Bayesian Network Inference with Qualitative Expert Knowledge for Group Decision Making

Wichian Premchaiswadi
Graduate School of Information Technology in Business
Siam University
Bangkok, Thailand
wichian@siam.edu

Nipat Jongsawat[1], Walisa Romsaiyud[2]
Graduate School of Information Technology in Business
Siam University
Bangkok, Thailand
nipatj@yahoo.com[1], walisa.r@siamu.ac.th[2]

***Abstract*—In this paper, we propose a practical framework and a methodology for transforming final group decision making statements or expert knowledge into a set of qualitative statements and probability inequality constraints for inference in a Bayesian Network. First we present several techniques in the decision process to produce a group preference ranking and a final group solution. Second, we describe a detailed method to transform the final group solution into a set of qualitative statements and an "a priori" distribution for Bayesian probabilistic models. The decision-making assumptions and the mathematical equation for Bayesian inference are derived based on data and knowledge obtained from experts. We also propose a simplified method for constructing the "a priori" model distribution. Each statement obtained from the group's solution is used to constrain the model space to the subspace which is consistent with the statement provided. Finally, we present qualitative knowledge models and then show a full formalism of how to translate a set of qualitative statements into probability inequality constraints.**

***Keywords*—Bayesian network, Bayesian network inference, decision-support systems, group preference schedule, group preference ranking, group qualitative expert knowledge, probability inequality constraints**

I. INTRODUCTION

In group decision making, different experts often think about the same problem in quite different ways. They frequently have different opinions for decision making about the same situation. Using a Bayesian network structure for optimizing problems, different experts who work as a group for projects may have different solutions for indentifying the causal relationships among variables in the BN model and quantifying graphical models with numerical probabilities. For example, expert-1 may state that "making a decision in situation A causes situation B and making a decision in situation B causes situation C". But expert-2 may state that "making a decision in situation B causes situation A and making a decision in situation A causes situation C". Even in a simple case of decision making, the expert knowledge obtained from different experts is quite different. It is typically not possible to avoid contradictions among different expert's solutions in group decision making.

In this paper, we propose a practical framework and a methodology for transforming expert knowledge or final group decision making statements into a set of qualitative statements and probability inequality constraints for inference in a Bayesian Network. First, we need to identify a set of alternatives on which the experts have opinions and then consider the problem of constructing a group preference ranking. If such a group preference ranking can be created, then one could utilize the alternative at the top of the ranked list the alternative preferred by the group. Second, after we obtain the most preferred alternative or statement such as "A causes B and then B causes C" from the group decision making, we propose a formal method to transform knowledge, represented by a set of qualitative statements, into an a priori distribution for Bayesian probabilistic models. The mathematical equation for Bayesian inference is derived based on knowledge obtained from the final group decision statements. The set of model parameters, consistent with the statements, and the distribution of models in the structure-dependent parameter space are presented. We also propose a simplified method for constructing the "a priori" model distribution. Each statement obtained from the experts is used to constrain the model space to the subspace which is consistent with the statement provided. Finally, we present qualitative knowledge models and then show a complete formalism of how to translate a set of qualitative statements into probability inequality constraints. Several cases of Bayesian influence are classified and the probability inequality constraints presented in each case are described.

This paper is organized as follows: Section II presents more detail about the background of Bayesian networks and some perspectives of qualitative causal relationships in the Bayesian approach. Section III addresses the method of constructing a group preference ranking and group decision making from the individual preferences obtained from the experts performing group work. Section IV addresses the methods to transform a final solution or expert knowledge into an "a priori" distribution for Bayesian probabilistic models in more detail. Section V describes the method used to translate a set of qualitative statements into probability inequality constraints and presents different cases of influences in Bayesian network model. Section VI presents a conclusion and discusses some perspectives and ideas for future work.

