

直觉模糊近似推理中的可信度传播

夏博龄, 贺正洪, 雷英杰

XIA Bo-ling, HE Zheng-hong, LEI Ying-jie

空军工程大学 导弹学院 计算机工程系, 陕西 三原 713800

Missile Institute, Air Force Engineering University, Sanyuan, Shaanxi 713800, China

E-mail: summerylittle@sina.com

XIA Bo-ling, HE Zheng-hong, LEI Ying-jie. Confidence factors spreading in intuitionistic fuzzy approximate reasoning. Computer Engineering and Applications, 2009, 45(15): 160-162.

Abstract: To the confidence factors spreading in IFS approximate reasoning, the related formulas for finding the confidence factors of conclusions are exposed in this paper. First, with confidence factors been introduced, the effects of confidence factors spreading abroad in rules on those of conclusions are analyzed. To the IFL approximate reasoning with confidence factors, including the generalized modus ponens, generalized modus Tollens, generalized hypothetical syllogism on IFL, the related formulas are exposed. The correctness and validity of the proposed formulas are verified in a particular instance.

Key words: intuitionistic fuzzy approximate reasoning; confidence factors; confidence factors spreading

摘要: 针对直觉模糊近似推理中的可信度传播, 提出计算结论可信度的相关算法。首先介绍了可信度因子的概念, 分析了规则中的可信度因子传播对结论可信度的影响。然后针对直觉模糊近似推理的三种基本模式(取式、拒式、假言式), 给出了计算结论可信度值的相关公式, 最后通过实例验证了算法的正确性。

关键词: 直觉模糊近似推理; 可信度因子; 可信度传播

DOI: 10.3778/j.issn.1002-8331.2009.15.046 **文章编号:** 1002-8331(2009)15-0160-03 **文献标识码:** A **中图分类号:** TP182; TP391

1 引言

模糊性与随机性是现实世界中的两种主要的不确定性, 一方面, 直觉模糊方法是近年来处理信息的模糊性方面的热点方法, 直觉模糊集最初由保加利亚学者 Atanassov 于 1986 年提出^[1], 是对 Zadeh 模糊集理论最有影响的一种扩充和发展。Atanassov 系统提出并定义了直觉模糊集及其一系列运算和定理, 发展了直觉模糊逻辑的若干基本概念。同时, 国内外许多学者对此开展研究。Kevin Lano 提出了近似推理的基本框架^[2], 雷英杰研究了基于直觉模糊逻辑的近似推理^[3]等。

另一方面, 在处理信息的随机性方面, 可信度方法是肖特里菲等人在确定性理论(theory of confirmation)的基础上, 结合概率论等提出的一种不确定性推理方法^[4], 在专家系统MYCIN中得到了成功的应用。该方法简单直观, 效果也比较好, 受到人们的重视。

在可信度与直觉模糊推理的结合方面, 雷英杰研究了直觉模糊条件推理的可信度传播^[5], 进一步将可信度因子引入直觉模糊近似推理, 并将两种方法的优势结合起来, 满足同时处理模糊性和随机性的需求。

2 可信度

人们根据经验对一个事物或现象为真的相信程度为可信度^[6]。

证据的不确定性是用可信度因子(Certainty Factor, CF)表示的, 例如 $CF(E)=0.6$, 表示证据 E 的可信度为 0.6。证据 E 的可信度在 $[-1, 1]$ 上取值。

证据可通过析取或合取进行组合, 组合证据的可信度为:

$E=E_1 \text{ AND } E_2 \text{ AND } E_3 \cdots E_n$ 则

$$CF(E)=\min\{CF(E_1), CF(E_2), \dots, CF(E_n)\}$$

$E=E_1 \text{ OR } E_2 \text{ OR } E_3 \cdots E_n$ 则

$$CF(E)=\max\{CF(E_1), CF(E_2), \dots, CF(E_n)\}$$

在一般不确定性推理中, 知识是用产生式规则表示的, 其形式为:

if E then $H(CF(H, E))$

其中: $CF(H, E)$ 是知识的可信度。 $CF(H, E)$ 在 $[-1, 1]$ 上取值, 它指出当前提条件 E 所对应的证据为真时, 它对结论 H 为真的支持程度。

在 CF 模型^[6]中, 把 $CF(H, E)$ 定义为:

$$CF(H, E)=MB(H, E)-MD(H, E)$$

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.60773209); 陕西省自然科学基金(the Natural Science Foundation of Shaanxi Province of China under Grant No.2006F18)。

