

基于混合方式的贝叶斯网弧定向算法

贾海洋, 陈娟, 朱允刚, 刘大有

(吉林大学计算机科学与技术学院, 吉林大学符号计算与知识工程教育部重点实验室, 吉林长春 130012)

摘要: 贝叶斯网是不确定知识表示及推理的主要方法之一, BNs 结构中的因果关系在知识建模中起到十分关键的作用, 因此确定 BNs 中弧的方向是一重要问题. 目前已有的方法存在以下问题: (1) 算法计算复杂性高; (2) 将统计不可分的弧定向, 可能与领域知识不符. 本文提出一种结合条件独立测试和打分搜索的 BNs 弧定向方法. 该方法仅执行零阶和一阶条件独立测试, 执行次数为多项式级; 打分搜索可分解为局部子图的搜索, 提高了算法的效率. 算法输出结果为最大链图, 该图仅对统计可分的弧进行定向, 对统计不可分的弧保留无向的特性. 这种结果更准确的表现了数据中蕴含的因果关系, 便于结合领域知识进行建模.

关键词: 贝叶斯网; 弧定向; 马尔科夫等价类; 链图

中图分类号: TP11 **文献标识码:** A **文章编号:** 0372-2112 (2009) 08-1842-06

A Hybrid Method for Orienting Edges of Bayesian Network

JIA Hai-yang, CHEN Juan, ZHU Yun-gang, LIU Da-you

(College of Computer Science and Technology, Key Laboratory for Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, Jilin 130012, China)

Abstract: Bayesian network is one of the most important methods for representing and inferring with uncertainty knowledge, causal relation between variables is a key property for modeling the knowledge, so it is an important problem to orient the edges. There are some problems in the exist methods: (1) the computational complexity of the algorithms is high; (2) orienting the statistical indistinguishable edges may inconsistent with the domain knowledge. This paper presents an algorithm which combining the constraint approach and score search approach to orient the edges. The time of zero order and first order conditional independent test is polynomial; The search space can be decomposed to sub graph, which improving the efficiency of the algorithm. The output of the algorithm is largest chain graph, which just orienting the statistical distinguishable edges, keep the indistinguishable edges undirected. This graph present the causal relation more accuracy, and more convenient for combining domain knowledge.

Key words: bayesian network; orienting edges; Markov equivalence class; chain graph

1 引言

贝叶斯网 (Bayesian Networks, BNs) 是概率模型的一种图形化表示方法, 其坚实的理论基础、自然的知识表示方式使其在不确定知识表示及推理方面表现出卓越性能, 为不确定人工智能提供了有力工具. BNs 的知识表示包括两部分: 定性部分 (网络结构) 和定量部分 (条件概率表). BNs 的网络结构是有向无环图 (Directed Acyclic Graph, DAG), DAG 不仅表示节点间的条件独立关系, 有向边也可表示节点间的因果关系. 因果关系的确定在知识建模过程中起到十分重要的作用, 因此确定 BNs 弧的方向是一重要研究问题.

目前, BNs 弧定向方法主要包括: 基于约束的方法^[1,2]和基于打分搜索的方法^[3]. 基于约束的方法将该问题看做约束满足问题, 通过条件独立 (CI, Conditional Independence) 测试找出各变量间的独立关系, 然后构造与这些独立关系一致的弧定向方案; 基于打分搜索方法将该问题看做组合优化问题, 运用评分函数对网络模型进行评价, 在所有候选结构中找出一个得分最高的弧定向方案. 两类方法均存在一定问题: (1) 算法执行效率较低: 约束的方法需要互信息测试次数为指数级, 且随着 CI 测试阶数的增加, 测试误差迅速增大, 打分搜索的搜索空间为指数级; (2) 将统计不可分的弧定向, 可能与领域知识不符: 给定概率模型条件下, 某些弧的方向无

收稿日期: 2007-07-31; 修回日期: 2009-01-20

基金项目: 国家自然科学基金重大项目 (No. 60496321); 国家自然科学基金 (No. 60373098, No. 60573073, No. 60603030, No. 60503016); 国家 863 高技术研究发展计划 (No. 2006AA10Z245); 吉林省科技发展计划重大项目 (No. 20020303); 吉林省科技发展计划 (No. 20030523); 欧盟项目 TH/Asia Link/010 (No. 111084)

