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

**Public document**

|  |  |
| --- | --- |
| **N658** | 2022/04/20 |
| **Source** | Requirements (SPG) |
| **Title** | MPAI-SPG Status report  |
| **Target** | MPAI Members |

# Project status

The month was devoted to test all the components of the project. Here you can find the project figures and states for the components of MPAI-SPG

1. Refine the SPG Digital Twin implementation inside the Online version of Pong, switching it from synchronous to asynchronous definition
2. Create a bigger dataset for Neural Networks training
3. Study trained LSTM Neural Networks behaviour inside the Online Pong to find a suitable configuration.



Figure 1 - Pong game



Figure 2 - MPAI-SPG Architecture

*LEGENDA*:

**Green**: new improvements

**Red**: TODO

**Black**: general status of the project

Both clients and server use Unity 3D as core game engine.

Features created so far:

* **Client**:
	+ Process C can send controller data to Process S
	+ Defined a CSV template to make the client log file
	+ Process C can obtain the ownership of one paddle sending an explicit request to the Process S
	+ Process C can send notifies to Process S through the so-called “RPCs” (Remote Procedure Calls), to specify some actions or some sort of communication
	+ Process C implements Client-side Prediction related to his paddle
	+ Process C implements interpolation (paddle position fixed with Lerp function)
	+ Process C implements server reconciliation (tick number for each data request)
* **Server:**
	+ Data exchange explanation between Game State Engine and game engines (Physics, Rule and Behavior)
	+ Defined a CSV template to make the server log file (game messages and game states)
	+ Game Server use the Photon architecture as space to instantiate each game
	+ Process S is a service program, running in “Batch/Headless mode” as Unity instance
	+ Process S can receive data from process C and acknowledge receipt of this data
	+ Process S instantiate both paddles and ball, then send (through RPCs) their ID to the clients so that they have a reference to those object
	+ Process S can handle the data (CD) sent by the process C to update the GS
	+ Process S is able to manage the physics both of the ball and the client paddles, sending the resulting data (GS) to both clients
	+ Process S is able to synch both ball and client paddles
	+ Better management of lag compensation
	+ Process S has
		- Game State Engine (GSE)
		- Physics Engine (PE)
		- Behaviour Engine (BE)
		- Rules Engine (RE)
	+ Game State Engine can send data to other three engines (PE/BE/RE)
	+ The three engines process the data correctly and send data to GSE
* **AI:**
* ML Agents developed, to simulate games to train the Neural Networks, using two techniques:
	+ - Imitation Learning
		- Reinforcement Learning
* Imitation learning Idea:
	+ - Build a Demo with human input (Teacher)
		- Define config.yaml with hyperparameters specifics (GAIL = GAN like learning)
		- Start learning. Agent acts as student and learns how to behave as similar as his Teacher. Time required is significantly shorter than other methods (in terms of minutes not hours)
		- NN produced will be the brain of our Client in Pong Game Online
* Integration of “ml-agents” framework inside Pong online version. This way will be possible to move paddle in an automatic way using inference
* Game simulation tests produce coherent data (log file) usable in any Neural Network
* Reinforcement Learning Idea:
	+ - A set of observed data is chosen (player position, opine position, ball velocity…)
		- Using its NN under development and observed data as input, the player executes an action and gets a reward. The reward is positive if it is correct in order to win the game, negative if is wrong.
* After multiple matches executed, paddle will be able to hit the ball
* Learning optimization (both Imitation and reinforcement learning) in terms of:
	+ - Physics accuracy
		- Hyperparameters selection
* Introduction of a raycasting system to improve Reinforcement Learning:
	+ - When the ball collides with a paddle, a series of rays is drawn on the game field representing ball trajectory ending on the opponent’s side of the field. This way the opponent will know where the ball will go and will been able to execute the correct action
* Introduction of a refined architecture (see the following chapter) that assigns “Divider” and “Composer” roles to Game State Engine AI
* Choice of the “Long Short-Term Memory” Neural Network Architecture to develop the 3 Simple Engines inside the MPAI-SPG Architecture (Physics Engine, Rules Engine, Behavior Engine) and we started the first training of them, using a small dataset of real game states taken from a Pong simulation
	+ - Made a comparison with the MLP (MultiLayer Perceptron) to analyze performances
* Creation of the Prediction System and application on the Pong Offline version:
	+ - Integration using Barracuda Framework
* Creation of Collector and Dispatcher for input and prediction data management
* Design of hardware requirements and availability at the end of December 2021 of the virtual architecture to create data and to train neural networks
* Received the hardware resources of Computer Science Department, University of Torino (CSD-UniTo)
* Collection of logs using CSD-UniTo hardware:
	+ - Using the trained AI Pong paddles, able to automatically play Pong, we collected around 15 million game states
* Neural Networks training:
	+ - A first set of training experiments was made to find a suitable predictive architecture using the 15 million record dataset generated at the previous step
		- The experiments were executed testing different kinds of architectures and changing network hyper parameters, like learning rate, batch size, optimizers, number of nodes inside LSTM layers
* Prediction System integration into Pong Online version:
	+ - To make inference using trained neural networks, a set of scripts representing the “refined architecture” was implemented using the Unity Barracuda Framework. In this way we were able to use Physics, Behavior and Rules Predicted engine inside the Pong Online game
		- This system is primarily used to test the trained neural networks previously described.
* Tests about different kinds of neural network architectures to find a setup able to correctly predict game states
	+ - Different hyperparameters were studied, like number of nodes, number of layers, optimizer, dropout values to cite some of them
		- Initially working with NNs with 500 or 1000 nodes for layers, we created an architecture able to make predict the game state with 50 nodes per layer
* Switch of the SPG prediction system implementation from synchronous to asynchronous:
	+ - Studying the server application behaviour, we found out some freezes and deadlocks that couldn’t permit the correct operation of the server. We found out that the Barracuda inference method takes some hundredths of a second to be completed. Switching the system into an asynchronous computation enables the inference method to be executed on another thread, avoiding graphic rendering freezes during server application execution. Note that this kind of approach executes the inference operation at a lower rate (about 12 frames/sec).
		- Studying the behaviour of the NNs created before this change, we noted that they were not very precise inside their predictions, forecasting ball’s position far from the correct one. This happened because the inference method was executed at a lower rate than the one used to collect the training data
* We computed an average measure of the time needed by the inference method to be executed
* Change inside the data collection scripts in Pong Offline to execute a game state sampling at the same rate of the inference execution calculated at the previous step
* Generation of a new training set from Pong Offline using this new script implementation
* Neural Network training that uses this new dataset
* Study of the Neural Networks behaviour inside Pong Online:
	+ - Neural Networks behave quite good with correct input data, while they are still not quite precise to execute predictions based on its old predictions as inputs
		- Neural Networks with a different number of timesteps in input were studied, noting that, increasing this quantity, the Barracuda inference method will require increasing time to be completed
* We have a set of 3 predictive Long-Short Term Neural Networks trained with about 28 million of records, corresponding to 16500 Pong scores
* These networks can forecast the next game state at time N+1 given a predefined number of states up to time N
* In case of packet latency or loss, SPG starts and uses the past game states received from the online game server as input to forecast the next game state. This means that it can forecast, in an approximate way, the next game state, given what happened before.
* Evaluate the prediction quality with an objective measure considering:

