Adaptive E-learning using Genetic Algorithms

Samia Azough[†,††], Mostafa Bellafkih[†] and El Houssine Bouyakhf [††],

†National Institute of posts and Telecommunications Rabat, Morocco
††Faculty of Sciences Rabat Agdal, Morocco

Summary
In this paper, we describe an adaptive system conceived in order to generate pedagogical paths which are adapted to the learner profile and to the current formation pedagogical objective. We have studied the problem as an "Optimization Problem". Using Genetic Algorithms, the system seeks an optimal path starting from the learner profile to the pedagogic objective passing by intermediate courses. To prepare the courses for adaptation, the application creates a descriptive sheet for resources, in XML format, while its integration in the database. Some experiments are added to this paper.

Key words:
Adaptive e-learning, Genetic algorithms, LOM standard, Resources Modelization.

1. Introduction

With the growth of the internet, E-learning has gained a lot of importance in many fields. This can be explained by the multitude of advantages that it offers. In particular, the Web enables the user to access to various resources of formation (image, audio, video, simulations, etc), with a good quality and from various locations. E-learning systems exploit these resources in order to organize flexible formation schedules more adapted to the users' activities and preoccupations.
Nevertheless, the majority of the existing formation platforms are generally conceived as contents distribution systems, with few care about the interests of singular learners. Actually, E-learning concept suffers from many inherent lacks:

- the teacher is not permanently present;
- the teacher is not directly confronted with the learner;
- the management of the immediate reactions of the learner is difficult.

The earlier efforts in this way were tutoring systems. They were concerned with the interaction between the learner and the system. They produced adaptive Intelligent Systems (IS) able to generate didactic material according to the total behavior of the learner. These systems allowed, using Artificial Intelligent Techniques, creating a completely interactive process adapted to the formation. These systems are particularly conceived to meet the need of formalizing or of representing the course topics, the learners and the pedagogic strategies. Among the main suggested adaptations, we find educational system GAITS [12]. In GAITS, supervised teaching techniques are employed where the teacher assigns a pedagogical objective to each student before teaching begins.
In order to achieve this objective, the tutor interacts with the learner using predefined dialogues. Dialogues are composed of a Header and a Body. The Header contains prerequisite knowledge that must be mastered by the student to consider this dialogue as a candidate, taught knowledge that comprises the actual lesson, a learning level, and a learning strategy to present the lesson. The Body represents the teaching materials that define the type of the interaction (e.g., exposition, question/answering, game, etc.)[11].
Essential policies should be adopted in order to conceive learner centric systems taking in consideration its own needs. Asking the trainer to conceive course according to each leaner profile and to each pedagogic aim and demanding permanent managing of learner reactions are possible solutions to achieve the adaptation, but they are very expensive in terms of time and cost. Instead, the following parameters should be improved to build an adaptive system of e-formation:

- profitability to produce more significant results in terms of quality and costs;
- reusability to reuse the already existing pedagogic resources;
- flexibility easy to implement at several various platforms;
- adaptability to adapt for several profiles of learners;
- interactivity to manage the learner reactions during the formation.

In our work, we have conceived an adaptive educational system based on the modelization of the description of pedagogic resources. This system will be able to provide the path the most adapted to the learner profile by using algorithms of optimization. For this reason, we propose to model the description of pedagogic resources in a descriptive sheet under XML format. This model will

Manuscript received July 5, 2010
Manuscript revised July 20, 2010

facilitate the search and the proposal for resources adapted to the learner profile in progress and the pedagogical goal of the formation. Providing an E-Learning platform with adaptive utilities needs thorough reflection. The learner must be guided when consulting his courses, however he should be enabled to better understand his steps of formation according to his capacities and thus to be able to auto-evaluate. Hence, the adaptive e-learning will be able to help the learner to be more autonomous, to have a better comprehension of the course and also to better apprehend and manage his learning process.
The article is structured as follows: In section 2, we present the architecture used for designing the system to be implemented. In section 3, we describe the modelization module and we define multimedia Learning Objects (LOs) as well as LOs metadata. The section 4 relates to the adaptation module, it defines Genetic algorithms and adapter functionality. The section 5 present the scenario implemented in modelization and adaptation modules. Finally, we discuss results in section 6 and we conclude in the last section.

