

# 人工智能简史

# 你一定爱读的



人工知能はどのようにして  
「名人」を超えたのか？

山本一成

北京出版

山本一成

纪晓雨 任轶(译)

从人类智慧的大视野，探索人工智能三大核心技术——机器学习/深度学习/强化学习的历史、本质与未来

# 你一定爱读的人工智能简史



[日]山本一成〈著〉 纪晗雨 任轶〈译〉

北京日报出版社

## 图书在版编目 ( C I P ) 数据

你一定爱读的人工智能简史 / (日) 山本一成著 ;  
纪晗雨, 任轶译. -- 北京 : 北京日报出版社, 2019.5  
ISBN 978-7-5477-3253-3

I. ①你… II. ①山… ②纪… ③任 III. ①人工智  
能-技术史-世界 IV. ① TP18-091

中国版本图书馆 CIP 数据核字 (2019) 第 047195 号

北京版权保护中心外国图书合同登记号: 01-2018-5096

JINKOCHINO WA DONOYONISHITE "MEIJIN" WO KOETA NO KA?

by Issei Yamamoto

Copyright © 2017 Issei Yamamoto

Simplified Chinese translation copyright © 2019 by Sunnbook Culture & Art Co.,Ltd.

All rights reserved.

Original Japanese language edition published by Diamond,Inc.

Simplified Chinese translation rights arranged with Diamond,Inc.

Through BARDON - CHINESE MEDIA AGENCY.

## 你一定爱读的人工智能简史

---

出版发行: 北京日报出版社

地 址: 北京市东城区东单三条 8-16 号东方广场东配楼四层

邮 编: 100005

电 话: 发行部: (010) 65255876

总编室: (010) 65252135

印 刷: 北京荣泰印刷有限公司

经 销: 各地新华书店

版 次: 2019 年 5 月第 1 版

2019 年 5 月第 1 次印刷

开 本: 880 毫米 × 1230 毫米 1/32

印 张: 9

字 数: 172 千字

定 价: 45.00 元

---

版权所有, 侵权必究, 未经许可, 不得转载

●●● \*E:\你一定爱读的人工智能简史\山本一成\卷末附录.exe

卷末附录  
世纪之战的意义何在?

●●● \*E:\你一定爱读的人工智能简史\山本一成\结束语.exe

结束语  
conclusion

\*整理: 崎谷秀穗

●●● \*E:\你一定爱读的人工智能简史\山本一成\第4章.exe

第4章  
当奇点出现时.....

●●● \*E:\你一定爱读的人工智能简史\山本一成\第3章.exe

Chapter  
第3章  
强化学习

●●● \*E:\你一定爱读的人工智能简史\山本一成\第2章.exe

第2章  
深度学习  
人工智能 脱离人

●●● \*E:\你一定爱读的人工智能简史\山本一成\第1章.exe

第1章  
机器学习——人工智能  
的热潮已

Chapter<2>

●●● \*E:\你一定爱读的人工智能简史\山本一成\前言.exe

前言  
preface

●●● \*E:\你一定爱读的人工智能简史\山本一成\目录.exe

目录  
Contents

“E:\你一定爱读的人工智能简史\山本一成\卷末附录.exe”

卷末附录  
世纪之战的意义何在?

“E:\你一定爱读的人工智能简史\山本一成\结束语.exe”

结束语  
conclusion \_

·整理: 崎谷美穗

“E:\你一定爱读的人工智能简史\山本一成\第4章.exe”

第4章  
当奇点出现时.....

