Also, these state variables are vital to make an analysis of the plant status and, by monitoring them to check threshold's violations, the system is capable of suggesting a strategy to deal with

the situation. Such suggestions, if taken into consideration, are likely to have a positive impact in the life cycle management of the plant.

The system's main function is to provide recommendations to the human user of the system regarding which actions to choose for achieving a specific goal. These recommendations result from the understanding of the plant operation through modelling cause-effect relations (cause events/actions to consequence effects) achieved by an effective state variable monitoring. The relations are used for the establishment of cases that describe the effects of a certain cause. The collection of these cases is then used to select the best option, from the set of available ones, for a specific situation.

Each time an event is fired, which in the system corresponds to a state variable that has trespassed some predefined threshold, the system performs the following set of operations:

1. Identifies the state variable associated to the fired event;
2. Searches for previous cases, that are associated to the same state variable;
3. Checks the status of those cases in order to identify if they were treated successfully and considered as solved;
4. From the ones that were considered solved collects the causes for them;
5. Checks which might be the consequences associated to the identified causes;
6. Applies a risk assessment strategy to identify which course of action may represent higher costs;
7. Selects the cause associated to the course of action with lower cost suggests a plan to eliminate that cause.

Each time this sequential process occurs, a new case is generated and, if after plan implementation, the problem was eliminated, then the case is considered as being solved and it will be used in a future situation as an additional source of information.

This ability highlights another relevant aspect of this approach which is the adapting capability of the system. In fact, each time new and valid information is added, the system will use it for future consideration. This mechanism will drive to more appropriate and refined results.

This characteristic of the system can be compared to the one found in adaptive systems where the computation of the control law<sup>20</sup> needs to be adapted, at each iteration, to the continuously changing conditions.

Adaptive control is a technique of applying some system identification method to obtain a model of the process and its environment from input-output experiments and using this model to design the controller. The parameters of the controller are adjusted during operation of the plant as the amount of data available for plant identification increases. Extensive work has been done in the

---

<sup>20</sup> According to Encyclopedia Britannica (<http://www.britannica.com/EBchecked/topic/135472/control-law>) a control law is the function of the state, which determines the control action that is to be taken at any instant over a system.

area since Kalman (Kalman, 1960) established the analytical form of the concept especially in the decades of 1970 and 1980 where the theoretical aspects of adaptive control were developed. The work of Sastry and Bodson (Sastry & Bodson, 1989) and Aström and Wittenmark (Aström & Wittenmark, 1989) provide good insight about the details of the technique. More recently the work in the area of adaptive control is still active particularly in what concerns its application to a diversity of areas ( (Neves-Silva, 1999), (Nunes, Mendonça, Lemos, & Amorim, 2007), (Tao, Chen, Tang, & Joshi, 2004)).

The main obstacle to the direct application of adaptive control to the problematic selected in this thesis is the difficulty in finding a unique mathematic relation, e.g. a transfer function, to be used as the industrial plant model. Thus, the work here developed will use the adaptation concept but, instead of using a mathematical model, a knowledge model to express plant behaviour will be used.

The aim is then to collect sufficient information that enable building a set of Cases (which are nothing more than problems which status is “solved”), reflecting the history of the plant. This information, together with the knowledge provided by the experts on the plant operation, enables representing *a priori* knowledge, expressing knowledge acquired over time with high level of certainty. Figure 4.2 shows the proposed concept for the intended risk-based decision support system.

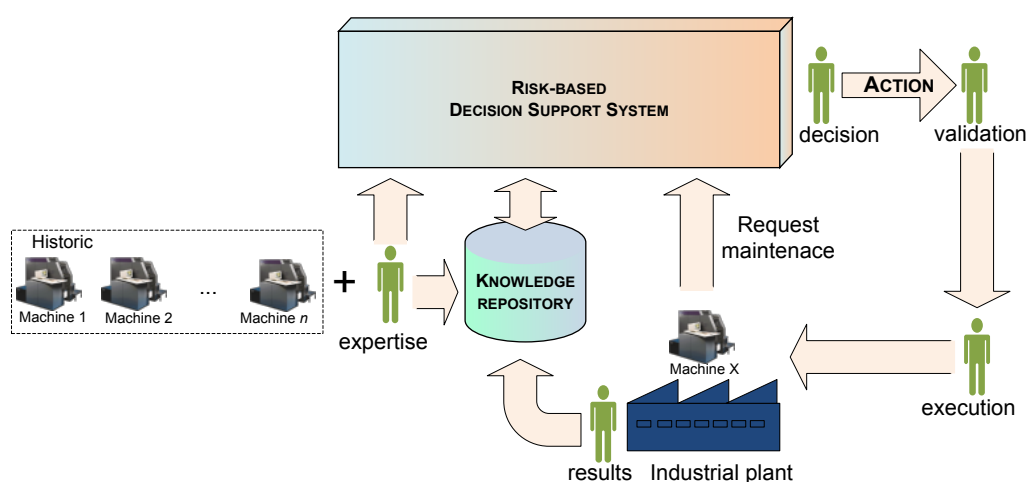


Figure 4.2. Concept for Risk-based Decision Support System

The cases are stored in the knowledge repository of the system to serve as a knowledge base for the rest of the process. Then, each time an abnormal value is detected on a state variable a request for maintenance will be generated which is analyzed and the best option to deal with problem presented.

After implementing the decision (in the form of a maintenance action), the result is measured in terms of success, i.e. solved or not the problem. Note that over the entire process it is requested the interaction with human actors, at several steps. Starting with the expert contribution for building the knowledge model and for defining the decision strategy, it is also requested checking

the option suggested by the system, validate it in the form of an action, implement that action and finally return the results to the system. In some cases it is possible that these different roles can be represented by the same actor. These human interventions guarantee the level of reliability of the developed methodology.

#### **4.2.1. Information collection**

The approach proposed is based on the collection of information and data about the industrial plant along its operational phase. The idea is to collect information coming not only from the machines but also from experts that have deep knowledge on the specific production processes.

This knowledge generally comes from previous experiences where similar plants were involved and from which experts could collect relevant knowledge. Thus, the idea of previous knowledge that may be used to deal with present and future problems is the cornerstone of this work.

First of all it is important to highlight that the richest source of information is when something goes wrong. Thus, to understand how the plant works and how maintenance should be developed (to maintain the work level) the focus should be on the problems the plant had in the past, how they were solved and which consequences aroused from them. Indeed, it is important to concentrate on problems that affect the state variables of the plant, understanding which of them should be monitored and how should be defined the normal behaviour of those variables.

The establishment of rules to settle thresholds associated to the normal values (desired for those variables) could be useful for detecting an abnormal situation. That situation would occur each time a variable violates the defined thresholds.

Additionally, it is also important to understand what happens when the situation occurs in terms of what could have caused it and what could be its consequences. Thus to achieve the overall stated objective the following aspects must be covered:

- List of state variables important to assess the overall status of the plant;
- List of rules associated to the normal behaviour of the state variables;
- List of common problems that occur each time a rule is violated;
- List of causes for those problems;
- List of actions to deal with the causes, i.e. to eliminate the problems

Once all this information on the plant is collected it is possible to build a set of *Cases* that represent what happened in past situations and which decisions were made at that time. The establishment of correlations between the situations, the actions developed to deal with them and their impact in overall life cycle of the plant will help clarifying the following aspects:

- Identification of what caused the abnormal value;
- Assess the risk of the abnormal situation;
- Assess the impact of the maintenance action in the plant;
- Assess the costs of the maintenance action.



## 4.2.2. Knowledge model

To fit the above mentioned needs this thesis proposes the development of a knowledge model to be used as structure for the knowledge repository.

At this point, the focus is on assuring the existence of the necessary entities to address the aspects of risk analysis and decision support. Thus, the simplicity of the model derives from this orientation line. Figure 4.3 presents the developed model showing the entities specified in order to implement the risk analysis and the support to decision process.

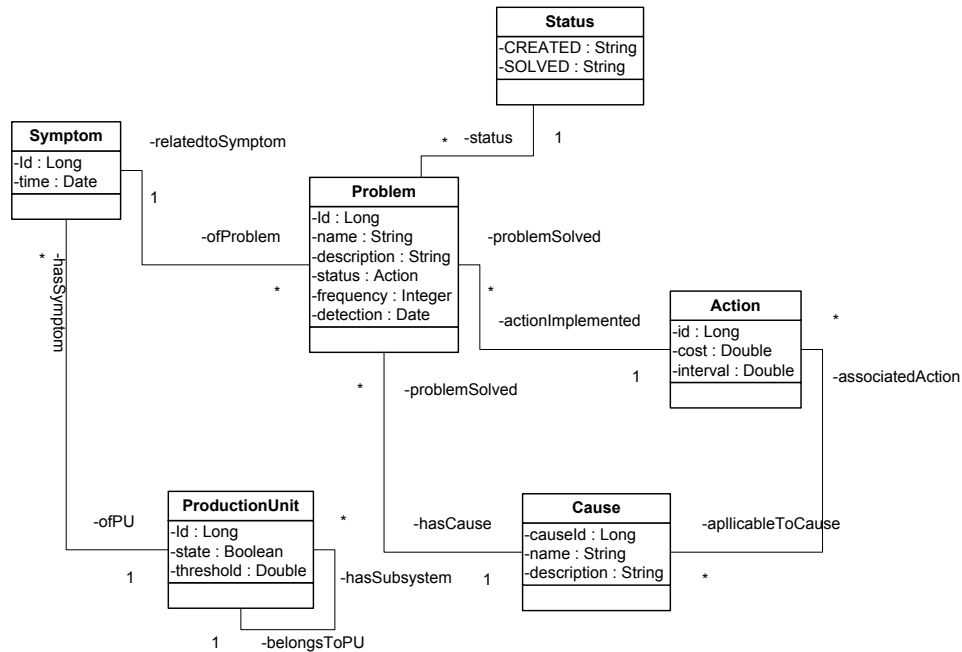


Figure 4.3. Developed model for the Knowledge Repository

The entity **Problem** is the central part of the entire system, since any deviation on the defined normal behaviour of the plant will be stored as an instance of this entity. This entity is connected to **Symptom**, representing the firing event for the risk assessment process, and to **Cause**, which represents the actual cause of the problem. The **Symptom** entity makes the connection between the problems and the operating parts of the plant, i.e. **ProductionUnit**, which is responsible by representing machines, and their subsystems, involved in production.

The idea is that each machine, and associated subsystems, has a specific cause when they breakdown. Nonetheless those causes are not immediately identified since the subsystems do not possess a unique symptom for each of them. Thus each time a **Symptom** is detected the only information is on which is the group of subsystems that may be involved in the **Problem**.

The use of the risk analysis strategy helps in the process of identifying the most probable cause taking into consideration the risks associated. This is done by collecting similar solved cases, check their causes, and compute which action should be applied in order to minimise the overall costs.

However, all these entities have several properties that are used to store the associated information. The entities referenced are the ones directly used in the context of this work.

Finally the connection between **Cause** and **Action** reflects the actions that may be developed to deal with a specific cause. Note that the strategy developed assumes that several actions may be developed to deal with one cause and the decision system will help the user in identifying the one that is most adequate for the current situation considering the risk analysis developed.

The property *frequency* in **Problem** entity represents the number of times a specific problem has occurred and normally this value is 1. Problems with higher frequency represent a *a priori* knowledge, and intend to symbolize a high level of reliability about that knowledge. When a similar problem occurs the probability of the cause to be the same defined for the problem with higher *frequency* is higher than for other different problems that only occurred once since these problems have more influence in the risk analysis strategy.

### 4.2.3. Information correlation and aggregation

To demonstrate how variables correlation is determined it is considered that state variable  $x_1$  describes the status of a specific machine. When used normally the status has a regular behaviour (see left side of Figure 4.4).

To avoid critical damages in the machine structure it is considered that the maintenance strategy defined includes the observation of a threshold. When the threshold is violated a decision point is achieved and a set different options must be considered in terms of maintenance actions (see right side of Figure 4.4):  $A = \{A_1, \dots, A_n\}$  (e.g. perform action  $A_1$ , or perform action  $A_2$ , where  $A_1$  stands for a quick action with low cost, and  $A_2$  represents a more profound solution to the problem thus having higher cost than  $A_1$ ).

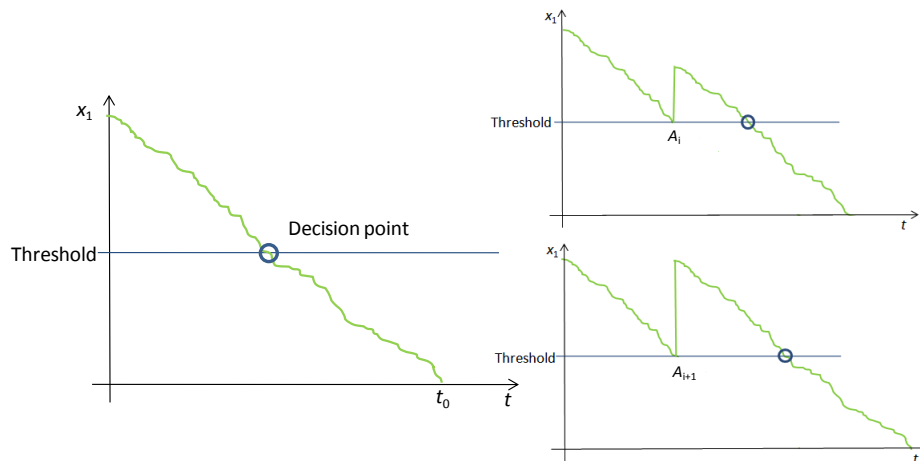


Figure 4.4. Expected behaviour of state variable  $x_1$  considering different maintenance strategies

The results can be distinguished by the type of solution applied in each case: actions which represent lower costs are normally associated to quick patches and will, eventually, lead to a new problem (i.e. violating the threshold) sooner than actions representing profound corrective

measures. From a long term perspective it is possible that actions with lower costs reveal themselves as more costly than the others. On the other hand, continuous application of profound corrective measures may also represent wasting a lot of potential. Thus, how should the decision be made?

Since the time elapsed between problems was identified as being affected by the selected action, then it is considered that there is a degree of correlation between the machine state and this time interval. By this, the effect of different manipulations of  $x_1$  (which is affected by the actions executed) is measured in terms of the impact to the machine (and consequently to the plant). This impact is the time elapsed between problems, varying accordingly with the action applied, which should be as long as possible to maintain the methodology objective (see Figure 4.5).

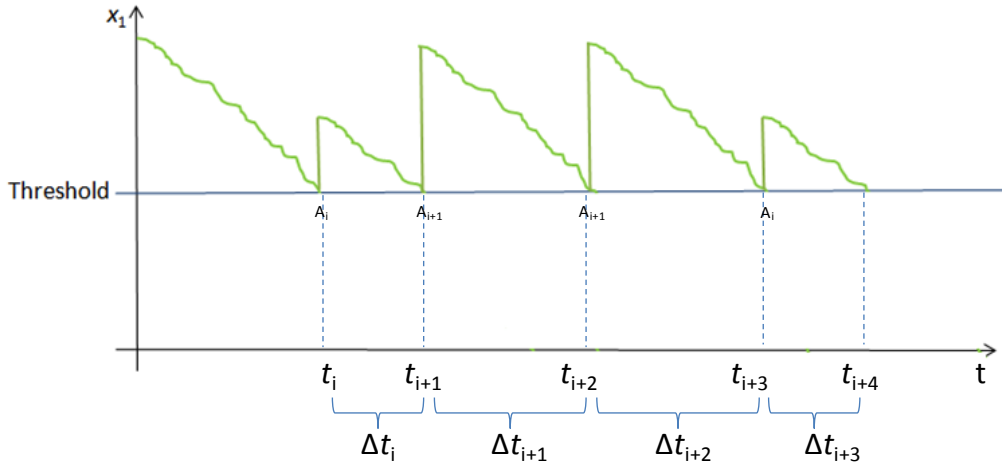


Figure 4.5. Behaviour of  $x_1$  according with the performed actions

The assumed similar effect of the actions enables to consider consequent similarity between the resultant  $\Delta t$ , thus:

$$(A_i \Rightarrow \Delta t_i \wedge A_i \Rightarrow \Delta t_{i+3}) \Rightarrow \Delta t_i \approx \Delta t_{i+3} = \Delta T_i \quad (4.1)$$

$$(A_{i+1} \Rightarrow \Delta t_{i+1} \wedge A_{i+1} \Rightarrow \Delta t_{i+2}) \Rightarrow \Delta t_{i+1} \approx \Delta t_{i+2} = \Delta T_{i+1} \quad (4.2)$$

As expressed in Figure 4.6 this approach also enables the aggregation of the cases,  $\mathcal{C}_a$ , in order to facilitate their comparison and the extrapolation of conclusions. The use of  $\Delta T$  enables the aggregation of the cases into subsets accordingly to the value of  $\Delta t_i$ .

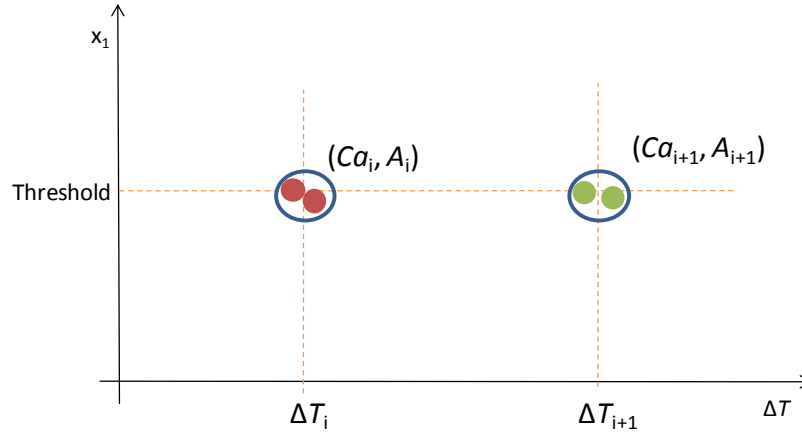


Figure 4.6. Correlation between  $x_1$  and  $\Delta T_i$

This reinforces the proposed approach stating that different actions will lead to different outcomes, and those outcomes can be measured in terms of frequency of occurrence of future problems.

Assuming that, instead of considering only one machine, a set of machines is accessible, contributing with cases that are related with the behaviour of  $x_1$  and  $\Delta t_i$  then the aggregation of these cases accordingly to their similarity, would enable the statistical analysis of the problem providing insight about what are the most probable outcomes of each possible action.

#### 4.2.4. Dealing with uncertainty

Maintaining the assumption that similar actions may have similar effects on the behaviour of a state variable, this may not be true in some situation. In fact, cases exist in which this state variable may present different temporal behaviour, i.e. the effect of similar actions is not certain.

This means that there are cases in which the application of a profound corrective measure may not lead to the expected result and others where the application of a quick fix may result better than expected. These cases are due to the uncertainty existing in this type of problems which is related to a series of factors such as:

- defects in materials;
- incorrect action development;
- external factors (e.g. temperature, humidity, etc.);
- operational factors (e.g. different operators using the machine), etc.

The existence of an uncertain component affects the behaviour of  $x_1$  and consequently  $\Delta t_i$ . One way of looking at the problem is considering that the differences in the behaviour of  $x_1$  are reflected in not re-setting the state to 1 as illustrated in Figure 4.5.

This way the previous proposal of having several  $\Delta t_i$  and simply aggregate them by action performed would represent loss of information since the uncertainty effect would be diluted in that major aggregation. Instead, and to deal with this issue, it is proposed to consider that each action

has a recommended  $\Delta t$ , provided by the manufacturer. This time interval represents the recommended medium time to repeat one action in order to have the system running on good condition. Although it is assumed that actions are only needed when a problem occurred, this interval will be used to categorize actions in terms of what might be expected when performing them. Thus, for each action, there will be a recommended  $\Delta t$ . Each time a problem occurs within that  $\Delta t$  it is considered that it is occurring when expected. On the other hand, problems that occur much sooner than  $\Delta t$  or much later must also be considered to guarantee good statistical treatment.

At this point it is proposed to stratify in three levels, short, medium and long, the  $\Delta t$  associated to each action. The number of problems occurring at each level will contribute to compute the cost of each action along time. Thus each time a problem occurs one has to check the action performed last time ( $A_i$ ), measure the time since that action and the current problem, and classify the problem accordingly with Table 4.1.

Table 4.1. Stratification of  $\Delta t$

Action ( $A_i$ )	
Time Interval	Level
Problems occurring in time interval $\in ]0, \Delta t - \frac{\Delta t}{3}]$	Short
Problems occurring in time interval $\in ]\Delta t - \frac{\Delta t}{3}, \Delta t + \frac{\Delta t}{3}[$	Medium
Problems occurring in time interval $\in [\Delta t + \frac{\Delta t}{3}, \infty[$	Long

With this classification it becomes easy to understand that an action registering more problems at “Short” may not be the best option to deal with the current problem. On the other hand, and since cost is here a major variable, the comparison must always be made considering both perspectives.

#### 4.2.5. Risk analysis

Up to now the influence that actions performed over the industrial plant may have in its overall status has been clarified. Nonetheless, the fact that those actions are only implemented in order to eliminate a detected problem must not be forgotten. Here the difficulty arises when there is the need to decide not only which action to perform, but also to deal with uncertainty in the actual cause of the detected problem.

In the proposed approach a set of state variables is defined which are monitored in order to assess its status. In case the status is abnormal (i.e. threshold has been violated), a symptom on the existence of a problem is generated. This symptom is nothing more than a request for maintenance (see Figure 4.2) which is sent to the Risk-based Decision Support System. When the symptom is received, the system searches for similar cases where that symptom was involved and it might come out with a set of cases that were caused by different causes. For example, imagine that  $x_1$  is being monitored by means of a sensor. When the system detects that  $x_1$  has

reached the threshold it is not possible to guarantee that the problem is indeed on  $x_1$  and not in the sensor. In this case the system will receive the same symptom, but without total confidence about its cause.

Having this in mind, the problem shifts towards the identification of which might be the cause of a specific abnormal situation that has occurred in the plant. This thesis proposes the identification of similar cases, trying to match the current situation with others about which the possible causes and actions to eliminate them are known.

The identification of the cause that should be considered is made by means of risk assessment strategy which, at the end of the process, provides the user with information on possible outcomes. If no optional actions and only one course exist then the outcomes will depend only on the cause of the problem and on a defined probability of occurrence for each outcome.

To use this approach the common definition of risk  $R$  is used which is the result of the product between the probability of occurrence,  $p$ , of a specific outcome,  $E$ , and the impact,  $I$ , of that outcome (i.e. the cost):

$$R = p(E) \times I(E) \quad (4.3)$$

If any of the relevant parameters defined for the plant crosses its threshold, the information provided by the system allows the estimation of the probability of a specific outcome to occur. Also, the estimated outcomes allow the computation of the level of risk that we are dealing with. The functions that describe outcome probability and associated impact can be introduced at set-up phase but they will be continuously refined during operation. The calculated risk is presented in cost units representing the cost of solving the problem, i.e. eliminating the cause (Marques & Neves-Silva, 2009).

Figure 4.7 shows the process of calculating the probabilities of outcomes associated to the detection of a specific symptom ( $S$ ). The symptom is the entity that enables the detection of some abnormal situation. As already mentioned this is not the cause of the situation or the situation itself. In fact the symptom just signals an abnormal situation, whose source can be on several causes ( $C$ ), each of them with one or more outcomes ( $Q$ ). The measurement of impact ( $I$ ) is associated with these outcomes.

The identification of the situation with the lower risk (i.e. lower cost) is done by navigation on the tree following the probabilities. Thus, based on (3.10) the calculation of the risk is done by:

$$R_i = p(C_i|S_i) \times p(Q_i|C_i) \times I(Q_i) \quad (4.4)$$

For each combination the method computes a value of risk which corresponds to the one associated with each cause. Note that it is assumed the independence of probabilities, i.e. if the cause is  $C_i$  then it will not be  $C_j$ , and even though it can derive to the same consequence,  $Q_i$ , they will not contribute simultaneously to it.

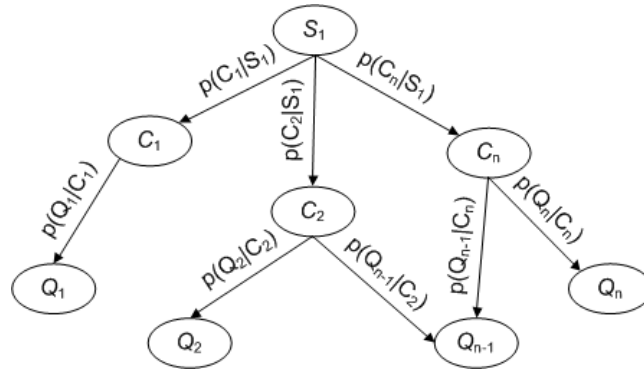


Figure 4.7. Risk analysis based on probabilities

The cause selected by the methodology is the one associated with the lower risk value calculated, since it is seen as the situation that represents the lowest cost in terms of maintenance.

#### 4.2.6. Decision model

Once available information about the industrial plant has been identified and correlated it is possible to establish its influence relations. At this point, and from a general point of view, information on symptoms, problems, causes and actions has been correlated. Nonetheless, even in cases where the methodology has identified the most probable cause there are still have several possibilities to deal with it... So, how can the user choose?

The existence of sufficient information about the behaviour of similar machines is a key aspect to enable the estimation of the probability of each outcome. In what concerns the outcome, and since the concept relies on the idea of minimizing the number of problems in the plant, then the course of action should be the one leading to the maximum “time to next problem”.

Thus, and recalling the tree analysis performed above, some important modifications are here introduced. In fact, instead of causes being directly selected by their probable outcomes, they become now mere possibilities which must be evaluated in a wider perspective, that includes the actions the user might choose to deal with the causes and their possible outcomes, at a longer term, considering the possibility of new problems arise (Marques & Neves-Silva, 2011).

Using the decision tree theory (Breiman, Friedman, Olshen, & Stone, 1984), (Murthy, 1998), (Quinlan, 1993)) it is possible to establish the model presented in Figure 4.8, where the user may choose between different courses of action considering the possible cause that might be associated to the detected symptom. Each path will lead to a different result which, in the proposed approach case, is measured in terms of “time to next problem”.

The introduction of time perspective in the tree has been studied in several domains, from stock market (Sap & Khokhar, 2004) to medicine (Yamada, Suzuki, Yokoi, & Takabayashi, 2003), representing the influence that time may have in some application. Time-series data consist of a set of time sequences each of which represents a list of values sorted in chronological order (Keogh, 2001). This perspective seems suitable to solve the structuring problem at this phase.

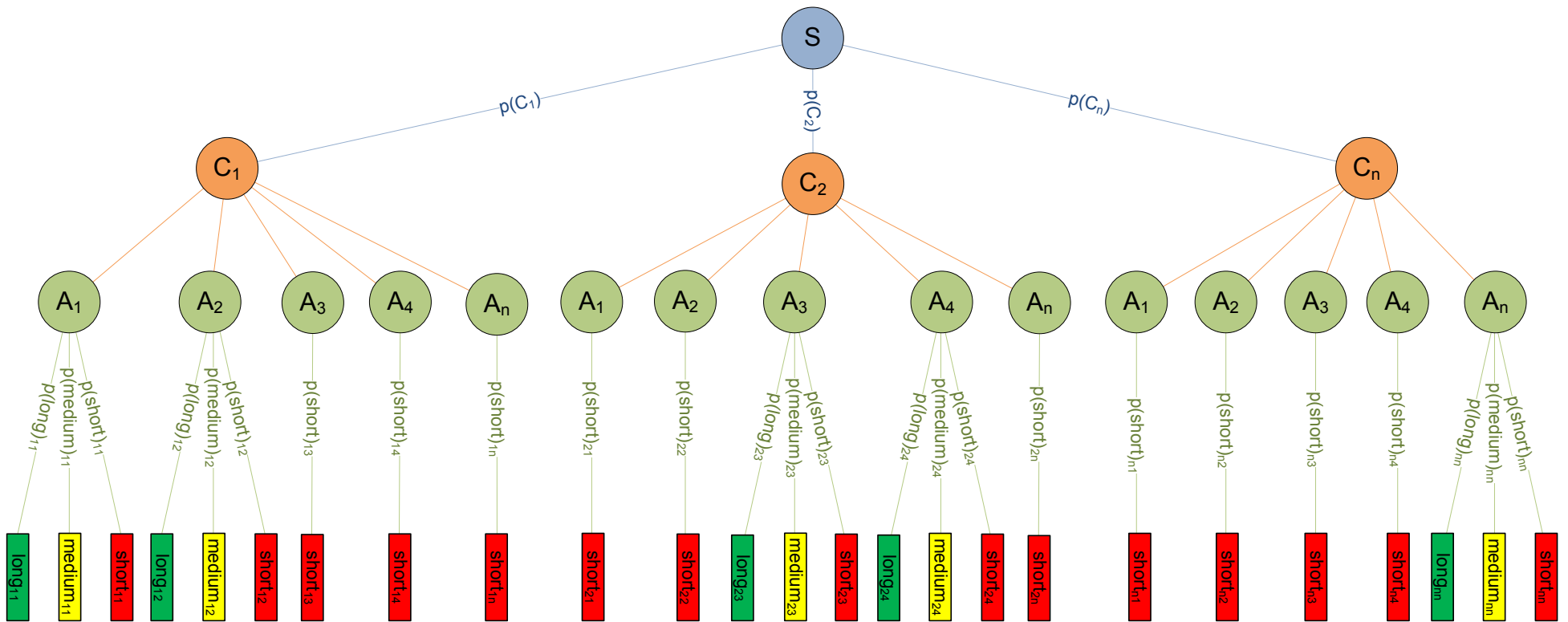


Figure 4.8. Decision model



Due to the characteristics of the problematic under analysis, this thesis proposes the use of the time-based clusters to structure the information on the tree. Following the results obtained in sections 4.2.3 *Information correlation and aggregation* and 4.2.4 *Dealing with uncertainty*, it is proposed that the clustering should be performed based on the time sequence of the problems together with an internal attribute which, in this case, will be the cause of the problem.

