International Symposium on Chinese Spoken
Language Processing (ISCSLP 2004)
Hong Kong
December 15-18, 2004

THE DISAMBIGUATION STRATEGIES OF SEMANTIC ANALYSIS IN CHINESE SPOKEN DIALOGUE SYSTEM

Bei Liu, Limin Du

Center for Speech Interaction Technology Research
Institute of Acoustics, Chinese Academy of Sciences, Beijing 100080
{liub, dulm}@ siit.ac.cn

ABSTRACT

Semantic frame analysis is one of the most commonly used semantic analysis methods in Chinese Spoken Dialogue System research. And the two typical ambiguous structures commonly encountered in semantic analysis are relation-ambiguity and structural–ambiguity. According to the features of these two ambiguous structures, this paper puts forth the Semantic PCFG model based disambiguation strategy to solve structural-ambiguity and the Expectation Model (EM) based disambiguation strategy to solve relation-ambiguity. Efficient algorithms of the two methods are also provided. The experimental results show that applying these two disambiguation strategies can most greatly improve the performance of the language understanding in base-line system. Especially, Sentence Accuracy is improved from 75.7% to 91.5%, and the three targets of Semantic Unit Understanding Rate--Correction, Recall, and Precision are also improved 10% averagely.

1. INTRODUCTION

Ambiguity is one of the most difficult problems faced by computational linguistics. And actually it is caused by the contradiction between meanings and formations [1]. Moreover, as the main input form of spoken dialogue system (SDS), spoken language contains more ambiguities. It is mainly composed of simple sentences, and contains multiple ungrammatical phenomena and ill formed sentence segments, such as ellipsis, repeat and correction, which always lead to semantic analysis failures and ambiguities. However, if we analyze these sentences within dialogue context, correct semantic relations can be restored then most ambiguities can be solved. Therefore, this kind of ambiguity is called semantic-relation ambiguity.

Besides, with the expansion of analysis rule set, the interactions between rules are strengthened, so ambiguity is inevitable to occur during analysis. Such kind of ambiguity that results from rules competition is called structural ambiguity. This paper aims at putting forth reasonable disambiguation strategies to solve these two kinds of ambiguities.

The description granularity of rule set and its coverage rate are contradicting factors. However, rule methods integrating with probabilistic methods will provide a good solution to this problem [2][3], especially in syntactic analysis. Therefore, this paper uses such method in semantic analysis disambiguation. We will make probabilistic transform to semantic rules, and build up the Semantic PCFG (Probabilistic Context Free Grammar) model. Then in the bottom up analysis, probabilistic values will be calculated for each of the competing rules, and used in semantic tree pruning.

Relation ambiguity disambiguation mainly depends on context information, such as dialogue history, context recognition hint words and co-occurring probability [4]. But these parameters haven't provided enough constrains in analyzing sentence segments. Then we use another special context feature—expectation [5]. In every turn, dialogue manager will make expectations of the user dialogue act in next turn according to the domain knowledge model, dialogue history and certain dialogue strategy. Expectations can provide strong constrains to semantic analysis, and help to restore sentence segments into complete dialogue act. The main task is to determine the case roles of the semantic concepts and user intentions, so as to disambiguate.

2. THE SEMANTIC PCFG MODEL BASED DISAMBIGUITION

The preprocessing of semantic analysis in the baseline system is to map terminal words into semantic concepts. And the semantic CFG rule set defines the semantic relations between concepts. However, spoken language is flexible both in sentence style and order, so we can not define exact analysis rules. Then when there isn't enough information to support decision-making, the choice and application of rules will meet confliction and competition. In such case, the priority of rules defines the application order of rules in confliction, which is rather important in determinate analysis algorithm.

2.1. The Semantic PCFG Model

The semantic CFG rules are different from syntactic rules in the following aspects.

Fist, syntactic analysis are based on the part of speech tags of words, but semantic analysis are based on semantic concepts and the rules mainly describe the hierarchical structure of concepts.

Second, ungrammatical sentence will result in an incomplete syntactic tree with its probability being zero. However, semantic analysis emphasizes on partial analysis, and generate partial semantic tree, which is suitable to spoken language.

Third, after syntactic analysis, the syntactic tree with the highest probability will be the optimal one. But in semantic analysis, probability will be calculated for each path when ambiguities occur, so as to aid pruning and ensure the real-time computability.

