A Tool to Support the Moderation of Asynchronous Online Discussions

Rodolfo C. Raga Jr.[1]
Jennifer D. Raga[1]

Maria Rosario D. Rodavia[2]
Arlyn D. Orense [2]

Raymund Sison[3]

Jose Rizal University[1]
80 Shaw Blvd. Mandaluyong City
Philippines
{jundy.raga, jennie.raga}@gmail.com

Arellano University[2]
2600 Legarda, Manila, Philippines
{rose.rodavia, arlyn.orense}@gmail.com

Dela Salle University[3]
2401 Taft Avenue, Manila
Philippines
raymund.sison@delasalle.ph

ABSTRACT

Asynchronous Online Discussion (AOD) forums are increasingly becoming a ubiquitous tool for promoting interaction and learning among students in open and distance learning environments. However, generating high-quality interactions in the non-visual and text-based setting of this tool requires huge and constant efforts on the part of the teacher-moderators. With this requirement, come major concerns about the need to develop tools that can provide automated assistance to teachers in utilizing and adopting these forums. Our study seeks to address two issues related to this concern. The first relates to monitoring forum contributions and their relevance to the topic of discussion and the second examines the feasibility of assessing message contributions relative to specific concepts. In this paper we aim to share our work regarding the former issue. We present an overview of the strategy we adopted in the development of our system which simply combines the semantic properties of words with some statistical properties of text to develop a method which can be used to compare the topic relevance of message contents. We also present the features of our proposed system that make it suitable for educational application, in particular, its ability to generate acceptable performance using as few human resources as possible. This paper also describes the experiments we conducted to analyze the performance of our system by running it through several messages of an actual forum transcript and discusses how the results obtained compare to a human coded set of results. Although we are only at the beginning of our work, the initial results, we believe, are encouraging enough to merit further investigation.

SUBTHEME

Future Directions, Spaces and Possibilities in ODeL

KEYWORDS

Discussion Forums; Moderation Support Tool; Topic Drifting; Message Content Analysis

1. INTRODUCTION

Although asynchronous discussion groups might afford or support opportunities for productive interaction and learning among students, they do not guarantee it (Murphy, 2004). Empirical evidence suggests for instance that maintaining topic coherence, which is a critical prerequisite for achieving high-quality interaction, is problematic and very difficult to achieve in this environment (Potter, 2007); because the topic focus of online discussions constantly changes from one topic to another (Beaudin,1999). This phenomenon is what (Potter, op cit) referred to as *topic drifts* or the tendency of online discussions to stray from their announced topic. Minimizing the occurrence of this problem often levy much burden which may be too high for the teachers that act as moderators.

Furthermore, teachers often do not have sufficient access to vital information that can help describe how the topic of discussion flows.

In an attempt to remedy this problem, a number of researches over the past decades have focused on developing software tools that can help the teachers monitor the progress of discussion. Many of these tools however, rely on huge volumes of training text and/or large numbers of annotated examples to accomplish their task. An attribute which we believe bears distinct disadvantages in educational application.

The purpose of the study reported on in this paper was to develop a method that could be used to measure the topic relevance of message contributions in asynchronous online discussions. The primary focus was to design a tool which requires as little human effort and resources as possible in order to ensure its value and practicality. The paper begins with an overview of the strategy adopted in the development of the tool. This is followed by a presentation of the features of the tool which highlights its advantages for educational application. A discussion on the experiments conducted to evaluate and analyze the performance of the tool follows. Afterwards, a discussion on the value and practicality of the tool is presented. Finally, implications for future research are then discussed in the conclusions.

2. DEVELOPING THE DISCUSSION SUPPORT TOOL

2.1. Working Definition of Topic

While looking for a working definition of the term topic, we've found that there is much disagreement about the precise meaning and notion of this term and that there is no universal definition of "*topical relevance*" in forum discussions. In our work, we chose to follow closely Brown and Yule's (1986) proposed concept of '*discourse topic*'. This notion defines topic as being represented not just by any single grammatical term or noun phrase, as often happens in the treatment of sentential topics, but by information that can be gathered from all relevant linguistic constituents (i.e., sentences, paragraphs, etc.) of the discourse. Since our work limits the application of the notion of topic to those of textual data, we thus provide our definition of 'topic' as follows: "*a set of words whose inter-relationship with each other form the core of the subject matter of a text or a discussion*"[1]. In line with this definition, we therefore view topic relevance between two texts as a simple overlap between the topics they represent.

