Merging domain knowledge and task analysis in an ontology

Alexander Staller*

Institut für Informatik, Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany

Current adaptive web-based educational systems rely on a subject domain ontology, specifying the domain concepts and the relations among them. In contrast, task analysis for instructional design is based on learning outcomes, including their taxonomic classification, information processing analysis, and prerequisite analysis. It is described how subject domain knowledge and task analysis results can be combined in an ontology, increasing didactic precision compared with purely concept-based models. Basic points with respect to modelling in the Web Ontology Language OWL are presented. Application areas include the Educational Semantic Web, metadata for reusable learning objects, and the Bologna process promoting a European higher education area.

Keywords ontology; domain knowledge; task analysis; learning outcomes; Educational Semantic Web

1. Introduction

The Educational Semantic Web is a vision for the further development of e-learning. Concept-based Adaptive Web-based Educational Systems are roots from which the Educational Semantic Web can evolve. These systems rely on a *subject domain ontology*, specifying the domain concepts and the relations among them. This ontology is used for annotating learning resources, building a learner model, and implementing adaptive behaviour [1].

However, from an instructional design perspective, domain concepts are not specific enough to guide learning, instruction, and assessment. Rather, task analysis for instructional design includes the taxonomic classification of *learning outcomes*. Further, there are established task analysis methods, such as information processing analysis and prerequisite analysis [2].

This paper is a proposal to combine subject domain knowledge and task analysis results in an OWL ontology. It is structured as follows: Section 2 presents information about task analysis for instructional design. In section 3 basic points with respect to merging domain knowledge and task analysis in OWL are described. Section 4 outlines potential application areas. Future work directions are provided in section 5.

2. Task analysis for instructional design

Task analysis is an important component of the instructional design process. Smith and Ragan [2] propose to formulate *learning goals* and to perform an *information processing analysis* and a *prerequisite analysis* for each learning goal. Information processing analysis specifies the mental steps that are performed to complete a task. Prerequisite analysis identifies for each step what the learner needs to be able to do in order to achieve the step. Instruction and assessment depend on the nature of the intended learning outcomes. So *classifying the learning outcomes* corresponding to the learning goal and the prerequisites is an essential part of task analysis. Further, *learning objectives* (i.e., performance objectives with observable behaviour) are formulated for the learning goal and the prerequisites. Finally, for each learning objective an *assessment instrument* needs to be specified.

For the classification of the learning goal and the prerequisites with respect to the intended learning outcomes, Smith and Ragan [2] use the taxonomy by Gagné [3]. There are many other taxonomies, each serving different instructional purposes. Anderson and Krathwohl [4] propose a revision of Bloom's well-known taxonomy of educational objectives in the cognitive domain, taking current findings of cog-

* Corresponding author: e-mail: staller@in.tum.de, Phone: +49 89 289-17354

nitive psychology into account. The revised taxonomy is based on the view that the structure of an objective (i.e., a learning outcome) in the cognitive domain has two dimensions: a *knowledge* dimension and a *cognitive process* dimension. Each dimension is divided into categories forming a two-dimensional taxonomy table. The main categories of the taxonomy are shown in Table 1. The knowledge types are divided further into subtypes (e.g., *Knowledge of classifications and categories* as a subtype of *Conceptual knowledge*). The categories of the cognitive process dimension represent increasing cognitive complexity and contain concrete processes. For example, the category *Understand* contains the process *exemplifying*.

Table 1 Main categories of the taxonomy by Anderson and Krathwohl [4].

