

第2章 计算机伦理基本原则与伦理分析方法

董延杰，助理教授

Email: ydong@szu.edu.cn

- 2.1 伦理学基本概念
- 2.2 伦理分析方法





- 2.1 伦理学基本概念
 - 2.1.1 什么是伦理学
 - 2.1.2 伦理、道德、法律
- 2.2 伦理分析方法

2.1.1 什么是伦理学



- 伦理学 (Ethics)
 - 研究宇宙万物之间的**关系**
 - 研究宇宙万物之间维持一个良好秩序的**维持方法**
 - 在西方也叫道德科学，是研究**道德行为**的学科
 - 属于**应用哲学**（哲学 → 应用哲学 → 应用伦理学 → 计算机伦理学）
- 道德 (Morality) 一词来源于**古希腊**，指**风尚、习俗**
 - Morality (from the Latin *moralis* "**manner, character, proper behavior**") is the **differentiation** of intentions, decisions, and actions between those that are distinguished as **proper** and those that are **improper**.
 - Morality can be a body of standards or principles induced by a code of conduct in a particular philosophy, religion, or culture
 - Morality can be derived from a standard that a person believes should be universal
 - Synonymous: **goodness** and **rightness**



2.1.1 什么是伦理学

- **亚里士多德 (384B.C.-322B.C.)**：最早把他关于人的**道德品质**的学问称为**伦理学**，并确立为一门专门的**学科**
- **Nicomachean Ethics** (《亚里士多德伦理学》，商务印书馆，1933年)
 - 该书是**西方伦理学史上第一部伦理学专著**
 - 书中系统阐述**德性在于合乎理性的活动、至善就是幸福**等观点
- **例：谈谈父母**从小对你的品德 (character) 教育和培养

2.1.1 什么是伦理学



- “德”、“道”：《尚书·商书》中已是一个重要的政治和道德概念
- 孔子、墨子、孟子、荀子等也都有丰富的伦理道德思想
 - 《论语》、《孟子》、《荀子》、《礼记》等
- 秦汉之际 (221B.C.-220A.D.)：形成了“伦理”的概念，产生了包含系统的道德理论、行为规范和德育方法的《礼记》、《孝经》等著作
- 有人把我国古代的哲学称为伦理类型的哲学

2.1.1 什么是伦理学



- 伦理学具体研究对象
 - 研究道德的本质、形成及其发展规律
 - 关于**理由的理论**，即做或不做某事的理由，认为某个行动、规范、做法、制度、政策和目标好坏的理由
 - 伦理学的任务：寻找和确定与行为有关的动机、态度、判断、理想和目标
 - 伦理学的特点：以人为中心（如何做人、如何对待他人，对待自然，对待万物.....）
- 伦理 vs 道德：先有道德现象，后有伦理学
 - 观点1：**伦理**强调**人与人、人与社会**的协调关系，**道德**局限于**个人**，是个人处事和修养的法则
 - 观点2：**伦理**是**人与人、人与社会、人与自然**关系的法则，**道德**是对伦理的态度
- 伦理和道德的紧密联系
 - 伦理存在的目的和意义：可为道德**提供依据或理由**，以指导人们的实践
 - 道德存在的目的和意义：在于可把伦理**落实在人心、落实在行动**

2.1.1 什么是伦理学



- 美国计算机协会发布的“计算机伦理十诫”
 - (1) 你不应该用计算机去伤害他人
 - (2) 你不应该去影响他人的计算机工作
 - (3) 你不应该到他人的计算机文件里去窥探
 - (4) 你不应该用计算机去偷盗
 - (5) 你不应该用计算机去做假证
 - (6) 你不应该拷贝你没有购买的软件
 - (7) 你不应该使用他人的计算机资源，除非你得到了准许或者为此做出了补偿
 - (8) 你不应该剽窃他人的精神产品
 - (9) 你应该注意你正在编写的程序和你正在设计的系统的社会效应
 - (10) 你应该始终注意，你使用计算机时是在进一步加强你对你的人类同胞的理解和尊敬

