

意识的原理:系统的体系结构

谈加林^{1,2}

(1.元系科技(深圳)有限公司,广东 深圳 518052;2.深圳职业技术学院,广东 深圳 518055)

摘要:意识系统由意元网络和意元网络串行活动机制以及外部设备三个基本部分构成。一是基于意元网络实现对意元及由意元构成知识的“存—算—控”一体化活动,并在觉醒状态下,处于一定准激活水平的各个意元由彼此之间存在的相互联系实现互相定义与说明,从而实现对事物的理解与定向;二是在意元网络串行活动机制的支持下,系列内外输入驱动意元网络相应的系列叠加态知识相继产生坍塌,构成有目的、有中心、有顺序的物理符号系统活动或形式化活动;三是通过外设实现与环境的交互。这是意识系统实现各种意识活动普遍的基本方式。

关键词:基本单位知识;意元;知识叠加态;人工智能;知识叠加原理;串行活动机制;意识系统

[中图分类号]TP18;TP11 [文献标识码]A [文章编号]1672-934X(2022)06-0111-11

DOI:10.16573/j.cnki.1672-934x.2022.06.012

Principles of Consciousness: System Architecture

Tan Jialin^{1,2}

(1.Yuanxi Science and Technology (Shenzhen) Ltd., Shenzhen, Guangdong 518052, China;

2.Shenzhen Polytechnic, Shenzhen, Guangdong 518055, China)

Abstract: The consciousness system is composed of three basic parts: chunk network, its sequential activity mechanism and peripherals. Firstly, based on chunk network, the integral events of "storage-calculation-control" of chunks and knowledge composed of chunks is realized, and in the awakening state, each chunk at a certain prospective activation level realizes its mutual definition and explanation through their mutual connection, further achieving understanding and orientation on matters. Secondly, with the support of sequential activity mechanism, a series of internal and external inputs drive the superposition state knowledge of the corresponding series of chunk networks to collapse one after another, forming a purposeful, centred and sequential physical symbol system and formalized activities. Finally realizes the interaction with the environment through peripherals. This is the universal and basic way for the consciousness system to realize various consciousness activities.

Key words: basic unit knowledge; chunk; knowledge superposition state; artificial intelligence; knowledge superposition theory; sequential activity mechanism; consciousness system

意识的产生是四大科学之谜中唯一尚未得到科学解答的自然之谜。

20 世纪 90 年代中期,谈加林已经证明存在

基本单位意识现象^[1-2];为解释这一现象,其在 2021 年申请的专利中,又提出基本单位知识概念且建立相应的模型^[3-4];并在 2022 年公开发表

收稿日期:2022-09-21

作者简介:谈加林(1961—),男,副教授,博士,主要从事心理学基础理论、人工意识、科技哲学等研究。

的论文中提出知识叠加原理,指出“意元及意元网络是实现强人工智能和意识活动的基础”^[5]。

意元网络整体并行地围绕众多事物,对各方面的大量信息同时进行处理。这时,由各种关于事物的叠加态知识和混合态知识所构成的意元网络活动,实际上是处于大规模并行活动下某种不确定的混沌状态,而并非确定的有目的、有重点、有条理、有顺序地进行某种针对性活动的状态,更别说是有计划、有逻辑的形式化活动。而确定的、有目的、有重点、有条理、有顺序地进行某种针对性活动,是系统的智能活动的有效性,尤其是意志活动和形式化逻辑活动的根本标志。意元网络系统如何以具有并行性和不确定性的意元和意元网络活动为基础来实现这种具有确定的中心和顺序等的意识活动,就是我们要讨论的意识系统的体系结构原理。

一、意元网络串行活动机制

不能有目的、有选择、有重点、有顺序地进行活动,就不能以最重要的事物为目标来有针对性地进行处理或应对。要基于大规模并行活动的意元网络,有目的、有选择、有重点、有顺序地进行各种活动,包括对事物的感知、记忆、想象、思维等各种认知活动和言语活动,对情绪情感的认知、体验等情感活动,特别是确定目的、明确任务、制定计划、执行计划和维持活动方向等各种意志活动,都需要某种串行活动机制予以支持。这些活动具有两个共同的特性:各瞬间中心内容的单一性和顺序性。可以将各瞬间中心内容的单一性看作是意识的确定性,中心内容的具体形式是关于事物的本征态知识(也叫符号知识)。所谓围绕一定的目的或任务或内容进行有选择、有方向、有重点、有中心的活动等,都是基于它的中心内容的确定性来实现的,而条理性、逻辑性、程序性等,则都是基于它的顺序性来实现的,这是由一个接一个的意元或以它们为中心的一系列意元网络串行活动所组成的。这样的活动的表层过程,可以抽象为图灵机活动或物理符号系统活动。底层过程则

正是意元网络叠加态知识和混合态知识的分布式并行活动。这里的所谓物理符号系统活动,泛指图灵(Alan Mathison Turing)、冯·诺依曼(John von Neumann)和西蒙(Herbert Alexander Simon)、纽厄尔(Allen Newell)等人,分别从不同角度描述的以一定物理符号处理方式进行的可以形式化的活动,这些都是对人类个体各种活动表层过程的抽象。

