Robust Language Pair Independent Sub-tree Alignment

John Tinsley & Ventsislav Zhechev
NCLT Seminar Series 09/05/07
Overview

- Motivation
- Related Work
- Sub-Tree Alignment Algorithm
- Evaluation Setup
- Experiments & Results
- Future Work
Motivation
Motivation

• What are we proposing?
  - Automatically align trees
Motivation

• What are we proposing?
  – Automatically align trees

• Why do we need it?
  – We need it for Data-Oriented Translation (DOT) but...
Motivation

• What are we proposing?
  − Automatically align trees

• Why do we need it?
  − We need it for Data-Oriented Translation (DOT) but...

• What other uses has it?
  − What can you do if you have sub-sententially aligned phrase-structure tree pairs?
Motivation

• What are we proposing?
  − Automatically align trees

• Why do we need it?
  − We need it for Data-Oriented Translation (DOT) but...

• What other uses has it?
  − What can you do if you have sub-sententially aligned phrase-structure tree pairs?

• Has something like this not been done before?
  − ...
Related Work
Related Work

• What has been done
  – Alignment of ps-trees, dependency structures
Related Work

• What has been done
  - Alignment of ps-trees, dependency structures

• What they did
  - isomorphic alignment, simple sentences, word-alignment
Related Work

• What has been done
  – Alignment of ps-trees, dependency structures

• What they did
  – isomorphic alignment, simple sentences, word-alignment

• What they were aiming for
  – resolve parse ambiguity, extract transfer templates
Related Work

• What has been done
  - Alignment of ps-trees, dependency structures

• What they did
  - isomorphic alignment, simple sentences, word-alignment

• What they were aiming for
  - resolve parse ambiguity, extract transfer templates

• Why it wasn't suitable for our purposes?
  - altered trees, language-pair specific, many external resources
Related Work

• Matsumoto et al. (1993) – aligned dependencies to resolve parse ambiguities

• Lu et al. (2001) – imposed source tree structure on target for extraction of transfer templates

• Eisner (2003) – tree mapping for dependencies as part of translation

• Gildea (2003) – tree structures altered to impose isomorphism

• Groves et al. (2004) – rule based approach which is language specific
Sub-tree Alignment Algorithm
Alignment principles
Alignment principles

• independence with respect to language pair and constituent labelling schema
Alignment principles

- independence with respect to language pair and constituent labelling schema
- preservation of the given tree structures
Alignment principles

- independence with respect to language pair and constituent labelling schema
- preservation of the given tree structures
- minimal external resources required
Well-formedness criteria for the alignments
Well-formedness criteria for the alignments

- a node can only be linked once
Well-formedness criteria for the alignments

- A node can only be linked once.
- Descendants of a source linked node may only link to descendants of its target linked counterpart.
Well-formedness criteria for the alignments

- a node can only be linked once
- descendants of a source linked node may only link to descendants of its target linked counterpart
- ancestors of a source linked node may only link to ancestors of its target linked counterpart
Basic algorithm
Basic algorithm

- Initialisation
Basic algorithm

• Initialisation

for each source non-terminal s do
Basic algorithm

• Initialisation

\[\text{for each source non-terminal } s \text{ do}\]
\[\text{for each target non-terminal } t \text{ do}\]
Basic algorithm

• Initialisation

\[
\text{for each source non-terminal } s \text{ do} \\
\quad \text{for each target non-terminal } t \text{ do} \\
\quad \quad \text{generate scored hypothesis } \gamma(\langle s, t \rangle) \\
\quad \text{end for} \\
\text{end for}
\]
Basic algorithm

• Initialisation
  
  for each source non-terminal $s$ do
    for each target non-terminal $t$ do
      generate scored hypothesis $\gamma(\langle s, t \rangle)$
    end for
  end for

  block all zero-scored hypotheses
make sure the scanner is connected to the HomeCentre
make sure the scanner is connected to the HomeCentre.
make sure the scanner is connected to the HomeCentre.
Basic algorithm

