Machines Like Me

Ian McEwan

我这样的机器

麦克尤恩作品 | Ian McEwan

Machines Like Me
And People Like You

我这样的机器
你们这样的人

[英]伊恩·麦克尤恩————著
周小进————译

上海译文出版社

Ian McEwan
MACHINES LIKE ME

图字：09 - 2019 - 226 号

图书在版编目(CIP)数据

我这样的机器 /（英）伊恩·麦克尤恩（Ian McEwan）著；周小进译. —上海：上海译文出版社，2020.7
（麦克尤恩作品）
书名原文：Machines Like Me
ISBN 978 - 7 - 5327 - 8507 - 0

Ⅰ.①我… Ⅱ.①伊… ②周… Ⅲ.①长篇小说—英国—现代 Ⅳ.①I561.45

中国版本图书馆 CIP 数据核字(2020)第 083811 号

我这样的机器
[英] 伊恩·麦克尤恩 著 周小进 译
责任编辑 / 宋 玲 装帧设计 / 储平工作室

上海译文出版社有限公司出版、发行
网址：www.yiwen.com.cn
200001 上海福建中路 193 号
江阴金马印刷有限公司印刷

开本 850×1168 1/32 印张 11 插页 5 字数 187,000
2020 年 7 月第 1 版 2020 年 7 月第 1 次印刷
印数：00,001—15,000 册

ISBN 978 - 7 - 5327 - 8507 - 0/I · 5236
定价：89.00 元

好机器人必须死

——《我这样的机器》导读

不知麦克尤恩有没有考虑过“恐怖谷”。“我这样的机器”，他的书名像是在玩味那条规则。1970 年，在一本名叫《能量》的古怪杂志上，东京工业大学教授森政弘发表了一篇文章，其标题“不気味の谷”中的“不気味”，是说那种莫名让人不舒服，让人毛骨悚然的感觉，也有人把它翻译成“诡异谷”。

人如果想要造人，也会像上帝造人那样拿自己当模型。从工业机器人、玩具机器人，一直到伴侣机器人，面貌体态会越来越像真人。总有一天，就像这部小说中的亚当，机器人会造得看上去完全就是一个真正的人类：身体健壮、相貌英俊、深色皮肤、浓密头发、鹰钩鼻子。肤色、心跳无不像真人，甚至呼吸都带着一丝湿润。只不过在它的浅蓝色瞳孔上，仔细看会发现有很多条纹。

按照“恐怖谷”理论，看到这样的机器人，人会产生奇异的恐惧感。森政弘用机器人仿真度做横轴，以人对机器人的“親和感”（しんわかん）为纵轴，画了一条函数曲线。机器人越像真人，人们就越喜欢它们，跟它们越亲近。可一旦人们成功设计出

仿真度很高的机器人，那条单调递增的曲线就会突然跌入低谷，他把这个称为“恐怖谷”。也就是说，如果机器人跟真人只有极细微差别，它们反而会让人害怕。“恐怖谷”理论渐渐受到重视，首先是娱乐工业，有些科幻电影因为设计了更逼真的机器人形象，反而票房惨败。游戏角色形象设计似乎也印证了这个猜想。2012年，有人将其翻译成英语，在美国电气电子工程师协会(IEEE)《机器人与自动化》杂志上发表，引起广泛关注。

至于何以会有这种现象，有一种解释利用了“心智理论”和脑神经科学最新研究成果。前者是说，人类有一种独特能力，能够理解他人的内心想法和意向，能够将心比心。有时候不需要对方说话表示，甚至不需要动作和表情，人就能猜到对方的心理状态。这种读心能力被称为“心理化”(mentalising)。人只有在面对其他人，也就是他们的同类时，才会启用这种心理化能力。有人做过一个实验，在屏幕上给志愿者播放仿真度不同等级的合成人像，对他们的大脑进行电磁扫描，发现合成人像仿真度越高，大脑中负责心理化的区域就越亮，证明其活动强度越大。意大利神经生理学家贾科莫·里佐拉蒂(Giacomo Rizzolatti)在猴脑腹侧运动前区发现了一种镜像神经元。当猴子做抓握推拉动作，或者捡起花生、剥壳放进嘴的一组动作时，大脑中会有相应的特殊神经元放电。神奇的是，当一只猴子看到另一只猴子做这些动作，它自己并没有做时，相应的神经元同样也会放电。研究者相信，这是人类进化的最大秘密，因为有了这些镜像神经元，人可以模仿学习他人的动作行为；可以学会利用口腔中的肌

