Human-Like AI in RTS Games

EVALUATION DOCUMENT

CMP401.2016-7.S1 - Feasibility Demo

Duncan Bunting - 1302739

Word Key / Word Context

Player(s) - refers to both human and artificial intelligence in a game.
User(s) - refers to human only players

Introduction

Although this project's end goal is to achieve a better user experience in a real-time strategy game, it is extremely hard to measure enjoyment properly when it comes to evaluating a human playing against an AI, especially when the human does not know what is going on behind the scenes or how the AI is thinking in detail. In addition to this, players enjoyment can often be mostly influenced by the look and style and smoothness of a game, this is not the research area of this honours project and in addition, non-gamers and non-RTS fans may be asked to play the game, pre-existing opinions on gaming and skill difference between test subjects will drastically corrupt results if simply asked to rate their experiences. I have therefore decided not to host a simple questionnaire instead to have an in-game test represented by a version of the Turing Test suitable for my RTS environment. In addition to this test, general feedback will be collected in the form or written responses rather than a rating system. Finally, some non-human data will be collected in form of the information collected when pitting AI's against each other (or against humans) to see how they performed in comparison to each other.

Evaluation Methods

The Turing Test

This section is outlined in the game design document, and at this moment is in a very early concept form. It is currently undecided in which form the test will be presented, a few options exist:

1. Simulate searching for a multiplayer match with other players and have the user play against who they will, hopefully, believe are human-players before revealing at the end of the game that they played against AIs.
2. Tell the user that they will be taking part in The Turing Test, have the user connect to a lobby with one other human and several AI's. It is then up to the user to find out which of the players is human.
3. Perform the Turing Test as mentioned in option 2, without telling the user that they are taking part in the test and then get reaction from the user on if they suspected anything or if they knew which player was human.

At the moment the most likely approach to be taken is option two, as that properly reflects the Turing Test, although a decision may not be made until the base game has been fully completed and a better idea of the game's play-style is established. It is also a possibility that multiple of the options may be tested, depending on the amount of test subjects available.

User Feedback

As previously mentioned simply asking the user to rate their experience I am afraid will be too broad of a question and may be intercepted in the wrong way. Instead, after a user tests the game and the AI's a face to face discussion will be had, if face to face is not possible (ie I decided to allow people to take part remotely, or I am unavailable) a series of questions will be presented to the user and they will be asked to input their thoughts. All feedback will be recorded in writing and relevant information will be collected from the feedback given. The feedback questions will be in relation to the AI and not of the actual game and having users express themselves freely opposed to rating on a scale I feel will be a better way to collected only the relevant information. Questions may be presented after facing each AI and then again at the end of the play-test to find users opinion across all AIs.

Example Questions:

1. How did you find the game experience in general?
2. Do you have any thoughts on the AI?
3. How did you find this AI compared to the others?
4. Do you believe the actions performed by the AI were human-like?
5. Were there any areas that you thought were explicitly unhuman-like? (see below)

In addition to these questions and the others that may be asked during the testing, a tool may be implemented into the game to allow the user to make note that something happened at this time that was notable. This will allow me to go back through the data collected around that time and figure out how the AI came to make this unhuman-like decision.

Performance Data

As briefly outlined in the game design document, the game will record a large number of stats at set intervals during the game, as the games pace is unclear at the moment it is hard to say at what interval relevant information will be able to be established. The intervals will likely be every 5 or 10 seconds, recording information from the what planets are owned, to the reputation of each player and their opinions on other players. In addition to general stats at interviews, it is a hope to also record major events within the game which may provide a clearer outlook on the overall flow of the game and most important help record why the AI decided to do certain things.

In addition to this recording of game statistics, throughout the games length an output will be generated and, hopefully, be able to be viewed live as the game progresses. This output will display vital information on why decisions were made, at vital parts in the code which will end up making the decisions of the AI there will be code to output the variables that went into make these decisions and these will be outputted.

An example of the output:

<00:18:27:02> [Player 2] [opinion of] [Player 3] | 20 > -35 (-55) | [Player 3 Attacked Ally]

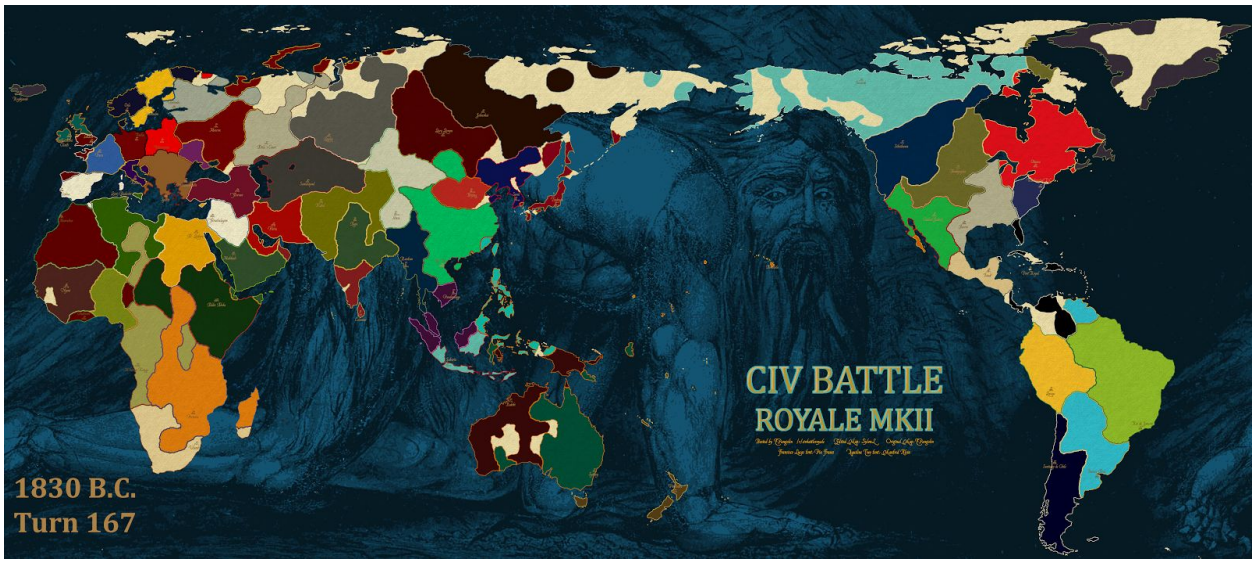
This can be broken down to the core information that will be important to the data collection:

