A Model and Sensitivity Analysis of the Quality Economics of Defect-Detection Techniques

Stefan Wagner
Software & Systems Engineering
Technische Universität München
Germany
wagnerst@in.tum.de

July 18, 2006

Introduction

Motivation

- Analytical software quality assurance (defect-detection techniques like reviews, tests, . . .) is still the most important method to improve the quality of software.
- The costs for those techniques are significant:
  - Myers (1979) estimates that 50% of the total development costs can be attributed to testing.
  - Jones (1987) states that 30–40% of the development costs are for quality assurance and defect removal.
  - The *National Institute of Standards and Technology* (2002) even assigns 80% to the detection and removal of defects.
- There is a huge opportunity for cost savings in that area.

Problem

- This results in three questions:
  1. Which techniques should be used?
  2. In what sequence?
  3. With what effort for each technique?
- Ntafos (2001): "Cost is clearly a central factor in any realistic comparison but it is hard to measure, data are not easy to obtain, and little has been done to deal with it."
- Rai et al. (1998): "A better understanding of the costs and benefits of SQA and improvements to existing quantitative models should be useful to decision-makers."

Overview

Introduction

Software Quality Economics

Example

An Analytical Model

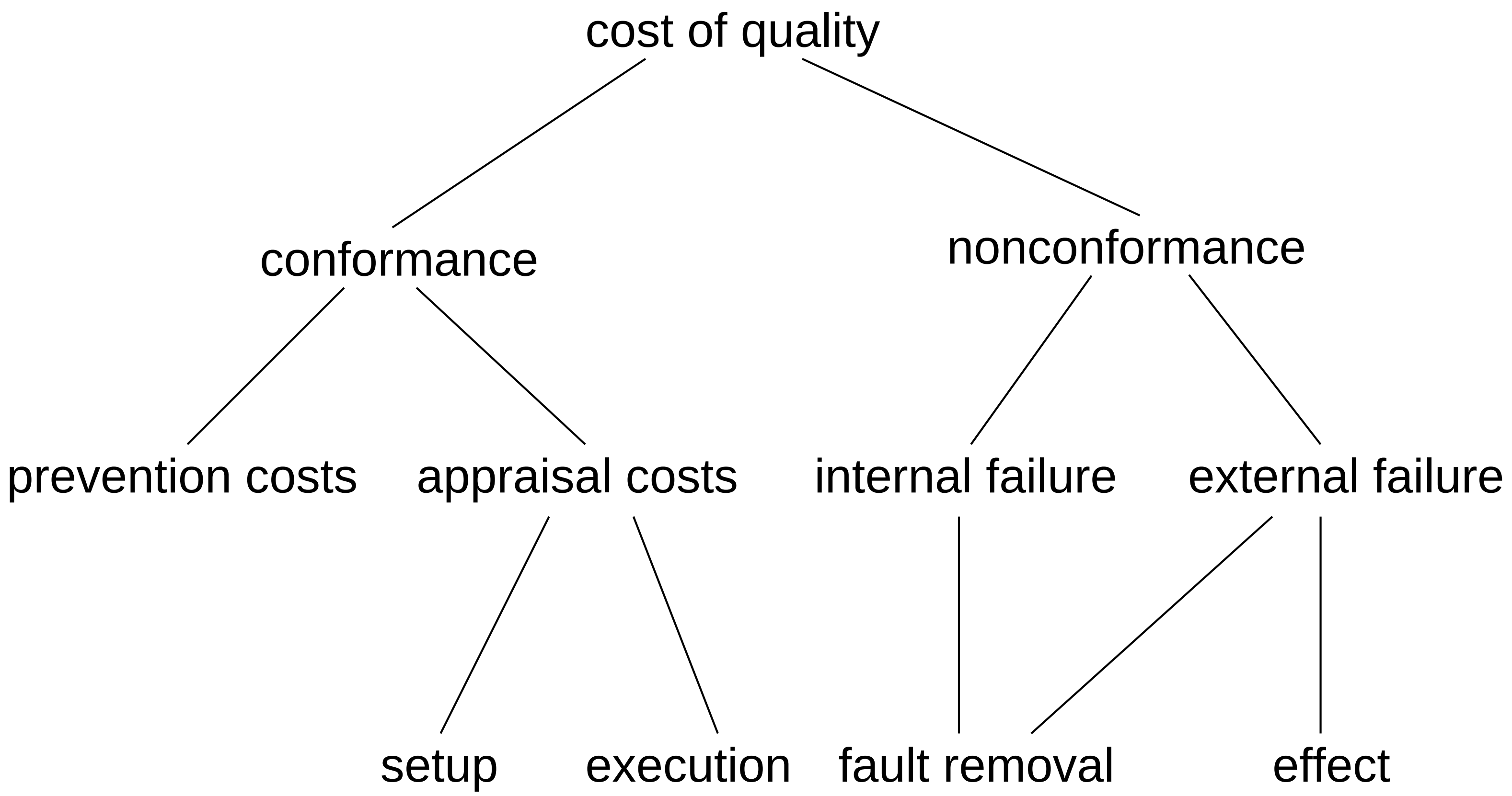
Sensitivity Analysis

Conclusions

Software Quality Economics

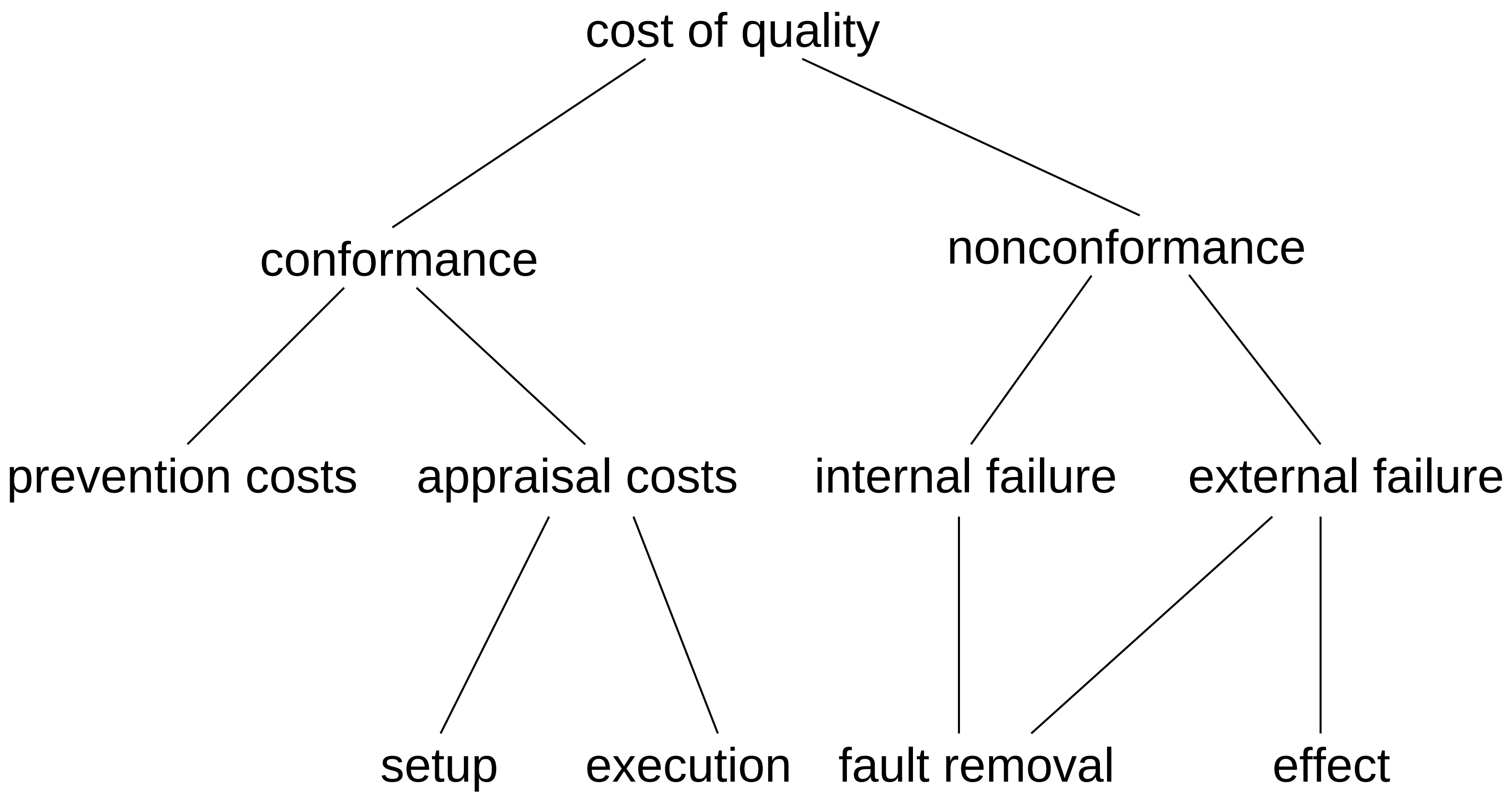
Quality Economics

- Quality economics is the cost/benefit analysis w.r.t. quality
- The benefits are the saved costs



Quality Economics

