华中师范大学

硕士学位论文

塞尔的意向性理论与人工智能的未来

姓名：谢锦江

申请学位级别：硕士

专业：心灵哲学

指导教师：高新民

20090401

内容摘要

人工智能作为一门学科诞生于 20 世纪 50 年代，在其后短短的 20 多年里得到了飞速的发展，取得了令人惊叹的成就。然而在以后的发展中遇到了前所未有的困难，也产生了很多的争论。机器能否思维？人工智能能否具有真正意义上的智能？很多哲学家和科学家对这些问题进行了深入的理论探讨和工程学实践，但至今仍无定论。近些年来，一些哲学家把人工智能观点大致分为两类：强人工智能（强 AI）和弱人工智能（弱 AI）。作为强人工智能的代表人物图灵和作为弱人工智能的代表塞尔分别提出了两种截然不同的人工智能标准。图灵通过他的“模拟游戏”提出只要机器通过测试就说明能思维，能思维就是具有智能。塞尔则认为意向性才是智能的标准。

在与人工智能有关的诸多哲学问题中，意向性是最重要的方面之一。之所以这样，是因为意向性是人工智能成为真正的智能必须具备的条件。塞尔认为，意向性是一种自然或生物现象，是自然生命史的一个组成部分，所以自称他的理论为“生物学的自然主义”。而要造出人工大脑，只仿造输入输出过程是不行的，而要仿造意识过程，而意识又是意向性的基础。由此可以看出，要想达到人类智能的水平，解析意向性是极其关键的一环。

塞尔认为意向性可以分为内在的和派生的意向性两类，只有人才具有内在的意向性，其它非人系统所具有的只是派生的或隐喻的意向性。而意向性是自然进化的结果，人的意向性更是具有内在的特性，即使我们承认某些机器具备意向性，而这类意向性在塞尔看来也只是派生的或隐喻的意向性。由此可见，人工智能的发展要想达到智能的层次，就必须要解析意向性，搞清楚意向性的起源、本质特征及其实现条件等问题，而要解决这一问题，就必须要向大自然这一最好的导师学习、取经，才能最终跨越人工智能的意向性鸿沟，达到真正意义上的智能。

关键词： 意向性；塞尔的意向性理论；智能；人工智能

Abstract

As a discipline, Artificial Intelligence(AI) was born in the 1950s and in the later short span of 20 years it had a rapid development and amazing achievement. Then the further development of Artificial Intelligence met unprecedented difficulty and caused a lot of controversy. Can machines think? Can Artificial Intelligence have a real sense of intelligence? Many philosophers and scientists had prosecuted a profound theoretical inquiry and engineering practice into these problems, however it still remained controversial. In recent years, concept on AI was divided into two kinds by some philosophers, strong Artificial Intelligence (strong AI) and weak Artificial Intelligence (weak AI). As the representative of strong AI and weak AI, Alan Turing and Searle put forward two distinct standards of AI respectively. In accordance of simulation game, Turing considered that machine can think as long as it passed the text. Capability of think is equated with intelligence. However, Searle believed that intentionality should be the standard of intelligence.

The intentionality is the most aspect of Philosophical questions about Artificial Intelligence. The reason is that intentionality is the prerequisite which making Artificial Intelligence to a real sense of intelligence. Searle believed that intentionality is a natural or biological phenomenon, and it is a part of natural life history. So, he named his theory " biological naturalism". It is not the sufficient conditions to create Artificial brain that only making sure correct input-and-output. But the imitation the process of consciousness must be based on intentionality. Therefore, analyzing intentionality is the key to create a human intelligence.

Searle considered the intentionality can be divided into two kinds: the inherent intentionality and the derived intentionality. Human has the inherent intentionality, and the other un-human system only possesses the derived intentionality. The intentionality is the production of natural evolution. Even though some machine can present a intentionality, in Searle's oppion, it is only derived intentionality or metaphor intentionality. From the above we can see, in order to pressing ahead the development of Artificial Intelligence to the standard of real intelligence, we must know how the

intentionality originated, and make explicit the essential quality and of the conditions of intentionality.

Key words: intentionality Searle's theory of the intentionality intelligence Artificial Intelligence

华中师范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：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的研究成果。除文中已经标明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对本文的研究做出贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

作者签名：谢锦记　　日期：2009 年 6 月 2 日

学位论文版权使用授权书

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：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，允许论文被查阅和借阅。本人授权华中师范大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。同意华中师范大学可以用不同方式在不同媒体上发表、传播学位论文的全部或部分内容。

作者签名：谢锦记　　导师签名：

日期：　年　月　日　　日期：2009 年 6 月 2 日

本人已经认真阅读“CALIS 高校学位论文全文数据库发布章程”，同意将本人的学位论文提交“CALIS 高校学位论文全文数据库”中全文发布，并可按“章程”中的规定享受相关权益。同意论文提交后滞后：□半年；□一年；□二年发布。

作者签名：谢锦记　　导师签名：

日期：2009 年 6 月 2 日　　日期：2009 年 6 月 2 日

导 论

自古以来，人们就对智能问题很感兴趣，经常提起的就是“人如何比其他动物聪明”、“某人比某人更有头脑”等话题。不仅如此，人类从来没有停止过在劳动工具上的创新，从最低级的石斧到近代的机器，无不体现着对智能的追求。“但是它们都没有像计算机那样给人类带来如此广泛、快速而深刻的变化。这是因为，计算机是伸延和放大人类智能的机器。”[①]而人工智能作为计算机科学的一个分支，产生于 20 世纪 50 年代，在其后的短短 20 多年间，人工神经网络、进化计算、专家系统、人工生命、人工免疫等迅速涌现，并取得了一系列惊人的成绩：造出了代替人类完成各种复杂任务的机器人，证明了尚未证明的数学定理，进行某些医学手术，下棋、甚至打败国际象棋冠军……然而在很多方面人工智能又显得非常笨拙、稚嫩，无法完成对人类来说十分简单的任务，如：正确理解人的日常对话，在日常环境中决定自己该做或不做什么……人工智能发展中为什么会存在着这样的现象呢？究其原因就是由于人工智能是一门特殊的科学，它的研究对象不同于一般科学研究的对象，而是人的“智能”，面对这种研究对象，我们应该怎样研究它，智能是什么，它是怎样起作用的？人工智能的未来发展路径在哪里？在此情况下，除了工程学、自然科学领域的研究之外，人工智能哲学也为其发展提供重要的理论依据。

在与人工智能有关的诸多哲学问题中，意向性是最重要的方面之一。之所以这样，是因为意向性是人工智能要成为真正的智能必须具备的条件。人的一言一行在多数情况下都是由自我意识引导完成的，积极、主动，而机器所做的每一件事情都需要事先做出规定，机械、被动。意向性在主动与被动之间划出一条分明的界线，人和机器的许多差别都可以在这里找到解释的根据。

意向性问题不仅是是心灵哲学的重要问题之一，同时也是许多学科关注的一个热点问题，如人工智能、认知科学、心理学、语言哲学等都很重视这一问题，并取得了大量富有成效的研究成果。美国当代著名哲学家约翰•塞尔认为：意向性问题是心灵哲学和语言哲学的交汇点，并且在意向性与人工智能方面提出了智能的标准问题。虽然在人工智能研究方面，心灵哲学与其它科学有着迥然不同的思路，可最后都把目光集中到了意向性这一问题上来。国内对于意向性与人工智能的研究方面，如刘西瑞、王汉琦翻译了英国玛格丽特.博登的《人工智能哲学》一书，以及郭斌的《塞尔“中文屋”思想实验的哲学意蕴辨析》、唐热风的《心的本质是计算吗?》、

[①] 冯天瑾，《智能学简史》，北京：科学出版社，2007 年，第 95 页。

蔡曙山的《哲学家如何理解人工智能——塞尔的“中文房间争论”及其意义》、张云台的《人工智能及其前景的哲学思考》等文章，对意向性与人工智能等方面的问题进行了一些有益的探讨。而真正对这一问题进行深入研究的还是我校政法学院哲学系高新民教授，高老师早在《现代西方心灵哲学》（武汉出版社）一书中就对意向性问题进行了专题讨论，并对于非人系统意向性与人的意向性之间的差异进行了论述。在 2008 年出版的《意向性理论的当代发展》一书中，更是把意向性与人工智能单独作为一章深入加以探讨。另外，还在多篇论文中专门探讨过这一问题。这些都对于我国在意向性与人工智能方面的研究作了具有开创性的探索。

意向性的现代意义是由布伦塔诺和胡塞尔赋予的。布伦塔诺以意向性作为区分心理现象和物理现象的标志，胡塞尔延用了这一概念，并不断拓展其内涵，最终将它改造为一个纯哲学用语，还把它作为他本人创立的哲学体系——现象学的重要基石。另外，由于意向性与意义的产生有着密切的关系，也在语言哲学中占有重要的地位。尤其在计算机出现后，机器主体化、语言指称等问题使意向性问题受到更多的关注，其中哲学家塞尔的观点很具代表性。

塞尔是近年来活跃在人工智能领域中的哲学家，他十分重视意向性，并著有《意向性：论心灵哲学》一书，对意向性进行了详细的解析与论证。关于意向性，他是这样说的：“意向性是为许多心理状态和事件所具有的这样一种性质，即这些心理状态或事件通过它而指向或关于或涉及世界上的对象和事态。”[①]

塞尔认为，意向性是一种自然或生物现象，所以自称他的理论为“生物学的自然主义”。而要造出人工大脑，只仿造输入输出过程是不行的，而要仿造意识过程，而意识又是意向性的基础。由此可以看出，要想达到人类智能的水平，解析意向性是极其关键的一环。

塞尔在探讨意向性的时候，把意向性分为内在的和派生的两类，而他认为只有人类才具有内在的意向性，其它非人系统所具有的只是派生的或隐喻的意向性。即使我们承认某些机器具备意向性，而这类意向性在塞尔看来也只是派生的或隐喻的意向性。由此可见，人工智能要想达到真正意义上的智能，就必须要解析意向性，搞清楚意向性的起源、本质特征及其实现条件等问题，这些问题的解决在塞尔看来，就必须要向大自然这一最好的导师学习、取经，才能最终跨越智能的意向性鸿沟，达到真正意义上的智能。

---

① 约翰·塞尔，《意向性：论心灵哲学》，刘叶涛译，上海：上海人民出版社，2007，第 1 页。

第一章　图灵测试与中文屋论证

人工智能自从 20 世纪 50 年代创立以来，取得了令人惊叹的成就。如 IBM 公司的吉宁特（Herbert Gelernter）建造的几何定理证明机，能够证明很多连数学系学生都感到头疼的定理，还有跳棋、象棋程序等都比它的设计者的水平更高。人工智能的成果以其快速的发展不断逼近人类智能的领地而备受关注，但在人工智能的背后，有很多难以解决的工程学及理论问题，从而引起了很多争议，其中最多的就是机器能否思维、人工智能能否真正实现（极限）的问题。对于人工智能的研究，图灵（A. Turing）做出了开创性的贡献，被誉为“人工智能之父”。

