ISSN 1870-4069

Integrating CBR with Data in Bayesian Networks for Decision Making in an Echelon Supply Chain Distribution Solution

Adrián Francisco Loera Castro[1], Alberto Ochoa Zezzatti[2], Jaime Sánchez[3], Humberto García Castellanos[3]

[1] Tecnológico Nacional de México Campus I.T.C.J., Depto. de Ingeniería Industrial y Logística, Cd. Juárez, Chih. Mexico

aloera@itcj.edu.mx

[2] Universidad Autónoma de Cd. Juárez, Depto. de Ingeniería Industrial y Logística, Cd. Juárez, Chih. Mexico

alberto.ochoa@uacj.mx

[3] Tecnológico Nacional de México Campus I.T.C.J., División de estudios de posgrado e investigación, Cd. Juárez Chih., Mexico

jsanchez@itcj.edu.mx, hgarcia@itcj.edu.mx

Abstract. When developing a causal probabilistic model, that is, a Bayesian network (BN), it is common to incorporate expert knowledge of factors that are important for decision analysis, but there are models where historical data is not available or difficult to obtain, or it is difficult to have a human expert nearby to help. This document explains how data is developed from a discrete/continuous simulated variable through a BN and mixed integer-linear programming (MILP), and the impact of this variable is measured as an important element for the decision-making model. Consider as an additional expert variable. The CBR model and the variable in question is contextualized to support in the decision-making process in a supply chain through two stages, the first is considered multiple factories, with multiple distribution centers (DC) and second, from the multiple distribution centers as it reaches multiple points of sale. As a design of a decision support system for the construction of a supply chain network (SCN) for a range of multiple end products, as well as the determination of factories and distribution centers, it also helps in the design of the distribution network strategy that satisfies all the capacities and requirements of demand of the product imposed through the points of sale. At the end of the work, an evaluation of the performance of two Bayesian networks is carried out, where one of them represents the incorporation of the expert variable using two methods, one of them the receiver operating characteristic (ROC) curve and two a method proposed by Constantinou et al. [2]., Where in both cases the Bayesian network gave a better performance with the expert variable.

Keywords: making decisions, Bayesian networks, case-based reasoning, supply chain networks, mixed integer-linear programming.

1 Introduction

Bayesian networks (BNs) [6] are rapidly becoming a leading technology in applied Artificial Intelligence. By combining a graphical representation of the dependencies between variables with probability theory and efficient inference algorithms, BNs provide a powerful and flexible tool for reasoning under uncertainty.

It has been argued that developing an effective BN requires a combination of expert knowledge and data [2]. Yet, rather than combining both sources of information, in practice, many BN models have been learned purely from data, while others have been built solely on expert knowledge.

Supply chain management is a complex domain where experienced manager practitioners hold much of their knowledge implicitly, making an appealing target for expert systems development, using Case-based reasoning (CBR). The efficiency of case retrieval algorithm is determined and affected directly by the used method for case representation. As a result, it is more logical to introduce case retrieval methods after surveying the representation methods to link them together Accuracy in obtaining the beliefs of experts, it is often unrealistic to expect the expert to provide precise probability values. In this document we present an application of a methodology proposed by [2] to a case of a BN using the learning cause of an Expert System (ES) in combination to model problems of distribution in the Supply Chain Network (SCN).

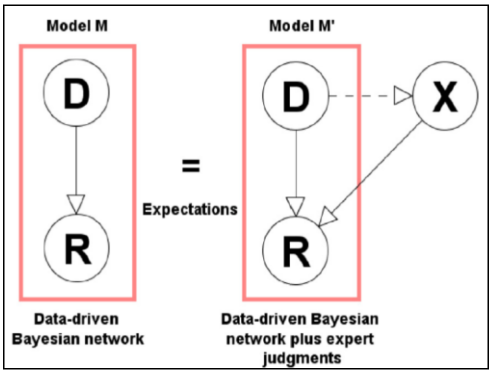

Fig. 1. Illustration where the Model M, with the data variables D and R, extends to the alternative Model M that incorporates the non-human expert variable X. Source: Constantinou et al., Integrating expert knowledge with data in Bayesian networks: Preserving data-driven expectations when the variables remain unobserved, 2006 [2].

Constantinou et al. [2] proposed a method for the evaluation of Bayesian networks, which is described below. The model *M* represents empirically observed data about the influence of D. In the example in figure 1 , the states of D are the investment options {bonds, shares, properties} and R is the *Network objective* , expressed as an observed distribution of values for each different option.

