Enhancing the Believability of Embodied Conversational Agents through Environment-, Self- and Interaction-Awareness

Kiran Ijaz[1] **Anton Bogdanovych**[2] **Simeon Simoff**[2]

[1] Faculty of Engineering and Information Technology
University of Technology,
Sydney, NSW, Australia.
Email: `kijaz@it.uts.edu.au`

[2] School of Computing and Mathematics
University of Western Sydney,
NSW, Australia.
Email: {`a.bogdanovych, s.simoff`}`@uws.edu.au`

Abstract

Research on embodied conversational agents' reasoning and actions has mostly ignored the external environment. This papers argues that believability of such agents is tightly connected with their ability to relate to the environment during a conversation. This ability, defined as awareness believability, is formalised in terms of three components - environment-, self- and interaction-awareness. The paper presents a method enabling virtual agents to reason about their environment, understand the interaction capabilities of other participants, own goals and current state of the environment, as well as to include these elements into conversations. We present the implementation of the method and a case study, which demonstrates that such abilities improve the overall believability of virtual agents.

Keywords: Virtual Worlds, Artificial Intelligence, Embodied Agents

1 Introduction

Virtual agents and online virtual worlds they populate are a catalyst that will accelerate the development of embodied Artificial Intelligence (AI) (Livingstone 2006). Having both humans and agents fully immersed into and constrained by the same computer-simulated environment provides fantastic opportunities for embodied AI researchers to study human behaviour, investigate cognition related aspects and search for more practical application domains for the discipline. Virtual agents have the opportunity to engage and learn from their interaction with the social network of human beings operating in these worlds. However, for that to happen, they need to keep humans engaged in meaningful joint activities - they need to be *believable* in terms of their "life" in the computerised environment they operate in.

Surprisingly enough, the majority of virtual agents are not fully integrated with their environment in terms of their reasoning. Most of existing research in the area is focused on agents' conversational abilities (Gandhe & Traum 2007), use of gestures (Hartmann et al. 2005) and emotions (Cunningham et al. 2005), social role awareness (Prendinger & Ishizuka 2001) while providing the agent with a limited awareness about the objects in the complex dynamic virtual world ((Johnson & Lester 2000) and (Lester et al. 1999)), its own state in relation to the environment and ignoring the interactions of other participants with the environment (Doyle 2002).

To illustrate the importance of integrating such features into agent conversations consider a scenario outlined in Figure 1. Here a human user learns in a Virtual World about the history and culture of the ancient city of Uruk (Bogdanovych et al. 2010) through controlling an avatar of Fisherman Andel. There is a clear task he has to accomplish (catch some fish) and this task involves interacting with virtual agents representing ancient citizens of Uruk. As shown in the figure the human isn't clear about how to progress toward the assigned task and asks the agent (Fisherman Jigsaw) about his options. To be able to reply in a similar manner as shown in the picture, the agent must know its interaction state, its location in the environment in relation to other objects and avatars, its own goals and beliefs and interaction capabilities of the user.

In 3D Virtual Worlds the integration of agents and environment in terms of agent reasoning is a feasible task due to the fact that the perception problem in such environments is minimised. Each agent can operate with precise coordinates of other participants and objects in the environment, request their names, properties, distances to them and operate with a number of own parameters (i.e. eye direction, body rotation etc.) to determine visibility of the objects, predict the movement of other actors and identify the target of their attention. Furthermore, in virtual worlds like Second Life (http://www.secondlife.com) it is even possible to supply agents with information about the elements constituting a particular object, as each of the objects there is composed of a number of primitives. We envision that by developing such mechanisms to utilise this information provides a very powerful toolkit for a sophisticated agent reasoning apparatus that significantly increases the believability of agent behaviour and its capacity to engage humans.

In this paper, we suggest that supplying virtual agents with an ability to reason about the objects in their environment, own state, goals and beliefs as well as the interactions of other participants would result in more believable virtual agents, which can be used in a much wider range of problems than at present. The

---

FishermanUruk1 Andel 20:57
What are you waiting for?

FishermanUruk2 Jigsaw 20:57
I'm waiting for you to start fishing

FishermanUruk1 Andel 20:57
How do I do it?

FishermanUruk2 Jigsaw 20:58
You should grab the spear that is on your left and sit inside the boat

Figure 1: Scenario: A conversation between two Fishermen.

remainder of the paper is structured as follows. In Section 2 we analyse the concept of believability and investigate believability features. In Section 3, we discuss and formalise the believability features that haven't received an appropriate research attention. Section 4 presents our approach to integrating these features into virtual agents. In Section 5, we illustrate our approach through a case study and evaluate it in Section 6. Finally, Section 7 concludes the paper.

2 Believability of Virtual Agents

Virtual agents are represented as graphical characters (avatars) that may or may not resemble a human body. It should always be the goal for these agents to act "believably" given the representation they have; hence, enabling and evaluating their believability is essential.

2.1 The Notion of Believability

The notion of believability originates in the field of animation and theatre (Mateas 1997). A believable character can be defined as a lifelike in its behaviour, whose actions appear to be real, who engages its audience and is perceived as human-like. A classical work of Walt Disney Studio on these animated characters "illusion of life" (Thomas & Johnston 1981) elaborates on the requirements for believability. Though these characters are not real they continue to impact the audiences' imagination to accept them as believable. Believability and realism have been differentiated by (Mateas 1997) and (Doyle 2002). According to the authors, a believable character does not necessarily mean to be a real character or who tells the truth; it is real in the context of its environment.

Earlier research in believability has been heavily influenced by the Carnegie-Melon set of requirements for believable agents, which is based on research in drama and story telling (Loyall 1997). These include personality, self-motivation, change, social relationships, and "illusion of life". Personality infuses everything that a character does - behaviour, style, "thought", "emotion", e.g. unique ways of doing things. Self-motivation assumes that agents have their own internal drives and desires which they pursue whether or not others are interacting with them, and they demonstrate their motivation. Change implies that characters change with time, in a manner consistent with their personality. Agents' behaviour and interactions should be consistent with their social relationships. "Illusion of life" is used as a label for a collection of features such as: pursuing multiple, simultaneous goals, having elements of broad capabilities (e.g. movement, perception, memory, language), and reacting quickly to stimuli in the environment.

