

A Virtual Communication Interface Called E-Mask

P. Gonçalves¹, G. D. Putnik²
pgoncalves@ipca.pt, putnikgd@dps.uminho.pt

(recebido em 11 de Abril de 2008; aceite em 12 de Maio de 2008)

Resumo. Este artigo apresenta um novo Interface de Comunicação Virtual (ICV), baseado em agentes, denominado *E-Mask*, que esconde a verdadeira aparência dos seus interlocutores. Depois de contextualizar a tecnologia subjacente aos Agentes da *Microsoft*, abordamos a implementação da virtualidade no *BM_Virtual Enterprise Architecture Reference Model (BM-VEARM)* com personagens “humanas”. São, também, apresentadas outras aplicações de agentes bem como do ICV. O *E-Mask* pode ser um potenciador na reconfiguração dinâmica de redes de Empresas Virtuais uma vez que, com este interface o tempo de comunicação entre utilizadores é mais curto.

Palavras-chave: Interface, comunicação virtual, reconfiguração dinâmica.

Abstract. This paper presents a new agent-based Virtual Communication Interface (VCI), called *E-Mask*, which hides the real appearance of its interlocutors. After a contextualization of Microsoft Software Agent technology, we focus on the implementation of virtuality in the *BM_Virtual Enterprise Architecture Reference Model (BM-VEARM)* with human-like characters. Other applications of agent software as well as the VCI are also covered. We find that the *E-Mask* interface may be a dynamic reconfiguration enabler in Virtual Enterprise (VE) Networks since the communication time between end-users is shortened.

¹ Polytechnic Institute of Cávado and Ave, Higher School of Technology, Barcelos, Portugal

² University of Minho, School of Engineering, Dept. of Production and Systems Engineering, Guimarães, Portugal

Keywords: Interface, virtual communication, dynamic reconfiguration.

1. Introduction

Today organizations are flatter, have a more diverse employee base and a greater use of teams. The domestic workforce is constituted by employees who reflect differences in age, ethnic heritage, race, physical abilities, gender and sexual orientation. Diversity is not just a matter of social responsibility; it is also an economic issue. Companies are realizing the advantages of combining the talents, experiences and perspectives of a diverse employee base.

The ability to compete in the global economy is the single greatest challenge facing business today. Organizations want to negotiate, buy and sell overseas, consider joint ventures, market and adapt products for an international market. This makes communication essential to organizational success.

State-of-the-art information technology and communication technology is enabling cross-boundary business activities. Several factors appear as supreme factors of competitiveness, namely: (1) the organizations' capability to achieve and explore competitive advantages in synergy, by using or integrating the optimal available resources for the functions that the organization undertakes; (2) the capability of fast adaptability to the market, together with (3) the capability of managing all business processes independently of distance and (4) the ability to learn with history. The recent Virtual Enterprise (VE) organizational models traduce these factors (Cunha & Putnik, 2002).

The proliferation of new technology has created new business models and has made virtual teams flourish. Virtual teams are groups of geographically and/or temporary dispersed individuals brought together via information and telecommunication technologies (DeSanctis & Poole, 1997; Jarvenpaa & Leidner, 1999; Lipnack & Stamps, 1997; Powell et al., 2004). Their main advantages include quicker responses, new potentials and greater adaptability.

Effective and efficient virtual communication is becoming crucial in VE networks. Even though computers extend what people want to communicate through wide

area network (WAN) technologies in a short period of time, sometimes written, verbal and body language is momentarily misinterpreted causing the direct communication process to be inefficient or delayed (Gonçalves et al., 2006).

The main goal of this article is to propose a “virtual” communication interface (VCI), which is able to minimize communication interruption time between end-users in a dynamic VE environment. We believe that the VCI may turn out to be more efficient than traditional face-to-face communication.

