V4XR

Intelligent Verification/Validation for XR Based Systems

The iv4XR solution - exploitation and valorisation opportunities and corresponding business models

The rise of extended reality (XR) systems relies on the benefits they provide in both the digital and the real world. A notable benefit of these systems is that they allow users to interact in simulated environments of different domains - such as healthcare, education, entertainment, automotive and aviation. The development, verification and validation of XR systems is an ever increasing challenge because of the requirements for iterative development, the integration processes and the current high need for human testing. The increased requirement for quality assurance provides new opportunities for automation.

Today, testing of XR systems is mainly done manually. Testers are assigned tasks for testing within manually created scripts that are crafted to validate component parts of the XR system functionality. Testing therefore requires a lot of human time and effort - thereby making the testing process overly expensive. Moreover, this requires testers to interact with a huge interaction space whilst also looking to assure the user experience (UX).

The iv4XR project presents a toolset to improve the quality assurance of XR systems by using intelligent agents [1]. The iv4XR toolset enables automated testing of XR systems and evaluation of UX by using intelligent agents - Functional Test Agents (FTA's) and Socio-emotional Test Agents (SETA's).

This white paper presents a brief description of the problem being addressed by the iv4XR project and the solution proposed by iv4XR. Beyond this the project team have engaged with likely customer prospects with a view to establishing commercial exploitation routes and business benefits that may be exploited by prospect companies developing XR systems.

The Problem

The XR industry relies on iterative development methodologies that require a diverse toolset. Such toolsets must support multiple needs, including, design and authoring of the experience, development of the software and integration, and testing for quality assurance (QA).

The testing is particularly critical to assure both a high quality user experience (UX) and to drive the design and authoring throughout the iterative development process. Testing will often continue after a project is released, because updates, patches and improvements may be required. This is particularly common in server-based, online experiences such as multi-user games.

Good testing practices involve engaging users frequently, however, this procedure is costly and can be complex to manage [2]. In general, the toolsets available to XR developers currently lack structural testing technologies to support such practice because existing, automated QA techniques are unable to handle XRs' high interactivity and realistic environments.

This has been impeding the industry's growth and profitability and its ability to deliver timely products to market which may require sophisticated virtual and augmented environments alongside high quality user experiences. The iv4XR project's solutions in development should unlock many new commercial opportunities if we remove this impediment. So the clear goal of the project is to develop advanced and powerful testing tools that will 'power up' the toolsets available to the XR industry.

Overview of the solution

Approach

At its core, the iv4XR Framework is an agent-based system. A testing task is formulated, essentially, as a pair of goals and tactics for a test-agent to perform. The goal declaratively specifies what the task should accomplish, and the tactic is a program expressing a heuristic for accomplishing/solving the goal [3].

This is fundamentally different from the traditional way of testing, in which a testing task is formulated as a step-by-step program formulating the sequence of steps that are involved in conducting the test.

When testing a highly interactive system like an XR system, the sequence would consist of many interactions. A test that runs for 30 seconds can easily consist of over 1000 interactions. For example and in contrast, tests on web or applications rarely exceed 50 interactions. XR systems also often behave non-deterministically or even adversarially (computer games and simulators are typical examples).

The figure below shows the top-level architecture of the iv4XR Framework and its typical workflow.