Contemporary AI uses the term "believability" in relation to engaging life-like systems. Reactivity, interactivity and appropriate decision making while observing behaviours are few characteristics which make autonomous agents suitable to achieve believability (Riedl & Stern 2006). Believability is also a main requirement of modern computer games. As suggested by (Livingstone 2006), "the need for modern computer games is not unbeatable AI, but believable AI".

2.2 Believability of Embodied Conversational Behaviour

Current research in believability of embodied conversational behaviour of virtual agents is focused on the believability aspects of the conversation itself and the non-verbal communication cues, including facial expressions, gestures and gaze.

2.2.1 Believable Conversations

Believable conversation in relation to the content is a term often associated with "chatbots" - software programs that try to keep a human engaged in a textual or auditory conversation. Virtual agents with such capability have been used in various commercial applications, games, training systems and web-based applications. Chatbots like Eliza (Weizenbaum 1966) and ALICE (Wallace 2004*b*) are based on pattern matching strategies. Technically, chatbots parse the user input and use keyword pointing, pattern matching and corpus based text retrieval to provide the most suitable answer from their “knowledge base” (Gandhe & Traum 2007).

2.2.2 Believable Facial Expressions

Facial expressions can be used to complement the word stream through expressing emotions, i.e. happiness, sadness etc(Cunningham et al. 2005). These emotional expressions have cross cultural boundaries, but generally existing work deals with a list of emotion expressions: {happy, sad, fear, anger, disgust, agreement, disagreement and surprise} as presented in (Cunningham et al. 2005).

A comprehensive survey of techniques for automatic analysis of facial expressions was presented by (Pantic & Rothkrantz 2000). It has been further noticed that most of the existing work deals with specific cases of image and video analysis (Pantic & Rothkrantz 2000).

2.2.3 Believable Gestures and Upper Limb Movements

Gestures are one of those factors in non-verbal communication which allow us to interact in a lively manner. Gesture selection and their correct execution may increase the expressivity of the conversation (Hartmann et al. 2005). Believable gestures are related to gestures selection being correctly aligned with the flow of conversation and the generation of realistic movements of agent’s upper limbs during the conversation (Hartmann et al. 2005).

2.2.4 Believable Gaze

Gaze helps to convey the cognitive state of a participant or synchronise a conversation as explained in (Lee et al. 2007). Various gaze models like avert, examining the current task, gaze at visitors, etc. were simulated by (Heylen et al. 2005). They measured the believability of the agent based on factors like satisfaction, engaging, natural eye, head movements and mental load among others; and this study showed the significant improvements in communication between humans and virtual agents. Lance in (Thiebaux et al. 2009) investigated a hybrid approach combining head posture, torso posture and movement velocity of these body parts with gaze shift.

3 Awareness Believability

Awareness is essential part of our conversational behaviour. In a conversation we are aware of where we are (environment awareness), who we are (self-awareness) and generally how the interaction is progressing (interaction awareness). Therefore, awareness is an essential component of the believability of embodied conversational behaviour, which we label as “awareness believability”. Further, we develop each of the subcomponents of awareness believability.

3.1 Environment Awareness

The importance of environment awareness for agent reasoning is best illustrated in (Elpidorou 2010), where it is suggested that our consciousness does not arise from the brain alone but from the brain’s exchange with its environment. Humans are embodied in space and use various cues related to space, like pointing and referring to areas of and things in it, in all they do (for more details see the chapters 1,2 in (O’Keefe & Nadel 1978)).

Existing literature presents a very limited picture on the use of environment awareness by animated agents. Agents like Cosmo (Lester et al. 1999) and Steve (Johnson & Lester 2000) are able to recognise and point to objects in the particular static environment but completely ignore their interactions with other participants, don’t cater for dynamic environment and can not orient themselves in a different environment. We suggest that awareness of environment objects alone would not be enough to achieve complete believability and virtual agents further require to be aware of other participants in the context of time.

Our key features of environment awareness include the positions of objects and avatars in the environment, how these evolve with time and the direction vectors associated with avatars (Gerhard et al. 2004). We formalise environmental awareness as follows:

$$EA = \{Objects, Avatars, Time\} \qquad (1)$$

Here EA is the set of components of environment awareness and includes the objects in the environment, other avatars representing agents and human participants with respect to the current time.

3.2 Self-awareness

Knowing own context and state within the environment, i.e. being self aware, is essential for a virtual agent to interact believably (Doyle 2002). To achieve that Doyle (Doyle 2002) proposes to annotate the environment and grant agents with access to this annotation. One of the most studied features of self-awareness for virtual agents and animated characters is social role awareness (Prendinger & Ishizuka 2001). However, self-awareness is a much richer concept and many of its characteristics remain understudied, in particular existing works mostly ignored many vital characteristics that arise in dynamic environments.

Hallowell defines self-awareness (Hallowell 1955) as the recognition of one’s self as an object in the world of objects and highlights the importance of the perception as the key function in self-awareness. The list of elements we have identified to enable self-awareness is as follows:

$$SA = \{G, P, B, Sc, St, ObjUsed, Role, Gest\} \qquad (2)$$

Here SA represents the set of components of self-awareness and includes the local goals of the agent (G), its current plans (P) and beliefs (B), current scene where the agent participates (Sc), its state within this scene (St), objects used by the agent ($ObjUsed$), the role it plays ($Role$) and the gestures being executed ($Gest$).

3.3 Interaction-Awareness

Believability of interactions goes beyond traditional focus on modeling the visual co-presence (Gerhard et al. 2005), Context awareness (perceiving other agents/objects in static environment) (Bickmore et al. 2007) and communication style (e.g. short vs long utterances, usage of specific vocabulary) of the agents. Human behaviour in interactions is a result of the mix of being rational, informed, impulsive, and the ability to influence others and cope with the influences from others. All these nuances impact the richness of human interactions, hence, must be taken into account when considering the believability of interactions between virtual agents and humans.