This paper is organized in six parts: in the first part, we present the BM_Virtual Enterprise Architecture Reference Model (BM-VEARM) and explain the birth of the VCI; in the second part, we uncover how Microsoft Agent Technology works; in the third part, we discuss antropomorphism and relate it to our study; in the fourth part, the basic role of Microsoft Software Agents in the VCI is addressed; in the fifth part, we present the E-Mask as a solution for virtual communication and dynamic reconfiguration in VE. Finally, in the sixth part, we explore some other agent and E-Mask applications.

2. The BM_VEARM: the VCI Parent

Even though we are living in a digital society, traditionally, people communicate face-to-face, ear-to-ear or even finger-tip to finger-tip. However, in the VE context, where companies are extended overseas and over various time zones and where business and manufacturing operations must be managed through WAN technologies and protocols, such as the Internet, in split second time, it is not always feasible to have the same people “communicate” on the “other end” or wake up the next morning and check the e-mail.

In our research on the implementation of “virtual communication” in the BM_VEARM model, at the University of Minho, we decided to develop and test, in a laboratorial environment, a user-friendly and intuitive communication interface which “masks” end-users from each other. We will now explain the model and clarify why we created the interface.

The BM_VEARM, proposed by Putnik (2000a; Putnik, 2000b) is defined as a hierarchical structure of multiple levels of inter-enterprise processes which should satisfy four fundamental requirements: *integrability*, *distributivity*, *agility* and *virtuality*. We will focus on the latest for we consider it to be the VCI parent.

Virtuality provides the system with the capability of on-line system reconfigurability without the interruption of any process (Putnik, 2000a). To implement virtuality in the enterprise, the introduction of an interface layer between the Control level i (principal, manager, “client”) and the Control level $i+2$ (agent, worker, “server”) is proposed. It now becomes Control level $i+1$. This level manages the resources which will carry on the process ordered by the upper level. Therefore, there has to be a resource manager or broker.

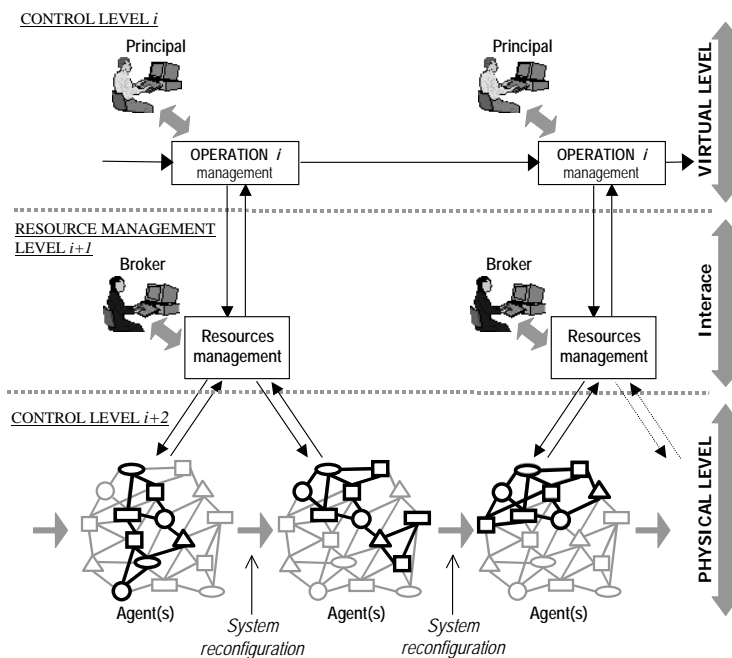


Figure 1 – Virtual Enterprise operation scheme – elementary structure
(Putnik, 2000a)

In this model, the Control level i (principal, manager) and the Control level $i+2$ (agent, worker) communicate through the Resource Management level $i+1$, i.e., through the *resource manager*, or *broker*. During the operation, the manager (the principal, the “client”) does not have direct contact with the worker (the operator or agent, the “server”), who provides the service (or production). The broker must provide the transition from one physical structure to another in a way that the “principal” cannot be affected by the system reconfiguration, in which case the operation would be interrupted and split in two, implying some lost time. The lost time can have two components: by interruption of the operation itself (e.g. set-up time for restarting the operation), and the principal’s adaptation time to the new specific organizational structure. The principal only sees a “virtual structure”.

