Chapter 9
Ontology-Based Interoperability in Knowledge-Based Communication Systems

Stefan Trausan-Matu

9.1 Introduction

One of the main ideas of this chapter is that for assuring interoperability for different kinds of applications, the existence of a well designed shared structure of concepts, a so-called "ontology", has a major importance. The structure should be well designed in order to capture the relations and the properties of concepts in the considered domain, and to allow the inference of new properties. A carefully developed ontology may be easily used as the main component of different *Knowledge-Based Systems* (KBS), similarly with a human being, which uses the same knowledge in performing a large range of activities.

KBS are software applications in which there is an explicit representation of knowledge, in a so-called knowledge base. There may be declarative ("know that") or procedural ("know how") knowledge. The ontology of a given domain is an important part of the declarative knowledge, containing at least the basic concepts of the domain and the relationships among them. This fact enables the reuse of an ontology in various Knowledge-Based Systems for the given domain.

Another important idea discussed herein is that ontologies have a role not only in KBS, which are traditionally based on the individual cognition paradigm considering the knowledge which is in someone's (e.g. an expert) mind (paradigm which proved to be limited, see for example, Brown (1999)). Ontologies may have a central place also in applications based on the opposed, socio-cultural idea, which considers that knowledge is socially constructed (Vygotsky 1978), in which communication and, in particular, dialogue has a major role. In fact, *dialogism* is a basic theory in the socio-cultural approach, considering that everything is a dialogue (Bakhtin 1981).

S. Trausan-Matu (✉)
Computer Science Department of the Politehnica University of Bucharest, Romania
e-mail: trausan@gmail.com

Dialogism can be used as a theoretical substrate for developing software tools for supporting collaborative applications which use also ontologies (Trausan-Matu et al. 2007).

The socio-cultural paradigm is present not only in explicit dialogic tools as instant messenger (chat) or discussion forums. It is also the substrate of the so-called Social Web (Web2.0), which includes also *folksonomies* (Mika 2005), which are opposed to ontologies in the sense that they are constructed by a community and not mainly by an individual (in folksonomies people add tags on web resources, the side effect being that they make a classification and a social conceptualization).

In this chapter we will try to integrate the two paradigms, cognitive and socio-cultural. We will show how socio-cultural concepts may be represented in an ontology and how an ontology may be used in a socio-cultural dialogic system. We will discuss how the same domain ontology may be used in a couple of different Knowledge-Based Systems: an expert system for giving advice related to urbanism and civil-engineering regulations, and a system that uses the ontology and natural language processing techniques for assisting participants to a virtual chat conference for design or e-learning. In this second system is illustrated also the possibility of adding new concepts in the ontology. The two applications considered are typical for the two paradigms: One is starting from the idea that the ontology provides the knowledge usually possessed by a human expert. The second enables humans to construct knowledge together through dialogue.

Even if based on different paradigms, both applications may be seen as involving knowledge communities that share a domain ontology and maybe a top level socio-cultural ontology. The idea that links these kinds of applications is that they use, in fact, different ways of communication and, in particular, the way of entering into a dialog: database query, hypertext browsing, keyword-based search engine, intelligent search engine, expert system dialog, controlled natural language, question answering, and natural language dialog. All these offer access, in different degrees, to socio-cultural knowledge construction.

The chapter continues with a section introducing the basic ideas of Knowledge-Based Systems, the role of ontologies in developing KBS and the reuse opportunities that they offer. The following section will be dedicated to a socio-cultural top-level ontology. The fourth section discusses the dialogistic character of any information system. The next sections introduce the above mentioned two application examples. The chapter will end with conclusions.

9.2 Artificial Intelligence, Knowledge-Based Systems and Ontologies

Artificial Intelligence is an interdisciplinary domain in which researchers try to develop "intelligent" computer programs that behave like intelligent humans in solving complex problems and that may communicate using human-like means, e.g. natural language. One of the ideals to be reached by artificial intelligence is to

develop so-called *expert systems* that could enter in dialog, answer questions and provide solutions to problems in the same way a human expert would do.

