Artificial Intelligence and Robotics Research 人工智能与机器人研究, 2019, 8(3), 140-157
Published Online August 2019 in Hans. http://www.hanspub.org/journal/airr
https://doi.org/10.12677/airr.2019.83017

Hans汉斯

Artificial Intelligence Upper Limit Demonstration

Shaowei Zhang

Hangzhou Xuanjia Intelligent Technology Co. Ltd., Hangzhou Zhejiang
Email: gquizhang@126.com

Received: Jul. 29th, 2019; accepted: Aug. 16th, 2019; published: Aug. 23rd, 2019

Abstract

With the development of artificial intelligence technology, the whole society has a heated discussion on the extent of artificial intelligence replace human beings and whether artificial intelligence can surpass human beings. This paper will demonstrate the upper limit of AI by comparing Turing Machine (the principle of AI) with Epistemology (the principle of human brain). On this basis, it will determine the orientation of AI in the future society and the principles that AI development should follow. With the solution of these problems, the principle of AI will be determined. The significance of this work lies in: 1) avoiding the social risks brought by AI to the maximum extent; 2) avoiding the development direction which AI cannot achieve to avoid the huge waste of social resources; 3) defining the direction of AI discovery, maximizing the advantages of AI to bring the human society to the greatest extent.

Keywords

Artificial Intelligence, Turing Machine, Epistemology, Construction of Artificial Intelligence Theory

人工智能上限论证

张绍伟

杭州玄价智能科技有限责任公司，浙江 杭州
Email: gquizhang@126.com

收稿日期：2019年7月29日；录用日期：2019年8月16日；发布日期：2019年8月23日

摘 要

随着人工智能技术的发展，全社会在人工智能这一技术究竟能在多大程度上替代人、人工智能是否能超

文章引用: 张绍伟. 人工智能上限论证[J]. 人工智能与机器人研究, 2019, 8(3): 140-157.
DOI: 10.12677/airr.2019.83017

越人的问题上展开了热烈讨论。本文将通过图灵机(人工智能的原理)和认识论(人脑的认识原理)的对比来论证人工智能的上限在哪里，在此基础上确定人工智能在未来社会中的定位以及人工智能开发所应遵循的原则，随着这几个问题的解决人工智能的理论建设才真正完成。这一工作的意义在于1) 能最大限度的避免人工智能带来的社会风险；2) 避开人工智能无法实现的开发方向以避免对社会资源造成的巨大浪费；3) 明确了人工智能的发现方向，最大限度的发挥人工智能的优势以带给人类社会最大的福祉。

关键词

人工智能，图灵机，认识论，人工智能理论建设

Copyright © 2019 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

随着人工智能技术的不断成熟，整个社会对于人工智能及其未来发展的讨论越来越热烈。早些年拍摄的人工智能题材的影片与一些有社会影响力的人如马斯克、已故的霍金等都向人们宣扬这样一种观点即随着人工智能技术的发展人工智能最终将超越人类甚至统治人类，一些研究人员也受到了这种观点的影响试图从理论角度证明人工智能可以超越人类。目前社会基本形成了这样的共识，人工智能最终将超越人类，它将是一种比人类更高等级的存在。

我们可称此观点为“超越论”，此刻姑且先不讨论这种论点是否正确，这种观点的存在起码说明了一个事实即社会要求对人工智能有一个整体完整的认识，我们需要对人工智能做好理论建设以指导人工智能未来的发展。

不知大家是否对超越论有过质疑？是否考虑过超越论如果错误会给社会带来的危害？举个例子：

有的人在开发人工智能技术来治疗抑郁症，而且从事这项研究的是在人工智能技术领域非常资深的专家，但是从目前病人的反馈来看这项技术不仅对病人没有帮助而且会加重病情。

如果超越论是错误的即这项技术是不可实现的，那么我们想想从事这项研究的科学家是否就把自己宝贵的研发生命浪费在这一个根本无法实现的技术上？这对于科学家个人和社会都是一种巨大的资源浪费，同时那些接受治疗的病人也会耽误宝贵的治疗时机而加重病情。

此例说明了为人工智能做理论建设的重要性和紧迫性，毕竟我国已经将人工智能提升为国家战略，不远的将来必定会大规模的应用于社会，那么像这样的例子就会成千上万，对社会造成的浪费与危害将是十分巨大的。

时代要求我们必须对这一问题作出回答，因此本文通过对人工智能上限研究驳斥超越论，为人工智能做完整的理论建设，其内容包含以下三个问题的回答：

1) 人工智能的上限在哪里？
2) 人工智能在未来社会的定位是什么？
3) 开发人工智能技术的原则是什么？

通过对以上几个问题的回答，我们将解决以下几个问题：

1) 最大限度的避免人工智能可能带来的风险，使技术进步带来的风险在社会绝对可承受范围内。
2) 明确人工智能的发展方向，使人工智能最大限度的造福人类社会，避免走进错误方向带来的巨大

资源浪费

3) 明确人工智能技术的改进方向，使人工智能技术能够快速的完善与成熟。

自从09年大学时期意识到人工智能在未来社会的重要意义后，10年来一直在为以上问题寻找答案，古人云：“愚人千虑，必有一得”，接下来我将尽我所能为这些问题寻找一个答案。

2. 图灵机和认识论介绍

2.1. 人工智能理论来源

说到人工智能不能不说的一位科学家就是图灵，他于1937年在《伦敦数学会文集》第42期上发表了一篇题为《论数字计算在决断难题中的应用》(《On Computable Numbers, with an Application to the Entscheidungs problem》)的论文。对于图灵的伟大成就和这篇论文的伟大意义我就不再赘述了，为了本文论述需要只明确一点，在这篇论文里提出的图灵机模型是我们现今所有计算机技术的理论基础，如手机、PC机、服务器、超级计算机、神经网络算法、量子计算机等，所有这些已经实现和未实现的计算机技术都没有超出图灵机理论的范围[1]。

1950年10月，图灵在论文《计算机器与智能》《Computing Machinery and Intelligence》中提出制造一台能够达到人类智慧水平的机器，当然这个机器也是基于图灵机的理论基础，并提出了“图灵测试”[2]。

正是基于图灵的这些工作，使很多人确信人 工智能能够最终像人类一样思考，甚至最终超越人类、取代人类，加上后来出现 了一些很成功的基于这个观点的科幻小说和科幻电影，使这一观点在社会中得到广泛传播并深入人心。

图灵的《论数字计算在决断难题中的应用》论文是非常严谨的数学论文，不仅与邱奇等人提出的理论相互印证而且经过了实践的充分检验。如果图灵想要在理论上证明有一台能够达到人类智慧水平的机器，那么他最好的办法就是以图灵机理论为基础，拿机器的智慧和人的智慧做对比。对于人的智慧的认识就是哲学中的认识论，可是西方哲学对于认识论众说纷纭，从古希腊的希腊三哲到笛卡尔、洛克、休谟再到康德、黑格尔，始终没有一个人能提出一个完美的认识论获得大家的公认，因此对于只接触到西方哲学理论的图灵来说最好的办法就是做一个实验，在同等条件下分别用人和图灵机做对比测试，这就是图灵在《计算机器与智能》中提出的图灵测试[2]。

(**注**：

图灵测试：图灵测试(The Turing test)由艾伦·麦席森·图灵发明，指测试者与被测试者(一个人和一台机器)隔开的情况下，通过一些装置(如键盘)向被测试者随意提问。

进行多次测试后，如果有超过30%的测试者不能确定出被测试者是人还是机器，那么这台机器就通过了测试，并被认为具有人类智能。

30%这一数值是图灵预期人工智能在2000能达到的水平。

)

但是图灵测试只是一个测试，并没有对人工智能是否能超越人类的问题给出一个确切答案，如今人工智能已经逐渐进入应用，我们不能继续用图灵测试作为这一重大问题的答案，原因我在引言中做了详细说明。

