

知识论框架

通向信息 - 知识 - 智能统一的理论

钟义信

(北京邮电大学, 北京 100876)

[摘要] 知识是人类所创造的宝贵财富,但至今没有形成系统的知识理论。文章旨在提出和建立知识论的框架体系,包括它的基础和主体两部分。基础部分主要给出知识的概念、定义、表示、度量、推理和决策规则;主体部分的核心是阐明由信息提炼知识(知识生成)以及由知识形成智能(知识激活)的机理。知识论的建立将为信息论 - 知识论 - 智能论的统一理论奠定坚实的基础,促进人们在更高的水平上利用信息和知识,研究、设计和应用各种智能机器,推动经济和社会的发展。

[关键词] 知识;知识量;知识生成;知识激活;信息 - 知识 - 智能的统一理论

1 引言

苏联科学院院士哈尔凯维奇在他1955年出版的《通信论简述》中指出:信息学存在一个如同物理学“能量守恒与转换”那样的基本定律,可惜至今没有被发现。把信息提炼成知识并把知识激活成智能,是信息学的核心和灵魂。因此,“信息 - 知识 - 智能转化理论”应当是信息学的“能量守恒与转换定律”。

科学技术的主要任务在于启迪人们:怎样利用各种资源创制先进的生产工具,扩展人类认识世界和改造世界的能力,改善和优化人类生存和发展的环境。

从工程科学的视角看,整个科学技术进步的历史,就是一部沿着“利用外部资源、创制生产工具、扩展人的能力”方向不断向深度和广度进军的历史:

——古代人类主要凭借古代初等的材料科学技术把物质资源加工成为材料,创造了人类社会第一代社会生产工具——人力工具,如锄头镰刀等等,扩展了人的体质功能,推动了农业社会生产力的发展,创建了农业时代的文明;

——近代人类借助于近代材料科学技术和能量科学技术把物质和能量两种资源分别转换成为材料和动力,创制了第二代社会生产工具——动力工具,如机床机车等等,扩展了人的体力功能,推动了工业社会生产力的发展,创建了工业时代的文明;

——现代人类正在利用现代材料科学技术、能量科学技术和信息科学技术把物质、能量和信息三种资源转变成为材料、动力和知识,创制第三代社会生产工具——智能工具,如机器人和专家系统等等,扩展人的智力功能,推动信息社会生产力的发展,创建信息时代的文明。

由此可以理解,“创制智能工具”已经成为我们时代科学技术的核心任务。当今世界,只有在国民经济各个部门和社会活动各个领域广泛地充分地采用智能工具,才能有效地实现国家工业、农业、科学技术、国防、文化教育以及其它各行各业的现代化。

现代科学技术表明,创制智能工具的基本问题是:怎样使机器(而不只是人)能够把“信息”加工成“知识”,又怎样使机器(而不只是人)能够把“知识”激活成“智能”。而这正是本文研究的

主题——知识论。它是信息论－知识论－智能论统一理论的重要组成部分，是信息科学的灵魂。

信息论和智能论已经分别在40年代和50年代问世，但是长期以来，它们各自独立发展，互相没有联系。本来，在信息论与智能论之间应当存在知识论作为桥梁。然而，知识论却长期成为一段空白。这种状况已经成为瓶颈，使信息论和智能论的发展也受到了严重的制约。因此，研究和建立知识论已经成为一项紧迫的任务。知识论的建立，将可以沟通信息论与智能论之间的联系，形成信息、知识、智能的统一理论，为创制智能工具提供坚实的理论基础。

2 概念与定义

一切科学理论都建立在自己的科学概念基础上。知识论也不例外。因此，本节首先将定义一組关于知识理论的基本概念：信息，知识，智能。考虑到信息和智能的理论虽然还不完整，但毕竟已经存在，于是，研究的重点将放在知识理论本身。只是为了揭示信息、知识和智能之间的联系，才对照地研究信息与智能的定义。

首先给出信息的定义。所谓得到了关于某个事物的信息，通常就是指：知道了这个事物现在处在什么样的运动状态以及知道了这个状态会按照什么方式变化。于是有：

定义2-1 从本体论意义上说，信息是指事物运动的状态和状态变化的方式^[1]。

例：对于某个随机型信息 X ，如果它有 N 种可能的运动状态，即其状态空间为：

$$x_1, x_2, \dots, x_n, \dots, x_N,$$

又如果这些状态的变化方式是按照某种概率分布的规则进行的：

$$p_1, p_2, \dots, p_n, \dots, p_N.$$

那么，由状态空间与概率分布结合而成的概率空间，就充分地刻画了这个信息。

需要说明的是，定义2-1所说的“事物”可以是物质客体，也可以是精神现象。物质和精神都处在不断的运动之中，都具有一定的运动状态，而且这些运动状态都会按照某种方式发生变化，因此都会产生信息。控制论的创始人维纳曾经说过“信息就是信息，不是物质，也不是能量”^[2]。信息论的奠基人仙农也曾经说过“信息是用来消除不定性

的东西”^[3]。显然，定义2-1与维纳和仙农的信息定义是相通而且等效的。

根据文献[1]，信息也有认识论层次的定义，称为“全信息”。不过，为了使知识论的讨论更清晰更精炼，此处从略。

现在就来讨论知识的概念和定义。一般认为：

定义2-2 知识是人们实践经验的结晶。

一切经验的共同特征是：它告诉人们什么样（状态）的事情应当用什么样的办法（状态变化的方式）去做才能达到目的（新状态）。可见，经验也是指“事物运动的状态和状态变化的方式”，因此也是信息。定义2-2表明，知识是信息的结晶。

为了便于与信息的定义进行对照研究，也可以把知识重新定义为：

定义2-3 知识是认识论范畴的概念，是关于事物运动的状态和状态变化的规律。

例如，在牛顿力学中， $F = ma$ 是一个知识，它告诉人们：质量为 m 的物体，受到大小为 F 的力作用后会产生加速度为 a 的加速运动。在这里，知识所告诉人们的，正是一类受力作用的物体的运动状态以及状态变化的规律。

又如，量子力学的德布罗意波函数 $\Psi(x, t) = \Psi_0 \exp[-i \frac{2\pi}{h} (Et - px)]$ 也是一个知识，它告诉人们的是能量为 E 、动量为 p 具有波粒二象性的实物自由粒子的运动状态和状态变化的规律。

再如，化学反应知识告诉人们的是如何由两种（或多种）物质（状态）化合成新物质（状态）的规律；生物遗传工程学知识告诉人们的是如何由父代（状态）衍生成子代（新状态）的规律；控制论告诉人们的则是使一个受控系统由原有状态演进到新状态（目标）的规律，等等。

可见，信息和知识概念互相贯通：由具体的“状态变化方式”（信息）到抽象的“状态变化规律”（知识），其间所经历的变化正是人们对信息所进行的提炼和加工。信息作为一种原材料，经过加工提炼之后，就可能形成相应的抽象产物——知识。由此可以得到：

推论2-1 知识是信息加工的规律性抽象产物。

不难看出，推论2-1与定义2-2完全一致的。现在给出智能的定义。

定义 2-4 在给定的问题、问题环境、主体目的的条件下,智能就是有针对性地获取问题与环境的信息、恰当地对这些信息进行处理以达到认知、然后在此基础上结合主体的目的信息合理地产生解决问题的策略信息、并利用所得到的策略信息在给定的环境下成功地解决问题达到主体的目的的能力^[4]。

定义 2-4 表明,作为能力,智能包含三个基本方面:首先必须有认知的能力,即获取有关信息和恰当处理这些信息并由此生成相应知识(达到认知)的能力;同时,必须具有决策的能力,即根据所得到的知识结合主体的目的生成解决问题的策略信息(决策)的能力;最后,还必须具有实施的能力,即利用所生成的策略信息在给定环境下实际解决问题达到目的的能力。

图 2-1 示出了智能过程中信息-知识-智能三者相互依存、共为一体的关系:

