软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2012,23(6):1531−1541 [doi: 10.3724/SP.J.1001.2012.04090]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

CP-nets 的完备性及一致性研究*

刘惊雷[1,2], 廖士中[1+], 张 伟[2]

[1](天津大学 计算机科学与技术学院,天津 300072)
[2](烟台大学 计算机科学与技术学院,山东 烟台 264005)

On the Completeness and Consistency for CP-nets

LIU Jing-Lei[1,2], LIAO Shi-Zhong[1+], ZHANG Wei[2]

[1](School of Computer Science and Technology, Tianjin University, Tianjin 300072, China)
[2](School of Computer Science and Technology, Yantai University, Yantai 264005, China)
+ Corresponding author: E-mail: szliao@tju.edu.cn

Liu JL, Liao SZ, Zhang W. On the completeness and consistency for CP-nets. *Journal of Software*, 2012,23(6):1531−1541. http://www.jos.org.cn/1000-9825/4090.htm

Abstract: CP-nets (conditional preference networks) is a simple and intuitive graphical tool for representing conditional preference statements over the values of a set of variables. It has been a studying hotspot in artificial intelligence recently. The algorithm of strong dominance with respect to any binary-valued CP-nets has not been given; preferences completeness of CP-nets have not been studied by anyone, This paper makes a study of completeness and consistency of CP-nets by designing a strong dominance algorithm. First, by constructing induced graph of CP-nets and studying its properties, the study solves the problem of strong dominance with respect to any binary-valued CP-nets by Warshall algorithm to get the transitive closure of flip relation. Second, by solving the preference number of three kinds of CP-nets (separeble-structured, chain-structured, tree-structured), the study gives preferences incompleteness theorem and counting number formula of separeble condition preference networks. Finally, the study deals with consistency problem, and consistency judgment theorem and algorithm are given. The method not only solves some difficult problems proposed by Boutilier and Goldsmith, but also deepens the basic theory researching of CP-nets.
Key words: strong dominance; preference completeness; preference consistency; transitivity closure of flip relation; separeble condition preference network; judgment theorem and algorithm

摘 要: CP-nets 是一种简单而又直观的图形化偏好表示工具,成为近几年人工智能的一个研究热点.然而,任意二值 CP-nets 上的强占优算法还没有给出,CP-nets 可表示的偏好的完备性还无人研究,CP-nets 所能表示的偏好是否一致也还未彻底解决.基于 CP-nets 上的强占优运算研究 CP-nets 的完备性和一致性.首先,通过构造 CP-nets 导出图及其性质的研究,得出强占优的本质是求取翻转关系的传递闭包,从而利用 Warshall 算法求出可判断任意 CP-nets 的强占优;其次,通过求取 3 种不同结构(可分离的、链表结构和树形结构)的 CP-nets 的偏好个数,给出了 CP-nets 可表达的偏好的不完备性定理,并给出了可分离的 CP-nets 中偏好的计数公式;最后,研究 CP-nets 的一致

* 基金项目: 国家自然科学基金(61170019); 天津市自然科学基金(11JCYBJC00700)
收稿时间: 2010-07-27; 定稿时间: 2011-07-04

性,给出了 CP-nets 的一致性判定定理及其算法.所做工作不仅解决了 Boutilier 和 Goldsmith 提出的一些难题,还深化了 CP-nets 的基础理论研究.

关键词: 强占优;偏好的完备性;偏好的一致性;翻转关系的传递闭包;可分离的条件偏好网;判定定理及算法

中图法分类号: TP181 **文献标识码**: A

偏好是 Agent 系统进行决策的一个核心概念,它能帮助人们获取 Agent 的目标.然而,即使对于单个 Agent 来说,其所提出的偏好也未必能够满足.例如,Agent *a* 想买一个便宜的、但速度不高油耗更低的法拉利轿车,其实这样的轿车是不存在的[1].由于偏好引导人们的选择,因此,理解偏好的表示方法和处理技术,就对构建代表用户意愿的决策系统起着重要的作用[2].目前,偏好不仅在社会选择(social choice)、决策论(decision theory)和对策论(game theory)中有应用,在人工智能中的协同过滤(collaborative filtering)、推荐系统(recommender systems)和产品配置(product configuration)也出现了大量的应用[3].

偏好是人工智能中一个重要的问题,《Artificial Intelligence》已于 2011 年 5 月出版一本专辑“Representing,Processing,and Learning Preferences:Theoretical and Practical Challenges”.结合文献[4],可将偏好研究归为如下 3 类:

(1) 偏好的表示(preference representation):研究如何将 Agent 的偏好用图形方式[1−3,5−7]或逻辑方式[8−10]加以描述,并重点探讨该表示的数学性质.哲学大师 Benthem[8]利用模态逻辑设计了反映 Ceteris Paribus(all else being equal)语义的逻辑系统,Lang[9]给出了偏好的逻辑和组合投票之间的关系,张志政[10]提出并构造了一个能够描述和推理多种类型偏好的逻辑系统;

(2) 偏好的提取(preference elicitation):研究如何通过与 Agent 交互来学习和推理 Agent 的偏好.该方向是以机器学习的方式来获取 Agent 的偏好.其中:Lang[11]研究最简单的一类属性可分离的 CP-nets 的学习问题;Koriche[12]提出了利用等价查询和成员查询来获取 Agent 的条件偏好;Chevaleyre[13]给出了多属性域偏好学习的通用准则,给出了哪些实例是可学习的,哪些实例是不能学习的.这些偏好学习方法[11−13]的主要缺陷是,在没有探明相应偏好表示模型(如 CP-nets)基本数学性质的前提下进行偏好学习,特别是没有得出 CP-nets 表达能力和学习能力关系时,对 CP-nets 学习就会出现困难;

(3) 偏好的聚合(preference aggregation):研究如何将多个 Agent 的偏好聚合成集体偏好,从而实现集体决策.Conitzer[14]给出了偏好聚合的一些应用场景,例如慈善捐款(charitablegiving)[15]、组合拍卖(combinatorial auctions)[16]、投票和排序聚合(voting and rank aggregation)[17,18]等.其中:文献[15]给出了给慈善机构捐款的一些算法问题;文献[16]的最优联盟结构值的本质就是组合拍卖的优化问题;Conitzer[17]给出了在具有少量被选举人的情况下,什么时候操纵选举是困难的问题,其目的是避免选举中的舞弊行为.Pini[18]给出了在不完全偏好的情况下,如何确定投票系统中的必然赢家和可能赢家的算法问题.