Knowledge Dimension	**Cognitive Process Dimension**					
	Remember	*Understand*	*Apply*	*Analyze*	*Evaluate*	*Create*
Factual						
Conceptual						
Procedural						
Metacognitive						

3. Merging domain knowledge and task analysis in OWL

3.1 Consistency through modularisation

In the fields of didactics and psychology, there is a multitude of terms that are not used consistently. For example, Smith and Ragan [2] use the term “objective” in the sense of performance objectives specifying observable behaviour, while Anderson and Krathwohl [4] use it more generally to refer to intended student learning outcomes. Consistency in educational OWL ontologies can be accomplished through the following modularisation guideline: Self-contained didactic and psychological models (such as the task analysis model by Smith and Ragan [2] and the taxonomy by Anderson and Krathwohl [4]) are represented in separate OWL files. The terms of the model are transferred to the ontology in a traceable way by citing the reference in the literature. Thus, consistency within the OWL file is ensured. For building an ontology based on a specific didactic or psychological model, the corresponding OWL file merely needs to be imported. Then the namespace mechanism allows an unambiguous reference to the terminology of this model. In this paper the namespace prefix SR stands for Smith and Ragan and AK stands for Anderson and Krathwohl.

3.2 Modelling task analysis results

The analyses of information processing and prerequisites are usually performed graphically, lending themselves to formalisation. Figure 1 illustrates classes (shown as circles or ovals) and properties (shown as arrows) derived from the description by Smith and Ragan [2]. The individuals (shown as diamonds) are included for illustration purposes and represent results of task analysis for a specific learning goal (represented by the individual lg).

It is good design practice that transitive properties have non-transitive subproperties. For example, the prerequisite relation is transitive, so SR:has_direct_prerequisite is defined as a non-transitive subproperty of the transitive property SR:has_prerequisite (not shown in Fig. 1). In a larger instructional context, a network of individuals is formed, in which it is possible to retrieve direct prerequisites as well as all prerequisites resulting from transitivity. Of course there are other properties not shown in Fig. 1 (e.g., a datatype property for including a textual description of the task analysis results).

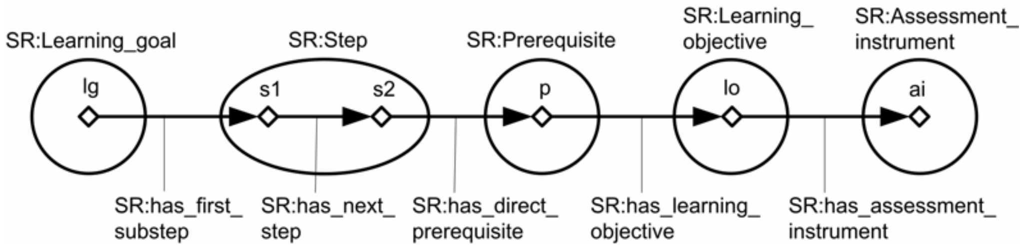

Fig. 1 *Representing results of task analysis according to Smith and Ragan [2].*

3.3 Modelling the classification of learning outcomes and incorporating domain knowledge

During task analysis the learning goal and the prerequisites are classified with respect to the intended learning outcomes. The modelling goal, however, is not only to represent taxonomic classification, but also to represent the relation between learning outcomes and subject domain knowledge. The revised two-dimensional taxonomy is the key for accomplishing this goal in an elegant way: The knowledge dimension actually provides the structure for a comprehensive subject domain ontology (excluding meta-cognitive knowledge, which is not restricted to a specific subject domain). Domain concepts belong to the knowledge type *Conceptual knowledge*, more precisely to one of its subtypes (e.g., *Knowledge of classifications and categories*). The categories of the taxonomy are transferred directly to OWL classes. Thus, domain concepts are represented as instances or subclasses of the classes representing the knowledge subtypes. In order to classify learning outcomes along the dimensions of the taxonomy, the properties AK:has_knowledge and AK:has_cognitive_process are provided. The fact that a learning outcome involves a specific domain concept can then be represented via the property AK:has_knowledge.

