Fostering Interaction in Computer-Supported Collaborative Learning Environment

Pratiksha Virendra Patil
Indian Institute of Technology Bombay
214380005@iitb.ac.in

Ashwin T S
Indian Institute of Technology Bombay
ashwindixit9@gmail.com

Ramkumar Rajendran
Indian Institute of Technology Bombay
ramkumar.rajendran@iitb.ac.in

ABSTRACT

Interaction is the key driving force behind the critical processes involved in collaborative learning. But novice learners find it difficult to effectively interact during collaborative tasks and need support. Speech data from the collaborative discourse is significantly used to monitor and assess the interaction among students. Our research goal is to foster interaction in computer supported collaborative learning environment. The initial plan is to design and develop a system for capturing speech data from collaborative discourse in real-time, and derive various verbal and non-verbal features from speech data to automatically detect and assess the collaboration quality. Also, we are planning to design and implement the real-time automated feedback based on the data captured and investigate the impact of the feedback provided.

Keywords

Collaborative Learning, Automatic detection of collaboration, Speech

1. INTRODUCTION

Collaborative learning is one of the 4C skills, considered to be the most important 21st-century learning skill. Much emphasis is being provided on enhancing collaborative learning skills in students and the workforce [1, 2]. The construction of shared knowledge, negotiation/coordination, and maintaining team function are critical processes involved in collaborative learning. Interaction among students is the driving force behind these meaning-making processes [3, 4, 5]. Collaboration is key to learning but it is not easy for novice learners to effectively interact in a collaborative task. Lack of interaction is due to challenges like students not being open in accepting opposing views, asking for help, building trust and giving elaborate explanations or providing feedback [6]. Cognition, affect, motivation, and metacognition of individual participants and group members also plays a role in interaction [7]. Interaction in collaborative learning is also influenced by factors such as group dynamics, pedagogy, task design, and process of evaluation as well [8]. To

address the lack of interaction, appropriate support needs to be provided to the students. And such support can be provided only after a timely and accurate diagnosis of the challenges faced by students [4, 8, 9, 10, 11].

Generally, the quality of collaboration is measured across five aspects as follows,1) communication/appropriate use of social skills; 2) joint information processing/group processing; 3) coordination/positive interdependence; 4) interpersonal relationship/promotive interaction; and 5) motivation/individual accountability. These five aspects are further mapped to the nine dimensions of collaboration, those are expertise, dominance, coupling, reflection, roles, engagement, coherence, misconception and uncertainty. These dimensions are measured largely using self-reports or conducting tests. But it has its own limitations. Quality assessment based on observation can be leveraged to address these limitations. Monitoring students' discourse can give clarity on student understanding and challenges faced by them while working on collaborative tasks [12]. To monitor the collaborative discourse, the existing researchers are largely dependent on manual observation, transcription, and analysis to identify challenges faced by learners. This is a time-consuming and laborious process that also causes delays in feedback. Moreover, it puts limitations on scaling collaborative learning activities [13, 14].

Recently there are a lot of research studies focused on providing adaptive support in computer-mediated/online collaborative activities by monitoring discourse from forum posts, chats, and log data [15]. Some researchers also have used multimodal data from online meetings to assess collaboration and provide feedback [16]. However, research in automated monitoring of collaborative discourse in physical spaces (collocated collaboration) to assess collaboration dynamically in real-time and provide adaptive feedback [17] is still in a nascent stage and emerging rapidly with the advancement in sensor technology and in the field of multimodal learning analytics. Multimodal data like gestures, posture, eye gaze, content, log data, self-reports, spatial data, facial expressions, and physiological indicators [18] can help us measure important collaboration indexes such as synchrony in
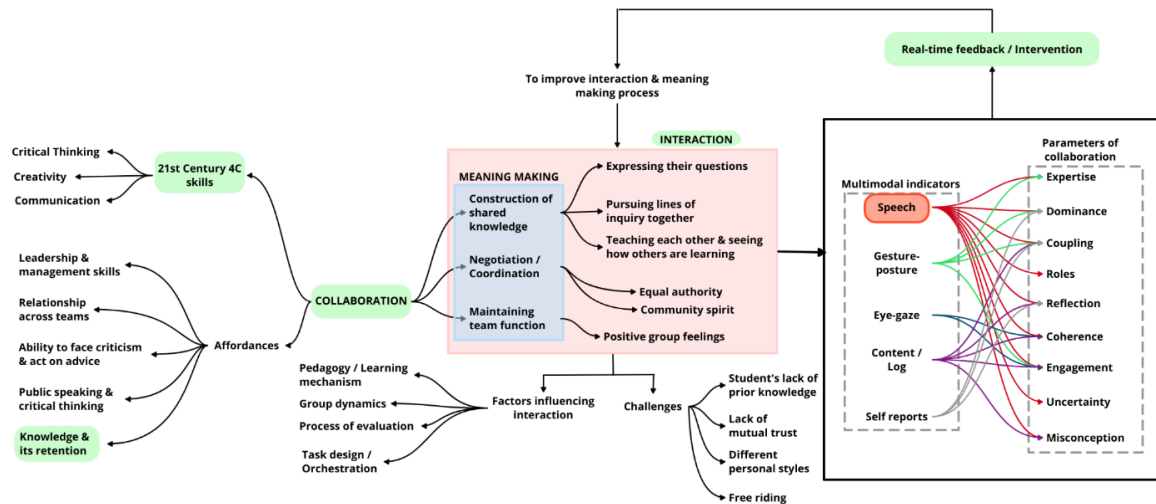
Figure 1. Conceptual representation of the proposed work.