一、“图灵测试”与计算主义

人工智能的开拓者、自然哲学家、数学家图灵 1950 年在《思想》杂志上发表震惊人工智能界的著名论文：《机器能思维吗？》，提出了判断一台计算机是否具有智能的测试方法，这也就是后来人们所说的图灵测试。

“机器能思维吗？”在当时的哲学界已经是一个引起激烈争论的问题了，而图灵则认为对这个问题的定义是不清楚的，要作出新的回答，就要找到合适的提问方法。他认为应该用另外一个与它密切相关，并且没有歧义的问题来替代它，即：计算机能否通过“模拟游戏”的测试？计算机能否和人一样无区别地回答提问者的问题？换言之，判断计算机有无智能，不应看其内部活动和过程，而应看它的实际行为表现。机器如果能完成人需要用智能完成的活动，如果能像人一样回答提问，那么它就具有和人一样的智能。

（一）“模拟游戏”

图灵的“模拟游戏”是这样的：游戏由三个人来做，一个是男人（A），一个是女人（B），还有一个提问者（C），其性别不限。C 待在与 A、B 分开的另一间房子里。提问者 C 的目的是，通过提问判断 A、B 两人中哪一个是男性，哪一个是女性，他以 X 和 Y 分别称呼 A 和 B。C 的问题可以是这样的：

C：X，请你告诉我你的头发长度，可以吗？

假定 X 实际上是 A，那么 A 必须做出回答。A 在游戏中的目标是，尽量使 C 作出错误判断。于是，他可能回答说：“我的头发是瓦盖式的短发， 最长的一束大

约长 9 英寸。”

为了不让提问者从声音中得到帮助，回答问题应当是写出来，最好打印出来。而游戏参与者（B）的目标是尽量帮助提问者。对她来说，最好的策略或许就是如实回答。她在回答时，可以加上这样的话：“我是女性，别听他的！”但是这种做法无济于事，因为那个男士也可以运用同样的方式。

图灵认为，现在我们要问的是：“如果在这个游戏中用一台机器代替 A，会出现什么情况？”在这种情况下做游戏时，提问者作出错误判断的次数，和他同一个男人和一个女人做这一游戏时一样多吗？这样就替代了原来的问题：“机器能够思维吗？” 图灵努力把原来的问题作一替代的目的就是试图把智能与人的其他能力区分开来，从而达到把他认为的与智能无关的因素在游戏中祛除。图灵由此断言：计算机肯定能通过上述测试，即让 C 分辨不出他是男性还是女性，那么该计算机就具有了智能。

（二）关于“机器”

在图灵测试中关于“机器”一词的含义至关重要，这也是理解图灵测试的关键所在。“本文关于‘思维机器’的兴趣是由一种通常称为‘电子计算机’或‘数字计算机’的特殊机器引发的……根据这一建议，我们只允许数字计算机参与我们的游戏。”[①]图灵认为我们人类的思维是遵循一定规则的，那么一台数字计算机应该由三个部分组成：存储器、执行单位和控制器。存储器是用来存储信息的，“相当于人类计算机使用的纸张，可以在这些纸上做演算也可以用它来印出充满规则的书。”[②]存储器中的信息一般被分成一个一个的数据包，而数字则按照某种规则被分配到存储器之中。执行单位是执行各种具体的运算的，运算的内容因机器的不同而不同。而供计算机使用的“规则书”则可以用机器中的部分存储所取代，这部分则称之为“指令表”。控制器的主要作用就是监督这些指令是否按正常的顺序依次执行。这就是图灵所设想的计算机模型，而这一机器的运行是从一个完全确定的状态，转变为另一个完全确定的状态，这些状态之间不可能混淆。也即是说：如果给定了一个初始状态和输入信号，那么就可以完全确定未来的状态是什么样的。“如果 A，那么 B”是图灵机的基本特征。图灵认为“即使我们考虑的是实际的物质机器，而不是理想化的机器，在某一时刻对于某一状态的足够准确的知识，在其后任何数量的

[①][①]A·M·图灵，《计算机器与智能》，载玛格丽特·博登：《人工智能哲学》，刘西瑞、王汉琦译，2006 年，第 47 页。
[②] 同上，第 48 页。

步骤上，仍能产生出足够准确的知识。”[1]由此“机器能思维吗？”这一问题就被“存在着可以想象得到的能够在模拟游戏中干得出色的数字计算机吗”所替代。如果从图灵机的普适性来说，也即是“我们只讨论一台特定的数字计算机C。如果经过改进，使这台计算机具备合乎要求的存储，动作速度也相应提高，同时还为它提供了恰当的程序，那么在模仿游戏中，由人担任B的角色，C就能够令人满意地扮演A的角色，这有可能成为事实吗？”[2]

从上述讨论可以看出“图灵测试”的重要意义在于：1、它给出了一个判断智能的标准，即行为标准，他所关注的是最后的行为表现，而不是内部的活动及过程，把内部的关于意识的东西先放在一边不予解决。2、从第一点我们可以得出这样一个结果：图灵不关注“内部”实际上也就是等于承认了机器与人在“内心”方面是不相同的。尽管“图灵测试”有着行为主义的倾向，但他的计算机模型仍然是现代意义上的计算机的一个基本思想，为现代计算机的诞生起到了奠基性的贡献，直到现在我们用的计算机仍然运用的是这一模式。

（三）计算主义

计算主义又称“强人工智能”。能够通过图灵测试的计算机就是具有智能的，图灵的这一观点成为人工智能走向计算主义的理论根据。所谓计算，就是把有句法结构的句子作为输入，然后根据符号加工规则，产生有句法结构的符号作为输出。计算机在其计算过程中并不理解符号的意义，它们只能理解句法属性。输入的句子具有语义属性，这是计算过程得以完成的前提。因此计算机是句法机。福多说：“计算过程是在有句法结构的对象内所发生的过程。从外延上看，计算是从符号到符号的映射；从内涵上看，它们是从句法描述下的符号到句法描述下的符号的映射。”[3]正如图林所说，如果我们有一种装置，它的操作是符号转换，它的状态变化是由它所转换的符号的句法属性驱使的，但仍有这样的可能，即该装置会可靠地将真的输入符号转换为同样真的输出符号。他认为人脑能做到的一切，机器都能模仿。

计算主义的心智学说有两种基本范式，一是传统的物理符号主义；二是20世纪80年代兴起的联结主义。尽管计算主义的理论形式发生着变化，但其核心观点都是将心灵看作是一台或多台能够执行计算任务的计算机。图灵认为，可求值计算的功能可以等同于一种在有限步骤之内能够求值的计算功能的实现过程。因此，完

[1] A·M·图灵，《计算机器与智能》，载玛格丽特·博登：《人工智能哲学》，刘西瑞、王汉琦译，2006年，第52页。
[2] 同上，第53-54页。
[3] Fodor,J.A.(1994):The Elm and the Expert . Cambridge,Mass:MIT Press,p.8.

全可以设计具有这样一种功能的计算机器。这样一来，形式与计算因此紧密交叉而产生了一个推论：能够被形式化的任何自然现象也能够通过合适正确的某种机器复制（或者模拟）。图灵认为，在这个意义上，机器能够“复制”人类计算者所做的一切，包括人的心智，这就奠定了“机械程序可以实现心智的计算，心灵的本质是计算”这一心灵的计算理论的基本思想。心灵的计算理论以揭示“心灵”的奥秘为己任，把产生精神现象的大脑比作计算机系统的硬件，把心灵比作计算机系统的软件。进而认为，人的心智与计算机程序具有一样的工作原理。因此思维的本质就是计算，人的心理过程就是一个计算执行过程。

图灵不仅利用伟大的数学成就开辟了一个崭新的领域，同时也把源自于数学界的形式与语义之争遗传给了他所开创的计算机科学、人工智能科学及其循此而勾画人类心灵图景的计算主义哲学家那里。关于心灵的计算理论的形式与语义之争大致包括了这样三个基本问题的争论：一是句法能够揭示语义吗？二是如果句法能揭示语义，那么在计算中，语义是如何获得了唯一的真值呢？三是即使获得了语义的真值，难道这等同于人类的理解吗？面对这样的问题，心灵的计算理论陷入了困境。即使我们肯定句法能够揭示语义，那么这种形式化的计算是否就是智能？我们来看塞尔的“中文屋论证”是如何来回应这个问题的。

二、“中文屋论证”与意向性

在哲学界，对于人工智能观点大致可以分为两大类：一类是强人工智能，认为“机器可以具有真正的智能”，如上所述图灵的观点；另一类是弱人工智能，认为机器智能只是一种模拟智能，即使某台计算机通过了图灵测试，它对问题仍然没有任何理解，计算机不可能具备真正意义上的智能，这也正是哲学家塞尔（John • R • Searle）的观点。为此，他提出了著名的“中文屋论证”的思想实验。虽然他的直接目的是批评尚克所设计的一个程序，但其意义远不在于此，它很重要的一点就是否定了图灵的智能观。他说：“这些论证同样适用于威诺格拉德的SHRDLU，魏曾鲍姆的ELLZA，当然还有图灵机对人类心理现象的各种模拟。”[①]

（一）“中文屋论证”

塞尔的思想实验是这样的：

[①] 约翰 • 塞尔，《心灵、大脑与程序》，载玛格丽特 • 博登：《人工智能哲学》，刘西瑞、王汉琦译，2006年，第74页。

假设塞尔本人被关在一个封闭的房子里，给了他一大批中文文本，且他本人对中文一窍不通，对他来说，中文文本与其它无意义的曲线一样。再假设，在第一批中文文本之后，又给了他第二批中文脚本，并带有一套规则，能使第二批与第一批发生联系。规则是英文的，用这些规则，塞尔可以把一组形式符号与另一组形式符号联系起来。再假设又给了他第三批中文符号，还有一些指令，指令是英文的，这些指令能使塞尔把第三批的中文符号同前两批联系起来，并指示他怎样送回某种特定形状的中文符号，作为对第三批送给他的那些特定形状符号的回应。在这里第一批符号叫“脚本”，第二批符号中“故事”，第三批符号叫“问题”，而送回回应第三批符号的符号叫“对问题的回答”，那套英文规则叫作“程序”。而塞尔又假设又给他一些他所能理解的英文故事（事实上他一定能理解），然后再给他一些关于这些故事的问题，并让他回答。

