

Micro-Counseling Dialog System based on Semantic Content

Sangdo Han, Yonghee Kim, Gary Geunbae Lee

Pohang University of Science and Technology, Pohang, Republic of Korea
{hansd, ttiti07, gblee}@postech.ac.kr

Abstract. This paper introduces a text dialog system that can provide counseling dialog based on the semantic content of user utterances. We extract emotion-, problem-, and reason-oriented semantic contents from user utterances to generate micro-counseling system responses. Our counseling strategy follows micro-counseling techniques to build a working relationship with a client and to discover the client's concerns and problems. Extracting semantic contents allows the system to generate appropriate counseling responses for various user utterances. Experiments show that our system works well as a virtual counselor.

Keywords: Dialog system, counseling dialog system, micro-counseling technique, semantic content, back-off strategy

1 Introduction

People often talk with other people to share their situation and to relieve stress. However, other people are not always available, and we may not want to reveal all information because some of it may be too personal; a micro-counseling dialog system can solve these problems. In our previous work, the system could not understand various user utterances because it used only lexical information to analyze them [4]. In this work, we developed a system that analyzes semantic information to achieve understanding of user utterances and to effectively respond to them for counseling.

In this paper, we measure the effect of our new information extracting method, new counseling information, and chat-oriented back-off strategy. Our system can extract information from a wider variety of utterances and get higher scores for counseling satisfaction than the previous system.

Relevant related work is presented in section 2. Micro-counseling techniques are summarized in section 3. Corpus data are introduced in section 4, and the micro-counseling dialog method is described in section 5. The experiments and results are shown in section 6 and conclusion is drawn in section 7.

2 Related Work

Han et al. [4] used a conditional random field algorithm to extract “who, what, when, where, why, how” (5W1H) information to counsel, but because the system only considers 5W1H information, some system utterances that consider time and place are not relevant in a counseling dialog. For example, the system could generate utterance like “Where did you mad?”. In addition, because the method is based on only lexical information, it needs a large corpus to understand various user utterances. Furthermore, this method could not detect various user emotions because it was based on only keyword matching.

Meguro et al. [8] introduced a listening-oriented dialog system based on a model trained by a partially observable Markov decision process using human-human dialog corpus. The system uses a listening-oriented dialog strategy to encourage users to speak, but the system utterances are limited because it selects responses from the corpus. It also cannot respond to utterances that are not in the specific domain.

In this work, we extracted emotion-, problem-, and reason-oriented information by extracting general semantic contents (subject, predicate, and object), then using this information to guide selection of appropriate counseling responses. By redefining counseling information from 5W1H, the system focuses on the user’s current situation and emotional state. The new method extracts this information by analyzing general semantic contents, so it can extract the information from various domain-independent utterances. However, not all utterances are relevant sources of semantic contents for counseling, and the counseling system should respond to all user utterances in order to encourage the users to continue talking; in this case the system should adopt a “back-off strategy” in which it uses a chat-oriented system to respond with a relevant sentence that has no counseling value, but which encourages the client to continue interacting. Most chat-oriented systems (e.g., ELIZA [9], ALICE¹) are based on the simple pattern matching technique, but several systems are based on a sentence similarity measure (Lee et al., [6]; Li et al., [7]); they select the most similar sentence to the user input among example sentence pairs and generate modified sentence as an output.

3 Micro-counseling Techniques

Micro-counseling techniques are basic counseling techniques that make clients feel that a counselor listens carefully and understands the clients [3]. Micro-counseling includes four main techniques: attending, paraphrasing, reflecting feelings, and questioning.

¹ ALICE: Artificial Intelligence Foundation Inc. <http://www.alicebot.org>

Attending is a technique to react naturally to an utterance. Attending utterances could follow any kind of user utterances. This technique makes a client feel that the system focuses on him or her, and encourages the client to continue talking to the system. Examples include “Please tell me more” and “Continue”.

Paraphrasing is a technique to make the user think the system is following what the user said. Unlike attending, paraphrasing utterance is dependent on a user utterance because the system should rephrase the client’s utterance. For example, when client says “I ate pizza”, the counselor could say “Oh, you ate pizza”.

Reflecting feelings organizes the user’s whole situation. This technique is similar to paraphrasing but whereas paraphrasing follows exactly what the client said in the previous turn, reflecting feelings follows all information that the user provides. For example “You don’t feel good because John deleted it to stop it”, or “Stopping it made you sad”.

Questioning is a technique to ask a user to provide more counseling information, e.g. “How do you feel about it?”, or “Why did John do so?”.