在串行活动机制的支持下,基于意元网络的大规模并行活动,有目的、有选择、有重点、有中心、有顺序地进行串行的物理符号系统活动或图灵机活动或形式化活动,是意识系统最重要的特点,也是为人工意识系统配置串行活动机制的主要目的。

(一)串行活动机制与串行处理机制

所谓串行活动机制,是指在并行的意元网络上实现串行活动的机制(如图 1 所示)。它由各种与意元网络相配合的串行处理机制构成。

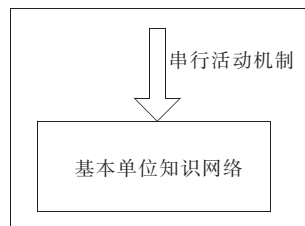


图 1 基本单位知识网络串行活动机制

所谓串行处理机制,是指各种有助于实现串行处理的特殊机制,作为一定串行活动机制的组成部分,包括不同部位数据精度等方面的差异化,比如具有一定功能的意元子网络或神经子网络的中心部分存在功能最强或更精细的加工,而其周边部分存在相对粗糙的加工。所述中心部分是指它的最精细或最敏感的部分,可以是相对的或大致的几何中心,也可以不是几何中心而只是功能较强的中心。中心部分相对于边缘具有一定优势,比如视网膜中央窝有更精细的加工,机械上的机械方式的听觉传感器的调谐机构,组织上的类似于视网膜及视觉传入神经子系统不同部位上存在的差异,以及因身体部位不同而产生的敏感性差异等。差异

化的作用是让中心部分进行较强或较精细的加工,从而使其更容易取得优势地位或中心地位。

在神经中枢也存在同样的现象。更重要、更广泛的串行处理机制是以一定方式使得只有某部分单元的功能进入活动状态或处于活动状态,或者相反,使其更不易于进入或处于活动状态等的机制,比如各种兴奋或抑制机制等。其中,抑制是指阻止激发或降低激发的水平或能力,包括侧抑制、拮抗性抑制、逐级抑制、超限抑制、习惯性抑制等。侧抑制是指对相近基本单位知识或单元的抑制;拮抗性抑制是指对立的基本单位知识或单元相互之间的抑制;逐级抑制是指较高级部分的基本单位知识对较低级部分的基本单位知识的抑制;习惯性抑制是因学习或训练而形成的一部分知识或单元对其他部分知识或单元及其作用的抑制,比如但不限于在空间认知中处于优势地位的视觉空间认知对触觉空间认知的抑制。这种对其他部分及其作用产生一定抑制作用的部分,常常处于优势地位。处于优势地位的部分或单元,对处于非优势地位的部分或单元常常具有抑制作用。超限抑制是指随着兴奋的持续时间延长或兴奋的次数增多而产生的对其自身的抑制。疲劳也会导致超限抑制。自19世纪末俄国生理学家、心理学家巴甫洛夫(Иван Петрович Павлов)等开始,生理学和心理学领域形成了关于神经系统的兴奋与抑制的大量实验研究成果,可供参考与借鉴。

一种串行活动机制可以结合多种串行处理机制。有些串行处理机制是相对独立的,但更多的是在不同形式、不同程度上与意念网络一起构成一体化的,甚至是意念网络所附带或所具备的机制(如图2所示)。

(二)实现串行活动机制的方法

实现串行活动机制的方法是多种多样的,尤其是在工程上,由于对系统水平的要求不同,形式更是可以灵活多样。

一是专门的意识机方式。在意念网络规模不大且相对简单的意识系统里,最适合采用专

门的意识机方式,它接受来自意念网络的各种意念的输入并向其输出,以驱动意念网络实现串行活动。这种串行活动机制类似于电子计算机中的控制器,是一种相对独立的机制,但对大规模意念网络而言则不是最合适的。

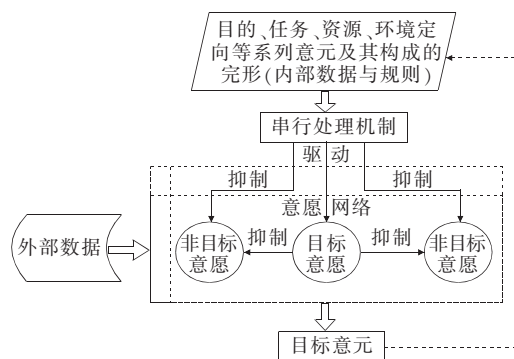


图2 串行活动机制及作用

二是与意念网络相融合的方式。这种具有更强适应性且可能应用更广泛的方式,是与意念网络相融合的串行活动一体化机制。这种串行活动一体化机制的实现方式,是从数据采集到基本单位知识分析再到内心世界的心智操作的各个部分,结合相应的串行处理机制,比如前述视网膜中央窝机制、听觉调谐机制、肤觉敏感性差异机制,以及在意念网络中利用意念之间丰富的相互联系,使得当时及其前后处于意识中心的意念,以参与构成一定完形的方式,驱动或激发相应的目标意念的活动,和(或)对其他非目标意念产生抑制的机制等(如图2—图4所示)。

