

New approaches for the use of the classical tools of scenario planning

Ricardo Baliero Fischer

Master in Business Management from Getúlio Vargas Foundation (FGV), Brazil

RESUMO

The future is to be built – is multiple and uncertain. Within the social sciences, scenarios can be defined as a description of a future situation and a course of events that allow move from a primary position toward this future situation. Currently, there is a multiplicity of methods and tools available for building scenarios, including methods of an essentially rationalist approach, as Michel Godet's method. The purpose of this work is to use the hypothetical-deductive method to reduce, starting from Michel Godet's Scenario Method and its tools, the complexity of the scenario-building process while maintaining the robustness of the findings. For this purpose, it is proposed two different approaches: (1) to integrate, in one step, the structural analysis and the cross-impact matrix so the first one derives automatically while filling the last one; (2) to use the concept of Bayesian networks as a method to integrate the cross-impact matrix and the morphological analysis. Both approaches aim to reduce the amount of information needed to feed the tools and improve the feedback criteria, resulting in greater flexibility during the process and better holistic view of the system. Scientifically, these approaches open a new field of studies in scenario planning as it appropriates the concept of Bayesian networks, widely used in other areas of knowledge (artificial intelligence, geological studies, medical diagnostics, pattern classification, etc.), and bring it to the field of social sciences.

KEY-WORDS: Scenarios. Structural analysis. Cross-impact analysis. Bayesian networks. Morphological analysis.

Novas Abordagens para o Uso das Ferramentas Clássicas de Planejamento de Cenários

ABSTRACT

O futuro está para ser construído – é múltiplo e incerto. Dentro das ciências sociais, cenários podem ser definidos como descrição de uma situação futura e de um curso de eventos que permita o movimento de uma posição original para essa situação futura. Atualmente, existe à disposição uma enormidade de métodos e ferramentas para construção de cenários, entre eles métodos de abordagem essencialmente racionalista, como o de Michel Godet. A proposta deste trabalho é utilizar o método hipotético-dedutivo para reduzir, a partir do Método de Cenários de Michel Godet e de suas ferramentas, a complexidade no processo de construção de cenários, mas ao mesmo tempo manter a robustez das conclusões. Para isso, foram propostas duas abordagens: (1) integrar em apenas uma etapa a análise estrutural e a análise de impactos cruzados, a primeira resultando automaticamente do preenchimento da última; (2) utilizar o conceito de redes *bayesianas* como forma de integrar a matriz de impactos cruzados e a análise morfológica. Ambas as abordagens visam reduzir a quantidade de informações necessárias para alimentar as ferramentas e melhoram o critério de *feedback*, resultando em maior agilidade no processo e melhor visão holística do sistema. Cientificamente, essas abordagens abrem um novo campo para estudos de planejamento de cenários já que se apropriam do conceito de redes *bayesianas*, muito utilizado em outras áreas do conhecimento (inteligência artificial, estudos geológicos, diagnósticos médicos, classificação de padrões, etc.) e o trazem para o campo das ciências sociais.

PALAVRAS-CHAVE: Cenários. Análise estrutural. Análise de impactos cruzados. Redes *bayesianas*. Análise Morfológica.

1 INTRODUCTION

The desire to know the future has existed since the beginning of humanity. In the pursuit of this knowledge, the man believed in anyone who could predict the future. It is possible to realize, in different moments in history, rulers in search of information that could minimize the risk of their decisions (Marcial & Markersbach, 2005).

But, after all, what is the future? The truth is that the future is multiple, uncertain and not written anywhere - it is to be built. If the man had the certainty of future events, they would lose their freedom and their purpose: the hope of a future you want (Godet & Roubelat, 1996). However, to deal with the future, we must accept that you are dealing with uncertainties. In the short term, projections of trends tend to work very well. However, in the medium to long term, the uncertainties increase and only the study of trends has not proved to be very effective. Assuming the acceptance of this inability of the man to predict the future, the studies of scenarios arose, which explored configurations of variables in order to create multiple possible futures. Schnaars (1987) compared results of econometric studies with studies that use methods of scenarios and ended up concluding that the method of scenarios has clear advantages over the traditional projective methods. Godet and Roubelat (1996) analyzed forecasts made years before and found that the errors were based primarily on two items: the overestimation of the impact of technological advances and the underestimation of inertial factors, such as the behaviors and the social structures.

Within the social sciences, scenarios can be defined as a description of a future situation and a course of events that allow to move from a unique situation to this future situation (Godet & Roubelat, 1996). Or even, to be able to view possible futures, particularly those derived and presented in systematic methods and those who define themselves by the holistic view of the circumstances in question (Miles, 2005).

The planning for scenarios (or prospective) is more than a method, it is a process. Its major characteristic is the need to be done in a participatory manner in which the main role in the process is the

responsibility of experts, the only ones able to handle practical and theoretical knowledge and use their sensitivity to develop coherent visions of the future (Marques, 1988). When assembling this simplified model of reality by means of structured information from experts, in the prospective it is not despised the other forms of analysis. On the contrary, it is tried to use all types of analysis so they can contain variables of quantity and quality.

Currently, there is a plethora of methods and tools to build scenarios. There are methods of intuitive approach, such as that of Pierre Wack (1985Rd; 1985b) and his disciples Peter Schwartz (2000) and Krugman (2004), as there are methods of approach essentially rationalist, as Michel Godet (1993). Also there are many appropriate tools for these methods, as the morphological analysis, more intuitive characteristic, and the countless variations of use of the cross-impact matrix, which has a more quantitative approach.

The existence of such a large variety of methods and tools can be explained by taking into account that, due to the nature fundamentally deterministic of the human being, it is always complicated to give credit to methods that claim to exist in several possible answers. Therefore, it ends up with a relentless pursuit for more robustness, in the process and results, in each new method presented, with the aim of giving it more credibility. Many times, in the search for this robustness in a same method several tools of distinct characteristics may be present, which can raise the complexity of the analysis. Added to this is the fact that there is available in the world wealth and incomplete information (Godet & Roubelat, 1996), but decisions must be made quickly. For all these reasons, methods of construction of scenarios should have a selective character. Starting a job of scenarios constructions with a focus very close to the model greatly increases the chances of losing points-key in the analysis of the determinants of future. To work with scenarios, it is necessary to have the vision of macro model for after getting into the details (Duncan & Wack, 1994). Therefore, the tools and the methods used must deal with the minimum amount of information possible and be simple enough so that the

results can be easily assimilated by decision-makers. At the same time, they should confer robustness and credibility to the results achieved.

Within this context, the proposal of this work, taking as conceptual basis the structure of the method of scenarios proposed by Michel Godet due to being the most structured and rigorous among the methods most known, is to present approaches that allow to reduce the complexity in the process of building scenarios while maintaining the robustness of the analysis and the results.

In this work, an exploratory research in which you use the hypothetical-deductive method with a methodological approach, the processes and tools contained in steps of the structural analysis, morphological analysis and method of specialists and probabilization are described in a comprehensive manner with the aim to propose two different innovative approaches:

1. Initiating the process with the completion of the cross-impact matrix and automatically extract from it a structural analysis. In this way, it is obtained, in only one analysis, two important pieces of information for the process: the hierarchy of variables and the mapping of the probabilistic space of the set of relationships between the variables;
2. To develop the concept of *Bayesian* networks as an assessment criterion in a morphological analysis to create a morphological subspace quantified, using an array of probabilistic crossed impacts for data feeding and processing.

For the analysis of the effectiveness of these approaches the following criteria were used:

- The theoretical foundations provide reliability and credibility to the method developed;
- The method does not give margins to interpretations, being clear about the objectives and the roles of each step;
- There is easiness to conduct the feedback in the process;
- The amount of information needed to feed the tools is reduced.

This research becomes especially relevant because, in relation to the original method of Michel Godet, proposes to reduce the amount of information needed to feed the tools and greater integration of stages of the method, improving the feedback process. Managerially, this means that decision-makers will have, with this approach, a method that allows greater flexibility in analysis and best holistic view of the system, maintaining the level of robustness of the analysis and credibility of results. Scientifically, they are innovative approaches which open a new field for studies of scenarios, since they make use of the Bayesian concept, widely used in other fields of knowledge (artificial intelligence, geological studies, medical diagnosis, classification of patterns, etc.), and bring it to the field of social sciences.

