

Extracting Knowledge from On-Line Forums for Non-Obstructive Psychological Counseling Q&A System

Yuanchao Liu¹, Ming Liu¹, Zhimao Lu², Mingkai Song¹

¹School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China

²Information and Communication Engineering College, Harbin Engineering University, Harbin, China

Email: ycliu@hit.edu.cn

Received February 29, 2012; revised March 20, 2012; accepted March 29, 2012

ABSTRACT

Psychological counseling Q&A system is enjoying a remarkable and increasing popularity in recent years. Knowledge base is the important component for such kind of systems, but it is difficult and time-consuming to construct the knowledge base manually. Fortunately, there emerges large number of Q&A pairs in many psychological counseling websites, which can provide good source enriching the knowledge base. This paper presents the method of knowledge extraction from psychological consulting Q&A pairs of on-line psychological counseling websites, which include keywords, semantic extension and word sequence. P-XML, which is the knowledge template based on XML, is also designed to store the knowledge. The extracted knowledge has been successfully used in our non-obstructive psychological counseling system, called P.A.L., and the experimental results also demonstrated the feasibility and effectiveness of our approach.

Keywords: Knowledge Base Construction; Keyword Extraction; Q&A System; Chatterbot

1. Introduction

It is well known that people are usually troubled by some psychological problems, which may be about family, human relationship, marriage, life pressure, learning, working and etc. [1-3]. Clients may choose consorting to human counselors to gain solutions for these problems. Psychological counseling has been defined as a working relationship between two individuals in which the counselors challenges the client to deal with problems by finding solutions based on the client's own value system [4, 5]. A significant amount of research has been conducted over the last two decades into the theory, method and the practices to advise the human counselors [6-9]. Besides, it has been well accepted that psychological counseling is different from psychological therapy [10-12].

Although providing much support for clients, the counseling conducted by human experts still has some shortcoming: 1) clients worry to divulge their privacy. Most clients may be reluctant to see human counselors as they are unwilling to tell something private to others; 2) Counseling cost are usually high, thus clients cannot obtain the help of human counselor very frequently. And they will go to see human counselor only when the problem has accumulated or upgraded to severe extent, thus the human counselors' help may be limited; 3) The level of human counselor. Human counselors cannot own all know-

ledge; some of them may have less knowledge. In addition, as one people, human counselor may exhibit some own feeling or sentiment during counseling process. Sometimes they may be impatient or weary. These will have adverse impact on clients.

Non-obstructive psychological counseling emphasizes the everyday-life FAQ consulting, and it is a social assistant measures which can provide early-preventive intervention. The clients are basically healthy, thus have no obvious psychological conflicts and can basically adapt to the environment. The purpose of consultation is to help clients develop their potential ability, improve the quality of learning and life, advice them on how to reduce the psychological stress, enhance adaptive capacity and so on.

In fact, non-obstructive psychological counseling is widely needed by most people, e.g., when people have trouble in life, they may talk to relatives or friends for some language comfort. In most situations it is not necessary to consult human experts for these troubles. Similar to other robots which can assist people in performing dangerous or repetitive work, Computer Q&A system can accumulate lots of knowledge, make rapid retrieval, provide real-time service, and can even perform some preliminary work for human experts. Although robots can not completely replace the work of human counselors, it can at least alleviate the anxiety, press and bad

mood of clients to some extent. It can help remind or make clients change their mind from negative status to positive active status.

The knowledge base is vital to the performance of Q&A system [13]. Currently, there are many online counseling forums [14,15] and a wealth of Q&A pairs has accumulated over time. We find that, the answers for many similar problems are often also very similar. This inspired the design of robot counselors which is based on the knowledge base extracted from corpus.

The remaining content of this paper is organized as follows. Section 2 presents the related work. Section 3 describes the basic principle of Q&A counseling system. P-XML: the knowledge base template of our psychological counseling system was presented in Section 4. For the next 3 sections, the methods to form the knowledge base are presented separately. Section 5 describes keyword extraction method, Section 6 describes semantic extension method, and Section 7 describes the extraction of sequence pairs. Section 8 presents the corresponding experimental results for above work. We draw conclusions and discuss some potential extensions of our framework in Section 9.

2. Related Work

Many researchers has dedicated to the application of computer and artificial intelligence in psychiatry [16-22]. Two clinical psychologists, Jaap Hollander and Jeffrey Wijnberg, developed the first on-line computer robot program psychologist Mind-Mentor [23], which maintains a conversation with a user by using a set of static and predefined rules. Around 1600 people tried the service, and 47% of them are satisfied with the counseling. Besides, a software program called Virtual Space Station was also developed to provide assistance to astronauts on future space missions. This interactive software suite is intended to address the psychosocial challenges of long-term space travel. Lihua Fan developed a psychological consultation expert system which is based on man-machine conversation [24]. The system designed 145 rules to diagnose anxiety. Aiming at the practical psychological counseling demands of college students, Mei-ling Zhang designed a psychological counseling expert system realized in java environment [25]. The system is composed of intelligent psychological counseling module, psychological test module, information managing module, assistant decision module and facial expression recognition module, which can gain students' information and perform counseling. Besides, there are also some computer-aided expert system researches for psychological counseling [26,27].

