

《计算科学导论》

课程总结报告

学生姓名：

学 号：

专业班级： 计算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% | 总分 | 评阅教师 |
|  |  |  |  |  |  |

2020年12月29日

1. 引言

《计算科学导论》这本书带我重新认识了计算机的世界，通过对这这本书的深入学习，我重新认识了计算机，对计算机的起源与发展、计算机体系结构、程序设计、算法、软件工程、操作系统、人工智能以及计算机专业的培养目标都有了更深入更全面的认识，同时在学习这本书的过程中，我对计算科学的兴趣也得到了培养，为以后的学习也奠定了基础。我对计算科学的基本概念和基础知识有了更深的理解，不仅对计算科学的意义，内容和方法有了进一步的理解，而且对计算科学的分类与分支以及计算机专业的培养规格和目标有了更清晰的认识。计算已经成为继理论、实验之后的第三种科学形态。计算机是当今社会发展不可或缺的重要元素,它自问世以来一直走在科技前沿,几乎个个领域都离不开计算机，计算机无时无刻不在推动者社会发展。作为计算机专业的学生未来的发展前景无比广阔。《计算科学导论》这门课为我们敲开了专业之门，我从中受益匪浅。

1. 对计算科学导论这门课程的认识、体会

对计算科学学科历史渊源、学科特点、学科知识组织结构、学科发展规律和趋势等内容有了一些整体的把握。计算科学是对描述和变换信息的运算过程，包括其理论、分析、设计、效率分析、实现和应用的系统研究。学习了科学哲学，科学认识论，科学方法论，学科方法论，计算模型，二进制，通用数字计算机系统结构与工作原理，数字逻辑与集成电路，机器指令与汇编语言，算法，过程与程序，高级语言与程序设计，程序设计方法与技术，系统软件与应用软件，计算机图形学，图像处理与模式识别，逻辑与人工智能，计算机组织与体系结构，并行计算机系统，通道与并行计算，计算机网络与通信，高性能计算。

通过学习“计算机科学与技术导论”课程，我对计算机发展史和每一代计算机又有了新的认识。

二进制

德国人莱布尼茨约在1672-1676发明了二进制，用0和1两个数码来表示的数。

逻辑学

更准确的说是数理逻辑,乔治布尔开创了用数学方法研究逻辑或形式逻辑的学科，既是数学的一个分支，也是逻辑学的一个分支，简单地说就是与或非的逻辑运算。布尔代数实现了从一组逻辑公理出发，依靠代数演算来推导逻辑定理，用数学的方法来研究人的思维结构，对应于命题演算。皮尔斯和米切尔扩展到谓词演算。1930年，哥德尔完成谓词演算系统的完备性证明，新的数学分支数理逻辑形成。数理逻辑又称符号逻辑、理论逻辑，是用数学方法研究形式逻辑的学科。主要分支包括：逻辑演算（包括命题演算和谓词演算）、模型论、证明论、递归论和公理化集合论。

逻辑电路

我们知道在布尔代数里面X表示一个命题，X=0表示命题为假;X=1表示命题为真。如果用X代表一个继电器和普通开关组成的电路。那么,X=0就表示开关闭合，X=1就表示开关打开。不过他当时0表示闭合的理念跟现代正好相反。

(1)开关的闭合与打开对应命题的真假，0表示电路的断开，命题的假，1表示电路的连通，命题的真。

(2)X与Y的交集，交集相当于电路的串联，只有两个都联通，电路才是联通的，两个都为真，命题才为真。

(3)X与Y的并集，并集相当于电路的并联,有一个联通，电路就是联通的，两个有一个为真，命题即为真。

这样逻辑代数上的逻辑真假就与电路的连通断开，完美的完全映射。而且，所有的布尔代数基本规则，都非常完美的适合开关电路。

1945年，冯·诺依曼和他的研制小组在共同讨论的基础上发表了一个全新的“存储程序通用电子计算机方案”——EDVAC（Electronic Discrete Variable Automatic Computer），一篇长达101页纸洋洋万言的报告，即计算机史上著名的“101页报告”。这份报告奠定了现代电脑体系结构坚实的根基。报告广泛而具体地介绍了制造电子计算机和程序设计的新思想。这份报告是计算机发展史上一个划时代的文献，它向世界宣告：电子计算机的时代开始了。最主要是两点:其一是电子计算机应该以二进制为运算基础，其二是电子计算机应采用存储程序方式工作。并且进一步明确指出了整个计算机的结构应由五个部分组成：运算器、控制器、存储器、输入装置和输出装置,并描述了这五部分的职能和相互关系。

