

湖北大学

硕士学位论文

人工智能能否超越人类智能——德雷福斯“联结论”思想研究

姓名：黄严磊

申请学位级别：硕士

专业：科学技术哲学

指导教师：舒红跃

20100501

摘 要

自 1956 年人工智能诞生起，关于人工智能发展趋势的争论就异常激烈。有人认为人工智能仅仅是实现人类智能的工具，只是作为人类智能的附庸和补充在局部上摹仿人类行为，并不能了解人类智慧的实质对人类智能构成挑战，更不可能取代和超越人类智能；但也有人从达尔文的进化论进行类比推断，认为人工智能与人类智能并无实质性差别，人工智能战胜人类智能只是时间问题，人工智能不仅能替代，而且将超过人类智能。

德雷福斯被看作是怀疑论者和悲观主义者的领袖和代表人物。德雷福斯认为，维特根斯坦、海德格尔、梅洛·庞蒂这些哲学家的一些哲学结论已经给人工智能的可能性提出了疑问，他断言人工智能肯定是有限度的。

本文以德雷福斯对人工智能的研究为依据，从人工智能研究的两大分支（传统 AI 和联结论 AI）出发，讨论联结论思想对人工智能的影响；并从生物学假想、心理学假想、认识论假想、本体论假想这四个方面入手证明人工智能的发展确实存在着局限性；最后对德雷福斯联结论思想进行评析，得出人工智能虽能促进人类社会的发展，但并不能超越人类智能的结论。

关键词：人工智能；德雷福斯；传统 AI；联结论 AI；

Abstract

Since 1956, the arguments on developing trend of Artificial Intelligence was sharply. Some people think that Artificial Intelligence which could only local on the imitation of human behavior is just a tool for achieving Human Intelligence, without understanding the essence of human wisdom. Artificial Intelligence is only a vassal of Human Intelligence and supplements, it can't pose a challenge to Human Intelligence, as well as mention to replace and surpass Human Intelligence; But some analogy from Darwin's theory of evolution inferred believe that Artificial Intelligence has no substantial difference between Human Intelligence, Artificial Intelligence beyond Human Intelligence is in the area of time, Artificial Intelligence can will exceed Human Intelligence and so does it.

Dreyfus is thought to be leaders and representatives of the skeptics and pessimists. He believes that Wittgenstein, Heidegger, Merleau • Ponty first of these philosophers, some of the philosophical conclusions have been given the possibility of Artificial Intelligence into question, He asserted that Artificial Intelligence is certainly limited.

In this essay, the writer discuss the impact of connectionism on Artificial Intelligence from the two branches of Artificial Intelligence research (Traditional AI and Associative AI) on the basis of Dreyfus' artificial intelligence research. And the essay will discuss from the following four-pronged approach biological hypothesis, psychology hypothesis, epistemological assumptions, ontological assumption, which prove that the development of Artificial Intelligence dose exist limitation; And thinking through the couplet concluding assessment shows that the impact of Artificial Intelligence on human society; Then the conclusion that Artificial Intelligence can't exceed Human Intelligence.

Key words: Artificial Intelligence; Dreyfus; Traditional AI; Associative AI;

湖北大学学位论文原创性声明和使用授权说明

原创性声明

本人郑重声明：所呈交的论文是本人在导师的指导下独立进行研究所取得的研究成果。除了文中特别加以标注引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写的成果作品。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律后果由本人承担。

论文作者签名：黄平磊

日期：2010年6月1日

学位论文使用授权说明

本学位论文作者完全了解学校有关保留、使用学位论文的规定，即：按照学校要求提交学位论文的印刷本和电子版本；学校有权保留并向国家有关部门或机构送交论文的复印件和电子版，并提供目录检索与阅览服务；学校可以允许采用影印、缩印、数字化或其它复制手段保存学位论文；在不以赢利为目的的前提下，学校可以公开学位论文的部分或全部内容。（保密论文在解密后遵守此规定）

作者签名：黄平磊

日期：2010.6.1

指导教师签名：金红红

日期：2010.6.2

引言

人工智能（AI）出现的背景是：1950 年，英国科学家图灵发表了一篇题为“计算机与智能^①”的文章，第一次提出了“机器能不能思维”这一重要课题，就此拉开了人工智能研究的序幕。作为对图灵这一课题的响应，1956 年夏季，美国一些从事数学、心理学、计算机科学、信息论和神经学研究的年轻学者汇聚在达特茅斯大学，举办了一次长达两个月的学术讨论会，会议讨论用机器模拟人类智能的问题。这次会议第一次使用“人工智能”这一术语，以代表有关机器智能这一研究方向。这是人类历史上第一次人工智能研讨会，标志着人工智能学科的诞生。

经过 50 多年的发展，虽说人工智能在专家系统、机器学习、自然语言理解、人工神经网络、智能决策支持系统等诸多领域取得了一定的成就，但始终有一类问题从人工智能诞生之日起就与哲学紧密联系着，引起了无数人工智能学者的探讨和争论，这就是人工智能的基础问题，即人工智能能否超越人类智能的问题。

（一）人工智能的定义

什么是人工智能？要回答这一问题，首先要明确人工智能究竟是属于科学范畴还是属于技术范畴，或是介乎二者之间的学科。大多数研究人员认为人工智能是介于科学和技术二者之间的一个学科，研究的主要目标是创造有用工具。

由于不同学科背景的学者对人工智能的理解各不相同，人工智能有时被定义为研究怎样制造计算机并为其编程，使其能做心灵所能做的事情。也有不少人偏爱这一更有争议的定义，即把人工智能看作是一般性的智能科学，或更确切地说，看作是认知科学的智力内核。这样，它的目标就是提供一个系统的理论，该理论既可以解释意向性的一般范畴，也可以解释以此为基础的各种不同的心理能力。它不仅包括地球上各种生物的心理，而且还包括全部可能存在的心灵。^②

（二）人工智能的发展

人工智能经过 50 多年的发展大体可分为孕育、形成、知识应用和综合集成四个阶段。

^① 图灵. 计算机与智能[J]. 心灵, 1950.10: 433-460

^② 博登. 人工智能哲学[M]. 刘西瑞 王汉琦译. 上海世纪出版集团, 2005: 1-2

50年代到60年代初是人工智能发展的初级阶段，也叫孕育期。这一时期的研究主要集中在修道士过河这样一类智力难题的求解、下棋程序以及用计算机证明逻辑学和几何学定理等方面。

1963年以后，人工智能的发展步入形成期。这一时期人们开始注意运用自然语言（如中文、英文）通讯的能力，这是人类智能的重要标志。于是如何能使计算机理解自然语言、自动回答问题和进行图像分析就成了当时AI研究所追求的主要目标，这也是人工智能研究进入第二阶段的标志。