三是意念网络内部的各种相互控制与层级关系方式。尽管意念网络中各个意念都处于知识的基本单位层次,但从信息加工意义和逻辑抽象意义等方面来说,它们又可以处于不同的层级。这里的意念网络的不同部分,一方面在相互控制和相互影响上存在等级关系,另一方面在功能上也存在层次关系。在控制层级上处于较高级的部分具有相对的优势地位,比如它具有对较低级部分的基础抑制作用。处于较高级中枢抑制下的低级中枢的活动,在缺乏来自较高级中枢的激励输入的情况下,较难激发成为活动的中心,也较难进入串行活动过程。也

就是说,这种多级中枢系统的各级中枢之间存在的基础控制关系,是串行活动机制的一部分。意元网络各部分因为功能上的相互关系,也表现出层级关系。比如在思考和想象活动中,各种程序单元具有相对于数据单元的优势地位。尽管感知活动中的感知输入具有驱动作用,但有关特征的各基本单位知识单元的活动,其实仍然是在更高级的关于物体知觉的基本单位知识单元的综合控制下进行的。

意元网络各部分之间的这种相对“优势—劣势”关系,也是对意元网络各部分相应意义的反映。一般来说,物理意义越复杂、越综合、越抽象,相应的基本单位知识或子网也就越高级;与目的或意图、操作程序或过程等相关,且在功能意义上具有更高地位的基本单位知识或子网,相较于一般物体的基本单位知识也处于更高级的优势地位。总之,意元网络的各个意元除了是知识的基本单位外,以及各部分之间除了知识性信息的传播外,较高级的部分具有相对优势地位,对较低级部分产生常态性抑制作用和激发时的去抑制作用,因而彼此之间存在具有控制性意义的相互关系。

这里的目标意元,即为所要驱动的意元;非目标意元,则是指某种意义上不要驱动的意元。其中的驱动,是指激发一定的意元产生相应本征态知识输出,也就是触发叠加态知识的坍塌。

产生抑制作用的单元的被抑制或被破坏或抑制通路被阻断,可以解除其对其他单元的抑制作用,比如高级部分或优势部分的被抑制或被破坏或抑制通路被阻断,可以解除或消除它产生的对低级部分或非优势部分的抑制。

来自不同部分的信息在信息处理中的权重也反映了这种“优势—劣势”关系。比如在空间认知中,对正常人而言,视觉一般处于优势地位,而表现为对矛盾的触觉的抑制作用。

图 3 是一个多级中枢系统中各级中枢间的基础控制关系的简化示意图。因为主要描述中枢间的基本控制关系,本应该有的逐级输入的部分被省略,且各部分的网络关系也被简化,最

后抽象为一个简化的框图。其中,系统的高级中枢不仅实现高级的或抽象的认知功能,更重要的是还处于更高的控制水平实现对次高级中枢的控制和调节,再由次高级中枢实现对低级中枢的控制和调节,以及由低级中枢直接控制、支配与调节系统的各功能部件(包括外部设备和机体内部设备的执行部件)执行相应操作。系统各部件可以是分立的,也可以是联结在一起的。较高级的中枢相对处于优势地位,较低级的中枢受到来自较高级中枢的抑制作用,这是串行活动机制的一部分。意元网络各部分之间不仅存在这种控制关系,而且根据目的和任务等相关知识可以灵活打开或选择相应的“兴奋—抑制”通路,其宏观表现为兴奋和抑制的扩散与集中。在意元网络上,意识中心的兴奋轨迹类似于闪电轨迹。内部知识的输入参与驱动随后的串行活动机制,从微观来看,可以是意元或意元网络的不同子网都接受来自其他子网的输入,产生向其他意元和意元子网的输出,并成为其他意元或子网的内部知识输入。其中,高级子网,尤其是有关目的、任务、操作及程序等的意元子网都处于相对优势的地位。

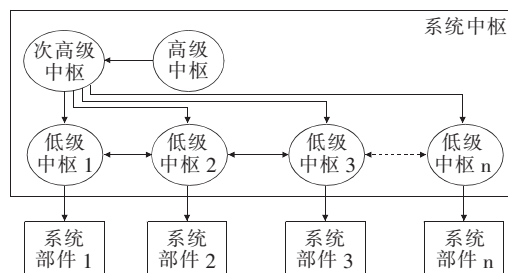


图 3 多级中枢的基础控制关系

(三)进入串行活动的几种方式

在意元网络大规模的并行活动中,哪一种会优先进入串行活动,这也是串行活动机制的一部分。

一方面,在围绕一定目标的活动中,比如图灵机活动或物理符号系统活动,符合相应活动要求或规则的系列意元活动构成串行活动。这时,各种心理活动的形式化规则和活动规律发挥着作用。另一方面,在围绕一定目标的活动

中,若有与目标无关且具有新异、独特、对比强烈等特点,或具有特殊意义、符合系统要求或需要的其他外部刺激或事物的出现(信息输入),则串行活动将可能转向这样的事物,相应意元的激活水平将得到加强,而其他竞争性单元可能被抑制或活动减弱,由此,其基本单位知识的活动将优先进入串行活动,或随后的串行活动更易于或优先展开。

也就是说,心理学中的注意规律可以在这里综合地发挥作用。其中,意志作用随着个体越来越成熟,而外界信息的作用相应地也会越来越受意志的调控。

二、意识系统的结构与工作原理

(一)意识系统的结构

要基于意元网络进行有目标、有秩序的活动,特别是图灵机的形式化活动,需要有支持意元网络进行串行活动的某种串行活动机制。所以,完整的意识系统包括三个基本部分(如图4所示):一是意元网络;二是与意元网络相匹配的串行活动机制;三是各种外部设备,包括输入和输出、预处理设备和执行操作或动作等各种其他功能设备。

