Supporting Tools for Blended Animators in Complex Learning Communities

Berardina De Carolis, Stefano Ferilli, Enrica Gentile, and Paola Plantamura
Dipartimento di Informatica, Università degli Studi di Bari “Aldo Moro”, Bari, Italy

Abstract – The society model of the last years has given a key role to knowledge in terms of economic and social development. On this background, it is allocated the success of the communities of practice, and more recently, of the Complex Learning Community model, whose strength is represented by the ability to provide students, educators, and professionals with a common creative space to develop not only knowledge and expertise, but also ideas, synergies, chances. In both cases, particular emphasis is given to the interactions between users as a “place” where knowledge emerges, is built and delivered. Our former paper was dedicated to the communities of practice, and in particular, to the use of an intelligent agent to control and lead the interaction among users; now our research concerns environments for collaborative learning, where the role of the animator is fundamental, but we take into account a blended animator, supported by specific tools to manage the community. Such tools are necessarily connected to the comprehension and elaboration of the natural language which highlights the appropriate knowledge elements to be used. In this paper we intend to analyse the available techniques in order to identify those that are more suitable for designing the supporting tools required by the blended animator. Particular attention will be given to the domain ontology.

Index Terms — Learning communities, blended animators, ontologies.

I. INTRODUCTION

The expectations of the European Council to transform the Old Continent, until 2010 [1], into an “economy based on the most competitive and dynamic knowledge of the world”, have been disregarded because of the economic and financial crisis of the last few years.

At the same time, this crisis really seems to have strengthened the firm belief that the only useful antidote is to invest in knowledge as the primary resource for the economic and civil growth. Indeed, in the new ten-year strategic plan, the European Committee established that the principal aim for Europe in 2020 is a “clever growth”, founded on the development of “an economy based on knowledge and innovation [2]”.

Once more, the learning process is emphasized, and in particular the demand to conform it to the needs of a society deeply changed by the diffusion and continuous development of communication and information technologies, highlighting the necessity to move into models, procedures and formative artifacts the research of the last decades in the Tic field.

On the other side, experience has often shown that the adoption of new technological tools is not a sufficient condition to determine the paradigm change required by the knowledge society, that is, the overcoming of a disciplinary learning model based on progressive increasing of contents and expertise [3].

In fact the failure of some e-learning experiences was determined by the lack of synchronization between technological and didactic innovation.

In order to reach this balance and according to the existing literature in the field of collaborative learning in the net, we want to face the theme of new knowledge generated through the interaction between the users of a community of practice, reaching the definition of adequate tools and methodologies to support the animators who have the task to moderate or facilitate the collaborative activity inside these communities.

II. KNOWLEDGE MANAGEMENT IN LEARNING COMMUNITIES

Three factors characterize the knowledge society: change acceleration, globalization and complexity [3]. Knowledge is not a static and codified entity anymore, but a dynamic one.

A fundamental contribution comes from Edgar Morin’s studies on complex learning. According to the French philosopher «order, disorder and complexity coexist» [4] and the comprehension and management of such complexity must be an essential competence of the trainers. In fact, complexity is not an obstacle, but a value. Then, it is necessary to acquire the competences to understand the positive aspects of the complexity and manage it. So doing complex learning re-shapes the learner’s role from a user to a builder of his own knowledge, through continuous, not a-priori defined interactions.

The learner’s role, as well as the role of learning as a social phenomenon, were also discussed by Etienne Wenger [5] in his studies on the communities of practice, spontaneous aggregations of individuals who share the same practice and the same formative interest, where processes of knowledge transmission, building and increasing are fed by continuous interactions between the participants.

Communication and cooperation activated among the participants, in fact, aim to formalize and capitalize, inside a shared repertoire, the knowledge [6], also the so-called tacit knowledge [7], so difficult to be identified, communicated and shared. Exchange of experiences, reciprocal help, comparison between pairs and sharing make these aggregations real “communities which learn” [8], supporting the growth and professional updating of the whole group.

