Bringing Relevant Topics to Foster Learning with the Aid of Ontologies

Elaine Harada Teixeira de Oliveira¹, Erika Handa Nozawa¹, Rosa Maria Vicari², Alberto Nogueira de Castro Júnior³

¹ Programa de Pós-Graduação em Informática, Universidade Federal do Amazonas (UFAM), Brazil
² Programa de Pós-Graduação em Computação, Universidade Federal do Rio Grande do Sul (UFRGS), Brazil
elaine@dcc.ufam.edu.br, erika.handa@gmail.com, rosa@inf.ufrgs.br, alberto@ufam.edu.br

Abstract. Ontologies are powerful structures that may help organize and share knowledge considering its contexts. This work addresses the issue of modeling the Japanese Language Proficiency Test (JLPT) domain through an ontology. The ontology was developed as an attempt to provide options for personalized learning oriented to context, based on some relevance principles. The most recent results come from the evaluation process in which the presented ontology was submitted to evaluators to assure its computationally consistency and adequacy to the Relevance Theory.

Keywords: Ontology, cognitive model, JLPT, Relevance Theory, context.

1 Introduction

Ontologies have been useful specifications to share conceptualization in a certain context [1]. They are especially useful to avoid misunderstandings and ambiguities. With them, experts may create a collaborative structure that may organize, share and make it a consensual knowledge. Ontologies are not restricted to specific area; they may be developed and applied to any area of knowledge.

The idea of using ontologies on the educational domain is not new, as we can see in [2, 3]. The use of domain ontologies for educational environments requires a semantic mapping of the knowledge from a taxonomy of formalized goals and concepts (information structure), developing a cognitive model that carries, implicitly, heuristics. For heuristics, we think on the understanding of the strategies (techniques) that comprise the individual who is guided by his pre-determined goals, purposes and values, taking into account what seems significant in order to get to new discoveries and learning.

For that, we have sought for systematic exploitation of finite semantic rules and properties of inferences of ontologies for the mining of relevant and guided contexts. This provides evidence to show the current state of the environment and allow the pedagogical agent to make decisions to positively support students in their study.
The proposed scheme in this work becomes interesting as it tries to develop more focused learning environments. Learning Environments, Learning Management Systems and Intelligent Tutoring Systems are some of the main interests in our research group as can be seen at [4, 5, 6]. Most times, contextual navigation is overloaded with information, and if not well directed may become boring or harmful to the student’s goals. The pedagogical, animated and affective agent used for this work uses rules of perception and reaction to change emotional states such as sadness and joy, empathy and sympathy. Our research on pedagogical agents is also well established as it can be seen in [7, 8]. This work shows the process of learning from the integration of semantic-cognitive and affective domains, applied on the field of Artificial Intelligence.

For ontology, we mean an explicit specification of a shared conceptualization in a particular context. Our context is the Japanese Language Proficiency Test domain. In this case, the shared conceptualization models the necessary definitions and relations to enable people working in that domain to clearly communicate and exchange knowledge.

It is certainly useful and, most of the times, essential to have students paying attention to relevant information and studying relevant topics. We do not want students lost in a great amount of educational material, searching unnecessarily for relevant material, loosing time with irrelevant information. Considering this, our proposal is to bring relevant topics of certain subject to foster learning. This fostering of learning will be done with the aid and support of domain ontologies.

This scenario, validated within the multidisciplinary environment, presents interesting contributions, which motivated this work: a computational implementation of the formal representation of an ontology for JLPT opens possibilities for new services for the environment; the prospect of Maximum Relevance in the context of JLPT is essential, both to motivate the student to persist in his goal as to foster and direct his learning, important issues in proficiency courses. The use of pedagogical agent with an approach of affective tactics may strengthen the motivation of the student in the environment.

So the focus of this paper is to present the evaluation process of the conceived JLPT ontology. For now, it is our main result as it has taken close contact, interviews and collaboration with different domain experts, in a total of eight evaluators.

