<table>
<thead>
<tr>
<th>Title</th>
<th>Temporal analytics with discourse analysis: Tracing ideas and impact on communal discourse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author</td>
<td>Alwyn Vwen Yen Lee and Seng Chee Tan</td>
</tr>
<tr>
<td>Source</td>
<td>LAK '17 Proceedings of the Seventh International Learning Analytics &amp; Knowledge Conference, pp 120-127</td>
</tr>
<tr>
<td>Organised by</td>
<td>Simon Fraser University</td>
</tr>
</tbody>
</table>

Copyright © 2017 The Authors

This document may be used for private study or research purpose only. This document or any part of it may not be duplicated and/or distributed without permission of the copyright owner.

The Singapore Copyright Act applies to the use of this document.


This document was archived with permission from the copyright holder.
Temporal Analytics with Discourse Analysis: Tracing Ideas and Impact on Communal Discourse

Alwyn Vwen Yen Lee  
Nanyang Technological University, Singapore  
50 Nanyang Avenue  
Singapore 639798  
alwynlee@ntu.edu.sg

Seng Chee Tan  
Nanyang Technological University, Singapore  
50 Nanyang Avenue  
Singapore 639798  
sengchee.tan@ntu.edu.sg

ABSTRACT
This paper presents a study of temporal analytics and discourse analysis of an online discussion, through investigation of a group of 13 in-service teachers and 2 instructors. A discussion forum consisting of 281 posts on an online collaborative learning environment was investigated. A text-mining tool was used to discover keywords from the discourse, and through social network analysis based on these keywords, a significant presence of relevant and promising ideas within discourse was revealed. However, uncovering the key ideas alone is insufficient to clearly explain students’ level of understanding regarding the discussed topics. A more thorough analysis was thus performed by using temporal analytics with step-wise discourse analysis to trace the ideas and determine their impact on communal discourse. The results indicated that most ideas within the discourse could be traced to the origin of a set of improvable ideas, which impacted and also increased the community’s level of interest in sharing and discussing ideas through discourse.

CCS Concepts
• Applied computing → Education → Computer Assisted Instruction.

Keywords
Learning Analytics; Discourse Analysis; Temporality; Social Network Analysis; Idea Measurement

1. INTRODUCTION
Learning analytics as a nascent field has been gathering broad interests in educational research and practice [13], and recent research has contributed significantly to further our understanding of learning processes [30]. Discourse analysis has long been conducted to analyze language beyond literal language use, and detailed analysis of online discussions can inform the design and improve the productivity of knowledge creation [6]. The analysis of discourse over time can provide additional insights of agent interactions and specifically students’ level of understanding in discussions. However, this form of temporal analysis is often an undervalued process that is not studied in greater detail [18], resulting in broad and ambiguous conclusions to be drawn from studies. By combining the usage of discourse analysis with temporal analytics, we propose to integrate data sources and analysis from interactions and discourse among students during learning activities. The results obtained from such analysis could provide clearer indications on how temporal analysis of discourse could provide feedback that is useful for learning.

To conduct an analysis of temporal features within group interactions, emergent methods and measures are required for characterizing temporal dynamics [24]. Usage of networks, especially social networks can provide substantial insights into the quantity and types of relationships that exist between participants of the network. The problem of analyzing and understanding complex networks is understandably a challenging one, but it could also be seen as an area of research that contains utmost potential. In this regard, social network analysis (SNA) [29, 32] has become one of the most commonly applied methods in learning analytics research [4]. With the prevalent usage of online databases and ease of access to an abundance of data, little temporal data are being used by analysts to provide a deeper understanding of the content [24]. Even with the abundance of platforms and tools that are used with SNA [17], few studies can make full use of network analysis, and there has yet to be significant breakthroughs that assist researchers in furthering the understanding of learning.