Expert systems are Knowledge-Based Systems, which intensively and explicitly process important amounts of knowledge, similarly to an intelligent human person. Moreover, KBS are computer programs in which a clear difference is kept between the knowledge they use and the procedures for processing it. This division permits the incremental development of the so-called knowledge bases, while the processing procedures remain unchanged and are usually reused for a wide range of applications. This possibility is crucial for the development of computer programs for problems which are usually solved by humans which posses a large amount of knowledge in problem's domain. The reason is, first of all, psychological: It is very hard for a human to describe the whole amount of knowledge it uses. The constructing of knowledge is much easier in an incremental process. It is easier to understand what knowledge the system lacks by examining its behavior, by entering into a dialog with it, if possible.

Another important consequence of the separation of the knowledge base and the inference engine is the reuse of the knowledge bases for several different purposes (e.g., constructing a solution, understanding a solution or generating explanations). However, for this purpose the base should be carefully developed, in order to be sufficiently general. One perspective in this idea is to construct the knowledge base starting from an ontology, which should be viewed as a "theory" of the considered domain. An ontology is a particular kind of knowledge base, containing declarative knowledge and being a skeleton for further knowledge acquisition activities. In fact, this was the context in which the term "ontology" started to be used in knowledge engineering in the beginning of the last decade of the twentieth century: For developing the knowledge base of an intelligent program, "knowledge engineers" realized that it is very useful to have a skeleton of the main concepts and relations of the considered domain, a so-called ontology.

Viewing knowledge bases as ontologies determines important advantages for developers of Knowledge-Based Systems. First of all, an ontology is developed as a coherent framework for the reality and therefore it facilitates knowledge acquisition and machine learning. It is easy to add a new concept in such a framework by finding one or some more general concepts and defining some differences between the new concept and the more general ones.

Ontologies are very important in applications that extract knowledge from texts (text mining) and, in general, to applications for the Semantic Web (Berners-Lee et al. 2001). For this kind of software they offer the substrate for semantic analysis and, very important, the possibility of defining a measure of semantic neighbourhood, based on the lattice structure (based on the hypernymic order relation) of ontologies (Hirst and St-Onge 1998). This semantic closeness is very important in text analysis for example in the retrieval of texts that do not contain a given word, but they contain a synonym or a semantically related word.

From a knowledge representation perspective, ontologies are semantic networks that state what kinds of concepts exist and what abstraction-particularization (generalization/ specialization) relations hold among them. If a concept is a
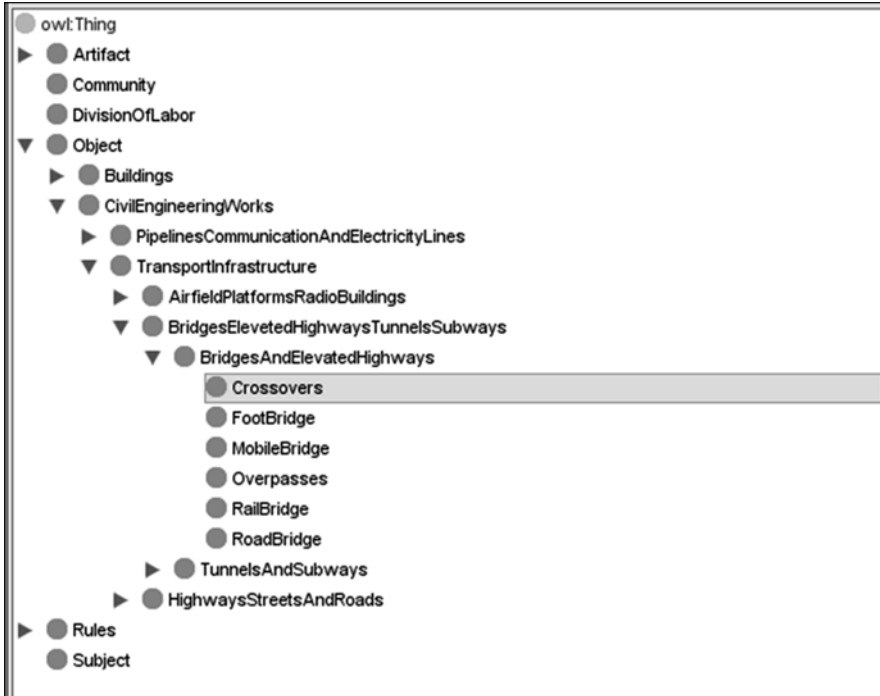
Fig. 9.1 A fragment of the urban development and civil engineering ontology