The early research in automated QA or dialogue system is in the 1960s, and for the recent years, much work has been done in this area [28,29]. Many open-domain

Q&A systems answer questions by first harvesting many candidate answers, and returned the most possible one [30]. Research activities involving QA has been greatly motivated by the introduction of the QA Track in TREC (Text Retrieval Conference) evaluations [31,32]. Leila Kosseim etc. present a method for acquiring semantic-based records automatically for question answering work [33]. In the context of question answering on free text, Tiphaine Dalmas assessed the value of answer comparison and information fusion [34]. Rafael M. Terol presented a restricted domain question answering (QA) system [35]. The knowledge is acquired through the use of Unified Medical Language System (UMLS) and WordNet. Clinical questions are often complex and with many forms. Yonggang Cao built a clinical Q&A system named AskHERMES to perform semantic analysis on clinical questions and can give question-focused answers [36].

Two famous chatterbots are Eliza [37] and ALICE [38]. The XML dialect called AIML was developed and formed the basis for "ALICE" (Artificial Linguistic Internet Computer Entity), which has won the annual Loebner Prize Contest for Most Human Computer three times. The data and knowledge of Eliza is predefined and formatted, Whereas Mauldin try to get knowledge from outside and utilize the record matching, which is more dynamic and thus is different from Eliza. Mauldin has created two chatterbots: Sylvie and Julia [39].

In the past years, many Psychological Counseling websites has emerged, and there are some human experts to provide the answer to the questions of clients on-line or off-line. In essence, it is still one to one service. Although more convenient service has been provided, there are still some shortcomings for on-line psychological counseling: 1) users must access the service by internet or mobile service, thus extra fee will be needed. When network error occurs, the service will be affected. 2) The service is still provided by human psychologists, thus depend much on the existence of human expert. Clients may still worry their privacy known by other people, and can not be relaxed fully.

Most often, the question of clients are very similar, and experts may give same or very similar answer again. In this paper, we first collected enormous counseling Q&A pairs, and extract some knowledge to teach robots. Basically, advances in automatic language analysis and process such as language understanding, generation, and speech synthesis have paved the way for the emergence of complex, task-oriented psychology consulting. There is some difference from traditional dialog system, as the purpose of automatic psychology Q&A consulting is to help solve problem of clients and not just talk freely. When someone has trouble, what he need is psychology consulting, not just chattering. So the conversation is more like that with a psychologist.

3. The Basic Principle of Our Q&A Counseling System

We developed a dialog system called P.A.L. (Psychologist of Artificial Language). It can provide answer for clients' problems and also have some basic chatterbot-like characteristics. The question answering process is based on entry searching and solution amendment. The basic principle is shown in **Figure 1**. For the problem inputted by user, P.A.L. first make word segmentation, POS tagging, stop words removal and semantic analysis, then it check the matched entry by searching in the knowledge base. Once matched, the corresponding solution will usually be slightly amended to suit for the user's problem, and return the response to the user. The search is based on both index and some kinds of keywords. The index can be used for narrowing possible range, and select a candidate set, then the system compare the input with each candidate records in knowledge base one by one.

Two important part of knowledge base are the record index and the XML documents. The former is used for searching records in XML documents quickly. The index is generated by Clucene, whereas the XML documents store the Q&A pairs and some kinds of keywords information for each record. Although the Lucence index is very efficient means for retrieve the matched record, it still has some drawbacks as there maybe some slight difference between users' input and the matched questions. So we also generate some keywords to capture some trivial features of each record. Thus when users give their questions as input, the system first search the matched records according to index, then for the matched records, the system check the similarity. There are 3 matching method: keyword matching, extension keyword matching, and sequence matching.

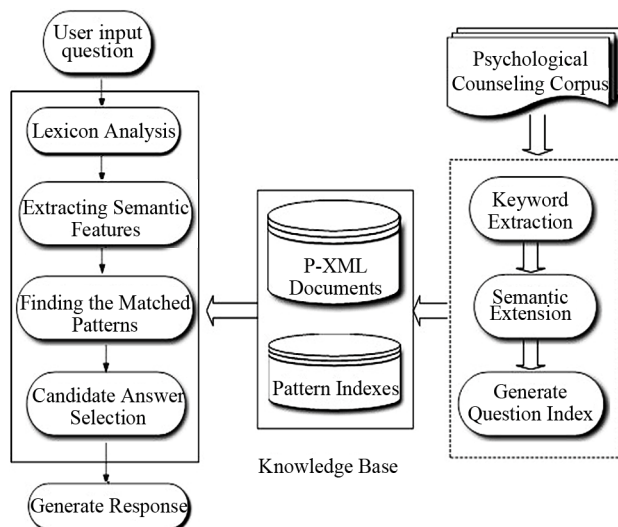


Figure 1. The basic principle of P.A.L. counseling system.

Generally, there are two big steps to find the matched records; one is based on index, another is based on keyword. The system will check whether the key words intersection between the users' input and the entry in the collection are above a certain threshold. If there are some name entities in users' input, the name entity similarity must be big enough, or the system will turn to the next matched record by index. If there are no suitable records, the system will give a "not found" message. The system checks whether the intersection between the users' input and the corresponding entry in the collection are above a certain threshold, and find the possible entry to give the solution. Suppose the number of intersection elements are N' , whereas the number of the elements in the current entry is N , then the percentage can be calculated as N'/N . In general, if the users' questions are very concrete (e.g., they are bigger than 100 words), then the word sequence method are used to find the best solution; otherwise, the extended keywords may be adopted.

