Design of an Intelligent Agent for Stimulating Brainstorming

Chun-Hsiang Wang
Computer Science Department,
National Chengchi University,
Taipei, Taiwan
102703039@nccu.edu.tw

Tsai-Yen Li
Computer Science Department,
National Chengchi University,
Taipei, Taiwan
li@nccu.edu.tw

ABSTRACT

In recent years, brainstorming has gradually become a mainstream approach used by groups of people to collect ideas before making decisions. This approach is useful in handling the predicament of lacking ideas on specific topics for individual members, and breaking through the limitations of their own experiences. However, in the process of group brainstorming, there exist some common problems such as the lack of comprehensive discussion contents due to the limited experiences recalled by the members, the stagnation in the progress, and so on. Therefore, in this paper, we propose the design of an intelligent agent system that plays the vital role of facilitator to participate in the brainstorming process; in other words, our system can discuss a designated topic on an on-line chatroom with other members in a brainstorming discussion. This system collects textual data and establishes the knowledge model in a specific domain with machine learning methods prior to the brainstorming. Then, in the brainstorming process, it attempts to conjecture the topic of the current dialogues, determine the progress, and generate responses to the brainstorming chatroom in order to come up with diverse ideas or extend the present topic. The goal of this system is to increase the diversity and profundity of discussions, and make the whole process smooth and fruitful. The results of our experiments show that our intelligent agent is effective in helping on-line brainstorming, especially idea generations.

Keywords

Brainstorming; Intelligent Agent; Knowledge Model; Machine Learning; Natural Language Processing.

1. INTRODUCTION

Brainstorming has been widely used to collect ideas since it can help resolve the problem of encountering a bottleneck in generating ideas for solving a specific problem by an individual or by a group in the setting of traditional discussion. However, not all brainstorming sessions are fruitful. Some unsuccessful sessions may result from the absence of a facilitator who can lead the discussions such that divergent ideas can be elicited and converged at a later stage. Hence, it is important to have a discussion facilitator who can guide the members in the brainstorming process and ensure the smooth progress and useful results of the discussion.

Recently, the studies of natural language processing and artificial intelligence has become prevailing. In this work, we have made use of the recent advances in these research topics to design an intelligent agent system to facilitate online brainstorming. As shown in Figure 1, the intelligent agent acts as a team member, named Jack, in a brainstorming session and proactively send out textual messages at appropriate times to guide other members to participate in the discussion in the right direction according to the function of a given session. This intelligent agent system relies on a knowledge model built from the data collected from social media to stimulate the team members to increase both the depth and the breadth of their thoughts and compensate for the missing elements of the current solution.

In this paper, we have used novel machine learning algorithms to cluster the textual information obtained from the social media as the knowledge model of the system in the brainstorming process. In addition, we have chosen the Chinese language as the target language to be used in the brainstorming session. Thus, our intelligent agent system will take advantage of Chinese natural language processing techniques including text segmentation, word vector, and information extraction.

Based the established knowledge model, our intelligent agent system uses a rule-based model to recognize and control the progress of the brainstorming process effectively. This system can identify the degree of divergence and convergence in the conversations to determine the current progress and the succeeding subgoal of the discussion. Moreover, this system also collects dialogues from online discussion forums, calculates word vectors, and conducts sentiment analysis. Then, we manually create lists of sentences for specific contexts such that the system can directly reply to other members with sentences which match particular situations, increasing the diversity of responses. In order to generate appropriate responses, the system establishes a number of dialogues in advance whose keywords are hollowed out. When the system knows the situation from the discussion progress model, it will consider the emotional intention and fill in the vacancy of the an appropriate sentence chosen from the candidate list.

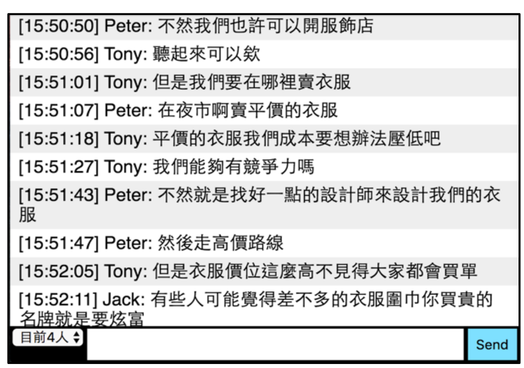

Figure 1. System interface

2. RELATED WORKS

In previous years, chatbots established on instant messaging software or web-based Q&A systems has attracted much attentions about research in the design of interactive agents in both academics and industries. For example, the Q&A system for psychological problems used a large amount of dialogues with questions and answers originally from the web to respond to users with similar questions [1]. In addition, there are researches aiming to analyze the data from Wikipedia to design an interactive intelligent tutoring system [2]. Furthermore, a previous work attempts to construct an intelligent agent system to answer students' open-ended questions on Earth Science, allowing students to learn deeper concepts during the process of discussion [3].

