Virtual Distance Education in a Collaborative Hypermedia Environment – the VIENA Classroom

Werner WINIWARTER, Osami KAGAWA, Shin’ichi KONOMI and Yahiko KAMBAYASHI
Department of Information Science, Kyoto University, Sakyo, Kyoto, 606-01 Japan

Abstract. VIENA Classroom is a virtual distance education system that addresses a wide range of applications from intensive on-line teaching in small groups to wide area teaching via global networking for large communities. In VIENA Classroom the teaching material is designed as hypermedia documents and presented to the students within a collaborative environment. An important feature of the system is its user-friendly question support facility. By means of an adaptive natural language interface the students can formulate their questions directly in Japanese which are either answered automatically from a FAQ knowledge base or gathered for later processing by the teacher.

1. Introduction

The last years have witnessed a rapidly growing interest in “learner-centered” approaches to education which aim at replacing passive lecture attendance and text-book reading by active exploration and construction [1]. This recent renaissance of well-established problem-based learning strategies such as “learning by doing” or “training on the job” is due to the now available advanced computer technologies. Hypermedia and collaborative environments are two such compelling technologies that make it possible to engage students in solving interesting tasks which results in a more active participation and higher motivation.

Another important factor for the fast promotion of the problem-based teaching paradigm is global networking and the internet. Especially the World-Wide Web attracted many researchers to write tools to make education systems available to a large number of users [2, 3, 4].

The VIENA Classroom represents a virtual distance education system with a wide range of target applications from intensive on-line teaching in small groups including group discussions and solving problems by team-work to off-line teaching for large communities such as night schools or wide area teaching via global networking with time differences. In VIENA Classroom the teaching material is designed as hypermedia documents and presented to the students within a collaborative environment. In order to provide the required flexibility as concerns the personalized display of shared information, we apply the deputy object model which allows to create multiple deputy display objects from a certain media object not only by selecting but also by adding and modifying attributes and methods.

An important feature of the system is its user-friendly question support facility which overcomes many limitations of real classroom situations. We provide the students with a natural language interface so that they can formulate their questions directly in Japanese. The computed semantic representations are applied to the retrieval of answers from a FAQ knowledge base or to the gathering of semantically equivalent new questions. FAQ frequency data is used for displaying new questions, as input for improving the teaching material and for the automatic generation of context-sensitive FAQ lists which can be consulted by the students as valuable assistance for formulating questions. In order to guarantee successful portations of the interface to new application domains and the efficient treatment of the constant changes in the dynamic
environment of VIENA Classroom, we make use of methods from machine learning as well as existing linguistic resources so that we result in an adaptive interface architecture.

The rest of this paper is organized as follows. After a short survey of related work we give a brief overview of the features of the VIENA Classroom system, especially of the applied techniques for the personalization of shared information. Then we deal in more detail with the functionality of the question support facility and finally we present the applied techniques for the realization of the adaptive natural language interface architecture.

2. Related Work

The recent development of high speed computer networks and the increasing number of multimedia computers created new opportunities for innovative applications of computer systems including the emergence of Computer-Supported Cooperative Work (CSCW) systems. CSCW systems support groups of people engaged in the solution of a common task by providing an interface to a shared environment [5].

Most of the existing systems are combinations of video systems, communication facilities and basic window sharing facilities. Commercial products such as desktop video conferencing systems have the capacity of distance presentation by utilizing video phones, white boards, telepointers etc., however mostly without any appropriate database support.

Database support for collaborative work and the efficient management of multimedia data are two major issues in designing multiuser hypermedia systems [6] which enable flexible locking and change notification by using databases for sharing. Since traditional database view mechanisms cannot support the required high degree of flexibility for sharing information, new concepts such as proxies [7], MultiView [8] or adapters [9] have been proposed.

Whereas the last years already witnessed a growing number of emerging education systems utilizing hypermedia techniques (e.g. see [10, 11]), only recently research started to focus on the high potential of CSCW systems for providing the necessary framework for collaborative education. Such systems aim at the solution of more interesting and complex problems as well as the promotion of communication skills and the articulation of arguments to other group members. Examples of first successful prototypes are CoVis [12], SpeakEasy [13] or CSILE [14].

