Software Engineering II
Theory

Christian Rossi

Academic Year 2023-2024

Abstract

The objective of the course is to teach the principals methods and processes of software engineering needed to develop complex and qualitative software. The course covers the following arguments:

- Software process and its organization.

- Modelling languages.

- Requirements analysis and definition.

- Software development methods and tools.

- Approaches for verify and validate the software.

Contents

Chapter 1

Introduction

1.1 Definition

The field of software engineering aims to find answers to the many problems that software development projects are likely to meet when constructing large software systems. Such systems are complex because of their sheer size, because they are developed by a team involving people from different disciplines, and because they will be modified regularly to meet changing requirements, both during development and after installation.

Definition. Software engineering is a methodological and managerial discipline concerning the systematic production and maintenance of software products that are developed and maintained within anticipated and controlled time and cost limits.

The programmer develops a complete software and works on known specifications individually. Instead, the software engineer identifies requirements and develop specifications, designs components that will be combined with others and works in a team. The main skills of a software engineer are: technical, managerial, cognitive, organizational.

1.2 History

Initially, the software was considered as an art. The computers were used for computing to solve mathematical problems and the designers were also the users. The first programs were created with low-level languages and had high resources constraints.

When the request for new custom software exploded the art became a craft: the developer started to create programs also for the people with

new high-level languages. At the end of this period there were a "software crisis" due to increasing software complexity and lack of effective software development techniques.

To solve this problem in 1968 was defined the term *software engineering* in a NATO conference. The main focuses of this conference was on:

- Development of software and standards.

- Planning and management.

- Automation.

- Modularization.

- Quality verification.

1.3 The process and product

The developing of a software program needs a process. Both software and processes have a quality and the software engineer needs to reach the optimal quality because the process modifies the final output.

The software is different from traditional types of products because it is:

1. Intangible (difficult to describe and evaluate).

2. Malleable.

3. Human intensive (does not involve any trivial manufacturing process).

The quality of the software is influenced by the following variables: development technology, process quality, people quality, cost, time and schedule. The software quality attribute are:

- Correctness: software is correct if it satisfies the specifications.

- Reliability: probability of absence of failures for a certain time period.

- Robustness: software behaves reasonably even in unforeseen circumstances.

- Performance: efficient use of resources.

- Usability: expected users find the system easy to use.

- Maintainability.

- Reusability: similar to maintainability but applies to components.

- Portability: adaption to different target environments. Interoperability: coexist and cooperate with other applications.

The process quality attribute are:

- Productivity.

- Unity of effort (person month).

- Delivered item (lines of code and function points).

- Timeliness: ability to respond to change requests in a timely fashion.

1.4   Development process

Initially there were no reference model, so it was simple code&fix. As a reaction to the software crisis mentioned before it became necessary to have a model. The first complete model was the "waterfall". The key requirements of this model are:

1. Identify phases and activities.

2. Force linear progression from a phase to the next (without returns).

3. Standardize outputs from each phase.

4. Software is considered like manufacturing.

After this model many other flexible processes were proposed: iterative models, agile movement and DevOps. The main phases shown in the image are:

1. Feasibility study and project estimation: determines wheather the project should be started, the possible alternatives and needed resources. This phase produces a *Feasibility Study Document* which contains: preliminary problem description, scenarios describing possible solutions, cost and schedule for the different alternatives.

2. Requirement analysis and specification: analyze the domain in which the application takes place, identify requirements and derive specification for the software. This phase produces *Requirement Analysis and Specification Document*.
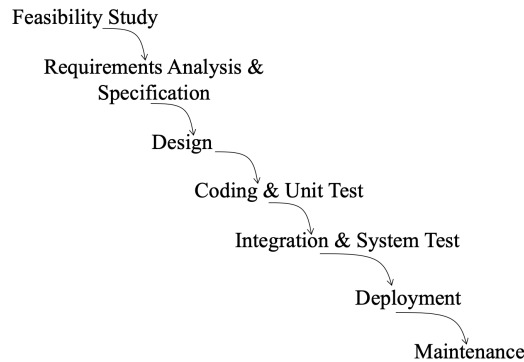
Figure 1.1: Waterfall process model

3. Design: defines the software architecture (components, relation and interactions among components). The goal is to support concurrent development and separate responsibilities. It produces the *Design Document*.

4. Coding and unit test: each module is implemented and tested. Inspection can be used as an additional quality assurance approach. Programs include their documentation.

5. Integration and system test: the modules are integrated into systems and integrated systems are tested. This phase and the previous may be integrated in an incremental implementation scheme.

6. Deployment.

7. Maintenance: the maintenance can be:

   - Corrective: deals with the repair of faults or defects found.
   - Adaptive: consist of adapting software to changes in the environment.
   - Perfective: deals with accommodating to new or changed user requirements.
   - Preventive: concerns activities aimed at increasing the system's maintainability.

