Rigorous Approaches to Software Measurement

Sandro Morasca
Università degli Studi dell'Insubria
Via Valleggio 11
I-22100 Como, Italy
sandro.morasca@uninsubria.it
Motivations

Software Engineering strives to be like any other engineering discipline

Goals of (software) production

• high quality product
• within budget constraints
• by a specified deadline

These goals have been achieved in other production processes by using scientific principles

• hypothesis setting (based on observation)
• hypothesis verification (based on empirical studies)
Motivations

So far, software development improvement has been carried on a mostly *ideological* basis.

Software Engineering needs empirical investigations to
- substantiate claims—like in any *scientific* discipline
- enable continuous, quantifiable improvement—like in any *engineering* discipline

Software Engineering is a human-intensive business
- *rigor* and *precision* are indispensable, even more than in other disciplines …
- … but common sense should always rule
Motivations

Measurement is a way of formalizing knowledge
- maybe the most extreme one

Real World

Measurement

Values

• C

60dB

Introduction
- Measure
- Definition
- Process for Defining Measures
- The Future

Rigorous Approaches to Software Measurement

Sandro Morasca
Benevento, April 5, 2004
Motivations

Software measures have often been defined
  • without any rigorous foundations or
  • with rather shaky intuitive justifications

These measures are difficult to interpret and use
  • it is not clear what they really measure

Only recently has the need for rigor in measure definition been recognized as a central issue
  • formal abstractions for artifacts
  • formal description of attributes
  • formal measure definition process
Rigorous Approaches to Software Measurement

Attributes and Measures

- **Usefulness**: Are we *measuring the right attribute*?
  - size, complexity, cohesion, coupling, connectivity, functionality, maintainability …

- **Construct validity**: Are we *measuring the attribute right*?
  - thousands of measures have been defined for software attributes

Many techniques have been defined to improve software based on software attributes, e.g.,
- decrease coupling/increase cohesion

Too many “complexity” measures
Relevant concepts

- **Entity**
- **Attribute** (e.g., size, complexity, maintainability, etc.)
  - internal ones
  - external ones
- **Empirical Relational System**
- **Numerical Relational System**
- **Measure**
- **Representation Condition**
- **Scale**
Five scale types/levels of measurement are usually identified:
- nominal: least informative
  - mode
- ordinal
  - median
- interval
  - mean
- ratio
  - geometric mean
- absolute: most informative
Often, concepts are described in Mathematics by means of sets of axioms

Example: Distance

- Given a set S, distance is a function
  - \( d: S \times S \rightarrow \text{Real} \)
  - such that

Distance.1: Non-negativity
\[ \forall x, y \in S \ (d(x,y) \geq 0) \land (d(x,y) = 0 \iff x = y) \]

Distance.2: Symmetry
\[ \forall x, y \in S (d(x,y) = d(y,x)) \]

Distance.3: Triangular Inequality
\[ \forall x, y, z \in S \ (d(x,y) + d(y,z) \geq d(x,z)) \]
A number of axiomatic approaches have been proposed in the literature for ratio scale complexity or coupling measures

- Prather
- Zuse
- Weyuker
- Fenton (coupling)
- Lakshamanian et al.
- Briand, Morasca, Basili
- Poels and Dedene
The measurement of software has specific difficulties, because software

- is a human-intensive business
- has immaterial artifacts (e.g., specifications, code)
- has elusive attributes (e.g., complexity, coupling, maintainability)
- has elusive relationships between attributes (e.g., complexity \(\Rightarrow\) maintainability)
- is a new discipline

An organized approach is required, where measures are derived from

- corporate goals
- empirical assumptions
- environmental characteristics
Define Measurement Goals

- From corporate goals
  - “improve product quality”
- To tactical goals
  - “increase the number of faults found prior to delivery”
- To measurement goals
  - “study the relationship between coupling and fault-proneness”

Each kind of goal needs to be prioritized
- higher payoff to the corporation, which leads to
  - adequate support to the measurement program
  - available information about the studied development environment (e.g., weaknesses, problems)
  - resources available to carry out the empirical study
GQM template can be used to set measurement goals

- **Object of study**: products, processes, resources, ...
- **Purpose**: characterization, evaluation, prediction, improvement, ...
- **Quality focus**: cost, correctness, defect removal, changes, reliability, ...
- **Viewpoint**: user, customer, manager, developer, corporation, ...
- **Environment**: team, project, product, ...

These five goal dimensions have a direct impact on the remaining steps of the measure definition approach and the data collection program.
Rigorous Approaches to Software Measurement

High-level Structure of a Process for Defining Measures

- **Measure Definition Process**
  - Setting of the Empirical Study
  - Definition of Measures for Independent Attributes
  - Definition of Measures for Dependent Attributes
  - Hypothesis Refinement and Verification
  - Literature

- **entities + independent attributes**
- **properties + abstractions + measures**
- **measures**
- **knowledge about the environment**
- **project teams**
- **experience factory**
- **environment-specific (properties + models + data + lessons learned + confirmed hypotheses + disconfirmed hypotheses)**
- **corporate goals**
- **experience**
- **factory**
- **properties + properties**
- **measures**
- **measures**
- **dependent measures**
- **independent measures**
- **empirical hypotheses**
- **independent attributes**
- **dependent attributes**

- **Introduction**
- **Measure Definition**
- **Process for Defining Measures**
- **The Future**
The Future

- More rigorous definition of attributes
- More rigorous definition of measures
- Better organized processes for the definition of measures
- Tool support and automation