

# 3DML: Exploring Microsoft Hololens for Interactive Visualization of UML Diagrams

Anders Mikkelsen  
Westerdals Oslo ACT, Oslo, Norway  
Mobile Technology Lab  
mikand@westerdals.no

Sondre Honningsøy  
Westerdals Oslo ACT, Oslo, Norway  
honson16@student.westerdals.no

Tor-Morten Grønli  
Westerdals Oslo ACT, Oslo, Norway  
Mobile Technology Lab  
tmg@westerdals.no

George Ghinea  
Westerdals Oslo ACT, Oslo, Norway  
Brunel University, Uxbridge, United Kingdom  
gg@westerdals.no

**Abstract**—Large diagrams of Unified Modeling Language (UML) are complex. In this paper we propose a novel approach to reducing this complexity by not focusing on UML itself but how we interact with it. We lift UML into an interactive 3D space and apply concepts and design guidelines of Natural User Interfaces (NUI). The goal is to present UML in a way that is easy to understand and that scales well regardless of the size and complexity of a software architecture. This research is aimed at a business landscape and collaboration between several architects. We start by describing the interactive system with a small introduction to the technologies in use and then look at the current state of UML and different interpretations of virtual realities. We follow up with our design rationale for the 3DML system and then evaluate it based on established design guidelines for NUI. We finally elaborate on valuable outcomes in the short and long term as well as a conclusion on the 3DML project in general.

## 1. Introduction

Unified Modeling Language (UML) is a modeling language used by software engineers to visualize the characteristics of system. The core features of UML is to provide a means to visualize elements such as activities (jobs), individual elements of a system and how they can interact with other system components, runtime behavior and so on. Over the years, system engineers are challenged by increasingly complex systems with a great demand for quality. As the complexity of UML increases when designing large scale systems, our ability to comprehend them is reduced (Singh and Sood, 2009). Issues with formal semantics, technical and visual complexity and adaption issues are all relatable issues that have been confirmed in recent research (Singh and Sood, 2009), Torchiano (2017), (Rodrigues et al., 2016), (Schattkowsky and Forster, 2007).

Based on these challenges, this research use features of emerging technologies for 3D visualizations, particularly

mixed reality (MR) to explore if there exist any visual benefits to support the understanding of software architecture models by developing the 3DML (software architecture visualization in MR 3D) approach. With 3DML, we aim to provide the user with a new, attractive experience to manipulate, analyse and view UML diagrams through multiple perspectives provided by the Microsoft Hololens.

### 1.1. Interactive System Description

The technical artifact from this research is 3DML, an approach that has been designed to provide viewing, creation and manipulation of UML objects in mixed reality (MR) 3D space. Our main goal with the prototype is to give users the ability to explore different perspectives, relationships and abstractions in a way that feels more intuitive, with the added goal of reducing the complexity of comprehensive UML diagrams. In contrast to traditional 2D UML tools, where the users have to interact with monolithic 2D UML diagrams, pairing UML with MR allows the user to exploit the depth in their field of view. The ability to use depth in UML diagrams allows for features such as collapsing and expanding UML elements that naturally belong together on demand. This will not only save physical space in such diagrams, but hopefully reduce the perceived complexity and issues of confusing semantics (Torchiano, 2017) as well.

#### 1.1.1. Microsoft Hololens.

The prototype is supported by the Hololens, which is a head-mounted display unit, also known as a pair of smart glasses developed and manufactured by Microsoft. Placed on the front of the glasses are most of the sensors and related hardware, including four infrared cameras, four microphones, light sensors and see-through holographic optics to name a few. In addition, the Hololens comes with the personal voice assistant, Cortana, and a manual clicker.