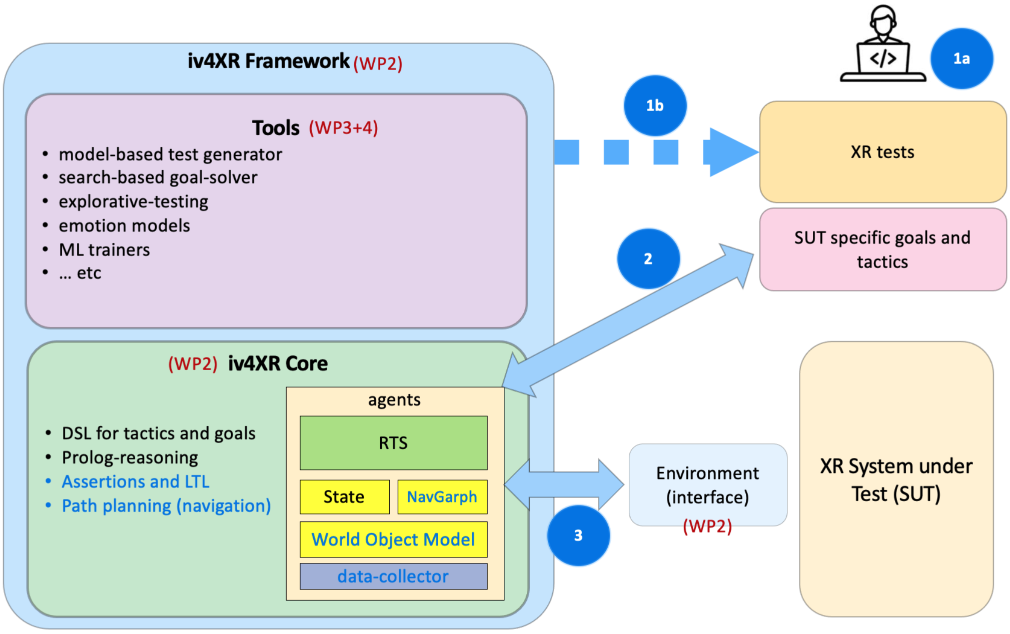

Fig.1: Top-level architecture of the iv4XR framework. WP refers to workpackages on the project

It is important to note that iv4XR strives to remain neutral towards the XR technology used in the system to be tested (System Under Test or SUT). However, this neutrality means that developers must first implement an interface (called Environment in Figure 1) between the system they wish to test and the iv4XR agents which control or work within the SUT.

Additionally, a library of domain specific abstract/high-level actions need to be defined and implemented. An example of such an action might be to toggle an in-world switch. Depending on the SUT, this action may first require the agent to approach the switch until it is close enough to it, and only then it can toggle the switch. Such knowledge must come from the SUT domain itself, as there is no way the agent could otherwise know this.

Accordingly, a programmer (or several) will be required to write an appropriate interface and any scripts which will need to be run to start test agents on the SUT. We refer to these programmers here as Test Engineers.

Solution application

Generally speaking, testing regimes may be described as Functional Testing or Non-Functional Testing. Functional testing is usually carried out first and focuses on functional requirements to validate the behavior of the application and verifies that each function of the SUT operates in conformance with its specification. Non-Functional Testing, on the other hand, is more associated with checking software performance, user user experience and usability and, particularly in the case of games, entertainment and fun value. The iv4XR agent-based approach should accommodate both forms of test regime [4] [5] [6] [7], as the Socio-Emotional Test Agents under development in WP4 can be used for certain non-functional tests [8].

In Figure 2 we show how iv4XR test agents fit within the test cycle of a typical development which, here, is based on a multiplayer online game (Space Engineers) which is frequently updated with new features and virtual environments. Even though we explain it for a specific case, the diagram is generic and can be applied to several other cases.

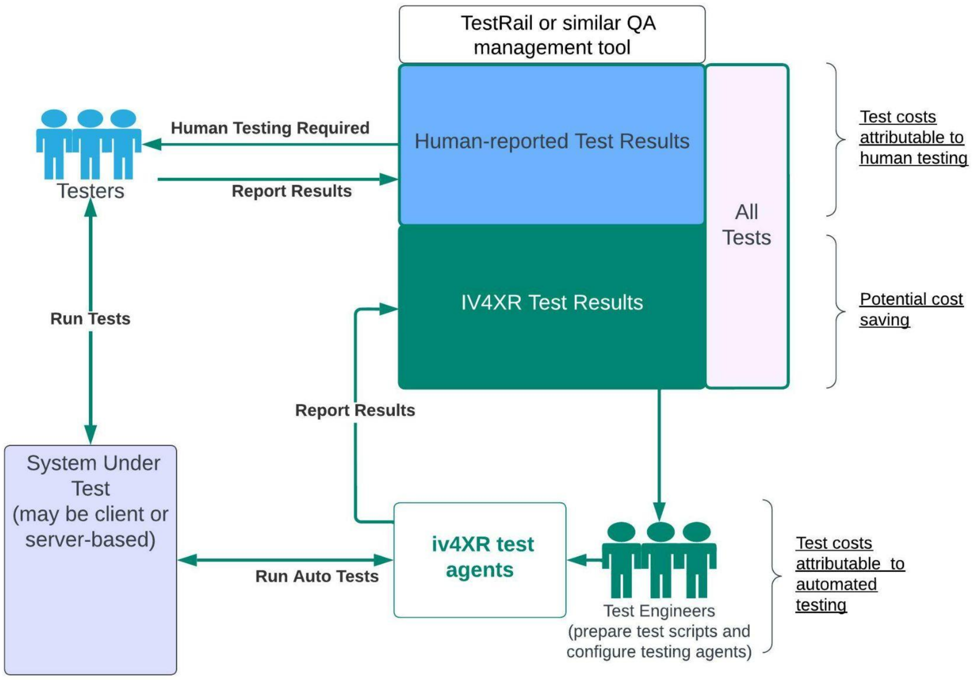

Fig.2: Example application of iv4XR framework.