“E:\你一定爱读的人工智能简史\山本一成\第3章.exe”

Chapter  
第3章  
强化学习——

“E:\你一定爱读的人工智能简史\山本一成\第2章.exe”

第2章  
深度学习——人工智能

脱离人

“E:\你一定爱读的人工智能简史\山本一成\第1章.exe”

第1章  
机器学习——人工智能  
Chapter<2> -

深度学习的热潮已

“E:\你一定爱读的人工智能简史\山本一成\前言.exe”

前言  
preface \_

回溯历史，从2008年至今，这10年时间或许可以说是人工智能（Artificial Intelligence，简称AI）发展历程中最重要的时期。

随着深度学习技术的大幅度发展，过去无法在计算机领域内解决的问题逐步被纳入可计算的范围。一些曾被看作天方夜谭的技术，如自动翻译、自动驾驶等，如今也正在一步步走向现实。

同时，在这过去的10年里，电脑在围棋、将棋（又称日本象棋）这类“象征智力”的游戏中多次创造了战胜人类顶级棋手的辉煌战绩。

谷歌所开发的围棋机器人“阿尔法狗”与本人所开发的将棋机器人“PONANZA”都是其中的杰出代表。

2016年，阿尔法狗在与世界顶级围棋手的对弈中一战成名，举世瞩目。

PONANZA则在“将棋电王战”的出场权争夺赛中与来自全世界的机器人选手进行对决，并豪取4连胜，赢得了出场权。它更在随后的职业将棋选手赛“电王战”中交出了5战5胜的惊人成绩单。2017年2月，PONANZA已当之无愧地登上了“最强将棋程序”的宝座。并且PONANZA已确定，将于接下来的4月对战顶级职业棋手佐藤天彦。

我所开发的PONANZA与谷歌公司的阿尔法狗从功能及性质上来看并不相同。然而，不论是将棋还是围棋，其人工智能都将致力于超越人类并不断升级至更高水平。从这点上来说，二者又是相同的。

阿尔法狗（据我所知）自诞生之日起，不过数年时间便取得了今

日的成就。PONANZA 虽诞生于 10 年前，然而其飞跃性的发展却是在近几年才得以实现的。

那么，为何近几年人工智能的发展势头会如此迅猛呢？本书的目的之一，便是向读者揭示这一现象产生的原因。

同时，我亦希望借此书向大家传达我（现阶段）关于“什么是智能”这一问题的思考。

PONANZA 在这 10 年里的不断成长给我带来了无数的惊叹和喜悦，在这期间我多次问自己，究竟“何为智能”“何为智力”。

其实，我们每天都在用智力解决各种问题，不过我们并没有刻意对此进行深入思考。

人类究竟是如何提升自己的将棋、围棋水平的呢？

人类究竟是如何作出决策的呢？

包括此刻正在阅读的你，究竟是如何做到理解文意的呢？

只有在特意被问到时，我们才会觉得不可思议。但在日常生活中，我们很少会留意这些问题。

在我每天为人工智能绞尽脑汁、魂牵梦萦的思考中，我反而渐渐弄明白了另一个问题的答案——何为人类的智慧。我认为，对这个问题的探讨可以让我们从更本质的角度来理解人工智能。

本书将讨论的绝不是晦涩难懂的话题，请诸位放心阅读。我们将以将棋、围棋为中心，继而探索“机器学习”“深度学习”“强化学习”

这三大人工智能核心技术的本质，并尽可能通俗地对其进行解读。此外，我将继续改进 PONANZA 的设计，并在本书中与大家同步分享设计过程中的直观感受，这将是绝无仅有的。