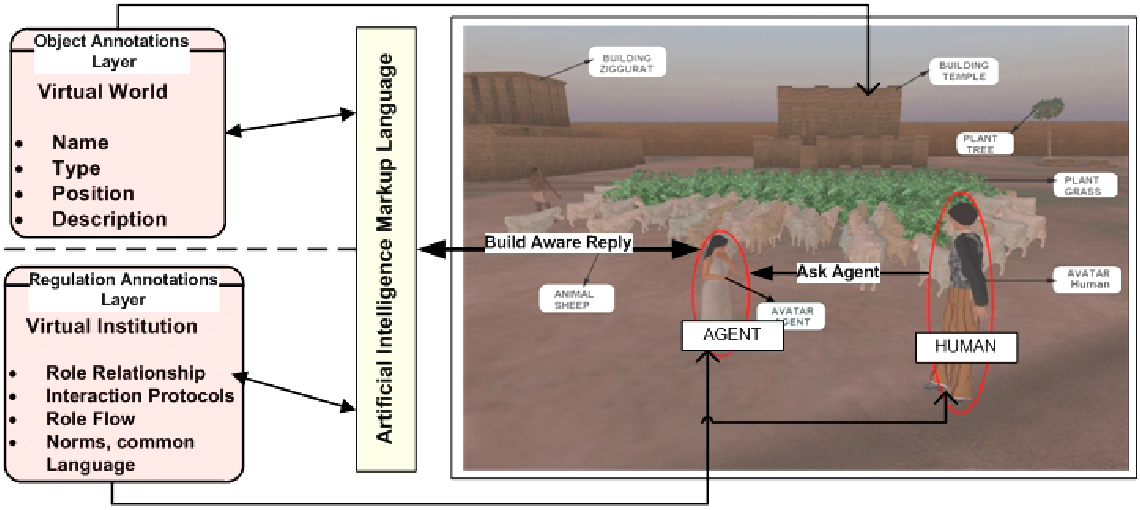


Figure 2: Layered Architecture.

Thus, interaction-awareness is defined as the state of an agent who is "able to perceive important structural and/or dynamic aspects of an interaction that it observes or that it is itself engaged in" (Dautenhahn et al. 2003). The components of the interaction-awareness model are outlined below.

$$IA = \{AV_{vis}, AV_{sc}, Act, Obj, State, Pos, Or\} \quad (3)$$

Here IA represents the set of components included in our interaction awareness model. AV_{vis} corresponds to the set of currently visible avatars. The AV_{sc} is a set of all avatars within the scene where the agent participates in a given moment of time. Act represents the set of actions each of the agents in the current scene is able to perform given its state. Obj refers to the list of objects the avatar can use. $State$ is the state of the avatar in the world. Pos is the position of the agents in the virtual world and Or is agent's orientation vector in the virtual world space. In the next section we outline the implementation details of our approach.

4 Implementation of Awareness Believability

In order to incorporate the aforementioned believability features in regulated interactive environments we propose to conduct two levels of environment annotation as shown in Figure 2: (i) *object annotation* - the annotation of objects in the environment with appropriate names, object types and descriptions (such annotations are fully supported by the Second Life technology that we use for design and visualisation of the dynamic virtual world); and (ii) *regulation annotation* - annotation of the social norms, interaction protocols, roles and other kinds of regulations of interactions. The Virtual Institutions (VI) technology (Bogdanovych 2007) is used to enable regulation annotation.

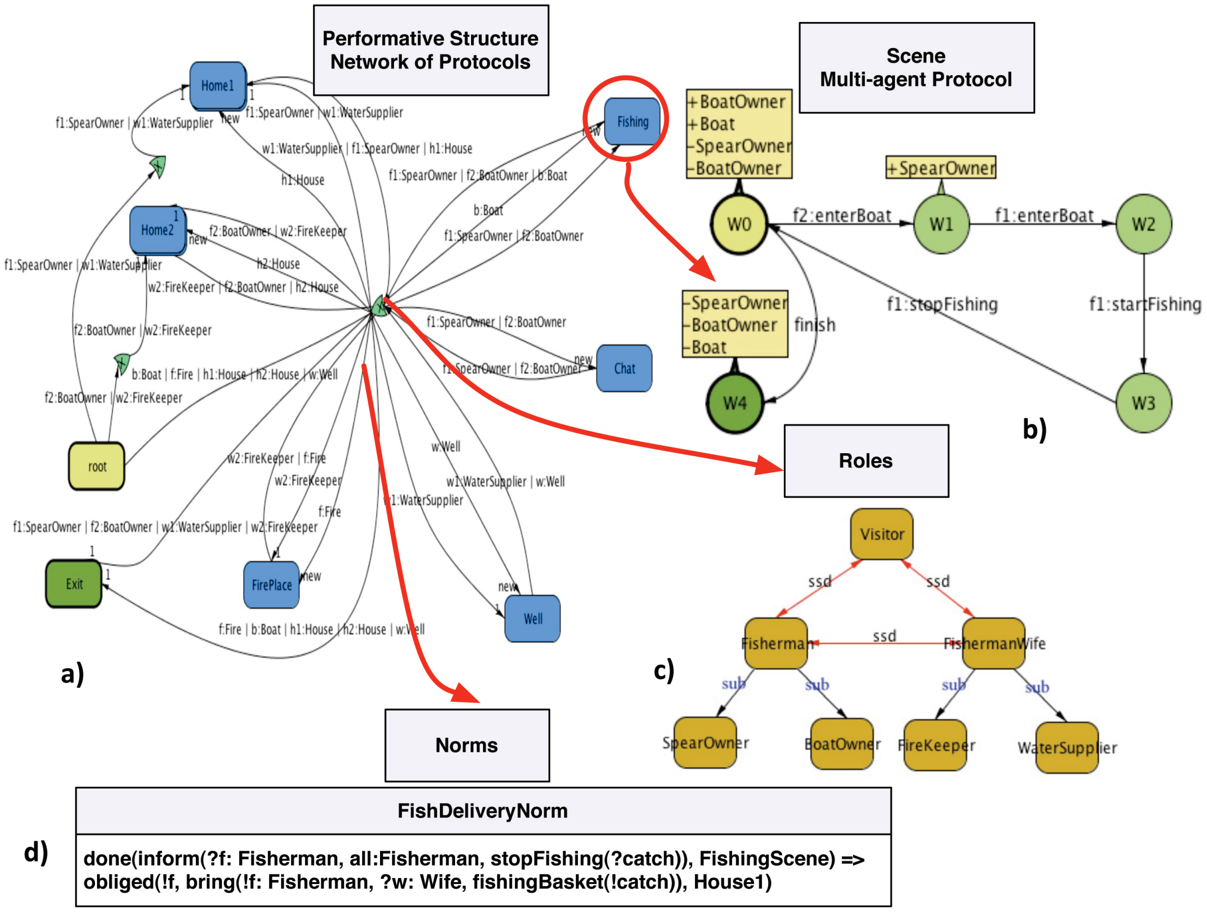
In our system human users interact with virtual agents present in the virtual world through a communication layer and these agents rely on two further layers to generate an intelligent response as shown in Figure 2. If the query asked was about the agent's environment, object annotations layer would be requested by the agent to generate a response which contain the objects' information. Whereas, queries regarding the agent's interactions and self awareness would be passed to the regulation annotations(VI) layer by the agent. VI layer passes interaction annotations (i.e. agent's goal, plans, objects used etc.) back to the Artificial Intelligence Markup Language (AIML) module to further build a response. This communication module has the responsibility of generating the text reply for the agent based on the information received from the virtual world or the virtual institution layer. Subsequent sections would provide a detail insight about these layers.

