Structuring and merging Distributed Content

Luca Stefanutti, Dietrich Albert, Cord Hockemeyer

Structuring and Merging Distributed Content

Luca Stefanutti, Dietrich Albert, Cord Hockemeyer
Department of Psychology, University of Graz
Universitätsplatz 2/III, 8010 Graz, AT
luca.stefanutti, dietrich.albert, cord.hockemeyer@uni-graz.at

A flexible approach for structuring and merging distributed learning object is presented. At the basis of this approach there is a formal representation of a learning object, called *attribute structure*. Attribute structures are labeled directed graphs representing structured information on the learning objects. When two or more learning objects are merged, the corresponding attribute structures are unified, and the unified structure is attached to the resulting learning object.

Keywords: distributed learning objects, knowledge structures, structuring content

1. INTRODUCTION

In order to decide which object comes next in presenting a collection of learning objects to a learner, one might establish some order. Given a set O of learning objects, a *surmise relation* on O is any partial order '$\preccurlyeq$' on the learning objects in O. The interpretation of '$\preccurlyeq$' is that, given any two learning objects o and o' in O, $o \preccurlyeq o'$ holds if a learner who masters o' also masters o. The concept of a surmise relation was introduced by [5] as one of the fundamental concepts of a theoretical framework called Knowledge Space Theory. According to this theory the *knowledge state* of a learner is the subset K of all learning objects in O that (s)he masters. A subset $K \subseteq Q$ is said to be a knowledge state of the surmise relation '$\preccurlyeq$' if $o \in K$ and $o' \preccurlyeq o$ implies $o' \in K$ for all learning objects $o, o' \in O$. Then the collection $\mathcal{K}$ of all knowledge states of '$\preccurlyeq$' is called the *knowledge space* derived from '$\preccurlyeq$'. Knowledge spaces are used for representing and assessing the knowledge state of a learner (see, e.g.,[5] and [6] in this connection).

The construction of a surmise relation may follow different approaches. After a brief presentation of an existing approach based on vectors of components of a learning object, we extend this approach to a more flexible representation called *attribute structure* [2]. The mathematical properties of attribute structures make it possible to compare distributed learning objects in terms of how much informative and how much demanding they are.

2. THE COMPONENT APPROACH

According to the *component* approach [1, 7], every content object o in O is equipped with an ordered n-tuple $A = \langle a_1, a_2, \ldots, a_n \rangle$ of *attributes* where the length n of the attribute n-tuple A is fixed for all objects. Each attribute a_i in A comes from a corresponding attribute set C_i called the i-th *component* of the content object. In this sense, given a collection $\mathcal{C} = \{C_1, C_2, \ldots, C_n\}$ of disjoint attribute sets (or components), each object $o \in O$ is equipped with an element of the Cartesian product $\mathcal{P} = C_1 \times C_2 \times \cdots \times C_n$. Usually each component C_i is equipped with a partial order '$\leqslant_i$' so that $\langle C_i, \leqslant_i \rangle$ is in fact a partially ordered set of attributes. The partial order '$\leqslant_i$' is interpreted in the following way: for $a, b \in C_i$, if $a \leqslant_i b$ then a learning object characterized by attribute a is *less demanding* than a learning object characterized by attribute b. To give a simple example, it might be postulated that 'computations involving integer numbers' (attribute a) are less demanding than 'computations involving rational numbers' (attribute b). A natural order $\leqslant$ on the elements in $\mathcal{P}$, the so-called coordinatewise order [4], is derived from the n partial orders '$\leqslant_i$' by

$$\langle x_1, x_2, \ldots, x_n \rangle \leqslant \langle y_1, y_2, \ldots, y_n \rangle \iff \forall i : x_i \leqslant_i y_i$$

4th International LeGE-WG Workshop:
Towards a European Learning Grid Infrastructure
Progressing with a European Learning Grid

1

If $f : O \to \mathcal{P}$ is a mapping assigning an attribute n-tuple to each learning object, then a surmise relation '$\preccurlyeq$' on the learning objects is established by

$$o \preccurlyeq o' \iff f(o) \leqslant f(o')$$

for all $o, o' \in O$. The mapping f can easily be established even when the learning objects are distributed (see, e.g., [8]).