### 1.1.2. Microsoft Hololens - Manipulation techniques and gestures.

Hololens uses sensual and natural user interface commands such as gaze, gesture and voice inputs Microsoft (2017) for interaction. It has five preconfigured hand gestures that serve as the main gaze commands, as depicted in figure 1. Using head tracking, the user places a pointer over the objects or icons that he wants to interact with. These objects and icons can be "pinned" to specific locations, physical structures, objects, or "follow" in fixed relation to the users field of view as he moves around. In the same way you would use left-click on your mouse to select an icon, the "select" gesture 1 is the equivalent for the Hololens. Alternatively, a "clicker" is included so that the user does not have to rely entirely on hand or voice commands, but is slightly limited as you cannot use it to resize or reshape 3D objects in an efficient way.

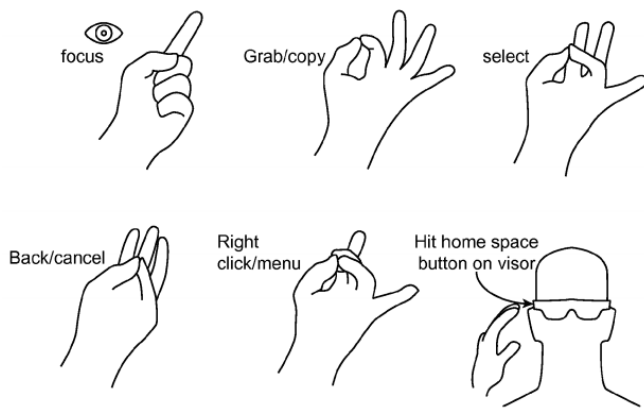


Figure 1. Hololens Gestures, Microsoft (2017)

### 1.2. Related Work

To the best of our knowledge there are no similar research projects that have attempted to implement UML, or similar notations in an mixed reality context. However, several approaches to UML diagrams in 3D exist, but none of them are well established in software design. Previous research and experiments has shown that environments that facilitate 3D visualizations of UML obtained positive evidence of its ability to improve the understanding of UML, such as the VisAr3D framework proposed by Rodrigues et al. (2016). The main difference with our approach is how users can interact with 3D UML representations in a MR space to provide a deeper understanding of the complex system models. When working in MR, and particularly with the Microsoft HoloLens, we also considered how the preconfigured hand gestures could impact the interactions with our proposed prototype. The idea is to have the user move, highlight, and manipulate 3D UML objects without using physical objects (other than the hand), which has been given the term "tangible MR" by Zhou et al. (2008).

The mixed reality domain has come a long way since Zhou et al. (2008) made their article on future research trends in AR from 2008. In some of their findings, hand gesture recognition was recognized as "one of the most natural ways to interact with an AR environment, particularly when facing partial occlusion of patterns". Now that AR has recently become more mainstream, Chang et al. (2017) took previous research on AR interaction a step further to explore several ways to manipulate and draw objects using hand gestures with the Hololens. Their experiment results indicate that for AR tasks that require accurate and beautified annotations, "surface drawing" yielded the best results. Surface drawing (SD) define the drawing position as the intersection between the detected surface mesh and cast from the user's head through the fingertip position. Consequently, as the user is drawing, moving of manipulating annotations, the user can easily verge on the object of interest since the annotation is displayed at the detected surface.

The UML notation has become widespread in both system design and business process modeling, and is often regarded as a convenient alternative for both activities Schattkowsky and Forster (2007). In model driven architecture (MDA) for example, the introduction of UML by the Object Management Group (OMG) had a major impact on modern software development Singh and Sood (2009), where modeling is the primary focus of the software development process. However, Singh and Sood (2009) argues that the approaches that support developers in providing enterprise centric solutions have been falling short in handling some of the most trivial issues like changing requirements, platform interoperability and that the technical complexity of UML is responsible for modeling adaption issues.

Further, he concludes that the UML standards are not up the level of present software industry requirements, and that "lack of adequate and suitable tool support could become an adoption barrier to MDA Singh and Sood (2009). Schattkowsky et al. mentions several semantic issues that he describes as "making UML more complicated than it needs to be", such as lack of formal semantics, excessive supply of contents, lack of methods for formal verification and a vast amount of alternative presentation options that describe the same properties in the model elements Schattkowsky and Forster (2007).

