AISB Journal **1**(1), © 2001
http://www.aisb.org.uk

Designing Agents for a Virtual Marketplace

Kaveh Kamyab, Frank Guerin, Petar Goulev and Ebrahim Mamdani

Department of Electrical and Electronic Engineering,
Imperial College of Science, Technology and Medicine
London, SW7 2BT, United Kingdom
k.kamyab@ic.ac.uk ; f.guerin@ic.ac.uk ;
p.goulev@ic.ac.uk ; e.mamdani@ic.ac.uk

Abstract

This paper is concerned with the design of animated agents for applications in e-commerce virtual marketplaces. This brings challenges such as personalisation, believability, conversational skills and multimodal interaction. We have developed a multi-agent architecture for a virtual sales assistant. Agents use high level abstractions for representing, reasoning and planning human-like dialogs. These include mental abstractions such as belief, desire and intention as well as social abstractions such as role relationships and commitments. We design interactions through the development of protocols which specify conventions of behaviour and endow agents with social awareness. Fuzzy modelling is used to model a user's preferences and subjective perception of product attributes. This synergy between existing technologies has the potential to fulfil the requirements for the next generation of virtual marketplaces.

1 Introduction

A portal is an access point to internet resources and services. The resources and services provided can include, for example, a directory of services, a search facility, news, weather information and stock quotes. Portals also provide personalisation, changing the content and appearance to suit the needs of the user. Current portals such as Yahoo, Excite and Netscape, present the user with a 2D interface. However, a great deal of effort is being placed on developing 3-D virtual environments, for example virtual marketplaces and chat rooms, that the next generation of portals could offer. Users would be able to navigate through this virtual space and enter shops to inspect goods and interact with other users (via their avatar) or with embodied software agents. The technology required to implement these embodied agents includes the realistic animation of synthetic faces and bodies as well as adequate conversational skills. This integration of technologies is being undertaken in the SoNG project (PortalS Of Next Generation). Our project partners have been working on designing the virtual world, the user interface and the facial and body graphics and animation (for synthetic characters). The synthetic characters are driven by agent technology; the design of the agents is the subject of this paper. SoNG is ongoing and is currently in its final integration phase. In this article we comment on our architecture for the agents, in particular how the inputs to an agent are interpreted according to conversation protocols and appropriate output is generated in the form of speech and animations for the agent. In addition we outline how a user profile is built to model the user's preferences in terms of products and interaction.

Within the virtual marketplace of SoNG several demonstration applications shall be developed including a theatre booking system, a clothes shop and a telephone shop. Here we report only on the latter. This is primarily considered to be an online information site, where the user has access to several services, which present typical telecommunication equipment. Products will be presented by an Embodied Conversational sales Agent (ECA). This paper is primarily concerned with the design of ECAs with particular emphasis on providing enhanced interaction to human users.

Animated embodied agents capturing some of the emotional qualities and related behaviours associated with humans are being increasingly employed in the development of a new generation of interfaces for a wide variety of applications (e.g. online tutoring, storytelling, help, e-commerce and so on). There are many researchers who are advancing this new generation, André and Rist (André et al., 2000), Bates (Bates, 1994), Lester (Lester et al., 1997b) and others. Embodied agents promise to improve the user's subjective experience while interacting with computer applications. These improvements include providing the user with intuitive multimodal interfaces consisting of natural language understanding and generation augmented by the use of realistic facial and body animations. In addition, the use of affect, "computing which relates to, arises from or deliberately influences emotions" (Picard, 1997), can play a key role in ensuring that such characters are able to generate animation appropriately selected for the emotional context of the interaction. Finally, personalisation of services enables the simplification of the interface by adapting behaviour to the user and presenting the user with information which he or she is genuinely interested in.

The use of embodied agents introduces a whole new set of social implications. In particular, a debate has flared up between advocates of embodied interface agents and advocates of Direct Manipulation (DM) interfaces. The latter maintain that over anthropomorphising can result in the user generating false mental models and false expectations of the actual capabilities of the system (Norman, 1997), (Maes and Shneiderman, 1997), (Dehn and van Mulken, 2000). This is due to a human tendency to respond socially to agents perceiving them as competent and capable (Reeves and Nass, 1996). In addition, DM advocates believe that the use of animated graphics can be more of a *distraction from* rather than an *enhancement of* interaction. This would lead to more time-consuming applications and less user satisfaction. On the other hand, embodied agents may help lower the 'getting started' barrier by allowing for the emulation of human-human communication (Adelson, 1992). Furthermore, some compelling results have been found supporting the positive impact embodied agents have on factors such as likability, engagingness, perceived intelligence, and many more. The relationship between users and applications is discussed in section 2.1.

Aside from the presumed advantages of using embodied agents, they introduce a multimodality into the interface that provides flexibility to the user, but also introduces considerable design and implementation overheads to the designers of such systems. In particular, we consider contextually aware, believable agents with the capability of conversing with the user via natural language. Thus support for context sensitivity, synchronised face and body animation and Text To Speech (TTS) output and dialogue capabilities must be considered at the design stage. In this paper we describe a design approach which addresses these issues, linking agent behaviour, user dialogue and personalisation. Communication is considered from a social perspective through the design of interaction protocols.

Here is a short description of the content. In Section 2 we give a brief overview of the issues arising in the design of agents with personality and personalisation. We also discuss issues of particular relevance in an e-commerce application. In Section 3 we give an overview of the agent architecture employed. In Section 4 we describe how interactions are designed, beginning with a storyboard and progressing to an interaction protocol. Section 5 describes the low level implementation details and includes a short walkthrough describing how the various different components interact to produce a behaviour. Section 6 discusses some related work. Finally, Section 7 outlines issues that must be tackled in the future and then summarises and draws some conclusions.

2 Agents with Personality and Personalisation Know-How

The use of ECAs in an e-commerce application such as the one developed on SoNG requires designers to consider how agent behaviour must be correlated with the underlying functionality of the e-commerce system. In addition, how will agents behaviour be generated and how might this effect the user of the system?

The agent's function is to provide users with personalised product suggestions matching their needs. In order to provide this information, the agent must learn about user's preferences in terms of the products available. The agent's behaviour must also be appropriately selected to support the function of a shopkeeper. As shown in section 2.1, there is evidence that a correlation exists between users' perception of agent behaviour (e.g. quality and frequency of interactions and animations) and their liking of and trust [1] in the underlying system. We can think of such behaviour as being a function of the agent's 'personality'. This result may call for the need to personalise the agent's personality to the user's preferences.

2.1 User Modelling and Personalisation

User modelling is a technique used to build and maintain user profiles containing information about the user which is relevant to the application domain. User profiles may contain user data such as name, gender and address, but they may contain what the system presumes are the user preferences, knowledge and goals.

The importance of user modelling emerged throughout the eighties largely from research in dialogue systems. One of the earliest such systems was ELIZA, which responded to keyword matching. Often its human communication partners assumed they were communicating remotely over Teletype with another human (Weizenbaum, 1966). However, ELIZA didn't take context into account, which sometimes caused the system to generate inappropriate responses when it was unable to "understand" the user.

At a later stage in the development of dialogue systems it was recognised that a system should generate assumptions about the user that may be relevant to the task domain at hand. In particular these include the user's goals, the plans with which the user intends to achieve these goals and the knowledge or beliefs the user has about the domain. The assumptions were to aid the system to converse more naturally (e.g. to supply additional relevant information, avoid redundancy in answers and explanations and detect the user's wrong beliefs about the domain (Kobsa, 1990)).

Throughout the 70's and 80's another AI application that was generating a lot of interest was that of Intelligent Tutoring Systems (ITS). ITSs' predecessors, Computer Assisted Instruction (CAI) systems, already recorded and evaluated the student's interactions, but no emphasis was put on educational psychology or, more importantly, the student's level

[1] Here we use the everyday definition of the word trust, i.e. the user's reliance on the ability and integrity of the system.

of expertise. (Sleeman and Brown, 1982) first coined the term ITS to distinguish these systems from the traditional CAIs. ITSs brought the concept of knowledge representation to student modelling to aid the tutoring process.

