

软件工程与实验研究

周晓聪

中山大学计算机系

2018年9月

<http://sdcs.sysu.edu.cn/space/zxc>
isszxc@mail.sysu.edu.cn

- 1 什么是实证软件工程?
- 2 实验研究的基本思想
- 3 度量的基本概念
- 4 进一步阅读与课程论文

内容提要

- 1 什么是实证软件工程?
- 2 实验研究的基本思想
- 3 度量的基本概念
- 4 进一步阅读与课程论文

什么是实证软件工程？

- **实证软件工程**(Empirical Software Engineering, ESE)
 - 又称为**实验软件工程**(Experimental Software Engineering)
 - 软件工程研究的分支，关注在软件工程中如何应用和实践实证研究方法[1]
 - 广义的实证研究方法包括**调查**(survey)、**案例研究**(case study)和**实验**(experiment)等
 - 狭义的实证研究方法主要指实验

[1] Wikipedia: http://en.wikipedia.org/wiki/Experimental_software_engineering

什么是实证软件工程？

- 狭义地理解实证软件工程指研究在软件工程领域中如何应用实证研究方法
 - 如何设计实证研究、如何收集实证研究的数据、如何分析这些数据等
 - 特别地，研究如何在软件工程领域中进行实验研究
 - 软件工程领域的实验与其他工程领域的实验有什么不同？
- 广义地理解实证软件工程包括在软件工程中应用实证研究方法所进行的研究和实践
 - 例如，用实验研究哪些因素会影响结对编程(pair programming)的软件生产效率？
 - 使用实证研究方法针对软件开发过程、工具和产品所进行的一切研究都可属于实证软件工程的范围

课程内容

- 教学内容：结合例子讲授有关实证软件工程的基本概念、技术和方法
 - 为什么要在软件工程中进行实证研究？
 - 实证研究的基本思想、术语与过程
 - 系统性文献综述(systematic literature review)
 - 实验研究(experiment)
 - 基本的试验设计方法和基本的试验分析方法
 - 概率统计基础知识概要
 - 描述性统计方法
 - 统计假设检验
 - 试验设计与试验分析
- 考核内容：通过阅读文献理解并报告更多的实证软件工程的研究例子
 - 了解实证软件工程在软件工程中的应用例子
 - 了解软件工程各个领域的一些最新进展
 - 特别是一些运用实证研究方法而取得的进展

为什么需要实验研究？

- 科学研究根本目的是增加人类的知识
 - 人类的知识按被人接受程度可分为如下几种
 - 由所有人接受的基本事实(fact)
 - 无可争辩的命题(undisputed statements)
 - 有争议的命题(disputed statements)
 - 猜想(conjecture)或假想(speculation)
 - 科学研究对猜想或有争议的命题进行探索
 - 实验是其中最重要的一种探索方式或步骤
- 广义地，实验是验证猜想的一个过程[1]
 - 实验的基本目的是验证猜想或已有的理论、模型
 - 实验是可重复进行的
 - 实验结果必须进行有逻辑地分析以达到实验目的

[1] <http://en.wikipedia.org/wiki/Experiment>

为什么需要实验研究？

● 一种研究（或学习）的过程[1]

- 开始由可用数据归纳出一个初始的模型(假设、理论或猜想)
- 当模型演绎出的进一步结果与实际观察的数据（事实、现实）不一致时，通过一个归纳过程导致对模型的修正
- 进一步演绎模型的结果并再一次和旧的或新获得的数据进行比较，依次下去，可能得到进一步的改正并获得知识

[1] George E. P. Box, J. Stuart Hunter, William G. Hunter. *Statistics for Experimenters: Design, Innovation, and Discovery* (Second Edition), John Wiley & Sons, Inc, 2005

为什么需要实验研究？

● 归纳-演绎学习：日常体验

- **模型**: 今天和往常每天一样
- **演绎**: 我的车会停在它的停放位置
- **数据**: 它不在那个地方
- **归纳**: 一定有人将它开走了

- **模型**: 我的车被人偷了
- **演绎**: 我的车会不在那个停车场
- **数据**: 不，哦，它在停车场的那边
- **归纳**: 有人开走又开回来了

