Supplementary Material for Neural Chinese Address Parsing

Hao Li and Wei Lu
StatNLP Research Group
Singapore University of Technology and Design
hao.li@mymail.sutd.edu.sg
luwei@sutd.edu.sg

Pengjun Xie and Linlin Li
DAMO Academy
Alibaba Group
chengchen.xpj@alibaba-inc.com
linyan.lil@alibaba-inc.com

<table>
<thead>
<tr>
<th>Transition Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-gram; 4-gram; 3-gram; bigram; unigram</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N-gram Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>digit, Chinese digit, punctuation characters; URL; direction words; English characters</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Character Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>segment string; segment length; segment suffix</td>
</tr>
</tbody>
</table>

Table 1: Features used in two baselines ℓCRF and sCRF.

1 Experimental Setup

We describe the discrete features used in two baselines ℓCRF and sCRF respectively.

- ℓCRF is the standard first-order linear CRF model (Lafferty et al., 2001) with discrete features for sequence labeling tasks. Such discrete features include transition features, n-gram features and character features mentioned in Table 1.

- sCRF is based on the standard semi-Markov CRF (Sarawagi and Cohen, 2004) with discrete features. Such discrete features include transition features, n-gram features, character features and segment features mentioned in Table 1.

References