In addition, research on intelligent help systems, game playing and in the field of Human-Computer Interaction (HCI) also identified the need for user modelling. As a result of such a proliferation of use, user modelling evolved into a discipline in itself. More recently it has been applied to the broad discipline of personalisation, where services and application content are tailored in order to satisfy user preferences. These include Information Retrieval (IR) and Filtering (IF) systems, customisable interfaces and so on.

Despite its promise to solve the development issues in many disciplines, user modelling came under a lot of criticism in the early 90's for failing to deliver (McCalla, 1992) and this was attributed to a number of causes. Intrusiveness of data gathering methods, privacy issues and failure to take the wider user context into account resulted in systems that often made erroneous assumptions and couldn't win over user confidence.

However, utilizing personalisation and the underlying 'one-to-one' marketing paradigm is of great importance for today's businesses (Peppers and Rogers, 1993). For example, good sales assistants have the ability to make accurate assumptions about the user and use the information gathered to tailor their service. Furthermore, regular customers receive an even more personalised service, not only because of their tastes, but because the sales assistant knows how best to interact with them. For instance, a good restaurant waiter is one who knows which is your favourite table and knows what your favourite dish is. However, this knowledge comes from passive observation of the customer's behaviour and it comes with the awareness that some people like to choose their own table and change dish at every visit. In order to provide this quality of service, information must be gathered regarding the customer's tastes, background, personality and habits. Not surprisingly, great advances have been made in the area of personalisation for use on e-commerce websites (e.g. Amazon.com, CDnow, eBay, etc) and there are indications that there are substantial commercial benefits (Fink and Kobsa, 2000). Thus a user-modelling component is an essential part of an e-commerce system such as the SoNG application. However, measures need to be taken to avoid the pitfalls highlighted in previous user modelling systems. In particular, we need to address issues of data collection, privacy and the wider user context.

2.1.1 Data Collection and Privacy

These first two issues are intrinsically related. Acquiring data about a user is not a trivial task. A lot of data can be acquired from a user's conversation and by observing a user's behaviour (Orwant, 1996). (Kay, 1993) identifies three distinct categories of user model information.

1. The information *explicitly given* by the user to the system, either by request or voluntarily. For example the user states that he/she is a novice user of the system.

2. The information gathered by *observation* including monitoring user actions and dialogue history.

3. The information that the user has been *told about* and hence is aware of.

Indeed, requesting information about a user *from* the user can have extremely detrimental effects. Firstly, it may appear to be an invasion of the user's privacy. Users may be wary of how and by whom the information will be used. Secondly, users actually have a poor ability to describe their own preferences. Often, what they say they like and what they really like are not the same thing. Finally, as stated in the paradox of the active user (Carroll and Rosson, 1987), people are more motivated to start *using* things than to take the initial time to learn about them or to set up a lot of parameters. So information should be gathered through the observation of user interaction (what does the user ask for, what does the user need) or if a user explicitly states he or she likes a specific product attribute.

A non-invasive method of data collection has the advantage that users are less prone to be aware of an invasion of their privacy. This helps maintain a relationship of trust between the user and the agent. In any case, a user will be more willing to reveal information to an entity it trusts rather than one it mistrusts. The user's willingness to give up his/her private information helps a user modelling system, but does not address the problem of privacy. Deeper considerations must be made. Many commercial websites are beginning to self-regulate their own handling of confidential user information by adhering to and contributing to initiatives such as (TRUSTe, 2001) and (P3P, 2001). This shows that privacy must be guaranteed in order to attract customers establishing relationships of trust.

2.1.2 User Context

Just as the first dialog systems were criticised for not taking user context into account at all, many user modelling systems can be criticised for not taking *enough* of the user context into account. User needs and interests depend on many parameters including the personal and the sociological (personality, momentary needs, moods, etc) and they change over time (Hanani et al., 2001). Traditionally, the context modelled has been strictly that relating to the application. In Information Filtering systems user modelling has been content-based, in ITS systems it has been knowledge-based, in e-commerce systems it has been product-based, and so on. However, recently this trend began to curb. For example, the realisation that learning is greatly dependant on motivation and emotions has lead to the application of Affective User Modelling (AUM) to the area of ITS (Elliot et al., 1999), (Paiva and Martinho, 1999). It is possible that AUM can be applied to e-commerce systems too (a satisfied customer is a happy customer, or more to the point, a happy customer is a satisfied customer), but its use needs to be evaluated. The difficulty lying not only in gathering the data, but also in knowing what to do with it. It is clear, though, that each application domain necessitates its own unique considerations with regard to the content of the user model.

Due to its deeper social implications, the use of embodied agents introduces a new set of dimensions that weren't considered for DM interfaces. These dimensions are bound to have an effect on the user models employed. Many studies have been carried out which attempt to evaluate the effect of likability, engagingness, perceived intelligence, believability, adaptability and motivational qualities of an embodied agent on a user (Koda and Maes, 1996), (Moon and Nass, 1996), (King and Ohya, 1996), (Lester et al., 1997b), (Charlton et al., 2000). Some compelling results have been found supporting embodied agents. (Koda and Maes, 1996) found that the use of animated faces in an interface improved the likability of the system and engaged users more leading them to attribute higher levels of intelligence to the agent. The results were more pronounced if the face used a realistic animation of a human face. (Charlton et al., 2000) found a strong relationship between likability and perception of intelligence (the more a system is liked, the more it is perceived to be intelligent) indicating a need for personalisation of the embodied agent. In fact, (Moon and Nass, 1996) suggest that user's prefer adaptable animated interfaces that become more similar to the user over time, a result which perhaps stems from the natural human tendency to present oneself in a manner which pleases the person we are interacting with (Sproull et al., 1996). (King and Ohya, 1996) found that animated agents were more likely to be seen as intelligent, but this could give a false impression of the real system intelligence. Take, for example, the SoNG embodied agent. The agent's personality and behaviour can influence the user's perception of it and the service it provides. Thus, as found by (Charlton et al., 2000), if a user likes the character it is interacting with he or she will consider it to be more intelligent and competent and as a result the user will trust the agent more. It is thus important to model the user's preferences in terms of personalities and behaviours which they like to interact with.

2.2 Personality

Personality is widely studied in human psychology, yet no single defining theory exists. Two broad approaches are taken when defining personality: explanatory and descriptive. The latter and most popular approach concentrates on creating taxonomies of lexical descriptions of behavioural traits, generally associated with personality, which can help identify individual differences between people. This is referred to as the trait theory. The lexical approach assumes that socially relevant personality characteristics have become encoded in our language (Fujita, 1996) and is hence limited to the description of surface behaviour rather than the underlying mechanisms of personality. Many variations in the trait theory have been studied such as the Big Five (John, 1990) and the Giant Three (Eysenck, 1991). In the SoNG project we make use of a subset of the Big Five theory, a description of which is given below. For an example of explanatory approaches to defining personality see (Moffat, 1997).

2.2.1 The Big Five Theory:

The most widely used of the aforementioned theories is the Big Five, which is generally agreed to define the following categories:

1. Extroversion: activity and energy level traits, sociability and emotional expressiveness.

2. Agreeableness: altruism, trust, modesty, prosocial attitudes.

3. Conscientiousness: Impulse control, goal directed behaviour.

4. Neuroticism: emotional stability, anxiety or sadness.

5. Openness: Breadth, Complexity, and depth of an individuals life.

Although the names given to the five categories vary according to different theorists, their meanings remain largely unchanged. (John, 1990) presents a survey of a range of approaches to the Big Five theory and the variations in terms used.

The Big Five theory has been successful in categorising personalities empirically by the use of a personality questionnaire, although its success is only measured relative to other taxonomies of personality. In fact, of 300 adjectives used in the personality check only 112 have been successfully associated with one of the five categories (John, 1990). Thus large parts of the personality sphere are not considered by the theory. Other problems highlighted by (Fujita, 1996) include high correlation between the five categories and, most importantly, the criticism that the theory may only be the structure of personality lexicon and not personality. However, no competing theories have proved to be any more complete than the Big Five theory. In fact, theories such as the Giant three are essentially a subset of the Big Five, where Agreeableness and Conscientiousness are replaced by Psychoticism and Openness, sometimes linked with intelligence, is omitted. In particular, the Big Five theory is a progressive research program and has been subject to a much greater extent of practical experimentation than other theories, making the Big Five theory the most widely used.