作者简介: 夏博龄(1985-), 女, 硕士生, 研究方向: 智能信息处理与信息融合; 贺正洪(1966-), 男, 副教授, 硕士生导师, 研究方向: 指挥自动化信息处理与辅助决策; 雷英杰(1956-), 男, 教授, 博士生导师, 研究方向: 智能信息处理与智能决策等。

收稿日期: 2008-03-20 **修回日期:** 2008-05-19

其中, MB (Measure Belief) 称为信任增长度, 它表示因与前提条件 E 匹配的证据的出现, 使结论 H 为真的信任增长度。MB 定义为:

$$MB(H, E) = \begin{cases} 1, & \text{若 } P(H) = 1 \\ \frac{\max\{P(H/E), P(H)\} - P(H)}{1 - P(H)}, & \text{否则} \end{cases}$$

MD (Measure Disbelief) 称为不信任增长度, 它表示因与前提条件 E 匹配的证据的出现, 对结论 H 为不信任增长度。MD 定义为:

$$MD(H, E) = \begin{cases} 1, & \text{若 } P(H) = 0 \\ \frac{\min\{P(H/E), P(H)\} - P(H)}{-P(H)}, & \text{否则} \end{cases}$$

上式中, $P(H)$ 表示 H 的先验概率; $P(H/E)$ 表示在前提条件 E 所对应出现的情况下, 结论 H 的条件概率。

根据 $CF(H, E)$ 的定义及 MB 与 MD 的互斥性, 可得到 $CF(H, E)$ 的计算公式:

$$CF(H, E) = \begin{cases} MB(H, E) - 0 = \frac{P(H/E) - P(H)}{1 - P(H)}, & \text{若 } P(H/E) > P(H) \\ 0, & \text{若 } P(H/E) = P(H) \\ 0 - MD(H, E) = -\frac{P(H) - P(H/E)}{P(H)}, & \text{若 } P(H/E) < P(H) \end{cases}$$

其中, $P(H/E) = P(H)$ 表示 E 所对应的证据与 H 无关。

3 可信度因子的传递

由于证据和知识中的可信度因子传递, 结论可信度的计算方法为:

$$CF(H) = CF(H, E) \times \max\{0, CF(E)\}$$

若由多条不同知识推出了相同的结论, 但可信度不同, 则可用合成算法求出综合可信度。

设有如下知识:

$$\text{IF } E_1 \text{ THEN } H \quad (CF(H, E_1))$$

$$\text{IF } E_2 \text{ THEN } H \quad (CF(H, E_2))$$

则结论 H 的综合可信度可分如下两步算出:

(1) 首先分别对每一条知识求出 $CF(H)$:

$$CF_1(H) = CF(H, E_1) \times \max\{0, CF(E_1)\}, CF_2(H) = CF(H, E_2) \times \max\{0, CF(E_2)\}$$

然后用下述公式求出 E_1 与 E_2 对结论 H 的综合影响所形成的可信度 $CF_{1,2}(H)$:

$$CF_{1,2}(H) = \begin{cases} CF_1(H) + CF_2(H) - CF_1(H) \times CF_2(H), & \text{若 } CF_1(H) \geq 0, CF_2(H) \geq 0 \\ CF_1(H) + CF_2(H) + CF_1(H) \times CF_2(H), & \text{若 } CF_1(H) < 0, CF_2(H) < 0 \\ \frac{CF_1(H) + CF_2(H)}{1 - \min\{|CF_1(H)|, |CF_2(H)|\}}, & \text{若 } CF_1(H) \text{ 与 } CF_2(H) \text{ 异号} \end{cases}$$

4 带有可信度因子的直觉模糊近似推理

直觉模糊近似推理基本的推理模式可归为三大类^[3]: 直觉模糊取式推理模型 (Generalized modus ponens)、直觉模糊拒式推理模型 (Generalized modus tollens)、直觉模糊假言式推理模型 (Generalized hypothetical syllogism)。

直觉模糊取式推理规则陈述的是, 给定两个直觉模糊命题

“ x is $P \rightarrow y$ is Q ”和“ x is P' ”, 可推出一个新的直觉模糊命题“ y is Q' ”。直觉模糊拒式推理规则陈述的是, 给定两个直觉模糊命题“ x is $P \rightarrow y$ is Q ”和“ y is Q' ”, 可推出一个新的直觉模糊命题“ x is P' ”。直觉模糊假言推理规则陈述的是, 给定两个直觉模糊命题“ x is $P \rightarrow y$ is Q ”和“ y is $Q \rightarrow z$ is L ”可推出一个新的直觉模糊命题“ x is $P' \rightarrow z$ is L' ”。

