An Intelligent Agent and an Adaptive Search Engine to support tutoring activities on-line

Orlando De Pietro[a], Carmelo Piu[b], Maurizio De Rose[a] and Giovanni Frontera[a]

[a]Department of Economics and Statistics;
[b]Department of Education Sciences – University of Calabria

{depietro, mderose, gfrontera}@economia.unical.it;
carmelopiu@yahoo.it

Abstract

Concentrating our attention on the interaction tools in a web learning environment, an open source software system, called TutorBot which can be implemented in any e-learning platform supporting the tutoring activities on line, is described in the present work. This software is included among the category of intelligent agents and can be used as an interface between student and knowledge base to allow flexible and effective forms of tutoring on line. The essential aspect is moreover the integration of an Adaptive Search Engine (ASE) with TutorBot, that allows to provide for possible lack of information that may be found in the knowledge base, but above all thanks to the used adaptivity techniques, guarantees high levels of personalisation taking into account the user's profile during the interaction phases.

Edizioni Erickson – Trento
www.erickson.it

Je-LKS Journal of e-Learning and Knowledge Society — *Vol. 2, n. 1, marzo 2006 (pp. 97-110)*

1. Introduction

An e-learning system represents a particular informative system based on the web (WIS – Web Based Information System), that allows to classify, store, manage and assign teaching contents and in which the synergy between human and technological resources becomes fundamental and peculiar. In the course of time, e-learning has changed its proprieties which considered it as a simple supply system far from the teaching contents, in which the student and the teacher are physically in different places, still evolving until allowing, on the one hand, the students to reach high levels of commitment and interactivity, thus encouraging the main corollaries of the «constructivist» and «collaborative» approach, and on the other hand to put at the teaching Team's disposal — teacher, his/her assistants, tutor —, new tools that permit them to realise and manage the Learning Objects (LO), as well as the monitoring and assistance of the student's learning activity. Everything is obviously coming in line with a simplification of the interface man-machine, that becomes essential in an e-learning environment and has to be as direct and simple as possible. Moreover, as we first said, the public diffusion of the web-based services is dealing with a continuous increase of the users' number and typology, which is also very different either from the subjective point of view (socio/demographic features) or from the objective one (aims and purposes), determining a certain lack of balance between typology of use and typology of the accessible information, content management systems, information research tools (search engine), leading to characterize tools able to guarantee a personalisation at the user's level, even more superior concerning the contents and the available resources. The studies concerning the so-called «adaptivity» of the systems and of the computer services to the user's specific profile are found in this optics. As a matter of fact, we find it much more effective, in a context characterized by a high differentiation of the use and the contents, a system able to «adapt itself» to the requirements of every single individual to be able to give a personal answer to his/her univocal requirements; the limits of being able to reach compromises in terms of identification of the optimal lay-out (Web-Adaptive) and presentation of the interest contents (Search-Adaptive) of the user are overcome in this way (Appratto, De Pietro et al., 2003). Specifically, considering the influence that technology and in a special way the communication technologies can have in an e-learning context, the attention of the authorized personnel (scholars/experts) is lately directing towards the development of the technological resource through the project and implementation of tools that promote the new paradigms which are asserting themselves in the field of training on-line, such as the just-in-time learning, the collaborative learning, but above all this attention is oriented to the realisation of automatic systems, in open source modality, able to develop these activities that mainly consider the use of human figure (De Pietro, Appratto, 2002). In an e-learning context, one of the figures who plays an important and at the same time strategic role is the «tutor», since he/she is determining for the fulfilment of an efficient learning level from the students, which the success or the failure of a FAD type procedure very often depends on.