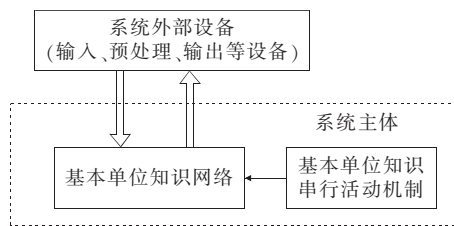


图4 意识系统的结构

(二)意识系统的工作原理

意识系统的工作原理,即系统基于主要以并行方式工作的基本单位知识网络,或具有基本单位知识网络本质的人工神经网络,实现知识的一体化存储与计算以及相应的控制,并且在串行活动机制的支持下,进行串行方式的各種图灵机活动或物理符号活动或形式化活动,进行各种心智活动或意识活动,以及借助外部设备实现和环境的交互作用。

为系统配置一定的基本单位知识网络,或具有基本单位知识网络本质的神经网络,可以是一个单一的网络,也可以是一个由各种基本单位知识子网或神经网络构成的统一的基本单位知识网络,或具有进行基本单位知识形式活动能力的神经网络,或是一个由分置的相对独立的具有分布式关系的若干种基本单位知识子网构成的基本单位知识网络,或由分置的且相对独立的若干具有进行基本单位知识形式活动能力的人工神经子网构成的人工神经网络。由不同部分组成的基本单位知识网络,其不同部分具有相对独立性。特别是,因为能得到相关辅助机制的支持不同,可能具有的性质及能实现的活动也会有所不同。

所述具有基本单位知识网络性质的人工神经网络,是指任何包含基本单位知识单元、基本单位知识活动装置、实际可以实现基本单位知识活动的人工神经网络,比如各种经过适当改造或改进的或带有相应串行活动机制的深度神经网络。谈加林的《意识的原理:基本单位知识与知识叠加原理》一文对此已有深入讨论^[5]。

(三)串行活动机制的作用过程

串行活动机制的作用过程,如图2—图3所示。首先,串行活动机制在任一瞬间只允许基本单位知识网络上的大规模并行活动中的一个意元处于它的中心。其次,最近处于中心的意元、新的输入将驱动的意元以及它们所构成的各种知识等,包括对事物的认识以及目的、任务、结构、程序、操作、数据等,作为内部信息返回输入意元网络,和外部输入的信息一起,驱动或触发意元网络上的混合态知识坍塌,使得坍塌到的一个新的意元成为意元网络新的活动中心和串行活动新的一环。最后,如此循环往复,将实现各种不同目的和内容的具有图灵机活动性质的心智活动。所述知识可以是关于事物数据的知识,也可以是关于事物程序或规则的知识。图灵机活动或物理符号活动等的形式化活动规则,是形式化活动的一部分,且是以符号知识或形式化符号形式存在的,一般称之为程序

性知识或操作性知识。这种借助串行活动机制的支持且基于意元网络来进行串行的图灵机活动方式,被称为“焊枪方式”:串行活动机制就像焊枪,它指向基本单位知识网络上的哪里就把进行中的串行的意识活动“烧”到哪里。

(四) 内部输入与外部输入

与传统人工神经网络方法不同,对意元网络系统而言,输入的不仅可以是来自外部的数据形态的信息,也可以是来自内部的知识形态信息;不仅可以是来自外部的物理空间结构形态的信息,也可以是来自内部知识结构形态的、逻辑结构形态的、任务结构形态的、语言结构形态的信息;不仅可以是本征态符号形式的信息,也可以是叠加态完形形式的信息。内部与目的、任务、结构、程序等有关信息的输入,发挥着组织思维等串行的意志活动的作用,包括对外部信息的选择性处理、理解以及反应等。

同时存在的信息输入,无论是外部输入还是内部输入都是多方面的,任何时候意元网络都在相应地进行着大规模并行的活动,只是由于串行活动机制的作用,选择了其中的一部分成为串行活动的中心,不仅成为意识到的内容,而且成为返回输入中具有决定性意义的成分,主导着系统接下来的活动。内部输入及其选择性作用与主导性作用,是串行活动机制的核心。所谓内部输入的选择性作用,从宏观来看,是在大量并行的意元活动中,选择哪一部分作为串行活动内容的问题;从微观来看,是意元网络中意元之间相互作用的问题。

不断的内外输入驱动相应的知识活动产生本征态知识的输出;新的输出又作为输入的一部分驱动下一部分知识的输出,如此往复,构成系统有目的、有重点的活动。尽管这依然是一般意义上的信息加工或数据处理过程,但它又凸显了新的实质:它是知识形态的信息加工活动,不是一般的数据性信息加工活动。意识系统中发挥主导作用的是知识形式的活动。

(五) 意识系统实现图灵计算的方式及意义

意识系统在其串行活动机制的支持下,基

于意元网络的大规模并行活动,不仅可以有目的、有选择、有重点、有中心、有顺序地进行一定的活动,更可以实现图灵机活动或形式化活动或物理符号系统活动。这不仅是人工意识系统技术可以继承现有知识智能方面已有成果的前提,而且一种系统如果是可以进行形式化活动的意识系统,必定能通过图灵机活动的检验。

