A Concept Map Extractor Tool for Teaching and Learning

Luiz Claudio Duarte Dalmolin, Silvia Modesto Nassar, Rogério Cid Bastos, Gustavo Pereira Mateus
Informatics and Statistics Department (INE)
Federal University of Santa Catarina (UFSC)
Florianópolis, Brazil
dalmolin@labtrans.ufsc.br, {silvia, rogerio, gpmateus}@inf.ufsc.br

Abstract — This work presents a tool to help concept maps’ design process. This tool was developed to organize contents from thematic modules of a course, applying concept maps techniques. In this tool, teachers can build their own concept maps based in textual reference documents. Teachers also can use a visual editor to design maps and linking learning objects to concepts. Students can access learning objects and make their own notes in their own concept maps improving teaching and learning process. The concept map extractor is an algorithm based in text mining techniques for term extraction. This algorithm extracts relevant terms that can be considered concepts or links, making agile the concept map building process.

Keywords - Concept Maps; Term Extraction; Learning Objects;

I. INTRODUCTION

Nowadays teachers have a lot of information from numerous sources such internet websites and other digital media libraries. These sources have contents which can useful to plan a course or simple class to any learning level. Otherwise, organizing contents and identifying the main topics of a thematic module of a course or a class is not a trivial job. Finding out functional forms for representing, organizing and sharing knowledge to be taught using computers is an issue for computer aided learning.

Concept maps (CM) are considered a useful manner to represent knowledge in a concise and accessible form. Frequently CMs are used in learning management, since they offer the possibility to personalize learning and re-enforce learning to improving learn skills, promising vehicle for knowledge sharing [1][2].

Concept mapping can be used either as a planning and teaching tool for teachers or as a learning tool for students. Unfortunately, generating such knowledge models may require considerable effort to determine which concepts and relations to include [4][6]. In such cases, a concept map extractor tool can be helpful on teaching and learning process, mainly, when it is used by a teacher to build a concept map to represent a course program or a thematic module. Teachers can use CM to highlight key concepts and principles to be taught in a highly concise manner [3].

This work presents a tool whose main features are identify relevant concepts in textual documents, design concept maps and store multi-disciplinary learning objects linked to concepts added in a map.

To make it possible, an algorithm for term extraction was developed. This algorithm is based in a hybrid approach that analyses some statistical and grammatical variables to calculate a relevancy of a term.

II. CONCEPT MAP EXTRACTOR TOOL

The concept map extractor consists in a desktop tool for design CM. The CM’s design process proposed in this work involves an algorithm that analyses a set of text documents about a determined subject (e.g. courses program, class tests, exercises, academic articles, book’s chapters) and extracts relevant concepts and their semantic connectors.

More than a term extraction algorithm, this tool can be used to represent knowledge of a domain. The combination of CM and learning objects theories, allow to teachers and students to choose a more precise and structured knowledge, thus improving learning processes [2].

The figure 1 describes the steps which user must follow to use all tool’s resources. In green is showed actions allowed only for students, in red actions allowed for teachers, and blue actions allowed for both kinds of users.

![Figure 1. Steps to use the proposed tool.](image)

Starting the tool, users must firstly fill a cadastre to gain access to the system. This is a very simple form used to get information about users, providing important data to control courses’ statistics and module’s access control.

Students and teachers have different holes at the system. Teachers have access to concept map extractor module and course creation module, while student users can access learning objects and concept map resources.
By creating a course is possible to group concept maps within in a same subject. So, all lessons of a course can be planned in different concept maps.

When a teacher creates a concept map of a course, he may link some kind of learning objects to any added concept. The tool allows adding links to websites, textual documents, images and other multimedia objects. Shortcuts call Google and Wikipedia websites to easily show additional information about selected concept.

The interface allows different ways to insert concept map objects, using mouse click, textual insertion and automatically using concept mapping extractor module. Semantic links are created dragging the mouse over objects that represent concepts. The map’s objects can be configured of many ways, such as: changing text color, background images and background forms.

A. The Concept Map Extractor Algorithm

The proposed term extraction algorithm was developed to work with BR-Portuguese idiom. All heuristics are based on BR-Portuguese grammar rules. Different idioms can be used changing some rules and some input files. In this algorithm every word separated by a space is considered a term. It was defined that a concept can be formed by one two or three terms to get a best performance.

Every document added to CM Extractor must have a relevancy measure defined by the user. This relevancy can assume values between one and five. This value is used to calculate the relevancy of a term.

When the user starts the concept extraction module, he must include reference texts to the system. Different kinds of documents may be inserted; some of them are semi structured, like academic articles, course programs. There are specific fields to each structure of these documents. Each field must be valued with relevance measure which will be used on term ranking.