2. BIBLIOGRAPHIC REVIEW

2.1 SCENARIOS METHOD

As already exposed earlier, among the most celebrated methods, which present the most robust structure, rationalist and less dependent on the intuition is the one used by Michel Godet. In its methodology, Michel Godet (1993) integrates the method of construction of scenarios the two tasks: the diagnosis of the organization and support for strategic choices. To Michel Godet, the objectives of the method of scenarios are three:

- Identification of key variables that characterize the system studied, establishing relations between them by exhaustive analysis;
- Determination of key actors, their strategies and the means available to achieve their goals;
- Description of the evolution of the system studied taking into account the changes more likely of key variables when related to each other and influenced by the games of the actors.

The method of Godet (Figure 1) is characterized by being the most robust and less flexible, because it believes in the accuracy of using the analysis tools. It is one of the few methods that uses the concept of subjectivity probabilities to analyze the scenarios developed.

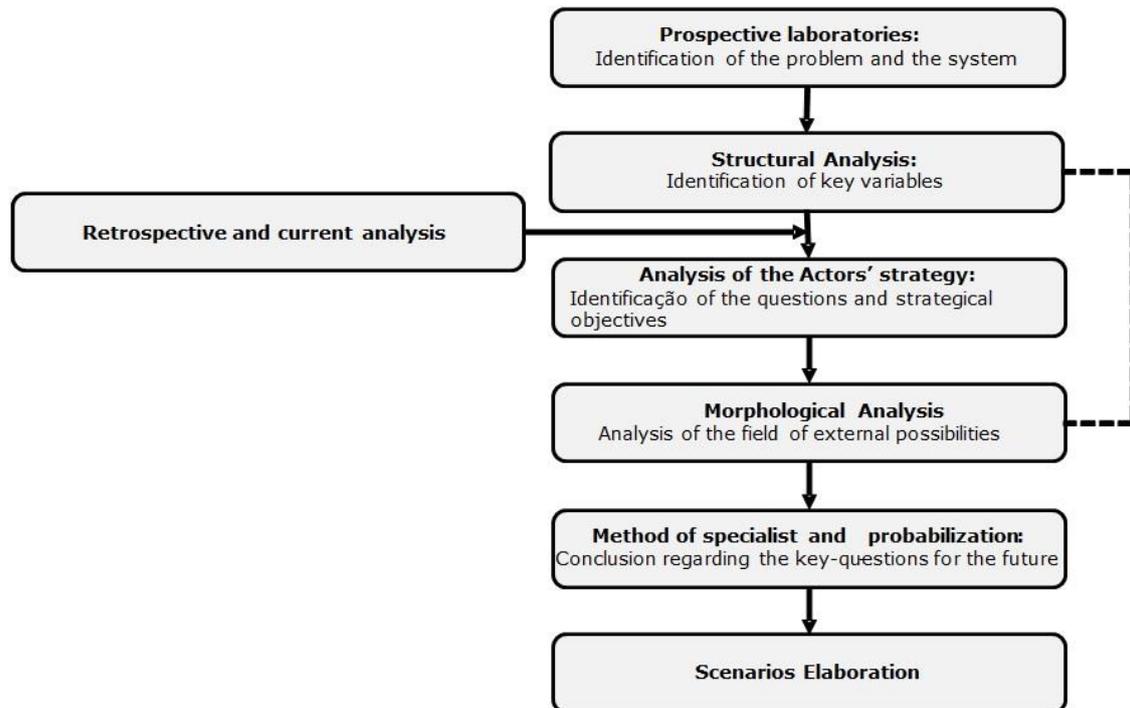


Figure 1: Method of Scenarios of Michel Godet

Source: Adapted from Godet (1993)

The first stage, in which occurs the delimitation of the system and the environment, it serves to specify the scope of the study. In it, it is defined the object of study, the time horizon and the geographic area, i.e., the focus of the study. Usually it is a concern of the company.

The second stage begins with the preparation of a comprehensive list of variables that somehow can explain the behavior of the system composed. Then there is the structural analysis (MICMAC), which allows the classification of variables in relation to the number of fundamental parameters to characterize their role in the system, putting in evidence a hierarchy of variables and facilitating the identification of key variables.

In the next step, the games of actors are analyzed. The examination of their relations of force is essential to highlight what is the evolution of the strategic challenges and to put the key issues for the future, since the actors will tend to manipulate the variables in accordance with its strategic objectives. These actors and their strategic objectives are confronted in their convergences and divergences in an array called MACTOR, whose

output is the identification of the most influential actors in the variables of the system.

At the subsequent step, it is applied the morphological analysis and it has the beginning of the exploration of the field of possible developments. It is begun the process by bringing together the key variables of structural analysis. Following are defined the constraints between events for space reduction morphology. In the end, it is explored the combination of these settings using the key issues of the game of actors. In the method of scenarios, it is used the tool called Morphol.

The next step is the use of a method of cross-impacts of binary character (tool SMIC-Prob-Expert), in which is entered a smaller number of hypotheses, already filtered from the morphological analysis. This phase is based in consultation with a range of experts, for which they are proposed issues regarding the simple probabilities (*a priori*) of an event and the conditioned probabilities (*a posteriori*) of a certain event to take place in relation to any other. Then it is performed an analysis of sensitivity and robustness, which can result in a feedback for the beginning of the process.

In the last step, finally, the elaboration of scenarios occurs. This step is performed in two main phases: improvement of the final images of scenarios within the timeframe of the study and construction of a narrative that corresponds to the final images drawn.

2.1 SCENARIOS TOOLS

2.2.1 Structural analysis

The structural analysis, according to Arcade, Godet, Meunier and Roubelat (1994), is one of the most frequently used tools in studies of the future and has two complementary objectives (Godet, 1993): during the initial phase, to obtain the best possible representation of the system under study; and to reduce the complexity of the system by choosing the main variables.

The structural analysis, a tool of structuring of collective reflection, provides the possibility to describe a system with the aid of a matrix that

lists all of the components of this system (Godet, Monti, Meunier & Roubelat, 1999). The objective of this method is to make emerge the main variables influent and dependent, thus determining the variables essential for the evolution of the system. It has as main stages: the identification of variables, the description of the relations between the variables and the identification of key variables.

In the first stage, it is listed the set of variables that characterize the system studied. At this stage it should be the most exhaustive possible and not exclude, *a priori*, any form of research. At this stage it is desirable to feed the collection of variables by means of not directive interviews with representatives of actors of the system being studied. It is obtained, finally, a list of internal variables and external to the system under consideration.

In the second stage, it is sought to identify the existing relations between the variables using an array of structural analysis. The filling is qualitative, being questioned for each pair of variables if there is a direct influence of the variable i at the variable j . In response to this question, you can fill in this matrix in two ways: *Boolean* (with 1 if there is a relationship and 0 in case it does not occur) or quantifying the relationships (for example, defining the relations as: non-existent=0, weak=1, average=2, strong=3). In Figure 2, it is illustrated the process of populating a *Boolean* array with data generated at random.

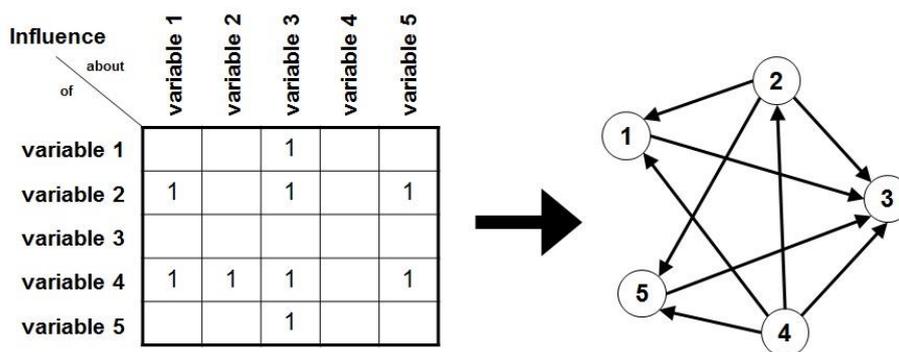


Figure 2: Structural matrix of direct relations and the adjacent graph

Source: Arcade et al. (1994)

This phase of filling helps to put, for n variables, $n.(n-1)$ issues, some of which would have been omitted if there had not been a reflection so systematic and comprehensive. The visual result of the structural matrix

graphs whose nodes (or vertices) are the variables; and the edges, their inter-relationships nurtured on the matrix of structural analysis (Arcade et al., 1994).