2. System Architecture

2.1 Approach presentation

The current proposed adaptive systems require rebuilding the courses according to a specific format defined by the researcher. In instance, they use adaptive hypermedia [8].
In fact the trainer must construct new courses and he can not reuse earlier developed courses. Additionally, the system does not permit the trainer to choose the form according to his experience. To satisfy such need, we propose to model the description of the resources, whatever are its formats, basing on the available knowledge.
On one hand, the learner must own a precise intellectual baggage to be able to understand a definite course. This knowledge set, which is called pre-requisite concepts, is commonly defined by either the author of the pedagogic resource or an expert. On the other hand, each course is associated to a definite knowledge set that should be acquired after the course (pedagogic goal of the course); we call them the post-acquired concepts (post-concepts).
Thus, the approach used to describe the course is based on the study of the state envisaged of a virtual learner before and after taking this course. To model the pedagogical resources, we have built for each resource a describing sheet based on the pre-requisite and acquired concepts intervening in the latter.
From a semantic point of view, several concepts intervene in the contents of the pedagogical resources, either as preliminary knowledge (pre-requisite), or as knowledge to be acquired (post-acquired). The pre-requisite ones and post-acquired of each course will be presented in the sheet. A good description of the contents makes possible to present to the learner the course most adapted to his profile.
To discern concepts intervening in each pedagogic resource, we class them into modules. The resources belonging to the same module deal with various knowledge (concepts) according to the pedagogic objective of the formation.
In order to define the module in the database, the expert registers the various concepts approached by the resources belonging to the same module of formation as well as the relations between them. The system forms the concepts tree and records it in the database as a characteristic of each module of formation.
The choice of the description based on the concepts enables us to remain independent of any pedagogic approach and any resource format. The modelization of the resources description [2] in XML format will allow their reuse. It allows also the automatic design of various pedagogic courses and thus the selection of the course adapted to the learner profile [1]. The research of the adapted courses is transformed into the optimization problem of researching the most optimal path from a starting point (learner profile) to a final point (pedagogic objective) while passing by intermediate points (courses). The result will be in a form of courses list which has to be attained.

2.2 General Architecture

In this part, we present various diagrams of application design for an adaptive e-learning. The objective is to conceive a system which can model the description of pedagogic resources and guide the learner in his formation according to his assets and to the pedagogic objective that is defined by the trainer. This pedagogic objective presents the capacities that the learner must have acquired at the end of the formation activity.

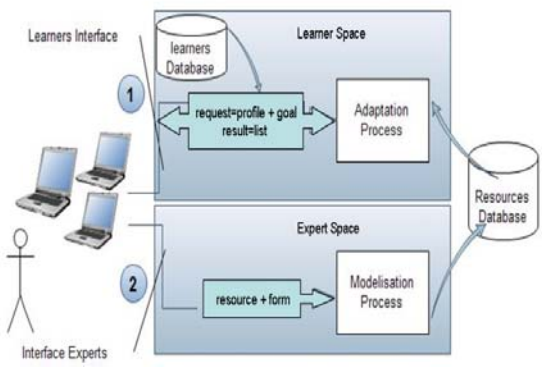

Fig. 1. Structure of the system

Our application is made up of two principal modules.
Learner Space (1) (fig.1) accommodates the identifiers of a learner, selects his profile from the Learners database and returns it to the adapter as well as the goal of this formation.
The adapter (adaptation process) uses optimization algorithms to seek the optimal strategy while selecting the courses in the resources base and provides them to the user interface. The Learner database contains the identifiers of the learner and his knowledge or asset. As a result, the system provides an optimal courses list to achieve the current goal by applying the genetic algorithms to seek the intermediate states.
The part (2) related to the modelization of pedagogic resources prepare them to be used by the adapter. In Expert Space, the teacher or the expert, who seeks to integrate new resources in the base, describes them by filling a form.
Registered information will be recorded in XML sheet by the modelisator (modelization process). The sheet, recorded in the database, facilitates the reuse of these resources. The pedagogic resources base contains the courses that have to be attained. These courses are associated with a descriptive sheet containing preliminary knowledge called pre-requisite concepts and acquired knowledge called post-acquired concepts.
This database is composed of modules; each module is composed of several courses.

