ETM
ENGINEERING & TECHNOLOGY MANAGEMENT

ETM 665: Research Methods in Engineering Management

Instructor: Tim Anderson PhD

Final Report

Study of Transforming Causal Maps to Bayesian Causal Maps

Author: Yulianto Suharto

Spring 2011
June 15 2011

ABSTRACT

The objective of this paper is to introduce the concept of Bayesian causal mapping which is build from causal maps (CMs). CMs provide a rich representation of ideas, through the modeling of complex structures --representing the chain of arguments-- as networks. However, CMs is not easy to define and the magnitude of the effect is difficult to express in numbers. Hence, Bayesian causal maps can be used to make inferences in CMs.

Index Terms— Bayesian Causal Mapping, Causal Maps, Network.

1. INTRODUCTION

In a real-world situation, decision maker(s) utilize the available information for making analysis and reaching decisions. The process of data analysis and decision making can be considered as a prediction process. Liu [1] mentions that in general there are two types of tasks in this process, which require different approaches: (1) classification which is concerned with deciding the nature of a particular system given the features, which usually produces *labeled* data; (2) causal prediction which is concerned with the effect of the changes in some features to some other features in the system.

The later process --causal prediction-- is more related to causal inference which is concerned with the degree of change of feature(s) in the prediction process. This change would directly or indirectly alter some of the features in the data. Hence, Causal Maps (CMs) considered being applicable for this purpose [2]. According to Nadkarni [3], since CMs represent domain knowledge more descriptively than other models such as regression or structural equations, they are more useful as a decision tool.

A Bayesian networks (BNs) is a graphical model that encodes relationships among variables in the system. Spigehalter *et.al* [4] argue that BNs has several advantages for data analysis: (1) BNs readily handles situations where some data entries are missing, (2) BNs can be used to model causal relationships, and hence can be used to gain understanding about a problem domain and to predict the consequences of intervention, and, (3) BNs is an ideal representation for combining prior knowledge (which often comes in causal form) and data since the model has both causal and probabilistic semantics.

BNs can be constructed through two different approaches -- data-based approach and knowledge-based approach [5]. The data-based approaches use conditional independence between variables of interest of Bayes nets to induce models from data. The knowledge-based approach uses expert's judgement in constructing Bayesian networks. The knowledge-based approach is especially useful in situations where domain knowledge is crucial and availability of data is scarce.

2. CAUSAL MAPS (CMs)

In order to understand the effect of the change(s), decision maker(s) must have some mechanisms that can discover the cause and effect relations from the data set. Causal Maps (CMs) is widely known to approach such a problem. Eden et al. [2] defines CMs as a ``directed graph characterized by a hierarchical structure which is most often in the form of a means/end graph". In the last decades, CMs have been widely used to construct a framework and represent major factors, knowledge and conditions that influence decision making process [5,6].

Causal relationships can be either positive or negative, as specified by a '+', respectively a '-', sign on the arrow connecting two variables. The variables that cause a change are called cause variables and the ones that undergo the effect of the change are called effect variables [7].

CMs provide a rich representation of ideas, through the modeling of complex structures --representing the chain of arguments-- as networks [2,3]. Often times the last stage of intervention process is to identify and agree to a set of potential strategic options. In some cases, the preferred direction may emerge naturally from a process of negotiation; in others further, more or-less formal, analysis to evaluate the options and to understand their impacts on the goals could be helpful [8]. CMs can provide us to look at the problem more extensively than other decision tools which consider causal relations, such as regression. CM has been widely used in international relations, administrative science, political science, sociology, policy analysis, organizational behavior and management [1-3,5,8-10].

One major concern that needs to be addressed in CMs is that CMs are not easy to define and the magnitude of the effect is difficult to express in numbers. In general, CMs are constructed by gathering information from experts. These experts are more likely subjectively express themselves in qualitative rather than quantitative terms [7] . Kosko [11] introduced the concept of Fuzzy CMs (FCMs) to overcome the problem. FCM represents the concepts linguistically with an associated fuzzy set. FCM is a signed directed graph that allows feedback and employs concepts (nodes) and weighted edges between concepts [12]. The degree of relationship between concepts in an FCM is either a number in [0; 1] or [-1; 1], or a linguistic term, such as ’often’, ’extremely’, ’some’, etc [7].