图灵机假定有一条无限长且分成了一个一个小方格的纸带,有一个读写头在纸带上左右移动,读写头有一组内部状态和一些固定的程序。那么,在每个周期,读写头都要从当前纸带上读入一个方格信息,然后结合自己的内部状态与程序,一方面输出信息到纸带的当前方格;另一方面转换自己的内部状态并进行移动。

具体来说,图灵机由以下几个部分组成。

一是一条无限长的纸带。纸带被划分为一个接一个小格子,每个格子上包含一个来自有限字母表的符号,字母表中有一个特殊的符号表示空白。纸带上的格子从左到右依次编号为 $0, 1, 2, \dots$, 纸带的右端可以无限伸展。因为现实中的任何机器不可能有图灵机所需无限长的纸带,可见图灵机只是一个理想设备。

二是一个读写头。读写头可以在纸带上左右移动,并能读出当前所指格子上的符号,还可以改变当前格子上的符号。

三是一套控制规则。根据当前机器所处的状态以及当前读写头所指格子上的符号,来确定读写头下一步的动作,并改变状态寄存器的值,令机器进入一个新的状态。

四是一个状态寄存器。它用来保存图灵机当前所处的状态。图灵机的所有可能状态的数目有限,并且有一个特殊的状态,称为停机状态。就任意一个程序在有限时间内结束运行进行判断的问题,称为停机问题。但是停机问题已被证明是一个不可解的问题。同样,哥德尔也证明了一阶形式逻辑系统是不完备的。

由意元网络及其串行活动机制加上外部设备构成的意识系统是如何实现图灵机活动的呢?一是意元网络相当于图灵机无限长的纸

带,从某种意义上说,它是真正的无限长,意元就是纸带上的符号(如图5所示)。二是读写头的作用由串行活动机制来实现。串行活动机制从意元网络上读出的符号(知识),包括数据(描述性知识或数据性知识)和程序(程序性知识或操作性知识),以及根据程序对数据进行相应操作的结果(描述性知识)。而所谓操作,实际上仍然是借助串行活动机制从意元网络上“读出”相应操作的结果,即某种知识,它以一定的形式已经储存在意元网络。不断重复这样的过程,可以持续得到新的前述各种知识,并相继返回“写入”意元网络,从而驱动新的知识输出,这相当于寄存器当前状态的不断被改变和在纸带上读写符号。三是控制规则,也是储存在意元网络上的知识。这里的规则包含两种:其一为底层的完形驱动规则和符号驱动规则,主要是以符合知识叠加原理的方式存在的,称为底层规则,是在意元网络上读写符号(知识)的规则。其二是表层的图灵机活动要实现的活动的形式化规则,称为表层规则,是完成一定任务的活动形式化规则。表层形式化规则也以底层基本单位知识的形式储存在意元网络上。图灵机根据表层形式化规则并按一定程序就一定数据所进行的计算(操作)和计算的结果,都是在底层规则支持下,通过意元网络一系列串行符号知识的活动来实现的,即基于意元网络的串行活动,其表层是图灵机形式化活动,而底层是以知识叠加原理等方式进行的活动。如果把底层当作“黑箱”,不考虑底层的活动,意识系统的图灵机活动,就是纯粹知识的形式化活动。四是意元网络及其串行活动机制等辅助运行机制本身具有实现状态寄存器的作用。意识系统不必拥有专门的状态寄存器。五是意元网络储存的程序性知识,包括有关程序结束的知识。得到的是否为最终的结果,程序是否结束,如果根据图灵机形式化计算来确定,就称为算法式。但图灵机对停机问题不可解,图灵机是不完备的。对意识系统而言,更多的情况是,在缺少相应程序知识和算法知识时,或没有那么多时间用来

计算时,或没有必要得到精细的程序或精确的结果时,都可以采用经验的办法,称为经验式,即基于知识叠加原理的直觉方式或知识突现方式,来确定是否已经得到启动时所期待的最后结果,程序是否可以结束。经验方法背后的本质是知识叠加原理,它是图灵机形式化原理之外的另一种知识处理原理。意识系统进行图灵机计算时,可以基于知识叠加原理对结果(包括对每一个步骤的正确性)进行(直觉式、经验性)判断。因此,对结合了形式化原理和叠加原理的意识系统而言,停机不再是问题。采用图灵机形式化原理的算法式,或采用知识叠加原理的经验式,意识系统必定可以根据有关程序是否结束的知识来确定程序是否结束。意识系统是自治且完备的。

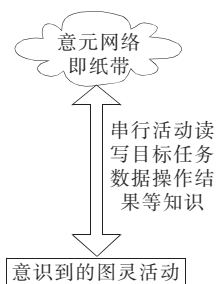


图5 意识系统的形式化活动

总之,意识系统的图灵机活动的表层,是基于意元网络进行符号知识的读取与存储等形式化的活动,底层则主要是基于知识叠加原理等来实现符号知识的读取与存储的完形知识的活动。其中,读取的符号知识,也可以是对外界事物识别的结果。意识系统的所有形式化活动或物理符号系统活动都是基于相应知识的活动来实现的。在意识系统里,如果要进行数字化活动,它也处于表层,是基于一系列数字化知识的活动来实现的。比如,十进制数据表示为二进制数据,二进制数据的基本运算,在意识系统里都是相应知识的活动。以十进制的“10”表示为二进制数为例,规则知识是:在二进制的各个位上,0是“0”,1是“1”,逢“2”进1。意识中十进制“10”向二进制转换的活动,可以一步步读取外部“十进制转为二进制”任务及数据“10”得到

