

Ch2R: A Chinese Chatter Robot for Online Shopping Guide

Peijie Huang, Xianmao Lin, Zeqi Lian, De Yang, Xiaoling Tang, Li Huang, Qiang Huang, Xiupeng Wu, Guisheng Wu and Xinrui Zhang

College of Informatics, South China Agricultural University,
Guangzhou 510642, Guangdong, China

pjhuang@scau.edu.cn

Abstract

In this paper we present a conversational dialogue system, Ch2R (**C**hinese **C**hatter **R**obot) for online shopping guide, which allows users to inquire about information of mobile phone in Chinese. The purpose of this paper is to describe our development effort in terms of the underlying human language technologies (HLTs) as well as other system issues. We focus on a mixed-initiative conversation mechanism for interactive shopping guide combining initiative guiding and question understanding. We also present some evaluation on the system in mobile phone shopping guide domain. Evaluation results demonstrate the efficiency of our approach.

1 Introduction

Spoken dialogue systems are presently available for many purposes, such as, Airline Travel Information System (ATIS) project in the early 1990s (Price, 1990), customer service (Gorin et al., 1997), weather inquiry system (Zue et al., 2000), campus navigation system (Zhang et al., 2004), bus schedules and route guidance (Raux et al., 2003), stock information inquiry (Huang et al., 2004), restaurant recommendation system (Liu, et al., 2008), drug review system (Liu and Seneff, 2011), and spoken route instruction (Pappu and Rudnicky, 2012). These systems have been well developed for laboratory research, and some have become commercially viable.

The next generation of intelligent dialogue systems is expected to go beyond factoid question answering and straightforward task fulfillment, by providing active assistance and subjec-

tive recommendations, thus behaving more like human agents (Liu et al., 2010). For example, in the scenario that we envision, on online e-commerce site, an intelligent dialogue system which roles play a conversational shopping guide may suggest which digital camera is a better choice, considering brand, price, pixel, etc.; or suggest which mobile phone is the most popular among teenagers or highest rated by users.

In this paper, we describe our development effort on a Chinese chatter robot, named Ch2R (**C**hinese **C**hatter **R**obot) for shopping guide with both intelligent ability and professional knowledge. The challenges of developing such a information guiding dialogue system in Chinese includes: 1) how to provide active assistance and subjective recommendations; 2) how to deal with the diversity and flexibility of Chinese language in question understanding; 3) how to ensure the system with great adaptability which can be easily applied to be a shopping guide in a certain new specialized field.

To tackle the first problem, we propose a mixed-initiative framework. The proposed framework is able to take initiative to obtain users' need, perform passive analysis and understanding of users' questions, and switch between the two modes self-adaptively.

Our solution to the second challenge is to analysis Chinese questions by combining grammar and semantic (Huang et al., 2014). First, hand-crafted sentence compression grammar bases including grammar rules and question type patterns are added to the robot. By sentence compression, the diversity and flexibility of Chinese utterances can be recognized and categorized into limited sentence structures. Then, a question understanding method is proposed by combining grammar based question type pattern recognition and semantic based information extraction and organization.

Finally, we address the last problem by using the design concept of building professional knowledge based on intelligent ability, which combines a variety of human language technologies and intelligent technologies. Such design enables Ch2R with great adaptability, making it easily applied to the shopping guide in a new restricted domain, by adding the semantic knowledge and the detailed commodity information of that specialized field.

An example scenario of Ch2R in mobile phone domain is shown in Figure 1.

- | |
|---|
| (1) Ch2R : 您好! (Hello!) |
| (2) User : 你好! (Hello!) |
| (3) Ch2R : 您想买什么品牌的手机呢?
(Which brand of mobile phone do you want to buy?) |
| (4) User : 我要三星的。(I want Samsung.) |
| (5) Ch2R : 您打算买什么价位的?
(How much you are going to buy?) |
| (6) User : 2000 块左右的吧。(About 2000 RMB.) |
| (7) Ch2R : 你想要什么输入方式的?
(Which input method do you want ?) |
| (8) User : 触屏手机有什么好介绍?
(Can you recommend some mobile phones with touching screen?) |
| (9) Ch2R : 有E120L和I9050。(E120L and I9050.) |
| (10) User : 我也不知道哪个好。
(I don't know which one is better.) |
| (11) Ch2R : 你可以考虑屏幕大小或者摄像头像素。
(Maybe you can consider the screen size or the camera pixel.) |
| (12) User : 哪个屏幕大一点?
(Which one's screen is bigger?) |
| (13) Ch2R : E120L。(E120L.) |
| (14) User : 这款摄像头像素高吗?
(Does it has high megapixel camera?) |
| (15) Ch2R : 是的, 800 万像素。
(Yes, it has a 8 mega pixel camera.) |
| (16) User : 好的, 就要这款。(Ok, I will buy it.) |
| (17) Ch2R : 欢迎下次光临! (Welcome your next visit!) |

Figure 1. Example dialogue of Ch2R.

