

基于 Agent 的知识表达度量理论^{*}

The Measure Theory of Knowledge Representation Based on Agent

李凡长

(苏州大学计算机工程系 苏州215006)

Abstract Agent is like human or intelligence object, which has ability to autonomy, guess, repercussion, cooperate, self-techer, mutual learning, and so on. Based on Agent, we have tried various ways to quantitative analysis for knowledge representation. From measure theory of knowledge, using measure principle of information theory, we study knowledge representation thoroughly, and regard Agent as carrier of knowledge, and give the measure theory of knowledge representation.

Keywords Knowledge representation, Quantitative analysis, Measure theory, Agent

一、前言

Agent 是指具有智能的人,或其它智能物或相当于智能物的实体,以及能进行种种工作的软件等^[1]。这种 Agent 对给定介质环境中的事件变化能迅速作出反映,具有自治、推理、反响、协作、自学习和相互学习的能力等,知识是人类智慧认识世界的结晶,从知识的起源来看,它来自于人类对客观世界的认识和实践,是人们对客观事物及其联系的认识和知道如何去改变客观环境的能力。知识表达是智能计算技术研究的最基本问题之一。目前在人工智能领域,对知识表达的研究取得了相当大的进展,不仅提出了很多行之有效的知识表达方法,如语义网络、框架、逻辑等,而且开发了很多实际的智能系统进行实现验证和评价,同时对这些知识表达方法进行了一定的分析比较^[2~4]。从以上可以看出,过去人们研究的知识表达方法仅仅是一种推理表达模型,不是一个完善的智能行为,本文抓住 Agent 的特点,从智能体的特点出发,把知识当作是智能体的一部份进行研究,给出基于 Agent 的知识表达度量理论。

二、Agent 与知识的相关概念

为了以后研究的方便,我们先讨论 Agent 与知识的一些基本概念。

约定1 Agent 是一个智能信息系统,表示为“Ag”。

约定2 信息是研究系统的所有要素,表示为

“Inf”。

约定3 知识是人类智能的结晶,是信息的子集,表示为“Kn”。

由此有: $\{Kn\} \subseteq \{Inf\} \subseteq \{Ag\}$

假设1 可测假设:客观世界知识集 Kn,信息集 Inf, Ag 均是可测集。

假设2 Agent 表示:设 X 是 Ag 的一个子集,若存在一类集 $\sum = \{x_i, i \in J\}$, 索引集 J 为有限集或可列集,满足:

$$X = \bigcup_{i \in J} x_i$$

那么,称集类 $\sum$ 为 X 的一个表达,特别地,如 $\sum$ 中集会互不相交的,即

$$i_1, i_2 \in \sum, i_1 \neq i_2, i_1 \cap i_2 = \emptyset$$

则称 $\sum$ 为 X 的一个无二义性的理想表达, x_i 则为 $\sum$ 的一个表达基元。

假设3 Agent 测度:设 X 是 Ag 的一个子集,即 $X \subseteq Ag, A: \sum \rightarrow [0, 1]$, 是定义在 $\sum$ 上的集函数,若 A 满足:(1) $A(\emptyset) = 0$; (2) $A(x) = 1$; (3) $\forall a, b \in \sum$, 如果 $a \subset b$, 那么 $A(a) \leq A(b)$, 则称 A 为 X 上的一个表达测度,当 $X = Ag$ 时,则称 A 为 Ag 上的一个表达测度。

显然这种表达测度和人们通常所说的抽象度、粒度、细度是一致的。

三、基于 Agent 的知识表达度量理论

本节主要讨论 Agent 表示知识的度量理论框架。

^{*} 本文的研究得到国家自然科学基金和苏州大学科研处基金资助。李凡长 教授,主要研究方向是人工智能、动态模糊智能理论,智能软件等。

1. 知识量

定义1 设 X 为 Ag 的某一邻域, $\sum = \{x_i, i \in J\}$ 均为有限集或可列集, $j=1, 2, \dots, N$, (N 为有限或 $+\infty$, 下同) 为 X 的一个表达, μ 为 Ag 上的一个表达测度, 对 $\forall x_i \in \sum$, 令

$$I(x_i) = \begin{cases} \log_b \mu(x_i), & \mu(x_i) > 0 \\ +\infty, & \mu(x_i) = 0 \end{cases}$$

