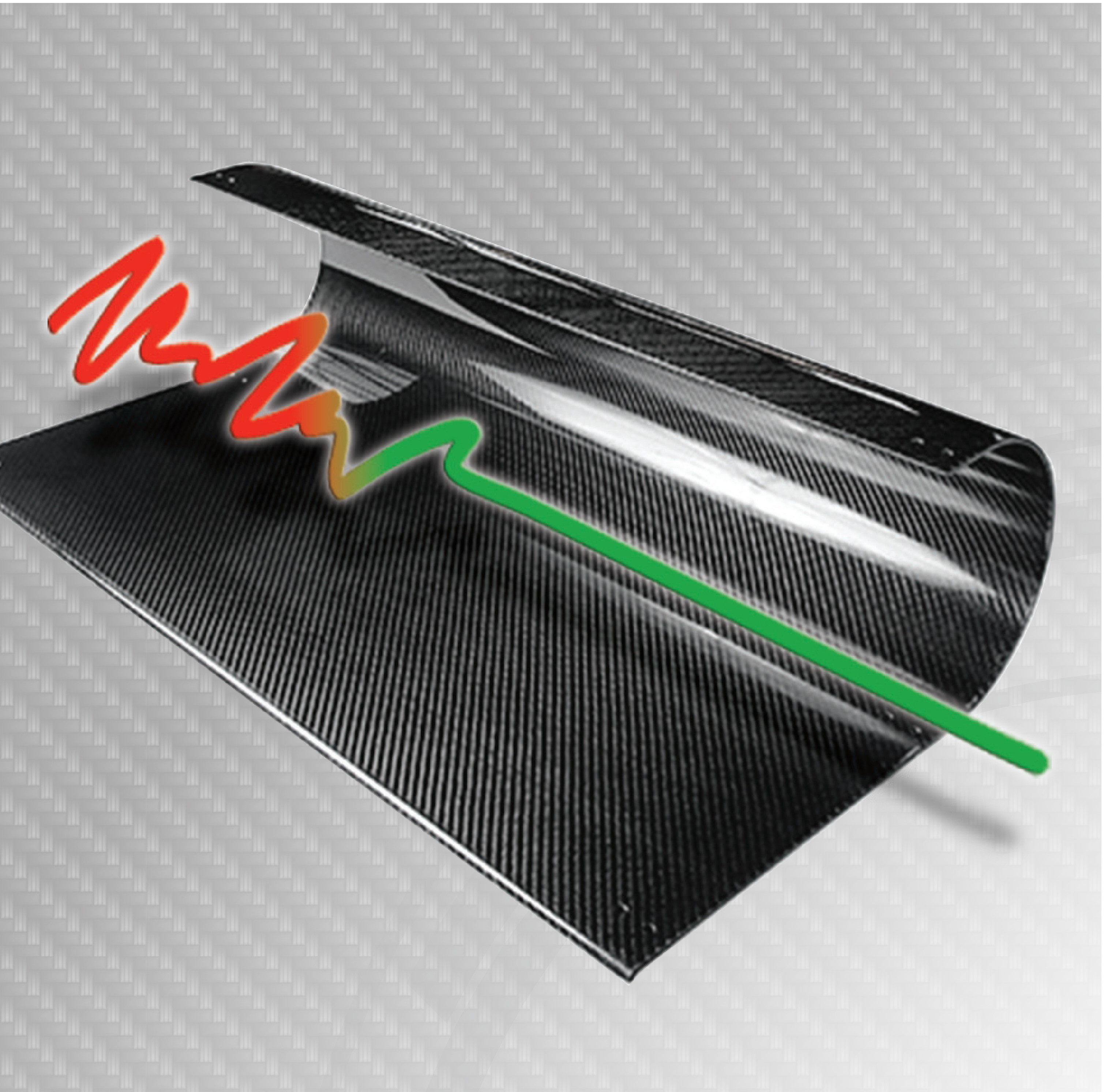
OPERATIONS INTELLIGENCE FOR COMPLEX RULES AND SYSTEMS

WHITEPAPER



CONTENTS

SUMMARY

The implementation of a complex process, as with the assembly of a complex system, requires full control of many competing action levers that are all vying for attention. In a context that is continuously changing and in the absence of a universal model that would take into account any operating variation identified in practice, it is extremely important to take advantage of prior experience by facilitating an empirical understanding of what must or must not be done according to the context and by clarifying the resulting return on experience.

