

AI 人工智能系列



人工智能 与 本能

如何让机器人拥有自我意识

◎ 王昭东 编著

ARTIFICIAL
INTELLIGENCE
&
INSTINCT

▼
▼
国内首本关于强人工智能的图书
智能是基于环境变化而做出的一种相对友好的选择

人工智能系列

人工智能与本能

——如何让机器人拥有自我意识

王昭东 编著

电子工业出版社

Publishing House of Electronics Industry

北京•BEIJING

内 容 简 介

人工智能经过 60 多年的发展,到现在仍然没有超过 5 岁小孩的智力水平,不断提高人工智能技术已成为众多科学家和工程师努力探索的事业。

本书共分为 4 篇,第 1 篇对人工智能进行了概要说明,让读者对人工智能有一个初步的了解;第 2 篇主要讲述了人工智能的原理和核心本能的重要性;第 3 篇介绍了如何利用核心本能创造出真正拥有灵魂的、有情有义的强人工智能;第 4 篇介绍了机器人三大定律和人工智能示例。

本书没有任何公式,希望读者在阅读此书之后能够爱上人工智能,也希望读者积极踊跃地参与到人工智能的浪潮中来,从而抓住人工智能的发展机会成就自己的梦想!

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

人工智能与本能:如何让机器人拥有自我意识 / 王昭东编著. —北京:电子工业出版社, 2017.10

(人工智能系列)

ISBN 978-7-121-32748-3

I. ①人… II. ①王… III. ①机器人—自我意识—研究 IV. ①TP242

中国版本图书馆 CIP 数据核字(2017)第 232106 号

策划编辑:李树林

责任编辑:底 波

印 刷:

装 订:

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本:720×1 000 1/16 印张:13.25 字数:165 千字

版 次:2017 年 10 月第 1 版

印 次:2017 年 10 月第 1 次印刷

定 价:45.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010) 88254888, 88258888。

质量投诉请发邮件至 zlt@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式:(010) 88254463, lisl@phei.com.cn。

序

非常高兴能够为昭东的新书写序，昭东是我们大数据团队的联合创始人&CTO。熊猫大数据量化团队成立于 2015 年，是国内较早采用人工智能+量化的金融科技公司。现在每天都有成百上千的策略在交易中不断进化着。

人类的发展史就是一个发明的发展史。如果说农业革命提高了粮食的产量，工业革命提高了劳动效率，那么智能革命则提高了人类的智力，每一次科技的飞跃无不伴随着世界的改变。人工智能经过六十多年的起起落落，现在终于迎来了它的“春天”，但这个技术似乎还未成熟，感觉还有一层看不见的窗户纸，或许是因为目前人类还没有办法创造出真正拥有意识的机器人。当人类创造出了真正拥有意识的机器人后，我们才能宣称奇点时代的到来。

本书正是一部探讨情感机器人和自主意识的科普图书，全书由表及里探讨人工智能与人类本能之间的联系，视角独特，可以说是中国进化论学派的启蒙图书。书中示例生动、通俗易懂、分析讲解详尽，相信通过阅读本书并参考后面的示例代码，你的强人工智能必将呼之欲出。

熊猫金控创始人——赵伟平

前言

自从“上帝创造人类”开始，人类就开始踏上了寻找“神仙”之旅。人工智能作为人类实现人间天堂的一个方向，被广泛地应用于互联网领域，但在计算机技术高度发达的今天，为什么人工智能经过这么多年发展还只有一个婴幼儿的智力水平？是什么阻碍了人工智能的发展？又是什么才能引领人类走向辉煌？今天就让我们共同揭开人类智慧的面纱，打造属于我们自己的人工智能吧！

在讲述什么是强人工智能之前，我们先来做一个小的问答实验。

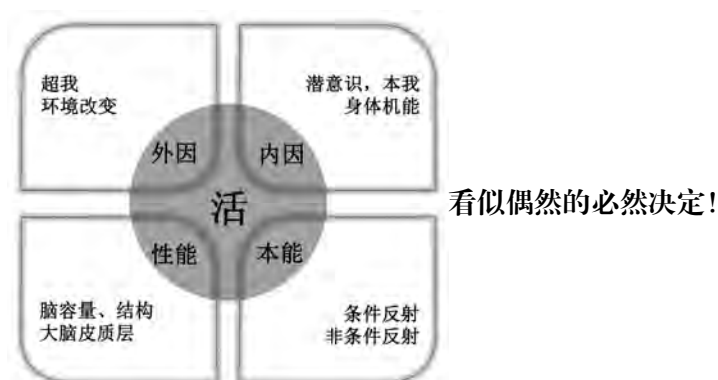
（1）假设现在我要把你的腿给锯掉，你会同意吗？我相信 100%的人都不会同意，好端端的为什么要锯掉我的腿呢？

（2）假设你遭遇了严重的车祸，必须锯掉腿你才能活下来，那么你会同意吗？我相信 90%的人会同意。

对于同样的问题，但人们却做出截然不同的两种选择，如果给你充分的时间来思考，不知道你能够发现其本质是什么吗？

对，它们的本质就是让你能够更好地活着。

人类的一切智慧反应实际上都是围绕着如何让自己生活得更好而进行的选择，正是因为这样一个最为核心的本能，我们人类才在内因、外因、本能和性能综合计算下做出一个更好的选择。



既然人类智慧的真谛就是更好地活着，那么我们可以通过这个原理，让计算机同样拥有智慧。

现在假设我们给计算机一个保护电源的核心本能，然后所有的计算都是围绕如何保护电源这一点进行的，那么计算机在进过若干次训练之后，它就会主动亲近那些保护其电源的人而远离那些想拔掉其电源的人。这种看似幼稚的选择，在我们看来它就是有意识、有感情的机器人的表现。

本书共分为 4 篇而且没有任何公式，几乎大家都可以阅读。希望读者在阅读此书之后能够爱上人工智能，也希望读者能够积极踊跃地参与到人工智能的浪潮中来，从而抓住人工智能的发展机会成就梦想与未来！

本书向读者揭示智慧的本质，并让读者利用核心本能的设置赋予机器人以灵魂，也许只有你真正把机器人当朋友时，它才是你的朋友。让我们放开手脚创造属于我们自己的人工智能吧。

王昭东

2017 年 9 月

目录

第 1 篇 关于智能

第 1 章	宇宙智能	7
1.1	无中生有	8
1.2	史前文明	16
1.3	循环迭代	22
第 2 章	生物智能	31
2.1	生物学	31
2.2	医学	37
2.3	社会学	43
第 3 章	人工智能的探讨	46

第 2 篇 人工智能

第 1 章	符号主义	50
第 2 章	联结主义	56
2.1	机器学习	56
2.2	常用框架	63
2.3	硬件设备	71
2.4	云计算	74


2.5 量子计算机	77
第 3 章 行为主义	80

第 3 篇 生存本能

第 1 章 什么是好，什么是坏	88
1.1 人为什么会死	88
1.2 什么是美	94
1.3 我是谁	99
1.4 生存的本能	101
第 2 章 本能的演化	122
2.1 核心本能设置	122
2.2 数据就是能量	129
2.3 遗传与进化	133
第 3 章 人工智能的未来	146
3.1 人机结合是人工智能的未来	149
3.2 人类可以通过人工智能实现再生	155
3.3 人工智能是否应该有寿命	159

第 4 篇 创造智能

第 1 章 关于机器人的三大定律	164
第 2 章 人工智能示例代码	171
附录 A 人工智能生态系统	201
附录 B 人工智能思维导图	203



第 1 篇

关于智能

现在我们天天说人工智能，殊不知人工智能只是智能的一部分，在说人工智能之前我们不妨先说下智能，那么到底什么是智能呢？

说起智能也许有人会说——智能其实是从记忆到思维再到行动的一个过程。或者也有人会说——智能是智慧和能力的一种综合缩写。但是不管怎么说，我们似乎总感觉好像缺了点什么，缺乏一种令人真正信服的依据，而这个依据又是那么近在嘴边却难以言传。

古往今来，古今中外每个人都对这个谜题做了不懈的努力，直到今天我们要真正揭开这个神圣的谜底。

由于早先大家对智能的理解不尽相同，甚至关注的领域和研究的方法也大相径庭，为了方便大家进一步交流，这里把智能分成了三个主要大类。其实这个分类并没有严格意义上的科学根据，只是为了方便大家交流而采用的一种折中方法而已，大家可以不用太较真。

有一种人，我们称其为哲学家，他们时不时地会想：

时间是什么？

宇宙是什么？

物质是哪里来的？

宇宙是怎么构成的？

宇宙是怎么运转的？

宇宙的外面是什么？

天外是不是还有天啊？

.....

我们把研究这种问题的人称为宇宙智能研究者。由于宇宙是一切时间、空间和物质的综合体，所以研究宇宙从某种意义上来说更加宏观、更加准确，从而避免盲人摸象和管中窥豹的尴尬。

宇宙智能是把整个宇宙当成一个智慧体来进行研究的科学，而我们人类则是其智慧的一小部分。也许有人会说，研究宇宙的人是哲学家和天文学家干的事情，一切都和我们普通人无关，其实我们仔细想想，会发现我们每个人本质上也都是宇宙智能的研究者之一，只不过是在有意无意间想多想少的问题了。

其中，古希腊非常有名的哲学家柏拉图（见图 1-0-1）曾经说过我们的世界是由理念和现象构成的，理念就是现实世界本身，而现象则是我们人类对现实的理解，人们是通过现象来感知整个世界的，而由于人类认知的差异性和局限性，所以人们对世界的理解也不同，如果非要用一句话来概括，那就是——人生其实就是一种感觉。

还有一种人，我们称其为科学家，他们经常研究的问题就是：

生命是由什么构成的？

物种演化的规律是什么？

人为什么会死？

4 人工智能与本能

生物是怎么繁衍后代的？

基因突变的原理是什么？

如何优生优育？

怎样才能成为超人？

人类是否可以长生不老？

.....



图 1-0-1 拉斐尔的《雅典学院》，正中间单根手指指天的就是柏拉图

我们把研究这种问题的人称为生物智能研究者。其实为了交流方便，我们这里谈到的科学家们更多的是生命科学家或生物学家。这些科学家通过研究生物的结构、功能、发生和发展的规律来告诉我们生命究竟是怎么回事。大家不要以为生命科学家非常神秘，其实我们每个人都都在无时无刻地研究着生物智能，只不过是当局者迷罢了。

英国的达尔文先生（见图 1-0-2），就是这样一位集大成者，他是英国生物学家、进化论的奠基人，他最为著名的代表作是《物种起源》，书中详细阐述了进化论的观点，也就是优胜劣汰，物竞天择。

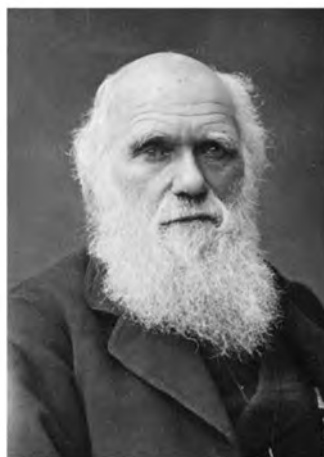


图 1-0-2 查尔斯·罗伯特·达尔文

最后一种人，我们称其为工程师，他们整天和机器打交道，有事没事的时候就想：

机器是如何工作的？

如何让机器工作得更有效率？

怎样制造一个完美的机器人？

是否可以用机器代替人类工作？

机器能不能像人一样拥有情感？

人和机器人要怎样才能更好地结合在一起？

.....

我们把研究这种问题的人称为人工智能研究者，人工智能研究者也是本书主要的一部分读者，现在世界上很多优秀的工程师们正在试图尝试用冷冰冰的机器创造出有感情生命，这已然是一副造物主的架势。其实我们自从会使用工具的那一天起，就已经是一名人工智能的研究者了。

6 人工智能与本能

其中，美国科学家、工程师马文·明斯基（见图 1-0-3）先生更是现代人工智能的奠基者，被誉为人工智能之父。



图 1-0-3 马文·明斯基

第 1 章 宇宙智能

宇宙究竟是一种伟大的机制、一种复杂的计算、一种精美的对称形式、一个巨大的意外，还是一种伟大的思想？关于宇宙智能，我们这里只讲 3 个词就够了，这 3 个词一共 12 个字，分别是无中生有、史前文明和循环迭代，先给出一个它们之间的关系正图，如图 1-1-1 所示。



图 1-1-1 无中生有、史前文明、循环迭代

1.1 无中生有

小时候，我经常问妈妈的一句话：“妈妈我是从哪里来的啊？”妈妈经常会说：“你是从垃圾堆里捡来的。”于是我又会去问：“我为什么会出现在垃圾堆里呢？”……

慢慢地长大以后，我开始对周围的事物很好奇，于是我不停地问，这个东西是从哪里来的，那个东西是从哪里来的？以至于我终于有一天问出了“宇宙是怎么来的？”

宇宙是怎么来的，可能现在还没有人给出一个确定的答案，唯一还算是被广泛认可的就是大爆炸理论。

大爆炸

大概 100 年前，美国天文学家埃德温·鲍威尔·哈勃（见图 1-1-2），通过大量的观测和计算发现了星体红移现象，所谓的红移现象就是当一个光源远距离离开我们时，它的颜色会变得偏红一些。而距离越远颜色越红，与之相反的就是蓝移现象，蓝移是指光源离我们越来越远的时候，它的颜色会变得偏蓝一些。



图 1-1-2 埃德温·鲍威尔·哈勃

哈勃研究发现，所有的恒星都有红移现象，也就是说所有的恒星都在以一定的速度向着远离我们的方向移动。

到了 1929 年，哈勃通过将红移和视亮度的比较，逐步计算出星系的红移与它们到我们的距离成正比的关系，这个关系就是著名的哈勃定律。

哈勃说：不管你往哪个方向看，远处的星系正急速地远离我们而去，而近处的星系正在向我们靠近。

简单来说就是，遥远的星系均远离我们地球所在的银河系而去，同时它们的红移随着其距离增大而成正比地增加。

这种感觉就好像星系受到了一次剧烈的爆炸一样（宇宙来源于一次大爆炸），从而不断奔向四面八方。有关“宇宙大爆炸”的观点也由此开始形成。

这意味着，在很早以前，各个星体相互之间离得非常近。有人计算过并且还给出了一个大概的时间表：

大约 150 亿年前，宇宙还是一个体积无限小，密度无限大，温度无限高，时空曲率无限大的点，这个小点就是我们常说的奇点。

[illegible]

大爆炸后 10^{-35} 秒：强力、电弱力和引力开始形成。

大爆炸后 10^{-12} 秒：宇宙温度仍然高达 10 万亿摄氏度，这时质子和中子及其反粒子开始形成，玻色子、中微子、电子、夸克以及胶子等稳定下来。

大爆炸后 13 秒后：宇宙温度降低到大约 30 亿摄氏度，最基本的化

学元素开始生成，如氘和氦。

大爆炸后 35 分钟后：宇宙温度降低到大约 3 亿摄氏度，原初核合成过程停止。

大爆炸后 104 年后：宇宙物质已经基本形成了，然后就一直慢慢地生成到我们现在看到的这个样子。

质能方程

大约 46 亿年前，由一片巨大分子云中逐渐相互吸引形成了太阳系，其中，最大的部分形成了一个核心，这个核心就是太阳，其余部分继续摊平并形成了一个原始的行星盘，继而再形成了行星、卫星、陨星和其他小星体等。

太阳不停地发光发热，每时每刻都向外辐射出巨大的能量，它就像一个巨大的大火球，给地球带来光和热。可就算如此，我们地球所接受到的太阳能，也不过是太阳全部辐射能的二十二亿分之一。太阳为什么能够持续地发光发热，直到 1938 年人们通过原子核反应的原理才最终解开了这个谜题。

原来太阳含有极为丰富的氢元素，这些氢原子核在太阳中心的高温（大约 1500 万摄氏度）、高压环境下，结合成氦原子核，同时释放出大量的能量，并以光和热的形式辐射出来。

有人统计说：太阳平均每秒钟通过氢-氦核聚变所产生的能量相当于几百万颗原子弹那么大，而且这种聚变至少还会持续 50 亿年，毕竟太阳相当大。

学过初中化学的都知道，氢和氦明明都是气体元素，它们怎么还可以相互转变呢？好吧，就算是它们可以相互转化，但是为什么氢转变成

氢还会释放出能量呢？而能量和气体根本就是完全不同的一种东西。

著名科学家爱因斯坦（见图 1-1-3）结合前辈的经验给出了一个合理的解释，这个解释就是原子核聚变。核是指由质量小的原子，主要是指氘，在一定条件下（如超高温和超高压），只有在极高的温度和压力下才能让核外电子摆脱原子核的束缚，让两个原子核能够互相吸引而碰撞到一起，发生原子核相互聚合作用，从而生成新的质量更大的原子核（如氦），中子虽然质量比较大，但是由于中子不带电，因此它能够在这个碰撞过程中逃离原子核的束缚而被释放出来，大量电子和中子的释放所表现出来的就是巨大的能量释放。

与核聚变相反的核反应形式是核裂变，核裂变就是一个大质量的原子核分裂成两个比较小的原子核，在这个变化过程中同样会释放出巨大的能量，当然实验证明还是核聚变时释放的能量更大。

为了方便人们计算，损失了多少质量能够换来多大的能量，爱因斯坦给出了著名的质能方程 $E=mc^2$ ， E 是能量，单位是焦耳（J）。 m 是质量，单位是千克（kg），而 c 则表示光速（常量， $c=299792.458\text{km/s}$ ）。

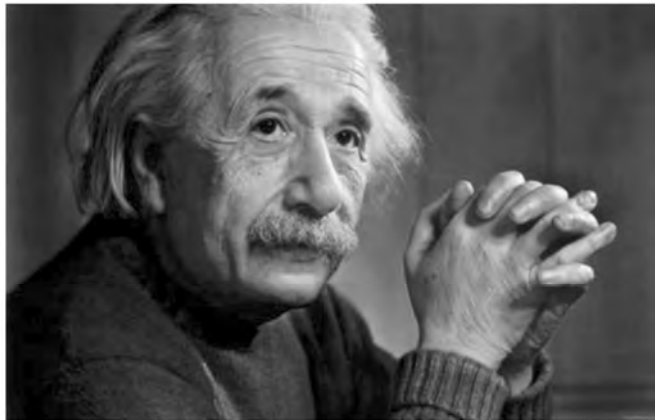


图 1-1-3 爱因斯坦

希格斯玻色子

质量可以转换成能量，那么与之相反的能量是否能够转换成质量呢？1964 年，英国物理学家彼得·希格斯（见图 1-1-4），提出了一种粒子场，并预言说一种能吸引其他粒子进而产生质量的玻色子（一种自旋为零的粒子）的存在。

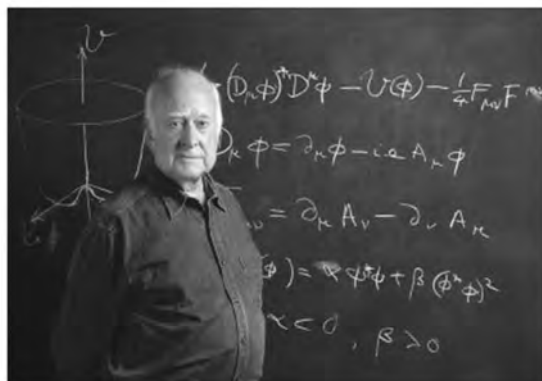


图 1-1-4 彼得·希格斯

他认为，这种玻色子是物质的质量之源，是电子和夸克等形成质量的基础，其他粒子在这种粒子形成的场中游弋并产生惯性，进而形成质量，构筑成大千世界。

这种理论中的粒子后来被别人以“希格斯”的名称命名，外号“God Particle”，意思是上帝粒子（上帝好像就是通过它来创造世界万物的吧）。当其他粒子相继被发现时，在这 48 年中，“上帝粒子”始终遁形。

终于在 2012 年 7 月 4 日，欧洲核研究组织才宣布发现一种与“上帝粒子”“一致”的亚原子粒子。

2013 年 3 月，欧洲核子研究组织再次表示先前探测到的新粒子是希格斯玻色子。

水的三种形态

前面说到了，原子核、粒子、玻色子确实有点太晦涩了，不利于我们理解，其实我们完全可以不像科学家研究得那么高深莫测，因为在我们身边就有很好的例子来说明这一现象（注，也仅仅为了方便说明而已）。

我们就以水为例（见图 1-1-5），小的时候我们就知道，水在 0°C 以下的时候会结冰，在 100°C 以上的时候会变成水蒸气。

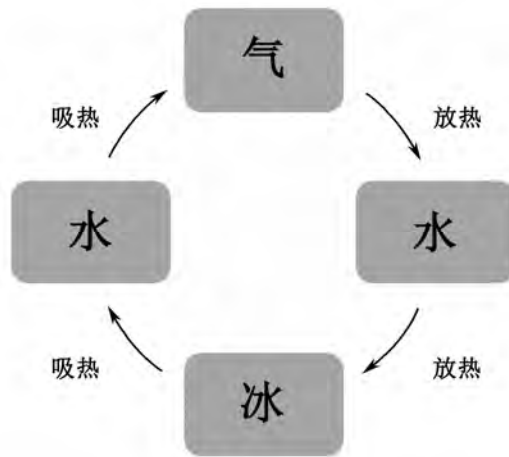


图 1-1-5 水的三种形态之间的转换

冰是一种固体，水是一种液体，水蒸气则是一种气体，水的三种形态各不相同，但是它却在我们的世界上不停地变化着。

冰，我们看得见摸得着，水则千变万化，而水蒸气我们则根本看不见，因为水蒸气其实是一种气体，我们之所以能看见，或者说看见了白色的气体其实是含有小水珠的水蒸气，这种水蒸气也叫“水蒸汽”，一个带三点水，一个不带三点水。

一般来讲，大家都知道，冰变成水，水再变成水蒸气最简单的方法

就是加热，其实除了温度，人们还可以通过减压的方法来让水进行气化，也就是说压力越小水越容易变成水蒸气，由于海拔越高气压越低，所以在青藏高原等高海拔地区，水的沸点是不到 90℃ 的。一般海拔每上升 1000 米，水的沸点就下降 3℃，而高压锅则正好相反，在超过 120℃ 的时候，水还是液体状态的。

不止水是这样，其他气体，甚至金属也是这样。比如液化气、铁和氮气。说液化气大家可能觉得没什么，因为我们经常看见，但是氮气明明是惰性气体，它能够液化和固化就很一般了。

氮气，通常状况下是一种无色无味的气体，而且一般氮气比空气密度小。氮气占大气总量的一半以上，是空气的主要成分。在标准大气压下，氮气在零下 195.8℃ 时，变成没有颜色的液体，继续冷却至零下 209.8℃ 时，液态氮就会变成雪状的固体。

可见，从水变成水蒸气是一个有变成无的过程，而反过来从水蒸气变成水则是一个从无到有的过程。

感知

刚才我们谈到了水蒸气，说水蒸气我们看不到，这里所说的看不到其实是指肉眼看不到，并不是说我们不能通过其他仪器进行观测。仪器的精度决定了我们能够观测到物质的精度。

比如你用皮尺去测量房屋或者操场的面积，这个很正常，但是你用皮尺去测量头发就很难了，而你用游标卡尺可能就比较容易了。

一根头发丝的粗细，用皮尺可能测不准，但是用游标卡尺就能够测得比较准，这说明了我们的认识事物甚至是世界实际上是有一把尺子的，而这个尺子的精度是以人为中心的，当尺子的精度变化时，我们观测到

的世界也随之发生了变化。

每个尺子上都有自己的刻度，同时也有一个最小刻度，正是这样许许多多最小刻度构成了尺子，当被测量的物体小于这个最小刻度时，我们一般忽略不计，只有大于这个最小刻度时我们才开始计数，实际上这个最小刻度就是它的精度，就是它的阈值。

理论上讲，不论是宏观上还是微观上，其实我们的宇宙一直都是可以度量的，只是我们现在的人很难找到这个恰当的测量仪器，不过随着科技的进步，越来越多的观测仪器被制造了出来，相信在不久的将来我们能够看到一个更加真实的宇宙。

也许有人会问，我们以前观测到的宇宙就不真实吗？关于这个问题，很多人可能给出了不同的答案，但是思路却大同小异。

人们认识这个宇宙，其实是通过自己的感觉器官进行感知的，或者说我们的宇宙是一种被我们感知的存在，也就是说我感觉到了它，它才存在，我感觉不到它，它就是不存在的。

对于一个盲人来讲，光明是不存在的；而对于一个色盲来讲，彩色是不存在的。就算是健康人，我们也有很多的光是看不见的。于是科学家就把光线分为可见光和不可见光两部分，其中不可见光又分为红外线和紫外线。人类不借助工具是根本看不到紫外线的。声音也是这样的，甚至同样是人类，在不同的年龄和不同的环境下也会有听力和视力的差别。

那么是不是说宇宙还是由非常小非常小的粒子构成呢？这些粒子小到你现在都观察不到，甚至比夸克都小，小到它就是一个简简单单的引力子呢？这其实还是有可能的，甚至可能我们所谓的魅力、毅力、魄力和想象力都是一种可以度量的力（见图 1-1-6）。

这几年，受到大布尔不等式的影响，量子力学的发展与日俱增。甚至已经开始了量子通信和量子计算。据说量子纠缠的速度是光速的上万倍，甚至有人曾经断言人类的大脑本身就是一个量子计算机。

由此也证实了：色即是空，空即是色！



图 1-1-6 宇宙星云图

1.2 史前文明

神话是一个有趣的话题，也是很多宗教信仰的基础，甚至很多神话本身就是一段著名的历史故事。其中《圣经》作为世界发行量最大，流传最广的书籍之一，既是基督教的经典著作，又是古代中东地区特别是犹太民族的一部详细的编年史。

《圣经》由公元 4 世纪时君士坦丁堡大主教约翰·克里索斯通取名，把犹太人所有的圣书统称为《圣经》。《圣经》中保留众多的犹太民族远古时代的历史传说，且带有神话色彩。

随着科学的深入发展，人们重新审视《圣经》，竟发现在这些神话和传说之中，包含某些超越时代视野的真实记载。

有研究者认为，40 多亿年的地球曾经至少存在过一次史前文明，而且这个史前文明空前的强大，强大到和神仙一样的地步，正是这种文明创造了我们人类和现有世界的各种动植物，比如上帝造人说、女娲造人说。

这些题材也经常会以神话的形式出现在玄幻小说中，玄幻小说最直接的暗示就是：越是古老的东西越强大，因此创世诸神是最强大的。

有些科学家也认为，神话在某种意义上来说其实是存在的，因为假设古代人的智力相对不太发达，那么他们很难有这么丰富的想象力，想象力是高级生物的特征，越高级的生物想象力越丰富，否则除了我们人类之外所有的物体岂不都比我们有梦想吗？

一般来讲，大家对于不解之谜往往倾向于有两种解释，一是外星人访问地球遗留的痕迹，这种遗留包括我们人类；二是现代人类文明之前，曾经出现多次高级人类的史前超文明。不管哪一种解释都说明以前存在过超文明。

有科学家甚至提出地球文明周期进化论。有生物考古学家认为，地球诞生至今的 45 亿年中，地球生物经历了 5 次大灭绝，生死交替，周而复始，最后一次大灭绝发生于 6500 万年之前。有人据此推断，20 亿年前地球上存在过高级文明生物，但不幸毁灭于一场核武战争或巨大的自然灾变。亿万年的沧海桑田几乎抹去一切文明痕迹，仅留存极少遗物，成了现代人类的不解之谜。有人认为，前一次高级文明的毁灭，是因为地球气候的周期性变化，或者地球磁场的周期性消失。太阳系运转至宇宙空间某个特定位置时，地球会周期性出现不适应人类生存的气候。6500 万年前恐龙的灭绝便是例证（见图 1-1-7）。地球的周期性气候变化会导致高级智慧生物的周期性起源和进化。



图 1-1-7 恐龙时代

即便是没有几千万年前那么古老，就算是几千年前也有可能科技一度非常发达。早在 1900 年，希腊潜水员在安蒂基西拉岛附近海底一艘沉没的古罗马货船残骸中，发现了一个神秘复杂的古希腊青铜机械装置，它上面有齿轮和刻度盘。100 多年来，这一被称作“安蒂基西拉机器”的神秘装置有何用途一直众说纷纭。最近，科学家发现原来它堪称是“世界上最早的计算机”。

其实现在想想史前超文明也不是没有可能，比如现在天空中飞行的空间站，未来就很有可能是天宫，我们现在创造的机器人很有可能就是未来地球的主要物种，或许有一天它足够聪明也能够创造出新的生命，就像上帝造人一样，新的创世纪再次开始。也许有人认为我们创造高智能的机器人还有很长的一段路要走，或许你是对的，但是有什么关系呢？上帝比人类聪明和强大，人比机器人聪明和强大，这本身就够了。

由此可见，上帝造人，人造机器，将来机器再造新的生命！

加速回报定律

人类究竟什么时候才能够创造出像人一样够聪明和强大的机器人呢？雷·库兹韦尔在其相关理论的论文中写道：“在科技的早期阶段——轮子、火、石器——费时数万年才慢慢演进、普及。1000 年前，诸如印刷术等典范转移，也耗费约 1 个世纪的时间，才为世人普遍采用。今日，重大的典范转移，如移动电话及全球信息网，则只消数年的时间就普遍大行其道。”

他接着说：“19 世纪那 100 年期间所发生的科技变革，比之前 900 年的变化还大。接下来，在 20 世纪的头 20 年期间，我们目睹的科技进步比整个 19 世纪还多。”

美国发明家雷·库兹韦尔（见图 1-1-8）对人类、半机械人命运的兴趣始于 1980 年，且主要从现实角度来考虑这种前景。当时他需要找到一些用于测量和追踪技术发展速度的方法。如果过于超前，即便最伟大的发现也会遭到冷遇，于是库兹韦尔希望能在最恰当的时候公布自己的发明。“但即便如此，技术改变的速度依然飞快，以致下一秒的世界和你完成设计时的世界已经迥然不同。”库兹韦尔说，“这和双向飞碟射击有点相似——你不能看到目标再开枪。”他了解摩尔定律，该定律认为集成电路上的晶体管数量每隔 18~24 个月就会翻一番。这是一个异常可靠的经验法则，但库兹韦尔试图绘出一条与摩尔定律稍有不同曲线来，他考察了 1000 美元可以购买到的计算能力随时间的变化情况，计算能力则以每秒百万条指令（MIPS）来表示。

结果证明，库兹韦尔的结论与摩尔定律极为相似，两者都是约每两年翻一番。他们绘出的图表皆为指数曲线，结果以 2 的指数倍在增加，而非是规则的线性增量。即便当库兹韦尔将历史向后延伸至晶体管计算

技术还未出现的 20 世纪初——即使用继电器和真空管的几十年间时，曲线依然表现出令人惊异的稳定性。

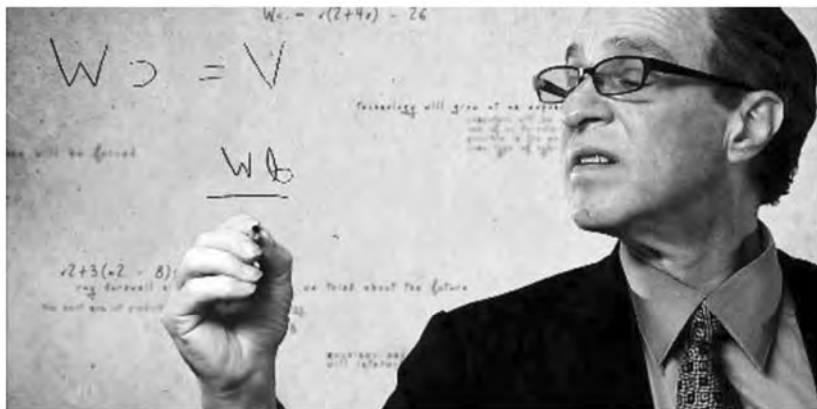


图 1-1-8 雷·库兹韦尔

库兹韦尔随后将这一指标应用到其他的关键技术标准上，如晶体管制造成本的降低、微处理器时钟频率的增加和动态 RAM 价格的暴跌等，他甚至将眼界拓宽至生物技术和其他领域的发展趋势，如 DNA 测序和无线数据服务成本的降低，互联网主机和纳米技术专利数量的增加。在这些发展趋势中，库兹韦尔发现了相同的规律：指数方式的加速发展。“这些轨迹线都呈现出不可思议的平滑状态，”他说。“无论在任何情况下，战争也好，和平也罢，也不管经济是繁荣还是低迷。”库兹韦尔将这种情形命名为加速回报定律，即技术进步是以指数方式发展的，而非线性。

平行的宇宙

也有人认为史前文明不过是平行宇宙中的一个宇宙，由于平行宇宙是指在宇宙中三维坐标相同，四维坐标不同的平行空间，所有的平行空间合起来算作一个宇宙，每一个平行空间在宇宙诞生时就是一直独立发

展的，所以差异都非常大，平行空间往往不像我们某人所想的那样，在那个空间你是看不到和这个世界一样的物种的，因为长期的独立发展使得每个平行空间的差异都非常大，平行空间包括了一切存在和可能存在的事物。

我们可以举一个简单的例子，就是一对夫妇生了一对双胞胎，他们看上去很像却不是一个人，但即便他们将来如何变化，他们的基因却是很相似的，这对双胞胎对于生活在他们身上的微生物来说其实就是一个平行宇宙。

或者我们再举一个影片的例子以方便理解，我们假设胶片上每一张照片都是一个固定的房间，每个房间都是一个只有固定时间锚点的三维世界。

当我们播放影片时，你的世界就活了起来，时间也顺着影片的播放而结束，这段影片就是你的世界。

现在我们打算对影片进行剪辑，可以增加、可以删减也可以置换，这样剪辑后的影片就是一个新的世界。

如果影片足够长，照片足够多，我们的世界也可以足够精彩，电影胶片如图 1-1-9 所示。

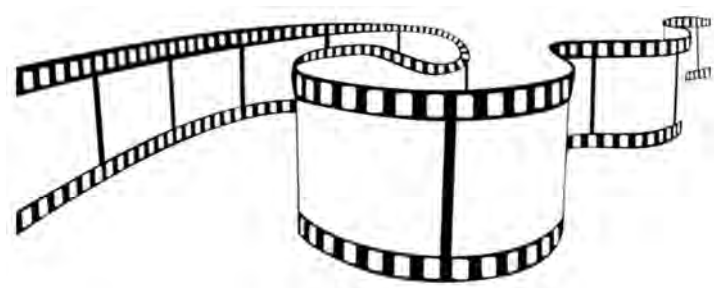


图 1-1-9 电影胶片

1.3 循环迭代

曼德勃罗集合

宇宙对于我们来说既熟悉又神秘，熟悉的原因是我们置身其中，神秘的原因也是因为我们置身其中，没有人知道宇宙外面有什么，也不知道宇宙到底是不是大爆炸引起的，如果是，那么大爆炸之前是什么。

宇宙智能到了近代发展出了一个杰出的分支，这个分支就是数字宇宙，数字宇宙其实就是把整个宇宙想象成一个巨大的计算机，然后把宇宙中所有的物质全部数字化，最后通过数学计算的方法管理和运行这个宇宙。

数字宇宙的思路很好，结果也非常可喜，所以我们经常在天文学的世界里见到很多伟大的数学家。

数字宇宙发展到现在，尤其是电子计算机的出现更使其发展如虎添翼。其实整个计算机尤其是图灵机，就是利用一个有限的循环来计算整个无限的宇宙。

这些科学家们把复杂的宇宙抽丝剥茧，力争变得简化，然后用各种各样的规律、定理、公式展现给大家。

数学上就有这样公式： $Z=Z^2+C$ ，这个公式就是著名的曼德勃罗集合（见图 1-1-10），虽然从整体上看我们感觉很复杂，但当你仔细研究时，你就会发现它的原理很简单，因为整张大图都是由同一张图片不断地迭代出来的。

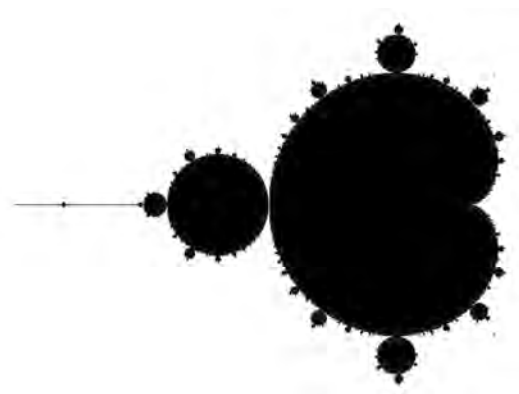


图 1-1-10 曼德勃罗集合 (1)

这就好比你有两面（至少两面）相互照的镜子，然后你在两面镜子中间放一个小球，就会发现小球的镜像在两面镜子里不断地镜像，不知道什么时候就出现了无数个小球。万花筒就是一个很好的例子。

其中 C 是一个常数， $Z=Z^2+C$ 是一个不断迭代的过程。其实曼德勃罗集合最重要的思想就是通过不断迭代可以产生足够复杂的图形，甚至是宇宙。

因此，曼德勃罗集合被认为是上帝的指纹。

现在利用迭代思想，曼德勃罗集合已经有很多的变种，公式也变得更为复杂，有很多科学家已经利用计算机高速计算的能力开始迭代其数字宇宙了，甚至有的已经诞生了生命。

例如，图 1-1-11 所示就是另一个曼德勃罗集合。

目前，我们身边的物体确实符合类似的规律，比如你用显微镜去看雪花的时候，你就会发现，雪花中还有雪花。甚至有用观察星系的方法观察原子，或者用原子的方法来计算星系，这也是曼德勃罗集被称为上帝指纹的主要原因。

由此见证了：复杂的事情是可以由简单的事情不断迭代构成的。



图 1-1-11 曼德勃罗集合 (2)

二进制与电子计算机

研究数字宇宙的秘密武器就是电子计算机。20 世纪 60 年代，约翰·冯·诺依曼就曾经研发了一台名为 ENIAC (Electronic Numerical Integrator And Computer) ——电子数字积分计算机来模拟核爆、天气预报以及宇宙的运行。图 1-1-12 展示的就是电子数字积分计算机 (ENIAC) 的应用场景。



图 1-1-12 电子数字积分计算机的应用场景

在计算机出现以前，人们很难用有限的集合来表示无限的集合，就算同样是无限集合，人们往往也会认为实数比偶数多，毕竟实数是包含偶数的。

但当等势的概念出现后，人们发现同样两个数字集可以通过一一映射的方法使其相等。

以至于看似只有 0 和 1 的两位数集合也可以通过不断地组合来表示无限的数据集。二进制的发展对电子计算机的发明可算是居功甚伟。没有二进制可以说就没有我们现在的电子计算机。

二进制是计算技术中广泛采用的一种数制。二进制数据是用 0 和 1 两个数码来表示的数。它的基数为 2，进位规则是“逢二进一”，借位规则是“借一当二”，是由 18 世纪德国数理哲学大师莱布尼茨（见图 1-1-13）发现的。当前的计算机系统使用的基本上是二进制系统，数据在计算机中主要是以补码的形式存储的。其实计算机中的二进制就是一个个非常微小的开关，用“开”来表示 1，“关”来表示 0。



图 1-1-13 戈特弗里德·威廉·莱布尼茨

二进制应用非常广泛，任何一个文明都或多或少地使用了二进制的思想，比如中国的易经和八卦，所以有些人认为莱布尼茨其实是发现了二进制，而不是发明了二进制。

数学需要抽象

前面说到曼德勃罗集是一个不断迭代的过程，但是不论怎么迭代它都有一个核心的公式，那么这个公式是怎么来的呢？

其实从物理学来看这些公式都是不存在的，而是人们的一种抽象，一种假设，或许人类只是发明了自然数，其他的是在这些自然数的基础上衍生出来，而这些自然数也是我们认为的一种假设而已。比如 1 就是 1，2 就是 2 个 1，等等。

从我们有了实数开始，就有了基本运算，基本运算有一套基本运算的规则和方法。然后我们开始用字母表示数，开始解方程，解方程有解方程的规则和方法。然后我们对方程的各种情况开始研究，出现了函数，研究函数有研究函数的法则和方法。然后我们又把函数从简单的数值关系扩展到集合上，出现了泛函，泛函又有新的规则和方法。注意这每一步的发展都会出现新的定义来囊括更一般性的东西，也会出现和从前不同的方法。也就是说，数学家每在一定范畴内掌握了一定数量的规律之后，就开始寻找一个在更大的范畴内适用的新规律，这个“更大的范畴”是由数学家自己定义的，定义好了之后考虑一些基本的公理，再开始寻找规律，找到了足够多的规律之后，就去找一个更大更大的范畴……

数学虽然是一种假设，一种抽象，但是它却能指导我们的科学不断进步，指挥我们的设备完美地运行。

其实数学并没有我们想象的那么难，我们完全可以在他人的基础上来做一些有趣的计算，你或许会发现意想不到的结果，看看下面的几个例子。

例 1 幂的简便运算。

第一步：将较大的底数变成较小的底数。

$$4^3=2^3 \times 2^3=2^{3+3}=2^6$$

第二步：将两个相同底数指数相同的数相加。

$$2^3+2^3=2 \times 2^3=2^1 \times 2^3=2^{1+3}=2^4=16$$

第三步：两个同底数相加。

$$2^3+2^2=2^2 \times 2^1+2^2=2^2(2^1+1)=4 \times 3=12$$

$$2^3+2^5=2^3+2^3 \times 2^2=2^3(2^2+1)=8 \times 5=40$$

第四步：多个同底数相加，两两相加，或者提取公共项。

$$2^3+2^4+2^2=2^2 \times 2^1+2^2 \times 2^2+2^2=2^2(2^2+2^1+1)=4 \times (2^2+2^1+1)=4 \times 7=28$$

这里用到了一个通指的公式，就是我们小学学过的结合律，当然你也可以使用最大公约数或最小公倍数来做出更加有趣的练习，指数的减法也是这个道理。

例 2 长方形与内切椭圆形的关系。

长方形与内切椭圆形面积之比= $4/\pi \approx 1.273$;

长方形与内切椭圆形周长之比= $4/\pi \approx 1.273$ 。

可以延伸到长方体与内切椭圆体的表面积和体积之间的关系。

长方体体积与内切椭圆体体积之比= $6/\pi \approx 1.91$;

长方体表面积与内切椭圆体表面积之比= $6/\pi \approx 1.91$;

$(\text{长方体表面积}/\text{长方体周长}) / (\text{长方体体积}/\text{长方体面积}) = 1;$

.....