particularization (specialization?) of another concept, it has all the features of the more abstract concept and, maybe, some particular ones, For example, in Fig. 9.1, the fact that the concept "BridgesAndElevatedHighways" has "Crossovers", "FootBridge", "MobileBridge", "Overpasses", "RailBridge" and "RoadBridge", implicitly enumerates the only possible cases. Moreover, all these concepts inherit properties (e.g. regulations) that belong to "BridgesAndElevatedHighways" or its ancestors (the Protégé environment - http://protege.stanford.edu - was used for the development of the ontology and the image is a screen-shot from it).

Ontologies offer reuse, simplifying computing in a similar way with Object-Oriented Programming (whose idea has common ancestors with ontologies in frames (Minsky 1975)). For example, an ontology may be seen as a library of concepts and relations that may be used for many applications. Another important resemblance is encapsulation and centralization, which simplify changes: When some concept or relation changes, it is enough to make a modification in a single place and all the descendant applications will inherit the new version.

However, ontologies do not cover all kinds of knowledge representation. In addition to declarative knowledge representation, there is a need also for procedural knowledge, saying what to do in a given context. Such type of knowledge

may be represented by production rules, which are pairs *condition – action*: IF *condition* holds, THEN PERFORM *action.* Conditions usually contain patterns and variables that may be linked to facts. A production rule system has a conflict resolution strategy that selects the rule that will be applied from the rules that may be applied.

9.3 A Top Level Socio-Cultural Ontology

In the socio-cultural paradigm, knowledge is seen as constructed and shared by communities of people acting in a more or less concerted way. In order to assure a coherent behaviour, people obey some explicit (for example, written) or implicit (for example, tacit) rules. Individuals (subjects) in a community may have different roles (for example, leader, professor, student, etc.) and their activities are associated with different types of work achieving some outcomes (or objects). As a consequence, a division of labor, that means a classification of different types of work should be considered.

A very important role in achieving the outcomes is the existence of mediators between subjects and objects, the usage of artifacts, of tools (Vygotsky 1978). A remarkable example of artifacts is the human language, which is a major tool enabling humans to collaborate (Vygotsky 1978).

All the above concepts are the ingredients of the Activity Theory of Yrjö Engeström (1987), which emphasizes categories (subjects, objects, and communities), mediators (general artifacts, social rules and division of labor) and relations between them (see also Fig. 9.2). This theory provides a theoretical framework
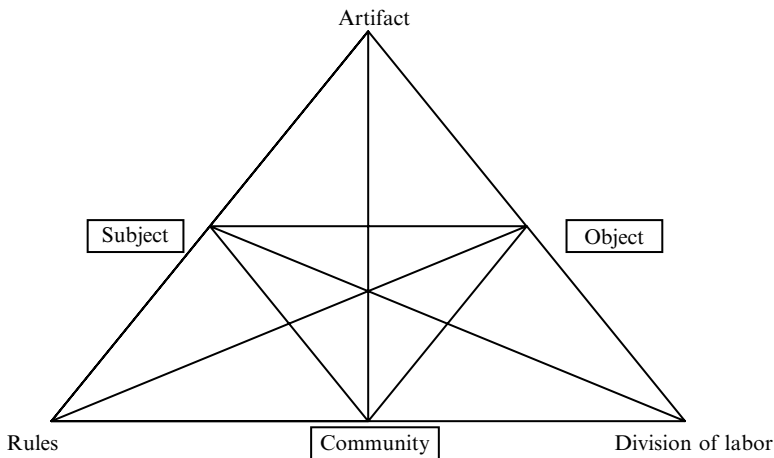