- **Initialisation**
  
  for each source non-terminal \( s \) do
  for each target non-terminal \( t \) do
    generate scored hypothesis \( \gamma(\langle s, t \rangle) \)
  end for
  end for

  block all zero-scored hypotheses
Basic algorithm

- **Initialisation**
  
  ```
  for each source non-terminal s do
    for each target non-terminal t do
      generate scored hypothesis $\gamma(\langle s, t \rangle)$
    end for
  end for
  block all zero-scored hypotheses
  ```

- **Selection**
Basic algorithm

• Initialisation
  
  for each source non-terminal \( s \) do
  
  for each target non-terminal \( t \) do
  
  generate scored hypothesis \( \gamma(\langle s, t \rangle) \)
  
  end for
  
  end for

  block all zero-scored hypotheses

• Selection

  while non-blocked hypotheses remain do
Basic algorithm

• **Initialisation**

  
  ```
  for each source non-terminal s do
    for each target non-terminal t do
      generate scored hypothesis \( \gamma(s, t) \)
    end for
  end for
  ```

  block all zero-scored hypotheses

• **Selection**

  ```
  while non-blocked hypotheses remain do
    link and block the highest-scoring hypothesis
  ```
Basic algorithm

• **Initialisation**
  
  for each source non-terminal $s$ do
  
  for each target non-terminal $t$ do
    
    generate scored hypothesis $\gamma(\langle s, t \rangle)$
  
  end for
  
  end for
  
  block all zero-scored hypotheses
  
• **Selection**
  
  while non-blocked hypotheses remain do
  
    link and block the highest-scoring hypothesis
    
    block all contradicting hypotheses
  
  end while
make the scanner is connected to the HomeCentre
make sure the scanner is connected to the HomeCentre

make sure le scanner est connectée à le HomeCentre
make sure the scanner is connected to the HomeCentre
D-17 and D-13
1.272599798786210317e-05
D-17 and D-4
1.272599798786210317e-05
D-8 and D-13
1.272599798786210315e-05
D-8 and D-4
1.272599798786210315e-05
NP-7 and NPdet-3
1.22703015051e-05
NP-16 and NPdet-12
1.18886738508e-05
PP-14 and PPcase-10
1.00410744905e-05
D-17 and D-13
1.2725599798786210317e-05

D-17 and D-4
1.2725599798786210317e-05

D-8 and D-13
1.2725599798786210315e-05

D-8 and D-4
1.2725599798786210315e-05

NP-7 and NPdet-3
1.22703015051e-05

NP-16 and NPdet-12
1.18886738508e-05

PP-14 and PPcase-10
1.00410744905e-05
Selection
Hypotheses with equal scores
Selection
Hypotheses with equal scores

• skip1
Selection
Hypotheses with equal scores

• $skip_1$

  while at least one non-blocked hypothesis with no tied competitors remains do
Selection
Hypotheses with equal scores

• \textit{skip1}

\textbf{while} at least one non-blocked hypothesis with no tied competitors remains \textbf{do}
\textbf{while} the highest-scoring hypothesis has tied competitors \textbf{do}
Selection
Hypotheses with equal scores

- skip1
  while at least one non-blocked hypothesis with no tied competitors remains do
  while the highest-scoring hypothesis has tied competitors do
    skip
  end while
Selection
Hypotheses with equal scores

• skip

while at least one non-blocked hypothesis with no tied competitors remains do

while the highest-scoring hypothesis has tied competitors do

skip

end while

link and block the highest-scoring not-skipped hypothesis
Selection
Hypotheses with equal scores