本文要对这一问题给出答案，基本方法和图灵在《计算机器与智能》中用的方法一样，就是以图灵机理论为基础用人和图灵机做对比，不同的是我将用认识论(人的智慧)和图灵机(机器的智慧)直接做对比，以得到一个确切的答案。

说到这里我们就需要对哲学中的认识论有一个起码认识以方便我们后面的论述，因此接下来我将先简单介绍下认识论。

2.2. 认识论

认识论是西方哲学的根本问题之一，用柏拉图的话说是“关于知识的知识”，即我们如何认识这个世界的问题。这个问题的产生是随着对本体论的思考产生的，因为从古希腊开始哲学家就开始思考这个世界的本质是什么，在思考本体论的过程中哲学家发现我们要认识这个世界是怎么样的就首先要弄清楚我们是怎么认识这个世界的，因此认识论就产生了。关于我们是如何认识这个世界的，近代西方哲学家分为两派即理性主义和经验主义。

理性主义者认为我们的知识来源主要是靠理性即我们的逻辑推理，从感官获得的感性经验是不可靠的。

经验主义者认为我们的知识来源主要是靠感官获得的感性经验，并在经验中得到验证。

(**注**：

以上是经验主义和理性主义最根本的分歧，实际上经验主义者之间、理性主义者之间的观点也不完全相同，我们此处只关注主要矛盾明白二者的根本区别即可，在西方产生经验主义和理性主义分歧之前中国的宋明期间爆发了心学和理学的分歧，其主要矛盾也基本相同，如有兴趣深入了解可查阅相关资料。

)

后来康德对这二者进行了调和，他认为“感性无知性则盲，知性无感性则空”[3]。这对感性和知性的总结无疑是客观的，这也是康德在西方哲学界享有盛誉的原因。

以上只是对认识论的简单介绍，除了让大家对认识论有个大概认识外主要的目的是引出康德的观点(在 3.1.3 中我会对此观点再加以完善)。下面我对康德的观点做一个解释以便更好的理解。

康德认为我们认识世界首先是通过感官如眼、口、鼻等获得颜色、味道、气味、声音等感性材料，再经知性对感性材料做分析整理(如抽象、综合等方式的处理)最终获得知识。这些感性内容是获得知识的材料，知性是获得知识的方式方法。所以如果没有感性给知性提供材料知性就没有可认识的内容，我们的大脑将一片空白，因此“知性无感性则空”。反之如果没有知性给感性材料做整理，感性材料就是一堆乱糟糟的混合物，我们的大脑就会一片混乱而无法明晰事物的规律、发展方向等，因此“感性无知性则盲”。

2.3. 整体论述逻辑概括

以上两节对图灵机和认识论做了基本介绍，这为我们接下来的论证打好了基础。

在 2.1 节中，我介绍过本文将通过对认识论(人的智慧)和图灵机(机器的智慧)直接做对比的方式求得人工智能是否超越人类这一命题的答案。可惜康德的认识论虽然客观但是不完整，假如以这样的认识论为基础那就一定会得出目前最流行的错误结论即超越论。

那么本文论证的最好的方法是先全面完整的总结认识论，然后用图灵机和完整的认识论做比较从而得出结论。但是这种方法的困难在于全面总结认识论所要涉及的问题实在太多，不仅会造成篇幅过大的问题而且很不利于非哲学专业人士的理解。那么可以采用另外一种简化方法：首先对认识论的两个关键点做总结，这两个认识论的关键点一个是人工智能能达到的另一个是人工智能无法达到的，通过人工智能与认识论的这两个关键点的对比就能清晰的看到人工智能的上限。这两个关键点分别是感性和感觉。

因此下文的论证逻辑是：第一步认识并总结感性和感觉，第二步给感性和感觉建立数学模型，第三步总结《论数字计算在决断难题中的应用》中图灵机模型的能力范围，第四步将图灵机理论与人的感性、

感觉能力作对比得出结论即人工智能的上限在哪里？第五步对引言中的所有问题给出完整解答。

3. 人工智能上限及其证明

3.1. 感性和感觉

佛家曰：“苦乐随缘，心无增减”，中庸：“喜怒哀乐之未发，谓之中，发而皆中节，谓之和”[4]。中国的哲学家主要以感觉作为认识的素材，西方的哲学家主要以感性 作为认识的素材，这是中西方哲学最主要的差别之一(造成这种差别的原因暂不讨论，因与本文论题无关)。接下来我们来深入认识感性和感觉。

3.1.1. 感性

感性是指感官获得感性材料的能力，感官指耳、眼、口、鼻等，感性材料指声音、颜色、味道等。

以上是对感性严格的科学定义，是以感官存在和物质存在为基础。

在哲学上这么定义可能会有人有异议，因为唯心主义者不承认客观物质的存在，在他们看来声音、颜色这些认识材料要么是假象要么是由某个终极存在(如上帝)直接植入人的意识中。本文遵循科学性原则，认物质存在为不证自明的。

(**注：**

客观物质本即存在，关于唯心主义者对客观存在的怀疑，我将来会在详细完整的认识论中给以解答。

)

因此，根据以上定义感性是人所具有的一种能力，感官是我们这种能力的载体，感性材料是我们这种能力的结果。

接下来让我们来分析一下我们所能获得的感性材料如声音、味道、颜色等，请看以下示例：

颜色有红、橙、黄、绿、青、蓝、紫等，可由三原色(品红、黄、青)按比例调出，自然每种颜色都可以是一个确定的比例值。

味道有酸、甜、苦、咸等，酸度、甜度等均可以通过测量得到准确数值。

声音虽有多种多样，但是声音的要素音调、音色、音量等均是可测量的。

其他的感性材料我不再一一列举，通过以上的举例我们可以归纳出感性材料的特点即感性材料均是可测的，通过工具测量可得到一个确定值。随着现代科技的发展我们已经深入事物内部发现这些感性材料的产生是客观事物作用于人体神经的结果(如颜色是可见光波刺激视觉神经产生，声音是声波刺激听觉神经产生，气味是分子刺激嗅觉神经产生)。古希腊哲学家普罗泰戈拉说：“人是万物的尺度”，我们的感官就像一把“尺子”一样精准的测量着这个世界，随着科技的发展人类不断的发明测量工具来代替我们的感官来测量这个世界，这些测量工具不但种类越来越多而且精确度越来越高，其准确度已经远远超越了人的感性。

(**注：**

康德将时间和空间看作先验感性[3]，此处对时空是否先验先不做讨论，但是时空首先是感性因此本节对感性的归纳总结必定同样适用时空。

)

3.1.2. 感觉

感觉的种类有很多，最基础的有饥饿、疼痛、美味、舒服等等，基本上我们的每一种感性都会对应一种感觉。比如味道的酸甜苦辣对应“美味”和“难吃”的感觉，声音有“悦耳”和“难听”，按摩如果力道重了会“疼痛”力道适中就“舒服”(对应的感性是触感)，那么我们的感觉是建立在感性的基础之

上吗？

显然不是，感觉和感性一样是人从外界获得的认识材料。比如说当人遭受核辐射时人是没有对应的感性可以“看见”这种核辐射的，但是人经过一段时间的核辐射后各种感觉如恶心、疼痛等就会向我们袭来。再比如有时候人会莫名的有一种危机感，这种感觉动物(尤其是野生动物)比人更具“天赋”，这种看不见任何威胁但是在危险来临时直接到达动物体内的危机感是野生动物生存的重要保障。

自然界的原始动物(如环节门的蚯蚓、水蛭等)具有的神经细胞是较早进化出的，这样的动物没有复杂的感性(未进化出)，运动主要靠神经反射，这种神经反射可看做动物最初的感觉，感觉“好坏”决定反射的动作。

通过以上例子说明感觉不仅同感性一样是人的认识材料而且感觉是比感性更原始的认识材料，感觉对于人的意义更重要。此外相比于感性的确定性感觉具有不确定性和变化性，接下来通过一些感觉的实例进行分析。

