Case study research in information systems engineering

Roel Wieringa
University of Twente
The Netherlands

Case study research in information systems engineering

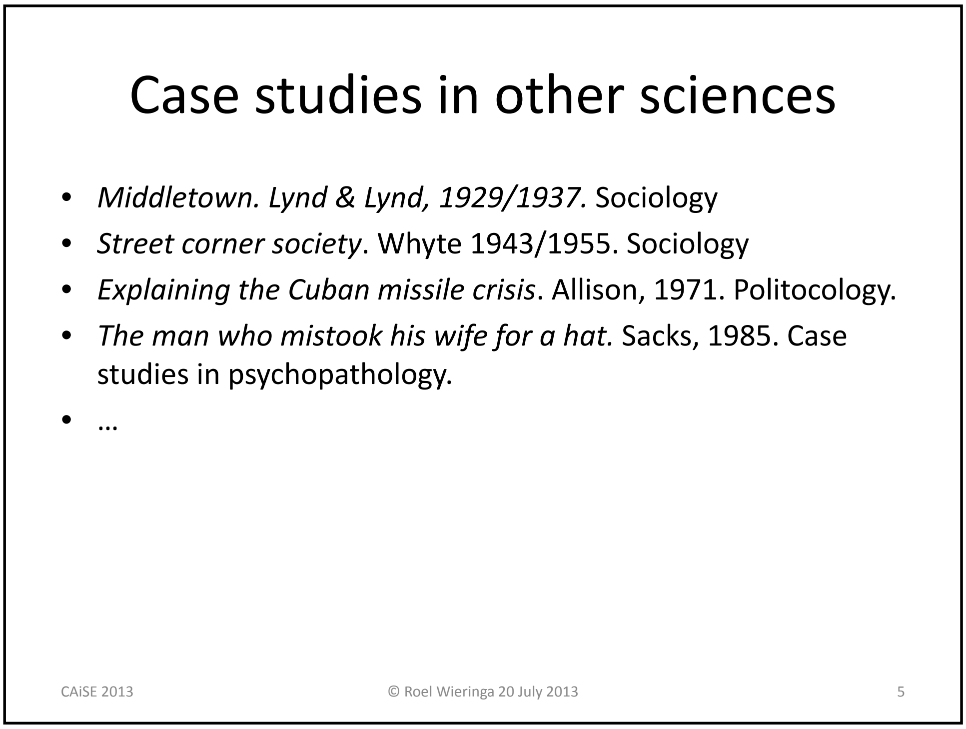
How to generalize, how not to generalize, and how not to generalize too much

Two things

1. Case study research
2. Generalizing from cases

Definition

- In a case study, a case is studied in its real world context
- (TBD: What is a case?)



Case studies in other sciences

- *Middletown. Lynd & Lynd, 1929/1937.* Sociology
- *Street corner society*. Whyte 1943/1955. Sociology
- *Explaining the Cuban missile crisis*. Allison, 1971. Politocology.
- *The man who mistook his wife for a hat.* Sacks, 1985. Case studies in psychopathology.
- …

Case studies require a **journalistic attitude**

- Reporting the facts
  - Detailed, concrete, accurate
  - Lack of control over events by reporter
  - Detached, neutral
    - Eye-witness reports
    - Independent fact-checking
    - More than one measurement instrument
    - Reporter has no opinion

Case studies require a **scientific attitude**

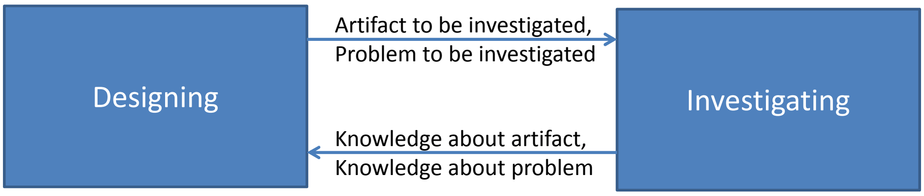
1. Using scientific theory
   - To ask research questions
   - To describe facts
   - To explain facts
   - To generalize to other cases
2. Acknowledging **fallibility** of theory
   - Submit case study to the critique of others: **Peer review**.
   - Submit theory to the test of observations: **Test** in future case studies

To prevent theories turning into opinions

What, why and how of case studies in Information Systems Engineering

ISE is a design science

- Design science is the design and study of artifacts in context
  - Artifacts: notations, techniques, methods, etc.
  - Context: IS engineering

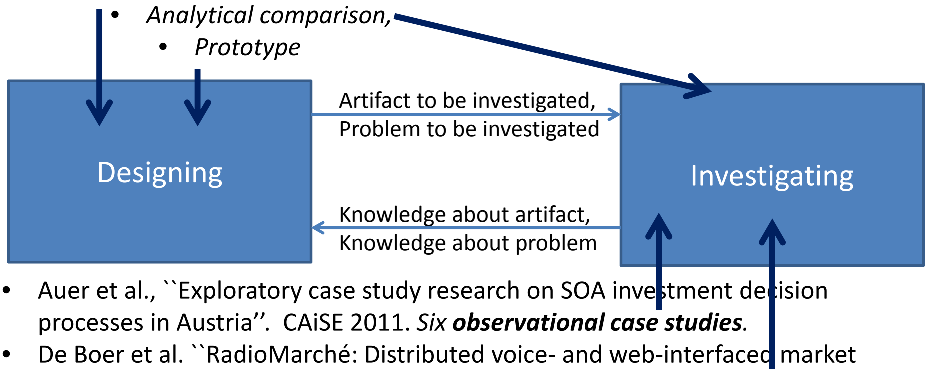


Designing an **artifact,**
Solving a stakeholder problem,
Achieving their goals

Answering a knowledge **question,**
Describing, explaining, generalizing

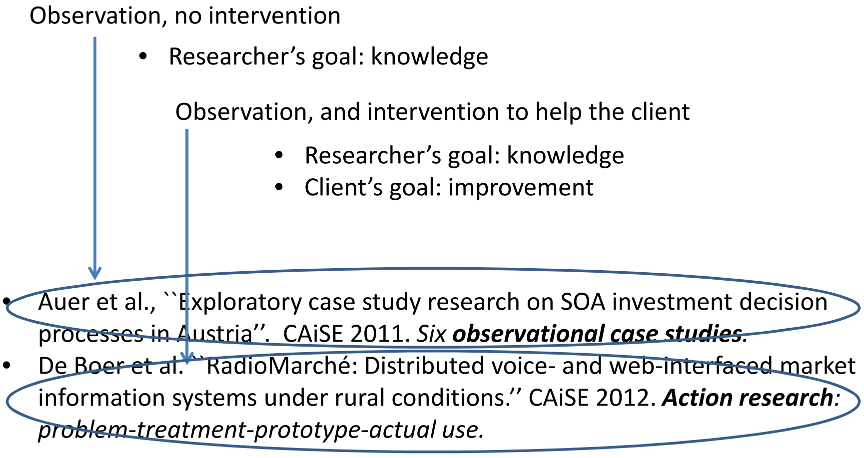
Examples

- Ly, Rinderle-Ma, Dadam. ``Design and Verification of Instantiable Compliance Rule Graphs in Process-Aware Information Systems'', CAiSE 2010.
  - *Design an artifact that improves something for stakeholders,*
    - *Analytical comparison,*
      - *Prototype*