2.2. Development Strategy

One of the challenges of building discussion support tools is the task of determining how to build on the components, functions, and principles of currently available technologies to develop a suitable framework that will guide the overall strategy of development. Our starting point for the strategy we adopted in our work was the opinions provided in the literature.

According to Luppicini (2002), some of the serious methodological stumbling blocks that exist in the analysis of text-based conference transcripts can be overcome by implementing better corroboration between latent and manifest content analysis. Schrire (2006) also agrees with this suggestion and quoted several authors' advise (particularly, (Henri, 1992; Hillman, 1999)) not to restrict research in the study of online forum environment to the measurement of manifest variables only but to also consider latent variables. *Manifest content* according to Rourke et al (2000) is content that resides on the surface of communication and is therefore easily observable; whereas *latent content* refers to the hidden facet which involves the imputation of meaning which have to be inferred from the text. Although there is a seemingly shared belief that combining the manifest and latent content of text can

[1] This perspective is somewhat also aligned with the definition provided by Blei & Lafferty (2009) where a topic is viewed "*as a distribution over a fixed vocabulary of terms*".

be beneficial to the analysis of text-based conference transcripts, a computational equivalent of this belief has not yet been realized. The nature and methods for merging these two features are still open questions. To contribute some insight into this issue, the proposed work seeks to intersect the following essential techniques:

1. The use of surface-level text statistics to represent manifest content, and
2. The use of word-space modeling technology to represent the latent content in text.

2.2.1. Surface-Level Statistics

In computational linguistics, word frequency is perhaps the most widely used surface-level text statistics, where frequency here usually means the number of occurrences in a given text or corpus. Word frequency has long been regarded as one key criterion in deciding which are the more or less important words relative to the topic of a text. However, reliance on simple word frequency can sometimes be misleading, especially when dealing with topicality. Two words, for example with different levels of topical importance can still occur in equal number within the text and will therefore be treated as equal using simple frequency. In order to circumvent this problem, in our work, we conflated the simple word frequency technique with other more complex surface-level text statistics, more particularly word density and word burstiness to ensure that a unique value, reflecting topical relevance, can be generated for each word.

2.2.2. Word Space Modelling

Word Space Modeling (WSM) is a technology for inferring the semantic relationship between two words based on their distributional patterns or context usage (Sahlgren, 2006). Word-space models have received considerable attention in recent years, and have begun to see employment in many different types of applications. There are many variants of this technology found in the literature. For our purpose, we utilized a technique known as Random Indexing (RI). As a Word Space Modelling technique, the literature describes RI as incremental, have been found to be more robust on small data sets, and most importantly, it can be used even after just a few examples have been provided to it.

2.2.3. Merging the Two Technologies

As a general strategy, our study explores the merging of the functionalities of surface-level statistics and Random Indexing, the details of which can be found in our previous work (Raga & Raga, 2010). Whereas Random Indexing can provide a measure of how semantically related two words are based on their context usage, the utility of surface level statistics is that it can provide an estimate of how important each word is relative to the topic of a specific text. Using this strategy, we were able to explore a novel means of representing the ‘topic’ of a text, one which we believe can be used to detect the relative topical alignment of even small fragments of texts.

3. FEATURES OF THE DISCUSSION SUPPORT TOOL

In our work, the product of merging the two technologies described in the previous section is a Transcript Analysis Tool (TAT) which mechanizes the topic alignment detection of the contents of message contributions. The TAT is aimed to provide assistance to teachers in their moderation task by offering them suggestions as to which messages are topically relevant and/or irrelevant to the topic of discussion. The system actually hinges on three important modules: (1) The LexNet module, (2) The DynaLex module, and (3) the ConLex module. These three modules, we believe, distinguishes our system from other support tool implementations in that they provide the built-in mechanisms that allow the teacher to conveniently use the tool with minimal effort and with as little data resources as possible. In this section, we will describe the specific functions and purposes of these modules.

