Model and algorithm of the conceptual scheme formation for knowledge domain in distance learning

Oleg Zaikin *, Emma Kushtina, Przemysław Różewski

Szczecin University of Technology, Faculty of Computer Science and Information Technology, Zobniwerska 49, 71-210 Szczecin, Poland
Available online 18 March 2005

Abstract

The article describes a knowledge model, oriented to the asynchronous mode of distance learning. The formalization of the knowledge model for a given domain, the operations on the knowledge and the algorithm of the knowledge model creation are submitted. All received decisions can be realized in a program environment compatible with the SCORM standard. The described methodology, based on a generalized knowledge model, enables to develop a distance learning course mainly for the fundamental knowledge. In this paper we describe the methodology and illustrate its use through a project to develop a distance learning course for a queuing system. Moreover, a practical application is proposed based on the eQuality project.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Knowledge-based systems; Education; Human-centred technology; Ontology; Distance learning

1. Introduction

Distance learning (DL) is widely developed in traditional educational institutions as well as in large corporations [55,63]. The modern approach to DL is based on two different methods of training: synchronous and asynchronous [67]. The asynchronous method is based on the student’s access to the training system at any time and at any place by means of the information network. Considering the cost of training and elasticity of application, in the reality of corporations and universities the asynchronous method is more popular. Therefore, in the given article, under the process of training we understand distance learning in the asynchronous mode. Moreover, the asynchronous systems of a DL training should achieve:

- High flexibility of representation and contents of didactic materials;

* Corresponding author. Tel.: +48 91 449 55 08; fax: +48 91 449 55 40.
E-mail addresses: ozaikine@wi.ps.pl (O. Zaikin), ekushtina@wi.ps.pl (E. Kushtina), prozewski@wi.ps.pl (P. Różewski).

0377-2217/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.ejor.2005.02.014
• Fast adaptation to different social and cultural conditions;
• Clarity of the incorporation and students access to the didactic materials;
• Effective knowledge transfer from an artificial computer system to the student.

Distance learning systems give students good opportunities to cooperate with the system and to have a great initiative at the same time. Flexibility of the content’s expression, rapid adaptation to different social and cultural issue and open nature on the implementation and access level are main indicators of DL asynchronous system. In such conditions a less creative student automatically becomes a less successful one, since the DL system allows the student to organize the learning process on his/her own. This means that DL systems have high requirements as to the creative and initiative abilities of the student.

The process of “teaching–learning” in an asynchronous mode of DL is characterized by the new function of the teacher, who loses the dominant position and influence on the educational process and transfers this role to the didactic materials and learning process organization. In the process of interactive co-operation with the DL system the student extracts and assimilates the subject knowledge independently. Therefore, didactic materials in a DL system should contain the necessary portion of knowledge, as well as the cognitive, built-in scheme, compensating this way the absence of the teacher in the learning process.

The continuous development and success of DL technologies essentially depend on the joint consideration of its pedagogical and economic aspects [12,30,58]. This is caused by the necessity to increase the efficiency of the DL process and, on the other hand, to reduce the expenses of its organization. The pedagogical aspects of DL are mainly related to the model of knowledge representation and transfer and the used pedagogical tools [13]. The economic aspects are related to the development of telecommunication space ensuring interactive access to common information resources for the teacher and the students in a time-sharing mode. The problem includes unification of pedagogical and economic aspects within the framework of the integrated models of a DL system design and support. Until now, the solution has been based on the approaches oriented to models of the knowledge representation and transfer for traditional environment and organization of teaching. However, a DL system functions in a new teaching environment, with other standards and forms of education and it dictates the necessity of other approaches.

In traditional learning we deal with a situation, when the knowledge of a subject domain and the technique applied to teaching are continuously accumulated and organized in the mind of the teacher and are referred to didactic modes of knowledge representation and transfer. The teacher realizes the process of knowledge adaptation to objectives and conditions of teaching in real time. In the knowledge adaptation process a teacher in his/her own, not formalized way, forms a set of concepts and defines their semantic depth and mutual relations. Keeping in direct touch with the students, the teacher has an opportunity to change the technique of teaching and also the depth, structure and borders of the used concepts, depending on knowledge accumulated by the students.

In the DL process there is no direct contact between the teacher and the students. This important function of the learning process is lost, and that results in the loss of learning efficiency. Thus, in DL conditions the question appears—how to change the role of the teacher and what replaces his functions as the knowledge “provider”, minimizing his/her absence in the learning process?

The solution to this problem is possible only via automation of intellectual processes, which are carried out by the teacher in the teaching process. At the same time, it is known that only formally described processes can be the object of automation. Automation of such processes demands solving the following tasks:

• definition of basic operations, which form the teaching process;
• control of results of these operations, in other words, setting a dependence between the input and output data for each operation;
• optimization and choice of the best parameters of the learning process;
• comparison of the final results of the traditional and distance learning processes.

The paper is organized in the following way: in the second section the knowledge issue is discussed in the perspective of the DL process. In the third section the concept network creation algorithm and didactic materials compilation algorithm are introduced. The 4th section covers some field of knowledge engineering, mainly focused on the ontology engineering. In the next two sections the given algorithms are discussed in detail. In Section 7 the algorithms are compared with the requirements of the SCORM 2004 standard (Sharable Content Object Reference Model)—one of the most significant distance learning standards [51]. Based on the eQuality project [2] in the 8th section the algorithms application is demonstrated.

