近似推理的理论（I）

L. A. Zadeh

摘要

本文中近似推理的理论阐述的是从一组不精确前提推演出可能是不精确的结论。

近似推理的理论是建立在模糊逻辑（FL）基础上的。它的真值是语言性的，即是像真，不真，很真，或多或少为真，假，不很假等这种形式的真值。推理的法则是近似的而不是精确的，并且假定前提采取下列模糊命题的形式。例如："(X is much smaller than Y) is quite true"，" If X is small is possible then Y is very large is very likely。"等。通过使用可能性分布（而不是概率分布)的概念，这种命题被翻译成PRUF的表达式。PRUF是(Possibilistic Relational Universal Fuzzy）的缩写，它是一个自然语言的意义表示语言。

一个PRUF的表达式是一个过程，它是用于计算由一个用自然语言表示的命题引入的可能性分布。通过对这种分布应用PRUF的推理法则，我们得到其它的可能性分布，它们是从模糊前提出发演绎得到的结论，然后对它们进行再翻译，用语言近似地表示出来。

模糊逻辑推理的基本法则是投影原理，特指/合取原理和必含原理。最后通过一些例子来说明这些法则在近似推理中的应用。

1 引言

非形式地说，所谓近似推理或者等价地称为模糊推理，是这样一种方法或过程，用这种方法或通过这一过程，我们可以从一组不精确的前提出发，导出一个可能是不精确的结论。这种推理大部分是定性的，而不是定量的，而且几乎全部落到了古典逻辑的适用范围之外。关于模糊推理的基础的全面阐述，可以在Gaines的（1976 a，b，c）文章中找到。

人所具有的一些显著的特点，如能够理解自然语言，辨认潦草的手写文字，下棋，还有人能够在复杂的、不确定的条件下作出合理的决策，这些活动都需要动脑筋思考和物理的技能，而近似推理是这种能力的基础。事实上，用定性的，不精确的术语进行推理的能力是区分人的智能与机器智能的一个重要标志。近似推理还没有在心理学、哲学、逻辑、人工智能和认知科学的其它分支中引起重视，主要原因在于它与人们在科学研究中使用的传统的精确推理方法很不一致，两者之间有一条很深的壕沟。同时它与人们广泛持有的——精确、定量的推理能够解决在人文研究的系统分析中遇到的极其复杂和不定的问题——这种观念相违背。

在早期文章（Zadeh 1973，1975a，b，c，1976，1977a，b）中，已经概括地给出了近似推理这个概念的框架，它是建立在语言变量和模糊逻辑的概念基础上的。这篇文章有一个新颖之处，就是后面将要介绍的，它引入了一个可能性分布的概念。（见Zadeh 1977）随后我们将会看到，可能性分布的概念为表示用自然语言表达的命题的意义，提供了一个自然的基础。并以此作为出发点，把不精确的前提翻译成PRUF语言的表达式，而且对它还可以应用与这个语言相关的推理法则。

介绍近似推理的理论之前，首先简单讨论一下可能性分布的概念以及这个概念在翻译用自然语

言形式表达的模糊命题的工作中所起的作用。第三节引入了语言变量的概念，作为近似地表示变量的值及变量之间的相互关系的方法。第四节和第五节讨论作为近似推理的基础的模糊逻辑的一些基本内容，并且引入了语义等价和语义必含的概念。最后在第六节，将列出在模糊逻辑中的基本推理法则，并且通过一些简单例子说明它们在近似推理中的应用。

2 可能性分布的概念

在讨论近似推理之前，有一个基本的假定就是，不精确性是自然语言固有的性质，而且这种不精确性主要是可能性的，而不是概率的。术语可能性的（possibilistic)是由B．R．Gaines和L．J．Kohout在他们有关可能的自动机方面的文章中创造的。(1975)

为了说明上述观点，考虑命题$p \triangleq X$ is an integer in the interval[0, 8]。即X是闭区间［0，8］中的一个整数。符号≜代表“定义为”或“表示”。显然，命题中的X不是与这个区间里的一个唯一的整数相关，它表示，在[0，8]这个区间里的任何一个整数都可能是X的一个值，而不在这个区间里的任何整数都不可能是X的值。