本文研究偏好的图形表示——CP-nets[1−3,5−7]的相关问题.因为有关 CP-nets 的研究中存在如下一些基础问题:一是任意 CP-nets 结构上的强占优算法还没有给出[3,5],且其复杂度的求取要么没有[3],要么复杂[5];二是 CP-nets 完备性,特别是一些特殊结构[11]的 CP-net 可表达的偏好是否完备还无人研究;三是 CP-nets 的一致性问题的研究[3,5]才刚刚开始.本文的特色在于:

(1) 基于 CP-nets 导出图性质的研究,指出 CP-nets 强占优关系的判断,本质上是求取 CP-nets 导出图翻转关系的传递闭包.从而基于 Warshall 算法设计了一个任意二值 CP-nets 上的强占优算法,以初等方法解决了文献[3]没有解决或没有彻底解决[5]的一个难题;

(2) 基于所设计的强占优算法,研究 CP-nets 表达偏好的完备性问题,不仅给出 CP-nets 表示偏好的不完备定理,还给出了 SCP(separeble conditional preference)[11]结构可表达的偏好个数的计数公式和完备程度;同时还研究了 CP-nets 的一致性问题,给出了任意结构 CP-nets 的一致性判定定理和算法;

(3) 将 CP-nets 的完备性和一致性判断归约到强占优算法的框架下求解,文献[3,5]则割裂算法和导出图性质之间的联系,因此其求解结果只能是部分结果,而且有的性质如完备性还无法求解.

1 CP-nets 的基本概念

1.1 偏好的相关定义

定义 1. 设 $V=\{X_1,X_2,\ldots,X_n\}$是决策属性集,$Dom(X_i)$代表 X_i 的有限定义域,则决策空间$\Omega=\times_{i=1}^{n}Dom(X_i)$表示所有属性组合.若配置 $o,o'\in\Omega$仅有一个属性值不同,而其他属性值都相同,则称 o 和 o'为可交换的配置(swap outcome).若 Agent 对属性 X_i 的偏好取决于 X_j,则称 X_j 是 X_i 的一个父亲,用 $Pare(X_i)$表示 X_i 的父亲集.

定义 2. $\geqslant$是决策空间Ω上的二元关系,

(1) 若$\geqslant$自反、反对称和传递,即$\geqslant$是偏序关系[19]时,称$\geqslant$为Ω上的偏好关系.

(2) 若存在配置 $o,o'\in\Omega,o\ngeqslant o'$且 $o'\ngeqslant o$,则称 o 和 o'不可比(incomparability).当Ω中存在不可比的两个配置时,称偏好关系$\geqslant$不完备.

(3) 若配置 $o\geqslant o'$而 $o'\ngeqslant o$,称 o 和 o'具有严格偏好关系,记作 $o>o'$.

1.2 条件偏好图——CP-nets

定义 3. 设 $CPT(X_i)$为属性 X_i 的条件偏好表(conditional preference table),它表示 X_i 在其父属性 $Pare(X_i)$的不同取值下,Agent 对 $Dom(X_i)$集合的一个排序.在 $Pare(X_i)$的所有取值下,Agent 对 X_i 取值的排序构成 $CPT(X_i)$.

定义 4. CP-nets 是一个有向图 $N=\langle V,CE\rangle$,其中,V 是顶点集;CE 为有向边集,代表所有属性之间的依赖关系,每个顶点 X_i 都有 $CPT(X_i)$与其关联.

例 1:"晚会穿衣"实例[3].某人参加晚会的穿衣主要考虑夹克、裤子和衬衫,分别用 J,P 和 S 来代表.对于夹克和裤子来说,他喜欢黑色胜于白色.而衬衫颜色则取决于夹克和裤子颜色的搭配.当夹克和裤子是同一种颜色,则他偏好红色衬衫而不是白色衬衫;当夹克和裤子颜色不同时,则偏好白色衬衫而不是红色衬衫.

该例子的 CP-nets 图 $N=\langle V,CE\rangle$如图 1 所示.其中,$V=\{J,P,S\}$,$Dom(J)=\{J_b,J_w\}$,$Dom(P)=\{P_b,P_w\}$,$Dom(S)=\{S_r,S_w\}$;$CE=\{\langle J,S\rangle,\langle P,S\rangle\}$.

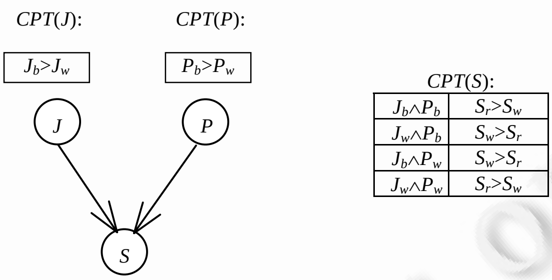

Fig.1 CP-nets for "evening dress"

图 1 "晚会穿衣"的 CP-nets

定理 1. 设 $N=\langle V,CE\rangle$是 CP-nets,若配置 o_1 与 o_2 是可交换的配置,则在 $O(n)$内可确定出 $o_1>o_2$ 或者 $o_2>o_1$.