- **模型**: 我太太用过我的车
- **演绎**: 她可能在车里留下了纸条
- **数据**: 是的，纸条就在这里

为什么需要实验研究？

● 归纳-演绎学习：日常体验

- **模型**：今天和往常每天一样
- **演绎**：我的车会停在它的停放位置
- **数据**：它不在那个地方
- **归纳**：一定有人将它开走了

- **模型**：我的车被人偷了
- **演绎**：我的车会不在那个停车场
- **数据**：不，哦，它在停车场的那边
- **归纳**：有人开走又开回来了

- **模型**：我太太用过我的车
- **演绎**：她可能在车里留下了纸条
- **数据**：是的，纸条就在这里

为什么需要实验研究？

● 归纳-演绎学习：日常体验

- **模型**：今天和往常每天一样
- **演绎**：我的车会停在它的停放位置
- **数据**：它不在那个地方
- **归纳**：一定有人将它开走了

- **模型**：我的车被人偷了
- **演绎**：我的车会不在那个停车场
- **数据**：不，哦，它在停车场的那边
- **归纳**：有人开走又开回来了

- **模型**：我太太用过我的车
- **演绎**：她可能在车里留下了纸条
- **数据**：是的，纸条就在这里

为什么需要实验研究？

● 归纳-演绎学习：日常体验

- **模型**：今天和往常每天一样
- **演绎**：我的车会停在它的停放位置
- **数据**：它不在那个地方
- **归纳**：一定有人将它开走了

- **模型**：我的车被人偷了
- **演绎**：我的车会不在那个停车场
- **数据**：不，哦，它在停车场的那边
- **归纳**：有人开走又开回来了

- **模型**：我太太用过我的车
- **演绎**：她可能在车里留下了纸条
- **数据**：是的，纸条就在这里

为什么需要实验研究？

● 归纳-演绎学习：日常体验

- **模型**: 今天和往常每天一样
- **演绎**: 我的车会停在它的停放位置
- **数据**: 它不在那个地方
- **归纳**: 一定有人将它开走了

- **模型**: 我的车被人偷了
- **演绎**: 我的车会不在那个停车场
- **数据**: 不，哦，它在停车场的那边
- **归纳**: 有人开走又开回来了

- **模型**: 我太太用过我的车
- **演绎**: 她可能在车里留下了纸条
- **数据**: 是的，纸条就在这里

为什么需要实验研究？

● 归纳-演绎学习：日常体验

- **模型**: 今天和往常每天一样
- **演绎**: 我的车会停在它的停放位置
- **数据**: 它不在那个地方
- **归纳**: 一定有人将它开走了

- **模型**: 我的车被人偷了
- **演绎**: 我的车会不在那个停车场
- **数据**: 不，哦，它在停车场的那边
- **归纳**: 有人开走又开回来了

- **模型**: 我太太用过我的车
- **演绎**: 她可能在车里留下了纸条
- **数据**: 是的，纸条就在这里

为什么需要实验研究？

● 归纳-演绎学习：日常体验

- **模型**: 今天和往常每天一样
- **演绎**: 我的车会停在它的停放位置
- **数据**: 它不在那个地方
- **归纳**: 一定有人将它开走了

- **模型**: 我的车被人偷了
- **演绎**: 我的车会不在那个停车场
- **数据**: 不，哦，它在停车场的那边
- **归纳**: 有人开走又开回来了

- **模型**: 我太太用过我的车
- **演绎**: 她可能在车里留下了纸条
- **数据**: 是的，纸条就在这里

为什么需要实验研究？

● 归纳-演绎学习：日常体验

- **模型**: 今天和往常每天一样
- **演绎**: 我的车会停在它的停放位置
- **数据**: 它不在那个地方
- **归纳**: 一定有人将它开走了

- **模型**: 我的车被人偷了
- **演绎**: 我的车会不在那个停车场
- **数据**: 不，哦，它在停车场的那边
- **归纳**: 有人开走又开回来了

