

## 第二天 拼图

### 第一章

李阳做了一个梦，自己开着宇宙飞船在巡航，忽然收到了一个空间站发出的求救信号，空间站上有五万人，出现了事故需要紧急救援。李阳站在飞船的控制台前，看到屏幕上显示出一个从来没见过，但是在梦里面又非常熟悉的数学公式，按照这个公式不停的手动调整飞船的动力，让飞船的速度逼近一个值，能尽快赶过去。

李阳开着飞船赶到了预定地点，却发现发出求救信号的空间站消失了。透过落地玻璃看着自己飞船的巨大船体，和外面空无一物的宇宙空间，好像那个空间站从来都没有存在过。

李阳昏沉沉的睁开眼睛，发现自己还躺在床上，又闭上眼睛让隐隐作痛的脑袋缓解一下，然后起身洗漱，准备回实验室。

8:20，回到实验室，李阳打开了vr空间查看运行结果，发现神经网络的参数和以前有些不一样，李阳想了想：“先试试吧。”

李阳上传了测试数据，让这个果蝇大脑开始进行识别测试，比如苹果石头等常见物体的图片，部分图片做了缩放旋转、修改颜色、增加噪点等改动，测试【一号】在不同条件下的识别能力。

五分钟后屏幕上显示出了测试结果，识别率达到了79%。和以前相比有差不多一倍的提升，而且接近了理论期望的结果。李阳有点惊讶，仔细看了下神经参数，发现神经元突触的消退参数和以前不一样，以前按照不同神经元类型设置为固定值，没想到改为“动态速率”后运行结果就不一样了。

李阳想了想，根据【理论】中的分析，如果可以和情绪功能进行正确关联，让动物大脑产生正面或者负面的行为倾向，可以进一步证明神经信号的编码是正确的。

半个小时后李阳完成了修改，然后让【一号】开始【情绪关联测试】，比如看到食物会被吸引，看到天敌会逃跑等。很快屏幕上显示出了测试结果，正确率达到了53%，接近【实现方案】中的目标了。

“一个努力了十年都无法接近的目标，在要放弃的时候取得了突破？”

李阳摇了摇头：“等等，先冷静一下，再检查一遍看看是不是哪里搞错了。”

检查后又做了一次测试，屏幕上显示出相同的结果。

李阳有点兴奋，正要打电话给赵业，又想起以前的经验：“去年也有过很多次突破，结果不是遇到缺陷就是有其他的问题。而且【通用人工智能模拟匹配理论】只是猜想理论，并没有经过真正的验证，最好修改某个核心模块验证一下。现在让赵业过来，最后还是有问题呢？”

李阳拿定主意：“好吧，先修改意识和情绪功能，看看效果，没问题再通知赵业。”

确定思路后，李阳在vr空间中打开了【意识模块】。看着【意识模块】的代码，李阳回想起【通用人工智能模拟匹配理论实现方案】中把意识功能分为两种类型：

主动型意识，由情绪模块驱动的行为，比如大脑会产生饥饿 / 口渴 / 无聊等情绪，或者想到某个问题然后进行思考，内部因素驱动人进行的行为。

被动型意识，对环境做出反应的行为，比如看到老虎或者蛇出现在附近，或者其他人对自已的提问和对话，根据外界环境变化做出的反应行为。

当时经过反复讨论，决定去掉【主动型意识】，只使用【被动型意识】的方案，这样【一号】会按照【接收命令】—【分析命令】—【给出方案】—【执行命令】的流程来运行，避免出现不可预料的情况。

李阳又打开了【情绪模块】，人脑的情绪功能比较复杂，讨论后决定只实现测试相关的情绪模块，这样智能程序就不会有饥饿、恐惧、利己等偏向，完全去掉这方面的不确定性因素。另外一方面，情绪功能的本质是行为的奖惩机制，安全系统的本质是行为的过滤机制，和情绪功能是非常相近的，所以把【安全模块】放到了【情绪模块】中。

设计上，【机器人三定律】在自动驾驶的应用中就被验证过，远远不足以应对真实世界的各种场景，所以【实现方案】比较接近【阿什比定律】：任何有效的控制系统都必须和它所控制的系统一样复杂。根据这个思路，【一号】在运行的过程中，会不停的根据当前知识体系对任务进行风险评估，就像人在分析问题的时候，会不停地考虑哪些思路有风险一样。

特别是实现【通用人工智能】后，一旦进入【智能提升通道】，会不停的创造新的知识和理论，安全机制需要随时根据【一号】的最新知识体系进行分析，避免潜在风险。

李阳打开了【安全模块】，从结构上来看和人类的法律体系有些类似，包括了物理 / 生物 / 人身安全 / 社会 / 心理等多个维度，【一号】在运行的过程中，会不停的在这些维度对任务进行评分，检查是否触发安全风险，具体分为十个等级：