We assume that, from relevant data:

(a) $P(D_i) = d_i$ is known for each $i = 1, \ldots, n$,

(b) $f\,(R \mid D_i)$ is a known distribution for each $i = 1, \ldots, n$,

Hence, these are the parameters of the model *M*. Let the *expected value* $E(f(R \mid D_i)) = r_i$ for each $i = 1, \ldots, n$. For simplicity, we write this as $E(R \mid D_i) = r_i$. Hence, in model *M* the expected value of *R* is:

$$E_M(r) = \sum_{i=1}^{n} E(r|D_i)P(D_i) = \sum_{i=1}^{n} r_i d_i \tag{1}$$

Now consider the revised BN model *M'*, as shown in Fig. 1. Here *X* is an expert supplied variable with *m* states $X_1, \ldots, X_m$. We assume the expert provides the prior probabilities for *X*, i.e. $P(X_j \mid D_i) = p_{ij}$ for each $i = 1, \ldots, n$ and for each $j = 1, \ldots, m$. When *D* and *X* are not linked, then instead of $n \times m$ priors we only need *m* priors $P(X_j) = p_j$ for each $= 1, \ldots, m$. The challenge for the expert is to complete the conditional probability table (CPT) for R in M' in such a way as to preserve all of the conditional expected values of R given D in the original model M, and also preserve the marginal expectation. Specifically, we require:

$$E_{M'}(R|D_i) = E_M(R|D_i) = r_i \ for\ each\ i = 1, \ldots, n \tag{2}$$

Note that, if we can establish Eq. (2), then it follows from Eq. (1) that:

$$E_{M'}(R) = E_M(R)$$

Specifically, Eq. (2) is also sufficient to prove that the unconditional expected value - the expected value of R when D is unobserved- of R is preserved in M'.

Table 1. The CPT for R in M'.

D	D_1				…	D_i				…	D_n			
X	X_1	X_2	…	X_{m-1}	X_m	…	X_1	X_2	…	X_{m-1}	X_m	…	X_1	X_2	…	X_{m-1}	X_m
R	f_{11}	f_{12}	…	f_{1m-1}	f_{1m}	…	f_{i1}	f_{i2}	…	f_{im-1}	f_{im}	…	f_{n1}	f_{n2}	…	f_{nm-1}	f_{nm}

The general form of the CPT for R in M' can be written as a function f_{ij}, whose expected value is r_{ij} for each i = 1 , . . . , n and j = 1 , . . . , m , as shown in Table 1. Specifically,

$$E(f_{ij}) = E_{M'}(R|D_i, X_j) = r_{ij} \ for\ each\ i = 1, \ldots, n\ and\ j = 1, \ldots, m$$

Since each X_i is conditioned on D_i we can use marginalization to compute:

$$E_{M'}(R|D_i) = \sum_{j=1}^{m} E(R|D_i, X_j)P(X_j|D_i) = \sum_{j=1}^{m} r_{ij} P_{ij} \tag{3}$$

Since by Eq. (2) we require:

$$E_{M'}(R|D_i) = E_M(R|D_i) = r_i \ for\ each\ i = 1, \ldots, n$$

it, therefore, follows from Eq. (3) that we require:

$$\sum_{j=1}^{m} r_{ij} P_{ij} = r_i \quad for\ each\ 1 = 1, \ldots, n \tag{4}$$

Eq. (4) thus expresses the necessary constraints on the expert elicited values for r_{ij}. We can use Eq. (4) as a consistency check on the expert elicited values if the user wishes to provide them all. However, in practice we would expect the user to provide a subset of the values and so use Eq. (4) to solve for the missing values. There is a unique solution in the case when the expert is able to provide $m - 1$ of the required m values

$$r_{11}, r_{12}, \ldots, r_{1m-1}, r_{1m}.$$

To prove this, without loss of generality suppose that r_{im} is the 'missing value'. Then we can compute the value of r_{im} necessary to satisfy Eq. (4) . We know, by Eq. (4), that:

$$r_i = \sum_{j=1}^{m} r_{ij} P_{ij}$$

So:

$$r_i = \left(\sum_{j=1}^{m} r_{ij} P_{ij} \right) + r_{im} P_{im}$$

Thus:

$$r_{im} = \frac{r_i - \left(\sum_{j=1}^{m} r_{ij} P_{ij}\right)}{P_{im}} \tag{5}$$

For each $i = 1, \ldots, n$ Eq. (5) thus provides the formula for computing the missing CPT values necessary to preserve in the model *M'* all of the conditional expected values of *R* given *D* in the original model *M* .

2 The Role of Distribution in the Chain of Supply

A supply chain is defined as a process with a complete set of activities wherein raw materials are transformed into final products, then delivered to customers by distribution, logistics, and retail. All inter-organizational practices such as planning, purchasing, distribution, delivery process, and reverse logistics are considered as a supply chain management system [21].

3 Distribution Decisions

Development of the new theories and methodologies in logistics and supply chain management can lead to the higher level intelligent and advanced systems. Such kind of systems enable supply chain experts to facilitate information-sharing, highly qualified decisions and to increase the value to products and services by internal coordination. Over the last decades, the direction of decision support systems has changed drastically. To monitor the materials cost in a garment manufacturer, a decision support model has assisted decision-makers in selecting efficient ways to reduce total manufacturing costs. Decision making is influenced by the characteristics and context of decision situations [37] and it is viewed that understanding the characteristics of different types of organizational decision-making contexts is a prerequisite for understanding the nature of decision-making processes and requirements for decision support within different types of decision-making contexts. There are several ways to characterize different types of decision situations and their associated decision-making contexts within organizations.