根据这个明显的事实，提出了下面的关于p的解释。命题“X is an integer in the interval[0, 8]”引入了一个可能性分布Π_X，它把每个整数n与它可能是X的一个值的可能性相联系。因此对上述的命题有

$Poss\{X=n\}=1$　对$0\leqslant n\leqslant 8$

和　$Poss\{X=n\}=0$　对$n<0$或$n>8$．

其中$Poss\{X=n\}$是“X取值为n的可能性”的缩写形式。注意由p引入可能性分布是均匀的，是指在区间[0，8]里的n是X的值的可能性均等于1，这个区间之外的n是X的值的可能性均等于0。

下面来看模糊命题 $q \triangleq X$ is a small integer，即X是个小的整数．小的整数是个模糊集合，定义为：

$$\text{small integer}=1/0+1/1+0.8/2+0.6/3+0.4/4+0.2/5 \tag{2.1}$$

这里＋表示并而不是算术和，一个具有形式0.8/2的个体表示整数2在模糊子集small integer中的隶属度是0.8。（见A．Kaufmann(1975)，C．V．Negoita and D．Ralescu(1975)和L．A．Zadeh，K．S．Fu，K．Tanaka和M．Shimura(1975))

作为对非模糊命题p的解释的一个推广，我们现在这样来解释q。命题$q \triangleq X$ is a small integer引入了一个可能性分布Π_X，它使X取一个值n的可能性等于在模糊子集small integer 中的n的隶属度。因而

$Poss\{X=0\}=1$

$Poss\{X=2\}=0.8$

$Poss\{X=5\}=0.2$

$Poss\{X=6\}=0$

更一般地，对一个具有$p \triangleq X$ is F这种形式的模糊命题，（其中X是一个在论域U中取值的变量，F是U的一个模糊子集。）引入一个可能性分布Π_X，它等于F，即

$$\Pi_X=F. \tag{2.2}$$

因此，在本质上X的可能性分布是一个模糊集合，它用来定义在U中X可能取任何指定的值的可能性。更具体地说，如果$u\in U$且$\mu_F:U\to[0, 1]$是F的隶属函数，那么给定“X is F”$X=u$的可能性是

$$Poss\{X=u \mid X \text{ is } F\}=\mu_F(u),\quad u\in U. \tag{2.3}$$

由于可能性分布的概念与模糊集合的概念一致，因此我们可以用那些决定模糊集合操作的规则，对可能性分布进行操作。更具体地，就是利用模糊约束进行操作。一个模糊约束是一个模糊集合，可以把它当作是关于可能赋予一个变量哪些值的一个弹性约束。一个变量如果它与一个模糊约束相关联，或者等价地与一个可能性分布相联系，则它是一个模糊变量。下面我们将集中讨论与近似推理有关的可能性分布的一些问题。

可能性与概率

可能性与概率之间的区别是什么？在

观上，可能性是与我们对可实行的程度和达到的难易程度的感觉有关，而概率是与信任度，似然性，频率或比例有关的。有这样一句话，可能的事情也许机会不大，而机会不大的事情未必是不可能。这个关系的一个更具体的说明包含在可能性/概率一致性原理之中（Zadeh，1977a）。然而可能性与概率之间更重要的区别在于：在求并运算下，决定它们的组合方式的规则是不同的。更明确地，如果A是U的一个非模糊子集，而Π_X是由命题"X is F"引入的可能性分布，那么A的可能性测度$\Pi(A)$定义为

$$\Pi(A) \triangleq \mathrm{Poss}\{X \in A\} \triangleq Sup_{u \in A} \mu_F(u). \quad (2.4)$$

用（2.4）定义的可能性测度是在Sugeno（1974）和Terano与Sugeno（1975）这两篇文章定义的模糊测度这个更一般概念的一个特殊情形。更一般地，如果A是U的一个模糊子集，则

$$\Pi(A) \triangleq \mathrm{Poss}\{X \text{ is } A\} \triangleq Sup_u[\mu_F(u) \wedge \mu_A(u)] \quad (2.5)$$