In our case, the “virtual structure” is implemented through the designed VCI, which is based on MS Agent technology. In this interface, we call E-Mask, the end-user can not see or hear the human interlocutor because he/she is “masked” by one of the chosen software agents. In other words, the end-user always sees the same software agent/mask and hears the same (mask’s) voice.

When it comes to communicating in a VE, and as dynamic reconfiguration occurs accompanied by many interlocutor changes, the end-user must have the capability to quickly adapt to these changes and continue to trust the person who appears on the screen. The VCI appears to play an important role in shortening the virtual communication time since there is an increased concern in reducing the end-user’s adaptation time to the new interlocutors, as opposed to traditional face-to-face communication, and to eliminate physical appearance and language as barriers.

3. What is Microsoft Agent Technology?

Microsoft Agent enables Web sites and conventional applications to include an enhanced form of user interaction. It provides several user interface components that enable users to access and interact with the character, know the character’s

status and change global settings. This section describes these basic elements of the Microsoft Agent user interface.

Microsoft Agent, by itself, is really just a software technology rather than a complete, integral software application that automatically does something. That is why it is not possible to see a character appearing on the computer screen as soon as Microsoft Agent is installed on any one of the following operating systems: Microsoft Windows Vista, Windows XP, Windows 2000 or Windows Me. Developers need to first program the Microsoft Agent technology into their applications and Web sites. A separate software application or scripting code has to be run and it is this that actually calls upon the services of Microsoft Agent. When the applications are run, it is possible to see an animated character appear and do as it has been programmed. Microsoft Agents are 2D animated characters that may simulate human beings (actors) through basic text-to-speech tasks.

This agent technology is compatible with a number of environments, including Microsoft .NET, C#, Microsoft Visual Studio, Microsoft Visual Basic, Visual C++, Microsoft Foundation Classes (MFC), Visual J++, HTML, Microsoft Office, and Microsoft Visual FoxPro. Various technologies and software packages can be used in the implementation. Thus, the final design usually consists of a combination of several different programming languages.

For the implementation of this technology it is necessary to download the following components from:

<http://www.microsoft.com/products/msagent/downloads/user.asp>:

Agent Requirements

- Microsoft Agent core components
- Microsoft Speech Recognition Engine
- Text-to-speech engines
- Microsoft Agent character files

All the requirements are freely available. The Microsoft site also offers downloads for development, such as:

- SDK documentation for Microsoft Agent
- Code samples
- Agent Character Editor
- Linguistic Information Sound Editing Tool
- Character data files for Genie, Merlin, Peedy and Robby

In Figure 2 it is possible to see four famous Microsoft Agents.



Figure 2 – Microsoft Agents: Genie, Merlin, Peedy and Robby

Many other characters have been developed all around the world³. We decided to use Robby, E-Man and E-Woman (Figure 3) in our work due to their “resemblance”.

Each character has its own programmable “personality”. E-Man may look like he is thinking or looking to the side and E-Woman may blink or wave. Even though Robby has less animations we decided that he would make a perfect “trio” with E-Man and E-Woman. You can see E-Man’s character details below (Figure 4).

³ Some examples may be seen in <http://www.msagentring.org/chars.aspx>.