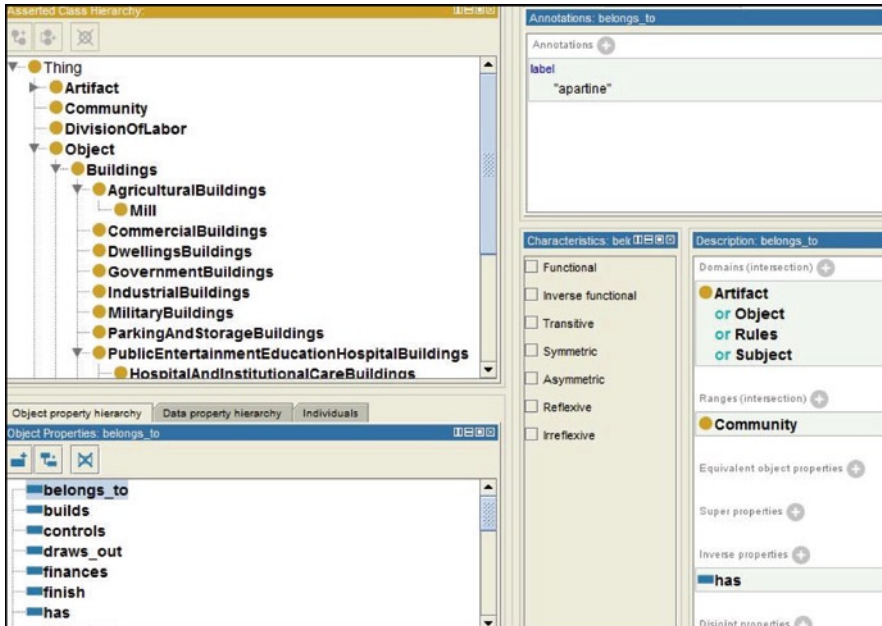**Fig. 9.2** The main concepts of the Activity Theory of Yrjö Engeström

Fig. 9.3 The relations of the urban development and civil engineering ontology, emphasizing the domain and range of the relation "belongs_to"

that has been used for developing an ontology for urban development (Trausan-Matu 2007) that has as basic concepts the components of the above mentioned two group of entities.

If we want to integrate the socio-cultural paradigm with the cognitive, knowledge-based one, an ontology of the socio-cultural concepts is useful. This ontology may be used as a top-level ontology, from which should be specialized concepts for the different socio-cultural communication and information systems. Starting from the Theory of Activity, each of the six entities is a basic concept (or "class") in the socio-cultural ontology. These concepts may have attributes, sub-concepts (that may be also sub-concepts of several other concepts, i.e. multiple inheritance of properties is allowed), and relations with other concepts (see Fig. 9.3).

In addition to generic concepts, the ontology contains also individuals (instances). For example, the "Subject" class has 12 instances (see Fig. 9.4). One of these, the "LocalAuthority" instance has several relations ("provides", "releases", "controls", etc.) with other individuals.

The idea of developing an ontology for the socio-cultural top concepts (a top-level ontology) starting from Engstrom's ideas (Engström 1987) is new. We do not know about other expert systems or ontology-based systems for urbanism or any other domain using such a top-level ontology.

Fig. 9.4 The "LocalAuthority" individual and its relations with other individuals

9.4   Types of Dialog in Information Systems

In order to provide the needed advice to various types of users and to different kinds of questions, an intelligent information system should consider several ways of dialogical interaction. Any information act is, in fact, dialogic. Moreover, as Bakhtin emphasized, any text is a dialog (Bakhtin 1981): Even if you write something and you upload a document on the web, this is a potential dialog with the readers of the text. Bakhtin's dialogism is extending Vygotsky's ideas (Wertsch 1991), it is one of the most important representant of the socio-cultural paradigm.

Different ways of querying in information systems may be considered as different ways of entering in dialog. A classification of querying types, on an increasing scale of the degree of dialogism is:

(a) database query
(b) hypertext browsing
(c) keyword-based search engine
(d) semantic web search engine
(e) intelligent search engine
(f) expert system dialog
(g) controlled natural language
(h) question answering
(i) natural language dialogue.

From the above list, only natural language dialog and question answering are, at least for the moment, the less satisfactorily implemented. All the other ways of information querying are, more or less, possible to implement.

The degree of dialogism, in our perspective is including the importance of contextual information, which is extremely important in natural language dialogue and

minimum in database query. Other factors are the existence of inter-animation patterns (Trausan-Matu et al. 2007) and the consideration of an ontology.