The purpose of this paper is to describe our development effort, and to present some evaluation results on the system. The remainder of this paper is organized as follows: Section 2 presents the detail framework of Ch2R. The implementation of the mixed-initiative conversation mechanism combining intelligent guiding and understanding within the architecture of our shopping guide robot is proposed in Section 3. The professional knowledge kept by the robot is briefly presented in Section 4. Section 5 shows the preliminary evaluation results of Ch2R in mobile phone shopping guide domain. The paper concludes by outlining future developments and possible applications in Section 6.

2 System Framework

Figure 2 shows the system framework of Ch2R.

In the *sentence preprocessing* component, Chinese word segmentation and part-of-speech tagging are processed by ICTCLAS (Zhang et al., 2003). Wrongly written characters, internet language, nickname of a product, etc. , are also dealt with in this step.

Then, two main parts enclosed by bold dotted lines denote the initiative guiding (branch ①, e.g. line (3), (5) and (7) in Figure 1) and the passive understanding (branch ②, e.g. line (8), (12), and (14) in Figure 1) modes, remaining the situation when no valid information exists after the process of *information extraction* component (branch ③, e.g. line (2), (10), and (16) in Figure 1). In that case, *AIML process*, a process based on the Artificial Intelligence Markup Language (AIML) (Wallace, 2003) is used to handle some simple conversations beyond domain knowledge, some cases potentially to switch to initiative guiding and the state of end. The out of domain utterance processing based on AIML will be discussed in detail later in a separate paper.

In the *information extraction* component, semantic information is extracted from the source utterance by using with the *semantic base*. The extracted semantic information is converted into well organized semantic knowledge in *knowledge organization* component.

For interrogative sentence, i.e. a user question, we use the hand-crafted sentence compression grammar rules to perform sentence compression and employed question type patterns in *question structure recognition* component, matching the only one question type pattern taking along with information for semantic organization and answering for any input question (Huang et al., 2014).

The Ch2R architecture embodies the combination of intelligent ability and professional knowledge. From the intelligent perspective, *sentence compression* and *question structure recognition* components show the ability to understand and analyze questions, *live-table* stands for the ability to memorize, Reinforcement Learning (RL) to update the sequence of the attributes in *live-table* embodies the ability of self-learning, and Case Based Reasoning (CBR) of *AIML process* provides the capacity for logical reasoning. From the professional perspective, *info-table* provides detailed commodity information of a certain specialized field. The semantic knowledge of that field is stored in *semantic base*.