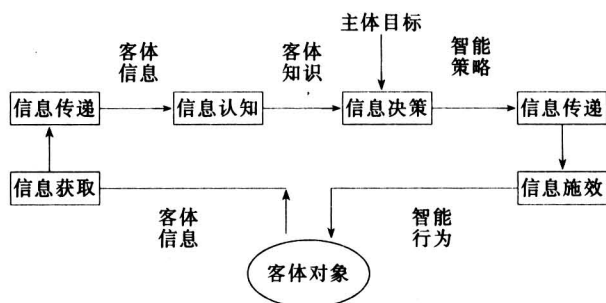


图 2-1 信息-知识-智能的关系

Fig. 2-1 Relations among information, knowledge and intelligence

总结起来,信息、知识、智能之间的关系可以这样来表述:信息是基本资源;知识是对信息进行加工所得到的抽象化产物;智能是利用信息资源加工生成知识、进而激活知识生成解决问题的策略信息并在策略信息引导下具体解决问题的能力。

图 2-1 所示的信息、知识、智能的关系,正好符合人类自身认识世界和优化世界活动过程中由信息生成知识、由知识激活智能的过程:其中,获取信息的功能由人的感觉器官完成,传递信息的功能由人的神经系统完成,处理信息和再生信息的功能由人的思维器官完成,施用信息的功能由人的效应器官完成。由此,可以体会信息-知识-智能统一理论的重要意义,也可以得到:

推论 2-2 信息经加工提炼而成知识,知识被目的激活而成智能。

3 知识的表示

上面已经阐明,知识是认识论范畴的概念,是相对于认识主体而存在的。因此,与本体论意义上的信息相比,知识的概念更为复杂,具有更丰富的内涵。

为了研究知识理论,必须解决知识表示的问题;而为了更好地表示知识,又必须研究知识的分类,以便在此基础上对知识进行分门别类的表示,而不是笼统的表示。但是,人类迄今所拥有的知识已经构成一个极其庞大的学科体系。因此,这里所关注的知识分类不可能是按学科来划分的类,而是为了研究知识的一般理论、针对一切知识所共有的性质而提出的具有普遍意义的分类。

一切知识,无论是数学、物理学、化学、天文学、地学、生物学的知识,还是工程科学的知识,它们所表达的“运动状态和状态变化的规律”必然具有一定的外部形态,与此相对应的知识可以称为“形态性知识”;同时,知识所表达的运动状态和状态变化的规律必然具有一定的逻辑内容,与此相对应的知识可以称为“内容性知识”;最后,知识所表达的运动状态和状态变化的规律必然对认识主体呈现某种价值,与此相对应的知识可以称为“效用性知识”。需要强调的是,形态、内容、效用三者的综合才构成知识的整体,见图 3-1。

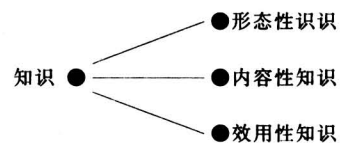


图 3-1 知识的三位一体

Fig.3-1 The trinity of knowledge

公理 3-1 任何知识都由相应的形态性知识、内容性知识、效用性知识构成。这种情形称为知识的三位一体。

3.1 形态性知识的描述

对于形态性知识的描述,就是对事物运动的状态及其变化规律的形式描述,主要回答的问题是:“它有多少种可能的运动状态,这些状态变化规律的形式特征是什么”。

一般来说,从形式上来描述事物运动状态的特征就是直接对这些状态进行编号,每个编号对应于一种实际的运动状态。

而为了从形式上来描述状态变化规律的特征，通常要针对事物运动的不同性质，采取不同的描述方法：例如，如果事物状态变化规律是随机的性质，它的状态变化规律的形式特征可以用概率分布 P 来描述，如图 3-2 所示。

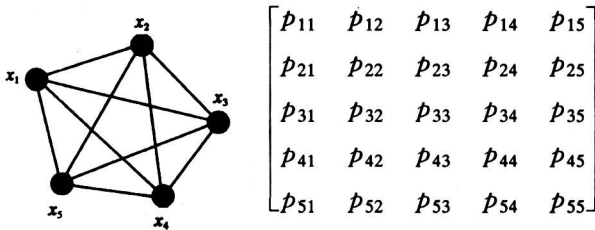


图 3-2 状态及其随机变化规律

Fig.3-2 states and their random varying law

如果事物是半随机变量（也称为“偶发变量”），它的状态变化规律的形式特征是可能度分布 Q 。如果事物是确定性模糊变量，它的状态变化规律的形式特征则是隶属度分布 M 。

根据莱布尼兹、贝努里和文献 [1] 的分析，概率分布、可能度分布、隶属度分布的统一概括是肯定度分布 C 。因此，有：

定义 3-1 设事物 X 具有 N 种可能的状态： $x_1, \dots, x_n, \dots, x_N$ ，那么，状态 x_n 在形态上呈现的肯定程度称为状态 x_n 的肯定度，记为 $c_n, n = 1, \dots, N$ 。由 X 的全部状态的肯定度所构成的集合，称为 X 的肯定度的（广义）分布，记为 C 。它刻画了该事物状态变化的形式规律。

注意到概率和可能度的归一性、隶属度的不归一性^[1]，肯定度 C 应当具有如下的性质：

$$0 \leq c_n \leq 1, \forall n \text{ 和 } \sum_{n=1}^N c_n \geq = \leq 1 \quad (3.1)$$

式 (3-1) 中的符号 $\geq = \leq$ 表示全部状态的肯定度之和可以大于、等于、小于 1，不一定能够归一。具体来说，当给定的事物是随机型或偶发型变量时，和式归一；当给定的事物是模糊型变量时则不归一。

因此，可以用事物 X 的状态集合及其肯定度分布 $\{X, C\}$ 来描述事物 X 的形态性知识。

3.2 内容性知识的描述

透过任何形态性知识，必然蕴含着相应的逻辑内容。因此，应当进一步来讨论内容性知识的描述问题。显然，事物 X 的状态所代表的实际内容将

随不同的具体事物而千差万别，无限丰富多彩，不可能对它们一一做出具体的描述。

逻辑学的原理证明：一切科学定律和定理都可以用一串真伪选择序列来表达。因而，关于内容性知识的比较合理的共性描述，是各个状态在逻辑上的真实性或真伪性。

定义 3-2 设事物 X 具有 N 种可能的状态： $x_1, \dots, x_n, \dots, x_N$ ，那么，状态 x_n 在逻辑上真实的程度称为状态 x_n 的真实度，记为 $t_n, n = 1, \dots, N$ 。 X 的各个状态的真实度所构成的集合，称为 X 的真实度的（广义）分布，记为 T 。

按照定义 3-2，显然有

$$0 \leq t_n \leq 1, \forall n \text{ 和 } \sum_{n=1}^N t_n \geq = \leq 1 \quad (3.2)$$

因此，可以用事物 X 的状态集合及其真实度分布 $\{X, T\}$ 来描述事物 X 的内容性知识。

3.3 效用性知识的描述

根据事物 X 各个状态 x_n 相对于主体目标所显示的价值来定义相应状态的效用度 $u_n, n = 1, \dots, N$ ，即

定义 3-3 设事物 X 具有 N 种可能的状态： $x_1, \dots, x_n, \dots, x_N$ ，那么，状态 x_n 相对于主体目标所显示的价值称为状态 x_n 的效用度，记为 $u_n, n = 1, \dots, N$ 。 X 的各个状态的效用度所构成的集合，称为 X 的效用度的（广义）分布，记为 U 。

按照定义 3-3，也有

$$0 \leq u_n \leq 1, \forall n \text{ 和 } \sum_{n=1}^N u_n \geq = \leq 1 \quad (3.3)$$

因此，可以用事物 X 的状态集合及其效用度分布来描述事物 X 的效用性知识。

3.4 综合内容性知识和综合效用性知识

注意到，从认识论的程序上考虑，在形式、内容和效用三要素之间，形式是最先被观察或感受到的要素，内容是要透过形式的分析才能进一步感受到的要素，效用则更是要针对一定的形式、内容和主体才能表现出来的要素。因此，除了如上所述分别给出单纯形态性知识、单纯内容性知识和单纯效用性知识的描述之外，还有必要讨论形态性与内容性知识的综合描述以及形态性、内容性和效用性知识的综合描述方法。为此：

定义 3-4 状态的肯定度与状态的真实度的结合称为状态的综合真实度，记为

$$\mathcal{J}_n = \alpha c_n \cdot \beta t_n \Rightarrow c_n t_n, \forall n$$

$$\mathcal{J} = \{\mathcal{J}_n\} \quad (3.4)$$

称 \mathcal{J} 为综合真实度分布。式中 \Rightarrow 表示“可简化为”。显然有

$$0 \leq \mathcal{J}_n \leq 1, \forall n \text{ 和 } \sum_{n=1}^N \mathcal{J}_n \geq = \leq 1 \quad (3.5)$$

