Carlo Ghezzi

DEIB-Politecnico di Milano

Being a researcher

Research in practice

Outline

- An ideal reference research process
- Essential activities/skills
  - Reading
  - Writing
  - Presenting
  - Building
  - Getting connected
- Have I achieved my goals? Validating research results

Ideal simplified process

Realize there is a problem

↓

Look to see if anyone has solved it

↓

Develop a plan to solve it

↓

Solve it

↓

Evaluate your solution

↓

Disseminate your solution

Abstract process

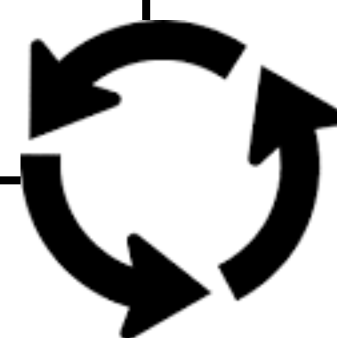
Research idea

Support engineers to write formal specifications of mobile robot missions through a catalog of structured natural language patterns (to avoid specification errors)

What do I want to achieve?

Recurrent specification come from published papers and from a selected number of industrial cases

Where do input data for the problem come from?



What do we do with the data?

Classify specifications to extract recurring patterns and provide a formal definition for the patterns

Have I achieved my goals?

How much do our patterns cover the body of specifications we started from? How useful are they for practitioners? Did they reduce the amount of incorrectness/ ambiguity? What are the limitations in all these cases?

Real process

- Highly iterative
- Non sequential
- Lots of dead ends, trial and error
- You need to convince (yourself) that the problem is worthwhile
- You need to convince (yourself) that you have a chance to solve the problem and prepare to handle risks
- Progressively formalize to guide your work (and not only)

Research area

- Unless you are a starting researcher, research ideas emerge after you choose a general research area in which you position yourself
- A research area guides your activities over a period of time: problems are "in the air"
- How to choose?
  - relevance, intellectual challenges, mature area, area engaging in interaction with real users, area engaging in systems implementations, hot area, fashion (volatile), …

Research problem

- A research problem has to be relevant
- Can be decomposed in further research problems, until it may become a workable unit that has its own coherence, provides a new contribution, and is worth disseminating
- Try to balance between inherent intellectual challenge and feasibility

Research proposal

- When you start the research process, often you start with a "problem area", which is not yet a research problem
  - if you are a beginner, don' t start from a too ill-defined problem area!
- After some wandering around (read papers, talk to people, …) you come out with a research problem
- Often the research problem is formalized in a research proposal for approval
  - from your supervisor or a committee of your school, e.g. as part of your PhD
  - for funding from a funding agency
  - for funding from industry or other bodies

Where do research ideas come from?

- Somebody else
  - senior researcher you work with
  - research customer
- A known problem that already attracted other researchers
- Emerges from intuition and ingenuity
  - knowing the field and being in the field helps
  - often part of the continuous incremental improvement research process
    - be humble: most research is of incremental nature

https://cacm.acm.org/blogs/blog-cacm/109579-long-live-incremental-research/fulltext
https://cacm.acm.org/blogs/blog-cacm/199207-whats-your-research/fulltext

Equip yourself for the journey: needed skills

- Read
- Write
- Connect
- Present
- Build

Equipping yourself: reading

- You need to know both the general area and the work that has been done in the specific subject area

- Although most reading is done in the initial stage of the research, more reading becomes necessary as you proceed

- Often background knowledge in other areas is also necessary

Equipping yourself: reading

- General ICST background beyond your area is always necessary (theory, architectures, systems, languages, AI, …), but you may also need a specific in-depth inspection
- Mathematical background often key in ICST research, but you may also need a specific in-depth inspection
    - algebra, graph theory, logic, statistical methods, stochastic methods
- Others: cognitive psychology, social science, philosophy, neuroscience, linguistics, physics, control theory, …
- MOOCs can be extremely useful

Equipping yourself: reading

- How much reading?
- Start from solid background knowledge and work incrementally, pause at stages
  - avoid the "I am not ready to do my work yet" (Peter Pan) syndrome
  - avoid the "work in isolation" syndrome

Opportunistic reading

- Often you may follow three phases, answering the following questions (if answer is no, exit)
  - is the paper interesting for me?
    - often the abstract is sufficient, but may need to quickly read a bit here and there (especially Intro and Conclusions)
  - where is the useful stuff for me?
  - read the entire paper for details

Equipping yourself: writing

- Writing is an essential part of the process
- Two main forms of relevant writing
  - research notebook
  - papers

Research notebook

- A very useful way of keeping record of ideas, how/why they develop and how/why they get discarded
  - if you don't write them down, you will forget
  - you may return to the same problem again
- Keep track of todo's, log postponed problems
  - What you write may be future problems to work on, which will generate future papers
- Go back to read what you wrote

Useful tips: How to write progress reports

https://homes.cs.washington.edu/~mernst/advice/progress-report.html

Writing papers

- Writing is fundamental for you
  - **you will never be sure that everything works until you carefully write it down for others to read**
    - *what seemed clear in your head does not survive as you try to formalize it*
- Writing is fundamental for communicating with others
  - **communicating results is a basic research responsibility**
- Avoid writing too late and writing too early

Writing papers

- Various forms:
  - Internal memo
  - Report for controlled circulation
  - Paper submitted for publication (we will talk about publication later)

