THE DEVELOPMENT OF A CHATTERBOT FOR ENVIRONMENTAL EDUCATION

Vinicius Tonelli de Oliveira
Elvio Gilberto da Silva
Patrick Pedreira Silva
Universidade do Sagrado Coração
Brazil

Abstract

With the accelerated growth of cities around the world, the need of learning about our environment becomes inevitable. In the face of such a situation, it would be of great importance to implement an educational software, easy to use and understand, so that children and adolescents could acquire an awareness in a practical and objective way. On this basis, the present research is aimed the development of a computational tool – a Chatterbot, with an educative and informative potential that has information related to Environmental Education as knowledge base.

Introduction

The relation between humanity and nature is understood through mankind exploring natural wealth and resources to satisfy its will. Since the time when records were ca**r**ved in cavern walls, man**'s** exploration of the environment has intensified**,** and the need of consumption has noticeably increased. However, the search for these resources interferes directly and indirectly on how we live now: soil degradation, pollution (rural and urban), waste creation, toxic and nuclear waste and global warming are examples of man's interference with nature. In the past decades, we have witnessed the emergence of numerous movements in favor of the environment. In several countries, programs and strategies have been undertaken in order to curb environmental degradation and / or to find new alternatives for production processes and less impacting consumption (Rodrigues & Colesanti, 2008).

Technology, seen as responsible for the environment degradation, can be one of the alternatives to minimize the environmental problem. By using new mechanisms of communication between man and machine, we can introduce basic concepts of nature preservation, in which the learning is of utmost importance to inform and clarify the current situation of the environment. However, the use of new technology must be thought **of in** a practical way for learning. It has to be at the same time didactic and with low complexity, which is hampered by the instability of the language.

Through an application with a concept of ***a**rtificial **i**ntelligence* (AI), in which the program thinks and interacts as a human being, a quick and didactic learning about the environmental issue becomes possible.

According to Fernandes (200**5**), ***a**rtificial **i**ntelligence* is the modeling of intelligence treated as a phenomenon. Understanding intelligence is not an easy paradigm to unravel, since there is no complete theory about the human mind and the reasoning processes.

A Chatterbot would be a way to promote awareness, because mechanisms of human language formation (semantic, syntactic, pragmatic and morphologic analysis) are used

simulating a normal conversation, in an almost human way, which could replace a human being. Unlike a human being, the Chatterbot has the power to save all the information that is imposed to it.

Russel and Norvig (2004) affirm**,** "If we are going to say that a given program thinks like a human, we must have some way of determining how humans think" (p.5).

Environmental Education

By definition, ***e****nvironmental education* can be understood as "[...] the processes by which the individual and the collectivity build social values, knowledge, abilities, attitudes and capabilities geared to the conservation of the environment, population's asset of common use, essential to a healthy quality of life and sustainability" (Política Nacional de Educação Ambiental - Lei nº 9795/1999, Art 1º **[**free translation**]**)**.**

According to the Environmental Education Program (2005), environmental education requires a perspective of complexity to the social level in which its implication must interact in different layers of reality: objective, physical, abstract, cultural and affective. In addition, it becomes the fundamental instrument of environmental management along with social interaction of its rights and social inclusion.

Informatics Applied to Environmental Education

There are many ways to understand technology. For some people, it is composed of scientific knowledge applied to human**s**' benefit. It should be understood widely, like any other artifact, method or technique created by human**s** to make their work lighter, their locomotion and communication easier, or simply to make their lives more satisfactory, pleasant and fun. In this sense of amplitude, technology is not something new; it is as old as humankind (Chaves, 2007). Technological advances have been created in order to have more knowledge and seek better life conditions (Damasio, 2007).

Currently, environmental education is necessary to change the increasing state of social-environmental degradation, in which the educator has a mediator function in the construction of environmental references and should know how to use them as techniques for the development of a social practice focused on the concept **of nature**. The information is not how it used to be. The large expansion of the Internet, media outlets, cyberspace, **and** education for citizenship represents the possibility to motivate and sensitize people to transform the many forms of participation in the defense of quality of life. Therefore, environmental education increasingly takes on a transforming function**,** becoming an essential objective to promote sustainable development (Silva, 2007).