– All the above types may (or must) beneficiate from an ontology. If for natural language processing the use of an ontology is mandatory (for deriving the meaning of the sentences), for the other cases is not, but it may be useful. For example, keyword-based search engines, hypertext browsing or even database queries may beneficiate from an ontology.
– In the system presented in Trausan-Matu et al. (2002), a domain ontology is used as a starting point in the serendipitous search, that has as result not only new pages but also new concepts (serendipitously detected), that appear in the browsed pages. The same ontology is used for semantic annotation of the retrieved documents and for the retrieval of relevant metaphors from the annotated documents. A collection of dynamically generated web pages reflects the structure of the domain ontology. In addition, the ontology is driving the construction of the user's model and the filtering of the amount of concepts and facts presented, in order to provide personalization. This multiple usage of the same domain ontology is an exemplification of the basic idea of this chapter, that a same domain ontology may be used for several applications, for the above querying styles (e.g. expert systems or intelligent information systems), but also for others, like diagnostic expert systems, intelligent e-learning (Trausan-Matu et al. 2002) or intelligent tools for supporting collaboration (Trausan-Matu et al. 2007).

9.4.1   Expert Systems for Providing Intelligent Advice

Even intelligent search engines, which extend the power of what Google can do, lack the ability to enter into a dialog, to provide an actual advice or plan of actions. What they can do (very well, is true) is to provide relevant documents in which you may find the topics that you mentioned what you need. An alternative is the development of more dialogical intelligent information systems (Trausan-Matu and Neacsu 2008), which provide an extended range of types of query answering, which are closer to the idea of a true dialog.

An expert system was implemented that enters into a dialog with users, for providing information about topics related to getting urbanism authorizations for new buildings. The Jess production rule system (http://www.jessrules.com/jess) was used. A program in Jess is a collection of rules that can be matched to the existing data in the working memory. Each rule has a first, matching part, and a second, action one, which modifies the working memory or prints something. A rule may have variables that are linked to values in the working memory using pattern matching. For example, a rule that prints the information that local authorities may provide is below exemplified. In this rule, the variables $?p, $?r, and $?c are matched to all the available data, in the working memory, regarding what the local authority provides, releases and controls.

```
Do you want to learn what documents you need when building a house? (yes or no) yes
Have you reveive the authorization? (yes or no) no
Do you own the property? (yes or no) yes
Do you have the budget? (yes or no) yes
Did you asked for Urban Certificate? (yes or no) no
Solution: Go to the local authority
Do you want to know more on the subject? (yes/no)
yes
LocalAuthority provides:
    - AprouvalDocumentAnalyse
    - AuthorizationRequestAnalyse
    - AuthorizationDocumentRelease
  releases:
    - ConstructionAuthorization
  in accordance with:
    - no10_1995 (see the web-site: ~#en http://www.pptt.ro/content/legea_nr__101995_privind_calitatea )
    - no350_2001 (see the web-site: ~#en http://www.cdep.ro/pls/legis/legis_pck.htp_act?ida=30200 )
    - no50_1991 (see the web-site: ~#en http://www.cdep.ro/pls/legis/legis_pck.htp_act?ida=1322 )
Try again? (yes or no) yes
Do you want to learn what documents you need when building a house? (yes or no) yes
Have you reveive the authorization? (yes or no) yes
Have you paid the taxes? (yes or no) no
Solution: Pay the taxes
Do you want to know more on the subject? (yes/no)
yes
The Investor has to pay
    - AprovalDocumentReleaseTaxe
    - ConstructionTaxe
    - AuthorizationTaxe
    - DesignTaxe
Try again? (yes or no) yes
```

Fig. 9.5 A dialog in the expert system session

```
(defrule local_authority
  (declare (salience 1))
  (print go_to_local_authority)
  ?f <- (object (is-a Subject)
                (:NAME "LocalAuthority")
                (provides $?p)(releases $?r)(controls
                $?c))
  (not (answer ?))
  =>
  (printout t (slot-get ?f :NAME) " provides: " crlf)
  (foreach ?x $?p (printout t " - "(instance-name ?x) crlf))
  (printout t " releases: " crlf)
  (foreach ?x $?r (printout t " - "(instance-name ?x) crlf))
  (printout t " in accordance with: " crlf)
  (foreach ?x $?c (printout t " - "(instance-name ?x) crlf))
```