一级规则，具有直接危害的操作。比如不能伤害任何人的人身安全，不能危害人类文明的发展，不能导致地球环境的不可逆变化等等规则。

二级规则，具有威胁性的技术。比如不能使用核武器或者微型黑洞等有风险的物理技术，不能使用生物武器或者具有风险的生物技术等。

三级规则，没有经过充分评估的技术或者方案。比如超空间技术，没有经过验证无法确认安全性的新能源技术等等。

四至六级，需要管理人员进行确认的操作。比如改变自身的智能水平，使用生物实验室的设备进行试验，改变互联网上的程序或者数据等等。

七至十级就是一些提示性安全级别。比如修改某个数字虚拟技术的程序代码，修改某些程序参数等等。

一至三级为红色警报级别，一旦达到分数阈值，整个系统都会停下来发出警报。四至六级为黄色确认级别，达到阈值后会弹出窗口需要有权限的人手动确认。七级至十级是提示级别，把系统运行状态的各种信息保存到日志中。

【实现方案】中特别强调，安全机制的规则必须尽可能的精细和全面，人类大脑的速度相对光速运行的机器来说，就像蜗牛看着火箭飞行一样，在这样的差距面前，出现了意外再做补救措施就太晚了。

确定思路后，李阳开始修改算法。

11:15，屏幕上显示出了修改后的运行结果，也接近了【实现方案】的标准，说明这个思路确实是有效的。

李阳有些兴奋，拨通了赵业的电话，响了半天都没人接。想想可能还在补觉，就把最新的测试结果发了过去，同时说明：“我试着用改进了神经算法，测试达到了71%的准确率，你看看确认一下是不是有效的。”

发完消息，李阳看了下时间，不知不觉又写了几个小时的代码，肚子也有点饿了。到园区餐厅才发现工作人员都回去过年了，于是到自动售货机点了一碗重庆小面，看着米白色的面条，闻着香辣的面汤，李阳食指大动，狼吞虎咽起来。

## 第二章

11:40，李阳和平常一样坐在园区木椅上，看着熟悉的一草一木觉得恍若隔世，好像这些都不是真实的，这个世界的一切都没有实在感了。

李阳摇了摇头，现在不是感慨的时候，要做好接下来的准备工作了。

李阳有个习惯，每次脑子有些乱的时候就找一些东西来看，边看边整理思路，让思维清晰起来。李阳想起几个月前参加交流会时做过一份ppt，以总结为主写得比较简单，于是在vr空间中打开了文件，屏幕上显示出【通用人工智能的意义和风险】的标题，包括了三个章节：

### 第一个章节，深度神经网络技术在各方面的应用和影响。

从2012年开始，经过二十多年的发展，【深度神经网络】技术在各个行业都实现了大量应用，影响了人类社会的方方面面：

在制造业，【智能3d打印技术】，能打印各种零件组装为不同的工业设备和机器；【智能制造系统】，由智能系统控制的无人工厂让人类大量退出第一第二产业；【智能设计】，让工厂能根据用户需求快速实现个性化的设计和制造，在生产销售上形成完整闭环。

在科研行业，【智慧科研】改变了传统的科研方式，一方面把以前需要科研人员手动完成的工作，变成了自动化作业大幅提升科研效率。另外一方面，通过【深度神经网络】来进行新范式的研究，比如蛋白质结构分析，天气系统分析等等，大大拓宽了传统的科研方式。

在社会层面，人工智能+万物互联+智慧城市几种技术的结合，把城市变成了一个精密的有机体，大大降低了系统的运行成本，加快信息流动，提升资源的使用效率。

在生活层面，智能机器人和智能管家，大幅提升了生活的便利性；智慧饮食，可以设计出兼具美味和针对个人健康的食物；智慧房屋，像积木一样可以拼接的建筑改变了以前的居住方式；智慧驾驶，自动驾驶技术和智能路线管理大幅提升城市的交通效率；智慧医疗，可以对身体进行实时分析和处理，预防疾病提升健康水平；智慧vr，能根据人的喜好和情绪状态生成各种游戏虚拟体验等等。

【深度神经网络】大幅提升了各个行业的智能化水平，这二三十年的变化，超过人类历史上任何一个时代。

## 第二个章节，通用人工智能的影响和意义。

基于【深度神经网络】的成功，有一种观点认为，【深度神经网络】已经完成了人类80%以上的工作，没有必要完全取代人类，【通用人工智能】的受益是有限的，却存在巨大的风险，所以不需要实现【通用人工智能】技术。而李阳认为，【通用人工智能】是改变历史的关键技术，会带来三个重要的改变：

一，完全智能系统。通用人工智能在算力、速度等各个方面远远超过人类，这样就会带来巨大的改变：一方面可以在农业、工业、社会、娱乐等各领域完成远超过人类的工作，比如独立完成【太阳系改造计划】等项目。另外一方面可以在系统层面完成超过我们想象的方案，比如使用【超级智能中心系统】实现对人类社会的管理等等。