The techniques developed for natural language understanding is a crucial part in these applications. In [4], the authors discussed several approaches to extracting useful information from texts, and focused on the ways of handling unstructured documents [4]. Besides, in [5], the authors targeted on the customer services on social media, and aimed to classify emotions concealed from textual reviews of customers [5].

Regarding dialogue generation, one research based on information retrieval and phrase-based statistical machine translation techniques presented several data-driven approaches to automatically generating responses to open-domain linguistic stimuli on Twitter [6]. Moreover, in [7], the authors designed a recurrent neural network architecture for data-driven response generation trained from large quantities of unstructured Twitter conversations [7].

In most of the current applications of conversational agents, researchers usually train the domain-specific data as the background knowledges of systems. And then, the systems intend to return different dialogues according to the similarity between the current messages from users and the corpora from their background knowledges. In these applications, the agents are usually rather passive and triggered only by the questions from the users and then respond to questions with the most appropriate answers. However, the intelligent agent in this work needs to be more proactive since we expect it to act as a facilitator in the brainstorming process. In other words, our agent system is not triggered by the keywords in the dialogue from a specific user, but determines the current progress according to the dialogues from all members and then generate responses at proper moments in order to extend the current discussion on the present topic.

3. SYSTEM

3.1 System Overview

In this paper, we have designed an intelligent agent system to take part in the process of group brainstorming consisting of several people discussing a specific topic. Our intelligent agent system stimulates members to discuss efficiently and intends to increase the completeness of the discussion result. The architecture of our intelligent agent system (See Figure 2) comprises three information models including *agent knowledge model* (*AKM*), *discussion progress model* (*DPM*), and *response generation model* (*RGM*). These three models are used in different situations in a collaborative way to generate suitable responses at appropriate moments.

Our agent system utilizes the data collected from social media automatically to construct the corpora in specific domains as the background knowledge. Then, at run time, our agent system takes the conversations from all other members and processes them with natural language processing techniques in order to understand the contents that the members are discussing and determine the progress of the whole brainstorming process. And, according to the judged result, our system automatically generates responses which are sent to the chatroom to increase the diversity of aspects and the completeness of the discussions. The experiments in this paper used a web-based online chatroom designed by ourselves as the brainstorming interface, and our intelligent agent system acted as one of members in the discussion to facilitate brainstorming.

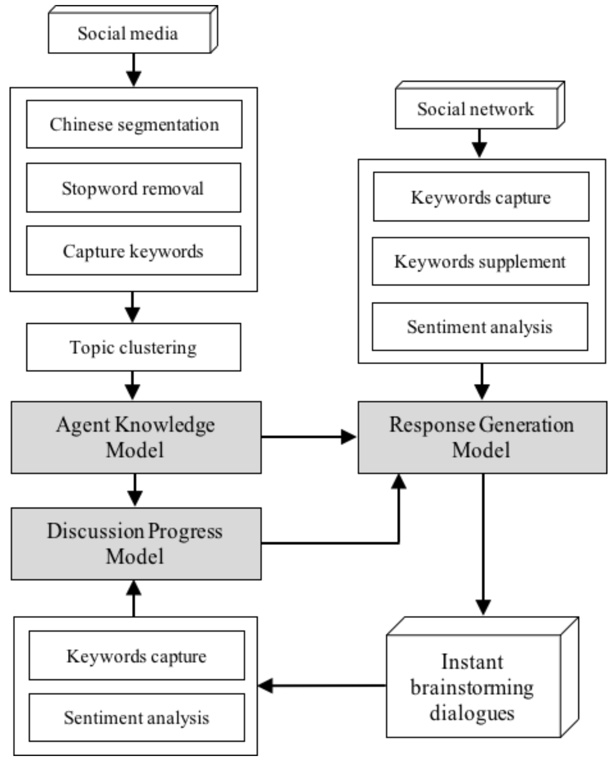

Figure 2. System architecture

The discussion procedure that a traditional brainstorming session follows is shown in Figure 3. We formulate system tactics for all procedures on the basis of discussion procedures depicted in Figure 4. In the beginning, the system agent proposes some introductory remarks to warm up the session, and then the members start to brainstorm by bringing up ideas to solve the proposed problem. For the discussion process, we divide the entire session into two stages. The first stage is considered as the topic divergence stage, in which all members can raise different ideas and have rudimentary discussions. The second stage focuses on topic convergence. In this stage, all members debate to find more feasible and practical solutions based on the proposed ideas. If there exist disagreements, the members may also return to consider other possible solutions. The members will repeat this discussion cycle until consensus is reached and the convergence stage ends. Our intelligent agent takes part in the group discussions, and leads the brainstorming process such that the session can be run in a smooth and effective matter.

