Believable Interactions Between Synthetic Characters

Nuno Costa
Instituto Superior Técnico
May 2015

Abstract

The human player is often required to interact and cooperate with synthetic characters, which also cooperate and interact with each other. However, unless the action is tightly linear and scripted, the expression of interaction is often confusing and difficult to understand by the human player. This work explores how traditional animation principles can be applied to the expression of interaction between the actors, both synthetic and human, and make the communication and cooperation more believable and the experience more immersive. To validate our work, we implemented our model in a multiplayer sports-game, where each character is an artificial player, and asked participants to evaluate videos of the interactions. The data we collected suggests that our approach not only significantly improves believability, but also makes the interactions between agents easier to understand and the action easier to interpret.

Keywords: believability, agent interaction, videogames, intentions, anticipation.

1 Introduction

High profile games keep pushing the boundaries of their supporting hardware as far as visual fidelity goes. Their strive for believable environments and for photo-realistic presentation is relentless. It is hard not to be amazed when playing a game such as Call of Duty Ghosts1. Its frenetic campaign features impressive, over-the-top scenes on par with Michael Bay’s2 most famous setpieces! The player has the strong illusion of cooperation with his synthetic team-mates, and the enemies seem very clever. However, this approach works because the action is very linear and highly scripted. Generally, in open-world games that feature companions that interact with the human controlled characters - such as in Skyrim3 - such attention to detail is usually lacking, and often we have trouble cooperating with these synthetic characters. As a result, the battles - which are critical aspects of these action-adventure games - come out as overly simplistic and mechanical.

1.1 Motivation

If we observe great action flicks we observe a whole level of interaction and cooperation between characters in the action setpieces. The final stand in Spielberg’s Saving The Private Ryan4 is a great example. It is as if the action scene is a story within a bigger story that is the movie, in such a way that the viewer can describe it as a relevant part of the drama. As such, the status quo regarding interaction and cooperation between synthetic characters in battles in adventure video-games is clearly hindering their approach for a cinematic experience.

Action/adventure isn’t the only genre where believable communication and cooperation is needed. In sports games such as the Fifa5 series, players often complain about moments in the game where poor behaviour by the AI breaks the immersion and frustrates the player. This is particularly true in the Career Mode, where the player is given control of a single character, and must work is way up to the starting eleven of his club and the national team. Unlike the other game modes where the player is given control of the whole team, alternating between each player depending on the circumstances - such as distance from the ball, and always controlling the player of his team in possession - in this game mode the player has to rely heavily on his teammates judgement. In the real world football, there is constant communication happening: The coach corrects his pupils positioning, the play-

---

2Michael Bay - American film director best known for high-budget action films such as the Transformers (2007-ongoing) film series
ers shout and gesticulate to each other to ask for the ball, a run from a teammate, to express their intentions, and to warn each other about the presence of a dangerous adversary or play. In Fifa 14 however, the player can request the ball, request a shot at goal, or to pressure the opponents in which the teammates blindly and unrealistically oblige, but cannot express his intents and neither can his synthetic teammates. This form of indirect control over the teammates is, in our perspective, not good enough. Moreover, if a player chooses to let his teammates decide by themselves, it is anyone’s guess what’s on their minds, alienating the player from a useful role in the context of the team, once again detracting from the experience and frustrating the player.

1.2 Our proposal

Drawing inspiration from the fundamental principles of traditional animation[6] that infuse believability in fictional characters, we believe that, by separating an action in two stages - anticipation, execution and execution - we can improve the cooperation and coordination between agents and the human player in a believable way. To achieve this goal, in each action the agent should first express his intent (anticipation stage), then execute the action (execution stage). As a result, the human player will be aware of the agent’s intentions, minimizing frustration and providing efficient coordination while keeping the player immersed in the experience.

2 Related Work

2.1 Principles of Traditional Animation

According to John Lasseter[6] the understanding of fundamental principles of traditional animation is essential to produce great animation, both traditional animation or computer animation. These principles were first introduced by the Disney animators Frank Thomas and Ollie Johnston in their book The Illusion of Life: Disney Animation[10], and are based on the early standardized practices followed by the animators working at Disney. These principles allowed for more believable and lifelike animation.