The purpose of keyword extension is that, for same question, users may give different input, so we use CILIN to make an extension for some typical keywords, e.g. name entity, or problem word. It should be noted that semantic extension are done on keyword set, thus the searching may be more intelligent.

4. P-XML: The Template for the Knowledge Base of P.A.L.

One important thing for P.A.L. is to extract and preserve knowledge in order to facilitate search and matching. It will be an ideal situation if the visitor's problem can be exactly matched with one record in the Knowledge Base; however, this is very rare in practice. For the same problems there may be many different expression ways for visitors. Thus the knowledge base of psychological counseling system should capture key information from many cases, so as to accurately provide support for finding the most similar answer.

The main scheme for the knowledge base of P-XML is based on keyword extracted from the Q&A pairs. Keywords are the important source to determine the key factors in the visitors' problems of the knowledge base. If one visitor's problem has great agreement with these proposed keywords of one entry in P-XML, which means their problems can be broadly similar. In theory, if the system can retrieve a very similar situation with the current visitor and return the corresponding answer to the user, then their requirements can be similar to a certain extent.

Although P-XML is inspired by ALICE, which uses XML to store knowledge base, there is much apparent difference between them. **Figure 2** is the format of P-XML. The records of the knowledge base in P.A.L. are extracted mostly from the Q&A pairs in BAIDU com-

munity. By analyzing a large number of psychological counseling Q&A examples, we found that the concerns of visitors can be divided into two categories: 1) relatively common problems, such as how to deal with the relation between work and family; 2) the problem about visitor's own actual situation or experiences. The "intersection_number" attribute in the answer node **Figure 2** means if the answer is general. if one answer has few intersection words with the question, that usually means the answer is general, when user input a question which has big similarity with the record, the general answer may be more "secure" than a concrete answer as there are usually no personal information. For example, the answer "no problems" is a general answer. Whereas "you can do what your wife told you." is not general, as the word "wife" is personal information which means the user is male, and has married. But the current user of P.A.L. may be female. So we also try to extract the speak role of each question in P-XML, which is the attribute of node "qapair" node in **Figure 2**. The "intersection_question_answer_keyword" node means the words which have shown in both the question and answers. Such words are usually the key problem of users, cared by both the questioners and answer users.

5. Extract Keywords from Psychological Counseling Q&A Pirs

Beside from the index which is generated from the ques-

```
<?xml version="1.0" encoding="GB2312"?>
- <domain name="*">
  <signon>*</signon>
  <nullresponse>*</nullresponse>
  <inputrepeat>*</inputrepeat>
  - <qapair speaker="*" Number="*">
    <question_id>*</question_id>
    <question_category>*</question_category>
    <question_username>*</question_username>
    <propose_time>*</propose_time>
    <solve_time>*</solve_time>
    <question_title>*</question_title>
    <question_content>*</question_content>
    <question_link>*</question_link>
    <question_relatedquestion>*</question_relatedquestion>
  - <question_title_sentences>
    - <sentence text="*">
      <entity_problem_word>*</entity_problem_word>
    </sentence>
  </question_title_sentences>
  - <question_title_and_content_highfre_keyword>
    <keyword>*</keyword>
  </question_title_and_content_highfre_keyword>
  <best_answer intersection_number="*">*</best_answer>
  <other_answer intersection_number="*">*</other_answer>
  - <intersection_question_answer_keyword>
    <keyword>*</keyword>
  </intersection_question_answer_keyword>
  - <answer_highfre_keyword>
    <keyword>*</keyword>
  </answer_highfre_keyword>
  - <domain_keywords>
    <keyword>*</keyword>
  </domain_keywords>
  <title_extension>*</title_extension>
</qapair>
</domain>
```

Figure 2. The format of XML templates in P.A.L. counseling system.

tions by Clucene, Keywords are the main knowledge form of P.A.L. system. Basically, when one user enters a problem and candidate records has be found, counseling system can compare the description with the keywords to determine which records are most consistent with that problem. The corresponding solution will be returned as the output to the user after language generation. Keywords must represent the main skeleton content of the psychological counseling Q&A pairs. Some type of keywords is listed here:

1) Name entity. The name entities mentioned in questions are important features of users' question, e.g. "boyfriend", "girlfriend", "wife" and etc. P.A.L. maintains a name entity list for recognizing such kind of keywords. If one answer in the current record contains name entity, whereas there are no same name entities in the current user's input, the answer can not be selected as candidate response.

2) Problem word. Some other specific types of keywords for problem type, e.g. "regret", "angry", "breaking up" and etc. P.A.L. maintains a problem word list for recognizing such kind of keywords.

3) Terms shown in both the question and answer. Another important word is the words which appear both in the questions and the answers. Such words are the focus of the users and have attracted the mention of answer providers. They are denoted as "intersection_question_answer_keyword" node in P-XML.

4) All the words after stop word removal shown in the question title can be keywords. For all Q&A pairs, there are titles given by visitors, which can abstract the questions well. Most of pairs have question contents, which help explain the problem in more detail. For example, one title is "boyfriend often paralyzed himself by playing game". Although there are many terms in the question content, the title will be more salient.