3. BAYESIAN NETWORKS (BNs)

3.1 Definition

In the eighteenth century, Bayes’ Theorem is developed by Thomas Bayes (1702—1761); since then the theory had a major effect on statistical inferences. The probability of a cause is inferred by Bayes Theorem when effect of cause is observed. The theorem was expanded in time. It has been used as a cause and effect diagram since the end of twentieth century [13]. Some of the advantages in using Bayesian Networks (BNs) are: (1) BNs can handle incomplete data sets (2) BNs focus on causal relationship and then facilitate the combination of background knowledge and experimental data in a way that the process can avoid over fitting problem [3,4].

BNs is a model in which events are connected to each other with probabilities. This model can be anything; for example, economic reasons, vehicle parts, ecosystem etc.., which can be modeled with Bayes. If the probabilities of events which affect each other are known exactly, the achievements are closer to the true results [14].

BNs is a directed acyclic graph (DAG) which means there are no cycles. In other word, BNs is a probabilistic graphical model that restricts the graph to be directed and acyclic. Other models such as Markov random fields (MRFs) have no such restrictions [15,16]. If there is a link between A and B (A→B) we say that B is a child of A and A is a parent of B [17,18]. In BNs, a link from node A to node B does not always imply causality. It implies a direct influence of A over B and the probability of B is conditioned on the value of A [19,20].

3.2 Conditional Probabilities in BNs

The direction of the arrows in BNs can be explained with causality as long as arrows do not cause an endless loop. The advantage in comparison to other statistical models such like regression is that casualty can supply missing information and details as well as bringing the priorities and key factors into focus [10]. Besides, the network is constructed in such a way that in the beginning all factors have the same certainties.

If A and B are the occurrences of two factors Bayes rule is defined as follows:

$$P(B|A) = \frac{P(A|B) * P(B)}{P(A)}$$

Where, P(A) gives the probability of the occurrence of factor A and P(A/B) is the probability of the occurrence of A when B event is occurred. Hence, the link from node A to node B means that factor A has a direct affect on factor B. Furthermore, the probability of B depends on the probability of A [21]. Each node of the network is annotated with a conditional probability distribution (CPD) that represents $P(X_i \mid Pa(X_1))$, where $Pa(X_i)$ denotes the parents of X_i. The pair (G, CPD) encodes the joint distribution $P(X_1, \ldots X_n)$. A unique joint probability distribution over X from G is factorized as:

$$p(X_1,\ldots,X_n) = \prod_i (p(X_i \mid Pa(X_i)))$$

BNs helps us to observe whole structure of factor interactions from a graph. This is the way marginal and conditional probabilities of the factors can be computed by marginalizing over the joint [22].

3.3 Probabilistic inferences

Probabilistic inferences about variables in the model can be drawn once a BNs is constructed. The conditionals given in a BNs representation specify the prior joint distribution of the variables. If we observe (or learn about) the values of some variables, then such observations can be represented by tables where we assign 1 for the observed values and 0 for the unobserved values. Then the product of all tables (conditionals and observations) gives the (un-normalized) posterior joint distribution of the variables. Thus, the joint distribution of variables changes each time we learn new information about the variables.

A simple system such as in Figure 1 illustrate the concept of probabilistic inferences in BNs. If there is an arc pointing from A to B, we say A is a parent of B. For each variable, we need to specify a table of conditional probability distributions, one for each configuration of states of its parents. Figure 1 shows these tables of conditional distributions -- P(A), P(B|A), P(C) and P(D|B,C).

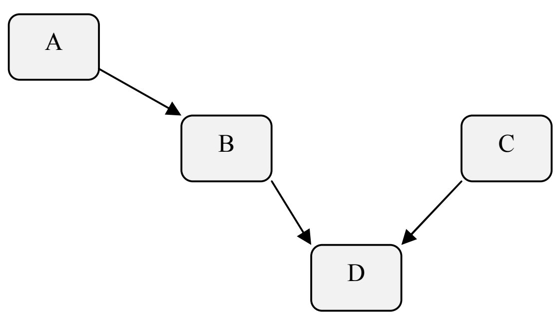

Figure 1 Graphic Representation of BNs

A fundamental assumption of a BNs is that when we multiply the conditionals for each variable, we get the joint probability distribution for all variables in the network. For Figure 1 we make assumption that: $P(A,B,C,D) = P(A)*P(B|A)*P(C)*P(D|B,C)$.