- Quality economics is the cost/benefit analysis w.r.t. quality
- The benefits are the saved costs



- Focus on *analytical* QA in the following

Example

A Software Project

- Suppose we have to do the quality assurance for a software project
- Maybe some control unit in an automobile
- It consists of 100 KLOC of C code

A Software Project

- Suppose we have to do the quality assurance for a software project
- Maybe some control unit in an automobile
- It consists of 100 KLOC of C code
- Usually we use the techniques
  - code inspection (A),
  - unit test (B),
  - integration test (C), and
  - system test (D)

A Software Project

- Suppose we have to do the quality assurance for a software project
- Maybe some control unit in an automobile
- It consists of 100 KLOC of C code
- Usually we use the techniques
  - code inspection (A),
  - unit test (B),
  - integration test (C), and
  - system test (D)
- *How do you plan the quality assurance?*

- We expect about 100 defects

- We expect about 100 defects
- *What is the best sequence of the techniques?* ($ABCD$ or $BACD$?)

- We expect about 100 defects
- *What is the best sequence of the techniques?* ($ABCD$ or $BACD$?)
- *How much effort is spent on each technique?*
  - More on the inspection or the unit test?
  - How much on unit, integration, and system test?

A Quality Assurance Plan

- We expect about 100 defects
- *What is the best sequence of the techniques?* ($ABCD$ or $BACD$?)
- *How much effort is spent on each technique?*
  - More on the inspection or the unit test?
  - How much on unit, integration, and system test?
- *How effective are the techniques on the defects?* (Do inspections detect more defects than unit tests?)

