|  |  |
| --- | --- |
| Immagine | Moving Picture, Audio and Data Coding  by Artificial Intelligence  www.mpai.community |

**Public document**

|  |  |
| --- | --- |
| **N1451** | 2023/10/24 |
| **Source** | Requirements (SPG) |
| **Title** | MPAI-SPG Status report |
| **Target** | MPAI Members |

# Progresses

No updates to report on the Pong Use Case (UC)

We performed an experimental study with 14 users to evaluate the early results of the implemented solution, quantitatively and qualitatively. In the experiment users are asked to play several runs of the Racing Game against 4 self-driving cars. In each run the MPAI-SPG server-side prediction can be activated with different levels of intensity (L0, L1, L2). Each intensity level is defined by an amplitude (i.e., the duration of the prediction) and frequency (i.e., how often the prediction is performed on a given client). At the end of each run users complete a questionnaire to assess their perception of the quality of the game and its graphical fidelity. To quantitatively evaluate the quality of the predictions we record for each play run data from the car driven by the user and its predicted version. The main results are as follows.

From a qualitative perspective the results are positive as we can see from the below graphs. The first shows a general positive appreciation for the game experience and the input responsivity regardless of the presence/absence of the predictions (L1 vs L2 and L3) and among different intensity and frequency of the predictions (L2 vs L3).

A graph of different colored bars

Description automatically generated

In the following graph we can see the evaluation provided by users on several graphical artefacts visible thought the game play. Overall, the presence of the predictions caused some visible glitches, however all results fall below the score of 3.

A graph of different colored bars

Description automatically generated

Finally, in the below graph we can see several graphs, each showing the discrepancy between a predicted car and one driven (either by the human player or by an AI). The three dimensions on which the discrepancy is evaluated are: position, rotation, and velocity. As a general trend we can see that the error increases over time, meaning that the longer must be the length of the prediction, the greater the error. However, when the prediction time falls within 0.5 seconds such error can be manageable. What most importantly emerges from the graphs is the difference in results between the car driven by a human (blue) and the one driven by an AI agent (green). This mean that our predictive network is showing a domain shift problem as the network has been training on synthetic data.

A graph of a graph

Description automatically generated with medium confidence

Based on these initial findings we can conclude that the first experiment yields overall positive results. Showing that from a player’s perspective the activation of the MPAI predictions were detected by users in the form of glitches and other graphical artifacts. However, the overall game experience and input responsivity were positively evaluated despite the activation of the car’s predicted states on both the L2 and L3 levels. In future iterations the overall quality and reliability of the predictions will require improvement. Most importantly the similarity between the predictions associated with the cars driven by real human players should be closer to the one associated with the AI agent based cars.

We released an updated version of the document Technical Report: Mitigation of data loss and cheating in online multiplayer gaming WD0.1.2, where we outlined the main steps of the process which led to the conclusion described in this document.

Our main thesis student working on the project has concluded his Master’s Thesis. However, a new student has just signed up.

# Future Plans

* Make progress in the Technical Report document.
* Organize the new student work and establish key goals oriented to the improvement of the current implemented solution.