Non disclosed papers

- They all aim at getting feedback
- Memos
  - used to interact within a research group, typically with your supervisor
  - may be direct and highly focused
- Reports
  - may be directed to a specific audience and/or preliminary to submission for publication
  - may explicitly say "draft, do not circulate", "do not cite"

Writing effort

- Writing papers requires effort
- Most common mistake: underestimate writing effort
- Especially critical because there are always deadlines
  - soft
    - if you do it later, the value is less
  - hard
    - most publication venues have a hard deadline

Equipping yourself: connecting

- *No researcher is an island entire of itself* —paraphrasing J. Donne
- Getting feedback is part of getting connected, which is fundamental in the modern world of research
- You may wish to get feedback from you colleagues in the lab or from leading experts working on the same or similar problems

Being connected

- Enables you to get in touch with other researchers and know what they are working on before it gets published
- Helps you generate research ideas, or redirect what you are doing in interesting directions
- Makes you visible in the research community, which is fundamental for your progress as a researcher
  - e.g., you need to be known when others will write recommendation letters for you
- You need to be proactive: look for others before others look for you, *do ut des*
- Conferences are the ideal networking environment (see later)

Equipping yourself: presenting

- Presentations are another form of communication
- Both useful for yourself and for others
- The presentation may range from "an idea" to a "worked out piece of research"
- It may be public (like in conferences) or (semi-)private
- Presenting is also essential for connecting

Equipping yourself: building

- Building artifacts is fundamental in ICST
- An artifact is an implementation of a research outcome, normally in a prototype form
- The artifact is often an integral part of the research effort
  - it may be the main final outcome
  - it may be necessary to support the validity of the work
- Making it available for others to use and enabling them to make further progress is key per a healthy research world

Validating the results of your research

you cannot publish something that has not undergone a thorough (internal) validation

Different approaches to validation: theory

- Theoretical work mostly validated by mathematical proofs
- Falsification may take the form of counter-examples, which dismiss validity of a claimed result
- Often (but not necessarily) validation takes the form of soundness and completeness
- In most cases, research motivated by practical problems, but not necessarily validated on them
    - can be too hard, or even unfeasible

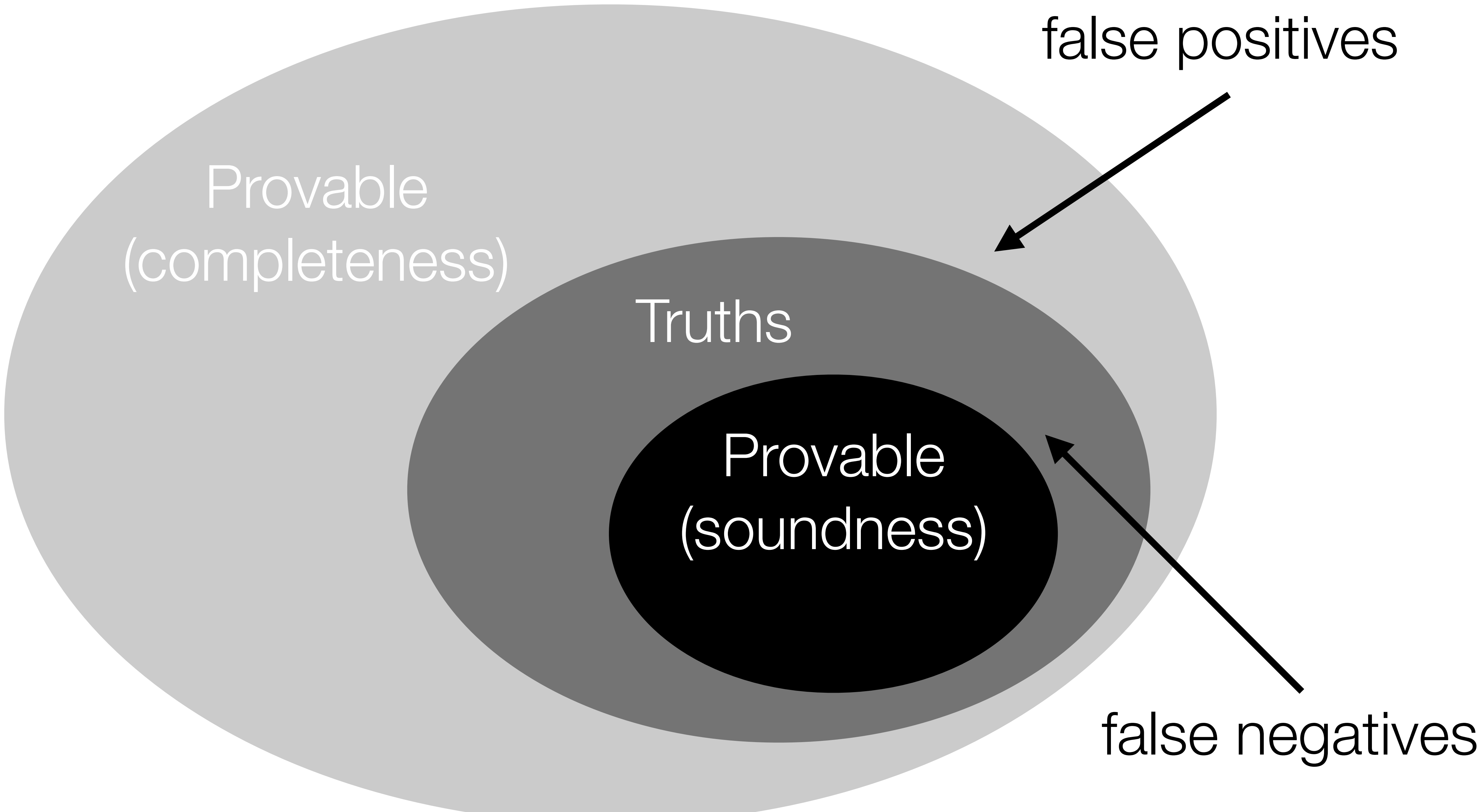
Theory: soundness and completeness

- Soundness (from logic):
  - a formula that is provable is true
    - for any set Φ of wff, and any wff α:
      - $\Phi \vdash \alpha \Rightarrow \Phi \models \alpha$ If there is a deduction of α from assumptions in Φ, then Φ logically implies α
    - *you cannot prove anything that is wrong*