In the next step, the graph resulting from the previous step can be represented in order to contribute to decipher the inter-relationships of the system and, if possible, to classify the variables in successive levels of spread of influences (Arcade et al., 1994), as shown in Figure 3.

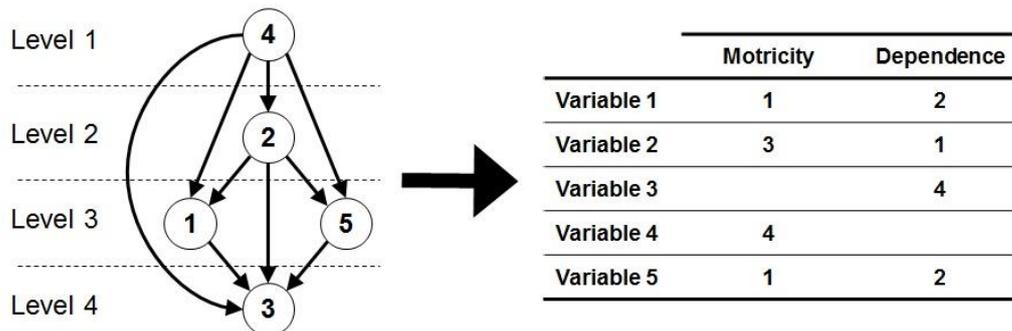


Figure 3: Graduated graph and its motricities/Dependence

Source: Adapted from Arcade et al. (1994)

In a very intuitive approach, the direct influence (motricity) of a variable can be assessed being considered the lines in the matrix structure. A variable acting on a smaller number of variables exerts its influence in a limited part of the system. Also, the direct dependence of a variable is obtained considering the columns of the matrix, i.e., the sum of direct influences exercised in that variable. Therefore, adding systematically all the elements of each row and column of the matrix structure, it is possible for each variable have the measure of the motricity and dependence for the system as a whole (Arcade et al., 1994).

In the system used as an example, it is sufficient to carry the analysis only taking as information their direct relationships, since the system is acyclic, i.e. has no circuits. However, in more extensive and complex analyses, it is necessary to create algorithms for, when the system presents circuits, consider the relationship between indirect variables. The MICMAC, tool of the Method of scenarios of Godet, is the most well-known, but has some limitations (Perestrelo & Caldas, 1998): indetermination of drivability and indirect dependence, overvaluation of the feedback, lack of

stability, multiplicative effects, separation between direct and indirect effects. To solve these problems of MICMAC, there are alternative methods, such as the Maximum flow and spread of effects.

2.2.2 Method of crossed impacts

Method of crossed impacts is the generic name for a family of techniques that try to assess changes in the probability of occurrence of a certain event when related to the previous occurrence of some other. The method begins with a list of events and their associated probabilities. The basic hypothesis of the method is that many times at the individual probabilities it is already considered the interactions among the events, but only in incomplete form. Take into account the interdependencies allows having a system of initial probabilities not processed for a set of "net" probabilities, i.e., corrected. The remains of the method consist of examining the sensitivity of the system in the construction of scenarios, highlighting the more likely final image (Godet, 1993).

Regardless of the type of approach, the method of crossed impacts has a well-defined methodological basis. It is a method of analytical approach, applied to the probabilities of an item within a system in anticipation. Its probabilities may be adjusted as a result of decisions concerning the possible interactions between the items. It is known by experience, that most of the events are in any way related to other events. From this interconnected flow, the events occur increasingly widely while interacting with other events, forming a large network of interconnections. It is this inter-relationship between events that is called "crossed impacts" (Gordon, 1994).

The first step to an analysis of crossed impacts is to define the events to be included in the study. This first step can be crucial to the success of the exercise. Any event not included in the system will, of course, entirely be excluded from the study. However, the inclusion of not relevant events may complicate unnecessarily the analysis.

After defining the set of events, the next step is to estimate the probability of each event, which indicates how each one of them might

behave for some years to come. Next follows the next step in the analysis, which is to estimate the probability conditional of each pair of events. Typically, the impacts are estimated in answer to the question: "If the event V_j occurs, what is the probability of occurrence of the event V_i ?" Thus, if the probability of the event V_i was originally judged in $P(V_i)$, and if it is considered that V_j occurred, a new probability $P(V_i/V_j)$ can be attributed to the event V_i . Every array of impacted cross is filled by placing this issue to each pair of events in relation to the occurrence and non-occurrence of events (Figure 4).

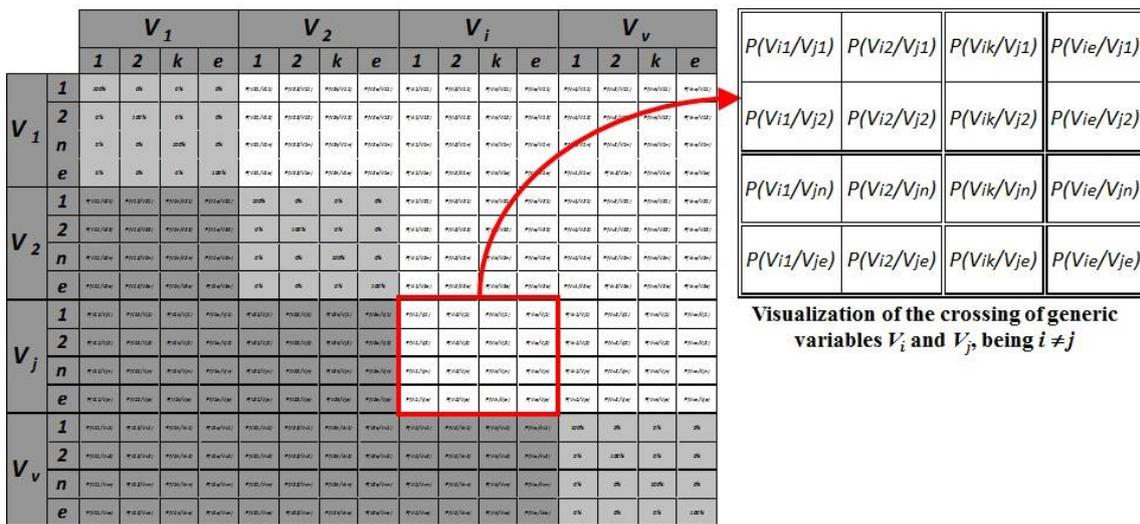


Figure 4: Matrix of generic crossed impacts

Source: Elaborated by the author

When the initial probabilities are estimated with reference to other events, some additional information enters in the estimate of the system. For each combined event, there are limits so that the conditional probabilities may exist (Gordon, 1994). The main issue is that, even if an expert could answer with probability simple conditionals for several pairs of events, it is practically impossible for their responses to meet the classical axioms governing the Probability Theory, as the rule of the sum and the rule of the product (Godet, 1993).

From that point, the results obtained depend on the procedures and the transitional formulae to calculate the final probability. It is possible, for example, that the participants of the analysis to decide that a judgment

must be changed, while observing the dynamics of the system. This learning process, which occurs while the array of crossed impacts is being estimated, is one of the main benefits of implementing this approach (Gordon, 1994).

2.2.3 Morphological analysis

A análise morfológica (morphological analysis – MA), the oldest technique of structuring of uncertainty for use in scenarios, was born and used for decades primarily in areas totally different from science. According to Ritchey (2002), developing future scenarios presents a series of methodological difficulties, as to quantify factors that contain strong socio political dimensions and games between actors. In this context, the morphological analysis shows itself as an alternative to mathematical methods of formal and causal modeling: a not quantified way of modeling, based on processes of judgment and internal coherence. The causal modeling can be used as an aid to judgment.

MA is a sophisticated title to a simple method that proved very useful to stimulate the imagination, helping to identify products and processes up to then unknown and to explore the field of possible future scenarios (Godet, 1993). The basic principle is the decomposition of the system studied in sub-systems or components, which should be the as independent as possible and cover the whole of the system being studied. This forms the so-called morphological space (Figure 5).