A Quality Assurance Plan

- We expect about 100 defects
- *What is the best sequence of the techniques?* ($ABCD$ or $BACD$?)
- *How much effort is spent on each technique?*
  - More on the inspection or the unit test?
  - How much on unit, integration, and system test?
- *How effective are the techniques on the defects?* (Do inspections detect more defects than unit tests?)
- *How much does it cost to use the techniques?* (How many people? Tools? Defect removal effort?)

A Quality Assurance Plan

- We expect about 100 defects
- *What is the best sequence of the techniques?* ($ABCD$ or $BACD$?)
- *How much effort is spent on each technique?*
  - More on the inspection or the unit test?
  - How much on unit, integration, and system test?
- *How effective are the techniques on the defects?* (Do inspections detect more defects than unit tests?)
- *How much does it cost to use the techniques?* (How many people? Tools? Defect removal effort?)
- *What does it cost me not to detect and remove some of the defects?* (Support, annoyed customers, . . .)

A Quality Assurance Plan

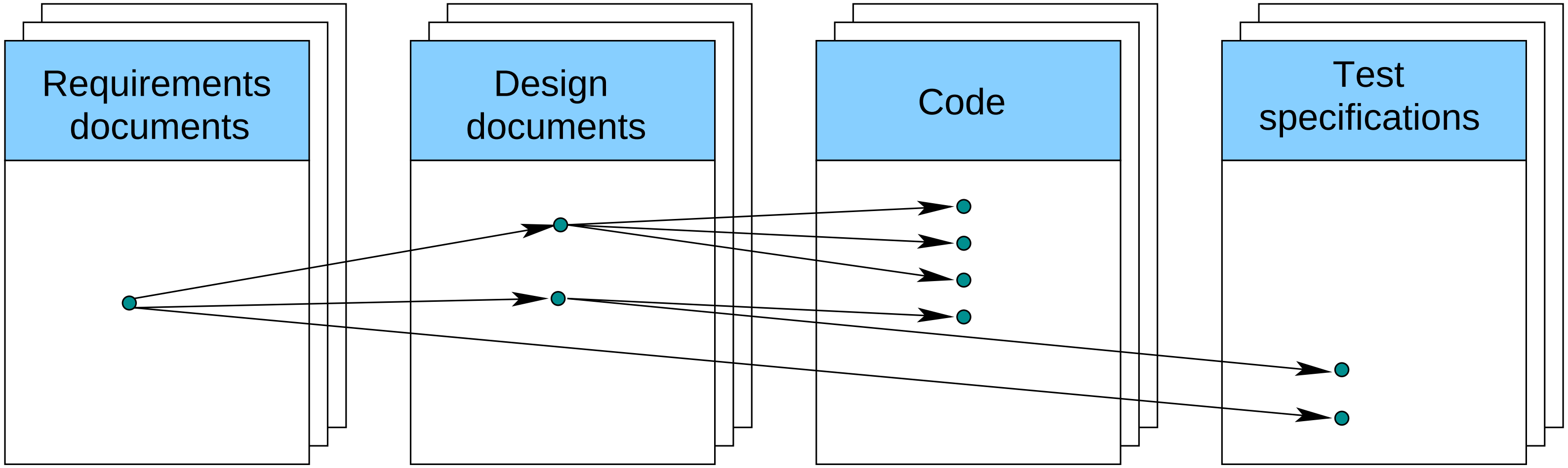
- We expect about 100 defects
- *What is the best sequence of the techniques?* ($ABCD$ or $BACD$?)
- *How much effort is spent on each technique?*
  - More on the inspection or the unit test?
  - How much on unit, integration, and system test?
- *How effective are the techniques on the defects?* (Do inspections detect more defects than unit tests?)
- *How much does it cost to use the techniques?* (How many people? Tools? Defect removal effort?)
- *What does it cost me not to detect and remove some of the defects?* (Support, annoyed customers, . . .)
- *How probable is it that they occur at the customer?*

