LEARNING OBJECT SEMANTIC DESCRIPTION
FOR ENHANCING REUSABILITY

Perla Señas and Norma Moroni

Laboratorio de Investigación y Desarrollo en Informática y Educación (LIDInE)
Instituto de Investigación en Ciencias y Tecnología Informática (IICTI)
Departamento de Ciencias e Ingeniería de la Computación
Universidad Nacional del Sur – Bahía Blanca – 8000 – Argentina

ABSTRACT
In the design of activities for Web-Based Education Systems, the concept of Didactic Object or Learning Object, has appeared as the axis of a new paradigm concerned with the reusability of contents and activities, due to its organization from a conceptual point of view, the use of metadata in known formats and the setting of new standards. In this article, the concept of reusability is analyzed within the context of the systems mentioned. In addition, Sowa-Style Conceptual Maps as well as Hypermedia Conceptual Maps are proposed as schemes capable of representing knowledge, since they allow for a clear visualization and tend to enable known automated functionalities accurately. The above mentioned resources tend to complete the information in metadata with regard to the semantic aspect.

Key Words: Web-Based Learning – Learning Object – Semantics – Reusability

1. INTRODUCTION
Web-Based Education has evolved in the last few years following an axis that consists in the so-called learning objects (LO), around which a set of technologies and standards are structured. Although there is no agreement in only one definition for the LO concept, it can be seen that in all definitions there is some kind of reference, whether implicit or explicit, to the reusability and reuse of objects, and thus becoming one of the most important potentials. “From an economic point of view, it is easy to build cost-benefit models to justify LO design, similar to a certain extent to the models that have justified component-based Software Engineering” [27].

The possibility of structuring new LOs from some other more elemental LOs and according to decisions made dynamically allows us to develop complex learning strategies from simpler ones. For this to be possible, we have to ensure a correct connection and compatibility among the different LOs and, thus, the use of standards for their description becomes essential. It is important to ensure a common metadata language that is wide and rich enough to express all the information needed that shall then allow us to make all compositions required. In the same way, a system based on rules is necessary in order to express the relations among LOs and to be able to build learning sequences or itineraries from them [33].

2. ON LEARNING OBJECTS
At present, the design of Web-Based Education Systems (WBES) is largely based on the use of LOs, essential pieces that are organized to create educational experiences. Such pieces, characterized as fine grain, are part of the courses structured pursuant to an algorithmic didactical planning (WebQuest). Closed WBES only contain computer LOs, while in open systems there could be different types of LOs.

The idea is that in order to create an educational experience and then make it available for users, we should have LOs, specially created or obtained in some of the existing stores, and develop them allowing for more complex educational resources. The use of elementary pieces and the possibility to assemble them freely to build with them aggregate models of a structure superior to the style of the pieces of a meccano is one of the most attractive features of this technology. Now, if in order to conform an educational experience we have to assemble LOs, some fundamental questions like the following may arise:

– LOs should be easily accessible and reused: they should be developed independently from the context in which they shall be used in a first stage. The ideal thing is that these LOs are built as normalized reusable components, and this turns into a benefit both for educational material developers and for those who assemble the different WebQuests.
– The structuring of the experience should respond to a didactic planning that takes into account the different individual characteristics of participants and that could be algorithmically represented.

Several attempts to develop standards on LO creation and use can be quoted; several organizations and institutions got involved in it, among which we can mention: the group LTSC from IEEE, the Advanced Distributed Learning (ADL) initiative from the US Defense Department, the Global Learning Consortium (IMS), the Aviation Industry Computer Based Training Committee (AICC), ISO (ISO-SC36), the Dublin Core Metadata Initiative (DCM); and several projects such as GESTALT, PROMETEU, ARIADNE, CEN-ISSS or GEM [1], [9], [35].

Learning Objects: Reuse and Reusability

The reuse of a LO is a fact that can be observed and that can take place within the organization itself or may involve several organizations. The reusability of a LO is one of its attributes and can be used as an a priori quality measure. “In the case of reusability – as it happens in Software Engineering – there are no precise measures, only indicators, that may or may not be confirmed by reuse rates a posteriori, without changing at all the potential reusability” [27].
The reusability of a LO is a concept that mainly entails three aspects: format, interpretation, and adequacy from a didactic point of view. Current standards and specifications cover the first aspect, however, in the other two aspects, improvements and new concepts are necessary. On the other hand, there are different points of view – including value and quality perspectives –, that justify the extra effort required by metadata production for LOs. Course design within the framework of the LO paradigm is clearly oriented by reuse. LO design and technical-pedagogical description and organization criteria are based on this reuse concept both for the authors and for those who organize them for WBE.

