An Empirical Investigation of Methods for Maximizing Knowledge and Performance in Framework Documentation Using Use Case Maps (UCM)

1Hee Wei Loon, 1Tan Chai, 2Tan Chuie Hong and 1Ho Sin Ban  
1Faculty of Computing and Informatics, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Selangor, Malaysia  
2Faculty of Management, Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Selangor, Malaysia

Abstract: In today's world of software engineering, timeframe and design consistency of a project have become a major, if not, crucial concern among developers. When developing an application, developers cannot afford to exceed the timeframe that has been allocated but at the same time, need to maintain the consistency of design throughout the whole development process. One way to support this is to have good framework documentation. Framework documentation is a process whereby a framework is documented in a proper manner so that it could be easily reused in any future application development. Two main documentation styles have been included in this study-patterns and minimalist styles. It is not easy to document a framework well. Therefore, well-established theories or rules to follow when documenting are helpful. The aim of this study is to discover a new method that can improve the learning curve and to make mass learning possible among users in framework documentation. It focuses on minimalist documentation and pattern documentation but with a new add-on approach augmented with Use Case Maps (UCM). Patterns have the idea of providing the solution to a problem in its context. It has contextual information which is not minimal, unlike minimalist. Minimalist documentation promotes minimal reading and random access which means that the learner can proceed in a self-directed manner as well as using small pages that can be read as per the learner's desires on the reading order. The use of UCM in these documentation styles has been examined in detail using a quantitative research method. We found UCM to be more effective in allowing learners to perform efficiently and understand the framework better when used in conjunction with Minimalist documentation.

Key words: Framework documentation, patterns, minimalist, UCM, Malaysia

INTRODUCTION

A framework is a collection of reusable parts that can be reused for a particular entity (Zamir, 1998). This framework can be reused or manipulated into new framework designs when there is a need to do so. Eventually, this framework allows tasks at hand to be executed effectively, efficiently and consistently without any major flaw.

This framework concept works effectively in Object-Oriented Programming (OOP). The framework is expressed in code. In order to reuse a framework to build new applications, one has to document them in a proper manner so that they can be easily understood when needed later for application development. In the world of software engineering today, timeframe and consistency of a project have become a major, if not, crucial concern among developers. When developing an application, developers cannot afford to exceed the timeframe that has been allocated and at the same time, should maintain the consistency of design throughout the whole development process. If one of the above criteria is not fulfilled, the whole project development might be in jeopardy and finally fail. So, the question is how to make the application development process a success with minimal flaws in terms of timeframe usage and consistency of design? In order to achieve that, framework documentation plays an important role in today's OOP development to ensure the application development process meets its timeframe and achieves consistency in its design as well (Fayad and Schmidt, 1997). This is the most effective and efficient way to develop an application.

Last but not least, there is a lot of research being carried out throughout the world today in order to
improve or further enhance current framework documentation theories such as John Carroll’s Minimalist theory, etc.

The focus of this study is to maximize knowledge and performance in framework documentation using Use Case Maps (UCM). The use of UCM in these documentation styles has been examined in detail using a quantitative research method. We found UCM to be more effective in allowing learners to perform efficiently and understand the framework better when used in conjunction with Minimalist documentation.

**Literature review:** It is not easy to create good framework documentation (Aguiar and David, 2000). A framework is basically useless if the learners are not able to understand and utilize it to meet their objectives. Thus, well-established theories or guidelines to follow are needed when documenting a framework. By having good framework documentation, the learners will be able to understand and know exactly how to apply it in their work, hence, making the learning process a success. This research uses several pedagogical framework documetations methods. The two main documentation styles being used are patterns and minimalist. We add Use Case Maps (UCM) to these styles to study if it enhances the pedagogical utility. The benefits and reasons of having UCM will be discussed later in this chapter. Therefore, in this research, we will have four documentation styles-patterns, patterns with UCM, minimalist and minimalist with UCM.