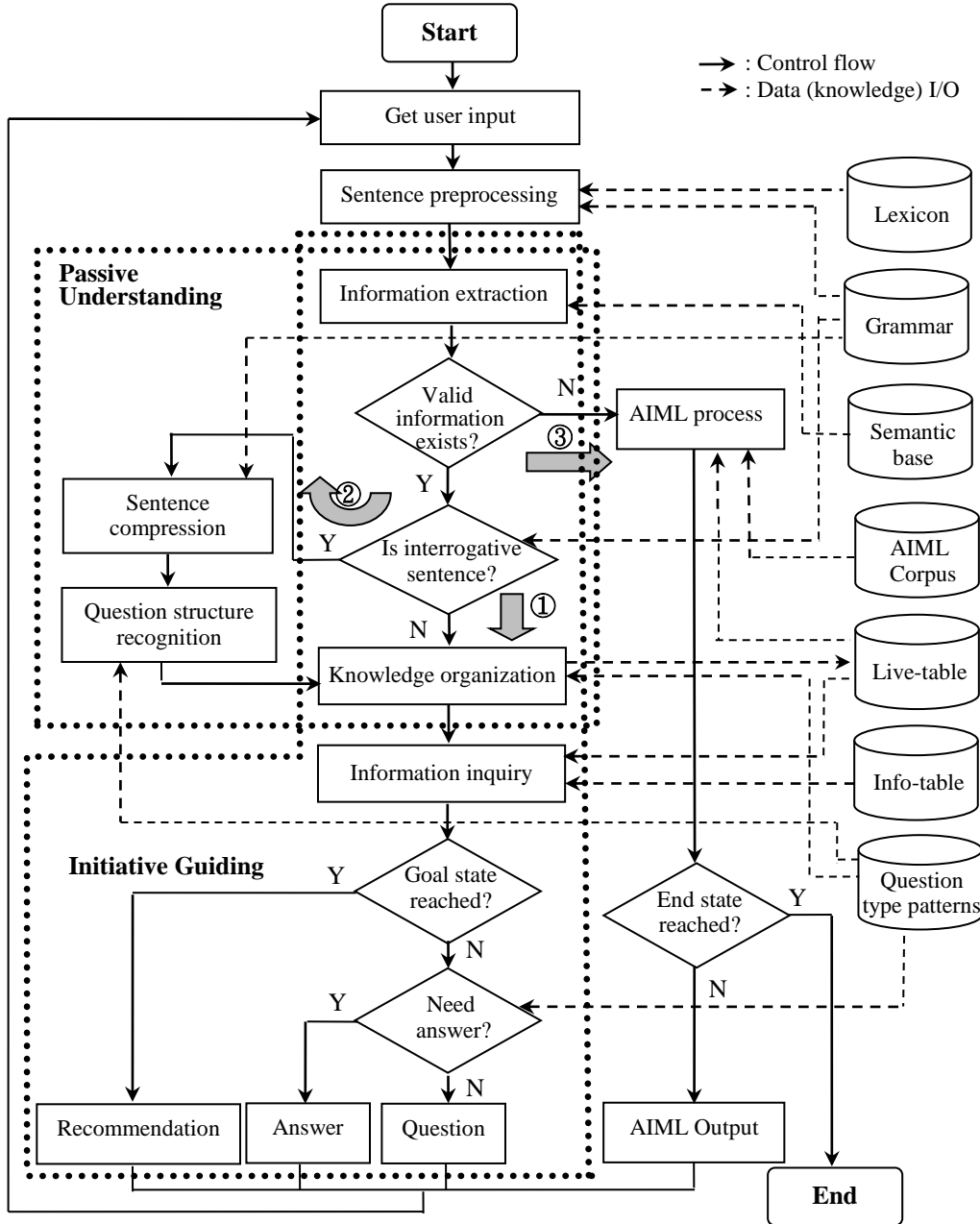


Figure 2. System framework of Ch2R.

3 Mixed-initiative Conversation

3.1 Initiative Guiding

One of the major benefits of Ch2R is that it can provide initiative guiding. We first introduce the *live-table*, and then briefly propose the guiding and recommendation mechanism based on *live-table*.

Live-table: the Ability to Memorize

Live-table is the message storage that acts as the memory of Ch2R. The information of *live-table* is live in the sense that it is active during the

whole process of shopping guide. There are three kinds of active information in *live-table*, including the attribute values, the context of the dialogue, and the recommendation list. The meaning representation of Ch2R is similar to other frame-based dialogue system, in which frame had pre-defined slots that were appropriate for task. Understanding in these systems amounted to extracting specific fillers for each slot (e.g. Brand). Figure 3 shows the update process of *live-table* according to the example dialogue in mobile phone shopping guide from Figure 3 (a).

In our current design, we only keep the last sentences of both chatbot and user to support

answering. The long-term memory is kept by the attribute values in *live-table*. So, *live-table* is the basis when Ch2R checks the candidate mobile phones in Information inquiry component.

.....
Ch2R:您想买什么品牌的手机呢?
 (Which brand of mobile phone do you want to buy?)
User :给我看看三星的吧。(I want Samsung.)
Ch2R:好的, 您打算买什么价位的?
 (How much you are going to buy?)
User :3000 块以下的吧。
 (Less than 3000 Yuan.)
Ch2R:您需要高清的摄像头吗?
 (Do you need a camera with high megapixel?)
User :500 万以上。(More than 5 mega.)

(a) Example dialogue in mobile phone shopping guide

属性(Attribute)	值(Value)
品牌(Brand)	三星 (Samsung)
价格(Price)	(0, 3000)
摄像头像素 (Pixel_camera)	[5,)
输入方式 (Input_method)	
...	...
颜色(Colour)	
型号(Model)	
Ch2R 最新的话 (Last_S_Ch2R)	您需要高清的摄像头吗? (Do you need a camera with high megapixel?)
User 最新的话 (Last_S_User)	500 万以上。 (More than 5 mega.)
推荐列表 (Recomm_list)	

(b) Live-table

Figure 3. Example of the update process of *live-table*.