4 Data Collection

We generated 512 utterances as a counseling corpus (Table 1). Because micro-counseling dialog is based on problems, feelings, and specific facts [3], our generated utterances focus on user’s problem, emotion and reason based on micro-counseling techniques. We generated the corpus based on 42 counseling situation (Table 2) and micro-counseling techniques. This corpus used to select micro-counseling utterances.

Table 1. Counseling Corpus Example

Speaker	Utterance
System	Hello. How are you today?
User	I feel bad because I fought with my boyfriend.
System	You fought with your boyfriend. Why does it happen?
User	He didn’t remember my birthday.
System	I see. You feel bad because he didn’t remember your birthday.
User	That’s right

A general chatting corpus was generated based on seven domain-independent dialog acts; it includes 11,328 user utterances. The corpus was generated by collecting chatting dialog between two people. It was used for micro-counseling utterance detection.

To generate counseling information extraction rules, we used Movie-Dic, which is a movie script corpus from 753 movies [1]. It includes 132,229 utterances, which we assume represent natural dialogs.

Table 2. Example of counseling situations

Emotion	Problem	Reason
Angry	I fought with John.	John yelled at me.
Sad	My dog died.	He fell from cliff.
Happy	My dad won the prize.	He got the best score.

5 Method

5.1 Architecture

Our system consists of four components: counseling utterance understanding (CUU), counseling strategy managing (CSM), counseling response generating (CRG) and a chat-oriented back-off dialog system. CUU understands what a user says, CSM decides what kind of strategy to use, and CRG decides how to generate counseling utterances. The chat-oriented dialog system is used to respond to general user utterances for which counseling utterances are difficult to generate (Fig. 1).

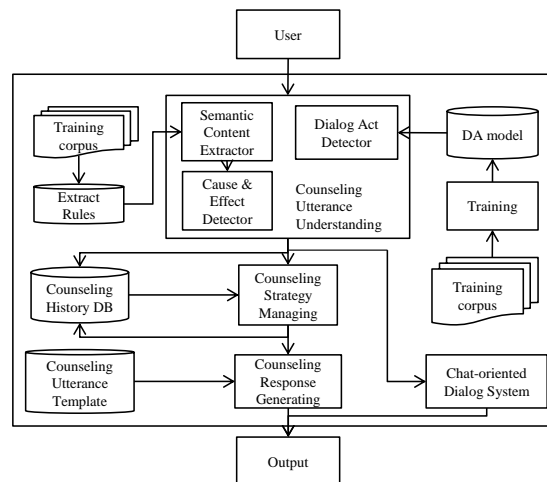


Fig. 1. System Architecture

5.2 *Counseling Utterance Understanding*

In the CUU module, the system first decides whether a user utterance is appropriate for micro-counseling dialog, then extracts counseling information. If the user utterance is not appropriate for a micro-counseling reaction, the chat-oriented dialog system generates a general response as back-off strategy.

Our system treats the utterances whose dialog act is a statement as appropriate utterances for micro-counseling dialog. Semantic contents to generate counseling response are mostly included in utterances whose dialog act is a statement because their purposes are to deliver information. To detect a statement dialog act, we used the MaxEnt algorithm [2] using a chatting corpus which is labeled with dialog act. We trained a model with word and Part of Speech (POS) bi-gram features to train the model.

As a second step, we check whether or not our system can extract semantic contents from the user utterance. If it cannot, the utterance is passed to the chat-oriented dialog system because we cannot generate a micro-counseling utterance. To extract semantic content, we use the dependency pattern matching method that is used in WOE^{parse} [10]. The dependency pattern is a partial dependency graph in which each node has a POS tag and each edge has a dependency label. Among those nodes, three nodes are marked as subject, predicate, and object. If a dependency pattern is found in the dependency graph of the user utterance, its corresponding subject, predicate, object phrases are extracted. We manually collected 360 dependency patterns from dependency graphs of the Movie-Dic corpus.

During a micro-counseling dialog, the system asks the user three types of questions: problem questions, reason questions, and emotion questions. Through the system questions, the system can detect a user utterance as the one that seeks counseling. For example, when a system asks the user about a problem, the user's answer is assumed to identify the problem.

Some user utterances can provide more than one type of counseling information. For example, "I feel sad because my dog died," includes two semantic contents: "I feel sad" is emotion information, and "my dog died" is problem information. To extract counseling information from this kind of double content utterance, we should consider the relationships between the types of counseling information (Fig 2). We split the user utterance's semantic contents into cause and effect by comparing the location of semantic contents and classifying the conjunction. For example, in case of "I am sad because my dog died", "my dog died" causes "I am sad" because it is after the conjunction 'because'. We generated 14 rules to split semantic contents into cause and effect.

