

# 人工智能相关研究六问

宗成庆

中国科学院自动化研究所

cqzong@nlpr.ia.ac.cn

随着计算机硬件性能的提高和计算机网络技术的快速发展与普及，大数据、互联网+和人工智能等一批新老术语如八面来风横扫神州，尤其在中国大陆，这些术语所掀起的技术浪潮正以不可阻挡之势席卷华夏每一寸土地，可谓“忽如一夜春风来，千树万树梨花开”。作为这些术语所涉及领域的学习者和研究者，理所当然地为之振奋、欢呼，并摩拳擦掌。然而，当沸腾的热血稍稍降温 and 澎湃的心潮渐趋平静之后，无数的问题如挥之不去的天外来音萦绕在脑海，让我辗转反侧，难以入眠。我在兴奋与迷惘中追寻这些问题的答案，但我始终无功而返或者原地打转。我知道我的无知和愚钝，我也知道自己的浅薄与狭隘，于是，我把这些问题记下来，公示于众，万望赐教。

## 1. 关于人工智能

上个世纪 80 年代中期，在我读大学和研究生的时候人工智能曾一度如日中天，直到 90 年代中期，它仍是一颗耀眼之星而备受宠爱，之后便日渐黯淡，最终“臭不可闻”。30 多年后的今天，究竟缘何被人们再度热捧？是人工智能有了新的内涵，还是“新瓶装老酒”？如果说在过去的 30 多年里，计算机科学技术有了突破性进展，那么，除了硬件性能的飞跃性提高，人工智能的贡献到底有多大？究竟什么是人工智能研究的核心内容呢？当他“受宠”的时候，人工智能好像是一个框，什么内容都可以往里装，似乎做什么都属于人工智能。当他“失宠”的时候，人人唯恐避之不及。如果抛开自然语言理解、图像和视频识别、搜索算法、知识工程等（而这些方向也都有自己独立的学科），还有什么人工智能独家的研究内容呢？

## 2. 关于机器学习

近年来机器学习方法研究发展迅速，在不到 10 年的时间里迁移学习(transfer learning)、增强学习(reinforcement learning)、概率图模型(graph-based model)和深度学习(deep learning, DL)等一系列模型和方法被相继热捧，而深度学习方法出现之前的每一种方法都只不过仅持续两三年的时间便黯然失色。是科学家“喜新厌旧”，还是那些被冷落的方法和模型真地无能为力？我们知道，无论哪一种统计学习方法都是建立在大规模训练样本之上的“赌博”模型，难以做到举一反三。所以很多模型在实际应用的复杂场景中往往不如一个三岁的小孩。而儿童在学习和理解

某些问题时，并非都基于大规模样本（丰富的经验）。那么，目前的统计学习方法研究确实是在正确的道路上，沿着正确的方向前进吗？

深度学习方法与人工神经网络（artificial neural networks, ANN）方法在数学上没有本质的区别。难道仅仅是网络层数的增加，就能够从根本上解决实际问题吗？如果是这样，除了提高计算机的硬件性能，是否还需要研究其它的新模型和新方法呢？网络层数达到多少才算达到了“足够的深度”？对于某些特定的任务，如语音识别、图像分割等，理论上有没有 DL 方法的性能上限呢？

### 3. 关于自然语言理解

目前在自然语言处理(natural language processing, NLP)中很多问题被归为分类问题或序列标注问题，如汉语自动分词、语块识别、语义角色标注(semantic role labelling, SRL)和词义消歧(word sense disambiguation, WSD)等，所采用的分类器模型如支持向量机(support vector machine, SVM)、条件随机场(conditional random field, CRF)、最大熵(maximum entropy, ME)和贝叶斯分类器(Bayesian classifier)等，也都是其它领域中常用的分类模型，如在人脸识别、指纹识别、虹膜识别、图像分割等中也都在使用这些模型。直觉上讲，自然语言文本与图像是完全不同的表现形式，NLP 与机器翻译(machine translation, MT)等毕竟要解决的是语义理解问题。对于完全不同的问题却采用几乎相同的解决方法，自然难以得到很好的处理结果。那么，当前的自然语言处理研究是否走进了机器学习的“迷魂阵”？真正实现自然语言从“处理”到“理解”的出路何在？