In today’s information technology world, each piece of software has its own dedicated documentation. Most documentation tells people how to perform the task, instead of telling them the purpose of the task itself and why they need to perform the task (Dix et al., 2004). This indirectly decreases the understanding of the readers and eventually prevents them from having the ability to reuse the same experience in other computer software. This is one of the main drawbacks of traditional documentation. To address this concern, patterns were introduced. One of the objectives of Patterns is to allow readers to understand the rationale behind a solution-why a particular task is required, so that they can have the ability to make their own judgement on when to apply the pattern. Patterns are proven successful solutions for recurring problem in a context (Schumacher, 2003). According to Meszaros and Doble (1996), patterns should have these attributes:

- **Pattern name:** for tracking purposes so that reader can refer to the pattern easily
- **Problem:** a precise and detailed description of the problem it solves
- **Solution:** a detailed description of the solution to the problem and a prescription on how the solution works to solve the problem
- **Context:** the circumstances of the problem that provide constraints on the solution
- **Forces:** common contradictory considerations that must be taken into account when choosing a solution to a problem

Minimalist documentation was pioneered by Dr. John Carroll based on two concepts, firstly, information which is irrelevant to the task at hand is not needed, and secondly, step-by-step instructions are hard to follow. Carroll (1998) created this Minimalist documentation to give the reader a minimal amount of information to read to accomplish the task. At the same time, it uses short pages or index cards of information and allows readers to read them in any order they prefer. Carroll gives some guidelines for minimalist documentation as follows:

- **Training on real tasks:** readers are more determined and motivated to complete a task if it is something they want to achieve
- **Getting started fast:** readers will lose interest easily if there is too much to read before performing a task and this will cause them to miss things easily
- **Reading in any order:** topics are brief and short which allow readers to choose the order they prefer and suit them the best
- **Coordinating system and training:** allow readers to interact with the system instead of providing detailed steps to guide them
- **Supporting error recognition and recovery:** instead of assuming readers follow step-by-step instructions flawlessly, expect them to fail and provide resources for them to understand how to recover
- **Exploiting prior knowledge:** exploiting prior knowledge by using analogies of readers’ prior experience to help them understand, instead of using jargon
- **Using the situation:** take advantage of the readers’ expectations that are brought to the situation

In today’s world of programming, requirement gathering and precise software engineering are very crucial to ensure software runs correctly and more importantly, meet the objective of creating the program. According to Amyot and Mussbacher (2001), several common issues have been identified for requirements gathering:
• Early focus on low-level abstractions
• Specific requirements and high-level decisions buried in the details
• Change of functionalities hard to manage and
• Introduction of new services are delayed
• They have also identified common issues faced in software engineering
• Difficulties to come out with requirements or analysis models required to support new dynamic systems
• The need to change from requirements or analysis models to design models seamlessly

Use Case Maps (UCM) tackles these issues. UCM is a notation to illustrate a scenario path relative to optional components that may exist in a scenario. It describes causal relationships between responsibilities of use cases in a map-like diagram.

MATERIALS AND METHODS

The experiments use learning Java ME as a test bed task. These experiments test to see if having UCM embedded into a documentation style will actually help the learner to perform and understand better. All of the documentation styles were converted and presented online. Thus, all of the subjects participated and completed the work task by following one of the sets of online documentation. There are four sets online documentations: Patterns, Patterns with UCM, Minimalist and Minimalist with UCM (Table 1).

There are two types of variables-independent (factors) and dependent. In statistical analysis, it is a common practice (Ho, 2008; Ho et al., 2007) to denote independent variables as x values such as x1, x2, x3, and so forth until xn and dependent variables as Y values from Y1, Y2, Yk, until Ym. Using formula expansion, we can show the explanation above as follows:

\[ Y_i = c_0 + c_1 x_1 + c_2 x_2 + c_3 x_3 + ... + c_n x_n \]

where, c represent any constant to fulfill the equality between Yi and xi. Table 2 shows the independent and dependent variables derived from the experiments.