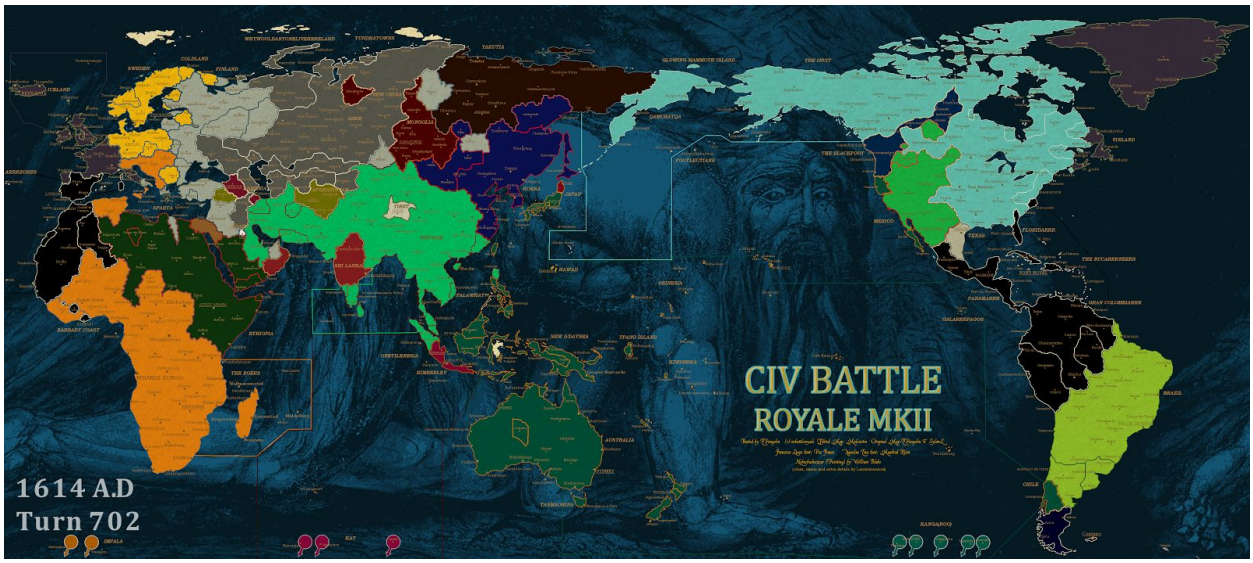
<hh:mm:ss:msms> [Player Affected][Stat Changing][Extra Info - ie who else it affects]
| [Previous Value] > [New Value] (Difference) |[Reason For Action]

This is merely a concept of how the information will be displayed, it is likely that the output will be less user-friendly as I will be the only one handling the data and equally as likely that more information than what is above will be displayed, this will be dependant entirely on the complexity of the AI and what is possible to be outputted and make reasonable sense.

Battle Royale

Finally, after completing all the AI's intended to be created, I will simulate a large scale game between all the AIs developed (and possible different instances of each technique with slightly adjusted algorithms), and idea inspired by the large scale AI battle even in Sid Meier's Civilization V (Firaxis Games 2010) dubbed the "The Civilization Battle Royale" where 60+ AIs compete against each other over a very long timescale. This simulation will be performed multiple times to ensure that data is accurate and a fair and even playing field will be created to ensure accurate results.





Summary

In conclusion I have came to the decision that a ranking system is not optimal for the type of interactions users may experience when testing the game's AI for feedback purposes. The amount of information needed to be collected can range from person to person and peoples pre-existing opinions, skill and general understanding may corrupt results if taking in a ranking based feedback form. Therefore, feedback from test subjects will be acquired via open ended discussion which will hopefully display similarities between people which can help judge how the AI performed against humans in general. In addition to the human testing, a more detailed statistic gathering will take place during game run-time to acquire all the necessary information to figure out why the AI makes the decisions it makes, and allow discussion of human-like behaviour, or lack of. Although the key of this project is to study and record the human interaction with the AI, the statistics behind each game will also provide a window into the world in-which the AI thinks. Finally, although the goal is not to create an skillful or unbeatable AI, several instances of each AI technique will be pitted against each other to figure out who performs best in this game environment and matches will be recorded and analysed to figure out exactly why.

References

Civilization V. 2010. [computer game]. PC. Firaxis Games.