Then we give the definitions of semantic rules.

$A :=< c_1,[c_2],\cdots,c_n >$ or $A := \{c_1,[c_2],\cdots,c_n\}$

Where, <> denotes order, {} denotes unordered and [] denotes optional. The right part of the rule defines all the probable concept sequence composition of concept A. Then the application of rules is a kind of fuzzy match. When all of the indispensable elements exist, the rule can be matched successfully.

The method of probabilistic transform is to add a probability $P(c_i \mid r)$ to each of the elements of the right part of the rule. It embodies the weight of this element in this rule or when this element exists, the probability of that the whole rule can be matched. Moreover, probabilities should satisfy the following unitary requirement.

$$\sum_{i=1}^{n} P(c_i \mid r) = 1 \qquad (1)$$

Therefore, CFG model is expanded into PCFG model. The probability of using the rule on the concepts sequence $c_k \cdots c_l$ can be calculated by the following equation.

$$P(r) = P(r \mid c_k,\cdots,c_l) = \sum_{i=k}^{l} P(c_i \mid r) \qquad (2)$$

On the other hand, to make the rule matching process more reasonable, we set a threshold $\theta(r)$ for each rule. Then the rule can be added into the candidate rule set only when $P(r) \geq \theta(r)$. Generally, $\theta(r)$ can be the sum of the weight of the indispensable elements of the rule.

2.2. Parameter Set of the PCFG Model

The procedure of computing the parameter value of the semantic PCFG model can be divided into two phases. Firstly, we can set initial value to each of the parameter, and then adjust them dynamically during corpus training process.

The principles of setting initial values are as follows,

1. The probability of any indispensable element will be larger than the total probability of all the dispensable ones.
2. The total probability of all the elements in a rule should be 1.

Then we can estimate the parameters of PCFG model using most-probable method through corpus training.

Suppose $C(\cdot)$ denotes frequency count, and $rhs(r)$ denotes the right part of rule r, then we define,

$$P(c_i \mid r) = \frac{P(c_i, r)}{P(r)} = \frac{C(c_i, r)}{\sum_{c_i \in rhs(r)} C(c_i, r)} \qquad (3)$$

2.3 The Rule Selection Algorithm

First, all the rules whose right part includes $c_1,\cdots,c_n$, and satisfies the order requirement constitute the rule set $RS(c_1,\cdots,c_n)$. If $RS(\cdot)$ has more than one rule, then the candidate rules will cause ambiguity during semantic analysis. It is obvious that $RS(c_k,c_l,c_m) \subseteq RS(c_k,c_l)$, so the fewer the rule set has rules the less ambiguous they cause. This is the first principle for rule selection, which is called "Max Coverage Best Principle".

In the bottom-up analysis, suppose current concept sequence is $c_i,\cdots c_j$, then the initial candidate rule set should be $RS(c_1,c_2)$. If it is not vacant, we will use the rule selection algorithm to disambiguate.

Algorithm 1 Rule Selection Algorithm

1. Suppose the initial candidate set is $E = E_1 = RS(c_1,c_2)$
2. Generate $E_i = RS(c_1,c_2,\cdots,c_{i+1})$ (i>=2)
3. If $E_i \neq \phi$, then let $E_{i-1} = E_{i-1} - E_i$, $E = E_i$, i=i+1
4. Repeat step 2-3 until $E_i = \phi$ (from E_1 to E_i ,the rule set's priority is increasing)
5. Set E_i with the highest priority to be the current rule set, that is $E = E_i$, and let $E_0 = \phi$
6. Compute probability $P(r)$ for each rule in E, and let $\hat{r} = \arg\max_{r \in E}(P(r))$
7. If $P(\hat{r}) \geq \theta(\hat{r})$, Then $\hat{r}$ is the best rule and go to step 9, else continue
8. Delete E_i and let $E = E_{i-1}$, if $E \neq \phi$, then repeat step 6-8, else continue
9. End

In the algorithm, E_i generation actually applies a looking-forward technique. And the concept composition of the candidate rule can be a kind of heuristic knowledge. It effectively avoids the situation that the rules with lower initial probabilities but have high matching probabilities later are pruned rather earlier. So this algorithm can guarantee local-optimization.

On the other hand, if there is no matchable $\hat{r}$, then $RS(c_2, c_3)$ will be set as the initial candidate rule set, and repeat the algorithm.