3. Features of VIENA Classroom

One of the main requirements to a collaborative environment is its high flexibility as concerns the personalization of shared information. The student may want to perform personalization along the following three dimensions:

- **personalized hypermedia**: adding personal comments and notes to the presented teaching material,
- **personalized display**: changing fonts, colors, emphasizing styles etc.,
- **personalized navigation**: revisiting previously presented material via links or referring to related material in a separate window.

In order to provide a sufficient degree of flexible sharing to deal efficiently with all these requirements, we developed the concept of deputy objects [15] which makes it possible to create multiple deputy display objects from a certain media object not only by selecting but also by adding and modifying attributes and methods.
However, there exists a trade-off between personalization and sharing, i.e. it is sometimes necessary to impose constraints on the personalization of hypermedia material so that a certain degree of common perception is guaranteed, e.g. to focus the attention on a certain position in the teaching material.

Therefore, we introduce *display environments* in which it is possible to declare defaults for display as well as constraints on personalization. Display environments in VIENA Classroom may include all participating students or can be also restricted to small seminar groups engaged in the collaborative solution of a common task.

Besides the development of these personalization mechanisms and the question support facility which is dealt with in Sect. 4, the realization of VIENA Classroom involved also the following list of other innovative features:

- *design aids* for the creation of the hypermedia teaching material,
- techniques for the *symbolic representation* of status and information about students which varies in detail as a function of the number of participants of the lecture,
- methods for monitoring behavior and actions of students to create corresponding individual *student models* as well as their abstraction and condensation to collective models, for other recent work on student modeling see [16, 17],
- the automatic generation of personalized teaching material based on feedback data from the students so that displayed information can be more or less detailed according to the amount and type of previously asked questions, for an overview of these so-called *intellimedia* applications which aim at adapting the presentation of hypermedia material according to the specific needs of the individual user see [18].

The CSCW functionality of VIENA Classroom is implemented under *GroupKit*, a toolkit based on Tcl/Tk (University of Calgary). GroupKit provides message and event handling mechanisms as well as widgets that are particular to real-time groupware applications such as shared telepointers or multiuser scrollbars. Therefore, it represents a powerful framework for the realization of the collaborative features of VIENA Classroom.

For the efficient administration of the hypermedia data we use *Ode* (AT&T Bell Lab), an object-oriented database system with multimedia support, versioning and triggers. Another convenient feature of Ode is OdeFS, a file system interface that makes it possible to manipulate Ode objects by using normal file manipulation commands.

### 4. Question Support Facility

In a conventional classroom setting the teacher is confronted with the following problems as concerns the answering of questions:

- limited number of questions that can be answered,
- considerable delay in answering,
- missing possibility to select relevant questions,
- repeated questions.

These problems are even more severe in the case of a virtual classroom because the number of students can be much larger. Therefore, one of our central issues was to supplement VIENA Classroom with a powerful question support facility, a topic which is not covered by most of the existing systems.
In VIENA Classroom students are free to ask questions at any time which are either answered directly from the FAQ knowledge base or transferred to the teacher for later processing. A question is always formulated in a specific context in that the student highlights the corresponding passage in the teaching material. It is the responsibility of the teacher to define a uniform context model during the design process of the teaching material so that a student can easily choose the corresponding context by double-clicking in the concerned area. However, the student is still free to mark a specific term or phrase, in such a case the question is resolved according to the context.

New questions are collected and grouped for each context on the basis of their semantic representation. For each group of semantically equivalent questions a representative question is computed as input for an informative display to the teacher. Figure 1 shows an example which was translated into English, original questions with context resolution are marked with an asterisk. If such a question is selected, the concerned term or phrase is highlighted in the context window. The answers of the teacher are not restricted to text but can consist of a sequence of animated actions, e.g. pointing, highlighting, presenting additional clarifying material etc.

The teacher can also choose between an off-line answering mode by clicking on the Answer button and an on-line answering mode (Connect button) in that he/she directly enters into discussion with the concerned student via a shared window. Furthermore, by way of the Filter button the display of new questions can be configured as filter that gathers questions until the next answering period which can be defined on the basis of a certain time period, break-points in the teaching material or thresholds on the frequency data for new questions.

