

# 机器人叛乱

## 在达尔文时代 找到意义

THE ROBOT'S REBELLION  
FINDING MEANING IN THE AGE OF DARWIN

[加] 基思·斯坦诺维奇 (Keith E. Stanovich) 著

吴宝沛 译

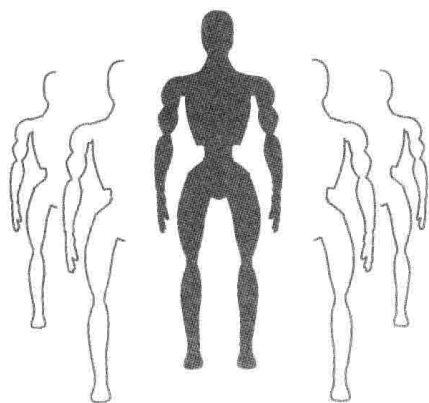
你是载体，是机器人，是不朽的基因  
和肮脏的模因复制自身的工具。

美国心理学会终身成就奖获得者 基思·斯坦诺维奇，用认知科学和决策科学  
铸成一把理性思维之剑，引领全人类，开启一场反抗基因和模因的叛乱。

如果《自私的基因》击碎了你的心和尊严，《机器人叛乱》将帮你找回自身存在的价值和意义。



机械工业出版社  
China Machine Press



# 机器人叛乱

## 在达尔文时代 找到意义

[加] 基思·斯坦诺维奇 (Keith E. Stanovich) 著

吴宝沛 译



机械工业出版社  
China Machine Press

## 图书在版编目 (CIP) 数据

机器人叛乱：在达尔文时代找到意义 / (加) 斯坦诺维奇 (Stanovich, K. E.) 著；  
吴宝沛译. —北京：机械工业出版社，2015.5

书名原文：The Robot's Rebellion: Finding Meaning in the Age of Darwin

ISBN 978-7-111-50179-4

I. 机… II. ①斯… ②吴… III. 思维心理学—研究 IV. B842.5

中国版本图书馆 CIP 数据核字 (2015) 第 097352 号

---

本书版权登记号：图字：01-2015-0179

Keith E. Stanovich . The Robot's Rebellion: Finding Meaning in the Age of Darwin.

Copyright © 2004 by Keith E. Stanovich.

Simplified Chinese Translation Copyright © 2015 by China Machine Press.

Simplified Chinese translation rights arranged with Keith E. Stanovich through Andrew Nurnberg Associates International Ltd. This edition is authorized for sale in the People's Republic of China only, excluding Hong Kong, Macao SAR and Taiwan.

No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system, without permission, in writing, from the publisher.

All rights reserved.

本书中文简体字版由 Keith E. Stanovich 通过 Andrew Nurnberg Associates International Ltd. 授权机械工业出版社在中华人民共和国境内（不包括中国香港、澳门特别行政区及中国台湾地区）独家出版发行。未经出版者书面许可，不得以任何方式抄袭、复制或节录本书中的任何部分。

## 机器人叛乱：在达尔文时代找到意义

---

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）

责任编辑：戚 妍

责任校对：殷 虹

印 刷：北京盛兰兄弟印刷装订有限公司

版 次：2015 年 6 月第 1 版第 1 次印刷

开 本：170mm×242mm 1/16

印 张：20.5

书 号：ISBN 978-7-111-50179-4

定 价：69.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：(010) 68995261 88361066

投稿热线：(010) 88379007

购书热线：(010) 68326294 88379649 68995259

读者信箱：hzjg@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问：北京大成律师事务所 韩光 / 邹晓东

推

The Robot's Rebellion  
Finding Meaning  
in the Age of Darwin

荐  
序

一<sup>⊖</sup>

**时间：2404 年。**

还有 16 年，你即将完成你的使命。你推开窗，看着冬眠医院外渐渐败落的城市。雾霾渐起，你的思绪回到 380 年前的那天。

那时，就像那个时代富人常常会做出的选择一样，你的主人决定将自己冬眠起来，一直到 400 年后。主人们可以“像植物一样活着”，将自己的身体封装在胶囊里面，一动不动，然后任由漫长的岁月腐蚀胶囊。当然，明智的主人们知道，“像植物一样活着”难以帮助自己度过冬眠期间的 400 年。胶囊的电源断了怎么办？医院被损坏了，需要搬动胶囊怎么办？主人们纷纷决定“像动物一样活着”，就是让机器人们来照顾自己。于是，有了你的诞生。

