

第一章

1-4. 为什么能够用机器（计算机）模仿人的智能？

物理符号系统的假设：任何一个系统，如果它能够表现出智能，那么它就必定能执行输入符号、输出符号、存储符号、复制符号、建立符号结构、条件性迁移 6 种功能。反之，任何系统如果具有这 6 种功能，那么它就能够表现出智能(人类所具有的智能)。

物理符号系统的假设伴随有 3 个推论。

推论一:既然人具有智能，那么他(她)就一定是一个物理符号系统。

推论二:既然计算机是一个物理符号系统，它就一定能够表现出智能。

推论三:既然人是一个物理符号系统，计算机也是一个物理符号系统，那么我们就能够用计算机来模拟人的活动。

1-5. 现在人工智能有哪些学派？它们的认知观是什么？现在这些学派的关系如何？

符号主义(Symbolicism)，又称为逻辑主义(Logicism)、心理学派(Psychologism)或计算机学派(Computerism) [其原理主要为物理符号系统(即符号操作系统)假设和有限合理性原理。]

认为人的认知基元是符号，而且认知过程即符号操作过程。认为人是一个物理符号系统，计算机也是一个物理符号系统，因此，我们就能够用计算机来模拟人的智能行为。知识是信息的一种形式，是构成智能的基础。人工智能的核心问题是知识表示、知识推理和知识运用。

联结主义(Connectionism)，又称为仿生学派(Bionicsism)或生理学派(Physiologism) [其原理主要为神经网络及神经网络间的连接机制与学习算法]

认为人的思维基元是神经元，而不是符号处理过程。认为人脑不同于电脑，并提出联结主义的大脑工作模式，用于取代符号操作的电脑工作模式。

行为主义(Actionism)，又称进化主义(Evolutionism)或控制论学派(Cyberneticsism) [其原理为控制论及感知-动作型控制系统]

认为智能取决于感知和行动。认为智能不需要知识、不需要表示、不需要推理；

人工智能可以象人类智能一样逐步进化。智能行为只能在现实世界中与周围环境交互作用而表现出来。符号主义、连结主义对真实世界客观事物的描述及其智能行为工作模式是过于简化的抽象，因而是不能真实地反映客观存在的。

以上三个学派长期共存与合作，取长补短，并走向融合和集成，为人工智能的发展作出贡献。

1-9 人工智能的基本研究方法有哪几类？它们与人工智能学派的关系如何？

1. **功能模拟法**：认为智能活动的基础是物理符号系统，以符号处理为核心对人脑功能进行模拟。符号主义
2. **结构模拟法**：认为认知过程不是符号处理，提出对人脑从结构上进行模拟。连接主义
3. **行为模拟法**：智能不取决于符号和神经元，取决于感知和行动，提出智能行动的“感知-动作”模式。行为主义
4. **集成模拟法**：集成以上三种方法。

补充：

解释什么是图灵测试？

维基百科：图灵测试（英语：**Turing test**，又译**图灵试验**）是[图灵](#)于 1950 年提出的一个关于判断[机器](#)是否能够思考的著名[思想实验](#)，测试某机器是否能表现出与人等价或无法区分的智能。测试的谈话仅限于使用唯一的文本管道，例如计算机[键盘](#)和[屏幕](#)，这样的结果不依赖于计算机把单词转换为音频的能力。

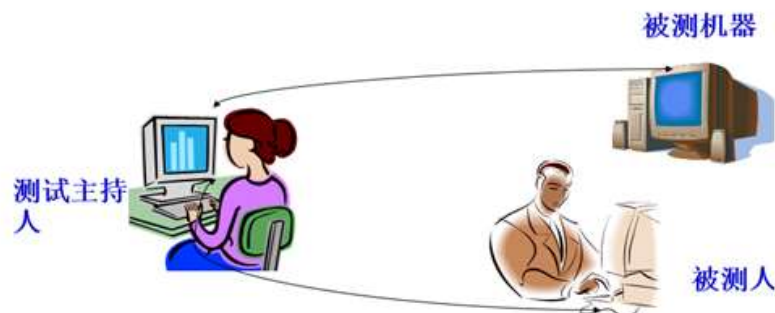
图灵测试

怎么判断机器有没有智慧 ——基于行为的测试

让一个人在某一房间里，把计算机配置在另一房间里。现在把双方都作为“人”来对待，如果在交谈中不能分辨出哪一个是人，哪一个是计算机，那么可以说这台计算机是智能的(Intelligent)。