II. BACKGROUND

Bayesian networks (also called belief networks, Bayesian belief networks, causal probabilistic networks, or causal networks) are acyclic directed graphs in which nodes represent random variables and arcs represent direct probabilistic dependencies among the nodes [1]. Bayesian networks are a popular class of graphical probabilistic models for research and application in the field of artificial intelligence. They are motivated by Bayes' theorem [2] and are used to represent a joint probability distribution over a set of variables. This joint probability distribution can be used to calculate the probabilities for any configuration of the variables. In Bayesian inference, the conditional probabilities for the values of a set of unconstrained variables are calculated given fixed values of another set of variables, which are called observations or evidence. Bayesian models have been widely used for efficient probabilistic inference and reasoning [1], [3], and numerous algorithms for learning the Bayesian network structure and parameters from data have been proposed [4], [5], [6]. The causal structure and the numerical parameters of a Bayesian network can be obtained using two distinct approaches [7], [14]. First, they can be obtained from an expert. Second, they can also be learned from a dataset or data residing in a database. The structure of a Bayesian network is simply a representation of independencies in the data and the numerical values are a representation of the joint probability distributions that can be inferred from the data [8, 9]. In practice, some combination of these two approaches is typically used. For example, the causal structure of a model is acquired from an expert, while the numerical parameters of the model are learned from the data in a database.

For realistic problems, the database is often very sparse and hardly sufficient to select one adequate model. This is considered as model uncertainty. Selecting one single model can lead to strongly biased inference results. On the other hand, in science and industry, there is an enormous amount of qualitative knowledge available. This knowledge is often represented in terms of qualitative causal relationships between two or more entities. For example, in the statement: "smoking increases the risk of lung cancer," the two entities: smoking and lung cancer are related to each other. Moreover, the smoking entity positively influences the lung cancer entity since lung cancer risk is increased in the case of smoking. It is therefore desirable to make use of this body of evidence in probability inference modeling.

III. Group Preference Ranking and Group Decision Making

In this section, we present the first step which is identifying the group solution for a BN model of the proposed framework (see Fig. 1). Several methods are described for experts to make decisions for identifying the relationship between variables in a Bayesian network model and arriving at a final BN solution representing the group.

The general case is one in which we have a group of experts and a set of alternatives, for example "*A activates B and B activates C*", "*B activates A and A activates C*", and "*C activates A and A activates B*", on which the experts have opinions. We assume that each expert has a preference ranking on the set of alternatives. That is, using these preferences, each expert can order the alternatives in a list such that if alternative A activates B, and B activates C are in the list, then the experts have an agreement with that alternative. A set of individual preference rankings, one for each expert in the group, is called a group preference schedule. One goal of the first portion of our proposed practical framework is to consider the problem of constructing a group preference ranking from the individual preferences (that is, from the group preference schedule). If such a group preference ranking can be created, then one could call the alternative at the top of the group list the alternative selected by the group of experts. However, such a group ranking may not be possible, and moreover, even if it is possible, the alternative at the top of the list may not be one that would win the majority selection in an election among all options. Thus the second goal of our work in the first step is to consider other possible ways of picking the most preferred choice, especially if none of the alternatives would receive a majority selection in an election among all alternatives. We will identify the properties of the decision process that corresponds to our ideas about the characteristics such decision processes should have.

Example 1. Suppose that we have a group of three experts, labeled expert-1, expert-2, and expert-3, and a set of three variables, labeled A, B, and C. For this example, assume the individual preference rankings are as follows:

Expert-1: *A→B→C*; **Expert-2**: *B→C→A*; **Expert-3**: *C→A→B*

Using pairwise comparisons and a simple-majority rule, we see that both expert-1 and expert-3 agree that "*A causes B*", and therefore, because the vote is 2 to 1, the group should agree with "*A causes B*". Therefore, on the basis of this information, we would propose that the group preference ranking should be "*A causes B and then B causes C*; (*A→B→C*)". However, both expert-2 and expert-3 agree with "*C causes A*", and therefore the group should agree with "*C causes A*". We conclude that the proposed group preference ranking in this example is not transitive: The experts agree with *A→B→C→A*. This cyclic, or intransitive, behavior is normally considered unacceptable for a preference ranking. We conclude that even in this simple situation, the majority rule decision process can lead to unacceptable preference rankings. The intransitive phenomena do occur when the number of variables and alternatives increase. That is for many groups and sets of preferences, the group preferences determined by the pairwise majority rule voting are intransitive. What are some ways to cope with the results of this example?