70年代人工智能发展到了第三阶段，即知识应用期。其标志是一批具有专家水平的程序系统在经过大量的科学推理探索之后相继问世。

80年代进入了以知识为中心发展的第四阶段，也叫综合集成期。这一时期针对不同应用领域研制了大量的专家系统，从而使越来越多的人认识到知识在模拟智能中的重要作用。^①

由于人工智能是一门交叉性学科，不同学科背景的学者对人工智能的理解各不相同，因此出现了多种不同的研究学派。不同学派研究的内容和途径各不相同，彼此间的争论异常激烈，这些为人工智能的发展做出了巨大贡献。例如以J. McCarthy和N. J. Nilsson为代表的逻辑学派；以A. Newell和H. A. Simon为代表的认知学派；以E. A. Feigenbaum为代表的知识工程学派（研究知识在人类智能中的作用与地位，提出了知识工程的概念）；以J. L. McClelland和J. D. Rumelhart为代表的联结学派（研究神经网络）；以C. Hewitt为代表的分布式学派（研究多智能系统中的知识与行为）；以及以R. A. Brooks为代表的进化论学派等。

从人工智能的研究途径来看，目前公认的研究学派主要有符号主义、联结主义和行为主义三大学派^②。

符号主义又称逻辑主义、心理学派或计算机学派，是基于物理符号系统假设和有限合理性原理的人工智能学派。符号主义认为，人类智能的基本单元是符号，认知过程就是符号表示下的符号计算，从而思维就是符号计算。它们主张运用计算机科学的方法进行人工智能的研究，通过研究逻辑演绎在计算机上的实现方法，实现人类智能在计算机上的模拟。

联结主义又称为仿生学派或生理学派，是基于神经网络及网络间的联结机制与学习

^① 史忠植 王文杰. 人工智能[M]. 国防工业出版社, 2007: 5-12

^② 张仰森 黄改娟. 人工智能教程[M]. 高等教育出版社, 2008: 6-8

算法的人工智能学派。联结主义认为，人类智能的基本单元是神经元，认知过程是由神经元构成的网络的信息传递，这种传递是并行式分布进行的。它们主张用仿生学的方法进行研究，通过研究人脑的工作模型，搞清楚人类智能的本质。

行为主义又称进化主义或控制论学派，是基于控制论和“感知动作”型控制系统的人工智能学派。行为主义认为，人工智能起源于控制论，提出智能取决于感知和行为（所以称为行为主义），取决于对外界复杂环境的适应，它不需要知识，也不需要表示和推理。智能行为只能在与实现世界的环境交互作用中表现出来，即人工智能也会像人类智能一样通过逐步进化而实现，而不需要知识表示和知识推理。

一、联结论的产生及其争论

（一）联结论的产生

早期人工智能的研究者都是用符号主义思想探究人工智能,主张运用计算机科学的方法,通过逻辑演绎以实现人类智能在计算机上的模拟。理论上我们可以基于数学逻辑对知识进行表示和推理,进而解决现实世界中的问题。但是逻辑推理并不是解决问题的唯一途径,有时非逻辑推理在求解问题的过程中起着更重要的作用,甚至是决定性的作用。比如说,人的感知过程主要是形象思维,这是逻辑推理做不到的,因而无法用符号方法进行模拟。另外,用符号表示概念时,有效性在很大程度上取决于符号表示的正确性,当把有关信息转换成推理机构能进行处理的符号时,将会丢失一些重要信息。比如说它不能准确地处理带有噪声的信息和不完整的信息。因此单凭符号方法解决智能中的所有问题是不可能的。

为了解决这一难题,一些学者尝试从生物学的角度来解决这一问题。仿生学派,即联结主义因此应运而生。联结论的学者对物理符号系统假设持反对意见,他们认为人工智能可以通过模拟人脑的结构来实现。人脑是人类一切智能活动的基础,人的思维单元是神经元,而不是符号处理过程。人脑不同于电脑,要进行人工智能的研究,就要对人脑的工作模式进行研究。从研究大脑神经元及其联结机制入手,搞清楚大脑的结构及其进行信息处理的过程与机理,揭示人类智能的奥秘,从而真正实现人类智能在机器上的模拟。

联结论的代表性成果为 MP 模型^①,它是 1943 年由生理专家 W. McCulloch 和数理逻辑学家 W. Pitts 创立的脑模型。MP 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。20 世纪 60 至 70 年代,以感知器为代表的脑模型的研究曾出现过高潮,但是当时由于受到理论模型、生物原型和技术条件的限制,脑模型研究在 70 年代后期至 80 年代初期落入低潮。直到 1984 年 Hopfield 教授提出用硬件来模拟神经网络,1986 年 J. D. Rumelhart 等人又提出多层网络中的反向传播(BP)算法时,联结主义才迎来春天。此后,从模型到算法、从理论分析到工程实现的研究,都为神经网络计算机走向市场打下了基础。

^① 张仰森 黄改娟. 人工智能教程[M]. 高等教育出版社, 2008: 7

联结主义学派的研究方法是以网络连接为主的联结机制方法，属于非符号处理范畴，研究的内容实际上就是神经网络。该学派的代表人物有：J.L.McClelland 和 J.D.Rumelhatr 等。

尽管目前人工神经网络的研究仍然热火朝天，但我们应该看到，以网络连接为主的联结机制方法其实并不适合模拟人的逻辑思维，而且就目前神经网络的研究现状来看，有固定的体系结构与组成方案所构成的系统还达不到开发多种多样知识的要求。因此单靠联结机制方法来解决人工智能中的全部问题也是不现实的。

（二）联结论的发展及其影响

联结机制认为，人类的认知过程并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，由类似于神经元的大量处理元素相互联结在一起。每个单元都有不同的活性，既可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，而这个动态过程被看作是认知能力。

在联结机制看来，突触是神经冲动从一个细胞跳到另一个细胞的联结点，这些突触上的变化是学习和记忆的基础。如感官的视觉冲动会以加强沿途所有突触的方式在网络上留下痕迹。差不多的情形同样会发生在由耳进入的听觉神经系统或大脑内其它脑际活动。结果是随意启动的网络会迅速将自己组织起来，通过某种正反馈积累经验，强健的经常被使用的突触会变得更加强健；而弱小的不经常使用的突触会萎缩。经常被使用的突触最后强健到一定程度后，记忆就被锁定了。这些记忆反过来又会布满整个大脑，每一个突触都与一个复杂的突触结构模式相对应，这些突触结构模式包含了成千上万个神经元，这种描述就被称之为联结论或联结机制。

鉴于以上的推论，联结机制认为人的思维基元是神经元，而不是符号；人的认知过程是权值的自组织而不是符号操作过程。因此人工智能应着重于结构模拟而不是功能模拟，即模拟人的生理神经结构。

（三）关于人工智能的争论

前面已经提到，在人工智能的发展过程上出现了多种不同的派别，不同的派别之间一直存在着激烈的争论。就人工智能能否实现一直存在着两个对立的派别——乐观主义派别和悲观主义派别。

