Multi-Engine Machine Translation by Recursive Sentence Decomposition

Bart Mellebeek  Karolina Owczarzak  Josef van Genabith  Andy Way

National Centre for Language Technology
School of Computing
Dublin City University

7th Biennial Conference of the Association for Machine Translation in the Americas
Outline

What is Multi-Engine Machine Translation?

MEMT with Recursive Sentence Decomposition.

Experiments, Results and Analysis.
What is Multi-Engine Machine Translation?

MEMT with Recursive Sentence Decomposition.

Experiments, Results and Analysis.
Multi-Engine Machine Translation (MEMT)

- Use several MT engines for same input.
- Combine outputs into consensus translation.
- Why? → ‘Three heads are better than one’ (Frederking and Nirenburg, 1994)
- Errors committed by a system are independent of errors committed by other systems.
- Other domains: Speech Recognition, Text Categorization, POS tagging.
- The units for comparison are not defined a priori for MT.
Different Approaches to MEMT

Previous Approaches

- Translate entire input sentence → individual MT system does not improve.
- Find consensus through output alignment.

Our Approach

- Translate chunks in context → individual MT systems can improve.
- Prepare input for processing.

References

Bangalore et al., IEEE 2001.
Nomoto, ACL 2004.
Jayaraman and Lavie, EAMT 2005.
van Zaanen and Somers, MT Summit 2005
Matusov et al., EACL 2006
Different Approaches to MEMT

Previous Approaches

- Translate entire input sentence → individual MT system does not improve.
- Find consensus through output alignment.

Our Approach

- Translate chunks in context → individual MT systems can improve.
- Prepare input for processing.

References

- Bangalore et al., IEEE 2001.
- Nomoto, ACL 2004.
- Jayaraman and Lavie, EAMT 2005.
- van Zaanen and Somers, MT Summit 2005
- Matusov et al., EACL 2006

What do we have in common?

- Majority voting.
- Language models.
- Confidence score.
Outline

What is Multi-Engine Machine Translation?

MEMT with Recursive Sentence Decomposition.

Experiments, Results and Analysis.
1. Decomposition.
2. Translation.
3. Selection.
4. Composition.
Decomposition of Input into Optimal Chunks

- Input = syntactic parse of string.
- Decompose into pivot and satellites.
- Pivot = nucleus (+ additional material).
- Satellite = argument/adjunct chunk

\[
\text{Chunk} \quad \rightarrow \quad SAT_1 \quad \ldots \quad SAT_i \quad pivot \quad SAT_{i+1} \quad \ldots \quad SAT_{i+r}
\]
The chairman, a long-time rival of Bill Gates, likes fast and confidential deals.
Example of decomposition

[The chairman, a long-time rival of Bill Gates,]$_{ARG_1}$ [likes]$_{pivot}$ [fast and confidential deals]$_{ARG_2}$. 
Skeletons and Substitution Variables

\[
\begin{align*}
[SAT_1] & \ldots [SAT_i] \text{ pivot } [SAT_{i+1}] & \ldots & [SAT_{i+r}] \\
\downarrow & \\
[V_{SAT_1}] & \ldots [V_{SAT_i}] \text{ pivot } [V_{SAT_{i+1}}] & \ldots & [V_{SAT_{i+r}}]
\end{align*}
\]

- \( V_{SAT_i} \) = a simpler string substituting \( SAT_i \)
- Reduce complexity original constituents \(\rightarrow\) better translation pivot.
- Track location satellites in target.
- Substitution Variables: static vs. dynamic

Example

[The chairman, a long-time rival of Bill Gates,]_{SAT_1} \text{ [likes]}_{pivot} \text{ [fast and confidential deals]}_{SAT_2}.
Skeletons and Substitution Variables

\[
[SAT_1] \ldots [SAT_i] \text{ pivot } [SAT_{i+1}] \ldots [SAT_{i+r}]
\]
\[
[\text{V}_{SAT_1}] \ldots [\text{V}_{SAT_i}] \text{ pivot } [\text{V}_{SAT_{i+1}}] \ldots [\text{V}_{SAT_{i+r}}]
\]