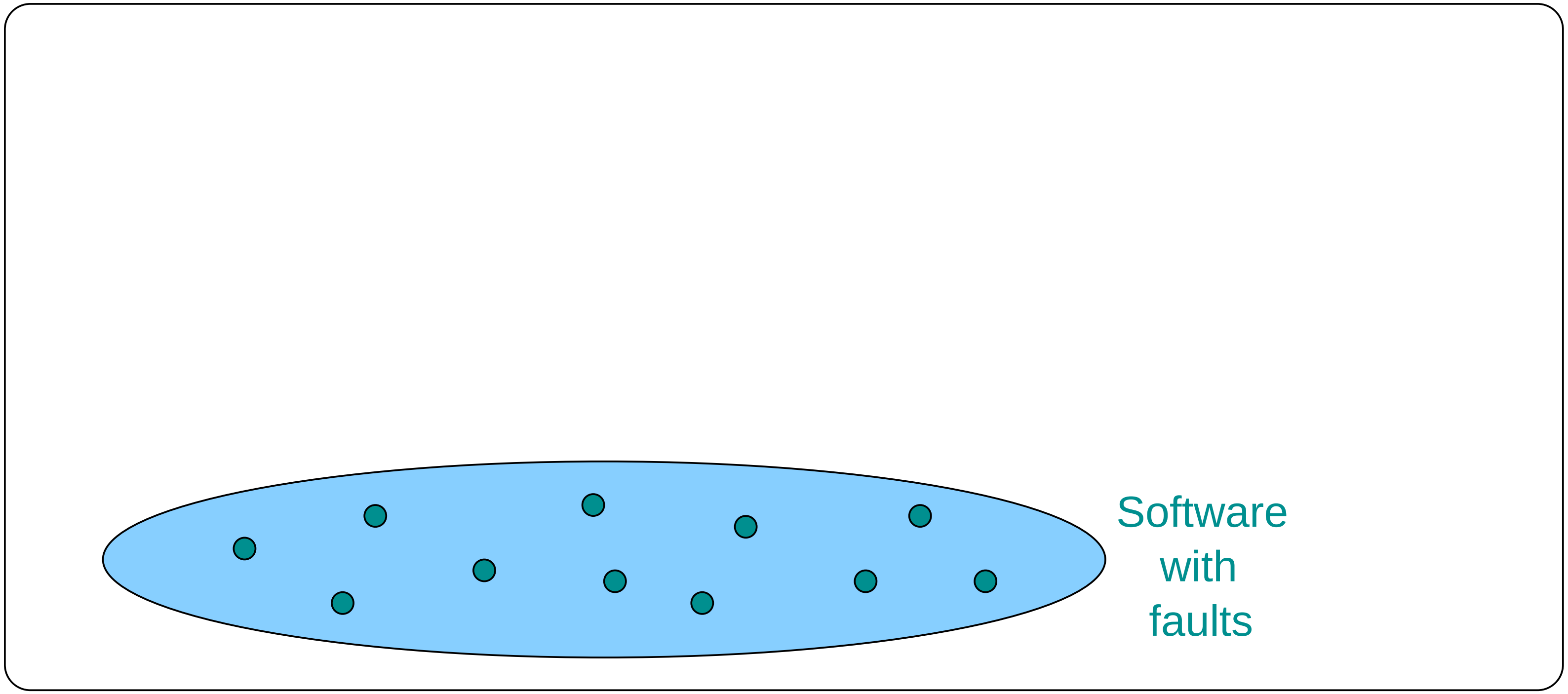
An Analytical Model

- For all types of defect-detection techniques
- Three components:
  - Direct costs d_A
  - Future costs o_A
  - Revenues / saved costs r_A
- Assumptions:
  - Found faults are perfectly removed
  - The effort for a technique can be freely varied
- Not covered:
  - Time to market
  - Net present value

Difficulty and Defect Propagation

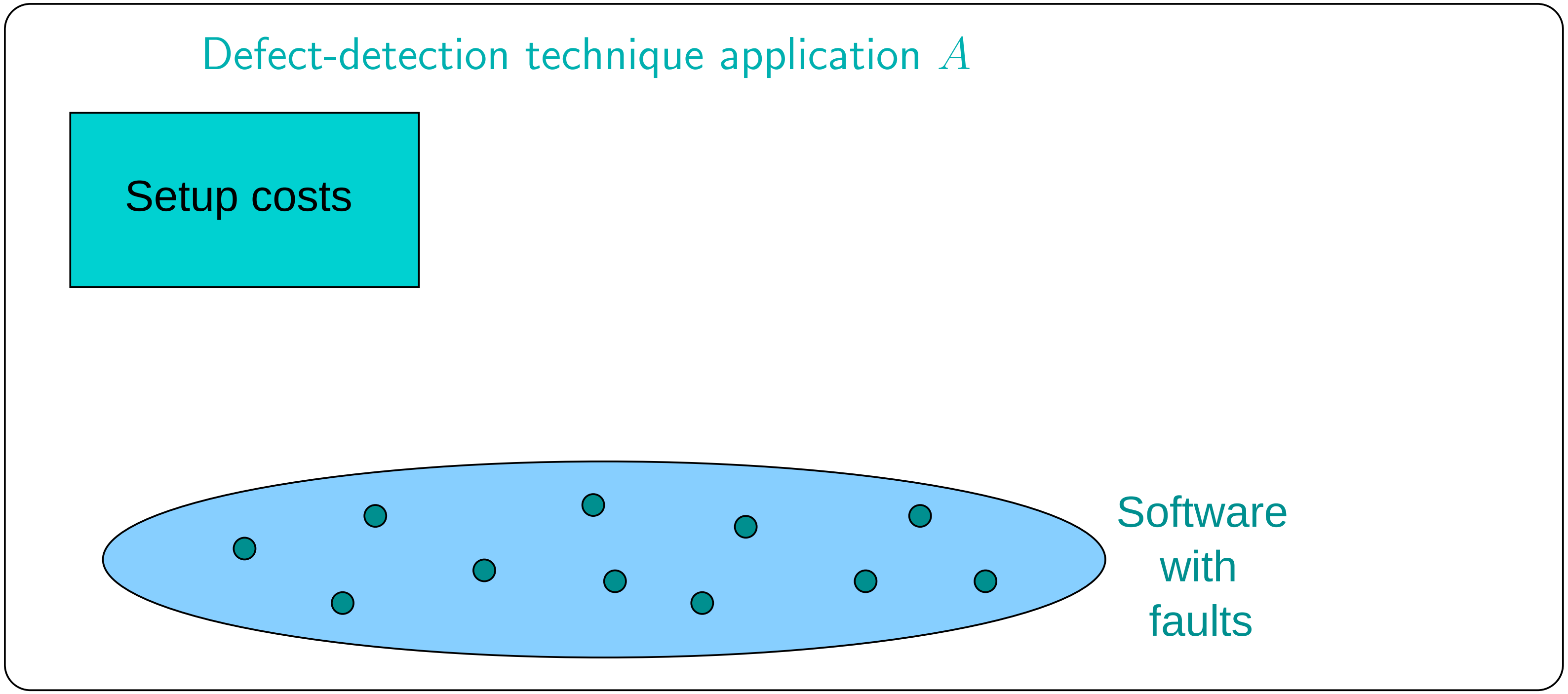
- Difficulty function
  - Adapted from Littewood et al. (2000)
  - Describes the probability that a specific defect is not detected by a specific technique
  - Extended by the dimension *effort*
- Defect propagation