## 2. Interactive system: Design

3DML is designed to highlight the beneficial impact of planning a system in UML before implementing it, while at the same time keeping that process fun, engaging and interactive. The biggest detractor of fun and interaction is the ever increasing complexity of a system and the UML that describes it. This challenge is where the initial idea came from and the primary focus of the design in the first iteration is to reduce the complexity to a level that is acceptable across any size of UML.

The first fundamental step was deciding to go with a 3D environment. This is something that underpinned the initial idea, and we also found that reducing complexity in

UML by utilizing 3D environments was already established in research Rodrigues et al. (2016). With 3D you can expand on the understanding of the diagram by employ depth into the diagram as well as complex informative animations and other effects that are not as easily employed in a 2D environment Rodrigues et al. (2016). Even though we have established that 3D can lead to a better understanding we do not think this is enough. You should not just view 3D through a 2D screen, we want the user to be able to walk around and directly interact with the model. For that reason we decided to go for a virtualized interface. The two choices here are Augmented Reality (AR)/Mixed Reality (MR) and Virtual Reality (VR).

VR gives great possibilities in creating whatever space you think is necessary for this kind of visualization but this is intended as a business product and in that context you should be able to introduce real objects into the virtual space and vice versa. You should for example be able to use any room in your office space to do the work you need as well using documents and other real life objects during work. This places us firmly in MR territory.

To be able to do complicated tasks we want to combine gesture recognition, speech recognition and traditional input styles. Hence, we thought the idea of implementing it for the Microsoft HoloLens was a natural choice. Firstly because we have access to it, but it also supports all these input paradigms and has great connectivity. In addition, it has great developer support though Unity, Visual Studio and the community. Its greatest drawback is a relatively limited field of view, but we consider this a generational challenge which will be solved in later iterations.

### 3. Interactive System: Development

Due to a time constraint of around 1 1/2 weeks the initial prototype for the 3DML system is a static display constructed in Unity. We have however decided implement models in Unity in much the same way it would be done in the 3DML system. Elements are dynamically added as GameObjects which will have some sort of representative size and function, but is modifiable by the user.

We will be using screen captures of these static representations in section 6 to highlight the capabilities of a future concrete implementation.

### 4. Interactive System: Evaluation

As designers, our role is to facilitate the task at hand for the user and make sure that the user is able to make use of the product with as little effort as possible (Chadia Abras1, 2004). Graphical user interfaces are well researched and documented when it comes to formal best practices for design, interactions, spatial patterns and so on (Wigdor and Wixon, 2011). MR is arguably still in its infancy, hence few specific MR guidelines and best practices exist as to what is the best ways to design these types of applications (Zhou et al., 2008). Since MR is in the domain of natural

user interfaces (Zhou et al., 2008), 3DML follows 8 design principles from four of the chapters proposed by (Wigdor and Wixon, 2011) in their book “Brave NUI World”.

(Wigdor and Wixon, 2011) wrote in their introduction chapter about natural user interfaces that “you can forget past interaction styles. Don’t simply transcribe an application rendered in a traditional medium (web or GUI) as a NUI”. 3DML is no exception, which is why we have based our design on principles that concern gestures and its properties, natural user interfaces, and social environments rather than focusing on specific design elements alone. The design principles taken from (Wigdor and Wixon, 2011)’s book “Brave NUI World” are as follows:

#### 4.1. Guideline: The Natural User Interface

- Create an experience that is authentic to the medium — do not start by trying to mimic the real world or anything else
- Avoid falling into the trap of copying existing user interface paradigms

At its current state, 3DML relies on the preconfigured gestures and functionalities of the HoloLens. We did not find it feasible, nor necessary to come up with new ways to interact with the 3D holograms other than what comes “out of the box” with the HoloLens. With that said, we do not to our knowledge know of any similar applications that merge UML and MR together, meaning that we do not have much previous work to compare ourselves with other than traditional GUI based 2D UML applications.

