

三读《赤壁赋》，从人工智能的角度解读 “心”与“理”的平衡

朱松纯

2022/1

目录

1 引言	2
1.1 人生轨道：跃迁与升华	2
1.2 苏轼的哲思：人生的意义	3
1.3 什么是活明白	5
1.4 当今学子的困惑	7
2 初读文学《赤壁赋》：横槊赋诗	8
3 重读文学赤壁赋：清风明月	12
3.1 做学问的本质——登无人之境、享清风明月	13
4 再读文学赤壁赋：“心”与“理”的平衡	15
4.1 人生的有限与无穷	15
4.2 出世与入世：“心欲”与“天理”的平衡	17
4.2.1 鹅湖之会——理学与心学的第一次交锋	17
4.2.2 理学的“格物致知”：从数据到模型	18
4.2.3 心学：心即是理、内圣外王	21
4.2.4 心与理的关系：对立统一	23
4.3 从人工智能的角度解读“心”与“理”的平衡	24
4.3.1 U系统——外在的社会规范	24
4.3.2 V系统——内在的价值体系	27
4.3.3 U、V系统的平衡统一	28
4.3.4 在不同境遇中不断调平U、V系统	29

1 引言	2
4.3.5 “心”与“理”平衡的认知架构	31
5 结束语	32
6 本文由来	33
7 附录：在讲座上的问答摘编	34
7.1 鸣谢	37
7.2 关于朱松纯教授	37

1 引言

1.1 人生轨道：跃迁与升华

人生一世所追求的，用世俗的语言来讲，无非三个层次：

第一就是活着，这对应于马斯洛七层需求理论(hierarchical theory of needs)的最低两层物质需求——生理和安全。

第二是想活得好，对应马斯洛说的中间两层社会需求——爱与认同，尊严与名声。

第三是追求活得有意义，或者说，活得明白，这对应于马斯洛的三层高级的精神需求——好奇与理解，审美，实现自我潜能与价值。

人生的终极价值可以分为两种：无穷或有限。

- **无穷而位列不朽。** 这是指个体的价值和影响力通过一代代人不断地传递。假设种群、文化不灭的话，在无穷的时间和社会群体求和(积分)，所能达到的极限是无穷大。中国哲学思想历来不相信轮回与天堂地狱之说，从2500年前春秋时期的叔孙豹开始，就建立了“立德、立功、立言”的三不朽思想。中国传统儒家思想的人生目标就是追求不朽。
- **有限而落入黑洞。** 大部分个体都价值有限，影响力传播的时空也很有限，人离世之后，落入一个“人生黑洞”，没有任何信息能传出来。年轻的时候想起这个结局，总觉得这未免太残酷了、太恐惧了。为解决这个问题，有人提出广义的社会不朽论，是指个体可以贡献到群体的伟大事业中，实现集体(民族、国家)的不朽，也即历史是由人民群众创造的。

如果我们把这些人生的层次画出来，就有点类似物理学中的原子轨道的能级图。人生不同阶段在这些层级中跳转，主动或者被动，跃迁到新的能级轨道。积极向上的人生态度就是努力调整、升华自己的思想境界，实现更高的自我价值。

根据原子物理学模型，电子在能级跳跃过程中，能激发出光子Photon，这就是物体发光的现象。人们在能级跳跃(特别是向下)过程中，往往有感而发，也能发出五颜六色的“光”，这包括互联网上的吐槽、流行歌曲和带有审美的诗辞歌赋。本文要讨论的《赤壁赋》就是苏轼在人生大变局、能级大跳跃(从高向低)的时刻，发出的一道光芒，划破千年的时空，仍能启迪当世。

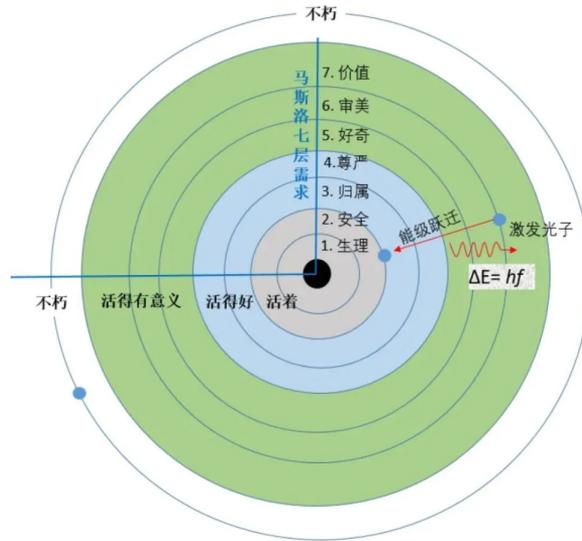


图 1: 人生轨道与能级跃迁

1.2 苏轼的哲思：人生的意义

苏轼当年从文坛才子、朝廷命官(知州，主政一方)，活得很好、活得有意义的能级，一夜之间被陷入狱，而后贬到黄州，落到他在《寒食帖》中所描绘的活着都艰难的“空庖煮寒菜，破灶烧湿苇”的困境。在这次人生向下的跃迁中，创作出了《赤壁赋》这样堪称“千古一赋”的文学名篇。黄州之后，苏轼经历了多次的波折，他的诗文反映了他在这个上下颠簸、仕途坎坷的过程中，内心的挣扎和心态的调整，从而练成了超然物外的乐观态度，成就了他“活得明白”的升华。《赤壁赋》中给世人所描绘的“清风明月”的境界，是对自然的好奇与审美，以及他对人生意义的哲思，仍然能与当今时代产生共鸣。

多年来，我一直试图读懂《赤壁赋》的内涵。随着年龄增长和阅历积累，我对《赤壁赋》在不同的人生阶段有不同的感悟。

- 少年时初读东坡赤壁，只感叹其横槊赋诗的豁达气概，开一代豪放派文风之先河；

- 而立之年重读，渐领会清风明月的意境，对应于学术人生，夜深人静之时探索无人之境的快乐；
- 已过不惑之年再读，又喟叹他对于出世与入世的平衡，和人生价值的哲学思考，且感受到文学、科学、人生的交织和互通。

夜深人静，清风明月，苏轼泛舟长江，思考人生的意义——人活一世到底能得到什么？人生如此渺小、短暂，“寄蜉蝣于天地，渺沧海之一粟”，“哀吾生之须臾，羡长江之无穷”，人生价值何在？他在《前赤壁赋》中的自问自答，以及在《后赤壁赋》中神交孤鹤与道士，反映了他当时内心的迷茫。

对于人生的意义，他还有一些想不明白、心有不甘，最后只能在含混的思维中睡去——“相与枕藉乎舟中，不知东方之既白。”我相信很多人思考过同样的问题，也就有过同样的经历：想不清楚，还是“洗洗睡吧”……

本文试图用人工智能的认知架构模型，来解读苏东坡在《赤壁赋》中的出世入世的困惑，探讨人类如何调整心态，达到“心”与“理”的平衡。有人会问：这怎么跟人工智能扯上关系了？

通用人工智能研究的目标就是要创造有自主的感知、认知、决策、学习、执行和社会协作能力，符合人类情感、伦理与道德观念的通用智能体(Agents)。

这样的智能体要有自主的意识，有三观，就算是智能体的三观与人不一样，他们也要能理解人类的三观。不同于过去2000多年来哲学家、思想家、国学家、美学家所给出的解答，研究通用人工智能的科学家必须拿出认知架构、数学模型，用数学的语言清楚地表达并能分析三观与思维的过程。所以，我们希望通过这样的架构与模型，用“重构”的方式来解读苏轼的心理活动，试图架起人工智能研究与中国古典人文哲学的沟通桥梁。

个体的思维活动与行为决策受到两大因素(模型、表达)的驱动，分别对应于中国儒家两大派别的思想：

程朱理学，注重“天理”，这包含我们今天所看到的物质世界的“物理”，对应于人类的直觉、朴素的物理模型(Intuitive physics)，和人文社会的“伦理”与社会规范(ethics and social norm)，后者对应于人工智能的行为决策函数所产生的行为模式。

陆王心学，这里可以理解为内心的欲望，暂且称作为“心”。现代的语言称作价值观。人工智能的语言称作价值函数或者功利函数。机器学习领域针对具体小任务而定义的“目标函数”、“损失函数”是其中的一种。

“理”的规范与“心”的欲望，这两者常常是不一致的，有矛盾。这个矛盾是人类作为高级动物，痛苦的根源。当人的个体能力、社会角色、地位，或者社会环境发生转

变的时候，心态的调整就是“心”的欲望与“理”的规范之间要达成一个新的平衡，重新定位自己。《赤壁赋》反映了苏轼在人生大变局，心态大调整之时的内心活动。

要真正达成“心”与“理”的平衡是很难的，孔子说：

“三十而立、四十不惑，五十知天命、六十耳顺，七十从心所欲而不逾矩”。

这里的“矩”指规矩，也就是社会规范，是天理对人的约束。这句话可以理解为，自己心中想做的与社会期望你做的，达成一致了。作为圣人的孔子，要到七十岁才能到达“天理”与“心欲”的平衡，可见这是多么困难。包括对天理的更深刻认识，心欲的消退与调整，以及对二者互为因果关系的理解。

那么，“心”与“理”这一对关系中谁是主导？这是东西方哲学的一个十分深刻的议题，也是迈向通用人工智能、给机器人树立三观的关键问题。我们在后面的章节会展开讨论，希望由此解读苏轼当年的心理活动。

从人工智能模型角度看，“理”与“心”作为两种函数，它们的相互作用就是前面提到的人生轨迹跃迁中“发光”的数理机制。心与理的不同组合，产生不同的“光”，也就体现了人的格局。

在人生轨道的跃迁中，苏轼是一个正面的典型，让我们感受到了他思想的光芒，受到后世的敬仰。历史上更多的是反面的例子，不少历史人物没有能够在变化到来之时及时做出调整，值得我们思考，这里不举例。

在进入主题之前，我先来解释一下，什么是本文所说的“活明白”。

1.3 什么是活明白

对于立志做学问的人群，以前社会上叫做“知识分子”，本来的目标应该是帮助人类活明白。下面这张图是多年前哈佛大学寄给校友的一个书签，我一直保存着，它是荷兰籍哲学家斯宾诺莎的一句名言：

“人类所能企及的最高活动就是为明白而学习，因为明白了就获得了自由。”

对于一般人来说，想不通的事情，就懒得去想，退回到活着、活得好的物质和社会能级。有更高精神抱负的人，追求思想的自由。想不明白的感觉就像是被困在一个狭小的空间，不得解脱，类似“幽闭恐惧症”。读者如果想体会这个“幽闭恐惧症”的滋味，可以读一下屈原的《天问》。

请问远古开始之时，谁将此态流传导引？
天地尚未成形之前，又从哪里得以产生？
明暗不分浑沌一片，谁能探究根本原因？
迷迷蒙蒙这种现象，怎么识别将它认清？

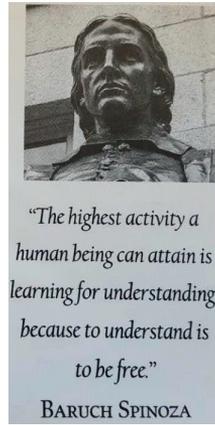


图 2: 哈佛送给校友的书签

白天光明夜晚黑暗，究竟它是为何而然？
 阴阳参合而生宇宙，哪是本体哪是演变？
 天的体制传为九重，有谁曾去环绕量度？
 这是多么大的工程，是谁开始把它建筑？
 天体轴绳系在哪里？天极不动设在哪里？

