One Tense per Scene: Predicting Tense in Chinese Conversations

Tao Ge[1,2], Heng Ji[3], Baobao Chang[1,2], Zhifang Sui[1,2]
[1]Key Laboratory of Computational Linguistics, Ministry of Education,
School of EECS, Peking University, Beijing, 100871, China
[2]Collaborative Innovation Center for Language Ability, Xuzhou, Jiangsu, 221009, China
[3]Computer Science Department, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
getao@pku.edu.cn, jih@rpi.edu,
chbb@pku.edu.cn, szf@pku.edu.cn

Abstract

We study the problem of predicting tense in Chinese conversations. The unique challenges include: (1) Chinese verbs do not have explicit lexical or grammatical forms to indicate tense; (2) Tense information is often implicitly hidden outside of the target sentence. To tackle these challenges, we first propose a set of novel sentence-level (local) features using rich linguistic resources and then propose a new hypothesis of "One tense per scene" to incorporate scene-level (global) evidence to enhance the performance. Experimental results demonstrate the power of this hybrid approach, which can serve as a new and promising benchmark.

1 Introduction

In natural languages, tense is important to indicate the time at which an action or event takes place. In some languages such as Chinese, verbs do not have explicit morphological or grammatical forms to indicate their tense information. Therefore, automatic tense prediction is important for both human's deep understanding of these languages as well as downstream natural language processing tasks (e.g., machine translation (Liu et al., 2011)).

In this paper, we concern "semantic" tense (time of the event relative to speech time) as opposed to morphosyntactic tense systems found in many languages. Our goal is to predict the tense (past, present or future) of the main predicate[1] of each sentence in a Chinese conversation, which has never been thoroughly studied before but is extremely important for conversation understanding.

Some recent work (Ye et al., 2006; Xue and Zhang, 2014; Zhang and Xue, 2014) on Chinese tense prediction found that tense in written language can be effectively predicted by some features in local contexts such as aspectual markers (e.g. 着 *(zhe)*, 了 *(le)*, 过 *(guo)*) and time expressions (e.g., 昨天 *(yesterday)*). However, it is much more challenging to predict tense in Chinese conversations and there has not been an effective set of rules to predict Chinese tense so far due to the complexity of language-specific phenomena. Let's look at the examples shown in Table 1.

In general, there are three unique challenges for tense prediction in Chinese conversations:

(1) Informal verbal expressions: sentences in a conversation are often grammatically incorrect, which makes aspectual marker based evidence unreliable. Moreover, sentences in a conversation often omit important sentence components. For example, in conversation 1 in Table 1, "如果*(if)*" which is a very important cue to predict tense of verb "废*(destroy)*" is omitted.

(2) Effects of interactions on tense: In contrast to other genres, conversations are interactive, which may have an effect on tense: in some cases, tense can only be inferred by understanding the interactions. For example, we can see from conversations 2, 3 and 4 in Table 1 that when the second person (你*(you)*) is used as the object of the predicate "告诉*(tell)*", the predicate describes the action during the conversation and thus its tense is present. In contrast, when the third person is used in a sentence, it is unlikely that the tense of the predicate is present because it does not describe an action during the conversation. This challenge is unique to Chinese conversations.

(3) Tense ambiguity in a single sentence: Sentence-level analysis is often inadequate to disambiguate tense. For example, it is impossible to determine whether "告诉*(tell)*" in conversations 3 and 4 in Table 1 is a past action (the speaker already told) or a future action (the speaker hasn't told yet) only based on sentence-level contexts.

[1]The main predicate of a sentence can be considered equal to the root of a dependency parse

Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Short Papers), pages 668–673,
Beijing, China, July 26-31, 2015.

