

Contents

Abstract and Exposéé	iii
1 Introduction	1
2 Background	3
2.1 Fundamentals of Object-Oriented Programming	3
2.1.1 Encapsulation	3
2.1.2 Inheritance	4
2.1.3 Polymorphism	5
2.1.4 Single Responsibility Principle	5
2.1.5 Open Closed Principle	6
2.2 Design Patterns and Architectural Patterns	6
2.2.1 Repository	7
2.2.2 Model-View-Controller	7
2.2.3 Model-View-Presenter	8
2.2.4 Client-Server	11
2.2.5 Peer-to-Peer	11
2.2.6 Multi-Tier and Multi-Layer Architecture	12
2.3 Software Project Management	14
2.3.1 Waterfall	14
2.3.2 Scrum	14
2.3.3 Extreme Programming	14
2.4 Technologies	14
2.4.1 Swing	14
2.4.2 JavaFX	14
2.4.3 PostgreSQL	14
2.4.4 MongoDB	14
2.4.5 Cached Controller	14

Contents

3	Methods	15
3.1	Domain Description	15
3.1.1	Use-Cases	16
3.2	Phase 1: One Databse and UI	18
3.2.1	Monolithic Code Samples	18
3.2.2	Best-Practice Code Samples	18
3.3	Phase 2: More Databases / UIs	18
3.3.1	Monolithic Code Samples	18
3.3.2	Best-Practice Code Samples	18
4	Results	19
4.1	Lines of Code	19
4.2	Programming Effort	19
4.3	Efficiency and Performance	19
4.4	Touched and Edited Files	19
4.5	Readability	19
5	Discussion	21
5.1	Programming Overhead	21
5.2	Use of OOP in General	21
6	Conclusion	23
7	Future Work	25
7.1	Physical Seperation of Tiers	25
7.2	MVC-Frameworks	25
7.2.1	Spring Framework	25

Abstract and Exposéé

There are many many different approaches to achieve good software design, however often it is not clear whether the benefits of the best practices outweigh the additional effort of implementing them. Little research is done on comparisons between the designs in terms of efficiency and programming effort. Many books do show sample code of their proposed designs, however they do not cover how the code would look like if implemented in a more straight-forward and monolithic way.

This bachelor's thesis tries to cover this topic by implementing the same application two times: as a monolithic program and as a program that pays attention to the information provided about software design and architecture. In the first step, the program will use only one data source and one graphical interface (GUI). In the second step in each program another data source and GUI was added.

The sourcecode will be available on Github ¹ under GPLv3 ². The application will be written in Java using Swing as a front-end and PostgreSQL as database. In phase two JavaFX should be supported as a front-end, XML should be supported as a database. The user will be able to choose between all options using parameters when starting the application.

¹<https://github.com/Vallant/design-patterns-study.git>

²<https://www.gnu.org/licenses/gpl-3.0.de.html>

1 Introduction

The topic of design patterns and architectural styles are vital to the subject of software development. They empower inexperienced programmers to write maintainable applications and provide guidelines that may be adopted but are implemented in each application in one way or another.

Many books are available on the topic of software design and architectural styles, on how to write maintainable and extensible applications. The books are written by software engineers with many years of experience , like Martin Fowler with his book *Patterns of Enterprise Application Architecture* ¹, Robert C. Martin with *Agile Software Development. Principles, Patterns, and Practices* ² or *Design Patterns. Elements of Reusable Object-Oriented Software* by John Vlissides, Richard Helm, Ralph Johnson and Erich Gamma ³.

The same topics are to a certain degree taught at universities however they all lack one substantial flaw: While they propose some pattern, explain when it is applying and how to implement it, they do not provide any evidence that when using the pattern in a real-world application there is an advantage to gain. Many patterns claim that they improve the extensibility and increase the decoupling of code however in most of the example code only the basic implementation is showed. The impact on the code when a extension is actually implemented is not being shown by the books which leaves the reader to believe the promises made by the author.