By using novel methods in conjunction with SNA [17], students can learn deeply to understand the meaning behind written text and spoken words [1] through feedback and assistive tools. The standard social network measures (eg., degree, closeness, betweenness, centrality) can represent learning processes and interactions within the group to a certain extent using network graphs, but they are restricted in the ability to provide the descriptive and quantitative feel of individual student’s or a community’s level of engagement. There is a limit to what SNA can achieve by inference, as most network variables and relationships are heavily dependent on user or analyst’s interpretations and can be tedious and challenging especially with large scaled-up networks. Also, research findings have so far been unable to robustly conclude or determine which types or combination of network measures can predictably determine levels of understanding within the group [32]. Therefore, by making use of the current prevalence of learning analytics deployed alongside discourse platforms, SNA can be combined with temporal analysis to determine sections of discourse where feedback and interventions could prove to be most helpful in helping to stimulate and increase students’ engagement in the discourse. By seeking to automate the procedure of discourse analysis with temporal analytics, larger social networks can be tackled and similarly analyzed without requiring educators to sift through all the contents within the discourse manually. The following research question is to be addressed: How could temporal
analytics assist discourse analysis in tracing ideas and impact on discourse to gauge students’ and community’s level of understanding?

2. BACKGROUND

2.1 Social Network Analysis

Social Network Analysis (SNA) is the process of investigating social structures using network and graph theories, and has been adopted since the 1930s [29], with the focus mainly on three standard social network measures: degree, closeness, and centrality [10]. The degree of a graph indicates the number of edges incident to a vertex or referred in this paper as a node. Metrics such as closeness centrality is often used to denote average lengths of the shortest path between all nodes in the graph. Studies have shown that it is often unclear how these network measures can be associated with learning outcomes, for example, whether a higher social centrality could lead to higher academic performance [10, 32]. In addition, it can be difficult to generalize models and try to apply them to scaled-up experiments or deployments within schools.

Advancement in technology has provided an array of newer technology and tools that can be used with SNA [17], which were previously technologically unfeasible but now integrated into social network analysis. By analyzing the interactions as key measurements within the social network, educators who have been using online platforms for teaching (e.g., online discussion forums, learning management systems) could then also focus on the content that students generate as part of their learning process. As the majority of online content are mainly text, it is only natural that content analysis programs such as text mining and image comparison are integrated to assist in deeper analysis. Further, there are works that made use of neural networks [3], which are computer systems modeled on the human brains, and fuzzy clustering algorithm [25] that allow data to belong to two or more clusters. These are possible content analysis methods, indicating the large range of technology that could provide assistance in the analysis of educational data using more efficient manners. However, the results from these analyses tend to be inconclusive in helping us to answer our research question of how ideas can provide the meaning of context and impact discourse within a community. As most current tools and analysis platforms perform specific roles in summarization work [17], they are therefore not suitable to be used for analysis and scaffolding for deeper learning.

2.2 Online discourse analysis

Discourse analysis has long been a framework and means of exploring the imbrications between language and social-institutional practices [9]. Pioneers Gilbert and Muckay [11] started by examining scientific dispute and analyzed how claims were made and positions were defended. Discourse analysis has since evolved with different interpretations, but the focus has always been explicitly on language as a social action, with discourse and argumentation as tools which participants use to compare thinking, shaping of agreements, exploring and identifying of ideas [7]. With technological advancements, written discourse has been moved onto online platforms and expressed within online learning environments, with accompanying analysis of online discourse becoming much more prevalent and ubiquitous.

2.3 Temporal analysis

The emergence of temporal analytics provides a different dimension in which data could be analyzed and also allows researchers to pay attention to details that transpire in parallel to other main events within the timeline. A common understanding of how time entails for different parties involved in a social network is, therefore, crucial for navigating through series of events that unfold over time [19], and procedures often require segmentation of the timeline to be analyzed. There is also a need to bridge the gap between temporal details which are mostly collected at the micro level [18] and underlying theory which operates on a macro level. Temporal patterns [5] are also able to provide insights of students’ actions from clickstream data, showcasing the potential for determining predictive actions and potential interventions.

The process of analyzing ideas is, however, complicated, could consist of events and content that overlap with each other and transpire over short periods of time. The commonly used “code and count” methods that aggregate results over time [31] tend to ignore fine-grained details and would, therefore, mask minute yet possibly significant details that could indicate the level of understanding within an individual or community at a particular time juncture within the discourse. Moreover, understanding of context within any discourse is heavily dependent on the context and is often not generalizable [12], especially when the textual content is framed within a certain structure that requires deeper analysis so as to understand the intentions and motive of the writer truly. Hence there needs to be an integrated and scalable method where ideas can be automatically identified, be verified to be part of learning outcomes as pre-determined by the teachers, while not losing contextual information during the process of determining ideas within the discourse.

