

经验研究

(Empirical Study)

什么经验研究

- 通过直接的观察和实验，获得相应的经验证据，从而总结出相应知识的研究方法
- 强调真实的观察和实验背景
- 不轻信理论推导

经验证据

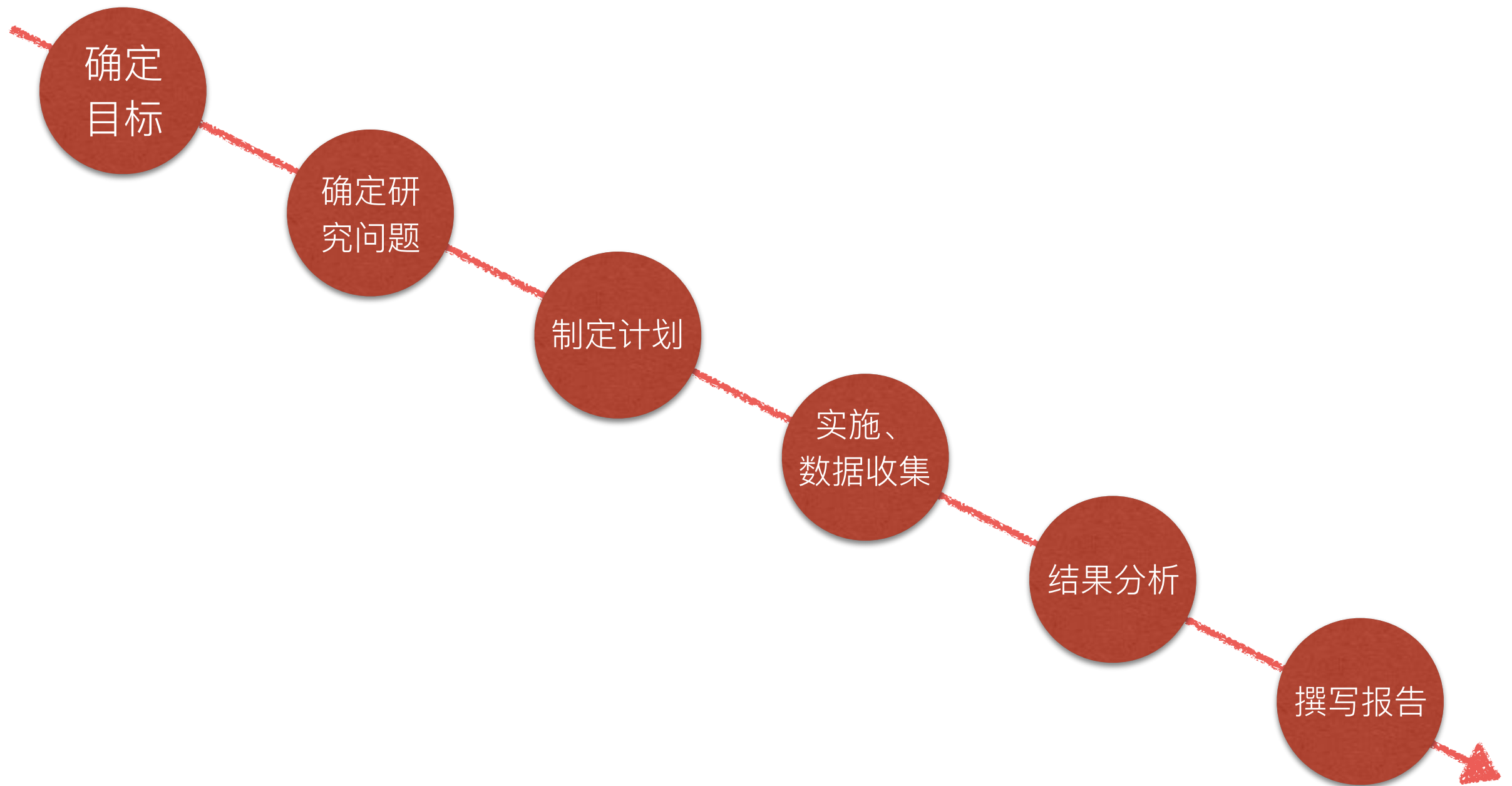
- 一种能够用来支撑判别断言真假的信息
- 经验科学：一个断言为真，当我们拥有支撑其为真的经验证据
- 理性科学：可以通过推导，判断某个断言的真假