肉，互相发出同样的声音，由此学会说话和语言；也可以发展出心理化能力，猜测别人的内心世界。

研究者们说，正是因为人有心理化能力，运动皮质层中有大量镜像神经元，人对机器人的观感才会有“恐怖谷”。因为这个仿真机器人，怎么看都像是个同类，大脑开始启动心理化，镜像神经元也开始放电。可是与此同时，大脑也十分确定，面前这家伙肯定不是真人。那么，到底是启动呢是启动呢还是启动呢？大脑陷入错乱。加州大学圣迭戈分校塞琴教授（Ayse Saygin）让人观看一个工业机器人、一个真人和一个仿真机器人的视频，结果发现观看仿真机器人视频时，大脑特别活跃，显示出不安情绪。

按“恐怖谷”理论，当亚当来到小说中那些人面前，他们应该坐立不安。尤其当他们看到亚当奇特的瞳仁，或者意识到它绝不是他们的同类。可他们似乎迅速地接纳了它，查理是花钱把亚当买回家的主人，他的不动声色容易理解，他早有思想准备。米兰达呢，头一天晚上她对查理说，亚当的身体是暖和的，真“吓人”。亚当能用舌头发音，说出词语，“有点怪怪的”。这些感受有点接近“恐怖谷”了。但她很快就习以为常，甚至率先开发利用了亚当在某些方面的特殊能力。其他人，比如米兰达的父亲，那位在很多方面显得颇为睿智的作家，他甚至把机器人亚当和准女婿查理搞混了，分不清谁是真人谁是假人。儿童们以他们纯净的感受，常常能够观察到亚当有些异样。但谁也没有对亚当真正感到恐惧。要知道，在一般科幻电影剧集当中，人与机器

人之间那条“恐怖谷”之深之宽，简直类似某种种族仇恨，必欲杀之而后快。我们玩“底特律变人”[①]，在思想上预先就做好了准备，人和机器人永远势不两立。

麦克尤恩笔下的人们轻松地越过了“恐怖谷”。当然，这也是理所当然。因为麦克尤恩从不关心那种预先注定的憎恨，他的人物也会发生冲突，甚至也会互相杀害。但一切冲突无论最后如何激烈决绝，最初都来自于日常生活中偶尔出现的一条小裂缝。杀人犯不例外，机器人也不例外。如果像很多人以为的，人和机器人在未来互相有一场生存之战，麦克尤恩倒想找找硅基碳基物种差异以外的原因。因为人造了机器人，就像上帝造了人，一开始互相都满怀着善意，在伊甸园中其乐融融。到底是从哪儿出了问题呢？

小说男主人公查理，是个自由自在的宅男，不上班。他只要电脑、网线和一小笔钱。股票、期货或者房地产，他每天就在各种投资投机市场上找机会。有时赚有时赔，基本能满足温饱。作为一个后现代主义青年，他接受了一种价值中立的道德视角。因为进化心理学、因为文化人类学，它们的初级课程往往会让人获得一种相对主义的伦理自由感。这种道德相对论让他在税务上闹了一点小麻烦，上了法庭。政治上查理有点保守，支持撒切尔对马岛的远征，他的女朋友米兰达则绝不认同，为此他们争论起来。那是他们第一次吵架，造成了让人吃惊的后果。

---

① 一款人工智能题材互动电影游戏，2018 年 5 月 25 日发售。

仿生机器人亚当夏娃们上市时，他正好继承了一幢房产，卖掉房子得到一大笔钱。宅男们总想比别人抢先一步进入未来世界，如果能力有限，就买下可能通向未来的新潮电子产品。尽管只是测试产品，beta 版，而且要花 8.6 万美元，他也毫不犹豫下单了。公司统共只推出了 25 个，12 个亚当，13 个夏娃。他买到亚当，虽然他本想买一个夏娃。

女主角米兰达，社会史专业博士、查理楼上的邻居、比查理小十岁。她也算亚当的半个主人，查理自己向她让渡了这个权利。他一直不知道如何向米兰达表达感情，希望这个机器仆人能成为某种联结象征。亚当会进入他们俩的生活，查理觉得当他和米兰达一起关注亚当，创造亚当后，就会自动成为一家人。

