ABSTRACT
In recent years, a concentration of effort to design adaptive web application has arisen: generally, each user has different information needs, depending on her/his social role, culture, etc. Especially in the field of web-based learning, it has become progressively clearer that these needs do not only depend on a long-term characterisation of the user, but also on the contingent situation the user lives.

In this paper, a model for short-term content adaptation is proposed, whose aim is to satisfy contingent needs of users by adjusting the information a web-based learning system provides on the basis of a short-term user profile. The model results in the design of an adaptive filter that acts in the interface between the data and the logic layer of the system. The designed filter is provably correct in the model, that is, it is guaranteed to exhibit a coherent short-term adaptive behaviour.

KEY WORDS
Web-based learning, Information filtering, User modelling

1 Introduction

What is the difference between a good waiter and a good web application? Although the question appears bizarre, it leads to some understanding of the difficulties in the design of state-of-the-art web applications. In fact, both a good waiter and a good web application provide a service, and both of them are smart, precise and accurate. A good waiter is friendly and efficient, and a well designed web application has to meet usability and accessibility goals.

But a good waiter understands the customer, and the purpose of this work is to explain how to mimic this understanding of the user in a web application.

In fact, the behaviour of a web application is functional: two different users asking the same question will receive an identical answer. In contrast, a good waiter, asked for the same service by two distinct customers, will answer differently, adapting his behaviour to the customers personalities.

To mimic the waiter behaviour, a web application should adapt its answers to the user needs, trying to discover them by observing the user behaviour.

The personalisation of web application services is not a novelty. In fact, the most interesting approaches have been investigated in the field of web-based learning services, see Section 5 for an overview.

However, they rely on methods of individualisation concealing two fundamental drawbacks: first, the user is forced to register and to provide personal information (like demographic characters, feedbacks, or web and user logs); second, the discovery of user needs requires a huge amount of time, since enough statistical data must be collected to individuate a pattern, or enough feedback from the user must be analysed to classify him.

In the waiter analogy, these methods correspond to his work experience, but they cannot describe an essential ability of the waiter, namely, to guess the user needs by observing his behaviour on a very short amount of time. Therefore, the kind of adaptive behaviour here considered is devoted to codify this short-term understanding of the user needs.

The context where the adaptive behaviour is experimented is the VICE project, see Section 2, an e-learning system with Semantic Web support. In the framework of VICE, our attention is concentrated on the adaptive filter, composed of a profile database and a pair of algorithms, the analyse and the transformer.

The design of the filter is based upon a strict analysis of the mathematical aspects leading to the derivation of its main algorithms, that are certified by construction, i.e., they are guaranteed to exhibit the intended adaptive behaviour.

Henceforth, the main contribution of this work is the definition of the mathematical model of the filter, that ensures a correct short-term behaviour, as detailed in Section 3.
2 The VICE project

The VICE (Virtual Continuing Education) project [1] aims to develop an innovative methodology and platform to build high-level e-learning applications with a strong technological support.

2.1 Goals

The VICE project is strictly based on the Semantic Web technology [5] and it uses Artificial Intelligence techniques to cope with the adaptive aspects of Learning Objects (LOs) fruition. In this perspective, the VICE project wants to develop an innovative platform that allows the creation, storage and access to LOs, enriched with suitable metadata along with a formal semantics that gives them meaning. Various intelligent agents use this meaning, e.g., to support semantic navigation of the repository content, to automatically construct lessons, and to verify that they are composed of LOs satisfying some pedagogical criteria. In this way, the VICE platform supports the creation of courses and lessons tailored to the cognitive needs of learners. Moreover, paying a special attention to reusability of learning components, the repository is founded on a powerful multimedia database specifically developed to represent and to manipulate LOs represented in the XML technologies and following the SCORM standard.