The time-based clustering uses a sort of fuzzy approach by means of the definition of crisp membership functions where the levels applied for the  $\Delta t_i$  stratification are used as linguistic terms (Short, Medium, Long). However, this clustering methodology is used only for the final leafs of the tree representing the past cases and by enabling the observation of the impacts those different actions may have. Next section details the way these clusters are build.

From a general point of view the modifications between this proposition and the one presented at Figure 4.7 were made at outcome level since what was introduced was the concept of different outcomes associated to the different courses of action. In fact and as already mentioned, the former strategy did not consider the action performed to affect the outcome. Instead those were calculated based on previous analysis of what could happen if nothing was done. In this approach the methodology tries to congregate the effect of the action in the future behaviour of the plant. This is done by following the probabilities attributed to each path. In the end, the solution will come in the form of an action to eliminate a cause, being that action the most cost-efficient for that cause, at that moment.

From a formal point of view the model accumulates aspects from the previously presented risk analysis model together with standard decision tree approach. In fact here the possibility of the user to follow different courses of action was added, which is a characteristic of the decision tree models. Nonetheless, and to maintain resemblance, the model was kept as similar as possible to the one presented in Figure 4.7.

The model was built following the logic of the problem, ensuring that at each probabilistic node probabilities along any outgoing branch sum to one. The expected result is achieved by rolling the tree backward (i.e., starting at the bottom and working towards the root).

On the tree, the value of a node can be calculated when the values for all the nodes following it are available. The value of a node is the expected value of the nodes following it, using the probability of the arcs. The tree can grow in complexity considering the number of causes that might be involved as well as the number of possible actions that may be selected.

#### **4.2.7. Proposed decision algorithm**

The algorithm proposed is then based on a twofold combination of probabilistic risk analysis and decision trees. It is assumed that the combination of these two methodologies enables to identify the most probable cause of a problem as well as best strategy to eliminate it, to be followed, i.e. the most appropriate action to be executed.

The starting point is the analysis of the collection of previous cases which represent the history of the system in terms of symptoms, causes and actions. Then, the probability of a specific cause is established by the number of previous cases, associated to a symptom, whose source was proved to be that cause. On a second stage the time elapsed, between cases with the same symptoms and causes, is used to estimate the effect that a specific action may have on the system behaviour. Having this result the risk analysis is performed in order to identify the option with lower risk level in terms of costs.

The use of the methodology can be put in a step-by-step procedure, as follows:

1. Once a symptom is detected, state vector  $x$  is generated, containing the set of state variables with relevant information for the characterization of the state of the plant in some particular instant of time when a *problem* is to be reported, such that:

$$x = [x_1 \dots x_n]^T \quad (4.5)$$

2. Collect all the cases,  $Ca$ , associated to the symptom detected,  $S$ :

$$Ca(S) = \{Ca_1, Ca_2, \dots, Ca_n\} \quad (4.6)$$

In this step, the methodology finds the stored cases which present appropriate similarity characteristics with the current problem,  $P$ . Nearest-neighbour retrieval is a simple approach that computes the similarity between stored cases and new input case based on weight features,  $w$ . A typical evaluation function presented by Kolodner (Kolodner, 1993) is used to compute nearest-neighbour matching:

$$Sim\{Ca_i, P\} = Sim(x^{Ca_i}, x^P) = \frac{\sum_{k=1}^n w_k Sim(x_k^{Ca_i}, x_k^P)}{\sum_{k=1}^n w_k} \quad (4.7)$$

Details on the calculation of  $Sim(x_k^{Ca_i}, x_k^P)$  can be found in annex C (Campos, 2010).

3. Discard all the cases which similarity level is below a specified threshold. The remaining ones are considered using the similarity level,  $Sim\{Ca_i, P\} = w_i$  as a weighting value.
4. Collect all the causes,  $C$ , of the remaining cases:  $C = [C_1 \dots C_n]^T$
5. Arrange the remaining cases in a matrix that relates  $Ca$  with  $C$ :

$$M_{CaC} = \begin{bmatrix} Ca_{1c_1} & \dots & Ca_{1c_n} \\ \vdots & \ddots & \vdots \\ Ca_{nc_1} & \dots & Ca_{nc_n} \end{bmatrix} \quad (4.8)$$

6. For each cause calculate the probability of that cause to be the correct one, considering the similarity level of each case,  $w_i$ , the frequency of the case,  $f$ , and the total number of cases associated to that symptom considering their frequency:

$$p(C|S) = \frac{\sum_{i=1}^n w_i f_i}{\sum_{i=1}^n f_i} \quad (4.9)$$

7. For each action,  $A$ , collect the time intervals,  $\Delta t$ , established by the manufacturers as the recommended ones to perform that action. Use that value,  $\Delta t_A$ , as a reference value to stratify the time into the following intervals:

$$L_{\Delta t_A} = \begin{cases} ]0, \Delta t_A - \frac{\Delta t_A}{3}] \\ ]\Delta t_A - \frac{\Delta t_A}{3}, \Delta t_A + \frac{\Delta t_A}{3}[ \\ [\Delta t_A - \frac{\Delta t_A}{3}, +\infty[ \end{cases} \quad (4.10)$$

8. Organize each element of  $M_{CaC}$  in chronological order from the oldest case to the most recent one.
9. Calculate the time elapsed,  $\Delta t_{Ca}$ , between two consecutive cases:

$$\Delta t_{Ca} = Detection(Ca_i) - Detection(Ca_{i-1}) \quad (4.11)$$

10. Aggregate the cases into three levels: *short*, *medium* and *long*, considering the executed action  $A_i$  and accordingly to the calculated  $\Delta t_{Ca}$  and  $L_{\Delta t_A}$ , as follows:

$$\begin{aligned} [A_{iCa_{\Delta t}}] = \\ = \begin{cases} \text{if } \Delta t_{Ca} \leq \Delta t_A - \frac{\Delta t_A}{3} \Rightarrow Ca_{i-1} \in [A_{iCa_{Short}}] \\ \text{if } \Delta t_{Ca} > \Delta t_A - \frac{\Delta t_A}{3} \wedge \Delta t_{Ca} < \Delta t_A + \frac{\Delta t_A}{3} \Rightarrow Ca_{i-1} \in [A_{iCa_{Medium}}] \\ \text{if } \Delta t_{Ca} \geq \Delta t_A + \frac{\Delta t_A}{3} \Rightarrow Ca_{i-1} \in [A_{iCa_{Long}}] \end{cases} \end{aligned} \quad (4.12)$$

Note that the cases are aggregated starting from the oldest one to the most recent. The last case in each element of  $M_{CaC}$  is not aggregated until a new case (with the same characteristics) occurs enabling the calculation of  $\Delta t_{Ca}$ .

11. For each action compute the probability of that action to result in short term problem, a medium term problem or a long term problem:

$$\begin{aligned} p((Short, Medium, Long)|A_i) = \\ = \begin{cases} p(Short|A_i) = \frac{\|[A_{iCa_{Short}}]\|}{\|[A_{iCa_{Short}}]\| + \|[A_{iCa_{Medium}}]\| + \|[A_{iCa_{Long}}]\|} \\ p(Medium|A_i) = \frac{\|[A_{iCa_{Medium}}]\|}{\|[A_{iCa_{Short}}]\| + \|[A_{iCa_{Medium}}]\| + \|[A_{iCa_{Long}}]\|} \\ p(Long|A_i) = \frac{\|[A_{iCa_{Long}}]\|}{\|[A_{iCa_{Short}}]\| + \|[A_{iCa_{Medium}}]\| + \|[A_{iCa_{Long}}]\|} \end{cases} \end{aligned} \quad (4.13)$$

12. Calculate the costs of the action  $A_i$  diluted along the accumulated time for each vector:

$$PartialCost(A_i) = \begin{cases} ShortCost(A_i) = \frac{\| [A_{iCaShort}] \| * Cost(A_i)}{\sum \Delta t_{CaShort}} \\ MediumCost(A_i) = \frac{\| [A_{iCaMedium}] \| * Cost(A_i)}{\sum \Delta t_{CaMedium}} \\ LongCost(A_i) = \frac{\| [A_{iCaLong}] \| * Cost(A_i)}{\sum \Delta t_{CaLong}} \end{cases} \quad (4.14)$$

Note that these partial costs represent the impact of applying  $A_i$ . In this sense they correspond to the outcomes presented in (4.4).

13. Calculate the partial risk associated to  $A_i$ :

$$\begin{aligned} PartialRisk(A_i) &= p(Short|A_i) \times ShortCost(A_i) \\ &+ p(Medium|A_i) \times MediumCost(A_i) \\ &+ p(Long|A_i) \times LongCost(A_i) \end{aligned} \quad (4.15)$$

14. Calculate the total risk considering the cause probability calculated in (4.9). Thus, for each cause, the total risk is given by:

$$TotalRisk(C_i) = P(C_i|S) \times \sum_{k=1}^n PartislRisk(A_k) \quad (4.16)$$

15. Select  $C_i$  that presents lower *TotalRisk*.  
16. Select  $A_k$  that presents lower *PartialRisk*.

The results aim at minimising the costs of intervention and for that reason the minimum risk is selected as the appropriate maintenance strategy to be suggested.

#### 4.2.8. Numerical example

Consider the following example which is focused on the risk analysis and decision tree part of the algorithm. Let's assume the existence of two causes  $C_1$  and  $C_2$  that can be associated to the detection of symptom  $S_1$ . Additionally, consider a set of four actions, two for each cause, that can be followed to deal with the cause, i.e. to try to eliminate the generated problem (see Table 4.2).

Table 4.2 Specifications for numerical example: Symptom, Causes, Actions and Costs

Symptom	Cause	Action	Cost
$S_1$ – Temperature above desired limit	$C_1$ – sensor failure	$A_1$ – clean sensor	15€
		$A_2$ – replace sensor	40€
	$C_2$ – cooling system failure	$A_3$ – replace cooling liquid	20€
		$A_4$ – replace cooler	60€

Since there are doubts about the exact cause that originated the problem with Symptom  $S_1$ , then the situation constitutes a risk management problem, which must be solved considering:

1. What is the probability of selecting the right cause;
2. Which cause is the most risky one;
3. Which are the costs of selecting the wrong cause.

When trying to identify the most probable cause of a problem one has to look at the history of the system and find the common causes, associated to a specific symptom, during its life time.

Here the method congregates both local information (i.e. information that was generated by the machine in which the problem was detected) and information that was generated by other similar machines in order to work up on a more complete data set.

Let's consider a set of 20 cases previously observed, stored and eliminated, i.e. their cause was well identified and the action performed succeeded in solving the problem. Additionally, let's consider that from those 20 cases, half were caused by  $C_1$  and half were caused by  $C_2$ , and that a new problem  $P$  has just occurred and needs to be solved. If all the past cases have the same weight then it is possible to directly compute the following probabilities:

$$p(C_1|S_1) = 0.5$$

$$p(C_2|S_1) = 0.5$$

This result is not very helpful in terms of selecting the best action to eliminate the problem. Note that at this point the example reached step 6 of the algorithm presented above. Thus, if no more information could be used let's apply the entire methodology proposed to help identifying the best option (i.e. action), in terms of cost, to deal with this problem. In order to find this best option let's compute the costs associated to each possible path.

Thus, the next step consists in using the defined recommended time interval for each action based on manufacturer recommendations. This interval is used for problem classification in terms of occurrence. Table 4.3 presents the intervals defined for each action and the resultant stratification levels for Short, Medium and Long term.

Table 4.3. Specifications for numerical example: intervals for maintenance actions

Action	$\Delta t$	Short $(]0, \Delta t - \frac{\Delta t}{3}])$	Medium $(] \Delta t - \frac{\Delta t}{3}, \Delta t + \frac{\Delta t}{3}[)$	Long $([ \Delta t + \frac{\Delta t}{3}, \infty[)$
$A_1$	10 days	$]0, 6.67]$	$]6.67, 13.33[$	$[13.33, \infty [$
$A_2$	30 days	$]0, 20]$	$]20, 40[$	$[40, \infty [$
$A_3$	20 days	$]0, 13.33]$	$]13.33, 26.67[$	$[26.67, \infty [$
$A_4$	40 days	$]0, 26.67]$	$]26.67, 53.33[$	$[53.33, \infty [$

At this point there is the need of grouping the problems according to their causes and actions, bearing in mind that the first problem that has occurred is used as a starting point to measure the time interval between problem occurrences for each action. Moreover, at this point, is time to aggregate problems in three sub-groups tagged as **short** (if they occurred in the time interval  $]0, \Delta t - \frac{\Delta t}{3}]$ ), **medium** (if they occurred in the time interval  $] \Delta t - \frac{\Delta t}{3}, \Delta t + \frac{\Delta t}{3}[$ ) or **long** (if they occurred in the time interval  $[ \Delta t + \frac{\Delta t}{3}, \infty[$ ).

Table 4.4. Cases caused by  $C_1$  and associated sub-group

Group 1			
		N° of problems (%)	Sub-group
C1	A1	33%	Short
		33%	Medium
		33%	Long
	A2	67%	Short
		0%	Medium
		33%	Long

Table 4.5. Cases caused by  $C_2$  and associated sub-group

Group 2			
		N° of problems (%)	Sub-group
C2	A3	33%	Short
		33%	Medium
		33%	Long
	A4	33%	Short
		33%	Medium
		33%	Long

Aggregation is made considering the actions that were performed to solve the problem when it occurred (see Table 4.4 and Table 4.5). For example, problem  $P_1$  occurred on 01.Jan and was treated with action  $A_1$ . When a new problem  $P_2$  occurs on 10.Jan it will be allocated under the form:  $C_1 \rightarrow A_1 \rightarrow \text{short}$ . If problem  $P_2$  was solved using  $A_2$  then, when a new problem occurs, let's say on 10.Feb, it would be allocated under the form:  $C_1 \rightarrow A_2 \rightarrow \text{medium}$ .

In the considered example the existing problems and the time frame associated to their occurrence led to the probabilities presented in the above tables.

Since  $A_3$  and  $A_4$  are not suitable to deal with  $C_1$  (as well as  $A_1$  and  $A_2$  are not suitable to deal with  $C_2$ ) they will not solve the problem which will result in 100% of probability of a new problem at short time. According to the proposed decision algorithm the cost of the action must be divided by  $\Delta t$  to compute the total cost, thus, and for the sake of significance, it was considered  $\Delta t = 1$  in these cases. This means that the total cost of an action is wasted when the cause is not identified correctly.

With this information it is possible to build the appropriate tree which can be seen in Figure 4.9. The analysis of the results leads to a double step approach to reach the final decision. Note that the goal is minimize the risk and this risk is, due to the assumptions made above, the cost of implementing a specific action. Thus, to achieve the most proper solution there is the need to

consider not only the cost of the action but also the probability of that action solve problem, i.e. the probability of choosing the correct cause to eliminate, as well as the time along which the cost of a specific action can be diluted.

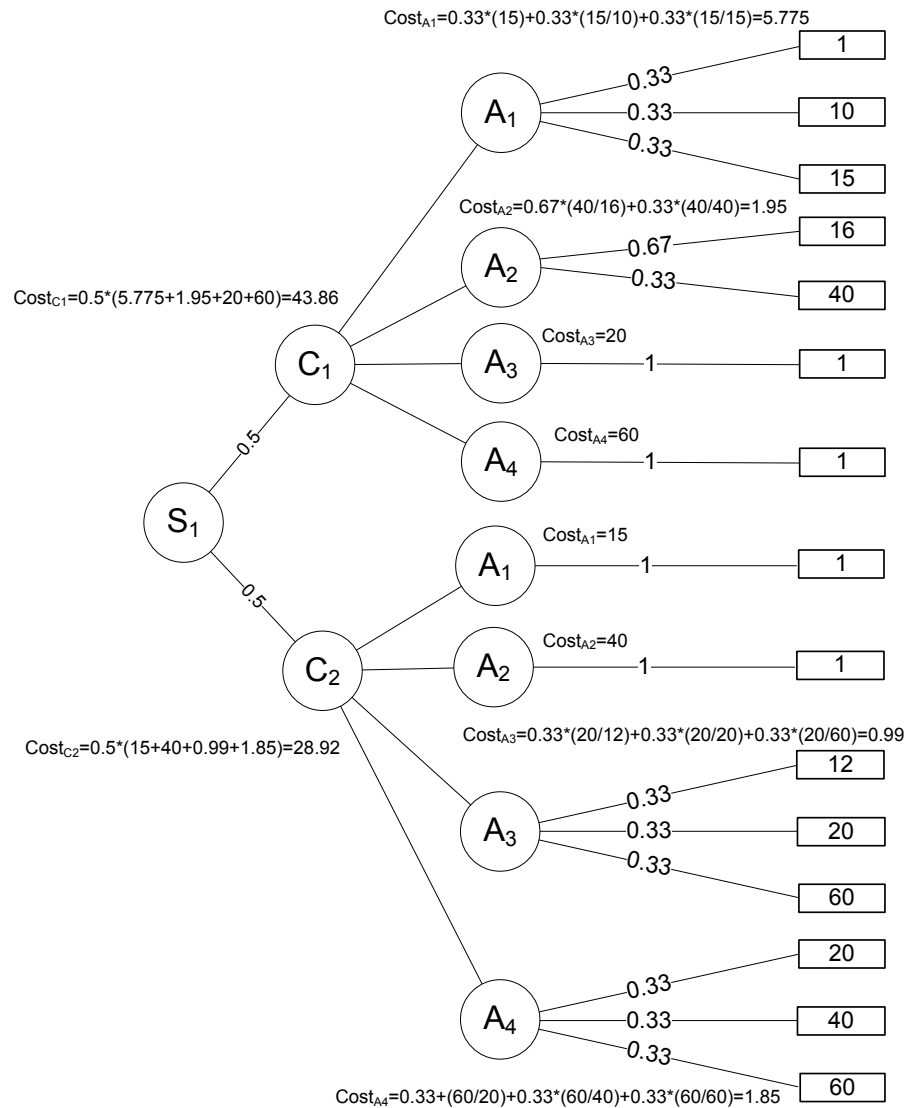


Figure 4.9. Decision Model for the current example

In the example considered the method suggests the use of action A<sub>3</sub> to eliminate cause C<sub>2</sub> since this is the path that presents the lower associated costs. Nonetheless, and with the continuous operation of the plant, this recommendation may change as the system learn more about the plant and its reaction to the applied actions.

### 4.3. Conclusions

The work presented in this chapter aimed at detailing the theoretical methodological aspects of the approach proposed in this thesis. Starting with the identification of the problematic to be treated, the concept to improve life cycle management of industrial plants through the minimisation of the life cycle maintenance costs was proposed.

A model for the knowledge repository to be used as the base for the decision process in industrial plants is proposed. The model congregates the aspects identified as crucial for the development of the proposed methodology.

The aimed cost minimisation is achieved by means of a risk analysis strategy which takes into consideration the different impacts of each action over the plant and uses those impacts to compute the possible outcomes.

Accordingly with these outcomes the most adequate action is proposed to eliminate the cause with lower associated costs. Note that the results of the algorithm are highly dependent on the knowledge inserted, thus the introduction of appropriate knowledge at set-up phase contributes to improve algorithm performance.

The algorithm proposed was formalised in order to facilitate its implementation and testing. Additionally the numerical example presented helped to illustrate how the algorithm effectively works and how the results are obtained. Nonetheless, without further testing, it becomes hard to assess the impact of the approach in the life cycle of an industrial plant and validate the proposed algorithm. Unfortunately, testing this algorithm in real industrial environment is difficult since its results are highly dependent on the number of occurred problems, thus depending on the number of undesired events. For this reason the development of a simulator, where problems could be freely created, was identified as a mandatory step of algorithm validation.

On the other hand, testing the algorithm in real industrial environments remained as a crucial part of the methodology validation. For this reason, and to enable testing some parts of the algorithm, namely the parts related with the model and its relations, an industrial prototype was also developed.

Both validation methodologies are described in the following two chapters, constituting the validation of the entire methodology.



## 5. *Validation through simulation*

---

Simulation is a common method when trying to assess the eventual real effects of alternative conditions and courses of action. The need for a simulator mainly relies on two factors:

1. the impossibility of having access to the real system, which may be due to several reasons, namely, assuring system safety and prevent eventual costs;
2. the difficulty of realizing identical tests, assuring repeatability, when dealing with real environments, which ends in lack of consistency in the results obtained.

These reasons are clearly valid when dealing with industrial plants. Additionally, and since the objective is to achieve results in terms of the plant life cycle, the simulator had to developed accordingly. In fact, the real results in terms of plant life cycle would be hard to obtain since they require long operational periods that overcome the normal duration of a PhD thesis.

Thus, and to test the proposed methodology, a simulator was developed congregating the simulation of the plant with the risk-based decision support system. The simulation was also used to test the learning capabilities of the algorithm. The following sections describe the details of the simulator together with the simulation results obtained.

### **5.1. *Development of the simulator***

The development of the simulator followed the general concept for the architecture of intelligent decision support systems presented in Figure 5.1, which recalls the concept specified in this work and presented in Figure 4.2. The main objective of Figure 5.1 is however to highlight the physical difference between the decision support system and the plant. When trying to develop a simulation model with these characteristics, i.e. with a physical part and a control part (many times performed by software) that acts over the physical one, it is important to keep this difference in mind so that, in the end, it results in a realistic model.

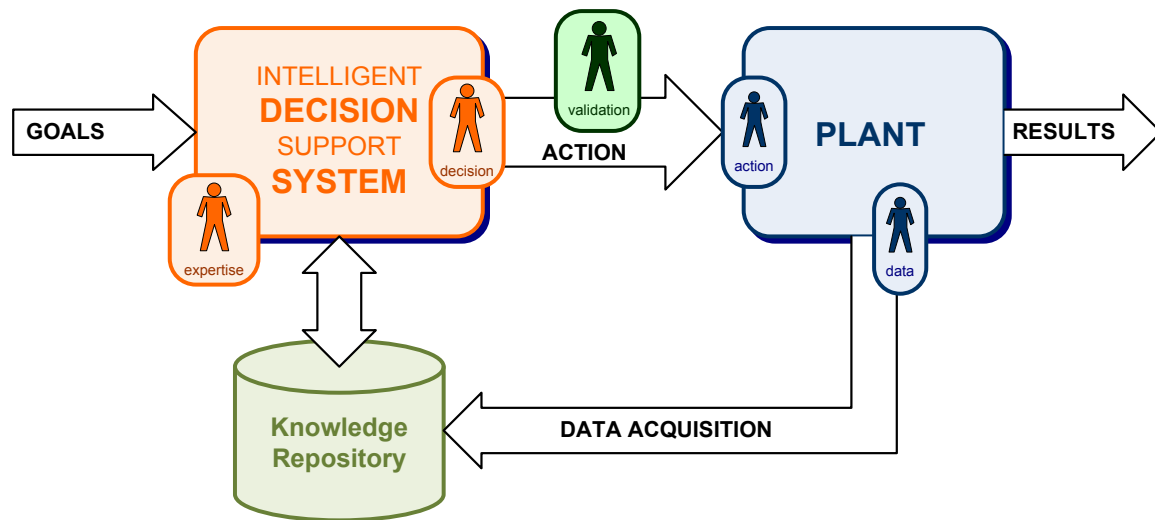


Figure 5.1. Conceptual architecture of an Intelligent Decision Support System

The workflow presented in Figure 5.2 illustrates, from a simple point of view how the simulator works. After pressing the “Start” button, the first action is to check if there is any information at the knowledge repository that can be used to build a first decision model. This initial decision model is not critical for the system to work, but it is crucial when one wants to achieve results with good confidence level in a shorter term.

Once the initial decision model is finished, the system will start plant simulation, which will continue until a threshold is trespassed and a symptom is generated. This is the driver to the risk analysis process that will generate a suggestion on which is the best course of action considering overall cost. This course of action is presented to the user, and after the final decision is made and executed, the decision model is updated with the new information. If the user does not stop the process it will continue returning back to the plant simulation.

As already mentioned and in order to have the system installed and running in several different industrial plants there was the need for developing a knowledge model, general enough that comprises all the relevant information on plant operation. The developed knowledge model, presented in section 4.2.2 *Knowledge model* was used as the structure of the Knowledge Repository representing the entities needed for developing the proof-of-concept. The result was thought to be a good balance between simplification of some aspects and the complexity that still is associated to the production process.

The simulator was developed using a set of software tools which are presented in Table 5.1. To fulfil the development needs a programming language, an Integrated Development Environment (IDE) and a database server were selected. The selection of the tools was mainly guided by the open source criteria.

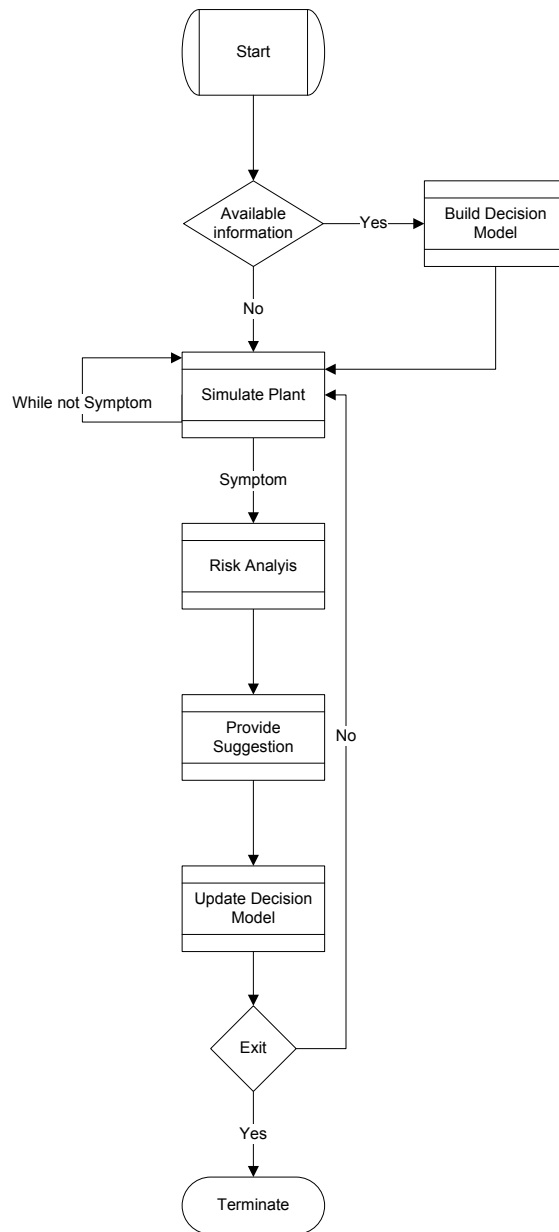


Figure 5.2. System workflow

Table 5.1. Software tools

Task	Tool	Motivation
Programming Language	Java	Its flexibility, portability and expressivity were the main criteria for choosing it
IDE	Eclipse	Its plug-in architecture and its extensibility together with the fact that it is free
Database Server	MySQL	One of the most popular Open Source database servers. It has proved its high performance and reliability in numerous development and production systems during the years

In the following sections details on the development strategy in terms of plant and decision support system are presented.

### 5.1.1. Industrial plant simulation

As previously mentioned the plant is the physical part of our system. The development of any model should try to capture the important parts of the plant in order to use them for testing purposes. Thus, models can vary according to their objective.

The approach here presented is based on the ideas about the state function presented in section 4.2.3 *Information correlation and aggregation*. From a general point of view this function goes from 1 to 0, over time, with the form illustrated in Figure 5.3. In this work it is assumed that each time a maintenance action is performed over the plant it resets the value of state to 1. Nevertheless, it is also assumed that not all the actions represent the same “level of heal”, for the plant. Once again, questions about uncertainty on the result of a maintenance action are here considered.

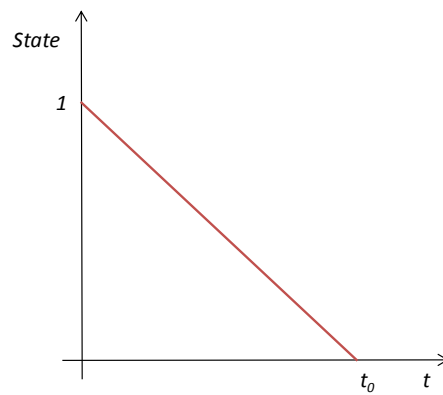


Figure 5.3. General state function

In fact, from the point of view of the plant, uncertainty must also be considered, since it will contribute for the behaviour of the plant each time an action is implemented. To comply with these requirements the following approach is proposed.

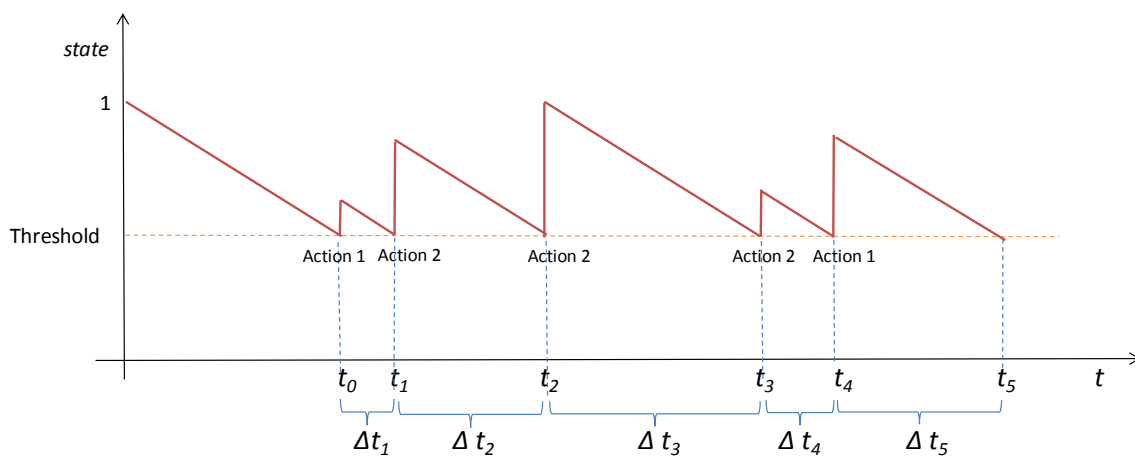


Figure 5.4. Uncertainty on the result of an action using an approximated state function

Consider Figure 5.4, where Figure 4.5 is recalled, using the proposed state function. Considering that two actions can be used to deal with one problem caused by an identified cause and that

both actions were used in different occasions with different results in terms of  $\Delta t$ . These considerations are reflected in the different re-set results (i.e. sometimes the state is re-set to one whereas other it is not).

