近似推理的理论(Ⅲ)

L.A.Zadeh

6 推理法则和近似推理

像在其它逻辑中一样，在模糊逻辑中，一个命题q从一组前提$\{p_1,\cdots p_n\}$出发的演绎也是由推理法则支配的。所不同的是，在模糊逻辑中，前提和结论均允许是模糊命题。而且由于在再翻译的过程中使用了语言近似，所以从前提$p_1,\cdots p_n$得出的最终结论，一般是$p_1,\cdots p_n$的近似的而不是精确的结果。

在模糊逻辑中的基本推理法则是下列几个。

1.投影原理

令p是一个模糊命题，它的翻译表示成

$$p \to \Pi_{(X_1,\cdots X_n)} = F.$$

令$X_{(s)}$表示变量$X \triangleq (X_1,\cdots X_n)$的一个子变量，即

$$X_{(s)} = (X_{i_1},\cdots X_{i_k}) \qquad (6.1)$$

其中下标序列$s \triangleq (i_1,\cdots i_k)$是序列$(1,\cdots n)$的一个子序列。

令$\Pi_{X_{(s)}}$表示$X_{(s)}$的边缘可能性分布，即

$$\Pi_{X_{(s)}} = Proj_{U_{(s)}} F \qquad (6.2)$$

其中U_i，$i=1,\cdots n$是与X_i相关的论域；

$$U_{(s)} = U_{i_1} \times \cdots \times U_{i_k}, \qquad (6.3)$$

并且F在$U_{(s)}$上的投影是用下面的可能性分布函数来定义的，

$$\Pi_{X_{(s)}}(U_{i_1},\cdots U_{i_k}) = Sup_{U_{j_1},\cdots U_{j_m}} \mu_F(U_1,\cdots U_n) \qquad (6.4)$$

其中下标序列$s' \triangleq (j_1,\cdots j_m)$是$s$的补，$\mu_F$是$F$的隶属函数。

令q是下列可能性赋值方程的再翻译

$$\Pi_{X_{(s)}} = Proj_{U_{(s)}} F. \qquad (6.5)$$

则投影原理断言q可以从p推得。用一种示意图的形式，这个断言可以表达得更明显。

$$p \to \Pi_{(X_1,\cdots X_n)} = F \qquad (6.6)$$

$$q \leftarrow \Pi_{X_{(s)}} = Proj_{U_{(s)}} F.$$

投影原理的陈述，采取了$n=2$这种特别简单的形式。在这种情况下，把X_1，X_2，U_1，U_2分别写成X，Y，U，V，我们有：

$$p \to \Pi_{(X,Y)} = F \qquad (6.7)$$

$$q \leftarrow \Pi_X = Proj_U F \qquad (6.8)$$

对在V上的投影也照这样处理。

当$\Pi_{(X,Y)}$是正规模糊集的笛卡尔乘积时，我们就得到了(6.6)式的一种特殊情形。因而，如果

$$p \to \Pi_{(X,Y)} = G \times H \qquad (6.9)$$

则由p我们可推得q和r，其中

$$q \leftarrow \Pi_X = G \qquad (6.10)$$

且 $$r \leftarrow \Pi_Y = H. \qquad (6.11)$$

作为一个简单例子，如果

$p \triangleq$ John is tall and fat（约翰高而且胖）则由p我们可以推得

$q \triangleq$ John is tall（约翰是高的）

$r \triangleq$ John is fat（约翰是胖的）

2. 特指/合取原理

令p是一个模糊命题，它的翻译表示成

$$p \to \Pi_{(X_1,\cdots X_n)} = F,\quad F \subset U_1 \times \cdots \times U_n \qquad (6.12)$$

则由p我们可推得r，其中r是$\Pi_{(X_1,\cdots X_n)}$的一个特指的再翻译，即

$$r \leftarrow \Pi_{(X_1,\cdots X_n)}[\Pi_{X_{(s)}} = G] = F \cap \overline{G} \qquad (6.13)$$

其中$X_{(s)}$是X的一个子变量，$\overline{G}$是G的柱形扩张，$G \subset U$，$\Pi_{(X_1,\cdots X_n)}[\Pi_{X_{(s)}} = G]$表示对于$G$特指$X_{(s)}$产生的$n$元可能性分布。等价