The principle which will serve as groundwork for our model is the Anticipation principle. Anticipation is the preparation for the action. Anticipation serves as a declaration of intent. It "warns" the audience for the upcoming action, having it expect it before it actually occurs. When an animated character is about to throw a stone, it usually rotates his hip, leans back and draws his arm backward before actually throwing the stone. Without anticipation the action would seem unnatural. Anticipation also leads the audience attention to certain artifacts, so if a character stares happy at a particular item, the audience will also focus on that item because there is an expectation that the character will interact with it.

2.2 Emotions and believability

Taking inspiration from traditional animation to improve agent’s believability is not something new. In his paper[1] The Role of Emotion in Believable Agents, Bates argues that artists, not scientists, have come closest to capture the qualities that are able to make a creature appear alive. One quality that is particularly important is the ability to express emotion, which, in traditional animation, according to Thomas and Johnston, is what gives Disney’s animated characters the illusion of life[10]. Bates further argues that an emotionless character is merely a machine, and that the agent should express itself in a way that makes us believe that it has fears and desires, and that it cares about the world around it. In Eurico Doirado’s creation DogMate[3] the dog companion communicates with the human player by expressing emotions such as anger, fear or joy.

2.3 Cooperative agents

There has been progress as far as believability goes regarding cooperation with synthetic companions. Eurico Doirado investigated how believability could be enhanced by anticipating the human player behaviour[3]. He used an affective anticipatory mechanism called Emotivector[7] that made possible to infer, based on distance and velocity variations between the human player and the other entities (such as enemies, objectives...) surrounding him, his underlying intent. He implemented this model in a synthetic sidekick - a virtual dog named Rusty - that, with this information, was able to react accordingly and in a believable way based on the sensations it experienced, for example, it would express his worry if he predicted the player was going to engage an enemy if the player’s character had low health. Conversely, if Rusty believed the human player was able to take down an enemy and he predicted confrontation (based on distance reduction between the human player and the synthetic enemy) he would get exited. His reactions in the different scenarios would serve as an advice to the human player. The result of this study was positive, yielding 61.11% intent recognition.

Jonathan Tremblay and Clark Verbrugge investigated the field of adaptive companions applied to first person shooter games genre[11]. They used
Behaviour Trees to model a sidekick that cooperates with the player in a believable way, learning and adapting itself based on the player’s experience, and shaping its own behaviour. They start by making a clear and important distinction between two kinds of companions: The **fully autonomous companions**, which are independent companions over whom the player has no direct control over. The authors use Skyrim as an example of a game which implements this kind of companion, which often breaks the suspension of disbelief by walking in front of the player in fights, engaging in combat when the player intends to sneak, etc. Ideally, the fully autonomous companion should help the player by using a matching or complementary approach to battle that would potentiate its strength and the player’s, but open games often struggle to achieve this. And the **semi-autonomous companions**, which allow the player some degree of indirect control over the companions, such as in games where the player can give orders to the companions (which usually they obey blindly), or select the companions stance such as aggressive, or defensive, as in Dragon Age.

The model developed for the adaptive companion attempted maximized player experience by optimizing game intensity. Game intensity pacing is crucial to a good game experience[2], as excessively long periods of high or low intensity tend to respectively frustrate and bore players[9]. Using this notion, the model measures the current intensity of the game (mainly based on variation of player health), and select the optimal companion behaviour from a pool of three distinct options - Cautious, support or aggressive.

Besides game intensity, other metrics were developed to measure the performance of the companion: The personal space, and combat load ratio metrics. The personal space metric is helpful at determining if the companion is so close to the player that gets in the way of his movement and fire and thus breaking the suspension of disbelief, and the combat load ratio, which determines who’s fought more, the player or the companion.

