

## 摘要

物理符号系统假设自 20 世纪 50 年代提出以来,已经有了近半个世纪的历史。作为认知科学发展史上较早提出的一个重要理论,在认知科学的领域中发挥了极大的作用,被称为“经典的”认知科学研究纲领。该理论的成功提出,吸引了大批学者进入这一研究阵营,并在以其为核心的符号系统范式的指导下,取得了一系列令人瞩目的工作成果,极大的推动了认知科学的发展。随着研究中任务域的逐渐扩大,该理论在指导解决诸如常识知识等问题时出现了一些困难,近年来开始受到了一些挑战。但其生命力依然顽强,它所提倡的“符号处理”基础假定的合理性以及所具有的解释力显示了它仍然具有极高的存在价值。

本文试图循着物理符号系统假设的发展历史,对物理符号系统假设展开历史回顾与思考。论文共分 4 章:第一章主要考察其历史渊源、产生背景,回顾了历史中与物理符号系统假设关联的人类思想和实践,旨在说明使用符号代替思维有其哲学思辨理由和科学实证根据。第二章主要通过阅读、整理纽威尔和西蒙的原著,详尽地阐释了物理符号系统的核心内容和附带推论,并通过例举来自人工智能、认知心理学领域的有效证据,合理地解释了该理论在提出后迅速获得成功的原因。第三章聚焦于物理符号系统假设的主要思想来源:亚里士多德的逻辑思想为物理符号系统假设形成打下了最根本的基础,弗雷格的人工语言为使用符号模拟人类思维过程解除了障碍,图灵关于计算机和智能以及脑和计算机在普适图灵机中的抽象数学表征的理论统一,是物理符号系统假设的先驱。这一章还描述了该理论建立所历经的五个阶段。第四章介绍了物理符号系统假设提出以来的一系列主要应用,分析了它所面临的主要困难,并论述它所受到的攻击和挑战。

正确理解物理符号系统假设,有助于认知理论的深化,在认知科学史中也具有价值。

**关键词:** 物理符号系统, 认知理论, 科学史, 范式, 纽威尔, 西蒙

## Abstract

Physical symbolic system hypothesis can be dated back to over 50 years ago when it was proposed in 1950s. As a vital theory in the early period of cognitive science development, physical symbolic system hypothesis, which is regarded as classical cognitive science, has made great contribution to this field. Thanks to the theory, a plenty of researchers joined in the research club. They gained series of remarkable achievements under the symbolic system paradigm, the core content of the theory, and thus paved the way for cognitive science development. With the broadening range of study tasks, the theory has been challenged by some problems, such as in resolving common sense problems. However, it is no doubt that the reasonability of the basic hypothesis and the capability of explanation of this theory has well demonstrated compellingly high value.

This thesis aims to review and reflect the history of physical symbolic system hypothesis. The paper is divided into four parts. The first part is mainly considering its history and background and reminding us of the human thoughts and practices related with physical symbolic system hypothesis in history so as to illustrate that using symbols instead of thinking has its own reasons of speculative philosophy and scientific empirical research bases. The second part provides detailed outlines for the core content and incidental inferences of physical symbolic system hypothesis by reading and collecting Newell and Simon's original works. Besides that, we also list out effective evidence of artificial intelligence and cognitive science to explain reasons why the theory succeeded as soon as it was brought forward. The third part focuses on main sources of physical symbolic system hypothesis: Aristotle's logical thoughts have made a fundamental basis for physical symbolic system hypothesis; Frege's artificial language has unblocked the obstacles of using symbols to simulate human thinking process; Turing has integrated the theories, which is the pioneer of physical symbolic system hypothesis. This part also describes the five steps of theory development. The fourth part introduces a series of applications, analyzes troubles and illustrates the opponency and challenges the hypothesis facing.

Understanding physical symbolic system hypothesis correctly can help us deepen the comprehension of the cognitive theory and add values to the history of cognitive science.

**Keywords:** Physical Symbolic System Cognitive Theory  
History of Science Paradigm Newell Simon

## 引言

“物理符号系统假设”(physical symbolic system hypothesis, 简称 PSSH), 主要由美国著名学者纽威尔(又译纽厄尔, A. Newell)和西蒙(又译司马贺, Herbert A. Simon)在上世纪五十年代提出的, 其在认知科学领域具有极大的影响。

认知科学作为一门研究认知的本质和规律的前沿性尖端学科, 跨越心理学、人工智能、神经科学、语言学、人类学和哲学这六大领域, 被认为是 20 世纪科学标志性的新兴研究门类, 受到了全世界科学家的广泛关注。熊哲宏在《认知科学导论》一书中曾写道, 认知科学是指科学地研究认知过程的任何一种学科, 它包含对心、脑产生智能行为的任何尝试。<sup>1</sup>目前学术界对认知科学有众多的定义, 唐孝威院士认为, 这在一定程度上是由于讨论的范围不同引起的, 狭义的认知科学指对脑处理知识的过程的研究, 广义的指对智能的研究, 包括人、动物和机器的智能的研究, 更广义的指对心脑的研究。唐孝威院士指出, 在认知科学的发展的历史上, 曾经有过许多关于认知的理论, 它们分别对认知过程提出不同的观点, 这些理论中就包括物理符号系统假设, 按唐孝威院士的看法, 物理符号系统假设是在认知科学发展和正式成立阶段较早提出的一个重要理论。<sup>2</sup>

纽威尔和西蒙的物理符号系统假设具有丰富的内容, 这种理论认为, 认知活动是以物理符号来表征的, 认知过程是个体对这些物理符号进行计算。纽威尔和西蒙说, 脑内存在对外部世界符号(称为物理符号)的表征, 而认知是在离散的时间对符号表征进行计算操作。这种认知理论实际上建立在计算机隐喻的基础之上: 把人脑比喻成是计算机, 把人的心智比喻成是计算程序, 认知过程像计算机的计算过程那样, 是对输入的符号进行信息加工然后输出的过程。在这个认知理论中, 基本的概念是表征和计算。他们对认知的理解基于表征和计算, 认为认知的内在信息加工等同于计算机按一定规则进行计算操作。<sup>3</sup>简单地来说, 就是企图将大脑的思维过程和认知活动转换成赋值、复制、书写等 10 种抽象符号操作, 从而把人类思维过程表述为一般物质运动的形式, 并借以在计算机上实现输入、

<sup>1</sup> 熊哲宏. 认知科学导论. 武汉: 华中师范大学出版社, 2002. 24-25.

<sup>2</sup> 唐孝威. 统一框架下的心理学与认知理论. 上海: 上海人民出版社, 2007. 129.

<sup>3</sup> 唐孝威. 统一框架下的心理学与认知理论. 上海: 上海人民出版社, 2007. 129.

输出、存储等6种功能，通过编程来实现对人类认知的模拟；这种理论尚待实践和时间的验证，所以创建者本着严格的科学精神，把它称为 PSSH(在纽威尔和西蒙合著的《作为经验探索的计算机科学:符号和搜索》和西蒙的《人类的认知:思维的信息加工理论》中，作者都将它称为 physical symbolic system hypothesis)，即尚处于假设(hypothesis)阶段。在我国一些文献里，通常将这一假设直接称为“物理符号系统理论”。

物理符号系统假设自提出直至20世纪80年代中期，在认知科学领域发挥了重要的作用，被称为“经典的”认知科学研究纲领，受到了较高的评价。美国认知科学家魏格曼(M. Wagman)在他的《认知心理学与人工智能:认知科学的理论和研究》一书中评价物理符号系统假设时写道:“纵观大部分心理学史，一般智能的概念主要来自心理测量技术的方法和结果。在近十多年，作为区别于智能测量应用技术的智能科学的可能性，来自认知处理的计算理论。就理论连贯性和经验结果来说的这些计算理论的多数影响，是被纽威尔和西蒙所系统化的物理符号系统假设。”<sup>4</sup>

在物理符号系统假设的指导下，认知科学领域中出现了一系列令人瞩目的研究、实践成果，这些成果极大地推动了认知科学的发展，同时也吸引了大批学者加入到这一研究阵营，并最终形成了以物理符号系统假设为核心，以西蒙和纽威尔这些人的初创性研究成果为范例的“符号系统范式”(symbolic system paradigm)。

20世纪80年代中期，随着日本“第五代计算机”研究计划的流产，物理符号系统假设的作用和地位受到了挑战，并渐渐转入低谷。但其生命力顽强，它所提倡的“符号处理”基础假定的合理性以及它所具有的解释力显示了它仍然具有极高的存在价值。

因此，研究物理符号系统假设是如何起源的？它的形成受到哪些思想的启发影响，经历了哪些过程？它的核心内容和主张是什么？它对后来的认知科学领域有怎样的影响？它自身理论的哪些局限，导致了它受到何种挑战？进一步厘清这些问题，在认知理论中具有一定意义，在认知科学史中也具有价值。本文试图循着物理符号系统假设的发展历史，对物理符号系统假设展开历史回顾与思考，进

---

<sup>4</sup> Morton Wagman. Cognitive Psychology and Artificial Intelligence: Theory and Research in Cognitive Science. Westport, Conn. : Praeger publishers,1993. 131.

而为认知科学的研究与讨论提供历史借鉴和启示。

本文的研究对象在科技史学界属于较前沿的研究领域。据前期资料搜集，国内相关研究较少，主要有以熊哲宏教授为代表，著有《认知科学导论》一本，以及发表有《关于符号处理范式在认知科学中的地位和前景》等文。《认知科学导论》一书，全面系统综合、概括、提炼认知科学自诞生以来，特别是20世纪90年代以来发展的新成就，并在此基础上进一步探讨制约认知科学发展的深层哲学问题。本书对于我了解物理符号系统假设的产生背景、影响、应用及其自身存在的哲学问题和面临的挑战有很大的帮助；同时，本书也为我研究物理符号系统假设的思想来源方面提供了一些材料和线索。但是，该书缺少物理符号系统假设的形成过程的相关研究，同时该书旨在从认知心理学的角度出发来概述整个认知科学的发展过程，并未对物理符号系统假设进行系统的梳理和分析。《关于符号处理范式在认知科学中的地位和前景》一文在概述了“符号处理范式”的理论框架和基本原理的基础上，就联结主义对该范式的挑战以及符号处理论者的应答与辩护作了客观的述评。作者主张符号处理范式仍具有较强的生命力，并不因为它受到联结主义的挑战就会失去自身存在的价值。该文对于帮助我理解物理符号系统假设的核心内容和主张有相当大的帮助；同时它部分论述了物理符号系统假设所面对的联结主义提出的挑战，对于进一步理解物理符号系统假设所存在的局限提供了帮助和线索；最后，它也简略地提到了物理符号系统假设的数学和哲学背景是图灵的初创性工作，这对于研究物理符号系统假设的思想来源提供了重要线索。但除了简略提到图灵，该文对于物理符号系统假设的其他思想来源没有过多论述，所以要厘清来龙去脉，还需要靠研读西蒙和纽威尔的原始文献。

本文通过精读西蒙和纽威尔的相关原始文献，提出问题；阅读研究文献，扩大线索；最后，进行深入研究，完善论述。旨在对物理符号系统假设形成一定深度的理解，努力寻找以下5个问题的答案：

- 1、物理符号系统假设的思想来源是哪些？
- 2、它的核心内容和主张是什么？
- 3、它的形成具体经过了哪些过程？
- 4、它对于认知科学有什么影响？
- 5、该理论的自身局限在哪里，以至于它受到了何种挑战？

论文的创新点在于，从科技史、认知科学史角度出发，将物理符号系统假设作为研究对象提出问题，并搜集材料进行系统的整理和研究。首先，以新的视角(人类使用符号代替人类思维)较为全面地论述了物理符号系统假设形成的背景，以便人们了解该假设背后的坚实的思想基础；其次，着重分析了它的主要思想来源；第三，以新的视角(人类试图通过编程让计算机去模拟人类的认知过程)去解读物理符号系统假设所产生的影响和所面临的困境；最后，对其合法性受到的攻击和应用层面受到的挑战，给出了清楚的分析。目前，该工作在目前国内学术界属于起步阶段，是一个具有活力的较新的研究方向。

## 第1章 物理符号系统假设(PSSH)产生背景

艾伦·纽威尔(Allen Newell)是美国科学院、工程院院士,美国人工智能学会的发起人之一,曾任美国人工智能学会、美国认知科学学会主席,1971年美国信息处理协会联合会授予古德纪念奖,1975年荣获计算机科学最高奖——图灵奖,1992年荣获美国总统科学奖,代表作:《IPLV语言手册》、《计算机结构》、《计算机与数字系统设计》、《认知的统一理论》、《SOAR:一般智能的体系结构》。

赫尔伯特·A·西蒙(Herbert A. Simon)是美国著名心理学家、管理学家,曾任卡内基—梅隆大学计算机科学与心理学教授,美国伊利诺伊理工学院教授,1975年荣获计算机科学最高奖——图灵奖,1978年荣获诺贝尔经济学奖,1986年荣获美国总统科学奖,代表作:《管理行为》、《经济学和行为科学中的决策理论》、《管理决策的新科学》等。他在心理学与认知理论方面也做出了突出贡献。

物理符号系统假设作为认知科学领域的最为重要学术思想之一,由纽威尔、西蒙和肖(C. Shaw)三人于20世纪50年代末首次提出。要考察其产生背景,我们就需要对人类发展历史中与物理符号系统假设关联的片段进行回顾,我们可以追溯到古代、近代和现代的人类相关思想和实践,这对了解物理符号系统的产生有一定的帮助。本章主要从人类使用符号代替人类思维的角度来提出历史材料,重点在于展示这段历史中,人类是如何发展用符号来替代人类思维的。这区别于以往计算机史站在机械、存储单元形成发展的角度来描述历史。

本章按照时间维度提出历史材料,包括三个部分:萌芽时期,形成时期和发展时期,时间分界点为19世纪末和20世纪50年代末。以19世纪末为第一个时间分界点,是根据纽威尔和西蒙在著名论文《作为经验探索的计算机科学:符号和搜索》(Computer science as empirical inquiry: symbols and search)中曾明确提到,物理符号系统假设的形成经历了“形式逻辑”、“图灵机与数字计算机”、“存储程序概念”、“表处理”、“LISP语言”这五个阶段(相关内容将在本文第三章详细论述),并强调“形式逻辑”阶段从弗雷格(G. Frege)、怀特海(A. N. Whitehead)和罗素(B. Russell)就形式化逻辑提出方案开始,而这个时间就是19世纪末。以20世纪50年代末为第二个时间分界点,是因为纽威尔、西蒙和肖在20世纪50年代末



首次提出物理符号系统假设，西蒙在1986年出版的《人类的认知：思维的信息加工理论》一书中曾写道：“物理符号系统假设提出来已经30年了”<sup>5</sup>，也再一次肯定了物理符号系统假设的提出时间为20世纪50年代末。

## 1.1 萌芽时期(19世纪末之前)

在古希腊的众多学说中，我们就可以找到一些与物理符号系统假设有着丝丝关联的思想。毕达哥拉斯学派在思考世界本源的时候，就主张“思维可计算”，他们曾提出“万物皆数”的这一观点，他们认为，万物归于数，观念(逻各斯)与数、思维与数的演算是同一的，现在我们在物理符号系统假设中也可以隐约地看到这一思想的影子。其后，古希腊思想家亚里士多德作为演绎推理理论的创始人，研究了思维的形式和规律，第一个把意识的理性部分形式化，他在其主要逻辑著作《工具论》中提出了一种非形式化的三段论逻辑系统和一些重要的逻辑规律，允许在初始前提的条件下机械地推导出结论。公元前1世纪伊壁鸠鲁学派明确的谈到了逻辑和计算的同一性。这些思想为之后物理符号系统假设的形成打下最为根本的基础。

12世纪末至13世纪初年间，西班牙神学家和逻辑学家拉蒙·鲁尔(Ramon. Lull)在一定程度上发现了概念的组式，他开始用字母即数字符号来表达概念，用“+”、“-”、“×”等来表示概念的运算，试图得到一种逻辑运算，并在1274年创制了机械装置Ars Magna用来向异教徒证明基督教真理。14世纪，意大利著名的艺术家、科学家达·芬奇(Leonardo da Vinci)试图设计机械计算器，虽然最后没有成功，但是极大地推进了计算自动化的思想形成。

17世纪，英国哲学家霍布斯(T. Hobbes)提出推理就如同数字计算，他在《利维坦》中写道，“当一个人进行推理时，他所做的不过是在心中将各部相加求得一个总和，或是在心中将一个数目减去另一个数目求得一个余数。这种过程如果是用语词进行的，他便是在心中把各部分的名词序列连成一个整体的名词或从整体及一个部分的名词求得另一个部分的名词。”“推理就是一种计算，也就是将

---

<sup>5</sup> [美]司马贺(Herbert A. Simon). 人类的认知：思维的信息加工理论. 荆其诚、张厚粲译. 北京：科学出版社，1986. 14-15.