#### 4.2. Ecological niche: Computing, the social environment, and ways of working

- Consider the context of use and the new possibilities that the interface brings to interaction in that context

With 3DML, we want to make the experience of creating, sharing, and viewing UML to be more exciting and intuitive. Some of the new possibilities that we are exploring with 3DML, that otherwise would not be possible, or at the very least limited on traditional 2D applications is the ability to directly interact with UML on the z-axis (depth) in a real physical environment. In a social context this could be very efficient during barnstorming and learning sessions, on-boarding of new employees or to simply make complex UML models more comprehensible to the viewers.

#### 4.3. Guideline: The anatomy of a gesture

- Minimize the load required to transition between registration and continuation phases. It should be immediately clear to the user how to specify the command she wants to execute.
- Provide clear feedback for the user at each step, ensuring she understands when she has transitioned from registration to continuation, and how to terminate the gesture

Since we rely on using the HoloLens to project 3D UML holograms, one of the prerequisites for our prototype is to have the user learn how to operate the HoloLens and learn every standardized gesture. The “cost of learning” these gestures is relatively small, as it is only a matter of learning what the different hand gestures and voice commands do (figure 1), which is manageable for most users. The transition between the different phases (registration and continuation) is handled by the HoloLens operating system, which is based on Windows 10. The 3DML holograms are encapsulated by an invisible “box” which is highlighted when the cursor (controlled by the users head motions) hovers over it. As an example, the user can select the hologram (registration), and the edges of the invisible box will be highlighted. Once the user moves the hologram (continuation) and “pins” it to another location in the room, the highlighted box will be invisible again (termination), making the hologram immersed in the room again.

#### 4.4. Properties of a gesture language

Identity of objects are permanent unless explicitly deleted, and an action on a given object in a given context always yields the same result. Further negation of any action can be reversed midcourse, and that reversal will return the system to its previous state. Lastly reciprocal must be in place once an action is completed, a side effect of that action can be undone by another action. For example, horizontal stretching of a graphic object will change its width and its aspect ratio. A subsequent vertical stretching of the same object will not undo the change in width but will restore the aspect ratio.

We identified the importance of these three properties in our design. Since our current prototype is limited to displaying pre-made static 3DML holograms because of time constraints, most of the gesture implications are already handled by the built-in features of the HoloLens. For example: when a user selects an object, the object will be highlighted (identity). If the user selected the wrong object, the user can get back to its previous state by using the “back/cancel” gesture (negation). When a user accidentally places a hologram on top of another hologram, the system responds with an appropriate error message so that the holograms do not intervene with each other (reciprocal). Given the nature of interaction with MR and HoloLens, the ability to build and create 3D UML diagrams from scratch requires precise considerations of all three properties.

#### 4.5. The Fundamentals Have to Work

A development team would be wise to start with an application that is both simple and different from existing applications on the GUI or the web. Simple games are often a good starting point for a team to learn how to design and implement a NUI. The mechanics must also be right first. If the mechanics of the interaction, the gestures with their feedback, are not fluid and enjoyable, the system will fail. This has been understood in game design for years. The Mario Brothers franchise would not have been as successful if the mechanics of jumping were not fun. It is wise to look at other NUI applications that are in the same domain. It will give the team insight on what works and does not.

Since we are dealing with a relatively new interface paradigm, there are a lot of unknowns as to what we should, or should not do. What may seem like a good idea today, may not be a good idea at all in the foreseeable future as the MR domain keeps evolving. Hence, some risks and “gut feeling” decisions have to be made. Since the availability of the HoloLens is currently limited to a small audience of early adaptors, our approach is to make the interaction similar to that of the standard and available applications on the HoloLens. This allows us to focus more on the implementation of 3D UML visualisations themselves.