They proceeded to develop a third-person shooter prototype in which the goal of the player was to collect blue boxes. There are several enemies traversing a pre-defined path that try to prevent the player from achieving his goal by attacking him, and his companion, if they are seen. The enemies will also chase them to their last known position if they run away. The companion will try to defend the player by helping him in combat. To test, the authors created three scenarios in the prototype and the role of the human player was filled by a basic AI several times, emulating different players with different skill. The results showed the adaptive companion to have a positive impact on game intensity levels, specially during intensity ”spikes”, hence, theoretically improving the game experience. The companion also spent less time in the player’s line of sight and in his personal space displaying more believable cooperation.

### 2.4 Communication between and with agents

In order to create believable behaviour, the creator must strive for more than just believable cooperation. The way the agent expresses itself and communicates with others is equally important. In multiplayer games with other human players, specially in games where cooperating and coordination between players is crucial, the player can usually use natural language, either by typing messages or through speech, so communication is usually not an issue in this scenario. However, as Gorniak and Roy point out[5], this same level of cooperation and coordination may be equally needed in games that feature synthetic characters that team up, and, in this case, communicating in natural language is very hard due to noise and also the semantic and syntactic complexity, so cooperation is often achieved through less desirable point and click interfaces. Gorniak and Roy attempted to overcome this barrier by contextualizing speech with the action, i.e., depending on where the speaker is located, her surroundings and the timing of speech and other aspects, the speech input is disambiguated. While the speech recognizer achieves a 50% word error rate it was augmented to deliver possible alternatives from each word spoken which are then shown in order of decreasing probability. The parser will then consider each alternative. By identifying the speaker intention, the speech recognizer errors are not as important as they would be otherwise, making this a robust approach in the subject.

### 2.5 User Testing

In order to evaluate our approach, however, user validation is required in some form of quantifiable metrics. Paulo Gomes, Ana Paiva, Carlos Martinho and Arnav Jhala developed a set of metrics with the goal of measuring perceived believability[4]. These metrics, which the authors call believability dimensions, are the following:

- **behaviour coherence:** The audience will evaluate the behaviour of the agent regarding coherence, which is a key aspect of believability[8].

- **change with experience:** This dimension quantifies how the character is influenced by story events.
• **awareness:** The audience should perceive an agent to be aware of the world around it.

• **behaviour understandability:** The participant should be able to understand the character’s behaviour. The agent must express itself in a way that its motivations and thoughts are clear to the spectator.

• **personality:** The agent should be perceived as an individual. Its behaviour should suggest personality traits that make it unique.

• **emotional expressiveness:** The agent should express its emotions.

• **social:** The social relationships between the agents must be clear to the audience.

• **visual impact:** The agent should draw the attention of the participant.

• **predictability:** A very predictable agent will harm believability just as much as a completely unpredictable agent, as it affects behaviour coherence[8], so either one of these extremes should be avoided.

The audience perception is assessed using Likert scales, using one scale per dimension. The templates for the phrases (except emotional expressiveness) to be rated by the subjects are:

- **awareness:** \(< X >\) perceives the world around him/her.

- **behaviour understandability:** It is easy to understand what \(< X >\) is thinking about.

- **personality:** \(< X >\) has a personality.

- **visual impact:** \(< X >\) ’s behaviour draws my attention.

- **predictability:** \(< X >\) ’s behaviour is predictable.

- **behaviour coherence:** \(< X >\) ’s behaviour is coherent.

- **change with experience:** \(< X >\) ’s behaviour changes according to experience.

- **social:** \(< X >\) interacts socially with other characters.

For emotional expressiveness the participants are asked what emotions are displayed by the agent in specific moments, such as joyfulness or sadness. If participants frequently identified the same emotion the system was aiming to reproduce, this dimension would score higher. Conversely, if they interpreted a different emotion, the score would be lower.

The authors put forward the hypothesis that, apart from predictability, higher dimension values will result in a higher sense of believability.

3 Solution

We are proposing a new, clearer approach to agent communication and cooperation. The core of our approach consists in dividing an action in two stages: anticipation and execution.

The anticipation stage, just as in traditional animation, serves the purpose of communicating the intent so every other agent and human player is expecting it and can prepare for it accordingly. The human player should be able to “read” the intention of any other agent and to interact with it as well. After the anticipation stage is complete, the next step of the action will be its execution. In this stage, as the name implies, the action (such as passing the ball) will be executed.