地，特指原理可以表示成示意图的形式：

$$p \to \Pi_{(X_1, \ldots X_n)} = F \quad (6.14)$$

$$q \to \Pi_{(X_{i_1}, \ldots X_{i_k})} = G$$

$$\overline{\qquad\qquad\qquad}$$

$$r \leftarrow \Pi_{(X_1, \ldots X_n)} = F \cap \overline{G}.$$

对于$n=2$的特殊情形，特指原理可以简单地叙述为：从

$$p \triangleq (X, Y) \text{ is } F \quad (6.15)$$

和 $q \triangleq X$ is G

我们可推得

$$r \triangleq (X, Y) \text{ is } F \cap \overline{G}. \quad (6.16)$$

例如，从

$p \triangleq X$和Y近似地相等 (6.17)

和 $q \triangleq X$是小的

我们可推得（没有应用语音的近似）

$r \triangleq X$和Y是（近似地相等∩（小的×V）） (6.18)

如上所述，特指原理可以看成是稍微更一般化的合取原理的一种特殊情形。更明确说，假定$p \to \Pi^p_{(Y_1, \ldots Y_k, X_{k+1}, \ldots X_n)} = F$

(6.19)

$$q \to \pi^q_{(Y_1, Y_k, \ldots Z_{k+1}, \ldots Z_m)} = G \quad (6.20)$$

其中$Y_1, \cdots Y_k$是在Π^p和Π^q中都出现的变量，U_i，V_j和W_l是与X_i，Y_j，Z_l相关的论域；令S是U_i，V_j和W_l的包含笛卡尔积$V_1 \times \cdots \times V_k \times U_{k+1} \times \cdots \times U_n$和$V_1 \times \cdots \times V_k \times W_{k+1} \times \cdots \times W_m$的最小的笛卡尔积；并且令$\overline{F}$和$\overline{G}$分别是在$S$中的$F$和$G$的圆柱形扩张。则由$p$和$q$我们可推得$r$，（用示意图形式）其中

$$p \to \Pi^p_{(Y, X)} = F \quad (6.21)$$

$$q \to \Pi^q_{(Y, Z)} = G$$

$$\overline{\qquad\qquad\qquad}$$

$$r \leftarrow \Pi_{(X, Y, Z)} = \overline{F} \cap \overline{G}$$

$Y \triangleq (Y_1, \cdots Y_k)$，$X \triangleq (X_{k+1}, \cdots X_n)$和$Z \triangleq (Z_{k+1}, \cdots Z_m)$。

（6.21）式的一个特殊但很重要的情形，是在取$n=3$，$k=1$时得到的，后面我们将要用到。在这种情形下，（6.21）式可表示成

$$p \to \Pi^p_{(X, Y)} = F$$

$$q \to \Pi^q_{(Y, Z)} = G$$

$$\overline{\qquad\qquad\qquad}$$

$$r \leftarrow \Pi_{(X, Y, Z)} = (F \times W) \cap (U \times G).$$

虽然特指原理是包含在合取原理之中的，但是由于特指原理较简单，使用的更频繁，而且它给人的印象更直观一些。因此，在大多数应用中，我们使用特指/合取原理这个名字指的是特指原理，在某些应用中，指的是合取原理。应当注意，在谓词逻辑中（Lyndon 1966），这个原理包含一般化法则

3．必含原理

非正规地说，必含原理断言，由任意的模糊命题p，我们可推得模糊命题q，如果由p引导的可能性分布包含在由q引导的可能性分布之中。因此，示意地我们有：

$$p \to \Pi_{(X_1, \cdots X_n)} = F \quad (6.23)$$

$$\downarrow$$

$$q \leftarrow \Pi_{(X_1, \cdots X_n)} = G \supset F.$$

例如，从$p \triangleq$“X是非常大的”，我们可推得$q \triangleq$“X是大的”。

推理合成法则 一般地，上述的推理法则是按顺序地或者结合地使用。一种特别有效的结合是先应用特指/合取原理，然后又接着应用投影原理。这个结合被称作推理合成法则（Zadeh 1973）。正像我们在后面将要看到的那样，推理合成法则包括广义假言推理作为一种特殊情形。

为了我们的目的，将推理合成法则表示成示意图的形式是很方便的。

$$p \to \Pi_{(X, Y)} = F \quad (6.24)$$

$$q \to \Pi_{(Y, Z)} = G$$

$$- - - - - - - - -$$

$$r \leftarrow \Pi_{(X, Z)} = F \cdot G$$

其中X，Y和Z分别是在U，V和W中取值的变量，F是$U \times V$的一个模糊子集，G是$V \times W$的一个模糊子集，并且$F \cdot G$是F和G的合成，定义为，

$$\mu_{F \cdot G}(u, w) = Sup_v[\mu_{F(u, v)} \wedge \mu_{G(v, w)}] \quad (6.25)$$