2.4 Analyzing Ideas and Relevant Tools

2.4.1 Definition and Focus on Ideas

John Locke used the word idea to represent the most basic unit of human thought, and these immediate objects of perception are interesting to him as they point beyond themselves [16]. The general concept of an idea is more than just a unit of thought, but one which allows room for improvement. The crucial feature in ideas is not so much about the content of ideas, but rather what the ideas are capable of doing, such as providing the epistemic function to represent something else, and the ability to improve beyond itself in order to provide a different higher level of understanding.

We propose to understand the effects of ideas on discourse, as knowledge is mediated by ideas and the search for improvable ideas subsequently leads to the creation of new knowledge. Within a traditional instruction-based classroom, most ideas within the community are provided by teachers and are also mostly limited in improvability, due to the factual nature and pre-assigned sources of content, such as structured syllabus designated by authorities and institutions. Communications within classrooms are also often constrained by a one-way instruction from teacher to students, with little room for feedback or inquiries from students, due to limited time for cramming of syllabus into the academic year. Therefore, by identifying and searching for improvable ideas, we are striving to discover communally interesting ideas present in the community discourse that can be improved through sharing of knowledge and may prove elusive to the human eye but can be detected using temporal analytics with social network analysis in a time-efficient manner.

2.4.2 Engaging Ideas through Knowledge Building

Ideas are complex entities, representation of opinions and knowledge, especially difficult to interpret when represented by textual format within online discussion communities. Problems that have to be addressed are also difficult to solve, and the use of either quantitative or qualitative approaches alone is inadequate to address this complexity. Small talk (30%) and logistic information
are major sources of information overload on discussion forums which do not enhance learning experience [2], and hence there is a need to identify and filter out main ideas from discourse, which are essential for maintaining interest within the community.

In the present study, we choose to focus on using knowledge building [27] as our model of knowledge creation, among other models such as the transformation of activity systems and the creation of new products at organizational level [8, 20]. Knowledge building is a pedagogical approach that engages students in continual production and improvement of ideas useful to the class community through the collaborative efforts of its members [27]. Using knowledge building, the learner’s capability of natural idea generation can be leveraged for continual improvement of ideas, and for teachers to maintain student engagement in idea improvement. The identification of ideas within discourse is a natural approach towards efficient learning, and discourse itself often plays a creative role in encouraging improvements on ideas [14]. We acknowledge the concept of idea improvement [26] as being crucial to students who are often required to have motivation in their search for solutions to handle significant challenges. Students have to be aware of emergent themes of inquiry from multiple sources of inputs, acknowledge knowledge gaps, and participate in collaborative idea improvement [34].

Therefore, to ensure that students continue to stay engaged in the learning process throughout their learning journey, they are encouraged to participate in knowledge building discourse to share and continuously improve their ideas. This paper investigates and analyzes content from knowledge building discourse, which allows learners to engage in collaborative inquiry to enable creation, contribution, and advancement of community knowledge. By engaging students in knowledge building discourse, students could collaboratively build on each other ideas so as to meaning-make and also improve the community’s level of understanding through deeper learning.

2.4.3 Analyzing Ideas using SNA and Keywords
To analyze temporal aspects of knowledge building discourse, we used an analysis platform, Knowledge Building Discourse Explorer (KBDeX) [22]. KBDeX was developed with technological affordances for visualizing student interaction networks, discourse unit networks, and keyword networks. As KBDeX was created with SNA in mind and the backend analytical processes are catered for this specific purpose; standard social network measures are offered and measurements such as density and centrality are often chosen for determining the level of interaction between students [33]. KBDeX is a unique analysis platform that allows users to conduct a step-by-step analysis of social networks through visual displays of networks and the respective social measures. This unique step-wise functionality of assessing discourse allows us to pause the discourse at any time junctures of the discourse, conduct the necessary analysis before moving onto the next timeframe of discourse. Other than visual indicators and prompts to assist users, standard social network measures can be retrieved for a more detailed form of analysis that provides insights into the interactivity between different nodes and related networks. Overall, the role of KBDeX is to provide a discourse analysis platform that fulfills the role of idea identification and also provide step-wise analysis of a continuous discourse through social network analysis using keywords.