3.1.The LexNet module – Generating a network of words to represent the topic of discussion.

The LexNet or Lexical Network module is the core of the system. This module generates a network of content-bearing words rated based on their combined semantic and topical relationships with each other. This module works by first extracting all the important keywords within the given text or a fragment of it. These words are then formed into a network by connecting them to each other with edges. The values associated with the edges describe not only how two words in a text are semantically related but also, how important their relationship is relative to the topic of the text. Figure 1 shows how a small fragment of words in the LexNet are connected using the values of their topical relationship as edges. This type of structure is what the system uses to represent the topic of any given text and can be used to judge the topic relevance of incoming message contributions to the topic extracted from the contents of a reference document.

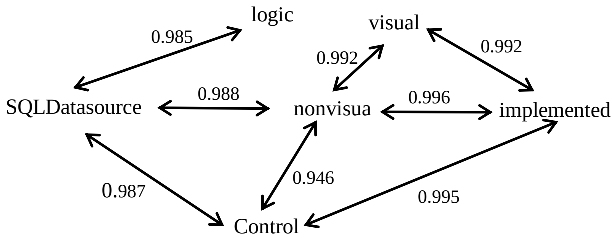

Figure 1: A Fragment of the Lexical Network

The only input requirement for utilizing this module is any raw text document that the teacher deems sufficient enough to serve as the reference document for the discussion. This reference document can actually be extracted from the existing class lectures of the teacher, and since it requires no annotation, the implementation of LexNet module enables the teacher to easily adjust and setup the TAT to support a variety of topics and subject areas. The teachers can also incrementally add more references if they deem it to be necessary.

3.2.The DynaLex module – Dynamic Learning through relevant messages

Although the LexNet module is designed to capture the most detail from the input text and use this to judge the topic relevance of incoming message contributions, the complex process of inference and idea generation in human discourse can easily generate new ideas or responses to the existing discussion topic that cannot be properly handled by the limited contents of any input text.

Figure 2: Snapshot of the ConLex Module Interface

To address this problem, the TAT is also designed to dynamically learn new concepts on its own and this is done through the DynaLex module. The learning in the system is achieved by processing the student's message contributions and applying procedures adopted from the field of machine learning. This learning is done through three phases. First, the relevance value of the message is assessed based on the current LexNet configurations. Second, if the content of the message is found to be sufficiently relevant, all the new keywords found in it are extracted and processed. These extracted keywords are assumed to provide relevant subtopic references. For the third phase, these keywords are used to update the structure of the LexNet, enabling it to pick-up on the keywords the next time it encounters it on succeeding messages.

The entire process is completely autonomous; the only intervention required on the part of the teacher is the setting of the DynaLex learning threshold parameter. This parameter defines the learning rate of the TAT, that is, if the computed relevance score of a given word is within or above the given value for this threshold then it is incorporated into the LexNet structure otherwise it is discarded.

3.3.The ConLex module – Knowledge construction through direct teacher intervention

In addition to the words and concepts learned through the DynaLex module, the teacher can also broaden the scope of the TAT's topic understanding through manual knowledge construction and intervention. This is done by simply editing the existing LexNet structure through the ConLex module. The construction itself is largely discretionary and depends wholly on the teacher's personal opinions and preferences. This way the teacher is able to customize the decisions that the TAT generates to his/her own personal style of teaching because the structure of LexNet is shaped to some extent by the teacher's intuition. For example, through automated processing, the system might get to associate a certain word with another word, the teacher can manually remove this association if he/she thinks that it is not that important to the topic of discussion and redirect the word to link to a another word to form a network that reflect more his/her own understanding. The concepts contributed by the teachers take higher precedence and consideration in the decision making process of the TAT. Figure 2 shows a snapshot of the interface used to implement the ConLex module in the TAT.

4. EVALUATION OF THE PERFORMANCE OF LEXNET

To explore how the features we incorporated into the system may come into use in practice, we performed several experiments that aim to apply the method for the analysis of a transcript of an AOD. The overall goal is to compare the output generated by the tool with the decisions made by human annotators. This section of the paper describes the experiments we conducted and presents the results we gathered including our interpretations of these results.

