

## Chinese Temporal Information Processing Framework Research Progress Report

10th COE Postdoctoral and Doctoral  
Researchers Technical Presentation

Yuchang CHENG

Computational Linguistics Laboratory  
Graduate School of Information Science

Nara Institute of Science and Technology

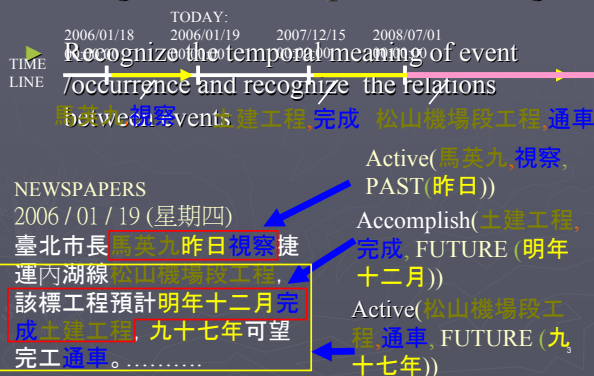
January 26th, 2006

## Goal of Chinese Temporal Information Processing System

- ▶ Using the results of dependency analysis for processing temporal information in Chinese
- ▶ Recognize the temporal meaning of event/occurrence and recognize the relations between events
- ▶ Applications:
  - Q&A, planning, temporal information database, machine translation.....

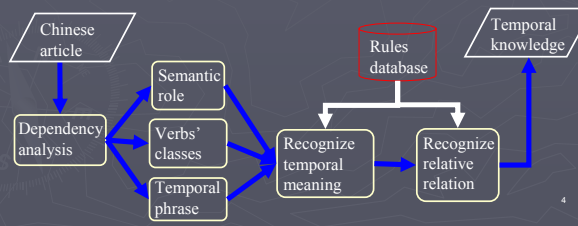
2

## Recognize the Temporal Meaning



## Temporal Information Processing System

- ▶ Using the results of dependency analysis for processing temporal information in Chinese



4

## Temporal Information Processing System

- ▶ Word dependency and semantic role analysis
- ▶ Extract temporal phrase and classify verb
- ▶ Define the rules
- ▶ Recognize temporal meaning and the relative relation between events

5

## Progress on the Dependency Analyzer

- ▶ Try other machine learning methods (kernel)
  - Our system: polynomial kernel (SVMs)
  - Try: tree kernel (SVMs) -> unsuccessful
  - Try: a Boosting Algorithm for Classification of Trees (BACT) -> unsuccessful
- ▶ Try an other chunker
  - Our system: SVMs based chunker (YamCha)
  - Try: Conditional Random Fields based chunker -> improved the PP chunking

6

## Progress on the Dependency Analyzer

- ▶ Participating in the 10<sup>th</sup> CoNLL Shared Task: Multi-lingual Dependency Parsing
  - Word dependency and semantic role
- ▶ Semantic role
  - Our dependency analyzer only analyzes word dependency
  - 松山機場段工程預計明年十二月完成(土建工程完成, FUTURE(明年十二月))
    - Subject (construction)
    - Accomplishment Verb (complete)
    - Object (earthwork)

7

## Temporal Information Processing System

- ▶ Word dependency and semantic role analysis
- ▶ Extract temporal phrase and classify verb
- ▶ Define the rules
- ▶ Recognize temporal meaning and the relative relation between events

8

## Temporal Information in Chinese

- ▶ Verbs don't change the morphemes according to tense
  - ▶ 我去東京-> I (will) go to Tokyo.
  - ▶ 我昨天去東京-> I went to Tokyo yesterday.
- ▶ Temporal information
  - Verb class (temporal feature)
    - ▶ static / dynamic, durative / instantaneous, telic / non-telic
  - Temporal phrase
    - ▶ 在第二次世界大戰期間 (During the WWII...)
  - Temporal adverbs
    - ▶ 在 (being), 已經 (have), 將要 (will)
  - Aspect auxiliary words
    - ▶ 了 (been), 過 (was)

9

## Verb Classification

- ▶ The temporal features of a verb
  - Static / Dynamic
    - ▶ 知道 know/損壞 damage
  - Durative / Instantaneous
    - ▶ 排隊 queue /爆炸 explode
  - Telic / Non-telic
    - ▶ 興建 build / 跑步 run
- ▶ Verb classification
  - 知道 (know) {static, durative, non-telic}
  - 興建 (bulid) {dynamic, durative, telic}

10

## Verb Classification

- ▶ Collect verbs
  - From a linguistics book (現代漢語動詞)
    - ▶ About 300 verbs
    - ▶ Includes temporal features
  - From Penn Chinese Treebank, and Sinica Treebank
    - ▶ About 10,000 verbs
    - ▶ Needs manual classification
    - ▶ Automatically classify?

11

## Extract and Recognize Temporal Phrases

- ▶ Recognize meaning of various temporal phrases
  - Ex. 1950s=1950年代=約五十年前=二次大戰後10年
- ▶ Extract temporal phrases
  - Use chunker to extract phrases
- ▶ Recognize Temporal phrases
  - Use rules to recognize meaning

12

## Extract and Recognize Temporal Phrases

- ▶ Extract temporal phrases
  - From Penn Chinese treebank (TMP) (8254)
  - NP-TMP (49%)
    - ▶ “一九九五年九月” (September 1995)
    - ▶ Many of the NP-TMP include numbers
    - ▶ Regular expressions can be used
  - PP-TMP (22%)
    - ▶ “自成立以來” (after it established)
    - ▶ Need the semantic information
  - LCP-TMP (21%)
    - ▶ “近年來” (in recent years)
    - ▶ Need the semantic information
- ▶ Adopt machine learning method (CRF) to extract temporal phrases
  - Features: word, POS, head word and head POS
  - Precision: 92% Recall: 86%

13

## Extract and Recognize Temporal Phrases

- ▶ Recognize temporal phrases
  - Use rules to recognize meaning
  - Need meaning of number and temporal words
- ▶ Ambiguous of temporal phrase
  - “94年” means “中華民國94年(2005)” or “1994”
  - Need the contextual information
- ▶ Define the meaning of other temporal words
  - Collect temporal words from treebank and linguistics book
    - ▶ “國慶(National Foundation Day)”
    - ▶ About 300 words
  - Manually define the meaning of words
    - ▶ “國慶”={year=xx, month=10, day=1 (or 10)}

14

## Conclusion

- ▶ Participate CoNLL-X Shared Task
  - Word dependency and semantic role
- ▶ Extract temporal phrase and classify verb
  - Classify verb
    - ▶ manual classification
  - Recognize the meaning of various temporal phrase
    - ▶ Regular expression and word definition
- ▶ Define the rules
- ▶ Recognize Temporal information and the relative relation between events

15