However, founding the knowledge production and diffusion principally on spontaneous, not codified relationships – carried out in virtual environments
essentially based on the use of written words – if on one side, can be considered a winning factor of the communities of practice, on the other represents a critical point in the learning process. The knowledge created, developed and delivered through informal modes, or that is not a-priori classified, may be dispersed, or unlikely retrieved and shared. Instead, the value of the knowledge fundamentally depends on its capitalization, transmission and tracking inside a universe which changes constantly and rapidly.

For this reason, the technological endowment of the environment is determinant. Synchronous and asynchronous communication tools, repositories, databases and knowledge management tools are required. Furthermore, people are required, having the task to keep the interactions between participants useful to the learning goals, and capitalize and manage the generated knowledge.

Wenger himself, although pointing out the spontaneous and informal nature of the communities of practice and the absence of any hierarchy between the users, recognizes to the coordinator a "crucial role" [7]. He/she is not an external character, but one of the participants who has the task to focus on the thematic field, support the relationships and develop the practice.

According to Wenger, another important character is the "Community Librarian" [7]. His/her function is fundamental to make the community a "legitimated and influential organizational actor", by promoting collection and evaluation of materials on the basis of a taxonomy which is consistent with the community interests, and helping in organizing information, writing summaries, and schemes.

Describing the organizational structure of the community of practice, Salima Salis et al. [9] point out different roles, among which particular relevance is given to the "community manager". He/she is the person who coordinates the publishing activities, by filtering and gathering the contents of the interactions among the members of the communities.

Guglielmo Trentin [10] proposes a *vademecum* for the "facilitator of activities in the net", who has particular relevance in orienting the community members. According to Trentin, his/her role is not only determinant in distinguishing quality and quantity of the participation, but also in continuously feeding the interaction. By pointing out the difficulties, but also the undeniable advantages, of the net communication, which is made, above all, by written texts supported by technologies, the author underlines the necessity of a wide involvement of the participants.

This is the principal task of the facilitator, whose action must focus on improving interactions by stimulating the discussion with creative ability.

### III. Conceptual Graphs as a Support

In previous work of the authors the topic of knowledge management in learning communities has been already analyzed [11], focusing on the monitoring of interactions aimed at evaluating their actual effectiveness for the creation and diffusion of knowledge. In particular, that work focused on the definition of a method for modeling the ability of participants in collaborative activities in a network, combining linguistic analysis methods aimed at identifying the communicative contribution of single statements through interaction pattern analysis methods aimed at analyzing the comprehension abilities of the participants in a discussion.

After realizing that the amount and significance of information in textual form is, in the context of learning communities (such as the Communities of Practice), extremely relevant, it becomes evident the need to support the users that play the role of animator, moderator or facilitator, with suitable Natural Language Processing tools that help them in analyzing and constantly keeping under control the information content of the system and its evolution(s). Actually, the huge amount of data to be handled makes a pure manual approach clearly unpractical, and represents a pressing motivation toward the identification of automatic or semi-automatic solutions.

In practice, a first form of processing of texts in natural language, that is useful to extract the information that is more immediately evident, but also preliminary to subsequent processing tasks, consists with no doubt of the lexical and syntactic analysis of the texts themselves. To this aim, a wide variety of linguistic resources and tools have been developed in the last years, and are nowadays widely available as state-of-the-art. The available tasks include identification of the language in which a text is written, filtering of less significant words (such as articles, prepositions, conjunctions), identification of the grammatical category of words, extraction of the root of terms. The whole landscape of such resources and techniques has undergone a significant evolution, as regards quantity and quality, mainly for the English language, for obvious motivations. Hence, the interest is still high toward the possibility of having equivalent items for the other languages, as well, and possibly to obtain them automatically rather than spending time and money to manually produce them through the work of experts in linguistics and computer science. Once these resources and tools are available, an additional set of tools can be applied to index the texts according to the most relevant terms they contain, in order to allow an effective and efficient retrieval thereof based on queries made up of sequences of terms.