乐观主义派别认为人工智能的发展前途无量,人工智能与人类智能并无实质性的差别,它不仅能替代,而且最终将超过人类智能。悲观主义派别不这样认为,他们把人工智能贬低为现代炼金术,认为它目前的成果不过是零敲碎打和拼拼凑凑,仅仅是在局部表面上对人类的摹仿,远未达到对人类智慧的实质了解。这一争论在某种意义上已经成为历史,因为人工智能并没有朝着单独支持哪一派的立场发展,它既向着深入认识、反映人类智慧的方向前进,又未能完全实现当初的既定目标。面对这种状况,争论双方冷静下来,意识到对方观点中的合理因素,争论因而趋于平淡。但这并不意味着派别之争就此消失,取而代之的是强 AI 和弱 AI 之争。强 AI 指出,人工智能不仅仅是实现人类智能的工具,事实上它就等同于人类智能;弱 AI 对此持反对意见,他们认为,人工智能仅仅是实现人类智能的工具,对它的评价不应越出这一范围。^①

除了上述几种派别之争外,不同学科背景的学者对人工智能研究的分歧还表现在技术和哲学这两种语境上。技术专家和哲学家在各自的领域里都可以自圆其说,但两种语境之间缺乏通约性,只能自话自说。这两种语境之别主要表现在人工智能的两大分支上,即传统 AI 和联结论 AI。传统 AI 是以符号逻辑为基础的算法系统,是由图灵和冯·诺伊曼规定出的那一套方法;联结论是建立在统计分布规律之上的并行分布式系统,包括对大脑神经网络的模拟。

下面简单回顾一下人工智能学者们在各自的研究领域对其理论的辩解。

首先,图灵在“计算机与智能^②”一文中提出了机器能否思维的问题,并对这一问题给出了肯定的回答。图灵的这一观点立刻受到了来自不同方面的攻击,其中最具影响的例子就是塞尔的反驳。塞尔在“心灵、大脑与程序^③”一文中运用图灵自己的计算概念来反驳他的一台配有适当程序的计算机是有智能的观点。塞尔认为,人工智能程序和计算机模型都应该是纯形式句法的(和图灵机一样),基于这一点,一个系统不可能纯粹借助完成计算而达到理解的目的。计算心理学绝不可能理解我们的心理能力,任何一个程序更加不可能将智能赋予计算机。智能或意向性不仅需要心灵式的行为,同时还需要作为这一行为的基础的正确的因果推导能力。

但是纽厄尔和西蒙反对塞尔的这一理论,他们在“作为经验探索的计算机科学:符号和搜索^④”一文中指出,在计算机同心灵哲学的关系上,心灵是一个计算系统,大脑

^① 博登. 人工智能哲学[M]. 刘西瑞 王汉琦译. 上海世纪出版集团, 2005: 487-488

^② 图灵. 计算机与智能[J]. 心灵, 1950.10: 433-460

^③ 塞尔. 心灵、大脑与程序[J]. 行为和大脑科学, 1980.3: 417-424

^④ 纽厄尔 西蒙. 作为经验探索的计算机科学: 符号和搜索[J]. 计算机协会通讯, 1976.3

事实上是在执行计算的职能，它与可能出现在计算机中的计算是完全等同的。人类智能可通过一组组控制着行为和内部信息处理的输入输出规则得到解释。由于计算机具备了正确的因果能力，它们就可以成为智能的。一台计算机（像大脑一样）是一个物理符号系统，而一个物理符号系统对于一般智能行为来说是必要的和充分的手段。

玛尔在“人工智能之我见^①”中就纽厄尔和西蒙的这一看法提出了指责。他认为，许多人工智能工作与科学无关，人工智能程序常常建立在一些互不联系的、理论上无依据的见解和毫无定责的经验摸索之上；由于纽厄尔和西蒙忽略了基础的、无意识的信息处理任务，他们的工作缺乏科学性（对算术来说模式匹配可能是关键性问题，而他们却视模式匹配为理所当然的）。

斯洛曼在“动机、机制和情感^②”中指出，动机、情感和意向仅仅是与大脑相关联并且最终在大脑中完成的多层虚拟机器集合的一些方面。尽管冯·诺依曼计算机（图灵机的近似形式）在原则上可以执行任何计算，但对于那些有多重目标而资源有限的相关类型的计算很可能需要由根本不同的方式（并行机）来实现。可见实现动物心理过程的机器（大脑）与计算机有很大的区别。

从上述人工智能学者间的争论中可以发现，联结论在很大程度上弥补了传统 AI 的不足，具有容错能力和较强的学习功能。德雷福斯在“造就心灵还是建立大脑模型：人工智能的分歧点^③”一文中指出：对于那些耳熟能详的批评，联结论可以在一定程度上为传统 AI 作辩护。但是他坚定地认为，语言和常识是不可能由人工智能获取的，即使是联结论类的也不行。他对人工智能的怀疑源于这样的观点：人们并不使用关于日常世界的理论，科学也不能表达这样的理论，因为并不存在一组与语境无关的理解元素。我们的知识是熟练的技能，不同于过程规则、表述方式或知识内容，甚至我们关于形式系统的知识也不能不借助于有关怎样延续数学序列或应用逻辑规则的背景直觉。

德雷福斯在概述人工智能两个分支激烈对抗的历史时还谈到了形式主义人工智能的初始成功、神经网络研究的早期成就和暂时沉寂、以及用常识性理解编程时遇到的困难等问题。他特别谈到了在传统 AI 中不能获得的整体知觉、语境敏感性和对族的相似性及相关性识别这些问题在联接论中都能得到了较好的处理。但是德雷福斯对智能理论，即使是建立在联结论信息加工基础之上的理论仍旧感到失望。他认为，只有大小和回路结构与大脑近乎完全等同的联结论系统才能做出人类智能的模型，这种系统还要具

^① 玛尔. 人工智能之我见[J]. 人工智能, 1977.9: 37-48

^② 斯洛曼. 动机、机制和情感[J]. 认知和情感, 1987.1: 217-233

^③ 德雷福斯. 造就心灵还是建立大脑模型: 人工智能的分歧点[J]. 人工智能, 1988.1

备人类动机、文化追求和血肉躯体。

事实证明了德雷福斯的预言是正确的，人工智能的研究越来越沿着他预言的方向前进。德雷福斯对人工智能的这一系列的批评引起了整个人工智能研究领域专家们的注意，很多研究人工智能的专家都承认他们从德雷福斯的批评中获得许多灵感，他的批评对他们的研究工作起到了积极作用。

二、人工智能为什么不能超越人类智能

20 世纪 60 年代以后的数十年时间，人工智能的研究并没有取得预期的成就，它在许多方面进展缓慢或停滞不前。这些挫折引发了种种怀疑的情绪，德雷福斯是那些怀疑论者和悲观主义者的领袖和代表人物。他认为维特根斯坦、海德格尔、梅洛·庞蒂等哲学家的哲学结论已经对人工智能的可能性提出了质疑，并断言人工智能肯定是有限度的。