证明:可交换配置 o_1 与 o_2 只有一个属性值不同,因此可在 $O(n)$内确定出哪个属性值不一样;当找到 o_1 与 o_2 不同取值的 X_i 后,则容易求出 $Pare(X_i)$,显然,从 X_i 的条件偏好表中就可确定 o_1 与 o_2 的偏好强弱.

2 CP-nets 的强占优

CP-nets 可对 Agent 的偏好进行描述,基于强占优可实现偏好的强弱判断.

2.1 CP-nets的导出图

定义 5. 设 $N=\langle V,CE\rangle$是一个 CP-nets,则有向图 $N'=\langle\Omega,IE\rangle$是 N 的导出图,其中,IE 是可交换配置所构成的

有向边集,且对有向边终点的偏好强于对起点的偏好.

根据图论知识[19]可得,IE 是一个关系集,称 IE 为翻转关系(flip relation),因为对于二值 CP-nets,若 $o'\ IE\ o$,则将 o'的一个属性值翻转(取两个中的另一个),则得到配置 o.

定理 2. $N'=\langle \Omega,IE\rangle$是 CP-nets 图 $N=\langle V,CE\rangle$的导出图,且$|V|=n,|Dom(X_i)|=2$,则:

(1) $|\Omega|=2^n,|IE|=n\times 2^{n-1}$;

(2) 2^n 个配置之间偏好关系个数为 $2^{n-1}(2^n-1)$.

证明:

(1) 因为$|Dom(X_i)|=2$,而$\Omega=\times_{i=1}^{n} Dom(X_i)=Dom(X_1)\times Dom(X_2)\times\ldots\times Dom(X_n)$,因此$|\Omega|=2^n$.

由于 2^n 个配置中的每个配置均有 n 个可交换配置,因此总共可交换的配置是 $n\times 2^n$ 个;而两个可交换的配置有一条有向边相连,故总共边数是 $n\times 2^n/2=n\times 2^{n-1}$.

(2) 2^n 个配置可构成的偏好关系个数为 2^n 个顶点所构成的无向完全图的边的数目,根据文献[19]知,m 个顶点的无向完全图具有 $m(m-1)/2$ 条边,因此,2^n 个配置之间的偏好关系个数为 $2^{n-1}(2^n-1)$. □

定义 6. 设 $N'=\langle \Omega,IE\rangle$是 CP-nets 的导出图,对于 $o,o'\in\Omega$,若存在着一条路径连接顶点 o'和 o,则称 o 强占优 o',记作 $o>o'$.

$o>o'$说明配置 o 比配置 o'优,即在 CP-nets 导出图上存在一个翻转序列,使得该序列连接顶点 o'和 o[3].为了判断 $o>o'$是否成立,文献[3]使用了 Suffix Fixing 技术、Forward Pruning 等技术[3],但其只能求出表 1 中几种结构的 CP-nets.

Table 1 Strong dominance complexity for binary-valued CP-nets

表 1 二值 CP-nets 的强占优复杂度

CP-nets structure	Complexity	Remark
Directed tree	$O(n^2)$	Lower bound
Single connected (polytree)	$O(2^{2k}n^{2k+3})$	k-maximal indegree
Single connected DAG	NP-Complete	Reduction from 3SAT
δ-connected DAG	NP-Complete	Minimal flipping sequences are polynomially bounded
General DAG	?	**Harder than NP**?

为了给出一般结构的强占优算法,下面从强占优的深层含义出发,即考虑 CP-nets 导出图中翻转关系的传递闭包,来进行 CP-nets 的强占优判断.

2.2 二值CP-nets的强占优算法

定义 7. 设 R 是集合 A 上二元关系,如果 R'满足:(1) R'是传递;(2) $R\subseteq R'$;(3) A 上任何包含 R 的传递关系 R''都有 $R'\subseteq R''$,则称 R'为 R 的传递闭包[19],记为 $t(R)$.

定理 3. 设 R 是集合 A 上的二元关系,则 $t(R)$为

$$t(R)=R\cup R^2\cup R^3\cup\ldots\cup R^n\cup\ldots^{[19]};当|A|=m,t(R)=\bigcup_{i=1}^{m}R^i.$$

CP-nets 导出图中的边集是翻转关系的集合,对该翻转关系求传递闭包后,会得到所有的强占优关系.我们可对 IE 的关系矩阵 A 进行运算,其求解方法是定义传递闭包矩阵 $B=A\vee A^2\vee\ldots\vee A^n$,其中,$\vee$为逻辑运算符,矩阵中的元素 1 和 0 看成命题的真和假.

Warshall 算法可求关系的传递闭包[19],当关系矩阵 A 是 $m\times m$ 阶时,算法复杂度为 $O(m^3)$[19].由于 IE 的传递闭包就是 CP-nets 上的强占优关系,因此如下所示的 Warshall 算法可认为是求 CP-nets 的传递闭包,进而可判断两个配置是否具有强占优关系的算法.

Warshall Algorithm. Strong *dominance*(CP-nets *V*,Outcome o_1,o_2)

Input:CP-nets 图 N 及任意两个配置 o_1 与 o_2;

Output:o_1 与 o_2 是否具有强占优关系.

结合定理 1 和定理 2 求取导出图 N'中翻转关系 IE 的关系矩阵 $A[m][m]$

```
Int B[m][m]    //传递闭包的关系矩阵,其中,m=2^n 为二值网的配置数
For i=1 to m Do
  For j=1 to m Do
    B[i][j]=A[i][j]    //将关系矩阵 A 赋值给 B
    End For
End For
For k=1 to m Do
  For i=1 to m Do
    For j=1 to m Do
      If B[i][j]≠ 0 Then B[i][j]=A[i][k]∧A[k][j]
    End For
  End For
End For
求配置 o1 与 o2 在关系矩阵 A 的位置于变量 p 和 q 中
If (B[p][q]=0 and B[q][p]=0) Then Printf (“o1 与 o2 无强占优关系,它们不可比较”)
If (B[p][q]=1 and B[q][p]=0) Then Printf (“o2>o1”)
If (B[q][p]=1 and B[p][q]=0) Then Printf (“o1>o2”)
```

例 2:利用 Warshall 算法判断图 2 中配置“$J_w\ P_b\ S_r$”与“$J_w\ P_b\ S_w$”的关系.