While all of the above has significantly simplified a large part of the procedures, and satisfied many needs, of the users of information in textual form, it cannot be denied that stopping at the lexical level is strongly limiting, due to the well-known ambiguity problems that are intrinsic to natural language, above all synonymy and polysemy.

To overcome this limitation, one needs to proceed in at least two directions. On one hand, one must perform syntactic and logical analysis of the texts, that is able to exploit not just the terms, but also the relationships among terms, in order to better understand the role that each term plays in the text. On the other hand, one must also associate a meaning to each term in the text. We believe that the step from the purely terminological level to the conceptual one is fundamental and priority with respect to the objectives of the moderator of a Community of Practice, and hence in this paper we will focus on this aspect in particular. Indeed, having available a structured representation of the concepts underlying a given domain may help in better organizing the texts, in retrieving them in a way that is more focused and tailored on specific needs, in contextualizing and cross-relating them among each other and with the various users.
Also at this level, some resources are available in the literature that allow to step from the terms to their underlying meaning, also providing an apparatus of taxonomic and non-taxonomic relationships that relate terms, meanings and both. The most famous and outstanding example is WordNet [13]. In some cases it is additionally possible to obtain formal descriptions of the concepts, this way reaching the level of real formal ontologies. Again, the availability of these facilities is significantly wider for English than for other languages. However, a general conceptual taxonomy/ontology, such as WordNet, can hardly capture the peculiarities of a community that lives around more restricted or specialist domains, and hence can offer a quite limited support to the aforementioned objectives. On the other hand, developing this kind of resources requires efforts that are out of the capabilities of each single community, and this represents another motivation for the research of automatic solutions.

In particular, we propose the development of techniques and tools that, starting from the very texts that make up a specific corpus (e.g., those handled in the various respects of a Community of Practice), are able to identify the concepts underlying the collection, to define their degree of relevance and their inter-relations, to discover more frequent and significant usage patterns, and possibly their abstract definitions. In short, these tools should be able to build (domain-specific) conceptual taxonomies or ontologies starting from texts in natural language. First of all, the outcome would be specifically tailored on the specific domain of interest. Secondly, the extracted concepts would have a direct relationship with the portions of text expressing them. Third, the taxonomy would enable higher-level analysis of the text contents. Last, but not least, one might save the time and money needed to manually produce these resources. A desirable characteristic of these techniques should be incrementality, that would allow to keep continuously up-to-date the ontology with respect to the corpus evolution, and hence to ensure that it reflects at every time the state of the world and suitably supports the users (both the moderator and the participants).

Extraction of taxonomies or ontologies by mining (large quantities of) texts is a research branch of Artificial Intelligence that has already resulted in the publication of several solution proposals in the literature. Basically, these proposals can be divided in two main categories: the former considers only what is contained in the text, while the latter exploits the availability of external resources, as well, to fill the gap between the purely syntactic level and the semantic one.

The approaches of the former kind are specifically suited when there exists no external knowledge structured in a format that can be directly exploited by computers. In this landscape, [12] builds conceptual hierarchies using Formal Concept Analysis to group objects based on their attributes, that are in turn determined starting from the text and using the links between nouns and verbs. On the other hand, for our purposes it would be more useful to build conceptual graphs that exploit the whole network of available concepts and relationships. Another approach, proposed in [16], defines a language for the construction of formal ontologies by deductive discovery, as in the Logic Programming setting. More specifically, the Author defines both a specific language for the manipulation of Web pages, and a logic program that discovers the concept lattice. In our case, it would be desirable that the nature or kind of relationships to be extracted is not limited a priori to a pre-defined set, but we would rather like to create and add new relationships any time a new (i.e., not yet encountered before) verbal relation is encountered.

