

# Counseling Dialog System with 5W1H Extraction

Sangdo Han, Kyusong Lee, Donghyeon Lee, Gary Geunbae Lee

Department of Computer Science and Engineering, POSTECH, South Korea

{hansd, kyusonglee, semko, gblee}@postech.ac.kr

## Abstract

In this paper, we introduce our counseling dialog system. Our system interacts with users by recognizing what the users say, predicting the context, and following the users' feelings. For this interaction, our system follows three basic counseling techniques: paraphrasing, asking open questions, and reflecting feelings. To follow counseling techniques, we extracted 5W1H information and user emotions from user utterances, and we generated system utterances while using the counseling techniques. We used the conditional random field algorithm to extract 5W1H information, and constructed our counseling algorithm using a dialog strategy that was based on counseling techniques. A total of 16 adults tested our system and rated it with a higher score as an interactive communicator compared with the baseline system.

## 1 Introduction

Over the past 45 years, suicide rates have increased by 60% worldwide.<sup>1</sup> To prevent suicide, suicide people need to counsel with counselors. However, counseling with a human counselor requires a substantial cost, and in addition, there is a location restriction. Developing a counseling dialog system could be an effective solution to address this problem because the system has no limitations with respect to time and location.

In this study, we present a counseling dialog system. The system interacts with users by recognizing what the users say, predicting the context, and following the users' feelings. We used three counseling techniques for our system, to interact with the users. The system performs paraphrasing, asks open questions, and reflects feelings.

---

<sup>1</sup>

[http://www.who.int/mental\\_health/prevention/suicide/suicideprevent/en/](http://www.who.int/mental_health/prevention/suicide/suicideprevent/en/)

Paraphrasing is a technique that paraphrases user utterances. For example, when a user utterance is "My dog picked up the ball", then it could be paraphrased by "Oh, your dog picked up the ball". The technique of asking open questions is to ask some questions to the user, to obtain more information. For example, when a user says "I played computer games", then the counselor could say "When did you play?" or "Where did you play?". Finally, reflecting a feeling is a similar technique to paraphrasing, but it includes emotional comments. For example, when a user says "My dog died. I'm so sad", then the counselor could say, "Oh, your dog died. You look depressed." or "You look so sad".

In our approach, we extract 5W1H (who, what, when, where, why, how) information and four basic emotions (happy, afraid, sad, and angry) from user utterances. We generate system utterances using 5W1H information and basic emotions.

## 2 Counseling Techniques

Counselors show empathy with clients by listening and understanding them. Clients feel comfortable by a counselor's attention. Counselors listen, ask questions, answer questions, and concentrate on clients. Attention and empathy is important for counseling. Counselors show interest and care about the clients' emotions. Our counseling dialog system also focused on attending and empathy.

Many counseling techniques are used in counseling. Basic attending, self-expression, and micro-training skills are introduced in Theron et al. (2008). Basic attending and self-expression skills are about non-verbal behavior, such as tone of voice and eye contact. Micro-training skills are the basic verbal counseling techniques that are learned for counseling beginners: open and closed questions, minimal encouragement, paraphrasing, reflection of feelings and summarization.

We chose three micro-training skills to attend and show empathy with clients. These skills are open questions, paraphrasing, and reflection of feelings because they are basic techniques to show emphasize effectively.

### 3 Related Work

The SEMAINE project aims to build a Sensitive Artificial Listeners (SAL) – conversational agents that are designed to interact with a human user through robust recognition and the generation of non-verbal behavior (Schröder et al., 2008). This system detects user emotions by multimodal sensors (camera, microphone). A virtual face in this system shows facial expressions based on user emotions, and it encourages the user to speak by reacting and asking questions. These techniques could show empathy with users. However, it has limited verbal skills because SEMAINE does not have language understanding module. In our research, our system follows user utterances and generates system utterances based on user’s 5W1H.

### 4 Data Collection

We generated 4,284 utterances by using fifty-three 5W1H information sets and four basic emotions (Figure 1). Each utterance could be generated by using part of the 5W1H information and four emotions.