与此相应的知识,称为综合内容性知识。

定义3-5 状态的肯定度、真实度与效用度的结合称为状态的综合效用度,记为

$$\eta_n = \alpha c_n \cdot \beta t_n \cdot \gamma u_n \Rightarrow c_n t_n u_n, \forall n \quad (3.6)$$

$$\eta = \{\eta_n\}$$

称 η 为综合效用度分布。显然也有

$$0 \leq \eta_n \leq 1, \forall n \text{ 和 } \sum_{n=1}^N \eta_n \geq = \leq 1 \quad (3.7)$$

与此相应的知识,称为综合效用性知识。

在研究逻辑推理问题的场合,综合内容性知识的描述非常有用;在研究基于知识的决策问题的时候,综合效用性知识的描述非常有用。

应当指出,知识描述的方法不是唯一的。但是不管何种知识描述方法,都应当能够描述事物的运动状态以及状态的变化规律。这是基本的要求。另外,这里所采用的知识描述方法不是针对某几种具体领域的知识的描述,而是针对一切知识的共性——形态性知识、内容性知识和效用性知识——所做的描述,因此具有广泛的普遍适用性。

4 知识的处理

对于知识所进行的任何操作过程,都可以称为知识的处理。但是,实际上,知识处理的最基本和最重要的内容应当包括:各种知识的定量度量、基于综合内容性知识的推理、以及基于综合效用性知识的决策。下面就分别来讨论知识的度量、推理和决策问题。

4.1 知识的度量

定义4-1 知识的数量称为知识量。

如同知识本身一样,知识量也可以进一步分为形态性知识量、内容性知识量、效用性知识量、综合内容性知识量和综合效用性知识量。其中,最基础的知识量是形态性知识量。

研究知识度量的一个直观而合理的思路是:用“所能解决的问题量”来度量相应的“知识量”。因此,“知识量”的研究就转化为“问题量”的研究。

进一步,按照一般的度量理论,首先应当设计

一种合理的标准问题,把它所包含的问题量作为问题量的“单位”,然后,任何一个实际问题的问题量就可以同这个单位相比较,从而得出这个实际问题的问题量。

注意到,一个形态性问题的题量与两个因素有关:一是问题的可能状态数;二是问题各状态的肯定度分布。一方面,在同样的肯定度分布条件下,问题的可能状态数越大,问题量也越大;另一方面,在同样的可能状态数的条件下,肯定度分布越均匀,问题量也越大。

因此,最容易被接受的合理标准问题是“标准二中择一问题”,即一个问题只有两种可能的状态,且这两种状态的肯定度相等。这里所谓“标准的”二中择一问题,是指两种可能状态的肯定度分布为均匀分布。标准二中择一问题的模型如图4-1所示。

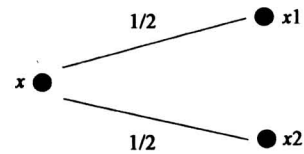


图4-1 标准二中择一问题模型

Fig.4-1 A standard alternative problem model

定义4-2 标准二中择一问题所包含的问题量定义为一个单位问题量,单位为奥特。如果一个知识正好解决了一个单位问题,那么,它的知识量也就等于1奥特知识量。

定义4-2是很自然的,因为按照这个定义,所谓1单位知识量就是解决一个标准的“是或非”问题、一个标准的“正或负”问题、一个标准的“男或女”问题、一个标准的“有或无”问题、一个标准的“好或坏”问题、一个标准的“输或赢”问题等等所需要的知识量。

其中单位“奥特”是“二中择一”问题的英文(alternative)前三个字母“alt”的译音,也可以简称为“奥”。单位的中文“奥”含有“深奥”和“奥妙”的意思,与“问题”的含义正好相通,因此用作问题量的单位,不无美感。在随机信息场合,“奥特”可退化为“比特”。

有了问题量单位,一个具体问题的问题量就等于它所包含的标准二中择一问题的数量。例如,如果某个实际问题有4种可能的状态,且4种状态的

肯定度都等于 1/4，那么，这个问题实际上包含了两个“标准二中择一问题”，它的问题量为两个单位，即 2 奥。又如，如果一个问题有 8 种可能的状态，而且这 8 种状态的肯定度都等于 1/8，那么，这个问题所包含的问题量就是 3 奥，如此等等。

虽然利用单位问题量来测度实际问题量的方法在概念上非常直观自然，但是，实际使用却并不方便。一方面，小于 1 奥的问题量不好度量，另一方面，当问题所包含的可能状态数目不是正好等于 2 的 n 次方或肯定度不是均匀分布时，问题量也不好度量。因此，还是要寻求一般的问题量的度量方法。

为此，还需要引入新的概念和定义。

注意到式 (3.1)，肯定度存在归一和不归一两种情形。这里首先研究归一的情形。

定义 4-3 均匀分布的肯定度和 0-1 型分布的肯定度代表肯定度分布的两种极端情形，分别把它们记为

$$C_0 = \{c_n \mid c_n = \frac{1}{N}, \forall n\} \quad (4.1)$$

$$C_S = \{c_n \mid c_n \in \{0,1\}, \forall n\} \quad (4.2)$$

定义 4-4 定义在肯定度分布 C 上的平均肯定度由下式给出：

$$M_\varphi(C) = \varphi^{-1}\left\{\sum_{n=1}^N c_n \varphi(c_n)\right\} \quad (4.3)$$

式 (4.3) 中的 φ 是待定的单调连续函数， φ^{-1} 是它的逆函数，也单调连续。

定义 4-5 两个问题 X 和 Y 具有相同的状态数 N，各自的肯定度分布为 C 和 D，若满足条件

$$\varphi^{-1}\left\{\sum_{n=1}^N c_n \varphi(c_n d_n)\right\} = \varphi^{-1}\left\{\sum_{n=1}^N c_n \varphi(c_n)\right\} \cdot \varphi^{-1}\left\{\sum_{n=1}^N c_n \varphi(d_n)\right\} \quad (4.4)$$

则称它们互相 φ -无关。于是有：

定理 4-1 满足定义 4-4 和 4-5 条件的待定函数 φ 必为对数形式。

这是一个很重要的结果。不过它的证明可在文献 [5] 和 [6] 找到，这里从略。由此可以得出：

系 4-1 状态数为 N 且肯定度分布为 C 的事件 X 的平均肯定度为

$$M_\varphi(C) = \prod_{n=1}^N (c_n)^{c_n} \quad (4.5)$$

系 4-2 这样定义的平均肯定度的值界于 1/N 与 1 之间

$$\frac{1}{N} = M_\varphi(C_0) \leq M_\varphi(C) \leq M_\varphi(C_S) = 1 \quad (4.6)$$

系 4-1 和 4-2 的证明是直截了当的。

系 4-2 的结果表明，肯定度为均匀分布时，平均肯定度最小；肯定度为 0-1 分布时则平均肯定度最大。前者是最不肯定的情形，相当于无知识的情形；后者是完全肯定的情形，相当于拥有充分知识的情形。

由此，可以自然地引进一个重要的概念：某个观察者对于某个事物是否拥有知识，或拥有多少知识，可以用这个观察者对于这个事物所具有的平均肯定度的大小来判断。平均肯定度越大，拥有的知识越充分。

如果把最小平均肯定度作为一个比较的基准，就可以建立一个相对的形态性知识度量：

定义 4-6 观察者 R 关于事物 (X, C) 的形态性知识量，可以用下式测度：

$$K(C) = \text{lb} \frac{M_\varphi(C)}{M_\varphi(C_0)} = \text{lb} N + \sum_{n=1}^N c_n \text{lb} c_n \quad (4.7)$$

定义 4-7 观察者 R 在观察某个事物 X 之前所具有的关于 X 的肯定度分布称为他关于 X 的先验肯定度分布，通常记为 C；观察之后的肯定度分布则称为后验肯定度分布，记为 C*。