3. ATTRIBUTE STRUCTURES

An *attribute structure* is used to represent structured information on a learning object or an asset and in this sense it represents an extension of the attribute n-tupel discussed in Section 2. From a mathematical standpoint attribute structures correspond to the feature structures introduced by [3] in computational linguistics. Let C be a set of components and A a collection of attributes, with $A \cap C = \emptyset$. An attribute structure is a labeled directed graph $\mathcal{A} = \langle Q, \bar{q}, \alpha, \eta \rangle$ where:

- Q is a set of nodes of the graph;
- $\bar{q} \in Q$ is the *root node* of the graph;
- $\alpha : Q \to A$ is a partial function assigning attributes to some of the nodes;
- $\eta : Q \times C \to Q$ is a partial function specifying the edges of the graph.

As an example, let

$$C' = \{\text{picture}, \text{topic}, \text{subtopic}, \text{text}, \text{language}\}$$

be a set of components, and

$$A' = \{\text{PICTURE1}, \text{TEXT1}, \text{ENGLISH}, \text{MATH}, \text{MATRIX\_INVERSION}\}$$

be a collection of attributes. Suppose moreover that a simple learning object is described by the asset structure $\mathcal{A}_1 = \langle Q_1, \bar{q}_1, \alpha_1, \gamma_1 \rangle$, where Q_1 is the set of nodes, $\bar{q}_1$ is the root node, and α_1 and γ_1 are defined as follows: $\alpha_1(0)$ is not defined, $\alpha_1(1) = \text{PICTURE1}$, $\alpha_1(2) = \text{TEXT1}$, $\alpha_1(3) = \text{MATH}$, $\alpha_1(4) = \text{ENGLISH}$, and $\alpha_1(5) = \text{MATRIX\_INVERSION}$; $\eta_1(0, \text{picture}) = 1$, $\eta_1(0, \text{text}) = 2$, $\eta_1(0, \text{topic}) = 3$, $\eta_1(1, \text{topic}) = 3$, $\eta_1(2, \text{topic}) = 3$, $\eta_1(2, \text{language}) = 4$, and $\eta_1(3, \text{subtopic}) = 5$. The digraph representing this attribute structure is displayed in Figure 1. The structure $\mathcal{A}_1$ describes a very simple learning object containing a picure along with some text explanation. Both text and picture have MATH as topic and MATRIX_INVERSION as subtopic. The root node of the structure is node 0 and it can be easily checked from the figure that each
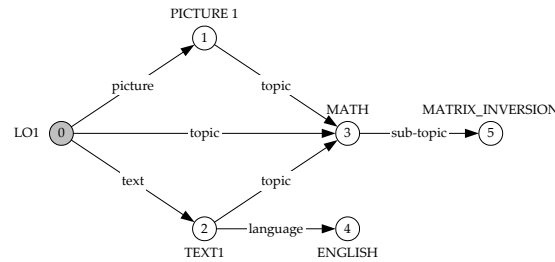
FIGURE 1: The attribute structure $\mathcal{A}_1$ describes a simple learning object on 'matrix algebra'

node can be reached from this node following some path in the graph. The root node is the entry node in the asset structure of the learning object, and the edges departing from this node specify the main components of the learning object itself. Thus, in our example, the learning object represented by $\mathcal{A}_1$ is defined by three different components: picture, text and topic. The values of these three components are the attributes given by $\alpha_1(\eta_1(0, \text{picture})) = \text{PICTURE1}$, $\alpha_1(\eta_1(0, \text{text})) = \text{DESCRIPTION}$, $\alpha_1(\eta_1(0, \text{topic})) = \text{MATH}$.

Observe, for instance that node 5 can be reached from node 0 following the path $\langle \text{picture}, \text{topic}, \text{subtopic} \rangle$. The fact that, in this example, each node is reachable from the root node through some path is not a coincidence. It is explicitly required that every node in an attribute structure be reachable from the root node.