- \(\text{V}_{SAT_i}\) = a simpler string substituting \(SAT_i\)
- Reduce complexity original constituents \(\rightarrow\) better translation pivot.
- Track location satellites in target.
- Substitution Variables: static vs. dynamic

Example

[The chairman, a long-time rival of Bill Gates,.]_{SAT_1} [\text{likes}]_{pivot} [\text{fast and confidential deals}]_{SAT_2}.
Skeletons and Substitution Variables

\[ [SAT_1] \ldots [SAT_i] \text{ pivot } [SAT_{i+1}] \ldots [SAT_{i+r}] \]
\[ \downarrow \]
\[ [V_{SAT_1}] \ldots [V_{SAT_i}] \text{ pivot } [V_{SAT_{i+1}}] \ldots [V_{SAT_{i+r}}] \]

- \( V_{SAT_i} \) = a simpler string substituting \( SAT_i \)
- Reduce complexity original constituents \( \rightarrow \) better translation pivot.
- Track location satellites in target.
- Substitution Variables: static vs. \textbf{dynamic}

\textbf{Example}
\[ [\text{The chairman}]_{V_{SAT_1}} [\text{likes}]_{pivot} [\text{deals}]_{V_{SAT_2}}. \]
Skeletons and Substitution Variables

\[
[SAT_1] \ldots [SAT_i] \text{ pivot } [SAT_{i+1}] \ldots [SAT_{i+r}]
\]

\[
[V_{SAT_1}] \ldots [V_{SAT_i}] \text{ pivot } [V_{SAT_{i+1}}] \ldots [V_{SAT_{i+r}}]
\]

- \( V_{SAT_i} \) = a simpler string substituting \( SAT_i \)
- Reduce complexity original constituents → better translation pivot.
- Track location satellites in target.
- Substitution Variables: static vs. dynamic

Example

\[ [\text{The boy}]_{V_{SAT_1}} [\text{likes}]_{pivot} [\text{books}]_{V_{SAT_2}} \]
Translation of Input Chunks

- If complexity $SAT_i$ not OK → recursion.
- If complexity $SAT_i$ OK → translate.
- Embed $SAT_i$ in context for optimal translation.
- Context: static vs. dynamic.

Example

[The chairman, a long-time rival of Bill Gates,]$_{SAT_1}$ [likes]$_{pivot}$ [fast and confidential deals]$_{SAT_2}$.
Translation of Input Chunks

- If complexity $SAT_i$ not OK $\rightarrow$ recursion.
- If complexity $SAT_i$ OK $\rightarrow$ translate.
- Embed $SAT_i$ in context for optimal translation.
- Context: static vs. dynamic.

Example

[The chairman, a long-time rival of Bill Gates,]$_{SAT_1}$ [likes]$_{pivot}$ [fast and confidential deals]$_{SAT_2}$.
Translation of Input Chunks

- If complexity $SAT_i$ not OK $\rightarrow$ recursion.
- If complexity $SAT_i$ OK $\rightarrow$ translate.
- Embed $SAT_i$ in context for optimal translation.
- Context: static vs. dynamic.

Example

[The chairman likes]_{Context} [fast and confidential deals]_{SAT_2}.
Translation of Input Chunks

- If complexity $SAT_i$ not OK $\rightarrow$ recursion.
- If complexity $SAT_i$ OK $\rightarrow$ translate.
- Embed $SAT_i$ in context for optimal translation.
- Context: static vs. dynamic.

Example

$[\text{The boy sees}]_{Context} [\text{fast and confidential deals}]_{SAT_2}$. 
The selection of the best translation $C_i^{best}$ for each input chunk $C_i$ is based on 3 heuristics:

1. Majority Voting: identical translations are better.
2. Language Modeling: 213M words, trigram model with Kneser-Ney smoothing, SRI.
3. Confidence Score: choose overall best engine if no clear winner.
Recursive procedure.

Only syntactically simple chunks sent to MT engine.

Substitution variables track location of translation chunks in target.

