



第1章 计算机伦理学概述

董延杰

致腾楼929室, ydong@szu.edu.cn

Syllabus



- 案例教学（理论课和讨论课穿插进行）
 - 初步计划：每两节理论课后，安排一节讨论课，讨论内容将提前一周通知
 - 目的：教会学生多角度的理解计算机工程中的伦理问题
- 分数占比
 - 期末大作业40% + 平时成绩50% + 考勤10%
 - 平时成绩：随堂测试10%、两次作业20%、课内发言10%、课堂报告/课堂辩论赛10%

- 1.1 计算机伦理基本概念
- 1.2 计算机伦理学的研究方法及其发展



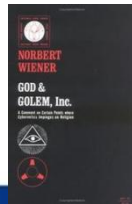
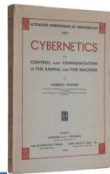


- 1.1 计算机伦理基本概念
 - 1.1.1 计算机伦理的提出
 - 1.1.2 计算机伦理学研究的基本问题
 - 1.1.3 计算机伦理学的定义
- 1.2 计算机伦理学的研究方法及其发展

1.1.1 计算机伦理的提出



- **Norbert Wiener.** 《控制论》 *Cybernetics: Or Control and Communication in the Animal and the Machine.* **1948**
 - ... in the presence of another social potentiality of unheard-of importance for good and for evil
- **Norbert Wiener.** 《人有人的用处》 *The Human Use of Human Beings.* **1950** (伦理学奠基性作品)
 - an account of the purpose of a human life
 - four principles of justice
 - a powerful method for doing applied ethics
 - discussions of the fundamental questions of computer ethics
 - examples of key computer ethics topics
- **Norbert Wiener.** 《上帝与机器人》 *God & Golem, Inc.: A Comment on Certain Points Where Cybernetics Impinges on Religion.* **1964.** -> 讨论了**控制论**与宗教的关系
- 以上这些思想为计算机伦理学奠定了**基础**
- **Norbert Wiener**是信息和计算机技术伦理的**鼻祖**



1.1.2 计算机伦理学研究的基本问题



1. 哲学中的伦理概念

- 伦理，也称道德哲学（moral philosophy），研究**关系（relations）**和**秩序（orders）**
 - 人与人、人与社会、人与自然之间的**关系**
 - 社会生活应该是什么样的**秩序**
- 伦理学是研究**道德（morality）**的学科
- 不同学派对伦理学的**研究对象**有很大差异
 - 例如：“善”；人类的道德行为；人类的幸福；道德原则和规范

1.1.2 计算机伦理学研究的基本问题



- 伦理学是哲学的一个分支学科，是对人类道德这个特定现象进行哲学思考的产物
 - 哲学遵循理论理性，伦理学遵循价值理性
- 伦理学
 - 理论伦理学：伦理学基本理论的学科
 - 应用伦理学：将伦理学的基本理论应用于社会生活的学科
- 哲学 → 应用哲学（applied/practical philosophy）→ 应用伦理学（applied ethics）→ 计算机伦理学（computer ethics）

1.1.2 计算机伦理学研究的基本问题



2. 伦理学的基本问题

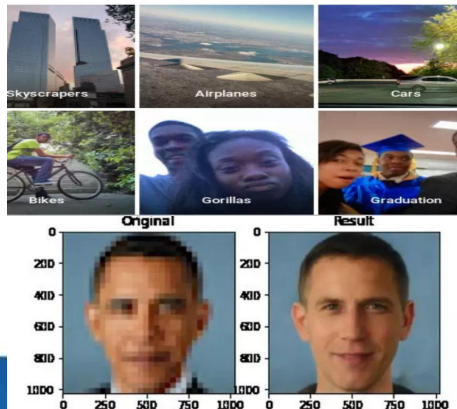
- 1) 基本问题：道德和利益两者的关系问题
 - 即：在道德框架内做对的事情
- 2) 研究方向：经济利益和道德的关系问题；个人利益和社会整体利益的关系问题
 - 例：Cross-silo联邦学习中，不同的客户拥有的数据量不一样。数据量多的客户不想和数据量少的客户合作。个人利益or社会整体利益？

