Evaluation of Discussions in Online Classrooms

Aiman Badri, Floriana Grasso, and Paul Leng

Department of Computer Science University of Liverpool, UK {uscasd,floriana,phl}@csc.liv.ac.uk

Abstract. Classroom discussion is an important element in many forms of online collaborative learning. We here discuss the design of systems to support tutors in the tasks of moderating such discussions and evaluating the contributions of participants. We describe the architecture of a system that will use techniques of conversation analysis and dialogue game theory to provide the instructor with information relevant to these tasks. The research draws on our experience in the running of an online degree programme.

1 Introduction

In recent years computer supported cooperative learning (CSCL) has become one of the fastest growing research areas in education technology. CSCL involves the study of how information and communication technology can be exploited as a tool alongside more traditional methods of collaborative learning (e.g. peer tutoring, simulation, games etc.). Notwithstanding the interest in this area within both the Education and the Artificial Intelligence communities, few systems have been implemented so far which fully exploit the benefits of more sophisticated AI techniques, such as user modelling, or natural language processing.

In many forms of cooperative learning, classroom *discussion* plays a key role. Research in education has a long tradition of study into pedagogical strategies and conversational mechanisms (see e.g. [1]), but these have rarely been exploited in computerised tutoring. The importance of the issue has been recognised, however [2], prompting investigations into the application of conversation analysis, especially threaded discourse, negotiation and computer-mediated argumentation for learning. Research has primarily focussed on the modelling of students, as the recipients of the education materials, and few have emphasised the importance of the role of the moderator of the classroom discussion [3]. In our research, conversely, we take the instructor's perspective, rather than the student's, as we seek to produce both a theoretical account of how discussion in a virtual classroom takes place, and practical tools to assist the instructor in the moderating activity. Methodologically, we base our research on the analysis of a corpus of classroom interactions from an on-line Master's degree in Information Technology. In this paper we present proposals for a framework for the analysis of classroom discussions in this context, intended as the basis of a system to support the role of the instructor/moderator.

V. Palade, R.J. Howlett, and L.C. Jain (Eds.): KES 2003, LNAI 2774, pp. 193-200, 2003.

[©] Springer-Verlag Berlin Heidelberg 2003



Fig. 1. A Virtual Classroom

2 Scenario

The context for the work we describe is an MSc degree programme that is being delivered entirely online. Within this programme, described more fully in [4], modules are taught to classes of no more than 20 students, who interact asynchronously with each other and their instructor through email exchanges using the SoftArc FirstClassTM system [5]. Each module is subdivided into *seminars*, each taking place over a period of one week, during which a topic of study introduced by the instructor is explored by the class. This exploration may include practical work, sometimes in groups, and coursework assignments of various kinds. In every case, however, the seminar includes also a discussion of aspects of the topic being examined, carried out through email contributions to an open folder in the module virtual classroom (Fig. 1).

This classroom discussion has a central role in the teaching and learning paradigm used: it demands the active engagement of students in the learning process, and promotes collaborative learning and sharing of experiences among the class. Much research, however (e.g: [6]), while underlining the value of this kind of interaction, has pointed out that without direction from a moderator, a proportion of students will not participate in the discussion. For this reason, contribution to the classroom discussion is a requirement on students, and a component in the assessment of the module. It is, of course, the responsibility of the instructor both to guide the discussion, where necessary, and finally to evaluate the contributions of students for assessment. The aim of our research is to explore the design of a system to support the instructor in these tasks, by providing him/her with a structured and objective analysis of the discourse.

3 A Framework for Monitoring Classroom Discussion

We present in this section a proposal for a framework to monitor classroom discussions, in the form of a system's architecture (see Fig. 2). Our purpose in this paper is to identify the main issues involved, and the techniques required to approach each of them. As opposed to other systems, that we survey below, we mainly concentrate on the analysis of the discussion dynamics.