One can read these conditional independence assumptions directly from the BNs graph as follows. Suppose we pick a sequence of the variables such that for all directed arcs in the network, the variable at the tail of each arc precedes the variable at the head of the arc in the sequence. Since the directed graph is acyclic, there always exists such a sequence. In Figure 1, one such sequence is A B C D. Then, the conditional independence assumptions can be stated as follows. For each variable in the sequence, we are assuming that it is conditionally independent of its predecessors in the sequence given its parents. The essential point here is that missing arcs (from a node to its successors in the sequence) signify conditional independence assumptions. Thus the lack of an arc from A to C indicates that C is independent of A; the lack of an arc from B to C indicates that B is independent of C; and the lack of an arc from A to D indicates that D is conditionally independent of A given B and C [3].

4. TRANSFORMING CMS TO BNS

Both BNs and CMs are causal models that represent cause - effect beliefs of experts. However, there are some differences in the two approaches to modeling that need to be addressed if we are to transform CMs to BNs. These differences are discussed in the following paragraphs.

4.1 Conditional independencies

Pearl [23] states that a network model can be either a dependence map (D-map) or an independence map (I-map). In a D-map, an arrow between two variables in the model implies that the two variables are related. However, a lack of an arrow between variables does not necessarily imply independence between the two variables. An I-map, on the other hand, implies that concepts found to be separated are indeed conditionally independent, given other variables. Hence, CMs is a D-map since CMs is a directed graph that depicts causality between variables and also in CMs an arrow between two variables implies dependence. However, the absence of an arrow between two variables in CMs does not imply a lack of dependence. There is a possibility that the absence of the arrow resulted from the lack of articulation of the expert's judegement. It does not necesary imply that the expert believes that the variables to be independent [5].

BNs, on the other hand, is an I-map. Hence, an absence of arrow from a variable to its *child* indicates conditional independence between the variables. Thus, when we want to transform CMs to BNs, it is important to ensure that the lack of links between the concepts in the causal maps implies independence and the presence of links between concepts implies dependence [3][5].

4.2 Reasoning underlying cause–effect relations

It is believed that from a logic or reasoning process standpoint, individuals perceive cause - effect relationships based on two types of reasoning: deductive and abductive [3]. A reasoning is called deductive if we reason from causes to effects. Abductive reasoning, on the other hand, happen when we reason from effects to causes.

A distinction between deductive and abductive reasoning behind the causal linkages is essential to establish accurate directions of linkages in CMs. The emphasis in deriving CMs should be on the causal theory underlying the causal statements rather than the language used [8][3].

4.3 Direct Vs. Indirect Relations

In CMs a direct link between two variables does not guarantee a direct relationship between the two variables. It just implies a relation between the two variables that can be either direct or indirect. This distinction is important to identify conditional independencies in the CMs [3][5][8]. Figure 2 ilustrates the distinction of direct and indirect relationship and how a lack of distinction affects conditional independence assumptions in a CMs.

Figure 2 Direct Vs. Indirect Relations

In Figure 2 on the Original CMs, both A and B affect C while in the Bayesian CMs, there is no linkage between A and C, implying that A affect C strictly through B. If we have complete information of B, any additional information of A will be irrelevant in making inferences about C.

4.4 Circular loop and reciprocal influences are not permitted in BNs
As explained in the previous paragraph, CMs is directed graphs that characterized by an acyclic structure. However, circular relations or causal loops destroy the hierarchical form of a graph. Circular relations in the CMs violate the acyclic graphical structure required in BNs. It is therefore essential to eliminate circular relations to make CMs compatible with BNs. Causal loops can exist for two reasons. First, they may be coding mistakes that need to be corrected. Second, they may represent dynamic relations between variables across multiple time frames [2-5,8,24].