公认为标示或表明思想的普通名词所构成的序列相加减。”<sup>6</sup> 同时期的,被西方哲学界称为“理性主义之父”的法国数学家、科学家和哲学家笛卡尔(R. Descartes)提出了“思维可计算”设想。1642年,法国物理学家和数学家帕斯卡尔(B. Pascal)制成了世界上第一台会演算的加法器,在当时获得了广泛的实际应用,他认为“数学机器得出的结果,要比动物所做出的一切更接近于思想。”<sup>7</sup>17世纪在这方面最具成就的是德国数学家和哲学家莱布尼兹(G. W. Leibniz)。1670年,莱布尼兹在帕斯卡尔加法器的基础上制造四则运算机。之前他还设想建立一种代替自然语言的“普遍的符号语言”,通过字母和符号进行逻辑分析和综合,按照代数的原则,把逻辑推理变成数学演算,他确信所有依靠推理的问题通过数字的变换演算,就会直接促进完美答案的发现,他说:“如果有人怀疑我的答案,我将告诉他:‘先生,让我们演算一下’。这样通过笔墨,我们就会立刻解决这个问题。”<sup>8</sup> 莱布尼兹这种“万能符号”和“推理计算”的思想和实践是现代计算机模拟人类认知的最早萌芽。

1823年,英国研究数学、计算技术和力学的科学家巴贝奇(C. Babbage)开始制造差分机(difference engine),差分机可以五阶差分表示来计算任何函数。随后巴贝奇又提出分析机(analytical engine)的设计,预期实现按序执行、分支选择、循环控制功能,他的儿子在他去世后完成了分析机的主要部件,并进行了某些运算试验,这台机器实际上是一台非完全自动的博弈机,是近代智能机器的最高成就。1847年至1854年,英国数学家乔治布尔(G. Boole)致力于使“思维规律”具有形式化和实现机械化的研究,最终他发明了布尔代数又称为布尔逻辑,这既为现代计算机的电路设计预备了数学工具,又为人类逻辑思维的形式化处理奠定了基础。

## 1.2 形成时期(19世纪末至20世纪50年代末)

1897年,德国著名的数学家、逻辑学家弗雷格发表了著作《概念文字:一种

<sup>6</sup> [英]霍布斯(Thomas. Hobbes). 利维坦. 黎思复、黎廷弼译. 北京:商务印书馆,1986. 27-28.

<sup>7</sup> [法]帕斯卡尔(B. Pascal). 思想录. 何兆武译. 北京:商务印书馆,1997. 156.

<sup>8</sup> [德]莱布尼兹(G. W. Leibniz). 莱布尼兹自然哲学著作选. 祖庆年译. 北京:中国社会科学出版社,1985. 13.

模仿算术语言构造的纯思想的形式语言》，创立了概念语言，日常语言表达严密思想的障碍在一定程度上得到了解决。同时，意大利数学家皮亚诺(G. Peano)发明了一种表意语言，以解决语言含混而致使的数学发展缓慢的问题，这种语言符号简单清晰，对于剖析算术、几何的原理，辨别其基本从属概念、定义、公理和定理有极大的帮助，也能够很好陈述日常语言几乎无法表达的复杂思维过程。之后，英国哲学家、数学家罗素初步建立了一个自足的、完全的逻辑系统。

1927年，怀特海在《符号主义》一书中就提出：“人的心智是符号在进行活动，即经验的某些成分发出意识、信念和情感，而这些活动又与经验中的其他成分相关联。这里，前一种成分的集合就是‘符号’，后一种成分的集合则构成了符号的‘意义’。这种将符号与它的意义相联结的过程就是‘符号的所指’。”<sup>9</sup> 他的思想对于日后物理符号系统假设的建立有着很大的影响。1936年，图灵在论文《论应用于解决问题的可计算数学》中给出算法概念严格的数学表达，首次产生了符号处理的概念，并开始转向了“可改变的编码程序”。1937年，他在论文《理想计算机》中提出了一个独特的设想，即著名的图灵机——理想计算机模型，在论文中图灵对计算理论作了精辟的论述。同年，丘奇(A. Church)和图灵分别发展了丘奇—图灵论题，该论题认为人类所能解决的所有问题都可归结为一套算法，机器智能与人类智能基本上是相等的。他们的工作建立了一个严格、精确、适用的逻辑工具，并将它与现实世界的可逻辑化和符号化的思想结合起来。1938年德国工程师朱斯(Zuse)研制成了第一台纯机械结构式的累计数字计算机。这些思想和实践都为以后的物理符号系统假设的诞生奠定了基础，从此以后人类存储和处理信息的方式开始发生革命性的改变。

20世纪40年代，随着二战的爆发，战争对科学技术的需求极大地刺激了计算机科学的发展，用计算机来模拟人类智能行为的研究开始兴起，大批物理界、数学界精英都参与到新型智能机器的研究中来。1945年，匈牙利数学家、博弈论的创立者冯·诺依曼(John. Von. Neumann)在《关于离散变量自动电子计算机的草案》一文中提出了存储程序的概念，并提出了相关的电子计算机模型——“冯·诺依曼机”，这一思想不但成为现代电子计算机的范式，更宣告了电子计算机时代的开始。与此同时，图灵也形成了相似的观点，并亲身参与了第一台能完全执行

<sup>9</sup> 车文博、朱新明、李亦菲. 架设人与计算机的桥梁：西蒙的认知与管理心理学. 武汉：湖北教育出版社，1999. 63.

存储程序的电子计算机模型机的设计，即“曼彻斯特机”，并于1949年底取得成功。

1948年，美国数学家香农(C. Shannon)发表《通讯的数学理论》，创立信息论，该理论用数理统计方法来研究信息的度量、传递和变换规律，主要是研究通讯和控制系统中普遍存在着信息传递的共同规律以及研究最佳解决信息的获限、度量、变换、储存和传递等问题的基础理论。同年，美国数学家维纳(N. Wiener)发表了关于信息理论的创新性著作《控制论》一书，并提出控制论(主要包括信息论、反馈理论)。这在历史上可以认为是实现机器智能在哲学、理论及方法方面的一次全面探讨。控制论成为人们研究比较简单(相对而言)的系统以及系统运行环境并不复杂情况下的一面旗帜。维纳在反馈理论上的研究最终让他提出了一个论断：所有人类智能的结果都是一种反馈的结果，通过不断地将结果反馈给集体而产生动作，进而产生智慧。比如，家里的抽水马桶就是一个十分好的例子，水之所以不会常流不断，正是因为有一个装置在检测水位的变化，如果水太多了，超过了规定的容量，水管就会自动关闭，这就实现了反馈，它是一种负反馈。维纳认为：“如果连我们厕所里的装置都可以实现反馈了，那我们应该可以用一种机器实现反馈，进而实现人类智力的机器形式重现。这种观点和想法对于认知科学早期的理论有着重大的影响。”<sup>10</sup>西蒙曾经这样评价：“反馈概念对心理学有相当大的、但相对来说非特定的影响，但香农—韦弗的信息论对心理学的影响则是清楚明确的。”<sup>11</sup>

1950年，图灵在其论文《计算机器和智能》中提出了“机器能够思维吗”<sup>12</sup>的问题，对“机器思维”进行了定义，明确地提出了计算机能够被程序控制及可以显示智能行为的观点，并对人工智能的研究提出了初步的设想。之后，米勒(G. Miller)在《魔法般的数学7，加上或减去2：我们信息处理的能力之局限》的论文中，把信息处理概念用于人类的注意和记忆研究，提出关于短时记忆容量有限的理论。米勒提出，思维过程的信息容量，尤其是短时存储，最好是按照有意义的信息单位——语义块来测量，而不是照抽象的位来测量。<sup>13</sup>这些成果，使“思维

<sup>10</sup> 武秀波、苗霖、吴丽娟、张辉. 认知科学概论. 北京: 科学出版社, 2007. 6.

<sup>11</sup> [美]赫尔伯特·A·西蒙(Herbert A. Simon). 我生活的种种模式——西蒙自传. 曹南燕, 秦裕林译. 上海: 东方出版社, 1998. 252.

<sup>12</sup> A. Turing. Computing machinery and intelligence. Mind, 1950(59). 433.

<sup>13</sup> 熊哲宏. 认知科学导论. 武汉: 华中师范大学出版社, 2002. 16.

可计算”设想部分地成为现实。

1956年夏季, 麦卡锡(J. McCarthy)等人在达特茅斯(Dartmouth)召集了人工智能夏季研究会, 会上正式使用“人工智能”(artificial intelligence)一词, 这标志着人工智能学科的正式诞生。作为认知科学重要组成部分的人工智能学科正式成立以后, 直至20世纪80年代初期, 这门学科开始迅速发展。

1956年, 纽威尔和其他的研究人员一起开发了人工智能的历史上有重要地位的“逻辑专家”(the logic theorist)程序, 它采用书中搜索, 寻找与可能答案接近的树的分枝进行搜索。IBM公司研究小组进行电脑程序模拟人类学习过程的相关研究, 并开发出了具有自学习、自组织、自适应能力的西洋跳棋程序。该程序可以模拟优秀棋手, 通过看几步来弃棋, 还能学习棋谱, 在分析大约17500个不同棋局后, 可以猜测出书上所有的推荐的走步, 准确率可以达48%。同年, 西蒙完成了编制人工智能程序的表处理技术。之后他还设计与实现了表处理语言IPL(information processing language)。在人工智能的历史上, IPL是所有表处理语言的始祖, 也是最早使用递归子程序的语言。其基本元素是符号, 并首次引进表处理方法。IPL最基本的数据结构是表结构, 可用以代替存储地址或有规则的数组, 这有助于将程序员从繁琐的细节中释放出来而在更高的水平上思考问题。IPL的另一特点是引进了生成器, 每次产生一个值, 然后挂起, 等待被调用, 在调用时从被挂起的地方开始。早期的很多人工智能程序都是用表处理语言编制而成的。1957年, 纽威尔、西蒙和肖等人所在的心理学小组编制出了一个称为“逻辑理论机”LT(the logic theory machine)的数学定理证明程序, 当时该程序证明了罗素和怀特海的《数学原理》一书第二章中的38个定理(后来在1963年改进后的LT程序证明了第二章中全部的52个定理)。后来他们又揭示了人在解题时的思维过程大致可以归纳为三个阶段: 1、先想出大致的解题计划; 2、根据记忆中的公理、定理和推理规则组织解题步骤; 3、进行方法和目的分析, 以便发现问题, 进而修正解题计划。这种思维过程不仅解数学题是如此, 解决其他问题时也大致如此。<sup>14</sup>20世纪50年代末, 纽威尔、西蒙和肖三人首次提出了物理符号系统假设, 物理符号系统假设正式诞生。

<sup>14</sup> 林尧瑞、郭木河. 人类智慧与人工智能. 北京: 清华大学出版社, 2001. 80.

### 1.3 发展时期(20世纪50年代末至20世纪80年代初)

物理符号系统假设正式提出之后,纽威尔、西蒙等人就问题表示、搜索技术和通用启发等一系列问题进行了探讨,利用符号表示方法和逻辑推理方法,通过计算机的启发式编程方法,按心理学和人类认知的过程建立一些人类求解问题的过程模型,并把它们运用于计算机程序中来解谜、博弈和检索信息。1960年,纽威尔、西蒙和肖三人又共同编写了著名的“通用问题求解程序”GPS(*general problem solver*),该程序可以解决十种类型不同的问题,包括简单的智力测试、命题演算定理的证明和机器编程等,它表明用计算机程序的确可以准确地模拟人类的一类智能行为,这是在认知科学应用领域的一个成功。其后,他们还开展了其他方面有关的研究工作,以实现模拟人类的认知过程,如人的口语学习和记忆模型,自然语言理解程序和启发式求解方法等。

1972年,纽威尔和西蒙发表了《人类的问题解决》一书,提出了认知的物理符号系统假设,认为认知活动是以物理符号来表征的,认知过程是个体对这些物理符号进行计算。<sup>15</sup>

纽威尔和西蒙在书中宣称:“由计算机操作的二进制数串能够表达包括现实世界的任何东西,大脑和心灵与计算机一样,都不外是一种物理符号系统,无论它们在结构和动力机制上有多大不同,在计算机理层次上都是具有产生、操作和处理抽象符号的能力。”“计算机本身的活动在某些方面类似与认知过程。计算机接受信息、处理符号,在记忆中存储一件件事情,再检索它们、分类输入、识别模式等。它们是否像人一样地做这些事情并不重要,重要的是它们完全做到了。计算机的出现提供了必要的证据,证明认知过程是实在的。”“计算机不仅是处理数值符号,而且是处理任何一种符号的设备。因此计算机成为一种指定任意的处理方法。这种类型的理论可以称为信息处理理论,它们基本上是非数量的(可能完全不包括数目)。虽然和经典的数学理论比起来,它们不够精确也不够严密。”

16

1958年,麦卡锡在“意见采纳者”的系统中建议采用谓词演算这种语言来

<sup>15</sup> 唐孝威. 统一框架下的心理学与认知理论. 上海: 上海人民出版社, 2007. 129.