The principal problems with software evolution are:

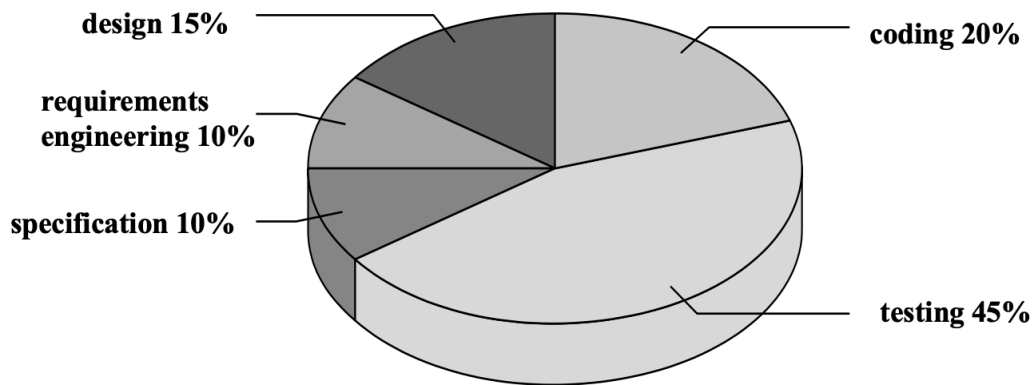- It is almost never anticipated and planned.

Figure 1.2: Effort in each phase

- Software is very easy to change (changes applied directly to the code that causes inconsistent state of project documents).

To face properly the evolution we need a good engineering practice that consist in two main steps: modify the design and then change the implementation and apply changes consistently in all documents. In fact, one of the main goal of software engineering is to create software that must be designed to accommodate future changes reliably and cheaply.

Waterfall model is a black-box system because the company that requests the software makes requirements and doesn't interact during the development phase. If we need more transparency with the customer we need to use a different development model (that allows the customer to give feedback regularly). With every interaction with the customer is possible to check two main things:

- Validation: check if the product follows the customer's requests.

- Verification: check if the product works in the right way.

The idea of flexible process is to adapt to changes, in particular the requirements and specification. The idea is to have incremental processes and be able to get feedback on increments. They exists in many forms, for example: SCRUM, extreme programming, incremental releases and rapid prototyping, DevOps, ...

Chapter 2

Requirements engineering

2.1 Definition

The primary measure of success of a software system is the degree to which it meets the purpose for which it was intended.

Definition. Software systems *requirements engineering* is the process of discovering that purpose, by identifying stakeholders and their needs, and documenting these in a form that is amenable to analysis, communication, and subsequent implementation.

The important issues of this phase are: identify stakeholders, identify their needs, produce documentation and analyse, communicate and implement requirements. Another possible definition is the following.

Definition. Requirements engineering is the branch of software engineering concerned with the

- real-world goals for,

- function of, and

- constraints on

software systems. It is also concerned with the relationship of these factors to precise specifications of software engineering behaviour, and to their evolution over time and across software families.

2.2 Importance and difficulties

The requirements given from the customer can be classified in three main types:

Figure 2.1: Some relevant QoS characteristics

- Functional: describes the interaction between the system and its environment independent from implementation. They are the main goals that the software has to fulfill.

- Nonfunctional: user visible aspects of the system not directly related to functional behaviour.

- Constraints: imposed by the client or the environment in which the system operates.

The nonfunctional requirements are constraints on how functionality has to be provided to the end user. They are independent of application domain but the application domain determines: their relevance and their prioritization. They are also called Quality of Service attribute.

2.3 The requirement engineering process

Poor requirements are ubiquitous. Requirement engineering is also hard and critical because a problem with the initial phases can be up to two hundred times costly in the final phase. Requirement engineering is so complex

Relative Cost of Error Correction

200

The cost of correcting
an error depends on
the number of
subsequent decisions
that are based on it

50

Errors made in
understanding
requirements have
the potential for
greatest cost,
because many other
design decisions
depend on them