则称 $I(x_i)$ 为表达 $\sum$ 中的一个基元 x_i 的自蕴含知识量, 其中对数底数 $b > 1$ (以下, 底数 b 从略)。

对于表达 $\sum = \{x_i, i=1, 2, \dots, N\}$ 令

$$\mu_j = \mu(x_j), j=1, 2, \dots, N, \mu = \sum_{i=1}^N \mu_i,$$

$$P_i = \frac{\mu_i}{\mu} = \frac{1}{\mu_i} \sum_{j=1}^N \mu_j, j=1, 2, \dots, N$$

$$I(\sum) = \sum_{i=1}^N P_i \cdot I(x_i),$$

则称 $I(\sum)$ 为表达 $\sum$ 的平均知识量, 其中约定 $0 \cdot \log \infty = 0$ 。

2. 熵

定义2 设 X 为 Ag 的一个邻域, $\sum = \{x_i, i \in J\}$, J 为有限集或可列集, $j=1, 2, \dots, N$, μ 为 Ag 上的表达测度, $\mu_j (j=1, 2, \dots, N)$, $P_j (j=1, 2, \dots, N)$, μ 的含义同上, 那么表达 $\sum$ 的熵为:

$$H(\sum) = I(\sum) - I(X) = \log \mu(x) - \frac{1}{\mu} \sum_{i=1}^N \mu_i \log \mu_i$$

$$= \log \frac{\mu(x)}{\mu} - \sum_{i=1}^N P_i \log P_i$$

显然, 熵 $H(\sum)$ 恒大于等于 0。

如果 $\sum$ 为 X 的一个理想表达, 则其熵记为

$$H_D(\sum) = - \sum_{i=1}^N P_i \log P_i$$

3. 等价知识基元个数

定义3 设 X 为 Ag 的一个子集, $\sum$ 为 X 上的一个表达, μ 为 Ag 上的表达测度, 则称

$$\tilde{N}(\sum) = \mu(x) / \hat{\mu}(\sum)$$

为 $\sum$ 在平均知识量定义下所包含的等价的表达基元的个数, 即等价知识基元个数, 并称平均知识意义下的表达基元为知识基元。其中

$$\hat{\mu}(\sum) = \prod_{i=1}^N \mu_i^{P_i}, P_i = \mu(x_i) / \sum_{i=1}^N \mu(x_i)$$

显然, 根据定义2给出的熵的定义, 有 $H(\sum) = \log \tilde{N}(\sum)$ 。

定义4 设 M 为某一表达 Agent, N 为 M 所能表

达的知识基元个数, 则该表达 Agent M 的表达熵为:

$$H_e(M) = \log N,$$

定义5 设 X 是 Ag 的一个邻域, $\sum_x, \sum_y$ 为 X 上的两个表达, M_1, M_2 为两个表达 Agent, 令

$$D(M_1, M_2) = H_e(M_2) - H_e(M_1)$$

$$1) (\sum_x, \sum_y) = H(\sum_y) - H(\sum_x)$$

则称 $D(M_1, M_2)$ 为 Agent M_1, M_2 熵差, 称 $D(\sum_x, \sum_y)$ 为表达 $\sum_y$ 到表达 $\sum_x$ 的熵差。

对于两个表达 Agent M_1, M_2 , 如果 $D(M_1, M_2) = 0$, 则称 M_1, M_2 在知识量意义下表达能力等价。如果 $D(M_1, M_2) > 0$, 则称在知识意义下 M_2 的表达能力大于 M_1 的表达能力。反之则在知识意义下 M_1 的表达能力大于 M_2 的表达能力。

根据 Agent 的表达熵及熵差的概念, 就可以对单个表达式及多个表达式进行分析比较。

例1 一个 Agent M_1 表达“一幅图画”, 另一个 Agent M_2 表达“一千字文”, 现比较如下:

设一幅图画 I 有 400×640 个像素, 每个像素可取 10 个像素值, 则其所能表达的知识基元的个数 $N_I = 10^{400 \times 640} = 10^{256 \times 10^3}$, 其表达熵 $H_e(I) = \log N_I = 2.56 \times 10^5 \times \log 10$ 。另一个千字文的文本 S , 每个字可从一个万字表中选取, 则其所能表达的知识基元的个数 $N_S = 10000^{1000}$, 其表达熵为 $H_e(S) = \log N_S = 4000 \times \log 10$ 。显然, $D(S, I) > 0$, 因此, 在知识意义下, M_1 的表达能力远远大于 M_2 , 但对于某一个邻域 X 上的一个具体表达 $\sum$, 其等价知识基元个数为 $\tilde{N}(\sum)$, 只有当 $\tilde{N}(\sum) \leq N_S$ 时, $D(S, I) = 0$, 它们的表达能力才相等, 否则, 图画 I 的表达能力均大于千字文 S 的表达能力。

