**[UCLA朱松纯]浅谈人工智能：现状、任务、构架与统一（本地）**

强人工智能难度很大，常识，推导

小结一下，现在的人工智能和机器人，关键问题是缺乏**物理的常识**和**社会的常识**“Common sense”。 这是人工智能研究最大的障碍。那么什么是常识？常识就是我们在这个世界和社会生存的最基本的知识：（1）它使用频率最高；（2）它可以举一反三，推导出并且帮助获取其它知识

**其一、人工智能是一个非常广泛的领域**。当前人工智能涵盖很多大的学科，我把它们归纳为六个：

 （1）计算机视觉（暂且把模式识别，图像处理等问题归入其中）、

 （2）自然语言理解与交流（暂且把语音识别、合成归入其中，包括对话）、

 （3）认知与推理（包含各种物理和社会常识）、

 （4）机器人学（机械、控制、设计、运动规划、任务规划等）、

 （5）博弈与伦理（多代理人agents的交互、对抗与合作，机器人与社会融合等议题）。

 （6）机器学习（各种统计的建模、分析工具和计算的方法）

真正的智能体：感知、认知、推理、学习、和执行

第一阶段：前30年以数理逻辑的表达与推理为主。这里面有一些杰出的代表人物，如John McCarthy、Marvin Minsky、Herbert Simmon。他们懂很多认知科学的东西，有很强的全局观念。这些都是我读大学的时候仰慕的人物，他们拿过图灵奖和其它一堆大奖。但是，他们的工具基本都是基于数理逻辑和推理。这一套逻辑的东西发展得很干净、漂亮，很值得我们学习。大家有兴趣，可以参考一本最新工具书：The Handbook of Knowledge Representation，2007年编写的，1000多页。但是，这些符号的知识表达不落地，全书谈的没有实际的图片和系统；所以，一本1000多页的书，PDF文件只有10M，下载非常快。而我现在给的这个讲座，PPT差不多1G， 因为有大量的图片、视频，是真实的例子。

要学的还有很多，学的内容可查看文章

1984年，他发表了一篇文章，开创了computational learning theory。他问了两个很简单、但是深刻的问题。第一个问题：你到底要多少例子、数据才能近似地、以某种置信度学到某个概念，就是PAClearning；第二个问题：如果两个弱分类器综合在一起，能否提高性能？如果能，那么不断加弱分类器，就可以收敛到强分类器。

广义模式理论

多领域融合，统一的系统，数理逻辑

我们正在进入这么一个大集成的、大变革的时代，有很多机会让我们去探索前沿，不要辜负了这个时代。这是我演讲的第一个部分：人工智能的历史、现状，发展的大趋势。

人类环境当中的很多任务

大数据、小任务范式

小数据、大任务范式（更复杂的环境、更多的任务和价值，更少的单任务数据）

人的各种感知和行为，时时刻刻都是被任务驱动的

案例：imageNet上迁移学习能够取得很好的结果，imageNet有很多二分类任务。多任务生成的模型能更好迁移到特定的下层问题。

**任务塑造了智能**

如果我们把整个发展的过程都考虑进来，智能系统的影响可以分成三个时间段：（1）亿万年的进化，被达尔文理论的一个客观的适者生存的pheontype landscape驱动；（2）千年的文化形成与传承；（3）几十年个体的学习与适应。 我们人工智能研究通常考虑的是第三个阶段。

测量比例和精度根据任务改变

模型拼接

意识的分配，任务状态，感官的尺度

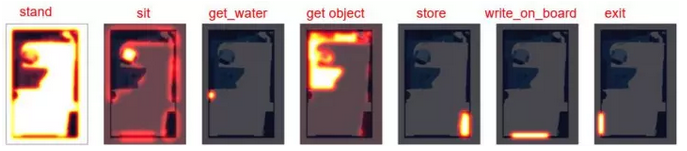
可以延展，有侧重点的模型

再往里面走一步，就进入人与动物的内心世界Mind, 内心世界反映外部世界，同时受到动机任务的影响和扭曲。研究内涵包括：

* Ta看到什么了？知道什么了？什么时候知道的？这其实是对视觉的历史时间求积分。
* Ta现在在关注什么？这是当前的正在执行的任务。
* Ta的意图是什么？后面想干什么？预判未来的目的和动机。
* Ta喜欢什么？有什么价值函数？这在第九节会谈到具体例子。

认知需要对象，考虑对象的反应学习认知能力

拓扑学是什么意思？就是说图象空间，语言空间，就是一个大集合，全集。我们的每个概念往往是它的一个子集，比如说，所有的图象是一个集合，一百万个象素就是一百万维空间，每张图像就是这百万维空间的一个点。人脸是个概念，所有的人脸就是在这一百万维空间的一个子集，但是这个子集和其它个子集要发生关系，这个关系叫拓扑关系。计算机的人把它叫做语法，对应于代数拓扑。比如，头和脖子在肩膀上是合规的，概率很高。这个图像空间的结构其实就是语法，这个语法就是STC-AOG，时空因果的与或图。语法可导出“语言”，语言就是一个符合语法的句子的总的集合。