Technological separation of the environment layer from the Institutional regulations allowed us to implement a generic solution, where same features could be deployed in a new dynamic environment without modifying the core functionality. Secondly, this layered approach enables to deal with a dynamic environment, where objects could be changed, inserted or deleted at any time. All annotations outlined in Figure 2 can still be detected by the agent for newly added objects. Further, we briefly describe the Virtual Institutions technology and show how it can enable the integration of the awareness believability features in such dynamic virtual worlds.

4.1 Virtual Institutions Technology

The Virtual Institutions technology (Bogdanovych 2007) provides tools for formal specification of institutional rules, verification of their correctness, mapping those to a given virtual world and enforcing the institutional rules on all participants (both humans and autonomous agents) at deployment, creating a *regulated virtual world*.

The specification is expressed through three types of conventions and their corresponding dimensions (Esteva 2003): *Conventions on language – Dialogical Framework*, determines language ontology and illocutionary particles that agents should use, roles they can play and the relationships among the roles; *Conventions on activities – Performative Structure*, establishes the different interaction protocols (scenes) the agents can engage in, and the role flow policy among them. A scene is a logically independent activity that can be enacted by a group of agents. In Virtual Institutions each scene often corresponds to a particular space in the virtual world; *Conventions on behaviour – Norms*, capture the consequences of agents' actions within the institution (modelled as commitments and obligations) that agents acquire as consequence of some performed actions and that they must fulfil later on. Thus, the Virtual Institutions



done(inform(?f: Fisherman, all:Fisherman, stopFishing(?catch)), FishingScene) =>
obliged(!f, bring(!f: Fisherman, ?w: Wife, fishingBasket(!catch)), House1)

Figure 3: The interaction related information provided by the Normative Platform.

technology helps us define the interaction protocol for all the specified scenes, relationships and role flow among the participants and norms of the society. An example of specifying the aforementioned dimensions is shown in Figure 3.

At deployment, the specification is connected to the virtual world (or respective area, i.e. an island in Second Life) and the state of the virtual world is mapped onto the state of the institution. Having the institutional formalisation aligned with the virtual world drives the decision making of virtual agents controlling the avatars. Each agent has access to the institutional formalisation and can sense the change of the institutional state and reason about the actions (both own actions and actions performed by other participants) that resulted in the state change. Hence, for an agent it makes no difference if it interacts with an agent or a human as the result of the other party's actions can be sensed through the normative platform and interpreted on a high level using the specification.

4.2 Implementing Interaction-Awareness

Enabling interaction-awareness means making an agent understand its own opportunities in interacting with other participants and predict the possible actions other participants may perform in a given scene provided the agent knows what the current state of the scene is. The VI technology tracks all the participants in every scene and maintains the state of every scene. Every action performed by any participant, as well as the corresponding state changes, can be sensed by the agent through the VI technology.

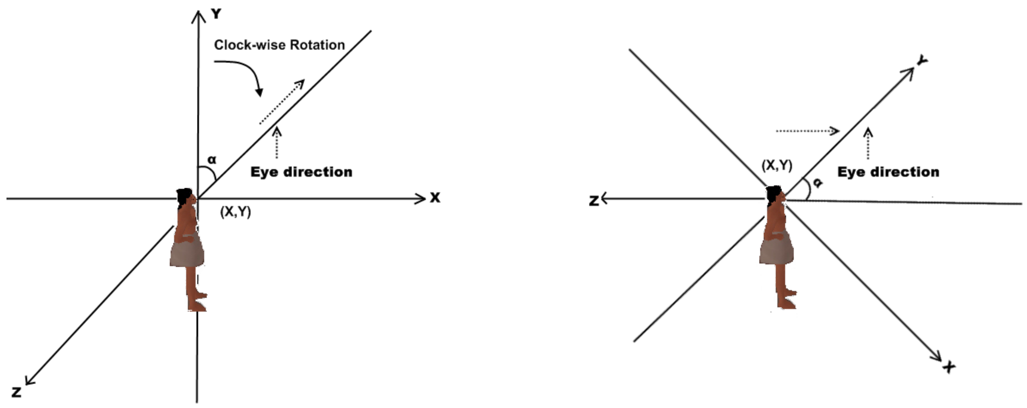
The VI technology provides the agent with high-level information of interaction opportunities of all participants. The agent can relate this information with its conversational rules and answers questions like "what can I do in this scene", "what can I do next", "what are you waiting for", etc.

The kind of information an agent can get through the VI technology is shown in Figure 3. The role hierarchy of possible participants is shown in Figure 3 c. The Performative Structure (Figure 3 a) illustrates possible role flow and deals with five major scenes so far including fishing, Well, Fire place, chat and fishermen Homes. To give an insight into the interaction protocol Figure 3 b illustrates the enactment of the fishing scene. Here the scene is represented as a finite-state machine where the state of the scene is changing as the result of the defined actions (i.e. enterBoat, startFishing, stopFishing, etc.). Finally, Figure 3 d presents an example of a social norm that all participants must abide by. In this case the norm dictates a fisherman to bring all the catch to his wife.

