Conflict Resolution in Probabilistic Multi-agent Systems

S. K. M. Wong and Tao Lin
Department of Computer Science
University of Regina
Regina, Saskatchewan
Canada S4S 0A2
{wong, taolin}@cs.uregina.ca

Abstract

Knowledge representation in the Multi-agent systems (MAS) can be characterized by Probabilistic Network. Coordination among probabilistic agents can be achieved on account of their common knowledge, while agent can obtain his own knowledge locally and variously. To construct the common knowledge for the MAS, the conflict among local knowledge of all agents has to be identified and resolved. We propose the necessary and sufficient conditions for their knowledge structure. We also address that the redundancy has to be eliminated at the beginning in case of misleading conclusion. An algorithm is suggested to construct the structure of common knowledge in terms of Bayesian DAG.

Introduction

The Multi-agent system (MAS) is a system consisting of multiple autonomous computational components, in which each agent deploys his own asynchronous computation based on local data to solve the domain problems coordinately (Jennings et al., 1998). There implemented various MAS applications which attract the public interest (Jennings et al., 1998; Weiss, 1998).

It was noticed that the probabilistic semantics can be utilized as a *communication protocol* in the MAS (Wong et al., 2001; Xiang, 1996). In this paper, we only consider a system of cooperative agents collaborating with each other to draw inference from their knowledge in a shared environment. We interprate the semantics of knowledge in terms of *conditional independencies* (CIs). Each agent has his own knowledge explicitly represented in terms of *probabilistic network* such as *Bayesian Network* (BN) or *Markov Network* (MN) (Pearl, 1988), which can be learned from the sample data coming from the local environment, or directly specified by the domain expert. The global coherence in such MAS is maintained through the belief updating on *common knowledge* (Fagin et al., 1995) among agents. The semantics of common knowledge received intensively study as a communication protocol in the MAS (Fagin et al., 1995). There are several advantages to introduce common knowledge as an organization of local knowledge for all agents. Firstly, it enables agents interprate and obtain their own local knowledge independently by various methods and techniques. Secondly, agents are able to share the knowledge of other agents for global reasoning and acting. Thirdly, uncertainty can be handled consistently and propagate correspondingly through the whole system.

To construct the common knowledge in the MAS, as soon as agents pool their knowledge together to yield to a set of CIs, we have to settle down the structure of common knowledge represented by a Bayesian *directed acyclic graph* (DAG), and then update the corresponding *conditional probability tables* for each agent by applying conventional propagation technique (Jensen, 1996). Intuitively, as agents may have different local knowledge overlapped, some of them could be *conflicted* with others. It means that not all CIs, which are coming from probabilistic networks of all agents, can be *faithfully* represented by one single Bayesian Network for the whole system. The most difficult task in the construction procedure for common knowledge is to identify conflict CIs for a given set of CIs (Verma et al., 1992). It is necessary to characterize a set of CIs which can be faithfully represented by a Bayesian DAG. By characterizing the given set of CIs, various options or disciplines can be provided and followed to determine which CIs should stay or not, so as to construct the common knowledge for the MAS. In this paper, We define the hierarchical structure for a set of CIs. After redundant CIs are removed firstly, we point out that a set of CIs must posses a particular hierarchical structure in order to construct a Bayesian DAG. They are necessary and sufficient conditions. An algorithm is consequently suggested to construct a Bayesian DAG for a set of CIs if they satisfy those conditions.

We include a brief review of basic concepts in section 2. The definition of hierarchical CI is presented in section 3. The method for removing redundant CIs from the input CIs set is described in section 4. In section 5, we describe the necessary and sufficient conditions to construct a Bayesian DAG for a given set of CIs and propose an algorithm for the construction if those conditions are satisfied. The conclusion is found in section 6.

Definitions

Let the *universal context* $U = \{A_1, A_2, \ldots, A_m\}$ denote a finite set of variables. We say an agent *knows* Y is *condition-*

ally independent of Z under the circumstance X, denoted by

$$I(Y, X, Z), \tag{1}$$

if and only if

$$P(Y|XZ) = P(Y|X), \tag{2}$$

where X, Y and Z are disjoint subsets of U, and XY stands for the union of X and Y. The Equation (1) is called a *conditional independence statement*, where the left hand side X is the *key* of this CI. The union of all of variables involved in a CI, XYZ namely, is defined as the *context* of this CI.