其中μ_A是A的隶属函数，$\wedge \triangleq min$。

由可能性测度的定义，马上可得，对U的任意子集A和B，由下式给出A和B的并的可能性测度：

$$\Pi(A \cup B) = \Pi(A) \vee \Pi(B) \quad (2.6)$$

其中$\vee \triangleq max$。因此，可能性测度不具有概率测度的基本的可加的性质，即如果A和B是不相交的，则

$$P(A \cup B) = P(A) + P(B) \quad (2.7)$$

其中$P(A)$和$P(B)$分别表示A和B的概率测度。

与概率不同的是，可能性的概念无法包含重复试验的思想。因此可能性的概念本身不是统计性的，它是在我们所要研究的现象中，存在不精确性或不确定性而用统计的方法分析和描述又不允许时，使用的一个自然的概念。

可能性赋值方程

为什么可能性分布的概念在近似推理中起这么重要的作用？原因在于，我们假定一个自然语言的命题，可以解释成一个模糊集合对一个可能性分布的赋值。更明确说，如果p是一个自然语言的命题。我们将说p翻译成一个可能性赋值方程：

$$p \to \Pi(x_1, \cdots x_n) = F \quad (2.8)$$

其中X_1，$\cdots X_n$是p中明确地包含或隐含着的变量；$\Pi(x_1, \cdots x_n)$是n元变量$X \triangleq (X_1, \cdots X_n)$的可能性分布；$F$是一个模糊关系，就是笛卡尔积$U_1 \times \cdots \times U_n$的一个模糊子集，其中$U_i$，$i=1, \cdots n$，是与$X_i$相关联的论域。在这种情况下，可能性赋值方程

$$\Pi(x_1, \cdots x_n) = F \quad (2.9)$$

被称作是p的翻译，相反地，称p是（2.9）式的再翻译，在这种情况下，它与（2.9）的关系将表示成

$$p \longleftarrow \Pi(x_1, \cdots x_n) = F. \quad (2.10)$$

一般地，一个形式为$p \triangleq X$ is F的命题，其中X是一个对象的名字或者一个命题，可翻译成

$$p \longrightarrow \Pi_X = F \quad (2.11)$$

而翻译成

$$p \longrightarrow \Pi_{A(X)} = F \quad (2.12)$$

其中$A(X)$是一个X的暗示的属性。例如，

$$\text{Joe is young} \to \Pi_{Age(Joe)} = young \quad (2.13)$$

$$\text{Maria is blond} \to \Pi_{colour(Hair(Maria))} = blond \quad (2.14)$$

$$\text{Max is about as tall as Jim} \to \Pi_{(Height(Max), Height(Jim))} = approximately\text{-}equal \quad (2.15)$$

其中$young$，$blond$和approximately-equal分别是各自论域上的一元的和二元的模糊关系。更具体说，如果u是Joe的年龄的一个数字的值，那么（2.13）蕴含

$$Poss\{Age)(Jce) = u\} = \mu_{young}(u) \quad (2.16)$$

类似地，如果u是头发颜色识别的标号，那么（2.14）蕴含

$$Poss\{colour(Hair(Maria)) = u\} = \mu_{blond}(u), \quad (2.17)$$

而（2.15）表示

$$Poss\{Height(Max) = u, \ Height(Jim) =$$

$v) = \mu_{APPROXIMATELY.EQUAL}(u, v)$ (2.18)

其中u和v分别是变量$Height(Vera)$和$Height(Jim)$的隶属值。

投影和特指

对可能性分布可以施加的操作中，有两个操作：投影和特指，是与近似推理特别相关的。

令$\Pi_{(X_1, \cdots X_n)}$表示一个n元的可能性分布，它是一个$U_1 \times \cdots \times U_n$的模糊关系，具有$\Pi_{(X_1, \cdots X_n)}$的可能性分布函数（即$\Pi_{(X_1, \cdots X_n)}$的隶属函数）用$\pi_{(X_1, \cdots X_n)}$或更简单地用$\pi_X$表示。