Let's consider again the simple situation of three experts and three alternatives. Then each expert has 6 different preference rankings-that is, 6 ways in which the 3 alternatives can be listed: 3 choices for the alternative listed first, 2 choices for the alternative listed second, and 1 choice for the alternative listed last. Because there are 6 experts, there are 6 x 6 x 6 = 216 different preference schedules for the group. The likelihood of intransitive group preferences depends on how the experts select their individual preference rankings. For instance, if we know that two of the experts have the same preference ranking, then that preference ranking will be the preference ranking for the group, and intransitivity will not occur. As another example, if two experts have alternative Z as the top choice, then intransitive group preferences will never occur. However, intransitive group preferences can still occur if experts select their individual preferences at random. This situation is more complicated but it is not considered in this paper because we assume that the experts use their own experience to make their own decisions. They will not make a decision at random.

In light of this discussion about the difficulties encountered with simple-majority voting, we look for other ways to achieve our primary goal of finding ways for groups to make decisions. We introduce the concept of sequential voting or selection: a sequence of votes where at each vote, a choice is to be made between two alternatives. In any situation with an odd number of experts, this process always yields a result, and this result can be taken as a most preferred alternative. However, as we show in an example below, this method also has problems.

Example 2. Suppose that the relationship between variables is to be identified by first considering, for example, A and B and then considering the impact on the last variable. Expert-1 considers A and B first and states that "*B causes A and then A causes C*": *B→A→C*. Expert-2 considers "*A and C first and state that A causes C and then C causes B*": *A→C→B*. Expert-3 considers "*B and C first and state that C causes B and then B causes A*": *C→B→A*. The results in this example show that we are in the unfortunate situation of having a group preference that depends on the sequence in which the selections were taken.

We have illustrated some of the problems with simple-majority rule and sequential selecting decision processes. We turn next to another approach to the problem: assigning points to a pair of variables of each order on the basis of their relative rankings and defining a group preference ranking by adding the points assigned to each alternative by all experts.

Example 3. We will illustrate the technique by considering five experts and four variables (See Table I). Each expert makes a series of decisions at each order-level. For example, expert-1 makes a decision that "*making a decision in situation A causes situation C*" in a first order level, "*C causes B*" in a second order level, and "*B causes D*" in a third order level. Each expert assigns 3 point to the first order level, 2 point to the second order level, and so on. For a specific alternative, add the points assigned by all experts. The alternative with the most points is the most preferred, the alternative with the second largest number of points is the second most preferred, and so on.

This method is known as the Borda count group decision process [15]. We observe that this decision process has an implicit relative strength of preferences. The relative strengths of all preferences are the same.

TABLE I. A GROUP OF FIVE EXPERTS AND FOUR ALTERNATIVES

Order	Expert1	Expert2	Expert3	Expert4	Expert5	Points
1	A→C	D→A	B→A	C→B	A→C	3
2	C→B	A→C	A→C	B→D	C→D	2
3	B→D	C→B	C→D	D→A	D→B	1

Remark: The group preference ranking is obtained by adding the points assigned to each alternative. (*A→C*:10 points, *C→B*:6 points, *D→A*:4 points, *B→D*:3 points, *B→A*:3 points, *C→D*:3 points, *D→B*:1 points)

We conclude that the group preference ranking is "*A causes C, C causes B, and B causes D*". The alternative "*D→A*" has 4 points but it is not included because A is a parent node in the first order level so that D cannot cause A.