A LO should tend to reuse, this is an essential and intrinsic property of its characterization. One way or another, this aspect appears in the definitions of the term that are currently in use. For example, for Wiley a LO is “any digital resource that can be reused in order to supply learning” [21]; for Polsani, a LO is “an independent, self-contained learning content unit that is ready to be reused in multiple learning contexts” [21]. Obviously, from an economic perspective, “the repeated use is the source of value and scale economy in the case of content suppliers” [27].

It is important to take into account the fact that since WBE is becoming more and more widespread and since this trend is markedly positive, LO reuse shall depend on the ease of LO access and assembly; it shall not depend only on its content but also on what is known as megadata, specially if we think that in the near future LO access shall be achieved almost exclusively through software tools. Metadata are specifications that allow us to find the LOs that we need. However, those descriptions should be mainly machine-oriented instead of human reading-oriented. “To think in large scale reuse without specialized software mediation is to lose the perspective of the phenomenon that we intend to characterize” [25]. If we look for reusability, we have to think in metadata that go beyond a LOM-compatible register (IEEE, 2002) or a SCORM-compatible package [27]. It is true that LOM or SCORM metadata creation is useful, however, this does not guarantee reusability, since they allow for metadata that, although complete and correct, are not necessarily capable of the automated processing needed if we think in the universalization of WBE. It is well known that not all formats allowing for human processing necessarily allow for computer management. Although some formats – such as SCORM – allow for content exchange among platforms qualified for WBE, this is no more than a file passage, i.e. just a technical reuse. Obviously, since the exchange is possible, the format supplies reusability, although we can have the case of a SCORM content reused just a few times because it is semantically too general or too specific for the educational experience in question. That is, adhering to LOM or SCORM, LOs with a high degree of reusability can be achieved, although this is not a sufficient condition, since in order to achieve it, it is necessary for metadata to have formats adequate for their automated processing. One example in this regard is the LO “design by contract” that does its best to supply a clear metadata semantics so that software tools can select and combine such objects. These are real reuse tasks that, to the present, some educational experience designers and many tutors and professors still solve in a homemade fashion – i.e. cut and paste. These are the major reusability aspects to be considered:

- A technical format aspect that implies that materials should be formatted according to certain rules and conventions. There has been great progress with current standards in this regard.
- A technical interpretation aspect regarding the fact that reused metadata should allow to enable, in an automated way and with accuracy, certain known functionalities. LOM is not enough in this area, but it can be extended by means of special techniques and practices.
- An instructional design aspect, so that content design and its granularity is reuse-oriented, thinking in possible environments of future uses. On this topic, a concept characterization draft has been proposed in which reusability in different LO educational contexts is related to total reusability [28].

Learning Objects Design and Reusability Evaluation

The LO paradigm is more valuable than other existing approaches for the design of contents and educational activities. We can highlight the following items:

From the meaning of a value generation: value can refer to elements of an economic nature or to the service capacity, among others. If expressed in terms of cost-benefit ratio it would be the value of the acquisition or production of the LOs needed for the educational experience vs. the increase in competencies and/or knowledge of recipients. Didactic activities in the organizational context are part of an adaptation cycle, therefore the value is conceptualized as a final increment in the organization’s competitive capacity. In addition, scale economy that can be reached by standardized LO manufacturers will, for sure, lead to a reduction in production global costs.

- From the technical conformity perspective: for e-learning systems, current standards are the basis of the interoperability of contents and educational activities. Their capacity in this area is proved and specification organisms develop an intense activity in order to cover more conformity areas.
- From the point of view of pedagogical adequacy: for educational experience designers in search of LOs for a concrete situation, the availability and composition of automated search tools save them design time and increase their possibilities of finding a LO that is adequate to their concrete needs. It is worth mentioning in this item that effective reuse possibilities increase proportionally to the descriptive quality of metadata. It is important to notice that the use of structured and non-ambiguous interpretation metadata opens up a completely different scenario for the construction of search tools.
- From the quality derived from repetitive use: Framed within a research-action domain, contents and educational activities are evaluated and improved with practice, in this case with the repetitive use, i.e. permanent evaluation and repetitive experience allow to increase LO quality. There are some endeavors that add a meta-information dimension on their quality and adequacy, since this is a very important intrinsic value.