The stages may be overlapped at any time. As such, an agent might make a pass and immediately after start a new anticipation stage, such as requesting the ball in an open area. The action may also be interrupted before the execution stage, such as when an agent expresses the intent to make a pass, but the opponent marks the teammate before it gets the chance to. The agents must be aware of each others current stage of the action, so it will be broadcast to every agent, otherwise they wouldn’t be able to effectively cooperate. Conceptually, it will also be broadcast to the human player through gestures and ”body language”.

3.1 Agent architecture

The **Decision Making** module will select an **Intention** according to both **Beliefs** and **Desires**, as expected in a BDI architecture. The intention will have associated a confidence value which will be generated according to the agents beliefs. When the action enters the anticipation stage, the agents in the environment may react in a way that will influence the confidence of the subject. If the confidence value goes below the confidence threshold, the agent will be frustrated, the action interrupted and the process starting all over again. Table 3.1 illustrates how positive and negative feedback can impact the confidence of an agent.

<table>
<thead>
<tr>
<th></th>
<th>Positive feedback</th>
<th>Negative feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expects to score</strong></td>
<td>Confidence += 30%</td>
<td>Confidence -= 15%</td>
</tr>
<tr>
<td><strong>Does not expect to score</strong></td>
<td>Confidence += 15%</td>
<td>Confidence -= 30%</td>
</tr>
</tbody>
</table>
The percentage values in table 3.1, as the confidence threshold value, are dependent on the personality of the agent. Just as feedback impacts different people differently in the real world, it should impact different agents differently. With this approach, we aim to provide the possibility to easily develop agents with distinct personality types.

If the agent has a confidence value above the confidence threshold after all feedback is given, the action will enter the execution stage. The agent, as well as its teammates, will sense the outcome of the action and react accordingly. The outcome of the action can also influence the confidence associated with future intents. As such, if an agent is consistently having a bad game, it may be more reluctant to ignore negative feedback and thus more prone to execute actions that have its teammates consent.

The decision making module however, is beyond the scope of this work, and was emulated by a script. This script features both the actions and the confidence associated with the actions that each agent should perform sequentially. It may also set the desire of an agent. The script does not take care of the communication or synchronization between the characters, nor does it feed the agents the emotions it should display, as these items are the focus of our work and are accomplished dynamically.

Not every action is divided in two stages however, otherwise the observer would no doubt be overwhelmed and confused. We decided that there would be no anticipation stage for the simple action of moving from one place to another, or dribbling the ball from one place to another.

The following figure illustrates how the different modules interact with each other:

To summarize the role of the main modules of the agent architecture:

- **Decision Making**: Has the role of selecting an intent. To select the intent it considers the agent’s desires and beliefs.
- **Intention**: The current intent selected by the Decision Making module. The intent has a confidence value associated with it.
- **Sensors**: The mechanism which will enable the agent to receive stimuli from the environment and its teammates.
- **Expectation**: The agent will have an expectation associated with the action at hand. It may believe he won’t be able to score, for instance, or vice-versa. It is subject to the agent’s base confidence value and confidence associated with the action.
- **Affective Appraisal**: The affective appraisal module will interpret the sensed stimuli and decide how it will influence the confidence of the current action if the agent is performing an action. It may raise the confidence or it may lower the confidence. The impact which will have on the confidence will be decided along with the agents expectation regarding the action.
- **Emotional State**: Is the mechanism which will decide the emotion to be displayed at a given time. It will be influence by the Affective Appraisal module.
- **Action**: The action is divided in two stages:
  - **Anticipation**: Before executing the action, the agent will broadcast its intention to the world.
  - **Execution**: The action is executed in this stage, if the action’s confidence value is above the agent’s base confidence value after the teammate’s feedback.
- **Reaction**: The agent will express its emotions reactively depending on the events that occur in the environment. It will express reactively if a teammate missed an easy goal, or he does not agree with an intention of a teammate, for example.

