We propose a new formal model of UML use case diagram using Z notation to address some of its shortcomings. UML use case diagram has therefore become commonly used to structure functional requirements and the greatest challenge facing the software developer nowadays is to deliver a high quality product meeting customers’ requirements. However, the major disadvantage of UML models is their imprecision. In addition, they are basically in a form of semiformal modelling representations and associated natural language requirements and lack any mechanism to rigorously check consistency which results in its models being subject to ambiguity. This paper reports on the first formal modelling approach of use case diagrams in a multiview context. The approach is divided into two steps. In the first step, a formal model of UML use case diagram is proposed using Z notation. Then, a multiview consistency checking is presented. This approach guarantees software consistency, improving requirements quality.

1. Introduction

Today, people’s daily life is strongly dependent on software systems which become increasingly large and complex. "In software, requirements are the foundation for everything we do, so the quality of requirements is paramount" [1]. As the software in today’s systems grows larger, requirements documents become enormous and ambiguity is inevitably introduced. Many software development projects fail due to problems in requirements. Accordingly, it has become very important to be able to improve software quality by paying more attention to requirements activities. UML use case diagrams [2, 3] have therefore become commonly used during early stages of deriving software requirements and have been incorporated into most major object-oriented development methods thanks to its simplistic designs promoting good software engineering practices. However, on the one hand, the lack of formality in UML use case diagrams definition provides ample room for misunderstandings. On the other hand, UML models are not directly analysable and lack the ability to be type checked mechanically.

In this article, we propose a formal model of use case diagrams using Z notation [4, 5] to benefit from the rigour that comes from formal methods. By this formalism concept, we overcome UML use case diagram limitations allowing the ability to use type checkers and theorem provers on Z specification such as Z/EVES system. It is noteworthy that this paper reports on the first formalization of use case diagrams in Z notation that considers the consistency in multiview modelling.

This article is a continuation of [6–8], where they first described a formal model of class, sequence and state machine diagrams and checked the consistency between the different views. The extensions in the current paper are as follows: (1) a formal model of use case diagrams using Z notation, (2) a multiview consistency checking by means of Z theorems. As was pointed out above, "Formal specifications can be very useful artifacts at various stages of software development" [9]. The purpose of this article is to profit from the expressiveness and the mathematical basis of formal methods to improve requirements quality. More than this, “formal methods contribute to demonstrably cost-effective development of software with very low defect rate” [10].

The rest of this paper is organized as follows. Section 2 reviews relevant existing work that has been done in use case diagrams formalization. Section 3 presents a formal model
of use case diagrams using Z notation. Section 4 discusses multiview consistency between class, sequence and use case diagrams using Z theorems. Finally, Section 5 contains some concluding remarks and outlines some future directions of our work.

2. Related Work

Formalizing UML use case diagrams has drawn great attention from research community because it provides a considerable potential to help developers analyze whether the specification capture the intended requirements. In this section, we present some of the research literature showing approaches that define UML use case diagram semantics and discuss its similarity to and difference from our approach. Two main research schools have been working on this problem. The first is to propose techniques and tools to transform use case diagrams into different formalisms. For example, Junior et al. [11] propose a formalization of use cases using graph transformation models with the goal of running tool supported analyses on them and revealing possible errors. Sinnig et al. [12] propose an integrated development methodology based on non deterministic finite state machine formalism which defines a common semantics for use cases and task models. Butler et al. [13] define the notion of use case and the related terminology by means of a specification written in Z. The focus is on formalizing and clarifying the basic notions of use case and its related concepts. In our work, the focus is on specifying the different components and constraints imposed on/by use cases. Also formalizations of actors and inheritance relation between them and extension points are added. Barajas [14] presents the formal specification for a tool that models the functional requirements of a system based on use case models using Alloy. Our model is similar to that work but using Z language and with slight differences in addition to our multiview consistency checking. Scandurra et al. [15] propose the framework asmetaRE to automatically transform use cases models into ASM executable specifications and then validate systems requirements through simulation and scenario based simulation of the generated ASM. Shen and Liu [16] propose a new rigorous review technique including execution and testing which can be applied to software requirements models using a new language HCL (high level constraint language). Murali et al. [17] describe an approach for embedding a formal method within UML use case modeling using event-B. They extend use case modeling to allow for the explicit representation of safety concerns. Mostafa et al. [18] propose a formalization of the syntax of the popular UML diagrams (use case diagram, class diagram and state machine diagram) using Z specification. They formalize each basic notion of use case diagram without specifying it as a whole. In our model, the use case diagram is well explained and formalized. The model was improved by including the notion of scenario in case of extension.