Theory: soundness and completeness

- Completeness (from logic):
  - every true formula is provable
    - for any set Φ of wff, and any wff α:
      - $\Phi \models \alpha \Rightarrow \Phi \vdash \alpha$ If Φ logically implies α, then there is a deduction of α from assumptions in Φ
    - *you can prove anything that is right*
  - Ideally one wants soundness and completeness, but sometimes we are happy with unsound and or incomplete theories
  - generalized to approaches that are valid to all instances of the problem it aims to deal with

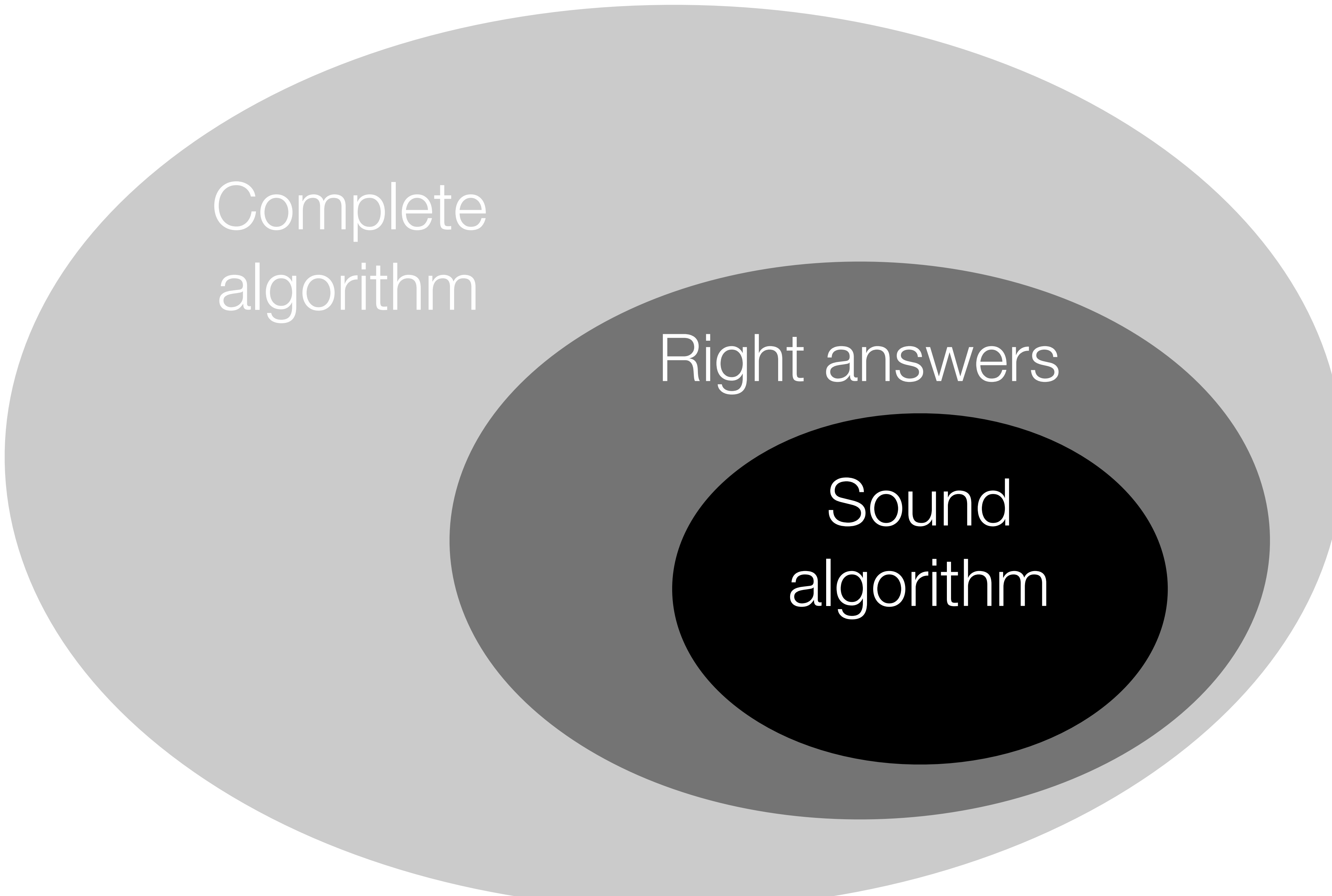
Soundness vs completeness in logic



Soundness and completeness used ubiquitously

- Used to refer to "approximations" of an approach, imperfect (but maybe very useful and clever) solutions
  - A **sound algorithm** never includes a wrong answer, but it might miss a few right answers => not necessarily "complete"
  - A **complete algorithm** gets every right answer in S (it include the complete set of right answers) but it might include also wrong answers => not necessarily "sound"
-