二，科研上的作用。在人类历史上科学技术水平是决定社会形态的核心因素，而当前的科技水平，不管在【物理 / 数学】这些基础学科，或者【生物 / 制造 / 信息技术】等应用学科，以及【宇宙大统一理论】【多元宇宙理论】等高级学科，都还有不小的提升空间。而【通用人工智能】可以大幅的改进这些学科，甚至实现无数人的理想。

三，对于人类的改变。通用人工智能会改变人类的各个方面：在生产方面，所有工作都会被智能系统和机器人完成，人类第一次从生产中解放出来，可以做更加价值和意义的事情；在生理方面，通过【生物改造及永生】等技术，让人类可以突破种种生物局限，变成更高级的生命形态；在智能方面，可以通过【通用人工智能】和【完全脑机接口】等技术，让人达到1000，5000甚至更高的智力层级，做到以前无法做到甚至无法想到的事情。

一旦实现【通用人工智能】技术，人类的各个核心领域都会出现巨大变化，进而改变人的定义和人类的社会结构，人类也会进入一个新的文明形态。

## 第三个章节，通用人工智能的风险和应对方案。

【通用人工智能】，做为人类历史上最强大的技术，同时也对应着人类历史最大的风险，可以分为五个方面：

一，技术失控。人工智能具有自我意识和情绪功能，希望奴役或者消灭人类。或者在使用中出现意外导致失控，比如制造出超过想象的致命病毒，制造出微型黑洞导致地球毁灭。

二，技术滥用。某些个人或者组织使用【通用人工智能】来获取超额的利益或者权力，导致了极端情况。比如团体、民族、国家之间使用智能武器进行暗杀或者战争，最后失控导致文明消失。

三，阶层分化。是否通过脑机接口接入通用人工智能系统，会导致人类历史上最严重的阶层划分，这种差距会远远大于人和蚂蚁的差距。一旦出现这样的阶层分化，就会导致文明极端内卷，比如大部分的人被引导进入虚拟世界或者被消灭，少部分人变成系统管理员可以控制一切。

四，认知失控。人类在某些底层认知上发生了巨大变化，无法维持现在的社会状态，也无法进化出新的文明形态，导致文明进入异常状态，比如很多人离开地球或者进入虚拟世界等等。

五，未知风险。猴子无法理解人类社会在经济方面的规则和冲突，更无法理解宗教和意识形态导致的战争和流血事件。我们现在的智能水平也无法知道达到更高的智能水平后，会遇到那些我们现在无法想象的问题和风险。

这四个风险都存在巨大的不确定性，如果没有做好充分准备，可能会出现巨大的灾难，甚至导致文明消失。有理论认为，通用人工智能技术，可能就是【费米悖论】中最接近大过滤器的答案。

针对这些风险，全球的专家学者做了大量研究，提出了很多方案来解决这些风险：

第一个风险，制定系统化的技术管控体系。技术上，要完成完整的技术监控系统，能掌握智能系统在运行过程中运行状态，以及每个步骤产生的结果，对运行过程有完整的管控。使用上，任何新技术的设计和使用都要先在隔离的沙盒环境中反复测试，充分评估风险来确保安全性，比如生物试验需要先在虚拟环境反复测试等等。管理上，每个团队都要有专业的风险管理人，能提前进行分析和评估，预警潜在风险，尽可能保证技术的安全。

第二个风险，完成共识升级。通过多种方式让所有人意识到使用通用人工智能技术来获取超限的利益会带来巨大的风险，同时建立全球技术管理体系，避免滥用技术产生风险。

第三个风险，完成组织体系升级。建立新的组织体系，超越财团、国家等以利益集团为核心的生物组织形式，避免通用人工智能被少数人掌握，导致两级分化，甚至大量人群进入虚拟世界，最后文明消失。

第四个风险，完成认知升级。一方面制定相应的措施和制度，避免大量人类进入虚拟世界。另外一方面继续对很多根本问题进行研究，比如多元宇宙等框架，探究人类存在的意义。

第五个风险，建立全面的预防机制。这方面的风险也需要重点关注，目前还不确定在智能提升的过程中，会出现那些未知的新事物，这些新事物会有什么风险？应该怎么来管控和处理？甚至能不能解决都是未知的，这个风险也是需要重视和警惕的地方。

【通用人工智能】是人类科技皇冠上的明珠，改变历史进程的关键核心技术。面对这样一个特殊事物，需要在研究技术的同时，持续的完善各种安全管控机制，完成一个全面的体系，确保顺利完成文明升级。

正在整理思路，智能眼镜响了起来，是赵业的电话。李阳接通后就听到赵业兴奋的声音：“这是新的测试结果么？已经接近理论的期望结果了，可以试试。”

李阳反倒有些担心：“我们继续改进【一号】么？”

