Hierarchical Qualitative Inference Model with Substructures

Zehua Zhang[1,2], Duoqian Miao[1], and Jin Qian[1]

[1] Department of Computer Science and Technology, Tongji University, Shanghai, China 201804
zehua.zzh@gmail.com

[2] College of Computer Science and Technology, Taiyuan University of Technology, Shanxi, China 030024
{miaoduoqian,qjqjlqyf}@163.com

Abstract. Qualitative propagation influences in qualitative inferences are unlike and interrelated on the different hierarchy of knowledge granules, and quantitative information loss easily results in reasoning conflicts. This paper presents a hierarchical qualitative inference model with substructures which to some extent can eliminate the qualitative impact of uncertainty and solve trade-off problems by metastructures with basic decomposition and coarse-grained mesoscale substructures with edge-deletion. The substructural inferences could not only reduce computational complexity, but provide an approximate strategy for modular reasoning on large-scale problems. The example respectively illustrates the two substructural methods are both effective.

Keywords: qualitative inference, substructures, hierarchical structure, granular computing.

1 Introduction

Uncertainty problems with inaccurate, incomplete and incorrect information extensively exists in the reality. As a kind of tools to efficiently acquire knowledge and reason for uncertain knowledge in complex problems, Bayesian networks (BNs)[6] can visually reveal the structure of problems via graph theory, and analyze the structure according to the probability of problems.

Due to the network structure is still quite complex to build BNs according to the actual problems, Wellman[13] proposes qualitative probability networks (QPNs), which reason with qualitative signs propagation way of conditional probability. QPNs as the abstract description on general BNs can improve reasoning abilities and enhance reasoning efficiency. And some polynomial time sign propagation QPNs algorithm[2,12] extend qualitative reasoning from singly connected to general multiply connected networks. But whatever qualitative methods are taken, qualitative abstract with information lost may lead to reasoning results with conflicts. In other words, the inference with conflicts would get uncertainty results. Thus, in recent years the related work mostly

J.T. Yao et al. (Eds): RSKT 2011, LNCS 6954, pp. 753–762, 2011.

adopts the integrated strategy with qualitative and quantitative information to solve conflicts, such as the incremental qualitative and quantitative combined method[5], sign propagation based on context[9], enhanced qualitative propagation strategies[10], rough-set-based weights QPNs[16] and interval probability parameters as indicators of influence strengths[17]. Moreover, network decomposition measures in QPNs is taken[4], but the method is not involved in utilization of the prior knowledge, and weak as before in solving the large-scale complex network.

Most of all, QPN algorithms have never considered that qualitative propagation influences are not identical but interrelated upon the different hierarchy of knowledge granules. For example, the people possibly have different opinions based on their perspective for the same problem. As a matter of fact, the different cognitive description of the problem are to some extent correlative. Hence, whether the existing knowledge of different levels are benefit to each other or not, merely depends on how the knowledge is used.

The motivation of the work with the hierarchy theory of granular computing comes from making up above shortages of QPNs. Granular computing (GrC), as a popular research field in recent years[14], concerns the processing of complex information entities, which arise in the process of data abstraction and derivation of knowledge from information. The hierarchy theory of granular computing which is a multi-disciplinary and cross-disciplinary study provides a multi-layered framework based on levels. Yao[15] considers different levels involve different types of processing. The lowest level concerns numeric processing, the intermediate level concerns larger information granules, and the highest level concerns symbol-based processing. Feng and Miao[3] draw the hierarchy theory into the decision-making domains, and puts forward a hierarchical decision rule mining model based on multidimensional data and a hierarchical rough set model based on the concept of attributes. Pedrycz[7] elaborates on the design of information granules for machine learning techniques.

Consequently, we propose a hierarchical qualitative inference model with substructures including inference on metastructures with basic decomposition and coarse-grained mesoscale structures with edge-deletion. Qualitative inference on substructures in different levels could avoid conflicts, but benefit to find the suitable inference model and the optimal structure of reasoning aiming at the problems. Because global reasoning partly depends on local reasoning, the distributed inference method with substructures could accelerate global reasoning. In addition, substructural inference methods provide an approximate strategy of modular reasoning on large-scale problems.