那么，第一种情况下，当塞尔送回的中文符号答案与一个讲中文母语的人的答案毫无区别时，我们就认为他通过了图灵测试，但是这时的塞尔对中文一无所知，更不可能理解他所送回的中文符号了。而在第二种情况下，塞尔的回答与其他讲英文母语的人的回答一样时，显然是理解英文故事的。所以在中文场合下与在英文场合下，对于塞尔来说显然是不同的。在中文场合下，对于讲中文母语的人来说塞尔就像是一台计算机，他是根据规定好的形式规则来执行计算操作的。由此塞尔认为：计算机不可能有任何理解，不可能具备真正的智能。他说：“如果认为尚克那台编程计算机理解故事的意义，只是就门能理解那种比喻的意义而言，而不是就我理解英文的意义而言……我将论证：从严格的意义上讲，编程计算机所理解的，正是汽车和加法机所理解的，就是说，恰恰什么都不理解。计算机的理解并不只是（像我对德语的理解那样）局部的或不完全的，而是零。”[①]

（二）智能的根本特征：意向性

塞尔认为人工智能可以被复制是不可能实现的。他认为相同的程序可以通过各种没有任何的意向性的形式实现。例如可以使用一卷手纸和一堆小石头来构造一台计算机。同样地，中文故事的理解程序也可以被编入一系列的水管和一套风机或者是只懂得说英文的人中，这些没有一个需要理解中文。石头、手纸、风、水管都是没有意向性的错误的材料，只有那些和大脑具有相同因果能力的事物才可以有意向性，并且即使讲英语的人具有产生意向性的正确的材料，也很容易看到他并没有通

[①] 约翰·塞尔，《心灵、大脑与程序》，载玛格丽特·博登：《人工智能哲学》，刘西瑞、王汉琦译，上海：上海人民出版社，2006年，第78页。

过记住程序而得到额外的意向性，因为记住程序也不能教他中文。塞尔认为计算机程序完全是形式化的，而意向状态不是形式化的。它们是根据内容来定义的。

由此可见，塞尔之所以认为计算机不可能具有真正的智能（理解）是因为他认为计算机所运行的程序是形式化的句法实体，而人脑所具有的智能是有语义性的。“理解”不仅仅是句法的转换，它背后隐含有语义，也即是句法对于“理解”是来说不是充分条件。塞尔在1990年对此提出：1、计算机程序是形式化的句法实体；2、头脑具有精神内容（或语义）；3、语法自身对语义而言是不充分的；4、精神源于大脑。[①]而要使计算机具有真正的智能，最为关键的就是计算机要有“中文屋论证”中提出的意向性这一特征。只有真实的大脑过程对于意向性的产生才是充分的，而基于运行程序的计算机来说永远不可能成为产生意向性的充分条件。

总之，在塞尔看来，图灵机和尚克的程序所做的事情并不能说明它们有智能。因为智能不在于形式转换，而在于能处理内容或有意向性。人类的智能之所以是真正的智能，根本原因在于：它有意向性。他说：“大脑产生意向性的那种因果能力，并不存在于它例示计算机程序的过程中，因为无论你想要什么程序，都能够由某种东西来例示这个程序，而它并不具有任何心理状态。无论大脑在产生意向性时所做的是什么，都不可能存在于例示程序的过程中，因为没有一个程序凭借自身而对于意向性来说是充分的。”[②]

[①] 冯天瑾，《智能学简史》，北京：科学出版社，2007年，第196页。
[②] 约翰·塞尔，《心灵、大脑与程序》，载玛格丽特·博登：《人工智能哲学》，刘西瑞、王汉琦译，上海：上海人民出版社，2006年，第95页。

第二章　塞尔的意向性理论

从“中文屋论证”可以看出，人工智能要成为真正的智能就必须具备意向性，这也是近年来人工智能、脑科学、认知科学等很多涉及智能领域的科学家们的共识。在西方心灵哲学中，从布伦塔诺、迈农，到胡塞尔、梅洛—庞蒂都把意向性问题作为其理论的核心问题。而约翰·塞尔对于意向性的探讨独特而新颖，在心灵哲学中影响巨大，对人工智能的发展有着非常重要的哲学意义。

一、意向性的概念界定

要了解塞尔的意向性理论的全貌，首先要了解“意向性”这一语词。“意向性”一词在西文里的原意是“渴望”、“意愿”，在此基础上引申为意向。对于意向性问题，历史上很多哲学家都作过研究，最早可以追溯到亚里士多德。“根据亚里士多德的观点，在知觉中心接受所知觉物体的形式而并没有接受其物质。当我知觉一匹马的时候，我知觉的是马的形式。”[①]也就是说这个形式在我心中和在马的思维中具有不同的存在形式，在马的思维中，这个形式具有自然的存在，而我在关于马的思维中，马的形式就是意向性的存在了。而在中世纪，一些经院哲学家，如：圣·安瑟伦（St.Anselrn）托马斯·阿奎那都对意向性问题进行过研究。而意向性的现代意义是由布伦塔诺和胡塞尔赋予的。塞尔从心灵哲学的角度对意向性问题提出了新颖而独特的见解。

“意向性”一词是一个用法比较混乱的概念。不仅心理学、哲学等学科广泛使用它，而且日常生活中也有它的多种不同的用法，如“某人有做某事的意向”、“双方有合作的意向”、“在这一点上，他是有意向的”等等。尽管如此，但有一点可以肯定，即意向性的确是心理现象的某种真实的、根本性的方面或特征。因为许多心理状态或事件都具有关系性特征，如“想到”“围绕”“关于”等，意向性体现的正是一状态与别的外在的事态的一种关系。鉴于此，在哲学上，一般都倾向于这样定义意向性，即认为它指的是心灵关于（about）或指向（directat）别的事物、状态、属性等的特征。

约翰·塞尔是这样定义意向性的：“意向性是为许多心理状态和事件所具有的

[①] 唐热风，《心身世界》，北京：首都师范大学出版社，2001 年，第 128 页。

这样一种性质，即这些心理状态或事件通过它而指向或关于或涉及世界上的对象和事态。”[①]也就是说一心理状态或事件如果指向、关于或涉及其他事态或对象的性质特征，那么它就具有意向性。如我有一个愿望，那一定是要做某事的愿望，或是关于某事应该发生的愿望。

二、意向性的本质特点与实现条件

在《意向性——论心灵哲学》一书中，塞尔提出并论证了他的意向性理论，而意向性理论也是他整个哲学体系的基础。只有详细解析塞尔的意向性理论，了解意向性的起源、结构、运行机制、本质等问题，才能对人工智能的发展提供理论依据和实践基础。

（一）意向性的起源

从笛卡尔以来，心身问题一直是很多哲学争论的焦点之一，也是哲学史上的一大难题。塞尔认为，这个问题之所以难解，首先是由于我们用一套过时陈旧的范畴体系来说明这一难题。也就是说，“你要么是一个一元论者，要么就是一个二元论者。如果你是一个一元论者，那么，你要么是一个唯物论者，要么就是一个唯心论者。如果你是一个唯物论者，那么，你要么是一个行为主义者，要么就是一个物理主义者，等等。”[②]这样我们在讨论心身问题时，只能在这些范畴的束缚下进行抉择，所以也就不可能得到解答。再者，我们在讨论此问题时，提前就预设了这样两类事物：一是心理事物，如思想、情感，把它们看成是主观的、有意识的、非物质的；二是物理事物，具有质量、广延的。然后解决的方法就是否定其中一类事物的存在，或在一定程度上去贬低一类事物的地位。

塞尔则认为心身问题其实解答起来相当简单。“这种解答既与我们所具有的神经生理学知识相一致，又与我们关于心理状态——痛觉、信念、欲望等等的性质的常识的概念并行不悖。”[③]他总结了心理现象的四个特征：意识性、意向性、主观性、因果性，其中意向性又是意识的基础，所以解决了意向性何以可能的问题，心身问题也就随之而解决了。那么意向性是何以可能的呢？真空中的原子怎么会具有意向性？它们又是怎么能涉及、指向外物的呢？

要解开意向性之谜，就要去揭示其怎样实际地存在着，“就是要尽可能详细地

---

① 约翰·塞尔，《意向性：论心灵哲学》，刘叶涛译，上海：上海人民出版社，2007，第1页。
② 约翰·塞尔，《心、脑与科学》，杨音莱译，上海：上海译文出版社，2006年，第7页。
③ 同上，第7页。

去描述那些怎样由生物的过程引起，同时又体现在生物系统中的现象。”[①]塞尔认为：“全部心理现象，不论是有意识的或无意识的、视觉的或听觉的，不论是痛觉、角痒觉或思想，乃至我们全部心理生活，都是由在脑中进行的过程产生的。”[②]也就是说像视觉、听觉、痛觉等都是由脑过程引起的，并体现在脑结构之中，它们都是具有意向的心理现象，它们都实际地存在着，就像万有引力、光合作用、消化过程一样的存在着，“其神秘性不多也不少”。如现代科学已经揭示的某些类型的渴觉就是由发生在中枢神经系统中，主要集中于下丘脑中的一系列活动引起的，并在下丘脑中体现出来。具体来说是在下丘脑中由连续的神经传导引起的，这些连续的传导依次由下丘脑中血管紧张肽的活动引起，而血管紧张肽依次由肾脏分泌出的高血压蛋白原酶合成。另一个例子就是塞尔所说的患有某种肿瘤的病人由于肿瘤压迫下丘脑以至于总是感到口渴，不论喝多少水也解决不了他的口渴感。而下丘脑部位受到损伤的病人则从来不感到口渴。上述例子充分说明了，像渴觉等有意向性的心理现象是神经生理学的事实，而意向性则是生物进化的结果，是一种生物学现象。这就是塞尔生物学自然主义的观点。

（二）意向性的结构

要解析意向性，首先就是要把握意向性本身的结构，只有这样我们才能解决心身关系问题，才能了解心灵如何通过意向性与世界相联系？才能对人工智能的发展提供理论研究和实践基础。

1、意向内容

意向内容就是意向状态或事件所带的命题。意向内容使得心的状态得以指向一个对象，如“天在下雨”这里的意向内容是整个命题“天在下雨”。但并不是所有的意向状态都有一个完整的命题作为意向内容。如“我打算今天去学校。”这里“今天去学校”可去可不去。也就是说，此意向内容可真可假。由此也就引出了意向性状态的另外一个结构特征——满足条件。

2、满足条件

满足条件就是意向状态或事件得以成真的条件。塞尔强调，理解意向性的关键就是满足条件。“如果世界就是以某意向状态所表示的方式存在，那么这种意向状态就是被满足的。信念可以是真的或假的，愿望可能实现也可能受挫，意图可以被实行或不实行。在每一种情况中，意向状态是否被满足取决于命题内容和被表示的

[①] 同上，第16页。
[②] 同上，第10页。

实在之间实际上是否一致。”“具有命题内容的意向状态的一个一般特征就是它们具有满足的条件。实际上，如果需要一个用来分析意向性的口号，我相信，这个口号应当是：‘根据它们的满足条件就会知道它们。’”[①]如，“我相信天在下雨。”如果天确实正在下雨，那么我关于天正在下雨的信念是真的，意向状态就被满足了；如果天没有下雨，那么我关于天正在下雨的信念就为假，意向状态就没有被满足。所以我们要想知道一个人的意向状态是什么，只要问一问，在什么条件下，它确实得到满足或没得到满足。