其中$u \in U$，$v \in V$，$w \in W$；μ_F和μ_G分别是F和G的隶属函数；虚线表示，由于在再翻译时使用了语言近似，因此r一般是p和q的一个近似的而不是精确的结果。另外应当注意，推理合成法则类似于，由

X的概率分布和给定X的情况下Y的条件概率分布可产生Y的概率分布的法则。

容易证明，推理合成法则可以看成是，应用特指/合取原理，之后接着应用投影原理的一个结果。更明确说，由于对（6.24）应用（6.21），我们得到：

$$p \to \Pi_{(X,Y)} = F \quad (6.26)$$
$$q \to \Pi_{(Y,Z)} = G$$
$$\overline{\qquad\qquad\qquad\qquad\qquad}$$
$$S \leftarrow \Pi_{(X,Y,Z)} = (F \times W) \cap (U \times G)$$

其中$\mu_{(F \times W) \cap (U \times G)}(u, v, w) = \mu_F(u, v) \wedge \mu_G(v, w)$ （6.27）

然后再对S应用投影原理和把$\Pi_{(X,Y,Z)}$投影到$U \times W$上，我们有

$$Proj_{U \times W}[(F \times W) \cap (U \times G)] = \int_{U \times W} Sup_v [\mu_F(u, v) \wedge \mu_G(v, w)] / (u, w) \quad (6.28)$$

将它与（6.25）比较表明，所得命题与（6.24）是一致的，并且可表示成

$$r \leftarrow \Pi_{(X,Z)} = F \cdot G \quad (6.29)$$

当p和q具有下述形式时，即$p \triangleq X$ is F，$q \triangleq$ If X is G then Y is H，我们得到了推理合成法则的一个重要的特殊情形，在这种情形下，由（4.20）和（6.24）产生了合成假言推理：

$$p \to \Pi_X = F \quad (6.30)$$
$$\underline{q \to \Pi_{(X,Y)} = \overline{G'} \oplus \overline{H}}$$
$$r \leftarrow F \cdot (\overline{G'} \oplus \overline{H})$$

它可以看成是推广的古典假言推理。它的一个特殊情况是在F，G和H是非模糊的，且$F = G$时得到的，这时（6.30）就简化为

$$p \to \Pi_X = F \quad (6.31)$$
$$\underline{q \to \Pi_{(X,Y)} = \overline{F'} \oplus \overline{H}}$$
$$r \leftarrow \Pi_Y = F \cdot (\overline{F'} \oplus \overline{H})$$

而且由于

$$F \cdot (\overline{F'} \oplus \overline{H}) = H$$

因此就有$r \leftrightarrow Y$ is H

它表明由$p \triangleq X$ is F和$q \triangleq$ If X is F then Y is H，我们可推得$r \triangleq Y$ is H，它与假言推理的陈述是一致的。

上面介绍的推理法则，为在解答问题和用模糊命题进行推理时使用近似推理提供了一个基础。下面我们将通过对几个典型的问题应用这些法则，来说明基于这些法则的方法是如何使用的。

语义等价 举一个简单例子，假定由前提

$p \triangleq$ Ellen is vnot ery tall

我们希望推出"Is Ellen tall? τ"这个问题的答案，其中符号? τ表示期望对问题的回答具有如下形式

$q \triangleq$ Ellen is tall is τ

其中τ是一个语言的真值。

为了得到这个问题的答案，我们将要求p和q是语义等价的，这隐含着由p引导的可能性分布等于由q引导的可能性分布。

因此，通过使用翻译法则（4.5），（4.6）和（4.56），我们得到

$$\text{Ellen is not very tall} \to \pi_{Height(Ellen)}(u) = 1 - \mu^2_{tall}(u) \quad (6.32)$$

$$\text{Ellen is tall is } \tau \to \pi_{Height(Ellen)}(u) = \mu_\tau[\mu_{tall}(u)] \quad (6.33)$$

其中假定μ_{tall}是给定的tall的隶属函数。那么，由（6.32）和（6.33）则有想要的隶属函数μ_τ满足恒等式

$$1 - \mu^2_{tall}(u) \equiv \mu_\tau[\mu_{tall}(u)], \ u \in [0, 200] \quad (6.34)$$