DELMIA Operations Intelligence is the first software offering that provides this type of approach towards formalizing and capitalizing on expertise. The concept of explanatory rules – both simple and universal – is at the heart of the software and allows it to be used by all: with a minimum of training, it is possible to deploy this knowledge within the organization.

1. PROBLEMS AND CHALLENGES

In the context of continuously increasing pressure on costs, deadlines and quality, businesses in charge of implementing complex processes or assembling complex systems need to resolve – as rapidly and cheaply as possible – the multiple operational problems that they will face. Wherever this is possible, the teams responsible for the design of these systems and processes are working to refine their models and to anticipate any operational difficulties using simulation-based techniques on these models. The fact remains, however, that the implementation of these systems and processes often meets with real and unexpected problems that are difficult or impossible to foresee.

1.1. IMPLEMENTATION LIFECYCLE FOR PROCESSES OR SYSTEMS

Implementing a process or system is a challenge that is not limited to the industrialization phase of a product or the assembly phase of a system. In fact, the challenge is present throughout the life span of the system or process: during the initial design and experimentation phase, during industrialization, when being used or during preventive maintenance. Of course, return on experience is extremely useful during all these various phases for improving the processes and systems.

As soon as the design phase gets underway, it is possible to test a system or process by simulating it: control rules can be refined and a good range of settings and operating conditions can be defined – all perfectly attainable objectives at this stage.

During the industrialization phase of a process or the operational start phase of a system, industrialization will require numerous adjustments: this is when theory is put into practice and the model is compared to real results. This means that as much as possible needs to be learned about the processes and systems, and thus as rapidly and cheaply as possible (i.e. with a minimum of tests), so that stable operating conditions can be identified.

During the usage phase, conditions will change continuously, requiring operators to make adjustments to the process or system. Sometimes, these changes in conditions can suddenly lead to a blocking situation even though the changes themselves do not seem to be that different from those that have already occurred. At this stage, it is absolutely critical to be able to rapidly understand what is causing the difficulty and how it can be solved or how a workaround can be found.

When the system or process is being used, empirical knowledge is gathered with regard to successes or failures. The configurations that are deemed unsatisfactory, as well as the workarounds and maintenance actions that have enabled the problem to be successfully solved can thus be identified. It is, therefore, possible to preventatively detect risky situations and to take good corrective action. In addition, feedback to design will enable the reliability of the systems and processes to be improved for the next generation.