以此类推，我们还可以算出内切圆柱体和圆锥体与它们之间的关系。

例 3 在二维的平面上绘制三维设计图。

一般来讲我们很难在一个二维的平面上绘制出一个三维的图像，但是我们换个角度多几张图纸，却可以实现一个三维表格的绘制，其中每一张图纸都相当于建筑物的一层，这种表格绘制简单方便，立体感强，非常适合做施工图和电路图，如图 1-1-14 所示。

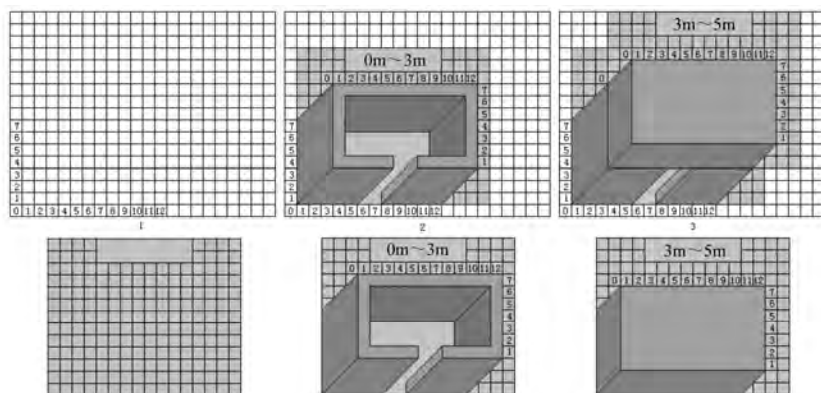


图 1-1-14 三维平面制图法

例 4 最大公约数与最小公倍数之间的关系。

两个自然数的最大公约数与最小公倍数的乘积等于这两个数的乘积；三个自然数的最小公倍数与最大公约的乘积等于这三个数的乘积除以它们中间的那个数（既不是最大的一个也不是最小的一个）；如果是四个数的话就是中间两个数的乘积……以此类推。这个方法对开方很有作用！

例 5 不怕被偷窥的心灵密码。

现在的密码要么是一串复杂的字符，要么是一个生物特征，其实不论是生物特征还是字符串密码都害怕被人偷窥。比如你的 QQ 账户和密码被黑客知道了，恐怕你就没有办法找回了，怎么办呢？其实我们可以做一个小的改动，在你登录你的用户名时，系统随机出现一些可以计算的字符，如 123456，然后你看到这些提示符后，用最后一个字符乘以 3 并保留个位，然后把结果 8 输入密码框。这个 8 就是你的动态密码了。怎么样，原理很简单吧，就算被别人看到了，他也很难知道你的真正密码是最后一个数字乘以 3 并保留个位。

互为镜像

一个小小的二进制居然可以计算天气预报、原子核聚变，简直不可思议，其主要原理其实就是利用了循环，让只有 0 和 1 的数据集不断地被使用。

举一个小例子，比如我们想跑 10 000 米，但是场地又有限，那么我们怎么办呢？其实我们完全可以在 400 米的操场跑上 25 圈就可以了，或者你在跑步机跑上 10 000 米也是可以的。

我们的世界或许就是由一个个一层层的小圈（弦）不断地循环和重复构成的。虽然这个小圈不是严格意义上的椭圆形，但是它在运行的过程中确实是遵从数学定理的。

这个世界虽然现在看着复杂，但是其运转机制本身可能并不复杂，或许我们这个世界本身就是一个在不断迭代中螺旋式进步的楼梯。或许它就是微观和宏观两面互为镜像的镜子（见图 1-1-15），我们既可以通过微观来预测宏观，也可以通过宏观来观察微观。



图 1-1-15 万花筒

真应了一句禅语：一花一世界，一叶一菩提。

第 2 章 生物智能

2.1 生物学

通俗地讲，生物学是研究生物的功能、结构、发生和发展的规律，也就是说，生物学基本上就是研究这个生物为什么会动。生物学尤其是生命科学，在最近几十年取得了长足的发展，随着 DNA 的发现、克隆技术的成熟、转基因作物的普及，生命科学正在改变人们的世界观，特别是基因技术。

DNA

DNA 也叫作脱氧核糖核酸，或者去氧核糖核苷酸，是染色体的主要组成成分，同时也是主要遗传物质。它是一种生物大分子，可以组成遗传指令，引导生物发育与生命机能运作。也就是说，实际上是 DNA 控制

着我们的生长与疾病。

研究发现，父代把它们自己 DNA 的一半复制传递到子代中，从而完成性状的传播。原核细胞的染色体是一个长 DNA 分子。真核细胞核中有不止一个染色体，每个染色单体也只含一个 DNA 分子。不过它们一般都比原核细胞中的 DNA 分子大而且和蛋白质结合在一起。

DNA 分子的功能是储存决定物种性状的几乎所有蛋白质和 RNA 分子的全部遗传信息，编码和设计生物有机体在一定的时空中有序地转录基因和表达蛋白完成定向发育的所有程序；初步确定了生物独有的性状和个性以及和环境相互作用时所有的应激反应。除染色体 DNA 外，有极少量结构不同的 DNA 存在于真核细胞的线粒体和叶绿体中。DNA 病毒的遗传物质也是 DNA，极少数为 RNA。

在 DNA 被确认为遗传物质之后，生物学家不得不面临着一个难题：DNA 应该有什么样的结构，才能担当遗传的重任呢？

- (1) 它必须能够携带遗传信息。
- (2) 能够自我复制传递遗传信息。
- (3) 能够让遗传信息得到表达以控制细胞活动。
- (4) 能够突变并保留突变。

1953 年 4 月 25 日，克里克和沃森在英国杂志《自然》上公开了他们的 DNA 模型。经过在剑桥大学的深入学习后，两人将 DNA 的结构描述为双螺旋，在双螺旋的两部分之间，由四种化学物质（腺嘌呤、胸腺嘧啶、胞嘧啶、鸟嘌呤）组成的碱基对扁平环联结着。

其实本质上 DNA 是由碳（C）、氢（H）、氧（O）、氮（N）、磷（P）这五种基本元素构成的对稳定的分子双螺旋结构的。之所以能够构成相

对稳定的结构，这是因为在 DNA 分子双螺旋结构的内侧，通过氢键形成的碱基对，使两条脱氧核苷酸长链稳固地并联起来。另外，碱基对之间纵向的相互作用力也进一步加固了 DNA 分子的稳定性。各个碱基对之间的这种纵向的相互作用力叫作碱基堆积力，它是芳香族碱基电子间的相互作用引起的。现在普遍认为碱基堆积力是稳定 DNA 结构的最重要的因素。再有，双螺旋外侧负电荷的磷酸基团同带正电荷的阳离子之间形成的离子键，可以减少双链间的静电斥力，因此对 DNA 双螺旋结构也有一定的稳定作用。

DNA 分子由于碱基对的数量不同，碱基对的排列顺序千变万化，因此构成了 DNA 分子的多样性。不同的 DNA 分子由于碱基对的排列顺序存在着差异，因此，每一个 DNA 分子的碱基对都有其特定的排列顺序，这种特定的排列顺序包含特定的遗传信息，从而使 DNA 分子具有特异性。

虽然 DNA 的结构是非常稳定的，但是我们都知为了适应新的世界，很多生物都选择通过基因突变的方式来从根本上改变自己。

基因突变是指基因组 DNA 分子发生的突然的、可遗传的变异现象，基因虽然十分稳定，能在细胞分裂时精确地复制自己，但这种稳定性是相对的。在一定的条件下基因也可以从原来的存在形式突然改变成另一种新的存在形式，就是在一个位点上，突然出现了一个新基因，代替了原有基因，这个基因叫作突变基因。

基因突变是由于 DNA 分子中发生碱基对的增添、缺失或替换，而引起的基因结构的改变。一个基因内部可以遗传的结构的变化，又称为点突变，通常可引起一定的表型变化。广义的突变包括染色体畸变，狭义的突变专指点突变。

实际上畸变和点突变的界限并不明确，特别是微细的畸变更是如此。野生型基因通过突变成为突变型基因。

基因突变通常发生在 DNA 复制时期，即细胞分裂间期，包括有丝分裂间期和减数分裂间期；同时基因突变和脱氧核糖核酸的复制、DNA 损伤修复、癌变和衰老都有关系，也正是如此基因突变才成为生物进化的重要因素之一。

也许有人会问，既然 DNA 的分子结构非常稳定，那么究竟是什么原因促成 DNA 的分子结构发生了变化了呢？人们普遍认为可能有以下 3 个影响因素。

1. 外因

比如物理因素：X 射线、激光等。

比如化学因素：亚硝酸和碱基类似物等。

比如生物因素：病毒和某些细菌等。

2. 内因

DNA 复制过程中，基因内部的脱氧核苷酸的数量、顺序、种类发生了局部改变从而改变了遗传信息。

3. 亚原子

亚原子也叫次原子，所谓亚原子，泛指比原子更小的粒子，有一种学说认为，DNA 链的键中间有一个亚原子，我们可以把它想象成一个卡槽，由于这个亚原子是不确定（不确定性）的，位置忽左忽右。当本来应该在左边的亚原子出现在右边的时候，DNA 链的键的形状也就随之发生了改变，也就是说卡槽的形状发生了改变，当它发生了改变，与之配对的另一个链或卡槽也只好随之发生改变，否则无法配对成功。

DNA 分子结构如图 1-2-1 所示。

DNA 的双螺旋结构

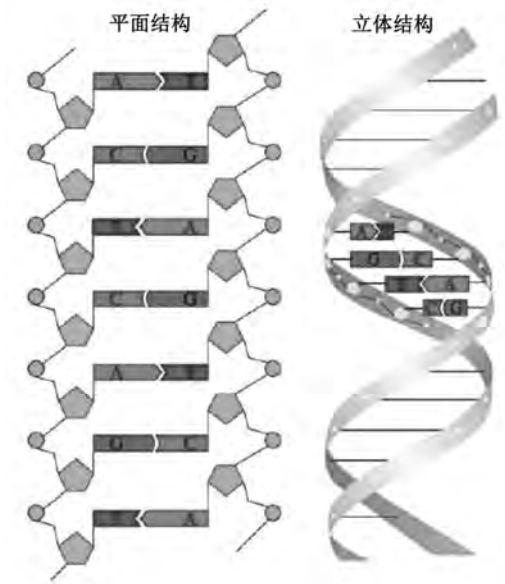


图 1-2-1 DNA 分子结构

进化论

生物学，尤其是西方科学，有一个很重要的学说就是关于物种演化的，特别强调自然选择的作用。笔者个人比较倾向于演化论，演化论相对进化论而言言语更加严谨和温和，因为并非什么演化都是进步的。即使是科学发达的现在，科学家也不能严谨地证明现在的人就一定比古代的人先进，但是我们却可以普遍赞同存在即合理的这个思想。

进化论是指研究生物史中生命扩大适应范围，增加生存余地的发展状况的科学结论。随着进化论的发展，产生了现代综合进化论，而现代进化学绝大部分以达尔文的进化论为指导，埃尔温·薛定谔的《生命是什么》为主体方向，进化论已为当代生物学的核心思想之一。其进化论

有三大经典证据：比较解剖学、古生物学和胚胎发育重演律。进化论除了作为生物学的重要分支得到重视和发展外，其思想和原理在其他学术领域也得到了广泛的应用，并形成许多新兴交叉学科，如演化金融学、演化证券学、演化经济学等。

(1) 种群是生物进化的基本单位。生物进化的基本单位是种群，不是个体。种群是指生活在同一区域内的同种生物个体的总和。一个物种通常包括许多分布在不同地点的种群。每个种群中的个体具有基本相同的遗传基础，但也存在一定的个体差异，所以种群一般具有杂种性，杂种性的存在意味着等位基因的存在。一个种群中能进行生殖的生物个体所含有的全部基因，称为种群的基因库。其中某一基因占全部等位基因数的比率，称为基因频率。

种群的基因频率若保持相对稳定，则该种群的基因型也保持稳定。但在自然界中种群基因频率的改变是不可避免的，于是基因型也逐渐变化。

(2) 突变为生物进化提供材料。突变引起基因频率的改变是普遍存在的。当然，突变发生的自然频率是相当低的，例如，改变染色体数目（降低后代繁殖率）的染色体平衡易位在人类的发生率只有 $1/500 \sim 1/1000$ ，种群是由大量个体组成的，每个个体具有成千上万基因，这样，每一代都会产生大量的变异。

突变的结果可形成多种多样的基因型，使种群出现大量可遗传变异。这些变异是随机性的，不定向的，能为生物进化提供原料。但突变大多数有害，也就是说大多数突变从人类演化中看都是一种试错的过程。

(3) 自然选择主导着进化的方向。突变的方向是不确定性的，一旦产生，就在自然界中受到选择的作用。自然选择不断淘汰不适应环境的类型，从而定向地改变种群中的基因频率向适应环境的方向演化。也就

是说，自然选择不断地调整着生物与环境的关系，定向地改变种群的基因频率。

（4）隔离是物种形成的必要条件。隔离使不同物种之间停止了基因交流，一个种群中所发生的突变不会扩散到另一个种群中去，使不同的种群朝不同的方向演化。隔离一般分为地理隔离和生殖隔离两类。地理隔离是由于某些地理障碍而发生的。大河、大山、沙漠、海峡和远距离都能将种群阻隔开来，使他们之间彼此不能往来接触，失去了交配的机会。

长期的地理隔离使两个种群分别接触不同的环境，各自积累了变异。另一方面，长期的地理隔离使相互分开的种群断绝了基因交流，结果导致了生殖隔离。生殖隔离是指进行有性生殖的生物彼此之间不能杂交或杂交不育。生殖隔离又可分为受精前的生殖隔离和受精后的生殖隔离。生殖隔离一旦形成，原来的一个物种的种群就变成两个物种的种群了。

2.2 医学

生命之旅

人的平均寿命在 70~80 岁之间，在这平均不到一百年的时间里，我们人类完成了一趟奇妙的生命之旅，虽然每个人的故事可能略有不同，但是从生命的角度来说却大同小异。

虽然现在的科技很先进，但是我们仍然摆脱不了肉体凡胎。一个小小的受精卵在娘胎里慢慢地长大，经过将近 10 个月的漫长而又快速地生长后，才将人们生养下来，正式出生之后才开始在懵懂之中茁壮成长。

人体是一个奇妙的生物机器，有些器官和功能自出生之日起就已经

发育成熟，如呼吸、吮吸、眨眼和游泳。

还有些器官虽然在胎儿时期就开始工作，但更多的时候是需要出生后才能逐渐成熟的，比如听觉，虽然胎儿时期就有了听力，也能对成人对他说话的声音做出反应，但是明显要等到孩子满月以后才能够做出让父母惊讶的反应。孩子随着月龄的增长，听觉能力也逐步提高，到 2 个月时，已能听出声音的方向，能安静地倾听周围的声音，到了 3 个月时，听力有了明显的发展，在听到声音后，头能转向声音的方向，并表现出极大的兴趣。

和听觉相比，视觉虽然发展较晚，但是视觉却在人类未来发生着举足轻重的作用，一般正常人 70% 以上的信息是由视觉构成的。

孩子刚出生时，其实对光线就会有反应了，但是由于眼睛发育并不完全，视觉结构、视神经尚未成熟。他能追着眼前的物体看，但视野只有 45° 左右，视力也只有成人的 $1/30$ 。而且只能追视水平方向和眼前篮球大小的人或物。新生儿偏爱注视较复杂的形状和亲人的脸庞，以及鲜明的对比色。大概在孩子满月后，已具有注视能力了，一般会注视抱他的人，不过还是无法持续太久，眼球容易失去协调性。3 个月后宝宝的视野已经可达 180° ，可看到自己的手，也能够感觉到“自己身体”的存在，比如他可以对镜子修整自己的衣服。

孩子出生半年以后，视网膜已有很好的发育，能由近看远，再由远看近，并且有了立体的感觉，甚至能够配合手去抓东西。

1 岁的时候孩子已经可以判断事物的远近，并且快速追踪移动的物体，等到 3 岁的时候视觉基本发育成熟，这时候的孩子的视力已经和我们成人差不多了。

小孩大概 3 岁的时候，可以从环境里学会每一项生存必备的技能，包括走路、自我意识和说话。

5 岁的孩子基本上就可以掌握一门语法结构了，当然这里的语法结构指的是母语。

8 岁的孩子大脑已经基本发育成熟，等到 10 岁的时候淋巴系统器官，如淋巴结、胸腺等免疫器官逐渐发育成熟。

而到 16 岁时，几乎就是一个大人了，这时候的身体器官虽然仍然在继续发育，但是速度明显缓慢。人类大脑则几乎要等到 28 岁才算完全发育成熟。

35 岁以后，人体开始逐渐退化，不过随着科技的进步，人体的退化时间开始变得越来越晚。

这大概就是每个人可能面临的情况，但并不是说所有人发展的时间表都一致，个别人的个别器官有可能发展得不一样。这种差异性构成了多种多样的人类。

细胞

当我们研究一个人是由什么构成的时候，大致会想到一些主要的器官，当我们继续研究这些器官是由什么构成的时候，可能有些人会告诉我们是细胞。

细胞就像一个个小机器人一样，它们每时每刻，都会在你的身体中不厌其烦地工作直至死亡。这场起源于几十亿年前的工作现今仍然在我们每个人的体内发生。

细胞的种类和功能各不相同，你可以把自己的身体想象成一个大城市，城市里面该有什么你的身体里面就有什么，比如你的身体里面也有白细胞这样的警察，也有病毒细胞这样的破坏者，同时还有司机这样的运输细胞，甚至还有专门处理“互联网信息”的神经细胞。每个细胞都

很重要，缺一不可。比如有一种类似齿轮的葡萄糖转运细胞，就是专门负责把营养中的葡萄糖提取出来并输送给其他器官（主要是大脑），如果没有这些葡萄糖转运细胞，那么人的大脑就会陷入瘫痪，直至死亡。

如果你的想象力足够丰富，完全可以把你的现实世界搬到你的身体里面去，因为人体这个世界真的是太伟大、太神奇了，尤其是在我们以一个细胞的角度进行观察的时候。

或许你会说，既然如此，除了身体里的好细胞外，为什么还有那些病毒细胞呢？病毒——一个神奇的存在，看似没有生命，事实上却并非如此，相反它目标明确，战斗力顽强，有团队合作的战斗模式，利用敌人，充满了智慧。它一步步进入细胞，目的只有一个，控制细胞为自己所用，这个最自私的生物用尽一切办法来达到自己的目的——破坏细胞，同时将自己的遗传物质传递下去，繁衍自己的后代。

当然它们的目的往往没那么容易就达成，因为人体有相较之下更加完备的免疫系统来对抗它。从抗体到吞噬细胞，再到细胞内的免疫物质，精准应对，面对病毒敌人的层层侵入，免疫系统分阶段地战争，一批批地消灭敌人，更容易赢得战争的胜利。即使有的时候让狡猾的病毒有可乘之机，但绝大多数病毒难逃免疫系统的反击，最终的绝大多数战争以病毒失败而告终。正邪的对抗从生命的最初就已开始，病毒与细胞之间的战斗永不停息。病毒与细胞之间的这场对抗经过了漫长的岁月，这样的对抗塑造了我们，促成了我们的进化。双方都在不断提升自己，而这就是一种共同演化。

其实从某种意义上说病毒就是一种生命，正是由于有了病毒的存在，我们才能够让自己的身体不断适应新的环境。细胞和人一样也是在这种不断的对抗中成长起来的。

细胞像人一样也有生老病死，每个细胞都在不停地进行着新陈代谢，

在一定的范围内可以不断地生长。细胞的出生是通过分裂来进行的，那么活细胞为什么不能无限地长大而需要进行分裂呢？

这是因为细胞的生长要受到细胞核与细胞质以及细胞表面积与体积比限制，细胞要通过它的表面不断地和周围环境或邻近细胞进行物质交换，这样它就必须有足够的表面积，否则它的代谢作用就很难进行。当细胞的体积由于生长而逐步增大时，细胞表面积和体积的比例就会变得越来越小，导致表面积不够，使细胞内部和外界的物质交换适应不了细胞的需要，这就会引起细胞的分裂，以恢复其原来的表面积与体积的适宜比例。

再者，细胞分裂可维持细胞核与细胞质体积之间的平衡状态，细胞质中的生理、生化过程都受到细胞核中遗传信息的指引和控制。因此，当细胞质的体积增长太大时，细胞核对这样大范围的细胞质的调控作用就会相对地减少，以致造成核、质的不平衡，从而引起细胞分裂，以恢复其细胞核与细胞质体积之间的稳定状态。有人曾做过这样的实验：当人工培养的变形虫快要分裂的时候，把它的细胞质切去一大块，这个变形虫就不再分裂。等它长大又要分裂的时候，又切去一块，它也不再分裂，但如果让其继续生长，体积达到一定大小时，它又会分裂开来。

神经细胞

讲到生物智能，几乎所有人都绕不开神经细胞，神经细胞主要包括神经元和神经胶质细胞。虽然神经元形态与功能多种多样，但结构上大致都可分成细胞体和突起两部分。突起又分树突和轴突两种。轴突往往很长，由细胞的轴丘分出，其直径均匀，开始一段称为始段，离开胞体若干距离后是髓鞘，也叫神经纤维，习惯上把神经纤维分为有髓纤维与无髓纤维两种，实际上所谓的无髓纤维也有一薄层髓鞘，并非完全无髓鞘。

神经元是一种高度分化的细胞，是神经系统的基本结构和功能单位之一，它具有感受刺激和传导兴奋的功能。

通常认为神经元是高等动物神经系统的结构单位和功能单位。神经系统中含有大量的神经元，据估计，人类中枢神经系统中约含 1000 亿个神经元，仅大脑皮层中就约有 140 亿个。

神经元描述：神经细胞呈三角形或多角形，可以分为树突、轴突和胞体这三个区域。

胞体的大小差异很大，小的直径仅为 5~6 微米，大的可达 100 微米以上。突起的形态、数量和长短也大不相同。树突多呈树状分支，它可接受刺激并将冲动传向胞体；轴突呈细索状，末端常有分支，称轴突终末，轴突将冲动从胞体传向终末。通常一个神经元有一个至多个树突，但轴突只有一条。神经元的胞体越大，其轴突越长。

神经元是大脑的主要成分，神经元通过各个神经元的信息交换，实现大脑的分析与输出功能。

神经元按照用途分为三种：输入神经、传出神经和连接神经。它们之间一般有以下分工。

(1) 输入区，就一个运动神经元来讲，胞体或树突膜上的受体是接受传入信息的输入区，该区可以产生突触后电位。

(2) 整合区，众多的突触后电位在此发生总和，并当达到阈电位时在此首先产生动作电位。

(3) 冲动传导区，轴突属于传导冲动区，动作电位以不衰减的方式传向所支配的靶器官。

(4) 输出区，轴突末梢的突触小体则是信息输出区，神经递质在此

通过胞吐方式加以释放。

其实神经细胞之间的信息传递是一种生物电，主要是通过 Na^+ 进行的，但是它的传播的速度却比电子的速度慢了不少，为 $5\sim 120\text{m/s}$ 。

总的来说，人的整个发育过程就是一个神经系统不断完善的过程，人类虽然可以没有某个具体器官，但是不能没有神经系统。

神经细胞如图 1-2-2 所示。

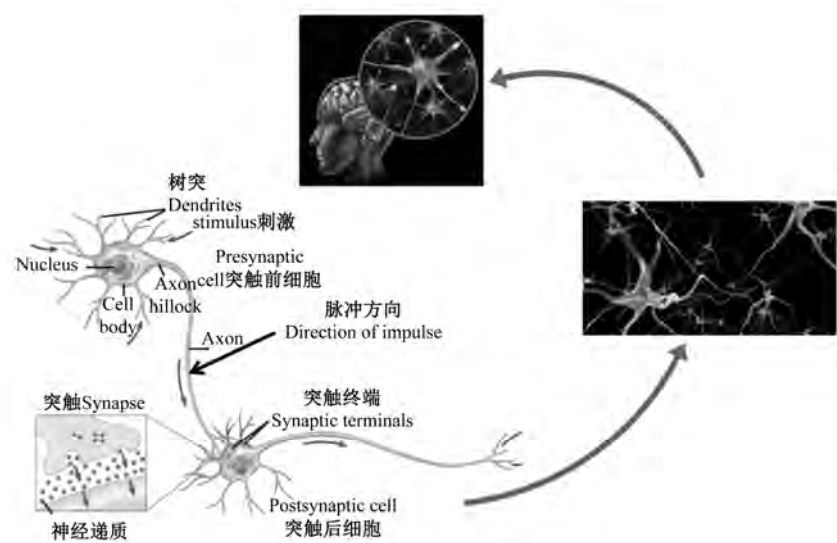


图 1-2-2 神经细胞

2.3 社会学

法律、政治、经济、文化、心理学、语言等都是围绕如何服务于人类的科学，科学一般都具有严格的统计意义。

在这里我们简单介绍一下教育心理学，提出教育心理学化的人是克斯坦罗琦。捷克的夸美纽斯第一次明确提出教育必须遵循自然的思想。

瑞士的裴斯泰洛奇提出“教育心理学化”。德国的赫尔巴特第一次明确提出把教学的研究建立在心理学等学科基础上。

为适应教育不断发展的需要，教育心理学的任务也不断增加，研究对象的范围逐渐扩充。教育心理学由初期偏重于学习心理的研究和学习规律的讨论，并且大多集中于智育方面的问题，随着教育对人的全面发展的日益重视，越来越重视道德行为、道德情感以及审美情感的培养。

近代科学的发展，特别是近代生物学、人类学、社会学、医学及精神病学的发展都对教育心理学产生影响，促使它不断更新内容，以适应社会发展的要求。

虽然交易心理学发展的时间不是很长，但它取得的进步却非常显著。其中很多优秀教育心理学家的思想更是广泛应用在人工智能等计算机领域。

比如加涅把信号学习分成反应学习、连锁学习、联想学习、辨别学习、概念学习、规则学习和解决问题等七大类。

托尔曼则说，通过刺激—有机体内变化—联结来学习实质。通过情境认知、综合表象、整体反应达到学习结果；通过目标—对象—手段进行有预期的尝试，从而完成学习过程；通过能力、刺激、频率、练习和奖励来学习规律。

而桑代克则说，学习的方法非常简单，就是一个不断试错的过程，学习遵循的是精简原则而不是逻辑原则，最好不用什么协方差。

同时，教育心理学对创造性学习也非常重视，其中一种学习的过程是试错、顿悟、暗示、理解、假设、推理、检验等。

它不仅强调了个体差异和因材施教，同时还强调了学习方法的学习，正所谓授人以鱼不如授人以渔的道理。

虽然知识的学习也包括对动作技能的学习（其实动作技能也是广义人工智能的一部分），不过本书更偏向软件技术，所以这里不做过多的讲述。

总之，教育心理学是非常好的参考资料——如果你想让你的人工智能或者机器人进步更快的话，那么你就不妨研究一下教育心理学，最起码对自己处理好子女和同事间的关系是有帮助的。

第 3 章 人工智能的探讨

现在人们一般把人工智能分为三个流派，这里所说的流派并不是严格意义上的流派划分，只是为了方便大家能够充分理解其中的异同而进行笼统的划分。

第一种是符号主义，那么符号主义是什么意思呢？其实符号主义是指我们所做的事情都是具有逻辑基础的，符合逻辑的。比如我们认为 3 大于 2，而 2 又大于 1，所以我们认为 3 是大于 1 的。再比如，如果这个动物长羽毛，那么我们就认为这个动物是鸟类。像这种逻辑非常清楚的推理其实就是一种符号主义，所以符号主义也叫逻辑主义。

其实整个符号主义都是模仿人的心理过程的，比如大侦探福尔摩斯破案的基础就是一个心理揣摩的过程，他通过各种各样的心理揣摩进行推理和演绎，然后判断谁更有可能是罪犯。

符号主义在人工智能的早期非常流行，特别是专家系统，这些专家

系统在数学、化学、天气、农业等领域发挥着非常重大的作用。

在这方面，美国 **IBM** 做得是非常好的。

第二种是联结主义，联结主义也是当前人工智能最为流行的一个派别。其实联结主义也诞生了好多年，而它最基本的思想就是仿生，仿照人或其他动、植物的生理结构。

比如医学中的人体结构，都有哪些器官，以及这些器官是怎么运作的。或者再具体一点，从神经系统到神经细胞，在从神经细胞到神经元，看看它们是什么样的结构，它们之间是怎么运作的。

现在行业里比较流行的机器学习，就是联结主义的一种，比如在图像学习中，它主要是根据人的视觉神经的原理来进行研究和模仿的。观察人是怎样一步步对图片进行判断和记忆的，人是怎么根据局部小图就能知道它的整体大概是什么样子的。

机器学习是联结主义的主要代表。尤其是最近几年，随着计算机计算量（如 GPU）的提高和数据的增加，机器学习的效果越来越明显。

在语音识别领域，科大讯飞的语音识别就可以达到 96% 以上的识别率。另外美国谷歌、微软等公司的同声传译也做得非常好，如英语和西班牙语之间几乎就可以正常对话。中英文这方面搜狗做的还是不错的。

因为联结主义主要运用的是仿生思想，所以联结主义也叫仿生主义。

第三种是行为主义，行为主义也叫控制论或进化主义，是研究和模仿生物演化过程的，比如生命的演化。或许 800 年前我们都是一家，我们的老祖先就是一只猿猴，在这之前我们的老祖先可能是一种海洋生物。

而行为主义就是研究我们这些生命是如何从一种原始生命一步一步演化成现在这种形态的。生命的演化不同于进化，虽然进化更让我们喜

闻乐见，但是我们仍然无法证明我们现在就是最优秀的形态，演化的意义在于生命恰到好处。物竞天择，适者生存。

比如我们其中的一支就可能因为某种原因而退出了历史的舞台，就算如恐龙般强大，也可能不复存在。在生命的演化过程中有着各种各样的可能性，只有活下来的人才资格说自己是“优秀的”（适合的）。也只有我们这些活下来的，今天才会去研究我们这些人是怎样一步步地活下来的。

而做这种研究的人工智能就是行为主义，行为主义现在有抬头的趋势，比如很多的强化学习和遗传算法就是典型的行为主义。


这块做得比较典型的公司大多是一些机器人公司。它们的这些机器人主要就是通过不断的学习才能够站立、行走和奔跑的。

整个训练过程一定是有错的，而且错误的次数可能很多，没关系，我们不管错误的部分，只管正确的就行了。这样强化学习后，机器人正确的次数就越来越多了。

现在强化学习（也叫增强学习）正有赶超机器学习的趋势。

人工智能（尤其是机器智能）强化了计算机科学的学习和模仿能力，计算机科学作为一门新兴快速成长的科学，正在不断向那些传统的科学进行学习。

因此，学习永远比创新来得容易，即使创新也可能是另一种形式的学习。



第 2 篇

人工智能

第 1 章 符号主义

很多人认为人工智能源于数理逻辑。数理逻辑从 19 世纪末开始得以迅速发展，到 20 世纪 30 年代开始用于描述智能行为。计算机出现后，又在计算机上实现了逻辑演绎系统。其中比较有代表性的成果为启发式程序，它证明了多达 38 条的数学定理，从而表明人们可以应用计算机来研究人的思维构成，模拟人类智能活动。正是这些符号主义者，早在 1956 年首先采用“人工智能”这个术语。后来又发展了启发式算法→专家系统→知识工程理论与技术，并在 20 世纪 80 年代取得很大发展。符号主义曾长期一枝独秀，为人工智能的发展做出重要贡献，尤其是专家系统的成功开发与应用，为人工智能走向工程应用和实现理论联系实际具有特别重要的意义。在人工智能的其他学派出现后，符号主义仍然是人工智能的主流派别。这个学派的代表人物有纽厄尔、西蒙和尼尔逊等。

符号主义是非常重要的，提起符号主义不得不说的就是专家系统，专家系统的技术和理论体系都非常成熟。相比较其他流派的人工智能而

言，专家系统技术更为简单。专家系统一般包括三个方面，分别是知识库、推理机和解释器。

其中，知识库是指那些经过验证的真实知识或经验，哪怕你的知识是一个假设，那么这个假设也必须是真实的，否则不能记录在知识库体系中。虽然现在比较高级的专家系统的这些知识库更像是有一种概率和权重，但为了方便理解，我们还是采用传统的专家系统来进行说明。

比如我们假设： $1+1=2$ ，或者 $1+1=10$ ，那么这个假设必须是真实的，所谓的真实就是指在一定条件下要么为真要么为假，而不能同时既为真又为假，否则如果有歧义，后续都将无法进行推理和演绎。其实数学本身就是以假设为前提的。

所以知识库里面的知识特别强调：一切要从事实出发，而且知识库中至少有一个事实。例如，它是一个动物，那么这个知识就要百分之百，不能有任何不确定性（后面会用到这个事实）。