在德雷福斯看来，整个人工智能的计划建立在生物学、心理学、认识论和本体论^①四个假想之上。

生物学上的假想：在某一运算水平上（一般认为是神经元的水平上）大脑以离散的运算方式加工信息，采用某种生物学的办法，等价于开关的闭合与断开办法。

心理学的假想：大脑可被看作一种装置，这一装置按照形式规则加工信息单位。在心理学看来，计算机的作用相当于大脑的模型，就像休谟这样的经验主义者所设想的那样（那些信息单位作为基本印像），也像康德这样的唯心主义者所设想的那样（由程序提供规则）。无论经验主义者还是唯心主义者都为把数据加工作为思维模式的想法提供了背景，这是一种第三人称的加工过程，在这一过程中，“加工者”并无实质性的作用。

认识论的假想：一切知识都可被形式化，也就是说，凡是能理解的都可通过逻辑关系来表达；更确切地说，通过布尔函数来表达，即表达为支配信息单位按规则互相关联的方式的逻辑演算。

本体论的假想：输入数字计算机的一切信息必须采取信息单位（比特）的形式，因此计算机大脑模型预设了关于外部世界的有关信息，以及对智能行为的产生有实质意义的一切原则上都必须可以分解为一组与局势无关的确定元素。这是关于存在就是一组在逻辑上完全互相独立的事实的本体论假想。

德雷福斯在逐个分析上述四种假想的合理性后得出的结论是：每个假想看似都是自明的、从未被提出疑问的公理，然而事实上这些假想只不过是可供选择的有疑点的假说，没有相应的经验性和先验性的论点来支撑。在德雷福斯看来，人工智能的发展是有局限的。

^① 德雷福斯。计算机不能做什么[M]。三联书店，1986：165-235

（一）生物学假想

计算机能不能像人脑一样工作？对此，有人选择用仿生学的方法从生物学的角度进行探索。从事认知模拟（CS）和人工智能（AI）的研究者坚信人类智能行为是数字计算机进行信息加工的结果。他们认为，既然自然创造了这种加工形式的智能行为，正确的程序设计就能为数字机器设计这种行为，采用的方式可以是对自然的模仿，也可以在程序设计上不受自然约束。

德雷福斯对这一观点提出了质疑。他首先提出把人脑看成一部电子计算机的假想，即认为人是成功的数字计算机程序活的实例。接下来引用现代数字计算机发明人 John von Neumann、麻省理工学院 Jerome Lettvin 和 Walter Rosenblith 等人的观点，最后得出了从当前有关人脑的经验证据中无法引出人工智能的可能性这一论点。德雷福斯的具体论证如下：

首先，即使大脑在某一层级真的像数字计算机那样工作，这也不一定就使 CS 和 AI 的研究者受到鼓舞。因为大脑可能是由许多随机相连的神经元连接成的庞大阵列，就像由明斯基当作早期控制论者拒之门外的研究小组提出的视感控器那样。用某种程序可以模拟这些神经网络，但在任何意义上讲，这种程序都不是启发式程序。因此不能仅仅以人脑或许是台数字计算机这个事实作理由就产生乐观主义看法，认为西蒙和明斯基所定义的人工智能定会成功。

其次，人脑中的基本信息加工过程能否从数字计算机模拟的角度得到最好的解释，这是经验性的问题。人脑的“信息加工”方式可能完全不同于数字计算机。比如说信息可能像电阻模拟机在通过一个网络求解最短路径问题那样，按全局的方式被加工。从经验论的角度看，当前的证据的确说明人脑神经元开关模型再也站不住脚。早在 1956 年，现代数字计算机的发明者之一，冯·洛伊曼就提出了他对这个模型的怀疑：

“现在，具体讲到人的神经系统，这是个庞大的系统——至少比我们所熟悉的任何人工制品要大 10^6 倍——而且它的活动也相应的复杂多样。他的职业包括对外部的感觉刺激，即传到大脑的物理条件和化学条件的解释，对运动肌的活动和体内化学条件的制约，通过极其复杂的信息搜索和转换而实现的记忆功能，当然还有不断地传送编码指令及或多或少的数量信息。通过数字方式（即，使用以二进制数表达的、或者经过更多的编码花样而用十进制以及其他进制表达的数目）控制所有这些过程。用代数的（基本是算术的）方法来加工数字化（通常是数值化）的信息，这都是可能的。这也许是人类设

计者目前愿意采取的解决这一问题的方式。目前已有的证据尽管不多，而且还不恰当，却可以说明人类神经系统的原理和步骤与此不同。因此，传递消息的脉冲流看来是用某种模拟式的特性来表达意义的（用脉冲的术语来说——似乎是一种半数字半模拟的混合系统），如果一个通道中的脉冲时间密度，一族通道内各脉冲时间序列的相关性等。”^①

因此，把人脑看成一个通用符号处理装置，像计算机那样工作，这是一种经验论的错误假说。从当前有关人脑的经验证据中，无法引出人工智能的可能性。事实上，大脑组织的这种“强烈相互作用”的性质，和机器组织的非相互作用性之间的差别，意味着只要生物学上的这些证据是中肯的，那么他就否认了用计算机产生智能的可能性。

（二）心理学假想

由于从生物学的角度无法解决大脑是否像数字计算机那样工作这一经验性问题，德雷福斯尝试从神经心理学的角度来解决这一问题，即解决思维能否像计算机那样工作，也就是说实用心理学上的计算机模型是否合理。大脑显然是一个使用物理过程转化各种信息的物理客体。如果说心理学和生物学是不同的，那么心理学家就必须在物理化学反应以外的某一层次上对大脑功能做出描述。德雷福斯引入信息加工这一假想，从认知模拟的角度对思维像计算机那样工作这一说法的正确性提出怀疑，并从经验性和先验性两方面入手找到了答案。

要对心理学假想作经验性的证实，就必须对经验性证据做出正确评价。心理学假想既不能从计算机和人的行为之间笼统的相似性得到证实，也不能因不能详细表述这种相似性就加以拒绝。对于心理学假想的检验，需要详细比较人类和机器信息加工所包含的各个步骤。

1975 年纽厄尔和西蒙曾预言心理学理论将在十年之内采取计算机程序的形式，并着手这一预言的实现，就是说编写一系列旨在模拟人类认知的程序，用来模拟一个人在完成某一特定的认知过程之前有意无意所经过的步骤，即“物理符号系统”假设。这种假设认为，无论是人还是计算机都是具有输入符号、输出符号、存储符号、复制符号、建立符号结构、条件性转移六种功能的完善的信息加工系统。信息加工系统也就是“符号操作系统”，或叫做“物理符号系统”。这个假说认为，任何一个系统，如果它的行为表现出智能的话，就必须具有前述的六种功能；反之，任何一种系统，如果能执行这六种功能，那么它的行为就会表现出智能的特征。从这一假设出发，他们又作出了三个

