

解释学习中模糊概念的学习*

于 津

石纯一

(汕头大学计算机系, 汕头 515063)

(清华大学计算机系, 北京 100084)

摘要 本文提出了一种模糊知识的表示模式, 给出了在模糊意义下的基于解释的学习的一种描述, 以及解释、学习机制, 使在解释学习中能够学习到一些带有模糊修饰词的新概念. 本文描述形式同算子模糊逻辑而语义与其不同.

关键词 模糊概念, 解释机制, 学习机制.

目前的机器学习研究多是进行符号概念的学习, 对所涉及到的模糊概念的模糊性语义以及模糊概念的描述形式大多采取回避的态度. 本文首先借鉴了算子模糊逻辑(OFL)的表示方式, 但使用了与 OFL 不同的语义表示方法, 在基于解释的机器学习(EBL)中, 提出了一种显式的模糊知识的表示模式, 这使得领域知识的描述更丰富、更细致. 进而, 给出了在模糊意义下的 EBL 描述, 在 EBL 中引入了能处理模糊概念的解释机制、学习机制. 从而使得系统能够学习到一些带有模糊修饰词的新概念. 这是对 EBL 的一种推广.

1 模糊概念

以下给出模糊概念描述、内部约束、真值计算模式、微调模式以及修饰词模式.

1.1 模糊概念的描述

模糊概念从本质上来说就是语言变量的语言值, 而语言值多用模糊集合描述. 模糊集合一般由论域和隶属函数构成. 为了减少领域知识构建的复杂性, 对论域的描述只需给出直观的区间, 然后, 根据系统阈值(ω)和论域的值调整、算出要代入隶属函数的值.

同一个模糊概念, 在不同论域、不同系统阈值的情况下, 其结果是很不相同的, 但其隶属函数一般只需用一个, 不同的模糊概念对应不同的隶属函数.

对于同一个模糊概念, 如大、小, 既可以是大小杯子, 亦可以是大小箱子, 所以对这类模糊概念还要给出它所属概念的信息.

定义 1. 语言变量. 一般由两个原辞构成, 其值是由原辞通过修饰词和连接词派生出的不同的语言值. 语言变量的所有语言值的集合称为该变量的辞集.

用派生规则能生成辞集中的所有元素.

* 本文 1993-07-06 收到, 1994-03-03 定稿

本研究是国家自然科学基金资助项目. 作者于津, 1960 年生, 讲师, 主要研究领域为自动推理, Fuzzy 逻辑, 机器学习. 石纯一, 1935 年生, 教授, 博士导师, 主要研究领域为人工智能应用基础, 知识工程.

本文通讯联系人: 石纯一, 北京 100084, 清华大学计算机系

原辞是上下文相关的,而修饰词、连接词上下文无关,依赖于原辞产生作用.

定义 2. 一般概念(非模糊)的描述形式是: $P(t_1, \dots, t_n)$ (1)

模糊概念的描述形式是: $f-P(t_1, \dots, t_n)$ (2)

$\lambda f-P(t_1, \dots, t_n)$ (3)

$\lambda f-AP P(t_1, \dots, t_n)$ (4)

其中(1)、(4)式中的 P 是非模糊概念,(2)、(3)式中的 P 是模糊概念; λ 表示修饰词; AP 是修饰 P 的模糊概念; f —表示其后紧邻被修饰的是一个模糊概念,且有内部约束如下(称为模糊关系数据库 $FRDB$):

模糊谓词名	所属概念名	L	H	$POINTER$

其中 L, H 表示模糊谓词的论域区间的下、上界, $POINTER$ 指示是隶属函数还是模糊关系. 为方便,非模糊概念和模糊概念统称为概念.

对具体的模糊概念计算其真值时,还需给出其所对应的隶属函数或模糊关系.

1.2 模糊概念的值

模糊概念的值依赖于其隶属函数或模糊关系.

1.2.1 隶属函数的选择

隶属函数选择是重要的,模糊在很大程度上都表现为主观的,隶属函数的选择亦是主观的.但一些常用的在实轴上均匀分布的一类模糊概念的隶属函数还是有一定规律的.