5. CONSTRUCTING BAYESIAN CAUSAL MAPS

A procedure to construct Bayesian causal maps can be summarized into 4 steps [3][5]:
1. Data elicitation
2. Derivation of CMs
3. Modification of CMs to construct Bayesian causal maps
4. Derivation of the parameters of Bayesian causal maps

Data-based or knowledge-based approaches or a combination can be used for data elicitation purpose. Hence, literature review and/or expert's opinion are used to determine the variables of interest for constructing the original CMs. Based on the elicitate data, the second step is to construct an original CMs. In the third step, the CMs of the expert is modified-- using the approaches of transforming CMs to BNs as explained in the above paragraphs which include: conditional independencies, reasoning underlying the link between concepts, distinction between direct and indirect relations, and eliminating circular relations -- to eliminate biases that result from the use of textual analysis and to make the structure of the CMs compatible with BNs. In the final step, the parameters of the Bayesian causal maps are derived using probability-encoding techniques [3].

In step three--modification of CMs to construct Bayesian causal maps--two most widely used methods are structured interviews and adjacency matrices [3,8,13]. In structured interviews, the experts are provided a list of paired concepts as well as different alternative specifications of the relation between the concepts in the original map. The experts are then instructed to choose an alternative to specify the direct relation between the pair of concepts. Adjacency matrices, on the other hand, experts are provided the concepts in the form of an adjacency matrix, where the rows represent causes and columns represent effects. The experts are asked to enter '0'(no relation), ' +1 ' (positive relation), or '-1' (negative relation) in each cell to specify the relation between two concepts in the matrix. These two structured methods help in removing the four modeling biases relating to the construction of Bayesian causal maps.

For the last step, once the structure of the Bayesian causal maps is constructed, numerical parameters of this modified structure need to be assessed so that the propagation algorithms in the Bayesian network can be used to make inferences [5][13]. For this purpose, data-based (historical data) and knowledge-based approaches (expert's opinion) can be utilized to get the parameter (prior and conditional probabilities of the variables of interest).

6. A CASE STUDY: CLEAN ENERGY INVESTMENT [25]

This section describes the construction of Bayesian causal map for a specific case study in clean energy investment [25]. First, the paper illustrates how starting from a CMs, then constructed the qualitative structure of a Bayesian causal map. Additional information -- collected from experts to address the modeling issues discussed in Section 4.1 as well as to derive the numerical parameters of the Bayesian causal map -- was also presented here. Second, Bayesian network software is also introduced to draw probabilistic inferences in a Bayesian causal map.

6.1 Decision Context

The decision faced for the presented case study here is whether decision maker(s) should invest in the nuclear energy or other renewable energy resources in regards to so many variables affecting the decision option. The role of the system analyst was to analyze the decision and suggest a recommendation.

6.2 Procedure for constructing a Bayesian causal map

6.2.1 Step 1: Data Elicitation

Through literature review and expert's validation, the authors observed that there are twenty-nine decision driving forces covering ecological, economical, technological and social variables. These factors are eliminated to eleven factors by a cognitive map study.

6.2.2. Step 2: Derivation of CMs

Adjacency matrix --as explained in Section 5-- is presented to the experts to specify the relation between two concepts in the matrix. By taking the mode of the expert's responses. The complete adjacency matrix is presented in Table 1. This process resulted in the original of CMs as shown in Figure 3.

Table 1 Interrelation of Energy Investment Criteria.

	Renewable Energy Investments	Nuclear Energy Investments	Primary Energy Consumption	Primary Energy Import	Renewable Energy Production	Fossil Fuel Production	GDP per Capita	Population	Urbanization	Industrialization	Greenhouse Emission
Renewable Energy Investments	0	-1	0	-1	1	-1	0	0	0	1	-1
Nuclear Energy Investments	-1	0	0	1	-1	-1	0	0	1	1	-1
Primary Energy Consumption	1	1	0	1	1	1	1	0	1	1	1
Primary Energy Import	1	1	1	0	1	1	0	0	1	0	1
Renewable Energy Production	1	-1	0	-1	0	-1	1	1	1	1	-1
Fossil Fuel Production	-1	0	1	1	-1	0	1	1	1	1	1
GDP per Capita	1	1	1	1	1	1	0	1	1	1	-1
Population	1	1	1	1	1	1	-1	0	1	1	1
Urbanization	1	1	1	1	1	1	1	1	0	1	0
Industrialization	1	1	1	1	1	1	1	1	1	0	1
Greenhouse Emission	1	1	0	-1	1	-1	-1	-1	0	-1	0