Guiding and Recommendation

Effective guiding is achieved by looking up the unconfirmed attributes in the *live-table* to present a question. It is worthwhile to note that the sequence of the attributes in *live-table* is variable, which means the priority of the attributes used in initiative guiding can be changed for users' preference in shopping. A process of Reinforcement Learning (RL) (Kaelbling et al., 1996) is used to achieve such flexibility, according to the analysis of the users' questions collected in passive understanding mode.

Ch2R would confirm every attribute, however, this would lead to too many times of interactions with user, and it looks a little mechanical. We address this problem by offering a recommendation in the initiative guiding process once there

are only limited numbers of candidates, e.g. 1, 2 or 3 candidates. Such recommendation limits the interactions and can embody the profession of Ch2R in shopping guide.

3.2 Passive Understanding

We now turn to the passive understanding mode. The word passive means that when face to a user's question, the chatbot has to analyze and answer. Actually, we can simply observed that human being, even a three-year-old child can understand many sentences with different structures and make different responses according to different structure of the questions. It reveals that learning and recognition of sentence type structure is the foundation of the ability to understand and analyze questions. After mastering the grammar sentence type structure, once learning semantic knowledge of a certain domain, one can then dialogue in that specialized field.

Therefore, first, hand-crafted grammar bases in Chinese including *grammar rules* for sentence compression and *question type patterns* for question structure recognition are added to the shopping guide robot, which act as the "language acquisition device" suggested by Chomsky (2005). And then we simplify a complex sentence leaving only the structure by sentence compression based on *grammar rules*. Finally, we employed *question type patterns* for question structure recognition, matching the only one question type pattern for any input question (Huang et al., 2014).

Sentence Compression

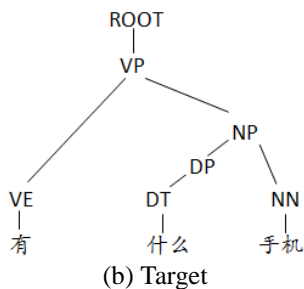
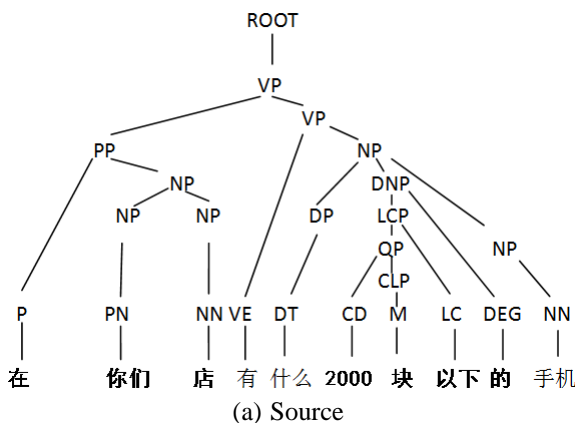
For all kinds of languages, sentences are diversified and innumerable, but the sentence structures are limited. By sentence compression, the diversity of user inputs can be recognized and categorized into limited sentence structures, i.e. question type pattern. Given an input source sentence of words $x = x_1, x_2, \dots, x_n$, a target compression y is formed by removing any subset of these words (Knight and Marcu, 2002). The aim of the sentence compression in our system is to produce a summary of a single sentence that retains the most important structure information while remaining grammatical.

Tree based representation is used in sentence compression. The *Stanford Chinese Parser*¹ (Levy and Manning, 2003) is employed for the tree-based parsing. In order to get a correct syn-

¹ <http://nlp.stanford.edu/software/lex-parser.shtml>.

tax tree from Stanford Parser, we have to formalize the sentence because Stanford Parse can't understand some sentence structures. Then, we use the hand-crafted sentence compression rules, and rely on recent work in text-to-text generation methods (Cohn and Lapata, 2009; Cohn and Lapata, 2013) to perform sentence compression.

Hand-crafted grammar rules for sentence compression are obtained by analyzing hundreds of question examples with different sentence structures. Because our system use tree based representation in sentence compression, the grammar rules take the forms like $(NP(DNP, NP1)) \rightarrow NP(NP1)$, which states that a NP consisting of a DNP and another NP, denoted as NP1, can be rewritten as NP just consisting of NP1 (without the DNP). Taking a wh-question question, “在你们店有什么 2000 块以下的手机? (Which mobile phones are less than 2000 RMB in your store?)”, as example, the sentence compression result is “有什么手机? (Which mobile phones?)”. Figure 4 illustrates such example based on the hand-crafted *grammar rules base* used in our system.