于是，所谓观察者 R 通过观察获得了关于事物 X 的形态性知识，就是指他在观察之后关于 X 的后验平均肯定度比观察之前的先验平均肯定度增大了。

定义 4-8 观察者 R 通过观察 X 所获得的形态性知识量可以用下式测度：

$$K(C, C^*; R) = K(C^*) - K(C) = \sum_{n=1}^N c_n^* \text{lb} c_n^* - \sum_{n=1}^N c_n \text{lb} c_n \quad (4.8)$$

这是在肯定度分布归一的情形下关于形态性知识量的重要结果。只要知道了观察者在观察某一事物或实验的先验和后验肯定度分布，就总是可以利用式 (4.8) 计算出观察者在观察过程中所得到的形态性知识量。

由式 (4.8) 可知，当先验肯定度为均匀分布而后验肯定度分布为 0-1 分布时，观察者所获得的形态性知识量达到最大值。一般，只要观察者的后验平均肯定度大于先验平均肯定度，就意味着他能够在观察过程中获得某种程度的形态性知识。另

一方面, 不管先验和后验肯定度分布的形式如何, 只要两者相同, 观察者在观察过程中所获得的形态性知识量就总是为 0。反之若观察者的平均后验肯定度小于平均先验肯定度, 就意味着他在观察过程中丢失了形态性知识量。若观察者的先验肯定度分布为 0-1 形式而后验肯定度分布为均匀分布, 那么, 观察者在观察过程中所丢失的形态性知识量达到最大值。这些都是与人们的直觉相一致的结果, 因而是合理的结果。

不难看出, 定义 4-2 规定的单位形态性知识量与定义 4-8 的理论结果是完全一致的。只要在式 (4.8) 中令状态数 $N=2$, $c_1=c_2=1/2$, C^* 为 0-1 分布, 就可以得到:

$$K(C_0, C_S^*; R) = 0 - \text{lb} \frac{1}{2} = 1$$

其中对数的底取为 2, 单位为奥 (特)。

在理想观察条件下, 后验肯定度为 0-1 型分布。此时, 若假定先验分布为均匀分布, 那么由式 (4.8) 可以得到

$$K(C_0, C_S^*; R) = \text{lb} N$$

这时, 形态性知识量与状态数目呈对数函数关系。

如果对这一关系作进一步的人为简化, 把 $\text{lb} N$ 简化为 N , 就可以直接用状态数目来近似计算知识量。这就是为什么情报界和文化界通常都用字数来估计情报量的道理。显然, 这只是一种非常粗糙的估计。

现在再来考虑肯定度不归一的情形, 即模糊试验的情形。由于肯定度不归一, 不能直接应用前面的结果。但是, 对于肯定度集合的任意元素, 总可以构造新的分布:

$$\{c_n, (1-c_n)\}, \forall n \quad (4.9)$$

显然, 式 (4.9) 永远是归一的集合, 因此, 可以应用上面的结果。于是由式 (4.5) 有

$$M_\varphi(C_n) = (c_n)^{c_n} (1-c_n)^{(1-c_n)} \quad (4.10)$$

由式 (4.7) 可以写出第 n 分量的先验形态性知识量:

$$K(C_n) = c_n \text{lb} c_n + (1-c_n) \text{lb}(1-c_n) + \text{lb} 2 \quad (4.11)$$

根据 (4.8) 可以进一步写出第 n 分量的形态性知识量公式:

$$K(C_n, C_n^*; R) = c_n^* \text{lb} c_n^* + (1-c_n^*) \text{lb}(1-c_n^*) - [c_n \text{lb} c_n + (1-c_n) \text{lb}(1-c_n)] \quad (4.12)$$

对于确定性的模糊试验来说, 可以直接写出相

应的平均知识量:

$$K(C, C^*; R) = \frac{1}{N} \sum_{n=1}^N K(C_n, C_n^*; R) \quad (4.13)$$

这样, 就建立了形态性知识量的计算或测度的方法。

注意到逻辑真实度 T 、综合逻辑真实度 \mathcal{J} 、效用度 U 、综合效用度 η 又都具有模糊集合的性质, 因此, 式 (4.9) ~ (4.13) 的演算过程可以直接应用。只要把公式中的模糊肯定度参量换成相应的逻辑真实度、综合逻辑真实度、效用度、综合效用度, 同样可以建立内容性知识、综合内容性知识、效用性知识、综合效用性知识的度量公式。(详细过程略)

$$K(T, T^*; R) = \frac{1}{N} \sum_{n=1}^N K(T_n, T_n^*; R) \quad (4.14)$$

$$K(\mathcal{J}, \mathcal{J}^*; R) = \frac{1}{N} \sum_{n=1}^N K(\mathcal{J}_n, \mathcal{J}_n^*; R) \quad (4.15)$$

$$K(U, U^*; R) = \frac{1}{N} \sum_{n=1}^N K(U_n, U_n^*; R) \quad (4.16)$$

$$K(\eta, \eta^*; R) = \frac{1}{N} \sum_{n=1}^N K(\eta_n, \eta_n^*; R) \quad (4.17)$$

4.2 知识的推理

推理, 属于对于综合内容性知识的处理。它的具体方法可以归结为三类: 相对粗糙的经验推理、比较严谨的逻辑推理、更为高级的辩证逻辑推理。不过, 由于辩证逻辑推理在目前还不成熟, 这里暂不涉及。

在一般情况下有:

推理规则 4-1 推理的基本规则形式是:

若 {前提事物的运动状态和状态变化方式为真}, 则 {结果事物的运动状态和状态变化方式为真}。 (4.18a)

或 IF {前提} THEN {结论} (4.18b)

也可用符号表示为:

$$P \Rightarrow C \quad (4.18c)$$

其中符号 P 表示推理的前提, 可以是单一前提或多个前提; C 表示推理的结论, 可以是单一结论或多个结论; 双箭头表示推理的方向和过程。容易看出, 无论是推理前提还是推理结论都是某种综合的内容性知识。因此, 确实是基于综合内容性知识的

推理。

类比推理是最常用的一种经验性推理，它的基础是“相似性原理”，其含义如下：

若事物 A 有 N 个状态，事物 B 也有 N 个状态；且事物 A 和 B 的前 N-1 个状态都相同或相似，于是，如果 A 和 B 的状态变化规律也是相同的或基本相似的，那么，结论“B 的第 N 个状态也与 A 的第 N 个状态相同或相似”也应当为真。于是有：

推理规则 4-2 类比推理规则的符号表达式是

$$\begin{aligned} & \text{IF } \{A = \{A_n\}, B = \{B_n\}, n = 1, \dots, N; \\ & \quad A_n \approx B_n = 1, \dots, N-1, C(A) \approx C(B)\} \\ & \text{THEN } A_N = B_N. \end{aligned}$$

式中符号 C 表示事物运动状态的变化规律。

严格地说，基于经验类比的推理一般不能保证推理结果必然正确，但是在经验和常识范围内却往往有效。而且经验类比推理容易理解，容易操作，容易被接受，因此，仍然是一种很有用的初级逻辑推理方法。

比较高级、比较规范、也比较理论化的逻辑推理是数理逻辑推理。它包括命题逻辑和谓词逻辑两种规范。数理逻辑推理的基本特征是：基于事物之间存在的某种因果性联系或者某种默认的关系，建立了一套严格的推理公式和推理程式。

推理规则 4-3 典型的命题逻辑和谓词逻辑推理规则包括：

$$P_1, P_2, \dots, P_n \Rightarrow \bigwedge_{n=1}^N P_n \quad (4.19)$$

$$P, P \rightarrow Q \Rightarrow Q \quad (4.20)$$

$$(\forall x)P(x), A \Rightarrow P(A) \quad (4.21)$$

运用这些基本推理规则和其它有关规则，可以相当有效地进行许多重要的推理。例如，人工智能理论中著名的“三段论”逻辑推理就可以很容易表达出来：

大前提 人都是要死的。
小前提 苏格拉底是人。
结论 苏格拉底也是要死的。

或写成数理逻辑表达式则是：

$$(\forall x)\{MAN(x) \rightarrow MORTAL(x)\}$$

$$MAN(SOCRATE)$$

$$MORTAL(SOCRATE)$$

而且可以很容易运用推理规则和归解原理证明：结论逻辑为真。

同样可以看出，在数理逻辑推理公式中，前提和结论也都是综合的内容性知识，因此，也是基于综合内容性知识的推理。

经典逻辑推理虽然已经自成体系，但是，也还有许多问题没有充分考虑。于是就导致了一系列新的逻辑系统（即所谓非标准逻辑）的陆续问世。

首先，知识通常都通过语言表达，而语言学常常包含大量的模糊现象，因此，作为对于经典标准逻辑的补充，模糊逻辑得到了越来越多的关注。模糊逻辑是通过在经典逻辑的基础上引入模糊因素而形成的。模糊因素可以表现为或者前提模糊、或者结论模糊、或者推理模糊、或前提结论和推理均为模糊的情况。例如下面的推理就是一个典型的模糊推理：