When the system asked a problem question, cause is assumed to be reason information and effect is assumed to be emotion information.

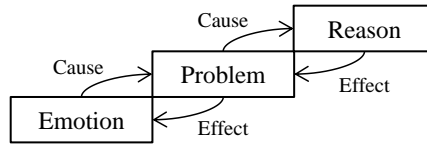


Fig. 2. Counseling Information Relationship

5.3 Counseling Strategy Manager

Table 3. Counseling Strategy Table. ‘O’: information should exist; ‘X’: information should not exist

Counseling Strategy	Information in Current User Utterance			Information in Dialog History		
	Emotion	Problem	Reason	Emotion	Problem	Reason
Attending	O					
Paraphrasing	O					
Reflect Feeling	X	O	O	O	X	X
	X	X	O	O	O	X
	O	X	X	X	O	O
		⋮			⋮	
Emotion Question	X			X		
Problem Question	X			X		
Reason Question	X			O	X	X

Our micro-counseling dialog system has four counseling strategies: attending, paraphrasing, reflect feeling, and question. We defined a counseling technique table that consists of strategies and the required conditions of each strategy (Table 3); the required conditions mean the existence of information in user’s current utterance and dialog history. The system selects the best strategy based on counseling technique table.

- **Attending:** Attending utterances could follow after any kind of user utterances. So attending technique does not consider information extracted from the current or dialog history.
- **Paraphrasing:** Paraphrasing should follow a user utterance that includes at least one counseling information.
- **Reflecting:** Reflecting feelings should be used when information in current user utterance and information in dialog history includes whole counseling information.
- **Questioning:** Questioning techniques should be used to request information that has not been provided; emotion, problem, and reason. In case of reason questioning, problem information should exist in information in dialog history because reason should be asked after problem already known.

As an example of counseling dialog strategy (Table 4), the system asks an emotion question or a problem question at the beginning of dialog to induce the user speak. Questioning can be in a dialog turn with others.

Table 4. Counseling Strategy Example.

Speaker	Utterance	Counseling Strategy
System	Hello. How are you today?	Emotion Question
User	I feel bad because I fought with my boyfriend.	-
System	You fought with your boyfriend. Why does it happen?	Paraphrase, Reason Question
User	He didn't remember my birthday.	-
System	I see. You feel bad because he didn't remember your birthday.	Paraphrase
User	That's right	-

5.4 Counseling Response Generation

Our system utterances are generated by using a counseling response template. We choose a system template by checking the counseling information extracted from the dialog, and use extracted contents to fill slots in a counseling response template (Table 5). Each technique has its own templates, and each template has its own counseling information slots to fill.

Table 5. Counseling Response Template. Slots: <es> subject of emotion; <ep> predicate of emotion; <eo> object of emotion; <ps> subject of problem; <pp> predicate of problem; <po> object of problem; <rs> subject of reason; <rp> predicate of reason; <ro> object of reason

System Template	Counseling Strategy
Oh I see.	Attending
You feel <eo>.	Paraphrasing
<es> <ep> <eo> because <ps> <pp> <po>.	Paraphrasing
You feel <eo> because <rs> <rp> <ro>.	Reflect Feeling
Please tell me about your problem.	Problem Question
How do you feel about <ps> did so?	Emotion Question
Why did <ps> do so?	Reason Question

5.5 Chat-Oriented Dialog System

The chat-oriented dialog system can respond to any kind of user input sentence whether or not it is related to the counseling purpose. The system selects the most appropriate response from the chatting cues given the user input. This is based on the EBDM [6] framework; detailed description is beyond the scope of this paper. We only explain the example matching method. An example is a pair of a user-side sentence u and a system-side response s . We adopt a sentence similarity score with POS weights (simPoS) to find the most appropriate responses as follows:

$$\text{sim}_{\text{PoS}}(u, s) = \frac{2 \cdot |u \cap s|}{|u| + |s|}$$

The intersection is the set of words that occur in both sentences. When finding a matching word, coarse-grained POS tags and lemmatized words are used to ignore inflectional changes of the words. We also define POS weights and assign the word weight according to its POS. Finally, $|u|$, $|s|$ and $|u \cap s|$ are defined as the sum of all word weights in u , s and $u \cap s$ respectively.

6 Experiment & Discussion

We first tested the performance of dialog act detection and semantic content extraction modules. Our 5-fold cross validation experiment test dataset includes a chatting corpus and a counseling corpus. The whole 11,840 utterances are labeled with dialog act, and semantic contents that can generate a counseling response. Our experiment achieved > 89% statement dialog act detection performance, and > 95% semantic content extraction performance as shown in Table 6.