1   5   10   20

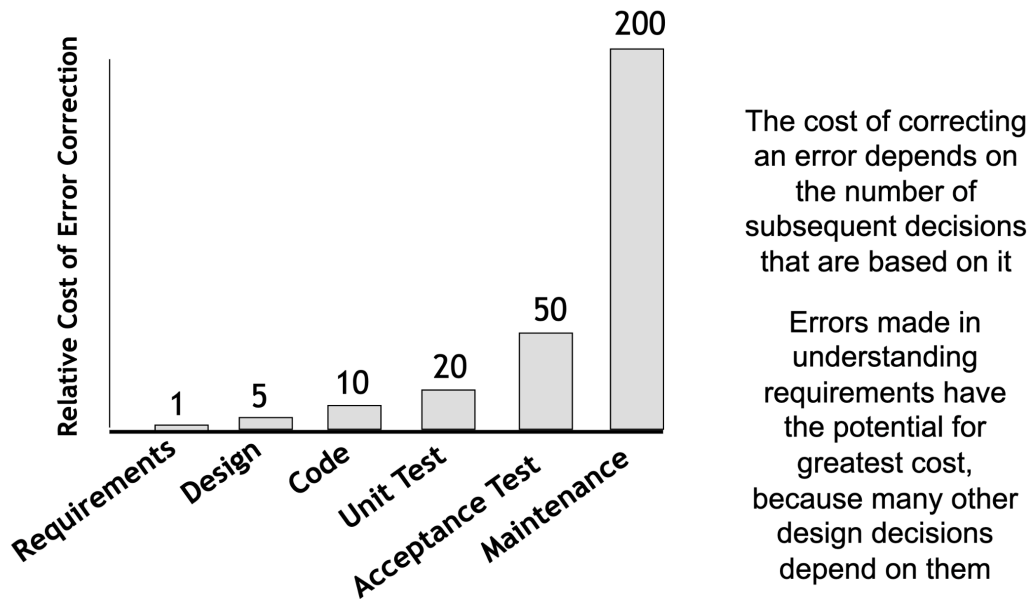Requirements   Design   Code   Unit Test   Acceptance Test   Maintenance

Figure 2.2: Cost of late correction [Boehm, 1981]

because of: composite systems. more than one system, multiple abstraction
levels, multiple concerns and multiple stakeholders with different background.
The requirement engineers needs to:

- Eliciting information (project objectives, context and scope; domain
  scope and requirements).

- Modelling and analysis (goals, objects, use cases and scenarios).

- Communicating requirements (analysis feedback, RASD document, sys-
  tem prototypes).

- Negotiating and agreeing requirements (handling conflicts and risks;
  helping in requirement selection and prioritization).

- Managing and evolving requirements (managing requirements during
  development: backward and forward traceability; managing require-
  ments changes and their impacts).

2.4   Understanding world-machine relationship

The machine indicates the portion of system to be developed, while world
indicates the portion of the real-world affected by the machine. The purpose
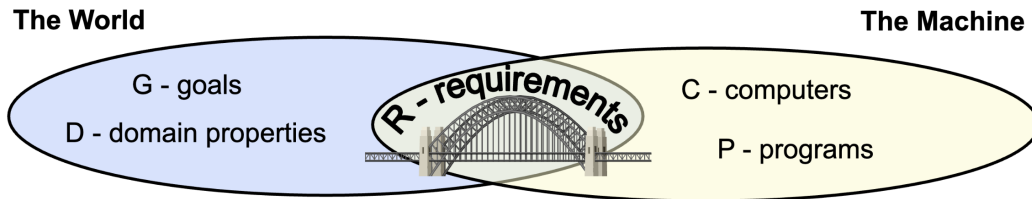of the machine is always in the world.

Figure 2.3: Goals, domain assumptions, and requirements

Requirements engineering is concerned with phenomena occurring in the world as opposed to phenomena occurring inside the machine. We can say that requirements models are models of the world.

Some world phenomena are shared with the machine. This type of phenomena can be:

- Controlled by the world and observed by the machine.

- Controlled by the machine and observed by the world.

Goals are prescriptive assertions formulated in terms of world phenomena (not necessarily shared). Domains properties/assumptions are descriptive assertions assumed to hold in the world. Requirements are prescriptive assertions formulated in terms of shared phenomena.

The requirements R are complete if:

- R ensures satisfaction of the goals G in the context of the domain properties D, this means that $R \wedge D \models G$.

- G adequately capture all the stakeholders' needs.

- D represent valid properties/assumptions about the world.

2.5 Elicitation of requirements

The complexity in requirement engineering can be coped with:

- Adopting different approaches and strategies and combining the results reached with all of them.

- Being as close as possible to stakeholders.

- Letting stakeholders describing their viewpoints.

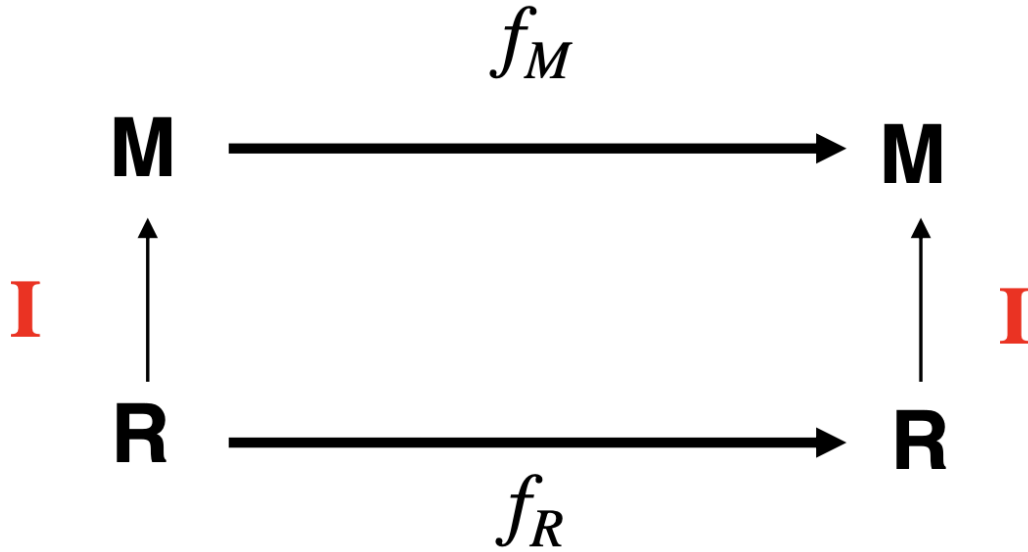The scenario can be generalized in term of:

Figure 2.4: Relationship between model M and reality R

- Participation actors.

- Describe the entry condition.

- Describe the flow of events.

- Describe the exit condition.

- Describe exceptions.

- Describe special requirements.

2.6 Modeling requirements

Definition. A *model* is a representation in a certain medium of something in the same or another medium. The model captures the important aspects of the thing being modeled and simplifies or omits the rest.

The reality R is composed by: real things, people, processes and relationship. The model M is an abstraction of things, people, processes and relationship between these abstraction.

The reality needs to be interpreted (I) with a mapping function. To have a good model the relationships that are valid in the reality R need to be valid also in the model M. The software models are used for:

- Capture and precisely state requirements and domain knowledge.

- Think about the design of a software system Generate usable work products.

- Give a simplified view of complex systems Evaluate and simulate a complex system.

- Generate potential configurations of systems.

The principal modeling issues are coherence (different views of the system must be coherent) and variation in interpretation and ambiguity (define where different interpretation of the model are acceptable).

In requirements engineering we should only model:

- The objects and people that are of interest for the given problem.

- The relevant phenomena.

- The goals, requirements, and domain assumptions.

The tool that we can use for modeling are:

- Natural language (English, Italian, . . .):

  - Pros: simplicity of use.
  - Cons: high level of ambiguity, it is easy to forget to include relevant information.

- Formal language (FOL, Alloy, Z, . . .):

  - Pros: possibility to use tool to support analysis and validation, the approach forces the user in specifying all relevant details.
  - Cons: you need to be expert in the use of the language.

- Semi-formal language (UML):

  - Pros: simpler than a formal language, impose some kind of structure in the models.
  - Cons: not amenable for automated analysis, some level of ambiguity.

- Mixed approach: use a semi-formal language for the basics. Comment and complement the semi-formal models with explanatory informal text. use a formal language for the most critical parts.
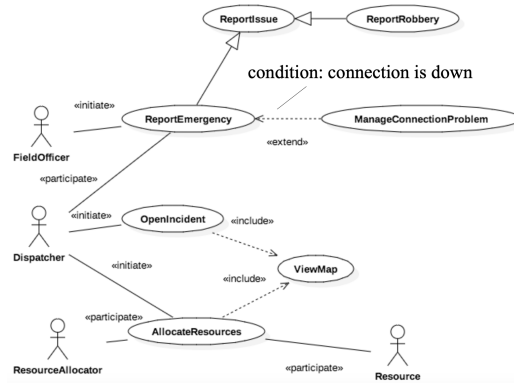
Figure 2.5: Use case model example

2.7 Use cases and requirements

The main steps when formulating use cases are:

1. Name the use case

2. Find the actors: generalize the concrete names to participating actors.

3. Concentrate on the flows of events, entry and exit condition using natural language.

4. Focus on exceptional cases and special requirements.

Each use case may lead to one or more requirements.

A use case is a flow of events in the system, including interaction with actors. The use cases are initialized by an actor and has a termination condition.

Definition. The *use case model* is the set of all use cases specifying the complete functionality of the system.

Definition. A *use case association* is a relationship between use cases. The principal types of use case association are:

- Include(a use case uses another use case).

- Extends (a use case extends another use case).

- Generalization (an abstract use case has several different specializations).

Example	Grammatical construct	UML Component
"Monopoly"	Concrete Person/Thing	Object
"toy"	Noun	Class
"board game"	Noun	Class
"6 years old"	Adjective	Attribute
"enters"	Verb	Operation/Association
"depends on ..."	Intransitive Verb	Association
"is a", "either ... or", "kind of ..."	Classifying Verb	Inheritance
"Has a", "consists of"	Possessive Verb	Aggregation
"must be", "less than ..."	Modal Verb	Constraint

Figure 2.6: Abbott Textual Analysis example

2.8 Requirements-level class diagrams

The requirements-level class diagrams are conceptual models for the application domain. They may model objects that will not be represented in the software-to-be. Usually, the do not attach operations to objects: it's best to postpone this kind of decisions until software design.