5) Domain-keyword: The top high-frequency terms in the psychological counseling corpus for one topic are an important source of keyword extraction. For example, in the field of marriage, frequent terms include "love", "marriage" and etc. For each field, the domain frequent keywords are generated by finding most frequent terms, and a domain keyword lexicon can be formed. For each Q&A pair, the domain words which have been shown in question (both title and content) are put into the node of domain-keyword in **Figure 2**.

6. Semantic Extension of Keywords

By extracting keywords according to the strategies in Section 5, psychological counseling system can identify the main information in visitors' problem and has created favorable conditions for case-based matching and retrieval. But the user's language is diverse and complex, and sometimes the users' expression is in different ways although

the meaning is the same. So another approach is to utilize semantic knowledge to make extension for the extracted keywords. In this paper, we mainly use K4, which is all the keywords in the question, to make semantic extension, and generate the new index.

TONGYICILIN [40] are used as the knowledge source to find words with same meaning and enrich the knowledge base. In CILIN, there are many word groups; the words in each group share the meaning although they are different in expression form. There are 3925 group titles in total. For example, words in group Ae07A01 are all have meaning of “farmer”. In this paper, the mapping from words to small class is utilized for keyword extension.

Keyword Extraction Expansion Algorithm

Input: keyword set S

Output: the semantic extended keyword set S'

BEGIN

- 1) loading the CILIN and all keywords;
- 2) For each keyword W ,
 - a) search W in CILIN,
 - b) If found, then get the list which has same meaning with W ;
 - c) push all terms in the list found into the semantic extended keyword set S' ;
- 3) End For

END

7. Extracting Word Sequence

Apart from the keywords extracted, another factor to be considered is the information conveyed by the combination of some words. For example, one question is about “how to choose between love and work”. Although some important keywords like “love” or “career” can be identified and form the keyword set, the semantic relation between these terms may be hidden. When there are lots of sentences in questions, the actual performance may be affected as their relations are not distinguished in keyword set. Based on this consideration, we try to find

some skeleton words in question sentence and extracted them as word pairs to represent the sentence. By contrast, although the keywords extracted in the former section can represent some key information in the whole question, they are independent with each other. Word sequences may grasp the relation of words to some extent.

So in this paper, we also focused on identifying keyword sequences which have direct dependent relationship with each other. For example, in the sentence “因为他自信心不足” (because he is lack of confidence) in **Figure 1**, the pairs are {“自信心” (confidence), “他” (he)} and {“自信心” (confidence), “不足” (inadequate)} (**Figure 3**). The element in each pairs must be keywords extracted in Section 4. We use the LTP platform for extracting the direct dependency relation between keywords.

Word Sequence Extraction Algorithm

Input: keyword set S

BEGIN

Output: the set of S' , each element is a pair with form <keyword 1, keyword 2>

- 1) loading the question and keyword set S ;
- 2) parse the question, get all the dependency pairs;
- 3) For each pair P ,
- 4) if both the words in P are in S , then put pair P into S' ,
- 5) End For

END

8. Experimental Result and Analysis

8.1. Experimental Setting and Evaluation Method

The corpus for P.A.L. dialog system knowledge base construction is mainly from BAIDU ZHIAO, which is one of the biggest Chinese communities for all kinds of questions and the corresponding answers. Seven different fields are covered in the knowledge base, as shown in **Table 1**. They are all about users' trouble and many solutions.

Table 1. Some high-frequency words for each domains.

	D 1	In English	D 2	In English	D 3	In English	D 4	In English	D 5	In English	D 6	In English
1	他	He	他	He	她	She	长高	High	他	He	治疗	Treatment
2	她	She	她	She	他	He	女生	Girls	她	She	抑郁	Depression
3	结婚	Marry	妈妈	Mom	有	Have	岁	Years	有	Have	症状	Symptom
4	老公	Husband	有	Have	想	Want	男生	Boys	什么	What	患者	Patient
5	离婚	Divorce	爸爸	Father	你	You	他	He	你	You	症	Disease
6	想	Want	什么	What	好	Good	怎么	How	怎么	How	怎样	How
7	孩子	Children	妈	Mom	什么	What	现在	Now	感觉	Feeling	发病	Pathogenesis
8	年	Years	父母	Parents	现在	Now	好	Good	为什么	Why	心理	Psychological
9	你	You	现在	Now	爱	Love	她	She	能	Can	病情	Condition
10	老婆	Wife	他们	They	朋友	Friend	人	People	觉得	Feel	描述	Description
11	好	Good	怎么	How	能	Can	怎样	How	不知	Unknown	帮助	Help
12	怎么	How	爸	Dad	时候	Time	多	Many	时候	Time	医院	Hospital
13	爱	Love	孩子	Children	年	Years	发育	Development	朋友	Friend	怎么	How
14	家	Home	家	Home	觉得	Feel	身高	Height	心情	Mood	她	She
15	生活	Life	钱	Money	女朋友	Girlfriend	青期	Puberty	喜欢	Like	男	Male

The knowledge base construction results and some related information are shown in Experiment 1. And the performance of obtained knowledge base in P.A.L. system is evaluated from the perspective of both breadth and sensitivity, as shown in Experiment 2. Breadth evaluation is to examine if the system can answer as many questions as possible. Sensitivity evaluation is to examine if the system can “understand” users’ question. Sometimes, although there are records which can meet users’ input in knowledge base, the expression may be different.