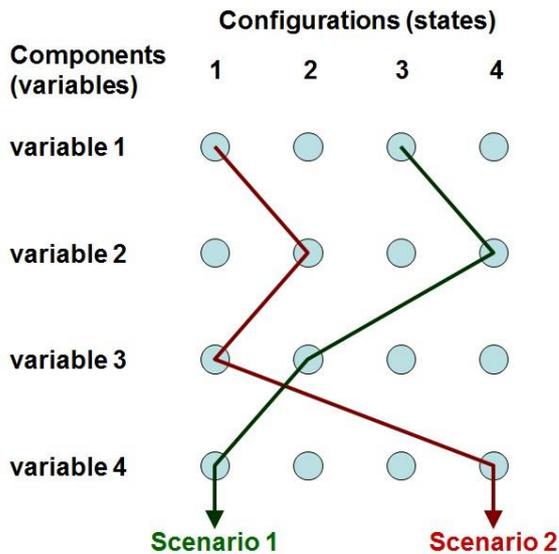


Figure 5: Morphological space

Source: Elaborated by the author

Finally, the next step in the process of analysis-synthesis is to reduce the number of possible configurations of a morphological space to a smaller set of settings internally consistent representing a "solution space", or a morphological subspace.

2.2.4 Bayesian networks

The graphs have proved to be a very intuitive language to represent states of dependence and independence, and therefore provides a great way to communicate and discuss these dependencies and independence between variables within the domain of a given problem. According to Pearl (2000), the roles of the graphs in probabilistic modeling are: to provide convenient means of expressing important assignments, facilitate economic representations of functions of joint probabilities and facilitate efficient inferences generated by observations.

For working so well with the visualization of dependencies in a given system of variables, the concept is an excellent tool to represent causal networks. According to Kjaerulff and Madsen (2005), the concept of causality plays an important role in the process of building models of probabilistic networks. To contextualize this link, this work uses the

approach of *pensamento causal* (or *raciocínio causal*, in English *causal reasoning*) and not "*causação*" (in English *causation*). This means that the statement of the model would be: "The event *A* causes for sure x the event *B*". Based on this, the causal thinking would be: "If it is known that *A* happened, then *B* it can happen with certain x ".

As redes bayesianas (*Bayesian networks* – BN) are one of the most popular methods of reasoning in probabilistic networks among practitioners of Artificial Intelligence. This name was coined by Judea Pearl in 1985, but the ideas and concepts came from many sources. There are several applications of this method (medical diagnosis, learning maps, understanding of the language, vision on machines, heuristics search, etc.), but, in spite of the apparent importance, the ideas and techniques have not been disseminated far beyond the borders of the Community responsible for them (Charniak, 1991).

In the universe of the BN, the graph representing the causal relations among the variables is the directed graph Acyclic (*Directed Acyclic Graph* – DAG). Its main characteristic is the absence of arches that represent feedback in the model. DAG is a class of graphs that can represent a very compressed, large and important set of assignments of relations of dependence and independence expressed in representations of factored joint probability distributions teams (Kjærulff & Madsen, 2005).

The BN offers an approach to the reasoning probabilistic that includes on the one hand, the graph theory for the establishment of causal relations between sentences (qualitative part) and, on the other, the theory of probabilities, for allocation of levels of reliability (quantitative part). The best way to understand BN is to imagine the modeling of a situation in which the link has important role, but there is complete understanding of what is happening. In other words, the uncertainty is inherent to the model and it is necessary to explain it using probabilities. Pearl (2000) provides three aspects that should be emphasized in BN: the subjective nature of the information that feeds the model; confidence in rule of Bayes as the basis for updating the information; the distinction between forms of causal reasoning and "evidencial". These aspects demonstrate that not only the construction of a model graph in *Bayesian* network is qualitative, but also

not necessarily the probabilities must be based on frequencies. Probability can also be estimated completely subjective, on the expectation about an event. This in no way invalidates the use of the rules of calculation of probabilities (Jensen, 1996).

In BN, the causal relations among the variables are expressed in the form of conditional probabilities. That is, given an event A , it has a probability $P(A)$ of the occurrence. However, it is known that A is conditioned upon the occurrence of B , then the probability of A occurring, given that B occurred is $P(A/B)$. In other words, the probability of occurrence of A is changed when knowing new data. Within the theory of probabilities, the most important axiom for BN is called basic rule, rule of multiplication or rule of Bayes Theorem: to two events A and B , the probability that both of them occurs $P(A \cap B) = P(A, B) = P(A/B) \cdot P(B) = P(B/A) \cdot P(A)$. This axiom explains that the probability of occurrence of two events simultaneously can be obtained by multiplying the probability of occurrence of one of the events by the probability conditional of occurrence of another event if the first happens. The result $P(A, B)$ is known as a joint probability of A and B .

The biggest problem in networks and effect lies in the fact that there is a lot of information to be treated. For example, a complete DAG containing v variables, each variable containing e states will have $v \cdot e^2 \cdot (v-1)/2$ probability conditional pair to pair and e^v values possible in the distribution of joint probabilities. Performing an exact inference consists of knowing all these values in *Bayesian* discrete network. There are also various methods (e.g.: Monte Carlo) for carrying out the approximate inference, which should be used when there is lack of information about several variables in the system studied (Ben-Gal, 2007).

There is a way to simplify the calculations in *Bayesian* network, reducing the number of information only to relevant information, based on how information propagates in the network. This criterion is called *separação-d* (*d-separation*) and defines the relationship of dependence or independence of any pair of variables in a causal network, given that a new evidence entered in the network (Jensen, 1996). One of the advantages of *Bayesians* networks is that they admit *d-separation* if A and B are *d-*

separation when an evidence ε enters, then $P(A/B, \varepsilon) = P(A/\varepsilon)$. This means that you can use d-separation to find conditional independence.

Whereas the concept of conditional independence, it is possible to express the distribution of the joint probability of a *Bayesian network* through multiplication of probabilities of conditional convergence connections and independent on the network for example, for a complete DAG representing a system of four variables (Figure 6), the joint probability would be $P(V_1, V_2, V_3, V_4) = P(V_1) \cdot P(V_2/V_1) \cdot P(V_3/V_2, V_1) \cdot P(V_4/V_3, V_2, V_1)$.

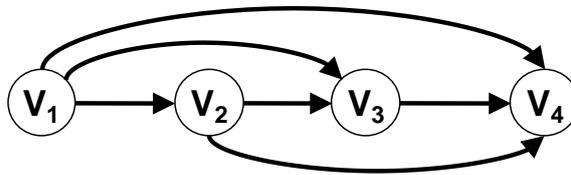


Figure 6: Graphic representation of a *Bayesian* network of four variables

Source: Elaborated by the author

Generalizing for n , it is reached:

$$P(V_1, \dots, V_n) = P(V_n/V_{n-1}, \dots, V_1) \cdot P(V_{n-1}/V_{n-2}, \dots, V_1) \dots P(V_2/V_1) \cdot P(V_1)$$

[1]

or simply:

$$P(V_1, \dots, V_n) = \prod_i^n P(V_i/Pa(V_i)) \quad [2]$$

This calculation of the probability is called rule of chain, with $Pa(V_i)$ representing the predecessors of V_i . One of the issues in BN is how to perform probabilistic inference for $P(V_i/Pa(V_i))$. One solution is to simply assume that the inputs are conditionally independent of each other, simplifying the approach of the problem without significantly compromising the precision of the result (Lacerda & District, 2004). This approach is known as rule *naïve* of Bayes. It is called *naïve* (*algo como ingênuo*) precisely because he considers the assembly (V_{i-1}, \dots, V_1) mutually independent given the variable V_i . The model *Naïve* of Bayes is a special

form of the *Bayesian* network used a lot in classifications and groupings (*clustering*), but little explored in probabilistic modeling (Lowd & Domingos, 2005). Developing the concept, it is originated in $P(V_n/Pa(V_n)) = P(V_n/V_{n-1}, \dots, V_1)$, it is applied the rule of Bayes and it is considered that the variables are conditionally independent:

$$P(V_n/V_{n-1}, \dots, V_1) = \frac{P(V_{n-1}, \dots, V_1/V_n) \cdot P(V_n)}{P(V_{n-1}, \dots, V_1)} \quad [3]$$

or

$$P(V_n/V_{n-1}, \dots, V_1) = \frac{P(V_{n-1}/V_n) \cdot \dots \cdot P(V_1/V_n) \cdot P(V_n)}{P(V_{n-1}, \dots, V_1)} \quad [4]$$

or simply:

$$P(V_n/V_{n-1}, \dots, V_1) = \frac{P(V_n)}{P(Pa(V_n))} \cdot \prod_{j=1}^{n-1} P(V_j/V_n) \quad [5]$$

Although somewhat unrealistic (that is why it is called *naïve*), this assumption works remarkably well in practice. *Naïve* Bayes has already proved its effectiveness in many practical applications, including classification of text, medical diagnosis and management of performance of systems (Rish, 2001).

