

基于 d-分隔的 Credal 推理网络约简算法研究

瞿 英^{1,2} 吴祈宗¹ 崔春生^{1,3}

(北京理工大学管理与经济学院 北京 100081)¹ (河北科技大学经济与管理学院 石家庄 050081)²

(河南财经学院信息学院 郑州 450002)³

摘 要 针对 Credal 网络推理应用中出现的 Credal 集顶点组合爆炸等问题,系统地提出了 Credal 推理网络约简的概念,并利用 Credal 网络中变量间的 d-分隔特性,通过变量间独立性验证,设计了求解大规模 Credal 推理网络约简的算法。应用实例表明,该算法简化了目标网络,有效规避了 Credal 集顶点组合爆炸问题,对特定推理问题具有适用性和可行性,提高了 Credal 网络推理的效率。

关键词 d-分隔, Credal 网络, 推理网络约简

中图法分类号 TP301.6 **文献标识码** A

Credal Network Inference Reduct Algorithm Based on d-separate

QU Ying^{1,2} WU Qi-zong¹ CUI Chun-sheng^{1,3}

(School of Management and Economics, Beijing Institute of Technology, Beijing 100081, China)¹

(School of Economics and Management, Hebei University of Science and Technology, Shijiazhuang 050081, China)²

(School of Information, Henan Institute of Finance & Economics, Zhengzhou 450002, China)³

Abstract According to the problem such as vertexes of Credal set combination explosion in the inference of Credal Network, inference network reduction was proposed. By using d-separate among the variables, an algorithm to compute inference network reduction of large scale network was designed. The application case indicated that the algorithm could simplify the inference network and improve the efficiency of inference. It was useful and feasible for some kinds of inference.

Keywords d-separate, Credal network, Inference network reduct

1 引言

Credal 网络是贝叶斯网络的扩展,通过概率分布集合或凸集分布理论在参数概率值上引入不确定性,来表示有关参数概率的知识不完备性和模糊性,符合人类专家认识问题的客观规律,从而支持研究概率模型下模糊不确定信息对决策结果的影响,并为可靠性决策提供理论依据。Credal 网络的本质是放宽对参数概率严格要求的贝叶斯网络,使其不必为每个节点指定在其父节点一定配置下的唯一的概率分布,而是可以用概率分布集合、概率分布区间、不等式约束等不同的概率测度方法来表达专家或数据信息形成的对参数概率值的不确定性表达^[1]。

Credal 网络的推理建立在贝叶斯网络推理的基础之上,计算在一定扩展和证据条件下的某一变量的分类值的最大/最小概率(即区间的上下界)。强扩展类型下的查询变量的后验概率计算问题可以看作在 Credal 集约束下的贝叶斯网络推理的全局优化问题^[2]。解决这种优化问题的精确推理算法主要有枚举法、分支定界法、2U 算法。然而,这些推理算法遇到的困难在于各边缘 Credal 集或条件 Credal 集给 Credal 网络强扩展带来了大量的顶点组合,即使对于小型的网络也容易

易出现顶点组合爆炸问题,因此人们尝试考虑近似推理算法。Cano 等提出采用启发式搜索算法^[3-8],如模拟退火算法、遗传算法、粒子群算法及蚁群算法等,这些算法的思想是通过概率技术在人工变量配置的集合中选择配置好的解。目前的推理算法无论是精确推理算法还是近似推理算法,都是在这个目标网络中进行的,推理概率传播涉及全部变量,而在实际推理中,有些变量对目标变量后验概率的求解是不起作用的。因此在进行推理时可以不考虑这些变量,从而使推理网络得到约简。文献[8]提出了面向目标的 Credal 网络转化方法,将目标网络转化为由查询变量和目标变量的祖先节点组成的简化网络进行推理。本文尝试利用 Credal 网络中变量间 d-分隔的特性对该简化网络进行进一步的约简,这样就把对复杂网络的推理转化为在相对简单的网络上进行推理,从而提高推理的效率。