Next section presents the problem addressed in this work, the fostering of Japanese language proficiency. Section 3 explains briefly the Relevance Theory as part of the solution for our problem. Section 4 presents the ontology modeled to solve the issue of representing knowledge and some use cases. Section 5 presents the evaluation process achieved so far along with the evaluation obtained from different experts. Last section comes with some final considerations of this work with suggestions for future works.

## 2 Fostering Japanese Language Proficiency

As for other languages, Japanese language has a specific proficiency test, the Japanese Language Proficiency Test (JLPT). JLPT is held once a year in several large
cities worldwide. This test aims to assess and certify the foreigner’s proficiency level of knowledge of the Japanese language. The test is applied only to those whose Japanese is not the native language.

The test is divided into four different levels. The applicant shall submit to the test equivalent to his Japanese language level. In each level, the test is divided into three parts, so called subtests: writing/vocabulary, listening and text comprehension/grammar.

There is not an official public domain agenda for JLPT, but the characteristics of each subtest were verified through examples of previous tests, within the context of content and independent of JLPT level:

- **Writing/vocabulary**: evaluation in the field of reading and writing Japanese ideograms (kanji) as well as of vocabulary domain;
- **Listening**: evaluation on the ability of listening comprehension of spoken Japanese. Language tricks and expressions are common in this type of test;
- **Text comprehension/grammar**: evaluation in the field grammar, reading and text comprehension.

There are many preparatory courses for JLPT. The majority of these courses are offered in a traditional paper-based form and only a few of them are available online. In order to solve this problem, [9] proposed the eJLPT simulator, an educational hypermedia tool for students willing to practice their skills and knowledge of the Japanese language through an online simulation of JLPT.

eJLPT system has been utilized by a community of users interested in JLPT and some need for improvements and new functionalities, such as automatic scenarios, has become necessary. The main concerns were towards having some guidance through learning process, resources to adapt learning needs to JLPT context, special tools for teachers and students to avoid unnecessary effort on certain subjects which may not be important to get approved in JLPT.

To address these and other difficulties a new architecture for eJLPT is proposed. The existing system should not be presented as a unique hypermedia document for every user. It should be as flexible as an adaptive learning management system. In order to provide adaptativity of content and interface, one of the necessary issues to be considered is having an adaptation model which supports the user model. In this work, the user model is based on a cognitive perspective. The intention is to model the domain in a way that it should represent part of the cognitive context of an individual. Therefore, our approach aims to model the domain using an ontology based on a cognitive perspective, referred to as the Relevance Theory [10]. Another issue that needs to be considered is the adaptation mechanism. In our schema, this will be done by an intelligent pedagogical agent. Details of this mechanism will be treated in future works.

Next section focuses on the usage of the Relevance Theory as the support to understand, model and represent, in some way, cognitive contexts. It will also give directions on the attempt to achieve the maximum relevance of a certain subject according to pre-defined goals.
3 Relevance Theory

One of the keywords of this work is the word ‘context’. It may recall a lot of different definitions, ideas and meanings and may be used in almost all areas of knowledge. Many definitions of context may be found in the literature [11, 12, 13, 14, 15]. In Computer Science it is usually used in the area of Ubiquitous Computing, in which most of the usage of contextual information regards to circumstantial context. The research relies mostly on how to manage contextual information of the individual’s identity, his location and the moment in which the individual is interacting with the system. This is not the intention of this work neither the main concern of the adaptativity.

One of the first challenges is the attempt to model a knowledge domain to reflect, in some way, the representation of the cognitive context of a person over such domain. For cognitive context, we mean the set of assumptions used to interpret a statement and, for cognitive context of a person over a particular domain, we mean the person’s subset of assumptions on such domain.

Therefore, there is not the intention to model a circumstantial context neither the entire cognitive context of an individual, which would probably be impossible, but only those relevant to the domain of JLPT. An ontology was developed to represent this domain of JLPT. It consists of a class that represents the major context and subclasses representing subordinated contexts. In this case, the ontology plays both the role of representing the general area, acting as a course agenda, as something more specific, such as the representation of the knowledge already acquired by a particular individual. In this case, the ontology of the individual is a subset of the general ontology of the domain.

