Acknowledgement

This version of the contribution has been accepted for publication, after peer review (when applicable) but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: https://doi.org/10.1007/978-3-031-64299-9_36.

Citation

Please cite this manuscript as:

de Araujo, A., Martens, J., Papadopoulos, P. M. (2024). Enhancing Student Dialogue Productivity with Learning Analytics and Fuzzy Rules. In *Proceedings of the 25th International Conference on Artificial Intelligence in Education (AIED 2024)*. Recife, Brazil. https://doi.org/10.1007/978-3-031-64299-9_36

Enhancing student dialogue productivity with learning analytics and fuzzy rules

Adelson de Araujo, Jara Martens, and Pantelis M. Papadopoulos

Department of Learning, Data, and Technology, University of Twente, Enschede, the Netherlands
a.dearaujo@utwente.com, j.martens@student.utwente.nl, p.m.papadopoulos@utwente.nl

Abstract. This study explores the use of the Collaborative Learning Agent for Interactive Reasoning (Clair) in a digital collaborative learning activity where interaction takes place via chat. Clair is designed to adaptively facilitate productive student dialogue using "talk moves" based on the Academically Productive Talk (APT) framework, a popular approach in related conversational agent studies. In this paper, we detail how Clair, powered by learning analytics, machine learning, and a fuzzy rule-based system, can adaptively trigger talk moves in student dialogue. In an experimental study conducted with n = 9 university student dyads, we assess the impact of Clair's presence on student dialogue productivity. We analyzed the within-subjects differences (with/without Clair) in four key goals of student dialogue productivity: the frequency of (a) students sharing thoughts, (b) orienting and listening, (c) deepening reasoning, and (d) engaging with others' reasoning. Our findings indicate a notable improvement in deepening reasoning (p = .047), highlighting Clair's capability to prompt students to engage in more critical thinking and elaborate on their ideas. Yet, the impact on other goals was less pronounced, suggesting the complexity of facilitating all goals of productivity. This paper demonstrates the potential of integrating learning analytics and fuzzy rules into triggering approaches for collaborative conversational agents, offering a novel approach to adaptively trigger talk moves in student dialogue. The results also underline the need for further refinement in the design and application of such systems to comprehensively support productive student dialogues in collaboration settings.

Keywords: conversational agents, collaborative learning, student dialogue, academically productive talk, learning analytics.

1 Introduction

In the 21st century, *Critical Thinking*, *Creativity*, *Collaboration*, and *Communication* skills—collectively known as the 4Cs—have become essential for navigating the fast-changing global landscape. However, productive dialogue, where students build on each other's ideas constructively (Chi & Wylie, 2014), often requires explicit guidance; otherwise, participation can be uneven and superficial (Gillies, 2019). For example, an eloquent student contributes more while the partner just agrees, an idea presented is too vague, or students may seldom build on each other's contributions.

Teachers play a key role in fostering the 4Cs and encouraging productive student dialogue. The Academically Productive Talk (APT), or 'Accountable Talk' (Michaels et al., 2008), is a method that enhances student dialogue through structured 'talk moves', reflective prompts that teachers can make to guide students, such as "*Could someone summarize what we have discussed so far?*" (i.e., Recapping). In student dialogue, these content-independent strategies encourage peer interaction and reasoning, fostering a deeper understanding of the subject being discussed. Yet, effective APT application requires teachers to skillfully listen to conversations and pose timely questions. Thus it can be challenging, time-consuming, and not feasible for the teachers to handle multiple groups at once.

To address this, researchers are exploring APT strategies delivered by collaborative conversational agents (CCAs) as scalable solutions to facilitate student dialogue. For instance, Tegos et al. (2016) found that the presence of a CCA in student dialogue was linked to a higher frequency of explicit reasoning behaviors and more balanced participation. Nguyen (2023) found a positive effect, compared to a control condition, on students' transactive exchanges, i.e., explanations that are directed towards building on a partner's contribution. Adamson et al. (2014) reported on several studies, demonstrating that CCAs' effects can largely depend on the audience and learning material. These studies employed learning analytics, e.g., through natural language processing, and a rule-based system to trigger interventions in student groups. Yet, studies demonstrating whether and how CCAs promote productive dialogue are limited.