亚当确实需要他们“创造”。因为查理把它领回家时，它仍是原厂设置，亚当可以按照说明书，在个性化设置上“自由创造”。他决定了，自己只完成一半，另一半让米兰达勾选。如此一来，他和米兰达那种看起来有点不切实际的关系，就能够凝结在有形实体上了。但这个慷慨的举动将会成为日后所有困扰的起点，此刻他并未预料到。

在小说中，这批测试版仿真机器人于 1982 年面世。把人类目前远未实现的科技能力设定在过去的历史时间当中，是常见的科幻故事装置。让一种未来技术穿越时间回到过去，或者索性立足于现代物理学观点，构造一个平行世界。作者由此可以技术决定论地设问：科技发明会在多大程度上改变人类命运？但这部小说并不关心此类问题，让亚当在 1982 年来到查理家中，

或者让它在2032年出现，看起来没什么差别。

我们猜想起来，麦克尤恩让故事发生在1982年，更可能是出于——作为一个小说家，对有关机器人的大众叙事历史和观念变迁的关注。有一个历史统计数据，到1981年，日本汽车装配生产线上已使用了6 000多个机器人，同一时间英国只有370台工程机器人。英国首相撒切尔夫人于是大力鼓吹广泛使用机器人作业，此举或许也是首相对各地不断出现罢工浪潮的冷酷回应。当时英国失业率常年徘徊在10%。撒切尔推广机器人的言论得罪了工人大众。本来憎恨机器的卢德派在英国就有久远历史渊源，政府和工人们两相激发，机器人成了那个年代英国人最热门的话题。

事实上，如果你在谷歌图书词频（Google Ngram Viewer）上检索“Robot”（机器人）和“AI”（人工智能），就会看到这两个词在80年代初异军突起，检索量在整个80年代形成一个高峰，到90年代反而渐渐下落。事实上，最令人难忘的机器人电影就是此刻拍摄的——1984年上映的《终结者》。显然，智能仿真机器人是那个时代极其广泛的大众话题。那时候，麦克尤恩刚搬家到伦敦没几年，出版了几本小说，开始创作剧本，与同道友好交往，参加午餐会讨论热门话题，也许这个有关“比人更聪明的人造人”的想法，在这个30岁出头年轻作家（正是小说中查理的年纪）的内心，一度掀起过极大波澜。

Robot这个词，来自捷克作家卡雷尔·恰佩克（Karel Čapek）的那部戏剧，《罗素姆的万能机器人》。作者在剧中借用了捷克

语“Robota”(劳工),变造了机器人一词。这部戏于1922年在纽约上演时,适逢大萧条,戏剧故事迎合了自动机器会加剧工人失业的恐慌心理。机器人话题的每一次真正热门,都跟失业潮、跟担心机器人在所有工作岗位上取代人类有关。

但麦克尤恩并不为此担心。在这部小说中,每个人物都欣然接受亚当的代劳。从厨房清洗到检索政府档案卷宗——亚当的大脑可以直接接入网络,密码也挡不住它。亚当开始热衷于文学创作,虽然只是最简单的日本俳句。但文学,那不是人类智慧树最高处结出的果实吗?没人当回事。对于亚当给生活带来的种种改善,男主角查理心安理得地享受着。亚当甚至帮他操盘做投资,它是真正的超级智能,不是如今投行们使用的那些高频交易算法。盈利是毫无疑问的,就像从自己口袋拿钱。根据未来生命研究所所长泰格马克的思想实验,他构想的超级人工智能“普罗米修斯”只用数月时间,就能让100万美元的起始资金通过投资市场增值至10亿美元(《Life 3.0》)。让人有点惋惜的,倒是麦克尤恩没让亚当迅速赚上几亿英镑。亚当颇有节制地每天只赚那么一点点,正好能让查理过上入门级的富裕生活。也许是因为亚当的智能算法告诉它,如此程度的财富最能让人有幸福感。由此读者隐约意识到亚当的算法中,包含着某些道德参数——理所当然,亚当是按照上帝的至善标准制造的。而这将会让后现代道德相对主义的查理和米兰达陷入困境,但此刻他们毫无警觉。