有人把疼痛分为十二级仿佛疼痛和感性一样有确定值，然而所谓的在医学上把疼痛分为十二级是一些网友的杜撰，科学上从不这么分级。其实每个人对疼痛的感受都不同，比如大家经常看到的分娩体验仪有的人在 2 级疼痛就已经受不了但是一些“天赋异禀”的人 12 级疼痛也没什么感觉。就如同每个人的口味偏好(口感)、音乐品味(乐感)、服装审美(美感)等都不同一样，每个人的疼痛感受也不同。人不仅对疼痛的感受大相径庭而且同一个人在不同时间对疼痛的感受也不同。就如同一个人的口味会变一样，身体条件、心情等各项因素都会影响对疼痛的感受，举个极端的例子如一个人受到剧烈伤害骨折时初期感受不到任何疼痛(交通事故中有很多这种例子)。因此每个人的疼痛感受具有不确定性而且同一个人的疼痛感受也在不断变化。

相比于疼痛其他感觉的不确定性和变化性更加明显，比如拿我们日常的吃饭来说，在我们十分饥饿时吃红烧肉或者其他肉类是最美味的，当我们吃的快饱时吃点青菜是最舒服的，吃饱时再吃肉类就会感觉到恶心。

(**注：**

当然人的口味是不确定的，此处是以喜吃肉为前提举例的，主要是为了说明一餐之间人的口味的变化之大。

)

我们从小到大口味都在不断变化，长大后有时去怀念童年买一些童年爱吃的食物但是却怎么都吃不出小时候的感觉了。

除了痛感、口感，还有乐感、美感等等都具有相同的规律，我就不一一举例了。

通过以上的大量事实我们可以对感觉做一个总结：感觉和感性和相互独立的，二者都是我们认识事物的基础材料，假如一定要区分二者对于人的重要性的话明显感觉更加重要。相比于感性的确定值，我们的感觉虽然也有正反强弱大小，但是这个值对每个人都是不同的，即使同一个人其值也是随时随地都在变化的。

(**注：**

虽然在科学上对于感觉已经做了大量的分析研究尤其是对于痛感，但是我们先不要着急分析感觉为什么不同为什么变化，我们在此处首先是从宏观上对所有的感觉做一个总结。

)

我们以上总结了感觉的不确定性和变化性，有的人一定会问为什么我见了多年不见的好友去了多年未去的故地当年的感觉又涌上心头，这不是说明感觉有时候是不变的吗？但事实上我们遭遇的多数情况如同中年的鲁迅遇到了儿时的玩伴闰土，感觉的变化远多于不变。斗转星移物是人非是人生常态，此处

可以用崔护的一首诗来形象表达：

去年今日此门中，人面桃花相映红。

人面不知何处去，桃花依旧笑春风。

变中有不变，不变中有变，就如同动中有静，静中有动，静是相对，动是绝对，同样变化是绝对，不变是相对。因此相比于感性的完全确定性，感觉具有不确定性和变化性。

3.1.3. 感觉和感性在认识论中的完善

前文引用康德的“知性无感性则空，感性无知性性则盲”时说过，这个观点对感性和知性的关系认识是客观的，但是这个认识论是不完整的。那么此刻我们就可以将之补充完整了，完整的认识论应该是：知性无感性和感觉则空，感性和感觉无知性则盲。为了更好理解，举个例子来说明：

当我们教一个小孩子认识小狗时，通常我们指着一条小狗对小孩子说“这是小狗”，小孩子通常会好奇的打量这个小狗有的小孩还会去抚摸小狗以及和小狗玩，那么下次再看见小狗时小孩子就会认识小狗。而且不论这条小狗是毛色变了还是体型变了，此小孩都知道这是小狗。那么为什么小孩子的学习能力这么强呢？答案很简单，因为人认识物质的材料既有感性又有感觉，小孩子观察小狗除了看他的外型颜色等感性信息更重要的是认识小狗带给他的感觉信息。如果这个小狗带给小孩子是“可爱”的感觉，那他就会记住这种可爱的感觉，如果这个小狗带给小孩子是“可怕”的感觉，那他就会记住这种可怕的感觉，下次再见到这种动物，上次那种感觉就会立马向他袭来(感觉信息判断)，加上外形的相似性(感性信息判定)，他自然就知道这个物体是上次大人告诉他的“小狗”。而人工智能认识事物为什么要靠大数据训练呢？我们从没有见一个小孩认识“小狗”需要看过成千上万的小狗图片才能认识。而且人工智能经过大数据训练，“看过”成千上万的小狗图片后形成了对小狗的认识后，给它一张瘸腿的或者少一只耳朵的小狗图片它就无法辨认。但是为什么小孩第一次看到一只少两条腿的或者少耳朵的小狗依然能够认识这是小狗呢？我想大家心里已经隐约有了答案，让我先卖个关子后文给出答案。

3.2. 感性和感觉的数学模型

在上文中对感性和感觉都做了分析，在这一节我将把感觉和感性抽象为两种数学模型以让大家有一个更清晰的认识。

感性是十分容易抽象的，很明显感性可以抽象为一把“尺子”，是一个白箱模型。虽然每一个种感性“尺子”的刻度和计量单位都不同，但是相应的刻度和计量单位都是确定的。用这把尺子测量所有物体都是一个确定值，我们发明的所有测量工具也可以看做是我们感性的延伸。因此感性是一把“尺子”，所有物质都是确定值。

感觉如何抽象呢？感觉首先无疑是一个数，我们的感觉有“好坏”两个方向(正负)，并且“好坏”都有“大小”的不同。但是明显感觉不是一把确定的“尺子”，是一个黑箱模型，我们就把它抽象为一个“黑箱”。每个人均是一个独一无二的“黑箱”，每个黑箱的“正负”与“刻度”都大不相同。对于某一个黑箱来说，所有的物体在这个黑箱面前都是一个？是一个未知数，直到这个物体通过这个黑箱的时候才显示一个数字。比如说我们不吃一样东西永远不知道它好不好吃以及多美味或多难吃直到入口的那一刻，我们不穿一件衣服永远不知道它穿在身上好不好看以及多漂亮或多丑直到它穿上身的那一刻，我们不滑一次雪永远不知道滑雪的感觉好坏以及多刺激或多难受直到我们尝试过。而且同一物体通过同一个黑箱每次显示的数值都可能不同，就如同我上文的举例一样，同一餐时间的同一种食物的感觉都完全不同甚至相反。至于痛感、审美等感觉的变化就不再赘述了。还有感觉变的绝对性和不变的相对性也要辩证认识。因此感觉是一个“黑箱”，所有物质都是未知值、变化值。

3.3. 图灵机模型及其能力上限

图灵在《论数字计算在决断难题中的应用》中提出了可计算性问题，并通过抽象出的图灵机模型来论述这一想法[1]。

在图灵发表此论文几个月之前，阿隆佐·邱奇在“关于判定性问题的解释”(A Note on the Entscheidungs problem)一文中证明出了一个相似的论题，但他采用递归函数和 Lambda 可定义函数来形式地描述有效可计算性。

在图灵、邱奇等人工作的基础上，邱奇图灵论题被提出来。

邱奇图灵论题：

以任何常规编程语言编写的计算机程序都可以翻译成一台图灵机，反之任何一台图灵机也都可以翻译成大部分编程语言大程序，所以该论题和以下说法等价：常规的编程语言可以足够有效的来表达任何算法。该论题被普遍假定为真，也被称为邱奇论题或邱奇猜想和图灵论题。

邱奇图灵论题表明，所有计算或算法都可以由一台图灵机来执行。也就是说图灵机的能力上限是逻辑和数学中的有效或机械方法可由图灵机来表示，“有效方法”这个想法在直觉上是清楚的但是我们无法严格定义，因此下文中我在引用这个定理的时候一定是谨慎的、直觉上完全正确的。

(**注**：

邱奇图灵论题之所以直觉上清晰但是无法严格证明就是因为涉及到了人的认识，如果要证明必定需要全面完整的总结认识论，前文说过这样会带来篇幅过大和难以理解的问题。因此我们此处就以图灵邱奇论题为理论基础，虽然不完美但是首先可避免前面所说的难题其次在直觉上大家是能够接受的。