4 Case-Based Reasoning (CBR)

A CBR system should be organized with some basic elements: the knowledge representation, to depict the cases, and the similarity measure to define how much a case is similar to another one [16, 17].

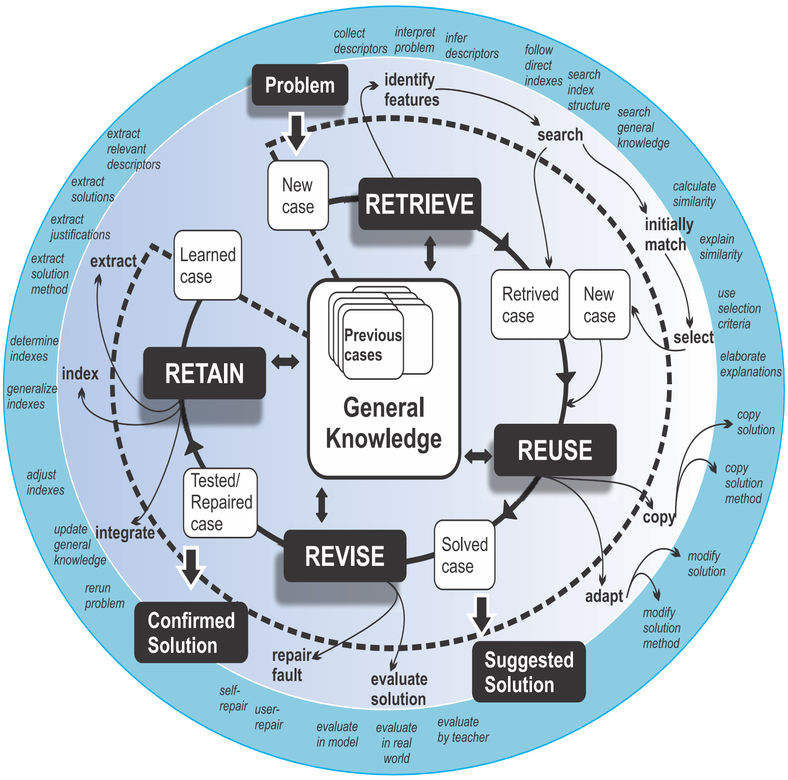

Fig. 2. A process-oriented vision of the CBR adaptation cycle based on Aamodt and Plaza (1994) Source: Loera et al. Implementation of an Intelligent Model for Decision Making Based on CBR for Supply Chain Solution in Retail for a Cluster of Supermarkets.

In figure 2, the tasks are shown with the names of the nodes in bold, while the methods are in italics.

5 R based Framework for Distribution Planning

5.1 Model Formulation

The foundation of a CBR system is the representation and definition of a case. So far, there is no uniformed standard to represent a case [5]. The constituted model represents two echelons, multi-factories, multi-warehouse or distribution centers (DC), and multi-sales points. Decision maker wishes to design of supply chain network (SCN) [35] for the end product, determine the factories and DCs and design the distribution network strategy that will satisfy all capacities and demand requirement for the product imposed via sales points. The problem is a single-product, multi-stage SCN design problem. We formulated the SCN design problem as a Mixed-Integer Linear Programming model (MILP), [18]-[21], as is shown in figure 3.

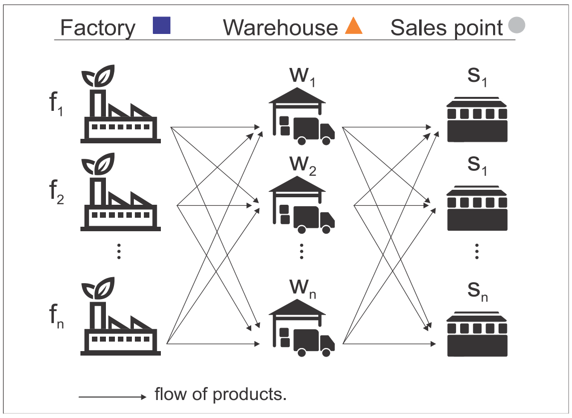

Fig. 3. Simple network of two-stages in supply chain network.

5.2 Model Nomenclature

The indices, parameters and decision variables of the mixed integer linear programming model are listed below:

Index

$f = 1,2,3, \dots F$	Set of production facilities,
$p = 1,2,3, \dots P$	Set of product type,
$w = 1,2,3, \dots W$	Set of warehouse facilities (distribution centers),
$s = 1,2,3, \dots S$	Set of sales points.