Different modalities to access the repository are available: (1) through the content, the user may search LOs matching some criterion on metadata, or even on the data content, e.g., text in slide notes; (2) querying the LOs - since they are semantically related by means of their metadata, e.g., prerequisites or level of difficulty - the user can browse the repository following these semantic links; (3) through pedagogical criteria, the user may ask for courses, i.e., sequences of LOs, satisfying some pedagogical goal that covers not only the general topic, but also the preferred pedagogical style.

Moreover, the access to the repository is adaptive: whenever the user poses a query or follows a link, the system analyses this request to individuate a profile. Then the system transforms the retrieved answer according to the inferred interests and preferences.

2.2 The pedagogical wizard

In the VICE project, the pedagogical wizard is the component that realises the pedagogical querying activity. It forces the LOs representation to support pedagogical and semantic metadata that are interpreted according to the wizard knowledge, organised as a coordinated set of ontologies written in the OWL language. The pedagogical wizard interprets LOs as the necessary information to support a didactic process. Hence, by means of a pedagogical taxonomy, it can classify LOs according to their style, to their level of difficulty and to their relative importance in a specific learning process.

Therefore, the pedagogical wizard benefits of this information to compose LOs in lessons and courses, with the warranty that the composition is uniform and balanced with respect to the pedagogical criteria specified by the user. An intelligent agent, that acts by planning the lesson or course with Artificial Intelligence techniques, performs this composition activity.

A side effect of the knowledge available in the ontologies involved in the planning activity is that the repository can be inspected, i.e., queried or browsed, following the semantic categories (the ontological relations) that describe the knowledge. Thus, the ontological representation provides the support for the semantic navigation feature. In addition, the knowledge developed to support planning helps the generation of the adaptive behaviour that is the focus of this paper.

The represented knowledge, although coded in the form of OWL ontologies, wants to embody the common knowledge an expert may have of the e-learning process in general, and of the specific instructional domains the LOs in the repository cover. It is evident that this sort of knowledge is mostly qualitative, sometimes incomplete, and often imprecise or approximated. Therefore, the ontologies have been developed in a reasoning framework that is intrinsically fuzzy.

2.3 The adaptive filter

In order to construct automatic agents that search and compose LOs in an effective and coherent way, it is necessary to have an integrated ontological system that refers to all elements constituting the application domain. Specifically, the needed ontologies cover (i) the contents, to correctly extract and navigate the objects in the repository, (ii) the users, to personalise the instructional process, (iii) the learning goals, (iv) the taxonomies of the LOs and (v) the pedagogical criteria codifying the teaching styles. The adaptive behaviour is based on this knowledge, and specifically it relies on the ontologies to provide the set of possible interests that it aims to infer. Moreover, the ontologies are used to interpret queries in the mathematical model, as detailed in Section 3.

The adaptive agent is a filter that intercepts queries to the repository and uses them to construct a user profile. Then, it transforms the answer of the repository according to the adaptive strategy. The profile is a model of the querying user, which is devoted to code his interests and characters. While interests are intended to be short-term goals the user wants to fulfil, characters are considered as long-term information describing the user preferences and capabilities.

In this paper, we will focus on the short-term behaviour that constitutes the real novelty in our adaptive approach. In fact, the long-term adaptive model is a natural extension of the short-term model, and is comparable with many existing approaches, see Section 5.

The first goal the adaptive filter wants to meet is the
convergence of the transformation process. Precisely, in the hypothesis that a sequence of queries represents an attempt to satisfy an interest of the user, we expect that every query in the sequence refines the interest, until a satisfying result is found. Hence, the transformation process should result in a series of answers that are closer and closer to the satisfaction of the user interest, and this series of answers should converge with the same speed as the sequence of queries.

Therefore, the filter must be able to segment the flow of queries from a user in sequences each of one converging to the full expression of a single interest, until the interest gets satisfied. In other words, the filter should sense the changes of interest in the flow of queries issued by a user.