)

3.4. 人工智能上限确定

经过了前面如此多的准备工作，本节我们终于可以对人工智能的上限做一个确定了，我们在前文中已经抽想出了感性和感觉的数学模型，接来下我们分析那些以感性和感觉为基础的数并确定他们的可计算性，从而最终确定人工智能的上限。

3.4.1. 感性为基础的数

在我们分析以感性为基础的数之前，首先确认一下实数、有理数、可定义数和不可定义数的定义。

实数：在数学上实数定义为与数轴上的点相对应的数，实数和数轴上的点一一对应。

有理数：是指两个整数的比，有理数是整数和分数的集合。

可定义数：可定义数是指能够有限的文字描述出来的数。自然数，有理数，代数数，圆周率它们都有明确的定义，都属于可定义数的范畴。

事实上，整个人类历史上所有文献提到过的所有实数都是可定义的，因为它们都已经被我们描述出来了。

下面定义一个特殊的可定义数。

蔡廷常数：

圆周率的小数展开看上去似乎是完全随机的，但毕竟是有办法算出来的。如果你想知道 π 的小数点后第一亿位是多少，我总能在有限的时间里算出答案来。

1975 年，计算机科学家格里高里·蔡廷(Gregory Chaitin)研究了一个很有趣的问题：任意指定一种编程语言中，随机输入一段代码，这段代码能成功运行并且会在有限时间里终止(不会无限运行下去)的概率是多大。他把这个概率值命名为了“蔡廷常数”(Chaitin’s constant)。

这听起来有点不可思议，但事实上确实如此——蔡廷常数是一个不可计算数。也就是说，虽然蔡廷常数是一个确定的数字，但现已在理论上证明了，你是永远无法求出它来的。

不可定义数：并不是所有的数都能够用有限的文字描述出来的。原因很简单，因为长度有限的文字段落是可以逐一枚举的(虽然有无穷多)，而全体实数是不能枚举的，因此总存在一些不可能用语言描述出来的数。这种数就叫做不可定义数。

从以上的定义可得出各种数的包含关系，实数包括不可定义数和可定义数，我们已经定义的自然数，有理数，圆周率，蔡廷常数等可定义数中有些是不可计算的如蔡廷常数。

包含关系见图 1 所示。

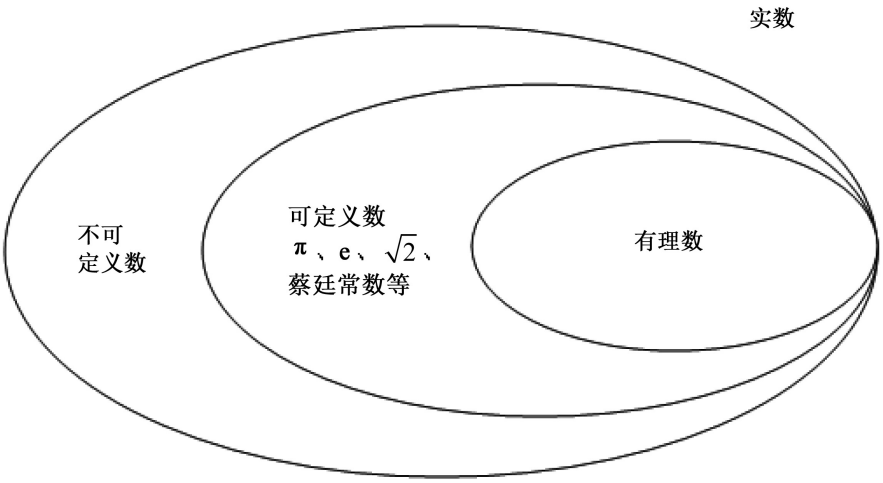

Figure 1. Inclusion relations of various numbers in real number range
图 1. 实数范围内各种数的包含关系

接下来我们分析以感性为基础的各种数。

我们在 3.2 中得出的结论是感性是一把“尺子”，任何物质在这把尺子前都是一个确定数，这个确定数就是我们直接测得的数有理数(整数和分数)。以有理数为基础我们又发现了 π，$\sqrt{2}$ 等无理数。

π 是圆的周长与直径的比值，$\sqrt{2}$ 是边长为 1 的正方形的对角线长度，e 的计算为当 $n\to\infty$ 时 $\left(1+1/n\right)^n$ 的极限，这些无理数的定义以有理数为基础，自然是以感性为基础的数。根据蔡廷常数的定义它也是以感性为基础的数，因此这些已经定义的无理数和有理数均是以感性为基础的数，它们的集合称为可定义数。

(注：

1) 理发师悖论：

在某个城市中有一位理发师，他的广告词是这样写的：“本人的理发技艺十分高超，誉满全城。我将为本城所有不给自己刮脸的人刮脸，我也只给这些人刮脸。我对各位表示热诚欢迎！”来找他刮脸的人络绎不绝，自然都是那些不给自己刮脸的人。可是，有一天，这位理发师从镜子里看见自己的胡子长了，他本能地抓起了剃刀，你们看他能不能给他自己刮脸呢？如果他不给自己刮脸，他就属于“不给自己刮脸的人”，他就要给自己刮脸，而如果他给自己刮脸呢？他又属于“给自己刮脸的人”，他就不该给自己刮脸。

2) 蔡廷常数的定义过程

我们来分析蔡廷常数的定义过程以确定它是否是以感性为基础的数。

首先理发师问题是以感性判断为基础定义的，以理发师悖论为代表的逻辑悖论引发了第三次数学危机，图灵的停机判定同样属于理发师悖论类逻辑问题，而根据蔡廷常数的定义它是以停机判定为基础定义的数，因此蔡廷常数也属于以感性为基础的数。

)

因此我们可以说可定义数均是以感性为基础的数。

通过以上分析我们发现可定义数均是以感性为基础的数，那么接下来我们是否可以得出结论在可定义数范围内人与图灵机的计算能力等价？

对于这一问题目前只有邱奇图灵论题可以提供理论支持，虽然这一论题还没有得到严格证明，但是该论题被普遍假定为真。在可定义数范围内那些算法清晰的数人与图灵机均可计算，如那些人可计算的有理数和无理数图灵机也可计算，而像蔡廷常数这样的数以逻辑悖论为基础人与图灵机均不可计算。因此邱奇图灵论题说明在可定义数范围内人与图灵机计算能力等价，这在直觉上是清晰的。

通过以上的分析我们得出结论，可定义数是以感性为基础的数，在可定义数范围内人与图灵机计算能力等价。

接下来我们关注和实数并列的另外一类数虚数。

虚数：在数学中，虚数就是形如 $a+b*\mathrm{i}$ 的数，其中 a, b 是实数，且 $b \neq 0, \mathrm{i}^2 = -1$。虚数这个名词是17世纪著名数学家笛卡尔创立，因为当时的观念认为这是真实不存在的数字。后来发现虚数 $a+b*\mathrm{i}$ 的实部 a 可对应平面上的横轴，虚部 b 与对应平面上的纵轴，这样虚数 $a+b*\mathrm{i}$ 可与平面内的点(a, b)对应。

虚数是以−1(整数)开平方和角度旋转为基础定义的数，因此虚数是以感性为基础的数。从理论上说，虚数的算法逻辑清晰，根据邱奇图灵论题在虚数计算上人与图灵机能力等价。从实践上说，早就用程序编写出了虚数计算器，事实证明在虚数计算上人与图灵机能力等价。

因此虚数也是以感性为基础的数，在虚数的计算上人与图灵机等价。

未来我们在感性的基础上或许还能抽象出其他类型的数，根据邱奇图灵论题均可证明在这些数的计算上人与图灵机的能力等价。

总结，以感性为基础的数不论是可定义数、虚数还是未来可能发现的其他数，人与图灵机的计算能力均等价，这正是人工智能强大的原因所在。

3.4.2. 感觉为基础的数

前面我们总结了感觉数的数学模型，它是一个“黑箱”，任何物质都是一个未知数直到通过这个黑箱时而且每次通过黑箱时显示的数都可能不同。这造成了感觉数与感性数的最大不同即感觉数是不可定义的而感性数是可定义的。