2.1.2 伦理、道德、法律



- 康德---法律是道德的外壳
 - 道德是肯定性的，积极地推动人们的行为
 - 法律是否定性的，消极地限制人们的行为
- 法律的这种否定性和消极性对于道德来说，却又起着积极的维护作用
- 例：去做道德鼓励的事情容易，还是不做法律否定的事情容易？

2.1.2 伦理、道德、法律



- 1. 两者所借以维持的力量不同
 - 道德：内心的信念(自律)
 - 法律：外在的强制力量(他律)
- 2. 两者作为行为规范的形式不同
 - 道德：约定俗成的方式
 - 法律：条文规定
- 3. 两者对人们行为的规范和调整范围不同
 - 违背道德的行为不一定是违法的
 - 违法的行为一般来说也是违背道德的



- 2.1 伦理学基本概念
- 2.2 伦理分析方法
 - 2.2.1 常用的伦理学理论
 - 2.2.2 伦理抉择5个基本原则
 - 2.2.3 伦理分析的一般框架

2.2.1 常用的伦理学理论



1. 相对主义 (relativism)

- 认为不存在普遍的道德准则，强调各种文化中行为的差异，认为关于对与错问题是相对的
- 更多的是对一种行为的描述，而非研究该怎么做的规范伦理
- Relativism is the belief that there are **no universal moral norms of right and wrong**. ... **Moral relativism** is the idea that each person decides what is right and wrong for them.
Anthropological relativism (人类学相对主义) is the concept that right and wrong is decided by a society's actual moral belief structure
- 例：在是非观念模糊的情况下，有较多人持这种观点
- 例：留意某些人通过这种观点来美化自己的错误行为

2.2.1 常用的伦理学理论



2. 美德论 (virtue ethics)

- 认为**某一行为的对错问题**取决于**品德高尚的人在相同情况下决定**
- 品德高尚的人：践行美德的人，如岳飞、范仲淹
- 美德：能够使社会繁荣，使全人类满足的行为或特性 (character traits)，如**爱国、敬业、诚信、正直、友善、忠诚**
- **例**：在发生家庭变故之前，该员工一直对大家都很好，会经常邀请大家去他家作客，为大家点外卖。但是，该员工由于家庭变故几乎缺席了项目开发过程中的组会，拖慢的整组进度。将此事上报给企业领导，会导致他失去当年的所有年终奖（六个月月薪）。项目组成员该怎么做？诚信、正直

2.2.1 常用的伦理学理论



- 例：孟德斯鸠《论法的精神·序言》
 - “...才能将包括博爱(Fraternity)在内的普世美德付诸实践...”
 - 所谓博爱，指的是己所不欲，勿施于人；己所欲者，常施于人
 - 所谓法，指的是“浅显理性与各种存在物之间的关系的综合”，同时也体现着所有客观存在物彼此间的关系”->美德是法的精神要义
 - 与我国“道法自然”的哲学思想很像
- 例：美德论思想在我国传统文化中也有很深刻的体现，《墨子》强调“举义”、“利人”、“利天下”；历史上无数志士仁人和民族英雄（文天祥、岳飞、范仲淹、顾炎武、林则徐等）；儒家的五常“仁、义、礼、智、信”
- 例：家庭和工作，哪个更重要？

2.2.1 常用的伦理学理论



3. 功利主义（或结果论，utilitarianism）

- 认为某一行为对错问题取决于这一行为的好处是否大于坏处
- 行为功利主义（Act utilitarianism）：量化某一行为的好处和坏处，进而确定某一行为的对错
- 规则功利主义（Rule utilitarianism）：将某一行为普遍化，如牵涉各方总效用提升，该行为即为对的行为
- 例：关于以好处和坏处的量化关系决定一个决定的对错，你认同这个观点吗？
- 例：是否应该占用农田修高速路？
- 例：某病毒可以通过操作系统漏洞感染电脑，是否应该开发另外一款病毒同样利用该漏洞清除第一种病毒？