4 **Soccer Pucks**

The model is going to be implemented in a multiplayer sports game called Soccer Pucks. Soccer Pucks is a physics based game where each player controls one character (unlike other sports games where the player controls the whole team) in a team of two or three elements with the objective of scoring more goals than the opponent. There are no pre-defined roles in Soccer Pucks, so each player must choose his positioning in the pitch based on
his own preferences and also on his judgement of circumstances, dynamically, during the match.

We have also built a new character, more suitable for communication, that will use his hands to communicate and which we named Tom.

4.1 Agent’s expressions

Tom is able to communicate with big hands that will show up when necessary. We developed four animations that communicate three different desires: Pointing, waving, reject pass, and thumbs up. These animations are exaggerations of real world gestures in a football match.

Tom also expresses itself through its emotions. The reactive module in our architecture is able to deliver three distinct emotions: frustration, anger and joy. These emotions are used in a simple and straightforward way: If an agent misses a goal it shows frustration, when an agent asks for a pass but it is ignored and its partner misses the goal it will be angry, and when a goal scored it will display joy.

- **Agent’s gestures**

  - **Pointing:** The agent uses its index finger to point to the area where it wants to shoot at. It is an animation used in the action’s anticipation stage.
  
  - **Waving:** The agent distressfully waves its hands in the air, drawing attention to itself, displaying an intent to receive the ball. It may serve as negative feedback if the teammate had already expressed a different intent, such as shooting at goal.
  
  - **Reject pass:** The agent interprets the teammate’s intent to pass the ball and rejects it, pointing to wherever he feels its teammate should place the ball at instead. This animation is similar to the pointing animation, except the agent swings his arm up and down in order to draw the teammates attention not to itself, but to wherever it is pointing. Because it contradicts the teammate’s intent of passing the ball, it delivers a negative feedback.
  
  - **Thumbs up:** The agent acknowledges its teammate intent and agrees with it. It represents a positive feedback.

- **Agent’s emotions**

  - **Frustration:** This emotion is displayed when the agent misses a goal. A drop of sweat of exaggerated proportions slowly falls from Tom’s face. We drew inspiration from cartoons.
  
  - **Anger:** Pressurized steam is leaked from both sides of the agent, where ears would have been. This animation is also a borrowed concept from many cartoons.

4.2 Experience

In order to have more control in the evaluation, we decided that the user should observe videos rather than watching a real-time demonstration. Due to the unpredictability of the physics, it was the only way we could assure every user would be watching the exact same scene.

We developed a questionnaire that would ask the participants questions regarding each one of the videos shown. The questionnaire was filled in 20 to 30 minutes. In order to avoid bias, we developed eight variants of the questionnaire with different question ordering. The participants were sent a link, and the link would redirect the participants to a random version of the questionnaire. We opted for a Within Subjects Design as we wanted the participants to experience every video so the different perceptions accross every video could be compared.

4.3 Videos

We developed a scene with approximately fifteen seconds. In the scene there are two agents - two Toms - that are from the same team (Red). Their objective is to score a goal. There are no opponents in order to focus the user’s attention in the interaction of the only two agents in the pitch. The spectators play no role other than to make the experience more visually appealing. In our following descriptions we will be referring to the agents as ”the top agent” and ”the bottom agent” to the agent which will always be in the left side of the pitch (hence nearer to the top of the screen) and to the agent which will always be at the right side of the pitch (bottom of the screen) respectively.

4.3.1 Scenario

In our scenario, the top agent dribbles the ball to the left flank and initiates the anticipation stage of the action by showing its intent in shooting at goal. The bottom agent, which intended to be passed the ball waves its arms in the air requesting the pass, to which the top agent relents. The bottom agent receives the ball and scores the goal.
4.3.2 Additional videos

We have also used two additional videos similar to the one that describe our scenario, but one is without the anticipation module, and the other without both the anticipation module and the reactive emotion module. To dissipate the suspicion that the users may prefer the videos with both modules not because it is more believable but simply because it has more animations and thus is more appealing, we decided to introduce another video, similar in result as the others - The top agent passes to the bottom agent which scores - but with random animations that have nothing to do with the agents intentions.