• skip

  while at least one non-blocked hypothesis with no tied competitors remains do
  while the highest-scoring hypothesis has tied competitors do
    skip
  end while
  link and block the highest-scoring not-skipped hypothesis
  block all contradicting hypothesis
• skip1

  while at least one non-blocked hypothesis with no tied competitors remains do
  while the highest-scoring hypothesis has tied competitors do
      skip
  end while
  link and block the highest-scoring not-skipped hypothesis
  block all contradicting hypothesis
  re-enable all non-blocked skipped hypotheses
end while
1.2725599798786210317e-05

D-17 and D-13
D-18 and D-4
D-8 and D-13
D-8 and D-4
NP-7 and NPdet-3
NP-16 and NPdet-12
PP-14 and PPcase-10
1.18886738508e-05
1.18886738508e-05
1.00410744905e-05
1.00410744905e-05
1.22703015051e-05
1.22703015051e-05
D-17 and D-13
1.2725599798786210317e-05
D-17 and D-4
1.2725599798786210315e-05
D-8 and D-13
1.2725599798786210315e-05
D-8 and D-4
1.2725599798786210315e-05
NP-7 and NPdet-3
1.22703015051e-05
NP-16 and NPdet-12
1.1888638508e-05
PP-14 and PPcase-10
1.00410744905e-05
make sure the scanner is connected to the HomeCentre.
1. D-17 and D-13
   1.27255997987921607e-05
2. D-8 and D-4
   1.272559979876210315e-05
3. NP-16 and NPdet-12
   1.18886738508e-05
4. PP-14 and PPcase-10
   1.00410744905e-05

1. The scanner is connected to the HomeCentre.
2. The scanner is est connecté à le HomeCentre.
3. D-17 and D-13
4. D-8 and D-4
5. NP-16 and NPdet-12
6. PP-14 and PPcase-10
7. D-17 and D-13
8. N-14

The numbers 1 to 19 are placed in the table on the page.
make sure the scanner is connected to the HomeCentre
make sure the scanner is connected to the HomeCentre
Selection
Hypotheses with equal scores
Selection
Hypotheses with equal scores

• $\text{skip}_2$
V-2 and P-14
8.29702211770749522363e-05

V-2 and P-5
8.29702211770749522363e-05

NX-19 and P-5
8.1330915823403189978e-05

NX-19 and P-14
8.1330915823403189911e-05

A-14 and N-23
5.4622895082041485328e-05

...
V-2 and P-14
8.29702211770749522363e-05

V-2 and P-5
8.29702211770749522363e-05

NX-19 and P-5
8.13309158823403189978e-05

NX-19 and P-14
8.13309158823403189911e-05

A-14 and N-23
5.46228915082041485328e-05

...
Selection
Hypotheses with equal scores
Selection

Hypotheses with equal scores

- \textit{skip2}
Selection
Hypotheses with equal scores

• $skip^2$
  
  while at least one non-blocked hypothesis with no tied competitors remains do
Selection

Hypotheses with equal scores

- \textit{skip2}

\begin{verbatim}
while at least one non-blocked hypothesis with no tied competitors remains do
  if the highest-scoring hypothesis has tied competitors then

\end{verbatim}
Selection

Hypotheses with equal scores

• \textit{skip2}

\begin{verbatim}
while at least one non-blocked hypothesis with no tied competitors remains do
    if the highest-scoring hypothesis has tied competitors then
        mark the constituents of all competitors as \textit{skipped}
    end if
\end{verbatim}
Selection
Hypotheses with equal scores

• \textit{skip2}

\begin{verbatim}
while at least one non-blocked hypothesis with no tied competitors remains do
  if the highest-scoring hypothesis has tied competitors then
    mark the constituents of all competitors as \textit{skipped}
  end if
while the highest-scoring hypothesis has a skipped constituent do
\end{verbatim}
Selection
Hypotheses with equal scores

• skip2
  
  while at least one non-blocked hypothesis with no tied competitors remains do
    if the highest-scoring hypothesis has tied competitors then
      mark the constituents of all competitors as skipped
    end if
  while the highest-scoring hypothesis has a skipped constituent do
    skip
  end while
Selection
Hypotheses with equal scores

