

沉默的艺术

李剑锋

著

$$D=|M\rangle\otimes|W\rangle$$

意识科学基础理论

了解“你”是谁 /

了解你心智的奥秘是什么 /

知道什么样的系统可以产生意识 /

大脑如何产生意识和自由意志 /

未来的人工智能会有意识吗 /

如何用意识科学理论改进人工智能 /

如何走出第二困境 /

玛丽黑白屋 /

拒绝含糊——
——拒绝纯粹思辨



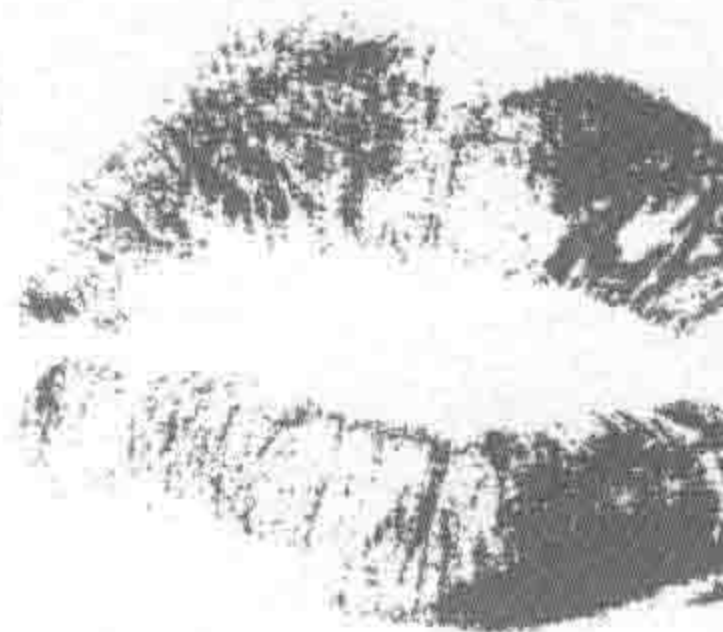
社会科学文献出版社
SOCIAL SCIENCES ACADEMIC PRESS (CHINA)

非外借

沉默的艺术

意识科学基础理论

$$D=|M\rangle\otimes|W\rangle$$



李剑锋

著



社会科学文献出版社
SOCIAL SCIENCES ACADEMIC PRESS (CHINA)

图书在版编目(CIP)数据

沉默的艺术：意识科学基础理论 / 李剑锋著. --
北京：社会科学文献出版社，2018.8

ISBN 978 - 7 - 5201 - 2979 - 4

I. ①沉… II. ①李… III. ①意识论 IV.

①B016.98

中国版本图书馆 CIP 数据核字 (2018) 第 141946 号

沉默的艺术

——意识科学基础理论

著 者 / 李剑锋

出 版 人 / 谢寿光

项目统筹 / 袁清湘 张博群

责任编辑 / 赵怀英 张馨月

出 版 / 社会科学文献出版社·独立编辑工作室

(010) 59367202

地址：北京市北三环中路甲 29 号院华龙大厦

邮编：100029

网址：www.ssap.com.cn

发 行 / 市场营销中心 (010) 59367081 59367018

印 装 / 三河市东方印刷有限公司

规 格 / 开 本：880mm × 1168mm 1/32

印 张：8 字 数：151 千字

版 次 / 2018 年 8 月第 1 版 2018 年 8 月第 1 次印刷

书 号 / ISBN 978 - 7 - 5201 - 2979 - 4

定 价 / 59.00 元

本书如有印装质量问题，请与读者服务中心 (010 - 59367028) 联系

 版权所有 翻印必究

前 言

1

中国经过近四十年的改革开放，政治、经济、科技皆取得了巨大的进步，但创新能力仍然不足。因此，近年来国家一直号召鼓励大家勇于开拓创新，特别鼓励在新的领域大胆提出自己的见解与理论；当然，这里有一个前提，提出的科学理论必须可证伪。