One way of dealing with this issue when modelling the plant is computing different re-set values depending on the action implemented and on random event responsible for introducing uncertainty component in the process. Another option, which is thought of producing similar results, consists in maintaining the slope and recalculating the threshold. Let's consider an action  $A_i$  with the conditional probability expressed in Table 5.2. This means that each time  $A_i$  is executed there is  $p_{short}$  probability of a new problem occurring in short time,  $p_{medium}$  probability of the problem occurring in medium time and  $p_{long}$  probability of the problem occurring on long time. Also, these probabilities are not known by the decision support system. In fact they have to be defined in the plant model using *a priori* knowledge on the effect of each action.

Table 5.2. Example of conditional probability for action  $A_i$

$p((Short, Medium, Long) A_i)$	
$p(Short A_i)$	$p_{short}$
$p(Medium A_i)$	$p_{medium}$
$p(Long A_i)$	$p_{long}$

Where:  $p_{short} > p_{medium} > p_{long}$ .

Since it was assumed that state could vary between  $[0;1]$ , where 1 represents maximum life-expectancy and 0 stands for break, let's assume the same variation interval for threshold. Additionally, and since  $\Delta t$  was stratified into three levels, let's do the same with the threshold (see Figure 5.5). Note that  $p(x)$  is a uniform distribution, generated at the beginning of the simulation, which unique objective is to add a stochastic component to the plant behaviour.

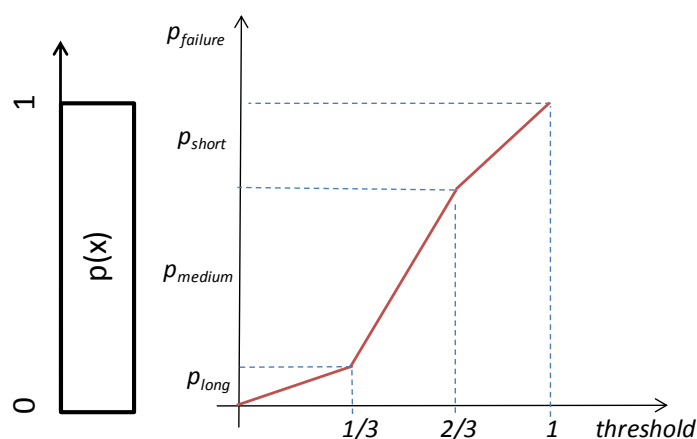


Figure 5.5. Calculation of the threshold accordingly with the action performed

This approach for modelling the generation of failures is consistent with the one proposed by Park (Park, 1993) which defines a model for the probability of developing a failure based on a vector of

risk factors. The model corresponds to a normal distribution where the probability depends on the logit of risk<sup>21</sup> function (see Figure 5.6).

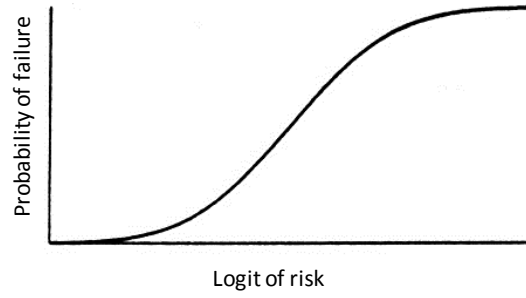


Figure 5.6. Probability of Developing a Failure vs the Logit of risk

In the approach proposed in this thesis the vector of risk factors is composed by one element, which is the running hours of equipment, i.e. time. This point of view is congregated in the threshold that is recalculated each time an action is performed over the plant.

With these conditions, each time a problem occurs and it is eliminated by means of  $A_i$ , a new threshold value must be computed using the following approach:

$$\begin{aligned}
 x_0 &= 0 \\
 x_1 &= p(short|A_i) \\
 x_2 &= p(short|A_i) + p(medium|A_i) \\
 x_3 &= p(short|A_i) + p(medium|A_i) + p(long|A_i) \\
 y_0 &= 0; y_1 = \frac{1}{3}; y_2 = \frac{2}{3}; y_3 = 1
 \end{aligned}$$

$$threshold = \begin{cases} \frac{1/3}{x_1} x, & \text{if } x \leq x_1 \\ \frac{1/3}{x_2 - x_1} x + \left( \frac{1}{3} - \frac{1/3}{x_2 - x_1} x_1 \right), & \text{if } x_1 < x \leq x_2 \\ \frac{1/3}{x_3 - x_2} x + \left( \frac{2}{3} - \frac{1/3}{x_3 - x_2} x_2 \right), & \text{if } x > x_2 \end{cases} \quad (5.1)$$

The result is a set of different thresholds that are settled each time a problem occurs and an action is implemented. Figure 5.7 congregates the resultant set of thresholds for the example of Figure 5.4.

<sup>21</sup> The logit is a measure of the total contribution of all the independent variables used in the model. In the model proposed by Park (Park 1993) it congregates all the identified risk factors.

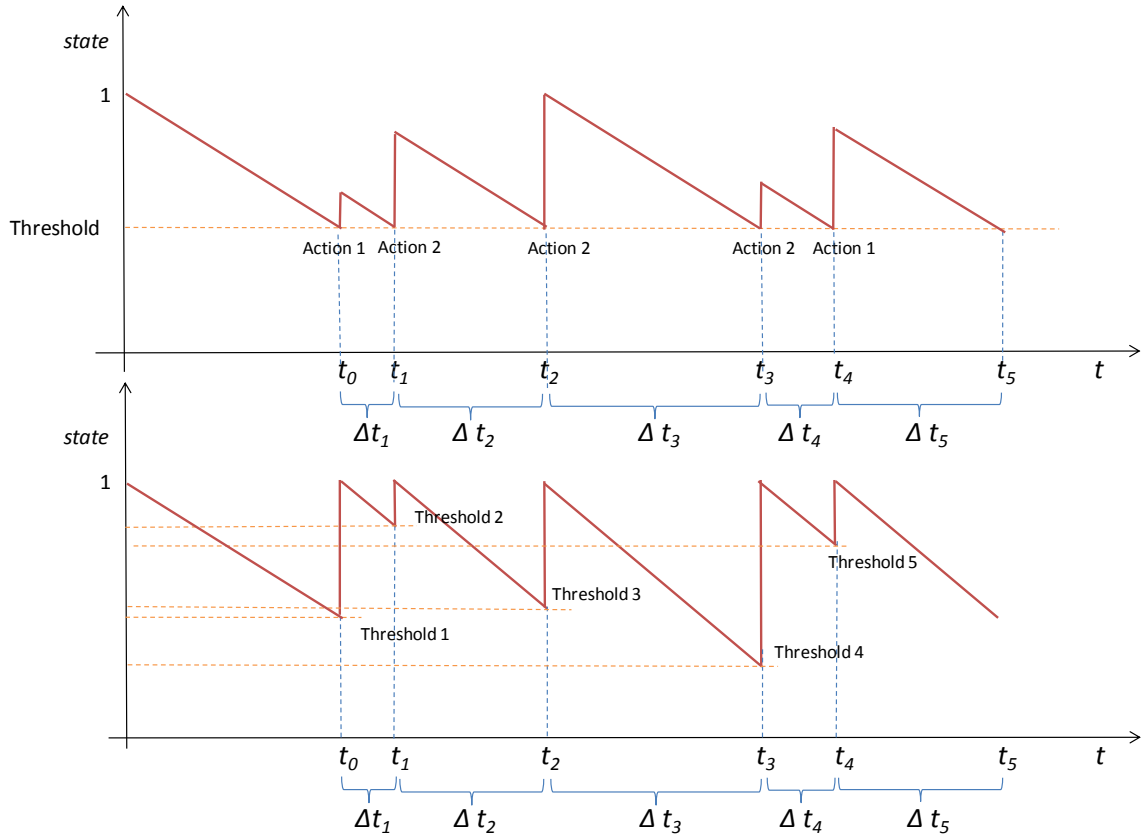


Figure 5.7. State function and its relation with threshold establishment

It is easy to conclude that both by recalculating the re-set value or by recalculating threshold similar results in terms of  $\Delta t$  are obtained.

As a final remark it is important to highlight that since the plant model was developed to be simulated on a computer there is the need to treat the state function not as continuous function but as a discrete one (see Figure 5.8). Thus, the discrete state function is given by:

$$state_{k_i} = state_{k_{i-1}} - \Delta k \quad (5.2)$$

Where  $\Delta k$  is established accordingly with the number of iterations needed to reset the state:

$$\Delta k = \frac{1}{n_{iterations}} \quad (5.3)$$

Each time  $state < threshold$ , a symptom is fired and a problem is generated.

Figure 5.9 presents the general workflow for plant simulation. Recalling the general simulation workflow presented in Figure 5.2, here it is addressed the part responsible for the behaviour of the plant.

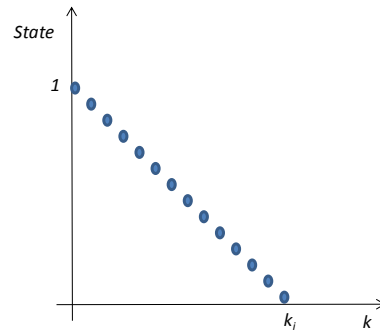


Figure 5.8. Discrete state function

When the process is initiated the first step is to check the thresholds set for the production units belonging to the plant. Once one of those thresholds is trespassed a symptom is generated and stored at the knowledge repository and the Risk analysis is initiated.

Note that each time this occurs the plant simulation is halted and the simulation is resumed once the problem is solved. This means that all production units will maintain their state during this halt and continue their normal behaviour afterwards. This is true except for the production unit which threshold was trespassed. In this one, and since a maintenance action was developed on it, the state is reset to 1 and the threshold is recomputed accordingly with the action performed.

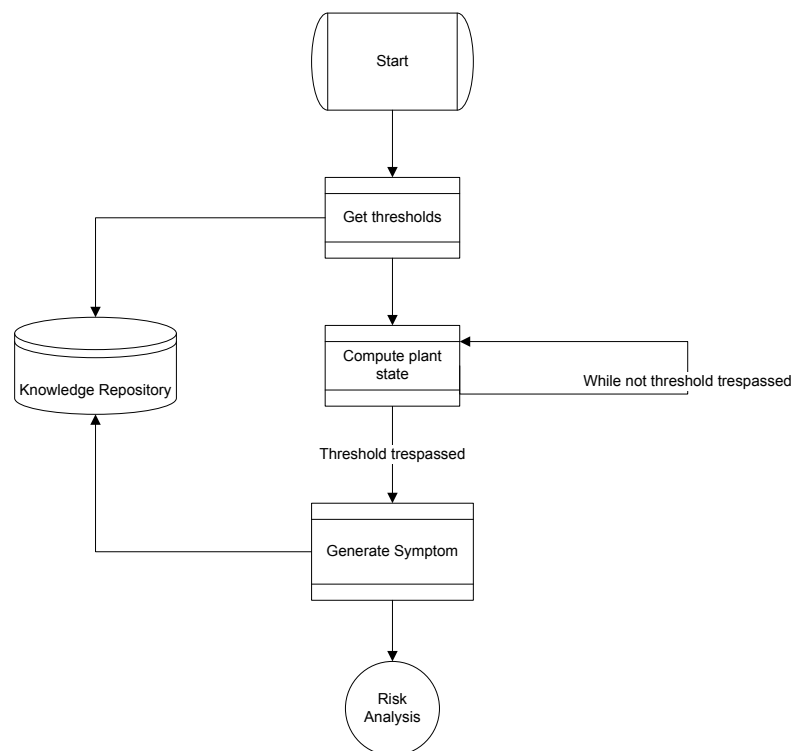


Figure 5.9. Plant workflow

### 5.1.2. Decision support system

The system intends to support the user in developing an improved maintenance strategy to help reducing the life cycle costs of machines and thus contribute for an improvement of life cycle



management of industrial plants. This support is provided by means of knowledge relative to past similar situations and specifically regarding the strategy followed in those cases, as well as the possible impact that a specific decision might have in the current situation.

The information is presented to the decision maker taking into consideration criteria, specifically the impact, in terms of costs, of a possible course of action. Once the information is provided, the user should be able of choosing the most appropriate strategy.

The system congregates both an automatic part and a user interface that requires input from the human operator to incorporate feedback. The automatic part of the process treats the problem in terms of the risk that the detected situation might represent for the plant, whereas the user interface registers the decision actually made and implemented together with its result, i.e. successful or not.

#### **5.1.2.1. Building the Decision Model**

The decision model is built at the starting phase of the overall process presented in section 5.1 *Development of the simulator* based on the information available, i.e. in the knowledge stored at the repository that can be used to build the model. This knowledge is mainly the existence of previous problems whose cause has been identified and the action implemented was efficient to eliminate that cause.

Since the focus is on analysing the time between problems it is imperative that the stored problems have a detection date associated so that the system is able of computing the time elapsed between them. Nonetheless, and if nothing else was previously introduced, the system needs at least one problem to start simulation. This problem is used as a starting point so that time between future problems can be computed.

Figure 5.10 presents the workflow for the process of building the decision model. Note that although using the same symbolic representation for the “Knowledge Repository” and the “Decision model” they have different implementations in reality. In fact the “Knowledge Repository”, as already explained, is a knowledge base *de facto* where both a priori knowledge and operational knowledge are stored under a specific format. This knowledge is maintained after system stops and can be used later when system starts. On the other hand the “Decision model” acts like a dynamic memory which is built at starting phase, updated during operational phase, but destroyed in the end.

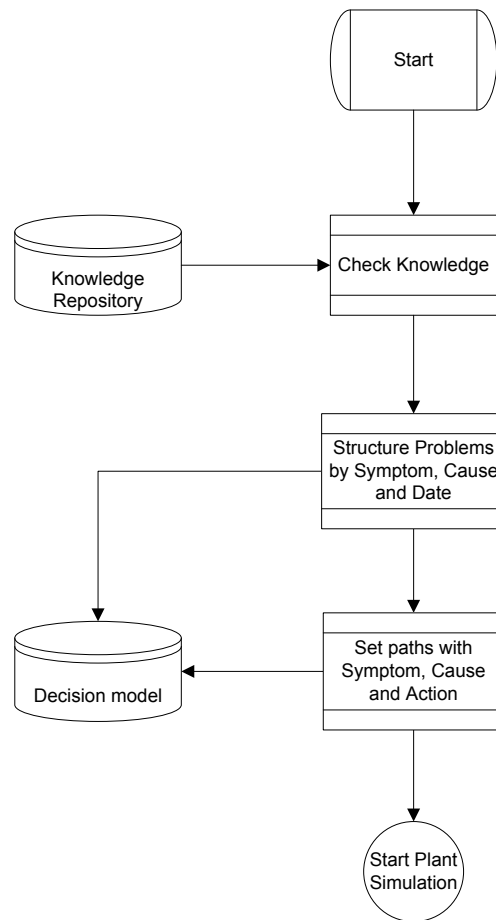


Figure 5.10. Workflow for building the decision model

#### 5.1.2.2. *Performing Risk Analysis*

Figure 5.11 illustrates the workflow for risk analysis which is initiated each time a threshold is trespassed at plant simulation. When this occurs the process will start by getting the structures that describe the decision model, using them to compute outcomes and risk for each path.

Using a minimize cost strategy the system finds the path with lower risk which comprises a cause and an action to deal with the situation. This is the path that will be provided as suggestion.

However, before providing the suggestion risk analysis will still be responsible for creating the new problem at the “Knowledge Repository” using for that all the information available including the detection date.

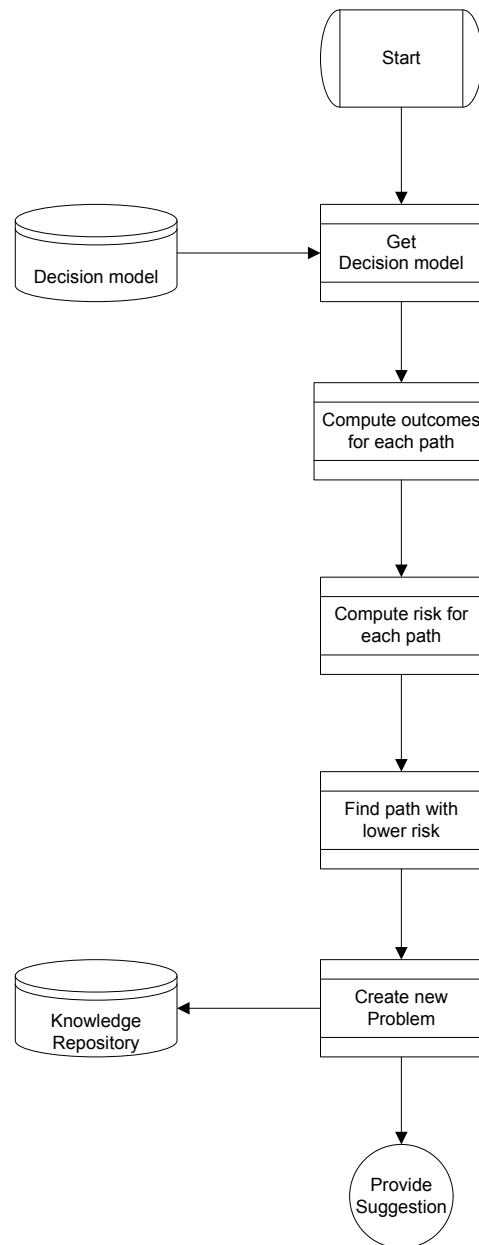


Figure 5.11. Risk Analysis workflow

#### 5.1.2.3. *Provide suggestion*

The provision of suggestion is initiated once risk analysis is finished. Figure 5.12 shows the workflow for this process.

The suggestion comprises both the cause and the action that were selected as the ones belonging to that path minimise costs, and they are presented to the user. Assuming that the user accepts the provided suggestion and implements the action with success, the problem will be eliminated and its status is updated accordingly.

Additionally, the plant status is also updated in terms of the threshold to be used in the production unit involved in the problem. This threshold will be used once the simulation plant is resumed.

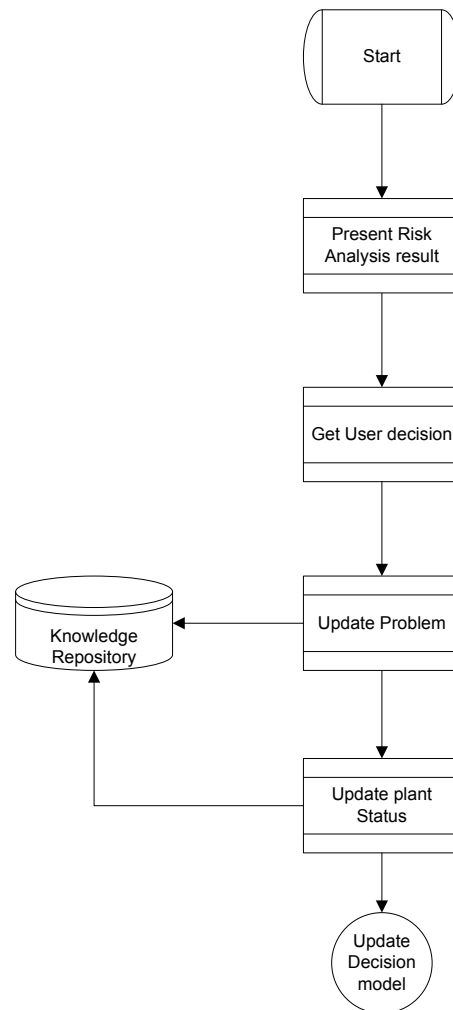


Figure 5.12. Workflow to provide suggestion

Note that, even though not detailed in Figure 5.12, different options from the user are possible to be considered. In fact the user can accept the suggestion or can force the selection of a different cause and/or action. This is useful in cases where the user is sure about which is the cause of the problem and the system is suggesting a different one. Additionally, it can also be useful when it is not possible to implement the suggested action and the user decides to go for a different one.

The process finalises by requesting an update to decision model, based on user confirmation that the problem was eliminated by the implemented action (which guarantees the correctness of cause identification).

#### 5.1.2.4. **Updating the Decision Model**

Figure 5.13 presents the workflow for updating the decision model. This process starts by getting updates from knowledge repository, which consist in the new problem added to the knowledge base.

Based on the structure of the decision model, presented in section 4.2.6 *Decision model*, the process allocates the new problem in the correct structure and path. The new decision model is

stored to be used from that point on. At the end of the process the system resumes plant simulation if no request to terminate the simulation is received from the user.

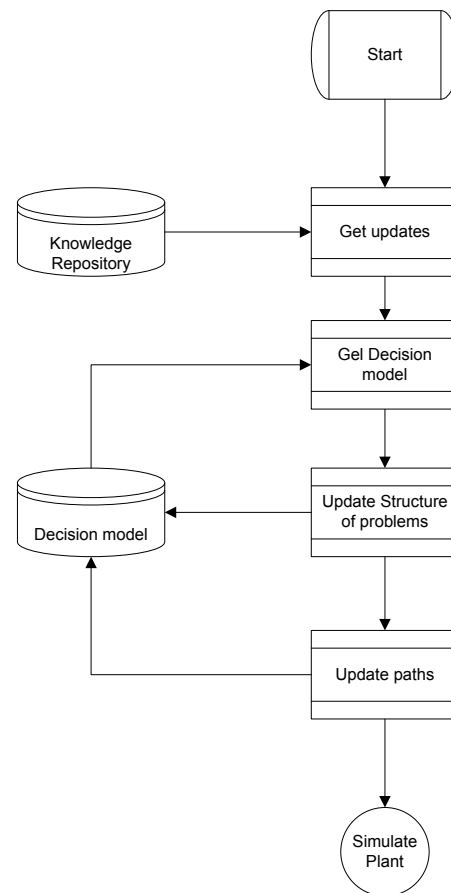


Figure 5.13. Workflow for updating the decision model

#### 5.1.2.5. User interface

In order for the system to work properly the introduction of information by the user is a major issue, starting from the set-up phase where experts on plant operation provide their knowledge until the stage where the success of a specific action must be communicated to the system.

To support these tasks appropriate interface should be developed providing information suited to user requirements and needs. Nevertheless, at this stage the focus is not on developing set-up interface. Instead attention is paid in developing appropriate interface for providing life cycle information. Thus the interface congregates information about the impact of specific actions in the plant, which is measured in terms of costs. The suggested action is the one that minimizes those costs. Using this strategy it is envisaged to achieve a minimization of the life cycle cost of the industrial plant.

As soon as a result is achieved, it is presented to the user in the form of a tree where it is possible to observe the different causes considered together with the possible actions to deal with them. Additionally, and to highlight the result, the (cause, action) pair suggested by the system is also presented in the results area (see Figure 5.14).

At this stage the user has three options:

- 1- Accept the result: is a confirmation that the provided suggestion did eliminate the problem;
- 2- Exclude one of the causes: useful in cases where a cause that was suggested is known as not being the correct one;
- 3- Select Action: useful when for some reason, it is not possible to implement the suggested action.
- 4- Start: restarts the simulation using the information stored;
- 5- Exit: exits the application.

Note that option 2 is meant to be used by people with extensive knowledge on the plant behaviour since it is based on the idea that the human operator is sure about the impossibility of that cause to be the correct one. Additionally, note also, that option 1, 2 and 3 resume the current simulation whereas option 4 restarts a new one.

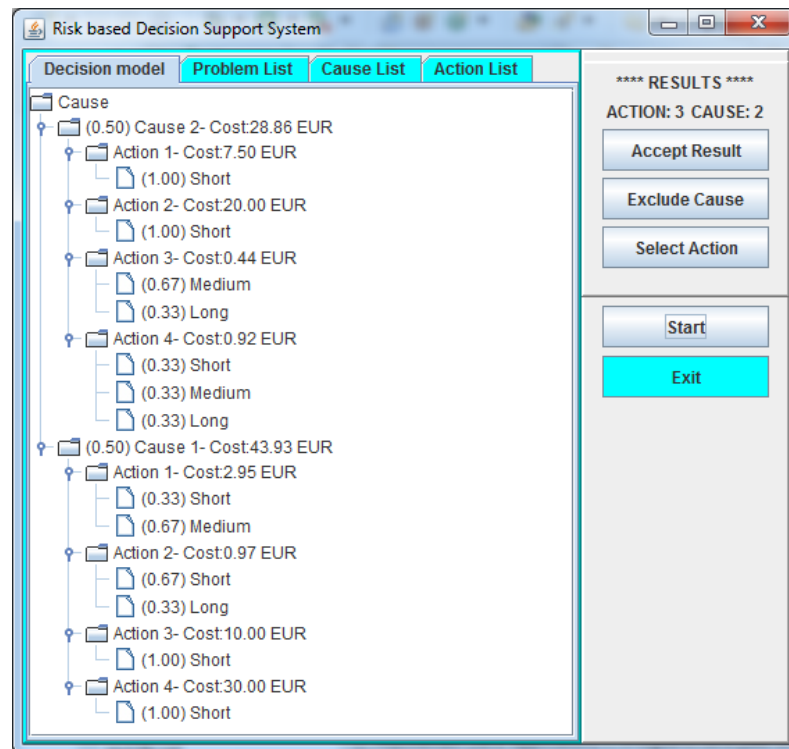


Figure 5.14. Risk based Decision Support System: Results Interface

For user information the simulator presents the status of the knowledge base in the tabs “Problem List”, “Cause List” and “Action List”. During simulation “Problem List” is updated with the new information being stored. The other two tabs are constant and cannot be changed using the simulator, instead the user must access the knowledge base by other means (e.g. set-up tool) to introduce new information on them.

## 5.2. Simulation results

The system was thoroughly tested in order to congregate solid results on its soundness. The learning capabilities were the first to be tested in order to ensure reliability of the results achieved. Afterwards the tests were concentrated on comparing the developed strategy against a set of standard well known (and widely used) maintenance methodologies. In the following sections the main results of these tests are presented.

### 5.2.1. Learning capabilities of the algorithm

At this stage the intention is to test the learning capabilities of the system which are based on “learn from examples” strategy. An incremental approach was used, in which the only input was a sequence of instances, each represented by a set of attribute-value pairs.

Also, to guarantee usability of validation results the tests were focused in two generally used performance indicators:

- **Accuracy:** The degree of closeness of measurements of a quantity to its actual (true) value.
- **Learning Rate:** The speed at which classification accuracy increases during training. It is a more useful indicator of the performance of the learning algorithm than is accuracy for finite-sized training sets.

To ensure repeatability the tests of the learning properties of the implemented algorithm were developed using the following methodology:

- 1- Definition of outcome probabilities for each possible action, considering each cause, in the plant side of the system;
- 2- Preparation of a set of starting examples, i.e. problems at the knowledge base, which can be reduced to 1;
- 3- Build the decision model considering the starting examples;
- 4- Each time a problem occurs:
  - a. Run the decision model to select the path with lower cost, which corresponds to most suitable action for the most probable cause;
  - b. Select that cause as the cause for the problem, and assume that action as the one implemented to eliminate it;
  - c. Update the decision model to incorporate the information on that new case.

Note that, in this testing strategy the user is not supposed to interact with the system. In fact the intention of the test is to assess the capability of the decision support system to correctly estimate the probable outcomes for each action defined in the plant system. This estimation is based on the time elapsed between problems each time a specific action is implemented, and on the number of problems associates to that action (i.e. problems that occurred each time that action was implemented to eliminate a cause).

The first step consisted in the definition of the outcomes probabilities for a generic action  $A_1$  which were set to:

$$p(short|A_1) = 0.8$$

$$p(medium|A_1) = 0.15$$

$$p(long|A_1) = 0.05$$

Additionally a single problem was introduced at the knowledge base, to be used as a starting point for the calculation of time between problems.

The algorithm was tested using two different horizons in terms of problem generation in order to highlight the refinement of results with the increase of the information available.

#### 5.2.1.1. Accuracy testing

##### Generating 50 instances

The first test was set to produce 50 instances of problems and the results found for the outcomes probabilities were:

$$p(short|\widehat{A_1}) = 0.92$$

$$p(medium|\widehat{A_1}) = 0.04$$

$$p(long|\widehat{A_1}) = 0.04$$

Comparing these values with the ones defined previously, they represent a medium error of approximately 15% for  $p(short|A_1)$ , of more that 70% for  $p(medium|A_1)$  and of 20% for  $p(long|A_1)$ . Graphically, it is possible to observe in Figure 5.15 the distribution for time between problems illustrating the results achieved with this test.

