

# 软计算和知识获取 Soft Computing and Knowledge Acquisition

苏运霖  
Su Yunlin

(广西大学梧州分校 梧州 543002)  
(Wuzhou Branch of Guangxi University, Wuzhou, 543002)

**摘要** 讨论在人的浩瀚知识中, 哪些属于不精确和不确定性的知识, 并且说明这些知识可以通过软计算来获取的。然后以示例形式, 论述如何使用软计算的方法获取这些不精确、不确定的知识。

**关键词** 软计算 知识获取 不精确性 不确定性

**中图法分类号** TP301.4

**Abstract** It is discussed which knowledge belongs to imprecise and indeterminate ones among the vast illumination of knowledge, and is illustrated that these pieces of knowledge may be acquired via soft computing. Subsequently, it is exemplified how to obtain the imprecise and indeterminate knowledge by way of soft computing.

**Key words** soft computing, knowledge acquisition, imprecision, indetermination

## 1 问题的提出

自从 20 世纪 90 年代初, 模糊逻辑的奠基者 Lofti A Zadeh 在他的模糊逻辑理论的基础上, 创造性地提出软计算的概念以来, 软计算已经获得了极其引人注目的发展<sup>[1]</sup>, 而且成为一个逐渐扩展的方法学联合体。这些方法学都以处理不精确性和不确定性以及部分正确性等等作为对象, 旨在实现问题求解, 并且达到必要的健壮性、可处理性以及解决问题的低费用。在软计算刚刚被提出时, 它被认为是由模糊逻辑、神经逻辑和概率推理组成的。但软计算提供的以现实的不确定性来表示人的思想的二义性, 为各种各样类似的方法提供了机会。因此人们在采用软计算所涉及的模糊逻辑、神经网络(计算)和概率推理(计算)的同时, 又纷纷提出进化计算、混沌计算以及免疫网络等等的方法学, 并且把它们归入软计算的范畴, 使软计算成为范围更加广阔, 应用更加有效的方法学联合体。

因此软计算正在引起工程和科学领域的一个突破, 因为它能解决由传统的分析方法所不能解决的问题(可解性)。此外, 软计算还产生了丰富的知识表示(符号和模式), 灵活的知

识获取(通过数据以及通过和专家的交互的机器学习),以及灵活的知识处理(通过符号和模式知识界面的推理),它们使得以低的费用来构造高智能系统(产生高的机器智商(High Machine Intelligence Quotient,简称HMIQ))<sup>[2]</sup>成为可能。

由于软计算仍然是一个不断扩展的方法学联合体,而且它们都特别适合于处理现实生活中的不精确性和不确定性,因此人们自然预期在可能的条件下,把这些具有共同特性的方法学组合起来加以应用,以进一步增强它们解决问题的能力。如模糊计算的基础特性是通过对于模糊事物的判定,得出事物的本质和问题的结论。它的基本功能是实现符号知识和模式知识的高度匹配。神经网络的基本特性是通过学习,有能力来标识非线性的模型。概率推理的基本特性是运用概率推理确定出最有可能的某种模式。进化计算的基本特性是对于非线性优化问题的有效求解能力。至于混沌计算的基本特征,是在非线性和混沌的状态下,确定一种可能的知识模式。由此可以看出,它们确实都具有不精确性和不确定性的特征。

反观知识,知识有两种类型,一种是确定和精确的,这要求运用严格的分析和推理的方法来获取,以及对它们的确定和精确的记忆。但还有一种类型,它本身具有不精确和不确定的特性,因此对它的获取和记忆,也要运用不精确和不确定的方法。可以认为,对这类知识的获取,软计算是可行和有效的方法。本文旨在分析这样一类知识。文中首先分析这类知识的特性,然后讨论运用软计算方法来获取这些知识的步骤和过程,最后,提出了若干思考的问题。