Soundness vs completeness of an algorithm



Example: program analysis leads to approximations

- Ubiquitous undecidability (e.g, termination, null pointer dereferencing, ..)
- Program analysis must be approximate in practice
  - May report errors where they are really not there
    - False positives
  - May not report errors that really exist
    - False negatives
- A sound approach (tool) has no false negatives
  - Never misses an error in a category that it checks

Sound analysis

- An analysis of a program P with respect to a formula F is sound if the analysis returns True only when the program actually does satisfy the formula
- If satisfaction of a formula F is taken as an indication of correctness, then a sound analysis is the same as a safe or conservative analysis
- If the sense of F is reversed (i.e., if the truth of F indicates a fault rather than correctness) then a sound analysis is not necessarily conservative
- (Note that use of the term sound has not always been consistent in the literature)

Complete analysis

- An analysis of a program P with respect to a formula F is complete if the analysis always returns True when the program actually does satisfy the formula (it may return True also if it does not)

- If satisfaction of a formula F is taken as an indication of correctness, then a complete analysis is one that is subject to optimistic inaccuracy

Example

- A static analysis tool S analyzes the source code of a program P to determine whether it satisfies a property. It can be "wrong" in one of two ways:
    - S is sound
        - it will never miss any violations, but it may say that P violates ϕ even though it doesn't
            - *all incorrect are programs detected (but it may incorrectly classify a correct program as an incorrect program—report as possible an inexistent null pointer dereferencing)*
    - S is complete
        - It may miss violations of ϕ
            - *all correct programs are detected, but possibly also incorrect programs*

How does this affect research?

- Advances of research must show improvements approximations (perhaps in well specified and significant cases)

- Reduce false positives

- Reduce false negatives

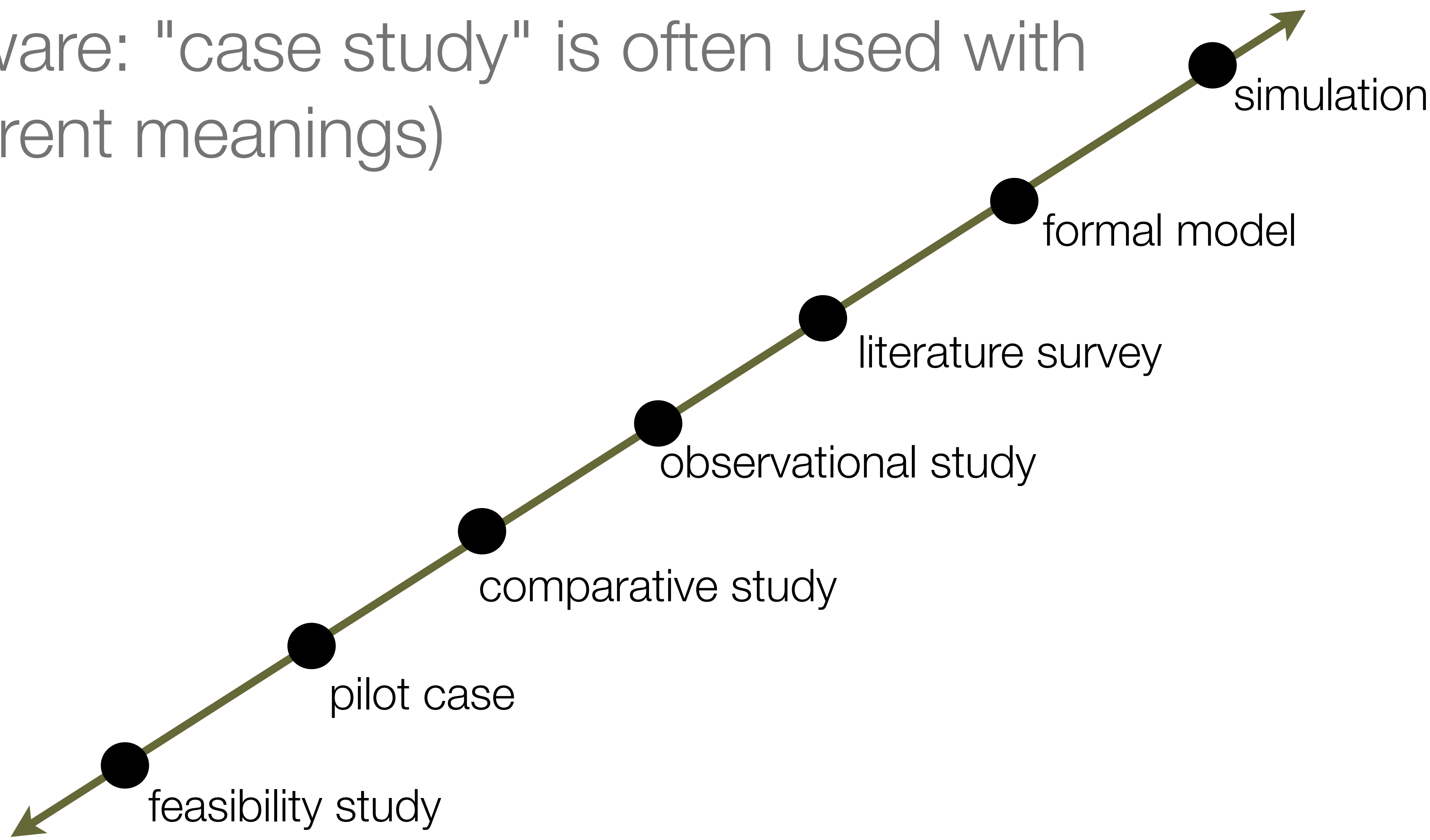
Constructive (synthetic) research

- In ICST it is common in research to create something new.
- Novelty is not enough to establish a contribution