3. PROPOSAL FOR LO SEMANTIC DESCRIPTION

This proposal focuses on the use of Hypermedia Conceptual Mapping (HCM) and Sowa-Style Conceptual Mapping (CM³) as knowledge representation schemes [24]. When the computational limit on automated reasoning and its effect on knowledge representation are examined, we notice that individuals do not reason
correctly and with the same ease on the different representation languages. In addition, the degree of difficulty generally increases in parallel with the expressive power of languages. The knowledge representation scheme presented is flexible enough for human management and it is rigorous in order to perform automated reasoning.

On Hypermedia Conceptual Mapping and Sowa-Style Conceptual Mapping

HCM is based on Novak’s Conceptual Maps (CM) and add the flexibility and richness of hypermedia technology. In the educational field, both schemes have been successfully tested as powerful structures capable of contributing for significant learning construction in individuals. The value of the hypermedia resource is highlighted, not only in the operational aspect, but also in the fields related to perception and abstraction [39]. We noticed some failures when we tried to extend the model for automatic management on a semantic basis. For this reason, a CM extension is performed adding elements from Sowa’s Conceptual Graphs [31] [32]. Then, the CM and an architecture for representation of a knowledge base with capacity for automatic reasoning are defined.

Knowledge Base Scheme

HCMs are a successful representation among human agents, however, they are incomplete as a representation scheme in mixed learning environments –i.e. consisting in human agents and software. Such failures are centered in concept hierarchy, in the definition of classes and individuals and in managing relation aridity. In order to solve this, the model is enriched in the following way: a class scheme and proposition representation scheme are created [24]. The class scheme is a reticulate represented by means of a HCM based on Novak’s CM model. For proposition representation, there is a migration to a model based on Sowa’s CGs [32], for which the problems posed before are already solved. The CG model is chosen because it is intuitive and because of the simplicity of its notation, its visual impact, its capacity to be viewed and its underlying logics. CGs make up a strong basis for logical reasoning, the resulting relations and concepts can be used and the consistency kept. Thus, CM and an architecture for representation of a knowledge base with capacity for automatic reasoning are defined.

Class Scheme

Inheritance is a natural tool to represent knowledge in a taxonomically structured way. This organization guarantees that all members in a class may inherit the adequate properties, ensuring consistency with the class definition. With this strategy the size of the knowledge basis is reduced and the implementation of default values is allowed. Default values are simply inherited from the appropriate superclasses. A model capable of representing those hierarchies that allow a multiplicity of parent classes is more expressive. Although these multiple inheritance hierarchies may introduce difficulties in the definition of representation languages, benefits outnumber disadvantages [12]. Reticulates are a common form for the multiple inheritance case. A partial order in the set of classes is established, indicated by the symbol \( \subseteq \) (\( \subseteq \) represents class inclusion). Subclass and superclass concepts are defined and, since it is a reticulate, the classes may have multiple parents and multiple offsprings. However, each class pair must have a minimum common superclass and a maximum common subclass. The minimum common superclass of a class collection is the appropriate place to define the properties that are common only to those classes. In order to solve the problem that arises when there are classes that do not have natural common superclasses or subclasses, two special classes covering those functions are added. Thus, the following is achieved: the class \( \subseteq \) is a real reticulate. In this proposal, class hierarchy is represented by a HCM with two standard classes: a Universal class as superclass of all classes and an Absurd class as subclass of all classes. Following CM conventions, if \( C_2 \) is a subclass of \( C_1 \), \( C_1 \) appears in the representation at a level higher than \( C_2 \). Classes are linked through the relation is a, therefore it is necessary to explicitly draw an arrow as shown in Figure 1.

![Figure 1](image1.png)

For hierarchies of a large number of classes, the use of HCM marks an important difference in the operational field.

Proposition Representation

Each simple proposition is represented by means of a CM that is a directed finite graph characterized by:
- The map nodes represent concepts. Graphically, concepts are drawn as labeled ellipses. All relations are binary. The traditional representation for CM relations is kept, i.e. they are represented by arches labeled with the name of the relation.
- The nodes represent objects from the discourse universe; they may be concrete or abstract. Concrete concepts include both generic and specific concepts.
- Verbal propositions are represented in the following way: the verb concept is the root of CM that represents the proposition. For example, for the proposition The bear drinks water, the agent relation links the concept drinks with the concept bear and the object relation links the concept drinks with the concept water as can be seen in Figure 2.