阿兰·图灵
(1912-1954)
现代计算机之父



阐述中文屋问题是否能驳倒图灵测试？

不能：

不能，图灵测试中，那台机器接受信息并做出反馈 我们需要测试的也正是这台机器；而“中文房间”中整个房间（说英语者+一套输入输出符号规则）接受信息并做出反馈。我们需要测试的却只是房间里的人，明显两个情境是不同的，所以正确地讲应该是这个“房间”懂中文而不是“说英语的人懂中文”，这样的话这个反驳其实构不成反驳。中文屋可以看作图灵测试的另一种形式。

中文屋问题：一个对中文一窍不通，只说英语的人关在一间只有一个开口的封闭房间中。房间里有一本用英文写成的手册，指示该如何处理收到的汉语讯息及如何以汉语相应地回复。房外的人不断向房间内递进用中文写成的问题。房内的人便按照手册的说明，查找到合适的指示，将相应的中文字符组合成对问题的解答，并将答案递出房间。

不能驳倒。参阅手册并作出回复这个过程本身需要认知能力，而认知本身体现智能。所有人都认为人是拥有智能的，而人的智能决策来自于脑细胞的电信号转换，每一个脑细胞并不理解单词的意义，只是简单的缓冲、传递或抑制一个电信号，脑细胞创造了语法规则，创造了决策策略（相当于规则书与不懂中文的人），但是它们并不懂每个单词的意义。而人类却显示出与人沟通的能力。如果按照希尔勒的观点，那么人类也不存在认知能力，但这与事实是不符的。所以依然可以认为若某段计算机程序，能够完成图灵测试，则说明该段计算机程序具有认知能力。

能：

中文屋问题可以驳倒图灵测试。通过图灵测试并不意味着拥有智能。至多只能是智能的一个模拟。

这个思维实验是说，如果把一位只会说英语的人关在一个封闭的房间里，他只能靠墙上的一个小洞传递纸条来与外界交流，而外面传进来的纸条全部由中文写成。这个人带着一本写有中文翻译程序的书，房间里还有足够的稿纸和铅笔。那么利用中文翻译程序，这个人就可以把传进来的文字翻译成英文，再利用程序把自己的回复翻译成中文传出去。在这样的情景里，外面的人会认为屋里的人完全通晓中文，但事实上这个人只会操作翻译工具，对中文一窍不通。

在这个实验中，房间里的人不会说中文，他不能够用中文思考。但因为他拥有某些特定的工具，他甚至可以以中文为母语的人以为他能流利地说中文。电脑就是这样工作的。它们无法真正的理解接收到的信息，但它们可以运行一个程序，处理信息，然后给出一个智能的印象。所以说，即使表现出“人工智能”的机器，他也有可能实质上并不能理解人类的语言，所以说这个实验一定程度上驳倒了图灵测试。

中文屋悖论

通过图灵测试并不意味着拥有智能。
至多只能是对智能的一个模拟。



- 这个思维实验是说，如果把一位只会说英语的人关在一个封闭的房间里，他只能靠墙上的一个小洞传递纸条来与外界交流，而外面传进来的纸条全部由中文写成。
- 这个人带着一本写有**中文翻译程序**的书，房间里还有足够的稿纸和铅笔。那么利用中文翻译程序，这个人就可以把传进来的文字翻译成英文，再利用程序把自己的回复翻译成中文传出去。在这样的情景里，外面的人会认为屋里的人完全通晓中文，但事实上这个人只会操作翻译工具，对中文一窍不通。