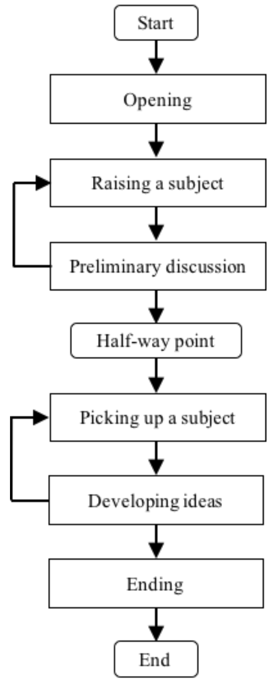

Figure 3. Flowchart for traditional brainstorming

3.2 Agent Knowledge Model (AKM)

There are three main steps in the construction of agent knowledge model. In this model, we collected a good number of Chinese documents from popular social media such as PTT, the largest bulletin board system in Taiwan, and Wikipedia. And then we conducted Chinese word segmentation after removing stopwords.

The collected documents are stored as term sets $\mathcal{D}_i$ consisting of a set of terms $\mathcal{P}_k$ representing the contents of original documents.

$$\mathcal{D}_i = \{\mathcal{P}_1|\mathcal{P}_2|\ldots|\mathcal{P}_t\}\ for\ i = 1\ to\ N \tag{1}$$
$$N\text{: amount of documents}$$
$$t\text{: amount of terms}$$
$$\mathcal{P}_i\text{: the } i^{th} \text{ term}$$
$$\mathcal{D}_i\text{: the } i^{th} \text{ term set}$$

Then, we used TextRank [8], a graph-based ranking model, to calculate the importance of every term from each document. After we stored all the values as the keyword scores $\mathcal{S}_k^i$, we sifted the keywords from all terms in each document according to their scores.

$$\mathcal{S}_k^i = TextRank(\mathcal{P}_k)\ for\ \mathcal{P}_k\ in\ \mathcal{D}_i \tag{2}$$
$$\mathcal{S}_k^i\text{: keyword score of } k^{th} \text{ term in } i^{th} \text{document}$$

In this paper, we picked up the top 10% scores from TextRank as the most representative terms for each document and kept them as the new term sets $\mathcal{D}_i'$ for all documents.

$$\mathcal{D}_i' = \left\{\mathcal{S}_k^i \middle| \mathcal{S}_k^i\ is\ the\ top\ 10\%\ in\ \mathcal{D}_i\right\}\ for\ i = 1\ to\ N \tag{3}$$
$$\mathcal{D}_i'\text{: the new } i^{th} \text{term set}$$

Subsequently, we combined all keywords from all new term sets as the corpus $\mathcal{D}_c$ for the background knowledge of our intelligent agent system.

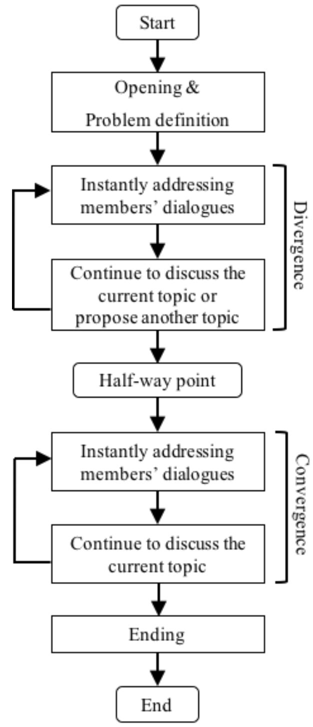

Figure 4. Flowchart for system strategies

$$\mathcal{D}_c = \mathcal{D}_1' \cup \mathcal{D}_2' \cup \ldots \cup \mathcal{D}_N' \tag{4}$$
$$\mathcal{D}_c\text{: corpus set}$$

Finally, we treated each term from the corpus $\mathcal{D}_c$ as a unit, and adopted the clustering algorithm, K-Means [9], to conduct topic clustering for all keywords and get the term set $\mathcal{T}_s$ for each topic cluster labelled s. We saved the clustering results including all keyword terms and their labels into our database as the agent knowledge model.

$$\mathcal{T}_s = \{\mathcal{P}_i|\text{kmeans}(\mathcal{P}_i) \in s\}\ for\ \mathcal{P}_i\ in\ \mathcal{D}_c \tag{5}$$
$$s\text{: topic cluster label}$$
$$\mathcal{T}_s\text{: term set for label } s$$

3.3 Discussion Progress Model (DPM)

In this paper, we established the discussion progress model by designing response strategies for all stages (See Figure 4) during the brainstorming session according to the flow chart depicted in Figure 3. First, our intelligent agent system greeted other members and described the topic to be discussed. Then, the discussion would go through the divergence stage as well as the subsequent convergence stage. The discussion progress model determined the progress and the current situation by the conversation records until this model thinks that the progress has reached the end. Lastly, the model made the conclusion and finished this discussion.