Consequently, concentrating our attention on the activities carried out by this figure and considering the suppositions outlined in the premise and the advantages offered, in our opinion, by the web adaptive systems, an open source software module called TutorBot, to support the tutoring activities on line is presented in this article. Belonging to the category of the intelligent agents, TutorBot is integrable in whatever e-learning platform. Using the AIML language (Artificial Intelligence Markup Language) for the data construction (Thomas, 2001; Wallace, 2001) and a proper Adaptive Search Engine (ASE), TutorBot can be used as an interface between student and knowledge base in an e-learning platform, to guarantee forms of tutoring on-line flexible in terms of space and time. The student, in his/her own language interacts with the intelligent agent to carry out the retrieving of the teaching contents and at the same time to receive «assistance», exactly as a human tutor would do during a chat session (Ali, Channarukul et al., 2001; Descamps, Prendinger et al., 2001). We have to stress the fact that the content is not only presented in a textual format but, thanks to the AIML structure, it is also possible to manage the multimedia contents, making the whole learning process more effective. TutorBot's aim is therefore to reduce the activities carried out by man — in our case by the tutors — in an e-learning platform with the advantage to be always available (ChenSession, Revithis et al., 2002; Ishizuka, Jatowt et al., 2003). It is obvious that TutorBot does not have to be considered as a substitutive element to the human figure and/or to all the other synchronic and asynchronic communication tools, but it has to be seen as a support and integration system for the improvement of all the e-learning processes.

2. The AIML Knowledge Base

The intelligent agent, reflecting the OpenSource philosophy, founds its main features on a description of the knowledge base in AIML language (Artificial Intelligence Markup Language). TutorBot AIML knowledge Base can also constitute a complement to the Knowledge Base of the E-Learning platform in which the agent is integrated. The construction rules of the data, specified by the AIML language, define classes of data objects in which the information are captured. The data, structured in AIML, are contained in ASCII files that entirely constitute TutorBot's knowledge bases. However, the agent does not directly question the AIML files, but in order to speed the information retrieving operations, the data contained in the AIML files are stored in an appropriate Database to which TutorBot can have directly access. The importation mechanism of the data in TutorBot's Database is completely automatic through appropriate modules. In this way, the agent can implement its own Knowledge Base with information belonging to different contexts, still if structured according to the AIML rules.

The process at the basis of TutorBot's knowledge management system is schematised in figure 1.

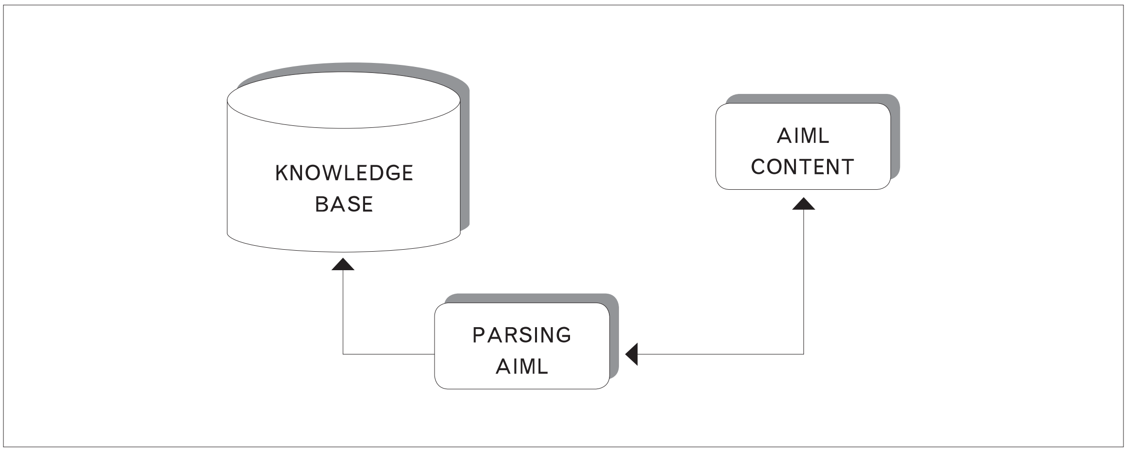

Figure 1 Import-Storing Metadata AIML.

The object Topic identifies the thematic classes and it is indispensable to better classify and index the information; it can contain 1 or more Categories and represents the element according to which the treated topic of these categories contained in it is defined. In its turn, the element Category contains the elements pattern and template; inside the Topic a lot of categories can be nested in, according to the complexity of the content that we want to describe. The metadata Pattern and Template are essential for the action stimulus-answer that guarantees the interaction between the artificial tutor and the learners, since the information that will serve for the matching with the requests in the question phase are contained in the first one, and the corresponding answers in the second one.