The listed variables above form the questions in the tutorial exercise. Dependent variables which have no significant values in homogeneity (Ho et al., 2009) during the statistical tests will be omitted.

Table 3 shows examples from the experiments with all the mapping between dependent variables and their exercise questions that serve as a base to establish test items for the framework.

Patterns documentation style includes the background of the problem or issue as well as the solution.

<table>
<thead>
<tr>
<th>Table 1: Describes the variables used in the experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>Independent variables</td>
</tr>
<tr>
<td>Dependent variables</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2: Independent and dependent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable type</td>
</tr>
<tr>
<td>Independent (x)</td>
</tr>
<tr>
<td>Dependent (Y)</td>
</tr>
</tbody>
</table>
Table 3: Mapping between dependent variables and their exercise questions

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Examples from experiment</th>
</tr>
</thead>
</table>
| Comprehension      | Q1: Which package(s) will be called does <the event>?  
|                    | Q2: Which abstract class to be extended when <the event>?  
|                    | Q3: Which interfaces to be implement when <the event>?  
|                    | Q4: Which line of code will be executed for <the event>?  
|                    | Q5: Which method in the program that does <the event>?  
| Workings           | Q1: Write the necessary Java ME code when importing classes  
|                    | Q2: Write the necessary Java ME code when creating main class  
|                    | Q3: Write the necessary Java ME code when defining variables  
|                    | Q4: Write the necessary Java ME code when creating classes and procedures  

<the event> refers to functionality achieved from code execution

![Tutorial: How to Write a Java ME Application]

Fig. 1: Patterns

needed to solve the problem or issue as highlighted in Fig. 1. Patterns documentation with UCM is the same as Patterns but with the addition of UCM (Fig. 2). UCM consists of a path that shows the causal relationships between responsibilities of one or more use cases in a map-like diagram.

Minimalist documentation contains as little information as possible for the reader to read in line with the idea of having people read the minimal amount of information to get the task done. Minimalist documentation does not state the background of a problem and the solution needed to solve the problem, unlike patterns. This is shown in the screenshot in Fig. 3.

Minimalist documentation with UCM is the same as the minimalist but with the addition of UCM (as shown in the second screenshot in Fig. 4).

Apart from setting up the sets of documentation in the various styles, a survey collected data from subjects. According to Trochim (2006), surveys can be divided into two main categories which are the questionnaire and the interview. The questionnaire method usually involves paper-and-pencil instruments by having a set of questionnaires prepared for the subjects to complete. While interviews are conducted by the interviewer recording what the subjects say. The questionnaire method can be divided into three types-mail survey questionnaire, group-administered survey questionnaire and household drop-off survey questionnaire. Interviews are divided into two types personal interview and telephone interview.

For this research, the group-administered survey questionnaire method is adopted and performed. By having this group administered method, the subjects involved in the survey are brought together to respond to a structured sequence of questions in a more controlled environment (i.e., classes) and are also more likely to have the same background or experience. The rest of the survey types are not suitable for this type of pedagogical research. The survey questionnaire form used to collect data from the subjects for this research is shown in Fig. 2.
How to Write a Java ME Application

Tutorial Topics:

1. First of all, you have to record Start Time: _______ AM/PM

2. After reading on how to get started writing a Java ME application, record Quarter Completion Time: _______ AM/PM

3. Please proceed to Task 1 (Checkpoint)

4. Upon completion of Task 1, record Semi Completion Time: _______ AM/PM

5. Please proceed to Task 2 (Checkpoint)

6. Upon completion of Task 2, record 2nd Quarter Completion Time: _______ AM/PM

Fig. 2: Patterns with UCM

Fig. 3: Minimalist
RESULTS AND DISCUSSION

The data analysis will be carried out to identify which documentation style allows the subjects to complete parts of the documentation (END1STQ, ENDSMEI, END2NDQ) and complete the overall documentation the fastest (ENDDFINAL) with the number of difficulties recorded by the subjects at the intervals (DIFF). The documentation style which allows the subjects to understand the most are also analyzed and recorded (COMPR) as well as their knowledge in the inner workings of Java ME (WORK).