3. Adaptation

3.1 Genetic Algorithms: Definitions and Process

The genetic algorithms are algorithms of optimization based on the theory of the evolution of the species of Charles Darwin [4]. The first work of John Holland goes back to the years 1960 and found a first result in 1975 with the publication of ’Adaptation in Natural and Artificial Systems’ [6]. It is however the work of David Goldberg who largely contributed to develop the genetic algorithms [5]. The purpose of a genetic algorithm is to seek an extreme of a function defined on a space of data. The algorithm rests on the following points:

- A coding principle of the population element; the quality of the data coding conditions the success of the optimum research. Two forms of coding are found: the binary coding and real coding.

- A mechanism of generation of the non-homogeneous initial population. The choice of the population conditions the speed of convergence towards the optimum.

- A function to be optimized which turns over a function of adaptation (fitness) of the individual.

- A mechanism of selection of the individuals allowing statistically to identify the best individuals of a population and to eliminate worse.

- Operators allowing to diversify the population during generations and to explore the space of the other possible solutions. The operator of crossing (crossing-over) recomposes genes of individuals existing in the population. The purpose of the operator of change is to guarantee the exploitation of the space of solutions.

Parameters of dimensioning such as cut population, criteria of stop, probability of application of the genetic operators.

3.2 Structure and Operation of the adapter

Sheets built during the modelization step and integrated in the database are used to produce the path most adapted to a learner profile. [1]

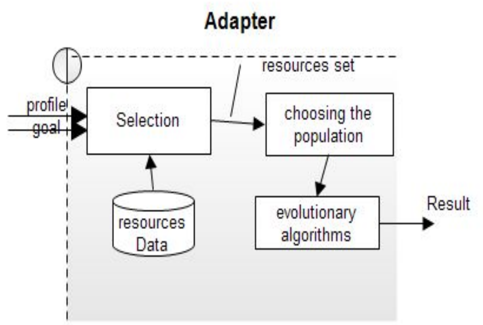

Fig. 2. Structure of the adapter

The adapter (fig.2) selects from the resources base the courses which can be intermediate states by comparing their sheets with the student profile and the pedagogic goal. We transformed the search for adapted course into a problem of optimization using the evolutionary algorithms. The goal, the learner profile, the pre-requisite and post-acquired concepts are seen at the entry of the adapter under the format of vectors.
Goal (1 1 1 0 1 1) means that after following this formation, the learner must have acquired the concepts (1, 2, 3, 5 and 6).
Profile (1 1 0 0 0 0) means that the learner has already acquired concepts 1 and 2.
For the courses, the pre-concepts enable us to check the conditions to reach the course in question. The post-concepts enable us to determine the probable state of the profile of virtual learner after the follow-up of the course in question.

In the genetic algorithm, coding used is binary. The fitness function is calculated according to adaptation of learner for the courses under various formats. The probability of crossing used is equal, and can be corrected by observing the actions of learner and the evolution from its profile. It means in our application the contribution of the capacity and the intelligence of learner in acquisition of concepts.

3.3 Adaptation Process and Implementation

To integrate the genetic algorithms into the centre of our system, we used package JGAP. JGAP is a Genetic Algorithms and Genetic Programming Component provided as a Java framework. It provides basic genetic mechanisms that can be easily used to apply evolutionary principles to problem solutions.
The idea is to transform the problem of searching for the adapted path into an optimization problem. The starting point is the learner profile, the arrival point is the teaching goal of the formation and the intermediate states are the evolutions of the profile after taking the available courses.
To implement the genetic algorithm we followed five stages:

- To predict our chromosome.
- To implement a function of fitness.
- To install an object of configuration.
- To create a population of the potential solutions.
- To evolve population!