Given a set G of CIs, we say a CI is *derivable* from G if it can be inducted from G by applying *semigraphoid axioms* (Pearl, 1988). We say G is *equivalent* to another set G' of CIs if and only if any CI derivable from G is also derivable from G' and verse versa. Any set of CIs is called a *cover* of G if they are equivalent. If any CI in G^+ can be identified from a graphic structure $\mathbf{G}$ and any CI identified from $\mathbf{G}$ is also in G^+, then we say G can be *faithfully represented* (Pearl, 1988) by the graphic structure $\mathbf{G}$.

Let Y be a subset of variables over the universal context U. We say a CI $I(V_1, X, V_2)$ *split* Y if and only if $V_1 \cap Y \neq \phi$ and $V_2 \cap Y \neq \phi$. Y is *split by a set* G of CIs if there exists a CI in G that splits Y. For a CI $I(Y, X, Z)$ with the context $U_i = XYZ$, if there exists another non-trivial CI $I'(YW, X, ZV)$ with the context $U_j = XYZWV$, $W \neq \phi$ or $V \neq \phi$, which is derivable from G by applying semigraphoid axioms, then we say I' is the *extension* of I, and I is the *projection* of I' under the context U_i. We say I' is *the maximal extension* of I if there is no extension of I'.

A Bayesian DAG can be *moralized* and *triangulated* (Jensen, 1996) into a *chordal conformal* undirected graph $\mathbf{M}$. This graph can be equivalently represented by an *acyclic hypergraph*, namely,

$$\mathbf{H} = \{h_1, h_2, \ldots, h_n\},$$

where the hyperedges $h_i \in \mathbf{H}$, $1 \leq i \leq n$, are exactly the maximal cliques of $\mathbf{M}$. The union of all variables involved in $\mathbf{H}$ is called the *context* of hypergraph. We call the ordering of hyperedges $h_1, h_2, \ldots, h_n$ *construction ordering* of $\mathbf{H}$, which satisfies the *running intersection property* (Beeri et al., 1983) such that

$$s_i = h_i \cap (h_1 \cdots h_{i-1}) \subseteq h_j,\ 1 \leq j \leq i-1.$$

We say s_i, $1 \leq i \leq n-1$, is the *separator* of $\mathbf{H}$ with respect to the hyperedge h_i.

By graphical separation method (Castillo et al., 1997), for each separator s_i of the acyclic hypergraph $\mathbf{H}$ with the context U, there exists a corresponding CI, denoted by $I(s_i)$, such that

$$I(\mathcal{H}_i, s_i, U - \mathcal{H}_i - s_i), \tag{3}$$

where $\mathcal{H}_i$ is the *connected component* including remaining part of h_i in hypergraph $\mathbf{H} - \{s_i\}$. The acyclic hypergraph $\mathbf{H}$ can be used to faithfully represent a set G_U^H of full CIs under the context U of hypergraph $\mathbf{H}$ such that:

$$G_U^H = \{I(s_1), I(s_2), \ldots, I(s_n)\},$$

where $s_1, s_2, \ldots, s_n$ are separators of $\mathbf{H}$. It should be noticed (Jensen, 1996) that some of CIs represented by a Bayesian DAG may get lost after this DAG being transformed into a chordal conformal undirected graph.

Hierarchical CIs

In this section, we introduce a concept called *hierarchical CIs* (HCIs) and will show how to construct such a Bayesian DAG from the given set of CIs, by taking advantage of this concept in the following sections.

Given a set G of CIs over the universal context U, we may partition these CIs respectively according to their contexts, $U_1, U_2, \ldots, U_n$ namely. It follows that :

$$G = G_{U_1} \cup G_{U_2} \cup \cdots G_{U_n},$$

where G_{U_i}, $1 \leq i \leq n$, is a subset of G in which each CI has the same context, $U_i \subseteq U$ namely. After partitioning an arbitrary set G of CIs, we say G is a set of HCIs if

(1) All CIs in each group G_{U_i}, $U_i \subseteq U$ namely, can be perfectly represented by an acyclic hypergraph denoted by $\mathbf{H}_{U_i}$;

(2) For each hypergraph $\mathbf{H}_{U_i}$, $U_i \subset U$ namely, there is a distinct hyperedge h in another hypergraph $\mathbf{H}_{U_j}$ such that $U_i \subseteq h$.