令$s \triangleq (i_1, \cdots i_k)$是下标序列$(1, \cdots n)$的一个子序列，并且令$s'$代表补子序列，$s' \triangleq (j_1, \cdots j_m)$。例如对$n=5$，$s=(1,3,4)$和$s'=(2,5)$。依据这种序列，一个具有形式$(A_{i_1}, \cdots A_{i_k})$的$k$元组可以用一个缩写形式$A_{(s)}$来表示。特别地，变量$X_{(s)} = (X_{i_1}, \cdots X_{i_k})$可以看成是$X \triangleq (X_1, \cdots X_n)$的$k$元子变量，它的补$X_{(s')} = (X_{j_1}, \cdots X_{j_m})$也是子变量。

$\Pi_{(X_1, \cdots X_n)}$在$U_{(s)} \triangleq U_{i_1} \times \cdots \times U_{i_k}$上的投影是一个$k$元的可能性分布，表示成

$$\Pi_{X(s)} \triangleq Proj_{U_{(s)}} \Pi_{(X_1, \cdots X_n)} \quad (2.19)$$

并且定义为

$$\pi_{X_{(s)}}(U_{(s)}) \triangleq Sup_{U_{(s')}} \pi_X(u_1, \cdots u_n) \quad (2.20)$$

其中$\pi_{X_{(s)}}$是$\Pi_{X_{(s)}}$的可能性分布函数。例如，对$n=2$，

$$\pi_{X_1}(u_1) \triangleq Sup_{u_2} \pi_{(X_1, X_2)}(u_1, u_2)$$

是$\Pi_{(X_1, X_2)}$在U_1上的投影的可能性分布函数的表达式。由于它与一个边缘概率分布的概念类似，$\Pi_{X_{(s)}}$可以看成为一个边缘可能性分布。注意在（2.21）之前我们用（2.19）中$\Pi_{X_{(s)}}$表示Π_X在$U_{(s)}$上的投影。

关于边缘可能性分布概念的重要性，是从可以把$\Pi_{X_{(s)}}$看成是子变量$X_{(s)}$的可能性分布这一事实得出的。因此正像投影原理（见第6节）阐述的那样，$X_{(s)}$和$\Pi_{X_{(s)}}$之间的关系可以表述如下。

从变量$X \triangleq (X_1, \cdots X_n)$的可能性分布$\Pi_{(X_1, \cdots X_n)}$，子变量$X_{(s)} \triangleq (X_{i_1}, \cdots X_{i_k})$的可能性分布$\Pi_{X_{(s)}}$可以从投影$\Pi_{(X_1, \cdots X_n)}$到$U_{(s)}$上推得。就是，

$$\Pi_{X_{(s)}} = Proj_{U_{(s)}} \Pi_{(X_1, \cdots X_n)}. \quad (2.21)$$

作为一个简单的例子，假定$n=3$，$U_1 = U_2 = U_3 = a+b$更习惯地为$\{a,b\}$，且$\Pi_{(X_1, X_2, X_3)}$表达成一个线性的形式

$$\Pi_{(X_1, X_2, X_3)} = 0.8aaa + 1aab + 0.6baa + 0.2bab + 0.5bbb \quad (2.22)$$

其中具有$0.6baa$这种形式的项表示的意义为

$$Poss\{X_1 = b, X_2 = a, X_3 = a\} = 0.6 \quad (2.23)$$

为了从（2.22）导出$\Pi_{(X_1, X_2)}$，只要将（2.22）式的每一项中X_3的值用空串Λ代替，这样产生

$$\Pi_{(X_1, X_2)} = 0.8aa + 1aa + 0.6ba + 0.2ba + 0.5bb = 1aa + 0.6ba + 0.5bb \quad (2.24)$$

类似地

$$\Pi_{X_1} = 1a + 0.6b + 0.5b = 1a + 0.6b. \quad (2.25)$$

接着我们讨论特指操作，令$\Pi_{(X_1, \cdots X_n)} = F$表示$X = (X_1, \cdots, X_n)$的可能性分布，并且令$\Pi_{X_{(s)}} = G$表示子变量$X_{(s)} = (X_{i_1}, \cdots X_{i_k})$的一个指定的可能性分布（并不要求必须是边缘分布）。