The meanings of the variables are:

- END1STQ: end of 1st quarter
- ENDSMEI: end of semi completion
- END2NDQ: end of 2nd quarter
- ENDDFINAL: end of completion
- DIFF: number of difficulties faced
- COMPR: comprehension
- WORK: workings

In order to identify whether the seven dependent variables above are normally distributed or not, the
normality of the variables was tested using the One-Sample Kolmogorov-Smirnov test (K-S test). The normality test show that four (out of seven) dependent variables are normally distributed as shown in Table 4.

**One-sample Kolmogorov-Smirnov test (K-S test):** K-S test is used to determine the normality of the dependent variables (D’Agostino et al., 1990). The calculation result is focused on the significance of the p-value. In particular, p-values that are <0.05 are considered “statistically significant” and interpreted as being small enough to justify rejection of the null hypothesis (Lane, 2015). Based on the K-S test being carried out for all of the dependent variables, we use the 5% level (i.e., p<0.05) as the indicators to determine whether it is normally distributed or non-normally distributed. If the p<0.05, it is statistically significant and thus, it is non-normally distributed. For non-normally distributed variables, we use the Kruskal-Wallis test (Bluman, 2004). If the p>0.05, it is not statistically significant which means that the data is normally distributed. This means that we can use Multivariate Analysis of Variance (MANOVA) for the analysis. The Kruskal-Wallis and MANOVA tests will be discussed in detail later in their respective sections. Table 5 shows the results of the normality test. The median is used as the expected values due to the non-parametric outcome of the dependent variables of END1STQ, COMPR and DIFF.

**Means and standard deviations:** Secondly, we present the means and standard deviations of the seven dependent variables for each of the four documentation styles. The purpose of this analysis is to observe which documentation style performs the best. The fastest END1STQ, ENDSEMI, END2NDQ and ENDFINAL are the lowest achievement time whereas the best COMPR and WORK are the highest scores. The lowest DIFF indicates the least number of difficulties the participants face in learning Java ME.

Table 6 shows that Minimalist with UCM has the best ENDFINAL. The fastest ENDFINAL in Minimalist with UCM can be due to the minimal reading materials within the documentation style. Apart from this, by having UCM (Amyot and Muesbacher, 2001) integrated with the traditional Minimalist style, the documentation is enhanced, resulting in a faster completion time (ENDFINAL) than traditional Minimalist documentation.

However, Minimalist documentation has the best WORK score, perhaps due to the fact that it has the least amount of material to read among all of the documentations; thus, it is slightly easier to follow. On the other hand, patterns documentation has the highest achievement for the COMPR score. However, by adding UCM, COMPR increases when we compare Minimalist to Minimalist with UCM. Patterns with UCM is the worst in terms of performance and understanding of the learners.

![Image](image.png)
Table 7: MANOVA results for the normally distributed variable

<table>
<thead>
<tr>
<th>Category</th>
<th>F-values</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENDSemi</td>
<td>4.822</td>
<td>0.003*</td>
</tr>
<tr>
<td>END2NDQ</td>
<td>4.477</td>
<td>0.005*</td>
</tr>
<tr>
<td>ENDFinal</td>
<td>4.692</td>
<td>0.004*</td>
</tr>
<tr>
<td>WORK</td>
<td>4.392</td>
<td>0.006*</td>
</tr>
</tbody>
</table>

*Statistically significant at 0.05 level

Table 8: Mean ranks of END1STQ, COMPR and DIFF

<table>
<thead>
<tr>
<th>Documentation Style</th>
<th>Pattern with UCM</th>
<th>Minimalist with UCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size (n)</td>
<td>30.00</td>
<td>30.00</td>
</tr>
<tr>
<td>Mean rank of END1STQ</td>
<td>57.92</td>
<td>73.65</td>
</tr>
<tr>
<td>Mean rank of COMPR</td>
<td>66.78</td>
<td>55.12</td>
</tr>
<tr>
<td>Mean rank of DIFF</td>
<td>50.65</td>
<td>71.18</td>
</tr>
</tbody>
</table>

