

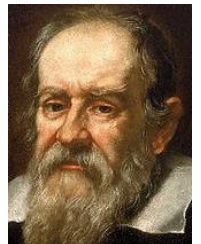
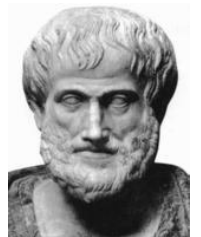
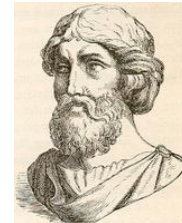
我对软件工​​程研究的认识

熊英飞，2017

什么是科学研究方法？
科学研究方法的特征是什么？

科学的历史

- 发源于古希腊自然哲学 (Natural Philosophy)
- 科学之父：泰利斯(Thales, 640-546BC)
 - 万事万物都应有自然的解释
- 古希腊三圣
 - 苏格拉底(Socrates, 469BC-399BC)
 - 批判性思维(Critical thinking)
 - 苏格拉底方法(Socratic Method)
 - 柏拉图(Plato, 424BC-348BC)
 - 演绎推理(Deductive Reasoning)
 - 亚里斯多德(Aristotle, 384BC-322BC)
 - 归纳推理(Inductive Reasoning)*
- 近代科学之父：伽利略(Galileo, 1564-1642)
 - 科学方法(Scientific Method)



* 存在争议

批判性思维 (Critical thinking)

- 反复问：“我是否应该相信你？”
- 批判性对话
 - 甲：我这学期的成绩肯定比上学期更好。
 - 乙：为什么？（质问理由）
 - 甲：我选了一门喜欢的课。（选了喜欢的课 -> 成绩会更好）
 - 乙：那又怎么样呢？（质问推理过程）
 - 甲：如果喜欢一门课，我会更努力。
 - 乙：更努力就一定成绩更好吗？（质问推理过程）
 - 甲：好吧，至少有很大的可能我的成绩会更好。
 - 乙：哪门课？我去查查选课系统看看你是否列在里面了。（质问前提）
 - 甲：《高级逻辑学》。结果还没出来，但我觉得我能选上。
 - 乙：那么只有你选上了，你才有很大的可能成绩会更好。（修正甲的声明）

批判性对话

- 学生：我设计了一个排序算法，比堆排序快100倍。
- 老师：怎么知道的？（要求理由）
- 学生：我根据数据的分布概率，计算了进行比较运算的平均次数。
（数据分布概率->该算法比堆排序快100倍）
- 老师：你的分布概率从哪儿得来的？（质问前提）
- 学生：根据我电脑里的文件统计的。
- 老师：这样的数据有代表性吗？
- 学生：.....
- 老师：你的算法只进行比较运算吗？（质问推理过程）
- 学生：还需要计算一些加减乘除。
- 老师：那比较运算的次数能代表算法速度吗？
- 学生：.....
- 老师：你这门课挂了，明年重修！

演绎推理 (Deductive Reasoning)

- 保证推理过程是正确的
- 今日形态：数理逻辑学
- 数学证明题都是演绎推理的过程
 - 前提
 - 推导步骤1
 - ...
 - 推导步骤n
 - 结论

归纳推理 (Inductive Reasoning)

- 如何保证前提是正确的
 - 柏拉图：真理在我心中
 - 亚里斯多德：真理来自对世界的观察
 - 称为经验主义(Empirism)
- 归纳推理
 - 从证据中提取假设
 - 用更多的证据对假设进行检验和修正
 - \neq 数学归纳法
- 归纳推理不能保证获得真理
 - 科学也是一种信仰，和宗教信仰类似

科学方法 (Scientific Method)

1. 提出一个问题

- a) 物体下落速度和质量有多大关系?
- b) 存在比堆排序快的排序算法吗?

2. 提出猜想 (归纳推理)

- a) 物体下落速度和质量无关
- b) 我的算法比堆排序快100倍

3. 从猜想进行推理 (演绎推理)

- a) 如果从比萨斜塔上扔两个不一样的铁球, 一定会同时着地
- b) 对我的电脑上的文件名进行排序, 我的算法应该比堆排序快100倍

4. 测试

- a) 真的同时着地了
- b) 悲剧了, 我的算法比堆排序还慢

5. 分析结果

- a) 一个实验太少了, 需要再做更多的实验
- b) 需要修正猜想, 我的算法可能只在某些限定条件下比堆排序快

6. 跳转第二步或者发表结果

为什么古代中国没有发明出飞机军舰？

科学的价值

- “开复，你是想一辈子写一堆像废纸一样的学术论文呢？还是想用产品改变世界？”
– 苹果副总裁对李开复说的话
- 没有产生产品的学术论文就是废纸吗？
- “勿在浮沙筑高楼” —— 侯捷
- 单个部件的可靠性为90%，系统包含10个部件，可靠性为35%
- 必须保证每一步都是非常可靠的