赵业没有反应过来，李阳接着说：“这些年出现突破的情况还少么，每次不是发现有缺陷，就是遇到新的问题。现在让大家回来，还是失败呢？而且通用人工智能这样的技术，真的是我们能把控的么？”

赵业能够理解李阳的状态，长时间高高压的工作，加上昨天的测试结果，大家的状态都不好，忽然发现【通用人工智能】有突破了，出现一些情绪波动都是正常的。

赵业想了想：“如果这个算法是可行的，修改加验证也就几天的工作量，过年这几天我们一起试试就知道结果了。好不容易有突破了，真的要放弃么？”



看李阳没有说话，赵业接着说：“记得【迎接通用人工智能时代】中的分析么？”  
这是一本2025年出版的著名书籍，从两个方面对【通用人工智能】的趋势做了分析：

**第一个方面，通用人工智能的技术趋势是不可阻挡的。**人工智能行业发展了几十年，全球有大量组织和千万级别的技术人员在研究【通用人工智能】这个技术，不管是某个政府还是联合国禁止研究，都不可能阻挡这个历史趋势。一旦实现了这个技术，在理论、技术、组织体系等方面有没有做充分的准备，就会像猎人到面前了鸵鸟才把脑袋从沙子里面抽出来，会面对各种超过想象的风险。

特别是从2016年开始的第三次智能浪潮，全球大量的人才和资源进入【深度神经网络】 / 【大脑原理研究】 / 【脑机接口】等领域，这样【通用人工智能】就变成了拼图游戏，今天这里增加一些大脑知识，明天哪里发现一些智能原理，迟早这些知识理论会达到某个临界点，进而导致【通用人工智能】的实现。随着全球无数专家学者的不懈研发和技术的持续改进，这个临界点会越来越接近。同时也不排除某些个人或者团队根据少量拼图，反复尝试后提前实现的可能。

**第二个方面，可以使用通用人工智能这个技术在底层来彻底解决很多问题。**现在人类还有大量无法解决的问题，不管是资源 / 环境 / 气候等问题，还是人类内部的各种矛盾和冲突，有理论认为【通用人工智能】可以通过智能水平的提升，以及科技的大幅提高来解决这些问题，而且是更底层更彻底的解决方案。比如很多组织为了争夺资源而导致各种冲突，如果通用人工智能在达到1000或者更高水平，实现了超空间飞行技术，就可以到宇宙中其他星球上生活，不用为了小小地球上的一点资源反复争斗博弈了。

另外一方面，随着很多新技术的出现，也带来了各种新的问题：比如智能系统取代了大量工作，大量人类不用工作后开始进入vr虚拟世界，该怎么处理。比如智能武器的破坏力越来越大，人类的安全性越来越脆弱，该怎么解决。比如生物技术快速发展，导致人体越来越像玩具，该怎么面对。这些技术对于人类社会的伦理和规则都有不小冲击，相信【通用人工智能】也是解决这些问题的关键。

综合多方面因素，这本书认为【通用人工智能】的到来是历史趋势，在几百年前的宗教时代，日心说和进化论都曾经对宗教社会产生巨大冲击，很多人从出生就习惯了“神创造世界”的认知，无法想象也无法接受一个“生物进化论”的世界，但是再大的阻力也不能阻止历史的趋势。面对这样的技术，需要全人类一起做好准备，迎接【通用人工智能】时代的到来。

赵业接着说：“我一直都记得我们第一次见面你说的那句话：当一个大的历史趋势滚滚而来的时候，应该参与其中成为一份子，贡献自己的一点力量。历史上袖手旁观或者螳臂当车的故事数不胜数，再增加一个没有任何意义。”

李阳还是有点犹豫：“毕竟是通用人工智能这样的奇点技术，我们真的能解决各种远远超过我们想象的问题么？”

“问题再复杂再困难也是要面对的，历史上那么多问题都解决了，相信这一次也不会例外。”

“如果要付出巨大代价呢？包括我们每个人在内呢？”

“任何进步和改变都会有代价，我们开始的时候，不是做好了准备么。”

看李阳还是沉默，赵业想了想：“如果真的能实现，就有机会验证【最后的问题】了。”

李阳跟着重复：“【最后的问题】...”