To find objects and classes we need to:

- Analyze any description of the problem and application domain you may have.

- Analyze your scenarios and use cases descriptions.

Finding objects is the central piece in object modeling. A possible tool to use in the analysis is the Abbott Textual Analysis also called noun-verb analysis: nouns are good candidates for classes and verbs are good candidates for associations and operations.

2.9 Dynamic modeling

The purpose of the dynamic modeling is to supply methods to model interactions, behaviours of participants and workflow. This can be done with: sequence diagram, state machine diagram and activity diagram. Some objects can be found whilst completing those diagrams.

The sequence diagram is created following the flow of events in the use case diagram. A sequence diagram is a graphical description of objects participating in a use case scenario using a Directed Acyclic Graph notation. The principal rules to create a sequence diagrams are:

- An event always has a sender and a receiver.

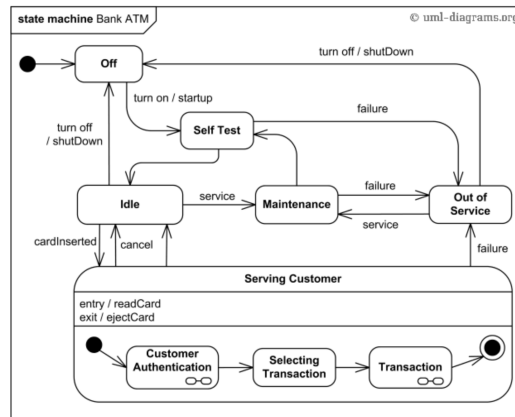- The representation of the event is sometimes called a message.

Figure 2.7: Example of a state diagram

- Find sender and receiver for each event.

For a good dynamic modeling we ha to construct a model only for classes with significant dynamic behaviour and consider only relevant attributes. We have also to look ate the granularity of the application when deciding on actions and activities and reduce notational clutter.

Chapter 3

Alloy

3.1   Definition

Alloy is a formal notation for specifying models of systems and software. It looks like a declarative object oriented language, but also has a mathematical foundation.

Alloy comes with a supporting tool to simulate specifications and performs property verification.

Alloy has been created to offer an expressive power similar to the Z language as well as the strong automated analysis of the SMW model checker.

Alloy can be used in requirement engineering to formally describe the domain and its properties, or operations that the machine has to provide. In software design Alloy can be used to formally model components and their interactions.

3.2   Introduction

Alloy is a mixture of first order logic and relational calculus. The main elements of this formal language are:

- Carefully chosen subset of relational algebra (uniform model for individuals, sets and relations; no high-order relations).

- Almost no arithmetic.

- Modules and hierarchies.

- Suitable for small, explanatory specification.

- Powerful and fast analysis tool.

3.3   Syntax

Alloy shows bounded snapshots of the world that satisfy the specification. Allow does bounded exhaustive search for counterexample to a claimed property using SAT.

Definition. Atoms are Alloy's primitive entities (indivisible, immutable and uninterpreted).

Definition. Relations associate atoms with one another (set of tuples; tuples are sequence of atoms).

The relations in Alloy are typed. The relation type is determined by the declaration of the relation. The basic Alloy relations type are: *none* (empty set), *univ* (universal set) and *iden* (identity relation). The logic operators in Alloy are:

- Union ∪.

- Intersection &.

- Difference −.

- Subset *in*.

- Equality =.

- Cross product → (similar to a natural join).

- Dot join . or [ ] (the last element of the first relation joins on the corresponding first elements of the second relation, and then removes the element from the relation).

The possible binary closures on the relations are:

- Transpose ∼ (inverts the order of the elements in the relation).

- Transitive ∧ ($^\wedge r = r + r.r + r.r.r + \dots$).

- Reflexive transitive * ($^* r = iden +^\wedge r$)

The possible restrictions are:

- Domain restriction <:, that restricts the elements on the left side to the set on the right side.

- Range restriction :>, that is same as before but the relations are inverted.

- Override ++, that removes the tuples on the left that are in the right relations and adds all the remaining relations of the right relation.

The Alloy Boolean operators are the following: negation (! or *not*), conjunction (& or *and*), disjunction (|| or *or*), implication ($\implies$ or *implies*), alternative (, or *else*) and bi-implication ($\iff$ or *iff*). The alloy logic quantifiers are:

- *all*: holds for every element.

- *some*: holds for at least one element.

- *no*: holds for no elements.

- *lone*: holds for at most one element.

- *one*: holds for exactly one element.

To define a relation with a singleton we can use the following declaration:

$$x \ : \ m \, e$$

Where x is the name of the relation, m is the multiplicity of the element (that can be: *set*, *one*, *lone* or *some*) and e is the name of the element in the relation. If the relation is composed by couple the declaration became like this:

$$r \ : \ A \, m \ \to \ n \, B$$

Where r is the name of the relation, A and B are the name of the elements with multiplicity m and n respectively.

  let is used to define a formula or expression that can be reused. Other useful operators are: $\#r$ (that define the number of tuples in r), $0, 1, \ldots$ (integer used to define the value of some variables or constants), + (plus), − (minus), all the comparison operators ($<, <=, =, =>, >$). There is also the operator *sum* that adds all the elements in the selected tuple.

