Models of Synchronous Grammar Induction for SMT

Workshop 2010

The Center for Speech and Language Processing
Johns Hopkins University

June 21, 2010

Team members

Senior Members
Phil Blunsom (Oxford)
Trevor Cohn (Sheffield)
Adam Lopez (Edinburgh/COE)
Chris Dyer (CMU)
Jonathan Graehl (ISI)

Graduate Students
Jan Botha (Oxford)
Vladimir Eidelman (Maryland)
Ziyuan Wang (JHU)
ThuyLinh Nguyen (CMU)

Undergraduate Students
Olivia Buzek (Maryland)
Desai Chen (CMU)

Statistical machine translation

Urdu → English

اس حملہ کے بعد بڑی تعداد میں مقامی باشندوں نے علاقوں کو خالی کردیا ہے .

↓

- Statistical machine translation: Learn how to translate from parallel corpora.

Statistical machine translation:

Urdu → English

اس حملہ کے بعد بڑی تعداد میں مقامی باشندوں نے علاقوں کو خالی کردیا ہے .

↓

After this incident, a large number of local residents fled from these areas.

- Statistical machine translation: Learn how to translate from parallel corpora

Statistical machine translation: Before

Urdu → English

اس حملہ کے بعد بڑی تعداد میں مقامی باشندوں نے علاقوں کو خالی کردیا ہے .

↓

In this attack a large number of local residents has should vacate areas.

- Current state-of-the-art translation models struggle with language pairs which exhibit large differences in structure.

Statistical machine translation: After

Urdu → English

اس حملہ کے بعد بڑی تعداد میں مقامی باشندوں نے علاقوں کو خالی کردیا ہے ۔

↓

After this attack, a large number of local residents have to vacate the areas.

- In this workshop we've made some small steps towards better translations for difficult language pairs.

Statistical machine translation: limitations

Structural divergence between languages:

English	**Who wrote this letter?**
Arabic	من الذي كتب هذه الرسالة؟ (function-word) (who) (wrote) (this) (the-letter)
Chinese	这封信是谁写的？ (this) (letter) (be) (who) (write) (come-from) (function-word)

Statistical machine translation: limitations

Structural divergence between languages:

English	**Who wrote this letter?**
Arabic	من الذي كتب هذه الرسالة؟ (function-word) (who) (wrote) (this) (the-letter)
Chinese	这封 信 是 谁 写 的 ？ (this) (letter) (be) (who) (write) (come-from) (function-word)

Statistical machine translation: limitations

Structural divergence between languages:

English	**Who wrote this letter?**
Arabic	من الذي كتب هذه الرسالة؟ (function-word) (who) (wrote) (this) (the-letter)
Chinese	这封 信 是 谁 写 的 ？ (this) (letter) (be) (who) (write) (come-from) (function-word)

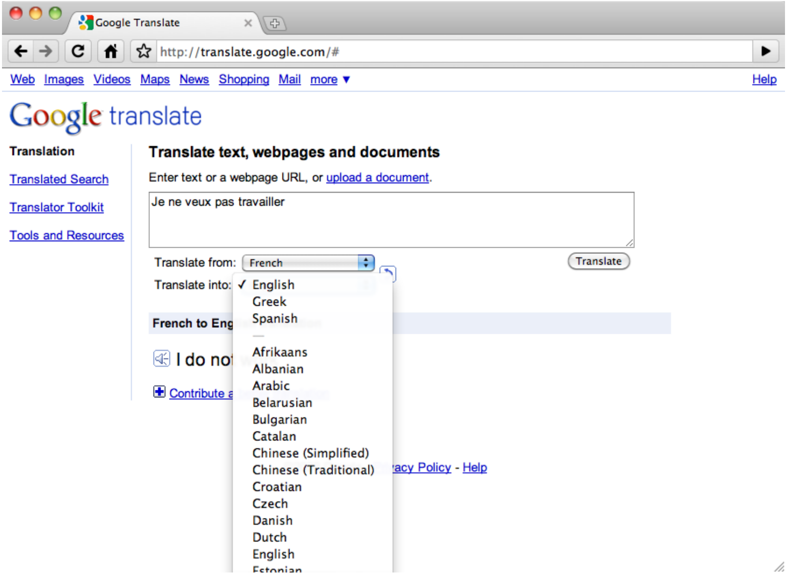
Statistical machine translation: limitations

Structural divergence between languages:

English	**Who wrote this letter?**
Arabic	من الذي كتب هذه الرسالة؟ (function-word) (who) (wrote) (this) (the-letter)
Chinese	这封 信 是 谁 写 的 ？ (this) (letter) (be) (who) (write) (come-from) (function-word)

- Phrasal translation equivalences (existing models)
- **Constituent reordering (this workshop!)**
- Morphology (Next year?)