那么，让我们一起开始吧！

**山本一成 (2017 年 2 月)**

▷前言 \_ <5>

▽第 1 章 \* 机器学习——人工智能寒冬期宣告结束 \_ <001>

必须由将棋专家来编程吗? \_ <002>

计算 + 存储 = 电脑? \_ <006>

搜索 + 评估 = 智能? \_ <009>

将棋游戏中的搜索与评估 \_ <016>

评估体系的建立 \_ <019>

人工智能的“寒冬” \_ <022>

为什么杯子是杯子? \_ <027>

电脑将棋为何比电脑国际象棋晚了 20 年? \_ <033>

盘面复杂并不是输给人类的理由 \_ <035>

对电脑来说,将棋的难点在于…… \_ <038>

于电脑而言,将棋与国际象棋的本质区别 \_ <040>

电脑将棋中的机器学习 \_ <044>

机器学习的缺点 \_ <051>

PONANZA 的成长 \_ <057>

电王战 \_ <060>

从程序员的环节毕业 \_ <066>

▽第 2 章 \* 深度学习——人工智能的热潮已经到来 \_ <069>

黑魔法的影响力似乎越来越强了 \_ <070>

失败已成为家常便饭 \_ <072>

黑魔法之一:懒散的非列性 \_ <075>

深度学习下的人工智能热潮 \_ <080>

深度学习的前世今生 \_ <083>

脱离——深度学习的黑魔法 \_ <087>

几乎,每天都有新技术发表 \_ <092>

智能的本质即为图像? \_ <098>

还原主义科学的落幕 \_ <100>

▽第 3 章 \* 强化学习——人工智能可以脱离人类模板吗? \_ <103>

人工智能已经完全超越人类? \_ <104>

人类无法直观地理解“指数级的成长” \_ <106>

人类将面对与职业棋手相同的经历 \_ <109>

“守破离”——智能学习三阶段 \_ <112>

什么是“强化学习”? \_ <116>

PONANZA 流派的诞生 \_ <120>

“输给电脑了,真是没办法呀”	<124>
阿尔法狗的隆重登场	<129>
为何围棋游戏如此特殊?	<131>
救世主蒙特卡洛法	<135>
骰子也有智能吗?	<139>
蒙特卡洛围棋的成长	<142>
阿尔法狗所揭示的道理:围棋是图像	<144>
阿尔法狗的三大武器	<149>
合奏效果	<156>
科学变为宗教的刹那	<159>
智慧的本质在哪里?	<161>
第4章·当奇点出现时……	<165>
智能与智慧	<166>
水平越高的人,越有能力进行正确的目标设定	<169>
意义和故事:只为人类所有?	<176>
越接近人脑,越容易成功	<179>
PONANZA 2045	<183>
人工智能已在模仿人类的伦理观	<187>
当奇点出现时	<190>
>结束语	<199>
卷末附录·世纪之战的意义何在?	<207>
战胜人类的阿尔法狗是怎样炼成的?	<208>
在图像识别上,电脑超越了人类	<213>
“这真是一个力气活儿”	<218>
阿尔法狗的第六感真是太强了	<223>
“对不起,没能讲解好这一局”	<229>
并非天生强大的终盘战斗力	<234>
所有人都开始承认阿尔法狗的实力	<238>
与人类不同,电脑的情绪是不会动摇的	<242>
“这种下法也太没品了”	<247>
“之前明明那么强大,现在好像一下子变笨了!”	<253>
“围棋的本质是快乐”	<260>
人工智能的无限可能	<267>
科幻漫画里的世界早就到来了	<272>

●●● "E:\你一定爱读的人工智能简史\山本一成\卷末附录.exe"

# 卷末附录

谷歌人工智能与人类

●●● "E:\你一定爱读的人工智能简史\山本一成\结束语.exe"

# 结束语

conclusion \_

——  
在?

●●● "E:\你一定爱读的人工智能简史\山本一成\第4章.exe"

# 第4章

当奇点出现时.....

谷实穗 \_

●●● "E:\你一定爱读的人工智能简史\山本一成\第3章.exe"

Chapter

# 第3章

强化学习

●●● "E:\你一定爱读的人工智能简史\山本一成\第2章.exe"

# 第2章

深度学习

的热潮已

Chapter<2> \_

脱离人

人工智能

●●● "E:\你一定爱读的人工智能简史\山本一成\第1章.exe"

# 第1章

机器学习——人工智能

寒冬期宣告结束

Chapter<1> \_

# 1 必须由将棋专家来编程吗？

10年前，正在读大二的我不小心留了级，那时的我整日泡在东京大学的将棋俱乐部里消磨时间。

由于留级的缘故，我平时没有太多课，反而有很多的闲暇时间。我思考着未来，开始阅读此前从未接触过的编程类书籍，并为之深深着迷。

要想开发出能打败将棋专家的程序，究竟应该怎么做呢？

必须由将棋专家来编程吗？把他关于将棋的所有知识一一录入电脑，这样的程序才有可能打败他自己，不是吗？

当时，我的将棋水平是在业余五段。虽不能与职业棋手相提并论，但在全国数百万的将棋爱好者中，我也可以跻身前 1000 位。此外，我还曾获得关东地区学生赛的亚军，可以说实力尚可。