Chapter 4

Requirement analysis and specification

4.1 Structure of a RASD document

The RASD has the following purposes:

- Communicates an understanding of the requirements (application domain and the system to be developed).

- Contractual (may be legally binding).

- Baseline for project planning and estimation (size, cost and schedule).

- Baseline for software evaluation (support system testing, verification and validation activities; should contain enough information to verify weather the delivered system meets requirements).

- Baseline for change control (requirements change, software evolves).

The RASD is used by:

- Customers and users (interested in high level description of functionalities).

- System analyst and requirement analyst (specification of other system to inter-relate).

- Developers and programmers (implementation).

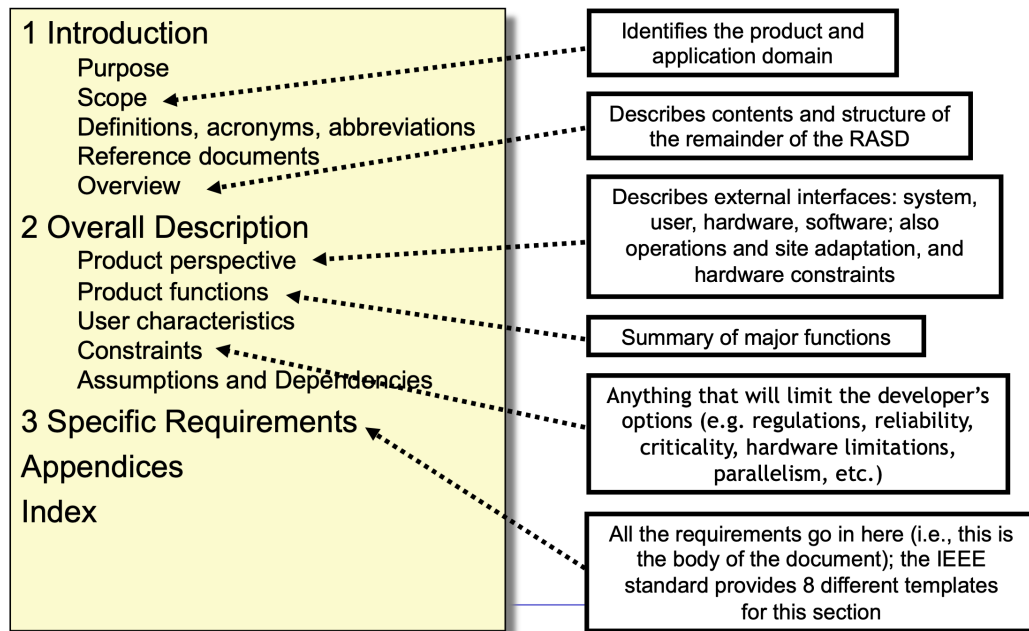- testers (check if requirements are met).

Figure 4.1: IEEE standard for RASD

- Project managers (control the development process).

The target qualities for a good RASD are:

- Completeness:

    - w.r.t. goals: the requirements are sufficient to satisfy the goals under given domain assumptions:

    $$Req \wedge Dom \models Goals$$

    This means that all Goals have been correctly identified, including all relevant quality goals and that the Dom represent valid assumptions; incidental and malicious behaviours have been anticipated.

    - w.r.t. inputs: the required software behaviour is specified for all possible inputs

    - Structural completeness: no TBDs

- Pertinence:

    - Each requirement or domain assumption is needed for the satisfaction of some goal.

- – Each goal is truly needed by the stakeholders.

  – The RASD does not contain items that are unrelated to the definition of requirements.

- Consistency: no contradiction in formulation of goals, requirements, and assumptions.

- Unambiguity:

  – Unambiguous vocabulary: every term is defined and used consistently.

  – Unambiguous assertions: goals, requirements and assumptions must be stated clearly in a way that precludes different interpretations.

  – Verifiability: a process exists to test satisfaction of each requirement.

  – Unambiguous responsibilities: the split of responsibilities between the software-to-be and its environment must be clearly indicated.

- Feasibility: the goals and requirements must be realisable within the assigned budget and schedules.

- Comprehensibility: must be comprehensible by all in the target audience.

- Good structuring: every item must be defined before it is used.

- Modifiability: must be easy to adapt, extend or contract through local modifications and impact of modifying an item should be easy to assess

- Traceability:

  – Must indicate sources of goals, requirements and assumptions.

  – Must link requirements and assumptions to underlying goals.

  – Facilitates referencing of requirements in future documentation.

Chapter 5

Software design

5.1 Software architecture

Definition. The *architecture* of a software system defines that system in terms of computational components ad interactions among those components.