The many taxonomies and theories of personality leave us in the dark when it comes to stating what personality is. According to (Picard, 1997), personality can be thought of as a predisposition governing the likelihood of focused emotional responses. The stance taken on SoNG is that personality accounts for consistent patterns of behaviour (Pervin and John, 1993) and thus has a bearing on the agent's behaviours. We use a subset of the big five traits to describe the personality of embodied agents, which includes Extraversion (activity, sociability and expressiveness), Agreeableness (prosocial attitudes) and Openness (depth of the agents background encapsulated in a persona). Conscientiousness and Neuroticism are not considered as the agent must exhibit impulse controlled, emotionally stable behaviour.

3 Architecture Overview

Here we describe the general architecture employed for our agents. There are three different agents involved in a typical interaction with the user: an embodied sales agent, a disembodied search agent and a disembodied database agent. The user sees only one, the embodied sales agent. The user may ask for a product in natural language, the sales agent then sends a request to a search agent who queries one or more database agents and returns with the required information (see Figure 1). The sales agent is the most complex since it will interact with the virtual world to find out about changes in the world (such as a user speaking or moving) and to effect changes (facial and body animations and speech output for the agent's virtual body).

As depicted in Figure 1, the architecture of agents is divided into three parts, these are clearly labelled for the sales agent. Firstly there is the *Observation* part (top left) which is responsible for interpreting natural language from a human user and messages from other agents through the *Event Interpreter*. It also interprets other events in the virtual world such as the movement of the human's avatar. The interpretation of these events is recorded as a change to the agent's *Mental State* (which constitutes the second part). The mental state contains an *Internal State* which uses an explicit representation of beliefs, desires and intentions (Rao and Georgeff, 1992). Among the agent's beliefs are beliefs about the beliefs, desires and intentions of other inhabitants of the virtual world. For example the preferences of the human users (see section 5.0.1). In addition to representing its internal mental attitudes, an agent also maintains a copy of the *Social State*. The social state represents common knowledge and endows the agent with an awareness of its social context (see section 4.1). Finally, the *Action Execution* part uses the current mental state to decide what actions the agent should do next. In addition to selecting actions the *Action Planner* may also modify the mental state, for example, it may decide to add a new desire or intention.

The architecture of the search and database agents contains some of the same modules as the sales agents, but does not require those modules which deal with human interaction. The modular architecture allows each of our agents to be built with re-usable components. For example, both embodied and disembodied agents will use the same module for interpreting the meaning of events in the world, but embodied agents will use a natural language processor while disembodied ones will not. The agents communicate

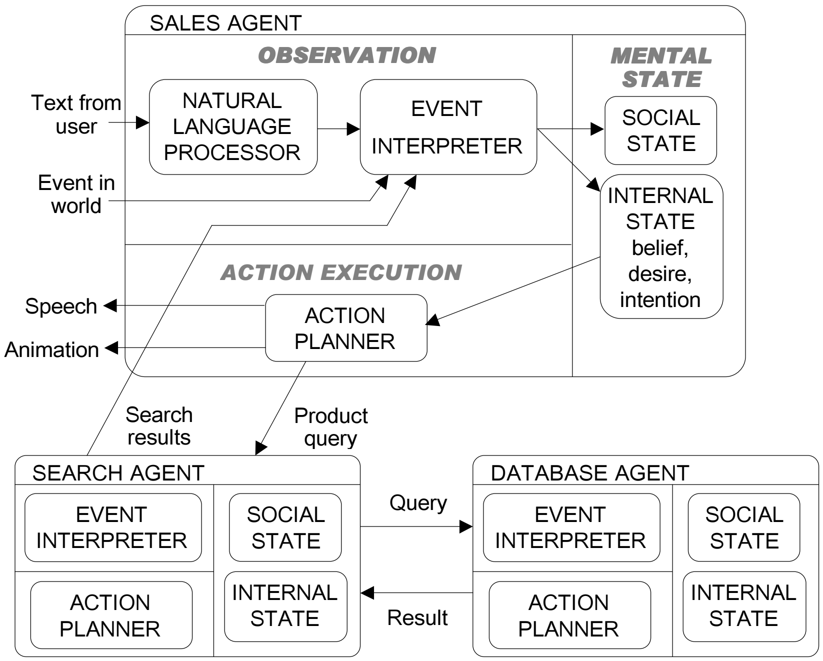

Figure 1: Modular Agent Architectures.

among themselves using an agent communication language (ACL). The natural language processor translates a natural language sentence into an ACL message so that the sales agent's event interpreter treats this messsage in the same way as a message from another agent. An aim of this design is to raise the level of agent-agent interactions. An agent will process a message from another synthetic agent in the same way as a message from a human user, considering its effect on mental attitudes and also the social perspective of the society of agents. Thus agent communication involves high level abstractions and is not simply the exchange of procedural directives. This is important since we are designing agents for an e-commerce application and this technology might be applied in scenarios where different vendors contribute their own database agents. In such competitive markets self interested agents might not be trusted to give the search agent the lowest price for example (Singh, 1998).

4 Designing Interactions

Our embodied agents interact with humans in the virtual world by observing natural language text input and responding using both verbal and non-verbal communication. We wish to go beyond a simple stimulus-response behaviour with keyword matching. In order for an artificial agent to be able to communicate effectively with a human it must be able to interpret the meaning of incoming communications, update its mental state and plan its own communications. In fact it must understand something of the human-level meaning of the communication. Clearly, achieving a human level understanding is beyond our reach, but we can take some steps towards this goal by identifying certain high level mental abstractions (such as belief, desire and intention) and social abstractions (such as role relationships and commitments) which can be used to characterise the meaning of human communications. Thus the semantics of communications is defined in terms of these high level abstractions. The agent maintains an explicit representation of these abstractions (in its mental state) and reasons using this.

In addition, the interpretation and planning of communication should be sensitive to the context. We concentrate on giving our agents an awareness of three aspects of context as follows:

1. The domain of discourse: The agent will be aware that both the communications it observes and those it generates should be relevant to the shopping domain in which it is situated.

2. The history of the discourse: The agent maintains a record of the discourse which helps it to interpret and plan communication. For example, if the agent has just proposed to demonstrate a product for the human and the human responds positively, then the agent can interpret this response as an acceptance of the proposal.

3. The social roles occupied by the conversational partners: The agent is aware of its role as a sales assistant and its associated obligations, for example it must greet a human user who enters the shop and it must respond to queries.

4.1 Developing Social Rules for an Interaction

We begin our development by inventing storyboards for interactions. A story can be analysed and generalised to a class of interactions which follow a similar pattern. This pattern constitutes a protocol for the interaction which can be encoded in terms of a set of rules which describe the social conventions governing such interactions. These rules can be encoded logically for our sales agents and will constrain the agent's planning process while it is executing such a conversation. In this way we simplify the choices available to an agent and ensure that it will behave in a way compatible with the expectations of the human users who are accustomed to dealing with socially aware individuals.

We identify roles for each participant, in our story (see Figure 2) there are two, the sales agent is playing the role of *Shopkeeper* and the human user is playing the role of *Customer*. We chart out the paths the conversation can take as shown in the protocol diagram (see Figure 3), taking care to reuse existing states as much as possible. The diagram shows a UML-style statechart diagram where states are states of the conversation (partial social states) and the arcs between them describe the action in the virtual world which effects the state change. The lower part of each state bubble lists actions to be performed whilst in that state.

We now develop rules to describe the effects of communication. It has long been recognised by philosophers that the rules governing communication are *constitutive rules*. Searle describes these rules as follows

> "Constitutive rules constitute (and also regulate) an activity the existence of which is logically dependent on the rules." (Searle, 1965)

These rules take the form "X counts as Y", for example making a promise counts as an undertaking to do some action simply because this is the convention of usage. Thus it is a social convention that a certain speech act has a certain meaning, these conventions have evolved over time. It is Searle's contention that

> "…the semantics of a language can be regarded as a series of systems of constitutive rules…" (Searle, 1965)

<*user enters shop*>	
Phone Agent :	"Good Morning John." "Welcome back to the phone shop." "We have some new phones for you: the Nokia 6110 and the Nokia 3210." "Would you like to see them?"
User :	"Yes, the Nokia 6110 please."
<*Demonstration*>	
User :	"hhhmm. . . I would like something a little lighter"
Phone Agent :	"How about the Nokia 3330?"
<*Demonstration*>	
User :	"How much does it cost?"
Phone Agent :	"It costs 150 pounds."
User :	"OK, I'll take it."
Phone Agent :	"One Nokia 3330 has been added to your shopping cart."
User :	"Do you have any matching leather case ?"
Phone Agent :	"No, I'm sorry, we don't sell leather cases." "However, we have some nice faxes."
User :	"No thanks." "Bye!"
Phone Agent :	"Good-bye! I hope you come back soon."