2.2.1 常用的伦理学理论



4. 康德主义 (Kantianism)

- 康德主义是道义论 (Deontology) 的一种
- 认为某一行为的对错问题取决于普适的道德准则
- 普适道德准则第一判定条件：某一行为普遍化，不会推导出逻辑矛盾
- 普适道德准则第二判定条件：在决定做某事时，应该尊重自己和他人，承认他人自治独立的个体，而非利用他人
- 例：一公司资金链出现问题，急需新投资弥补资金缺口。该公司法人，在隐瞒公司财务状况情况下，与他人签订合约。这种行为是对是错？
- 例：某一半导体公司招聘，该公司是不告知新员工一年后需要搬去其他国家工作的情况下招聘新员工，这种行为是对是错？

2.2.1 常用的伦理学理论



- 例：说谎
 - 义务论/道义论：认为始终是错的
 - 功利主义/结果论： ???

2.2.1 常用的伦理学理论



- 例：攻击别人的网站
 - 义务论/道义论：认为始终是错的
 - 功利主义/结果论： ???

2.2.1 常用的伦理学理论



- 问：初创企业的产品被头部企业抄袭
 - 义务论/道义论：认为始终是错的
 - 功利主义/结果论： ???

2.2.2 伦理抉择5个基本原则



1. 尊重生命原则

- **最基本**的道德原则
- 尊重他人生存的权利，对自己生命的热爱
 - 例：康德曾经论述过自杀是不道德的
 - 例：正当防卫（尊重生命不代表任何情况下都不可以伤人）
 - 例：医生向危重病人隐瞒病情，树立病人战胜疾病的信心

2.2.2 伦理抉择5个基本原则



- 广义的“尊重生命原则”可以扩展为“无害原则”
 - 分析伦理问题的逻辑起点：首先要确定谁是受害者、谁是迫害者；迫害者通过什么手段造成了这些损害；损害程度如何

– 例：熊猫烧香病毒



2.2.2 伦理抉择5个基本原则



2. 社会公正原则

- “公正”主要指人们按照某种**公认合理的规则**处理问题的方式（如法律、行为规范、协定、游戏规则等）
- 公正是对有限资源都的合理分配
 - 公正是对个人自由的**某种限制**
 - 公正意味着人们都可以**平等**地享受权利和履行义务
- **例：**在法律面前人人平等
- **例：**政府出台政策时须关注公信力，如机动车限购、房屋限价问题

2.2.2 伦理抉择5个基本原则



3. 自主原则

- 当一个人**意志自由**时，既有**自主决定**采取何种行动来维护自身权益的能力，也有**尊重他人**拥有同样权益的能力→自尊和尊重他人
- **自尊**：对自己权益的**维护**、对自己行为的**负责**、对自己感性冲动的**克制**
- 尊重他人与自尊是一致的
- **例**：“己所不欲，勿施于人（treat someone else as you want to be treated）”，“爱人如己”
- **例**：保护商业秘密和个人隐私
- **例**：苹果公司与联邦调查局“撕破脸” 你支持谁？

<http://news.sina.com.cn/w/zg/2016-02-20/doc-ifyprucs6278920.shtml>

2.2.2 伦理抉择5个基本原则



4. 诚信原则

- 在各民族传统道德中都把诚信（**integrity**），如**不说谎、不做伪证、信守诺言**等作为基本的道德规范
- **例：**学术生涯中诚信（**credit**）的重要性，直接影响声誉（**reputation**）
- **例：**市场经济中诚信的重要性（如，信用卡/银行个人征信）
- **例：**精致的利己主义者

Please note that the final version of your manuscript **MUST NOT DIFFER** from the accepted version of the paper, except for very minor final edits such as typo corrections, or updating the information in existing references. You may not remove or add any content to the paper (including references or proofs) without the explicit permission of the Editor and Area Editor handling the paper. This will be checked by the IEEE Editorial Office, and the acceptance of papers changed inappropriately may be rescinded. If you have questions about whether any desired changes are appropriate, ask the Editor.