- Costs that can be measured directly from the application of a technique
- Dependent on the effort t

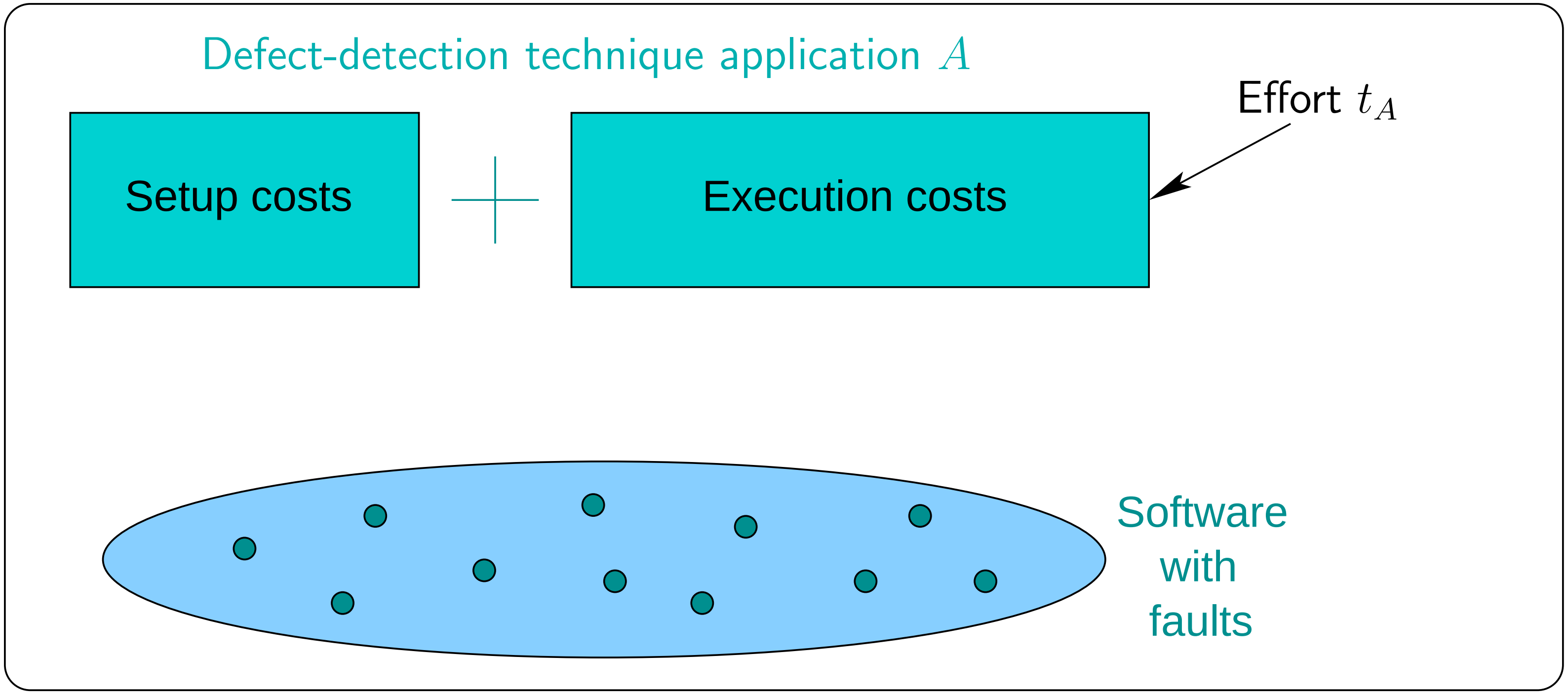
$$d_A = \tag{1}$$



- Costs that can be measured directly from the application of a technique
- Dependent on the effort t