### 5. Five Actionable Events

- 1) User locks onto an UML element
  - This will transition the interface to a different context for this user. Gestures with a certain action in a global context will have different uses here for a more fine grained control of the element in question. A contextual menu with clickable and speech activated options will be available. Methods and Fields for a class element should f.ex. have an option of being dictated to the system by the user.

- 2) An options menu is prompted after an UML element is chosen
  - If a user simply selects an element, and does not lock onto it, a contextual menu will be displayed with relevant information. For example you might get an option of doing relational connections to other close elements in a database modeling scenario.
- 3) Other team member pushes update to SCM (notification – do you want to pull it?)
  - Whenever a member of a distributed team of architects pushes an update to an ongoing design to source control, the other members should get a notification with a couple of options. Either merging the update directly, or being shown an overlay on the design that shows what the update is proposing, displaying any notes and comments the remote team member has written to detail a specific feature.
- 4) User does an illegal activity (example: drops one element directly into another element)
  - System responds with an appropriate error message, preferably by serving up some alternatives for the user.
  - System reverts to previous state.
- 5) After placing two or more UML elements: suggest an action to be performed for the user.
  - The system should routinely monitor the users actions to recommend appropriate actions for simplifying the process for the user. This is a configurable property to differentiate between users that want more automation or more manual control.

## 6. Result and Discussion

### 6.1. Simple MVC

Figures 2 and 3 display simplified examples of an Model-View-Controller (MVC) architecture. We can see that the individual elements can be viewed from both the front and the back, so not everyone has to use the same viewing angle, and the view is simply mirrored on the alternate side.

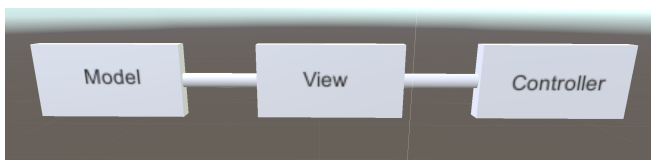


Figure 2. Simplified MVC Front



Figure 3. Simplified MVC Back

### 6.2. Simple Layers

Figures 5 and 4 display simplified examples of how we can present a layered architecture in a 3D environment.

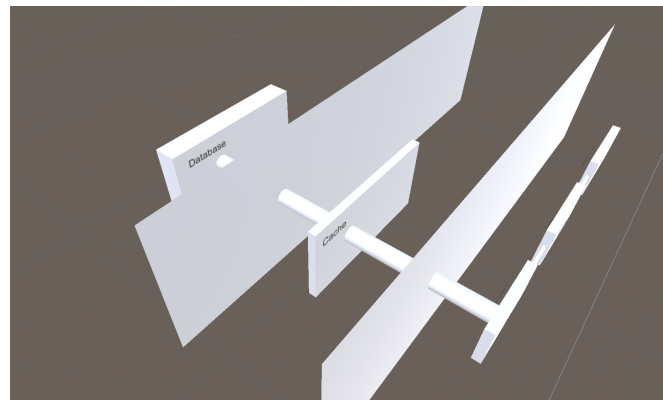


Figure 4. Layers Front

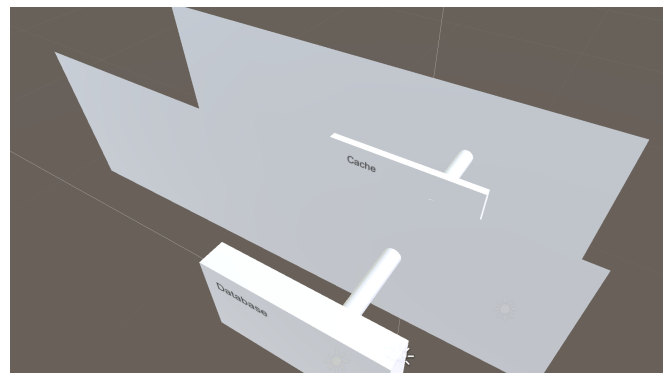


Figure 5. Layers Back

Looking at Figure 6 we see how this is already a much more natural representation than what you can see in 2D UML. We also see how we can easily show that only specific components communicate between the layers. That kind of complexity will increase in 2D but will still be direct and natural in 3DML.