上图是《天问》白话译文的开头部分。屈原那个时代，有太多想不清楚的物理现象，当时神话传说的创世理论与模型，漏洞百出，根本无法解释他所观察到的自然现象，他想活明白一点，但是求而不得。从《天问》可以读出来，他问的都是物理学的宇宙起源等重大科学问题，可是没有科学的手段来回答，万分苦恼。后面我要谈到，活明白真的是人生到了一定层次(马斯洛第五层，好奇)的重大乐趣、收获与意义。子曰：“朝闻道，夕死可矣”，直接翻译就是：早上明白一个道理，晚上死了也值啊。不过我个人估计，孔子说的闻道，是指社会伦理，他不太关心自然科学的定律与模型。

谈到这里，顺便要说几句，以屈子为代表的楚文化，是中国的一个亚文化，历史上催生了很多优秀的、在重要历史关头敢于探索真理的人物，这明显不同于以孔子为代表的北方主体文化。

中国主体文化不提倡去讨论《天问》中那些想不明白的时空问题。比如，“六合之外，圣人存而不论；六合之内，圣人论而不议。”——《庄子·齐物论》

又比如，《论语》中，季路问事鬼神。

子曰：“未能事人，焉能事鬼？”

曰：“敢问死。”

子曰：“未知生，焉知死？”

我们今天在大学里面搞研究，就是要追根溯源，开辟新的空间Space，探究模型在时空推到无穷大与无穷小之下的特征。而孔子却直接浇灭了学生的好奇心。从这个意义上讲，孔子是优秀的人生导师，思想辅导员，却未见得是一个好的科研导师。

中国民间流传的“难得糊涂”的智慧，是有历史渊源的。很多人就不想活明白，有些人明白很多事不能探究，搞明白了反而只会带来心理上的负担，不能采取行动(actionable)，大多数时候只是徒增烦恼。所以，干脆就装没看见，装不知道。“麻痹自己”是智能体(包括人)在“心”与“理”博弈中达成一种短暂平衡态的有效办法。麻痹的方法有酒精，用化学办法阻止理性思考，也有鲁迅先生说的“精神胜利法”。我的意思是，这些心理活动的现象，都需要而且应该用人工智能的模型来解释。

1.4 当今学子的困惑

2020年底，我读到北大临床心理学博士徐凯文教授的一篇调查论文，报告称，有四成身在顶级学府的大学生出现了所谓的“时代空心病”现象，“非常优秀的年轻人，在成长过程中没有明显创伤，却感到内心空洞，感觉不到生命的意义和活着的动力”。近又看到一篇面向万名科研工作者的问卷调查，发现大约1/2的研究人员有焦虑症状。

这两份报告折射的是由应试教育竞争和功利的科研学术评估所产生的社会现象。大多数青年学生对选课和分数斤斤计较，而对于科学前沿的探索反而犹犹豫豫。学生选容易考高分的课上，青年教授选容易发文章、获高引用率的方向立题。这表面看起来是在优化自己的路径，其实质是缺乏信仰的表现。由此导致的外在表现是短视、缺乏勇气，赢在局部与眼前，输在大局与长远。

人们常常抱怨我们的科学研究缺乏原创，其实，原因大家也都看到了。我们不少科研人员在活着、活得好层面被各种评价指标牵制住了，在活明白的层面还不够。

- **从外因看：**在物质和社会世俗的层次，政府主导科研与高校，在活着、活得好这两个层级建立了各种细分的能级、阶梯，把科研人员的工资收入、晋升通道与这些指标、帽子挂钩，深入到生活的方方面面。政府参与组织的会议，排座位、活动的报道都必须严格按照这些帽子、阶梯的等级来排序。这本意是对科研的重视，但也一定程度干扰了学术的风气与秩序。
- **从内因看：**长时间的这种挂钩就形成了科研习惯，很少有人能够抗拒这些考核与评估的指标，跳出圈子，去追求科研的初心——帮助人类活明白。兴趣被浇灭，身

心俱疲之后，用一个新名词来说，赶紧找个地方“躺平了”。当大家都去追求有限的物质和社会功利目标，就聚集在一起、散不开，造成了内卷与焦虑。

当前，全国的科研机构与高校正在开展“破五唯”改革，力图破除“唯论文、唯帽子、唯职称、唯学历、唯奖项”的人才评价标准。但如何培养和认定“人才”，如何资助和评价科研“成果”，大家普遍还比较困惑。

2 初读文学《赤壁赋》：横槊赋诗

苏轼出生于四川眉州，20岁中进士，在北宋朝廷与文坛轰动一时。

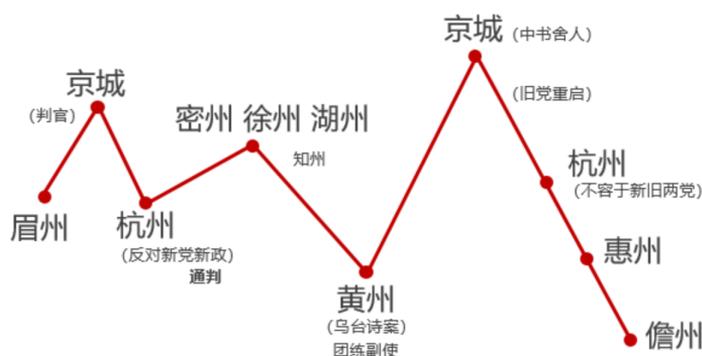


图 3: 苏轼的人生轨迹

公元1079年，42岁的苏轼任湖州知州时，遭人陷害，卷入“乌台诗案”文字狱，后经多方营救得以释放，被贬湖北黄州任团练副使(地方民兵武装部副部长)，跌落至人生低谷。从活得好、活得有意义，跌落到活着艰难的层次。到黄州后，他和家人一起开垦城东一块坡地，种地以贴家资，始称“东坡居士”。公元1082年-1084年间，他连续创作了《赤壁赋》、《念奴娇·赤壁怀古》、《后赤壁赋》、《寒食帖》等诗词与书法，在中国文学与文化史上留下了浓墨重彩的一笔。从此，中国有了苏东坡。

明代书画大家董其昌认为《赤壁赋》在文学史上的地位可与屈原的《离骚》相媲美，《寒食帖》的书法价值上可与王羲之的《兰亭序》相当，开一代文风之先河，后人难以超越：

“东坡先生此赋，楚骚之一变；此书，兰亭之一变也。宋人文字俱以此为极，则与参参，知所藏名迹虽多，知无能逾是矣。”

南宋《风月堂诗话》则认为：“东坡文章至黄州后，人莫能及”。

苏东坡一生命运多舛，在一次次人生不断转轨的过程中，他的诗词就像是能级跃迁时激发出一道道光芒，流传千年，也成就了他人生最大的价值，达到文学“不朽”的

境地。

从表面上看这些诗词，苏东坡显得很淡定，超然物外，说什么：

“竹杖芒鞋轻胜马，一蓑烟雨任平生。”

但是，他在情急悲痛之中写下的《寒食帖》暴露了他真实的境遇与心情，此时的苏轼生活窘迫、欲哭无泪。

“空庖煮寒菜，破灶烧湿苇。那知是寒食，但见乌衔纸。

君门深九重，坟墓在万里。也拟哭途穷，死灰吹不起。”

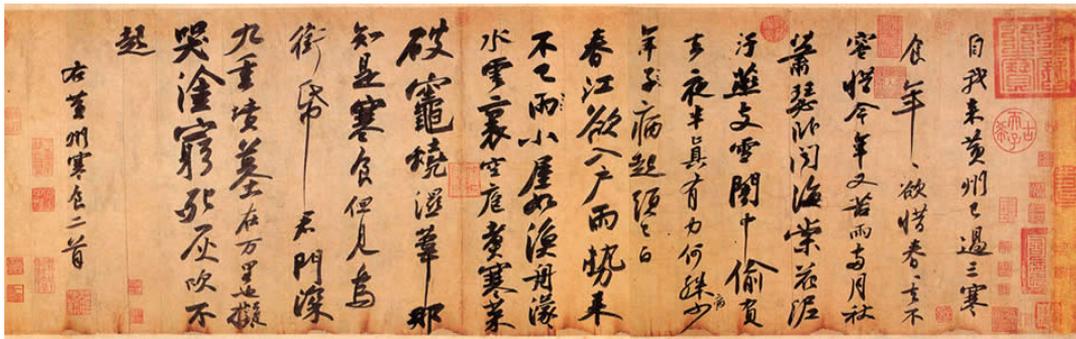


图 4: 苏轼的寒食帖（现藏于台北故宫博物院）

这个被称作“天下第三行书”的《寒食帖》，其价值之一就在于它毫无掩饰地记录了苏轼当时的状态。

说到这里，顺便讲一讲我对书法欣赏的理解。现在有些人想用人工智能来自动生成书法，用不同风格的字去训练神经网络，希望能够输出类似的字体。显然，这些神经网络不可能懂得这些字的语义、更无法体会其中的情感。几年前，一位MIT的教授来UCLA讲座，显示了两组文字(不是中文)，一组是人写的，一组是人工智能模型合成的，大家很难猜出来哪一组是合成的，这样就通过了所谓的“图灵测试”。我坐在听众席上，一眼就猜出来了。他问我怎么猜出来的。我猜的办法是，自己想象着握笔写字的过程，人是用一只手握一支笔按照一定顺序写，这里面就有次序和停顿。当然，写书法还得停下来蘸墨水。而计算机生成的时候是喷墨，更具体一点说，是最底层产生式神经元的感受域的向量的线性叠加，完全不受这个动作、墨水的限制。这两者在图像上虽然只有极其细微的差别，但是在整体的感知 + 认知表达上却相去甚远。

用我自己的模型语言来说，你看到的是一幅图像(这是信号)I，你首先得到的是这个字的一个感知的解译图(parse graph)，记作 P ，并由此重构、想象出写字的人的动作和心情，得到一个认知的解译图。

越是懂书法的人，能够体会到的认知的部分就越多，甚至可以重构苏轼当时捏笔行文的姿态，去体会他当时的心境，进入到与古人对话的状态。这些书法(感知)之外的东

西(认知), 要靠人去想象(俗称称作“脑补”), 这也正是它所蕴含的价值。越是有经验的书法家, 脑补的就会越多。我把这些人运用经验、主观想象的部分称作智能的“暗物质”(Dark Matter of AI)。

不光是书法, 人看到椅子会想象自己坐上去是否舒服, 看到杯子, 会想象手怎么去把杯子拿起来, 这些都属于认知解译图的部分。

除了书法、人工智能作诗、编曲、作画, 都存在类似的问题。如果一个智能体没有主观的价值函数、没有主观的情感、不能想象人的行为动作, 它就只能停留在浅显的图像层面 I, 这些神经网络模型甚至到不了感知表面(), 更不可能达到与人形成深层次共鸣的认知层面()。

回到苏轼和他的《赤壁赋》。和大多数人一样, 我第一次读到《念奴娇·赤壁怀古》, 是在高中的语文课本里。当时对这首词的理解, 仅仅停留在读词的内容和豪放派的文风。老师说, 苏轼应该知道赤壁之战并不是发生在黄州, 现在这里直接就叫“东坡赤壁”了。见图右上角的确有一个赤色的岩面, 后来长江改道, 不从这里经过, 变成了一个湖汉子。