$r1: (VP(PP, VP1)) \rightarrow (VP1)$
$r2: (NP(DNP, NP1)) \rightarrow (NP1)$

(c) Grammar rules used in this example

Figure 4. Example of sentence compression showing the source and target trees, and the grammar rules used. The bold source nodes show the terminals that need to be removed to produce the target string.

Question Type Pattern Recognition

Question type pattern recognition is important for the later steps in information organization and answering. However, building an effective knowledge base of question type patterns is a challenge, especially in Chinese language, which is unlike English, in which question word can basically represents the classification of the interrogative. We design a 4-set question type pattern as {interrogative sentence type, interrogative word type, interrogative phrases, sentence structure}. Taking the compressed sentence, “有什么手机? (Which mobile phones?)”, as example, its question type pattern is “(特指问, 什么, 什么/哪*, $VP+\sim+NP$) (wh -questions, which, which, $VP+\sim+NP$)”, where \sim stand for the interrogative phrase. Other questions, such as “有哪些手机? (Which mobile phones?)” and “有哪款手机? (Which mobile phones?)” will be recognized as this question type pattern. Notice that a more complex question, such as “有什么 2000 块以下的大屏幕的手机? (Which mobile phones are less than 2000 RMB and with big screen?)” and an informal user input, such as “有什么 2000 块左右的? (Which are about 2000 RMB?)”, will be also recognized as the same question type pattern after sentence compression and question type pattern recognition, which shows good robustness of our design. It also leads to good performance with limited question type patterns (30 question type patterns in our current dialogue system) (Huang et al., 2014).

The procedure of question type pattern recognition is shown in Figure 5.

Due to the inaccuracy of Chinese word segmentation in Stanford Parser, to raise the matching rate, we remark the word tags by employing a more satisfied word segmentation interface in the first step of question type pattern recognition. Given a compressed interrogative sentence, ICTCLAS (Zhang et al., 2003) is introduced to remark the word tags and get the syntax sequence.

In the 4-set question type pattern, the interrogative word type is not used as recognition factor. The similarities of interrogative sentence type, the interrogative phrases and the sentence structure are taken as the three factors for computing the similarity between the compressed interrogative sentence and any question type pattern in *question type patterns base*.

Figure 6 shows how to calculate the similarity between the source syntax sequence and the target syntax sequence, i.e. the sentence structure in a certain question type pattern.

Algorithm 1: Question type pattern recognition
Input: Compressed interrogative sentence s
Output: Question type pattern with the highest similarity
Data depend on: *Question type patterns base*

Procedure :

1. get the syntax sequence ss of s
2. for all question type patterns in data:
 - a. judge if s has the features of interrogative sentence type in a certain question type pattern. Here is the score c_1 .
 - b. judge if s has the interrogative phrases in a certain question type pattern. Here is the score c_2 .
 - c. calculate the similarity between ss and the sentence structures in question type patterns. The similarity is c_3 .
 - d. calculate the final similarity $c=c_1*w_1+c_2*w_2+c_3*w_3$, w are the weight between 0 and 1.
3. find the highest score c , and corresponding pattern is the question type pattern of s .
4. return the question type pattern matched.

Figure 5. Algorithm of question type pattern recognition.

Algorithm 2: Sentence structure similarity calculation
Input: Source syntax sequence x , target syntax sequence y
Output: Similarity score, 100 for maximum
Reference algorithm: Classical Edit Distance

Procedure :

1. for syntax in x , do one of the four operation below:
 - e. replace a syntax α in x with β by score $SR(\alpha) + SR(\beta)$, which is always negative.
 - f. add a syntax α to x with score $SA(\alpha)$, which is always negative.
 - g. delete a syntax α in x with score $SD(\alpha)$, which is always negative.
 - h. do nothing with a syntax α and get $SN(\alpha)$, which is always positive.
2. transform x into y with the operations above, figure out the highest score s .
3. transform y into y with the operations above, figure out the highest score s_{max} .
4. return the final similarity score, $100*s/s_{max}$.

Figure 6. Algorithm of sentence structure similarity.