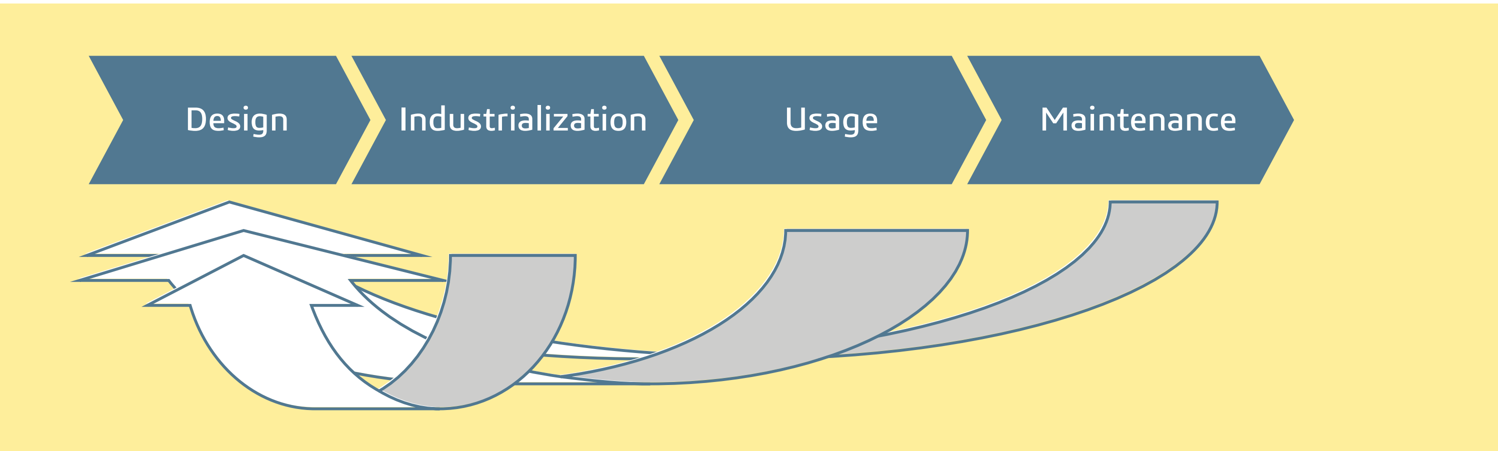


Figure 1: Lifecycle

1.2. DIFFICULTIES DURING IMPLEMENTATION

The theoretical difficulty of implementing a process or system is based on the highly combinatory nature of parameter interactions. For example, when a parameter changes value, there will be countless different setting possibilities (changes in the value of one or more parameters) that will allow compensation for the change. Generally, however, these various possibilities are not particularly simple because more often than not, a change to several parameters will be required. As a result, where a large number of contributing parameters exist and in order to respond to any change, there are a vast number of possibilities and basic setting combinations and it is impossible to explore them all.

Classically, a complex process or system can have many hundreds of parameters – some are influenced by the process or system, some can be controlled and others are measured as results.

Ideally, a good design should allow the operating quality of the process or system to be controlled deterministically. However, in reality, this is never completely true. In fact, it is impossible to guarantee that all the factors that influence a process or system operation have been identified and can be controlled. Thus, there is still a difference between the model and reality.

It is, therefore, extremely important to develop – as and when experiments are carried out – concrete know-how that will make up for the inadequacies in the model.

1.2.1. SYSTEM OR PROCESS RESPONSE

This can be represented diagrammatically as follows:

If we call:

- O (for "Output") the vector for exit parameters that measures the quality of the results or the performance of the system
- I (for "Input") the vector for identified entry parameters
- E (for "Environment") the vector for environment parameters
- S (for "Settings") the vector for the settings parameters
- U (for "Unknown") other non-identified variables (or those that are identified but not measured)

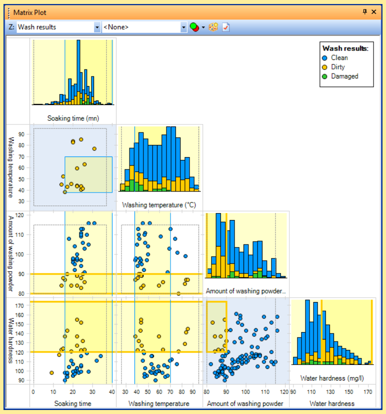
If the phenomena that underlie the functioning of the process or system are deterministic, there is a function f that (with a given parameter configuration) associates an exit value:

O = f (I, E, S, U)

In practice, because U is unknown and supposing that the influence of these variables is sufficiently weak, then we can assume that there is a function g that links O to I, E and S:

O = g (I, E, S)

This function g is more commonly known as a response function.



Visualization of Rules and Related Samples:
A graphical view enables users to understand how samples are distributed within and outside of a rule.

1.2.2. FROM DESIGN TO IMPLEMENTATION

In reality, even if the design has allowed for the estimation of a theoretical response function, real-life implementation or use always reveals significant differences between theory and reality because the theory cannot take into account all the actual constraints. **It is, therefore, necessary to complement the theoretical knowledge about the process or system operation with empirical knowledge** that allows the actual response function to be estimated – supposing of course that the function exists (i.e. supposing that the values I, E and S determine O deterministically).

Supposing that the objective in terms of quality or performance can be described in the form: Obj(O)> 0, where Obj is an objective (and possibly multi-criteria) function, the challenge of implementing the process or the system is as follows: identify a policy that allows the quality or the performance for the objective to be retained despite the variations in the included variables (I and E), by varying the controllable variables or settings (S) variables.

In practice, the solution that is found must also check a certain number of constraints, which can be expressed in the form of an inequality:

$$C(I, E, S) > 0$$

In other words, a function h needs to be constructed:

$$S = h(I, E)$$

Such as: $\forall I, E, \quad Obj(g(I, E, h(I, E))) > 0$ **and** $C(I, E, h(I, E)) > 0$

However, the ambition to precisely understand such a function h (i.e. a function that checks itself over time and experience) is virtually impossible: more often than not, there are many hundreds of parameters and the possibility scope is too complex. In actual fact, it comes down to compensating for this combinatory difficulty through a pragmatic approach. Because it is impossible to sufficiently explore the scope of the possible variations of the included parameters and because, despite this drawback, an exploration that is statistically sufficient enough to begin production or service operations cannot be expected to have been carried out, another form of knowledge and understanding will need to be constructed.