Artificial Intelligence

In the last decades, with the increasing complexity of the problems to be treated computationally and the amount of data generated by different sections, the need of more sophisticated computational tools, became clear, reducing the need of human intervention and dependence o**n** specialists. Thus, these techniques would have to be able to create themselves from experiences, a hypothesis, or function, able to solve a determined problem.

The word *intelligence* comes from the Latin inter (between) and legere (choose). Intelligence means what allows the human being to choose between two things. Intelligence is the ability to accomplish on an efficient form a determined task. On the other hand, the word *artificial* comes from the Latin artificiale, which means something not natural, in other words, which was produced by the human being. Therefore, artificial intelligence is a kind of intelligence produced by the human being to endow machines of any kind of ability simulating the human intelligence (Fernandes, **200**5).

Artificial intelligence developed many ways of simulating human conversation, bringing about, among other technologies, the appearance of the chatterbots, which are simulator programs of a conversation with another person.

They are scheduled to talk about the most various subjects (…) they are also computer sophisticated programs which can understand and answer correctly sentences and questions done by users as **if** they were people in a chat room (Siqueira, 2005).

Chatterbot

There are many kinds of intelligent agents, a chatterbot can be classified as a reactive agent because it executes an operation according to the perception of **its** sensors. This kind of agent can contain learning systems, with the use of the tag "learn" on AIML (Silva, 2012).

The term *bot* is an abbreviation of the word *robot*, an agent that works for a system or user, simulating human activity. Chetterbot or chatbot is one of the categories of bot that has as function to simulate conversations with users from its environment (Tarouco et al., 2003).

According to Othero and Menuzzi (2005), in 1966 Joseph Weinzenhaum developed Eliza, the first chatterbot able to simulate dialogues. It was a simple program **that** stimulated patients to deepen more details of their problems during the simulation, thus creating, a kind of psychologist, or sentimental counselor. Initially, Weinzenhaum**'s** aim was to create only a conversation program using a system based on types to construct sentences through the patient's answer.

Methodology

Exploratory research helps the researcher to know which of the various options are applied to the research problem. Besides that, it may also help to establish the priorities for researching. These priorities may be established because a particular explanatory hypothesis that emerged during the exploratory research may seem more promising than others. Furthermore, exploratory research can generate information about the conduction practice possibilities of specific researches (Mattar, 2012).

Initially, this project was characterized as exploratory research, which attempted to study the system's integration using three languages (AIML, MySQL and PHP), for the construction of a specific prototype, more specifically, a chatterbot specialized to help in environmental education. The research proposal was developed in two distinct stages: a stage to analyze the theoretical aspects and a practice stage in which there was the system implementation.

In the first stage, theoretical materials from various subjects were studied, which are related to the work scope, involving areas such as artificial intelligence, natural language processing, AIML, chatterbots and. As a product of this stage, the developing of a type was proposed, for the development of the conservational agent.

In the second stage, there was the proposed chatterbot implementation. In other words, the conservational agent's installation and customization, its knowledge base manipulation, and then, the necessary tests.

The tags AIML used for the chatterbot implementation were created starting from the type in order to maintain the consistency of the knowledge base, as well the coherence of answers provided by the agent. The tags used are the same described in the paper regarding to the language AIML.

Study Place and System Structure

The AIML knowledge base is indispensable for the chatterbot Ema execution because all the bot knowledge is present in it. As mentioned before, it was divided in two parts: one with a standard knowledge base, generated to attend ordinary context entries with more frequency of use, and the other with specific knowledge in which the conversational agent will act. Access to the conversational agent occurs according to Figure 1.