<sup>16</sup> [美]科恩(R. P. Cohen)、费根鲍姆(A. E. Feigenbaum). 人工智能手册(第三卷). 周少柏、黄汛译. 北京: 科学出版社, 1991. 4.

表示和运用知识。从此，谓词演算和一些它的变形构成了人工智能知识表示的基础。1959年麦卡锡的MIT小组专为符号表示方法发明了一种表处理语言LISP，成为了人工智能程序设计的主要语言，直至今日还被广泛采用。与此同时，其他研究者还给出了一系列包括象一阶谓词逻辑、产生式系统、语义网络、框架表示方法在内的知识表示方法。1972年，美国麻省理工学院的威诺甘德(T. Winograd)研制出“自然语言理解系统”SHEDLU，这是一个在“积木世界”中进行英语对话的自然语言理解系统。威诺甘德的积木系统把句法、语义、推理、上下文和背景知识灵活的结合于一体，模拟一个能够操纵桌上的一些玩具积木的机器人手臂，用户通过人机对话方式命令机器人放置那些积木块，机器人系统能够接受人的自然语言指令，进行积木的堆叠操作，并能回答或者提出比较简单的问题。系统通过屏幕给出回答并显示现场的响应情景。该系统的成功为自然语言在计算机上处理做出了巨大贡献。同时，法国马塞大学Colmerauer创建了PROLOG语言，美国人香克(R. Schank)提出了概念从属理论。在符号系统范式下进行的相关研究开始进入兴旺时期。

1968年，被誉为“专家系统和知识工程之父”的美国著名人工智能专家费根鲍姆(E. A. Fergenbaum)设计了第一个能演示具体领域知识的专家系统DENDRAL，它可以根据所提供的化学分子式和质谱分析图来预测有机物分子的结构。之后，其他“专家系统”也相继出现，包括医疗诊断专家系统、计算机结构设计专家系统、符号积分与定理证明专家系统、钻井数据分析专家系统和电话电缆维护专家系统等。这些系统的性能可以与同类专家的智能相匹敌，这是物理符号系统假设当时在应用研究方面的重大进展。从GPS诞生到专家系统的不断出现，表明以逻辑为基础的符号计算(处理)方法，无论在智能模拟上，还是在智能系统的建造上都同样能取得成功。1977年，费根鲍姆进一步提出了“知识工程”(knowledge engineering)的概念，这一思想成功地指导了许多智能系统的研制，并被广泛地应用于各个领域。<sup>17</sup>

20世纪80年代初期，物理符号系统假设正式确立，西蒙在1980年宣称：“在过去的25年里，就我们对人类思维过程的理解方式而言，任何社会科学的发展都比不上这次革命(即所谓信息处理革命)那么激进。”<sup>18</sup>

<sup>17</sup> 王振友、谢青松. 人工智能学科的发展. 山东工程学院学报, 第15卷, 第4期, 2001.

<sup>18</sup> [美]黎黑(T. H. Leahey). 心理学史: 心理学思想的主要趋势. 李维译. 杭州: 浙江教育

## 第 2 章 PSSH 核心内容和有效证据

1976年, 纽威尔和西蒙在《计算机协会通讯》(Communications of the ACM) 上发表了著名的论文《作为经验探索的计算机科学: 符号和搜索》。在论文中, 纽威尔和西蒙对智能的本质进行了深入的分析。他们认为“因为所有信息都是为一些目的服务而由计算机加工的, 而我们衡量一个系统的智能水平, 是看它在面临任务环境所设置的种种变动、困难和复杂性时, 达到规定目的的能力。当所完成的任务范围有限时, 计算机科学在实现智能过程中的这一总的投入并不引人注目, 因为这时可以准确地预见这一环境中的全部变动。当我们将计算机扩展到更为综合、复杂和知识密集型的任务中去时, 亦即我们试图让它们作为我们的代理者, 能够独自处理自然界中的全部偶发事件时, 它就变得较为醒目了。”<sup>19</sup> “正如不存在能通过自己的特殊性质表示生命实质的‘生命原理’一样, 也不存在任何‘智能原理’。然而, 没有简单明了的解决办法, 并不意味着只能在构造上没有任何必备条件, 一个这样的必备条件就是存储和处理符号的能力。”<sup>20</sup> 因此, 他们认为“符号是智能行动的根基”, 而智能水平则取决于系统处理符号的能力。“我们这里的目标并不是从外部世界做出理解。我们是要考察这门科学的一个方面——通过经验探索而形成的新的基本理解。”<sup>21</sup> 这个“新的基本的理解”就是物理符号系统假设。

### 2.1 科学理论层次说明和界定

#### 2.1.1 科学理论层次说明

在《人类的认知: 思维的信息加工理论》一书中, 西蒙对于科学理论的层次

---

出版社, 1998. 754.

<sup>19</sup> Allen Newell, and Herbert A. Simon. Computer Science as Empirical Inquiry: Symbols and Search. Communications of the ACM, Vol.19, No.3, 1976. 114-115.

<sup>20</sup> Allen Newel, and Herbert A. Simon. Computer Science as Empirical Inquiry: Symbols and Search. Communications of the ACM, Vol.19, No.3, 1976. 115.

<sup>21</sup> Allen Newell, and Herbert A. Simon. Computer Science as Empirical Inquiry: Symbols and Search. Communications of the ACM, Vol.19, No.3, 1976. 115.



和规律进行了说明。西蒙以成熟的物理学为例，他这样写道：“物理学有不同的层次，例如，对于一个杯子，物体力学可以研究它的原子结构，也可以研究它与其它物体的相互关系，还可以研究它的离子运动。我们不必等待原子结构问题解决之后再研究整个物体力学；要研究一个杯子的下落，也不必同时去研究构成杯子的原子结构。”<sup>22</sup>

根据上面的论述，西蒙对于物理符号系统假设的合理性进行了说明。西蒙认为，就如同物理学那样，人工智能、心理学也要从不同的水平来研究对象的行为。对于人工智能的研究而言，不一定要完全对于计算机的硬件机构有了清楚的研究，才开始研究软件，我们可以脱离硬件研究，单独的进行软件研究；而对于心理学的研究而言，也没有必要非要等到对神经元、神经突触有了清楚的了解和系统的认识之后，才提出生理学理论，同样，也可以在生理学理论没有完全弄清楚之前，单独着手研究人类的高级复杂行为。西蒙将人工智能、心理学的研究分为三个层次和水平：“第一级水平是研究复杂性行为(在人工智能中就表现为如何模拟人类的复杂性行为)，例如研究问题解决、概念形成和语言现象；第二级水平是研究简单的信息加工过程(在人工智能中就表现为如何模拟人类的简单的信息加工过程)，例如对光点的感觉、图形知觉的形成都是简单的信息加工。”“第三级水平是生理水平，例如对中枢神经工程、神经结构的研究。”<sup>23</sup>物理符号系统假设被定义为第一级水平的科学理论和研究纲领。

### 2.1.2 西蒙对于 PSSH 的界定

在对物理符号系统假设所处的科学理论层次进行了阐述之后，西蒙又对物理符号系统假设的性质进行了界定，进一步维护了物理符号系统假设的合理性。首先，西蒙将科学理论分为“定量结构定律”和“定性结构定律”。“定量结构定律”主要以成熟的物理学为代表，其中牛顿(I. Newton)的万有引力定律就是典型的“定量结构定律”这种定律有定量的分析和描述，有定量的公式，非常具有说

<sup>22</sup> [美]司马贺(Herbert A. Simon). 人类的认知：思维的信息加工理论. 荆其诚、张厚粲译. 北京：科学出版社，1986. 1.

<sup>23</sup> [美]司马贺(Herbert A. Simon). 人类的认知：思维的信息加工理论. 荆其诚、张厚粲译. 北京：科学出版社，1986. 2.

服力。但是，除了“定量结构定律”，在其它很多学科中还存在着这“定性结构定律”，物理符号系统假设就是一个“定性结构定律，”西蒙说，“我们曾设想心理学(也包括人工智能)要有这种规律(定量结构定律)才能成为真正的科学。但这种观点比较狭隘，而且不切实际。其他科学领域的规律也不一定都是定量的分析。”<sup>24</sup>接着，西蒙在《作为经验探索的计算机科学:符号和搜索》一文中举例了“生物学中的细胞学说”、“地质学中的板块构造学说”、“细菌致病理论”来支持他的观点：

“生物学中的细胞学说”，该学说是指“一切活的有机体的基本组成单元是细胞”。通过后续的研究，该学说进一步细化为：无论各种形式的细胞，都由中间的细胞核、周围的细胞质和最外层的细胞膜组成。这个学说是很好的定性结构定律的一个例子，西蒙在评价该学说时写道：“这一定律给生物学带来的影响是巨大的，在这一学说被逐渐接受之前，该领域丧失的活力是相当可观的。”<sup>25</sup>

“地质学中的板块构造学说”，该学说认为：地球表面是由数十块大板块拼集而成的，它们以地质学的速率运动着，相互重叠、背离，或是向下进入地心直至失去原貌。大陆和海洋的形状、相对位置以及火山、地震活动区域等现象都能通过这个学说得到很好的解释。另外该学说还可以成功的解释许多细节，包括西非与南美东南部之间动植物群、底层的一致性等问题，近10年来，它得到了承认，西蒙在评价该学说时，认为它是一个“令人感兴趣”的例子。

“细菌致病理论”，该理论认为：大多数疾病是由于身体中的微小的、活的单细胞有机体的存在和繁殖而引起的，疾病的传染是由于这些有机体从一个宿主传播到了另一宿主。在该理论获得承认后，很多研究在它的指导下来进行工作，以确定特定疾病与特定有机体的联系，并对这些有机体的情况进行描述。西蒙在评价该理论时，认为它是一场引起医学界革命的定性结构定律，虽然有很多例外情况——很多疾病不是由细菌引起的，但这并没有降低它的重要性。此外，西蒙还强调，人工智能研究深受细菌致病理论范式的启发。细菌致病理论的基本范式是：确定疾病，然后寻找细菌；而人工智能研究是确定需要智能的任务域，然后

<sup>24</sup> [美]司马贺(Herbert A. Simon). 人类的认知：思维的信息加工理论. 荆其诚、张厚粲译. 北京：科学出版社，1986. 3.

<sup>25</sup> Allen Newell, and Herbert A. Simon. Computer Science as Empirical Inquiry: Symbols and Search. Communications of the ACM, Vol.19, No.3, 1976. 115.

为数字计算机构造程序，使其能完成这一任务域中的任务。

最后，西蒙还例举了“原子论学说”来证明“定性结构定律”的重要性和合理性。“原子论学说”一开始是一种“定性结构定律”，它指出：元素是由微小的、均匀的颗粒组成的，一种元素的颗粒不同于另一种元素。但是，基于原子种类的简单性和变化范围的有限性，它很快的转变为一种“定量结构定律”，但是它“吸收了原来定性假设中的所有一般性结构”。西蒙将物理符号系统假设界定为一种“定性结构定律”，并再一次强调“定性结构定律”是整个科学赖以运作的说明方法。

## 2.2 PSSH 核心内容及其附带推论

### 2.2.1 核心内容

在说物理符号系统假设的核心内容前，我们首先要对“物理”、“符号”、“操作”和“表达式”进行说明。西蒙曾写道：“有的符号系统由玻璃和金属制成(计算机)，有的由血肉组成(大脑)。过去，我们比较习惯于认为数学和逻辑的符号系统是抽象的、没有形体的，而不考虑纸、笔、人脑这些东西。事实上，必须有它们才能赋予符号系统生命。计算机将符号系统从柏拉图的理念天国(platonic heaven of ideas)输运到发生着实际过程(完成这些过程的是机器或大脑，或协同工作的两者)的经验世界来了。”<sup>26</sup>在符号系统前冠以“物理”两字，主要是强调所研究的对象是一个具体的物质系统，如计算机的构造系统，人的神经系统、大脑神经元等。有人将人比喻成“符号”的动物，比如卡西尔(E. Cassirer)就曾经说过：“除了在一切动物种属中都可看到的感受器系统和效应器系统以外，在人那里还可发现可称之为符号系统的第三环节，它存在于这两个系统之间。这个新的获得物改变了整个的人类生活。与其他动物相比，人不仅生活在更为宽广的实在之中，而且可以说他还生活在新实在之维中。”<sup>27</sup>在一般意义上，“符号”是泛指代表、表示或意指他物的一切东西，可以理解为人与自然、人与社会、

<sup>26</sup> [美]司马贺(Herbert A. Simon). 人工科学：复杂性面面观. 武夷山译. 上海：上海科技教育出版社，2004. 22.

<sup>27</sup> [德]卡西尔(E. Cassirer). 人论. 甘阳译. 上海：上海译文出版社，1985. 33.

人与人之间发生联系的意义系统，是人类独有的特殊标志。但是在物理符号系统假设中，“符号”有其特定的意思。西蒙是这样定义的：“所谓的符号就是模式(pattern)，任何一个模式，只要它能和其他模式相区别，它就是一个符号。”<sup>28</sup>。一页书上不同的英文字母就是不同的符号。符号既可以是物理的符号，这包含各种载体上的文字符号、声波和光波等；也可以是头脑中的抽象的符号，这包含语言、表象等；还可以是计算机中的电子运动模式和头脑中神经元的某种运动方式。“物理”、“符号”两词指明了该系统的两个重要的特征：首先，物理符号系统显然是遵循物理学定律的，它们可由工程化工量构成的工程化系统(如人脑和计算机等)来是实现；其次，虽然在使用了符号这一术语，预先勾画的是意向式的解释，但是它并不仅仅局限于人类符号系统(语言)，任何可以被认知器官或者认知功能所辨别的有意义的模式都可以归入符号这一范畴。而“表达式”(expression, 或符号结构)则由一些以某种物理方式(如一个符号实体紧接着另一个符号实体)相联系的符号实体(或标记)组成。在物理符号系统假设中，还对符号的“操作”进行了定义，所谓“操作”就是对符号进行比较，即找出哪些符号是相同的，哪些符号是不同的。

对表达式、符号和对象而言，纽威尔和西蒙提出了两个核心观念：“指称”和“解释”。所谓“指称”就是说，在已知表达式的情况下，一个系统能够对对象本身施加影响或者能够以取决于对象的方式规范其行为。其实质是，在每种情况下，结果都是经由表达式抵达对象。所谓“解释”就是说，已知一个表达式，系统就可以完成所指定的过程，也即是，系统能够根据指称这些过程的表达式而再现和执行它所拥有的过程。一个具备指称和解释能力的系统也必然满足若干新增的必然条件，具有完备性和封闭性：1、符号可以用来指称任何一种表达式；2、存在着一些表达式是指称计算机所能完成的每一过程的；3、存在着一些过程是以任意方式建立任何表达式并修正任何表达式的；4、表达式具有稳定性，它们一旦建立，在明确地被修改或被取消之前会一直存在下去；5、系统所能拥有的表达式的个数基本上是无限制的。

关于“物理符号系统”，纽威尔和西蒙是这样定义的：“一个物理符号系统是由一组叫做符号的实体组成的，这些实体是一些物理模式，可以作为另一种叫

<sup>28</sup> [美]司马贺(Herbert A. Simon). 人类的认知：思维的信息加工理论. 荆其诚、张厚粲译. 北京：科学出版社，1986. 10.

