



中国石油大学 (华东)
CHINA UNIVERSITY OF PETROLEUM

《计算机科学导论》课程总结报告

学生姓名: 陈泊蓂

学 号: 2007010207

专业班级: 本研人工智能 2001

学 院: 计算机科学与技术学院

课程认识 30%	问题思考 30%	格式规范 20%	IT 工具 20%	Latex 附加 10%	总分	评阅教师

2021 年 1 月 2 日

1 引言

时光飞逝，在不经意间就已经进入石大里学习了快一个学期。从考完高考之后的放松，再到填报志愿时候的迷茫，我最终选择了计算机这门专业，并选择了位于山青岛的中国石油大学来学习深造，从那时起我也与计算机之间连接起了一条看不见也摸不着，但却十分牢固的一条线。我对计算机最早的了解是在我小学一二年级的时候，在学期结束之后漫长的暑假之中，我跟着我的父亲到他所任职的学校去看看。在一番闲逛之后，我进入到了父亲的办公室之中，这是我第一次了解到所谓的电脑这一事物。在父亲的指导之下，我学会了如何启动与关闭计算机和如何用计算机来上网浏览各种各样的消息以及玩一些网页小游戏，从此我也算正式了解到计算机这一项新奇的事物。

就这样一直到了填报计算机为我的首选志愿，在此前我对计算机的认识就只是用来上网游戏和用来做一些 PPT，Word 文档这些简单的办公软件，其他的关于计算机究竟是如何组成以及其究竟是如何来运行的都未曾去了解过。在选择专业时我也是没有打定好自己的主意的，但从父母中了解到了计算机的工资高以及好就业的前景，并在我平时经常使用电脑的情况下积累出的对计算机的兴趣，我最终选择了计算机这门专业来学习。

在刚进入石大时，面对老师们的提问：你为什么选择计算机啊？你对计算机的哪方面感兴趣啊？诸如此类的一系列问题，在我仔细思考后我陷入了迷茫，开始疑惑自己选择计算机是不是正确的选择。但随着这几个月来对计算科学的学习，我受益匪浅。在老师们的讲解之下，让我渐渐的明白了究竟何为计算机，我也从当初一点不懂的小白到现在最起码对计算机有这样一定的认识。在计算科学导论的学习之中，我对我自身有着更加深刻的理解，还要对计算机这个行业有了更好的认识，对自己未来的就业方向也有了大致的了解。

2 对计算科学导论这门课程的认识、体会

在刚拿到计算科学导论这本书时，出于好奇心我翻开目录来看看里面究竟讲了哪些内容，发现了里面的内容十分多样，包含的知识涉及到了计算机的各个方面。在第一次上课时，孙老师就和我们说了在我们有空的时候多读读这本书，虽然以我们现在专业知识水平并没有办法能够将其理解明白，但老师还是希望我们能够将其看完并鼓励我们大三大四后再回来将其读一遍，到那时候我们的收获会更大。导论可以说是对于我们这些刚刚接触计算机的小白来开设的课程，在我们还没对计算机这整个学科有着深入了解的情况下，从计算科学的概念和知识来让我们建立起对计算机这一学科的基本框架，并在介绍时引入了很多有趣的例子来加深同学们的印象，让同学们有着更好的兴趣来学习计算科学导论。

2.1 计算科学与计算机科学

在课堂上，孙老师提出了一个非常有意思的问题：为什么书是计算科学而不是计算机科学，这两个词之间究竟有何区别呢？这个问题让我产生了很大的好奇心。在《术语在线》[2] 中，将计算科学定义为 (1) 计算机科学的又称 (2) 用计算的方法研究（物理、化学、生物等）其他学科领域的科学，如计算物理学、计算化学、计算生物学等。将计算机科学定义为在电子计算机上实行计算所涉及的数学理论和逻辑方法体系，以及其应用于解决实际问题的方法和技术。在我

的认识之中，计算机是为了满足人们对于计算数据的需要而设计出来的。计算机是由二进制来实现的，可以说计算机无论是运行程序还是执行某一项任务，都离不开计算这一核心问题，而计算机科学可以说是具体到了计算机的各种领域之中，计算科学是将计算机科学包含在其中的。计算科学可以说是为计算机的诞生与发展以及人工智能技术的发展提供了理论支撑。计算科学也为人们来更好的认识我们这个世界提供了理论依据。