Figure 2: A Storyboard for an Interaction in the Phone Shop.

The idea of basing the semantics of communication on conventional meaning is now receiving attention in the field of multi-agent systems (Singh, 1998). The necessity to endow agents with social role awareness (Prendinger and Ishizuka, 2001) and to represent social commitments (Traum and Rickel, 2001) has also been recognised by those designing embodied agents for virtual worlds.

Our approach uses an explicit representation of the *social state* which represents social facts that are currently true. Social facts include commitments and *expressed mental attitudes* (Guerin and Pitt, 2000). The notion of expressed mental attitudes allows us to distinguish between an agent's publicly expressed mental attitudes and its personal internal mental attitudes which need not be identical. For example, if an agent makes a *request* we will say that the agent has expressed a desire, but we do not really know if the agent has that desire internally. Thus the semantics of our speech acts are defined as social rules which capture the conventional meaning of communicative acts in our virtual society and describe the social relations (for example roles and commitments) that hold between its inhabitants. Table 1 lists the speech acts that will be needed and their semantics.

The left side of Table 1 gives the speech act name only (we assume the act is successfully performed and contains a sender, receiver and content) the right side gives the semantics of the act. The expressions on the right are written in a social facts language. E-DESIRE and E-BELIEF denote a publicly expressed desire and belief respectively (the E- pre-

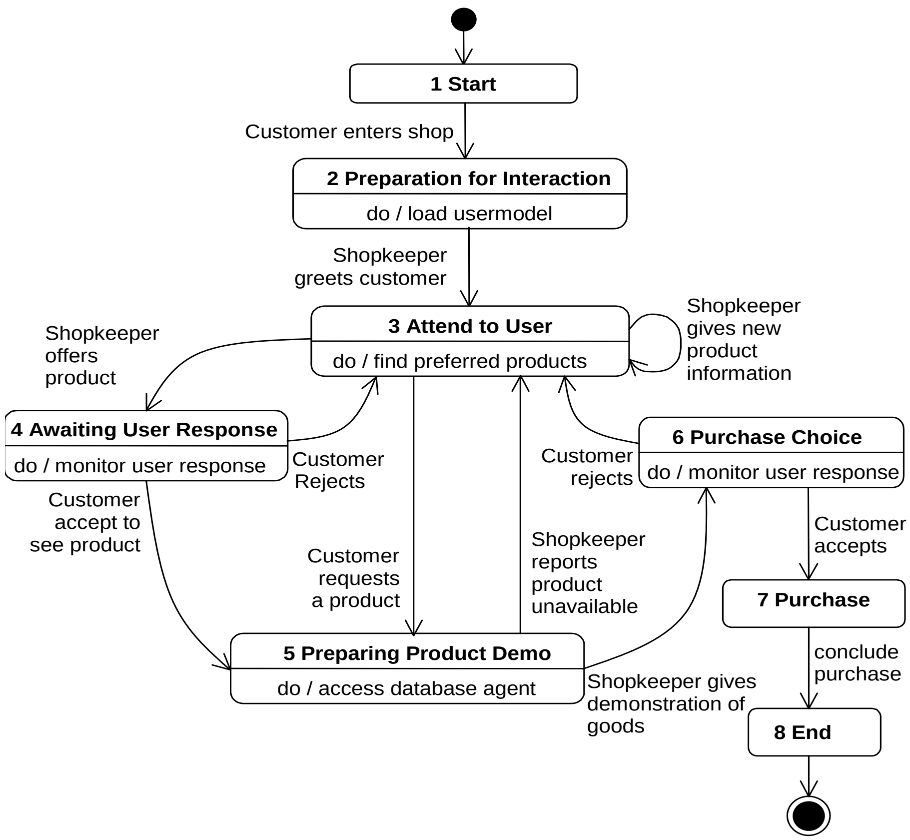

Figure 3: Shopkeeper-Customer Protocol.

Speech Act Name	***Expressed Mental Attitude***
query	E-DESIRE *sender* KNOW *sender content* *proposition*=*content*
request	E-DESIRE *sender* DO (*receiver*, *content*)
offer	E-WILLING *sender* DO (*sender*, *content*) E-DESIRE *sender* KNOW *sender* DESIRE *receiver* DO (*sender*, *content*)
inform	E-BELIEF *sender content*
confirm	E-BELIEF *sender proposition*
disconfirm	E-BELIEF *sender* NOT *proposition*
salute	E-HAPPY *sender* DONE (*receiver*, *enter shop*)

Table 1: Speech Act Semantics

fix stands for *expressed*). E-HAPPY denotes an expression of the state of happiness. E-KNOW is for knowledge, meaning the agent believes that a proposition is true or that it is false. E-WILLING describes the willingness of an agent x to adopt an intention provided that another agent y desires it. The semantics describes the facts that are added to the social state after the performance of the speech act. Both the sender and receiver of a message maintain a copy of the social state and update it as messages are sent or received. The state is updated first with the speech act semantics and secondly with the protocol rules (see section 4.2). The terms *sender*, *receiver* and *content* are replaced by their values in the speech act. The **query** speech act adds to the social state an expressed desire of the sender to know if the content is true. It also assigns the term *proposition* a certain value in the social state. When this term occurs in the semantics of an act it is replaced by the existing social fact. Thus the acts **confirm** and **disconfirm** (which need no content) are only meaningful after a **query**. The **offer** speech act adds two propositions to the social state. Firstly the sender expresses a willingness to do the action specified in the content. Secondly the sender expresses a desire to know if the receiver desires this action to be done. Where *content* appears within a DO (...) or DONE (...) expression it is an expression in the action language described in the Appendix, otherwise it is a true or false expression in the social facts language. This type of semantics can be formalised using denotational semantics (Guerin and Pitt, 2001).

Note that the social state is merely a set of propositions which can be represented by text strings. There is no requirement that these propositions are logically consistent and logical implication does not follow automatically, it must be explicitly specified by a rule if desired. The absence of a proposition from the social state does not entail the validity of its negation. Thus if an agent does not believe a proposition or its negation, its attitude with respect to that proposition is in an undefined state. This can be formalised with a three valued logic, but in this paper we focus only on implementing a system. The propositions in the social state will be explicitly used by the protocol rules to define new states when a communication happens.

4.2 Protocol Constraints

The speech act semantics above define speech acts as changes to the social state. The protocol can define a set of commitments for each conversational participant based on the current social state. An analysis of our storyboards leads to the identification of those social facts which are relevant to a description of the state of the conversation and hence constrain the possible future paths. The social facts identified in different storyboards can combine in different ways to describe the social state for a more general class of conversations. Since the total number of possible social states can be large we restrict our attention to partial states of the conversation. A partial state describes some aspect of the conversation which is relevant to the commitments that arise in a given state. In this way we generalise from our stories to produce a set of rules which describe the social conventions governing such interactions. These rules are encoded logically for our sales agents and constrain the agent's planning process while it is executing such a conversation. This simplifies the choices available to an agent and attempts to make it behave in a way compatible with the expectations of human users. The rules are described informally in Table 2. The left hand side shows the condition characterising the partial social state and the right hand side describes the social commitments which arise in such partial states. All these commitments are for the Shopkeeper, in this protocol we cannot constrain the actions of the human Customer. Each entry on the right hand side should be prefixed with "COMMITTED *in role:Shopkeeper*" meaning that the agent playing the role of Shopkeeper is committed.