The paper is arranged as follows. Section 2 briefly introduces preliminaries and their qualitative abstractions of QPNs. In Section 3 the work elaborates the hierarchical qualitative inference model with substructures and its two substructural reasoning strategies. Section 4 demonstrates the both methods are feasible by solving an inference example. The paper ends with conclusions in Section 5.

2 Preliminaries

A Bayesian network encodes probabilistic knowledge about a problem domain through a dependence structure in the form of a directed acyclic graph (DAG) and Conditional Probability Tables (CPTs) associated with nodes of the graph. Qualitative probabilistic networks[13] are introduced as qualitative abstractions of Bayesian networks, and thus bear a strong resemblance to their quantitative counterparts. A qualitative probabilistic network also is described to a DAG model by variables and the probabilistic relationships between them, denoted $G = (V(G), A(G))$. However, a qualitative probabilistic network associates with its digraph *qualitative influences* and *qualitative synergies* instead of conditional probability distributions.

Definition 1. ***qualitative influences[13]:*** *Let $A \to B \in A(G)$ be an acyclic digraph, and $A, B \in V(G)$. A positively influences B, denoted $S^+(A, B)$. If* $\Pr(B|Ax) - \Pr(B|Ax) \geq 0$ *for any combination of values x for the set $\pi(B)\backslash\{A\}$ of predecessors of B other than A.*

This definition expresses observing a higher value of A makes higher value of B more likely, regardless of any other direct influences on B. Similarly, a negative qualitative influence and a zero qualitative influence, respectively denoted S^- with $\leq$ and S^0 with $=$ in the above function. A non-monotonic or unknown influence of A on B is denoted by $S^?$, called ambiguous influence. The set of all influences of a qualitative network exhibits various important properties. The property of symmetry states that, if the network includes the influence $S^\delta(A, B)$ and $S^\delta(B, A)$, $\delta \in \{+, -, 0, ?\}$.

Table 1. The $\oplus$-and$\otimes$-operators

$\oplus$	+	-	0	?
+	+	?	+	?
-	?	-	-	?
0	+	-	0	?
?	?	?	?	?

$\otimes$	+	-	0	?
+	+	-	0	?
-	-	+	0	?
0	0	0	0	0
?	?	?	0	?

Besides, a qualitative probabilistic network includes *synergies* that express how the values of one node influences the probabilities of the values of another node in view of a value for a third node[13].

Definition 2. ***qualitative synergies:*** *Let G be as before, with $A, B, C \in V(G)$ and $\pi(C) = \{A, B\}$. A negative product synergy of node A on node B (and vice versa), given the value c for node A , denoted $X^-(\{A, B\}, c)$, expresses that, given c, a higher value for C renders the higher value for B less likely, that is,*

$$\Pr(c|a\overline{b}) \cdot \Pr(c|\overline{a}b) - \Pr(c|ab) \cdot \Pr(c|\overline{a}\overline{b}) \geq 0 \qquad (1)$$

Positive, zero, and ambiguous product synergies are defined analogously. The product synergy $X^{\delta}(\{A, B\}, c)$ serves, upon observing c, a qualitative *intercausal influence* with sign δ is induced between A and B.

Qualitative probabilistic inference propagates the sign between related nodes, bases on the idea of tracing the influence of observed node on others nodes. From table 1, we could draw some conclusions that the combined influences only result an ambiguous sign on V type structure by $\oplus$-operator. Once the conflict happens, it would spread the ambiguous result to the whole network. So it is vital to research on avoiding conflicts in qualitative inference.

3 Hierarchical Qualitative Inference Model

Granular computing theory can enhance the process of data abstraction and derivation of knowledge from information, while QPNs problems are required to look for the balance between the conflict-free and efficiency. The multi-view and multi-level structure models reflect the internal knowledge structure of problems, and make the structured solution possible. Therefore, we propose a hierarchical qualitative inference model with substructures as Fig.1. The model could respectively reason on coarse-grained mesoscale structures with edge-deletion and metastructures with basic decomposition.