In this model, we modularize the discussion procedure with a rule-based approach. Two algorithms are used during the divergence stage and the convergence stage, separately.

With the discussion progress model, the system reads conversations among all members, marks the order of sentences, and saves all dialogue records during the brainstorming process as dialogues set $\mathcal{I}$. In addition, in this model, the system considers the r latest sentences as the recent dialogues, and congregates the recent dialogues into the recent dialogues set $\mathcal{R}$. The system updates both the dialogues set and the recent dialogues set continuously during the discussions.

$$\mathcal{I} = \mathcal{I}_1 \cdot \mathcal{I}_2 \cdot \ldots \cdot \mathcal{I}_i \tag{6}$$
$$\mathcal{I}_k\text{: the } k^{th} \text{ dialogue}, i\text{: the amount of dialogues now}$$
$$\mathcal{I}\text{: the dialogues set}$$
$$\mathcal{R} = \mathcal{I}_{i-r+1} \cdot \mathcal{I}_{i-r+2} \cdot \ldots \cdot \mathcal{I}_i \tag{7}$$
$$r\text{: the amount of dialogues seen as recent dialogues}$$
$$\mathcal{R}\text{: the recent dialogues set}$$

The system computes the number of terms belonging to a specific label for all terms in the overall dialogues set and the recent dialogues set, respectively, with the agent knowledge model. The labeling results of the two sets are represented as the overall label result $\mathcal{L}_{\mathcal{I}s}$ and the recent label result $\mathcal{L}_{\mathcal{R}s}$. The overall label result $\mathcal{L}_{\mathcal{I}s}$ indicates the distribution of topics in the overall discussion, while the recent label result $\mathcal{L}_{\mathcal{R}s}$ indicates the current composition of topics in the more recent discussion. This system uses different rules to interpret and make use of the two results in different stages of the brainstorming.

$$\mathcal{L}_{\mathcal{I}s} = |\{i \mid AKM(i) \in s\}| \text{ for } i \text{ in } \mathcal{I} \tag{8}$$
$$\mathcal{L}_{\mathcal{R}s} = |\{i \mid AKM(i) \in s\}| \text{ for } i \text{ in } \mathcal{R} \tag{9}$$
$$s\text{: topic cluster label}$$
$$AKM\text{: get label of the term from AKM}$$
$$\mathcal{L}_{ks}\text{: the label result of dialogues set } k \text{ for cluster label } s$$

In Figure 5., we proposed an algorithm to choose the label that the next response should belong to. In the divergence stage, we design a threshold τ_1 to determine if the discussion needs to switch to another topic in order to help the group members to consider problem from different aspects. If not, this model extends the current topic by responding with related sentences of the same topic so as to assist others in having further discussions on the current topic.

In the divergent stage, the system arranges the topics discussed in recent dialogues in a sorted order according to the number of dialogues for the current topic. If the number is less than the threshold, this system will continue to lead the discussion on the same topic by proposing a related response. Otherwise, this system will change to another topic with the number of dialogues smaller than the threshold.

We consider the behavior of changing discussion topic when the current topic has accumulated enough dialogues as the autonomous divergence of DPM. The topic switching will continue until the accumulated numbers of dialogues for all topics have exceeded the threshold τ_1. In this case, we consider that the brainstorming has reached the half-way point, and the model will start to enter the convergence process.

In the course of convergence, we have designed another algorithm, as shown in Figure 6, to deal with the process, and two threshold τ_2 and τ_3 are used to the determinate the progress. τ_2 is compared with the topic labels of the recent dialogues while τ_3 is contrasted with those of the overall conversations. The main strategy in this model is to respond with the contents most related to the topic at this moment during the convergence process until the extent of the discussion on the primary topic is sufficient.

In DPM, the system first selects the topic with the highest number of labels in the specific cluster as the dominant subject of the recent conversation. And we determine if the extent of the agitating topic has achieved the ideal condition by comparing the two abovementioned quantities with the two thresholds separately. If the two conditions are met, this system will bring the brainstorming session to an end with some concluding dialogues. Conversely, if any of the threshold is not satisfied, the system will extend the discussions to help members have deeper communication.

Algorithm 1. Divergence process

1. $\mathcal{L}'_{\mathcal{R}} \leftarrow$ label queue records the result by sorting all $\mathcal{L}_{\mathcal{R}s}$;
2. $for\ \ell\ in\ \mathcal{L}'_{\mathcal{R}}$:
   a. $if\ \mathcal{L}_{\mathcal{I}\ell} == 0$:
      1) $\ell \leftarrow$ randomly select a label from ℓ to the last label of $\mathcal{L}'_{\mathcal{R}}$;
      2) return label ℓ;
   b. $if\ \mathcal{L}_{\mathcal{I}\ell} < \tau_1$:
      1) return label ℓ;
3. end divergence process and enter the next stage.