Building on the foundations laid by previous studies, we developed a CCA that combines various learning analytics with a fuzzy rule-based approach to improve the flexibility required to trigger a variety of talk moves in student dialogue. This paper uniquely contributes to existing knowledge by (i) describing in detail our novel CCA triggering technique and (ii) reporting on a new case study evaluating our CCA across various dimensions of productive dialogue among university students.

1.1 Research question

By timely prompting students with a variety of talk moves, we anticipate that the presence of CCAs can encourage productive student dialogue. Michaels & O'Connor (2015) elaborated on essential goals of teacher guidance that align with observable student behaviors, which are particularly relevant for assessing the impact of CCAs. These goals are outlined as the Four Goals for Productive Discussions (FGPD):

1. Helping students share their own thoughts.
2. Helping students orient to and listen carefully to one another.
3. Helping students deepen their reasoning.
4. Helping students engage with each other's reasoning.

Accordingly, this paper presents a case study formulated around the following research question:

RQ: To what extent does the collaborative conversational agent make written student dialogue more productive compared to when the agent is not present in terms of (a) sharing their thoughts, (b) orienting and listening to one another, (c) deepening their reasoning, and (d) engaging with each other's ideas?

2 Collaborative learning agent for interactive reasoning (Clair)

In designing collaborative conversational agents for online student dialogues, finding the right type and time for a talk move can be challenging (Adamson et al., 2014; Nguyen, 2023; Tegos et al., 2016). Clair uses eight talk moves (see labels and examples in Table 1 in Sect. 2.2) crafted based on APT guidelines (Michaels et al., 2015). Clair's talk moves can either target the last student who spoke, another discussant, or both students. Clair uses three alternative phrasings of each talk move to avoid repetition. By employing various talk moves, we aim to comprehensively support students in the FGPDs based on what happens in the dialogue.

Figure 1 illustrates the process for using learning analytics and fuzzy rules to trigger talk moves in student dialogue. The process starts with Clair's configuration, which requires the fuzzy rules chosen and configuration details of the task at hand including the topic keywords.

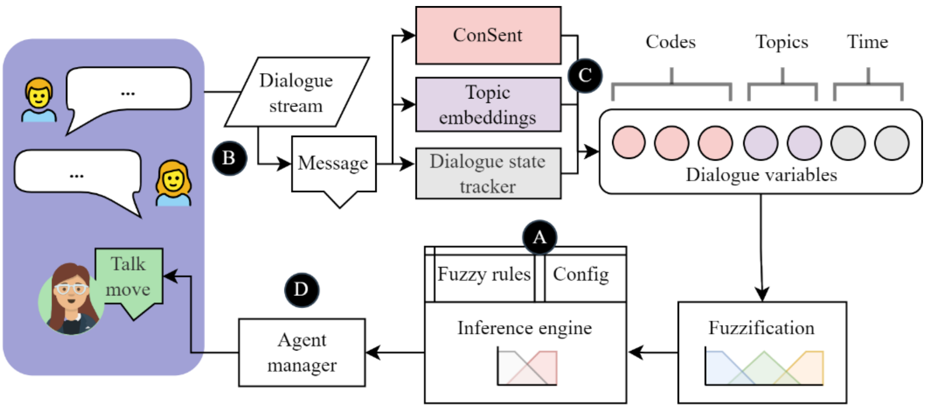

Fig. 1. Flowchart of Clair's internal components. Clair's triggering mechanism must be prepared with configuration adjusted to the task at hand and the fuzzy rules (A). Subsequently, Clair is ready to receive messages (B), calculate its dialogue variables (C), and send talk moves to the dialogue (D).

2.1 Dialogue variables

To make decisions on the timing of interventions, Clair first employs learning analytics instruments hereby called 'dialogue variables'. In the current version of Clair, there are a total of twelve dialogue variables, calculated for each message. Eight of them are probabilities of collaborative behavior categories, outputs of a ConSent model, powered by the pre-trained multilingual Universal Sentence Encoder (mUSE, Yang et al., 2019) which we evaluated in previous work and found satisfactory to moderate levels of reliability (de Araujo et al., 2023b). Three of these are related to the message' '*focus*' (at κ = 0.60), which are Domain (L1_DOM), Coordination (L1_COO), and Off-task (L1_OFF); whereas the five others are related to its '*intent*' (κ = 0.61), which are Informative (L2C_IN), Argumentative (L2C_AR), Asking for information (L2C_AI), Active motivating (L2C_AM), and None of the specified (L2C_NOS).