Our protocol is initiated by a non-speech action, the entrance of a user avatar in the shop. When the interaction begins there is a blank social state and this creates a commit-

State Condition	***Social Commitments Arising***
Initial (empty) state	greet the Customer AND inform the Customer that this is the phone shop
elapsed time exceeds *limit*	give new product information OR offer to demo a product for the Customer
Customer E-DESIRE to know about the availability of a product	(disconfirm AND offer an alternative) OR (confirm AND demo the product)
Customer E-DESIRE product demo	demo the product
Customer E-DESIRE to purchase	inform that the product was added to cart

Table 2: Speech Act Semantics

ment for the Shopkeeper to express happiness that the Customer has entered and to tell the Customer that this is the phone shop. Included in the social state is a timer variable, each agent's copy of this fact is periodically updated by the agent platform. Another social fact *elapsed* keeps track of how much time has elapsed since the last speech act in the conversation. As shown in the table, if this value exceeds a certain predefined *limit* then the Shopkeeper is committed to perform a speech act so as to avoid a long silence. In some states the Shopkeeper is committed to more than one action, for example if the Customer queries about the availability of a product and the Shopkeeper decides that it is available then the shopkeeper is committed to confirm the availability and offer to give a demo.

5 Implementation

The implementation is built using JADE (JADE, 2000), a FIPA (FIPA, 2000) compliant agent platform that provides the infrastructure for agent communication using the FIPA Agent Communication Language (ACL) although it does not implement the FIPA semantics. SoNG agents implement a cyclic behaviour which involves checking new inputs, followed by an internal planning phase after which chosen actions are executed. To implement internal planning and reasoning we use JESS (JESS, 2000) to build a rule based action planner for each agent. JESS operates on a Knowledge Base (KB), which represents each agent's public data (social model) and private data (beliefs about the state of the world, the user and the discourse, as well as desires and intentions). Each agent is able to make inferences about events in the environment and update its beliefs. Furthermore, each agent makes decisions according to its internal state. A rule base is constructed and loaded into the JESS engine at runtime. A rule is activated if its pre-conditions are true. In addition, the JESS KB can be stored as data in relational databases to maintain product information in the e-commerce application.

5.0.1 User Model

The use of user modelling in SoNG is apparent in two contexts: the personalisation of content, i.e. which goods match the user's preferences, and the personalisation of interaction. Modelling a user's product preferences allows the agent to tailor suggestions to the user's specific requirements and modelling a user's interaction preferences allows the agent to vary the intensity and quality of conversation and animation. The user model is constructed as a set of JESS facts, which are loaded into the agent's knowledge base at the start of a session. Three categories of facts are stored: firstly, those that represent a user's beliefs, goals and preferences, secondly, those that represent circumstantial information that the system gathered about the user and, finally, those that represent fuzzy variables whose membership functions are determined by the user.

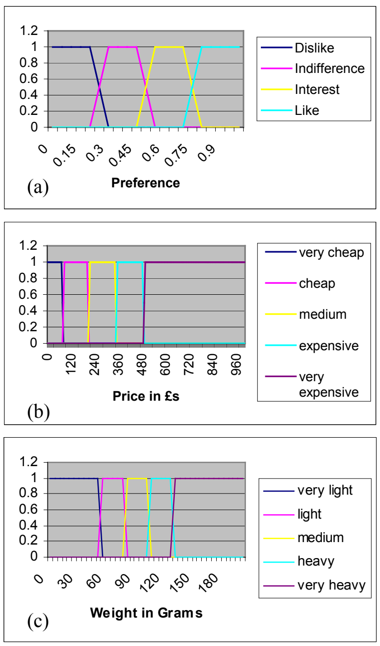

Figure 4: Examples of Fuzzy Set and constituent membership functions for (a) user preference, (b) product price and (c) product weight.

The first category is represented as a set of beliefs that the agent holds about the user. These include beliefs about the user's presumed knowledge about the domain, i.e. products or places that the agent has told the user about during the course of the conversation and hence the user knows exist. Any object that is stored in this category can be referred to directly in conversation, without the need for any introduction. In addition, the user's presumed goals will be stored as well. Finally, the agent holds beliefs about the user's preferences. As the agent's interaction with the user will be mainly through natural language, a linguistic representation of a user's preferences with respect to specific product attributes (e.g. brand, model, price, colour, etc.) was selected. An ideal candidate for such a representation is a fuzzy variable for which linguistic expressions can be used to describe fuzzy concepts in an english-like manner. Four such terms are used: like, interest, indifference, and dislike (see Figure 4(a)). Two types of preferences are modelled: preferences with respect to particular product attributes and preferences with respect to the personality of the agent. The latter will determine the intensity and quality of the agent interaction (see Section 5.0.4). Currently SoNG agents support an explicit model of users' domain specific goals and knowledge and product and personality preferences as in Appendix B.

Updates to a user's product preferences will assign a fuzzy value from the abovementioned fuzzy variable to specific product attributes. When suggesting a product to the user, the agent will select the product with the most "liked" attributes. In the event that more than one product is selected by this method, the agent will compare the remaining attributes of each product, selecting the product with the most attributes of "interest" to the user, and so on. Attributes that the user "dislikes" are not to be recommended.

Another factor that needs to be modelled is the user's interaction pattern. In particular, when was the last time the user visited the shop. Was it recently or did some time elapse. Also, what product did the user buy at his or her last visit, if any. This sort of information will allow the agent to greet the customer with utterances such as: "Hi John, I haven't seen you for a long time. How are you finding the Motorola WAP phone you bought last month?". A user's response will allow the agent to gather feedback about its model of the user's preferences. The second category of facts stored in the user model represents circumstantial data such as this collected about the user. In particular, three types of information are stored: the user's profile information (e.g. gender, age, nationality, etc.), the time and date of the user's previous sessions and the products that the user has bought in the past. These are represented as per the following example:

```
(user-profile (user-id id001) (name John) (gender male)(age 23)
(address nil) (country UK) (nationality UK))

(session-info (arrival-time "2001-09-19 05:35:26")
(departure-time "2001-09-19 05:37:11") (session-id 6))

(purchase (product-type "mobile") (model "7110") (brand "Nokia")
(price 200) (features "WAP" "large display") (session-id 6))
```

The user profile serves the purpose of identifying possible interaction preferences the user might have (i.e. as a function of gender, age and nationality). Keeping track of the regularity of a user's visit can indicate to the agent the relative satisfaction or dissatisfaction with the service provided. Finally, the purchase information is used to update user's preferences for various attributes of the product.

Additional fuzzy sets are used to model qualitative descriptions of products referring to attributes such as price (ranging between 0 and 1000) and weight (ranging between 0 and 1000 grams). For example, a product can be described by linguistic terms such as "cheap" or "expensive" and "light" or "heavy". Users' perception of the meaning of these qualitative product descriptions may differ and can be modelled by varying the distribution of membership functions within a given fuzzy set. A wealthy user will have a different perception of a "cheap" phone than a poorer user. As the membership functions for these concepts are determined by the user, their fuzzy set data is stored in the user model and loaded at the beginning of each session. Examples of the membership functions can be seen in Figure 4(b) and (c).

Amongst other things, we have identified the need to model the user's preferences about the personality and behaviour exhibited by the agent. The agent's personality is represented in JESS by means of three fuzzy variables, each describing a personality trait (e.g. closed, neutral, open), defined in a JESS fact as follows:

```
(personality (openness "open") (extraversion "extravert")
(agreeableness "agreeable"))
```

This is thus stored as a fact in the user model, to be loaded at the start of a session. Two methods are available for inferring the user preference. Firstly, by monitoring the user's utterances, the agent identifies utterances such as ”Go away”, ”You're annoying!” as indications that the user thinks the agent is trying to interact too frequently and in an annoying manner. This will cause a decrease in extraversion and agreeableness in the agent's personality. On the other hand, utterances such as ”I need help!” and ”You're too grumpy?” will have the opposite effect. Secondly, as it is not necessarily the case that the user will provide the agent with such information, the agent can make use of the Law of Similarity Attraction (Reeves and Nass, 1996). This states that people tend to like personalities similar to their own, in particular when people try to become like them. Thus by monitoring user utterances and their behaviour in the virtual world the agent can attempt to adapt its personality to the user. A user who speaks a lot and moves frenetically around the shop (as a measure of the average number of *position_changed* event notifications per minute) is deemed to be more extraverted and a user who is polite is deemed to be agreeable. Also, a user who is willing to talk about matters outside the shop domain is deemed to be more open. As mentioned above, (Moon and Nass, 1996) found this method of personality adaption to be effective.