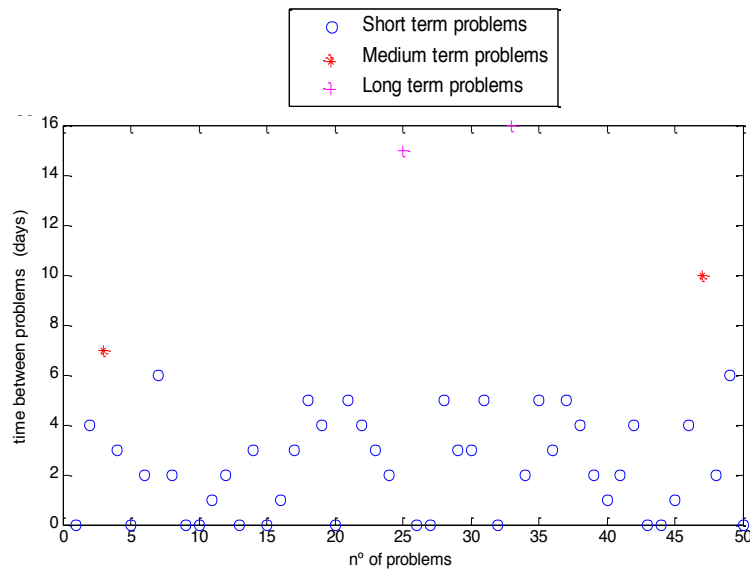


Figure 5.15. Testing accuracy with 50 instances



### Generating 300 instances

The second test was set to produce 300 instances of problems and the results found for the outcomes probabilities were:

$$p(\widehat{short|A_1}) = 0.82$$

$$p(\widehat{medium|A_1}) = 0.14$$

$$p(\widehat{long|A_1}) = 0.04$$

In this trial the values obtained represent an error of around 3% for  $p(\widehat{short|A_1})$ , of 7% for  $p(\widehat{medium|A_1})$  and of 20% for  $p(\widehat{long|A_1})$  when compared with the ones defined at starting phase. Graphically, it is possible to observe in Figure 5.16 the distribution for time between problems illustrating the achieved accuracy for this test.

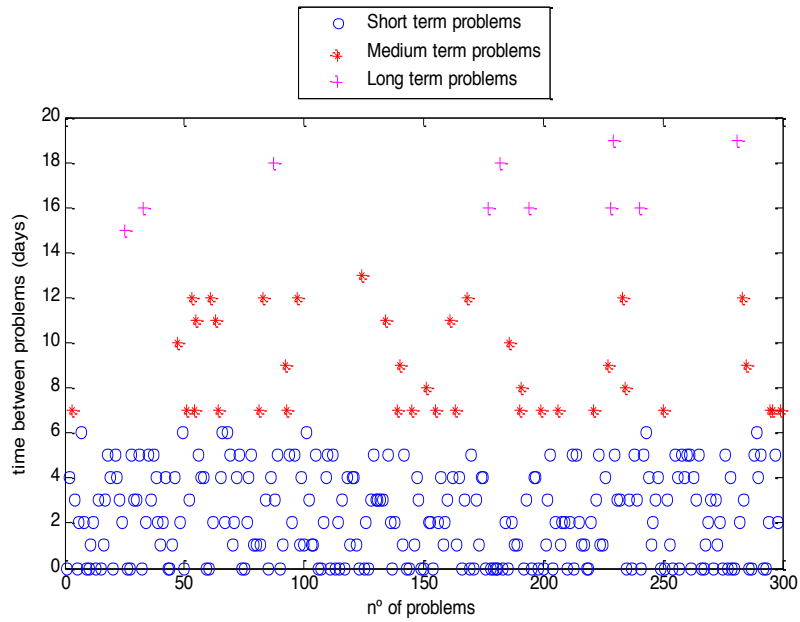


Figure 5.16. Testing accuracy with 300 instances

Note that besides  $p(\widehat{long|A_1})$  presents the same value as in the first test this second test is thought of being much more reliable than the previous one. This is due to the decrease of the error registered in the other outcomes.

#### 5.2.1.2. Learning Rate testing

For testing learning rate a comparison between two simulations is performed:

- 1- **Test with no previous knowledge:** the knowledge base is empty except for one single problem to be used as the starting point, to provide insight on the refinement of the outcomes probabilities calculation along time;
- 2- **Knowledge introduced at set-up phase:** made using previous knowledge on the action result, coming from past problems introduced at the knowledge base.

### Test with no previous knowledge

This test intends to simply show the results achieved when a substantial number of instances are generated. In this case we generated around 350 problem instances which were used to compute the outcomes probabilities of action  $A_1$ . Figure 5.17 illustrates the results achieved with this test.

At the end of the test it is possible to observe a refinement on the estimated values of the outcomes probabilities. It is reasonable to conclude that the results become more refined, i.e. reducing error, with the increase on the number of generated instances.

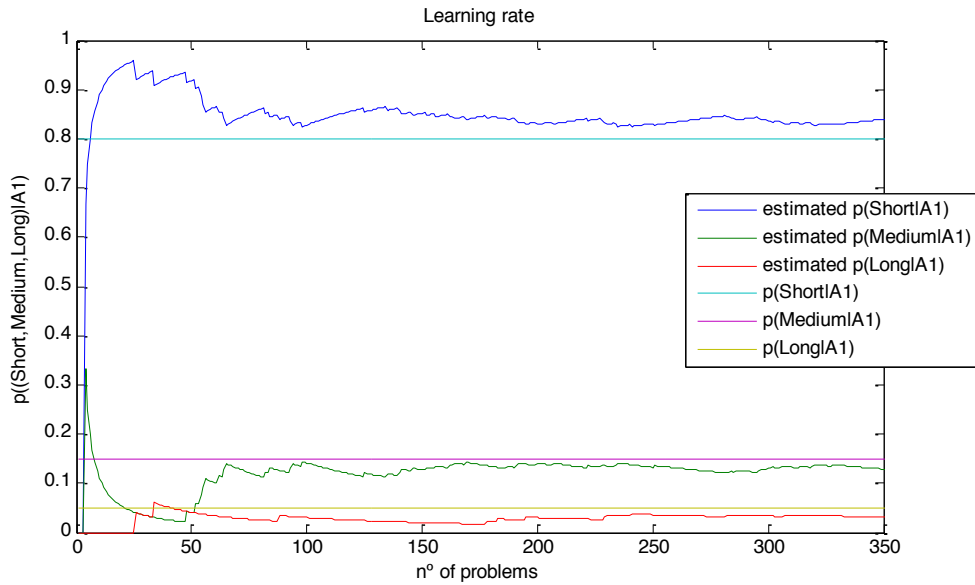


Figure 5.17. Learning curve with no previous knowledge at set-up

### Test with knowledge introduced at set-up phase

This test was elaborated to generate 150 instances of problems but, at set-up phase 50 instances were already introduced which were used to compute the starting point of outcomes probabilities for action  $A_1$ .

In Figure 5.18 it is possible to observe the behaviour of the three estimated curves for the outcomes probabilities associated to action  $A_1$  when 50 instances were available on the knowledge base at starting phase.

As expected the system presents a smoother behaviour at starting phase, converging rapidly for the predefined probabilities values. This is an important aspect in terms of system reliability as a whole since it represents the amount of information that the system needs to start producing reliable results. The observation of Figure 5.17 allows concluding that the system needs approximately 50 cases stored at the knowledge to provide reliable information to the user.

Nonetheless note that the final results are not very different from the ones observed in the above test. This means that, after a very fast starting phase, where the system rapidly converges to the predefined values, it enters a slower phase where successive refinements are made.

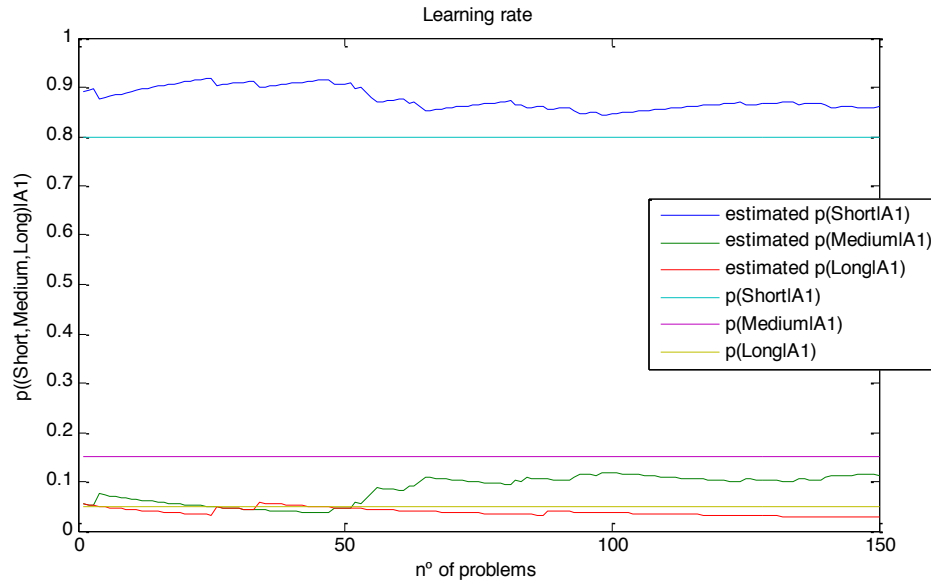


Figure 5.18. Learning curve with 50 problems introduced at set-up

### 5.2.2. Comparison with standard maintenance methodologies

In chapter 2 *Maintenance for life cycle management* the main aspects of a series of maintenance methodologies were detailed. In this section, and to test the effective results of the approach developed, a comparison of its results, in terms of accumulated costs, is performed. The comparison uses three of the methods described at that chapter as benchmarks, namely:

- run-to-failure;
- condition-base maintenance; and
- preventive maintenance.

To make the testing scenarios as real as possible there was the need to introduce some modifications in the system, namely in the part where the action to be executed is selected. Note that in the approach proposed, the action is selected after performing the risk analysis providing indication about the best action in order to minimize costs. This is not the case for any of strategies used for comparison thus the actions to be implemented follow a specified plan which is set for each scenario. Additionally, and to enable comparison, the same  $p(x)$ , representing the distribution used for problem generation, is the same in all simulations as well as the starting data at the knowledge base.

To compare the approaches a specific set of parameters concerning the actions to be applied were chosen, specifically three possible actions,  $A_1$ ,  $A_2$  and  $A_3$  were defined. The characteristics of the three actions are presented in Table 5.3.

Table 5.3. Characteristics specified for actions

Action	Cost (€)	Recommended Interval (days)	Description
$A_1$	15 €	10	A simple action (e.g. clean a part)
$A_2$	40 €	30	A more complete action (e.g. replace a part)
$A_3$	65 €	40	A complete repair

The costs and time intervals for action  $A_3$  were selected so that they sum the ones specified for  $A_1$  and  $A_2$  to guarantee a degree of consistency in the results.

Additionally, for each action set of outcomes probabilities were defined to transmit the effect they might have in the behaviour of the plant considering the specified characteristics. Once again, and for the sake of consistency, the probabilities for  $A_3$  were chosen based on the ones specified for  $A_1$  and  $A_2$ . The result can be seen in Table 5.4.

Table 5.4. Outcomes probabilities for specified actions

$A_1$		$A_2$		$A_3$	
$p(short A_1)$	0.8	$p(short A_2)$	0.05	$p(short A_3)$	0.425
$p(medium A_1)$	0.15	$p(medium A_2)$	0.15	$p(medium A_3)$	0.15
$p(long A_1)$	0.05	$p(long A_2)$	0.8	$p(long A_3)$	0.425

All the details about the test cases are explained in the following sections.

#### 5.2.2.1. Test case 1: Maintenance based on Run-to-failure strategy vs Maintenance supported by the Risk-based decision support system

As already mentioned the logic of run-to-failure is simply based on: do nothing if it is working. Thus, if the machine is not broken and continues producing, then no maintenance procedures should be implemented. The problem is that each time the machine is broken a complete repair is needed which is normally more costly than to apply corrective measures when something seems to be on the wrong path.

As already mentioned both cases were simulated separately but using the same  $p(x)$  which means that the generation of problems was based on the same data set. The difference is on the thresholds that are calculated when one action is applied.

Note that using run-to-failure is almost like playing a 50%-50% probability “game” in which the player might be lucky and get the most out of the machine (this is guaranteed by setting the threshold to 0, which stands for the maximum rupture point), or, in the other hand, can face a new problem right after making a total repair.

The resultant accumulated costs for this test case can be seen in Figure 5.19. The difference between the two approaches is clear with the maintenance supported by the risk-based analysis presenting much lower accumulated costs. This is due to the probabilities associated to the

complete repair of the machine which results in having short term problems even after performing a complete repair on the machine.

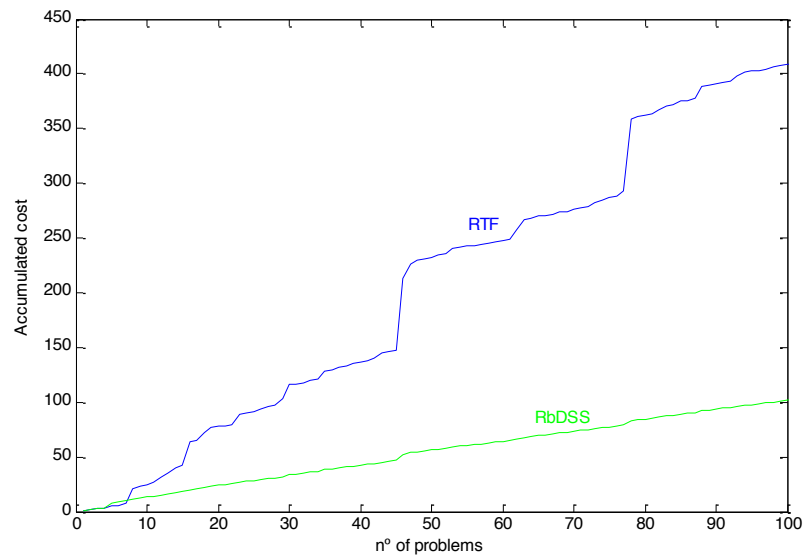


Figure 5.19. Accumulated costs for test case 1

In the run-to-failure curve it is possible to observe some fast cost increases (around problems 15, 45 and 75). These are due to complete repairs that ended in a new problem on short term. On the other hand the results obtained through the risk-based decision support system present a smoother behaviour since the action to be applied is being continuously selected based on the risk analysis strategy. This enables the system contribution for optimising the life cycle costs of the plant.

#### **5.2.2.2. Test case 2: Maintenance based on Condition-based strategy vs Maintenance supported by the Risk-based decision support system**

Condition-based maintenance relies on the assumption that the plant is equipped with a set of adequate sensors in order to detect any important deviation of the normal operating values. From this point of view, this assumption resembles the one adopted at the approach proposed in this thesis. However, the difference is when the detected problem may be solved using two, or more, actions. And this is where the approach proposed in this work may help, i.e. to support the decision that has to be made.

Nonetheless, one can say that this decision is only important if there are doubts in identifying the real problem. For instances if the plant has a sensor to detect if a part is dirty or if it is broken, then, recalling Table 5.3, there would be no doubts in which action should be applied. In this case, the decision would be on whether the investment on a comprehensive sensor net, which might be extremely costly, should be performed or not. This decision should be evaluated in each case for there is not a single answer for this issue.

In order to make the two approaches comparable it is assumed that this is not the case, i.e. the plant is not equipped with such an extended sensor net. Thus, for the Condition-based strategy each time a problem occurs, and in the absence of any additional information, the maintenance procedure follows a pre-established plan, based on the recommended intervals presented in Table 5.3, stating that:

- $A_1$  should always be executed three times in a row;
- When  $A_1$  was executed three times then execute  $A_2$ .

If the occurrence of problems was deterministic and problems only occur when the recommended interval has passed, this plan would lead to execute  $A_1$  three times, completing 30 days, and then execute  $A_2$ . The setback is that this is not the case, and problems will occur in unexpected times. Figure 5.20 presents the accumulated costs for both strategies (the same  $p(x)$  was used in both simulations).

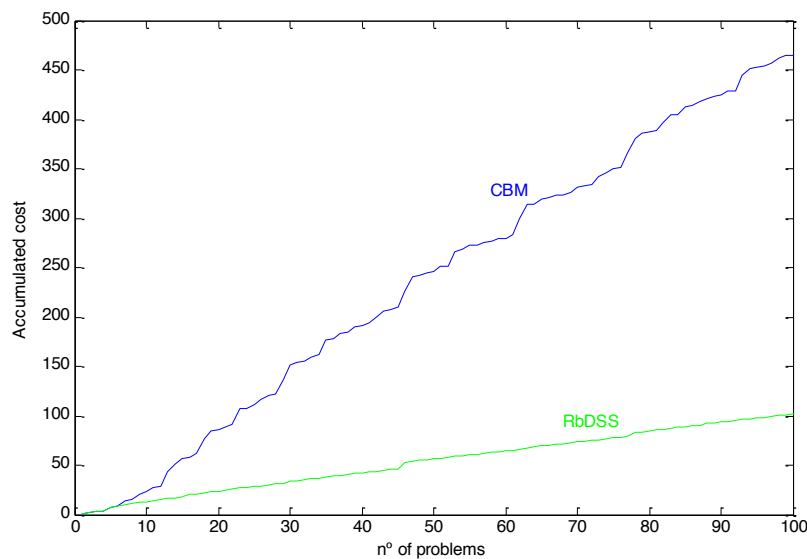


Figure 5.20. Accumulated costs for test case 2

Once again the difference between the two strategies is patent and once again the use of the proposed approach produces better results in terms of cost minimisation.

Nonetheless one can discuss the maintenance plan proposed and wonder about what would be the result of using a different one. For that reason three additional simulations were performed with different maintenance plans:

- using  $A_1$  and  $A_2$  in an alternate way;
- using only  $A_1$ ; and
- using only  $A_2$ .

The results obtained are presented in Figure 5.21. The worst strategy is the one that uses only  $A_1$ , which means that using the less costly action results in higher accumulated costs. The

alternate strategy also results in high costs although they are much lower than the ones achieved with the previous one. Finally, almost neck and neck, the costs achieved by the risk-based decision support system and the ones obtained by the exclusive use of  $A_2$  can also be observed. Note that, although their similarity, the strategy of using only  $A_2$  did perform better than the one provided with the risk analysis support which can be explained the following two factors:

1. the probabilities defined for  $A_2$  privileged its behaviour in the long term, making it a very good option in terms of cost minimisation;
2. Although the fast learning rate of the system, it still presents a starting phase in which its results are not so good. This implies that at the beginning phase the system is still trying to find the best option, between the ones available. After this phase, the system learns that the best options is indeed  $A_2$ , and from that point on the results of both approaches are similar. Nonetheless, the initial phase will affect the overall behaviour of the system, which is reflected in the accumulated costs. One way of solving this issue would be the implementation of a forgetting mechanism which would, at some point, ignore the cases which detection date was too far from present, even when their similarity is above the specified similarity value.

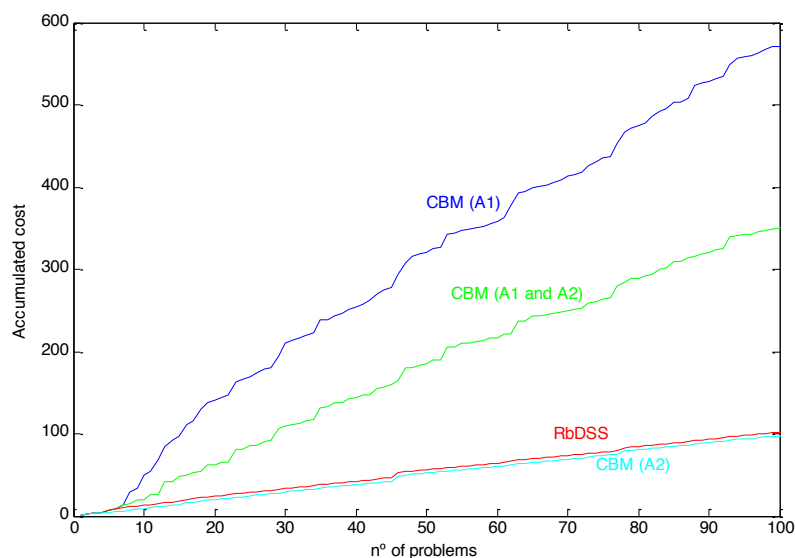


Figure 5.21. Accumulated costs for test case 2 with different maintenance plans

### 5.2.2.3. Test case 3: Maintenance based on Preventive strategy vs Maintenance supported by the Risk-based decision support system

On the opposite side of run-to-failure it is the preventive maintenance strategy. Much more conservative, and sometimes considered as a huge waste, for defending the implementation of maintenance strategies even in cases where any problem is detected.

Thus, when using preventive maintenance the aim is to prevent complete repairs of the machine and prefer to apply maintenance actions from time to time. And this notion of time driven actions must be introduced when simulating the system for comparison.

To maintain comparison settings the same specifications for the possible actions presented in Table 5.3 and Table 5.4 have been used. Additionally three simulations were performed, all of them based on the implementation of the following maintenance plan:

- $A_1$  should always be executed three times in a row;
- When  $A_1$  was executed three times then execute  $A_2$ .

The first simulation was made considering that the maintenance plan, if executed at the specified intervals, would prevent further problems. This means that, similarly to what was explained above,  $A_1$  would be executed each 10 days, for three times, then  $A_2$  would be executed one time, 30 days later the plan returns to  $A_1$  and so on. Note that it is assumed that no problems would occur and thus costs are limited to the implementation of the planned actions. The results can be observed in Figure 5.22.

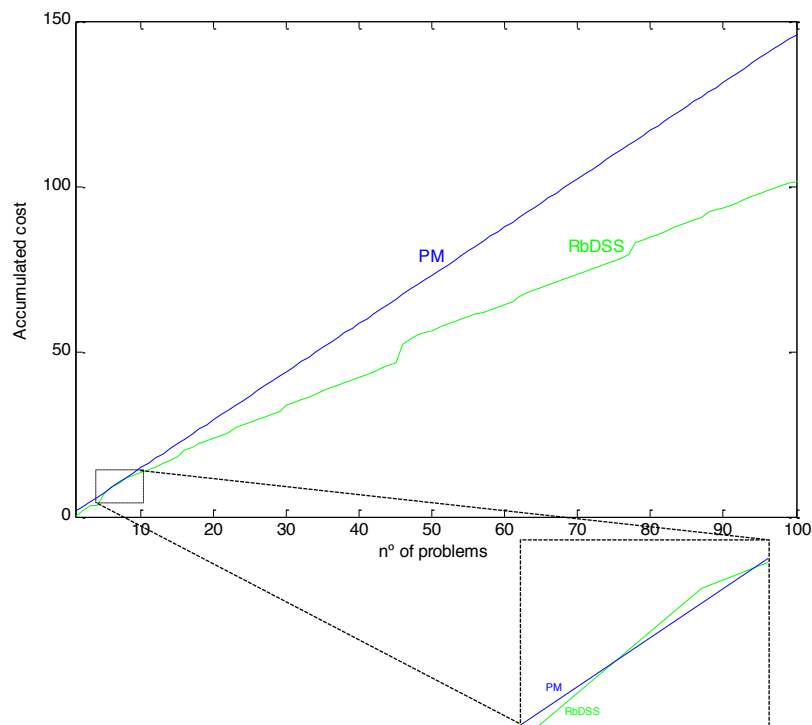


Figure 5.22. Accumulated costs for test case 3

The results obtained are interesting in terms of comparison of the approach proposed with the one here used as benchmark, especially due to the time during which the system is still learning and adapting to the knowledge being generated. During that time, the results obtained by means of preventive maintenance are better than the ones obtained through the risk-based decision support system. However, after the initial period, the results obtained through the proposed approach become more adequate and the cost minimization becomes also more efficient.



Still, this comparison was performed based on the assumption that the implementation of preventive measures would restrain problem occurrence. Unfortunately, this is not true in many cases, and for this reason, there was the need to introduce the stochastic problem generation component in the simulation. Thus, and keeping the maintenance plan specified above, the next simulations were performed assuming that:

- The interval specified within the action corresponds to achieving machine's half-life, which is represented by a threshold value;
- This is the minimum threshold allowed by the preventive maintenance approach, meaning that even in cases where the computation of threshold provides a lower threshold the approach imposes the implementation of a maintenance action when the specified value for machine's half-life is achieved.

The system was tested with two different minimum threshold values allowed:

- A more conservative approach in which threshold was set to 0.5;
- A less conservative approach in which threshold was set to 0.1.

The results obtained with threshold on 0.5 are presented in Figure 5.23 and the ones obtained with threshold on 0.1 can be seen in Figure 5.24.

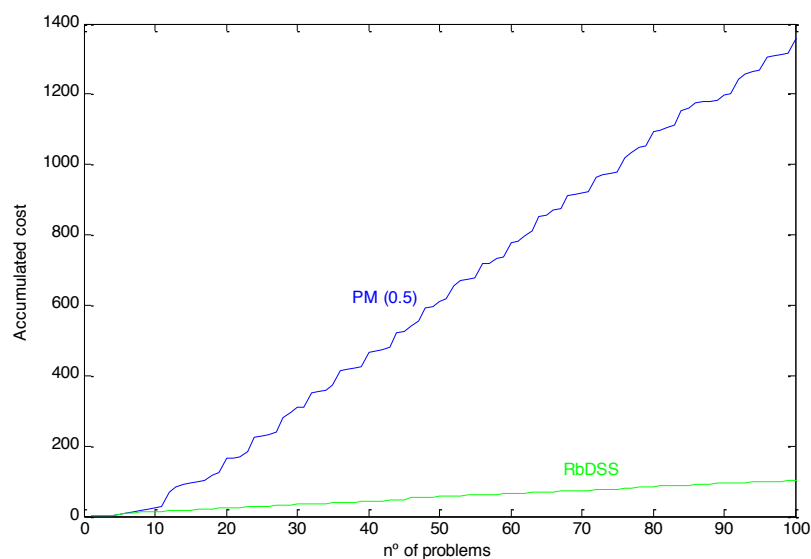


Figure 5.23. Accumulated costs for test case 3 with minimum threshold on 0.5

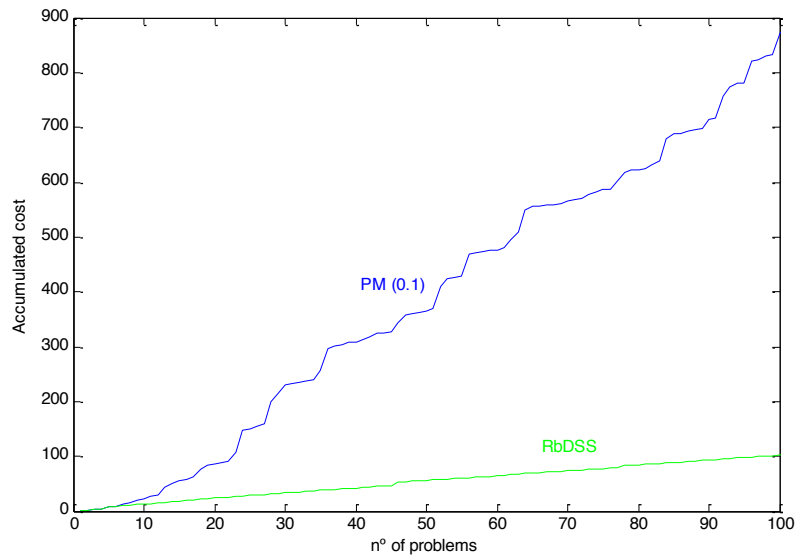


Figure 5.24. Accumulated costs for test case 2 with minimum threshold on 0.1

It is easy to conclude that both approaches result in aggregated costs higher than the ones achieved by the approach based on the use of the risk-based decision support system. Additionally it can also be observed that costs get lower with lowering the threshold which is a natural result since the threshold influences the time between problems when a specific action is executed.

### 5.3. Conclusions

The development of the simulator enabled the intensive testing of the approach proposed in order to understand its complete potential. Additionally, and since the approach assumed that the results of the algorithm would contribute for improving the life cycle management of the plant (specifically in terms of life cycle costs), the simulator enabled to test this impact in an appropriate time frame.

In terms of results, and despite its simplicity, the simulator proved the efficiency of the proposed approach and enabled to understand the impact that different decisions may have in the life cycle of the plant. The life cycle impact was assessed by means of the cost calculation, and subsequent minimization of the cost through the use of risk analysis strategy.

The tests performed both in terms of learning rate and accuracy, established the characteristics of the algorithm showing that it presents good level of accuracy and a fast learning rate, especially when the knowledge stored at the knowledge base increases. This way, the existence of knowledge at set-up phase is particularly important to improve the results by decreasing the number of new cases needed before the system starts delivering reliable results.

In terms of comparison with other maintenance methodologies the approach proposed in this thesis also presented advantages, especially when compared with methods such as preventive

maintenance and run-to-failure. The three test cases performed enabled to confirm the good behaviour of the algorithm in terms of cost minimisation along the life cycle of the plant.

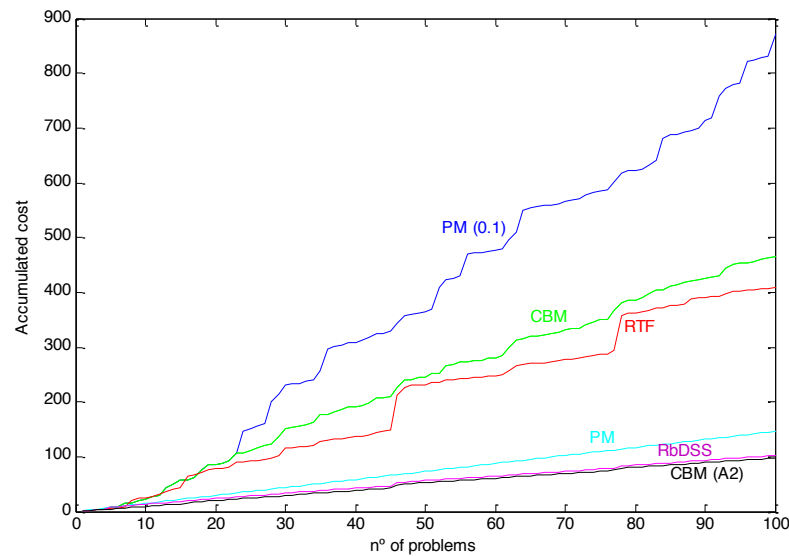


Figure 5.25. Comparison of the three test cases

To highlight the differences between the three test cases, and the results obtained for each of them, Figure 5.25 presents a general comparison, where it is possible to conclude that the results may be divided in three major groups:

1. Preventive maintenance with stochastic problem generation and threshold establishment is the one presenting higher costs;
2. Condition-based maintenance with the same maintenance plan used for preventive maintenance and Run-to-failure, present similar results in terms of accumulated costs, and these are around one half of the ones achieved by the previous result;
3. Preventive maintenance with no problems, Risk-based Decision Support System and Condition-based Maintenance with maintenance plan constituted only by  $A_2$  present the best results in terms of costs.

