

意识机:理论与模型

钟义信

(北京邮电大学信息工程学院,北京 100876)

摘 要: 智能机器的研究已经成为日益迫切的课题和普遍关注的焦点.但是,由于智能问题本身的巨大复杂性,使智能机器的研究面临着许多理论上的困难.研究发现:智能系统的基础和核心在“意识”;而意识起源的机制则正好是“全信息”的生成过程.基于这一重要发现,本文提出了一种基于全信息理论的意识机理论与模型,论证了意识机模型的技术可行性以及意识机与智能机的联系与区别.这一研究结果有助于人们对于意识问题的奥秘进行深入的探索,也为智能理论的研究开拓了一条新的途径.

关键词: 智能机; 全信息理论; 意识机; 归纳学习; 神经网络

中图分类号: TP183 **文献标识码:** A **文章编号:** 0372-2112 (2000) 10-0041-04

Consciousness Machine: Theory and Model

ZHONG Yi-xin

(College of Information Engineering, University of Posts and Telecommunications, Beijing 100876, China)

Abstract: The research of intelligent machines has been an increasingly concerned subject these days. Due to the great complexity of the subject, however, the research faces many difficulties for the time being. From a long period of study we discover that the foundation and the nucleus of intelligence roots in the consciousness and the mechanism of consciousness emergence is just the same as the process of comprehensive information production. As a result of this discovery, a theory and the related functional model of consciousness based on comprehensive information theory are established, and the technical feasibility of the model implementation is also discussed. The results reported in the paper may lay certain foundation for the exploration of the mysterious consciousness and a new, and a promising, approach to the research of intelligent machines may also hopefully be initiated.

Key words: intelligent machines; comprehensive information theory; consciousness machines; inductive learning; neural networks

1 引言

发展科学技术的目的在于扩展人类认识世界和优化世界的能力,主要包括扩展人的体质能力、体力能力和智力能力.其中,扩展智力能力是最为引人入胜然而也是为困难的问题.

人们一直努力尝试使机器具有智能,期望机器不仅可以代替人类进行一般的体力劳动,而且可以帮助人们从事一些智力工作.沿着这个方向,人工智能研究已经取得了一些令人鼓舞的进展.各种专家系统的问世,为人们提供了许多各有特色的信息服务:咨询、预报、管理、诊病、助教、辅助决策等等.不久前,IBM研制的“Deeper Blue”计算机还几度击败了国际象棋世界冠军,给人以深刻的印象.

尽管某些人工智能系统表现了相当出色的性能,但从总体上看,所有的专家系统都具有明显的“窄台阶”效应:只能在非常固定的专业领域内按照严格的逻辑知识来工作,一旦越

出原先设计的专业领域,它们的表现就变得不合逻辑.

这样,我们就不能不对现有智能机器的基本理论进行反思.

早在人工智能理论问世之前,人们就认识到:智能是由信息加工提炼出来的.因此,作为信息论创始人的Shannon被邀请为1956年人工智能学会成立大会的几位主要发起人之一.但是随后,人们很快就发现,Shannon信息论不能为人工智能的研究提供有效的帮助,于是便迅速同Shannon信息论分道扬镳,并把自己定位于计算机科学的一个分支.

智能理论发展的这段历史有两个误解,发人深省.第一,当初人们把信息理论看作是智能理论的基本前提,是完全正确的;但是,当初人们把Shannon信息论误认为也是面向智能问题的信息理论,结果事与愿违.第二,Shannon信息论不能支持智能问题的研究,本来只能说明Shannon信息论有限

性,需要建立新的面向智能问题的信息理论;但是,结果却误认为人工智能的研究可以不需要信息理论的支持.这两个误解提供的共同启示是:需要建立和发展一种新的、面向智能研究的信息理论.

2 智能、意识与信息

什么是智能?尽管在有关的教科书中已经存在不计其数的智能定义,人们在实际的工作中也提出过许多不同的理解,为了深入研究问题,我们还是从智能问题的根本出发点做起,重新探讨智能的定义.这里,给出如下的定义^[1]:

智能,就是“对于给定的问题-环境-目的,有针对性地获取有关信息、恰当地处理这些信息达到认知、并在此基础上合理地再生策略信息、从而成功利用策略信息在给定环境下解决问题达到预定目的”的能力.