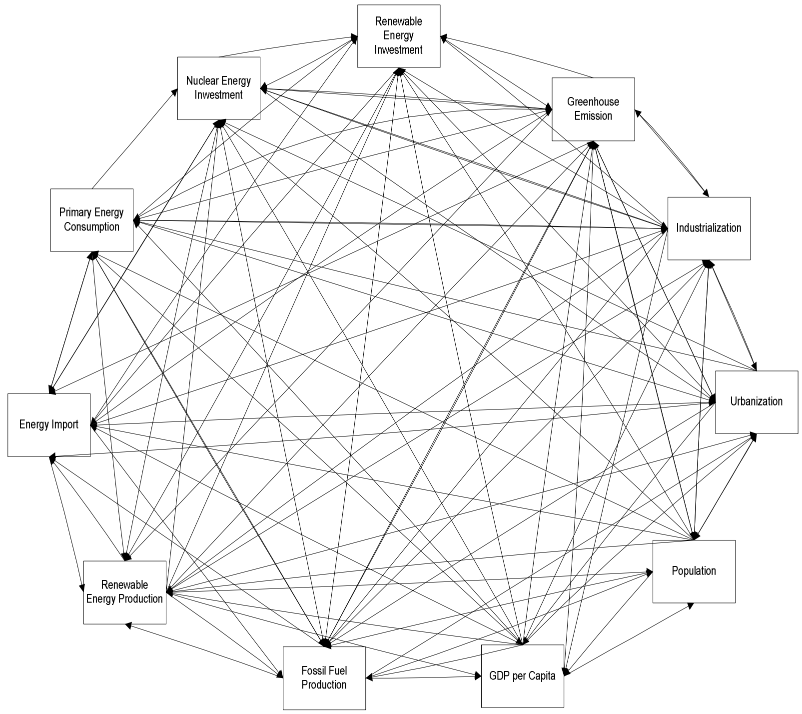

Figure 3 Original CMs

6.2.3. Step 3: Modification of original CMs

Utilizing approaches as explained in Section 4 when we want to convert the CMs into BNs--by looking at four basic requirements: conditional independencies; reasoning underlying the link between concepts; distinction between direct and indirect relations; and eliminating circular relations--, the orginal CMs then converted into BNs. The final result of this processes is shown in Figure 4.

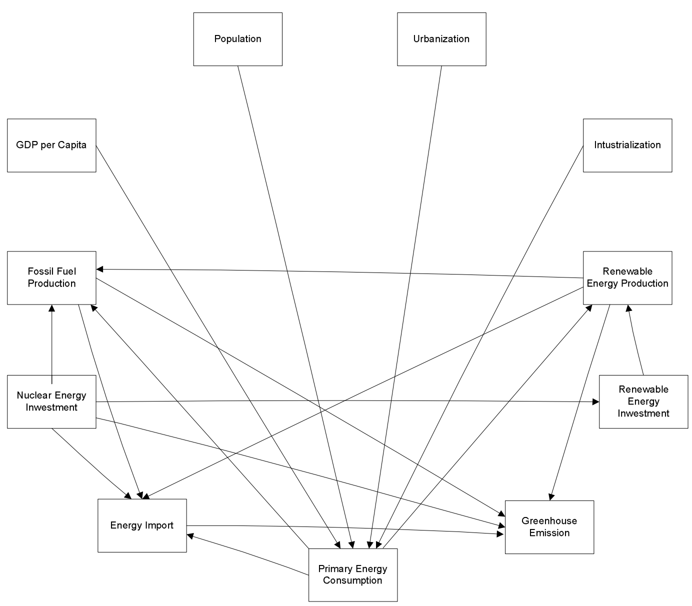

Figure 4 Modified CMs

6.3.4. Step 4: parameters Assessment

In this step, the parameters of the Bayesian Causal Map --which consist of marginal probabilities and conditional probabilities-- were assessed using discrete scale of 0-10.