The tests were developed aiming to compare the approach proposed with a set of different maintenance methodologies, which are commonly used in industrial plants. The results showed that the continuous calculation of the accumulated costs, enable to support the decision for the best action to eliminate a specific problem.



## 6. *Validation in Industrial Applications*

---

Industrial validation and acceptance is a key aspect of any research work produced. In fact, if industry does not approve the results, the most probable scenario is to assume that those results will never be used. This way, the research work developed within this thesis was scoped through different Small and Medium Enterprises (SME) from different industrial sectors.

Additionally, and to facilitate acceptance, a prototype was also developed and used by the three involved in two different scenarios. This enabled to test the prototype in different situations.

This way, the validation methodology applied consisted in a three level approach:

- Level 1: development of the prototype designed as life cycle management system to demonstrate the **feasibility** of the concept;
- Level 2: feed the prototype with appropriate information to demonstrate **adequateness** to real industrial applications;
- Level 3: test the prototype and collect the results to demonstrate **applicability**.

In the next sections the entire methodology is described, including the three levels mentioned, starting by the description of the prototype, followed by the specific aspects that led the process of collecting adequate information, and finally the methodology that was applied to assess applicability. The chapter ends with the description of the two scenarios used for validation, as well as the results of using the prototype in each of them.

## 6.1. Validation methodology

### 6.1.1. Development of the life cycle management system prototype

#### 6.1.1.1. Concept and workflow

Developed in the context of InLife project (Marques & Neves-Silva, 2008), the Life Cycle Management System (LCMS) is constituted by several modules that are used in an aggregate mode to build the set of services provided.

The objective of InLife was to provide an approach for life cycle optimisation based on the computation of life cycle parameters using the information obtained on the status of the industrial plant. This information was collected both by using Ambient Intelligence (Aml) systems scattered along the plant as well as by the results of using the services provided by the system prototype.

The life cycle optimisation was achieved by the decision support capabilities of the system. In fact, more than just computing a set of statistics regarding the life cycle of the plant, LCMS intended to support the user in making more informed decisions at each step. This approach is thought of having high impact in the entire life cycle of the industrial plant. For a better understanding Figure 6.1 shows the rationale of the project, and Figure 6.2 presents the concept behind LCMS.

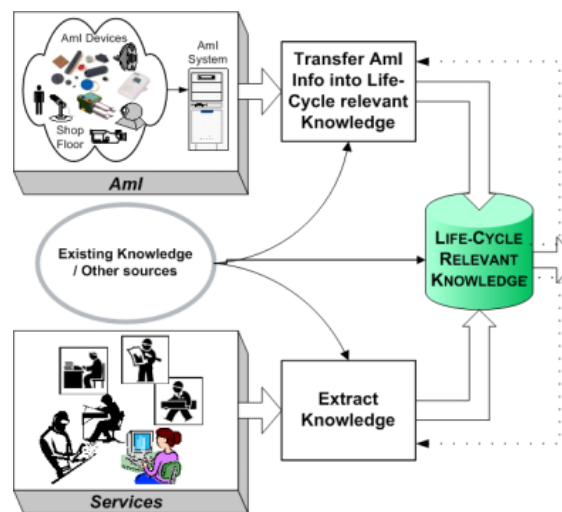


Figure 6.1. InLife rationale

The LCMS key modules were:

- **Aml processing module**, to process information from different Aml systems to support calculation of LCP and different services;
- **LCP monitoring and Product Lifecycle Management (PLM) module**, to process information from different sources to compute different LCP serving for life cycle management;
- **Service platform**, constituted by a subset of modules, to provide services to manage the industrial plant, namely:

- Installation and ramp-up support;
- Online remote maintenance and diagnostics;
- Prevention of hazardous situations;
- Condition-based maintenance; and
- Continuous improvement;
- **Knowledge Management Infrastructure**, serving as a set of tools for supporting other modules activities;
- **Common Repository**, as central repository gathering data for all modules.

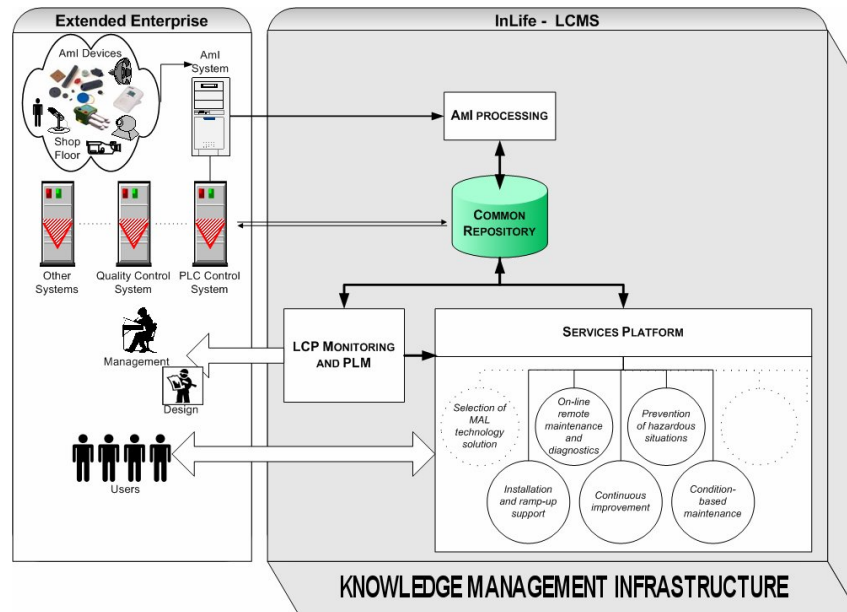


Figure 6.2. Life Cycle Management System concept

The system can be divided in two operating modes, automatic and on-demand, and this behaviour is highly dependent on the services that are going to be used. The automatic mode is responsible by all the part related with the daily operation of the plant, whereas the on-demand behaviour is only activated by some specific request of the user. Services “*Condition-based Maintenance*”, “*Online Remote Maintenance and Diagnostics*” and “*Prevention of Hazardous Situations*” constitute the automatic part of the system and their utilization is suggested when some abnormal situation is detected. On the other hand, services “*Installation and Ramp-up Support*” and “*Continuous Improvement*” are part of the on-demand style since they are only activated through a specific request from the user.

The automatic part of the system is orchestrated by a specific module, designated “*Risk Assessment*” module, which is responsible for the assessment of the situation criticality and, if possible, for the identification of the cause that is the root of the problem. Additional details on the complete architecture of the system can be found in Annex D. However, from a general point of view, the automatic part of the system, works as follows (see Figure 6.3):

- The LCPs related to plant operation are continuously monitored by the set of sensors (Aml) spread along the plant;
- The information coming from the Aml devices is used to calculate and measure the defined LCP;
- Each time a LCP value crosses any defined threshold, the system applies a risk assessment strategy to identify the kind of risk associated to that situation. Here, three options are possible:
  - If the situation is due to any previously defined critical LCP, then service “*Prevention of Hazardous Situations*” will be suggested to the user to help the user dealing with that potentially critical situation;
  - If the LCP associated to the situation is not critical, but there is not enough information available to automatically solve the problem, then the user is suggested to use “*Online remote Maintenance and Diagnostics*” which will guide the user through the process of finding the cause of the problem by means of case-based reasoning techniques;
  - If the LCP associated to the situation is not critical and there is sufficient information on that kind of situation, then the cause of problem is automatically identified and the user is suggested to use service “*Condition-based Maintenance*”, informing the user on how to proceed to eliminate the cause.

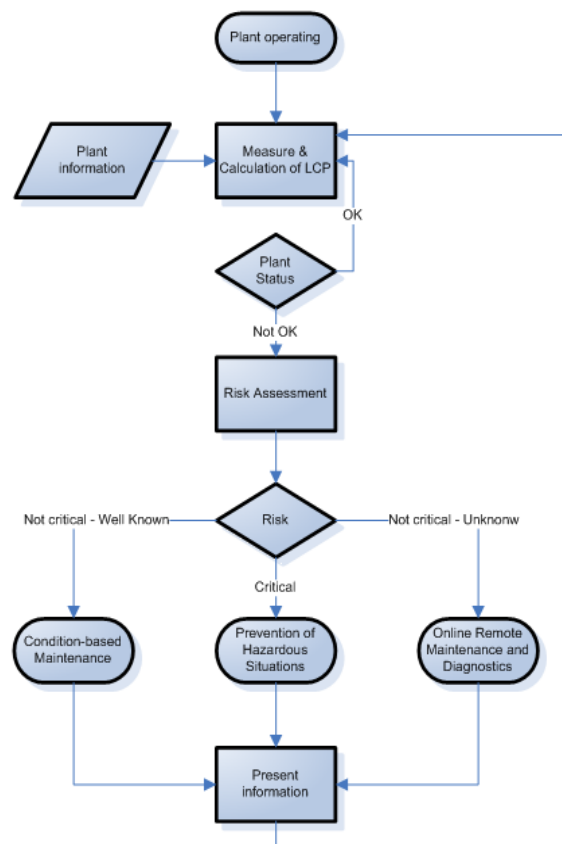


Figure 6.3. Flowchart for automatic behaviour of LCMS



The on-demand mode is controlled by the user, through the available user interface, by selecting the desired service. Both on demand services are based on searches performed on the history of the plant (previous problems, their causes, etc.) but the fundamental idea behind each of them constitutes the main difference. Thus, if the user intends to get support on ramp-up phase, the search is performed among the problems that occurred, during that phase, in similar installations. On the other hand, if the user is interested in finding more about recurrent problems in the plant, then the “*Continuous Improvement*” service will provide information which can be helpful to eliminate those problems in future plants.

For a better understanding of the kind of interface provided to the user, Figure 6.4 shows the aspect of the “*Condition-based Maintenance*” service home. Here the user has access to information regarding the calculated risk (in terms of possible costs) and the most probable cause for problems pending for elimination.

Condition Based Maintenance home

logged in as: admin

Main menu

User Administration

PHS

Online Remote Maintenance and Diagnostics

Installation and Ramp-Up Support

Condition-Based Maintenance

Continuous Improvement

Search

LCP

Check Maintenance Status

Check Maintenance Status

3 items found, displaying all items.

Problem Id	Name	Description	Status	Risk	Cause
20	Problem generated from rule(s): Valores normais de temperatura apos o compressor; Valores normais de baixa pressao;	++ Risk Test +- -Risk: 49EUR- -Risk: 45EUR- -Max Risk:232.10EUR	SOLVED	232.1 EUR	Vibracoes nas tubagens
21	Problem generated from rule(s): Valores normais de alta pressao;	++ Risk Test +- -Risk: 108EUR- -Risk: 43EUR- -Max Risk:132.0EUR	SOLVED	132 EUR	Fuga de agua
22	Problem generated from rule(s): Valores normais de temperatura apos o compressor; Valores normais de baixa pressao;	++ Risk Test +- -Risk: 49EUR- -Risk: 45EUR- -Max Risk:232.10EUR	SOLVED	232.1 EUR	Vibracoes nas tubagens

Export options: CSV | Excel | XML | PDF

Fields marked with an asterisk are required

Help

Figure 6.4. LCMS Condition-based Maintenance service home.

#### 6.1.1.2. Scope and simplified model

From all the functionality presented in the previous section, the author of this thesis was responsible by the development and implementation of the risk assessment strategy, which was based on some of the original contributions of this thesis. Additionally the author also developed the “*Condition-based Maintenance*” service that includes a scheduling module to help the user in scheduling the necessary procedures to eliminate the detected problems. Nevertheless, and since this particular module does not make part of the scope of this thesis it will not be detailed here.

As already motioned previously, condition-based maintenance is a technique to assess the general health of industrial equipment by regularly measuring and analyzing the data obtained during plant operation. With robust, regular and consistent data, the maintenance operator can be able to schedule maintenance using knowledge of the condition of the equipment rather than the

running hours or time alone. This way, condition-based maintenance uses real-time data to prioritize and optimize maintenance resources and allows determining the equipment's health to act only when maintenance is actually necessary.

Traditionally maintenance has followed the philosophy of either run-to-failure or planned maintenance at regular intervals. Each of these approaches has been found to be more expensive and time-consuming when compared to condition-based methods, under which the condition of a machine is monitored and maintenance is only undertaken if condition requires it. This method equally applies to industrial plant processes where the settings of some machines or components may need to be altered based on the monitored condition of the process.

Note that, apart from the direct functionality provided by “*Conditioned-based Maintenance*”, all the information gathered by the use of the service is stored so that it can be re-used by the other services, especially in what concerns the on-demand services.

In order to have the system installed and running in several industrial plants it was identified the need for developing a model, general enough that comprises all the relevant information about plant operation. Nevertheless, and despite the generic character of the model it was also required that it included several aspects of the operation. Thus, the result requires a good balance between simplification of some aspects and the complexity that still is associated to the production process.

Since the developed model to support the entire system includes some aspects that are not relevant for the work described in this dissertation, we will focus on the entities that are vital for the application of the proposed approach, i.e. “*Risk Assessment*” module and “*Condition-based Maintenance*” service. Figure 6.5 shows part of the developed model simply showing the entities that are directly involved in these two elements. Additional information on the entire system model can be found in Annex D.

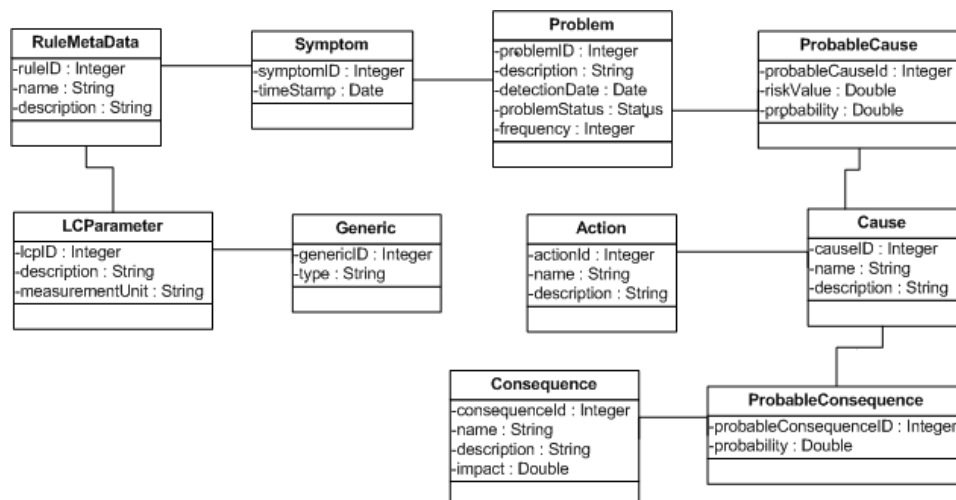


Figure 6.5. Simplified model

Although their similarity, the model here presented is not the same as the one presented in Figure 4.3 (which was the core of the development of the risk-based decision support system). This is

due to the difficulty in aggregating the impact of an action in the life cycle of the plant. Thus, in this model, outcomes are modelled based on the idea that one action eliminates one cause, and that cause may result in different consequences which were identified by experts and introduced at set-up phase.

The entity **Problem** is the central part of the entire model, since any deviation on the defined normal behaviour of the plant will be stored as an instance of this entity. This entity is connected to **Symptom**, representing the firing event for the risk assessment process, and to **ProbableCause**, which is a set of possible causes for the occurred situation (among a list of predefined causes). The **Symptom** entity makes the connection between the lower level part of the model, i.e. representing the equipment, and the upper one, i.e. representing the operators' knowledge.

In the lower level are visible the entities that represent the operating parts of the plant like **Generic** (representing machines involved in production, production processes or product parts) and **LCParameter** (representing the defined parameters that are seen as relevant for the plant). Additionally, **RuleMetaData** is used to represent the set of rules that will govern the symptom firing. In the upper level of the model, the aim is to store the information associated to causes and actions that are related to the problems. Additionally the information on which are the probable causes for a specific problem and the probable consequences for each cause are also stored.

All these entities have several attributes that are used to store the associated information. However, once again, and for the sake of simplicity, here are only referred the ones that are used in the context of this work. In what concerns **ProbableCause**, the attributed *probability* and *risk* are used. The first one will store the probability of that cause to be the correct one, whereas the second will store the value of the calculated risk during risk assessment. Note that a **ProbableCause** is unique but two probable causes can be associated to the same cause. The **ProbableConsequence** has an attribute to store the *probability* of a specific cause to induce a specific consequence. Finally, **Consequence** has an attribute to store the *impact* that consequence will have (if occurs).

### 6.1.2. Gathering and correlating information

In order to have the system working properly it was critical to make a correct assessment of the plant operation, both in terms of its normal status as in terms of failures/problems.

To acquire this kind of information it was required to combine the information coming from three sources:

- the **plant** itself (through the available monitoring instruments);
- the **historical information** available (e.g. failure history); and
- the **information collected through the users** of the plant (i.e. operators, maintenance staff, engineers and designers) which is of major importance since it reflects

knowledgeable opinions of people technically involved on the operation, as well as of those who have studied the failure mechanisms and how they affect plant performance.

The method used to combine the gathered information is based on a Bayesian approach, which is strongly oriented to give more consideration to high reliable data. This approach follows the philosophy that performance projections are strengthened by multiple technical viewpoints. (ASME, 2003).

Details about the most common sources of information are available in Table 6.1. These sources are somehow clear especially in what concerns the experts involved in the process. In fact the personnel directly involved in plant operation are indeed the best option for estimating plant problems.

Table 6.1. Common information sources

Source	Description
Monitoring systems	Provide information on the current status of the plant.
Stored operating information	Provide information on normal and abnormal plant operation, especially in terms of failure history.
Expert Opinions (Designers, operators, maintenance staff and engineers)	Based on interviews to people who work at the plant in a daily basis, together with specific plant information (provided mainly by designers and engineers) that can be used to locate potential problem sources.
Similar plant information	Provide information on similar plants. Only useful when similar operating conditions can be guaranteed.
Engineering analysis	Define potential failure operating modes and causes.

There are several methods that could be used to conduct the interviews so that the extracted information is as clear and free of misleading information as possible. The idea was that, at the end of the interviews, one is able to estimate a curve that reflects failure probability over time, as well as collected the necessary information about the causes for those failures and their related consequences. Additionally it was also important to have a clear understanding of the actions that need to be developed to eliminate the failure causes, i.e. to solve the problems.

The analysis of the information collected also enabled the estimation of the curves for causes probability, describing the probability of that cause to be the correct one (based on the number of times that it was identified has being the solution for the problem), and consequences probability, describing the probability of that consequence to occur (based on the number of times that it has occurred).

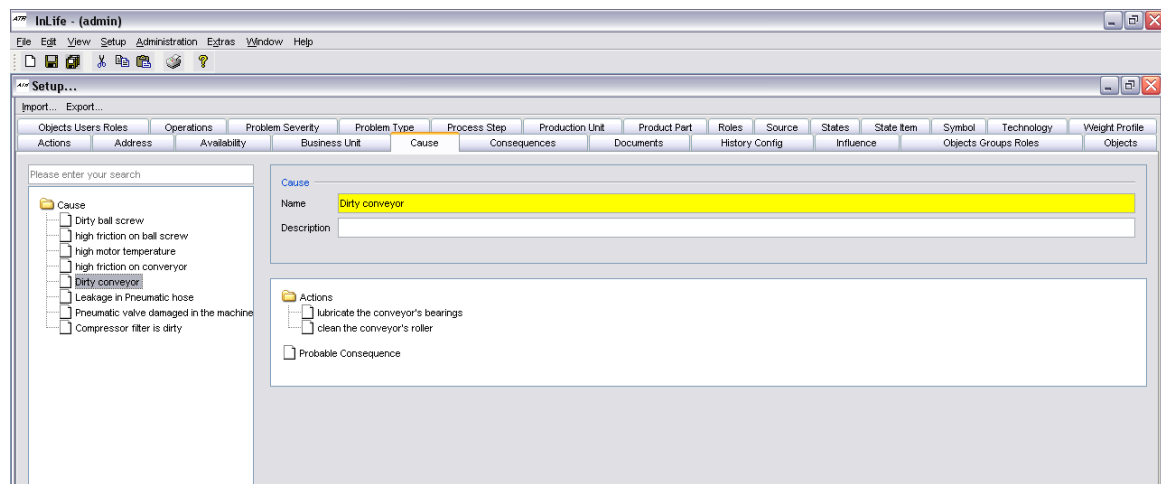


Figure 6.6. InLife set-up tool

The aim was then to collect sufficient information that enabled to build a set of **Cases** (i.e. problems which status is “solved”) that reflected the history of the plant, representing the “*a priori*” knowledge. After having the available information collected the system needed to be fed with it.

To facilitate this task, a set-up module was developed which enabled the introduction of the collected information into the system knowledge base. Figure 6.6 shows the front-end of the developed set-up tool.

### 6.1.3. Assessment methodology

The testing and verification of the developed prototype was performed in real industrial operating environments and used for measurement of the business benefits achieved by the implementation of the system. Therefore, in order to perform this validation two possibilities were previously analyzed:

- a quantitative measurement method, that could lead to the generation of statistics of the use of the different functionalities and scenarios.
- a qualitative method where the users of the system would describe their experience regarding the most useful functionality, how it is used, etc.

In the first case, the data obtained would be used to evaluate, by comparison, the current and the previous situations. However, the quantitative method has shown to be very difficult to carry out. In fact it would be very hard to compare directly the previous ways of doing things with the new ways that are available by the use of the prototype, since the previous ways of storing and reusing knowledge lacked consistency and could lead to poor conclusions.

In the second case, users would describe the advantages of using the system and give an estimation of the gains obtained. In order to do this, it was proposed to perform the evaluation of the prototype by the users, and to organize several workshops to record that material and to assess the expected benefits by comparing to the experience of users performing the same tasks without the system.

This second method was considered to be more convenient for validation purposes. According to this, the method includes the description of the following topics:

- Definition of **what we are measuring**;
- Description of **how we are measuring** – a qualitative measurement, describing the experience from the involved users;
- Description of the **expected benefits**, describing the advantages and estimated gains.

In order to describe the application of this method in a common way by the different scenarios, the following structure was used:

- **Measurement of success indicators:** the specific objectives that had been defined for each scenario;
- **Measurement process:** the measures were obtained and registered by testing the prototype during typical situations of operation. In the different scenarios several approaches were carried out according to the specific scenarios and working environments
- **Expected Benefits:** obtained from the qualitative analysis of the expected advantages and improvements achieved from the use of the prototype functionality compared to previous processes.

According to the selected validation methodology the different companies have performed the testing during typical situations of operation. Each has performed the assessment according to the specific scenarios and working environments.

## **6.2. Scenario I: condition-based maintenance of a manufacturing plant**

### **6.2.1. Description and objectives**

The first validation scenario is composed by two German SMEs. The first one, founded in 1982, plans and designs assembly lines for the automotive industry world-wide. Their focus is on chassis component assembly, putting together engines, cylinder heads, axles, transmissions and steering systems with the highest level of accuracy. The range of supply comprises individual solutions as well as complex and fully automatic production lines. In addition to assembly systems, which excel in flexibility, precision and reliability, services such as process planning, simultaneous engineering, customer training and after-sales service form an important part of the product range.

The second company was founded in 1992 and, within a very short time, evolved into one of the best-known service providers in the mechanical engineering sector with more than 100 employees. Their competence makes them an adequate partner in technology and service for the automotive industry, machine tool manufacturers and many other industries. Their main focus is on reconstruction and process optimization, new installation of manufacturing and assembly lines and provision of maintenance services.

The two companies have been working together for many years following a business model where the second one provides maintenance services to the manufacturing lines produced by the first one. Along this cooperation, they identified some problems that could be solved, or at least mitigated, with the use of LCMS, namely:

- High maintenance frequency;
- problem causes are sometimes not correctly identified leading to difficulties in efficient and time effective maintenance as well as in components availability; and
- availability of data about the plant status to improve life cycle management.

Thus, as main objective the two companies highlight the need for a tool that supports the maintenance service together with the optimisation of the life cycle of the plant. To achieve these objectives LCMS provides the following functionality:

- Measurement and calculation of LCP enabling to assess plant status;
- Support to identification of reasons of problems which include information on previous similar problems; and
- Collection of knowledge from the maintenance staff.

## 6.2.2. Testing scenario

### 6.2.2.1. Data model

To enable proper behaviour from the system it is necessary to introduce appropriate information regarding the manufacturing plant that is going to be used for testing. In this case the manufacturing plant is composed by three machines: A10, A20 and A30 which were modelled and introduced in the knowledge base.

The modelling task was mainly focused in aspects such as the important life cycle parameters that are associated to the status of the machines. Additionally the modelling task also included the identification of the main causes for common problems, the actions to deal with those causes and the consequences that might occur if the corrective actions are not applied.

Therefore, firstly the information related with the manufacturing plant was introduced. For that the *Generic* entity was used to define its relevant parts (see Table 6.2).

Table 6.2. Data model for the manufacturing plant of scenario I

Production units		
A10	Assembly Tool	Power screwdriver
		Feeding unit
	Lifting unit	Mounting unit
		Gear drive

	Tilting unit	Hydraulic pump
		Motor
		Motor fuse link
	Sliding unit	Pneumatic unit
	Safety system	Control switched
<b>A20</b>	Suspension unit	Spring Mounting
		Suspension Base
	Lifting unit	Mounting unit
		Gear drive
	Tilting unit	Mounting unit
		Hydraulic pump
	Sliding unit	Pneumatic unit
	Safety system	Control switched
<b>A30</b>	Linear unit	Laser scanner
		Gear drive
	Drive unit	Mounting unit
		Gear drive
		Motor
	Sliding unit	Pneumatic unit
		Compressor
	Tilting unit	Mounting unit
		Hydraulic pump
		Motor
	Safety system	Control switches
<b>Conveyor</b>	Holding system	Mounting unit
	Drive unit	Gear drive
		Motor



Then, for each machine two life cycle parameters were defined: the *takt time*, representing the pace for each machine of the manufacturing plant, and the *air flow* needed for each machine to work properly (see Table 6.3).

Table 6.3. Data model for life cycle parameters of scenario I

Name	Description	Measured in (Units):
TT A10	Takt time for machine A10	sec
TT A20	Takt time for machine A20	sec
TT A30	Takt time for machine A30	sec
AF A10	The air flow rate needed for machine A10	litre/hour
AF A20	The air flow rate needed for machine A20	litre/hour
Af A30	The air flow rate needed for machine A30	litre/hour

Next step consisted in the definition of the list of causes and associated actions to be introduced in the model (see Table 6.4).

Table 6.4. List of causes and associated actions for scenario I

Cause	Actions
Dirty conveyor	Lubricate the conveyor's bearings
	Clean the conveyor's roller
Leakage in Pneumatic hose	Change the hose
Pneumatic valve damaged in the machine	Change the pneumatic valve

Finally the list of consequences for each cause was also introduced (see Table 6.5).

Table 6.5. List of causes and associated consequences for scenario I

Cause	Consequences
Dirty conveyor	Increased takt time
	Damages in the product
Leakage in Pneumatic hose	Damages in the product
Pneumatic valve damaged in the machine	Damages in the machine
	Damages in the product

#### 6.2.2.2. Runtime phase

Each time the rule associated to any of the defined life cycle parameters is violated it represents a problematic situation. The rule must therefore include the selected parameter compared with the threshold. If the condition specified in the rule becomes true this will generate a symptom.

This scenario focuses on the application service *Condition-based Maintenance* (CBM). The scenario describes a situation where an abnormal value occurs, generating a problem that, due to its characteristics, will have a calculated risk value and a set of corrective actions to eliminate the problem cause. Table 6.6 details the steps of this scenario.

The scenario here presented was based on the analysis of '*takt time*'. The risk assessment's input consists of a symptom which has been associated to a set of solved problems, with well known causes, as well as a set of corrective actions defined. Thus, the service to be selected is the *Condition-based Maintenance* service. The generated **Problem** contains the occurred symptom and associates a calculated risk value to it.

Table 6.6. Scenario for testing Condition-based Maintenance

Step no.	Title	Description
1	Fire Monitoring Rule	Monitoring system detects an abnormal value for the "takt time" A10 and fires the appropriate rule. This rule has information on: <ul style="list-style-type: none"><li>the condition that was violated</li><li>criticality of the occurrence - for <i>Condition-based Maintenance</i> cannot be critical</li></ul>
2	Generate symptom	Based on the rule fired the system generated a symptom which comprises information on: <ul style="list-style-type: none"><li>time instant of the occurrence</li><li>involved generic</li><li>fired rule</li></ul>
3	Generate problem	Based on symptom information and on the information on previous (similar) problems already stored in the knowledge base the system creates a new problem.  For <i>Condition-based Maintenance</i> the problem is marked as <i>SOLVED</i> since the system is able of finding, among previous problems, the solution for this new occurrence.

Step no.	Title	Description
4	Available problem Information	The information available on the created problem is presented including the specific cause which in this case is: <ul style="list-style-type: none"> <li>• dirty conveyor</li> </ul>
6	Involved generics	The list of generics that are involved in the problem which, is presented.
7	Actions needed	The list of the actions needed to eliminate the problem is presented. In this scenario the action involved are: <ul style="list-style-type: none"> <li>• Lubricate the conveyor's bearings</li> <li>• Clean the conveyor's roller</li> </ul>

### 6.2.3. Test results

The available prototype was tested by both companies, using as benchmark the available information related with previous problems and the way they were treated and the new operating mode enabled by the use of LCMS. The prototype was assessed based on the requirements previously defined.