2. Knowledge in the traditional and distance learning process

The natural language (NL) is the basic mode of knowledge representation and transfer in conventional teaching. However, various types of symbolic/artificial languages can also be used (e.g. [34]), what gives additional possibilities to enhance the semantic expressiveness of the NL. From the point of view of ergonomics and psychology, the computer environment sets bounds on the NL usage [11]. On the other hand, computer environment gives unlimited possibilities for using symbolic and artificial languages and media.

The question appears: how to achieve, in the DL environment, the level of NL accuracy in the knowledge manipulation? The solution to the problem is developing a knowledge model of representation and delivery. Knowledge, as an object and purpose of study, can be divided into two basic kinds: fundamental (theoretical) and operational (procedural). Each kind of training activity, irrespective of the subject domain, demands the assimilation of both theoretical knowledge and practical skills.

Fundamental knowledge reflects conceptual thinking and contains new paradigms, problem statements, principles of behaviour etc. Procedural knowledge is necessary for the development and realization of scenarios, algorithms and the performance of various kinds of operations. Analyzing various problems and situations, examined in the learning process, demand simultaneous use of both kinds of knowledge but in different proportions, depending on the complexity of the problem that is being solved. New methods of procedural knowledge representation and DL technologies allow to create effective systems of computer training (simulators, expert systems, etc.). Such systems set a relationship between practical skills (speed, accuracy, conformity to standard) and the available theoretical base in real time.

Therefore, conceptual thinking is shown in the following:

• abilities of abstraction and generalization,
• setting associative links,
• inference of a new knowledge on the basis of the available one,
• formulation of paradoxical knowledge which exceeds the existing paradigm.

There are no objective quantitative methods of evaluating the conceptual knowledge acquired during training. Therefore, assessment of the quality of learning the conceptual knowledge is based on the objective properties of the domain and techniques of studying it. With the same amount of knowledge in the subject domain and other equal conditions the various techniques of training can affect the development of conceptual thinking in a different degree. The responsibility for the choice of the training technique lies in the hands of the teacher. Therefore, in the conditions of DL, it is necessary to use the methods of theoretical knowledge representation that give opportunities to present objective knowledge in the subject’s domain as well as techniques of studying it.

The analysis of existing models of knowledge representation shows that none of them satisfy the demands of DL. It is shown in [48] that the structure of knowledge, which differs from the Rules Systems, can be presented on the basis of the semantic networks. The semantic networks can represent both abstract categories and certain objects. The big obstacle for using the semantic
network is the difficulty with formal representation of the semantic relations. Semantic networks are the starting point for such knowledge representation models as Mind Maps [9], Conceptual Maps [56,44] and Topic Maps [4,32]. The models are oriented towards a specific kind of knowledge and a concrete user. Each of them is dedicated to a specific kind of tasks and a specific domain. The level of universality in the DL knowledge model should be high. The appropriate model must at least work with one kind of knowledge (e.g. theoretical) in all its contexts. The final knowledge representation system should merge the knowledge manipulation language with the corresponding pedagogical approach, which is used to learn about a subject.

The mind stores knowledge in the form of a concepts [5,41]. The concept is considered as a structure of mind representation, which consists of the proper description of a single-meaning class. The methods of conceptual knowledge teaching take advantage of the concept’s communication level between the students. Teacher can use metaphors or other pedagogical methods [13]. The most important issue is to establish the cohesion state, what is achieved when both, teacher’s and student’s minds refer to the same concepts prototypes simultaneously.

The proposed knowledge model plays the role of a collaboration tool between the teacher and the subject matter expert. In the phases of the learning objectives’ creation and student’s knowledge discovery, the following activities are performed based on the knowledge model:

- learning object creation,
- semantic border localization,
- synonyms chain creation.

These operations give the possibility to edit the concept semantic from the expert’s side as well as the teacher’s side. The knowledge model is the basis for the modular didactic materials formation. The model of each concept is formulated as a logical unit with references to the media metaphor and other concepts that are stored in the repository. Such approach is compatible with the requirements of the Learning Management Systems (LMS) and Learning Content Management Systems (LCMS), which are widely used in DL [29]. LMS is a strategic solution designed for planning, delivering and managing all the training events in the company, considering virtual classes as well as the ones taught by an instructor [25]. The Learning Content Management System [8] is what we call a system used for creating, storing, making available (sending) personal educational content in the form of a learning object.

The structure and construction of the learning object are covered in detail by the SCORM standard [51]. However, there is a lack of information or knowledge about the content of the learning object in the SCORM standard. The international research society has been investigating the problem of the learning object for several years [15,27,33,47], however, the general solution has not been found yet. Until now the set of guidelines and rules has been published [27]. The learning object problem is also discussed in [39,53,66] and European Union programmes [1,3]. The authors have proposed the above approach as the next stage of the universal distance learning system creation.

3. Idea of algorithms of the knowledge domain representation based on the concepts network formulation

The next two sub-sections present the general idea of the concepts network creation algorithm and didactic materials compilation algorithm. The algorithms and associated concepts will be discussed in detail in Sections 4–6.