It means that faced with the impossibility of precisely modeling the actual behavior of the process or system, it is necessary to develop, as soon as experiments are carried out, concrete knowledge that will provide enough information to find out how to compensate for the variations in the included variables (entry and environment) through the controllable parameters.

1.3. THE ROLE OF EXPERTS

Traditionally, for solving this problem, businesses have relied on system and process experts who, from prior experience, understand the effect of the numerous parameters that they measure or control. In order to satisfy the required demands despite the continuous variations in the included variables, in general they know how to find the right combinations of actions to be carried out on the controllable parameters. They also know from experience which combinations are actually feasible.

The difficulty with this problem is the necessity to combine various different aspects:

- the capacity to identify solutions to these combinatory problems that involve many tens, even hundreds of variables simultaneously
- the understanding of what is feasible and what actually makes sense in a professional business environment
- the inclusion of the business's policy with regard to the implementation of the processes and systems

With this in mind, it is easy to understand why this work has been left to the experts up until now and why the classic software approaches that were able to solve the first point were not capable of being deployed through a lack of ability to deal with the second point.

2. THE CONCEPT OF EXPLANATORY RULES

In all phases of a process or system's life, it is very important to learn as much as possible from experience. This interaction between reality and the models must take place on a permanent basis – it incites the exchange of complementary expertise within the business and requires that each participant has a perfect understanding.

This is why **it is very important to make sure that a simple and universal modeling format for this knowledge is used. This will incite dialog** between the decision-maker, the design, development and industrialization engineers, the method engineer and the operator or those responsible for usage.

2.1. NOTHING MORE SIMPLE THAN A RULE

When an expert is requested to formalize his knowledge and understanding, he/she will automatically use rules. For example, "When the outside temperature is above 20°, when the acidity level of the raw material is between 3 and 4.5 and when the lubrication used is type A, then potentiometer P1 must be set to between 10 and 13 and potentiometer P2 must be set to 'medium' in order to guarantee that quality objectives will be met."

If an order needs to be transmitted to an operator, a rule is also the correct format. Likewise, when a decision-maker transmits his decisions, this will generally be done in the form of rules.

The advantages of this method of formalizing knowledge and understanding are:

- its simplicity
- its universal nature
- the ease with which it can be carried out

2.2. RULES THAT PROVIDE EXPLANATIONS

There are many different software technologies that use rules: expert systems, "business rules" engines, etc. These systems manipulate the rules that trigger actions or create information that can be re-incorporated into new rules, and so on.

Explanatory rules are much simpler: they formalize hypothetical combinations that provide an explanation for a result.

Explanatory rules are reports based on facts:

> ***"Experience shows us that each time such-and-such hypothesis were checked simultaneously, then such a result was obtained."***

However, explanatory rules are not limited to explaining each individual fact. In fact, they generalize, i.e. they identify the rationales behind a certain behavior in the system or process. Furthermore, a set of rules is not simply an accumulation of basic reports, but rather a description of the operation of the system based on facts.

User Interface: View shows a rule of Best Practice discovered in the data.

2.3. THE NEED FOR MULTIPLE LOCAL EXPLANATIONS

Moreover, explanatory rules can address the most complex needs of manufacturing teams who are challenged to look for a unique setting of their process that would be robust to any environment change or supply decision.

By offering process and systems experts a way to extract, optimize and validate an entire set of explanatory rules, one can explain all local behaviors of the system while preserving simplicity to keep the solution directly operable at low cost.

As rules can easily be implemented, it makes possible to create "agility" in the way processes or systems are piloted. This agility lies mainly in giving operational staff the technology to locate the best validated rules that apply in their current context, help them anticipate risk or performance issues and proactively decide process or systems settings changes.

Important benefits can be obtained through this approach, since adapting the manufacturing settings in a process of standard quality is generally cheaper than guaranteeing a high level of quality all over a big rule.

3. ALGORITHMIC PRINCIPLES

The use of explanatory rules in an industrial context, as we have already seen, is a subject for which DELMIA has developed a unique and innovative approach that ensures performance and efficiency.