In step 1 of Figure 6, different syntax has different weight of score. For instance, modal particles, adverbs, punctuation will have a lower weight in score calculation, but nouns, verbs and interrogative words will have a higher weight. In

step 2, the highest score is calculated by Edit Distance algorithm (Ristad and Yianilos, 1998).

4 Professional Knowledge

The detailed commodity information and semantic knowledge of a certain restricted domain is the professional knowledge that should be added to Ch2R when applied it to the shopping guide in that specialized restricted domain.

4.1 Info-table: Detailed Commodity Information

Info-table, which is the basic professional knowledge base of Ch2R, provides detailed commodity information of a certain specialized field. There are totally 89 attributes in the Info-table of mobile phone domain, including one as the primary key, and other 88 attributes provide the commodity information in detail. These attributes are selected from major mobile phone e-commerce sites. *Info-table* acts as the complete product information of robot. In other words, it gives the robot more sufficient information than real human in online shopping guide.

4.2 Semantic base: Domain Semantic Knowledge

In *semantic base*, DSem-table stands for the domain semantic knowledge kept by the chatbot. There are totally 77 attributes in the DSem-table of mobile phone domain. Table 1 shows a fragment of it.

Currently, there are three kinds of domain semantic knowledge in Ch2R. The semantic knowledge of the price value is as follows:

1. Conceptual knowledge: “价格/价钱/价位 (price)”, “多少钱/多少块钱/多少元钱 (how much)”.

2. Qualitative knowledge: “便宜 [0,1000) (cheap [0,1000))”, “一般 [1000,2000) (moderate price [1000,2000))”, “较高 [2000,) (expensive [2000,))”.

3. Quantitative knowledge, which is further divided into two types:

- Approximate number knowledge: “以上/大于 (much than)”, “以下/小于 (less than)”, “左右/大概 (about/ approximately)”, “不超过 (not higher than)”, “少/低/便宜一点 (less/lower/ cheaper)”, “高/贵一点 (higher/more expensive)”.

- Exact quantity knowledge: “数字 ($0 \leq \text{数字} < 20000$) + 块/元 (($0 \leq \text{digits} < 20000$) + RMB/ Yuan)”, “数字 ($0 \leq \text{数字} < 20000$) + 到 + 数字 ($0 \leq \text{数字} < 20000$) (($0 \leq \text{digits} < 20000$) + RMB / Yuan to ($0 \leq \text{digits} < 20000$) + RMB/ Yuan)”.

属性 (Attribute)	领域语义知识 (Domain semantic knowledge)
品牌 (Brand)	品牌, 牌子, 三星, 苹果, 华为, 索尼, 诺基亚, 摩托罗拉, 小米, 魅族, 中兴, 黑莓, 步步高, ... brand, Samsung, Apple, Huawei, Sony, Nokia, Motorola, HTC, Meizu, ZTE, BBK...
价格 (Price)	价格, 价钱, 价位, 多少钱, 多少块钱, 多少元钱, 便宜[0,1000], 一般/适中/普通[1000,2000], 较高/贵[2000,], 以上, 以下, 大概/大约, 左右, 大于, 小于, 不超过, 少/低/便宜一点, 高/贵一点, 数字(0<数字<20000)+块/元, 数字(0<=数字<20000)+到+数字(0<=数字<20000). price, how much, cheap [0,1000], moderate price [1000,2000], expensive [2000,], about, approximately, much than, less than, not higher than, less/lower/cheaper, higher/more expensive, (0<=digits<20000)+RMB/ Yuan, (0<=digits<20000)+RMB/ Yuan to (0<=digits<20000)+RMB/ Yuan.
输入方式 (Input method)	输入方式, 手写, 键盘, 触屏, 触控, 语音输入 input method, handwriting, touch screen, keyboarding, voice input
.....

Table 1. Fragment of DSem-table (Mobile phone domain).

Another semantic table is ESem-table, which is the evaluating semantic knowledge kept by the service robot. There are totally 11 attributes in the ESem-table of mobile phone domain in current system. Table 2 shows a fragment of it.

属性 (Attribute)	评价性语义知识 (Evaluating semantic knowledge)
外形 (Outline)	外形, 好看, 一般, 难看, 漂亮, 时髦, 潮流, 时尚, 经典, 过时, 老土. outline, good-looking, bad-looking, beautiful, fashion, popular classical obsolete.
性价比 (Rate of quantity and price)	性价比, 高[100, 85], 中[85, 70], 低[70, 0]. rate of quantity and price, performance-price ratio, high [100, 85], moderate [85, 70], low[70, 0].
热门程度 (Hot degree)	热门, 流行, 受欢迎, 最热. hot, fashion, popular, most popular.
.....