Table 9: Kruskal-Wallis test on END1STQ, COMPR and DIFF

<table>
<thead>
<tr>
<th>Category</th>
<th>Chi-square</th>
<th>df</th>
<th>Asymptotic significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>END1STQ</td>
<td>8.855</td>
<td>3</td>
<td>0.031*</td>
</tr>
<tr>
<td>COMPR</td>
<td>1.780</td>
<td>3</td>
<td>0.619</td>
</tr>
<tr>
<td>DIFF</td>
<td>6.320</td>
<td>3</td>
<td>0.047</td>
</tr>
</tbody>
</table>

*Statistically significant at 0.05 level

Multivariate Analysis of Variance (MANOVA): The dependent variables identified as normally distributed during the K-S test are further analyzed using the MANOVA test. This is to identify whether the differences between the four documentation styles are significant or not.

Table 7 shows that all of the categories, i.e., ENDSemi, END2NDQ, ENDFinal and WORK are significant at the 0.05 level. This means that Minimalist with UCM can be beneficial in terms of performance when guiding the learners to learn to develop the Java ME phone calendar application.

Kruskal-Wallis test: Lastly, for the remaining three dependent variables, i.e., END1STQ, COMPR and DIFF, we carried out the Kruskal-Wallis test (Bluman, 2004). This is because of their non-parametric results from the previous K-S test. The mean ranks shown in Table 8 are consistent with Table 6 with END1STQ being the best for Minimalist with UCM while COMPR and DIFF are the best for Patterns.

Table 9 shows that END1STQ is statistically significant at the 0.05 level while COMPR and DIFF have no significant differences. The non-significance of COMPR and DIFF may due to not much of comprehension or understanding being required to write a simple phone calendar application in Java ME. However, the significant difference in END1STQ shows that Minimalist with UCM style is effective in guiding learners to follow through the documentation section of the phone calendar application. This supports the hypothesis that using UCM increases performance (i.e., efficiency).

CONCLUSION

This empirical study evaluated the effectiveness of Use Case Maps (UCM) in helping mass learning when used with two different framework documentation styles, i.e., Patterns and Minimalist. Hence, four Pedagogical Framework Documentations (PFD) have been tested, Patterns vs. Patterns with UCM and Minimalist vs. Minimalist with UCM.

This study focuses on teaching readers how to use frameworks with the help of Use Case Maps (UCM) within pedagogical documentation. The motivation of the research is to close the gap between the requirements gathering and design stage by introducing UCM. However, UCM is not necessarily applicable to any documentation style. Thus, this study helps to identify which documentation style suits UCM better - Patterns or Minimalist.

The analysis results show that adding UCM to pedagogical documentation increases knowledge (understanding) and performance of the learners in developing Java Micro Edition (JME) applications. However, the amount to read in pedagogical documentation affects or impacts the effectiveness of UCM. This is due to the fact that more reading (i.e., Pattern with UCM) causes more confusion to the learners. This will indirectly cause the understanding of the learners to decrease and therefore, impacts their performance in completing the framework. Based on the above, we can conclude that UCM works better in Minimalist documentation (less to read) than Patterns (more to read).

This study solves several questions but at the same time raised a few more. First, there is a need to run these experiments using other programming language than Java ME, to see if the type of programming language used in the framework documentation has an impact on the overall effectiveness and efficiency of the documentation. Second, how did the subjects know how to answer the comprehension questions correctly if they skipped the section in the documentation style describing the internal workings? Third, if we apply this programming framework documentation to other different frameworks such as project management frameworks etc., will the observations still be valid? Another possible future work is to investigate other alternatives than UCM - how would other types of diagrams such as entity relationship diagrams (Jain et al., 2004; Schuch, 2005) or flowcharts affect the outcome of the experiments?
REFERENCES