设计模式是经验研究的产物

例子

- 问题1:
 - 非经验型：自动化的软件错误定位能够提高效率
 - 经验型：自动化的软件错误定位能够提高多少效率
- 问题2:
 - 非经验型：应用设计模式能够提高代码质量
 - 经验型：设计模式和代码质量的关系有待研究

经验研究的基本过程



经验研究的类型

- 文献研究 (secondary research)
- 案例研究 (case study)
- 实验 (experiment)
- 意见调查 (survey)

确立目标

- 目标的构成要素
 - 研究对象：你要研究的对象
 - 研究目的：分析、评估、理解.....
 - 关注点：某些性质，可能是研究对象的性质，或者它导致的影响
 - 利益相关者：从谁的角度出发进行研究？
 - 上下文：展开研究的上下文、所处的环境等等

确立目标

- 例子
 - 研究对象：结对编程
 - 研究目的：评估
 - 关注点：对开发效率的影响
 - 利益相关者：软件工程师
 - 上下文：某某公司

确立目标

- 例子
 - 研究对象：设计模式
 - 研究目的：评估
 - 关注点：对代码可理解性的影响
 - 利益相关者：研究人员和从业人员
 - 上下文：开源项目

分解研究问题

- 从目标出发，分解若干可以回答的研究问题
 - 研究问题必须是个问题
 - 研究问题必须和目标直接相关
 - 研究问题不应超过研究目标划定的范围
 - 研究问题应该有意义

分解研究问题

- 例子
 - 使用了设计模式的代码是否比不使用设计模式的代码可读性更高？
 - 对于同样复杂的代码，与没有使用设计模式的代码相比，开发人员是否更容易理解使用了设计模式的代码的意图？
 - 对于实现了同样功能的代码，开发人员是否能够更快（更准确）地理解基于设计模式的实现？
 - 一个模式中扮演不同角色的类是否具有不同的可理解性？
 -