- **模型**: 我太太用过我的车
- **演绎**: 她可能在车里留下了纸条
- **数据**: 是的，纸条就在这里

为什么需要实验研究？

● 归纳-演绎学习：日常体验

- **模型**：今天和往常每天一样
- **演绎**：我的车会停在它的停放位置
- **数据**：它不在那个地方
- **归纳**：一定有人将它开走了

- **模型**：我的车被人偷了
- **演绎**：我的车会不在那个停车场
- **数据**：不，哦，它在停车场的那边
- **归纳**：有人开走又开回来了

- **模型**：我太太用过我的车
- **演绎**：她可能在车里留下了纸条
- **数据**：是的，纸条就在这里

为什么需要实验研究？

● 归纳-演绎学习：日常体验

- **模型**：今天和往常每天一样
- **演绎**：我的车会停在它的停放位置
- **数据**：它不在那个地方
- **归纳**：一定有人将它开走了

- **模型**：我的车被人偷了
- **演绎**：我的车会不在那个停车场
- **数据**：不，哦，它在停车场的那边
- **归纳**：有人开走又开回来了

- **模型**：我太太用过我的车
- **演绎**：她可能在车里留下了纸条
- **数据**：是的，纸条就在这里

为什么需要实验研究？

● 归纳-演绎学习：日常体验

- **模型**: 今天和往常每天一样
- **演绎**: 我的车会停在它的停放位置
- **数据**: 它不在那个地方
- **归纳**: 一定有人将它开走了

- **模型**: 我的车被人偷了
- **演绎**: 我的车会不在那个停车场
- **数据**: 不，哦，它在停车场的那边
- **归纳**: 有人开走又开回来了

- **模型**: 我太太用过我的车
- **演绎**: 她可能在车里留下了纸条
- **数据**: 是的，纸条就在这里

为什么需要实验研究?

● 归纳-演绎学习过程

- 开始点: 初始模型 M_1
- 演绎过程(induction):
 - M_1 有哪些期望结果?
 - 需要怎样的数据来比较 M_1 的一些期望结果?
- 设计实验来获得数据
 - 实验是受控制的, 目的是获得所需的数据
 - 实验数据中有噪音, 也不一定与 M_1 的期望结果完全相关
- 分析获得的数据与 M_1 的期望结果比较
 - 如果一致, 则从某种程度上验证了模型 M_1
 - 如果不一致, 则从新的数据进行归纳以修正模型 M_1 得到修正模型 M'_1

为什么软件工程领域需要实验研究？

● 软件工程

- 使用系统化、科学化和量化的方式去开发、运行和维护软件
- 软件不仅仅是一个产品，而且是一个过程
- 软件工程不仅仅是自然科学领域的工程而且是社会科学领域的工程
 - 软件的开发、运行和维护不仅仅受到参与人员的影响还受到整个社会环境的影响

● 软件工程还处于一种前科学化的状态(pre-scientific status) [1]

- 存在大量可选方法与技术，但方法与技术的可行性由时间验证而非实验验证
 - 软件工程领域的方法与技术发展很快，许多方法与技术很快被淘汰
- 使用人的经验而非精确的模型预测采用某种方法与技术的后果
 - 人们选择某种方法和技术多数随潮流或赶时尚

● 实验研究是使得软件工程领域科学化、量化的必然途径

[1] N. Juristo et al. *Basics of Software Engineering Experimentation*. Kluwer Academic Publishers, Boston, Massachusetts, USA, 2001. 🔍 ↻

为什么软件工程领域需要实验研究？ [1]

Example (软件危机是否真的存在？)

- 软件危机只是开发人员和用户的一个强烈主观印象而非证实的现象
 - 所谓的“软件危机”是提出“软件工程”这个词汇的源动力之一
 - 多数软件工程教科书都会提到软件危机及其表现
 - 但到底软件产品普遍出现故障的概率达到多少时才算是“危机”呢？
 - 有研究表明汽车产业的产品故障率也会高达25%以上

Example (面向对象方法是否真的提高软件质量？)

- 人们普遍相信面向对象方法会提高软件开发效率甚至质量
 - 但Hatton 1998年报告说有强烈证据表明C++在程序编写效率与质量有负面作用

[1] N. Juristo et al. *Basics of Software Engineering Experimentation*. Kluwer Academic Publishers, Boston, Massachusetts, USA, 2001. 🔍 ↻

软件工程领域进行实验存在的挑战

- 软件工程不仅仅是自然科学领域的工程而且是社会科学领域的工程
 - 软件工程领域的实验比自然科学领域工程的实验更为复杂
 - 不同于化学、农业科学领域的实验而更似医学领域的实验
 - 实验成本比较高，实验过程可能比较长
 - 实验主体（参与人员）的因素对实验结果有非常重要的影响
 - 实验结果缺乏普遍意义而难以为人们广泛接受
- 软件工程领域的方法和技术发展与更新非常快
 - 实验需要一个相对稳定的环境，对方法与技术不断地重复实验与验证
- 软件工程领域的人员尚未意识到实验对软件工程的重要性
 - 尚未有充足的资金投入软件工程领域的实验研究

实证软件工程的发展[1]