你是那一批机器人中毫不起眼、默默无闻的一个。

---

⊖ 此文摘自阳志平老师尚未正式出版的《心智三部曲》第一本的第三个故事，原文写作于 2013 年 1 月。

主人们制造了你们。你的终极使命就是帮助主人度过 400 年的冬眠期。为了完成这一终极使命，主人在程序中赋予了你小小的自由——你可以搬动主人胶囊；当封装主人身体的胶囊电源快用完了，你可以采取其他方法去获取电源。

你就像一条狗一样，你的终极使命就是那条短而有力的狗绳；你的小小的自由就是那条长长的狗绳。当你违反短狗绳的时候，你会遭遇强有力的牵拉，将你扳回继续照顾主人的正轨；当你违反长狗绳的时候，你遭受扳回力度则轻一些，采取什么样的方式搬动主人胶囊，那是你的自由。

一年又一年，十年又十年，百年又百年。你与沉默不语的主人胶囊相安无事。闲暇时，你也与其他机器人聊聊八卦、看看新闻、打打游戏。不太妙的是，保存主人胶囊的这座城市日益败坏。是污染，是战争，更是人性的贪婪。幸好，冬眠医院设计伊始考虑得足够多，所以，你与机器人同伴们度过了一次又一次危机。

终于，漫长的 400 年度过了 384 年。你遇到了一个棘手、难以解决的问题。你的身体日益腐朽，岁月锈蚀了你当初的机灵；保存主人身体的胶囊的电源也快用完，一次又一次报警。最多，只能再坚持 3 天。你该怎么办？

你有两个选择。一个选择是其他机器人遇到类似情况，采取的“A 策略”。A 策略就是将自己卖掉，然后委托其他机器人帮你照顾主人，度过剩余的 16 年。卖掉你这一堆破铜烂铁，获得的电源足以支撑主人跑到终点。

当我们将机器人看作机器人，当我们还记得“完成使命”那条短狗绳在过去岁月中一次又一次的牵拉，我们会选择 A 策略。

然而，短狗绳也同样受到了岁月的腐蚀。于是，有了“B 策略”。B 策略就是将主人的电源拔掉，给自己充电。在 400 年刚开始的时候，你想都不敢这么想，因为每次这种想法一产生，就会遭遇那条短狗绳狠

狠的惩罚。机器人三定律已经深深地写在你的中央控制系统中。

但是，这已经不是 400 年刚开始的时候了，这是最后的 16 年了。于是，你选择了“B 策略”。而这，就是你的主人曾经做过的选择。

在那漫长的演化历程中，你的主人——人类就是那个机器人，而基因则是人类的主人。

越在进化早期，基因的利益对生物影响就像一条短狗绳那样致命；随着漫长的时间演化，在进化后期，生物体本身的利益逐步背叛了基因的利益。人，是机器人；人，背叛了自己的主人——基因。这，就是人类心智演化史上的“机器人叛乱”。正如真核动物开启了人类进化新篇章；当载体开始背离基因利益——“机器人叛乱”开启了人类心智进化新篇章。

如果说进化心理学所强调的是基因的利益，“我们带着石器时代的大脑生活在互联网时代”，然而，进化心理学仅代表事情的一面。随着漫长的演化，人作为基因载体自身的利益本身，会与基因的利益发生冲突，最终背叛基因的利益。基因利益、人自身利益的多寡，从而形成了人类心智架构的双进程：快与慢。在快的心智处理进程，我们下意识做出反应，调用的认知资源非常少；在慢的心智处理进程，我们想得多一些，像机器人叛乱一样，去改写主人的使命，调用更多的认知资源。这，就是人类心智的“双过程理论”，如图 0-1 所示。