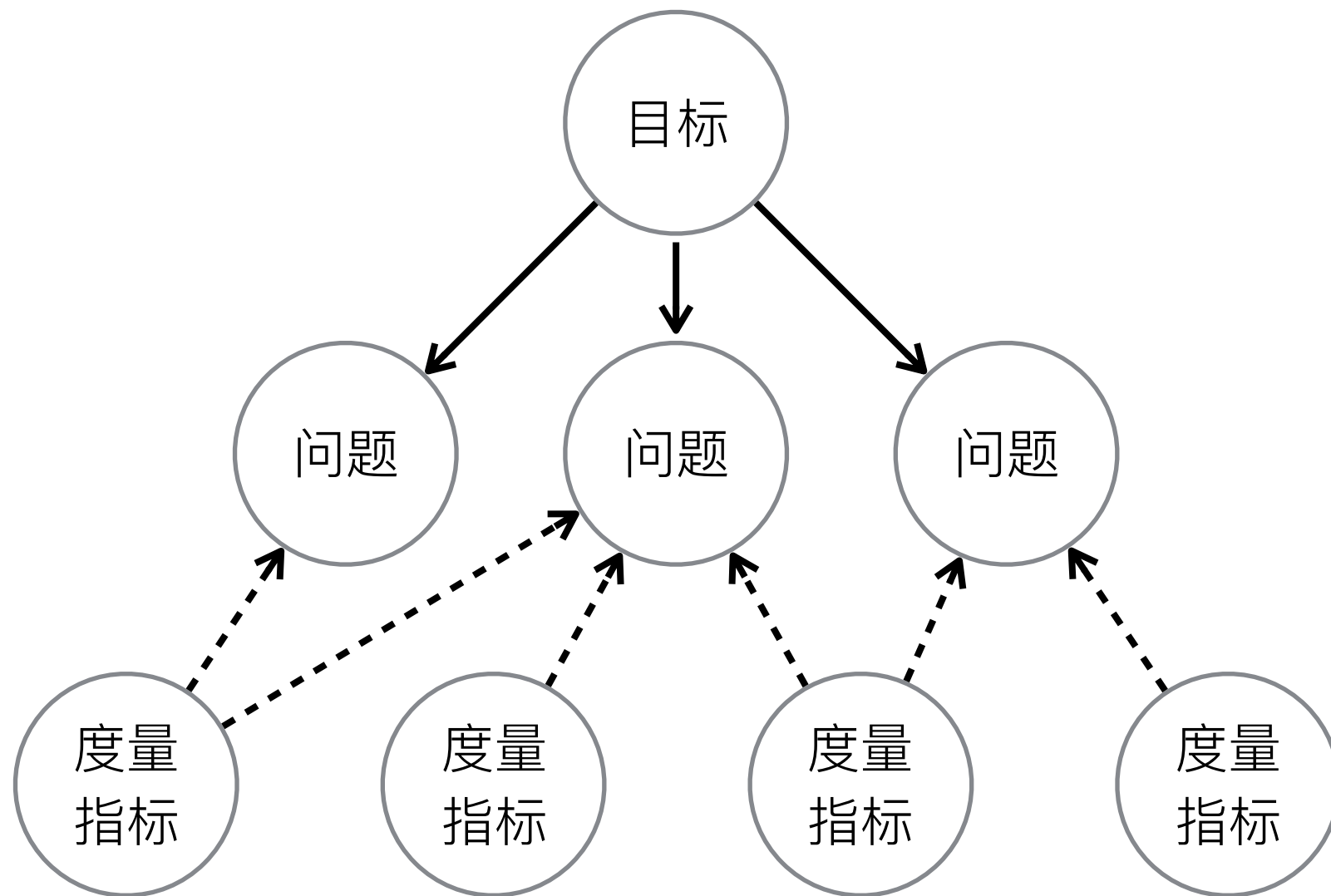
分解研究问题

- 反例
 - 设计模式的好处（不是问题）
 - 开发人员喜欢设计模式吗？（与目标无关）
 - 设计模式能够提高代码质量吗？（超出目标范围）

分解度量指标

- 度量
 - 为了回答研究问题，需要在研究中收集相关数据
 - 数据的定义、单位、收集方法（计算公式）
 - 研究问题和度量指标不是一一对应的关系

目标、问题、度量



分解度量指标

- 例子
 - CL: 代码长度
 - CC: 代码复杂性
 - UT: 代码理解用时
 - UA: 代码理解准确性
 -

制定计划

- 目的
 - 制定一个可操作（他人可重复）的过程
 - 研究实施时可参照的标准
 - 预先控制风险
 - 便于后期评估过程可靠性

结果分析

- 结果分析包含三个层次
 - 展示结果
 - 解读结果（你自己的解释）
 - 讨论结果（你自己的建议）

结果分析

- 有效性分析，分析可能影响研究结论的因素，包括：
 - 内部有效性
 - 如果研究因果关系，内部有效性涉及那些可能影响因果关系但没有被作为自变量的因素
 - 外部有效性
 - 研究结论是否被过分夸大了
 - 构造有效性
 - 实验设计中存在的误差，比如度量方法（工具）的误差
 - 可靠性
 - 对结果的分析 and 解释是否存在人为主观偏见，比如：你的方法真棒好！

文献研究

- 通过系统化的文献收集、分析、统计方法，对一个领域的研究现状进行调研
- 围绕具体的研究问题展开调研
- 通过对大量样本（论文）统计分析得出结论

文献研究

- 需要考虑的问题
 - 如何收集论文：
 - 怎样搜索（搜索关键字）
 - 搜索范围是什么（数据库、时间界限）
 - 滚雪球？
 - 如何筛选

案例研究

- 通过实际的案例，获取相关的数据，并得出结论
- 通过个例进行研究，目的在于获得可推广的结论
- 在真实环境下进行研究

案例研究

- 主要考虑的问题
 - 案例设计
 - 主体的选择
 - 你（作为观察者）是否对主体存在误导
 - 主体的经验是否会影响结果
 - 案例的实施环境是否具有代表性 / 特殊性
 - 数据收集
 - 数据的收集过程是否存在偏见（比如通过问卷调查）

实验

- 在受控环境下，分析因果关系的过程
- 分析在怎样的条件（自变量）下，可以得到怎样的结果
- 通常存在对照组

实验

- 主要考虑的问题
 - 怎么选择对照组
 - 比较是否“公平”
 - 实验量是否足够
 - 主观性误差