![Figure 2](image2.png)
In noun propositions, the concept to which a property is associated is the root concept of the CM^5. For example, for the proposition *Blue bird* the associated CM^5 is Figure 3.

![Figure 3](image)

- In CM^5 representing propositions, following the CG proposal, each noun concept is a unique individual of a particular class. It could be a generic or specific individual. The notation for the different cases in the corresponding ellipses is the following:

<table>
<thead>
<tr>
<th>generic individual</th>
<th>without generic marker</th>
<th>&lt; Class name &gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>with generic marker</td>
<td>&lt; Class name &gt; : *</td>
</tr>
<tr>
<td>specific individual</td>
<td>with name</td>
<td>&lt;Class name &gt; :</td>
</tr>
<tr>
<td></td>
<td>with marker</td>
<td>&lt;Class name &gt; : # &lt;Individual number &gt;</td>
</tr>
</tbody>
</table>

Each individual in the world of discourse has only one token associated, called numeric marker, that identifies it in full. This allows to indicate individuals that are specific but without a name. CM^5 allow for the use of name variables. These are represented by an asterisk followed by the variable’s name – for example, *X. This is useful when two different ellipses indicate the same individual, however it is a non-specified individual. The map on Figure 4 represents the statement *The kid rests his forehead on his knees.* Although there is no specification to which kid the proposition refers to, the variable *X indicates that the forehead and knees belong to the same kid. It also allows to have propositional nodes to represent subordinate or coordinate propositions. In such way, in addition to using the CM^5 to define relations among objects of the world, we can also define relations among propositions.

![Figure 4](image)

- A propositional node is represented as a map node that is labeled with a CM^5 representing a proposition, i.e. it is indicated as an ellipse containing another CM^5. For example, the sentence *John thinks that the bird is blue* is represented by the CM^5 shown in Figure 5. In this case, *thinks* is a relation that takes a proposition as argument. Each CM^5 represents a simple proposition. CM^5 may be arbitrarily complex, however they are always finite. A typical knowledge base shall contain a certain number of these maps, in addition to the HCM representing the class scheme. Propositional concepts can be used with appropriate relations to represent knowledge about propositions. Thus, we show how CM^5 with propositional nodes can be used to express the modal concepts for knowledge and belief.

![Figure 5](image)

**Creation of New CM^5**

In order to create new CM^5 from existing CM^5, operations that allow to deal with generalization and specialization – canonic formation rules – and logic operations are included. In the first group, there are operations such as copy, restrict, join, and simplify and, in the second, negation, conjunction and disjunction.

**Canonic Formation Rules**

Given CM^5 m_1 and m_2, the application of each of the rules gives as a result a new CM^5 as shown in Figure 6. The restriction rule can be used to make a correspondence between two concepts appear and, thus, then the join rule can be applied. The rules restrict and join together allow to implement inheritance. Replacement of a generic marker by an individual one implements the inheritance of one of the properties of the class to an individual. For example, in CM^5 m_3, Professor John inherits the Argentine nationality property originally defined in m_1 for a generic individual of the professor class. Replacement of the label of a class by the label of a subclass defines the inheritance between a class and a subclass. It is the case of the property inherited by the professor subclass in m_4 from the person class in m_2.

In addition, joining a CM^5 with another and restricting certain concepts, the inheritance of a variety of properties can be implemented. Since CM^5 are based on the CG model, matching and restriction can also be applied to them in order to implement plausible suppositions that play a major role in common language understanding, for example the sentence *Mary and Thomas went out together to eat pizza* can be modeled with CM^5.

As in the case of CG, CM^5 restriction and matching are specialization rules. They define a partial order over the set of derivable CM^5. If a CM^5 m_1 is an specialization of m_2 then we can say that m_2 is a generalization of m_1. As mentioned by Luger, generalization hierarchies are important in knowledge representation; they are used in many learning methods, together with the provision of bases for inheritance and other common sense reasoning schemes. Obviously, these are not inference rules, they do not guarantee that from true CM^5 shall always derive true
CM^5. In the restriction of m3 map of Figure 6, the result could not be true, for example if John is not a professor. Another representative example of non-preservation of the true is constituted by m6 in the same figure, since the professor that is in the office could be a different person from the professor correcting the exam.