As the commonly used type of SNA is insufficient in examining community knowledge advancement through students’ collaboration and interaction networks [21], text analysis processes such as identification of fundamental or advanced usage of keywords becomes increasingly and specifically important in determining ideas in discourse. This is so since keywords can be representative of knowledge that is recognized and understood by the students. At times, the usage of keywords within text might not be wholly representative of a student’s thoughts and ideas, or could also be misrepresented by students due to incorrect conceptual understanding. However, when these individual opinions are situated within discourse as part of the community’s sequence of thoughts, the usage of keywords and content can be monitored to understand further the overall usage flow of keywords and how they are used within the discussion threads in response to their peer’s ideas and thoughts.

Therefore, it is important to seek out tools that can determine keywords from a text. SOBEK is a tool that can identify relevant keywords and concepts within the text, through the frequency analysis of textual material and selection of important related keywords [23]. Text mining is often used for extracting relevant information from unstructured data, which in this case, from sources such as the knowledge building discourse. Graphs are constructed with connections between related nodes, for users to better understand relevant terms, with frequency and sizes of nodes indicating frequency of usage. Similar concepts are also run through an inbuilt thesaurus to consolidate common terms and ideas into a single node. The primary purpose of implementing SOBEK within this present study is not only for displaying graphs visually to users but also mainly for determining keywords from a statistical perspective, to be provided as inputs to social network analysis using KBDeX.

3. METHODS
3.1 Dataset and Setting
The dataset involved in this study was obtained from the discourse of a single graduate-level course, delivered by two instructors over a period of 13 weeks. The graduate class consists of 13 in-service teachers who were undergoing further training and professional development. Participants were encouraged to actively share their thoughts and opinions using a computer-supported collaborative learning environment called Knowledge Forum [28]. A total of 281 Knowledge Forum notes, referred as discourse units (DU) in this paper, were written with each of the 13 students providing at least a note of a reasonable amount of content. The role of the two teachers was to facilitate the lesson and co-create knowledge. Over the span of 13 weeks, students learned and applied the basic principles of knowledge building, and shared with the community on how knowledge building could be applied to affect, improve and apply to future learning opportunities for themselves and classes that they would instruct.

Each weekly lesson of three hours would consist of discussions and sharing sessions regarding readings that would be prepared for lessons. Instructors are available to provide face-to-face instruction during lessons and can provide consultation if help is required in understanding topics. The participants were able to use scaffolding and analytical tools to help each other improve ideas and to also reach a greater level of communal understanding. The participants were provided with credentials to use Knowledge Forum throughout the duration of the course, as this was the principal way in which teaching staff interacted with them during the weekly instructional sessions and outside of classrooms. This asynchronous mode of learning is a means of learning that allows the participants to engage in discussions transcending time and space, also ensuring knowledge building as a continuous process is
not restricted within lessons but can also be conducted outside of classrooms with inputs and considerations from authentic learning situations.

3.2 Procedures and Measures

3.2.1 Using SOBEK for finding keywords

Our previous design made use of expert opinion for determining keywords [15]. For the present study, we used text-mined keywords to represent ideas which are not obtained from domain experts, that is, the instructors for the course. There are a few reasons for following this course of action: the possibility that there could be inherent bias from experts regarding the field of knowledge when deciding keywords; the scope of knowledge regarding the topic could instinctively be narrowed down to only a specific range of keywords which the instructors would want the participants to learn; and keywords determined by experts often originate from their own perspectives and experience, often neglecting keywords that are representative of students’ views and opinions within discourse. The text-mined keywords would provide an unbiased source of ideas obtained from the participant-generated discourse. SOBEK allows us to filter out common words such as ‘a’, ‘the’, ‘with’, related prepositions and adverbs to ensure important keywords and concepts are captured during the analysis. Similar concepts represented by different words were also consolidated under a common keyword so that there would be less repetitive synonyms, while relationships between different keywords are also linked together to indicate inter-keyword relationships.

