Using Sentence Embeddings to Automatically Extract Cohesion and Alignment Metrics in Problem-Solving Tasks

Alejandro Andrade, ACTNext, ACT, Inc., laandrad@indiana.edu
Chris Georgen, Indiana University, cgeorgen@iu.edu
Michael Stucker, Indiana University, mstucker@indiana.edu

Abstract: We introduce an automated approach that builds on sentence embeddings, a novel natural language processing technique that extracts meaning from sentences, to create two quantitative measures that serve as proxies of collaborative learning. Cohesion is extracted as adjacent utterance similarity and represents the amount of overlap between contiguous conversational turns. Alignment is extracted as a similarity between a focus utterance and a reference text and represents the degree to which a conversation utterance aligns with the task. These two dimensions divide the quality of conversation in four quadrants.

Introduction

Pre- and post-collaboration scores measure the effect of collaboration on students' learning in a distal manner. For in-vivo measures of collaboration, however, researchers investigate the interaction among participants while they solve a problem as a group. The challenge is that metrics that assess a group's collaboration quality are hard to come by manually. In particular, producing hand codes of a group's collaboration requires sizable human resources. To tackle this issue, various studies have investigated the use of automated approaches to help researchers and instructors reduce the costs of hand coding collaboration transcripts (Mu, Stegmann, Mayfield, Rosé, & Fischer, 2012). Natural language processing (NLP) techniques offer solutions to automatically extract viable measures of collaboration. Some automatically extracted features can predict the quality of a collaborative activity by analyzing linguistic properties of the transcript (Andrade, Georgen, & Stucker, 2017). Various studies have explored computational techniques to extract the cohesion of a conversation (Luna Bazaldua et al., 2015). Cohesion refers to the similarity of adjacent utterances, and it measures the semantic alignment between interlocutors (Graesser, McNamara, Louwerse, & Cai, 2004). Cohesion is indicative of the development and maintenance of shared understanding, which has been shown to be a good predictor of learning in groups (Roschelle, 1992). Other studies have used similar techniques to extract a group's alignment with the task (Andrade et al., 2017). Task alignment is related to how similar a participants' discourse is to the disciplinary ways of talk—i.e., how discipline experts talk. Previous research has shown that task alignment is correlated to course grades (Barron, 2003). Our goal is to investigate the value of a new computational approach that potentially improves the measures of cohesion and task alignment as proxies of a group's collaboration quality. We ask: how is collaborative learning related to patterns of conversational cohesion and task alignment?

Measuring cohesion and task alignment

The development of word embeddings has facilitated the geometrical representation of words. Mathematically, a word embedding is a function that maps a word to a vector of numbers within a multidimensional space (Schnabel, Labutov, Mimno, & Joachims, 2015). For instance, word embeddings capture semantic similarities that can reproduce linguistic analogies such as "Insect is to Ant as Fruit is to Apple" in a geometrical space. This geometrical space allows for algebraic expressions such as: Insect – Ant + Apple = Fruit. Recently, Google released a Universal Sentence Encoder for sentence embedding (Cer et al., 2018). Instead of mapping single words, sentence embedding provides a mapping between a span of text and a vector of numbers. Cosine similarity examines the semantic similarity between two sentence embeddings. Computing a group's conversation cohesion is an iterative process in which a similarity value between adjacent utterances is calculated. Computing task alignment is an iterative process in which each utterance is compared to a reference text. In this case, the reference text is an expert's way of solving the problem. This comparison uses the cosine similarity between the focal text (a student's utterance) and the reference text (expert written solution).

Collaboration quadrants and predicting learning gains

Since cohesion and alignment values are captured at each communicative exchange, we hypothesize that there is a time dependency between these two dimensions of collaboration. A conceptual space with four quadrants represents the intersection between levels of high and low values in each dimension. For instance, there are moments in the conversation students have high levels of cohesion (their talk builds upon each other) as well as high levels of task alignment (their talk mostly refers to the problem at hand). On the other hand, there are moments when their conversation has high cohesion but low alignment (great off-task talk); low cohesion but high alignment (e.g., their on-task but do not seem to take up each other's ideas); and low cohesion and low alignment.

Results

Table 1 shows brief excerpts illustrating what these collaboration quadrants look like in practice. These excerpts were taken from a transcript, during those conversational exchanges our algorithm pointed out towards the corresponding quadrant. A quick appraisal of quadrants seems to validate the qualitative differences at these conversational moments.

Table 1: Collaboration Quadrants

	High Cohesion	Low Cohesion
High Alignment	S17: It is impossible for DC to flow across a capacitor. S06: And AC is like the sin wave right. Yeah AC is a sin wave. Wait let me zoom in even more on this. Well I mean there is some oscillation, so it is in between 3.5 and 4. So maybe it is not actually I mean it looks like AC.	S17: Oh, yep ok there we go. s06: Or test point 1. S17: Oh, you want to try it at 1. s06: And then connect this to ground and this timing control thing is so annoying.
Low Alignment	S1: So how can we prove that it is going to light up twice. S2: Isn't that just because in that one second period it is going to light up twice. S1: Yeah but how do we prove this on circuit wizard. S2: I'm not too sure about that one. S1: Here let's try this. Shucks. Let's throw in a circuit board or something.	S1: Trying to remember where that was. We did that in that last. S2: It is in. You find it. I can't see it. Go to the bottom. Virtual instruments oscilloscope. S3: Channel one is for the positive. Oh, come on.

Discussion and conclusion

The use of sentence embeddings allows for the geometrical representation of the meaning of whole sentences, which can be transferred to the encoding of utterances from collaborative transcripts. Our results show that these two measures of collaboration, automatically extracted from the transcripts, are highly correlated to learning gains. The development of cohesion and alignment values during a conversation create a quadrant space that informs the quality of the talk. Some of these quadrants are productive and some are not. For instance, successful groups spend more time in a kind of conversation that has high cohesion and high collaboration. Not so successful groups tend to spend more time in quadrants where there is low cohesion and low alignment.

References

Andrade, A., Georgen, C., & Stucker, M. (2017). Exploring a Text-Mining Approach as Rapid Prototyping Tool for Formative Assessments in Inquiry-Based Online Learning. In B. Smith, M. Borge, E. Mercier, & K. Y. Lim (Eds.), *Making a Difference: Prioritizing Equity and Access in CSCL* (Vol. 1, pp. 151-158). Philadelphia, PA: International Society of the Learning Sciences.

Barron, B. (2003). When smart groups fail. *The journal of the learning sciences, 12*(3), 307-359.

Cer, D., Yang, Y., Kong, S.-y., Hua, N., Limtiaco, N., John, R. S., . . . Tar, C. (2018). Universal sentence encoder. *arXiv preprint arXiv:1803.11175*.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior research methods, instruments, & computers, 36*(2), 193-202.

Luna Bazaldua, D. A., Hao, J., Khan, S., Liu, L., von Davier, A. A., & Wang, Z. (2015). *On convergence of cognitive and non-cognitive behavior in collaborative activity.* Paper presented at the Proceedings of the 8th International Conference on Educational Data Mining Conference.

Mu, J., Stegmann, K., Mayfield, E., Rosé, C., & Fischer, F. (2012). The ACODEA framework: Developing segmentation and classification schemes for fully automatic analysis of online discussions. *International journal of computer-supported collaborative learning, 7*(2), 285-305.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The journal of the learning sciences, 2*(3), 235-276.

Schnabel, T., Labutov, I., Mimno, D., & Joachims, T. (2015). *Evaluation methods for unsupervised word embeddings.* Paper presented at the Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing.