做表达式(或符号结构)的实体的分量而存在。”<sup>29</sup>物理符号系统是一架机器,存在于一个对象世界之中,这些对象的数量比符号表达式本身更多,它产生出一个随时间而演化发展的符号结构集合体,在任一瞬间它都包含一个符号结构的集合体,除此之外,它还包含一个由按照一些表达式运作,以产生出另一些表达式的过程,如创造过程、修正过程、再生过程和破坏过程等。物理符号系统的基本任务和功能就是对符号进行操作,即辨认相同的符号和区分不同的符号。为此,这个系统就必须能够辨别不同的符号之间的物理差别,比如辨别光波和声波的差别等。

纽威尔对物理符号系统定义了10种操作符:<sup>30</sup>

1、赋值符号(assign a symbol): 建立符号与项目之间的基本关系。对项赋值,称之为存取。存取一个操作符意味着存取它的输入、输出和唤配机制。

2、复制表达式(copy expression): 将表达式和符号加到系统里,新的表达式是输入表达式准确的复制,即在各种作用中具有完全相同的类型和符号。

3、写表达式(write an expression): 建立任何规定内容的表达式,它并不建立任何新的表达式,而是修改它的输入表达式。

4、写(write): 在给定的作用下建立一个符号。

5、读(read): 在规定作用下读符号。

6、执行序列(do sequence): 使系统按规定的序列执行任何工作。

7、条件推出和条件继续(exit-if and continue-if): 系统行为有条件地继续执行一个序列,或从中退出。

8、引用符号(quote a symbol): 控制自动地解释被运行的表达式。

9、外部行为(behave externally): 符号系统可控的外部行为之集合。

10、环境输入(input from environment): 利用新建立的表达式将外部环境的输入录入到系统中。

西蒙还以纸上的文字为例,他认为纸上的文字就是一个不完善的物理符号系统,因为它仅仅有存储符号的功能。而一个完善的物理符号系统应该有以下六种

<sup>29</sup> Allen Newell, and Herbert A. Simon. Computer Science as Empirical Inquiry: Symbols and Search. Communications of the ACM, Vol.19, No.3, 1976. 116.

<sup>30</sup> Allen Newell. Physical symbol systems. Cognitive Science4, 1980. 144-145.

Available at <http://www.cogsci.rpi.edu/CSJarchive/1980v04/i02/p0135p0183/Main.pdf>.

功能：<sup>31</sup>

1、输入符号(input):我们拿笔在纸上写字,可以给白纸输入信号;人类可以通过耳朵、眼睛、鼻子等五官感觉和触觉等来输入信号;计算机可以通过键盘、鼠标等输入信号。

2、输出符号(output):纸本身并不能输出符号,但是我们的眼睛可以使它输出,如我们在阅读时,文字符号就可以输出进入眼睛;人类可以通过说话、书写、肢体表达等方式输出信号;计算机可以通过显示器、打印机等终端输出信号。

3、存储符号(store):人类可以通过记忆,将输入符号保存在头脑中;计算机可以通过存储,把模式存进硬盘、光盘等配件。

4、复制符号(copy):人类通过感知外界的刺激,复制成为表征,存储在大脑里;计算机可以复制存储的材料。

5、建立符号结构(build symbol structure):通过找到各种符号之间的关系,在符号系统中形成符号结构。人类通过学习接收信息,然后对符号进行不同的组合,得出新的关,组成新的符号系统,人类可以建立各种知识之间的联系,构成知识系统;计算机可以通过各种符号之间的关系,形成符号结构。

6、条件性迁移(conditional transfer):依赖已掌握的符号而继续完成行为。换句话说,就是以已存储的信息、当前输入信息为根据进行一系列活动。可以用一个很简单的例子来说明人类的条件迁移。给被试一个只有四个指令的程序:“迈开左脚”,“再迈开右脚”,“重复做”,“一直走到屋子的一端就停止”。条件性迁移有个假定:如果满足了某种条件,即如果有了条件A,就去进行活动B;如果没有条件A,就不要进行活动B。在上例中,告诉被试“向前走”,“继续向前走”,如果没有“到屋子的一端就停止”的指令,被试就要撞到墙上。这里,走到屋子的一端就是条件A,有了这个条件就引出条件B,即停止。计算机在已存储的符号基础上,可依据目标来改变符号结构,如计算机辅助教学程序就具有条件性迁移的功能。正因为计算机有这种本领,所以它就获得了很大很大的能量和灵活性,可以完成多种功能。

物理符号系统假设,就是说任何一个系统,如果它能够表现出一般智能(表示与我们所看到的人类行动范围相同的智能:在任一真实情境中,对该系统目的

---

<sup>31</sup> [美]司马贺(Herbert A. Simon). 人类的认知:思维的信息加工理论. 荆其诚、张厚粲译. 北京:科学出版社,1986. 11-12.

来说是恰当的、并与环境要求相适应的行为,会在一定的速率和复杂性的限度之内发生),它就必须执行上述的六种功能。反过来说,任何系统如果具有了上述的六种功能,它就能表现出一般智能。也就是说对于一般智能而言,物理符号系统具有必要的和充分的手段。所谓“必要的”,就是任何表现出智能的系统都可以经分析被证明是一个物理符号系统。所谓“充分的”,就是任何足够大的物理符号系统都可以通过进一步组织而表现出一般智能。

### 2.2.2 附带推论

在2.2.1中,我们对物理符号系统进行了说明,并且根据物理符号系统系统所应有的六种功能,详细的说明人类和计算机(事实上,无论是现代的大型计算机或小型计算机)分别具备物理符号系统这六种功能。而物理符号系统假设认为,任何表现出智能的系统都是物理符号系统,反之,任何物理符号系统系统也都可以表现出智能。因此,我们可以很好的理解以下物理符号系统假设的三个附带推论,或者说附带条件:

第一个推论是,既然人具有智能(人类能够观察、认识外界事物、接受智力测验、通过考试等,这些都是人的智能的表现),那么他就一定是个物理符号系统。西蒙认为,人所以能够表现出智能,就是基于他的信息加工过程。虽然西蒙也承认证据还不是那么明确,因为我们只能从外表来观察,然后推测脑子里面发生了什么。或者说从行为去推到脑子里通过什么程序造成了这种行为。

第二个推论是,既然计算机是一个物理符号系统(西蒙曾经说过他所定义的物理符号系统是计算机科学家们所熟悉的,从属性来看它与一切通用机极为相似。如果采用符号处理语言来定义一台计算机,那么它们的亲缘关系就向同胞一样),它就一定能表现出智能,这是人工智能的基本条件。计算机和人类一样都可以接受信息、处理符号,它的活动与人类的认知过程相似,只要机器能做这些事情,至于能否像人类那样做就是次要的事情了。佩帕特(S. Papert)曾经说:

“人的思维方式显然不同于机器。人是生物。我们如果问机器能否会思维,就等于问我们自己是否愿意将思维这一概念的范围扩大,使它把机器可能做的事情都包括进去。只有在这一点上,我们问机器会思维吗?才有意义。当牛顿说地球受

到太阳的作用，它在力学方面引进了一个新的技术概念。人工智能的情况也一样，它在思维方面同样引进了一个新技术概念。”<sup>32</sup>

第三个推论是，既然人是一个物理符号系统，计算机也是一个物理符号系统，那么我们就能用计算机来模拟人的活动。当然，第三个推论不一定是第一、第二推论推导出来的必然结果。西蒙认为，虽然人和计算机都是物理符号系统并且具有智能，但是它们可以用不同的方式和原理来进行活动，因此计算机并不一定都是在模拟人的活动。计算机可以通过编程来实现复杂计算，人类的思维未必采用这种运算过程，但是计算机程序可以通过按照人类思维操作的过程来编制。

## 2.3 PSSH 成立的证据

西蒙在《作为经验探索的计算机科学：符号和搜索》一文中曾经提到，物理符号系统假设所涉及的领域还没有可以与之抗衡的专门理论出现，然而还是需要为它的成立提供有效证据。虽然要证明物理符号系统假设成立，即证明物理符号系统与智能之间存在联系，我们没有什么办法在纯逻辑的基础上做到，但是事实提供了大量的证据来帮助证明它的有效性。在他曾经这样写道：“现在，认知心理学所做的，也就是试图用物理符号系统假设中的基本规律来解释人类复杂的行为现象。物理符号系统假设提出来已经30年了，大量的经验材料都支持了这个假设和它的三个附带推论。”<sup>33</sup>西蒙将这些事实证据分为两类：一种用来证明理论成立的充分性，这些证据主要来自于人工智能领域；一种来证明理论成立的必要性，这些证据主要来自于认知心理学领域。

### 2.3.1 来自人工智能领域的支持

人工智能领域的证据主要用来证明“任何足够大的物理符号系统都可以通过进一步组织而表现出一般智能”。这些证据来自人工智能领域过往的大量实践和

<sup>32</sup> 章士嵘、王炳文. 当代西方哲学家评传·第二卷 心智哲学. 济南：山东人民出版社，1996. 322.

<sup>33</sup> [美]司马贺(Herbert A. Simon). 人类的认知：思维的信息加工理论. 荆其诚、张厚粲译. 北京：科学出版社，1986. 14-15.



成果，主要是构造人工智能系统方面所取得的成果。

在前面我们曾经提到西蒙说过人工智能的研究深受细菌致病理论的启发，就是确定需要智能的任务域，然后为计算机构造程序以完成任务域的任务。到目前为止，人工智能领域出现了大量的此类程序，这些程序都能在恰当的领域完成某种程度的智能行动，这些程序所体现的智能，从一开始的简单、低级不断向复杂、高级发展。程序所体现的智能的发展，推动了任务域不断的扩大，同时，任务域的不断扩大，也使得研究工作扩展到建立系统。下面我们来举出一些例子：

1956年，IBM公司研究小组进行电脑程序模拟人类学习过程的相关研究，并开发出了具有自学习、自组织、自适应能力的西洋跳棋程序。该程序可以模拟优秀棋手，通过看几步来弃棋，还能学习棋谱，在分析大约17500个不同棋局后，可以猜测出书上所有的推荐的走步，准确率可以达48%。之后通过不断的试验和修改程序，这个程序的智能水平不断地得到提高。1959年，这个程序战胜了该程序的设计者本人，1962年还击败了美国一个州的跳棋大师。与此相同的例子还有很多，包括1997年计算机“深蓝”以二胜一负三平的成绩击败了世界国际象棋冠军卡斯帕罗夫，这在当时引起了很大的轰动。

逻辑理论机LT和通用问题求解程序GPS则为人工智能的研究工作从建立简单程序到建立系统提供了证据。一开始，LT被设计出来是为了建立问题求解理论的形式化模型；后来设计者发现人在解题时的思维过程大致可以归纳为三个阶段（1、先想出大致的解题计划；2、根据记忆中的公理、定理和推理规则组织解题步骤；3、进行方法和目的分析，以便发现问题，进而修正解题计划），而并不使用和这个程序一样的控制过程，因此他们修改了LT，加进了他们了解的人在问题求解时的控制过程——“手段-目标分析”，这样就产生了著名的GPS。这个过程，我们可以简化为：首先建立LT，它是初步的理论和模型；接着，通过LT的改进和研究的发展，提出新的问题，改变任务域；最后，产生新的系统GPS。这类的例子，还包括机遇问题求解、各类专家系统等。

当然，所有着手研究的智能系统远远不止上面罗列的例子，所有这些工作都为证明“任何足够大的物理符号系统都可以通过进一步组织而表现出一般智能”提供了证据。

### 2.3.2 来自认知心理学领域的支持

认知心理学领域证据主要用来证明“任何表现出智能的系统都可以经分析被证明是一个物理符号系统”。这些证据来自于认知心理学领域过去的大量经验材料，主要是建立人类符号行为模型方面所取得的成绩。西蒙曾经写道：“过去的20年中，搜索根据符号系统对人类智能行为做出解释的做法，已在很大程度上取得成功，达到信息加工心理学成为认知心理学中当前的主导观点的脚步。尤其是在问题求解、概念获取和长时记忆领域中，符号处理模型目前居于支配地位”

34

认知心理学领域(主要是信息加工心理学)的证据往往可以分为两部分。第一部分是对人类完成需要智能的任务时的行为进行观察和实验。包括让人类受试者参与解算术难题、获得概念、记忆、向短时记忆存放信息、处理视觉刺激信号和完成采用自然语言的任等实验，对受试者进行观察，并通过观察对受试者的符号加工做出系统假设。第二部分是基于观察到的人类行为的系统假设，通过编制符号系统程序而建立模型，与人工智能中的对应活动十分相似。这方面的例子就包括我们上面提到的GPS，GPS就是在实验得出人类解题思维过程系统假设的思想基础上，修改LT程序而产生的。当然这样的例子还很多，包括记忆系统、EPAM系统等。这些例子都能很好的证明“任何表现出智能的系统都可以经分析被证明是一个物理符号系统”。

---

<sup>34</sup> Allen Newell, and Herbert A. Simon. Computer Science as Empirical Inquiry: Symbols and Search. Communications of the ACM, Vol.19, No.3, 1976. 119.