While the patterns are in extensive use and thus have empirically proven themselves useful, the scientific research on this topic is far from being exhaustive. A study of 2008 was published at IEEE that tried to cover the effect of using design patterns in software development. This study came to

¹IBAN: 978-0321127426

²IBAN: 978-0135974445

³IBAN: 978-0201633610

1 Introduction

the result that while the implementation of best practice methods in many cases enhanced some quality aspects while at the same time performing worse in other factors. It concluded that the use of patterns should be considered carefully as they lead to a certain trade-off.

Another study published covering a similar subject came to different results, concluding that the use of established design patterns in most of the cases led to a significant improvement of code quality, contradicting the previous study. <http://ieeexplore.ieee.org/abstract/document/4493325/?reload=true> and <http://www.sciencedirect.com/science/article/pii/S0950584911002151>;

It is hardly possible to perform scientific research on real-world software projects as there are only little ways for serious comparison. The patterns only provide a coarse guideline, the actual implementation largely depends on the use case and will vary substantially from application to application. Because of this most of the research is done via *Controlled Experiments*, as is this thesis. Therefore it has to be said that most of this kind of studies are considered at least partially theoretical as a number of assumptions have to be made. The development of an application is highly subjective and depends on the coding style and experience of the engineer, code review policies can influence the quality as well.

However this thesis intends to give insights on how the use of well-established architectural and design pattern that claim to improve extensibility impact the code quality when this extensions are in fact implemented. This is done by developing two applications that are both designed to meet the same use cases and requirements. One is programmed by using established patterns while the other is developed without considerations regarding code quality while nevertheless implementing the same basic program structure in order to improve comparability.

After satisfying the basic requirements new features are being implemented, namely another type of persistence, a different kind of user interface and lastly means for caching data in the application to reduce the need of database queries. After realizing these changes in both programs the costs of implementing can be compared in both quantitative and qualitative ways.

2 Background

This chapter is intended to provide information on some basic principles in terms of the Object Oriented Programming. The list is by no means exhaustive, however is meant to help understanding the focus of this thesis.

2.1 Fundamentals of Object-Oriented Programming

The paradigm of Object Oriented Programming (OOP) uses *classes* as primary mean to gather and structure data. The data within a class is mostly called *attributes*, means to interact with it are called *methods* **Castagna97**

While a class is the abstract definition of such a container an *object* is a concrete instance filled with actual data. Attributes that may differ between each instance are therefore also called *instance variables*. Variables that belong to the class itself and thus are only instantiated once per class are called *class variables*.

2.1.1 Encapsulation

To encourage refactoring ¹ each class should prevent direct access to its internals from the outside. It should however provide a well-defined interface in terms of methods for manipulating the data, as this allows the class to

¹Refactoring describes the process of re-writing classes to improve readability, performance or other criteria while preserving its signature, inputs and outputs.

2 Background

enforce invariants². This means that it hides all information not relevant to others as they are only implementational details and thus not relevant. As other classes now rely on an interface rather than concrete implementations the code is called *loosely coupled*. Modifying the internals now does not break interdependencies which encourages programmers to perform refactoring resulting in improved code quality.

Many programming languages provide different access levels varying from visible to all others, accessible only within the class or visible from within and derived classes. The latter access level can be problematic as they effectively break the encapsulation. TODO TODO

2.1.2 Inheritance

Inheritance describes the concept of a child class inheriting all attributes, methods and other properties from a parent class. The child class is connected with the base through a *is-a*-relationship. The child is therefore a superset of the base class as it can be extended to meet the new requirements. This concept is important as it encourages developers to reuse existing code and in that way lower the risk of programming errors **johnson91**

A prominent problem often mentioned in this context is the *Diamond Problem* in the sense of multiple inheritance. It describes a situation in which at least two parents of a derived class share a single base class **Truyen04** If now a method of the topmost class is overridden by both ancestors of the lowest class, the question arises which of the two possible methods should be called. Some languages, such as Java or C#, do not support multiple inheritance for this reason, while others explicitly allow it, such as C++ or Python. In these cases if a situation as described in the Diamond Problem arises the results can cause undefined behaviour.

²A invariant is some kind of requirement that must be fulfilled at all times before and after a method call. An example of such an invariant could be the descending sort order of an integer list.