非形式地，所谓$\Pi_{(X_1, \cdots X_n)}$的特指，其含义就是通过规定$\Pi_{X_{(s)}}$的可能性分布是G而产生的对$\Pi_{(X_1, \cdots X_n)}$的修正。更具体说，

$$\Pi_{(X_1, \cdots X_n)}[\Pi_{X_{(s)}} = G] \triangleq F \cap \overline{G} \quad (2.26)$$

其中左部很明显为X_i（即为属性），它们在$\Pi_{(X_1, \cdots X_n)}$中是被特指的，而右部定义的是特指的结果，用$\overline{G}$表示G的柱形扩张，就是在$U_1 \times \cdots \times U_n$上的柱状模糊集合，它在$U_{(s)}$上的投影是$G$。因此，

$$\mu_{\overline{G}}(u_1, \cdots u_n) \triangleq \mu_G(u_{i_1}, \cdots u_{i_k}), \quad (2.27)$$

$$(u_1, \cdots u_n) \in U_1 \times \cdots \times U_n.$$

作为一个简单的例子，考虑由（2.22）定义的可能性分布，并且假定

$$\Pi_{(X_1, X_2)} = 0.4ca + 0.9ba + 0.1bb$$

$$(2.28)$$

在这种情况下，

$$\overline{G}=0.4aaa+0.4aab+0.9baa+0.9bab+0.1bba+0.1bbb$$

$$F\cap\overline{G}=0.4aaa+0.4aab+0.6baa+0.2bab+0.1bbb$$

因此

$$\Pi_{(x_1,x_2,x_3)}[\Pi_{(x_1,x_2)}=G]=0.4aaa+0.4aab+0.6baa+0.2bab+0.1bbb \quad (2.29)$$

一般在一个特指的可能性分布（或一个模糊关系）中的一些变量在其各自的论域中被赋予固定值，而其它的变量与可能性分布相联系。例如，在一个模糊关系的情况下，该关系表示的是高个子，金发碧眼，名叫Smith的男人的模糊集合，特指的关系具有的形式为

$$MAN[name=Smith;\ \Pi_{Height}=TALL;\ \Pi_{Colour(Hair)}=BLOND] \quad (2.30)$$

（注意当想要强调它表示一个关系时，用大写字母来标识）。类似地，除了具有上述特征外，还有年龄大约30岁的男人的模糊集合，将表示成：

$$MAN[Name=Smith;\ \Pi_{Height}=TALL;\ \Pi_{Colour(Hair)}=BLOND;\ \Pi_{Age}=APPROXIMATELYEQUAL[Age=30]] \quad (2.31)$$

在这种情况下，与变量Age相关的可能性分布本身就是一个特指的可能性分布。

应当注意的是，在（2.30）和（2.31）中使用的这种表示形式与在语义网中普遍采用的形式及用自然语言表示的高阶谓词演算中命题的形式有点类似。有关这种表示的讨论可以从下面一些文章，Newell和Simon（1972），Miller和Johnson-Laird（1976），Bobron和Collins（1975），Minsky（1975）及文献目录中列出的其它的文章和书中找到。然而在（2.30）和（2.31）中为了表示模糊变量的值使用的可能性分布与用特指修饰的可能性分布的具体的说明方式之间有着本质的差别。

意义和信息

由（2.26）式给出的特指的定义在PRUF（一个用于表示模糊命题的意义的语言）中起着特别重要的作用。在Zadeh（1977b）的文章中对PRUF进行了简要的说明。将有另外的文章提供关于PRUF的详细的讨论。

简单地说，一般的在PRUF中的一个表达式P是一个计算可能性分布的过程。更明确地，令U是一个论域，$\mathscr{R}$是U中的一个关系的集合。那么，这两者

$$D\triangleq(U,\mathscr{R}) \quad (2.32)$$

组成一个数据库，用P定义在$\mathscr{R}$中的一个关系的子集。正如上面定义的那样，一个数据库的概念是与模态逻辑中的一个可能的域相联系的（Hughes and Cresswell, 1968; Miller and Johnson-Laird, 1976）。

如果p是某种自然语言的一个表达式，而且P是它在PRUF中的翻译，即$p\to P$，那么过程P可以看成是用它计算的可能性分布定义p的意义$M(p)$，组成由p传达的信息$I(p)$。（用PRUF中的一个表达式定义的过程和它产生的可能性分布类似于在二值逻辑中的一个谓词的内涵和外延。（Cresswell, 1975.）当使用的意义不很严格时，$M(p)$和$I(p)$之间没有什么差别。）