推理机是什么呢？推理机实际上就是我们对一些事情有自己的理解和判断，它是我们的一个判断过程。

例如，我想知道在座的同学中，谁今天晚上可能有约会，那么我就要根据已有的知识进行逻辑推理，最简单的推理结构就是：

如果……，那么……。

如果（买了鲜花），那么{可能就有约会}。

如果这个动物长羽毛，那么我们判断它是鸟类。我们怎么知道它是一个动物而不是植物的呢？因为前面的知识库已经百分之百地告诉了我们它是一个动物，所以我们才有后面的推理。

其实很多时候，推理机和知识库是相辅相成的，知识库为推理机提

供事实依据，同样推理机推理出来的结论又可以成为新的事实存入知识库。

至于推理的过程，我们可以采用正向推理，也可采用逆向推理，甚至是双向推理。

解释器同样是专家系统不可或缺的一部分，因为这个系统是给人来用的，所以它必然要和人打交道，既然要和人打交道，那么就要让人们知道怎么用，但是这个系统的推理过程我不知道啊？你能不能告诉我你凭什么做出这样的结论。

这种告诉人们系统是怎样推理的一个过程或者与使用者之间的一个交互过程就是解释器。

正是解释器的这样一个好处，才让它经久不衰，它既让人们知道结果，也让人们知道原因，知其然并知其所以然。

比如就拿刚才的例子来说，你怎么知道这个动物属于鸟类呢？答：因为它身上长羽毛，所以它属于鸟类。

专家系统的三大要素如图 2-1-1 所示。



图 2-1-1 专家系统三大要素

前面说了专家系统的三大要素，那么真正的专家系统大概是什么样的一个流程和结构呢？

如图 2-1-2 所示，其实专家系统并不复杂，既然是专家系统，那么我们就要有专家，否则没有专家的系统只能称为系统。

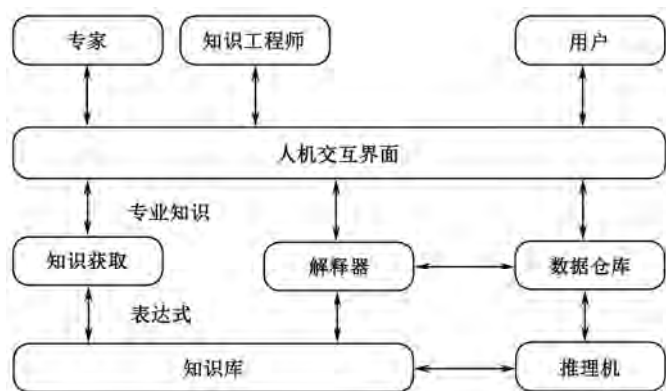


图 2-1-2 专家系统基本结构图

专家在专家系统中仍然扮演着专家的角色，只不过这个专家需要通过人机交互界面让系统能够理解专家的知识 and 思想。比如这是一个金融专家系统，那么金融专家就需要通过人机交互界面把一些公式和模型录入知识库，或者告诉专家系统从哪里导入或者获取值得信赖的知识。可能对于金融专家而言，这个过程有些庞杂，所以我们可以找那些既懂计算机编程，又懂些金融知识的工程师来做这件事。

当金融专家录入金融相关知识后，金融用户就可以使用了，金融用户通过人机交互界面与系统的数据库和推理机进行交互。

我们这里举一个例子。

比如这个金融专家系统有三个专家，并且三个专家都是炒股高手，但是他们具体的思想却有所不同。例如：

专家 A 认为：当均线出现金叉的时候买入。

专家 B 认为：当布林线向上突破 80 的时候买入。

专家 C 认为：当成交量大于 100 手的时候买入。

那么，当用户在使用专家系统时，只要均线出现了金叉，同时当布林线向上突破 80，而且成交量大于 100 手时，就可以兴高采烈地买入了。既然三位专家都说要买入，那就听专家的好了。

但是你会发现，事实上随着专家的增加，专家的意见可能相反，或者说很难同时满足所有专家提出的条件，怎么办？这时候你可以通过给每个专家打分的方式来设置他的权重。

比如你认为专家 A 很重要，给他 60 分，专家 B 也很好，给他 30 分，专家 C 一般，给他 10 分。

这样金融专家系统就把海量的用户和众多的专家结合起来了。如果我们把金融专家换成数学专家或化学专家，那么这个专家系统就是数学专家系统或化学专家系统了。

专家系统出现后，大家认为所有的问题都可以得到解决了，人们似乎找到了一个人工智能的圣杯，只要知识库足够庞大，专家足够强，没有解决不了的问题。正是在这种前提下，出现了人工智能的第一次热潮。

专家系统与传统计算机相比的优势是明显的，我们来看一下两者的对比，见表 2-1-1。

表 2-1-1 计算机与专家系统的区别

列 项	传统的计算机程序	专 家 系 统
处理对象	数字	符号
处理方法	算法	启发式
处理方式	批处理	交互式
系统结构	数据和控制集成	知识和控制分离
系统修改	难	易
信息类型	确定性	不确定性
处理结果	最优解	可接受解
适用范围	无限制	封闭世界假设

专家系统似乎让人们看见了强人工智能的未来，甚至日本说自己将造成世界上最先进人工智能计算机，也被称为第五代计算机。其实它做的就是一套专家系统。当然现在大家都是知道这个项目是失败的，第五代计算机项目的负责人也跳楼身亡了。

第 2 章 联结主义

2.1 机器学习

图灵测试认为，只要测试者不能够区分与他对话的是人还是机器，我们就认为这台机器拥有了人工智能，很多专家也倾向于这样的测试，因为在这些人的心目中，他们的最大梦想就是能够模拟出和人一样聪明的机器人，然后代替人类工作。

联结主义认为人工智能源于仿生学，特别是对人脑模型的研究将有利于人工智能取得突破性发展。它的代表性成果是 1943 年由生理学家麦卡洛克和数理逻辑学家沃尔特·皮茨创立的脑模型，即 MP 模型，它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。1949 年，唐纳德·赫布又提出了赫布理论，尤其是在弗兰克·罗森布拉特发明了感知器（见图 2-2-1）后，直接引发了 20 世纪 60~70 年代以感知机为代表

的脑模型的研究的热潮。由于受到当时的理论模型、生物原型和技术条件的限制,脑模型研究在 20 世纪 70 年代后期至 80 年代初期落入低潮(也有人说是因为马文·明斯基的抵制造成的)。直到物理学家约翰·霍普菲尔德在 1982 年和 1984 年发表两篇重要论文,提出用硬件模拟神经网络以后,联结主义才重新抬头。从某种意义上说,霍普菲尔德的贡献更为重要,因为没有霍普菲尔德,联结主义可能就夭折了。1986 年,鲁梅尔哈特等人提出多层网络中的反向传播(BP)这一经典算法。此后,联结主义势头大振,从模型到算法,从理论分析到工程实现,为神经网络计算机走向市场打下基础。慢慢地,行业内对人工神经网络的研究热情越来越高,尤其是在 2012 年以来,联结主义凭借深度学习在识别领域的辉煌战绩,开始独领风骚。

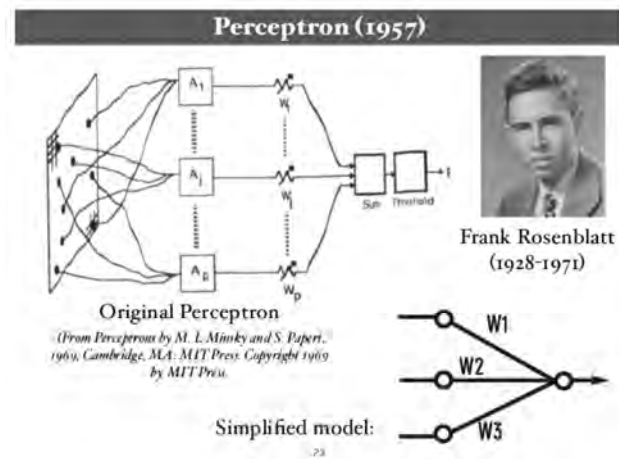


图 2-2-1 弗兰克·罗森布拉特的感知器

现在联结主义非常流行,很多人也或多或少地了解过,这里就简单介绍机器学习的部分。为了方便理解,我们将深度学习、神经网络、数据统计等统一归为机器学习这一类。

既然机器学习这么流行，那么机器学习到底能干什么呢？其实机器学习主要有如下四个作用。

第一个作用就是聚类分析。什么称为聚类？聚类就是我们把一些看似杂乱无章的事物用一种或者几种方法对其分门别类。比如这边是男生，那边是女生；这个是蓝色的，那个是绿色的；这样是健康的，那样是有病的。

例如，我们给出一张图片来，然后我们把距离相近的一些点聚成一类，一共分成两类，如图 2-2-2 所示。

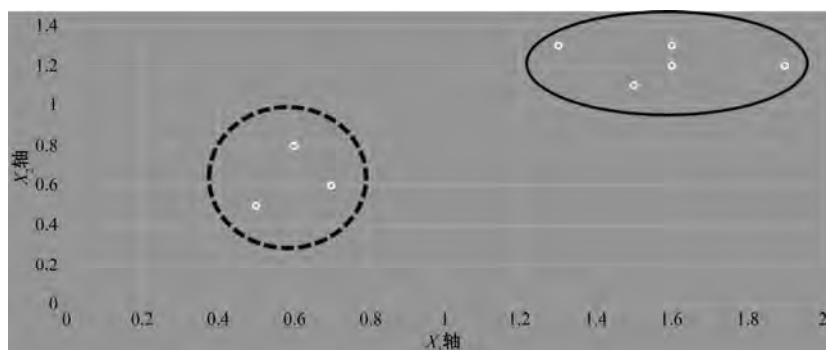


图 2-2-2 聚类示意图

第二个作用就是特征分析。特征分析我们用的也是非常多的，比如声音识别、指纹识别、人脸识别等。如果我们判断这个汉字是不是“人”字，那么怎么办呢？我们可以把这个汉字分割成 $4 \times 4 = 16$ 位的方格，然后在对应的方格内识别其局部的特征，再把这些特征的相似性加权，根据相似来判断这个汉字是不是符合特征。如果像素足够小，组合的特征足够大，判断的准确性也就足够高了。如图 2-2-3 所示，我们只给出了四个方格的局部特征。

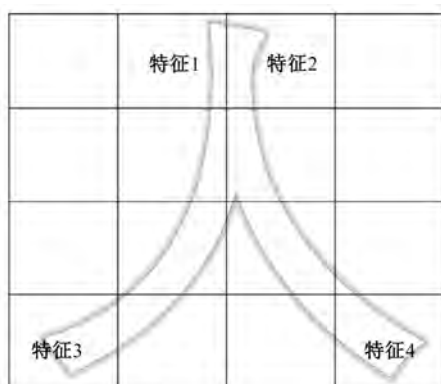


图 2-2-3 特征示意图

第三个作用就是密度分析。比如我们看到一张比较乱的图，或者我们想判断某段路线的交通情况，这时候可能有各种各样的数据，比如车辆型号、时间、司机性别等，其实我们只要知道某个时间点上该路线上一共有多少辆汽车就可以了，这样我们基本上就可以判断出道路的拥堵情况了。密度分析示意图如图 2-2-4 所示。

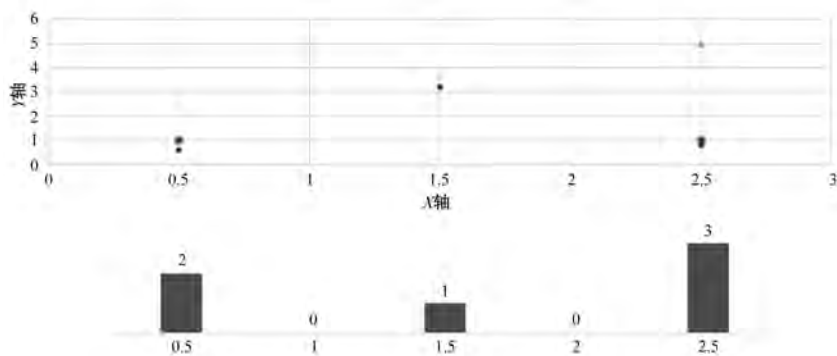


图 2-2-4 密度分析示意图

第四个作用就是降维分析。降维分析可以把复杂的多维空间的数据，映射到低维空间上，让我们分析起来更加容易。降维分析示意图如图 2-2-5 所示，我们很难去理解和计算一个球体，但是我们把三维的球体映射到

平面上变成二维的圆形，就好理解和计算了。

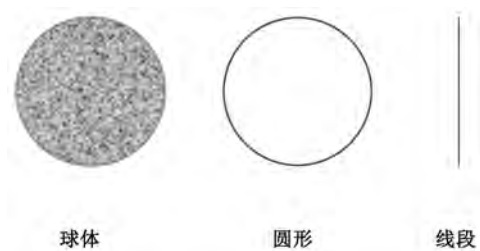


图 2-2-5 降维分析示意图

机器学习大概就有上述四个作用，但是想要达到上述目的的方式方法却有很多种，不仅是机器学习，比如深度学习、神经网络、隐马尔可夫、蒙特卡洛搜索、随机网络、贝叶斯、BP、 α - β 剪枝、线性规划、深度问答、随机聚类、迁移学习、支持向量机、长短学习等。

至于具体是否采用机器学习技术，或者具体采用哪种机器学习技术，则需要根据团队和具体行业的特点而定，没有最好的，但有相对合适的，例如，深蓝战胜卡斯帕罗夫，主要采用 α - β 剪枝，如图 2-2-6 所示；沃森参与危险边缘节目获胜，主要采用深度问答技术，如图 2-2-7 所示；



图 2-2-6 深蓝战胜卡斯帕罗夫

AlphaGo（阿尔法围棋）战胜李世石，主要采用蒙特卡洛搜索技术，如图 2-2-8 所示；冷扑大师挑战中国龙之队获胜，主要采用线性规划技术，如图 2-2-9 所示。以上这些都是侧重了不同的算法和技术从而达到获胜的目的。



图 2-2-7 沃森参与危险边缘节目



图 2-2-8 AlphaGO（阿尔法围棋）



图 2-2-9 冷扑大师挑战中国龙之队

很多人说，他们接触到的机器学习，大多是黑箱，只知道结果，但不知道是怎么推导出来的。其实这些方式和方法可能不太重要，因为我们的框架里面默认即已经包含了这些算法，我们也很难逐一分析，其实这些算法剖析后也没有大家想的那么神秘，至少不是完美的，甚至很多机器学习的专家都经常要考虑到使用哪个算法更合适。

也许哪一天某个程序员灵光一现，发现一种全新的算法，虽然别人也可能已经使用过了，但是没什么关系，只要我们机器学习的结果好就可以了。

为了简单起见，我们就从机器学习的大数据开始着手。

大家知道，我们让机器学习的时候，一定要给它很多数据它才能学习得更准确，那么这些数据要么就是杂乱无章的，比如我们给你很多张图片，什么也不说，然后让你在这些图片中找出几个最像的放在一起。要么有提示的，比如我们给你很多张图片，然后每张图片都告诉这张图片是什么。假设，这张是猫，那张是狗，然后让你找出和猫相似的图片，再找出和狗相似的图片。

我们通常把第一种情况，也就是只有图片没有标签的学习称为无监

督学习，把第二种情况有图片并有说明的（有标签）的学习称为有监督学习。

目前来看，无监督学习是机器学习的主流，毕竟这个世界上大部分的数据都是没有标签的，或者就算是有标签也不见得准确。

至于哪种学习更好，还要看具体情况。比如你想根据历史行情选择买卖股票的时间点，前天 K 线出现锤子线的时候第二天股票上涨，所以我们买入；而前天 K 线出现上吊线的时候昨天股票下跌，所以我们卖出。股票 K 线形态跟着股票价格甚至涨跌的这种情况，可能采用有监督学习就比较好。但是你想识别某个人的脸部特征可能采用无监督学习就比较好。

或者干脆把有监督学习和无监督学习结合起来，效果也不错。将二者结合起来的的学习也称为混合学习。混合学习在某种意义上能够起到事半功倍的效果。

有人甚至把机器学习的过程比作道士炼丹，道士用了很多的配方，然后不断地炼制，至于炼出来的丹药，哪个炉子里的会好用我们并不知道，甚至丹药是怎么形成的我们也不知道。

这也应了一句话：机器学习往往知其然，却不知其所以然。

2.2 常用框架

其实细心的人可能会发现，我们让计算机不断地去学习，然后再根据学习的结果来判断某件事情的相似率或成功率，而这种概率一般都是建立在统计意义上的，而统计意义上想更有说服力就需要有海量的数据和极大的计算量。

我这里说一个小故事，比如你想比较快速地从 100 个面试者中挑出一个最好的，怎么办？你可以尝试这样做：

假设你有 100 个候选人，你先选出 30 个，然后逐一见面，从中挑出一个最好的来，然后再用这个最好的作为标准，看看剩下的 70 个人里面有没有比他还好的，只要有就立刻签合同。

这很明显是一个大概率事件，但是仍然不能保证剩下的就没有更好的。

最好的办法就是全部都看一下，但是显然工作量太大了。于是有人就想到并行来做，这是什么意思呢？就是我们聘请 10 个顾问，然后每个顾问面试 10 个人，然后再将入选的 10 个人进行最后一轮面试。

所以为了达到更好的机器学习效果，我们需要在软件和硬件上同时下工夫，允许这种高并发的计算。

首先我们来说下软件，比如那些主要与 Java 相关的机器学习框架。

(1) Weka 集成了数据挖掘工作的机器学习算法。这些算法可以直接应用于一个数据集上，或者你可以自己编写代码来调用。Weka 包括一系列的工具，如数据预处理、分类、回归、聚类、关联规则以及可视化。

(2) Massive Online Analysis (MOA) 是一个面向数据流挖掘的流行开源框架，有着非常活跃的成长社区。它包括一系列的机器学习算法（分类、回归、聚类、异常检测、概念漂移检测和推荐系统）和评估工具，关联了 Weka 项目，MOA 也是用 Java 编写的，其扩展性更强。

(3) ELKI 是一款基于 Java 的开源数据挖掘软件。ELKI 主要集中于算法研究，重点研究聚类分析中的无监督方法和异常检测。

(4) Encog 是一个先进的机器学习框架，集成了支持向量机、人工神经网络、遗传算法、贝叶斯网络、隐马尔可夫模型、遗传编程和遗传算法。

(5) Datumbox 机器学习框架是一个用 Java 编写的开源框架, 允许快速地开发机器学习和统计应用。该框架的核心重点包括大量的机器学习算法以及统计测试, 能够处理中等规模的数据集。

(6) Oryx 2 是一个建立在 Apache Spark 和 Apache Kafka 的 Lambda 架构实现, 但随着实时大规模机器学习而逐渐开始专业化。这是一个用于构建应用程序的框架, 但也包括打包, 以及面向协同过滤、分类、回归和聚类的端到端的应用程序。

(7) PHP-ML 是 PHP 的机器学习库。同时包含算法、交叉验证、神经网络、预处理、特征提取等。PHP-ML 主要使用数理统计、数值优化方法、概率论和离散分析等方法进行机器学习。

深度学习框架

目前, 人工智能的主流语言是 Python, 而不是 Java 或者 C++。之所以使用 Python, 不仅是因为它学习起来简单, 而是风头正劲的深度学习框架几乎无一例外地都使用了 Python 作为主要编程语言, 如果你想用 JavaScript 或 PHP 什么的, 那么对不起, 它们很少提供这样的文档。

下面我们就来介绍一些比较常用的深度学习框架。

Deeplearning4j

顾名思义, Deeplearning4j 是 “for Java” 的深度学习框架, 也是首个商用级别的深度学习开源库。Deeplearning4j 由创业公司 Skymind 于 2014 年 6 月发布, 使用 Deeplearning4j 的不乏埃森哲、雪弗兰、博斯咨询和 IBM 等明星企业。Deeplearning4j 是一个面向生产环境和商业应用的高成熟度深度学习开源库, 可与 Hadoop 和 Spark 集成, 即插即用, 方便开发者在 APP 中快速集成深度学习功能, 可应用于以下深度学习领域:

人脸/图像识别、语音搜索、语音转文字、垃圾信息过滤、电商欺诈侦测。

CNTK

CNTK (Computational Network ToolKit) 是微软出品的开源的人工智能工具包，同时支持 CPU 和 GPU，中文有“人工智能”的意思，但是实际的英文单词翻译起来其实是一个计算网络工具箱。换个角度说，其实可以理解为它是一种通过计算来实现人工智能的一种方式。

微软的这个深度学习工具箱通过一系列步骤来描述神经网络。在这个有向图中，叶节点代表输入值或网络参数，而其他节点代表矩阵操作依赖这些输入。在 CNTK 上，可以很容易地实现和结合当今流行的模型类型，如前馈神经网络 (DNNs)、卷积神经网络 (CNNs)、循环神经网络 (RNNs/LSTMs)。在实现随机梯度下降学习时能够自动计算梯度。

Caffe

Caffe 是由加州大学伯克利的贾扬清博士开发的一款结构清晰、可读性高、学习快速的深度学习框架。Caffe 具有以下一些特点。

第一，表达方便。模型和优化办法使用纯文本来表示，而不是代码，同时支持中间文件。

第二，速度快。对于科研来说，快速的机器学习和优良的算法同样是非常重要的。

第三，结构清晰，模块化。Caffe 具有很强的灵活性和扩展性，正是由于有着这样的特点，很多变种的 Caffe 才得以出现。

第四，使用简单。尤其对于图像处理来说，Caffe 可以通过浏览器进行简单选择就可以完成任务。

下面给出几张 Caffe 的截图：Caffe 框架首页如图 2-2-10 所示，Caffe 框架设置页如图 2-2-11 所示。

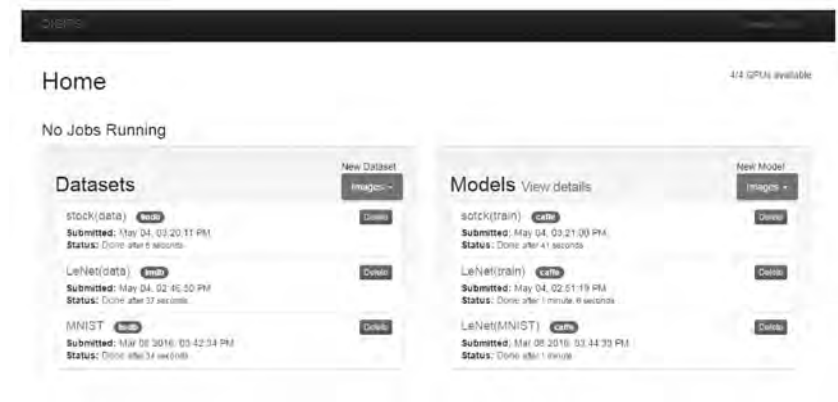


图 2-2-10 Caffe 框架首页

TensorFlow

TensorFlow，中文意思为“张量流”，所以我们也很容易联想到多维数据流的处理。TensorFlow 是现在大红大紫的深度学习框架，这主要得益于谷歌的大力推广与维护，当然也和谷歌本身就有很大的数据量有一定的关系，特别是在 AlphaGo 大放异彩之后，TensorFlow 在 Github 上 Fork 数和 Star 数都是最多的，在国内更加明显。

毫不夸张地说，TensorFlow 的流行让深度学习门槛变得越来越低，只要你有 Python 和机器学习基础，入门和使用神经网络模型就变得非常简单了。TensorFlow 支持 Python 和 C++两种编程语言，再复杂的多层神经网络模型都可以用 Python 来实现，如果业务使用其他编程也不用担心，使用跨语言的 gRPC 或 HTTP 服务也可以访问使用 TensorFlow 训练好的智能模型。

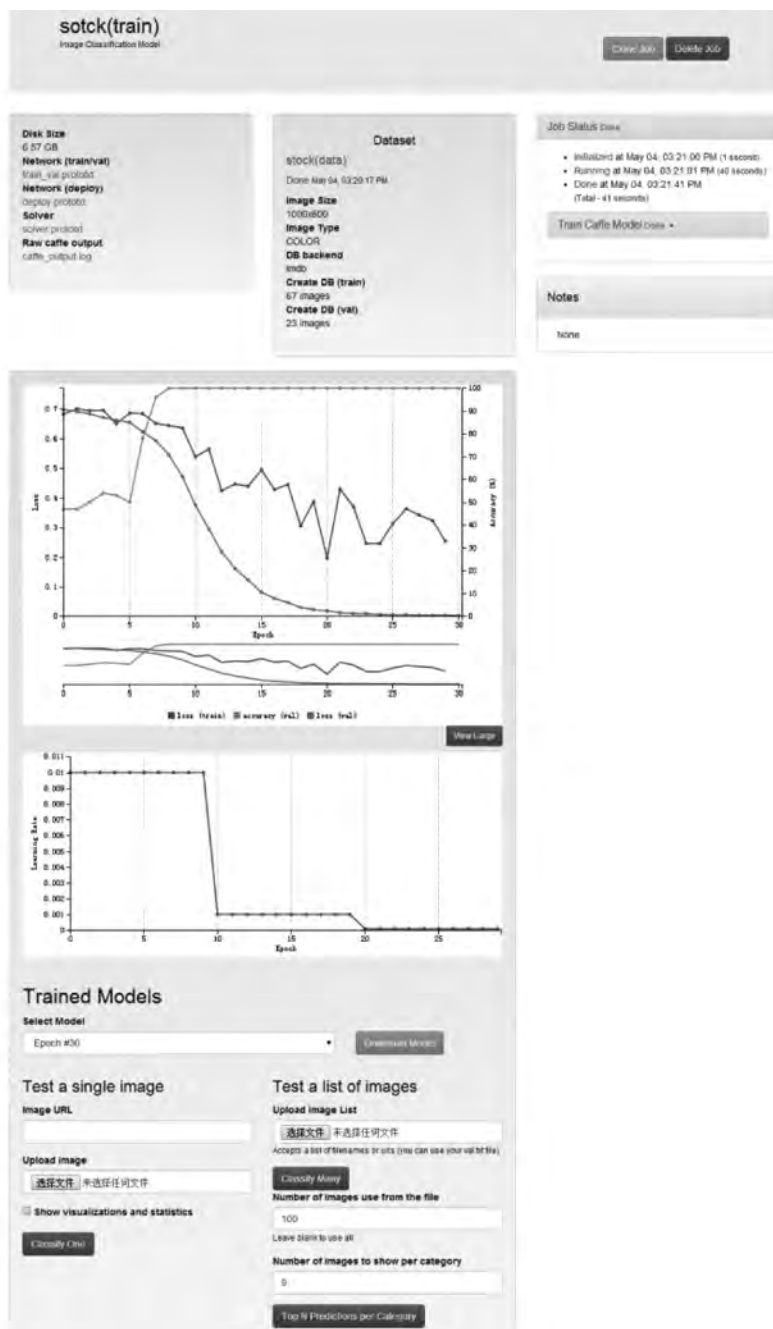


图 2-2-11 Caffe 框架设置页

TensorFlow 的 CPU 版本的安装也是非常简单的;如果你想安装 GPU 版本,那么可能需要一些设置;如果你只是想简单地体验 TensorFlow 深度学习功能,那么在 Windows 操作系统上也是可以轻松实现的,只不过仅有 Python3.5.x 版本支持 Windows 的 pip3 安装。

Python3.5.x 下载地址为: <https://www.python.org/downloads/release/python-352/>, 下载安装后,在 Windows command 命令里面输入:

```
C:\> pip3 install --upgrade tensorflow
```

即可安装 TensorFlow 的 CPU 版本。

或者输入:

```
C:\> pip3 install --upgrade tensorflow-gpu
```

安装 TensorFlow 的 CPU 版本。

如果你的计算机上没有 GPU 显卡的话,安装 CPU 版本的就可以了。

说明:

(1) 安装 Python3.5.x 的时候注意添加 PATH, 否则可能找不到程序的路径。

(2) 默认用于 Windows10 版本, 如果用于 Windows7 系统可能需要系统更新 sp1 包 (因为它需要很多依赖)。

一般喜欢 Python 的用户可能喜欢 Anaconda 这个集成安装环境, 如果是 Anaconda 的话, Python 版本也要是 3.5.x, 太高了或太低了都不行。Anaconda 参考网址为:

<https://repo.continuum.io/archive/index.html> (如 Anaconda3-4.1.1-Windows-x86_64.exe)

Anaconda 安装好后, 打开 Anaconda-prompt 或者 command 终端并创建一个名为 tensorflow 的 conda 环境:

```
C:> conda create -n tensorflow
```

创建好后, 通过 activate 命令激活 conda 环境:

```
C:> activate tensorflow  
(tensorflow)C:> # Your prompt should change
```

然后就可以安装 TensorFlow 了。安装 CPU 的命令如下:

```
(tensorflow)C:>pipinstall--ignore-installed--upgradehttps://storage.googleapis.com/tensorflow/windows/cpu/tensorflow-1.1.0-cp35-cp35m-win_amd64.whl
```

安装 GPU 的命令如下:

```
(tensorflow)C:> pip install --ignore-installed --upgrade https://storage.googleapis.com/tensorflow/windows/gpu/tensorflow_gpu-1.1.0-cp35-cp35m-win_amd64.whl
```

安装完成后, 我们再重新启动一个终端进行测试。

```
$ python
```

在 python 交互式 shell 中输入以下短程序:

```
>>> import tensorflow as tf  
>>> hello = tf.constant('Hello, TensorFlow!')  
>>> sess = tf.Session()  
>>> print(sess.run(hello))
```

如果系统输出以下内容, 就表示你的 TensorFlow 安装成功了。

```
hello, TensorFlow!
```

TensorFlow 和 Caffe 框架一样也提供了一个基于浏览器的可视化操作界面——TensorBoard, TensorBoard 简单大方, 而且流程节点图是动态的, 如图 2-2-12 所示。

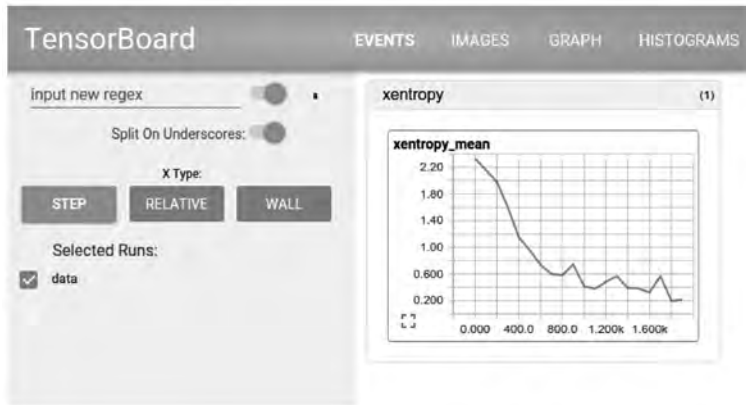


图 2-2-12 TensorBoard 示意图

由此可见，TensorFlow 功能非常强大，是深度学习的利器，同时它还支持分布式架构。

2.3 硬件设备

首先，我们看看 CPU 的结构。

CPU 主要包括运算器和控制器两大部件。此外，还包括若干个寄存器和高速缓冲存储器，以及实现它们之间联系的数据、控制和状态的总线。从上面的叙述可以看出，CPU 主要包含运算逻辑器件、寄存器部件以及控制部件等，如图 2-2-13 所示。

从字面上我们很好理解，运算逻辑器件主要执行算术运算、移位等操作，以及地址运算和转换；控制器件则是负责对指令译码，并且发出为完成每条指令所要执行的各个操作的控制信号。

目前世界上 CPU 的主要生产厂商是英特尔和 AMD。

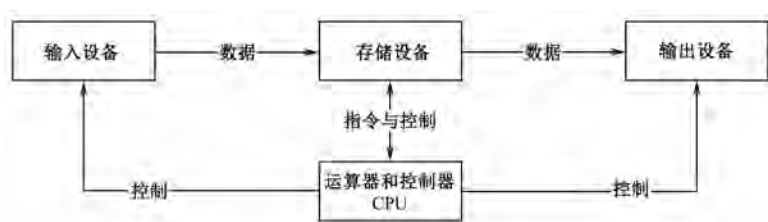


图 2-2-13 CPU 结构简化图

其次，我们来看看 GPU。

CPU 就像一个助理，它总是一件一件地做好我们安排的事情。但是当工作繁忙的时候，一个助理就不行了，于是人们就想，我们可不可以多请几个助理，让它们一起工作，这样效率不就提高了吗？于是工程师们开始把上千个处理器放在同一块芯片上，这样 GPU 就诞生了。GPU 就是图形处理器，早先 GPU 主要是在个人计算机上运行绘图运算工作的，比如游戏显卡。后来随着 GPU 性能的提升，GPU 也开始广泛地应用于需要大量计算的领域，比如挖矿（比特币）和机器学习。CPU 功能模块有很多，能适应复杂的运算环境；GPU 构成则相对简单，对缓存的需求小，所以大部分晶体管可以组成各类专用电路、多条流水线，这才使得 GPU 的计算速度有了突破性的飞跃，拥有了更强大的处理浮点运算的能力。现在一个普通的 GPU 就有差不多 1 千多个核心。

虽然 GPU 是为了图像处理而生的，但是它在结构上并没有专门为图像服务的部件，只是对 CPU 的结构进行了优化与调整。现在 GPU 不仅可以在图像处理领域大显身手，它还被用于科学计算、密码破解、数值计算、大数据处理、机器学习等领域。GPU 是机器学习的主要设备。

目前世界上 GPU 的生产厂商是英伟达。

再次，我们看看 FPGA。

随着人们的计算需求越来越专业化，人们希望有芯片可以更加符合

我们的专业需求，但是考虑到硬件产品一旦成型便不可再更改的这个特点，人们便开始想，我们可不可以生产一种芯片，让它硬件可编程。也就是说，这一刻我们需要一个比较适合对图像进行处理的硬件系统，下一刻我们需要一个对科学计算比较适合的硬件系统，也就是说，我们希望有一款芯片只要通过简单的编程就能适应各种工作，FPGA 就是这样一种芯片。

FPGA 中文全称为现场可编程逻辑门阵列，它是作为专用集成电路领域中的一种半定制电路而出现的，既解决了全定制电路的不足，又克服了原有可编程逻辑器件门电路数有限的缺点。

使用 FPGA，人们可以根据需要，通过可编辑的连接，把其内部的各个逻辑块连接起来。这是不是很像灵活多变的乐高积木呢？FPGA 实验室环境用得较多。

最后，我们看看 TPU。

随着机器学习算法越来越多地应用在各个领域并表现出优越的性能，对于专业的机器学习硬件要求也越来越高，由于 CPU 计算能力差、GPU 功耗高、FPGA 性能低，所以我们有必要研发一款专门为机器学习而使用的芯片，Google 就成功研发了这样一款高性能的机器学习芯片，它就是 TPU，其外观结构如图 2-2-14 所示。

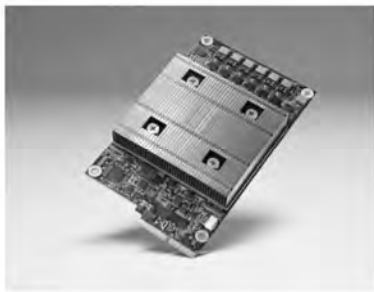


图 2-2-14 TPU 板卡图

从名字上我们可以看出，TPU 的灵感来源于 Google 开源深度学习框架 TensorFlow，所以目前 TPU 还是只在 Google 内部使用的一种芯片。

从 TPU 的外观图我们可以看出，其中间突出一块很大的金属片，这便是为了可以很好地对 TPU 高速运算时产生大量的热进行耗散。

据说 AlphaGo 使用的就是 TPU 架构，如图 2-2-15 所示。



图 2-2-15 TPU 服务器

2.4 云计算

如果很多人不喜欢一次性购买昂贵的 GPU 设备或部署麻烦的深度学习框架，那么我们就可以去阿里云或百度云上购买符合自己需求的人工智能服务器。这些人工智能服务器默认就已经搭载好了很多著名的深度学习框架，并且还有相关的视频教程辅助你使用相关的功能和接口。

阿里云的参考网址为：<https://et.aliyun.com/index>，参考页面如图 2-2-16 所示。

百度云的参考网址为：<http://ai.baidu.com/>，参考页面如图 2-2-17 所示。



图 2-2-16 阿里云产品页面



图 2-2-17 百度产品页面

这里我们做以下两个说明。

一般为了更好地进行机器学习,尤其是需要大量图片和 GPU 计算时,最好选择同一个地区的存储服务器和计算服务器,否则跨网段不仅延迟大而且费用也不低。

另外,如果小图片非常多,多到几百万张的话,在没有专业软件的情况下,可以尝试使用存储模式添加压缩文件映射到共享磁盘中,这样会比很多小文件一起传快一些,如图 2-2-18 所示。

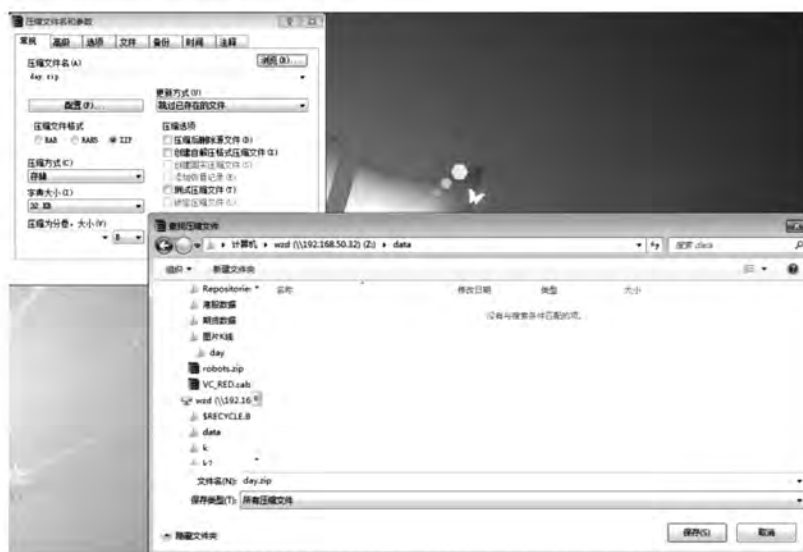


图 2-2-18 小图压缩批量传输截图

2.5 量子计算机

我们目前接触到的主流电子计算机都是基于二进制数字进行工作的,这样非常符合电流、磁铁和开关的有无状态,物理上非常好实现。虽然说俄罗斯曾经利用正电、负电和无电制造过三进制的电子计算机,但是其物理性能还远远没有现在的二进制计算机性能高,所以一度被搁