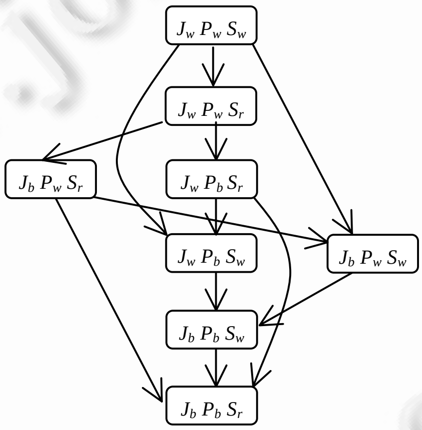

Fig.2 Induced graph of CP-nets
图 2 CP-nets 的导出图

解:图 2 中的配置集$\Omega=\{o_1,o_2,o_3,o_4,o_5,o_6,o_7,o_8\}$,$o_1$=“$J_w\ P_w\ S_w$”,$o_2$=“$J_w\ P_w\ S_r$”,$o_3$=“$J_w\ P_b\ S_r$”,$o_4$=“$J_w\ P_b\ S_w$”,$o_5$=“$J_b\ P_b\ S_w$”,$o_6$=“$J_b\ P_b\ S_r$”,$o_7$=“$J_b\ P_w\ S_r$”,$o_8$=“$J_b\ P_w\ S_w$”;翻转关系

$$IE=\{\langle o_1,o_2\rangle,\langle o_2,o_3\rangle,\langle o_3,o_4\rangle,\langle o_4,o_5\rangle,\langle o_5,o_6\rangle,\langle o_1,o_4\rangle,\langle o_1,o_8\rangle,\langle o_2,o_7\rangle,\langle o_3,o_6\rangle,\langle o_7,o_6\rangle,\langle o_7,o_8\rangle,\langle o_8,o_5\rangle\}.$$

$$IE \text{ 的关系矩阵 } A=\begin{bmatrix} 0&1&0&1&0&0&0&1\\ 0&0&1&0&0&0&1&0\\ 0&0&0&1&0&1&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&1&0&1\\ 0&0&0&0&1&0&0&0 \end{bmatrix}, \text{执行 Warshall 算法后}, B=\begin{bmatrix} 0&1&1&1&1&1&1&1\\ 0&0&1&1&1&1&1&1\\ 0&0&0&1&1&1&0&0\\ 0&0&0&0&1&1&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&1&1&0&1\\ 0&0&0&0&1&0&0&0 \end{bmatrix}.$$

由于 $B[3][5]=1, B[5][3]=0$,故 $o_5>o_3$,而 $B[3][7]=B[7][3]=0$,故 o_3 和 o_7 不可比.

定理 4. Warshall 算法对二值 CP-nets 图 N 进行强占优是充要的,且时间复杂度为 $O(8^n)$.

3 CP-nets 的完备性

CP-nets 是表示偏好的语言,一个自然的问题就是,它能否将所有偏好都表达出来,即 CP-nets 表示的偏好是否完备.如果一个 CP-nets 不完备,则对文献[11–13]中的 CP-nets 学习产生影响,然而文献[1–7,11–13]对偏好的完备性都未作讨论.

定义 8. 若二值 CP-nets 图 N 可表达的偏好关系个数为 $2^{n-1}(2^n-1)$,则称 N 是完备的.

从定理 2 知,二值无环 CP-nets 图 N 表达的 2^n 个配置之间的偏好关系的个数理论上应该为 $2^{n-1}(2^n-1)$个,显然,将 CP-nets 的完备性定义成"一个偏好表示语言实际所能表达的偏好关系个数等于理论上可能具有的偏好关系个数"是合理的.

3.1 可分离的CP-nets的完备性

CP-nets 可表达的偏好个数与其结构紧密相关,由于可分离的(各个属性没有依赖关系)CP-nets 最简单(只有顶点而没有边的零图),因此有关 CP-nets 的研究都从它开始.例如文献[11]就是如何学习可分离的 CP-nets.同样分析 CP-nets 的完备性,我们也从 SCP 结构(如图 3 左边)的 CP-nets 开始.

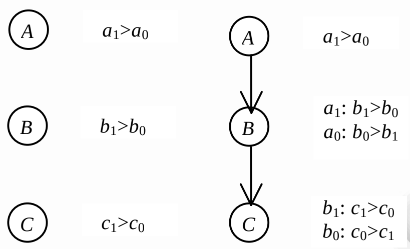

Fig.3 SCP and chain-structured CP-nets

图 3 可分离结构和链式结构的 CP-nets

定理 5. SCP 结构的二值 CP-nets 可表达的偏好个数是 3^n-2^n.

证明:对于 n 个顶点的二值 CP-nets 图共有 2^n 个配置,其可以表示成从$\langle 0,0,0,0,0,\dots,0,0\rangle$到$\langle 1,1,1,1,1,\dots,1,1\rangle$这 2^n 个二进制序列.第 i 位为 1,代表属性 X_i 的取值为 1(代表两个取值中的一个);第 i 位为 0,代表属性 X_i 的取值为 0(代表两个取值中的另一个).并假设可分离的 CP-nets 的条件偏好表为"1>0",其含义为对属性值为 1 的偏好大于对属性值为 0 的偏好.

(1) 当配置 $o=\langle 0,0,0,0,0,\dots,0,0\rangle$时,比 o 优的配置 $o'(o'>o)$的个数为 $2^n-1=C_n^0\times(2^n-1)$.