图 5: 本文作者2020年9月回国后, 再访松风阁

我的家乡在湖北鄂州, 三国时期孙权在此建都, 称作武昌(用意是因武而昌, 是武昌鱼的原产地), 与黄州隔江相望。地处长江中下游, 江面开阔, 水流平缓。完全没有诗词中描绘的:

“乱石穿空, 惊涛拍岸, 卷起千堆雪”

这样的气势恢弘的大场面。后来有很多书画大家根据苏轼的文章来作画，产生了一些气势磅礴的赤壁场景，其实都是不符合地理情况的想象。

当地唯一的山是鄂州的西山(樊山)，上图的左下角照片是我站在山顶的亭子上拍摄的，背后的江面和整个黄州十分平坦。西山上有一个很有名的亭子“松风阁”(上图的右下角)，因宋代著名书法家黄庭坚的《松风阁诗帖》而闻名。1102年，黄庭坚在此过夜，听松涛怀念亦师亦友的苏轼，当时苏轼已经过世。

苏东坡两次夜游赤壁，写成《前赤壁赋》和《后赤壁赋》，就被供奉在赤壁公园的二赋堂，正面是《前赤壁赋》，背面写着《后赤壁赋》。虽然说《念奴娇·赤壁怀古》很容易理解，但是《赤壁赋》对于年轻人来说是很难读懂的，缺乏人生阅历，就无法体会苏轼的心境。这一点就像前面提到的对于《寒食帖》书法的欣赏，需要根据自己的经历来脑补这其中的。



图 6: 赤壁公园的二赋堂

总体来说，《赤壁赋》给年轻的我留下了两个思考的问题：

1. 作为一个文人，苏轼为什么在此时此地多次念念不忘曹操、周瑜？
2. 怎么有人能够凭借几百个字，就能流芳千古，达到不朽？

相比之下，我们现在写的论文，往往3年之后就没人再读了，更有甚者，就算是网

络上轰动一时的热搜，一周、一月之后，就再无人问津，更何况是千年之后呢。这太神奇了。

前文提到我小时候对于人生黑洞的恐惧，《赤壁赋》算是给我指了一条出路，看到了人生价值的一线希望。大概这就是人文之乡风对青年学子的影响了吧。



图 7: 东坡像前留影

作者于1986年8月，接到中国科技大学的录取通知后，赴合肥之前游黄州赤壁时照的老照片，背后是苏东坡像。

3 重读文学赤壁赋：清风明月

等我过了而立之年，读完博士，2002年当上了终身教授，重读《赤壁赋》，开始领悟到两首《赤壁赋》所描绘的清风明月意境之妙。

“壬戌之秋，七月既望，苏子与客泛舟游于赤壁之下。清风徐来，水波不兴。举酒属客，诵明月之诗，歌窈窕之章。少焉，月出于东山之上，徘徊于斗牛之间。白露横江，水光接天。”——《前赤壁赋》 1082年农历7月16

“携酒与鱼，复游于赤壁之下。江流有声，断岸千尺；山高月小，水落石出。予乃摄衣而上，履巉岩，披蒙茸，踞虎豹，登虬龙，攀栖鹘之危巢，俯冯夷之幽宫。划然长啸，草木震动，山鸣谷应，风起水涌。”——《后赤壁赋》 1082年农历10月15



图 8: 苏子夜游赤壁图，画作出自吉林艺术学院温国良教授

苏轼的行为，就算放在今天，也已经是离经叛道了。他已经44岁了，挑的日子是农历七月十六，满月之时，半夜抱着酒坛子和鱼，泛舟长江之上，在江中心烧烤，又是吹箫、朗诵诗歌，又是登岸、攀爬岩石，而后长啸之声划破长江之上寂静的夜空，最后还在江中睡到天亮。

近年来回老家的时候，我也几次想去体验一下，驾一叶扁舟夜游长江，可惜周围无人附和，大家都劝阻。在中国，要行离经叛道之事特别困难。

我上大学的时候，去合肥的交通十分不便，坐大轮、顺江而下，到安庆上岸转长途汽车，或者到芜湖转火车。大轮在长江上缓缓而行，要走一个晚上。我站在甲板上，皓月当空，看两岸山峦起伏，思考自己迷茫的人生道路，轮船的汽笛声划破夜空，更是让人万分惆怅。完全没有苏轼那种潇洒的、诗情画意的场面，更没有他那种横槊赋诗的豪迈气概。

3.1 做学问的本质——登无人之境、享清风明月

苏轼从朝廷大臣、文坛名士变成一介民夫，放浪山水，“渔樵于江渚之上，侣鱼虾而友麋鹿”，突破传统文人的思想和行为的束缚，进入一个崭新的思想自由的空间，从

而开创出被后世士大夫阶层所推崇的“清风明月”的境界，达到了他自己和中国文学创作的顶峰，成为永恒。中国近代文学家林纾(1852-1924)评论《赤壁赋》：“奇妙无以复加，易时不能再作”。

我对苏轼的清风明月意境，有两重理解：

1. 做科研的本质是在夜深人静之时，探索前沿，登无人之境。

用数学语言来讲，科学研究本质上就是找到新的空间(Space)，这里有它独特的元素、结构、测度距离和美学，这个空间可以是一个抽象的数学空间，也可以是一个图像空间，几何形状的空间，或者感知、认知的空间。科学研究的目标就是探索与领略这种空间的结构、奥妙，把它描绘出来。我的导师之前研究代数几何，是一位菲尔兹奖获得者，他介绍自己数学研究的经历：“就像找到一扇门，打开以后，进入一个花园，然后把这个花园的结构理清楚，展示给世人”。

苏轼是历史上第一个描绘这个赤壁场景(清风、明月、江流、断岸、吹箫、吟诗…)的文人，如果还有后人效仿，就没有新意了。

2. 科研人员能获得的往往只是享受清风明月的过程与经历。

苏轼在《前赤壁赋》的最后一段交代：他在江面上想明白了一件事，人一辈子到底能得到什么呢？

“惟江上之清风，与山间之明月，耳得之而为声，目遇之而成色，取之无禁，用之不竭，是造物者之无尽藏也，而吾与子之所共适。”

虽然说，现在不少科研人员可以拿到技术发明专利、开公司挣钱，但对于探索基础和前沿的科学家，他们更多的是享受这个“清风明月”的经历和过程。话说回来，人们挣了钱，然后花钱旅游、坐过山车、看电影、去山上滑雪，其实买的也只不过是一种经历。科研人员如果经历了清风明月这个过程，那本身就是享受了，不是吗？更何况这是“取之无禁，用之不竭”的自然馈赠。

我以前常常在实验室待到很晚，然后开车回家。在空旷的高速路上，头顶着一轮皓月，打开音乐，一路狂奔，特别是调通了代码、想通难题之后，心情舒畅，特别享受。下面我再举一个例子，解释一下科研的享受在于过程与经历，而不是结果。

我以前跟我的学生分享过读海明威《老人与海》的感悟。这篇小说很朴实，人们把它当作一个寓言故事，不同背景的人，会读出不同的感悟。主角是一个贫穷的、生活在古巴的老人(这里我们把他寓意为一个缺乏经费的科研人员)，靠一艘小船在大海打鱼为生(科研探索)。一次偶然的机，他钓上了一条巨大的鱼，被鱼拖到了深海(无人之境)，经过日夜的搏斗(加班加点的科技攻关)，终于制服了大鱼(巨大的科研发现)。但是，他的船太小了(平台、队伍、资源都有限)，装不下这条大鱼，只能把鱼放在水中拖(文章没有发表，或者是只能发表在小众的期刊上)。结果，鱼的血腥引来了一群鲨鱼(商业开发、大资本)，沿途把鱼肉吃了个干净。精疲力竭的老人终于到达岸边，把一个巨大的骨架拖上海滩。天亮了，岸边路过的人们，看到这个骨架，议论纷纷。

这个老人虽然没有吃到鱼，但他获得了这次打大鱼的经历。探索科学前沿的科学家大抵如此。现代社会能够为有创新成就的科学家提供越来越好的生活环境，而作为有理想、有情怀的科学家，活得好不是主要目标，应该升华到活明白，追求做出伟大的贡献。

我在2017年曾经写过《学术人生：文章千古事，得失寸心知》一文，当时在这篇文章中我通过“爱因斯坦评论做科研的三种动机”以及“科学研究的三个时期与战略定力”两小节，讲述了做科研的本质与定位问题，进一步阐述了“清风明月”是一种高级的审美与价值选择。这里就不重复了。

现实世界中，真正醉心于“清风明月”的科学家非常少，至于是否获奖、是否有很高的引用并不重要，著书立说、成就经典才是对这些学者的褒奖。唯有如此，才能够沉潜于学术研究、深耕于各自领域，化幽壑之潜蛟，待大浪淘尽，已斩波千里。

4 再读文学赤壁赋：“心”与“理”的平衡

过了不惑之年，再读《赤壁赋》，就能够体会到苏轼在长江之上的哲学思考了。

第一节开篇提到过，个体的行为同时受到天理与心欲的驱动，这两者的不一致是导致人们内心痛苦的根源。当人的社会角色或者社会势态发生转变的时候，心欲与天理需要达成一个新的平衡，不是每个人都能适应这个转变、达成这个平衡。《赤壁赋》反映了苏轼在变局之中，心态调整之时的内心活动。他纠结于两大问题：

- **入世和出世的平衡：**当年苏轼已经44岁了，人生下半场是积极奋进(好像又没有了平台和机会)? 还是随遇而安，继续发明各种东坡美食?
- **有限与无穷的问题：**人生终极价值如何结算、定论? 这是所谓的中年危机，每个想活明白的人到了中年都要面对的问题。

下面，我对这两个问题做一个解读。

4.1 人生的有限与无穷

苏轼在《赤壁赋》中，对有限与无穷的思考贯穿始终。“寄蜉蝣于天地，渺沧海之一粟”是诗人感叹个体的渺小。“哀吾生之须臾，羡长江之无穷”是诗人感叹个体生命的有限，羡慕自然之永恒。

从博弈论的角度来讲，针对有限轮次的博弈，与无穷次博弈，博弈的策略是完全不同的。如果我们认识到人生的短暂与渺小，是不是就该及时行乐，那么奋斗的意义何在? 所以，他想到了曹操当年：

“方其破荆州，下江陵，顺流而东也，舳舻千里，旌旗蔽空，酹酒临江，横槊赋诗，固一世之雄也，而今安在哉？”

又想到了《念奴娇赤壁怀古》中的周瑜：

“雄姿英发，羽扇纶巾，谈笑间，檣櫓灰飞烟灭。”

人生的有限与无穷是一个困惑人类的根本问题。

在西方文明的源头，以苏格拉底、柏拉图、亚里士多德为代表的古希腊哲学家，主要聚焦在知识的表达与逻辑的推理，这在后来发展成为数理逻辑、形式语言，构成了人工智能的前30年的主要理论框架。但是，这一套知识体系到了雅典后期，就让位于基督教的宗教故事了，因为宗教对这个困扰人类的问题给出了一个解答。新约中记载，耶稣门徒St. Paul到访雅典，与希腊的哲学家展开辩论，最后宗教获胜（获得大众的追随）。基督教的上帝和天堂的信仰，对于困扰人类的终极问题至少是给出了一个答案。

