

Development of a Virtual Assistant for Alerts and Notifications in a Learning Environment

Ygor Amaral¹, Rodrigo L. Rodrigues^{2,4}, Alexandre M. A. Maciel³,

¹ Faculdade de Ciência e Tecnologia de Caruaru – Universidade de Pernambuco (UPE)
Rodovia BR 104, Km 62 – Nova Caruaru – CEP: 55.002-971 – Caruaru/PE

² Departamento de Educação – Universidade Federal Rural de Pernambuco (UFRPE)
Rua Dom Manoel de Medeiros, s/n, Dois Irmãos – CEP: 52.171-900 – Recife/PE

³ Escola Politécnica – Universidade de Pernambuco (UPE)
Rua Benfica, 455 – Madalena – CEP: 50.720-001 – Recife/PE

⁴ Centro de Informática – Universidade Federal de Pernambuco (UFPE)
Avenida Jornalista Aníbal Fernandes - Cidade Universitária, Recife - PE, 50740-560

ygor.amaral@gmail.com, rlr@ded.ufrpe.br, amam@ecom.poli.br

Abstract. *Increasingly, learning platforms have been part of the daily lives of teachers and students, either in a distance environment, either in a face to face learning. Several initiatives have been undertaken in order to provide resources that can enhance communication and interaction in educational contexts. Despite these efforts, confusing content offered by these platforms can often cause adverse effects such as lack of engagement, or in more severe cases, dropouts. In this scenario emerged the Avatar Education, an animated virtual assistant with synthetic voice to support online learning, integrated with LMS Moodle. This work aims to present the process of design, development and evaluation of a mobile module of Avatar Education as a way to improve communication and interaction with the platform. For its realization, a process of requisites survey and validation was conducted with users and experiments based on interviews served to validate the quality of the obtained solution. The results were promising and encourage the growth of new features to increasingly improve communication and interaction.*

1. Introduction

The ubiquitous availability of Internet, associated with the growing society interest to promote qualification through computational capabilities have changed the initial vision of distance learning, and the way it can deliberate on teaching and learning process (Flôres & Taroco, 2008). Multidisciplinary research involving computer science and education have contributed significantly in the planning, development and use of technological resources in order to offer new approaches of communication and interaction in educational contexts (Santos, 2009).

The Learning Management Systems (LMS), an important tool to support communication and interaction, provide several synchronous and asynchronous artifacts as a way to assist learning. Among these artifacts, discussion forum, chat and tool for activities and evaluation management are the most commonly used. In spite of this, the constant use of these environments can cause a feeling of high load of information for students, who normally feel confused given the great volume of content available.

This situation is considered quite usual in distance learning courses since it is part of its nature a high degree of students' interaction through the environment and a high daily flow of information. Despite this, the excessive load of information, presented in an unstructured way, can cause adverse effects such as lack of engagement in the actions of the course, or in more severe cases, dropouts (Santos, 2009).

In Maciel *et al.* (2014) was presented the design, development and evaluation of an animated virtual assistant with synthetic voice to support online learning, integrated with LMS Moodle, called *Avatar Education Plugin*. This study used a middleware called *Avatar Vocálico* presented in Rodrigues *et al.* (2012) responsible for generating the animation sequences and perform communication with the *Framework FIVE* (Maciel & Carvalho, 2010) responsible for generating synthetic voice.

Despite the good results obtained in Maciel *et al.* (2014), the use of the plugin in real environment presented some communication problems in order to provide a centralized mechanism for alerts and notifications. Thus, this paper presents the process of design, development and evaluation of a mobile module of Avatar Education integrated to LMS Moodle, as a way to improve communication and interaction with the platform. This work was conducted in cooperation with the Vocal Lab company and funded by Sebraetec Inova SEBRAE Program.

The next sections of this paper is organized as follows: section 2 presents the related technologies study necessary to conduct this research; section 3 shows the process of design and development of the mobile module and its architecture; section 4 presents experiments, results and discussion conducted to validate the app, and finally, section 5 presents conclusions and possibilities for future work.

2. Related Technologies Study

This section presents the study of related technologies necessary for the understanding of the research problem and for the development of the mobile version of virtual assistant.