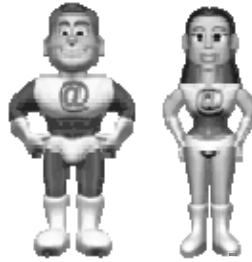


Figure 3 – Microsoft Agents: E-Man and E-Woman

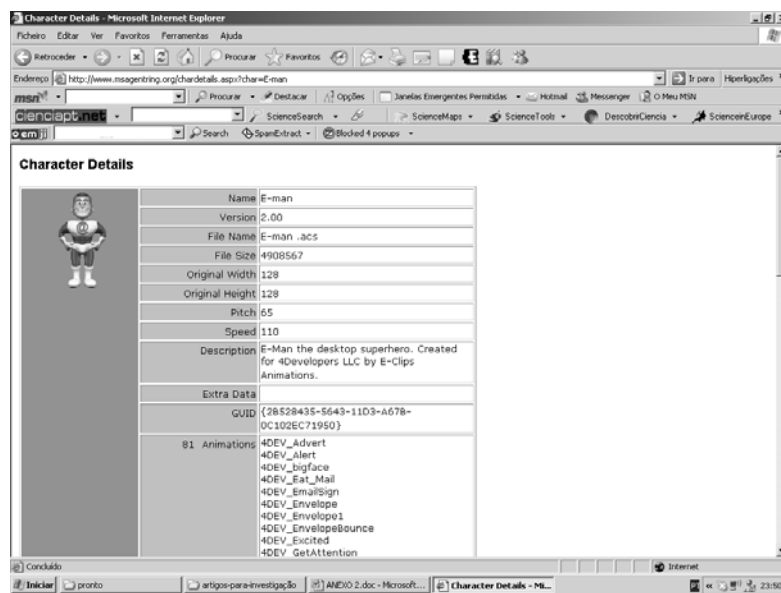


Figure 4 – Character details of E-Man

In our design, Microsoft Agent is used in the communication between the user and the computer. In communicating with users, Microsoft Agent does some of the basic text-to-speech tasks.

To make the character speak and move, the simple commands are as follows:

```
function Play() {  
  
AgentX.Characters.Load("Robby", "Robby.acs");  
  
Robby = AgentX.Characters.Character("Robby");  
  
Robby.Show();  
  
Robby.Play("Explain");  
  
Robby.Speak();  
  
}
```

4. Anthropomorphism

Should an agent be represented by a humanoid character on the screen? The argument for doing so is that it may facilitate the formation of a perceived relationship between the user and the computer. Once there is a relationship a person acts differently towards the computer – he/she may feel more engaged in their work; may try harder; may be able to focus his/her attention on something the computer is presenting.

The central issue regarding anthropomorphism is whether we are attracting a person's attention to the things they are actually trying to do or distracting them. If the character is entertaining or supportive, then anthropomorphism can be helpful. If on the other hand it takes a person's attention away from the things that they are focusing on, then it is not.

Agent enthusiasts all over the world have been developing all types of agent models (animal-looking, cartoon-looking, alien-looking and even human-looking agents).

We believe human character figures may be more suitable in the case of VE communication since people are used to “communicate” and “negotiate” with people; not animals, not cartoon characters and, hopefully, with sane minds, not even “aliens”. Robby the robot character is used in our interface because we believe people would respond better to him rather than to Genie, Merlin, or Peedy. Ideally, we prefer human-looking characters.

In Figure 5 it is possible to see some human-looking agent examples:

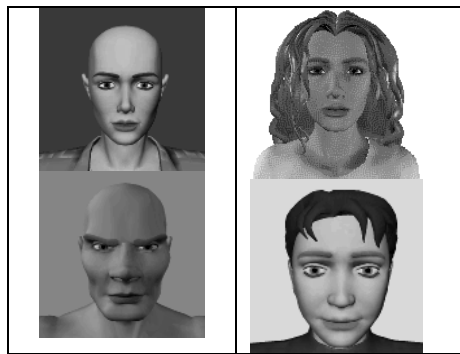


Figure 5 – Human-looking Agent Characters⁴

Of course, there is always the danger that people will treat the on-screen character as a real person, or their expectations for the human-like behavior of the system will inevitably be overextended. This is a real danger, and one that we must keep in mind and address as we are designing systems. A good example of this danger may

⁴ Images retrieved from: [<http://www.msagentring.org/chars.aspx>] and [<http://www.roblindman.com/characters.php>] on 17/02/2007.

be seen in the stories told by “Second Life⁵” users around the globe. In our interface, the intention is to have a real person at the end of the line, therefore, and in our case, the danger is unrealistic.