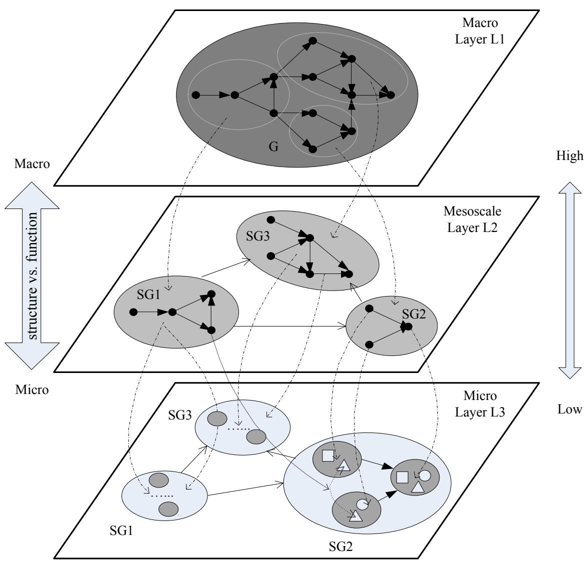

Fig. 1. Hierarchical qualitative inference model with substructures

To begin with, the hierarchy theory from high to low level demonstrates the cognitive style of human beings from the macro scale to the micro scale, and the cognitive process from coarse granules to fine granules. And then the analysis from the top to the bottom also reflects the relevance between knowledge structures and functions, which provides the theoretic support for solving structured and modular problems. Furthermore, we are supposed to observe that the computational cost is generally much higher from coarse granules to fine granules, the accuracy yet is not necessarily improved. It means the fine granular method is no more effective than the coarse-grained for actual problems. Our model could find the suitable method by hierarchical analysis.

In the next, the qualitative inference methods on metastructures with basic decomposition and coarse-grained mesoscale structures with edge-deletion would be introduced.

3.1 Mesoscale Substructures Inference with Edge-Deletion

Mesoscale is an intermediate scale described as the size between macro and micro structures, which is contained by macro structure and could include some micro substructures. The qualitative inference methods are based on a Bayesian network with known structure, while mesoscale substructures inference with edge-deletion attempts to reconstruct local network with conflicts to solve trade-off problems.

Edge deletion strategy: Let $U \rightarrow X$ be an edge in a Bayesian network $G = (V(G), A(G))$ and $U \rightarrow X \in A(G)$, and suppose that we would like to delete this edge. But the deletion operation will cause two problems. variable X will lose its direct dependence on parent U. On the other hand, variable U may lose evidential information received through its child X. To address these problems, Choi[1] shows recovering the deleted edges with high mutual information can improve the quality of approximations computed by belief propagation without incurring much additional computational cost. On the basis of above research, Renooij[11] considers the possible impact of removing a single, pre-selected arc in a different setting on the behavior of the network, both with and without evidence.

The above work on edge deletion with adding new parameters pay attention to the effect on the network, but not to approximating networks to make inference feasible. In the paper we focus on simplifying networks that observable variables are efficiently approximated to the certain sets. Especially on a sparser structure, domain experts can easily find the relationship between variables. Therefore, we propose a modified edge deletion method defined in the following:

Definition 3. ***Modified Edge Deletion(MED):*** *Let $U \rightarrow X \in A(G)$, $G = (V(G), A(G))$. We say that the edge $U \rightarrow X$ is deleted when it results in a conflict network that is obtained from G as follows:*

- $G' = G/(U \rightarrow X)$.
- $\exists X, Y \in V(G)$, *st.*$Y \rightarrow X$, $U \rightarrow Y$, $Y \neq U$, *and* $U' = U$.
- *Case1: If* $Pr(X|U') \geq Pr(X|Y)$,
  then $G' = G/(Y \rightarrow X)$ *and* $Pr(X|U'') = |Pr(X|U) - Pr(X|Y)|$.
- *Case2: Else delete* $U' \rightarrow X$, *and* $Pr(X|Y') = |Pr(X|U) - Pr(X|Y)|$.

– *$U \rightarrow E'_X$, E'_X replaces X as a child of U.*
 Due to E'_X cannot directly impact on inference result, E'_X is instantiated as soft evidence on U.