Table 2. Fragment of ESem-table (Mobile phone domain).

Semantic base is used for semantic information extraction. Taking the source sentence of “有什么2000块以下的大屏幕的手机? (Which mobile phones are less than 2000 RMB and with big screen?)” as example, the extracted semantic information is “价格: 2000块, 以下; 主屏尺寸: 屏幕, 大 (Price: 2000 RMB, less than; Screen_size: screen, big)”. Then, the extracted semantic information is converted into well organized semantic knowledge based on the corresponding question type pattern. The extracted semantic information in the above example is organized as “‘价格: 2000块, 以下’and ‘主屏尺寸: 屏幕, 大’ (‘Price: 2000 RMB, less than’ and ‘Screen_size: screen, big’)”. Generally, knowledge organization of the question type patterns with the same interrogative sentence type will be the same or at less similar.

5 Evaluation

5.1 Screenshot

Figure 7 shows a screenshot of the Ch2R in Web-based application form. Chat log is also shown in the web page which is convenient for customer to look over.



Figure 7. Screenshot of Ch2R in Web-based application form.

5.2 Preliminary Evaluation Results

We performed a preliminary system evaluation by logging the interactions of 6 subjects with the system. Each evaluator tests 3 times, i.e. total 18 dialogues. All of the evaluators were familiar with the Ch2R system capabilities, but did not have a detailed knowledge of how to constitute a correct reference answer.

The overall statistical results are shown in Table 3. Branch ①, ② and ③ stand for the turns of initiative guiding, passive understanding and out of domain utterance processing, respectively.

Turns (Avg/Max/ Min)	Turns of Branch ① (Avg/Max/ Min)	Turns of Branch ② (Avg/Max/ Min)	Turns of Branch ③ (Avg/Max/ Min)
13.6/24/5	5.8/10/3	5.2/11/0	2.6/6/1

Table 3. Overall statistical results.

Total turns	Successfully guiding	Success rate
105	100	95.2%

Table 4. Performance of initiative guiding (Branch ①).

Total turns	Correctly understanding	Success rate
93	84	90.3%

Table 5. Performance of question understanding (Branch ②).

Total turns	Correctly process	Success rate
47	42	89.4%

Table 6. Performance of out of domain utterance processing (Branch ③).

As we can see from Table 3, large gaps between the maximum and the minimum turns of both branch ① and ② show the diversity of the evaluators. Some of them like to ask questions, while some others enjoy system-initiative.

The results for the initiative guiding mode are given in Table 4. Our system provided successful active guiding for 100 of the 105 turns of initiative guiding, and made only 5 cases of failed guiding. One of the errors was due to the change of user intent, i.e., the intent of user was changed but the system failed to catch such change and update the *live-table*. The other four errors that the system made were due to the imperfection of the current *semantic base*, which resulted in incorrectly extracting attribute values in utterances, and thus affected the later guiding.

The results for question understanding are given in Table 5. There are total 96 user questions in the test, 3 of which incorrectly enter the out of domain utterance processing branch due to the imperfection of the *semantic base*. In 93 user questions that entering branch ②, 90.3% were correctly understudied (both correctly semantic information extraction and question type pattern recognition), including some utterances with typing mistakes or ellipsis. We also found that the 93 user questions are only distributed within 12 different question type patterns. That means of

the total 30 question type patterns in current system, 18 did not exist in the test. This is mainly because most of the questions in the test are wh-questions and yes-no questions. In 9 incorrect cases, most of them were due to the inaccuracy brought by part-of-speech tagging and the imperfection of the current *semantic base*. Only 2 errors were due to factors related to the design of the question type patterns.

The results for out of domain utterance processing are given in Table 6. Of the 5 sentences that provided incorrect answers, 4 were also due to the imperfection of the *semantic base*, and thus incorrectly leading the dialogue process entering out of domain utterance processing branch.

6 Conclusions and Future Work

This paper presents the development and preliminary evaluation of a Chinese conversational dialogue system named Ch2R with intelligent ability and professional knowledge for online shopping guide. As we can see from the evaluation results, it can perform well in the mobile phone shopping guide in all kinds of situations including initiative guiding, passive understanding, and out of domain utterance processing. Although still in its primary stage, by combining a variety of human language technologies and intelligent technologies into an integrated framework, it can dialogue like a human being and provide a professional service. Moreover, the design concept of building professional knowledge based on intelligent ability ensures Ch2R with great adaptability. We can easily apply it to the shopping guide of other specialized fields.

