IV Workshop of
the AI*IA Working Group on
“Artificial Intelligence & E-Learning”

Reggio Emilia

December 12, 2009
**Workshop Coordinator**  
Adorni Giovanni, University of Genoa

**Scientific Committee**  
Andronico Alfio, University of Siena  
Carbonaro Antonella, University of Bologna  
Casadei Giorgio, University of Bologna  
Coccoli Mauro, University of Genoa  
Colazzo Gino, University of Trent  
Esposito Floriana, University of Bari  
Maresca Paolo, University of Naples “Federico II”  
Pazienza Maria Teresa, University of Rome “Tor Vergata”  
Pirrone Roberto, University of Palermo  
Poggi Agostino, University of Parma  
Roselli Teresa, University of Bari  
Rossi Piergiuseppe, University of Macerata  
Sarti Luigi, ITD – CNR  
Sugliano Angela Maria, University of Genoa  
Tasso Carlo, University of Udine  
Vercelli Gianni, University of Genoa  
Visaggio Giuseppe, University of Bari
Table of Contents

Building an AIML Chatter Bot Knowledge-Base Starting from a FAQ and a Glossary
G. De Gasperis.................................................................1

IWT: A Semantic Web-based Educational System
N. Capuano, S. Miranda, and F. Orciuoli.................................11

Neighbor Selection and Recommendations in Social Bookmarking Tool
A. Dattolo, F. Ferrara, and C. Tasso......................................17

Social Tagging in eLearning. Recommendations from Virtual Fellow-Students
I. Torre ............................................................................27

Capturing the Semantic Foundations of an Application Domain: an Ontology Based Approach
S. Alvino, S. Bocconi, P. Boytchev, J. Earp, and L. Sarti.............37

Issues on Intelligent Web-Based Education: Structuring the Subject Matter
G. Adorni, D. Brondo, M. Coccoli, and G. Vivanet.....................51

Ecli-Law: A simulated environment for problem solving in the field of law
P. Maresca, G. Pascuzzi, and L. Stanganelli .........................63

A Multi-Agent Environment for Tracking and Monitoring Learning Activity
P.G. Rossi, S. Carletti, D. Bonura, C.A. Bentivoglio ..................71

WordNet-based Summarization to Enhance Learning Interaction Tutoring
A. Carbonaro .....................................................................79

Acquisition of New Knowledge to Manage Conversation in TutorJ
R. Pirrone, V. Cannella, A. Pipitone, and G. Russo ....................85

A Prolog based tool in support to the instructional design process
S. Battigelli, L. Stanganelli, A.M. Sugliano ...............................95
Issues on Intelligent Web-Based Education: Structuring the Subject Matter

Giovanni Adorni, Diego Brondo, Mauro Coccoli, and Giuliano Vivanel

E-Learning & Knowledge Management Laboratory - DIST,
University of Genoa, Viale F. Causa 13, 16145 Genoa, Italy
{giovanni.adorni, diego.brondo, mauro.coccoli, giuliano.vivanel}@unige.it

Abstract. The aim of this paper is to present a knowledge based model for the design and development of units of learning and teaching aids. The idea behind the proposed model comes from the analysis of the open issues in instructional authoring system and the lack of a well defined process that merges a system for the knowledge organization of the domain with a pedagogical strategy. Within this paper we define an Educational Concept Map, a logical and abstract annotation system, derived from fundamental theories of instructional design, created with the aim of guaranteeing the reusability of the teaching materials and of the whole knowledge structures. Authors can create lessons or entire courses starting from an ontological structure characterized by the integration of hierarchical and associative relationships between the educational objectives.