作为一个简单的例子，考虑命题

$$p\triangleq \text{John resides near Berkley} \quad (2.33)$$

在PRUF中，它翻译成

$$RESIDENCE[Subject=John;\ \pi_{Location}=Proj_{\mu\times City1}\ NEAR[City2=Berkeley]] \quad (2.34)$$

其中$NEAR$是一个模糊关系，有框架$NEAR\ \|City1\,|\,City2\,|\,\mu\,|$，表达式$Proj_{\mu\times City1}\ NEAR[City2=Berkeley]$代表靠近$Berkeley$的城市的模糊集合。一个模糊关系的框架列出它的名字及它的变量（即属性）的名字和μ—它是关系中的每个元组的隶属度。

在PRUF中，表示成（2.34）这种形式的表达式，实际上是一个计算John的居住地

表1

RESIDENCE	Subject	Location	π
	John	Oakland	1
	John	Polo Alto	0.6
	John	San Jose	0.2
	John	Orinda	0.3
	……	…	…

的可能性分布的过程。因此，给定一个关系 *NEAR*，它将回送一个表1这种形式的可能性分布（π≜可能性值），它可以看成是由命题"John resides near Berkeley"传达的信息。

PRUF在近似推理中是必不可少的，因为它是一个基础，通过它把用自然语言形式表示的模糊前提翻译成可能性赋值方程，可以以一种系统的方式应用近似推理中的那些推理规则。在第四节我们将较详细地讨论在模糊逻辑中使用的一些基本的翻译规则，它们组成了PRUF的翻译规则的一个小的子集。关于PRUF的这个简单的说明可以满足本文的要求。

下面我们将转向讨论语言变量的概念，这个概念是近似推理，模糊逻辑和系统分析的语言方法的基础。

3　语言变量的概念

在以人为中心组成的系统中，描述人的行为时，通常是使用词语而不是数字表示变量的值和变量间的关系。例如，描述人的年纪可以用"非常年轻"，描述人的智商可用"相当高"，还可以用"不是非常友好"说明与其他人的关系，用"相当迷人"描述人的容貌。

显然使用词语而不是数字表示一个变量的值，意味着精确度是较低的。但在一些情况下，我们仍选用不精确的词语，是因为这些情况不需要那么高的精度。然而，有很多情况下是因为某种对象的属性没有度量单位，没有像使用某种坐标来表示这种属性值的定量标准，只好用不精确的词语对它们进行描述。

从这个角度来看，语言变量的概念可以看成是，用某种自然语言或人造语言的词语或句子来表示变量的值和描述变量间的内在联系的一种系统化的方法。语言变量概念所起的作用是，它为近似推理中变量值的表示和模糊命题的真值，概率值和可能性值的表示提供了一个基本的函数。

在这一节，我们只集中讨论语言变量概念与近似推理直接有关的那些问题。关于语言变量概念及其应用的更详细的讨论，可以在Zadeh (1973，1975c)，Wenstop(1975，1976)，Mamdani and Assilian(1975)，Procyk(1976)及文献目录中其它文章中找到。

作为讨论的出发点，一个变量例如年龄，既可以把它看成是一个分布在[0，150]区间上的数字变量，也可以看成是一个取值为*young*，*not young*，*very young*，*not very young*，*quite young*，*old*，*not very young*和*not very old*等等这类值的语言变量。这些值中的每一个，都可以解释成是论域$U=[0，150]$的一个模糊子集的一个标名，它的基变量u是年龄的类属的数字值。

典型地，一个像年龄这样的语言变量的值是建立在一个或多个原始词（原始模糊集合的标名起的作用，有点类似于机械系统的物理单元）之上，加上一组修饰词和连接词，将它们作用于原始词上产生合成的语言值。通常这种词的个数是2个，其中一个是另一个的反义词。例如，在年龄这个情况下，原始词是年轻（*young*）和老（*old*）。