The general structure regarding these last data objects is represented as follows.

```
<!-- Category: aiml-category-elements -->
 <aiml:pattern>
  <!-- Content: aiml-pattern-expression -->
 </aiml:pattern>
  <aiml:template>
   <!-- Content: aiml-template-elements -->
  </aiml:template>
```

The above-mentioned mechanism stimulus-answer of AIML is elaborated according to construction rules of proper AIML data such as: symbolic reduction of text, sentences union and scission, synonyms management, etc. (De Pietro, Frontera, 2005).

3. Tutorbot: system functionality and implementation

In order to underline the potentialities of TutorBot in an e-learning environment, the interface of Tutorbot will be illustrated in the following paragraphs by focusing on its functionalities and implementation inside a didactic environment web based. In the specific, the phases of the interaction learner-Tutorbot will be described, by highlighting the potentialities offered by the agent in the cases of «informative lack». Finally, in order to underline the consequential potentialities from the use of the agent regarding the «profile» of the student, an example of interaction/dialogue will be introduced. This will allow to achieve such objective.

3.1 The Interface

TutorBot is introduced with a very simple interface, that the student uses during the interactive phases with the intelligent agent and through which he/she submits his/her own query in natural language; this latter is carried out by TutorBot, after appropriate operation analysis on the input (lexical reduction), by means of a software module (for an in depth study see De Pietro, Frontera, 2005). The Query can be in whatever form, that is it can explicitly refer to the information requirement relating to the knowledge base or it can be a conversation between the student and the robot.

In the figure 2, the interface of TutorBot, that can be implemented in any e-learning platform, is illustrated.

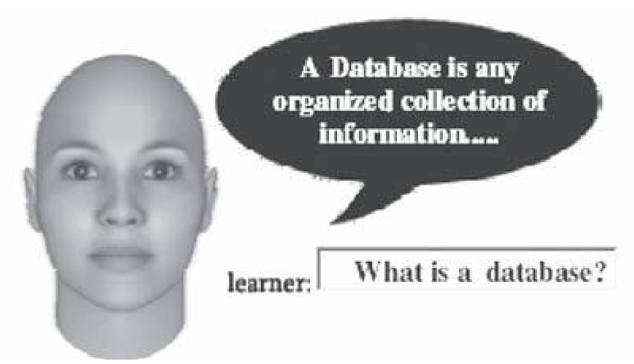

Figure 2 Simple Interface of TutorBot.

As it can be noticed by the interface represented in the figure above, the learner directly formulates his/her own question using a simple *text-box*, and the answer is given by TutorBot inside a special area destined to such purpose. In fact, the simplicity of the interface and of the interaction developed during the different phases of dialogue are set to simulate the tutoring process at its best as it has been conducted by a human tutor. It can be observed, that the *face* of TutorBot has been built on purpose using the *technology 3D*, to allow the movement of the mouth and the eyes, with the purpose to make the interaction more engaging. In fact, in contexts of IA it does not only need to pay attention to the technical aspects of the knowledge management, but it is also necessary to refer to the environment in which the interaction man-machine happens, making it more and more representative of the reality.

3.2 Functioning of the system

TutorBot has been tried out within the web-portal of the course of «Informatica per il turismo» [1] in the Faculty of Economics, University of Calabria. The integration of the agent inside a pre-existing platform allows to exploit some information found in the data base of the web-portal itself, as data which identify the profile of the students enrolled to the course. Such data are, for example, the courses which each learner belongs to, some data produced by *tracking* during the use of web-portal, etc.

All the information found trough the web-portal are used to profile the learner on the basis of personal characteristics, being their profiles of a great value to be able to manage adaptive systems.

On the basis of these preliminary remarks, it is possible to understand how TutorBot works, by taking into account (figure 3) of the previous paragraph, by analysing the output of the agent which is the result of what happens in background.

As shown in figure 3, to the question asked by the learner in natural language (TutorBot Knowledge Base is in Italian language): «Fammi vedere i database relazionali e la struttura» (in English language: «Let me see the relational database and the structure»), the agent answers with a PATTERN included in its own knowledge base, nevertheless, the agent, in order to do that and eliminate any chance of error, firstly analyses the data regarding the learner who has asked the question, data which are then used to profile the learner and match the final profile with the proper sections inside the AIML knowledge base identified by the topic (adaptivity).

