

5 分钟介绍强人工智能

甄景贤 (King-Yin Yan)

General.Intelligence@Gmail.com

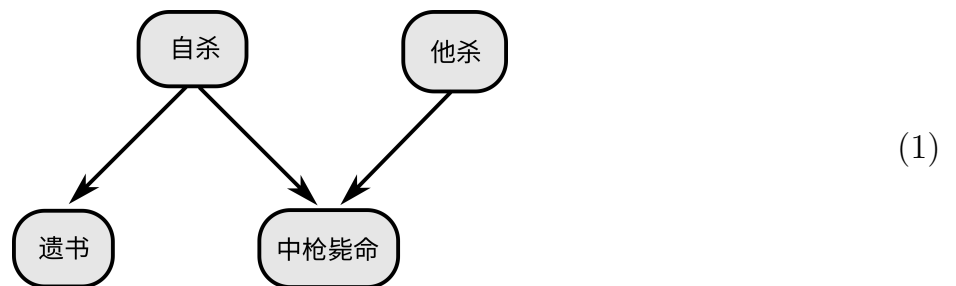
July 27, 2017

Abstract.

要研究 AGI, 最好熟悉一些 logic-based AI 的基础, 这些细节我打算出版一本书叫《强人工智能导论》, 正在找合作者写这本书, 有兴趣的可以联络。

Bayesian networks

首先由 Bayesian network 谈起, 它是一种命题之间的 graphical model, 用以计算某些 nodes 的未知的 probabilities:



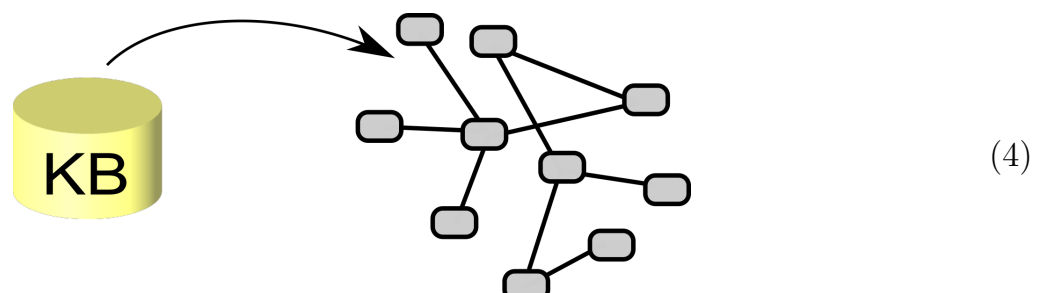
以下的过程叫 “lifting”:

$$\boxed{\text{propositional logic}} \xrightarrow{\text{lift}} \boxed{\text{Bayesian networks}} \quad (2)$$

AGI 的其中一个关键步骤是:

$$\boxed{\text{first-order logic}} \xrightarrow{\text{lift}} \boxed{\text{first-order Bayesian networks}} \quad (3)$$

这个问题基本上是解决了的, 其中一个方法叫 KBMC (knowledge-based model construction):



China vs USA

附带一提，即管在某些外国名校，如果能做到 lift Bayesian network to probabilistic case，那大概已经可以拿个 PhD，因为这已经算是一种“cutting edge”的工作（至少在大约 2000-2010 年这段时间，有很多 first-order Bayesian network 的论文出现）。当然我并没有贬低 PhD 的意思，但我想指出，在数学上 / 抽象的角度来看，这其实是一个很简单的动作。由此也可以看出其实 AGI 的数学化仍然处于颇为原始的阶段（例如，和现在高度数学化的 physics 相比之下）。

外国有足够的资源养活一班 PhD 研究员，而 lifting 这课题也确实算是 edge，但香港的科技颇为落后，资源也极缺乏（相对于人口比例），所以我和 Joseph Cheng 等很多香港人并没有 PhD。我们要在有限的资源下搞 AI。大陆的情况我不太熟悉，似乎你们也很少研究人材（相对人口比例，但总数量也很可观），然而问题是内地对 AGI 的接触似乎不多，教授们通常研究一些较传统的项目。

很多人关心「中国 vs 美国」的 AI 竞赛谁会赢，我觉得中国的研究人材虽然占人口比例小，但由于中国人多，所以和美国比也可以达到不相伯仲的地步。问题是其余的人口会不会将中国的人材「拖下去」其实是会的，例如百度，它比起 Google 是一个「较差」的搜寻引擎，而部分原因是，百度的搜寻质素被中国人民的平均知识水平拖低了。在 AI 的情况会是类似的，不过与其说中美较量，我更倾向支持国际化的合作模式。

而更重要的是：这不是瓶颈。

The bottleneck

AGI 的瓶颈是 learning。认识到这点是非常重要的，在 1980-90 年代有 AI winter 是因为逻辑 AI 系统的学习算法要依赖 ILP (inductive logic programming)，它是基于 combinatorial search 的算法，所以很慢。它慢到根本不能解决实际情况中出现的 problem size，所以只有很少的商业应用价值。而这一切在 deep learning 出现之后改变了；现在 deep learning 已经可以解决「翻译」和「视觉」这两大问题 — 曾经被认为是 AI-hard 的问题。

Using deep learning to do logic

PROLOG 的基本运作是这样的：

$$\boxed{\text{KB}} + \boxed{\text{query}} \xrightarrow{\text{PROLOG}} \boxed{\text{answer}} \quad (5)$$

它搜寻答案的算法是 unification 和 resolution，前者负责处理 variable substitutions，后者负责寻找逻辑命题的 proof tree。这些是逻辑 AI 的基础，在书里会有详细解释。

关键是用 deep learning 去模拟逻辑的 \vdash 运算， \vdash 是 **逻辑后果算子** (consequence operator)。

Bibliography