By considering a few examples, we have identified shortcomings of some common decision processes in group decision making. With the last technique, problems are still possible to occur when two alternatives at the same level have the same score. However, this section proposes several techniques in the decision process to produce a group preference ranking and a final group solution.

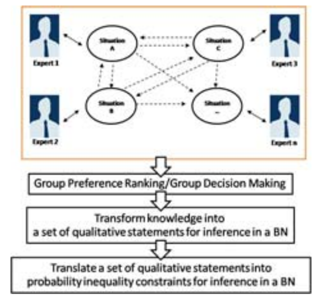

Fig. 1 A practical framework.

IV. METHODS

We describe a methodology to use qualitative expert knowledge obtained from the previous step for inferencing in a Bayesian network. We proceed from the decision-making assumptions and the general equation for Bayesian inference based on final group decision making statements obtained from the experts to a detailed method to transform knowledge, represented by a set of qualitative statements, into an a priori distribution for Bayesian probabilistic models.

For simplicity, let's consider a simple case of decision making in which the body of expert knowledge ω consists of a single statement ω = "*making a decision in situation A causes situation B*". We know that there are 2 random events or variables A and B, which we assume are binary, and we need to consider the set of all possible Bayesian models on A and B. The set of possible model structures are described in the following categories: 1) S_1: A and B have no causal relationship between them, 2) S_2: A and B have some causal relationship between them but the direction of influence cannot be identified, 3) S_3: A causes B, and 4) S_4: B causes A. "*making a decision in situation A causes situation B*" directly states a causal influence of A on B. We use the statement "*A activates B*" to constrain the space of structures: $P(S_3|\omega) = 1$; $P(S_n|\omega)=0$, n=1,2,4). The ω is represented as a qualitative statement described by the expert, A causes B. The graph structure (S_3) encodes the probability distribution

$$P(A,B) = P(B|A)P(A) \quad (1)$$

No further information on P(A) is available; however, P(B|A) can be further constrained. The corresponding Conditional Probability Table (CPT) is shown in Table II.

TABLE II. CONDITIONAL PROBABILITY TABLE

A	P(B=1|A)
0	θ_0
1	θ_1

The values of the conditional probabilities from the components of the parameter vector $\theta = (\theta_0, \theta_1)$ of the model class with structure S_3. θ_0 is the probability of B is active when A is not active. θ_1 is the probability of B is active when A is active. From the statement, we now can infer that the probability of B is active when A is active is higher than the same probability with A inactive. The P(B|A) when P(A) is available is higher than the P(B|A) when P(A) is not available. The inequality relationship is obtained as follows:

$$P(B=1|A=1) \geq P(B=1|A=0),\ \theta_1 \geq \theta_0 \quad (2)$$

Hence, the set of model parameters consistent with that statement is given by

$$\Theta_3 = \{(\theta_0, \theta_1)|\ 0 \leq \theta_0 \leq 1 \wedge \theta_0 \leq \theta_1 \leq 1\} \quad (3)$$

and the distribution of models in the structure-dependent parameter space becomes

$$P(\theta \mid S_3, \omega) = \begin{cases} 1, & \theta \in \Theta_3 \\ 0, & \text{else} \end{cases} \quad (4)$$

A Bayesian model m represents the joint probability distribution of a set of variables $X = X_1, X_2, X_3, \ldots, X_D$. The model is defined by a graph structure, which determines the structures of the conditional probabilities between variables, and a parameter vector θ, the components of which define the entries of the corresponding conditional probability tables (CPTs). Hence, a Bayesian network can be written as m = {s, θ}. Given some observations or evidence E, reflected by fixed measured values of a subset of variables, the conditional probability given the evidence in light of the model is described as P(X|E, m).