We envisage evolving of individuals (courses) according to the genetic operators to build a final individual solution of our problem.

- **Chromosome Course**: The chromosome course is presented in the form of Boolean genes, its size is the number of concepts on which the formation module is based. The values of genes present the envisaged profile of a virtual learner attaining the course in progress.
- **Fitness Function:** The function of fitness is implemented according to the learner profile and to the pedagogic goal of the formation. The individual maximizes the function in question.
- **Genetic Operator**: Starting from two individuals courses, the operator produce an individual course (solution) result of union of the two courses in question. To follow the trace of the operator, we add a gene to our chromosome in the form of a character string (name of the generated individual) formed of the concatenation of the names of the courses submitted to the operator.
- **Object of configuration**: We create a configuration object with our fitness function, we initialize the chromosome and we choose the size of the population to evolve.
    1) **Population:** The initial population is formed by selecting from the database the courses which can be used to reach the goal.
    2) **Evolution of the population** To evolve the population, we apply the designed genetic operator and we post the solution which maximizes the fitness function.

4. Modelization

4.1 LOM (Learning Object Metadata)

The purpose of this multi-part Standard is to facilitate search, evaluation, acquisition, and reuse of learning objects, for instance by learners, instructors or automated software processes. This multi-part Standard also facilitates the sharing and exchange of learning objects, by enabling the development of catalogs and inventories while taking into account the diversity of cultural and lingual contexts in which the learning objects and their metadata are reused. LOM (Learning Object Metadata) data element is a data element for which the name, explanation, size, ordering, value space, and datatype are defined in this Standard [7]. Data elements describe a learning object and are grouped into categories. The LOMv1.0 Base Schema consists of nine such categories:

1. The **General** category groups the general information that describes the learning object as a whole.
2. The **Lifecycle** category groups the features related to the history and current state of this learning object and those who have affected this learning object during its evolution.
3. The **Meta-Metadata** category groups information about the metadata instance itself (rather than the learning object that the metadata instance describes).
4. The **Technical** category groups the technical requirements and technical characteristics of the learning object.
5. The **Educational** category groups the educational and pedagogic characteristics of the learning object.

6. The **Rights** category groups the intellectual property rights and conditions of use for the learning object.

7. The **Relation** category groups features that define the relationship between the learning object and other related learning objects.

8. The **Annotation** category provides comments on the educational use of the learning object and provides information on when and by whom the comments were created.

9. The **Classification** category describes this learning object in relation to a particular classification system.

4.2 Resources Classification by Modules

The current e-learning rests on a model of static modules diffused via the Web. The course followed by learner is fixed once for all by the teacher at the creation phase of the module and, that, without caring of the preferences and results obtained. To allow an automatic adaptation of the module, we followed the process below (fig.3).