> When one discovers a fact about nature, it is a contribution per se no matter how small. Since anyone can create something new [in constructive research], that alone does not establish a contribution. Rather one must show that the creation is better.
> (F. Brooks)

Research validation through case studies

A whole spectrum of case studies exist (beware: "case study" is often used with different meanings)



Based on S. Demeyer's notes on "Research methods in CS"

Feasibility study

- Here is a new idea, is it possible?
  - Is it possible to solve a specific kind of problem…effectively?
    - CS perpective (P=NP, algorithm complexity, cryptography, cryptocurrency…)
    - engineering perspective (build efficiently, fast, small, cost-effective, profitably…)
  - Is the technique new/novel/innovative?
  - Proof by construction
    - build prototype
    - apply to a case
  - Conclusions
    - primarily qualitative; lessons learned
    - quantitative (engineering perspective)
      - cost-benefit, speed-memory

Pilot case (aka demonstrator)

- Here is an idea that has proven valuable; does it work or us?
  - proven valuable
    - ascertained merits, lessons learned from feasibility study
  - does it work for us?
    - context may be relevant
  - Demonstrated on a simple yet representative case
  - Proof by construction
    - build a prototype
    - apply on a case
  - Conclusions
    - primarily qualitative, lessons learned
    - quantitative, against predefined criteria

Comparative study

- Here are two (+) techniques, how do they compare?
  - not necessarily a ranking
  - positions may vary according to purpose
    - where do they differ? what are the tradeoffs?
  - Predefine criteria checklist
    - should be fair, not favor your technique
    - should be complete, and reusable
    - qualitative: how to be unbiased?
    - quantitative: data should represent what you want
  - Reusable checklists are important contributions, support replicated studies
  - Score criteria checklists, often by applying techniques to cases
  - Typically summarized in tables

Observational study (etnographic studies)

- Understand phenomena through observation
  - systematic collection of data from direct observation of everyday life
    - phenomena best understood in the real context
      - observation and participation
      - interviews, questionnaires, monitored data collection
        - example: how does a certain technology improve awareness of healthy behaviors
    - conclusions primarily qualitative: classifications/observations

Literature survey

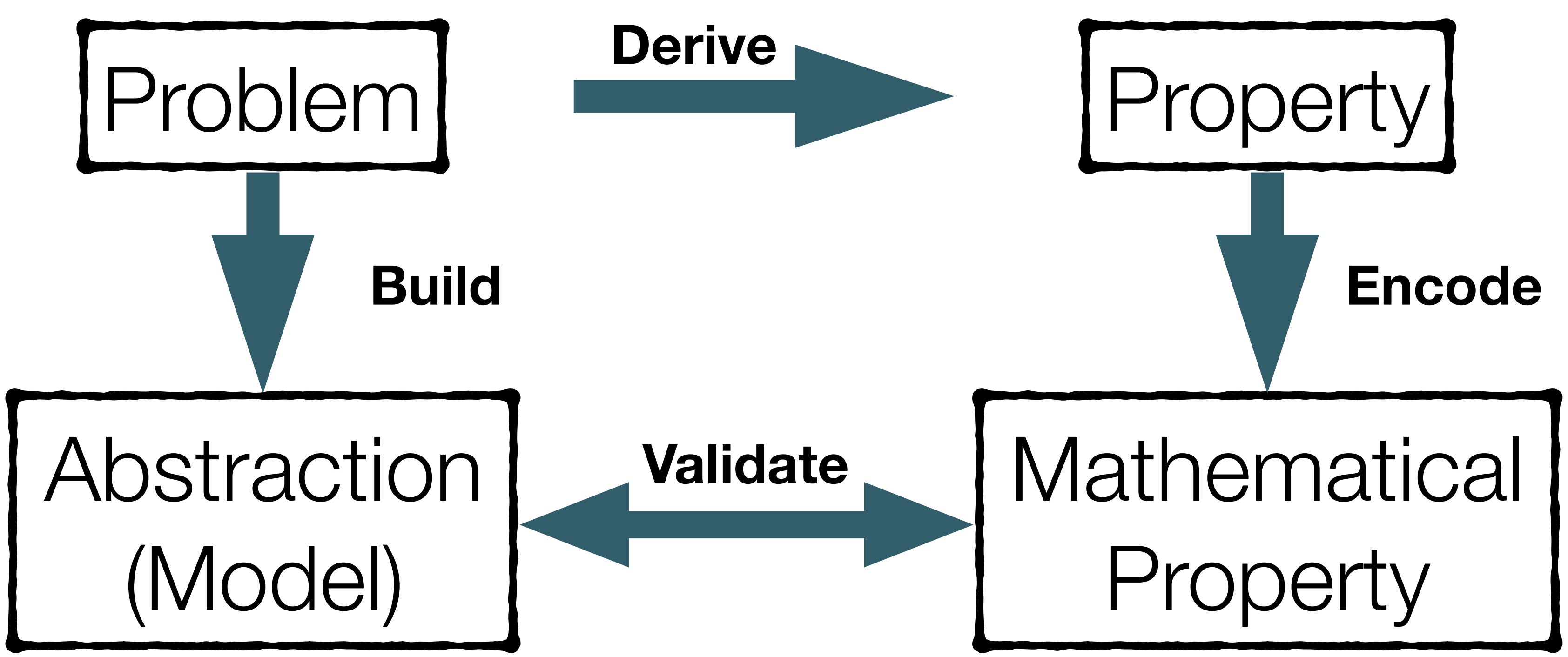
- What is known? What questions are still open?
  - Various degrees of completeness and formality of the approach
  - Must be performed as part of the validation
    - how does my approach differ from competing approaches?
    - answer may be provided both qualitatively and qualitatively, may require experiments or may be done through argumentation
  - Must be performed also as a prerequisite for the research effort, to understand the state of the art, to avoid re-addressing known problems with known solutions