Therefore composing translation is straightforward.
A Worked Example

| The chairman | a long-time rival of Bill Gates | likes | fast and confidential deals |
| The chairman | a long-time rival of Bill Gates | likes | fast and confidential deals |
|--------------|---------------------------------|-------|----------------------------|
| El presidente |                                |       |                            |
| El presidente |                                |       |                            |
| El presidente |                                |       |                            |
### A Worked Example

| The chairman | a long-time rival of Bill Gates | likes | fast and confidential deals |
|--------------|---------------------------------|-------|-----------------------------|
| El presidente |                                |       |                             |
| El presidente |                                |       |                             |
| **El presidente** |                                |       |                             |
### A Worked Example

| The chairman | a long-time rival of Bill Gates | likes | fast and confidential deals |
|--------------|---------------------------------|-------|----------------------------|
| El presidente | una largo - vez rival de Bill Gates | (-33.77) | |
| El presidente | un rival de largo plazo de Bill Gates | (-23.41) | |
| El presidente | un rival antiguo de Bill Gates | (-22.60) | |
### A Worked Example

| The chairman | a long-time rival of Bill Gates | likes | fast and confidential deals |
|--------------|---------------------------------|-------|----------------------------|
| El presidente | una largo - vez rival de Bill Gates | (-33.77) |                           |
| El presidente | un rival de largo plazo de Bill Gates | (-23.41) |                           |
| El presidente | un rival antiguo de Bill Gates | (-22.60) |                           |
| El presidente | un rival antiguo de Bill Gates |       |                           |
| The chairman                     | a long-time rival of Bill Gates | likes | fast and confidential deals |
|---------------------------------|---------------------------------|-------|-----------------------------|
| El presidente                   | una largo - vez rival de Bill Gates | le gustan |                            |
|                                 | (-33.77)                        | (-10.94) |                            |
| El presidente                   | un rival de largo plazo de Bill Gates | tiene gusto | de |
|                                 | (-23.41)                        | (-16.41) |                            |
| El presidente                   | un rival antiguo de Bill Gates | quiere |                            |
|                                 | (-22.60)                        | (-9.73) |                            |
| El presidente                   | un rival antiguo de Bill Gates |      |                            |
| English | Spanish | likes | fast and confidential deals |
|---------|---------|-------|-----------------------------|
| The chairman | a long-time rival of Bill Gates | likes | fast and confidential deals |
| El presidente | una largo - vez rival de Bill Gates | le gustan | |
| | (-33.77) | (-10.94) | |
| El presidente | un rival de largo plazo de Bill Gates | tiene gusto de | |
| | (-23.41) | (-16.41) | |
| El presidente | un rival antiguo de Bill Gates | quiere | |
| | (-22.60) | (-9.73) | |
| El presidente | un rival antiguo de Bill Gates | quiere | |
| The chairman | a long-time rival of Bill Gates | likes | fast and confidential deals |
| El presidente | una largo - vez rival de Bill Gates | le gustan | los los tratos rápidos y confidenciales |
| | (-33.77) | (-10.94) | (-28.13) |
| El presidente | un rival de largo plazo de Bill Gates | tiene gusto de | repartos rápidos y confidenciales |
| | (-23.41) | (-16.41) | (-22.16) |
| El presidente | un rival antiguo de Bill Gates | quiere | los tratos rápidos y confidenciales |
| | (-22.60) | (-9.73) | (-23.12) |
| English                     | Spanish                      | Phrase                        |
|----------------------------|------------------------------|-------------------------------|
| The chairman               | El presidente               | un rival de largo plazo de Bill Gates |
| a long-time rival of Bill Gates | una largo - vez rival de Bill Gates | le gusta los tratos rápidos y confidenciales |
| likes                      | El presidente               | un rival antiguo de Bill Gates |
| fast and confidential deals | El presidente               | un rival antiguo de Bill Gates |
|                            |                              |                              |
### A Worked Example