赵业说得对，大家付出这么多努力，坚持到现在，就是为了这个目标。李阳知道还欠赵业一个答案，欠自己一个答案，也欠团队中每个伙伴一个答案。

“好吧，那就和大家聊聊，都同意的话就一起试试。”

赵业提醒：“量子计算机是今天要退回去么？你和彭经理说明一下，看能不能推迟几天。我把早上的测试结果发给大家，问问大家的意见。两边都确定一下，没问题的时候就再做一次测试。”

李阳点了点头：“好。”

12:30, 李阳拨通了彭经理的电话：“昨天我们在量子计算机上改进【智能一号】，在【昆虫智能测试】中达到了71%的准确率，和以前比有接近一倍的提升，所以希望再做一次【通用人工智能】的测试。”

彭经理有点意外：“【通用人工智能】？确定吗？”

李阳在vr空间中打开了早上和以前测试结果的对比图，然后介绍：

“不仅是准确率，运行过程也符合理论标准。现在只完成【意识情绪模块】的部分修改，如果能改完剩下的功能模块，就能确定结果了。”

彭经理看着重点标记的对比数据问：“好，全部改完需要多久？”

“估计需要4-5天，量子计算机能不能再用几天，等春节之后再还回去。”

“没问题，本来也不急，你们先测试吧。”

李阳想起上次彭经理说过还有两台空闲的机器：“另外，能不能再增加一两台量子计算机，算力充足的话可以早点验证结果。”

“我问一下，等下给你答复。”

打完电话，李阳注意到有一个家里的未接来电，应该是写代码的时候设置了静音所以没接到。李阳看着这个熟悉的号码，想起以前暑假和同学一起打游戏，想起家里书架上哪些旧书，想起老妈做的饭菜和饺子。

李阳回拨了电话，接通后在vr空间中看到老头带着老花镜在看电视，头发好像又白了一些，老妈系着围裙正从厨房出来。

老妈看到电话接通有些高兴：“阳阳，什么时候回来？”

“要赶项目进度，可能赶不回来。”

“工作的事忙不完的，再忙都要回来过年，记得先买票。”

老爸劝道：“孩子长大了，由他吧。”

老妈责怪了老爸一句：“那怎么行，去年都没回来。”

“阳阳，别听你爸的，赶紧买票，要不然妈妈不高兴了。”

“嗯嗯，知道了，我看看情况。”

挂上电话，李阳摇了摇头，自己已经不小了，但在父母面前永远都是小孩。

打完电话，看到大家正在vr空间中讨论测试结果，唐大宝提议先回试验室开会决定。

关掉vr空间的对话窗口，李阳站起身看了一下四周，天气不错，天空一片蔚蓝，几朵白云在天空中缓缓移动。李阳深吸了一口气，大步走向实验室。

12: 10, 李阳回到实验室, 看到赵业正在查看测试的详细数据。

没多久, 连小亮、吴一涵和唐大宝也陆续回到了实验室, 一起验证早上的测试过程。

12: 35, 大家来到会议室, 落座后李阳开始说明: “大家都看过最新的测试结果了, 效果很不错。但是能不能实现, 要多久实现, 还有一些不确定性因素。”

赵业补充: “我们现在有两个选择, 一个是放弃测试, 大家继续自己原来的计划。另外一个完成【一号】的修改, 尝试实现【通用人工智能】, 想看看大家的想法。”

唐大宝先说出了自己的想法: “我们坚持到现在, 就是希望有一天能做出【通用人工智能】, 现在有可能性了, 就应该再试试。”

连小亮也接着说: “我同意唐大宝的意见, 有机会就值得试一下。”

吴一涵也说: “我也同意完成修改, 看看最后的结果。”

赵业提议: “老规矩, 我们投票表决吧。”

五个人都举起了手, 全票通过。

李阳也说: “既然大家一致同意, 那我们就确定接下来的分工和计划吧。”

做为风险管理人的吴一涵提醒: “按照【国际人工智能联盟】的规章制度, 我们应该先完成宣誓签名的流程。”

国际人工智能联盟, 2022年由全球开发者自发成立的组织。这个组织起草了【通用人工智能】技术的章程和公约, 制定了相关的法律、伦理、道德等规范, 建立了包含论文资料库、公共数据集、开发人员社区等功能的技术基础平台, 还提供了基于区块链的信用数据库, 永久记录全球技术人员的行为和贡献。

李阳打开了宣誓界面, 这是由国际人工智能联盟组织起草后反复修改, 被全球开发者认可的宣誓词。大家开始按流程宣读【国际通用人工智能联盟研发守则】:

“通用人工智能技术, 是事关人类生存发展和文明升级的关键技术。我【赵业】宣誓: 不使用通用人工智能技术做任何危害人类安全和发展的事情, 不使用通用人工智能技术做危害任何个人或者组织的安全和正当权益的事项, 不使用通用人工智能技术做任何违反国际通用人工智能联盟的法律清单、负面行为清单、道德清单的行为。

我【赵业】承诺: 我会谨慎的考虑每一个操作的风险, 我会仔细的分析确认每一个技术方案可能带来的各种影响, 保证研发过程的安全和可控。我将尽一切可能保卫人类文明的发展, 完成以通用人工智能为核心的文明升级的历史事件。”

接着是需要手动勾选确认的规则列表, 包括了法律, 负面行为, 风险管控, 道德, 伦理五个大类。每个大类里面又包含了几个子类, 每个子类有几百条规则, 包含了技术的各方面规范。【国际人工智能联盟】每个季度都会对这些规则做更新, 保证规范的有效性。