3.2.2 Using KBDeX for Social Network Analysis

Text-mining of discourse data using SOBEK would provide a list of keywords that could be provided as input for KBDeX to conduct SNA. KBDeX provides social network analysis on three different kinds of discourse data: (a) the participating students and teachers, (b) discourse units, and (c) the identified keywords within discourse. These three categories of data are considered as nodes within their respective networks, such as student network, discourse unit network, and keywords network. As time is a continuous variable within discourse analysis, it is inevitable that it has to be broken down into discrete segments to compare and contrast different portions of the discourse. KBDeX assists with this segmentation process by designating discourse from different agents into turns, such that all participants are included in a turn-based discussion. Analysts are therefore able to determine the total number of discussion units within the discourse, the initiating time of specific discourse units and the related keywords which are highlighted in color, together with the actual contents of the unit, all within a single view (see Figure 1).

As this study analyzes the relationships between sets of keywords within discourse units, we use the discourse unit network within KBDeX for analysis, with the help of underlying bipartite graphs that determine relationships between discourse units and keywords. For example, discourse units A, B and C in the discourse unit network contain texts of “What is learning?”, “Learning is a process.” and “I understand knowledge building.”. If “learning” is an identified keyword, discourse units A and B will be connected through the keyword. This sequence is then repeated for all discourse units and keywords in the respective discourse unit and keyword networks, to form a bipartite graph.

3.2.3 Using Betweenness Centrality Measure for Analyzing Temporality

The present study uses a conventional network measure, betweenness centrality (BC) coefficient to detect and identify important connections between words inside discourse units. We interpret the BC coefficient as the degree of importance of discourse units within discourse and consider the ideas within these units as relevant to the context and forum discussions at the point in time. Through graphical displays, we can identify and determine time junctures where the ideas are initiated, shared or fade over the entire discourse. By viewing the BC measure of a specific discourse unit that spans over multiple time junctures, the observed trend of BC values is representative of the unit’s significance and the subsequent impact on community discourse after its introduction into the discourse space.

4. RESULTS AND DISCUSSIONS

There are two major aspects of analysis in this study: 1) the use of auto-mining tools to help determine keywords and relationships within discourse, and 2) using keywords to find promising ideas, and making use of temporal analytics to assist and provide additional insights of the promising ideas in discourse analysis.

4.1 Keywords and Relationships

The discourse was collected and anonymized at the end of the course before being parsed through SOBEK for determination of the keywords. We used a frequency threshold setting of 50, which sets the criterion to display keywords that appear more than 50 times within the discourse. This setting was chosen, because it provides an average of 10 to 15 relevant keywords for the whole discourse, and this number is comparable to the number of keywords that experts provided to us in previous studies [15]. A total of 11 keywords were found, as shown in Table 1 with their respective frequency counts throughout the discourse.

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Frequency</th>
<th>Keyword</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>knowledge</td>
<td>325</td>
<td>kb</td>
<td>273</td>
</tr>
<tr>
<td>learning</td>
<td>172</td>
<td>students</td>
<td>151</td>
</tr>
<tr>
<td>knowledge</td>
<td>151</td>
<td>understanding</td>
<td>114</td>
</tr>
<tr>
<td>building</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>discourse</td>
<td>82</td>
<td>community</td>
<td>79</td>
</tr>
<tr>
<td>idea</td>
<td>67</td>
<td>based</td>
<td>63</td>
</tr>
<tr>
<td>information</td>
<td>55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Segmented Discourse View in KBDeX.

Table 1. Keywords with frequency from SOBEK
These keywords were used as indicators to determine ideas that students are generating and discussing regarding the topic of knowledge building. The usage of keywords in discourse often represents inquiry to understand the concepts behind keywords further, and could reflect learners’ possession of knowledge. However, it could also be construed as misconceptions. On the one hand, keywords could be unique words which are unlikely to be repeated without sufficient evidence of understanding. On the other hand, misconceptions could complicate processes of understanding, especially if it is propagated within a community without verification. Through discourse, misinterpretations can be corrected through knowledge sharing and critique by other learners, or through corrective guidance from the teacher.