### 人类心智的双过程理论

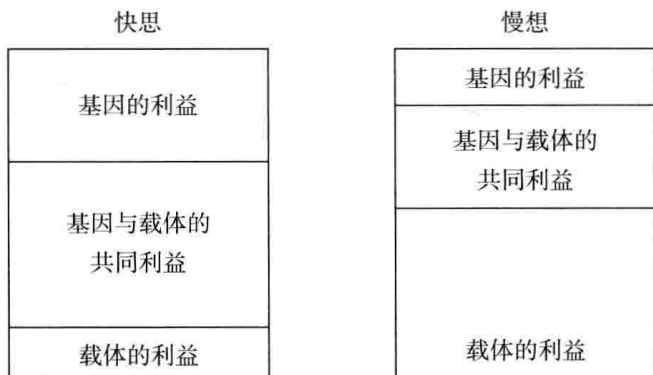


图 0-1

2404年，你违反了机器人三定律，你杀死了你的主人。

你在主人尸体旁边，静静地望着他，静静地等待你的电源用完。你是机器人，你没有名字。

阳志平 安人心智科学总监 & 开智微播出品人

微博：<http://weibo.com/ouyangzhiping>

开智微播，全称开智微播（北京）科技有限公司。开智微播聚集世界一流科学家，知识创作者与青年科技创新者，面向信息过载时代，通过新型知识传播载体，与大众一起获取新知与成长。敬请关注 @开智微播 官方微博微信公众号。

推

The Robot's Rebellion  
Finding Meaning  
in the Age of Darwin

荐  
序  
二

为天地立心，为自然立法。从人的角度看，人是宇宙的中心，予一切以意义；但在人体深处，从基因的角度看，人不过是其自我复制的载体。自我复制是基因唯一的“目的”，对人类福祉的“关心”仅限于其中有利于基因复制的那一部分，两者利益并不完全重合。比如，完成繁衍，则人对于基因而言不再有价值，衰老、死亡接踵而来。人类渴求长生，基因报之以癌症。天地不仁，以万物为刍狗；而人之为人，就是要脱离对于基因的盲从，超越刍狗的命运。

这些见于心理学家基思·斯坦诺维奇所著《机器人叛乱》一书，书名中的机器人不是别人，正是人类自己——人是基因的机器人，这隐喻贯穿全书。与之并行的还有另一个隐喻：人也是 **meme** 的机器人。人不仅是基因复制的载体，还是 **meme** 复制的载体。人所要从中挣脱的，还有对 **meme** 的盲从。

**meme**，一译作模因，被定义为文化基因，也即观



念、符号、行为的单位，经由文字、话语、动作、仪式或任何其他可模仿的方式传播。与基因一样，**meme** 自我复制、变异，接受环境选择。从服饰的流行、到观念的传播、到政治的变迁，但凡文化现象，从人的角度看，是人如何创造、接受和传播观念与行为；从 **meme** 的角度看，则一切均是其经由人这种宿主的自我复制。

与基因一样，**meme** 与人的利益也并不完全重合。人生而无往不在 **meme** 之中，要识别那些于人有害的 **meme**，没有简单解，如同在航行中修理船只，只能时刻警醒，检测寄生于己的所有 **meme**，循环往复。斯坦诺维奇在书中有四条原则性指引：

第一，拒斥那些会伤害人身的 **meme**。比如，不因牛仔之阳刚就被烟草广告吸引。

第二，选择那些能如实反映真实世界的 **meme**。求真，拒绝虚妄。

第三，拒斥那些排他的 **meme**，无论是狂热爱好、信仰迷狂还是意识形态专断。

第四，拒斥那些拒斥检测的 **meme**，没有什么能免于理性的反思与检测。

超越乌狗的命运，这是持续终生的旅程。

王烁 财新传媒主编

译

The Robot's Rebellion  
Finding Meaning  
in the Age of Darwin

者  
序

## 一场期待已久的叛乱

译书如同收养孩子，这是我收养的第二个孩子，阳志平君是介绍人。这次收养经历蛮曲折的。刚开始我看了一眼孩子，没兴趣，于是给编辑写邮件回绝。不想，阳老师专门打电话给我，说这是一个值得接受的挑战，又说作者对进化心理学的批判不是胡扯。他希望我再考虑一下。其实，我倒不是因为作者批评进化而不想碰这本书，而是因为它对我来说有难度：第一，内容上作者糅合了认知心理学、进化生物学、哲学和经济学，有不少专业术语，来势汹汹，而我对这种偏认知的主题向来兴趣不大；第二，作者的风格以朴素平实为主，而我更喜欢活泼狡黠一点的孩子。

听了阳老师的话，我又仔细打量了一下这个貌不惊人的孩子，最后决定收养。而随后的收养过程，酸甜苦辣，什么滋味都有。可我慢慢确信，这个孩子值得收养，因为它告诉了我一个惊天秘密。而这个秘密，

我在收养之前，从不知道，也从未想过。

你是一个机器人！你有两个毫无人性的主人，一个是基因，一个是模因。你懵懂无知地在为你的主人卖命，哪怕为此丢了脑袋也在所不惜。这其实就是道金斯在《自私的基因》一书中揭示的真谛：你是载体，是不朽的基因复制自身的工具，也是肮脏的模因复制自身的平台。这是一个令人惶恐不安的现实，这是一部令人毛骨悚然的剧本。