2.2 哲学家就餐问题

这是一个很有意思的问题，问题大致是这样描述的：五个哲学家共用一张圆桌，分别坐在周围的五张椅子上，在圆桌上有五个碗和五只筷子，他们的生活方式是交替的进行思考和进餐。平时，一个哲学家进行思考，饥饿时便试图取用其左右最靠近他的筷子，只有在他拿到两只筷子时才能进餐。进餐完毕，放下筷子继续思考。由于只有五根筷子，每根筷子只能给其相邻的两位哲学家共享的，所以当一只筷子被使用时，旁边的几个哲学家是不能使用。

这是计算机中的死锁问题。所谓死锁，就是每个线程都等待其他线程释放资源从而被唤醒，从而每个线程陷入了无限等待的状态。在哲学家就餐问题中，一种出现死锁的情况就是，假设一开始每位哲学家都拿起其左边的筷子，然后每位哲学家又都尝试去拿起其右边的筷子，这个时候由于每根筷子都已经被占用，因此每位哲学家都不能拿起其右边的筷子，只能等待筷子被其他哲学家释放。由此五个线程都等待被其他进程唤醒，因此就陷入了死锁。

那么该如何避免死锁问题呢？

1. 仅当一个哲学家左右两边的筷子都可用时，才允许他拿筷子 2. 给所有的哲学家编号，现制定一个规矩，规定奇数号的哲学家要先拿起他左边的筷子，然后再去拿他右边的筷子，而偶数号的哲学家则相反。

哲学家就餐问题是在计算机科学中的一个经典问题，用来演示在并行计算中多线程同步时产生的问题。在网上查阅资料了解到其他人的思考过程和解决方式，我觉得我的思考问题方式有了一定的提升。之前思考问题通常都是没有着比较整体的，系统的框架来逐层挖掘问题的解决方法。而在学习到别人解决问题的明确方向后，我今后也将多在这方面来进行努力以提高自己。

3 进一步的思考

我们小组的分组演讲主题是“人机对弈——围棋”，在演讲中我们主要讲述了计算机在学习如何下围棋这一方面的发展历程，讲述了计算机是如何学习围棋的并挑选出了其实现学习的一些算法来简要讲述，并简单分析了运用在让计算机学习围棋的方法在未来的应用以及其前景。在进行回答问题的过程中，老师提出的一些问题给我留下了深刻的印象，接下来就对老师对我们演讲内容所提出的疑问进行深入挖掘与思考。

3.1 强化学习

强化学习（Reinforcement Learning, RL），又称再励学习、评价学习或增强学习，是机器学习的范式和方法论之一，用于描述和解决智能体在与环境的交互过程中通过学习策略以达成回报最大化或实现特定目标的问题。强化学习是智能体以“试错”的方式进行学习，通过与环

境进行交互获得的奖赏指导行为，目标是使智能体获得最大的奖赏，强化学习不同于连接主义学习中的监督学习，主要表现在强化信号上，强化学习中由环境提供的强化信号是对产生动作的好坏作一种评价（通常为标量信号），而不是告诉强化学习系统（RLS）如何去产生正确的动作。由于外部环境提供的信息很少，RLS 必须靠自身的经历进行学习。通过这种方式，RLS 在行动-评价的环境中获得知识，改进行动方案以适应环境。

AlphaGoZero 使用新的强化学习方法，让自己变成了老师。系统一开始甚至并不知道什么是围棋，只是从单一神经网络开始，通过神经网络强大的搜索算法，进行了自我对弈。随着自我博弈的增加，神经网络逐渐调整，提升预测下一步的能力，最终赢得比赛。更为厉害的是，随着训练的深入，阿尔法围棋团队发现，AlphaGoZero 还独立发现了游戏规则，并走出了新策略，为围棋这项古老游戏带来了新的见解。

强化学习与有监督学习、无监督学习的比较

在强化学习初步介绍 [1] 中提到，（1）有监督的学习是从一个已经标记的训练集中进行学习，训练集中每一个样本的特征可以视为是对该情况的描述，而其标记可以视为是应该执行的正确的行为，但是有监督的学习不能学习交互的情景，因为在交互的问题中获得期望行为的样例是非常不实际的，代理只能从自己的经历中进行学习，而其中采取的行为并不一定是最优的。这时利用 RL 就非常合适，因为 RL 不是利用正确的行为来指导，而是利用已有的训练信息来对行为进行评价。

（2）因为 RL 利用的并不是采取正确行动的经验，从这一点来看和无监督的学习确实有点像，但是还是不一样的，无监督的学习的目的可以说是从一堆未标记样本中发现隐藏的结构，而 RL 的目的是最大化奖励信号。