4.4 Questionnaire

In our questionnaire we asked users to watch nine videos, each with approximately fifteen second length. After each video, the users were asked for their agreement with the following statements:

- The intentions of the artificial players are easy to understand.
- The artificial players are aware of their surroundings.
- It is easy to understand where each artificial player is focusing its attention on.
- The artificial players communicate with each other.
- The emotion of the artificial players are easy to understand.
- The artificial players clearly display their personality

These statements are an adaptation of the work of Gomes et al [4] described in section 2.5. The audience perception was assessed using Likert scales that went to 1 - Strongly disagree, to 5 - Strongly agree.

We also decided to create a question to assess the user’s interpretation of the agents intentions. In order to facilitate statistical treatment of the data collected, we decided to create two sets of multiple answer questions for each player, one for the top agent who passed the ball, and the other for the bottom agent who shot at goal. So, for the four videos we had:

- Related to the top artificial player:
  - The top player passed the ball because that was the wish of the bottom player.
  - The top player passed the ball because that was its own wish.

- Related to the bottom artificial player:
  - The bottom player intended to receive the ball.
  - The bottom player did not intend to receive the ball.
  - I can not decide, the video is not clear enough.

Lastly the questionnaire featured a page with the same four videos, and asked which video portrayed the agents as "life-like, believable creatures with interesting human-like qualities, and less like pre-programmed (ro)bots?" The user was asked to order those four videos by order of preference. The videos were the video with both the anticipation and reactive emotion module, without the anticipation module, without the anticipation and reactive emotion module, and the random video. After the ordering the user was asked, in open answer question, to briefly justify his choice and what would he suggest to improve his favourite approach.

5 Results

Because the video with the anticipation module was the richest in terms of interaction between agents and because it was conclusive in terms of results, we will proceed to make an in-dept analysis of this scenario, comparing it to its three homologous variants: Without the anticipation module, without both the anticipation and the reactive module, and the one with random animations. In tables and in figures, these four approaches are respectively labeled Anticipation, Reaction, None, and Random for simplicity.

5.1 Believability measuring

The Spielman Wilks test indicated that the data set for the six perceived believability measuring questions for each of the four videos are not modeled by a normal distribution therefore we will be using non parametric tests on our data. We then proceeded to apply the Friedman’s Test in four exemplars of videos each representative of our model with anticipation and reaction, reaction only, without both modules, and a random animations videos to detect bias, in order to conclude whether participants ranked our videos differently.

In order to detect which videos are ranked differently, we proceed to use a Wilcoxon signed rank test, for each of the six believability questions.
5.1.1 The intentions are easy to understand

- Wilcoxon signed rank test

<table>
<thead>
<tr>
<th></th>
<th>Anticip.</th>
<th>Reaction</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction</td>
<td>Z = -4.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Z = -5.1</td>
<td>Z = -2.1</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td>p = 0.0</td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>Z = -5.2</td>
<td>Z = -3.8</td>
<td>Z = -2.3</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td>p = 0.0</td>
<td>p = 0.0</td>
</tr>
</tbody>
</table>

Table 5.1.1 shows us that, between the video with the anticipation module and the remaining videos, there is a statistical relevant difference in user perception of the agents intentions, with the preference clearly residing in the video with the anticipation module. To our surprise, the same relevant difference exists in the video with neither module and the one with only the reactive module, even though the p-value is higher (0.038) it is still below the significance of 0.05.

5.1.2 The agents are aware of their surroundings

- Wilcoxon signed rank test

<table>
<thead>
<tr>
<th></th>
<th>Anticip.</th>
<th>Reaction</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction</td>
<td>Z = -1.6</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>p = 0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Z = -3.2</td>
<td>Z = -3.2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td>p = 0.0</td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>Z = -3.7</td>
<td>Z = -3.2</td>
<td>Z = -1.0</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td>p = 0.0</td>
<td>p = 0.3</td>
</tr>
</tbody>
</table>

In table 5.1.2 we can conclude, that for a significance of 0.05, there is not relevant statistical difference in terms of preference in the dimension “The artificial players are aware of their surroundings” between the video with the anticipation module and the video with only the reactive module. Our approach is preferred, however, when compared against the video with neither module and the one with random animations.