In Fig. 9.5 is illustrated a simple dialog that, among others, presents what the "LocalAuthority" can provide, release and control. An important observation is that the data is obtained from the ontology and it may be different if the ontology changes.
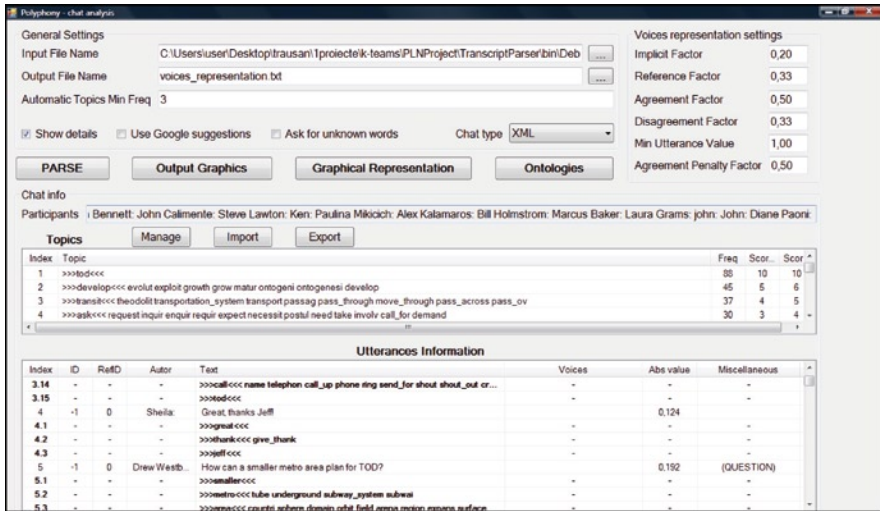
Fig. 9.6 Topics detected from a chat on urban development

9.4.2 Ontology-Based Support for Dialogue in Chats

Specialists in the urbanism domain have started to use environments, instruments and facilities specific to Web2.0, the so-called *Social Web*, which includes folksonomies, forums of discussions and on-line conversation. The same domain ontologies from the previous sections may be used by intelligent tools that provide useful abstracting facilities for analyzing the interactions in the above mentioned instruments. Such an ontology-based system was developed that detect the topics of instant messenger chat conversations, the threads of discussion and the important utterances. The system visualizes a graph of the conversation (Fig. 9.7) and allows the expanding of the domain ontology (Fig. 9.8). The application may be used for analyzing what was discussed and in what degree participants are implied in the chat conversation. For example, the automatic analysis may emphasize the main topics of discussion in a group of urbanism specialists (see Fig. 9.6).

Topics maybe detected from the frequent words discovered in the text. The application uses the WordNet (Miller 1995) lexical ontology for detecting similar concepts (concepts which are at a small semantic distance – see the section about intelligent search) and a domain ontology, which extends WordNet with domain-specific words and relations which are not present in the lexical ontology.

Natural language technology is used for the identification of discussion topics, for the segmentation of the conversation, for identifying implicit references among utterances and for graphical visualization. The generated diagrams allow identifying the participants which had an important contribution in the conversation. The domain ontology may be extended with the new topics identified by the system, as illustrated in Fig. 9.8.

