Challenges and Opportunities in Annotating a Multimodal Collaborative Problem-Solving Task

Mariah Bradford[1], Paige Hansen[1], Kenneth Lai[2], Richard Brutti[2], Rachel Dickler[3], Leanne M. Hirshfield[3], James Pustejovsky[2], Nathaniel Blanchard[1], and Nikhil Krishnaswamy[1]

[1] Colorado State University, Fort Collins, CO 80523, USA
[2] Brandeis University, Waltham, MA 02453, USA
[3] University of Colorado, Boulder, CO 80309, USA
mbrad@colostate.edu

Abstract. We present an annotation scheme for an upcoming data collection of small group work with the long-term goal of developing an AI partner for collaborative learning. The annotation scheme couples automatically extracted features and semantic information with collaborative problem solving (CPS) frameworks. We showcase our annotation framework with a novel collaborative problem solving task, The Weights Task, which elicits semantically rich information across all the components of the annotation scheme. The resulting annotated audiovisual recordings of small group work in a learning context will be leveraged to create robust machine learning models for AIEd technology.

Keywords: Collaborative problem solving · Group work · Multimodality · Annotation.

1 Introduction

Effective collaboration is a common and necessary skill seen in many areas of life, including education, community events, and the workplace. Collaborative problem solving (CPS) has been shown to be an effective pedagogical technique to improve learning outcomes in the classroom [5, 6, 9]. However, group work is not always advantageous to learning; groups that lack key features, such as clear goals and accountability, often miss out on the educational benefits of collaborative learning [10]. Teachers mitigate these losses through proper facilitation of groups, such as verifying on-task engagement and answering clarifying questions. However, teachers cannot constantly monitor every group closely, making CPS difficult to scale in the classroom. In these situations, an artificially intelligent (AI) partner can be an extremely useful tool for educators. The agent would ideally track group behavior and emulate the helpful interventions teachers provide while being readily available to each group. However, in order to train any models for use by such an agent, there is a marked need for annotated datasets which effectively demonstrate group work with quantifiable outcomes. These datasets must contain representative collaboration and multiple levels of annotation over features pertinent to CPS, including feature recognition, semantic events, and indicators of CPS. The relevant indicators can be defined through existing frameworks for CPS [1, 7, 12].

In this paper we discuss some of the challenges inherent in gathering and annotating such a dataset and opportunities it provides. We first outline our annotation scheme, and then provide our use case in upcoming data collection. We discuss how our methods connect existing AI technology with CPS frameworks and identify items we see as limitations.

2 Data Annotation

In order to train an AI partner for collaborative problem solving, data on which it is trained needs to be annotated in a way that is compatible with traditional collaborative problem solving frameworks. Here, we showcase how we couple a scheme for AI annotation with a collaborative problem solving framework.

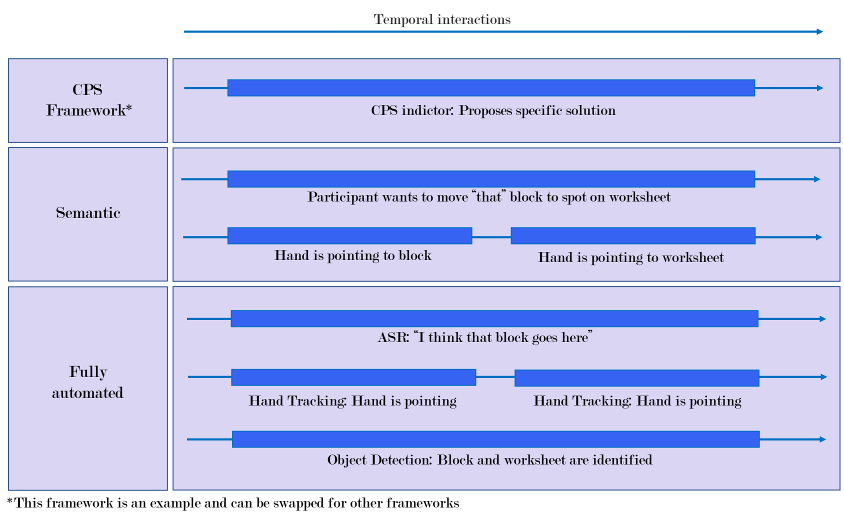


Fig. 1: Weight Task annotation scheme using CPS framework [12]