As is shown in Fig. 2, a list of all bugs, problems and issues is stored in a QA management tool (essentially a custom database) such as TestRail or Jira. Some tasks require human testing (the blue area) and other tests are considered suitable for automated testing (the green area). Test specifications are prepared by the Test Engineers. It is possible for many thousands of such tests to be performed in the time it would take human testers to perform a few hundred. Both manual and automated test results are reported back to the QA database for the developers to address. As revisions to the code are made, the process is repeated iteratively until the team lead is satisfied that the code is suitable for release.

Note that once the iv4XR test agent specifications have been prepared, there is little to no further requirement on the Test Engineers and so their cost can be considered as a 'once-off' cost. However, each test iteration which requires human testing will consume more resource costs. The same is true, indeed, for subsequent code updates and releases.

The potential for cost-saving is thus likely to be significant and may be substantial over a product's lifetime. The commercial challenge which iv4XR will face is in identifying which industries and, within these, which types of application development may yield the best opportunities to exploit such savings.

During the last few months, we have studied several software enterprise areas where we believe the automated testing approach offered by iv4XR could be most appropriate. In the next section we look at some early results from this research.

Business Opportunity

Software development is increasingly focused on XR technologies in several industries. We consider the following sectors to be of special interest for unlocking the commercial potential of iv4XR post-project outputs.

Automotive and Aviation

In the automotive industry, companies like Mercedes and BMW have been investing in Augmented Reality to boost navigation by improving accuracy and ease of use in navigation systems. For BMW, the front view camera sends live footage to a curved touchscreen where interactive arrows show the driver exactly where they need to go. And in Germany, they have been testing a system that helps drivers find parking spots in busy car parks.

But AR in the automotive industry goes beyond improving the navigation experience. It can significantly improve security by helping the driver perceive the environment better, for example, by highlighting pedestrians and other vehicles. These systems need a high level of accuracy. Testing procedures should be critical to ensure that the system is working to improve safety without being distracting or providing inaccurate information that could have adverse effects.

Another industry that has shown a growing interest in XR technologies is the aviation industry. For the aviation XR weekend, the International Air Transport Association (IATA) showed some examples of companies using XR technologies. These examples included training of ramp staff (Qatar airways), inflight entertainment (Iberia and Sri Lanka airline), and an AR bag sizer for a mobile app (LATAM and KLM). Projection of data information and design details to help construction and inspection (Airbus) and Qantas has designed their latest lounge using VR.

In the aviation industry, flight simulators have been used for several years to provide training for pilots. From Aviation Training Devices that train basic skills to full flight simulators that allow pilots to train emergency procedures without actually having to experience a malfunction or accident. The same occurs for healthcare applications where physicians can train surgery skills and how to respond to unexpected but possible problems without risking someone's life.

Education and Training

Aside from training simulations, XR technologies are seen as something that can revolutionize education, especially for younger children. Many studies and projects have focused on XR technologies to improve learning. The underlying belief is that XR technologies can provide more immersive learning experiences and improve children's engagement in class. It also allows students to experience things they may not have been able to experience before. For example, it may be difficult for many children in different countries to take a field trip to the Louvre, but a virtual tour of the museum using VR technology might allow them to have the next best experience. XR technologies can make learning more fun, and fun enhances learning.

These simulations and other education environments pose challenges to designers and developers. For example, simulations for training should be as complete as possible and allow users to train situations they might encounter. The system needs to react appropriately; if not, the simulation is missing stimuli that will impact the user's reaction if they encounter a similar situation in real life.

Gaming and Entertainment

Finally, the gaming and entertainment industries are, in many ways, the main engines of XR technology development, as many of these technologies started in those sectors. In fact, other industries sometimes use game engines like Unreal to create XR experiences. The gaming and entertainment industries do not have the same challenges and potential 'life or death' consequences as the other industries we presented above. But because they have been using and developing these technologies longer and at a speedier rate, they have encountered challenges that the other industries do not have.