1.1.2 计算机伦理学研究的基本问题



3. 计算机伦理学的研究对象

- 计算机伦理学的**研究对象**是在计算机的设计、开发、应用中，信息的**生产、存储、交换和传播**中，所涉及的**伦理道德问题**
- 个人隐私**（例：个性化推荐技术、用户画像技术）
- 知识产权**（例：淘宝网上售卖的盗版书籍等）
- 电子垃圾**（例：手机、硬盘、笔记本电脑等硬件）
- 信息垃圾**（例：垃圾邮件过滤）
- 算法公平**（例：将黑人识别为大猩猩）
- 网络攻击**
- 网络色情**



1.1.2 计算机伦理学研究的基本问题



4. 计算机伦理学研究的问题

- ① 计算机技术与**环境**的协调关系，这个关系中既有IT从业者和使用人员的问题，又有社会、文化、经济等因素
- ② **IT职业**，包括职业道德、社会责任产生的伦理问题
- ③ 技术如何为人类生活得更美好服务，**软件**品质和IT的风险及其管理问题、**算法公平性**
- ④ 软件等产品的**知识产权**保护，盗版及开放源代码运动问题
- ⑤ 网络、数据库等信息存储、传播带来的个人**隐私**及其保护的问题
- ⑥ IT的使用不当，如网络**成瘾**、网络**攻击**、网络**犯罪**，给人造成的身心健康问题
- ⑦ “数字鸿沟”加剧了不公平竞争、IT技术垄断等所涉及的相关**经济**问题

1.1.3 计算机伦理学的定义



- **James Moor**（美国计算机伦理学家）：计算机伦理学是“研究计算机技术的**本质（essence）**及其对自然和社会的**冲击（impact）**的分析，以及形成相应的**道德规范**与**评判政策**”的**应用伦理学科**
- **我国学者**：计算机伦理学是对计算机技术的各种**行为（behavior）**，尤其是计算机行为）及其**价值（value）**所进行的基本描述、分析和评价，并能阐明这些分析和评价的充足理由和基本原则，以便为有关计算机行为**规范**和**政策**制定提供理论依据的**一种理论体系**

1.1.3 计算机伦理学的定义



- **Terrell Ward Bynum**（美国南康涅狄格州大学教授、计算机与社会研究中心主任）：以技术进步并且保护**人类价值**的方式**整合计算机技术和人类价值** ...
- 《**Computer Essentials**》（国外某大学教科书）：是人类在社会生活中使用计算机的**可接受的道德行为**的指导，包括隐私、准确性、所有权和访问权。



- 1.1 计算机伦理基本概念
- 1.2 计算机伦理学的研究方法及其发展
 - 1.2.1 计算机伦理学的研究方法
 - 1.2.2 计算机伦理学的发展

1.2.1 计算机伦理学的研究方法



- 问题：“计算机技术到底给人类的生活带来什么样的变化”此类哲学问题
- 多学科合作（multidisciplinary collaboration）：计算机信息科学、哲学、社会学、心理学、法学、经济学等
 - 例：设计算法公平性机制是，我们可以参考经济学关于公平性的定义，如max-min公平、比例公平、以及alpha公平

1.2.1 计算机伦理学的研究方法



- 实证研究
 - 实地研究
 - 实验研究 (experiment)
 - 例：假设我们设计一种算法安全策略。在我们设计的安全策略投放市场之前，我们需要计算机仿真实验以确定我们的安全策略是能够防守住现时存在的网络攻击。
 - 调查研究 (Survey)
 - 例：当我们设计一种公平性机制时，我们需要针对设计的机制做调查问卷，以确定用户都对我们设计公平机制的满意程度。

1.2.1 计算机伦理学的研究方法



- 案例分析：案例是对一个包含了一个或多个**真实疑难问题**的**复杂情景**的描述（**例**：领导为节约成本要求员工把未按标准严格测试的产品交付给用户）

1.2.1 计算机伦理学的研究方法



- 社会调查：个案调查法、文献收集法、问卷法、访谈法

1.2.2 计算机伦理学的发展



1. 计算机伦理学先驱者（pioneers）

- Norbert Wiener
 - 奠定计算机伦理的著作：《控制论》、《人有人的用处》、《上帝与机器人》
- Walter Maner
 - 1976年首次提出“计算机伦理学”一词
 - 开拓了一个应用伦理学研究的新领域：计算机伦理学
 - 开始着手编写教学大纲并撰写文章

1.2.2 计算机伦理学的发展



1. 计算机伦理学先驱者(cont.)