3.1. SPECIFICATIONS

The specifications adopted by DELMIA include:

- The ability to handle a large number of parameters
- The ability to handle small amounts of experience data. It must be noted that it is difficult to produce both a valid explanation and the means to control the process or system using a very limited number of configuration observations. A small data sample signifies a miniscule amount of information in comparison to the size of scope that is possible and this, therefore, represents a major scientific challenge
- The ability to handle digital and symbolic data both continuously and discretely
- The ability to be robust to noise and missing data
- The ability to take into account the difference between external constraints, measurements and controllable variables and, more generally, the ability to take into account the specifics of the process (its structure, its temporal evolution, the direction it takes etc.)
- The ability to handle any kind of data distribution
- The ability to take into account dissymmetry and "local" phenomena

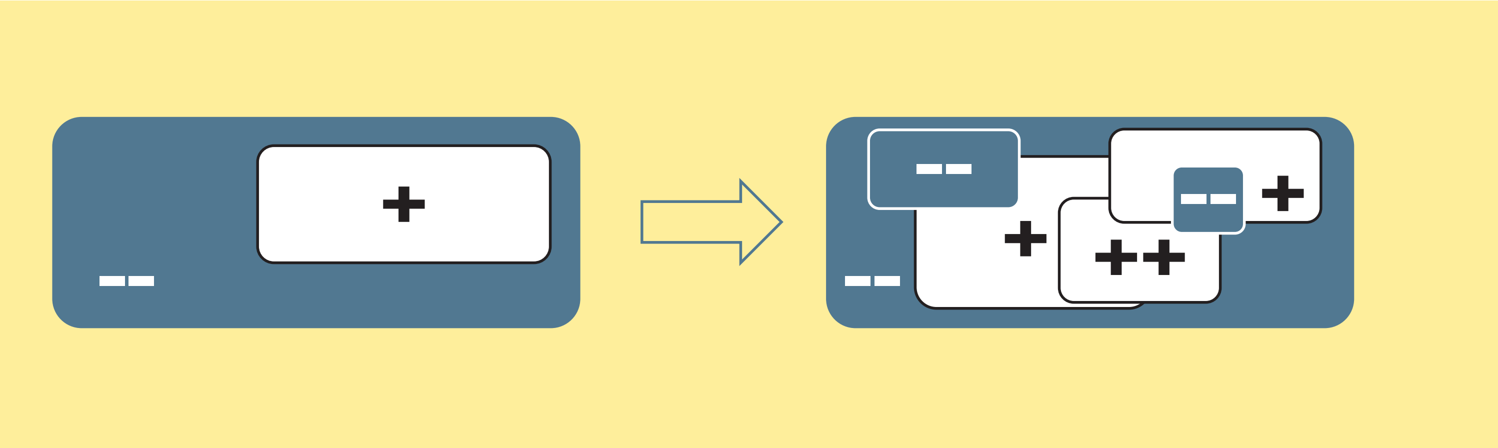


Figure 2: From robust to agile aproach

3.2. KNOWN SOLUTIONS

The challenge of using rules is a known problem about which a great deal of research has been carried out.

3.2.1. TYPOLOGY OF THE DATA ANALYSIS TECHNIQUES

More generally, among the classic approaches to data analysis, the use of rules is well established: it is seen as a data analysis technique based on the automatic production of an explicit and explanatory model. What does this mean? Among the known techniques, the implicit techniques (for example the case-based reasoning techniques) use the data itself to make predictions, whereas explicit techniques produce a model from the data (for example: a response surface, a decision tree, rules or a neural network) that can be used as a proxy instead of data itself.

Among the explicit techniques, some use a data representation model that does not allow any explanation (for example a neural network). This is known as a "black box" technique. On the other hand, those techniques that produce a model that can be easily understood are known as "white box" techniques or techniques based on explanatory models or knowledge models. This is the case particularly for models based on rules, decision trees or Bayesian networks.

In a complementary manner, there are so-called "statistical" methods that can also provide solutions to the problems being discussed here. Among these are: Regression, Principal Component Analysis and Partial Least Square techniques. These techniques highlight the parameters that seem to be the most influential and can be used as a complement to the use of rule-based techniques.

3.2.2. EXPLICIT EXPLANATORY MODELS

With regard to the possible explicit explanatory models, rules offer the advantage of combining power of expression with simplicity.