法通过统计方法辨别方向(弧方向不同,但对应相同的概率模型,此时需要领域知识进行定向)。

本文提出一种结合 CI 测试和打分搜索的弧定向方法.该方法 CI 测试阶段仅需进行多项式级的 CI 测试(最坏情况下 $O(n^2)$),打分搜索阶段在子图构成的搜索空间进行,提高了算法的运行效率.算法输出结果为最大链图,该图仅对统计可分的弧进行定向,保留统计不可分弧的无向特性.这种结果更准确的表现了数据中蕴含的因果关系,便于结合领域知识进行建模。

2 背景知识

符号约定:在论述随机变量时大写字母 X, Y, Z 表示随机变量;小写字母 x, y, z 表示随机变量的取值;在论述集合时大写字母 V, E 表示集合,小写字母 v, e 表示集合中的元素,圆括号表示无序对 (v_1, v_2) ,尖括号表示有序对 $\langle v_1, v_2 \rangle$ 。

2.1 贝叶斯网

贝叶斯网(又称为信念网、概率网络或因果网)是由路径分析,因果模式,影响图等逐渐演变而来,1988 年, Pearl 建立了贝叶斯网基础理论体系,将概率理论和图论有机结合,用一种紧凑的形式表示联合概率分布^[4]。

定义 1 贝叶斯网(Bayesian Networks, BNs)表示 n 个随机变量 $X = (X_1, X_2, \dots, X_n)$ 的联合概率分布,由两部分组成 $BNs = \langle G, \theta \rangle$,其中: G 是一个有向无环图(Directed Acyclic Graph, DAG),其节点对应于 X 中的随机变量 X_1, X_2, \dots, X_n ,有向弧表示因果关系,及节点间的条件独立关系; θ 代表用于量化网络的一组参数, $\theta = \{p(X_i | \pi_i) | 1 \leq i \leq n\}$ (π_i 为节点 X_i 的父节点集),称为条件概率表(Conditional Probability Table, CPT)。

G 是对联合概率分布的定性表示,表明随机变量之间的条件独立关系; θ 是对联合概率分布的定量表示,表明在已知 π_i 的情况下 X_i 发生的条件概率.利用 BNs 所表示的条件独立关系和条件概率可以紧凑的将联合概率分布表示为:

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | \pi_i)$$

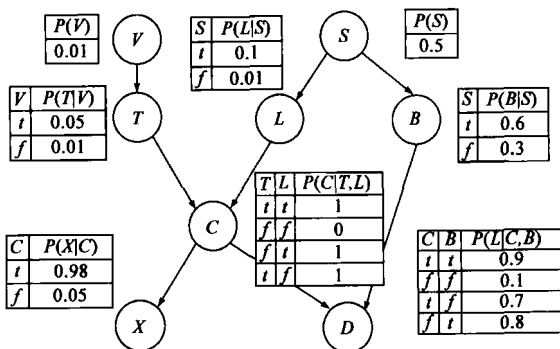


图1 贝叶斯网举例

图 1 是一个 BNs 的例子,有 8 个随机变量 V, S, T, L, B, C, X, D ,每个随机变量具有真、假 (t, f) 两种取值状态(为了表示简洁,依据概率和为 1 的原则 CPT 中的概率值省略了一半).其联合概率分布表示为(每一个条件概率都可以在 CPT 中查得):

$$P(V, S, T, L, B, C, X, D) = P(V)P(S)P(T|V)P(L|S)P(B|S)P(C|T, L)P(X|C)P(D|C, B)$$

2.2 马尔可夫等价

不同的 BNs 之间存在等价关系,等价的 BNs 具有相同的表示能力,可表示同一联合概率分布,这种等价关系称为马尔可夫等价(Markov Equivalence, ME)。