(2) 当配置 o 中含有 1 个 1 时,该 1 的位置共有 C_n^1 种,而每种情况下,比 o 优的配置 o'的个数为 $2^{n-1}-1$($n-1$ 位二进制所能表示的最大数).根据乘法原理,比含有 1 个 1 的配置优的配置共有 $C_n^1\times(2^{n-1}-1)$.

(3) 当配置 o 中含有 k 个 1 时,该 1 的位置共有 C_n^k 种,而每种情况下,比 o 优的配置 o'的个数为 $2^{n-k}-1$.根据乘法原理,比含有 k 个 1 的配置优的配置共有 $C_n^k \times (2^{n-k}-1)$.

(4) 当配置 o 中含有 n 个 1 时,该 1 的位置共有 C_n^n 种,而每种情况下,比 o 优的配置 o'的个数为 $0=2^0-1$.根据乘法原理,比含有 n 个 1 的配置优的配置共有 $C_n^n \times (2^0-1)$.

根据加法原理,SCP 的条件偏好网共有的强占优关系个数为

$$C_n^0 \times (2^n-1) + C_n^1 \times (2^{n-1}-1) + C_n^2 \times (2^{n-2}-1) + \ldots + C_n^k \times (2^{n-k}-1) + \ldots + C_n^n \times (2^0-1) =$$

$$\sum_{k=0}^{n} C_n^k \times (2^{n-k}-1) = \left(\sum_{k=0}^{n} C_n^k \times 2^{n-k}\right) - \sum_{k=0}^{n} C_n^k = \left(\sum_{k=0}^{n} C_n^k \times 1^k \times 2^{n-k}\right) - \sum_{k=0}^{n} C_n^k = (1+2)^n - 2^n = 3^n - 2^n \qquad \square$$

推论 1. 非平凡(n>1)的 SCP 结构的 CP-nets 是不完备的.

证明:二值 CP-nets 图 N 应表达 $2^{n-1}(2^n-1)$个偏好,而由定理 5 知,SCP 结构的 CP-nets 只能表达 3^n-2^n 个偏好,显然,当 $n>1$ 时,$3^n-2^n<2^{n-1}(2^n-1)$,因此推论成立. □

3.2 链结构的CP-nets的完备性

例 2 是一个树形结构的 CP-nets,关系矩阵 B 中元素 1 的个数是 23,代表 CP-nets 可表达的偏好个数为 23.而 3 个顶点的二值 CP-nets 应该表达 28 对偏好,但此时图 2 有 5 对配置间的偏好没有作阐述,此时,CP-nets 对偏好的表示不完备.

基于 Warshall 算法可判断图 3 右边所示的链式结构(chain-structured)CP-nets 的完备性,8 个配置按二进制排列可表示为 o_0=“$a_0\ b_0\ c_0$”,o_1=“$a_0\ b_0\ c_1$”,o_2=“$a_0\ b_1\ c_0$”,o_3=“$a_0\ b_1\ c_1$”,o_4=“$a_1\ b_0\ c_0$”,o_5=“$a_1\ b_0\ c_1$”,o_6=“$a_1\ b_1\ c_0$”,o_7=“$a_1\ b_1\ c_1$”,$IE=\{\langle o_0,o_4\rangle,\langle o_1,o_0\rangle,\langle o_1,o_5\rangle,\langle o_2,o_0\rangle,\langle o_2,o_3\rangle,\langle o_2,o_6\rangle,\langle o_3,o_1\rangle,\langle o_3,o_7\rangle,\langle o_4,o_6\rangle,\langle o_5,o_4\rangle,\langle o_5,o_7\rangle,\langle o_6,o_7\rangle\}$,则

$$A=\begin{bmatrix} 0&0&0&0&1&0&0&0\\ 1&0&0&0&0&1&0&0\\ 1&0&0&1&0&0&1&0\\ 0&1&0&0&0&0&0&1\\ 0&0&0&0&0&0&1&0\\ 0&0&0&0&1&0&0&1\\ 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&0&0 \end{bmatrix}$$,执行 Warshall 算法后,$$B=\begin{bmatrix} 0&0&0&0&1&0&1&1\\ 1&0&0&0&1&1&1&1\\ 1&1&0&1&1&1&1&1\\ 1&1&0&0&1&1&1&1\\ 0&0&0&0&0&0&1&1\\ 0&0&0&0&1&0&1&1\\ 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&0&0 \end{bmatrix}$$,B 中有 27 个 1,即 3 个顶点的链式结构的二值 CP-net 可表达 27 对偏好,此时有 1 对偏好没有表达.

3.3 CP-nets的完备性

定理 6(CP-nets 表示可交换配置的完备性). 利用 CP-nets 图 N 可表达二值 CP-nets 中所有的可交换配置的强占优关系.

证明:由定理 2(1)可知,二值 CP-nets 图有 $n*2^{n-1}$ 对可交换的配置.设 o_1 与 o_2 是任意一对可交换的配置,根据定理 1 可知,o_1 与 o_2 的强占优关系可判定,因此 $n*2^{n-1}$ 对可交换配置在顶点的条件偏好表中作了表达.□

定理 7(偏好表示不完备定理). CP-nets 图 N 对决策空间中的任意两个配置间的强占优关系不是都完备的,即存在两个配置间的强占优关系在 CP-nets 中无法表达.

证明:显然,当 n>1 时,SCP 结构、链式结构和树形结构的 CP-nets 都不完备.故 CP-nets 表示偏好不都完备. □

定义 9. 对于二值 CP-nets 图 N 来说,N 所能表达的偏好的个数是 $NumOfCom$,则 N 的完备程度:

$$\eta=NumOfCom/2^{n-1}(2^n-1),$$

其含义为 N 表达的偏好个数占应该表达的偏好个数的百分比.

基于定义 9 和强占优算法,很容易写出求 CP-nets 完备程度的算法,因为只需统计出 B 中 1 的个数后,除

以 $2^{n-1}(2^n-1)$就能得出完备程度.