Definition. The *software architecture* of a system is the structure of structures of the system, which comprise software elements, the externally visible properties of those elements, and the relationships among them.

In the second definition the most important issues raised are: multiple system structure and externally visible properties of components. This definition does not include: the process, rules and guidelines and architectural styles.

Software architecture is really important because:

- It is the vehicle for stakeholders communication.

- it manifests the earliest set of design decisions (constraint on implementation, dictates organizational structure and inhibits or enables quality attributes).

- it is a transferable abstraction of a system (product lines share a common architecture, allows for template-based development and basis for training).

5.2 Architecture life cycle

The iteration of this life cycle are mainly on functional requirements, only a few stakeholder are involved and no balancing of functional and quality
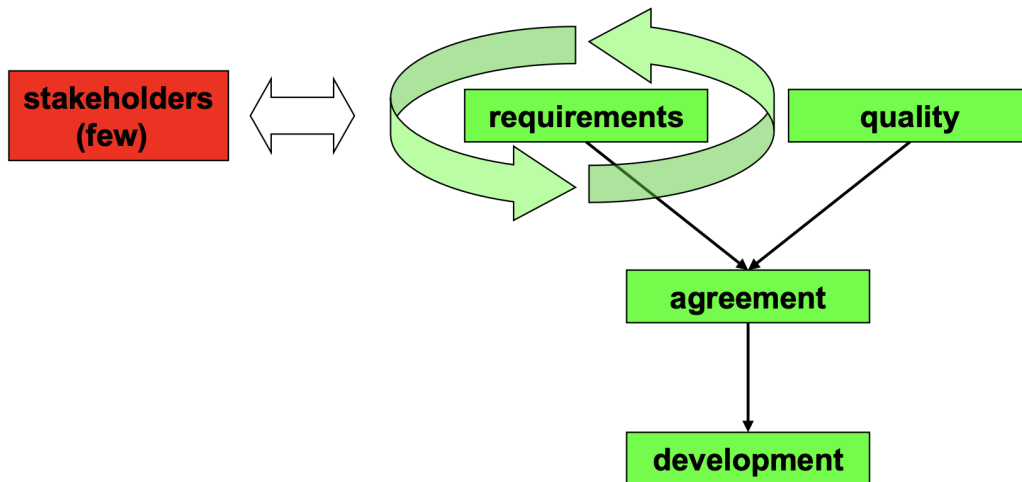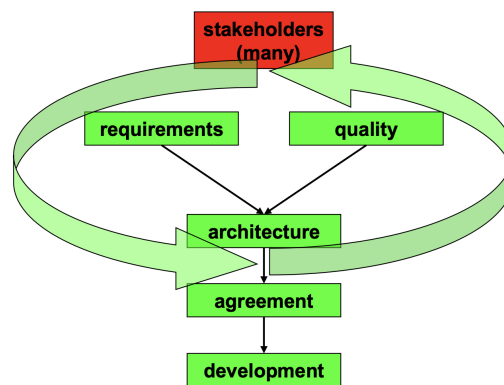
Figure 5.1: Pre-architecture life cycle

Figure 5.2: Life cycle considering architecture

requirements. If we add importance to architecture we have the following characteristics:

- Iteration on both functional and quality requirements.

- Many stakeholder involved.

- Balancing of functional and quality requirements.

In general, a software architecture is intrinsic in the software.

When changes are applied to the code we have an architectural degradation: the changes in the software system are not reflected in the corresponding architectural description. This happens because of: lack of a documented ar-
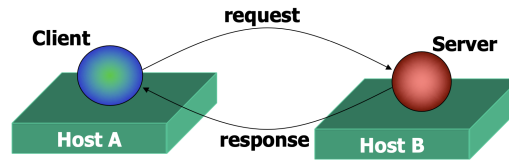
Figure 5.3: Client-server architecture general schema

chitecture, inadequate processes and supporting tools, developer sloppiness
and perception of short deadlines. The degradation can be classified in:

- Drift: the change is not included nor implied by the architecture but
  it does not necessarily imply its outright violation.

- Erosion: the change violates the described architecture.

5.3   Software architecture quality and style

The notion of quality is central in software architecting: a software archi-
tecture is devised to gain insight in the qualities of a system at the earli-
est possible stage. Some are observable via execution (performance, secu-
rity, availability, functionality and usability) and some aren't (modifiability,
portability, reusability, integrability and testability).

Definition. An *architectural style* determines the vocabulary of components
and connectors that can be used in instances of that style, together with a set
of constraints on how they can be combined. These can include topological
constraints on architectural descriptions. Other constraints—say, having to
do with execution semantics—might also be part of the style definition.

The architectural styles can be:

- Layered (operative system): the system is organized trough abstraction
  levels, as a hierarchy of abstract machines. Hierarchy is given by the
  use relation.