5.0.2 Natural Language Parser

The Natural Language Processor (NLP) is responsible for the accurate recognition and interpretation of a client's utterances. The NLP module creates an ACL message in order to represent the user's communication in a manner compatible with the SoNG specifications (Figure 5). The information flow through NLP is influenced by three main factors - the parser, rules and dictionaries. All of them can be used for tuning the recognition process. The parser, based on the Link Grammar (Lafferty et al., 1992) parser, is the most stable component, however, for some inputs it can generate multiple outputs. A rule base is used to identify which of these outputs is the most correct for the current conversational context. The products and service dictionaries make the specialisation of the agent possible. Generally speaking, if the NLP is more specialised it has a better chance of identifying the customers utterances. Therefore this dictionary has to be built carefully to achieve optimal functionality with a limited number of words. The NLP uses two different methods to recognise human input. The first one is based on ELIZA and uses keyword matching for simple categories such as salutation and confirmation. The main problem with this approach is that context cannot be taken into account. However, it can still be used if, for example, we assume that "yes" always means confirmation and "Hi" is always a salutation. Despite these problems this method is the fastest and most simple for programming and debugging. For more complex sentences a rule based system is implemented. For a better understanding an example is shown on Figure 5 which is based on the story board Figure 2.

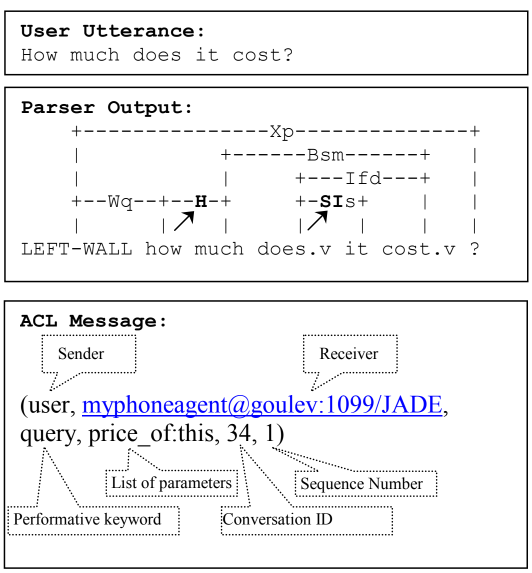

Figure 5: A User Utterance and its Coding Stages.

As shown in the example, the parser recognises six links (Xp, Wq, H, Bsm, Ifd and SIs, the meaning of the different links is beyond the scope of this paper) but we are only interested in two of them. They are marked with arrows and represent a question of type "HOW + ". The linkage "H" connects the word "how" to "much" or "many". "SI" means subject-verb inversion. In order to convert this linkage information to the ACL message showed in bottom right corner of Figure 5 we have to define a rule which is activated when both linkages H (associated with "much") and SI are present. The word on the right hand side of the SI linkage is taken to be the subject of the conversation. The rule actually inserts new facts that may be used later. A performative is identified (see Section 1) and a speech act is constructed corresponding to the user input. In some cases one NL utterance corresponds to more than one speech act.

5.0.3 Environment Awareness

The SoNG embodied agent is aware of events occurring in its 3D environment. Events are communicated by the SoNG MPEG4 player via an External Authoring Interface (EAI)

(VRML, 1997). The agent is equipped with four sensors, which alert it of events in the environment and some user states. These are:

1. Proximity Sensor: alerts the agent of the position and orientation of the user's avatar with respect to a fixed point in the 3D scene. This enables the agent to face and approach the user's avatar. In addition, it alerts the agent when the user's avatar enters or leaves the shop.

2. Visibility Sensor: alerts the agent when a particular object is in the user's field of view.

3. Collision Sensor: alerts the agent when the user's avatar collides with and object. This is especially useful for detecting when the user's avatar collides with the agent's.

4. Touch Sensor: alerts the agent when the user clicks on a object, for example a phone.

These sensors allow the agent to have awareness of its context. This is helpful in disambiguating utterances received from the user, which refer to objects in the 3D world. For example, when the user refers to "*this* phone", the agent has a knowledge of what object the user is clicking on and matches that with the reference. Furthermore, the agent can keep track of the user's interaction with objects and movements in the 3D space.

5.0.4 Action Planner

The agent uses an explicit representation of beliefs, desires and intentions (BDI architecture) to reason about its contextual information and plan communicative acts. In our system desires and intentions can make use of a language which we have developed for representing agent actions (see Appendix). We design our agents to respect their social commitments, which constrain the agent's choices and simplify the planning process by providing a limited set of possible actions. The decision making within these choices is made based on the values returned by other modules within the agent, for example the choice of product to suggest to the user is based on the products preferred by the user according to the user model. When the agent has decided what action it will take the speech and gestures it has chosen (for example pointing) will be performed with the characteristic behaviour of the agent.

In addition, the agent's personality has a bearing on the outcome of the action planner. Its effect is twofold: firstly the agents levels of extraversion will have a direct effect on the social fact *elapsed* time limit. The more extraverted the agent is the lower the limit, and viceversa, the less extraverted the agent is the higher the limit. In addition, within the limit, the frequency of agent utterances also vary in a similar fashion. As the agent's personality is a function of user preferences, it is assumed that higher levels of extaversion correspond to increase help and assistance towards the user.

Secondly, the personality also has an effect on the quality of the agent interaction. Increase levels of extraversion also increase the degree of emotional expressiveness in the agent's utterances, just as increased agreeableness will increase the politeness. Other than the agent's utterances, personality also has an effect on the intensity of the agent's animation.

When the agent needs to control its embodied representation the output of the Action Planner in terms of speech and behaviour scripts are merged into an Avatar Markup

Language (AML) (Kshirsagar et al., 2002) script. AML is an XML based format, also developed within SoNG, which allows for the timely synchronisation and merging of face animations, body animations, lip synching and text-to-speech. Furthermore, AML allows for the specification of parameterised behaviour calls such as walking, pointing and facing, which each take 3D coordinates as input. To support this functionality the AML Processor maintains an internal record of the current position of the agent's avatar in the 3D scene.

5.1 Current Status: System Walkthrough

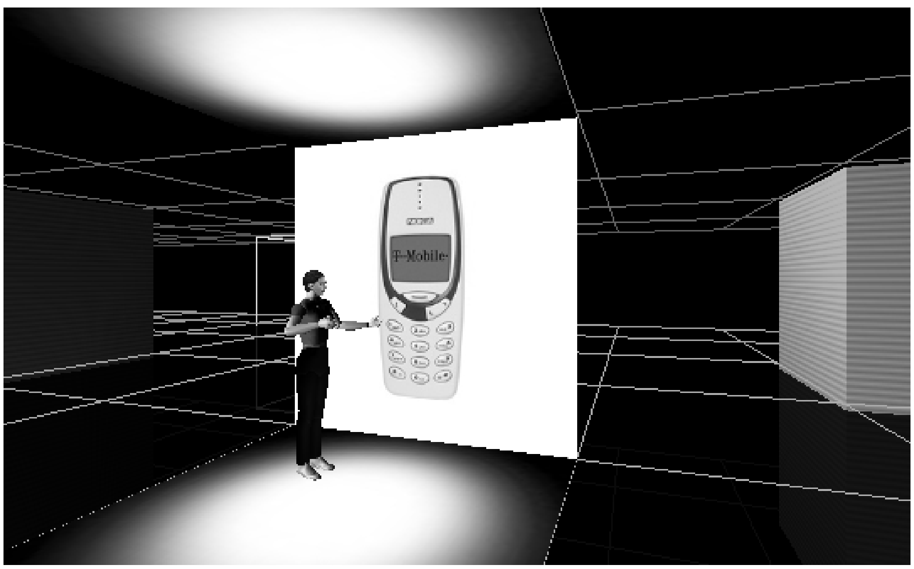

Figure 6: Sales Agent in Phone Shop.