相应任务知识“十进制转为二进制”和数据知识“10”,任务知识“十进制转为二进制”返回输入意元网络,搜索相应的关于操作及其结果的知识,采用算法式就是一步步进行相应操作和得出结果:1 的二进制数是“1”,2 是“10”,……,10 是“1011”。如果对不同进制之间数据的转换非常熟悉,则可以采用经验式直接在不同进制之间就同一个数的表示读取相应知识而实现转换:10 的二进制数是“1011”。所以,在意识系统中,知识的活动处于核心层,即便是计算活动,也由有关计算知识的存取活动来实现,而不是真正通过计算来实现。

以“ $1+2=?$ ”为例。“ $1+2=?$ ”的输入,相继产生对“1”“+”“2”“=”“?”以及任务“ $1+2=?$ ”的认知,即得到相应的知识。这些认知相继返回输入意元网络,从意元网络读出结果“3”。如果对“ $1+2=?$ ”还不太熟悉,则这个过程会分解为更多的同样结构的子过程。

万能逼近定理(Universal Approximation Theorem)^[6-7]意味着意元网络可以实现对任意事物的认知与表示,包括对任何结构形式的事物、事物之间如何相互作用、事物如何发展变化及过程等的认知与表示。而意元网络可以实现图灵机活动则意味着意识系统的形式化活动是自洽且完备的,从理论上来说,意识系统能够以形式化的方式认识任何事物的任何规律。

总之,不仅是不断地从意元网络读出有关事物的“数据性”知识或有关操作的“程序性”知识,以及操作结果(包括中间结果和最终结果)的知识,而且新读出的知识不断被追加到返回输入的序列之中成为其中最新的部分,并据此继续驱动所需的知识输出,实现系列的串行活动。意识系统这种以知识为基本单元进行形式化活动的方式,将成为对知识时代的最新也是最确切的定义。可以说,意识系统的实现,将是数字时代迈进知识型人工智能时代的标志。

(六)意志与始动问题

智能系统不应是孤立的系统。完备的智能系统要能自主适应各种内外要求才能保证自身

的生存并正常工作。系统的意志,涉及目的性、自主性、自治性等一系列问题,包括始动问题、解题问题和停机问题等。

有一种心理学理论采用二分法把人的心理分为意向和认知两个部分。意识系统要进行活动,首先要产生活动的意向(比如目的),并产生对内外要求的认知,而后才可以有进一步意向的图灵机活动,再根据结果是否已经满足了系统的要求来确定是否可以停机,这不仅使得停机问题在意识系统里不再是不能解决的问题,而且也是不难解决的问题。

始动问题是与图灵机的停机问题类似的一个问题,是指什么时候适应系统面临的内外要求而启动一个物理符号过程或形式化活动的问题。始动问题源于且基于需求问题,利用图灵机可以解决价值问题。如果系统不能认识内外要求,就不能明确目标并产生主观能动性,也就不能发动一个图灵机活动,而且系统哪怕是意元网络系统,即使能接受训练,形成对世界的摹写,它在一定程度上仍然是刻板的、机械的,不能真正明白包括它自己在内的各种事物的意义,白板系统也就无法得到全面而真正有意义的学习与训练。图灵机的停机意味着程序结束,活动的结果得到确定,同样也意味着需求是否得以满足和如何得以满足得到确定,以及活动各部分的价值都得以确定。需求问题和价值问题是紧密联系的,前者与始动相联,后者与停机相关,它们合称为“需求—价值”问题。

因为两个方面的原因,图灵机对始动问题不可解。一是图灵机所有可能的状态是有限的,但系统面临的对内外要求作出响应的启动状态却可能是无限的,特别是存在状态集里不包含新的需求状态的情况,这与图灵机的有限状态假设相矛盾,意味着反映当前的某个内外要求的启动状态可能不在图灵机的有限状态集里,或图灵机无法在有限的时间里搜索到与内外要求相应的启动状态。二是只要图灵机形式化系统是不完备的,图灵机就不能保证它的有限状态集里包含相应的启动状态,也意味着图

灵机的启动问题不可解。哥德尔不完备性定理已经说明了问题。

始动问题最根本的是:图灵机不能解决系统内外需求的无限可能性和新形式需求的不断产生。解决“需求—价值”问题的最好方式,是基于知识叠加原理的意元网络大规模并行的知识活动与在串行机制支持下基于形式化原理的物理符号系统活动的结合。意识系统中所有基本单位知识以多态的形式存在,它们遵循知识叠加原理,而知识叠加可以以学习与训练的方式,而不必以图灵计算的方式实现,这就使得意识系统进行图灵机活动时存在的不完备性得到直接弥补,始动问题和停机问题都可得到解决。