Figure 5. Algorithm for divergence process

Algorithm 2. Convergence process

1. $\ell \leftarrow$ the top label of the result by sorting all $\mathcal{L}_{\mathcal{R}s}$;
2. $if\ \mathcal{L}_{\mathcal{R}\ell} < \tau_2$:
   a. return label ℓ;
3. $if\ \mathcal{L}_{\mathcal{I}\ell} < \tau_3$:
   a. return label ℓ;
4. end the convergence process and enter the end of brainstorming.

Figure 6. Algorithm for convergence process

Furthermore, in DPM, the system record and analyze the dialogues in the brainstorming session to decide whether this model needs to stimulate further discussions no matter of the stage that the brainstorming is in. By asking the particular member to propose ideas or inquire all other members' thoughts about the current topic, our system attempt to help all members to deliberate upon the relevant contents for discussion topics and let the whole brainstorming process advance successfully.

3.4 Response Generation Model (RGM)

The Response Generation Model (RGM) aims to generate a good variety of responses to the chatroom based on the needed discussion topic in the current stage of a brainstorming session.

We collected a large amount of messages by crawling contents on websites, such as Dcard, a social network website for university students based in Taiwan. Then we conducted Chinese segmentation on each message $\mathcal{M}_i$ to separate the sentence or phrase into several terms $\mathcal{P}_k$. After the segmentation, we adopted the language model, TextRank [8], to calculate the importance score $\mathcal{S}_k^i$ of each term of all sentences. Lastly, we selected the most important term with the highest score as the keyword $\mathcal{K}_i$ for each message.

$$\mathcal{M}_i = \mathcal{P}_1 \cdot \mathcal{P}_2 \cdot \ldots \cdot \mathcal{P}_t \tag{10}$$
$$\mathcal{P}_k\text{: the } k^{th} \text{ term}$$
$$\mathcal{M}_i\text{: the } i^{th} \text{ message}$$
$$\mathcal{S}_k^i = TextRank(\mathcal{P}_k) \text{ for } \mathcal{P}_k \text{ in } \mathcal{M}_i \tag{11}$$
$$\mathcal{S}_k^i\text{: the importance score of } k^{th} \text{ term of the } i^{th} \text{ message}$$
$$\mathcal{K}_i = \mathcal{P}_k \text{ if } \mathcal{S}_k^i \text{ is the top in } \mathcal{M}_i \tag{12}$$
$$\mathcal{K}_i\text{: the keyword of the } i^{th} \text{ message}$$

In RGM, we replaced the keyword of each message with a symbol representing a vacancy to be filled by another word.

$$\mathcal{M}'_i = \mathcal{P}_1 \cdot \mathcal{P}_2 \cdot \mathcal{P}_{\mathscr{k}} \cdot \ldots \cdot \mathcal{P}_t \;\; let\; \mathcal{P}_{\mathscr{k}} = e \qquad (13)$$
$$where\; \mathcal{P}_{\mathscr{k}} = \mathcal{K}_i \; and\; 1 \le \mathscr{k} \le t$$
$$e\text{: } sign\; for\; empty\; mark$$

In addition, we utilized the contents of Chinese documents from the online encyclopedia, Wikipedia, as the training data for the neural networks in the Word2Vec algorithm [10]. The Word2Vec algorithm produced word embeddings, which can be used to find the terms $\mathcal{E}$ within a distance ℓ from the keyword of each message. The model assumed that the relationship between these terms and the keyword is closely interrelated for the given context. Consequently, these terms can be integrated with the keyword into a set of new keyword extension $\mathcal{K}'_i$ for every message.

$$\mathcal{E} = \{\mathcal{P} \mid |wTv(\mathcal{P}) - wTv(\mathcal{K}_i)| < \ell\} \qquad (14)$$
$$wTv\text{: } use\; word2vec\; to\; get\; word\; embedding$$
$$\ell\text{: } the\; distance\; limit$$
$$\mathcal{E}\text{: } the\; extension\; set\; for\; keyword$$
$$\mathcal{K}'_i = \{\mathcal{K}_i\} \cup \mathcal{E} \qquad (15)$$

On the other hand, in our RGM, we employed the traditional classification algorithm, Naïve Bayesian [11], to train the model for emotion analysis by the data from the rating reviews on websites. And the model predicted all messages for the hidden emotion values $\mathcal{N}$. Finally, our RGM incorporated the corresponding extension keyword set and emotion value into every message and kept them as the extension list $\mathcal{C}$ for possible responses in extending the current topic.

$$\mathcal{N}_i = NaiveBayes(\mathcal{M}_i) \qquad (16)$$
$$\mathcal{N}_i\text{: } the\; latent\; emotion\; value$$
$$\mathcal{C} = \{(\mathcal{M}'_i | \mathcal{K}'_i | \mathcal{N}_i)\} \qquad (17)$$
$$\mathcal{C}\text{: } the\; list\; for\; possible\; responses$$

List for opening:

1. 大家好 (Hello everyone)
2. 大家都準備好了嗎 (Are you all ready?)