2

2016 年，笔者参加一个国际大脑意识科学大会，会上许多学者认为在未来二十年里意识研究将会完成从哲学领域转向科学领域的大转身。当时笔者感到既兴奋又沮丧。

感到兴奋，是因为此情境让我联想到了三百多年前的那次激动人心的伟大转身，在伽利略、牛顿等人的努力下，物理学完成了从哲学领域到科学领域的大转身。那次转身的完成是以牛顿提出他的力学理论为



标志的。同样，这次意识研究的大转身必然也会以一个大家都认同的、能够经得起实验检验的关于意识的科学理论的提出作为标志。这是一个几百年一遇的大好机会。聪明的人都应该早早地走上台桌，等待着分一份蛋糕。

感到沮丧，是因为当时发现，立在桌子周围的大都是外国人，一千多个参会者中只有三四个中国人。这是因为过去二三十年里，无论国内国外，沾上“意识”，就会被贴上类似伪科学的标签，这种偏见令许多学者只能望而却步，只是最近，国外这种偏见略微少了些。

或许，未来将要发生的这次大转身是中国人在科学领域一展身手的最后机会；虽说科学无国界，但笔者仍希望我们后辈的教科书里能多些中国人的名字。

3

20 世纪人类最大的科学贡献或许就是建立了量子力学。量子力学相当神奇，一方面它是人类目前能得到的最精确的理论（比如其对精细结构常数理论的预测极为精确）；另一方面它也是目前仍被一些物理学家认为没有坚实基础的理论。

虽然，以前大家都认为量子力学只要具备强大精准的理论预测能力就足够好了，并认为这已经是物理的全部，以至于有科学家提出我们应该“shut up and

calculate”（不争辩埋头算），且得到了广泛的认同；但近些年越来越多的物理学家开始不太认同此实用主义哲学了，并投身于 Quantum Foundation（量子基础）领域。

有些科学家比较乐观，认为诸如退干涉理论和关系量子力学或许能为量子力学建立一个大家都能接受的基础，而且这些理论不需要将观测者真正考虑进来。然而，也有科学家对能否真正为量子力学找到理论基础感到悲观，他们认为若不预先搞清楚观测者的性质，或许无法真正建立量子力学的基础，即需要先建立一个关于观测者的基础理论，当然或许观测者的基础理论本身也是量子力学的基础。

可惜，物理学家除非万不得已，不会去触碰观测者本身，因为他们明白，所谓观测者的理论其实就是一个关于意识如何产生的理论。当然，这或许也是为什么量子力学建立快一百年了，这个数学结构并不复杂的物理理论，却一直找不到合适基础的原因。

那为什么大家会如此害怕去触碰意识这个概念呢？

4

意识或许是人类迄今为止遇到的最为诡异的一个概念：它既熟悉又陌生，离我们最近却又最远；此外，它还是一个既古老又新颖的研究课题。



在任何一个时代，总会有种魔力吸引当时人类最聪明的天才去思考意识的本质，现代亦不例外。比如20世纪，即便许多正统科学家对意识研究下了驱逐令，然而，还是有许多物理天才不顾禁令，“偷偷地”思索其本质。

然而，意识真正的神奇之处在于，即便这个星球古往今来大多天才都提出过各式方案、思路试图理解它，但似乎它有着一层神奇无比的保护层，令人们至今无法真正接近其本质。

似乎，每个时代的天才在意识这个对象面前都在不断地重复着同一轮思考过程：开始时觉得这个问题很有意思，觉得它似乎是可以理解的，觉得自己只要再努力多一点点，就能触碰到它的本质。然后不断地思考各种可能性，试图解开什么是“自由意志”，什么是“自我”的秘密。最后，还觉得自己似乎到达了终点，但在冷静的时候，却发现自己好像仍然在原点，并没有比古人前进多少。

就如梦魇般。

这里，似乎有一股无形的力量阻碍着人们在意识方面的思考取得进展，或许大自然为人类事先就准备好了一个巨大的思维陷阱，历史上和现代无数的天才都掉进了这个陷阱里，以至于有人怀疑这可能是大自然精心策划的阴谋。