Fig.2 illustrates the modified edge deletion operation. Firstly, the conflict happens on X in Fig.2(a), and then Fig.2(b) shows the method as[1]. The edge $U \rightarrow X$ is removed from the graph G. Y is one of other observed fathers of X. If Y is connected with U, then the smaller edge is alternative to be removed by influence on X. And a new variable U' replaces variable U as a parent of X, where U' is a copy of U, with the same state as U. E'_X as soft evidence on U could replace X as a child of U, where E'_X is instantiated and can help to choose the deleted edge. Fig.2(c) and (d) respectively show the result in the case1 and the case2 on merging the edges of X by the influence on X.

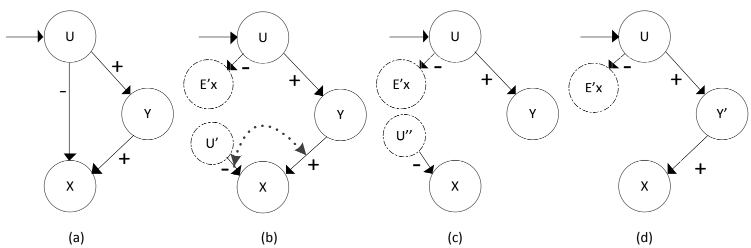

Fig. 2. Modified edge deletion operation

Only when the influence of U' on X is not less than the influence of Y on X, the modified edge deletion method could separate the whole network into some subnetworks without conflicts as Fig.2(c).

Algorithm 1: Sign-Propagation algorithm with modified edge deletion
Input: a Bayesian network with known structure.
Output: propagation sign.
Begin procedure PropagateSign(trail, to, messagesign):

– sign[to]$\leftarrow$ sign[to] $\oplus$ messagesign;
– trail$\leftarrow$ trail $\cup$ to;
– **for** each active neighbour V_i of to,
– **Do** linksign $\leftarrow$ sign of influence between to and V_i ;
– messagesign $\leftarrow$ sign[to] $\otimes$ linksign;
– **If** $V_i \notin$trail and sign[V_i]$\neq$ sign[V_i]$\oplus$ messagesign,
– **then** Delete the edge and update the influence by Definition 3 MED.
– PropagateSign (trail, to, V_i, messagesign).
– **endif**
– **endfor**

The idea of the algorithm is to follow the sign influence with breadth-first search. When the inference conflicts happen, the modified edge deletion method is taken until the whole network is decomposed into subnetworks without conflict.

3.2 Metastructures Inference with Basic Decomposition

The metastructure is one kind of micro structure which could provide extra semantic information about one or more aspects of the data set. The metastructure is defined as follow:

Definition 4. ***Metastructure:*** *Let S_G is one substructure of graph G and $V_i \in S_G$, so there are $S_G \subseteq G$ and $V_i \in G$. If any V_i meets $F_{S_G} = f(v_i)$, then S_G is called the metastructure with respect to F. Where F_{S_G} indicates the function of substructure S_G, and f is the mapping of v_i to F_{S_G}.*

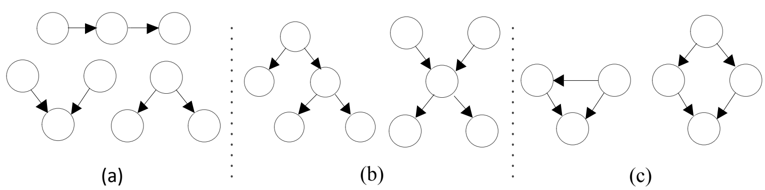

Fig. 3. basic structures and metastructures

The basic structures include three types in Fig.3(a), such as chain type, discrete type and convergent type also called V type structure. As the composite structures, Fig.3(b) shows some classic simply connected structures, and multiply connected structures are in Fig.3(c).

Sign propagation of metastructures inference with basic decomposition is similar with mesoscale substructures inference with edge-deletion, except that the former is bottom-up based on depth-first search and the latter is top-down basd on breadth-first search. At the beginning of the metastructures inference, we first find the nodes set with zero input degree and randomly pick up one of the nodes as the start point. Then subnetwork is generated with along one arc by depth-first search algorithm until contract all nodes through substructures. It is obvious that the time complexity of the algorithm is $O(|E|^2)$ at the worst case, where $|E|$ is the number of edges in the network.