members of the group, equality of contribution, information pooling, and so on. For example, the rise and fall of average pitch, intensity obtained from the audio signal; body position, head direction, pointing, using both hands, joint visual attention, etc.; these indicators can be used to measure synchrony. High synchrony indicates good collaboration. Total speaking time and the number of ideas and questions raised can tell us about equality in participation, which is indicator of good collaboration. Web search is the indicator for information pooling, similarly, different multimodal indicators are mapped to different collaboration indexes. These indexes can further help us understand high-level collaboration parameters like expertise, dominance, coupling, coherence, roles, reflection, uncertainty, misconception, and engagement. It is also evident that all these aspects of collaboration can be significantly detected using speech modality [18]. Mapping of audio indicators with different parameters of collaboration is shown in Table 1.

Table 1. Mapping of audio indicators to collaboration parameters [18].

Collaboration parameters	Collaboration Indexes	Audio Indicators
Expertise, Dominance,Coupling,Reflection,Roles,Engagement,Coherence,Misconception,Uncertainty	Synchrony	rise and fall of average pitch,intensity, amplitude
	Equality	jitter,total speaking time
	Mutual Understanding	dialouge management,verbal discourse,statements,questions

In the speech, the existing works used verbal and non-verbal features to understand various aspects of collaboration. It is clear that cognitive and socio-emotional interaction can be analysed significantly using speech. There are works related to understanding the semantics of the interactions and the automation of the same is not explored much in the literature. Moreover, capturing multimodal data in a live classroom to assess the quality

of collaboration dynamically in real-time and providing adaptive feedback is a challenging task. Such feedback can help students collaborate better in face-to-face settings [19]. Also, it will help in reducing the cognitive load on teachers/facilitators and enable them to effectively conduct collaborative classes on a large scale [16, 19, 20]. From the existing studies, we observe that very few works attempted automatic detection of collaboration using multimodal data and specifically using speech alone. And those are limited to very few aspects of collaboration. There are a lot of scopes to leverage non-verbal and verbal features of speech in the detection and assessment of collaboration. Non-verbal features like duration of speech, pause, turn-taking, pitch, jitter, intensity etc. can be used to detect low level collaboration indexes like equality and synchrony.

Non-verbal features extracted from speech can help in assessing collaboration quality in diverse contexts and tasks, while preserving the privacy of participants [28, 29]. On the other hand verbal/lexical features extracted from speech data captured during student collaborative discourse can be effectively used to create knowledge graphs and analyse them deeply to understand conceptual knowledge and transactivity between concepts [22, 23]. This understanding can lead to the design of effective feedback to foster cognitive interaction in the collaborative learning task. We are planning to address a few of these gaps in our proposed work. The conceptual framework of the proposed work is shown in figure 1. We are planning to design a system for capturing speech data from collaborative discourse in real-time, and derive various verbal and non-verbal features from speech data to automatically detect and assess the collaboration quality. Also to design and implement the feedback based on the data captured and investigate the impact of the feedback provided.

2. METHODOLOGY AND PROGRESS

In order to understand how students collaborate and also to learn the processes involved in our proposed framework, we conducted a study to capture speech data from a collaborative learning environment. We used data from 12 participants solving a programming problem with a shared screen using a python programming teaching environment. Students worked in dyads.