According to the rule *naïve* of Bayes shown above, each product of rule of the chain is factorized to only probabilities *a priori* and conditional pair to pair. In other words, it is possible to calculate the probability of a *Bayesian* network simply conditioning probabilistically the pairs of events (the arches of the DAG). Like this, the rule of the chain can be rewritten, the following way, being that for $i = j$, $P(V_j/V_i) = 1$ and for V_j and V_i independent, $P(V_j/V_i) = P(V_j)$:

$$P(V_1, \dots, V_n) = \prod_{i=1}^n \prod_{j=1}^i \frac{P(V_j)}{P(Pa(V_i))} \cdot P(V_j/V_i) \quad [6]$$

3 PRESENTATION OF THE PROPOSALS

The first step of any causal modeling is the choice of the set of variables to be studied. Normally, the process is begun at the construction of the model structure using the structural analysis that, most likely, will have the form of a cyclic graph. For the purpose of illustration, for this work it was built a *Boolean* model of four variables with values generated at random and whose adjacency matrix structure can be seen in Figure 7. Recalling that one of the properties of causal networks is that the relations between the variables can be considered transitive, to have a better view of the relationship between indirect variables, it was applied an algorithm of spreading of regulated effects, which performs the calculation of the junction of the direct and indirect effects of the relations between the variables of the system. The final result of this process is the construction of a new adjacency matrix in which the cells receive new values not *Booleans* representing their relationships directly or indirectly. The next procedure would be the ranking of variables by the concepts of mobility and dependence.

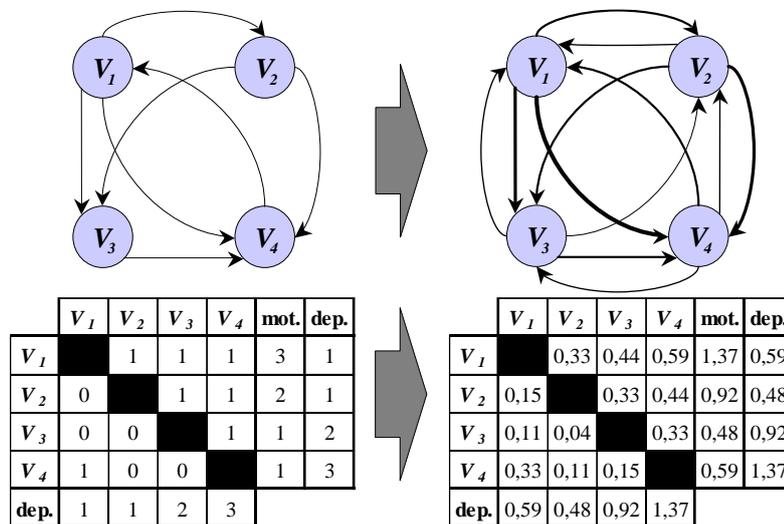


Figure 7: Structural analysis with propagation of regulated effects

Source: Elaborated by the author

3.1 INTEGRATION OF STAGES OF STRUCTURAL ANALYSIS AND CROSSED-IMPACTS

The first strategy proposed in this article is precisely to skip this first stage of valuation of the arcs of a graph only by determining their relations and go straight to the filling of a matrix of generic crossed-impacts (presented in Figure 4).

From that moment on, it is possible to calculate the values of each relationship between the variables of the system to interpret if the values of the adjacency matrix of a structural analysis is a measure of the influence that the predecessor variable performs in the variable conditioned. It is possible to say that from the moment in which it is supposed logically that this order of magnitude of the influence is proportional to the differences between the probabilities *a posteriori* (conditions) and the probability *a priori* of the states of a variable. That is, the higher the value of the arc, the more influence the variable receives from its predecessor, the greater are the deviations between the probabilities *a posteriori* and *a priori*, the more dependent it is of its predecessor. In reverse, the smaller the value of the arc is, the less influence it receives, the lower the deviations are between the probabilities, the more independent it is.

Taking this concept of influence as the behavior of the probabilities *some posteriori*, given the probability *a priori* and the value of the arc in the adjacency matrix, it is proposed here the classification of relations between two variables in four degrees of influence:

- Maximum influence: characterizes the total dependence of the variable successor in relation to its predecessor, setting up a relationship of casual certainty. In other words, if the condition of the variable predecessor is known, then it will be known the condition of a successor. In terms of probabilities, this means that $P(A/B) = 1$ and $P(A) = P(B)$;
- Strong influence characterizes the strong influence of the predecessor variable over the successor, that is, the conditional probabilities $P(A/B)$, due to the influence of the occurrence of B, distance to a great extent from $P(A)$. In other words, there is a

strong dependence of the variable successor in relation to the predecessor. With regard to the order of magnitude, the closer the value in the adjacency matrix of the value of maximum influence, the stronger this influence is;

- Weak influence: it characterizes the weak dependence between the variables and, therefore, a weak influence of variable predecessor on the successor. This means that the probabilities *a posteriori* $P(A/B)$ are relatively close to the probabilities *a priori* $P(A)$. The lower the value of the relationship between these variables, the less this influence is and more independent they are among themselves;
- No influence: it characterizes the independence between the variables. Translating into probabilities, it means that $P(A/B) = P(A)$. The value of this relationship is null.

To illustrate these concepts, generic probabilities were generated randomly and four sets were chosen as examples for each of these degrees of influence mentioned above. In these examples (Figure 8), the tables of probabilities consider three states (mutually exclusive) for each variable and its probabilities *a priori* and *some posteriori*. The calculation of the values of the degrees of influence is performed by summing the absolute values of the differences between the probabilities *a posteriori* and *a priori*, $[P(A/B) - P(A)]$, in each state of each variable and then normalizing the maximum possible so that these values are limited between 0 and 1. In the case of three states per variable, this maximum value is 4, which corresponds to the sum of all elements of the array of absolute values presented in Figure 8. Generally, this value is given by the formula $2 \cdot (e - 1)$, being e the number of states of a given variable.