$$d_A = u_A \tag{1}$$

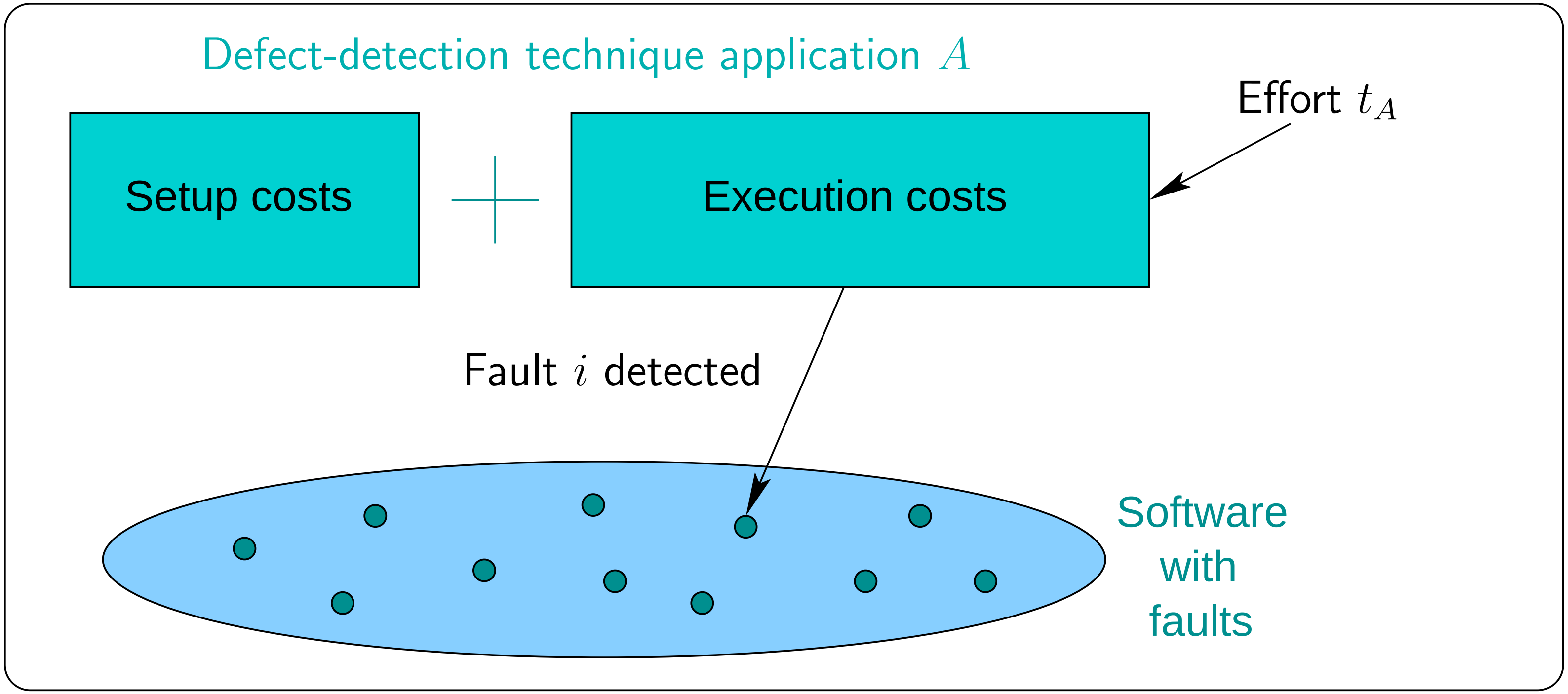
Example: Direct Costs



- Costs that can be measured directly from the application of a technique
- Dependent on the effort t