Statistical machine translation: successes



Google Translate

http://translate.google.com/#

Translate text, webpages and documents

Enter text or a webpage URL, or upload a document.

Je ne veux pas travailler

Translate from: French

Translate into: ✓ English

French to Eng

I do no

Contribute a

Workshop overview

Input:

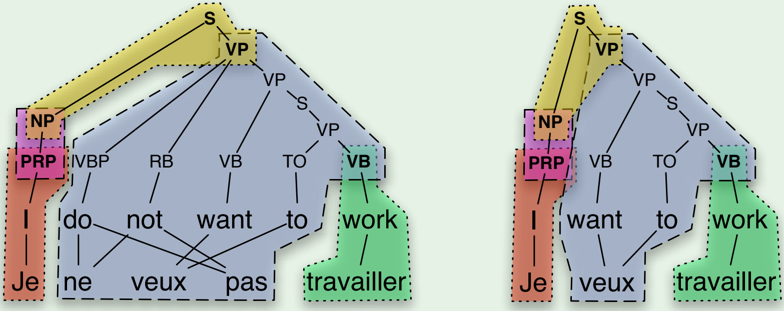
- Existing procedures for synchronous grammar extraction

Output:

- New unsupervised models for large scale synchronous grammar extraction,
- A comparison and analysis of the existing and proposed models,
- Extended decoders (cdec/Joshua) capable of working efficiently with these models.

Models of translation

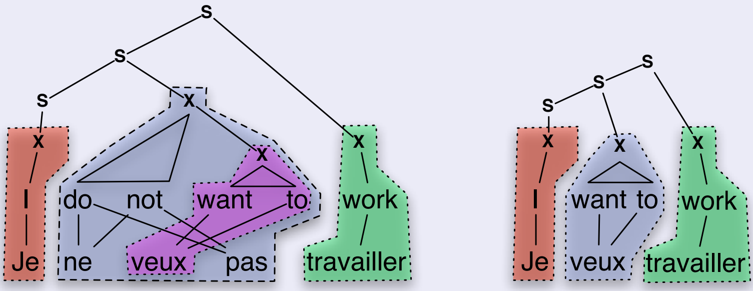
Supervised SCFG: Syntactic Tree-to-String



- Strong model of sentence structure.
- Reliant on a treebank to train the parser.

Models of translation

Unlabelled SCFG: Hiero



- Only requires the parallel corpus.
- But weak model of sentence structure.

Using syntax in Machine Translation:

Synchronous Context Free Grammar (SCFG)

$$S \rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle$$

$$X \rightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{2}}\ X_{\boxed{1}} \rangle$$

$$X \rightarrow \langle \mathit{Sie},\ \mathit{She} \rangle$$

$$X \rightarrow \langle \mathit{eine\ Tasse\ Kaffee},\ \mathit{a\ cup\ of\ coffee} \rangle$$

$$X \rightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{1}}\ X_{\boxed{2}} \rangle$$

$$X \rightarrow \langle \mathit{will},\ \mathit{wants\ to} \rangle$$

$$X \rightarrow \langle \mathit{trinken},\ \mathit{drink} \rangle$$

Example Derivation

$$S \Rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle \quad \Rightarrow \langle X_{\boxed{2}}\ X_{\boxed{3}},\ X_{\boxed{2}}\ X_{\boxed{3}} \rangle$$

$$\Rightarrow \langle \mathit{Sie}\ X_{\boxed{3}},\ \mathit{She}\ X_{\boxed{3}} \rangle \quad \Rightarrow \langle \mathit{Sie}\ X_{\boxed{4}}\ X_{\boxed{5}},\ \mathit{She}\ X_{\boxed{4}}\ X_{\boxed{5}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{5}},\ \mathit{She\ wants\ to}\ X_{\boxed{5}} \rangle \quad \Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{6}} X_{\boxed{7}},\ \mathit{She\ wants\ to}\ X_{\boxed{7}} X_{\boxed{6}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee}\ X_{\boxed{7}},\ \mathit{She\ wants\ to}\ X_{\boxed{7}}\ \mathit{a\ cup\ of\ coffee} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee\ trinken},\ \mathit{She\ wants\ to\ drink\ a\ cup\ of\ coffee} \rangle$$