- Auer et al., ``Exploratory case study research on SOA investment decision processes in Austria''. CAiSE 2011. *Six **observational case studies**.*
- De Boer et al. ``RadioMarché: Distributed voice- and web-interfaced market information systems under rural conditions.'' CAiSE 2012. ***Action research**: problem-treatment-prototype-actual use.*

Examples



Observation, no intervention

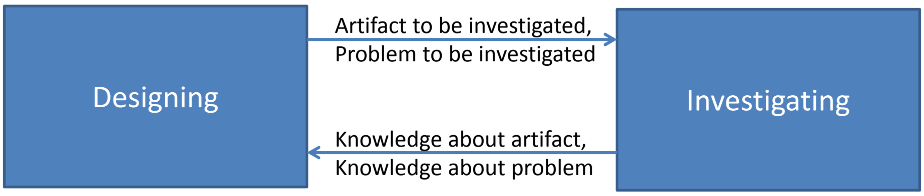
- Researcher’s goal: knowledge

Observation, and intervention to help the client

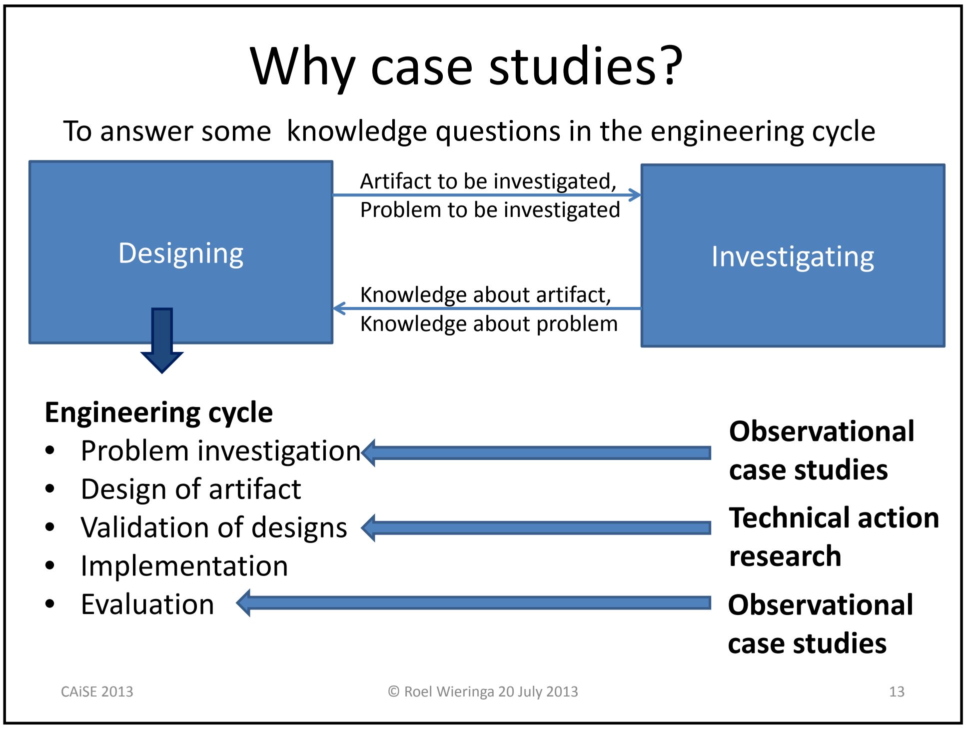
- Researcher’s goal: knowledge
- Client’s goal: improvement

- Auer et al., ``Exploratory case study research on SOA investment decision processes in Austria’’. CAiSE 2011. *Six* ***observational case studies.***
- De Boer et al. ``RadioMarché: Distributed voice- and web-interfaced market information systems under rural conditions.’’ CAiSE 2012. ***Action research:*** *problem-treatment-prototype-actual use.*

ISE is a design science



- Conceptual analysis
- Empirical research: surveys, experiments, **case studies, action research**
- Meta-research



Why case studies?

To answer some knowledge questions in the engineering cycle

Designing

Artifact to be investigated,
Problem to be investigated

Knowledge about artifact,
Knowledge about problem

Investigating

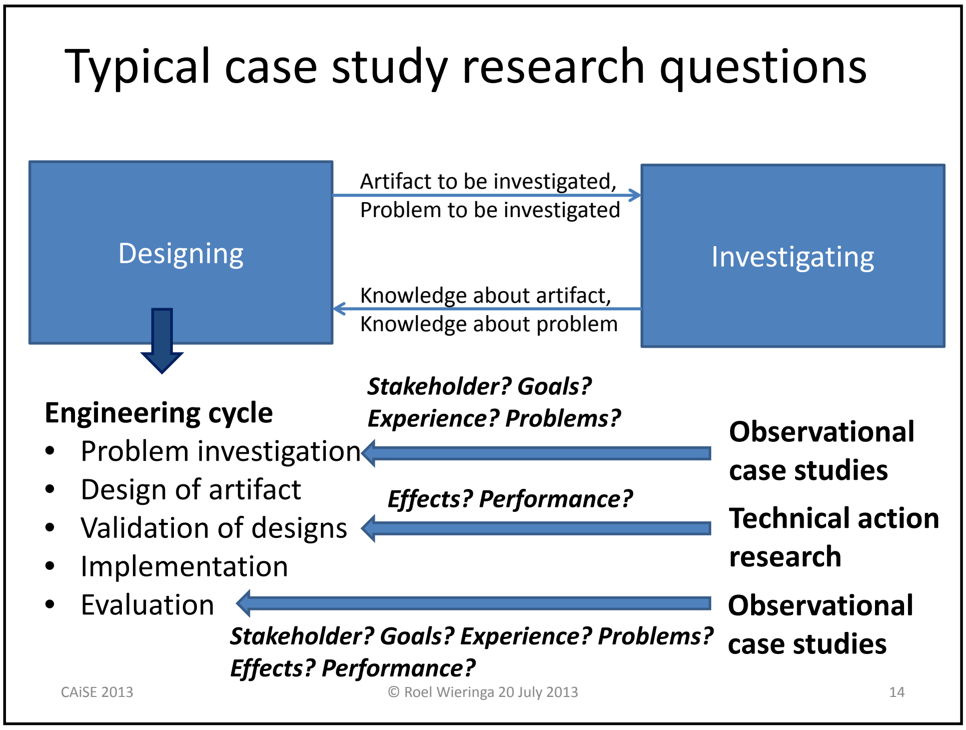
Engineering cycle

- Problem investigation
- Design of artifact
- Validation of designs
- Implementation
- Evaluation