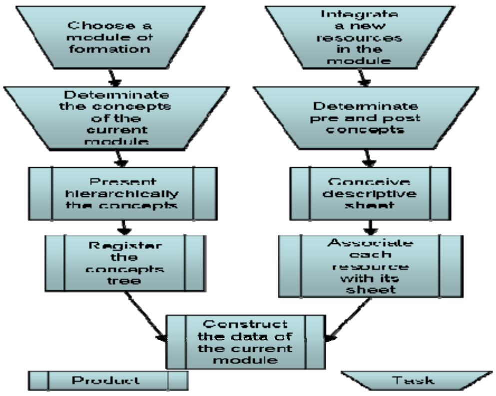

Fig. 3. Development process of the module

The cited products (fig.3) are implemented in the application and the tasks are done by the expert (teacher). To define a new module, concepts, which will intervene in resources belonging to the module in progress, are mentioned through a form.
Hence, the expert completes the form by the relations between concepts. The stage of determination of the concepts is carried out before the integration stage of new resources in the base.
Out of the integration of the resources, the expert selects among the concepts of the current module those intervening in the resource in progress as pre-requisite or post-acquired concepts. The system then produces the sheet based on the data provided by the user and records the result in the base.
The content scenarisation makes sense of its structure and determinates the arrangement of concepts approached by the learner in his formation process. The resources format or the pedagogy used while creating courses are not fixed. The application let the teacher to choose the convenient ones according to his experience. So that, he can add to the database any pedagogic resources he thinks will be useful to the student.
The integration of new resources in the database is done by filling the form presented by the Integration interface of new resources. The expert chooses the correspondent module.
If the module to be chosen was already designed by an expert college, a consultation of the module would be necessary to remain in analogy with the definition of the concepts presented and to be able to add others to them if it is necessary. The scenarisation content is done following the figure below (fig.4).

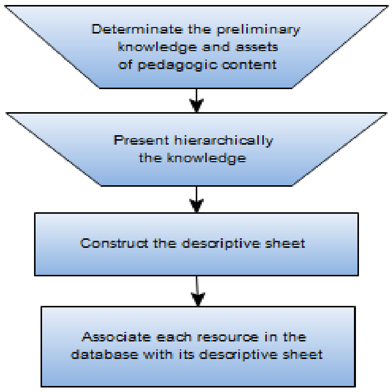

Fig. 4. Integration of the pedagogic resources

When we determinate concepts intervening in the pedagogic resource either as pre-requisite or assets, we can easily fill the form. After the data acquisition, they are provided to the system to build the sheet and to record it in the database.

4.3 Modelization Process

In order to implement such system and to test it, we choose a profile for applying the LOM standard. To design the sheet, we use XML language and the LOM standard which is most widespread in the description of the resources.
The choice of XML to present the metadata indicates the intention of opening and facilitating without predicting technical choices of implementation to be used.

The characteristics used to design the descriptive sheet of the resource are:

–identifier (LOM 1.1),

–title, (LOM 1.2)

–name(s) of the author(s),

–dates of realization and last modification, (LOM 8.1 and 8.2)

–key words, (LOM 1.5)

–description (LOM 1.4)

–format (LOM 4.1)

–size of the resource, (LOM 4.2)

–localization of resource (URL). (LOM 4.3)

–language (LOM 1.3)

–pedagogic type of resource (LOM 5.2)

–pre-concepts (LOM 9.4)

–post-concepts (LOM 9.1)

The sheet is designed according to the following format (fig.5).

```
<Sheet><name>Structures</name>
<description>Etudes des classes et interfaces du package UTIL</description>
<Keywords>liste, collection, calendrier</Keywords>
<format>pdf files</format>
<path>d:\java\str1.pdf</path>
<subject>Programmation Langage Java</subject>
<Concept num="0" name="syntaxe"  Value="1" Type="Is Prerequisite" >
<Concept num="1" name="types"  Value="1" Type="Is Prerequisite" >
<Concept num="8" name="matrices"  Value="1" Type="Is Prerequisite" >
<Concept num="13" name="package util"  Value="1" Type="Is Acquired" >
</Sheet>
```

Fig. 5. An example of created sheets

The sheet is designed automatically starting from the data provided by the expert. The following parameters mean:

- num: identifier of the concept;
- Name: name of the concept;
- Value: relevance of the concept (level of boarding of the concept: beginning/ intermediate /Advanced);
- Type: pre-requisite or post-acquired concept.

The data recorded in the card can be used to index resources and to facilitate the research and the automatic construction of the relations between them.

The sheet makes possible to describe several types of resources and thus to compare and evaluate the various formats and their impacts on the learner profile. To focus on certain parameters allows to reduce the size of the data at the treatment phase and consequently to minimize the execution time of the application.

In the article [1], the sheet is considered as available for each resource of the base. The aim was to be able to propose the most adapted path for the learner profile. The sheet is henceforth designed automatically during the integration of new resources in the base. The modelization part of our application makes possible to design the sheet and to organize the database.

In this application two principal actors intervene:

- The Learner selects a module. His knowledge is then evaluated to update his profile. He will choose a pedagogic objective among those predefined by the expert. The research of the adapted path is made thereafter and the result is in the form of a courses list to follow. These courses are recorded in a file called virtual portfolio to allow the consultation during each connection to the application. During the formation, the learner can test the evolution of his capacities thanks to the tests attached to the attained courses.