- skip2

while at least one non-blocked hypothesis with no tied competitors remains do
  if the highest-scoring hypothesis has tied competitors then
    mark the constituents of all competitors as skipped
  end if
while the highest-scoring hypothesis has a skipped constituent do
  skip
end while
link and block the highest-scoring not-skipped hypothesis
Selection
Hypotheses with equal scores

• skip2

while at least one non-blocked hypothesis with no tied competitors remains do
if the highest-scoring hypothesis has tied competitors then
mark the constituents of all competitors as skipped
end if
while the highest-scoring hypothesis has a skipped constituent do
skip
end while
link and block the highest-scoring not-skipped hypothesis
block all contradicting hypotheses
Selection
Hypotheses with equal scores

- skip\textsubscript{2}

\begin{verbatim}
while at least one non-blocked hypothesis with no tied competitors remains do
  if the highest-scoring hypothesis has tied competitors then
    mark the constituents of all competitors as skipped
  end if
  while the highest-scoring hypothesis has a skipped constituent do
    skip
  end while
  link and block the highest-scoring not-skipped hypothesis
  block all contradicting hypotheses
  re-enable all non-blocked skipped hypotheses
end while
\end{verbatim}
V-2 and P-14
8.29702211770749522363e-05

V-2 and P-5
8.29702211770749522363e-05

NX-19 and P-5
8.13309158823403189978e-05

NX-19 and P-14
8.13309158823403189911e-05

A-14 and N-23
5.46228915082041485328e-05

...
V-2 and P-14
8.29702211770749522363e-05

V-2 and P-5
8.29702211770749522363e-05

NX-19 and P-5
8.13309158823403189911e-05

NX-19 and P-14
8.13309158823403189911e-05

A-14 and N-23
5.46228915082041486326e-05

...
Selection
Selection

- span
Selection

- `span` while non-blocked non-lexical hypotheses remain do
Selection

- \textit{spani} \\
  \textbf{while} non-blocked non-lexical hypotheses remain \textbf{do} \\
  link and block the highest-scoring hypothesis
Selection

- \textit{span} \\
  \textbf{while} non-blocked \textit{non-lexical} hypotheses remain \textbf{do} \\
  link and block the highest-scoring hypothesis \\
  block all contradicting hypotheses
Selection

- \textit{spani}

while non-blocked non-lexical hypotheses remain do
link and block the highest-scoring hypothesis
block all contradicting hypotheses
if no non-blocked non-lexical hypotheses remain then
Selection

- *spani*
  
  while non-blocked *non-lexical* hypotheses remain do
  
  link and block the highest-scoring hypothesis
  
  block all contradicting hypotheses
  
  if no non-blocked *non-lexical* hypotheses remain then
  
  while non-blocked *lexical* hypotheses remain do
Selection

- \textit{spani}
  
  \textbf{while} non-blocked \textit{non-lexical} hypotheses remain \textbf{do}
  
  link and block the highest-scoring hypothesis
  
  block all contradicting hypotheses
  
  \textbf{if} no non-blocked \textit{non-lexical} hypotheses remain \textbf{then}
  
  \textbf{while} non-blocked \textit{lexical} hypotheses remain \textbf{do}
    
    link and block the highest-scoring hypothesis
Selection

- \textit{spani}
  
  \begin{verbatim}
  while non-blocked non-lexical hypotheses remain do
  link and block the highest-scoring hypothesis
  block all contradicting hypotheses
  \end{verbatim}

  \begin{verbatim}
  if no non-blocked non-lexical hypotheses remain then
    \begin{verbatim}
    while non-blocked lexical hypotheses remain do
      link and block the highest-scoring hypothesis
      block all contradicting hypotheses
    \end{verbatim}
  \end{verbatim}

  \begin{verbatim}
  end while
  \end{verbatim}

  \begin{verbatim}
  end if
  \end{verbatim}

  \begin{verbatim}
  end while
  \end{verbatim}
without \textit{spani}

\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|}
\hline
\hline
\end{tabular}

