Rodrigo, M. M. T. et al. (Eds.) (2021). Proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Reflection Support Environment for Creative Discussion Based on Document Semantics and Multimodal Information

Atsuya SHONO, Yuki HAYASHI* & Kazuhisa SETA
Graduate School of Humanities and Sustainable System Sciences, Osaka Prefecture University, Japan
*hayashi@kis.osakafu-u.ac.jp

Abstract: The purpose of this paper is to realize a mechanism of system processing that stimulates the reflection on a series of research activities with creative discussions (research meetings) as a starting point. In this study, as a clue for systems to capture a part of the semantic contents of the discussion interactions, we utilize *document semantics*, which represent the intentions and contents of the discussion materials, and *multimodal information* exchanged among discussion participants. We propose declarative rules for detecting discussion sections and generating advice that follows the discussion interaction context as a stimulus for reflection. Then, we develop a reflection support environment consisting of a rule creation support system and a reflection support system that embodies the proposed mechanism.

Keywords: Reflection support, multiparty multimodal interaction, document semantics

1. Introduction

An academic research meeting is a place where students belonging to a laboratory can share their research through creative discussions, and is a promising opportunity to cultivate thinking skills (Mori, Hayashi, & Seta, 2019). In general, the proposers (learners) examine the contents of proposals from a multilateral viewpoint (internal conversation), then summarizes the results as discussion materials before the meeting. After the discussion, it is essential that they should reflect on the discussion and examine the inadequate points raised in the meeting. In this reflection, it is desirable not merely to reflect on discussion topics, but also to focus on thinking activities at the time of preparation before the discussion. Such reflection activities contribute to cultivating important perspectives in achieving successive knowledge co-creative discussions, such as a refinement of the internal conversation toward the next research meeting and obtaining perspectives for correctly communicating their intentions to the discussion participants.

This study tackles a research question on how to realize a mechanism of system processing to stimulate the reflection on a series of research activities with such research meetings as a starting point. To approach this, we consider capturing a learners' reflection section by utilizing *document semantics*, which represent the intentions and contents of the discussion materials (the fruits of internal conversation), and *multimodal interaction information* (such as gazing and utterance information) during discussion. In this paper, we propose a novel reflection support environment composed of two systems: a rule creation support system that detects a reflection section aimed at an advisory presentation based on document semantics and interaction information and a reflection support system that applies the created rules and prompts reflection activity.

2. Functional Requirements of Reflection Support Environment

2.1 Requirement Definition

In this section, the functional requirements for the mechanism of the system processing required for our target reflection support environment are listed.

R1. *Mechanism for capturing the context of discussion interactions in a computer-readable format*: Research meetings generally share discussion materials that correspond to the outcome of the learner's preparation activity, and discussion progresses while confirming the same. To provide advice that follows the discussion interaction context as a stimulus for reflection, it is desirable that a system can grasp the semantic content (research content) of the content described in discussion materials as well as what kind of intention (logic composition intention) this is an attempt to explain.
R2. *Mechanism to detect the reflection section by declarative rules with guaranteed reusability*: To adaptively detect the discussion section contributing to the reflection, it is necessary to consider a mechanism with high reusability that does not depend on the session composed of specific discussion participants. Furthermore, by allowing to explicitly associate the defined declarative rules with each layer of the hierarchical interpretation model (Section 2.2), it is desirable that when a new rule is created in any layer, the existence of data detected by such upper or lower layer can be distinguished to create a rule. Here, it is desirable that the processing procedure for detection not be a form embedded in a program but a declarative form in which the intent is easy to understand.
R3. Mechanism for providing advice according to the focused section of the reflection: In the discussion reflection, not only is it possible to give the situation at the time of discussion according to the learner's focused section (reflection section detected by the rule satisfying R1 and R2), it is also necessary to be able to provide an introspective advisory toward the thinking activity at the time of preparation.

We focus on document semantics (Section 3.1) and multimodal information (Section 3.2) for R1, and consider a framework of rule creation support to satisfy R2 (Section 3.3). A rule creation support system equipped with this mechanism is then developed (Section 4.1), and a reflection support system is proposed to satisfy R3 (Section 4.2).