3.1. Concepts network creation algorithm

The goal of the concepts network creation algorithm (CNCA) (Fig. 1) is to identify the knowledge in a specific subject domain and convert the knowledge to the form of a concepts network (CN). The subject domain analysis is conducted on the first step of the expert’s work. The subject matter expert analyses the domain and the concepts are selected according to the rules of the concepts description. The identified set of concepts is
enhanced by adding a semantic relation. Simultaneously, the creation of the concepts multimedia representation is being performed. The main function of each metaphor is to describe the concept widely or the idea which is covered by the concept. The abstract form of the concept implies its digital representation. The complete processing information, including concepts, meta-information (each notion is characterized by the author, data, etc.), media files is stored in the repository. The criterion I of the concepts network algorithm decides whether the choice of the multimedia metaphor is correct. The expert, who is carefully considering the judgment process, can enlarge the depth of the concept metaphor basing on his/her knowledge of the subject. After the positive semantic verification the CN is applied to the repository structure. The whole system is evaluated again according to criterion II. The network of concepts should cover every bit of knowledge in the specified domain. The results of the algorithm are:

- Concepts network;
- Repository;
- Hypermedia network of concepts.

The last one maps the concepts onto their multimedia representations stored in the repository. The algorithm outcomes are fully compatible with the SCORM standard.

3.2. Didactic materials compilation algorithm

The didactic materials compilation algorithm (DMCA) (Fig. 2) adapts the didactic materials for the student by taking under consideration the

![Fig. 1. The algorithm of the concepts network designing.](image1)

![Fig. 2. Didactic materials compilation algorithm.](image2)
general educational standards and pedagogical conditions of the learning process. In general, the CN is adapted to a specific student by the content personalization. The goal of the algorithm is to develop a sequence of educational elements (in the sense of the SCORM 2004 standard), which is delivered to the student through the telecommunication network.

The CNCA algorithm creates a CN about the subject area. In the first step of the DMCA algorithm, the relationships in CN are ordered basing on the student’s profile. In the next step, the ordered CN is transformed into a multilevel hierarchical graph with the basic concepts at the top level. Concepts at the low level of the hierarchical graph are interpreted as the objectives of learning. Such graphic representation of knowledge determines uniqueness of each concept and the number of its semantic links. The links define semantic richness of each concept and its importance for other concepts.

At the next stage of the algorithm, the multi-level hierarchical network is decomposed into its own set of tree-like sub-graphs which are based on the methods described in this paper. The trees are covered by a ‘knowledge portion’ mechanism, which is constructed basing on cognitive regulations. The selected ‘knowledge portion’ is a basic element of the learning object. Transformation of the ‘knowledge portion’ into a learning object requires a representation of the concept in the repository content.

The data and its metadata are analyzed basing on criterion III. Criterion III estimates the ‘size’ of the data according to the telecommunication network bandwidth limitation, the copyright issue and the resources accessibility. The clustering process joins the concepts of every ‘knowledge portion’ into one node. The knowledge incorporated into the ‘knowledge portion’ together with the essential metadata description creates the learning object. The generated learning objects are impinged on the CN, which is after that converted to a learning object network. Afterwards the sequence of nodes (learning objects) is set. The established learning object chain determines the way of learning for each student. Before the learning objects sequence will be placed in the repository, all elements are converted to a SCORM-compatible form. The course instructor/teacher has to make the final decision about the didactic materials acceptance (criterion IV). Finally, the didactic materials are accessible for every registered student.

4. Ontology considered as a knowledge model

In general the ontology \( \Omega \) (called lightweight ontology) can be defined as a 3-tuple \([18,28,64]\):

\[
\Omega = \langle CD, ID, RD \rangle,
\]

where \( CD \) is a set of classes, which defines the concepts used to the real object description; \( ID \) is a set of instances, which represents the instance of the concept defined in the set of classes; \( RD \) is a set in relations on the set of classes.

In the paper the following assumptions are made:
- Ontology describes the fundamental knowledge;
- Ontology is build by the subject experts of the specified domain;
- Concepts are the ontology nodes, connected by relations;
- Ontology can be considered as an unordered graph;
- Concepts have to be unique along the ontology.

4.1. Semantic unit of the knowledge model description

Reality is defined by an unlimited and diverse set of information and stimulus, which attract human perception system. The cognitive science assumes the mind natural ability of conceptualization [23]. The conceptual scheme exists in every domain [16] and informs about the domain boundaries and paradigm. Basing on the ontological approach one can create a conceptual model of the knowledge domain, where the concepts are the atomic, elementary semantic structures [7,26,59,60]. On practical point of view, the ontology is a set of concepts \( \{c_1, \ldots, c_n\} \) from a specific domain. The modelling approach, considered as a cognition tool, assumes certain level of simplification of the examined object.
The following definition of a concept will be used in our research [36,38]: the concept is a nomination of classes of objects, phenomena, abstract category, for each of them the common features are specified in such way that there is no difficulty with distinguishing every class. Having of given definition makes possible the modelling of the knowledge model for any domain basing on the set of basic concepts which were specified by an expert in a verbal way.