Parameters

m_{fp} Production cost in the factory f for the product p,

D_{fw} Distance between the factory f and the warehouse w,
T_p Cost of transporting the product p from the factory f to the warehouse w,
d_{sw} Distance between the sale point s and the warehouse w,
t_p Cost of transporting the product p from the sales point s to the warehouse w,
r_{sp} The s point of sale demand r for the product p,
C_{fp} The f factory capacity for the product p,
c_w Capacity of the warehouse w,
O_p The turnover rate for each product p,

Decision variables

x_{pfw} The amount of product p that is transported from factory f to warehouse w,
y_{sw} A binary variable taking value 1 when sale point s is associated with warehouse w.

Mathematical model proposed by Adrian Loera et al. in his PhD Thesis (2019):

Minimize

$$Z = \left[\sum_{f=1}^{F}\sum_{p=1}^{P}\sum_{w=1}^{W} x_{pfw}(m_{fp} + T_p D_{fw})\right] + \left[\sum_{s=1}^{S}\sum_{w=1}^{W}\sum_{p=1}^{P} r_{sp} t_p d_{sw} y_{sw}\right] \quad (6)$$

Subject to:

$$\sum_{w=1}^{W} x_{pfw} \leq C_{fp} \quad \forall w \in W, \forall f \in F, \forall p \in P \quad (7)$$

$$\sum_{f=1}^{F} x_{pfw} = \sum_{s=1}^{S} (r_{sp} y_{sw}) \quad \forall s \in S, \forall w \in W, \forall f \in F, \forall p \in P \quad (8)$$

$$\left[\sum_{p=1}^{P}\sum_{s=1}^{S} \frac{r_{sp}}{O_p}\right][y_{sw}] \leq c_w \quad \forall w \in W, \forall s \in S, \forall p \in P \quad (9)$$

$$\sum_{w=1}^{W} y_{sw} = 1 \quad \forall w \in W, \forall s \in S \quad (10)$$

$$m_{fp} \geq 0 \quad \forall f \in f, \forall p \in P \quad (11)$$

$$D_{fw} \geq 0 \quad \forall f \in F, \forall w \in W \quad (12)$$

$$T_p \geq 0 \quad \forall p \in P \quad (13)$$

$$d_{sw} \geq 0 \quad \forall w \in W, \forall s \in S \quad (14)$$

$$t_p \geq 0 \quad \forall p \in P \quad (15)$$

$$r_{sp} \geq 0 \quad \forall p \in P, \forall s \in S \quad (16)$$

$$C_{fp} \geq 0 \quad \forall f \in F, \forall p \in P \quad (17)$$

$$c_w \geq 0 \quad \forall w \in W \tag{18}$$

$$O_p \geq 0 \quad \forall p \in P \tag{19}$$

$$x_{pfw} \geq 0 \quad \forall w \in W, \forall f \in F, \forall p \in P \tag{20}$$

$$y_{sw} \in \{0,1\} \quad \forall w \in W, \forall s \in S \tag{21}$$

5.3 Case Representation

Therefore, taking into consideration of characteristics of SCN, the process case of distribution can be defined as a collection of $three - tuple$:

$$CASE = \{H, D, S\},$$

where H is the case number, $D = \{m_{fp}, T_p, D_{fw}, r_{sp}, t_p, d_{sw}\}$ is the condition feature description of distribution and S=$\{x_{pfw}, y_{sw}\}$ is the corresponding solution of distribution planning. The case representation of distribution planning in terms of condition features is shown in Table 2.

Table 2. Distribution planning case presentation.

Case representation of Distribution planning Case number (H): X
Condition features of distribution problem(D) • Production cost in the factory f for the product p • Distance between the factory f and the warehouse w • Cost of transporting the product p from the factory f to the warehouse w. • Distance between the sale point s and the warehouse w. • Cost of transporting the product p from the sales point s to the warehouse w • The s point of sale demand r for the product p • The f factory capacity for the product p • Capacity of the warehouse w • The turnover rate for each product p
Solution (S) • The amount of product p that is transported from factory f to warehouse w. • A binary variable taking value 1 when sale point $s$is associated with warehouse w.

6 Building a BN

Irrespective of the method used, building a BN involves the following two main steps [2]:

1. Determining the structure of the network: many of the real- world application.
2. Determining the conditional probabilities (CPTs) for each node also, referred to as the parameters of the model.

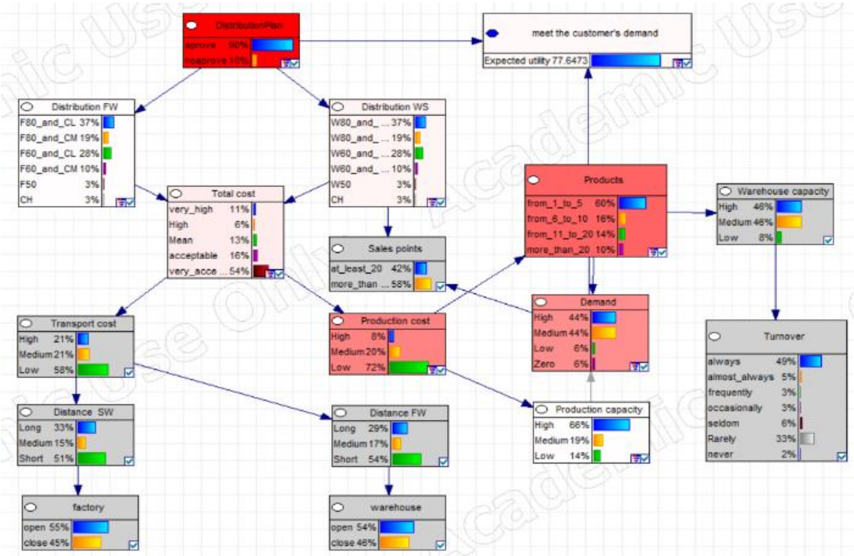

Fig. 4. Bayesian network model to SCN (M model).

