

《AGI in 5 minutes》

YKY

October 31, 2021

- 我预计 AGI 会在 5-6 年内出现。
- 理由是 因为 BERT/GPT 已经示范了
现时计算机的算力
已经进入了 AGI 的 “ballpark” 范围
- BERT/GPT 可以 回答问题、写文章、甚至写 code,
这些能力 毫无疑问是 通用智能 的特性
- 虽然很多人见到了 BERT/GPT,
但并不意识到它们就是 AGI 的雏型

AGI Early and Current Players

YKY and Ben Goertzel, founder of OpenCog. (2009 photo)



YKY and Pei Wang (王培), founder of OpenNARS.



- OpenCog 和 OpenNARS 都是基于 逻辑 的引擎, 但加入了某种 probabilistic 或 uncertainty logic.

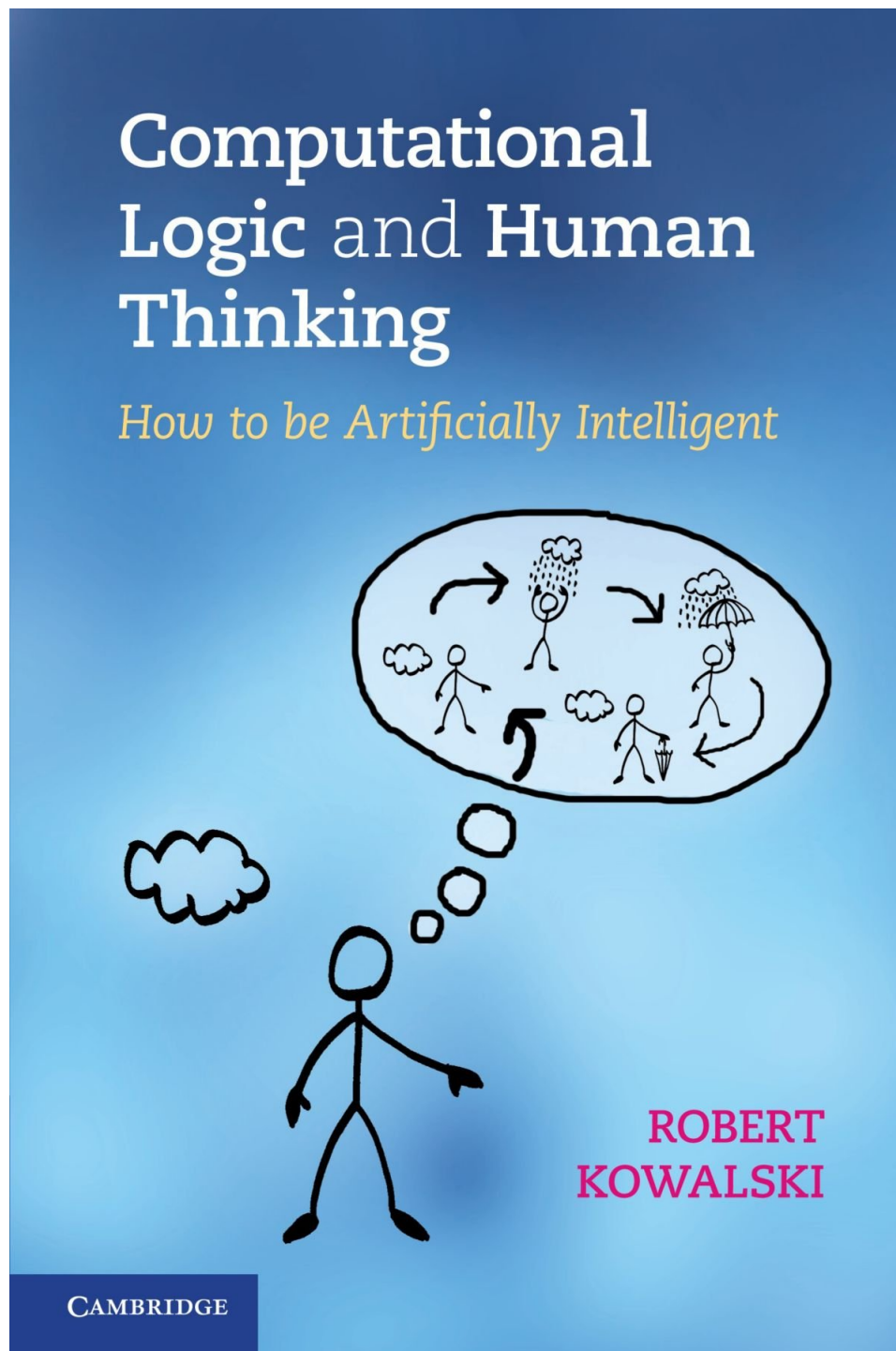
- 他们的 uncertainty 处理方式有点麻烦，于是我提出了一种 fuzzy-probabilistic logic (2012 paper) 但没有被采纳，然而这争拗 很可能已不再是重点所在。
- 他们是我的前辈也是朋友，也是竞争者，我非常欣赏他们不遗余力地推广 AGI，Ben 在全世界 AGI 圈子内很有名，王培 在国内也培养了很多 AGI 人材，而 OpenNARS 最近也有国际上的研发者参与
- 但我个人觉得 现代 AGI 已被 deep learning 占了主导地位
- DeepMind 是现时世界最强的 AGI 研究中心，他们有些早期成员来自 Ben 的 AGI mailing list. DeepMind 是一间对 AGI 很著迷的公司，他们绝对不止于玩玩围棋或 Atari 而已，正如他们的名称所示。
- Ben 早期的项目叫 Novamente，(他是天才，在南美长大，23 岁 PhD) Novamente 在西班牙语意思是“again”，但也可以理解为“new mind”。