（3）总的来说，RL 与其他机器学习算法不同的地方在于：其中没有监督者，只有一个奖励信号；反馈是延迟的，不是立即生成的；时间在 RL 中具有重要的意义；代理的行为会影响之后一系列的数据。

强化学习介绍中 [5] 提到，RL 采用的是边获得样例边学习的方式，在获得样例之后更新自己的模型，利用当前的模型来指导下一步的行动，下一步的行动获得奖励之后再更新模型，不断迭代重复直到模型收敛。在这个过程中，非常重要的一点在于“在已有当前模型的情况下，如果选择下一步的行动才对完善当前的模型最有利”，这就涉及到了 RL 中的两个非常重要的概念：探索和开发：探索是指选择之前未执行过的行为，从而探索更多的可能性；开发是指选择已执行过的行为，从而对已知的行为的模型进行完善。RL 非常像是“试错学习 (trial-and-error learning)”，在尝试和试验中发现好的规则。就比如下图中的曲线代表函数 $f(x)$ ，它是一个未知的 $[a,b]$ 的连续函数，现在让你选择一个 x 使得 $f(x)$ 取得最大值，规则是你可以通过自己给定 x 来查看其所对应的 $f(x)$ ，假如通过在 $[a,0]$ 之间的几次尝试你发现在接近 x_1 的时候的值较大，于是你想通过在 x_1 附近不断的尝试和逼近来寻找这个可能的“最大值”，这个就称为是开发，但是 $[0,b]$ 之间就是个未探索过的未知的领域，这时选择若选择这一部分的点就称为是探索，如果不进行开发也许找到的只是个局部的极值。“探索”与“开发”在 RL 中同样重要，如何在“探索”与“开发”之间权衡是 RL 中的一个重要的问题和挑战。

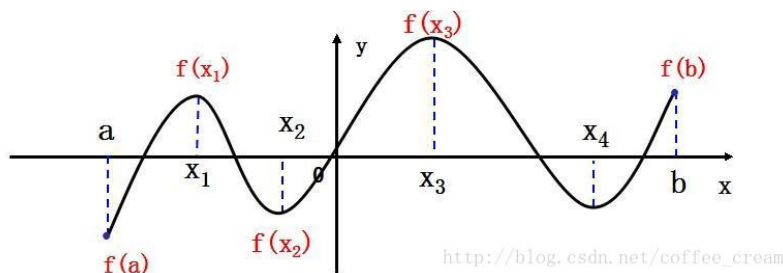


图 1: RL

3.2 蒙特卡洛树搜索法 (MCTS)

蒙特卡洛树搜索又称随机抽样或统计试验方法，属于计算数学的一个分支，它是在上世纪四十年代中期为了适应当时原子能事业的发展而发展起来的。传统的经验方法由于不能逼近真实的物理过程，很难得到满意的结果，而蒙特卡洛树搜索方法由于能够真实地模拟实际物理过程，故解决问题与实际非常符合，可以得到很圆满的结果。这也是以概率和统计理论方法为基础的一种计算方法，是使用随机数 (或更常见的伪随机数) 来解决很多计算问题的方法。将所求解的问题同一定的概率模型相联系，用电子计算机实现统计模拟或抽样，以获得问题的近似解。

蒙特卡罗算法采样越多，越近似最优解。举个例子，假如筐里有 100 个苹果，让我每次闭眼拿 1 个，挑出最大的。于是我随机拿 1 个，再随机拿 1 个跟它比，留下大的，再随机拿 1 个……我每拿一次，留下的苹果都至少不比上次的小。拿的次数越多，挑出的苹果就越大，但我除非拿 100 次，否则无法肯定挑出了最大的。这个挑苹果的算法，就属于蒙特卡洛算法。告诉我们样本容量足够大，则最接近所要求解的概率。

随机数

[3] 上述蒙特卡洛方法的关键在生成大量的随机数并对其范围进行统计，那么对随机数生成就有一个必备的条件“分布均匀”，随机数都是由随机数生成器生成的，那么我们的生成随机数的办法均匀吗？

1. 真随机数

真正的随机数是使用物理现象产生的：比如掷钱币、骰子、转轮、使用电子元件的噪音、核裂变等等，这样的随机数发生器叫做物理性随机数发生器，它们的缺点是技术要求比较高。真随机数是依赖于物理随机数生成器的。使用较多的就是电子元件中的噪音等较为高级、复杂的物理过程来生成。至于“宇宙中不存在真正的随机”这种言论已经属于哲学范畴，在此不做讨论。在此我们默认存在随机。使用物理性随机数发生器生成的真随机数，可以说是完美再现了生活中的真正的“随机”，也可以称为绝对的公平。