- 上世纪八十年代中后期Basili等开始提倡在软件工程进行实验研究[1]
- 上世纪九十年代中期前软件工程中的实证研究还相当缺乏
 - Tichy 1995年对软件工程领域的400多篇文章进行了综述
 - 其中40%需要实验验证但却没有
 - Zelkowitz 1998年对软件工程领域的600多篇文章进行了综述
 - 其中30%的文章没有实验验证, 只有10%的文章提供了一些实验
- 目前软件工程领域越来越重视实验研究
 - 1997年Springer开始出版ESE(Empirical Software Engineering)期刊
 - 期刊IEEE TSE、会议ICSE, FSE的文章几乎都包含实证研究的内容
 - 2002年IEEE/ACM开始召开ISESE会议, 2007年后改为ESEM(ESE and Measurement)会议
 - 大型公司如Microsoft 等也成立实证研究团队[2]

[1] V. R. Basili et al., Experimentation in software engineering, IEEE Transactions on Software Engineering, 12(7):733 – 743, Jul., 1986

[2] <http://research.microsoft.com/en-us/groups/ese/>

实证软件工程的发展

Example (软件测试领域的实证研究)

- 1994年到2003年发表在TSE/TOSME/ICSE/FSE/ISSTA/ ICSM 的软件测试领域的论文[1]
 - 34.5%的论文包含实验研究，而56%的论文包含案例研究
 - 实验研究论文的比例也在不断上升
 - 1994年到1998年占27.5%
 - 1999年到2003年则占38.8%
- 有关软件领域实验研究的综述和评论可进一步参阅[2][3]

[1] H. Do et al. Infrastructure support for controlled experimentation with software testing and regression testing techniques. ISESE, 2004

[2] Sjøberg, et al. A Survey of Controlled Experiments in Software Engineering. TSE, 31(9):733-753, 2005

[3] V. Kampenes, et al. A systematic review of quasi-experiments in software engineering. Information and Software Technology, 51:71-82, 2009

内容提要

- 1 什么是实证软件工程?
- 2 实验研究的基本思想
- 3 度量的基本概念
- 4 进一步阅读与课程论文

什么是实证研究？

● 实证研究(Empirical research) [1]

- 一种科学研究(scientific research)的方法
 - 不同于纯粹的数学研究或哲学思辨
- 通过直接的观察(observation)或者经历(experience)而获得知识的一种方法

● 实证研究的基本思想

- 观察(observation): 收集、整理和组织实证数据
- 归纳(induction): 归纳数据而形成假设(hypothesis)
- 演绎(deduction): 在假设的基础上演绎出可检验的结论
- 检验(testing): 使用数据检验假设（演绎出的结论）是否成立
- 评价(evaluation): 对检验结果进行分析和评价，并指示下一步的研究方向

[1] http://en.wikipedia.org/wiki/Empirical_research

什么是实证研究？

● 实证研究的对象(object)

— 考察问题领域不同现象之间的关系

- 定性关系(descriptive relationship): 各种现象之间有一定的行为模式(pattern)
- 关联关系(correlation): 各种现象之间存在一种相互影响的函数关系
- 因果关系(causal relationship): 一种现象与另外一些现象之间存在依赖关系

● 实证研究的结果(output)

— 定性模型(descriptive model): 描述现象之间定性关系的行为模式

- 可确定影响一种现象有哪些(which)可能因素(factor)

— 经验模型(empirical model): 描述现象之间一种经验统计函数关系

- 可确定可能因素如何(how)影响一种现象

— 理论模型(theoretical model): 给出现象与现象之间因果关系的定量描述

- 可解释可能因素为什么(why)影响一种现象

什么是实证研究？

● 实证研究的不同目的(purpose) [1]

- 探索性(exploratory)
 - 为新的研究探索新的想法或假设
- 描述性(descriptive)
 - 描述发生的现象或情景
- 解释性(explanatory)
 - 解释当前的现象或情景发生的原因
- 改进性(improving)
 - 试图对当前的现象或情景进行改进

[1] P. Runeson and M. Höst. Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering*, 14:131-164, 2009

什么是实证研究？

● 实证研究的两种风范(paradigm) [1]

— 定性研究(qualitative research)

- 在自然状态(natural setting)下关注研究对象
- 自然和非控制地观察研究对象
- 主观地(subjective)、面向发现(discovery-oriented)地进行研究
- 文字(text)、图形(graph)或表格描述而非数值描述

— 定量研究(quantitative research)

- 关注或比较多个现象之间的联系
- 进行受控的测量(measurement)
- 客观地(objective)，面向验证(verification-oriented)地进行研究
- 使用数值模型进行描述

[1] C. Wohlin et al. *Experimentation in Software Engineering: An Introduction*, Kluwer Academic Publishers, 2000