## 第3章 PSSH 主要思想来源和形成过程

在第1章,我们谈到了在人类发展历史中与物理符号系统假设相关联的一些人类思想和实践,应该说物理符号系统假设是建立在众多的人类思想和实践之上的。在本章,我们重点来探讨物理符号系统假设的主要思想来源,以及物理符号系统假设的形成所经历的主要过程。

### 3.1 主要思想来源

物理符号系统假设的形成是建立在众多思想基础之上的,它的主要思想来源是亚里士多德、弗雷格、图灵的相关思想。而纽威尔和西蒙在他们的论文《作为经验探索的计算机科学:符号和搜索》中,也曾提到了弗雷格、图灵等人的思想对物理符号系统假设形成的重要意义。

亚里士多德是逻辑的创始人,他编著了《工具论》。该书主要收集了《范畴篇》、《解释篇》、《前分析篇》、《后分析篇》、《论辩篇》、《辩谬篇》等六部名著,集中体现了其逻辑思想。亚里士多德认为“推理是一种论证,其中有些被设定为前提,另外的判断则必然地由它们发生。”<sup>35</sup>他在《解释篇》中提出了一个重要的观点,他认为口语是心灵的经验的符号,书面语是口语的符号。在《前分析篇》第二章他首次引进了变项,将A、B、C等符号作为词项变元代入具体内容。他在《工具论》中提出了一种非形式化的三段论逻辑系统和一些重要的逻辑规律,允许在初始前提的条件下机械地推导出结论。一个三段论由三个命题组成,两个是前提,一个是结论。其推理形式可公式化为: M 是 P, S 是 M, S 是 P。他指出“如果一个三段论除了所说的东西以外不需要其他什么就可明确得出必然的结论,那么,我们就称这个三段论是完满的;如果一个三段论需要一个或多个尽管可以必然从已设定的词项中推出但却不包含在前提中的因素,那么我们就称这个三段论是不完满的。”<sup>36</sup>亚里士多德认为:“三段论是一种论证,其中只要确定某些论断,

<sup>35</sup>亚里士多德. 亚里士多德全集(第一卷). 苗力田主编. 北京:中国人民大学出版社,1997. 353.

<sup>36</sup>王路. 逻辑的观念. 北京:商务出版社,2000. 38.

某些异于它们的事物便可以必然地从如此确定的论断中推出。所谓‘如此确定的论断’，我的意思是指结论通过它们而得出的东西，就是说，不需要其他任何词项就可以得出必然的结论”<sup>37</sup>虽然所使用的主要表述工具仍为自然语言，但在一些论述中力求过渡到使用符号，以代表自然语言。正如海森堡所说：“亚里士多德在他的逻辑中分析了语言形式，分析了与它们的内容无关的判断和推理的形式结构。”<sup>38</sup>在亚里士多德的逻辑思想中包含了“思维可计算”的思想萌芽和“符号”的概念，并允许在初始前提的条件下机械地推导出结论，这些思想是物理符号系统假设形成的最根本的基础。

弗雷格思想的基础来源 1879 年出版的《概念文字：一种模仿算术语言构造的纯思想的形式语言》。他最初想证明，逻辑包括算术以及可以规划为算术的数学分支。但是他在研究逻辑和算术的关系时意识到，最重要的是要保证推理的过程完美无缺，而他发现语言的缺陷和不完善是实现这种严格要求的障碍。他是这样说的，“因为在逻辑问题中，(日常)语言是不可靠的。指出语言为思想者设置的陷阱，确实是逻辑学家最重要的任务之一。”<sup>39</sup>他认为，语言不受逻辑规律的支配，遵守语法并不能保证思维活动形式的正确性，“语言中恰恰没有严格确定的推理形式的范围，以致无法将语言形式方面完美无误的进展与省略了中间步骤区别开来。”<sup>40</sup>“为了逻辑研究，为了演绎体系的构造，必须重新构造一种语言，这是一种自觉的有意识的努力。”<sup>41</sup>因此他要修正现行的语言，构造一种新的概念文字来满足他的这种严格要求。用他的话来说即是，一种模仿算术语言构造的纯思维的形式语言。其目的在于通过设立一些形式符号建立一套逻辑运算和推演系统，运用逻辑语言来推出数学。在《概念文字：一种模仿算术语言构造的纯思想的形式语言》这部著作中，弗雷格详细地介绍了他的概念文字，以及用这种形式语言构造的逻辑演算系统。在构造概念文字的过程中，弗雷格借鉴了两种语言，一种是传统逻辑的自然语言，另一种是算术的形式语言。他正式借鉴了数学的形式语言和传统的自然语言的表达方式，用表示逻辑关系的符号补充数学的形式语言，最终创造出他的概念文字。弗雷格打破了传统逻辑的概念、判断、推理的体

<sup>37</sup> 亚里士多德. 工具论(上). 余纪元译. 北京: 中国人民大学出版社, 2003. 85.

<sup>38</sup> 解丽. 符号学与逻辑学的关系. 中共郑州市委党校学报, 第一期, 2006. 151.

<sup>39</sup> [德]弗雷格(G. Frege). 弗雷格哲学论著选辑. 王路译. 北京: 商务印书馆, 1994. 148.

<sup>40</sup> 王路. 世纪转折处的哲学巨匠: 弗雷格. 北京社会科学文献出版社, 1998. 28.

<sup>41</sup> 王路. 世纪转折处的哲学巨匠: 弗雷格. 北京社会科学文献出版社, 1998. 28.

系结构，建立了现代逻辑的体系结构。它标志着现代逻辑的开端，为逻辑提供了一种可以精确表述推理形式的形式语言，这是一种用关系符号补充数学形式语言而构造的逻辑形式语言，它为人们提供了一套构造形式语言的方法，从而使逻辑走上了形式化的道路。正像弗雷格所说的：“在我看来，这种概念文字的发现本身就促进了逻辑的发展。”<sup>42</sup>此外，弗雷格的这种人工语言的建立为计算机准确理解和表达人的思想，成功地刻画和模拟人类的思维过程解除了障碍。而这个问题是物理符号系统假设形成必须解决的。

物理符号系统假设的最为主要的思想来源是图灵的初创性工作。1936年，图灵在《论应用于解决问题的可计算数学》中用严格的数学表达给出算法概念，其中首次提出符号处理的概念，并由此转向了“可改变的编码程序”。1937年，在《理想计算机》中提出了著名的图灵机——理想计算机模型，证明了只有图灵机解决的计算问题，实际计算机才能解决；如果某计算问题图灵机不能解决，则实际计算机也无法解决。1950年10月，图灵发表了著名论文《计算机与智能》(Computing machinery and intelligence)，当中提出了回答“机器能够思维吗”这类问题的测试方法——“模仿游戏”，即著名的图灵测试(turing test)：<sup>43</sup>图灵测试由三个人来做，一个男人(A)，一个女人(B)，还有一个提问者(C)，性别不限。提问者C呆在一间与另两人分开的独立房间里。参加游戏的三方各自有不同的目标任务。提问者C在游戏中的目标是，通过电传打字机与另外两位测试者进行文字问答，从而判断出哪一位是男性哪一位是女性。提问者可以用标号X和Y称呼他们，在游戏结束时，他可能说“X是A，Y是B”，也可能说“X是B，Y是A”。提问者可以提出这样的问题：“X，请告诉我你的头发长度，可以吗？”男性测试者A的目标是努力使得提问者作出错误的性别判断，即让提问者C以为他是女性，而女性测试者B的目标则是尽力帮助提问者作出正确的判断。图灵提出新的问题是：如果在游戏中用一台机器代替男性测试者A的位置，那会怎样？提问者错误判断的次数是否还会和原先一样多？图灵认为这些问题可以用来替代原先“机器会思维吗”这个不太明确的问题，同时认为新问题有利于区分人的身体能力和智力能力。如果机器代替男性测试者之后，提问者判断的准确率并无提高，那么机器就被认为通过了测试，并且说明该机器具有智能。由此原先“机

<sup>42</sup> [德]弗雷格(G. Frege). 弗雷格哲学论著选辑. 王路译. 北京: 商务印书馆, 1994. 4.

<sup>43</sup> A.Turing. Computing machinery and intelligence. Mind, 1950(59). 433-434.

器会思维吗”转为“是否存在能够在模仿游戏中表现出色的机器”这个问题。图灵希望撇开哲学争论，提出一个使得“机器会思维吗”这个问题的正反双方都能够接受的测试，特别鼓励正方建造出可以赢得“模仿游戏”而通过测试的机器。另外，图灵还在文中陈述了他的论题：有效的程序是一般智能必要而充分的条件。<sup>44</sup>图灵关于计算机和智能以及脑和计算机在普适图灵机中的抽象数学表征的理论统一，是物理符号系统假设的思想先驱。

### 3.2 形成过程

纽威尔和西蒙曾在《作为经验探索的计算机科学：符号和搜索》一文中提到，物理符号系统假设的建立经历了“形式逻辑”、“图灵机与数字计算机”、“存储程序概念”、“表处理”、“LISP 语言”这五个阶段。下面我们来详细描述这五个阶段：

1、“形式逻辑”阶段，纽威尔和西蒙也将它称作“形式符号处理阶段”。由于传统逻辑存在一些自身无法克服的缺陷，从近代开始，一些逻辑学家试图建立一门新的逻辑科学。莱布尼茨曾提出建立“普遍的符号语言”和“一般数学”的逻辑思想，建立一种代替自然语言的“普遍的符号语言”，并将推理的一般规则换作计算的规则，以修正传统逻辑的用自然语言表述的不精确性这一缺陷。他的“普遍的符号语言”即一种符号系统，一种通行于全人类的人工语言。莱布尼茨设想在这种语言中，初始概念是照着代数学的样子用数目或字母符号表达的，而复杂的概念应该被分解并表述为初始概念组合，命题则表述成等式。他还试图在这种普遍语言的基础上创立一种逻辑代数即他所讲的“推理演算”，推理就是根据这个演算的规则来计算，这样就能把人类知识化归为一个记号和记号的运算系统。他的构想揭开了形式逻辑的发展序幕。之后，以弗雷格、怀特海、罗素为代表的逻辑学家们付出了艰辛的努力，就形式逻辑提出了方案，使得莱布尼茨的设想在特定的意义上得以实现。他们的研究成果，使得以逻辑方式获取基本的概念式数学观念，把证明和演绎观念置于可靠的根基上。纽威尔和西蒙认为“形式

<sup>44</sup> 熊哲宏. 关于符号处理范式在认知科学中的地位和前景. 华中师范大学学报(人文社会科学版), 第 38 卷, 第 4 期, 1999. 59.

逻辑”阶段从弗雷格、怀特海、罗素就形式化逻辑提出方案开始。这种努力在数理逻辑中达到了顶点，这就是我们熟悉的命题逻辑，一阶和高阶逻辑。形式逻辑的最大特征就是人工语言的使用。所谓人工语言，是相对于自然语言而言的，指的是不同于自然语言的人造符号。形式逻辑的命题及其组成命题的词汇，均用人工规定的、单义的人造符号表示，整个推理实际上也就是符号之间的演算。在符号化的基础上，形式逻辑赋予人工符号系统的语形、语义解释、形成语形系统、语义系统并存的形式演算系统，并从公理化的角度要求系统具有无矛盾性、独立性、一致性等。形式逻辑为后来图灵机和数字计算机的发展提供了基础。纽威尔和西蒙曾经这样写道：“我们具有的是一个机械系统，尽管是非强制性的(现在我们也称为非决定论的)，而有关这一系统的种种事情都可以加以证明。这样的成功首先是通过一步步远离所有原来看来与意义和人类的符号相关的东西而取得的。”<sup>45</sup>

2、“图灵机和数字计算机”阶段，纽威尔和西蒙也将它称作“自动形式符号处理”阶段。1936年，图灵在论文《论应用于解决问题的可计算数学》中给出算法概念严格的数学表达，首次产生了符号处理的概念，并开始转向了“可改变的编码程序”。1937年，他在论文《理想计算机》中提出了一个独特的设想，即著名的图灵机——理想计算机模型，在论文中图灵对计算理论作了精辟的论述。图灵机由一个有限状态控制器和一根假设无限长的纸带组成，纸带上有数据，就是我们所共知的0和1。纸带起着存储器的作用，控制器可以在带上左右移动，控制带有一个读写头，读写头可以读出控制器访问的格子上的符号。工作带上印有一个个的方格，方格里可以印一个符号，也可以是空白。一个读写头“盯”任一个方格。读写头可以根据当时状态和看到的方格内的符号，采取下述三种行动之一：左移一格，右移一格，或者印一个符号(也可以印空白，把原有符号抹掉)。显然，读写头可以根据它所处的状态和所看到的符号依次修改纸带上各个方格的符号。这也就是符号逻辑推理了。当把方格内的符号视为数时，图灵机扫视整个纸带的过程，就是由纸带上的原始数据求解新的处理结果(仍在纸带上)，这也就是数值计算了。简单来说，图灵机可以在纸带上完成“读”、“写”和“扫描”操作。“读”操作作为控制状态提供条件分支，而控制状态则表现为读头下数据的函

<sup>45</sup> Allen Newell, and Herbert A. Simon. Computer Science as Empirical Inquiry: Symbols and Search. Communications of the ACM, Vol.19, No.3, 1976. 117.

数。图灵机可以处理各种由逻辑符号组成的东西，例如用文字书写的文件、数学表达式等。这种机器可以用通用逻辑运算器组装成专用的、算术运算所需要的数值运算器(numerical operator)。也就是说，这种图灵机主要是一种逻辑符号处理机，也可以做数值计算。这个理论模型具有重要的实际意义。图灵证明，只有图灵机能解决的计算问题，实际计算机才能解决；如果某计算问题图灵机不能解决，则实际计算机也无法解决。纽威尔和西蒙在回顾这个阶段是这样评价：“在这一阶段上所完成的只是解释原理的一半——证明机器可以根据说明运转。因而，这是自动形式符号处理的阶段。”<sup>46</sup>

3、“存储程序概念”阶段。1945年，匈牙利数学家、博弈论的创立者冯·诺依曼(John. Von. Neumann)在《关于离散变量自动电子计算机的草案》一文中提出了存储程序的概念，并提出了相关的电子计算机模型——“冯·诺依曼机”。从此，计算机程序作为数据的同时，也可以作为数据来运作。存储程序概念的出现，使得计算机可以对系统拥有的数据作出解释得以实现。这样它在图灵机的基础上，完成了物理符号系统假设理论中的解释原理的另一半。

4、“表处理”阶段。1956年，纽威尔和西蒙首次提出并成功应用了链表(list)作为基本的数据结构，完成了编制人工智能程序的表处理技术。在计算机科学的发展中，表处理同时表现为三件事情：(1)它是计算机真正的动态存储结构的创建，改变了之前计算机只有固定结构的情况。它在替代和改变内容操作之外，增加了建立和修正结构的操作。(2)它证明了“计算机是由一组数据类型和一组对这些数据类型来说是恰当的操作组成的”这一基本抽象方式。使得计算机系统能够不受硬件限制，运行任何一种对应用来说是恰当的数据类型。(3)表处理产生了一个指称模型定义符号处理。表处理概念的出现使得物理符号系统假设中的指称原理得以实现。纽威尔和西蒙曾这样评价：“以抽象方式出现的表处理概念的形成开创了一个新天地，在这里，指称和动态符号结构定义出许多特征。”<sup>47</sup>

5、“LISP语言”阶段。随着表处理方面不断发展的经验累积，1959年麦卡锡的MIT小组专为符号表示方法发明了一种表处理语言LISP，成为了人工智能程序设计的主要语言。LISP语言完成了对动作的抽象，是将表结构从它们具体

<sup>46</sup> Allen Newell, and Herbert A. Simon. Computer Science as Empirical Inquiry: Symbols and Search. Communications of the ACM, Vol.19, No.3, 1976. 117.