在语言变量概念下的一个基本的假设是，原始词意义是上下文相关的，而修饰词和连接词的意义是上下文无关的。而且一旦原始词的意义在一个给定的上下文中被指定（或被标准），合成术语像*not very young*，*not very young and not very old*等的意义，可以通过应用一个语义规则计算得到。

典型地，术语集合是语言变量的语言值的集合，它是由每一个原始词产生的值以及由每一个原始词的各种不同的组合产生的值组成的。例如，在年龄的情况下，年龄的语言值的一部分列表如下：

young	old
not young	not old
very young	very old
not very young	not very old
quite young	quite old
more or less young	more or less old
extremely young	extremely old
etc.	etc.

not young nor old

not very young and not very old

young or old

not young or not old etc.

注意到大多数的语言变量具有与*Age*相同的基本结构这一点很重要。例如，由于用*tall*替换*young*，用*short*替换*old*，我们就得到语言变量*Height*的语言值表。我们可以对语言变量*Weight*（*heavy*和*light*），*Appearance*（*beautiful*和*ugly*），*Speed*（*fast*和*slow*），*Truth*（*true*和*false*）等进行同样的替换。括号内词是代表原始词。

正如Zadeh在文章(1973，1975c)中指出的那样，一个语言变量可以利用属性文法来表示（见Knuth 1968；Lewis等1974），它产生变量的术语集合，并为计算合成的语言值的意义提供了一个简单的过程，它是根据出现在合成语言值的成分中原始模糊集合进行计算的。

现在来看下面给出的属性文法的例子，其中S，B，C，D和E是非终极符；*not*，a和b是终极符；a和b是原始词（也就是原始模糊集合）；下标符号是模糊集合，用非终极符来标记，且有$L \triangleq left$（即属于先行词），$R \triangleq right$（即属于结论）；一个具有下面这种形式的产生式

$$S \to S \text{ and } B,\ S_L = S_R \cap B_R \quad (3.1)$$

表示代表先行词S的意义的模糊集合是由分别代表结论S和B的意义模糊集合S_R和B_R的交构成。

$S \to B$，$S_L = B_R$ (3.2)

$S \to S$ and B，$S_L = S_R \cap B_R$

$B \to C$，$B_L = C_R$

$B \to$ not C，$B_L = C'_R$（$\triangleq C_R$的补集）

$C \to S$，$C_L = S_R$

$C \to D$，$C_L = D_R$

$C \to E$，$C_L = E_R$

$D \to Very\ D$，$D_L = D^2_R$（$\triangleq D_R$的平方）

$E \to Very\ E$，$E_L = E^2_R$（$\triangleq E_R$的平方）

$D \to a$，$D_L = a$

$E \to b$，$E_L = b$.

上述的文法生成的语言值由下表列出：

a	b
not a	not b
very a	very b
not very a	not very b
not very very a	not very very b
etc.	etc.

a and b

not a and b

not a and not b

not very a and not very b

etc.

一般地，为了计算由这个文法生成的语言值l的意义，需要利用(3.2)中的关系式，根据它的直接子孙的意义，计算l的语法树的每个结点的意义。然而，在大多数情况下，计算可以通过调查来完成，它包括翻译法则的简单应用，这些法则将在第四节阐述，因此我们可以很容易地得出结果。例如，

$$\text{not very } a \to (a^2)' \quad (3.3)$$

$$\text{not very } a \text{ and not very } b \to (a^2)' \cap (b^2)'$$

其中a'是a的补，a^2是由下式定义的

$$\mu_{a^2}(u) = (\mu_a(u))^2,\ u \in U. \quad (3.4)$$

为了表示原始模糊集合a和b。使用带有可调参数标准的隶属函数往往是较为方便的。一个这样的函数*S-function*（*S*–函数）$S(u;\ \alpha,\ \beta,\ \gamma)$定义为

$$\begin{aligned} S(u;\ \alpha,\ \beta\ \gamma) &= 0 && \text{对} u \leqslant \alpha && (3.5) \\ &= 2\left(\frac{u-\alpha}{\gamma-\alpha}\right)^2 && \text{对} \alpha \leqslant u \leqslant \beta \\ &= 1 - 2\left(\frac{u-\gamma}{\gamma-\alpha}\right)^2 && \text{对} \beta \leqslant u \leqslant \gamma \\ &= 1 && \text{对} u \geqslant \gamma \end{aligned}$$