For breadth evaluation, we use cross validation [41, 42], which can examine how well knowledge base construction model perform in practice. The raw sample data are partitioned into complementary subsets, performing the knowledge base extraction on one subset, *i.e.*, the training set, and test the performance on the other subset, *i.e.*, the testing set. The average of 10 rounds of cross-validation results is performed using different partitions to reduce variability. As for sensitivity evaluation, another evaluation method was used, *i.e.*, some existing records in knowledge base are randomly selected, and the questions were made a slight change, and then scores are assigned by examining the output.

8.2. Experiment 1. Some Statistical Information about the Knowledge Base

From Figure 4(a) to (f), the relationship between the number of keywords and frequency threshold on different knowledge base size for different domains are examined. In Figure 4, different series means the different knowledge base size, *i.e.*, Series n means $500 \times n$ records. E.g., series 7 means 3500 records are counted and series 10 means 5000 records are counted. X-axis means the frequency threshold. E.g. “x = 3” means the words which occur more than 3 times are counted. Y-axis means the number of keywords which meet these conditions.

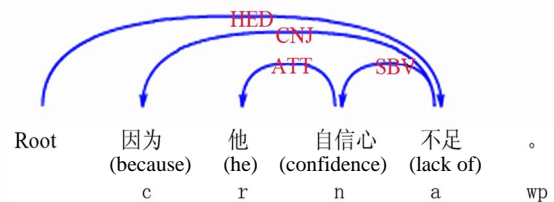


Figure 3. The dependency relation of one example sentence.

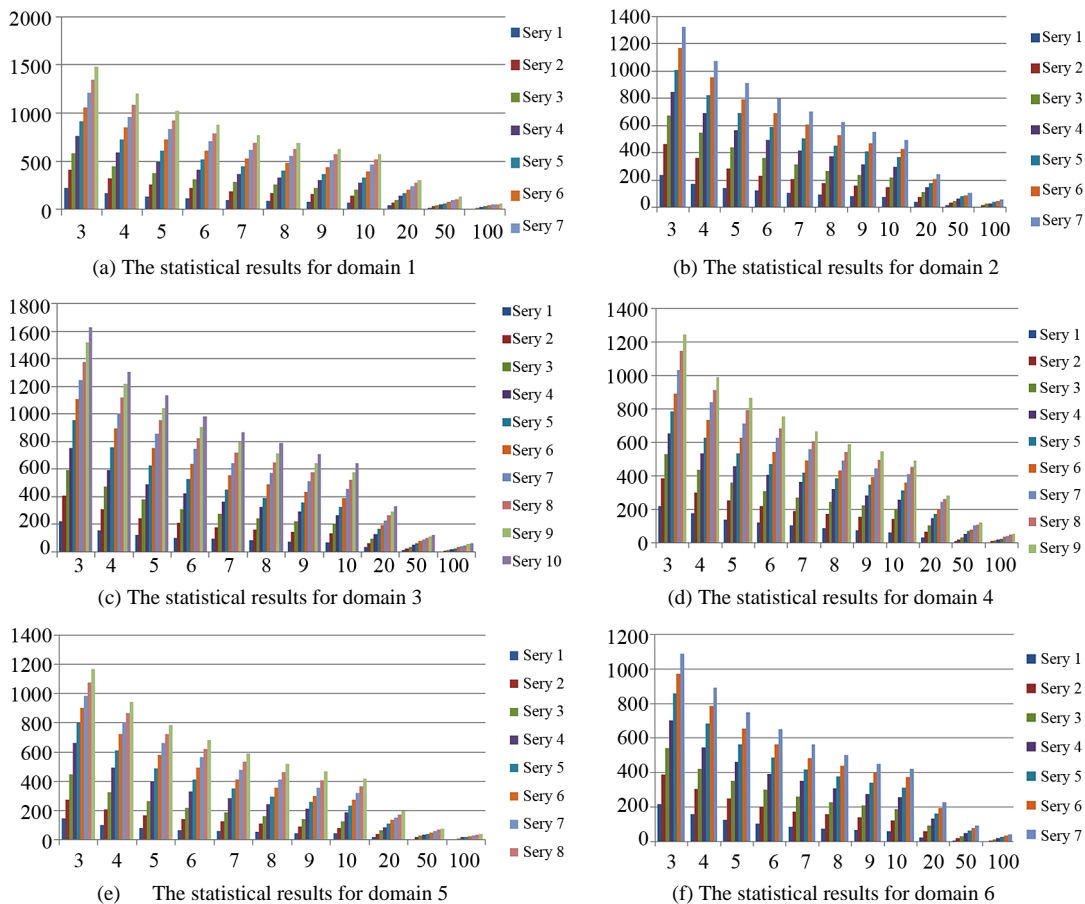


Figure 4. The relationship between the number of keywords and knowledge base size for different domains (X-axis: frequency threshold; Y-axis: the number of keywords. Different series means the different knowledge base size, *i.e.*, Series n means $500 \times n$ records).

It can be seen that, when frequency threshold is bigger, the number of extracted keywords increase not so sharply, as many records share same keywords. It may indicate that the user's problem distributions are relatively stable. This is closely related with the problem of psychological counseling, as users' concerns are relatively fixed in some kinds of problems. It can also be found that significant numbers of visitors have some repetitive questions and the answers also have greater similarity. Correspondingly, Table 1 gives some high-frequency words for each domain.

