WebXR, Connecting Realities

Fabien Bénétou

We are using the web hundreds of times per day. Usually we use it explicitly, opening our mailbox for example, but even more often unknowingly, using an "app" on our smartphone that is in fact relying entirely on web technologies (HTTP allowing traffic to flow, WebRTC for real time communication, HTML5 to structure documents, CSS to visually present documents). Seemingly far away from those banal activities we see increasingly often totally new interfaces like augmented reality (AR) and virtual reality (VR). Those interfaces have been gaining popularity during the last few years thanks to the significant price drop of sensors itself leading to the distribution of better and cheaper head mounted displays for mass market. The distance between the web and those new interfaces is shrinking thanks to the efforts of the lead browser vendors including Google Chrome, Mozilla Firefox, Microsoft Edge or even the new Samsung Internet Browser. AR and VR are indeed slowly but surely arriving on the web. In fact the role of the browser started displaying a simple page with solely text and links first. Shortly after images appeared and finally within the last few years 3D content. Moving from 3D to VR happened without much fireworks since 2016 at Mozilla then Google with basic experimentation. Since and thanks to the collaboration of all those actors but also of content producers or distributor the web could shift to become the first distribution channel of all immersive content: AR, VR, MR (Mixed Reality)... to XR today. XR recently became the umbrella abbreviation for all immersive technologies. We have AR, MR and VR today but one can imagine all those different "realities" as a spectrum. Consequently XR is {X}R where X is a variable which can be {A, M, V} but also anything that could appear in the future. Even more interestingly X can be the interplay of those different "realities" as we will explore here.

The World-Wide-Web, or more casually the web, is a well defined technological construct. It is measurable in term of size, electrical requirements, number of views per time period, etc. Yet, to simplify, we can start from the bare minimum: 2 pages with a link. This very small and simple construct is the foundation for the web as we know it today.

This deceptively simple construct comes with an important set of properties: each page can be stored physically anywhere as long as they are on the same network; the owner of each page can be different; there is no need to ask permission to link to another page; the format of the page is written in an open format allowing anybody to build and provide tools to create, modify, host a page; last but not least as it leads to permissionless innovation, you can create and share your page to your network, potentially the world today, without having to ask any authorisation first.

Those properties might seem very basic but they allow for quite unique behaviors: the network of pages can grow without a predefined structure; there is no required authority indicating which page is good, bad or even relevant; the set of tools to interact with the network can freely grow in complexity as long a consensus remains; the topology of the underlying network infrastructure can and does change constantly.

It is precisely those properties that allow the web to grow and transformed it to this platform so central to our daily lives. Consequently, the underlying technologies, even though not perfect, are important. The determining factor though, was the ability to grow based on the efforts of content creators and their desire to connect to other content.

This set of desire and efforts remain regardless of the media. Again, the web started as text, then adopted images, animations, videos but

is now gradually adopting XR. Those new media consequently automatically benefit from the properties of the web that allowed to grow so vast, so quickly.

Growth itself is interesting but not necessarily associated with value or usefulness. There might be a vast amount of content with very low diversity or quality. Without taking strong assumptions we can rely on usage as a metric of usefulness, if the tool is used we can assume it is useful at least in some ways to its users. This perspective is analogous to how we perceive our environment. We do not perceive through our imperfect senses an objective description of the world around us but rather an interface of the properties of the world based on their relative usefulness to us as an active agent of this world. We do not perceive objects but affordances to task relative to a goal.[1] We behave day to day as signal processing organisms avoiding threats and securing resources needed to maintain the homeostatis required for our survival. This perspective has the quite profound implication of transforming everything around us, not just man-made tools, into interfaces for our goals. It also simplifies the way we can think of new media like virtual reality, augmented reality, mixed reality or any other engineered reality. Those engineered realities are intermediary interfaces that, when successful, allow us to reach our goals more efficiently. We can conceptualize XR as pre-digested perceptions relative to a goal. The intelligence in the interface, like for any other piece of software, literally comes from the knowledge of the task made explicit during its development process.

Consequently we can analyze the future of engineered realities through the filter of usefulness relative to a goal. What the web allows in this context is an efficient way to create, test, share then curate those engineered realities. What it also allows, as we were able to connect the first 2 pages of the web together is to connect two engineered realities regardless of their owner. This might seem trivial but this constitutes the distinctive advantage of WebXR, making the web the platform to enable then efficiently share future thoughts.