Besides the use of frequency data for the ranking of new questions, also the frequency data of already answered questions is a very valuable source of information because it provides important feedback for the improvement of the teaching material. A quick first solution for the teacher is to simply add relevant questions and answers until he/she has more time for designing a completely revised version.

Finally, FAQ frequency data is also used for the automatic generation of context-sensitive ranked FAQ lists which can be consulted by the students before formulating a question in natural language. This is especially useful for situations where the student does not know how to formulate the question and would like to have some guidance. In this context FAQ are marked with an asterisk to indicate that they have not been answered yet by the teacher. The submission of such a question is thus an easy way to increase its priority of answering.

<table>
<thead>
<tr>
<th>CONTEXT</th>
<th>The virtual memory makes it possible to execute non-resident programs on real storage. It realizes a larger space and a larger process than physical storage.</th>
</tr>
</thead>
<tbody>
<tr>
<td>REP. Q.</td>
<td>QUESTION</td>
</tr>
<tr>
<td>What is the difference between process and space?</td>
<td>10</td>
</tr>
<tr>
<td>Which meaning has page table?</td>
<td>5</td>
</tr>
<tr>
<td>What is the advantage of global replacement?</td>
<td>3</td>
</tr>
<tr>
<td>ORIGINAL QUESTIONS</td>
<td>QUESTIONER</td>
</tr>
<tr>
<td>What is the difference between process and space?</td>
<td>Werner Winiwarter</td>
</tr>
<tr>
<td>What is the difference to process? *</td>
<td>Osami Kagawa</td>
</tr>
<tr>
<td>What is the difference between space and process?</td>
<td>Yahiko Kambayashi</td>
</tr>
</tbody>
</table>

| Answer | Connect | Filter | Notes | Quit |

**Fig. 1:** Example of display of new questions

In VIENA Classroom students are free to ask questions at any time which are either answered directly from the FAQ knowledge base or transferred to the teacher for later processing. A question is always formulated in a specific context in that the student highlights the corresponding passage in the teaching material. It is the responsibility of the teacher to define a uniform context model during the design process of the teaching material so that a student can easily choose the corresponding context by double-clicking in the concerned area. However, the student is still free to mark a specific term or phrase, in such a case the question is resolved according to the context.

New questions are collected and grouped for each context on the basis of their semantic representation. For each group of semantically equivalent questions a representative question is computed as input for an informative display to the teacher. Figure 1 shows an example which was translated into English, original questions with context resolution are marked with an asterisk. If such a question is selected, the concerned term or phrase is highlighted in the context window. The answers of the teacher are not restricted to text but can consist of a sequence of animated actions, e.g. pointing, highlighting, presenting additional clarifying material etc.

The teacher can also choose between an off-line answering mode by clicking on the Answer button and an on-line answering mode (Connect button) in that he/she directly enters into discussion with the concerned student via a shared window. Furthermore, by way of the Filter button the display of new questions can be configured as filter that gathers questions until the next answering period which can be defined on the basis of a certain time period, break-points in the teaching material or thresholds on the frequency data for new questions.

Besides the use of frequency data for the ranking of new questions, also the frequency data of already answered questions is a very valuable source of information because it provides important feedback for the improvement of the teaching material. A quick first solution for the teacher is to simply add relevant questions and answers until he/she has more time for designing a completely revised version.

Finally, FAQ frequency data is also used for the automatic generation of context-sensitive ranked FAQ lists which can be consulted by the students before formulating a question in natural language. This is especially useful for situations where the student does not know how to formulate the question and would like to have some guidance. In this context FAQ are marked with an asterisk to indicate that they have not been answered yet by the teacher. The submission of such a question is thus an easy way to increase its priority of answering.
5. Natural Language Interface

Although there have been many ambitious attempts of developing natural language interfaces, the achieved results were so far rather disappointing (for a recent survey see [19]). The two main reasons for this are missing customization, resulting in unexpected restrictions, and missing integration, responsible for insufficient performance and wrong interpretations. As concerns the former we stress the importance of empirically collecting the training and test set so that representative data is obtained which covers all relevant linguistic phenomena. With regard to missing integration we introduce the Deductive Object-oriented Approach (DOA) in which the natural language interface is designed as component of the deductive object-oriented database system ROCK & ROLL (Heriot-Watt University) by making use of the available powerful logic and object-oriented programming language. Especially the object-oriented inheritance mechanisms make it possible to structure the linguistic knowledge hierarchically which enables the flexible insertion of features at the appropriate level of abstraction.