Using syntax in Machine Translation:

Synchronous Context Free Grammar (SCFG)

$S \rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle$

$X \rightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{2}}\ X_{\boxed{1}} \rangle$

$X \rightarrow \langle \mathit{Sie},\ \mathit{She} \rangle$

$X \rightarrow \langle \mathit{eine\ Tasse\ Kaffee},\ \mathit{a\ cup\ of\ coffee} \rangle$

$X \rightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{1}}\ X_{\boxed{2}} \rangle$

$X \rightarrow \langle \mathit{will},\ \mathit{wants\ to} \rangle$

$X \rightarrow \langle \mathit{trinken},\ \mathit{drink} \rangle$

Example Derivation

$$S \Rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle \quad \Rightarrow \langle X_{\boxed{2}}\ X_{\boxed{3}},\ X_{\boxed{2}}\ X_{\boxed{3}} \rangle$$

$$\Rightarrow \langle \mathit{Sie}\ X_{\boxed{3}},\ \mathit{She}\ X_{\boxed{3}} \rangle \quad \Rightarrow \langle \mathit{Sie}\ X_{\boxed{4}}\ X_{\boxed{5}},\ \mathit{She}\ X_{\boxed{4}}\ X_{\boxed{5}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{5}},\ \mathit{She\ wants\ to}\ X_{\boxed{5}} \rangle \quad \Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{6}} X_{\boxed{7}},\ \mathit{She\ wants\ to}\ X_{\boxed{7}} X_{\boxed{6}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee}\ X_{\boxed{7}},\ \mathit{She\ wants\ to}\ X_{\boxed{7}}\ \mathit{a\ cup\ of\ coffee} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee\ trinken},\ \mathit{She\ wants\ to\ drink\ a\ cup\ of\ coffee} \rangle$$

Using syntax in Machine Translation:

Synchronous Context Free Grammar (SCFG)

$S \longrightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle$

$X \longrightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{2}}\ X_{\boxed{1}} \rangle$

$X \longrightarrow \langle Sie,\ She \rangle$

$X \longrightarrow \langle eine\ Tasse\ Kaffee,\ a\ cup\ of\ coffee \rangle$

$X \longrightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{1}}\ X_{\boxed{2}} \rangle$

$X \longrightarrow \langle will,\ wants\ to \rangle$

$X \longrightarrow \langle trinken,\ drink \rangle$

Example Derivation

$S \Rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle \quad \Rightarrow \langle X_{\boxed{2}}\ X_{\boxed{3}},\ X_{\boxed{2}}\ X_{\boxed{3}} \rangle$

$\Rightarrow \langle Sie\ X_{\boxed{3}},\ She\ X_{\boxed{3}} \rangle \quad \Rightarrow \langle Sie\ X_{\boxed{4}}\ X_{\boxed{5}},\ She\ X_{\boxed{4}}\ X_{\boxed{5}} \rangle$

$\Rightarrow \langle Sie\ will\ X_{\boxed{5}},\ She\ wants\ to\ X_{\boxed{5}} \rangle \quad \Rightarrow \langle Sie\ will\ X_{\boxed{6}} X_{\boxed{7}},\ She\ wants\ to\ X_{\boxed{7}} X_{\boxed{6}} \rangle$

$\Rightarrow \langle Sie\ will\ eine\ Tasse\ Kaffee\ X_{\boxed{7}},\ She\ wants\ to\ X_{\boxed{7}}\ a\ cup\ of\ coffee \rangle$

$\Rightarrow \langle Sie\ will\ eine\ Tasse\ Kaffee\ trinken,\ She\ wants\ to\ drink\ a\ cup\ of\ coffee \rangle$

Using syntax in Machine Translation:

Synchronous Context Free Grammar (SCFG)

$$S \rightarrow \langle X_{\boxed{1}},\; X_{\boxed{1}} \rangle$$

$$X \rightarrow \langle X_{\boxed{1}}\, X_{\boxed{2}},\; X_{\boxed{2}}\, X_{\boxed{1}} \rangle$$

$$X \rightarrow \langle \mathit{Sie},\; \mathit{She} \rangle$$

$$X \rightarrow \langle \mathit{eine\ Tasse\ Kaffee},\; \mathit{a\ cup\ of\ coffee} \rangle$$

$$X \rightarrow \langle X_{\boxed{1}}\, X_{\boxed{2}},\; X_{\boxed{1}}\, X_{\boxed{2}} \rangle$$