4.1. Discussion Transcript and Reference Documents

The AOD transcript we used in our experiment is a corpus derived from a public forum. Details of this transcript are presented in table 1 below. We deem this transcript as suitable for our purpose since the discussion that generated it revolved around an online article that had been cited in the first message. We treated this online article as our main reference document. In addition, we also downloaded two other online articles that we aim to use as supplementary reference documents. Table 2 provides details on our complete set of reference documents.

Table 1: Details of the Test Discussion Transcript

Total number of messages posted	87
Total number of participants	16
Average message length	86.3

Table 2: Details of the Reference Documents used in the experiments

Code	Text Title	Size	Topic Focus
refDoc1*	Server Side Debates	635 words	Debates on issues related to SQLDataSource
refDoc2	SQLDataSource Control	611 words	Features of the SQLDataSource control
refDoc3	SQLDataSource & ASP.Net	1294 words	Tutorial on how to use SQLDataSource

* Main reference document

4.2. Manual Message Annotation

The messages in the test discussion transcript were independently annotated by three people who were not otherwise involved in this research. The annotators classified each messages as either topically relevant (R) or topically irrelevant (NR) relative to their perception of the topic of discussion. The technical background of the annotators also varies; the first coder is a call-center agent, the second one is a computer programmer, and the third is a computer science instructor.

4.3. Research Questions

The experiments described below are meant to address the following questions:

1. When trained only with a small reference document, how efficient is the baseline performance of the TAT in identifying relevant and irrelevant messages in comparison with the decisions of human annotators? Will this performance increase with additional reference documents?
2. Given the maximum performance achieved in the previous experiment as baseline, will the performance of the TAT increase when supplemented with the DynaLex module?
3. Given the maximum performance achieved in the previous experiment as baseline, will the performance of the TAT increase if users are allowed to use the ConLex module.

4.4. Performance measure

We associate the performance efficiency of the TAT to the level of agreement of its rating with those of the human annotators. To measure this agreement, we use Cohen's Kappa (K). Cohen's Kappa is the most widely used measure to determine inter-coder agreement in online discourse. It computes the proportion of agreement actually observed between annotators after adjusting for the proportion of agreement expected by chance. The literature suggests an acceptable level of K >= 0.61 when used in analyzing discourse transcripts (Jeong, 2003; Rosé et al, 2008; Mikšátko & McLaren, 2008), but the results can also be interpreted using the table provided by Landis & Koch (1977) as follows:

Kappa (*K*)	**Strength of Agreement**
< 0.00	Poor
0.00 – 0.20	Slight
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Substantial
0.81 – 1.00	Almost Perfect

4.5. Evaluation Approach

To measure and evaluate the agreement between the decisions of the three annotators and the TAT we employed several evaluation strategies as follows.

4.5.1. Measuring Average Decisions

To investigate how the TAT's decisions compare with each individual human annotator, we measured the kappa between its output and the coding decisions of each annotator. We then compute the average Kappa of the TAT with all the annotators and compare this with the average Kappa of the annotators with each other. In essence, we believe this measure will capture the degree to which the TAT's output align with each individual annotator who may utilize particular insights and thus may show varying degrees of subjectivity in discriminating the relevance of each message.

4.5.2. Measuring Concensus Decisions

To account for the overall performance of the TAT, we devised what we refer to as a set of concensus decisions (CD). The CD set is generated by assigning an overall category to each message in the transcript by seeking the agreement of the most number of coders for each item, that is, a message is considered as "Relevant" if at least two annotators considered it as such and otherwise it is tagged as "Irrelevant" and vice-versa. We then compute the kappa of TAT's decisions with the CD set. Given this condition, 30 messages in the transcript were tagged as **R** and the rest as **NR**.