2.2 Hierarchical Interpretation Model Based on Interaction Corpus

Hierarchical interpretive model of the interaction (Sumi, Yano, & Nishida, 2010) is a conceptual approach that raises the interpretation from the data belonging to the low-order hierarchy to the high-order hierarchy during a multiparty multimodal interaction. The model consists of four layers: *raw data layer,* which includes simple data sequences such as eye coordinates and voice waveforms; *interaction primitive layer* that corresponds to interaction elements of individual participants such as those who are speaking or looking at someone; *interaction event layer* that combines interaction elements and interprets them such as a joint attention and a mutual gaze; and *interaction context layer* that builds to a higher-order interpretation of interactions regarding the conversational context.

Based on this model, this study considers capturing the discussion sections in the research meeting by combining and interpreting the document semantics and multimodal information of the discussion participants and provides the learners with these sections as reflection points.

2.3 Measurement Environment for Multimodal Information

This study utilizes a multimodal interaction-aware platform for collaborative learning that has a mechanism to capture several verbal and non-verbal information exchanged in multiparty interactions (Sugimoto, Hayashi, & Seta, 2020). The platform is configured to incorporate several learning support tools. Within the system, each participant's video image, utterance timing, interval section of gaze information (gazing at other participants / gazing at material parts) by setting gaze target regions corresponding to each participant's video image and material contents, and so on can be measured using sensing devices. These data are stored in the database for each discussion session, corresponding to the raw data and interaction primitive layers of the hierarchical interpretation model.

3. Rule Creation Based on Document Semantics and Multimodal Information

3.1 Document Semantics

Document semantics are computer-readable datasets that indicate the semantic content of areas on the discussion material. This allows the system to interpret information with the meaning, such as a learner is gazing at the 'research purpose' area from the eye coordinates on the corresponding discussion material area. In this study, we consider two types of semantic information, i.e., *research content semantics* representing the content of the research activity of the learner and *intention semantics* expressing the intention of the logical composition incorporated into the discussion materials.

Research Content Semantics: We use the research activity ontology proposed by Mori, Hayashi, & Seta (2019). In this ontology, the thinking activity required to conduct the ITS research and the requirements of the activity (sub-activities, inputs, and outputs) are structured in a way to take into account the general and specific nature of the research area. Each sub-activity is necessary for achieving another activity. We assume that there are discussion materials in which concepts of the research activity ontology (research content and its linkage) are corresponded by a learner who intends to sufficiently carry out an internal conversation referring to this ontology before the discussion.

Intention Semantics: We utilize the intent ontology of the logical composition about discussion materials, which is an extended ontology proposed by Matsuoka, Seta, & Hayashi (2019). In this ontology, the concept for clearly expressing the logical role of the research content is defined from the viewpoint of the discussion purpose.

In addition to the research content semantics, discussion materials given these intentions reflect the planning activities of the learners who try to design the discussion. Therefore, giving these semantics by learners themselves to the discussion materials as tasks before discussion not only increases the readability of a computer, it also enhances the readiness of learners, and becomes a significant activity from the viewpoint of enhancing the quality of the discussion itself.

3.2 Multimodal Information

In complex communications consisting of several persons, such as discussions, multimodal information such as utterance timing, back-channeling, and gazing information plays a crucial role in addition to the verbal content to be transmitted (Burgoon et al., 2017; Thiran, Marques, & Bourlard, 2009).

This study focuses on two types of modalities that can be detected using the multimodal interaction-aware platform: "utterance information," which is crucial in advancing the discussion, and "gaze information," which represents a part of the thinking of the activity subjects. These communication signals, exchanged in the discussions, assume that they can be treated as having role-based general-purpose and highly reusable interaction information that does not rely on a particular participant by giving the roles of actors such as "instructors" and "proposers."

3.3 Framework for Rule Creation Support System

Based on the multimodal interaction information toward discussion materials given the research content and intent semantics, we propose three types of declarative rules to realize a mechanism for detecting a reflection section that can be applied to sessions consisting of various discussion participants (R2) and a mechanism for generating advice on thinking activities at the time of preparation for discussion (R3).