2 d-分隔及 Credal 推理网络约简

d-分隔是 Credal 网络中的重要概念之一,d-分隔与顺连节点、分连节点、汇连节点^[9]密切相关。

2.1 d-分隔的概念

定义 1^[9](阻塞) 设 Z 为一节点集合, X 和 Y 是不在 Z

到稿日期:2009-01-06 返修日期:2009-03-16 本文受河北省教育厅基金项目(2007205),河北省科技厅软科学项目(074572214)资助。

瞿 英(1973-),女,博士生,副教授,主要研究方向为数据挖掘、信息管理理论与方法等,E-mail:quying1973@126.com;吴祈宗(1947-),男,教授,博士生导师,主要研究方向为信息管理理论与方法、管理决策的方法与技术等;崔春生(1974-),男,博士生,主要研究方向为信息管理理论与方法等。

中的两个节点,考虑 X 和 Y 之间的一条通路 α ,如果满足下面条件之一,则称 α 被 Z 阻塞

- (1) α 上有一个在 Z 中的顺连节点;
- (2) α 上有一个在 Z 中的分连节点;
- (3) α 上有一个汇连节点 W ,它和它的后代节点均不在 Z 中。

定义 2^[9] (d-分隔) 如果 X 和 Y 之间的所有通路都被 Z 阻塞,那么我们就说 Z 有向分隔(directed separate) X 和 Y ,简称 Z d-分隔 X 和 Y 。

定理 1^[9,10] (整体马尔可夫性) 设 X 和 Y 为贝叶斯网 $\mathcal{N}$ 中的两个变量, Z 为 $\mathcal{N}$ 中不包含 X 和 Y 的节点集合,如果 Z d-分隔 X 和 Y ,那么 X 和 Y 在给定 Z 时条件独立,即 $X \perp Y | Z$ 。

2.2 Credal 推理网络约简

Credal 网络的推理通常是用来计算在一定证据条件下所要考察的节点上的分类值的概率区间。即如果 X_q 是一个查询变量,而 X_E 是一个观察变量集合,那么“推理”就是要计算 X_q 上一个或者某几个分类值的条件概率的区间 $K(X_q | X_E)$ 上下界^[11-12]。如果要得到 $K(X_q | X_E)$,在给定 X_E 的条件下,仅仅需要与查询变量 X_q 有关的变量,其余的变量对于该推理是冗余的。也就是说,删除该条件下与 X_q 独立的变量,对该推理是无影响的,因此推理网络是可以约简的。

定义 3(推理网络约简) 设 $\mathcal{N}$ 为 Credal 网络, X 为 $\mathcal{N}$ 中所有变量集合, $X' \subseteq X$,若对于由 X' 构成的 Credal 网络 $\mathcal{N}'$ 满足

$$K_{\mathcal{N}}(X_q | X_E) = K_{\mathcal{N}'}(X_q | X_E)$$

其中, X_q 为查询变量, X_E 为证据变量,则称 $\mathcal{N}'$ 是相对于 X_q 和 X_E 的 $\mathcal{N}$ 的一个约简。设 $\mathcal{N}$ 中各变量是等质的,即各变量所对应的 Credal 集顶点数是相同的,在 $\mathcal{N}$ 的所有约简中,变量个数最少的约简称为最小约简。

由约简的定义自然而然想到 d-分隔的特性,约简问题就转化为求一个变量集合 Y , Y 是被 X_E 将其与 X_q d-分隔的所有节点集合。因为 X_E d-分隔 X_q ,由定理 1 可知, $X_q \perp Y | X_E$,因此

$$K_{\mathcal{N}}(X_q | X_E, Y) = K_{\mathcal{N}}(X_q | X_E) = K_{\mathcal{N}'}(X_q | X_E)$$

2.3 Credal 网络推理约简思路

推论 1^[9] 在一个 Credal 网络中,给定变量 X 的马尔可夫边界 $mb(X)$,则 X 独立于网络中所有其他变量。

约简 1 利用推论 1 可知,若 $mb(X_q) \subseteq X_E$,则由 $\{X_q \cup mb(X_q)\}$ 组成的 $\mathcal{N}'$ 即为原目标网络的一个约简。