While the first school has only targeted formalization issues, the second school focuses on multiview consistency between use case diagram and other UML diagrams. For example, Yue et al. [19, 20] propose a method and a tool called aToucan to automatically generate a UML analysis model comprising class, sequence and activity diagram from use case model. However, a use case diagram is not sufficient in its information to generate complete UML analysis model. While they chose to transform use case model into a UML analysis model, we propose a formal specification of use case diagram using Z notation providing a basis for checking consistency between use case diagram and a subset of formalized UML diagrams. Giese and Heldal [21] present a way to relate informal requirements in form of use cases to more formal specifications written in OCL. Ibrahim et al. [22] propose a set of consistency rules between use case diagram, sequence diagram and class diagram. The abstract syntax of the proposed consistency rules are formally defined using logical approach. Most of the consistency rules cited in this work are treated by our proposed model using Z notation.

Overall, these studies highlight the need for formalizing use case diagrams. However such studies remain narrow in their focus. Our approach includes transforming use case diagram into formal model using Z notation and multiview consistency checking between different UML diagrams. Furthermore, our work is one of very few existing works on multiview consistency checking and the first one using Z notation. Needless to say, through our approach both customers and developers can enjoy the benefits of formal languages to improve the quality of requirements and subsequently the quality of the obtained software.

3. Formal Model of Use Case Diagrams

Use case diagrams are considered for high level requirement analysis of a system for the purpose of capturing its dynamic view. They are essentially used to gather requirements and functionalities of a system captured in use cases and also to identify internal and external agents interacting with the system. These agents are known as actors. Figure 1 shows an example of use case diagram of ATM System [3].

So in brief, use case diagrams consist of actors, use cases and their relationships. Figure 2 illustrates the abstract syntax of use case diagrams [3].

The first step in the approach is to formalize the notion of use cases and relationships among them. To get a better understanding of the proposed formal model, we first define the different terms involved in the specification such as included use cases, extending use cases and extension point.

We will use Z notation to specify and describe the different notions of use case diagram. The basic concepts of Z notation [4, 5] are summarized below.

3.1. Summary of Z Notation. Z [4] is a formal specification language created by J.R. Abrial based upon set theory and mathematical logic. In Z notation, a specification uses the notion of schema to structure the underlying mathematics and allow an easy reuse of its subparts. Its graphical notation, called schema, is represented as below:
Figure 1: Example of the use cases and actors for an ATM system.

Figure 2: The abstract syntax of use cases package.
A schema contains a declaration part that may contain the declaration of state variables beside references to other schemas and a predicate part which consists in a set of predicates constraining the variable state values. These predicates express properties on the state variables and introduce relationships between them. The name of the Z schema enables its re-use. A Z schema may be used or re-used as a declaration, a type or a predicate. When the specification requires a composite type, a schema is used to denote it.

3.2. Use Cases Formalization. According to the UML standard, a use case is defined as a classifier. Actually the set of all unique identifiers throughout the specification is denoted by the given set CLASSIFIER.

Therefore, the variable UCASE is introduced as a classifier to specify the set of use cases’ identifiers.

\[\text{CLASSIFIER}\]

\[\text{UCASE} : \mathcal{P} \text{CLASSIFIER}\]

There are several different kinds of relationships between use cases. These are the include relationship, the extend relationship and generalization, all of which are described in this section.

Include is a relationship between two use cases which is used to show “that behavior of the included use case is inserted into the behavior of the including use case” [3]. Extend is a relationship “that specifies how and when the behavior defined in usually optional extending use case can be inserted into the behavior defined in the extended use case” [3].

The relationships include and extend are basically represented as a Z relation that relates two use cases. The relation include is used to record which use case is directly included by the original use case, whereas the relation extend is used to record which use case is directly extending the original use case.

The functions includedUsecases and extendingUsecases are respectively obtained with the transitive closure of include and extend. They are used to formally specify the function that returns the set of use cases that are directly or indirectly related with the relations include or extend.

\[
\begin{align*}
\text{include: } & \text{UCASE} \to \text{UCASE} \\
\text{extend: } & \text{UCASE} \to \text{UCASE} \\
\text{includedUsecases: } & \text{UCASE} \to \mathcal{P} \text{UCASE} \\
\text{extendingUsecases: } & \text{UCASE} \to \mathcal{P} \text{UCASE}
\end{align*}
\]

\[
\begin{align*}
\forall u1 : \text{UCASE} & \cdot \text{includedUsecases} u1 = \{ u2 : \text{UCASE} | (u1, u2) \in \text{include} ^{+} \} \\
\forall u1 : \text{UCASE} & \cdot \text{extendingUsecases} u1 = \{ u2 : \text{UCASE} | (u2, u1) \in \text{extend} ^{+} \}
\end{align*}
\]

The extension may occur at a specific point of the extended use case called extension point. An extension point is associated with a constraint indicating when the extension occurs. According to the UML standard, an extension point must have a name. Formally, an extension point is specified as follows:

\[
\begin{align*}
\text{ExtensionPoint} & \\
\text{name: } & \text{NAME} \\
\text{constraint: } & \mathcal{P} \text{BOOL} \\
\# & \text{constraint} \leq 1
\end{align*}
\]

If no constraint is associated, the extension is unconditional. In this case, the state variable constraint is an empty set and therefore its cardinality is equal to 0.