6.1 Conditional Probabilities

Once the topology of the BN is specified, the next step is to quantify the relationships between connected nodes – this is done by specifying a conditional probability distribution for each node**,** see table 3.

Table 3. Conditional probability data (M model).

final stage															
Turnover		***Warehouse capacity***		***Sales points***		***warehouse***		***Distance FW***		***factory***		***Distance SW***		***Transport cost***	
State	P(ϑ|x)*	State	P(ϑ|x)	State	P(ϑ|x)	State	P(ϑ|x)	State	P(ϑ|x)	State	P(ϑ|x)	State	P(ϑ|x)	State	P(ϑ|x)
Always	50%	High	44%	at_least_20	42%	open	55%	Long	37%	open	57%	Long	41%	High	26%
Almost_always	5%	Medium	44%	more_than_21	58%	close	45%	Medium	21%	close	43%	Medium	18%	Medium	31%
Frequently	3%	Low	12%					Short	42%			Short	42%	Low	43%
Occasionally	3%														
Seldom	5%														
Rarely	31%														
Never	2%														
	100%		100%		100%		100%		100%		100%		100%		100%
Distribution WS		***meet the customer's demand***		***Demand***		***Production capacity***		***Products***		***Production cost***		***Distribution FW***		***DistributionPlan***	
State	P(ϑ|x)	State	P(ϑ|x)	State	P(ϑ|x)	State	P(ϑ|x)	State	P(ϑ|x)	State	P(ϑ|x)	State	P(ϑ|x)	State	P(ϑ|x)
F80_and_CL	37%	Exp. utility	76%	High	41%	High	59%	from_1_to_5	49%	High	15%	F80_and_CL	37%	approved	90%
F80_and_CM	19%			Medium	41%	Medium	21%	from_6_to_10	16%	Medium	27%	F80_and_CM	19%	no_approved	10%
F60_and_CL	28%			Low	9%	Low	20%	from_11_to_20	19%	Low	58%	F60_and_CL	28%		
F60_and_CM	10%			Zero	9%			more_than_20	16%			F60_and_CM	10%		
F50	3%											F50	3%		
CH	3%											CH	3%		
	100%		76%		100%		100%		100%		100%		100%		100%
** Posterior marginal probability distribution P(ϑ|x)*															

6.2 Parameter Learning

The structure of the Bayesian network was imported into GeNie, a general-purpose Bayesian network commercial software [7] for the parameter learning stage.

Likelihood maximization with randomized initial values for the parameters was used so that the process could be repeated from different starting points to avoid local minima. The process was completed when the expectation maximization algorithm converted; that is when the negative log-likelihood had been minimized.

BNs model the quantitative strength of the connections between variables, allowing probabilistic beliefs about them to be updated automatically as new information becomes available [8]-[11], and this can be observed in tables 5 and 7.

Other outputs of the BN model are represented by the adjacency matrix, which represents a graph with $|V|$vertices, that is, it is a matrix of $|V| \times |V|$ of zeros and ones, where the entry in line i and column j is 1 if and only if the corner (i, j) is in the graph. Case one is represented with an X, as shown in tables 4 and 6.

Table 4. Adjacency Matrix fo M model.

Adjacency Matrix, Lower triangular	Turnover	Warehouse capacity	Sales points	warehouse	Distance FW	factory	Distance SW	Transport cost	Distribution WS	meet the customer's demand	Demand	Production capacity	Products	Production cost	Distribution FW	DistributionPlan
Turnover																
Warehouse capacity	x															
Sales points																
warehouse																
Distance FW				x												
factory																
Distance SW						x										
Transport cost					x		x									
Distribution WS			x					x								
meet the customer's demand																
Demand			x							x						
Production capacity											x					
Products		x									x					
Production cost												x	x			
Distribution FW								x						x		
DistributionPlan									x	x					x	

Table 5. Strength influence of M model.