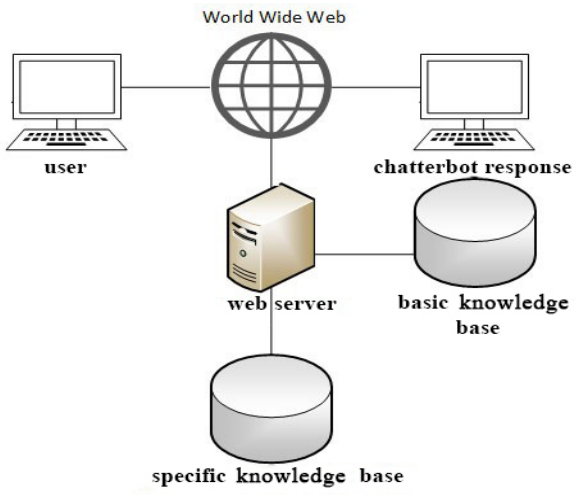

Figure 1. Simplified system architecture.

As it can be observed in Figure 1, the user**,** by means of a computer with internet access**,** can have a conversation with the bot using the HTTP protocol. Its communication is possible through requisitions and answers. The web server is composed of database servers MySQL and PHP. The PHP server is responsible for the parser AIML execution. The working of the two knowledge base**s** is done through MySQL server.

The Use Case Diagram presented in Figure 2 describes the scenario and**,** at the same time, shows the system functionalities.

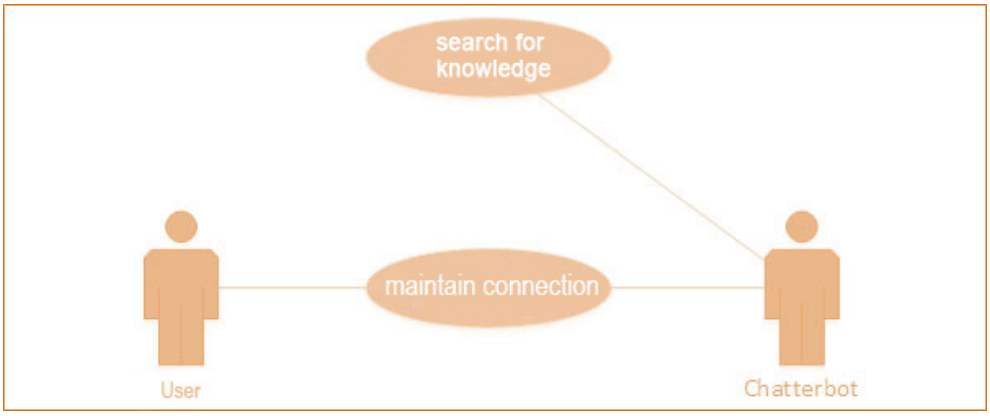

Figure 2. Use Case Diagram chatterbot.

As Figure 2 illustrates, the system is composed of two actors – the user, who has as objective to find a clarification for his/her doubts, and the chatterbot, which interacts with the user. There are also two use cases, one that illustrates a dialogue between the two actors and another that represents the information search on the knowledge base.

To separate the standard knowledge from the specific in conversational agents, two files were created in the "aiml" format. In the first, only the content regarding to common character questions was inserted, whereas the second contemplates sentences regarding the approached theme.

After the creation and edition of these files, their upload was carried out through the option Upload AIML, which is presented in the Program-O Administrative Area. Loading the file ". aiml**,**" a knowledge base is immediately created, so it is automatically erased and recreated. This way, the bot's knowledge could be worked using both files. It is worth highlighting that modifying the information exist**ing** in the files makes it necessary to upload the updated files again.

Updated Software

The *chatterbot* was developed using the platform Program-O. Both **k**nowledge bases were written in AIML language, which is based on XML language, and developed with the purpose to create conversations in a natural language. The updated software used to create *chatterbot,* to manipulate its *AIML* knowledge base and to edit the configuration files was the Notepad++6.8.1, which is a free text editor, versatile and in Portuguese, designated to editing written files in many languages.

We also used the XAMPP v3.2.1*,* which is a software that installs and automatically configures many servers, such as, APACHE (HTTP web server), MySQL (database server) and FTP (files server). Another tool available by XAMPP is the PHPMyAdmin, which is a database manager that facilitates the administration of databases on MySQL.

Developing the database knowledge**.** Portuguese was found to be a problem for a search mechanism simply because it has accentuation. Through a study of a *PLN* mechanism, it was planned that each pattern that had accentuation must be created in different ways, for example, the pattern *"VOCÊ ESTÁ ÓTIMO"* (YOU ARE GREAT) must have its variances with/without its respective accentuation, such as *"VOCE ESTA OTIMO"* (YOU ARE GREAT) or *"VOCÊ ESTA OTIMO"* (YOU ARE GREAT).

