Transformation and Decomposition for Efficiently Implementing and Improving Dependency-to-String Model In Moses

Liangyou Li

Dublin City University
liangyouli@computing.dcu.ie

Joint work with Jun Xie, Andy Way and Qun Liu

October 25, 2014
Motivation and Objectives

to make a dependency-based model available in Moses [Koehn et al., 2007]

Dependency-to-string (Dep2Str) model [Xie et al., 2011]
- to simplify the implementation of this model
  - transformation of dependency trees
- to improve this model to be comparable with Moses HPB [Chiang, 2005]
  - decomposition of dependency trees
    - sub-structural rules
    - pseudo forest
Bolivia holds presidential and parliament elections

**Figure:** Rule extraction in the Dep2Str model
Overview of the Dep2Str Model

Training

Head Rules

Boliweiya玻利维亚/NN
Juxing举行/VV
Xuanju选举/NN
Guohui国会/NN
Zongtong总统/NN
Yu与/CC

Bolivia holds presidential and parliament elections

Figure: Rule extraction in the Dep2Str model
Overview of the Dep2Str Model

Training

Boliweiya 玻利维亚
Juxing 举行
Zongtong 总统
Yu 与
Guohui 国会
Xuanju 选举
Bolivia holds presidential and parliament elections

Head Rules

Boliweiya 玻利维亚 → Bolivia
Yu 与 → and
Zongtong 总统 → presidential
Guohui 国会 → parliament
Xuanju 选举 → elections
Juxing 举行 → holds

Figure: Rule extraction in the Dep2Str model
Overview of the Dep2Str Model

Training

Boliweiya/NN → Bolivia

Zongtong/NN → presidential

Yu/CC → and

Guohui/NN → parliament

Xuanju/NN → elections

Juxing/VV → holds

Figure: Rule extraction in the Dep2Str model
Overview of the Dep2Str Model

Training

**Head Rules**

Boliweiya 玻利维亚/NN $\rightarrow$ Bolivia

Yu 与/CC $\rightarrow$ and

Xuanju 选举/NN $\rightarrow$ elections

**Head-Dependent Rules**

Boliweiya Xuanju Juxing (玻利维亚) (X1:选举) 举行 $\rightarrow$ Bolivia X1 holds

Figure: Rule extraction in the Dep2Str model
Overview of the Dep2Str Model

Training

**Head Rules**

- Boliweiya 玻利维亚 NN \(\rightarrow\) Bolivia
- Yu 与 CC \(\rightarrow\) and
- Xuanju 选举 NN \(\rightarrow\) elections

**Head-Dependent Rules**

- Boliweiya 玻利维亚 NN 举行 Juxing Xuxing 举行
- Xuanju 选举 NN \(\rightarrow\) elections

Bolivia holds presidential and parliament elections

**Figure:** Rule extraction in the Dep2Str model
Overview of the Dep2Str Model

Training

Head Rules

- Boliweiya 玻利维亚/NN \(\leftrightarrow\) Bolivia
- Yu 与/CC \(\leftrightarrow\) and
- Xuanju 选举/NN \(\leftrightarrow\) elections

Head-Dependent Rules

- Boliweiya 玻利维亚 (X1:选举) 举行 \(\leftrightarrow\) Bolivia X1 holds
- Xuanju 选举/NN \(\leftrightarrow\) X2 X1 holds
- Juxing 举行/VV \(\leftrightarrow\) holds

Figure: Rule extraction in the Dep2Str model
Overview of the Dep2Str Model

Training

**Head Rules**

- Boliweiya 玻利维亚/NN \(\rightarrow\) Bolivia
- Yu 与/CC \(\rightarrow\) and
- Xuanju 选举/NN \(\rightarrow\) elections

**Head-Dependent Rules**

- Boliweiya 玻利维亚/NN \(\rightarrow\) Xuanju 选举/NN \(\rightarrow\) Juxing 举行/VV
- Xuanju 选举/NN \(\rightarrow\) elections
- Juxing 举行/VV \(\rightarrow\) holds

**Figure:** Rule extraction in the Dep2Str model
Overview of the Dep2Str Model

Training

Head Rules

Boliweiya 玻利维亚/NN $\iff$ Bolivia
Yu 与/CC $\iff$ and
Zongtong 总统/NN $\iff$ presidential
Guohui 国会/NN $\iff$ parliament
Juxing 举行/VV $\iff$ holds