Strength influence				
Parent	***Child***	***Mean***	***Maximum***	***Weighted***
Demand	meet the customer's demand	0.56	1.00	0.56
Demand	Sales points	0.08	0.20	0.08
Distance SW	factory	0.40	0.60	0.40
Distance FW	warehouse	0.40	0.60	0.40
Distribution FW	Production cost	0.23	0.40	0.23
Distribution FW	Transport cost	0.16	0.45	0.16
Distribution WS	Transport cost	0.15	0.31	0.15
Distribution WS	Sales points	0.09	0.35	0.09
DistributionPlan	Distribution FW	0.40	0.40	0.40
DistributionPlan	meet the customer's demand	0.69	1.00	0.69
DistributionPlan	Distribution WS	0.40	0.40	0.40
Production capacity	Demand	0.00	0.00	0.00
Production cost	Products	0.74	0.84	0.74
Production cost	Production capacity	0.35	0.51	0.35
Products	Demand	0.24	0.35	0.24
Products	Warehouse capacity	0.19	0.29	0.19
Transport cost	Distance SW	0.51	0.73	0.51
Transport cost	Distance FW	0.61	0.84	0.61
Warehouse capacity	Turnover	0.53	0.78	0.53

7 Knowledge Engineering Bayesian Networks (KEBN)

Knowledge Engineering can be viewed as an engineering discipline that involves integrating knowledge into computer systems in order to solve problems normally requiring a high level of human expertise. Similarity assessment techniques (e.g., [17]). The combination is the focus of this article, with the peculiarity that the human expert is replaced by the expert machine, i.e. the CBR [7,9,14,15].

7.1 Integrated BN with CBR

The problem we are interested in solving is the general case where a discrete expert variable -CBR- is inserted into a BN model as a parent of a discrete/continuous data variable, see figure 5. However, when the data variable is discrete some limitations apply proposed per Constantinou et al. [2].

7.2 Data Analysis

Descriptive statistics, Bayesian networks, and Receiver Operating Characteristic (ROC) curve analysis are used in this study for further investigation of the relationships between variables; This is detailed below.

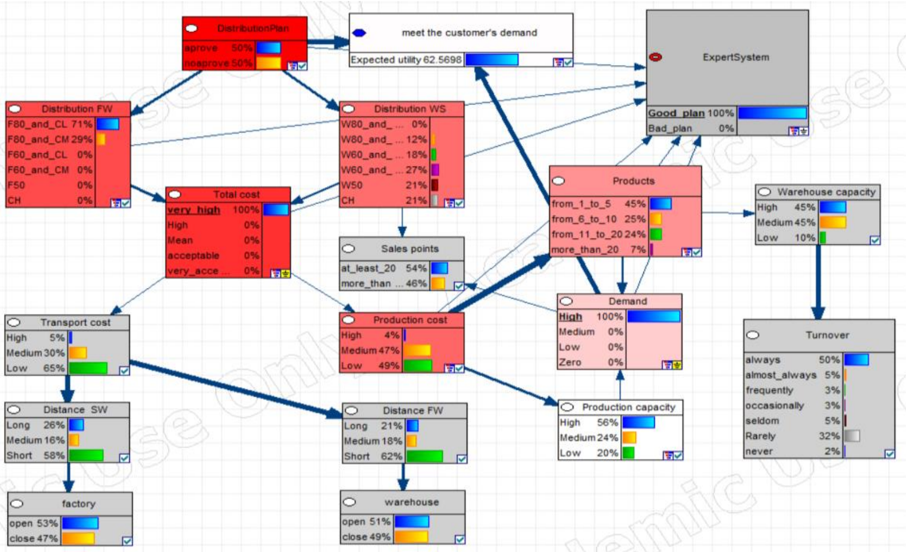

Fig 5. Integrated BN with CBR (M' model).

Table 6. Adjacency Matrix integrated BN and CBR (M' model).

Adjacency Matrix, Lower triangular	ExpertSystem	Turnover	Warehouse capacity	Sales points	warehouse	Distance FW	factory	Distance SW	Transport cost	Meet the customer's demand	Demand	Production capacity	Products	Production cost	Total cost	Distribution WS	Distribution FW	DistributionPlan
ExpertSystem																		
Turnover																		
Warehouse capacity		x																
Sales points																		
warehouse																		
Distance FW					x													
factory																		
Distance SW							x											
Transport cost						x		x										
Meet the customer's demand																		
Demand	x			x						x								
Production capacity											x							
Products	x		x								x							
Production cost	x											x	x					
Total cost	x								x					x				
Distribution WS				x											x			
Distribution FW	x														x			
DistributionPlan	x									x						x	x	

Table 7. Strength influence integrated BN and CBR (M' model).