The disadvantages in developing a knowledge base with these variations is that the information would be redundant. To solve this problem, patterns were created following this model *"VOCE ESTA OTIMO"* (YOU ARE GREAT) - with no punctuation. Thus, before searching for the information on the database, the sentences written by the users will go through a pre-processing in which the punctuation will be removed. The pattern found on the platform makes a treatment on the sentences typed by the users, removing symbols, retaining only the letters. The punctuation removal was added to this treatment.

The Standard Knowledge Base

The standard knowledge base was created from the examples of other categories found on the Internet; however, they were all written in English. After a study and analyses of these examples, it was possible to create similar patterns to the ones found on the Internet. They could be found in salutes, greetings, storage of name or specific knowledge base and treatment of unknown sentences.

The Specific Knowledge Base

The specific knowledge base was developed through base questions and answers on several themes about the environment. A key term was taken from each question; those were used to create categories and subsequent themes to facilitate the search, such as, in the question *"O que é meio ambiente?"* ("What is the environment?"), the words used to create a new category *"que"* (What) inside the theme *"meio ambiente"* (environment).

Tests

With the knowledge base inserted, little content related to the theme was attributed, due to the variety of subjects and contents diffusion found during the gathering of requirements. As the knowledge base was filled, the developer performed the functional test, that is, the system was being verified without having access to the source code, knowing only how it should behave through the access modes with the right questions.

Through the creation of themes referring to the covered topics, it was possible to provide the users with an easy access to the main questions treated by the theme, that is, even if the user does not know the right question to be asked. If a theme the conversational agent recognizes is mentioned, the user will be taken to the section of such theme. To avoid problems with the communication, a selection list was created, which can help the user decide what to ask to the bot, as illustrated in Figure 3.

Figure 3. Selection question list.

To allow the use of the illustrated list in Image 3, a table was created in the chatterbot's database containing questions pre-set by the developer. At every reload on the page, eight randomized questions are brought to the selection list.

Test Run of a Conversation

Some test runs were made with different contexts to present the bot's behavior before some questions. By default, in the beginning of every conversation, Ema – the *chatterbot* introduces itself in a friendly way and encourages the user to ask a question. Figure 4 illustrates the introduction screen.

Figure 4. Conversation opening screen.

Through the database pattern it is possible **t**o greet and ask personal questions to the conversational agent, as observed in Figure 5.

Figure 5. Greeting the conversational agent.

The topics of the covered themes about environmental education can be accessed in two ways: just saying the theme or writing you wish to talk about such **a** theme. Figure 6 demonstrates this situation.

Figure 6. Defining the theme that will be covered.

It is also possible to ask a question by using the selection list that can be found right beside the button *"Dizer"* ("Say"). When a help option is selected, the area in which the user would supposedly speak will automatically be completed, that is, taking the role to help the user to create the desired question. Figure 7 demonstrates this situation.