The relationship between text-mined keywords are displayed in the graph (Figure 2) and the graphical interpretation shows that ideas related to knowledge building as an approach to learning are naturally linked to the following keywords ‘knowledge’, ‘learning’ and ‘community.’ A closer qualitative look at the content in discourse units show that ideas regarding the understanding of ‘knowledge building’ are present within the discourse, and this is representative of the participants’ understanding regarding this field of knowledge. From the keyword graph, we can determine that participants’ perspectives are varied and the graph encapsulates a rudimentary level of understanding on the assigned topic.

We made further adjustments to SOBEK so as to determine if more advanced concepts and principles, such as knowledge building principles (e.g., ‘improvable ideas’ and ‘rise above’) are present within the discourse. These are some of the principles explored by researchers and were not present or linked within the initial graph in Figure 2. As there are interests in discovering the presence of such principles within discourse, we subsequently relaxed the frequency criterion in SOBEK to a frequency threshold setting of 15, that is, any keywords with at least 15 appearances within the discourse will appear in the graph. The keywords ‘idea improvement’ and ‘rise above’ were eventually found in our analysis, and ‘idea improvement’ was illustrated with linkage to the main keyword ‘kb’ with a frequency count of 16, which explains the relatively smaller size of the node (see Figure 3). Among a large amount of keywords in discourse, the presence of the ‘idea improvement’ principle was largely obscure and hidden in Figure 2, but with an expanded search, the principle can be found as shown in Figure 3. Hence we question if it is possible to detect lesser discussed but nevertheless important ideas using a more optimized method such as text mining for keywords, as compared to casting the net wide and capturing less important keywords that might deviate discussions. By using temporal analysis, previously unnoticed ideas can be found in an alternatively more visible manner, such as through analysis of BC trends over time.

Figure 2. Graph of relationships between keywords text-mined from discourse.

In addition to the process of relaxing this criterion during analysis, we also noticed that the number of identified keywords increased significantly and there are not many conclusions that could be made about the larger keyword network, apart from indicating presence or absence of keywords and possible ideas. We recognize that even though the analysis of social networks based on keywords provided relationships and detects presence of ideas within discourse, the process of identifying keyword relationships was, however, unable to adequately show the process of improving ideas and deeper understanding of students within discourse. Therefore, there is a need to conduct a more detailed analysis such as using temporal analytics to provide greater insights and also track the improvement of ideas throughout the entire discourse.

4.2 Analyzing Temporality of Discourse

By focusing on the temporal dimension of discourse, we were able to observe how the content, usage of language and connectedness of discourse units within the discussion space change over time. Although the frequency of keyword usage can be tracked, we focused on the usage of BC trends of individual discourse units. Prior work such as the FA methodology [15] has shown that ideas within discourse can be identified and classified to give an indication of the communal understanding and an approximate number of ideas within discourse. In essence, FA methodology involves visualization and step-wise analysis of discourse network measures using SNA and measurement coefficients of chosen keywords to trace the evolution of ideas within a discourse space. This methodology was validated with a qualitative analysis but previously only contains a brief temporal analysis. In this paper, we seek to provide a more detailed analysis using the present study, to show that identified ideas with potential to improve can be further worked on through tracking of ideas and understanding of their BC trends. With the understanding of BC trends, we are able to show how a certain idea was influencing or was influenced by other ideas, and subsequently how its introduction into the discourse impacted overall community discourse.

4.2.1 Finding promising ideas with FA methodology

Firstly, we applied the FA methodology with the help of KBDex, in searching through the whole discourse to identify groups of ideas. The methodology then classified the ideas according to their level of promisingness, a combined measure of the following attributes, namely 1) idea relevancy to context, 2) underlying motive of idea initiation, and 3) impact of idea on the community. These attributes were measured from the perspectives of participants within the community, be it students who are consumers of content or teachers who help co-create knowledge in the community. This methodology was similarly applied to the

Figure 3. ‘Idea improvement’ (highlighted rectangle linked to ‘kb’) within large network of keywords.
current study to obtain indications of promising ideas. Discourse units with high relevancy and significantly high community interest displayed significant peaks in their respective BC trends with consistent fluctuations, and are classified as containing promising ideas (see Figure 4). The y-axis in Figure 4 denotes the normalized BC values, while the x-axis indicates the number of discourse turns in the whole discourse.