Initialization Rule: {type, subject, target, rate, inequality, time} · · · (1)

This rule provides the initial settings for the interaction data captured by the multimodal information measurement environment. type is a parameter that specifies target signal out of *Speaking* for handling utterance information, *GazingAtUser* for handling gaze information of the other participants, and a *GazingAtObject* for handling gaze information of the discussion materials. subject is a parameter that specifies the role information of the activity subject (participant) set as a type parameter. If the type is *GazingAtUser*/*GazingAtObject*, the object of the action can be specified as a target parameter from the role information/document semantics. time parameter indicates the arbitrary time interval (e.g., 1, 5, or 10s). This rule detects the interval sections if the section includes the information specified by type, subject, and target more / less than a specific rate, where inequality (more/less) and rate (%) parameters are specified for the detection. The detected discussion sections are stored in the working memory as data corresponding to the interaction primitive layer.

Integration Rule: {layer, function, [arg], [constraint]} · · · (2)

This is a rule for accumulating interpretations in correspondence with each layer of the hierarchical interpretation model based on the data detected by the initialization rule.

layer is a parameter that sets which of the four layers of the hierarchical interpretation model corresponds to the interaction interpretation of the detected data section from this rule. With a function, it is possible to set a function that considers the temporal relationship of the discussion interval (e.g., *Overlap(P1, P2, ...)*, *All(P1, P2, ...)*, *Before(P1, P2)*, *After(P1, P2)*, *Switch(P1, P2)*). Without going into detail in this paper, the arguments of this function are specified in [arg] (e.g., *P1* as arg[0] and *P2* as arg[1]). The function is executed if the data intervals specified in other rules exist in the working memory. [constraint] is a parameter for giving detailed constraints on the role information of the subject/object for the data intervals specified in [arg]. In the integration rule, the data intervals are detected based on the forward chaining method.

Advive Generation Rule: {target_section, feedback} ・・・ (3)
This rule is used for generating advice considering the discussion interval detected by the integration rule as a reflection section. target_section indicates the data intervals (stored in working memory) detected by the integration rule as the reflection interval. In feedback parameter, the rule designer can set the advice that should be checked during this interval in a template format that mixes natural language with specified role information and document semantics (see Section 4.1).

4. Reflection Support Environment

4.1 Rule Creation Support System

Figure 1 shows the rule creation support system. This system consists of a rule creation area and a rule confirmation area. This system has a function to create the three rules explained in Section 3.3.
Function of initialization rule creation (the left side of Fig. 1): This function is used to specify the initialization rule, assuming that the rule corresponds to the interaction primitive layer of the hierarchical interpretation model (Fig. 1(a)). The type parameter for the target interaction data can be specified in the area of Fig. 1(b). Based on the set type, other parameters (i.e., subject, target, rate, inequality, and time) can be specified in the area of Fig. 1(c), respectively. The configured rule is displayed rule confirmation area along with the rule name entered in Fig. 1(d).
Function of integration rule creation (Fig. 2(i)): This function regulates the integration rule. The layer of the hierarchical interpretation model corresponding to the rule can be selected in Fig. 2(i-a). By specifying the function to be applied in Fig. 2(i-b), the corresponding parameters of the function (i.e., [arg] and its [constraint] if necessary) can be set in this area. Figure 3 is the situation where the function is set to 'Overlap,' which detects the intersection of arbitrary intervals. The set rule is listed at the bottom of the rule confirmation area along with the rule name entered in Fig. 2(i-d).

Here, for example, we can set up the initialization and integration rules to capture "*All participants became silent after proposer explains the experimental purpose* (integration rule (7))" by stacking the interpretation as shown in Table 1.

Table 1. *Example of Initialization Rules and Integration Rules*

Example of Initialization Rules	
(1)	{type='Speaking', subject='Proposer', target=null, rate=40(%), inequality='more', time=10(s)} => "*Proposer is speaking*"
(2)	{type='GazingAtObject', subject='Proposer', target='Experimental objectives', rate=60(%), inequality='more', time=10(s)} => "*Proposer is gazing at experimental purpose area in discussion materials*"
(3)	{type='Speaking', subject='Proposer', target=null, rate=10(%), inequality='less', time=10(s)} => "*Proposer is not speaking*"
(4)	{type='Speaking', subject='Teacher', target=null, rate=10(%), inequality='less', time=10(s)} => "*Teacher is not speaking*"
Example of Integration Rules	
(5)	{layer='Interaction-Event', function='Overlap', arg=["*Proposer is speaking*", "*Proposer is gazing at experimental purpose area in discussion materials*"], constraint=[arg[0].subject='Proposer(X)', arg[1].subject='Proposer(X)']} => "*Proposer is explaining experimental purpose*"
(6)	{layer='Interaction-Event', function='Overlap', arg=["*Proposer is not speaking*", "*Teacher is not speaking*"], constraint=[arg[0].subject='Everyone', arg[1].subject='Everyone']} => "*All participants are silent*"
(7)	{layer='Interaction-Event', function='Switch', arg=["*Proposer is explaining experimental purpose*", "*All participants are silent*"], constraint=null} => "*All participants became silent after proposer explains the experimental purpose*"