这一定义也可以用图1来形象地表述.

分析这个定义可以得到一些重要的启迪.

首先,这个定义指出:智能是一种“能力”,但并不是抽象的、神秘莫测的、无所不包的能力,而是一种具体的、可以实现的能力,一种在给定环境下成功地解决问题满意地达到预定目的的能力.

更具体地说,这种能力包括四个方面的要素:有针对性地获取与问题-环境-目的有关的信息的能力、恰当地处理这些信息达到认知的能力、合理地再生策略信息的能力、成功地利用这样产生的策略信息解决问题达到目的的能力.

定义揭示了:智能的核心,是获取-处理-再生-利用相关信息的能力;而且,四个环节环环紧扣,浑然贯通.获取信息是处理信息的前提;处理信息是为了再生信息;再生策略信息是为了用来解决问题.显而易见,如果没有信息,智能就无从谈起.

第二,定义强调了获取信息要具有“针对性”、处理信息要具有“恰当性”、再生信息要具有“合理性”、利用信息要具有“成功性”.这些要求表明:研究智能问题所需要的信息不应是 Shannon 信息,因为后者只包含了信息的语法因素(即信息的形式因素),不包含信息的语义(即信息的含义)因素和效用(即信息的价值)因素.可是,如果不了解信息的内容和价值因素,就无以为据来判断是否满足上述“针对性、合理性、明智性、成功性以及满意性”的要求,而不能作出这些基本的判断,就不可能产生真正具有智能水平的行为.可见,与通信的情形不同,研究智能问题必须综合考虑信息的语法、语义和语用因素.这种综合考虑信息的语法、语义和语用因素的信息理论,称为“全信息理论”.这就解释了为什么历史上的智能理论研究很快就与 Shannon 信息理论分道扬镳的原因,也说明了为什么要从全信息理论的立场来研究智能问题的原因.只是由

于篇幅的限制,这里不再介绍“全信息理论”本身,需要了解的患者可以参看文献[2].

第三,为了实现获取信息的针对性、处理信息的恰当性、再生信息的合理性、利用信息的成功性,关键对于给定的主客观条件(主观条件—目的;客观条件—问题与环境)的信息要具有理解能力.换言之,智能的关键在于对信息的理解能力.而理解力是与意识系统直接相关的功能.这就明确地启示了为什么要把注意力从一般的智能理论研究转向更基本的意识理论研究的道理.

为了深入研究智能问题,必须首先探讨它的基础——意识过程.

需要指出,目前,国内外学术界对“意识”的理解还存在许多差异.其中,哲学界通常把“意识”理解为与“物质”相对应的概念,泛指“人的头脑对客观物质世界的反映,是感觉、思维等心理过程的总和”,这是一种广义的理解,狭义的理解是指“人的觉知能力”.只要一个人有感觉、记忆的判断能力,就说他有意识.此外还存在许多不同层次的理解,这里不一一列举.

本文认为,虽然意识与行为的关系非常密切,但是,为了便于研究,还是把意识定义为一个认识的过程,不包括行为过程.意识只为智能提供一个认识基础.智能则必须在意识的基础上采取行动,解决问题,达到目的;在不能达到目的的情况下还要反过来纠正认识,再付诸行动,直至成功.这是“意识”与“智能”的联系与区别.

本文把“意识”的概念理解为“对于外部刺激的感觉、记忆、理解和判断”的能力.所谓某个系统对某种刺激“具有意识”,主要是指这个系统对于这种外部刺激具有感觉能力、记忆能力、理解能力和判断能力.

当然,意识能力不是铁板一块,系统的意识能力有高低强弱之分.有的系统对各种类型的刺激都具有意识能力,有的则只对某类或某些刺激有意识能力;有的系统对刺激的感觉、记忆和理解能力都很强,有的可能只有较强的感觉和记忆能力,理解能力比较弱;等等.但是这些情况都笼统地称为“有意识”.

3 意识起源和意识生成机理

在研究的初始阶段,目的是要探求意识起源的奥秘:研究人的意识能力的生成机制,即研究人的意识究竟是怎样从无到有生长起来,又怎样从弱到强发展起来的.这是意识问题研究的基础,也是智能问题研究的基础.然后,再研究如何将这种机制移植给机器,使机器也具有某种意识能力.