Observational case studies

Technical action research

Observational case studies



Typical case study research questions

Designing

Artifact to be investigated,
Problem to be investigated

Knowledge about artifact,
Knowledge about problem

Investigating

Engineering cycle

- Problem investigation
- Design of artifact
- Validation of designs
- Implementation
- Evaluation

Stakeholder? Goals? Experience? Problems?

Observational case studies

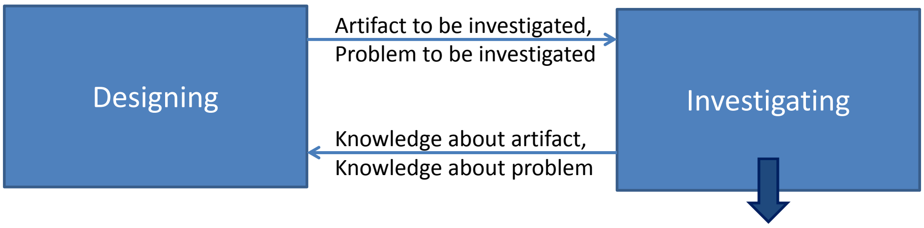
Effects? Performance?

Technical action research

Observational case studies

Stakeholder? Goals? Experience? Problems? Effects? Performance?

How to do case studies?



Empirical cycle
- Research problem
- Research design
- Validation of design
- Execution
- Analysis

The empirical cycle

- Research problem investigation
  - **Conceptual framework, theory, research questions, population to which you want to generalize**
- Research design
  - **Object of study, measurement instruments, treatment (if any), inferences to be done**
- Research design validation
- Research execution
- Analysis of results (inferences from the data)
  - **Descriptions, explanations, generalization**

Observational case study

Example 1
Damian & Chisan. RE introduction in a development company. TSE July 2006.

- Research problem
  - *What are the effects of RE improvements on productivity, quality, and risk management?*
  - *RE and SE concepts*
  - *No theory ….?*
- Research design
  - *Object of study: A global development organization*
  - *Interviews, questionnaires*
  - *No inferences planned….*
- Validity
- Execution
- Analysis
  - Observations
    - *Descriptive statistics, correlations of opinions*
  - Explanations
    - *Improvements in P, Q and RM were caused by RE improvement;*
    - *or by new management;*
    - *or by other processes*
  - Generalizations
    - *Same effects will occur in organizations with similar problems*

Technical action research

Example 2:
Morali & Wieringa: Confidentiality risk assessment in outsourcing. RE 2010

- Research problem
  - *Does the method CRAC++ (designed by Morali) support risk assessment?*
  - *Without disclosing confidential info?*
  - *Easy to use? Repeatable?*
  - *The method is the theory*
- Research design
  - *Objects of study: Two organizations*
  - *Treatment: Consulting using CRAC++*
  - *Measurement: Diary, interviews, work products*
  - *No inferences planned …*
- Validity
- Execution
- Analysis
  - Observations
    - *Work products, interview results*
  - Explanations
    - *Structure of the problem & structure of the method explain results*
  - Generalizations
    - *Same effects will occur in organizations with similar problems*

- These explanations seem almost analytically true!

- Yes.
- We need a real-world case to see if the analysis also stands in the real world

2. Generalizing from case studies

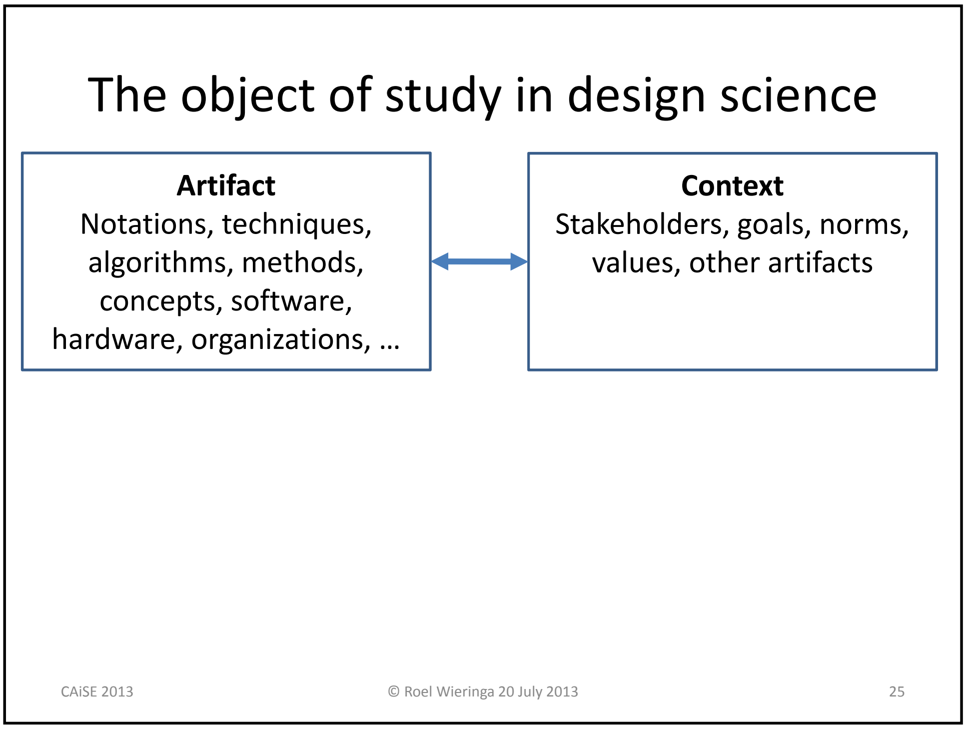
- If we could not generalize from a single case, then we could not learn from a single case.