事实上，之前几乎所有的科学家在思考关于意识的问题时都掉进了坑里，而哲学家的情形反而要好

点。大约有一半多一点的哲学家掉进了坑里，四分之一似乎觉察到了坑的存在并试图爬出来，而最后四分之一的哲学家则试图找到坑的位置，想在坑旁边立个牌子，提醒别人。

这最后四分之一的哲学家中，最杰出的一个代表就是查尔默斯，他明白无误地对意识科学中的问题进行了划分，认为在这其中存在简单问题和困难问题两种。简言之，简单问题是意识相关问题中可归结为解释结构与功能或可归结为事物间关系的问题，困难问题是意识相关问题、不能最终归结解释结构与功能的问题。

事实上，之前的天才们也或多或少感觉到了困难问题的存在，但是由于简单问题与困难问题之间存在一类诸如“什么是自我”这样的伪困难问题——这是真正的坑。这些伪困难问题极易被天才们误作为意识研究的核心问题。因此，天才们非常容易滑入此坑去解释“自我指称”这类难题，更致命的是这类伪困难问题或多或少与数学、计算机领域的著名悖论或著名命题有关联。因此，就会有許多天才们试图用罗素悖论、哥德尔不完备定理、停机问题等来启发，以理解意识的本质。当然，最后没人真正成功。

而当查尔默斯在 1995 年明白无误地将简单问题和困难问题间的界线划出来之后，就可以让新的研究者不至于不断地滑入这个陷阱里出不来，或浪费大量的时间与生命在这些伪困难问题上面。



知道问题所在，并不代表能够真正解决问题。本书试图直接提出理论解决这个问题。因此，人们也许会问：“为什么你认为你就不会掉进这个坑里而且还能真正解决这个问题呢？”

首先，这得益于查尔默斯的工作，使得笔者自从2009年之后就不再在伪困难问题上浪费任何时间与精力，而是直接面对困难问题。

其次，笔者在物理学里找到了描述不可结构化事物的物理学语言，即找到了谈论困难问题的方法。谈论困难问题的难点在于：困难问题都是不可结构化的，而物理学似乎又只能研究可结构化的问题，但现在又要求用物理学来解决困难问题，这就是意识科学的第一大困境。走出第一困境的关键是要学会维特根斯坦的沉默，他说：“凡是能够说的，都能够说清楚；凡是不能谈论的，就应保持沉默。”即能结构化的都能说清楚，不能结构化的就应该沉默。

但沉默不代表什么都不做，关键是如何沉默。这可参考康托在构建集合论时的做法，当时他用到了强大的理论工具：映射。或许我们也可用映射来处理沉默，将不可结构部分映射到其他事物，只要不去描述它即可。而物理学中量子态既包含了可结构化的部分也包含了不可结构化的部分，某量子态在某组基底量子态或本征态下的展开表达是可结构化的，但某量子态自身的本质是什么是不可言说，不可结构化，但可将其映射到某一意识体验上的。这种处理为走出第一

困境提供了一个突破口，这也是我们仍然能用物理学研究意识的根本原因。

能够谈论它，仍然不代表可真正解决它，况且目前多数学者认为在了解意识的本质或建立意识相关的理论之前，必须首先完全理解大脑工作原理。因此可以想象，若有人在脑科学还没真正开始时就宣称已建立了意识的基础理论，他们认为这不可信。

下面，笔者从两个方面来扭转人们这方面的成见。

首先，大脑的工作原理跟意识的基础理论并不是同一层面的原理。可以想象，外星人也可以有某种智能与意识，他们大脑的原理可能就跟我们的不一样，然而，他们意识产生的原理应该跟地球上意识产生的原理基本一致。这就好像光学原理只有一个，但具体制造天文望远镜的方案有好几种。一个功底深厚的物理学家会对着一台复杂的天文望远镜说，这台望远镜详细的工作原理可能我不懂，但它所用的基本光学原理我非常清楚。

其次，许多学者都认同意识问题相当根本，而通常物理学里有个信念，越根本之事物，其物理逻辑越简单（虽然数学操作可能极复杂）。因此坚持这个信念，找到意识与普通事物之间的物理逻辑，就有可能构建意识或观测者的基本理论。就好像爱因斯坦也相信越根本的事物其物理逻辑越简单，他认为时空和能量（物质）都是非常根本的物理，他猜测时空与能