[1] Application link: http://www.economia.unical.it/corsi/depietro8/modules.php?name=TutorBot.

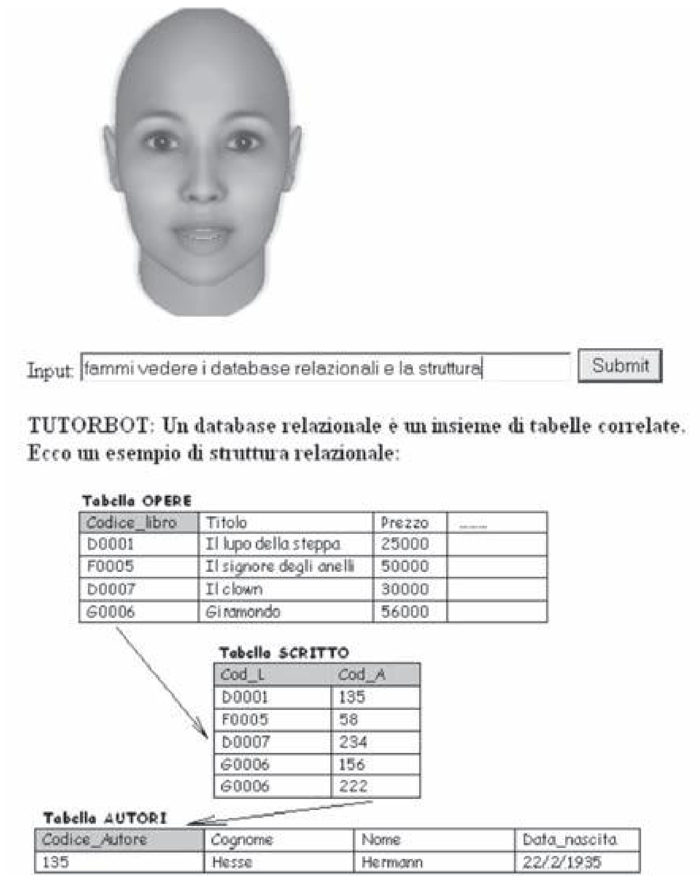

Figure 3 Interface of TutorBot in multimedia contents management.

To clarify the concept, it can be said that a learner who follows a course of computer science, will have a preference of correspondence inside the knowledge base AIML on topics associated to the subject «computer science», therefore it is possible to avoid confusion in getting a topic of conversation between TutorBot and the learner. Obviously, the topic identified may also be different from topics connected to computer science if the learner will shift the conversation on different topics.

Indeed the agent can autonomously pick up a topic of conversation, from (as it was pointed out before) the profile of the learner as well as from the same conversation. This is possible thanks to the rules entailed in AIML, which you can go to for further information.

Going back to the example in figure 3, to the question (in Italian language): «Fammi vedere i database relazionali e la struttura» (in English language: «Let me see the relational database and the structure»), the following AIML code can be identified inside the knowledge base (in Italian language):

```
<aiml>
<topic name=»Database»>
<category>
<pattern> * Database _ struttura
</pattern>
<template> Un database relazionale è un insieme di tabelle correlate.<br> Ecco un esempio di struttura relazionale:<br>
<img src=»database_schema.jpg» width=»100» height=»100» alt=»» border=»0» />
</template>
</category>
</topic>
<aiml>
```

Note 1 – The AIML Knowledge base of TutorBot is in Italian as it is full of informative contents. For this reason it was decided not to provide any translation.

The same example in English language:

```
<aiml>
<topic name=»Database»>
<category>
<pattern> * Database _ structure
</pattern>
<template> A relational database is a set of correlated tables.<br> This is an example of relational structure:<br>
<img src=»database_schema.jpg» width=»100» height=»100» alt=»» border=»0» />
</template>
</category>
</topic>
<aiml>
```

As it is understood, from both figure 3 and the same AIML code presented above, it is possible to manage the multimedia contents (images, videos, etc.) directly inside the code by inserting tag HTML inside the same tag TEMPLATE of AIML.