总之,系统可以基于意元网络实现对一般事物的认知,包括对内外要求的认知,而基于物理符号系统活动,实现对要求的选择、对满足要求的目标设定以及制定计划与按计划执行,可以基于学习和创新形成对事物新的认识和理解,进而形成新的心理需求并规划和实现相应的活动。这就是意识活动或精神活动。

三、对事物的理解与意识

(一)事物的映像和内心世界或内心小宇宙

对事物的理解是基于内心对世界的认识来实现的。除对事物自身的认知之外,对事物的理解还与对相关的人、事、时间、地点等方面的认知相联系,而相关的人、事、时间、地点等方面又与所处的世界相联系,这些都是一个人所处世界的一部分。

所谓事物的映像,是指以这个事物本身的意元为核心加上这个事物的不同部分的其他若干意元所构成的关于这个事物的心像。事物其他部分的意元越丰富,事物的映像就越饱满、丰富和鲜明。

所谓内心世界或内心小宇宙,是指由意元网络中所有意元以一定关系构成的关于世界的内心认识或内心映像。一般情况下,关于事物的映像和关于世界的认识处于知识叠加态和知识混合态,是在活动过程中且在串行活动机制

支持下,才以一定的图灵机活动或物理符号方式的活动或形式化活动的方式展现在人们眼前,如借助思维和语言等图灵机活动形式的结构化活动或描述与表达。所谓对事物的理解,就是关于该事物与周围世界关系的知识,也就是以该事物为核心的关于周围世界的知识。

(二)定向与意识系统对事物的理解

定向是指对一事物所具备的一定属性或特征等的确定,即对一事物与其他事物之间关系的确定,包括对事物的组成、结构、性质和相互关系等的确定。定向常常表现为关于系统自身或关于其他事物的一定说明。定向包括空间定向、时间定向、物体定向、目的定向、任务定向、结构定向及其他关系定向等。理解是定向的前提,定向是理解的表现。

如果把实现定向的机制称为定向机制,那么就某一方面进行定向的机制,就是指与相应的定向活动相关的机制。所谓的定向机制,除包括基本单位知识网络上各基本单位知识之间的相互联系、相互说明、相互定义之外,还包括基于基本单位知识网络在串行活动机制支持下围绕相关方面进行串行活动的机制,包括进行感知、记忆和思维等心智活动的各种机制,也就是说,不需要专门的定向机制,但人工系统可以根据需要配置强化的、专门的定向机制。

与一事物相关的事物是众多的,且是多层次的。对一事物的多个方面的属性或特征以及它和多方面事物之间所存在关系的确定,称为多重定向。其结果是可以给自己或相关事物相应多方面的说明。多层次定向是指由一事物对另一事物及其他更多层次事物的定向,或反过来的定向。多重定向和多层次定向以及人工意识系统对事物的理解,是意识系统内心世界所具有的一种功能或作用。

定向机制是系统具备一定理解力的基础,任何可以使系统具备、或增加、或扩大、或扩充、或增强定向能力的方法,比如增加或增强某种定向机制,都可以增强系统对事物的理解力。系统所能实现的定向越丰富、越清晰、越深刻,

理解也就越丰富、越清晰、越深刻。多重定向和多层次定向不仅使系统对事物的理解和相应的心智活动更加丰富、清晰和准确,而且也更加鲁棒。存在定向功能单一、定向功能不全或定向功能障碍的系统,将如人类患有老年痴呆症一般,不能完整地、准确地认知事物,或混淆事物,或错误地认知事物,包括它的来源、作用、事物之间的关系等。

(三)觉醒

不同于深度学习等人工神经网络系统,人工意识系统基本单位知识网络始终处于一个具有一定激活水平的状态,各种基本单位知识都同时处于一定水平的准激活状态,即意元网络中的意元处于相互关联的觉醒状态。

基本单位知识没有大小、没有内容、没有结构,除了它自身,再也没有任何可以说明或表明它自身的东西,只能由与之存在关系的其他基本单位知识来说明,并产生有关“它”的完整意义与理解。只有在处于觉醒状态的时候,基本单位知识网络的各种基本单位知识才能处于一个彼此完整关联的相互说明、相互支持的状态,才可能整体地对事物作出反映或反应,才可能构建系统内心关于事物的映像和系统内心关于世界的小宇宙,以更全面、完整、准确地理解一定的事物。这是一种由各种基本单位知识的大规模并行活动来实现的状态。在前面,我们讨论完形的概念、定向的意义、对事物的理解以及事物在内心的映像时,都需要领会这种觉醒的准备状态或准激活状态。所以,意识系统存在某种觉醒机制,以维持意元网络处于一定的准激活状态。

所谓觉醒状态,是指意元网络处于一定水平的准激活状态。还有两种与觉醒状态相关的睡眠状态:深睡眠状态和浅睡眠状态。深睡眠状态,即相对的不应状态;浅睡眠状态,即相对的可应状态(对强烈的刺激可以作出反应),觉醒状态则是在应状态(对任何刺激都可以作出反应)。觉察到的意元处于意识的中心状态,人工意识系统可以根据需要确定是否设置以及设