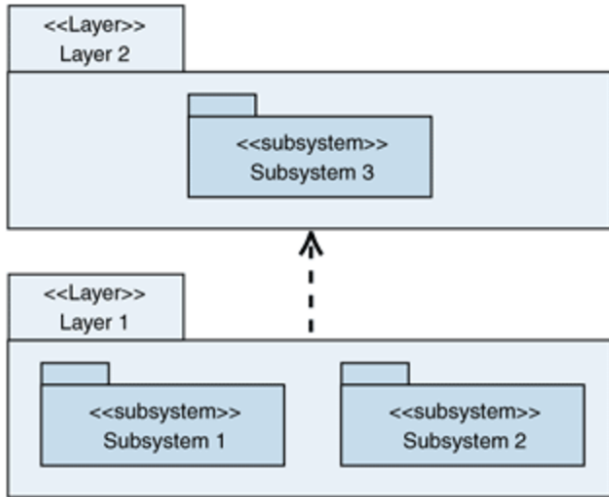


Figure 6. Layers 2D

### 6.3. Single Element Interaction

When locking onto or selecting an element you will be shown a contextual menu. This has not been constructed in Unity. Partly because we would need to decide all appropriate actions first and this should evolve during the course of detailing the design. It would however be loosely designed with the Hololens base menu as a reference so as to present the user with a familiar view to start with, as shown in Figure 7.

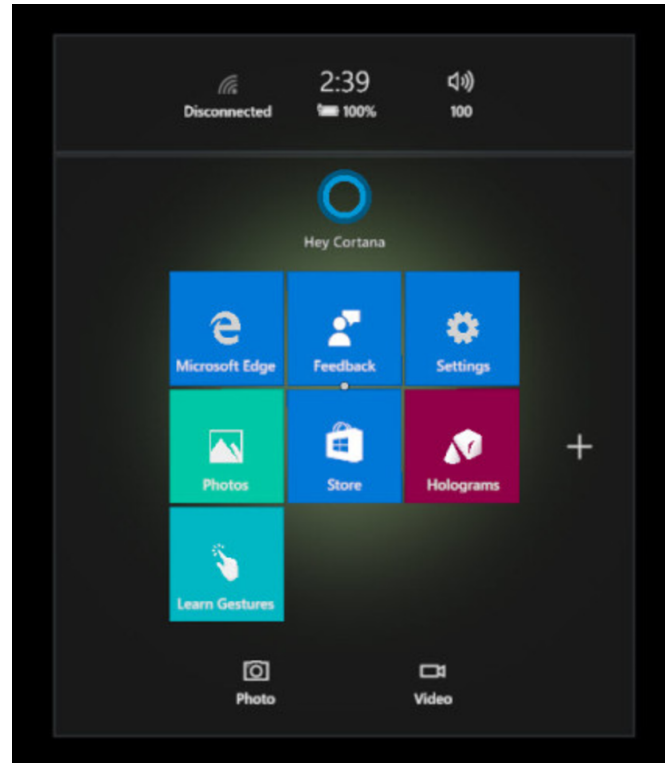


Figure 7. Hololens Menu

### 6.4. SCM Management

Figure 8 shows a simplified view of what a notification of a change in the projects Source Control Management (SCM) might look like. Some other architect has added a second level cache to the project and wants the chief architect to add it to the project UML. The chief architect will then have the option of rejecting or accepting it. An accept will produce what is seen in Figure 4 while a reject would result in what is seen in Figure 9.