The concept $\phi$, is defined as a tuple:

$$\phi = \langle N(\phi), X, T \rangle,$$

where $N(\phi)$ is the name of the concept $\phi$; $X$ is a set of information, which provided the concept description. The description is made basing on one of the metadata standards (e.g. DublinCore, IEEE LOM). $T$, the matrix of the concept depth $T = [t_{ij}]$ has the following meaning:

$$T[i_j] = \begin{cases} 
N(\phi), & \text{when } i = 1 \land j = 1, \\
\text{Attribute } i, & \text{when } i \neq 1 \land j = 1, \\
\text{Object } j, & \text{when } i = 1 \land j \neq 1, \\
t_{ij}, & \text{otherwise}.
\end{cases} \quad (3)$$

All elements of the matrix $T$ belong to the specified domain, as a result, they are included in a single ontology $\{N(\phi), \text{Object } i, \text{Attribute } j, t_{ij}\} \in \Omega$, for $i = 1, \ldots, I, j = 1, \ldots, J$.

Such concept definition allows a formal representation of its internal structure and makes it possible to shift from a verbal form to a formal description in the form of an abstraction, which can be organized as a matrix or another graphic structure [24,34]. In the case of the matrix (Fig. 3), the number of rows equals the number of attributes of the concept and the number of columns symbolizes the amount of objects merged into one concept class [62]. The number of elements in the matrix $t_{ij}$ is called the concept depth in the specified domain. The coloured field (Fig. 3) describes the context of the concept. The basic relations, which occurred between concepts and are defined by the subject’s matter expert, are the following: homonym, synonym, meaning extension and meaning deepening.

Aggregation ($PART_OF$) and generalization ($IS_A$) are semantic operations, which can be considered as a result of the abstraction creation method. The concept can be examined as an abstraction [22,57], what helping to understand a complex object by decomposing it into uncomplicated components.

4.2. Concepts network relations

Similar to the semantic network, the concepts network (CN) does not attempt to represent the real mind structure. The CN is no more than an abstract representation, which is an instrumental model adapted to the domain knowledge manipulation in the form of ontology engineering. Researchers believe that knowledge in the mind is organized according to a cognitive economy: concepts with a higher number of semantic relationships are situated on the upper level of the knowledge representation’s structure.

In general, a CN always covers knowledge from one, specific domain and for that reason can be considered as an ontology. The statements confirming such declaration come from previous assumptions.

The conceptual framework for a CN is based on the ontology characteristic. The concept $\phi = \langle N(\phi), X, T \rangle$ has been defined in the previous section. Semantic relation is defined as a relation between two concepts. Out of many other existing ones, the set of semantic relations: $R = \{IS_A, PART_OF, \emptyset \}$ is incorporated into the approach. According to [22,57] the inclusion, aggregation and association relations classes are the most important ones. The last class notices some information about the association between concepts. In CN this type of information is carried on by the inclusion and aggregation class. The remaining relations’ set $R$ is, on one hand, relatively small but, on the other hand, the cohesion with the abstraction mechanism is preserved.

The IS_A relation, according to [17], covers wide range of relations like: inheritance, implication and inclusion. The most popular IS_A application is taxonomy. If any elements of $T_1 = [t_{ij}]$ (for the concept name $N(\phi_1)$) can be described as a matrix $T_2 = [t_{ij}]$ (concept name $N(\phi_2)$) then the concept $\Phi_2$ is connected with the $\Phi_1$ by the IS_A relation.
The PART_OF relation is a base for several types of hierarchy (e.g. whole-part) [56]. If \( N(\phi_1) \subset N(\phi_3) \) and \( T_1 = [\tilde{t}_{ij}] \subset T_3 = [\tilde{t}_{ij}] \), then the concept \( \phi_3 \) is connected to the \( \Phi_1 \) by the PART_OF relation.

In presented approach the ontology is defined as a tuple:

\[ \Omega = (S, \Pi) \]

where \( S = \{ \phi_i \}, i = 1, \ldots, n \) is a set of concepts from a specific domain; \( \Pi : S \times S \to R \) is a mapping from an ordered pair of concepts to the set of connections \( R \) (similar to [10]).

In order to improve computer processing, the \( \Omega \) form is transformed into a non-oriented graph \( G_\Omega \). The ontology in form of the non-oriented graph \( G_\Omega \) has the following definition:

\[ G_\Omega = (V, E) \]

where \( V = \{ v_i \} \) is a set of graph nodes \( i = 1,2, \ldots, n \), every node matches one concept from the ontology; \( E = \{ e_j \} \) is a set of graph edges \( j = 1,2, \ldots, m \). The edge between nodes \( v \) and \( w \) is defined as a symmetric relation \( \{ v, w \} \).

The incidental matrix \( A = [a_{ij}] \) corresponded to the graph \( G_\Omega \) has the dimension \( n \times n \) (where \( n \) is the total number of ontology concepts). Each element of the \( A \) matrix can be computed from:

\[ a_{ij} = \begin{cases} 1, & \text{for } \{ i, j \} \in E, \\ 0, & \text{for } \{ i, j \} \notin E. \end{cases} \]

The transformation \( (\Omega \Rightarrow G_\Omega) \) from \( \Omega \) form to \( G_\Omega \) form is defined in the following way:

\[ a_{ij} = a_{ji} = \begin{cases} 1, & \text{if } \Pi(s_i, s_j), \\ 0, & \text{otherwise}. \end{cases} \]

The reverse transformation \( (G_\Omega \Rightarrow \Omega) \) from \( G_\Omega \) form to \( \Omega \) form is defined in the following way:

\[ \Pi(s_i, s_j) = \begin{cases} \emptyset, & \text{if } a_{ij} = 1, \\ \emptyset, & \text{otherwise}. \end{cases} \]

where \( \emptyset \) is any element from the set: \{IS_A, PART_OF, \emptyset\}.