2.1 Systems for Alert and Learning

The development of strategies that allow students to control their learning processes should encourage or promote metacognitive skills. So to help students self-manage their learning it is necessary that they are provided with educational artifacts that will allow them to review and monitor their performance and goals, in order to find satisfaction and success in certain learning environments (Araújo, 2012).

Thus, the control of the fundamental aspects of learning allows us to understand which aspects should be improved to achieve success. For Freire (2009), it is no longer sufficient to

convey information to students. It is necessary to "encourage them to be able of selecting, building and transferring their own knowledge."

Considering the large volume of content which is intrinsic to Learning Management Systems, it is necessary the research and development of alternatives to give support to students in order to lead them to self-manage their learning. Kampff (2009) states that more than good quality material and good technologies for courses' management, the presence of interventions becomes even more important in order to guide students.

In this sense, the development of alerting and notification systems that automate part of these interventions, such as: proposed tasks, available resources and the messages between students and teacher, boost the communication and interaction when coupled with virtual assistants, going to meet the human aspects that favor the feeling of teacher's presence during the execution of the course (Maciel *et al.*, 2014).

2.2 Avatar Education Plugin

The Avatar Education Plugin consists in a virtual assistant integrated to LMS Moodle to support online learning. With it, predetermined queries are realized in the environment's database, and according to student's status, personalized warning messages are presented through an animated character with synthetic voice (Maciel *et al.*, 2014).

The architecture established to facilitate the development of *Avatar Education Plugin* uses a web service to send requests for *Avatar Vocálico Middleware*, proposed by Rodrigues *et al.* (2012), responsible for generating the animation sequences and perform communication with *Framework FIVE* (Maciel & Carvalho, 2010), which is, in it turns, responsible for generating synthetic voice using *VL Synthesizer Engine*. Figure 1 shows the architecture overview.

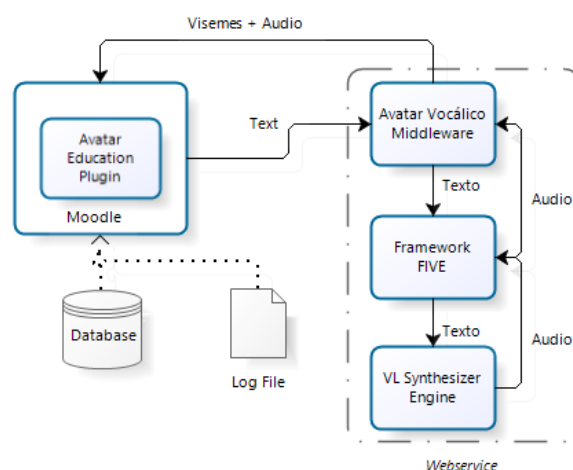


Figure 1. Avatar Education Plugin architecture (Maciel *et al.*, 2014).

2.3 Visemes based animation

According to Gratch (2002), animated virtual characters are audiovisual resources able to easily capture user attention and make the interaction activity more attractive and engaging. To have this happening effectively it is important to make a large multidisciplinary effort.

Knowledge in areas such as speech processing, natural language processing, artificial intelligence, computer graphics and animation are fundamental for the development of characters increasingly closer to reality.

Depending on the application given to the animated virtual characters, they can play different roles, such as instructors, assistants, presenters, attendants or vendors. Specifically in education area, these characters can be used as pedagogical agents in order to increase the effectiveness of education and training applications, or can also be used in interactive applications that help final users (Prola, 2003).

A trivial way to achieve modeling and animating faces with realism is through the use of visemes. According to Costa (2009), visemes are lip postures images visually contrasting each other, associated with the various speech sounds (phonemes). Figure 2 shows a general scheme of this process.

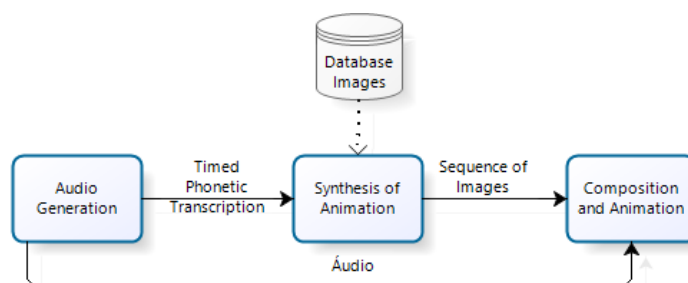


Figure 2. Visemes based on Animation process (Adapted from Costa, 2009).