比如有两个人，一个人喜欢吃辣另外一个不喜欢吃辣。我们让喜欢吃辣的人把吃辣的感觉告诉不喜欢吃辣的人，那么无论他怎样描述即使他用尽一切语言也无法将他体会到的感觉传达给不喜欢吃辣的人。反之亦然。

中国的哲学以感觉为基础，其所研究的感觉不仅是这些日常的感觉而是早已进入了对人生、世界的感悟。庄子说：“天地与我并生，而万物与我为一。既已为一矣，且得有言乎”，意思是说我与天地万物化为一体，这种感觉无法用语言描述[5]。佛家有拈花一笑的典故，佛祖拈花，迦叶一笑，这种心灵的感觉的传递只能心心相印而无法用语言描述，因此佛家说“不可说，不可说，一说即是错”。

此处所说的感觉数的不可描述不可定义性要与上节定义的实数范围内的不可定义数区别开来，感觉数不属于实数，感觉数更加的“不可定义”。

数学经历了几千年的发展历史不可能对感觉数没有涉及，事实上我们在数学中定义了一个特殊的感

觉数无穷数，符号为∞。无穷在科学、神学、哲学、数学和日常生活中有着不同的概念，但是其根本来源只有一个。接下来通过分析无穷数的根本来源来更好的认识以感觉为基础的数。

(**注：**

西方哲学对于感觉研究甚少，但是感觉作为人的认识的基本素材之一不可能不被讨论，很多西方哲学家讨论过无限并且以此为基础论证上帝的存在，康德将对无限的认识归于理性，并以此将人的逻辑划分为知性和理性。西方哲学因为不了解感觉从而对无限产生的一系列错误观点与本文论题关系不大故此处不做批驳。

)

我们是如何认识到无穷的呢？这要从认识的最初阶段开始寻找。我们知道即使一个没有多少课堂知识的人也知道无限的含义，比如一个青年人看到一位心仪的女子，他多半会想我要和她永远在一起，这个“永远”即是时间的无限。此处之所以会用“多半”这个词是因此他并不绝对会这么想，但是假如他经历过长时间的求而不得辗转反侧就肯定会生出想要和她永远在一起的想法。这就如同一个人必须经历过长时间的饥饿才会想要拥有“吃不完”的食物，一个人必须经历长时间的贫困才会想要拥有“花不完”的钱一样。这方面和珅是一个绝佳的例子，和珅之所以贪污了“永远”也花不完的钱还是管不住手，就是因为曾经经历的贫穷痛苦太深。

从上面的举例我们可以看到人之所以希望有“无限”是因为曾经经历的痛苦(感觉)。

当一个人意识到人会死亡的事实自然会感到恐惧，自然他就希望“永生”，这是生命的无限。当一个人极度痛苦并且意识到现实世界就是有苦有乐无法改变时，自然就希望有一个“极乐”世界，这就是快乐的无限。

我们分析永远、永生、极乐就可以发现，在它们背后其实都是在追求无限的“快乐”逃避无限的“痛苦”，一个人最初意识到的就是无限的痛苦然后才追求无限的快乐。经历过极度饥饿的人都有过这样的体验，当饥饿达到一定程度时无论怎么吃都吃不饱，仿佛这种饥饿感是无穷无尽的。在我们成长的过程中，最初对于无限的认识就是来源于那种无穷的“痛苦”。无限最初就是通过“痛苦”直接跳进我们的意识之中，然后在实践中理性将无限与其他概念相结合发展出了无限的时间(永远)、无限的快乐(极乐)、无限的生命(永生)，最后无限和数相结合发展出了数学中的无穷数。这个基于感觉的无穷数和基于感性的有限数(实数)是有本质区别的，这个区别就是我们前文所说的感觉数不可描述不可定义性。

(**注：**

正如《道德经》所说“名可名，非常名”，我们只是给无穷数起了一个名字“∞”，我们也只是起了一个名字，对于无穷数的真正含义(常名)，我们无法用语言严格定义[6]。

)

至此，我们认识了一个基于感觉的数无穷数，它和复数，实数并列为数的三大分支。

无穷数的不可定义性和实数中的不可定义数有本质区别，无穷数不属于实数。为了不让二者混淆我们就将基于感觉的数命名为感觉数，无穷数是一个感觉数。

感觉数是不可定义的因此我们无法在图灵机中表示这个数，编程人员都知道无论哪种计算机语言中都无法定义真正的无穷数，如果我们需要表示无穷数的概念只能自行定义一个较大的数，但是这个数无论再大都是一个“有限数”和无穷数有本质的区别。

和无穷数一样其他的感觉数都是不可描述不可定义的，自然感觉数就是图灵机不可计算的数。现在我们就可为 3.1.3 中所举例子给出一个答案，图灵机只能“认识”感性数而不能“认识”感觉数，所以相比于人脑同时“认识”感性数和感觉数就有了先天的差距。一个小孩可以同时运用这两种数“认识”小

狗，当小孩子看见小狗时就可以从外形(感性)和可爱或可怕(感觉)两种方式来辨认，所以不用所谓的大数据训练就可以轻松“认识”小狗。而人工智能只能从外形(感性)去“认识”小狗，所以必须经过大数据的训练(即海量的外形数据计算)，并且一旦“小狗”少一只耳朵或者少两条腿(即感性数出错)人工智能就必定出错，它就可能将小狗认成小鸡。当然理论上来说每种动物都或多或少的在视觉上有一些差别，只要训练量足够大(即人工智能付出足够的感性数计算代价)，人工智能最终也能“认识”各种动物。

但是人的概念认识中有时候“感觉”是必不可少的。比如说有一个成语“士别三日当刮目相待”，一个人经历了大的人生变故一夜之间变“成熟”“稳重”了，此时这个人的外表(感性)没有发生任何变化，那么这种变化是人工智能永远无法“认识”的。再比如一只猫突然失去了主人变成了流浪动物，它从“宠物”变成了“野生动物”，这种变化也是人工智能永远无法“认识”的。人在思维活动中感觉不可或缺的时候有很多，我们下一章会做更详细的分析以确定人的哪些思维活动是人工智能永远无法实现的从而确定人工智能在未来社会中的定位。

总结，感觉数是不可定义的图灵机不可计算，这是图灵机和人相比的“先天缺陷”。

3.4.3. 人工智能上限总结

此时我们可以得出结论，人工智能的上限是对人的感性的计算，可以实现人的“感性认识”，并且随着传感器技术的进步人工智能的“感性认识”不仅能达到人类水平而且会超越人类。但是对于人的感觉计算，人工智能的理论基础图灵机理论具有先天缺陷，由于感觉数的不可定义性，图灵机永远无法实现“感觉认识”。

(**注：**

相信大家现在就明白为什么西方会流行人工智能超越人类论，假如以西方哲学的认识论为基础将人的认识仅仅看做感性的计算，那么人工智能无疑有一天会超越人类。

)

4. 人工智能的社会定位

在上节中我们对人工智能的上限做了界定即人工智能可以实现对以感性为基础的数的计算而无法实现对感觉数的计算，并且知道在概念认识中某些认识不能实现，如“成熟”“稳重”等认识人工智能永远无法实现。接下来我们总结在思维活动中哪些“认识”是人工智能永远无法实现的，哪些工作是智能绝对无法胜任的，从而确定人工智能在未来社会中的定位。