Table 6. Dialog act and semantic content detection result

	Statement dialog-act detection	Semantic content extraction
Precision	88.9%	97.4%
Recall	89.6%	92.7%
F measure	89.3%	95.0%

We recruited 16 volunteers to evaluate the effectiveness of the counseling information extraction method, the counseling strategy, and the chat back-off strategy. The baseline system for comparison is a previous counseling dialog system that uses 5W1H extraction. We gave 20 counseling situations to each user and asked them to talk to each system for a total of 30 minutes.

Each volunteer scored six evaluation questions on a scale of 1(low) to 10. To assess the CUU module based on semantic content extraction, the questions were asked users how much they were satisfied by the system’s ability to understand their utterances. To assess the CSM module’s counseling strategy, the questions

were asked whether they were satisfied with its counseling strategy on the counseling information. To assess the back-off strategy we asked them to assess the relevance of its responses. Our system achieved a higher score overall than the baseline system (Table 7).

User satisfaction increased because the counseling information was extracted from various utterances. The redefined counseling information encouraged the user to interact intensively with the system. The chat-oriented back-off strategy increased overall satisfaction because it avoided interruption of dialogs.

Table 7. Experiment Result. ($p < 0.01$ for each question)

Question	Baseline	Proposed
System extracted appropriate information.	5.33	7.43
System understood my various utterances.	5.19	7.00
Information that system focused was appropriate.	5.90	7.19
System's dialog strategy was appropriate.	5.68	7.28
There was no interruption in my dialog.	6.43	9.19
I wanted to chat more with the system.	4.10	6.57

7 Conclusion

We developed a counseling dialog system that extracts semantic counseling information, defines counseling information, and uses a chat-oriented dialog system as a back-off strategy. Because the counseling dialog system was developed for various user utterances, it can be used for other research in human-computer interaction such as development of health informatics and companions for seniors. Our future work is to improve our system to generate various system utterances that use additional micro-counseling techniques [5].

Acknowledgments This work was partly supported by ICT R&D program of MSIP/IITP [10044508, Development of Non-Symbolic Approach-based Human-Like Self-Taught Learning Intelligence Technology] and National Research Foundation of Korea (NRF) [NRF-2014R1A2A1A01003041, Development of multi-party anticipatory knowledge-intensive natural language dialog system].

References

- [1] Rafael E. Banchs. 2012. Movie-DiC: a Movie Dialogue Corpus for Research and Development, *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, pp. 203-207, Jeju, Republic of Korea.

- [2] Adam L. Beger, Stephen A. Della Pietra, and Vincent J. Della Pietra. 1996. A Maximum Entropy Approach to Natural Language Processing, *Association for Computational Linguistics*, pp. 39-71
- [3] David R. Evans, Margaret T. Hearn, Max R. Uhlemann, and Allen E. Ivey. 2010. *Essential Interviewing*, Eighth edition. Cengage Learning.
- [4] Sangdo Han, Kyusong Lee, Donghyeon Lee, and Gary G. Lee. 2013. Counseling Dialog System with 5W1H Extraction, *In Proceedings of the SIGDIAL2013 Conference*, pp. 349-353, Metz, France.
- [5] Allen E. Ivey, Mary B. Ivey, and Carlos P. Zalaquett. 2013. *Intentional Interviewing and Counseling*, Eighth edition. Cengage Learning.
- [6] Cheongjae Lee, Sangkeun Jung, Seokhwan Kim, and Gary G. Lee. 2009. Example-based dialog modeling for practical multi-domain dialog system, *Speech Communication*, 51 (5), pp. 466–484
- [7] Yuhua Li, Zuhair Bandar, David McLean, and James O’Shea. 2004. A Method for Measuring Sentence Similarity and its Application to Conversational Agents, *The 17th International FLAIRS conference*, pp. 820-825, Florida, USA.
- [8] Toyomi Meguro, Yasuhiro Minami, Ryuichiro Higashinaka, and Kohji Dohsaka. 2013. Learning to Control Listening-Oriented Dialogue Using Partially Observable Markov Decision Processes, *ACM Transactions on Speech and Language Processing*, Vol. 10, No. 4, Article 15.
- [9] Joseph Weizenbaum. 1966. ELIZA - A Computer Program For the Study of Natural Language Communication Between Man and Machine, *Communications of the Association for Computing Machinery*, Vol 9, pp. 36-45.
- [10] Fei Wu and Daniel S. Weld. 2010. Open Information Extraction using Wikipedia, *In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, ACL ’10, pp. 118–127, Morristown, NJ, USA.