量间物理逻辑应该很简单，即能量（物质）会影响时空的弯曲，而时空会反过来影响物质的运动，然后他就此提出了广义相对论，解释了引力的起源及预测了引力波的存在。

这里需要特别指出的是，本书提出的意识的理论虽大量涉及量子理论，但它与以往的量子意识有根本区别。许多人或许以为本书的理论是另一个量子意识理论，其实不然。一方面之所以起用量子力学不是因为纠缠或波的坍塌，而是因为量子态本身可用于描述不可结构化的体验，之前没有一个量子意识理论注意到这点。另外，之前的量子意识理论从未真正阐述清楚意识体验具体由何产生，但本书的理论非常明确地指出它只能由某种自由度巨大的基本粒子的内在属性或现象属性产生。注意，本书的意识理论认为意识体验并不是一种量子效应，之前的理论似乎都将意识体验归结成一种量子效应、归结成一种物理过程，而笔者认为结构化的物理过程断然不能产生体验本身。最后，在本书的意识理论中，量子力学所起的作用是辅助性的，并不占主导。

下面简要介绍本书主线及各章节的主要内容。

本书主线：先重点介绍阻碍意识研究向科学领域进发的两大障碍，它们分别跟意识科学的困难问题和组合问题相关。然后，再提出三个层面的理论，用于扫除两大障碍，从而为将来的意识科学提供一个坚实的理论基础或一个候选的理论框架。

第一章引言通过总结人类思想史的方式引出意识研究，然后探讨现代科学研究的困境，指出它们的困境和意识研究的困境或许都同时指向同一类悖论。这类悖论本质上揭示了构造意识科学基础理论的普遍原则：由于属性只能通过系统间的关系或系统配对来体现，因此若一个系统包含了所有，它便没有属性。在新的哲学法则下，这一类悖论将自动消解。

第二章将详细介绍什么是意识科学中的困难问题，其定义、起源以及其他学者对困难问题的一些反应——其他学者如何解决困难问题。还会给出若干不同学派对此类问题的整体看法，以及他们如何面对困难问题。

第三章主要介绍意识科学中的组合问题，即普通物质碎片化的体验属性如何巧妙地组合成高度整合的意识体验？

以上为第一部分，第二部分是本书的主体部分。

第四章将给出意识科学一个普遍性的理论，这个普遍性理论从几个简单不容易错的原则出发，借助量子力学的描述语言展开，然后用最大信息获取原则来定义时间和观测者本身，同时也精确地界定了何为体验属性，何为物理属性。它也给出了如何从本理论还原到经典物理学的基本方略，以及目前此理论还有哪些重要的问题急需解决的答案。最后，还在此理论基础上，用计算机模拟构建了一个非常简单的宇宙模型，并尝试通过分裂它来得到可体验时间流逝的观测



者和相应意识主体。注意，或许这是人类首次模拟意识产生相关的逻辑结构（注意，是模拟这个逻辑结构，而不是模拟意识本身），这个模拟计算可让读者更为具象地理解笔者提出的意识理论。这一章主要解决了困难问题，使得我们能够从理论上定义和谈论意识体验。笔者不得不在此提醒读者，由于第四章内容较艰深，易让人困顿；但笔者仍希望读者能努力读懂它，因为它是本书核心。

在第四章中，时空一开始并不存在；而在第五章，将假设已经存在时间、空间和普通物质。然后在此基础上，探讨大脑产生意识的可能机理，探讨大脑如何巧妙地运用暂时全同粒子法则，捕获和稳定自由度极高的基本粒子，如何通过基本粒子的体验属性产生意识体验来感受世界万物。若此理论预测被实验证明，那么它便基本解决了意识科学中的组合问题：在此理论中，内在自由度巨大的基本粒子的体验属性本身就是意识，因此无需组合；另外，大脑通过调节前置系统的物理属性，借助全同粒子法则间接地调节意识主体的物理属性，调节它的内在属性，使得意识主体可体验到外面的客观世界。即解决组合问题的关键在于要认识到毋须解决组合问题。此理论是第四章普遍性理论的一个具体化形式，或一个实现方式。