The analysis of cost/benefit of the infrastructure was based on the use of the LCMS to optimize the life cycle impact during the operational phase of the manufacturing plant. In the case of Condition-based Maintenance it is expected that it leads to optimized maintenance activities, for the monitored manufacturing plant, based on its actual state.

Generally, and from companies' perspective, benefits can be expected in terms of reducing their service costs. From their customer point of view benefits will include reduced maintenance and production cost and higher performance. More specifically the following benefits are expected to be achieved:

1. Improved collecting of repair and maintenance data and structured knowledge;
2. Cut the maintenance service and travel costs by preventing break-down of components and having the right spare part at the right time;
3. Rapid collection of information from the manufacturing plant brings a quicker reaction of maintenance service to the problem occurred;
4. Immediate reaction of maintenance service to the problem occurred and reduced effort/time needed for problem solving;
5. A decrease in the number of required maintenance activities leading to a reduction of downtimes of the manufacturing plant and, therefore, to higher productivity on the side of the customer.

In order to assess these expected benefits, a basic scenario for the usage and testing of *Condition-based Maintenance* within the LCMS prototype has been defined.

The more classical approaches to maintenance, e.g. performing maintenance actions only when machine failures occur or time to time maintenance based on the manufacturer defined timetables, tend to cause undesired results both ending in interruptions of production with consequent increase in time and costs. Condition-based Maintenance allows preventive maintenance before production is stopped due to machine failures.

Note that some expected benefits could not be exactly measured since they are not short term objectives, which stands as a difficulty in the measurement process. Nonetheless, it was possible to find some correspondence between the above mentioned benefits and a set of related benefits which measure was possible:

- Problems registered in a structured way, appropriate for re-use: related with expected benefit 1;
- Flexibility of maintenance intervals: related with expected benefit 5;
- Percentage of spare parts consumption: related with expected benefit 2;
- Manufacturing plant availability and performance: related with expected benefit 3, 4, and 5.

The measurement process was based on internal testing since prototype was available, involving experts from both companies. The current measures are compared with the above listed objectives and benefits have been assessed within appropriate testing scenarios by experienced experts. For this scenario, the following measurement indicators were defined:

- number of solved problems registered,
- number of eliminated problems registered,
- time spent by each user operating the LCMS,
- efforts spent in the maintenance actions needed to eliminate each problem.

The expected benefits have been assessed based on experience and knowledge of experts on the operating mode of the manufacturing plant.

Table 6.7. Testing results for scenario I

Benefit	Measure before LCMS	Measure after using LCMS	Comment
Problems registered in a structured way, appropriate for re-use.	c.a. 50% of all problems	c.a. 100% of all problems	Measured from previous registers and the current stored problems
Flexibility of maintenance intervals	low	Medium	-
Percentage of spare parts consumption	c.a. 60%	c.a. 40%	Extrapolated from events of problem <i>SOLVED</i> and problem <i>ELIMINATED</i> .
Manufacturing plant availability and performance	Medium	High	Extrapolated from time between problem <i>SOLVED</i> and problem <i>ELIMINATED</i> .

When analysing Table 6.7 results, note that events such as problems *SOLVED* and problem *ELIMINATED* stand, respectively, for the problems which cause was identified and for the ones which were already treated with the corrective actions.

The results achieved showed the efficiency of the system when correctly modelled and used. It is expected that the results will be refined along usage time but current results show a higher level of structured information about problems in the manufacturing plant which has a strong impact in the diagnosis process and thus influence maintenance procedures.

The information gained is fundamental for establishing a continuous improvement strategy where knowledge can be re-used for the design of new plants.

### **6.3. Scenario II: condition-based maintenance from the product perspective**

#### **6.3.1. Description and objectives**

This validation scenario is constituted by an air conditioning company located in Portugal. The company has an average number of employees of 400 people. They are specialized in manufacturing and assembling of compression and sorption systems including: air-air machines, air-water machines and water-air acclimatization. The productive process considers the manufacturing of all the metallic components, copper kits (refrigeration), electric boards and the assembly of half-finished components

The company, with large experience in the area, concluded that the behaviour of their units, specially the large scale acclimatization ones, has a strong impact on their daily operation. Thus the following problems were identified as aspects where LCMS could help:

- High maintenance frequency;
- Incorrect maintenance due to misidentification of problems causes;
- Lack of information on air conditioning unit status.

The overall objective of the company is then to improve the life cycle of their large scale acclimatization units by means of a more agile maintenance process together with the collection of knowledge that will be used for designing future units.

To achieve these objective LCMS provides the following functionality:

- Measurement and calculation of LCP enabling to assess units performance;
- Automatic cause identification, based on previous similar problems; and
- Collection of knowledge about current units' condition.

## 6.3.2. Testing scenario

### 6.3.2.1. Data model

Once again and to enable proper behaviour from the system it is necessary to introduce the appropriate information. Nevertheless, in this case, this information regards not the manufacturing plant itself but the final product which are the acclimatization units.

The modelling task was, as previously, developed focusing in the aspects of important life cycle parameters that are associated to the status of the machines, main causes for common problems, actions to deal with those causes and consequences that might occur if the corrective actions are not applied.

The focus of this scenario will be the GAAS-200 acclimatization units, belonging to the production units family GAAS, which has several models as evidenced in Table 6.8:

Table 6.8. Data model for GAAS acclimatization units' family

Acclimatization Unit	
GAAS Chiller	GAAS 200
	GAAS 260
	GAAS 300
	GAAS 350
	GAAS 400
	GAAS 440
	GAAS 500
	GAAS 600

As in the previous scenario the first the information to be introduced was the one regarding the *Generic* entity. The main components of these units are the refrigerating circuit, the steel plate, the axial fan, the hydraulic system and the electrical switchboard and control system. The refrigerating circuit includes some of the main components of the acclimatization unit, like the evaporator, the condenser or the compressor. Table 6.9 presents the elements that constitute each GAAS production unit.

Table 6.9. Data model for GAAS acclimatization unit

Acclimatization Unit		
Refrigerating circuit	Compressor	Thermostat
		Oil display
		Oil pressure pump
		Motor
		Differential pressure switch
		Anti-vibrating support
		Liquid separator
	Evaporator	
	Condenser	
	Piping	
	Thermostatic expansion valve	
	Drying filter	
	Liquid display	
	High-pressure plug	
	Low-pressure plug	
	High-pressure switch	
	Low-pressure switch	
	Liquid deposit	
	Charging valve	
Electrical switchboard and control system	Control circuit	
	Servos	
	Switches	
	Connectors	
	Relays	
Axial fan	Fan blade	
	Sheave	

Acclimatization Unit		
	Damper	
	Ventilator motor	
	Belt	
Hydraulic system	Water pump	
Steel plate		

Depending on the available sensors a set of life cycle parameters was defined enabling to assess the overall status of the production unit. Figure 6.7 presents the functional scheme of the units, indicating the values read by the monitoring system that are available as life cycle parameters in the data model.

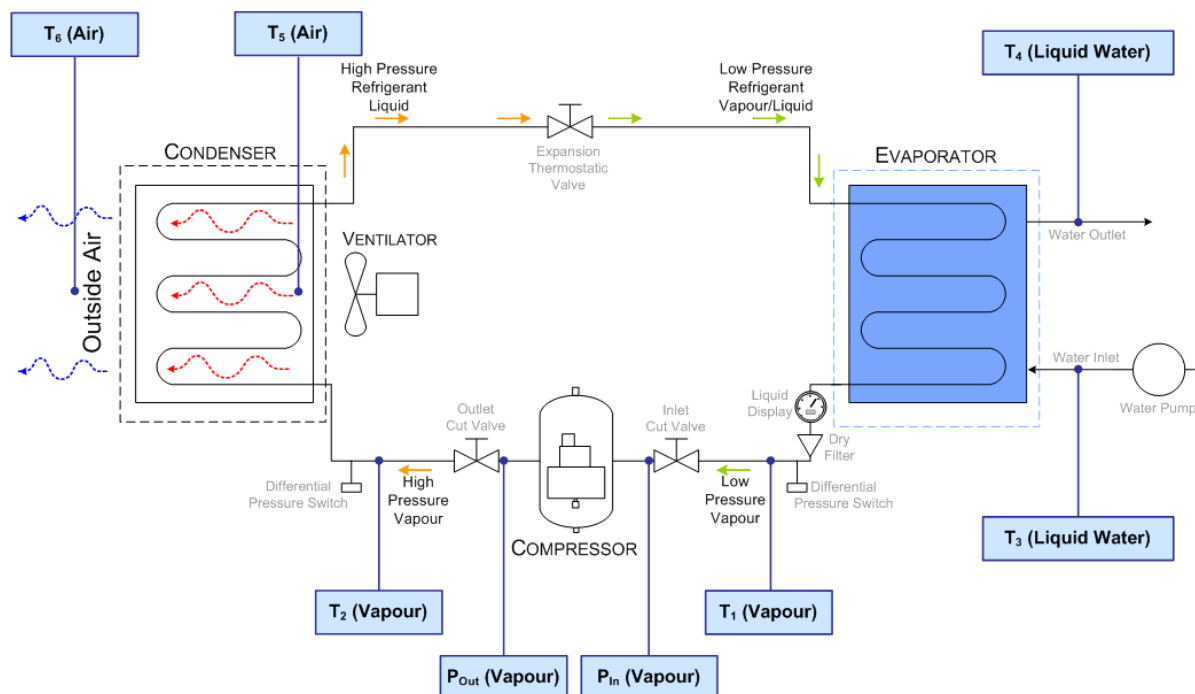


Figure 6.7. Functional scheme and life cycle parameters in scenario II data model

Table 6.10 describes the defined life cycle parameters for scenario II. Pressure values  $P_{in}$  and  $P_{out}$  are used to assess the overall performance of the production unit and the condition of the compressor. Temperatures  $T_1$ ,  $T_2$  reflect the behaviour of the compression process. They help prevent problems regarding excessive cooling and pressure. Temperatures  $T_3$  and  $T_4$  are used to analyze the water temperature before and after being refrigerated, thereby having parameters to evaluate the overall efficiency of the work being done by the unit. Temperatures  $T_5$  and  $T_6$  are used to compute the temperature difference of the air before and after the condenser, indicating the capacity of the system to refrigerate the water. Furthermore,  $T_6$  indicates eventual problems in the condenser.



Table 6.10. Data model for life cycle parameters of scenario II

Name	Description	Measured in (Units):
Pin	Gas pressure before compression	Bar
Pout	Gas pressure after compression	Bar
T1	Gas output temperature after evaporation	°C
T2	Gas output temperature after compression	°C
T3	Water temperature before evaporator	°C
T4	Water temperature after evaporator	°C
T5	Outside air temperature before condenser	°C
T6	Refrigerated air temperature after condenser	°C
$\Delta T = T6 - T5$	Temperature difference of the air at condenser	°C

Finally, after the identification of the life cycle parameters, a list of causes associated with corrective actions to eliminate them was elaborated (see Table 6.11) and a list of possible consequences associated to each cause was also introduced (see Table 6.12).

Table 6.11. List of causes and associated actions for scenario II

Cause	Action
Ventilator not working	Verify and/or replace the ventilator motor
Blocked expansion valve	Clean and/or replace the expansion valve
Excess of refrigerator	Eliminate the excess of refrigerator from the system
Drying filter blocked	Replace filter

Table 6.12. List of causes and associated consequences for scenario II

Cause	Consequence
Ventilator not working	Expansion valve malfunction
	Compressor breakdown
Blocked expansion valve	Condenser clogs
	Expansion valve blocks
	Expansion valve malfunction

	Compressor breakdown
Excess of refrigerator	Expansion valve blocks
	Expansion valve malfunction
	Compressor breakdown
Drying filter blocked	Compressor breakdown

#### 6.3.2.2. Runtime phase

This scenario focuses on the application service Condition-based Maintenance (CBM). The scenario describes a situation where an abnormal value occurs, generating a problem that, due to its characteristics, will have a calculated risk value and a set of corrective actions to eliminate the problem cause. Table 6.13 details the steps of this scenario.

The test here presented was developed based on '*differential pressure*'. The risk assessment's input consists of a symptom which has been associated to a set of solved problems, with well known causes, as well as a set of corrective actions defined. Thus, the service to be selected is the *Condition-based Maintenance* service. The generated Problem contains the occurred symptom and associates a calculated risk value to it.

Table 6.13. Scenario for testing Condition-based Maintenance

Step no.	Title	Description
1	Fire Monitoring Rule	Monitoring system detects an abnormal value for the "Differential pressure" of GAAS-400 and fires the appropriate rule. This rule has information on: <ul style="list-style-type: none"> <li>the condition that was violated</li> <li>criticality of the occurrence - for <i>Condition-based Maintenance</i> cannot be critical</li> </ul>
2	Generate symptom	Based on the rule fired the system generated a symptom which comprises information on: <ul style="list-style-type: none"> <li>time instant of the occurrence</li> <li>involved acclimatization unit</li> <li>fired rule</li> </ul>

Step no.	Title	Description
3	Generate problem	Based on symptom information and on the information on previous (similar) problems already stored in the knowledge base the system creates a new problem.  For <i>Condition-based Maintenance</i> the problem is marked as <i>SOLVED</i> since the system is able of finding, among previous problems, the solution for this new occurrence.
4	Available problem Information	The information available on the selected problem including the specific cause, is presented: <ul style="list-style-type: none"> <li>• inappropriate status of the drying filter</li> </ul>
5	Involved generics	List of the generics that are involved in the selected problem which, in this scenario, is the GAAS-400.
6	Select Production unit	Selection of the presented acclimatization unit as the one where a maintenance procedure is needed.
7	Maintenance Actions needed	List of the actions needed to eliminate the problem. In this scenario the action involved is: <ul style="list-style-type: none"> <li>• Replace filter</li> </ul>

### 6.3.3. Test results

Considering the expected benefits of the introduction of LCMS in the daily operation of the company, a list of measurement indicators was elaborated:

- Reduction in costs/efforts in MAL installation, ramp-up and guarantee phase achieved by better diagnostics of MAL using Aml information;
- Reduction in LCC (for service-life after guarantee period) achieved by better insight in the LCC based on Aml (and other) information;
- Increase in customers solving problems themselves (using on-line assistance);
- Reduction in number of equipment problems;
- Reduction in number of re-occurrence of same problem (re-visits);
- Reduction in equipment downtime;
- Reduction in hazard risks (accident rate);
- Increase in equipment/process improvement actions.

The objectives here stated are not short term objectives, thus they are extremely difficult to measure in a narrow time range since they are related with the behaviour of life cycle parameters, contributing to the complexity in the measurement process. To overcome this difficulty a list of

clearer measurement indicators was defined, focusing on the concrete use of the system and on what the company expects to achieve with it, namely:

1. Elimination of the time spent with the solution of well known problems and reduce the time for solving those which cause is still not known;
2. Reduction of the average time and efforts spent in maintenance actions;
3. Reduction on the maintenance cost.

A correspondence was established between the above mentioned benefits and a set of related measurable benefits:

- Problems registered in a structured way, appropriate for re-use: related with expected benefit 1;
- Average time to solve a problem (i.e. identify the cause of the problem): related with expected benefit 1
- Percentage of problems that require diagnostics travel: related with expected benefit 1;
- Number of customer complaints due to maintenance: related with expected benefit 3;
- Percentage of problems handled by customers: related with expected benefit 3;
- Number of maintenance actions scheduled and performed: related with expected benefit 2;
- Time spent in establishing a maintenance plan: related with expected benefit 2.

In order to draw conclusions on real benefits measured, the company identified the following measurable variables:

- number of solved problems registered,
- number of unsolved problems registered,
- time to solve an unsolved problem,
- time spent by each user operating the LCMS,
- efforts spent in the maintenance actions needed to eliminate each problem.

Since one of the main principles of the LCMS is to store any interaction with the system, to enable future statistical treatment of the information and tracking, some events are automatically stored in the knowledge base. This procedure also enabled to use the stored information as the measures here presented. Since occurred problems were normally registered in paper-based forms this information was used as benchmark, comparing the older information with the one acquired using the LCMS.

The use of the LCMS was seen has a major contribution in aspects such as structuring, collecting and searching for the adequate information. However, some of the measures were extrapolated, in terms of time range, enabling to draw some conclusions at a longer term.

Table 6.14. Testing results for scenario II

Benefit	Measure before LCMS	Measure after using LCMS	Comment
Problems registered in a structured way, appropriate for re-use.	c.a. 1 per month	c.a. 100% of all problems	
Average time to solve a problem (i.e. identify the cause of the problem).	c.a. 2 days	Up to c.a. 4 hours	Measured from LCMS events of problem created and solved.
Percentage of problems that require diagnostics travel.	c.a. 85%	c.a. 25%	Necessary travel was documented as part of problem.
Number of customer complaints due to maintenance.	c.a. 10 per year	-	Measurement not yet possible because of short period of time.
Percentage of problems handled by customers.	c.a. 0%	c.a. 25%	Concluded from events of problem scheduled and eliminated.
Number of maintenance actions scheduled and performed	c.a. 70%	c.a. 100%	Concluded from the LCMS events of maintenance action scheduled, in progress and finished
Time spent in establishing a maintenance plan.	1-2 hours	5-10 min	Concluded from the LCMS events of problem solved and problem scheduled

As stated before, these measured benefits do not account some highly significant intangible benefits. One of these benefits is the fact that the company achieved a structured knowledge base of problems commonly registered in acclimatization units, including their diagnosis and respective maintenance actions. This knowledge is now available for any employee, independently of the level of expertise in a specific product line or customer. This knowledge source is also quite valuable in the integration and learning process of new employees in the company. In addition, the company has gained vital information on how the units are really used by customers, and how they respond to different situations. This information is vital for the redesign process in a continuous improvement strategy. Additionally, the system also allowed receiving first-hand information, improving the relation with clients by providing an efficient and reliable maintenance support.

#### 6.4. Conclusions

The validation methodology defined at the beginning of this chapter proved to be efficient in demonstrating the proposed concept. In fact the development of the prototype considering several aspects of real industrial environments together with the tests performed *in situ*, contributed for the overall success of the approach proposed.

This success can be evaluated by the results of the tests performed especially in what concerns the objectives of the companies involved in the tests. In fact, and even that some benefits could

not be exactly measured due to their long term characteristic the results achieved demonstrated the efficiency of the system when correctly modelled and used. Additionally, and due to the continuous adaptation capacity of the system, it is expected that the results will be refined along usage time.

Unfortunately these long term objectives are extremely hard to measure since they are tightly connected with the occurrence of problems in the plants, i.e. depends on the existence of undesired events. Since most of the work on industrial plants is related with problem avoidance, it becomes quite difficult to gather enough information to extract long term conclusions in short term testing.

Nonetheless, current results show already an increase on the level of structured information which has a significant impact in the diagnosis process and thus influence maintenance procedures. The information gained is fundamental for establishing a continuous improvement strategy where knowledge can be re-used for the design of new plants.

Also companies achieved a structured knowledge base of problems commonly registered including their diagnosis and respective maintenance actions. This knowledge is now available for any employee, independently of the level of expertise in a specific product line or customer contributing, for example, to decrease the adaptation time of a new employee.

## *7. Conclusions and future work*

---

The objective of this thesis is to provide insight about the extent of impact that short term decision may have on the long term behaviour of an industrial plant. As stated in the research question defined at the beginning the notion that those decisions were not free of impact, guided the research work developed.

The need for answering to this question is not independent from the observations made along the years in industrial environments, especially in what regards every day procedures, which are sometimes performed without a complete understanding of their effective impact in the overall performance. The development of methods and tools to support decision makers should then be based in a holistic approach since without a comprehensive strategy it is very hard to extract conclusions about what could be improved and how. Accepting that all decisions have a consequence, and each course of action will lead to a different outcome is then the primary principle of this thesis.

One of the main problems in many industrial plants is maintenance management. Despite the developments being made, addressing new maintenance methodologies and sustainable maintenance procedures, maintenance still represents an important part of the costs. This thesis provided a survey on existing maintenance methodologies, analysing them in terms of characteristics and effects in the daily operation of industrial plants. Additionally, it recognises that besides normal maintenance procedures, there are some situations where inappropriate maintenance actions are implemented leading to extra costs which will be reflected in the overall production costs. These inappropriate actions are normally due to incorrect diagnostics and to lack of additional information for selecting the best procedure. Thus, selecting the best procedure is highly dependent on the existing knowledge about plant specificities.

Industrial environments are, due to their complex nature, difficult to represent in a straightforward mathematical expression. To deal with this problem this thesis proposes a model to structure plant information which derives from established correlations about plant behaviour, aiming to capture the essential nature of the plant. The model is used to store operational, as well as

maintenance, information enabling its reuse. Despite its simplicity, the proposed model was elaborated in order to highlight the relations between the different elements identified as crucial for the development of the methodology. The developed structure enables an easy understanding of the sequence of events that lead to a problem and the actions that may be applied in order to eliminate that problem.

Another problem that decision makers are well aware of is that decisions involve risks. Even in cases where this risk is not quantifiable, a better option may exist. This idea guides the problematic of decision making which normally involves a compromise between risk and objectives. Finding the best compromise is the goal of developing risk analysis strategies. Therefore, this thesis proposes a probabilistic risk analysis strategy to deal with the possible outcomes that different courses of action may have. The proposed strategy is based on the identified relations that were used to build the plant model, which considers that the detection of an event (the symptom) is related with a specific set of causes which will result in different outcomes depending on the actions performed to deal with each cause.

Together with the problematic of risk, decision also faces another difficulty, especially in some application domains where industrial plants are included. This difficulty is related with the huge amount of information that the industrial plant can produce through the available instrumentation. And if it is true that the existence of information is positive, it will have no effect if the extraction of knowledge from that information is not performed or if it is performed in an inappropriate way. The existence of decision support systems contribute for the organisation and correlation of the available information, especially in cases where this information is expected to increase over time. Moreover, the addition of intelligence notions to decision support systems increases their flexibility enabling the adaptation of their results to continuing changing conditions. The reutilisation of the knowledge stored in the model, enabling the system to learn from past experience and adapt its current behaviour accordingly to what was previously done, constitutes the base of the developed methodology. Thus, this thesis proposes the development of an intelligent decision support methodology which, by identification of similarities with previous problems enables the possibility of cause identification based on probabilities. Additionally, the methodology performs a time-based case aggregation, where the identified cause is used as second discriminator, to conclude about the possible outcomes of the different courses of action.

This two step approach is implemented using a decision tree model where risk calculation determines the selection of the path, i.e. the action, to be suggested. The similarity measure is performed each time a new problem occurs and results in the selection of a reduced number of problems to be used in the decision process. On the other hand, case aggregation is performed once the system has enough stored cases and is updated each time a new problem is considered solved. These two methods contribute for the implementation of the required intelligence notion of the developed decision support system.



Establishing that all decisions have a consequence is stating that each decision will have an impact in the performance of any system, and industrial plants are not the exception. This idea leads to the development of a comprehensive approach whose objective is to correlate decisions made on a daily basis with their impacts on a long term point of view. This concept is inspired in the life cycle perspective where it is defended that any action performed over the plant will have an impact, even if this impact is not clearly identified at action execution. The identification and monitoring of life cycle parameters associated to the plant behaviour contribute for reaching a broader view over plant long term behaviour. This thesis proposes the use of information related with maintenance life cycle costs, specifically maintenance costs, in order to extract conclusions about the evolution of this parameter over time. Additionally, and based on the developed risk analysis it guides the decision maker towards the most appropriate action in order to minimise the aforementioned costs. The elements collected regarding maintenance are also available for developing further strategies aiming the improvement of other life cycle parameters.

Finally, the validation of the methodology proposed in this thesis was made in two different stages, for which the type of validation method had to be adapted. The first validation method was performed by means of simulation. This need came with the difficulty in achieving life cycle results in an appropriate time frame since these required long operational periods to enable the extraction of realistic conclusions. Furthermore, the decision methodology developed is strongly dependent on the number of problems occurred in the plant. Obviously this is a situation that plant managers wish to avoid for several reasons, including system safety and cost increase. Thus, and to test the full potential of the proposed decision methodology, a simulator was developed congregating the simulation of the plant with the risk-based decision support system.

The learning capabilities of the algorithm were tested within the simulator. The algorithm proved to possess good accuracy and learning rate, especially when the number of occurrences increased. These results are consistent with what was expected for the algorithm performance. Regarding the simulation results in terms of efficiency of the decision method, they were very promising when compared with other maintenance methodologies commonly used in real industrial environments. In fact, the developed method proved its efficiency in terms of cost minimisation against all the selected maintenance methodologies used as benchmarks.

The second validation methodology was developed aiming the test of the proposed approach in real industrial environments. For this purpose two industrial test cases were developed, in the scope of an European funded project (InLife), both focusing on the application of some of the concepts developed in this thesis. The knowledge model developed for industrial testing was slightly different from the one used for simulation. This was due to the difficulty in collecting, in real industrial environments, enough information to establish the different outcomes for the possible actions. Nevertheless, the model used also included the notion of impact associated to a specific cause, and, although those impacts were static over time, it enabled the implementation of the developed risk analysis strategy. Due to these issues the tests performed in real industrial plants were mainly focused in testing the appropriateness of the model regarding the relations

between symptoms, causes and actions and exploring the possibility of performing risk analysis to identify the most probable cause of a new problem.

The SMEs involved in the test cases adopted the developed knowledge model and used it to structure their production process enabling an increase on the level of structured information, which had a significant impact in the diagnosis process and thus influence maintenance procedures. The risk analysis was tested in terms of selection of the most appropriate service to be used, from the set of available ones. The selection was dependent on the level of risk calculated for each situation. This strategy proved its efficiency in guiding the user to the utilisation of the correct service at each situation.

This thesis tried to give answer for the formulated research question, and in the end it is possible to conclude that the hypothesis considered turn out to be valid. In fact the use of simulation together with data models, instantiated with specific information of the industrial plant, proved their efficiency in providing knowledge about existing options and the impact of each of them in the life cycle of the plant. Note that, despite the system capacities it is not expected to apply it without human supervision. In fact, the correctness of the results and, by this, their continuous refinement, is strongly dependent on the degree of reliability of the knowledge under consideration. For this it is extremely important the validation of the system results which should be done by human intervention.

The existence of technological barriers for applying some of the theoretical concepts developed in this thesis in real industrial environments is recognised. The development of the simulator tried to cope with this difficulty, but the author acknowledges that the existence of real results in terms of life cycle would represent an immense contribution for the overall conclusions.

The author continues working on the development of decision support methodologies based on risk analysis with special focus in industrial environments. Regarding the risk analysis strategy the refinement of the method to perform risk analysis in order to increase results reliability is a subject to be considered in a subsequent development of this work. Additionally different methods to derive possible outcomes are also under reflection. In what concerns the life cycle perspective, the use of the information collected to derive knowledge about other life cycle parameters represents a possible path for the development of the work. Additionally the application of these concepts to new areas (e.g. medicine, energy efficiency, etc.) is also being considered.

## Bibliography

- Aldrich, C., Schmitz, G., & Gouws, F. (2000). Development of fuzzy rule-based systems for industrial flotation plants by use of inductive techniques and genetic algorithms. *The Journal of The South African Institute of Mining and Metallurgy* .
- Amstadter, B. L. (1977). *Reliability Mathematics: Fundamentals, Practices, Procedures*. McGraw-Hill.
- ASME. (1994). *Risk-based Inspection – Development of Guidelines* (Vol. 1). ASME Research.
- ASME. (2003). *Risk-based methods for equipment life management*. New York: ASME International.
- Aström, K., & Wittenmark, B. (1989). *Adaptive Control*. Addison-Wesley.
- Babka, O., & Whar, S. Y. (1997). Case-based reasoning and Decision Support Systems. *1st International Conference on Intelligent Processing Systems*. Beijing: IEEE.
- Badri, M. A. (2007). Dimensions of Industrial Location Factors: Review and Exploration. *Journal of Business and Public Affairs* , 1 (2).
- Baker, D., Bridges, D., Hunter, R., Johnson, G., Krupa, J. M., & Sorenson, K. (2002). *Guidebook to Decision-Making Methods*. From Department of Energy, USA: [http://emi-web.inel.gov/Nissmg/Guidebook\\_2002.pdf](http://emi-web.inel.gov/Nissmg/Guidebook_2002.pdf)
- Barlow, R. E., & Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing: Probability Models*. To Begin With.
- Barnard, G., Jenkins, G., & Winsten, C. (1962). Likelihood Inference and Time Series. *Journal of Royal Statistical Society* , 321–372.
- Barringer, H. P. (2001). *How To Justify Equipment Improvements Using Life Cycle Cost and Reliability Principles*. Humble: Barringer & Associates, Inc.
- Belbin, R. M. (2004). *Management Teams* ( 2nd edition ed.). Elsevier Butterworth-Heinemann.
- Benayoun, R., Roy, B., & Sussman, B. (1966). *ELECTRE: Une méthode pour guider le choix en présence de point de vue multiples*. SEMA-METRA International, Direction Scientifique.
- Benbya, H. (2008). *Knowledge Management Systems Implementation: Lessons from the Silicon Valley*. Oxford: Chandos Publishing.
- Berenji, H., Wang, Y., Saxena, & A. (2005). Dynamic Case Based Reasoning in Fault Diagnosis and Prognosis. *The International Conference on Fuzzy Systems* (pp. 845 - 850). Reno, Nevada: IEEE.
- Bernoulli, D. (1738 (1954)). Specimen Theoriae Novae de Mensura Sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae (Econometrica)* , 5 (22) (1) , 22-36. (L. Sommer, Trans.)
- Bernstein, P. (1998). *Against the Gods: The Remarkable Story of Risk*. John Wiley and Sons.