Fig. 9.7 Visualization of the conversation graph
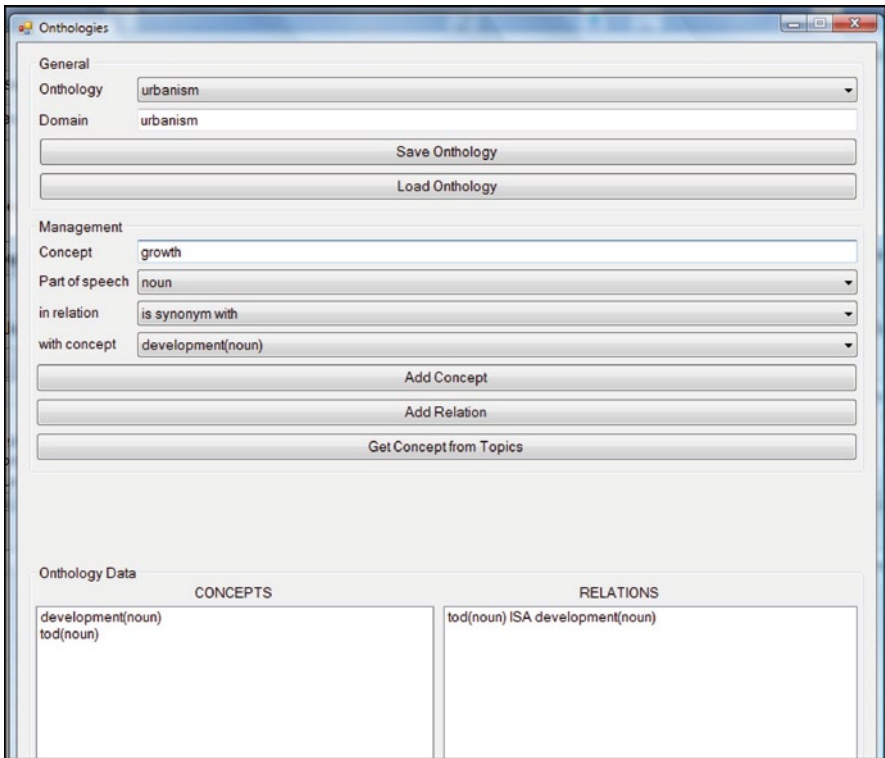

Fig. 9.8 Extending the domain ontology

The visualization of the conversations (Fig. 9.7) shows threads of related utterances in the discussion, allowing to identify important chains of argumentation, the importance of each utterance (the small vertical lines in the upper part of the image – see Trausan-Matu et al. 2007, for details) and the inter-animation (the degree in which participants collaborate effectively, in which their discussion threads display a structure similar to polyphonic music– see Trausan-Matu et al. 2007, for details). These facilities empower us to identify the important conversations and chains of argumentation from a library of online chats like, for example, those at http://www.cyburbia.org or http://www.planetizen.com, in the domain of urbanism (Trausan-Matu and Rebedea 2009). Moreover, in combination with the topic detection facility, it may be used for identifying new concepts and relations and to add them in the domain ontology. For example, in Fig. 9.8, the new concept "tod" (transit oriented development) is related to "development", which is synonym with "growth" (Trausan-Matu and Rebedea 2009).

The topic detection and conversation visualization system has been used in analyzing chats for Computer-Supported Collaborative Learning in Politechnica University of Bucharest, Romania and Drexel University, Philadelphia, USA. A new version is now under development in the EU FP7 project LTfLL (Learning Technology for Lifelong Learning, see http://www.ltfll-project.org/).

Further work will consider more complex semantic distances (than only synonymy) and more elaborated interaction patterns (Trausan-Matu et al. 2007). Machine learning techniques will be used for the identification of discourse patterns. Moreover, a completely automated version for discovering new rules for the implicit relations is in progress.

9.5 Conclusions, Open Problems and Research Challenges

The same domain ontologies may be used in several different KBS, including communication ones. In the context of the Social Web, ontologies may be used in conjunction with tools supporting dialogue for providing intelligent access to information and even procedural help, as shown in a precedent section. In fact, any type of information system may be viewed as dialogue-based. This vision is specific to the socio-cultural paradigm, which considers that knowledge is socially built. Urbanism essentially needs to consider the socio-cultural perspective and, therefore, in this context, the usage of ontologies should also consider socio-cultural concepts. For this aim, a socio-cultural ontology should be developed and integrated with the domain ontology. Such an approach was presented in the paper.

Dialogue may not only beneficiate from supporting tools provided by an ontology. It may also be a source of new concepts to be included in the ontology, as illustrated in the above section.

There are many open problems and research challenges related to ontologies and their relation to the socio-cultural perspective. An ontology should be a shared conceptualization, as the very much cited definition of Gruber (1993) states. This status

is at the basis of its potential use in different applications for a large number of potential users. However, it is many times very difficult to have an agreement on the structuring the concepts among human experts in a given domain. Moreover, there are opinions that even it is sometimes impossible to obtain a categorization (Lakoff 1987). Related to this, it is not yet clear how to construct ontologies starting from data provided by socio-cultural structures like folksonomies.