论文

- 多人构成的学术社区如何接近真理
- 苏格拉底方法
 - 通过循序渐进的讨论来接近真理
- 论文是这种讨论的书面反应

论文的内容

- 论文反映的是科学研究的结果
- 论文的内容应该能经得起批判性思维的检验（一定程度上）
- 论文三要素
 - 结论：论文提出的声明
 - 推理：归纳推理或演绎推理的过程
 - 前提：证据或者已经公认为真的命题

论文的例子

- 张路等. “Is Operator-Based Mutant Selection Superior to Random Mutant Selection?”. ICSE'10
 - 结论：基于操作符的变异选择方法并不比随机选择好
 - 推理：（反证法）如果基于操作符的方法比较好，那么它应该在实际数据集上选出更好的变异体
 - 前提：在西门子数据集上，基于操作符的方法并没有选出更好的变异体

基础研究 vs 应用研究

- 基础研究（科学研究）：认识世界
- 应用研究（工程研究）：改造世界

- 应用研究的例子
 - 结论：本文提出了一种比堆排序快**100**倍的算法
 - 推理：通过算法复杂度分析 或者 大量实际数据测试
 - 前提：已有的数据概率分布 或者 实际测试结果

Threats to validity

- 为什么要在论文中通常有这一部分
 - 自我批判论文中可能出现的问题
 - 保持科学的基本态度：诚实
- **Internal validity**
 - 在单一实例上推理可能出错
- **External validity**
 - 将单一实例泛化到多个实例上的时候可能出错
- **Construct validity**
 - 证据的收集可能出错

其他形式的科研——纯理论研究

- 经验主义论文
 - 研究方法：科学方法
 - 结论：新发现或者新发明
 - 推导：实验设计过程
 - 前提：实验结果或者已经被公认的证据
- 纯理论论文
 - 研究方法：纯演绎推理
 - 结论：新发现或者新发明
 - 推导：数学推理过程
 - 前提：已经被公认正确的命题

纯理论论文更应特别注意证据的有效性

其他形式的科研 – Grounded Theory

- 源于社会学领域
- 背景：做实验代价大、研究人员对实际情况不了解、研究社区有大量相关研究
- 步骤
 - 收集数据
 - 分析数据，建立真实世界的模型
 - 发表模型中有意义的结果
- 论文例子：
 - Miryung Kim, Vibha Sazawal, David Notkin, Gail C. Murphy: An empirical study of code clone genealogies. ESEC/SIGSOFT FSE 2005: 187-196
 - Thorsten Berger, Steven She, Rafael Lotufo, Andrzej Wasowski, Krzysztof Czarnecki: Variability modeling in the real: a perspective from the operating systems domain. ASE 2010: 73-82

常见误区

- 论文就是描述系统
- 论文就是广告

怎么判断研究的价值？

研究的质量属性

- 研究的可信度：基本要求
- 研究结论的interesting级别
 - 主观因素：取决于读者
 - 通用准则：如果该声明被接受，在多大程度上改变人们思考或者行动的方式？
- 例子
 - 相对论：改变相当长的历史时期内人们看待时间和空间的方式（大）
 - 操作系统开机速度对比方法：改变一定历史时期内操作系统测试方法（中）
 - Windows 7比Windows XP开机速度快：影响当前XP用户的升级选择（小）
 - Windows 7比Windows Vista开机速度快：强化当前绝大多数用户的已有观念（很小）
 - eCos 2.0比eCos 3.0开机速度快：影响人数非常少的eCos用户（很小）

OOPSLA会议的NICE准则

- Novelty
- Importance
- Clarity
- Evidence

软工顶级会议期望的结论类型

- 应用研究
 - 更好的解决了一个重大问题
 - 是否是实际存在的问题？
 - 是否是重大问题？
 - 该方法是否能解决？
 - 该方法是否比已有方法更优？是否有副作用？
 - 该方法的优势是否大于其实现代价？
- 基础研究
 - 发现了重要的事实、理论
 - 该事实/理论是否重要？对软工的实践和研究有何影响？
 - 发现是否有效可信？

科学方法是推动计算机发展的唯一途径吗？

- 早期大牛的论文
- MIT的策略：强调实现
- 软工界的反思

- 我的观点
 - 早期的计算机发展可依赖工程
 - 未来将越来越依靠科学