由它我们可立即断定，μ_τ是由下式给定的（见图5）

$$\mu_\tau(v) = 1 - v^2 \quad (6.35)$$

它的一个粗略的语言的近似可表示成

$$\tau \approx \text{not very true} \quad (6.36)$$

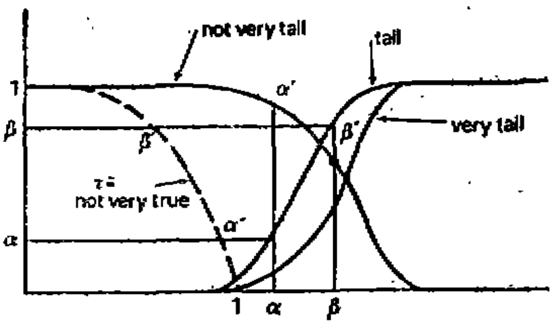

图5 利用语义等价求出的答案

通过接连使用分管否定，真值限定和修饰的应用的法则，可以得到相同的结果，这是很有启发性的。这样，我们可以断言：

John is not very tall

↔John is not very tall is u-true

（利用（4.60））

John is not very tall is μ-true

$\leftrightarrow$John is very tall is ant (u-ture)
(利用(5.21))

John is very tall is ant (u-true)

$\leftrightarrow$John is tall is$^{\frac{1}{2}}$ (ant(u-true))
(利用(5.23))

它隐含着

$$\tau = {}^{\frac{1}{2}}\ (ant(u\text{-}true)) \qquad (6.37)$$

即τ是 ($ant(u$-$true$)) 的"左平方根"，因此

$$\mu_{u\text{-}true}(v) = v \qquad (6.38)$$

我们有$\mu_\tau = 1 - v^2$ (6.39)

与 (6.35) 是一致的。

语义必含 假定我们希望从前提

$p \triangleq$ Most Swedes are tall

推出问题"How many Swedes are very tall?"的答案。

利用 (4.40) 翻译p，我们有

Most Swedes are tall $\rightarrow \pi_p(\rho) = \mu_{most}[\int_0^{200} \rho(u)\mu_{tall}(u)du]$ (6.40)

其中$\rho(u)du$是高度在区间$[u, u+du]$内的瑞典人的比例，π_p是ρ的可能性分布函数。（注意高度是以厘米为表示单位的。）

接着利用 (4.30) 个子是很高的瑞典人的比例是由

$$\gamma = \int_0^{200}\rho(u)\mu^2_{tall}(u)du \qquad (6.41)$$

给定的。因此，我们的问题是从 (6.40) 的右部给出的关于ρ的可能性分布的知识，找到γ的可能性分布。用一个变分原理（它由(4.48)可得），这个问题可表示成

$$\pi(\gamma) = Max_\rho \mu_{most}[\int_0^{200}\rho(u)\mu_{tall}(u)du] \qquad (6.42)$$

满足$\gamma = \int_0^{200}\rho(u)\mu^2_{tall}(u)du$ (6.43)

要找出这个问题的最大的ρ，它具有

$$\rho(u) = \delta\ (u - \alpha) \qquad (6.44)$$

这种形式（Baliman和Zadeh 1976）。其中δ是一个δ函数，α是区间[0，200]中的一个点。δ函数密度意指所有那些在该属性方面具有相同的值的成员。这样，由 (6.43) 我们有

$$\gamma = \mu^2_{tall}(\alpha) \qquad (6.45)$$

因此$\pi(\gamma) = \mu_{most}(\mu_{tall}(\alpha))$ (6.46)

$$= \mu_{most}(\sqrt{\gamma})$$

或等价地（见(5.30)）

$$\pi(\gamma) = \mu_{{}^2most}(\gamma) \qquad (6.47)$$

因此对"How many Swedes are very tall?"这个问题的期求的回答是（见图6）

$q \triangleq {}^2$most Swedes are very tall。 (6.48)

为了证实p语义必含q，我们注意

$$q \leftrightarrow \pi_q(\rho) = \mu_{{}^2most}[\int_0^{200}\rho(u)\mu^2_{tall}(u)du] \qquad (6.49)$$

$$= \mu_{most}[\sqrt{\int_0^{200}\rho(u)\mu^2_{tall}(u)du}]$$

但是利用Schwarz的不等式

$$\int_0^{200}\rho(u)\mu_{tall}(u)du \leqslant \sqrt{\int_0^{200}\rho(u)\mu^2_{tall}(u)du}, \qquad (6.50)$$