As concerns approaches of the latter kind, the literature has mainly focused on the construction of taxonomies and/or ontologies. In particular, [14, 15] builds ontologies by labeling taxonomic relationships only, while –as said– for our purposes it would be useful to extract also other kinds of relationships derived by actions expressed by verbs, and hence non-taxonomic. [17] builds taxonomies by considering only those concepts that specifically appear in a given domain, but not in others. We believe that this is a significant limitation, as well, because deeply understanding a corpus must necessarily be based, in addition to domain-specific concepts, on generic concepts that make up a kind of “glue” in which domain-specific ones are immersed, as well.

Thus, in our vision, we propose to start from a syntactic analysis of the texts that make up the corpus, to identify the structure of single sentences (including special forms, such as negative or passive ones), and then to subsequently apply some kind of logical analysis that is able to identify subjects and (direct or indirect) objects. Clearly, more traditional forms of text processing, such as stemming or, even better, lemmatization (that returns the basic form of inflected terms, whereas stemming just returns their stem), which is more indicative of their grammatical role and can be more easily read by humans, which are actually the very final users of the entire apparatus. The logical structure of sentences, then, will represent a base for building the actual taxonomy/ontology.

More specifically, in our view nouns are candidates for nodes of the taxonomy, representing concepts. Adjectives represent properties of concepts, and are useful to identify similarity among concepts in order to group them in classes. Verbs express relationships among concepts, that provide links among concepts that extend and complete the conceptual graph with relationships other than just generalization/specialization. Verbs are also useful to derive additional features of concepts, to be used again for grouping similar concepts. After extracting the basic graph directly expressing the content of the corpus, subsequent analysis will be started to enrich the conceptual graph with derived, high-level knowledge and relationships.

IV. CONCLUSIONS AND FUTURE WORK

We believe that the approach we propose may find a useful application field in the forum, that is one of the most widely exploited tools in the Community of Practice environment, in its double role: as a communication means and as a repository of the knowledge created through the users’ interactions. Beyond its flexibility, this tool has many points of weakness, among which the need to keep consistent the users’ interventions and to make easily retrievable the generated knowledge. This requires a significant cost, in terms of effort and time, by the moderator or facilitator. Using and underlying ontology, that is linked to each piece of text expressing the underlying concept(s), automatically identifying inconsistencies and retrieving the (parts of) documents that are associated to given concepts would become much easier.
Another possible application field that is particularly interesting can be identified by considering the role of the “Community Librarian”, as outlined by Wenger. In particular, we refer to the task of organizing the information from a large, and very often not homogeneous, database. Just think of the creation of summary documents concerning a specific problem or subject of interest for the community starting from chat conversations, forums or previously processed documents, or even of a simple indexing by relevant subject of the documents that are stored in the community database. Again, having a conceptual graph that is directly learned from the involved texts allows to immediately associate each text to the underlying concepts, and hence to easily retrieve all the pieces of knowledge that are relevant to a given subject.

REFERENCES


AUTHORS

Berardina De Carolis is with the Computer Science Department (Intelligent Interfaces group), University of Bari, 70125 Italy (e-mail: berardina.dec Carolis@uniba.it)
Stefano Ferilli is with the Computer Science Department and with the Inter-departmental Research Center on Logic and its Applications, University of Bari, 70125 Italy (e-mail: stefano.ferilli@uniba.it).
Enrica Gentile is with the Computer Science Department, University of Bari “Aldo Moro”, 70125 – Bari (BA) – Italy (e-mail: gentile@ di.uniba.it).
Paola Plantamura is with the Computer Science Department, University of Bari “Aldo Moro”, 70125 – Bari (BA) – Italy (e-mail: pplantamura@ di.uniba.it).

Manuscript received 26 March 2013. This work was supported in part by the U.S. Department of Commerce under Grant BS123456 (spenser and financial support acknowledgment goes here).

Published as submitted by the author(s).