### 4. 关于视频和图像理解

视频和图像分析作为模式识别领域的基础问题，在边界分割、模式匹配和检索等方面取得了若干优秀的成果，但是，这些成果在多大程度上体现了“智能”？离真正的“理解”还有多远呢？如下面的图 1 所示，人扫一眼就可以明显地看出这一副带有渲染色彩的美术作品，而不是一匹真正的骏马的照片，图像识别系统能够做到这一点吗？图 2 显然是一副带有艺术色彩的表现少女与骏马之间暧昧关系的真实的照片，也许图像识别系统不难切分出马脸与人脸之间的分界线，但是系统如何理解两者之间的深情含义呢？

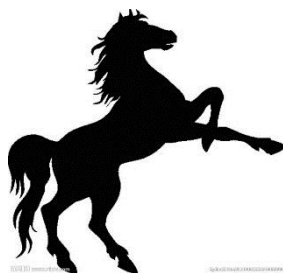


图 1



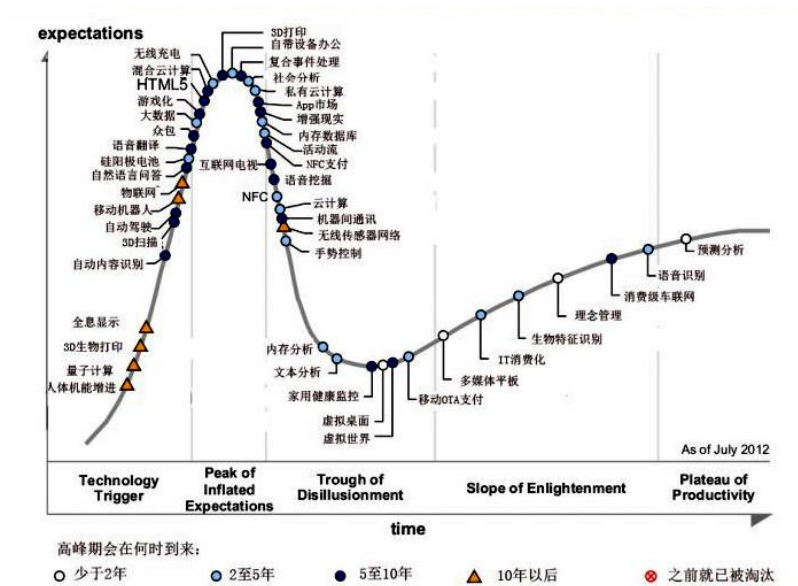
图 2

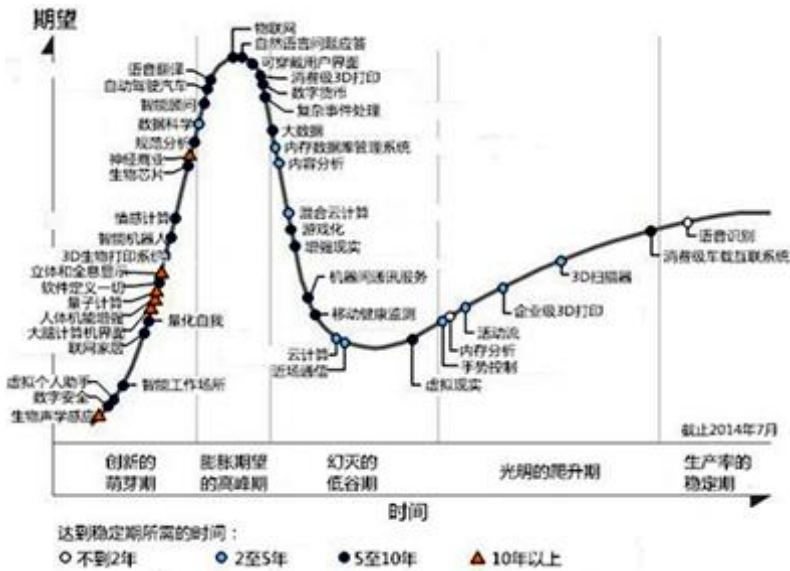
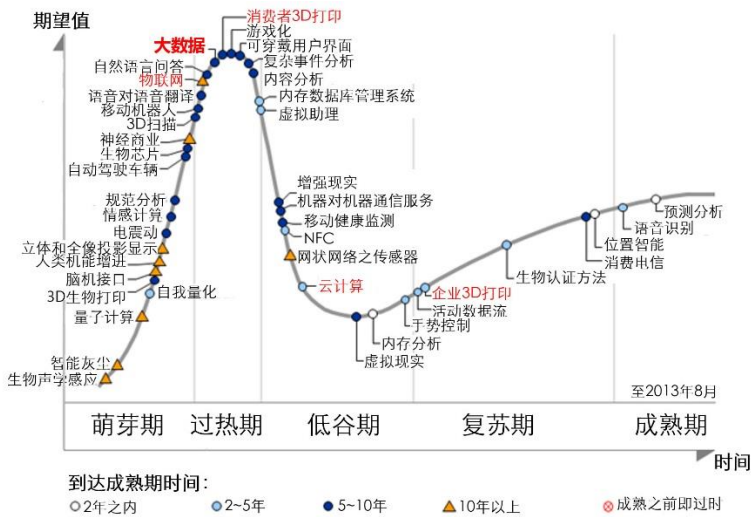
## 5. 关于智能系统

大多数实用系统都需要多种技术综合集成，而不是仅靠某个单项技术，如无人驾驶汽车。自动快速地辨识交通灯、车道线、周边车辆和障碍物以及汽车鸣笛声等，并进行主动避障，是保障无人车正常行驶的基本能力。每个单项的模式识别技术完成的只是类似人的眼睛和耳朵的功能。那么，是不是将这些单项技术简单地集成在一起就一定能够构成一个“智能系统”呢？