5. Scheme of the concepts network creation algorithm

This section presents some propositions for the concepts network creation algorithm (CNCA) implementation. The CNCA algorithm, briefly described in Section 3.1, can be put into service as a procedure shown in Fig. 4.

In order to make the discussion of presented methods more strict, the Queuing Systems theory example is examined basing on [35,46]. In the preliminary study step of the procedure the unstructured reality is limited to the field of the discussed domain. The domain’s boundaries have to be founded and specified. It will make possible to create the ontology. In case of the technical and mathematic science the existing taxonomies (e.g. The Mathematics Subject Classification (MSC 2000)) provide some additional help, e.g. for the Queuing System theory the following association can be used:
> 60-xx Probability theory and stochastic processes
  > 60Kxx Special processes
    > 60K25 Queueing theory

According to [49] on the basis of the knowledge identification step an expert decides what knowledge and information are useful in the ontology creation. The potential knowledge sources can be: a related expert, corresponding ontologies, digital sources. After that, the knowledge is already indexed.

At the verbalization step, the expert’s knowledge is recorded. The verbalization begins the knowledge coding process. The knowledge coding process adapts knowledge to the requirements of the environment (e.g. LMS/LCMS systems). The knowledge in DL environment is transformed into a collection of concepts \( \phi \) with their matrixes \( T = [t_{ij}] \). Table 1 shows the concept depth matrix for the Queueing Systems concept. The concept is described by several other Queueing Systems theory concepts (like \( M/M/1 \) Queueing System or \( M/M/1/S \) Queueing System). The concepts can be

---

**Table 1**

Example depth matrix for the Queueing Systems concept

<table>
<thead>
<tr>
<th>Queueing Systems</th>
<th>( M/M/1 )</th>
<th>( M/M/m )</th>
<th>( M/M/m/m )</th>
<th>( M/M/1/S )</th>
<th>( M/D/1/C )</th>
<th>( G/G/\infty/\text{prt} )</th>
</tr>
</thead>
</table>
| Arrival pattern        | Markovian (Poisson) arriving process = \( M(P) \) | \( M(P) \) | \( M(P) \) | \( M(P) \) | \( M(P) \) | General ...
| Kind of servicing      | Markovian (Exponential) servicing process = \( M(E) \) | \( M(E) \) | \( M(E) \) | \( M(E) \) | Deterministic General ...
| Number of servers      | 1           | \( m \)    | \( m \)    | 1           | 1           | \( \infty \) ...
| Discipline of servicing| First come first serve (FCFS) | FCFS | FCFS | FCFS | FCFS | Priority ...
| Population             | \( \infty \) | \( \infty \) | \( \infty \) | \( \infty \) | Limited by \( C \) \( \infty \) | ...
| Queue capacity         | \( \infty \) | \( \infty \) | \( m \)    | Limited by \( S \) \( \infty \) | \( \infty \) | ...

---
differentiated basing on their attributes (like the kind of servicing or population).

The knowledge is structured during the graph representation phase. This phase creates a network (e.g. conceptual maps), which plays the role of the formal knowledge representation model. During this phase the relations are set. The example of CN for the Queuing Systems domain is shown in Fig. 5.

At the last step of the procedure, the CN is changed into a computer ready form. The computerization phase adjusts the structures of the DL model to the digital form. The knowledge manipulation language defines the model inputs and outputs. The final ontology can be expressed in one of the well accepted onto-languages (e.g. FLogic [7]).

6. Proposition for the didactic materials compilation algorithm

The repository with the finished learning objects will be relatively large. A learning object is a strongly personalized piece of educational content and because of this, each student should possess his/her own set of learning objects. Basing on this assumption, the repository has to store an enormous number of learning objects in order to pro-

![Diagram](image_url)
vide common access for many students—what is not feasible.

As [6,31] shows, the conceptual learning process bases on the concepts manipulation process. The proposed idea is to store knowledge about a specific domain in an ontological, generic form. By taking advantage of the concepts network (CN) structure, the knowledge can be stored in the repository on two layers: ontology and SCORM (Fig. 6). The ontology layer accumulates the knowledge in the CN form. Based on the CNCA algorithm the expert creates the knowledge domain model, where CN concept is specified and the depth matrix is established. On the second layer the knowledge is compiled into SCORM-like structure. Nowadays almost any LMS/LCMS system can import and operate with SCORM-like structures.

Every student has his/her own learning capabilities. Hence, the system should adapt the learning process to the individual student profile, which consists of information about the student’s learning style [40], previous knowledge and experience in the specific domain. Furthermore, the didactic content is affected by the learning objectives. Based on this information, LMS/LCMS system should adapt the CN to the learning object. The authors propose the didactic materials compilation algorithm (DMCA) to perform this task. The DMCA algorithm consists of several steps, which will be discussed below (more details about the algorithm can be found in [50]).