5. Software Agents as VCI Actors

As mentioned earlier, our goal is to improve virtual communication in VE with the design and development of an effective and efficient user-friendly interface with Microsoft Agent representatives (actors). In this section, we will describe software agents in general, discuss some of their applications and mention how Microsoft Agent technology works and is applied to our work.

In our literature review, agent technology has been applied for over a decade. Authors like Bradshaw (1997), Nwana & Azarmi (1997) and Sycara et al. (1996) defend that intelligent agents have been applied to many Internet and Web applications. Over the years, many enhancements make on-line intelligent agents more human-like. Like a secretary, travel agent, stockbroker, or medical assistant, they can provide suggestions and help lessen burdens on those that employ them. The best ones provide personalized assistance, remembering past interactions, and using them to infer the user’s preferences. They automate commonly performed tasks, constantly improving their behavior and being always on the lookout for opportunities.

Simple agents handle question/answer problems by looking up the question from a database. However, everything has to be hard coded by the programmer and the agents do not work well when the questions are not in the database. More advanced intelligent agents reply with information that does not exactly match the question in the database, but has the same meaning and the reply is valid. They have improved ability in coming up with an answer, but they are still limited to only replying whatever the content in the answer field.

Collaborative agents can also work behind the scenes as well as directly in the user interface. Collaborative agents take information from many different sources and

⁵ For more information see: <http://secondlife.com/>.

locations. They collaborate when multiple agents act as a team in analyzing data and solving problems, bringing multi-perspective and multi-expertise capabilities to bear in solving problems.

In our work, Microsoft Agents play an important role in virtual communication since they may be the supporting technology for VE reconfiguration. As mentioned previously, the broker, in the BM_VEARM model, will be represented by an animated software character (Microsoft Agent). In the next section we will present the VCI.

6. E-Mask as a Solution for Virtual Communication and Dynamic Reconfiguration

In the quest of applying software agents as supporting technology for virtual communication and dynamic reconfiguration we created the E-Mask (virtual mask) application. It was developed according to the paradigm Client-Server in C# on the .NET platform. Figure 6 illustrates the global architecture of the software. It runs over the internet and, in theory, may have more than one client program communicating with the server program, though, for experimental purposes, only one client program can run over a local area network (LAN).

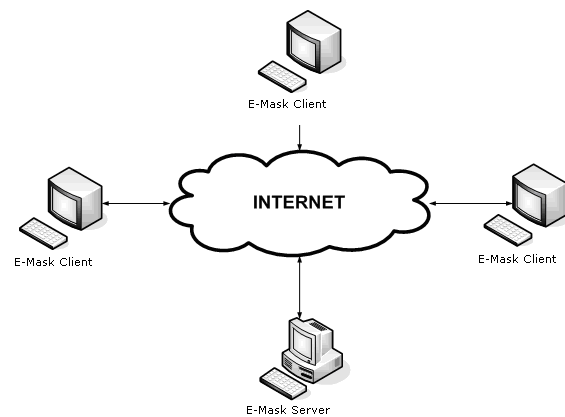


Figure 6 – Global Architecture of the Virtual Communication Architecture

In order for E-Mask to work we specifically needed:

- a *Text-To-Speech Portuguese (Brazilian)* Engine (it provides speech output capabilities for Microsoft Agent and the Agent Character Editor);
- MS Agent characters: Robby, E-Man and E-Woman;
- SAPI 4.0 runtime support (it is required for three reasons: first, E-Mask was programmed for the Windows XP operating system; second, because Robby is compiled to use the L&H TruVoice engine as the default speech output engine and the SAPI 4.0 runtime is necessary to set this characters' TTSMoDeID property to use the TTS3000 engines; and third, the other Microsoft Agent characters also require this SAPI 4.0 runtime support to set their TTSMoDeID property).