5.1.3 It is easy to understand where the agents are focusing their attention on

- Wilcoxon signed rank test

<table>
<thead>
<tr>
<th></th>
<th>Anticip.</th>
<th>Reaction</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction</td>
<td>Z = -2.9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Z = -4.4</td>
<td>Z = -2.5</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td>p = 0.0</td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>Z = -4.2</td>
<td>Z = -3.1</td>
<td>Z = -1.3</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td>p = 0.0</td>
<td>p = 0.2</td>
</tr>
</tbody>
</table>

Again, table 5.1.3 shows a clear preference on the video with the anticipation module against its peers in the dimension “It is easy to understand where each artificial player is focusing its attention on”. Regarding the video with random animations against the video with neither module there is no statistical relevance for the preference of the participants for a significance of 0.05.

5.1.4 The agents communicate with each other

- Wilcoxon signed rank test

<table>
<thead>
<tr>
<th></th>
<th>Anticip.</th>
<th>Reaction</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction</td>
<td>Z = -5.4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Z = -5.8</td>
<td>Z = -2.9</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td>p = 0.0</td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>Z = -4.9</td>
<td>Z = -1.2</td>
<td>Z = -3.1</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td>p = 0.2</td>
<td>p = 0.0</td>
</tr>
</tbody>
</table>

As for the dimension “The artificial players communicate with each other”, table 5.1.4 shows yet again a clear preference for the video with the anticipation module. The participants did not make a statistically significant distinction between the video with random animations and with neither modules.

5.1.5 The emotions of the agents are easy to understand

- Wilcoxon signed rank test

<table>
<thead>
<tr>
<th></th>
<th>Anticip.</th>
<th>Reaction</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction</td>
<td>Z = -3.3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Z = -5.4</td>
<td>Z = -4.6</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td>p = 0.0</td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>Z = -3.9</td>
<td>Z = -1.8</td>
<td>Z = -2.9</td>
</tr>
<tr>
<td></td>
<td>p = 0.0</td>
<td>p = 0.1</td>
<td>p = 2.9</td>
</tr>
</tbody>
</table>

Regarding the dimension ”The emotions of the artificial players are easy to understand.”, table 5.1.4 shows us the users felt the artificial players displayed their emotions more clearly in the video with random animations (we remind that these animations also include emotions) than in the video with neither module, which is understandable. The participants did not feel, however, that there was a difference between the video with random animations and the video with only the reactive module for this particular dimension give a significance of 0.05. Regardless, the video with anticipation is the best ranked video also in this dimension.
5.1.6 The agents clearly display their personality

- Wilcoxon signed rank test

<table>
<thead>
<tr>
<th></th>
<th>Anticip.</th>
<th>Reaction</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction</td>
<td>Z = -4.0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>p</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>Z = -4.7</td>
<td>Z = -3.2</td>
<td>-</td>
</tr>
<tr>
<td>p</td>
<td>0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>Z = -2.4</td>
<td>Z = -1.6</td>
<td>Z = -3.2</td>
</tr>
<tr>
<td>p</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 5.1.6 shows us that, for the dimension "The artificial players clearly display their personality", the participants did not distinguish the video with random animations from the video with only the reactive module, for a significance of 0.05. As expected, the video with anticipation, is the most preferred video for this dimension.

5.1.7 Questionnaire reliability

In order to assess the reliability of the believability measuring questions, we decided to do a Cronbach’s alpha test.

We found that our alpha value of 0.904 would lower if any dimension is removed. We conclude that no dimension should be removed.

5.2 Interpreting the agents intents

As we stated before, after the participant, for each video, ranked the six believability measuring items, he would proceed to answer multiple-answer questions to assess his interpretation of the scene.