\ldots

\begin{align*}
V-15 & \text{ and ADV-23} \\
& 4.6881584514e-11 \\
V_Paux-20 & \text{ and V_Paux[pass]-20} \\
& 2.4674517256e-11 \\
\ldots
\end{align*}
with \textit{span1}
with

\( \textit{span1} \)
Computing the scores
Computing the scores

• A link between two nodes in translationally equivalent trees indicates that:
Computing the scores

- A link between two nodes in translationally equivalent trees indicates that:
  
  (i) the substrings dominated by those nodes are translationally equivalent
Computing the scores

• A link between two nodes in translationally equivalent trees indicates that:
  (i) the substrings dominated by those nodes are translationally equivalent
  (ii) all meaning carried by the remainder of the source sentence is encapsulated in the remainder of the target sentence and vice versa
• Given a tree pair $\langle S, T \rangle$ and hypothesis $\langle s, t \rangle$, we compute the following strings:
Given a tree pair $\langle S, T \rangle$ and hypothesis $\langle s, t \rangle$, we compute the following strings:

\[
\begin{align*}
sl &= s_i \ldots s_{ix} \\
\bar{s}_l &= S_1 \ldots s_{i-1}s_{ix+1} \ldots S_m \\
t_l &= t_j \ldots t_{jy} \\
\bar{t}_l &= T_1 \ldots t_{j-1}t_{jy+1} \ldots T_n
\end{align*}
\]
• Given a tree pair \( \langle S, T \rangle \) and hypothesis \( \langle s, t \rangle \), we compute the following strings:

\[
\begin{align*}
  s_l &= s_i \ldots s_{ix} & \bar{s}_l &= S_i \ldots s_{i-1}s_{ix+1} \ldots S_m \\
  t_l &= t_j \ldots t_{jy} & \bar{t}_l &= T_i \ldots t_{j-1}t_{jy+1} \ldots T_n \\
\end{align*}
\]

\( s_i \ldots s_{ix} \) and \( t_j \ldots t_{jy} \) denote the terminal sequences dominated by \( s \) and \( t \) respectively.
Given a tree pair \( \langle S, T \rangle \) and hypothesis \( \langle s, t \rangle \), we compute the following strings:

\[
\begin{align*}
sl &= s_1...s_{ix} \\
\overline{s_l} &= S_1...s_{i-1}s_{ix+1}...S_m \\
t_l &= t_j...t_{jy} \\
\overline{t_l} &= T_1...t_{j-1}t_{jy+1}...T_n
\end{align*}
\]

\( s_i...s_{ix} \) and \( t_j...t_{jy} \) denote the terminal sequences dominated by \( s \) and \( t \) respectively.

\( S_1...S_m \) and \( T_1...T_n \) denote the terminal sequences dominated by \( S \) and \( T \) respectively.
• Given a tree pair $\langle S, T \rangle$ and hypothesis $\langle s, t \rangle$, we compute the following strings:

$$s_l = s_i...s_{ix} \quad \bar{s}_l = S_1...s_{i-1}s_{ix+1}...S_m$$

$$t_l = t_j...t_{jy} \quad \bar{t}_l = T_1...t_{j-1}t_{jy+1}...T_n$$

$s_i...s_{ix}$ and $t_j...t_{jy}$ denote the terminal sequences dominated by $s$ and $t$ respectively.

$S_1...S_m$ and $T_1...T_n$ denote the terminal sequences dominated by $S$ and $T$ respectively.