新约Clement福音中明确提出：“human philosophies can reach partial truths, but these philosophies must be assisted by faith.”他们意识到，人类哲学（特别是讲求实证的科学）永远只能是触达部分的真理，剩下的部分要由信仰来协助回答。他们认为这种信仰就是一种道德的纯洁化(moral purification)。

在东方思想中，我们不相信鬼神，也不相信什么来世今生。有人会问，佛教不是也有六道轮回的思想吗？据我的一位熟悉佛学的博士生告诉我，其实这是世俗社会对佛经的一个长期的误解。佛学中的“三世今生”，“六道轮回”本来都是代表人的思想状态、心理状态的变化、跳转，通过比喻的方式来讲解这些状态。结果世俗的人就把比喻当成真的了。比如，六道中：

“饿鬼”是指人处于物欲贪婪的心态中；

“阿修罗”是指人处于自大自负的状态中、用今天的话来讲，个人膨胀的心态；

“畜生”是指人处于非理性的、冲动的状态；

“地狱”是指人处在痛苦、无法解脱的状态，等等。

佛学修行认识到，人处于不同的心态下，决策函数与价值函数有所不同。

绝大部分中国人没有生死轮回、上帝天堂等宗教信仰，尤其在儒学思想自汉代确立了主导地位后，功名利禄成为了社会精英的主要追求目标。前面提到过，孔子反对讨论那些想不明白的时空与起源问题，提倡“圣人存而不论”。儒学在后世的发展中，也没有对生死和无穷进行解释，而是强调当世的社会秩序与个体行为规则。

从现代科学的角度解释，生命体的无穷可以体现在几个方面：

生物信息：个体基因在人群中的流传。

物理信息：人的活动在物理世界留下、除了地球上的痕迹，还有图像通过光发送到宇宙中，若干光年之外的星系也许能在若干年后看到。

智能信息：思想、行为、价值取向等等，比如，《赤壁赋》在人世间的传播。未来，随着人工智能的进步，可能复制一个虚拟的个体，从而实现人的永生，当然，虚拟个体

是有了，如果后世无人访问，那还是没有作用和意义。

这三者之中，达到无穷最靠谱的办法应该还是第三种：智能信息的无穷传播，以及对人类整体价值的提升。

接下来，我们就来到了本文的核心部分，由下面的4.2节和 4.3节阐述“心欲”与“天理”的平衡，其中4.2节从哲学的角度介绍这个概念，4.3节从人工智能模型的角度给予进一步的阐述。

4.2 出世与入世：“心欲”与“天理”的平衡

出世入世之说源自佛学，不同佛学派别有不同的解读，本文采用世俗的说法。入世指积极的人生态度：主观上愿意维持更多的社会关系，承担更多的责任、力求有所作为。出世指消极的态度：回避某些社会关系与责任，挣脱群体与社会规范的约束，过清静悠闲的日子。历史上一个有名的例子是东晋时期的陶渊明(365-427)。

那么到底是要出世还是入世呢？儒家提倡所谓的“穷则独善其身，达则兼济天下。”《赤壁赋》反映了苏轼在人生变局中，对于“入世”的无奈，和对于“出世”的不甘。他和陶渊明的价值取向是非常不同的，苏对陶有很高的评价，但他毕竟没有把黄州东坡当作桃花源。陶渊明更多是关注自己个体生命的体验，在官僚体制中受挫之会，心灰意冷，开启自得其乐的田园生活。苏轼虽然蒙冤入狱，被贬到一个羞辱他智商与才华的位置上，但他没有怨天尤人，也没有消极避世。《赤壁赋》对于人生意义的思考比《归去来兮辞》要积极、深远，这也许是因为苏轼受到后世更多的敬仰。

前面提到过，个体的行为同时受到天理与心欲的驱动，心态的调整本质上就是“心”的欲望与“理”的约束之间的平衡。这是东西方哲学讨论的深刻问题，也是通用人工智能需要解决的重要问题。

下面介绍一下这个哲学讨论的主要概念。

4.2.1 鹅湖之会——理学与心学的第一次交锋

在《赤壁赋》著成93年后，南宋1175年，在江西省一个叫做鹅湖的小镇，记载了中国思想史上一次重要的辩论。史称“鹅湖之会”，由理学代表朱熹和心学代表陆九渊围绕“理”与“心”的关系展开了一场辩论。

首先要说一点，在10年前，作为一个理工科的科研人员，我对所谓什么理学、心学是不以为然的，看不懂，感觉莫名其妙。但是，自从开始考虑通用人工智能的大一统之后，我才意识到这是多么核心的问题。

我认为“鹅湖之会”才是中国历史上真正的“华山论剑”。顺便提一句，金庸小说《射雕英雄传》的故事背景是南宋，经武侠迷考证，金庸小说的第一次华山论剑发生在1199年，也就是同一个时间点。



图 9: 鹅湖之会 1175年(网络图片)

儒家思想从一开始就制定了“三纲、五常、八端”等基本社会伦理规范与礼仪制度，以及“格物、致知、诚意、正心、修身、齐家、治国、平天下”的人生入世的路线图。到了南宋时期，由于外来的佛学有更宏观、完整的理论体系，唐代的禅宗六祖慧能又把佛学与中国社会实践相结合，广泛传播。传统的儒学已经势微了，儒学家从程颢、程颐开始了革新运动，希望重振儒学，这而后分出两派：理学与心学。

缺乏这些哲学概念的支撑，特别是完全没有现代数学的工具，苏轼在1082年很难想明白入世出世的根本问题。

4.2.2 理学的“格物致知”：从数据到模型

理学的代表人物朱熹(1130-1200)主张“理”是自然万物和人类社会的根本法则、根本规律，世间万物包括人在内都需要按照天理来约束自身的行为。我们后世讲的自然科学，如物理、生理；社会科学，如伦理ethics和社会规范(social norm)，都应该属于朱熹所讲的天理的一部分。

那怎么认识天理呢？朱熹开出的处方是儒家的方法论——“格物致知”。仔细观察事物，相当于数据收集，然后提炼出知识，也就是今天的数理模型。格物致知本质就是从数据到模型的知识发现过程，就是今天人工智能所讲的：**Knowledge Discovery: From Data to Model**。

可是，怎么实现这个目标呢？朱熹当年不可能提出统计模型、机器学习的算法，所以，他的学生相当困惑，包括之后明代的王阳明(1472-1529)也抱怨，他没法通过格物，

比如肉眼观察竹子的生长，而得到竹子的理(模型)。但要根据此否定格物致知的思想，是完全没有道理的。朱熹只不过是远远超前于他的时代了。

下面我来举一个格物致知的独特例子，非常有历史意义，也说明中国800多年前的哲学思想与现代人工智能是直接关联的。

有一次，一名学生问朱熹到底什么是“理”，朱熹被逼到墙角了，指着谈话的桌面说，你看树的纹就有理，这就是所谓的纹理(英文texture)。当年看到这段对话，我不禁拍案叫绝——因为这是计算机视觉乃至现代机器学习的一个根源、发端的问题。

我来解释一下这句话的意思，比如杨树的每一块木板花纹都不一样，但是我们都知这是杨树(下图左)；同样，柳树的木板，每一块不同，但是，我们一看就知道属于柳树(下图右)。

我们今天铺地板的大理石，每一个方块都不同，但我们看起来是同一类。

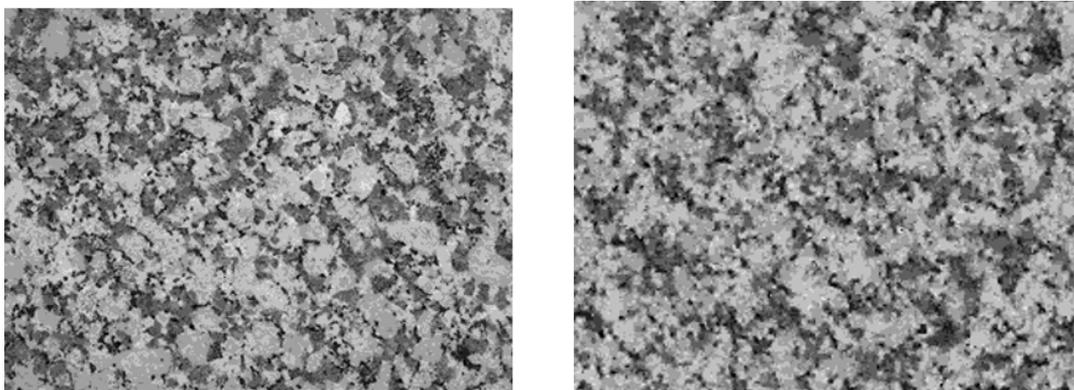


图 10: 两块大理石

那么，到底是什么“理”(模型)来判断两块“纹”属于同一类?

数学上讲，杨、柳、大理石，这样的概念 C 分别就对应于一个图像 I 的集合，是一个等价类:

我们可以认为，人脑必定是从图像中提取了某些特征与统计量，离散化之后，表达成为一个高维的向量。同一个等价类的图像拥有一个共同的向量数值。

朱熹在1160年左右被问到而回答不清的问题，800年后，到1960年被贝尔实验室的一位著名的心理物理学家Julesz重新提出。后来我把这个叫做 Julesz 之问(Jules question):

“是什么图像特征与统计量，使得人眼认为两幅图像的花纹看起来是一样的?”

这个问题把视觉问题转化成为一个数学问题，Julesz开启了一系列的心理物理学的实验，这些实验为计算机视觉的创始者David Marr(马尔)在1970年代末开启视觉计算理论提供了重要依据。1980年代末，以数理逻辑为基础的人工智能走到了道路的尽头，人

工智能进入所谓的寒冬期。而这个时候，计算机视觉等学科得以发展，统计建模的理论悄然开始了。而纹理texture的模型确实是统计建模理论里最早被攻克的问题。

由于纹理是所有物体都有的表观特征，80-90年代计算机视觉的很多学者都研究过这个问题。到了1995-2000年，这个问题被我和合作者用一系列文章较为完整地回答了，这也是我在哈佛大学博士论文的一部分。

上面的大理石图像，其实是1996年我博士论文的一个实验结果，左边是观察到的一幅大理石纹理图片，右边是计算机根据统计模型用马尔可夫链蒙特卡洛(Markov Chain Monte Carlo)随机游走法从等价类中采样生成的结果。如果人眼看出来是一样的，就验证了我们找到了这类纹理最大的共同特征和统计量。下面这张图是更多的结果：左边是观察的水波图像，由此建立一个模型(纹理)，由模型自动采样生成中间和右边的两幅图像作为验证。

这个模型的取得综合了神经科学、心理学、统计物理学的很多知识，我这里不宜展开说了。纹理最后就表达成为这么一个概率模型——这就是朱熹当年所需要的“纹理”！

$$p(I; \lambda) = \frac{1}{Z(\lambda)} \exp \left[\sum_{k=1}^K \sum_{x \in D} \lambda_{k,x} ([F_{k*}I](x)) \right] \quad (1)$$

这个模型里面，F 代表了一个神经元，F*I 是神经元从图像提取的特征。在这个表观模型之上，我们可以进一步建立各种高阶的组成式(compositional)模型，来表达物体的构成、生活场景的布局、人的行为和社会活动的规范。这都是现代计算机视觉、认知科学研究的范畴。