定义 2 称两个贝叶斯网 BN_1 和 BN_2 马尔可夫等价当且仅当

- (1) BN_1 和 BN_2 包含相同节点;
- (2) BN_1 所能表示的任意联合概率分布 BN_2 均能表示,反之亦然。

两个 BNs 之间是否马尔可夫等价,可以通过其网络结构判断,下面一组定义和定理给出判断方法^[5]。

定义 3 在有向图 $G = (V, E)$ 中,节点 $v_1, v_2, v_3 \in V$,如果 $\langle v_1, v_2 \rangle \in E, \langle v_3, v_2 \rangle \in E$ 且 $\langle v_1, v_3 \rangle \notin E, \langle v_3, v_1 \rangle \notin E$,则称节点 v_1, v_3 在 v_2 处形成 V 结构,记为 $V(v_1, v_3 | v_2)$ 。

定义 4 将有向图 G 的有向边变为无向边,得到的无向图称为图 G 的骨架(skeleton)。

定义 5 具有相同节点集的两个有向无环图 DAG_1 和 DAG_2 图等价,当且仅当

- (1) DAG_1 和 DAG_2 具有相同的骨架;
- (2) DAG_1 和 DAG_2 具有相同 V 结构。

定理 1 贝叶斯网 $BN_1 = (G_1, \theta_1)$ 和 $BN_2 = (G_2, \theta_2)$ 若其网络结构 G_1 和 G_2 图等价,则 BN_1 和 BN_2 马尔可夫等价^[6]。

图 2 给出马尔可夫等价的例子(线框内的 DAG 等价,为描述方便下文不再区分 BNs 之间的马尔可夫等价和图等价).大部分打分函数对等价的 DAG 赋予相同的分数,例如: AIC (Akaike Information Criterion)^[7], BIC (Bayesian Information Criterion)^[8] 和 BDe (Bayesian Dirichlet

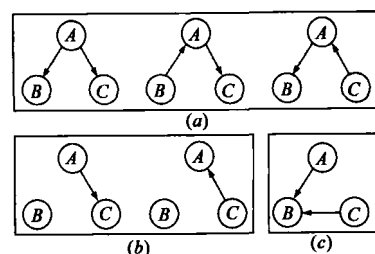


图2 马尔可夫等价举例

equivalent)^[9], 这类打分标准称为分数等价^[10]. 等价结构组成的集合称为马尔可夫等价类 (Markov Equivalence Class, MEC).

2.3 链图及最大链图

定义 6 图 $G = (V, E)$, 若 E 中边均为有向边, 则称 G 为有向图 (DG, Directed Graph); 若 E 中边均为无向边, 则称 G 为无向图 (UG, Undirected Graph), 若 E 中既包含无向边又包含有向边, 则 G 称为混合图 (HG, Hybrid Graph).

定义 7 混合图 HG 中存在一条路径 $e_0, e_1, e_2, \dots, e_n$, 满足: $e_0 = e_n, n > 2, e_i$ 和 e_{i+1} 之间存在一条无向边 (e_i, e_{i+1}) 或有向边 $\langle e_i, e_{i+1} \rangle$ 且路径上至少存在一条有向边, 则称该路径为部分有向环 (PDC, Partial Directed Cycle).

定义 8 图 $G = (V, E)$ 称为链图 (CG, Chain Graph), 当且仅当 G 中不包含部分有向环.

定义 9 删除链图中的有向边, 得到的无向图中每个最大连通子图称为一个链分量 (CC, Chain Component).

定义 10 链图 $CG = (V, E)$ 包含 m 个链分量 $CC_i = (V_i, E_i), 1 \leq i \leq m, bd(CC_i) = \{v_j | v_j \in V - V_i, \langle v_j, v_k \rangle \in E, \exists v_k \in V_i\}$ 称为 CC_i 的边界 (border), $cl(CC_i) = bd(CC_i) \cup V_i$ 称为 CC_i 的闭包 (closure).

CG 也是概率模型的图形表示方式之一, 马尔可夫网和贝叶斯网可以看作是 CG 的特殊情况 (关于 CG 可参阅文献[11]). 不同 CG 之间也存在马尔可夫等价, 定理 1 的判别方法同样适用于 CG^[12]. 例如图 3 中的 CG 与图 2(a) 中 BNs 等价.