4. 联合表达及知识量

定义6 (联合表达) 设 X 为 Ag 的一个邻域, $\sum_x, \sum_y$ 为 X 的两个表达, 其中

$$\sum_x = \{Z_{ij}, i=1, \dots, N\}, \sum_y = \{y_j, j=1, 2, \dots, M\}$$

令 $\sum_{xy} = \{Z_{ij}, i=1, \dots, N; j=1, 2, \dots, M, Z_{ij} = x_i \cap y_j\}$ 则称 $\sum_{xy}$ 为 $\sum_x, \sum_y$ 的 (二元) 联合表达, $\sum_x, \sum_y$ 则称为 $\sum_{xy}$ 的边际表达, 与联合表达、边际表达相对应的度量特征为联合熵和边际熵。

设 μ 为 Ag 上的表达测度, 令

$$\mu_{ij} = \mu(Z_{ij}), i=1, \dots, N, j=1, \dots, M,$$

$$\begin{aligned} \mu_x &= \sum_{i=1}^N \sum_{j=1}^M \mu_{ij}, \mu_y = \sum_{j=1}^M \sum_{i=1}^N \mu_{ij}, \mu_i = \sum_{j=1}^M \mu_{ij}, \mu_j = \sum_{i=1}^N \mu_{ij}, \\ &= \sum_{i=1}^N \mu_{i0}, \end{aligned}$$

再令 $P_{ij} = \mu_{ij} / \mu_i, i=1, \dots, N, j=1, \dots, M,$

$P_{ji} = \mu_{ji} / \mu_j, i=1, \dots, N, P_{ji} = \mu_{ji} / \mu_j, j=1, \dots, M$

则联合熵 $H(\sum_x, \sum_y)$ 及边际熵 $H(\sum_x), H(\sum_y)$ 分别为

$$H(\sum_x, \sum_y) = H(\sum_z) = - \sum_{i=1}^N \sum_{j=1}^M P_{ij} \log P_{ij},$$

$$H(\sum_x) = - \sum_{i=1}^N P_{i\cdot} \log P_{i\cdot}, H(\sum_y) = - \sum_{j=1}^M P_{\cdot j} \log P_{\cdot j},$$

如果对 $\forall i, j, i=1, \dots, N, j=1, \dots, M$, 恒有 $P_{ij} = P_{i\cdot} \cdot P_{\cdot j}$, 则设表达式 $\sum_x, \sum_y$ 为相互独立, 显然, 对于相互独立的两个表达式 $\sum_x, \sum_y$ 恒有

$$H(\sum_x, \sum_y) = H(\sum_x) + H(\sum_y)$$

定义7(条件表达) 根据联合表达的定义, 进一步令 $\sum_i = \{\sum_{i1}, \dots, \sum_{iM}\}, i=1, \dots, N$, 则 $\sum_i$ 构成了 X_i 的个表达, 称 $\sum_i$ 为 $\sum_i$ 的基元 X_i 在已知条件 $\sum_y$ 下的条件表达, 其对应的条件熵为:

$$H(\sum_i / \sum_y) = \log \mu_{i\cdot} - \frac{1}{\mu_{i\cdot}} \sum_{j=1}^M \mu_{ij} \log \mu_{ij}, i=1, \dots, N$$

其平均条件熵 $H(\sum_y / \sum_x)$ 为

$$H(\sum_y / \sum_x) = \sum_{i=1}^N P_{i\cdot} H(\sum_i / \sum_y) = \sum_{i=1}^N \frac{\mu_{i\cdot}}{\mu} (\log \mu_{i\cdot} -$$

$$\frac{1}{\mu_{i\cdot}} \sum_{j=1}^M \mu_{ij} \log \mu_{ij}) = \frac{1}{\mu} \sum_{i=1}^N \sum_{j=1}^M \mu_{ij} \log \frac{\mu_{ij}}{\mu_{i\cdot}}$$

类似地 $H(\sum_x / \sum_y) = \log \mu_{\cdot j} - \frac{1}{\mu_{\cdot j}} \sum_{i=1}^N \mu_{ij} \log \mu_{ij},$

$$H(\sum_x / \sum_y) = \frac{1}{\mu} \sum_{j=1}^M \sum_{i=1}^N \mu_{ij} \log \frac{\mu_{ij}}{\mu_{\cdot j}}$$