2.2.2 伦理抉择5个基本原则



5. 知情同意原则

- “同意”是某人对某事**自愿表示认可**，但要使同意有意义，**前提必须是某人对某事“知情”**，即他应知道即将发生的事件的**准确信息**并了解其**后果**
 - 个人隐私保护场景：为某一目的而采集到的**隐私信息**，在没有得到信息主体**自愿和知情同意**之前，**就不能用做其他目的**
- 建立在公平原则、自主原则和诚信原则之上
 - 只有知道实情，才能充分衡量后果，并作出相应选择，以维护自身利益
 - 只有充分知情，才能做到真正的自主，故意隐瞒真相违背了诚信原则
 - 在与尊重生命原则冲突时，具体问题须具体分析
- **例：**查看支付宝年度账单前，在首页发现了一行特别小的字：‘我同意《芝麻服务协议》’，不但字特别小，而且已经帮你选择好了‘同意’。而账单的查看和《芝麻服务协议》没有关联性，所以你若选择‘取消同意’，依然能够看到年度账单。但如果你没注意到，就会直接同意这个协议，允许支付宝收集你的信息，包括在第三方保存的信息。
- **例：**如果提示字特别大，但是仍然点确定，这算不算“知情同意”？

2.2.2 伦理抉择5个基本原则



- “尊重人的伦理学”：人们所遵守的行为规则应该把每个人都作为相互平等的道德主体来尊重，包括三种分析方式
 - 黄金法则：“己所不欲、勿施于人”（美德论）
 - 自我不利原则：如果我的某种行为普遍化，就会产生自我不利的情况（康德主义）
- 例：借钱不还
- 例：考试作弊
- 例：汽车停放不按规则

2.2.2 伦理抉择5个基本原则

- “尊重人的伦理学”要求不但平等对待他人，还要将他人作为道德主体来尊重。如何“尊重他人的道德主体”——赋予他人必要的权利
 - 权力：一种行动的权力，或者以某种方式行使个体行为的权力
- 哲学家Alan Gewirth提出的3个层次：
 - 最基本的权利（生命、身体的完整和精神的健康）
 - 维持个人实现自己的目标水平的权利（不被欺骗，对医疗实验的知情同意权，不被偷窃，不被诽谤，不遭毁约）
 - 实现提升个人的目标水平所必需的权利（设法获得财产的权利）
- 例：工厂为了节约资金而向外排放有高度致癌性的污染物，“侵犯”和“侵害”

2.2.3 伦理分析的一般框架



- 陈泽怀（上海师范大学哲学系教授）：底线伦理、共同信念、终极关怀
- **底线伦理**：所有人都应遵循的行为约束和规范
- 例：“勿杀人”、“勿盗窃”“勿说谎”“勿奸淫”等
- **共同信念**
- 例：“建设中国特色社会主义”、“贯彻社会主义核心价值观”、“促进全人类的和平与发展...”
- **终极关怀**：能化解生存与死亡、有限与无限的紧张对立，克服对于生死的困惑与焦虑

2.2.3 伦理分析的一般框架



• 案例分析1:

- 2018.08, 某司在其网站使用人工智能机器人 Dreamwriter 自动撰写并发表了一篇财经报道。同日, 盈某科技在其运营的“某贷之家”网站发布了与之一模一样的文章。某司发现此现象后将盈某科技告上法庭, 认为盈某公司的行为侵犯了某司 Dreamwriter 的著作权
- 2019.09, 法院受理了这一案件, 于 2020 年 1 月进行宣判。法院认为人工智能生成的作品属于著作权保护的范畴之内。法院的判案依据是《中华人民共和国著作权法实施条例》第二条关于作品的定义以及结合本案的事实, 法院最终认定某司 Dreamwriter 产生的作品具有独创性。本案例是全国首例认定人工智能生成的文章构成作品的著作权生效案件, 对于今后同类型案件的审理有一定的借鉴意义
- 著作权的归属应该如何判定?
- 人工智能创作是否会削减人类创作积极性; 人工智能的岗位替代是否会威胁人类就业安全; 人工智能创作的经济效益由谁来享用, 责任由谁来承担? (使用者vs写代码的人)