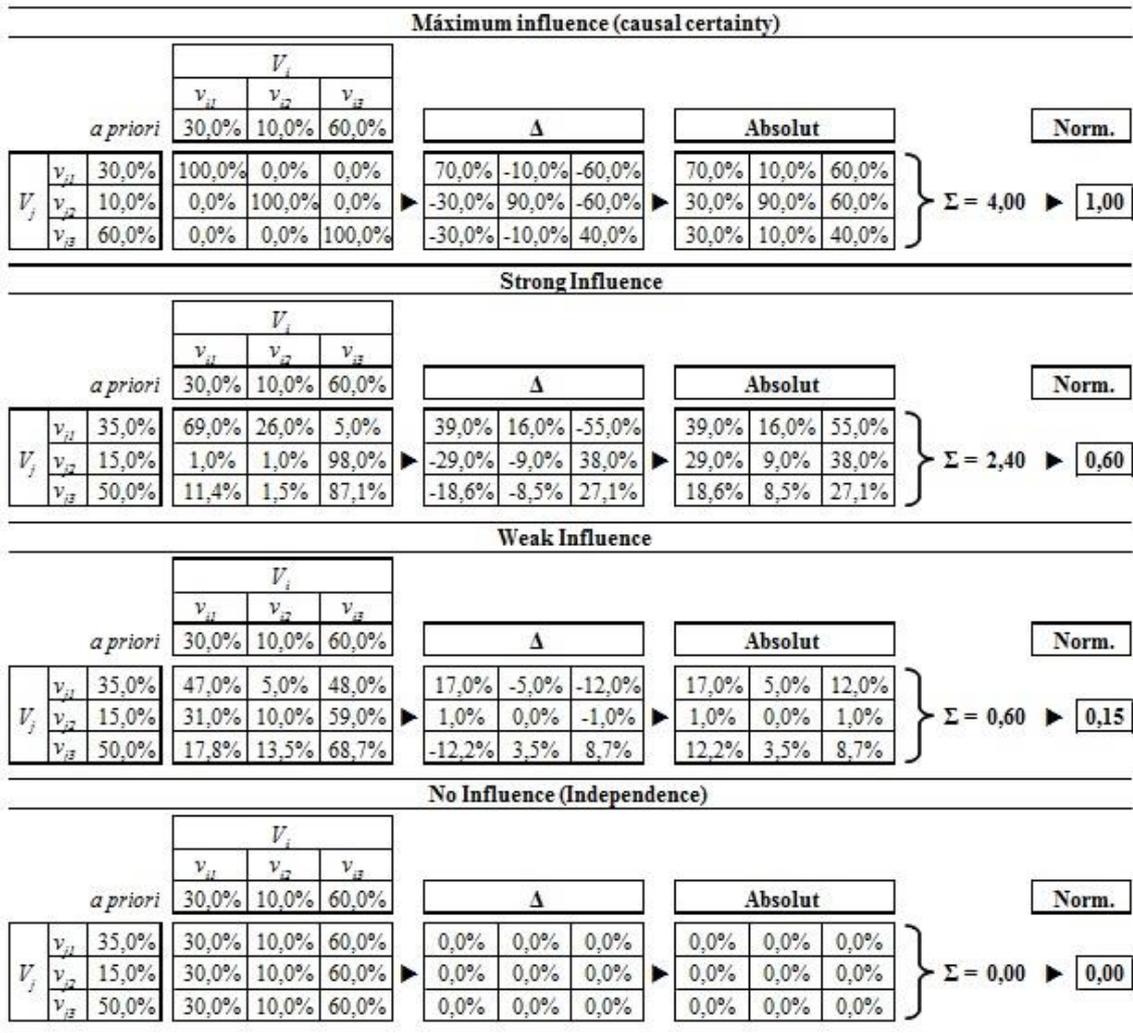


Figure 8: Deviations between probabilities as a measure of dependence among variables

Source: Elaborated by the author

The principle that serves as the basis for this method is the logical deduction that there is a correlation between these two ways of measuring the level of dependence (or influence) between two variables in a causal model:

- Calculation of the junction of direct and indirect relations generated by the analysis of spreading of effects is displayed from a Boolean array;
- Standardization of the sum of the absolute values of the deviations among probabilities *a posteriori* and *a priori* of all states mutually exclusive of the variables.

To prove that the loss of information is minimized when the structure of a cyclic model is simplified to a DAG when it is not considered relations of lesser degree of influence, Fischer (2010) conducted an empirical test that showed satisfactory results.

As for the model constructed, the information contained in the matrix of crossed-impacts would suffice to obtain the probability of all settings of variables that a model can provide. But one of the most important roles in the process of building scenarios is promoting experimentation and experience of users in the dynamics of the relationship between the variables. The morphological matrix allows such experience, since, as the estates of the variables are chosen, the probabilities of each level are emerging, guiding the user in the choice of the desired configuration, because the data processed within the matrix of crossed-impacts form a distribution of joint probabilities which, applied to the structure of the morphological analysis, end up creating a morphological subspace.

3.2 BAYESIANS NETWORKS AS A BASIS FOR THE MORPHOLOGICAL SUBSPACE

In the second approach of this work, is developed using the concept of *Bayesians* networks as a means of unification of the matrix of crossed-impacts and morphological analysis, as Ritchey (2005) (2005) had already proposed.

The concept of *Bayesians* networks comes within this framework as a method of calculating the data provided through consultation with experts in the form of subjective conditional probabilities concerning evaluations about the occurrence of events involved in the construction of scenarios. The role of the matrix of crossed-impacts is being used as the table of probabilities of the probabilistic model, serving as a tool by which the probabilities are fed, organized, checked for its consistency and prepared for the next step. The next step consists in the construction of the scenarios themselves, in which the tool used is the morphological analysis, which receives the information handled in the matrix of impacts crossed by the

principles of *Bayesians* networks, and transforms them into its morphological subspace.

3.2.1 Establishment of the model structure

It is already known that the first step in the process of creating a scenario is the achievement of structural analysis, which usually results in a cyclical system. As BN require acyclic causal, the cyclic system resulting from structural analysis needs to be simplified. There are several ways to achieve this simplification; one of these strategies is simply transform the cyclic graph in a DAG when removing or adding some edges until all loops are broken. The edges should be removed or added so that the difference in the results obtained in the comparison between the models is minimized and that none of the data dependencies in the loop graph are violated (Sandnes & Sinnen, 2004). This strategy is supported by the concept of stability (Pearl, 1999) and by the concepts of dependence and addiction (Pearl, 1988). Operationally, to carry out this step, you can think about the arcs of a graph as vectors, process summarized in Figure 9, in which complements the process initiated in Figure 7.

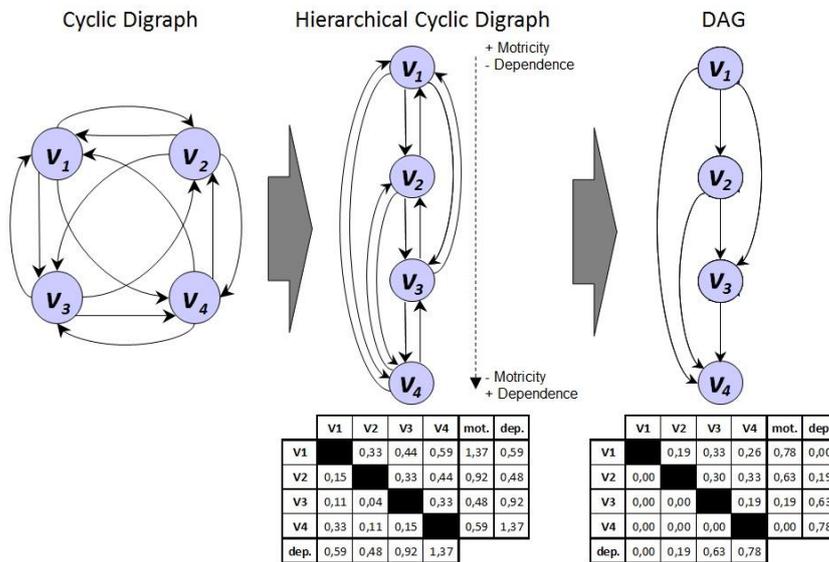


Figure 9: Reduction to a DAG by normal dependence

Source: Elaborated by the author

Considering a pair of variables, structurally (in the case of a causal modeling) it can contain a maximum of two arcs, each one in a sense. If

these arcs are valued, it is possible to assume they are "causal vectors"; one of them has a positive sign (direction of the hierarchy held) and the other, a negative sign (opposite direction). The junction of these two "vectors" in opposite directions forms a vector resulting in only one direction (in the direction of greater absolute value) and whose value is equal to the difference between the absolute values of each vector. When repeating this process for each pair of variables in the model, it is possible to obtain finally a DAG starting from a digraph cyclic.

3.2.2 Bayesians networks as a basis for the array of crossed-impacts

In the next step, it is started the quantitative analysis phase, in which the causal relations among the variables receive values corresponding to the probability of occurrence of each of its member, remembering that each variable V_i may present e mutually exclusive states ($v_{i1}, v_{i2}, \dots, v_{ie}$). It is important to remember that these probabilities may have multiple sources: Frequencies are already known, inferred frequencies, the user of the process, expert opinion, etc. as the focus of this work is to build scenarios, the probabilities are consulted to the experts who decide whether they will use frequencies or their intuition to give probabilities to events (i.e., likely subjective).

Resuming the graphical representation of a Bayesian network shown in Figure 6 and recalling the rule *naïve* of Bayes, to define the relations in a *Bayesian* network, it is necessary to confer only probabilities *a posteriori* pair to pair for each state of each variable in the model. With this information, you can start feeding the array of generic crossed-impacts, already presented in Figure 4. At this time, it is possible to perceive one of the great advantages of using *Bayesians* networks in scenarios. Knowing that the variables are already in hierarchical order in the matrix, it is easy to see that only the upper diagonal right (light part of the matrix in figure 4) has the need to be filled. This means a reduction of approximately half the number of information that would normally be required in other methods of traditional crossed-impacts - the total amount of information (probabilities *a priori* and *a posteriori*) in a matrix of cross-impacts can be

easily deduced in formula $v.e.[e.(v-1) +1]$, where v is the number of variables of the system and e is the number of states of each variable.