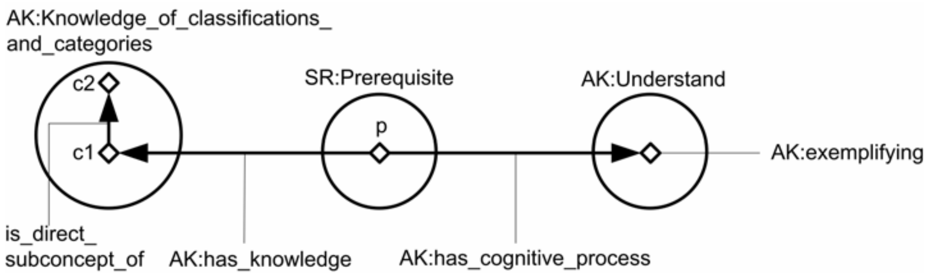

Fig. 2 *Representing a domain concept as the knowledge component of a learning outcome along with the taxonomic classification of the learning outcome.*

Figure 2 illustrates the representation of the following example: Prerequisite p (see Fig. 1) has *exemplifying* as the cognitive process component and the concept *c1* as the knowledge component. Since c1 is an instance of the class AK:Knowledge_of_classifications_and_categories, the taxonomic classification of the knowledge type is represented without any further effort. Concepts may be connected by many relations (e.g., the subconcept relation or *part-of*), represented by properties. But note that concepts do not stand in a prerequisite relation, which is represented at the level of task analysis results (see Fig. 1).

In an analogue way, domain knowledge outside conceptual knowledge is represented. For example, if a learning outcome involves the application of a specific technique (e.g., a proof technique in mathematics), an individual representing this technique is contained in the class AK:Knowledge_of_subject-specific_techniques_and_methods (a subclass of AK:Procedural_knowledge) and can be the value of AK:has_knowledge.

There are three species of OWL: OWL-Lite, OWL-DL (Description Logics), and OWL-Full. In contrast to OWL-Full, OWL-DL is a decidable fragment of First Order Logic and therefore amenable to automated reasoning. In the above example concepts are represented as individuals. In OWL, however, concepts are usually represented as classes. But using classes as values of the property AK:has_knowledge would lead outside OWL-DL, because properties connect individuals and in OWL-DL a class cannot be an individual. A possible way of staying in OWL-DL is to represent each domain concept by a class and a corresponding individual. In this way, two parallel conceptual structures are formed: In the first structure concepts are represented as classes. The subconcept (*is-a*) relation – one of the basic relations in subject domain ontologies – is easily modelled in this structure by building subclasses. In the second structure concepts are represented as individuals, and relations between concepts as properties (see Fig. 2 left). On the one hand, this approach increases the complexity of the domain ontology, on the other hand it is amenable to automated reasoning. DL reasoning has already been applied to the parallel conceptual structures, but details are beyond the scope of this paper.

4 Application areas

4.1 Educational Semantic Web

The Educational Semantic Web [1] is an important application area of the model. The reliance on concept-based domain ontologies has the following consequences (e.g., [5]): Firstly, learning and assessment material is connected to concepts. Secondly, the learner model is usually a numeric overlay on the concept structure, but the overlay value does not specify precisely what the learner can do with respect to the concept. Thirdly, prerequisite relations are defined between concepts.

In contrast, the model outlined in section 3 represents learning outcomes, having a cognitive process component in addition to domain knowledge. Further, there is a strict discrimination between relations among domain concepts (e.g., the subconcept relation) and relations among learning outcomes (e.g., the prerequisite relation). If task analysis as suggested by Smith and Ragan [2] is conducted completely, there are also individuals representing performance objectives and assessment instruments. This additional knowledge greatly increases didactic precision regarding the annotation of material, the information in the learner model, and the specification of prerequisite relations. For example, it is possible to represent that a student can recall the definition of a concept (category *Remember*), but has difficulties finding examples of the concept (category *Understand*). This knowledge is semantically richer than an overlay value for the concept. Thus, adaptation and personalisation can be more precise.

The effect of this increased precision is not restricted to adaptation and personalisation within a system. Representing learning outcomes also has implications for system interoperability, an important goal of the Educational Semantic Web. An ontology representing both task analysis results and domain concepts can serve as a reference which is more precise than a merely concept-based ontology. For example, a learner model in an Adaptive Web-based Educational System is highly system-specific, depending on the system's observations of the learner's actions. These observations are encoded in overlay values for domain concepts. However, another system has no information about the observations underlying these values. A learner model representing learning outcomes allows a more specific *system-independent* representation of the learner's abilities.