$$d_A = u_A + e_A(t) \tag{1}$$

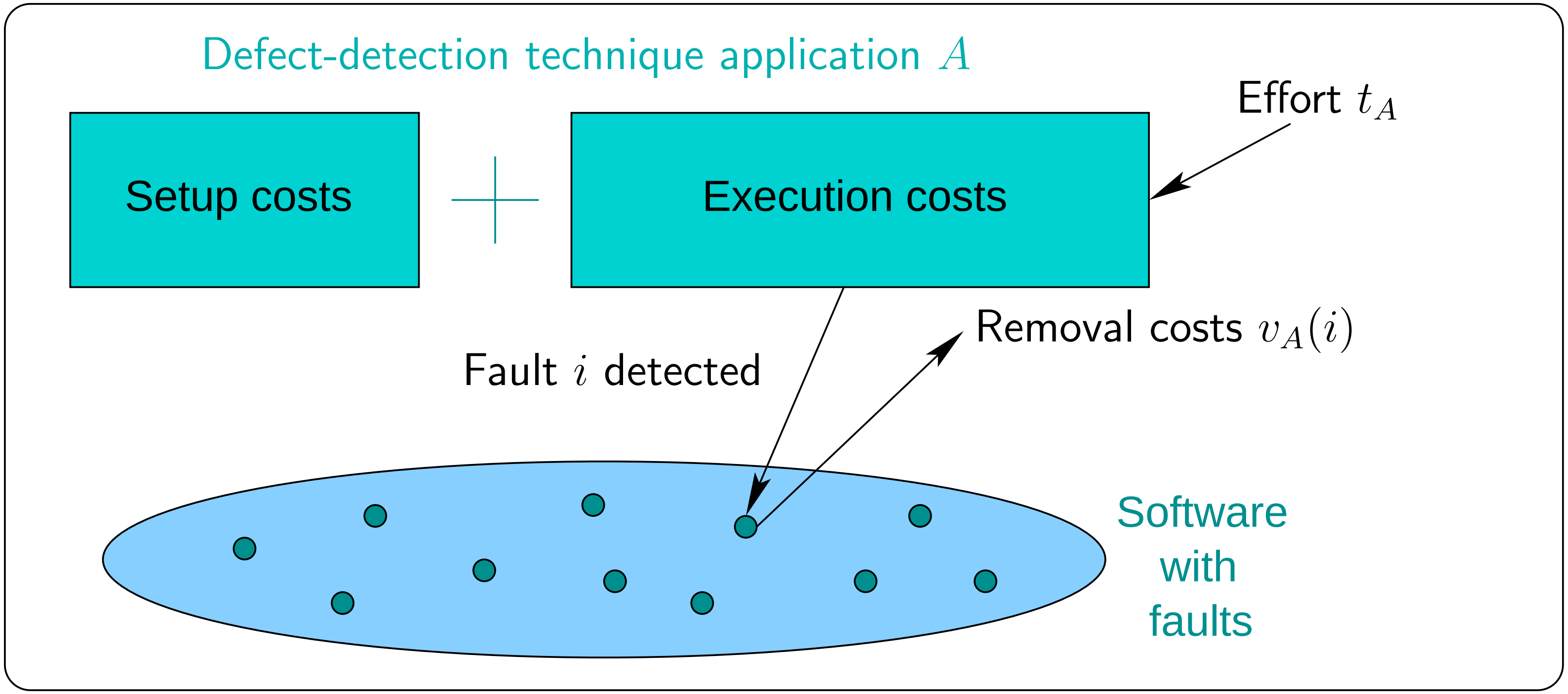
Example: Direct Costs



- Costs that can be measured directly from the application of a technique
- Dependent on the effort t

$$d_A = u_A + e_A(t) \tag{1}$$

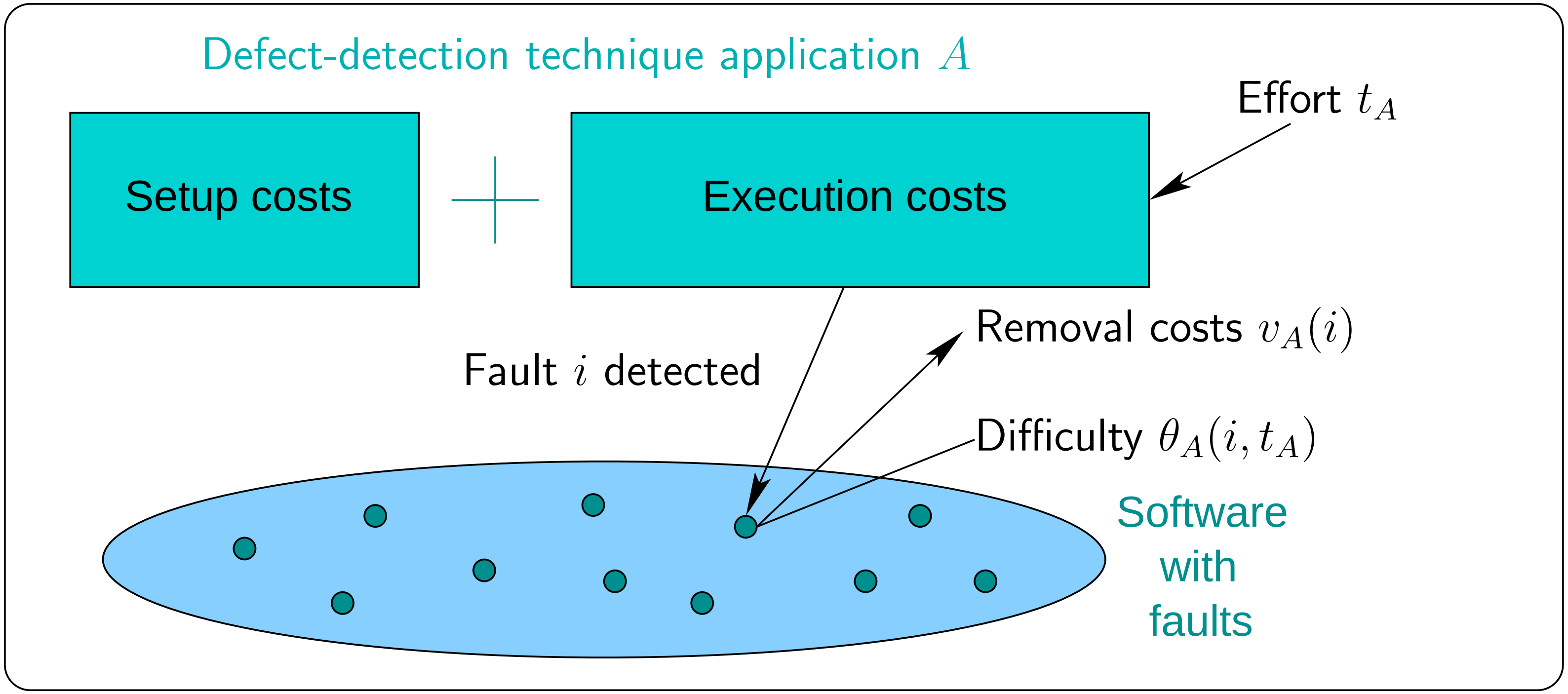
Example: Direct Costs



- Costs that can be measured directly from the application of a technique
- Dependent on the effort t

$$d_A = u_A + e_A(t) + \sum_i v_A(i) \qquad (1)$$

Example: Direct Costs



- Costs that can be measured directly from the application of a technique
- Dependent on the effort t

$$d_A = u_A + e_A(t) + \sum_i (1 - \theta_A(i, t_A))v_A(i) \tag{1}$$

Usage

- Theoretical analysis of defect-detection techniques and improvements
  - Which parts of the model are affected by improvements?
  - How do different techniques differ?
  - What are the most important factors?
- Practical optimisation with simplified model
  - Grouping of defects by defect types
  - Collecting data in a company
  - Predicting for new project based on old data
  - Using optimisation methods