However, with the information supplied in the array of crossed-impacts, it is possible to calculate the results of joint probabilities directly with the formula of the rule in the chain. In Figure 10, it is easy to see that (a) represents the information contained on the diagonal top right of the matrix and the information (b) is contained in the diagonal lower left (darkest part). This region of the matrix of crossed-impacts would be automatically populated by the rule of Bayes from the information contained in the diagonal top right of the array.

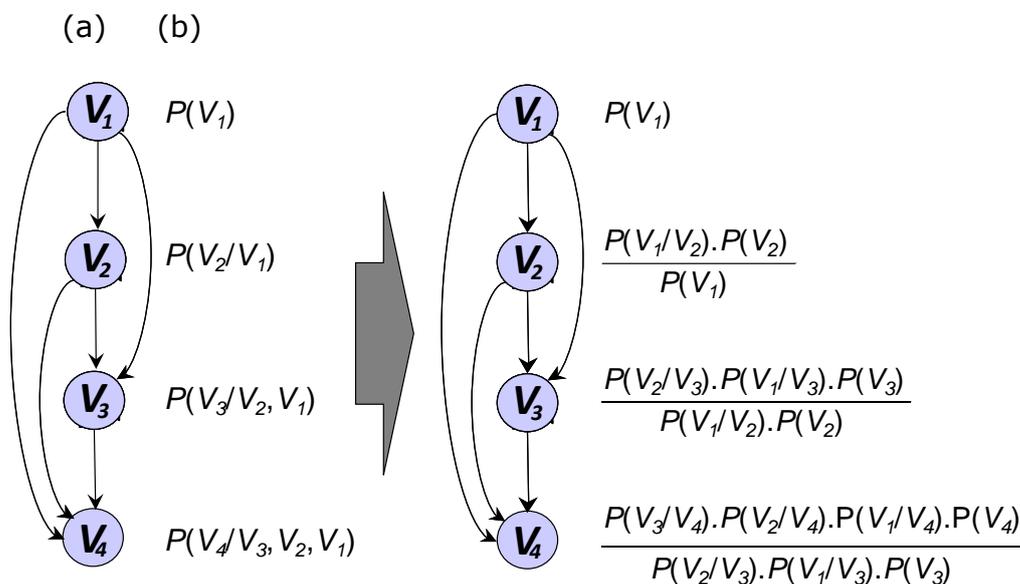


Figure 10: Graphic representation of the rule of chain for one model of four variables

Source: Elaborated by the author

With the information built diagonally left bottom of the array of crossed- impacts, it is now possible to calculate the distribution of joint probabilities of the model, containing the probabilities of all settings of the variables. It is always important to remember that, for every analysis of crossed-impact being used the probability reasoning (regardless of the calculation methodology used), it is necessary the coherence and consistency between the probabilities assigned to events,

And, finally, it is reached the stage of configuration and experience in the construction of scenarios: morphological analysis. Being the MA a

hierarchical tool, the concept of BN fits perfectly in creating a morphological subspace, as it can be seen in Figure 1.

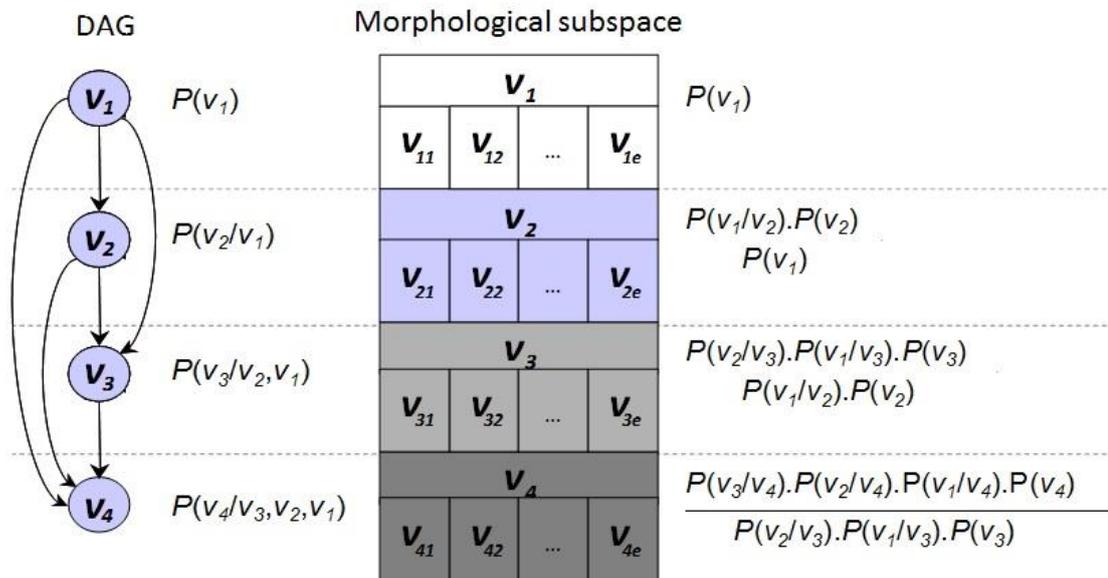


Figure 11: Structure of Bayesian network as morphological subspace

Source: Elaborated by the author

4 FINAL CONSIDERATIONS

In prospective analyses, the tools and the methods used must deal with the minimum amount of information possible and be simple enough so that the results can be easily assimilated by decision-makers. At the same time, they should confer robustness and credibility to the results achieved.

Based on the premises above, the proposals of this work were developed aiming at the goal of reducing complexity in the process of building scenarios, while maintaining the robustness of the analysis and the results. In accordance with the criteria already set out in the section Introduction, for the verification of the concepts of complexity and robustness in the context of this work, generally speaking, the conclusion is that the method proposed, as well as bringing the advantage of reducing the complexity of the analysis, it also proved to be robust, consistent, reliable and rigorous, since it is supported by honored theoretical foundations. The Method of the scenarios, by Michel Godet, now proposes

the practice of those tools in a complementary way, but not integrated. This is because they are tools of different types:

- The structural analysis is a tool that has as its main objective to study and map the dynamics between the variables and determine the degree of influence of each one in the system. Its limitation lies in the fact that this degree of influence is always measured in a very subjective way, not taking into consideration the possible states that this variable can present;
- The array of crossed- impacts is a quantified tool that receives and processes information, verifies its consistency and coherence, and returns the results. However, it has the deficiency of being difficult to show the characteristics of the model, understand the existing standards and understand the dynamics in the relations between the variables;
- The morphological analysis is a qualitative tool that has as main features the easy viewing of the behavior of the model and the possibility of the user to experience the construction of scenarios and understand the patterns and dynamics that exist. Its limitation lies in the fact that many times the construction of scenarios is based on subjective rules, which can generate a lot of discussion and distrust.

Working on the integration of these tools makes it possible to reduce its limitations while maximizing its benefits. One of these benefits relates to one of the most important roles in the process of building scenarios - the experimentation and the experience of users of the process in the dynamics of the relationship between the variables. This is an aspect that the integration allows efficiency gains, but may not misrepresent the individual objectives of each tool. It is important, therefore, that the user does not receive just the results of scenarios, but that experience and understand the dynamics between the variables.

Besides making it easy the navigation between the tools, the integration still makes it easy to give feedback to the process, allowing for a review in the structure of the model or data input (probabilities) and the view of the breaking points within the model. This *feedback* in the process,

being possible to go back to the beginning of the modeling and modify their probability of occurrence *a priori* and *some posteriori*, is performed if it is perceived that:

- It does not seem to be a lot of consistency in the dynamics between the variables. In this case, the team of scenarios should be very careful to conduct the exercise to take the correct arbitration, thus avoiding that this process of *feedback* takes the deconstruction of the ruptures and return to normative scenarios
- The action plan developed at the end of the process of "cenarização"[SETTING THE SCENARIO] has a direct impact on the system of variables and actors, modifying the structure of the system (main purpose of Strategic Foresight).