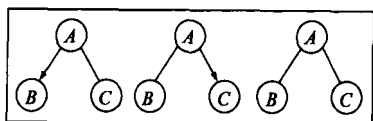


图3 马尔可夫等价的链图

定义 11 与概率模型 M 等价的所有链图中, 包含无向边最多的链图称为最大链图 (LCG, Largest Chain Graph).

定理 2 任何马尔可夫等价类都存在唯一最大链图与之等价^[13].

因此 LCG 可以作为 MEC 的图形化表示. 且任意弧 A-B 若在 LCG 中有向, 则在所有等价的 CG 中有向且方向与 LCG 中相同; 若弧在 LCG 无向, 则在与之等价的不同 CG 中会出现不同方向. 例如图 3 中最后一个 CG 为该 MEC 的 LCG.

3 弧定向算法

由马尔可夫等价的判别条件可知: 给定无向弧的集合的条件下, 若能确定所有 V 结构, 即可确定 MEC, 确定 MEC 则可以确定唯一的 LCG, 达到弧定向的目的.

本文提出的弧定向算法分为两个阶段: CI 测试阶段和局部搜索阶段.

3.1 CI 测试

BNs 结构可以看作是对变量之间条件独立关系的图形编码, CI 测试的目的是通过统计测试方法找出训练数据中各变量间的条件独立关系.

本文采用互信息和条件互信息进行 CI 测试. 变量 A 和 B 之间的互信息定义为:

$$I(A, B) = \sum_{a,b} P(a, b) \log \frac{P(a, b)}{P(a)P(b)}$$

变量 A 和 B 以变量 C 为条件的条件互信息定义为:

$$I(A, B | C) = \sum_{a,b,c} P(a, b, c) \log \frac{P(a, b | c)}{P(a | c)P(b | c)}$$

若 $I < \epsilon$ 则变量条件独立, 表示为 $Ind(A, B)$ 和 $Ind(A, B | C)$. ϵ 为阈值 (一般取 $\epsilon = 0.01$). 理论上, 通过 CI 测试可以得到全部 V 结构, 但需进行的 CI 测试数量为指数级, 且随着 CI 测试阶的增加 (条件集合 C 中变量个数为 CI 测试的阶数), 误差迅速增大. 文献[1]虽然将 CI 测试次数降为 $O(n^2)$, 但需要单调 DAG 可信 (monotone DAG faithfulness) 的前提假设, 对于实际问题这一假设很难被满足^[14]. 本文提出并证明如下 3 个命题, 基于这些命题设计的弧定向算法仅需要进行 0 阶和 1 阶 CI 测试, 调用次数取决于待定向弧的结构以及概率模型中包含的条件独立关系, 最坏情况下调用的次数为 $O(n^2)$. 算法如图 4 所示:

算法 $CIP(Data, SK, CG, UNV)$

/* 括号内用“.”分隔输入和输出变量, 用“,”分隔不同变量名

输入: Data 训练数据

SK 骨架 (无向图的邻接矩阵)

输出: CG 链图 (部分有向图的邻接矩阵)

UNV CI_0 和 CI_1 无法判定的 PV

(节点三元组的集合) */

INI[UNV 初始化为空, CG 初始为骨架 SK]

UNV = Φ

CG = SK.

CI[对所有 PV 进行 CI 测试]

FOR $\forall PV_{SK}(V_j, V_k | V_i)$

IF $((V_i, V_j) \in CG \vee (V_i, V_k) \in CG)$

/* 存在需要定向的无向边 */

IF $Ind(V_j, V_k)$

/* 若 $Ind(V_j, V_k)$ 则 $V(V_j, V_k | V_i)$ */

CG = $CG \cup \langle V_j, V_i \rangle \cup \langle V_k, V_i \rangle$.