- Terrell Ward Bynum:
 - 1983年，编写教学大纲并开设计算机伦理课
 - 1985年，在美国哲学杂志*Metaphilosophy*上的组织专刊*Computers and Ethics*，James H. Moor发表*What is Computer Ethics?*论文
 - 1987年，在美国南康涅狄格州大学成立“Research Center on Computing and Society”
 - 1991年，主持召开第一届“国际计算机伦理大会”
 - 1995年，与Simon Rogerson创立“国际计算机与信息技术伦理会议（ETHICOMP）”

1.2.2 计算机伦理学的发展



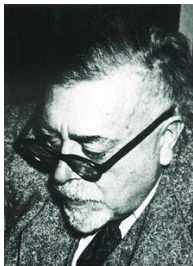
1. 计算机伦理学先驱者(cont.)

- James H. Moor: 1995年，总结计算机伦理学方面新的概念...
- Donald Gotterbarn: 计算机职业教育的领军人物、计算机伦理的实践者。领导开创了软件工程的伦理标准和职业实践规范；协助设计了软件工程师执业证书，担任美国ACM职业伦理委员会（Committee on Professional Ethics, COPE, <https://ethics.acm.org/>）主席
- Simon Rogerson: 欧洲研究计算机伦理的第一位教授，他在“作为一个技术人员应对社会承担的责任”方面做了很多工作

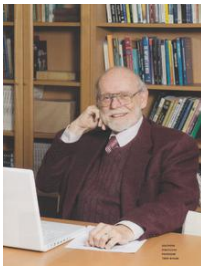
1.2.2 计算机伦理学的发展



1. 计算机伦理学先驱者(cont.)



Norbert Wiener



Terrell Ward Bynum



James H. Moor

1.2.2 计算机伦理学的发展



1. 计算机伦理学先驱者(cont.)

- 我国（始于20世纪90年代）
 - 《网络伦理》（1998年，严耕、陆俊、孙伟平）
 - 《终极市场—数字化经济时代》（严耕）
 - 《虚拟生存的意义》（陆俊）
 - 《西方计算机伦理学研究概述》（2000年，王正平）
 - 《信息化对21世纪经济和社会发展的影响》
 - 《信息化的哲学基础》（刘钢）
 - 《鼠标下的德行》（2002年，李伦 译）
 - 《网络伦理》（2007年，徐云峰）
 - 《计算机伦理学》（2011年，冯继宣、李劲东、罗俊杰）

1.2.2 计算机伦理学的发展



1. 计算机伦理学先驱者(cont.)

- 《着力构建中国特色的网络伦理》（2007年，王正平）：构建中国特色网络伦理的基本原则应包括：
 - 促进人类美好生活原则
 - 平等互惠原则（equality, win-win）
 - 自由与责任原则（freedom, responsibility）
 - 知情同意原则（informed consent）
 - 无害原则（no harm），例：Google公司的口号 “Don’t be evil”

1.2.2 计算机伦理学的发展



2. 计算机伦理学教育发展现状

- 1978年，美国最早开始在大学开展计算机伦理教学
- 20世纪90年代，慢慢发展到欧洲、澳大利亚等国家和地区
- 在21世纪初，我国教育部高等学校计算机科学与技术教学指导委员会编的《高等学校计算机科学与技术专业发展战略研究报告暨专业规范（试行）》，强调伦理教育必不可少
 - 计算机伦理学已经受到越来越多IT企业的重视（例：算法公平或许在未来某天将成为机器学习算法进入市场的一个准入条件、大数据杀熟立法。）

小节



- 哲学 -> 伦理学 -> 应用伦理学 -> 计算机伦理学
- 国内外学者
- IEEE-CS与ACM界认可的计算机伦理学课程的教学大纲

- 阅读Computer Science Curricula 2013 中的“Social issues and professional practice”部分内容，共12页。

https://www.acm.org/binaries/content/assets/education/cs2013_web_final.pdf

- 阅读2018 ACM Code of Ethics and Professional Conduct: Draft 1。

<http://ethics.acm.org/2018-code-draft-1/>

- 阅读**A very short history of computer ethics**

http://www.cs.utexas.edu/~ear/cs349/Bynum_Short_History.html

- 阅读Reading list中的内容

Reading list 中文章目录



- 阅读Reading list中的内容
 - a. How do fairness definitions fare? examining public attitudes towards algorithmic definitions of fairness
 - b. Equality of opportunity in supervised learning
 - c. AI Fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias
 - d. Joint Optimization of AI Fairness and Utility: A Human-Centered Approach