Head-Dependent Rules

Boliweiya 玻利维亚 (X1:选举) 举行 $\iff$ Bolivia X1 holds
Xuanju Xuxing (X1:选举) 举行 $\iff$ X2 X1 holds
Guohui 国会 (X1:NN) 举行 $\iff$ X2 X1 holds
Juxing 举行 (X2:NN) X3:NN $\iff$ X2 X1 X3

Figure: Rule extraction in the Dep2Str model
Overview of the Dep2Str Model

Training

**Head Rules**
- Boliweiya 玻利维亚/NN $\rightarrow$ Bolivia
- Yu 与/CC $\rightarrow$ and
- Xuanju 选举/NN $\rightarrow$ elections
- Zongtong 总统/NN $\rightarrow$ presidential
- Guohui 国会/NN $\rightarrow$ parliament
- Juxing 举行/VV $\rightarrow$ holds

**Head-Dependent Rules**
- Boliweiya 玻利维亚 (X1:选举) 举行 $\rightarrow$ Bolivia X1 holds
- Xuanju 选举 (X2:NN) (X1:选举) 举行 $\rightarrow$ X2 X1 holds
- Juxing 举行 (X2:NN) (X1:NN) 举行 $\rightarrow$ X2 X1 holds
- (X2:NN) (X1:NN) X3:NN $\rightarrow$ X2 X1 X3

**Figure:** Rule extraction in the Dep2Str model
Overview of the Dep2Str Model

Decoding

Figure: CYK decoder for the Dep2Str model

Liangyou Li (CNGL, DCU)
Overview of the Dep2Str Model

Decoding
Overview of the Dep2Str Model

Decoding

Figure: CYK decoder for the Dep2Str model

Liangyou Li (CNGL, DCU)
Dep2Str
October 25, 2014 4 / 17
Overview of the Dep2Str Model

Decoding

Figure: CYK decoder for the Dep2Str model

Boliweiya 玻利维亚/NN

Juxing 举行/VV

Boliweiya 玻利维亚/NN

Zongtong 总统/NN

Yu 与/CC

Guohui 国会/NN

Xuanju 选举/NN

presidential and parliament elections
Overview of the Dep2Str Model

Decoding

Bolivia holds presidential and parliament elections

Boliweiya 玻利维亚/NN
Juxing 举行/VV
Zongtong 总统/NN
Yu 与/CC
Guohui 国会/NN
Xuanju 选举/NN

Liangyou Li (CNGL, DCU)
Bolivia holds presidential and parliament elections.

Boliweiya 玻利维亚
Juxing 举行
Zongtong 总统
Yu 总统
与 与
Guohui 国会
Xuanju 选举

Bolivia holds presidential and parliament elections.
Figure: An example of transforming a dependency tree into a constituent tree
Figure: An example of transforming a dependency tree into a constituent tree.
Figure: An example of transforming a dependency tree into a constituent tree.
**Figure:** An example of transforming a dependency tree into a constituent tree.
Figure: An example of transforming a dependency tree into a constituent tree.
Table: BLEU score [%] of the Dep2Str model before (XJ) and after (D2S) dependency tree being transformed. Systems are trained on a selected 1.2M Chinese–English corpus.
Figure: An example of decomposition on a head-dependent fragment.
Decomposition

Formula

\[ L_i \cdots L_1 HR_1 \cdots R_j = L_m \cdots L_1 HR_1 \cdots R_n + L_i \cdots L_{m+1} HR_{n+1} \cdots R_j \]

subject to

\[ i \geq 0, j \geq 0 \]
\[ i \geq m \geq 0, j \geq n \geq 0 \]
\[ i + j > m + n > 0 \]

where \( H \) denotes the head node, \( L_i \) denotes the \( i \)th left dependent and \( R_j \) denotes the \( j \)th right dependent.
Bolivia holds presidential and parliament elections

Figure: Extracting sub-structural rules
Figure: Extracting sub-structural rules
Decomposition
Training: Sub-structural Rules

**Head Rules**
Boliweiya
玻利维亚/NN  $\longrightarrow$ Bolivia  ......  

**Head-Dependent Rules**
Xuanju  Juxing
(X2:NN) (X1:选举) 举行  $\longrightarrow$ X2  X1 holds  
......

**Sub-structural Rules**
Yu Guohui
(与) 国会  $\longrightarrow$ and parliament  
Guohui
(X1:CC) 国会  $\longrightarrow$ X1 parliament  
(X1:CC) X2:NN  $\longrightarrow$ X1 X2  
......