4.1. 创新

在我们的思维活动中创新是重要部分，从创新对象上人的创新主要分为产品创新、艺术创新和理论创新，而创新时刻离不开感觉的参与。

首先来说产品创新，在创新产品时考虑的因素很多但是首先考虑的是产品的使用价值，这是产品创新的方向，产品带给人的感觉好坏是判断商品使用价值的根本且唯一标准。商品吸引人购买首先在于它能带给人“愉悦”的享受，没人会花钱买“罪”受。比如说设计一款衣服，衣服满足人的需求有两个一个是保暖另一个是审美，在保暖设计上设计师有一个一般标准来满足春夏秋冬的保暖要求，但是每个人的冷暖感觉不同，尤其是春天大街上既有穿羽绒服也有穿单衣的人。在保暖设计上设计师可以参考一般标准不用考虑太多但是在外观设计上需要设计师深刻把握“美”感。一件衣服如何才是美的？这是一个感觉标准，不同的人有不同的标准，同一个人不同时期有不同的标准。那么如何设计一件“漂亮”或者“帅气”的衣服就是对服装设计师的能力考验。

人天生具有感觉，服装设计师的感觉更加敏锐，他需要靠“灵感”(感觉)来设计出一件“美”的服装。

没有任何设计师能把自己的设计“灵感”用语言精确定义出来，关于这一点我们前文总结过感觉数是不可描述不可定义的。

现在很多人认为人工智能具有创新能力是因为它有深度学习技术。我们接下来分析一下深度学习技术到底能在多大程度上实现创新。

深度学习(Deep Learning, DL)或阶层学习(hierarchical learning)是机器学习的技术和研究领域之一，通过建立具有阶层结构的人工神经网络(Artifitial Neural Networks, ANNs)，在计算系统中实现人工智能。

深度学习最有代表性的例子就是阿尔法狗，2016 年 3 月，阿尔法围棋与围棋世界冠军、职业九段棋手李世石进行围棋人机大战，以 4 比 1 的总比分获胜；2016 年末 2017 年初，该程序在中国棋类网站上以“大师”(Master)为注册账号与中日韩数十位围棋高手进行快棋对决，连续 60 局无一败绩；2017 年 5 月，在中国乌镇围棋峰会上，它与排名世界第一的世界围棋冠军柯洁对战，以 3 比 0 的总比分获胜。围棋界公认阿尔法围棋的棋力已经超过人类职业围棋顶尖水平。

在我们前面举例的对小狗的“认识”以及此处阿尔法狗对下棋的“学习”中运用的深度学习技术都是建立在我们前文分析的图灵机对感性数计算的理论基础之上，因此人工智能能够实现所谓的“认识”与“创新”。

人工智能“认识”小狗以小狗的外形标准(感性数)为目标，“学习”下棋以输赢标准(感性数)为目标。学习算法设计者以这些标准为前提可以设置学习方向，通过大量的感性数计算不断的“修剪”错误认识“加强”正确认识，最终实现所谓的“认识”与“创新”。如何设置“修剪”和“加强”的标准是这些学习算法的根本区别，所以与其说这是人工智能的创新不如说是人工智能的实践，这个实践检验的是算法设计师对学习标准(包括修剪和加强的标准)的设置是否合理。因此，深度学习不是“创新”而是“验证”，是对算法设计师设置“学习标准”(感性数)的合理性的验证。

因此所谓的阿尔法狗与人下棋就是赤裸裸的作弊，就如同一个普通人拿着超级计算器和人类的顶尖心算大师比赛一样，这个普通人只需知道计算规则就可以轻松碾压最顶尖的心算高手。阿尔法狗赢了比赛它本身没有任何“感觉”，它只是一台机器运行了一段程序，“兴奋”的是那个躲在阿尔法狗后面“藏头露尾”的算法设计师。要想进行一场公平的对决那就给围棋大师配备一个人工智能团队，双方都设置一个“棋路学习”标准，让各自的机器“学习”相同的时间后对决，谁的机器赢了说明谁的学习标准设置的更合理以及谁对围棋的理解更深刻。

(**注：**

传统算法是编程人员编写确定的函数映射关系，而深度学习是编程人员在确定因变量(学习目标)后利用学习算法(决策树、线性回归和神经网络等)从大数据中确定自变量和映射关系。虽然深度学习是一种巨大进步，但是其计算的函数自变量和因变量无法超越感性数的范围，感觉数始终是不可定义不可计算的。

深度学习如何“进化”都只是学习算法的改进都只是对感性数的计算，其进步意义在于它模仿了人的理性对感性数的处理方式，在下文我们将为人工智的社会应用做范围界定，深度学习技术在它的应用范围内实现了自身的价值。

)

明白了深度学习不是真正的创新而只能算是理论验证后，我们再来看深度学习如果进行产品设计会陷入怎样的困境。

假如我们让人工智能来设计一件衣服则必然涉及到衣服的“漂亮”“帅气”等感觉数的处理，但是此感觉数无法像感性数(如小狗外形、输赢标准)一样精确定义，因此人工智能没有“学习方向”，自然就无法实现“设计创新”。

在可能的实践中，我们把当前社会流行的各种服装领口、袖口、布料、花纹、衣兜等设计素材设置为某深度学习算法的参数，然后计算机将这些素材随机组合成成千上万款服装，接下来算法需要对这些衣服进行挑选，去掉“丑”的款而给“美”的款加权重。“美”这一感觉数无法定义而不像“学习”下棋那样棋路“好坏”“是否能赢”能够明确定义，因此人工智能可以通过“学习”下棋而棋路越来越好赢率越来越高却无法通过“学习”设计越来越“美”的衣服。

同样的，创新“美味”的食物，设计“漂亮”的建筑，设计“好用”的其他产品都不是图灵机能够胜任的工作。因此，产品创新人工智能无法实现。

(注：

在商业活动中创新既有产品创新也有商业模式的创新，模式创新可简单看做日常生活中所说的方法创新，人工智能无法实现这二者的原因一致，因此不再赘述人工智能在商业模式创新上的局限性。

)

其次来说艺术创新，艺术的创新标准同样是感觉，相信大家都听说过伯牙子期的故事，“伯牙鼓琴，钟子期听之，方鼓琴而志在高山，钟子期曰：‘善哉乎鼓琴！巍巍乎若泰山。’少选之间，而志在流水，钟子期曰：‘善哉乎鼓琴！洋洋乎若江河。’钟子期死，伯牙破琴绝弦，终身不复鼓琴，以为世无足复为鼓琴者。”，钟子期说“若泰山”就是像泰山，怎么像呢？感觉像，听曲子的感觉和看泰山的巍峨雄壮的感觉一样，钟子期说“若江河”就是说听曲子的“感觉”像看江河宽阔奔腾的“感觉”。

再比如说书法艺术，蔡邕的《笔论》对书法艺术做了精彩总结，将书法艺术与感觉的关系论述的淋漓尽致。

书者，散也。欲书先散怀抱，任情姿(恣)性，然后书之；若迫于事，虽中山兔豪不能佳也。夫书，先默坐静思，随意所适，言不出口，气不盈息，沉密神采，如对至尊，则无不善矣。为书之体，须入其形，若坐若行，若飞若动，若往若来，若卧若起，若愁若喜，若虫食木叶，若利剑长戈，若强弓硬矢，若水火，若云雾，若日月，纵横有可象者，方得谓之书矣。

《笔论》中除了“若迫于事”的每一个“若”都说的是感觉像，无论是音乐、书法、雕刻还是其他的艺术形式，艺术的好坏标准只有一个那就是感觉，至于这个艺术美感背后的标准究竟是什么，古今中外的哲学家进行了大量讨论。西方哲学家从柏拉图到康德、黑格尔、鲍姆加登在对“美”认识时均未意识到美的直接标准是“感觉”，他们发展出一个哲学子学科美学来直接讨论“美”的背后标准，因此其观点犹如隔靴搔痒、莫衷一是。中国的哲学是以感觉为基础的，因此中国的哲学天然包含美感，中国的儒家讲究仁义之美包含为人处世修身齐家之美，道家讲究天人合一之美，禅宗讲究明心见性、解脱之美。中国的艺术如诗歌、绘画、书法、武术、建筑、服装等直接从中国哲学中汲取了对美的感悟从而发展出了高超的艺术水平。我特别要赞叹的是中华服装设计艺术堪称绝伦，《春秋左传正义》：“中国有礼仪之大，故称夏；有服章之美，谓之华。”，汉服之美，震撼心灵！