The process of facial animation synthesis starts with the definition of the audio content, which may be the result of a recordable voice of an announcer or a speech synthesis system. From there we get the timed phonetic transcription from the audio. This transcription is composed of a sequence of phonemes that make up the phrase and their durations. This information is used to select the corresponding visemes on the database images, and after processing them, they generate a sequence of frames that will reproduce in the virtual face the articulatory movement. Finally, the composition and animation build the movement through various animation techniques such as metamorphosis, concatenative synthesis and trajectory of visemes (Costa, 2009).

2.4 Speech Synthesis

Speech synthesis is the process of creating a synthetic replica of a speech signal so as to transmit a message from a machine to a person, with the purpose of conveying the information in a message (Tabet & Boughazi, 2011).

According to Simões (1999), two characteristics are the most important on strategy determination: the signal quality – some applications require only a intelligible speech signal, others require that the signal quality must be as close as possible to natural speech; and the vocabulary size – simple systems working with fixed and small size vocabularies, on the other hand, more complex systems, like the Text-To-Speech (TTS) systems, require a very large vocabulary.

The process of synthesizing voice consists of two main steps. The first one is the Natural Language Processing (NLP) which consists in the transcription of an input text into a phonetic representation, and the second step is the Voice Generation in which waveforms are generated and the acoustic output is produced. These two phases are respectively called the high-level and low-level synthesis (Maia, 2006). A simplified version of the process is shown in Figure 3.

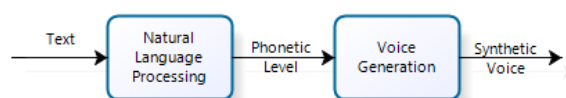


Figure 3. Speech Synthesis process (Adapted from Maia, 2006).

Two important resources designed for generating speech synthesis using HMM are: HTS (HMM-based Speech Synthesis Toolkit) and MARY TTS. In Maciel and Carvalho (2012) was presented the training of two synthetic voices using HTS to generate the VLSynthesizer Engine, while in Souza *et al.* (2014) was presented the training the same voices using MARY TTS. Both works were incorporated into the Framework FIVE (Maciel & Carvalho, 2010) allowing the use of male and female voices in Portuguese spoken in Brazil on multiple platforms.

3. Design and development

The conception and development processes of mobile version of Avatar Education were based in the Human-Computer Interaction (IHC) literature, specifically the method proposed in Interaction Design in Preece *et al.* (2013). It is composed of four activities: (1) identification of user needs; (2) development of design alternatives; (3) construction of iterative version; and (4) design evaluation.

The phase of identification of user needs involved a survey of the current situation of the available technologies to identify needs and opportunities for a tool development. This phase provided an elicitation of the main features of the application that generated as result the specifications of functional and non-functional requirements. Table 1 shows the main raised requirements.

Table 1. Main requirements for mobile module.

Requirements	Description	Actor
[login]	Use the same account registered in the Moodle environment.	User
[configurations]	Offer ways to configure profile and notifications.	User
[alerts]	Check status of students per course and send notification in accordance with the standard of the mobile platform.	System
[animation]	Each alert will be converted to synthesized voice and will be synchronized with the animation of the character.	System
[visualization]	The character must be shown in full screen with the options of the play, forward and rewind buttons.	User

In the design alternatives phase, caricatured scenarios technique (prototyping) was used to create a real use situation based on raised requirements. In this phase user tests were

conducted to identify possible problems in interaction flow interface, as well as potential usability problems. Figure 4 shows the prototypes built.



Figure 4. Prototypes built for mobile module.