“如果（能说一口流利的普通话），那么（准是在中国住了相当长的时间。）”

在这里，推理的前提“能说一口流利的普通话”（有多流利？）和结论“在中国住了相当长的时间”（有多长？）都是模糊的；而且，由前提到达结论的推理本身也是模糊的，因此才会有“准是”（有多大的可能性？）这样的模糊表述。一般地，有：

推理规则 4-4 模糊推理的公式可以表示为：

$$\text{IF (模糊条件)} \rightarrow \text{THEN (模糊结论)} \quad \text{置信度 } (b) \quad (4.22)$$

对照前例，式中的含义一目了然，而 $0 \leq b \leq 1$ 是推理置信度（模糊度）的表示。

其次，经典的标准逻辑不能表示语言的情态，因此，作为补充，人们又提出了所谓的模态逻辑。它是通过在一阶谓词逻辑的基础上引入“必然”算符和“可能”算符而形成的。此外，为了表达逻辑推理中的时间概念，又引入了时序逻辑等等，就不一一讨论了。

最后还要特别指出，有时，一个推理规则可能会得出多个不同的推理结论，这时需要计算由前提到达各个不同结论的综合内容性知识量，选择其中最大者作为优选的结论，即：

$$\text{若有 } K(\mathcal{J}_{k_0}) = \max_k \{K(\mathcal{J}_k)\} \quad (4.23)$$

则选择第 k_0 个结论作为优选的推理结论。

可见，知识推理和知识度量之间密不可分，后者是前者必不可少的理论基础。

4.3 基于知识的决策

基于知识的决策，就是面对给定的问题、环境

和目标，利用所得到的知识选择适当的策略来求解问题，以期达到最大的得益或最小的损失。在这里，直观地看，直接发挥作用的是效用性知识，不过，由于效用性知识总是建筑在形态性知识和内容性知识的基础之上，因此，实际上发挥作用的应当是综合效用性知识。

如果把问题的“综合效用性知识量”定义为求解问题的目标函数，那么，最有利的求解策略就是能够使综合效用性知识量达到最大值的策略。这种策略就被称为“最大综合效用性知识量”决策准则。这就是一种典型的基于知识的决策。

值得指出的是，在这里，作为目标函数的“综合效用性知识量”不但直接体现了决策问题的目标——效用性知识量，而且也把环境的约束条件——问题状态的肯定度分布和逻辑真实度分布——同时表达出来了，实际上集目标函数和约束条件于一身。

具体地，假设决策问题 X 有 L 种可能的状态 $\{x_l\}_{l=1}^L$ 。又假定这些状态的肯定度分布和逻辑真实度分布分别为： $\{c_l\}_{l=1}^L$ 和 $\{t_l\}_{l=1}^L$ 。进一步，假定存在 K 种不同的求解策略 $\{a_k\}_{k=1}^K$ 。显然，不同策略不会改变状态肯定度和真实度这些客观参数，因此应当有： $c_{kl} = c_l$ ， $t_{kl} = t_l$ ，对所有 k 和 l ；但是，状态的效用度却必然因策略的不同而不同，即 u_l 应与策略有关，因而必须改写为 u_{kl} ， $\forall l$ ， $\forall k$ 。这样就可以得到一个决策矩阵：

$$\begin{matrix} x_1 & \cdots & x_l & \cdots & x_L \\ c_1 & \cdots & c_l & \cdots & c_L \\ t_1 & \cdots & t_l & \cdots & t_L \\ a_1 & \begin{bmatrix} u_{11} & \cdots & u_{1l} & \cdots & u_{1L} \\ \vdots & & \vdots & & \vdots \\ a_k & \begin{bmatrix} u_{k1} & \cdots & u_{kl} & \cdots & u_{kL} \\ \vdots & & \vdots & & \vdots \\ a_K & \begin{bmatrix} u_{K1} & \cdots & u_{KL} & \cdots & u_{KL} \end{bmatrix} \end{bmatrix} \end{matrix} \end{matrix} \quad (4.24)$$

由式 (4.24) 所示的决策矩阵可以建立相应于各个策略和状态的综合效用度：

$$\eta_{kl} = c_l t_l u_{kl}, \forall l, \forall k \quad (4.25)$$

至此就可以利用式 (4.11) 分别构造它们的先验综合效用性知识量。这里要根据 $\{\eta_{kl}\}$ 是否归一区分两种情况：如果归一，则有

$$K(\eta_k) = \sum_{l=1}^L \eta_{kl} \text{lb} \eta_{kl} + \text{lb} L, \forall k \quad (4.26)$$

否则有

$$K(\eta_{kl}) = \eta_{kl} \text{lb} \eta_{kl} + (1 - \eta_{kl}) \text{lb}(1 - \eta_{kl}) + \text{lb} 2, \forall l \quad (4.27)$$

$$K(\eta_k) = \frac{1}{L} \sum_{l=1}^L K(\eta_{kl}), \forall k \quad (4.28)$$

由此可以得到如下的：

决策规则 4-1 根据“最大综合效用性知识量”准则，基于综合效用性知识量的决策规则可以表示为：

$$\text{若 } K(\eta_{k_0}) = \max_{\forall k} \{K(\eta_k)\} \quad (4.29)$$

则 选择策略 k_0

由公式 (4.26) ~ (4.29) 所描述的基于综合效用性知识量的决策规则是具有普遍意义的决策规则。可以看出，有了综合效用性知识，就有可能做出科学合理的决策。

不过也要看到，一方面，公式中有一些参数（如肯定度和真实度）是通过统计或其它方法得到的，具有一定的客观性；但是，另一些参数（特别是效用度）的赋值则存在很强的经验色彩。因此，一个好的决策者不但要具有决策理论的科学知识，也要具有比较丰富的决策经验和决策艺术，这样才能真正做出好的决策。这就是为什么人们经常说“决策是一门科学，同时又是一门艺术”的道理。

颇有意思的是，如果把这个规则作一些简化处理就可以发现，历史上已有的许多重要和著名的决策规则实际上都可以在一定意义上看作“基于综合效用性知识量的决策规则”这个普遍规则的简化规则。

先来简化基于知识的决策规则。例如，把式 (4.26) 简化为

$$K(\eta_k) \approx \sum_{l=1}^L \eta_{kl} = \sum_{l=1}^L c_l t_l u_{kl} \quad (4.30)$$

进一步，若令 $t_l = 1, \forall l$ ，即认为所有状态都为真，同时假设决策问题是随机性事件，那么，肯定度就退化为概率，式 (4.30) 简化为：

$$K(\eta_k) = \sum_{l=1}^L p_l u_{kl}, \forall k \quad (4.31)$$

于是，决策规则 4-1 简化为：

决策规则 4-2

$$\text{若有 } \sum_{l=1}^L p_l u_{k_0 l} \geq \sum_{l=1}^L p_l u_{kl}, \forall k \quad (4.32)$$

则选择策略 k_0 ；否则就选择其它策略。

这恰好就是决策学著名的 Bayes 决策规则。显然，它是“基于综合效用性知识量决策规则”的一个特例。

同样，容易证明，著名的极大极小决策规则也是基于知识的决策规则的简化规则。

决策规则 4-3 极大极小决策规则是：

$$u_{\text{opt}} = \max_k \{ \min_l [u_{kl}] \} \quad (4.33)$$

它的基本操作过程是：在决策矩阵 $\{u_{kl}\}$ ， $k=1, \dots, K$ ， $l=1, \dots, L$ 中，首先找出每一行各个元素 u_{kl} ， $l=1, \dots, L$ ， $\forall k$ 的最小值（即每个策略的最小得益），然后，在所有各行的最小值之间挑选最大者（即挑选“最佳”策略）。可见，极大极小决策规则连每个策略的平均得益都不计算，只关心每个策略的最小得益（因此是最保守的规则）。它所选择的策略实质是极大的极小得益，因而有“极大极小”决策规则的名称。

