Chang, M. et al. (Eds.) (2019). Proceedings of the 27th International Conference on Computers in Education. Taiwan: Asia-Pacific Society for Computers in Education

Visualizing the language of teamwork

Elizabeth KOH[a*], Simon YANG[a], Dhivya SURESH[a], Yi Huan TEE[a] & Claire LEE[a]
[a]*National Institute of Education, Nanyang Technological University, Singapore*
*elizabeth.koh@nie.edu.sg

Abstract: As 21st Century problems require increasingly complex and creative solutions that are achieved through collaboration, the ability to work effectively in teams has become more important over time. To help students collaborate better, a digital formative assessment tool, My Groupwork Buddy, was developed to build teamwork competency using analytics. The scope of this paper focuses on the recent development of teamwork chatlog visualizations. This paper describes the building of the teamwork text classification system, system flow and two visualizations. Work is underway to evaluate the system, and future work includes refinements to integrate these new visualizations in the system to enable a more holistic formative assessment.

Keywords: Teamwork, discourse analytics, visualization, formative assessment

1. Introduction

As 21st Century problems require increasingly complex and creative solutions that are achieved through collaboration, the ability to work effectively in teams has become more important over time. Formative assessment is one key means to help students develop their teamwork. Utilizing the power of learning analytics, a digital formative assessment tool, My Groupwork Buddy (MGB), was developed. In MGB, four teamwork competency dimensions are focused on to help students during the process of collaborating in a team - coordination (COD), mutual performance monitoring (MPM), constructive conflict (CCF), and team emotional support (TES). These teamwork dimensions, based on extant literature on what good teamwork skills are, have been assessed using self and peer ratings and a visual dashboard from disposition analytics shown to trigger students' awareness and sensemaking (Koh, Hong, & Tan, 2018). A concurrent and more tedious development has been the assessment of teamwork competency dialogue from a chatlog (Suresh, Lek, & Koh, 2018). In this paper, we detail our most recent endeavor to visualize chatlog text in real-time of the four dimensions. The use of discourse analytics will complement the existing disposition analytics and provide a more holistic understanding of students' teamwork processes.

The next section explains the teamwork text classification system (TTCS). This is followed by an elaboration of the discourse (chatlog) analytics visualization process on MGB. A brief discussion concludes the paper.

2. Establishing the Teamwork Text Classification System

In order to generate a model that will automatically classify student chat data into the four teamwork dimensions, NLP techniques such as Named Entity Recognition (NER) and feature engineering were carried out followed by a multi-label classification. The TTCS has been developed using Python 3.7.x and the process is as follows.

First, data pre-processing is performed to handle the irregularities in the chat text. This includes emotions and punctuation tagging, chat abbreviation expansion, local terms replacement, and replacement of names in the text. The names were identified using Stanford NER model and then replaced with a generic {{NAME}} tag, thereby grouping similar features together, resulting in a better predicting classifier. The feature engineering process follows this; ten new features were created with context-sensitive rules using indicative terms dictionary, POS tagging and regular expressions. Next, using a tf-idf vectorizer from sci-kit learn, the messages are vectorized. All features extracted including the ten new features are passed to a classifier to carry out the multi-label classification (See Suresh, Lek,

& Koh, 2018). The classifier for our TTCS has been trained extensively and has the following benchmark metrics - precision: .855, recall: .651, accuracy: .701.

Through this development, TTCS has undergone revisions to enhance its reliability and execution time. In earlier versions, each step was designed as an independent module but in order to automate the process, all the modules were integrated. Also, an earlier version included a spelling correction module but this code was incompatible with 64-bit operating systems. This module was excluded and further edits were made to account for the exclusion. Additionally, the NER was a bottleneck due to the tagger.tag method. Therefore, revisions were made to reduce the execution time of this code.

3. Creating Chatlog Analytics Visualizations of Teamwork Competency

The MGB platform is built on NodeJS with VueJS as the framework. To create the chatlog analytics visualizations, two additional node packages were utilized - python-shell and highcharts packages. Figure 1 illustrates the flowchart of how MGB visualizes teamwork competency dimensions from chatlog dialogue.