1	a: [如果(if)]你(you)动(touch)我(my)儿子(son)一下(once)，我(I)先(first)废(destroy)了你(you)。 (If you touch my son, I'll destroy you.)
2	b: 我(I)告诉(tell)你(you)一声，航班(flight)取消(cancel)了。(I'm telling you: the flight is canceled.)
3	c:你 (you)刚刚 (just now)和他(to him)说(say)什么(what)了？ (What did you say to him just now?) d: 我(I)[刚才(just now)]告诉(tell)他(him)一声，航班(flight)取消(cancel)了。(I told him the flight is canceled.)
4	e: 你(you)要(will)干(do)吗(what)去(go)? (What are you going to do?) f: 我(I)[要(will)]告诉(tell)他(him)一声，航班(fight)取消(cancel)了。(I'll tell him the flight is canceled.)
5	a: 发生(happen)了什么(what)事情(event)? (What happened?) b: 我(I)跟吴清(Wu Qing)一起(with) (I was with Wu Qing) b: 我们(We)在(keep)监视(surveilance)一批货(a cargo) (We were keeping surveillance on a cargo...) b: 我们(We)怀疑(suspect)那些(thoses)是(are)偷来的(stolen)文物(antiques) (We suspected those were stolen antiques) b: 那些人(those guys)，突然(suddenly)就走(walk)出来(out)打(beat)我们(us) (Suddenly, all those guys walked out to beat us up!) b: 我(I)要(want)报警(call the police)他们(they)才停手(stop) (They stopped only when I tried to call the police)

Table 1: Five sample conversations that show the challenges in tense prediction in Chinese conversations. a,b,c,d at the beginning of each sentence denote various speakers. The words in square brackets are **omitted content** in the original sentences and the underlined words are main predicates.

In fact, the sentence in conversation 3 omits "刚才*(just now)*" which indicates past tense and the sentence in the conversation 4 omits "要*(will)*" which indicates future tense. If we add the omitted word back to the original sentence, there will not be tense ambiguity.

To tackle the above challenges, we propose to predict tense in Chinese conversations from two views – sentence-level (local) and scene-level (global). We first develop a local classifier with linguistic knowledge and new conversation-specific features (Section 2.1). Then we propose a novel framework to exploit the global contexts of the entire scene to infer tense, based on a new "One tense per scene" hypothesis (Section 2.2). We created a new a benchmark data set[2], which contains 294 conversations (1,857 sentences) and demonstrated the effectiveness of our approach.

ble

2 Method

2.1 Local Predictor

We develop a Maximum Entropy (MaxEnt) classifier (Zhang, 2004) as the local predictor.

Basic features: The unigrams, bigrams and trigrams of a sentence.

Dependency parsing features: We use the Stanford parser (Chen and Manning, 2014) to conduct dependency parsing[3] on the target sentences and use dependency paths associated with the main predicate of a sentence as well as their dependency types as features. By using the parsing features, we can not only find aspectual markers (e.g., "了") but also capture the effect of sentence structures on the tense.

Linguistic knowledge features: We also exploit the following linguistic knowledge from the Grammatical Knowledge-base of Contemporary Chinese (Yu et al., 1998) (also known as GKB):

- Tense of time expressions: GKB lists all common time expressions and their associated tense. For example, GKB can tell us "往年 *(previous years)*" and "中世纪 *(Middle Ages)*" can only be associated with the past tense.
- Function of conjunction words: Some conjunction words may have an effect on tense. For example, the conjunction word "如果*(if)*" indicates a conditional clause and the main predicate of this sentence is likely to be future tense. GKB can tell us the function of common Chinese conjunction words.

Conversation-specific features: As mentioned in Section 1, different person roles being the subject or the object of a predicate may have an effect on the tense in a conversation. We analyze the person roles of the subject and the object of the main predicate and encode them as features, which helps our model understand effects of interactions on tense.

2.2 Global Predictor

As we discussed before, tense ambiguity in a sentence arises from the omissions of sentence components. According to the principle of efficient information transmission (Jaeger and Levy, 2006;

[2] http://nlp.cs.rpi.edu/data/chinesetense.zip
[3] We use CCProcessed dependencies.

Jaeger, 2010) and Gricean Maxims (Grice et al., 1975) in cooperative theory, the omitted elements can be predicted by considering contextual information and the tense can be further disambiguated. In order to better predict tense, we propose a new hypothesis:

One tense per scene: Within a scene, tense in sentences tends to be consistent and coherent.