Fig. 2. System's Architecture

We assume that the system operates on "contributions", i.e. textual messages posted by the various students to the virtual classroom. The system analyses and evaluates these contributions, with the final aim of providing the instructor with a summary of the virtual class activities. This process involves various stages, each of which is represented by a separate component, described below.

3.1 The Filter Component

The Filter selects the contributions which are relevant to the job at hand, by filtering out, for instance, repeated or out of topic contributions. Our analysis of the corpus of class conversations has revealed that it is not uncommon for students to mistakenly send more one copy of a contribution to the class, and the discussion folder inevitably also contains a proportion of messages that are not relevant to the topic of discussion. There are several ways to approach the problem of relevance. Teufel and Moens [7] propose a mechanism to decide whether a text is aligned with another, by determining the characteristic properties of the aligned sentences according to a number of features: presence of particular phrases, location in the text, sentence length, occurrence of thematic words (characteristic to the content of the document) and occurrence of proper names. The Filter module will operate on the basis of knowledge of previous contributions to the classroom discussion and of the current discussion topic, both of which will be included in the Dictionaries the system will use. Care must be taken in defining and using this knowledge: Burstein & Marcu [8] argue it is important that the classifier is topic independent for positional, lexical and semantic features, so as to be portable to other topics, something which is of particular significance in our scenario. They adopt this strategy in their work in analysing English essays [9]. Makagonov & Alexandrov [10] use domain oriented

dictionaries for a quasi-semantic analysis of a text for its relevance. Caron [11] applies Latent Semantic Analysis, which uses statistical correlation between words within a body of documents to infer underlying semantic connections between both the document and the words. The Filter will apply techniques of these kinds to produce a list of relevant contributions that can proceed to the next stage of the process. The irrelevant contributions will form part of the report to be addressed to the instructor, to help, for instance, identify students who constantly disrupt the conversation with off-topic contributions.

3.2 The Tag Supplier Component

The aim of the Tag Supplier is to identify and categorise the single discussion elements in each contribution, tagging them with extra information based on their content. Well known techniques from machine learning are available, for instance which treat the text as a "bag of words" and compare it to a predefined source [12]. We are interested, however, not only in identifying what the contribution is about, but especially in understanding how it relates to the rest of the discussion, and to create a dialogue meta-structure in which all the contributions are represented. Annotation schemes can be used which mark the elements relevant to the process, for example the structure of the text, or the type of activity the text is describing [13, 14]. We are especially interested in schemes that view a contribution as encoding one of a set of moves, expressing transitions in a dialogue [15]. The task is not simple. First it must be decided what are the "basic units" of the contribution: for example, single clauses, prosodic units, dialogue turns, sentences or intentionally defined discourse segments [16]. It has been argued that errors in discourse segmentation greatly influence the quality of the interpretation [17]; larger basic units, while computationally better, may leave significant rhetorical information out [18]. Secondly, it must be decided which types of rhetorical relation can exist among textual units [19]. There is debate on the number and the nature of the relations that should be used [20]. For instance, [21] reports a taxonomy of as many as 29 relations used for tagging the interaction between a tutor and a student.

While the content of the single contribution is crucial to the analysis, our aim is primarily to model the discussion meta-structure, so we plan to use shallower natural language processing techniques to retrieve the contribution content [13], and will focus on how the student's contribution relates to others in the discussion [15]. We wish to capture the intensity of the discussion, the chain of responses, discussion turns and responses to responses created in the conversation [22]. The notion of a dialogue move, as an abstraction which captures the participation of an "agent" in the discussion, helps in this task [23]. Dialogue moves are defined on the basis of factors such as cognitive plausibility, ease of coding, reliability and computational tractability [15]. The analysis of the corpus will guide us on how many and which dialogue moves are needed, but we expect these to include at least:

- 1. comments, that is original contributions;
- 2. questions, either addressed to the classroom, or to a particular individual (a fellow student, or the instructor);
- 3. reference to other messages, either to agree, or disagree, or provide extra information in support, etc.
- 4. reference to external sources, like web pages, excerpts of the text book, etc.