1936年，艾伦·图灵(1912-1954)提出了一种抽象的计算模型 —— 图灵机 (Turing Machine）。图灵机指一个抽象的机器，它有一条无限长的纸带，纸带分成了一个一个的小方格，每个方格有不同的颜色。有一个机器头在纸带上移来移去。机器头有一组内部状态，还有一些固定的程序。在每个时刻，机器头都要从当前纸带上读入一个方格信息，然后结合自己的内部状态查找程序表，根据程序输出信息到纸带方格上，并转换自己的内部状态，然后进行移动。图灵机的基本思想是用机器来模拟人们用纸笔进行数学运算的过程，他把这样的过程看作下列两种简单的动作：（1）在纸上写上或擦除某个符号；（2）把注意力从纸的一个位置移动到另一个位置。它证明了通用计算理论，肯定了计算机实现的可能性，同时它给出了计算机应有的主要架构。图灵机模型引入了读写与算法与程序语言的概念，极大的突破了过去的计算机器的设计理念。图灵机模型理论是计算学科最核心的理论，因为计算机的极限计算能力就是通用图灵机的计算能力，很多问题可以转化到图灵机这个简单的模型来考虑。

至此计算机的硬件结构(冯诺依曼)以及计算机的自然科学理论(图灵）已经比较完全了。计算机经过了第一代电子管计算机的时代。

第一代：电子管计算机，1946-1957

1. 元器件：电子管
2. 主存储器：电子射线管
3. 运算速度/每秒：几千次到几万次
4. 软件：机器语言、汇编语言
5. 应用：军事研究、科学计算

第二代：晶体管计算机，1958-1964

1. 元器件：晶体管
2. 主存储器：磁芯
3. 运算速度/每秒：几十万次
4. 软件：监控程序、高级语言
5. 应用：数据处理、事务处理

第三代：中、小规模集成电路计算机，1964-1971（1964年4月7日，在阿姆达尔的带领下，历时三年，耗费50亿美元，第一台IBM大型机SYSTEM/360简称S/360诞生）

1. 元器件：中小规模集成电路
2. 主存储器：半导体
3. 运算速度/每秒：几十万次到几百万次
4. 软件：操作系统、编译系统、应用程序
5. 应用：有较大发展、开始广泛使用

第四代：大规模及超大规模集成电路计算机，1971年至今

1. 元器件：大规模、超大规模集成电路
2. 主存储器：集成度更高的半导体
3. 运算速度/每秒：上千万次到上亿次
4. 软件：操作系统完善、数据库系统、高级语言发展、应用程序发展
5. 应用：渗透到社会各个领域
6. 进一步的思考

个性化推荐是根据用户的兴趣特点和购买行为，向用户推荐用户感兴趣的信息和商品。随着电子商务规模的不断扩大，商品个数和种类快速增长，顾客需要花费大量的时间才能找到自己想买的商品。这种浏览大量无关的信息和产品过程无疑会使淹没在信息过载问题中的消费者不断流失。为了解决这些问题，个性化推荐系统应运而生。

我们不难看出，不管是电商还是资讯，或者音乐电影等内容，个性化推荐都离不开用户的数据；有了数据，系统才知道用户的偏好以及特征是什么，所以收集用户的数据非常重要。一种为显式反馈，即用户那个把自己喜欢的和想要的信息告诉你；第二种为隐式反馈，即系统会收集用户的行为数据，这种情况用户一般是不会察觉的。