定理 2^[9,10] 给定一 Credal 网络 $\mathcal{N}$,设 $P(X_q | X_E = e)$ 是需要计算的后验概率。令 $\mathcal{N}'$ 为从 $\mathcal{N}$ 中除去在 $an(X_q \cup X_E)$ 之外的所有节点后所得的贝叶斯网络,则

$$K_{\mathcal{N}}(X_q | X_E = e) = K_{\mathcal{N}'}(X_q | X_E = e)$$

约简 2 利用定理 2,可以得到目标网络的一个从 $\mathcal{N}$ 中除去在 $an(X_q \cup X_E)$ 之外所有节点后所得的约简 $\mathcal{N}'$ 。

约简 3 在上述约简基础上,继续利用 d-分隔的特性,还可以找出使得被 X_E 将其与 X_q d-分隔的所有变量集合,因此还可以对已得到的约简 $\mathcal{N}'$ 进一步简化。但在这个过程中,需要将 $\mathcal{N}'$ 转化为端正图(moral graph),求解被 X_E 将其与 X_q d-分隔的所有变量集合,即 $\mathcal{N}'$ 上被 X_E 将其与 X_q d-分隔的所有变量集合,从 $\mathcal{N}'$ 中将其删去,得到 $\mathcal{N}'$ 的一个约简。

3 Credal 推理网络约简算法设计

输入:一个 Credal 网络 $\mathcal{N}(X, A(X))$, X_q, X_E

输出: $\mathcal{N}$ 的一个约简 $\mathcal{N}'$

Step1 计算 $mb(X_q), an(X_q \cup X_E)$

//计算查询变量及证据变量的马尔可夫界和祖先节点。

Step2 $X_m = an(X_q \cup X_E)$,

$$\mathcal{N}'(X', A(X')) = \mathcal{N}'(X_m, A(X_m))$$

Step3 if $mb(X_q) \subseteq X_E, X_r = X \setminus X_q \cup mb(X_q)$,

For $X_i \in X_r$,

If $X_i \in X_m$ Then $X_m = X_m \setminus X_i$

$$\mathcal{N}'(X', A(X')) = \mathcal{N}'(X_m, A(X_m))$$

End//约简 1

Step4 //约简 3

建立 $\mathcal{N}'$ 的道德图 $\mathcal{N}^0$;

在 $\mathcal{N}^0$ 上 delete X_E 及 $A(X_E)$;

$$X_i = X' \setminus (X_q \cup X_E)$$

For every X_i ,

在 $\mathcal{N}^0$ 中找连接 X_i 与 X_q 的通路,

If not find Then $X' = X' \setminus X_i$

Output $\mathcal{N}'(X', A(X'))$

4 算法应用实例

这里采用文献[7]中的有关港口生产安全评价的 Credal 网络模型对基于推理网络约简的近似推理算法进行评估,网络的拓扑结构如图 1 所示。Credal 网络模型拥有 30 个节点变量,属于中等规模。对于这个实例来说,强扩展下的顶点组合数是 3072 个。如果采用传统的枚举算法进行求解,那么将花费很长的时间。

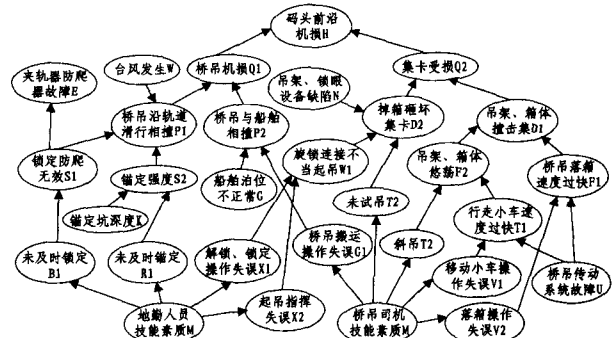


图 1 港口装卸作业安全评价 Credal 网络拓扑结构

选取 $G1, P1$ 为证据节点, $Q1$ 为要考察的目标节点。

Step1 $mb(X_q) = mb(Q1) = \{P1, P2, H, Q2\}$

$$an(X_q \cup X_E) = an(Q1, G1, P1) = \{P1, P2, G1, M, G, S2, W, S1, B1, K, R1, L\}$$