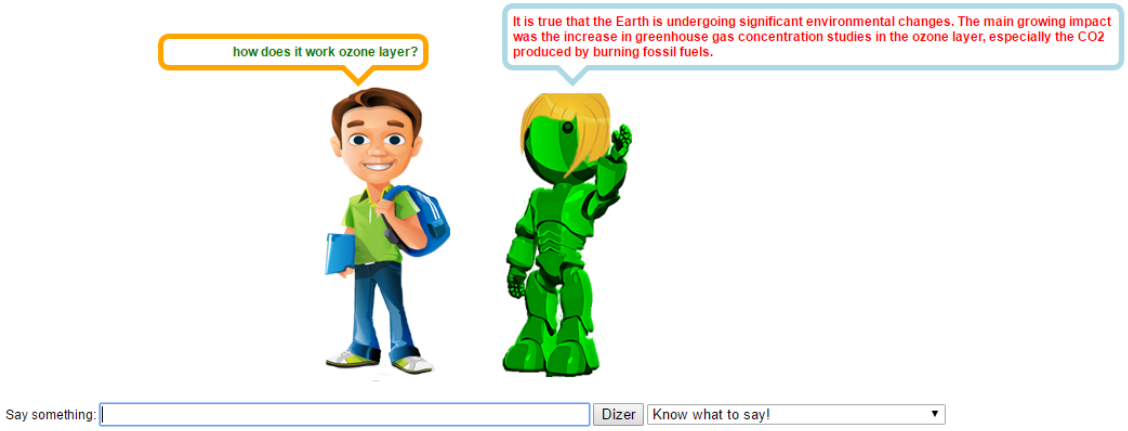

Figure 7. Using the help option to form a question.

Final Considerations

This study presented the stages of the development of a chatterbot to assist in environmental education, highlighting the interaction between the virtual agent and the student, providing a higher interactivity between educators and professionals of the area, stimulating the learning in a didactic level.

With the development of a chatterbot it was possible to conclude that the field of **artificial intelligence** study is an area that can be more explored to assist other areas and**,** where it is possible to create a source of artificial knowledge, which can be applied to other activities.

In addition, some adaptations are necessary to develop an intelligent agent that works with the Portuguese language, as it is recommended that the agent can be capable of performing a complete recognition of the spelling.

Another factor to be noted is that the process of developing a knowledge base requires more creativity than knowledge, because it is fundamental to imagine what can be asked to the agent, stimulating the creativity. Another relevant issue is that, in order for the agent to be more precise in the answers, the database must be subjected to a continuous process of improvement and refinement, which requires considerable time and study.

For this refinement, it is possible to use a logging mechanism that contains all the conversations already made by the agent and the user, which assists in clarifying the doubts about the improvement.

This project contributed as a multidisciplinary proposal, which included computer science as a support tool in the environmental education, thus, promoting greater interaction between educators in this area and consequently improving the results in the treatments. Computer science does not replace real human interaction yet, but it offers benefits helping students learn better.

References

Chaves, E. O. C. (2007). A Tecnologia e Educação. [Technology and Education]. Disponível em: <http://smeduquedecaxias.rj.gov.br/nead/Biblioteca/Forma%C3%A7%C3%A3o%20Continuada/Tecnologia/chaves-tecnologia.pdf>

Damasio, L. M. (2007). Tecnologia na educação. [Technology in Education]. Disponível em: <http://nte-estrela.pbworks.com/f/Tecnologia+na+educa%C3%A7%C3%A3o.doc>

Fernandes, A. M. R. (2005*). Inteligência artificial: Noções gerais.* [*Artificial intelligence: General notions*]. Florianópolis, Brazil: Visual Books.

Mattar, F. N. (2012) Pesquisa de marketing, edição compacta. [Marketing research, compact edition (5th ed.)]. Rio de Janeiro, Argentina: Elsevier.

Othero, G. Á, & Menuzzi, S. M. (2005). *Linguística computacional: Teoria e prática*. [Computing linguistics: Theory and practice (2nd ed.)]. São Paulo, Brazil: Parábola Editorial.

ProNEA - Programa Nacional de Educação Ambiental. [National Program of Environmental Education]. (2016.). Disponível em: <http://portal.mec.gov.br/dmdocuments/publicacao1.pdf>

Rodrigues, G. S, & Colesanti, M.C.M. (2008). Ducação Ambiental E As Novas Tecnologias de Informação e Comunicação [Environmental education and information and communications technology] *Society and nature, Uberlândia, 20*(1), 51-66.

Russel, S., & Norvig, P. (2004). *Inteligência Artificial [Artificial Intelligence* (2. ed.)]. (2004). (Tradução: V. D. de Souza) Rio de Janeiro, Brazil: Elsevier.

Siqueira, R. A. (2005). Robôs com Inteligência Artificial. [Robots with Artificial Intelligence]. Robô Ed. Disponível em <<http://www.ed.conpet.gov.br/materias/2005_comciencia.php>

Author Details

Vinicius Tonelli de Oliveira
vinicius.tdo@hotmail.com

Elvio Gilberto da Silva
egilberto@uol.com.br

Patrick Pedreira Silva
patrickpsilva@gmail.co