The gaming industry deals with increasingly complex environments where sometimes entire worlds are simulated, and users have demanding expectations for these environments. For example, they expect the possibility of interacting with other users without it affecting the game's performance. They expect the system to save changes they made to the environment or game progress. Besides, entertainment and games are very fast-paced industries. A slow, unengaging product may mean that the users will not buy the product or stop subscribing, which significantly impacts the companies. So games and entertainment are industries where people want constant novelty with a great user experience. This constant demand for novelty and good user experience means that testing the systems is critical, and the faster and more accurate it is, the better.

To learn more about current industries using XR technologies, we approached companies in 7 different industries to ask about their use of XR technologies and the cost of testing those technologies (see Figure 3).

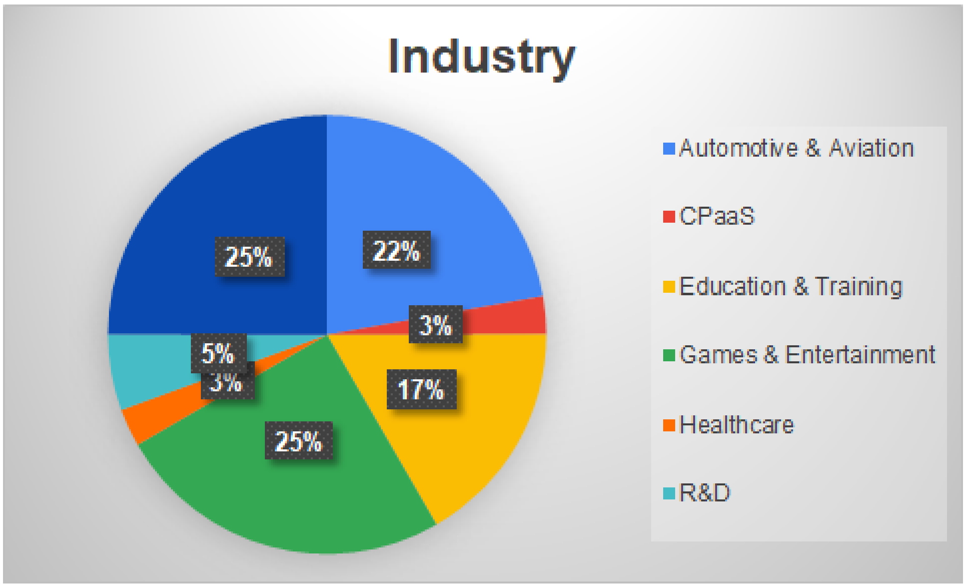

Fig.3: Surveyed Industries using XR technology

The companies were mostly micro companies (30.6%) or large companies (33.3%), and the majority of the companies developing their own systems, reported spending 10 to 20% of their budget on testing (Figure 4).

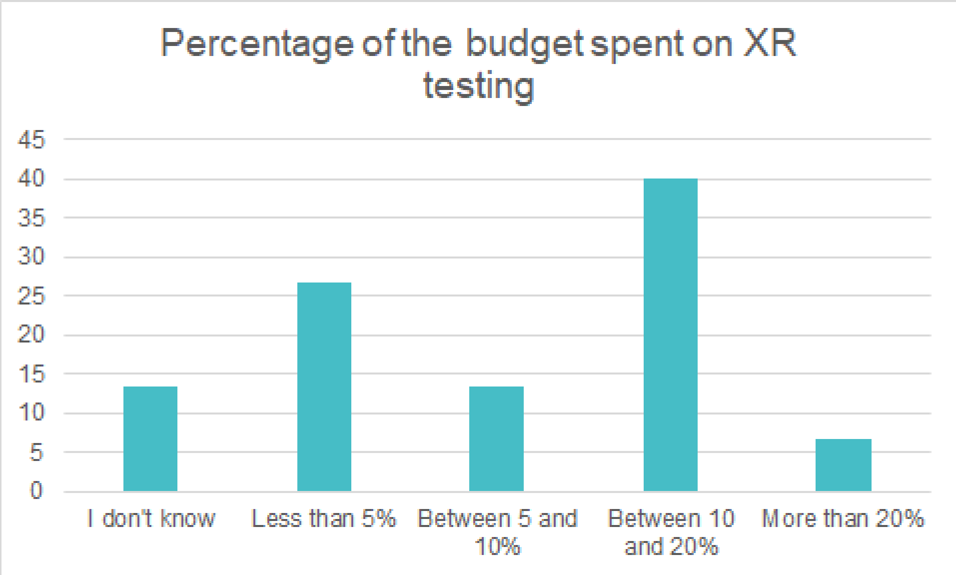

Fig.4: Percentage of software development cost spent on testing

We also asked the companies why they were using or developing XR technologies. The replies were diverse and were naturally dependent on the industry but many companies reported using XR technologies for the audio and visual features. Some were taking advantage of those features to improve teaching, including offering a more immersive experience to trainees. Others are using it to let clients visualise what the final product will look like. Games and entertainment companies are also using XR to improve immersion and offer more immersive experiences.