定义8(互知识量) 设 X 为 A_g 的一个邻域, $\sum_x, \sum_y$ 为 X 上的两个表达, 则表达式 $\sum_x$ 与 $\sum_y$ 之间的互知识量 $I(\sum_x, \sum_y) = H(\sum_z) - H(\sum_y / \sum_x)$, 即

$$I(\sum_x, \sum_y) = \frac{1}{\mu} \sum_{i=1}^N \sum_{j=1}^M P_{ij} \log \frac{P_{ij}}{P_{i\cdot} P_{\cdot j}},$$

由对称性可知:

$$I(\sum_y, \sum_x) = H(\sum_z) - H(\sum_x / \sum_y)$$

由知识量, 联合熵, 边际熵的定义, 可以得到如下推论:

推论 $I(\sum_x, \sum_y) = H(\sum_z) + H(\sum_y) - H(\sum_x, \sum_y)$

根据二元联合表达的定义, 可进一步推广为 K 元联合表达, 设 X 是 A_g 的一个邻域, $\sum_1, \sum_2, \dots, \sum_k$

为 X 上的 k 个表达, 其中 $\sum_i = \{x_i^j | j=1, \dots, N^i\},$

令 $\sum_z = \{Z_{i_1 i_2 \dots i_k} | Z_{i_1 i_2 \dots i_k} = x_{i_1}^{j_1} \cap x_{i_2}^{j_2} \cap \dots \cap x_{i_k}^{j_k}, i_1=1, \dots, N^1, \dots, i_k=1, \dots, N^k\}$, 则称 $\sum_z$ 为 $\sum_1, \sum_2, \dots, \sum_k$ 的 k 元联合表达, 其对应的联合熵为:

$$H(\sum_1, \sum_2, \dots, \sum_k) = H(\sum_1) + H(\sum_2 / \sum_1) + \dots + H$$

$$(\sum_k / \sum_1, \sum_2, \dots, \sum_{k-1}) = \sum_{i=1}^k H(\sum_i / \sum_1, \sum_2, \dots, \sum_{i-1})$$

熵在知识表达的定量分析研究中起着重要作用, 熵增, 意味着获取了更多的知识, 熵减, 则意味着知识的丢失。接下来给出最大熵定理和熵增定理。

定理1 单一表达最大熵定理: 设 X 是 A_g 的一个邻域, $\sum = \{x_i | i=1, \dots, N\}$ 为 X 的一个表达, μ 为 A_g 上的测度, 令 $\mu_i = \mu(x_i), i=1, \dots, N$. 则当 $\mu_1 = \mu_2 = \dots = \mu_N$ 时, $H(\sum)$ 有最大值 $\log N$.

证明: 略

定理2 联合表达最大熵定理: 设 X 是 A_g 的某一邻域, $\sum_x, \sum_y$ 为 X 上的两个表达, 如果 $\sum_x, \sum_y$ 相互独立, 则 $H(\sum_x, \sum_y)$ 的熵达到最大, 且等于 $H(\sum_x) + H(\sum_y)$.

证明: 略

定理3 复合熵增定理: 设 X 是 A_g 的一个邻域, $\sum_x = \{x_i | i=1, \dots, N\}$ 为 X 的一个表达, 对于每一个 x_i , 又存在一个表达 $\sum_i = \{x_{i1}, x_{i2}, \dots, x_{iM}\}$ (对于不同的 i, M 可以不同, 但可以通过增加若干空集, 使所有 M 都相同), 这样便形成 $\sum_x$ 的一个复合表达:

$$\sum' = \{x_{11}, x_{12}, \dots, x_{1M}, x_{21}, x_{22}, \dots, x_{2M}, \dots, x_{N1}, \dots, x_{N2}, \dots, x_{NM}\}$$

那么 $H(\sum') \geq H(\sum_x)$

证明: 略

定理4 联合熵增定理: 设 X 是 A_g 的一个邻域, $\sum_x, \sum_y$ 为 X 上的两个表达, $\sum$ 为 $\sum_x, \sum_y$ 所形成的一个联合表达, 则

$$H(\sum) \geq H(\sum_x), H(\sum) \geq H(\sum_y)$$

证明: 略

四、相关工作比较

目前人们用于表达知识的表示方法比较典型的有语义网络, 框架表示和逻辑。

1. 语义网络表达^[8] 语义网络是由 Quillian 在 1968 年首先提出的, 它是一种有向图的知识表达方式,

结点、表示客体、客体的性质、概念、事件、行为等,有向边表示结点间的关系,语义网络知识库的刷新为有向图中“结点-连接边”的增减或修改。

2. 框架表示^[5] 框架是由 Minsky 1975 年提出的,本意是将语义网络过程等知识表达方式结合起来描述一些固定的环境和常规行为,一般将它归类为一种聚集语义网^[1]。其中的语义网聚集结点对应于框架名,其聚集元素为框架的槽(slot),这些槽及其相应的值构成了框架的一个联合表达。因此,根据联合最大熵定理,在使用框架表达方式时,应使其各个槽及相应的值相互独立,这个结点基本与人们的实际使用经验相符合。