^① 诺依曼。概率逻辑与从不可靠成分到可靠组织的综合[J]。选集。1963

附带的推论：第一，既然人具有智能，它就一定是个完善的物理符号系统。他们认为，人类的智能就是在人对符号和符号结构的操作过程中，在对信息的加工过程中体现出来的；第二，既然计算机是一个完善的物理符号系统，它具有输入、输出、存储、复制符号以及建立符号结构、实行条件性转移六种功能，所以计算机的行为就能够表现出如同人脑一样的智能。这一推论，可以说是他们选择心理学途径进行人工智能研究的基本前提；第三，既然人和计算机都是完善的物理符号系统，二者工作的基本原理是一致的，所以，我们完全可以用计算机来模拟人的认知过程，包括对复杂问题的解决过程。也就是说，我们可以按照人类心理活动的操作过程来编制计算机程序，把人类记忆、理解过程的特点和解决问题的步骤、策略和经验编进计算机程序中去。这样，一方面模拟了人类的认知活动，就可以进一步深入了解人类认知活动的特点和规律。因为用计算机来模拟人的认知活动，就可以利用计算机程序形式化地描述人的认知过程，建立人类各种认知过程的模型或理论。另一方面，用计算机来模拟人的认知过程，将会使机器的行为表现出人类那样的智能。^①

用计算机模拟人类的心理活动的操作过程，无疑有助于我们对人类自身认知活动的特点和规律的了解与掌握，由此而建立的认知模型也能表现出某些如同人类大脑那样的智能，甚至能完成众多用人脑无法完成的工作，但是如仿生学途径一样，心理学模型无论如何也不可能像人的大脑那样地完成各种工作。其理由有三：

第一，把人的认知活动仅仅解释成一种信息操作活动缺乏严格的科学依据。把人脑比作一种信息处理装置，这至少在目前还只是一种假设。这就是说，像人脑的功能是否都可以通过信息处理活动来描述这样的问题，无论从理论上或实验上都远未得到证实。而选用心理学途径来从事人工智能研究的人，就是以这一“远未得到证实”的东西作为其推论的依据。而从一般的论证要求来看，这一作法显然违反了论证规则，犯了“预期理由”的错误。

第二，即使我们承认大脑是如同计算机等人工智能机那样一种信息处理装置，但是信息处理层面显然还不是认知层面，二者之间还存在着过渡的环节或层面。就是说，把信息处理视为从大脑的生理活动到达思维活动的必经之路或唯一环节是没有现实根据的。没有根据的东西作为推理的根据，其结论的科学性是很难令人信服的。也许正是由于这一问题的存在，使不少研究者开始注意对形象思维（以及顿悟）的特殊规律的探索。这从一定意义上也可以看出，用信息处理作为人脑的功能作用方式而作出描述和解释的

^① 纽厄尔 西蒙. 计算机与人的信息加工[J]. 美国科学家. 1964.9: 282

假设遇到了很大的困难。

第三，用计算机模拟所形成的心理学模型，也许能使机器表现出一定的智能行为，但这种机器决不可能像人的大脑那样完成各种智力活动。在建立心理学模型时，尽管不必把问题形式化，但是当用计算机程序模拟这个心理学模型时，还是少不了一个符号化、形式化的环节。因为计算机只能做符号处理工作，况且计算机本身无法完成由非形式化的心理学模型到形式化语言的变换，而只能按人们输入的一定程序进行刻板式的推演运算。且不说计算机自身无法实现对心理模式的模拟，就是提出一个心理模式也无法做到。它不可自动地应用实验心理学的方法去搜集和考察人在各种情境下解决各类问题时的言语材料和行为表现，不会对这些实验材料进行认真的甄别、分析与研究，了解人在知觉、记忆理解或解决各种问题时采用什么样的步骤和策略，有什么窍门或捷径，怎样规划自己的行动，等等。在总结出一些规律的基础上，提出心理学模型，这一切都只能由人来完成。计算机等人工智能系统不具有学习的能力，它不能从现实世界的实例与现实中获取并总结出知识。尤其对一些知识背景不清楚、推理规则不明确、环境信息十分复杂的知识处理，或是算法难以提取的信息处理任务更是显得无能为力。正是由于这一致命的缺陷，使得某些于人而言极易做到的事，如图像识别，计算机却显得困难重重，其识别过程远远超过人脑的视觉系统所花费的时间。

由此我们可以发现：鉴于实际成果和目前的停滞状态，人类在目前的技术条件下要造出智能机器在很大程度上是不现实的。因为这一认知模式偏离了目前公认的科学依据。

（三）认识论假想

认识论假想指的是，虽然不能用假设人实际上在一系列无意识运算中遵循启发规则的方法解释人类活动，但是在智能行为仍然可用这种规则加以形式化，因此又可用机器加以复制。

德雷福斯认为，明斯基的乐观主义观点是一种典型的认识论假想。明斯基坚信所有非随意性行为都可以形式化，数字计算机能适用所得到的形式化系统复制这一行为。正是出于这种信念，明斯基满有把握的断言：没有理由认为机器有人所没有的局限。然而德雷福斯认为在支持明斯基这一主张之前，首先应该弄清楚明斯基所说的机器是什么意思。

数字计算机是一种照柏拉图认为可用来理解有序行为的那种标准运转的机器。明斯

基是根据图灵的定义作出机器的这种定义的，他把机器定义为一种“服从于规则的机制”。正如图灵所说的那样，“这种——计算机应该遵循固定的规则——控制系统的作用是确保正确地并按恰当的顺序服从这些指令。控制系统是按照能起这种作用的目的构造出来的。”^①这里所说的机器，是那种加以限定了的但又是基本的机制，按照确定的、非二义数据的严格规则，对这些数据进行运算。于是明斯基就提出了这样的主张：这种机器（图灵机）体现了数字计算机本质的机器，原则上能做到人能做到的任何事情——原则上只有人所具有的那些局限。

通过明斯基对图灵机的这种理论化的设想，我们似乎找到了认识论假想成立的理论依据，但是从行为规律的二义性和信息加工上我们很容易发现这一理论是站不住脚的。因为当人类行为被看作运动——物理信号的输入和输出时，可能具有形式主义者所要求的意义上的有规律的性质。但是这丝毫没有支持见于明斯基和图灵那里的形式主义假想的作用。当明斯基和图灵提出人是图灵机的主张时，他们并不含有人是一种物理系统的意思。从另一方面看，说行星或船只是图灵机倒是恰当的。他们的行为也可以用数学公式化的规律加以描写——把它们纳入与输出的能量联系起来——至少在原则上能够在数字计算机上复制到任意一种准确的程度上。然而当明斯基和图灵说人也是一部图灵机时，他们指的一定是数字计算机能够复制人类行为，不是靠解算物理方程，而是靠以逻辑运算加工，代表有关世界事实的数据，这种方式就能简化为匹配、分类和布尔运算。这也就是明斯基所说的，“心智过程类似于——在计算机程序中见到的那些种类的过程：任意的符号联想，树形存储，条件转移之类的过程。”^②