| The chairman | a long-time rival of Bill Gates | likes | fast and confidential deals |
|--------------|---------------------------------|-------|-----------------------------|
| El presidente | una largo - vez rival de Bill Gates | le gustan | los los tratos rápidos y confidenciales |
|               | (-33.77)                         | (-10.94) | (-28.13) |
| El presidente | un rival de largo plazo de Bill Gates | tiene gusto de | repartos rápidos y confidenciales |
|               | (-23.41)                         | (-16.41) | (-22.16) |
| El presidente | un rival antiguo de Bill Gates | quiere | los tratos rápidos y confidenciales |
|               | (-22.60)                         | (-9.73)  | (-23.12) |
| El presidente | un rival antiguo de Bill Gates | quiere | repartos rápidos y confidenciales |
A Worked Example

- Our approach is not limited to a blind combination of previously produced output chunks.
- An individual MT system might improve its own translation, which can contribute to a better MEMT score.

Source
The chairman, a long-time rival of Bill Gates, likes fast and confidential deals.

Baseline MT
El presidente, rival de largo plazo de Bill Gates, gustos ayuna y los repartos confidenciales.

TransBooster
El presidente, un rival de largo plazo de Bill Gates, tiene gusto de repartos rápidos y confidenciales.
A Worked Example

- Our approach is not limited to a blind combination of previously produced output chunks.
- An individual MT system might improve its own translation, which can contribute to a better MEMT score.

**Source**
The chairman, a long-time rival of Bill Gates, **likes** fast and confidential deals.

**Baseline MT**
El presidente, rival de largo plazo de Bill Gates, **gustos** ayuna y los repartos confidenciales.

**TransBooster**
El presidente, un rival de largo plazo de Bill Gates, **tiene gusto de** repartos rápidos y confidenciales.
A Worked Example

- Our approach is not limited to a blind combination of previously produced output chunks.
- An individual MT system might improve its own translation, which can contribute to a better MEMT score.

Source
The chairman, a long-time rival of Bill Gates, likes fast and confidential deals.

Baseline MT
El presidente, rival de largo plazo de Bill Gates, gustos ayuna y los repartos confi denciales.

TransBooster
El presidente, un rival de largo plazo de Bill Gates, tiene gusto de repartos rápidos y confi denciales.
Outline

What is Multi-Engine Machine Translation?

MEMT with Recursive Sentence Decomposition.

Experiments, Results and Analysis.
Experimental Setup

- English → Spanish.
- 800-sentence test set Penn-II treebank.
- One set of 800 reference translations.
- Baseline systems: LogoMedia, Systran, SDL.
- Three different syntactic analyses as input.
  1. Original Penn-II structure.
  2. (Charniak, 2000)
  3. (Bikel, 2002)
|       | BLEU  | NIST  | GTM  |
|-------|-------|-------|------|
| LogoMedia | 104.9 | 104.8 | 103.1 |
| Systran  | 109.7 | 107.1 | 104.4 |
| SDL      | 108.4 | 105.5 | 102.4 |

MEMT1: relative BLEU increase of 8.4%-9.7%.

|       | BLEU  | NIST  | GTM  |
|-------|-------|-------|------|
| LogoMedia | 102.1 | 103.5 | 102.0 |
| Systran  | 106.8 | 105.8 | 103.4 |
| SDL      | 105.6 | 104.2 | 101.4 |

MEMT2: relative BLEU increase of 2.1%-6.8%.

|       | BLEU  | NIST  | GTM  |
|-------|-------|-------|------|
| LogoMedia | 101.2 | 103.2 | 101.8 |
| Systran  | 105.8 | 105.5 | 103.2 |
| SDL      | 104.6 | 103.9 | 101.2 |

MEMT3: relative BLEU increase of 1.2%-5.8%.
Conclusions

- Novel approach to MEMT.
- Input preparation instead of output alignment.
- Recursive sentence decomposition.
- Individual MT systems can improve their output before Multi-Engine selection.
Future Research

- Experiment with a variety of language models (Nomoto, 2004).
- Experiment with similarity measure.
- Implement word graph-based consensus at level of output chunks.
Questions?

Thanks for your attention.

http://www.computing.dcu.ie/research/nclt