显然，极大极小决策规则是基于知识的决策规则的进一步退化结果。

5 知识的生成：知识与信息的关系

“知识论”的两个主题——“知识生成”和“知识激活”，应当说，乃是人类科学实验活动的基本问题：人们通过科学实验来发现新的知识（知识生成）；并运用所获得的知识来解决所面临的各种复杂问题（知识激活）。因此，本节所研究的内容正是人类科学实验活动的基本规律。

本文的主要着眼点是机器。换言之，这里研究的目的是试图使机器能够具有“知识生成”和“知识激活”的能力。这里所采用的研究思路是：借鉴人类科学实验活动中的“知识生成”和“知识激活”机制，寻求能够在机器中实现的机理。

知识的生成有两个基本途径。一个是从实践中逐步积累，把所观察到的现象的共性核心升华成为概念，把所积累经验的精华上升成为理论；另一个途径是由已有的知识通过推断产生新的知识。前者是由信息到知识、由具体到抽象的过程，在逻辑上称为归纳；后者是从知识到知识、由抽象到抽象的过程，在逻辑上称为演绎。

在人类社会进步的早期阶段，人类所拥有的知识量很少，此知识和彼知识之间难以建立相互的联系，演绎的途径难有大的作为。因此，在科学发展的初期阶段，人类主要通过归纳的方法一点一滴地生成知识和积累知识。随着人类所拥有的知识越来越丰富，知识之间的距离越来越接近，通过理论思维从已有知识演绎出新知识就逐渐成为知识生成的重要手段。这就是近代科学理论能够得到迅速发展

的原因。当然，从发现新知识来说，无论演绎怎样重要，归纳永远是基本的途径，而且是具有永恒生命力的途径。

实际上，演绎和归纳是相辅相成的知识生成手段。在现代科学技术发展的条件下，任何复杂规律的发现和任何新知识的生成都不可能仅仅是单纯的归纳或单纯的演绎，而必然是归纳和演绎的辩证互动。这是现代一切科学实验的基本特征。

只是考虑到演绎问题已经在人工智能理论中得到深入的研究，而本节的目的是要探讨信息与知识的关系，因此，将主要关注归纳型的知识生成问题。

从第二节关于信息和知识定义的讨论中可以知道，信息表达的是事物运动的状态以及状态变化的方式，知识表达的是事物运动的状态以及状态变化的规律。因此，由信息生成知识的归纳过程本质上就是一个由个别事物运动状态的具体变化“方式”升华为一类事物运动状态的普遍变化“规律”的抽象化过程，即飞跃过程。

显然，不应当笼统地研究知识生成的问题，因为不同类型知识的生成有着各不相同的规律。

5.1 形态性知识的生成机制

现在，人们已经普遍公认，知识是由信息通过加工提炼生成的。问题在于，究竟通过什么具体机制才能实现把信息加工成为知识的过程？

经验表明，由信息生成知识的过程通常非常复杂，而且生成的机制也远不是唯一的。这里将分别探讨形态性知识、内容性知识和效用性知识的生成机制。可以想见，形态性知识的生成机制相对简单。另外，概念是知识的基元，知识生成必须从概念生成开始。

研究发现，由信息生成形态性概念知识的基本算法可以归结为如下的

生成机制 5-1 形态性概念知识生成机制是一类形式对比归纳过程。具体算法如下：

1) 观察一个信息样本 S_1 ，尽可能多地提取它的形式特征，包括它的运动状态的形式特征（如大小、高度、重量、形态、颜色、所发声音的频谱以及其它可以作为信息特征的参量）和状态变化方式的形式特征（如随机性、偶发性、模糊性等）。

2) 观察第二个信息样本 S_2 ，也提取它的形式特征，并与第一个信息样本的信息特征相比较，保留二者共有的特征，去掉不一致的特征。

3) 把这个步骤重复 N 次, N 是一个足够大的正整数, 从而得到一组具有共性意义的信息特征 $\{f_k\}$, $k=1, 2, \dots, K$, 它们构成一个“共性特征集合”, 其中 K 也是一个正整数, 随着 N 的增大, K 可能会增大, 也可能会减小, 但总远小于 N 。

4) 排除不具备 $\{f_k\}$ 的信息样本; 保留的信息样本数为 N' 。当观察的信息样本数 N 和 N' 再进一步增加而这些共性特征集合 $\{f_k\}$ (包括特征的内容和数目 K) 保持不变的时候, 就认为共性特征集合已经稳定。

5) 称所有满足共性特征集合的信息样本为同“类”信息样本。给这样形成的“类”命名, 从而形成了一个具有“类”名称的概念, 完成了一个概念的生成。

同样, 也可以建立其它的“类”概念。当然, 由于观察样本数量的有限性, 在以后的应用过程中, 也存在对“类”概念的“共性特征集合”进行局部调整的可能性。但是, 通过这样的调整, “共性特征集合”一般应当能够稳定下来。否则, 这个概念就有问题, 需要返回第一步重新建立。

概念是知识的基元, 因此, 由信息样本建立“类”概念的过程, 就可以看作是基本知识单元的生成过程。有了“类”的概念, 就具有了分类的知识, 也就具有了利用“类”的“共性特征集”对新的对象进行分类的能力。

我们可以这样来表述利用生成机制 5-1 所生成的形态性知识: 某个概念(知识)的内涵就是与之相联系的“共性特征集合”, 而这个概念(知识)的外延则是满足这个“共性特征集合”条件的信息样本全体。在这里, 在所观察的信息样本中发现和建立它们之间的“运动状态和状态变化方式的共性形式特征集合”是由个别现象到一般规律的抽象化的关键环节。

形态性概念知识的生成机制比较简单。但正因为它比较简单, 就比较容易看清楚其中最基本的机理, 因此也就成为了研究其它知识生成机制的基础。而且, 作为“知识论”的框架性研究, 阐明概念知识的生成机制是很有意义的。

除了“概念”生成之外, “关系”生成也是形态性知识生成的重要组成部分。这正是当前正在兴起的“数据挖掘(Data Mining)”的研究内容。由于这方面的资料比较容易找到, 又由于本文的篇幅

有限, 这里不再展开。

5.2 效用性知识的生成机制

按照知识的层次关系, 在讨论形态性知识的生成机制之后, 应当接着讨论内容性知识的生成机制。但是, 这里却需要先讨论效用性知识的生成机制。原因是: 从认识过程来看, 形态性知识和效用性知识都可以从外部信息中直接感知, 而内容性知识则只能通过对形态性知识和效用性知识的分析得到。因此, 需要先讨论效用性知识的生成机制。

为了具体明确“效用”的概念, 本文约定: 所谓一个信息对某个主体有没有“效用”, 是指这个信息“对于实现主体的目标有没有贡献”。而“效用度”则是用来度量这种贡献程度的一个参量。可见, 效用是与主体目标相联系的概念。

在一般的情况下, 效用性知识可以通过对信息的加工提炼直接得到, 在一些情况下也可以通过经验推理或理论演绎间接得到。同样, 由于本文特别关注信息与知识之间的关系, 因此这里将主要论及前者。

不难理解, 由信息生成效用性知识的过程要比生成形态性知识的过程复杂。原因是这里不但要利用形式化的比较, 更重要的是, 按照上面关于“效用”概念的约定, 生成效用性知识的主体(人、生物或机器)必须具有明确定义的目标, 还要有能力来判断某个信息究竟是有利于实现目标还是有碍于实现目标。有利于实现主体目标的信息, 具有正的效用度; 有碍于实现目标的信息, 具有负的效用度。当然, 要真正判断一个信息对于实现主体目标究竟是“有利”还是“有害”, 也不是一种简单的形式上的比较, 而更可能的是要在实践中直接承受它所引起的真实后果——或者是收到某种得益, 或者是付出某种代价。

下面是由信息生成效用性知识的过程的一个算法描述。

生成机制 5-2 由信息生成效用性知识的算法包括如下的步骤:

1) 明确定义主体追求的总体目标 G ; 并且把这个总体目标分解成为一系列容易检验的具体目标 $G = \{G_n\}_{n=1}^N$ 。任何一个具体目标都可以看作是目标状态空间中的一个状态 G_n , $\forall n$ 。