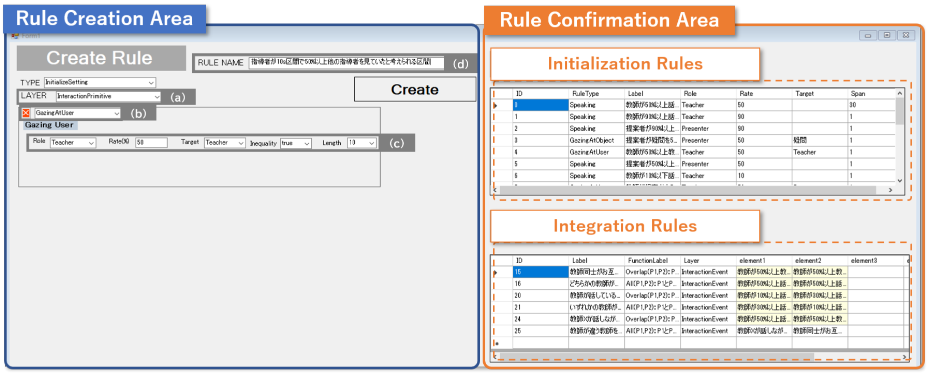
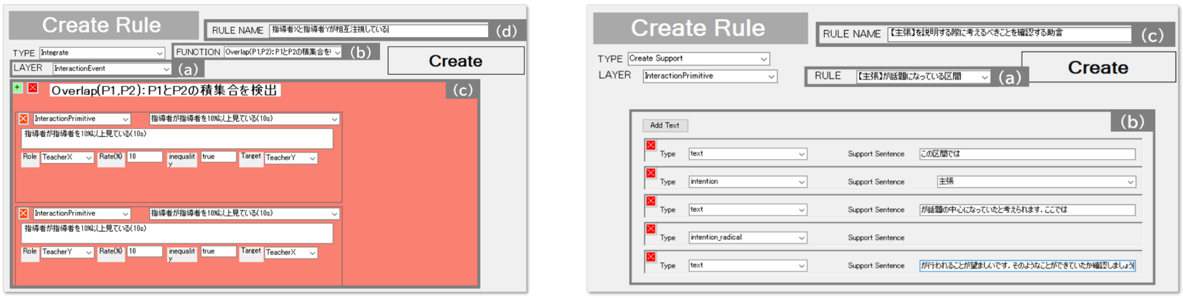


Figure 1. Rule Creation Support System.

Function of advice generation rule creation (Fig. 2(ii)): This is a function used to interpret the detected discussion interval as reflection sections by selecting a rule name as target and setting the concrete advice as feedback. For the target data, the rule name specified in the integration rule can be selected from the drop-down list in Fig. 2(ii-a). The feedback can be set in the area shown in Fig. 2(ii-b). In this area, the rule designer can specify the advice as an array of template formats, including free description text, role information, document semantics, and rational relationships of document semantics (e.g., sub-activity of the research content semantics and consistency of the logical composition of the intent semantics). Here, as advice corresponding to the data interval detected by the integration rule described above ("*All participants became silent after proposer explains the experimental purpose*"), it is possible to create templates as shown below as an example, in order to encourage reflection focusing on the thinking activity in the timing of the preparation for the discussion.

- In this discussion section, all participants became silent after you (Proposer) explained ["experimental purpose" (*concept of research content semantics*)]. In ["proposal" (*concept of intent semantics*)] of ["experimental purpose" (*concept of research content semantics*)], it is desirable to examine ["thinking about experimental subjects", "thinking about evaluation methods" (*sub-activity of "experimental purpose" in research content semantics*)] in advance.
- It is also desirable to examine ["thinking about the validity of proposals and assumptions" (*concept of intent semantics*)] before the discussion. Let's reflect on these points whether you examined them enough before the discussion.