In the construction of iterative versions phase a survey of the main technologies for application development was realized, and it was decided to use Android platform due to its largely use and low development cost. The requirement [*login*] was implemented as it had been prototyped, and the access validation is being performed directly accessing Moodle platform. The requirement [*settings*] was implemented as context menu, a native Android feature. The requirements [*alerts*] and [*animation*] were implemented to perform, in off-line mode, according to the period of time specified in the settings, being available to meet the requirement of [*display*] when requested by the user. For each one of these requirements, specifications of use cases were implemented. Figure 5 shows the interactive screens version.

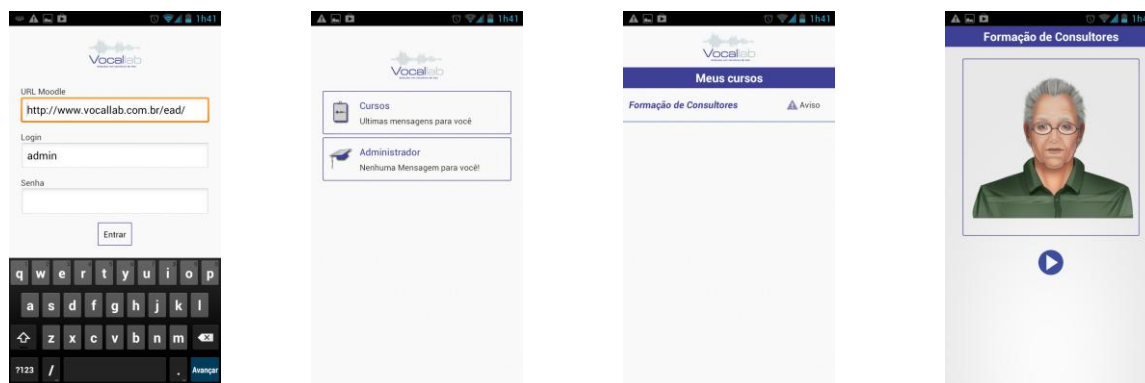


Figure 5. Interactive screens for mobile module.

In the design evaluation phase was realized an experiment with undergraduate students in a real teaching environment which will be detailed in the Experiments section.

3.1 Architecture

The architecture established to facilitate the development of mobile module of Avatar Education was based on the plugin architecture as described in Figure 6, integrating to it a communication manager module between the app and the Moodle environment. Figure 6 shows the integration architecture.

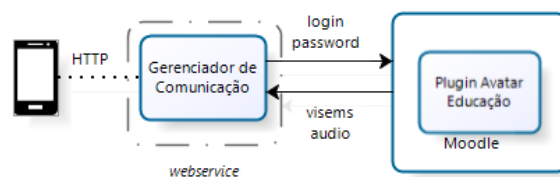


Figure 6. Integration architecture for mobile module.

The animation process of mobile module started from periodic requests as specified in requirement [alerts]. These requests sent the login and password information through a web service to the communication manager, which performed a query on the Moodle database aiming to know the student's situation in the course, and returned a JSON file with information about the visemes and synthetic audio. With this information, the app, which had previously stored images of characters, synchronized through concatenative synthesis the animation with the synthetic audio.

For the optimization of communication, the communication manager generated a security token for users, with a particular expiration deadline, to avoid recurring authentication in Moodle environment. Another feature implemented was an update flag to prevent recurrent downloads of visemes and synthetic audio. Measures like this prevent the use of data transmission.

4. Experiments

In order to evaluate the effectiveness of the proposed solution, a quantitative and qualitative study was conducted in which it was possible to make use of the app in a real learning environment. The experiment was conducted with 20 students from Information System discipline of Computing Engineering course at Universidade de Pernambuco. This discipline was conducted as a blended learning course in which textbooks were available, the completion of activities was required, and teacher's messages were available through the app.

4.1 Results and Discussions

The quantitative experiment was conducted through a questionnaire containing five issues, using a satisfaction scale numbered from one to five, where one was the worst and five was the greatest. These issues sought to assess the level of user satisfaction with respect to the facilitation of information management and to the identification of factors related to the potential application.

Due to the small number of participants involved in the experiment, it was decided to use analysis which non-parametric statistics such as the Wilcoxon test with continuity correction with significance level of 5%. This test replaces the t-Student test for paired samples, when the data are measured on an ordinal scale relying on posts (ranks) of intra pairs differences (Campos, 1987). Table 2 shows the results of average and test values.