Logic as the basis of AGI

很多人怀疑：符号逻辑 可以作为 人工智能的基础吗？

可以看看这本书，作者 Robert Kowalski 就是 Prolog 语言的创始

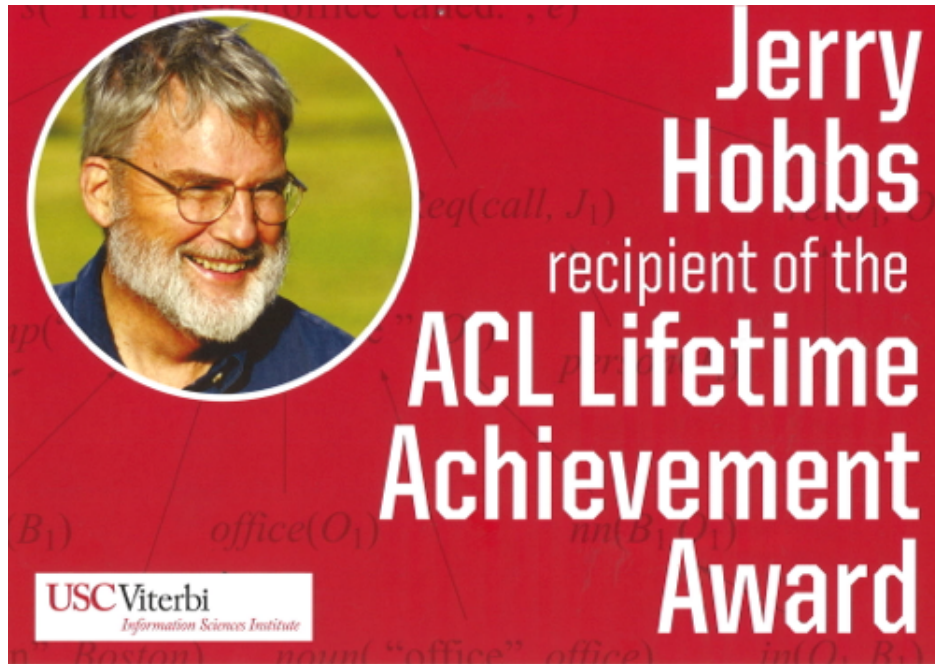
人之一：



- 这本书解释，在逻辑里主要有三种推理模式：
 - deduction（前向推理）
 - abduction（溯因推理，找出命题 A 的解释，即 $B \Rightarrow A$ ）
 - induction（归纳推理，从事例中总结出，例如「所有书呆子都是四眼的」）

- 这些逻辑模式，可以涵盖所有人类思维

Jerry Hobbs 提出 abductive interpretation of natural language
这是用 经典逻辑的方式 解决 **自然语言理解** 的问题：