We say $\mathbf{H}_{U_i}$ is the *refined hypergraph* of hyperedge h.

By the definition of HCIs, there exists a collection of acyclic hypergraphs corresponding to a set of HCIs. We may construct a new hypergraph called a *combined hypergraph*, written as:

$$\mathbf{H}(G) = \{h_1, h_2, \ldots, h_m\},$$

where h_j, $1 \leq j \leq m$, is a hyperedge of $\mathbf{H}_{U_i}$, $U_i \subseteq U$, $1 \leq i \leq n$, such that h_j has no refined hypergraph. The union $\mathcal{H}$ of any connected hyperedges is called a *hyperedge cover* of the combined hypergraph $\mathbf{H}(G)$. We may replace those hyperedges h_{i_1}, h_{i_2}, …, h_{i_k}, by hyperedge cover $\mathcal{H}$, to obtain a new hypergraph, such that,

$$\mathbf{H}(G) \cup \mathcal{H} - \{h_{i_1}, h_{i_2}, \ldots, h_{i_k}\}.$$

We refer to this new hypergraph as the *hypergraph cover* of $\mathbf{H}(G)$.

Example 1 *Let G be a set of CIs as follows:*

$$G = \left\{ \begin{array}{c} I(B, A, C) \\ I(D, B, CE) \\ I(E, C, BD) \\ I(A, BC, DEF) \\ I(F, DE, ABC) \end{array} \right\}. \tag{4}$$

After we group each CI according to their contexts respectively, it follows that

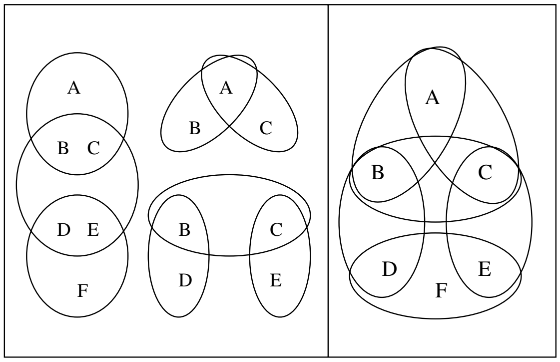
$$G = G_{ABC} \cup G_{BCDE} \cup G_{ABCDEF},$$

where

$$\begin{array}{l} G_{ABC} = \{I(B, A, C)\}, \\ G_{BCDE} = \left\{ \begin{array}{c} I(D, B, CE) \\ I(E, C, BD) \end{array} \right\}, \\ G_{ABCDEF} = \left\{ \begin{array}{c} I(F, DE, ABC) \\ I(A, BC, DEF) \end{array} \right\}. \end{array} \tag{5}$$

The acyclic hypergraphs corresponding to each group of G in equation (5) are shown in Figure 1(a). G is a set of HCIs according to the definition.

The combined hypergraph $\mathbf{H}(G)$ *of G is shown in Figure 1(b).*



(a) (b)

Figure 1: A collection of acyclic hypergraphs faithfully representing a set G of HCIs, given in the Example 1, is shown in (a). The combined hypergraph $\mathbf{H}(G)$ is shown in (b).

Reduced HCIs

When agents pool their CIs together, there might exist some redundant CIs which could mislead conclusion on the structure of common knowledge. For example, consider an agent in a MAS who has two CIs such as: $\{I(B, A, C), I(B, A, D)\}$. In his mind, neither of them are redundant. But if he comes to cooperate with another agent, who only has one CI such as $\{I(BC, A, D)\}$, it appears that $I(B, A, D)$ is redundant. If this redundant CI is kept, then the set G including all CIs, $G = \{I(B, A, C), I(BC, A, D), I(B, A, D)\}$ namely, is not hierarchical, although it is equivalent to a set G' of HCIs, $G' = \{I(B, A, C), I(BC, A, D)\}$ namely. In this section, we introduce an algorithm on how to remove the redundant CIs in a given set G of CIs. The output of this algorithm is referred to the *reduced cover* of G, denoted by G^*. If G^* satisfies our *hierarchical* conditions, then we say G^* is a set of *reduced HCIs* of G. The complexity of this algorithm is dominated by Step 2 which is discussed in (Ozsoyoglu et al., 1987).

Algorithm 1 *To find the reduced cover.*
Input: A set G of CIs.
Output: The reduced cover G^* of G.