6.4 Validating the Bayes net Model

A graphical software package, Netica [14] is utilized to make probabilistic inferences using sum propagation. The sum propagation computes the marginal probabilities of all the model variables and updates the marginals with all additional evidence received about other variables [3,5,13]. In our case study, we can evaluate each energy investment alternatives under different scenarios.

The scenarios were defined in consultation between the authors and the experts, and they represent situations in which there are unambiguous prescriptions for energy investment decision in the energy literature.

One example of those scenarios is presented here to demonstrates how probabilistic inferences is drawn from the model by utilizing Netica. İn Figure 4 of the modified CMs, suppose that two nodes, population and industrialization are predicted to be at their highest states, what will happen to greenhouse emission and energy import when the decision maker(s) decided to invest whether in nuclear energy or other renewable energy resources? Table 2 summarized the prior and posterior probabilities for all variables.

Table 2 Prior and Posterior Marginal Probabilities under Different Scenarios

Variables	States	Prior	Posterior Marginal on Scenario 1*	
			With Nuclear Investment	With Renewable Energy Investment
GDP	low	10	0	0
	medium	60	0	0
	high	30	100	100
Population	low	28.5	28.5	28.5
	medium	43	43	43
	high	28.5	28.5	28.5
Urbanization	low	50	50	50
	high	50	50	50
Industrialization	low	50	0	0
	high	50	100	100
Primary Energy Consumption	low	34.6	19.6	19.6
	medium	25.8	10.7	10.8
	high	39.6	69.6	69.6
Renewable Energy Production	low	44.8	73.5	6.48
	medium	8.63	3.55	3.55
	high	46.5	23	90
Fossil Fuel Production	low	43.7	47.5	36.3
	medium	11.6	4.83	4.83
	high	44.8	47.7	58.9
Energy Import	low	**50**	**42.5**	**51.1**
	high	**50**	**57.5**	**48.9**
Greenhouse Emission	low	**50.5**	**46.9**	**57.6**
	high	**49.5**	**53.1**	**42.4**

*Scenario 1: GDP and Industrialization is at its highest states

It is noticed from Table 2 that once we exposed our scenario to the original/initialized Bayesian causal maps, sum propagation method will compute the posterior probabilities of the nodes/variables of interest so we can compare the result to the prior probabilities. In this case study, it is noticed that our scenario with the nuclear energy investment will result in the increase of both greenhouse emission and energy import--these 2 variables are affected by more variables compared to other variables as shown in Figure 4--, a result that obviously not in favor for the decision maker(s). While the same scenario running with renewable energy will result in lower energy import and greenhouse emission. The same approach can be used with different scenarios to find the most favorable policy for the decision maker(s).

7. SUMMARY AND CONCLUSİON

Most of the focus in CMs has been in its use for knowledge representation [3]. This study enables decision-makers to use causal maps for decision making by converting the map into Bayesian causal maps. Once a Bayesian network is constructed, it can be used to make probability inferences about the variables in the model. The modified map was validated qualitatively, through expert's consensus, and quantitavely through examining the posterior probabilities of decision variables under different scenarios. BNs is quite an effective method to conclude the plans which have complex structure.

Bayesian causal maps provide a framework for representing the uncertainty of variables in the map as well as the effect of variables. Bayesian causal maps combine the strengths of causal maps and Bayesian networks and reduce the limitations of both. Using concepts from the literature on causal modeling and logic [8,13,23], Bayesian causal maps clarify the cause - effect relations depicted in the causal maps. They depict dependence between variables based on causal mapping approach (D-map) as well as a lack of dependence between variables based on the Bayesian network approach (I-map) [3,5,8,11,13,23].

References

[1] Z.-qiang Liu, "Causation, bayesian networks, and cognitive maps," *Zidonghua Xuebao/Acta Automatica Sinica*, vol. 27, 2001, pp. 552-566.

[2] C. Eden, "Analyzing cognitive maps to help structure issues or problems," *European Journal of Operational Research*, vol. 159, Dec. 2004, pp. 673-686.

[3] S. Nadkarni, "A Bayesian network approach to making inferences in causal maps," *European Journal of Operational Research*, vol. 128, Feb. 2001, pp. 479-498.