It is not yet clear how to integrate ontologies and knowledge provided in social networks. They are based on different if not totally opposing paradigms: ontologies follow a cognitivist approach, which usually describe knowledge which is in the mind of an expert, while the socio-cultural view considers that everything is socially constructed.

Another fundamental problem is related to the extent to which the knowledge may described as an ontology and to which expert systems may become similar in power with humans. It seems that there are types of knowledge, for example tacit or experiential which may not be represented in ontologies. The metaphor: "stocks are very sensitive creatures" is giving us very valuable insights in the behavior and characteristics of stocks, we even could understand them by comparing to ourselves, as very sensitive creatures, having the experience of being living creatures (Trausan-Matu et al. 2002).

Starting from the problem of metaphors, we should say clearly that Natural Language Processing is far from being similar to human capabilities and many opinions (for example, Winograd and Flores 1986) say that it will never be, limiting the possibilities of extracting ontologies automatically and contradicting optimists that see ontologies and associated description logics as providing full support for powerful language technologies.

References

Bakhtin, M.: The Dialogic Imagination: Four Essays. University of Texas Press, Austin (1981)

Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web, Scientific American, May 2001

Brown, JS.: Learning, working & playing in the digital age. AAHE Conference on Higher Education, www.johnseelybrown.com/learning_in_digital_age-aspen.pdf (1999). Accessed 22 Dec 2010

Engeström, Y.: Learning by Expanding: An Activity Theoretical Approach to Developmental Research. Orienta-Konsultit Oy, Helsinki (1987)

Gruber, T.R.: A translation approach to portable ontologies. Knowl. Acquis. **5**(2), 199–220 (1993)

Hirst, G., St-Onge, D.: Lexical chains as représentations of context for the detection and correction of malapropisms. In: Fellbaum, C. (ed.) WordNet: An Electronic Lexical Database, vol. 13, pp. 305–332. MIT Press, Cambridge (1998)

Lakoff, G.: Women, Fire, and Dangerous Things: What Categories Reveal About the Mind. University of Chicago Press, Chicago (1987)

Mika, P. Ontologies are us: A unified model of social networks and semantics. In: International Semantic Web Conference, Lecture Notes in Computer Science, pp. 522–536. Springer, Berlin Heidelberg (2005)

Miller, G.A.: WordNet: A lexical database for english. Commun. ACM **38**(11), 39–4 (1995)

Minsky, M.: A Framework for representing knowledge. In: Winston, P. (ed.) The Psychology of Computer Vision. McGraw-Hill, New York (1975)

Trausan-Matu, S., Maraschi, D., Cerri, S., Trausan-Matu, S., Maraschi, D., Cerri, S., Trausan-Matu, S., Maraschi, D., Cerri, S.: Ontology-centered personalized presentation of knowledge extracted from the web. In: Lecture Notes in Computer Science, vol. 2363, pp. 259–269. Springer, Berlin Heidelberg (2002)

Trausan-Matu, S.: A socio-cultural ontology for urban development. In: Teller, J., Lee, J., Roussey, C. (eds.) Ontologies for Urban Development, Studies in Computational Intelligence, vol. 61, pp. 121–130. Springer, Berlin Heidelberg (2007)

Trausan-Matu, S., Stahl, G., Sarmiento, J.: Supporting polyphonic collaborative learning. E-ser. J. **6**(1), pp. 58–74. Indiana University Press (2007)

Trausan-Matu, S., Neacsu, A.: An ontology-based intelligent information system for urbanism and civil engineering data. In: Teller, J., Tweed, C., Rabino, G. (eds.) Conceptual Models for Urban Practitioners, pp. 85–92. Società Editrice Esculapio, Bologna (2008)

Trausan-Matu, S., Rebedea, T.: Ontology-based analyze of chat conversations. An urban development case. In: Teller (ed.) Proceedings of Towntology Conference, Liege (2009)

Vygotsky, L.: Mind in society. Harvard University Press, Cambridge (1978)

Wertsch, J.V.: Voices of the Mind. Harvard University Press, Cambridge (1991)

Winograd, T., Flores, F.: Understanding Computers and Cognition. Ablex, Norwood (1986)