- Bhargava, H., & Power, D. J. (2001). Decision Support Systems and Web Technologies: A Status Report. *American Conference on Information Systems*. Boston.
- Birnbaum, A. (1962). On the foundations of statistical inference. *Journal of the American Statistical Association* , 57 (298), 269–326.
- Bouchon-Meunier, B. (2009). Interpretable Decisions by Means of Similarities and Modifiers. *International Conference on Fuzzy Systems* (pp. 524 - 529). Korea: IEEE.
- Bouza-Fernandez, J., Gonzalez-Filgueira, G., de las Heras Jimenez, S., & Vazquez-Gonzalez, D. (2010). Expert system of a sewage treatment plant for wood industry. *Emerging Technologies and Factory Automation (ETFA)*. Bilbao: IEEE.
- Brans, J., & Vincke, P. (1985). A preference ranking organization method. *Management Science* , 31, 647-656.
- Brans, J., Vincke, P., & Marechal, B. (1986). How to select and how to rank projects: The PROMETHEE method. *European Journal of Operational Research* , 24, 228- 238.
- Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and Regression Trees*. Belmont: Chapman & Hall.
- Buchanan, L., & O'Connell, A. (2006). A brief history of decision making. *Harvard Business Review* .
- Campos, A. R. (2010). *Intelligent Decision Support Systems for Collaboration in Industrial Plants*. Lisbon: FCT – Universidade Nova de Lisboa.
- Chang, C., Wang, D., Liu, Y., & Qi., B. (2006). Fuzzy similarity measure based case retrieval method. *International Conference on Machine Learning and Cybernetics*. Dalian: IEEE.
- Chen, Y., Kilgour, D. M., & W., H. K. (2006). Multiple criteria classification with an application in water resources planning. *Comput. Oper. Res* , 33 (11), 3301–3323.
- Chen, Z. (2000). *Computational Intelligence for Decision Support*. CRC Press LLC.
- Choo, E. U., & Wedley, W. C. (2008). Comparing Fundamentals of Additive and Multiplicative Aggregation in Ratio Scale Multi-Criteria Decision Making. *The Open Operational Research Journal* , 2, 1-7.
- Choy, K., Lee, W., Lau, H., & Choy, L. (2005). A knowledge-based supplier intelligence retrieval system for outsource manufacturing. *Knowledge-Based Systems* , 18 (1), 1-17.
- Cotaina, N., Matos, F., Chabrol, J., Djeapragache, D., Prete, P., Carretero, J., et al. (2000). *Study of existing Reliability Centered Maintenance (RCM) approaches used in different industries*. Madrid: Universidad Politécnica de Madrid.
- Damasio, A. (1994). *Descartes' Error: Emotion, reason, and the human brain*. G. P. Putnam's Sons.

- Diasio, S. R., & Agell, N. (2009). The Evolution of Expertise in Decision Support Technologies: A Challenge for Organizations. *13th International Conference on Computer Supported Cooperative Work in Design*. Chile.
- Doumpos, M., & Zopounidis, C. (2001). A preference disaggregation decision support system for financial classification problems. *European Journal of Operational Research* , 130 (2), 402-413.
- Doumpos, M., & Zopounidis, C. (2002). Multicriteria classification and sorting methods: A literature review. *European Journal of Operational Research* , 138 (2), 229–246.
- Doumpos, M., & Zopounidis, C. (2002). Multi-criteria classification methods in financial and banking decisions. *International Transactions in Operational Research* , 9, 567–581.
- Doumpos, M., & Zopounidis, C. (2002). *Multi-criteria Decision Aid Classification Methods*. Dordrecht: Kluwer.
- Drucker, P. F. (1986). *The Frontiers of Management : Where Tomorrow's Decisions Are Being Shaped Today*. Penguin Group.
- Ellis, H. (2003 йил April). *Your maintenance costs are too high*. From Plant Maintenance Resource Center.
- Etzioni, A. (1989). Humble Decision Making. *Harvard Business Review* , 4.
- Ferguson, R. L., & Jones, C. H. (1969). A Computer Aided Decision System. *Management Science* , 15 (10), 550-561.
- Figueira, J., Greco, S., & Ehrgott, M. (2005). *Multiple Criteria Decision Analysis: State of the Art Surveys*. New York: Springer- Verlag.
- Figueiredo, M. (1998). Bayesian Methods and Markov Random Fields. *Computer Vision and Pattern Recognition*. Santa Barbara: IEEE.
- Fishburn, P. (1970). Utility Theory for Decision Making. *Operations Research* , 18.
- Follett, M. P. (1918). *The New State: Group Organization the Solution of Popular Government*. Pennsylvania State University Press.
- Galton, F. (1886). Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland* , 15, 243-263.
- Gauss, C. F. (1809). *Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium*. (C. H. Davis, Trans.) New York.
- Gigerenzer, G. (1995). *Simple Heuristics That Make Us Smart*. Anchor Books.
- Gitzel, R. (2010). *Statistical failure prediction: Take a pragmatic approach to the repair/replace*. Putman Media, Inc.
- Goel, S., & Salganik, M. (2009). Respondent-driven sampling as Markov chain Monte Carlo. *Statistics in Medicine* , 28, 2202–2229.
- Golding, A. R., & Rosenbloom, P. S. (1991). Improving Rule-based systems through Case-based reasoning. *9th National Conference on Artificial Intelligence*. Anaheim.

- Gorry, A., & Scott-Morton, M. (1971). A Framework for Information Systems. *Sloan Management Review* , 13 (1), 56-79.
- Grünig, R., & Kühn, R. (2005). *Successful Decision-Making: A Systematic Approach to Complex Problems*. (M. Matt, Ed.) Springer.
- Harris, R. (2009). *Introduction to Decision Making*. From VirtualSalt: <http://www.virtualsalt.com/crebook5.htm>
- Heckerman, D., Geiger, D., & Chickering, D. (1995). Learning bayesian networks: The combination of Knowledge and statistical data. *Machine learning* , 20, 197-243.
- Henley, E. J., & Kumamoto, H. (1981). *Reliability Engineering and Risk Assessment*. Prentice-Hall.
- Holsapple, C. W., & Winston, A. B. (2000). *Decision Support Systems: A Knowledge-based Approach*. Minneapolis/St Paul: West Publishing Company.
- Holt, C. C., & Huber, G. P. (1969). A Computer Aided Approach to Employment Service Placement and Counseling. 15 (11), 573-595.
- Holton, G. A. (2004). Defining Risk. *Financial Analysts Journal* , 60 (6), 19-25.
- Ishizuka, M., & Matsuda, T. (1990). Knowledge acquisition mechanisms for a logical knowledge base including hypotheses. *Knowledge-Based Systems* , 3 (2), 77-86.
- Jakulin, A., & Rish, I. (2006). Bayesian Learning of Markov Network Structure. *7th European Conference on Machine Learning* (pp. 198-209). Berlin: Springer.
- Janis, I. (1972). *Victims of groupthink: a psychological study of foreign-policy decisions and fiascoes*. Boston: Mifflin.
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* , XLVII, 263-291.
- Kaiser, K. A., & Gebraeel, N. Z. (2009). Predictive Maintenance Management Using Sensor-Based Degradation Models. *Transactions on Systems, Man and Cybernetics, A Systems and Humans* , 39 (4).
- Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME Journal of Basic Engineering* , 82, 35-45.
- Keeney, R. L., & Raiffa, H. (1976). *Decision with Multiple Objectives: Preferences and Value Tradeoffs*. New York: John Wiley and Sons.
- Keogh, E. J. (2001). Tutorial on Mining and Indexing Time Series Data. *International Conference on Data Mining* (p. IEEE). San Jose: IEEE.
- Kim, H. H., Ha, K. N., & Lee, K. C. (2009). Resident Location-Recognition Algorithm Using a Bayesian Classifier in the PIR Sensor-Based Indoor Location-Aware System. *Transactions on Systems, Man and Cybernetics, C Applications and reviews* , 39 (2).

- Klutke, G., Kiessler, P. C., & Wortman, M. A. (2003). A Critical Look at the Bathtub Curve. *Transactions on Reliability* , 52 (1).
- Knight, F. (1921). *Risk, Uncertainty, and Profit*. Boston: Houghton Mifflin Co.
- Kobashikawa, C., Hatakeyama, Y., Dong, F., & Hirota, K. (2009). Fuzzy Algorithm for Group Decision Making With Participants Having Finite Discriminating Abilities. *Transactions in Systems, Man and Cybernetics. A Systems and Humans* , 39 (1).
- Koller, G. (1999). *Risk Assessment and Decision Making in Business and Industry: a practical guide*. CRC Press LLC.
- Kolodner, J. (1993). *Case-based reasoning*. Morgan-Kaufmann.
- Lapin, L. (1987). *Statistics for Modern Business Decisions*. Harcourt Brace Jovanovich.
- Lewin, K. (1951). *Field theory in social science*. New york: Harper.
- Lewin, K. (1848). *Resolving social conflicts*. New York: Harper.
- Lindley, D. (1991). *Making Decisions* (2nd Edition ed.). John Wiley & Sons.
- Linkov, I., Varghese, A., Jamil, S., Seager, T., Kiker, G., & Bridges, T. (2004). Multi-criteria decision analysis: A framework for structuring remedial decisions at the contaminated sites. In I. A. Linkov (Ed.), *Comparative Risk Assessment and Environmental Decision Making* (pp. 15-54). New York: Springer.
- Little, J. D. (1975). Brandaid, an On-Line Marketing Mix Model, Part 2: Implementation, Calibration and Case Study. *Operations Research* , 23 (4), 656-673.
- Little, J. D. (1970). Models and Managers: The Concept of a Decision Calculus. *Management Science* , 16 (8), 466-485.
- Lucas, P. (2010). *Knowledge Representation and Reasoning – Logic meets probability theory*. Nijmegen: Radboud University.
- Macharis, C., Springael, J., De Brucker, K., & Verbeke, A. (2004). PROMETHEE and AHP: The design of operational synergies in multicriteria analysis: Strengthening PROMETHEE with ideas of AHP. *European Journal of Operational Research* , 153, 307-317.
- Madson, A. L., & Jensen, F. V. (1998). Lazy propagation in junction trees. *14th Conference on uncertainty in Artificial Intelligence*. Madison: AUAI Press.
- Malakooti, B., & Yang, Z. Y. (2004). Clustering and group selection of multiple criteria alternatives with application to space-based networks. *Transactions in Systems, Man and Cybernetics. B, Cybernetics* , 34 (1), 40-51.
- Marques, M., & Neves-Silva, R. (2008). Decision Support based on Services for Complex Manufacturing Systems. *8th Portuguese Conference on Automatic Control*. Vila Real, Portugal: IFAC.

- Marques, M., & Neves-Silva, R. (2009). Decision support system using risk assessment for life cycle management of industrial plants. *13th Symposium on Information Control Problems in Manufacturing*. Moscow: IFAC.
- Marques, M., & Neves-Silva, R. (2011). Risk Based Decision Support System for Life Cycle Management of Industrial Plants. *9th International Conference on Industrial Informatics*. Caparica, Lisbon: IEEE.
- McCalley, J., Voorhis, T. v., & Jiang, Y. (2003 йил October). Risk-Based Maintenance Allocation and Scheduling for Bulk Transmission System Equipment. *PSERC Publication* , 3 (26).
- Mi, C., Qian, H., Liu, S., & Chang, Z. (2008). Study on case retrieving in case-based reasoning based on grey incidence theory and its application in bank regulation. *International Conference on Fuzzy Systems* (pp. 1530 - 1533). Hong Kong: IEEE.
- Mobley, R. K. (2002). *An Introduction to Predictive Maintenance* (2nd Edition ed.). (Butterworth-Heinemann, Ed.) Elsevier Science.
- Mora, E. (2010 йил 20-February). *From the Origins of Maintenance to The Total Productive Maintenance Concept*. Retrieved 2011 йил 20-April from TPMonLine.com: <http://www.tpmonline.com/TPM-001.htm>
- Mora, E. (2011 йил 8-May). *Maintenance in History*. Retrieved 2011 йил 10-May from TPMonLine.com: [http://www.tpmonline.com/articles\\_on\\_total\\_productive\\_maintenance/tpm/tpmprocess/maintenanc einhistory.htm](http://www.tpmonline.com/articles_on_total_productive_maintenance/tpm/tpmprocess/maintenanc einhistory.htm)
- Murthy, S. K. (1998). Automatic Construction of Decision Trees from Data: A Multi-disciplinary Survey. *Data Mining and Knowledge Discovery* , 2, 345–389.
- Nelson, W. (1990). *Accelerated Testing Statistical Models, Test Plans, and Data Analysis*. New York: Jon Wiley & Sons.
- Netherton, D. (1998). Standard To Define RCM (Part 2). *Maintenance Technology* .
- Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behaviour* (2nd Edition ed.). Princeton: Princeton University Press.
- Neves-Silva, R. (1999). *Dual and predictive control of processes with accessible disturbances*. Lisbon: UTL-IST.
- Nikolova, N., Shulus, A., Toneva, D., & Tenekedjiev, K. (2005). Fuzzy rationality in quantitative decision analysis. *Journal of Advanced Computational Intelligence* , 9 (1), 65-69.
- Nja, O., & Rake, E. L. (2009). A discussion of decision making applied in incident command. *International Journal of Emergency Management* , 6 (1), 55-72.
- Nunes, C., Mendonça, T., Lemos, J., & Amorim, P. (2007). Control of depth of anesthesia using MUSMAR--exploring electromyography and the analgesic dose as accessible disturbances. *29th*



- Annual International Conference in Engineering in Medicine and Biology Society* (pp. 1574 - 1577). Lyon: IEEE.
- Osborne, M. (2009). *MT: History and Rule-based Systems*. University of Edinburgh, School of Informatics.
- Paine, J. I. (1996). *Expert Systems*. Oxford University.
- Park, S. K. (1993). Condition-Based Predictive Maintenance by Multiple Logistic Function. *Transactions on Reliability* , 42 (4), 556-560.
- Pearl, J. (1985). Bayesian Networks: A model of self-activated memory for evidential reasoning. *Conference of the Cognitive Science Society*, (pp. 329–334). Irvine.
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers.
- Pendse, N. (1997). *Origins of today's OLAP products*. From The OLAP Report: [www.olapreport.com](http://www.olapreport.com)
- Pomerol, J., & Barba-Romero, S. (2000). Multicriterion Decision in Management. *International Series in Operations Research & Management Science* , 25.
- Price, R. (1764). An Essay towards solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society of London* .
- Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. San Mateo: Morgan Kaufmann.
- Raiffa, H. (1982). *The Art and Science of Negotiation*. Cambridge: Harvard Univ. Press.
- Rausand, M., & Høyland, A. (2004). *System Reliability Theory: Models, Statistical Methods and Applications* (2nd Edition ed.). New York: John Wiley and Sons.
- Raymond, R. (1966). Use of the Time-sharing Computer in Business Planning and Budgeting. *Management Science* , 12 (8).
- Remmen, A., Jensen, A. A., & Frydendal, J. (2007). *Life Cycle Management: A business guide to sustainability*. UNEP/SETAC.
- Rockwell Automation. (2004). *Reaping the Rewards of Strategic Maintenance*. Milwaukee: Rockwell International Corporation.
- Roy, B. (1968). Classification and choice when there are many points of view - The ELECTRE method. *French Review of Computational Science and Operational Research* , 8, 57-75.
- Roy, B. (1996). *Multicriteria Methodology for Decision Aiding*. Dordrecht, The Netherlands: Kluwer.
- Russel, S., & Norvig, P. (2003). *Artificial Intelligence – A modern Approach*. Prentice Hall.
- Saaty, T. (1996). *Decision Making with Dependence and Feedback: The analytic network process*. Pittsburgh: RWS Publications.
- Saaty, T. (1999). Fundamentals of the analytical network process. *10th Interbational Symposium on the analytical hierarchy process*. Kobe.

- Saaty, T. L. (1980). *Analytic Hierarchy Process*. New York: McGraw-Hill.
- Sales, M. B., Schwaab, A. A., & Nassar, S. M. (2010). Application of Bayesian Networks to Assist the Expansion of the Digital Inclusion Of Elderly People. *Latin America Transactions* , 8 (3), 275-279.
- Sap, M. N., & Khokhar, R. H. (2004). Fuzzy Decision Tree for Data Mining of Time Series Stock Market Databases. *5th International Conference for the Critical Assessment of Microarray Data Analysis* . Durham: Academia.
- Sastry, S., & Bodson, M. (1989). *Adaptive Control: Stability, Convergence and Robustness*. New Jersey: Prentice Hall, Inc.
- Schank, R. (1982). *Dynamic Memory: A Theory of Learning in Computers and People*. New York: Cambridge University Press.
- Scott-Morton, M. (1967). *Computer-Driven Visual Display Devices - Their Impact on the Management Decision-Making Process*. Harvard Business School.
- Scott-Morton, M., & McCosh, A. M. (1968). Terminal Costing for Better Decisions. *Harvard Business Review* , 46 (3), 147–156.
- Scott-Morton, M., & Stephens, J. A. (1968). The impact of interactive visual display systems on the management planning process. *IFIP Congress*. 2. IFIP.
- Scully, J. K. (2006). Case Based Reasoning in the plenary diagnostic environment. *Autotestcon* (pp. 591 - 598). Anaheim: IEEE.
- Simon, H. (1991). Bounded Rationality and Organizational Learning. *Organization Science* , 2 (1), 125-134.
- Sondalini, M. (2008). *Life Time Reliability Solutions*. Retrieved 2011 йил 20-April from <http://www.lifetime-reliability.com/free-articles/maintenance-management/condition-based-maintenance.html>
- Stamatelatos, M. (2000). *Probabilistic Risk Assessment: What Is It And Why Is It Worth Performing It?* NASA Office of Safety and Mission Assurance.
- Steuer, R. E. (1986). *Multiple Criteria Optimization: Theory, Computation, and Application*. New York: John Wiley and Sons.
- Takata, S., Kimura, F., Houten, F. v., Westkamper, E., Shpitalni, M., Ceglarek, D., et al. (2004). Maintenance: Changing Role in Life Cycle Management. *Annals of the CIRP* , 53 (2).
- Tan, A. W., & Kumar, A. (2006). A decision-making model for reverse logistics in the computer industry. *International Journal of Logistics Management* , 17 (3), 331 - 354.
- Tao, G., Chen, S. H., Tang, X. D., & Joshi, S. M. (2004). *Adaptive Control of Systems with Actuator Failures*. Springer.

- Tenekedjiev, K., Nikolova, N., & Dimitrakiev, D. (2004). Application of the triple bisection method for extraction of subjective utility information. *2nd International Conference on Management and Engineering*, 2, pp. 115–117. Sofia.
- Thibault, A., Siadat, A., & Martin, P. (2006). A Framework For Using A Case Based Reasoning System Applied To Cost Estimation. *Conference on Cybernetics and Intelligent Systems* (pp. 1-6). Bangkok: IEEE.
- Turban, E. (1967). The Use of Mathematical Models in Plant Maintenance Decision Making. *Management Science*, 13 (6).
- Tversky, A., & Kahneman, D. (1981). The Framing of Decisions and the Psychology of Choice. *Science*, 211 (4481), 453-458.
- U.S. DEPARTMENT OF TRANSPORTATION. (1998). *FAA ORDER 8040.4. FEDERAL AVIATION ADMINISTRATION*.
- UNEP/SETAC, Life Cycle Initiative. (2009). *Life Cycle Management: How business uses it to decrease footprint, create opportunities and make value chains more sustainable*. Ireland: Power Editing.
- Urban, G. (1967). SPRINTER: A Tool for New Products Decision Makers. *Industrial Management Review*, 8 (2), 43-54.
- Vandaie, R. (2008). The role of organizational knowledge management in successful ERP implementation projects. *Knowledge-Based Systems*, 21 (8), 920-926.
- Vroom, V. H., & Yetton, P. W. (1973). *Leadership and decision-making*. University of Pittsburg Press.
- Wang, Z., Kuruoğlu, E., Yang, X., Xu, Y., & Huang, T. (2010). Time Varying Dynamic Bayesian Network for Nonstationary Events Modeling and Online Inference. *Transactions on Signal Processing*, 59 (4).
- Watson, I. D. (1997). *Applying Case-Based Reasoning: Techniques for Enterprise Systems*. Morgan Kaufmann Publishers Inc.
- Watson, I. D. (1995). *Progress in Case-Based Reasoning*. Berlin: Springer.
- Wattayay, W., & Peng, Y. (2004). A Bayesian network based framework for multi-criteria decision making. *XVII-th Multiple Criteria Decision Making International Conference*. Whistler.
- X-Stream LEAN. (2007). *Total Productive Maintenance: How and Why to Implement TPM in Your Factory*. Queen Street, Pottstown: X-Stream LEAN, LLC.
- Xu, Y., & Bernard, A. (2011). *Knowledge-Based Systems*, 24 (1), 166-175.
- Yamada, Y., Suzuki, E., Yokoi, H., & Takabayashi, K. (2003). Decision-tree Induction from Time-series Data Based on a Standard-example Split Test. *20th International Conference on Machine Learning*. Washington DC: Springer.

Yoon, K., Hwang, C., & Yoon, K. (1995). *Multiple Attribute Decision Making: An Introduction*. Sage Publications.

## *Annexes*

## **A. Decision-making methods**

When speaking about decision making in real world, especially if when dealing with business world, many problems are still solved using a simple Cost-Benefit Analysis (CBA) which evaluates the costs and benefits of the alternatives on monetary basis. CBA is a great technique to solve many problems but it has some problems in incorporating different factor (e.g. environmental impact) to improve the quality decision making. Thus, there are situations where different techniques that enable the aggregation of such factors, by considering other criteria besides monetary valuation, are needed.

In fact, the most interesting decision problems, and the ones about which more research work has been developed, are the ones that involve a finite set of criteria and a finite set of alternatives. In the next sections some of the most used methods stating their intrinsic characteristics are presented.

Methods like Pros and Cons analysis, Maximin, Maximax, Conjunctive and Disjunctive as well as Lexicographics are part of the so-called elementary methods which are simple approaches where, in principle, no computational support is needed. As stated by Linkov (Linkov, Varghese, Jamil, Seager, Kiker, & Bridges, 2004) these methods are best suited for problems with a single decision maker, few alternatives and criteria.

Here an overview of these methods as well as of others with higher level of complexity, like Multi-Attribute Utility Theory, Analytic Hierarchy Process and Outranking methods is provided.

### **Pros and Cons analysis**

Pros and cons analysis is a method based on a qualitative comparison method. For each alternative the positive aspects (pros) and the negative ones (cons) are identified. The lists of the pros and cons are then compared one to another for each alternative and the alternative with the strongest pros and weakest cons is the selected one. The method does not require any mathematical skill and its implementation is straightforward.

### **Maximin and maximax methods**

The Maximin method is based on a strategy that tries to avoid the worst possible performance, maximizing the minimal performing criterion. The selected alternative is the one that presents the highest score for the weakest criterion.

The maximin method can be used only when all criteria are comparable so that they can be measured on a common scale. According to Linkov (Linkov, Varghese, Jamil, Seager, Kiker, & Bridges, 2004) this can constitute a limitation for the application of the method.

On the other hand the Maximax method selects the alternative by its best criterion rating. As explained by Yoon (Yoon, Hwang, & Yoon, 1995) in the Maximax method only a single criterion represents an alternative and all the others criteria are ignored. Thus the method also requires some degree of comparability among criteria.

### **Conjunctive and disjunctive methods**

Conjunctive and disjunctive methods are based on the concept of satisfaction rather than on achieving best performance in each criterion. The conjunctive method imposes that the selected alternative should achieve minimal performance level for all criteria. On the other hand the disjunctive method requires that the alternative should exceed the given threshold for at least one criterion.

In both methods the alternatives that do not meet the conjunctive or disjunctive rules are ignored in further consideration. The disjunctive method is usually used together with the conjunctive one as a prior alternative selection approach. The resultant subset of alternatives can afterwards be analysed by other decision making methods.

### **Lexicographic methods**

These methods are based on the idea of weighted vectors, one for each criterion, representing the behaviour of each alternative against criteria.

Pomerol (Pomerol & Barba-Romero, 2000) divides the lexicographic methods into five categories: basic, multi-criterion, semiorder, permutation and ordering with aspirations.

Despite their intrinsic differences the basic idea is to rank criteria in order of their importance. Then the criterion with the highest rank is used to rank the alternatives. The alternative with the best performance score on the most important criterion is chosen. Where there are ties they are solved using the following criterion until no ties persist and a unique alternative is found.

### **Multi-Attribute Utility Theory (MAUT) and Multi-Attribute Value Theory (MAVT)**

Keeney and Raifa (Keeney & Raiffa, 1976) have introduced the differentiation between the notions of preference based on certainty and choice under risk. The first ones rely on the idea of value functions to represent preferences based on the notions of ordinal comparisons and strength of preference, whereas the latest incorporate the concept of risky choices which is represented using utility functions. Since decision under risk is, in many applications, a central part of the problem, here we focus on MAUT. Nevertheless, the main concepts ruling MAUT and MAVT are the same.

Following this notion MAUT extends the concept of expectation to include explicit modelling of risky preferences. As explained by Figueira (Figueira, Greco, & Ehrgott, 2005) MAUT tries to assign a utility value to each action. This utility is a real number, representing the degree of preference of the considered action. The idea is to use utility functions, relative to each single criterion, which can be applied to transform the raw performance values of the alternatives against different criteria, both factual (objective, quantitative) and judgmental (subjective, qualitative), into a universal range. Additionally utility functions convert the raw performance values, to reflect preference. In this case, a performance with higher preference obtains a higher utility value. The formalization of the expected utility theory by von Neumann and Morgenstern (Neumann & Morgenstern, 1947) is seen as the most important contribution for MAUT. This process was based on the idea that a person:

- has several objectives;

- can make a general assessment of alternative behaviours taking into account those objectives; and
- can classify alternative behaviours' according to own preferences.

If all these conditions are satisfied then it is possible to express the preference using a utility function as an indifference curve that can be defined.

These principles are reflected in the three axioms, proposed by Fishburn (Fishburn, 1970) that rule MAUT: Ordering, Independence and Continuity. The expected utility theory states that these axioms are only valid when there is a real-valued function  $u$  such that for all  $p, q$  in  $P$ ,  $p$  is preferable to  $q$ , if and only if:

$$\sum_{x \in X} p(x)u(x) \geq \sum_{x \in X} q(x)u(x) \quad (\text{A.1})$$

where  $P$  is a convex set of simple probability distributions on a non empty set  $X$  of possible outcomes.

The construction of the global utility function  $U(X)$  starts with the construction of partial utility functions for each attribute,  $u_i(x_i)$ , satisfying the expected utility hypothesis for variations in  $X$ . This process may present some technical problems related to its axiomatic basis due to the need of guarantee that those axioms are not violated. In fact if Independence cannot be guaranteed then the expected utility model will not be linear which must be treated by using nonlinear utility models.

Nevertheless, and despite its simplicity MAUT, constitutes a reliable method to build preferences where the major problem can be on aggregating the different  $u_i(x_i)$  into  $U(X)$  still satisfying the expected utility hypothesis. The most common option to solve this problem is to use an additive or a multiplicative form to treat the set of  $u_i(x_i)$ . However these are not universal solutions since their application is limited to some conditions. A comparison of the advantages and limitations of these two aggregation methods in terms of their fundamental properties is provided by Choo and Wedley (Choo & Wedley, 2008).

### **Analytic Hierarchy Process (AHP) and Analytic Network Process (ANP)**

In what concerns the Analytic Hierarchy Process the basic idea is to convert subjective assessments of relative importance to a set of overall scores or weights in order to build utility functions. AHP is one of the most widely applied multi-attribute decision making methods. It is based on a theory of measurements that uses pair wise comparisons, along with expert judgements, to deal with the measurement of qualitative or intangible criteria. The pair-wise comparisons are made using a nine-point scale (see Table A.1):



Table A.1. AHP nine-point scale

Grade	Description
1	Equal importance or preference
3	Moderate importance or preference of one over another
5	Strong or essential importance or preference
7	Very strong or demonstrated importance or preference
9	Extreme importance or preference

The intermediate values are normally not used but, when needed, they can be applied to transmit the idea of half distance between two values.

The aggregation of the results is done by multiplying each normalized alternative score by the corresponding normalized criterion weight, summing the results for all. The preferred alternative will have the highest total score.

A recent extension of AHP is the Analytic Network Process (ANP) also proposed by Saaty (Saaty T. , Decision Making with Dependence and Feedback: The analytic network process, 1996), (Saaty T. , Fundamentals of the analytical network process, 1999). The process allows the inclusion of all the elements and criteria, tangible and intangible involved in making the best decision. In fact ANP allows both interaction and feedback within elements of a cluster (inner dependence) and between clusters (outer dependence). Additionally the method provides a way to input judgments and measurements to derive ratio scale priorities for the distribution of influence among the elements and groups of elements in the decision. The feedback capability enables the capture of complex effects of interplay in human society, especially when risk and uncertainty are involved. The ANP has been applied to a large variety of decisions: marketing, medical, political, social, forecasting and prediction, etc.

### **Outranking methods**

Most outranking methods assume that alternatives and criteria are identified, since these are needed to apply the method which uses the same data required to build the decision table.

The methods are based in the idea that alternative a outranks alternative b if, on a great part of the criteria, a performs at least as good as b (concordance condition), while its worse performance is still acceptable on the other criteria (non-discordance condition). This exercise is made for all alternatives (in a pair-wise mode) to achieve a combined ranking. According to Figueira (Figueira, Greco, & Ehrgott, 2005) an outranking relation is a binary selection  $S$  defined on the set of potential actions  $A$  such that  $aSb$  if there are enough arguments to decide that  $a$  is at least as good as  $b$ .