Issues (1) (2) (3) e (5) presented significant differences between the Web version and Mobile Version means. Among these, the issues (1) and (5) presented higher average in Web version showing that these aspects are better accepted by students through the Moodle interface itself. On the other hand, the issues (2) and (3) were better evaluated in Mobile version. Thus, it can be argued that aspects related to notifications in the flow of learning are

better accepted when made through the Android App. Regarding the question (4), there were no significant differences between the two versions.

Table 2. Wilcoxon tests

Issue	Web	Mobile	<i>p</i> -valor
(1) Facilitation of information management.	4.833	3.750	0.02145
(2) Identification of possible backlogs activities.	3.250	4.416	0.02546
(3) Identification of new resources posted by professor.	3.333	4.500	0.03418
(4) Reception of personalized messages from the teacher.	4.083	4.416	0.20290*
(5) Sufficiency of information to guide the course.	4.833	4.000	0.00994

* There was no significant difference in average

For qualitative evaluation, a questionnaire with open-ended question was used to capture the suggestions about the use of virtual assistant to the management of information. Table 3 shows the main results.

Table 3. Suggestions for improvements from users.

Suggestions
(1) "I suggest improvements in the quality of synthetic voice".
(2) "There could be information identifying the number of notifications not heard".
(3) "It would be interesting if Avatar could specify the activity that is pending".
(4) "I think there could be a specific notification for courses' grades".

The results of this evaluation showed evidence of significant improvements in features of the Mobile Version. Regarding suggestion (1) about the quality of the synthetic speech, in Souza *et. al* (2014) new voices were presented and preliminary tests indicated that the voices sounded more natural. Regarding suggestion (2) which deals with modifications in the interface, this type of improvement involves prototyping and user involvement, to be held in subsequent experiments. Finally, the tips (3) and (4) involve the construction of new queries to the database and reformulating the text set for speech synthesis. Based on the architecture developed for this application, the incorporation of new queries and texts sets is perfectly configurable through the administrator system.

5. Conclusions and Future Works

This paper introduced the process of design, development and evaluation of a mobile module of Avatar Education integrated to LMS Moodle, as a way to improve communication and interaction with the platform. The design of application, based in method proposed in the Interaction Design of Preece *et al.*, provided an iterative development, with user involvement that led to a final solution that met real needs. The current status of the development, result of previous work, has modeled a previous architecture that was adapted to the requirements of the application. This provided a gain in productivity and maintainability. The experiments performed, even if on a small scale, allowed identification of the best features in relation to information management and significant improvements to the mobile version.

Therefore, we concluded that this study offered important contributions in two aspects. On the conceptual aspect we believe that the provision of an integrated LMS mobile Moodle module favored the combination of the concepts of self-management of learning and multimedia learning through a virtual assistant. On the technical aspect, this work has become innovative in order to add an alert system based on an existing architecture with an easy scalability and adaptability.

As for future works some initiatives are already being designed in order to perform experiments with larger samples in totally online learning environments. This test aims to assess the robustness of the architecture designed. Another ongoing initiative also proposes the use of Data Mining techniques to create an integrated intelligent module to Moodle environment. This module aims to provide personalized follow-up messages according to the usage profile of the students. In Cavalcanti *et al.* (2014) the first results were obtained and are the basis for modeling this type of message.

