（様式４）

博士論文内容要旨

氏名（本籍）	中山 祐貴（栃木県）
学位の種類	博士（理工学）
学位記番号	
学位授与年月日	
学位授与の要件	学位規則第 10 条第 1 項該当
研究科	福島大学大学院共生システム理工学研究科（博士課程）
学位論文題目	Algorithms for Extraction of Discussion Processes from Communication Media and Their Applications (コミュニケーション媒体中の議論過程抽出手法とその応用)
論文審査委員	主　査　中村 勝一（福島大学・准教授） 副　査　董　　彦文（福島大学・教授） 副　査　神長 裕明（福島大学・教授） 副　査　藤本 勝成（福島大学・教授） 副　査　宮寺 庸造（東京学芸大学・教授）

論文目次

論文内容要旨

Discussions using communication media (e.g., e-mail and BBS) are widely carried out in various intellectual activities such as research, cooperative learning, and software development. It is important but generally difficult for participants to successfully grasp the processes of discussions and their results since the number of utterances in communication media continuously increases as discussions progress. To overcome these issues, many methods have been proposed for discussion support. However, they have not been effective enough since most target only one of the three main phases of discussion support: “grasp discussion topics”, “manage descriptions that build discussions”, or “grasp discussion results”. There is thus a strong need for a method that supports grasping the processes of discussions and their overviews in association with these phases. A promising approach is to extract and visualize “descriptions that build a discussion and their relationships”. Although a few research projects have taken this approach, they impose strong restrictions, such as description style, on users. Moreover, analysis granularity of most existing methods is too rough. Consequently, the goal of this research was to develop a support method for grasping discussion processes by using a novel approach that extracts discussion processes from utterances in communication media in a way capable of handling the diversity of response granularity without severe restrictions and visually presents the extracted discussion processes to users.

Initially, this research developed the algorithms for extracting the discussion processes from utterance data accumulated in communication media. These algorithms extract the discussion processes by analyzing the attributes of utterances and quotation relationships between utterances without excessively depending on natural language technology. Specifically, at first, the descriptions that build a discussion are clustered by topic on the basis of analyses of structural features of utterances and quoted descriptions. Then, valid utterances are extracted on the basis of analyses of punctuation marks, characteristic words and phrases while the invalid utterances are removed. Next, the pairs of one description and another responding to it (a "response pair") are diagnosed. Here, a discussion process is complicated since it generally includes the exchanges of responses built in diverse description granularity such as a single utterance, multiple utterances, and one message (e.g., one whole e-mail). In this research, to identify such description granularity of each exchange and its reply relationships, styles of response descriptions were sorted into six typical patterns on the basis of careful examinations of the actual discussions using communication media. To diagnose response pairs, each unit message (e.g., an e-mail) is analyzed and its form of appearance made up of quotations and utterances in reply is extracted as a bit sequence (an "utterance set"). Then, response granularity of respective exchanges is identified by comparing the extracted utterance set with the typical patterns of response description styles. On the basis of this diagnosis, the response pairs are extracted. Moreover, the discussion structure is extracted by recursively connecting response pairs on the basis of an analysis of a shared description. Consequently, the proposed method extracted the discussion processes at a more detailed level than existing methods for responses with diverse styles without imposing severe restrictions on users.

Next, a system for visualizing the discussion processes was designed and developed by introducing the proposed algorithms. This system targets discussions using e-mail which is one of the main communication media. It visualizes the discussion processes in e-mails to support observation and comprehension of discussion processes. The system requirements were identified on the basis of observation of the actual discussion situations and used to carefully design the system. Then, a prototype was implemented. This prototype is equipped with support functions to manage the discussion topic, manage descriptions that build each discussion, visualize the discussion processes, give a bird's-eye view of the discussion structure, and provide the relevant user interfaces. In particular, the system visually shows the discussion processes as a graph in which the descriptions building a discussion and their connections are expressed as nodes and edges, respectively. Moreover, the system visually suggests a user's viewing position in a big discussion structure and provides an environment in which to browse the descriptions that build a discussion in multiple viewpoint levels (particle size). In this dissertation, the actual images of grasping of the discussion processes using the system are also presented with execution examples of the main functions. Consequently, the system enables both the overall discussion structure and details of the focused descriptions to be cooperatively grasped.

Experiments using actual utterance data were conducted to verify the accuracy of the extraction algorithms and practical effectiveness of the system for visualizing the discussion processes. Specifically, extraction accuracy was calculated and the time required for the work was measured. In addition, experimental collaborators were interviewed and actual situations of descriptions were scrutinized. Here, elaborate verifications were conducted on the basis of collation of these results. Consequently, the algorithms were confirmed to be reliably accurate. Excellent values were obtained for recall ratio in particular, which is significant in this research. Moreover, the system was apparently effective at grasping the discussion processes. These characteristics of the proposed method were identified through careful examination with little precedent in previous researches. Furthermore, findings were obtained that include the tendencies of how people pick out the utterances and hints for improving the system.