Two ways to generalize

- Case-based
  - By analogy
  - From **case** to **case**

Not for case studies

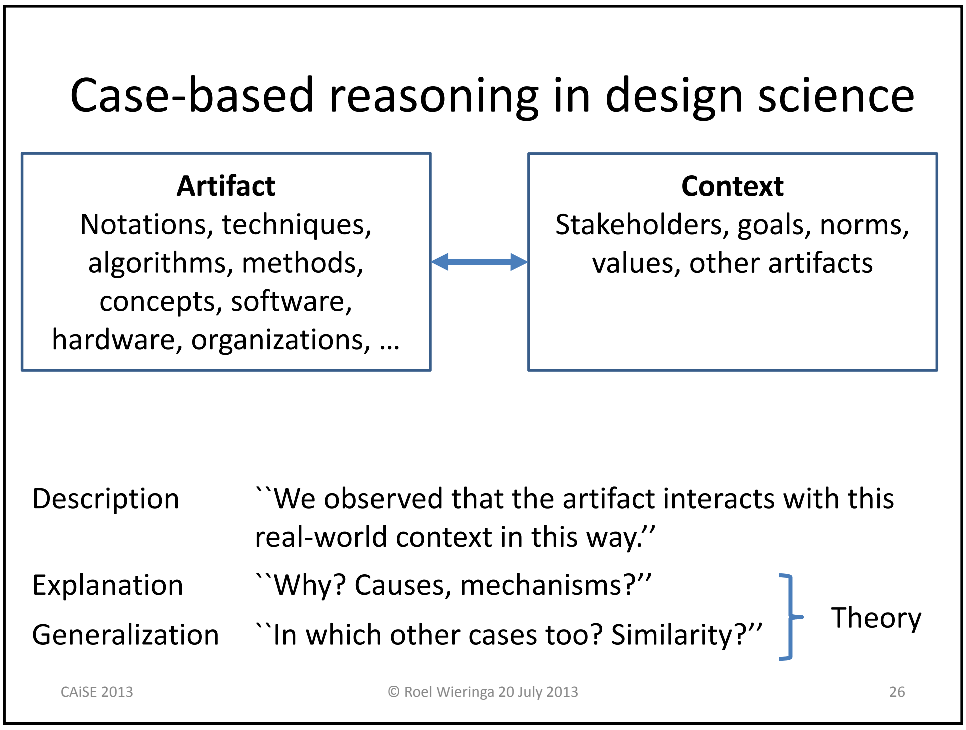
- ~~Sample-based~~
  - ~~Random sample of at least 30 cases. 100 is better.~~
  - ~~Few variables~~
  - ~~From averages, variances, correlations of variables over cases in the **sample** …. generalize to the **population**~~



The object of study in design science

Artifact
Notations, techniques, algorithms, methods, concepts, software, hardware, organizations, …

Context
Stakeholders, goals, norms, values, other artifacts

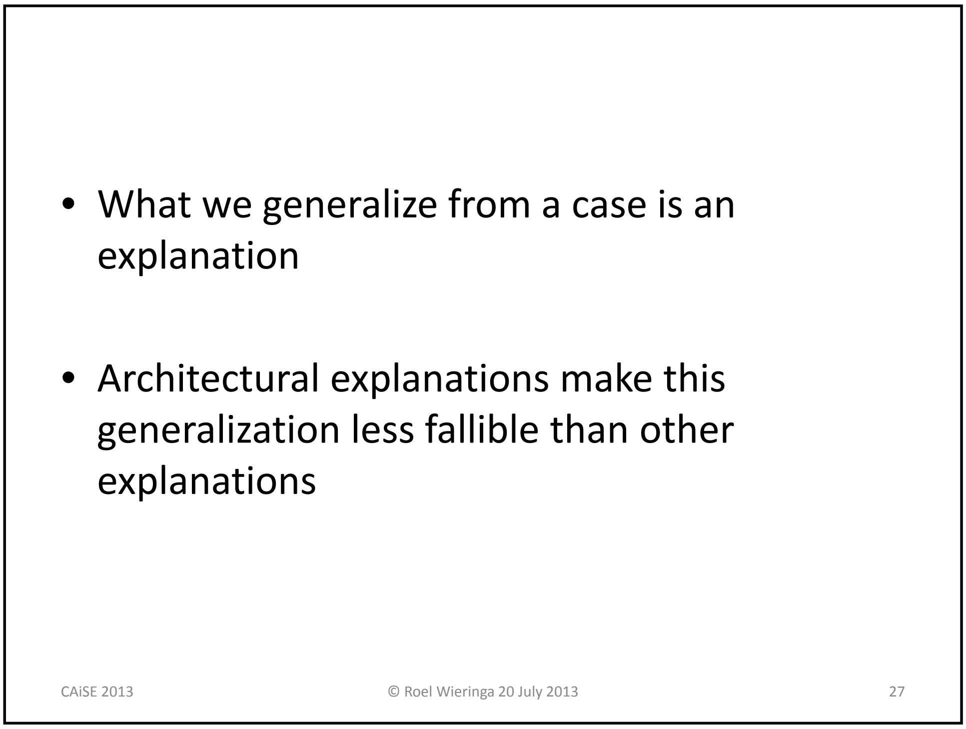


Case-based reasoning in design science

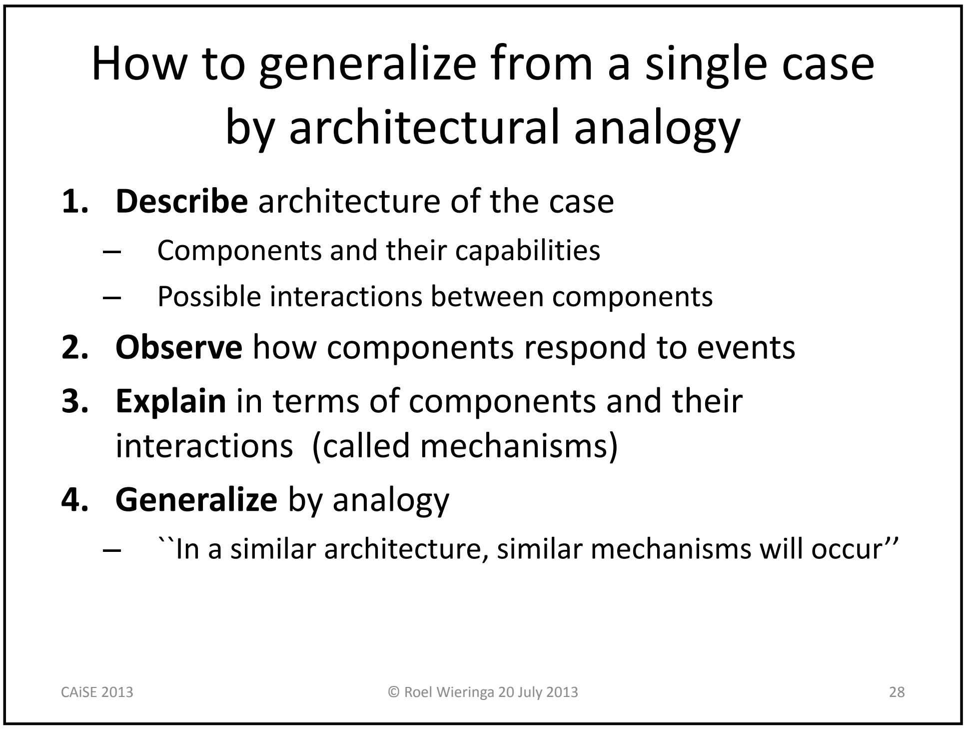
Artifact
Notations, techniques, algorithms, methods, concepts, software, hardware, organizations, …