ELSE

/* 若 $\neg Ind(V_j, V_k)$ 则需 CI_1 测试 */

IF $Ind(V_j, V_k | V_i)$

/* 若 $Ind(V_j, V_k | V_i)$, 则 $\neg V(V_j, V_k | V_i)$ */

IF $\langle V_j, V_i \rangle \in CG$

```

CG = CG ∪ ⟨Vi, Vk⟩.
ELSEIF ⟨Vk, Vi⟩ ∈ CG
CG = CG ∪ ⟨Vi, Vj⟩.
ELSE
/* 将三元组 (Vj, Vk, Vi) 加入 UNV */
UNVi,j = UNVi,j ∪ (Vj, Vk, Vi).
ACYC[若存在仅含一条无向弧的部分有向环,定向该弧使 CG 中
不存在有向环]
FOR ∀ PCyc(Vi, Vi+1...Vm-1, Vm) ∧ (Vj, Vj+1) ∈ CG ∧ ⟨Vk,
Vk+1⟩ ∈ CG
Vi = Vm, ∃ j ∈ [i, m-1], ∀ k ≠ j, k ∈ [i, m-1]
CG = CG ∪ ⟨Vj+1, Vj⟩
    
```

图 4 利用条件独立测试完成部分弧定向

定义 12 在骨架 SK 中,称 V_k, V_j 在 V_i 处构成可能 V 结构(PV),当且仅当: $(V_i, V_j) \in SK, (V_i, V_k) \in SK, (V_j, V_k) \notin SK$ 记为 $PV_{SK}(V_j, V_k | V_i)$ (在不引起歧义的情况下省略下标 SK).

命题 1 若 $PV(V_j, V_k | V_i)$ 且 $Ind(V_j, V_k)$, 则 $V(V_j, V_k | V_i)$.

证明 由 $PV(V_j, V_k | V_i)$ 可知在 LCG 上 V_j, V_k, V_i 存在两种可能的局部结构: $V_j \rightarrow V_i \leftarrow V_k$ 和 $V_j - V_i - V_k$. 在 $V_j - V_i - V_k$ 等价结构下若 V_i 未知, 则 V_j, V_k 通过 V_i 相互影响, 即 $\neg Ind(V_j, V_k)$; 由 $Ind(V_j, V_k)$ 可得 $\langle V_j, V_i \rangle \in E \wedge \langle V_k, V_i \rangle \in E$ 即 $V(V_j, V_k | V_i)$

命题 2 若 $PV(V_j, V_k | V_i)$ 且 $Ind(V_j, V_k | V_i)$, 则 $\neg V(V_j, V_k | V_i)$.

证明 若 $V(V_j, V_k | V_i)$, 则在 V_i 已知的情况下, V_j, V_k 通过 V_i 相互影响, 即 $\neg Ind(V_j, V_k | V_i)$; 由 $Ind(V_j, V_k | V_i)$ 可知 $\neg V(V_j, V_k | V_i)$.

命题 3 若 $PV(V_j, V_k | V_i)$ 且 $\neg Ind(V_j, V_k), \neg Ind(V_j, V_k | V_i)$, 则 V_j, V_k 间在骨架上至少存在一条不同时包含 $(V_i, V_j), (V_i, V_k)$ 的路径. 此时仅通过 CI_0 和 CI_1 测试无法确定是否 $V(V_j, V_k | V_i)$.

证明 分别对如下两种情况讨论证明:

(1) 若 $V(V_j, V_k | V_i)$, 则在 V_i 未知的情况下, V_j, V_k 不能通过 V_i 相互影响, 又已知 $\neg Ind(V_j, V_k)$, 则一定存在一条不同于 $(V_i, V_j), (V_i, V_k)$ 的路径使 V_j, V_k 相互影响;

(2) 若 $\neg V(V_j, V_k | V_i)$, 则在 V_i 已知的情况下, V_j, V_k 不能通过 V_i 相互影响, 又已知 $\neg Ind(V_j, V_k | V_i)$, 则一定存在一条不同于 $(V_i, V_j), (V_i, V_k)$ 的路径使 V_j, V_k 相互影响.