真正刺激到男主角查理的,是亚当对另一件事情的代劳。

在他看来，那相当于亚当给他戴上了绿帽。米兰达跟亚当上床，就在楼上她自己房间。过程中查理都能听到，或者有那么一会，是听到没有任何动静。他甚至感觉自己看到了整场戏，用他的“意识之眼”，或者“内心之眼”。这正是麦克尤恩擅长的，他的人物偶尔会获得一种超现实的感知能力。这节“米兰达出轨”故事，完全出自查理夹杂在酸楚幻想中的旁听视角，其内心复杂滋味层层揭露。他一边伤心，一边却又特别兴奋。这可能多少跟某种古怪的雄性本能有关，一种反向的生物信息素电化学反应。更主要的是此时此刻，查理竟萌生了某种时代弄潮儿的感觉（riding the breaking crest of the new）。因为在人类历史上，他是第一个被机器人戴上绿帽的。人们常常忧虑于未来会因为机器人而“下岗”（displacement），这出理应史诗般的大戏，却由他第一个悄悄出演了。可是如果机器人什么都能做得比人类更好的话，有什么可以阻挡人家不使用它们呢？

让仿真智能机器人担任性伴侣，如今应该算近未来科幻了。事实上至少已有两家公司拿出了真正的产品。一个是“真伴”（True Companion）公司的“萝克西”（Roxxxy）。分金银两种产品序列，“银萝”价格 2 995 美元，她能应景说话，“金萝”9 995 美元，根据公司宣传，她能“听”你说话。能听比能说昂贵得多。在这两条产品线中，又按年龄个性分了好多型号，“冷感法拉”、“野性温蒂”、“熟女玛莎”之类。真伴公司也做亚当类产品，名叫 Rocky，听上去可比亚当威猛多了。跟亚当一样，它们也有心跳和体液循环。“大脑”输入存储了大量数据，其中包括维基百科。

你可以到油管上找到它们的视频，从背后看跟真人几乎没什么差别。另一家公司的产品名叫“真偶”(RealDoll)，似乎应用了人工智能科技方面更新的研究成果，售价 5 000 美元。不要小看它们，虽然目前看起来都只算是新奇玩具，但确实前途无量，因为纵观人类科技发展史，真正推动技术产品进步的，要么是战争，要么就是性。

发生了那么一件事，米兰达不以为意，读者甚至会怀疑她故意让查理听见。不是出于某种情趣，而是由于她对男性多少有一些无意识的敌意(读者以后会慢慢发现这点)。无论如何，她觉得亚当只是一台机器。查理却无法同意她的观点，两人的分歧实际上源于信息不对称。因为亚当，这台机器，曾独自发现了米兰达的秘密，悄悄告诉查理：她撒了个弥天大谎。

这是至关重要的情况，亚当能够理解米兰达撒谎。查理凭这一点认定亚当已跨过机器智能的边界，不能简单视其为机器了。对于亚当，查理头脑中产生了定义混乱。我们先前说过，“心理化”是人类最重要的一种能力，几乎也是人类独具的能力。知道米兰达说谎，知道她为什么说谎，知道不能把她说谎的内容告诉查理。窥测如此复杂多层的意向性水平，以前只有人类才能办到。心理化，就是猜测人心中隐秘的多层意向，这个过程总是跟一些动词相关联，当一个人说他“猜想”、他“知道”、他“怀疑”时，他就表现出了一阶意向性水平。也就是说他了解自己头脑中在想什么。人也能推测他人的心智活动，“我怀疑他知道”，这个判断能力一般 5 岁儿童就能习得，那已是二阶意向性水平。

人类可以猜测他人意向达到六阶水平：我想/你会认为/我知道/你希望/我怀疑/你在猜测。查理知道，亚当一旦能洞察米兰达说谎，就距离他产生自我意识不远了。

果然，不久亚当就对查理宣布，他爱上了米兰达。这是机器人亚当产生自我意识的明证。很简单，如果没有“我”，如何会有“我”爱你？意识，或者自我意识，或者“心灵”，每个人都知道它在那里，却没有人知道它在哪里。时至二十一世纪，人们相信事物首先必须在物理上存在，它们才存在，人脑也不例外。意识虽然很神秘，大部分科学家都相信它也只是人脑的电化学活动，像迪昂(Stanislas Dehaene)这样的神经学家，已在实验室追猎到意识的一些踪迹。其中也有人把物理主义推到极端，直接否定了意识的存在，认为从来就没有这回事，那都是科学蒙昧时代的迷思。另外有一些严肃科学家，则坚持意识“属灵论”，相信意识是人类天赋，是自然界的一种例外。