2.1.3 Polymorphism

Polymorphism describes the ability to tie the same interface to different belonging types. ³ There are two main kinds of polymorphism: The *overriding* polymorphism, which is tied closely to inheritance and describes the ability to choose at runtime between equally-called methods and attributes of a base class and its child class. For example, if a base class `Animal` has a method `speak`, each derived class `Dog` and `Cat` both inherit this method. With overriding polymorphism if the method is called the two subclasses are able to behave in different ways while providing the same programming interface. It is determined at runtime which method should be executed for an object.

The other important kind is *overloading* polymorphism which is used to provide methods with the same name but different signatures (and thus attributes). An example could be two methods called `add`, one taking a number, one taking a text as a parameter. Here it is determined at compile time which method will be used. ³.

2.1.4 Single Responsibility Principle

This important principle states that each class should only full-fill one particular purpose and as a result does only have one reason to change. The computer scientist D.L. Parnas wrote that in software development each design decision which is likely to change should be placed in a single, independent module and hides this decision from others **srp** When followed it avoids side effects on other responsibilities when changing the class.

A example of a class violating this concept could be a class that reads two numbers from the user, calculates the sum and prints the result. While this program seems quite simple three different responsibilities are placed in the same class. If either the means to provide the input, for presenting need to be modified or the algorithms should support other data types, the class need to be changed. The modifications may introduce errors to the other

³<https://docs.microsoft.com/en-us/dotnet/csharp/programming-guide/classes-and-structs/polymorphism>, accessed 31.08.2017

2 Background

functionality, which could be avoided if the program could conform to the principle. This would require placing each responsibility in an independent module.

2.1.5 Open Closed Principle

The Open-Closed-Principle states that each class should be open to extension and closed for modification **ocp**. As already written code is assumed to be well tested and working as intended, it should be avoided to modify it afterwards to add new functionality. Every change could lead to unwanted side effects that may only occur in very specific situations and are therefore difficult to detect.

Inheritance addresses this problem as it empowers the programmer to add new features while preserving all of the old code. Even when overriding methods in terms of polymorphism, the principle is not violated as the original class is preserved as-is.

2.2 Design Patterns and Architectural Patterns

In opposite to the previous chapter which describes rather abstract guidelines in object-oriented programming, this section focuses on widespread methods which try to help solving common problems that might occur in bigger projects. They are called *design patterns* and are subject to a large amount of books, of which the most famous is probably *Design Patterns - Elements of Reusable Object-Oriented Software* by Erich Gamma, Richard Helm, Ralph Johnson, and John Vlissides and *Patterns of Enterprise Application Architecture* by Martin Fowler. While the former book is more focused on general programming patterns, the latter one is more relevant for bigger businesses where a multi-tier architecture (see section 2.2.6) is more common.

The main difference between architectural and design patterns lies in different phases they are applied at. Architectural patterns provide high-level strategies providing a big picture often describing the interaction between so-called *components*, which can vary widely in its complexity and each of

2.2 Design Patterns and Architectural Patterns

them providing a well-described service and fulfilling a certain purpose. Design patterns focuses more on the actual implementation and describes a way to solve a problem on class and interface level.

2.2.1 Repository

The repository design pattern is a programming strategy that is useful when operating with some kind of data persistence service. It provides an abstraction layer between the business code and the place where data of the application is actually saved, providing a simple and clean interface for the business code and leads to a looser coupling. All actions dealing with specialized database operations, object retrieval and others are hidden by a repository, because of this it is considered part of a data-abstraction-layer (see section 2.2.6). A repository is often used if only basic operations with the data source are necessary. Other patterns and frameworks supporting advanced functionality and are available however they are not relevant to this thesis.

To create one repository per business object (data class) is the most basic way of implementing this pattern. Besides from the basic *CRUD*-Operations ⁴, each repository should provide specialized object retrieval methods that are only valid for the business model it was programmed for.

For example, in a booking system there may be a data class `Order`. A repository for this class may then provide a method called `getOrdersByUsername`. In general only methods that are actually needed should be implemented, which is described in the *YAGNI*-principle ⁵.