List for stimulating discussion:

1. 關於這個有什麼想法嗎 (Any idea about this?)
2. 聽起來不錯，還有類似的想法嗎 (It sounds good, and any similar idea about this?)

List for convergence:

1. 如果是___這方面呢 (How about ___ ?)
2. 關於___，大家有任何想法嗎 (Any idea about ___?)

Figure 7. Artificial lists for various situations

We also manually built up several lists of sentences which will be used when the brainstorming session is in the opening, ending, or stimulating stages, and this model can directly reply with the whole sentences to the chatroom at their corresponding situations (See Figure 7.). One of the lists is used in the convergence stage to change discussion topics. This list consists of several open-ended sentences with empty symbols like the sentences in the extension list, and a keyword from another cluster is chosen and inserted to form a complete response for further discussions.

In the brainstorming process, our RGM takes different steps in accordance with the given conditions as shown in Figure 8. According to the computed result from DPM, our system gets the cluster label that needs to be replied at that moment as well as the corresponding strategy for topic extension or shift. If the RGM decided to apply responses from the lists for specific situations, such as opening or stimulating the discussion, this model will select randomly a complete sentence like "大家好" (Hello everyone.), and "關於這個，有人有其他想法嗎" (Does anyone have any idea about this?), to be sent to the chatroom.

Algorithm 3. Response generation

1. ℓ、$\mathscr{s}$ ← get the label and current situation from DPM;
2. if $\mathscr{s}$ is not for continuation and not for divergence：
   a. return a response from the specified set $\mathcal{A}_{\mathscr{k}}$ randomly;
3. if $\mathscr{s}$ is for divergence：
   a. t ← get a term from the target ℓ cluster randomly;
   b. combine a sentence from the set $\mathcal{A}_d$ with t and return;
4. if $\mathscr{s}$ is for continuation：
   a. accumulate the average emotion value $\mathcal{N}'$ from $\mathcal{R}'$;
   b. $\mathscr{k}$ ← get the keyword from $\mathcal{R}'$;
   c. if $\mathscr{k}$ is in ℓ cluster：
      1) $t \leftarrow \mathscr{k}$;
   d. $else$：
      1) t ← get a term from the target ℓ cluster randomly;
   e. $\mathscr{r} \leftarrow \{(\mathcal{M}'_i|\mathcal{K}'_i|\mathcal{N}_i) | t\; in\; \mathcal{K}'_i\}$ for all 3-tuple in $\mathcal{C}$;
   f. $\mathcal{M}'$、d ← get $\mathcal{M}'_i$ and the distance by finding the least distance $|\mathcal{N}_i - \mathcal{N}'|$ from $\mathscr{r}$;
   g. if $d < \tau_4$：
      1) combine $\mathcal{M}'_i$ with t and return;
   h. $else$：
      1) pass this time.

Figure 8. Algorithm for response generation

However, if the progress is in the divergence stage, this model will use the AGM to find the term in the target cluster, insert it into the open-ended sentence, and generate the response.

Moreover, in the situation of topic extension, this model adopted an approach similar to the way for determining the progress in DPM. This model picks up fewer dialogues from the recent dialogues set as the immediate dialogues set $\mathcal{R}'$ that is regarded as a reference to the current atmosphere and emotions.

$$\mathcal{R}' = \mathcal{J}_{i-\mathscr{s}+1} \cdot \mathcal{J}_{i-\mathscr{s}+2} \cdot \ldots \cdot \mathcal{J}_i \; where\; \mathscr{s} < \mathscr{r} \qquad (18)$$
$$\mathscr{s}\text{: } the\; amount\; of\; dialogues\; seen\; as\; a\; reference\; to\; emotions$$
$$\mathcal{R}'\text{: } the\; immediate\; dialogues\; set$$

In order to generate responses to extend recent subjects, RGM predicts the emotions of all sentences in the immediate dialogues set $\mathcal{R}'$ and stores their average as the current average emotion $\mathcal{N}'$.

$$\mathcal{N}' = AVG\big(NaiveBayes(\mathcal{J}_{\mathscr{k}})\big)\; for\; \mathcal{J}_{\mathscr{k}}\; in\; \mathcal{R}' \qquad (19)$$
$$\mathcal{N}'\text{: } the\; current\; average\; emotion$$