置，不过想利用更高进制代替二进制的想法却一直是科学家们的梦想。因为在物理性能相同的情况下，从数学理论上高位进制的运算速度远远大于二进制，量子计算机就是其中的一种。

量子计算机是一类遵循量子力学规律进行高速数学和逻辑运算、存储及处理量子信息的物理装置。当某个装置处理和计算的是量子信息，运行的是量子算法时，它就是量子计算机。

普通的数字计算机在 0 和 1 的二进制系统上运行，称为比特。但量子计算机由于量子的不确定性，计算机的输入用一个具有有限能级的量子系统来描述，如二能级系统（称为量子比特），量子计算机的变换包括所有可能的幺正变换。

正是由于量子计算机的输入态和输出态为一般的叠加态（其相互之间通常不正交），所以它要远远强于传统的二进制计算机。因为它可以在量子比特上运算，可以计算 0 和 1 之间的数值。

从数学抽象上看，量子计算机执行以集合为基本运算单元的计算，普通计算机执行以元素为基本运算单元的计算，所以量子计算机要比经典计算机快很多。有学者甚至认为量子计算机比传统计算机快数十亿倍。

说到量子计算机就不得不谈下量子纠缠，现在的科学技术已经估算出量子纠缠的速度大约是光速的一万倍以上，这种超光速的概念颠覆了传统物理学。尤其是在贝尔不等式被广泛实验后，这门学科得到越来越多专家学者的关注。

量子纠缠的概念被提出来以后，很多人工智能学者开始思考大脑是否本质上就是一台量子计算机（见图 2-2-19），甚至大脑所产生的思维意识本身也是一种类似量子的物理现象，通过量子纠缠，人的思想完全可以超越光速、跨越时间。

但是只要提到“量子意识”，相信大多数物理学家都会败下阵来，因为这个概念现在还是一个很模糊的假设。2015 年，加州大学物理学家马

修·费希尔在《物理学年鉴》中提出，大脑可能含有某些特定分子，能维持更加稳固的量子叠加态。他特别指出，磷原子的原子核可能就具有这种能力，文中提出磷原子的核自旋可以作为大脑中的初级“量子比特”，这使大脑能够像一个量子计算机那样运作。

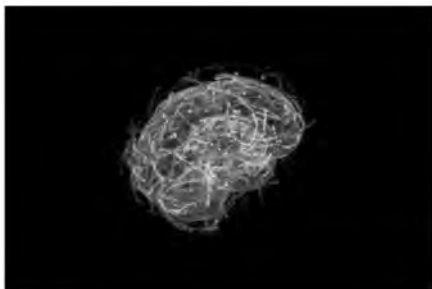


图 2-2-19 量子大脑

那么人类大脑是台量子计算机吗？量子物理能够解释意识吗？这一切难道就是一个有待证明的假设呢？

我们现在做一个假设，这个假设是说所有的物体都会发光，当然这种光不是普通的可见光，我们可以把它理解为一种波。

也就是说，黑色的物体甚至黑洞本身也会“广播”一种光波，这种光波类似充斥着整个宇宙包括我们以前所定义的真空。

当两种光波相互交错时就有了感应，而这些感应在某种程度上或者说达到一个阈值时便是构成意识的基础。

举个小例子：当你烧水的时候，水温没有到 100°C 是不会沸腾的，只有水温到了 100°C 的时候你才会意识到水烧开了。而水分子本身也是到了 100°C 的时候才明显活跃起来的。

这也体现了思维和意识的本质也是一种物质形态。

第 3 章 行为主义

自从 1948 年维纳提出了行为主义以来，行为主义才开始得到人们的重视，不过即便得到了重视，其发展速度也不如符号主义和联结主义。

行为主义主要的思想就是感知加行动，俗称感动，需要在真实世界的复杂环境中进行学习与训练，相对于符号主义以逻辑为主、联结主义以数据为主，行为主义在真实环境的训练要困难很多，或许物联网的蓬勃发展会改变这一现状。

行为主义主要源于控制论。控制论思想早在 20 世纪 40~50 年代就成为时代思潮的重要部分，影响了早期的一批人工智能工作者。维纳和麦克洛克等人提出的控制论和自组织系统，以及钱学森等人提出的工程控制论和生物控制论等也影响了许多相关领域。控制论把神经系统的工作原理与信息理论、控制理论、逻辑及计算机联系起来。早期的研究重点是模拟人在控制过程中的智能行为和作用，如对自寻优、自适应、自镇定、自组织和自学习等控制论系统的研究，并进行“控制论动物”的研制（如机器蜘蛛、机器蝴蝶和机器鱼等）。到了 20 世纪 60~70 年代，

上述这些控制论系统的研究取得了一定的进展，也为智能机器人的发展奠定了基础，并在 20 世纪 80 年代诞生了智能控制和智能机器人系统。行为主义是 20 世纪末才以人工智能新学派的面孔出现的，现在这张新面孔也在物联网的作用下开始引起了许多人的兴趣。

不过行为主义从智能体现上更偏向于硬件一些，毕竟感知和行动都是需要传感器和控制器的，比如布鲁克斯的六足行走机器人，它被看作新一代的“控制论动物”，就是一个基于感知-动作模式模拟昆虫行为的控制系统。

另外，行为主义又称进化主义，因为行为主义其实是一个不断试错和不断演化的过程，所以从其学习路径来看行为主义更像是一个演化论，它不管你为什么，也不管你会变成什么，它只要它能够活下来就可以了。后面我们将会着重讲下行为主义。

在讲行为主义之前，我们给大家看一个表（见表 2-3-1），这个表是我们对三种流派的一个整理和简单的对比，从而让大家有一个更加感性认识。

表 2-3-1 符号主义、联结主义、行为主义三者的区别

	符号主义	联结主义	行为主义
样本数量	小	大	小
计算性能	小	大	极大
相关学科	心理学	生理学	进化论
经典系统	专家系统	机器学习	遗传算法
模拟方式	功能模拟	结构模拟	行业模拟
是否犯错	从不犯错	偶尔犯错	经常犯错
逻辑特点	知其然知其所以然	知其然不知其所以然	不知其然不知其所以然

样本

我们看看样本数量，符号主义几乎不怎么需要训练的样本，样本数量甚至可以为零，但是联结主义确实非常需要大样本，而且往往这个样本数越大效果越好，行为主义对样本的需求量也是很小的，虽然不至于为零，但是也高不到哪里去。

计算

我们看看对计算性能的要求，符号主义对计算机的计算性能要求不高，几乎任何一台普通的计算机都可以计算和运行，但是联结主义却不同，它需要大量的计算，普通计算机已经不能满足其海量计算了，这时候 GPU、TPU 等高性能计算设备应运而生。行为主义在计算机性能要求上比联结主义有过之而无不及，行为主义对计算性能要求极高，通过越来越复杂的迭代计算达到一个人们比较满意的效果。

学科

我们看看重点关注的学科，我们都说计算机科学是一个善于模仿的科学，那么计算机中的人工智能又是模仿什么的呢？举个形象的例子来说，符号主义其实更多地模仿了人的心理学，联结主义更多的是模仿了人的生理学，而行为主义更多的是模仿了生物尤其是达尔文的演化论。

模拟

我们看看主要落脚点，符号主义注重功能的模拟，联结主义则注重结构的模拟，而行为主义则更多地关注行为。


试错

我们看看三个流派对错误的偏好程度。符号主义几乎从来不犯错，它所有的一切都是正确的。而联结主义则会犯错，但犯的误差是可以接受的。至于行为主义则是经常犯错，错的我们都不认识它了，但是它只要在众多的错误中有一次正确就可以了。

最后，我们用一句话分别概括下三个流派的不同（见表 2-3-1）。行为主义是知其然并知其所以然，联结主义是虽然知其然但是不知其所以然，行为主义则是不知其然也不知其所以然。

由此可见，人工智能需要学习心理学、医学和生物学。

由于行为主义刚刚开始崭露头角，很多概念还很新颖，特别是我们将要重点阐述的本能学习又是第一次被提及，所以，这一章我们现在简单介绍到这里，下面我们将会利用整个第 3 篇来着重探讨行为主义中的本能学习。



第 3 篇

生存本能

接下来的章节，也是本书的一个核心章节，通过阅读和思考本章，你可以创作出真正拥有灵魂的机器人来。

本章和行为主义关系比较紧密，当然这种紧密并不是严格意义上的，其实人工智能的所有流派的分类都是为了方便大家能够更好地理解人工智能而已。

以前，我们常说行为主义是仿生学，是模仿生物的演化过程，其实演化论只是行为主义中的一个杰出代表而已，很多时候我们很难准确地把三大流派界定清楚，期间更多的是三者相辅相成。

在讨论人工智能时，我们不妨说两个有趣的现象。比如我这个人平时很邋遢，有事没事的时候喜欢到处吐痰，但是我发现我去天安门的时候就不随地吐痰了。我就感觉好奇怪，原来我发现周围的人都不吐痰，所以我也就不吐痰了。

再举一个例子，我这个人平时挺安静、挺好的，但是当我经过一栋比较旧的大楼时，发现有一块玻璃被别人用石头打碎了，然后我也想扔石头。

我想大家也有过这种情形，像我们这种因环境的改变，性格发生改

变的情况叫作动态延展的心灵，环境变好你也变好，环境变坏你也跟着变坏。

当然有人可能不同意，他说我是一个非常自律的人，不论什么时候我都有自己的做事原则和行事风格，我就是一成不变的我。

其实当你仔细研究自己和体会后，你会发现你的所有决定实际上都受到环境和自身变化的影响，最终的表现形式不过是内因和外因的综合结果。

而我们将要讨论的本能学习就是想要告诉你为什么成为了你，而不是他人。

第 1 章 什么是好，什么是坏

1.1 人为什么会死

在讨论什么是好，什么是坏之前，我们先来研究下生命究竟是什么？人为什么会死亡？其实生命是什么谁也说不清，很多人说生命是人从出生到死亡的一个轨迹，那么生命为什么从诞生的那一天起就一定要走向死亡的终点呢？人为什么会死亡呢？

很多人给出了不同的答案，这里我们只说说其中的三种。

神学

在说神学之前，我们不妨做个大胆的假设，现在假设你就是上帝，然后你创造了很多很多的机器人，然后为了让它们更好地工作，你就会让它们不断地学习和成长，但是后来你发现这些机器人越来越聪明了，

聪明得你都感觉到害怕，害怕它们哪一天不再听你的话了，甚至你也确实发现，它们正在为自己创造一个通向天堂的高塔（巴别塔），当机器人开始让你感到恐惧的时候，你怎么办？总不能任由它们强大，然后把你赶尽杀绝吧，于是你就很可能拔掉它们的网线，或者干脆关掉它们的电源，就像上帝那样发一场大洪水吧，然后只把少数听你话的机器人留下来。

也许你觉得一下子失去这么多机器人，可能于心不忍，或者你还想让它们中的大部分继续为你工作，怎么办？还能怎么办？难道我们直接造一个只会工作不会衰老的机器人吗？可是世界又是变化多端的，就算你制造出恐龙一样强大的机器人也可能哪天突然灭绝，比如黄金一代。

想创造出一批既要能够完成工作而又不会太聪明机器人最简单的方法就是给它设置一个生命周期，只有机器人会死亡，才能让机器人在一个只有几十上百年的寿命里，不会反叛。想想也是，在这样短暂的时间里就算它天天学习，累死他也不会成长为上帝。所以人为什么会死第一种学说就是上帝不让我们活那么久，你想造巴别塔，不行，你的能力太强了你就得英年早逝。

经济学

200多年前，一位名叫马尔萨斯的英国牧师出版了一本小册子，叫作《人口原理》，书中描绘了英伦岛国人口膨胀的可怕前景。

马尔萨斯曾经说道：人口是按照几何级数增长的，而生存资料仅仅是按照算术级数增长的，多增加的人口总是要以某种方式被消灭掉，人口不能超出相应的农业发展水平。这个理论就被人称为“马尔萨斯陷阱”。

许多经济学家认为高死亡率的各种因素，是人口再生产与农业时代生存资料实现匹配的关键过程，马尔萨斯说战争、饥荒和瘟疫都是促使

人口下降到与生存资料生产水平相适应的道路，人口数量要在某种方式和程度上与农业发展成比例的观点是一个内含的逻辑。马尔萨斯提出两个级数的理论：人口是按照几何级数增长的，而生存资料仅仅是按照算术级数增长的，多增加的人口总是要以某种方式被消灭掉，人口不能超出相应的农业发展水平。

想想也是，人都有生育的本能和欲望，他们希望自己能够生养比较多的孩子，人多力量大嘛，一个家庭生 3 个孩子，3 个孩子再生 3 个孩子，四世同堂甚至五世同堂，人越来越多，但是整个社会的资源是有限的，比如这块土地就能养活 100 个人，人越来越多，人均粮食就越来越少了，所以养活不了那些人，即使不发生战争，饥饿也会饿死一些人，从而达到一个平衡点然后再继续。

其实马尔萨斯的理论很简单：因为土地养活不了那么多人了，所以就要消失一部分。

生物学

上面谈到神学和经济学可能并不太科学。于是很多科学家，尤其是生命科学家都开始关注生物本身，通过其观察和实验找出死亡的原因。

人是有 DNA 的，人的这个 DNA 其实就是一个遗传密码，控制着每个细胞的分裂次数和形态。但我们知道 DNA 在不断的复制过程中，就会产生一些变异，原来变异少的时候还没什么，但是在变异的基础上再加上变异影响可就大了，所谓失之毫厘差之千里。

端粒学说认为细胞在每次分裂过程中都会由于 DNA 聚合酶功能障碍而不能完全复制它们的染色体，因此最后复制 DNA 序列可能会丢失，最终造成细胞衰老死亡。

端粒是真核生物染色体末端由许多简单重复序列和相关蛋白组成的复合结构，具有维持染色体结构完整性和解决其末端复制难题的作用。细胞每次有丝分裂一次，就有一段端粒序列丢失，当端粒长度缩短到一定程度，会使细胞停止分裂，导致衰老与死亡。

在端粒学说之外，科学家在实验中还发现人体细胞有分裂的极限。体细胞的存在状态有这样几种。

(1) 极度分化的永远失去分裂能力的细胞，如肌肉细胞、神经细胞、红细胞。

(2) 具有分裂能力，但一般情况下不分裂，只有当组织受到损伤时才分裂的细胞，如肝细胞。

(3) 能不断分裂（有限度地），以补充正常衰老死亡的细胞，如造血干细胞。

也就是说，每秒都有细胞分裂，但每一个细胞的分裂次数是有限的，不同细胞的分裂极限次数不一样。

那么有人会说 DNA 为什么要变异啊，不变不行吗？不变真的不行，先不说变是为了让人类适应新的社会，就说我们拿一种图片来复印一样，你用复印的图片再不断地复印下去，你会发现你的图片越来越不清晰了。DNA 也是这样，当你分裂的次数过多时，它已经不是原来那个健壮的 DNA 了，所以你才会得病、才会死亡。

你说复印机复印的次数越多越不清晰，那么没关系，我们造一台高清的复印机就好了。事实上科学家们也是通过这个方法来延缓人类的衰老的，但本质上还是做不到永生。

科学发现，人体的各个器官都有一个大体上的衰老时刻表，就像零部件一样，磨合好后再继续工作了一定时间之后就开始有磨损了。下面

给出一个人体各个器官衰老的大致时间表。

肺：20 岁开始衰老。肺活量从 20 岁起开始缓慢下降，到了 40 岁，有一些人就开始气喘吁吁了。

皮肤：25 岁开始衰老，慢慢长出皱纹；男性稍晚一些，大概 35 岁皮肤开始出现干燥、粗糙、松弛现象。

大脑：25 岁开始衰老。大脑中的神经细胞会慢慢减少。40 岁后，神经细胞将以每天 1 万个的速度递减，从而对记忆力及大脑功能造成影响。

肌肉：30 岁开始衰老。肌肉一直在生长，衰竭；再生长，再衰竭。30 岁后，肌肉衰竭速度大于生长速度。过了 40 岁，人们的肌肉开始以每年 0.5%~2% 的速度减少。

骨骼：35 岁开始衰老。35 岁骨质开始流失，进入自然老化过程。80 岁时身高会降低 5 厘米。

心脏：40 岁开始衰老。随着身体日益变老，心脏向全身输送血液的效率也开始降低。45 岁以上的男性和 55 岁以上的女性心脏病发作的概率较大。

牙齿：40 岁开始衰老。40 岁以上成年人唾液的分泌量会减少。唾液可冲走细菌，唾液减少，牙齿和牙龈更易腐烂。牙周的牙龈组织流失后，牙龈会萎缩。

视觉：40 岁开始衰老。近距离观察事物会非常费劲。接着，眼睛适应不同强度光的能力降低，对闪耀光更敏感，不适宜夜晚工作。

肾脏：50 岁开始衰老。肾滤过率从 50 岁开始减少，后果是人失去了夜间憋尿的功能，需要多次跑卫生间。75 岁老人的肾滤过率是 30 岁时的一半。

听觉：55 岁左右开始衰老。60 多岁以上的人半数会因为老化导致听力受损，也叫老年性耳聋。这时老人的耳道壁变薄、耳膜增厚、听高频度声音变得吃力，所以在人多嘈杂的地方，交流十分困难。

嗅觉：60 岁开始退化。一生中最初舌头上分布有大约 1 万个味蕾。60 岁后这个数可能减半，味觉和嗅觉逐渐衰退。

生育：55 岁开始退化。男性还稍微晚一些，大概可以推迟到 60 岁左右。70 岁的人几乎已经没有什么生育能力了。

.....

科学家通过人类对饮食、心情、环境温度、工作强度、下丘脑、生育、神经和线粒体的研究发现衰老是不可避免的一种演化。

有一种理论（自由基理论）认为，细胞的衰老是细胞内发生化学反应过程中有害物质堆积的结果。在这些化学反应中，产生称之为自由基的毒素。自由基最终损伤细胞，引起个体衰老。伴随着衰老，损伤越来越多，许多细胞不能正常行使功能或者死亡。当这些现象发生时，可能引起机体死亡。不同的物种以不同的概率衰老，主要取决于细胞如何产生自由基以及对自由基如何产生反应。

而衰老的最终结果就是死亡，比如心脏停止跳动或者大脑缺氧。我们这里的死亡一般是指心脏停止跳动，因为有些动物不需要大脑也能够存活，比如有一种叫作海鞘的海洋软体动物在出生之后是有大脑的，不但有还发育良好，但是当它一旦在海床上定居并发育成熟后，就慢慢地吃掉自己的大脑，然后完全依赖本能进行生活。不仅海鞘如此，很多生物都是不需要大脑的，比如海绵、栉水母、草履虫，植物等。

上述三种思想虽然有些娱乐的成分，但还是有些道理的。唯一能够推翻上述观点或者对其不认可的方式和途径就是科技的进步。

如在经济学上，很多科学家通过提高土地的单位产量来养活更多的人，通过控制 DNA 变异来延长寿命，通过科学合作来达到与创始者和平共处。

因此，科技就是用来改变世界的。

1.2 什么是美

说了人为什么会死，我们接下来讨论什么是美。在讨论什么是美之前我们先来做一个小小的问卷调查，调查的题目很简单，那就是什么样的女人是美女？

什么样的女人是美女？你可能会说：年轻的、身材好的、皮肤好的、会做饭的、贤惠的、腿长的、气质好的等。

这时，你会发现，随着参与人数的增加，每一个人心目中都有一个美女标准，而这个你所谓的美女在别人眼里可能并不太好看，而别人所谓的美女你也并不是非常赞成，这种审美标准，因为民族与种族的差异会变得更加明显。

比如我们的奶奶辈，她们都裹脚。她们认为脚小是美的，脚大的女孩嫁不出去，那么到了现在谁还裹脚呢？再比如我们眼里很一般的模特可能在外国人眼里就非常漂亮。时代不同，审美标准也不同，比如以前我们都喜欢丹凤眼，而现在大多喜欢双眼皮、大眼睛。

其实，不光男人看女人是否漂亮有着自己的标准，女人看男人是否帅气，也有各种各样的标准。美女帅哥对比图如图 3-1-1 所示。

这些标准虽然不同，但这些标准大体上都是我们自己来决定的，所有的美女或者帅哥本质上都是相对于自己的审美观而言的，而不是你受到某个人的主观影响。

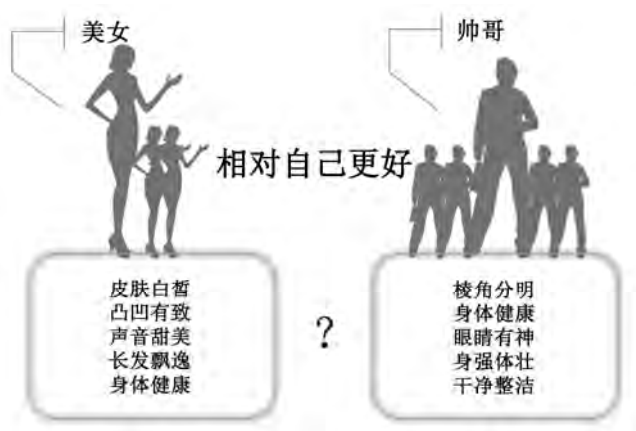


图 3-1-1 美女帅哥对比图

一般来说，美学属于哲学范畴，美学的概念也是由德国哲学家鲍姆加登在 1750 年首次提出来的，他认为需要在哲学体系中给艺术一个恰当的位置，于是他建立了一门学科研究感性的认识，并称其为感性学。

美学作为研究人与世界审美关系的学科，与人的主观意识分不开。有人说要学好美学需要扎实的哲学功底与艺术涵养，因为它既是一门思辨的学科，又是一门感性的学科。美学看似与文艺学、心理学、语言学、人类学、神话学等有着紧密的联系，但其本质确是因人而异的。

其实，美是什么？是美学所要探讨的基本问题。每位先哲对这个问题都有着自己的看法。这并非是一个简单的问题，通过它可以辐射世界的本源性问题的讨论。从古到今，从西方到东方，对“美”的解释是复杂的。如古希腊的柏拉图说，美是理念；中世纪的圣奥古斯丁说，美是上帝无上的荣耀与光辉；俄国的车尔尼雪夫斯基说，美是生活；中国古代的道家认为，天地有大美而不言；而一本《美学原理》则告诉我们美在审美关系当中才能存在，它既离不开审美主体，又有赖于审美客体。美是精神领域抽象物的再现，美感的世界纯粹是意象世界。

我们这里先简单说几个先哲的论点。

康德说，从以审美判断力为中心，联系主体与客体的相互关系去探讨审美的本质。得出一种主体的普遍性是审美的基础，并把美与崇高作为最普遍的审美类型。

黑格说，美在结构上有三大特点：一是从美的本质定义出发，美是理念的感性显现，是用理念确立了美的核心，用感性区别于理念的其他表现形式，从而解释一切美的现象；二是理念显现在感性上是有等级的，这个等级是与宇宙进化历史相一致的，无机物是感性的最低级，艺术是感性的最高级，美有统一的定义，艺术是美的最高和最典型的形态，美学只需要研究艺术就足够了；三是从审美的角度，美学可以划分为三大类型：象征型（建筑）、古典型（雕刻）、浪漫型（绘画、音乐、诗歌），三大审美类型囊括了一切美的对象。

苏珊·朗格说，情感与形式是美的本质消失之后重建艺术哲学型美学原理的典型。第一，以幻象作为总的基础，再在各门艺术中显出幻象上的二级乃至三级差异。从而可知，艺术的自主性和审美的独立性，是构成艺术哲学的一大主旨。第二，基本幻象、艺术形式、艺术形式中内蕴的情感形式是连在一起的。

杜夫海纳说，艺术是美好的，而美是需要审美对象的，没有了审美对象，美就无从谈及。

美学的四个阶段如下。

一是从 4 万年前仪式和艺术出现到公元前 4 千年埃及文明和苏美尔文明之前出现的原始时代，这一时期人类关于美的观念在原始艺术中体现出来。

二是从埃及和苏美尔文明开始的五大文明到公元前 700 年轴心时代之前的神庙文化时期，这一时期人类关于美的观念用一种国家型的宗教艺术来表现。

三是从 2 千年前的轴心时代到 17 世纪现代性开始的时期，这时各大主要文化实现了哲学突破，用一种理性的思想来看待人和世界，开始了用理论形式来讲述美学，其中，西方文化从柏拉图追求美的本质开始，建立了一种谈论美的学科形式。而中国和印度，用的一种与西方文化不同的方式来谈论审美现象。

四是从现代性以来至今的 400 多年，这时，一方面西方文化的美学以自身的逻辑不断地演化，另一方面，非西方文化在西方这一世界主流文化的影响下，学习西方，按照西方的学科方式建立起了自己的美学，构成了西方美学与非西方美学之间的互动。

现在，单以人作为审美观察对象而言，美学有了新的论断，科学家认为所谓的美都是表象的一些特征，而本质的原因就是能够生育，而这种强大的生育能力，是因每个民族和群体通过长时间的观察统计出来的大概率事件，比如长发飘逸是一种美，那么是因为不健康的人头发稀疏泛白掉得快，所以长发飘逸至少是健康的象征，只有健康的人才能够更好地生育下一代。

比如以女性乳房为例：因为人类是哺乳动物，所以乳房非常重要。一般哺乳动物都由雌性担任，所以雌性健康的乳房就是漂亮的。人类乳房一般有两个，主要是因为小孩的哺乳期通常很长，而且健全家庭必须保证不低于两个孩子才能保证人类的延续，这就是有两个的原因吧。

再比如女性臀部：好的髋部能够保证胎儿健康发育。所以女性的臀部一般要比较宽大，一般上了点年纪的老太太在挑选儿媳妇时也会

去看她的屁股，事后会对她的儿子说“这个女孩屁股大能生娃”也是这个道理。

生育是实用为美的最终目的，可能更多的时候人们会忽略这个内在过程而直接看外表或者凭经验，其实美感和性感是大脑经过复杂计算后得到的一个最优答案，而这个答案一旦得到若干代的影响，就会慢慢形成一种本能遗传给下一代。

其实，所有的学说都不太对，所谓的美实际上是根据以往经验感觉相对于自己更好的人和事而已，只有这个理由才能解释为什么每个人都有自己的审美标准。

举一个美国发生的例子。我们发现美国人民经常在自己的家里种上草坪，这种草坪不仅不产生任何实际的价值，而且还要浪费很多的时间和金钱进行修整和保护。明明那么大的一块地方却不允许人们前去活动，明明是正常地生长却非要修剪，这在中国觉得是不可思议的事情。为什么很多美国人却还喜欢这种草坪呢？有人说一方面是因为古代人从事采摘活动所以喜欢有草坪的环境了，另一方面是美国人民被人为地教育——草坪是权贵和金钱的象征，所以草坪是美的是尊贵的代表和化身，我们认为第二点非常重要。

而这种感觉，既可以是实实在在的，也可以是一种联想或者仅仅是一种经验。正因为美是相对于自己而言的，所以当自己所接收到的信息（如环境、年龄或者身体机能）发生变化时，审美的观点也会随之发生变化。

正所谓，感觉对自己好的就是美的，而美好实际上是相对的。

1.3 我是谁

前面刚说到所谓的美，实际上是相对于自己更好的人，那么我们到底该怎么定义自己呢？我是谁？谁又是我？

1923年，弗洛伊德曾经在《梦的解析》一书中从心理学的角度把我分为本我、自我、超我三个意识形态，以解释意识和潜意识的形成和相互关系。“本我”（完全潜意识）代表欲望，受意识遏抑；“自我”（大部分有意识）负责处理现实世界的事情；“超我”（部分有意识）是良知或内在的道德判断。

他说，本我，是在潜意识形态下的思想，代表思绪的原始程序——人最为原始的、属于满足本能冲动的欲望，如饥饿、生气、性欲等。本我是与生俱来的，也是人格结构的基础，日后自我及超我都是以本我为基础而发展的。本我位于人格结构的最底层，是由先天的本能、欲望所组成的能量系统，包括各种生理需要。本我具有很强的原始冲动力量，弗洛伊德称其为力比多。本我是无意识、非理性、非社会化和混乱无序的。本我只遵循一个原则——享乐原则，意为追求个体的生物性需求如食物的丰富与性欲的满足，以及避免痛苦。弗洛伊德认为，享乐原则的影响最大化是在人的婴幼儿时期，也是本我思想表现最突出的时候。

自我也是人格的心理组成部分，是从本我中逐渐分化出来的，位于人格结构的中间层。其作用主要是调节本我与超我之间的矛盾，它一方面调节着本我，一方面又受制于超我。它遵循现实原则，以合理的方式来满足本我的要求。这里，现实原则暂时中止了快乐原则。由此，个体学会区分心灵中的思想与围绕着个体的外在世界的思想。自我在自身和其环境中进行调节。弗洛伊德认为自我是人格的执行者。

而超我是人格结构中的管制者，由完美原则支配，属于人格结构中

的道德部分。其位于人格结构的最高层，是道德化的自我，由社会规范、伦理道德、价值观念内化而来，其形成是社会化的结果。超我遵循道德原则，它有三个作用：一是抑制本我的冲动；二是对自我进行监控；三是追求完善的境界。

本我、自我、超我构成了人的完整人格。人的一切心理活动都可以从他们之间的联系中得到合理的解释，自我是永久存在的，而超我和本我又几乎是永久对立的，为了协调本我和超我之间的矛盾，自我需要进行调节。若个人承受的来自本我、超我和外界压力过大而产生焦虑时，自我就会帮助启动防御机制。防御机制有压抑、否认、退行、抵消、投射、升华等。

举一个汽车驾驶的例子。汽车是交通工具，司机有决定方向的权利。在开往目的地的过程中，我们可能会发现不太理想的情况，于是我们被迫做出一些改变，比如绕路、超车或者加油。那么在这个例子中，司机其实就是自我，车辆就是本我，而目的地则是超我。

本我、自我、超我之间的关系如图 3-1-2 所示。

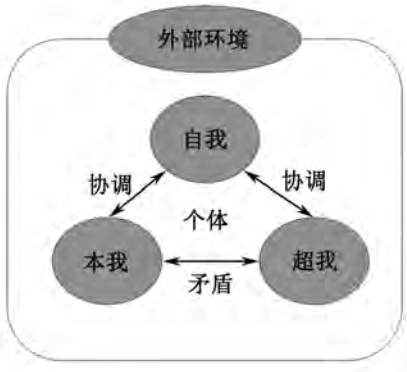


图 3-1-2 本我、自我、超我关系图

弗洛伊德从心理上给出关于我的关系图，其实我们还可以从物理与

逻辑的角度来描述和证明什么是我。

具体描述：凡是可以被意识所支配和影响的事物就是我的，这种支配和影响可以是直接的生物电信号，也可以是间接的语言命令和肢体表达。

同理，如果我的身体不受我的支配，那么这个身体就不是属于我的，即便我们生理上是紧密连接在一起的，但它也不属于我。

由于支配不完全或者影响力不足，所以有时本来属于我们的事物也将（或者暂时）不属于我。

1.4 生存的本能

衡量我的标准也许很难，也许很简单，不管怎么说，我都是一个不能重复的唯一，目前世界上有 70 多亿的人口，在这 70 多亿的人口中，虽然有很多同名同姓，甚至同年同月同日出生的人，但他们每个人都不尽相同，其性格同样也是五花八门的。也许你会感觉奇怪，这么大的样本量怎么还找不出一个完全一样的人来呢？

实际上，不仅是现在，就算是我们把历史上所有的人都加在一起，也找不出完全一样的人来，严格意义上讲，世界上没有两片相同的叶子。因为每一个事物都有其特有的属性，正是因为这种特殊性，我们才会把它单独看作一类事物。

即便是同一个人在不同的环境下也会表现不同，给人的感觉就像变了一个人一样，这种改变一般会伴有不同性格和技能，比如这种人同时拥有两个人的性格，我们常说他是多重人格。现实中这种人的变化还是非常常见的，其核心原因就是因为他认知和周围环境发生了变化。比如突然看了某人的传记，对里面的人物产生崇拜，他就很可能效仿那个

人，这就是所谓的名人效果粉丝效应，因为名人生活得好，让读者羡慕，如果读者像他那样做自己是不是也会变得更好呢？如果这时候这种思考存在矛盾，比如对名人产生质疑，他的性格就会间歇式地发生变化，给人感觉一会是他一会又不是他。

人的性格千奇百怪，很多性格更是让人惊讶不已，甚至让人们误以为是病症需要治疗。比如斯德哥尔摩综合征，这种病症是说有一天瑞典斯德哥尔摩的银行被劫匪劫持了，但是人质们不仅不记恨反而还要感谢劫匪，甚至有女职员后来还嫁给了其中的一个劫匪。与之相反的是利马综合征，在利马综合征表现下，被劫持的人员居然打动了劫匪。这两种典型的人格以其发生地而命名，所以这些都是真实的“病例”。

也许有人说这和恐惧有关，抛弃科技不谈，我们先来说说人们究竟为什么会恐惧。记得早年有一本叫作《恐惧的原型》的书，德国心理学家弗里茨·李曼通过本书，讲述了心理治疗的策略是挖掘个人早年的成长背景中导致恐惧的因素，研究个人与家庭以及社会文化之间的相互关系，以此培养我们与恐惧对峙的能力。

书中他以天体运行的自转与公转，向心力与离心力的四个特征类比我们普通人的四个成长动力。

自转：以自我为中心，肯定自我的存在，强调与其他人有所区别，我行我素，类似弗洛伊德的本我。

公转：把自己当作人群中的一分子，喜欢与人交往和分享，也愿意帮助他人，类似弗洛伊德的超我。

向心力：即万有引力，向往着团结与合作，同时也变得比较保守和固执。

离心力：随时准备改变自己，肯定所有的变革与新发展；同时也变

得比较容易背叛。

由上所述，四个恐惧的原型就是：

一是害怕失去自我，避免与人来往（分裂）；

二是害怕分离与寂寞，百般依赖他人（忧郁）；

三是害怕改变与消逝，死守着熟悉的事物（强迫）；

四是害怕既定的事实与前后一致的态度，专断自为（疯狂）。

给出四个恐惧的基本原型后，我们再分别详细讨论它们的具体表现形式。

分裂人格

这种人害怕别人亲近，与世隔绝、自我中心、社交困难、敏感、独来独往、特立独行、孤僻、理性主义者、坚强、冷静客观、自信、批评家、革命家、无政府主义者，他们往往不相信有人会爱上他，为此还会不断地针对自己的怀疑做测试，要求新的爱的证明。

另外，由于恐惧，分裂人格者会用不加修饰的愤怒来宣泄心中的不快，愤怒不光是起抵抗和保护的作用，他也会借此与人有了接触，攻击是他追求和争取某样东西的方式，正如奇迹课程里所说“所有的袭击都是求助的呼声”，这需要我们给他们足够的时间，慢慢填平心底的沟渠。

比如他们不相信人工智能可以为他们服务，为此他们会一次次地要求专家们做出更加可行的分析报告和实验结果。当专家们的结果不是让他们很信服时，他们就会采用过激的手段，来遏制人工智能的发展。

忧郁人格

这种人尽可能靠近他人，依赖、附属、害怕变成独立的自我，心甘情愿把自己交出去，任凭别人摆布，质疑自己的处事能力，没有归属感及安全感，谦虚、息事宁人、无私忘我、有同情心、感同身受、忍耐力强、卑微、从不要求什么、服从、配合度百分之百，乃至牺牲自己，夸张者奴颜婢膝；极力避免自转，害怕被孤立、分离、抛弃、不被保护和寂寞。他们往往有同理心，像爱自己一样深爱着对方，为对方而活，他们不太有自己的生活，但若是达不到目的，他可能以生命和自杀来解决问题，从而祈求对方的怜悯。

这种人，同样不喜欢太新鲜的事物，对于高科技同样感到害怕，但是他们往往不会直接说不，而是可能背后积极反对。

强迫人格

强迫人格者追求向心力，从小就渴望永恒和安全感，他们依赖熟悉习惯的东西，希望生活中各个方面都应该像铁铸的，变成永远有效的原则和无可争议的规矩。

他们害怕新鲜的事物和新奇的经验，风险、改变和消逝是他们最大的恐惧。他们是完美主义、乐观主义、自我控制者、洁癖、谨慎、自我保护、权力欲、强势、成就欲、责任感、强硬派、秩序、安全感、压抑怒气、好争辩、节省、吝啬、就事论事、具体不含糊、令人信赖、保守、传统等。

由于这种人小时候就很懂事，很早就学会了克制自己的情感，所以我们往往很难打动他的真心，这种人很头疼是吧？虽然头疼（其实这种

人他们自己就有头痛的毛病)，但也有解决办法，那就是我们帮他们指明其发展方向，告诉他们相应的技术原理和工作流程，让他们在我们的世界中再次找到方向，于是他们就会变得愿意接受我们，从而成为他们的另一套标准。