During a conversation, a speaker/listener can know the tense of a predicate by either a tense indicator in the target sentence or scene-level tense analysis. **A scene** is a subdivision of a conversation in which the time is continuous and the topic is highly coherent and which does not usually involve a change of tense. For example, for the conversation 3 in Table 1, we can learn the scene is about the past from the word "刚刚 *(just now)*" in the first sentence. Therefore, we can exploit this clue to determine the tense of "告诉*(tell)*" as past.

Therefore, when we are not sure which tense of the main predicate in a sentence should be, we can consider the tense of the entire scene. For example, the conversation 5 in Table 1 is about a past scene because the whole conversation is about a past event. For the sentence "我们(We)在(keep)监视(surveillance)一批货(a cargo)" where the tense of the predicate is ambiguous (past tense and present tense are both reasonable), we can exploit the tense of the scene (past) to determine its tense as past.

Global tense prediction

Inspired by the burst detection algorithm proposed by Kleinberg (2003), we use a 3-state automaton sequence model to globally predict tense based on the above hypothesis. In a conversation with n sentences, each sentence is one element in the sequence. The sentence's tense can be seen as the hidden state and the sentence's features are the observation. Formally, we define the tense in the i^{th} sentence as t_i and the observations (i.e., features) in the sentence as o_i. The goal of this model is to output an optimal sequence $\boldsymbol{t^*} = \{t_1^*, t_2^*, ..., t_n^*\}$ that minimizes the cost function defined as follows:

$$Cost(\boldsymbol{t}, \boldsymbol{o}) = \lambda \sum_{i=1}^{n} -lnP(t_i|o_i) + (1-\lambda) \sum_{i=1}^{n-1} \mathbf{1}(t_{i+1} \neq t_i) \quad (1)$$

where $\mathbf{1}(\cdot)$ is an indicator function.

As we can see in (1), the cost function consists of two parts. The first part is the negative log likelihood of the local prediction, allowing the model to incorporate the results from the local predictor. The second part is the cost of tense inconsistency between adjacent sentences, which enables the model to take into account tense consistency in a scene. Finding the optimal sequence is a decoding process, which can be done using Viterbi algorithm in $O(n)$ time. The parameter λ is used for adjusting weights of these two parts. If $\lambda = 1$, the predictor will not consider global tense consistency and thus the optimal sequence $\boldsymbol{t^*}$ will be the same as the output of the local predictor.

Figure 1 shows how the global predictor works for predicting the tense in the conversation 5 in Table 1. The global predictor can correct wrong local predictions, especially less confident ones.

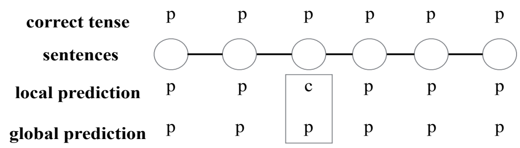

Figure 1: Global tense prediction for the conversation 5 in Table 1.

3 Experiments

3.1 Data and Scoring Metric

To the best of our knowledge, tense prediction in Chinese conversations has never been studied before and there is no existing benchmark for evaluation. We collected 294 conversations (including 1,857 sentences) from 25 popular Chinese movies, dramas and TV shows. Each conversation contains 2-18 sentences. We manually annotate the main predicate and its tense in each sentence. We use ICTCLAS (Zhang et al., 2003) to do word segmentation as preprocessing.

Since tense prediction can be seen as a multi-class classification problem, we use *accuracy* as the metric to evaluate the performance. We randomly split our dataset into three sets: training set (244 conversations), development set (25 conversations) and test set (25 conversations) for evaluation. In evaluation, we ignore imperative sentences and sentences without predicates.

3.2 Experimental Results

We compare our approach with the following baselines:

- Majority: We label every instance with the majority tense (present tense).