With this model, our system takes the current average emotion as the standard for the most appropriate response and selects some keywords from the immediate dialogues set as the probable terms with high priorities. If all the probable terms are not satisfactory due to the disparities between cluster labels and the target label, our system will arbitrarily choose the term in the target cluster. After ensuring the keyword, our system searches for the most befitting combination with the extension keyword set containing the keyword and the possible responses with the smallest difference between emotion values. Furthermore, we set up a threshold τ_4 to ensure that the difference between two emotion values is not too high so as to avoid members' indescribable feelings. Therefore, if the

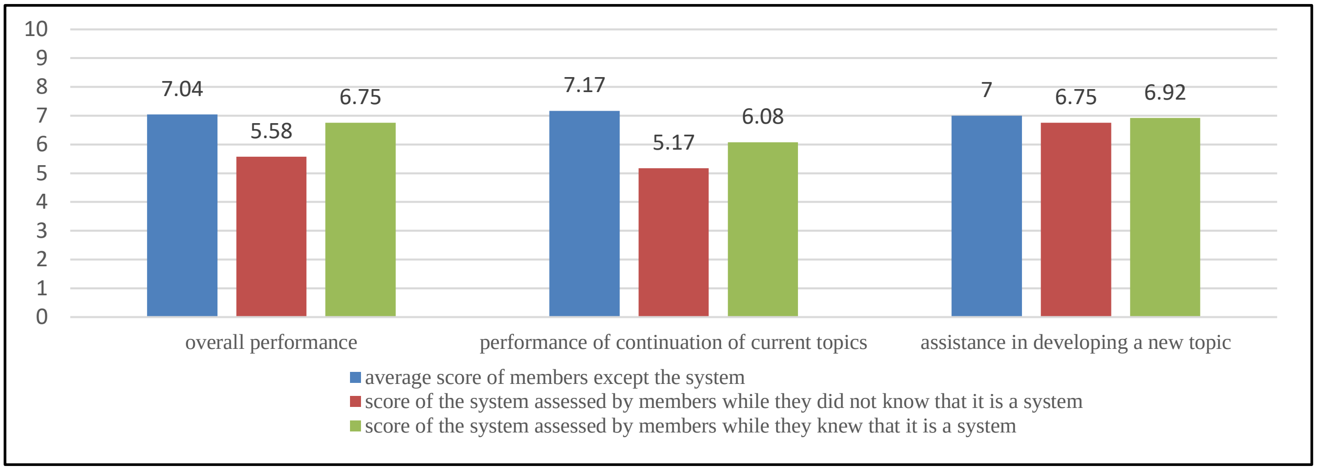

Figure 9. Experimental results on human evaluation

difference between the emotion of the most proper sentence and the current average emotion exceeds the threshold τ_4, this model will choose to skip this time.

4. EXPERIMENTS

4.1 Experimental Settings

In this paper, we have designed an online chatroom as our experiment interface, and all users can communicate with other members anonymously to participate in the brainstorming. We have conducted four experiments and, in each experiment, we invited three subjects to attend the brainstorming session with our intelligent agent system in the circumstance that they did not know who other members were. We assigned identity numbers to the four members at random in the entire discussion.

After the brainstorming sessions were over, we requested three subjects to fill out a two-stage questionnaire based on the performance of the other three members in three aspects including the overall performance, the continuation of the same topic, and the competence of topic convergence, respectively. Each subject are asked to evaluate other members separately with scores from zero to ten.

All members answered the questionnaire in the first stage immediately after the brainstorming session; however, before the second stage, all members were informed that one of the members in the brainstorming was played by the system and asked them to guess which member our system was. Then, members evaluated the system on the three aspects again. Finally, we presented the result with three values including the average score of the other three members, and the scores by other members on grading our system in two conditions: knowing and not knowing the identity of the system agent.

4.2 Experimental Results

As shown in Figure 9, the experiment result reveals that when the subjects did not know one of the members is a system agent, the score of the system for the extension of current topics is 5.17. On the other hand, the score for the assistance in the topic divergence is 6.75, which is close to the average score of the other three members. Therefore, we can conclude that the performance of our system for divergence is effective. We think the main reason is that the way to generate responses by directly inserting selected keywords into the empty slots of the sentences is useful in expressing new topics during the divergence process. The subjects can easily catch the new topic from the responses of our system agent and, hence, can successfully switch to discuss other topics. However, for the continuation of the same topic, our system did not do very well probably because of the approach that the response generation model picks up the most appropriate sentence only by emotions may not be sufficient. The meanings of the terms in Chinese often varied from sentence to sentence, so the responses by our system agent may not be coherent all the time.