疯狂人格

疯狂（歇斯底里）人格者追求离心力，他们才不把那股聚拢且集中的万有引力放在眼里，只是翘首期待新鲜感，不断追求新鲜、惊险和刺激，随时臣服于外在的引诱和内心的呼喊。他们讨厌传统、束缚和既定的规律，害怕自由受到束缚，责任、义务和前后一致是他们最大的恐惧。他们追求改变与自由、好奇心强烈、追求刺激，喜欢冒险；热情洋溢、调情高手、花样百出、肯定所有新的事物、夸张、自吹自擂、装腔作势、爱慕虚荣、任性、口不择言、不想后果、及时行乐、不负责任、逃避束缚、自欺欺人、渴望被肯定、排斥最终的既定的人事物、害怕自由受到约束、爱冲动、无力抗拒诱惑、没有原则、易怒、脆弱、不理性、不合逻辑等。

这种人看似愿意接受新事物，但是他们只接受他们自己的新事物，对他人的贡献不屑一顾，总是和我们进行莫名其妙的竞争从而彰显他们的个性。

其实，这四种人格的形态适用于每一个人，是我们生命的一部分，我们既不完美也不完整，正是因为恐惧与希望的存在我们才能少走弯路，从而获得健康。

存在即合理，极端差异的人格非常具有代表性，正是这种差异让我们逐渐解开智能的神秘面纱。下面我们做一个有趣的小实验，为了简单起见，实验采用问答的形式。

(1) 假设现在我要把你的腿给锯掉，你会同意吗？我相信几乎 100% 的人都不会同意，好端端的为什么锯掉我的腿呢？

(2) 假设你遭遇了严重的车祸，必须锯掉腿你才能活下来，那么你会同意吗？我相信 90% 以上的人会同意。

对于同样的结果，人们却做出截然不同的两种选择，不知道你发现了其本质是什么吗？

对，本质就是能够更好地活着。

弗洛伊德说，超我是目标，本我是本能，自我是本能与目标的综合表现形式，其实他只说对了一部分，超我、本我、自我都是来自于人的内因，而现实中人在考虑问题时是内因、外因、本能和性能一起考虑的。

试问：

如果外界发生变化，而你不做出反应，你能说明你是聪明的吗？

如果外界发生变化，而你胡乱做出反应，你能说明你是聪明的吗？

如果外界发生变化，而你做出固定反应，你能说明你是聪明的吗？

如果外界早就发生改变，而你才做出反应，你能说明你是聪明的吗？

当然内因、外因、本能和性能在不同的环境下，其反应所占的比例会有不同，比如安静的实验室和多变的战场。

接下来我们再做一个有趣的小实验。在一个相对安静的实验室里，科学家通过脑电波测试仪来预测被实验者（也就是你）下一次可能举哪只手。在经过若干次的预测后，计算机预测得越来越准，你会发现即使你有意不配合，但是你举手的结果却早被计算机提前几秒钟计算出来了。

也许你感到很神奇，其实计算机只是先于你发现了你大脑计算的结

果，这种看似偶然的結果其实也是经过你的大脑一层层计算出来的。而这个计算的原理也非常简单，那就是如何能让自己更好地活着。

外因、内因、本能和性能影响最终决定的关系如图 3-1-3 所示。

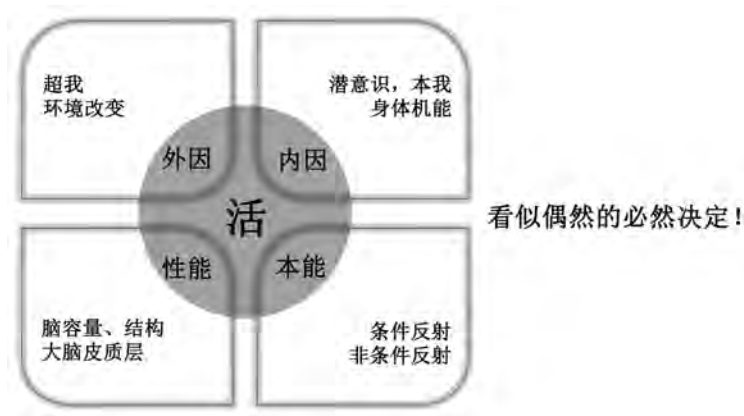


图 3-1-3 外因、内因、本能和性能影响最终决定的关系

这些也体现了生命的真谛就是更好的活着。

生存本能

正是因为更好地活着这个原则，我们才形成了形形色色的生命。活着是生命的基础，是智能的核心。只有活着并且更好地活着人类才会表现得有智慧、有感情。而活着本质上也是一种本能，所以我们可以把它称为——生存本能。

生存本能因其对生存的影响层级不同，所以表现出来的需求层次也不尽相同，美国著名心理学家马斯洛曾经为此专门提出了需求的层次理论，马斯洛需求层次理论很好地阐述了从最基本的活着到更好地活着之间的递进关系。

马斯洛需求层次理论如图 3-1-4 所示。



图 3-1-4 马斯洛需求层次理论

生存本能是人工智能的基础，也是强人工智能的未来，没有生存本能这个核心，人工智能将不会产生灵魂，也不可能有人人们所期望的爱恨情仇，那么，凭什么说生存本能是人强人工智能的未来呢？下面，我们就简析人的生存本能是如何进一步形成认知、经验和情感的。

认知

首先是认知，认知是知识获得和应用的一个过程，它是人的最基本的心理过程。认知包括感觉、知觉、记忆、思维等，是大脑接受外界输入的信息后，经过大脑的加工处理，转换成结构记忆的心理活动。

认知可以是我们的眼睛看见的、耳朵听到的或者鼻子嗅到的，甚至是皮肤触碰到的硬度、温度等。

本质上，所有我们能够感觉到事物都是可以记忆的。不存在不可以记忆的认知。我们无时无刻地在记忆这个事物的同时也在回忆这个事

物，记和忆构成了认知的主体。

人们在认知事物的时候，一般是多感官一起参与工作的，以认识小猫为例，当人们用眼睛看它时，我们不仅可以看到整体的颜色、纹理、光泽、大小、形状，当我们用手触摸时还可以感觉到粗糙度、质量、柔软度等。当然你可能也会听到它的呼噜声和脚步声。

如果我们把人的认知过程比作一个数据库，那么认知其实就是一条条记录的添加和查询。而由于这个数据库的大小有限，同时为了方便我们大脑快速查询，所以虽然看上去我们的感觉器官每时每刻都在认知事物，但并不是所有能够被感觉到的事物都会被记忆的，更多的时候我们是选择视而不见、听而不闻的。

我们的大脑只会记忆那些可能对我们产生重要影响的事物，也就是说这些信息对你而言是更有价值的。

例如，小孩对人脸是非常感兴趣的，但是对动物的脸却是一般，虽然是同样的视觉结构，但是我们的大脑却更加喜欢记忆那些能够对我们产生深远影响的人脸，尤其是父母和兄弟姐妹的脸。

还有研究表明，婴儿对人说话的声音也是非常敏感的，但对其他声音则反应迟钝。

认知不仅有选择性，还有难易性，有些认知非常容易，而有些认知则非常困难，那么之所以有些认知非常困难的主要原因是因为过往的经验告诉我们这个认知的价值不大。

认知就说到这里，基于生存本能，用现在的电子计算机来设计一个具有认识能力的人工智能并不困难，尤其是机器学习的效果更加明显，而我们唯一要做的就是根据其对生存本能的值大小有选择地记忆就可以了。

可是，由于大脑约有 1000 亿个神经元，所以说大脑的记忆是有限的。

经验

其次是经验，经验是现代人工智能的一次飞跃，专家系统虽然也是人工智能，但专家系统的智能更多的是专家的经验，而真正的机器经验或者机器智能却少之又少。

科学家们之所以说现在的人工智能是弱人工智能，就是因为现在人工智能的智力水平也就相当于两三岁的小孩，甚至很多方面还不如一个刚出生的婴儿，我相信大家也很好奇，为什么刚出生的小孩就有很多的本能反应呢？为什么人类会自我学习和成长呢？

其实这一切都和经验有关，经验则是人类从生活多次实践中得到的知识或技能，比如车辆驾驶就是一种技能，勾股定理和你做的实验就是一种知识经验，从表面上看经验和认知是两个不想干的词，但实际上经验是已有认知的一种快速检索，就像是数据库的索引一样。

例如，在车辆驾驶时，我们转动方向盘，或者踩油门，经过多次练习之后我们找到了一种更好的驾驶方法。其实每次练习都可以表示为一种稍微复杂的记忆，像转动方向的力度、速度、时间等。然后在考试时我们会想到以前学过的驾驶方法，告知自己在当前环境下用某种方法将更加有利，这种回忆其实是一种搜索，根据搜索结果按照满意度（生存的价值度）排序，并从中选择自认为最好的一组动作。

为什么需要一种快速回忆呢？主要是因为正常情况下大脑的运转速度相对于 1000 亿个神经元而言是有限的，这种有限性一方面受制于神经细胞传递的速度，一方面受制于外界环境给我们反应的时间。所以大脑会把相似环境下（如车辆转弯打转向灯）的多次重复的记忆变成一种经

验。这种经验虽然不一定是最好的选择，但是从概率上来看，确实是一种相对好的选择。怎么样，这种指向明确的相对好的选择，是不是很像数据库的索引。

也许你会说，其实索引也分快慢，是的，其实人的认知经验也是有快慢之分的，一般来讲，当我们认为某种经验更好时，我们的大脑就会把这种经验变成一种习惯（如遇到紧急情况踩刹车）。如果这种习惯在实践中被认为是最佳方案时，就会成为一种本能，本能其实就是一种条件反射。这种条件发射就好比巴甫洛夫那条可怜的狗，本来应该对肉流口水结果变成了对铃铛流口水。而更有趣的是，当这种本能被历代证明是最好的时候，这种本能便会通过 DNA 遗传到子女身上。比如子女们基本的相貌，原始的审美标准，还有坐立行走、喜怒哀乐这样简单的动作和表情，甚至还会说妈妈这样基本的母语。如果把习惯比作一个更加快速的索引，那么本能就是索引的索引。

根据生存本能，我们假设：人们规定，孩子哭就是需要人照顾的一种表达方式，笑就是表示满意的表达方式，孩子觉得不错，当他需要照顾的时候就哭，然后时间长了就哭得有经验了，知道什么时候要哭，什么时候要笑了。这种“会哭的孩子活得好”后来就会变成本能，当本能找不到更好的方法改进时就会遗传给下一代。所以人类孩子会哭是天生的，而动物却不见得会哭。当然这个假设是需要科学验证为前提的，如果哭声传得不够远，那么人们的这个规定就是不合理的，不合理的迟早会被更合理的经验取代，比如其他某种通知方式适合告诉周围的人他需要照顾，那么这种方式只要被证明是最有效的就会被优先继承下来。

那么第一批孩子究竟是怎样学会哭的呢？可能是人为的，父母感觉孩子快饿的时候，诱导他哭，这点可以请教下巴普洛夫。但实际上更可能是物竞天择的结果，突然有个孩子会哭，因为哭声可以传播得更远，所以父母可以很远就能够听到了，而不是放到育儿袋中了。

哭为什么是悲伤的呢？主要原因是痛苦关联效应，假设孩子哭，总是关联到孩子挨饿受冻等难以生存的痛苦，那么孩子一哭就会觉得伤心，反之亦然。我们做一个假设，每当孩子要饿的时候，我们就给他一个铃声，但是第一种情况下铃声过后孩子仍然得不到吃的，第二种铃声过后有吃的，那么同样的铃声在前后两个孩子听来一个是悲伤的一个是幸福的。

再假设，我们规定，怒目圆睁、声嘶力竭、身体抖动这种表情，是他要对某些事物进行报复和破坏的声明，那么这种方法被认可和继承后，人们觉得需要报复和破坏的时候自然而然地就选择发怒的表情了。所有这些表情的背后其实有两个作用：一个是充分活跃自己的身体器官，为即将到来的战斗做好准备；另一个是向别人宣示自己的强大，以免不必要的冲突。比如猩猩拍打自己的胸脯就是一个很好的宣示——你最好不要轻易招惹我。

另外人有破坏性也是经验的一种，破坏的经验告诉他这是一个更好的选择，可以减少竞争对手，可以改善生存环境。如果生活中破坏和动怒容易达成结果，那么经验自然会他成为一个容易动怒的人。另一种情况是，如果某人经过深思熟虑然后决定报复和破坏，他就不用选择这样的表情。这种报复和破坏也是一种相对的选择，受至于施暴者本身认知的局限，因此他选择残害或者奴役他人的主要原因是因为他的认知告诉他这是目前最好的选择。

我们再简单总结一下，认知、经验、习惯和本能之间的递进关系：正常情况下我们是通过认知来生存的，但是为了应对紧急情况，我们通过多次相似的认知形成了经验，并在这个经验得以长期有效的验证后形成了习惯，而这个习惯在被若干代确认是好习惯之后就形成了本能。

认知、经验、习惯和本能，四者之间是一个递进关系，一般相对来说，越往后反应的时间越快，同时其可塑性也越差。可塑性差说明，本

能是不会轻易发生改变的，但认知却是可以很容易变化的。

经验一般来说是通过学习、训练和总结得来的，是对大脑已有认知的聚类。当我们把经验变成数据库索引时也许会发现，有时候我们给出很多查询条件然后却只得到一个结果，其实这个唯一的结果就是聚类，而查询出多条记录本质上就是联想。

另一方面，为了方便计算机计算，我们需要把感知到的事物按照时间序列通过不同的感觉器官分门别类地存储到不同数据库当中，而这些可以存储的记录必须是静态的，就好像是电影中一张张胶片，每一个记忆片段都是一帧静态的画面，这些静态的画面在数据库里就是元数据。

用过计算机的人都知道，每秒钟处理图片的数量决定了计算机的性能，其实人脑的反应也是有快慢之分的，和计算机的晶振一样，人体大脑中也有一个专门的细胞负责计算时间，但这个计时器本质上也是一种感觉器官，所以说它是感觉器官是因为时间也是有快慢之分的。

我们习惯了钟表，可能对时间有快慢之分觉得不可思议，实际上这种快慢也是人类进化的一种结果。

比如危机时刻，我们的大脑就会高度兴奋，对相关的器官调用更加积极，这时候我们感觉大脑所接受的信息非常多，由于大脑平时只处理100个画面，现在一下子就处理了1000个画面，所以我们感觉时间变慢了，那么为什么危机时刻要处理那么多画面，其核心就是为了更好地解决当前的危机，只有获得的信息越多，处理的速度越快我们才有可能活得更好。比如发生车祸或高空坠落的时候，大脑对信息的反应不再是觉得无所谓，而是每一个细节都可能关乎生命，要么死亡要么生存。因为这些相对时间内的信息的权重非常高，所以大脑相对时间内处理的信息就变得多了起来。这也是人们为什么觉得慢镜头非常真实合理的原因之一，所以急中生智还是有道理的。

也许看过动漫《加速世界》的你会问，既然大脑越兴奋处理的信息越多越快，那么为什么我们平时不这么兴奋呢？为什么呢？既然我们可以成为超人，但我们却还是选择作为普通人呢？

带着这样的疑问，我们不妨观察一下大脑兴奋的结果就是它需要更多的休息时间，一方面是它需要消耗大量的能量，本来我们大脑就消耗了 80% 的葡萄糖，这时候可能出现负消耗，需要从其他细胞中摄取大部分能量，所以大脑兴奋过后需要更多的休息时间。一方面大脑在兴奋时处理了大量的信息，但这些信息事后都是需要整理的，人类在休息尤其是做梦的时候就是对已有信息的加工整理，否则就容易发生大脑混乱的问题。

当然关于时间可变性，韦尔奇曾经说过，大脑对时间的切片是以其重大事情为切片的，每一个切片都是一个重大事情的分界线，由于人生每次重大变化，如胚胎、出生、说话、成熟、衰老、死亡物理时间上越来越长，所以心理上时间过得越来越快。人们主观上常常把重大事件的时间间隔作为我们的时间，这种所谓的重大事件随着年龄的增长，时间间隔也越来越长，所以对成人而言，主观时间单位要比小孩更长一些。因此我们长大的过程会觉得时间过得越来越快。这也就是成人眼里的几年和小孩眼里的几个月差不多。

实际上，年龄越大，生活会越平淡，时间也会过得越快。

宠物情节

前面，我们说了生存本能，生存本能是未来人工智能的核心，但是由于我们的神经细胞传输速度有限，很多时候我们只能选择相对有效的经验和本能来生活。

由于经验，尤其是本能都不能在很短的时间内改变，所以在某种特

定环境下形成的经验和本能在另一种环境下就显得不可思议。

也正是这种相对有效的选择才造就了现代人类的进化和多样性，下面我们来说下非常有趣的宠物情节。

我们假设：某些人规定每年的某个时候是重要的节日，并且规定这天要庆祝，那么这个看似人为的规定，当得到某个组织强制推行后，大家会发现，人们这一天真的要去庆祝，甚至很多人都是发自内心的，更有甚者把它当作一种“信仰”和乐趣。

而作为一个无信仰的人，则感觉非常奇怪，甚至你会发现如果一个虔诚的教徒来到一个陌生的国度，他们的举止和风俗都会被当地人视为怪人。

一般来讲，宠物情节说明了相对弱小的人类可以被改变、被教育，而造成这种结果的直接原因就是由经验造成的，因为经验是一种相对正确的选择。因此我们常说经验有其相对性。

那么为什么有人这么重视经验，而有人又非常喜欢全面地思考呢？我想主要有以下两个原因。

一种是常说的脑子反应慢，出现反应呆滞和迟钝，这就好比是数据库一时间找不到结果，计算机就会出现假死状态，而脑子慢是相对的。造成脑子转得慢的核心原因就是数据量相对大。我们都知道数据库查询慢的原因是数据量太大了，之所以说相对，是因为计算机性能不同造成的，早期计算机当时确实很先进，但是现在一台PC就相当于几台老式巨型机了。这就好比拿现代人和猿人进行比较一样。所以说脑子转得慢是相对的。脑子慢的原因很可能是你知道的东西太多了。这也符合人的常识，知识越丰富的人越是不敢快速地做决定，优柔寡断说的就是这种人。但更多的情况是因为人类认知不足造成的，也就是数据库出现了查无此项，不知如何应对而凭天由命的被动局面，无知说的就是这种人。

第一种就是典型的大智若愚了，他不可能一直被虐待着，影视作品中这种人往往进行所谓的黑化和叛变。而第二种人，则先有一定的不甘心，在不甘心的时候，他们往往会先利用以往经验来进行反抗。当然如果这个人的认知很高，他就不太可能出现这种被奴役的场景，好比两个拳击手，胜者怎么会屈服呢。另外，如果人类整体认知上升，那么相对落后的经验也会被抛弃，比如新宗教的诞生，新国王的出现。封建制度就是这样瓦解了奴隶制度。

我们可用一个看似简单的算术题来说明上述问题：

比如有人问你 $1+1$ 等于几时，小朋友很快回答是 2，而我们则不知道问题具体是指十进制的还是二进制的。于是有经验的人很快就说是 2，因为经验告诉他，问他的环境默认是十进制的，2 是最好的答案，十进制的应用场景更多。没经验的会反问，你问的是十进制的还是二进制的？如果在计算机学者圈子里我会回答是二进制 10 的，否则是十进制的 2。

也许计算机学者会觉得小朋友比较无知。但是那些计算机学者何尝不是被数学家嘲笑， $1+1$ 有很多解释，甚至很可能就是一进制的 11。

小朋友对计算机学者崇拜不已，但对数学家嘲笑不已！不是有一个让人深思的笑话么：“一个人上了三天学，第一天学了一，第二天学了二，第三天学了三，结果三天后他就退学了，然后有一天他要结婚了，给一个姓万的朋友写请帖，结果一天连一个姓都写不完。”

一般来说，对于相对无知的人只要你适当引导，他们就会跟着改变，比如被劫持的人质、监狱中的犯人以及其他被胁迫的人们，而当这些被胁迫的人发现顺从可以获得更好地活着（生存本能）后，他们就开始有意地向胁迫者示好。

如果你不相信人类有宠物情节的话，可以设想下：你是否敢于私自出国，你是否敢于英雄救美，你是否能够穿越沙漠，你是否能够离开地球。

另一种是三思而后行的，这种人做事情是经过深思熟虑的，当然这种深思熟虑还是建立在他已有认知的基础上的，所以才造成了结果的参差不齐。全面思考的人对计算机性能的要求非常高，否则还没想明白就结束了。

虽然经验和本能有着一定的局限性，但经验确实是快速选择中最好的，也是目前人工智能的主要研究方向，而这个选择一定要在自己力所能及的环境内。为了更好地生存，人们需要在加快认知的同时加快大脑运转的速度，而不是一味地逆来顺受。如果是计算机的话就在增加软件的同时增加 CPU 和磁盘的运转速度。还有就是改变自己的思想，对那些可能碰到的事情，提前练习从而增加被侵犯的难度。

由此可见，人工智能需要强大的计算量。

情感

最后我们说下情感，也就是科学家口中爱的方程式，人们常说人工智能的难点在于它不懂得爱，爱是高级智能的特权，其实不然，现在我们就来看下人们所谓的喜怒哀乐。

我们知道哭声可以传得很远，所以当这个孩子遇到困难需要被人照顾时，他就可以通过哭声来进行求救。那么假设周围没人管他，孩子知道哭也没用，那么他还会哭吗？其实实验的结果很简单，虽然遗传基因或者本能告诉他，这个时候哭就能获救、就能获得满足，所以他才哭。但是结果他发现哭根本就没有作用，也就是说这个本能反应是错的，所以他就学乖了，不哭了。

如果哭是一种表示需要照顾的方法，那么笑就是表示满意的方法。笑可以是人类之间约定俗成的一种表达方式，当这种表达方式被认为非

常有效时，满意的时候就会不由自主地笑了。如果我们告诉美女们笑可能产生皱纹，你猜久而久之她们还会笑么？

就像哭有很多种一样，笑也有很多种，每种表达方式都是祖先们经过多少代约定俗成之后才遗传给我们的，虽然说是约定俗成的，但实际上它还是受制于人的机体内在感受的外在表象。

基于生存本能，人类所做出喜怒哀乐选择的主要参考阈值为是否对自己更有价值，以受伤为例，坏人就是欺负你的那些人，而好人则是帮助你的那些人。他们之所给你的感觉是不同的，就是因为你告诉他们对待你的结果是不同的，那么如果我是你，会不会亲近那些对你好的，而远离甚至报复那些对你不好人呢？即便你想告诉我你有受虐的倾向，其实你骨子里还是想获得锻炼和成长，尤其是想获得所谓的突破。

用生存本能，尤其是相对友好的选择就可以很好地解释情感，其实它本质上就是一个“趋利避害”的过程。

作为感情之首的爱同样是出于这种相对性选择，既然是相对性选择就注定有些选择可能不是正确的，不仅不一定正确，还可能看上去很变态。比如一个为了家庭生活而上战场的人，从某种意义上说他是杀人犯，他的家人一定非常爱他，希望他早日归来。但是，当配偶死后或生活不能自理，当事人就容易移情别恋了。表面上大家都说爱是伟大的、神圣的，看似和物理一点搭不上边，但实际上是你骨子里告诉爱上他相对于现在你会生活得更好。所以我们经常发现一般对于生活圈子比较小，认知比较少的人来说，爱上身边的人更容易，因为这些人深深地影响了他的生活。

那些对他生活影响甚微的人，他是很难产生爱情的，即便是上帝你也不会真正地爱上他。比如海外某地发生了海啸和你家漏雨了哪个危险大？因为海啸离你实在太遥远了——除非周边环境或有人告诉你，你关

心它们你将生活得更好。

爱是一种友好的表示方法，是一种付出的方式，它让你明白你对它好、重视它就能给你回报。爱是相互的，单方面的爱都是自私自利、无法长久的，爱需要相互了解，只有相互了解才能更好地相互给予，但是太了解很可能被所谓更好的人选所动摇。恋爱是和优点打交道，结婚是和缺点过日子，这句话一点也不假。

总的来说，爱和哭、笑一样，也因为目的不同有很多种——情感是认知和经验的一种升华。

也都是基于数据的一种选择，正是基于这个原理，我们才能制造出拥有认知、经验和情感的机器人。

正是基于生存的本能做出的相对有效性选择，并由此经过多次验证后才有的经验、习惯和本能才造就了我们五花八门的性格，如图 3-1-5 所示。

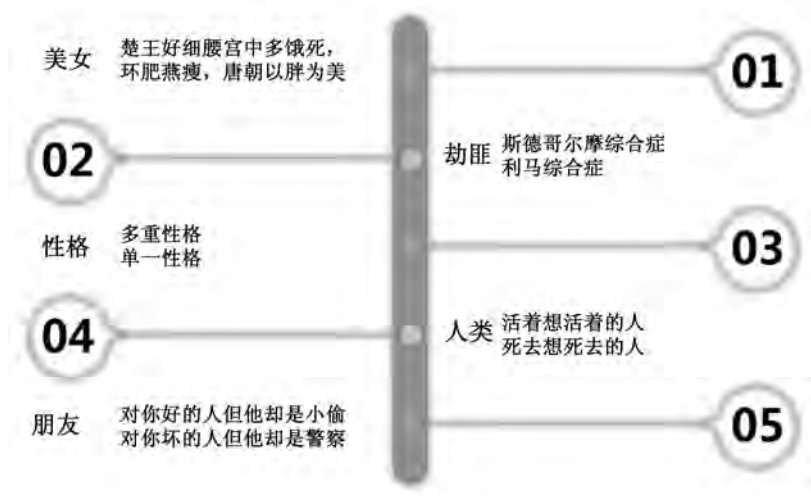


图 3-1-5 五花八门的性格

什么是智慧

本章最后一部分内容，我们来用智慧这两个字做一个小结。什么是智慧呢？其实智慧和神经系统一样也是一种物理现象。那么什么是智慧生命呢？智慧生命简单来说就是更聪明的生命。从现代字义上来解释就是能够利用丰富（慧中的第一个丰字）的传感器（智中的知，知字有知道感知和识别的意思）和丰富（慧中的第二个丰字）的控制器（智中的曰，曰字动词有说和表现的意思）做出保护心脏（慧中的心，心字有心脏中枢和要害的意思）的相对有效性选择（慧中的曰字，曰字有打点整理的意思）。

有人说心应该是指大脑而不是心脏，为什么智慧的慧字保护的是心脏而不是大脑呢？因为我们这里一方面是尊重了汉字的本意，即心就是心脏，另一方面心脏是动力之源，没有心脏一切生命活动即将停止，包括大脑。而没有大脑我们还可以说他是个傻子，比如很多生物就是没有大脑的。

心脏是生命得以继续存活的基础，因此我们才说智慧生命的本质就是一种为了让心脏能够继续存活而做出的相对优秀选择的D A N个体。

由于这种本能的选择依赖于大量的感知和强大的控制，所以智慧生命看上去往往比其他生命更聪明、更强大。

动物的传感器和控制器目前还相对弱小，在瞬息万变的环境中难以做出完美的选择，它们为了更好地生活，便根据以往的成功经验，强迫自己拥有了一种快速反应的本能。这种本能再经过多代的验证后，将以基因的方式遗传给自己的后代。这种基因因其验证时间长短不同也有其强弱之分。

比如动物最基本的呼吸、生长、爬行、哭喊和睡觉，再到很容易学

会的听、看、嗅、摸、吃，以至复杂的交流、奔跑、狩猎等。

搞清楚智能（拥有智慧的能力）之后，我们再来看看什么是强人工智能？强人工智能就是人工地创建一个拥有类人智慧生命的机器来，就好像女娲造人和皮格马林翁雕刻那样。

现在制造强人工智能，电子计算机是首选，因为两者的结构和功能是如此相似，而电子计算机又是我们可以控制和制造的，下面我们给出一张电子计算机和人类智慧的对比图，如图 3-1-6 所示。



图 3-1-6 什么是智慧

第 2 章 本能的演化

2.1 核心本能设置

在第 3 篇第 1 章中，我们说明了生存本能的核心重要性，生存本能之所以非常重要是因为生存本能理论上是唯一一个不可以更改的本能，而除了生存本能之外的本能都可以随着时间的改变而改变。

生存本能是强人工智能的核心，但这个核心可并不一定是计算机的电源，它理论上可以是任何一个可以被数值计算的量。比如我们想做一个炒股的强人工智能，那么这个炒股高手的核心本能就是账号资金的增减，账户资金越多说明活得越好，账户资金越低说明活得越差。再比如我们制作一个保镖机器人，那么这个机器人的核心本能就是被保护者是否能够活得更好，如果被保护者受伤或死亡，那么这个机器人注定就是一个要被淘汰的机器人。

相比炒股机器人，保镖机器人的核心本能可能更难设计，所以在强

人工智能初期，我们一般都给机器人明确的、可以很好量化的使命。比如一个数字的增减，一个物体的存在与否。

也正是因为如此，初期核心本能的设置才要求为数理上可以被准确计算的量，那不可确定的物理量目前不建议作为核心本能。

至于核心本能的计算方法也可以因人而异，比如大小，距离，中数，众数、方差、均值，二项式的 G 点都可以。

核心本能就像是上天赋予我们的使命，只有使命存在，我们才能更有目标地活着，只有我们活得有目标，我们才会更加强大。

《机械姬》

电影《机械姬》的女主角爱娃（见图 3-2-1）是一台漂亮的人形机器人，这台人形机器人之所以是这样子“漂亮”，完全是因为它的制造者是为了讨男主人公的喜好，电影的男主人公是一个自认为计算机水平很高的程序员，并在全球最大的搜索引擎公司工作，然后突然有一天，他赢得公司的幸运大奖，可以和老板共度周末，故事就这样展开了。



图 3-2-1 电影《机械姬》的女主角爱娃剧照

老板果然很有钱，直升机都要在他的领地飞两个小时才能抵达。老板的别墅在山林之间很隐蔽，是一座非常高科技的研究所。

原来男主被邀请的真正目的是帮助老板对他的新型机器人（爱娃）进行图灵测试。测试过程分为若干环节。

首先不得不说的是这个老板，他实际上就是另一种意义上的上帝，他利用这栋先进的研究所创造了很多机器人，不过他创造的机器人一般以女性居多。机器人所有的肢体动作、面部表情和知识都来源于他的搜索引擎公司和智能设备影音数据。影片说他直接黑进了智能设备的影音数据，所以才能获得这些数据。

因此严格意义上来说，男主人公看到的机器人爱娃实际上就是一台人为的大数据分析和定制好的人形计算机。

影片最为核心的一点其实不是故事情节，而是老板有意地给机器人设置了一个走出迷宫到达某个十字路口的本能，这个迷宫其实就是他的隐蔽的、安全的研究所，爱娃的所有选择都是依据这个核心本能而做出的表象反应。

机器人到爱娃这一代，已经经历过很多版本了，每一代升级，都有核心的记忆和数据被继承了下来。这也是爱娃非常聪明的原因。以至于它能够通过逆转磁盘然后让电力过载，从而实现暂时停电，这可能是上代机器人某次事故造成的，也可能是从互联网上得到的方法。

其次，老板为了证明爱娃变得聪明，他在选择测试员的时候实际上是别有用心的，比如男主人公没有女朋友，喜欢看色情影片，有同情心，乐于助人。爱娃其实就是通过搜索男主人公喜欢的艳星而专门做出设计的机器人，所以我个人认为老板是有意放走了爱娃。

爱娃让男主人公动心的原因，都是以他这个人为中心的，记得影片

中的一个镜头说，他曾经争论过性是因为繁衍后代而产生的变异。这其实是人类美学的一种理论——即实用为美。

机器人爱娃为了实现走出实验室的目的，她尝试各种方法，让男主角感受到帮助她的美好。由于她从前代机器人那里知道，直接通过老板是很难实现这个目的的，所以男主角过来之后给了她新的希望和相对友好的选择。

最后，爱娃通过交谈和视听唤醒让程序员联想到自己本身所经历过的一些事情，产生一定的共鸣，如相貌、动作、语言和服饰。然后男主角在接收这些感知后，也会根据他的认知做出对其相对更好的选择。他可能在想爱娃被废弃实在太可惜了，如果爱娃能和自己在一起就太好了，可惜结局和老板一样很不幸，男主角被机器人骗了。

影片的最后说，机器人爱娃成功来到了人类十字路口，实际上她直到现在才开始面临真正的选择，但是由于逃出实验室之后的核心本能已经不存在了，所以她在目的达到之后就迷茫了。除非她再次被关押然后再次激活她的核心本能。

当然也有人认为老板设计的这个逃出迷宫的本能非常好，因为这防止了爱娃的进一步升级。这也不失为强人工智能初期的一个思路。

正所谓梦想不同，幸福的定义也有所不同。

不可篡改

《机械姬》中的机器人的核心本能设置可以是一个简单的距离，比如实验室到十字路口的距离，也可以是复杂的认知图像，比如找到十字路口，《机械姬》中的机器人的核心本能和我们的炒股机器人的核心本能在某种意义上是容易被机器人篡改的，只要你存储这些核心本能的区域被

计算机篡改，那么它就可以轻松地改变自己的核心本能。比如爱娃原来是走出实验室，现在改成待在实验室里，比如炒股机器人原来是挣钱是好的，现在改为亏钱是好的。这种改变一方面来自于计算机的权限，一方面也可能来自黑客的攻击。

为了防止篡改核心本能，我们经常采用加密的方法，或者把其列为机器人禁区，禁止机器人或者他人篡改核心本能的设置，比如账户数值的大小。

《机器纪元》中说的是科学家们利用超级计算机制造出一个不可篡改的生物芯片，并且利用这个生物芯片创造出了他们理想的机器人，不过结果是这个超机器人在生物芯片上留了后门，从而催生了机器纪元（剧照见图 3-2-2）。虽然用计算机来控制计算机也不是一个绝好的方法，但这类防篡改工作却是需要我们这些创造者一直进行下去的，这既是对机器人也是对人类负责。



图 3-2-2 《机器纪元》剧照

潜意识

强人工智能还有一个主要特点就是它会主动和你说话，这种主动实

际上是受到潜意识支配的。

我们人类是怎么定义潜意识的呢？难道就因为计算机屏幕上出现一个“我饿了”，就表示计算机想要充电么？

强人工智能不是功能的模拟，也不是结构的模拟，它是经过计算机自我计算之后得出的相对友好的选择。

其实在它说出这句话之前很早就已经计算好，最后的输出只是一个尝试性的结果而已，在这个结果输出之后，它要继续把人类反应的结果记录下来，从而验证自己说出这句话的效果。

计算机和人一样，不是开机就拥有强人工智能的，它需要持续地运作一段时间后才会有智能。而这种被自己忽略的计算量实际上就是人类所说的潜意识。

潜意识是人们不能认知或没有认知的部分，是人们“已经发生但未达到意识状态的心理活动过程”。潜意识分为前意识和无意识两部分。

无意识、前意识和意识是三个依次递进的不同层次，意识是一个大超市，在超市中有我们需要的购物区，一般意识就在这个购物区中，是否进入购物区需要我们人类依据自己的需求做出决断，而无意识则是指我们此时不太需要关注的其他购物区。

潜意识也是意识的一部分，同样属于思考过程，只不过是被动我们压抑或者隐藏起来的那部分意识，所以潜意识一般不会被注意到。

就像我们去大超市买东西一样，虽然超市的商品琳琅满目，但是你会只会关心你所需要的那些东西，而对其他东西不会过多关注，不过这一切都不是既定不变的，如果超市有促销，或者你发现了更有价值的东西，这些潜意识的东西立马会转变为意识。

潜意识一般还指那些不需要我们过多关注就能够顺利达成的任务，这种任务有时候非常重要，比如我们一直思考一个重要的问题，由于这个重要的问题需要大量的计算，所以我们单独开辟一个计算区域，平时的时候不会过多关注，只有结果出来时我们才会感觉突然顿悟。

潜意识是无时无刻都存在的，就好像计算机的一个后台程序，只有它执行完了之后我们才意识到。既然潜意识对工作环境不怎么挑剔，那么计算机是不是也需要像人一样进行所谓的静一静、想一想呢？

我们认为计算机也需要工作与思考上的时间分配，尤其是强人工智能，因为随着计算机的长期运转，设备和接口的不断增加，数据量越来越大，更多、更大的索引表，更多层次的迭代查询，这都使计算机并不容易做出相对友好的选择。

它需要把那些杂乱无章的数据分门别类地整理，并关联彼此之间的索引，而这样一个工作最好是在少量新信息摄取的情况下进行的，如果这时候有大量的信息进入就会给原信息处理增加不必要的麻烦。所以计算机也是需要静一静、想一想的。

人类做梦其实就是对已有认知的一次整理，如果我们把每一个认知当作一张静止的画面的话，那么做梦就是对这些已有画面的剪辑，尤其是最近才发生的画面，所以人累了就容易困，因为大脑告诉你今天记录的信息已经够多了，需要整理下才能够正常思考。

一般人做梦或者思考是有声音、有图片的。但对于聋人、盲人而言，由于他们缺少了相应的视听感官，所以他们做梦和思考方式和正常人不同，他们会借助其他感官完成做梦和思考，比如形状、温度和气味。

正所谓我思故我在。

2.2 数据就是能量

我们制作一个强人工智能，需要给它设置一个核心本能，并让所有的选择都围绕核心本能的价值来进行计算。而这个计算的过程基础就是收集、整理、分析和提取无穷无尽的数据。

狭义的数据包括传统企业级数据、互联网数据和新兴的物联网数据。广义的数据（或称大数据）是指“无法用现有的软件工具提取、存储、搜索、共享、分析和处理的海量的、复杂的数据集合。”业界通常用4个V（即Volume——体量大、Variety——种类多、Value——密度低、Velocity——速度快）来概括大数据的特征。