3、适应指向

“不同类型的意向状态可以说是用不同的适应职责（obligations of fitting）.把命题内容与实在世界相关联。”[②]如一信念是否得到满足取决于世界是否实际上如信念所表示的那样。因此，塞尔认为，信念具有心灵向世界的适应指向。而意图、愿望等具有世界向心灵的适应指向（world -to- mind direction of fit）。塞尔用 G •E •M •安斯库姆所举的例子生动地说明了这两种适应指向。

一个男人按照他夫人给他的购买清单去市场买物品：啤酒、黄油和香肠。男人于是到超市按照清单上所列的物品，把它们一一放到购物车中。男人的这种行为其实就是让世界适应清单。“单子所起的作用就像一道命令或一个愿望，因而具有世界向单子的适应指向（world-to-list direction of fit）。”[③]假设在男人购物时有个侦探跟踪他，并记录了他所买的东西：啤酒、黄油和香肠，形成了另外一张单子。这时这个男人和侦探就各有了一个相同的单子。但是这两个单子的功能是截然不同的，即具有不同的适应指向，妻子给丈夫的清单具有改变世界从而使它与单子符合，也就是实现世界向心灵的符合；而侦探的单子是对实际发生的事件的一种描述，即心灵向世界的符合。塞尔认为：“这种区别对于任何意向性理论来说都是至关重要的。”[④]另外，并不是所有意向状态都具有“世界向心灵”或“心灵向世界”的适应指向。如：因为做错了事而感到后悔，取得了成绩而感到高兴，对别人的成功表示祝贺等，在上述情况下你所具有的意向状态都是以命题内容已经得到满足为前提。“对于这些情况，我认为，它们具有零适应指向（the null direction of fit）。”[⑤]

---

① 约翰·塞尔，《心灵、语言和社会》，李步楼译，上海：上海译文出版社，2006 年，第 101 页。
② 同上，第 98 页。
③ 同上，第 99 页。
④ 同上，第 100 页。
⑤ 同上，第 101 页。

（三）意向性的种类、特征、本质

1、意向性的种类

意向性的归属有不同的种类，正是由于这些不同才造成了对意向性认识上的一些混乱，因此我们有必要对它们作一下区分。

（1）有意识的意向性和无意识的意向性

从字面就能看出，所谓有意识的意向性就是有意识的心理状态或事件指向、关于或涉及其它客体或事态，例如，我相信鸟会飞（清醒时）；相反，无意识的意向性就是心理状态或事件指向、关涉对象时，我没有意识到，如，我在睡觉时相信鸟会飞。

（2）内在的意向性和派生的意向性

为了探讨内在的和派生的意向性，塞尔举了下面的例子作一区分：

A、罗伯特相信罗纳德·里根是总统。

B、比尔看见天正在下雨。

C、“Es regent”是下雨。

塞尔认为：A 句完全把一种有意向的心理状态归属于某一个人；而 B 句则不仅仅如此，它包含了更多的内容，因为“比尔看见天正在下雨。”意味着他不但具有了某种意向性的形式，而且这种意向性还实际获得了满足条件，即得到了满足。因为，“看见”用塞尔的话是一个“成功动词”，它表明看见的某事已经实际存在着。“我们完全可以说一个睡着的人相信这样或那样的事，但不能说他看见了这样或那样的事。信念和视觉经验都是存在于行为者心和脑之中内在的意向现象。说它们是内在的仅仅是指这种状态和事件确实存在于行为者的心脑之中。”[①]虽然 A 句和 B 句存在着区别，但不管是相信或是看见都是某人大脑中内在的意向性现象。C 句虽然确实有意向性，但它的意向现象不是内在固有的，而是派生的。也就是说，这个语句其实可以指别的事件或什么也不指，它的意向性是从说德语的人更基本的内在意向性形式中派生出来的。

（3）真实的意向性和隐喻的意向性

所谓真实的意向性就是心理状态所应有的属性，其真实性取决于某种意向现象的存在。不论是有意识的、无意识的，还是内在的、派生的意向性都是真实的意向性。而隐喻的意向性则是把意向性赋予本来没有意向性的东西。如：

D、我车子里的恒温器感觉到了发动机温度的变化。

① 约翰·塞尔，《心、脑与科学》，杨音莱译，上海：上海译文出版社，2006 年，第 111 页。

在这句话中，不存在真实的意向性，因为车子的恒温器是没有感觉的。塞尔在《心灵的再发现》一书中对此种意向性作了详细的说明：

A、我现在很渴，真的很渴，因为我已有一整天没喝过东西了。

B、我的草坪很渴，真的很渴，因为已经一星期没浇水了。

第一个句子描述的是某人的真实的、意向性的心理状态。如果我说了这句话并且为真，那么是我的渴的感觉使得它为真，因为它包含了对喝水的欲望。（在上面分类时已讨论过）而第二句没有真正的意向性归属，描述中的“意向性”只是仿佛（as-if）的而不是真实的。虽然我的草坪确实需要浇水，但是草坪只是好像、仿佛很“渴”，“好像、仿佛”的意向性归属只是一种比喻，而非真实的。“仿佛意向性不是一种意向性，一个系统具有仿佛意向性（as-if intentionality）是说它仿佛具有意向性。”[①]

2、意向性的特征

（1）意向性的首要特征就是指向性、关于性

意向性是大多数心理状态或事件所具有的一种属性，这一属性使得心理状态或事件直接与外部世界的对象和事态相关。也就是说，一种状态如果有指向、关于、涉及其他事态的特征，那么它就具有意向性。比如，信念、愿望、害怕、希望等这样的心的状态，它们总是指向某对象的。如果你希望或想要，那么就一定是希望或想要某个事物。如果我害怕，那么必定是对某种东西害怕或害怕某事会发生。总之，它们都是关于、涉及世界上的对象和事态的，指向性就是许多心理状态和事件的一种性质，心灵凭借这种性质从而指向世界上的客体。

（2）意向性不是一切而只是有些心理状态和事件的一种属性

比如信念、愿望都是有意向性的，因为我的信念或愿望必定总是关于某种东西的。但也有一些心理状态不具有意向性，如莫明的焦虑、高兴、紧张，此时却只是感觉焦虑、高兴、紧张，但不指向或涉及到它们之外的什么东西。塞尔对意向性这一特征的论述与布伦塔诺的观点是有很大区别的，布伦塔诺明确指出意向性是心理现象的独有特征，从而把它作为区分心理现象和物理现象的标准。塞尔认为：“而某些非心理的事件和状态也可能有意向性，例如计算机就没有心理状态，但其计算也可能有意向性。”[②]

（3）意向性不同于意识

在很多关于意向性的理论中都认为意向性只是人所特有的特征，只存在于人的

---

① 约翰·塞尔，《心灵的再发现》，王巍译，北京：中国人民大学出版社，2005年，第69页。
② 高新民，《现代西方心灵哲学》，武汉：武汉出版社，1986年，第596页。

心理现象之内。塞尔则认为许多有意识的状态都不是意向性的，如上文所提到的不因任何东西的高兴或得意，而许多意向状态也不是有意识的，如我相信我生活在中国，但直到现在为止，我从没有有意识的表述或考虑过这个信念。所以，我们不能到有意识的心理状态中去寻找意向性，也不能根据意识说明意向性，即不能认为只要有意识就会有意向性。把意识与意向性等同的人模糊了一个关键性的区分："当我具有一种对忧虑的有意识的经验时，实际上就会存在某种东西，我的经验就是涉及它的经验，也就是涉及忧虑的经验，但'涉及'的这种意义与意向性的'涉及'大不相同，比如后者出现在我有一种对蛇的有意识的恐惧这个陈述当中；因为就忧虑来说，忧虑的经验和忧虑是相同的；但是，对蛇的恐惧却不等同于蛇。按照我对这个概念的用法，下面这一点是意向状态所特有的：在这种状态和该状态所指向或关于或涉及的东西之间存在着区别。"[①]总之，不仅有意识的心理状态或事件，而且无意识的心理状态或事件，甚至一些非心理现象也都是有意向性的。所以，"有意识的状态这个类和意向心理状态这个类交叉，但它们并不全同，也并不是一个包含在另一个当中。"[②]

3、意向性的本质

从上述对意向性的产生、特征、种类的分析，我们可以看出意向性的本质：它不是什么超时空的、非存在的东西，也不是什么副现象，意向性是真实存在的，是我们自然生物进化的结果，是自然生物生命史的组成部分，是和呼吸、消化、胆汁分泌一样发生在生物机体内的真实的现象。塞尔在继承前人的基础上，对意向性的范围进行了扩展，认为不仅人的心理状态有意向性，甚至一些非人的自然事物（树的年轮等）也有隐喻意义上的意向性。但这类意向性从根本上说源于人类心灵的作用，它们没有内在的、固有的意向性，其原因是它们没有意识。内在的意向性体现了心理状态与对象的关系，心理状态总有一个对象。因此，从根本上说意向性包含着意识机制，意向性离不开意识。例如，一块石头的纹理像一个人的画像，那是由于我们"看"的作用，它只是一种派生的意向性。"一切派生的意向性都是从内在意向性派生出来的。""我们需要把人类和动物内在的具有的意性同语词、句子、图画、图表和图形那种引申出来的意向性区别开来。此外，我们还要把这两种意向性同比喻的意向性归属区别开来。"[③]总之，人的意向性是原始的、内在的、固有的，而人之外的事物以及非心理状态的意向性都是根据心理状态的意向性，它们可以在比喻、派生的意义上具有意向性。

---

① 约翰·塞尔，《意向性：论心灵哲学》，刘叶涛译，上海：上海人民出版社，2007年，第2-3页。
② 同上，第3页。
③ 约翰·塞尔，《心灵、语言和社会》，李步楼译，上海：上海译文出版社，2006年，第90-92页。

（四）意向性的基础、实现条件、运行机制

1、意向性的基础

塞尔认为：意识和意向性都是心理现象的两个重要特征，两者不是相互分离、互不相干的，而是相互依赖的。但两者的地位和作用又是不同的。从时间上说，意识先于意向性；从作用上说，意识优于意向性，前者是后者的基础，后者是前者的产物。