但第五章的理论仍然不够具体，它只回应了意识如何可能的问题，没有回答意识主体如何具体与神经系统作用从而产生体验的问题，以及它能否影响控制

外界和是否有自由意志等问题。第六章的理论将结合深度学习或人工智能的一些研究成果来初步探讨这个问题。本章将提出一个大脑的最简约模型，这个模型可以模拟一个深度学习网络加上意识的影响，意识将在其中起到协调作用。我们将通过模拟计算，运用识别手写数字来说明意识对这个模型的作用。事实上，第六章部分地回应了“意识的生物学基础是什么？”的问题。

最后一章将讨论一些验证本书的理论预测的实验方案。笔者非常希望从事相关专业的读者读了此书之后能够去验证这些理论预测。

最后，讨论本书的使用方法。

大家可以先大致浏览一下目录与附录，再从第一章开始看起。

喜欢哲学的朋友，可将注意力集中于第二、三、四章，侧重理清其中逻辑。特别是第四章的内容可能会对你们以后的思考有所启发。

喜欢物理学的朋友，可将注意力集中于第四章和第五章。其中若能体会到第四章的理论的妙处，你会发现里面有许多定义得非常好的物理问题值得探究。而第五章提出的大胆的“暂时全同粒子猜想”也可能会触动您的物理神经，这猜想对吗？如果对的话，那么它不就可以同时解决制造量子计算机、捕获暗物质的难题了吗？

喜欢生物学或来自神经科学领域的朋友，可将注

意力集中于第六章，里面提出了意识的可能的生物起源，你们或许可从中得到启发，尝试在以后的研究工作中，采用其中提到的思路思考问题，甚至尝试用实验去验证本章的理论假设及预测。

同样，对于来自于人工智能领域和计算机领域的朋友们，你可以在第六章找到你们熟悉的深度学习的代码，或许你也可以在以后的训练学习中尝试将意识的影响加进来。

.....

总之，真诚地希望来自不同领域、不同年龄段的朋友们都能从本书中获益。

在此，特别感谢我博士期间的导师杨玉良院士，感谢他的包容和对我研究的支持。诚挚地感谢联科集团的帮助。感谢我家人对我研究的默默的支持。感谢微博上各位网友，与你们的讨论让我获益良多。

目 录

前 言 / 1

第一章 意识科学导论 / 1

- 1.1 人类科学思想简史 / 1
- 1.2 科学的尽头? / 7
- 1.3 问题的症结 / 15
- 1.4 转向科学 / 19

第二章 意识科学的第一大困境 / 24

- 2.1 困难问题 / 25
- 2.2 物理学的局限性 / 27
- 2.3 意识科学第一大困境及可能的解决方案 / 31
- 2.4 玛丽黑白屋 / 33
- 2.5 各方对待困难问题的观点 / 35

第三章 意识科学的第二大困境 / 44

- 3.1 充分理由原则与体验属性的普遍性 / 45
- 3.2 詹姆斯的组合问题 / 51
- 3.3 内在属性的不能组合性与
图灵测试的局限性 / 53
- 3.4 如何走出第二困境 / 58



第四章 意识科学的普遍性理论 / 62

- 4.1 无时空量子观测者理论的简介 / 63
- 4.2 基本原则 / 71
- 4.3 背景时空问题 / 75
- 4.4 无时空的观测者理论 / 81
- 4.5 重构物理学 / 98
- 4.6 体验和现象属性的数学定义 / 104
- 4.7 显含观测者模型宇宙的模拟计算 / 106
- 4.8 待解决的问题 / 113

第五章 大脑的基本物理模型 / 115

- 5.1 意识主体不可分猜想 / 116
- 5.2 如何稳定意识主体与暂时全同粒子法则 / 119
- 5.3 大脑的物理模型 / 133
- 5.4 意识的进化 / 138
- 5.5 意识相关的计量学 / 142

第六章 意识的生物学基础 / 151

- 6.1 大脑结构 / 151
- 6.2 意识的生物学基础 / 158
- 6.3 意识理论增强机器学习 / 166
- 6.4 意识协调的神经网络 / 170
- 6.5 再论自由意志 / 184
- 6.6 意识科学和人工智能的未来 / 186