Constituency to Dependency Translation with Forests

Haitao Mi and Qun Liu

presented by Yang Liu

Institute of Computing Technology
Chinese Academy of Sciences
Constituency Trees in SMT

(string-to-constituency)  (constituency-to-string)  (constituency-to-constituency)

Bush held a talk with Sharon

(Galley et al., 2006; Liu et al., 2006; Zhang et al., 2007)
Dependency Trees in SMT

- **string-to-dependency**
- **dependency-to-string**
- **dependency-to-dependency**

(Shen et al., 2008; Xiong et al., 2007; Quirk et al., 2005)
Constituency-to-Dependency

IP
  VP
    PP
      NPB
        bushi
    NPB
      yu
    NPB
      shalong
    VP
      VS
        juxing
      AS
        le
        huitan
  VPB

Bush

held

a

talk

with

Sharon
Synchronous Generation

[Diagram with IP, VP, NPB nodes and arrows indicating relationships]
Synchronous Generation

NPB

IP

VP

NPB

bushi

Bush
Synchronous Generation

IP

NPB
bushi

VP

NPB
bushi

Bush

Bush
Synchronous Generation
Synchronous Generation
Synchronous Generation

[Diagram showing tree structure with labels such as IP, NP, PP, VP, and X1, X2]
Synchronous Generation
Synchronous Generation

bushi

P

NPB

IP

VP

PP

NPB

Bush

yu

with

P

Synchronous Generation

ACL 2010, Uppsala, Sweden
Synchronous Generation

bushi yu
NPB P NPB
IP PP VP VPB

Bush with

NPB shalong
Sharon
Synchronous Generation
Synchronous Generation
Synchronous Generation

ACL 2010, Uppsala, Sweden

bushi yu shalong juxing le

VPB

PP

NPB P NPB VS AS NPB

bushi yu shalong juxing le

Bush held with Sharon

handed

VPB

VS AS X1:NPB

a

X1

held
Synchronous Generation

Bush held with Sharon a huitan talk

bushi yu shalong juxing le

Bush held with

NPB P NPB VS AS NPB

PP VP VPB

NPB

ACL 2010, Uppsala, Sweden
Synchronous Generation

bushi yu shalong juxing le huitan

Bush held a talk with Sharon

NPB P NPB VS AS NPB

huitan

talk
Packed Forest

(Billot and Lang, 1989; Klein and Manning, 2001; Huang and Chiang, 2005)
Rule Extraction

“Bush … with (Sharon) … held … (Sharon) … held ((a) talk)”

Bush

with

Sharon

hsitas
Rule Extraction

“(Bush) … (Sharon)"

“Bush □ with (Sharon)”

“with (Sharon)”

“held … (Sharon)"

“held ((a) talk)”

NPB

bushi

Bush

“Bush”

“with”

“with”

“Sharon”

“held”

“”

“talk”

bushi

yu

shalong

juxing

le

huitan

Bush

held

talk

with

Sharon

7/22/10

ACL 2010, Uppsala, Sweden
Rule Extraction

IP_{0.6}“(Bush) … (Sharon)"

NP_{0.3} “Bush □ with (Sharon)” VP_{1.6} “held … (Sharon)"

PP_{1.3} “with (Sharon)” VPB_{3.6} “held ((a) talk)"

NPB_{0.1} “Bush” CC_{1.2} “with” P_{1.2} “with” NPB_{2.3} “Sharon” VS_{3.4} “held” AS_{4.5} “” NPB_{5.6} “talk”

bushi yu shalong juxing le huitan

Bush with talk with
held

VPB X_{1:VS} AS X_{2:NPB}
le ↓ a
X_{2} X_{1}

7/22/10 ACL 2010, Uppsala, Sweden 23
Rule Extraction

"Bush \( \sqcap \) with (Sharon)" "held \( \sqcap \) (a) talk"

- IP_{0,6}
- NP_{0,3}
- VP_{1,6}
- PP_{1,3}
- VPB_{3,6}
- NPB_{0,1}
- CC_{1,2}
- P_{1,2}
- NPB_{2,3}
- VS_{3,4}
- AS_{4,5}
- NPB_{5,6}