6.1. Concepts network dimension reduction

The basic generic ontology (expert’s ontology $\Omega_E$) generated by the expert in many cases is beyond the learning objectives. The teacher, who carries in mind the course intention, student’s profile and his/her characteristics, is obligated to set the course aim and objectives. Therefore, the ontology has to be limited in order to achieve the teacher’s expectations. In other words, the teacher’s ontology $\Omega_T$ enters the expert’s ontology $\Omega_E$. The result ontology $\Omega_{T(E)}$ (usually $\Omega_{T(E)} < \Omega_E$) covers the required student’s course knowledge, and is called a reduced expert’s ontology $\Omega_{T(E)}$.

6.2. Basic concepts selection using the student’s profile

A set of concepts $S_S$, represents the knowledge which has been learned and understood by the student in a specific domain and it is defined in the student’s profile. The learning process will be more effective if the relation between the acquired knowledge and the new one is emphasised. The common knowledge space allows student to do knowledge assimilation. Therefore, such shared concepts can play the role of starting points for the new learning process.

The idea is to integrate the $S_S$ set with the reduced expert’s ontology $\Omega_{T(E)}$. For this, the basic concept set should be obtained and the ontology $\Omega_{T(E)}$ has to be updated. If we move the discussion...
from the conceptual level to the mathematical one, the non-oriented ontological graph $G_\Omega = (V, E)$ should be converted into an oriented graph $\tilde{G}_\Omega$, where the orientation is made on the base of the $S_X$ set. The supported transformation $\Omega \Rightarrow G_\Omega$ (from the abstract ontological form into the mathematical one) is discussed in Section 4.2.

In the Queuing System example, the basic set of concepts $S_X$ has the following form $S_{S1} = \{\text{Erlang Distribution, General Distribution, Deterministic Distribution, Exponential/Poisson Distribution, First Come First Serve (FCFS), Last Come First Serve (LCFS), Random servicing, Priority Discipline}\}$. The oriented graph $\tilde{G}_\Omega$ provides additional information about learning process. The student (teacher) can deduct the direction of learning by simply following the oriented relations. In Fig. 7 an example of the oriented graph corresponds to Fig. 5 and $S_{S1}$ set is represented.

Student begins with the basic concepts, which are included in the $S_{S1}$ set, and then explores every concept according to the oriented relationship. The learning process finishes after the student has visited every concept (node of graph). The oriented relationship is made on the basis of the most likely direction the real learning process would occur.

![Fig. 7. Queuing System concepts network converted to an oriented graph using the $S_{S1}$ set.](image-url)
The result of the discussed step is the oriented graph:

\[ \tilde{G}_{\Omega} = (V, \tilde{E}), \]

where \( V = \{v_i\} \) is a set of nodes of the graph \((i = 1, 2, \ldots, n)\). Every node matches one concept from the ontology \( \Omega_{\Omega} \); \( \tilde{E} = \{e_j\} \) is a set of edges of the graph \((j = 1, 2, \ldots, m')\). The edge between nodes \( v \) and \( w \) is defined as a relation \((v, w)\).

The incidental matrix \( \tilde{A} = [a_{ij}] \) represents the oriented graph \( \tilde{G}_{\Omega} \). The matrix elements can be computed from:

\[
a_{ij} = \begin{cases} 
1, & \text{for } (i, j) \in \tilde{E}, \\
0, & \text{otherwise}.
\end{cases}
\]

6.3. Hierarchically ordered concepts network

Next phase of the DMCA algorithm transfers the oriented graph \( \tilde{G}_{\Omega} \) to the hierarchical ordered concepts network \( \tilde{G}_{\Omega} \). Each concept of the hierarchically ordered network is characterized by a corresponding level. The network nodes cannot be connected within the same level.

The concepts, which belong to the basic knowledge set \( S_{\Omega} \), are located on the first level (i.e. level 1 in Fig. 8). These concepts are called basic ones. The other levels consist of the concepts included in the course’s objective and the student should memorize them (levels 2–10 in Fig. 8). The hierarchical location gives information about relations between concepts.
The result of this phase is the hierarchically ordered concepts network which can be represented as a 3-tuple:
\[ \overrightarrow{G} = (L, \overrightarrow{V}, \overrightarrow{E}), \tag{11} \]
where \( L = \{l\}, \ l = 1, \ldots, l' \) is a set of levels; \( \overrightarrow{V} = \{v_k\} \) is a set of network nodes; \( l \) is the level index; \( k \) is the current node number on the level \( l \); \( \overrightarrow{E} = \{e_j\} \) is a set of graph edges \( (j = 1, 2, \ldots, m', \overrightarrow{E} = \overrightarrow{\hat{E}}) \).

The incidental matrix \( \overrightarrow{A} = \|a_{ij}\| \) reflects the hierarchical network \( A = \overrightarrow{A} \).

6.4. Transformation of the hierarchically ordered concepts network

The main aim of the DMCA algorithm is to create the sequence of learning objects. In the DL environment rooted in the SCORM standard (SCORM 1.2 or better) the content sequence is determined, based on the dedicated SCORM structure called Activity Tree (SCORM standard is discussed in detail in Section 7). The next step of the DMCA algorithm converts the hierarchical network \( \overrightarrow{G}_\Omega \) into a forest \( \hat{G}_\Omega \). The forest consists of a set of trees \( \hat{G}_\Omega \). Every tree can be considered as a SCORM’s Activity Tree. Finally, the set of Activity Trees makes up a course for LMS/LCMS systems.