2) 输入信息 X , 主体首先记住 X 的形态描述 $D(X)$ 。然后计算在目标状态空间中信息 X 所反映的现实状态与主体目标状态之间的距离, 如:

$$d_n(t_m) = \{[X(t_m) - G_n(t_m)]^2\}^{1/2},$$

$$\forall n, m = 0, 1, 2, \dots \quad (5.1)$$

其中 $d_n(t_0)$ 是信息的当前 ($t = t_0$) 状态与第 n 个具体目标状态之间的距离, m 是时间坐标点。 $X(t_m)$ 和 $G_n(t_m)$ 分别是在第 m 时刻的信息状态和第 n 具体目标状态。

3) 根据信息 X 的状态变化方式, 考察信息所反映的未来状态与主体目标状态之间的距离的平均变化趋势, 若有

$$E\{d_n(t_{m+i})\} \leq E\{d_n(t_m)\}, \forall i, \forall m \quad (5.2)$$

则赋值

$$u_n \geq 0$$

反之, 则有 $u_n < 0$ 。式 (5.2) 中的符号 E 表示对所有的状态 n 取平均, 符号上面的横杠表示在时间域上的平均。

4) 由效用度分量 $u_n, n = 1, \dots, N$, 可以计算出整体效用度 u 。

5) 对于新信息, 如果它们的形态描述与 $D(X)$ 相同或相近, 就给效用度赋予相同的值。否则, 就返回第 2 步重新判断。

在效用性知识的生成机制 5-2 中, 第 2 步是关键。在这里, 信息状态和主体目标状态之间的差异被定义为欧几里德距离。实际上, 距离的定义要根据具体情况的情况确定。在有些情况下, 也许很难给出距离的确切定义和算法。但是, 无论如何, 生成机制 5.2 的基本原则——通过评价信息与主体目标之间的差异以及这个差异的发展趋势来判断信息对于主体的效用, 这在逻辑上是完全合理的, 而且在实践上是有意义的。

例如, 一个儿童头一次看见“狗”这种动物。首先, 他看到并且记住了“狗”的形象。同时他就会捉摸, 具有这种形象的狗同自己的目标——在这里是自身的安全——之间存在的某种联系(计算这之间的距离): 是一种保护还是一种威胁? 如果狗露出了锋利的牙齿, 眼里射出了凶光, 这种状态就会使他算出负的效用度, 决心逃避或求援。此后, 如果再见到这种动物, 他心中就会有“负效用度”这个数了——获得了效用性知识。

在这个例子中, 谁也不知道这位儿童究竟是怎样定义他的目标(安全)与他看到的信息状态(狗)之间的距离的, 也不知道他是怎样进行具体的计算的。但是, 毫无疑问, 他确实进行了这样的“捉摸”, 并且得出他自己的结论。

5.3 内容性知识的生成机制

如前所说, 内容性知识是一种抽象的知识, 它的生成机制也更为复杂。图 5-1 是内容性知识生成机制的一种可能的实现途径。

图 5-1 表明, 内容性知识的生成要以形态性知识和效用性知识为前提条件。那么, 在生成了形态性知识和效用性知识的基础上, 怎样进一步生成相应的内容性知识呢?

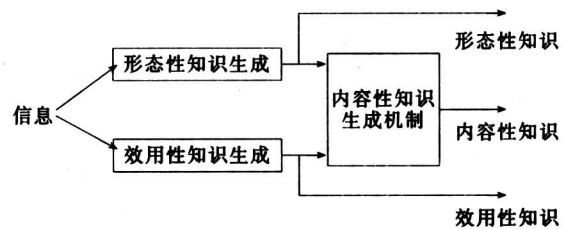


图 5-1 内容性知识生成的功能性模型

Fig.5-1 The functional model of content knowledge production

内容性知识的表现方法多种多样, 一种最简单的方法就是通过逻辑“与”的直接陈述。如果要陈述一个概念的内容, 就把与这个概念相联系的形态性知识和效用性知识按照逻辑“与”的规则直接联系起来。具体的形式可以表达为:

“它的概念类属: 它具有如此这般的形态, 且具有这般如此的效用。”

比如, 怎样描述“苹果”这一概念的内容? 按照逻辑“与”的规则可以表示如下:

”苹果是一种水果(概念类属), 它具有红润的颜色, 球状的形态, 拳头般大小, 肉质甜脆而多汁(形态描述), 且可以食用, 有益健康(效用描述)。”

由此, 可以构造如下的

生成机制 5-3 由信息生成内容性知识的过程可以描述如下:

- 1) 利用生成机制 5-1 由信息生成形态性知识 K_C ;
- 2) 利用生成机制 5-2 由信息生成效用性知识 K_U ;
- 3) 内容性知识的最简单生成机制是: 给信息归类, 并在形态性知识与内容性知识之间建立一个映射关系:

$$CONT: K_C \rightarrow K_U. \quad (5.3)$$

可见，它把整个知识的生成机制都集成到一起了。

当然，同形态性知识和效用性知识的情形一样，内容性知识也可以通过逻辑演绎推理的方法生成。不过，本文主要的兴趣在于探讨直接由信息生成知识的机制，探讨信息 - 知识 - 智能的关系。关于由知识生成知识的演绎方法，将在另文讨论。

应当强调说明的是，知识论博大精深，决无可能在一篇论文中穷尽。以上关于知识生成的论述是框架性的。给出这些结果的目的是希望由此激起更多的热忱，引出更多的研究，逐步丰富和完善知识理论的内容。

6 知识的激活：知识与智能的关系

知识的重要性人所共知。但是，从人类活动的总体目标来说，也应当承认，获得知识并不是我们的最终目的。这是因为，知识本身并不能直接促进社会的发展。知识只有在被用来解决各种复杂的实际问题的时候，才能真正发挥出积极的作用。

按照本文第二节给出的定义，针对给定的问题、问题的环境和目标，有效地获得与问题和环境相关的信息，恰当地处理这些信息以生成相应的知识，并在目标的引导下由知识再生成求解问题的策略，然后利用所生成的策略成功地解决所面对的实际问题，这正是“智能理论与技术”的任务。可见，知识与智能之间的转化是通过形成求解问题的策略来实现的。而求解问题的策略又是在求解问题的目标的引导下由相关的知识所生成（称之为“再生”）。

为了确证这一点，不妨来考察一下图 6-1 所示的人工智能系统的功能结构的模型。

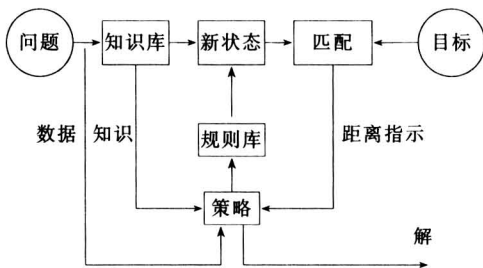


图 6-1 人工智能系统模型

Fig.6-1 Artificial intelligence system model

图 6-1 描述的是一个人工智能系统的典型工

作过程：在控制策略的调度下，规则库给出一个具体的规则，作用于体现问题状态的知识库，使知识库形成一个新的状态，这个新状态与问题的目标状态之间的差距及其增大减小的状况则由差距测度指示出来。如果差距缩小了，说明控制策略的效果是积极的；反之则是消极的。把这个差距测度反馈给控制策略单元，后者就可以根据这个效果对控制策略不断进行改进，使差距测度不断朝减小的方向变化。一旦这个差距的测度变成了零，或者变得足够小了，就认为问题得到了满意的解决，系统就可以把这个策略以及相应的规则和知识作为问题的解答输出给用户。

在人工智能系统中，知识库的知识和规则库的规则是由系统设计者事先设定的，是一类先验的知识；数据库的数据和求解目标则是随着问题的输入而给定的。作为知识的存储载体，知识库、规则库的确都是人工智能系统不可缺少的组成部分，但是，面对具体的问题，更为要紧的是需要有一个巧妙的控制策略来调度这些知识：确定在哪个进程上，调用哪一个规则，作用于数据库和知识库中的哪个知识状态，才能一步一步地把问题从当前的状态转变到求解的目的状态，使问题得到最终的解决。