因此如果要论艺术的影响力，西方的艺术随着近代以来资本主义的扩张而影响力更大，但是如果要论艺术的高度，中华艺术冠绝天下。这就是为什么毕加索赞齐白石：“他是中国最了不起的一位画家，他的画技真的是太神奇了。画鱼时，明明只是用水墨，连颜色都没有上，但人们却从他的画中看到了活的鱼和流动的河水。我一直搞不明白的一件事，为什么那么多人都要来西方学习艺术，其实，最好的艺术在中国。”

关于艺术之“美”各家提出了不同的见解但是只能定性不能定量，其原因如我们前文所述感觉数是不可描述不可定义的。

(注：

一定会有人认为可以通过大数据给服装的“美”或者艺术的“美”等感觉下定义，但是西方的哲学家经过千年的实践总结(数据量够大，总结者够聪明)，美学也只是一个大概的定性且众说纷纭，没有人通过读美学著作就能成为艺术家的而那些有天分有“灵感”的不读也能成为艺术家。还是佛家悟的最透彻“不可说，不可说，一说即是错”。

因此艺术家总是独一无二的，从古至今天下只有一个王羲之只有一幅《兰亭集序》，即使王羲之本人也没有那种感觉再写出第二幅《兰亭集序》，将王羲之的《书论》和蔡邕的《笔论》对照会发现，对于书法艺术的感觉描述极限就是《笔论》而无法更进一步的精确描述。相对应的自然科学是以感性为基础的，牛顿虽然伟大但是他的发现都可以精确描述，自然后世就会有成千上万的人在对物理的理解上达到和超过牛顿。这就是为什么科学总是能不断前进，艺术却经常成为绝唱的原因。

)

因此，艺术创新是人工智能永远无法实现的。

最后来说理论创新，或许有人认为产品创新、艺术创新涉及到人的感觉的参与，自然科学理论创新涉及的对象是完全客观的，只是对自然规律的总结，只涉及到感性数的计算，人工智能总能实现吧。但是事实并非如此，因为自然科学的规律总结(如万有引力的发现)虽然只涉及到可定义数可计算数但是具体涉及的数的种类却需要人的感觉去发现。牛顿在发现万有引力时，首先他需要感觉(灵感)去发现事物之间的“运动”有共同性，其次要发现这个“运动”的共同性与物体的“质量数”和“距离数”有关系才能最终总结出万有引力定律。这个世界能够产生的数据是无穷的，我们统计所有物体距离地面的高度也是一种数据，统计所有牛的眼睫毛数量也是一种数据，只要我们想就能测出无数种数据，但是在这些数据中哪些是有关的以及有什么样的关系却需要我们的感觉(灵感)来把握，很多科学家一辈子都没有重大的成果，就是因此没有敏锐的“感觉”去发现那些关键的数以及数之间的正确关系，因此我们自然无法用人工智能去发现自然规律。现实可行的方法是科学家有了感觉(灵感)后把相关的数据测量出来，推测各项数的关系，最后把数的计算工作交给计算机，计算机只是一个辅助工具。

(**注**：

1) 未来在自然科学的理论创新上我们可以采取人发现关键数据、采集数据而人工智能“发现”数之间的正确关系的方式去发现新理论，也就是说人工智能这个工具在理论创新中发挥更多的辅助作用，但是让人工智能完全独立发现新理论是不可能的。

2) 理论创新的对象除了自然科学还有法律、经济、哲学等社会科学，人工智能无法实现自然科学理论的创新更无法实现社会科学理论的创新。

)

因此，人工智能也无法实现理论创新。

综上所述，创新类工作人工智能永远无法胜任。

4.2. 决策

人在生命中经常要做决策，如读什么专业、找什么工作、寻找另一半等都是人生的重要决策，人做决策的决定性依据就是人的价值观，人的价值观既有感性标准又有感觉标准。比如说一个女孩子找另一半，她会告诉介绍人想找什么学历，什么身高外貌，什么家庭背景，什么工作等等这些感性标准，但是按照这样的标准找来的候选人她通常会说我对这个人没感觉，最终她找了一个各项条件都不符合的反而非常相爱，虽然这样的例子不是全部但是在生活中经常发生，这就是感觉标准在起作用。对另一半的学历、家庭、外貌、工作等有要求也是为了婚姻生活更“幸福”，通过分析可以发现这些感性标准的制定依据是感觉标准，这就是人做决策时通常感觉会战胜感性的原因。因此，人作决策既有感性标准又有感

觉标准但归根结底是感觉标准，婚姻如此，学习、工作同样如此。

对于个人如此对于一个群体同样如此，一个城市做什么规划如何发展，一个民族一个国家如何发展既有理性规划(感性标准)又有感性考量(感觉标准)，这要求决策者既考虑理想又考虑实际，在做决策时灵活变通。这样的决策永远都只能靠人来做，而不能靠没感觉的人工智能。

因此，社会管理类工作人工智能永远无法胜任。

(注：

严格来说人的决策也包含只须感性参与的决策如在两点之间有多个路径时寻找一个最短路径的决策，本文中的决策特指有感觉参与的决策。

)

4.3. 人工智能的社会定位

通过前两节的分析发现创新与决策都是人工智能无法实现的思维活动，但是我们也不要对人工智能太失望，人工智能能够实现人的感性“学习”无疑是一种巨大的进步。当前社会的大部分工作都是只涉及感性的重复性劳动(熟练工)，不管是工厂还是公司，90%以上的工作只需人的感性参与就可以做的很好，在这些岗位上人工智能将大有作为。

以与我们生活息息相关的衣食住行的产品生产为例，未来消费者买衣服只需在网上挑选好款式，然后智能工厂即可按照款式和消费者的身体数据生产好服装，接下来的仓储、物流、签收等环节均可实现无人化(目前已经部分实现)，整个过程唯一需要人做的就是服装设计。未来的餐厅中掌厨、洗刷、上菜、收费等均会实现智能化，唯一需要人参与的就是菜品设计。未来的房屋建设中搬运、砌墙、封顶、粉刷等均会实现无人化，唯一需要人做的就是建筑设计、装修设计。至于出行，现在的无人驾驶汽车已经被各大公司争相测试，未来唯一需要人做的就是交通工具设计。

(注：

当然那些实现智能化的环节不是绝对的无人，比如机器的维修保养、制定经营规划等还是会有人参与，但是相比于现在的从业人数可视为无人化。

)

在只需感性参与的重复性劳动岗位，人工智能不仅高效而且高质量。比如拿汽车驾驶来说，无人驾驶汽车经过大数据训练后其驾驶熟练度要远超人类司机，而且醉驾、疲劳驾驶、路怒症等等情况都不会发生，其安全性、高效性都将大大提升。又比如在无人工厂机器人大幅提升生产效率而且生产质量会更高(出错率更小，精确度更高)。

所以在未来社会生产中，人工智能负责“制造”人负责“创造”。人工智能和封建社会的铁器农具、资本主义的机器一样依然是我们的生产工具，只是这个生产工具更加的安全高效高质，随着人工智能的全面应用未来社会的生产力将会得到空前提高。未来会有大量的人得以从繁复的工作中解放出来从事更有价值的创造性劳动，整个社会的创造力将得到极大的释放，人类在艺术、科学、文化、体育等领域将实现空前的发展，如果说当代社会创造了人类辉煌的物质文明的话那么未来社会将会创造人类璀璨的精神文明。

(注：

我们可以将未来的人工智能按照应用领域划分为生活机器人、生产机器人和军事机器人，在每一个领域人工智能都有巨大的应用潜力，在具体实施中本着需要感性参与就行需要感觉参与就不行的原则来确定其具体应用环节。

)

5. 人工智能的使用与开发原则

5.1. 第一署名原则

在上一章中确定了人工智能的社会定位是生产工具，但是这个生产工具和我们以往的工具毕竟有很大不同，最大的区别就是此工具的应用范围、自主灵活性要大大超越以往的工具，因此我们要确定它有可能带给社会的风险并加以防范。