Context
Stakeholders, goals, norms, values, other artifacts

Description	``We observed that the artifact interacts with this real-world context in this way.’’	
Explanation	``Why? Causes, mechanisms?’’	Theory
Generalization	``In which other cases too? Similarity?’’	



- What we generalize from a case is an explanation
- Architectural explanations make this generalization less fallible than other explanations



How to generalize from a single case by architectural analogy

1. **Describe** architecture of the case
   - Components and their capabilities
   - Possible interactions between components
2. **Observe** how components respond to events
3. **Explain** in terms of components and their interactions (called mechanisms)
4. **Generalize** by analogy
   - ``In a similar architecture, similar mechanisms will occur''

How to continue by analytical induction

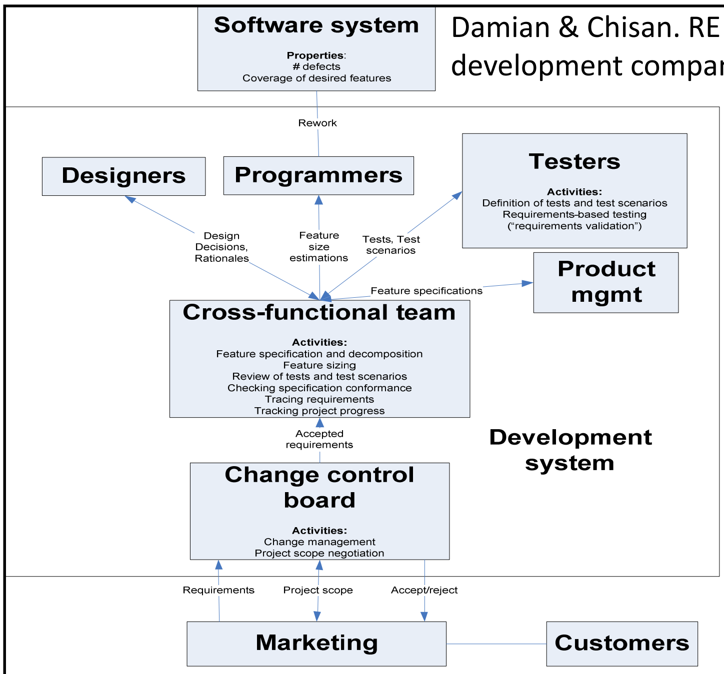
5. To **confirm**
   - Find a similar case
   - Test if prediction is correct
6. To **falsify**
   - Find a dissimilar case
   - Test if prediction occurs anyway
7. In both cases, **refine**
   - The conceptual framework and/or
   - The generalization

How to generalize from a single case by architectural analogy

1. **Describe** architecture of the case
   - Components and their capabilities
   - Interactions between components
2. **Observe** how components respond to events
3. **Explain** in terms of components and their interactions (which we call mechanisms)
4. **Generalize** by analogy
   - ``In a similar architecture, similar mechanisms will occur’’

Explanation

- **Causal explanation** refers to variables
  - ``Temperature increased because pressure increased''
  - ``Account balance increased because interest is high''
- **Mechanical explanation** refers to components and interactions
  - ``Pressure increase produces more collisions between gas particles, which raises their kinetic energy by which the gas becomes hotter''
  - Mechanisms can explain causality between variables
- **But are there social mechanisms?**
- **Yes!**



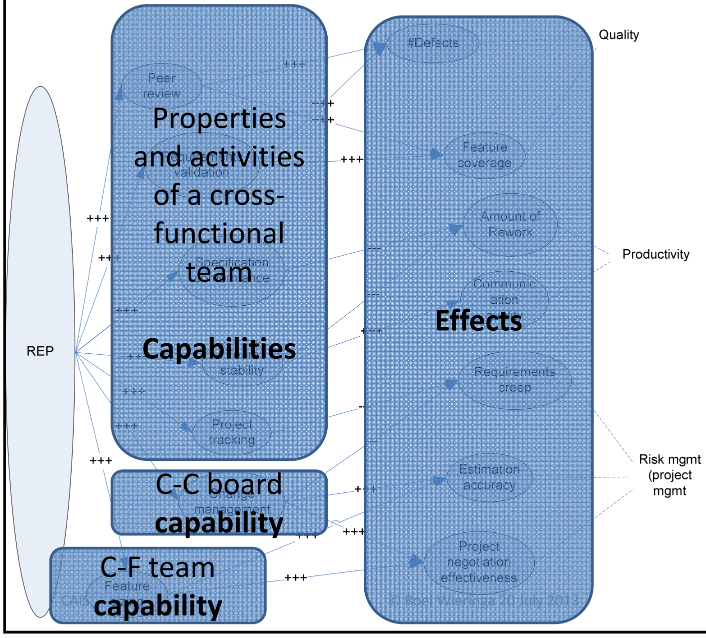
Damian & Chisan. RE introduction in a development company. TSE July 2006.

An architectural model. Mechanisms that explain improved productivity, quality, risk mgmt:

C-F team
- improves shared feature understanding,
- reduces rework

C-C board
- prevents requirement creep

Damian & Chisan. RE introduction in a development company. TSE July 2006.