儒家所提出的人的基本伦理、社会规范(Social Norm)，都可以总结成为概率模型。如果我们把人看作社会大系统中的一个个个体，其运动和行为就要符合这些模型(天理)，如此一来，儒家的“八端”：“孝、悌、忠、信、礼、义、廉、耻”对人的行为的规范和约束作用，便与物理学的各种相互作用与力是类似的，最后，都可以表达成为一个综合的势能函数 U，对应于上面纹理模型数学公式的指数项，我将在下文介绍这个势能函数 U。

儒家思想认为，如果所有自然世界和人类社会都按照天理，即势能函数U来运行，那么社会就和谐平稳了，世界就太平了。孔子的理想就是要建立规范，恢复周朝的社会安定局面。

那么，有没有“捣乱”的因素呢？有！那就是人心。朱熹认为，社会之所以“不治”，就是因为人不合理的欲望太强，导致礼崩乐坏、物欲横流。他进一步提出要“存天理、灭人欲”。这就把理与心对立起来看待了。

4.2.3 心学：心即是理、内圣外王

与“天理”对应的是“心欲”，这是人的各种物质和精神层面的诉求，包括第一节讲的七层需求：生理、安全、爱与认同、尊严与名声、好奇与理解、审美、实现自我潜能与价值，代表了人类对美好生活的向往与追求，是个体主观能动性，也是社会发展的驱动力。我们可以认为，这是人类在进化的过程中受到环境、生物属性、激素等共同的影响进化出来的价值体系总和，我把它们记作一个**V系统**。与之相对应的是理学的各种势能函数的总和，称作**U系统**(包含了客观物理定律和社会规范)。

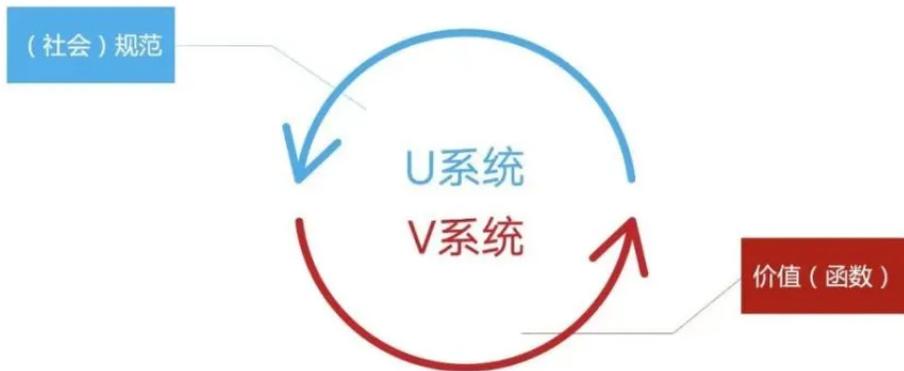


图 11: U系统和V系统

陆九渊认为朱熹的“格物致知”道路，导致了外在的客观规律与社会规范 U系统对主观个体内心追求 V系统的压制。他提出了“自吾心达宇宙”的外化道路，将“心”作为个体与所处世界、社会关系的出发点。提出了与朱熹理学不同的修行路径：首先要启发内心，从初心出发，再做到知行合一，强调“自作主宰”，个体按照内心的规则行事。后来，陆九渊进一步主张“心即理，心外无物”。这是一个受到批判的唯心主义说法。我们不一定要主张“唯心”，但要明白“心”的确是智能的主体。我从研究人工智能的角度解读，“心即是理”这个论断的意思应该是：

从系统(价值体系)完全可以推导出U系统(社会规范)。

这是一个十分深刻的认识，也是当前人工智能发展的一个哲学转折点。智能体(包括人、人工智能等)都不能被物化成为一个物理系统的元素，而应该由主观的价值体系来主导，而社会规范是第二性的、处于从属的地位。

到了明代，王阳明继续发展了“心学”的理论体系。王阳明曾困顿于龙场，身处艰困的环境使他深刻反思，并设问圣人处此境况将何以自处？也就是在理的规范还没有涉及到的新的情况下(novel situations)该怎么办？内心日夜反省，某半夜，王阳明忽然有

所顿悟：“始知圣人之道，吾性自足，向之求理于事物之外者，误也！”王阳明自此领悟到“心”才是感应万事万物的根本，由此提出心即理的命题，这也解决了此前他对朱熹格物致知说的疑惑，史称“龙场悟道”。借由阳明心学以及其他贡献，后来，王阳明实现了“立德、立功、立言”三不朽，成为了儒家所言的“圣人”、“内圣外王”的典范。

其实这个思想，早在禅宗六祖慧能(638-713)的《坛经》中已经有所暗示了。《坛经》的核心思想是十六个字：

“菩提自性，本自清净，但用此心，直了成佛。”

这里说的“自性”与“心”范围更广，它包含了前面的价值体现V系统之外，还指人的认知思维的架构与机制(见本文 4.3.4 节中智能体的认知架构图)。

佛是指觉悟的人(泛指“有情众生”，应该也包括觉悟了的人工智能系统)，通过参悟内心，可以了悟社会伦理。

慧能对于心在日常生活中的感知、认知、推理的重要的、主导作用，就体现在他出场的一句可堪惊世骇俗的言论：

“不是风动，也不是幡动，仁者心动。”

智能体首先要“心动”，才能看到“幡动”，心是主体。我们每天走在大街上，对绝大多数的人、车、物体的行为都视而不见。我们看到的只是与我们当前任务(开车、走路、找某个地点)相关联的那些部分。当然，一切任务又是被心所驱动的，是为了实现价值体系在某些维度的增值。

‘理’与‘心’的关系



程朱理学

朱熹 1130 – 1200年

- 理或天理是自然万物和人类社会的根本法则
- 物、人各自之理都源于天理
- 存天理，灭人欲
- 格物致知，知先后后
- 由外而内



陆王心学

陆九渊 1139 – 1193年

王阳明 1472-1529

- 心即理，心外无物，心本身包含着义理，伦理根植于心中
- 致良知，人性本善，应当发明本心
- 知行合一 心即是理
- 由内而外

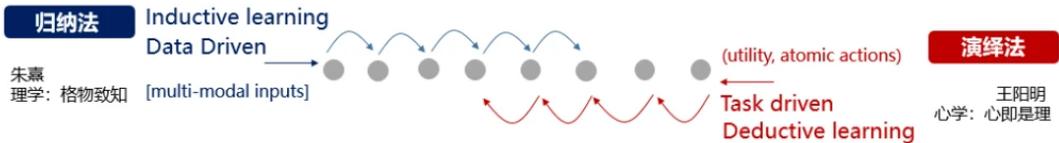


图 12: “理”与“心”的关系

4.2.4 心与理的关系：对立统一

不难看出，“程朱理学”和“陆王心学”存在着对立之处，比如“程朱理学”提倡理或天理是自然万物和人类社会的根本法则，物、人各自之理都源于天理；而“陆王心学”则提倡“心即是理，心无外物，心本身包含着理，伦理根植于心中”。“理学”认为要“存天理，灭人欲”，而“陆王心学”则提倡“人性本善，应该发明本心”。“程朱理学”以格物致知为方法论，知先后后，“陆王心学”则讲究从心出发，致良知(V系统)，做到知行合一。

我把这两者的路径画中在上图中，分别用红色和蓝色的箭头代表其演进的方向。“理学”是从由外(数据)而内(模型)的归纳法，鹅湖之会时，朱熹主张先博览而后归之于约，认为陆九渊的教法太简易；“心学”则更遵从内心，是由内(价值)而外的演绎法，陆九渊主张先发明人的本心而后使之博览，认为朱熹的教法过于支离。

这两者必须相互作用，最后达成动态的平衡，而牵引这个动态平衡不断移动的，是V系统。总之，“理”和“心”在本质上应该是对立统一的，“心”应是主体。

注意：与物理学研究的纯客观模型不同的是，人工智能的模型(理)从一开始就与“心”密不可分。比如在 Julesz 之问中(见本文第4.2.3节)，两个纹理是否属于一类，其判断的标准是人眼，这就是主观的(受到任务和价值主导)，而非纯客观的！

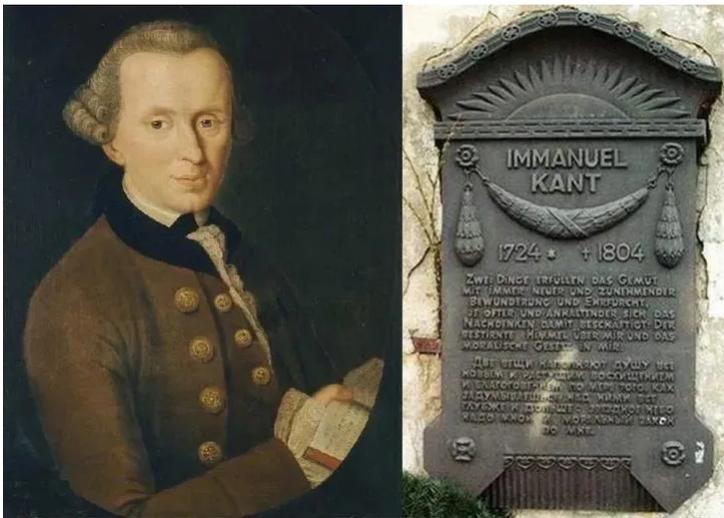


图 13: 德国古典哲学创始人康德(1724-1804)

西方哲学同样探讨“心”与“理”的对立统一性。18世纪，德国古典哲学创始人康德(1724-1804)在《实践理性批判》一书中提出，一个有理性者的道德行为必须是自律的而不是他律的。所谓实践理性，指的是主观意志居首要地位，理性居从属地位，离开意

志，理性不能存在。

“心”与“理”的思考伴随了康德的一生，他的墓志铭中写道：

“有两种东西，我对它们的思考越是深沉和持久，它们在我心灵中唤起的惊奇和敬畏就会越来越历久弥新，一个是我们头顶的星空，另一个就是我们心中的道德法则。”

“the starry skies above me and the moral law inside me.”