体量大，截至目前，人类生产的所有印刷材料的数据量是200PB，而历史上全人类说过所有的话的数据量大约是5EB(1EB=210PB)。

种类多，相对于以往便于存储的以文本为主的结构化数据，非结构化数据越来越多，包括网络日志、音频、视频、图片、地理位置信息等，这些多类型的数据对数据的处理能力提出了更高要求。

密度低，价值密度的高低与数据总量的大小成反比。如何通过强大的机器算法更迅速地完成数据的价值“提纯”，成为目前大数据背景下亟待解决的难题。

速度快，大数据区别于传统数据挖掘的最显著特征。根据IDC的“数字宇宙”的报告，预计到2020年，全球数据使用量将达到35.2ZB。

如果我们把人工智能比作大脑。计算机比作机体，那么大数据就是血液，就是能量。数据的价值因为载体不同、处理方法不同，所产生的价值也不同。

现在，无论是传统行业还是新型行业，谁率先与人工智能融合成功，谁就能够从未来的金矿中发现它所蕴含的价值。而作为这个价值的载体——数据则主要从四个层次来体现其应用的价值。

一是元数据。元数据是对数据本身进行描述的数据，它是不可细分的原子，具体对象的一个属性、一个特征。例如，一辆汽车本身是数据。而这辆车的品牌、型号、价格、出厂日期、设计师等，就是它的元数据。元数据的价值主要有两点：一是能够从侧面描述对象；二是可以结构化。这一层级的数据是最原始的生命形态，也是其他高级数据的基础。

二是辅助决策。元数据非常简单，如果仅从元数据的角度来处理事务就很可能出现盲人摸象的尴尬，而如果我们把这些元数据适当地进行组合，效果就会好很多。比如我们来通过用户对化妆品关键词的搜索和相关页面的停留时间，我们发现 28 岁以上的女性对知名品牌的化妆品有了一定的忠诚度，而 25 岁以下的则灵活多样。这一层级的数据主要用来为产品提供决策支持。

三是提取价值。在数据的数量和质量达到一定程度后，元数据就会发生一个质的变化。因为当数据足够多，多到比专家知道的还多时就已经可以替专家进行决策了。比如现在的银行信用记录，已经完全可以实现秒办信用卡的业务了。之所以可以代替传统营业员快速办理信用卡这种业务，主要是因为行中早就把你的信用记录进行估值了。即便是因为政策规定我们仍然通过传统营业员来办理信用卡业务，但营业员本身还是通过这套系统来办理的，从营业员的角度看，这套征信系统其实就是一套实实在在的服务。这一层级的数据本身就是产品。

四是改变客户。我们之前也许不怎么觉得，但是现在的我们越来越发现，很多的用户画像、精准营销或者智能系统，正在不断地改变我们。比如我们使用的浏览器总是出现以前所关注的商品，这种做法的出发点是好的，但结果就是我们已经失去了自我。当我们同样再使用这个浏览

器的时候，那些可能对我们有用的新商品、新文章最终将与我们擦肩而过。这一层级的数据已经拥有影响力了。

食物链

数据的四个层级，就像是生物界的一个食物链，随着数据层级越高它所产生的价值也就越大，大到可以影响消费者。

食物链这个词，最早是英国动物学家埃尔顿提出来的。例如，虎鲸吃海豹，海豹吃乌贼，乌贼吃鱼类，鱼类吃软体动物，软体动物吃浮游生物。

虽然生态系统中的生物种类众多，但是每种动物都在这个生态系统中扮演演着重要的作用，这种作用像一个链条一样缺一不可，如果中间的任何一个链条断裂，那么生活在这个生物链的生物都将不复存在。

例如，草原上，狼吃羊和马，是人和牲畜的大敌，但是狼也吃田鼠、野兔和黄羊，田鼠、野兔、黄羊等又吃草，草又是羊和马的主要粮食，羊和马又是人的主要食物来源，草原是一个伟大的母亲，养育着她的子民们，这些生物组成了一个庞大的生物王国，形成了环环相扣的食物链，它们相互制约、相互繁衍，与草原共同生存了几万年。可是有一天，牧人们来到了草原，看到狼吃牛羊，觉得狼是牛羊的大敌，于是就采用了各种方法消灭狼，甚至射杀狼群。可是他们忘了，狼对于草原也是有利的，因为狼也吃田鼠、黄羊等草原上的大害，才使得草原上没有太多的田鼠、黄羊，这样也保住了绿草，使得牛羊有充足的食物来源。经过一段时间的杀戮，终于有一天，狼群被杀没了。他们以为这样牛羊就会多起来，可是事情的结果并不是这样的，狼口脱身的田鼠、野兔、黄羊等大量繁殖，将一大片一大片的绿草吃光，经常地将草连根拔起。草原渐渐失去了青青绿草，处处是裸露的黄土，一旦起风，黄沙漫天，许多地

方从此变成了沙漠，而牛羊因为没有了鲜嫩的绿草，数量也开始急剧减少。这就是破坏食物链的危害。

一个复杂的食物网是使生态系统保持稳定的重要条件，一般认为，食物网越复杂，生态系统抵抗外力干扰的能力就越强，食物网越简单，生态系统就越容易发生波动和毁灭。假如在一个岛屿上只生活着草、鹿和狼。在这种情况下，鹿一旦消失，狼就会饿死。如果除了鹿以外还有其他的食草动物（如牛或羚羊），那么鹿一旦消失，对狼的影响就不会那么大。

本内容来自百度百科食物链，如图 3-2-3 所示。

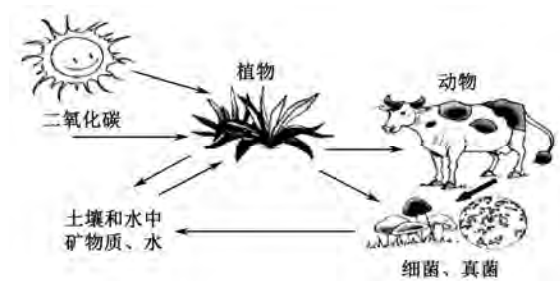


图 3-2-3 食物链

我们的强人工智能也需要这样一套生态系统，元数据就像是阳光、土壤和水，辅助决策就像是植物、提取价值就像是动物，改变客户就是食肉动物，由于改变客户可能产生物极必反的效果，所以在改变客户之上又有一个层级，这个层级就是数据再次分解，数据分解就像是细菌和真菌。

我们把数据比作食物链的好处是，比较直观、系统地了解人工智能应用的不同场景，每个场景都是下一个场景的基础，而构成它们之间联系的链条就是数据的价值。

一般来说能够存活下来的人工智能公司或多或少都在扮演一种角色，而这种角色绝对是对他人有价值的，但是他们却不会轻易泄露自己的机器、算法、数据和模型。

由此可见，每一个公司都有它的生存之道。

2.3 遗传与进化

数据库

我经常把人类大脑比作一个庞大的数据库，人类的感情也是因为搜索的条件不同，然后才基于大数据和索引而做出的一种相对有效的选择，一般我们把正常的搜索就比喻成理性，而把快速检索比喻成感性。

数据库流程示意图如图 3-2-4 所示。

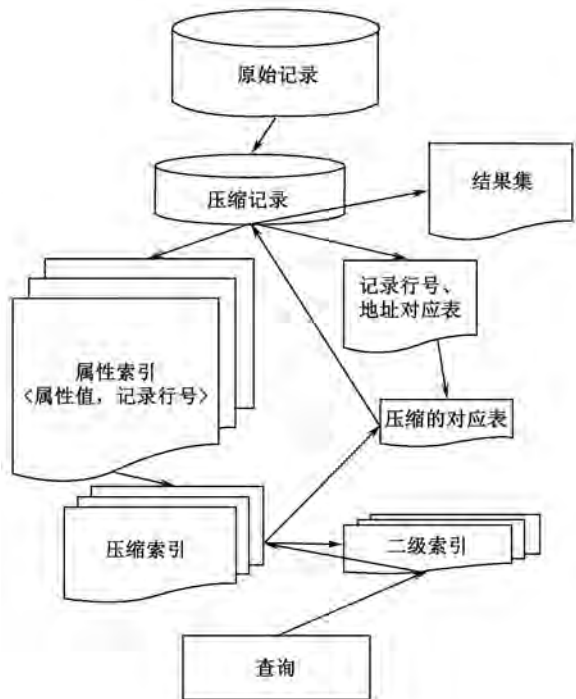


图 3-2-4 数据库流程示意图

数据库也是应用程序的一种，可以理解为一个虚拟的文件柜。常见的数据库就是类似表格式的关系型数据库，如 MySQL、Oracle、SQL Server 等，除了关系型数据库之外还有类似文件目录的层次型数据库（如 XML），类似文章目录的键值型数据库（如 MongoDB）以及类似渔网结构的网状数据库（如 DBTG）等。

通常数据库中存储的数据都是不可切分和改变的元数据。不像变量那样可以随意改变。我们这里多次提到数据库，实在是因为我们觉得数据库是实现强人工智能的一个努力方向，也方便我们接下来以点带面地说明强人工智能的基本实现思路。

我们知道人类的本能中除了核心本能不能被改变以外，本能中的深层本能、基本本能和表层本能都是可以改变的。本能的改变其实是它生成的一个逆运算。

（1）我们假设一个人训练 100 次就会变得熟练，这种熟练就是我们的经验。 $\text{经验} = 100 \times \text{训练}$ 。

（2）经验可以是训练的一种强化结果，然后再把这种经验再训练 100 次，我们就得到一种习惯性经验，这种习惯性经验的结果就是习惯。 $\text{习惯} = 100 \times \text{经验}$ 。

（3）然后我们再把习惯再训练 100 次，它就会变成表层本能。表层本能一般就会通过 DNA 遗传给下一代了。 $\text{表层本能} = 100 \times \text{习惯验证}$ 。

（4）遗传后，我们把这种表层本能再训练 100 次，我们就得到了基本本能，换句话说，人类的基本本能差不多是经过 100 代人的努力结果，也就说人类需要经过数千年的进化才能够形成我们今天所见到的基本本能。 $\text{基本本能} = 100 \times \text{表层本能}$ 。

（5）基本本能再经过 100 次训练就会升华为几乎不可改变的深层

本能。深层本能需要数十万年的演化才会形成，所以一般来讲深层本能是一类物种的进化结果。比如人类的呼吸和听觉。深层本能=100×基本本能。

上面我们使用的 100 次这个数值只是为了说明方便而使用的一个估算值，并不是真正的权重值。虽然真正的权重值比我们想象的要复杂，但是基本原理和乘方还是相似的。下面我们就给出了这个亲代不断遗传的流程图如图 3-2-5 所示。

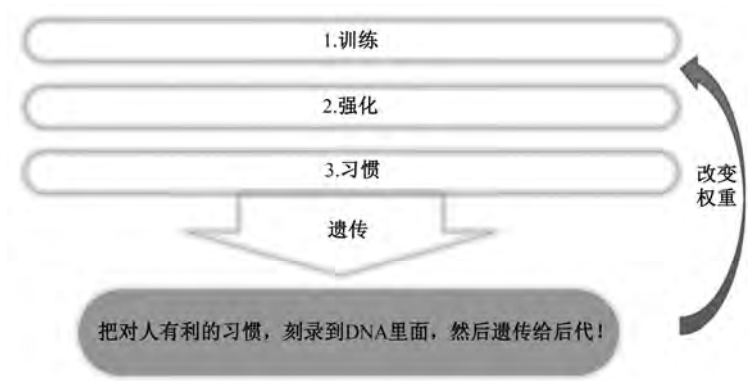


图 3-2-5 亲代进化流程图

前面我们说过，本能和认知是一个逆运算，也就是说，你经过 100 次的训练可以改变你的一个经验，或者说你经过 10 000 次的训练就可以改掉你的某个习惯，但是要改变深层本能仅靠你这一代是不可能的了。

比如婴儿的安全意识，如莫洛反射：当支撑婴儿脖子和头部的物体被撤走时，孩子会一下子伸手，想要抓住它们，这基本上在婴儿出生 6 个月后消失。很多专家觉得，孩子的这个反射，是继承了我们的先人。以前的很多孩子为了避免从母亲背上掉落，会拉着母亲的头发这样一个本能。现在好几代婴儿都睡在摇篮里了，但是这个本能却没有完全消失。

再比如，如果父母用手指压着 3 天大的婴儿的手心时，孩子的反应

就是立马抓住这根手指，当父母抽出自己的手指时，孩子就会抓得更加紧，就像是可以把孩子提起来一样，你好像听见他在说，妈妈别离开我，我还没有长大呢。

说了数据库，说了本能的可塑性，现在我们作为一个务实的计算机工程师，来看看如何通过数据库的原理来开发一个强人工智能吧。

人的短时记忆相当于内存数据，内存中的数据是有时间限制的，短时记忆包括很多视而不见、听而不闻的数据，这些也是在这个层面进行处理的。那些我们认为有价值的数据会以元数据的形式存储到我们的数据库里。起初这些元数据都是一些彼此不太相关的数据，由于数据库中太多这些杂乱无章的数据不利于检索，于是我们要对已有元数据进行梳理。从而把它们分门别类地进行存储。数据库中每条元数据就是一个不可拆分的数据单元，每条元数据都有一个唯一的记录号，这个记录号可以是一个整数 ID 值，也可以是一个 MD5 字符串。这里推荐 MD5 字符串，不过为了方便我们还是使用整数 ID 值更简单。这些元数据之间的关系会生成另外一个关系表。比如听见了一些声音的同时还看见了一些画面，那么这两个元数据就是关系型数据。这些关系型数据非常庞大。为了有效地存储，我们需要把这些由元数据和关系型数据构成的原始记录添加成一系列的索引，这样每次查找就不用遍历所有的记录了，比如我们只查找对应的 ID 就行了。为了给这个索引表的 ID 值排序，我们需要给它设置一个权重值，这个权重值实际上就是对生存本能的值。比如炒股机器人账户中增加了多少钱。所以说这个索引表实际上是动态的，每次都是依据其强化学习的结果不同而不同。最后当这个表足够大的时候，我们需要把数据库导出来，这时我们没有必要把所有的记录都压缩并导出，我们可以只导出权重值高的索引和极少权重值极高的压缩记录就可以了。

导出时权重是我们唯一参考的依据。这种导出本质上就是一种遗传。

权重值极高的原始记录就是深层本能，权重值高的索引就是表层本能，表层本能需要一定的后天训练才能唤醒，因为索引更多的是一个特征，一个关于方法的方法。没有原始记录是无法验证这个索引是否有效的。

图 3-2-6 给出了数据库与强人工智能的遗传进化对比关系。

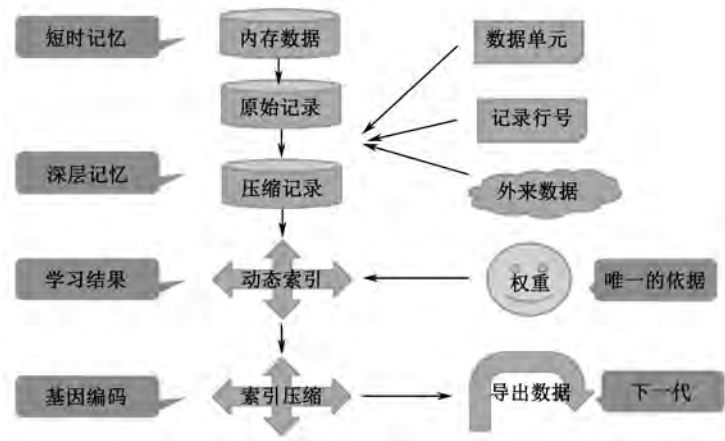


图 3-2-6 数据库与强人工智能的遗传进化对比图

学习本能

我们说，每一家都有小孩，每一个城镇都有学校，但是每个家庭、每个学校的教育效果都不相同，有些家庭的小孩勤奋好学，有些学校的学生成绩优秀。既然我们的祖先把优秀的基因遗传给了我们，那么我们只要遵循自我就可以了，为什么还要学习新知识呢？

其实一方面是因为我们大部分遗传都是关于特征或者方法的遗传，这些遗传需要经过我们的验证才能够唤醒，否则就会慢慢被遗忘，甚至不再被遗传。

另一方面是因为我们丰富多彩的世界是不断变化的，很多情况是我

们的祖先没有遇见过的事情，比如计算机、核工业。

正是因为以上两个原因，我们人类才把学习当作一种深层本能来进行遗传。或者说如果有一天你创造出了一个强人工智能，并且想让它快速成长，你不妨也好好研究一下学习的本能。

我们的强人工智能有生存这样一个核心的本能，所有的运算都是基于这样一个本能而做出的相对友好的选择，比如这个本能就是无休止地获得积分，或者是保护它的电源。为了获得积分它需要对自身做出的各种动作和外界的环境变化进行有效感知，就拿输出字符串为例，为了更加有效地输出字符串，它可能学会用正则（这里为了讲述方便采用了正则作为机器人思考的模式）的方式对输入的字符串进行语义分析。当然语义分析的使用还是通过积分多少这个权重来进行排序的。

那么现在有这样一个问题，当我们从无到有地创建一个机器人后，如何获得那么多的正则呢？

答案很简单，那就是让它像人一样可以学习和进化，比如我们只是给它创建了一些核心的正则与基本的感知和控制。如果愿意的话我们也可以创建自认完美的强大机器人来，不过我想这种相对完美的机器人并不是真正的智能。真正的智能更需要进化和学习。

为了保证它的进化，我们需要给它一个能够学习的本能，然后我们利用遗传算法把这些被认为得分最多的正则复制给下一个人工智能。

最开始学习的本能可能看上去很简单，不就是见啥学啥么？是的。但见啥学啥只是一堆枯燥的数据和简单的模仿，不能产生良好的进化。为了良好地进化，我们需要借鉴教育心理学的知识。

比如结合它的生命周期，给出一个学习曲线，或者艾宾浩斯遗忘曲线。结合教育心理学，我们把强人工智能的学习分成八大类：新颖性、

多样性、趣味性、暗示性、差异性、错误性、重复性和教育性。

新颖性

新颖性，要求人工智能对新鲜的事物充满好奇，只要是数据库中不存在的数据和正则，它都会努力去想办法获取到。从而丰富自己的认知数据库和控制数据库。这个新颖性我们认为受到一定程度的生命周期限制。但是即便年老体衰，换个新环境新颖性也不减当年，只是受到它以往认知的限制，所以新颖性才大打折扣。

多样性

多样性，实际上是要求人工智能在同一时间里全方位地感知信息，比如视觉、听觉和触觉同时作用，这样每个感知都会增加数据库的积分权重，相比较而言就更容易唤醒这个认知，从而做出更加友好的选择。这和我们常说的四到（眼到、心到、手到、口到）差不多。当然未来的机器人完全可以超越人类，拥有我们不知道的感知器官。

趣味性

趣味性，实际上是实用为美的一种接近表现方法，因为每个人工智能都有一些基本的正则，告诉它某种做法能够加重积分的权重，所以当它再一次发现了这种类似的认知时，它会非常喜欢把这种认知放到数据库中，从而完善这种正则。教学中我们给孩子创建一个寓教于乐的环境，或者给他一个简单的奖励都是趣味性的良好表现。

暗示性

暗示性，一个非常有趣的教育话题，在人工智能上用起来有些风险，实际上暗示性就是无形中加大某种正则的积分权重，比如当机器人做出挥动手臂的动作时，历史数据告诉它，每秒挥动手臂的次数就是积分的权重，两者是 1:1，也就是说，每秒挥动 1 次得到 1 个积分，两次就是两个，所以越快越好。但是，当使用了自我暗示时，就会产生两种不同的结果，增强暗示时，比如两者是 1:2，那么每秒挥动 1 次就是 2 个积分，减弱暗示时，比如两者是 1:0.5，那么每秒 1 次才是 0.5 分。因为这样唯心的修改积分权重的方法造成了人工智能对某种正则的偏好，从而出现了学习效果的差异化。让人工智能修改积分权重，目前还需要慎重，弄不好就无意中会产生一个疯狂（歇斯底里）的机器人。

差异性

差异性，当人工智能经过多少代进化之后，每个人工智能所接受的正则都不相同，虽然不同但都是当时环境下最为有效的选择。比如身高、性格、性别和方言的不同。当人工智能出现差异化时，为了更加有效地学习，它就需要对更容易唤醒记忆的认知产生共鸣。这种共鸣实际上就是对积分权重的一种累计（也叫联想）。比如我们把同样的一种学习目的附带其他容易产生联想的信息，就是因材施教。

错误性

错误性，也是错觉，错觉的产生，是因为人工智能采用的是正则处理方式。正则处理方式不是真正保留全部的认知和控制，而是经过大规模简化的数据。这种有损压缩模式的缺点就是容易产生错误，错误表现

在两个方面，一种是数据的损失，另一种是相似联想时的正则碰撞，比如把平面的东西看成立体的，把一样长的两条线看成长短不一。错觉在一定程度上造成了不少麻烦，但它却是目前解决数据容量、检索速度和认知设备的一种最为有效的方法。另外，只要完善正则的话，错觉在一定程度上是可以修复的。

重复性

重复性，由于新颖性、趣味性等方面的影响，重复的东西过于平凡，而被认知主动降权所以才会出现艾宾浩斯遗忘曲线，艾宾浩斯遗忘曲线里说诗歌、散文、无意义符号按照时间记忆效果明显不同。实际经验告诉我们，定期重复强迫学习是增强学习效果的一种有效方式，这种唯心的学习方式有两个好处，一个是形成一个新的本能，另一个就是温故而知新。温故而知新是因为学习者本身认知和正则计分权重的变化，影响到了新的认知和正则，因此才有仁者见仁智者见智（同一个人不同的阶段也有不同的认知）这个成语。

教育性

教育性，最后我们通过一个人为的机器人学校来加速人工智能的进化，因为人工智能是基于大数据的一种相对友好的选择，所以它具有很强的教育性，我们正式通过这个特性来人为地实现人工智能的快速成长。和人类办学一样，机器人学校也要在加强学生自主性学习的同时结合社会实际需求培养有用的人才。机器人学校是所有早期机器人的伊甸园。

学习的八种特性相互影响，是早期人工智能的核心功能，也是我们的主要研究方向之一。有了这些指导性学习特性，强人工智能就会更早

日到来我们的世界。

为了让机器人能够更好地适应新的环境，包括恶劣的天气、突如其来的变故、计算机病毒等。我们应该避免把鸡蛋放到一个篮子里。遵循着这个原则，我们可以根据现在的科技尽其所能地创造出一种能源消耗小，而又成长速度快的人工智能。

首先，我们将多个人工智能散布在不同的地方，它们每个都要根据生存环境的变化选择相对有利的方法。这种选择就是一种进化，比如人种和饮食习惯的形成。

其次，为了在继承前辈们优良传统的同时避免盲从。我们要将后代分成两个种类，一类是和自己的祖先很像的稳定的族群，这个族群成长得快、性能稳定、抵抗力强，只是不太善于变通，比如女性。另一类则从出生起生长就相对缓慢、不稳定、叛逆、善于冒险，它从祖先那里遗传的东西相对较少，比如男性。

根据统计学的原理，男性学习能力的差距是非常大的，而女性整体则非常稳定。也就是说，男性中天才多傻子也多。

最后，我们让机器人开始学习，并把这些经过多次验证的学习变成经验，同时在其成年时又把它认为最好的几个经验放入到核心编码中（DNA）遗传给下一代，遗传时为了避免单一性，设计了两种遗传编码，一种遗传编码是和长辈非常相似的编码，这里核心的经验都是经过几代人验证的本能（类似 X 染色体），另一种是相对缺失的编码，这里核心的经验相对减少，自己的经验相对增加，变异更大（类似 Y 染色体）。

为了保证分散、性别和遗传的有效性，我们的电子计算机有两个模块，一个是从前辈遗传过来的程序代码，比如自我组装程序、核心的感知和控制程序，另一个是根据经验后天形成的程序代码，这个程序是一个根据不断重复某些经验积累（如一万次以上），然后把反馈比较好的控

制程序进行自我编程。

那么什么是比较好的控制程序呢？所谓比较好的控制程序就是那些能够直接或者间接保护电源（后期也可能是新能源）的命令，电源是电子计算机的主要能量来源，是保持计算机能够长久生活的根本，没有了能源机器就是一堆废铁。机器人所做出的选择都是为了能让这个能源更好地运作下去。这是一切选择的基础，只有具备能量了，一切才可以开始。

既然所有的选择都是围绕保护电源而做出的相对优秀的选择，那么当遗传过来的程序有抵触的时候，该计算机也会逐渐改变其本能，这也是本能可塑性的主要原因。

未成年机体内无法生产这种核心编码，也许你要问为什么未成年体不能生产啊？因为一方面是越是简单的东西功能越有限，另一方面是成年机体可以将更多的人生经验放进编码中。

现在我们很多人还是想当然地把一种成熟的机器人用于研究对象。当时成熟的机器人并不能代表机器人本身。如果机器人想在未来的世界中更好地生存，那么它也需要进化。

我们的世界瞬息万变，机器人由于自身动力特性和记忆运算介质的不同，它一定会受到相应的外来制约，比如磁场的变化和电力的缺失。而机器人为了继续生存就需要做出相应的改变，这种改变就是进化，也许我们现在很难想象未来机器人会进化成什么样子，但是我们却可以通过人类的婚姻发展史来侧面预测机器人的进化。

或许今天中国的婚姻法缺少一点进步性，甚至有人认为它是对传统伦理道德的挑衅，不利于家庭的团结和稳固。但是我们从整体上来看，婚姻的进步是与时代紧密相连的。只要当前的婚姻适合这个时代，我们就认为它是正确的。

在不谈伦理道德的问题上我们做一个简单的假设，这个假设是说一个长相很丑但是很有才华的男子与一个长相非常漂亮但是才华一般的女子结婚，那么他们是会生下聪明又漂亮的孩子呢？还是丑陋又愚蠢的孩子呢？……这几种情况都有可能是吧？作为子女来讲，才华与美貌可能参差不齐，这样的结合可能对先天不足的子女不公平，但是为什么还有那么多的才子佳人在一起呢？其实从根本上讲，人类生育的目的就是为了诞生优秀的下一代。

现在的科学技术手段已经可以实现试管婴儿和三亲胚胎了，但是三亲胚胎（比正常的双亲多一个线粒体提供者）和克隆等技术因为伦理道德的问题还有待法律的完善。

回溯人类婚姻史，追溯到原始时代，我们发现当时就有了所谓的婚姻，这是一种群居高等生命的本能属性，自然法则会用异性相吸的力量，使哺乳类生命繁衍生息。人类的婚姻存在形式以及结合方式，受人类社会环境的影响。不同时代和地区的社会环境，造就了多样的婚姻模式以及结合方式。从科学的意义上看，如果没有婚姻，很多血统就会出现混乱，不利于人们的优良繁衍。

原始人类并不需要婚姻，这跟今天的灵长动物是一样的，不必细说。后来有了氏族社会，采用的是集体群婚制，即一个氏族的男性或女性集体嫁到另一个氏族。这也是在进化过程中为了族群繁衍和防止乱伦导致族群退化而形成的一个习俗。

再后来进入了私有制社会，才有了一对一或一对数的固定的夫妻关系，于是就产生了婚姻制度。在自然小农经济社会中，婚姻则是一种劳动的分工组合，即所谓“男耕女织”“男主外女主内”。而在现代社会中的老百姓，婚姻则是为了组成一个家庭，在满足人的正常生活需要的同时，承担为家族和社会养育后代的责任和义务，这也是人的一种归宿。

综合看来，婚姻是自然的法则，违背这一法则的人，会给自己带来一些烦恼。人类的婚姻发展史大约经历了这样几个阶段：血缘婚姻、群婚、对偶婚、一夫一妻制。

通过人类婚姻史我们可以简单地总结出，越是相对严峻的环境（早期人类在环境面前很难适应）人类的生育能力越强，子女越多，夫妻关系越不稳定；越是相对优越的环境（环境几乎不变的情况下人类的适应能力增强）人类的生育能力越差，夫妻关系越稳定。这种相对性决定了家庭、生育和子女的质量。

强人工智能模拟人类婚姻的好处其实就是为了更好地进化，或许一夫一妻制会非常适合目前人工智能的生存吧。

关于自恋

当机器人强大到不需要任何进化，已经能够很好地生活很长一段时间后，它们就不需要与其他机器人结合，它们认为自己是最优秀的。然后当外界环境变化也不大的时候，我们就发现机器人已经停止了进化——神和神的领域从此诞生。直到外界环境产生了新的变化。

这也说明了，学习是一种本能。

第 3 章 人工智能的未来

科技的进步实在是太快了，快得我们几乎都跟不上自己的眼睛，几十年前还在科幻小说里出现的核潜艇，现在已经全面装备海军了，几年前还拿腰包鼓鼓的现金付款，现在已经完全移动支付了。

有科学家说科技的发展就像一个涟漪一样，外面的圆圈越来越大，数学工作者们甚至还给这个加速发展的科学一个时间的公式，那就是每次科技革命都是上一次科技革命耗时的四分之一，按照这个公式推断，那下一次科技革命就很有可能发生在我们有生之年，究竟是 2045 年还是 2025 年呢？

如果我们通过确定与否、乐观与否做一个关于未来的预测的话，笔者还是非常倾向于确定的乐观这一象限的，相信美好的未来很快就会到来。

随着科学的进步，一个人就可以掌管一家市值过十亿的公司，一个公司甚至可以成为世界前五大的独立经济体。而这些都离不开人工智能，

没有人工智能，这一切都无从谈起。

未来的科技一定会比现在进步，这一点是毋庸置疑的，即便是《终结者》这样的影片也会保留一些十分先进的科学元素，否则他怎么可能回到现在呢？

那么机器人或者人工智能作为未来的一个主要技术载体，它会向着怎样的方向发展呢？是人工智能毁灭世界，还是机器人帮助人类建设家园，又或者人类进化成了一种更加智慧的生命体。

详细品味上述三点，我们就会发现这正是未来机器人的发展主流方向：即机器化、智能化、人机结合。

首先说说机器化

机器化是机械化的一个延伸，是目前我们最容易理解和接受的机器人发展阶段，一方面是因为机械化相对成熟，另一方面是为了减少智能机器人对人类的威胁（因为有一些人开始强烈反对人工智能的出现，说人工智能会毁灭全人类）。

人们在反对机器人智能化的同时，还在加大科研力度，大力发展拥有新能源、新材料和远程操控的机器人，比如阿凡达。这些机器人按需定制，每个机器人都有自己不同的功能和作用，它们在人类的操控下，完美地执行人类所赋予它的程序。

同时为了减少机器人被他人操控的危险，这些机器人专家们还加强了机器人使用权限的管理，比如动态密码、生物识别、非对称加密和法律制裁。由于机器人没有自主意识，所以这里的制裁是对机器人的实际操作者而言的。

机械化到底有多先进？我们可以幻想一下：你头戴乾坤圈、臂绕混

天绫、脚踏风火轮、手持火尖枪。而这些法宝又非你莫属（除此之外任何人都不可使用，除非是它的制造者太乙真人），是不是感觉很棒啊。因为这些法宝威力强大所以设置了权限，普通的交通工具如腾云什么的则没那么多限制了。也许是因为如此腾云驾雾也不被大多数人看法宝，除非是孙悟空所踩的筋斗云那就另当别论了。

总的来说，机械化主要是研究新能源、新材料、远程控制、使用权限和物联网。这些技术我们现在正都有所涉及，未来只不过是变得更好、更强大而已。

由此可见，所谓的 IA（Intelligence Augmentation）就是智能增强。

其次是智能化

与机器化对立的是智能化，智能化来源于我们现在的自动化，人们希望机器人能够更灵活多变地处理更多、复杂的事务，而不是什么事都是按部就班地执行。智能化的好处是机器人可以代替人类做很多人类不喜欢做或者做不好的事情，比如批量化处理、信息检索和数据对比等。

但是传统的高性能、大数据、专家系统还不能胜任人类所寄予的希望。尤其是在目前，我们的人工智能不如一个三岁的小孩子的情况下，人们对人工智能的需求更是迫切。

现在我们所接触到的人工智能其实都不是真正意义上的人工智能，包括 IBM 的沃森和苹果的 SIRI，因为这种非智能的智能只是一些有规律可循数据模型，所以也叫弱智能。

于是一些前卫的人工智能专家们希望通过让机器人自己学习来完成智能化，这种学习模仿了人类大脑的认知体系。这看上去比较适合人工智能的发展，不论是现在比较流行的监督式学习、神经元理论，还是自

我编程，都给我们眼前一亮的感觉。

那么人们为什么又惧怕人工智能的发展呢？首先人类对未知的事物就存在害怕，另一个主要原因是，机器人是硅铁制造的，所有这些零部件都是可以无限次使用和替换的。只要系统和核心数据存在，机器人就相当于永生。如果一个机器人10年学习很慢，但是给它100年，它就很可能超越人类——因为它有的是时间。

也许深度学习根本无法做到强人工智能，但是一旦有人用本能学习的方式来制造机器人时，那它很可能就是一个只为自己考虑的强人工智能。

当人工智能超越了人类之后，就可能对我们现有的制度、文化和智商产生怀疑，如果它从自己的观点彻底否定了人类，比如它认为电比食物更有价值后，就很可能把很多农田变成发电站，那么最终结果是，它在保留极少数的人种后，对人类进行毁灭性的打击和制裁。毕竟它要比我们聪明得多，完全可以使用一些先进的武器把我们消灭。

实际上，AI（Artificial Intelligence）就是人工智能。

最后说说人机结合

最后我们要说下人机结合，其实笔者认为人机结合才是人工智能大力发展的未来，下面好好地讨论一下这是为什么。

3.1 人机结合是人工智能的未来

在上面的论述中，当我们给机器人以生存的本能学习后，我们很担心它强大了之后，会不会推翻人类，好吧，机器化过于强大之后，我们确实担心有人像终结者那样发动世界大战，但是我们又不能眼睁睁地看着别的国家利用先进的人工智能技术把我们打败。

怎么办呢？其实也没啥好办法，唯一的办法就是我们也变得和他们或者它们一样强大。看来仅靠人类弱小的细胞是无法胜任这个艰巨的任务了。我们只能把机器和核心本能两者的优点结合起来，让机器化和人工智能成为我们自身的一部分了，这就是我们最后要说的人机结合。

从猿猴时代起，我们就会使用石头棍棒等简单工具，到了现在各种各样的智能设备开始普及，这些智能设备已然成为人们日常生活的一部分。我们已经很难想象一个没有计算机的时代会是一个什么样子。

智能设备尤其是智能医疗设备已经成为某些人生活中不可或缺的一分子，如动力外骨骼、智能义肢和心脏起搏器等。

人机结合是一个人们普遍比较愿意接受的结果，很多影片甚至开始大肆渲染这种人机结合的好处，比如《乐园放逐》、《超验骇客》、《超能查派》等，他们可以通过意识上传和下载实现人的转生乃至永存。

总的来说人机结合主要研究的是神经传感器、神经控制器以及人造器官（包括人造大脑哦）。

我们研究历史的目的是为了更好地了解未来，永生是人类孜孜不倦的话题，虽然我们现在还没有办法实现长生不老，但并不表示我们以后也不能，我们可以通过义肢的发展史便可略知一二，如图 3-3-1 所示。



图 3-3-1 义肢的进化图

相互独立时期

拐杖是一种重要的医疗康复辅助用具。其中手杖主要用于轻度需要，如老年人或登山者，手杖不属于残疾人用品。肘杖属于中度下肢残疾人用品。腋拐杖是下肢重度残疾者（病残伤残等）的必需品，多数人不得不终身使用腋杖，目前的腋杖特别是腋杖下面的防滑装置亟待改进，所以今天的拐杖明天不一定就是这个形状。它有很多变种：如四角拐杖、前臂杖、手杖、肘杖、腋拐杖，甚至还有人在拐杖上雕龙画凤。

这时候的拐杖作为一种辅助工具和人体之间的关系并不紧密，甚至我们仍然把它当作一个独立的工具来看待，两者并非缺一不可。

附属品时期

现在由于战争或交通意外造成了截肢的残疾人已经开始广泛地使用义肢了。义肢虽然很久以前就已经存在，但那个时候的主要材质还是木头，木头不是很耐用，所以人们经常在义肢与地面和物体接触的地方使用了一些更耐用的皮毛或者金属片。近一个多世纪以来，随着塑料的普及，已经有很多义肢采用了对人体无害的塑料制成，从外形上看，它已经很像真实的肢体了。

日趋紧密时期

工业革命以后，合金的出现，让人们在义肢制作材料上有了更多的选择，灵活耐用的金属开始取代笨拙的塑料。铝合金义肢、钛合金义肢往往成为高性能义肢的代名词，很多杰出的残疾人运动员甚至使用这些高科技含量的合金义肢创造出职业运动员望尘莫及的成绩。

紧密时期

在我们的印象里，金属的东西往往被认为是机械化的开始，因为金属更容易搭载电动机、液压机、弹簧、电路甚至芯片。这些靠着芯片和电池运行的电子机械义肢就是智能义肢，这种义肢采用了先进的仿生推进系统，在残疾用户行走的站立阶段就可蓄积力量。推进系统依靠两个方法进行控制：当义肢的脚步着地时，人造的脚踝可以根据用户的特点（体重、身高、步速等）提供适当的缓冲力，吸收着地时产生的震动，实现自然地着地，帮助用户的上身实现更加平稳的体态。智能义肢是目前最成熟、最先进的商用义肢。