3. 逻辑表达 逻辑是由一组严格定义的表达知识的语法符号,加上一些严格定义的对这些符号表达的知识进行解释和操作的形式化方法组成,一般说来,任何满足以下条件的知识处理系统即可看作是一个逻辑“系统”^[1]。

(1)有一种严格定义的知识表达语言(语法结构);

(2)有一种严格定义的模型理论(语义),给这种语言的语句赋值的语义;

(3)有一种严格定义的证明理论及形式推理方法,对这种语句进行语法操作,从一些语句推导出另一些语句。

4. 本文基于 Agent 的知识表达方法和前述的方法相比主要区别有以下几点:(1)Agent 它本身是一个智能体,它自己有推测、反响、自学习和相互学习等能

力;(2)Agent 它本身就可以作为一个知识库或信息库来处理,贮存知识;(3)用 Agent 表示知识真正从本质上解决了智能化问题。

结论 本文抓住 Agent 的特点,研究 Agent 的知识表达度量理论,归纳起来作出如下贡献:

1. 给出了 Agent 知识量的概念,为进一步研究 Agent 的度量理论奠定了基础;

2. 利用熵这个概念,对 Agent 的知识进行量化处理,找到了 Agent、Agent 与 Agent 之间的量化理论依据;

3. 给出了联合表达最大熵定理,复合熵增定理,联合熵增定理等基本理论。

通过本文的研究,进一步丰富和发展了 Agent 系统理论的研究内容,为 Agent 走向实用化、商品化打下坚实的理论基础。

参考文献

- 1 Hu SHanli, et al. The logic tool used in the formalized research on agent. Computer Science, 1999, 26(12): 1~4
- 2 Davis R, Shroke H, Szlovits P. What is a knowledge representation? AI Magazine, 1993, spring: 17~33
- 3 Fu-Tong, et al. Knowledge Engineering. Beijing, Science Press, 1992
- 4 Dietterech T G. Exploratory research in machine learning. Machine Learning, 1990(5): 5~9
- 5 Li Fanzhang, et al. Principle and method and applications of artificial intelligence. Kunming, Yunnan Science Press, 1997

(上接第113页)

if (weight = heavy) & (comp = medium) then
(mileage = low)

表2是表1所导出的最大泛化规则集。

表2 一组最大泛化规则

make.model	weight	power	comp	trans	mileage
-	heavy	-	medium	-	low
USA	medium	high	-	-	medium
USA	medium	-	medium	-	medium
-	medium	-	-	auto	medium
USA	-	light	-	-	medium
-	heavy	-	high	-	medium
-	-	medium	high	manu	high
Japan	-	-	-	-	high
-	light	-	-	-	high

小结 发现分类规则是分类知识发现的核心任务。在基于 RS 理论的发发现方法中,可以从信息表的约简直接写出规则,并进一步求出最大泛化规则。在求最大泛化规则的过程中,尝试删除条件属性,在不降低规则的分类精确度的前提下使规则长度尽可能地短,而

删除条件属性的顺序由条件的属性有效度来决定。本文引入 J-measure 作为属性的有效度量,在开始删除操作前计算出属性的有效度并排序,可以生成较优化的最大泛化规则。

参考文献

- 1 Cercone N, Hamilton H, Hu X, Shan N. Data Mining Using Attribute-Oriented Generalization and Information Reduction. In: Rough Sets and Data Mining: Analysis of Imprecise Data Mining. Eds. Lin T Y and Cercone N. Kluwer Academic Publishers, Boston/London/Dordrecht, 1997. 199~227
- 2 Smyth P, Goodman R M. Rule Induction Using Information Theory. In: Knowledge Discovery in Databases. Eds. Piatetsky-Shapiro G and Frawley W J. Cambridge, MA: AAAI/MIT Press, 1991. 159~176
- 3 Smyth P, Goodman R M. An Information Theoretic Approach to Rule Induction from Databases. IEEE Transactions on Knowledge and Data Engineering, 1992, 4(4): 301~316
- 4 徐如燕. 基于粗糙集理论的知识发现研究:[硕士学位论文文] 空军雷达学院, 2000