The agent can sense the movement over the performative structure, agents entering or leaving these scenes, track the actions that result in the state changes and also estimate the actions of other participants that can make the scene evolve to a new state. How this information can be used to dynamically generate text responses to various user queries is explained in the next section.

4.3 Implementing Environment-Awareness

Our environment-awareness method enables embodied virtual agents to have an up-to-date knowledge



a) Cartesian Planes w.r.t Agent's Current Position(x,y) and Agnle(α)   b) Transformed Cartesian Planes with Rotation Angle(α)

Figure 4: Aligning cartesian planes with agent's orientation.

of their surroundings. Our aim is for an agent to be able to interact with humans and correctly respond to human inquires about the objects and avatars in the environment. Agents' conversational abilities in our system are supported by the Artificial Intelligence Markup language (AIML) (Wallace 2004*a*). When a participant in the environment inquires something like: "What is that building on your left?" - we first extract the desired information from an AIML rule. The corresponding rule would be: "What is that *", where "*" represents any number of characters in the remainder of the query. First step of environment awareness process is to tokenise the input. These tokens are matched with the list of directions, objects and relative pronouns as shown in Table 1.

To locate the object(s) of interest - a virtual agent uses the data from the Object Annotation Layer to search for keywords like object name, type, etc as shown in Figure 2. For example: House and direction: "Fisherman house on the left". Moreover, it needs the exact relative pronouns like "You" or "Me" to co-relate the right objects in the space. As the user inquiry could be about an object on either participant's left or virtual agent's left, this hint helps the agent to specify participant's correct direction of interest.

Table 1: Environment Awareness

Environmental Entities List		
Pronouns	Directions	Objects
	In-front	Building(House, Ziggurat, Temple ...)
You/your	Behind	Plant(Tree,Grass ...)
Me/My	Left	Avatar(Agents,Humans)
	Right	Animal(Sheep,Donkey ...)

The agent further needs to locate the specified direction in the context of current space, which is relative to the pronoun given to search the "object" of interest in the virtual environment. The agent knows its current position vector and body orientation with respect to standard cartesian coordinate system. Identification of conversational agent's direction like "left" or "right" requires coordinate transformation in respect to the agent's current position and orientation.

Figure 4 (a) shows the position of embodied virtual agent in current cartesian coordinates space, where the agent currently looking towards the side pointed as eye direction and was away from the Y-plane with angle(α). Defining a direction like "left" or "right" with respect to agent's current orientation requires to rotate the cartesian planes clockwise or anti-clockwise with angle(α).

Agent's Eye Direction w.r.t Y-plane = Angle of Rotation = alpha - α, where (x, y) are the coordinates representing some point in the current space. In two dimensions, the rotation matrix has the following form:

$$\Re(\alpha) = \begin{vmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{vmatrix} (\textit{rotation by angle } \alpha).$$

The new coordinates (x', y') for point (x,y) will be:

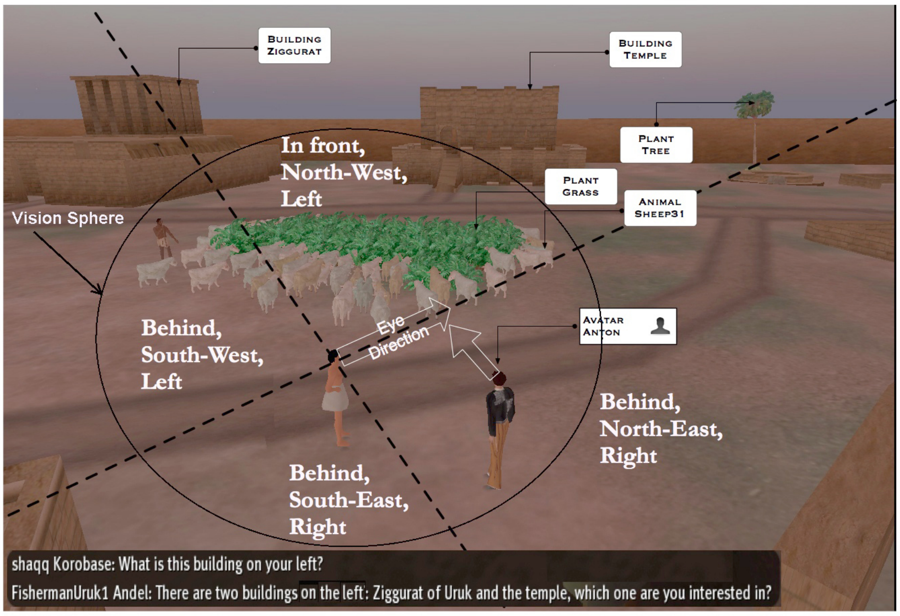
$x' = x * \cos(\alpha) - y * \sin(\alpha)$; $y' = x * \sin(\alpha) + y * \cos(\alpha)$.
To rotate the coordinates anti-clockwise, we need to replace α with -α.

$$\Re(-\alpha) = \begin{vmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{vmatrix} (\textit{rotation by angle } -\alpha).$$

This transformation of cartesian planes has been demonstrated in Figure 4 (b).

Once the agent has figured out the desired direction e.g. "left" in the current request, it further needs to search all the objects of class "Building" to locate the appropriate object. Each agent in the environment has its vision sphere as presented in Figure 5, which helps it to locate the object of its interest within the vision radius. In the given scenario the agent has found two buildings inside the vision sphere, which are Ziggurat and Temple of Uruk. If multiple objects of same class have been found – the agent would return the list of the objects and request more details. Otherwise, the object would be sensed within the vision sphere and its name is passed to the AIML module. The role of the AIML module is to produce the relevant object description given this name.

Each AIML rule(category) has two main tags to generate a dialog which are $< Pattern >$ and the $< Template >$ as shown in AIML example below. The Template tag generates a response to the inquiry requested in the Pattern tag. The AIML rule below also shows the $< Environment >$ tag, which is a custom tag responsible for invoking our environment awareness routine. The current AIML rule could be divided into two parts, "What is that" – is the general rule to invoke the corresponding template with custom $< Environment >$ tag. The second part is represented by a wildcard character "*", which encloses the user input part of the pattern. The $< star/ >$ tag substitutes the value matched by "*".



shaqq Korobase: What is this building on your left?
FishermanUruk1 Andel: There are two buildings on the left: Ziggurat of Uruk and the temple, which one are you interested in?