The full Bayesian network model does not attempt to approximate the true underlying distribution. Instead, all available information is used in an optimal way to perform inference, without taking one single model for granted. To formalize this statement for our purposes, let us classify the set of available information into an available set of data D and a body of nonnumeric expert knowledge ω. The probability distribution of model m is given by

$$P(m|D, \omega) = \frac{P(D|m)P(m|\omega)}{P(D, \omega)} \quad (5)$$

The first parameter value D of P(D, ω) is the likelihood of the data given the model, which is not directly affected by nonnumeric expert knowledge ω, the second parameter value ω denotes the model a priori, whose task is to reflect the background knowledge. For simplicity, the numerator P(D, ω) of P(m|D, ω) will be omitted from the equation (5). The term P(D|m) contains the constraints of the model space by the data, and the term P(m|ω) contains the constraints imposed by the expert knowledge. Hence, given some observation or evidence E, the conditional distribution of the remaining variable X is performed by integrating over the models.

$$P(X|E, D, \omega) = \int P(X|E, m)P(m|D, \omega)dm \quad (6)$$

$$= \int P(X|E, m)P(D|m)P(m|\omega)dm \quad (7)$$

In this paper, we consider the case of no available quantitative data; D is assigned a null value. The term D and P(D|m) will be omitted from equation (6) and (7). Even in this case, it is still possible to perform a proper Bayesian inference.

$$P(X|E, \omega) = \int P(X|E, m)\, P(m|\omega)dm \quad (8)$$

Now, the inference is based on the general information (contained in ω) obtained from experts, and the specific information provided by the measurement E. In order to determine P(m|ω), we need a formalism to translate the qualitative expert knowledge into an a priori distribution over Bayesian models. The following notations are adopted for a Bayesian model class. A Bayesian model is determined by a graph structure s and by the parameter vector θ needed to specify the conditional probability distributions given that structure. The parameter vector θ is referred to by one specific CPT configuration. A Bayesian model class is then given by 1) a discrete set of model structures $S = \{s_1, s_2, s_3, \ldots, s_K\}$ and for each structure s_k, a set of CPT configurations Θ_k. The set of member Bayesian models m Є M of that class is then given by $m = \{(s_k, \theta)|k \in \{1, \ldots, K\}, \theta \in \Theta_k\}$. The model distribution is shown in (9).

$$P(m|\omega) = P(s_k, \theta|\omega) = \frac{P(\theta|s_k, \omega)P(s_k|\omega)}{\sum_{a=1}^{K} \int_{\Theta_a} P(\theta|s_a, \omega)d\theta\, P(s_a|\omega)} \quad (9)$$

In (9), the set of allowed structures is determined by means of ω, followed by the distributions of the corresponding CPT configurations. The model's a posterior probability P(m|ω) is calculated as shown in (9). Inference is carried out by integrating over the structure space and the structure-dependent parameter space.

$$P(X|E, \omega) = \sum_{k=1}^{K} \int_{\Theta_3} P(X|E, s_k, \theta)P(s_k, \theta|\omega)d\theta \quad (10)$$

It is common to express nonnumeric expert knowledge in terms of qualitative statements about a relationship between entities. The ω is represented as a list of such qualitative statements. The following information is essential to determine the model a priori (10): First, each entity which is referenced in at least one statement throughout the listed is assigned to one variable X_i. Second, each relationship between a pair of variables constrains the likelihood of an edge between these variables being presented. Last, the quality of the statement such as activates or inactivates affects the distribution over CPT entries θ given the structure. The statement can be used to shape the joint distribution over the class of all possible Bayesian models over the set of variables obtained from ω in the general case.

We propose a simplified method for constructing the a priori model distribution. Each statement is used to constrain the model space to that subspace which is consistent with that statement. In other words, if a statement describes a relationship between two variables, only structures s_k which contain the corresponding edge are assigned a nonzero probability $P(s_k|\omega)$. Likewise, only parameter values on that structure, which are consistent with the contents of that statement, are assigned a nonzero probability $P(\theta|s_k, \omega)$. If no further information is available, the distribution remains constant in the space of consistent models.