Still on the topic complexity, in the case of method of crossed impacts, the richer in information the system is, the more complex and long the process becomes. This is because the amount of information needed to feed an array of crossed-impacts grows exponentially in relation to the increase of the variables used in the model. At this point, there are two gains. The first is the proposal to remove the structural analysis and use the cross-impact matrix to perform directly the ranking of variables. The second is to use the concept of *Bayesians* networks as a way to integrate the method of raw impacts used to the morphological analysis. In the latter case, number of information, in this case, probability, is reduced almost by half. In other words, the model becomes easier to be built, taking less time and enabling greater agility in driving the process. As an example, a system of 16 variables (v) with two states each (e would have the need to be fed with 992 information of probabilities (*a priori* and *a posteriori*) to complete a cross-impact matrix, given the formula presented previously, $v.e.[e.(v-1) + 1]$. Using *Bayesian* networks, this number is reduced to 512, because the formula for calculating the number of information evolves to $v.e. [0,5. e.(v-1) + 1]$.

In the field of the search for robustness, we can conclude that there is direct gain in the use of *Bayesians* networks when there is a need for introduction of new variables in the system or greater detail of the variables. Using BN, it is possible to do so without raising the complexity.

Taking again the example from the previous paragraph, if you keep the number of information in the same order of magnitude, it is possible to have 22 variables instead of 16 (968 probabilities) or reduce the model for 15 variables, but considering three states by variable (990 probabilities). In the case of the proposal of this work, which makes use of the MA in an integrated manner, the more stratified in states the variable is, the better. Therefore, the approach of enriching the model through the increase in the number of states of the variables makes more sense (it is simply not possible to develop a morphological analysis minimally rich with two states per variable).

One of the bases of the development of the proposed use of Bayesian networks is the belief in the reliability of information generated by consultation with experts, what constitutes a limitation of the proposed approach. The data collected can have different levels of reliability, but are treated in a similar manner. In addition, as presented in the course of this work, there is no way to say that subjective probabilities *some posteriori* is more reliable than the probabilities *a priori*. One way to resolve this limitation is to consider that the experts already have in their minds throughout the chain of events to estimate the probabilities *a priori*, which can be handled directly by methods that use, for example, the distribution Beta. The comparison between these two approaches for the treatment of views (in the form of probabilities) collected from experts can be configured as a clue to future works.

Another limitation of this study, now in the approach proposed of integration of structural analysis with crossed impacts, is the assumption that there is a correlation between the degrees of influence of variables calculated by structural analysis and through the gaps between the probabilities *a priori* and *a posteriori* of the variables, indicating that this enhancement represents the level of dependence between the variables (the bigger, the more dependent; the smaller, the more independent). This statement is conceptual and arises only from logical deduction, without any empirical confirmation - the structural analysis arises from null values to define independence, as well as the standard deviation of zero between

probabilities *a priori* and *some posteriori* indicates independence. It may reside in this fact another evidence for future studies.

Finally, it was said in the course of this work, that the proposed approach reduces the complexity of the analysis, maintains the robustness and allows users to easily understand the methodology and understanding of the dynamics of the scenarios. These statements were made based on the belief that the integration of these tools would lead the user of scenarios, to these conclusions. Therefore, another clue to future work would be to measure the perception of professionals working with scenarios on how it behaves the approach proposed in terms of the complexity of the analysis, robustness of the process, reliability of the result and understanding of the methodology.

REFERÊNCIAS

- Arcade, J., Godet, M., Meunier, F., & Roubelat, F. (1994). Structural analysis with the MICMAC method & actors' strategy with MACTOR method. In J. C. Glenn, & T. J. Gordon (Orgs.), *Futures research methodology*. Paris: AC/UNU Millennium Project.
- Ben-Gal, I. (2007). Bayesian networks. In F. Ruggeri, R. S. Kennett, & F. W. Faltin (Orgs.), *Encyclopaedia of statistics in quality and reliability*. UK: John Wiley & Sons.
- Charniak, E. (1991). Bayesian networks without tears. *AI Magazine*, 12(4), 50-63.
- Duncan, N. E., & Wack, P. (1994). Scenarios designed to improve decision making. *Planning Review*, 22(4), 18-46.
- Fischer, R. B. (2010). *Desenvolvendo o conceito de redes bayesianas na construção de cenários prospectivos*. Dissertação de Mestrado, Escola Brasileira de Administração Pública e de Empresas da Fundação Getulio Vargas (EBAPE/FGV), Rio de Janeiro, RJ, Brasil.
- Godet, M. (1993). *From anticipation to action: a handbook of strategic prospective*. Paris: UNESCO.
- Godet, M., & Roubelat, F. (1996). Creating the future: the use and misuse of scenarios. *Long Range Planning*, 29(2), 164-171.
- Godet, M., Monti, R., Meunier, F., & Roubelat, F. (1999). *A caixa de ferramentas da prospectiva estratégica*. Lisboa: CEPES.
- Gordon, T. J. (1994). Cross impact method. In J. C. Glenn, & T. J. Gordon (Orgs.), *Futures research methodology*. Paris: AC/UNU Millennium Project.
- Heijden, K. (2004). *Planejamento de cenários: a arte da conversação estratégica*. Porto Alegre: Bookman.
- Jensen, F. (1996). *An introduction to Bayesian networks*. London: UCL Press.
- Kjærulff, U. B., & Madsen, A. L. (2005). *Probabilistic networks: an introduction to Bayesian networks and influence diagrams*. Recuperado em 30 de junho, 2009, de <http://www.cs.aau.dk/~uk/papers/pgm-book-I-05.pdf>

- Lacerda, W. S., & Braga, A. P. (2004). Experimento de um classificador de padrões baseado na regra naíve de Bayes. *INFOCOMP Journal of Computer Science*, 3(1), 30-35.
- Lowd, D., & Domingos, P. (2005). Naïve Bayes models for probability estimation. *Proceedings of the International Conference on Machine Learning*, 22, Bonn, Germany.
- Marcial, E. C., & Grumbach, R. S. (2005). *Cenários prospectivos: como construir um futuro melhor* (3a ed.). Rio de Janeiro: FGV.
- Marques, E. (1988). *Prospec: modelo de geração de cenários em planejamento estratégico*. Recuperado em 11 de abril, 2007, de http://www.bndes.gov.br/conhecimento/livro_ideias/livro-11.pdf
- Miles, I. (2005). Scenario planning. In *UNIDO Technology foresight manual* (vol. 1, pp. 168-193) Vienna: Unido.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: network of plausible inference*. San Francisco, CA: Morgan Kaufmann.
- Pearl, J. (1999). Graphs, structural models and causality. In C. Glymour, & G. F. Cooper (Orgs.), *Computation, causation and discovery*. Cambridge, MA: AAI/MIT Press.
- Pearl, J. (2000). *Causality: models, reasoning, and inference*. New York, NY: Cambridge University Press.
- Perestrelo, M. M. M. C., & Caldas, J. M. C. (1998). *Instrumentos de análise para utilização no método dos cenários: I – análise estrutural*. Lisboa: Dinâmia.
- Rish, I. (2001). An empirical study of the naïve Bayes classifier. *Proceedings of the International Joint Conference on Artificial Intelligence*, 17, Seattle, USA.
- Ritchey, T. (2002). *General morphological analysis: a general method for non-quantified modelling*. Recuperado em 20 de abril, 2007, de <http://www.swemorph.com/pdf/gma.pdf>
- Ritchey, T. (2005). *Futures studies using morphological analysis*. Recuperado em 20 de abril, 2007, de <http://www.swemorph.com/pdf/futures.pdf>
- Sandnes, F. E., & Sinnen O. (2004). Stochastic DFS for multiprocessor scheduling of cyclic taskgraphs. *Proceedings of the International Conference on Parallel and Distributed Computing: Applications and Technologies*, 5, Singapore.

Schnaars, S. P. (1987). How to develop and use scenarios. *Long Range Planning*, 20(1), 105-114.

Schwartz, P. (2000). *A arte da visão de longo prazo: planejando o futuro em um mundo de incertezas* (4a ed.). São Paulo: Best Seller.

Wack, P. (1985a, September-October). Scenarios: uncharted waters ahead. *Harvard Business Review*, 63(5), 72-89.

Wack, P. (1985b, November-December). Scenarios: shooting the rapids. *Harvard Business Review*, 64(6), 139-150.