Decision trees are used on a large scale and have the advantage of being highly simplistic. By assimilating each "leaf" in the tree with a rule, decision trees can be reduced to specific rule systems. However, decision trees can hardly find very complex models (say, a reliable tree with more than 7 levels) because statistically a lot of samples are eliminated at each level. Also, decision trees are deterministic, therefore less hypotheses will be explored. Random forests can be used to introduce variability but in this case the model would be less understandable.

Bayesian networks allow chains of events to be modeled with cause and effect probability relationships. This means that "a priori" expertise can be taken into account and it is possible to interpret the results obtained through experience. However, the interpretation, as well as the implementation requires excellent knowledge of statistics and very good optimization skills. In addition, they have very limited levels of efficiency where there are many parameters and very little experience data.

On the other hand, models based on rules are extremely proficient at taking into account expert feedback, as they offer users a simple metaphor that favors interaction, and thus the appropriation of results without necessitating "a priori" scientific knowledge and understanding.

3.2.3. TECHNIQUES FOR USING RULES

Classically, there are various approaches for building rules:

- Descending approach ("top-down"): this is the classic approach for models based on "decision-making rules" inspired by decision trees. This approach consists of segmenting the scope iteratively into sub-scopes.

- Ascending approach ("bottom-up"): this is the approach DELMIA has adopted. Contrary to the "top-down" approach, this approach consists of using examples and attempting to generalize them.

The advantage of ascending approaches is their ability to detect and explain minority phenomena and dissymmetry, which descending techniques cannot do as easily. In fact, rather than using one single model to explain everything, this type of approach offers various models (rule set) that include different parameters.

3.3. THE LEARNING ENGINE

The rules created by DELMIA Operations Intelligence's learning engine are as follows:

If **Hypothesis** ***Then*** **Exit = Value**

where:

- **Hypothesis is a conjunction of basic hypotheses**
- **Each basic hypothesis represents the limits of the variable's variation field**
- **Exit is a discrete exit variable**
- **Value is a possible form for the Exit variable**

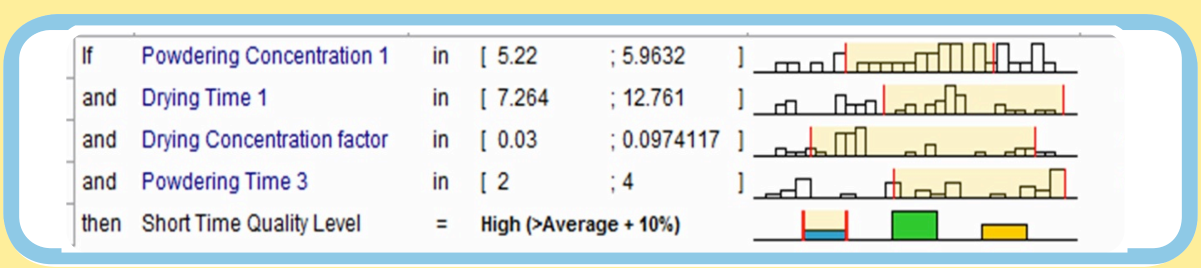
DELMIA Operations Intelligence's rule learning engine is based on a supervised stochastic learning algorithm that identifies zones containing identical experience data (micro-rules), known to be locally pure, and then extend, generalize and aggregate them according to selected quality criteria.

It is important to note that because the algorithm explores the combination of the example sets and not the parameter sets, its response time grows linearly with the number of parameters that are taken into account. In addition, this gives the algorithm some protection from irrelevant parameters.

3.3.1. RULE INDICATORS

The creation of a set of rules takes into account a group of optimization criteria that depend on rule quality indicators. Among these criteria there are:

- Purity: The uniformity of the behavior of a system or process within the conditions described by a rule is measured using the notion of rule purity. Supposing that a rule revolves around an exit variable that can take two values: good or bad and that the value good is targeted. The level of purity is calculated as the ratio of the number of experiences that satisfy the conditions of the rule as a hypothesis or in conclusion (thus judged as being good) over the number of experiences that satisfy the conditions of the rule as a hypothesis only (thus both good and bad). To illustrate this, a rule whose conclusion is "good" that only contains good experiences could be said to be 100% pure and a different rule that only contains 75% of good experiences could be said to have a purity level of 75%.
- Adjusted Purity: estimates the purity the rule would have if it was applied to new samples in conditions similar to those present when the dataset was created. The Adjusted purity is expressed as a percentage, the higher the better. The Adjusted purity is greater than the overall purity of the dataset but less than the Purity of the rule. This is because if a rule is based on too few examples and too many variables, it could actually be a coincidence.
- Relative Size: indicator that qualifies how exhaustive the explanation given by the rule is. The Relative size is the percentage of samples in the rule that match the class as the rule, compared to the total number of samples of this class in the dataset. For example, if the Relative size of a rule is "65%", the rule contains 65% of the samples of the dataset that match the output class of the rule. If the Relative size is "100%", the rule contains all the samples of its output class.
- Complexity: indicator that qualifies the number of variables included in the rule and, in particular, the number of uncontrollable variables that generally need to be kept to a minimum.