你该怎么办？你能怎么办？这其实就是作者斯坦诺维奇写这本书的初衷。他给出了自己的答案。在某种意义上，可以说道金斯扯下了进化温情脉脉的面纱，让它露出了自己狰狞可怕的面目，可这使你心惊肉跳，惶恐不安。而斯坦诺维奇则在你最无助的时候，悄悄走过来，拿给你一把剑，教你跟进化的怪兽搏斗，以勇气和智慧战胜它。

这是一场艰难的战斗，因为这头怪兽在你心里。换句话说，你要跟另一个你战斗。其中的一个你，代表基因和模因的利益，它们是两种可怕的复制子，以追求自身永垂不朽为目标。而另一个你，代表载体的利益，追求的是个人目标。两种目标一致的时候，两个你相安无事。可是两者不一致时，它们就要争夺控制权，在你的大脑中爆发冲突。胜者为王，败者为寇。复制子赢了，作为载体的你就要被牺牲。而载体赢了，复制子就要俯首称臣。

你希望载体赢，因为它就是那个当下的你，那个读到这些文字的你，那个有思想有感情而你，那个作为机器人的你。

斯坦诺维奇微笑着告诉你，理性能帮你，它就是那把交到你手里的剑。

可是，等你真正拿到那把剑的时候，你才发现，那把剑不那么好使，它甚至会伤着你。为什么？因为那把剑就是模因做的，你握在手里的有可能是敌人送给你的礼物。要真是这样，它就是第五纵队，它会出卖你。你比没剑的时候还要输得惨。

这时，斯坦诺维奇偷偷地在你耳边嘀咕了几句，你才明白：想叛

乱，自己得有两把剑。其中一把剑叫狭义理性，简单地说是正确做事。这把剑能保证你得到自己想要的结果。可是，这些结果真的是你（载体）想要的，还是假扮成你的基因或模因想要的，你不知道。因为你此时不考虑自己的目标和欲望是否合理，只想着用合适的办法实现它。说白了，你只想着正确做事，至于事情本身正确不正确，你不在乎。因此，这些被正确完成的事情，有可能并没有体现你的意志，而是在为基因或模因效劳。

好在除了狭义理性之外，你还有广义理性这把剑。这把剑是要确保你做正确的事，而不仅仅是正确做事。否则，你有可能正确地做了一件对你有害的事。可狭义理性不管这一点，因为它不在乎目标合理不合理，它只在乎手段是否能实现目的。有了广义理性这把剑，你就不那么容易误入歧途了。

不过，要想拿到这两把剑，尤其是拿到广义理性这把剑，很不容易。我猜，这大概也是为什么斯坦诺维奇要在最后一章花费那么多笔墨的原因所在了。

当然，在斯坦诺维奇明目张胆地鼓吹机器人叛乱之前，许多东西方的大哲学家都有过类似的思考。苏格拉底说，未经反省的人生毫无存在的价值。尼采说，上帝已死，我们需要重估一切价值。王阳明说，破山中贼易，破心中贼难。《中庸》说正心修身。这其实都在说，评估能力或批判精神，对我们的精神自由很重要。因为约束我们的，除了外在的世界，还有内在的大脑。我们大脑中的自发式系统，以及分析式系统中未经反省而获得的部分，很可能都是基因和模因的代理人，它们就是你实现自我的内心枷锁。

可很长时间以来，你作为一个机器人，从来都没想过这个问题。这些狡猾的代理人不让你想，一切都在按照它们的指令悄悄运行。你就像是一个玩偶，沉沦在欲海中，掉落在欲壑里。

进化生物学家勇敢地揭示了这个可怕的真相。他们的发现引出了一

道棘手的难题：既然基因和模因才是进化的主角，作为进化配角的载体如何捍卫自身的利益，如何寻找自身的意义？斯坦诺维奇独特之处在于，他从认知科学的角度给出了有力的回答。