- The expert defines the modules of formation by filling a form. He integrates the pedagogic resources into the database. He attaches to each resource a kind of marked exercise (tests) to allow the evaluation of learner knowledge. Also, he defines pedagogic objectives to guide the learner during his formation.

5. Experimentation Remarks

5.1 Profile Modelization

Student level is considered in experimentation tests as predefined data. To automatise the action of evaluating knowledge students we propose to use the following process. To get informed about the learner level, we propose for a learner the tests to pass. These tests are in fact added by the expert and attached to the resources to allow the follow-up of the formation and the update of the profile.

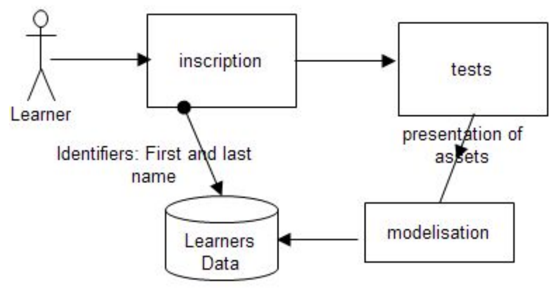

Fig. 6. Stages followed for the modeling of the profile

The teacher proposes questionnaires allowing to evaluate the learner knowledge and thus to direct his formation. Out of the inscription, the learner selects a module, answers the associated questionnaires and passes at the stage of research of the adapted path.

5.2 Adaptation Result

Let us consider the sheet presented in fig.5. The block which defines the concepts is used by the adapter for the research of the adapted course. It is transformed in binary format. For the course presented above (fig.5):

- (1100000010000): pre-concepts or preliminary state of the profile of a virtual learner who will be able to follow this course.
- (1100000010001): post-concepts or preliminary state of the profile of a virtual learner who has followed this course.

These two vectors will be supplemented by using the tree of concepts of the module in question.
The use of such sheet makes possible to consider its integration in several interfaces and the extension towards other characteristics (or attributes) of the resources. It can also be used to describe the contents of different form resources either than the courses (exercises, examinations, articles bound, software,) by adding other attributes. The design of the sheet in format XML makes possible to consider a regeneration of a meta-sheet with other attributes making possible to study quality of pedagogic resources. The attribute 'Value', used to indicate the relevance of the concept in the course, can be equal to a percentage by adding to the application a fuzzificator (fuzzy logic) to define the rules.
Through the Student Interface, the learner can choose an objective for seeking adapted path. He will consult his virtual portfolio and write messages. In order to evaluate his assets, he can access to tests attached to resources contained in his portfolio.

By choosing an objective, the system calls the adapter for using the genetic algorithms. The obtained result is a list of courses to attain. These courses are copied in the portfolio, and the linked exercises (tests) are put in the auto-evaluate part to allow the learner to evaluate his capacities.
The adapter return results such as presented in fig. 7. The number of boolean genes in the chromosome is equal to the number of current module based concepts.

```
Fittest Chromosome is: Size:19, Fitness value:50.0,
Alleles:[BooleanGene=true, BooleanGene=true,
BooleanGene=false, BooleanGene=true,
BooleanGene=true, BooleanGene=true,
BooleanGene=false, BooleanGene=true,
BooleanGene=true, BooleanGene=false,
BooleanGene=true, BooleanGene=false,
BooleanGene=false, BooleanGene=false,
BooleanGene=false, BooleanGene=false,
BooleanGene=false, BooleanGene=false,
   StringGene=C6C10C16]
```

Fig. 7. Solution obtained using the adapter

The 'StringGene' (fig.7) indicates the path followed to obtain the fittest chromosome. This string is transformed to a list of courses to take by using indexes.
Until now, the results don't care of the student feedback.
Actually, we try to perform results through fitness function. We are looking for a function which can choose courses according to last feedback of the learner in progress.

5.3 Application

In order to examinate the results of this approach, the expert defined a module through the application interface following the steps below.