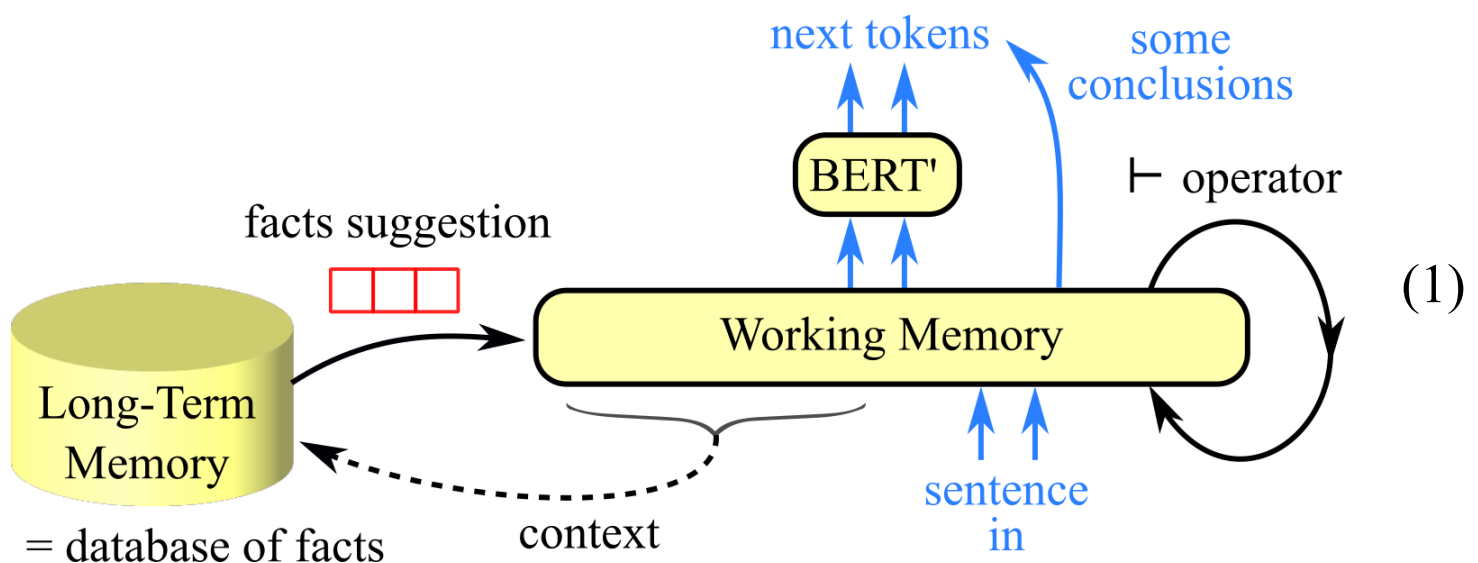
- 例如：
 - door-knob 是「门的手把」
 - school-girl 是「上学的女孩」
 - machine-gun 是「用机械发子弹的枪」这些组合词的诠释，各有不同，
它们需要用 abductive reasoning 解释。

The Problem of Logic-Based AI

- 经典逻辑 AI 有个 **严重问题**：
就是 缺乏 快速 / 高效率的 **学习算法**，
而学习算法是 AI 的灵魂，
这直接造成了历史上所谓的“AI Winter”。

- 经典逻辑 AI 要学习 巨量的 逻辑式子,
eg. $\forall x. \text{Human}(x) \Rightarrow \text{Mortal}(X)$
而这是在一个 discrete lattice 里面搜寻,
换句话说是一种 combinatorial search,
而 search tree 是所有逻辑句子的 syntax,
它的 branching factor 很大,
而在没有 oracle 「神喻」的情况下,
就 exactly 像 $P \neq NP$ 那种难解的情况。
- 深度学习的 **革命性**意义,
在于打破了 learning algorithm 的困境.
- 注意: 即使现在有了 深度学习,
逻辑 AI 仍然欠缺快速的 learning algorithm,
而这件事没有因为 深度学习的出现而解决。

Hot-swap architecture



- 问题是 BERT 和 Sym 的 state 是否可以混合?

- 主要目的是想将 BERT states 变成 ground states, 则以后可以「接地应用」
- 如果可以将 BERT 的知识内容 转移到 Sym, 似乎是有益的
- 问题是 如果用 RL + Sym 训练 masked language model, 凭什么 觉得 这样 学到的知识 是 grounded?
- 所谓 grounded 就是依照 internal consistency 训练, 但如果依照 token prediction 会如何?
- Next sentence / token prediction 不是真的知识, 因为缺乏深层理解, 有时甚至是错的
- 先前说过 如果要学习 logical consistency 需要 量度 压缩, 但深度学习没有 压缩 metric, 只能量度 prediction accuracy.
- 所以 token prediction 其实也没有影响 groundedness

BERT 的奇怪点

- N -in N -out 表示 它每次可以输出 N 个结论
- 但 Q, K, V 是同一组矩阵, 是不是表示 只有 **1 条** rule?
- 其实每个 概念 可以代表 1 条 rule, 类似 $Q \Rightarrow V$
- 从 Neural Turing Machine 的角度, 它是某种 content-addressed memory retrieval
- 以前也猜测过 Q-K-V 是一种 universal computation.
- 最有趣问题是 BERT 如何做到 \forall quantification?