Here we analyse part of a simple interaction which shows how all the modules described above interact to produce the desired behaviour. The first event is the entrance of a user in the phone shop. The event semantics for this initiates an interaction following the `Shopkeeper-Customer` protocol with the user that entered playing the role of Customer and the agent playing the role of Shopkeeper. The existence of an empty partial state leads to the creation of a commitment for the Shopkeeper to greet the Customer. A planning rule in the agent is triggered by this commitment and the belief that the agent knows the name of this user, creating the intention to send a speech act which results in "Good morning John." being sent to the output terminal. The agent loads the user model for the Customer that entered, this includes loading the fuzzy preferences of the user. To make it an effective sales agent, the agent has the following desire:

```
(in shopkeeper-customer
    (find list_of: telephone: date:
        last_interaction(in role: customer));
    (tell in role: customer list_of: telephone: price:
        preference(in role: customer))
)
```

This is an action expression. The syntax of this language is given in Appendix A. When the agent is engaged in a Shopkeeper-Customer protocol the desire is asserted for this specific interaction giving a composition of two actions to be executed sequentially:

```
(find list_of: telephone: date: last_interaction(in role:customer))
```

The agent playing the role of Customer is substituted and the date of his last interaction is found, giving:

```
(find list_of: telephone: date: 2000-02-27 01:52:27)
```

Now the agent does not have knowledge about the required subject, but it believes that the search agent does, so it sends a query to the search agent:

```
query,"list_of: telephone: date:2000-02-27 01:52:27"
```

This starts an interaction with the search agent who is now committed to reply to the phone agent. Since it doesn't have the required information itself it forwards the query to a database agent who it believes knows about phones. In the database agent the query is asserted and a list of phones is generated, this is then passed back to the search agent who replies to the phone agent with the list as content. This is a list of phones released since the user last visited but not necessarily matching the user's preferences. When the phone agent receives the list of phones it asserts each phone as a fact, matching each parameter of the phone with the user's fuzzy description for that parameter. For example where the price is represented by a numerical value in the phone list received from the search agent, the phone agent will assert this as a price "expensive" or "cheap" according to the user's fuzzy terms stored in the user model. Now that the agent has completed the find part of the action expression it proceeds with:

```
(tell John list_of: telephone: price: preference(John))
```

The agent refines the phone list according to the user's preferences and informs the user of any new phones which match his preferred price range. In the meantime, the user model states that John's last session was more than two months ago. The agent formulates a speech output by applying its personality traits to a template expansion mechanism that selects an appropriate output for the current dialogue context. Finally, the agent fulfils its social obligation and utters:

"Good Morning John. Welcome back to the phone shop. We have some new phones for you: the Nokia 6110 and the Nokia 3210. Would you like to see them?"

6 Related Work

The design of embodied conversational agents is a topic which has attracted a great deal of interest from the research community in recent years. Embodied agents are used in a number of application domains including online tutoring, interactive storytelling, help, news reading and e-commerce systems. Although, these agents take on a variety of forms, some of the underlying techniques are the very similar, however, the research emphasis is often quite different.

Peedy (Ball et al., 1997) the parrot is a character developed by Microsoft Research in a Project called Persona to assist a user in selecting and playing musical tracks. Its implementation is based on the Microsoft Agent API, which supports the development of desktop based animated agents. Its developers outline three requirements for assistive interfaces which motivate the work on Persona: 1) ”Support interactive give and take” through support for mixed initiative interaction. 2) ”Recognise the cost of interaction and delay”. An assistive interface should only interact with the user if the cost of making a mistake is higher than that of interrupting the user. Finally, 3) ”Acknowledge the social and emotional aspects of interaction”. Just like human assistants learn what the appropriate behaviour is depending on time of day, task and the mood of the people they are interacting with.

REA (Cassell et al., 1999) is a fully immersed animated agent that acts as a female real estate advisor. It is built on experience gained from Gandalf and Animated Conversation and combines the qualities of both those systems. In particular, aside from generating verbal and non-verbal output and supporting turn taking, REA is able to recognise verbal and non-verbal input directly from the human user by means of sensors. In addition, REA is able to give signals that indicate the state of a conversation. Empirical studies have been carried out to identify appropriate (Cassell et al., 2001).

REA’s architecture is based on the Functions, Modalities, Timing and Behaviour (FMTB) conversational framework. Here (conversational) functions can be of two types: interactional, i.e. serving the purpose of regulating conversation, and propositional, i.e. functions that convey the content of the conversation. Modalities include speech as well as facial animations and gestures such as hand waves and head movements. Timing highlights the need for prompt response as well good synchrony between output behaviours. Finally, behaviours are identified as those building blocks used to convey conversational functions. Many other systems have attempted to map gestures and animations to conversational functions. COSMO (Lester et al., 1997a), for example, uses pointing as a means of resolving ambiguity in conversations when an object is referred to in the environment.

Although most efforts have been spent on studying the relationship between animated agents and human users, some research has been looking at the relationship between multiple animated agents conversing and interacting with each other. Indeed, (André et al., 2000) have been looking at using teams of animated agents in a car sales scenario interacting with each other and a human user simultaneously. This metaphor enforces the need for explicit models of social roles and personalities in order to vary the behaviour of the participants in the conversation.

7 Conclusion and Future Work

The work described in the paper is concerned with the design and development of 3D embodied agents which are capable of carrying out conversations with users and tailoring services to user preferences, where conversations are enriched by meaningful facial expressions and body animations.

We have presented an architecture and implementation for agents situated in a virtual marketplace. We have designed a distributed modular architecture for agents, which means that the various different agents can be built from re-usable blocks. The issue of believable interaction is central to our study. In order to achieve this and create the impression of intelligence we need to break away from prescripted interaction and develop an architecture, which allows emerging behaviour. This is made possible by incorporating a model of personality in the agent, which varies along the dimensions of extraversion, agreeableness and openness.

In our implementation we have successfully integrated all modules described in this paper to create an architecture for a believable embodied agent. This agent was inserted in a multi-agent system the purpose of which is to provide database search facilities to the end user. Finally, we interfaced the system with an MPEG4 3D player to which animations are streamed and from which event notifications are received. The contribution of this work is twofold. Firstly to identify the issues involved in the development of embodied agents for e-commerce applications. Secondly to demonstrate a design which effectively tackles these issues.

Now that the system is fully integrated we must study the effectiveness of this type of system on real users. In particular, tests will be carried out to evaluate the use of personalisation and personality traits to adapt the behaviour of the embodied agent to the user. Also, in the next stage of development we plan to investigate the possibilities of enriching the agent's personality model by developing it on top of an explicit emotion engine. By the same token we need to be able to acquire users' emotional states and we plan to develop the NLP module to allow for this functionality.

Acknowledgements

We are very grateful for the extensive comments given by the anonymous reviewers which led to significant mprovements in the paper. This work has been undertaken with the financial support of the SoNG project (IST-1999-10192), part of the EU-funded Information Societies Technology (IST) programme.

References

Adelson, B. (1992). Evocative agents and multi-media interface design. In *Proceedings of the UIST'92 (ACM SIGGRAPH Symp. on User Interface Software and Technology), Monterey*, pages p352–356.

André, E., Klesen, M., Gebhard, P., Allen, S., and Rist, T. (2000). Integrating models of personality and emotions into lifelike characters. In *Proceedings of Affect in Interactions*. Springer-Verlag.

Ball, G., Ling, D., Kurlander, D., Miller, J., Pugh, D., Skelly, T., Stanosky, S., Thiel, D., and Wax, T. (1997). Lifelike computer characters: The persona project at microsoft research. In Bradshaw, A., editor, *Software Agents*, pages p191–222. The AAAI Press/The MIT Press, Cambridge, MA.

Bates, J. (1994). The role of emotion in believable agents. *Communications of the ACM*, 37(7):122–125.

Carroll, J. and Rosson, M. (1987). The paradox of the active user. In Carroll, J., editor, *Interfacing Thought: Cognitive Aspects of Human-Computer Interaction*. MIT Press, Cambridge, MA.

Cassell, J., Bickmore, T., Camphell, L., Vilhjalmsson, H., and Yan, H. (1999). The human conversation as a system framework: Designing embodied conversational agents. In Cassell, J., editor, *Embodied Conversational Agents*. MIT Press.

Cassell, J., Nakano, Y., Bickmore, T., Sidner, C., and Rich, C. (2001). Annotating and generating posture from discourse structure in embodied conversational agents. In *Proceedings of the Workshop on Multimodal Communication and Context in Embodied Agents, Autonomous Agents*.

Charlton, P., Kamyab, K., and Fehin, P. (2000). Evaluating explicit models of affective interactions. In *Proceedings of the Workshop on Communicative Agents in Intelligent Virtual Environments*.