因此 V_j, V_k 通过除 $V_j - V_i - V_k$ 外的其它路径相互影响, 此时仅由 CI_0 和 CI_1 无法确定是否 V 结构. 例如图 5(a)、(b) 中均满足 $\neg Ind(V_j, V_k), \neg Ind(V_j, V_k | V_i)$ 且 $PV(V_j, V_k | V_i)$.

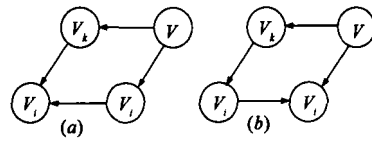


图 5 CI_0 和 CI_1 无法判别的局部结构

3.2 搜索局部结构

经过 CI 测试阶段若仍存在无向弧, 则需进一步处理. 设无向弧 (V_i, V_j) , 若其不在 UNV 中的任意三元组内出现, 则说明其参与的所有 PV 都在 CI 测试阶段被判别, 因此该无向弧无需进一步判断. 否则该无向弧将通过打分搜索判断其方向. 打分搜索方法运用评分函数对网络结构进行评价, 在搜索空间内找到一个得分最高的结构. 评分准则主要有以下几类: 贝叶斯统计方法 BD/BDe^[9,15]、基于最小描述长度 (Minimum Description Length, MDL) 方法^[3] 和等价的贝叶斯信息标准 (Bayesian Information Criterion, BIC) 方法^[16]. 本文采用 BIC 作为评分标准:

$$Score(DAG) = \sum_i Score(X_i, \pi_i) = \sum_i \left(\sum_{l=1}^M \log_2 P(X_i | \pi_i) - \frac{1}{2} \|\pi_i\| (\|\pi_i\| - 1) \log_2 M \right)$$

$\|X\|$ 是 X 的不同状态的个数, M 是样本集合的大小, X_i^k 是 X_i 的第 k 个状态.

在搜索过程中, 搜索可分解为局部子图的搜索: CI 测试阶段得到 CG, 生成 CG 的链分量闭包. 每个包含无向弧的 CC, 其闭包对应的所有可能 LCG 即搜索空间. 搜索空间上的状态转移算子是 UNV 中的可能 V 结构, 转移过程即 V 结构的生成和消除, 通过评分函数找到得分最高的 LCG, 完成弧定向. 搜索空间的上界为 $O(2^k)$, k 为无向弧的个数. 若经搜索得到的 CG 不是 LCG, 可通过文献^[11]中的转换算法生成与之等价的 LCG.

4 实验

本实验利用多组训练数据对算法进行测试. 首先以 Alarm 网为例, 说明弧定向算法的执行过程. Alarm 网是关于病人监护的贝叶斯网, 包含 37 个节点和 46 条边, 其网络结构为图 6 所示^[17].

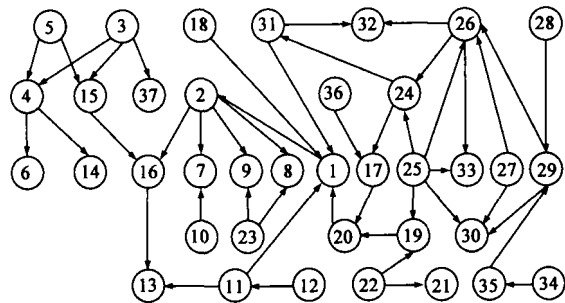


图 6 Alarm 网的网络结构

CI测试阶段 Cl_0 执行 62 次, Cl_1 执行 22 次. CIP 算法执行结束后, 32 条 (70%) 无向弧正确定向. 剩余的 14 条无向弧中 10 条需要进入局部搜索, 4 条由于对应的 UNV 集合为空无需处理. CIP 算法输出的 CG 如图 7 所示.