There are four other dialogue variables. The first two are about how students addressed the topic: Topic similarity (TSIM) measures semantic similarity to a list of topic keywords (transformed into mUSE embeddings; Yang et al., 2019) and Topic accumulation (TACC) measures the ratio of the speaker's accumulated TSIM value compared to dialogue partners. The other two are metrics from the dialogue state: Time spent (TIME) measures the time since the dialogue started and Messaging speed (PACE) measures the proportion of messages per unit of time. For more information and examples on these dialogue variables, we refer to our previous work (de Araujo et al., 2024).

2.2 Triggering mechanism

The triggering mechanism, based on a fuzzy expert system, decides whether to select a talk move in response to a chat message and its dialogue variables. This approach was chosen for its balance between clarity and managing uncertain data, as outlined by Zadeh (1983), and implemented using the Scikit-Fuzzy library (Warner et al., 2019).

The first step, 'fuzzification', changes clear-cut dialogue variable values into "fuzzy sets" such as *high*, *medium*, and *low*. These sets represent intensity levels that are not fixed but cover a range, allowing for more flexible decision-making with the dialogue variables. To tailor these fuzzy sets for each dialogue variable, we analyzed chat message data using K-Means clustering (with K = 3). This helped us define what *high*, *medium*, and *low* mean for each variable (e.g., see Figure 2a). For each talk move output, we define two levels: *active* and *not-active* (e.g., see Figure 2b) which guide whether a response should be chosen or not.

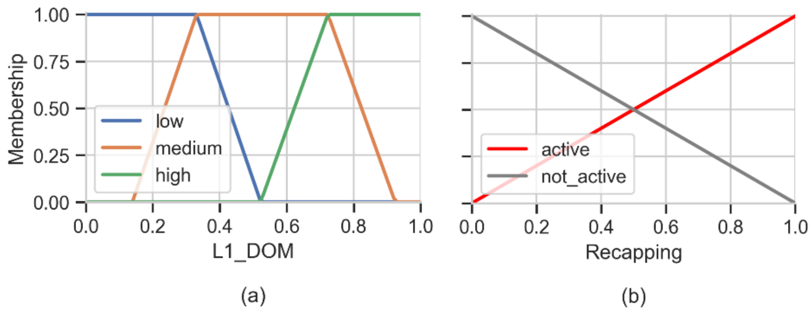

Fig. 2. Membership functions for the dialogue variable L1_DOM (a) and the talk move Recapping (b).

After the fuzzification step, the 'inference engine' step is responsible for determining outputs for each talk move by relying on fuzzy rules applied to the inputs that indicate the likelihood that each talk move should be triggered. The fuzzy rules are in the form of “IF-THEN” statements that relate to states of inputs with a desired level of output for each talk move. More specifically, fuzzy rules are defined in the form of “*IF* $(x_1$ *is* $A_1)$ *and* ...$(x_n$ *is* $A_n)$ *THEN* $(y_1$ *is* $B_1)$ *and*...$(y_m$ *is* $B_m)$”, where A_i and B_j are fuzzy sets to describe the x_i dialogue variable and y_j talk moves, respectively. To ensure that the rule base is simpler to interpret and adjust, we currently employ one rule to determine when each talk move’s *active* level should be *high*, i.e., moments when Clair could intervene, and another general rule to determine when all talk moves’ *not-active* level should be *high*, i.e., moments when Clair should not intervene at all.

The triggering mechanism’s final step is translating the fuzzy output into a binary decision. This step, usually known as ‘defuzzification’, is implemented by Clair’s ‘agent manager’. The agent manager initially chooses the primary talk move candidates based on the active values (higher than 0.75). Candidates used in the last three interventions are excluded to avoid repetition, and talk move frequency is monitored, favoring less used ones when active values and frequencies match. In a tie, the selection is random and the final utterance variation is randomly chosen and sent to the dialogue.

The fuzzy rules for all talk moves utilized in the current version of Clair are detailed in Table 1. The formulation of rules and refinement were conducted by expert evaluation and interviewing secondary school teachers (de Araujo et al., 2023a).

Table 1. Clair’s talk moves with examples and associated fuzzy inference rules used in the triggering mechanism.