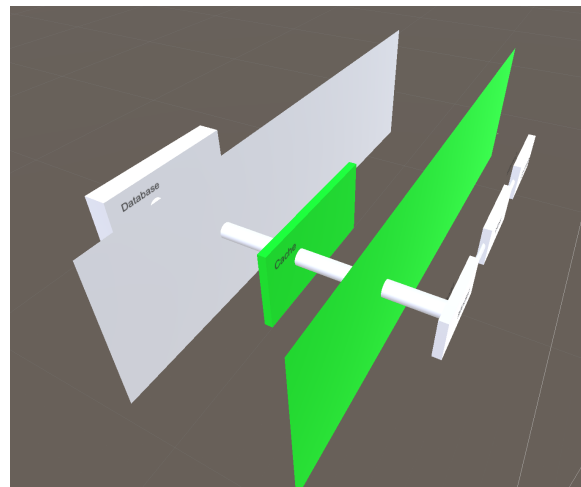


Figure 8. Cache Added SCM

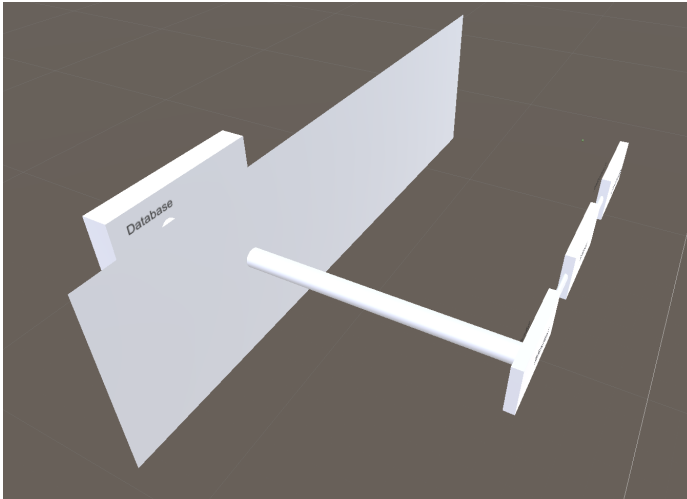


Figure 9. Cache Addition to SCM Rejected

### 6.5. Error in UML interaction

Figure 10 shows a simplified view of what would happen when a user attempts to place an element where it cannot be placed, in this example on top of another element. This would highlight the element red and then replace it to its previous location if it is pre existing (11) or simply remove it when it is an addition.

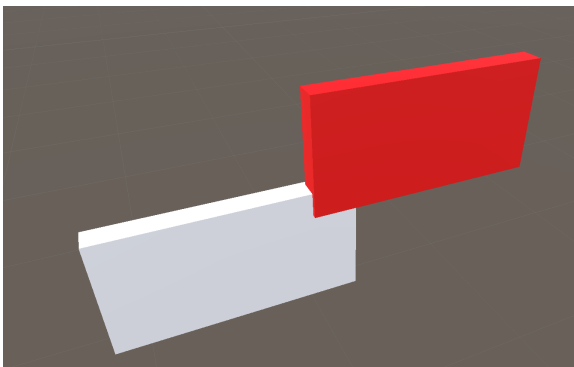


Figure 10. Red Highlight

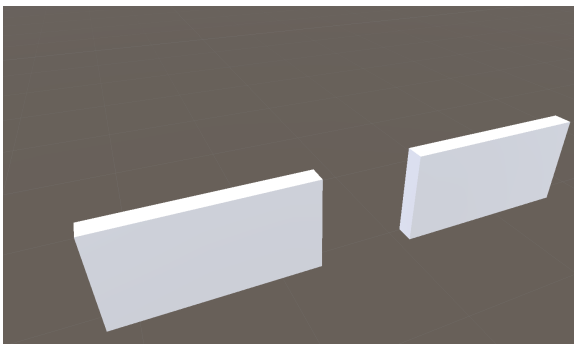


Figure 11. Element replaced

## 7. Two Valuable Outcomes

### 7.1. Short Term

3DML is a great collaborative platform for both small and large teams of architects. For a company implementing 3DML the short term benefits would primarily be related to visualizing the existing architecture by converting 2D diagrams to 3D and reorganizing them to maximize the value of this new way of understanding the UML. Instead of dealing with layered UML in 2D with static information, they will now be able to analyze an interactive model of their architecture with the ability to dive into any specifics. This will make issues with high coupling or low cohesion much more apparent Rodrigues et al. (2016). When the architecture has been converted to a 3D visualization and the senior members are proficient it will be a great tool for on-boarding new junior team members. Instead of trying to grasp a monolithic 2D UML hanging on a wall the architecture can be examined in detail with a 3DML model. You can go through it piece by piece and see its application flow in a much more intuitive way Rodrigues et al. (2016).