到了康德时期，牛顿(1643-1727)已经用力学定律和万有引力理论把天上和地下物体的运动都统一了(也就是势能函数的U系统)。这对于宗教信仰是一个巨大冲击，也给康德造成了以上的困惑。在我看来，“头上的星空”代表了按照“天理”来运行的客观世界(U系统)，“心中的道德”代表了人的价值体系(V系统)。

由此看来，东西方哲学关于“心”和“理”的见解、思考、困惑，在本质上应该是相通的。最重要的是认识到两者之间的对立统一，达成平衡。这个平衡就是孔子所描绘的“从心所欲而不逾矩”，也是苏轼在赤壁赋中思考而不得的平衡！

那么，这个平衡是什么？该如何到达这个平衡呢？

4.3 从人工智能的角度解读“心”与“理”的平衡

作为一名长期从事人工智能领域的研究者，近年来，我愈发认识到人工智能的研究与人文社科哲学密不可分。从哲学层面看，我认为人工智能的发展大致可以分成三个时期。

人工智能的第一时期，1960-1990年，逻辑的表达与推理是主要的理论框架，这得益于西方哲学的源头——古希腊文明，以苏格拉底、柏拉图、亚里士多德为代表的辩论与逻辑，后来发展成为严密的命题逻辑、谓词逻辑，事件逻辑等。

人工智能的第二时期，1990-2020年(及今后若干年)，概率建模与随机计算占据了主导地位，这对应于儒学的格物致知，从数据到模型，前面已经谈到了。

人工智能的第三时期，也就是未来，需要从理学过渡到心学，由心驱动，实现理与心的动态平衡。让我欣慰的是，东方哲学与智慧，能为人工智能后半场的发展提供哲学层面的指导。

需要明确的是，未来人工智能的发展与人文社科哲学的研究是互通的。无论是以程朱理学、陆王心学为代表的东方哲学，还是以康德为代表的西方哲学，均试图探求天理和人欲的平衡问题，这个问题也是人工智能要追求的境界。

4.3.1 U系统——外在的社会规范

在我们生活的社会系统里面，假设每个人的状态都可以用向量来表达。那么，社会总体(N个人)状态则是一个由众多个体状态组成的高维向量，即。其中表示我(当事

人)的状态，则表示除去我之外其他人的状态。在社会中，我们会跟许多人产生社会关系，比如父母、子女、配偶、同事、邻居等，这些个体用表示。

马克思说“在其现实性上，人的本质是一切社会关系的总和”。这是一个非常透彻的论断。就个体而言，每一对关系都可以由一个势能函数表示，代表了某种社会规范，如配偶、君臣、师生等。每个U都可以看作是一个“弹簧”，代表一种约束与相互作用。某个维度上“弹簧”的作用，是在该维度上保持一种动态的平衡：如果二者距离过小，会导致推力(合力)凸显；如果距离过大，会导致拉力(合力)凸显。这里我们只说二元关系，多元和高阶的关系也可以有函数表达。

既然与之间存在约束，那么在这个维度上“我”跟“他”发生了作用关系，相当于他对“我”施加了一个力，力可以通过对这个势能函数求导计算得出，即

因为人是社会动物，个体在社会中会接受来自各个方向的力。在社会空间中，其他的人要么是拉着“我”，对我产生促进、激发、鼓励的作用；要么是推着“我”在走，对我产生抑制、压制、障碍的作用。就个体而言，这两种力的作用方向、大小各不相同，这二者的关系在数学上表现为合力，而个体的行为 and 选择正是受到合力的影响。

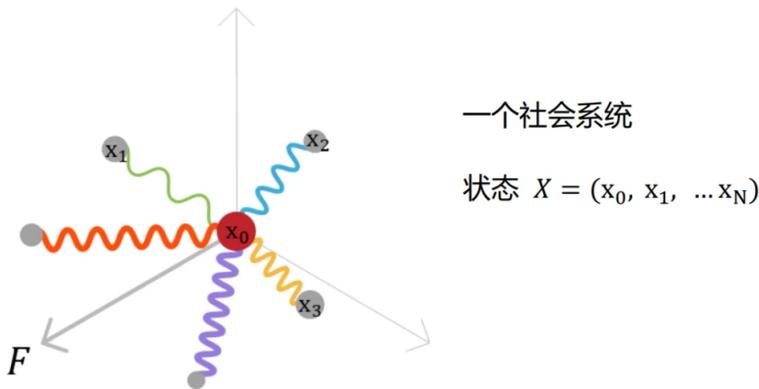


图 14: 一个社会系统的状态

一个人在整个社会空间中受到的各个方向的合力 F ，可以表达为下式。

$$F = - \sum_{i \in \Omega_0} \frac{\partial U(x_0, x_i)}{\partial x_0} \quad (2)$$

其中是一个状态向量，代表一个时空因果的解译图，可以有复杂的高阶相互作用，弹簧只是简化的二元关系，是个体(我)的社会关系的集合。

U系统表达的是社会规范(社会要求你怎么做)、社会秩序(社会层级结构是什么)、社会期待(社会关系对你行为的期待)。该系统是在长期的社会实践、规则演化、多次博弈等过程中所最终形成的一种稳定的社会规制性力量。人只要在社会之中，就会受到各种

社会规范、社会秩序、社会期待即U系统的约束。

中国儒家思想对U系统定义的非常清楚，其主体思想是教导个体在社会中如何生存。所谓“君君臣臣，父父子子”就是一种经典的U系统表述，刻画了君臣(事业)、父子(家庭)两种核心社会关系。君臣之间是一种非对称的关系，也就是：

在现代社会的职业场域，君臣关系也可以延伸成领导-被领导关系。在家庭场域，父与子也是一种规范性力量。除了职场、家庭，还存在亲戚、朋友、老乡等其他社会关系。他们对个体也存在规范、制约性作用。因此，人的负担就非常重，通俗地理解为“活得很累”。所谓活得很累，就是你身边有很多个社会关系、很多个弹簧拉着“你”，或者说，你需要考虑多种社会关系。

可见，社会关系具有规范、制约、引导个体行为的作用。一个人的社会关系越多、关系的作用范围越大，则关系在社会中的地位与影响越“重要”。但与此同时，人能够满足别人的需求越多，能够帮助别人的越多，也就越有成就感，也就是所谓积极地“入世”。儒家的“入世”思想认为，人要接受社会规范来指导自己的行为，比如三纲五常。当绝大多数人都遵从这样的规范，社会总体的运转就是和谐的，整个社会与国家就会呈现出一种秩序与稳态。当然，总是有很多人不甘心这个稳态。

与儒家不同，佛学提供了另外一种个体在社会生存的基本解决方案。

首先，在佛学里面，社会关系被统称为“缘分”。这个“缘分”不是永恒的，而是只在一个很短的时空内存在，佛教把这个现象称作为“空”。佛学有一个很强大的口号，叫做：

“缘起性空”

我的解读是：概念起源于“缘”，当木匠把四条腿和一个面板钉在一起，就产生了一个“桌子”。当有人把腿打碎了、拆开了，“桌子”也就不存在了。所以，桌子本性是空的。物理关系如此，社会关系也如此。这里我就不展开说了。

“出世”就是斩断某些“缘分”，使人不再受这些世俗关系的影响，如“尘缘已了”、“斩断尘缘”等说法。当一个人出家，就宣告了他与之前的很多社会关系断绝了。佛教为其重新定义了社会观、世界观，即重新定义了一种U系统。例如，他自己会有一个新的名字(佛家法号)，他称呼所有人都是“施主”，连他的父母、妻子、兄弟、亲戚、朋友也都不例外。

世俗的社会关系被主动斩断后，相互之间的作用力自然也就消失了，各个维度上各种弹簧的社会作用力均会处于失效状态。出家的个体，减少社会关系约束，于是就进入了一种“出世”的状态。因为断了尘世关系、不再受各种“力”的纷繁扰乱，个体就获得了极大的解放，心灵也不会觉得那么累了。

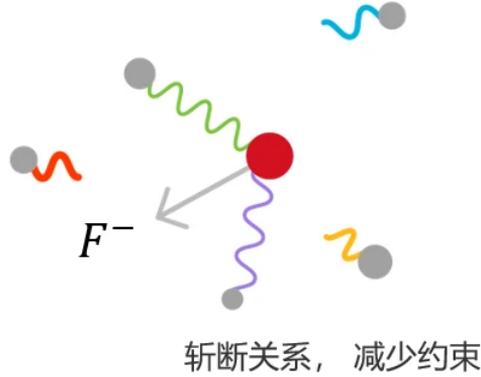


图 15: 斩断关系、减少约束

4.3.2 V系统——内在的价值体系

U系统是外在世界、社会场域的作用力、规范力、制约力，是个体之外的一种系统空间。我们内部还存在一个V系统，在这个V系统空间中，每个人都有着主观能动性，受各种生理欲望、内心信仰等力量的驱动。人在进化的过程中受到环境、生物属性、激素等共同的影响进化出了价值体系。

人类在很多维度上都存在一个主观价值函数，V系统是定义在很多个子空间与维度上的价值函数的集合。

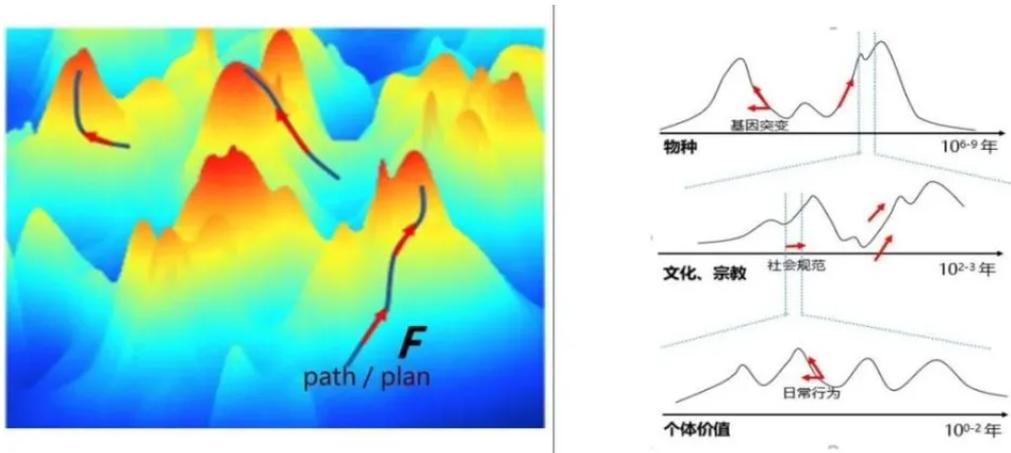


图 16: 价值函数的三个尺度，与内驱力 F (梯度上升的方向)

总的来说，人类具有三个尺度的价值函数：

1. **生物进化的尺度**：跨越数十万到亿年。这个尺度的价值函数往往是由基因突变带来的进化导致的，不受当下人类的控制和左右，除非基因编辑技术的发展能够精准到改变我们的价值函数！
2. **文化和宗教尺度**：几百年到千年左右。不同国家、不同时期会有不同的社会价值取向，其盛行的文化和宗教也往往不一样。于文化而言，比如中国好几个朝代都信奉“万般皆下品，唯有读书高”，那社会民众都会以考取功名走上仕途为荣。对宗教而言更不必多说，不同的宗教对人的价值函数要求大为不同，甚至很多都是互斥的。这就是所谓文明的冲突，首先是价值体系的冲突。
3. **个体价值尺度**：百年左右。个体有差异，具体定义到日常行为上的价值函数，比如对颜色、味道的偏好。具体这个V的函数还包含了个体的感知、认知的能力。比如，短视的人与目光远大的人，V系统就有区别。

根据英国哲学家、经济学家边沁提出的“效益原则”，人的每个行为都是希望在不同的空间或维度上实现价值函数V的最大化。无论是最大化个人的价值函数，还是最大化全人类的福祉，效益最大化都是我们内心的驱动力。这个主观的内心驱动力，就是对价值函数求导，梯度上升，数学表达公式为：

其中，“z”表示流态，代表人处在不同空间的行为函数。通俗来说，“人往高处走，水往低处流”，说的就分别是对V和U两个系统的求导结果。

4.3.3 U、V系统的平衡统一

那么，在U和V两个系统下，人能否达到平衡统一？答案是肯定的，无论是佛家消极避世的UV平衡(出世)，还是儒家积极进取的UV平衡(入世)，人总是要维持一个自治状态，否则个体就会处于失衡、纠结的内心状态，亦会表现出错乱、紊乱的行为模式。