The filter provides families of transformations, such as ordering the elements of a multiple answer according to their closeness to the user interest. These transformations are parametrized with the user profile. The property the filter is called to preserve is informally stated as "to a small variation in the profile, a small variation in applied transformation corresponds", hence the need for a mathematical model. In fact, the model is more precise and what it does really preserve is the convergence speed.

3 Mathematical aspects

The adaptive filter is based upon two algorithms: the analyser, extracting information from the query and updating the user profile, and the transformer, choosing a transformation, parametrised by the user profile, to apply to the result.

The goal of the adaptive strategy performed by the filter is to exhibit two behaviours: convergence and changes of interest. Convergence means that if the sequence of queries approximates the user interest, then the series of results must be closer and closer to the satisfaction of that interest. On the other hand, the notion of change of interest models the situation when the user stops a sequence of queries and starts a new one, which focuses on a different interest.

3.1 The user space model

The user model is composed of two parts: the former contains the information about the user characters, while the latter models his interest. Therefore, the former part is persistent, that is, not affected by a change of interest, while the latter is transient. Thus, the user is modelled as a point in the profile space, denoted as $\mathcal{U}_p$, that is defined as the Cartesian product of two spaces $\mathcal{U}_C$ and $\mathcal{U}_I$. Both $\mathcal{U}_C$ and $\mathcal{U}_I$ are $n$-dimensional cubes of edge $[-1, 1]$.

Every coordinate of these spaces is uniquely associated with an attribute. Thus, the user profile, a point in the profile space, is a set of pairs of the form $(a, w)$, where $w$ is the weight of the attribute $a$, denoting how strong is the user interest in the attribute: $-1$ denotes an absolutely strong interest, $-1$ denotes a complete negative interest, and $0$ denotes a lack of interest. The attributes correspond to the ontological classes used by the pedagogical wizard to reason about the user properties, distinguished in interests and characters.

The profile space is also a metric space, being isomorphic to a subspace of an Euclidean space. Thus the profile space is equipped with the usual Euclidean norm and the associated distance function is $d(a, b) = ||a - b||$, as usual, see, e.g., [10].

Moreover, since the profile space is the Cartesian product of $\mathcal{U}_C$ and $\mathcal{U}_I$, for every $p \in \mathcal{U}_p$ there is a vector $p^0 = (p', 0) \in \mathcal{U}_p$ such that $p = (p', p'')$. The $p^0$ vector represents the neutral profile, that models the user deprived from his immediate interests.

The important point in our model is that a query can be interpreted as a profile. In fact, the query asks for LOs that do posses or do not posses some metadata. These metadata are used to mark the nodes of the semantical network, derived from the ontologies of the pedagogical wizard, with a weight: $1$ if the metadata occurs positively in the query, $-1$ if it occurs negatively, $0$ otherwise. These ontologies, being fuzzified, have a weight for every relation among classes. By means of standard algorithms, see [3], the initial marking of the net is extended through its arcs until the marking of nodes becomes stable. The query profile is then composed of the weights of the nodes associated to their corresponding classes (attributes).

Therefore, given a sequence of queries $q_1, \ldots, q_n$, it is possible to associate to every query a corresponding profile, obtaining a sequence of query profiles $\hat{q}_1, \ldots, \hat{q}_n$. Thus, given an initial user profile $p_0$, the sequence of queries $q_1, \ldots, q_n$ generates the sequence of user profiles $p_0, \ldots, p_n$ with the rule that $p_{i+1} = \alpha p_i + \beta \hat{q}_i$, that is, the weighted mean of the current profile and the query profile.

Given a sequence of queries and an initial user profile, one should consider when the generated sequence of user profiles terminates. In fact, since the user may change his interests after issuing a few queries, the profiling activity of the filter should be restarted. As previously said, this restart is called a change of interests.