如分段平方函数 S -函数:

$$S(U; \alpha, \beta, \gamma) = \begin{cases} 0 & U \leq \alpha \\ 2((U-\alpha)/(\gamma-\alpha))^2 & \alpha \leq U \leq \beta \\ 1-2((U-\gamma)/(\gamma-\alpha))^2 & \beta \leq U \leq \gamma \\ 1 & U \geq \gamma \end{cases} \quad (5)$$

其中 $\beta = (\gamma + \alpha) / 2$ 是过渡点. 即若 $U = \beta$ 则 $S(U; \alpha, \beta, \gamma) = 0.5$, α, γ 对应定义 2 中的用 L, H 和 ω 算出的值.

模糊概念的描述若存在多方面因素时,多选用模糊关系来描述,这样更方便、直观.对不能用 S -函数作隶属函数的模糊概念要依具体情况选择合适的隶属函数.

1.2.2 修饰词

修饰词没有上下文关系,它是去修饰模糊概念,构成新模糊概念,其隶属函数是由原模糊概念通过修饰词的语义构成的新隶属函数.

如给出原辞:年轻;修饰词:有点,稍微,非常, ..., 则一般情况可规定:

$$\mu_{\text{非常年轻}} = \mu_{\text{年轻}}^2 \quad \text{称为“集中”}$$

$$\mu_{\text{有点年轻}} = \sqrt{\mu_{\text{年轻}}} \quad \text{称为“扩张”}$$

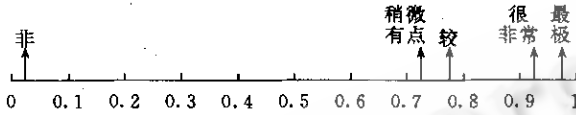
由此定义了修饰词非常,有点的语义,以及新概念非常年轻,有点年轻.

我们采取在 $[0, 1]$ 区间的实轴上充填修饰词的方法,通过对几个常用修饰词的语义定义,取出它在 $[0, 1]$ 区间中的值,按程度描述的相对性,把一些不太常用的修饰词插入 $[0, 1]$ 区间.如讨论年轻的概念:

$$\mu_{年轻}(U) = 1 - S(U; \alpha, \beta, \gamma)$$

若本系统中给出系统阈值 $\omega = 0.85$, 则只要模糊值 ≥ 0.85 就认为 U 是年轻的, 那么在年轻的集合中, 值大于多少才能算作很年轻, 有点年轻, ..., 由 $\mu_{很年轻} = \mu_{年轻}^2$ 得出 $\mu_{年轻}(U) \geq 0.92$ 时, 则 U 应被认为是很年轻的; 由 $\mu_{有点年轻} = \sqrt{\mu_{年轻}}$ 得出 $\mu_{年轻}(U) \geq 0.72$ 时, 则 U 应被认为是有点年轻的.

所以依修饰词之间程度的相对差异在 $[0, 1]$ 区间上插入许多其它修饰词, 如下:



将 $[0, 1]$ 区间上的修饰词与模糊谓词的真值匹配, 还原为语言值. 这是本文的一个有特色的处理, 使得学到的知识在形式上更接近于自然语言的描述.

修饰词模式: 依据模糊概念的值与 $[0, 1]$ 区间上的修饰词相匹配, 得到一对应的修饰词 λ , 返回形如 $\lambda f - P(t_1, \dots, t_n)$ 的模糊概念.

修饰词的增删在具体系统中调整、掌握. 且还可有一个对应的英文修饰词图尺.

1.2.3 模糊概念求值

解释和学习机制中需要使用下列模式, 以完成求值和学习的功能.

1. 模糊模式

需完成以下计算:

- ①判断触发概念提供的信息是否完全;
- ②在 $FRDB$ 中查找相应信息计算真值. 先计算要代入隶属函数的 α 和 γ 的值:
 - 若是 S -函数, 则 $\alpha = (\sqrt{\omega/2}H - L) / (1 - \sqrt{\omega/2}), \gamma = H$ (6)
 - 若是 $1-S$ -函数, 则 $\alpha = L, \gamma = (H - \sqrt{\omega/2}L) / (1 - \sqrt{\omega/2});$ (7)
- ③若此概念是修饰目标概念的则标志此概念.

2. 微调模式

需完成以下计算: (λ' 是目标概念的修饰词 λ 对应的值, δ' 是新规则头的修饰词 δ 对应的值, f_{FP} 是被修饰的模糊概念名)

- ①如 $|\lambda' - \delta'| \leq \epsilon$ 则结束 (ϵ 是系统给出的能使 λ 和 δ 在同一水平的小正数);
- ②否则, 询问微调原则 (一般用 λ 替 δ , 特殊的要按新原则处理);
- ③若需要则调整 f_{FP} 对应的模糊概念的 $FRDB$;
- ④按上述原则调整得到一新修饰词 λ'' , 作为新规则头的修饰词.