- Definition of based concepts of the module in process;
- Definition of courses;
- Defintion of pedagogical objectives of the formation.

All this data are introducing to the database through a form. After defining the module of formation, the learner can begin the adaptation process by choosing an objective from those defined in the database. Then, the adapter will use the genetic algorithm and generate an adaptive path to follow to attain the chosen objective

6. Conclusion

The application was conceived to assistance of free tools in the form of dynamic Web site. That facilitates its use and reduces the finances. The use of the genetic algorithms enabled us to automate the search for courses adapted to the learner.
Thus, we do not solidify the pedagogic course to take.
The fact of modeling the description of the resources makes possible to benefit from the pedagogic experiment of the expert (teacher) in designing the courses. Separation between the contents and its presentation allows the application of generic models and the integration of specific pedagogic models, also the use of XML format makes possible to really consider the generalization of the logical structuring of numerical information.
In prospect, we will define and implement the rules of fuzzy logic to treat the levels of a beginner learner, intermediate or expert in the acquisition of the concepts. We also hope to study the user feedback and the impact on the profile of the learner.

References

[1] Azough S., Bellafkih M. (2007) Reuse of resources for adapted formation to the learner. 3rd International Symposium on Computational Intelligence and Intelligent Informatics ISCIII IEEE Agadir, Morocco, March 28-30.

[2] Azough S., Bellafkih M., Bouyakhf El H. (2008) E-learning adaptatif : modélisation des ressources en XML, CINQUIEME CONFERENCE SUR LES SYSTEMES INTELLIGENTS : THEORIES ET APPLICATIONS (SITA'08), Institut national de Postes et Télécommunications Rabat, Maroc, 5-6 Mai.

[3] Azough S., Bellafkih M., Bouyakhf El H. (2009) E-learning adaptatif : modélisation des ressources et adaptation au profil de l'apprenant, Systèmes Intelligents: Théories et Applications Edités par europia Productions. ISBN:978-2-909285-55-3. pp51-68.

[4] Darwin Charles (1980) L'origine des esp`eces, petite collection maspero, Paris.

[5] Goldberg, David Edward (1989) Genetic algorithms in search, optimization, and machine learning, Addison-Wesley, 412 p.

[6] Holland J.H. (1992) Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence, Cambridge, Mass. : MIT Press.

[7] LTSC-IEEE LOM Working Group (2001) Draft for Learning Object Metadata (LOM 6.1), http://ltsc.ieee.org/wg12/

[8] Michel C., Rouissi S., (2002) Etude de l'organisation et caractérisation de l'information pédagogique pour construire des documents hypermédias adaptatifs diffusés sur le web, 5ème Colloque International sur le Document électronique (CIDE'05).Hammamet, Tunisie. 308p - pp153-167.

[9] Quafafou, M. GAITS: Fuzzy Sets-Based Algorithms for Computing Strategies Using Genetic Algorithms. Proceedings on Fuzzy Logic in Artificial Intelligence FLAI 1993, in The Lecture Notes in Artificial Intelligence, Springer Verlag.

[10] Semet, Y. Lutton, E. Collet, P. Ant colony optimization for E-learning: observing the emergence of pedagogic suggestions. Swarm Intelligence Symposium, SIS '03. Proceedings of the 2003 IEEE. 24-26 April 2003.

Samia Azough born in Casablanca, Morocco. She received the Research Master's in Computer Engineering, Telecoms and Multimedia from Mohammed V University of Rabat in 2003. She is conducting, since 2004, doctoral research at e-learning systems, modelization of pedagogic resources and adapting formation to the learner profile.

Mostafa Bellafkih born in Oujda, Morocco. He received the PhD thesis in Computer Science from the University of Paris 6, France, in June 1994 and Doctorat Es Science in Computer Science (option networks) from the University of Mohammed V in Rabat, Morocco, in May 2001. His research interests include the network management, knowledge management, A.I., Data mining and Database. He is Professor in The National Institute of Posts and Telecommunications (INPT) in Rabat, Morocco since 1995.