3. THE EXPECTATION MODEL BASED DISABIGUATION

3.1. The Generation of Expectation Model

Case Frame Representation is one of the commonly used semantic representation forms. User speech acts comprising user intentions and information items can be expressed by case frame structure, in which, information items corresponds to the slot-value structure, and can be classified as a case according to its semantic role. Generally, an utterance can be parsed into a series of concepts. Then the case role of each concept is decided by the semantic relations between concepts, especially by the case identify words which are also a kind of context feature. One example is showed in Figure 1.

Sentence: 订一张从北京到上海的火车票
(Reserve a train ticket from Beijing to Shanghai)

User intention: *Reserve-Ticket*
Intention identifier: reserve
Case role:
TrainTicket.ticketnumber: one **case identifier**: 张
TrainTicket.DepStation: Beijing **case identifier**: from
TrainTicket.AriStation::Shanghai **case identifier**: to

Fig 1. An example

Case identifiers will help to correctly recognize user intentions and information items. However, ungrammatical sentences, like ellipsis, repeat and correction segments usually cause identify words missing or misunderstanding. Then Expectation Model provides an efficient way to generate reasonable expectation for the case roles of information items and user intentions.

Expectation Generation is based on three aspects: 1 system task model; 2 current user input; 3 dialogue history. EM can generate case role and intention expectation candidates dynamically with dialogue progress, then sort them by priority. This paper uses the algorithm described in reference [6] to build EM, and the shoot rate can reach 90%. Besides, to meet the requirement of mix-initiate strategy, EM can help to recognize user intention transitions in the whole system tasks range.

3.2. The EM Based Disambiguation Algorithm

Case role expectations are set in terms of system inquiries. If user hasn't provided the necessary information items to complete system task, system will start query sub-dialogue to ask their values, at the same time set correspondent case role expectations.

Another important function of EM is to generate user intention expectation e_i, and assign a unified priority $r(e_i)$ for each expectation candidate. If the user intention hasn't been recognized, the input concept sequence has to be matched with expectation candidates, so as to disambiguate. Then system will select the most matchable candidate as the user intention. This can be described by equation (4).

$$\hat{I}_n = \arg\max_I P(I_n \mid C_q, E) = \arg\max_I \{B(e_i)\} \quad (4)$$

where, E is the expectation candidates set and $B(e_i)$ is the belief value of expectation candidate. $B(e_i)$ is decided by two aspects, one is the priority of e_i, the other is the matchability between e_i and the input concept sequence C_q. Since every user intention has been defined a standard concept composition C_r in semantic analysis rules, then the matchability can be calculated basing on the similarity between C_q and C_r.

The similarity can be calculated using equation (5).

$$M(e_i) = Sim(C_q, C_r) = \sum_{c_i \in C_q} \sum_{c_j \in C_r} K(c_i, c_j) T(c_j) \quad (5)$$

where,

$$K(c_i, c_j) = \begin{cases} 1 & if\ c_i = c_j \\ T(c_i) & if\ c_i \in Sub(c_j) \\ 0 & otherwise \end{cases} \quad (6)$$

Here, $Sub(c_j)$ denotes the lower layer concepts set of c_j in the hierarchical model of concepts. $T(c_j)$ and $T(c_i)$ can be calculated using the parameters of the semantic PCFG model as follows,

$$T(c_j) = P(c_j \mid r)$$

$$T(c_i) = P(c_j \xrightarrow{r'} c_i) = P(c_i \mid r')$$

Therefore,

$$B(e_i) = \lambda r(e_i) + (1 - \lambda) M(e_i) \quad (0 < \lambda < 1) \quad (7)$$

The value of λ is determined through experiments, and the experimental result is showed in Figure 2. When λ is between 0.6 and 0.9, the shoot rate of expectation reaches its best performance 90%, so we choose λ =0.08.