The mathematical model captures a change of interest by observing, step by step, the sequence of generated profiles. When a change of interest occurs, the filter restarts its action by posing the current profile to the neutral profile $p^0$. The rule that identifies a change of interest is a lack of convergence in the sequence of generated profiles. This is achieved by measuring the convergence speed of the generated profile sequence, that is, to observe if $||p_{i+1} - p_i|| > \delta / f(i)$, where $\delta$ is a constant threshold value, and $f$ is a monotone increasing function. For example, if $f(i) = i$, then a sequence that does not violate the constraint is converging with linear speed.
3.2 The transformation space model

The transformer algorithm is designed according to the principle that a small variation in the user profile induces a small variation in the applied transformation. Thus, we assume that the filter is equipped with a set of transformations, and one of them is applied to the result according to the query and the user profile.

Usually, the transformation applied to the result depends on the query. To capture this dependency, the set $Q$ of possible queries is partitioned by an equivalence relation $\rho$. Hence, the transformation space is the disjoint union of transformations: $\Theta = \bigcup_{x \in Q/\rho} \{ \theta_x(p) \mid p \in U_P \}$.

The functions $\theta_x: U_P \to (U_A \to U_A)$ are the transformations: they act by generating a concrete function from an user profile, that, when applied to a result, produces the transformed result. The result is considered to be a point in the space $U_A$, the set of possible answers, identified with the set of valid XML documents.

A transformation $\theta_x$ is said to be admissible if it is associated with a distance function $d_x$ on the answer space that ensures that $U_A$ forms a Banach space [10] and that satisfies the following constraint: let $a \in U_A$ be any answer from the repository, then, for every $p, q \in U_P$,

$$d_x(\theta_x(p)(a), \theta_x(q)(a)) \leq K d(p, q),$$  \hspace{1cm} (1)

where $d$ is the distance function in the profile space and $K$ is a (small) constant, possibly depending on $a$ and $x$.

If the transformation space contains only admissible transformations it is said to be adaptive. It is possible to prove that a distance function $d_x$, respecting the constraint (1), forces an homeomorphism between the user space and the answer space via an adaptive transformation space. Moreover, one can show that the convergence speed is preserved by the homeomorphism. The proof is based on the observation that the user space is a Banach space.

As an important example of admissible transformations, we consider orderings. An ordering transformation operates by sorting the elements of an answer in the form of a list, according to their closeness to the user profile. Precisely, every element $r_i$ of the list $(r_1, \ldots, r_n)$ is associated with a point $\tilde{r}_i$ in the profile space, as described for query profiles. The sorting key is calculated by means of a continuous function $f: U_P \times U_P \to \mathbb{R}$ applied to the pair $(\tilde{r}_i, p)$, $p$ being the user profile.

The ordering transformation satisfies the constraint (1), since, posing $d_x(a, b)$ as the minimal number of element swappings needed to transform $a$ in $b$, it follows that $d_x$ is, indeed, a distance function. Since $f$ is continuous, but sorting is a discrete function, a sufficiently small value of $d(p, q)$ implies that $\theta_x(p)(a) = \theta_x(q)(a)$, thus, calling $\epsilon$ such value, it follows that $d_x(\theta_x(p)(a), \theta_x(q)(a)) \leq n/\epsilon$, because, no more than $n$ swaps are needed to transform a list of length $n$ into any of its permutations. Thus, posing $K = n/\epsilon$, the constraint (1) immediately follows.

4 Preliminary experiments

The filter model as presented is an abstraction on the development experience accumulated in the VICE project, at the moment still ongoing; thus, the measures that are discussed in the following are not quantitatively relevant, being performed on test cases in laboratory, and not yet by monitoring the real usage on the field. Despite this fact, our belief is that the presented experimental results possess enough qualitative value, useful to understand the validity of the followed approach, although it cannot be a faithful measure of its effectiveness.