Having derived the Bayesian model class (s_3, Θ_3) consistent with the statement, we can now perform inference by using an equation (10). Under the condition of A is set to active ($E = \{A = 1\}$), let us ask what is the probability of having B active. We can determine this by integrating over all models with nonzero probability and averaging their respective inferences, which can be done analytically in this simple case

$$P(B = 1|E, \omega) = \sum_{k=1}^{K} P(s_k, \omega) \int_{\Theta_3} P(B|A, s_k, \theta)P(\theta| s_k, \omega)d\theta$$

$$= \omega \int_{\Theta_3} P(B=1|A=1, \theta)d\theta \qquad (11)$$

$$= \omega \int_0^1 \int_{\theta_0}^1 \theta_1 d\theta_1 \, d\theta_0 = 2/3$$

where $\omega = 2$ is the normalizing factor in the parameter space of $\theta = (\theta_0 , \theta_1)$ such that

$$\omega \int_0^1 \int_{\theta_0}^1 d\theta_1 d\theta_0 = 1 \qquad (12)$$

It is worth noting that, as long as simple inequalities are considered as statements, the problem remains analytically tractable even in higher dimensions. In general, integration during Bayesian inference can become intractable using analytical methods.

V. Probabilistic Representation of a Qualitative Expert Knowledge Model

The model from the previous section is derived to provide a full formalism of how to translate a set of qualitative statements into probability inequality constraints. Several qualitative models have been proposed in the context of qualitative probabilistic networks. Qualitative knowledge models describe the process of transforming qualitative statements into a set of probability constraints. The proposed Bayesian inference method outlined above is independent of the qualitative knowledge model. The model's a posterior probability is independent of the set of qualitative statements used, once the set of probabilistic inequality constraints which are translated from qualitative statements is determined. Three existing qualitative models are the Wellman approach [10], the Neufeld approach [11], and the orders of magnitude approach [12]. In this paper, we utilize the Wellman approach, where qualitative expert knowledge involves influential effects from parent variables to child variables which are classified according to the number of inputs from parent to child and their interaction. For reasons of simplicity, binary-valued variables are used in our examples. The values of a variable or node defined as "present" and "absent" or "active" and "inactive" are represented as logical values "1" and "0" (as synonyms A and $\overline{A}$). For multinomial variables, similar definitions can be applied.

Qualitative influences with directions can be defined based on the number of influences imposed from parent to child. There are three cases of influences, namely, single influence, joint influence, and mixed joint influence. In addition, there are recurrent statements and conflicting statements. The first issue can be solved by using a Dynamic Bayesian Network (DBN) [13], [14] and the second issue can be solved by adopting a voting scheme. The definitions of influence presented in this paper are refined based on the QPN in [10]. They are used to translate the qualitative expert statements into a set of constraints in the parameter space which can be used to model the parameter distribution given the structure. For a more general understanding of the explanation in this section, we assume that we obtained a set of final group decision making statements, transformed them into a set of qualitative statements, and explained those using different case studies in each criterion of probability inequality constraints for inference in a Bayesian Network. The BN model of each case study in each criterion is shown in Fig. 2.

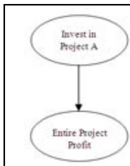

Fig. 2A. Example of single positive and negative influence.

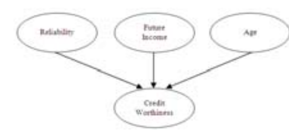

Fig. 2B. Example of plain synergy influence. Reliability, future income, and age synergically influence credit worthiness.

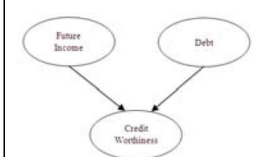

Fig. 2C. Example of mixed joint influence. Debt and future income influence on credit worthiness.