What is already currently possible is for a content creator to build a virtual world, for instance a simulation of the social structure during medieval time in continental Europe. In itself being able to build a world and let anybody experience it instantly is incredibly powerful and, until tried once, arguably impossible to fully comprehend. It is important here to distinguish experience from simple consommation. Experiencing immersive content means being part of it, feeling as if we are there. This specifically allows us to embody another role, our environment affording a totally new set of actions otherwise impossible. What is also possible is for somebody else entirely unknown from the content creator to build another virtual world allowing to explore the social structure during the same time period in China. What is unique to the web though is the ability for the second creator to then link back to the first virtual world. This action affords entirely new insights, namely the ability to compare, itself allowing to identify commonalities and differences. Beyond that comes the ability to step back and to identify a pattern, are those commonalities and differences applicable to other objects within the same group? If we create a third virtual world from another culture, will it exhibit all those properties; some of them, are they predictable and if so, why? This again is what is making engineered realities on the web so unique, they are becoming manipulable objects that can now be created, shared but also remixed, linked and then indexed.

The ability to conceive realities as nearly trivial manipulable object is very powerful in the sense that what is arguably the most personal to us, our worldviews, can then be more efficiently changed and, ideally, improved. Even though as claimed before we only see the world around us as interfaces, the quality of those interfaces directly impact our agency and thus our well being. If we have incorrect interfaces our efficiency and thus our livelihood might be at risk. Consequently being able to efficiently manipulate engineered realities is not only possible but very valuable. Consequently here it is important to understand XR not just as a lowly tool that will be

misused for better advertisement, even though it surely will, but rather also as a tool for experimental philosophy. Indeed ideas about what reality is have been a common thread of most main philosophical efforts: questioning what reality is; if reality is atteignable via our senses only or via our intellect and finally if reality was atteignable at all. Here XR can help to falsify century old hypothesis and isolate new concepts so dear to our selves and our position in the world e.g. consciousness or our self-model theory.[2]

This example showcased how multiple virtual reality experiences can be linked today but it is already possible to do more: in addition to going from a virtual world to another, list them, etc., it is possible to bring content from a virtual world to another. Let's imagine that in the first example of a virtual world in continental Europe during medieval time, one gets randomly attributed resources that will impact its social status. One can then imagine being able to bring those to the virtual world representing China at the time. Consequently it is possible to create a causal link between actions in one world and another. Those virtual worlds become de facto causally linked.

Another possibility is to link the virtual reality example to an augmented reality example. Say you are visiting a castle in the French countryside after having visited the virtual world on the same topic. You can now use your AR device to see superimposed on the physical environment around you. This time, though based on the status you previously had according to your allocated resources, the behavior of virtual characters will be adapted. If you had a very low social status most places will be banned or at least with a thorough unpleasant set of questions and verification.

Those links and their ability to make the experience continue from one virtual world to another or to an augmented reality overlay is not trivial. Those links allow an isolated experience to permeate within a larger context.

How the social structure in medieval time changes one's worldview might not be obvious (even if resources are allocated randomly the extremely high chance of being of lower status would most likely make one truly appreciate their current life situation) but one can imagine quite a few other experiences from the mundane to the rare awe evoking moments. What the web once again allows is indexing (or more casually bookmarking) which is itself a both personal and thoughtful process. By indexing engineered realities that were indeed useful to me I can create a curated selection of experiences that can be lived again in the appropriate context. Curating useful engineered realities is already becoming a new skill but it is creating a meta-interface. Our curated list of useful engineered realities is basically becoming a proxy for our constantly evolving worldview. One can even imagine sharing bookmarked engineered realities as an explicit way to provide to someone else a glimpse through the set of interfaces used to navigate and interact with the world.

To conclude, the web is not just yet another distribution channel for new immersive content of AR or VR. The web is already the most efficient way to build test and share content. It allows and fosters permissionless innovations that are so crucial for an emerging new medium. The web allows to link content and to connect content creators and consumers in a way that no other solutions allows in XR. The web will consequently be the medium thanks to which the interconnection of XR content will let blossom a future we can barely imagine today.

Finally, the ongoing discussion on the WebXR specification and the implementation in all majors browsers in 2018 make it possible to create and use a link in immersive content. This very link becomes then a primitive allowing us to make engineered realities manipulable objects, affordances to think beyond what we know today.

[1] D. HOFFMAN, M. SINGH, "Natural selection and shape perception: Shape as an effective code for fitness", in S. DICKINSON, Z. PIZLO (eds.), *Shape Perception in Human and Computer Vision: An Interdisciplinary Perspective*, New York, Springer, 2013, 171-185.
[2] T. METZINGER, O. BLANKE, "Full-body illusions and minimal phenomenal selfhood", *Trends in Cognitive Sciences* 13(1), 2009, 7-13.