什么是实证研究？

● 实证研究的一般过程

- 定义(definition): 确定研究问题的范围和研究的具体目标
- 规划(planning)
 - 确定研究对象（总体）及抽样方法
 - 确定具体的研究范围：哪些现象与因素
 - 确定数据收集和分析的方法
- 执行(execution): 根据规划执行（观察）并收集数据
- 分析(analysis)与解释(interpretation)
- 结论(conclusion)与报告(presentation)

● 两个主要的关注点

- 如何获得客观的、合理的数据作为分析的基础
- 如何科学地分析数据以作出正确的结论

什么是实证研究？

● 常见的几种实证研究策略(strategy)

— 普查(survey)

- 一种收集信息以描述、理解或解释知识、态度或行为的综合系统[1]
- 主要目的是理解问题领域的整体情况

— 案例研究(case study)

- 在其自然状态下研究当前所发生现象的一种实证研究方法[2]
- 可以在复杂的实际项目中探索和验证理论，但对定量研究支持不足

— 实验(experiment)

- 一种将实际收集到的数据与假设模型进行比较的一种实证研究方法[3]
- 可以建立事物之间的定量关系，但在很多因素难以控制时难以应用实验研究

[1] S. Pfleeger and B. Kitchenham. Principles of Survey Research Part 1. *ACM SIGSOFT Software Engineering Notes*, 26(6):16-18, 2001

[2] P. Runeson and M. Höst. Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering*, 14:131-164, 2009

[3] N. Juristo et al. *Basics of Software Engineering Experimentation*. Kluwer Academic Publishers, Boston, Massachusetts, USA, 2001. 🔍 ↻

什么是实证研究？

● 不同实证研究策略的共同点

— 研究步骤基本一致

- 大体上都有定义、规划、执行、分析和结论等几个阶段

— 都需要使用统计分析方法

- 使用统计抽样方法确定研究的总体和样本
- 使用描述性统计或推断性统计分析所获得的数据

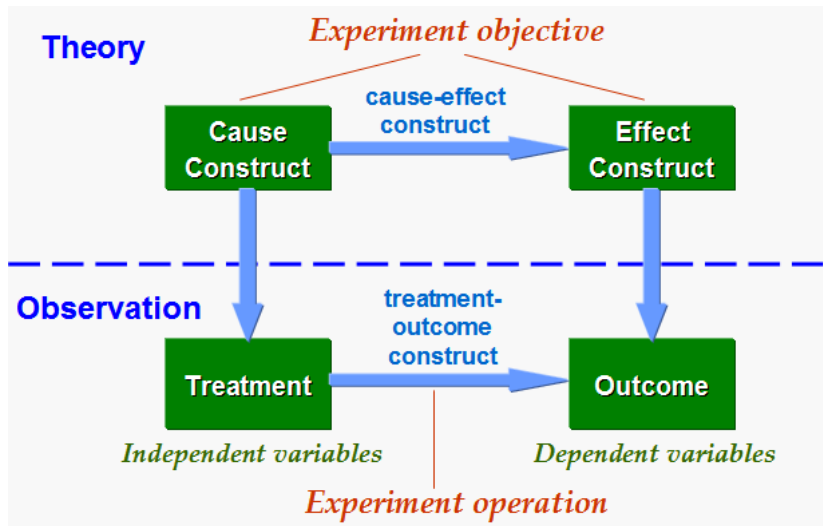
● 不同实证研究策略的主要区别

- 在主要的研究目的、执行过程与环境、对数据度量的控制、研究成本、重复的容易度等方面不同

什么是实证研究？

	普查	案例研究	实验
主要研究目的	偏重于描述性目的，理解总体的情况	偏重于探索性目的，探索研究对象是否具有某种性质	偏重于解释性目的，验证某种假设是否成立
执行过程环境	在产品（工具、方法）应用之前或应用之后；普查的过程难以控制	在实际的环境中考察产品（工具、方法）的应用情况；案例研究的过程难以控制	在实验室画境下考察产品（工具、方法）的应用情况；实验是完全受控制的过程
数据的度量	定性数据为主	定性和定量数据	定量数据为主
研究成本	低	较高	很高
重复的容易度	容易	很难	较易