塞尔强调应根据意识来解释意向性，因为意识先于意向性，比意向性更根本，“意识对于我们得以在在上生存下去的最重要的特征就是它使我们得以了解我们自身的意识状态之外的世界。”[①]因此是解释意向性的最好根据。从起源和存在上说，意识是由人脑这一生物系统最先产生出来的心理现象，同时又贯穿于各种心理生活和心理形式之中，甚至是主观性、意向性、合理性、心智因果关系的基础和根源。他承认：主观性、意向性、合理性、自由意识和心智因果关系等是心灵的特征，但它们都依赖于意识。“意识和意向性之间在这样一个关键性方面有着本质的联系：一个主体的心理状态的归属性要么是一种意识状态的归属，要么是一种可能成为有意识的状态的归属。”[②]也就是说，任何意向状态要么现实地、要么潜在地是一种有意识的意向状态。如，塞尔说：我们在谈论琼斯时说到“琼斯相信克林顿是美国总统”，即使琼斯在睡觉时我也可以这么说。因为此时归之他的并不是此刻的一个有意识的信念，而是他头脑的一种能力使他能具有关于克林顿是美国总统的有意识的信念。无意识的状态由于它的能力使它在原则上能够产生有意识的心理状态。

2、意向性的实现条件

每一种意向状态都有各自的意向内容，并且是通过无数其他意向状态的关系来决定其满足条件的。即意向状态不是独立地起作用，而是和其他意向状态密切相关。例如，某人想去超市买东西，他就必须不要具有关于是否有超市、超市在什么地方等很多其他的意向状态，如果没有这些意向状态，他就不会形式“去超市买东西”这样的意向。所有这些信念、愿望等意向状态只是其他心理状态组成的一个更大的复合体的组成部分，塞尔把这个完整的整体性网络简称为“网络”（Network）。虽然意向性网络中的许多意向状态（是否有超市）对于某一意向状态不是至关重要的，但只有把这一意向状态放在与其相关的意向状态网络中，它才能有其满足条件。而背景则是指“能够使所有表征得以发生的一个由非表征心理能力组成的集合”。意向状态只在其本身并不是意向状态的能力的背景下才具有它们的确具有的满足条

[①] 同上，第74页。
[②] 同上，第75页。

件，也就是说背景是“前意向性的”，它不是意向性的一种形式，而是意向性的一个在先条件或集合。总之，意向状态并不是作为一个孤立的心理能力而起作用的。意向状态实际起作用要有一套预设的背景能力，背景提供了一个起作用的集合，它使得意向性发挥作用成为可能。就像塞尔所说的：美国宪法能够使某人潜在的候选人形成变成总结的意向，背景提供了一个“起动条件集”。“意向性发生在行动与感知的协调流之中，而背景是该协调流采纳形式的可能性的条件。”[①]例如，你在吃饭，在工作，在打球，每种情形中，实施的可能性的条件就是背后的背景能力。

3、意向性的运行机制

意向性与因果性之间有无关系？意向状态能否因果地起作用？塞尔认为，意向性与因果性是协调统一的，在因果性是有意向性的参与。“意向内容必定是一个相关的层面，而且它必定例证一种可计划的规律性。”[②]塞尔从传统的因果关系入手，即传统因果性的图式：“台球 A 撞击台球 B 这个事件引起了 B 的运动这个事件”。按传统观点我们实际观察到的是一个事件跟随着另一个事件，并没有观察到有任何因果关联。只是能够观察到类似的事件反复出现使得我们认为这两个事件因果地相关。即：一、因果关系本身不可观察；二、只要存在一对关系为原因和结果的事件，就会例证某种一般性的规律；三、因果规律性不同于逻辑规律性。塞尔则认为因果性都必然是意向因果性，如，由于“我口渴”而引起了“渴水”这一行为，这时我无需任何不同于所讨论的事件的经验的进一步观察而便可知道因果问题的答案。再则因果性不必与因果律相关，因为我知道是什么让我去喝水：口渴了，但当我说出这一点时，我并没有保证任何因果律的存在。另外，原因与结果之间好像存在着逻辑的关系，上述例子说明了这一点。但是我们：“不用考虑口渴是如何被描述的，它都包含喝水的愿望，而这种愿望就将一个人的喝水作为其满足条件了。”[③]即对原因的描述和对结果的描述之间好像存在着逻辑的关系，是由于原因本身和结果本身之间存在着逻辑关系，也就是存在着一种与其满足条件因果相关的意向内容。在塞尔看来，意向因果性是因果性的初始形式，因果性是是意向的因果性。“不能够做出行动或进行感知的存在物，不可能具有我们对于因果性的经验。”[④]“在每一个场合，我们都既成发现因果的成分，也发现意向的成分。心灵的表现能力和对世界的因果关系应当以某种系统的方式协调一致，这一点对于意向性发挥作用是很重要

① 约翰·塞尔，《心灵的再发现》，王巍译，北京：中国人民大学出版社，2005 年，第 161 页。
② 约翰·塞尔，《意向性：论心灵哲学》，刘叶涛译，上海：上海人民出版社，2007 年，第 142 页。
③ 同上，第 124 页。
④ 同上，第 135 页。

的，并且确实对于我们在世界上的生存也是很重要的。”[①]

（五）意向性与意识

意向性不同于意识。英美大多数哲学家所接受的看法：意向性并不是一切心理现象的共有特征，有些心理现象的独特特征是现象意识，还有一些现象同时具有这两个特征，因此布伦塔诺对心理现象的统一理解是值得怀疑的。塞尔认为意识是意向性的基础，所以要搞清意向性就要搞清什么是意识及意识与意向性的关系。

1、意识

大多数哲学认为意识就是“发生在我们身上、伴随着我们的某些心理活动的一种内在的过程或状态或现象。”[②]我们每个人都能经验自己当下的心理活动。当我醒着时，我是有意识的；当我睡觉（不做梦）时，意识就中止了。不过要对意识进行说明却不是一个容易的事。首先，意识与观察的关系。我们平时能够观察到很多东西，如课本、桌椅等，但却不能以同样的方式来观察意识。对于意识本身来说，我们不能像对其他观察对象那样把观察与被观察的东西区分开来。再者，意识经常被认为是在世界之外的东西，而不是普通物理世界的一部分。塞尔则认为，意识是像任何其他生物学现象一样的生物学现象。

2、意识与意向性

意识是意向性的基础，意识发挥它的作用时，它就具有了意向性，即指向存在或不存在的东西。大部分意识都是有意向性的。塞尔认为，意向性是心理状态的一种属性，以此心理状态指向、关于世界上的对象和事态。也就是说意识是意向性的基础，“在很多情形中，意识实际上是意识到某物，而‘意识到’中的‘到’是关于意向性。”[③]虽然不是所有的意识状态都具有意向性，但至少一部分或很大部分是这样的，所以要解决意向性就必须考虑到意识。按照意向性理论，每一意识活动指向某物或事态时，意识就处于意向状态，意识所指向的事物或状态就是意向对象，意识活动就与意向对象之间就形成了一种意向关系。“只有拥有有意识的意向状态，才能拥有意向状态，所有无意识的意向状态至少潜在地是有意识的……任何关于意向性的讨论，如果遗漏了意识问题，将是不完备的。”[④]

---

① 约翰·塞尔，《心灵、语言和社会》，李步楼译，上海：上海译文出版社，2006年，第102页。
② 高新民，《现代西方心灵哲学》，武汉出版社，1986年，第552页。
③ 约翰·塞尔，《心灵的再发现》，王巍译，北京：中国人民大学出版社，2005年，第110页。
④ 同上，第111页。

（六）意向性与意义

塞尔认为语言哲学是心灵哲学的一个分支，他试图用意向性理论来解决意义问题。他说：意义是一种意向性，意义就是头脑之中。

塞尔的“中文屋论证”思想实验的目的就是为了说明计算机并不具有思维，不了解形式符号的意义。在实验中，塞尔本人通过了图灵测试，在房子外的中国人以为房子里的人懂汉语，而实际上塞尔根本就不懂，甚至分不清楚那些符号是中文还是日文。这个思想实验表明：计算机程序不能构成真正的理解。因为计算机只是根据形式规则来操作符号，并不理解所操作符号的意义。人们操作的符号具有意向性这样的语义性，这是基于我们与世界的交互，而计算机的表征是独立于外部世界的，缺少人的意向性，不能赋予符号以意义。房子里的人并不懂中文，更不理解中文的意义，但根据程序规则，也能和懂中文的人一样得到问题的相同答案。这说明对于意义的获得并不需要凭借外在的东西。

塞尔坚持内在主义的观点，认为语言的意义是由使用者大脑的心理事件以内在的方式决定的。“某种形式的内在论必定是正确的，因为没有什么别的东西有此作用。大脑是所有我们所具有的、用来完成把世界表征给我们自己的这一任务的东西，我们所能使用的每一样东西必定都在大脑当中。”[①]

三、意向性的自然化

“意向性”这个词究竟有没有所指？是否存在？意向性在自然界中有没有地位？当代很多对意向性的研究持有物理主义的观点（唯物主义者、物理主义者、功能主义者等），认为世界上的一切事物和现象都是物理的。但塞尔认为，这些人其实都心照不宣地假定有意识的心理事件的特征不可能像世界上任何其他较高层次的特征那样作为世界的普通物理特征，他们最终都要否定世界内在心理特征的存在。而塞尔则坚持认为，的确存在如内在心理现象这样的东西，它们不能被还原为另外的东西，或通过某种类型的重新定义而被取消，心理状态与任何其他生物现象一样真实，“和哺乳、光合作用、大便或者消化一样实在。和这些其他的现象一样，心理状态由生物现象引起，并反过来引起其他生物现象。”[②]塞尔称为“生物自然主义”(biological naturalism)。塞尔既否定二元论，因为它不能解释心理现象；同时又认为唯物主义把意识还原为物质的现象也是错误的。他不满唯物主义没有给

[①] 约翰·塞尔，《意向性：论心灵哲学》，刘叶涛译，上海：上海人民出版社，2007年，第236页。
[②] 同上，第271页。

予心灵在世界中以正确的地位。他说：我们人和其他高等动物像其他有机体一样是自然生物系统秩序中的一部分，人与自然中的其他部分是连续的。因此，“对心—身问题的正确解决，并不在于否认心理现象的实在性，而在于正确把握它们的生物本性。”[①]

（一）意向性在自然界中的地位

关于意向性在自然界中的地位问题，有很多观点。“当代心智哲学的各个学派都根据各自的理论给出了各自的回答。物理主义认为心理状态其实就是一种物理状态；行为主义则认为意向心理状态只不过是人类有机体的一系列刺激——反应行为；唯物主义又认为物理学的陈述和心理学的陈述是谈论同一事实的两种方式；功能主义则认为意向状态只是有机体的功能状态；如此等等。”[②]所以，无论是物理主义、行为主义还是功能主义他们都否认内在心理特征的存在，不承认意向性是一种生物学现象，是一种客观实在。而塞尔则认为，心理现象是由大脑的动作引起的，同时又在大脑的结构当中实现。意向性和消化或血液循环一样，都是人类生物学的组成部分，是自然界的一种客观实在。所以意向性在自然界中肯定有其地位。“总之，某些生物体具有内在的意向状态，这些意向状态是由这些生物体神经元系统中的过程导致的，并且它们是体现在这些神经元系统的结构之中的。”“按照达尔文的意思，脑导致心理现象并不比物体具有引力更神秘。”[③]