- Client-server (distributed applications): they are different processes
  with well-defined interface (accessible only trough interfaces and they
  can be defined by a set of hardware/software modules).

- Three-tier architecture: the mid level deals with the application logic.
  the main advantage is the decoupling of logic and data, logic and pre-
  sentation.

- Event-based systems: the components in this architecture can register to send/receive events that are always broadcasted to all registered components (they don't know who is the sender nor the receiver). Those systems are often called publish-subscribe. This type of architecture allow easy addition/deletion of components and is increasingly used in modern integration strategies. The main problems are about scalability and ordering of events. The main characteristics are:

  - Asynchrony (send and forget).
  - Reactive (computation driven by receipt of message).
  - Location/identity abstraction (destination determined by receiver, not sender).
  - Loose coupling (senders/receivers added without reconfiguration one-to-many, many-to-one, many-to-many).

- Service-oriented architecture (SOA), where the main pros are:

  - Enables reuse of registries and providers across organization boundaries.
  - Relies on run-time discovery and dynamic binding.
  - Service offered through a well-defined Application Programming Interface.

  But it has also many problems like: dynamic orchestration of discovered services in sot trivial and SOAP protocol is considered too heavy.

- Micro-service architecture: monolithic systems are decomposed into small specialized services and deal with a single bounded context in the target domain. The micro-service architectural style is an approach to developing a single application as a suite of small services, each running in its own process and communicating with lightweight mechanisms, often an HTTP resource API. This type of architecture solves the following problem with the monolithic systems: issues with frequent deployment, overloaded containers, obstacle to scaling development, may lead to availability issues and requires long-term commitment to a technology stack. With this architecture it's possible to use different technology for each part of the complete system. This architecture is resilient because every module (that can use different technology and developing teams) is detached partially from the others and also highly scalable. The main problems of this method are that we deal with distributed systems and that the decomposition of monolithic systems are complex and needs collaboration.
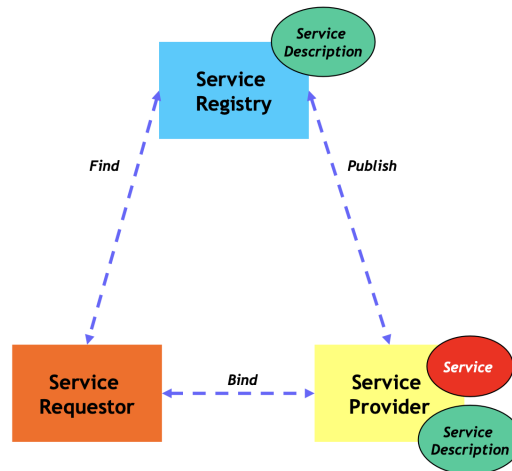
25

Figure 5.4: Service-oriented architecture schema

5.4 Software design description and principles

The IEEE defines two standards for the architectural model:

- IEEE Standard for Information Technology — Systems Design — Software Design Descriptions.

- IEEE Standard for Systems and software engineering — Architecture description (manner in which architectural descriptions of systems are organized and expressed).

According to IEEE standards a SDD needs to have:

- Identification of the SDD (date, authors, organization,...).

- Description of design stakeholders.

- Description of design concerns.

- Selected design viewpoints.

- Design views.

- Design overlays.

- Design rationale.

The eleven design principles are:

1. Divide and conquer: simplify the problems with smaller ones.

2. Keep the level of abstraction as high as possible: ensure that your designs allow you to hide or defer consideration of details, thus reducing complexity.

3. Increase cohesion where possible.

4. Reduce coupling where possible.

5. Design for reusability: design the various aspects of your system so that they can be used again in other contexts.

6. Reuse existing designs and code: design with reuse is complementary to design for reusability.

7. Design for flexibility: actively anticipate changes that a design may have to undergo in the future, and prepare for them.

8. Anticipate obsolescence: plan for changes in the technology or environment so the software will continue to run or can be easily changed.

9. Design for portability: have the software run on as many platforms as possible.

10. Design for testability: take steps to make testing easier.

11. Design defensively: be careful when you trust how others will try to use a component you are designing.

5.5   The design process

Definition. Defining the software architecture or design is a problem-solving process whose objective is to find and describe a way to implement the system functional requirements while respecting the constraints imposed by the quality, platform and process requirements (including the budget) and while adhering to general principles of good quality.

The possible approaches are:

- Top-down: first design design the very high level structure of the system, then gradually work down to detailed decisions about low-level constructs and finally arrive at detailed decisions such as the format of particular data items and the individual algorithms that will be used.

- Bottom-up: make decisions about reusable low-level utilities then decide how these will be put together to create high-level constructs.

- Mixed approach: create the structure with top-down and define reusable components with bottom-up methods.

The designer is faced with a series of designed issues and needs to select the best option for every problem that encounters. The space of the possible designs that could be achieved by choosing different sets of alternatives is often called the design space.