The multimedia content, inserted in the TEMPLATE, is shown by TutorBot in output alongside the textual content that usually constitutes the answer of the agent to the learners' questions. For the same reason, links might be introduced, being possible to exploit hyper-textual as well as multimedia potentialities. In such a way, the interaction with the tool is highly improved, just as a human tutor would do, TutorBot can draw the user's attention to further informative material, to have more detail on the topics of conversation.

Another characteristic of TutorBot has to be emphasised. If the question raised by the learner cannot be met because a correspondence does not exist inside the knowledge base AIML, the agent, through a search engine expressly implemented for this purpose, ends up questioning external knowledge bases. Such search engine is of an «adaptive» kind and has therefore been denominated ASE (Adaptive Search Engine) (see paragraph 4) (Appratto et al., 2003; Brusilovsky, 2004; Laforcade et al., 2004; McTear, 1993; Mendel et Jerry, 1970); this indexes the resources on the basis of two fundamental criteria:

- key-words present in the document;
- logical context of reference which the topic of the indexed resource and user's profile correspond to (through the WEB platform).

With these two parameters, the resources can be found not only by means of key-words taken out from the learner's questioning sent to TutorBot in natural language, but also by means of the logical context which the learner is referring to during the conversation and strictly linked to his/her profile. The coherence and the actual logical context are determined by the recursiveness of some key-words, that during the conversation, clearly identify the object of the conversation. In the example in figure 4, we can synthesize what happens in background after a conversation of this kind.

In Italian language:

- *Learner: «Cercami database relazionali».*
- *TutorBot: «Per sapere di più sui database puoi verificare su queste risorse che ho cercato per te».*

The same example in English language:

- *Learner: «Search relational database».*
- *TutorBot: «If you want to know more about the database you can check through these resources that I searched for you».*

In other words, in the example, to the user's request to carry out a research, identified by the key-word «search» (along with others included in the knowledge base), TutorBot gets started by explicitly questioning (in this case following the learner's input) the search engine (ASE) which it is connected to, the latter, as illustrated in the following paragraph in figure 5, provides an output related to the search on key-words taken out from the PATTERN of conversation, in our case «relational database» (learner's input). In case the agent finds a lack of information in his AIML knowledge base, he automatically activates the research to external sources, always through ASE (this is clear for the user, see paragraph 4).

The output of TutorBot, if we carry on with the above example and still refer to figure 4, ends up being composed by two elements:

- a section in natural language, where the research which has been carried out, is explained to the learner;
- a section which contains the results of the research carried out by ASE, with the related links that you can click on to the corresponding resources.