（二）意向性——一种被自然化的生物学现象

塞尔对意向性的讨论是生物学自然主义的，即把意向心理现象当作我们自然生物生命史的一个组成部分，是生物机体的真实内在特征。人为什么有意向性？塞尔从渴是怎么一回事说起：机体系统中由于缺少水分便引起肾脏分泌凝乳酶，凝乳酶作用于被称之为“血管紧张肽”的循环液，产生“血管紧张肽 2”。这种物质进行大脑，作用于下丘脑区域，使该区域神经元放电频率增加，从而引起动物产生一种有意识的喝水的欲望。这就是生物学上最原始的意向性形式的神经生物学过程，由此过程引起如渴这类有意识的意向性状态，是自然运作的结果。“感觉口渴、具有视觉经验、怀有愿望、恐惧和期待与呼吸、消化、睡眠一样都是一个人生物生命史的组成部分。”[④]因此，他认为意向性根源于一种生物学机制。意向性是心灵与世界相

---

① 同上，导言第 3 页。
② 刘高岑，《语言哲学的心理意向性理论及其哲学意义》，科学技术与辩证法，2004 年 04 期。
③ 约翰·塞尔，《心、脑与科学》，杨音莱译，上海：上海译文出版社，2006 年，第 114-115 页。
④ 同上，第 113 页。

联系的中介，心灵指向、关于、涉及世界上的事态，心灵是自然的一部分，可以表现为信念、愿望等等不同的意向状态。塞尔通过生物学语言对心理现象进行分析，从而使意识和意向性得以自然化。认为意向性不是什么超验现象，它只不过是有机体大脑的一个属性。他从内在意向性对自然界的依赖性以及它在自然界中的地位，认为意向性作为一种现象不是存在于自然界之外，而是自然界的一个组成部分。“内在意向现象是由在脑中进行的神经生理过程导致的，并且是在脑结构中发生和实现的。”[①]也正是由于有了意向性，人才能够跨越时空关于那些在场或不在场、存在或不存在的事态。而意向性的这一生物学机制在当代已经被神经科学、脑科学、生物学所证实。

[①] 同上，第 113 页。

第三章　塞尔的意向性理论对人工智能的启示

通过对塞尔意向性理论的详细解析，可以看出：人工智能要想有重大突破，即要具备真正意义上的智能就必须具备产生意向性的能力，而意向性在塞尔看来是生物长期进化的结果，是生物生命史的一个组成部分。因此，只有向大自然学习，才是人工智能发展的必经之路。

一、智能的源泉——生物进化

塞尔在解决心身问题时，提出了生物学自然主义的观点。他认为人类具有的意识、意向性是生物进化的结果，是一种生物学现象。塞尔的生物学自然主义是一种既不同于二元论，又不同于传统的唯物主义的意识理论。他提出了一种既不违背唯物主义的基本原则，同时又能够肯定意识在世界中的地位与作用的观点，即生物学自然主义。“生物学的”是指对精神现象的解释方式是生物学的，“自然主义”指心灵是自然的一部分。

塞尔认为，唯物主义和二元论之所以陷入两难选择，根本原因在于一个共同的错误前提——“如果意识真的是主观的、质的现象，那么它就不可能是物质的物理世界的组成部分”[①]。塞尔抛弃了这些陈旧的形而上学包袱，从生物学的角度去思考心灵问题，认为“意识是和其它生物学现象一样的一种生物学现象”[②]。意识的生物学定位轻而易举地揭示了唯物主义和二元论错误的根源：唯物主义错在忽略了意识的存在，二元论错在否认意识是生物学的、自然界的组成部分。明确了这一点后，塞尔进一步阐述“具有第一人称的本体论”的意识不能还原为具有“第三人称本体论”特征的物质的事物或过程。“但这并不妨碍意识是大脑的更高层次的特征。正如消化是胃的更高层次的特征”[③]一样。这样他就完成了意识的自然化，并把这种观点称为“生物学的自然主义”。他的观点可以大致概括为：“意识是由大脑过程所引起的，它是大脑系统在更高层次上的特征”[④]。

塞尔通过生物学语言对心理现象进行分析，从而使意识和意向性得以自然化。认为意向性不是什么超验现象，它只不过是有机体大脑的一个属性。他从内在意向

[①]约翰·塞尔，《心灵、语言和社会》，李步楼译，上海：上海译文出版社，2001年，第50页。
[②]约翰·塞尔，《心灵、语言和社会》，李步楼译，上海：上海译文出版社，2001年，第51页。
[③]约翰·塞尔，《心灵、语言和社会》，李步楼译，上海：上海译文出版社，2001年，第54页。
[④]同上。

性对自然界的依赖性以及它在自然界中的地位，认为意向性作为一种现象不是存在于自然界之外，而是自然界的一个组成部分。也正是由于有了意向性，人才能够跨越时空关于那些在场或不在场、存在或不存在的事态。而意向性的这一生物学机制在当代已经被神经科学、脑科学、生物学所证实。也正是因为意识、意向性是生物进化的产物，是人脑的属性，是人类智能的基础，因此，机器要具备智能就必须也具有意向性。塞尔承认机器可以具有派生的意向性，但这种意向性并不足以让机器产生智能。在他看来，心理状态之所以能主动地指向、关联于世界上的事态，是因为它有意识，即既能有意识地去指向，同时在指向的过程中和之后，又能意识到指向了什么。意识是区分原始意向性和派生意向性的关键。塞尔的论证得到了许多人的肯定，因此成了人工智能等领域引用频率最高的经典。有的人还在此基础上从语义学的角度作了发挥，认为，现有的计算机之所以不令人满意，一点也看不出超越人脑的希望，根本原因在于：人脑不仅是句法机，而且是语义机，而计算机仅仅只是句法机。塞尔不赞成后一观点，强调：它连句法机都算不上，因为它内部真实发生的是电脉冲的转换，而并没有句法转换。句法转换是我们人为了解释计算机的行为而归属给它的。因此要使计算机具备真正的智能，就要去研究大自然，向大自然取经，进行人工进化、育种。只有这样才有可能塑造出意向性，也才有可能使我们的机器成为真正意义上的智能机。

二、人工智能的新标准——意向性

塞尔认为意向性在生物进化的结果，所以他不同意所谓“强人工智能”的观点。让我们简单回顾一下“中文屋”论证：

设想你坐在一间有两个小孔的屋子里，从一个小孔递给你一些你根本不认识的中文字符，也就是说你完全不知道这些字符的意义。但是你有一套操作规程，你根据这个操作规程可以把递给你的那些中文字符转换成另一些中文字符，然后将这些新的字符从另一个小孔递出去。简单地说，我们所进行的活动只有一下三个步骤：中文字符被送入房间——按照操作规程将输入的中文字符转换为另一些中文字符——将新的中文字符送出房间。

塞尔的“中文屋”模型，其实是一个精心设计的、带有任意程序的机器。从本质上讲，塞尔中文屋是模拟了一段计算机程序：输入一段中文字符，经过计算机运算，输出另一段中文字符。塞尔的这个思想实验是让我们自己参与了机器的操作和程序的运行，然后，让我们自己得出结论：机器到底是否具有智能。通过这个“中

文屋”，我们可以按照操作程序与屋外的人中中文交谈。现在设想测试者提问：“你能说中文吗？”即使你根本就不会中文，但你在房屋里仍然可以按照事先编好的程序回答：“我能，当然能。”尽管你的回答不过是按规程操作的结果，但房间外面的人却相信了你是一位精通中文的人，哪怕你对中文是一窍不通。由于这个中文屋模拟了一段理解中文的计算机程序（这个计算机程序就是参与试验的你手中的那一套操作规程），而该程序则模拟了通晓中文的人，所以它能够保证房屋里不懂中文的人看起来像懂中文的人。因此，这恰恰说明了计算机理解的只是我们编好的操作程序，并不具备在真正的智能，因此“强人工智能”就是不存在的。

塞尔的“中文屋”论证确实对包括图灵测试在内的已有的人工智能标准提出了挑战。按照塞尔“中文屋”，给定任意的操作规程，机器并不需要理解所处理字符的意义，而只须按照程序操作即可。这样塞尔就成功地制造了一部能完成类似人的智力行为的却不具有一点智能的机器。并且，程序的改变仅仅是操作规程的改变，这种改变丝毫也不会增加机器的智力。也正是在这一点上，塞尔的中文屋比图灵测试更有意义，它为人工智能提供了一个动态的的标准：计算机能够完成某种智能行为，仅仅是因为它执行了人们按照一定目的事先编制的程序，或者说，是人类智能决定了机器智能。

三、诘难与回应

塞尔对“图林测试”以及建立于其上的物理符号系统理论的批判，从生物自然主义的观点出发，强调人类智能具有意向性和意识之类的本质特征。塞尔的这一智能标准一经提出，就引出了很大震动，立即遭到了许多人的批评和否定，但同时也得到了许多人的喝彩和支持。

对于塞尔的诘难，“机器人”的例子与大脑模拟两个最具代表性。

（一）“机器人”诘难与塞尔的反驳

“设想我们编写了一个与尚克程序类型不同的程序，同时，我把一台计算机放入机器人中，这台计算机不只是接收形式符号作为输入，发送形式符号作为输出，它实际上是在操作机器人，通过操作，使机器人做出与感知、步行、走来走去、钉钉子、吃东西、饮水十分相像的事情——以及你希望的任何事情。例如，机器人配有一台摄像机，它就能‘看见’，有了手臂和腿，它就能‘动作’，而这一切都由计算机‘大脑’来控制。这种机器人与尚克的计算机不同，具有真正的理解力和其他

心理状态。”[①]

面对“机器人”的诘难，塞尔首先肯定了这一假设的合理性。“机器人”的例子承认了认知不仅仅是形式符号操作，而是一系列与外部世界的因果关系。然后，塞尔将中文屋的思想实验和机器人的例子结合起来，又进行了一次新的思想实验。他假设将一个微型的塞尔（微型计算机）放在机器人里，充当机器人的大脑。其后的步骤跟中文屋实验基本一致，给微型的塞尔更多的中文符号和更多的英文说明。微型塞尔从机器人的感知装置接受信息，而后将信息传达到自动转装置，机器人开始四处移动。假设微型塞尔并不知道那些被给的中文符号来自于电视摄像机，而且“我”也不知道反馈回的中文符号能够让预先安置于机器人里的自动装置移动机器的四肢，这整个过程中只存在符号的操作，则充分可以说明机器人不具有意向性。微型塞尔也不具有意向性，它的行为只是遵循有关操作形式符号的形式指令而已。