Example: this image shows a rule that has been created with DELMIA Operations Intelligence (Rules Discovery product)

3.3.2. RULE EDITION AND OPTIMIZATION

As a complement to the learning algorithm in a set of rules, it is possible to edit and optimize a specific rule according to the same criteria and taking into account certain constraints. The optimization is carried out by an algorithm based on "genetic algorithms".

This offers the "what-if" simulation and user dialog capabilities that are essential to the "acceptability" of the proposed rule, and therefore, to the success of their projects.

3.4. DELMIA'S UNIQUE TECHNOLOGY

In response to the specifications listed below, DELMIA's patented solution has no direct equivalent:

- It can be used by anyone – in particular by those with little or no data mining and analytics background
- It can take into account a user objective expressed in terms of a combination of quality indicators
- It is robust to noise and missing data and it can handle qualitative and quantitative data equally as well
- It allows the creation of rules that are directly operational, by favoring the use of controllable variables during the learning, editing and optimization phases
- It incites constant dialog with the users, thus making it possible to create rules that are both optimal and operational from an expert's point of view
- Through simple interaction, it provides a means to understand analyzed phenomena and to carry out simulation – to the point where it constitutes real support for decision-making
- It is suited to analyzing problems that include many hundreds of different parameters

4. VISION

By combining:

- Algorithmic power that integrates the best statistics, optimization and automatic learning
- The simplicity and ergonomic nature of a user interface that exploits the power and universality of the concept of explanatory rules as much as possible
- The integration of these functionalities into a business software solution that is both open to and conforms with current standards

DELMIA Operations Intelligence allows men and women who are responsible for the implementation and industrialization of complex systems and processes to better understand how they function, to accelerate their implementation and to improve their reliability.

The application of technology based on explanatory rules to complex systems and processes, allows the integration of quantitative approaches to continuous improvement and qualitative methods based on knowledge capitalization. By attaching an operating "map" (that takes into account favorable or risky situations for achieving good performance) based on explanatory rules to a given process or system, it is possible to offer the various people involved in the industrialization of these systems and processes a common vision that facilitates arbitration and decision-making. Because these decisions are based on fact, they are easy to execute and understandable. This in turn incites group adhesion more easily, and makes the decisions de facto more effective.

Delivering Best-in-Class Products

CATIA
Virtual Products

GEOVIA
Model and Simulate our Planet

SOLIDWORKS
3D Design

EXALEAD
Information Intelligence

SIMULIA
Realistic Simulation

NETVIBES
Dashboard Intelligence

DELMIA
Digital Manufacturing and Production

3DSWYM
Social Innovation

ENOVIA
Collaborative Innovation

3DVIA
3D Communication

About Dassault Systèmes

Dassault Systèmes, the **3D**EXPERIENCE Company, provides business and people with virtual universes to imagine sustainable innovations. Its world-leading solutions transform the way products are designed, produced, and supported. Dassault Systèmes' collaborative solutions foster social innovation, expanding possibilities for the virtual world to improve the real world. The group brings value to over 150,000 customers of all sizes, in all industries, in more than 80 countries. For more information, visit www.3ds.com.

Europe/Middle East/Africa
Dassault Systèmes
10, rue Marcel Dassault
CS 40501
78946 Vélizy-Villacoublay Cedex
France

Asia-Pacific
Dassault Systèmes
Pier City Shibaura Bldg 10F
3-18-1 Kaigan, Minato-Ku
Tokyo 108-002
Japan

Americas
Dassault Systèmes
175 Wyman Street
Waltham, Massachusetts
02451-1223
USA

Visit us at
3DS.COM/DELMIA

Item: 3DS.WP.TRI.2012x.0417