Besides the representation of the concepts, the ontology should allow navigation in a context from the perspective of the Relevance Theory. According to this theory, for any assumption to be relevant in a context, there should be connections among new assumptions and existing ones, which are already part of the context.

Relevance can be characterized in terms of contextual effects. To modify and improve a context is to have some effect in that context. There is no change in the context where the information is completely duplicated or when it is not related to any old information. There must be an interaction between old and new information. “The context used to process new assumptions is a subset of the old assumptions of an individual, with the new assumptions which combine to generate a variety of contextual effects” [10].

To ensure relevance, certain conditions must be met, which leads us to a comparative definition. An assumption is relevant in a context in the proportion of: 1) the increase of the contextual effects and, 2) the decrease of the effort required to process it in that context.

One of the main goals of this work is to meet the principle of maximum relevance by assessing the relationship of best cost-benefit between the contextual effect and processing effort.

Considering the graphical representation of ontology, assuming that we are in the original node, the more we move towards the children nodes, the greater the contextual effect obtained. To have this move forward, and consequently, the
contextual effect, the move should occur by the links between nodes. Each move is valued in accordance to the specified value in each link in the ontology modeling.

The processing effort, similarly, also increases as we move through the graph. As the value of contextual effect, the value of the effort of processing is not an absolute value, does not have a unit of measure, but is a relative value, for comparative purposes, given by experts who, based on domain principles and parameters, can identify topics with greater weight and importance to the learning of subsequent topics.

Currently, the values for each move for both contextual effect and processing effort are unitary values. In the future, we intend to attach values to the links in accordance to principles and parameters of second language learning.

Next section will present the fundamentals of ontology, specially the application of this knowledge conceptualization in education and the concepts of the JLPT ontology.

4 JLPT Ontology

It is important to evidence that there is no official and available public agenda neither from Japan Foundation\(^1\) nor Japan Educational Exchanges and Services\(^2\), entities responsible for the administration, organization and dissemination of JLPT.

Once it was decided that an ontology would be a good solution for our problem, we started a search on the main ontology repositories (Swoogle\(^3\), Protégé\(^4\), DAML\(^5\), Tones\(^6\)). We’ve found some ontologies on Japanese Language, but none of them was complete or related to JLPT.

Therefore, the concepts for the development of this ontology were obtained through the consulting on the grammar adopted by Japanese school books, available preparatory courses and interviews with an expert, a Japanese language teacher for JLPT certification. The presented ontology is, then, an agenda, result of the compilation of several materials related to the Japanese language grammar because it expresses and defines the rules for a language as well as for the vocabulary. Figure 1 represents the class ‘Grammar’ and its is-a relations.

Japanese language is rich in vocabulary and expressions. Japanese grammar clearly defines its elements and rules essential to the Japanese. Gobi is an example that we can cite as trivial. Trivial in natural language, from the perspective on human reasoning. But how to express and make Gobi instances to be automatically computationally identified as trivial? Gobi is our use case and its use is explained further. One of the possibilities is to use description logic to facilitate the representation of knowledge based on logic.

The main characteristic of description logics is that concepts (or classes) can be defined in terms of descriptions. These descriptions specify the properties whose

\(^1\) http://momo.jpf.go.jp/jlpt/overseas/index_en.html
\(^2\) http://www.jees.or.jp/jlpt/en
\(^3\) http://swoogle.umbc.edu
\(^4\) http://protegewiki.stanford.edu/index.php/Protege_Ontology_Library#OWL_ontologies
\(^5\) http://www.daml.org/ontologies
\(^6\) http://owl.cs.manchester.ac.uk/repository/browser
objects must satisfy the domain of the concept. The language to be used shall allow the construction of descriptions of composition, including restrictions on binary relations that connect objects. This language defines a set of instances with syntax, to build the description, and semantics, representing concepts in multiple hierarchies, organizes in a structure of subconcepts, or taxonomy. Such formalization allows identifying attributes such as multiple inheritance, restriction of values, limits and also roles (transitivity, inversion, etc.).

In first order logic, the predicate is a language feature that can be used to create a statement or give a property. Thus, properties and instances (individuals) can be represented in predicate logic. Predicate logic has expressiveness from existential (at least some individual/body) and universal (all subjects/body) quantifiers. A quantifier is a logical symbol that quantifies the instances (individuals).