（二）大脑模拟的诘难与塞尔的反驳

大脑模拟的观点认为，“假定我们设计一个程序，它并不表述我们具有的外部世界的信息，如尚克脚本中的那些信息，但是它模拟一个讲中文母语的人在理解中文故事和对故事作出回答时在他大脑突触上的神经元激发的实际顺序。这台机器接收中文故事和有关故事的问题作为输入，模拟实际懂中文的大脑在加工这些故事时所具有的形式结构，并以中文回答作为输出。我们甚至可以设想，这台机器动作时，用的不是单一串行程序，而是并行动作的整组程序，也就是我们认为实际人类大脑在加工自然语言时采用的那种操作方式。毫无疑问，在这种情况下，我们不得不说，机器理解这些故事。如果我们拒绝承认这一点，我们是不是也要否认讲中文母语的人理解这些故事呢？在突触层次上，计算机程序和懂中文大脑的程序会有、或者能有什么差别呢？”[②]

塞尔认为强人工智能的核心观点是在不了解人脑的结构及工作原理的情况下，仍然可以模拟心的工作。假设构成计算过程的神经操作通过各种不同的大脑过程来实现，而且这些神经操作超过了组成神经本质的形式元素，以同样的方式任何计算机程序可以通过不同的计算机硬件来实现，于是心灵之于大脑相当于程序之于硬件，这样就可以在不理解神经物理学的情况下去理解、模拟心灵了。然而塞尔回答则是否定的，他认为即便是如此地贴近大脑的操作，对于理解的产生也起不到实质

[①] 约翰·塞尔，《心灵、大脑与程序》，载玛格丽特·博登：《人工智能哲学》，刘西瑞、王汉琦译，2006年，第82-83页。
[②] 约翰·塞尔，《心灵、大脑与程序》，载玛格丽特·博登：《人工智能哲学》，刘西瑞、王汉琦译，2006年，第84页。

作用。塞尔指出，问题的关键在于，这只是模拟了脑的形式结构或特点，而不是脑的因果特性和它产生意向性的能力。他是这样论证他的观点的。他将中文屋中那个只懂一种语言的人看作成一个正在操作一组以阀门相连的精巧水管的人。当这个人接收到中文符号信息时，他按照英文说明打开和关闭阀门，完成这一系列的规范动作，这就相当于讲中文的人脑中神经突触的活动。水管的连接处则相当于人的大脑突触，当所有突触正确接上后，输出的中文符号就顺着管道从终端流出。在整个过程中，塞尔认为这个系统什么也没有理解，那个人不会汉语，而水管也同样不会汉语。另外，在神经元结构层，大脑皮层约有 10^{10} 个神经元，它们以一层薄片的形式盘绕在大脑的半球上。这些问题都是神经元细胞及其结构与计算机系统的区别。因此，我认为计算机无法模拟真实的大脑，也就不足以产生意向性。

（三）塞尔的支持者——来自科学界的回应

对于塞尔的观点持肯定态度的有著名诺贝尔奖获得者克里克。他在批评强人工智观时，所得出的结论与塞尔的几乎没有区别。他说："这种比较"，即把人脑比作计算机，"如果陷入极端的话，将导致不切实际的理论"。[①]在批评功能主义时，他指出："他们认为，了解脑的细节永远得不到任何有用的东西。这一观点如此古怪，以至大多数科学家都惊讶它为什么能够存在。"[②]计算机的"智能"为什么不能与人的智能相提并论呢？他说："计算机的操作是序列式的，即一条操作接着一条操作。与此相反，脑的工作方式则通常是大规模并行的。例如从每只眼睛到达大脑的轴突大约有 100 万个，它们全都同时工作。"还有，计算机的运行高度稳定，而人脑的神经元则变化多端，甚至边"计算"边改变。另外，在计算机中，信息被编码成由 0 到 1 组成的脉冲序列。通过这种方向，它把信息从一个地方传递到另一个地方，而脑中没有这种精确性，神经元的脉冲模式可能携带信息，但不存在精确的由脉冲编码的信息。最后，计算机的软件与硬件的联系不那么紧，因为编程时，可以不考虑硬件，而人脑的软件与硬件之间没有明显的差异。[③]计算机与人类智能存在这些差别根本的原因在于，脑能自觉、有意识地处理各种外来信息，用哲学的话来说，它有其固有的意向性。

著名脑科学家埃德尔曼等人也认为，从接受的信号上看，脑能从各种变化多端的信号中归结出一些模式来，它不需要预行规定的代码就能把它们分成一些相关的

---

① 克里克：《惊人的假说》，汪云九等译，湖南科技出版社 1998 年，第 181 页。
② 同上，第 78 页。
③ 同上，第 181-183 页。

类，而计算机没有这种能力。[①]可见，他们也看到了人类智能与计算机的根本差别之所在，即一个有意向性，一个没有。许多人工智能专家也承认意向性对于智能的核心地位。斯蒂沃特·威尔逊是这一领域有影响的科学家，他说：已有的智能机器的问题在于："不会直接从周围环境中汲取所需，而只能坐在那儿，直到人们给它们信号，然后也仅仅是复制这些信号而全然不知它的意义。"[②]

另外，冈德森对符号加工和联结主义研究后，对人工智能的未来也作了探讨，并得出了与塞尔相似的结论，认为有无意向性是人类智能与人工智能区分的根本标志，今后发展的迫切任务是探讨怎样让人工智能表现出真正的意向性或语义性。除激进的取消主义之外，大多数心灵哲学家都倾向于塞尔关于智能的观点：人类智能之所以是真正的智能，根源在于它有意向性。

四、从意向性看人工智能的极限及发展路径

人工智能是20世纪中期产生的新兴综合性学科，是现代科学技术的一项卓越成就。在人工智能成就的背后，人们开始对其中的许多理论问题进行争论，对人工智能的极限问题进行深入的思考。1997年5月，IBM的"深蓝"战胜世界象棋冠军卡斯帕罗夫。人们深深感受到AI向人类智能发出的挑战，不禁想问智能模拟到底能走多远，它究竟有没有极限？人工智能至多只是部分地超越人类智能，在整体上是不能最终代替和战胜人类智能的。塞尔的"中文屋"论证可以充分地说明这一点。塞尔说："我的输入和输出，与讲中文母语的人没有区别，而且你想要任何形式的程序，我都可以有，但我仍旧什么也不理解。根据同样的理由，任何故事，无论它们是中文的，英文的，或是其他文字的，尚克的计算机一概都不理解，因为在中文的情况下，计算机就是我，在计算机不是我的情况下，与我什么都不理解的情况相比，它同样也无知。"[③]

（一）机器与人脑

机器可以代替人脑，但绝不会超越人脑。人工智能的出现是人类文明上划时代的发现，机器（计算机）极大地突破并弥补了人脑的局限性。电脑克服了人脑在接受和处理信息速度上的局限性、存储和记忆上的局限性、意识活动过程中可靠性和精确性上的局限性，具有惊人的"记忆力"、敏捷的运算速度、精确的逻辑推理能

① 埃德尔曼、托诺尼著：《意识的宇宙》，顾凡及译，上海科技出版社2004年，第55页。
② 转引自弗里德曼：《制脑者》，第10页。
③塞尔：《心灵、大脑与程序》，载玛格丽特·博登：《人工智能哲学》，第95-96页。

力，有力地延伸和扩展了人脑的意识活动功能。计算机甚至可以代替和超越了人类的部分思维能力，使人类可以从那些机械性的、重复性的脑力劳动中解放出来，去进行更加富有创造性的思维活动。

那是不是由此就可以认为电脑可以完全代替甚至超过人脑，或认为将来的世界会是机器人统治的世界？从辩证唯物主义观点看来，机器能替代人(在某些时候)，但不能超越人。首先，人类制造工具，就是为了使自身得到解放。从产生那天起，人类就会制造比人的自然器官强得多的工具，这些工具的使用反过来又加强人的自然器官。其次，人工智能只能对人的思维过程中那些可以形式化的部分进行模拟，这些模拟只是人的思维活动的手段和工具，不能完全代替人脑的功能。因此，从整体上看，电脑永远不如人脑。因为“机器思维”与人脑思维、人工智能与人类智能存在着本质的区别。机器思维是人们利用电子管、晶体管、集成电路等电气元件和线路所组成的机械的、物理的装置，并用软件方法等模拟人的思维活动。机器思维不是人类的精神活动，而是纯属无意识的机械的物理的过程。人类意识是物质世界长期发展进化的产物，是人类在生理基础上的心理过程，是由人类的情感、直觉、想象、猜测等心理活动所构成的精神世界。人在行动时要考虑到由此引起的这样或那样的效果，机器智能只执行特定的指令，并不探求任务本身的社会意义，不会考虑到社会后果，即不具有社会性和创造性。人类意识是社会的产物，具有社会性，创造性。机器智能也不具有主体性，它只能被动地接受人类的“指令”，在程序的指引下工作。两者之间，总是人脑的思维在前，电脑的功能在后。“机器思维”与人脑思维、人工智能与人类意识的这些区别，在实践中已得到充分证明。

（二）意向性鸿沟

在塞尔看来，人类的智能之所以是真正的智能，根本原因在于它有意向性。智能不在于形式转换，而在于能处理内容或有意向性。他说：“大脑产生意向性的那种因果能力，并不存在于它例示计算机程序的过程中，因为无论你想要什么程序，都能够由某种东西来例示这个程序，而它并不具有任何心理状态。无论大脑在产生意向性时所做的是什么，都不可能存在于例示程序的过程中，因为没有一个程序凭借自身而对于意向性来说是充分的。”[①]

意向性对于智能来说为什么如此重要？正如塞尔在前面曾说过的，意向性就是心理状态对它之外的对象的指向性、关于性。例如人谈到某件事，说出某个语词，能在谈和说的同时，与世界建立了一种关联，并自己意识到了被关联的事物。机器

[①]塞尔：《心灵、大脑与程序》，载玛格丽特·博登：《人工智能哲学》，第 119 页。

就不一样了，它只能进行符号转换，什么也不能“关于”，什么也不指涉，它的工作就是一种纯形式转换。人类正是由于心理状态的这种属性，它才能置之身外，把自身内部所作的“运算”、加工与有关的对象关联起来。人才成了人，心灵才成了心灵，人和心灵才能有对世界的认识、利用和改造；也正是有了这种关联，世界及其发生的变化才能起到影响和改变人的反作用。

有人会说，其他的事物之间也有相互作用，但那都是不自觉的，没有心灵的主动参与，而人及其心灵与世界的关联由于心灵的意向性这种特殊本质，才能以主动、自觉的形式进行。从信息论的观点看，好像也能按人的方式与世界发生关系，好像与人没有多大区别。例如我想知道一组数据相加的总和是多少，我只需要将这些数据输入计算机，它经过运算计算机可以既准确又迅速地告诉我答案。从表面上看，这计算结果是关于这一组数据的。其实不然，它完全没有这种关于性。它的这种关联是我“加给”或“归属”给它的。计算机处理的始终是纯形式的是符号或句法，它不知道把它与指称或外物关联起来。人的计算尽管没有它准确和迅速，但人在计算时，从始至终都是关联着计算过程之外的实在的。由上可知，人类的智能之所以是真实的智能，根源在于一开始就有能关涉世界的能力，而计算机智能之所以是比喻性或仿佛性的，就是它没有这种与世界主动关联的功能，因此它永远只能是一个死板的工具，而不能成为活生生的智能体。造成这种差异的原因：就是一个有意向性，一个没有。