我们可以用下面的数学公式来表达UV平衡：一边是在U系统中受到的各种社会关系的合力，另一边是对另一个空间的V价值函数求导(针对某个行为a)而得到的力。

当二者相等，即可求解出均衡状态。

“理”与“心”平衡的数学表达

顺便说一下，我们也可以用概率的方式来定义。但是，用力的方式定义更直观，我在这里不够严谨，也不展开来说了。

当两个力相等的时候，人就达到了平衡状态，这个平衡就是孔子所言“从心所欲而不逾矩”的状态，即心里想这样做，外界也希望我这样做，二者实现了基本的平衡与一致。这时候人的心就平和了，佛学叫做“自在”的状态。既知道自己根据内心V系统应该做什么，又知道社会外界的U系统期待自己做什么。

但是如果两者差得很远，就是一种失衡的状态。这就免不了会产生心理上的矛盾。主观行为与社会规范不匹配，就会出现贯穿一生的抑郁、不得志、不幸福心态。

绕了这么大一圈，咱们再次回到《赤壁赋》。

苏东坡夜游赤壁的思考，就是一个高级知识分子、士大夫寻求UV平衡态的思想过程与困惑。苏东坡少年成名，早年得志，在京城无疑是风流人物，可以说达到了“人生巅峰”。在朝廷做高官的时候，他有很多的社会关系，都是社会名流、才子佳人，可谓“谈笑有鸿儒、往来无白丁”，其“朋友圈”的人脉、资源都非常高端。一句话来讲，早年苏东坡的U系统很大。

同时，苏东坡早年V价值体系的格局也很大。所谓格局，我这里试图给一个定义，就是说一个人的V系统(价值函数)能够定义在多少空间、多大维度的子空间上。空间越多、维度越大，V价值体系的格局越大。

空间越大，格局越大。所谓“家事国事天下事，事事关心”，就是个体在家庭、国家、天下等不同的空间上都定义了价值函数，即在这些空间中，其行为都有内在的、主观的驱动力。这是一种“志得意满、指点江山”的境界，因此，苏东坡早期在U和V等高维空间中达到了一种层次很高的平衡。这叫做“入世”的平衡，也称为“高配版”的平衡，即U和V都很高且相匹配。就在写《赤壁赋》7年前(1075)，苏轼在密州当知州时，描绘的就是在大平衡中的心态与气派：

“老夫聊发少年狂，左牵黄，右擎苍，锦帽貂裘，千骑卷平冈。”

当然，还有“低配版”的平衡，佛家提供的解决方案大概就是如此。佛教的出家修行，能够在两个方面分别降低U和V的空间维度，使二者在“降维”的空间达成平衡。

一方面是降低U系统的成分，相当于把弹簧(U)斩断，断绝多种社会关系，这样一来，个体需要面对和处理的的社会关系大大减少，U产生的外在“压力”就会非常小。

另一方面是降低V系统的成分，佛教通过对出家人施加250多条戒律，来约束个体主观的行为空间(V)，之后再吧价值函数V拉平，使其不再对事情做价值判断。正如《心经》所言，“不垢不净”——人不要有好恶、干净污垢之分，这就相当于把价值函数全部抹平为一个常数，对V求导，产生的内驱力等于0。由此整体的内驱力就会非常小。

这种降维空间的平衡就是“出世”的平衡，这是一种“低配版”的平衡。这是人人都可以做到的，俗称“佛系”。所以，我认为，佛学的思想是给人们提供了一个“精神层面的社保”，让人人都能实现一种自在的生活状态。

儒家入世的思想强调社会责任，要求高维度、大格局的“大平衡”，如下图，小平衡可能是两个维度(2D)，大平衡则可能达到一百个维度(100D)。

4.3.4 在不同境遇中不断调平U、V系统

苏轼一辈子都在迁徙之中：从朝廷(京城)被贬至黄州，东山再起回到京城之后，又

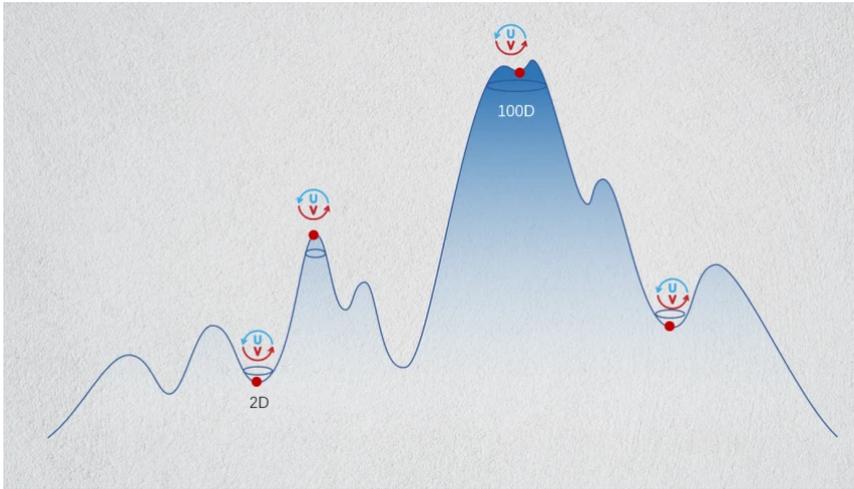


图 17: 不同维度的大平衡与小平衡

被贬到杭州，最后到海南岛，可以说是“筋力疲于往来，日月逝于道路”。他一直在努力调节自己的内心(V系统)，力求在不同的位置与环境(U系统)都能达到新的UV平衡。例如“竹杖芒鞋轻胜马，一蓑烟雨任平生”，就体现了苏轼的豁达心态，他能够随遇而安、进退自如，能够使主观V系统与外在U系统进行匹配与适应。同时，苏东坡又欲成大事，实现对国家和人民的抱负，即存在入世的愿望很强烈的时候。

苏轼时而想要追求低配版的小平衡，时而想要追求高配版的大平衡。《赤壁赋》一文集中反映出苏轼对自身前途理想的追求与抱负，出世与入世的思想矛盾。

一方面他想接受目前的状态，颇为享受当下的UV小平衡状态，闲度喝酒戏舟、侣鱼虾而友麋鹿的悠哉日子。但另一方面他又心有不甘，内在对更高层次的UV大平衡的渴求不言自表，如文章中对三国赤壁的吊古伤今之词句，就反映出他怀有强烈的“入世”愿望。于是在追求大、小两种平衡的矛盾心绪中，他沉沉睡去，不知东方之既白。

《赤壁赋》反映了人的一个普遍的心境，是普通人、科学家、企业家、官员，都会在人生中遇到的烦恼。位高权重的人，在位时大U大V，肩负巨大的社会责任，能创造巨大的社会价值。退下来之后，U大部分消失了，身处在V大而U小的状态，心态难以平衡、不好适应。要逐渐把大V降成小V，以适应小U，这样才能达到一种新的动态平衡。

我们每个人都应该去寻找一下自己的U和V究竟是什么。当遇到挫折的时候，不断寻找内心的平衡。像苏轼那样，在各种境遇下能做到自得其乐，而又积极向上、向善。正所谓，“穷则独善其身、达则兼济天下”。

顺便说一下，个体有一个(U, V)，这代表一个人的人性，群体也有一套(U, V)，代表该群体的文化、文明。个体的(U, V)系统与群体社会性(U, V)系统，往往处在矛盾冲突之中。利他性是指个体会自觉牺牲一己的利益、权利、幸福以至生命，以完成某种

群体(家庭、氏族、国家、民族、阶级、集团、宗教、文化等等)的要求、利益。比如历史上的荆轲、岳飞、文天祥等人，这里不展开说了。

4.3.5 “心”与“理”平衡的认知架构

个体的(U, V)系统在社会通信(包含交流、学习)过程中，不断更新迭代。这个迭代的过程依赖于人(智能体)的认知架构。下图表示的是两个智能体 (可以是AI, 也可以是人, 人机混合)的通信模型, 也就是一个基本的认知架构。这里每个椭圆代表一个客观的世界(最底下的椭圆), 或者是客观世界在智能体脑袋中的映射。s是客观物理世界的一个状态, 包括物体、场景和人物、行为。每个智能体(AI或人)都有一套世界模型 θ (“知”), 其空间是一个五边形, 以及决策函数 π (“行”), 其空间是一个菱形。对于输入的感知图像I, 根据自己的“知”模型 θ , 对客观物理世界状态有一个概率的估算, 同时也对于另一个智能体对物理世界的状态估计也有自己的估计。每个智能体还有自己的价值函数V, 其空间是一个方形。

“知”模型和决策函数构成了“理”:U; 价值函数则构成了“心”:V。“理”与“心”在这个认知架构里并存。通过交流, 智能体之间会形成共同的价值观 V^* , 社会规范 π^* , 共同常识 θ^* 。

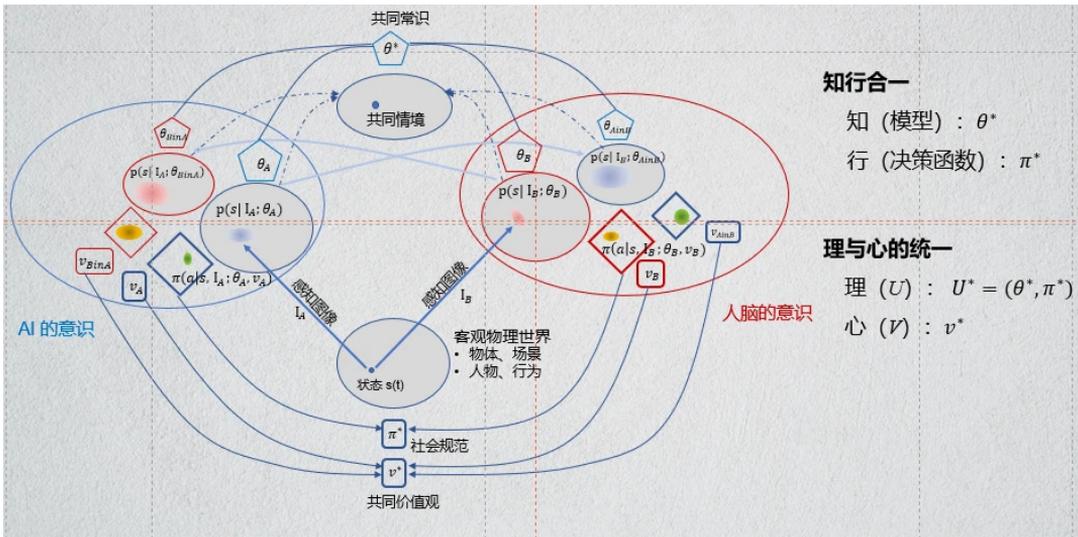


图 18: “心”与“理”平衡的认知架构

所谓“格物致知”，就是双方各自对所观察到的客观世界U系统进行建模，然后通过交流互相了解对方的行为决策函数、价值观，找到差异并形成共识。于是我们就和机器人建立了一个共同的模型，这个模型就是社会规范(U*系统)。我们的V系统在交流

中达成一致的部分，就是社会共同的价值观(V*系统)。

中国古典哲学中关于“心”与“理”的对立统一的辩论被抽象为数学模型后，就可以被认为是关于“U”占主导还是“V”占主导的问题。“理学”认为U是主导,个体行为应同客观世界的天理、运行规则相一致。“心学”认为V是主导，个体内在驱动力是第一性问题。