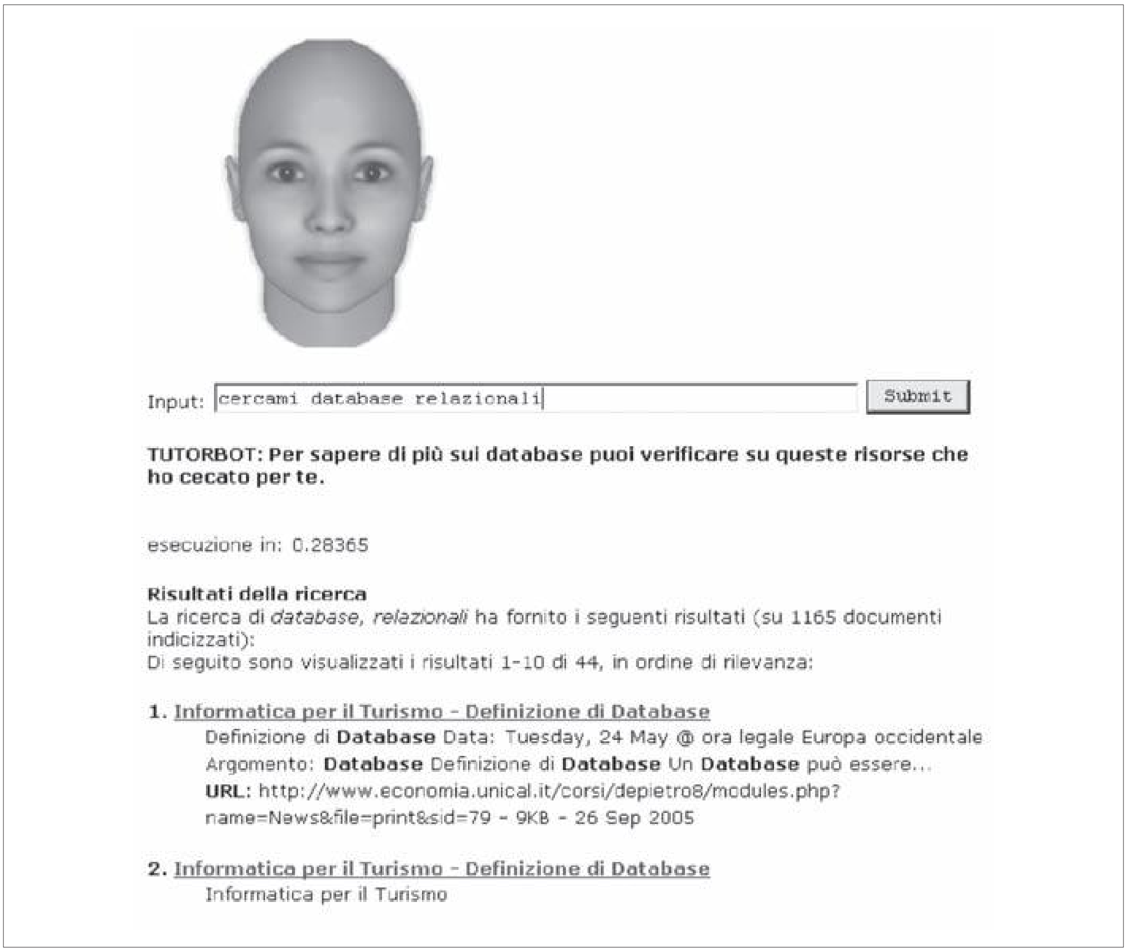

Figure 4 Interface of TutorBot with the integration of the search engine.

Both sections are completely integrated, therefore creating a homogeneous set which is by and large the answer of the agent to the learner's questioning.

4. Description of the Adaptive Search Engine (ASE)

The adaptive search engine (Appratto, De Pietro et al., 2003) is used as a support to the TutorBot's information retrieving. It is used by TutorBot, as we previously pointed out, when the student's query is not directly satisfied from TutorBot AIML knowledge base. In practice, it is TutorBot itself that «questions» the adaptive search engine (ASE) to be able to make up for an information lack in its knowledge base and consequently, to be able to return to the student the results researched in external sources; the student could find information approximately near the asked query. The essential element to have TutorBot work together with the «search engine» is constituted by the AIML tag <search>, these latter send the element template and the reference topic to the ASE that provides for the information retrieving. In this way, TutorBot, using the AIML rules (De Pietro, Frontera, 2005), proceeds to a «lexical reduction» of the student's input, whenever it uses its own natural language, to allow the ASE to carry out the research and consequently to present the possible links to external resources or the ones present in the knowledge base indexed by it. The sending of the query, corresponding to the pattern characterized in the AIML knowledge base, associated to the topic which corresponds to the topic object of the conversation between the student and TutorBot, turns out the results of the ASE research to be the most relating to the question asked by the student.

An example of AIML code that activates the questioning towards ASE is presented as follows (see figure 4):

```
<aiml>
 <category>
      <pattern>* cercami _ Database *</pattern>
      <template> Per sapere di più sui database puoi verificare su queste
                      risorse che ho cercato per te.
                      <br>
      <search query=»http://160.97.33.198/cgi-bin/search/search.pl?»>
                <topicstar/>
                <star/>
      </search>
      </template>
</category>
</aiml>
```

Note 2 – The AIML Knowledge base of TutorBot is in Italian as it is full of informative contents. For this reason it was decided not to provide any translation.

