Annotating "tense" in a tenseless language

Nianwen Xue, Hua Zhong and Kai-Yun Chen

Computational Language and Education Center
University of Colorado at Boulder

May 28, 2008
LREC 2008
Marrakech, Morocco

Inferring Chinese "tense"

- 他 这 学期 　教　中文　　。
  He this semester teach Chinese .
  "He is teaching Chinese this semester"
- 他 上 学期 　教　中文　　。
  He last semester teach Chinese .
  "He taught Chinese last semester."
- 他 下 学期 　教　中文　　。
  He next semester teach Chinese .
  "He will teach Chinese next semester."

- Chinese does not grammatically mark tense

Why is this problem interesting?

他　明天　　　返回　上海　　　。
He tomorrow return Shanghai .
"He will return to Shanghai tomorrow."

我明天　　　下班　　　　后　就　回家　　。
I tomorrow get off work after then go home .
"I will go home right after I get off work."

他　过去　　　　常常　在　海边　散布 。
he in the past often on beach walk .
"He often walked on the beach in the past."

他　喜欢　看　　功夫片。
He like watch kung fu movie
"He likes to watch kung fu movies."

Why is this problem interesting?

他 明天 　　 返回 上海 　　 。 　　　　　　　　 future

He tomorrow return Shanghai .

"He will return to Shanghai tomorrow."

我明天 　　 下班 　　　 后 　 就 　 回家 　　 。 　 future

I tomorrow get off work after then go home .

"I will go home right after I get off work."

他 过去 　　　　 常常 　 在 海边 　散布 。 　　　　 past

he in the past often　on beach walk .

"He often walked on the beach in the past."

他 喜欢 看 　　 功夫片。 　　　　　　　　　 present

He like　watch kung fu movie

"He likes to watch kung fu movies."

Why is this problem interesting?

他 明天　　　返回 上海　　　。　　　　　　　future
He tomorrow return Shanghai .
"He will return to Shanghai tomorrow."
"He returned to Shanghai tomorrow.' Google Translate BETA

我明天　　　下班　　　　　后　　就　　回家　　。　future
I tomorrow get off work after then go home .
"I will go home right after I get off work."
"I go home after work tomorrow.' Google Translate BETA

他 过去　　　　常常 在 海边 散布 。　　　　　past
he in the past often on beach walk .
"He often walked on the beach in the past."
"In the past he often walks on the beach Google Translate BETA

他 喜欢 看　　功夫片。　　　　　　　　　present
He like watch kung fu movie
"He likes to watch kung fu movies."
"He liked to watch kung fu." Google Translate BETA

6

Google Translate BETA

Text and Web | **Translated Search** | **Dictionary** | **Tools**

Translate Text

Original text:

他明天返回上海。
我明天下班后就回家。
他过去常常在沙滩上散布。
他喜欢看功夫片。

Automatically translated text:

He returned to Shanghai tomorrow.
I go home after work tomorrow.
He often in the past to spread on the beach
He liked to watch kung fu.

Chinese to English ▼   Translate

⊞ Suggest a better trans...

Translate a Web Page

http://   Chinese to English ▼   Transl

Done

Research questions

- Is it possible for human annotators to infer tense consistently?

- Assuming the answer to the first question is yes, is it possible for the computer to infer tense with a high accuracy?

- Will only try to address the first question in this talk

Previous work

- Ye (2007), 11 tags taken directly from the English tense system
  - Simple present
  - Simple past
  - Simple future
  - Present perfect
  - Past perfect
  - Future perfect   ITA: 43%
  - Present progressive
  - Past progressive
  - Future progressive
  - Present perfect progressive
  - Past perfect progressive

Reasons

➢ Why you might not want to annotate Chinese text with the English tense system:

- ❖ Same tense, different temporal locations
  - ▪ He is coming.
  - ▪ He is playing video games
- ❖ Same temporal location, but different tenses
  - ▪ I expect that I will come.
  - ▪ I expect to come.
- ❖ Might want to translate Chinese into other languages
- ❖ Might be difficult to find annotators who know both languages well enough to do this.

A simplified "tense" annotation framework

- ➢ **Absolute tense: relative to the moment of speech**
  - ❖ Present: a situation is assigned "present tense" if it is true at an interval that includes the present moment.
  - ❖ Past: a situation is assigned "past tense" if it happens anterior to the moment of speech
  - ❖ Future: a situation is assigned the "future tense" if it happens posterior to the moment of speech
- ➢ **Relative tense: relative to another (reference) situation**
  - ❖ Future-in-past: the reference situation is temporally located in the past and the target situation happens posterior to the reference situation

Absolute tense

> 他 这 学期 **教** 中文 。 present
  He this semester teach Chinese .
  "He is teaching Chinese this semester"

> 他 上 学期 **教** 中文 。 past
  He last semester teach Chinese .
  "He taught Chinese last semester."

> 他 下 学期 **教** 中文 。 future
  He next semester teach Chinese .
  "He will teach Chinese next semester."

Relative tense

- ## Relative to a verb situation

  科索沃　　独立　　　　可能　引起 骚乱，

  Kosovo independence might cause riot   .

  联合国　人员　　　已　　**准备**　**撤离**。

  UN　　　personnel already prepare leave .

  "Kosovo independence might cause a riot. UN personnel already prepared to leave."

- ## Relative to a noun situation

  科索沃　　独立　　　　可能　引起 骚乱，

  Kosovo independence might cause riot   .