Step2 $X_m = an(X_q \cup X_E)$

建立祖先闭集上的约简,结果如图 2 所示。

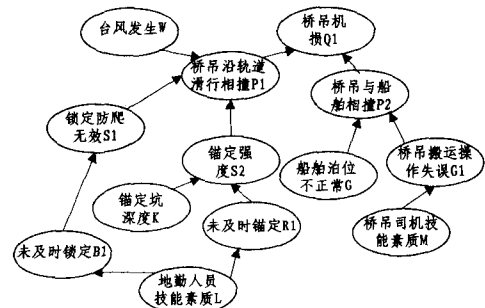


图 2 祖先闭集上的约简

Step3 由于 $mb(Q1) = \{P1, P2, H, Q2\}$, $mb(X_q) \not\subseteq X_E$, 故 Step3 跳过。

Step4 $M, S2, S1, W, B1, K, L, R1$ 均被 $P1$ 与 $Q1$ d-分隔, 所以可以删去, 得到图 3 所示的约简结果。

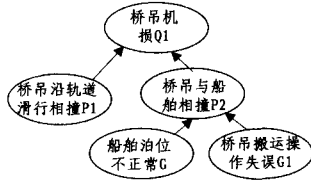


图 3 基于 d-分隔的约简

由此可见, 本算法将一个具有 30 个节点的推理网络缩减为仅有 5 个节点的推理网络, 缩小了网络规模, 大大提高了推理效率。

结束语 利用 Credal 网络进行推理在预测和诊断等领域具有广泛的应用前景, 而如何解决大规模 Credal 网络的推理效率问题, 目前仍是机器学习领域中的热点和难点问题。利用 Credal 网络中变量间 d-分隔的特性, 本文设计了一种求解大规模 Credal 推理网络约简的算法。应用实例表明, 本算法有效缩小了 Credal 推理网络的规模, 规避了 Credal 集顶点组合爆炸问题, 对特定问题的处理具有良好效果。

参考文献

- [1] Cozman F. CredalNetworks[J]. Artificial Intelligence, 2000, 120(2): 199-233
- [2] Fagioli E, Zaffalon M. 2U: An exact interval propagation algorithm for polytrees with binary variables[J]. Artificial Intelligence, 1998, 106(1): 77-107

- [3] Cozman F. Robustness analysis of Bayesian networks with local convex sets of distributions[C]// Proceedings of the Thirteenth Annual Conference on Uncertainty in Artificial Intelligence. Morgan Kaufmann, San Mateo, 1997: 108-115
- [4] Cano A, Cano J E, Moral S. Convex sets of probabilities propagation by simulated annealing[C]// The Fifth International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems, Paris, France, 1994: 4-8
- [5] Cano A, Moral S. A genetic algorithm to approximate convex sets of probabilities[C]// Proceedings of Information Processing and Management of Uncertainty in Knowledge-based Systems Conference (IPMU'96). Granada, Spain, 1996: 859-864
- [6] Cano A, Cano J E, Moral S. Simulation algorithms for convex sets of probabilities[R]. TR-93-2-xx. DECSAI, Universidad de Granada, 1993
- [7] 郑恒. Credal 网络及其在安全评价中的应用研究[D]. 北京: 北京理工大学, 2006
- [8] 李玉玲. Credal 网络及其应用研究[D]. 北京: 北京理工大学, 2007
- [9] 张连文, 郭海鹏. 贝叶斯网引论[M]. 北京: 科学出版社, 2006
- [10] Acid S, De Campos L M. An Algorithm for Finding Minimum d-Separating Sets in Belief Networks[C]// Proceedings of the Twelfth Conference of Uncertainty in Artificial Intelligence. 1996
- [11] 马少平, 朱小燕. 人工智能[M]. 北京: 清华大学出版社, 2004
- [12] Antonucci A, Zaffalon M, Ide J S, et al. Binarization algorithms for approximate updating in Credal nets[C]// Proceedings of The third European Starting AI Researcher Symposium. Amsterdam (Netherlands): IOS Press, 2006: 120-131