Strength of influence				
Parent	**Child**	**Average**	**Maximun**	**Weighted**
Demand	meet the customer's demand	0.56	1.00	0.56
Demand	Sales points	0.08	0.20	0.08
Demand	ExpertSystem	0.02	1.00	0.02
Distance SW	factory	0.40	0.60	0.40
Distance FW	warehouse	0.40	0.60	0.40
Distribution FW	Total cost	0.31	0.76	0.31
Distribution FW	ExpertSystem	0.01	1.00	0.01
Distribution WS	Total cost	0.23	0.58	0.23
Distribution WS	Sales points	0.09	0.35	0.09
DistributionPlan	Distribution FW	0.40	0.40	0.40
DistributionPlan	Distribution WS	0.40	0.40	0.40
DistributionPlan	meet the customer's demand	0.67	1.00	0.67
DistributionPlan	ExpertSystem	0.01	1.00	0.01
Production capacity	Demand	0.00	0.00	0.00
Production cost	Products	0.74	0.84	0.74
Production cost	Production capacity	0.35	0.51	0.35
Production cost	ExpertSystem	0.01	1.00	0.01
Products	Demand	0.24	0.35	0.24
Products	Warehouse capacity	0.19	0.29	0.19
Products	ExpertSystem	0.00	1.00	0.00
Total cost	Production cost	0.19	0.41	0.19
Total cost	Transport cost	0.20	0.33	0.20
Total cost	ExpertSystem	0.02	1.00	0.02
Transport cost	Distance SW	0.51	0.73	0.51
Transport cost	Distance FW	0.61	0.84	0.61
Warehouse capacity	Turnover	0.53	0.78	0.53

8 Evaluation Method

The expression "evaluation of a BN" could, in short, be defined as "estimation of the performance of a BN" or "estimation of the quality of recommendations obtained by using a tool based on a BN"[12]. Evaluation constitutes a requisite for the practical application of BNs. Conventional BN evaluation consists of obtaining a set of cases from records or from experts, querying the network for a diagnostic or predictive recommendation for each case, and determining how well the recommendations agree with the actual results known for the cases [1]-[6], [12]-[13]. There are two important issues with regard to the evaluation process of a BN: on the one hand, the selection of the cases and, on the other hand, the method for measuring the performance. The cases can be obtained in two different ways:

- From the BN itself, or
- From a database or with the help of an expert in the domain.

The assessment of performance can be addressed following two distinct strategies:

- By relying on expert opinion to judge the results produced by the BN, or
- By executing a mathematical method whose entries are the cases available and the inferential results.

9 Empirical Results with Constantinou Method

The BN model M and the data-based background are shown for the type of distribution planning (D) and conditional distribution for the solution variables in Table 2 (R). We assume that there is an expert glider. Then R is represented by a set of Condition features of distribution problem described in Table 2. Suppose expert node X now includes states or $u_1, \ldots, u_k$, (where k ≥1) that have been observed. In this case, the problem is that, instead of having to keep the expected value so that:

$$E_M(R|D) = E_{M'}(R|D, X),$$

we only have to ensure that:

$$E_M(R|D) = E_{M'}(R|D, X \text{ not equal to any of } u_1, \dots, u_k).$$

So, Eq. (5) needs only to preserve the data-driven network in model M' under the states of X for which the expert assumes that they are indirectly captured by data and hence, ignore any $u_1, \dots, u_k$. This implies that the states $u_1, \dots, u_k$, which are assumed not to have been captured by data, will now have added impact on R .

Therefore, equation (2) and calculating each of the factors that we obtain the results in terms of a conditional probability:

$E_M(R|D) = 0.35$, $E_{M'}(R|D, X) = 0.27$, thus, $E_M(R|D) \neq E_{M'}(R|D, X)$.

Therefore the final reasoning is that the network based on data in the M' model under the states of X for which the expert assumes that they are captured by a CBR involving the states $u_1, \dots, u_k$ of X. There is no evidence to say that they are the same, so it is concluded that there is a different impact on R in each M and M' model. In our case the conditional probability for the approval of a distribution plan obtained in the M' model is adjusted, so it is realistic that the conditional probability of the M model.

10 Empirical Results with ROC

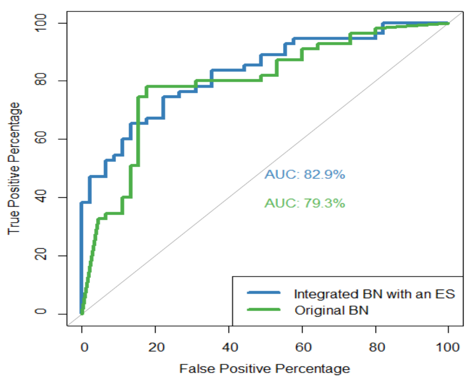

Fig. 6. Comparison of ROC and AUC.

The original Bayesian network obtained an overall model accuracy of 79.3% on validation data and the integrated Bayesian network obtained a slightly higher accuracy of 82.9% (figure 6).

11 Discussion and Conclusions

This research presented an application of a methodology to construct decision support models for BN through the incorporation of a CBR. The main contribution of this application was the incorporation of an ES to a Bayesian network. Applying the integrated Bayesian network learning method could obtain greater accuracy than the original Bayesian network learning method, due to the fact that there are more interconnections between nodes compared to the more dispersed network of the original method. However, the improvement in overall accuracy was only 3.6%.

