

不要再怪我没教你： Logic-based AI 入门

YKY 甄景贤

Independent researcher, Hong Kong

generic.intelligence@gmail.com

June 30, 2019

Talk summary

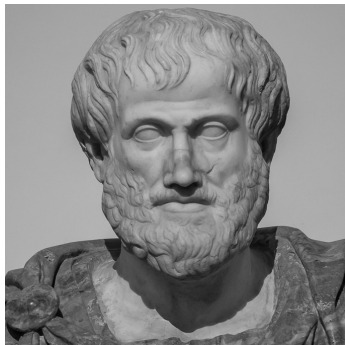
- 1 历史
- 2 绪论
- 3 各种 逻辑
- 4 逻辑 AI 系统的 架构
- 5 推导 算法
- 6 概率 / 模糊 逻辑
- 7 学习算法

Section 1

历史

Aristotle (ca. 300BC)

- 逻辑 是 研究 (人类) **思考方式** 的学问
- 亚里士多德 研究
三段论 (syllogism)
- eg:
所有人都会死
苏格拉底 是人
⇒ 苏格拉底 会死



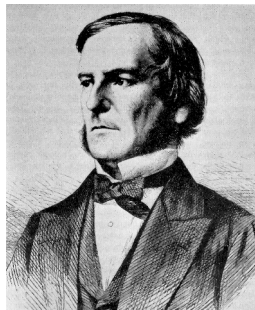
Leibniz (1646-1716)

- 他预言 在未来 人类的争论 可以用数学逻辑 透过 **计算** 解决 (*calculus ratiocinator*)
- 逻辑学中的大部分规律, 早已在他的著作中出现
- 可惜著作失传, 直到近代才翻译, 对逻辑学发展的影响不大



George Boole (1815-1864)

- 企图用 **代数** (algebra) 模拟 逻辑
- eg: $x = \text{人}$, $y = \text{动物}$
 $xy = \text{人是动物}$
- 他发明的逻辑 和我们今天讲的 Boolean 代数有些不同



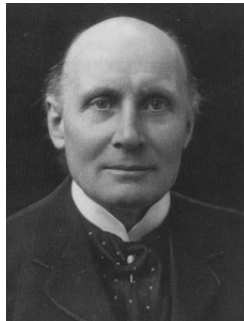
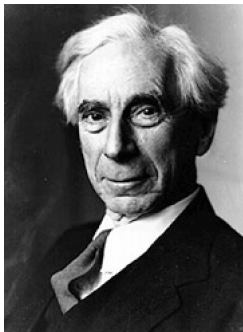
Gottlob Frege (1848-1925)

- 发明 **predicate logic** (谓词逻辑), 第一种可以表述大部份思想 / 语言的逻辑
- Russell 写信告诉他, 他的逻辑中存在 **悖论** (paradox)
- 他说: 「一个作者最惨就是, 正当付梓印刷时, 才发现书的基础有问题」!



Principia Mathematica (1910's)

- Russell & Whitehead
巨著
- 奠定了
谓词逻辑
在数学中的
地位



- John McCarthy 在 1950's 将 数理逻辑 引入 AI

Emil Post (1897-1954)

- 犹太裔 天才儿童
- 12 岁时 车祸 失去左臂
- 开创 递归论、可计算性理论
- Manic-depressive 精神病
- 接受 电击疗法 时 心脏病发致死，死时 57 岁



Haskell Curry (1900-1982)

- 他的姓和名 都变成了编程语言!
- 发展 **combinatory logic**, 一种不需用 **变量** 的逻辑
- Currying 的意思是 $f(x, y)$ 变成 $\hat{f}(x)(y)$
- Curry-Howard correspondence, **逻辑** 与 **计算** 之间的深邃关系



Alfred Tarski (1901-1983)

- 开创 **模型论**
(model theory)
- 开创 **cylindric algebra**,
揭示 谓词逻辑 的 代数
/ 几何 结构
- 研究 theory of truth
(这部分我不太清楚)



Alonzo Church (1903-1995)

- 发明 λ -calculus
- 目的是研究 代入 (substitution) 的机制
- 是任何 可计算函数 的 universal form (此即 Church-Turing hypothesis)



Kurt Gödel (1906-1978)

- 受 纳粹 迫害，逃到美国
- 晚年有 被害妄想症，怕被人下毒，只吃妻子煮的食物
- 妻子是护士，有一次她生病入院 6 个月，结果 Gödel 在家**饿死**了！
- **不完备性** (incompleteness)
似乎不妨碍 AGI 的发展
“not a show-stopper”



Richard Montague (1930-1971)

- Tarski 的学生
- 将英语语法的一个 fragment 转译成 谓词逻辑
- 由此证明了 自然语言 可以翻译成 逻辑 (形式语义学)
- 同性恋者, 喜欢 性派对
- 某次他带了 3-4 人回家, 第二天被发现勒死在浴室, 死时 40 岁
- 疑犯将他的汽车撞毁并焚烧
- 这案件至今未侦破



Section 2

绪论

为什么要学 logic-based AI?

- **神经网络 / 深度学习** 是现时 (2019) 最强的学习算法
- 某种意义上说，逻辑 AI 已经 **过时**
- 很多人（包括外国）不熟悉 逻辑 AI，所以看不懂我的 理论，还有 OpenCog、OpenNARS 这些「第一代」AGI
- 逻辑 是 强人工智能 的基础

什么是「命题」？

- 命题 是 任何可以 赋予 真值 的表达式
- 「山楂饼」不是命题
- 「我爱吃山楂饼」是命题
- 命题 由 概念 (concepts) 组合而成

以上似乎是 逻辑 最基本的 预设

反对 逻辑，论点 #1

有没有 不能用 逻辑 表达的思想 (thoughts) ?

- 如果有，你必须把它用文字表述出来
- 基本上，任何语言表述 可以 转化成 逻辑形式
- 所以，一定讲不出来

事实上，一些足够强的逻辑，包括一阶谓词逻辑，已经可以表达任何 可计算函数

神经网络 vs 逻辑



反对 逻辑，论点 #2



Section 3

一些 逻辑 理论

命题逻辑 vs 谓词逻辑

- blah

一阶逻辑 vs 二阶 / 高阶逻辑

- blah

范畴 逻辑 (categorical logic)

- blah

- blah

Section 4

逻辑 AI 系统的架构

- blah

Section 5

推导 算法

消解 (resolution) 算法

- 1965 年 J A Robinson 发明
- 它是 **推导** (deduction) 的最基本形式
- 在一堆 conjunctions 之中, 将两个 **互补** 的项消去, 例如 A 和 $\neg A$
- $\neg A \vee B, A \vee B \Rightarrow B$
- 可以将欲推导的 **结论** 反转, 加入前提中, 再用消解法 导出矛盾 (反证法)

消解 (resolution) 算法 (例子)

- 移民外国，寄人篱下，活得没有尊严
→ 惨
- 留在本国，制度压抑，不能舒展抱负
→ 惨
- $A \rightarrow B \equiv \neg A \vee B$ 所以有：
- $\neg \text{移民} \vee \text{惨}, \text{移民} \vee \text{惨} \Rightarrow \text{惨}$
- 移民也惨，不移民也惨 \Rightarrow 惨！

同一化 (unification) 算法



加速：“rete”算法



Section 6

概率 / 模糊 逻辑

贝叶斯 (Bayesian) 网络

- blah

- blah

- blah

Section 7

学习算法

- blah

- blah

?



多谢收看 😊