在这里，控制策略体现了系统的智能，成为整个人工智能系统的指挥者与组织者，使系统能够面对具体的问题有效地组织知识和调度知识，使知识得以激活。从这个意义上可以说，没有控制策略的调度和激活，知识原本是刻板的。控制策略使死的知识变成了活的知识，变成了真正有用的知识。因此，把知识转变成智能策略的过程，就是知识激活的过程。

当然，控制策略本身也可以看作是一种知识，是一种利用知识的知识，组织知识和调度知识的知识，或称为“元知识”。这种“元知识”与其它的知识不相同的地方就在于它是一种主动的知识。因此，我们特别感兴趣的是，怎样才能把一般的知识激活成为能够有效地解决问题的主动的知识。这是本节要探讨的中心问题。

直到现在，控制策略（元知识）的形成还没有满意的方法。比较成熟的是搜索方法，包括盲目搜索方法和启发式搜索方法。

盲目搜索方法通过逐一的尝试，才能最终找出解决问题的控制策略——求解路径。它的最大缺

点，就是没有利用与问题有关的任何知识，因此是智能度最低的方法。

启发式搜索方法则利用问题的效用性知识来建立控制策略。它巧妙地利用了图 6-1 的“差距测度”。在这里，“差距测度”不仅要给出“有没有差距”的指示，同时还给出“差距有多大”的评估测度，从而可以由此判断现有策略的优劣；并且还可以根据“差距增大或缩小”的变化情况，指示出策略改进的方向。显然，这种评估测度实际上就是本文前面所讨论过的一种效用性知识的测度。可见，知识测度方法也是研究智能的重要理论基础。

这里启示了一个重要的、带有普遍意义的机理，这就是：效果（效用性知识）的反馈是非常重要的。一旦系统能够把自己的行为和这种行为所导致的效果联系起来，系统的工作就有了明确的方向。

图 6-2 示出的就是这种反馈调节的机制。通过“控制—反馈—判断—调整”的动态机制，控制策略便逐步得到了优化，知识也由此得以被调度被激活，求解问题的智能也在其中得以形成，问题便在这个过程中得到了满意的解决。

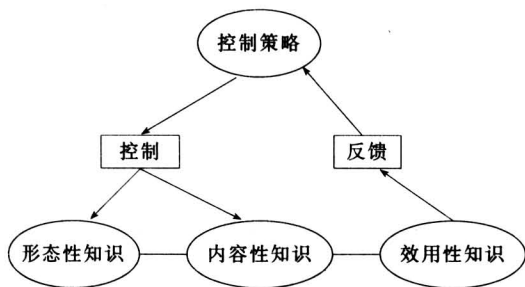


图 6-2 知识与策略（知识激活）

Fig. 6-2 Knowledge and strategy (knowledge activation)

总结图 6-1 和图 6-2，可以得出“知识激活”的基本原理：

第一，按照智能的基本定义，控制策略集中体现了求解问题的智能，因此，把知识激活成为智能，就是要在目标引导下由相关知识生成求解问题的控制策略。

第二，在目标引导下由知识生成策略的基本方法，是不仅要在控制策略单元与体现问题求解状态的知识单元之间建立控制联系，还要在知识单元

与控制策略单元之间建立控制策略效果（即效用性知识）的反馈联系。

第三，控制策略的效果的好坏，可以利用综合效用性知识测度的方法来测度。

第四，知识单元的知识要表示出形态性知识、内容性知识和效用性知识。在这里，形态性知识和内容性知识主要用来显示问题求解的现时状态，效用性知识则用来反馈给控制策略单元，用以判断和修正控制策略。

可见，形态性知识、内容性知识、效用性知识的表示与度量，对形态性知识和内容性知识的调节与控制，以及效用性知识的反馈，构成了激活知识生成策略的基本要素。没有知识，没有对于知识的控制和反馈，生成策略就成为空中楼阁。

当然，作为一种完整的理论，知识激活是智能理论要研究的一个很大的课题，这里提出的激活原理只是其中一种可行的机制，更丰富的结果有待进一步的研究。但是，从图 6.2 的分析确实已经可以看出，只要给定明确的问题和目标，知识是可以被激活的，知识是可以通过控制策略的形成而转变成成为有用的智能的。

反过来，我们也确实可以清楚地看到：没有信息，便不可能有知识；没有知识，便不可能有智能。

7 结语：信息－知识－智能的统一理论

不无遗憾的是，在此之前，学术界还没有建立起“知识论”。

虽然，人工智能领域曾经一度探讨和建立过“知识工程”，然而，正像它的名称所表明的那样，“知识工程”主要关心的是知识的实际工程操作方法，没有全面涉及和解决知识本身的理论问题，更谈不上揭示“信息－知识－智能”三者之间内在的本质联系。这方面最明显的例证之一是：在知识工程的典型成就——专家系统的研究中，系统的知识仍然主要依靠人工的方法获取，因而成为专家系统研究中的一个明显的瓶颈。另一个明显的例证是，虽然知识工程研究了各种各样的人工智能专家系统，然而，并没有从中提炼出如何把知识激活成为智能的一般性理论。

当今世界科学技术的迅猛发展，已经对知识理论的研究提出了紧迫的需求。研究和建立知识理论

的任务,已经时不我待。有鉴于此,本文的工作集中到一点,就是从定性和定量两个方面为研究知识理论建立起必要的基础,进而在此基础上阐明知识理论的两个核心命题:第一,如何通过归纳把信息提炼成为“知识”;第二,如何通过激活把“知识”升华成为“智能”。由于知识论框架的建立,使信息论与智能论之间长期隔断的相互联系得以沟通。因此,这个结果也必然会成为“信息—知识—智能”统一理论的核心和灵魂。当然,本文提出的知识论框架仅仅是一个初步的尝试,还有大量的工作需要进一步研究。

令人欣慰的是,不久前,作者和他的学生们把知识论的思想成功地应用于国家“八六三”计划项目“智能型机器自动文献摘要系统——LADIES”的研究,取得了令鉴定委员会专家们甚感满意的成果^[7]。我们衷心希望,通过学术界的共同努力,“信息—知识—智能的统一理论”会很快地日臻完善,成为 21 世纪这个信息—知识—智能社会的重要学术支柱。

致谢 本文的研究有幸得到中国工程院院长宋健院士的关注、鼓励和指导,文中关于科学定律和定理的逻辑判断以及科学实验与知识论的关系等重要问题都是经他提示才注意到。本文的研究还得到王守觉院士、陈太一院士、涂序彦教授、以及国际“系统学—控制论—信息学”大会主席 N. Callaos 教授等的帮助,作者在此一并致以衷心的感谢。

参考文献

- [1] 钟义信. 信息科学原理 [M]. 福州: 福建人民出版社, 1988; 再版. 北京: 北京邮电大学出版社, 1996
- [2] Wiener N. Cybernetics [M]. Elsevier Press., 1948
- [3] Shannon C E. Mathematical Theory of Communication [M]. BSTJ, 1948
- [4] 钟义信, 等. 智能理论与技术: 人工智能与神经网络 [M]. 北京: 邮电出版社, 1992
- [5] Aczel J. Lectures on functional equations and their applications [M], Academic Press, New York, 1966
- [6] Hardy G H, Littlewood J E, Polya G. Inequalities [M]. Cambridge University Press, London, 1973
- [7] 杨晓兰, 钟义信. 基于理解的自动文摘系统的研究与实现 [J]. 电子学报, 1998, (7): 96~99

A Framework of Knowledge Theory: Toward a Unified Theory of Information, Knowledge and Intelligence

Zhong Yixin

(University of Posts and Telecommunications, Beijing 100876, China)

[Abstract] Knowledge has been very important wealth to the mankind but there has not a knowledge theory existed yet till the present time. An attempt is thus made in the paper to present a framework of knowledge theory that includes two parts: fundamentals and the main body of knowledge theory. The first part is to deal with a series of basic issues such as the related concepts and definitions, the methods of representation, the measurements, the reasoning and decision rules. The second part is to explore the mechanism of knowledge formation based on information processing and the mechanism of intelligence formation based on the activation of knowledge. It is believed that the establishment of the knowledge theory will lay a solid foundation to the unified theory of information, knowledge, and intelligence and will greatly facilitate the effective utilization of information and knowledge, leading to the growth of the research in the field of intelligent machines.

[Key words] knowledge; amount of knowledge; knowledge formation; knowledge activation; unified theory of information-knowledge-intelligence