"Bush" "with" "Sharon" "held" "talk"

bushi yu shalong juxing le huitan

Bush talk with

held

ACL 2010, Uppsala, Sweden
Bush held a talk with Sharon

Bush with (Sharon)

held ((a) talk)

(Bush) … (Sharon))

Rule Extraction
Decoding
Decoding
Decoding

bushi  CC  bushi
yu     yu     with
shalong CC  shalong

NPB  CC  NPB
0,1  1,2  2,3

PP  VPB
1,3  3,6

IP  IP
0,6  1,6

NPB  CC  NPB
0,3  1,2  2,3

bushi  CC  bushi
yu     CC  with
shalong CC  shalong

NPB  CC  NPB
0,1  1,2  2,3

PP  VPB
1,3  3,6

IP  IP
0,6  1,6

NPB  CC  NPB
0,3  1,2  2,3

bushi  CC  bushi
yu     CC  with
shalong CC  shalong

NPB  CC  NPB
0,1  1,2  2,3

PP  VPB
1,3  3,6

IP  IP
0,6  1,6

NPB  CC  NPB
0,3  1,2  2,3

NPB  CC  NPB
0,1  1,2  2,3
Decoding

NPB | bushi
↓  Bush

CC | yu
↓  with

NPB | shalong
↓  Sharon

NPB | huitan
↓  talk

NPB₀,₁  CC₁,₂  P₁,₂  NPB₂,₃  VS₃,₄  AS₄,₅  NPB₅,₆  NPB₀,₁  CC₁,₂  NPB₂,₃  NPB₅,₆
Decoding
Evaluation

- Task: Chinese-English
- Training set: FBIS (6.9M+8.9M words)
- LM: 4-gram trained on Gigaword Xinhua
- Development set: NIST 2002
- Test set: NIST 2005
- Metric: case-insensitive BLEU4
## Results

<table>
<thead>
<tr>
<th>System</th>
<th>Rule</th>
<th>Type</th>
<th>Well-formed?</th>
<th>Number</th>
<th>BLEU</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Constituency-to-String</strong></td>
<td>C2S</td>
<td>No</td>
<td>31.9M</td>
<td></td>
<td>34.17</td>
</tr>
<tr>
<td></td>
<td>S2S</td>
<td>No</td>
<td>77.9M</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>C2S</td>
<td>Yes</td>
<td>13.8M</td>
<td></td>
<td>32.48</td>
</tr>
<tr>
<td></td>
<td>S2S</td>
<td>Yes</td>
<td>9.0M</td>
<td></td>
<td>32.48</td>
</tr>
<tr>
<td></td>
<td>C2S</td>
<td>Yes</td>
<td>13.8M</td>
<td></td>
<td>34.03</td>
</tr>
<tr>
<td></td>
<td>S2S</td>
<td>No</td>
<td>77.9M</td>
<td></td>
<td>34.03</td>
</tr>
</tbody>
</table>

**Constituency-to-Dependency**

<table>
<thead>
<tr>
<th>Rule</th>
<th>Type</th>
<th>Well-formed?</th>
<th>Number</th>
<th>BLEU</th>
</tr>
</thead>
<tbody>
<tr>
<td>C2D</td>
<td>Yes</td>
<td>13.8M</td>
<td></td>
<td>33.25</td>
</tr>
<tr>
<td>S2D</td>
<td>Yes</td>
<td>9.0M</td>
<td></td>
<td>33.25</td>
</tr>
<tr>
<td>C2D</td>
<td>Yes</td>
<td>13.8M</td>
<td></td>
<td>34.88</td>
</tr>
<tr>
<td>S2D</td>
<td>No</td>
<td>77.9M</td>
<td></td>
<td>34.88</td>
</tr>
</tbody>
</table>
Conclusion and Future Work

- Constituency-to-dependency combines the advantages of constituency-to-string and string-to-dependency:
  - Decoding is fast
  - Better target side generation

- Future work
  - Larger data and more language pairs
  - Comparison with constituency-to-constituency
  - Dependency forest
Thank you!
Improving Rule Coverage

bushi yu shalong juxing le huitan

NPB0,1 CC1,2 P1,2 NPB2,3 VS3,4 AS4,5 NPB5,6

IP0,6 VP1,6 VPB3,6

PP1,3

NP0,3

VPB

juxing le huitan

held a talk

ACL 2010, Uppsala, Sweden