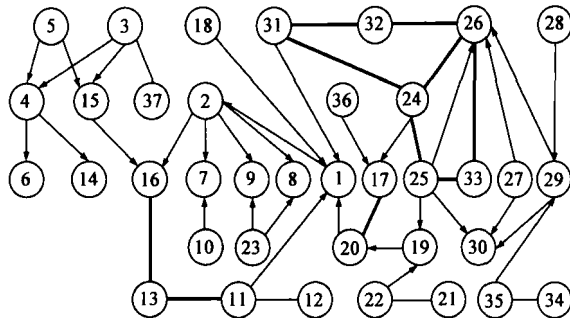


图7 CIP运行之后得到的CG

该 CG 包含 3 个 CC, 每个 CC 的状态转移算子通过 UNV 集合获得. 在应用算子之前, 算子集合可以进行约简. 形如 $V_j - V_i < -V_k$ 和 $V_j - V_i < -V_m (k \neq m)$ 的算子只需保留一个即可. 图 8 列出了三个 CC 闭包和相应的状态转移算子.

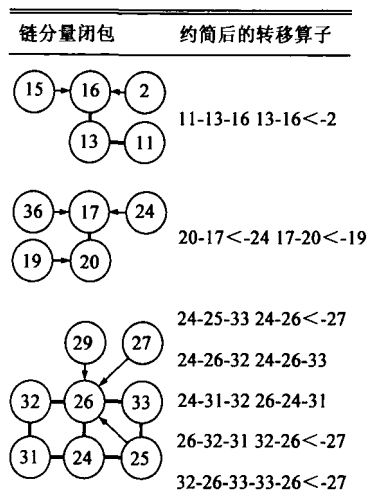


图8 链分量闭包及状态转移算子

局部结构	算子状态	得分
	$V(13,2 16)$ $\neg V(11,16 13)$	$-2.14 \cdot 10^4$
	$\neg V(13,2 16)$ $\neg V(11,16 13)$	$-1.98 \cdot 10^4$
	$\neg V(13,2 16)$ $V(11,16 13)$	$-1.83 \cdot 10^4$

图9 对CC闭包局部搜索过程

每一个局部结构通过对算子选择不同状态获得 (所选状态之间不能存在矛盾), 运用评分函数选择得分最高的局部结构作为弧定向结果. 图 9 列出了第一个 CC 闭包的搜索过程: 通过打分搜索, 将 CIP 输出的无向弧 (V_{13}, V_{16}) 和 (V_{13}, V_{11}) 定向为 $\langle V_{16}, V_{13} \rangle$ 和 $\langle V_{11}, V_{13} \rangle$.

搜索结束后, 10 条无向弧均被正确定向. 本实验得到的 CG 即为 LCG, 若非 LCG 可将其转换为 LCG. 图 10 显示了通过本算法定向后得到的 LCG.

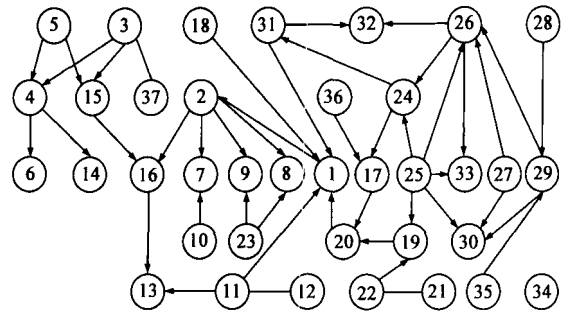


图10 弧定向算法得到的最大连通图

表 1 列出算法在几组训练数据下的运行状况 (关于训练数据的详细说明可参阅 <http://www.cs.huji.ac.il/labs/compbio/Repository/>). 由于训练数据中存在噪声或由于 CI 测试的参数选择问题, 在 CIP 过程中会出现同一条弧在不通的 PV 中被定向为相反方向. 此时将此弧标记为无向弧, 在搜索阶段再对其进行定向.

表 1 算法运行数据

训练数据	节点数	待定边数	Cl_0 次数	Cl_1 次数	CIP 定向边数/比例	搜索定向边数/比例
Alarm	37	46	62	22	37/80%	10/20%
Barley	48	84	126	29	77/92%	7/8%
MILDEW	35	46	64	19	40/87%	6/13%
Insurance	27	52	92	25	42/81%	10/19%

5 结论

本文提出的弧定向算法结合独立性测试和打分搜索两种方法, 利用多项式级的 Cl_0 和 Cl_1 测试判断是否存在 V 结构, 对部分弧进行快速定向; 对无法通过低阶 CI 测试判断的 V 结构, 在链分量闭包上打分搜索, 有效的缩小了搜索空间, 提高了弧定向的执行效率. 通过命题证明和测试实验说明, 本算法输出的 LCG 对统计可分的弧正确定向, 对统计不可分的弧保留其无向特性, 从而更准确的体现了数据中的概率因果关系.