The number of trees in the forest depends on the learning method selected in the initial phase. For the inductive learning style the number of trees in the forest is equal to the number of leaves in the hierarchical network. On the contrary, in the deductive learning style the number of trees in the forest is equal to the number of leafs. For the inductive learning style the number of trees in the forest is identical to the number of nodes. In the presented approach, learning objects accumulate similar concepts into one object. In the next step, the tree \( \hat{G}_\Omega \) will be covered with the portions, which will be converted further into a learning object. Portion is a sub-graph of the initial tree graph \( \hat{G}_\Omega \). All portions of the discussed tree are included in the portion set \( g_p = \{\hat{G}_p\} \), where \( i = 1, \ldots, p^* \).

The portion \( \hat{G}_p \) definition is the following:
\[ \hat{G}_p = (L_p, \hat{V}_p, \hat{E}_p), \tag{13} \]
where \( L_p = \{l\} \) is a level set for the portion \( \hat{G}_p \); \( L_p \subset L \); \( \hat{V}_p = \{v'_k\} \) is a node set for the portion \( \hat{G}_p \); \( \hat{V}_p \subset \hat{V} \); \( \hat{E}_p = \{e'_j\} \) is an edge set for the portion \( \hat{G}_p \).

The portion creation rules are the following:

1. Every portion \( \hat{G}_p \) is a sub-graph of \( \hat{G}_\Omega \), \( A_p \subset A \).
2. Sum of the sub-graphs \( \hat{G}_p \) gives the \( \hat{G}_\Omega \) graph, in which the portions unrelated to \( \hat{G}_\Omega \) cannot be allowed.
3. Every portion \( \hat{G}_p \) has a limited number of nodes.
4. Every portion has to contain at least one node common with another portion, i.e. \( \hat{G}_p \cap \hat{G}_p' \neq \emptyset \), where \( \{\hat{G}_p, \hat{G}_p'\} \in g_p, i \neq j \).
Structural elements

Input

Process

Waiting Queue

Server facility

Servicing Process

Queuing Systems

Arriving Population

Arriving Process

Deterministic Distribution

Erlang Distribution

General Distribution

Distribution Law

Kind of servicing

Queuing Discipline

Service Time Distribution

FCFS

LCLS

Random Servicing Priority Discipline

Exponential/Poisson distribution

Exponential Distribution

Deterministic

General

Erlang Distribution

General Distribution

Exponential/Poisson distribution

Fig. 9. Tree with portions ($P_{max} = 4$).
As we already mentioned above, learning objects accumulate similar concepts into one object. The question is—how many concepts should be incorporated into one learning object? The answer comes from the cognitive science. Mind research provides a useful memory model, which consist of the following memory types: sensor memory, short-term (ST) memory and long-term memory. The information from the environment is converted into an electrical signal by the sensor memory. After that, basing on the knowledge obtained from the unlimited long-term memory, the information is analyzed in the ST memory. Finally, some part of information is encoded in the ST memory to the form of a cognitive structure and then sent to the long-term memory as knowledge.

The most important type for our study is the ST memory, due to its limited capacity. The idea is to relate the “size” of the learning object to the ST memory capacity. The ST memory capacity is determined by the Miller Magic number \((7 \pm 2)\) [43], which was updated by Cowan (about 4) [14]. The newest research [19] brings up new estimation (about 3). Another research [61,65] is focused on the empirical experiments and localizes the ST memory as a tiny spot in the posterior parietal cortex.

An additional significant indicator is the unique nature of the ST capacity. The limitation is counted by chunks, not bites. A chunk [20,21,54] is a set of concepts which are strongly related to each other and, at the same time, weakly related with other concepts. In case of the concept definition from Section 4.1, the concept is considered as a chunk.

Therefore, the number of concepts in \(\hat{G}_p\) should not overcome \(P_{max}\), where \(P_{max}\) is the ST memory capacity set by a psychological survey. In Fig. 9 there is shown an example of a tree, which was obtained from the hierarchical ordered Queuing System concepts network (Fig. 8).

As a result, the discussed step generated a set \(g_p\), which is a sub-set of \(\hat{G}_p\) for every tree in the forest \(\hat{G}_D\).

### 6.6. Graph clustering

The last phase of the didactic materials compilation algorithm is the graph clustering. At this step any portions of any tree \(\hat{G}_D\) from the forest \(\hat{G}_D\) are integrated into the form of a consolidated node \(W_s\). The consolidated node has a name, which comes for example from the largest integrated concept. The clustered graph \(S\) has the following description:

\[
S = (N_s, W_s, E_s),
\]

where \(N = \{n\}\) is a set of levels for the clustered graph \(S\); \(W_s = \{w_i\}\) is a set of nodes for the clustered graph \(S\), \(i = 1, \ldots, \text{card}(g_p)\); \(E_s = \{e_i\}\) is a set of edges for the clustered graph \(S\).

The clustered graph \(S\) can be converted into the SCORM activity tree structure. The sequence of learning objects in the SCORM nomenclature is

---

![Fig. 10. Clustered graph.](image-url)
called an activity tree. The clustered graph obtained from the tree (Fig. 9) is demonstrated in Fig. 10. The clustered graph is converted to a SCORM activity tree structure (shown in Fig. 11). Basing on the activity tree, the LMS/LCMS systems generate a learning objects sequence. In general, LMS/LCMS systems use a pre-order strategy. The sequence in Fig. 11 is as follows: A, B, BA, BAA, BB, BBA, BC, BCA, BCAA, BCB, BCBA, BCC, BCCA. Moreover, the SCORM standard can use a script language in order to prevent repetition (e.g. Distributed Law in Fig. 10). The build-in script language gives advanced sequence control options. For instance the following rule \textit{IF node\_4.state = completed THEN go.previous} can be applied.