(上接第 178 页)

$\succ$ -核为 $\{a_1, a_4\}$ 。

$\Delta^< = a_1 \wedge a_3 \wedge a_4$, $\prec$ -约简是 $\{a_1, a_3, a_4\}$, $\prec$ -核为 $\{a_1, a_3, a_4\}$ 。

$\Delta^> (x_1) = \bigwedge \{ \bigvee D_{ij}^> : D_{ij}^> \in M_0^> \} = a_1 \wedge (a_2 \vee a_3) = (a_1 \wedge a_2) \vee (a_1 \wedge a_3)$, x_1 的 $\succ$ -约简是 $\{a_1, a_2\}$ 与 $\{a_1, a_3\}$ 。

$\Delta^< (x_1) = a_4$, x_1 的 $\prec$ -约简是 $\{a_4\}$ 。

类似地可得:

x_2, x_4 的 $\succ$ -约简是 $\{a_1\}$, x_3 的 $\succ$ -约简是 $\{a_1, a_4\}$, x_5 的 $\succ$ -约简是 $\{a_3\}$ 与 $\{a_2, a_4\}$, x_6 的 $\succ$ -约简是 $\{a_2, a_3\}$; x_2, x_3 的 $\prec$ -约简是 $\{a_3\}$, x_4 的 $\prec$ -约简是 $\{a_2\}$, $\{a_1, a_3\}$ 与 $\{a_3, a_4\}$, x_5 的 $\prec$ -约简是 $\{a_1\}$, x_6 的 $\prec$ -约简是 $\{a_1, a_4\}$ 。

结束语 集值信息系统反映了信息的不确定性, 亦即信息是有噪声的。在信息不确定的情况下, 不依赖于领域专家清除噪声, 直接建立知识发现方法, 具有一定现实意义。基于此, 本文利用属性集值的取值大小在集值信息系统上定义了两种不同类型的优势关系, 它们都是经典信息系统上优势关系的推广。然后根据粗糙集方法讨论了在这两种优势关系下集值信息系统的知识约简与对象的相对约简, 给出了知识约简与对象的相对约简的判定, 并利用布尔推理方法得到了集值信息系统在这两种优势关系下的知识约简与对象的相对约简的具体操作方法。至于集值决策信息系统在这两种优势关系下的知识约简与决策规则的提取与优化, 我们将另文叙述。

参考文献

- [1] Pawlak Z. Rough set[J]. International Journal of Computer and Information Science, 1982, 11: 341-356
- [2] Pawlak Z. Rough set: Theoretical Aspects of Reasoning about Data[M]. Boston: Kluwer Academic Publishers, 1991
- [3] 张文修, 梁怡, 吴伟志. 信息系统与知识发现[M]. 北京: 科学出版社, 2003
- [4] Greco S, Matarazzo B, Slowinski R. Rough sets theory for multi-criteria decision analysis[J]. European Journal of Operational Research, 2001, 129: 1-47
- [5] Kryszkiewicz M. Rough set approach to incomplete information systems[J]. Information Sciences, 1998, 112: 39-49
- [6] Kryszkiewicz M. Rules in incomplete information systems[J]. Information Sciences, 1999, 113: 271-292
- [7] Leung Y, Li D Y. Maximal consistent block technique for rule acquisition in incomplete information systems[J]. Information Sciences, 2003, 153: 85-106
- [8] Guan Yan-Yong, Wang Hong-Kai. Set-valued information systems[J]. Information Sciences, 2006, 176: 2507-2525
- [9] 宋笑雪, 解争龙, 张文修. 集值决策信息系统的知识约简与规则提取[J]. 计算机科学, 2007, 34(4): 182-184
- [10] 陈子春, 秦克云. 集值信息系统基于变精度相容关系的属性约简[J]. 计算机工程与应用, 2008, 44(9): 27-29
- [11] Skowron A, Rauszer C. The Discernibility Matrices and Functions in Information Systems [M]. Boston: Kluwer Academic Publishers, 1992: 331-362