2.2.2 Model-View-Controller

The Model-View-Controller (MVC) is a pattern that describes a separation of a program in three layers, it is considered both a design and a architectural pattern, depending on the point of view. Its goal is to improve flexibility and

⁴Create, read, update, delete

⁵You aren't gonna need it

2 Background

extensibility of an application which can help to improve maintainability. It consists of three components, from which the *model* contains the business logic, the *controller* processes interactions and the *view* presents some kind of user interface.

This pattern is not very well defined and is subject of constant discussions. While it is commonly agreed upon that the application should be split in the three layers, the data flows and connections between the components are described in many different ways. The original concept of MVC developed for Smalltalk as described before the model holds logic and data of the application. Views subscribe at the model for notifications on changes of data in order to be able to update and refresh themselves. The controller was responsible of managing user interaction, which includes switching between views and forwarding requests to the model. This situation can be seen in Figure ?? . This direct interaction between model and view is often not desirable, especially in the context of web applications, which leaves it deprecated today.

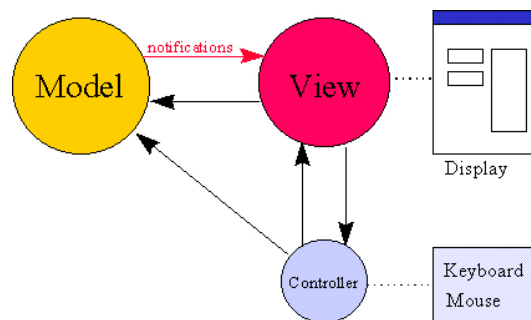


Figure 2.1: Classical Model-View-Controller concept as used in Smalltalk. Model and View know each other.

Source: <https://www.mimuw.edu.pl>, accessed 5.9.2017

2.2.3 Model-View-Presenter

The Model-View-Presenter (MVP) pattern is similar to the MVC-model as it shares the concept of separating the application into three layers. However the responsibilities differ considerably, as can be seen in Figure ?? . The view

2.2 Design Patterns and Architectural Patterns

only communicates with the presenter and does not know about the model. The presenter handles user input and requests, gathers data from the model, updates it and finally is responsible for refreshing the view with new data. This implies that the model does not know about the existence of view or presenter and at least in the original concept of MVP all business logic is placed in the presenter classes. It is important to note that it is vital for the use of this pattern to implement it using interfaces. This ensures that its main goal, easy testability, can be reached easily.

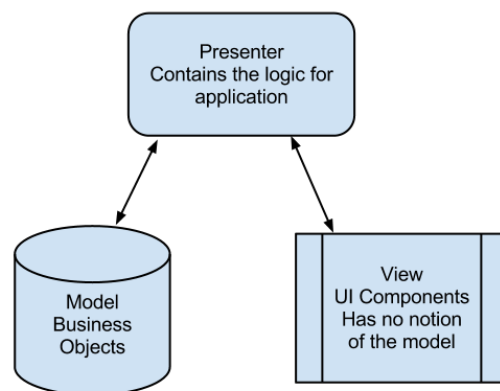


Figure 2.2: Diagram of Model-View-Presenter. While the view again only holds the means for displaying the data and interacting, in opposite to the MVC-model it does not know about the model but only corresponds with the presenter. The presenter is responsible for gathering the necessary data from the model and forwards it to the view. The model does not know about the presenter.

Source: <https://de.wikipedia.org/wiki/Model-View-Presenter>, accessed on 5.9.2017, by Daniel.Cardenas - Own Content, CC BY 3.0, <https://commons.wikimedia.org/w/index.php?curid=19794348>

After some years Martin Fowler drafted two design patterns that are closely connected to MVP. In his blog entry ⁶ he explained that he felt the need to split the MVP-pattern into two resulting in the following concepts, briefly described in Figure ??.