7. SCORM framework

The SCORM layer model consists of four conceptual layers (Fig. 12). At the bottom is the data layer (1). The data layer includes all data files, which are used to the concept’s metaphor creation. Furthermore, the data layer contains all meta-data files used at higher layers of the SCORM model (i.e. as a part of IEEE LOM declaration for the Sharable Content Object). The second layer (2) covers the structures intended to data integration. There is a number of SCORM integrated structures: Sharable Content Object, Asset, Content Aggregation. Each of them has its own, special role. The structure Content Aggregation is the most important one due to the integration features. The Sharable Content Object and Asset can be integrated into a learning object-like structure. The third layer (3) is a sequence layer. On this level the transformation Content Aggregation $\rightarrow$ Activity Tree occurs. Simultaneously, the transformation Asset, Sharable Content Object $\rightarrow$ Learning Activity takes place and the learning object content is made. On the highest layer (4) the mechanisms of LMS/LCMS system create the learning object basing on the activity tree analysis.

The goal of the proposed algorithms is to enhance the SCORM standard by adding two additional layers (X, Y). The concepts network creation algorithm (CNCA) is implemented between layer 1 and 2 (layer X in Fig. 13). The CNCA layer works with the ontology and concept structures. Concepts can be seen as Sharable Content Objects. The matrix of concept depth

\begin{center}
\begin{tabular}{|c|c|}
\hline
Layer & Data \hline
1 & Learning Activity, Activity Tree \hline
2 & Sharable Content Object, Asset, Content Aggregation \hline
3 & Learning Activity \hline
4 & Learning Object \hline
\end{tabular}
\end{center}
expressed in XML is adapted to the SCORM standard. Moreover, ontology can be considered as a generic form of the SCORM Content Aggregation.

Another proposed algorithm i.e. didactic materials compilation algorithm (DMCA) enters the second additional layer. This layer is placed between the second and third layer. The main task of layer Y is to adapt the concepts network (CN) to the requirements of SCORM sequencing theory. Based on the proposed algorithm the network ontology structure is transformed to the forest.

Fig. 13. Enhanced SCORM layer model.

Each of the forest trees has the nature of an activity tree and the tree nodes are the learning activities.

8. Application

The quality issue is a difficult problem, especially in the human-centered technology [52]. Around the world the quality research projects are running in order to find a methodology or procedures that allow achieving a high level of quality. One of the biggest issues is the quality in distance learning [42,45]. Therefore, a deep investigation is required. The authors have done some research in the framework of the European eQuality project [2].

The eQuality project objectives are to produce a common reference methodological framework for Quality in ODL in Europe, taking into account cultural differences, experiences, needs and best practices, joined with the ongoing normalization work. The project incorporates several European universities: European University Pole of Montpellier and Languedoc-Roussillon (France), University Montpellier 2 (France), Open University of Catalonia (Spain), University of Tampere.

Fig. 14. Architecture of the on-line simulator.
The algorithms (CNCA, DNCA) presented in this paper are being successfully applied in a training session organized to provide feedback upon the eQuality project methodology and tools. The training session is set in the Simulation Web Portal (SWP) environment. The SWP consist of four main modules/sections: simulation principles, on-line tests, on-line simulation, personal files management [37]. The first section is a simple textbook on computer simulations. Its aim is to introduce students to the computer simulation principles. The second section “on-line tests” is a set of exercises where students can verify their knowledge. In order to solve some of the exercises, student is asked to build a simple simulation model, carry out a simulation experiment and give its results. These tasks can be done in the next section “on-line simulation”. The On-line Simulator (Fig. 14) is based on the SIMAN language compiler from Rockwell/C213s ARENA simulation package. This system is highly specialized in serving educational purposes and was designed to demonstrate the abilities and advantages of computer simulation. In the last module, “Personal files management”, a registered user can save and open simulations’ results and parameters of all models developed in the “on-line simulation” section.

9. Summary

The proposed approach provides a high quality of interaction between all participants of the distance learning process, namely:

• shares more precisely the functions of an expert of the subject domain and the teacher which prepares methodical materials for DL;
• allows to carry out an exchange between the student and the teacher at the level of knowledge.

The model of knowledge representation and transfer can serve as a platform for cooperation between the teacher and the expert. The new teacher’s role in DL provokes a more creative approach to the analysis of the conceptual scheme of the subject domain and the choice of a teaching method depending on the objectives of learning and the initial knowledge of students.

The developed algorithms allow uniting the knowledge of the subject domain and methods of teaching, thus providing high quality methodical materials. Apart from that, algorithms can also be used as a methodical toolkit for the teacher, working in both traditional and distance learning systems.

Acknowledgement

This research has been supported by the eQuality Socrates/Minerva project no. 110231-CP-1-2003-1-MINERVA-MPP.

References


[43] G.A. Miller, The magical number seven, plus or minus two: Some limits on our capacity for processing information, Psychological Review 63 (2) (1956) 81–97.