4.2 Metadata for reusable learning objects

Reuse of learning objects is aimed at reducing costs for the development and maintenance of learning content. However, there is still no satisfying solution because of the inherent conflict between context independence of reusable material and context necessary for effective learning and instruction [6]. Modularisation of content at different levels of granularity is important for solving this conflict. For example, the Autodesk Learning Object Content Model [7] is a hierarchy starting with raw content items and continuing via information objects and learning objects to lessons and courses. Information objects clarify an issue (e.g., a fact, concept, or procedure). Learning objects are composed of information objects and serve exactly one learning objective.

Annotation of the modular components with metadata is crucial for achieving reusability. The model presented in this paper allows the annotation of information objects *and* learning objects. Information objects clarify domain knowledge and are therefore annotated with individuals modelling subject domain knowledge. For example, a text defining a specific concept is annotated with the individual representing this concept in the class AK:Conceptual_knowledge. Learning objects are annotated with instances of the class SR:Learning_objective (see Fig. 1). With these annotations the information and learning objects are already embedded in an instructional context defined by the ontology. The ontology allows the inference of knowledge that is not specified directly as metadata. For example, if a learning object is annotated with an instance of the class SR:Learning_objective, the following may be inferred from the represented knowledge: an appropriate assessment instrument; the knowledge component of the learning outcome (e.g., a fact, concept, or procedure); information objects referring to this knowledge component; the cognitive process component of the learning outcome, hence its cognitive complexity; the (transitive) prerequisites of the learning outcome and corresponding information and learning objects.

A specific model of reusable learning objects [6] proposes that learning objects result from the combination of information objects and educational objects. Information objects are related to subject matter content, while educational objects include a learning objective and an educational activity. It is proposed to use a subject ontology for annotating the information objects and the taxonomy by Anderson and Krathwohl [4] for annotating the educational objects. So far, the model has been described only informally. The approach presented in this paper is the basis for a formal model in OWL.

4.3 The Bologna process

The Bologna process promotes a European higher education area. One of its cornerstones is the specification of learning outcomes, which are assumed to have highly beneficial implications for curriculum design, teaching, learning, assessment, quality assurance, and consistency across study modules and programs. An OWL ontology representing task analysis results and domain knowledge can be the basis for services supporting students, teaching staff, and accreditation agencies.

5 Conclusion and future work

This paper lays the foundation for the enrichment of subject domain ontologies using task analysis results, thereby increasing didactic precision compared to purely concept-based models.

This approach is going to be applied to online teacher education in informatics. A further research goal is the development of a tool that liberates the user from redundant work (e.g., the creation of parallel conceptual structures). It should also allow subject domain experts and instructional designers to conduct their analyses while the underlying OWL representation is created automatically.

Acknowledgements The helpful comments by Peter Hubwieser are gratefully acknowledged.

References

[1] L. Aroyo and D. Dicheva, Educational Technology & Society, **7** (4), 59 (2004).
[2] P.L. Smith and T.J. Ragan, Instructional design (3rd edition), Hoboken, NJ: Wiley (2005).
[3] R.M. Gagné, The conditions of learning (4th edition), New York: Holt, Rinehart & Wilson (1985).
[4] L.W. Anderson and D.R. Krathwohl (eds.), A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives, New York: Longman (2001).
[5] P. Brusilovsky and J. Vassileva, International Journal of Continuing Engineering Education and Lifelong Learning, **13** (1-2), 75 (2003).
[6] B. Krämer, Reusable learning objects: Let's give it another trial, Keynote at the EADTU (European Association of Distance Teaching Universities) Conference, Rome, Italy, November 10-11 (2005).
[7] MASIE Center, Making sense of learning specifications & standards: A decision maker's guide to their adoption (2nd edition), http://www.masie.com/standards/s3_2nd_edition.pdf (2003).