4.5.3. Measuring Unanimous Decisions

Finally, to give us an idea of how limited the TAT's decisions are, we compared its output to a set of decisions which we refer to as the Unanimous Decisions (UD). The UD set is generated by extracting from the transcript all the messages where the annotators showed complete agreement with each other. In this case, a total of 17 relevant and 49 irrelevant messages were extracted from the transcript. We believe that the high confidence implicitly expressed by the aligned decisions of the annotators in these messages will enable us to better evaluate the weaknesses in the TAT's decisions.

4.6. Results

Of the 87 messages that were tagged by the annotators, 66 or 75.9% were tagged by them with the same category. That left 21 messages which were tagged with different categories by each annotator. This gives us a baseline of how humans generally concur on the relevance/irrelevance value of messages. Table 3 displays the Kappa agreement generated by the human coders with each other.

Table 3: Average Agreement between the human annotators.

	Coder1 with Coder2	Coder2 with Coder3	Coder3 with Coder1	Ave. Kappa of Agreement Reached
Kappa(K) Agreement	0.58	0.65	0.67	0.635

Table 4 present the results of the analysis of the average decisions between the TAT and the annotators when only the main reference document was used as input to TAT and when the input size was gradually increased using the two supplementary references.

Table 4: Kappa Agreement between the decisions of the TAT and the human annotators.

Setup of the Text Input provided to the TAT	Kappa agreement with the Decisions of			Ave. Kappa of Agreement reached
	Coder1	Coder2	Coder3	
TAT1 (refdoc1 only = 635 words)	0.40	0.31	0.46	0.40
TAT2 (refdoc1+refdoc2 = 1246 words)	0.43	0.38	0.50	0.44
TAT3 (refdoc1+refdoc2+refdoc3 = 1750 words)	0.55	0.39	0.46	0.47

The data in table 4 shows that using only the main reference document as basis, the TAT was already able to achieve a moderate level of agreement (k = 0.40) with the annotators. Given the size of the training text used, we considered it already noteworthy that the TAT was even able to demonstrate this level of correlation with human judgment. Furthermore, the table also shows that it is possible to increase the efficiency of the system by gradually increasing the size of the training set. Unfortunately, we also found that arbitrarily increasing the size of the training set will not necessarily generate an increase in performance for the TAT. A possible reason is because, although two texts may superficially focus on similar or related topics, the way the authors distribute words in each text is always different. If the word distribution implicitly embedded in each text is not compatible, combining them will only tend to confuse the system and it will generate lower performance scores.

To be able to generate increase in performance, the text used to augment the LexNet structure must be compatible with the initial training data used preferably from the same author. Results using the Concensus dataset and Unanimous dataset are shown in table 5 below. The data presented in this table also shows an increase in performance in both dataset as the input size was gradually being increased.

Table 5: Kappa results generated with the Concensus and Unanimous Datasets

	TAT1	TAT2	TAT3	Ave. Kappa of Agreement reached
Concensus Decisions	0.44	0.47	0.49	0.467
Unanimous Decisions	0.50	0.56	0.62	0.560

In evaluating the performance of the TAT relative to the DynaLex module, we only focused on the setting that generated the maximum performance in the previous experiments (average k = 0.47). Our aim is to test the performance of the TAT with DynaLex using various settings of the learning threshold. Table 6 shows the results we gathered in this experiment. From the data presented in this table, we can clearly see that the DynaLex module, although highly dependent on the learning threshold, can also increase the efficiency of the TAT. In our experiment, the most effective learning threshold found is between 0.3 – 0.6.

Table 6: Kappa Agreement generated by activating the DynaLex with various learning thresholds.

	Learning Threshold applied to TAT3								
	0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
Average Decisions	0.34	0.44	0.44	0.44	0.44	0.44	0.427	0.427	0.427
Concensus Decisions	0.36	0.47	0.55	0.55	0.55	0.55	0.47	0.47	0.47
Unanimous Decisions	0.43	0.57	0.66	0.66	0.66	0.66	0.56	0.56	0.56

Finally, for the last experiment, we conducted several sessions of knowledge construction with the lexical network and for each session determined the effects of the editing we made to the performance of the TAT. The editing actions done for each session are described in table 7.

Table 7: Editing Sessions with the ConLex module.