We can conclude that, in the video with the anticipation module, 76.9% of the participants identified the intentions of the top player as we intended. It is interesting to note that, the only means of interaction between agents was through the anticipation module, but that did not deter the participants from interpreting the interaction. As such, in the videos without anticipation and without both the anticipation and reaction, the majority of the users believed the top player passed the ball to its teammate because it wanted to in the first place, even though, as we can see by the bar graph, the participants are much less decided about it. This leads us to conclude that the participants tend to assume a behaviour to be voluntary and intentional unless some form of conflict is made explicit.

In regards to the bottom player, the advantages of the anticipation module are even clearer, with 92% of decisiveness.

5.3 User ordering

In the question that asked users to order the videos by believability, the video which featured agents with both the anticipation and the reactive emotion module (Video A) was in first place in 71.2% of the answers. The video which featured the reactive emotion module but not the anticipation module (Video B) was tied with the video which featured random animations (video D), both being placed first by 13.5% of the users. The video which featured agents with neither anticipation no reactive emotion module or random animations was put in first place by only 1.9% of the users. The video with random emotions was ordered in fourth and last place by 50.0% of the users.

5.4 Conclusions

Our analysis of the data we collected through the questionnaire led us to conclude that the video with the anticipation module was consistently ranked higher across every dimension, which means that, according to the work of Gomes et al[4](see section 2.5) the participants perceive the video featuring
our approach to be more believable than the remaining videos.

The participants also found the video with the anticipation module much clearer and easier to understand than the remaining videos. Only 9.6% were left undecided about the intention of the top player in the video with the anticipation module, against 26.6%, 40.4% and 48.1% in the videos with only the reaction module, without the anticipation and reaction module and with random animations respectively. Regarding the bottom agent the results were even more pronounced, with 1.9% of uncertainty in the video featuring our approach against 36.5%, 50.0% and 42.3% in the remaining videos respectively.

Lastly, when asked which approach portrayed agents as “life-like, believable creature with interesting human-like qualities, and less like pre-programmed (ro)bots?”, 71.2% of the participants selected our approach as their preferred approach.

6 Conclusions

With this work we aimed to create a new level of interaction between synthetic characters. A new level which would permit clearer communication, conflict and cooperation. These synthetic characters have their own expectations and desires which they seek to fulfill. If they fail they get frustrated. They also possess other human-like qualities to them, like confidence which is circumstantial.

Our approach consists primarily in dividing an action in two stages: anticipation and execution and we implemented our solution in a video-game called Soccer Pucks. We described the agent architecture, and how each module would be interacting with another. The agent uses gestures to communicate with its teammate and displays emotion regarding the outcome of its actions or the actions of its teammate. It is able to ask for the ball, reject a pass, and show an intent to pass or to shoot. It may get angry, frustrated and display joy.

To validate our approach we developed a scenario that would feature two agents that communicated with each other to score a goal. This scenario led to four videos: a video with our anticipation module, a video without our anticipation module, another video without both the anticipation module and the reactive module, and, lastly, a video with neither module but featuring random animations.

Analysing our data gathered from the questionnaires led us to conclude that our approach is highly successful. The participants showed a clear preference for the video with the anticipation module in every believability dimension as well as in the ordering. The participants also felt the video with the anticipation featured clearer interactions that were easier to interpret that the other two approaches.

6.1 Future Work

There is still a lot of work that can be done around this approach. Early in our work we intended to develop a version of Soccer Pucks which would enable the player to cooperate and play against the artificial characters and a full featured match. However, midway through development, in early playtesting, we found that this approach was not the best to objectively evaluate our model. We decided instead to go for a more controlled approach, that would not be dependent on player skill and experience with video-games. Now, after having validated our model, we feel that it would be interesting to see it implemented in a playable scenario.

The transition to a real-time sports game isn’t trivial however. The decision making module would have to be fully developed, and believable decision making is not straightforward, as any mistake may break the player’s suspension of disbelief. Nevertheless, it would be interesting to adapt the model to other genres of games that could benefit a new layer of believable interaction, such as in adventures games like Skyrim, where the player’s attention was not focused on his companion at all times, so a mixture of body language and vocalization could be required to draw the player’s attention when needed.

References