⁶<https://martinfowler.com/eaDev/ModelViewPresenter.html>

2 Background

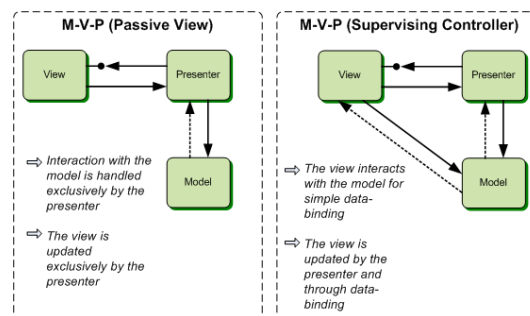


Figure 2.3: Basic diagrams of the Passive View (left) and Supervising Controller patterns (right). In the former as little logic as possible is placed in the view making it easy replaceable by a mock object for testing. Synchronization logic needs to be placed in the Presenter. In the latter view and model are connected through data bindings empowering these two to perform synchronization tasks on their own. This simplifies the role of the presenter.

Source: <https://msdn.microsoft.com/en-us/library/ff709839.aspx>, accessed on 5.9.2017

Supervising Controller

In this concept it is the views responsibility to perform data synchronization using shared classes between it and the model. This simplifies the tasks to be performed by the presentation layer while reducing the level of testability possible.

Passive view

In the concept of passive view the model is not connected with the view, which in terms only corresponds with the presenter. It is desired to minimize the logic placed in the view, leaving it (as the name describes) passive and therefore easy to mock. This is especially useful for automated testing, as it allows the tester not only to test the basic logic placed in the presenter but also its synchronization capabilities. The drawback of this clearly lies in the additional responsibilities that need to be met by the presenter.

7

⁷<http://www.wildcrest.com/Potel/Portfolio/mvp.pdf>

2.2.4 Client-Server

The Client-Server model describes a way to separate tasks within a network. While the client, which can be any kind of computer program, requests a resource from the server. It is designated to be used within the internet and is still one of the cornerstones of current web infrastructure. A prominent example of the client-server model is the Domain-Name-System (DNS). When a user types in the domain name of a service he wants to access, the browser needs to translate the name into a IP-Address he can work with. In order to do that he sends a request to a DNS-Server, making him a client. While he does not need to know how exactly the resolving takes place he requires the server response to have a certain format described in the communication protocol.

There are several benefits of using this architecture. For starters it centralizes the administration which lowers the effort to managing access rights, accounts and so on. Because the important data is in one place it is easier to create backups and ensure security policies. However a server is considered a *Single Point of Failure*, meaning that a whole service will be unavailable if the server for any reason is not reachable, making it a worthy target of attacks. While the impact depends on the actual service built, scalability and extensibility are important factors. It is easier to extend the features provided by a service when using the client-server-model it reduces scalability as the hardware needs to be upgraded from time to time to meet a increased number of requests directed to the server.

2.2.5 Peer-to-Peer

Peer-to-Peer (P2P) infrastructure is the counterpart of client-server-models. No central server exists which means that each of the clients are equally featured, each client has the means to request and provide resources. P2P-Networks are considered more robust and scalable than Client-Server set-ups, as the failure of one node does not lead to the breakdown of the whole system.

2 Background

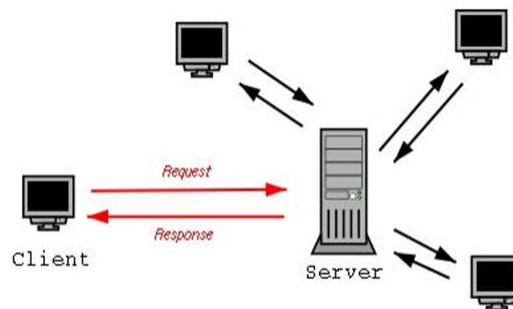


Figure 2.4: Client-Server setup. The client sends a request to a server by messages defined in the protocol. The processing of the request is transparent to the client, it does not know about the actions taken by the server and awaits a response in a certain format.

Source: <https://abcnetworking.wikispaces.com/Client+Server+Networking>, accessed on 6.9.2017

2.2.6 Multi-Tier and Multi-Layer Architecture

Multi-Tier architecture is mostly mentioned in the context of web applications. In general it describes the separation of responsibilities on different *physical* places⁸. This is the main difference to a multi-layer-system where the responsibilities are separated but placed at the same process which indicates the use of interfaces.