- Local predictor with basic features (Local(b))
- Local predictor with basic features + dependency parsing features (Local(b+p))
- Local predictor with basic features + dependency parsing features + linguistic knowledge features (Local(b+p+l))
- Local predictor + all features introduced in Section 2.1 (Local(all))
- Conditional Random Fields (CRFs): We model a conversation as a sequence of sentences and predict tense using CRFs (Lafferty et al., 2001). We implement CRFs using CRFsuite (Okazaki, 2007) with all features introduced in Section 2.1.

Among the baselines, Local(b+p) is the most similar model to the approaches in previous work on Chinese tense prediction in written languages (Ye et al., 2006; Xue, 2008; Liu et al., 2011). Recent work (Zhang and Xue, 2014) used eventuality and modality labels as features that derived from a classifier trained on an annotated corpus. However, the annotated corpus for training the eventuality and modality classifier is not publicly available, we cannot duplicate their approaches.

	Dev	**Test**
Majority	65.13%	54.01%
Local(b)	69.74%	66.42%
Local(b+p)	70.39%	67.15%
Local(b+p+l)	71.05%	69.34%
Local(all)	71.05%	69.34%
CRFs	69.74%	64.96%
Global	**72.37%**	**72.26%**

Table 2: Tense prediction accuracy.

Table 2 shows the results of various models. For our global predictor, the optimal λ (0.4) is tuned on the development set and used on the test set.

According to Table 2, n-grams and dependency parsing features[4] are useful to predict tense, and linguistic knowledge can further improve the accuracy of tense prediction. However, adding conversation-specific features (interaction features) does not benefit Local(b+p+l). The first reason is that the subject and the object of the predicates in many sentences are omitted, which is common in Chinese conversations. The other reason, also the main reason, is that simply using the person roles of the subject and the object is not sufficient to depict the interaction. For example, the subject and the object of the following sentences have the same person role but have different tenses because "警告(warn)" is the current action of the speaker but "教(teach)" is not. Therefore, to exploit the interaction features of a conversation, we must deeply understand the meanings of action verbs.

我(I)警告(warn)你(you)。 (I'm warning you.)

我(I)教(teach)你(you)。 (I'll teach you.)

The global predictor significantly improves the local predictor's performance (at 95% confidence level according to Wilcoxon Signed-Rank Test), which verifies the effectiveness of "One tense per scene" hypothesis for tense prediction. It is notable that CRFs do not work well on our dataset. The reason is that the transition pattern of tenses in a sequence of sentences is not easy to learn, especially when the size of training data is not very large. In many cases, the tense of a verb in a sentence is determined by features within the sentence, which has nothing to do with tense transition. In these cases, learning tense transition patterns will mislead the model and accordingly affect the performance. In contrast, our global model is more robust because it is based on our "One tense per scene" hypothesis which can be seen as prior linguistic knowledge, thus achieves good performance even when the training data is not sufficient.

3.3 Discussion

There are still many remaining challenges for tense prediction in Chinese conversations:

Omission detection: The biggest challenge for this task is the omission of sentence components. As shown in Table 1, if omitted words can be recovered, it will be less likely to make a wrong prediction.

Word Sense Disambiguation: Some function words which can indicate tense are ambiguous. For example, the function word "要" has many senses. It can mean 将要*(will)*, 想要*(want)* and 需

[4]We also tried adding POS tags to dependency paths but didn't see improvements because POS information has been implicitly indicated by dependency types and thus becomes redundant.

要*(need)*, and also it is sometimes used to present an option. It is difficult for a system to correctly predict tense unless it can disambiguate the sense of such function words:

- 一会儿(later)他(he)要 **(will)**过来 (come)。(He**'ll** come here later.)
- 我 (I)要 **(want)** 吃 (eat) 苹果 (apples)。 (I **want** to eat apples)
- 你(you)要**(need)**多 多(much)锻 炼(exercise) (You **need** to take more exercises.)
- 为什么(why)你(you)要**(opt)**救(save)我(me)? (Why did you save me?)

Verb Tense Preference: Different verbs may have different tense preferences. For example, "以为*(think)*" is often used in the past tense while "认为*(think)*" is usually in the present tense:

- 我(I)以 为**(think)**他(he)不 会(won't)来(come) (I **thought** he would not come.)
- 我(I)认 为**(think)**他(he)不 会(won't)来(come) (I **think** he won't come.)