实验的基本原理



实验的基本原理

- **实验(experiment)**是将理论（假设）与实际观察相比较的过程
- 实验的起点是有一个理论(theory)、假设(hypothesis)或模型(model)
 - 描述了问题领域的一些现象（或事物）之间的关联关系（或因果关系）
 - 例如，不同程序设计语言编写的程序的质量不同
 - 实验的目的是验证这个理论（假设、模型）是否成立
- **因变量(dependent variables)或响应变量(response)**
 - 模型中随其它事物变化而变化的事物的一种观察结果（度量结果）
 - 因变量是实验所关注的目标，是试验(test)的输出结果(outcome)
- **自变量(independent variables)**
 - 模型中影响因变量变化的事物的观察结果（度量结果）

实验的基本原理

● 因子(factor)

- 模型中所关注的与因变量有关联关系或因果关系的自变量
 - 引起因变量变化的自变量很多，但模型只关注其中一些自变量
- 模型：不同编程语言编写的程序质量不同
 - 因变量：程序质量
 - 因子：编程语言
 - 其他自变量：程序员经验、编写环境、所编写程序的功能等等

● 处理(treatment)

- 因子所取的不同的值或水平(level)
 - 因子：编程语言
 - 处理：C++语言、Java 语言、Python 语言

实验的基本原理

● 实验的基本原理

- 在控制其它自变量的基础上，考察因子的不同水平（处理）对因变量的影响
- 自变量和因变量都需要细化为可度量的变量
 - 例如，程序质量、程序员经验需要细化（转换）为可度量的变量
 - 对各种变量如何度量是细化模型时需要考虑的核心问题之一
- 因变量的取值受到多种自变量的影响，因而是随机变量
 - 在一次（组）实验中观察到的因变量的值是样本观测值
 - 如何使得因变量的样本观测值具有更好的代表性是实验设计的核心内容之一
- 除因子外，其他自变量需控制在特定的水平
 - 如何控制其他自变量而突出因子对因变量的影响是实验设计的核心内容之一

实验的基本原理

● 实验主体(subject)

- 实验中促使自变量影响因变量变化的人
 - 例如程序员、软件开发人员、项目管理员等等
- 实验主体的某些属性可能是重要的自变量
 - 例如人员的经验

● 实验对象(object)或实验单元(unit)

- 实验主体所作用的对象，作用的过程或结果可度量因变量的值
 - 例如程序、项目文档等等都可能是实验对象
- 实验对象（除因变量外）的某些属性也可能是重要的自变量
 - 例如所编写程序的功能等

实验的基本原理

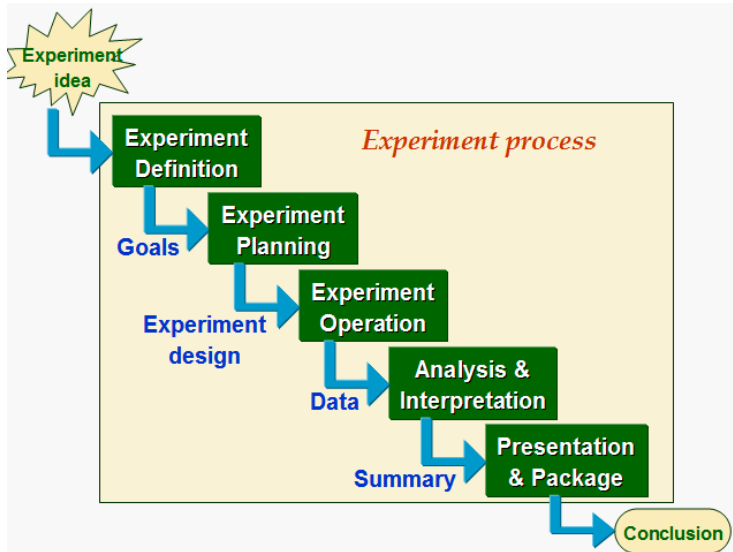
● 试验(test)

- 实验主体、实验对象与因子处理的一个组合称为一次试验
 - 从另外一个角度看，一次试验是对主体和对象的一次抽样
 - 一次试验结果是（因变量）的一个样本观测值
- 实验由一系列的试验组成
- 实验设计就是设计要做哪些试验，因此也称为试验设计

● 现代汉语词典中试验和实验是同义词

- 实验是检验某种科学理论或假说而进行的活动
- 试验是检验某种产品的性能而进行的操作或活动
- 实验比试验的含义更广泛
 - 因此我们将experiment翻译成实验，而test翻译成试验（或检验）