Step 1:	*Group CI in G on account of their contexts respectively, that is,* $G = G_{U_1} \cup G_{U_2} \cup \ldots, G_{U_n}$;
Step 2: [1]	*Remove redundant CIs from each* G_{U_i}. *That is, remove g from* G_{U_i} *if* $(G_{U_i} - \{g\})^+ = G^+_{U_i}$;
Step 3:	*Replace every g in G by its maximal extension;*
Step 4:	*Go to step (1) until no more change can be made in G;*
Step 5:	*Remove g from G if* $(G - \{g\})^+ = G^+$.
Step 6:	*Output G.*

Example 2 *Consider a set of CIs in equation (4) as example 1. Its reduced set G^* is given as follows:*

$$G^* = \left\{ \begin{array}{c} I(BD, A, CE) \\ I(D, B, ACE) \\ I(E, C, ABD) \\ I(F, DE, ABC) \end{array} \right\}.$$

G^* can be represented by a collection of acyclic hypergraphs shown in Figure 2. It is easy to verify that G^* is a set of HCIs.

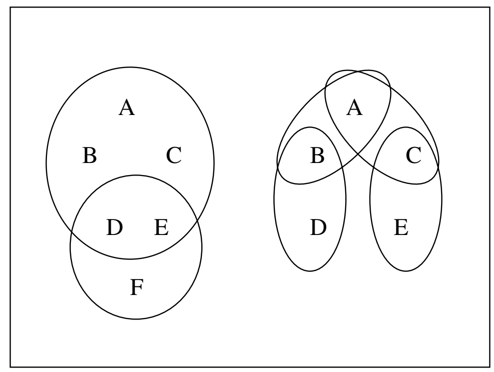


Figure 2: A collection of acyclic hypergraphs corresponding to the reduced HCIs of G, given in Example 2.

After its reduced cover G^* is obtained by removing redundant CIs from a given set G of CIs, we have to find out which CIs are *conflicting* with others to construct a Bayesian DAG and therefore should be eliminated. We may directly take the criterion as to preserve given CIs as much as possible, but most likely, this criterion should leave to the practical applications.

The Bayesian DAG Construction

Consider the combined hypergraph with respect to a set G of HCIs. There exist a split-key pattern called *opposing triangle pattern* (OTP). In this section, after the OTP is formally defined, we show its existence determines sufficiently and necessarily whether or not a set of reduced HCIs can be faithfully represented by a Bayesian DAG. An algorithm is also suggested on how to construct a Bayesian DAG based on this claim. An example is provided to show the construction procedure.

Definition 1 *For the combined hypergraph with respect to a set G of reduced HCIs, if it has a hypergraph cover in which more than one split keys are contained by a single hyperedge, then we say G has the OTP.*

[1]Removing redundant multi-valued dependencies with fixed context was discussed vigorously in (Ozsoyoglu et al., 1987). Its idea can be transferred to the probabilistic implication problem by our discussion in (Wong et al., 2000).

Example 3 *Given a set G of reduced HCIs as follows:*

$$G = \left\{ \begin{array}{c} I(BC, A, EFD) \\ I(D, EF, ABC) \\ I(B, A, C) \\ I(E, D, F) \end{array} \right\}. \quad (6)$$

Its combined hypergraph is shown as Figure 3. The hyperedge $\{BCEF\}$ already contains two split keys, BC and EF namely. It follows that G has the OTP.

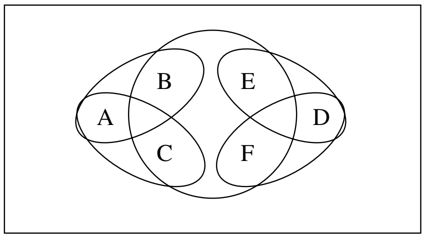


Figure 3: Opposing triangle pattern in the combined hypergraph $\mathbf{H}(G)$, where $I(B, A, C)$ and $I(E, D, F)$ are two CIs in G.

Theorem 1 [2] *Given a set G of reduced HCIs, it can be faithfully represented by a Bayesian DAG if and only if it has no OTP.*

After redundant CIs are eliminated, Theorem 1 implies that some of CIs have to be removed if all CIs can not be faithfully represented by one single Bayesian DAG. It provides a criterion on how to resolute conflicting CIs by characterizing the structure of common knowledge. As long as those CIs coming from different agents have no OTP, they will stay together for coordination and become the basic structure of the common knowledge among agents, which is faithfully characterized by a Bayesian DAG. After a Bayesian DAG is fixed, the belief updating for each agents can be naturally followed by traditional propagation techniques. As the complement of this argument, the following algorithm suggests a method to construct a Bayesian DAG if the given set of CIs satisfies Theorem 1.