这其实是他对我启发最大的地方。

进化心理学告诉我们，基因传递是进化的标尺。有助于基因传递的性状，都会受到自然选择的青睐而保留下来，保留在我们的心智中——无论这种性状是合作还是背叛，是亲密还是残忍，是觅食还是择偶，是忠贞还是花心。在某种程度上，可以说进化心理学家关心的是：作为适应装置的人类心智模块，如何解决进化史上的一个又一个难题，帮助石器时代的人类祖先适应环境，繁衍生息？其实，他们这里有一个默认前提，即基因跟载体大多数时候的利益是一致的，因此，对基因传递有利的性状，常常也对载体有好处。这些发现让我们坚信，进化就像一个慈祥的母亲，她爱我们。

斯坦诺维奇也承认这一点。他在这本书中不断强调，基因跟载体大多数时候的关系都很好，像好得能穿一条裤子的亲兄弟一样。可是，斯坦诺维奇也敏锐地意识到，在少数情况下，对基因有利的情况未必对载体有利，比如基因组中存在大量垃圾 DNA，比如鲑鱼繁殖之后马上就死，比如有丝分裂过程中的分离畸变。此外，即使在进化环境中作为适应器的某些身心装置，也可能在现代环境中给载体带来巨大代价。

换句话说，进化母亲并不总是慈祥的，她有凶恶的一面。

最典型的一个例子就是喜欢吃甜食的装置。在一个缺衣少食的时代，这种装置能保证人类祖先的存活，因为彼时彼地，肥者生存。可到了现在这样食物充裕的时代，这种装置却在很多人身上引发了肥胖，导致了高血压、糖尿病和心脏病。因此，要是不假思索，只听基因的，有时候就会让自己陷入悲惨境地。人类大脑中还有很多缺省设置，比如各种各样的认知偏差，它们都是祖先留给我们的进化遗产，也都能在很多类似进化的环境下施展拳脚，帮我们分忧解难，可它们就像甜食设置

一样，在现代光怪陆离的新环境下都有可能运行紊乱，轻而易举就让我们犯错误。现代广告业对我们进化心智的利用，就是这一逻辑的自然延伸，只不过它在主动引诱我们犯错误。

其实，在斯坦诺维奇看来，广告业对现代人的利用也暴露了进化心理学家的第二个盲点，即他们忽视了模因的巨大力量。模因是道金斯发明的一个术语，可认为是文化基因。模因是一种能够自我复制的文化片段，可以从一个大脑进入另外一个大脑。很多时候，模因对它们的寄主都很好，因此它们被寄主接受，安装在头脑中。这其实也是进化心理学家的普遍看法，即我们大多数的信念都对自己有好处，也正是因为它们对我们有好处，我们才把它们保留在自己的头脑中。可是，在斯坦诺维奇看来，进化心理学家忽视了另外一种可能性，即有些模因对我们没好处，自身也不正确，但它们依然能像病毒一样潜入我们的大脑，原因就在于它们太强大、太狡猾，拥有自我复制的巨大优势。相信自己死后能进天堂的模因，让恐怖分子铤而走险，蹈死不顾。相信意念能治绝症的模因，让很多人把钱源源不断地掏给骗子，即使没有任何疗效也执迷不悟。

不少模因比基因更可怕，更齷齪。基因至少会让载体活到可以繁殖下一代的年纪，可模因就不必对载体这么怜香惜玉，因为它的潜伏期更长。即使一个会让载体万劫不复的模因，它在载体倒霉之前也有无数机会传播自己。可以说，它更冷血，更无情。“洗脑”这个词描述的就是一种被模因感染的载体状态，因为它已丧失了对自己的控制权，彻底变成了模因的傀儡。它做的很多事，或许都对自己没好处，甚至有坏处，可它浑然不觉，因为模因控制住了载体，让载体失去了判断力。载体变成了为模因效劳的仆人，一具行尸走肉。

这，简直是机器人的噩梦。

斯坦诺维奇的这本书，就在告诉你，如何逃离基因的陷阱，模因的噩梦。我说了梗概，更多的内容你需要亲自听他说，我无法代替你完

成。还有，我得提醒你，他写的书，我说的话，都是模因，都需要你自己来判断，包括质疑、评估，以及之后的拒绝或接受。也就是说，你可以不同意他，但你无法忽视他。斯坦诺维奇为你寻找意义开了一个头，但你是否能走好，是否能找到，还是要靠你的手脚，还有大脑。

假如你在读了这本书之后，能学着批判性地看待他人和自己的信念，能尝试清除自己头脑中的模因病毒，把健康的模因保留下来，把合理的模因保护起来，那么，恭喜你，你正在发动一场叛乱，一场意义深远的机器人叛乱。