使用工具的是人，而人具有使用工具犯罪的可能性，在未来社会用人工智能犯罪和用刀、枪支等工具犯罪在法律上没有任何区别(我们通过论证已经知道人工智能进化出人格是无稽之谈，企图用人工智能给人的犯罪当替罪羊是痴心妄想)，而人工智能这种工具又通常能实施使用者不在场的犯罪。

因此每一个人工智能机器出厂时必定要有有效的生产设计者的署名，交付使用后必定要有有效的使用者署名，否则将不能运行。如果人工智能机器造成了社会危害我们就有了追责的对象，若是使用者不遵守使用规范造成的危害就追求使用者责任，若是产品本身的问题导致的危害就追求生产设计者的责任。对于生产设计者的责任又可视情况分为故意犯罪和生产事故，这二者的区别定性我们可通过可追溯原则来实现。

5.2. 第二可追溯原则

有三种情况可视为生产设计者的故意犯罪。

1) 生产设计者为人工智能设定危害社会的任务目标，将人工智能作为犯罪工具实施某种犯罪。

2) 生产设计者设定不切实际的开发目标造成不可预知的风险。

虽然我在前文中不断论述感觉是不可定义的，但是我相信一定会有人不服气，一定会有人妄图靠着自己对感觉的“一知半解”但是自我感觉良好的去造个所谓的“人”出来，就好像能量守恒定律不管多正确都一定有人想要造出个“永动机”一样。

接下来让我从三个方面再说说为什么造“人”一定不可行。

首先人认识世界的唯一方式就是理性，我们对感觉的总结是理性对感觉的结论。理性得出的结论一定是建立在人已经表现出来的所有感觉现象之上，无论一个人多么天才的把人类感觉已经表现出来的所有现象总结出完美规律，我们都不知道人的感觉是否还会表现出新的特点。比如在古希腊时毕达哥拉斯学派将万物总结为数，无论前期理论再完美可是一旦根号 2 被发现“万物皆数”的理论就会崩溃，再比如牛顿的经典力学再完美的将地面运动和天体运动统一起来，一旦光速运动被发现经典力学理论也会崩溃。感性理论如此，感觉理论同样如此。我们的理性总结永远都是对已经出现的感觉现象的总结而无法包含那些未来可能出现的感觉现象，所以只要人类的实践一天没结束理论完善就一天不会停止。

其次建立在理性对感觉的认识之上的“模拟”和感觉本身是两回事。比如一个人假装高兴和他真正高兴完全是两码事。中庸说“天命之谓性”“喜怒哀乐之未发谓之中”[4]。真正的高兴来自“天命”即物本身，假装的高兴来自理性的认识，这二者的来源有本质不同，我们每个人的感觉都能分辨“真”“假”的不同。同样的人工智能无论怎样实现感觉都是建立在理性对感觉的认识上，所以人工智无论怎样模拟感觉，它都是程序而不是感觉，都是“假”而不是“真”。

最后，人可以实现对感觉的超脱。佛家的最核心思想就是“应无所住而生其心”“苦乐随缘，心无增减”，说明人是可以完全断绝欲望，抛弃痛苦与快乐这些感觉的。人工智能要如何实现？

(注：

中国哲学对感觉进行了大量探讨，比如儒家的中庸思想，道家的天人合一，佛家的不可说，迄今为止探讨成果只能算略知皮毛。我们当然不能总结了感觉数的不可定义性就什么都不做，我们可以从理论

上对感觉进行更多探讨。感觉虽然不能定量但是首先可以定性，中国的哲学对感觉的探讨进行了几千年有很多定性的成果，其成果对西方哲学以感性为基础导致的虚无主义(本体论的危机，休谟对因果律的质疑等)有重大意义，具体内容我会在完整的认识论中论述。

)

通过以上进一步的分析总结，我希望每个开发人员都不要贸然去造所谓的“人”，因为你无法预测你的行为结果。驱使一个人造“人”的无非是人的欲望，但是凭着一知半解造出来的不仅不是“人”反而是一个不可控的机器，就如同造了一个没刹车的汽车或者没保险的步枪一样，必定会给社会带来潜在危害。那么最终你不仅无法“功成名就”反而会像那个做基因编辑的贺建奎一样身败名裂成为人人喊打的过街老鼠。

3) 开发者不遵守行业开发守则

前文分析过人工智能不能做创新和决策类的工作，那么像社会管理、产品创新类工作我们当然可以交给人来做，但是人工智能在重复性工作中有时难免会碰到感觉数的处理(包含但不限于情绪模拟、决策模拟等)，比如无人汽车在路上行驶时前方有人或者动物，这时肯定不能为了高效省时直接撞过去而需要避让，但是肯定也不能因为有一只蚂蚁在路上也要避让，那么避让算法的相关原则就需要制定统一的行业标准，制定行业统一标准可以群策群力最大限度的避免社会危害，无人汽车的设计者在写避让算法时必须首先遵守这些原则。

因此每个行业都应成立组织制定本行业的统一标准，如开发者不遵守行业守则造成社会危害则属于故意犯罪。

因此，开发人工智能的第二个原则是可追溯原则，可追溯即程序执行过程的可追溯，通过制定统一的开发规则如写注释、返回每段程序执行结果等使人工智能的程序运行过程具有可读性，如果人工智能对社会造成危害就可以通过追溯程序确定人工智能的生产设计者是否涉及以上三种故意犯罪的情形。假如以上三种均不涉及则可视为生产事故。

生产事故分为两种，一种是行业守则不够完善导致的，那么就可以根据事故原因完善守则条款以避免下次发生此类事故。另一种是产品质量问题，那么追究生产者责任。

通过以上两个开发原则即可最大限度避免人工智能可能带来的社会危害。

最后回顾引言中的问题，为什么人工智能治疗抑郁症不仅不会好转反而病情会加重？原因显而易见，用人工智能治疗抑郁症，病人得到的都是机器人“有逻辑的”“例行的”询问，机器人模拟的感情都是“假的”，这种情况下抑郁症病人会强烈感受到被“应付”被“抛弃”的感觉，自然病情会加重。因此需要感觉参与的工作不是人工智能的开发方向，人工智能应该专注于只需感性参与的重复性的工作。从事熟练工作的人工智能，其价值必定是越高效越高质越好，而市场竞争会优胜劣汰将那些故障高或不遵守行业准则或不够高质高效的生产设计方淘汰掉，最终人工智能不仅更高质高效而且更安全，在这条道路上人工智能会给我们的社会带来最大的福祉。

参考文献

[1] Turing, A.M. (1937) On Computable Numbers, with an Application to the Entscheidungs Problem. The London Mathematical Society, London.

[2] Turing, A.M. (1950) Computing Machinery and Intelligence. *Mind*, **59**, 433-460.

[3] 康德. 纯粹理性批判[M]. 李秋零, 译. 北京: 中国人民大学出版社, 2004: 45-70.

[4] 四书[M]. 陈晓芬, 王国轩, 蓝旭, 万丽华, 译注. 北京: 中华书局, 2017: 150.

[5] 庄子. 庄子[M]. 孙通海, 译注. 北京: 中华书局, 2016: 39.

[6] 老子. 道德经[M]. 富强, 译注. 北京: 作家出版社, 2016: 3.

知网检索的两种方式：

1. 打开知网首页：http://cnki.net/，点击页面中“外文资源总库 CNKI SCHOLAR”，跳转至：http://scholar.cnki.net/new，搜索框内直接输入文章标题，即可查询；
   或点击“高级检索”，下拉列表框选择：[ISSN]，输入期刊 ISSN：2326-3415，即可查询。
2. 通过知网首页 http://cnki.net/顶部“旧版入口”进入知网旧版：http://www.cnki.net/old/，左侧选择“国际文献总库”进入，搜索框直接输入文章标题，即可查询。

投稿请点击：http://www.hanspub.org/Submission.aspx
期刊邮箱：airr@hanspub.org