A 3-Tier-Architecture nowadays refers mostly to the typical set-up of a web application as it can be seen in Figure ?? . It has to be noted that it is very common to use multi-tier and multi-layer architecture mutually. For example, the logic tier in the figure could use a data-access-layer to abstract the communication with the data tier.

⁸In this context physical separation means that there has to be going on any kind of communication between independent processes within the same machine or connected trough the internet.

2.2 Design Patterns and Architectural Patterns

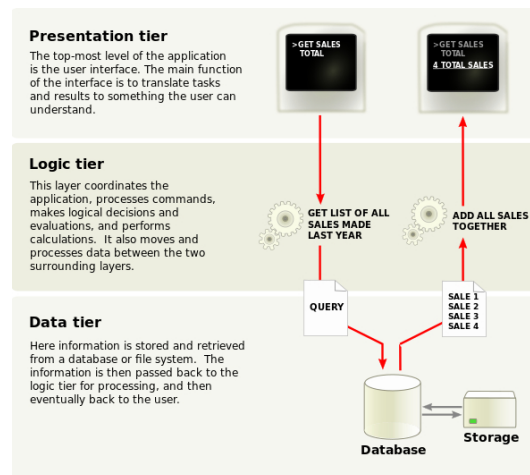


Figure 2.5: Multi-Tier set-up. The tiers are physically separated.

Source: https://en.wikipedia.org/wiki/Multitier_architecture, accessed on 6.9.2017

2 Background

2.3 Software Project Management

2.3.1 Waterfall

2.3.2 Scrum

2.3.3 Extreme Programming

2.4 Technologies

2.4.1 Swing

2.4.2 JavaFX

2.4.3 PostgreSQL

2.4.4 MongoDB

2.4.5 Cached Controller

3 Methods

3.1 Domain Description

As stated before the application should provide a time-tracking solution for commercial projects. Therefore it is divided into three general areas:

- **Project:** A project is the overall term for a distinguishable amount of work that is done for a specific customer.
- **Phase:** Each project exists of one or more phases that are used to monitor the progress of the project.
- **Activity:** A activity represents a workload that is done by a specific member in a specific phase for a specific project.

In this application, two basic roles exist.

- **Project Leader:** A project leader has the responsibility of maintaining project specific data, such as description, phases and assigning project members. He is *not* allowed to create new users. After creating a new project, one is automatically project leader. A project leader is allowed to monitor the statistics of his projects including the workloads of the project members
- **Member:** As a simple member one is allowed to add, delete and modify activities created by the user. A member is only allowed to perform this actions for projects he is assigned to. He is also only allowed to monitor his own statistics.

3 Methods

3.1.1 Use-Cases

General Use-Cases

1. **Login:** After starting the program, a login prompt shall show up. The user provides his credentials which will be compared to the saved credentials in the database. If they are correct, he is forwarded to the main program.
The user has the possibility to reset his password by activating a button and providing his email-address. A new password will be sent to the address, if a user with this address is known by the system.
2. **Starting / Stopping activities:** Through a banner at the top of the program each user can select a project he is assigned to and furthermore a phase of this project. After providing a name of the activity he can start the activity by clicking on a start button. The same button can be used to stop the activity. After stopping (or closing the application), the activity is complete and is added to the database and project statistics.
3. **Personal statistics:** Utilizing the associated button of the left side bar the main panel lists all project the user is assigned to and the time he spent working for these projects. At the top of the main panel he can choose the period he wants to see. After clicking on one of the entries a statistic of the working time for each phase of the project is shown. By clicking at a phase all activities in this phase are listed and can be modified and deleted, furthermore a new activity can be created here too. By clicking on a button at the top the user can switch the view to the previous view.
4. **Create and manage projects:** Each user can create new projects by using the associated button of the left side bar. The main panel then lists all projects which project leader he is. After clicking on a entry, the main window shows the details of the project, such as the description, the created phases and the project members. He is able to perform all necessary operations associated with the project in this view, as well as declaring a new project leader.
5. **Settings:** Utilizing the associated button of the left side bar the main panel provides means to change the password, email-address and name.