Session Activity Description		
Session	**Activity Done**	**Effects Observed**
1st session	Created a node for the concept [*advocate*] and linked it to the node of the concept [*Microsoft*]	Converted Message #8 from NR to R
2nd session	Created a link between the concepts [*time*] and [*development*]	Converted Message #64 from NR to R
3rd session	Created a link between the concepts [*project*] and [*sqldatasource*]	Converted Message #87 from NR to R
4th session	Created a link between the concepts [*application*] and [*program*]	Converted Message #35 from NR to R
5th session	Created a link between the concepts [*control*] and [*developer*]	Converted Msg #6 from NR to NR

Finally, the Kappa Agreement we gathered for each session are shown in table 8. This table shows that after only 5 knowledge construction sessions, the performance of the system was dramatically boosted well within the acceptable range. This clearly shows that the idea behind the ConLex module works; direct user feedback, in this case the teacher, can be used to instruct the system to expand or modify its LexNet configuration to broaden its understanding of the topic of discussion. Furthermore, this process can be done in a gradual approach; while the teacher is browsing the messages he can instruct the system to change its perception of the important terms encountered in each message. The only disadvantage we found is that utilizing the ConLex module may initially require a steep learning curve. To address this problem, a built-in phrase analyzer was incorporated into the system. This analyzer provides the users with a summarized as well as step-by-step view of the decision process of the system, this will enable the teachers to quickly focus on important keywords which require editing and construction sessions.

Table 8: Kappa Agreement generated after each session with the ConLex interface.

	Kappa results measured when Knowledge Construction Sessions was applied to (TAT3 with DynaLex)				
	1st	**2nd**	**3rd**	**4th**	**5th**
k =	0.50	0.53	0.56	0.59	0.62

5. DISCUSSION

This section discusses the value and practicality of the TAT as a tool that can assist teachers in their AOD moderating task. The value of the tool relates to its effectiveness in generating results that can provide insight into the topic drifting behavior of the discussion. The practicality of the tool, on the other hand, relates to its ease of use by the implementing teachers.

In relation to value, in this one case of testing, manual evaluation revealed that the output that the TAT generated can be useful for identifying when students are sharing relevant contributions, and when the discussion are starting to gear off topic. The graph in figure 3, for example, visually depicts the R and NR decisions of the TAT, in this case, several clustering of NR points in this graph actually corresponds to phases in the discussion where some participants are flaming each other. We believe that from a moderator's perspective, such visualizations could serve not only as an early warning mechanism, signaling the teachers that their intervention to re-align the topic is needed; but also as a general indicator to the students that their overall engagements are not being productive. In the latter sense, the graph provides the students the capability to monitor their own engagement and enhance their own participation.

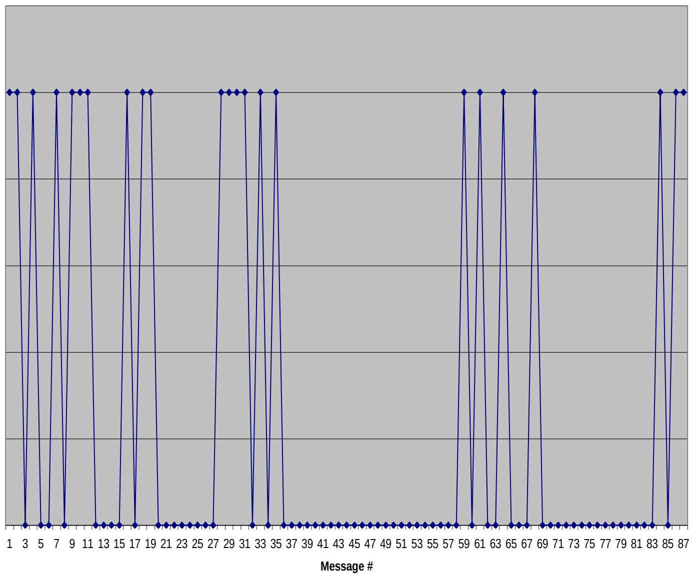

Figure 3: A Graph of the output of the TAT

In terms of practicality, these results are also significant because they were generated based on a very simple algorithm that require little human intervention and do not use any manually annotated training examples. The texts used to train the system are even small enough to be taken from the reference text of class lectures. In this regard, the teachers can easily customize the system to suit their own pedagogical approach. While the knowledge construction process may take some time to master, implementing them is quick and easy once the teachers have gotten used to it. In addition, the system incrementally retains all the knowledge it learns from one forum to another making it more useful for teachers wanting to utilize AOD forums in several classes.