Formal model

- How can we understand/explain the world?
  - build mathematical abstraction of a problem
    - analytical model (equations) a logical (specification) or discrete model (automata), a stochastic model (DTMC, SPN, QN)
    - prove important properties, based on inductive reasoning a deriving theorems
    - derive simulations

Early validation via formal models

- A FM supports validation prior to construction
- A FM is a mathematical formulation of an abstraction
- ICST is permeated by the notion of abstraction



Abstraction

[verb] trans.
Design and build (a machine or structure)
Oxford American Dictionaries

- ICST commonly builds abstract machines
- The tools used are themselves abstractions
- Languages are abstractions of the hardware
- Computer programs embed abstractions of the environment with which the abstract machine it realizes interacts

Model

- Abstractions often take the form of a (formal*) model

[noun]
A system or thing used as an example to follow or imitate a simplified description, esp. a mathematical one, of a system or process, to assist calculations or predictions

Oxford American Dictionaries

* formal synonym of mathematical

Usefulness of models

- To communicate
  - They embody a shared lexicon
    - E.g., the notions of state, transition
- To simplify descriptions and help focus, ignoring details that distract from the essence of the problem
- To reason about the modeled reality
  - Mathematics makes reasoning formal
  - Through models we can predict properties of the real system before it exists

What makes a good model?

- A model is good if it provides the right amount of information you need
  - It stays at the right level of abstraction
- A model abstracts from details
  - Make sure that they are details, not the essence
  - Be aware of the approximations
- A model serves a purpose
  - Different models for different purposes
- Expert judgment always needed!!!

Iterative model calibration

Models and validation

- Different models enable different kinds of validation
  - Logical models enable theorem proving
  - State models enable model checking (exhaustive state exploration)
  - Stochastic models enable simulations

Simulations

- What happens if…?
  - simulated because real world too expensive or even impossible in practice
  - often in conjunction with some real tests
  - may provide hints on scalability
  - require careful predefinition of what is included and what is excluded from the simulation scenario, which determine threats to validity of conclusions
  - examples: distributed systems, network protocols, modeling notations
    - performance evaluation

More on case studies

Spectrum of cases

created for *explanation*,
illustrate *differences*
foo/bar program excerpts
simple model

Toy example

accepted illustrative example
(e.g in teaching)
standard textbook example
simple but illustrates relevant issues

Exemplar

Case studies

real-life example
industrial system
open source system

Real case

competition
approved by community
as a way to compare

Community case

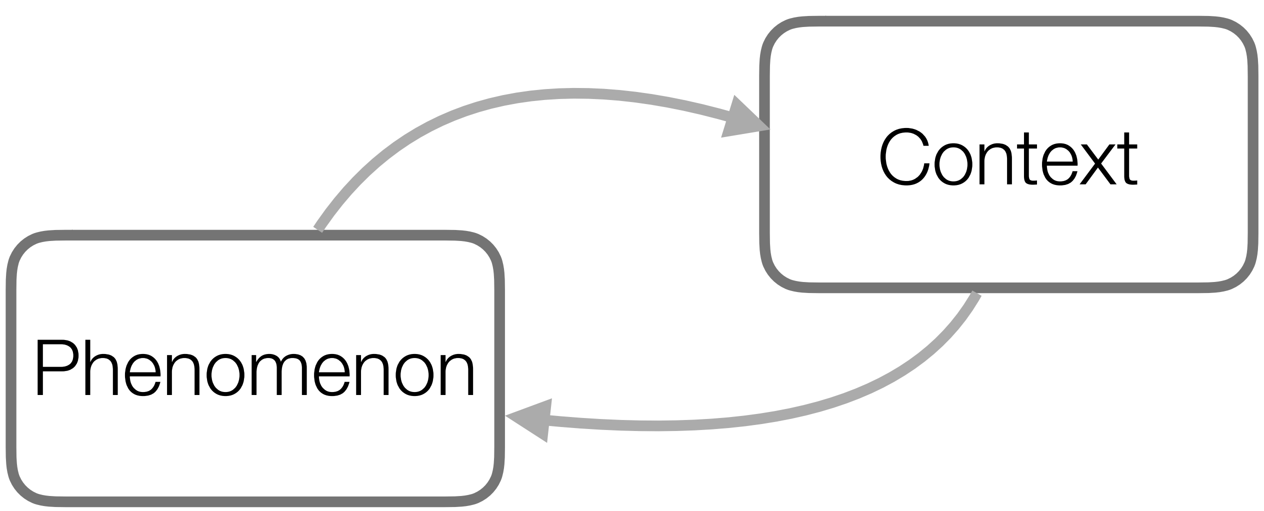
standard point of reference
against which to compare
approved by community
known context

Benchmark

Case study

- A case study is an empirical inquiry that investigates a contemporary phenomenon with its real-life context, especially when the boundaries between the phenomenon and context are not clearly evident [R.K Yin, Case Study Research: Design and Methods]