置什么样的“觉醒—睡眠”机制。

给人工意识系统中的意元网络配置一定的“觉醒—睡眠”机制,可以激励并维持意元网络整体或其中的一些部分处于一定的觉醒状态,和(或)抑制以使得意元网络整体或其中的另一些部分处于一定的睡眠状态。因此,“觉醒—睡眠”机制具有对意元网络的“觉醒—睡眠”状态进行调节与控制的作用。意元网络处于一定的觉醒状态时,系统基于对事物的理解而产生对相应事物的意识,即被激发的一定的基本单位知识,同时也处于被其他相关基本单位知识所说明或所定向的状态,而不仅仅是可被说明、可被定向、可被理解的状态。处于觉醒状态时,围绕一定事物而组织在一起的相关意元同时处于一定水平的激活状态,它们一起构成了关于该事物的映像或意识,其中关于该事物的基本单位知识处于意识的中心,是串行活动中相应那一时刻的内容或节点。“觉醒—睡眠”机制可以和基本单位知识网络融为一体,也可以是相对独立的装置,或是部分融为一体而部分相对独立的装置。“觉醒—睡眠”机制可以和串行活动机制融为一体,也可以是相对独立的装置;或是部分融为一体,部分相对独立的装置。可以说,基本单位知识网络和它的“觉醒—睡眠”机制以及串行活动机制三方面,既可以是相对独立的,也可以是以某种形式相互融合或结合在一起的,或是部分相对独立而部分融合的。

四、其他

(一)融合与扩展

意识系统是人工神经网络系统和物理符号系统的完美融合,也是知识叠加原理和形式化原理的完美融合。在人工智能系统中,知识叠加模块和形式化模块不仅可以完全融合在一起,而且可以以一定的方式彼此植入到不同原理的模块中,这样能使系统更加灵活且有效。

同样的原理可以应用于控制论系统,也可以和以控制论系统为代表的行为主义人工智能进行完美融合,以发挥它们各自的独特功能和

优势。在控制论系统中,控制问题被当作知识问题处理,可以采用同样的前述知识的原理与方法。也就是说,基于意识系统原理,可以实现与连接主义、符号主义和行为主义智能的完美结合。

依据意识系统原理,人工意识系统可以向知识叠加技术和形式化技术两端不断扩展,以发挥它们各自的优势,包括纵向的同一技术方法的深度扩展和横向的不同技术方法的扩展。

(二)继承

由于意识系统的性质,人工意识系统可以继承人工神经网络的成果,也可以继承形式化物理符号系统的成果。同时,由于具有继承的性质,分别作为人工意识系统一方面内容的知识叠加技术和形式化技术,不仅可以融合发展,也可以相对独立地发展。

(三)私有知识和共有知识

不同系统关于同一事物的叠加态知识是各自独立且不同的,是它们各自的私有知识,而本征态知识则是相同的共有知识。共有知识可以实现共享,而私有知识可以保证各自的独立和私有。

(四)从数字时代走向知识时代

在目前的人工神经网络系统中,输入的数据直接驱动人工神经网络产生结果,表现出的是“数据输入—结果输出”的数据驱动的工作方式,而在意识系统中或在基本单位知识网络系统中,主体是“知识 A—知识 B”的知识驱动的工作方式,尽管它也可能包含对输入数据进行知识的处理。

觉醒状态下意元网络处于一定水平的激活状态,因而各意元处于相互说明的状态中,从而使得串行心智活动可以由内部信息或知识来相互驱动,而不必由外部输入的信息来驱动,并由此可以在内心维持各种心智活动,使得系统随时可以就外界输入进行处理并给出说明或定

向,或作出其他反应。显然,这提供了一种以知识为核心的系统原理和方法,由此也可以真正定义“知识时代”。

(五)主观体验

目前这样的系统,尽管可以产生关于客观事物的映像、产生自己的小宇宙以及对内心事物进行各种心智操作,但却仍然是缺乏主观体验的系统。同样地,将其应用于人工情绪,除可以比较容易实现对情绪的认知成分和行为成分的处理,一样会缺少主观体验的成分。

有关系统的主观体验,还需要在未来的系统,特别是在本体觉中进行模拟探索。只有能够进行主观体验的系统,才是一个完整的意识系统。

本文所介绍的只是意识系统的体系结构的最基本构成,而要使机器真正可以产生类人意识,许多具体的技术方法和理论课题还有待深入探索。

[参考文献]

- [1] 谈加林.“整体优先”等说明了什么:论意识的基本单位:意元[J].现代教育研究(现名湖南师范大学教育科学学报),1995(4):80-86.
- [2] 谈加林.意原学说[J].现代教育研究(现名湖南师范大学教育科学学报),1996(1):60-65.
- [3] 谈加林.一种使机器产生知识的方法、装置、系统及存储介质[P].中国:202111275955.4,2021-10-29.
- [4] 谈加林.一种使机器产生意识的方法、装置、系统及存储介质[P].中国:202110584886.9,2021-05-27.
- [5] 谈加林.意识的原理:基本单位知识与知识叠加原理[J].长沙理工大学学报(社会科学版),2022(3):110-121,136.
- [6] Hornik K,Stinchcombe M B,White H.Multilayer Feed-Forward Networks Are Universal Approximators[J].Neural Networks,1989,2(05):359-366.
- [7] Cybenko G. Approximation by Superpositions of a Sigmoidal Function[J].Mathematics of Control Signals & Systems,1989,2(04):303-314.