8.3. Experiment 2. The Evaluation of Knowledge Base Performance

The effective work of dialog system depends much on the breadth of the knowledge base. In order to evaluate the breadth, M questions are selected as the input ($M = 100$ in this test) and sent to the system, and then examine if there are closely related questions (other than the input itself) automatically, as shown in Section 8.1. In Figure 5, we set a similarity threshold, and examine the number of matched records which are more similar (above the threshold) with the input. The results of Figure 5 demonstrated that when user input one problem, a similar problem (e.g. the similarity is above 0.5 or 0.8) in the knowledge base may be found with about 5% - 25% possibility.

The system sensitivity is evaluated and shown in Figure 6. N existing questions ($N = 20$ in this test) in knowledge base are selected. Their paraphrases are input to the system again to get the response. For one response, if it can be found in the original answers of the input, and at the same time the personal information shown in the response appears in the input either, the answer will be deemed as correct, otherwise it will be incorrect. In Figure 6, method A is the traditional IR technology which only use index to find response. Method B is the method used in P.A.L., which use both IR and personal information extracted to find suitable response. Generally, Figure 5 is to examine the question breadth, whereas Figure 6 is to examine the answer sensitivity. It can be seen the accuracy of method B is apparently higher than method A. which verified the importance of processing personal information in P.A.L.

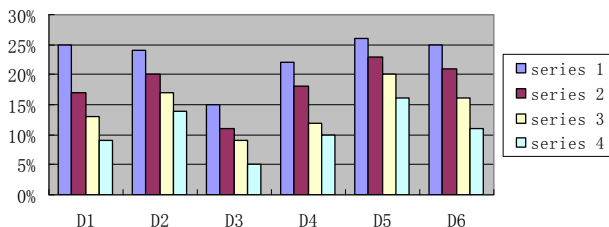


Figure 5. The breadth evaluation results for different domains (the similarity threshold for series 1 to 4 is 0.5, 0.6, 0.7 and 0.8 separately; X-axis is the domain ID, Y-axis is the breadth evaluation value).

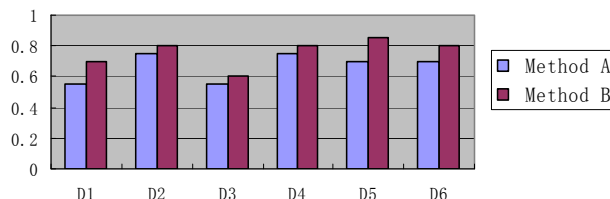


Figure 6. The system sensitivity evaluation results for different domains (method A: IR only. Method B: IR plus considering personal information. X-axis is the domain ID; Y-axis is the sensitivity evaluation value).

9. Conclusions and Future Work

We propose an approach of extracting knowledge from on-line forums, which is essential for non-obstructive psychological counseling Q&A system. In fact, such corpus and knowledge can also help human experts by providing rich language and cases. Some strategies are designed for finding the important keywords. Besides, it was found that many problems have same or very similar answer, thus we can mine the common words or records in these questions, and add new strategies for extracting knowledge.

The drawback of our knowledge base is that by use of XML and more knowledge used, storage space requirement also increase. On mobile phone or other PDA platforms, this phenomenon may be much apparent. To solve this problem, the system may accommodate more Q&A pairs by reducing some semantic knowledge. In this way, the system performance may not be adversely affected as it can answer more questions, *i.e.*, the breadth increase can complement the decrease of sensitivity.

10. Acknowledgements

This research was supported by Youth Funds of China social & humanity science (10YJCZH099), the Fundamental Research Funds for the Central Universities (Grant No.HIT.NSRIF.2009065) and Key Laboratory Opening Funding of China MOE—MS Key Laboratory of Natural Language Processing and Speech (HIT.KLOF.2009022), the project of The National High Technology Research and Development Program (863 program) of PR China under a research Grant No.2007AA01Z172.

REFERENCES

- [1] S. Guney, F. Akca and G. Sahin, "The Interrelation between Traumatic Life Events and Mental Health in Turkish University Students," *Procedia-Social and Behavioral Sciences*, Vol. 12, 2011, pp. 122-125.
- [2] E. J. R. David, S. Okazaki and A. Saw, "Bicultural Self-Efficacy among College Students: Initial Scale Development and Mental Health Correlates," *Journal of Counseling Psychology*, Vol. 56, No. 2, 2009, pp. 211-226. [doi:10.1037/a0015419](https://doi.org/10.1037/a0015419)