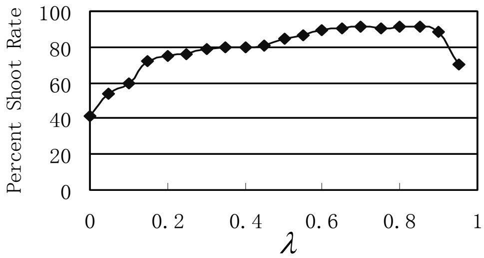

Fig.2. The shoot rate changing curve

4. EXPERIMENTS

The baseline system is Chinese SDS development platform BEST (Beijing Railway Station Ticket Information System), in which the semantic PCFG model totally defines 142 rules. The user inputs mainly apply textual form, but they can simulate all the phenomena in spoken language. In the experiments, we test the performance of semantic analyzer with three different disambiguation strategies. First, we use only the Expectation Model; second, use only the PCFG model; third, use the previous two strategies together.

The evaluation targets are defined as follows [7],

1. Sentence Accuracy (SA)

$$SA = \frac{\sum \# \, of \;\, correctly \;\, parsed \;\, sentences}{\sum \# \, of \;\, total \; user \; sentences \; in \; one \; dia\log ue} \qquad (8)$$

2. Semantic Unit understanding rate (SU), which includes these three targets--correct, recall and precision.

$$SUcorrect = 1 - \frac{\sum(SUS + SUD + SUI)}{\sum SU} \qquad (9)$$

$$SUrecall = \frac{\sum SUC}{\sum SU} \qquad (10)$$

$$SUprecision = \frac{\sum SUC}{\sum(SUC + SUS + SUI)} \qquad (11)$$

where，*SU* denotes the number of semantic units user inputs, *SUC* denotes the correctly recognized units, *SUD* denotes the number of delete errors, *SUS* denotes the number of substitution errors, *SUI* denotes the number of insertion errors. And *SU=SUC+SUD+SDS*.

The test corpus includes 30 dialogues, 163 user utterances. Table 2 gives the experimental results contrast.

Table 1. Experimental Result Contrast

Disambigation Strategy	SA(%)	SU(%)		
		Correct	Recall	Precision
Baseline	75.7	81.3	82.6	87.5
＋PCFG	83.2	84.5	85.7	90.2
＋EM	89.6	90.5	91.7	94.4
＋EM ＋PCFG	91.5	92.4	93.6	96.4

We can see from table 2 that, after applying the three disambiguation strategies, SA and SU are improved greatly, so does the robustness of semantic analysis. However, by only applying the PCFG model, system's performance is improved not so dramatically as by only applying the Expectation Model. This is because the baseline system applies a online developing pattern like develop-test-revise-test. Then the rule set has been revised and optimized continuously during developing and testing. So the disambiguation between rules has been reduced to a low degree. Altogether, using the two disambiguation strategies simultaneously can derive the best performance (SA reaches 91.5%).

5. CONCLUSION

In Chinese SDS, the two typical ambiguities frequently met with in semantic analysis are relation ambiguity and structural ambiguity. Then this paper puts forth the Expectation Model based disambiguation strategy and the PCFG model based disambiguation strategy. The flexibility of spoken language and speech recognition errors will result in a great deal of ungrammatical segments which bring about semantic analysis difficulties. However, integrating dialogue context with the EM model will solve the problems of case role determining and user intentions reasoning. While building PCFG model can solve the structural ambiguities produced during semantic analysis, at the same time improve the coverage of rule set. Experimental results show that applying the EM and PCFG model together can acquire the best robustness in semantic analysis.

11. REFERENCES

[1] Z.W. Feng，*Fundamental Computational Language*，Business Publishing House, Beijing, 2001.

[2] Young S. "Talking to machines (statistically speaking)". in Proc. ICSLP'02, Denver, USA, pp. 9-16. 2002.

[3] A. Stolcke ,"An Efficient Probabilistic Context-Free Parsing Algorithm that Computes Prefix probabilities", Computational Linguistics, MIT Press for the Association for Computational Linguistics, 1995.

[4] Nobuo INUI, Toshiaki EBE, et. al., "A Case-Based Language Dialogue System Using Dialogue Act", IEEE International Conference on Systems, Man and Cybermetics, pp.193-198, 2001.

[5] Smith B W. *Spoken Natural Language Dialogue System: A Practical Approach,* Oxford University Press, London, 1994.

[6] Liu B, Du L.M, Yu S.Y, "The Method of building Expectation Model in Task-oriented Dialogue Systems and Its Realization Algorithms", IEEE NLP-KE'03, Beijing, China, pp.174-179, 2003.

[7] Michael F. "Spoken Dialogue Technology: Enabling the Conversational User Interface", ACM Computing Surveys, vol.34, pp.90 – 169, 2002.