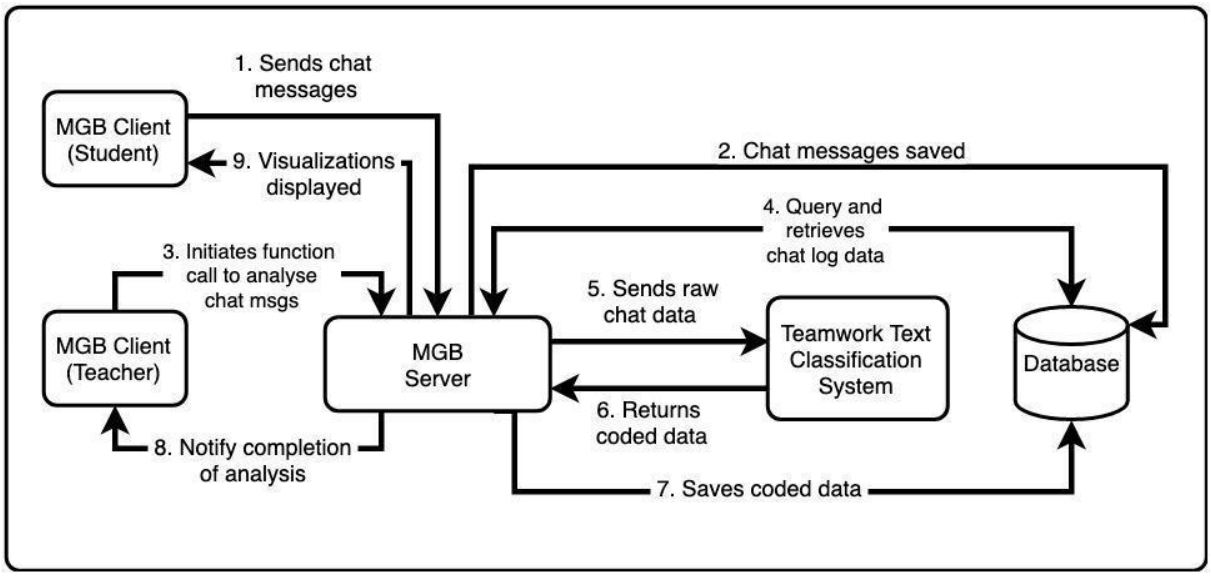

Figure 1. Flowchart of chatlog analytics visualizations

The process begins as students start chatting on the team chat on MGB. All chat messages are saved into the database as raw chat log data. To enable the display of the visualizations, the teacher performs a function call from a designated page in MGB. This enables the raw chat log data to be retrieved, formatted into a proper JSON format, and sent over to TTCS. Only uncoded chat messages will be sent while those coded, not coded again. The idea behind having a teacher to manually trigger the analysis process is to achieve better system performance and reduce the running of multiple calls to the python script.

TTCS will process the raw chat data as described in the earlier section. After all chat data are processed, each row of chat data will be coded into one or more dimensions, or none at all. The returned result consisting the chat message, the dimensions breakdown, the userid, the message timestamp and the coding timestamp will be formatted and sent back to the MGB server as a JSON string. The coded dataset is then saved in the database and the teacher notified of the completion of the analysis process.

Once coded datasets are available in the database, students are able to access a dashboard in MGB to view the visualizations. Two variations of visualizations have been designed: bar chart with slider (Figure 2) and streamgraph (Figure 3). For the former, the data is displayed with four bars, each bar representing a teamwork dimension from the student's chat text. The slider at the bottom is positioned from the earliest timestamp up till the latest timestamp of all the chat messages sent by the student. Depending on where the slider position is located, it shows the accumulated count of the teamwork dimensions up till that point in time.