Discourse units in this paper are named with ‘DU’ followed by a number that represents the turn in which the discourse unit was introduced into the discourse. We identified discourse units that contain promising ideas using FA, and choose to discuss a few such as DU6, DU18 and DU66. These discourse units are frequently engaged by the community and consistently garnered community interests during the discourse. The ideas were traced to understand how and when were they initiated, how did the ideas change over time, and whether the ideas created impact or influenced other discourse units within the discourse.

On the one hand, with regards to the first phase, where the idea was initially introduced and the BC values peaked at a relatively high value, we know that there is an immediate interest by the community in the contents of the mentioned DU that could lead to further debates. The ideas within this kind of DU tend to be thought-invoking, novel or disruptive, which in any case naturally leads to more discussions among a larger part of the community. An example that exhibits this type of BC trend is DU66, which was introduced in the early-middle part of discourse at turn 66 but was able to invoke interest from the community rapidly.

On the other hand, if the DU does not exhibit any significantly drastic BC changes when the discourse unit was initially introduced into the discourse, it is safe to then label the contained ideas within the DU as the typical type of assertion that does not require too much attention by the community at that particular juncture in the discourse. However, subsequent increasing BC values and trends that arise in later parts of the discourse mean that there was a delayed pick-up in communal interests, signaling a possible return of interests to a previously discussed topic or concept. Both discourse units DU8 and DU18 exhibit such similar behavior, and are therefore classified as having low initial community interests upon introduction but contain significantly promising ideas that impact later stages of the discourse.

We show DU8 in Figure 5 with low initial community interests when initiated at Turn 8, and significant communal interest picked up later during the discourse at Turn 24. DU66, however, initiated at Turn 66 and immediately registered a spike in communal interests when the discourse unit was introduced.

The two different types of BC trends belonging to DU8 and DU66 suggest that the temporal pattern of communal interests in ideas could be predicted using social network measures such as BC. We verify the observations in this study by scrutinizing the qualitative contents of the DUs to determine if the nature and contents of the DUs are consistent with how the community perceives the discourse units. As both DU8 and DU66 are lengthy and descriptive in nature, the important parts of the discourse units are mentioned in the following quotations.

**Excerpt of DU8 from Student S7:** knowledge is a very complex entity. philosophically, knowledge refers to justified true belief. in other words, knowledge is a form of beliefs which one may possess. what differentiates knowledge from a belief is the process one takes to reason to oneself why s/he thinks that this belief is true. ....... Is also called propositional knowledge, there exist other forms of knowledge, ...... learning refers to a process where one acquires new knowledge with respect to the learner ...... knowledge building (kb) can be seen as a way of learning, where learners gain new knowledge in the process. this happens in a community setting, where students come to gather to generate their inquiry, contribute their different views (questions, ideas, theories, etc.) about the topic of interest, it is an interactive process where students may perform individual research before coming back to revisit some of these ideas that were originally proposed, the purpose is to improve the ideas, or perhaps if i may call it knowledge to the group of learners. ......
DU8 was written by student S7 in response to generic questions that were posed to the learning community, regarding knowledge, learning and knowledge building. The general concepts were explained with details using examples and DU8 was considered a well-written note, despite being initially uninteresting to the community due to its lengthy response and the community’s lack of familiarity and metaknowledge about the topic in the early stages of discourse. However, the community started paying more attention to DU8 after subsequent DUs (DU 21 to DU26) consisting of inquiries from the community required a clear explanation, and students started seeking for answers and explanations from DUs that transported earlier in the discourse. We detected students retrospectively referring to DU8 after Turn 26, resulting in a higher BC trend for DU8, even though there were other similar DUs that were present in later stages of the discourse. Students realized that other students’ DUs were unable to adequately address their inquiries and concerns as succinctly and clearly as S7’s note (DU8). Another discourse unit that we analyzed is from Student S6 and the following excerpt is quoted summarily from DU66.