正象上述智能定义所表明的那样,任何智能系统都是有目的的系统;反之,没有目的的系统不可能有智能,尤其不可能有高级的智能.作为智能系统的核心,意识系统也是有目的的系统;没有目的的系统也不可能有意识.

当然,系统的目的有大有小,有简单,有复杂.不过,任何复杂目的都可以分解为一系列简单目的.因此,为了研究方便,我们总是从简单目的(单一目的)系统开始.

观察表明,当一个有意识能力的系统接受到某个具体的外来刺激的时候,通常会产生两方面的效应:一方面会观察到

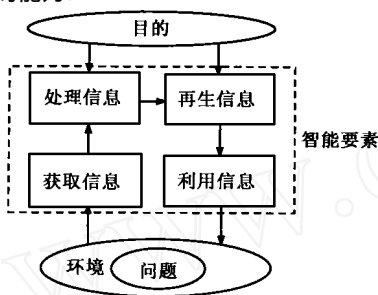


图1 智能定义的解释

这个刺激的外部形象;另一方面会感受到这个刺激对自身目的的影响。前一种效应比较简单,意识系统可以通过自己的感觉器官来实现对刺激形象的观察,并通过自己的记忆系统把这种观察的结果保存下来;后一种效应比较复杂,需要把刺激所产生的后果与自身的目的相比较才能判断这种影响究竟是有利还是有害。最初,系统不可能判断这种利害,往往要等这种利害实际发生之后才能有所体验。有了这种切身的体验之后,就可以成为经验性的判断力。因此,可以把意识系统的外部刺激在概念上人为地分解为两个基本分量:

(1) 刺激的“形象”分量,指系统所感知的这个刺激的外部形象和姿态,包括它的外貌、形状等静态形式参数和动作、情态等动态形式参数;

(2) 刺激的“效用”分量,指对于系统所设定的某个目的而言,这个刺激可能会带来什么样的影响,是有利、有碍、还是无碍于实现系统所设定的目的。

于是,有理由设想:意识系统首先将分别对“形象”刺激分量和“效用”刺激分量进行一些必要的预处理,然后对它们进行综合分析,具体包括以下功能过程:

—信息检测功能,即应当能够感知这个刺激的存在,包括感知它的“形象”参数和“效用”参数;

—模式分类与记忆功能,即提取这个刺激的模式特征,从而能够把这个刺激同其它不同的刺激加以分类;

—编码与记忆功能,即以比较简洁和易于检索的形式,将这些特征参数存储在短期或中期记忆系统内;

—理解和判断功能,即根据所提供“形象”编码和“效用”编码,对刺激进行综合理解,产生相应的判断,并将理解和判断的结果存入记忆系统备用。

此后如果系统再接受到这种刺激的形象(不一定要同时接受到效用刺激),系统就会立即唤起(在记忆系统中检索出)相应的理解,作出判断。

这样,就可以得到意识系统的功能模型,如图 2 所示:

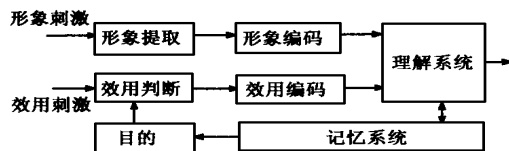


图 2 意识系统的功能模型

图 2 所示的模型中,预处理单元主要包括信息检测、特征提取、效用判断与信息编码。除了“效用判断”之外,他们都属于常规系统。在最简单的情况下,“效用判断”的任务只需给出一个三值的模糊描述:有利,无所谓,有害。当然,进一步要求则还需要给出效用的数值大小。

相比之下,一个需要特别关注的环节是其中的理解系统。

先来考察其中一种最简单的理解方式——初级理解方式的工作机制。它可以表示为直接在“形象 F ”与“效用属性 U ”之间建立映射联系:

$$\{F\} \rightarrow \{U\} \quad (1)$$

意即:“形象特征为 F 的刺激”具有“属性为 U 的效用”。

如果外来刺激的形象可以随时间而变化,而每种具体形

象对系统的目的又可能产生不同的影响,那么,初级理解机制的更具体表达式可以写为:

$$\begin{aligned} F(1) &\rightarrow U(1,1), U(1,2), \dots, U(1,N(1)) \\ F(2) &\rightarrow U(2,1), U(2,2), \dots, U(2,N(2)) \\ &\dots \\ F(M) &\rightarrow U(M,1), U(M,2), \dots, U(M,N(M)) \end{aligned} \quad (2)$$