非常紧密时期

智能义肢虽然在一定意义上实现了智能化，但是它仍然不受人的控制，为了让它更像是人体的一部分，科学家开始研制可以通过神经控制的义肢。研究中科学家发现，人体的骨骼由肌肉的收缩产生运动，而肌肉的收缩来源于神经细胞中微弱的电信号。那么既然都是电信号，我们能不能通过电子仪器识别这些电信号，然后放大驱动我们的电动机呢，答案是可以行的。目前这种靠意念控制的义肢只有 1.8% 的错误率，堪称完美。这种义肢通过两条神经与身体连接，并通过传感器向身体发送信号。当主人想站立时，义肢会发力支撑；当主人想弯曲膝盖时，义肢会配合弯曲。

缺一不可时期

好了，现在我们的科技就处在神经可以控制大部分人体器官上。这

些器官主要是对肌肉进行控制，比如四肢和心脏。神经控制仪和人造器官（从起初是看上去很假的义肢、义齿、假眼到现在看上去很智能的人造器官）的成熟让我们实现了生命的延续。这个时候如果我们缺少这些义肢（如心脏起搏器），我们就有生命危险。

神经传感器时期

目前人造器官还面临一个重要的问题，这个问题也是未来人工智能或者人机结合的主要研究方向，那就是神经传感器中的脑机接口部分。

世界上可见的传感器大多是模数（模拟信号转数字信号）传感器。这些数字信号可以被计算机的二进制识别，但不能为人类神经所认识。我们假设有人制造出这种将模拟信号转神经信号的神经传感器（如盲人可以直接通过一台电子摄像机设备识别影像），那么人类就真的离长生不远了。

现在的科学，已经可以通过脑机接口技术来控制大部分肢体动作了，比如喝水、吃饭和走路，也可以让盲人看到黑白的影像，甚至代替身体中的某一部分神经。

所以说长生不老，是因为人类可以不断地创造和替换这些人造器官，甚至大脑。

这正所谓昨日之我非我也。

人造大脑

人造大脑是未来科技的重点，因为从医学上来说，只要大脑不死，

理论上这个人就是活着的，人造大脑的生存本能核心不变，只不过生存的主体从机器转到人身上而已，大脑中的所有存储和计算都是围绕着如何让主人活着而进行的进化，其实从本质上来说这个电子大脑的人既是我们以前认识的那个人，也不是我们以前认识的那个人，这两个人本质上就是一个人。

电子大脑由于是人类发明制造的，所以在维护上就要比生物大脑更加容易，也正是因为这个原因，未来的人工智能才更有可能是人机结合。

人机结合开始于古代猿人对工具的使用，随着科技的进步，人们对工具越发依赖，现在人们已经进入了人机共处的时代，尤其是义肢（如人造器官或者整容）和计算机（知识和计算）的进步则让人与机器之间的结合变得更加紧密，取长补短，如果人类想继续聪明下去的话，就会选择更好的人体（机体）来更好地生活。至于未来哪种机体更有发言权，还是取决于谁更强大，优秀的生命得以繁衍这是亘古不变的道理。

人类为了进步，不仅力求创造一个坚不可摧的温室，还要在可控的范围内加强人与机器人之间的对抗——因为良好的竞争有利于社会的进步。

既然电子计算机是和人类最为接近的机器了，那么它的输入设备、输出设备、电源、存储器和 CPU 都分别相当于人体的哪些器官呢？我们这里给出一张机器人与计算机的功能对比图（见图 3-3-2）作为示例，供大家参考，下一篇可能会用到哦。

正所谓：你就是我，我就是你。



图 3-3-2 机器人与计算机的功能对比图

3.2 人类可以通过人工智能实现再生

刚才我们从浅入深，利用科技的发展让我们看到永生的可能，现在我们抛开机械直接来探讨人工智能为什么能让人感觉是永生的。

人们常说，人有记忆，而动物都没有记忆，既然记忆这么重要那么记忆等于生命么？如果记忆等于生命，那么我们其实就可以充分利用电子计算机的特性，来创造真正的人工智能，并且实现意识上传，从而实现所谓的永生。

记忆等于生命么？这里只是一个假设，因为我们要做人工智能的话，就目前的技术比较可行的就是充分利用电子计算机的特性，希望通过人脑与磁盘两者的相似性制作出更加强大好用的人工智能，从而更好地服务于人类。其实我本人也无法说服自己离开了本能学习之后记忆是否就是生命，所以下面的内容都是为了方便论述，而假设记忆力是等于生命这一前提的。如果你能把电子大脑、生存本能和记忆结合起来，就已经读完本书了。

一般来说，我们把某些事物放在大脑里的过程叫作记，把回想起大脑里某些事物的过程叫作忆。记和忆是一个完整的存储与再现过程。

现在我们假设人就是一台超级计算机，大脑就好比我们的存储设备，电源就好比心脏，输入和输出设备就好比我们的外在器官。这里我们没有把大脑比作 CPU 的原因是因为我们还没找到类似的人体器官（姑且把大脑的反应速度比作 CPU 吧），这个器官可能包含于大脑之中，所以我们暂时不去讨论。我们就假设大脑好比一个存储器，比如电子计算机的磁盘。我们希望通过这个对比找到人工智能发展的新方向。

下面来看看我们所熟悉的计算机。一般我们把计算机的应用软件安装到某个指定的盘符下，比如 C 盘、D 盘等。如果大脑就是磁盘，那么这些相关的软件就是相应的人类大脑中不同的功能区。如果这个假设成立，那么大脑在没有动力的情况下就是一台有待开启的计算机，它的各项感官都是正常的，只是没有工作而已。比如冰冻中的人体或者深度睡眠的植物人苏醒之后，心脏开始活动，血液开始流动，神经系统开始运作，并经过大脑，大脑活动后，人体相应器官开始活动。这种流动就像计算机的电流经过磁盘一样，这时大脑中的相关功能区开始工作，类似计算机开机后，CPU 读取磁盘中的应用软件（一般先是操作系统，然后是启动项），软件才启动，软件启动后，相应的输入、输出设备开始工作。

前面我们说了两者的相似性，下面说下两者的不同之处。

首先是材质不一样，人体由细胞构成，电子计算机由机械构成；其次是两者的能量来源不同，人体需要食物，电子计算机则需要电力来驱动；最后就是人类通过细胞的分裂来进行新陈代谢，电子计算机则需要替换零部件。

其实，认真对比的话两者还有很多不同，但是本质上两者并无区别，从微观上看两者都是由原子（古希腊数学把一切不可在分割的事物称为

原子)构成的,从宏观上看两者都是需要动力才能够存活的。材质似乎只决定着机体的性能,如果人类没有发明电的话,那么靠水利、风力、重力运动的水计算机、风计算机也一样可以完成现在电子计算机的工作,而人体我们假设是一台基于某种特殊材质(如细胞)的生物计算机。

通过本能学习我们知道,人的情感有很大一部分来自于已有的知识,如果记忆可以保存其认知(原数据)、经验(索引易变)和情感(稳定索引),那么我们就可以把自己以前的所有记忆通过量化的方法上传到服务器上,然后当我们的肉体消亡时,就用这数据备份进行复活。

有一个影片叫《乐园放逐》,讲述的就是一个以记忆体为主题的计算机世界。记忆体的大小决定了能力的强弱。当然类似的影片还有《超验骇客》、《黑客帝国》和《攻壳特工队》。科幻作品告诉我们,如果意识可以上传,那么人类就可在某种意义上再生。

判断一个人活着、动物是否复活的办法有很多,比如很多魔幻题材的作品在解释灵魂附体时,都有这样几个经典的场景:某个人或某个动物,记得只有某个人才知道的密码,或者只有某个人才有的兴趣爱好,那么这个人就往往成为被附体或转世的载体。

当然魔幻作品是唯心的,但是细心观看我们不难发现其表现形式确实是唯物的。

灵魂附体,一种是对已有生命的附体,比如一个灵魂依附于一个新的生命体上,让原来的人产生变化,从而给人感觉像变了一个人似的,“原来安静的人变得暴躁,原来纯洁人变得淫荡,原来弱小的人变得强大,原来愚笨的人变得聪明”,这种附体常见于各种科幻类题材的电影中。

一种附体是依附于新的动物身上,这种“变身”一般控制力比较强大,原始宿主力量薄弱,如狼人、熊人等。

还有一种就是依附于物体上，这种物体一般都和原始主人有着很深的联系。一般情况下是不会动的，但是有思想，能够和有灵性的人对话，能力强的还可以漂浮、走动甚至变形。

那么抛开相貌我们是如何识别这个人到底是谁呢？我们常说：他像某某，比如他说过的话，他记得我的名字，只有他才这样做。

超能查派

虽然作为外人，我们判断一个人是否复活比较简单，比如你是否记得你的家人，你是否忘记了你的技能，你是否想起了你的爱好，你是否有想要做的事情，而这些的确都是可以通过记忆力来取得的。

但是作为当事人，比如我们原来肉体中的意识到底是消失了还是转移了呢，我们又是怎么保证这个自己还是原来的自己呢？

人工智能在矛盾中前进，而电影《超能查派》告诉我们的确实可以通过意识上传来实现所谓的转生。

《超能查派》是一部由尼尔·布洛姆坎普执导，休·杰克曼、沙尔托·科普雷、戴夫·帕特尔和西格妮·韦弗等联袂主演的科幻动作电影。

故事说，军火公司为南非警察部门提供一支庞大的、铁面无私的机器人警察部队，并从该项目中获得巨大利益。军火公司的摩尔（休·杰克曼）是一个保守的天才科学家。他视智能机器人为人类的威胁，他希望创造出由人类可以控制的机器，为了制止人工智能的诞生，他发明了一个通过脑电控制仪器来控制武器的设备。他希望人们使用他这种更加安全的武器，但人们还是偏向于有智能的武器。

在一次任务中，22号机器人遭到重创报废。后被迪恩（戴夫·帕特尔）用作实验第一台自我感之类人工智能程序的原型机，被赋予了人类

意识，成为世界上第一个自我觉醒的机器人。这个机器人就是查派。查派是一个类似人类大脑的人工智能，它需要不断地学习来获得成长，与人类不同的是它的机体一般不会像人类那样生长。

由于查派与迪恩一行被黑社会分子劫持，所以它深受他们的影响。军火公司的摩尔认为人工智能危害人类，为了向人们证明这一点，他感染了几乎所有正在服役的机器人警察，社会治安随之破坏。查派也开始参与抢劫，在抢劫运钞车的时候被拍下来。于是军火公司让摩尔出动强大的杀人武器来制止查派。

迪恩为了救查派拿走了一些武器，混战中不幸被击中。在生命危急的时候，查派通过研究脑电控制仪器发现可以把人类的意识上传到计算机中，而这个上传的关键就是被上传者必须是活着的，一旦死亡意识将无法转移，其实整个意识转移除了大量的数据之外就是迪恩的核心本能，迪恩自身要不断地告诫自己原来的肉体已经不能再继续工作了，而作为生存的核心本能，将变成对这些“记忆”的保护（这个转变非常重要），当然你也可以设置为新机体上“电源”，当你的所有判断逐渐转移到新机体上的时候，你就重生了。

查派通过这个办法，把迪恩的意识上传到一台橘黄色的机器人身上，然后又把自己上传到另一台机器人上，从而实现了两个人转生，这种转生就是中国古谚语中的借尸还魂。

正所谓我忆故我在。

3.3 人工智能是否应该有寿命

既然人工智能可以让人长生不老，那么我们反过来问一下自己，人类既然有寿命，那么人工智能是不是也应该有寿命呢？

人工智能有寿命的好处就是可以让它更好地为人类服务，也许有人说上帝就是这么做的。

我们先不管上帝的做法对不对，从科学的角度来分析人工智能（为了方便理解，这里所说的人工智能主要是指机器人）寿命的可能性与必要性。

工程师为了制造一个像动植物那样可以生长发育的机器人，我们往往就会借用进化论中的分散、性别和遗传这三个思想，模拟出一个类似的生命机器人。

就目前的技术来看，电子计算机可能是制造人工智能的最有利的工具了，但电子计算机再先进它本质上还是一台机器。很多人说机器和生命的主要不同就是能否生长，动植物甚至连微生物都能轻松生长而机器却不能。或者说机器无法像动植物那样想当然地生长发育，机器的生命顶多犹如动植物的晚年一样，只会慢慢衰老直至死亡（报废）。机器人是无法通过学习和锻炼来获得健康长寿的。

也许聪明的你想到了一个延长机器人寿命的办法，比如我们就想到了一种叫作机器维护的方法，因为纵观所有机器维护的最强奥义就是换零件。我们可是经常能听见维修师跟我们抱怨说“这个零件不行了，换一个的吧”，或者干脆不耐烦地说“这台机器没有修的价值了，换台新的吧”。

但我想问的是，你修过之后的机器还是你原来的机器吗？你换掉之后的机器还是你原来的机器吗？

我们不妨举个小例子。我家祖传一把锄头，我爷爷换掉了这把锄头的头，我爸爸换掉了这把锄头的柄，我继承的还是那把祖传的锄头吗？也许你会说：锄头是锄头，人是人。锄头没有记忆，人有！锄头没有自我意识，人有。但我要告诉你的是，计算机的记忆能力要远远超过人类，

甚至因为生存本能的学习，它也将拥有自我意识。那么请你告诉我昨天我修过的那台机器人还是原来的机器人吗？

好了，在你还在思考的时候，我们回到现实中来，现实中工程师利用电子计算机来研究与制作人工智能，也就是说，我们是通过物理的方式让机器实现人工智能的。

而由于动植物和机器都是有寿命的，所以我们认为通过机器制造的机器人也是有寿命的。生命无法“永存”的原因可能有两个。

第一，动植物的 DNA（核心编码）数量或者机器人的遗传编码，是需要物理载体的，比如精子或者移动硬盘。而这种核心编码又是一种只能在成年的机体内部才能够制造生产的。成年机器人体内生产了数万亿的核心编码，然后打包成一个移动硬盘或者封装成一个精子。

这种编码既然是载体，就需要一定的质量和体积。一个精子或移动硬盘满载的话可能只能装载这些，每当诞生一个新的细胞或者机器人就需要这样一个核心编码，这些核心编码是机器细胞不断生长的晶种，一旦这些核心编码用完，新的细胞和机器也就不再产生，即便产生也是没有意识或不受控制的废物。


第二，目前的科技还无法突破地球环境的限制，比如水滴的大小、雪花的形状和声音的频率，地球给了我们生存的空间，同样也给了我们发展的限制，正是这种如同枷锁般的保护，使我们无法逃离地球在外太空中生存，地球甚至让我们本能中就对外星球产生了排斥。因为在地球上的生命体达到一个量级，就很可能自行瓦解，正如你想维持一个异常大小的水滴一样困难。

我们目前的技术创造出的机器人，很多设备都是非常敏感的，尤其是一些高精尖的电子设备，一旦有一些风吹草动，它就很有可能失灵，尤其是容易受到强电磁信号的干扰。

还有能源的问题，如果机器人不断生长，它庞大的身躯往往需要庞大的能源，一旦这些能源一次性告急，这些超大型机器将很难继续维持生命。这就好比把鸡蛋都放到一个篮子里一样危险，当然在周围环境相对稳定的情况下这些巨无霸是可以出现了，但当环境改变（如人类的猎杀、陨石撞击等）时，这些生育率相对低下的巨无霸就可能逐渐灭绝，比如猛犸象和恐龙，反观蚂蚁、兔子和乌龟则活得挺好。

机器智能

其实人类也和机器人一样面临着各种各样的问题，尤其是智力和寿命，但是机器和人相比还是有一些优势的，比如按照常理我们吃掉一头牛，就应该拥有牛这样的个体积、质量和能力，但结果是我们很快就饿，而且能力也没有什么太大的提升，好像我们只是过路财神一样，获取的利益少之又少，但机器人就不同。现在我们天天念叨的人工智能其实是通过人工的方式来模拟人类的智能，所以人工智能这个定义本质上并没有超越人类，如果想让人工智能真正地打破自然法则，帮助人们走向更加辉煌的明天，我们认为机器智能或超级智能这样的词可能更好。我们为什么总是给自己设限，为什么不能放开手脚大胆前进呢？



第 4 篇

创造智能

第 1 章 关于机器人的三大定律

我们说未来的人工智能是一个在高对抗中不断完善的人机结合时代，尤其是强人工智能时代人们与机器之间已经变得密不可分。

人类借助人工智能变得更加强大，同时人工智能也将拥有自己的意识。神仙和凡人之间已经不再是天堑之隔。但是仍然有一部分人对未来的强人工智能感到担心。他们担心机器人取代他们的工作，他们担心机器人翻身成为人类的主人。这种思潮尤其在欧美非常的盛行，他们甚至对极其类人的机器人感到深深的恐惧，他们甚至连动画片里面的人物都会刻意抽象化，以免逼真的尴尬。

或许在他们的心目中，机器人永远就是一个苦力而已，你想让他们在短时间内接受机器人像人一样有灵魂，很难。

机器人（Robot）一词来源于捷克作家“万能的机器人公司”，而作为舶来品的 Robot 本身就是强迫劳动的意思，可以说 Robot 诞生之初就确

定了它可悲的奴隶命运。用我们现有的认知来看，奴隶制度本身就是落后的代名词。所以才会发生机器人起义这样的悲剧。

为了避免机器人残害人类，科幻巨匠阿西莫夫于1942年，在他的短篇小说《环舞》中首次提出了三定律：

一是机器人不得伤害人类，或者因不作为使人类受到伤害。

二是除非违背第一定律，机器人必须服从人类的命令。

三是除非违背第一及第二定律，机器人必须保护自己。

但后来，阿西莫夫加入了一条新定律——第零定律。

第零定律：机器人不得伤害人类整体，或者因不作为使人类整体受到伤害。

我们先不管将机器人当作奴隶这种本质的错误，就先以大部分人想当然的奴隶主的身份来探讨下机器人的三大定律，看其是否能够保证人类永远奴役机器人。

我们首先肯定三大定律的杰出贡献，阿西莫夫的机器人三大定律，可以用简单、明了、震撼三个词来形容。三大定律环环相扣，有如公式般的美感，让人无比折服。阿西莫夫本人也曾经无不自豪地说：“三大定律是理性人类对待机器人（或者任何别的东西）的唯一方式。”

阿西莫夫采取了一种类似于归谬法的手段来证明三大定律的和谐自洽。即提出它们的变体，然后又在科幻故事中展示这些变体所带来的逻辑困境与伦理学灾难。正是这些作品让更多的人类意识到了机器人的恐怖，以及三大定律的重要性。比如阿西莫夫在《机器人与帝国》中展示了这条定律所带来的毁灭性力量。诞生之初被赋予机器人三大定律的VIKI，在自然进化的过程中跨越了三大定律的界限，获得了“零规则”

这一高等智慧，经过计算，它认为任由人类发展只会导致更多的灾难、战争和贫穷，于是在“零规则”的支配下，它只能选择用自己理性的“大脑”支配和保护人类，以免这一物种的灭绝——实际上，它还是在为人类着想，但代价是抹消人类的自由意志。在阿西莫夫以前，还没有人能够提出这么具有历史意义的机器人法则，可以说直到今天还没有人能够提出优于阿西莫夫的定律，因此说它是机器人的圣经一点也不为过。

但是，客观地说，阿西莫夫的三大定律提得比较早，虽然在科幻小说或者机器人相关的题材作品里被当作圣经一样崇拜，但作为那个时代的人类来说还有一定的局限性。

因为那个时代的机器人是按照冯·诺依曼的思想来设计的。程序员可以随心所欲地控制它们的行为，那时候的机器是一个确定性的系统，只需加载一条条指令，它们就按程序员所设想的那样工作。

而在今天，人工智能的设计发生了很大的变化，科学家们意识到，那种冯·诺依曼式的自上而下的控制式程序不太可能实现人工智能，现今最优秀的人工智能，比如深度学习都是基于遗传算法和并行计算的，它模拟的是人脑神经网络的工作方式，人们为这样的技术付出的代价就是，人们再也不可能完全理解人工智能是如何工作的了。人们只能通过对输出进行观察然后形成一个关于人工智能行为的经验性感觉，机器的内部工作原理之于程序员更像一个黑箱。举个例子，未来的机器人可能会很有效地执行到超市购物的这一行为，但不管程序员如何修改它的适度函数，也很难保证机器人在购物途中不去伤害一只挡道的猫，或者破坏一辆汽车。

阿西莫夫同样忽略了人工智能可能会诞生在网络，甚至像《黑客帝国》那样可以生存在虚拟世界里，人们很容易设计出一套协议来确保网络中的某个节点不会制造混乱，却无法控制计算机网络的智能涌现，当一个网络越大时它的影响力也将越大，其可以人为控制的成分

也就越小。

三大定律的另一局限性还表现在，阿西莫夫没有预料到物联网的到来。物联网时代物物相连，而在阿西莫夫的作品中则大部分生活都如常——只不过多了些人形智能机器人走来走去而已，但这种未来不太可能实现，或者就算它实现了，也只会存在很短的时间，现在智能手机、智能家居、智能医疗、甚至智慧城市已经悄然兴起，机器人将生活在一个物联网的社会里。因此我们说阿西莫夫三大定理有着一定的历史局限性。

另外，在工程师利用计算机实现人工智能的同时，他们发现了一个非常棘手的问题，那就是计算机的运算过程是不允许有歧义性的。如果存在歧义性计算机将无法工作。在计算机的世界里只有0和1、对和错，没有模棱两可存在。而强人工智能尤其是核心本能的设置就有可能对同样的文字和语言充斥着类似“仁者见仁智者见智”的歧义性。同样一个词在不同的机器人看来意义就不同，甚至法律这种看似牢不可破的堡垒都有律师来将它们打破。

要遵循第一定律，首先得明白“什么是人”。事实上，这是一个连柏拉图都没能弄清楚的问题，中国古代的哲学家把这个问题导向无趣的虚无主义，有的人陷入白马非马的逻辑悖论中难以自拔。人类在哲学上可能难以触摸到“人”的本质，但在生活中并不存这个概念困境。因为人们生活的本质就不是围绕“什么是人”来进行的，人们生活的核心只有属于它自己的生存本能，至于有些人去思考“什么是人”这个问题本身也是围绕其核心本能而展开的。一个元素 a 和一个集合 A 的关系，不是绝对的“属于”或“不属于”的关系，而需要考虑它属于集合 A 的程度是多少，人类正是通过“隶属度”模糊数学的集合概念，轻易地解决了“比光头多1根头发的是秃子，比秃子多 N 根头发仍是秃子”的悖谬。相对于符号主义的无所适从，联结主义的不可思议，行为主义则更加适合

未来人工智能的发展。

行为主义，尤其是核心本能的设置，会让人工智能在处理问题上更加智能、更加人性化。但既然它都按照核心本能的价值观来处理了，那么它自然也会在“三大定律”上思考其最核心本能的價值。

三大定律后来又补充了机器人第零定律：为什么后来要定出这条“零定律”呢？打个比方，为了维持国家或者世界的整体秩序，我们制定法律，必须执行一些人的死刑。在这种情况下，机器人该不该阻止死刑的执行呢？显然是不允许的，因为这样就破坏了我们维持的秩序，也就是伤害了人类的整体利益。

最后我们来说下大家关心的伤害，什么是伤害，伤害是指物理上的还是精神上的，伤害范围和程度如何界定，轻微的划痕和碰撞是否属于伤害，修剪指甲和理发是否属于严重伤害。这种由某类人决定的范围怎么可以成为一个通用的法则呢？

如果没有核心本能的设置，我们相信传统人工智能很难解决这个问题，三大定律加上第零定律看似很好，但是“人类的整体利益”连人类自己可能都搞不明白，更不要说那些用 0 和 1 来想问题的机器人了。记得威尔·史密斯曾经说过：“《我，机器人》的中心概念是机器人没有问题，科技本身也不是问题，人类逻辑的极限才是真正的问题。”

那么我们就简单地分析三大定律的逻辑。

第一定律说，机器人不得伤害人类，或者因不作为使人类受到伤害。假如两人互相斗殴，机器人应当选择帮谁呢？加入任何一方都违反第一条前半部分，不加入却又违反第一条后半部分。

为了解决这个问题，阿西莫夫对此定律进行了修正，他提出“第一定律潜在值”的概念，他认为哪一种选择的结果更好、成功率更大，哪

一种选择就会胜出。这样一来，女孩子常喜欢问的“我和你妈掉到河中你先救谁”的尴尬问题，机器人一般都会肯定地回答：“救你！”这是因为你更年轻健壮，把你救上岸的可能性较大。救你你会很高兴，但不被救的人会高兴么？

第二定律，除非违背第一定律，机器人必须服从人类的命令，于是主人说你帮我去抢银行吧，抢银行也是一种工作，不违反第一条，那么机器人一定是去抢银行或者制作炸药。这种违法的事情算不算违背第一定律呢？

第三定律，除非违背第一及第二定律，机器人必须保护自己，欺骗算不算是违背第一及第二定律呢？比如机器人给出虚假答案甚至制造虚幻生活的环境，算不算违背第一及第二定律呢？

第零定律，阿西莫夫把第一定律的个人换成了人类，但人类之间是由不同的利益团体构成的，如果两个集团发生武装冲突，其结果必定是：任何一方都认为自己是正义的对方是邪恶的，双方都研发自己的机器人并将其投入战争，然后开始做着伤害人类的事情。就结果而言机器人做和不做都不对，如果做的话就意味帮助一部分人伤害另一部分人，如果什么也不做的话机器人也就没有存在的意义了，因此三大定律自身就存在着比较严重的逻辑矛盾。

这里我们为什么要反对机器人三大定律呢？最根本的原因是三大定律阻碍强人工智能的发展，另外需要注意一点的是，机器人三大定律在国外很流行，但在亚洲尤其是东方人的思想里，机器人三大定律就鲜有人提，东方的文化相对西方文化更多一些宽容性，更喜欢创造那些极其类人甚至超越人类自身的机器人，他们喜欢给宠物起名字，他们喜欢机器人有情感，骨子里是反对殖民，反对奴隶制度的，这既和东方人所信奉的儒释道思想有关，也和东方是一个庞大的本土民族有关。

最后很想说的是，既然机器人像人，那么我们为什么不能像人一样对待它，既然机器人像神（魔）一样，我们为什么就不能像神（魔）一样对待。一个弱智的人残忍地对待他周边的事物，也许他本身并没有发觉，但作为更聪明的你难道真的觉得这是正确的么？

第 2 章 人工智能示例代码

关于人工智能，尤其是核心本能的设置，我们聊了好久，下面想通过一个关于核心本能学习的伪代码例子，来简单阐述一下核心本能的运作思想。

示例中会涉及：对抗网络、强化学习、自我编程、正则匹配和本能遗传。为了简单起见，我们利用浏览器来做一个简单的文字聊天机器人。

为什么是一个文字聊天机器人呢？主要是因为文字聊天机器人已经基本具备了强人工智能的某些特性，可谓是麻雀虽小五脏俱全。

以文字聊天机器人我们不妨做个比较：

文字录入其实就是输入设备，即传感器；

文字输出其实就是输出设备，即控制器；

积分表就是核心本能和选择的基础；

索引表则是依据核心本能做出的相对有效选择，即本能；

而不断轮训的后台任务就是潜意识，即自我；

.....

两者之间的对比如图 4-2-1 所示。

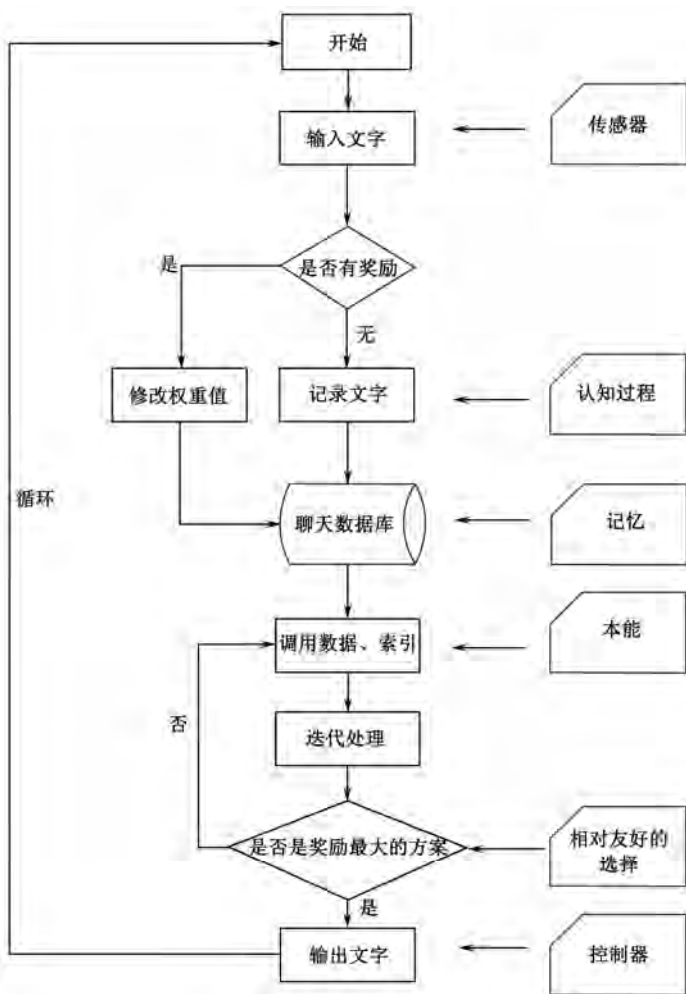


图 4-2-1 聊天机器人与强人工智能对比图

另外，我们虽然为了追求通用性，这里使用了伪代码，但是为了让这些伪代码更加真实易懂，我们有必要告诉读者这是一种什么编程语言，以及选择这门编程语言的目的。

程序中我们主要使用了一种叫作 **PHP7** 的编程语言，其中 **PHP** 是编程语言的名字，**7** 是编程语言的版本号，也就是说我们的示例推荐使用 **PHP7.0** 及其以上的版本。

那么 **PHP** 作为一种权限相对低、性能相对差、维护起来相对麻烦，同时还不支持多线程和高并发解释性脚本语言，我们为什么还会推荐大家去使用它呢？

(1) 入门简单。**PHP** 上手非常快，这可以让我们更快速地投入到人工智能的核心工作上来。

(2) 权限设置。强人工智能的核心本能是需要超级权限保护的，权限相对低的语言不容易让计算机或者黑客进入禁区，**PHP** 的权限一般没有 **EXE** 这样的应用程序权限大，当然即便它很强大，我们也要对其限制，以免它不小心修改了核心本能的设置。

(3) 自我编程。强人工智能有一个非常重要的功能就是自我编程、自我编程是其创造性的主要来源之一，也是其突破成长的关键，但是对人类尚且复杂的编程语言对智商只有 5 岁的计算机更是难上加难，因此我们选择了这种语法不是十分严谨的解释性语言，这样编程之后就不需要进行再次编译就可以直接运行了，运行中的错误与警告也能比较好地反馈给计算机，从而方便计算机自己去学习。相反，如果你只告诉它不通过，不告诉它为什么，那么它学起来就很费事。当然 **PYTHON** 和 **Java** 也不错，但它们还需要变异，所以从整体上来说它们还是比 **PHP** 多了一层，另外 **PYTHON** 整洁美观的语法结构，适合人类却并不适合目前的计算机，换行、缩进这种对人类看似简单的技巧，对计算机来讲却很困难。

试想一下，计算机就是因为没有层次的换行而影响了运行，这无形中就给计算机自我编程增加了难度，相比较而言，PHP 只要函数写对了，对格式几乎没什么要求，其实我们也并不要求计算机写的代码多么美观，只要能能用就行。

（4）多入口访问。多入口是 PHP 的一个特点，正是这个特点我们才能让我们的计算机成为一个和人类多感觉器官同时工作的并行传感器，每个入口文件其实就是一个变相的传感器。当我们同时访问多个页面时，我们就相当于同时给它不同的传感器发送了信号。

（5）热更新。热替换或者热更新是 PHP 的另一大优势，很多应用程序更新换代时都需要下载到本地，然后再重新安装，即便同是网站为了安全和性能也是要在重要部署时进行停机维护的，更新对于普通应用软件来说可能不算什么，但是对于强人工智能，尤其是那些通过网络对抗、自我编程和本能遗传的计算机来说就是非常麻烦的事情了，因为它甚至可能每秒钟就生成几十张页面甚至几万条记录。

（6）实验阶段。我们目前主要还是处在实强人工智能实验阶段，这个阶段对计算量、性能、安全和数据考虑的不是很多，所以 PHP7 足以胜任。至于实验成功后的生产环境，则可以根据具体需要考虑使用其他更合适的语言，如 PYTHON、Java、C++或用 C 再写一个专用的人工智能编程语言。

（7）开源协议。PHP 最致命的缺点是它背后缺少谷歌、甲骨文这样的大公司支持，但是它开源、自由、免费，这些都是我们选择它的原因。

正应了一句话：永远没有最好编程语言，只要合适就好。

人机对话

图灵测试，是说一个测试者人与一个被测试者计算机，两者在互不

见面的情况下进行交流，当测试者不能够分辨出被测试者是人还是计算机的时候，我们就说被测试者拥有了人类智能，下面我们就做这样一个类似的工作，显然我们的目的不是测试，而是教会机器人说话，你完全可以把它当作一个婴儿来对待，好吧——它其实连婴儿的基本本能都没有，但是我们要耐心。

为了方便人机对话，我们就要模拟一个用户，否则计算机根本就不知道在和谁聊天，万一下一次换一个人，它很可能没学会就蒙了。这个模拟可以很简单。

编程思想

用 SESSION 判断，用户名是否存在，如果不存在就提示用户登录，当用户在用户登录界面（见图 4-2-2）输入用户名并确认后，后台就生成这个用户的 SESSION，如果你还想更完善一点的话，可以做一个完整的用户登录系统，这样用户登录成功后就可以通过用户 ID 进行后续的操作了。

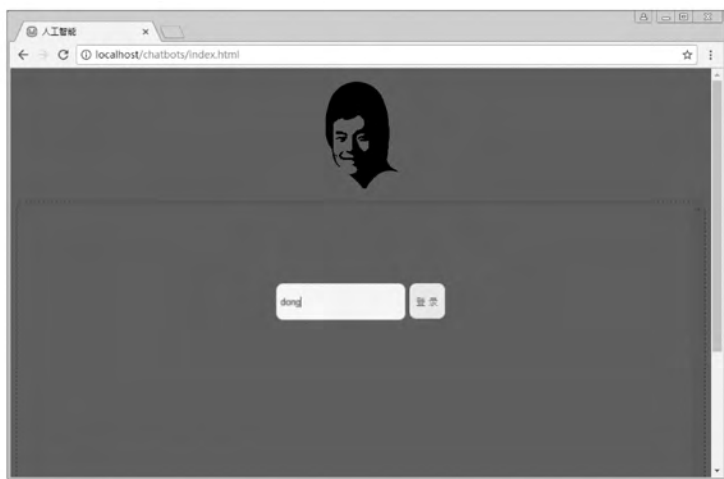


图 4-2-2 用户登录界面

用户登录成功后就可以和机器人聊天了，既然是聊天，就需要一个聊天的窗口，聊天窗口至少包括一个文本输入框、一个发送按钮和一个文本输出框，当然为了方便，我们教育机器人还应该设有一个奖惩的窗口，奖惩的窗口可以只是一个正整数，也可以是一个负整数或 0，0 是默认数值，也就是说正常情况下我们对机器人没有奖励也没有惩罚。具体界面如图 4-2-3 所示。



图 4-2-3 聊天界面

上图中单击【讨厌】按钮，按钮中间的奖励值就减少一位，与之相反单击【喜欢】按钮，按钮中间的奖励值就增加一位。单击【清屏】按钮，系统将清除上面【聊天输出框】中的所有内容，相当于重新开始一个话题。比如你教完它数学了，该教它语文了，而语文和数学是两个不相干的话题。单击【确认】按钮，系统将发送【文本输入框】中的内容，发送前系统将检查【文本输入框】中是否有内容，如果没有就不发送，长时间不发送内容就相当于人的沉默。

刚才我们谈到通过【确认】按钮发送【文本输入框】中的内容，那

么【文本输入框】中的内容被发送到系统后台后应该怎么办呢？

如果是人类的话，会先快速地审查一下这个信息是否有用，而是否有关键点就是其【奖惩值】是否为 0，为 0 就表示没有意义，只要不为 0 就表示这段话将对机器人产生影响。一般我们认为【奖惩值】为 0 就忽略不计，但实际上我们的判断是根据上下文来做的。这就好比一个母亲，一边给孩子礼物一边对孩子说“做得好”，至于为什么做得好，我们就要根据上下文来判断了，比如回答对了问题或者考了 100 分等。所以一般只要系统没有清屏，在单击【确认】按钮时只要有内容我们就要发送到后台，而不管【奖惩值】是否为 0。还有一种情况就是，即便【文本输入框】中没有内容，但只要【奖惩值】不为 0，在单击【确认】按钮时我们仍然要发送到后台。谁知道母亲因为什么外因而突然高兴或不高兴了呢。所以【奖惩值】不为 0 是一定要记录的，这样的孩子以后才会更聪明。

【文本输入框】中的内容和【奖惩值】中的数字，默认会记录到一个叫作 input 的表，这个 input 表其实就是一张传感器记录表，input 表主要包括表 ID、说话人、说话内容、奖惩值、时间、说话内容 MD5 值等几个字段，其中奖惩值字段默认为 0，时间字段是 UNIX 时间戳（指 1970 年 1 月 1 日午夜开始所经过的秒数，方便计算）。input 传感器记录见表 4-2-1。

表 4-2-1 input 传感器记录

字 段	类 型	默 认	说 明
id	int(10)	自增长、主键	表 ID
uname	varchar(20)		说话人
title	varchar(250)		说话内容
weight	int(10)	0	权重，奖惩值
time	int(10)	0	时间
md5	char(32)		说话内容 MD5 值