12 Future Research

Bayesian networks are now well established as a modeling tool for expert systems in domains with uncertainty. The reasons are its powerful but conceptual transparent representation for probabilistic models in terms of a network, there is no doubt of applicability, the persistent problem is the lack of data, so the recommendation is to apply methods that allow incorporating rare or never seen events and give them a treatment so that in the expectations based on the data of the model, under the assumption that these rare or not observed events known are not established as false within the model.

References

1. Cano, A., Masegosa, A.R., Moral, S.: A Method for Integrating Expert Knowledge When Learning Bayesian Networks from Data. IEEE Trans. Systems, Man, and Cybernetics, Part B 41(5), 1382–1394 (2011)
2. Costa Constantinou, A., Fenton, N.E., Neil, M.: Integrating expert knowledge with data in Bayesian networks: Preserving data-driven expectations when the expert variables remain unobserved. Expert Syst. Appl. 56, 197–208 (2016)
3. Hsu, F.-M., Lin, Y.-T., Ho, T.-K.: Design and implementation of an intelligent recommendation system for tourist attractions: The integration of EBM model, Bayesian network and Google Maps. Expert Systems with Applications 39(3), 3257–3264 (2012)
4. Guo, Y., Chen, W., Zhu, Y.-X., Guo, Y.-Q.: Research on the integrated system of case-based reasoning and Bayesian network. ISA transactions 90, 213–225 (2019)
5. Jiang, Z., Jiang, Y., Wang, Y., et al.: A hybrid approach of rough set and case-based reasoning to remanufacturing process planning. J Intell Manuf 30, 19 (2019)
6. Flores, M.J., Nicholson, A.E., Brunskill, A., Korb, K.B., Mascaro, S.: Incorporating expert knowledge when learning Bayesian network structure: a medical case study. Artificial intelligence in medicine 53(3), 181–204 (2011)
7. Tang, K., Parsons, D.J., Jude, S.: Comparison of automatic and guided learning for Bayesian networks to analyse pipe failures in the water distribution system. Reliability Engineering & System Safety, vol. 186, pp. 24–36 (2019)
8. Korb, K.B., Nicholson, A.E.: Bayesian Artificial Intelligence. Second Edition, CRC Press Taylor & Francis Group (2011)
9. Liu, M., Stella, F., Hommersom, A., Lucas, P.J.F., Boer, L., Bischoff, E.: A comparison between discrete and continuous time Bayesian networks in learning from clinical time series data with irregularity. Artificial Intelligence in Medicine 95, 104–117 (2019)
10. Qiu, J., Gu, W., Kong, Q., Zhong, Q., Hu, J.-L.: The emergency response management based on Bayesian decision network. In: Computational Intelligence (SSCI), IEEE Symposium Series on (2016)
11. Neapolitan, R.: Learning Bayesian Networks. Prentice Hall, Series in Artificial Intelligence (2004)

12. Galán, S.F., Arroyo-Figueroa, G., Díez, F.J., Sucar, L.E.: Comparison of Two Types of Event Bayesian Networks: A Case Study. Applied Artificial Intelligence 21(3), 185–209 (2007)
13. Yet, B., Perkins, Z.B., Rasmussen, T.E., Tai, N.R.M., Marsh, D.W.R.: Combining data and meta-analysis to build Bayesian networks for clinical decision support. Journal of biomedical informatics, vol. 52, pp. 373–85 (2014)
14. Na, Y.C., Yang, J.: Distributed Bayesian network structure learning. In: IEEE International Symposium on Industrial Electronics (2010)
15. Yang, Y., Gao, X., Guo, Z., Chen, D.: Learning Bayesian networks using the constrained maximum a posteriori probability method. Pattern Recognition, vol. 91, pp. 123–134 (2019)
16. Aamodt, A., Plaza, E.: Case-based reasoning: foundational issues, methodological variations, and system approach. AI Communications 7(1), 39–59 (1994)
17. López De Mántaras, R., Mcsherry, D., Bridge, D., Leake, D., Smyth, B., Craw, S., Faltings, B., Maher, M.L., Cox, M.T., Forbus, K., Keane, M., Aamodt, A., Watson, I.: Retrieval, reuse, revision, and retention in casebased reasoning. The Knowledge Engineering Review, vol. 00:0, 1–2. Cambridge University Press, United Kingdom (2005)
18. Senoussi, A., Dauzere-Peres, S., Brahimi, N., Penz, B., Kinza Mouss, N.: Heuristics Based on Genetic Algorithms for the Capacitated Multi Vehicle Production Distribution Problem. Computers and Operations Research (2018)
19. Pant, K., Singh,, A.R., Pandey, U., Purohit, R.: A Multi Echelon Mixed Integer Linear Programming Model of a Close Loop Supply Chain Network Design. Materials Today: Proceedings, vol. 5(2), pp. 4838–4846 (2018)
20. Barlett, M., Cussens, J.: Advances in Bayesian Network Learning using Integer Programming (2013)
21. Bartlett, M., Cussens, J.: Integer Linear Programming for the Bayesian network structure learning problem. Artificial Intelligence, vol. 244, pp. 258–271 (2017)