  - Boundaries between the phenomenon and context are not clearly evident, as opposed to *experiment*

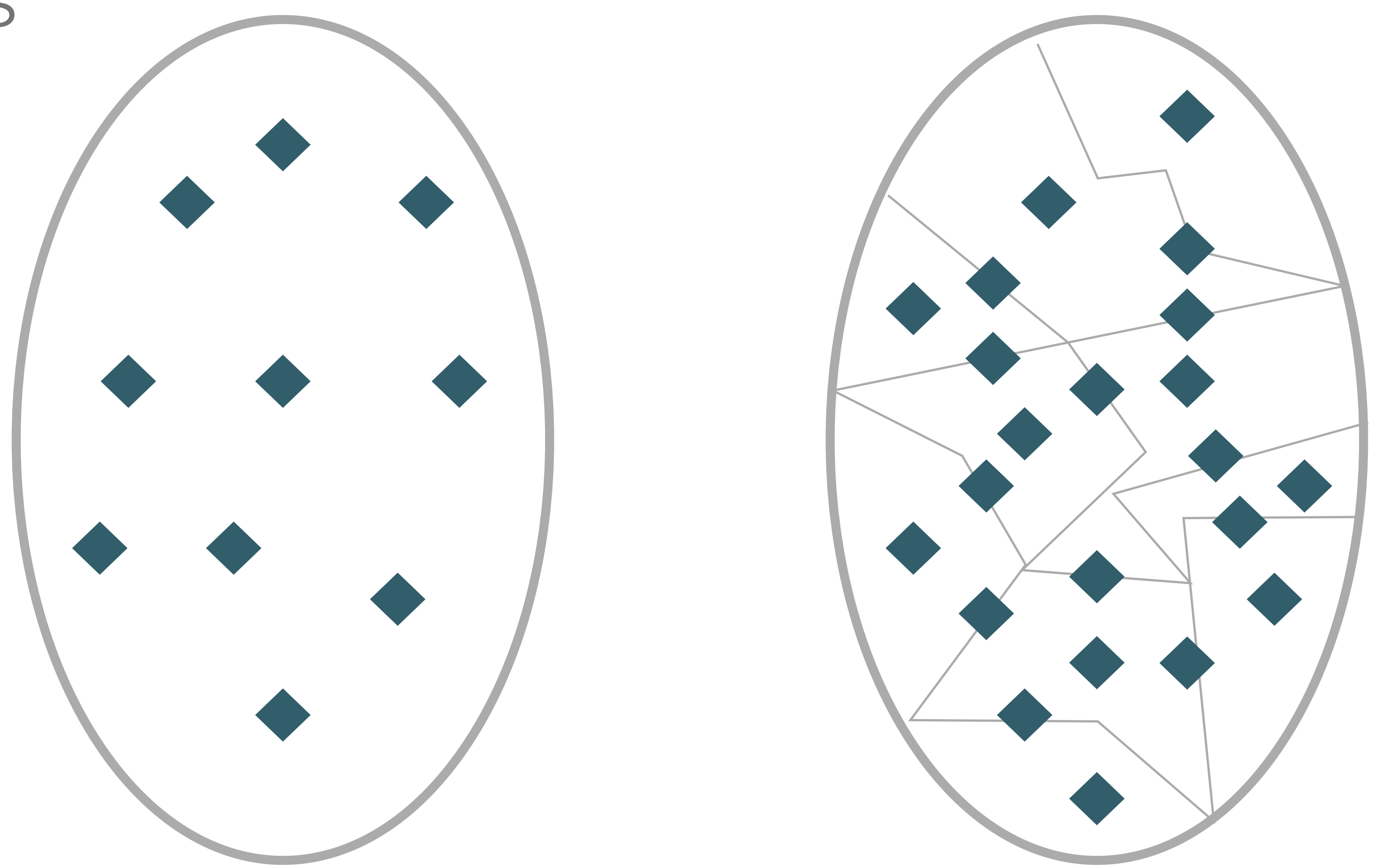


Critical issues

- Careful about generalizations on the basis of individual cases
- Case studies help understanding
- Look for counterexamples (null hypothesis)
  - try to falsify assumptions
    - just one black swan falsifies "all swans are white"
    - sampling logic vs. replication logic

Sampling logic vs replication logic

- Sampling logic analyzes randomly selected cases
- Replication logic carefully selects "modes" and "corner" cases