Talk move	Example	Fuzzy inference rule
Recapping	*“Can someone give a brief summary of what we've covered so far?”*	L1_DOM is *medium/low* and L2C_AR is not *high* and TSIM is *low* and TIME is *high*
Add-on	*“<discussant>, could you add some new perspective to what <speaker> just said?”*	(L1_DOM is *high* or L1_COO is *high*) and L2C_AR is *high* and TSIM is not *low* and TACC is *high*
Rephrasing	*“<discussant>, can you rephrase what <speaker> said so that everyone is on the same page?”*	L1_DOM is *high* and L2C_AR is *high* and TSIM is *high* and TACC is *high*
Agree/Disagree	*“<discussant>, can you explain to <speaker> if there is something you disagree with?”*	(L1_DOM is *high* or L1_COO is *high*) and L2C_AR is *high*/*medium* and TSIM is *medium* and TACC is not *low*
Linking contributions	*“<discussant>, can you link your ideas to what <speaker> said?”*	L1_DOM is *high* and (L2C_IN is *high* or L2C_AR is *medium*) and TSIM is *medium* and TACC is *high*
Build on prior knowledge	*“<speaker>, can you explain to <discussant> how this fits into the bigger picture?”*	L1_DOM is *high* and (L2C_IN is *high* or L2C_AR is *high*) and TSIM is *high* and TACC is *low* and TIME is *high*/*medium*

Example	*"<speaker>, can you give an example or a real-life scenario for <discussant> that can help illustrate the concept better?"*	L1_DOM is *high* and L2C_AR is *high* and TSIM is *medium* and TACC is *medium/low*
Expand reasoning	*"<speaker>, can you explain to <discussant> how you got this idea?"*	L1_DOM is *high* and (L2C_IN is *high* or L2C_AR is *high/medium*) and TSIM is *high/medium* and TACC is *medium/low*

3 Case Study

To evaluate Clair's impact and address RQ1, we conducted a within-subjects repeated-measures study with university students on a previously covered topic. In total, 18 Psychology Bachelor's students volunteered to participate (13 females, 5 males; age range: 19 to 23, M = 21.6, SD = 1.10). Participants received an anonymized username and were randomly assigned into pairs (n = 9).

After login, the participants provided their consent and demographical information in a form. The activity started with a familiarization phase that lasted 8 minutes, engaged the pairs in chat discussion of a topic on climate change, and helped the participants to get a better understanding of the overall task goal for the upcoming phases.

Next, participants discussed the two main topics (i.e., classical and operant conditioning). Each topic was discussed for 15 minutes and finished with the pair's final answer in the chat. In the second topic, all participants interacted with Clair. Finally, the participants filled out a questionnaire indicating their satisfaction with the task and with Clair which was reported on a preliminary report by Martens (2023).

3.1 Data analysis

To analyze our RQ, we employ learning analytics with the dialogue variables from ConSent models, i.e., message's focus and intent, to measure FGPDs. Using these variables, we applied sequential pattern mining to track target behaviors, determining outcomes by the frequency of behavior-indicating patterns in dialogues. We analyzed dialogues in n-message windows, identifying patterns within these message sequences. Conditions for a pattern match within these sequences are assessed in three steps, and the behavior does not need to appear in consecutive messages. We found n = 7 windows more consistently capture FGPD-related patterns, including three condition-meeting messages and four peripheral ones. Smaller windows, e.g., n = 5, seem to not capture behaviors of Goals 3 and 4.

In particular, the pattern conditions used the following proxies to operationalize dependent variables of FGPD: Goal 1's pattern involves identifying sequences of task-related informative and argumentative statements, indicating thought sharing; Goal 2 focuses on recognizing responses to peers' task-related questions; Goal 3 on expanding one's own task contributions; and Goal 4 on collaborative argument discussion. Thresholds were adopted using each dialogue variable's limit of *high*, *medium*, and *low* from the triggering mechanism. Ultimately, each goal's outcome variable is measured as the count of unique pattern matches. This approach is further described in our previous work (de Araujo et al., 2024).

3.2 Results

Wilcoxon's signed rank test indicated varied within-subjects impacts of Clair on the FGPDs. We observed a statistically significant improvement in Goal 3 (deepening reasoning) from Phase 1 to Phase 2 (p = .046, r_{rb} = 0.905; for α = .05). This suggests that Clair effectively facilitated students in elaborating and expanding their reasoning over the dialogue. However, the results for Goals 1 (sharing thoughts, p = .159), 2 (orienting and listening, p = .891), and 4 (engaging with others' reasoning, p = .074) did not show significant differences between phases.