Next, we present part of the taxonomy of JLPT ontology, the Grammar class. Just after, we present some use cases from the stage of formalization of the ontology.

Fig. 1. Classes of JLPT ontology.
Use case: Gobi class

The Gobi class describes a common and useful resource in sentence endings. Gobi is very used in the Japanese language. It literally means “language tail” and it simply refers to anything that comes at the end of a sentence or a word, such as a verb (the minimum of a complete sentence in Japanese language). Most common Gobi usage is presented in this class. Some examples: yo-gobi, ne-gobi, yone.

Necessary and sufficient condition:
Instances (Individuals) are inferred as Gobi instances if they are instances of the subclass Basic_Grammar:

\[ \text{Gobi} \equiv \text{Basic}\_\text{Grammar} \]

A necessary condition can be identified in this class:
Gobi is a subclass of Basic_Grammar:

\[ \text{Gobi} \subseteq \text{Basic}\_\text{Grammar} \]

Instances of the mapped classes allow the inference of relevant contexts to JLPT agenda, from the Japanese Grammar formally expressed as:

\[ \text{YO-GOBI} \subseteq \text{Gobi}, \quad \text{YO-GOBI} \equiv \text{Basic}\_\text{Grammar} \]
\[ \text{NE-GOBI} \subseteq \text{Gobi}, \quad \text{NE-GOBI} \equiv \text{Basic}\_\text{Grammar} \]

Use case: Verb class and its subclasses

Verb class covers the basic properties of verbs. In Japanese language, verbs can be divided in:

- Intransitive Verb
- Causative Verb
- Causative Passive Verb
- Passive Verb
- Transitive Verb

Some necessary conditions may be identified in this class:
Verbs are Basic_Grammar subclasses:

\[ \text{Verbs} \subseteq \text{Basic}\_\text{Grammar} \]

Instances of Verbs class relates to each other through hasParts property with at least one instance of Particle class:

\[ \text{Verbs} \subseteq \exists \text{hasParts}\_\text{Particle} \]

Instances of Verbs class relates to each other through hasParts property with at least one instance of Gobi class:

\[ \text{Verbs} \subseteq \exists \text{hasParts}\_\text{Gobi} \]

Instances of Verbs class relates to each other through isPartOf property with at least one instance of Expression class:

\[ \text{Verbs} \subseteq \exists \text{isPartOf}\_\text{Expression} \]
Instances of Verbs class relates to each other through isPartOf property with at least one instance of Sentence class:
Verbs ⊆ ∃ isPartOf.Sentence

Instances of Verbs class relates to each other through isPartOf property with at least one instance of Essential_Grammar class:
Verbs ⊆ ∃ isPartOf.Essential_Grammar

The types of verbs used in Japanese Grammar classification (Intransitive Verb, Causative Verb, Causative Passive Verb, Passive Verb) are subclasses of Verb class. As consequence they inherit its properties and necessary condition.

Intransitive_Verb are Verbs subclasses:

Intransitive_Verb ⊆ Verbs

If there a instance of Intransitive_Verb, it consequently cannot be a instance of Causative Verb, Causative Passive_Verb or Passive_Verb. Formally, this can be expressed using disjoin:

Intransitive_Verb ⊆ (∼CausativeVerb, ∼CausativePassiveVerb, ∼PassiveVerb)

Some properties used to specify the classes:

hasParts is an inverse property of object of isPartOf, which domain and codomain is the class Context.

hasParts ∈ P

hasParts = isPartOfR ⊆ ∀ hasParts.Particle . . R ⊆ ∀ (hasParts.Particle)
hasParts = isPartOfR ⊆ ∀ hasParts.Gobi . . R ⊆ ∀ (hasParts.Gobi)
hasParts = isPartOfR ⊆ ∀ hasParts.Verb . . R ⊆ ∀ (hasParts.Verb)
hasParts = isPartOfR ⊆ ∀ hasParts.Expression . . R ⊆ ∀ (hasParts.Expression)
hasParts = isPartOfR ⊆ ∀ hasParts.Sentence . . R ⊆ ∀ (hasParts.Sentence)
hasParts = isPartOfR ⊆ ∀ hasParts.Basic_Grammar
. . R ⊆ ∀ (hasParts.Basic_Grammar)

Use case: Complete_Sentence class

A complete sentence is a specific sentence in Japanese grammar. It has, at least, one verb.