- [3] L. Brosan, L. Hoppitt, L. Shelfer, A. Sillence and B. Mackintosh, "Cognitive Bias Modification for Attention and Interpretation Reduces Trait and State Anxiety in Anxious Patients Referred to an Out-Patient Service: Results from a Pilot Study," *Journal of Behavior Therapy and Experimental Psychiatry*, Vol. 42, No. 3, 2011, pp. 258-264. [doi:10.1016/j.jbtep.2010.12.006](https://doi.org/10.1016/j.jbtep.2010.12.006)
- [4] G. Corey, "Theory and Practice of Counseling and Psychotherapy," 5th Edition, Brooks Cole, Pacific Grove, 1996.
- [5] S. Martin, A. Akers and A. W. Jackson, "Male and Female Athletes' and Nonathletes' Expectations about Sport Psychology Consulting," *Journal of applied sport psychology*, Vol. 13, No. 1, 2001, pp. 18-39. [doi:10.1080/104132001753155930](https://doi.org/10.1080/104132001753155930)
- [6] J. Owen, M. M. Leach, B. Wampold and E. Rodolfa, "Client and Therapist Variability in Clients' Perceptions of Their Therapists' Multicultural Competencies," *Journal of Counseling Psychology*, Vol. 58, No. 1, 2011, pp. 1-9.
- [7] W. M. A. Ludwikowski, D. Vogel and P. I. Armstrong, "Attitudes toward Career Counseling: The Role of Public and Self-Stigma," *Journal of Counseling Psychology*, Vol. 56, No. 3, 2009, pp. 408-416. [doi:10.1037/a0016180](https://doi.org/10.1037/a0016180)
- [8] D. L. Vogel, N. G. Wade and A. H. Hackler, "Perceived Public Stigma and the Willingness to Seek Counseling: The Mediating Roles of Self-Stigma and Attitudes toward Counseling," *Journal of Counseling Psychology*, Vol. 54, No. 1, 2007, pp. 40-50. [doi:10.1037/0022-0167.54.1.40](https://doi.org/10.1037/0022-0167.54.1.40)
- [9] M. Schoenberg and S. Shiloh, "Hospitalized Patients' Views on In-Ward Psychological Counseling," *Patient Education and Counseling*, Vol. 48, No. 2, 2002, pp. 123-129.
- [10] G. R. Jiang, "The Theory and Practice of Psychological Consulting," High Education Press, Beijing, 2005.
- [11] B. D. Locke, J. S. Buzolitz, P.-W. Lei, J. F. Boswell, A. A. McAleavey, T. D. Sevig, J. D. Dowis and J. A. Hayes, "Development of the Counseling Center Assessment of Psychological Symptoms-62 (CCAPS-62)," *Journal of Counseling Psychology*, Vol. 58, No. 1, 2011, pp. 97-109. [doi:10.1037/a0021282](https://doi.org/10.1037/a0021282)
- [12] J. Owen, M. M. Leach, B. Wampold and E. Rodolfa, "Multicultural Approaches in Psychotherapy: A Rejoinder," *Journal of Counseling Psychology*, Vol. 58, No. 1, 2011, pp. 22-26.
- [13] Q. L. Guo and M. Zhang, "Question Answering Based on Pervasive Agent Ontology and Semantic Web," *Knowledge-Based Systems*, Vol. 22, No. 6, 2009, pp. 443-448. [doi:10.1016/j.knosys.2009.06.003](https://doi.org/10.1016/j.knosys.2009.06.003)
- [14] M. J. Mallen, "The Practical Aspects of Online Counseling Ethics, Training, Technology, and Competency," *The Counseling Psychologist*, Vol. 33, No. 6, 2005, pp. 776-818.
- [15] J. R. Alleman, "Online Counseling: The Internet and Mental Health Treatment," *Psychotherapy Theory Research Practice Training*, Vol. 39, No. 2, 2002, pp. 199-209.
- [16] S. Chattopadhyay, D. K. Pratihari and S. C. de Sarkar, "Developing Fuzzy Classifiers to Predict the Chance of Occurrence of Adult Psychoses," *Knowledge-Based Systems*, Vol. 21, No. 6, 2008, pp. 479-497. [doi:10.1016/j.knosys.2008.03.006](https://doi.org/10.1016/j.knosys.2008.03.006)
- [17] A. P. Oliveira and A. Cardoso, "A Musical System for Emotional Expression. Knowledge-Based Systems," Vol. 23, No. 8, 2010, pp. 901-913. [doi:10.1016/j.knosys.2010.06.006](https://doi.org/10.1016/j.knosys.2010.06.006)
- [18] R. Schmidt and O. Vorobieva, "Case-Based Reasoning Investigation of Therapy Inefficacy," *Knowledge-Based Systems*, Vol. 19, No. 5, 2006, pp. 333-340. [doi:10.1016/j.knosys.2005.11.016](https://doi.org/10.1016/j.knosys.2005.11.016)
- [19] O. Kelly, K. Matheson, A. Martinez, Z. Merali and H. Anisman, "Psychosocial Stress Evoked by a Virtual Audience: Relation to Neuroendocrine Activity," *Cyber Psychology & Behavior*, Vol. 10, No. 5, 2007, pp. 655-662. [doi:10.1089/cpb.2007.9973](https://doi.org/10.1089/cpb.2007.9973)
- [20] L. Gorini and G. Riva, "Virtual Reality in Anxiety Disorders: The Past and the Future," *Expert Review of Neurotherapeutics*, Vol. 8, No. 2, 2008, pp. 215-233. [doi:10.1586/14737175.8.2.215](https://doi.org/10.1586/14737175.8.2.215)
- [21] M. Kotlyar, C. Donahue, P. Thuras, M. G. Kushner, "N. O. Gorman, E. A. Smith and D. E. Adson, "Physiological Response to a Speech Stressor Presented in a Virtual Reality Environment," *Psychophysiology*, Vol. 45, No. 6, 2008, pp. 1034-1037. [doi:10.1111/j.1469-8986.2008.00690.x](https://doi.org/10.1111/j.1469-8986.2008.00690.x)
- [22] K. Kim, C.-H. Kim, S.-Y. Kim, D. Roh and S. I. Kim, "Virtual Reality for Obsessive-Compulsive Disorder: Past and the Future," *Psychiatry Investigation*, Vol. 6, No. 3, 2009, pp. 115-121. [doi:10.4306/pi.2009.6.3.115](https://doi.org/10.4306/pi.2009.6.3.115)
- [23] J. Hollander. <http://www.mindmentor.com>
- [24] L. Fan, "The Research and Realization on the Expert System of Psychological Consultation," Master Thesis, Chengdu University of Technology, Chengdu, 2008.
- [25] M.-L. Zhang, H. Pan, W.-S. Wang and J. Dong, "Study of Expert System for College Students Psychological Counseling Based on Agent," *Computer Engineering and Design*, Vol. 30, No. 7, 2009, pp. 1735-1739.
- [26] H.-H. Shi, F. Yan, D.-M. Li and Y.-F. Hu, "Research on On-Line Psychology Consultation Expert System Based on Man-Machine Interaction Technique," *Computer Engineering and Design*, Vol. 26, No. 12, 2005, pp. 3307-3309.
- [27] D. Jing, "Research of Sentiment Model Based on HMM and Its Application in Psychological Consulting Expert System," Master Thesis, Hunan University, Changsha, 2009.
- [28] M. Nakano and Y. Hasegawa, "A Multi-Expert Model for Dialogue and Behavior Control of Conversational Robots and Agents," *Knowledge-Based Systems*, Vol. 24, No. 2, 2011, pp. 248-256. [doi:10.1016/j.knosys.2010.08.004](https://doi.org/10.1016/j.knosys.2010.08.004)
- [29] M. Y. Sun and J. Y. Chai, "Discourse Processing for Context Question Answering Based on Linguistic Knowledge," *Knowledge-Based Systems*, Vol. 20, No. 6, 2007, pp. 511-526. [doi:10.1016/j.knosys.2007.04.005](https://doi.org/10.1016/j.knosys.2007.04.005)
- [30] S. Schlobach, D. Ahn, M. de Rijke and V. Jijkoun,