当机器人重建了三维场景后（在谈视觉的时候提到了，这其实是一个与任务、功能推理的迭代生成的过程），它就带着功利和任务的眼光来看这个场景。如下图所示，哪个地方可以站，哪个地方可以坐，哪个地方可以倒水等等。下面图中亮的地方表示可以执行某个动作。这些图在机器人规划中又叫做Affordance Map。意思是：这个场景可以给你提供什么？

有了这些单个基本任务的地图，机器人就可以做任务的规划。这个规划本身就是一个层次化的表达。文献中有多种方法，我还是把它统一称作一种STC-PG。这个过程，其实相当复杂，因为它一边做，一边还要不断看和更新场景的模型。因为我前面介绍过，对环境三维形状的计算精度是根据任务需要来决定的，也就是Task-Centered视觉表达。

这个动作计划的过程还要考虑因果、考虑到场景中别人的反应。考虑的东西越多，它就越成熟，做事就得体、不莽莽撞撞。

前面谈的五个领域，属于各个层面上的“问题领域”，叫Domains。我们努力把这些问题放在一个框架中来思考，寻求一个统一的表达与算法。而最后要介绍的机器学习，是研究解决“方法领域”（Methods），研究如何去拟合、获取上面的那些知识。打个比方，那五个领域就像是五种钉子，机器学习是研究锤子，希望去把那些钉子锤进去。深度学习就像一把比较好用的锤子。当然，五大领域里面的人也发明了很多锤子。只不过最近这几年深度学习这把锤子比较流行。

学习的极限与“停机问题”

包括以下七种学习模式（每种学习模式其实对应与图中的某个或者几个箭头），这里面还有很多模式可以开发出来。

（1）被动统计学习passive statistical learning：上面刚刚谈到的、当前最流行的学习模式，用大数据拟合模型。

（2）主动学习active learning：学生可以问老师主动要数据，这个在机器学习里面也流行过。

（3）算法教学algorithmic teaching：老师主动跟踪学生的进展和能力，然后，设计例子来帮你学。这是成本比较高的、理想的优秀教师的教学方式。

 (4) 演示学习learning from demonstration：这是机器人学科里面常用的，就是手把手叫机器人做动作。一个变种是模仿学习immitation learning。

（5）感知因果学习perceptual causality：这是我发明的一种，就是通过观察别人行为的因果，而不需要去做实验验证，学习出来的因果模型，这在人类认知中十分普遍。

（6）因果学习causal learning：通过动手实验， 控制其它变量， 而得到更可靠的因果模型， 科学实验往往属于这一类。

（7）增强学习reinforcement learning：就是去学习决策函数与价值函数的一种方法。

智能科学要研究的是一个物理与生物混合的复杂系统。智能作为一种现象，就表现在个体与自然、社会群体的相互作用和行为过程中。我个人相信这些行为和现象必然有统一的力、相互作用、基本元素来描述。其实这些概念对我们搞计算机视觉的人来说一点也不陌生。我们的模型与物理模型是完全相通的，当你有一个概率分布，你就有了“势能函数”，就有了各种“相互作用”， 然后就有了各种“场”与“力”。

回到前面乌鸦的例子，我在第四节讨论到，我们研究的物理与生物系统有两个基本前提：

**一、智能物种与生俱来的任务与价值链条。**这是生物进化的“刚需”，动物的行为都是被各种任务驱动的，任务由价值函数决定，而后者是进化论中的phenotype landscape，通俗地说就是进化的适者生存。达尔文进化论中提出来进化这个概念，但没有给出数学描述。后来大家发现，基因突变其实就是物种在这个进化的、大时间尺度上的价值函数中的行动action。我前面那个叠衣服的价值函数地形图，就是从生物学借来的。

二、**物理环境客观的现实与因果链条。**这就是自然尺度下的物理世界与因果链条，也就是牛顿力学的东西。

说到底，人工智能要变成智能科学，它本质上必将是达尔文与牛顿这两个理论体系的统一。

第一节    现状：正视现实

第二节   未来：一只乌鸦给我们的启示

第三节    历史：从“春秋五霸”到“战国六雄”

第四节    统一：“小数据、大任务”范式与认知构架

第五节    学科一：计算视觉 --- 从“深”到“暗”

第六节    学科二：认知推理 --- 走进内心世界

第七节    学科三：语言通讯 --- 沟通的认知基础

第八节    学科四：博弈伦理 --- 获取、共享人类的价值观

第九节    学科五：机器人学 --- 构建大任务平台

第十节    学科六：机器学习 --- 学习的终极极限与“停机问题”

第十一节  总结：  智能科学 --- 牛顿与达尔文的统一