Dehn, D. M. and van Mulken, S. (2000). The impact of animated interface agents: a review of empirical research. *International Journal of Human-Computer Studies*, 52:p1–22.

Elliot, C., Lester, J., and Rickel, J. (1999). Lifelike pedagogical agents and affective computing: an exploratory synthesis. In Wooldridge, M. and Veloso, M., editors, *AI Today*.

Eysenck, H. (1991). Dimensions of personality. *Personality and Individual Differences*, 12(8).

Fink, J. and Kobsa, A. (2000). A review and analysis of commercial user modeling servers for personalization on the world wide web. *User Modeling and User-Adapted Interaction*, 10:p209–249.

FIPA (2000). Fipa specification. In *Foundation for Intelligent Physical Agents. http://www.fipa.org/specs/fipa2000.tar.gz*.

Fujita, F. (1996). The big five taxonomy, http://www.iusb.edu/%7Effujita/Documents/big5.html.

Guerin, F. and Pitt, J. (2000). A semantic framework for specifying agent communication languages. In *Fourth International Conference on Multi-Agent Systems (ICMAS-2000)*, pages 395–396. IEEE Computer Society, Los Alamitos, California.

Guerin, F. and Pitt, J. (2001). Denotational semantics for agent communication languages. In *Autonomous Agents 2001, Montreal*, pages 497–504. ACM Press.

Hanani, U., Shapira, B., and Shoval, P. (2001). Information filtering: Overview of issues, research and systems. *User Modeling and User-Adapted Interaction*, 11:p203–259.

JADE (2000). Jade project home page available at http://sharon.cselt.it/projects/jade.

JESS (2000). JESS, the Java Expert System Shell. Sandia National Laboratories. http://herzberg.ca.sandia.gov/jess/.

John, O. P. (1990). The big five factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In Pervin, L., editor, *Handbook of personality: Theory and research*.

Kay, J. (1993). Pragmatic user modelling for adaptive interfaces. In M. Schneider-Hufschmidt, T. K. and Malinowski, U., editors, *Adaptive user Interfaces: Principles and Practice*.

King, J. and Ohya, J. (1996). The representation of agents: Anthropomorphism, agency and intelligence. In *Proceedings of CHI'96*.

Kobsa, A. (1990). User modelling in dialog systems: Potentials and hazards. *AI and Society*, 4:p214–240.

Koda, T. and Maes, P. (1996). Agents with faces: The effects of personification of agents. In *Proceedings of HCI'96*, London.

Kshirsagar, S., Guye-Vuillème, A., and Kamyab, K. (2002). Avatar markup language, submitted to the 7th international conference on 3d web technology, web3d 2002.

Lafferty, J., Sleator, D., and Temperley, D. (1992). Grammatical trigrams: A probabilistic model of link grammar. In *Proceedings of the AAAI Conference on Probabilistic Approaches to Natural Language, October*.

Lester, J., Voerman, J., Towns, S., and Callaway, C. (1997a). Cosmo - a life-like animated agent with deictic believability. In *Working notes of the IJCAI workshop on animated agents: Making Them Intelligent*.

Lester, J. C., Converse, S. A., Kahler, S. H., Barlow, S. T., Stone, B. A., and Bhogal, R. S. (1997b). The persona effect: Affective impact of animated pedagogical agents. In *CHI*, pages 359–366.

Maes, P. and Shneiderman, B. (1997). Direct manipulation vs. interface agents: A debate. *Interactions*, 4(6).

McCalla, G. (1992). The central importance of student modelling to intelligent tutoring. In Costa, E., editor, *New Directions in Intelligent Tutoring Systems*, pages p108–131. Springer-Verlag, Berlin.

Moffat, D. (1997). Personality parameters and programs. In Trappl, R. and Petta, P., editors, *Creating Personalities for Synthetic Actors: Towards Autonomous Personality Agents*. Springer-Verlag.

Moon, Y. and Nass, M. (1996). Adaptive agents and personality change: Complementarity versus similarity as forms of adaptation. In *Proceedings of CHI'96*.

Norman, D. (1997). How might people interact with agents. In Bradshaw, A., editor, *Software Agents*, pages p49–55. The AAAI Press/The MIT Press, Cambridge, MA.

Orwant, J. (1996). For want of a bit, the user was lost. *IBM Systems Journal*, 35:p398–416.

P3P (2001). Platform for privacy preferences (p3p) project, http://www.w3c.org/p3p.

Paiva, A. and Martinho, C. (1999). A cognitive approach to affective user modeling. In *Proceedings of the Workshop on Affect in Interactions Towards a New Generation of Interfaces, held in conjunction with the 3 rd i3 Annual Conference*.

Peppers, D. and Rogers, M. (1993). *The One-to-One Future*. Doubleday, New York.

Pervin, L. and John, O., editors (1993). *Personality: Theory and Research*. Wiley and Sons.

Picard, R. (1997). *Affective Computing*. The MIT Press.

Prendinger, H. and Ishizuka, M. (2001). Social role awareness in animated agents. In *Autonomous Agents 2001, Montreal*, pages 270–277. ACM Press.

Rao, A. S. and Georgeff, M. P. (1992). An abstract architecture for rational agents. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Third International Conference (KR '92)*, pages 439–449. Morgan Kaufmann Publishers, San Mateo, CA, USA.

Reeves, B. and Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press, Cambridge, England.

Searle, J. R. (1965). What is a speech act ? In *Philosophy of Language. edited by A.P. Martinich, Third edition.* Oxford University Press.

Singh, M. (1998). Agent communication languages: Rethinking the principles. *IEEE Computer*, 31(12):40–47.

Sleeman, D. and Brown, J. (1982). *Intelligent Tutoring Systems*. Academic Press, London, UK.

Sproull, L., Subramani, M., Kiesler, S., Walker, J., and Waters, K. (1996). When the interface is a face. *Human-Computer Interaction*, 11:p97–124.

Traum, D. and Rickel, J. (2001). Embodied agents for multi-party dialogue in immersive virtual worlds. In *Autonomous Agents 2001 Workshop on Multimodal Communication and Context in Embodied Agents, Montreal.*

TRUSTe (2001). Truste, http://www.truste.org.

VRML (1997). Vrml97, http://www.web3d.org/technicalinfo/specifications/vrml97/index.htm.

Weizenbaum, J. (1966). Eliza: A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9:p36–45.

Appendix A: Action Expression Language

```
<Action expression>   ::= <Actor>, <AE>
<Actor>               ::= <who>
<AE>                  ::= <AE1> ; <AE2> |
                          (if <cond> then <AE>) |
                          (in <prot> <AE>) |
                          (assign var=<value>) |
                          (get position <who>) |
                          (point <coords>) |
                          (walk <coords>) |
                          (tell <who> <what>) |
                          (find <what>) |
                          (demo <object description>) |
                          (buy <object description>) |
                          (sell <object description>)
<cond>                ::= (protocol=<prot>) |
                          (var=<value>)
<prot>                ::= shopkeeper-customer
<who>                 ::= role:<role expression> |
                          in role:<role name> |
                          <agent identifier> |
                          me
<role expression>     ::= <role name>, <interaction id>
<role name>           ::= <string_value>
<interaction id>      ::= <number>
<value>               ::= <what>
<coords>              ::= in front <who> |
                          <number>,<number>
<what>                ::= is_there: <object description> |
                          list_of: <object description> |
                          price_of: <object description> |
```

	is: <object description> |
<object description>	::= <object type>: <parameters> | gcomparative this <parameter> | lcomparative this <parameter> | comparative this <parameters> | this
<object type>	::= this | <product type>
<product type>	::= mobile | fax | telephone | pda
<parameters>	::= <parameter>:<parmvalue> +
<parameter>	::= manufacturer | model | weight | price | date | colour
<parmvalue>	::= <string_value> | last_interaction(<who>)

Appendix B: JESS Belief Structure

(belief
    (agent User)
    (content <Content>)
)

Content	::= exists <Concept> preference <FuzzyPreference> <Concept> goal <Goal> <Concept>
FuzzyPreference	::= like interest indifference dislike
Goal	::= buy enquire see
Concept	::= product <product type> manufacturer <string value> price <string value> personality <string value>