Virtual Worlds (VWs) such as Second Life are becoming increasingly popular for many interaction/collaboration activities. Just as in real world, language barriers exist in VWs. We develop machine translation services to facilitate multilingual communication in VWs. Context plays an important role in disambiguation and generate correct translations. Context information is easier to capture in VWs than in real world. We study context-aware machine translation in VWs.

Virtual World Applications

- Virtual Education
- 3d Reality Classes
- Museum
- Seminar and discussion
- Sightseeing

Machine Translation for Virtual Worlds

- Intercept incoming/outgoing messages and redirect to translation server.
- Phrase-based Statistical Machine Translation system based on CMU's SMT project.

Context-aware Machine Translation

- Chinese sentence: 我想买几张电影票张。tickets, admission, votes? (what?)
  MT output: Can I buy tickets for the opera here?

- Chinese sentence: 舞蹈! the term?
  MT output: There are red it?

- Chinese sentence: 一位是九 saat(points, hours, etc/)
  MT output: The next is a few points?

<table>
<thead>
<tr>
<th>Example</th>
<th>Chinese sentence</th>
<th>Reference translation</th>
<th>MT output:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1</td>
<td>我想买几张电影票张。</td>
<td>Can I buy tickets for the opera here?</td>
<td>can I buy tickets for the opera here?</td>
</tr>
<tr>
<td>Example 2</td>
<td>舞蹈!</td>
<td>There are red it?</td>
<td>There are red it?</td>
</tr>
<tr>
<td>Example 3</td>
<td>一位是九</td>
<td>The next is a few points?</td>
<td>The next is a few points?</td>
</tr>
</tbody>
</table>

Table 2: Examples of incorrect machine translations generated by Google’s online MT system.

- Current MT systems look at surface form only
- Correct meaning of a sentence depends also on non-verbal context information such as time/location/users’ gender/social background …
- Conversational speech tends to be more concise and more context-dependent.
- Context information is more accessible in VWs than in real world which requires physical sensors and sensing algorithms (GPS for location; accelerometers for gesture; gender?)

Ongoing Work

- Improving translation service performance and robustness
- More conversational data collection
- Analyzing context information
- Compare context-independent MT vs. context-aware MT