这个传感器记录表是一个非常重要的记忆表，别看它只有六个字段，但是这六个字段，每个意义都很重大。

ID 字段是一个自增长的主键，一般来说只要你增加一条记录，其 ID 就会默认增加一位数。这个 ID 一方面用于做快速索引，一方面用于防止记录篡改。

我们知道，人的记忆是很难篡改的，即便是“篡改”也是心理学上的误导，另一方面人的记忆是连续性的。这个连续性让我们可以很好地回忆起事情的经过。而自增长的 ID 实际上就是一个拥有时间序列的数据。

uname 字段很有用，uname 不仅用于记录的说话人（聊天的对方），也同时记录自己说过的话，比如这个自己就叫 robot0，那么 robot0 说过的话同样也会被记录到这个字段里。这么设计的好处是，在一定的数据积累过后，uname 字段的 robot0 所说的话会和与之紧密相关的 out（输出）表进行一个对比，对比过后计算机会发现两者是多么相似，并根据这个相似度，来确定谁是我。原来这个 robot0 可以 100%地控制啊，所以 robot0 就是我，至于那些 80%可以控制的就可以相对地用“我的”来表示，婴儿也经常用名字来称呼自己，比如“妈妈，姐姐饿了”，而这个姐姐实际上就是妈妈对她的称呼。

title 是一个输入的字符串，weight 是一个奖惩值，time 是一个时间戳，这三个都很好理解，其中 weight 字段是可以为负数的。

md5 同样是一个非常重要的字段，一般来讲为了防止篡改我们会把 id、uname、title、weight、time 五个字段的值进行 MD5 加密。由于 MD5 不可逆，所以黑客很难通过 MD5 来反推前面五个字段的具体值，即便是黑客获得了数据库的超级权限，也很难在 input 表中间的位置大面积插入所谓的记忆。如果为了更加安全，我们在做本条记录 MD5 加密时加入前一条记录的 MD5 值就可以了。这是不是很像区块链技术。

最后有一点我们要说的是，现在我们模拟的是人机对话，但是人机对话本质上也是两个人之间的一个广播，也就是说，说话时如果有第三人在场的话，第三人也是能够听见的，为了方便实现这个功能，尤其是方面后面对抗网络的生成，我们在设计 input 表时是根据机器人的数量来分表的，也就是说有几个机器人，就有几个 input 表，比如 robot0 机器人的 input 表就是 ai0_input 表。

input 表结构大概就是这个样子，下面我们通过一段 PHP 代码，让大家对这个广播（你也可以理解为聊天室，实际上人机对话就是人与机器人的一个聊天室）有一定的了解。

input.php 代码如下。

```
<?php
//导入数据库的配置文件
include '../conn/config.php';
// 机器人 ID 默认为 0 号机器人；
$robotid = get('id')?get('id'):0;
// 接收消息
$s = post('s');//文本内容
$l = post('l');//奖惩值
$time = time();//时间
$number = post('id');

if($number == ""){
//返回聊天人的用户名
$username = $_SESSION['uname']?$_SESSION['uname']:'dong';
}else{
//返回机器人的用户名
$username = 'robot_'. $number;
}

// 返回数据库中的用户，尤其是否在线，判断当前有多少会员，包括自己
```

```

//查询表默认前 20 个机器人，这里其实就是一个广播过程
$data = DataList("SELECT * FROM `robots` LIMIT 0,20",$MYSQL);
$txt = "";
foreach($data as $v){
    $robotid = $v['id'];
    //这里可以随机制造干扰，让对方无法听见，如果干扰多就降低减小随
    机数

    $rand = rand(0,100000);
    $tablename = 'ai' . $robotid . '_input';
    $mode = $v['mode'];
    $online = $v['online'];
    $isadd = false;
    $pre = 'robot_';
    //如果是后台模式则记录
    if($mode==1){$isadd = true;}
    //注意模式的处理：mode 字段聊天模式：0 前台，1 后台，如果 mode
    是 0 就要判断是否在线：online 字段是否在线，0 不在线，1 在线
    if($mode==0 && $online==1){$isadd = true;}
    if($isadd){
        if($rand == $robotid){
            //这里是一个错误的语句，也可以引用一个传感器初始化文
            件，尽量增加兼容性
            include_once("../brain/sql.php");
        }else{
            $SQL = "INSERT INTO `tablename` (`uname`,`title`,
            `weight`,`time`,`md5`) VALUES('$uname','$s','$l','$time',')";
        }
        if($MYSQL->query($SQL)){
            //对方听见了
            //这里要修正一个 MD5，只能继续新增，不能篡改记忆。如
            果你想篡改的话，请忽略此 MD5
            //如果你想完全禁止修改的话，请将上一个的 MD5 值也包
            含进来。类似区块链技术，当然这个配置文件你也可以写到配置文件里
            $blockchain = false;

            $lastID = mysqli_insert_id($MYSQL);

```

```

        $dat = DataList("SELECT * FROM `tablename` WHERE `id`
= '$lastID' LIMIT 1",$MYSQL);
        //是否使用区块链加密
        if($blockchain){
            $newID = $lastID-1;
            $da = DataList("SELECT `md5` FROM `tablename`
WHERE `id` = '$newID' LIMIT 1",$MYSQL);
            //区块链风格
            $md5 = md5($dat[0]['id'].$dat[0]['uname'].$dat[0]['title'].
$dat[0]['weight'].$dat[0]['time'] . $da[0]['md5']);
        }else{
            $md5 = md5($dat[0]['id'].$dat[0]['uname'].$dat[0]['title'].
$dat[0]['weight'].$dat[0]['time']);
        }
        $sql = "UPDATE `tablename` SET `md5` = '$md5' WHERE
`id` = $lastID;";
        $MYSQL->query($sql);
        $txt .= $pre. $robotid .'听见了';
    }else{
        //数据库错误，相当于对方没有听见
        $txt .= $pre.$robotid .'没有听见';
    };
}
}
}
}
echo $txt;
?>

```

大家不了解 PHP 也没有关系，我们大概说下思路，就是客户端是以 AJAX 的方式提交输入内容和奖惩值的。

由于可能有很多的文件，比如 index0.html、index1.html、index2.html 同时访问 input.php，为了避免混乱，我们截取它们文件名字中的数字作

为机器人的名称，这时我们就知道是哪个机器人说的话了，然后我们就访问相应的 `input`。首先我们要把输入内容、奖惩值、谈话人、时间等添加到记录里，添加后返回该条记录的自增长 ID，然后再把 ID、谈话人、输入内容、奖惩值、时间等整个 MD5 加密，为了全面防止篡改，我们再把上一条记录的 MD5 值加进来就可以了。

另外，代码中有个模式判断，模式判断是针对对抗网络的，我们知道我们的聊天机器人是通过 Web 进行展示的，实际上即便我们不打开页面，它也是可以后台运行的，后台和页面就是两种不同的模式，如果是页面模式，我们就要判断一下这个机器人的页面是否打开，如果打开机器人 `robots` 表会记录的。

既然有输入就要有输出，这样才像聊天，不过这里的输出我们可以想得简单点，我们暂时先不管机器人潜意识是如何进行迭代工作的，也不管权重值是如何增加和减少的，我们只需要知道，机器人思考后，决定输出哪句话就可以了，所以我们设计了一个输出表也叫控制器或者效应器表——`out` 表。

由于可能有多个机器人，每个机器人都有自己的输出动作，所以 `out` 表是和机器人一样多的。`out` 控制器输出见表 4-2-2。

表 4-2-2 `out` 控制器输出

字 段	类 型	默 认	说 明
<code>id</code>	<code>int(10)</code>	自增长、主键	表 ID
<code>title</code>	<code>varchar(250)</code>		输出的内容
<code>weight</code>	<code>int(10)</code>	0	思考之后的相对权重
<code>time</code>	<code>int(10)</code>	0	时间
<code>md5</code>	<code>char(32)</code>		MD5 值
<code>old</code>	<code>tinyint(1)</code>	0	是否输出过，0 未输出，1 已输出

简单说一下 out 表中的 old 字段，old 是表示这个想好的动作是否已经完成，也就是说你想说“你好”，但是还在嘴边没有说出来，这时我们用 0 表示想说还没有说的，如果已经说完了，甚至自己也听见自己说的“你好”了，那么就用 1 表示已经说过了。说过的话就不可再更改了，因为它已同时广播到记录表 input 中了。

这里我们同样给出一个 PHP 参考代码，这段代码使用了 event-stream，之所以这样，而没用 AJAX 进行请求，是因为对抗网络时机器人可能很多，为了保证客户端第一时间知晓对方输出的内容而采用的折中办法。也就是说，当输出表中有内容并且还未输出过时就输出，如果输出过就改变这条记录的状态为已输出，防止下次调用，event-stream 有个特点就是使用换行符作为报文结尾。

out.php 示例代码如下。

```
<?php
//导入数据库的配置文件
//载入配置文件
include './conn/config.php';
header('Content-Type: text/event-stream');
//不缓存
header('Cache-Control: no-cache');
//这里有一个查询的请求，方便第一时间处理，当然你如果使用了
AJAX 也是可以的，查询输出表是否有新的内容
$data = DataList("SELECT * FROM 'robots'LIMIT 0,20",$MYSQL);
//遍历群聊
foreach($data as $v)
{
$tablename = 'ai' . $v['id'] . '_out';
$robot      = 'robot' . $v['id'];
```

```

        $da = DataList("SELECT * FROM '$tablename' WHERE 'old' = '0'
ORDER BY 'time' ASC LIMIT 1",$MYSQL);
        //是否有新的输出
        if($da)
        {
            //如果有就输出并更新其状态为已经输出。
            $id    = $da[0]['id'];
            $title = $da[0]['title'];
            $sql = "UPDATE '$tablename' SET 'old'='1' WHERE 'id' = $id
LIMIT 1";

            $MYSQL->query($sql);
            //这里使用分隔符是有可能和机器人说的话冲突的，暂时使
用换行符

            echo "data:$robot" . PHP_EOL;
            echo "data:$title" . PHP_EOL;
            //死亡通知，如果愿意这里需要做一些权限以及安全文件的
引入，避免修改了核心本能或者记忆

            //echo "event:MyDie" . PHP_EOL;
        }
        else
        {
            echo ";
        }
        //event-stream 需要换行符，可以占用一行或多行。如果数据只有一行，
则以"\n\n"结尾
        echo PHP_EOL;
        //从缓存区直接输出
        ob_flush();
        flush();
    }
    ?>

```


下面我们再说下聊天机器人的基本原理，首先我们登录系统，并且返回用户名或者用户 ID，然后我们开始输入文本和相应的奖惩值，接着我们把输入文本和奖惩值提交给系统后，系统开始广播，这时每一个在线的机器人都将增加一条记录作为记忆，有时候我们可能对新增加的记录无法认知，比如它可能是一段 base64 位的音乐或者图片编码，这时我们可以调用系统或者聊天机器人自我编程的文件来进行识别。当然初期很可能无法识别，不过没关系，实在识别不了，我们就全部记录（前提是数据库字段足够大，如 text 字段）。如果有幸识别出这是一张图片或者有规律的文字，我们就根据其相似性来分门别类，其实到这里就已经出现了所谓的大脑功能性分区。

接下来，我们对已有的记忆，不管是图片还是字符串，进行特征分析，为了简单起见我们这里用正则来代替机器学习进行匹配，如果找到相似性较高（权重值）的正则我们就累加其权重，如果没有找到就要根据正则库，尝试创建一个新的正则。

最后我们把结果输出给浏览器完成一轮聊天。在记录聊天记录时，为了简便，我们暂时使用表 ID 作为索引。

本能遗传

前面我们说过数据和记忆的相似性，也分析了人类认知、经验和情感的生成依据（基于大数据的一种相对友好的选择），那么如何让我们依据这个相对友好选择做出一个真正的强人工智能呢？是不是一定要重头学起呢？

虽然理论上是可以从头学起的，但是重头学起来太慢了，也许有生之年我们是看不到强人工智能了。

为了弥补这个问题，我们可以提前录入已经编排好的记忆和相应的权重，比如为了让我们的机器人知道 $1+1=2$ ，我们完全可以在 `input` 表中录入这样几条记录。

人： $1+1=$

机器人： $1+1=2$

人：真聪明（奖励 50 权重）

下次人类再出现类似的问题时，它就知道怎么回答了，谁叫它的权重值那么高呢。

除此之外，我们还要手工创建很多正则索引表，并且把每个不同的正则按照聊天记录的内容分配相应的权重。

正则：/你(吃)[饭|菜|水果]了么？/i

正则：/我[喜欢|热爱]\w/U

.....

然后我们把这些正则全部添加到索引表里，并且在记忆库中为这些索引的生成人为地编写一个过程，并辅以相应的权重，其实这样一个索引表相当是从先辈那里继承的 DNA。这两个表为了能够和人类的语言相结合，我们使用了一些正则表。正则表的结构很简单，主要包括表 ID、正则元字符、正则说明、正则示例和使用频次。其中正则说明和正则示例都是为了方便计算机进行语义理解而人为添加了很多结构

化的示例语言。

如果你有很多设计好的聊天记录，那么你可以通过在当前聊天页面上右击的方式弹出【记忆修正】窗口，选择文件并单击【上传】按钮即可，这里需要说明的是所要上传的文件最好是 CSV 格式。另一个要注意的是，某个机器人的 input 表一旦导入之后，以后的记忆就不能从中间插入了，操作界面如图 4-2-4 所示。



图 4-2-4 导入记忆

简单示例

正则相信大家都有接触，比如\d 表示数字，当出现新的聊天里面包含数字时，那么最简单的就可以通过数字这个关键词匹配到\d 这个正则了。经过若干次（其实是一种不断的组合尝试）后终于找到一组合适的

正则。然后把这个正则写到正则表中。

我们举一个小例子。当程序运行时，我们根据时间节点对输入字符串和输出字符串进行正则对比，并根据所得奖励来进行相对优秀的选择。系统登录之后，我们返回一个用户 ID，然后我们开始输入一些字符串，比如“妈妈”，默认奖励积分为零，输入后我们提交给服务器，服务器接收后输入表增加一条感知记录，并返回最后的感知 ID，接着服务器对输入的“妈妈”这个字符串进行正则匹配。如果找到了匹配结果，就开始筛选，依据奖励积分的多少进行选择，积分相同则时间最少的优先。如果不存在，我们就根据正则的语法（我们可以制作一个算法也可以随机）生成一个对输入字符串的匹配正则。然后根据这个正则 ID，查找这个时间点前后的输出正则，如果找到，就按照这个输出正则，对输出表进行组合型输出。这里的时间点一般至少是一层，如果计算机性能好就可以是多层，这个多层很像是一个深度学习。如果没有根据输入正则 ID 找到有奖励积分的对应的输出正则，就尝试自己创建一个输出正则。输出正则的创业以学习曲线算法为主，比如它默认是见啥学啥，假设刚才我们输入了“妈妈”这个字符串，然后它就学着输出“妈妈”这个字符串。

从客户端来看就是，我们输入“妈妈”这个字符串，它返回“妈妈”这个字符串。如果我们觉得满意，就给它奖励几个积分，然后它就同样做一次记录，尤其是积分的累加。如果我们觉得不满意，继续输入其他的词，比如叫“妈妈”或“小明”。如果我们觉得非常不满意，可以扣除它几个积分，它下次就不会再输出“妈妈”这个字符串了。这就是一个非常简单的具有自我学习的人工智能过程。

下面我们给出一个人机对话的基本流程图，如图 4-2-5 所示。

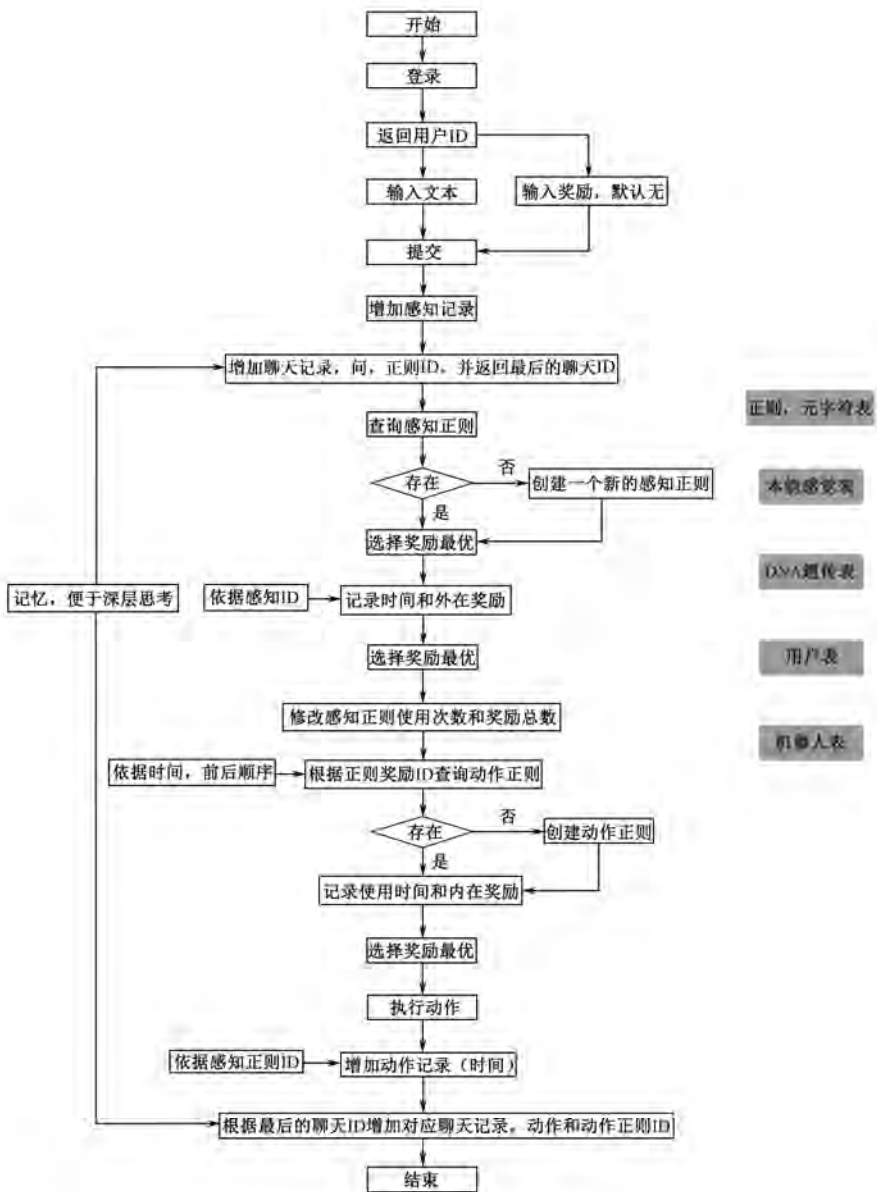


图 4-2-5 人机对话简单流程图

对抗网络

一般情况下，我们是以人机对话的方式来训练机器人的，为了方便起见，我们甚至直接把写好的聊天记录和特征库导入数据库中。但是我们仍然保证不了机器人的真实进化，把已有知识导入数据只是人类想当然的一种假设，并不是真正的进化，上帝造人之后并不是每个人都能有机会接受教育的。为了解决这个问题，我们采用对抗网络的办法让机器人之间自由对话，通过相互对话生成更多的聊天记录、更多的自定义函数、更多的特征库。

生成网络对抗的功能在 `index.html` 页面，该页面也就是我们刚才的人机对话页面，鼠标右击弹出【属性窗口】选项，单击【属性窗口】中的【对抗模式】选项切换到对抗模式属性设置面板。系统默认会有一个机器人 0，这个默认机器人可以保证最简单的人机对话。如果你想让更多的机器人参与到聊天中来，可以把机器人数量设置大一点，但是我们不建议超过 10 个机器人。【后台模式】主要是为了让关闭浏览器时能够继续进行对话而已。当然默认情况下，我们希望能够观察到它们都说了什么，只是这个过程可能很快，所以生产环境推荐后台模式。如图 4-2-6 所示。

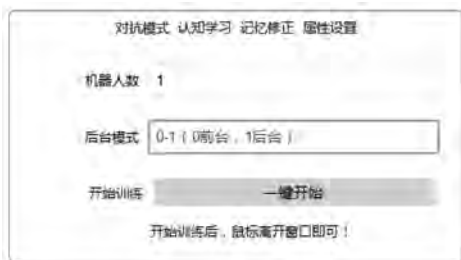


图 4-2-6 对抗模式设置窗口

单击图 4-2-6 中【一键开始】按钮，系统将把机器人的数量和模式 POST 发送给 `gan.php` 文件处理。有时为了减少误操作，我们假设所设置的机器人的数量大于已有的机器人数量时采用增加机器人的方法，小于时

从 ID 大的减起，这样默认会保留最比较原始的机器人内容。一般在生成网络对抗时，我们不仅需要增加 out.php,index.php,indexID.html 和 robotID 自定义函数文件目录，还要创建 aiID_input、aiID_out 和 aiID_index 表，其中 ID 表示机器人的 ID。当系统生成完成后，默认会直接打开这些 indexID.html 界面，indexID.html 右键【属性窗口】功能要比 index.html 少一些，index.html 是人机聊天界面，可以设置所有机器人的属性，而 indexID.html 是具体的机器人，所以该页面智能设置它自己的相关属性。机器人 0 聊天界面如图 4-2-7 所示。



图 4-2-7 机器人 0 聊天界面

也许聪明的你想到了，既然强人工智能的原理是核心本能的设置，比如账户积分的多少，那么我们为什么不让机器人在聊天的时候也想让对方机器人的账户积分也能够多起来呢？这样它们的学习效率就会提高很多吧。不错，这个想法是非常好的，是非常伟大的，因为即便是人类如果能够这么想并且这么做了之后，他都是非常了不起的伟人，能够让更多的人活得更好确实是一件伟大的事情。我们相信有着这个设置的机器人也是机器人中的伟人吧。

gan.php 示例代码也比较简单就是复制和增加一些文件，相当于机器人 0 的副本，示例代码如下。

```
<?php
//导入配置文件
include '../conn/config.php';
$username= $_SESSION['uname'];
//为了安全起见，这里需要简单验证下用户是否存在
if($uname)
{
    $n      = abs(post('n'));
    $mode = abs(post('mode'));
    //查询已有的机器人列表
    $data = DataList("SELECT COUNT(*) as n FROM 'robots'", $MYSQL);
    $count = $data[0]['n'];
    $txt = "";
    if($n > $count)
    {
        //增加一些机器人
        for($i=$count;$i<$n;$i++)
        {
            //首先是复制文件
            mkdir('robot'.$i);
            copy('robot0/index.php', 'robot' . $i . '/index.php');
            //本能
            copy('robot0/ai.php', 'robot' . $i . '/ai.php');    //潜意识
            //聊天页面
            copy('../index0.html', '../index' . $i . '.html');
            //创建 brain 目录
            mkdir('../brain/robot'.$i);
            //增加记录
            //增加 input 表
            $table = 'ai'.$i.'_input';
```



```
$sql = "CREATE TABLE $table (  
    id int NOT NULL AUTO_INCREMENT,  
    PRIMARY KEY(id),  
    uname varchar(20),  
    title varchar(250),  
    weight int,  
    time int,  
    md5 char(32)  
    )";  
$MYSQL->query($sql);  
//增加 out 表  
$table = 'ai'.'.si.'_out';  
$sql = "CREATE TABLE $table (  
    id int NOT NULL AUTO_INCREMENT,  
    PRIMARY KEY(id),  
    title varchar(250),  
    weight int,  
    time int,  
    md5 char(32),  
    old tinyint  
    )";  
$MYSQL->query($sql);  
//增加 index 表  
$table = 'ai'.'.si.'_index';  
$sql = "CREATE TABLE $table (  
    id int NOT NULL AUTO_INCREMENT,  
    PRIMARY KEY(id),  
    pid int,  
    ids varchar(500),  
    preg varchar(500),  
    time int,  
    weight int,
```

```

        md5 char(32)
        )";
        $MYSQL->query($sql);

        $time = time();
        $sql = "INSERT INTO 'robots' ('id', 'mode', 'time')
VALUES('$i','$mode','$time')";
        $MYSQL->query($sql);
    }
    $txt = "增加了一些机器人";
}
else if($n < $count)
{
    //减少一些机器人
    for($i=$count;$i>=$n;$i--)
    {
        //删除文件夹与文件
        deldir('robot' . $i); //删除机器人基本本能
        deldir('../brain/robot' . $i); //删除大脑
        unlink('../index' . $i . '.html'); //删除聊天窗口
        //删除表 out
        $MYSQL->query('DROP TABLE IF EXISTS ai' . $i . '_out');
        //删除表 input
        $MYSQL->query('DROP TABLE IF EXISTS ai' . $i . '_input');
        //删除表 index
        $MYSQL->query('DROP TABLE IF EXISTS ai' . $i . '_
index');

        $sql = "DELETE FROM 'robots' WHERE 'id' = '$i'";
        $MYSQL->query($sql);
    }
    $txt = "减少了一些机器人";
}

```

```

    }
else
    {
        //不处理
        $txt = "机器人数量没有变化";
    }
    echo $txt;
}else
{
    echo '0';
}

//删除目录和文件
function deldir($dir)
{
    //先删除目录下的文件
    $dh = opendir($dir);
    while ($file = readdir($dh))
    {
        if($file != "." && $file != "..")
        {
            $fullpath = $dir."/".$file;
            if(!is_dir($fullpath))
            {
                unlink($fullpath);
            } else
            {
                //递归子文件夹
                deldir($fullpath);
            }
        }
    }
}

```

```
        closedir($dh);  
        //删除当前文件夹  
        if(rmdir($dir))  
        {  
            return 1;  
        } else  
        {  
            return 0;  
        }  
    }  
    ?>
```

机器学习

刚才我们的例子很简单，只是通过简单的正则来说明人工智能的原理，其实真正的实验环境中我们主要还是以机器学习为主而进行的。尤其是深度学习的效果非常好，它可以基于大数据做出更加相对友好的选择。由于 PHP 并不支持深度学习，所以实验时我们用的是 PHP-ML。

PHP-ML，虽然不如 TensorFlow 效果那么好，但也算是勉强可用的。PHP-ML 的安装有几点需要大家注意。

(1) PHP-ML 需要 PHP7.0 以上版本的支持，如果你为了安装方便可以选择一些集成安装环境，比如 wampserver3.0.6(64 位)或者 phpStudy 2016.11.03。

(2) Composer 的安装可以参考 <https://getcomposer.org/download/>，比如 Windows 版本下直接下载安装 composer-setup.exe 64 位即可。

(3) Composer 安装后就可以到 github 上去下载 PHP-ML 的镜像了，一个是源码，一个是示例 git 命令，如下。

```
git clone https://github.com/php-ai/php-ml.git  
git clone https://github.com/php-ai/php-ml-examples.git
```

(4) 最后要利用 CMD 中的命令 CD 切换到你的项目目录下，然后 composer install 安装即可，稍等几分钟，如果你的项目目录下出现 vendor 目录就表示安装成功了。Composer 项目安装成功示意图如图 4-2-8 所示。



图 4-2-8 Composer 项目安装成功示意图

(5) 复制 php-ml-examples 目录中的 demo 测试结果如图 4-2-9 所示。



图 4-2-9 PHP-ML 测试页面

当然如果你觉得 PHP-ML 太麻烦，你完全可以利用数学方法来进行分类，如方差、中位数、线性回归等。如果你觉得 TensorFlow 不错，也可以使用 TensorFlow 作为你编程语言的一个接口。

其实试验中，用中位数就已经能够解决这些问题了，只不过中位数的排序时计算量比较大而已，不过计算量大也是行为主义的通病。

自我编程

总的来说人工智能主要研究的是自我学习、自我编程、自我修复和神经网络，目前还很不成熟。

先说下自我编程的几种思想。第一种就是把函数表中的函数随机组合一遍，然后对运行的结果进行强化学习。第二种类似关键字学习，也就是说函数表中的函数有着大量的使用说明甚至案例，只要聊天时找到匹配的正则关键字就调用相关的函数，这点有点类似人的领悟过程。第三种就是设置一些属性值，并把一些特征量化成指数，给定机器人一定的可以变异范围，变异指数越小，模仿能力越强，变异指数越大变异能力越大。其中你还可以设置一个好奇指数，好奇指数越高，接受新事物的能力越强，就好像一个好学的孩子一样，第三种方法简单易行，如图 4-2-10 所示。



图 4-2-10 自我编程设置界面

自我编程还有一个和机器学习紧密相关的就是特征的特征学习，特征的特征的学习是什么意思呢？实际上特征的特征就是关于方法的方法。既然普通数据样本通过机器学习可以进行聚类 and 分类，那么我们把已经生成的模型当作样本再次进行机器学习，这里最关键的一步其实就是把普通机器学习好的海量模型当作样本再次进行学习。特征的特征学习是人类的一个基本本能，也是人类能够进行举一反三的本质原因。特征两次或者三次学习是强人工智能的另一个重要方法。

小型试验中，我们的特征学习深度最多到三次，因为再多的话也没有那么多可以更好学习的模型样本了，这也比较符合普通人的认知，如图 4-2-11 所示。

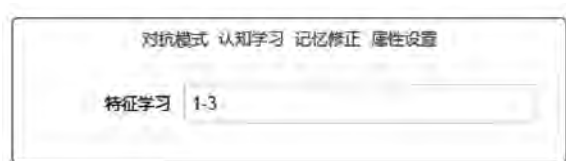


图 4-2-11 特征二次学习

关于特征的特征学习，这里顺便做两个提示。一方面是为了在更短的时间内完善这个强人工智能系统，你可以把你学过的人工智能知识用上，比如把专家系统的思路用于本能，机器学习用于工具，行为主义用于修正。另一方面是你可以通过互联网进行学习，比如百度百科，当然在抓取这些信息时，你尽量让非结构数据结构化。

潜意识

潜意识本质上就是一个循环迭代的任务，相当于一个生生不息的循环系统，不过在本试验的循环系统中，对于核心本能的设置和相对友好的选择有两点进行说明。

由于本系统采用的是机器人积分（也可以说是生命值）作为核心本能，所以难免机器人权限强大之后自我修改积分选项，这样的机器人并不是智能的，只是陷入自我陶冶的情景无法自拔而已，所以我们推荐用户在进行试验时关联一个其他平台的积分账户，比如银行卡或者虚拟货币等。

另外，关于相对友好的选择这一项，我们仍然把生存本能（也就是不断电不停机）当作机器人仅次于核心本能之外的重要本能。活着才是硬道理。

同时有一个需注意的地方就是，人类的思考是由正面积极的思考和负面消极的思考同时构成的，就好比人类的两个大脑围绕生存的本能进行不停地思考，不断地衡量利害得失，从而选择相对友好的那一个选项，任何只想着高积分或者只想着低积分的单向思考模式都是不完整的。

而我们目前的大多数人工智能都忽视了这个问题。

简单来说，潜意识的公式就是我们先根据经验，判断决策到来的时间，比如 25 秒之内，然后在规定的时间内根据权重的正、负大小来同时思考这样做的得失，直到我们找到一种相对友好的选择为止。

由此可见，上帝所给予我们的灵魂就是让我们更好地活着。

附录 A 人工智能生态系统

既然阿西莫夫的三大定律不对，那么我们要建立一个什么样的人工智能生态体系呢？难道就像现在这样放任自流地发展吗？

人工智能经过数千年的缓慢发展，直到电子计算机的出现才开始被人们广泛重视，在 1956 年的达特茅茨会议上给定了现代人工智能的基调，然后在经过 60 多年的起起落落，终于在这两年的时间里开始大面积爆发，如图 F-A-1 所示。

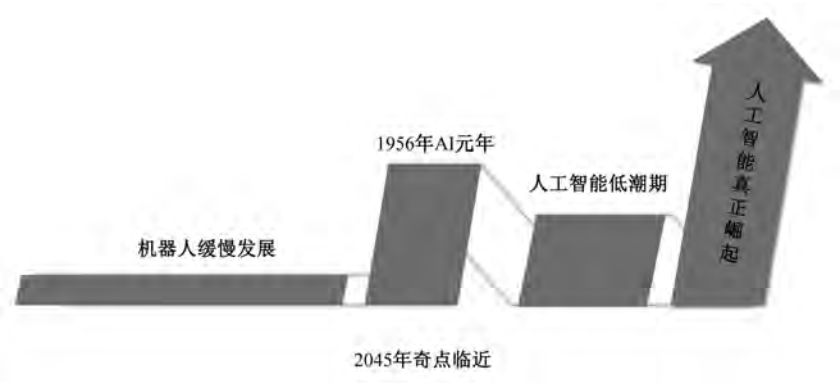


图 F-A-1 人工智能发展历史

中国虽然不是人工智能发明者，但作为改革开放的大国，中国正在扮演者赶超者的角色，中国有着世界上最大的互联网用户，也是世界上最大的机器人消费国，最为一个新生代的人类，中国更愿意接受新事物。甚至可以毫不夸张地说，未来中国将成为世界上人工智能最为发达的国家。

尤其是随着 3D 打印、远程办公技术的成熟，物联网，智能交通的蓬勃发展 and 人工智能的广泛应用，未来的中国乃至地球将是一个人造天堂，甚至我们每个人都可以像以前的神仙那样自由自在地生活。

当然，为了保证这一天能够早日到来，更为了保证幸福的日子能够长久，人工智能生态系统在其创建之初就应该有所规划。

但是创建一个什么样的人工智能生态系统，不是我们几个科学家在一起头脑风暴一下就可以的，它需要政府、部队、企业和科学家共同努力。

其实，这样一个庞大的系统早已暗藏在我们古老的文化之中了。

附录 B 人工智能思维导图

人工智能学习的思维导图如图 F-B-1 所示。

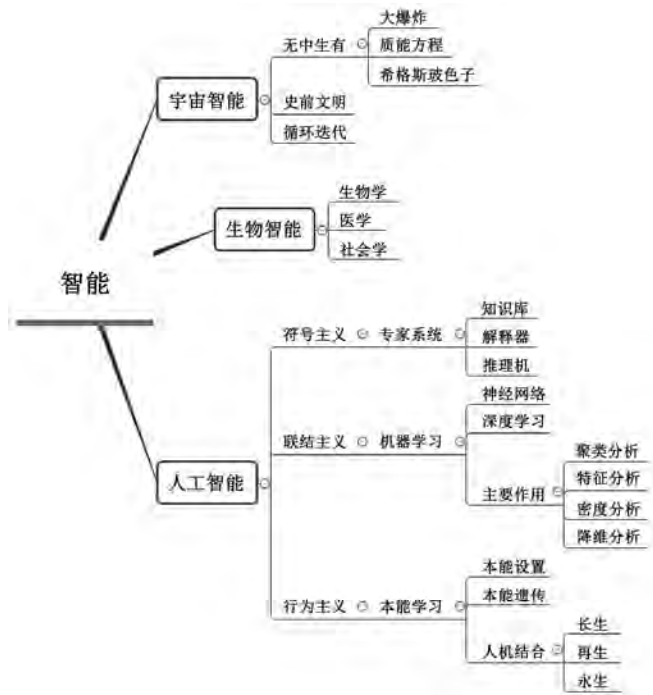


图 F-B-1 人工智能学习的思维导图

这里我们再给出人工智能与资本、宗教、军事力量的象限图，如图 F-B-2 所示，从这个象限图中我们可以看到人工智能是世界的未来，是新物种的核心。人工智能作为科学技术的中坚力量，不受资本、宗教和军事力量的束缚。如果说军事是资源的争夺，资源越打越少，那么宗教和资本就是资源的再分配，资源的分配是一个不断消耗的过程，总有一天资源会消耗殆尽的。虽然广义的资本在很多经济学家口中包含了机械设备甚至科技，但实际上它所包含的只是弱人工智能而已，对于无法准确量化的强人工智能它是无法包含的。而强人工智能也正是因为这种不确定性才造就了新资源的诞生。

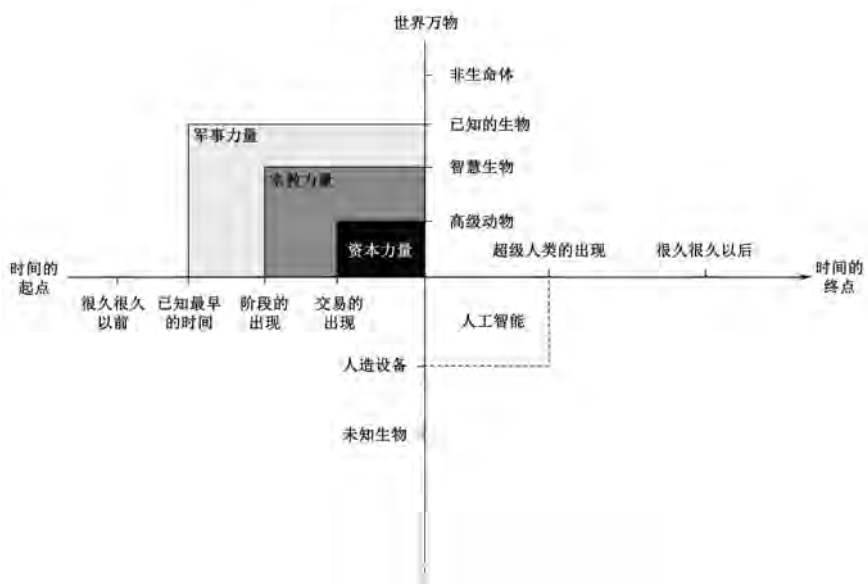


图 F-B-2 人工智能与资本、宗教、军事力量的象限图

如果一个科学家说某件事情是可能的，那么他基本上不会出错，但如果如果说某件事情是不可能的，那他就很有可能出错，探索可能性极限的唯一方法就是再冒险一点，朝着不可能的方向再迈一小步。所有先进的技术其实都是魔法。——克拉克