Generic and specific subject/object: Whether the subject/object is generic or specific has an effect on tense. For example, in the sentence "那场(that)战 争(war)太(very)残 酷(brutal)了", the predicate "残 酷(brutal)" is in the past tense while in the sentence "战 争(war)太(very)残酷(brutal)了", the predicate "残酷(brutal)" is in the present tense.

4 Related Work

Early work on Chinese tense prediction (Ye et al., 2006; Xue, 2008) modeled this task as a multi-class classification problem and used machine learning approaches to solve the problem. Recent work (Liu et al., 2011; Xue and Zhang, 2014; Zhang and Xue, 2014) studied distant annotation of tense from a bilingual parallel corpus. Among them, Xue and Zhang (2014) and Zhang and Xue (2014) improved tense prediction by using eventuality and modality labels. However, none of the previous work focused on the specific challenge of the tense prediction in oral languages although the dataset used by Liu et al. (2011) includes conversations. In contrast, this paper presents the unique challenges and corresponding solutions to tense prediction in conversations.

5 Conclusions and Future Work

This paper presents the importance and challenges of tense prediction in Chinese conversations and proposes a novel solution to the challenges.

In the future, we plan to further study this problem by focusing on omission detection, verb tense preference from the view of pragmatics, and jointly learning the local and global predictors. In addition, we will study predicting the tense of multiple predicates in a sentence and identifying imperative sentences in a conversation, which is also a challenge of tense prediction.

Acknowledgments

We thank all the anonymous reviewers for their constructive suggestions. We thank Prof. Shiwen Yu and Xun Wang for providing insights from Chinese linguistics. This work is supported by National Key Basic Research Program of China 2014CB340504, NSFC project 61375074, U.S. DARPA Award No. FA8750-13-2-0045 and China Scholarship Council (CSC, No. 201406010174). The contact author of this paper is Zhifang Sui.

References

Danqi Chen and Christopher D Manning. 2014. A fast and accurate dependency parser using neural networks. In *EMNLP*.

H Paul Grice, Peter Cole, and Jerry L Morgan. 1975. Syntax and semantics. *Logic and conversation*, 3:41–58.

TF Jaeger and Roger P Levy. 2006. Speakers optimize information density through syntactic reduction. In *Advances in neural information processing systems*.

T Florian Jaeger. 2010. Redundancy and reduction: Speakers manage syntactic information density. *Cognitive psychology*, 61(1):23–62.

Jon Kleinberg. 2003. Bursty and hierarchical structure in streams. *Data Mining and Knowledge Discovery*, 7(4):373–397.

John Lafferty, Andrew McCallum, and Fernando CN Pereira. 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data.

Feifan Liu, Fei Liu, and Yang Liu. 2011. Learning from Chinese-English parallel data for Chinese tense prediction. In *IJCNLP*.

Naoaki Okazaki. 2007. CRFsuite: a fast implementation of conditional random fields (CRFs).

Nianwen Xue and Yuchen Zhang. 2014. Buy one get one free: Distant annotation of Chinese tense, event type, and modality. In *LREC*.

Nianwen Xue. 2008. Automatic inference of the temporal location of situations in Chinese text. In *EMNLP*.

Yang Ye, Victoria Li Fossum, and Steven Abney. 2006. Latent features in automatic tense translation between Chinese and English. In *SIGHAN workshop*.

Shiwen Yu, Xuefeng Zhu, Hui Wang, and Yunyun Zhang. 1998. The grammatical knowledge-base of contemporary Chinese—a complete specification.

Yucheng Zhang and Nianwen Xue. 2014. Automatic inference of the tense of Chinese events using implicit information. In *EMNLP*.

Hua-Ping Zhang, Hong-Kui Yu, De-Yi Xiong, and Qun Liu. 2003. Hhmm-based Chinese lexical analyzer ictclas. In *SIGHAN workshop*.

Le Zhang. 2004. Maximum entropy modeling toolkit for Python and C++.