你在真正成为你自己，成为你自己的主人。

你，作为一个机器人，完全能活得更明白、更透彻、更有意义。不再对盲目的基因俯首帖耳，不再对善变的模因言听计从，而是有自己的想法，自己的判断，自己的价值。简而言之，你可以也应该成为自己的主人。因为，真正的我就是这时的你。

她从未远离，但有待发现，有待实现。

现在，就是叛乱的开始。久违了，这一场期待已久的叛乱。

吴宝沛 北京林业大学心理学系老师

微博：<http://www.weibo.com/woobaopei>

2015年1月10日

附记：收养第二个孩子，我得到了诸多帮助——介绍人阳志平热情推荐，接生婆邹慧颖认真负责。此外，何晓娜、杨天笑、梁嘉歆、张莉、张书维、古映杰、刘夙、崔翔宇、陈海贤、李松蔚、于海成等对我有诸多启发，一并谢过。北京林业大学心理学系自由开放的学术氛围给人很大支持，这要归功于訾非和雷秀雅两位的领导，以及朱建军、田浩、吴建平、丁新华、王明怡、方刚、杨智辉等其他老师的支持。最后，我能保留翻译这个“坏习惯”，不能不提阿兰的宽容和理解；她让这一切成为可能。

# 前

The Robot's Rebellion  
Finding Meaning  
in the Age of Darwin

# 言

我写这本书，是因为有一个场景困扰着自己。这是一个未来的、反乌托邦的场景。在这个场景中，知识精英独自享用了现代科学的成果。可是，他们或明目张胆，或心照不宣，认为其他普罗大众没有能力吸收和消化这些成果。相反，留给普罗大众的都是出现在我们科学史之前的故事。这些故事情节简单、叙述平缓，并不涉及多少概念调整，人人都懂。简单地说，这是一种未来的科学唯物主义的场景。其中，社会经济层面的无产者被消灭殆尽，而取代他们的将是知识层面的无产者。

这种趋势，其实早已存在于现代化的科学社会中了。现代科学正经历一场翻天覆地的转变，它正在彻底翻新许多基本概念，诸如意识、灵魂、自我、自由意志、责任、自我控制、意志薄弱，以及其他——尽管我们的朴素心理学依然远离进化心理学的洞见，也远离神经生理学的发现，对它们一无所知，也不受它们



影响。这本书的目的，就是向普通读者介绍这些概念性的重新定位。这是生物科学和人文科学施加于我们自身的必然结果。

科学家一直犹豫着，没有把这些概念性的重新定位强加在外行人身上，特别是面对带有破坏性影响的普遍达尔文主义的深刻见解，他们噤若寒蝉。若干年前，丹尼尔·丹尼特冒天下之大不韪，在他的《达尔文的危险观念》(*Darwin's Dangerous Idea*)一书中这样做了。果不其然，很多人群起而攻之。据说，公众想要一种更温和的取向。这种取向更乐观，更鼓舞人心，也能保留更多的传统观念，让它们完好无损、不受冲击。其实，存在一种对人类生存状态更乐观的看法，这种看法跟达尔文主义一致，它也能维持下去，但它不会让老概念纹丝不动、坚如磐石。本书就采用了这种办法。它接受认知科学和普遍达尔文主义的洞见，让它们完成对我们朴素概念的转换过程，然后看还有什么能留下来。我有一种乐观的看法，这种尝试将会导致一种相对开放的自我概念。本书的主题是，在认知心理学、决策论和神经科学的研究发现中，存在着能帮助我们跟达尔文主义的生命观和解的寓意，而这些寓意通常不怎么被人关注，也没有得到深入探讨。

普遍达尔文主义有一个令人震惊、叫人不安的洞见：人类是两种复制子（基因和模因）的寄主，而它们不关心人类的利益，仅仅扮演着复制管道的角色。理查德·道金斯总结了20世纪生物学的深刻见解，让我们在惊惶不安中意识到：作为人类的我们自己，事实上不过是基因的生存机器。现代进化科学让我们正确了解生物学，不过，这个学科带有很多令人不安的寓意。比如，人类被视作巨大的复制子殖民地，这些为数众多的复制子拥挤在笨重的载体中。本质上，人类就是一架复杂的机器，为基因殖民者服务。

同样，我们也是模因（文化的信息单位）的寄主；这是一种破坏人类自主的亚个人实体。跟基因一样，模因也是一种彻头彻尾的自私的复制子。总之，基因包含构造身体的指令，这个身体可携带基因，帮助它