4th International LeGE-WG Workshop:
Towards a European Learning Grid Infrastructure
Progressing with a European Learning Grid

2

4. COMPARING AND COMBINING ATTRIBUTE STRUCTURES

Attribute structures can be compared one another. Informally, an attribute structure $\mathcal{A}$ *subsumes* another attribute structure $\mathcal{A}'$ (denoted by $\mathcal{A} \sqsubseteq \mathcal{A}'$) if $\mathcal{A}'$ contains at least the same information as $\mathcal{A}$. In this sense an attribute structure can be thought as a class of learning objects (the class of all learning objects represented by that structure), and '$\sqsubseteq$' can be regarded as a partial order on such classes. Formally, an attribute structure $\mathcal{A} = \langle Q, \bar{q}, \alpha, \eta \rangle$ subsumes another attribute structure $\mathcal{A}' = \langle Q', \bar{q}', \alpha', \eta' \rangle$ if there exists a mapping $h : Q \to Q'$ fulfilling the three conditions

(1) $h(\bar{q}) = \bar{q}'$;

(2) for all $q \in Q$ and all $c \in C$, if $\eta(q, c)$ is defined then $h(\eta(q, c)) = \eta'(h(q), c)$;

(3) for all $q \in Q$, if $\alpha(q)$ is defined then $\alpha(q) = \alpha'(h(q))$.

In Section 2 the attribute sets were assumed to be partially ordered according to pedagogical criteria and/or cognitive demands. Similarly we assume now that a partial order '$\leqslant$' is defined on the set A of attributes so that, given two attributes $a, b \in A$, if $a \leqslant b$ then a learning object defined by attribute a is less demanding than a learning object defined by attribute b. Then the subsumption relation '$\sqsubseteq$' is made consistent with '$\leqslant$' if condition (3) is replaced by

(4) for all $q \in Q$, if $\alpha(q)$ is defined then $\alpha(q) \leqslant \alpha'(h(q))$.

According to this new definition, if $\mathcal{A} \sqsubseteq \mathcal{B}$ then $\mathcal{A}$ is either less informative than $\mathcal{B}$ or less demanding than $\mathcal{B}$ or both.

As an example consider the three attribute structures depicted in Figure 2. Assuming that $\mathrm{MATRIX\_PRODUCT} \leqslant \mathrm{MATRIX\_INVERSION}$, both mappings g and h fulfill conditions (1), (2) and (4), thus both attribute structures labeled by $\mathrm{LO2}$ and $\mathrm{LO3}$ subsume the attribute structure labeled by $\mathrm{LO1}$. However there is neither mapping from $\mathrm{LO2}$ to $\mathrm{LO3}$ fulfilling the subsumption conditions, nor the opposite, thus these last two structures are incomparable to each other. The
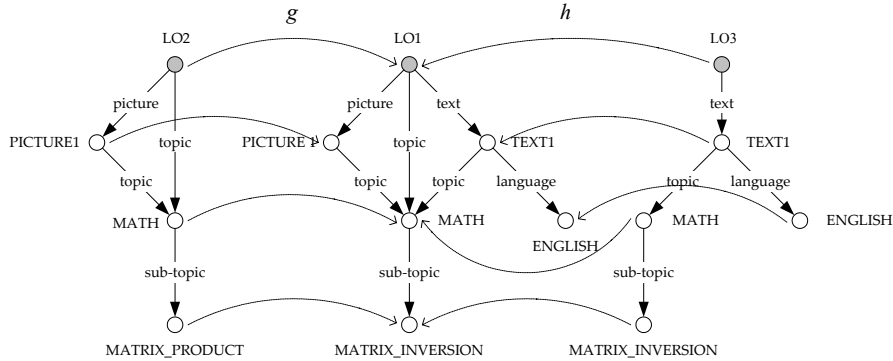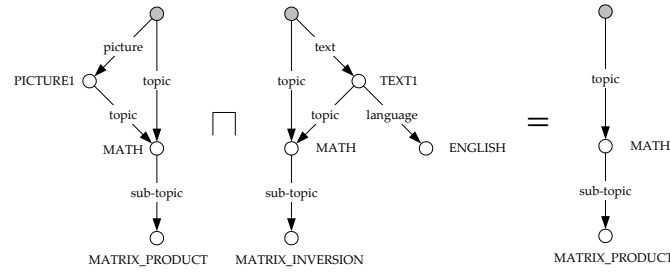