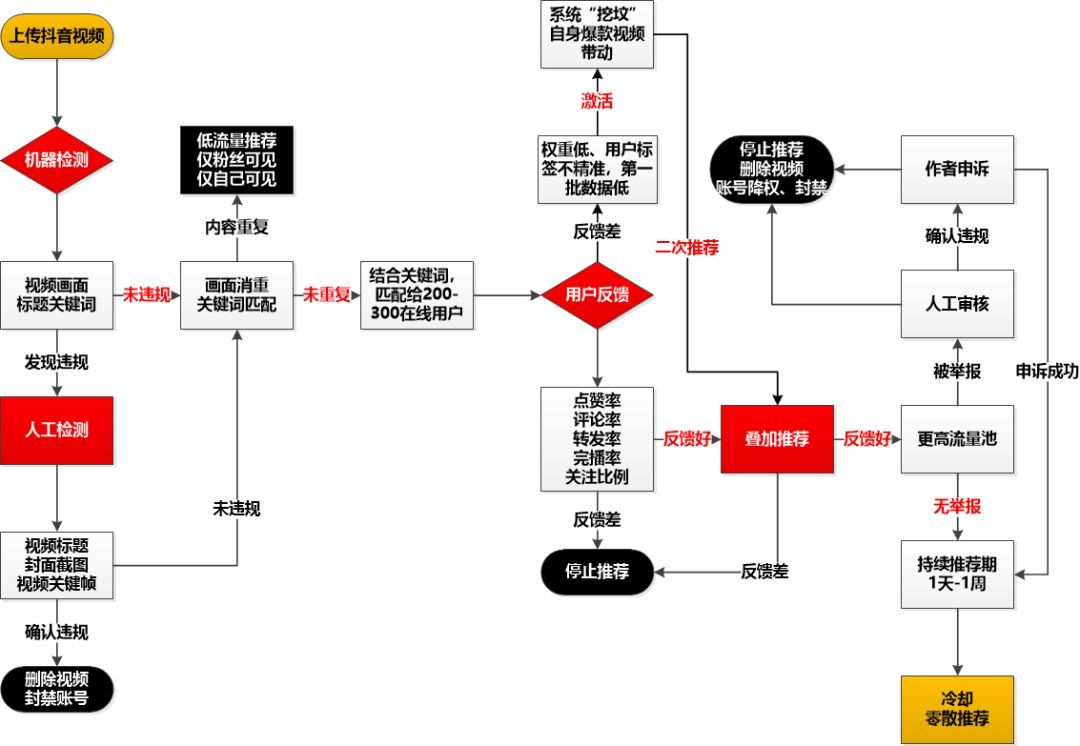
首先我们看一下显式反馈。最常见的一种方法就是让用户对自己体验过的用品进行评分，以此来体现出用户对此类物品的偏好程度，然后根据用户的反馈信息给出相关的推荐，例如淘宝的购物评分、网易云音乐的我的喜欢等功能；其次便是让用户自己选择自己感兴趣的分类，例如新闻客户端会让你选择感兴趣的新闻分类，豆瓣会让你选择感兴趣的影视类别等。这些用户主动提供信息的方式称为显式反馈。

除此之外，用户看不见的，还有隐式反馈。以阿里为例：在淘宝天猫等电商平台，会记录用户的购物行为，用户的偏好、年龄特征以及生活习惯都能够分析出来；饿了么会记录你的饮食习惯和偏好：你多久点一次外卖，喜欢吃什么食物会被记录；高德地图会记录你的出行信息.你的出行方式以及路线信息等都会被记录下来。单个信息不可怕，可怕的是把这些信息综合起来，就形成了一个完整的人物画像，你的年龄、性格、住址、习惯、兴趣等所有信息都会被提炼出来，而这些都是“个性化推荐”的基石。

有了人物画像，系统便可以根据其人物特征进行个性化推荐。想要达到非常好的推荐效果，系统需要采取非常多的推荐策略，各种复杂策略综合在一起，才能呈现出最完整的效果。

以下两个问题都与人物画像有关。

抖音的推荐算法问题



抖音的推荐算法机制是著名的信息流漏斗算法，也是今日头条的核心算法。通过审核后，第一步叫冷启动流量池曝光，比如你今天上传一个视频，通过双重审核的作品，系统将会分配给你一个初始流量池：200-300在线用户（也可能有上千个曝光）。抖音会根据这1000次曝光所产出的数据，结合你账号分值来分析是否给你加权，比如完播率、点赞、关注、评论、转发、转粉、游览深度等。以上这些都会对你的短视频数据造成影响，以及对你的短视频作出是否要加权的判断，然后会挑选前10%的视频，再增加1万次曝光。然后会给数据好的短视频进行更大的加权，并且会在第三步强化人群标签分发，让内容分发的更加精准，这类似猜你喜欢的打标，视频是有标签的，用户也是有标签的，两者之间会做标签匹配。最后进入精品推荐池，大规模曝光，一旦进入精品推荐后，人群标签就被弱化了