其中参数$\beta \triangleq \frac{\alpha+\gamma}{2}$是交叉点，即在这一点的$u$的值$S(u,\ \alpha,\ \beta,\ \gamma) = 0.5$。例如，如果$a \triangleq$ young，$b \triangleq$ old，我们或许有(见图1)

$$\mu_{young} = 1 - S(20,\ 30,\ 40) \quad (3.6)$$

和 $$\mu_{old} = S(40,\ 55,\ 70) \quad (3.7)$$

其中变元u为了简单起见被隐藏起来。这样，根据(3.6)，命题$p \triangleq$ Joe is young(见(2.13))的翻译，可以更具体地表示成

$$\text{Joe is young} \to \pi_{Age(Joe)} = 1 - S(20,\ 30,\ 40) \quad (3.8)$$

其中$\pi_{Age(Joe)}$是语言变量$Age(Joe)$的可能性分布函数，类似地

$$\text{Joe is not very young} \to \pi_{Age(Joe)} = 1 - [1 - S(20,\ 30,\ 40)]^2 \quad (3.9)$$

语言变量概念的一个重要方面是与这样的事实有关的，即一般这样一个变量的术语集合，在可对模糊集合施加的各种运算（如并，交，直积等）下是不封闭的。例如，如果l是一个变量X的一个语

软件攻坚术：速成原型与知识工程的结合

Pamela W. Jordan, Karl S. Keller,
Richard W. Tucker, David Vogel

摘要

本文提出一种叫做两阶段软件攻坚术的软件开发方法，它是速成原型法与知识工程相结合的产物。使用这种方法能生产功能更多的原型，而所需时间为四个月，而不像传统原型法那样需要2年。文中论及该方法与现有原型法的差别，并以开发美国陆军的MSES为例，说明了它的要领和主要做法。最后对该方法提出了几点建议。

没有任何一项革新称得上对软件开发过程的效率和方便性做出了重大改进。[1,2]软件开发方法分为两大门派：（1）从与机器打交道着手的软件开发方法，即高级程序语言和软件开发环境[3]；（2）着手软件的设计和规格说明的方法。[1,4,5]速成原型是这两大门派的混合。它是一种评价活动，试图概括一个系统解决目标系统中涉及的某些问题的能力，[6]同时引导对新暴露的问题做进一步研究。

在Mitre-Washington人工智能中心，我们试验了一种快速生产高功能原型的方法。我们对这一方法杜撰了软件攻坚术（software storming）这一名词，它涉及专家在紧张的开发工作中参与系统的初期设计与实现，把知识工程和软件开发技术与工作站硬件的最新成果相结合。这里所描述的试验是对人工智能领域中所发展的工具和技术的一种测试，也是软件攻坚术形成的第一步。使用软件攻坚术，我们花了不到四个月的时间，开发了一个比标准原型功能性强得多的软件原型。

言值，那么一般l^2不在X的术语集合中。

寻找X的一个语言值的问题（它的意义近似于U的一个给定的模糊子集）被称为语言近似的问题（Zadeh，1975c，Wenstop，1975；Procyk，1976）。在这篇文章中，我们将不讨论如何处理这一相当重要的问题，但是将假定，语言近似是隐含在一个可能性分布重新翻译成用自然语言表示的一个命题的过程之中。（见(2.10)）

（未完待续）

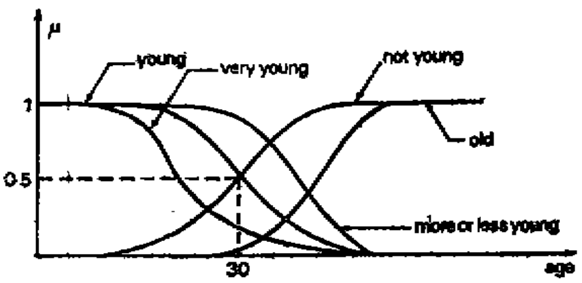

图1 Age的语言值的图形表示

〔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〕