Figure 2 shows the process model of our natural language interface architecture. The basic linguistic analysis consists of three main steps. Morphological and lexical analysis perform the tokenization of Japanese input, i.e. the segmentation of the individual words. By accessing a
domain-independent lexicon the input is transformed into a *deep form list* (DFL) which indicates for each token its surface form, category, sub-category and a set of associated deep forms. The domain-independent lexicon is constructed by adapting the *lexical approach*, i.e. we store only canonical forms and assign to them all syntactic and semantic features [20].

The second step includes *UVL-analysis and spelling error correction*. UVL-analysis (unknown value list analysis) deals with domain-specific terms contained in the input. This separation of domain-independent and domain-specific terminology guarantees an easy portation to new application domains and forms the basis for an efficient application of spelling error correction. The latter is restricted to domain-specific terms because they are much more susceptible to the occurrence of spelling errors and possess particular importance for the sentence meaning. We make use of a similarity measure which is applied to candidates for substitution retrieved from the domain-specific lexicon.

Finally, *semantic, syntactic and pragmatic analysis* generate the semantic representation of the sentence by accessing the semantic application model which provides *activation rules* to choose the correct *semantic category* on the basis of DFL and UVL. Syntactic analysis is only applied if necessary for disambiguation whereas pragmatic analysis deals with input referring to previously asked questions so that relevant information is missing. The *representative question* for display is computed by means of a *generation template* which may contain variables to be replaced by the domain-specific terms in the UVL. In order to increase the precision of semantic analysis, the semantic deep forms are also hierarchically structured. The resulting *conceptual hierarchy* can then be applied to the calculation of a distance measure which gives a more accurate characteristic for the degree of semantic similarity of two concepts. Figure 3 gives an example, for similar approaches see [21, 22].

One of the main difficulties for the practical use of natural language interfaces is the so-called “knowledge acquisition bottleneck” which refers to the often required high amount of manual knowledge engineering which has to be repeated again and again for each portation of the interface to a new application domain. In addition to that is VIENA Classroom in contrast to simple static target applications a dynamic environment which is subject to constant changes such as modifications to the teaching material from revised versions to completely new lectures, alterations to the language usage of students with advanced skills and knowledge or continuous changes in the user population which add constantly new students with different linguistic preferences.

Therefore, we extend the interface by *adaptive techniques* to deal efficiently with the high demands on dynamic knowledge engineering in that we apply methods from machine learning as well as existing linguistic resources to the automatic acquisition of linguistic knowledge (for a recent survey of this topic see [23]). In addition to the basic process model we add the following powerful adaptive techniques:

- the incremental learning of a consistent set of activation and syntactic disambiguation rules for selecting the correct semantic category on the basis of past input data by adapting the *SE-tree approach* [24], a generalization of decision trees especially well suited for domains with few training examples and noisy environments,
- the automatic extraction of candidates for domain-specific terms from new teaching material,
- the use of the freeware Japanese-English dictionary file *EDICT* (Monash University) as secondary lexicon for the retrieval of information about unknown words and the automatic generation and insertion of corresponding new lexical entries,
- the application of the pronunciation data contained in the Japanese character file *KANJIDIC* (Monash University) to improve the precision of the spelling error correction module.
6. Conclusion

In this paper we presented the VIENA Classroom system which aims at providing a new kind of educational environment that goes beyond the limitations of traditional lectures. In particular we address the problems of personalization of shared information and question support for students which reduces the hesitation to ask questions so that we result in a more active participation and higher motivation.

We have just finished the implementation of the adaptive natural language interface, the complete implementation of the other components is still in progress. Future work will concentrate on the important issue of designing and applying an appropriate evaluation scheme.
to measure the success of our system. The evaluation will also show how the new environment will contribute to the emergence of new teaching and learning styles. On the basis of these results we plan to develop a revised version of our system for public use via the World-Wide Web. We believe that VIENA Classroom will become a valuable tool for virtual distance education which will represent a tempting alternative to traditional text-book lessons by making the classes more entertaining and enjoyable.

References