如果不是，如何能够使得  $n$  项技术集成之后，其整体系统的性能远远超出  $n$  项技术的功能总和呢？“图灵测试”是测量一个实用系统智能性高低的最佳方法吗？

## 6. 关于坚守与跟踪

从近年来新技术发展的轨迹可以看出（见下面 2012 年至 2014 年三年间各项技术成熟预期随时间变迁的走势图），很多技术往往在短时期内被过高地抬举或炒作，有的只是虚拟的概念昙花一现，有的则是几十年前被冷落的技术由于硬件性能的改变而被再度推向期望的波峰，甚至多次轮回。回顾这些年来众多技术在峰回路转中所经历的命运变迁，我们不难看出，有时候科学家被舆论界所绑架，而有时候科学家又成为舆论界的“帮凶”，但似乎每个人都曾在这种忽上忽下、过山车似的变迁中感到兴奋和迷惘。而作为一个从事科学研究的学者（包括企业界从事技术研发的科学家），应该如何在这种热闹纷繁的世界中把握冷与热的平衡，既保持冷静的头脑，坚守自己的学术理想，潜心钻研，又不至于墨守成规、落伍于国际前沿，甚至被众多严苛的考核和竞争淘汰出局呢？





我确信，随着时间的推移和研究的加深，各类问题都将得到解决，或许有些问题根本不成其为问题。但在这些问题（包括“伪问题”）答案的追逐过程中我们所探索的和所获得的，都将成为这一学科发展的动力和内涵，我们自己也在为此而享受这一过程。

2015.9.6