它的意思是:某一事物可能在不同情况下呈现不同的形态(例如 M 种),而每一种形态又可能具有若干种不同的效用属性(例如 $N(m)$ 种, $m=1, \dots, M$)。一旦这个事物的所有形态和它们所对应的效用属性关系都建立起来,人们对这个事物就达到了某种理解,因而也就可以说系统形成了关于这个事物的某种意识。

比如,儿童对苹果的理解过程,就是在多次接受各种各样苹果的“刺激”的体验基础上逐渐形成苹果的“外部形象 F ”与苹果的“效用属性 U ”之间的联系的过程:

“丰满圆润成熟新鲜的苹果”

—丰富的营养,充足的水分,可口的味道”

“未成熟的苹果”—苦涩的味道”

“成熟过头的苹果”—腐烂的味道,可能致病”等等

一旦建立了这些联系,他就建立了关于“苹果”的理解和意识。这种理解方式的基本特征是:建立映射,枚举归纳,如图 3 所示。

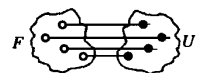


图 3 映射枚举的归纳理解

显然,只要能够把某种刺激“形象”的所有可能“效用属性”枚举得足够充分,而且如果这些“效用属性”之间不存在矛盾,那么,把它们直接归纳起来自然就可以形成对这种刺激的基本理解。这里所说的“直接归纳”,实际是一种简单的总括。

枚举归纳的理解方式看似简单,却是最基本最有意义的理解方式,人类群体意识起源的最基本机制就是建立在枚举与归纳的方式上,人类个体的意识起源(特别是幼儿阶段)主要也是通过这种理解方式来实现。

只有当通过“枚举与归纳”这种初级方式建立的理解积累到一定的程度之后,理解能力才有可能发生质的飞跃,以逻辑推理为主要特征的高级理解方式才有可能逐步建立起来。事实上,高级的理解方式也还要以“枚举归纳”理解为基础。

对照图 2 的模型可以发现:从“全信息理论”的观点来看,“形象”刺激所携带的正是刺激的“语法信息”,效用刺激携带的正是“语义信息”,而理解的结果所给出的则正是刺激的“语义信息”。因此,意识生成的机制正好就是“由语法信息和语义信息导出语义信息从而生成全信息”的机制,可以通过信息科学的原理和技术实现,如图 4 所示。

同上面的分析一样,从图 4 可以看出,在意识机的信息模型中,预处理和记忆系统都是常规信息系统,研究的难点也是归纳学习系统,其中的核心问题是如何从语法信息和语义信息归纳出语义信息,从而建立相应的全信息。

类似于上面的讨论,在最简单的情况下,由语法信息和语义信息归纳出语义信息的过程,也可以通过映射枚举的机制来实现:

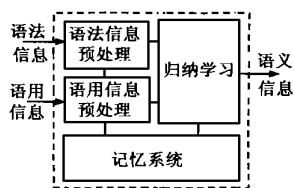


图4 意识机的全信息模型

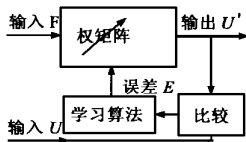


图5 用神经网络实现归纳学习

$$/ \text{语义信息} / : / \text{语法信息} / \rightarrow \gamma \text{语用信息} / \quad (3)$$

一旦建立了关于某个刺激的全信息,系统就可以根据语法信息在外部形态上把握这个刺激,根据语义信息来理解这个刺激,根据语用信息来判断这个刺激对自身目的的利害关系,并确定系统对这个刺激应采取的态度和应表达的感情:

语法信息—感知;语义信息—理解;语用信息—感情 (4)

显然,只要系统真正拥有了关于某个刺激的全信息,就应该说系统对这个刺激形成了基本的意识。系统拥有关于某个刺激的全信息越丰富,系统对这个刺激的意识能力就越强,系统对越多的刺激拥有充分的全信息,系统的意识能力范围就越广。