Neural networks configured and trained with various hyperparameter values:

1. NN4: 3 LSTM layers, each one of them containing 50 nodes, alternated with 3 Dropout layers, with a final dense layer that collects results and returns a vector representing the game state prediction. This network takes 20 timesteps as input and has been trained with 2,5 million records, equivalent to about 2400 Pong rallies.
2. NN6: equal to NN4 but considering 40 timesteps in input.
3. NN9: equal to NN4 but trained with 20 million records, which corresponds to about 16500 rallies.
4. NN10: equal to NN4 but trained through 51 million records, corresponding to 31000 rallies.

The best results in terms of loss and accuracy were obtained with NN4 and NN9, both using a reasonable number of timesteps; moreover, NN4 requires less training data.

* Replacement of Photon Lag Simulation GUI with an external program that manages the lag simulation: this step is necessary to overcome the limitation of Photon lag simulator which can be set only manually and not by scripting
* Analysis and feature modification of Pong Server to create the reference game in a deterministic way: this will be used with the set of played games of the demo with different network conditions and MPAI-SPG activations
* PLANS:
* Modification of Pong Server source code to replicate a reference game in a deterministic way
* Setup of the demo with the following schedule:
	+ 1. ***Reference game*** without network issues and SPG reference
		2. Reference game with a network packet (Controller Data) loss, introduced at second #n without SPG inference
		3. Reference game with a network packet (Controller Data) loss, introduced at second #n with SPG inference
		4. comparison of b) vs a), c) vs a) and b) vs c), qualitatively (game simulation) and quantitatively (Mean Squared Error, Mean Absolute Error)