The same example in English language:

```
<aiml>
 <category>
      <pattern>* database _ tries *</pattern>
      <template> For knowing more of the database you can verify on these resources
                      that I have tried for you.
                      <br>
      <search query=»http://160.97.33.198/cgi-bin/search/search.pl?»>
                <topicstar/>
                <star/>
      </search>
      </template>
</category>
</aiml>
```

The interaction process between TutorBot and ASE can be understood observing the scheme proposed in Figure 5.

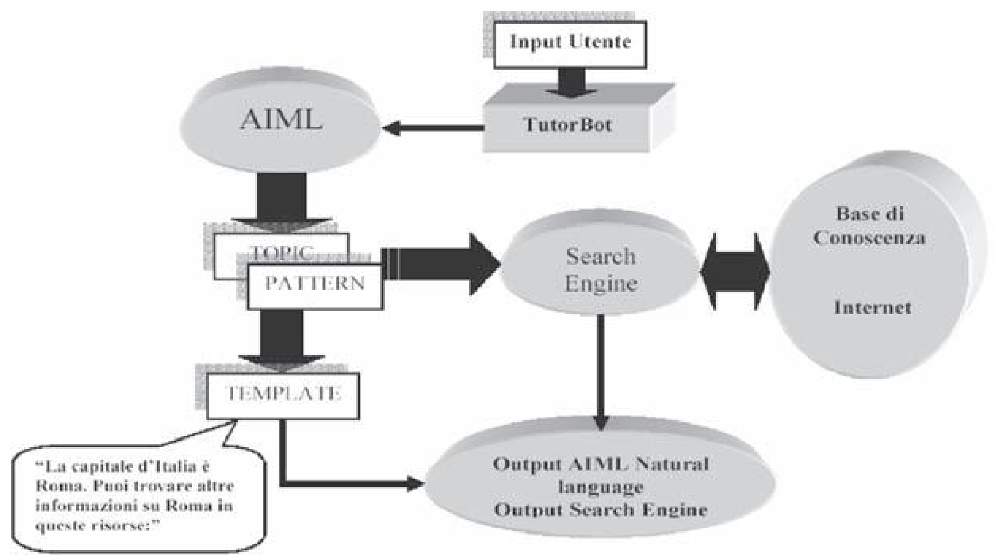

Figure 5 Interaction between TutorBot and ASE.

Moreover, rigorously following the adaptivity technology (Mendel, Jerry, 1970), the ASE is able to characterize the student's «profile» and consequently «adapting itself» to this latter, it only visualizes the results that exactly respect this profile. In fact, to a traditional search engine, such as Google, MSN or others, ASE integrates a particular module that carries out a further selection phase, which follows the traditional one, based on the matching of the research results and the features of the user's profile characterised by the adaptive-engine (adaptive module that deals with the acquisition of the student's information in the interaction phases with the e-learning platform) (De Pietro, Frontera, 2005; McTear, 1993; Sugar, 1995).

Finally, the adaptive search engine is based on two research criteria:

- *explicit:* research according to the keywords and to all the information and operations that the user inserts and carries out in a conscious way when he/she makes a research (in this way, the tool is identical to the traditional research engines);
- *implicit:* research according to the individual features, that is on the single user's profile, that the system through the adaptive-engine registers when every single user of the services surfs through the pages.

The implicit criterion is transparent to the user, however it is successful in capturing the peculiar aspects useful for the selection that not even the user can express, if questioned. The registration of the choices made by the user and the «paths» undertaken through the pages of an adaptive homepage, constitute a useful complement to the explicit results that are inserted as input. This allows to evaluate finally if a certain link, chosen by the engine through the traditional criteria, is included or not among the sphere of the user's usual interests that the research phase has directed through the insertion of a series of keywords.

5. Conclusions

The present work is included among the studies for the optimisation of the learning processes inside an LMS through the use of Intelligent Agents, and in particular wants to give a further tool for the fulfilment of a more effective learning. This tool presents remarkable strengths since, basing itself on the AIML language, it can be easily extended towards other applications, and in the specific context integrated in other LMS, encouraging the development of sharable repository.

Thanks to the standard in object, it is possible that AIML documents produced in other e-learning platforms, can easily be imported and used in its own knowledge base, increasing at the same time the interaction level with the students and knowledge.