Example Revisited

- Five different effort distributions
- Effort in staff-hours

Dist.	Inspection	Unit test	Int. test	System test	ROI
1	650	300	550	550	0.12
2	0	500	750	750	0.08
3	100	100	100	100	-0.31
4	50	50	50	2000	-0.04
Opt.	800	250	250	200	0.28

Sensitivity Analysis

- Evaluation of the model w.r.t. the influnce of the input factors on the output variance (in our case: ROI)
- Settings: *Factors priorisation* and *factors fixing*

- Evaluation of the model w.r.t. the influnce of the input factors on the output variance (in our case: ROI)
- Settings: *Factors priorisation* and *factors fixing*
- Factors priorisation
  - Which factors would reduce the output variance most if their "true" value would be known?

- Evaluation of the model w.r.t. the influnce of the input factors on the output variance (in our case: ROI)
- Settings: *Factors priorisation* and *factors fixing*
- Factors priorisation
  - Which factors would reduce the output variance most if their “true” value would be known?
  - Which factors are most benefical to investiagte further (in terms of good predictions)?

- Evaluation of the model w.r.t. the influnce of the input factors on the output variance (in our case: ROI)
- Settings: *Factors priorisation* and *factors fixing*
- Factors priorisation
  - Which factors would reduce the output variance most if their “true” value would be known?
  - Which factors are most benefical to investiagte further (in terms of good predictions)?
- Factors fixing
  - Which factors do *not* cause a significant change to the output variance if their value is fixed to one of their distribution?

Basics

- Evaluation of the model w.r.t. the influnce of the input factors on the output variance (in our case: ROI)
- Settings: *Factors priorisation* and *factors fixing*
- Factors priorisation
  - Which factors would reduce the output variance most if their "true" value would be known?
  - Which factors are most benefical to investiagte further (in terms of good predictions)?
- Factors fixing
  - Which factors do *not* cause a significant change to the output variance if their value is fixed to one of their distribution?
  - Which factors can be removed or fixed to simplify the model and data collection?

Approach

- Using the FAST method (based on Monte Carlo sampling)
- As implemented in the tool *SimLab*
- Literature review of the available empirical work on the input factors of the model[1].
- From this, the distributions of the input factors were determined.
- Hence, a sensitivity analysis on the basis of found distributions from industry and research

[1] Available as Technical Report and will be presented at ISESE '06

Results

First order			Total order		
Doc. type	c	0.4698	Doc. type	c	0.8962
Exec. eff.	t	0.1204	Remv. costs f.	v_f	0.4473
Avg. diff.	$\bar{\theta}$	0.0699	Avg. diff.	$\bar{\theta}$	0.4255
Remv. costs f.	v_f	0.0541	Setup costs	u	0.3916
Form function	ϕ	0.0365	Exec. eff.	t	0.3859
Setup costs	u	0.0297	Form function	ϕ	0.2888
Remv. costs i.	v	0.0264	Num. of pred.	ρ	0.2711
Num. of pred.	ρ	0.0256	Remv. costs i.	v	0.2546
Failure prob. f.	π	0.0158	Failure prob. f.	π	0.2068
Tech. seq.	s	0.0083	Tech. seq.	s	0.1825
Labour costs	l	0.0010	Labour costs	l	0.1489

First order			Total order		
Doc. type	c	0.4698	Doc. type	c	0.8962
Exec. eff.	t	0.1204	Remv. costs f.	v_f	0.4473
Avg. diff.	$\bar{\theta}$	0.0699	Avg. diff.	$\bar{\theta}$	0.4255
Remv. costs f.	v_f	0.0541	Setup costs	u	0.3916
Form function	ϕ	0.0365	Exec. eff.	t	0.3859
Setup costs	u	0.0297	Form function	ϕ	0.2888
Remv. costs i.	v	0.0264	Num. of pred.	ρ	0.2711
Num. of pred.	ρ	0.0256	Remv. costs i.	v	0.2546
Failure prob. f.	π	0.0158	Failure prob. f.	π	0.2068
Tech. seq.	s	0.0083	Tech. seq.	s	0.1825
Labour costs	l	0.0010	Labour costs	l	0.1489

- Other analyses (more differentiated factors, practical model) give only slightly different results

Conclusions

Related Work

- Efficiency models
  - Often without considering costs and benefits
  - Only for a specific type of technique
- Economic models
  - Very abstract
  - Data collection not clear

Summary

- Costs and benefits are a central factor in planning software quality assurance
- Models of those costs and benefits can help decision makers
- We propose an analytical model of the quality economics
- More detailed as common economical models
- But more general than specific inspection or reliability models
- Sensitivity analysis shows that the removal costs in the field, the document type, the difficulty, and the execution costs have the strongest effects on the output variation

Future work

- Evaluation of static analysis tools (bug finding tools) using the model with data from a company
- Evaluation of the predictive validity of the model
- Empirical research regarding the most important factors of the model
- Integrated quality modelling and management