## 2 一些知识的不精确性和不确定性

人类在其发展的历史长河中,已经经历了长达七八千年有文明记载的历史,因而也积累了浩瀚无际的各种各样的知识。在这些知识当中,有许多知识是精确的,确定的,但也有同样多乃至更多的知识是不精确和不确定的。既然知识本身是不精确和不确定的,获取这些知识的方法自然也不会是精确地和确定的。道理很简单,如果我们把不精确和不确定的知识当作一个目标集合,人们只要把箭射入这个集合中的任何位置,都算是击中了目标。那么就很简单,射进这个范畴的途径就不确定、不精确和不惟一的。因此,分析哪些知识属于不精确和不确定的,对于研究如何用软计算方法来获取它们是有用的,这里列举的,并不是穷尽的,但或许有相当的代表性,因为我们引用的许多例子取自于中国古代包括《道德经》在内的这样充满哲理的经典著作<sup>[3]</sup>,以及也被称为哲学教科书的法国著名作家巴尔扎克的论著<sup>[4]</sup>。

### 2.1 动态的知识

人类社会本身是发展的,动态的。《庄子秋水》中指出“物之生也,若骤者驰。无动而不,无时而不移。”由于这种动态性,变化性,人们通过“快照”摄取的知识,只能是某一个时间片断的知识。对这个时间片断稍作变更,则所得到的已是变化了的该过程的知识。因此由前后的2个时间片所得到的“快照”即动态的知识就不可能完全相同,也就谈不上精确和确定。

### 2.2 时态逻辑

和上述动态知识类似,涉及时态逻辑的知识也属于不精确和不确定的。因为根据时态本身的定义,可以以时间点来定义,也可以以时间范畴来定义。就后者而言,也就有一个不精确和不确定的问题存在。就时间点来说,在某个时间点来说不可能的事,在另外一个时间点却是可能的。我们经常听说的某个运动员是某项记录的保持者,这就是不精确和不确定的,只

要这项纪录是这个运动员创造的，而至今这项纪录仍旧未被其他人打破，那这句话就仍然是正确的。但所谓保持，可以理解为他仍然可以达到这项运动水平，但也同样可能，这仅仅是他过去的一个辉煌。又比如说2个男女，男子30岁，女孩2岁，这时他们当然不可能谈婚论嫁（假设2岁女孩确非男子的女儿或直系亲属），但是20年后，或之后某个时候，他们完全有可能结为良缘。在时态逻辑中，要描述的东西，就包括这样的现象，就是在某个时刻不可能的事件，在过了一段时间之后，它却完全可能。或者反过来，在某个时刻可能的事，再过一段时间之后，却成为不可能的了。

巴尔扎克在《乡村医生》曾说“一个22岁的青年设想所有的女人都是柔情脉脉的，他还不会区别爱情和情欲。”又在《猫打球商店》中说：“在18岁的年龄，爱情哪有不在一个少女眼睛和外部世界之间放上它的七色三棱镜的。”从这两句话来看，就可以知道时态逻辑的不精确和不确定性，18岁和22岁，并非完全确定的，只要处于该年龄段上，所描述的事实就大体上是正确的。同时，过了这个年龄段，或在这个年龄段之前，就大体上不成立了。

### 2.3 比喻

列宁曾经说过，“比喻是蹩脚的”。但是人类又有许多知识，是以比喻形式描述的，比如说，《道德经》上有“上善若水”这句话，把最高的善行比喻作好像水一样。然而，水并非什么时候都可以作为好的化身。《资治通鉴·唐记十三》中指出“水所以载舟，亦所以覆舟”。这大概就是列宁说的“蹩脚”的意思。然而比喻的用处，就在于通过所作的比喻，来增强所述道理的形象性和说服力。巴尔扎克在《人间喜剧》里说“忍耐，是支持工作的一种资本”，“逆运，不就是性格的试金石吗”，“只要你善于等待，一切将及时来到”。又说“不幸，是天才的进身之阶；信徒的洗礼之水；能人的无价之宝；弱者的无底之渊”，“苦难对于天才是一块垫脚石，对能干的人是一笔财富，对弱者是一个万丈深渊。”通过这些话，一方面人们会加深对于不幸和苦难的理解，但也可能造成一头雾水，茫然不知所云。