The environment in which the discussion takes place includes a number of features such as thread labelling that will be helpful in implementation.

3.3 The Discussion Manager Component

The Discussion Manager's role is to take the tagged discussion elements and use them to construct an image of the virtual classroom for the instructor. We envisage two main outputs: an Activity/Social Map and a Thread/Discussion Map.

The Activity/Social Map is a visualization of the activities of the students in the virtual class indicating the level of participation of each student. The map is typically constructed from statistical data on the contribution (date, time and owner) and from some evaluation of the dialogue moves that are assigned to the contribution. This evaluation can be approached as a holistic score based on the general quality of the contribution [9], or as a statistical analysis through predefined semantic values, rhetorical relations, inclusion of phrasal index terms, document-specific keyword frequencies, location heuristics, assessment of sentence similarity structure, and so on [24]. The best solution will be based on an analysis of how human instructors operate, and we plan to include a certain amount of personalisation to the instructor in this. Also, different graphical representations will be investigated for visualisation of the students' activities. A traditional way would depict data, whether raw statistics or results of the classification process, in a neutral quantitative way. We plan, conversely, to use a semantic approach, which makes use of knowledge about the social meaning of the categories defined to create a visualisation that reflects that knowledge [25].

The Thread/Discussion Map will map the level of discussion of the class for a certain thread of discussion. The map presents different types of dialogue acts between contributions as well as the evaluation of a contribution in terms of its content. Again, different graphical representations will be implemented to produce a map that best visualises the students' discussions. One approach is to track all student activities, such as number of postings, number of replies to posting, etc. An alternative is to track threads, recording each thread's temporal development and structure. We prefer a semantic approach here as well, which can represent the feel of the conversation as well as depicting its dynamics, by introducing assessments of the meaning and relevance of data into the visualisation process [25].

3.4 The Reporter Component

The final module, the Reporter, will be an alerting service for the benefit of the instructor, informing about events that happened and events that did not materialise. The Reporter could be activated each time the instructor logs to the system, or might also generate alerts to the instructor's mobile phone, for example. This is also intended to be personalised to the model of the instructors, to reflect his or her preferred attitude towards the classroom dynamics. A set of rules can be used to, for instance, signal inappropriate behaviour from a student, flag absence from the classroom, or draw attention to questions directed to the instructor, and so on.

4 Conclusion

The work described here is directed at providing intelligent support for instructors in online classes in which classroom discussion is a central feature of the pedagogy. We have presented the framework of a system that will help the instructor in the analysis and evaluation of these discussions, to assist in providing a more objective and consistent basis for the roles of moderation and assessment. We have identified relevant techniques, especially drawn from conversation analysis and dialogue game theory, that can be applied to the task. We now propose to develop a prototype implementation of this system, which we will evaluate in the context of our online degree programme. The system's output will be compared against the instructors' evaluation from the existing corpus of discussions. This corpus, together with the history of assessment and other instructor involvement, will provide us with data to validate and compare approaches experimentally.

References

- Brown, J., A.Collins, P.Duguid: Situated Cognition and the Culture of Learning. Educational Researcher 18 (1989) 32–42 193
- [2] Baker, M.: A model for negotiation in teaching-learning dialogues. Journal or Artificial Intelligence in Education 5 (1994) 199–254 193
- [3] Salmon, G.: E-Moderating: The Key to Teaching and Learning Online. Kogan Page (2000) 193
- [4] Grasso, F., Leng, P.: Issues of quality in on-line degree programmes. In Ghaoui, C., ed.: Usability Evaluation of Online Learning Programs. Idea Group Publishing (2003) 198-217 194
- [5] Persico, D., Manca, S.: Use of firstclass as a collaborative learning environment. Innovations in Education and Training International 37 (2000) 34–41 194
- Klemm, W., Snell, J.: Enriching computer-mediated group learning by coupling constructivism with collaborative learning. Journal of Instructional Science and Technology 1 (1996) 194
- [7] Teufel, S., Moens, M.: Sentence extraction and rhetorical classification for flexible abstracts. In: AAAI Spring Symposium on Intelligent Text Summarization, Stanford. (1988) 195