A use case generalization is a relationship from a child use case to a parent use case, specifying how a child can specialize all behavior and characteristics described for the parent. A use case generalization is translated into Z by adding, in the schema Usecase, the state variable parents defined as a set of use cases.

Once the related terminology of use cases has been clarified by means of specification written in Z, the notion of use case may thereafter be fully formalized.

Therefore, the notion of use case is clearly represented in the following schema:

\[
\begin{align*}
\text{Usecase} & \\
\text{parents: } & \mathcal{P} \text{UCASE}
\end{align*}
\]
The top half of the *Usecase* schema defines the different variables of a use case with their respective types. Each use case is identified by its identifier called *self* and its name. As explained before, the variable *parents* is used to define the set of use cases specialized by the current use case.

If a use case is extended, it is then composed of a sequence of fragments and an ordered list of extension points. An extending use case consists of one or more behavior fragments descriptions that are to be inserted into the appropriate spots of the extended use case. The extension points specify where the respective behavioral fragments of the extending use case are to be inserted [3]. If the condition is true when the first extension point is reached, then the extending use case will be executed. Otherwise the extension does not occur. If no constraint is associated with the extend relationship, the extension is unconditional.

As stated in the first predicate in the second half of the schema *Usecase*, a use case cannot include use cases that directly or indirectly include it.

After having defined the notion of use case and its related terminology, we are able to move to the next step, that of specifying actors and their relationships.

3.3. Actors Formalization. “An actor specifies a role played by a user or any other system that interacts with the subject” [3]. According to the UML standard, an actor is defined as a classifier. Actually the given set *CLASSIFIER* is used to define the set of all unique identifiers throughout the specification. The variable *ACTOR* is introduced as a classifier to specify the set of actors’ identifiers.

\[
\text{Actor}\text{TYPE} := \text{HUMAN} \mid \text{nonHUMAN}
\]

For actors, there are mainly two types of relationships: (1) a relationship between actor and use case and (2) a relationship between actors that is nothing other than generalization.

The relationship between actor and use case is basically represented as a Z relation that relates an actor to a use case. The relation *associationAU* is used to record which use case can be used by the actor, whereas the function *relatedUsecases* returns the set of use cases that are associated to the actor.

\[
\forall a: \text{ACTOR} \bullet \text{relatedUsecases} a = \{ uc: \text{Usecase} \mid (a, uc) \in \text{associationAU} \}
\]

Actor generalization is used to factor out and reuse similarities between actors. This relationship shows that one actor inherits the role and properties of another actor. To represent generalization, we chose to add a state variable to declaration part of the schema *Actor* as shown below.

This state variable is called *parents* and specified as a set of actors. The generalization relationship also implies that the descendant actor can use all the use cases that have been defined for its ancestor. This is clearly stated in the predicate part.
The function `getActorFromACTOR` takes an actor identifier and return the corresponding actor. This function will be used later in the use case diagram formalization.

\[
\text{getActorFromACTOR: ACTOR} \rightarrow \text{Actor}
\]

\[\forall a: \text{ACTOR}; ac: \text{Actor} \mid ac.\text{self} = a \implies \text{getActorFromACTOR } a = ac\]

Similarly, we define the function `getUsecaseFromUCASE`.

\[
\text{getUsecaseFromUCASE: UCASE} \rightarrow \text{Usecase}
\]

\[\forall UC: \text{UCASE}; uc: \text{Usecase} \mid uc.\text{self} = UC \implies \text{getUsecaseFromUCASE } UC = uc\]

### 3.4. Use Case Diagram Formalization

Based on the formalization of use cases, actors and their relationships, the concept of use case diagram can be fully defined. As shown in the following schema, a use case diagram has a name and it is composed of a set of actors and a set of use cases.

We also specify the scenario of execution of use cases in the case of extension. We propose the function `UCExecution` which takes a sequence of use cases fragments and extension points and returns a boolean type. This function indicates whether or not the extension will be executed.
The first predicate states that, for each actor, the inherited use cases are nothing other than related use cases of its parents. It is also important to note that use cases and actors identifiers are unique. The fourth predicate states that a use case cannot include use cases that directly or indirectly include it. The last predicate means that an extension occurs only if the constraint is true or if the extension is unconditional.

Now we concentrate on how to use the formal specification of use case diagram defined in this section to checking the multi-view consistency. In the following section, we discuss the consistency between use case diagram, class diagram and sequence diagram.