4 意识机模型:可实现性的讨论

通过以上对意识系统工作机制的探讨,发现“意识形成”与“全信息生成”存在很强的同一性。这种发现,使我们有可能利用信息科学的方法来研究人工意识系统的具体实现。

如前所述,意识机模型中的其它环节基本上是一些常见的信息系统,如检测、识别、编码和存储等等。因此,在讨论意识机模型可实现性问题的时候,重点将是基于语法与语用信息归纳语义信息的可实现性问题,即归纳学习问题。

描述语义信息归纳过程的一种比较自然的方法是状态空间方法:各种刺激的所有形式参数集合构成一个多维的形式状态空间 $/F/$,每种具体的刺激形式参数则是这个空间的点 F 。对应于任状态 F ,可以根据系统所体验的“效用分量”建立相应的效用参数 U ,形成效用属性空间 $/U/$ 。于是可以建立相应的状态肯定度和效用度的广义分布,从而在必要的时候也可以建立相应的语法信息和语用信息的度量。

在这种情况下,语义信息的表达就表现为在形式状态空间 $/F/$ 与效用属性空间 $/U/$ 之间建立正确的映射关系集合:

$$/ \text{若形式状态为 } F \text{—则效用属性为 } U / \quad (5)$$

实现这种理解机制的一种技术系统是人工神经网络。

例如,前向感知机的训练学习过程可以表示如下:当输入的训练样本集合为 $/F/$ 和 $/U/$ 的时候,人工神经网络的实际输出为 $/U' /$ 。用 E 来表示 U 与 U' 之间的误差,并根据这个误差的大小和性质来调整人工神经网络的权矩阵,如果学习算法设计得合适,通过调整网络权矩阵就可以使 U 与 U' 的误差 E 减小,直到满足要求为止。

经过训练,最终稳定下来的人工神经网络权联接矩阵就反映了与形象刺激分量 F 和效用属性分量 U 相联系的语义结构。一旦受到形象 F 的刺激,它就能够准确地产生相应的效用属性 U 。于是,这个神经网络的权联接矩阵就表达了这个

刺激的形式与效用之间的联系,“理解”了这个刺激的含义,形成了关于这个刺激的意识。

把图5所示的这种结构嵌入图2或图4的理解系统,就可以得到一种可能的意识机工作模型,如图6所示。

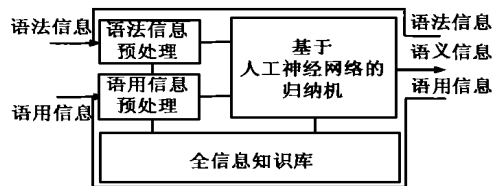


图6 意识机的实现模型

当然,归纳学习的算法还有其它实现形式,篇幅所限,这里就不一一讨论了。

当式(4)的映射集合比较庞大复杂的时候,上述简单的“枚举归纳”机制将显得十分笨拙。这时,需要借助于推理的归纳方法,而且还可能需要与演绎推理的方法结合起来才能解决问题。即,通过逻辑进行合并、化简和选择,在化简基础上归纳。在这些逻辑处理的过程中,需要借助于全信息的模糊逻辑推理来进行合并和化简,借助于全信息量计算来进行逻辑选择。

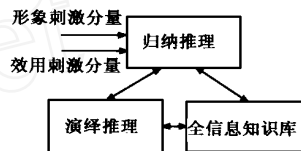


图7 高级归纳学习

这就是基于全信息理论的高级归纳学习方法,如图7所示。

总之,这些分析表明,具有以上意识功能的意识机是可以实现的。

5 结语

本文探讨了智能系统的核心问题——意识系统的工作机制,发现它原来就是通过语法信息和语用信息生成语义信息的机制;基于这一发现,找到了实现意识生成机制的技术途径,得到了可以实现的意识机模型,从而为探索意识问题的奥秘提供了一定的基础,也为智能系统的研究开辟了一个新的可能性。

另一方面,意识机的理论分析和实现技术还有许多问题需要进一步研究。例如,本文对“意识”的定位肯定还可以进一步商讨;“意识系统”的功能模型和信息模型都属于最简模型,还可以进一步展开;归纳学习系统特别是高级归纳学习系统的理论和实现本身还有许多值得探索的问题。对此,将在今后的研究中继续探讨。

参考文献:

- [1] 钟义信等. 智能理论与技术:人工智能与神经网络[M]. 邮电出版社,1992.
- [2] 钟义信. 信息科学原理[M]. 北京邮电大学出版社,1996.