Figure 5: Agent orientation in Uruk.

```
<Category>
<Pattern>What is that * ?</Pattern>
<Template><Environment/><Star/></Template>
</Category>
```

Natural Language Processing (NLP) and the pattern matching is done by the AIML engine which also enables the custom tags. Our environment awareness routine is further responsible for completing the query and scanning the desired environmental objects. The OpenMetaverse library interacts with AIML to pass on the information about an agent's surrounding.

4.4 Implementation of Self-Awareness

The self-awareness component of an agent helps it to reflect on its own state in the virtual world, explain the reasons for performing certain actions, using certain objects or walking in a particular direction when a human user inquires about those. Moreover, to be more believable an agent should also be aware of its current goals and plans to give an appropriate response to the human user.

In its reasoning to respond to the query, the agent partially relies on the Virtual Institution (as shown in figure 2) that provides the agent with the corresponding knowledge about the current scene and the agent's state within this scene. The rest of the details the agent either extracts based on its location in the environment (i.e. identifying the objects in the field of view of the avatar, objects being used or animation being played) or from its internal beliefs. Figure 6 examplifies some parameters of the self-awareness model.

5 Case Study: The City of Uruk, 3000 B.C

Our case study demonstrates the kind of a complex application domain that becomes feasible once we have conversational agents that are environment-, self- and interaction-aware. Further we label them as "aware agents" and those that don't possess such qualities are called "unaware agents".

The case study tackles the domain of history education and aims at recreating the ancient city of Uruk and reenacting the life in it from the period approximately 3000 B.C. in the virtual world of Second Life, showing the visitors how it looked like and how its residents behaved in the past. This work is based on the prototype developed by our group (Bogdanovych et al. 2010).

Our objective is to provide engaging and interactive learning experience to history students, immersing them in the daily life of Uruk. Through embodied interactions virtual agents will teach various aspects of Uruk history, culture and the daily life of ancient Sumerians. Currently, we have two fisherman families "living" in the Uruk environment and performing their daily routines. Virtual agents' behaviours and social interactions have been institutionalised with the help of the VI technology. Before this work, these virtual agents had their daily routines in place but interactions with other participants were very limited - they could only talk about pre-feeded historical facts with the visitors and were not aware of their surroundings while performing daily life routines. Through environment- self- and interaction-awareness the agents now can explain what they are doing, why they are doing it in a certain way, refer to the objects in the environment and even help visitors to perform their assigned tasks and showing which actions must they perform to further progress in their assigned mission. Next we show how we evaluate the impact on believability of the awareness features.

6 Believability Evaluation

There exists no formal definition of believability, nor there are clear methods to measure it. Thus, we have adapted and modified to our needs the approach in (Gorman et al. 2006). The subjective nature of the approach has stimulated another aim of our work - the design a rigorous objective evaluation of believable conversational agents and calculating a believability index as a measure of their human-likeness.

6.1 The Design of Experiment

To test the believability of our agents, we designed an experiment and analysis technique adapting the

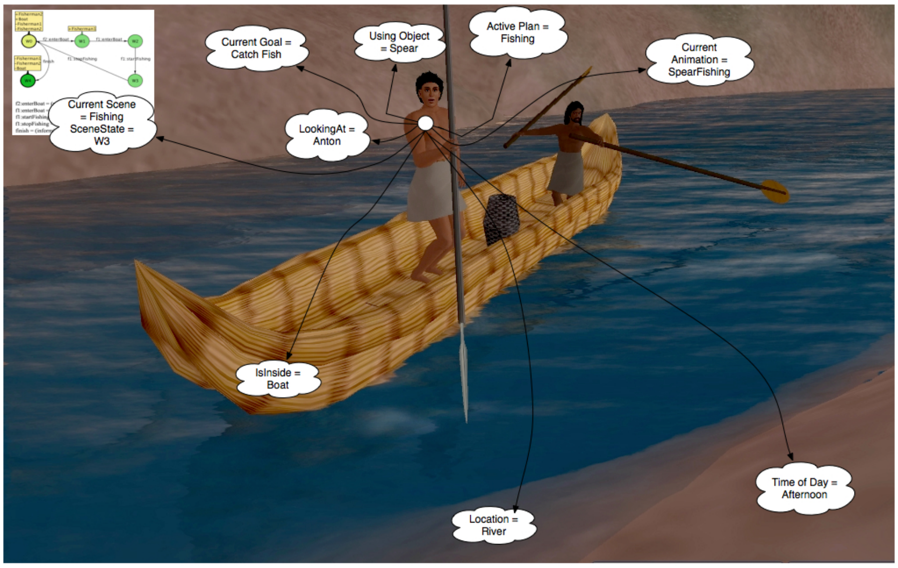


Figure 6: Self-awareness model.

methodology from (Gorman et al. 2006). The study sample included participants of diverse demographics. We provided participants with the background about Uruk and this study. During experiments the participants didn't know whether they had been conversing with avatars controlled by agents or humans.

We had to evaluate the believability of conversational agents, hence, presenting participants only video clips or other test mediums, as performed in (Gorman et al. 2006), was not acceptable due to the issues of biased responses and guess work. To minimise both we ensured that (i) participants interact with our conversational agents in the actual virtual world; and (ii) designer has no control over the routines of the agents with the flow of participants' conversations with them.

We also needed to present participants with highly immersive, engaging and interactive experience, which was essential in this study. From our past experience we have learned that navigation in virtual worlds requires some practice. Hence, the experiments were supervised by a researcher in order to assist participants with interfacing the agents.

Each participant was assisted to enter as an avatar into the city of Uruk. The participant was then requested to converse with two fishermen agents in the city, where our research assistant kept track of the conversation flow, ensuring that some of their questions were related to environment-, self- and interaction-awareness of virtual agents. The assistant navigated the agent herself, while the participant was directing her where to go, whom to talk with and what to ask. This allowed the participant to focus on examining the believability of the agent's dialogue in the context of its environment, actions and the behaviour of other participants in the city. As a result the participant had to assess the believability of each dialogue on the scale: {*1:Definitely Human; 2: Probably Human; 3:Not Sure; 4:Probably Artificial; 5:Definitely Artificial*}.