The considered test case has been a small repository with the related ontologies representing the metadata describing the on-line guide of Microsoft Word 2000® where every unit in the table of contents has been considered a LO. The filter was tested to ensure that an adaptive behaviour occurs, since this is a consequence of the adopted mathematical model, but, instead, to see if the user perceive the adaptive behaviour, and if they benefit from that behaviour.

We asked four students, skilled in the basic use of the word processor, but not experts, to write a very complex text, with serious formatting problems.

One of them had access only to the standard on-line help of Word; another one was enabled to use the search engine of our test case, but without the adaptive support, and, finally, the last two students could access our search engine with the adaptive filter.

As a result, the three students accessing the search engine gave a high mark (8.6 on average) to the help system, while the student that was limited to the standard on-line help gave (only) a good mark (7 points in a range 0 to 10). On the other hand, both students that had access to the adaptive version of the search engine gave a 9 as the mark to the help system, thus revealing that both had a benefit over the non-adaptive search engine. This benefit was measured in 1 point.

To recapitulate, the three students accessing the search engine reported that they found very useful the semantical search to locate in the on line guide the useful information to fix their problems; but, what is more peculiar, the two students using the adaptive search engine told us that they appreciated the apparent ability of the filter to follow their needs, although they were more impressed by the search ability than by the adaptive behaviour.

Thus we concluded that the benefit of the filter is sensible, but, probably not so high as marks measured.

The described experience clearly shows that, qualitatively, the adaptive behaviour is perceived by the users, and that it positively affects their feelings about the system. Obviously the limits of our test experience with students, both in the number of users involved, in the complexity of the test case, and in the precision of the questionnaire, prevent any quantitative evaluation.
5 Related works

Since the problems of students and teachers are paradigmatic for the class of adaptive behaviours under examination, the natural framework where to develop an adaptive support to content management is an e-learning environment. Moreover, as described in Section 4, the proposed model has been conceived to support an e-learning system, thus comparison with analogous experiences is due.

In general, although a huge amount of resources is available, often web-based learning systems are not able to support an effective and suitable utilisation of the Net. In fact, the majority of web-based courses address to a generic audience, without any concern of individual students, their peculiarities and specific educational needs. It can be done, instead, by the presence of some sort of sophisticated adaptive procedure, able to modify every single navigation path [13].

So, the Web has been extended [5] in order to turn from a simple information retrieval and communication tool into a smart and semantic space, where software agents are able to understand information and, then, to help users locate the wanted resources and to select the most appropriate presentation for the retrieved results. In this way, the Web would be able to meet user needs by understanding their specific information-seeking behaviour, thus knowing the goal of every query, in order to guide users towards really useful information, now easily accessible in ad hoc manners.

For all these reasons, web-based learning systems are expected to emphasise the freedom of accessing resources, and to enhance the possibility to choose the most appropriate learning or cognitive style [8].

Since the interest of users in different approaches to the educational resources is a fundamental subject matter of adaptivity, several applications have paid much attention to define a set of instructional strategies in order to meet users’ individual learning styles and preferences and, thus, to improve their performance.

Therefore, the mainstay of such applications is to provide students with different presentations of learning resources, as in TANGOW [6], or within an experimental virtual campus at the Open University of Catalunya [11], according to parameters strictly related to their profile, including the actions they perform while attending their tasks, and their individual preferences about learning strategies.

Another interesting solution is based on the complementary pair of authoring systems AHA! and MOT [14], whose main objective is to provide authors with specific tools that help them to supplement their adaptive applications by designing as many assortments and combinations of learning styles as possible (individual learning styles are assessed through a registration form where users can select their favourite instructional strategy).

What is most noteworthy in this project, the users’ browsing behaviour allows the system to improve its performance by inferring their preferences and keeping them up to date from time to time.