The most important outranking methods are organised in two main families: the ELECTRE and the PROMETHEE. The ELECTRE methods are based on concordance and discordance indices and, despite their inner characteristics, they all try to establish a partial rank of using qualitative criteria such as credibility. The methods are considered very important, especially from historical point of view, since ELECTRE I, introduced by Benayoun (Benayoun, Roy, & Sussman, 1966), was the first outranking method to be described.

On the other hand the PROMETHEE methods, introduced by Brans (Brans & Vincke, 1985) (Brans, Vincke, & Marechal, 1986), had ELECTRE as a starting point but they tried to bring more flexibility when modelling preferences as well as decrease the level of difficulty. The methods have the decision table as their starting point where weighting criteria is a very important step. Thus it is essential for decision makers to be able to see to what extent changes of the weights of the criteria will impact the rankings and PROMETHEE includes several tools to support this task.

Additionally, some developments have been made trying to find new approaches for outranking methods, especially taking advantage of their pair-wise base. The work developed by Macharis (Macharis, Springael, De Brucker, & Verbeke, 2004) to incorporate AHP ideas in the PROMETHEE methods is an example.

## B. Life Cycle Parameters

Life Cycle Parameter	Formula	Definition
Life Cycle Costs (LCC)	$LCC = A_c + O_c + M_c + D_c$	The term LCC is the total cost of a system during its life cycle from concept to scrap.
Acquisition cost ( $A_c$ )	$A_c = \text{Purchase cost}$ $+ \text{Administration}$ $/ \text{Engineering cost}$ $+ \text{Installation cost}$ $+ \text{Training cost}$ $+ \text{Conversion cost}$ $+ \text{Transportation cost}$	This is the initial cost which could be easily calculated during the conception phase.
Operating cost ( $O_c$ )	$O_c = \text{Direct labour cost} + \text{Utilities cost}$ $+ \text{Consumables cost}$ $+ \text{Waste handling cost}$ $+ \text{Spare parts inventory cost}$	This is the total cost which is consumed for operating the system during operational phase.
Maintenance cost ( $M_c$ )	$M_c =$ $= \text{Scheduled maintenance cost}$ $+ \text{Unscheduled maintenance cost}$	The cost for maintaining the system during operational phase
	$M_c =$ $= \text{all direct cost for maintenance activities}$ $+ \text{all indirect cost for maintenance activities}$	This is the total maintenance cost carried out only for the maintenance activities. Direct costs are costs charged to a maintenance budget as fixed costs (e.g. personnel, materials, subcontractors, and overhead). Indirect costs are related to loss of revenue due to

		unavailability
Decommissioning cost ( $D_c$ )	$D_c = \text{Decommissioning cost} - \text{Residual cost}$	The cost for decommissioning the system during disposal phase
Reliability	$\text{Reliability} = 100 - AFR$	The probability that a system will perform its intended function for a specified interval under stated condition
Average Failure Rate ( $AFR$ )	$AFR = \frac{8760}{MTBF}$	This the annual failure rate of the system. Note: 8760 is the total number of hours per year
Mean Time Between Failure ( $MTBF$ )	$MTBF = \frac{\text{Total accumulated operating time}}{\text{Number of failures}}$	Mean time between failures is calculated from the total accumulated operating time divided by the number of failures during the same period.
Availability		The ability of an item to be in a state to perform a required function under given conditions at a given instant of time or over a given time interval, assuming that the required external resources are provided
Inherent availability ( $IA$ )	$IA = \frac{MTBF}{(MTBF + MTTR)} \times 100$	Inherent availability reflects the percentage of time a

		product would be available if no delays due to maintenance, supply, etc. (i.e., not design-related) were encountered.
Mean Time To Repair (MTTR)	$MTTR =$ $= \text{Average total time duration required to physically repair or replace the failed item} + \text{average total time to restore operational functionality}$	MTTR refers to the average total time duration required to physically repair or replace the failed item and to reinstate the operational functionality. Part of the system downtime may be due to time delays (spares, resources), which are not included in the MTTR.
Operational availability (OA)	$OA = \frac{MTBM}{(MTBM + MDT)}$	Operational availability is similar to inherent availability but includes the effects of maintenance delays and other non-design factors.
Mean Time Between Maintenance (MTBM)	$MTBM =$ $= \text{Average time taken between all system maintenance actions}$	A basic measure of reliability for repairable fielded systems. The average time between all system maintenance actions. Maintenance actions are for repair or preventive purposes.
Mean Down Time (MDT)	$MDT = \text{Mean repair time} + \text{all delay time}$	The average time a

		system is unavailable for use, for example due to a failure. Mean downtime includes the mean repair time plus all delay time associated with a repairman arriving with the appropriate replacement parts.
Maintainability	<p>Quantitative parameters which are related to maintainability are:</p> <ul style="list-style-type: none"> <li>• Mean Time Between Maintenance</li> <li>• Mean Down Time</li> <li>• Total maintenance cost</li> <li>• Maintenance intensity</li> <li>• Maintenance cost intensity</li> <li>• Material cost intensity</li> <li>• Maintenance labour cost portion</li> <li>• Material cost portion</li> <li>• Turnover related maintenance ratio</li> </ul>	The relative ease and economy of time and resources with which an item can be retained in, or restored to, a specified condition when maintenance is performed by personnel having specified skill levels, using prescribed procedures and resources, at each prescribed level of maintenance and repair. In this context, maintainability is a function of design
Mean Time Between Maintenance ( <i>MTBM</i> )	$MTBM =$ <p>= Average time taken between all system maintenance actions</p>	A basic measure of reliability for repairable fielded systems. The average time between all system maintenance actions. Maintenance actions are for repair or preventive purposes.

Mean Down Time ( <i>MDT</i> )	$MDT = \text{Mean repair time} + \text{all delay times}$	The average time a system is unavailable for use, for example due to a failure. Mean downtime includes the mean repair time plus all delay time associated with a repairman arriving with the appropriate replacement parts
Total maintenance cost	$\begin{aligned} \text{Total maintenance cost} \\ = \text{all direct cost} \\ + \text{all indirect cost} \end{aligned}$	This is the total maintenance cost carried out only for the maintenance activities. Direct costs are costs charged to a maintenance budget as fixed costs (e.g. personnel, materials, subcontractors, and overhead). Indirect costs are related to loss of revenue due to unavailability
Maintenance intensity	$\begin{aligned} \text{Maintenance Intensity} = \\ = \frac{\sum \text{current maintenance costs}}{\text{Acquisition value of the equipment}} * 100\% \end{aligned}$	Maintenance intensity is defined as the ratio between the sum of the current cost equipment to the acquisition value of the equipment
Maintenance cost intensity	$\begin{aligned} \text{Maintenance Cost Intensity} = \\ = \frac{\text{Maintenance cost}}{\text{Production cost}} * 100\% \end{aligned}$	Maintenance cost intensity is the ratio between the maintenance cost and production cost of the

		equipment
Material cost intensity	$\text{Material Cost Intensity} = \frac{\text{Material cost}}{\text{Acquisition value of the equipment}} * 100\%$	Material cost intensity is defined as the ratio between the material cost for maintaining the equipment to the acquisition value of the equipment
Maintenance labour cost portion	$\text{Maintenance labour cost portion} = \frac{\text{Maintenance labour cost}}{\text{Total Maintenance cost}} * 100\%$	
Material cost portion	$\text{Material cost portion} = \frac{\text{Material cost}}{\text{Total Maintenance cost}} * 100\%$	Material cost portion is the cost spent for the material to maintain the equipment, it is calculated by the ratio of material cost to the total maintenance cost
Turnover related maintenance ratio	$\text{Turnover related maintenance ratio} = \frac{\text{Total Maintenance cost}}{\text{Turnover cost}} * 100\%$	Turnover related maintenance ratio is the ratio between total maintenance cost of all equipment to the turnover cost
Safety		The freedom from those conditions that can cause death, injury occupational illness, or damage to or loss of equipment or property, or damage to the environment.
Overall Equipment Effectiveness (OEE)	$\text{OEE factor} = \% \text{ of Availability (A)} * \% \text{ of Performance (P)} *$	OEE - Overall Equipment Effectiveness that



	<i>% of Quality (Q)</i>	measures the combination of three elements for the physical asset; equipment asset availability, performance and quality output
% of Availability	$A = \frac{\text{Total hours (24/7/365)}}{\text{equipment uptime}} * 100\%$	Percent of scheduled production (to measure reliability) or calendar hours 24/7/365 (to measure equipment utilization), that equipment is available for production
% of Performance	$P = \text{percentage of parts produced / time frame}$	Percent of parts produced per time frame, of maximum rate OEM rated production speed at. If OEM specification is not available, use best known production rate.
	$P = \frac{\text{Total breakdown time}}{\text{total production time}} * 100\%$	Performance is the ratio of total breakdown time to the total production time, it is showed in percentage
% of Quality	$Q = \text{percentage of good sellable parts out of the total parts produced / time frame}$	Percent of good sellable parts out of total parts produced per time frame.

Net Equipment Effectiveness ( <i>NEE</i> )	$NEE = \left( \frac{\text{Running time}}{\text{Running time} + \text{Breakdown time}} \right) \times \left( \frac{\text{Expected takt time}}{\text{Actual takt time}} \right) \times \left( 1 - \frac{\text{number of rejected parts}}{\text{Total number of parts}} \right) \times 100\%$	This is the another way of expressing and calculating the overall equipment effectiveness (OEE)
Overall Craft Effectiveness ( <i>OCE</i> )	<p><i>OCE factor</i></p> <p>= % of Craft Utilization (<math>C_U</math>)</p> <p>* % of Craft Performance (<math>C_P</math>)</p> <p>* % of Craft Service Quality (<math>C_{SQ}</math>)</p>	<p>The OCE Factor focuses upon craft labour productivity and measuring/ improving the value added contribution that people assets make. Just like OEE, there are three elements to the OCE Factor:</p> <ul style="list-style-type: none"> <li>• the effectiveness factor</li> <li>• the efficiency factor</li> <li>• the quality factor</li> </ul>
% of Craft Utilization	$\% C_U = \frac{\text{Total craft hours available \& paid}}{100\%}$	This element of OCE relates to measuring how effective we are in planning and scheduling craft resources so that these assets are doing value-added, productive work (wrench time).
% of Craft Performance	$\% C_P = \frac{\text{Total actual craft hours required}}{100\%}$	This element relates to how efficient we are in actually doing hands-on craft work when compared to an established planned time or performance

		standard.
% of Craft Service Quality	$\% C_{SQ} = \text{Quality of actual work} * 100\%$	This element includes quality of the actual work, where certain jobs possibly require a call back to the initial repair thus requiring another trip to fix it right the second time. However, Craft Service Quality can be negatively impacted due to no fault of the crafts person when hasty repairs, patch jobs or inferior repair parts/materials create the need for a call back
Failure rate ( <i>FR</i> )	$FR = \frac{\text{number of failure}}{\text{operating time}} * 100\%$	Number of failure per unit of gross operating period in terms of time, events, cycles or number of parts
Spontaneity Intensity ( <i>SI</i> )	$SI = \frac{\text{Total unplanned maintenance time}}{\text{Total maintenance time}} * 100\%$	This is one of the maintenance parameter, which is the ratio between total unplanned maintenance time to the total maintenance time
Breakdown Intensity ( <i>BI</i> )	$BI = \frac{\text{Total number of nbreakdown}}{\text{total production time}} * 100\%$	Breakdown intensity is defined as the ratio between the number of breakdown occurred for a period

		of total production time.
--	--	------------------------------

## C. Computation of Similarity

Consider a state vector  $x$  with the set of relevant information for the characterization of the state of a system in some particular instant of time when a *case* is to be reported.

$$x = [x_1 \dots x_n]^T \quad (C.1)$$

where each variable  $x_i$  can be of four different types:

- A) Boolean, i.e. take the possible values TRUE or FALSE;
- B) Discrete non numerable, e.g. colours;
- C) Discrete numerable, e.g. weekdays;
- D) Continuous, e.g. the temperature value from a range.

If  $x^A$  and  $x^B$  are the state vectors correspondent to case  $A$  and  $B$ , respectively, the *Similarity* between cases  $A$  and  $B$ ,  $S\{A, B\} = S(x^A, x^B)$  is defined fulfilling the following properties:

1.  $S\{A, B\} \in [0; 1]$
2.  $S\{A, A\} = 1$
3.  $S\{A, C\} \geq S\{A, B\} \cdot S\{B, C\}$
4. ...

Alternatively, the *Distance* between cases  $A$  and  $B$ ,  $D\{A, B\} = D(x^A, x^B)$  is defined fulfilling the following properties:

1.  $D\{A, B\} \in [0; \infty[$
2.  $D\{A, A\} = 0$
3.  $D\{A, C\} \leq D\{A, B\} + D\{B, C\}$
4. ...

A relation between these two measures is established as

$$S\{A, B\} = e^{-D\{A, B\}} \quad (C.2)$$

or in an inverse way (although not defined for  $S\{A, B\} = 0$ )

$$D\{A, B\} = -\ln (S\{A, B\}) \quad (C.3)$$

The *Similarity* is defined for two variables  $S(x, y)$  accordingly to each type of variable fulfilling the following properties (assuming that one of the variables is known and the other one is either known or given by its probability density distribution):

1.  $S(x, y) \in [0; 1]$
2.  $S(x, x) = 1$
3.  $S(x, z) \geq S(x, y) \cdot S(y, z)$
4. ...

### A. Similarity of two Boolean variables

Consider two Boolean variables  $x$  and  $y$ , the *Similarity*  $S(x, y)$  is defined as

$$S(x, y) = \begin{cases} 1 & \Leftarrow x = y \\ 0 & \Leftarrow x \neq y \end{cases} \quad x, y \in \{FALSE, TRUE\} \quad (C.4)$$

If one of the variables is defined as *unknown* then the *Similarity*  $S(x, y)$  is defined as the probability of the two variables being equal, i.e.

$$S(x, y) = P(x = y) = P(x = TRUE|x) \cdot P(y = TRUE|y) + P(x = FALSE|x) \cdot P(y = FALSE|y) \quad (C.5)$$

**Remark 1:** If the FALSE and TRUE results have the same probability for each unknown variable then the *Similarity* is equal to 0.5.

**Remark 2:** The definition given by (C.4) is included in the more general definition given by (C.5) since that if the values of the variables are known their conditional probabilities are either 1 or 0.

### B. Similarity of two discrete non numerable variables

Consider two discrete non numerable variables  $x$  and  $y$  the *Similarity*  $S(x, y)$  is defined as

$$S(x, y) = \begin{cases} 1 & \Leftarrow x = y \\ 0 & \Leftarrow x \neq y \end{cases} \quad x, y \in \{C_1, C_2, \dots, C_m\} \quad (C.6)$$

If some of the variables are defined as *unknown* then the *Similarity*  $S(x, y)$  is defined as the probability of the two variables being equal, i.e.

$$S(x, y) = P(x = y) = \sum_{i=1}^m P(x = C_i|x) \cdot P(y = C_i|y) \quad (C.7)$$

**Remark 3:** If the  $C_i$  results have the same probability for each unknown variable then the *Similarity* is equal to  $1/m$ .

**Remark 4:** Again the definition given by (C.6) is included in the more general definition given by (C.7) since that if the values of the variables are known their conditional probabilities are either 1 or 0.

**Remark 5:** Since the Boolean variable is a particular case of a discrete variable with only two possible results ( $C_1 = TRUE$  and  $C_2 = FALSE$ ) the definition given by (C.7) is general for the Boolean case also.

### C. Similarity of two discrete numerable variables

Consider two discrete numerable variables  $x$  and  $y$ , the *Distance*  $D(x, y)$  is defined as

$$D(x, y) = \gamma \frac{|x - y|}{N_m - N_1} \quad x, y \in \{N_1, N_2, \dots, N_m : i > j \Rightarrow N_i > N_j\} \quad (\text{C.8})$$

where  $\gamma$  is a scaling factor. The *Similarity*  $S(x, y)$  is defined from (C.8) as

$$S(x, y) = e^{-D(x, y)} \quad x, y \in \{N_1, N_2, \dots, N_m : i > j \Rightarrow N_i > N_j\} \quad (\text{C.9})$$

If one of the variables is defined as unknown with probability distribution  $p_y(y)$ , then the *Similarity*  $S(x, y)$  is defined as the average *Similarity* for all values in the set  $\{N_1, N_2, \dots, N_m : i > j \Rightarrow N_i > N_j\}$ :

$$S(x, y) = \sum_{i=1}^m p_y(N_i) \cdot e^{-D(x, N_i)} \quad y \text{ is unknown} \quad (\text{C.10})$$

**Remark 6:** Alternatively, the Boolean variable can be considered as a particular case of a discrete numerable variable with only two possible results ( $N_1 = FALSE$  and  $N_2 = TRUE$ ) resulting in a softer *Similarity* index than the one given by (C.4).

#### D. Similarity of two continuous variables

Consider two continuous variables  $x$  and  $y$ , the *Distance*  $D(x, y)$  is defined as

$$D(x, y) = \gamma \frac{|x - y|}{b - a} \quad x, y \in [a, b] \quad (\text{C.11})$$

where  $\gamma$  is a scaling factor. The *Similarity*  $S(x, y)$  is defined from (C.11) as

$$S(x, y) = e^{-D(x, y)} \quad x, y \in [a, b] \quad (\text{C.12})$$





## D. InLife System model and architecture

The main structure of the common repository is presented in Figure D.1 and briefly explained in Table D.1, Table D.2, Table D.3 and Table D.4.

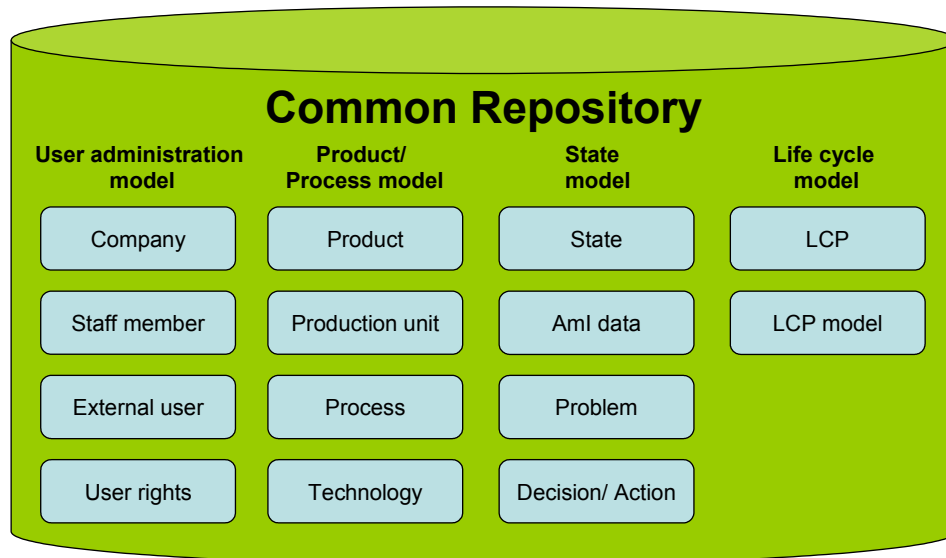


Figure D.1. Common repository

Table D.1. Characteristics/ description for User Administration Model

Object	Characteristics / Description
<b>Company</b>	Represents both the main company being modelled and companies connected to the enterprise (e.g. customers, suppliers, maintenance providers). Each one is specified by type of company and contact information and other relevant information.
<b>Staff Member</b>	Represents staff members of the main company. Staff members are defined by skills/technologies, responsibilities/competencies they hold and process steps they are involved in. They are further characterised by the company's department they belong to and according contact information.
<b>External User</b>	Models staff members of other companies connected to the main company (e.g. customers, suppliers, maintenance providers). Characteristics are equivalent to Staff Member.
<b>User Rights</b>	Models the level of access the user has to the InLife system. Rights are granted for reading, inserting and deleting data in the different modules of the system.

Table D.2. Characteristics/ description for Product/Process Model

Object	Characteristics / Description
<b>Product</b>	Models products in the scope of the enterprise. Products can be both outcome of production and assembly processes and input material. Products can be hierarchically structured and are further characterised by companies they are sold to or supplied by.
<b>Process</b>	Represents the steps in production and assembly processes. Process steps can

	be vertically and horizontally organised and are defined by input and output products as well as technologies/expertises applied in a step.
<b>Production Unit</b>	Machines/tools used in production and assembly processes to produce a product. Machines/ tools can be hierarchically organised and are connected to operational staff members. They are related to the technologies as well.
<b>Technology</b>	Represents methods/skills/expertises employed in the enterprise, which are known to staff members/external users.

Table D.3. Characteristics/ description for State Model

<b>Object</b>	<b>Characteristics / Description</b>
<b>State</b>	Represents the various (dynamically changing) variables defining the current status of products, processes and production units including LCP values (actual/predicted).
<b>Aml Data</b>	Models the context enriched information recorded by Aml devices and imported into the InLife system through the Aml Processing Module. (Raw data obtained by sensors and PLCs).
<b>Problem</b>	Describes actual and predicted problems occurring in the MAL. Problems are defined by different problem types, problem severities and products/production units and process steps they are related to.
<b>Decision / Action</b>	Represent decisions made and actions taken to resolve a given problem. Decisions and actions are further specified by the products/production units and process steps they affect as well as the amount of effort/costs that arise to perform them. Also, the degree of success in solving the given problem is modelled.

Table D.4. Characteristics/ description for Life cycle model

<b>Object</b>	<b>Characteristics / Description</b>
<b>Life Cycle Parameter</b>	Models the parameters needed for calculation of an LCP model. Parameters are defined by value and unit of measurement.
<b>Life Cycle Parameter Model (actual and predicted)</b>	Represents the aggregation of several Life Cycle Parameters into a Life Cycle Parameter Model (both actual and predicted) to compute the (general) status of the MAL by focusing on its life cycle impact. LCP models are thus connected to the products/production units/processes whose status they monitor. The models could be either for computation of the actual LCP values or for prediction of LCP values.

To provide a rough idea about the complexity of the developed model, Figure D.2 presents the entire UML class diagram.

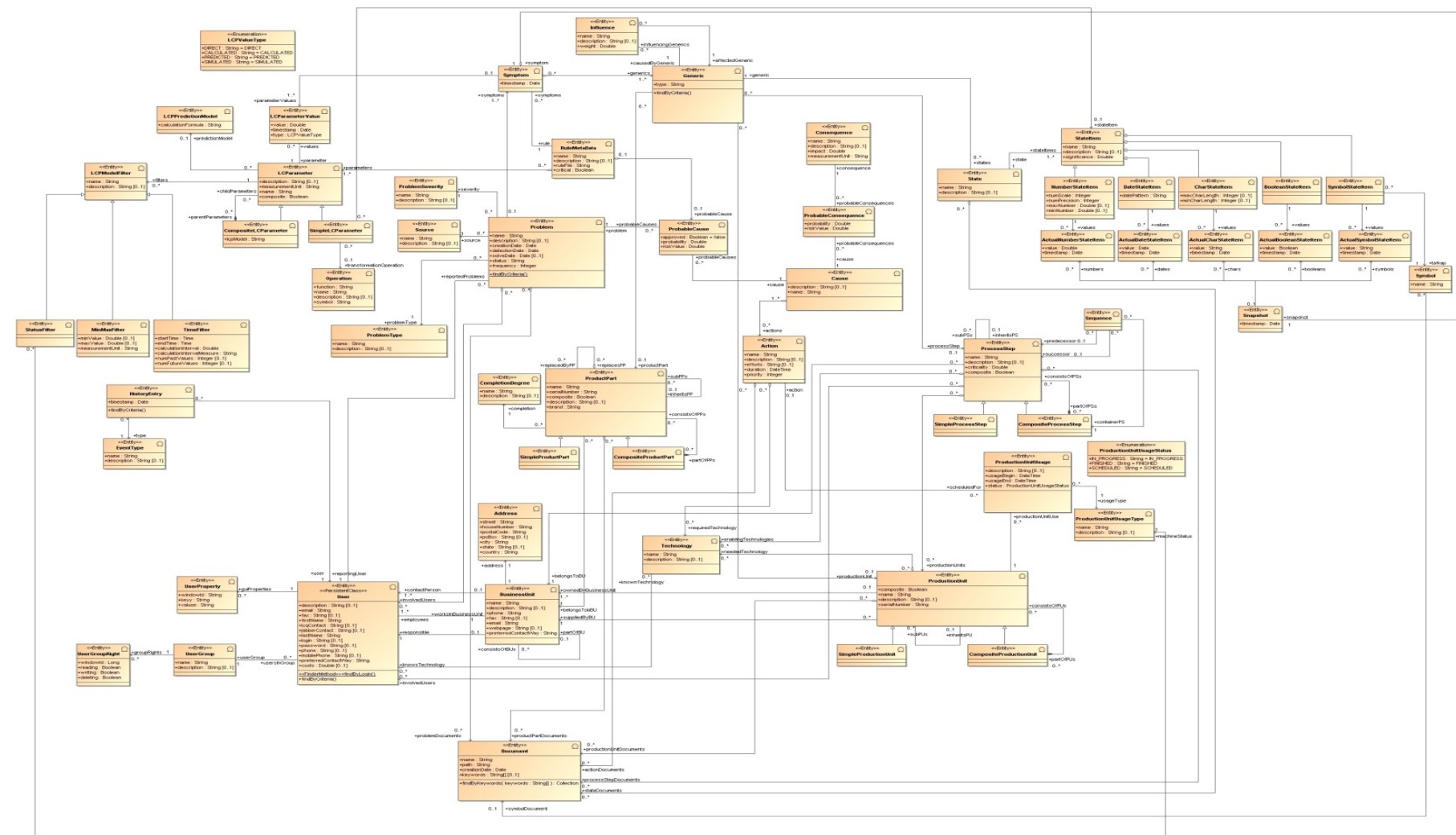


Figure D.2. . Common repository class diagram

The below Figure D.3 indicates the InLife system architecture.

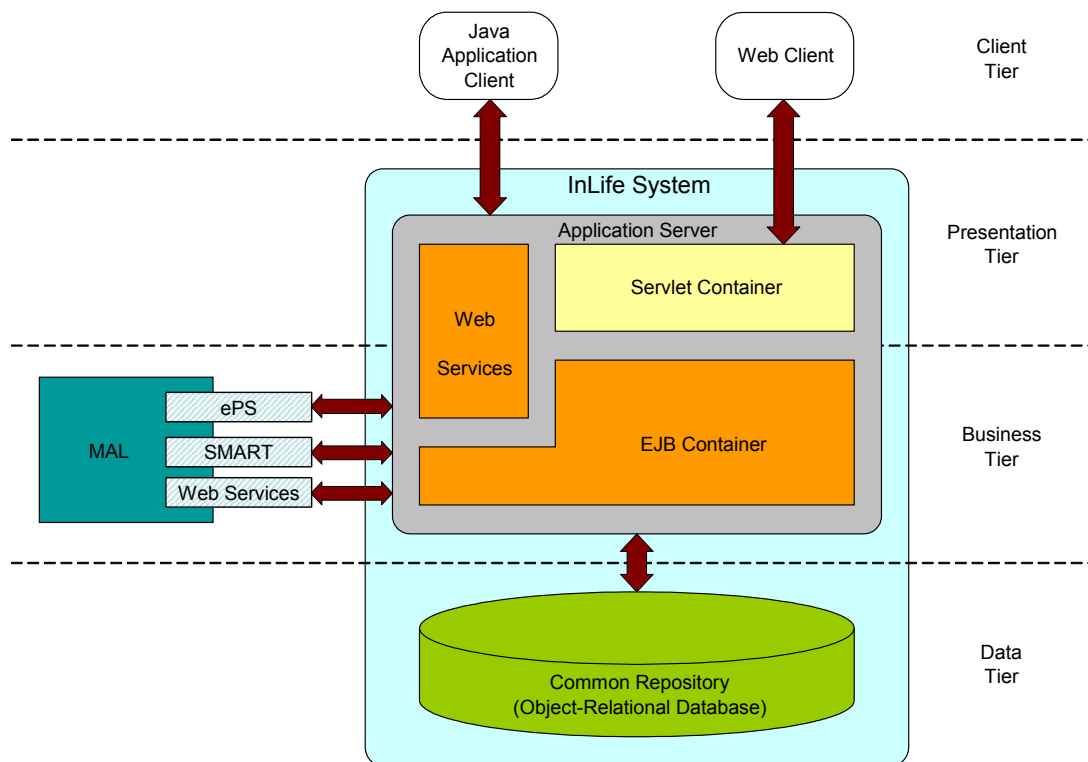


Figure D.3. InLife architecture

The InLife System was designed in common 4-tier architecture,

- Data tier;
- Business tier;
- Presentation tier; and,
- Client tier.

The data tier held all the static and dynamic data of the system. Common Repository was realised as an object-relational database.

The InLife System's whole business logic was encapsulated in the business tier, using Enterprise Java Beans (EJBs) as the implementing technology. Those EJBs were deployed in an EJB container, embedded in the application server. Furthermore, the InLife System exposed all of its functionality as WebServices to the outside world.

The connection between the MAL and the InLife system was realised using different technologies in namely the ePS system, the SMART agent platform and a combination of agent and WebService technology.

In the presentation tier the InLife system provides ways for a client to access the business logic's functionality. In particular, the InLife system offers two ways to do so: either through a java application client or an internet browser. While the java application client directly interacts with the InLife System, on the contrary access via internet browser requires a web server to serve static

web pages and a servlet engine to deliver web pages including dynamic content (Java Servlets, Java Server Pages) to the client.

For implementation, InLife relied primarily on open source software. Eclipse was used as the primary IDE. The backend database holding the common repository was realised in a MySQL database server while the EJBs and WebServices encapsulating the business logic were deployed in a JBoss application server. Also, the Tomcat servlet engine integrated in the JBoss application server provided necessary functionality to connect an internet browser based client to the InLife system.