Communication between the end-user and the agent is established through a server program and a client program for E-Mask.

The client application has the following implemented functions (see Figure 7):

- Connect to the Server;
- Select MS Agent;
- Create repository of files per session;
- Record Response (activated by the server).

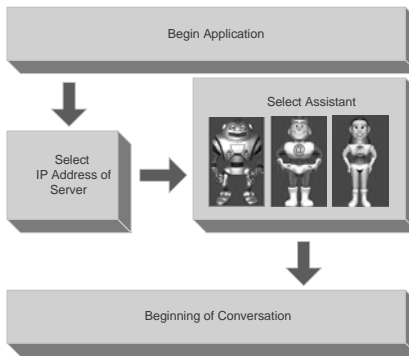


Figure 7 – Client Application Architecture

The “invisible” interlocutor, is behind the server program in order to type what he/she wants the agent to say (with the agent’s own voice) to the end-user after listening to the end-users’ responses. The end-user will have a sense that he/she is talking to the agent and that the agent is replying.

The interface of the application was designed in order to assure that the interaction with the client was through a Microsoft Agent. This technology allows end-users to have both a visual (they can see the agents) and a sound interaction (they can hear one agent speak at a time). This interaction enables a practical and friendly interface for the client.

E-Mask is integrated in MSN Messenger in order to use the same audio features, however, its function is to mask the interlocutor, or in other words, the end-user is not able to see a person as the interlocutor but rather as a MS Agent character (shown in Figure 8).

The client may choose the agent he/she wishes to speak to. The following options are MS Agents: Robby, E-Man and E-Woman.

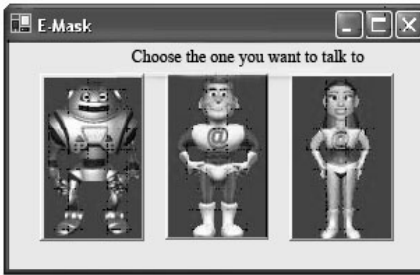


Figure 8 – VCI (E-Mask) Application Interface

The server interacts with the client through text that is converted into sound as the chosen agent reads out loud what the server wrote in the appropriate text box. In the following image you can see an example of an MS Agent introducing himself to the client (Figure 9). The text that is written in the balloon is actually read by the agent.

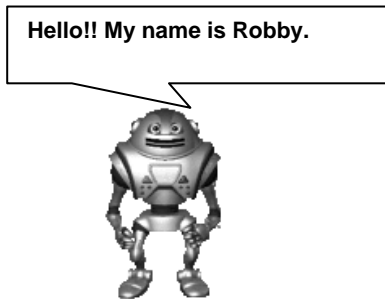


Figure 9 – Example of Robby, the agent, introducing himself

As Putnik et al. (2008) show, through experimental results in a laboratory environment and the subsequent statistical analysis, virtual communication time is shorter when software agents are used in a VE orientated interface. Therefore, the VCI appears to be a solution when it comes to shortening the dynamic reconfiguration time in VE networks.

7. Other Agent and E-Mask Applications

Computers are now getting released from the times of WIMP (windows, icons, menus, point-and-click devices) graphical user interface. Post-WIMP will be increasingly based on the way we naturally interact with our environment and with other humans (Dam, 2000). During the recent half century, skilled artists have achieved surprisingly life-like behaviors in animated characters. In the last few years, character agents, based on real video, cartoon-style drawings, or model-based 3D graphics, have become increasingly popular in user interfaces. For applications, they are a promising option since they make the application interface more lively and more appealing. However, while animated agents, or life-like characters, start populating the interfaces of numerous computer-based applications, their impact on human users is still largely unexplored (Prendinger & Ishizuka, 2003).