The cooperation between IBM and MIT gave rise to a pilot study of a system, whose core is the Dynamic Assembly Engine [9], to automatically assemble LOs in personalised paths, by ordering the learning resources in a logical sequence. Actually, the process flow provides that, once a query has been submitted, the system returns a list of the most relevant results matching the users’ requirements, according to their educational role. The main concern of this work is to find a solution not only to select the most appropriate resources to a specific learning path, but to connect them in a well-organised corpus too, tailored to specific cognitive needs. The framework of the Dynamic Assembly Engine is particularly close to VICE, see Section 4, since a XML federated repository, in which web resources are stored, and a strict correspondence of LOs content and their metadata to user needs work together to make this approach feasible.

The accuracy of the system is guaranteed by the knowledge the application has of users’ queries, and the coherence of the path is achieved by presenting objects that are closely related with one another, meeting the search parameters.

The German prototype ELENA [7], based on the Edutella P2P infrastructure [12], is another valuable research work in the domain of personalised supports for students in an e-learning network, in order to create an adaptive Semantic Web environment. The main contribution of ELENA is the creation of smart spaces for learning, defined as educational service mediators, based on Semantic Web and Artificial Intelligence technologies. These spaces provide students with an exhaustive description of those learning resources that meet their profile, by allowing them to access any kind of repository which is connected to a network.

Therefore, the German team aims at setting up an organised network of learning and knowledge management systems, so that students can access resources, lessons and courses according to their individual profile.

The keystone of this project, the Personal Learning Assistant (PLA), is expected to present a focused list of learning services. The action performed by the PLA, refers to the user dynamic profile, in order to extend queries by adding restrictions, variables, preferences, goals, etc., and to customise search results.

The close likeness between all these works and VICE is due to the fact that the adaptivity the filter provides in our work concerns the capability to take in account all the users’ features (gathered thanks to their navigational background and the actions performed while carrying out their searches), in order to customise educational contents and make them fulfil specific cognitive needs. This way, the filter allows users to orientate themselves in the guided usage of the knowledge stored in the repository, through typical technologies of AI and Semantic Web.

Noticeably, the filter model is conceived to address
the short-term behaviour of users, while the adaptive approaches described so far have been developed to exhibit a long-term understanding. A long-term approach to content adaptation requires a great extent of empirical study and analysis of the navigational behaviour of the users, by keeping track of their actions for a long amount of time (e.g., a whole day [15]), in order to obtain an accurate description of the users and their preferences, whereas our filter is expected to capture as soon as possible (less than a dozen of observations) a good approximation of users interests.

On the contrary, our short-term observation entails an imprecise deduction of information about users in order to understand their changes of interest (that is, the instants between the end of a sequence of queries and the start of a new one, while resources are browsed or sought) and, as inferred by the sequence of queries, to refine the acquired knowledge about users as quickly as possible.

As far as we know, no short-term approach to adaptivity appears in literature.

6 Conclusions

This work discusses the idea of content adaptation in a short-term perspective as a way to discover and to satisfy contingent user needs in the context of web-based learning systems. This result is obtained by means of a filter that operates between the logical layer and the repository. A mathematical model represents the filter action. The model is composed of the user space, representing the information about users, and of the transformation space, representing admissible manipulations the filter may apply when adapting the content. The profile space and the transformation space are related by a constraint that forces convergence of user profile sequences to be reflected in the convergence of the corresponding transformation sequences. This formal property, captures the informal requirement of a correct adaptive behaviour: small variations in the user profile induce small variations in the transformations.

With respect to our previous work on the same topic [4, 2], this article extends the mathematical model to consider characters of user in the middle-term range, and it presents for the first time a preliminary experience in the usage of the proposed techniques. The experience on the adaptive filter has shown that users effectively perceive the adaptive behaviour as a benefit. It is to remark that the displayed data are drawn from a measure on a toy example, and on an extremely exiguous population, because the project as a whole is still under development, therefore no test on the field can be performed in the actual stage. Despite the statistical irrelevance of the measures obtained so far, it is our belief that the suggested approach is of interest. More investigation is needed to quantify the impact of a short-term adaptive strategy in a real-world situation.

References