- [8] Burstein, J., Marcu, D., Andreyev, S., Chodorow, M.: Towards automatic classification of discourse elements in essays. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL'01). (2001) 195
- [9] Burstein, J., Marcu, D.: Towards using text summarization for essay-based feedback. In: La 7e Conference Annuelle sur Le Traitement Automatique des Langues Naturelles TALN'2000, Lausanne, Switzerland. (2000) 195, 197
- [10] Makagonov, P., Alexandrov, M.: Analysis of the contents of educational courses with special statistical and graphical methods. In Davis, G.B., ed.: Teleteaching '98 - Distance Learning, Training and Education, IFIP World Computer Congress 1998. (1998) 679–684 195
- [11] Caron, J.: Experiments with LSA scoring: Optimal rank and basis. In Berry, M., ed.: Computational Information Retrieval. SIAM Proceedings in Applied Mathematics. SIAM, Philadelphia (2001) 196
- [12] Lewis, D., Ringuette, M.: A comparison of two learning algorithms for text categorization. In: Proceedings of SDAIR-94, 3rd Annual Symposium on Document Analysis and Information Retrieval. (1994) 81–93 196
- [13] Riloff, E., Lehnert, W.: Information exteraction as a basis for high-precision text classification. ACM Transactions on Information Systems 12 (1994) 296–333 196
- [14] Zaenen, A., Larsson, S.: Annotations for instructional texts. Technical Report Deliverable D1.2, Task Oriented Instructional Dialogue (TRINDI) project (LE4-8314) (1999) 196
- [15] Cooper, R., Larsson, S., Matheson, C., Poesio, M., Traum, D.: Coding instructional dialogue for information states. Technical Report Deliverable D1.1, Task Oriented Instructional Dialogue (TRINDI) project (LE4-8314) (1999) 196
- [16] Carlson, L., Marcu, D., Okurowski, M.E.: Building a discourse-tagged corpus in the framework of rhetorical structure theory. In van Kuppevelt, J., Smith, R., eds.: Current Directions in Discourse and Dialogue. Kluwer Academic Publishers (to appear) 196
- [17] Marcu, D.: A decision-based approach to rhetorical parsing. In: Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL'99). (1999) 365–372 196
- [18] Marcu, D., Romera, M., Amorrortu, E.: Experiments in constructing a corpus of discourse trees: Problems, annotation choices, issues. In: Workshop on Levels of Representation in Discourse, Edinburgh, Scotland. (1999) 71–78 196
- [19] Marcu, D.: The rhetorical parsing of natural language text. In: Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics (ACL'97). (1997) 196
- [20] Hovy, E., Maier, E.: Parsimonious or profligate: How many and which discourse structure relations? available at http://www.isi.edu/naturallanguage/people/hovy/papers/93discproc.pdf (1995) 196
- [21] Marcu, D., Romera, M., Amorrortu, E.: Experiments in constructing a corpus of discourse trees. In: The ACL'99 Workshop on Standards and Tools for Discourse Tagging, Maryland. (1999) 48–57 196
- [22] Smith, M.: Tools for navigating large social cyberspaces. Communications of the ACM $45~(2002)~51{-}55~196$
- [23] Traum, D., Bohlin, P., Bos, J., Ericson, S., Larsson, S., Lewin, I., Matheson, C., Milward, D.: Dialogue dynamics and levels of interaction. Technical Report Deliverable D3.1, Task Oriented Instructional Dialogue (TRINDI) project (LE4-8314) (2000) 196

200 Aiman Badri et al.

- [24] Kupiec, J., Pedersen, J., Chen, F.: A trainable document summarizer. In: Proceedings of the 18th annual international ACM SIGIR Conference on Research and development in Information Retrieval. (1995) 68–73 197
- [25] Donath, J.: A semantic approach to visualizing online conversations. Communications of the ACM 45 (2002) 45–49 197