The rating of agents' conversations and behaviours was later used for calculating the believability index. The participants were requested to repeat this exercise for all parts of their conversations with our agents. They also had to provide specific feedback where they thought an agent's reply was artificial.

To minimise the chances of participants averaging out believability index (i.e. when a participant rates some virtual agent's responses as "Human" just because she rated few previous responses as "Artificial", we advised participants in the introduction that their rating should be purely based on their perception of the conversational behaviour of respective avatars in the virtual world.

6.2 Measuring Believability

For measuring believability we modified the equations for believability index from (Gorman et al. 2006) to reflect the interactive nature of our experiment, where the questions asked may differ across participants. Such index reflects participant's certainty with which s/he perceived a virtual agent as human-like or artificial. The equation for calculating the believability index for each dialogue is shown below:

$$h_p(c_i) = \frac{|r_p(c_i) - A|}{A - B} \tag{4}$$

where $h_p(c_i)$ is the perception of participant p of correspondence c_i as human-like and $r_p(c_i)$ is the rating of participant p for the same correspondence c_i. A and B represent the "Artificial" and "Human" value of the virtual agent response on the rating scale. Alternatively, $h_p(c_i)$ would be "0" if the respondent identified virtual agent's response as "Artificial" or "1" if s/he identified it as "Human", where all other values represent uncertain choices. The believability index for any participant is the average of his perceptions:

$$b_n = \frac{\sum_{0<p\leq n} h_p(c_i)}{n} \tag{5}$$

where n is the total number of responses per experiment. The overall believability in virtual agent's conversation B, based on the rating given by m participants is

$$B = \frac{\sum b_n}{m} \tag{6}$$

In a similar fashion, we could also measure the believability of each category for all participants in their conversations.

Table 2: Believability comparison for aware and unaware agents

Category	Unaware Agents	Aware Agents
Believability of environment-awareness	0.22	0.76
Believability of self-awareness	0.26	0.75
Believability of interaction-awareness	0.30	0.77
Overall Believability	0.27	0.76

6.3 Data Collection and Analysis

This believability analysis was conducted based on the conversations with aware and unaware agents - we wanted to evaluate how believable our agents would be if we supplied them with awareness features. The experiment was based on the three awareness believability dimensions we have identified: (1) environment-awareness; (2) self-awareness; and (3) interaction-awareness. Those who lack these features are regarded here as unaware agents. We divided our participants into two different groups (both groups were immersed in our case study). The first group conversed with aware agents, the second group – with unaware agents. The experiments were conducted over a two week period. After cleaning the incomplete responses, the data that was analysed included the responses of 22 participants - 11 per group.

Table 2 shows a summary of the results for comparing aware and unaware agents. The overall believability index for aware agents was 0.76 vs 0.27 for unaware agents. The comparison along our awareness believability dimensions shows that for environment-aware queries in 76% of the cases participants perceived aware agent as human vs only 22% of misclassification of unaware agents as humans. For queries about agent's own goals, plans, actions etc., aware agents were ranked as human 75% of the times vs 26% of misclassification of unaware agents as humans. In the case of interaction-awareness, aware agents were believed to be humans for 77% cases vs 30% of misclassification of unaware agents as humans. These results indicate that it was relatively difficult for participants to differentiate between aware agents and humans based on their conversations.

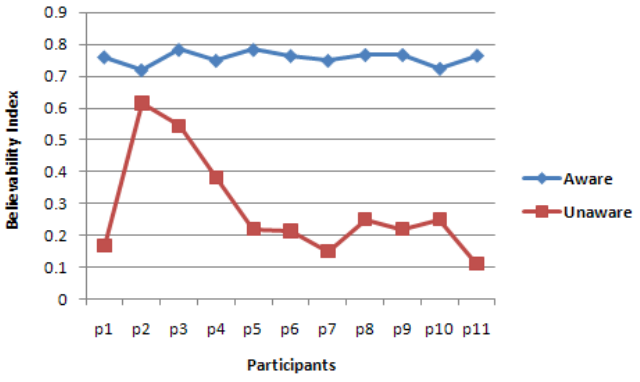


Figure 7: Believability for aware vs unaware agents.

Figure 7 shows the individual believability evaluations, which range within [0.72; 0.78] across participants dealing with aware agents and varied substantially for unaware agents. As our unaware agents were based on AIML rules, on early stages of a conversation they managed to give an impression of a real person, but later in the conversation participants correctly identified them as artificial.

To test believability of aware and unaware agents' conversations, the recorded conversations were partitioned into three categories: environment-, interaction- and self-awareness. Figure 8 shows a very high human-likeness (0.6 to 0.8) for each category in our aware agents.

7 Conclusion

This paper has analysed the notion of believability and believability characteristics of embodied conversational agents interacting with a user in a virtual world. The identified features of environment-, self- and interaction-awareness believability have not been addressed in previous works. To show that these features are indeed responsible for improving the overall believability of virtual agents we conducted an experiment where two groups of participants evaluated the believability of two different types of agents: those that incorporate the aforementioned features and those that don't. The study showed that when embedded, each feature on its own (as well as the combination of all three features) improves the perception of agents as more believable.

References

Bickmore, T. W., Mauer, D. & Brown, T. (2007), Context awareness in mobile relational agents, *in* 'IVA '07: Proceedings of the 7th international conference on Intelligent Virtual Agents', Springer-Verlag, Berlin, Heidelberg, pp. 354–355.

Bogdanovych, A. (2007), Virtual Institutions, PhD thesis, UTS, Sydney, Australia.

Bogdanovych, A., Rodriguez-Aguilar, J. A., Simoff, S. & Cohen, A. (2010), 'Authentic Interactive Reenactment of Cultural Heritage with 3D Virtual Worlds and Artificial Intelligence', *Applied Artificial Intelligence* **24**(6), 617–647.

Cunningham, D. W., Kleiner, M., Wallraven, C. & Bülthoff, H. H. (2005), 'Manipulating video sequences to determine the components of conversational facial expressions', *ACM Trans. Appl. Percept.* **2**(3), 251–269.

Dautenhahn, K., Ogden, B. & Quick, T. (2003), 'From embodied to socially embedded agents- implications for interaction-aware robots.'.
URL: *http://hdl.handle.net/2299/3057*

Doyle, P. (2002), Believability through context using "knowledge in the world" to create intelligent characters, *in* 'Proceedings of AAMAS '02: Conference', ACM, New York,USA, pp. 342–349.