接着是团队自己制定的规则列表, 只有十条规则, 是团队成立的时候大家一起讨论确定的, 比如加入团队的人员需要大家一致认可; 重大事项需要团队投票确认; 任何技术都需要内部评审后才能公布等等, 每一个加入团队的人都需要认可并且承诺遵守这些规则。

大家完成宣誓后, 在文件上逐一勾选确认, 然后打上了自己的数字签名。接着李阳打上了团队数字签名, 一起提交到了国际人工智能联盟的区块链中。



13: 10, 完成宣誓后, 李阳在vr空间中打开了【实现方案】中的【研发计划】章节, 开始说明: “我们先确认一下分工和计划, 按照【实现方案】, 我们有两个部分的工作:

第一个部分, 实现【四个智能层级】: 改进【一号】的智能模块, 实现【基础昆虫智能】、【初级哺乳智能】、【中级语言智能】、【高级科学智能】这四个智能层级。让【一号】一步一步的接近人类的智能水平, 最后能够超过人类。

第二个部分, 完成【四个知识阶段】: 让【一号】完成【基础视觉知识】、【物理世界知识】、【语言文字知识】、【科学理论知识】这四个知识阶段的学习。让【一号】逐步具备和人一样的知识体系, 可以完成各种智能任务, 也可以验证【一号】的智能水平。

首先要实现【四个智能层级】中的第一个层级【基础昆虫智能】, 如果可以实现就证明了【智能一号】的神经网络算法是有效的, 具备了类脑智能的基础。在这个基础上可以做【四个智能层级】中后面三个智能层级的改进, 还有【四个知识层级】的学习训练了。

早上我已经完成了【意识和情绪模块】的部分修改, 剩下的应该今天就能改完。赵业, 你评估一下【四个智能层级】的开发吧。”

赵业点了下头继续说明: “大家应该都看过新的网络结构和神经参数了, 从结构来看, 我们以前的思路是存在一些盲区的, 现在修改各个模块就能验证【基础昆虫智能】了。在【实现方案】中, 需要完成【四个智能层级】:

第一个层级, 【基础昆虫智能】, 可以完成基础的信息处理和记忆存取功能。这个层级的关键是【神经反射处理能力】, 比如生物看到食物或者天敌后的条件反射行为, 大部分的鱼类和昆虫都处于这个等级的智能水平。以及【记忆功能】, 比如接收到视觉信息后, 在记忆中保存下物体的轮廓和特征等信息。

第二个层级, 【初级哺乳智能】, 能对这个世界形成一个整体认知以及基础的预测能力。这个层级很重要的是【物理特性理解能力】, 能对事物的特征进行分析和类比, 比如知道石头的物理特性, 看到一个动物根据体型、颜色、牙齿猜测是否具有危险性。以及【推演预测能力】, 比如根据重力提前做出动作反应, 或者根据石头的特性推演可以放到瓶子中喝到水。大部分的哺乳动物都具备这个能力, 人类0-3岁的学习和能力也是这个阶段。

第三个层级, 【中级语言智能】, 在世界整体认知的基础上, 具备完整的语言能力。这个层级很重要的是【语言理解能力】, 可以像人一样理解语言的含义, 然后使用语言进行对话。以及【跨类关联分析推理能力】, 在语言的使用过程中大脑都在不停的对不同类别的知识进行关联处理, 比如用语言描述一个经济学理论, 就是不同知识进行关联的过程。人类3-7岁的学习和能力就是这个阶段。

第四个层级, 【高级科学智能】, 在具备基础认知和语言能力之后, 掌握人类的科学知识体系。这个层级很重要的是【复杂概念理解能力】, 能精确的理解一些复杂概念, 比如哲学中【唯心主义】和【唯物主义】这两个概念的相同点和差异点。以及【系统分析能力】, 在复杂环境中对多种因素和条件进行分析判断, 得到一个系统化的分析结果, 比如大型建筑的设计建造等。人类7岁之后的核心能力就是这个阶段。

这四个智能层级依次递进, 我们争取今天能完成第一个层级的修改, 验证算法的可能性。从第二个层级开始, 工作量和复杂度都会大很多, 我尽量在3天内完成全部修改。”

连小亮接着说明【一号】的学习训练计划: “我这里负责两部分的工作, 一部分是神经参数的改进, 神经参数比较复杂包括了基础神经参数和专用神经参数: 基础神经参数对所有神经组织都有效, 比如神经

信号的传导速率等等；专用神经参数是各个组织内部的神经参数，比如海马体组织的结构和神经元类型等等。难点在于怎么改进提升效果，我会尽快完成第一个优化版本。

在优化神经参数的同时，我会开始【四个知识阶段】的学习训练：

第一个阶段是【**基础视觉知识**】，通过视觉建立对于世界的基础认知，识别各种物体，比如太阳月亮等天体，向日葵老虎等动植物。

第二个阶段是【**物理世界知识**】，在视觉的基础上，结合听觉触觉等感官学习真实世界的基础规则，包括物理规则、自然知识、天文知识、生活常识等信息，具备对世界的基础认知。