  联合国 人员　　　　已　　做好 **撤离** 的　**准备**　　。

  UN　　　personnel already make leave DE preparation .

  "Kosovo independence might cause a riot. UN personnel already made preparations to leave."

A real example

➢ 17日奥巴马悄悄搭机**前往**(1)北卡罗来纳州，私下与爱德华兹**会晤**(2)。事后，奥巴马的发言人**证实**(3)了此次会谈，并**说明**(4)他们两人**讨论**(5)了目前的竞选大势，以及美国民众所面临的问题。但是这名发言人没有**说**(6)爱德华兹可能支持的人选。奥巴马在爱德华兹**退选**后，经常公开**称赞**(7)爱德华兹，似乎有心**要**(8)拉拢他作为副手。

➢ On the 17th, Obama quietly **FLY** (1) to North Carolina and **MEET** (2) privately with Edwards. Later, Obama's spokesperson **CONFIRM** (3) the meeting, and **INDICATE** (4) that the two of them **DISCUSS** (5) the political situation and the problems facing the American people. But this spokesperson **DO** (6) not reveal who Edwards might support. Since the halting of Edwards' campaign, Obama often publicly **PRAISE** (7) Edwards, and seemingly **INTEND** (8) to lure him to be his running mate.

A real example

➢ 17日 奥巴马悄悄搭机**前往**(1)北卡罗来纳州，私下与爱德华兹**会晤**(2)。事后，奥巴马的发言人**证实**(3)了此次会谈，并**说明**(4)他们两人讨论(5)了目前的竞选大势，以及美国民众所面临的问题。但是这名发言人没有**说**(6)爱德华兹可能支持的人选。奥巴马在爱德华兹**退选**后，经常公开**称赞**(7)爱德华兹，似乎有心**要**(8)拉拢他作为副手。

➢ **On the 17th**, Obama quietly **flew** (1) to North Carolina and **met** (2) privately with Edwards. Later, Obama's spokesperson **confirmed** (3) the meeting, and **indicated** (4) that the two of them **discussed** (5) the political situation and the problems facing the American people. But this spokesperson **did** (6) not reveal who Edwards might support. Since the halting of Edwards's campaign, Obama **has** often publicly **praised** (7) Edwards, and seemingly **intends** (8) to lure him to be his running mate.

A real example

> **On the 17th**, Obama quietly **flew** (1) to North Carolina and **met** (2) privately with Edwards. Later, Obama's spokesperson **confirmed** (3) the meeting, and **indicated** (4) that the two of them **discussed** (5) the political situation and the problems facing the American people. But this spokesperson **did** (6) not reveal who Edwards might support. Since the halting of Edwards' campaign, Obama **has** often publicly **praised** (7) Edwards, and seemingly **intends** (8) to lure him to be his running mate.

> (1,2): temporal expression

> (3,4): perfective viewpoint

> (5,6,7): bounded situation

> (8): state

Google translation

On the 17th, Obama quietly **flew** (1) to North Carolina and **met** (2) privately with Edwards. Later, Obama's spokesperson **confirmed** (3) the meeting, and **indicated** (4) that the two of them **discussed** (5) the political situation and the problems facing the American people. But this spokesperson **did** (6) not reveal who Edwards might support. Since the halting of Edwards's campaign, Obama **has** often publicly **praised** (7) Edwards, and seemingly **intends** (8) to lure him to be his running mate.

Obama quietly on the 17th fly to North Carolina, met privately with Edwards. Afterward, Obama's spokeswoman confirmed the upcoming talks, and that the two discussed the current election situation, the United States, as well as the problems faced by the people. However, the spokesman did not say that Edwards might support candidates. Obama in the Edwards to withdraw from the election, often publicly praised Edwards seems to want to cozy up to him as his deputy.

17

Boundedness

- Situation type: a situation can be bounded by virtue of
  - Being resultative:
    - 做完 do-finish, 煮黑 boil-black, 醒过来 wake-up
  - Being telic
    - 他在一小时内写了一封信。 "He wrote a letter in a hour."
  - Being modified by a temporal duration or quantifier
    - 他打了三小时游戏。 "He played video games for three hours."
    - 他打了几次游戏。 "He played video games a few times."
  - Being a DE-construction
    - 他 把 孩子 逗 得 笑 了 。 "He made the child laugh."
- Certain viewpoints are only compatible with (un)bounded situations
  - Bounded: perfective, experiential, delimitative, completive
  - Unbounded: progressive, durative, inceptive, continuative

Smith (2005)

- Default deictic pattern of temporal form and interpretation
  - Unbounded situations are located in the Present
  - Bounded situations are located in the Past
- Bounded event constraint
  - Bounded situation are not located in the Present
- Default interpretation can be overwritten with explicit lexical cues
  - Temporal adverbials
  - Conditional clauses
  - Future-oriented verbs (nouns)
  - Modal verbs

Preliminary annotation accuracy *(Xue et al, 2008)*

Tag	Ann1	Ann2	Agree	ITA
Present	276	123	106	.531
Past	501	480	424	.864
Future	180	320	171	.684
N/A	134	168	122	.61
Overall	1091	1091	823	.754

Summary of tense annotation

- Is it possible for human annotators to infer tense consistently?
  - Challenging annotation task, but not impossible
  - Need to further refine the guidelines and subject the annotators to additional training to achieve satisfactory annotation consistency
- The next step is to formalize the linguistic information as features in a machine-learning framework
  - Situation type and viewpoints seem to be crucially important to temporally locate a situation

Thank you. Questions?