Key words: Web-base education, ontologies, learning content design

1 Introduction

During the last years, Web-based education (WBE) has become a very important branch of educational technology. For learners, it provides access to information and knowledge sources that are practically unlimited, enabling a number of opportunities for personalized learning and collaboration, with clear advantages of classroom independence and platform independence; for teachers and authors of educational material, it provides access to Web resources and authoring tools for developing Web-based courseware. In the context of WBE, educational material is generally distributed over a number of educational servers [1]. The authors (teachers) create, store, modify, and update the material working with authoring tools on the client side. Likewise, learners use different learning tools to access, browse, read, and consult the resources fulfilling their learning tasks. The evolution of Artificial Intelligence (AI) during the first fifty years of this discipline has made the AI technologies an interesting framework and tool for supporting education processes. Intelligent Web-based education (IWBE) results from applying intelligent technologies to WBE. However, with reference to the application of Artificial Intelligence based technologies into IWBE, Mizoguchi and Bourdeau [2] highlighted some open issues, namely: a deep conceptual gap
between authoring systems and authors; authoring tools are neither intelligent nor particularly user-friendly; knowledge and components embedded in IWBEs are rarely sharable or reusable; the authoring process is not principled; there is a gap between instructional planning for domain knowledge organization and tutoring strategy for dynamic adaptation of the IWBE behavior. The importance of knowledge reusability and subject matter analysis have been long recognized by instructional designers with the aim of supporting meaningful learning processes through a careful knowledge selection, organization, and sequence [3]. A relevant problem concerns the fact that there are no canonical representation of knowledge structures (because any of them can be seen from different points of view, showing different structures). As Ohlsson [4] highlighted, this fact has such relevant implications for authoring systems that it should be stated as the “Principle of Non-Equifiability of Learning”, according to which "The state of knowing the subject matter does not correspond to a single well-defined cognitive state. The target knowledge can always be represented in different ways, from different perspectives; hence, the process of acquiring the subject matter have many different, equally valid, end states". It should be clear that a large number of possible instructional paths representations exists and the problem is to find the appropriate content and presentation for the specific educational purpose which has been defined [5].

In this paper, the authors’ attention will be focused on the formal representation of the subject matter structure (and the related educational goals) in the context of learning environments. In addition, an Educational Concept Map (ECM) model [6] for learning content design will be proposed.

The final goal of this research is the design and development of an authoring environment able to support also a formal representation of the subject matter structure. This system, CADDIE (Content Automated Design & Development Integrated Editor), is under development at the E-Learning & Knowledge Management (ELKM) Laboratory of the Department of Communication Computer and Systems Science of Genoa University [7]. It will exploit the ECM model presented herein for the design and the development of learning paths and related teaching aids. CADDIE can help teachers and instructional designers at different stages of their work: macro-design of courses or single units of learning; micro-design, linking learning resources to related topics into the ECM; editing and/or developing learning resources; semantic-based indexing of resources. CADDIE will enable instructional designers and teachers to organize courses in respect of the non-linear structure of knowledge representation that humans have in mind so that meaningful learning processes can be significantly supported [8].

2 Related Works

Since their origins, Knowledge Representation theories have been applied in the context of Computer-Assisted Instruction (CAI) with the aim of supporting Intelligent Tutoring Systems (ITS) [10][11] by means of a formal representation of: the subject matter to be taught; educational goals; the learner’s state at the
beginning and during the instructional path; and the learning model, including learning strategies and instructional context [12]. With regards to the first issue, Stelzer and Kingsley [13], whose work has inspired part of this study, proposed a comprehensive theory for organizing and describing subject matter structures. In their theory, founded on the paradigm of axiomatics, these structures are composed by content (content elements consist of Primary Notions, Secondary Notions, Basic Principles and Established Principles) and tasks (in order to specify tasks, the behavioral objectives of the instruction must be known at the beginning of any instructional design process). Moreover, they introduce the notion of dependency between content and task components so as to restrict the order in which contents can be presented in the course of learning processes.

Over the last decade, we have witnessed a revival of studies related to the representation of learning content structures. The major contribution to these topics has come from studies on Educational Modeling Languages (EMLs) [14] and the Semantic Web field [15]; it is worth to notice that in this latter case it has been oriented to the problem of the domain knowledge representation rather then to knowledge structures representation for learning purposes.

An EML provides a formal conceptualization of the learning process regardless of instructional theories, describing learning units not only to allow software applications to interpret an EML script, but also to promote the reuse and the exchange of these descriptions among different e-learning environments [16]. Nowadays, IMS Learning Design [17], which evolved from OUNL-EML Educational Modeling Language (OUNL-EML) by Open University of the Netherlands (OUNL), is probably the most widespread EML. It aims at representing the "learning design" of "units of learning" in a semantic, formal and machine interpretable way. IMS LD specification focuses its attention on the learning scenario as a whole and not specifically on subject matter structure representation.

Martínez-Ortiz et al.[16], classified EMLs into the following categories: evaluation languages (languages allowing designers to describe learning process stages in which problem-solving or question-answering are involved in an abstract way); content structuring languages (languages allowing designers to arrange the instructional resources in sequence, always taking into account learner’s needs and performances in order to enhance learning experiences); and, finally, activity languages (languages focused on the activities during the learning process). They defined as "content structuring languages": Learning Material Markup Language (LMML) by University of Passau [18]; TARGETed REuse and GENeration of TEAching Materials (Targetteam) by Universität der Bundeswehr [19]; AICC Course Structure Data Model (AICC-CSDM) [20]; IMS Simple Sequencing (IMS-SS) [21]; and, finally, ADL Sharable Content Object Reference Model 2004 (ADL SCORM) [22].