2 模糊解释学习

2.1 在模糊意义下的 EBL 描述

给出:

- * 系统阈值 STV(System Threshold Value): ω ,刻划模糊概念的真值水平, ϵ ,小正数,刻划同一层次修饰词的差异;
- * 目标概念 GC(Goal Concept):要学习概念的高级描述,可以是定义 2 中的四种形式;
- * 训练例 TE(Test Example):目标概念的一个例子;
- * 领域理论 DT(Domain Theory):一组规则和事实的集合,构造解释树的依据;
- * 可操作准则 OC(Operationality Criterion):凡被描述的模糊概念可操作性是可判定的,真值 $\geq\omega$ 可操作;非模糊概念以能有效识别为准则规定其表示形式.

2.2 领域理论

知识形式如: $A \leftarrow B$,其中, A 是概念,称为规则头; B 是概念的合取形式,称为规则尾;

2.3 解释机制

构造解释树.

1. (1)型概念 $P(t_1, \dots, t_n)$ 在 DT 中匹配:用熟知的 EBL 解释机制.

2. (2)型概念 $f-P(t_1, \dots, t_n)$ 在 DT 中匹配:

若条件满足,则触发 P 对应的模糊模式,计算真值;若不能触发,则匹配 $f-P(t_1, \dots, t_n)$ 规则的尾部,直到返回相应所需条件再触发该模式;真值 $\geq\omega$ 则匹配;触发修饰词模式.

3. (3)型概念 $\lambda f-P(t_1, \dots, t_n)$ 在 DT 中匹配:

若完全匹配,则去匹配此规则的尾部;否则,用 $f-P(t_1, \dots, t_n)$ 在 DT 中匹配,方式同 2.

4. (4)型概念 $\lambda f-AP P(t_1, \dots, t_n)$ 在 DT 中匹配:

若完全匹配,则去匹配此规则的尾部;否则,用 $f-AP P(t_1, \dots, t_n)$ 去匹配 DT 中的 $\delta f-APP(t_1, \dots, t_n)$ 形式的概念,若 $|\lambda-\delta| \leq \epsilon$,则去匹配此规则的尾部;否则,用 $P(t_1, \dots, t_n)$ 在 DT 中匹配,方式同 1.

2.4 概括

对解释进行概括,就是要将解释树所有结点上的被匹配的概念中的常量改为变量,但对做过标志的模糊概念中与模糊模式相关的常量不作改变,且将上述模糊概念下所属的分枝剪掉,如此得到概括后的解释结构.

2.5 学习机制

将上述概括后得到的结构中的所有叶结点作合取,作为新规则的尾,以解释树的根结点上的概念作为新规则的头,然后再进一步学习.分几种情况考察此规则.我们可以学到不同的概念表示形式如下:

1. 目标概念不是模糊概念,且规则尾中亦不含模糊概念,则学习过程结束;

2. 目标概念不是模糊概念,且规则尾中含模糊概念,但不是修饰目标的,则将模糊概念中的常量变变量,学习过程结束.

3. 目标概念不是模糊概念且规则尾中有修饰目标的模糊概念 AP 及算出的修饰词 λ ,则产生新概念 $\lambda f-AP P(t_1, \dots, t_n)$ 作为规则的头,规则尾部不变,此即新学到的概念.

4. 目标概念是形同模糊概念 $f-P(t_1, \dots, t_n)$ 的形式,将计算出的修饰词 λ 加在 $f-P(t_1, \dots, t_n)$ 之前,成为新的概念 $\lambda f-P(t_1, \dots, t_n)$,尾部不变.

5. 目标概念是形同模糊概念 $\lambda f-P(t_1, \dots, t_n)$ 或 $\lambda f-AP P(t_1, \dots, t_n)$ 形式,触发微调模

式,以使从知识库匹配得到的结果与 λ 一致.

2.6 可操作性

对模糊概念,通过触发其模糊模式,计算模糊真值,对于真值大于等于系统阈值 ω 的那些模糊概念,认为是可操作的,这样,我们明确给出了这一类概念的可操作性准则.