FIGURE 2: Both LO2 and LO3 subsume LO1. However LO2 and LO3 are incomparable.

derivation of a surmise relation for the learning objects parallels that established in section 2. If $s : o \mapsto s(o)$ is a mapping assigning an attribute structure to each learning object, then a surmise relation '$\preccurlyeq$' on the learning objects is derived by

$$o \preccurlyeq o' \iff s(o) \sqsubseteq s(o')$$

for all $o, o' \in O$.

Two binary operations are defined on attribute structures: *unification* and *generalization*. Mathematically, the unification of two attribute structures $\mathcal{A}$ and $\mathcal{B}$ (denoted by $\mathcal{A} \sqcup \mathcal{B}$), when exists,

4th International LeGE-WG Workshop:
Towards a European Learning Grid Infrastructure
Progressing with a European Learning Grid

3

FIGURE 3: Generalization of two asset structures

is the least upper bound of $\{\mathcal{A}, \mathcal{B}\}$ with respect to the subsumption relation. Dually, generalization (denoted by $\mathcal{A} \sqcap \mathcal{B}$) is the greatest lower bound. In particular, for any two attribute structures $\mathcal{A}$ and $\mathcal{B}$ it holds that

$$\mathcal{A} \sqsubseteq \mathcal{A} \sqcup \mathcal{B}, \quad \mathcal{B} \sqsubseteq \mathcal{A} \sqcup \mathcal{B}$$
$$\mathcal{A} \sqcap \mathcal{B} \sqsubseteq \mathcal{A}, \quad \mathcal{A} \sqcap \mathcal{B} \sqsubseteq \mathcal{B}$$

When two different learning objects are merged together, or when different assets are assembled into a single learning object, the corresponding attribute structures are unified, and the resulting attribute structure is assigned to the resulting learning object. On the other hand, the generalization operation is used to find the common structure of two or more learning objects or, stated another way, to classify learning objects. An example of the generalization operation applied to two attribute structures is shown in Figure 3. Here, the resulting structure shows that two learning objects have in common topic and $\mathrm{subtopic}$. Generalized attribute structures can also be used e.g. for searching a distributed environment for all learning objects whose structure is consistent with a certain 'template' (for instance to find out all learning objects that are 'problems' involving, as cognitive operation, 'recognition' rather than 'recall').

REFERENCES

[1] Albert D., Held T. (1999) Component-based knowledge spaces in problem solving and inductive reasoning. In D. Albert and J. Lukas (Eds.), *Knowledge Spaces. Theories, Empirical Research, Applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

[2] Albert D., Stefanutti L. (2003) Ordering and Combining Distributed Learning Objects through Skill Maps and Asset Structures. *Proceedings of the International Conference on Computers in Education (ICCE 2003).* Hong Kong, 2-5 December.

[3] Carpenter B. (1992). The logic of typed feature structures. *Cambridge Tracts in Theoretical Computer Science.* Cambridge University Press, Cambridge.

[4] Davey B.A., Priestley H.A. (2002). *Introduction to lattices and order. Second edition.* Cambridge University Press.

[5] Doignon J.-P., Falmagne J.-C. (1985). Spaces for the assessment of knowledge. *International Journal of Man-Machine Studies, 23*, 175–196.

[6] Doignon J.-P., Falmagne J.-C. (1999). *Knowledge Spaces.* Berlin: Springer-Verlag.

[7] Schrepp M., Held T., Albert D. (1999). Component-based construction of surmise relations for chess problems. In D. Albert and J. Lukas (Eds.), *Knowledge Spaces. Theories, Empirical Research, Applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

[8] Stefanutti L., Albert D., Hockemeyer C. (2003). Derivation of knowledge structures for distributed learning objects. *Proceedings of the 3rd International LeGE-WG Workshop*, 3rd December.

4th International LeGE-WG Workshop:
Towards a European Learning Grid Infrastructure
Progressing with a European Learning Grid

4