

# HW1

范潇 2254298

2024 年 3 月 5 日

图灵的论文建立在开篇设定的游戏及其规则之上，例如“询问者与被询问者隔离开来”、“对话以文字形式开展”、“测试对象是‘电子计算机’”等。这些规则使得图灵测试的目标聚焦于“机器能否思考”这一问题背后的本质。在此基础之上，图灵认为：到 2000 年，人类应该可以用 10GB 的计算机设备，制造出可以在 5 分钟的问答中骗过 30% 成年人的人工智能。

这篇论文成为经典的原因之一在于图灵从 9 个不同的角度，全面地考虑了可能的对于他的观点的反对意见，并逐一给出回复。我认为，这 9 个不同的角度可以概括为三个层面：1) 神学层面；2) 哲学层面；3) 科学层面。

图灵指出的第一条关于来自神学角度的反对，以及最后一条关于特异功能的讨论可以归为神学层面。图灵在他的论文中专门对从这一层面展开讨论有其历史背景，但是在科学文化发达的今天，来自这一方面的异议，尤其是对于人工智能的异议，已经毫无分量了。

”The ‘Heads in the Sand’ Objection”，”The Argument from Consciousness”，”The Argument from Informality of Behaviour” 这三方面的论述可以归于哲学层面。

我相信即使是在今天，仍有一部分人坚信机器是无法产生思想的。无论他们给出的原因是“人类科技还没有足够发达”，又或者只是毫无根据地坚信机器无法产生思想，背后是图灵所说的“鸵鸟思想”在作祟，这一点在人工智能领域快速发展的今天看来更加明显。图灵对于”The Argument from Consciousness”的讨论则体现出了图灵测试的合理性和灵活性，巧妙地避免陷入哲学中的唯我论。

在对于”The Argument from Informality of Behaviour”的论述中，图灵以自己编写的程序为例，从反面论证了“如果行为规律存在，我们无法找到它”。但是，如今机器学习的很多应用实际上便是从大量的数据中寻找规律。因此，通过大量程序输出，我们也可以在一定程度上寻找到该程序的行为规律，从这一方面来看，图灵的反驳还可以进一步完善。

在科学层面上，我认为来自数学方面的异议最有分量。首先，这一异议

来自己已经被证明了的哥德尔完备性定义以及其他数学定理——它们的正确性是不会随时间而改变的。其次，图灵对于该异议的反驳是基于人们对于人类智慧认知的局限性，而至今人类仍在对其进行探索之中，尚无定论。

另一个有分量的异议是”Lady Lovelace’ s Objection”。

正如拉芙莱斯女士受到当时的技术条件的影响而仍为机器无法创新一样，图灵受到当时科技水平的约束，并没有在论文给出具有说服力的实例来论证机器能让人“耳目一新”。但是，如今 Chat-GPT,Sora 等大模型的发布让大众切身体会到了机器的创新力，乃至产生了艺术工作者会因此而被淘汰的言论。于此同时，也有人认为这些大模型所生成的文章、图像和视频本质上只是输入给它的素材的堆砌和拼接，并无创新可言。由此可见，这个异议至今仍有分量。

综上，图灵所指出的一些异议直至现在仍有分量。回到他的观点本身，事实证明，图灵较为乐观——人工智能在 2012 年才第一次通过图灵测试。而至于“当今的计算机有多少可能性”这一问题，我认为需要进一步对明确这个问题。这里的“计算机”指的是人类集中当前顶尖的科学家所研制的计算机，还是一个普通家庭所能承担的商用计算机？如果是前者，我认为答案是肯定的。事实上，以我的个人经验而言，我极有可能无法分辨出 Chat-GPT 4 和人类。而对于后者，以这几年的发展速度来看，我认为 50 年之内便将成为现实。