实验过程



实验过程

- 实验过程分为定义、规划、执行、分析与报告几个阶段
 - 普查或案例研究也可分为这些阶段
- 这些阶段不是瀑布式而是迭代式
 - 在一些阶段可能要回溯到上一阶段进行细化
- 后面我们将详细展开论述每个阶段的主要任务
 - 规划阶段的试验设计和分析阶段的试验结果分析需要较强的概率统计知识
 - 我们将在讲授实验过程模型之后进一步展开试验设计和分析的内容
 - 我们也将讲授试验设计和分析之前复习有关概率统计的基础知识
 - 这些知识也是学习数据挖掘等课程必需的基础知识
 - 规划阶段的实验有效性(validity)分析也是比较难以掌握的内容

内容提要

- 1 什么是实证软件工程？
- 2 实验研究的基本思想
- 3 度量的基本概念
- 4 进一步阅读与课程论文

实验研究与度量

- 度量是实验研究的基石
 - 度量是控制变量水平的基础
 - DeMarco: “You cannot control what you cannot measure”
 - 实验研究必须度量自变量和因变量的水平（取值）
- 度量(measurement)是现实世界物体与理想世界之间的一种映射[1]
 - 长度的米制就是一种度量（法则）
 - 建立现实世界物体长度与理想长度之间的一种映射
- 度量值(measure)是在某种度量映射下赋予物体的值或符号[1]
 - 例如“10米”、“100千米”都是度量值

[1] C. Wohlin et al. *Experimentation in Software Engineering: An Introduction*, Kluwer Academic Publishers, 2000

度量的基本概念

● 客观度量与主观度量

- 客观度量(objective measurement): 对于度量结果没有主观的判断
 - 度量结果只决定于度量对象
 - 在误差允许范围内, 不同次的度量将得到相同的结果
 - 例如, 程序代码长度是一种客观度量
- 主观度量(subjective measurement): 度量结果受到度量者的影响
 - 度量结果不仅取决于度量对象, 而且收到度量者度量角度的影响
 - 不同次的度量可能得到大不相同的度量结果
 - 例如, 程序的可理解性是一种主观度量

● 直接度量与间接度量

- 直接度量(direct measurement): 与物体其他属性无关的度量
 - 例如, 程序代码长度是一种直接度量
- 间接度量(indirect measurement): 通过对其他属性的度量计算而得到
 - 例如, 程序员产出量= (所编写代码总长度/编写代码的总时间)

度量尺度的类型

- 度量尺度(scales)的类型：变量（特征）取值精确程度的层次
 - 定名尺度(nominal scale)
 - 定序尺度(ordinal scale)
 - 定距尺度(interval scale)
 - 定比尺度(ratio scale)

度量尺度的类型

- 度量尺度(scales)的类型：变量（特征）取值精确程度的层次
- 定名尺度(nominal scale)
 - 对变量（特征）的取值进行平行分类或分组，每一类或组给一个标识
 - 人的性别、职业、民族等；软件的错误类型等
 - 定名尺度的数据之间不能进行任何计算

度量尺度的类型

- 度量尺度(scales)的类型：变量（特征）取值精确程度的层次
- 定名尺度(nominal scale)
 - 对变量（特征）的取值进行平行分类或分组，每一类或组给一个标识
 - 定名尺度的数据之间不能进行任何计算
- 定序尺度(ordinal scale)
 - 不仅将变量取值分类，还能确定这些类的等级差别
 - 软件的复杂度、可重用性、软件机构的成熟度(CMM)等；
 - 定序尺度的数据之间可以比较大小

度量尺度的类型

- 度量尺度(scales)的类型：变量（特征）取值精确程度的层次
- 定名尺度(nominal scale)
 - 对变量（特征）的取值进行平行分类或分组，每一类或组给一个标识
 - 定名尺度的数据之间不能进行任何计算
- 定序尺度(ordinal scale)
 - 不仅将变量取值分类，还能确定这些类的等级差别
 - 定序尺度的数据之间可以比较大小
- 定距尺度(interval scale)
 - 能测度变量取值类别或次序之间的间距，即能表达出各个类别的大小和多少
 - 摄氏温度；百分制的学习成绩等
 - 定距尺度的数据能做加减运算，但不能做乘除运算

度量尺度的类型

- 度量尺度(scales)的类型：变量（特征）取值精确程度的层次
- 定名尺度(nominal scale)
 - 对变量（特征）的取值进行平行分类或分组，每一类或组给一个标识
 - 定名尺度的数据之间不能进行任何计算
- 定序尺度(ordinal scale)
 - 不仅将变量取值分类，还能确定这些类的等级差别
 - 定序尺度的数据之间可以比较大小
- 定距尺度(interval scale)
 - 能测度变量取值类别或次序之间的间距，即能表达出各个类别的大小和多少
 - 定距尺度的数据能做加减运算，但不能做乘除运算
- 定比尺度(ratio scale)
 - 不仅能表达出各个类别的大小和多少，还有一个有意义的绝对零点作为起点
 - 程序代码长度；软件无故障时间等
 - 定比尺度的数据能做加、减、乘、除运算