并且由于μ_{most}是单调函数，那么

$\pi_p(\rho) \leqslant \pi_q(\rho)$ 对所有的ρ和μ_{tall}

它隐含着p强语义必含q

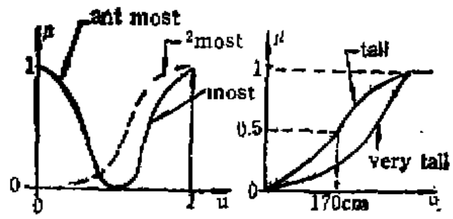

图6 关于most，tall以及它们的修饰词的表示

特指和投影原理 下面我们通过解决一个简单的问题来说明特指和投影原理的应用

假定前提是：

$p \triangleq$ John is very big

$q \triangleq$ John is very tall

其中big是$U \times V$的一个给定的模糊子集（即$Height$的值(以厘米为单位($\times Weight$的值(以公斤为单位))，tall是U的一个给定的模糊子集。问题是"What is John's weight?"

让我们假定对这个问题的回答是具有$r \triangleq$ John is w这种形式，其中w是John的体重的语言值（如重($heavy$)，很重($very\ heavy$)，不很重($not\ very\ heavy$)等等）则通过使用翻译法则 (4.6)，特指原理和投影原理，我们得到了再翻译关系

$$r \leftarrow Proj_{u \times Weight}BIG^2[\Pi_{Height} = TALL] \qquad (6.51)$$

它表示了对提出的问题的回答。

更具体地说，假定模糊集合BIG，$TALL$，$HEAVY$是用表格的形式来定义的。（这里列出的是不完全的表，见下表。）

由于将这些表代入(6.51)，我们就得到了属性$Weight$的可能性分布的一个近似的形式

$\Pi_{Weight}=0.5/60+0.7/65+0.8/70+0.9/75+1/80$ (6.52)

根据再翻译（以及语言的近似）产生的回答是：

$r\triangleq$ John is very heavy。

BIG	Height	Weight	μ
	165	60	0.5
	170	60	0.6
	175	60	0.7
	170	65	0.75
	175	65	0.8
	180	65	0.85
	170	70	0.8
	175	70	0.85
	180	70	0.9
	170	75	0.85
	175	75	0.9
	180	75	0.95
	180	80	1

TALL	Height	μ
	165	0.7
	1.70	0.8
	175	0.9
	180	1
	185	1

HEAVY	Weight	μ
	60	0.7
	65	0.8
	70	0.9
	75	0.95
	80	1
	85	1

作为另外的例子，考虑下面的前提

$p\triangleq$ Romy lives near a small city

$q\triangleq$ Arnold lives near Romy

由它们我们推出"Where does Arnald live？"这个问题的答案。

假定在p和q中记录的关系具有下面列出的框架，

$NEAR_p$‖City1｜City2｜μ｜　　$NEAR_q$‖City1｜City2｜μ｜

SMALL CITY‖City｜μ｜

其中$NEAR_p$和$NEAR_q$分别指的是在p和q中 NEAR 关系。按照这些关系，p和q的翻译可表示成

$$p\rightarrow\Pi_{Location(Residence(Romy))} \tag{6.53}$$

$$=Proj_{\mu\times city1}NEAR_p[\Pi_{city2}=SMALL\ CITY]$$

$$q\rightarrow\Pi_{(Location(Residence(Romy)),\ Location(Residence(Arnald))} \tag{6.54}$$

$$=NEAR_q$$

将（6.53）代入（6.54）并投影到属性$Location(Residence(Arnald))$上，我们得到

$$r\leftarrow Proj_{\mu\times city2}NEAR_q[\Pi_{city1}=Proj_{\mu\times city1}NEAR_p[\pi_{city2}=SMALL\ CITY]] \tag{6.55}$$

它作为这个提问的回答的表达式，

推理合成法则　当前提中涉及的变量是分布在有限集合上或者可以用分布在这种集合上的变量近似地表示的时候，应用推理合成法则是特别方便的。

作为一个简单例子，考虑前提

$p\triangleq X$ is small

$q\triangleq X$ and Y are approximately equal（X和Y近似相等）

其中X和Y是分布在集合$U=1+2+3+4$上，*small* *approximately equal*定义为

$$small=1/1+0.6/2+0.2/3$$

$$approximately\ equal=1/[(1,1)+(2,2)+(3,3)+(4,4)]+0.5/[(1,2)+(2,1)+(2,3)+(3,2)+(3,4)+(4,3)]。$$