- “Data-Driven Type Checking in Open Domain Question Answering,” *Journal of Applied Logic*, Vol. 5, No. 1, 2007, pp. 121-143. [doi:10.1016/j.jal.2005.12.001](https://doi.org/10.1016/j.jal.2005.12.001)
- [31] <http://trec.nist.gov/>
- [32] O. S. Goh, C. C. Fung and A. Depickere, “Domain Knowledge Query Conversation Bots in Instant Messaging (IM),” *Knowledge-Based Systems*, Vol. 21, No. 7, 2008, pp. 681-691. [doi:10.1016/j.knosys.2008.03.055](https://doi.org/10.1016/j.knosys.2008.03.055)
- [33] L. Kosseim and J. Yousefi, “Improving the Performance of Question Answering with Semantically Equivalent Answer Patterns,” *Data & Knowledge Engineering*, Vol. 66, No. 1, 2008, pp. 53-67. [doi:10.1016/j.datak.2007.07.010](https://doi.org/10.1016/j.datak.2007.07.010)
- [34] T. Dalmas and B. Webber, “Answer Comparison in Automated Question Answering,” *Journal of Applied Logic*, Vol. 5, No. 1, 2007, pp. 104-120. [doi:10.1016/j.jal.2005.12.002](https://doi.org/10.1016/j.jal.2005.12.002)
- [35] R. M. Terol, P. Martínez-Barco and M. Palomar, “A Knowledge Based Method for the Medical Question Answering Problem,” *Computers in Biology and Medicine*, Vol. 37, No. 10, 2007, pp. 1511-1521. [doi:10.1016/j.combiomed.2007.01.013](https://doi.org/10.1016/j.combiomed.2007.01.013)
- [36] Y. G. Cao, F. F. Liu, P. Simpson, L. Antieau, A. Bennett, J. J. Cimino, J. Ely and H. Yu, “AskHERMES: An Online Question Answering System For Complex Clinical Questions,” *Journal of Biomedical Informatics*, in Press, 2011.
- [37] ELIZA, “A Friend You Could Never Have Before,” 2010. <http://www-ai.ijs.si/eliza/eliza.html>
- [38] R. P. Schumaker, Y. Liu, M. Ginsburg and H. Chen, “Evaluating Mass Knowledge Acquisition Using the Alice Chatterbot: The AZ-ALICE Dialog System,” *International Journal of Human-Computer Studies*, Vol. 64, No. 11, 2006, pp. 1132-1140. [doi:10.1016/j.ijhcs.2006.06.008](https://doi.org/10.1016/j.ijhcs.2006.06.008)
- [39] K. Zdravkova, “Conceptual Framework for an Intelligent Chatterbot,” *Proceedings of the 22nd International Conference on Information Technology Interfaces*, Pula, 16 June 2000, pp. 189-194.
- [40] W. Che, Z. H. Li and T. Liu, “LTP: A Chinese Language Technology Platform,” 2010. http://delivery.acm.org/10.1145/1950000/1944288/p13-ch e.pdf?ip=27.17.15.18&acc=OPEN&CFID=75994051&CFTKEN=66856637&__acm__=1333611434_e22c99383ce7e350b16908ac53fa9f4a
- [41] S. Geisser, “Predictive Inference: Seymour,” Chapman and Hall, New York, 1993.
- [42] K. Ron, “A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection,” *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, San Mateo, 20-25 August 1995, pp. 1137-1143.