2.2.3 伦理分析的一般框架



- **案例分析2:**
- 小鹿咖啡财务问题始末：
 - 2020.02著名做空机构-浑水发布做空报告，小鹿咖啡股价盘中最大跌幅26.5%，随后小鹿咖啡回应论证有缺陷，两个月时间股价又涨回来了
 - 2020.04小鹿咖啡业绩造假22亿美元，股价暴跌80%，美律师事务所对小鹿咖啡发起集体诉讼
 - 2020.04小鹿咖啡董事长发声，表示承担责任。中国证监会回应“小鹿财务造假”，表示强烈谴责
 - 2020.04小鹿咖啡董事长道歉，表示羞愧心痛。小鹿咖啡财务造假高管员工被停职调查
 - 2021.09小鹿咖啡与股东达成和解
 - 2021.12小鹿咖啡的债务重组计划获得批准
- 小鹿咖啡事件违背了伦理学中的那些原则？
- 从企业运营者角度分析，如何避免财务造假重演？

2.2.3 伦理分析的一般框架



- **案例分析3:**
- **2021.11** 某公司发布（疑似泄露）一则公司内部纪律通报，根据有关资料显示，记录8月到9月期间，员工的各种个人流量详细使用信息，该文件旨在罗列员工的各种“摸鱼事件”同时该文件包含了涉事员工的个人信息如真实姓名等，该文件记录了如网易云音乐、腾讯视频、抖音、淘宝、字节跳动、快手等APP的详细流量使用记录。同时根据该文件，某公司将对涉事的相关员工予以警告甚至部分员工予以清退处罚
- 该事件的讨论引起了如知乎、微博、网易等多个社交平台的激烈讨论并登上这些平台的热搜，同时该事件引出人们对“摸鱼”和员工隐私的讨论，第三方软件 and 平台也积极响应了该事件，如“网易云音乐”就及时的推出“摸鱼计算器”功能狠狠蹭了热度
- **2021.11** 某公司继“摸鱼通报”事件后，对“泄露”的文件发出追究内部责任的公告，再一次登上社交平台的热搜。该事件也以某公司进行内部整顿而暂时结束，但互联网上的激烈讨论仍在继续
- 员工隐私和企业利益该如何保护？
- “摸鱼文化”在年轻人群体中得到推崇，究竟是年轻人缺乏上进还是现代社会的系统性问题？

- **案例分析4:**
- 外卖作为快节奏生活的必需品，得到了数百万外卖骑手和强大的人工智能系统的支持。2020年9月8日，网络上出现了一篇题为《外卖骑手，困在系统里》的热议文章。文章指出，美团在2016年至2019年间多次向配站站长发送加速通知。3公里送餐时限一再降低至38分钟；相关数据显示，2019年中国外卖订单的行业平均交付时间比3年前缩短了10分钟。在系统算法和数据的驱动下，外卖骑手逐渐成为高危职业。骑手无视交通规则，在算法规定的最晚送餐时间内完成送餐任务，继续加速车辆。
- 从企业角度，如何解决效率和安全的折中问题？
- 有哪些改进方案

小节



- 伦理、道德、法律
- 4个理论、5个基本原则及1个分析方法、1个一般框架

- 阅读 **Ten Commandments of Computer Ethics**
 - <http://cpsr.org/issues/ethics/cei/>
- 阅读 “美国COMPAS系统对黑人和白人重复犯罪率预测不一导致的歧视问题”
 - <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>
 - 根据第1-2章中有关伦理学和计算机伦理学的基本概念、伦理学的分析方法等，对这一事件进行分析。