其中, x, y, z 是事物的名称, P, P' 是 U 上的直觉模糊集, Q, Q' 是 V 上的直觉模糊集, L 和 L' 是 W 上的直觉模糊集, 且 P 与 P', Q 与 Q', L 和 L' 都很近似。 Q 与 Q' 之间的差异越大, P 与 P' 之间的差异也越大, L 和 L' 同样。

对应这三种基本推理, 典型的含有可信度因子的直觉模糊推理及可信度的传播情形:

(1) 已知“ x 是 $P \rightarrow y$ 是 Q, CF_1 ”和“ x 是 P', CF_2 ”, 那么, 结论“ y 是 Q' ”的 CF 为:

$$CF(\text{“}y \text{ 是 } Q'\text{”}) = f(\delta(P, P'), CF_1, CF_2)$$

其中 $f(u, v, w)$ 为 u, v, w 的非减函数, $f(u, v, w) \leq v$ 和 $f(u, v, w) \leq w$, 且:

$$f(u, v, w) = u \cdot v \cdot w = u \cdot \min(v, w)$$

$\delta(P, P')$ 为衡量 P 与 P' 相似度或贴近度的函数。此处相似度函数取 Hong 和 Kim^[6] 提出的修正公式:

$$\delta(A, B) = 1 - \frac{1}{n} \sum_{i=1}^n \frac{|\mu_A(x_i) - \mu_B(x_i)| + |v_A(x_i) - v_B(x_i)|}{2}$$

此函数是在 Li 和 Cheng^[7] 的基础上修正得到, 弥补了原相似度计算公式中的缺陷。

(2) 已知“ x 是 $P \rightarrow y$ 是 Q, CF_1 ”和“ y 是 Q', CF_2 ”, 那么, 结论“ x 是 P' ”的 CF 为:

$$CF(\text{“}x \text{ 是 } P'\text{”}) = f(\delta(Q, Q'), CF_1, CF_2)$$

(3) 已知“ x 是 $P \rightarrow y$ 是 Q, CF_1 ”, “ y 是 $Q \rightarrow z$ 是 CF_2 ”且“ x 是 P', CF_3 ”, 那么结论是“ z 是 L ”的 CF 为:

$$CF(\text{“}z \text{ 是 } L'\text{”}) = f(\delta(P, P'), CF_1 \times CF_2, CF_3)$$

5 实例分析

假定内科专家有如下的两条临床经验^[8]:

(1) 如果 A (血清结合胆红素) 有或多或少的升高; B (血清非结合胆红素) 有较强地升高; C (尿胆红素) 无变化; D (尿胆元) 有很强地升高。那么患者患有 Q_1 (溶血性黄疸) 的确定性为 0.7。

(2) 如果 A 很强地升高; B 有较强地升高; C 中度地升高; D 无变化。那么患者患有 Q_2 (阻塞性黄疸) 的确定性为 0.8。

根据上述语句所含有的直觉模糊量词, 可以给它们赋予相应的隶属度, 于是这两条经验知识可表示成如下的规则形式:

规则 1

$$\begin{aligned} & \text{IF } (A, (0.30, 0.56)) \text{ and } (B, (0.85, 0.05)) \\ & \text{and } (C, (0.00, 1.00)) \text{ and } (D, (0.93, 0.01)) \\ & \text{THEN } Q_1 \quad CF_1: 0.7 \end{aligned}$$

规则 2

$$\begin{aligned} & \text{IF } (A, (0.93, 0.01)) \text{ and } (B, (0.10, 0.75)) \\ & \text{and } (C, (0.45, 0.36)) \text{ and } (D, (0.00, 1.00)) \\ & \text{THEN } Q_2 \quad CF_2: 0.8 \end{aligned}$$

假定患者的症状为:

$(A, (0.82, 0.08))$ and $(B, (0.00, 1.00))$ and $(C, (0.10, 0.75))$
and $(D, (0.07, 0.91))$

由上述已知条件,改写条件为:

规则 1 IF x is P_1 THEN y is Q_1, CF_1

规则 2 IF x is P_2 THEN y is Q_2, CF_2

其中 P_1, P_2 为论域 $U=\{A, B, C, D\}$ 上的直觉模糊集:

$P_1=$

$\{(0.30, 0.56)/A, (0.85, 0.05)/B, (0.00, 1.00)/C, (0.93, 0.01)/D\}$

$P_2=$

$\{(0.93, 0.01)/A, (0.10, 0.75)/B, (0.45, 0.36)/C, (0.00, 1.00)/D\}$

而患者症状同样为论域 $U=\{A, B, C, D\}$ 上的直觉模糊集 R :