There were six groups in total. Before the study test was conducted to check the knowledge of participants of basic Python Programming. It had 12 multiple choice questions and 2 questions, for which they were required to write complete code. This study was conducted in a technology-enhanced collaborative learning classroom. (Refer to section 5.2)

2.1 Data Collection and Automated Speech Transcription

Introduction to the learning environment and all task-related instructions were provided by the instructor. Students had access to the study material while solving the task. Students' video and speech and log data were captured using the Open Broadcaster Software (OBS) installed on the computer they were using. Students' speech data captured using the OBS tool, is then converted into transcripts using a web-based tool for transcription, i.e., Otter.ai. Each group's audio files were automatically transcribed using the web-based Speech to text service. The service generates a transcript with a timestamp for every utterance spoken along with speaker diarization. It also provides a summary of keywords and the total speaking time of each speaker in percentage. It is observed that automated transcription has some limitations like, not capturing overlap between two speakers, jumbling with similar pronouncing words, and not fully transcribing long sentences. However, as we are interested in knowing who the speaker is and what topics are discussed, these errors will not have a significant impact [5]. We generated a transcript for an audio file containing spoken interaction between dyads. It was a 50 minutes collaborative problem-solving (Python Programming) activity. Further, we have coded the transcripts by marking co-occurrences of keywords in each utterance. We divided transcripts based on its timestamp into 5 scenes and compared the data for each speaker. In our study, we focused on the semantics/content of verbal data. We looked at the contribution of each group member in terms of keywords/concepts discussed by them, Speaker identification and timestamp in order to map turn-taking and overlap of speech. The Epistemic Network model for this data is created using web-based tool [21, 22, 23, 24]. Creating knowledge graphs to understand unfolding collaboration automatically in real time is a challenging task. Speech data needs to be converted into text/transcripts using NLP to understand the semantics of data. We will be using speaker diarization and text data mining to understand who is speaking, when they are speaking and most importantly what they are speaking- the concepts and linkages between the concepts. We are exploring how epistemic network analysis can be effectively used for this task and the need to incorporate social network analysis.

2.2 Audio signal processing for feature extraction

In order to detect collaboration indexes and asses collaboration quality based on speech interaction, we have extracted acoustic features like fundamental frequency, Mel-frequency cepstral co-efficients (MFCCs) and pitch from audio file (wav file). This audio file contain 45 minutes of speech data captured during collaborative design task, in which 4 students are working on concept map and facilitator is providing the instructions. Audio signal is pre-processed by sampling it at the rate of 44 KHZ [30]. Then it is segmented for fixed time window and features are extracted using librosa, package in python for audio analysis. We

have also clustered the data for different speakers using extracted features. Our aim is to use automated speaker recognition in multiparty audio files, segment the audio file for each turn of the speaker, extract acoustic and verbal features from the segments, and use these features to detect collaboration indexes.

3. CONTRIBUTION AND FUTURE WORK

We explored several studies that collect multimodal data from collaborative learning environments, we observe that most existing works analyse the quality of collaboration and some of them also provide feedback based on the assessment. After analysing the current literature we have identified certain gaps. Very few studies provide real-time dynamic feedback to the learners or facilitator, especially in the collocated collaborative learning environment. Most studies using speech modality to assess collaboration quality and provide feedback consider acoustic and non-lexical/ non-verbal features to classify/ detect collaboration using machine learning. This can lead to concerns about the reliability of the results. There exist no studies that provide real-time feedback based on learners' verbal cues and their interaction log data. Most of the studies focus on detecting and supporting dominance or coherence. There is no existing work to automatically detect other collaboration indexes such as uncertainty, misconceptions, etc. which can be better mapped using verbal features of speech. Current studies are providing feedback to foster socio-emotional interaction in collaborative learning. To address the research gaps, we aim to develop a system to automatically assess the collaboration quality in real-time. This system will capture speech data from collaborative discourse in real-time, and derive various verbal and non-verbal features from speech data using state-of-the-arts methods. Further it will segment and annotate the stream of data automatically to map it with different collaboration assessment indexes. Also we aim design and implement the feedback based on the real-time assessment and investigate the impact of the feedback provided.

4. REFERENCES

[1]   Chiruguru, Dr & Chiruguru, Suresh. (2020). The Essential Skills of 21st Century Classroom (4Cs). 10.13140/RG.2.2.36190.59201.