[4] D.J. Spiegelhalter, A.P. Dawid, S.L. Lauritzen, and R.G. Cowell, "Bayesian Analysis in Expert Systems," *Statistical Science*, vol. 8, 1993, pp. 219-241.

[5] S. Nadkarni, "A causal mapping approach to constructing Bayesian networks," *Decision Support Systems*, vol. 38, Nov. 2004, pp. 259-281.

[6] S. Sahin, F. Ulengin, and B. Ulengin, "A Bayesian causal map for inflation analysis: The case of Turkey," *European Journal of Operational Research*, vol. 175, Dec. 2006, pp. 1268-1284.

[7] B. Lazzerini and L. Mkrtchyan, "Risk Analysis Using Extended Fuzzy Cognitive Maps," *2010 International Conference on Intelligent Computing and Cognitive Informatics*, 2010, pp. 179-182.

[8] G. Montibeller and V. Belton, "Causal maps and the evaluation of decision options—a review," *Journal of the Operational Research Society*, vol. 57, May. 2006, pp. 779-791.

[9] C. Eden, F. Ackermann, and S. Cropper, "The analysis of cause maps," *Journal of Management Studies*, vol. 29, 1992, p. 309–324.

[10] K. Siau and X. Tan, "Improving the quality of conceptual modeling using cognitive mapping techniques," *Data & Knowledge Engineering*, vol. 55, Dec. 2005, pp. 343-365.

[11] B. Kosko, "Fuzzy cognitive maps," *Int.J.Man-Machine Studies*, vol. 24, Jan. 1986, pp. 65-75.

[12] Y. Huang, L. Ni, and Y. Miao, "A Quantum Cognitive Map Model," *2009 Fifth International Conference on Natural Computation*, 2009, pp. 28-31.

[13] R.E. Neapolitan, "Chapter 5 - Foundations of Bayesian Networks," *Probabilistic methods for bioinformatics: with an introduction to Bayesian networks*, Burlington, MA: Morgan Kaufmann, 2009.

[14] Norsys Software Corp, "Netica™ Application," 2011, pp. 85-133.

[15] E.N. Mortensen, "Real-Time Semi-Automatic Segmentation Using a Bayesian Network," *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1 (CVPR'06)*, 2006, pp. 1007-1014.

[16] Shu-Chun Ho, "Construct a Sequential Decision-Making Model : A Dynamic Bayesian Network Perspective," *Proceedings of the 44th Hawaii International Conference on System Sciences*, 2011, pp. 1-10.

[17] Y. Zeng, Y. Xiang, and S. Pacekajus, “Refinement of Bayesian Network Structures upon New Data.pdf,” *2008 IEEE International Conference on Granular Computing, GRC 2008*, 2008, pp. 772-777.

[18] L. Jingjing, P. Yan, and Z. Ting, “The Application of Bayesian Network in the Performance Evaluation and Decision-making System,” *2008 IEEE International Conference on Networking, Sensing and Control*, 2008, pp. 180-183.

[19] G. Thibault, S. Bonnevay, and A. Aussem, “Learning Bayesian network structures by estimation of distribution algorithms: An experimental analysis,” *2007 2nd International Conference on Digital Information Management*, 2007, pp. 127-132.

[20] E. Alpaydin, *Introduction to Machine Learning (Adaptive Computation and Machine Learning)*, The MIT Press, 2004.

[21] X. Changliang and S. Zhanfeng, “Wind energy in China: Current scenario and future perspectives,” *Renewable and Sustainable Energy Reviews*, vol. 13, Oct. 2009, pp. 1966-1974.

[22] P. Trucco, “A Bayesian Belief Network modelling of organisational factors in risk analysis: A case study in maritime transportation,” *Reliability Engineering System Safety*, vol. 93, 2008, pp. 845-856.

[23] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, 1988.

[24] D. Heckerman, *A Tutorial on Learning with Bayesian Networks, in Learning in Graphical Models*, Cambridge, MA: MIT Press, 1999.

[25] G. Kayakutlu, Y. Bayram, T. Daim, and Y. Suharto, “Clean Energy Investment Scenarios using Bayesian Network,” *Seven*, 2011, pp. 1-16.