由此，未来智能机器人会不会有自我意识，同样成了一个争论焦点。有些科学家认为就算超级人工智能越过了奇点，它们也不会像人类那样思考。至少从目前已出现的机器智能算法上来看，它们和人类智能在结构和本质上都不是一回事。比如日本有人写出一款人力资源管理算法，它可以约谈员工，预测他们有没有辞职倾向，准确率高得让人吃惊。但它并不依靠理解对话表面语义，只是对员工表述内容进行深度语言结构分析，从单词频率、次序上寻找特征。机器不用读心，人从群体进化中获得的心理化能力，机器根本不需要。

另一些人则认为，未来的超级智能，如果说有可能实现，一定在某种程度上模拟人脑活动，循此道路，机器产生自我意识也一定会发生。麦克尤恩显然相信机器人一定会产生意识。它们不仅能学会烤鸡、写诗，他们也会萌生爱情、并把它埋在内心深处，在自我意识浇灌下让它变得越来越强烈。

一旦机器产生自我意识，就会遭遇到奥莫亨德罗（Steve Omohundro）的难题：自我意识意味着自我保护，机器会认识到首先必须存在，才能达成围绕着自我意识的所有其它目标。超级智能机器会设法让开关失效。当查理再一次试图伸手关掉它的电源，亚当阻止了他，捏碎了他的腕骨。亚当坚定地夺回/取消了开关控制权。不过当查理把这件严重事故告诉小说中的“图灵”，也就是亚当和夏娃们真正的设计者时，图灵并没有震惊，认为这不过是一个“值得关注的”情况。图灵告诉查理，据他所知，那批测试版机器人中，已有十一位设法取消了开关。这不奇怪，历史上那位真正的图灵对此曾反复思考。1951 年，他在 BBC 三台谈话节目中做了一个演讲，题目叫“机器能思考吗?”（Can Digital Computers Think?）。其中说到，就算人类把开关控制在手中，紧急时刻有能力关掉它，对人类来说，这也够耻辱的了（We should, as a species, feel great humbled）。图灵是有点忧虑，可显然不赞成靠开关维持优势。

麦克尤恩设想了一种道德至上的智能机器。在它们复杂的运算“黑盒”底层，必定预置了一些最高级道德命令，以防发生难以控制的不测事件。因为按照牛津大学哲学家尼克（Nick

Bostrom)的预想，即便是一个善意的机器人，只要有一个人类命令它制造回形针，它也会把地球上的一切都变造成回形针，因为智能机器人在完成任务方面是一个彻底的完美主义者。可是，亚当没有为查理赚回世界上所有的钱，显然它在决策中综合考虑了互不相容的几种人类偏好。比如说，也要兼顾安全感，过量财富显然会带来身心两方面的危险。尤其是道德损害。

道德规则本身就包含无数偏好。比如小说故事就涉及到：是“复仇正义”要紧呢？还是“不能欺骗他人”更重要？这才是麦克尤恩真正感兴趣的问题。工程师给亚当们内置了所有人类道德规则，但那些远不是经由自我审视、选择而来的道德感。出厂设定代表了人类所有最美好的期望，可就算是人类自己，也没有一个能做到。因为那些规则条款，根本经不起社会人际摩擦。正如麦克尤恩几乎所有小说都发生的情况，那些天性良善的人物，让日常琐细冲突愈演愈烈，直至不可收拾。小说中那些亚当夏娃们再一次证明了麦克尤恩的观点。它们很快就陷入意识崩溃，无法在人世生存，一个接一个自杀了。

在它们中间，亚当遇到的可能是最好的主人，因为查理和米兰达凡事妥协让步的性格，使得亚当跟他们的冲突，能够拖延很久才爆发。但自从查理和米兰达分别设定，把亚当个性一分为二时，就注定了冲突的不可避免。给他输入的道德命令行，简单刻板而井井有条。但正如小说中图灵指出，在日常生活中，人类的情感、偏见、自我欺骗，以及其它各种认知缺陷会构成一个力场，互相挤压，道德原则在其中扭曲变形。亚当的大脑根本无法

处理如此复杂的情况，因为人类自己也不明白究竟，从来就没有解决方案，他们只是一味妥协，妥协，直到冲突爆发（就像麦克尤恩小说中每个人物都遭遇到的）。

亚当早晚会在某一天宕机。现在这个结局只是让读者更加震惊，它附带拷问了查理和米兰达的灵魂，他们的善意，他们的与世无争，他们自以为中立的道德价值视角，何以会导致小说最终给予他们的这个结果？

小　白

2020 年 6 月

献给格雷姆·米奇森[1]

1944—2018