[2]   2015. PISA 2015 Collaborative Problem Solving Framework

[3]   Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

[4]   Gerry Stahl, Timothy Koschmann, Dan Suthers (2010). Computer-supported collaborative learning: An historical perspective. http://gerrystahl.net/cscl/CSCL_English.pdf

[5]   Stewart, A. E. B., Vrzakova, H., Sun, C., Yonehiro, J., Stone, C. A., Duran, N. D., Shute, V., & D'Mello, S. K. (2019). I Say, You Say, We Say. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), 1–19. https://doi.org/10.1145/3359296

[6]   Ha Le, Jeroen Janssen & Theo Wubbels (2018) Collaborative learning practices: teacher and student perceived obstacles to effective student collaboration, Cambridge Journal

of Education, 48:1, 103-122, DOI: 10.1080/0305764X.2016.1259389

[7] Taub, M., Azevedo, R., Rajendran, R., Cloude, E. B., Biswas, G., & Price, M. J. (2021). How are students' emotions related to the accuracy of cognitive and metacognitive processes during learning with an intelligent tutoring system?. *Learning and Instruction*, *72*, 101200.

[8] Dillenbourg P. (1999) What do yuo mean by collaborative leraning?. In P. Dillenbourg (Ed) Collaborative-learning: Cognitive and Computational Approaches. (pp.1-19). Oxford: Elsevier

[9] Kirschner, P.A., Sweller, J., Kirschner, F. *et al.* From Cognitive Load Theory to Collaborative Cognitive Load Theory. *Intern. J. Comput.-Support. Collab. Learn* 13, 213–233 (2018). https://doi.org/10.1007/s11412-018-9277-y

[10] Lin, Feng & Puntambekar, Sadhana. (2019). Designing Epistemic Scaffolding in CSCL.

[11] Yingbo Ma, Mehmet Celepkolu, Kristy Elizabeth Boyer. Detecting Impasse During Collaborative Problem Solving with Multimodal Learning Analytics. Proceedings of the 12th International Learning Analytics and Knowledge Conference (LAK22), 2022, pp. 45-55.

[12] Bressler, D. M., Bodzin, A. M., Eagan, B., & Tabatabai, S. (2019). Using Epistemic Network Analysis to Examine Discourse and Scientific Practice During a Collaborative Game. Journal of Science Education and Technology, 28(5), 553–566. https://doi.org/10.1007/s10956-019-09786-8

[13] Emma L. Starr, Joseph M. Reilly, and Bertrand Schneider ems,@mail.harvard.edu, josephreilly@g.harvard.edu, bertrand_schneider@gse.harvard.edu (2018). Toward Using Multi-Modal Learning Analytics to Support and Measure Collaboration in Co-Located Dyads. https://repository.isls.org/bitstream/1/888/1/55.pdf

[14] Nöel, René & Riquelme, Fabián & Lean, Roberto & Merino, Erick & Cechinel, Cristian & Barcelos, Thiago & Villarroel, Rodolfo & Munoz, Roberto. (2018). Exploring Collaborative Writing of User Stories With Multimodal Learning Analytics: A Case Study on a Software Engineering Course. IEEE Access. PP. 1-1. 10.1109/ACCESS.2018.2876801.

[15] Chua, Y. H. V., Rajalingam, P., Tan, S. C., & Dauwels, J. (2019). EduBrowser: A Multimodal Automated Monitoring System for Co-located Collaborative Learning. Learning Technology for Education Challenges, 125–138. https://doi.org/10.1007/978-3-030-20798-4_12

[16] Cornide-Reyes, H., Riquelme, F., Monsalves, D., Noel, R., Cechinel, C., Villarroel, R., Ponce, F., & Munoz, R. (2020). A Multimodal Real-Time Feedback Platform Based on Spoken Interactions for Remote Active Learning Support. Sensors, 20(21), 6337. https://doi.org/10.3390/s20216337

[17] Vogel, F., Kollar, I., Fischer, F. *et al.* Adaptable scaffolding of mathematical argumentation skills: The role of self-regulation when scaffolded with CSCL scripts and heuristic worked examples. *Intern. J. Comput.-Support. Collab. Learn* **17**, 39–64 (2022). https://doi.org/10.1007/s11412-022-09363-z

[18] Sambit Praharaj, Maren Scheffel, Hendrik Drachsler, Marcus Specht ; Literature Review on Co-Located Collaboration Modeling Using Multimodal Learning Analytics—Can We Go the Whole Nine Yards?.,IEEE Transactions on Learning Technologies,2021