4 CP-nets 的一致性

除了考虑 CP-nets 的完备性,还要考虑 CP-nets 的一致性,因为一致性是计算机系统的一个优良特性,倘若 CP-nets 不一致,则说明其所表达的偏好出现自相矛盾,而这恰是系统所忌讳的.

4.1 一致性及判定定理

定义 10. 设 N 为任意一个 CP-nets,其所能表达的偏好关系为$>_N$($>_N$ 为 IE 的传递闭包),$o_1,o_2,\ldots,o_m$ 为决策空间Ω中的任意 m 个配置,若不存在配置序列$\langle o_1,o_2,\ldots,o_m\rangle$使得 $o_1>_N o_2,o_2>_N o_3,\ldots o_{i-1}>_N o_i,\ldots,o_m>_N o_1$,即$\langle o_1,o_2,\ldots,o_m\rangle$构不成环形序列的话,则称该 CP-nets 一致(consistency).

定理 8(一致性判定定理). 设 CP-nets 图 N 所能表达的偏好关系为$>_N$,则如下说法等价:

(1) N 是一致的;

(2) 任意一个配置 $o,o>o$ 都不成立,即$\forall o\in\Omega$有$\neg(o>_N o)$;

(3) $>_N$ 是反自反的;

(4) $>_N$ 关系矩阵中对角线上的所有元素都是 0.

证明:由定义 10 可知 N 一致,则不存在配置序列$\langle o_1,o_2,\ldots,o_m\rangle$,使得 $o_1>_N o_2,o_2>_N o_3,\ldots,o_i>_N o_{i-1},\ldots,o_m>_N o_1$.由于$>_N$ 具有传递性(定义 2),则不存在配置序列$\langle o_1,o_2,\ldots,o_m\rangle$使得 $o_1>o_1$.可见,$\forall o\in\Omega$有$\neg(o>o)$.

根据离散数学关系的反自反特性[19]可知,定理 8(2)、定理 **8**(3)和定理 **8**(4)等价. □

CP-nets 的一致性与 CP-nets 的结构紧密相连,下面从有环图和无环图角度来探讨 CP-nets 的一致性.

定理 9. 任何一个无环 CP-nets 都满足一致性**,但有环 CP-nets 不一定满足一致性.

证明:

(1) 用数学归纳法来证明无环 CP-nets 满足一致性.

假设 N 是一个无环 CP-nets,由于无环图可以进行拓扑排序,因此对 N 中顶点拓扑排序后得到序列$\langle X_{\pi(1)},X_{\pi(2)},X_{\pi(3)},\ldots,X_{\pi(n-1)},X_{\pi(n)}\rangle$,其中,$\pi(1),\pi(2),\pi(3),\ldots,\pi(n-1),\pi(n)$是 $1,2,3,\ldots,n-1,n$ 的一种全排列.下面归纳证明 CP-nets 图中包含一个顶点 $X_{\pi(n)}$时一致,包含两个顶点 $X_{\pi(n-1)},X_{\pi(n)}$一致,包含 $n-1$ 个顶点 $X_{\pi(2)},X_{\pi(3)},\ldots,X_{\pi(n-1)},X_{\pi(n)}$时一致,包含 n 个顶点 $X_{\pi(1)},X_{\pi(2)},X_{\pi(3)},\ldots,X_{\pi(n-1)},X_{\pi(n)}$也一致.

当 CP-nets 图只含一个顶点 $X_{\pi(n)}$时显然一致,因为此时顶点 $X_{\pi(n)}$的条件偏好表 $CPT(X_{\pi(n)})$表示对 $Dom(X_{\pi(n)})$内$|Dom(X_{\pi(n)})|$个取值的一种全排列,而任何一种全排列都一致,即不会出现一个环形序列.

假设当 CP-nets 包含 k 个顶点 $X_{\pi(n-k+1)},X_{\pi(n-k+2)},\ldots,X_{\pi(n-1)},X_{\pi(n)}$构成的图 N_k 一致,则 N_k 中的$|Dom(X_{\pi(n-k+1)})|\times|Dom(X_{\pi(n-k+2)})|\times\ldots\times|Dom(X_{\pi(n)})|$个配置可构成一个全排列(位置靠后的配置不优于位置靠前的配置),现证明增加一个顶点 $X_{\pi(n-k)}$后,所形成的包含 $k+1$ 个顶点 $X_{\pi(n-k)},X_{\pi(n-k+1)},\ldots,X_{\pi(n-1)},X_{\pi(n)}$构成的图 N_{k+1} 仍然一致.

令$\lambda(N_k)=|Dom(X_{\pi(n-k+1)})|\times|Dom(X_{\pi(n-k+2)})|\times\ldots\times|Dom(X_{\pi(n)})|,\lambda(N_{k+1})=|Dom(X_{\pi(n-k)})|\times|Dom(X_{\pi(n-k+1)})|\times\ldots\times|Dom(X_{\pi(n)})|$,则需要证明 N_{k+1} 中的$\lambda(N_{k+1})$个配置仍然一致.因为$\lambda(N_{k+1})=|Dom(X_{\pi(n-k)})|\times\lambda(N_k)$,即 N_{k+1} 中的配置个数是 N_k 中的$|Dom(X_{\pi(n-k)})|$倍.设顶点 $X_{\pi(n-k)}$的条件偏好表的一行为“$x_1>x_2>\ldots>x_m$”($Dom(X_{\pi(n-k)})=\{x_1,x_2,\ldots,x_m\}$),则 N_{k+1} 中的配置是在 N_k 中的所有配置前分别添加 $x_1,x_2,\ldots,x_m$ 后而形成的,其中,首字母为 x_1 的$\lambda(N_k)$个配置为一堆(记作 x_1N_k),首字母为 x_2 的$\lambda(N_k)$个配置为一堆(记作 x_2N_k),…,首字母为 x_m 的$\lambda(N_k)$个配置为一堆(记作 x_mN_k).根据归纳假设可知,$x_i\lambda(N_k)(1\leqslant i\leqslant m)$堆中的所有配置都一致,而 x_1N_k 堆的配置排在 x_2N_k 堆之前,x_2N_k 堆的配置排在 x_3N_k 堆之前,…,x_mN_k 堆排在最后,这 m 堆的配置可看成是 m 个分支的图,其中,$x_{i-1}N_k$ 分支的最后一个配置排在 x_iN_k 分支的第一个配置前,其构造方式如图 4 所示.显然,N_{k+1} 是一个非环形序列,即 N_{k+1} 一致.