Over the years we have witnessed a growing interest in employing animated agents for tasks that are typically performed by humans. Some of the more prominent applications we found are:

- as virtual *tutors* in interactive learning environments (Conati, 2002),
- as virtual *sales persona* on the web and at information booths (André et al., 2000),
- as virtual *presenters* (Badler et al., 2000),
- as virtual *communication partners* in therapy (Marsella et al., 2000),
- as virtual *actors for entertainment* (Prendinger et al., 2002)
- as virtual *personal representatives* in online communities and guidance systems (Cassell & Vilhjálmsson, 1999), and
- as virtual *information experts* enhancing conventional web search engines (Kitamura et al., 2002)

Day by day people seek the Web for sales and customer support. The virtual agent is playing an emerging role in helping companies build smart and customer-friendly self-service on the Web, on the phone, on mobile devices also through video-call, thus integrating and assisting contact centre agents.

In order to maximize customer satisfaction, Web sites should be as informative and timely as communicating with a human. Staffing the site with intelligent agents as virtual service representatives is a viable alternative to having several employees working long shifts. With natural language processing and a knowledge base, the intelligent agent can answer questions effectively.

Virtual Assistants (VA) now offer online users entertaining, yet competent professional services: for example browser based consulting systems can be used from any computer to help in the area of customer care to reduce service inquiries. Natural language interaction, online translation, 3D-avatar technology and artificial intelligence create a powerful instrument to enable innovative services that are likely to find a wide acceptance among users. The experience of an online and live communication, simulating a personal one-to-one dialogue with an avatar as an advisor (e.g. interaction with an avatar in “Second Life”), can provide personal and entertaining opportunities of use. However, the simplicity of interaction, the possibility of seamless integration into existing platforms and workflows, the very human-like nature of the avatar users, as well as the use of natural language to navigate within the application is an extremely remarkable challenge.

The development of intelligent systems using innovative application designs often leads to new technological problems. The more the advisor looks and behaves human-like, the more users will expect in terms of flexibility, knowledge and intelligence. Increasing the intelligence of a system usually involves adding knowledge sources and processes to handle them. The result is another challenge: ways to improve the filtering, indexing, retrieval and presentation of relevant external knowledge, for example documents from websites, as a natural continuation of the dialogue.

In the case of the E-Mask interface, we believe that it may be applied in areas such as e-commerce, e-health, e-finances and e-learning, just to name a few. Costumers may seek information about certain products or services in the case of marketing campaigns or sales, or just need an expert “opinion” before ordering and final payment. A person may want to be “seen” by an online “doctor” to be able to express what he/she feels more openly and receive unbiased “treatment”. It could also be useful to talk to an account manager or to a teacher 24 hours a day. In this sense, the E-Mask may be useful in any online situation where virtual communication is required and quick adaptability to the “person” (agent) or even “privacy” is preferred. The agent size, voice, speed of the speech, balloon styles, etc. can be changed to meet different situations.

8. Conclusions

The paper contributes to the study of virtual communication in the context of VE by presenting a new VCI and explaining how it was conceived, using Microsoft Agent as the underlying technology, how it is structured and how it may enable dynamic reconfiguration by reducing the adaptation time of the end-users. We also discussed anthropomorphism and related it to our study. In the end, other agent and E-Mask applications were described.

However, further research is necessary:

1. On other technologies that may support virtual communication in VE;
2. On the development of our own software agents for the VCI;
3. On the study of the design and functionalities of the VCI;
4. On testing the VCI on people who work in enterprises and/or are geographically dispersed.

The study of other underlying technologies, the development of our own software agents, as well as the study of the design and functionalities should be realized within an interdisciplinary group. We plan to undertake these goals in the near future.