Algorithm 2 *To construct a Bayesian DAG.*
Input: A set G of reduced HCIs satisfying the conditions of Theorem 1;
Output: A Bayesian DAG.

Step 1: *For each subset of G with the same context U_i, Construct hypergraphs $\mathbf{H}_{U_i}$ respectively;*
Step 2: *Initialize a stack S to store hyperedges;*
Step 3: *Let $U_i = U$;*
Step 4: *Specify the construction ordering of $\mathbf{H}_{U_i}$ as follows: if separator s_i of hyperedge h_i is split by a CI with the context h_j, then $j = i - 1$;*
Step 5: *Push all of hyperedges of $\mathbf{H}_{U_i}$ into the stack S following the construction ordering, starting from the first one.*
Step 6: *If the stack is empty, output the DAG and stop. Otherwise, pop the first hyperedge h_k from the stack S.*
Step 7: *If the separator s_k is not split by CI in G, then draw arrows from the variables of separator s_k to the other variables in h_k, go to Step 5. Otherwise, let $U_i = h_j$, go to Step 3.*

Example 4 *Consider a set G of reduced HCIs over the universal context $U = \{ABCDEFGH\}$ as follows:*

$$G = \left\{ \begin{array}{c} I(B, A, C) \\ I(E, D, F) \\ I(A, BC, DEFGH) \\ I(ABC, D, EFGH) \\ I(ABCD, EF, GH) \\ I(ABCDEG, F, H) \end{array} \right\}. \quad (7)$$

It can be faithfully represented by a collection of corresponding acyclic hypergraphs, which are shown in Figure 4. Since G satisfies the conditions of Theorem 1, we may construct the corresponding DAG following Algorithm 2.

In the first round, the construction ordering is specified for the hypergraph $\mathbf{H}_U$, as shown in Figure 4. After hyperedges of $\mathbf{H}_U$ are pushed into the stack S, the content of S is $\{ABC, BCD, DEF, EFG, FH\}$, where left "$\{$" is the bottom of the stack and "$\}$" is the opening of the stack. According to Step 6 of Algorithm 2, after FH, EFG are popped and corresponding arrows are drawn, we only have to pop hyperedge DEF without doing anything, because its separator EF is split by $I(E, D, F)$. And then the construction ordering for hypergraph $\mathbf{H}_{DEF}$ is specified. After that, its hyperedges DE and DF are pushed into the stack S. The content of S becomes $\{ABC, BCD, DE, DF\}$. This procedure can be executed recursively until the stack becomes empty. A Bayesian DAG is yielded accordingly as shown in Figure 5.

Conclusion

The common knowledge of agents in MAS is the knowledge that *each agent agrees and knows other agents agree* (Fagin et al., 1995). Building common knowledge is a procedure to identify conflicting knowledge, resolve the disagreement among agents. Based on probabilistic semantics, the knowledge in MAS can be viewed as the Probabilistic Network. A Probabilistic Network has the qualitative structure, which can be faithfully represented by a graphical structure, or a set of CIs, and quantitative measurement, which can be configured by conditional probability tables. The common knowledge can be obtained in two steps: the first step is to fix the

[2] The detailed proof for Theorem 1 can be found in (Wong et al., 2002).

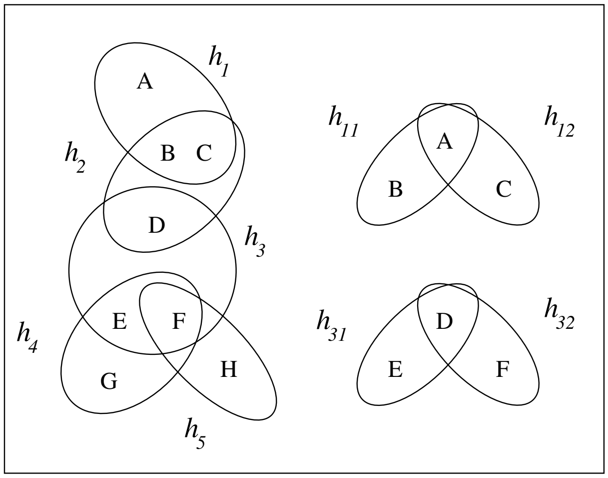


Figure 4: A collection of acyclic hypergraphs used to faithfully represent a set G of reduced HCIs listed in Equation (7). Their construction ordering are also specified in the figure.