These operations do not preserve the truth but hold the important property of preserving the significance condition, i.e., the canonic formation rules do not allow to form nonsense CM^5 from others that do have sense, thus becoming an important property. “Although it does not sound as inference rules, canonic formation rules are the bases for many plausible reasoning carried out in understanding of the natural language and common sense reasoning” [12].

Logical Operations

− Negation: The existence of propositional nodes in CM^5 makes it possible to easily implement a proposition’s negation. An operation called neg taking a propositional concept as argument and affirming that concept as false is defined. For graphical representation, the proposition that we try to negate is shown as a propositional node, and, in order to establish the negation, a fictitious node is used from which the neg relation goes to the propositional node. The use of such fictitious node is to the sole effect of treating the neg operation as binary.
− Conjunction: CM^5 representing disjunctive assertions can be formed. If each of the CM^5 representing the propositions to coordinate has a root node, the and relation can be established by linking both root nodes; otherwise, each of the propositions to be linked can be represented by means of a propositional node and then they can be related with and.
− Disjunction: According to the rules of logics, CM^5 representing disjunctive assertions can be formed by using negation and conjunction. In order to simplify this, an or relation can also be defined which takes two propositions and represents its disjunction in the same way as the conjunction is represented.

Variable Quantification

− It is assumed that in CM^5 generic concepts are existentially quantified. For example in the case of the CM^5 of Figure 3, the generic concept bird represents an existentially quantified variable. This CM^5 corresponds to the logical expression:

$\exists X \forall Y \left( \text{bird}(X) \land \text{color}(X,Y) \land \text{blue}(Y) \right)$

− Universal quantification can be represented by means of the use of negation and existential quantification. For example, for the CM^5 representing the negation for the proposition The bird is yellow we have the following logical expression:

$\forall X \forall Y \left( \neg \left( \text{bird}(X) \land \text{color}(X,Y) \land \text{yellow}(Y) \right) \right)$

A CM^5 that refers to a particular individual, for example the one represented by the proposition Simon bear is brown, corresponds to the following expression of the predicate calculus:

$\exists X_1 \left( \text{bear}(X_1) \land \text{color}(X_1,X) \land \text{brown}(X) \right)$

CM^5’s Expressive Power

As suggested by the previous examples, there is a direct correspondence from the CM^5 towards the predicate calculus notation. CM^5 are equivalent to the predicate calculus in its expressive power. The following algorithm allows us to obtain the predicate calculus expression equivalent to a given CM^5.

**Algorithm: Equivalent Logical Expression**

**Input:** MC^5 m

**Output:** predicate calculus expression equivalent to m

**Steps:**

1. Assign only one variable $X_1, \ldots, X_n$ to each of the n generic concepts in m.

2. Assign only one constant to each individual concept in m. This constant can simply be the name or marker used to indicate the concept’s referent.

3. Represent each concept node by a unary predicate with the same name of this node’s type of which the argument is the variable or constant assigned to that node.

4. Represent each conceptual relation in m as a binary predicate of which the name is the same as the name of the relation. This allows for each predicate argument to be the variable or constant assigned to the corresponding concept node linked to such relation.

5. Take the conjunction of all atomic sentences formed in items 3 and 4. This is the body of the expression of the predicate calculus. All variables in the expression are existentially quantified.
It is important to point out that although CM$^5$, as well as CG, can be reformulated using the predicate calculus syntax, they support a number of special purpose inference mechanisms such as matching and restricting that are normally part of the predicate calculus.

3. CONCLUSIONS

LO-based WBES paradigm emphasizes content and learning-oriented activity reusability. It is expected that LOs could be found, visualized, and added in order to be able to build educational experiences within the e-learning framework and, based on this paradigm, that this could be done by didactic designers and teachers with no specific education in the area of computing. In this article, we have dealt with the concept of reusability in the context of LO-based WBES and Sowa Style Conceptual Mapping and Hypermedia Conceptual Mapping are proposed as possible schemes for knowledge representation. Their potential for a clear visualization and for enabling automated functionalities in an accurate way, such as semantic-oriented searches or composition from the conceptual are shown. The above mentioned schemes are proposed to complete the information of current metadata with regard to the semantic aspect.

4. REFERENCES


[29] Sicilia, M. Reusability and reuse of learning objects: myths, realities and possibilities. www.um.es/ead/red/M2/