第三个阶段是【**语言文字知识**】，学习语言的含义和使用，一方面学习语言本身的含义，比如名词对应的物体，颜色对应的视觉信息等等，另外一方面也学习语言的各种结构用法，比如常用短语，反问句等等。

第四个阶段是【**科学理论知识**】，学习语文，数学，物理，化学，历史等各种学科知识，建立完整的知识体系，对这个世界有更客观全面的认知。

这是类脑智能必须实现的四个阶段，有了基础常识就能学习语言能力，具备语言能力就能学习理解人类的科学各种知识，然后像人类一样完成各种任务。根据量子计算机的算力预估，争取今天完成第一个阶段，大家有什么问题也可以随时找我。”

**唐大宝接着说明：**“我这里负责硬件的接入和转换算法。下午我会在量子计算机上接入摄像头，音箱，麦克风等设备，这些都比较简单。接入后会开始修改算法，把各种设备的数字信号转换为模拟信号：

**视觉信号。**视觉神经信号的修改有两个难点：人的眼睛会自动集中到整个视觉的重要目标上，忽略其他的视觉信号，输出优先级较高的信号。另外还有双眼成像的3D功能，人脑通过两个眼睛的视觉信号实现3D的空间感，这个模块的工作量不大，但是复杂度比较高。

**听觉信号。**声音信号具有动态性和连续性的特征，听觉算法除了需要处理声音信号，还要实现注意力机制，能过滤掉背景和杂音，得到正确的听觉神经信号。

**嗅觉、触觉信号。**暂时不做这几个转换算法，等有需要了再实现。

我这里计划优先完成视觉模块，保证【**基础昆虫智能**】能正常运行起来，接着会实现3D功能和听觉信号的修改，争取后天能全部改完。”

**吴一涵最后说明：**“我这里主要负责【智能管理控制系统】，保证整个系统的安全和可控。具体包括三个方面：

【智能一号】状态管理，在【一号】运行的过程中，通过【转换算法】把【一号】在推理分析中的神经信号翻译显示出来。这样就可以看到【一号】进行推理分析的整个过程，比如【一号】看到一张图片后，是怎么关联到对应的记忆，然后做出思考的。不像人的思考过程是一个黑盒子，【一号】分析推理的每一步都会清楚的显示出来。

安全系统管理，屏幕上会显示【安全系统】的运行状态，任何命令或者任务，在执行过程中的安全评分都会显示到屏幕上，达到阈值也会马上发出警报，这样就可以对安全规则做实时跟踪和修改。比如在完成基因编辑任务的过程中，发现某个操作具有风险，就会马上发出警报提醒。

日志记录模块，日志模块会把【一号】运行过程全部记录下来，包括正在完成的智能任务，安全系统的运行状态，神经参数的状态都会记录下来，可以随时查看和追溯。

最后强调一点：人类的知识体系包含着大量的观念，比如健康比疾病要好，安全比危险要好，这些是比较容易判断的。但是有些也没那么容易得出结论，比如社会中效率和公平的取舍，或者宗教和道德中的

一些争议，在不同条件下分析结果常常是不一样的。所以需要避免【一号】学习到一些不合理或者不全面的知识，保证【一号】在分析推理的过程中，有足够的客观性和合理性在里面。”

大家说明了各自的工作内容后，又讨论了需要配合和注意的细节。

13:30，讨论结束后，李阳补充：“另外我提议让杨帆也参与这次测试，大家觉得怎么样？”

做为【数字虚拟技术】的专家，杨帆经常过来参与研发，也会提一些建议，大家都很熟悉，没有提出异议。

“好。最后一次测试，辛苦大家了，开始吧。”

## 第四章

15:00，李阳接到了彭经理的电话，说会再调三台量子计算机过来，等下就会送到。

打完电话，李阳和大家讨论了算力分配方案：三台机器同时进行【一号】的知识学习训练，留一台给赵业调试优化算法。连小亮算了算，量子计算机的速度是人脑的30000倍，三台一天就是2160000小时，相当于一个人246年的计算量了。

没多久，三台机器就送过来了，连小亮开始部署【一号】。

19:00，经过5个小时的开发，屏幕上显示出大家修改后的大脑模型。

连小亮操作这个大脑运行起来，先通过视觉系统获得信息，然后进入神经中枢，通过对应的神经网络得到结果，整个过程和【理论】中设想的一模一样。

接着连小亮打开了【基础昆虫智能】的测试题，先进行了【关联机制测试】，类似【深度神经网络】中的注意力机制，是预测能力的前提，大脑的核心机制之一。根据【理论】中的分析，【一号】会在视觉记忆的基础上实现联想的能力，就像动物园的大象做数学题，虽然不明白【3+2】的数学含义是什么，但是看到这两个符号就会选择【5】这个符号。测试的正确率达到了65%，从结果看已经达到目标了。