按照这些集合，p和q的翻译可表示成

$$p\rightarrow\Pi_X=small \tag{6.56}$$

$$q\rightarrow\Pi_{(X,Y)}=approximately\ equal$$

并且由p和q我们可推得r，其中

$$r\leftarrow\Pi_Y=\pi_{(X,Y)}=small\circ approximately\ equal \tag{6.57}$$

small 和*approximately equal* 的合成，可以通过计算*small*和*approximately equal*的相应的关系矩阵的 maxmin 乘积很快地完成。这样，我们得到：

$$[1\ \ 0.6\ \ 0.2\ \ 0]\circ\begin{pmatrix}1 & 0.5 & 0 & 0\\0.5 & 1 & 0.5 & 0\\0 & 0.5 & 1 & 0.5\\0 & 0 & 0.5 & 1\end{pmatrix}$$

$$=[1\ \ 0.6\ \ 0.5\ \ 0.2]$$

它隐含着

$$\Pi_Y=\Pi_X\cdot\ \Pi_{(X,Y)}=1/1+0.6/2+0.5/3+0.2/4$$

根据再翻译产生的语言近似

$$r\triangleq Y \text{ is more or less small} \leftarrow \Pi_Y=1/1+0.6/2+0.5/3+0.2/4。$$

因此由$p\triangleq X$ is small和$q\triangleq X$ and Y are approximately equal，我们可近似地推得，$r\triangleq Y$ is more or less small。

作为合成假言推理的一个简单例子，像在Bellman和Zadeh(1976)的文章中一样，假定

$U=V=1+2+3+4$

$F=0.2/2+0.6/3+1/4$

$G=0.6/2+1/3+0.5/4$

$H=1/2+0.6/3+0.2/4$

有 $p\triangleq X \text{ is } F\rightarrow\Pi_X=F$ (6.58)

$q\triangleq \text{If } X \text{ is } G \text{ then } Y \text{ is } H\rightarrow\Pi_{(X,Y)}=\overline{G'}\oplus\overline{H}$

且 $r\leftarrow\Pi_Y=F\cdot(\overline{G'}\oplus\overline{H})$。

在这种情况下，

$$\overline{G'}\oplus\overline{H}=\begin{pmatrix}1 & 1 & 1 & 1\\ 0.4 & 1 & 1 & 0.6\\ 0 & 1 & 0.6 & 0.8\\ 0.5 & 1 & 1 & 0.7\end{pmatrix}$$

$$和\overline{F}\circ(\overline{G'}\oplus\overline{H})=[0\ 0.2\ 0.6\ 1]\cdot\begin{pmatrix}1 & 1 & 1 & 1\\ 0.4 & 1 & 1 & 0.6\\ 0 & 1 & 0.6 & 0.6\\ 0.5 & 1 & 1 & 0.7\end{pmatrix}$$

$$=[0.5\ 1\ 1\ 0.7]，$$

因此由p和q我们可推得

$r\triangleq Y$ is $0.5/1+1/2+1/3+0.7/4$。

上面这个例子仅仅是为了说明，当X和Y是分布在有限集合上的时候，应用合成假言推理是如何进行计算的。有关推理合成法则在所谓的模糊逻辑控制器的设计中的实际应用方面的详细讨论，可在Mamdani和Assilian(1975)，Mamdani(1976, a, b)，Kickert和van Nauta Lemke(1976)，Rutherford和Bleore(1975)和文献目录中列出的其它的文章中找到。

7 结束语

本文概要地阐述了近似推理的理论。不精确性是现实世界中普遍存在的一种现象，而近似推理的理论可以看成是为处理这种现象所做的一种尝试。

正因为近似推理是基于模糊逻辑的，所以它缺少精确推理的深度与普遍性。然而能够充分地证明，对于人文研究的系统中存在的复杂性和不定性的问题的理解与处理方面，它比精确推理更有效。因此，它能够对智能系统的设想与开发有所贡献。而这种系统能够在不确定性和不精确性存在的情况下，作出合理的决策方面，接近人的智能水平。

参考文献（略）

〔金雅芬译自《Machine Intelligence》(J.E.Hayes, D.Michie, and L.I.Mikulich, eds) Vol.9, Elsevier, New York, pp.149-194, 1979，声 钟校〕