2. 伪随机数

真正意义上的随机数（或者随机事件）在某次产生过程中是按照实验过程中表现的分布概率随机产生的，其结果是不可预测的，是不可见的。而计算机中的随机函数是按照一定算法模拟产生的，其结果是确定的，是可见的。我们可以这样认为这个可预见的结果其出现的概率是百分之百的。所以用计算机随机函数所产生的“随机数”并不随机，是伪随机数。从定义我们可以了解到，伪随机数其实是有规律的。只不过这个规律周期比较长，但还是可以预测的。主要原因就是伪随机数是计算机使用算法模拟出来的，这个过程并不涉及到物理过程，所以自然

不可能具有真随机数的特性。

蒙特卡洛树搜索法的应用

通常蒙特卡洛树搜索通过构造符合一定规则的随机数来解决数学上的各种问题。对于那些由于计算过于复杂而难以得到解析解或者根本没有解析解的问题，蒙特卡洛树搜索是一种有效的求出数值解的方法。一般蒙特卡洛树搜索在数学中最常见的应用就是蒙特卡罗积分。

蒙特卡洛树搜索在金融工程学，宏观经济学，生物医学，计算物理学 (如粒子输运计算、量子热力学计算、空气动力学计算) 等领域也应用广泛。

在森林采伐方面也有着蒙特卡洛算法的应用，森林管理的目标千差万别，其中包括保护受保护森林和自然保护区的资源，但主要目标往往是木材产品的生产，利用蒙特卡洛算法来实现采伐树木与保护树木之间的平衡，来使资源能一直利用下去。在《用于森林采伐计划的多目标蒙特卡洛树搜索》中 [4]，提出了一种蒙特卡洛树搜索方法，用于解决在明确区域受约束的双目标收获调度问题，总栖息地面积和栖息地内部核心总面积。这两个目标是使净现值和连接指数的概率最大化。提出该方法是一种帮助决策者估算有效替代解决方案和相应折衷方法。该方法已在数十个到超过一千个林分的森林实例以及三到八个时期的时间范围内进行了测试。使用该方法既在保护生物资源方面也起到了关键作用，也合理的采伐了树木，实现了资源的充分利用。计算机技术的发展，使得蒙特卡洛树搜索在最近 10 年得到快速的普及。现代的蒙特卡洛树搜索，已经不必亲自动手做实验，而是借助计算机的高速运转能力，使得原本费时费力的实验过程，变成了快速和轻而易举的事情。它不但用于解决许多复杂的科学方面的问题，也被项目管理人员经常使用。借助计算机技术，蒙特卡洛树搜索实现了两大优点：一是简单，省却了繁复的数学报导和演算过程，使得一般人也能够理解和掌握；二是快速。简单和快速，是蒙特卡罗方法在现代项目管理中获得应用的技术基础。

4 总结

在计算科学导论课程中，我大致了解到了计算机这一概念的大致内容，也对接下来的学习有了更好的理解。在制作这个报告中，我也从网络上的文献中解决了一些问题，从中了解到了在面对一个问题时应该如何分析处理的过程。计算科学导论就像是我学习路途中的指明灯，在我还停留在刚进入大学的迷茫之中，让我看清前方的道路，给我提供了奋斗的大致方向，也锻炼了我课堂上发表自己观点的能力，让我发现了自己的不足。计算科学导论这门课已经结束，但在这个期间所学习到的收获将会是我之后学习与前进的动力。在今后的学习生活中我也会将这收获运用到其中，以获得更大的收获与成果。也希望我能够在掌握好更多专业知识后再回过头来再浏览《计算科学导论》时能有着更多的收获和新的理解。

5

参考文献

[1] 强化学习 (reinforcement learning, rl) 初步介绍.

[2] 术语在线.

- [3] 随机数：真随机数和伪随机数.
- [4] Teresa Neto, Miguel Constantino, Isabel Martins, and João Pedro Pedroso. A multi-objective monte carlo tree search for forest harvest scheduling.
- [5] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*.