Excerpt of DU66 from Student S6: in the first lesson of the course, I posted the commonly asked question of “how different or similar are knowledge sharing, knowledge construction and knowledge creation? Subsequently, we were being tasked to read the two articles … further questions evolved after the reading: (1) how different or similar is justification of true beliefs (knowledge) from validation of knowledge? ….. (2) what are the differences and similarities of authentic knowledge building and creation of professional knowledge? Eventually, T1 mentioned in the second lesson that kb and kc are equivalent, however, how are they equivalent? Why are they equivalent? In the pursuit of deeper understanding, I read an article “distinguishing knowledge sharing, knowledge construction, and knowledge creation discourses that explains the differences between the terminologies … … personally, I think this question and line of thought would have implications on the implementation of kb in school settings where teachers have once explored the usage of discussion forums/bulletin boards, blogs etc. for instructional practice. How is kb distinctive from other pedagogies? How is KF different from other technological tools? Is the tool an essential or supplemental? these are practical questions on the “how” (implementation of kb), what do the rest of you think?

There was significant community interest in the discourse unit right from the moment in time when it was introduced at discourse turn 66. This level of interest was maintained throughout the discourse, leading us to believe that DU66 contained ideas and content that were sufficiently intriguing to the community while providing food for thought for the community. S6 pointed out differences in similar processes of knowledge sharing and knowledge building while encouraging the community with thought-invoking questions to further discuss and debate on Knowledge Forum. Upon further scrutiny, it was determined that the ideas and content within DU66 consist of deeper analysis built onto an earlier discourse unit DU25 by the same student S6. The student S6 had previously provided his/her perspectives in addition to thought processes, and have additionally conducted follow-up research to understand the topic in his/her own capacity further. The analyzed DU66 was written one week after the first discourse unit DU25, therefore serving as a report of S6’s findings over the week to the community, and also forms part of the documentation chain that S6 used for reflecting on his/her learning. The community acknowledged the effort put into DU66 in the following DUs by answering and discussing mentioned topics within DU66. Participants in the community were also challenged to improve on this group of ideas, by building onto S6’s perspectives regarding the topic, so as to further discussion regarding the practical implementation of knowledge building within the classroom.

Overall, the three analyzed discourse units (DU8, DU18 and DU66) attracted significant community interests in their content over the entire period of discourse within the study. Most ideas in the three discourse units either acted as sources which students continuously referred to, or served as inspiration for improving ideas within discourse. Even though DU8 and DU18 attracted delayed communal interest sometime after introduction in discourse, the ideas within the two mentioned discourse units were eventually broadly shared and built upon within the community. DU66 managed to invoke immediate interests from the community and sustained community engagement throughout the period of discourse, and this was sufficiently backed up by qualitative analysis of the discourse unit.

Temporal analysis of BC trends provided us with deeper insights of how ideas can be traced within discourse units and an alternative method of analyzing discourse. The usage of temporal analytics with discourse analysis aids us in further understanding of how ideas can influence and impact overall community discourse.

5. CONCLUSIONS AND FUTURE WORK

We examined the use of temporal analytics with a step-wise discourse analysis of an online discussion, which provided insights into the nature and evolution of ideas that were proposed by students. We used text mining techniques through SOBEK to obtain an unbiased set of keywords, which were used with social network analysis and FA methodology to identify and classify ideas in the study of student discourse. We traced the ideas and determined the impact of these ideas on subsequent discourse using the BC measure and temporal analysis of BC trends. Overall findings show that identified ideas within discourse can be improved through collaborative discussions on forums, and a more thorough understanding of ideas can be achieved through tracking and analyzing the impact of promising ideas on discourse.

We are currently building on this study with similar implementations of analytical tools and methodology presented within this paper on archived data. In addition, we intend to use data obtained from ongoing lessons and make use of outcomes obtained through analysis to provide educators with actionable feedback, so as to conduct interventions to improve online teaching and learning.

6. ACKNOWLEDGMENTS

The research reported here is supported by the Centre for Research and Development in Learning, Nanyang Technological University (CRADLE@NTU). The research team would also like to thank the teaching staff and students who participated in this study.

7. REFERENCES