（常用的用户标签包括用户感兴趣的类别和主题、关键词、来源、基于兴趣的用户聚类以及各种垂直兴趣特征（车型，体育球队，股票等）。还有性别、年龄、地点等信息。性别信息通过用户第三方社交账号登录得到。年龄信息通常由模型预测，通过机型、阅读时间分布等预估。常驻地点来自用户授权访问位置信息，在位置信息的基础上通过传统聚类的方法拿到常驻点。常驻点结合其他信息，可以推测用户的工作地点、出差地点、旅游地点。这些用户标签非常有助于推荐。）

广告联盟问题

当用户进入门户网站或者app时，不同的用户看到的是不同的广告，广告联盟的系统计算出了不同用户或者用户群体的不同需求，通过广告推荐引擎系统和数据仓库中的统计数据以及用户的需求，展示给对应需求的用户观看，点击。

DSP（Demand Side Platform），是广告需求方平台，DSP为广告主提供跨媒介、跨平台、跨终端的的广告投放平台，通过数据整合、分析实现基于受众的精准投放，并且实时监控不断优化。

RTB（Real Time Bidding）实时竞价是DSP、广告交易平台等在网络广告投放中采用的主要售卖形式，会在极端的时间内（通常是50~100毫秒以内）通过对目标受众竞价的方式获得该次广告的展现，RTB的购买方式无论在PC端或是移动端均可以实现。

程序化购买（Programmatic Buying）根据广告主定义的期望受众，系统帮助其找出优选的媒体来购买受众，为广告主提出最优媒介采买计划，通过程序化购买的方式执行，并按照期望的周期反馈监测结果，并对后续投放进行优化。包括但不仅限于RTB购买。

投放前DSP会要求在广告主网站布码，同时在DSP的业务平台中录入广告投放的需求，如投放金额，投放排期，投放定向（如地域，兴趣，年龄等），最高限价。

当访客（即潜在的消费者）从左上角访问广告主网站开始，访客在广告主网站上的行为会被收集，同时DSP会与ADX和SSP进行Cookie Mapping，形成日志进行处理，形成回头客相关的行为数据标签。

当访客完成对广告主网站的访问，去其他媒体网站进行访问时，相应的媒体广告位根据事先嵌入的广告代码向广告网络发起广告请求，广告网络会将广告请求封装成http头 pb体的格式向多个DSP发起竞价请求。

当DSP接到竞价请求时会根据与广告网络约定的pb格式进行解包，拆解出相关的字段进行匹配，根据之前相关媒体积累的点击率结合点击率预测模型对出价进行预测，找出平台内在此次竞价请求能让平台利益最大化的广告主的创意进行投放，返回给广告网络出价与广告代码。

广告网络会在特定时间内（通常是50~100毫秒）根据多个DSP的出价高低，以第二名价格多一分的价格让出价最高的dsp胜出，并将广告代码中的展现宏和点击宏进行替换（替换过程中会根据事先与dsp约定好的公钥对价格进行加密，以防止第三方篡改和窃听）

网络将广告代码返回给媒体，媒体会将广告代码放置在js对应的位置进行展现，展现和点击的过程中会先后触发广告网络和胜出DSP的展现代码，广告网络和DSP分别接收到展现请求会对相应的展现进行计费操作（月底会相互进行对账）。

DSP内部会根据收集到的展现和点击进行计费操作，形成相应的报表；而浏览、展现、点击的记录会分别进行收集形成日志，经过ETL由DMP进行抽取和分析，形成媒体数据，用户标签，CookieMatch数据以及回头客用户标签数据，这些数据会在投放过程中作为RTB竞价的参考依据。

广告联盟是由多家广告提供商提供形成的一个组织，提供了多个平台的收集到的数据进行整合，数据的分析、清理，计算、统计等，提供向需要投放广告的广告主提供了一个投放系统平台。