1、意向性——生物进化的结果

在意向性的起源中我们已经说过，意向性不是什么超时空的、非存在的东西，也不是什么副现象，意向性是真实存在的，是我们自然生物进化的结果，是自然生物生命史的组成部分，是和呼吸、消化、胆汁分泌一样发生在生物机体内的真实的现象。要想建构真正的人工智能，就得设法让它具有真正的意向性。既然塞尔说意向性是生命史的组成部分，那么我们就应该向大自然这位设计师和缔造者学习，研究它创造人类智能的历史过程。虽然塞尔没有明确指出一条怎么向大自然学习之路，但从中我们可以看出：要想获得类人的意向性就要向大自然学习，要进行自然选择与进化。而这也正是人工智能的极限所在。

2、人与非人的意向性——真实与隐喻的区别

塞尔从意向对象和意向活动之间的归属问题探讨了原始的意向性和派生的意向性。那么，计算机有无意向性呢？塞尔的答案是：有，但是派生的意向性并不是真实的意向性。通过对意向性的解析，我们知道，意向性是内在的，而派生的、隐喻的意向性都是由内在的意向性而产生的。我们平时所说“我的花要渴死了”“我

的相机可以记忆美好的时光”的这些意向性都是派生的、隐喻的意向性，而人工智能要成为真正的智能只具有这类意向性是不足够的，计算机只有具备了原始的、内在的意向性才有可能成为真正的智能体。“说某种实体具有‘好像’意向性不过是一种表述方式，说明它的行为表现得‘好像’具有意向性，而实际上并没有意向性。”“一切派生的意向性都是从内在的意向性派生出来的。着重强调这一点很重要，因为在当代有许多著作家把派生的意向性和‘好像’的意向性当作范式来看待，并企图按照它们来解释一切内在的意向性。这样，电脑操作这种派生的意向性被当作是研究人脑中的内在意向性的模型，而‘好像’的意向性归属有时竟被当作理解我们把内在意向性归于人时所作的那种的正确模型。”[①]由此我们可以看出，只具有派生、隐喻意向性的机器是不可能有我们所讨论的真正的智能的。如果我们不能把内在的和派生的意向性区别开来，那么我们所讨论的人工智能就会出现误解，从而得出一种强人工智能的观点。

3、意向性的基础——意识（主观能动性）

塞尔认为，意向性的基础是意识。而主观能动性是人的意识的独有特征，是人具有能动的认识世界和改造世界的能力。按照马克思主义的观点，人脑对外界事物的反映不是机械地、呆板地，而是积极地、能动地反映，它能把大量感性材料升华为理性认识，在思维中再现事物的本质和规律。人类智能的本质特征是能动性、创造性等不确定性思维。相比之下，计算机只会按规则行事，它总是机械地、精确地执行逻辑运算指令。即便是所谓的智能计算机，运算速度再快，也是一样。即便现代，甚至将来，计算机具有的自动性，并且现代很多的研究也正朝着这一方向发展，也取得了一定的成果，但计算机的这种自动性不是也不可能是计算机本身的主观意愿的能动性，而只不过是人的主观能动性的机器表现形式。

从上面的解析，我们可以看出：机器要想有真正的智能，就必须向大自然学习，必须全面揭示人脑结构、功能和活动规律。只有这样，才有可能模拟产生意向性的能力。目前的脑科学，特别是在认知脑科学领域，已经成为一个研究的热点问题之一。而在生命科学取得进展的同时，工程界及时地吸取这些成果，将其融入工程学领域的设计思想之中，就有可能由简单到复杂地设计出逐步逼近人脑特征的人工智能系统，但这种系统也只能是一个工具而已。目前的计算机系统，虽然具有一定的人脑功能特点，而且数字运算速度远远超过人脑，像“深蓝”这样的博弈机，能战胜国际象棋大师。但是，要想使它具有真正的智能，就要具有人脑的开放性、有序

[①]约翰·塞尔，《心灵、语言和社会》，李步楼译，上海：上海译文出版社，2006年，第92页。

性、层次性、可塑性、自组织能力和协同性等特点。而这些特点，在塞尔看来都是源自于自然生物进化的结果。所以人工智能的发展不仅要有脑科学的突破性进展，更重要的是要用自然进化的理论去指导工程学中的人工进化的方法。

再者，前面已经说过，人工智能的发展问题其实还在于我们如何去看待人工智能。需要是发明的原动力，人工智能也不例外。正因为我们有解放人类于更繁重工作的需要，所以我们才朝着既定的方向努力。但是，现在很多人看到了人工智能的极限，认为人工智能不能模仿人类的喜怒哀乐，不能像人类一样产生爱、恨等思想感情，更不会有复杂的心理状态就否定了人工智能的未来，这也是不正确的。我们要看到，人类社会需要的是能将人类从很多复杂的思维活动中解脱出来的工具，而不是一个与人类平起平坐的异类。目前人工智能工具在越来越多的领域发挥着重要的作用。我们有什么理由因为人工智能工具不具有喜怒哀乐而抱有悲观主义观点呢，难道只有一个经常发脾气、会罢工的智能机器才是我们所需要的吗？由此可见，对于人工智能（且用人工智能这一词语，实则为机器智能）的发展，我们要有一种乐观主义（不是文中所提到的强 AI 的乐观主义）的态度，即：在模仿人类智能方面，从工程学角度来说进展的空间还很大，我们也需要这样的智能系统工具来不断解放生产力，不断地解放人类自身。

参考文献

[1]高新民. 现代西方心灵哲学[M]. 武汉:武汉出版社, 1996
[2]高新民, 储昭华. 心灵哲学[M]. 北京:商务印书馆, 2002
[3]高新民. 人自身的宇宙之谜[M]. 武汉:华中师范大学出版社, 1989
[4]高新民, 刘占峰. 心灵的解构[M]. 北京:中国社会科学出版社, 2005
[5]高新民, 殷筱. 当代心灵哲学导论[M]. 北京:中国人民大学出版社, 2005
[6]高新民. 人心与人生[M]. 北京:北京大学出版社, 2006
[7]高新民. 意向性理论的当代发展[M]. 北京:中国社会科学出版社, 2008
[8][美]塞尔. 意向性[M]. 刘叶涛译. 上海:上海人民出版社, 2007
[9][美]塞尔. 心、脑与科学[M]. 杨音莱译. 上海:上海译文出版社, 2006
[10][美]塞尔. 心灵、语言和社会[M]. 李步楼译. 上海:上海译文出版社, 2006
[11][美]塞尔. 自由与神经生物学[M]. 刘敏译. 北京:中国人民大学出版社, 2005
[12][美]塞尔. 心灵的再发现[M]. 王巍译. 北京:北京人民大学出版社, 2005
[12][美]霍夫施塔特, 丹尼特. 心我论[M]. 陈鲁明译. 上海:上海译文出版社, 1988
[13][美]德雷福斯. 计算机不能做什么[M]. 宁春岩译. 北京:三联书店, 1986
[14][英]玛格丽特. 博登. 人工智能哲学[M]. 刘西瑞, 王汉琦译. 上海:上海译文出版社, 2001
[15][美]霍金斯, 布拉克斯莉. 人工智能的未来[M]. 贺俊杰, 李若子, 杨倩译. 陕西:陕西科学技术出版社, 2006
[16]埃德尔曼, 托诺尼. 意识的宇宙[M]. 顾凡及译. 上海:上海科技出版社, 2004
[17]克里克. 惊人的假说[M]. 汪云九等译. 湖南:湖南科技出版社, 1998
[18]卡尔文. 大脑如何思维[M]. 杨雄里等译. 上海:上海科技出版社, 1996
[19]彭罗斯. 皇帝新脑[M]. 许明贤, 吴忠超译. 湖南:湖南科技出版社, 1994
[20]冯天瑾. 智能学简史[M]. 北京: 科学出版社, 2007
[21]倪梁康. 面对实事本身[M]. 北京:东方出版社, 2000
[22]唐热风. 心身世界[M]. 北京:首都师范大学出版社, 2001
[23]车铭洲. 现代西方语言哲学[M]. 成都:四川人民出版社, 1989
[24]高新民. 意向性. 意义. 内容—当代西方心灵哲学围绕心理内容的争论及其

思考[J]. 哲学动态, 2003(2)
[25]刘景钊. 理解心智—塞尔意向性理论意义[J]. 光明日报, 2005(3)
[26]刘晓力. 哥德尔对心—脑—计算机问题的解[J]. 自然辩证法研究, 1999(11)
[27]毛华滨, 高新民. 从意向性看人工智能的发展方向[J]. 科学技术与辩证法, 2005(10)
[28]李亚宁. 关于人工智能极限研究的哲学问题[J]. 四川大学学报, 1999(6)
[29]赵泽林, 高新民. 计算主义在心灵哲学中的两大前沿论题评析[J]. 科学技术与辩证法, 2007(8)
[30]费多益. 人工意识是否可能[J]. 自然辩证法研究. 2005(7)
[31]赵泽林, 高新民. 人工智能的瓶颈问题与本体论语义学的回应[J]. 哲学动态, 2007(6)
[32]郑祥福. 人工智能的四大哲学问题[J]. 科学技术与辩证法, 2005(10)
[33]钱铁云. 人工智能是否可以超越人类智能[J]. 科学技术与辩证法, 2004(10)
[34]刘西瑞, 王汉琦. 人工智能与意向性问题[J]. 自然辩证法研究, 2001(12)
[35]郭斌. 塞尔“中文屋”思想实验的哲学意蕴辨析[J]. 自然辩证法研究, 2005(12)
[36]唐热风. 心的本质是计算吗?[J]. 自然辩证法研究, 1998[4]
[37]蔡曙山. 哲学家如何理解人工智能——塞尔的“中文房间争论”及其意义[J]. 自然辩证法研究, 2001[11]
[38]蔡曙山. 关于哲学、心理学和认知科学的12个问题———与约翰·塞尔教授的对话[J]. 学术界, 2007(3)
[39]陈世锋. 生物学的自然主义——塞尔论心身关系[J]. 新疆社会科学, 2005(5)
[40]张云台. 人工智能及其前景的哲学思考[J]. 科学技术与辩证法, 1995(12)
[41] 郭贵春, 郝宁湘. 丘奇—图灵论点与人类认知能力和极限[J]. 齐鲁学刊, 2004(5)