• The score for a given hypothesis $\langle s, t \rangle$ is computed according to the formula:

$$\gamma(\langle s, t \rangle) = \alpha(s_l|t_l)\cdot \alpha(t_l|s_l)\cdot \alpha(s_l|\bar{t}_l)\cdot \alpha(\bar{t}_l|s_l)$$
Given a tree pair \( \langle S, T \rangle \) and hypothesis \( \langle s, t \rangle \), we compute the following strings:

\[
\begin{align*}
sl &= s_i...s_ix \\
\bar{s}_l &= S_{i-1}s_{ix+1}...S_m \\
trl &= t_j...t_{jy} \\
\bar{t}_l &= T_{j-1}t_{jy+1}...T_n
\end{align*}
\]

\( s_i...s_ix \) and \( t_j...t_{jy} \) denote the terminal sequences dominated by \( s \) and \( t \) respectively.

\( S_{i-1}...S_m \) and \( T_{j-1}...T_n \) denote the terminal sequences dominated by \( S \) and \( T \) respectively.

The score for a given hypothesis \( \langle s, t \rangle \) is computed according to the formula:

\[
\gamma(\langle s, t \rangle) = \alpha(s_l|t_l) \cdot \alpha(t_l|s_l) \cdot \alpha(\bar{s}_l|\bar{t}_l) \cdot \alpha(\bar{t}_l|\bar{s}_l)
\]

the individual string correspondence scores \( \alpha(x|y) \) are computed using word-alignment probabilities derived by training the Moses system in both language directions.
Computing the scores
Two scoring functions
Computing the scores
Two scoring functions

\[ \alpha(x|y) = \prod_j \frac{\sum_i P(x_i|y_j)}{\sum_j P(x_i|y_j)} \]

**score1**
\[ \alpha(x|y) = \prod_i \frac{\sum_j P(x_i|y_j)}{|y|} \]

**score2**
Possible aligner configurations
Possible aligner configurations

<table>
<thead>
<tr>
<th>skip1_score1</th>
<th>skip1_score1_span1</th>
</tr>
</thead>
<tbody>
<tr>
<td>skip1_score2</td>
<td>skip1_score2_span1</td>
</tr>
<tr>
<td>skip2_score1</td>
<td>skip2_score1_span1</td>
</tr>
<tr>
<td>skip2_score2</td>
<td>skip2_score2_span1</td>
</tr>
</tbody>
</table>
Evaluation setup
• Intrinsic and extrinsic evaluation
• Intrinsic and extrinsic evaluation

• English – French section of the HomeCentre corpus:
• Intrinsic and extrinsic evaluation

• English – French section of the HomeCentre corpus:
  • Xerox printer manual
• Intrinsic and extrinsic evaluation

• English – French section of the HomeCentre corpus:
  • Xerox printer manual
  • translated by professional translators, sentence-aligned and annotated at Xerox PARC
• Intrinsic and extrinsic evaluation

• English – French section of the HomeCentre corpus:
  • Xerox printer manual
  • translated by professional translators, sentence-aligned and annotated at Xerox PARC
  • 810 parsed, sentence-aligned translation pairs
• Intrinsic and extrinsic evaluation

• English – French section of the HomeCentre corpus:
  • Xerox printer manual
  • translated by professional translators, sentence-aligned and annotated at Xerox PARC
  • 810 parsed, sentence-aligned translation pairs
  • manually aligned at sub-sentential level
• Intrinsic and extrinsic evaluation

• English – French section of the HomeCentre corpus:
  • Xerox printer manual
  • translated by professional translators, sentence-aligned and annotated at Xerox PARC
  • 810 parsed, sentence-aligned translation pairs
  • manually aligned at sub-sentential level

• We evaluate all 8 versions of the aligner
Intrinsic evaluation setup
Intrinsic evaluation setup

• Calculate the precision and recall of the automatically induces alignments, using the existing manual alignments as a gold standard
Intrinsic evaluation setup

• Calculate the precision and recall of the automatically induces alignments, using the existing manual alignments as a gold standard
  
  • the calculations are also done for the non-lexical links only – in order to better assess the quality of the alignments
Intrinsic evaluation setup

- Calculate the precision and recall of the automatically induces alignments, using the existing manual alignments as a gold standard
  - the calculations are also done for the non-lexical links only – in order to better assess the quality of the alignments
- Given a tree pair $T$, its automatically-aligned version $T_A$ and its manually-aligned version $T_M$: 
Intrinsic evaluation setup

- Calculate the precision and recall of the automatically induces alignments, using the existing manual alignments as a gold standard.
  - the calculations are also done for the non-lexical links only – in order to better assess the quality of the alignments.