At least one necessary condition can be identified in this class:

Complete_Sentence is a subclass of Sentence:
Complete_Sentence ⊆ Sentence

Some necessary and sufficient conditions to identify Complete_Sentence class can be expressed:

Instances of Complete_Sentence classes are related through hasVerb property with at least one instance of the class Verb:

Complete_Sentence ⊆ ∃ hasVerb.Verb
One instance of \texttt{Complete\_Sentence} relates to the property \texttt{hasVerb} at least once:

\begin{align*}
\text{Complete\_Sentence} \subseteq (\geq 1 \text{ hasVerb})
\end{align*}

As class \texttt{Complete\_Sentence} is a subclass of \texttt{Sentence}, it inherits from \texttt{Sentence}:

\begin{align*}
\text{Complete\_Sentence} \subseteq \exists \text{ hasOrderDirect.OrderDirect} \\
\text{Complete\_Sentence} \subseteq \exists \text{ hasOrderIndirect.OrderIndirect}
\end{align*}

Some properties in this class were defined and are used:

\texttt{hasVerb} is a property of functional objects, which domain is the \texttt{Sentence} class and the codomain is the \texttt{Verb} class:

\begin{align*}
\text{hasVerb} \in P_0 \\
R \subseteq (\leq 1 \text{ hasVerb}) \\
R \subseteq \forall (\text{hasVerb.Sentence}) \therefore R \subseteq \forall (\exists \text{Parts}. \text{Verb})
\end{align*}

In this case, the ontology plays both the role of representing the general domain, acting like a course agenda, as something more specific, such as the representation of the knowledge already acquired by a particular individual. In this case, the ontology of the individual is a subset of the general ontology of the domain. The shared attribute of the conceptualization also lead to a validation process which is explained on Section 5.

Finally, we refer to explicit specification as the formal language used to develop the ontology computationally. We adopted OWL (Web Ontology Language) as the formal language, since it is developed by the World Wide Web Consortium and is a de-facto standard for building web ontologies. Research on OWL and its inference properties lead to several implementations of fast and reliable reasoners. The availability of this kind of tool is important since the ontology is the base to our adaptive learning environment.

## 5 Evaluation Process

So far, there has been a partial evaluation process, accomplished in three stages. First the ontology was evaluated by four experts on the domain, through interviews. Their expertise is on Japanese language proficiency tests. This stage of the process helped to achieve the “shared” property of the ontology. It demonstrated that the ontology is somehow reflecting the common body of knowledge of the Japanese proficiency domain.

We are aware that more experts must evaluate the ontology, but this is part of the work evolution. As the ontology and our adaptive system are made public, we will be able to gather more feedback and improve the conceptualization.

Secondly, we based the evaluation stage on the work of [16]. That work defines and formalizes several metrics to evaluate the ontology from quantitative and qualitative perspectives. The definition is represented with a meta-ontology and the formalization establishes the mathematical formulae to calculate the quantitative part.
Three ontology experts have evaluated JLPT ontology using modularity, depth, breath and accuracy parameters. From this partial evaluation, the ontology is valid. Our next step is to evaluate the ontology against the full set of parameters defined in [16]. At this point we used more quantitative aspects since we are more concerned with the practical utilization of the ontology. The pragmatics aspect is related to the fulfillment of the requirements established at the specification stage with support from the utilization scenarios (more end-user-related issues).

Finally, the assessment of the extent to which the ontological knowledge base reflects the constructs prepared by the experts in the domain, is, in essence, the assessment of a conceptual ontology.