### 7.2. Long Term

The long term outcome of 3DML is making UML and planning architectures fun, interactive and intuitive. Multiple teams should be able to work on various parts of the greater architecture and senior architects should then be able to seamlessly merge these into a 3D model of all parts of their architecture. With a properly organized model you can scale the UML to almost any size, because its layout will be much more understandable than highly complex 2D UML. You can expand with depth, animations, collapsing figures, expanding with documentation in a separate browser within the MR environment, attach notes to components in the UML, fire up a virtual whiteboard to explain some detail etc. Future iterations could implement test visualizations directly in the model so you can run an integration test and see the entire flow as it spreads across the architecture.

## 8. Conclusion

This paper has reviewed challenges when using the Unified Modeling Language (UML) that has been recognized in recent research. Based on this research, we have proposed the 3DML approach, which provides the user with a tangible natural interface to interact with 3D representations of UML in a mixed reality environment. With 3DML, software architects no longer have to rely solely on static, monolithic 2D representations of UML, but can instead interact with UML that allows for analysis of interactive models of their architecture. The perceived complexity of UML is reduced as 3DML takes advantage of the Z-axis (depth) as illustrated in Figure ??.

## 8.1. Future Work

Future work will adopt 3DML in a lab setting to experiment with different types of 2D and 3D UML diagrams to get quantifiable results from different visualization techniques supported by the Hololens.

### 8.1.1. Future Opportunities.

Now that we have the system completely modeled in a connected and interactive 3D environment we should capitalize on that for daily work. We envision a future where operations can employ the 3DML system for viewing a system in realtime, based on the model. They can see the flow of traffic with metrics for individual components and immediately track the location of potential issues. They can then lock onto the element and pull up logs and other statistical information on that component. They can also control restarts, reprovisioning of resources and other resource management in a natural interface. This should be paired with strong speech recognition for solving complex tasks.

## References

- J. P. Chadia Abras<sup>1</sup>, Diane Maloney-Krichmar<sup>2</sup>. User-centered design. *Human-Computer Interaction. Thousand Oaks: Sage Publication*, 2004.
- Y. S. Chang, B. Nuernberger, B. Luan, and T. Hollerer. Evaluating gesture-based augmented reality annotation. *IEEE Symposium on 3D User Interfaces*, 2017.
- Microsoft. Development overview, 2017. URL [https://developer.microsoft.com/en-us/windows/mixed-reality/development\\_overview](https://developer.microsoft.com/en-us/windows/mixed-reality/development_overview).
- C. S. C. Rodrigues, C. M. L. Werner, and L. Landau. Visar3d. *Proceedings of the 38th International Conference on Software Engineering Companion - ICSE*, 2016.
- T. Schattkowsky and A. Forster. On the pitfalls of uml 2 activity modeling. *International Workshop on Modeling in Software Engineering (MISE'07: ICSE Workshop)*, 2007.
- Y. Singh and M. Sood. Model driven architecture: A perspective. *IEEE International Advance Computing Conference*, 2009.
- M. Torchiano. Empirical assessment of uml static object diagrams. *IEEE International Workshop on Program Comprehension*, 2017.
- D. Wigdor and D. Wixon. Brave nui world. *1st ed. Burlington, MA: Morgan Kaufmann*, 2011.
- F. Zhou, H. B.-L. Duh, and M. Billinghurst. Trends in augmented reality tracking, interaction and display: A review of ten years of ismar. *7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, 2008.