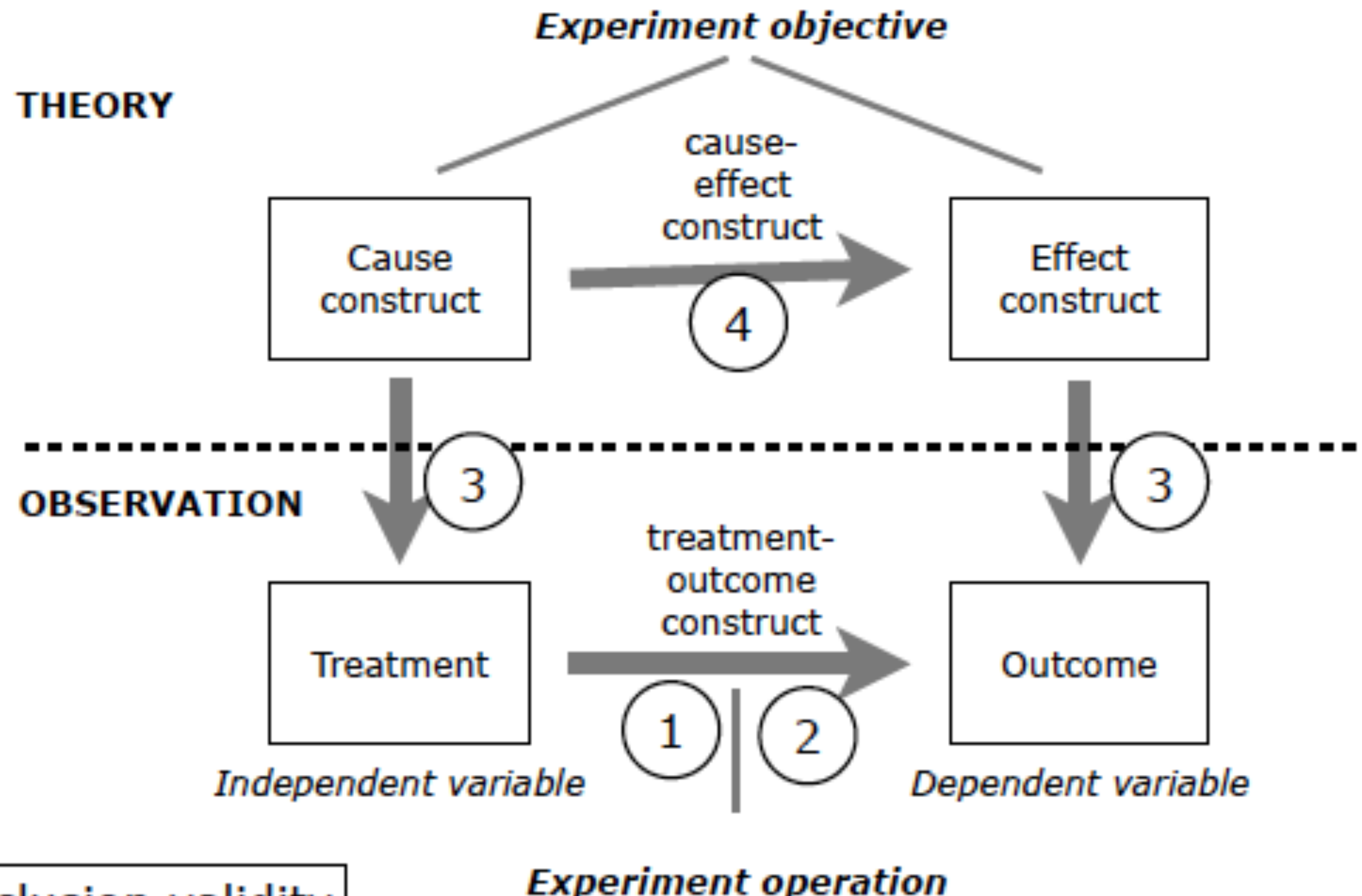


Purpose of case study

- **Exploratory**
  - Case studies used as initial investigations of some phenomena to derive new hypotheses and build theories
- **Confirmatory**
  - Case studies used to test existing theories. Especially important for refuting theories: a detailed case study of a real situation in which a theory fails may be more convincing than failed experiments in the lab

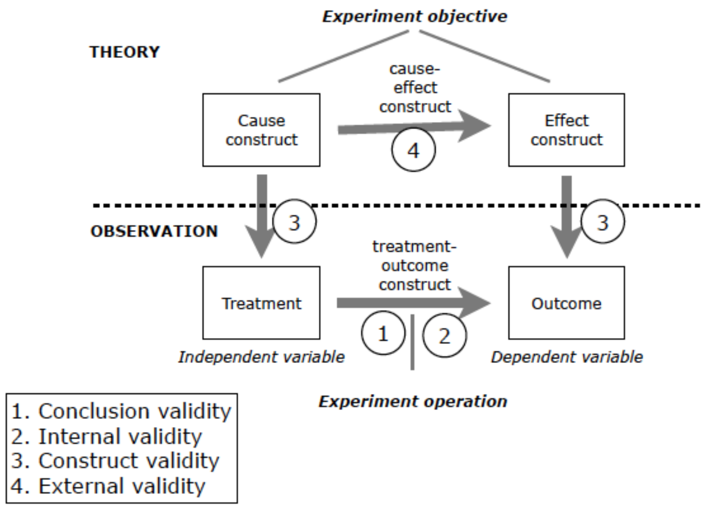
[S. Easterbrook et al. Selecting empirical methods for software engineering research]

Threats to validity for experiments



1. Conclusion validity
2. Internal validity
3. Construct validity
4. External validity

Threats to validity



- Construct validity
  - does the experimental setup reflect what the researcher had in mind? does it measure what it claims?
- Conclusion validity
  - does the way the experiment is conducted make the outcomes reliable? (e.g., is the tool used reliable?)
- Internal validity
  - extent to which a causal conclusion based on the study is warranted
    - causality vs correlation
- External validity
  - extent to which the conclusion is warranted to generalize to **other** contexts, validity of generalized (causal) inferences
    - a threat to external validity is an explanation of how you might be wrong in making a generalization

Risk management

- Threats to validity controlled by careful risk management
- Experiments are not perfect, but we must do our best to limit the chance of drawing wrong or inaccurate conclusions

Example to discuss

- The objective of the research experiment is to evaluate a set of programming languages used for introductory programming courses by assessing how they affect the programming mistakes made by students

Summary

- Research is essentially an iterative process
- The role of failures
- Essential skills and activities: reading, writing, presenting, building, connecting
- The essential role and multiple aspects of validating research results