6. CONCLUSIONS

The TAT tool presented in this paper is designed to be of use specifically with teachers interested in utilizing Asynchronous Online Discussions in their pedagogical approach. The results of the evaluation conducted showed that the system could offer support and guidance to teachers in maximizing the benefits of their interventions and can also possibly provide support to students. The tool is also sufficiently practical to implement. Requiring very little resources and less manual work, it could easily be embedded as a background process in forum discussion software.

Although the current paper reports only on the application of the tool for categorizing and describing the topic alignment of contributions, our work also focuses on assessing the value of each contribution relative to specific concepts. This can possibly assist teachers in interpreting and evaluating the learning and quality of participation reflected in student contributions. Future research also requires an expansive study involving larger transcripts, coming from different subject areas and/or generated in an actual classroom setting. In this case, we will need the aid of experienced online professors to help us in interpreting the results that will be generated.

7. REFERENCES

1. Beaudin, B.P. (1999).” Keeping Online Asynchronous Discussions on Topic”. Journal of Asynchronous Learning Network, Volume 3, Issue 2 - November 1999
2. Blei, D. and Lafferty, J. (2009). Topic Models. In A. Srivastava and M. Sahami, editors, Text Mining: Classification, Clustering, and Applications, 2009.
3. Brown, G. & Yule, G. (1983). Discourse Analysis. Cambridge: Cambridge University Press
4. Jeong, A. (2003). The Sequential Analysis of Group Interaction and Critical Thinking in Online Threaded Discussions. The American Journal of Distance Education, 17(1), 25–43
5. Landis, J. R. and Koch, G. G. 1977. The Measurement of Observer Agreement for Categorical Data, in Biometrics. Vol. 33, pp. 159–174
6. Luppicini, R.J. (2002). Toward a Conversation System Modeling Research Methodology for Studying Computer-Mediated Learning Communities. Journal of Distance Education 17(2): 87–101, 2002.
7. Mikšátko, J. & McLaren, B.M. (2008). What’s in a Cluster? Automatically Detecting Interesting Interactions in Student E-Discussions. In B. Woolf, E. Aimeur, R. Nkambou, S. Lajoie (Eds), Proceedings of the 9th International Conference on Intelligent Tutoring Systems (ITS-08), Lecture Notes in Computer Science, 5091 (pp. 333-342). Berlin: Springer.
8. Murphy, E. (2004). An Instrument to Support Thinking Critically about Critical Thinking in OAD. Australasian Journal of Educational Technology, 20, 295-315.
9. Potter, A. (2007). An Investigation of Interactional Coherence in Asynchronous Learning Environments. PhDCS dissertation, Nova Southeastern University
10. Raga, R. & Raga, J. (2010). A Corpus-Based Lexical Network for Measuring Topic Alignment of Forum Messages. Philippine Computing Journal. Vol V, No. 2, ISSN 1908-1995.
11. Rosé, C., Wang, Y. C., Cui,Y., Arguello, J.,, Stegmann, K., Weinberger, A., Fischer, F., (2008). Analyzing Collaborative Learning Processes Automatically: Exploiting the Advances of Computational Linguistics in Computer-Supported Collaborative Learning. International Journal of Computer-Supported Collaborative Learning, 3(3):237–271.
12. Rourke, L., Anderson, T., Garrison, D.R. and Archer, W. (2000). Methodological Issues in the Content Analysis of Computer Conference Transcripts. International Journal of Artificial Intelligence in Education (2000)
13. Sahlgren, M. (2006). The word space model. PhDCS dissertation, Stockholm University.

14. Schrire, S. (2006). Knowledge Building in Asynchronous Discussion Groups: Going Beyond Quantitative Analysis. Computers & Education 46 (2006) 49–70.