Specifically, the expected benefits can thus be synthesized:

- realisation of a tutoring activity on line usable at any time, anywhere and more and more complete;
- reduction of the management costs and the working time from the teachers/tutor, consequently a resource saving that can correspond to a better use for the fulfilment of a major quality of the learning at the global level;
- high level of interoperability between different systems, which is an expected aim from the main international organisations that are interested in e-learning.

One of the main advantages to be stressed, derives from the open source technology used for the implementation of the system, that makes it easily adaptable in whatever web-oriented context, and put the basis for the development of add-on in continuous evolution from the part of the community of developers who follow the same standard for further improvements. Another strength of the system is finally given by the integration of the Adaptive Search Engine (ASE), that allows as a first request to provide for possible lack of information in the existing knowledge base, and at the same time allows high levels of personalisation in the answering phases to the learners, thanks to the used adaptivity techniques that are mainly integrated with the AIML data objects, first amongst all with the object Topic, and consider the user's profile.

Moreover, we want to emphasize that the argumentations and the system developed in the article put, in our opinion, the bases for the implementation of mechanisms that are able to automatize the knowledge management processes according to standards recognised at international level, either to reduce human work or in contexts characterised by a high level of use and a variegated typology of the contents becomes very expensive.

BIBLIOGRAPHY

Ali S.S., Channarukul S. & McRoy S.W. (2001), *Creating natural language output for real-time applications.* Intelligence, Volume 12 Issue 2. June 2001.

Appratto F. & De Pietro O. (2002), *Web-learning: aspects of a new paradigm,* E-Learn 2002, World conference organized by AACE, Proceedings CD Rom, Montreal (Canada).

Appratto F., De Pietro O., De Rose M. & Frontera G. (2003), *Adaptive instruments for w-learning,* Tel03 Proceedings CD-ROM, International Conference, Milano (Italy).

Brusilovsky P. (2004), *Adaptive e-learning systems: KnowledgeTree: a distributed architecture for adaptive e-learning.* Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters.

ChenSession S., Revithis S. & Shi H. (2002), *6C: mobile embodied agents: An agent enabling personalized learning in e-learning environments.* Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 2.

De Pietro O. & Frontera G. (2005), *TutorBot: an application AIML based for Web-Learning* Advanced Technology for Learning, (ISSN# 1710-2251), Ed. ACTA Press, Calgary, Canada, Vol. 2, Issue 1, pp. 29-34.

Descamps S., Prendinger H. & Ishizuka M. (2001), *A Multimodal Presentation Markup Language for Enhanced Affective Presentation,* Advances in Education Technologies: Multimedia, WWW and Distant Education. In Proceedings of the International Conference on Intelligent Multimedia and Distant Learning (ICIMADE-01), Fargo, North Dakota, pp. 9-16, USA.

Ishizuka M., Jatowt A. & Mori K. (2003), *Enhancing conversational flexibility in multimodal interactions with embodied lifelike agent.* International Conference on Intelligent User Interfaces. Proceedings of the 2003 international conference on Intelligent user interfaces. ACM Press, pp. 270-272, NY, USA.

Laforcade P., Marquesuzaà C., Nodenot T. & Sallaberry C. (2004), *Adaptive e-learning systems: Model based engineering of learning situations for adaptive web based educational systems.* Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters.

McTear M.F. (1993), *User modeling for adaptive computer systems: a survey of recent developments.* Artificial Intelligence Review, 7, pp. 157-184.

Mendel & Jerry M. (1970), *Adaptive, learning, and pattern recognition systems; theory and applications*, edited by J.M. Mendel and K.S. Fu, Academic Press.

Sugar W. (1995), *User-centered perspective of information retrieval research and analysis methods.* Annual Review of Information Science, 44, pp. 413-427.

Thomas R. (2001), Contributing Authors: Dr. R.S. Wallace, A. Taylor, J. Baer, *Artificial Intelligence Markup Language (AIML) Reference Manual.* http://www.alicebot.org/documentation/aiml-reference.html.

Wallace R.S. (2001), *Symbolic Deduction in Artificial Intelligence Markup Language (AIML).* Disponibile all'indirizzo: http://www.alicebot.org/documentation/srai.html.