In the recent years, other different approaches to the learning content structure representation have been proposed [23][24][25][26]. Felix and Paloma [27] presented a framework based on an instructional application model, called Xedu, that provides entities representing instructional components and that will drive the instructional design process.
Pechenau et al. [28] proposed a theoretical model, named \textit{COUL-M} (COceptual Units’ Lattice Model), for representing the domain knowledge of an instructional system.

A great contribution to the problem of knowledge structure representation in learning design has come from the Semantic Web field (typically, these researches are based on an ontological approach to the conceptualization process) [2] [29] [30].

A learning content structure ontology should be based on several kinds of links that represent different relationships types among subject matter topics; the most common associations in these ontologies are "prerequisite links", which represent the fact that one concept has to be learned before another, other common relationships are traditional semantic links, such as "is-a" and "part-of" (categorizing topics into classes and subclasses). In more elaborate ontologies, the formal representation can include a vocabulary to classify educational units, such as \textit{definition}, \textit{example}, or \textit{exercise} [29]. Stojanovic et al. [31], referring to ontology-based metadata in the context of e-learning environments, discuss the use of metadata for describing the structure of learning materials.

According to them, several types of structuring relations between chunks of learning material may be identified, such as \textit{Prev}, \textit{Next}, \textit{IsPartOf}, \textit{HasPart}, \textit{References}, \textit{IsReferencedBy}, \textit{IsBasedOn}, \textit{IsBasisFor}, \textit{Requires}, \textit{IsRequiredBy}. On the other hand, it is interesting to highlight that such a kind of metadata representing relationships appears in conventional metadata models, such as the IEEE Learning Object Metadata [32] at the \textit{Relation} level.

3 The Educational Concept Map Model

The ECM is a logical and abstract annotation model (Fig: 1) created with the aim of guaranteeing the reusability not only of teaching materials, but also of knowledge structures (moving the generalization level from the contents to the definition of the contents’ schema). It derives from fundamental theories of instructional design which can only be briefly discussed here. According to Merrill’s [33] and Gagné’s [34] theories, once the learner profile is known, the instructional design process should start from the definition of a hierarchical organization of learning objectives describing what students should know or be able to do at the end of the instruction. This hierarchy of learning goals provides the general framework for contents selection. The preliminary organization of learning objectives and contents into a hierarchical structure (making logical order of learning content explicit) comes also from the Ausubel’s theory [8]. He introduced the notion of “advanced organizer” as a way to provide the cognitive structure, or, in other words, the mental scaffolding during meaningful learning processes. Referring to these theories, the ECM model has been developed in an ontological structure characterized by the integration of hierarchical and associative relationships. Firstly, it asks the authors to focus their attention on learners’ profile (in particular educational background, learning and cognitive styles) and objectives. Basing on these elements, the model suggests identifying, within the discipline’s subject matter, the key concepts and their relationships
so as to identify the most effective strategies of contents presentation and to support the activation of meaningful learning processes.

Educational objectives, according to the model, can be represented as SingleObjective or ComposedObjective, the former aren’t decomposable into sub-objectives, the latter are constituted by two or more sub-objectives. Single units of learning including topics (SingleUoL) will be associated to the first class; while composed units of learning (ComposedUoL) including SingleUoL or ComposedUoL will be linked to the second class. The model is based on a hierarchical and recursive organization (through the is-a relationship) of learning objectives to which corresponds a layered structure (n levels with n integer positive) of contents. Moreover, it is worth noticing that the relationship between an objective and a unit of learning is always necessarily a 1:1 association (a SingleUoL corresponds to a SingleObjective). An Educational Concept Map can be therefore defined by the syntax:

\[
\langle ECM \rangle ::= \langle UoL \rangle + \\
\langle UoL \rangle ::= \langle SingleUoL \rangle | \langle ComposedUoL \rangle \\
\langle SingleUoL \rangle ::= \langle Topic \rangle \\
\langle ComposedUoL \rangle ::= \langle ECM \rangle \\
\langle Topic \rangle ::= PrimaryTopic|SecondaryTopic
\]

In addition, it’s possible to define propaedeutic relationships between two objectives, and, as a consequence, there will be an equal association between the corresponding units of learning. Once the logical scheme of educational objectives is determined, it’s possible to identify and organize the schema of contents based on a taxonomic learning units organization (in fact, each objective is pursued through the corresponding UoL). Then, for each unit of learning, the ontology needs to specify the topics, the key-concepts on which the UoL is focused. Topics can belong to the two following classes [6]:

Fig. 1. Educational Concept Map Model
Primary Topic: this class identifies the "prerequisites", in other words the concepts that the student must know before attending the course (the set of these topics, which will not have instructional resources associated, represent the knowledge requirement of the course);

Secondary Topic: this class identifies the concepts which will be explained in the course of the unit of learning (this kind of topics will have specific learning materials associated).