For the streamgraph, the data is plotted against a timeline from the timestamp of the earliest to the latest sent messages. Each segment of the graph between two timepoints indicate the presence and count of the dimensions that are detected for the chat messages, as well as total messages count that falls within that time frame.

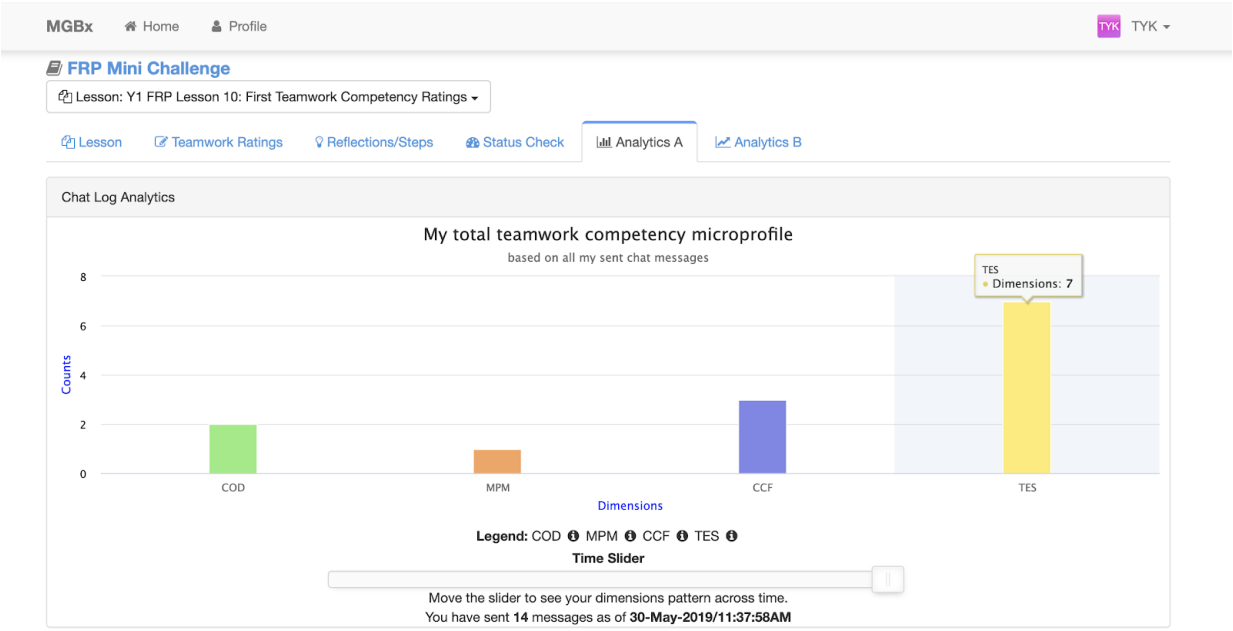

Figure 2. Bar chart with time slider

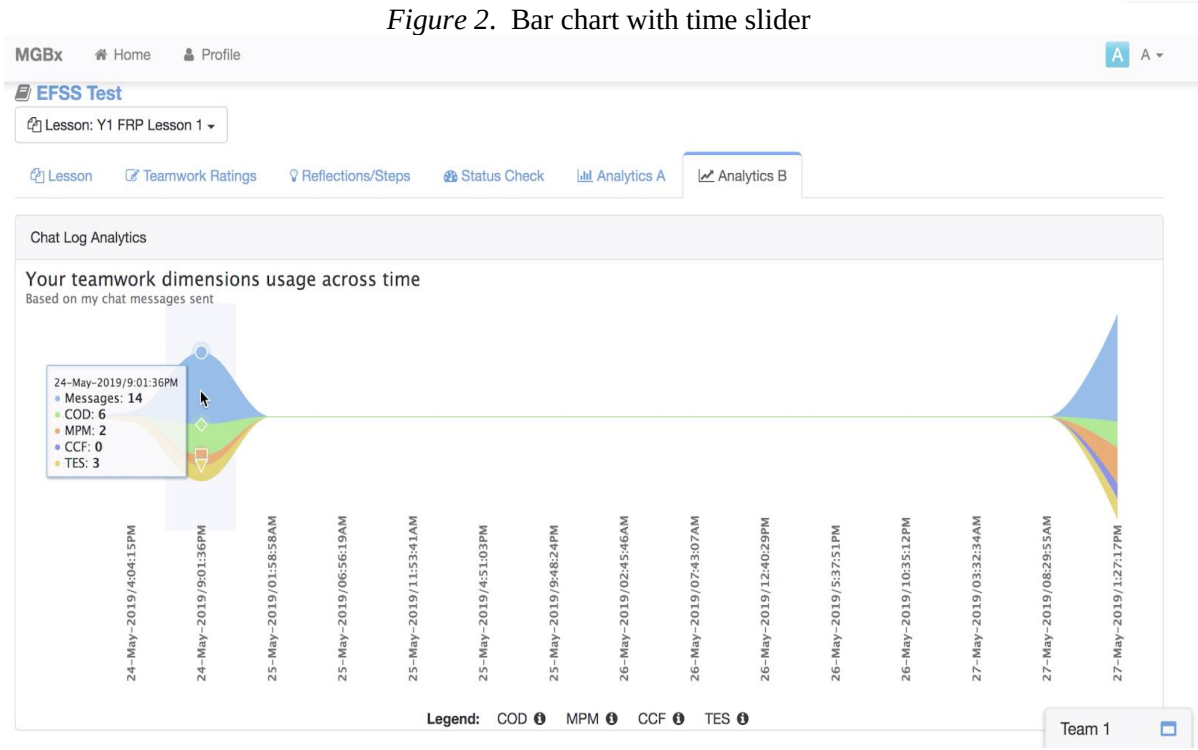

Figure 3. Streamgraph

These two visualizations were designed as they well represent the coded chat dataset of students' teamwork dimensions over time. The bar chart cleanly and clearly shows the cumulative counts for each dimension at one time-point, with the slider allowing for viewing of different time-points. On the other hand, the streamgraph shows more of the flow of chat messages, their dimensions over time and provides an overview of the intensity of the sent chat messages. The visualizations allow various interpretations of students' teamwork competency to trigger further reflections.

4. Discussion and Conclusion

This paper has made visible the language of teamwork as a means of formatively assessing students' teamwork competency. We have developed a classification system, process flow and two visualizations in order for students to learn teamwork skills based on the representations of their discourse on MGB. Students' are then made more aware of how their discourse contributes to working in a team and can learn to communicate better. An evaluation is currently underway to gather students' feedback