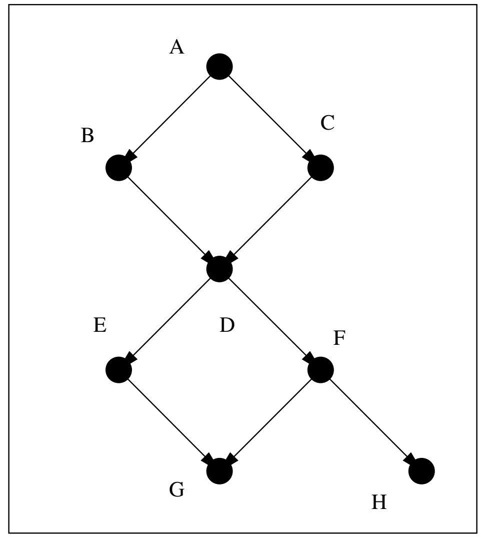


Figure 5: A DAG is shown in this figure to faithfully represent a given set G of reduced HCIs in Example 4.

structure of the common knowledge, and the second step is to configure the quantitative measurement for the common knowledge. In this paper, the first step is discussed intensively since the second step can be done by the conventional propagation technique. The question in this stage is: assume each agent has a set of CIs, after pooling these CIs together, how do we combine them to yield a Bayesian DAG over the problem domain? if not all of them can be faithfully represented by one single Bayesian DAG, how do we manipulate these CIs coming from different agents into a Bayesian DAG? At first we argue that the redundant CIs have to be removed at the very beginning in case of misleading judgment. And then we show that those non-redundant CIs have to come with a hierarchical structure. Furthermore, in the structure there should not exist a split-key pattern called "opposing triangle pattern". An algorithm is suggested to construct such a Bayesian DAG if all conditions are satisfied.

Acknowledgments

We would like to thank three anonymous referees for their helpful comments on the earlier version of this paper that improved the presentation of the paper considerably.

References

C. Beeri, R. Fagin, D. Maier and M. Yannakakis: On the Desirability of Acyclic Database Schemes, *J. ACM*, Vol. 30, No. 3, 479-513, July 1983.

Enrique Castillo, José Manuel Gutiérrez and Ali S. Hadi: Expert Systems and Probabilistic Network Models, Springer-Verlag, 1997.

R. Fagin, J.Y. Halpern, Y. Moses and M.Y. Vardi: *Reasoning About Knowledge* The MIT Press, 1995.

N. R. Jennings, K. Sycara and M. Wooldridge: A Roadmap of Agent Research and Development. *Autonomous Agents and Multi-Agent Systems*, 1:275-306, 1998.

Jensen, F.V.: An introduction to Bayesian networks. UCL Press, 1996.

Z.M. Ozsoyoglu and L.Y. Yuan: Reduced MVDs and minimal covers. *ACM Transactions on Database Systems*, Vol.12, No.3:377-394, September 1987.

J. Pearl: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers, San Francisco, 1988.

T.S. Verma and Judea Pearl: An Algorithm for Deciding if a Set of Observed Independencies Has a Causal Explanation. In *Eighth Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann Publishers, 1992.

G. Weiss, editor: *Multi-agent Systems*, The MIT Press, 1998.

S.K.M. Wong, C.J. Butz and D. Wu: On the implication problem for probabilistic conditional independency. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans* Vol. 30, No. 6, 785-805, November 2000.

S. K. M. Wong and C. J. Butz: Constructing the Dependency Structure of a Multi-agent Probabilistic Network. *IEEE Trans. on Knowledge and Data Engineering*, Vol. 13, No. 3, pp.395-415, 2001.

S. K. M. Wong, Tao Lin and Dan Wu: Bayesian DAG Construction from Conditional Independencies In *Seventh International Symposium on Artificial Intelligence and Mathematics*, 2002.

Y. Xiang: A Probabilistic Framework for Multi-agent Distribution Interpretation and Optimization of Communication. *Artificial Intelligence*, Vol. 87, No.1-2, pp.295-342, 1996.