4 An Inference Case

The qualitative inference case with conflicts as Fig.4(a) is introduced in [8]. Supposed that the value true has been given for the observed node H, now we

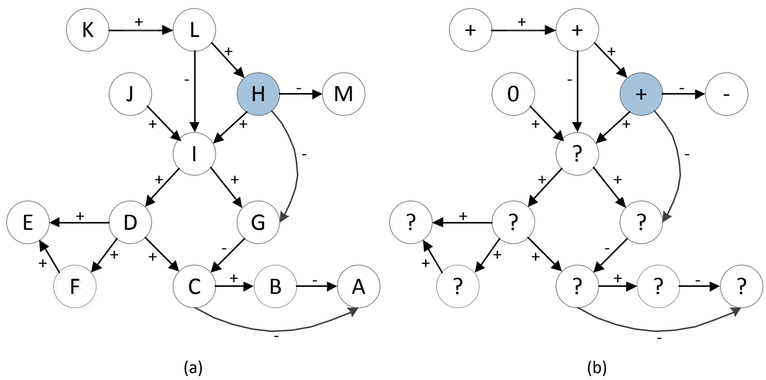

Fig. 4. An inference case with conflict

focus on its influence on the other nodes. By tracing the effect of observing every node's value on the other nodes in the network via message delivery between related nodes, the sign propagation result for each node is shown in Fig.4(b). The ambiguous signs reveal that there are at least one conflict existing in the network. For example, Node I results an ambiguous sign from the combination of the sign of node H to I and the sign of node L to I. And it is noticed that node J without input data has no influence on I. In the next, Fig.5 shows the qualitative inference methods based substructures how to solve the trade-off problems.

It indicates enhanced sign propagation in Fig.5(a), which can be used to distinguish sign strength. And Fig.5(b) illustrate how to solve the trade-off problems with edge deletion. Firstly, breadth-first search is carried out until conflict happens. Secondly. the edge $L \to I$ on the node I with conflicts is deleted in Fig.5(b). In this case, we take the enhanced sign propagation method to simplify calculation. The influence of node H on I is $+++$, while the effect of L on I is $+-$. So $L \to I$ is deleted, the composed influence on I is updated, and then the soft evidence E_I is added upon node L. Similarly, when the arcs $H \to G$ and $G \to C$ are removed, the network without conflict is separated into three inference submodules. In the end, we can easily draw conclusions on influence between nodes with lower computation cost. The time complexity is $O(|V|+|E|+|D|)$, where $|V|$ is the number of all nodes and $|E|$ is the number of edges in the whole network. As the cost of edge deletion operation, $|D|$ mainly is depended on the times of conflicts. When $|E|-1$ edges are deleted at the worst case, the time complexity of the algorithm is $O(|V|+2|E|)$.

According to the basic substructures in Fig.3, we can easily decompose the whole network into the structures S_G without conflict:(KL), (HM), (JI), (CBA), (DFE) and S_G with conflict:(LHI), (HIG), $(IDGC)$. In fact, the metasubstructures with semantics would be composed of the basic substrutures and problem domain or experts knowledge. And the metasubstructures do well in solving structured problems.

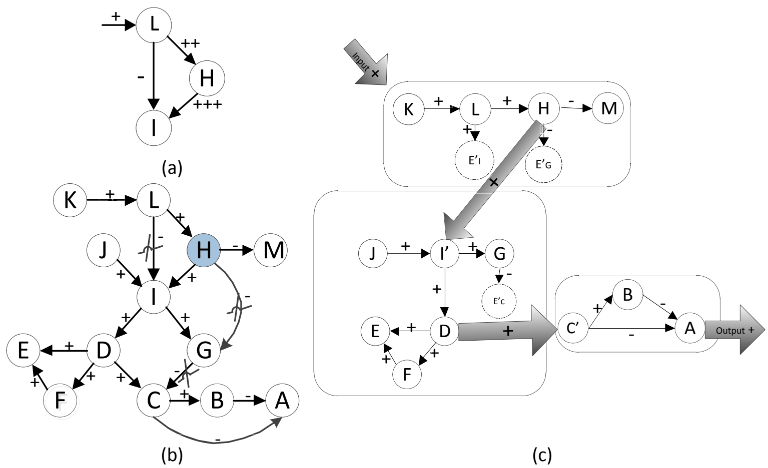

Fig. 5. mesoscale substructures inference with edge-deletion

Compare with the methods in the hierarchical qualitative inference model, mesoscale substructures inference solves trade-off problem better than micro and macro method, and its time complexity is much lower. However, the method with fine granules shows some conflict pattern or local knowledge structure to accelerate reasoning and benefit the network reconstruction. For example, substructure $S_G(CBA)$ both belongs to the mesoscale substructures and the metasubstructures. Furthermore, the two methods is alike in sparse networks, and but in dense problem substructures inference with edge-deletion is more effective because of lower network decomposition.