Fig. 2. The BN of each case study in each criterion

A. Single Influence

In the statement, "*investing in project A increases the profit of the entire project in such good economic situations*," *investing in project A* is the parent node which has a single positive influence on child node *the profit of the entire project*.

$$P(\text{Entire Project Profit}|\text{Invest A}) \geq P(\text{Entire Project Profit}|\overline{\text{Invest A}})$$

In another statement, "*investing in project A reduces the profit of the entire project in such a severe economic crisis*," *investing in project* A is the parent node which imposes a single negative influence on child node *the profit of the entire project*.

$$P(\text{Entire Project Profit}|\text{Invest A}) \leq P(\text{Entire Project Profit}|\overline{\text{Invest A}})$$

The graphical representation of the above qualitative statements from an expert is shown in Fig. 2A.

B. Joint Influence

Let's consider credit worthiness of individual causes. Several risk factors have been identified for credit worthiness. According to the credit bureau report, the three most prominent risk factors are reliability, future income, and age. The chance of getting credit worthiness increases as an individual gets higher future income, age, and reliability. This knowledge about credit worthiness factors can be encoded by a qualitative causality model.

According to the statements, the main risk factors that influence credit worthiness by positive synergy as shown in Fig. 2B.

The joint influence of these three factors together is more significant than individual influences from any of these factors alone. We can represent this synergy by the inequalities

$$P(CW|R, FI, A) \geq \left\{ \begin{array}{l} P(CW|R, \overline{FI}, \overline{A}) \\ P(CW|\overline{R}, FI, \overline{A}) \\ P(CW|\overline{R}, \overline{FI}, A) \end{array} \right\} P(CW|R, FI, A) \geq \left\{ \begin{array}{l} P(CW|R, FI, \overline{A}) \\ P(CW|R, \overline{FI}, A) \\ P(CW|\overline{R}, FI, A) \end{array} \right\}$$

and $\quad P(CW|R, FI, A) \geq P(CW|\overline{R}, \overline{FI}, \overline{A})$

If we assume these risk factors pair wise symmetric, we can further derive the following inequalities:

$$\begin{array}{l} P(CW|R, FI, \overline{A}) \\ P(CW|R, \overline{FI}, A) \geq \\ P(CW|\overline{R}, FI, A) \end{array} \left\{ \begin{array}{l} P(CW|R, \overline{FI}, \overline{A}) \\ P(CW|\overline{R}, FI, \overline{A}) \\ P(CW|\overline{R}, \overline{FI}, A) \end{array} \right\}$$

where CW, R, FI, and A stands for Credit Worthiness, Reliability, Future Income, and Age. Note that often but not always, the combined influence refers to the sum of independent influences from each parent node to each child node. Assume that parent nodes R and FI impose negative individual influence on child node CW, then the knowledge model can be defined as

$$P(\overline{CW}|R, FI) \geq \left\{ \begin{array}{l} P(\overline{CW}|R, \overline{FI}) \\ \\ P(\overline{CW}|\overline{R}, FI) \end{array} \right\} \geq P(\overline{CW}|\overline{R}, \overline{FI})$$

C. Mixed Joint Influence

Generally, the extraction of a probability model is not well defined if the joint affect on a child is formed by a mixture of positive and negative individual influences from its parents. Therefore, we adopted the following scheme: If there are mixed influences from several parent nodes on a child node, and no additional information is given, then these are treated as independent and with equal influential strength.

For example, future income imposes a positive single influence on credit worthiness and debt imposes a negative single influence on credit worthiness, then the joint influence can be represented by

$$\begin{array}{l} P(CW|FI, D) \geq P(CW|\overline{FI}, D), \\ P(CW|FI, \overline{D}) \geq P(CW|\overline{FI}, \overline{D}), \\ P(CW|FI, \overline{D}) \geq P(CW|FI, D), \\ P(CW|\overline{FI}, \overline{D}) \geq P(CW|\overline{FI}, D). \end{array}$$

A credit worthiness case study for a mixed joint influence is shown in Fig. 2C.