** 虽然文献[3,5]都直接或间接指出无环的 CP-nets 满足一致性,但还没有文献给出证明.

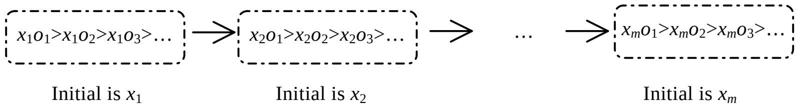

Fig.4 Outcome list of N_{k+1}

图 4 N_{k+1} 中的配置列表

(2) 通过构造反例证明有环 CP-nets 不一定满足一致性.

图 5(a)是有环 CP-nets,图 5(b)是给定的一种条件偏好表,其导出图如图 5(c)所示,此时,该 CP-nets 一致;当条件偏好表如图 5(d)所示时,其导出图则为图 5(e),此时,该有环图不一致.可见,有环图 CP-nets 的一致性与条件偏好表有关. □

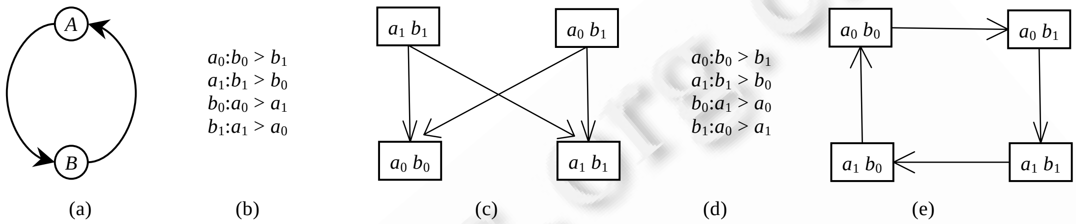

Fig.5 A CP-nets with cycle[3]

图 5 带环的 CP-nets[3]

两个有环 CP-nets 图的拓扑结构一样而条件偏好表不一样,则其一致性可能不一样.那么,一个有环图的条件偏好表具有什么样的特点,该有环图就一致呢?这是一个很有意义的开放问题.

4.2 CP-nets的一致性判定算法

无环 CP-nets 是一致的,而判断图是否存在环是一个可解问题,因此,无环 CP-nets 的一致性问题得到了彻底解决(即有判定定理及算法).虽然有环 CP-nets 的一致性不确定(与顶点的条件偏好表相关),但我们仍然可以利用强占优算法写出如下所示的一致性判定算法.至此,有环 CP-nets 的问题大部分得到解决(无直观的判定定理,但有直观的判定算法).

算法 1. *IsConsistent*(CP-nets N).

Input:CP-nets 图 $N=\langle V,CE\rangle$;

Output:若 N 一致,则输出 True;否则,输出 False.

Step 1

  If (N 是无环图) //利用图论算法来判断无环性

    Then Return True //输出该无环图是一致的

  Else //求有环图的传递闭包

    {

    构造 $N=\langle V,CE\rangle$的导出图 $N'=\langle \Omega,IE\rangle$

    调用 Warshall 算法求 IE 的传递闭包矩阵 B

    }

Step 2

  Consistent=0 //一致性的初始值

For i=1 to 2^n Do

  Consistent=*Consistent*+$B[i][i]$ //在矩阵 B 的对角线上找值为 1 的元素

End For

If (*Consistent*!=0) Then Return False; //若对角线有一个值为 1,则不一致

Else Return True

4.3 相关工作的对比

CP-nets 是一种重要的偏好表示工具,近几年,许多学者[1−7,11−13]都在研究与 CP-nets 有关的偏好处理问题.其中,文献[3,5]探讨 CP-nets 性质.为了理解本文特色,表 2 给出本文与经典论文[3]和利用复杂技术可求出强占优复杂度的文献[5]的对比.

Table 2 Comparison with Boutilier and Goldsmith's works about CP-nets

表 2 与 Boutilier 和 Goldsmith 关于 CP-nets 所作工作的对比

Related papers	Main work	Is strong dominance problem solved entirely?	Is strong dominance algorithm intuitive?	Did it discuss completeness?	Is completeness solved entirely? intuitive?	The solution of completeness and consistency
Boutilier's works[3]	Discuss the grammar, semantic and application of CP-nets	Solve partial (as shown in Table 1)	hard to understand[3]	No	No	No
Goldsmith's works[5]	Discuss the complexity of dominance and consistency of CP-nets	Only give complexity analysis, but not give algorithm	hard to understand, it uses complexity technique such as STRIPS planning[5]	Not discuss, but it is knew as the future work	No	The methods of strong dominance, completeness and consistency have not any relations
This paper	Give algorithm and complexity of any CP-nets, first study completeness, consistency algorithm	Solved entirely (have algorithm design and complexity analysis)	Simple and intuitive (solving the transitive closure)	Yes (gige the incompleteness theorem and counting formula of SCP)	Most are solved, and it is intuitive (enumerate the number of 1 in transitive closure matrix)	Reduction the completeness and consistency to strong dominance

由表 2 可知,本文的特色在于不仅给出了任意 CP-nets 的强占优的算法及复杂度,而且解决方法简单、直观.更重要的是,基于强占优算法给出了可分离 CP-nets 的计数公式,给出了 CP-nets 的一致性判定定理及算法.