### 2.4 辩证统一规律

作为一种科学学说，辩证唯物主义是我们认识事物的强有力的武器。按照这一理论，任何事物都是矛盾统一体。这种理论，是人类对于客观世界和主观世界认识的总结，是古今中外许多人都认同的。《道德经》里说“有无相生，难易相成，长短相形，高下相楹，音声相和，前后相随。”它把有无、难易、长短、高低、不同的声音以及前后加以归纳为一体。此外，当然还有远与近、善与恶、古与今、中与外等等，也都是辩证统一的。巴尔扎克在《幻灭》里也说“一切自然规律都有双重的相反的作用”。又比如，《道德经》里说“少则得，多则惑”。巴尔扎克在《驴皮记》里说“过多的知识同无知一样，得到的结果是否定”。这里，也都存在着不精确性和不确定性。因为对于相反的那两方面，如多与少，或善与恶等等，并没有明确的量化的含义。

### 2.5 进化论与逆进化论

达尔文所创立的进化论学说，也揭示了自然发展的规律，即“生物竞天演，适者生存”，任何事物都经历其发生、发展和消亡的不同阶段。这个理论，也为中外科学界所公认。巴尔扎克在《公务员》里指出“世间万物中，不论是赫赫有名的还是默默无闻的，也不论是个畜牲，还是个秘书长，都有其盛衰，沉浮，都有一个皮毛光润或是福星高照的时期。”

作为进化论的补充，可以提出逆进化论，即如果 $A$ 是由 $B$ 进化而来，表示为 $A \rightarrow B$ ，若 $B$ 是 $A$ 的逆进化，表示为 $B \rightarrow A$ ，或 $A \geq B$ 。我们古代的学者早就应用这一逆进化论。如《道德

经》上指出“合抱之木，生与毫末；九层之台，起于垒土；千里之行，始与足下。”又说“贵以贱为本，高以下为基。”而张华在《高以下为基》中，更明确地阐述“高以下为基，洪由纤起，川广自源，成人在始。”

然而，无论是进化论还是逆进化论，显而易见，也都存在着不精确性和不确定性。

## 2.6 相对论和绝对论

唯物主义者认为，世界上没有绝对真理，而只有相对真理。人类在认识的长河上发现的，总是一个个的相对真理，而只有其极限才是绝对真理。但人类永远也达不到绝对真理的阶段。由于真理的这种相对性，也就引起了不精确性和不明确性。巴尔扎克说：“从伟大到可笑，相差只有一步”。列宁也说过：“真理超越了一步，就成了谬误”。同样意思的话，出现在中国古代的《周子》，“失之毫厘，差以千里”。

巴尔扎克在《纲巴拉》中指出，“既是一，又是多，这是一切伟大的自然事物的特性。”这里所说的也是相对和绝对的关系，一不就是一，它还是多；多不就是多，它还是一。所以世界万物都是相对而存在。这同样包含有不精确性和不确定性。

## 2.7 思维数学

人的思维机理迄今仍是有待破解的问题。但是有一点是可以肯定的，即人们的思维作为知识获取的一个必不可少的途径，肯定也要采用软计算这一方法，至少是对于不精确和不确定的那些知识而言。因此有人推出了在思维过程中采用的数学这样一个概念。西北大学孟凯韬教授是思维数学的热心倡导者，他提出了“情商”这一概念，他的定义如下<sup>[5]</sup>：

假定某人所遇到的困难增大的幅度为  $m(m > 0)$ ，而其干劲增大的幅度为  $n(n > 0$  或  $n < 0)$  则称

$$Q = \frac{n}{m}$$

为其情商。

他接着指出，若  $Q \geq 1$ ，即干劲增大的幅度大于或等于困难增大的幅度，则遇到困难不会对其达到预期目的产生负面影响，甚至还会从反面起到促进作用，若  $0 < Q < 1$ ，即干劲增大的幅度小于困难增大的幅度，则遇到困难会对其达到预期目的产生一定的负面影响，但还不致使其改变初衷。若  $Q < 0$ ，即困难增大而干劲减少，则会使其一筹莫展，甚至悲观失望。