Once formulated, we can use a Monte Carlo sampling procedure to make sure that all inequalities are satisfied for valid models. Any additional structure can be brought into the conditional probability table (CPT) of the corresponding structure as soon as the dependencies between influences are made explicit by further qualitative statements.

VI. Conclusion and Future Work

In this paper, we presented several techniques in the decision process to produce a group preference ranking and a final group solution. After that we established mathematical equations for Bayesian inference based on a final group solution obtained from experts. We also described in detail a method to transform knowledge, represented by a set of qualitative statements, into an "a priori" distribution for Bayesian probabilistic models. The set of model parameters consistent with the statements and the distribution of models in the structure-dependent parameter space were presented. A simplified method for constructing the "a priori" model distribution was proposed. Each statement was used to constrain the model space to a subspace which is consistent with the statements. Next, we provided a full formalism of how to translate a set of qualitative statements into probability inequality constraints. Several cases of Bayesian influence were classified and the probability inequality constraints presented in each case are described.

For future research, we intend to construct multiple objective decision-making methods and its applications based on the concepts proposed in this paper. We will apply the concepts to a specific case study using a set of group decision making statements and report the simulation results.

References

[1] J. Pearl, "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference," San Mateo, CA, Morgan Kaufmann Publishers, pp. 552, 1988.

[2] T. Bayes, "An Essay Towards Solving a Problem in the Doctrine of Chances," Philosophical Trans. Royal Soc. Of London, 1763.

[3] S. L. Lauritzen and D. J. Spiegelhalter, "Local computations with probabilities on graphical structures and their application to expert systems, " J. Royal Statistical Soc., 1988.

[4] D. Heckerman, "A Tutorial on Learning with Bayesian Networks," Technical Report MSR-TR-95-06, Microsoft, http://research.microsoft.com/research/pubs/view.aspx?msr_tr_id=MSR-TR-95-06, 1996.

[5] D. Heckerman, "Learning Bayesian Networks: The Combination of Knowledge and Statistical Data," Proc. KDD Workshop, 1994.

[6] N. Friedman and M. Goldszmidt, "Learning Bayesian Networks with Local Structure, " Learning in Graphical Models, 1999.

[7] J. Cheng, R. Greiner, J. Kelly, D. Bell, and W. Liu, "Learning Bayesian Networks from Data: An Information-Theory Based Approach," Department of Computing Sciences, University of Alberta, Faculty of Informatics, University of Ulster, 2001.

[8] M. Singh and M. Valtorta, "Construction of Bayesian, Network Structures from Data: a Brief Survey and an Efficient Algorithm," Dept. of Computer Science, University of South Carolina, Columbia, USA, 1995.

[9] P. Spirtes, and C. Meek, "Learning Bayesian Networks with Discrete Varibales from Data," Proceedings of the Conference on Knowledge Discovery & Data Mining, 1995.

[10] M. P. Wellman, "Fundamental Concepts of Qualitative Probabilistic Networks," Artificial Intelligence, 1990.

[11] E. Neufeld, "A Probabilistic Commonsense Reasoner," Int'l J. Intelligent Systems, 1990.

[12] J. Cerquides and R. Lopez de Mantaras, "Knowledgge Discovery with Qualitative Influences and Synergies, " Proc. Second European Symp. Principle of Data Mining and Knowledge Discovery (PKDD), 1998.

[13] K. Murphy, "Dynamic Bayesian Networks: Representation, Inference, and Learning," PhD dissertation, University of California, Berkeley, 2002.

[14] J. Nipat and P. Wichian, "A SMILE Web-based Interface for Learning the Causal Structure and Performing a Diagnosis of a Bayesian Network," Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics, San Antonio, TX, USA , 2009

[15] T. E. María and M. Jose María, "Aggregation of Individual Preference Structures in Ahp-Group Decision Making," Journal of Group Decision and Negotiation, Springer Netherlands, Vol. 16, No. 4, July, 2007.