(i) Integration rule (ii) Advice generation rule

Figure 2. Integration Rule and Advice Generation Rule Creation Area.

4.2 *Reflection Support System*

Figure 3 represents the interface of the developed reflection support system. Before using the system, the user (learner) needs to select a target discussion session for reflection and assign the role information of each participant.

This system has a basic video reflection function that can confirm the synchronized participants' videos and the discussion materials used in the discussion. The user can check the discussion from a given time by operating the seek bar (Fig. 3(a)). In addition, the system has the following two characteristic functions that satisfy the mechanism for providing advice according to the focused reflection section (R3).

Detection function of the target reflection section: Based on the interaction data corresponding to the target discussion session and the role information assigned to the participants, the system applies the initialization and integration rules defined in the rule creation support system at the startup. Then, it

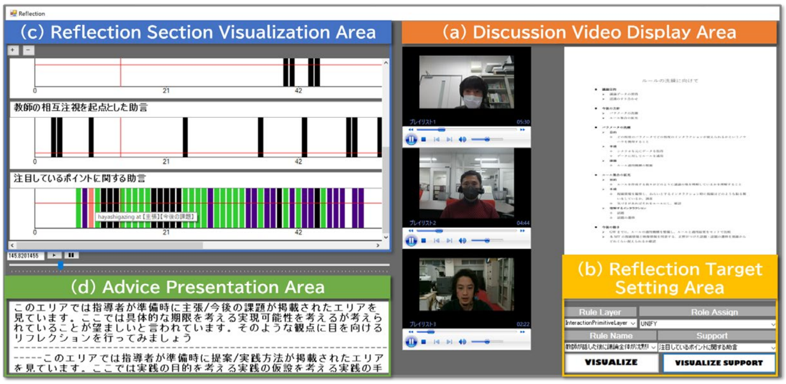


Figure 3. Reflection Support System.

stores the results of detected discussion intervals in the working memory. When there is a result that matches target_section of the advice generation rule, the system generates the instance of feedback of the corresponding rule based on the set template. The names of advice generation rules detected in this way can be available for selection in the area shown in Fig. 3(b).

Visualization function of the reflection target section: When a learner selects the names of advice generation rules as a reflection target section in Fig. 3(b), the detected target sections are displayed in the visualization area in a chart format alongside the discussion time series (Fig. 3(c)). The learner can check the details of the interaction occurring in the arbitrary target section by mouse over operation and can confirm the discussion video from that point by clicking on it. In addition, the advice corresponding to associated with clicked reflection target is displayed in the advice presentation area (Fig. 3(d)).

In this way, the system provides functions that allow learners to concentrate on reflection activities based on the advice, which follows the contents of discussion materials and discussion interaction context, encouraging reflection of thinking activities at the time of preparation.

5. Conclusions

In this study, we discussed a mechanism of system processing that provides advice that prompts a reflection on the discussion context of the research content, including thinking activities at the time of preparation for the creative discussion. At this initial stage, we have confirmed that the proposed system works properly by using experimental data from several sessions. Therefore, future tasks include verifying the effectiveness of the reflection support in the context of authentic research activities.

References

Burgoon, J. K., Magnenat-Thalmann, N., Pantic, M., & Vinciarelli, A. (Eds.). (2017). *Social signal processing*, Cambridge University Press.

Matsuoka, T., Seta, K., & Hayashi, Y. (2019). Internal self-conversation support system by iteration on reflective thinking and research documentation, *Procedia Computer Science*, Vol. 159, pp. 2102–2109.

Mori, N., Hayashi, Y., & Seta, K. (2019). Ontology-based thought organization support system to prompt readiness of intention sharing and its long-term practice. *The Journal of Information and Systems in Education*, Vol. 18, No. 1, pp. 27–39.

Sugimoto, A., Hayashi, Y., & Seta K. (2020). Multimodal interaction-aware integrated platform for CSCL. *Proc. of 22nd International Conference on Human-Computer Interaction*, Vol. 12185, pp. 264–277.

Sumi, Y., Yano, M., & Nishida, T. (2010). Analysis environment of conversational structure with nonverbal multimodal data. *Proc. of International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*, Article No. 44.

Thiran, J. P., Marques, F., & Bourlard, H. (Eds.). (2009). *Multimodal signal processing: Theory and applications for human-computer interaction*. Academic Press.