A variable-based model

Observed correlations

Variable-based and architectural explanations

- Variable-based explanations refer to causes
  - Change in X causes change in Y
- Architectural explanations refer to mechanisms
  - Components with capabilities and interactions
  - Mechanism = interaction between components that produces effects
- A mechanism can explain a cause-effect relation

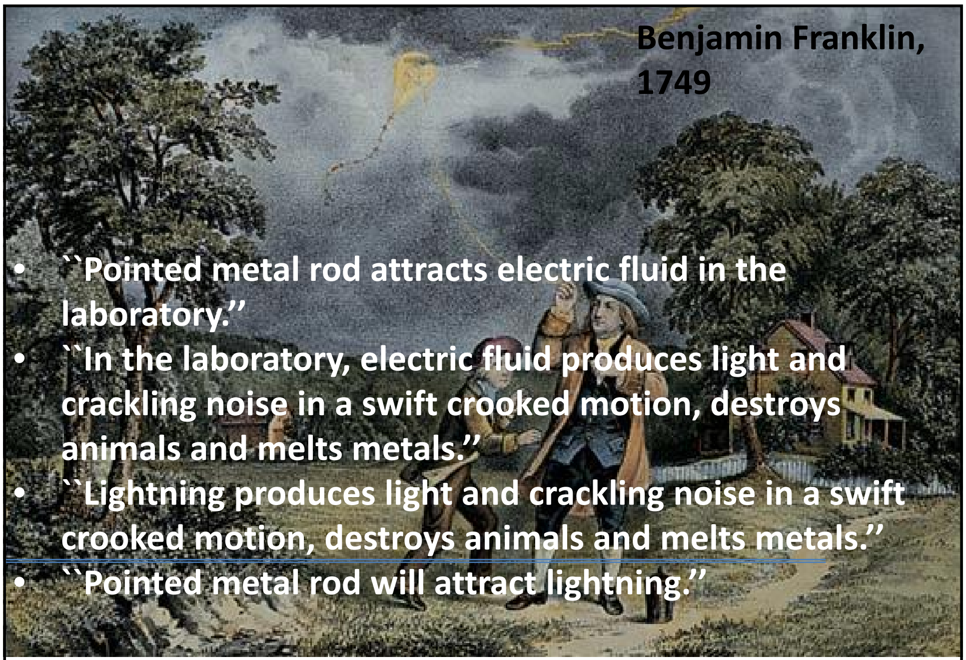
How to generalize from a single case by architectural analogy

1. **Describe** architecture of the case
   - Components and their capabilities
   - Interactions between components
2. **Observe** how components respond to events
3. **Explain** in terms of components and their interactions (which we call mechanisms)
4. **Generalize** by analogy
   - ``In a similar architecture, similar mechanisms will occur’’

Variable-based analogy

- ``Piet is a computer scientist from NL, is tall and likes cheese.’’
- ``Anne is a computer scientist from NL’’.

- ``Anne is tall and likes cheese’’



Benjamin Franklin, 1749

- ``Pointed metal rod attracts electric fluid in the laboratory.''
- ``In the laboratory, electric fluid produces light and crackling noise in a swift crooked motion, destroys animals and melts metals.''
- ``Lightning produces light and crackling noise in a swift crooked motion, destroys animals and melts metals.''
- ``Pointed metal rod will attract lightning.''

Variable-based analogy

- Looks at similarity in features
- It may lead to a correct generalizations, if we talk about a natural kind.
- ``If it walks like a duck and sounds like a duck, it is a duck.''

- But: it is very unreliable reasoning
  - Similar to sympathetic magic
- Mechanism that can support conclusion is unknown

Component-based analogy

- ``Project X
  - develops an information system,
  - it has more than 15 people in the delivery team, and
  - it is too late due to coordination and communication overhead among developers
- ``Project Y
  - develops an information system,
  - it has more than 15 people in the delivery team.''
- ``Project Y will be late due to coordination and communication overhead among developers too.''

Mechanism

We generalize the mechanism

Example 2: Morali & Wieringa: Confidentiality risk assessment in outsourcing. RE 2010

- Architecture
  - Manufacturing company A, outsourcing ERP administration to outsourcing service provider B
  - Employees of B have access to info in ERP system
  - Sarbanes-Oxley compliance requirement on A
  - Auditors of A do not have access to B's IT architecture
  - Security analyst has access to shared outsourcing architecture
- This results in a deadlock mechanism: auditors of A cannot give compliance judgment

- The CRAC++ method introduces a new mechanism by which IT confidentiality risks can be assessed,
  - which allows renegotiation of SLA,
  - which allows auditors to give judgment

- NB the method is the theory is the mechanism.
- Versus natural mechanisms

Summary of single-case generalization

1. Describe architecture of the case
   - Components and their capabilities
   - Interactions
2. Observe emergent effects
3. Explain effects in terms of mechanisms.
4. Generalize by analogy
   - ``In a similar architecture, similar mechanisms will occur’’

However ….

- Architectural analogy is fallible
  - What if we misunderstood the mechanism?
  - What if we misjudge similarity?
  - What if in the next similar case, other mechanisms defeat the one we observed?

How to continue by analytical induction

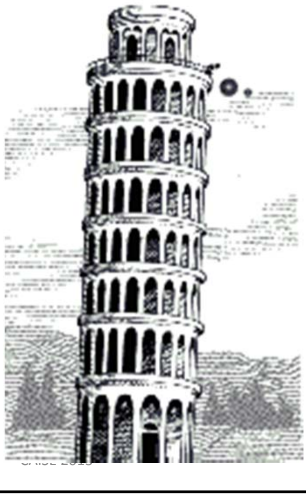
5. To **confirm** an architectural generalization
    - Find an analogous case
    - Test if mechanism has same effect
6. To **falsify**
    - Find a dissimilar case
    - Test if effect occurs anyway
7. In both cases, **refine** to match all cases so far
    - Improve the conceptual framework and/or
    - Improve the generalization

A. Mockus, R. Fielding, and J. Herbsleb,
“Two case studies of open source software development:
Apache and Mozilla,” TOSEM july 2002.

- Architecture: open source development project.
- After first case study:
    - Mechanism: A core developer team larger than 15 people gets overwhelmed by communication & coordination overhead
- Second case study falsified this.
    - Mechanism a: A core developer team larger than 15 people **without defined process as in Apache**, gets overwhelmed by communication & coordination overhead
    - Mechanism b: If core developer team has a **defined process, as in Mozilla**, it may consist of up to 36 developers without being overwehelmed by C & C overhead

- Why is case-based induction analytical?
  - The emergent effects follow analytically from the architecture

- Why is case-based induction analytical?
  - The emergent effects follow analytically from the architecture



- Galileo Galilei (1564-1642):
- ``Two cannon balls of equal weight fall equally fast.
- Put a string between them.
- The resulting object is twice as heavy but falls at the same speed as the original two cannon balls''

- Galileo never did the experiment
- He knew what the outcome would be because he had an analytical argument for it

- He unwittingly made an idealizing assumption: no air resistance
- Need to test the generalization in conditions of practice! Case studies needed

Galilean idealization

- Galileo used idealization to understand the real world
  - Point masses
  - Frictionless surfaces
  - …
- We do that in computer science too
  - Turing machines
  - Infinite data types
  - ….