- Given a tree pair $T$, its automatically-aligned version $T_A$ and its manually-aligned version $T_M$:

\[
\text{Precision} = \frac{|T_A \cap T_M|}{|T_A|}
\]

\[
\text{Recall} = \frac{|T_M \cap T_A|}{|T_M|}
\]
Extrinsic evaluation setup
Extrinsic evaluation setup

- Supply the automatically aligned data as training for a Data-Oriented Translation system
Extrinsic evaluation setup

- Supply the automatically aligned data as training for a Data-Oriented Translation system
- Translate several training/test data splits
Extrinsic evaluation setup

- Supply the automatically aligned data as training for a Data-Oriented Translation system
- Translate several training/test data splits
  - 80 test sentences in each split
Extrinsic evaluation setup

- Supply the automatically aligned data as training for a Data-Oriented Translation system
- Translate several training/test data splits
  - 80 test sentences in each split
  - translate in both language directions
Extrinsic evaluation setup

- Supply the automatically aligned data as training for a Data-Oriented Translation system
- Translate several training/test data splits
  - 80 test sentences in each split
  - translate in both language directions
  - evaluate all aligner configurations
Extrinsic evaluation setup

• Supply the automatically aligned data as training for a Data-Oriented Translation system

• Translate several training/test data splits
  • 80 test sentences in each split
  • translate in both language directions
  • evaluate all aligner configurations

• Evaluate the quality of the translations using BLEU, NIST and METEOR and
Extrinsic evaluation setup

- Supply the automatically aligned data as training for a Data-Oriented Translation system
- Translate several training/test data splits
  - 80 test sentences in each split
  - translate in both language directions
  - evaluate all aligner configurations
- Evaluate the quality of the translations using BLEU, NIST and METEOR and
- Compare the results with translations obtained using the manual alignments
Results

- Evaluation of the precision and recall of the induced alignments, over 8 configurations, with manual alignments as gold standard

<table>
<thead>
<tr>
<th>Configurations</th>
<th>all links</th>
<th>non-lexical links</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>precision</td>
<td>recall</td>
</tr>
<tr>
<td><code>skip1_score1</code></td>
<td>0.6045</td>
<td>0.7696</td>
</tr>
<tr>
<td><code>skip1_score2</code></td>
<td>0.6152</td>
<td>0.7842</td>
</tr>
<tr>
<td><code>skip1_score1_span1</code></td>
<td>0.6163</td>
<td><strong>0.8026</strong></td>
</tr>
<tr>
<td><code>skip1_score2_span1</code></td>
<td>0.6193</td>
<td>0.7926</td>
</tr>
<tr>
<td><code>skip2_score1</code></td>
<td>0.6103</td>
<td>0.7729</td>
</tr>
<tr>
<td><code>skip2_score2</code></td>
<td>0.6180</td>
<td>0.7840</td>
</tr>
<tr>
<td><code>skip2_score1_span1</code></td>
<td>0.6196</td>
<td><strong>0.8022</strong></td>
</tr>
<tr>
<td><code>skip2_score2_span1</code></td>
<td><strong>0.6219</strong></td>
<td>0.7923</td>
</tr>
</tbody>
</table>
Discussion
Discussion

• **All links** – *higher recall* – we're inducing more links per tree pair than the manual alignments
Discussion

- **All links** – *higher recall* – we're inducing more links per tree pair than the manual alignments