**Figure:** Extracting sub-structural rules
Figure: Extracting sub-structural rules
Figure: An example of translating a large HD fragment with the help of translations of its decomposed fragments
Decomposition
Decoding: Pseudo Forest

(a)

Guohui
国会/NN

Zongtong
总统/NN

Yu
与/CC

(b)

Guohui
国会/NN

Zongtong
总统

and parliament

Rule: (与)国会 → and parliament

Figure: An example of translating a large HD fragment with the help of translations of its decomposed fragments
Figure: An example of translating a large HD fragment with the help of translations of its decomposed fragments.
Figure: An example of creating pseudo-forest for a dependency tree.
Figure: An example of creating pseudo-forest for a dependency tree.
Figure: An example of creating pseudo-forest for a dependency tree.
Figure: An example of creating pseudo-forest for a dependency tree.
Figure: An example of creating pseudo-forest for a dependency tree.
## Experiments

<table>
<thead>
<tr>
<th>Systems</th>
<th>MT04</th>
<th>MT05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moses HPB</td>
<td>35.56</td>
<td>33.99</td>
</tr>
<tr>
<td>D2S</td>
<td>33.93</td>
<td>32.56</td>
</tr>
<tr>
<td>+ pseudo-forest</td>
<td>34.28</td>
<td>34.10</td>
</tr>
<tr>
<td>+ sub-structural rules</td>
<td>34.78</td>
<td>33.63</td>
</tr>
<tr>
<td>+ pseudo-forest</td>
<td>35.46</td>
<td>34.13</td>
</tr>
<tr>
<td>+ phrase</td>
<td>36.76*</td>
<td>34.67*</td>
</tr>
</tbody>
</table>

**Table**: BLEU score [%] of our method and Moses HPB on the Chinese–English task.
# Experiments

<table>
<thead>
<tr>
<th>Systems</th>
<th>Test12</th>
<th>Test13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moses HPB</td>
<td>20.44</td>
<td>22.77</td>
</tr>
<tr>
<td>D2S</td>
<td>20.05</td>
<td>22.13</td>
</tr>
<tr>
<td>+pseudo-forest</td>
<td>19.98</td>
<td>21.68</td>
</tr>
<tr>
<td>+sub-structural rules</td>
<td>20.52</td>
<td>22.76</td>
</tr>
<tr>
<td>+phrase</td>
<td>20.91*</td>
<td>23.46*</td>
</tr>
<tr>
<td>+pseudo-forest</td>
<td>20.25</td>
<td>22.24</td>
</tr>
<tr>
<td>+phrase</td>
<td>20.75*</td>
<td>23.20*</td>
</tr>
</tbody>
</table>

**Table:** BLEU score [%] of our method and Moses HPB on German–English task.
### Experiments

<table>
<thead>
<tr>
<th>Systems</th>
<th># Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CE task</td>
</tr>
<tr>
<td>Moses HPB</td>
<td>388M</td>
</tr>
<tr>
<td>D2S</td>
<td>27M</td>
</tr>
<tr>
<td>+sub-structural rules</td>
<td>116M</td>
</tr>
<tr>
<td>+phrase</td>
<td>215M</td>
</tr>
</tbody>
</table>

**Table:** The number of rules in different systems on the Chinese–English (CE) and German–English (DE) corpus. Note that pseudo-forest (not listed) does not influence the number of rules.
Conclusion

- implement the Dep2Str model without changing the decoder in Moses
- improve this model to be comparable with Moses HPB
- with resort to phase rule, this model is significantly better than Moses HPB

Download

This implementation is available at
http://computing.dcu.ie/~liangyouli/dep2str.zip
References


In Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions, pages 177–180, Prague, Czech Republic.

A Novel Dependency-to-string Model for Statistical Machine Translation.
The End
Thanks for Your Attention