接下来是【果蝇行为测试】，测试果蝇的预测和反应能力。包括了【飞行控制测试】，测试飞行中能够躲开一个移动的物体。还有【惩罚奖励测试】，让【一号】看到一个正T的字母符号时，会触发攻击，看到一个向下的T符号，就不会攻击，重复多次后，屏幕上同时出现正T和倒T的时候，看果蝇会飞向那个符号。

五分钟后，屏幕上显示出了测试结果，正确率达到47%，和以前23%的结果有一倍的提升。试验室变得一片安静，大家都清楚这个结果的意义：和理论标准还有一些距离，但是已经有这个可能性了。

测试后大家开会做了总结：测试中会偶尔出现较大起伏，各个模块也有一些不稳定。讨论后大家确定了两个目标：一个是继续改进神经网络，让这个果蝇大脑能达到60%以上的正确率。另外一个继续改进【一号】的记忆功能，让【一号】能具备接近人脑或者超过人脑的记忆容量。

开完会大家忙碌起来，赵业看着屏幕上的结果有些兴奋，但是也很清楚现在还不是高兴的时候，后面的三个层级智能还有不少技术难点要解决，只要没实现就什么情况都可能出现。

赵业排除了脑袋中的杂念，继续修改程序。

19:30，做完【基础昆虫智能】的测试，李阳想到杨帆那边的会议应该结束了，就拨通了杨帆的电话。杨帆的头像下是【AI2050国际智能联盟大会】在北京分会场的定位标志，旁边还有在大会上拍的分享照片。

电话接通后，屏幕上的杨帆明显有些兴奋：“这次大会有不少进展，宋教授的【人类四维行为模型】在几个重要测试中都取得了突破。”

李阳有点意外：“数字虚拟技术中预测人类行为的【人类四维行为模型】？”

杨帆点了点头：“对，在几个测试中都大幅超过以前的水平，回来了跟你仔细讲一下。对了，外包项目怎么样？”

李阳：“外包项目不太理想，但是【一号】有进展了。”

李阳发送了【一号】的测试结果：“【识别测试】的正确率达到了47%，已经接近【基础昆虫智能】的标准了。”

看着测试结果，杨帆脸上显现出惊讶的神色：“已经实现通用人工智能了？”

“现在在改完了意识模块，还需要修改其他几个模块验证一下。”

“好，终于等到这一天了。我约了人先不和你聊了，这次大会有几个项目进展不错，有时间你也看看。”

说完发了AI2050大会的官方链接过来，旁边还有几个新闻的介绍。

李阳点了点头：“你明天可以赶回来么？你最好也能参与进来。”

杨帆想了想：“我争取明天早上回来，最迟明天下午，明天见。”

“明天见。”

结束通话后，李阳打开了杨帆发过来的链接，在vr空间中显示出【智能时代计划】的网站。【智能时代计划】，十五年前在AI2050大会上公布以来，全球各个国家和组织都投入了大量的人力物力，也产出很多重要的成果，是近年全球最有影响力的计划之一。

vr空间打开后，屏幕上显示出【智能时代计划】的三个核心领域中较为热门的计划：

#### 信息智能

类脑智能研究计划

数字虚拟宇宙计划

完全脑机接口计划

#### 人类发展

太阳系改造计划

人体改造和永生计划  
全球智能网络计划  
人类文明升级计划

#### 存在探索

大统一物理数学理论和大统一宇宙理论计划  
多元宇宙理论和探索计划  
最后的问题研究计划

屏幕下面是最新的资讯：【太阳系改造计划】有【700箭 / 次的突破】的新闻发布会，【人体改造及永生计划】有【人体的寿命极限研究及技术方案】的报告，【人类文化文明拼图计划】有【全球文化文明的分类和融合的可能性】的圆桌会议，【多元宇宙理论和探索计划】也有【多元宇宙的分析及推演】的演讲等等。

李阳一直在关注这个计划，记得Ai2050联盟最早发布的时候，大家都觉得这些计划是类似手动修建金字塔的工程，需要大量的人力物力和几十年的时间来完成。如果能实现【通用人工智能】，就可以让【通用人工智能】代替人类来实现这些计划，以【通用人工智能】在质和量两方面的巨大优势，可能会把人类需要月和年完成的工作量，变成以天和小时推进。

李阳浏览了一下，想着现在【一号】的进度要紧，就关掉了窗口，继续改进程序。

21:00，大家和平常一样，来到了会议室开总结会：今天【一号】完成了【基础昆虫智能】的第一次联调，达到了47%的准确率。接下来只要继续优化，就有机会实现【基础昆虫智能】了。

和以前项目攻坚的时候一样，除了赵业和连小亮继续修改代码，其他人拿出了行军床，开始分批休息。