5 结论和未来工作

本文以一种统一的框架研究了 CP-nets 的两个重要属性——完备性和一致性.其中,完备性描述了 CP-nets 对多个偏好强弱的表达能力,一致性描述 CP-nets 所表达的偏好中是否出现矛盾的断言.所谓的统一框架,是指完备性和一致性的判断和求解可归约为 CP-nets 的强占优.

进一步研究方向是:(1) 研究链式结构和树形结构的 CP-nets 可表达的偏好个数的解析表达式,即进一步解决 CP-nets 可表达的偏好关系的完备程度问题;(2) 进一步研究有环图的一致性,探求 CP-nets 的条件偏好表与一致性关系.同时探讨 CP-nets 的其他数学性质,如可满足性[3]、连通性、有界性等.并利用这些性质来研究偏好的学习,给出比文献[11−13]更严密的学习方法,从而从理论上确定 CP-nets 表达能力与学习能力的关系.

References:

[1] Walsh, T. Representing and reasoning with preferences. AI Magazine, 2007,28(4):59−69. [doi: 10.1007/978-3-540-30227-8_1]

[2] Brafman R, Domshlak C. Preference handling—An introductory tutorial. AI Magazine, 2009,30(1):58−86.

[3] Boutilier C, Brafman R, Domshlak C, Hoos H, Poole D. CP-nets: A tool for representing and reasoning with conditional ceteris paribus statements. Journal of Artificial Intelligence Research, 2004,21(1):135−191.

[4] Domshlak C, Hüllermeier E, Kaci S, Prade H. Preferences in AI: An overview. Artificial Intelligence, 2011,175(7-8):1037−1052. [doi: 10.1016/j.artint.2011.03.004]

[5] Goldsmith J, Lang J, Truszczy´nski M, Wilson N. The computational complexity of dominance and consistency in CP-nets. Journal of Artificial Intelligence Research, 2008,33(1):403−432.

[6] Brafman R, Domshlak C, Shimony E. On graphical modeling of preference and importance. Journal of Artificial Intelligence Research, 2006,25(1):389−424. [doi: 10.1613/jair.1895]

[7] Bouveret S, Endriss U, Lang J. Conditional importance networks: A graphical language for representing ordinal, monotonic preferences over sets of goods. In: Proc. of the 21th Int'l Joint Conf. on Artificial Intelligence. San Francisco: Morgan Kaufmann Publishers, 2009. 67−72.

[8] Benthem JV, Girard P, Roy O. Everything else being equal: A modal logic for ceteris paribus preferences. Journal of Philosophical Logic, 2009,38(1):83−125. [doi: 10.1007/s10992-008-9085-3]

[9] Lang J. Logical preference representation and combinatorial vote. Annals of Mathematics and Artificial Intelligence, 2004,42(1): 37−71. [doi: 10.1023/B:AMAI.0000034522.25580.09]

[10] Zhang ZZ, Xing HC, Wang QQ, Ni QJ. A preference logic based on various kinds of preferences. Journal of software, 2007,18(11): 2728−2739 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/18/2728.htm

[11] Lang J, Mengin J. The complexity of learning separable ceteris paribus preferences. In: Proc. of the 21th Int'l Joint Conf. on Artificial Intelligence. San Francisco: Morgan Kaufmann Publishers, 2009. 848−853.

[12] Koriche F, Zanuttini B. Learning conditional preference networks. Artificial Intelligence, 2010,174(11):685−703. [doi: 10.1016/j.artint.2010.04.019]

[13] Chevaleyre Y, Koriche F, Lang J, Mengin J, Zanuttini B. Learning Ordinal Preferences on Multiattribute Domains: The Case of CP-nets. Berlin: Springer-Verlag, 2011. 273−296. [doi: 10.1007/978-3-642-14125-6_13]

[14] Conitzer V. Making decisions based on the preferences of multiple agents. Communications of the ACM, 2010,53(3):84−94. [doi: 10.1145/1666420.1666442]

[15] Conitzer V, Sandholm T. Expressive markets for donating to charities. Artificial Intelligence, 2011,175(7-8):1251−1271. [doi: 10.1016/j.artint.2010.11.007]

[16] Liu JL, Zhang W, Tong XR, Zhang ZR. $O(2.983^n)$ time complexity algorithm for optimal coalition structure generation. Journal of software, 2011,22(5):938−950 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3817.htm [doi: 10.3724/ SP.J.1001.2011.03817]

[17] Conitzer V, Sandholm T, Lang J. When are elections with few candidates hard to manipulate. Journal of the ACM, 2007,54(3): 1−33. [doi: 10.1145/1236457.1236461]

[18] Pini MS, Rossi F, Venable KB, Walsh T. Incompleteness and incompareability in preference aggregation: Complexity results. Artificial Intelligence, 2011,175(7-8):1272−1289.

[19] Qu WL, Geng SY, Zhang LA. Discrete Mathematics. Beijing: Higher Education Press, 2008. 118−121 (in Chinese).

附中文参考文献:

[10] 张志政,邢汉承,王蓁蓁,倪庆剑.一种基于多类型偏好的偏好逻辑.软件学报,2007,18(11):2728−2739. http://www.jos.org.cn/1000-9825/18/2728.htm

[16] 刘惊雷,张伟,童向荣,张振荣.一种 $O(2.983^n)$时间复杂度的最优联盟结构生成算法.软件学报,2011,22(5):938−950. http://www.jos. org.cn/1000-9825/3817.htm [doi: 10.3724/ SP.J.1001.2011.03817]

[19] 屈婉玲,耿素云,张立昂.离散数学.北京:高等教育出版社,2008.118−121.

刘惊雷(1970－),男,山西临猗人,博士生,副教授,CCF 会员,主要研究领域为人工智能,理论计算机科学.

张伟(1961－),男,博士,教授,CCF 高级会员,主要研究领域为分布式人工智能.

廖士中(1964－),男,博士,教授,博士生导师,CCF 会员,主要研究领域为人工智能,理论计算机科学.