The approach adopted was the one proposed by [16], known as assessment of the functional dimension of an ontology, with respect to its main purpose, i.e., the specification of a particular conceptualization or contextual assumptions on the domain or area of interest. According to [16] such specifications are always approximations, because the relationship between ontology and conceptualization (cognitive semantics) always depends on what the rational agent has conceived and on the semantic space that formally codifies it (formal semantics). Thus, according to the authors the functional assessment should focus on examining how these dependencies are implemented considering the ontology as a language that includes the object information and the required conceptualization.

In this context, [16] suggest evaluating the expressiveness of the ontology, with a computational perspective, according to its suitability and fitness to the concepts and relationships present in a referenced body of knowledge, or with experts of the involved areas. Regarding expressiveness, the ontology was evaluated in two phases, with experts in ontology engineering and cognitive science.

The ontology was formalized through description logic, which was implemented through the standard language called OWL DL (Web Ontology Language – Description Logic), developed by the World Consortium. Thus, since OWL was used for the development we were able to use two reasoners that logically verify OWL ontologies. The purpose of a reasoner is beyond this simple evaluation, but this is the part that better fits this stage of the work. Logical consistency tests were performed with the Pellet and RACER reasoners, and the results, analyzed by ontology experts. Both of them generated the same results confirming the validity of the ontology. Refinements in the ontology related to naming standards, modularity and connectivity between classes were identified and considered.

It should also be noted that the ontology was developed according to the concepts of the JLPT, using prior proficiency tests, grammar books and preparatory courses material widely recognized in the area, confirming the semantics of the ontology on JLPT domain considering the particularities of the Japanese language.

After validating the scenarios established in the use cases, through the symbolic description represented by the definitions of classes and instances in OWL-DL language, a validation of the domain ontology from the perspective of the Relevance Theory was accomplished. This was done through a document presenting the problem, the chosen approach and the proposed solution, with a questionnaire at the end, given to the cognitive science expert. In this step, the evaluation was done according to the glass-box method of functional dimension of [16]. The criteria were related to competency adequacy, such as compliance to expertise in the area,
flexibility, precision and recall related topic with specificity. The result of the evaluation of the ontology was considered highly satisfactory, responding to the four basic points based on the Theory of Relevance: the possibility of mapping the cognitive domain for the JLPT, reflection of ontological notions of contextual effect and processing effort, and finally, the conclusion that the ontology can meet the principle of maximum relevance through navigation on the ontology.

Therefore, given the assessments made by experts, it was possible to demonstrate that the ontology is appropriate and promising for the domain and desired goals. Suggestions to extend the purpose of the ontology to attend to other aspects of the Relevance Theory were identified by the expert, and contribute to future works.

After the current phase of implementation and partial evaluation, the intention is to have a prototype in order to verify experimentally the optimistic responses given by the experts, to compare, in groups of users, the use of the previous environment against the environment with the new implementation and the adapted navigation through relevant ontologies.

6 Final Considerations

The development of an ontology is a very hard task by itself. It requires a lot of research, dedication and, above all, close contact to knowledge experts of the proposed domain. The development process is exhausting and requires many interviews, many back and forth in each stage.

Another challenge was to develop it according to a perspective of the Relevance Theory. In many ways, it should reflect the notion of cognitive contexts. It was very important to have an expert who could evaluate it.

As the validation process shows, it seems that, so far, we have achieved a mature state of the ontology. With it, we may feel more confident that the domain is well modeled in many aspects – computationally, conceptually and cognitively.

With the preliminary validation of the experts, the knowledge base of the system is formalized. The adaptation of content, through the implicit heuristics in the structures of the ontology, allows the pedagogical agent to support and assist the student. These characteristics, necessary for an ITS, through an approach of adaptation of relevant content with semantic taxonomy, are interesting expected contributions to the AI research area.

For future works, the intention is to develop ontologies services so it becomes a real semantic web service, as a semantic repository, with other domain ontologies. Given the explicit semantics of knowledge representation, it is possible to instantiate a web service provided with features such as: semantic repository, recommendation services from expert perspective, or even though, with the integration of intelligent agents, among others.

There is also the intention to do some formalization on the Relevance Theory in order to make replication possible. Besides that, there should be the evaluation with other domain experts and an analysis of the results.
Acknowledgments

This work was partially supported by FAPEAM, CAPES, CNPq, Brazil.

References