Besides, although all members correctly guessed which identity our system is, the scores for three performances of our system increased significantly after members knew that some of the responses were generated by our system automatically during the whole brainstorming session. Especially, the score for the overall performance raised from 5.58 to 6.75, and the score for the continuation of the current topic increased to 6.08, which represented that all subjects generally thought that the responses of our system had positive influences on the brainstorming process.

We examined the results of the dialogues in the four experiments and analyzed the brainstorming processes. We found that there was one obvious deviation for the amount of dialogues from opening to the half-way point because the topics of the conversations between all members exceeded the background knowledge of our system. Therefore, no matter which topics that the members mentioned, the system could not effectively accumulate the amounts of predefined cluster labels. Additionally, from the half-way point to the ending, one experiment contains the least number of sentences because all members concentrated on two topics without frequent shifts and, consequently, the amount of dialogues for these two labels accumulated quickly.

5. CONCLUSIONS

In this paper, we have designed an intelligent agent system which actively takes part in the brainstorming discussions and automatically generate responses in the chatroom. Our system is different from the applications of traditional chatbots and question-and-answering systems because of its active role. Therefore, we believe that it is a novel application in the area of natural language processing. Our intelligent agent system utilizes agent knowledge model and discussion progress model to determine the current progress of the brainstorming and adopt different strategies in accordance with the situations for topic divergence and convergence. Our system can lead the discussions with our response generation model for different topics and, therefore, can stimulate discussions and help the brainstorming process to proceed smoothly.

However, our intelligent agent system has not focused on the semantic analysis very well. As a result, the contexts of responses by our system may sometimes deviate from the current conversation. Although the subjects still can understand which topic our system intends to talk about, but it is detrimental to the user experiences. In the future, we will consider to improve this problem and conduct more work on semantic analysis for our intelligent agent system. We aim to let our system create responses which are more corresponding to the contexts of current conversations.

6. ACKNOWLEDGMENTS

This work is financially supported by the Ministry of Science and Technology in Taiwan under project contract number MOST 105-2815-C-004-027-E.

7. REFERENCES

[1] Y. Liu, M. Liu, X. Wang, L. Wang and J. Li. "PAL: A Chatterbot System for Answering Domain-specific Questions." *ACL*. 2013.

[2] K. Lee, P.H. Seo, J. Choi, S. Koo and G.G. Lee. "Conversational Knowledge Teaching Agent that Uses a Knowledge Base." *16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. 2015.

[3] H.C. Wang, C.P. Rosé and C.Y. Chang. "Agent-based Dynamic Support for Learning from Collaborative Brainstorming in Scientific Inquiry." *International Journal of Computer-Supported Collaborative Learning* 6.3: 371. 2011.

[4] R.J. Mooney and R. Bunescu. "Mining Knowledge from Text Using Information Extraction." *ACM SIGKDD explorations newsletter* 7.1: 3-10. 2015.

[5] J. Herzig, G. Feigenblat, A. Rafaeli and D. Altman. "Classifying Emotions in Customer Support Dialogues in Social Media." *17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. 2016.

[6] A. Ritter, C. Cherry and W.B. Dolan. "Data-driven Response Generation in Social Media." *Proceedings of the conference on empirical methods in natural language processing*. *ACL*. 2011.

[7] A. Sordoni, M. Galley, M. Auli, C. Brockett, Y. Ji, M. Mitchell, J.Y. Nie, J. Gao and B. Dolan. "A Neural Network Approach to Context-sensitive Generation of Conversational Responses." *arXiv preprint arXiv:1506.06714*. 2015.

[8] R. Mihalcea and P. Tarau. "TextRank: Bringing Order into Texts." *ACL*. 2004.

[9] K. Wagstaff, C. Cardie, S. Rogers and S. Schroedl. "Constrained K-means Clustering with Background Knowledge." *ICML. Vol. 1*. 2001.

[10] T. Mikolov, K. Chen, G. Corrado and J. Dean. "Efficient Estimation of Word Representations in Vector Space." *arXiv preprint arXiv:1301.3781*. 2013.

[11] A. Go, R Bhayani and L. Huang. "Twitter Sentiment Classification using Distant Supervision." *CS224N Project Report, Stanford 1.12*. 2009.

Columns on Last Page Should Be Made As Close As Possible to Equal Length

Authors' background

Your Name	Title*	Research Field	Personal website
Chun-Hsiang Wang	Research Assistant	Natural Language Processing, Machine Learning	
Tsai-Yen Li	Professor	Artificial Intelligence, Motion Planning, Interactive Storytelling	http://www3.nccu.edu.tw/~li/

***This form helps us to understand your paper better, the form itself will not be published.**

***Title can be chosen from: master student, Phd candidate, assistant professor, lecture, senior lecture, associate professor, full professor**