4. Multiview Consistency Checking

Having discussed how to construct a formal model of use case diagram using Z notation, the final section of this paper addresses ways of checking consistency in multi-view context. Multi-view modeling is a methodology for describing different aspects of the system in terms of different views or models. This methodology presents several challenges, particularly those related to consistency, which could potentially put the integrity of the system at risk. "In general, a system is consistent when there is logical coherence (i.e., no contradictions) between its parts" [23]. For these reasons, it is essential that all views are shown to be consistent and coherent between them.

Once the formal model of use case diagrams has been defined, it is useful to state and prove theorems between different views. This helps to verify the system and check for errors. In this paper, we focus on consistency of three different models of a system, comprising class diagram, sequence diagram, and use case diagram. Note that class diagram and sequence diagram formal models and their constraints are, respectively, defined in Z notation in [6, 7].

4.1. Consistency between Use Case Diagram and Sequence Diagram. To formalize the notion of consistency between use case diagram and sequence diagram we consider a Z theorem called consistencyUsecaseAndSequenceDiagram. This theorem checks whether

(i) each use case name corresponds to a frame of sequence diagram;
(ii) actors in use case diagram are defined as objects in sequence diagram.

\[
\text{theorem consistencyUsecaseAndSequenceDiagram} \\
\forall ud: useCaseDiagram; uc: Usecase; a: Actor \\
\quad | uc \in ud.usecases \land a \in ud.actors \land uc \in a.relatedUC \\
\quad \exists sd: SequenceDiagram \land sd.Frame = uc.name \land a.self \in sd.Objects
\]

Here, for any use case diagram \( ud \), a use case \( uc \) and an actor \( a \) such that \( uc \) and \( a \) belong, respectively, to \( ud \) use cases and actors and \( uc \) is in the set of use cases related to \( a \); then there exists a sequence diagram \( sd \) such that its frame corresponds to the name of \( uc \) and the identifier of \( a \) is in the set of \( sd \) objects.

If the theorem is proved to be true, then every constraint expressed before could be satisfied and consequently this would guarantee the system always remains in consistent state.

4.2. Consistency between Use Case Diagram and Class Diagram. Let us now define the notion of consistency between use case diagram and class diagram.

\[
\text{op: Usecase} \rightarrow \text{OP}
\]

\[
\text{theorem consistencyUsecaseAndClassDiagram} \\
\forall ud: useCaseDiagram; a: Actor; uc: Usecase \\
\quad | a \in ud.actors \land uc \in ud.usecases \\
\quad \exists \text{class: CLASS; operation: OP} \\
\quad \exists a.self \in \text{ObjectsOfClass class} \land (uc, operation) \in \text{op}
\]

First, given an instance of class diagram and a use case diagram, we define a Z theorem called consistencyUsecaseAndClassDiagram. The theorem claims that the two views are consistent, if we assume that

(i) Actors correspond to defined objects of classes in class diagram.
(ii) Use cases in use case diagram are assigned to operations of classes in class diagram.

We define the function \( \text{op} \) which, given a use case, returns the corresponding operation.
Here, for any use case diagram \( ud \), a use case \( uc \) and an actor \( a \) such that \( uc \) and \( a \) belong, respectively, to \( ud \) use cases and actors; then there exists a class named \( \text{class} \) such that \( a \) identifier is in the set of objects of \( \text{class} \) and \( uc \) corresponds to one of its operations.

The proposed theorem ensures that the specification satisfy the constraints presented before. If the constraints are satisfied then the specification is consistent. The theorem should be proved to be true.

For further information on class diagrams and sequence diagrams formalizations, the reader is referred to [6, 7].

5. Conclusions and Future Work

“The specification of requirements is a key activity for achieving the goals of any software project and it has long been established and recognized by researchers and practitioners” [24]. In this paper, we have considered a formal specification of the major concepts of use case diagram. The objectives of our work are providing concise and unambiguous specification of use case diagrams using Z notation and consistency checking by means of Z theorems. Formal specification of use case diagrams has been done in the past using different languages but we believe that the expressiveness and the mathematical basis of Z notation provide a richer framework for such formalism. The main motivation behind our research was the fact that the use of formal specification can greatly enhance the quality of the produced software as well as to contribute significantly to cost savings in the software development process.

One important field for further research is to find ways to handle larger views in order to detect more inconsistencies. We also plan to extend our prototype to include use case diagram formalization. We reiterate that our prototype automatically translates a subset of UML diagrams into a Z formal specification to uncover most of the UML inconsistencies published to date. Furthermore, we will make our approach more practical using some industrial applications examples. All the presented specifications were syntactically checked and proven to be true using the Z/EVES system [25] which is a tool for analysing Z specifications supporting type checking and theorems proving.

**Data Availability**

No data were used to support this study.

**Conflicts of Interest**

The authors declare that they have no conflicts of interest.

**References**