<sup>47</sup> Allen Newell, and Herbert A. Simon. Computer Science as Empirical Inquiry: Symbols and Search. Communications of the ACM, Vol.19, No.3, 1976. 117.



机器内的潜入状态中取出，创建了一个新的带有 S 表达式的形式系统。

纽威尔和西蒙认为，上述的每个阶段的出现和发展，都是物理符号系统假设形成的不可或缺的部分。这些阶段的成果为物理符号系统假设的形成提供了宝贵的实践借鉴。

## 第 4 章 影响及其局限

纽威尔和西蒙的数十年的创造性研究为物理符号系统假设提供了有效的证据和强大的支持，并为符号系统范式提供了启示的范例。他们的努力和成功使得该理论成为认知科学的一个重要研究纲领。在 1958 年至 20 世纪 80 年代中期这一漫长的时期内，符号系统范式在认知科学的研究中一枝独秀，在实践中的进展也是一路高歌，可以说当时人工智能领域和认知心理学大多数研究主要是在符号系统范式的指导下开展的。

1982 年夏天，日本开展了雄心勃勃的“第五代计算机”计划。所谓“第五代计算机”就是把信息采集、存储、处理、通信同人工智能结合在一起的智能计算机系统。它能进行数值计算或处理一般的信息，主要能面向知识处理，具有形式化推理、联想、学习和解释的能力，能够帮助人们进行判断、决策、开拓未知领域和获得新的知识，人机之间可以直接通过自然语言（声音、文字）或图形图象交换信息。<sup>48</sup>日本通产省组织富士通、NEC、日立、东芝、松下等知名企业配合“新一代计算机技术研究所”（ICOT）共同开发，该项计划总投资预算达到 1000 亿日元，设计分三个阶段实施，10 年完成。该计划宣称以 Prolog 为机器的语言，其应用程序将达到知识表达级，具有听觉、视觉甚至味觉功能，能够听懂人说话，自己也能说话，能认识不同的物体，看懂图形和文字。人们不再需要为它编写程序指令，只需要口述命令，它自动推理并完成工作任务。据《日本经济新闻》报道，五代机计划最终目标是组装 1000 台要素信息处理器来实现并行处理，解题和推理速度达到每秒 10 亿次；与此相连接的是容量高达 10 亿信息组的数据库和知识库，包括 1 万个日语和外国语言的基本符号，以及语法规则 2000 条，可以分析 95% 以上的文章，自然语言识别率达到 95%。此外，还将配置语音识别装置和储存 10 万个图象的模式识别装置等等。1992 年，因最终没能突破关键性的技术难题，无法实现自然语言人机对话、程序自动生成等目标，导致了该计划最后阶段研究的流产。“第五代计算机”计划是符号系统范式指引下的产物，该计划的流产，揭示了物理符号系统假设其自身存在的一些问题，物理符号系统假设开始受到越来越多的攻击，德雷福斯甚至断言，物理符号系统假设将在

<sup>48</sup> <http://baike.baidu.com/view/907756.htm>

十年内消失。然而，德雷福斯的预言并未得到验证，物理符号系统假设本身仍然具有很多合理性和存在的价值，时至今日它仍然在许多方面指导人们展开相关研究。

## 4.1 主要应用

在纽威尔和西蒙看来心理学和人工智能之间的关系是相互关联、极其密切的，因为，一方面通过为人类智能建构模型启发指引了人工智能领域的发展，而另一方面人工智能系统的建立又反过来促进了心理学领域的进步。物理符号系统假设主张人脑是一种信息处理系统，人类智能就是由相互关联的符号结构来实现的符号系统。人类可以通过耳朵、眼睛、鼻子等五官感觉和触觉从外部接收编码的符号，并通过说话、书写、肢体表达等方式输出信号。通过复制和重组记忆符号、接收和输出符号、因同一和差异而比较符号结构，人类最终实现他们的思维。物理符号系统假设第三推论主张计算机可以通过编程来模拟人类的智能，物理符号系统假设支持者的研究工作也主要集中在建立和发展各种人类智能模型，通过计算机编程的方法来实现对这些智能过程的模拟，最后通过比较计算机的行为和人类的行为来验证理论的正确性。从 1958 年开始，物理符号系统假设主要被应用在各类模拟人类思维过程的人工智能系统构造方面，20 世纪 60 年代主要以问题求解的研究为主，70 年代以专家系统的研究为主，到了 80 年代自然语言理解等研究为主。探讨这些程序的本身，不是本文的写作目的，因此在这里我主要简单地谈一下物理符号系统假设在这些方面的应用以及这些方面的一些比较成功的例子。

物理符号系统假设首先被应用在问题求解方面，主要表现在相关学者设计了大量的程序来完成定理的证明或反证这一人类智能模拟任务(也包括将许多非形式的工作形式化，从而来完成证明或反正)，这要求这些程序不仅需要有能力根据假设进行演绎的能力，而且需要某些直觉技巧。其中最为著名的就是之前我们提到的 LT 和 GPS。纽威尔、西蒙和肖设计 LT 程序时，首先给该程序输入一些逻辑表达式，这些表达式是公理和已经证明的定理，然后给处另一些表达式，要求它根据逻辑规则和已有的公式或定理来证明这些表达式。LT 通过尝试各种方法，

去努力实现这些问题的证明,如果一种方法失败,它就继续下一种方法,如果一种方法获得了成功,它就将证明过程打印出来。最终,该程序成功地证明了《数学原理》一书中的大部分定理。LT 的成功具有极其重要的意义。首先,它成功地模拟了人类在解决该类型复杂问题时的思维活动,成为世界上第一个获得成功的智能系统,坚定了物理符号系统假设支持者的信念;其次,它的成功在一定程度上证明了 LT 的作业和人类问题解决的作业是一致的,从而为计算机模拟人类智能提供了一种重要方法,并成为计算机模拟人类智能的一个经典的范例。LT 之后,一系列类似的模拟人类智能行为的计算机程序涌现出来,GPS 就是其中较有影响力的一个。GPS 是纽威尔和西蒙在 LT 基础上建立起来的一个功能更为强大的人工智能系统,它包含了一个具有任务领域知识的子系统,并建立了一个与任务领域无关的通用问题解决机制,采用“如果……那么……”的启发式程序来模拟人类的问题解决行为,能够使用少量的加工过程来解决不定积分、三角函数等十多种类型的问题。GPS 创造性地对问题解决中的通用机制和与任务有关的机制进行了区分,并提出了“手段——目的”的分析方法作为描述人类问题解决行为的主要机制,这使得它作为一种有效而成熟的问题解决模型在多种其他问题解决程序中被广泛应用。

鉴于人脑的复杂性,实现计算机模拟人类智能的这一目的,需要首先来实现人类智能的一部分,通过某一特定领域的知识大富集来实现模拟人类专家在特定领域的行为也成了大量学者的一项重要工作。除了问题求解之外,物理符号系统假设在专家系统方面也有广泛的应用。应该说专家系统是问题求解程序的一个发展。所谓“专家”,一般都拥有某一特定领域的大量知识,以及丰富的经验。在解决问题时,专家们通常拥有一套独特的思维方式,能较圆满地解决一类困难问题,或向用户提出一些建设性的建议等。专家系统简单地讲,就是一个具有智能特点的计算机程序,它的智能化主要表现为能够在特定的领域内模仿人类专家思维来求解复杂问题。因此,专家系统必须包含领域专家的大量知识,拥有类似人类专家思维的推理能力,并能用这些知识来解决实际问题。例如,一个医学专家系统就能够像真正的专家一样,诊断病人的疾病,判别出病情的严重性,并给出相应的处方和治疗建议等。专家系统通常包含一个由特定领域的系统知识与专家经验组成的大型数据库,根据特定领域一个或多个人类专家问题解决方法进行推

理和判断,模拟人类专家的决策过程,以解决该领域那些需要专家决定的复杂问题。专家系统的任务范围往往局限在特定的领域之内,解决的问题一般包括解释、预测、诊断、设计、规划、监视、修理、指导和控制等。专家系统和之前问题解决程序的最大的区别是,它所需要解决的问题通常没有特定的算法解,需要计算机程序在不完全、不精确或不确定的信息基础上作出结论。目前专家系统在矿物勘测、化学分析、规划和医学诊断方面等方面取得了巨大的成功,甚至达到了人类专家的水平。著名的例子包括 PROSPECTOR、DENDRL 和 MYCIN。其中 PROSPECTOR 主要应用在矿物勘测领域,它成功的发现了一个价值超过 1 亿美元的钼矿沉积; DENDRL 可以被很好的应用于化学结构分析方面; MYCIN 则可以对血液传染病的诊断治疗方案给出意见,特别是在细菌血液病和脑膜炎的诊断方面, MYCIN 取得了巨大的成功,甚至被认为已经超过了这方面的专家。

自然语言理解是物理符号系统假设应用于实际领域的又一重要应用。要模拟人类使用语言来进行信息的传递和处理,就需要建立一个能够生成和理解自然语言的人工智能系统,这涉及到极其复杂的编码和解码问题。一个能理解自然语言信息的计算机系统,看起来就像一个人一样需要有上下文知识以及根据这些上下文知识和信息,用信息发生器进行推理的过程。目前语言处理研究的主要课题是:在翻译句子时,以主题和对话情况为基础,注意大量的一般常识——世界知识和期望作用的重要性。经过多年艰苦努力,这一领域已获得了一定的成果,出现了大量应用型软件,大家所熟知的“东方快车”就是汉化翻译软件的典型代表。它首先提出了“智能汉化”的概念,在一定程度上实现了模拟人类汉化翻译的功能,能够起到很好的辅助翻译的功能。

## 4.2 自身存在的问题

物理符号系统假设的实质性思想,就是把人看成一个“信息处理系统”,或者叫做“符号操作系统”,并认为计算机可以模拟人的活动。以纽威尔和西蒙为代表的一批物理符号系统假设的支持者,为了实现计算机模拟人类智能进行了很多研究和实践,上一节中我们曾经提到了物理符号系统假设在众多方面的广泛应用,以及取得的一些成果,这些方面的成功极大的鼓舞了该理论的支持者和追随

者，以至于在上个世纪60年代和70年代他们过分乐观的来看待他们所面对的众多困难。纽威尔和西蒙曾断言，由计算机操作的二进制数串能够表达包括现实世界的任何东西，人的大脑、心灵和计算机都是物理符号系统，那么就可以用计算机去模拟人的行为，完全可以在形式系统中通过用规则操作符号演算来生成智能。众多的物理符号系统假设的支持者认为，目前的该领域的研究方向和研究方法都已经完全确，他们所需要做的只要付出更多的努力和足够的耐心。他们甚至预言在未来的10年或者20年内计算机能够完全模拟人类智能。然而实践的结果却很难让人满意。随着研究的一步步深入，实现这一预想的问题变得越来越多，并且困难也越来越严重。日本雄心勃勃的“第五代计算机”研究计划的失败之后，物理符号系统假设主宰这一领域的时候宣告结束。人们开始审视物理符号系统假设本身，并重新思考实现机器模拟人类智能的可能行以及存在的困难、障碍。

通过分析之前物理符号系统假设支持者的众多工作，我们可以发现他们在解决人类智能模拟问题时，有一个很重要的指导方法：那就是先要把模拟的问题形式化；在将问题形式化之后，设法找出一个可以实现的算法。尽管这一方法具有非常重要的作用，但是它也具有很大的局限性。首先人脑是极其复杂的，它能够解决的问题可以说无所不包，虽然其中的某些问题可以被形式化，但是很多问题往往无法被形式化，比如常识问题、人类语言、人类情感等；其次，即使是已经被形式化的问题，要找出一个可以实现的算法也是极其困难的，比如组合爆炸难题。正如鲍亨斯基(J. M. Bochenski)所说的：“要牢牢记住形式系统总是抽象的，决不可把它与实在划等号。因此，决不应该把形式化方法当作惟一的方法，相反应结合其他方法一起使用形式化方法。”<sup>49</sup>下面我们来结合实例说明作为物理符号系统假设在具体应用中存在的问题。

#### 4.2.1 常识知识问题

所谓常识或者常识知识，就是指人类所掌握的日常的、操作性的基本技能、背景知识和直觉，它由练习、技能、能力、自然推理和直觉知识构成。人类往往

---

<sup>49</sup> [德]鲍亨斯基(J.M. Bochenski). 当代思维方法. 童世骏、邵春林、李福安译. 上海: 上海人民出版社, 1987. 45.

通过运用常识知识来处理日常生活事务。很难想象人类离开了常识知识，整个人类世界会是什么样子。“常识知识问题”(commonsense knowledge problem)是20世纪70年代由人工智能研究所提出的一个问题，具有极其深远的认知科学含义。如果不能有效解决常识知识问题，也就无法构造像人一样的智能系统或者像麦卡锡所宣称的“像人一样有智慧的机器”，那么物理符号系统假设所主张的——计算机可以模拟人的思维活动——最终也将无法得到证明。常识知识是如此的普遍，然而又是如此的复杂，复杂到我们很难描述它，甚至至今都无法对它做出解释。人类通过在物质世界中不断的训练来获得常识知识，而计算机要获得这些常识知识，就需要将通过形式化的方法人类的全部知识和信念得到有效表达，并将这些信息有效地组织起来以便能够灵活运用。然而要对常识执行进行有效地表达和组织运用是一项非常艰巨的任务，因此常识知识问题也成了目前符号系统范式研究纲领中最难以跨越的重大鸿沟之一。

很多学者在试图解决这一难题时进行了很多的努力和尝试。勒内特曾经主持设计了迄今规模最大的人工智能试验——CYC系统，CYC包含百科全书的意思，取自英文单词“encyclopaedia”的第3到5个字母“cyc”。他计划在计算机程序中输入一个见多识广的北美成年人所具有的全部知识，以此为计算机建立一个背景知识库以解决常识知识问题。然而这个工作量有多大呢？用勒内特自己的话来说要完成全部知识的输入“大概需要100万帧知识，10年时间。”<sup>50</sup>由此可见，要将所有这些知识全部输入程序，本身就困难重重。况且，一般的常识往往不包含在这些知识中，常识知识最为重要的部分恰恰不是知识本身，它是人类在成长中不断训练以获得的一种技巧，它通常隐藏在知识的联系与结构之中，常识知识本身无法被明确的陈述出来。而且要解决常识知识问题还需要考虑时间、空间、因果关系等诸多问题，因此最终CYC没有获得成果。

按照维特根斯坦当年的看法，世界的形式化原则上是可以做到的。物理符号系统假设的拥护者们也都认为，只要符号足够丰富，操纵这些符号的规则足够细致和严密，建立起一个能够为计算机识别和处理的形式化系统是可能的。然而到目前为止的事实却告诉我们，符号系统范式在常识知识问题上所处的困境，从根本点上来说与人类语言、人类情感这些方面所面对的尴尬是相同的，其根源和最

<sup>50</sup> [美]丹尼斯·萨沙(Dennis E. Shasha)、拉泽(Cath A. Lazere). 大师的智慧: 15位杰出电脑科学家的生平与发现. 刘军译. 保定: 河北大学出版社, 1998. 334.