The relationship between a primary and a secondary topic is named is-primary-topic-of. Primary topics have to be specified for each unit of learning, in order to allow the UoLs reusability in various instructional processes. In consequence of this the same topic could be considered as primary topic within a unit of learning, and as secondary topic in a different one.

The relationships among secondary topics can be [6]:

- is-requirement-of: it identifies a transitive and propaedeutic association among two or more topics (e.g., it may be used with the aim of specifying the logical order of contents);
- is-related-to: it identifies a symmetric association among closely related topics (e.g., it may be used with the aim of creating learning paths without precedence constraints);
- is-not-related-to: it identifies a symmetric relation of indifference between two or more topics (e.g., it may be used with the aim of making explicit the absence of association among topics);
- is-suggested-link-of: it identifies not closely related concepts (e.g. this relationship type may be used in order to suggest in-depth resources, internal or external to the contents repository).

These relation types have been selected with the aim of allowing teachers and instructional designers to create different learning paths (with or without precedence constraints among topics). More formally, a prerequisite is defined as follow: given a ECM and a set \{t_1, \ldots, t_k\} of topics \(T\):

\[ t_i \in T \text{ is defined } \text{"prerequisite" of } t_n, \Leftrightarrow \exists \text{ is-primary-topic-of } (t_i, t_n) \]

with \(n = 1, \ldots, i - 1, i + 1, \ldots, k\)

Similarly, it is possible to define the "learning outcomes" as follow: given a ECM and a set \{t_1, \ldots, t_k\} of topics \(T\):

\[ t_j \in T \text{ is defined a } \text{"learning outcome" } \Leftrightarrow \exists \text{ is-requirement-of } (t_j, t_n) \]

with \(n = 1, \ldots, j - 1, j + 1, \ldots, k\)

Subsequently to the ECM design, it will be possible to associate educational resources to the single nodes (Learning Objects, text documents, audio and/or video files, etc.) [6]. Once the resources are linked to the topics, the Educational Concept Map can be published on the Web and the relationships suggest the different navigation strategies of the underlying subject matter. More precisely,
each Web page related to a topic contains link to topic(s)/page(s) which is(are) considered prerequisite(s) on this ECM (by means of is-requirement-of relationship), link to related topic(s) (by means of is-related-to relationship), link to suggested readings/topic(s) (by means of is-suggested-link-of relationship).

Such maps can also be used to generate linearized path useful to produce a lesson or a document about a subject matter. In this latter case a Suggested Paths Strategy is necessary, by means of is-requirement-of relationships. Differently from some of aforementioned EMLs, the ECM model allows authors to define alternative learning content presentation strategies, in order to offer personalized learning paths. The definition of this strategy, introduced also in reply to the non-equifinality problem, derives from the Cognitive Flexibility Theory [35], according to which learning processes (especially in ill-structured and complex domains) requires a multiple representations of content and the criss-crossing of the conceptual landscape which should be revisited from different directions with the aim of acquiring an advanced knowledge [36].

Let us to consider the following example to explain such strategy, where only is-requirement-of relationships are used, represented as oriented arcs (see Fig: 2). Let us also consider the idea of preparing a lesson on a given argument, using the previous ECM, with \( R \) as learning outcome starting from a topic \( D \): then the following paths (string of nodes) are possible:

\[
(DGLPR) \\
(DGMPR) \\
(DGHIQR) \\
(DENMPR)
\]

![Fig. 2. Example of Educational Concept Map where only is-requirement-of relationship is used](image-url)
Where the maximum distance $\Delta_{\text{max}}$ (the number of topics) between the two topics on the graph is 4. We define, at this point, a Topic Aider (TA) as a chunk of knowledge (for instance, an example, an integrative reading, an exercise, a self evaluation test) somehow useful to complete the description of a topic or for a better understanding of the topic itself.