The complete annotation scheme is based on three tracks, separable by the complexity of each feature. Our first track — the atomic track — focuses on automated object detection, body pose detection, and automatic speech recognition using out-of-the-box packages discussed in [4]. The semantic annotation track combines those low-level features into meaningful patterns. One example of this is gesture semantics, which in our example is analyzed using Gesture AMR (GAMR) [3]. GAMR extends Abstract Meaning Representation (AMR) [2] — a graph-based meaning representation that encodes predicate-argument structure — with additional elements specific to gesture. One such element is *gesture act relations*, which identify how a gesture's meaning relates to its form. Gesture acts allow us to interpret the motions and objects identified in the atomic track. For example, a gesturer using their fingers to trace the outline of a square could be referring to a block; this is an *iconic* gesture, which describes an action or object by depicting its shape. A gesturer could also refer to the block by pointing to it; this is a *deictic* gesture, which identifies locations or objects in physical space. Applying this track of annotation to our dataset is necessary to connect our atomic features with abstract concepts by first assigning an interpretation to the features. Our highest track of annotation uses semantic features to detect events related to CPS based on existing frameworks, such as [12]. An example of this scheme being applied can be seen in Figure 1. Each track builds upon the lower track, adding progressively more information at greater levels of abstraction. This scheme highlights a process from automatic feature extraction using existing AI tools, to semantic interpretations of these features, to annotations drawn from frameworks developed for CPS, especially in the context of educational group work.

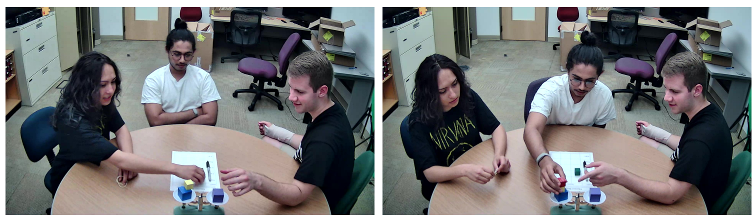
In the next section, we showcase our annotation scheme with a novel collaborative problem solving use case: The Weights Task.

3 The Weights Task

Here, we showcase how the annotations can be applied to a novel collaborative problem solving activity: The Weights Task. The task, depicted in Figure 2, is a novel collaborative problem-solving activity for a group of three. In the task, participants are asked to identify the weights of several blocks on the table in front of them. They are given a balance scale and the weight of one block to solve the puzzle. After completing this portion of the task, the participants are told the correct answers. Then, the researcher will remove the balance scale and give the group another "mystery" block, which they can identify through a pattern in the weights. They have two chances to identify the weight of this block. Participants are given a hint if their first attempt is incorrect. Next, participants are told there is another block and asked what the weight of this hypothetical block would be. Again, they have two chances to correctly identify the weight of the block and will receive a hint if their first answer is incorrect. To successfully complete the task, participants must identify the pattern of the blocks: their weights follow the Fibonacci Sequence (the weight of the next block is the sum of the previous two blocks).

Previous work has explored methods for eliciting collaboration between groups of people and annotating the results in the context of training an AI partner. [11] and [13] examine virtual collaboration in triads using the CPS framework developed by [12]. This framework describes three facets of CPS: construction of shared knowledge, negotiation and coordination, and maintaining team function. Figure 1 shows how a collaborative problem solving indicator from the framework of [12] couples with our annotation. This indicator corresponds with facets and sub-facets from the framework, but we have left these off the figure since they are easily inferred. Note that the annotation can be coupled with alternative collaborative problem solving frameworks, as long as those frameworks are focused around similarly time-dependent indicators.

The Weights Task is designed for in-person data collection and uses physical objects — this ensures the data is rich in both verbal and nonverbal communication, making both natural language processing and computer vision crucial elements of the pipeline. The annotation scheme encourages initial feature extraction, such as object detection and automatic speech recognition, and the combination of these multimodal elements to provide complete interpretations of the scene as action unfolds. Then, this scene can be analyzed for indicators of collaborative problem solving defined by existing frameworks.



(a) (b)

Fig. 2: Example of The Weights Task

4 Discussion

Ultimately, we seek to create an annotated dataset which effectively represents behaviors seen in real-world small group environments. This entails prompting the behaviors seen in classrooms, including verbal and nonverbal communication between group members, interaction with the physical objects, and the emergence of group dynamics. The proposed task provides an opportunity for a group to participate in complex problem-solving that encourages communication between members. We aim to use these tasks to build a dataset rich in multiple levels of annotation focused on indicators of successful and unsuccessful group work using existing CPS frameworks. As mentioned above, we will break down annotation into a bottom-up process which applies lower-level features and builds up to more complex concepts. To capture a useful dataset, we seek feedback from the educational community about the proposed approach, its strengths and weaknesses, and features that are important in identifying the condition of working groups.