4 Discussion and conclusion

Results analysis suggests that tools like Clair can partly aid productive student dialogue. In particular, the presence of Clair can help an audience of university students to deepen their reasoning more often than when Clair is absent. The significant impact can be attributed to Clair's ability to prompt students to build on each other's contributions and explain their reasoning. This is aligned with findings from Tegos et al. (2016), who found an increase in students' explicit reasoning behavior in the presence of a similar agent. Also, our result could be explained by the high responsiveness level of university-level participants (85% talk moves responded, 6% acknowledged, 9% ignored).

Educators may currently consider these tools as a supplement rather than a replacement for comprehensive guidance toward online productive dialogue. The results might not generalize to other audiences, e.g., K-12, as university students are usually more skilled in discussing with each other. In addition, our case study had a small sample size and further research may be required to uncover Clair's impact.

Regarding theory-building, we demonstrated a clear need for further research and development to more effectively support all FGPDs. Assuming human teachers can help student dialogue in all FGPDs, developing more advanced, human-like interaction features in CCAs could potentially facilitate their effectiveness. Furthermore, exploring the customization of talk moves, e.g., to the needs of different student audiences, may increase the chances of impactful results. Future research could explore incorporating large language models, to comprehensively guide productive student dialogue.

References

Adamson, D., Dyke, G., Jang, H., & Rosé, C. P. (2014). Towards an agile approach to adapting dynamic collaboration support to student needs. *International Journal of Artificial*

Intelligence in Education, *24*(1), 92–124. https://doi.org/10.1007/s40593-013-0012-6

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

de Araujo, A., Papadopoulos, P. M., McKenney, S., & de Jong, T. (2023a). Supporting Collaborative Online Science Education with a Transferable and Configurable Conversational Agent. *15th International Conference on Computer-Supported Collaborative Learning (CSCL)*.

de Araujo, A., Papadopoulos, P. M., McKenney, S., & de Jong, T. (2023b). Automated coding of student chats, a trans-topic and language approach. *Computers and Education: Artificial Intelligence*, *4*, 100123. https://doi.org/10.1016/j.caeai.2023.100123

de Araujo, A., Papadopoulos, P. M., McKenney, S., & de Jong, T. (2024). A learning analytics-based collaborative conversational agent to foster productive dialogue in inquiry learning. *Journal of Computer Assisted Learning*, *In press*.

Gillies, R. M. (2019). Promoting academically productive student dialogue during collaborative learning. *International Journal of Educational Research*, *97*, 200–209. https://doi.org/10.1016/J.IJER.2017.07.014

Martens, J. (2023). *Artificial Intelligence in Education: AI Conversational Agent for Online Collaborative Learning*. http://essay.utwente.nl/95259/

Michaels, S., & O’Connor, C. (2015). Conceptualizing talk moves as tools: Professional development approaches for academically productive discussions. In L. B. Resnick, C. Asterhan, & S. N. Clarke (Eds.), *Socializing intelligence through talk and dialogue* (pp. 347–362). American Educational Research Association.

Michaels, S., O’Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in Philosophy and Education*, *27*(4), 283–297. https://doi.org/10.1007/s11217-007-9071-1

Nguyen, H. (2023). Role design considerations of conversational agents to facilitate discussion and systems thinking. *Computers & Education*, *192*(104661). https://doi.org/10.1016/J.COMPEDU.2022.104661

Tegos, S., Demetriadis, S., Papadopoulos, P. M., & Weinberger, A. (2016). Conversational agents for academically productive talk: a comparison of directed and undirected agent interventions. *International Journal of Computer-Supported Collaborative Learning*, *11*(4), 417–440. https://doi.org/10.1007/s11412-016-9246-2

Warner, J., Sexauer, J., scikit-fuzzy, et al. (2019). *JDWarner/scikit-fuzzy: Scikit-Fuzzy version 0.4.2*. https://doi.org/10.5281/ZENODO.3541386

Yang, Y. et al. (2019). Multilingual Universal Sentence Encoder for Semantic Retrieval. *ArXiv:1907.04307*, 87–94.

Zadeh, L. A. (1983). The role of fuzzy logic in the management of uncertainty in expert systems. *Fuzzy Sets and Systems*, *11*(1–3), 199–227. https://doi.org/10.1016/S0165-0114(83)80081-5