That is to summarize, hierarchical qualitative inference model could effectively evade conflicts and find the better inference structure by comparison with the methods on different granules than those methods based on experts.

5 Conclusions

This paper presents a hierarchical qualitative inference model with substructures including metastructures with basic decomposition and coarse-grained mesoscale structures with edge-deletion, which to some extent can eliminate the qualitative impact of uncertainty and solve the trade-off problems. Moreover, multi-level analysis could not only find better structure of inference network, but provide an approximate strategy for modular reasoning on large-scale problems. In the future, we would devote to designing the strategy and construct the network according to the experts knowledge.

Acknowledgements. The research in this paper is supported by the National Natural Science Foundation of China (Grant No. 60970061, 61075056).

References

1. Choi, A., Darwiche, A.: A Variational Approach for Approximating Bayesian networks by edge deletion. In: Proceedings of the 22nd Conf. UAI, pp. 80–89 (2006)
2. Druzdzel, M.J., Henrion, M.: Efficient Reasoning in Qualitative Probabilistic Networks. In: 11st National Conference on AAAI, pp. 548–553 (1993)
3. Feng, Q., Miao, D., Cheng, Y.: Hierarchical decision rules mining. Expert Systems with Applications 37(3), 2081–2091 (2010)
4. Li, X., Liao, S.: Hierarchical Reasoning in QPNs based on Network Decomposition. In: IEEE International Conference on ICIP, pp. 97–100 (2010)
5. Liu, C.L., Wellman, M.P.: Incremental Trade-off Resolution in Qualitative Probabilistic Networks. In: Proc. of Conf. UAI, pp. 338–345 (1998)
6. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, Palo Alto (1988)
7. Pedrycz, W.: Hierarchies of Architectures of Collaborative Computational Intelligence. International Journal of Software Science and Computational Intelligence, 18–31 (2009)
8. Renooij, S., van der Gaag, L.C., Parsons, S., Green, S.: Pivotal Pruning of Trade-offs in QPNs. In: Proc. of Conf. UAI, pp. 515–522 (2000)
9. Renooij, S., van der Gaag, L.C., Parsons, S.: Context-specific Sign-propagation in Qualitative Probabilistic Networks. Artificial Intelligence 144(1), 207–230 (2002)
10. Renooij, S., van der Gaag, L.C.: Enhanced qualitative probabilistic networks for resolving trade-offs. Artificial Intelligence 172(12-13), 1470–1494 (2008)
11. Renooij, S.: Bayesian network sensitivity to arc-removal. In: Proceedings of the Fifth European Workshop on Probabilistic Graphical Models, pp. 233–240 (2010)
12. Van Kouwen, F.A., Renooij, S., Schot, P.: Inference in Qualitative Probabilistic Networks revisited. International Journal of Approximate Reasoning 50(5), 708–720 (2009)
13. Wellman, M.P.: Fundamental Concepts of Qualitative Probabilistic Networks. Artificial Intelligence 44, 257–303 (1990)
14. Yao, J.T.: A ten-year review of granular computing. In: Proc. of the IEEE International Conference on Granular Computing, San Jose, USA, pp. 734–739 (2007)
15. Yao, Y.Y.: Integrative Levels of Granularity. In: Bargiela, A., Pedrycz, W. (eds.) Human-Centric Information Processing Through Granular Modeling, pp. 31–47. Springer, Berlin (2009)
16. Yue, K., Liu, W.: Qualitative probabilistic networks with rough-set-based weights. In: Proc. of ICMLC, vol. 3, pp. 1768–1774 (2008)
17. Yue, K., Liu, W., Yue, M.: Quantifying Influences in the Qualitative Probabilistic Network with Interval Probability Parameters. Applied Soft. Computing 11, 1135–1143 (2011)