6 附录

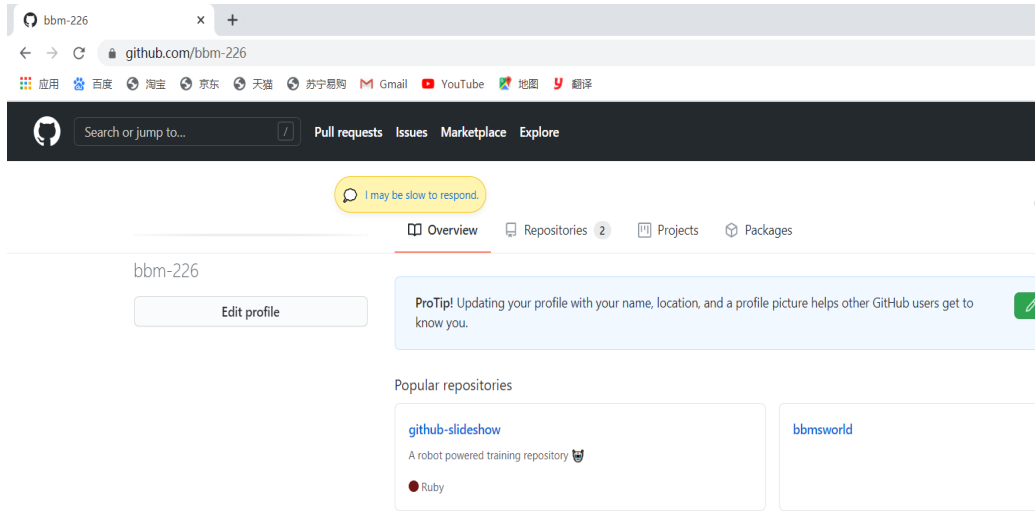


图 2: Github

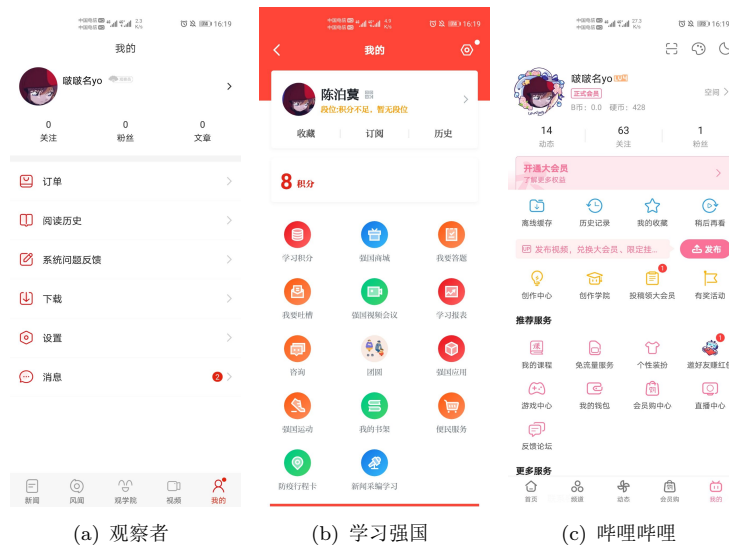
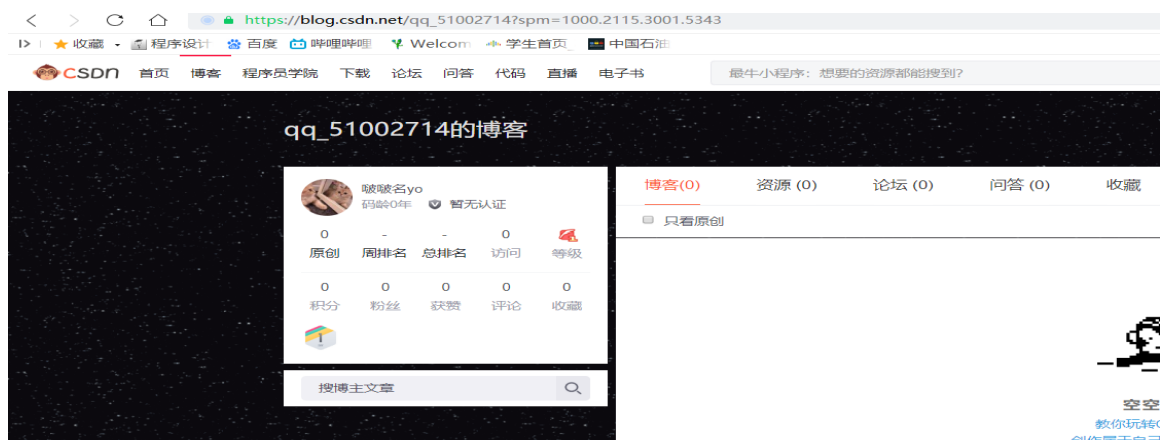


图 3: APP



(a) CSDN



(b) 博客园

图 4: CSDN 博客园



图 5: 小木虫