AI 研究者还坚持存在着可描写为数字形式的符号描写的心智层次。所有 AI 都热衷于使用逻辑运算来处理表达世界的的数据，而不热衷结算描写物理客体的物理方程。从物理学方面的考虑可以看出：只有能的输入和转换能量是发生的神经活动，债能原则上以数字形式描写和处理。还没有人想法或希望想办法用物理规律计算人体的运动。这在物理上是不可能的：“没有任何一种数据加工系统，无论是人工还是自然的，能加工每秒每克大于 (2×10^{17}) 比特的信息。”^③

如果这些计算是正确的，那么企图把人脑模拟成一个物理系统就是一件不可能的事。那些必不可少的大量计算可能被这种计算中包含的物理规律及信息理论所排除。通过以上的论证可以说，物理运动中的形式化并不是人工智能研究中所说的对人类行为的

^① 图灵. 计算机机械与智能[J]. 思维与机器. 1950: 8

^② 明斯基. 物质、思维与模型[J]. 语义信息加工: 429

^③ 布里默曼. 通过进化与重组实现优化[J]. 自我组织系统. 1962.1

形式化。也就是说，从物理学的成功中得出的是错误的论断。

麦卡锡也曾经尝试着在物理学与行为科学之间找出它们之间的关系。“虽然人们为表达数学中最重要的领域设计出了形式化的理论，并在形式化某些经验科学方面也取得一些成就，但是目前还没有能够表达那种在日常生活中使用的、中点——端点分析的形式理论——解决人工智能问题需要一种形式理论。”^①

乔姆斯基和转换语言学家们找到了一突破点。他们发现，对人的活动（运用）——在具体场合下对具体句子的运用——进行抽象化，就能使其余部分得到形式化，即使人的识别在语法上造得好的句子和排斥造的不好句子的能力被形式化。这就是说，它们能够提出一部有关大部分语言能力的形式理论。这一成功鼓励了 AI 研究者们主张人类行为无需还原到物理层次上即可形式化，因为这一理论至少确认了认识论假设中人类行为可形式化这一部分。在有序行为中，一个乍看上去好像是无规则的片段原来可用复杂的规则加以描写，而这些复杂规则又是那种可直接由数字计算机加工的规则（这里的意思是：不必经由对讲者声带运动或大脑中发生的生理化学过程做物理描写的阶段）。但是，这种形式仅仅证实了认识论假说的一半内容。语言能力不是 AI 研究者们期待进行形式化的东西。如果机器使用自然语言进行交际，那么机器的程序不只应符合于语法规则，而且还应包括语言运用的规则。换句话说，为了能使句法理论形式化而未加考虑的东西——人们能够运用它们自己语言的事实——正是也应得到形式化的东西。

从德雷福斯的论述中我们可以看出：证实认识论假说是否正确合理的问题就变成了检验性的问题，即是否有一种关于语用的形式化理论。实践表明这种句法理论不可能存在，原因在于从原则性上讲，要得到语用的形式化理论就必须有关于全部人类知识的理论，但这是根本不可能的。从描述性上看，并不是所有的语言行为是有规则的。如有些词语的表达方式很奇怪，违反规则，但是我们仍然能够理解。

（四）本体论假想

既然从生物学、心理学和认识论假想上都没有找到满意的答案，那么能否从本体论上找到突破口呢？德雷福斯首先追溯了本体论假设在哲学史上的渊源。“根植于我们文化中的哲学传统的目的就是要消除不确定性：道德的、智识的和实践的。”^②在柏拉图那里就已经有一个要求，那就是知识必须按照规则或定义来表达，我们能够不带解释上的

^① 麦卡锡. 具有常识的程序[J]. 语义信息加工: 410

^② 德雷福斯. 计算机不能做什么[M]. 三联书店, 1986: 218

风险来应用规则或定义。在莱布尼兹那里，传统的知识观念与世界可以被分析为离散元素的明斯基式观点的联系变得明确起来。根据莱布尼兹，我们在知性中将概念分析为更简单的元素。为了避免更简单和更简单元素的后退，一定有一些终极简单的概念，按照它们我们能理解所有复杂的概念。此外，如果概念要应用于世界，世界上一定有些被这些元素所对应的简单物。莱布尼兹憧憬“一种人类思想的文字”，这种文字的字符被用于证明时能够揭示在对像中也可以发现的某种联系、聚合和秩序。经验主义传统也被这种离散知识要素的观念所统治。对休谟而言，一切经验都是由印像构成的：印像是可孤立的、确定的经验原子。唯智主义（intellectualist）和经验主义学派汇合在罗素的逻辑原子主义中，这个思想在维特根斯坦的《逻辑哲学论》中达到极致。在《逻辑哲学论》中，世界是按照一组原子事实来定义的，原子事实可以用逻辑上独立的命题来表达。这就是本体论假设的最纯粹的表述，也是人工智能中一切工作的必要前提条件。

这段论述包含着一个重要思想，借用霍布斯的一句就是，推理即计算：——“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名词。人们在数字等方面虽然除开加减以外还用乘、除等运算，但这些运算法实际上是同一回事。因为乘法就是把相等的东西加在一起，而除法则是一个东西能减多少回就减多少回。——不论在什么事物里，用得着加减的地方就用得着推理，用不着加减的地方就与推论完全无缘。——根据以上所说，我们就可以定义或确定推理这一词在列为心理官能之一时其意义是什么。因为在这种意义下，推理就是一种对公认为标示或指明我们思想的一般名词的序列进行计算（就是说，加和减）；当我们私下计算时，我说的是标示，当我们证明或者向他人证明我们的计算时，我说的是指明。”^①

给出了本体论假想后，德雷福斯随即对本体论假想提出质疑。他认为，在我们的日常活动中我们并没有经验到世界是一组事实，我们可能贯彻这一分析不是自明的。比如说计算机要完成相关指令就要对数据进行运算。按照本体论的假想，即使信息可以用离散的数据来表示，但并不是所有的数据都要进行运算。合理的处理这些数据，语境起了决定性的作用。“为了理解一段话，把一个问题结构化，或者识别一个模式，计算机必须按照一个语境选择和解释其数据，可是，我们如何将这个语境赋予计算机呢？”^②

^① 霍布斯. 利维坦[M]. 黎思复 黎廷桐译. 商务印书馆, 1985: 27-28

^② 德雷福斯. 计算机不能做什么[M]. 三联书店, 1986: 230

人工智能在这里面临抉择。一方面总是存在大量的语境，另一方面肯定存在这一个不需要解释的终极语境。在人类日常生活中，一些事实与其他事实极其广泛地联系起来，以至于即使能对它们进行明确的表述，也很难对它们进行分类。人类的生活形式赋予事实以意义，使它们成为它们之所是，从来不会有通过大量由无意义的、孤立的数据构成的表进行存储和分类的问题。