6. References

- Araújo, T. S. Autonomia no estudo: Artefato para Planejamento e Monitoramento em Ambientes Pessoais de Aprendizagem Móveis. Master Dissertation. Universidade Federal de Pernambuco, 2012.
- Campos, H. de. Estatística experimental não-paramétrica. Esalq. Piracicaba-SP. 230p. (1987).
- Clark, R. C., Mayer, R. E. (2008). e-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning. 2 ed. San Francisco: Pfeiffer.
- Cavalcanti, L. C., Maciel, A. M. A., Rodrigues, R. L. (2014) Avaliação do Perfil de Uso no Ambiente Moodle Utilizando Técnicas de Mineração de Dados. In *Proceedings of Simpósio Brasileiro de Informática na Educação*. Dourados, MS, Brazil, November 06-10, 2014.
- Costa, P. D. P. (2009) Animação Facial 2D Sincronizada com a Fala baseada em Imagens de Visemas dependentes do Contexto Fonético. Master Dissertation, Universidade Estadual de Campinas, 2009.
- Flôres, M. L. P. & Taroco, L. M. R (2008). Diferentes Tipos de Objetos para dar Suporte a Aprendizagem. *Revista Novas Tecnologias na Educação* vol. 6.
- Freire, L. Auto-Regulação da Aprendizagem. *Ciências e Cognição*, v. 2, p.276-286, 2009.
- Gratch, J., Rickel, J., André, E., Badler, N., Cassell, J., Petajan, E. (2002) Creating Interactive Virtual Humans: Some Assembly Required. *IEEE Intelligent Systems*, 17(4), 54-63.
- Kampff, A. J. C. Mineração de Dados Educacionais para Geração de Alertas em Ambientes Virtuais de Aprendizagem como Apoio à Prática Docente. Master Dissertation. Universidade Federal do Rio Grande do Sul, 2009.
- Maia, R. S. (2006) Speech Synthesis and Phonetic Vocoding for Brazilian Portuguese based on Parameter Generation from Hidden Markov Models. *Doctoral Thesis*. Nagoya Institute of Technology, 2006.
- Maciel, A. M. A., Carvalho, E. C. B. (2010) FIVE – Framework for an Integrated Voice Environment. In *Proceedings of International Conference on Systems, Signals and Image Processing*. (pp. 239-242) Rio de Janeiro, RJ, Brazil, June 17-19, 2010.

- Maciel, A., Carvalho, E. (2012) Integration and Evaluation of an HMM-Based Text-To-Speech System To FIVE. In *Proceeding of International Conference on Systems, Signals and Image Processing*, (pp. 633-636) Vienna, Austria, April 11-13, 2012.
- Maciel, A. M. A., Rodrigues, R. L., Carvalho, E. C. B. (2014) Desenvolvimento de um Assistente Virtual Integrado ao Moodle para Suporte a Aprendizagem Online. In *Proceedings of Simpósio Brasileiro de Informática na Educação*. Dourados, MS, Brazil, November 06-10, 2014.
- Moreno, R., Mayer, R. E. (1999). Cognitive Principles of Multimedia Learning: The Role of Modality and Contiguity. *Journal of Educational Psychology*, v.92, n.2, 358-368.
- Preece, J., Rogers, Y., Sharp, H. (2013) Design de Interação – Além da Interação Homem-computador, São Paulo, Bookman, 2013.
- Prola, M. T. (2003) Modelagem de um Agente Pedagógico Animado para um Ambiente Colaborativo: Considerando Fatores Sociais relevantes. *Master Dissertation*, Universidade Federal do Rio Grande do Sul, 2003.
- Rodrigues, R. L., Maciel, A. M. A., Carvalho, E. C. B. (2012) Desenvolvimento de uma Ferramenta para a Produção de Mídias utilizando Personagem Animado com Síntese de Voz. In *Proceedings of Simpósio Brasileiro de Informática na Educação*. Rio de Janeiro, RJ, Brazil, November 26-30, 2012.
- Santos, L. M. A (2009). A Inserção de um Agente Conversacional Animado em um Ambiente Virtual de Aprendizagem a partir da Teoria da Carga Cognitiva. *Doctoral Thesis*. Universidade Federal do Rio Grande do Sul (2009).
- Simões, F. O. (1999) Implementação de um Sistema de Conversão Texto-Fala para o Português do Brasil. *Master Dissertation*. Universidade Estadual de Campinas (1999).
- Souza, D., Saturnino, L., Maciel, A. M. A. (2014) A Portability Evaluation of Brazilian Portuguese voices produced with MARY TTS. In *Proceedings of International Conference on Systems, Signals and Image Processing*, (pp. 95-98) Dubrovnik, Croatia, May, 12-15, 2014.
- Tabet, Y., Boughazi, M. (2011) Speech Synthesis Techniques – A Survey. In *Proceeding of 7th International Workshop on Systems, Signal Processing and their Applications*, (pp. 67-70) Tipaza, Algeria, May 9-11, 2011.