$R=$

$\{(0.82, 0.08)/A, (0.00, 1.00)/B, (0.10, 0.75)/C, (0.07, 0.91)/D\}$

则易计算得到 $\delta(P_1, R)$ 为 0.36, $\delta(P_2, R)$ 为 0.815。

此题中, CF_1, CF_2 为知识不确定性,其值分别为 0.7, 0.8, 由题意认为证据的完全真实,即 $CF(x \text{ is } R)=1$, 根据直觉模糊近似推理中的可信度传播算法,可得:

患者为溶血性黄疸的可能性为:

$CF(y \text{ is } Q_1)=f(\delta(P_1, R), CF_1, 1)=f(0.36, 0.7, 1)=0.36$

患者为阻塞性黄疸的可能性为:

$CF(y \text{ is } Q_2)=f(\delta(P_2, R), CF_2, 1)=f(0.815, 0.8, 1)=0.8$

与原算例使用距离测度匹配算法推理结果为患者患有阻塞性黄疸结果一致,本算法是可行的,合理有效的。

6 结论

首先介绍了可信度因子的概念,分析了规则中的可信度因子传播对结论可信度的影响。然后针对直觉模糊近似推理的三种基本模式(取式、拒式、假言式),给出了计算结论可信度值的相关公式。通过实例验证了算法的正确性,可以看出将可信度因子引入直觉模糊推理中,丰富了不确定信息处理算法,一定程度上满足了同时处理模糊性和随机性的需求。

参考文献:

- [1] Atanassov K. Intuitionistic fuzzy sets[J]. Fuzzy Sets and Systems, 1986, 20(1): 87-96.
- [2] Lano K. Formal frameworks for approximate reasoning[J]. Fuzzy Sets and Systems, 1992, 51(2): 131-146.
- [3] 雷英杰, 王宝树, 路艳丽. 基于直觉模糊逻辑的近似推理方法[J]. 控制与决策, 2006, 21(3): 306-311.
- [4] 王永庆. 人工智能原理与方法[M]. 西安: 西安交通大学出版社, 1998, 171-182.
- [5] 雷英杰, 王宝树, 王晶晶. 直觉模糊条件推理与可信度传播[J]. 电子与信息学报, 2006, 28(10): 1790-1794.
- [6] Hong D H, Kim C. A note on similarity measures between vague sets and between elements[J]. Information Science, 1999, 115: 83-96.
- [7] Li D F, Cheng C. New similarity measure of intuitionistic fuzzy sets and application to pattern recognition[J]. Pattern Recognition Letters, 2002, 23(2): 221-225.
- [8] 李凡, 徐章艳. 一个有效的区间值模糊推理方法[J]. 应用科学学报, 2000, 18(4): 327-330.
- [9] 姜守旭, 李建中. 一种 P2P 电子商务系统中基于声誉的信任机制[J]. 软件学报, 2007, 18(10): 2551-2563.
- [10] Li H Z, Mukesh S. Trust management in distributed systems[J]. IEEE Computer Society, 2007, 2(40): 45-53.
- [11] Chang J S, Wang H M, Gang Y. A dynamic trust metric for P2P systems[C]. Proceedings of 5th International Conference on Grid and Cooperative Computing Workshops. Changsha: IEEE Press, 2006: 117-120.
- [12] 侯孟书, 卢显良, 周旭, 等. P2P 系统的信任研究[J]. 计算机科学, 2005, 32(4): 113-115.
- [1] Xiong L, Liu L. PeerTrust: Supporting reputation-based trust for peer-to-peer electronic communities[C]. IEEE Transaction on Knowledge and Data Engineering, 2004, 16(7): 843-857.
- [2] Jøsang A, Ismail R, Boyd C. A survey of trust and reputation systems for online service provision[J]. Decision Support Systems, 2007, 43(2): 618-644.
- [3] Kamvar S D, Schlosser M T. EigenRep: reputation management in P2P networks [C]. Proceedings of the 12th Int'l World Wide Web Conference. Budapest: ACM Press, 2003: 123-134.
- [4] 奚文, 王怀民, 贾焰, 等. 构造基于推荐的 Peer-to-Peer 环境下的 Trust 模型[J]. 软件学报, 2004, 15(4): 571-583.
- [5] Song S S, Hwang K, Zhou R F, et al. Trusted P2P transactions with fuzzy reputation aggregation[J]. IEEE Internet Computing, 2005, 9(6): 24-34.
- [6] 袁巍, 李津生, 洪佩琳. 一种 P2P 网络分布式信任模型及仿真[J]. 系统仿真学报, 2006, 18(4): 938-942.

(上接 138 页)