To extract terms from a textual document is necessary to pre-process it [5]. The first step is changing all words to upper case. Then special characters and punctuation marks are removed. After this, a set of string lists are created to store distinct terms and its respective relevancy coefficient. Some heuristics are used to extract bad formed concepts (e.g. contain many stopwords, is a stopword).

A relevant term may be considered a concept. The term’s relevancy (TR) is determined by a ranking strategy that provides a term relevancy coefficient. Equation 1 describes the ranking strategy, where DC describes the total of documents added to the system, DR describes the relevancy’s average of documents which contain a Ti term. Tiocc describes the occurrences count of Ti in all added documents. T represents the total of distinct terms extracted of all added documents.

\[
TR = Tiocc \times \frac{1}{DC_{Ti}} \times DR_{Ti} \times \ln(T)
\]

This equation is based on total occurrences of a term, in all analyzed documents, (i.e. if a concept appears in a lot of documents it has a higher relevancy).

The algorithm’s main objective is to find semantic connections between relevant terms. A semantic connection can be described by neighbor terms being two concepts and one link. A link is a text that appears between two relevant terms. The textual analyses can extracts n semantic connections between two concepts. So, it is necessary to identify what is the best link to connect these two concepts.

The proposed algorithm implements a statistical model to calculate the link relevancy (LR) based on term’s relevancy, number of words, verbs and stopwords between these terms. The equation 2 shows how is calculated a link’s relevancy.

\[
LR = \frac{\ln(TR_a + TR_b) \times LS}{e^{(c+1)^2 + (v+1)^2 - (s+1)}}
\]

The system separates the stages of term extraction in two. If a user just want identify the most relevant concepts and make the link connections by himself is possible and faster. Identifying semantic connections on the text may
take more processing time and more there are more results to be analyzed before build the concept map extracted.

After use the extractor algorithm user can select in a grid what concepts and what links will be added to concept map. And then, the system automatically draw the CM.

III. VALIDITY AND RELIABILITY OF THE TOOL

The use of this tool supports a new manner to organize courses’ curriculum changing by a sequential learning to a concept map view, inciting meaningful learning and making possible to get an interdisciplinary knowledge.

The tool presented in this work suggests two different ways to use concept maps in teaching and learning process:

- Teachers use it to represent the knowledge of a course’s program or a course’s thematic module, they also can order to students making their own CM to analyze their previous knowledge about a subject.
- Students use concept maps created by teachers to access learning objects, reinforce their own knowledge analyzing concepts and links, drawing the map, they also can make their own notes, adding new concepts to an initial map, or still, build a new CM about any subject and share it with their colleagues.

A test of concept extractor where were used a set of five reference documents (summaries, courses programs, book index) of Statistic Courses from different graduation programs from five different universities of Brazil.

The documents were added to concept map extractor with all the same relevancy. Then the process was started. The system got more than 600 distinct terms on text, ordered by their relevancy.

The list of concepts was submitted to evaluation of 10 experts in Statistic (teachers). For all of them 70% of top 30 more relevant concepts extracted can be inserted in a Statistic Course’s concept map. The Semantic Links extracted has 40% of accuracy in this experiment.

New tests are being prepared for validate concepts and semantic connections extracted by algorithm.

IV. CONCLUSIONS AND FUTURE WORKS

The use of this tool supports a new manner to organize courses’ curriculum changing by a sequential learning to a concept map view, inciting meaningful learning and making possible to get an interdisciplinary knowledge.

Using this concept extractor and the developed heuristics to textual analyses had a interesting results once the most of semantic relations can be inserted in a new concept map.

In a pedagogic perspective, this tool intends to facilitate learning process, through accessing learning objects, making notes, identifying unknown concepts, focalizing studies at more relevant concepts.

We consider that adding learning objects to a concept map makes this tool a specialized knowledge repository, consequently, it may contribute to computer aided learning, allowing to teachers reuse learning objects and share interdisciplinary knowledge.

Next features to be implemented are: a new module to exporting contents to html template files for being applied directly in Distance Learning Environments; develop some agents to analyze student’s navigation.

Using this concept extractor and the developed heuristics to textual analyses had a interesting results once the most of semantic relations can be inserted in a new concept map.

We consider that adding learning objects to a concept map makes this tool a specialized knowledge repository, consequently, it may contribute to computer aided learning, allowing to teachers reuse learning objects and share interdisciplinary knowledge.

Next features to be implemented are: a new module to exporting contents to html template files for being applied directly in Distance Learning Environments; develop some agents to analyze student’s navigation.

V. REFERENCES