to provide easy sensemaking and actionable steps for them. There is also need to investigate the degree to which each dimension is important, which could vary according to the tasks required of the team. While there is much future work (e.g., integrating the visualizations and other enhancements based on the evaluation), this study has successfully built chatlog analytics visualizations of teamwork competency making the digital formative assessment tool of MGB more holistic and timely. This design uses learning analytics and has broken new ground in teamwork assessment, which will add to research and practice in the assessment of 21CCs.

Acknowledgements

This study was funded by Singapore Ministry of Education (MOE) under the Education Research Funding Programme (OER 09/15 EK and DEV 03/17 EK), administered by National Institute of Education (NIE), Nanyang Technological University, Singapore. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Singapore MOE and NIE.

References

Suresh, D., Lek, H. H., & Koh, E. (2018). Identifying teamwork indicators in an online collaborative problem-solving task: A text-mining approach. In Yang, J. C. et al. (Eds.), *Proceedings of the 26th International Conference on Computers in Education* (pp. 39-48*)*. Philippines: Asia-Pacific Society for Computers in Education.

Koh, E., Hong, H., & Tan, J. P. L. (2018). Formatively assessing teamwork in technology-enabled twenty-first century classrooms: Exploratory findings of a teamwork awareness programme in Singapore. *Asia Pacific Journal of Education*, *38*(1), 129-144.