The Suggested Paths Strategy works as follows:

1. For every path between two nodes, and for every node of the above path, reading the string from left to right do:
   (a) *insert* before a node on the string all the possible subpaths from Primary Topics to that node, reading the graph from top to bottom;
   (b) *prune*: reading the new string (as generated at step 1.a) from left to right, delete possible node repetitions (keep only the first letter of a node);
   (c) *compute the real distance* $\Delta S_{ij}$ on the string of every contiguous node (nodes with $\Delta = 0$) of the graph;
   (d) *insert suggestions* if a string exists with $\Delta S_{ij} > \Delta_{\text{max}}$, insert a suggestion for a Topic Aider of the node $i$ ($TA_i$) before the node $j$;

2. Order of the suggested paths: order the generated strings following the criteria of: number of $\Delta S_{ij} > 0$, and on equal terms for increasing $\sum \Delta S_{ij}$.

3. Allow the user to choose one of the path of the list at step 2.

Referring to the previous example we will have:

\[
\begin{align*}
A & D G L B E N M P C H I Q R \\
\end{align*}
\]

On the base of previous step 2 the $\sum \Delta_{ij} = 3 + 4 + 4 + 4 + 7 = 22$. On the base of previous step 1.d, the Suggested Path is:

\[(ADGLBENMPC[TA_G]HIQR)\]

And so on with the other possible paths shown in (3).

\[
\begin{align*}
A & D G B E N M L P C H I Q R \\
\end{align*}
\]

On the base of previous step 2 the $\sum \Delta_{ij} = 1 + 2 + 3 + 4 + 4 + 7 = 21$. On the base of previous step 1.d, the Suggested Path is:

\[(ADGBENMLPC[TA_G]HIQR)\]

\[
\begin{align*}
A & D G C H I Q L B E N M P R \\
\end{align*}
\]
On the base of previous step 2 the $\sum \Delta_{ij} = 1 + 4 + 4 + 6 + 7 + 8 = 30$. On the base of previous step 1.d, the Suggested Path is:

$$ADGCHIQLB[T_A D]EN[T_A C]MP[T_A Q]R$$

$ABDENGMLPCHIQR$

On the base of previous step 2 the $\sum \Delta_{ij} = 1 + 1 + 1 + 1 + 3 + 4 = 11$. On the base of previous step 1.d, the Suggested Path is:

$$ADBENGMLPCHIQR$$

The Suggested Path List among which the user could choose is:

1. $ADBENGMLPCHIQR$
2. $ADGBENMLPC[T_A C]HIQR$
3. $ADGLBNMPC[T_A C]HIQR$
4. $ADGCHIQLB[T_A D]EN[T_A C]MP[T_A Q]R$

We don’t have a single path (the "best" according to some minimum principle) but a list of paths to suggest to the author leaving the final choice to the author him/herself, to answer to the non-equifinality problem posed by Ohlsson [4]. The ”suggested” order lists is on the base of the principle of reducing as much as possible the distance of two topics of the list which are contiguous on the graph.

## 4 Discussion and Future Work

An Educational Concept Map allows creating learning paths and instructional resources more easily interoperable and reusable, because the concept representation is independent of its implementation. The same concept network could be used for designing different courses. It might be thought to the possibility of delivering a course whose content needs to be personalized according to different learners (with different educational objectives, skills and pre-knowledge). In this scenario, the ECM holds steady, and what changes are the specific instructional resources associated.

Once the ECM logical model has been defined, the ISO/IEC 13250 Topic Maps [9] standard has been selected as knowledge representation language for its implementation. Topic Maps (TM) are an ISO multi-part standard designed for encoding knowledge and connecting this encoded knowledge to relevant information resources. The standard defines a data model for representing knowledge structures and a specific XML-based interchange syntax, called XML Topic Maps (XTM). The main elements in the TM paradigm are: topic, association and occurrence. According to ISO definition, a topic is a symbol used within a topic map to represent one (and only one) subject, in order to allow statements to be made about the subject. An association represents a relationship between two or more topics. An occurrence is a representation of a relationship between a subject and an information resource. Therefore, two layers can be identified into the TM paradigm:
– the knowledge layer that represents topics and their relationships, allowing to construct the ECM model;
– the information layer that describes information resources, to be attached to the ECM topics.

Graph linearization is another relevant issue which needs to be considered, to take into account technological constraints regarding, for example, the respect of standards requirements (such as the ADL SCORM), and in order to support the learning process tracking of a LMS. The relationships used in the previously discussed model can be automatically processed and linearized, with the aim of extracting different learning paths.

References