哥德尔曾指出数学的“不可完成性”，在此对应的是“语境的不可最终形式化”，它们都体现了一种深刻的思想：形式化方法不可能穷尽对数学和实在的终极刻画。不存在任何体现“终极语境”的事实，因此也不存在“语境上独立的数据”。所以说世界可以按照语境上独立的数据而被详尽无遗地分析的本体论假设是错误的^①。也就是说世界不可能最终被形式化，因此人工智能是不可能的。

^① Stuart Goldkind. *Machines and Intelligence: A Critique of Arguments against the Possibility of Artificial Intelligence* [M]. Greenwood Press, 1987

三、人工智能的未来

（一）人工智能的局限

通过对以上四个假想的质疑，德雷福斯从现象学的角度发现了人工智能的缺陷。现象学本是为了克服传统笛卡儿主义的主客两分而产生的，主体和客体不能独立地存在，人是通过与世界的关系而认识自身和世界的。通过将现象学的方法运用于人工智能的分析，德雷福斯认为，现实的结构是由人的需要和活动共同决定的，获得知识的最基本的方式是知觉而不是理性。知觉有意向性结构，知觉总是对某物的知觉，知觉的产生是通过人的身体在具体的情境中产生的，这些都不是将一切都形式化和符号化的机器所能具有的。人工智能实现的前提还是笛卡儿的理性主义，认为现实有独立的理性结构，人用同样的理性结构来把握现实，人的智能就是理性的形式规则对现实进行符号操作的过程。德雷福斯认为，人的智力不仅存在于大脑中，人的身体对智力的形成和发展也有一定的影响。另外，知识的运用涉及到具体的情境和人的需要，需要将过去的知识和现实情景结合起来进行创造性地运用，而这些能力都不可能被机器形式化地处理。只有在能被形式化处理的知识才能被计算机掌握。因此人工智能是有局限性的。

德雷福斯认为，人工智能可以划分为认知模拟（CS）与人工智能（AI）两个领域。人工智能研究就是试图解决人在信息加工中是否真的像数字计算机那样遵循形式化规则，以及人类行为（无论是如何生成的）能否被描述为一种可由数字机实现的形式化系统。这是两个相互独立而又相互关联的问题。

在讨论这两个问题时可以发现，离开传统的哲学偏见，单纯从描述性现象学的分析中可以看出，所有形式的智能行为都包含着不可程序化的人类能力。只要人工智能能否实现的问题还是一种经验性的问题，那么我们就可以给出肯定的回答，在认知模拟或人工智能中要想取得进展是不可能的。如果我们对人工智能的发展仍持乐观主义观点，我们就必须证明，“尽管有经验性的困难，人工智能肯定是可能的”这一命题。但是事实表明，支持人工智能的先验性理由要比经验性理由更脆弱。那些理应证明形式化一定是可能的论断，不是支离破碎就是自相矛盾相反的事例却说明，除了某些已被一致意见所排除、极其不可能的经验性假想外，形式化是不可能的。这样一来，肯定形式化的先验性论点变成原则上有条件否定CS与AI可能性的论点。

前面在讨论CS时就发现人类在解决复杂问题、辨别相似性事物、使用修辞性或不

符合语法的语言等问题时并不遵循严格的规则，而是使用全局性的感知组织，在本质与非本质运算之间做实用性的区分，借助范式实例达到互相交流的目的。当然这些看似有序，显然又是非规则性的活动，可能仍是无意识遵循规则的结果。德雷福斯通过回归现象方法来理解这一行为。这种回归无法依靠求助于普通事实的方法终结，因为按照原有的主张，这些事实本身必须用规则来识别和解释。避免规则回归的唯一方法，是修改认为在最低层次上规则无需指令而自动使用的论点和主张。在经验中已知的东西，一方面是输给这个机体的连续性物理输入，另一方面是给予这个感知主体的普通物体世界。还没有哪位认知心理学家成功地在这两种提供运用规则的终极信息的输入之间界定另外一类输入。目前已有的全部描述结果，成了一种不协调的混合物，其中一方面有从能的方面所做的物理描述，另一方面有从粗糙定义的感觉数据方面所做的现象主义描述。^①

因此在这里的结论只能是：智能行为像数字计算机一样遵循固定形式化规则而产生的观点是不能从回归中解脱出来的，也就是说人工智能的实现确实是有局限性。

（二）人工智能对人类社会的影响

虽然人工智能存在着局限性，但经过 50 多年的发展，人工智能的研究成果已经融入我们的日常生活，在科学、技术、军事应用、经济、社会、文化等领域对当今社会的发展起到了积极作用。在科学领域，学习和记忆是人工智能研究的核心，这些研究成果对心理学有着显著的影响。人工智能更多的是作为一种工具用于解决多种科学和技术领域中的问题。在技术领域，人机交互是人工智能研究的一个重要任务。人工智能对自动化的发展起到了积极的作用，在生产过程中对改善工作环境、提高生产质量、节约能源等方面都起到了积极作用。在经济领域，人工智能最重要的影响就在于就业和就业结构。在社会领域，高技术的引入将人们从繁重的体力劳动中解救出来；智能的计算机辅助教学帮助我们完善教育计划，为每个学生提出指导性的建议，对教育的创新与发展起到了积极的作用。

（三）人工智能展望

尽管人工智能的发展存在着局限性，但是人们对人工智能的研究将会持续下去，人工智能还是有出路的。德雷福斯在《计算机不能做什么》一书中对人工智能的出路也作了一个预测和展望。我们可以分出下列四类智能活动，其中前两类适于数字计算机模拟，

^① 德雷福斯. 计算机不能做什么[M]. 三联书店, 1986: 294-295

而第三类只是部分可程序化，第四类全部不可驾驭。

第一个领域是 S-R^①心理学家最熟悉的领域，任何形式的条件反射都是这类行为的事例，目前编成程序的最好例子是对无意义音节的机械学习。有一种被称作“Geography”（查地图）的游戏就属于这个领域。语言翻译中的机器词典属于这一领域。问题求解中的纯试错搜索程序和模拟识别中的根据固定模板匹配模式也属于此类。

第二个领域是帕斯卡的数学思维的领域，也是最适于人工智能的领域。这一领域是由概念世界而不是感知世界构成的。在这一领域之中问题完全形式化了，并且完全可以计算。正因如此，最好称其为“简单形式化领域”。人工智能在这里原则上和实际中都是可能的。在这一“简单形式化领域”之中，形式语言取代了自然语言，逻辑是最好一例。游戏有精确的规则并可完全计算出来，比如三子棋游戏和捡石子游戏就是这样。模式识别在这一层次上按照确定的类别进行，问题求解则取靠反复使用形式化规则，以缩短中点与终点之间距离的形式。这一领域中的形式化系统很简单，完全可以用根本不需要搜索程序的算法来操作。启发法在这里不仅没有必要，而且正像王浩的算法论及程序优于纽厄尔、肖和西蒙的启发式逻辑程序所表明的，启发法在这里已成为障碍。在这一领域中人工智能取得了唯一名副其实的成功。