The iv4XR solution has the potential to reduce time and costs, since it automates a percentage of the tests. Companies could reduce the percentage of budget allocated for testing and redirect that money to other endeavors. Our framework also goes beyond testing functionality and aims to automate user experience testing, providing developers with a set of measures that indicate the quality of user experience without having to find a representative sample of end users to test a prototype.

Conclusion

We have established that there is significant interest in advanced automated XR testing tools to support labour-intensive manual testing. Our findings indicate that the key markets for automated XR testing are as follows: Automotive, Aviation, Education, Training simulations, Gaming and Entertainment. We also identify some potentially addressable markets at the leading edge of industries such as Construction, Healthcare and R&D. Here we have outlined the main reasons why the iv4XR solution is most applicable to these markets.

Whilst a number of commercial models are under consideration, digital distribution via engaged user portals/app stores such as Unity platform for Game Developers and digital creatives is high on the agenda for commercial review and consideration post project completion.

References

[1] Prada, R., Prasetya, I. S. W. B., Kifetew, F., Dignum, F., Vos, T. E., Lander, J., Donnart, J., Kazmierowski, A., Davidson, J., & Fernandes, P. M. (2020, October). Agent-based Testing of Extended Reality Systems. In *2020 IEEE 13th International Conference on Software Testing, Validation and Verification (ICST)* (pp. 414-417). IEEE.
[2] Alves, R., Valente, P., & Nunes, N. J. (2014, October). The state of user experience evaluation practice. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational* (pp. 93-102).
[3] Prasetya, I. S. W. B., Dastani, M., Prada, R., Vos, T. E., Dignum, F., & Kifetew, F. (2020, June). Aplib: Tactical agents for testing computer games. In *International Workshop on Engineering Multi-Agent Systems* (pp. 21-41). Springer, Cham.
[4] Vos, T. E., Aho, P., Pastor Ricos, F., Rodriguez-Valdes, O., & Mulders, A. (2021). testar–scriptless testing through graphical user interface. *Software Testing, Verification and Reliability*, *31*(3), e1771.
[5] Ferdous, R., Kifetew, F., Prandi, D., Prasetya, I. S. W. B., Shirzadehhajimahmood, S., & Susi, A. (2021, October). Search-Based Automated Play Testing of Computer Games: A Model-Based Approach. In *International Symposium on Search Based Software Engineering* (pp. 56-71). Springer, Cham.
[6] Shirzadehhajimahmood, S., Prasetya, I. S. W. B., Dignum, F., Dastani, M., & Keller, G. (2021, August). Using an agent-based approach for robust automated testing of computer games. In *Proceedings of the 12th International Workshop on Automating TEST Case Design, Selection, and Evaluation* (pp. 1-8).
[7] Prasetya, I. S. W. B., Shirzadehhajimahmood, S., Ansari, S. G., Fernandes, P., & Prada, R. (2021, April). An Agent-based Architecture for AI-Enhanced Automated Testing for XR Systems, a Short Paper. In *2021 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)* (pp. 213-217). IEEE.
[8] Fernandes, P. M., Jørgensen, J., & Poldervaart, N. N. (2021, December). Adapting Procedural Content Generation to Player Personas Through Evolution. In *2021 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 01-09). IEEE.

Further Reading

To learn more about the project and receive the latest updates, check out the website and follow us on Twitter and Facebook:

https://iv4XR-project.eu

@iv4XR

https://www.facebook.com/iv4XR

For our scientific publication check Zenodo:

https://zenodo.org/communities/iv4XR-project/?page=1&size=20 (check zenodo for sci pub)

Other resources:

Automotive and XR technologies https://europe.autonews.com/

International Air Transport Association (IATA) IATA - Home
Flight simulators simulator AVIATION NEWS International Aviation News, Airshow reports, Aircraft facts, worlds largest Aviation Museum database. Civillian, Military & Space, We cover it All