软件工程中的度量

● 软件工程中的度量对象

- 过程(process): 软件开发过程中有哪些活动
- 产品(product): 软件开发的产品, 代码、文档、数据等
- 资源(resource): 开发软件所需的人力、物力、财力等

● 软件工程对象的内部与外部属性

- 内部属性(internal attribute)
 - 只针对对象就可以考察与度量的属性
 - 内部属性通常是直接度量
- 外部属性(external attribute)
 - 需要通过考察与其他对象关系而考察与度量的属性
 - 通常需要利用内部属性而进行间接度量
 - 软件工程由于涉及人的因素较多, 因此需经常考察对象的外部属性

软件工程中的度量

类型	对象	内部属性	度量
Product	Specification	Size Reuse Functionality Syntactic correctness	类个数 不修改而可重用的类个数 功能点个数 语法错误个数
	Design	Size Reuse Coupling Cohesiveness	模块数 不修改而可重用的模块数 两两模块之间的关联数 功能内聚的模块数/总模块数
	Code	Size Complexity	非注释的代码行(NLOC) 控制流图中的节点数 McCabe 环形复杂度

软件工程中的度量

类型	对象	内部属性	度量
Process	Overall process	Time	总的开发月数
	Constructing specifications	Effort	需求阶段人月数
	Testing	Time Effort	总的测试月数 测试阶段人月数
Resource	Personnel	Cost	每月总薪酬额
	Teams	Size	项目组人数

类型	对象	外部属性	度量
Product	Specification	Comprehensibility Maintainability	项目组外人员理解规范所需时数 改变一个需求所需人月数
	Design	Maintainability	改变一个模块所涉及的模块数

软件工程中的度量

类型	对象	外部属性	度量
Product	Code	Quality	缺陷数/代码行数
		Usability	独立使用程序前的训练时数
		Maintainability	执行一个改变所花天数
		Efficiency	运行时间
		Reliability	时间段 t 内故障数 平均无故障时间
Process	Overall process	Schedule deviation	估计月数/实际月数
	Constructing spec.	Stability of req.	需求改变次数
	Testing	Cost-effectiveness Quality	检测出的错误数/检测成本 检查出的错误数/总的错误数
Resource	Personnel	Productivity Experience	已实现功能点数/人月数 有经验年数
	Team	Productivity	已实现功能点数/项目组总月数

内容提要

- 1 什么是实证软件工程？
- 2 实验研究的基本思想
- 3 度量的基本概念
- 4 进一步阅读与课程论文

进一步阅读

● 关于实证软件工程的基本概念

- Wohlin等的教材“Experimentation in Software Engineering: An Introduction”对实验研究的基本思想及过程做了非常清晰的介绍
 - 但是在试验设计和分析方面则介绍的比较简单
- Jursto等的教材“Basics of Software Engineering Experimentation”对试验设计和分析做了比较详细的介绍
 - 该教材还收集了许多软件工程领域的实验作为例子

● 关于普查可进一步阅读文献[1]，而关于案例研究可进一步阅读文献[2]

● 关于软件度量可进一步阅读文献[3]

[1] S. Pfleeger, B. Kitchenham. Principles of Survey Research, Part 1-6, *ACM SIGSOFT Software Engineering Notes*, 2001, 26(6), 2002, 27(1-3), 2002, 27(5), 2003, 28(2)

[2] P. Runeson, M. Host. Guidelines for conducting and reporting case study research in software engineering, *Journal of Empirical Software Engineering* 2009, 14(131-164)

[3] Christof Ebert and Reiner Dumke. *Software Measurement*, Springer, New York 2007

课程论文

- 学生必须撰写一篇课程论文作为评定课程成绩的主要依据
- 学生选择一篇有关实证研究的论文
 - 必须在期刊ESE、TSE、ASE、JSS 或IST的2010年以后的论文中选
- 阅读该论文并使用中文撰写有关论文的报告
 - 论文的背景与动机
 - 论文中的实证研究的定义、规划、执行与分析
 - 对论文的个人评价和理解
- 在规定的时间内以电子邮件的形式提交
 - 邮件标题必须包含“软件工程进展”字样
 - 邮件附件必须是.rar 文件，以“学号姓名.rar”命名
 - 包括论文原文以及学生撰写的课程报告