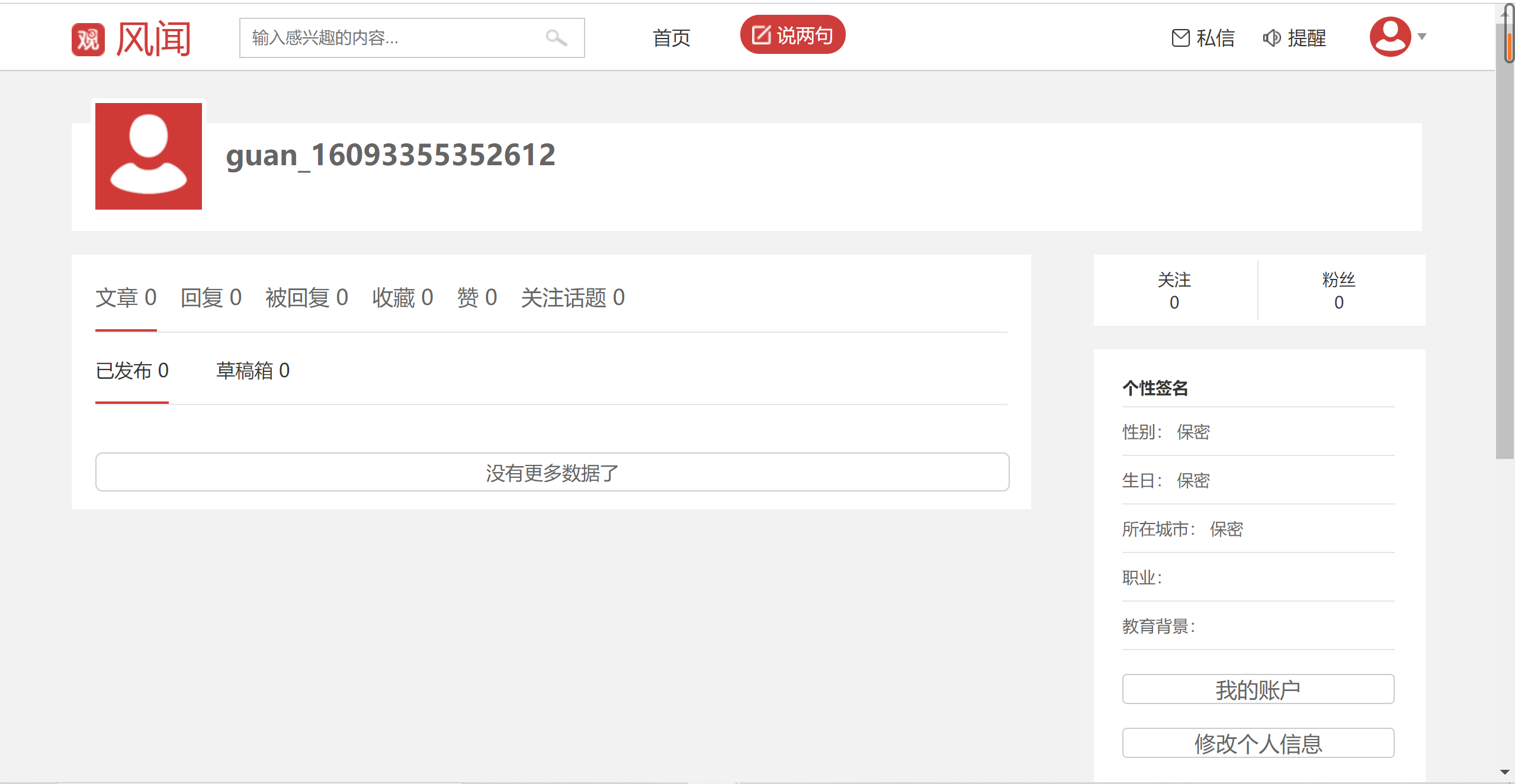
1. 总结

在学了计算科学导论之后，让我更深入的了解了我将来要从事的学科。通过课堂的学习，老师的讲解，我们开阔了眼界，知道了好多计算机领域的著作和奖项及名人等，这对我们人生道路也很有指导意义。他们就像我们的指路明灯，引领着我们不断前行。计算科学导论指导着我们该怎么学习计算机。让我更清楚的知道我们计算机科学与技术专业的方向。正如计算科学这座大楼一样，在不断的成长。计算机科学与技术也必将随着网络的进一步发展而更多的被人们重视。总之学习了这门课之后让我受益匪浅，也知道自己应该好好努力，争取在自己的专业领域上有所成就。

1. 参考文献
2. 赵致琢著，《计算科学导论》，科学出版社，2006年6月第3版。
3. 刘坤起著，《计算科学导论教学辅导》，科学出版社，2005。
4. 附录

<https://github.com/Zhidai-66CCFF/kexuedaolun>



<https://home.cnblogs.com/><http://muchong.com/bbs/space.php?uid=24886963>