这里，困难增大幅度及干劲增大幅度虽是数值，但这个数值如何得来，并未被具体说明，而且它们肯定也是不精确和不确定的。因而  $Q$  本身也是不精确和不确定的。

## 2.8 胡一舟现象

在现代中国的乐坛上，出现了一个奇怪现象，这就是一位弱智青年胡一舟，竟能成功地指挥一支训练有素、水平很高的乐队进行一场场威武雄壮气势轩昂的演出，并博得听众经久不息的热烈掌声。听众欣赏的不仅是整个乐队的整齐和谐的演出，而且当然也惊叹于这个弱智青年的杰出表现。在别的领域中，胡一舟的智力，经测定，大概不及一个四五岁的孩童的智力，然而在音乐，特别在指挥里，胡一舟却有着一般人所没有的智慧。这里我们关注的头一个问题是胡一舟这样的智力，他如何能获取指挥的知识呢？可以相当肯定地说，胡一舟只能通过软计算的方法来学习指挥的才能，同时指挥本身是否也属于不精确不确定的知识呢？答案似乎应该也是肯定的。

这里还要指出，胡一舟现象并不是个别的。在历史上，也曾出现过弱智的人成为出色的

画家或成为某方面的专家。因此可以说，那些领域的知识，一定也属于不精确不确定的。

### 3 通过软计算的知识获取

(1) 考虑通过软计算形成事物概念的过程。在形成符号（概念）知识之前，只有模式（事物）知识，通过对于这些模式特征的抽取，在特征空间中，得到模式集合。在有了这个模式集合之后，最后才取得了符号知识，图 1 表现了形成事物概念知识的过程。

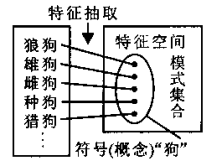


图 1 模式和符号

在图 1 中，特征抽取是通过软计算的方法获得的，所以得到的也就是一个模式集合。模式集合与概念之间是一对一的。但如果对于模式集合的元素而言，那是多对一的关系。图 2 进一步描述了模式集合与符号之间的关系。

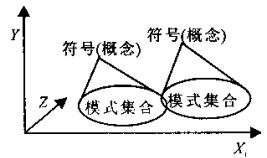


图 2 模式符号对

(2) 概念之间的关系，就形成命题、断言或描述，这也就是一个概念与另一个概念之间的联系。但这种联系可以是不确定和不精确的。因此有以下的表示（图 3）。这可以用于表现一个概念表示什么或类似什么，比如“上善若水”，图 3 下边的符号（概念）表示“上善”，而上边的符号（概念）表示“水”。这里给出的是“上善”和“水”的那些好的特征的对应。

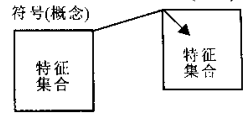


图 3 符号与符号之间的对应

(3) 上边是一对一的对应，也可以有多对一的对应。这种多对一的对应如图 4 所示。这可以表示，比如，图 4 上边的符号表示“苦难”，下边最左边的符号是“天才”，“天才”对应于“苦难”作为“垫脚石”的特征；中间一个符号是“能干的人”，它对应于“苦难”作为“财富”的特征；而下面最后边一个符号是“弱者”，它对应于“苦难”作为“无底深渊”的特征。

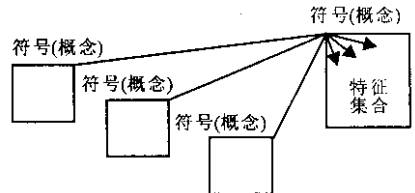


图 4 多个符号与一个符号的对应

(4) 通过进化算法比如基因算法获取知识的过程，也可以以算法流程图来描述。用基因算法来获得知识，首先要形成问题的系统阐述，其次要对它进行编码，然后生成初始分布，评价，进行基因运算，然后对于运算结果进行分析评价。如果已经达到满意程度，则认为已经获得了知识，否则则要继续运算。因此其流程图如图 5 所示。