- **Internal links** – *higher precision* – our phrase alignments are accurate, but conversely, our lexical alignment are inconsistent
Discussion

• **All links** – *higher recall* – we're inducing more links per tree pair than the manual alignments

• **Internal links** – *higher precision* – our phrase alignments are accurate, but conversely, our lexical alignment are inconsistent

• **span1** makes most (positive) difference
Results

- Evaluation of the DOT system using the 8 aligner configurations, and the manual alignments

<table>
<thead>
<tr>
<th>Configurations</th>
<th>BLEU</th>
<th>NIST</th>
<th>METEOR</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>manual</td>
<td>0.5444</td>
<td>7.0701</td>
<td>0.7302</td>
<td>70.4167</td>
</tr>
<tr>
<td>skip1_score1</td>
<td>0.5173</td>
<td>6.8577</td>
<td>0.7204</td>
<td>74.0625</td>
</tr>
<tr>
<td>skip1_score2</td>
<td>0.5254</td>
<td>6.8564</td>
<td>0.7280</td>
<td>73.5417</td>
</tr>
<tr>
<td>skip1_score1_span1</td>
<td>0.5243</td>
<td>6.8794</td>
<td>0.7295</td>
<td>75.3125</td>
</tr>
<tr>
<td>skip1_score2_span1</td>
<td>0.5249</td>
<td>6.8550</td>
<td>0.7244</td>
<td>73.6458</td>
</tr>
<tr>
<td>skip2_score1</td>
<td>0.5163</td>
<td>6.8685</td>
<td>0.7242</td>
<td>73.7500</td>
</tr>
<tr>
<td>skip2_score2</td>
<td>0.5280</td>
<td>6.8869</td>
<td>0.7297</td>
<td>73.5417</td>
</tr>
<tr>
<td>skip2_score1_span1</td>
<td>0.5252</td>
<td>6.8842</td>
<td>0.7285</td>
<td>75.3125</td>
</tr>
<tr>
<td>skip2_score2_span1</td>
<td>0.5258</td>
<td>6.8506</td>
<td>0.7268</td>
<td>73.6458</td>
</tr>
</tbody>
</table>
Discussion
Discussion

• Translation scores are very competitive wrt. the manual alignment scores, and always get better coverage (and more recent results have improved further)
Discussion

• Translation scores are very competitive wrt. the manual alignment scores, and always get better coverage (and more recent results have improved further)

• \textit{span1} makes most (positive) difference to translation quality
Discussion

• Translation scores are very competitive wrt. the manual alignment scores, and always get better coverage (and more recent results have improved further)

• \texttt{span1} makes most (positive) difference to translation quality

• No clear preference for other configurations, but \texttt{skip2} is more intuitive than \texttt{skip1} (and more recent results favour \texttt{score2})
Discussion

- Translation scores are very competitive wrt. the manual alignment scores, and always get better coverage (and more recent results have improved further)

- span1 makes most (positive) difference to translation quality

- No clear preference for other configurations, but skip2 is more intuitive than skip1 (and more recent results favour score2)

- The key to improvement is better lexical links, and there are a few ways to improve these…
Future Work
Future Work

- More data
- Different language pairs
Future Work

- More data
- Different language pairs
- Incorporation of Moses phrases
Future Work

- More data
- Different language pairs
- Incorporation of Moses phrases
- Use bilingual dictionary/lexicon to fix on things
- Investigate function word importance (risk loss of independence)
Future Work

- More data
- Different language pairs
- Incorporation of Moses phrases
- Use bilingual dictionary/lexicon to fix on things
- Investigate function word importance (risk loss of independence)
- Better (proper) probabilistic model
- Different search strategy
- Test suite of different applications
  - resolution of parse ambiguity
  - transfer templates
- String-to-Tree / Tree-to-String / String-to-String, how important are the trees?
References


Thank you for listening
¿Questions?