事实上，包括个人和团体在内，当时致力于电脑将棋开发的人员约有 50 人，我应该是所有人中将棋水平最高的。

因此，我理所应当认为，我写出的程序即使尚不能战胜将棋专家，也一定是所有程序中最强大的。

虽然直到大三为止，我都还只是一个完完全全的理科生，对电脑的基本操作都不甚熟悉。但在下定决心以后，我迅速地投身于对人工智能与编程的学习中，渐渐地，我将所有的时间都花在了将棋程序的设计开发上。

数月后，我完成了初代将棋程序“PONANZA”的开发工作。

带着 PONANZA，我向将棋俱乐部飞奔而去。

俱乐部的成员一直都知道我在写将棋程序。事实上，当时业界最先进的电脑将棋已经逼近了业余高级棋手的水平，因此所有人都做好了迎接劲敌的准备。

我启动了 PONANZA。

紧张的气氛随之弥漫开来。然而……

我的 PONANZA 一败涂地。其水平之低远远超出了所有人的想象。我还曾在对弈中故意让 PONANZA 8 个子，人为制造了非常悬殊的战力差来进行对抗。（图 1-1）

后 8 枚棋子意味着我（上）正在用极弱的战力与 PONANZA（下）对抗。一般而言，这样的战力差足以让稍有将棋知识的人轻轻松松赢过职业棋手。



图 1-1 我与 PONANZA 初次对战的情况

然而,即使让了 PONANZA 8 个子,我还是赢了。

我从未如此希望自己输掉这盘棋。我拼尽全力、倾注了所有热情编写的将棋程序,几乎是我那几个月全部生活目标的程序,惨遭失败。

当时我接触编程尚不足半年。不可否认,我的编程技术还十分稚嫩与笨拙,但这并不是问题的关键。

擅长下将棋的我误以为自己“真的懂将棋”了。但从某种意义上说,我其实对将棋一无所知。非但如此,当年的我对智能完全没有概念——智能究竟是什么呢?

# 2 计算 + 存储 = 电脑？

这个问题 10 年前就引起了我的注意，在本书对该问题进行深入展开前，我们先来了解一些基本的概念——什么是电脑，什么是编程？

在现代社会中，电脑是不可或缺的。小到人手一部的手机，大到飞机上的操控系统，生活中有着数不清的电脑装置。但从本质上讲，这些为丰富的活动提供支持、看似功能繁多的电脑装置实际上能做到的只有以下两点。

## 其一，“单纯的计算”。

在进行加减乘除及其他同级别的运算时，电脑的计算速度非常可怕，简单来说就是计算器的功能。其运作机制虽然简单，但光凭惊人的速度这一点，也足以满足很多需求了。

## 其二，“存储”。

大家对于电脑强大的存储功能并不陌生，它庞大的存储量可以帮助我们解决很多问题。例如电子词典，仅仅是将英语单词等相关内容存储进电脑，便可使我们的查询效率比翻阅纸质词典时更加高效。

除了不断重复“单纯的计算”和“存储”两大功能以外，电脑基本没有更多的功能了。我们之所以认为电脑有能力完成除此以外的更多任务，是因为程序员通过编写程序给电脑下达命令，使它能够更好地利用这两种功能去解决多样的问题。

那么，拥有此种特性的电脑在面对将棋时，该如何设计其程序呢？记录下所有职业棋手的棋局就可以变强吗？

遗憾的是，答案是否定的。与今天相比，10年前的电脑将棋虽然实力极弱，但要记录下所有职业棋手的对弈过程也并非难事。然而，不论10年前的电脑收录了多少棋谱，它都一直未能击败人类。