Project leader Use-Cases

1. **Project statistics:** Other than normal users, a project leader can monitor the statistics of all of his projects by using the according button on the side bar. The main panel lists all projects he is leader of. After choosing one project he can choose the period he is interested in at the top of the panel. At the left side of the main window now all project members are listed. After choosing one or more project members, the time they spent working for the different phases of the project is listed at the center.

The persistent storage has a quite simple structure, as can be seen in figure 3.1. It is notable that the roles are described as text and therefore are human-readable.

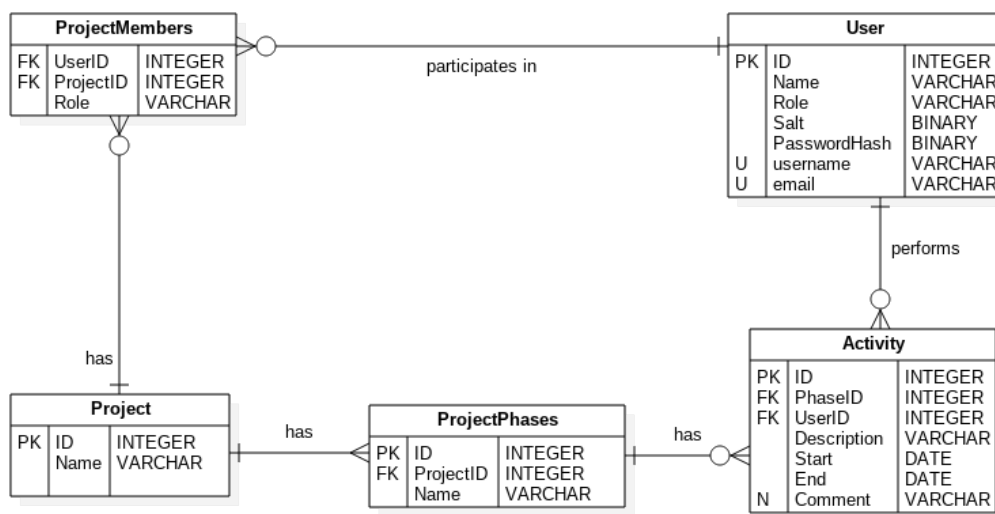


Figure 3.1: Database diagram

3 Methods

3.2 Phase 1: One Database and UI

3.2.1 Monolithic Code Samples

3.2.2 Best-Practice Code Samples

3.3 Phase 2: More Databases / UIs

3.3.1 Monolithic Code Samples

3.3.2 Best-Practice Code Samples

4 Results

4.1 Lines of Code

4.2 Programming Effort

4.3 Efficiency and Performance

4.4 Touched and Edited Files

4.5 Readability

5 Discussion

5.1 Programming Overhead

5.2 Use of OOP in General

6 Conclusion

7 Future Work

7.1 Physical Separation of Tiers

7.2 MVC-Frameworks

7.2.1 Spring Framework

Composing06

List of Figures

2.1	Classical Model-View-Controller concept as used in Smalltalk. Model and View know each other.	8
2.2	Diagram of Model-View-Presenter. While the view again only holds the means for displaying the data and interacting, in opposite to the MVC-model it does not know about the model but only corresponds with the presenter. The presenter is responsible for gathering the necessary data from the model and forwards it to the view. The model does not know about the presenter.	9
2.3	Basic diagrams of the Passive View (left) and Supervising Controller patterns (right). In the former as little logic as possible is placed in the view making it easy replaceable by a mock object for testing. Synchronization logic needs to be placed in the Presenter. In the latter view and model are connected through data bindings empowering these two to perform synchronization tasks on their own. This simplifies the role of the presenter.	10
2.4	Client-Server setup. The client sends a request to a server by messages defined in the protocol. The processing of the request is transparent to the client, it does not know about the actions taken by the server and awaits a response in a certain format.	12
2.5	Multi-Tier set-up. The tiers are physically separated.	13
3.1	Database diagram	17

List of Tables