[19] Lämsä, J., Hämäläinen, R., Koskinen, P., Viiri, J., & Mannonen, J. (2020). The potential of temporal analysis: Combining log data and lag sequential analysis to investigate temporal differences between scaffolded and non-scaffolded group inquiry-based learning processes. Computers & Education, 143, 103674. https://doi.org/10.1016/j.compedu.2019.103674

[20] Kasepalu, R., Prieto, L. P., Ley, T., & Chejara, P. (2022). Teacher Artificial Intelligence-Supported Pedagogical Actions in Collaborative Learning Coregulation: A Wizard-of-Oz Study. Frontiers in Education, 7. https://doi.org/10.3389/feduc.2022.736194

[21] Rolim, V., Ferreira, R., Lins, R. D., & Găsević, D. (2019). A network-based analytic approach to uncovering the relationship between social and cognitive presences in communities of inquiry. The Internet and Higher Education, 42, 53–65. https://doi.org/10.1016/j.iheduc.2019.05.001

[22] Praharaj, S.; Scheffel, M.;Schmitz, M.; Specht, M.; Drachsler, H. Towards Automatic Collaboration Analytics for Group Speech Data Using Learning Analytics. Sensors 2021, 21, 3156. https://doi.org/10.3390/s21093156

[23] David Williamson Shaffer and A. R. Ruis, Epistemic Network Analysis: A Worked Example of Theory-Based Learning Analytics, Wisconsin Center for Education Research, University of Wisconsin–Madison, USA DOI: 10.18608/hla17.015

[24] Gasevic, Dragan & Joksimovic, Srecko & Eagan, Brendan & Shaffer, David. (2018). SENS: Network analytics to combine social and cognitive perspectives of collaborative learning. Computers in Human Behavior. 92. 10.1016/j.chb.2018.07.003.

[25] Roschelle, J., Teasley, S.D. (1995). The Construction of Shared Knowledge in Collaborative Problem Solving. In: O'Malley, C. (eds) Computer Supported Collaborative Learning. NATO ASI Series, vol 128. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-85098-1_5

[26] Peter Guiney, Tertiary Sector Performance Analysis, Ministry of Education, Technology-supported physical learning spaces: An annotated bibliography,2016

[27] J. D. Walker; D. Christopher Brooks; Paul Baepler (2022). Pedagogy and Space: Empirical Research on New Learning Environments.. EDUCAUSE Quarterly 34. https://eric.ed.gov/

[28] S. A. Viswanathan and K. VanLehn, "Using the Tablet Gestures and Speech of Pairs of Students to Classify Their Collaboration," in IEEE Transactions on Learning Technologies, vol. 11, no. 2, pp. 230-242, 1 April-June 2018, doi: 10.1109/TLT.2017.2704099.

[29] Viswanathan, S. A., & VanLehn, K. (2019). Collaboration Detection that Preserves Privacy of Students' Speech. Artificial Intelligence in Education, 507–517. https://doi.org/10.1007/978-3-030-23204-7_42

[30] Rao, P. (2008). Audio Signal Processing. In: Prasad, B., Prasanna, S.R.M. (eds) Speech, Audio, Image and Biomedical Signal Processing using Neural Networks. Studies in Computational Intelligence, vol 83. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-75398-8_8

5. APPENDIX

In this section we have provided additional information which is important to understand the background and scope of the proposed work.

5.1 Definition of collaborative learning

Collaborative learning is a broad term applied to diverse learning situations. Inclusive definition of collaborative learning is as follows: "*It is a situation in which two or more people learn or attempt to learn something together*" [8]. This definition can be interpreted in many different ways. We created visual representation of the elements of the definition of collaborative learning as shown in figure 2.

Figure 2. Definition of "Collaborative Learning"- visual repres-enta-tion.

Another popular definition of collaboration is: "*... a coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem*" [25]. The PISA 2015 collaborative framework defines collaboration as "*a process whereby two or more agents attempt to solve a problem by sharing the understanding and effort required to come to a solution, and pooling their knowledge, skills, and efforts to reach that solution*" [2].

5.2 Collaborative learning spaces

Technology-enhanced collaborative learning classroom used for conducting the study is shown in figure 3.

Figure 3. The technology-enhanced collaborative learning classroom.

The major benefits of creating technology-supported physical learning spaces are more frequent and higher quality teacher-student and student-student interactions, increased student usage of, and satisfaction with, the learning space, and authentic learning experiences [26]. It is also observed that, the type of space in which a class is taught influences instructor and student behavior in ways that likely moderate the effects of space on learning [27].