为主要的的原因是，在面对这些问题时候我们难以将它形式化。正如17世纪莱布尼茨所说的，要将人类的有关技能进行观察和转化是极其困难的。威伦斯基甚至认为“没有人能指出常识问题是什么”。我们之前曾经提到，在符号系统范式研究纲领中，要有效解决一个问题，首先就需要对其进行形式化，然而正是因为常识知识的广泛存在和过于复杂，以至于人们始终无法对它进行有效的形式化，更不要说在通过形式化之后找出可实现的算法。要将常识知识问题形式化，那就涉及到提取问题，所以即使我们能够将所有这样的知识都输入进程序，我们还是不知道如何去提取它。因此，我认为常识问题在符号系统范式纲领的指导下难以得到有效的解决，正如德雷福斯所断言的，常识问题将引起人工智能的崩溃。

#### 4.2.2 组合爆炸难题

所谓“组合爆炸”，最初出现在一些指数函数中，当该函数自变量发生较小幅度的变化就会引起因变量的大幅度变化，特别在解决组合数学(combinatorics)问题中呈现爆炸性的增长，使得许多问题无解。在计算机解决问题方面，组合爆炸主要表现为，随着问题规模的不断增大，计算信息量出现爆炸性增长，填满所有的存储单元，直到溢出，同时运算频数和花费的时间达到计算机无法处理的程度，任何一个表示无穷大“ $\infty$ ”的运算在计算机上都是不能实现的。

物理符号系统假设支持者在研究实践中，最早面对“组合爆炸”是在通过编程实现棋类游戏的人类智能的过程中。在设计此类程序时，往往走出一步后，会出现多种选择，程序要对这些选择进行分析，并从中选出最佳的走法，然而随着步数的增多，出现的选择会越来越多，在对有些棋类游戏进行编程时，研究者发现随着游戏的进行，选择的可能性会多到计算机无法处理的地步。

在本文前面部分，我们曾经提到了该类程序在国际象棋方面的成功应用，然而当我们追溯国际象棋的产生时，会发现一个非常有趣的故事。据说在古印度的时候，娱乐的东西非常少，当时一个名叫达依尔的宰相为他们的国王舍汉王发明了国际象棋，这个国王非常喜欢，于是决定重赏这位大臣并许诺满足大臣提出的任何一个要求。这个大臣的要求非常简单，他拿出了自己发明的国际象棋棋盘，要求国王在第一个格子中放一颗麦粒，第二格放两颗，第三格放四颗，以此类推，



直到放满六十四格，每个格都是上一格的两倍，只要把以上的麦粒赏给他就行了。国王觉得这个要求很简单，于是欣然答应。结果经过计算，发现要满足这个要求，需要 $2^{64}-1$ 粒小麦，这是一个20位的数字，这么多的小麦能铺满地球表面厚厚的一层，即令举全国之力也无法办到。

从上面的故事中，其实我们就间接的感受到组合爆炸之所以成为一个难题的原因。当问题的因变量出现的爆炸性，导致一个系统无法承受，那么这个系统所面对的问题也就自然无法得到有效的解决。

在编写国际象棋智能程序时之所以取得了成功，很大的原因是因为国际象棋的规则设定，使得国际象棋的计算量得到控制，从而计算机可以进行有效处理。在国际象棋中当出现一步走法之后，一般大约只需要面对35种合理选择，计算机程序可以分析之后的几步棋的情况，以7步棋为例，计算机需要处理500亿到600亿的选择，那么按照目前大型计算机的运算能力，这个问题可以得到有效的解决。然而当程序算面对的对象复杂程度发生改变时，计算机程序就开始无法有效地解决所面对的问题。

以围棋为例，围棋的复杂性远远大于国际象棋，围棋盘上有 $19^2=361$ 个交叉点，每个节点上有白子、黑子、空位这三种可能性。宋朝大科学家沈括在《梦溪笔谈》的“棋局都数”中指出，所有不同的围棋局面的总数是 $3^{361}$ ，这是一个170位的数字，大得不可想象。围棋每走一步之后，大约会出现200种合理选择，这相对于国际象棋的35种来说，复杂性大大的增加。也就是说国际象棋在走七步之后，需要面对处理的选择是35的7次方，而围棋则是200的7次方。每多走一个回合，数字就会出现暴增，走到第七回合，国际象棋只是面对35的14次方选择，而围棋是200的14次方，即便是曾经击败过国际象棋冠军的“深蓝”来面对这个数字进行处理时，也需要花上一年半的时间。这样编写出的智能程序，在实际运用中就已经失去了意义。人类生活中，比围棋复杂的多的问题比比皆是，当符号系统范式所倡导的智能程序所面对的模拟对象的范围扩大之后，组合爆炸难题成了他们最大的困扰。

组合爆炸之所以成为难题，最主要的原因在于，当人们试图去编写程序模拟人类智能时，即使面对的问题已经被形式化，要找出一个可以实现的算法也是极其困难的。

### 4.3 面临的挑战

物理符号系统假设的研究纲领因为自身存在着的一些问题, 在实践过程中遇到了一系列目前难以克服的困难。进入20世纪80年代之后, 物理符号系统假设开始受到塞尔相关学说和联结主义学说的有力挑战。

#### 4.3.1 塞尔的“中文屋”的攻击

应该说纽威尔和西蒙的物理符号系统假设的最为重要的根基是图灵的思想学说。1950年10月图灵在哲学期刊《Mind》上发表的著名论文《计算机与智能》, 被公认为是一篇划时代的著作, 他在该论文中提到的著名的“图灵测试”为物理符号系统假设的成立了扫除了众多障碍。“图灵测试”的提出, 在一定程度上避免了关于“智能”概念的无止境的争论, 为“机器能否思维”找出了一个客观的准则, 这也为其后的物理符号系统假设的支持者们在从事研究工作时, 提供了合法的检验标准。然而随着以物理符号系统假设为核心的符号系统范式研究纲领陷入了困境, “图灵测试”也开始受到了越来越多的攻击, 其中塞尔的“中文屋”思想实验从理论上对“图灵测试”展开了摧毁性的打击。

美国哲学家塞尔在1980年、1984年分别发表了论文《心智、脑与程序》和专著《心、脑与科学》, 在其中塞尔提出了著名的思想实验“中文屋”, 从而引发了一场关于物理符号系统假设理论基础的论证, 这就是历史上非常出名的“中文屋论证”。塞尔对于“中文屋”实验是这样描述的: “假设你被关在一个房间里, 屋内有满满几筐汉语符号, 假如你(与我一样)对汉语一字不识, 但给你一本用英语写的用来处理这些汉字符号的规则, 这些规则是按照汉语的句法而不是语义, 对符号的处理加以纯形式的规定的。比如, 规则中可能有这样的话: ‘从一号筐中取出甲符号, 将它放在从二号筐中取出的乙符号之后。’ 假设这时又有另一些汉字符号被送进屋来, 你根据相应的规则将汉语符号送出屋外。假如你根本就不知道送进屋来的这些符号就是屋外人的‘问题’, 你也不知道你送出屋外的那些符号就是所谓‘问题的答案’。此外假设这些程序设计家们善于涉及程序, 而你又

善于处理符号,这样,你的答案就同一个地道的中国人做出的回答没有什么不同。你被这样关在屋子里,将你的汉字符号移来移去,用送出去的汉字符号回答送进来的汉字符号。按照我描绘的情况,你仅仅是在摆弄这些形式符号,不可能学到哪怕一丁点儿汉语。”<sup>51</sup>在塞尔的“中文屋”实验中,他用中文屋模拟了一个计算机程序,即通过输入一段中文字符,在事先输入的规则指导下,经过计算机运算,输出运算结果。在塞尔精心涉及的这个思想实验中,屋子里的一个根本不懂得汉语的人通过精确而熟练的使用规则以给出一个让人满意的答案,从而让屋外人相信你是一个熟悉汉语的人,但是,实际上他对这些输入的中文符号的意义没有丝毫的理解。之前在图灵看来只要计算机能够通过“图灵测试”,那么计算机所运行的计算机程序过程可以被看做和人类的心理过程是等同的。然而按照塞尔的观点,只要给定了有效的规则,计算机程序不需要理解它所接收的符号的意义,它所需要做的就是根据规则来操作符号并给出回应。

塞尔通过“中文屋”这个思想实验,构造了一个与人类智能没有任何联系但却可以模拟人类智能行为的机器,有力的证明了计算机程序的改变仅仅是规则的改变,而计算机本身根本不可能理解输入符号的意义,更不会具备任何人类所拥有的智能,这对于建立在“图灵测试”基础上的物理符号系统假设是一个非常严重的攻击。“中文屋”思想实验在根本上解除了物理符号系统假设所拥有的合法性,使得人们再一次思考机器“智能”与人类智能之间的本质性差异。

#### 4.3.2 联结主义的竞争

由于以物理符号系统假设为核心的符号系统范式在面对一些难题时,长时间无法给出有效的处理方法,在符号系统范式指导下的研究工作开始渐渐转入低潮,符号系统范式的地位和作用受到了严峻的挑战,<sup>52</sup>此时以网络理论为核心的联结主义开始慢慢进入人们的视野,并成为符号系统范式在认知科学领域的一个强有力的竞争对手。

联结主义又称仿生学派或生理学派。如同塞尔曾经在批判符号系统范式时说

<sup>51</sup> [美]约翰·塞尔(John.Searle). 心、脑与科学. 杨音莱译. 上海:上海译文出版社,1991. 23-24.

<sup>52</sup> 熊哲宏. 认知科学导论. 武汉:华中师范大学出版社,2002. 138.

的：“我们要放弃那种认为在心脑间存在着一种计算机程序的看法，还是把心与心理过程视作如同生长、消化、胆汁分泌一样基本的生物现象。”<sup>53</sup>联结主义主张要从仿生学角度、通过对人脑模型的研究来实现模拟人类认知，认为人类认知的基本元素不是符号而是神经元也就是神经细胞，人类认知过程是通过大量神经元的联结来实现的。联结主义研究纲领的核心概念是“并行分布处理”概念，即认知是从大量单一处理单元的相互作用中产生的，其主要对象为神经网络及神经网络间的联结机制与学习算法，该研究纲领以神经元为硬核，从整体主义思想出发，试图采用人类分布式并行计算的方式让计算机模拟人类自学习、自组织、自适应、联想等方面的行为。

联结主义的先驱是神经生物学家麦卡洛克(W. S. McCulloch)和数理逻辑学家皮茨(Pitts)，他们于1943年在《Bulletin of mathematical biophysics》上合作发表了《神经系统中所蕴涵的思想的逻辑演算》一文，总结了生物神经元的一些基本特征，提出了形式神经元(formal neuron)的数学描述与结构方法，并就形式神经元及其逻辑特征的有关思想进行了分析，随后又提出一种叫做“似脑机器”(mindlike machine)的思想，这种机器可由基于生物神经元特性的互连模型来制造，这就是神经学网络的概念。他们构造了一个表示大脑基本组分的神经元模型，对逻辑操作系统表现出通用性。神经网络最早的模型是M-P模型。在M-P模型中，神经元的性质是简单的，而网络的计算能力却是巨大的。M-P模型开创了用电子装置模仿人脑结构和功能的新途径，它从神经元开始进而研究神经网络模型和脑模型，开辟了人工智能的又一发展道路。

在之后很长的一段时间，由于符号系统范式的在认知科学领域取得的一系列令人瞩目的成果和巨大成功，以及联结主义自身理论的不完善，联结主义始终没有受到重视。1986年鲁梅哈特(D. E. Rumelhart)和麦克里兰(J. McClelland)合著出版了《并行分布式处理：认知的微观探索》两卷本，书中展示了包括物理学、数学、分子生物学、神经科学、心理学和计算机科学等许多相关学科的著名学者在不同方向和领域取得的成果。结合这些成果他们强调人类的认知不是简单的符号运算，而是网络整体结构活动的结果。网络是时间动态系统，由类似于神经元(节点)的大量处理元素(单元)相互联结在一起，每个单元都有其不同的活性，即可

<sup>53</sup> [美]约翰·塞尔(John.Searle). 心、脑与科学. 杨音莱译. 上海: 上海译文出版社, 1991.

以兴奋和抑制其它单元，也可以受其它单元的兴奋和抑制。认知过程就在于网络从初始状态到最后完成的稳定状态，而这个动态过程就可以被看做是认知能力。该书在认知科学领域产生了很大的影响，极大的推动了联结主义的发展，从此联结主义异军突起，声势日渐壮大，联结主义开始成为认知科学领域的另一个重要范式并受到了普遍的承认。

联结主义和符号系统范式对认知有着不同的解释，在理论框架、建模原则、加工方式等许多方面都存在着重要差别。在符号系统范式中采纳的是机能的或软件的描述水平，把人类的智能看作是一个物理符号系统，即符号的操纵系统，是对符号的接收、编码、储存、提取、变换和传递。在符号系统范式追随者看来，人类的认知是通过符号进行的表征的来实现的，符号又可以按照一定的规则进行加工，这一过程也就是符号的操作或计算。符号系统范式强调的是确定符号的形式结构和算法规则。符号系统范式赋予符号以核心性的地位，采纳符号表征和计算的理论，强调的是符号的序列或串行的加工，注重的是符号加工的逻辑基础。而联结主义则将人类的认知看成网络的整体活动。网络是个动态的系统，它由类似于神经元的基本单元或结点所构成。单元相互联结在一起，联结是加权的联结，权值为正的是兴奋性联结，权值为负的是抑制性联结。每个单元都有不同的活性，它既可以兴奋和抑制其他单元，也可以受到其他单元的兴奋和抑制。当网络有一初始的输入，其兴奋和抑制便在其单元之间扩散，直到形成一个稳定的状态。在联结主义看来，网络的信息加工不同于符号的串行加工，而是网络的并行加工。联结主义赋予网络以核心性的地位，采纳分布表征和并行加工的理论，强调的是网络的并行分布加工，注重的是网络加工的数学基础。我们可以通过以下几点对联结主义区别于符号系统范式的特点展开说明：

第一、内在并行性。区别于符号系统范式主张的串行加工的观点，联结主义强调，人工神经网络是高度并行的非线性系统，这一特点体现在其结构上和信息处理的过长中，在人工神经网络中不同的节点可以按不同规则同时运算，而规则之间体现的是竞争或协作关系，也就是说每个单元在特定的时间均有不同的活性状态，接受信息后，同时开始运算，并且相互影响和作用，结果是使某单元的活性状态发生改变，但计算结果只能是整体层次上的。

第二、分布式信息存储。符号系统范式认为信息存储在一个特定的存储区域，

信息存储是通过程序式记忆来实现的。而联结主义则强调分布式记忆，认为信息分布式地存储在整个网络系统中，在该系统中每个节点只包含了概念的一部分，对整体来说有贡献，但不能决定整体状态，存储的整体内容与存储的基本单元不在一个层次上。

第三、容错性。符号系统范式下产生的系统的容错性相对较差，一个细节的出错往往会引起整个程序或系统的崩溃。而联结主义由于采用分布式信息处理的方式，整个系统的行为往往不会因为其中部分单元出错而受到影响。联结主义的容错能力相对较强。

第四、自适应性。联结主义与符号系统范式最大的不同在于，在符号系统范式中主要采用编程来解决问题，而联结主义则通过学习来解决问题。人工神经网络具有学习功能，在学习的基础上，通过不断改变的网络单元之间的联结权重来实现记忆加工。而符号系统范式中所主张的符号系统需要通过外部对其修改程序来完成系统性能的改变。<sup>54</sup>

第五，打开“黑箱”。不同于符号系统范式的采用忽视硬件的“黑箱”研究方法，联结主义认为硬件是极为重要的，强调对硬件的研究和涉及，联结主义试图打开“黑箱”，去研究系统的内部状态，以实现对整个系统的结构和功能的认识。

联结主义在面对一些符号系统范式无法解决的问题时，体现了一定的优势，随着自身理论的不完善，它给出了一种与符号系统范式截然不同研究视角，并提供了一种新的研究方法，在认知领域对符号系统范式发起了全面的挑战。

---

<sup>54</sup> 熊哲宏. 认知科学导论. 武汉: 华中师范大学出版社, 2002. 188-190.