There are many possible and promising research directions for the near future. We are implementing new and funny interaction, such as communicate in voice using WeChat. Moreover, a separate component of dialogue management with explicit dialogue model will be added to the system. In addition, we also want to experiment with a larger number and various types of users which will make Ch2R more robust.

Acknowledgments

This work was supported by the Innovation Training Project for College Students of Guangdong Province under Grant No.1056412151, No.1056413096 and No.201410564290.

References

Chomsky N.. 2005. Three Factors in Language Design. *Linguistic Inquiry*, 36(1): 1-22.

- Cohn T. and Lapata M.. 2009. Sentence Compression as Tree Transduction. *Journal of Artificial Intelligence Research*, 34(1):637–674.
- Cohn T. and Lapata M.. 2013. An Abstractive Approach to Sentence Compression. *ACM Transactions on Intelligent Systems and Technology*, 4(3): 1-35.
- Gorin A. L., Riccardi G. and Wright J. H.. 1997. How may I help you?. *Speech Communication*, 23: 113-127.
- Huang P.J., Huang Q., Wu X.P., et al.. 2014. Question Understanding by Combining Grammar and Semantic for Chinese Dialogue System. *Journal of Chinese Information Processing*, in press. (in Chinese)
- Huang Y.F., Zheng F., Yan P.J., et al.. 2001. The Design and Implementation of Campus Navigation System: EasyNav. *Journal of Chinese Information Processing*, 15(4):35-40. (in Chinese)
- Kaelbling L.P., Littman M. L., and Moore A. W.. 1996. Reinforcement Learning: A survey. *Journal of Artificial Intelligence Research*, 4: 237-285.
- Knight, K. and Marcu, D.. 2002. Summarization Beyond Sentence Extraction: a Probabilistic Approach to Sentence Compression. *Artificial Intelligence*, 139 (1): 91-107.
- Levy R. and Manning C.D.. 2003. Is it harder to parse Chinese, or the Chinese Treebank?. In *Proceedings of the 41th Annual Meeting of the Association for Computational Linguistics (ACL 2003)*, pages 439-446.
- Liu J., Xu Y.S., Seneff S. and Zue V.. 2008. CityBrowser II: A Multimodal Restaurant Guide in Mandarin, In *Proceedings of the 9th International Symposium on Chinese Spoken Language Processing (ISCSLP 2008)*.
- Liu J., Seneff S. and Zue V. 2010. Dialogue-Oriented Review Summary Generation for Spoken Dialogue Recommendation Systems. In *Proceedings of the 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2010)*.
- Liu J., and Seneff S.. 2011. A Dialogue System for Accessing Drug Reviews, In *Proceedings of the IEEE Automatic Speech Recognition and Understanding Workshop (ASRU 2011)*.
- Pappu A., Rudnicky A.. 2012. The Structure and Generality of Spoken Route Instructions. In *Proceedings of the 13th SIGdial Workshop on Discourse and Dialogue*, pages 99-107.
- Price P.J., 1990. Evaluation of Spoken Language Systems: the ATIS Domain. In *Proceedings of DARPA Workshop on Speech and Natural Language*, Hidden Valley, PA.
- Raux A., Langner B., Black A., and Eskenazi M.. 2003. LET'S GO: Improving Spoken Dialog Systems for the Elderly and Non-natives. In *Proceedings of the Eurospeech 2003*.
- Ristad E.S. and Yianilos P.N.. 1998. Learning String-edit Distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(5):522–532.
- Wallace R.S.. 2003. *The elements of AIML style*. A.L.I.C.E. Artificial Intelligence Foundation, Inc..
- Zhang H.P., Yu H.K., Xiong D.Y., et al.. 2003. HHMM-based Chinese Lexical Analyzer ICTCLAS. In *Proceedings of the second Sighan workshop affiliated with the 41th Annual Meeting of the Association for Computational Linguistics (ACL 2003)*, pages 184-187.
- Zhang L., Gao F., Guo R., et al.. 2004. A Chinese Spoken Dialogue System about Real-time Stock Information. *Computer Applications*, 24(7):61-63. (in Chinese)
- Zue V., Seneff S., Glass J., et al.. 2000. JUPITER: A Telephone-Based Conversational Interface for Weather Information. *IEEE Transactions on Speech and Audio Processing*, 8(1): 85–96.