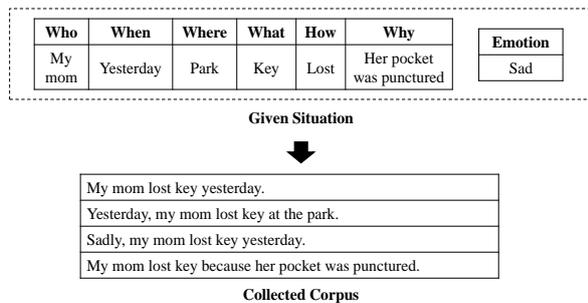


Figure 1. Counseling Corpus Collecting Process

We tagged each 5W1H element in each utterance and the user intention for each utterance (Table 1). The system’s actions were labeled by following counseling strategies which will be discussed in section 5.3.

Tagged Corpus	User Intention	System Action
<who>My mom</who> <how>lost</how> <what>a key</what> <when>yesterday</when>.	Inform_5W1H	Ask_Open_Question
<when>Yesterday</when>, <who>my mom</who> <how>lost</how> <what>a key</what> at the <where>park</where>.	Inform_5W1H	Paraphrase
<who>My mom</who> <how>lost</how> <what>a key</what> <when>yesterday</when>. I'm so sad.	Inform_5W1H_Emotion	Reflect_Feeling
I'm so sad.	Inform_Emotion	Reflect_Feeling
Thank you.	Thank	Welcome
Good bye.	Bye	Bye

Table 1. Corpus Tagging Examples

User intentions we defined can be separated in two groups: ‘counseling’ and ‘others’. Utterances in ‘counseling’ group include 5W1H information or emotional information. Utterances which do not including them are in ‘others’ group. Greetings, thanks, and farewells are included (Table 2).

Counseling group	Others group
Inform_5W1H, Inform_emotion, Inform_5W1H_emotion, ...	Thank, Bye, Greeting, Agree, Disagree, ...

Table 2. Two Separated Groups of User Intentions

## 5 Method

### 5.1 Architecture

Our system architecture is given in graph 2. When a user inputs a sentence, a natural language understanding (NLU) module understands the main action (the user’s intention) and extracts the 5W1H entities from the user’s utterance. The emotion detection module detects the user’s emotions using the emotional keyword dictionary. The dialog management module decides the system’s action from the main action and the 5W1H information from the trained module from the example dialog corpus. The natural language generation (NLG) module generates the system utterance using a system utterance template. We can generate the system utterance by replacing 5W1H slots with entities.

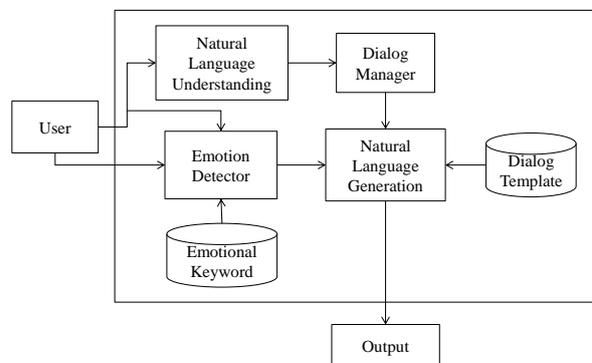


Figure 2. Counseling Dialog System Hierarchy

## 5.2 Natural Language Understanding

In our approach, the NLU module understands the user utterance by classifying the main action and the 5W1H entities from the user utterance. To classify user intention, we used maximum entropy model (Ratnaparkhi, 1998) trained on a linguistically motivated features. We used a lexical word features for the utterance model. The lexical word features are lexical trigrams using previous, current, and next lexical words. To extract 5W1H entities, we used a conditional random field (CRF) model (Laffery et al., 2001). We also used lexical word features (lexical trigrams) to train model.

## 5.3 Dialog Management with Counseling Strategy

When we extract 5W1H information or user emotions, the dialog management module keeps them in the emotion slot or in the six 5W1H slots. This slot information is discussed in a dialog.