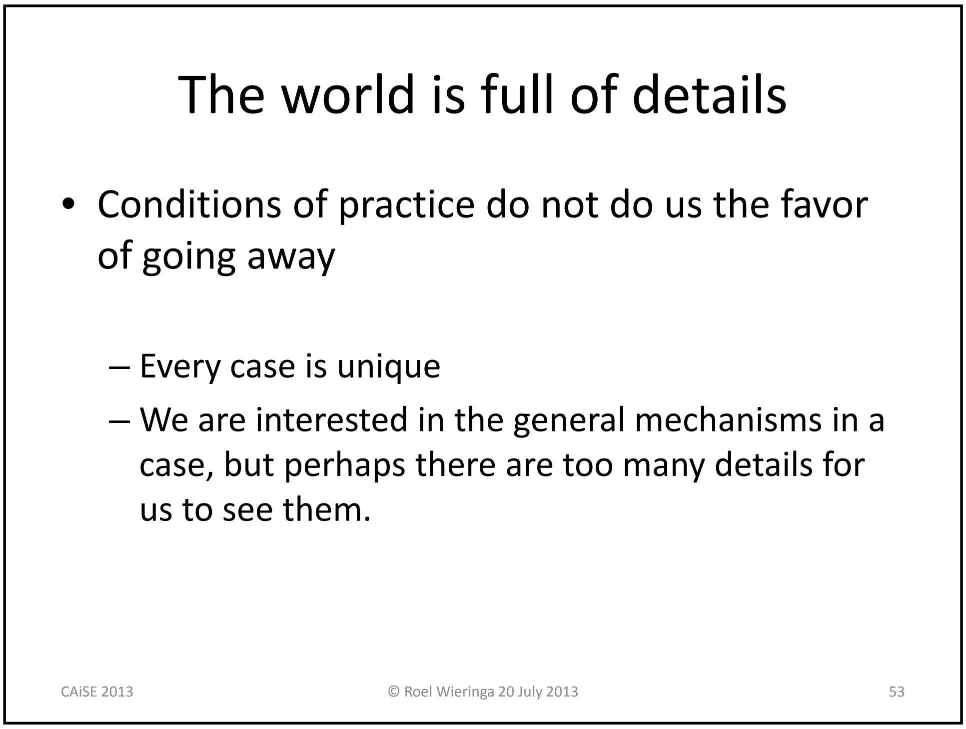
Idealization in basic science and design science

- Basic scientist approximate idealizations in the laboratory
  - Laboratory experiments are similar to the ideal case
  - Replications of idealized lab experiments are identical
  - Research budget spent on creating ideal conditions.

- Contrast with design science
  - We spend our budget on simulating real-world conditions
  - and therefore on doing case studies.

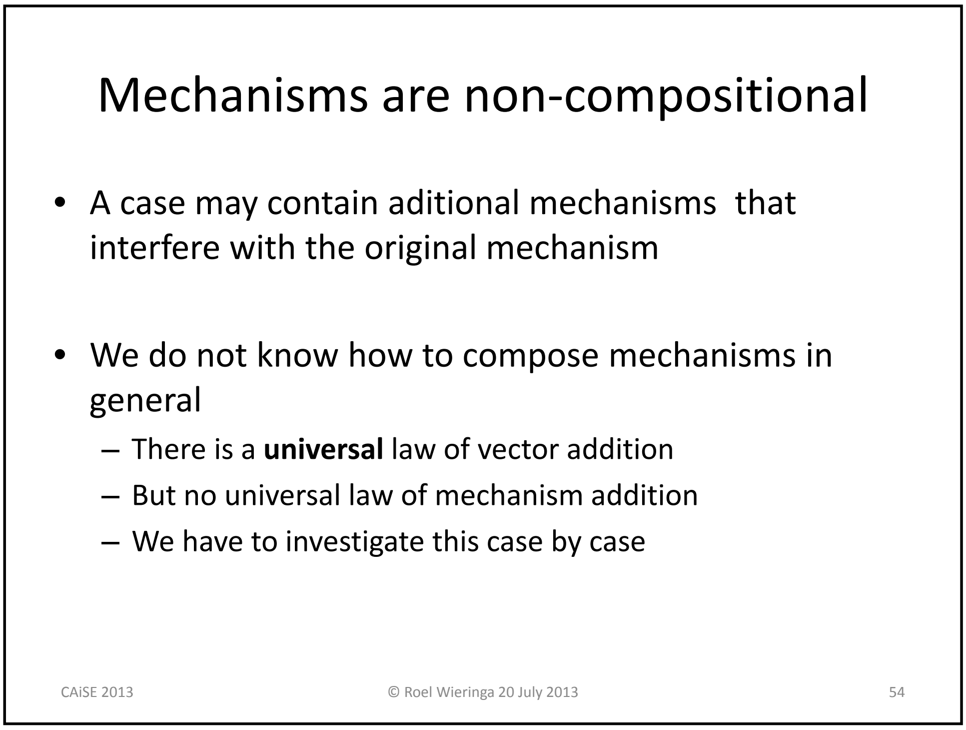
The real world is not ideal

- How to apply idealizing laws of nature?
  - You can’t.
  - First you must drop the idealizations



The world is full of details

- Conditions of practice do not do us the favor of going away
  - Every case is unique
  - We are interested in the general mechanisms in a case, but perhaps there are too many details for us to see them.



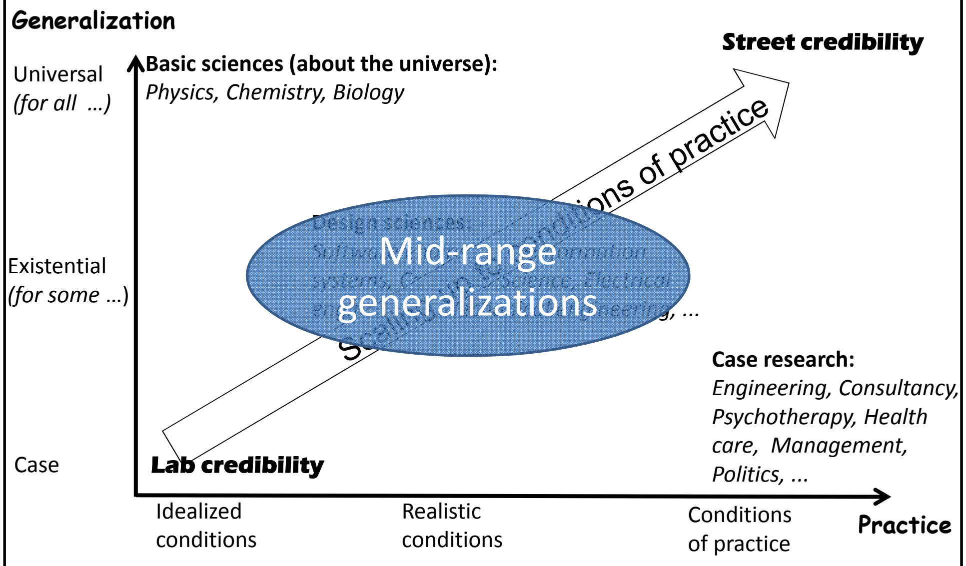
Mechanisms are non-compositional

- A case may contain aditional mechanisms that interfere with the original mechanism
- We do not know how to compose mechanisms in general
  - There is a **universal** law of vector addition
  - But no universal law of mechanism addition
  - We have to investigate this case by case