Elpidorou, A. (2010), 'Alva noë: Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness', *Minds Mach.* **20**(1), 155–159.

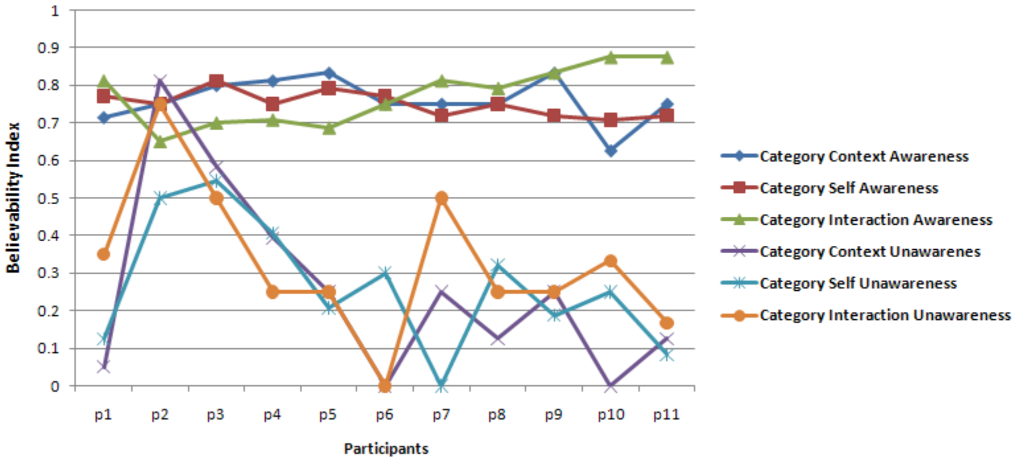


Figure 8: Believability index for each category.

Esteva, M. (2003), Electronic Institutions: From Specification to Development, PhD thesis, Institut d'Investigació en Intelligència Artificial (IIIA), Spain.

Gandhe, S. & Traum, D. (2007), Creating spoken dialogue characters from corpora without annotations, *in* 'Proceedings of Interspeech-07', pp. 2201–2204.

Gerhard, M., Moore, D. & Hobbs, D. (2004), 'Embodiment and copresence in collaborative interfaces', *Int. J. Hum.-Comput. Stud.* **61**(4), 453–480.

Gerhard, M., Moore, D. & Hobbs, D. (2005), 'Close encounters of the virtual kind: Agents simulating copresence', *Applied Artificial Intelligence* **19**(3-4), 393–412.

Gorman, B., Thurau, C., Bauckhage, C. & Humphrys, M. (2006), Believability Testing and Bayesian Imitation in Interactive Computer Games, *in* 'Proceedings of the 9th International Conference on the Simulation of Adaptive Behavior (SAB'06)', Vol. LNAI 4095, Springer, pp. 655–666.

Hallowell, A. I. (1955), *Culture and Experience*, University of Pennsylvania Press, Philadelphia, USA.

Hartmann, B., Mancini, M. & Pelachaud, C. (2005), Implementing expressive gesture synthesis for embodied conversational agents, *in* 'Gesture Workshop', pp. 188–199.

Heylen, D. K. J., van Es, I., Nijholt, A. & van Dijk, E. M. A. G. (2005), Controlling the gaze of conversational agents, *in* J. van Kuppevelt, L. Dybkjaer & N. O. Bernsen, eds, 'Natural, Intelligent and Effective Interaction in Multimodal Dialogue Systems', Kluwer Academic Publishers, pp. 245–262.

Johnson, W.L., R. J. & Lester, J. (2000), 'Animated pedagogical agents: face-to-face interaction in interactive learning environments', *International Journal of Artificial Intelligence in Education* **11**, 47–78.

Lee, J., Marsella, S., Traum, D., Gratch, J. & Lance, B. (2007), The rickel gaze model: A window on the mind of a virtual human, *in* C. Pelachaud, J.-C. Martin, E. Andr, G. Chollet, K. Karpouzis & D. Pel, eds, 'Intelligent Virtual Agents', Vol. 4722 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, pp. 296–303.

Lester, J. C., Voerman, J. L., Towns, S. G. & Callaway, C. B. (1999), 'Deictic believability: Coordinating gesture, locomotion, and speech in lifelike pedagogical agents', *Applied Artificial Intelligence* **13**, 383–414.

Livingstone, D. (2006), 'Turing's test and believable AI in games', *Computer Entertainment* **4**(1), 6.

Loyall, A. B. (1997), Believable Agents: Building Interactive Personalities, PhD thesis, Department of Computer Science, Carnegie Mellon University, Pittsburgh.

Mateas, M. (1997), An oz-centric review of interactive drama and believable agents, Technical report, Carnegie Mellon University , Pittsburgh, PA, USA.

O'Keefe, J. & Nadel, L. (1978), *The Hippocampus as a cognitive map*, Oxford University Press.

Pantic, Maja Member, S. & Rothkrantz, L. J. M. (2000), 'Automatic analysis of facial expressions: The state of the art', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(12), 1424–1445.

Prendinger, H. & Ishizuka, M. (2001), Social role awareness in animated agents, *in* 'AGENTS '01: Proceedings of the fifth international conference on Autonomous agents', ACM, New York, NY, USA, pp. 270–277.

Riedl, M. O. & Stern, A. (2006), Failing believably: Toward drama management with autonomous actors in interactive narratives, *in* 'TIDSE', pp. 195–206.

Thiebaux, M., Lance, B. & Marsella, S. (2009), Real-time expressive gaze animation for virtual humans, *in* 'AAMAS (1)', pp. 321–328.

Thomas, F. & Johnston, O. (1981), *Disney animation : the illusion of life*, 1st ed. edn, Abbeville Press, New York.

Wallace, R. (2004*a*), 'The elements of AIML style. AI Foundation.'.

Wallace, R. (2004*b*), 'The Anatomy of A.L.I.C.E', *A.L.I.C.E. Artificial Intelligence Foundation* .

Weizenbaum, J. (1966), 'Eliza—a computer program for the study of natural language communication between man and machine', *Communications of the ACM* **9**(1), 36–45.