2.7 示 例

给出:

- * 目标概念, $\text{cup}(c)$;
- * 系统阈值: $\omega=0.85, \epsilon=0.02$;
- * 训练例:

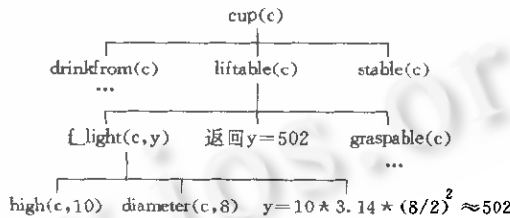
$\text{has}(c, b), \text{concavity}(b), \text{upward_pointing}(b), \text{bottom}(c, f), \text{flat}(f), \text{has}(c, h), \text{handle}(h), \text{high}(c, 10), \text{diameter}(c, 8)$.

* 领域知识:

$\text{cup}(x) \leftarrow \text{drinkfrom}(x), \text{liftable}(x), \text{stable}(x)$.
 $\text{drinkfrom}(x) \leftarrow \text{has}(x, y), \text{concavity}(y), \text{upward_pointing}(y)$.
 $\text{stable}(x) \leftarrow \text{bottom}(x, y), \text{flat}(y)$.
 $\text{liftable}(x) \leftarrow \text{graspable}(x), \text{f_light}(x, y)$.
 $\text{graspable}(x) \leftarrow \text{has}(x, y), \text{handle}(y)$.
 $\text{f_light}(x, y) \leftarrow \text{high}(x, y), \text{diameter}(x, d), y = h * \pi * (d/2)^2$.

* 对应的 FRDB 为:

谓词名	所属概念名	L	II	POINTER
light	cup	150	2000	1-S

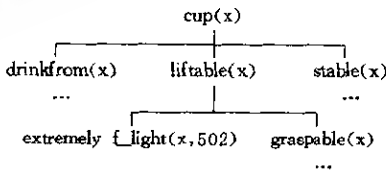


* 算出: $\alpha=150, Y=5434, \beta=2642$; 用公式(7). $\mu_{\text{light}}(502)=0.99$; 用公式(5).

* 由修饰词区间匹配得到, $\lambda=\text{extremely}$;

* 使 f_light 变为: $\text{extremely f_light}(c, 192)$;

* 概括:



* 学习到的新概念: $\text{extremely f_light cup}(x) \leftarrow \text{has}(x, y1), \text{concavity}(y1),$

$\text{upward_pointing}(y1), \text{bottom}(x, y2), \text{flat}(y2), \text{has}(x, y3), \text{extremely f_light}(x, 502)$.

3 结 语

本文是在模糊概念的学习中提出了一种新的模式,做了一些初步的工作,引入了模糊概念模式,使得对模糊概念的描述更细致,学习能力更强健,但同时也使领域知识的构建难度加大.进一步的工作还有待展开.象模糊性在规则上的传播,多个修饰词的处理,相反意义的修饰词如何用同一种模式表示等.

本文是作者于津在清华大学工作期间完成的.

致谢 感谢张旗、郝继刚对本文工作提出的有益建议和意见.

参 考 文 献

- 1 陈国权. 知识工程中自然语义的模糊表达. 北京: 科学出版社, 1989.
- 2 Dejong G. An introduction to EBL. AAAI-88 Tutorial, 1988.
- 3 Liu X H, Xiao H. Operator fuzzy logic and fuzzy resolution. In: Proc. of the 15th ISMVL, Kingston, Canada, 1985. 68-75.
- 4 Mani G. Some approaches to handle noise in concept learning. Int. J. Man. Machine Studies, 1992, 36(2), 167-181.
- 5 Mitchell T M *et al.* EBG: a unifying view. Machine Learning, 1986, 1(1): 47-80.
- 6 Mooney R J. A general explanation based learning mechanism and its application to narrative understanding. Morgan Kaufmann, 1990.

LEARNING FUZZY CONCEPTS IN EXPLANATION BASED LEARNING

Yu Jin

(Department of Computer Science, Shantou University, Shantou 515063)

Shi Chunyi

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

Abstract This paper presents a model of representing fuzzy knowledge. Based on this model, the authors give a description of explanation based learning (EBL), and next give explaining and learning mechanism of EBL. Thus it is possible to learn fuzzy concept in EBL. In this paper, the pattern in description is the same with operator fuzzy logic but semantics.

Key words Fuzzy concept, explaining mechanism, learning mechanism.