## 结语

物理符号系统假设的形成是建立在众多的人类智慧、思想基础之上的，自诞生以来，就产生了巨大的影响，并在认知科学的众多领域中产生了极大的作用。虽然该假设存在着一些局限性，并受到了以联结主义为代表的众多挑战，但是它仍有其存在的自身价值。正如熊哲宏所说的：“从PSSH的基础假定的合理性、PSSH的解释力以及它与联结主义模型的内在联系等方面来看，符号系统范式还不至于像德雷福斯所断言的那样，将会在10年内消失。”<sup>55</sup>上世纪80年代末，纽威尔还在物理符号系统假设的基础上提出了认知构建(cognitive architecture)理论，SOAR理论(State, Operator, and Result,即状态、操作和结果)，它是模拟人解决问题的计算机程序，能够从经验中学习。纽威尔把它当作认知的统一理论的一个范式，按照认知加工时间尺度的不同，分别讨论了认知的三个方面：即时行为，记忆，学习和技能，以及有意向的理性行为。他在介绍SOAR理论后，讨论了如何把SOAR应用于说明认知的这三个方面的符号加工过程。后来，Schultz讨论了计算主义新方向，认为可以用构建系统中不同部分之间的相互作用机制来说明认知过程。<sup>56</sup>应该说物理符号系统假设，及以它为核心的符号系统范式，还将在很长的一段时间内表现出较强的生命力。

本文第一章从物理符号系统假设的历史渊源入手，通过搜集、阅读和整理大量的背景资料，分三个时期论述了与物理符号系统假设相关联的人类思想和实践。区别于以往的相关研究文献中站在机械、存储单元形成发展的计算机史或机器智能发展史的角度来描述历史，本章主要从人类发展、使用符号代替人类思维的角度按照时间维度来提出历史材料，以给出物理符号系统假设所主张的使用物理符号表征认知活动的充分的哲学思辨理由和科学实证根据。

通过认真研究一手文献资料，和阅读二手文献，本人发现以往的相关研究文献中，对于物理符号系统假设的描述不够全面和详尽，不利于读者的理解。本文第二章主要通过认真研究一手文献资料，对于物理符号系统假设本身进行全面的解读。首先，在第一节中交代了物理符号系统假设创建者之一西蒙对于该理论所

<sup>55</sup> 熊哲宏. 关于符号处理范式在认知科学中的地位和前景. 华中师范大学学报(人文社会科学版), 第38卷, 第4期, 1999. 64.

<sup>56</sup> 唐孝威. 统一框架下的心理学与认知理论. 上海: 上海人民出版社, 2007. 130-131.

处的科学层次的说明和性质界定,这有助于理解物理符号系统假设的合理性和重要性;接着,在本章的第二节中详尽地阐释了纽威尔和西蒙所定义的物理符号系统假设以及由此推出的附带推论,这对于理解该理论的实践应用有重要作用,同时也为第三章中论述的主要思想来源作了铺垫;本章的最后一节主要例举了支持该理论的有效的事实证据,为该理论在提出后迅速获得广泛承认的事实作了合理的解释。

每个理论的提出,其背后必然有它最为主要的思想来源。物理符号系统假设的提出究竟主要受到哪些思想的影响和启发?解决这一问题对于更好地理解该理论具有重要意义,同时在认知科学史上也具有很大的价值。然而通过前期的调研工作,本人发现除了熊哲宏在《关于符号处理范式在认知科学中的地位和前景》一文中曾简略的提到了物理符号系统假设的数学和哲学背景是图灵的初创性工作,目前国内在这一方面缺少相关研究,而在纽威尔和西蒙合著《作为经验探索的计算机科学:符号和搜索》一文以及西蒙的自传中曾出现有相关的一些线索,本人遵循着这些线索进一步阅读相关材料,在本文的第三章的第一节中对于物理符号系统假设的主要思想来源进行了描述和分析。另外,第三章的第二节还根据一手文献对物理符号系统假设形成所历经的五个阶段进行了说明,应该说这几个阶段为物理符号系统假设的形成提供了主要的实践借鉴。

在分析了物理符号系统假设的主要思想来源和实践借鉴之后,本文的第四章主要分析、论述了该理论提出后的主要影响和面对的一些难题。理论的影响主要从两个角度来写,一方面是通过描述该理论在实际中的成功应用来正面体现其影响,另一方面其所遭遇的攻击和挑战也侧面地反应了它所具有的影响。在论述物理符号系统假设所面对的难题时,例举了常识知识问题和组合爆炸难题,来说明该理论在实际应用中面对的模拟问题的形式化和寻找可实现的算法的巨大困难。

本文将物理符号系统假设作为研究对象,以其发展历史作为主线,系统地回顾了物理符号系统假设的产生背景,详细地说明了物理符号系统假设的核心内容和主张,首次尝试分析了其主要思想来源,并整理、描述了其形成过程,最后论述了它的主要应用、局限以及面临的挑战,这有助于更好的理解和总结物理符号系统理论,在认知科学史中具有一定的意义。

由于本人知识结构的局限,在分析其主要思想来源时,肯定不够全面和深入;



同时对于物理符号系统假设的具体应用以及未来的发展,这主要依赖于计算机程序的发展和突破,这方面知识的不足,使得无法深入探讨程序本身内容,只能做些比较浅显的探讨;最后由于认知科学跨越了太多的领域,知识体系极为复杂,在本文中关于物理符号系统假设的影响只做了实践应用方面的正面论述,没有涉及该理论对于其他学术思想和认知理论是否存在影响,然而,这一问题也是极具意义和价值的研究。

以上是本文的总结,希望为以后的研究者提供借鉴。

## 参考文献

1. Allen Newell, and Herbert A. Simon. Computer Science as Empirical Inquiry: Symbols and Search . Communications of the ACM, Vol.19, No.3, 1976.
2. Allen Newell. Physical symbol systems. Cognitive Science4, 1980. Available at <http://www.cogsci.rpi.edu/CSJarchive/1980v04/i02/p0135p0183/Main.pdf>.
3. Allen Newell, and Herbert A. Simon. Human Problem Solving. Englewood Cliffs, New York: Prentice-Hall, 1972.
4. Herbert A. Simon. Models of Thought. New Haven, CT: Yale University Press, 1979.
5. Allen Newell, and Herbert A. Simon. Computer Simulation of Human Thinking. Science, New Series, Vol. 134, No. 3495, 1961.
6. Herbert A. Simon. Computers-Non-Numerical Computation. Proceedings of the National Academy of Sciences of the United States of America, Vol. 77, No. 11, 1980.
7. Herbert A. Simon, and Allen Newell. Heuristic Problem Solving: The Next Advance in Operations Research. Operations Research, Vol. 6, No. 1, 1958.
8. Herbert A. Simon. Invariants of Human Behavior. Annual Review of Psychology, Vol. 41, 1990.
9. Herbert A. Simon. The Future of Information Processing Technology. Management Science, Theory Series, Vol. 14, No. 9, 1968.
10. Herbert A. Simon. The Science of Design: Creating the Artificial. Design Issues, Designing the Immaterial Society, Vol. 4, No. 1/2, 1988.
11. Mie Augier. Models of Herbert A. Simon. Perspectives on Science, vol. 8, no. 4, 2000.
12. Robert M. Harnish. Mind, Brains, Computers: An Historical Introduction to the Foundations of Cognitive Science. Malden: Blackwell Publishers, 2002.
13. Morton Wagman. Cognitive Psychology and Artificial Intelligence: Theory and Research in Cognitive Science. Westport, Conn.: Praeger publishers, 1993.
14. A. Turing. Computing machinery and intelligence. Mind, 1950(59).
15. [美]司马贺(Herbert A. Simon). 人类的认知: 思维的信息加工理论. 荆其诚、

- 张厚粲译. 北京: 科学出版社, 1986.
16. [美]司马贺(Herbert A. Simon). 人工科学: 复杂性面面观. 武夷山译. 上海: 上海科技教育出版社, 2004.
17. [美] 赫尔伯特·A·西蒙(Herbert A. Simon). 我生活的种种模式——西蒙自传. 曹南燕, 秦裕林译. 上海: 东方出版社, 1998.
18. [英]玛格丽特·博登编著. 人工智能哲学. 刘希瑞、王汉琦译. 上海: 上海译文出版社, 2001.
19. [美] E. A. 费根鲍姆(E. A. Feigenbaum), P. R. 科恩(P. R. Cohen)主编. 人工智能手册 (第三卷). 周少柏, 黄汛译. 北京: 科学出版社, 1991.
20. 车文博、朱新明、李亦菲. 架设人与计算机的桥梁: 西蒙的认知与管理心理学. 武汉: 湖北教育出版社, 1999.
21. 休伯特·德雷福斯(Hubert L. Dreyfus). 计算机不能做什么: 人工智能的极限. 宁春岩译. 北京: 三联书店, 1986.
22. [德]鲍亨斯基(J. M. Bochenski). 当代思维方法. 童世骏、邵春林、李福安译. 上海: 上海人民出版社, 1987.
23. [美]丹尼斯·萨沙(Dennis E. Shasha)、拉泽(Cath A. Lazere). 大师的智慧: 15位杰出电脑科学家的生平与发现. 刘军译. 保定: 河北大学出版社, 1998.
24. [美]约翰·塞尔(John. Searle). 心、脑与科学. 杨音莱译. 上海: 上海译文出版社, 1991.
25. [英]霍布斯(Thomas. Hobbes). 利维坦. 黎思复、黎廷弼译. 北京: 商务印书馆, 1986.
26. [法]帕斯卡尔(B. Pascal). 思想录. 何兆武译. 北京: 商务印书馆, 1997.
27. [德]莱布尼兹(G. W. Leibniz). 莱布尼兹自然哲学著作选. 祖庆年译. 北京: 中国社会科学出版社, 1985.
28. [美]黎黑(T. H. Leahey). 心理学史: 心理学思想的主要趋势. 李维译. 杭州: 浙江教育出版社, 1998.
29. [德]卡西尔(E. Cassirer). 人论. 甘阳译. 上海: 上海译文出版社, 1985.
30. 亚里士多德. 亚里士多德全集(第一卷). 苗力田主编. 北京: 中国人民大学出版社, 1997.
31. 亚里士多德. 工具论. 余纪元译. 北京: 中国人民大学出版社, 2003.
32. [德]弗雷格(G. Frege). 弗雷格哲学论著选辑. 王路译. 北京: 商务印书馆,

1994.

33. 武秀波、苗霖、吴丽娟、张辉. 认知科学概论. 北京: 科学出版社, 2007.
34. 熊哲宏. 认知科学导论. 武汉: 华中师范大学出版社, 2002.
35. 林尧瑞、郭木河. 人类智慧与人工智能. 北京: 清华大学出版社, 2001.
36. 唐孝威. 统一框架下的心理学与认知理论. 上海: 上海人民出版社, 2007.
37. 唐孝威等. 脑与心智. 杭州: 浙江大学出版社, 2007.
38. 章士嵘、王炳文. 当代西方哲学家评传·第二卷 心智哲学. 济南: 山东人民出版社, 1996.
39. 王路. 逻辑的观念. 北京: 商务出版社, 2000.
40. 王路. 世纪转折处的哲学巨匠: 弗雷格. 北京社会科学文献出版社, 1998.
41. 熊哲宏. 关于符号处理范式在认知科学中的地位和前景. 华中师范大学学报(人文社会科学版), 第 38 卷, 第 4 期, 1999.
42. 王振友、谢青松. 人工智能学科的发展. 山东工程学院学报, 第 15 卷, 第 4 期, 2001.
43. 解丽. 符号学与逻辑学的关系. 中共郑州市委党校学报, 第一期, 2006.
44. 王黔玲. 论人工智能研究纲领的转变. 社会科学研究, 2004(6).
45. 刘瑞正, 郑延斌, 周振环. 人工智能中的符号机制与联结机制. 设计与研究, 1999(6).
46. 熊哲宏. 认知心理学“经典范式”的系统阐述. 心理科学, 2005, 28(6).

## 后记

两年的时间转瞬即逝，回望这两年的研究生学习生涯，收获很多，这将使我终生受益。

首先，要感谢我的导师黄华新教授。黄老师不仅在学业上给予了我指导，更在治学、为人方面为我树立了榜样，同时他对我生活上的关心使我倍感温暖。在整个论文写作过程中，黄老师给予我的关怀和帮助，我将终生难忘。

其次，要感谢我的导师王淼副教授。从最初的论文选题到最后定稿的整个过程中，王老师倾注了极大的心血和耐心，为我提供了大量有价值的资料，并进行了悉心地指导。

同时，还要感谢许为民教授、李磊副教授、张立老师。三位老师在论文开题、预答辩时给予了许多中肯意见，使我在之后的论文写作过程中能少走弯路，事半功倍。在此谨向三位老师致以诚挚的谢意和崇高的敬意。

最后，要特别感谢罗见今教授。在整个论文写作过程中，罗老师给了我极大的关爱和帮助，提供了极有价值的意见和建议。

师者如父，几位老师渊博的学识、严谨的治学态度、平和的性格、对学生的真切关爱，我都将铭记在心。

周挺

二〇〇八年五月九日于浙大西溪校区