Bibliography

- André, E., Rist, T., Mulken, S. v., Klesen, M., & Baldes, S. (2000). The automated design of believable dialogue for animated presentation teams. In J. Cassell, J. Sullivan, S. Prevost & E. Churchill (Eds.), *Embodied conversational agents* (pp. 220-255). Cambridge, MA: The MIT Press.
- Badler, N. I., Allbeck, J., Bindiganavale, R., Schuler, W., Zhao, L., & Palmer, M. (2000). Parameterized action representation for virtual human agents. In J. Cassell, J. Sullivan, S. Prevost & E. Churchill (Eds.), *Embodied conversational agents* (pp. 256-284). Cambridge, MA: The MIT Press.
- Bradshaw, J. M. (1997). *Software agents*: AAAI Press.
- Cassell, J., & Vilhjálmsson, H. (1999). Fully embodied conversational avatars: Making communicative behaviors autonomous. *Autonomous Agents and Multi-Agent Systems*(2), 45-64.
- Conati, C. (2002). Probabilistic assessment of user's emotions in educational games. *Applied Artificial Intelligence*, 555-575.
- Cunha, M. M., & Putnik, G. D. (2002). Discussion on requirements for agile/virtual enterprises reconfigurability dynamics: The example of the automotive industry. In L. M. Camarinha-Matos (Ed.), *Collaborative business ecosystems and virtual enterprises* (pp. 527-534). Boston: Kluwer Academic Publishers.
- Dam, A. v. (2000). Beyond wimp. *IEEE Computer Graphics and Applications*, 20(1).
- DeSanctis, G., & Poole, M. S. (1997). Transitions in teamwork in new organizational forms. *Advances in Group Processes*, 14, 157-176.
- Gonçalves, P., Putnik, G. D., Cunha, M., & Sousa, R. (2006). User-interface architectures for ve dynamic reconfiguration: An initial analysis. In e. a. Pham D. T (Ed.), *Intelligent production machines and systems (proceedings of the 2nd i*proms virtual international conference)* (pp. 150-157): Elsevier.
- Jarvenpaa, S., & Leidner, D. (1999). Communication and trust in global virtual teams. *Organization Science*, 10(6), 791-815.
- Kitamura, Y., Tsujimoto, H., Yamada, T., & Yamamoto, T. (2002). *Multiple character-agents interface: An information integration platform where multiple agents and human users collaborate*. Paper presented at the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-02), New York.
- Lipnack, J., & Stamps, J. (1997). *Virtual teams: Reaching across space, time, and organizations with technology*. New York: John Wiley and Sons, Inc.
- Marsella, S. C., Johnson, W. L., & LaBore, C. (2000). *Interactive pedagogical drama*. Paper presented at the Proceedings 4th International Conference on Autonomous Agents (Agents-2000), New York.
- Nwana, H. S., & Azarmi, N. (1997). *Software agents and soft computing: Towards enhancing machine intelligence*: Springer-Verlag.
- Powell, A., Piccoli, G., & Ives, B. (2004). Virtual teams: A review of the current literature and directions for future research. *The DATABASE for Advances in Information Systems*, 35(1).
- Prendinger, H., Descamps, S., & Ishizuka, M. (2002). Scripting affective communication with life-like characters in web-based interaction systems. *Applied Artificial Intelligence*, 16(7-8), 519-553.

- Prendinger, H., & Ishizuka, M. (2003). *Life-like characters. Tools, affective functions and applications.*: Springer Verlag.
- Putnik, G. D. (2000a). Bm_virtual enterprise architecture reference model. In A. Gunasekaran (Ed.), *Agile manufacturing: 21st century manufacturing strategy* (pp. 73-93). UK: Elsevier Science Publ.
- Putnik, G. D. (2000b). *Bm_virtual enterprise architecture reference model* (Technical Report RT-CESP-GIS-2000-<GP-01>). Portugal: University of Minho.
- Putnik, G. D., Gonçalves, P., Sluga, A., & Cunha, M. M. (2008). Virtual environments for dynamically reconfigurable concurrent/collaborative engineering “virtual” teams. In *Cirp annals* (Vol. 57).
- Sycara, K., Decker, K., Pannu, A., Williamson, M., & Zeng, D. (1996). Distributed intelligent agents. *IEEE Expert*.