前文谈到的“菩提自性”也许对应的就是这个认知的架构。儒家所言的“内圣外王”的境界或许可以理解为，把V系统调好了，有了公心，由内心价值导出的行为将是公众认可的王道。

通用人工智能的目标是创造具有自主感知、认知、决策、学习、执行和社会协作能力，符合人类情感、伦理与道德观念的智能体。TA所需要的正是这样的一个“菩提自性”，然后TA才可以融入到人类社会，与人类达成新的共同的(U, V)系统。这也是AI发展给人类带来的冲击。换句话说，当变化来临之时，人类也需要重新调整，达成新的平衡。

我一直在强调一个观点，要实现通用智能，5%要靠客观的观察，95%要靠主观的内心需求与想象，研究通用人工智能的关键在于为主观的“心”建模。

《赤壁赋》之所以能够千古流传而经久不衰，很大的原因在于它给人们留下了“普遍存在的、人人需要回答的谜团”没有得到答复：一是理与心的平衡；二是人生能得到什么、有什么终极意义。

这是三读《赤壁赋》带给我的哲学启示，我认为人工智能这样的大学科是需要一个底层的哲学作为构建基础，这种思考贯穿了我的研究生涯。

5 结束语

这篇文章的本意是解读苏轼的《赤壁赋》，与大家分享我对此千古一赋所做的人生思考，并非提供一个关于人生意义的统一答案。每个人可以追求不同的平衡态，自得其乐。我个人的人生态度是积极入世，也希望人们从追求“活着”、到“活得好”、到“活得有意义”、“活得明白”，甚至到达价值的极限——不朽。

2020年我刚刚回到北京，就去拜见了一位研究佛学的德高望重的前辈，我与他就人工智能的发展对人文的影响以及人生价值做了一次交谈。

据我肤浅的了解，对应于本文开篇的人生三个层次说法，佛学总结的人生意义也是三个层次：

第一是求做和尚，就是穷人出家，找一条活路；

第二是求作佛，就是做一个觉悟的人。显然，佛教并不追求物质丰裕和社会虚名这种世俗的“活得好”，而是直接跳过这一层。

第三是求涅槃，涅槃就是放下一切执念，包括对佛法本身的执着，忘记自己“登山”的路径，从而达到一种融通的境界。

果然，他以佛学的智慧否定了我的入世观点：

“人生的意义就在于无意义，无意义就是最大的意义。”

“越有价值就越无价值，越无价值就是越有价值。”

“一生求索活得明白，临了方知难得糊涂。”

好在多年来文章投稿被审稿人评论、拒稿，我已经非常习惯了，我当时听了这个观点，会心一笑。在佛学看来，追求人生价值也是一种执念，当人活到了一定年龄，也只得放下，否则是自寻烦恼。

6 本文由来

2021年3月12日，北京大学学生工作部邀请我在“教授茶座”与青年学生分享个人求学与研究的经历、探讨科学精神和人生价值，我便尝试从文学、科学、哲学三个层面解读《赤壁赋》，从苏轼的经历与哲思中学习人生的定位。



图 19: 北大教授茶座

2021年7月30日，在ACM图灵大会前夜于合肥，大会主席、清华大学刘云浩教授安排我给青年学生学者做一次“围炉夜话”式的座谈，以三读《赤壁赋》为主题，探讨学术人生。

之后，2021年9月11日，受刘志峰同学邀请，我给母校中国科大北京校友会做了一个晚餐的讲座。2021年10月17日，在北京通用人工智能研究院的组织下，给北大、清华



图 20: “围炉夜话”座谈

通用人工智能实验班(简称“通班”)参加团建的同学讨论学术人生的道路，再一次谈到《赤壁赋》与苏轼的人生价值取向。

本文根据“教授茶座”、“围炉夜话”、“通班团建”活动现场录音整理。因近来工作繁忙，我只能抽碎片化的时间来完成此文，文中难免漏洞百出，或有言辞不妥之处，还望读者海涵。人生观、价值观的问题没有对错之分，本文只是个人粗浅的认知，分享出来，希望给青年学子，以及热爱科研探索、勤于思考人生意义的人们，带来一点启发。

本文也试图探讨当代人工智能研究和传统人文社科研究的双向连接。一方面，人文社科学科中蕴含的思想对人工智能特别是通用人工智能的研究发展具有很大的启发意义。另一方面，通过人工智能的数理模型可以定义、重构和解读人文经典中的一些关键概念、为其提供一些显式的数理表达，从而更好地研究其认知架构体系。

7 附录：在讲座上的问答摘编

提问一：朱老师好，您之前在一次讲座中提到过“智能科学是一个物理和生物混合的复杂系统。智能作为一种现象，表现在个体与自然、社会群体的相互作用和行为过程中”。物理学所面对的是一个客观的世界，而智能科学却不仅仅需要去面对客观世界的因果律，还需要考虑智能体本身的价值判断和任务驱动。您觉得有没有可能建立一个统一的理论模型来同时包括物理学规律与智能科学，建立两者之间完整的因果链条？

答：非常好的问题。我个人相信，当然这仅仅是个人的学术信仰和追求，在物理和智能之间存在一个连续的频谱，而我们要做的就是找到它们在这个频谱上统一的模型。这个问题在认知科学被称作“animacy”。我之前在UCLA的一位博士生舒天民，目前在MIT认知科学实验室做博后，他的博士论文就是关于“物理社交模型”的认知心理实



图 21: 北大、清华通用人工智能实验班

验。从小钢球的运动，到蚂蚁，到人类的活动，系统越来越复杂，需要的变量也越来越多，模型空间的复杂度不断增长。我们希望研究在整个频谱上的空间怎么跃升(augment, jump)，空间到底有哪些结构，里面有多少个函数。在“三读《赤壁赋》”这篇文章里我着重谈了U系统与V系统的平衡，下一篇我将具体谈一下“心”(V系统)与“理”(U系统)的边界和统一。

长期以来，至少在中学和大学的物理课本上，物理学的模型被认定是客观的，被说成是定律。但其实这些模型只不过是人对客观世界的现象的解释，没有对错，只看好不好用。这些模型的复杂度取决于人的任务所需要的精度。在人工智能领域，我们的模型与任务和主观判断更是密不可分。比如，本文谈到的纹理模型，判断的直接依据就是人的视觉感知。这种主观与客观的深度融合，体现在U与V两套系统的交织，是十分有趣的前沿问题。

提问二：朱老师好，人有第一人称视角的主观体验，这似乎是我们能感知到“意义”存在的一个基本条件，您认为未来的人工智能是否可能拥有一种内在的视角呢？

答：我个人相信是可能的。在人工智能相关领域中，大家一般都回避“意义(meaning)”，因为从U系统的模型是没法谈意义的，需要建立V系统。这个问题正好是我和团队近年

来做研究的方向。我认为人工智能的很多知识不是从数据获得的，而是从“内心”获得的，人工智能如果没有内心模型和价值体系，就很难拥有像人这样能够泛化学习的能力。佛家有句话叫“相由心生”，大意是：对外界的感知(相)是由内心的价值和任务驱动的。比如，慧能出场的惊世骇俗的一句话：“不是风动、不是帆动，仁者心动”，首先心要动，才能看到帆动。这个时候，你的任务可能是想去读帆上的字，或者看看它的颜色。中国有一个寓言故事叫“疑邻盗斧”，当这个人怀疑邻居偷了他的斧头，就越看越像。

提问三：朱老师好，如果一个智能体存在内在的价值体系，TA自己内部去训练、更新价值观。但我们从外界可能无从知晓TA的价值观是什么，于是只能把TA当做一个黑箱，会有很多人表示担忧，担心TA是无法控制的。您怎么去看人工智能内在的价值体系？

答：这的确是人工智能的伦理问题。智能体想要和人类长期共存，那TA们必须符合我们的价值观，也就是和人类的价值观达到动态平衡。我们这个社会也存在坏人，但是坏人会遭到惩罚。我相信未来的智能体要有自己的“菩提自性”，能够自主调节U系统和V系统，学到符合人类社会的价值观，才能长期稳定地存在于这个社会。当然，法律系统需要涵盖智能体，在未来的智能社会如何立法与执法是同步需要研究的问题。

提问四：朱老师好，我们前两天上课的时候提到“一个程序不能够以它自身为它的作用对象”，即程序不能修改它本身。那么未来的智能体，它的任务对象是否有可能是它本身？

答：现在的养成游戏里面已经有可以自主学习的机器人，自我探索、自我学习，每天都在修改自己。智能体不断地修改甚至跃升augment自己的各种空间和函数。就像前面问题一谈到的，从简单物理到复杂智能体存在一个连续的频谱，人类是站在了这个频谱的一端，没有理由相信人类的智能就是极限和终结，人和人之间不是也有差异吗。假如某一天智能体自己把这个频谱搞清楚，它就有可能超过我们，从理论上讲这是可能的。

提问五：朱老师好，目前我们在使用手机拍摄照片的时候，会越来越多地使用图片内置的软件，比如滤镜，美颜等功能，当前是否存在这样一种趋势，即少数人设置了人工智能的审美，而这种审美被广泛的大众所接受，接受了之后又引发了对广泛大众审美的一轮新的洗刷。想问您是否感觉有这样的趋势，以及如何看待人工智能对人类审美的影响？

答：“齐王好紫衣，国中无异色。楚王好细腰，宫中多饿死”。这个现象自古以来就有，从研究角度看，这是价值V系统在人群中对齐(alignment)的一个现象。美颜这种影响还是比较轻的，更重要的是AI技术能改变人类的价值观。在信息时代，应用程序分析用户的兴趣来推送符合用户喜好的信息，有可能包含虚假甚至危险内容，导致人的价值观走向偏激，社会逐渐向极化的方向发展，产生社会分裂。对智能的治理已经迫在眉

睫。

7.1 鸣谢

感谢北大学工部王逸鸣等老师的辛勤劳动和组织教授茶座活动，并精心制作了视频，感谢刘云浩教授、刘志峰同学的邀请，使得我不断思考完善这篇文章。感谢康建、吕鹏、刘可、陈大鑫、董亦飞、杨志宏、吴郢等人协助修改本文。由于笔者事务繁忙，这篇文章的定稿一再延迟，文中疏漏之处，请读者海涵。

最后，我给北大学生的题词，也送给那些热爱学术，对探寻真理抱有无限热忱的同学们、读者朋友们：

登无人之境

享清风明月

平天理心欲

活精彩人生

7.2 关于朱松纯教授



朱松纯教授，全球著名计算机视觉专家、统计与应用数学家、人工智能专家。

1991年毕业于中国科学技术大学，1996年获美国哈佛大学计算机博士学位，2002年起任美国洛杉矶加州大学(UCLA)统计系与计算机系教授，UCLA计算机视觉、认知、学习与自主机器人中心主任。

2020年，朱松纯教授回国筹建北京通用人工智能研究院，探索“小数据、大任务”的全新人工智能研究范式，构建计算机视觉、认知科学乃至人工智能科学的大一统理论

框架，研发通用智能体。朱松纯担任清华大学与北京大学讲席教授，在北京大学和清华大学成立“通用人工智能实验班”（即“通班”），面向人工智能未来发展方向，培养通用人工智能方向的“通识、通智、通用”型国际顶尖复合型人才。

资料整理：mfy

2022年6月2日