Theory in the learning sciences [8] suggests that even providing the annotations for a single indicator, such as identifying off-task discussion, is very challenging because it is contextually and temporally dependent. These challenges are alleviated in the long term through the automation of as much of the lowest-level annotation as possible [4]. It is our hope that through connecting with the educational community we can mitigate these specific issues and gain feedback on possibly overlooked items in the data collection, as well as discuss ways to more effectively apply our annotation scheme to existing frameworks.

References

1. Andrews-Todd, J., Forsyth, C.M.: Exploring social and cognitive dimensions of collaborative problem solving in an open online simulation-based task. Computers in Human Behavior **104**, 105759 (Mar 2020). https://doi.org/10.1016/j.chb.2018.10.025, https://www.sciencedirect.com/science/article/pii/S0747563218305156
2. Banarescu, L., Bonial, C., Cai, S., Georgescu, M., Griffitt, K., Hermjakob, U., Knight, K., Koehn, P., Palmer, M., Schneider, N.: Abstract Meaning Representation for Sembanking. In: Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse. pp. 178–186. Association for Computational Linguistics, Sofia, Bulgaria (Aug 2013), https://aclanthology.org/W13-2322
3. Brutti, R., Donatelli, L., Lai, K., Pustejovsky, J.: Abstract Meaning Representation for Gesture. In: Proceedings of the Language Resources and Evaluation Conference. pp. 1576–1583. European Language Resources Association, Marseille, France (Jun 2022), https://aclanthology.org/2022.lrec-1.169
4. Castillon, I., VanderHoeven, H., Bradford, M., Venkatesha, V., Krishnaswamy, N., Blanchard, N.: Multimodal Features for Group Dynamic-Aware Agents. In: Interdisciplinary Approaches to Getting AI Experts and Education Stakeholders Talking Workshop at AIEd. International AIEd Society (2022)
5. Gillies, R.M.: The Effects of Cooperative Learning on Junior High School Students' Behaviours, Discourse and Learning During a Science-Based Learning Activity. School Psychology International **29**(3), 328–347 (Jul 2008). https://doi.org/10.1177/0143034308093673, https://doi.org/10.1177/0143034308093673, publisher: SAGE Publications Ltd
6. Graesser, A.C., Fiore, S.M., Greiff, S., Andrews-Todd, J., Foltz, P.W., Hesse, F.W.: Advancing the Science of Collaborative Problem Solving. Psychological Science in the Public Interest **19**(2), 59–92 (Nov 2018). https://doi.org/10.1177/1529100618808244, https://doi.org/10.1177/1529100618808244, publisher: SAGE Publications Inc
7. Hesse, F., Care, E., Buder, J., Sassenberg, K., Griffin, P.: A Framework for Teachable Collaborative Problem Solving Skills. In: Griffin, P., Care, E. (eds.) Assessment and Teaching of 21st Century Skills: Methods and Approach, pp. 37–56. Educational Assessment in an Information Age, Springer Netherlands, Dordrecht (2015). https://doi.org/10.1007/978-94-017-9395-7_2, https://doi.org/10.1007/978-94-017-9395-7_2
8. Langer-Osuna, J.M., Gargroetzi, E., Munson, J., Chavez, R.: Exploring the role of off-task activity on students' collaborative dynamics. Journal of Educational Psychology **112**(3), 514 (2020), publisher: American Psychological Association
9. Roschelle, J., Teasley, S.D.: The Construction of Shared Knowledge in Collaborative Problem Solving. In: O'Malley, C. (ed.) Computer Supported Collaborative Learning. pp. 69–97. NATO ASI Series, Springer, Berlin, Heidelberg (1995). https://doi.org/10.1007/978-3-642-85098-1_5
10. Slavin, R.: Research on Cooperative Learning: Consensus and Controversy. Educational Leadership **47** (Jan 1990)
11. Stewart, A.E.B., Keirn, Z., D'Mello, S.K.: Multimodal modeling of collaborative problem-solving facets in triads. User Modeling and User-Adapted Interaction **31**(4), 713–751 (Sep 2021). https://doi.org/10.1007/s11257-021-09290-y, https://doi.org/10.1007/s11257-021-09290-y

12. Sun, C., Shute, V.J., Stewart, A., Yonehiro, J., Duran, N., D'Mello, S.: Towards a generalized competency model of collaborative problem solving. Computers & Education **143**, 103672 (2020), publisher: Elsevier
13. Sun, C., Shute, V.J., Stewart, A.E.B., Beck-White, Q., Reinhardt, C.R., Zhou, G., Duran, N., D'Mello, S.K.: The relationship between collaborative problem solving behaviors and solution outcomes in a game-based learning environment. Computers in Human Behavior **128**, 107120 (Mar 2022). https://doi.org/10.1016/j.chb.2021.107120, https://www.sciencedirect.com/science/article/pii/S074756322100443X