参考文献:

[1] Jie Cheng, Russell Greiner, Jonathan Kelly, David Bell, Weiru Liu. Learning Bayesian networks from data: An information-theory based approach [J]. Artificial Intelligence, 2002, 137 (1

- 2); 43 - 90.
- [2] 李小琳, 何湘东, 苑森森. 基于信息论和遗传算法的 Bayesian 网络弧定向方法研究[J]. 复旦学报(自然科学版), 2004, 43(5): 717 - 720.
- [3] Bacchus F, Lam W. Learning Bayesian Belief Networks: an approach based on the MDL principle[J]. Computational Intelligence, 1994, 3(10): 269 - 293.
- [4] Jude Pearl. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference[M]. San Mateo, California: Morgan Kaufmann, 1988.
- [5] T Verma, J Pearl. An algorithm for deciding if a set of observed independencies has a causal explanation[A]. Proceedings of the Eighth Conference on Uncertainty in Artificial Intelligence[C]. Morgan Kaufman, 1992. 323 - 330.
- [6] Thomas Verma, Judea Pearl. Equivalence and synthesis of causal models[A]. Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence[C]. New York, NY, USA Elsevier Science Inc, 1990. 255 - 270.
- [7] Akaike H. A new look at the statistical model identification[J]. IEEE Transactions on Automatic Control, 1974, 19(6): 716 - 723.
- [8] Schwarz Gideon. Estimating the dimension of a model[J]. The Annals of Statistics, 1978, 6(2): 461 - 464.
- [9] Heckerman David, Geiger Dan, M Chickering David. Learning Bayesian networks: The combination of knowledge and statistical data[J]. Machine Learning, 1995, 20(3): 197 - 243.
- [10] D M Chickering. A transformational characterization of Bayesian network structures[A]. Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence[C]. Morgan Kaufmann, 1995. 87 - 98.
- [11] Liu Baijun, Zheng Zhongguo, Zhao Hui. An efficient algorithm for finding the largest chain graph according to a given chain graph[J]. Science in China, Ser. A, 2005, 48(11): 1517 - 1530.
- [12] Steen A Andersson, David Madigan, Michael D Perlman. A characterization of Markov equivalence classes for acyclic digraphs[J]. Annals of Statistics, 1997, 25(2): 505 - 541.
- [13] Martin Volf, Milan Studeny. A graphical characterization of the largest chain graphs[J]. International Journal of Approximate Reasoning, 1999, 20(3): 209 - 236.
- [14] David Maxwell Chickering, Christopher Meek. On the incompatibility of faithfulness and monotone DAG faithfulness[J]. Artificial Intelligence, 2006, 170(8 - 9): 653 - 666.
- [15] Herskovits Cooper. A Bayesian method for the induction of probabilistic networks from data[J]. Machine Learning, 1992, 9(4): 309 - 347.
- [16] Schwarz. Estimating the dimension of a model[J]. Annals of Statistics, 1978, 6(2): 461 - 464.
- [17] Beinlich IA, Suermondt HJ, Chavez RM, Cooper GF. The ALARM monitoring system: a case study with two probabilistic inference techniques for belief networks[A]. Proceedings of the 2nd European Conference on AI and Medicine[C]. Springer-Verlag, 1989, 247 - 256.

作者简介:

贾海洋 男, 1977 年生于吉林长春. 吉林大学计算机科学与技术学院讲师、博士. 研究方向: 知识工程、概率图模型、机器学习、E-learning. E-mail: jiahy@jlu.edu.cn

陈娟 女, 1980 年生于吉林长春. 吉林大学计算机科学与技术学院讲师、博士. 研究方向: 知识工程、空间推理、软件工程.