The dialog management module decides the system’s action by the main action, the 5W1H entities, and the user’s emotions. Dialog management follows the rules in figure 3, which is our dialog strategy for the counseling system. In figure 3, ‘Counseling group?’ node finds users intentions included in ‘others group’ (rejection or thanks could be included). The ‘User Emotion Detection’ node figures out whether the user utterance is to include emotional keywords or whether the user emotion is already known by the discourse. The ‘6 slot empty’ node checks whether the user utterance includes at least one of the 5W1H elements or whether the 5W1H entity is already known. The ‘6 slot full’ node decides whether the user utterance with a discourse has all six 5W1H entries. From this strategy, we can notice that we cannot reflect a user’s feeling without the user’s emotion. We cannot ask open questions when all of the 5W1H slots are filled.

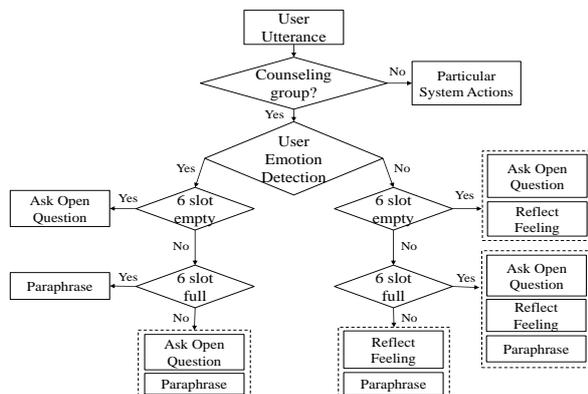


Figure 3. Dialog Strategy Architecture

## 5.4 Emotion Detection

The emotion detection module decides the user’s emotion with respect to the four basic emotions. To detect the user’s emotions, we find emotional keywords in the user’s utterances. If any emotional keyword appears in a user utterance, we decide that the user’s emotion which includes that keyword. For this approach, we made a dictionary of emotional keywords for each basic emotion. The dictionary has approximately 15 emotional keywords for each basic emotion (Table 3).

Emotion	Keywords
Happy	happy, joy, glad, pleasure, ...
Sad	sad, depress, grief, sorrow, ...
Angry	anger, rage, indignation, wrath, ...
Afraid	afraid, fear, terror, dread, ...

Table 3. Emotional Keywords for Emotions

## 5.5 Natural Language Generation

The NLG module generates the system’s utterances by selecting an utterance template from the template database. There are various system templates for each system action, for the 5W1H slot entities, and for the user’s emotions (Figure 4). When the system template is chosen, we can generate system utterances by replacing the 5W1H slot with a 5W1H discourse. ‘O’ means we should have the entity, ‘X’ means we do not have the entity, and ‘\*’ means we do not care whether we have the entity or not.

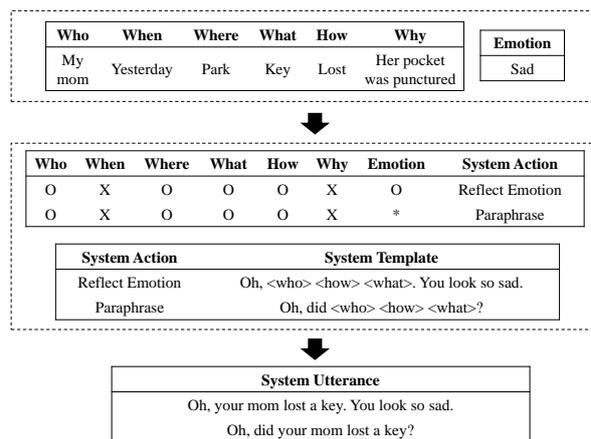


Figure 4. System Utterance Generation

Our overall counseling dialog example is shown in (Table 4).

<b>User</b>	My mom lost a key yesterday.
<b>System</b>	Oh, where did you mom lost a key?
<b>User</b>	At the park.
<b>System</b>	Your mom lost a key at the park, right?
<b>User</b>	Yes, that's right. I'm so sad.
<b>System</b>	You look so sad. Cheer up.
<b>User</b>	Thank you.
<b>System</b>	My pleasure.

