Welcome
to the Chinese temporal/discourse annotation workshop

Nianwen Xue
Brandeis University

June 1, 2010
Millennium Biltmore Hotel,
Los Angeles, USA
Why we are here?

• To map out a strategy for cross-lingual temporal annotation that has sufficient coverage and high annotation consistency (measured by inter-annotator agreement).
What is temporal annotation about and how is it different from other types of annotation?

- The ultimate goal of temporal annotation is the temporal structure of a document.
- The building blocks are events, time expressions, and the relations among them.
- Temporal annotation is genre-dependent:
  - Newswire: good for temporal annotation.
  - Tourism guide: bad for temporal annotation.
A strategic question

• Can all events be in a document be ordered temporally?
• Or should our goal to achieve the complete ordering of all events in a document?
A case study

- A Chinese newswire document consisting of 28 identified events
- There are 378 possible event pairs
- The question: how many event pairs are suitable for temporal relation annotation?

<table>
<thead>
<tr>
<th>Temporal relation</th>
<th># of instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclear</td>
<td>144</td>
</tr>
<tr>
<td>Clear but uninformative</td>
<td>86</td>
</tr>
<tr>
<td>Clear and informative</td>
<td>148</td>
</tr>
</tbody>
</table>
Strategic decisions

• Exhaustive annotation infeasible and unnecessary:
  – **Infeasible**: there is no clear temporal relation between many event pairs
  – This means that the “vague” category has to be taken seriously
  – **Unnecessary**: many temporal relations are transitive
  – Have to exploit temporal closure during manual annotation to achieve annotation efficiency

• How do we choose event (or event-time) pairs to annotate?
A (not necessarily new) proposal

- Follow the lead of other annotation layers when selecting event pairs to annotate
  - For intra-sentential temporal relations, follow syntactic and semantic (predicate-argument structure) structures
  - For inter-sentential temporal relations, follow discourse structures
- Infer other temporal relations via temporal closure or other linguistic layers
An example

- The temporal annotation workshop was held on June 1, 2010

Annotate an event pair if one is the modifier of the other or one event is an argument of another event.
Tactical questions

- How do we ensure cross-lingual compatibility?
- How do we ensure annotation consistency?
- How do we improve annotation efficiency?
- How do we make temporal annotation the most useful?
- What types of auxiliary annotation do we need to support temporal inference?
- ....
Cross-lingual compatibility

- Follow the same standard (e.g., ISO-TimeML)
- But also allow cross-lingual variations
  - Tense and aspect annotation
- “Principles and Parameters”
Annotation consistency

• Consistency in two aspects
  – Choosing the same annotation targets
  – Choosing the same temporal relation
• How do we choose?
  – Use guidance from other linguistic annotation layers
• Set up intuitive annotation categories in the guidelines
  – Allow the possibility of no meaningful temporal relation
• Proper annotator training
Annotation efficiency

• Exploit temporal closure in the temporal tool?
  – Are there available tools that do this?
• Explore alternative approaches of annotation
  – Mechanic Turk
What auxiliary annotations are needed for temporal inference?

• Main target of temporal annotation: events, times and relations among them
• Which event is linked to which time? Syntax
• Which events have intuitive temporal relations?
  – Syntax, predicate-argument structure, discourse relations
How to make our annotation useful?

- Annotation is expensive, so it’s desirable to have our annotation support as many NLP applications as possible
  - General-purpose annotation
- Also need to pay attention to specific applications, which might use different aspects of temporal annotation
  - For MT, temporal location is important to infer grammatical tense
  - What’s needed for cross-linguistic information extraction?
Agenda

• Background
  – Nianwen Xue, James Pustejovsky, Marc Verhagen

• Temporal annotation across languages
  – Yuping Zhou, André Bittar, Tommaso Caselli

• Interaction with other annotation layers
  – Joshi/Prasad, Martha Palmer, Chu-Ren Huang

• Annotation efficiency and Applications of temporal annotation
  – Steven Bethard, Heng Ji
Thank You

This workshop is supported by the National Science Foundation via Grant No. 0855184 “CRI CI-P: Building a Community Resource for Temporal Inference in Chinese”
Logistics

• For reimbursement of your travel expenses, contact
  – Jeanne DeBaie <jeanne@cs.brandeis.edu>

• Mailing address
  Michom School of Computer Science
  Volen Center for Complex Systems, Room 261
  Brandeis University
  Waltham, Massachusetts 02254

• Phone/Fax
  – (781) 736-2700 (Department Office)
  – (781) 736-2741 (Fax)