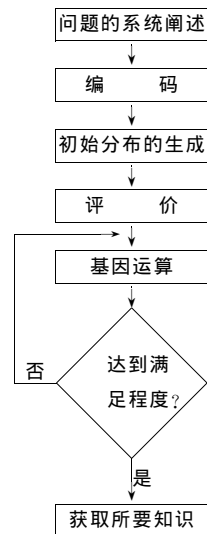


图 5 进化算法流程

(5) 对立统一规律知识的获取可以想象成一个集合  $S$ ，其中包含有许多概念（符号），与之对应的是另一个集合  $S^*$ ，它们包含的是这许多概念的对立物，如图 6 所示。在从集合  $S$  选取一个概念  $A$  后，为获得和  $A$  形成对

立统一的概念  $A'$ , 沿着对立统一律的路径, 对集合  $S^*$  进行检索, 找出  $A'$ 。还应注意, 对立统一律给出的“对立物”, 未必就是惟一的, 如果不惟一, 那还要通过适当的别的方法, 选出合适的对应物来。

#### 4 结束语

本文首先讨论了在人的浩瀚知识中, 哪些属于不精确和不确定的知识, 并且说明这些知识是可以通过软计算来获取的, 继而就以示例形式, 论述如何使用软计算的方法获取它们。但是, 运用软计算方法来获取这些不精确、不确定的知识途径还是比较初步的。当涉及到多个层次的概念之间关系时, 如何用软计算方法求得它们; 对于非线性的动态的模型, 在计算机中如何表现它们; 若干个不同的软计算方法如何加以组合来求解更为复杂问题以及在计算机上如何实现等, 这些问题都有待于继续深入研究。

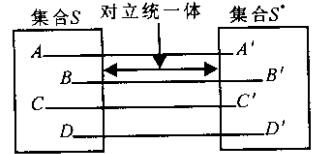


图 6 利用对立统一律获取知识

#### 参考文献

- 1 Furuhash T. Fusion of fuzzy/neuro/evolutionary computing for knowledge acquisition. Proceedings of The IEEE, 2001, 89(9): 1266~1274.
- 2 Bonarini A. Evolutionary learning reinforcement learning and fuzzy rules for knowledge acquisition in agent-based systems. Proceedings of The IEEE, 2001, 89(9): 1334~1346.
- 3 周 仪, 罗 平. 中国古代格言妙语(英译本). 南宁: 广西教育出版社, 1998.
- 4 纹 绮. 巴尔扎克妙语录. 兰州: 甘肃人民出版社, 1988.
- 5 孟凯韬. 哲理数学基础. 北京: 中国科学技术出版社, 1999.

(责任编辑: 黎贞崇)

### 广西参加第二十届全国青少年信息学奥林匹克竞赛取得较好成绩

由中国科学技术协会主管、中国计算机学会主办、复旦大学承办的 NOI2003 第二十届全国青少年信息学奥林匹克赛于今年 8 月 3 日至 9 日在上海举行。广西青少年信息学(计算机)竞赛委员会根据去年的联赛成绩, 经研究决定由柳州铁路第一中学的蒋程皓、林涛、李想、田翼四位选手组成广西代表队, 并由朱锦锋老师领队参加 NOI2003 竞赛。广西参赛选手不负众望取得较好成绩, 广西代表队获团体成绩第 9 名, 林涛、蒋程皓分别获得个人二等奖, 李想获得个人三等奖。林涛同学个人成绩排名第 20 位, 获得进入国家集训队的资格。全国有 32 个代表队共派出 131 名选手参加了本届全国青少年信息学奥林匹克竞赛。

(唐贵松)