Table 4. Counseling Dialog Example

## 6 Experiment

We measured how systems show empathy with users. Our baseline system is a Korean chat-oriented dialog system (Kim et al., 2012). The chat-oriented dialog system shows empathy by understanding user utterances and making a conversation. In our experiment, 7 basic situations are given for each person. Situations are explained by 5W1H, and users generated various utterances using that information. Each person generated approximately 100 utterances during 30 minutes and made estimates for each system. We recruited 16 volunteers to use our system and to estimate its effectiveness. Each user checked 17 questions from 1 to 10. The questions ask users how does each system understand the user utterance, is it appropriate for counseling, and does it satisfy the users (Table 5).

Question	Chat-Oriented	Counseling
1-1. The system used counseling techniques: paraphrasing, open question, reflect feeling.	3.50	7.06
1-2. The system knows my emotion.	3.44	6.88
1-3. There was no break in the conversation.	2.63	6.88
1-4. The system acts like a counselor.	2.88	6.69
1-5. The system shows empathy with me.	4.69	7.31
1-6. I feel the system understands me.	2.56	6.50
2-1. The system understands what I said.	2.88	6.81
2-2. The system understands 5W1H information.	4.13	7.44
2-3. System utterances are appropriate.	2.75	6.94
2-4. System utterances have no problem.	3.50	5.50
3-1. I could speak about various situations.	4.31	6.38
3-2. I had a casual conversation.	4.75	6.88
3-3. Scenarios look expandable.	5.50	7.63
4-1. I satisfied overall conversation.	3.10	6.56
4-2. I satisfied overall counseling.	2.38	6.56
4-3. The system looks appropriate as a counselor.	2.50	6.38
4-4. I'll recommend the system as a counselor to my friends.	2.31	5.38
Mean	3.40	6.69
Standard Deviation	0.96	0.59

Table 5. Experiment Results

Questions 1-1 to 1-6 ask users how each system is appropriate as a counselor. Counseling system rated 6.89 for mean. Questions 2-1 to 2-4

are about users' utterances understandability. In these questions, counseling system rated 6.67 on the average. Questions 3-1 to 3-3 show how various dialogs covered. Our system got 6.96 for mean. Finally, questions 4-1 to 4-4 are about overall satisfaction. These questions rated 6.22 for mean. Our p-value through t-test was  $3.77 \times 10^{-11}$ .

Counseling system got higher score than chat-oriented system because users felt empathy better with our system than baseline system. As a counselor, counseling system is much better than chat-oriented system. Our baseline system was not appropriate as a counselor because it rated 3.39 for average. However, our system scored over 6.5 overall. It means our system is valuable as a counselor.

## 7 Conclusion

In this study, we introduced counseling techniques that we used to implement counseling dialog system. The experimental results showed that our system shows empathy with users. Although the results of this study bring us a step closer to implementing counseling dialog system, the results are only valid with 5W1H information in Korean. Our future works are to improve our counseling dialog system using new NLU module which extracts 5W1H information from more general utterances, with new emotion detection method, and with more counseling techniques.

## Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(2012-0008835).

This research was supported by the MSIP(Ministry of Science, ICT&Future Planning), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency) (NIPA-2013-H0301-13-3002)

## References

- Kim, Y., Noh, H., & Lee, G. G. (2012). Dialog management on chatting system based on lexico-syntactic patterns and named entity types. *Proceedings of Spring Conference of Korean Society of Speech Sciences*, 41-42, Seoul, Korea.

- Lafferty, J., McCallum, A., & Pereira, F. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. *Proceedings of the 18th International Conference on Machine Learning*, 282-289.
- Ratnaparkhi, A. (1998). Maximum entropy models for natural language ambiguity resolution. *Computer and Information Science*, University of Pennsylvania, Philadelphia, USA.
- Schröder, M., Cowie, R., Heylen, D., Pantic, M., Pelachaud, C., & Shuller, B. (2008). Towards responsive sensitive artificial listeners. *Workshop on Human-Computer Conversation*, Bellagio, Italy.
- Theron, M. J. (2008). A manual for basic relational skills training in psychotherapy. *Masters of Arts in Clinical Psychology*, University of South Africa, South Africa.