Case-based generalization is limited

- Researchers have to reason case by case
  - Analytical induction is a way to check if we have dropped idealizations in the right way
  - Generalizations are limited and uncertain.
- Practitioners have to do a risk assessment case by case
  - What is the risk of applying the wrong generalization?
  - What is the risk of missing the right generalization?

Mid-range generalizations



Generalization

Universal
(for all …)

Existential
(for some …)

Case

Basic sciences (about the universe):
Physics, Chemistry, Biology

Street credibility

Mid-range generalizations

Case research:
Engineering, Consultancy, Psychotherapy, Health care, Management, Politics, …

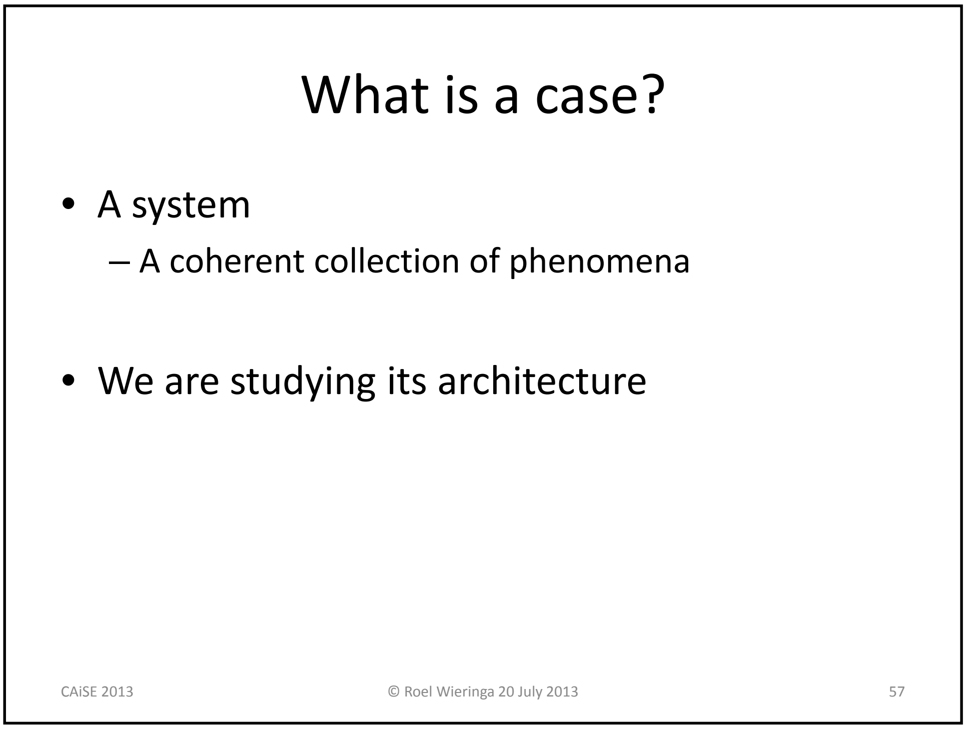
Lab credibility

Idealized conditions

Realistic conditions

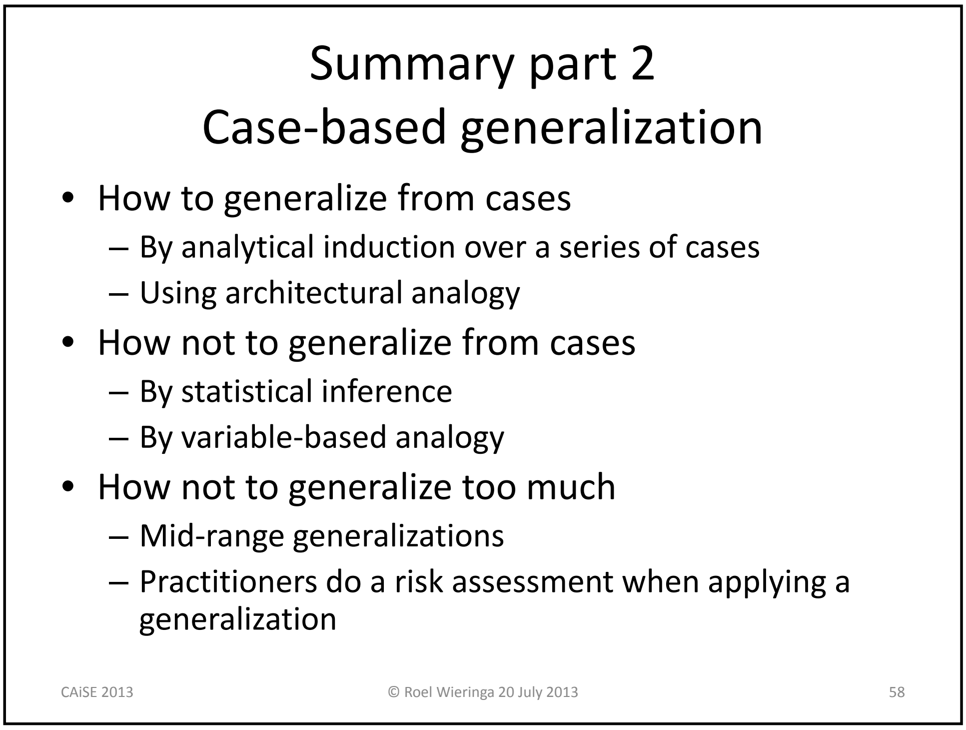
Conditions of practice

Practice



What is a case?

- A system
  - A coherent collection of phenomena
- We are studying its architecture



Summary part 2
Case-based generalization

- How to generalize from cases
  - By analytical induction over a series of cases
  - Using architectural analogy
- How not to generalize from cases
  - By statistical inference
  - By variable-based analogy
- How not to generalize too much
  - Mid-range generalizations
  - Practitioners do a risk assessment when applying a generalization

Thank you!