第三个领域是“复杂形式化系统领域”，这一领域最难被定义。人工智能领域内的误解和困难大多数是由这一领域造成的。这一领域包括原则上可形式化而实际上无法驾驭的行为。所需转换的数量随着其中元素数量按指数方式增长。复杂形式化在这里还包括那些实际上不能用穷举算法处理，因而需要启发程序的系统。

第四个领域可称为“非形式化行为领域”，包括有规律但无规则支配的、人类世界中所有的日常活动。这是受控制单步只有精确性的最明显个案，是我们对自然语言中歧义的识别。这一领域还包括规则不确定的游戏，如猜谜游戏。这一领域中的模式识别采用范式实例在识别种属或典型的基础上进行。在这一层次上问题的结构开放，在解决问题之前需要确定什么是相关的，什么是运算成为本质的洞察力。这一层次上的技巧通常使用从实例中加以概括的方法教授，而后被自觉地遵从，无需求助规则。我们可以采用帕斯卡的术语，把第四领域称作感知思维的归宿。在这一领域，为避免存储无限多的事实对全局性局势的了解是必不可少的，原则上不可能使用离散的技巧来直接复制成人行为。

^① S-R 为 stimulusresponse 的缩写，刺激-反应的意思。

结 论

现阶段人工智能在专家系统、机器学习、自然语言理解、人工神经网络、智能决策系统等诸多领域取得了引人瞩目的成就，但这都离不开人类智能的帮助。无论人工智能在局部表现出多么强大的功能，人工智能只是作为一种工具被人类智能限制在特定的范围之内，其智能水平远不能和人类智能相提并论。人工智能的长期目标仍是建立人类水平的人工智能，由脑科学、认知科学、人工智能等共同研究，形成交叉智能科学。

通过德雷福斯对人工智能的研究，以及我们对德雷福斯的各种观点的分析，并结合最新人工智能的成果我们可以得出这样的结论：虽然作为一门研究机器智能的学科，人工智能取得了一定的发展，但是在扩展人类智能方面它仍然存在着局限性。人类对人脑的功能会不断地进行认识，从而人工智能会不断的逼近人类智能。但从动态分析上，即人类智能也在不断的进化和发展，人工智能作为人类智能主体客体化的产物，人工智能既不能取得真正的主体性，更不可能超越人类的理性思维和价值主体性的界限。只能是无限的模仿、延伸和扩展人类智能，其作用和功能受到人类智能的制约，所以人工智能要低于人类智能，不能真正超越人类智能。

参考文献

- [1] 德雷福斯.计算机不能做什么[M].北京:三联书店,1986
- [2] 博登.人工智能哲学[M].上海:上海译文出版社,2001
- [3] 张守刚 刘海波.人工智能的认识论问题[M].北京:人民出版社,1984
- [4] 许万增.人工智能对人类社会的影响[M].北京:科学出版社,1996
- [5] 杰夫·霍金斯.人工智能的未来[M].西安:陕西科学技术出版社,2006
- [6] 史忠植 王文杰. 人工智能[M]. 国防工业出版社, 2007
- [7] 张仰森 黄改娟. 人工智能教程[M]. 高等教育出版社, 2008
- [8] 胡塞尔.现象学.重庆出版社[M],2006
- [9] 梅诺庞蒂.知觉现象学.商务印书馆[M],200
- [10] 潘晨光.哲学社会科学前沿思考[M].北京:中国社会科学出版社,2004
- [11] 黄顺基 刘大椿.科学技术哲学的前沿与进展[M].北京:人民出版社,1991
- [12] 傅京孙.人工智能及其应用[M]. 北京:清华大学出版社,1987
- [13] 周昌乐.心脑计算举要[M].北京:清华大学出版社, 2003
- [14] 霍布斯. 利维坦[M]. 黎思复 黎廷弼译. 商务印书馆, 1985: 27-28
- [15] 图灵. 计算机与智能[J]. 心灵, 1950.10
- [16] 塞尔. 心灵、大脑与程序[J]. 行为和大脑科学, 1980.3
- [17] 纽厄尔 西蒙. 作为经验探索的计算机科学: 符号和搜索[J]. 计算机协会通讯, 1976.3
- [18] 纽厄尔 西蒙. 计算机与人的信息加工[J]. 美国科学家. 1964.9
- [19] 玛尔. 人工智能之我见[J]. 人工智能, 1977.9
- [20] 斯洛曼. 动机、机制和情感[J]. 认知和情感, 1987.1
- [21] 德雷福斯. 造就心灵还是建立大脑模型: 人工智能的分歧点[J]. 人工智能, 1988.1
- [22] 诺依曼. 概率逻辑与从不可靠成分到可靠组织的综合[J]. 选集. 1963
- [23] 明斯基. 物质、思维与模型[J]. 语义信息加工
- [24] 布里默曼. 通过进化与重组实现优化[J]. 自我组织系统. 1962.1
- [25] 麦卡锡. 具有常识的程序[J]. 语义信息加工
- [26] 徐向东.联结主义、排除主义与思想的构成性[J]. 自然辩证法通讯. 1994.5

致 谢

三年前为了备考而挑灯夜读的情景还历历在目，如今马上就要毕业，离开美丽的湖大校园了。从决定考研到完成这三年的学业，一路走来，有过迷茫，有过焦虑，有过彷徨，再次回顾走过的点点滴滴，我感到满心感激和无悔。

首先，我要感谢我的导师舒红跃教授在学习和生活中给予了我无私的帮助。在毕业论文的写作中，舒红跃教授从资料收集和初稿修改等各方面都给了我悉心的教诲和重要的启迪，并给予了我许多有益的建议。如果没有舒老师的帮助，我很难独自完成这样一篇论文。在这里，我用最真诚的心向我的导师舒红跃教授表达我最衷心的感谢。

其次，我要感谢所有在我读研究生期间传授过课程和知识的老师，他们优秀的学识修养和品格风范都深深的感染了我。他们对我在学术研究上的启迪和指导将成为我一生的宝贵财富。在此，我深深地表示感谢。

最后，我要感谢一直支持我的爸爸妈妈和三年来在学习和生活中关心我，爱护我的朋友和同学，正是有了他们的帮助和鼓励，我才能够顺利的完成研究生三年的学业，为我的大学生活划上一个完美的句号。在此，我向他们表示我最真诚的谢意。

学习的道路永无止境，我感谢我在求学的路上所经历的一切。有了这些历练和经验，我才能在人生的旅途上更加自信从容地去面对各种未知的挑战和困难。我自感掌握的知识 and 能力有限，这篇毕业论文仍显稚嫩，但我深深明白学术之路任重道远，我会更加坚定地走下去。

黄严磊

2010年5月 武汉