$$X \rightarrow \langle \mathit{will},\; \mathit{wants\ to} \rangle$$

$$X \rightarrow \langle \mathit{trinken},\; \mathit{drink} \rangle$$

Example Derivation

$$S \Rightarrow \langle X_{\boxed{1}},\; X_{\boxed{1}} \rangle \quad \Rightarrow \langle X_{\boxed{2}}\, X_{\boxed{3}},\; X_{\boxed{2}}\, X_{\boxed{3}} \rangle$$

$$\Rightarrow \langle \mathit{Sie}\ X_{\boxed{3}},\; \mathit{She}\ X_{\boxed{3}} \rangle \quad \Rightarrow \langle \mathit{Sie}\ X_{\boxed{4}}\, X_{\boxed{5}},\; \mathit{She}\ X_{\boxed{4}}\, X_{\boxed{5}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{5}},\; \mathit{She\ wants\ to}\ X_{\boxed{5}} \rangle \quad \Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{6}} X_{\boxed{7}},\; \mathit{She\ wants\ to}\ X_{\boxed{7}} X_{\boxed{6}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee}\ X_{\boxed{7}},\; \mathit{She\ wants\ to}\ X_{\boxed{7}}\ \mathit{a\ cup\ of\ coffee} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee\ trinken},\; \mathit{She\ wants\ to\ drink\ a\ cup\ of\ coffee} \rangle$$

Using syntax in Machine Translation:

Synchronous Context Free Grammar (SCFG)

$S \longrightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle$

$X \longrightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{2}}\ X_{\boxed{1}} \rangle$

$X \longrightarrow \langle \mathit{Sie},\ \mathit{She} \rangle$

$X \longrightarrow \langle \mathit{eine\ Tasse\ Kaffee},\ \mathit{a\ cup\ of\ coffee} \rangle$

$X \longrightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{1}}\ X_{\boxed{2}} \rangle$

$X \longrightarrow \langle \mathit{will},\ \mathit{wants\ to} \rangle$

$X \longrightarrow \langle \mathit{trinken},\ \mathit{drink} \rangle$

Example Derivation

$$S \Rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle \quad \Rightarrow \langle X_{\boxed{2}}\ X_{\boxed{3}},\ X_{\boxed{2}}\ X_{\boxed{3}} \rangle$$

$$\Rightarrow \langle \mathit{Sie}\ X_{\boxed{3}},\ \mathit{She}\ X_{\boxed{3}} \rangle \quad \Rightarrow \langle \mathit{Sie}\ X_{\boxed{4}}\ X_{\boxed{5}},\ \mathit{She}\ X_{\boxed{4}}\ X_{\boxed{5}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{5}},\ \mathit{She\ wants\ to}\ X_{\boxed{5}} \rangle \quad \Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{6}} X_{\boxed{7}},\ \mathit{She\ wants\ to}\ X_{\boxed{7}} X_{\boxed{6}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee}\ X_{\boxed{7}},\ \mathit{She\ wants\ to}\ X_{\boxed{7}}\ \mathit{a\ cup\ of\ coffee} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee\ trinken},\ \mathit{She\ wants\ to\ drink\ a\ cup\ of\ coffee} \rangle$$

Using syntax in Machine Translation:

Synchronous Context Free Grammar (SCFG)

$S \rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle$

$X \rightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{2}}\ X_{\boxed{1}} \rangle$

$X \rightarrow \langle \mathit{Sie},\ \mathit{She} \rangle$

$X \rightarrow \langle \mathit{eine\ Tasse\ Kaffee},\ \mathit{a\ cup\ of\ coffee} \rangle$

$X \rightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{1}}\ X_{\boxed{2}} \rangle$

$X \rightarrow \langle \mathit{will},\ \mathit{wants\ to} \rangle$

$X \rightarrow \langle \mathit{trinken},\ \mathit{drink} \rangle$

Example Derivation

$$S \Rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle \quad \Rightarrow \langle X_{\boxed{2}}\ X_{\boxed{3}},\ X_{\boxed{2}}\ X_{\boxed{3}} \rangle$$

$$\Rightarrow \langle \mathit{Sie}\ X_{\boxed{3}},\ \mathit{She}\ X_{\boxed{3}} \rangle \quad \Rightarrow \langle \mathit{Sie}\ X_{\boxed{4}}\ X_{\boxed{5}},\ \mathit{She}\ X_{\boxed{4}}\ X_{\boxed{5}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{5}},\ \mathit{She\ wants\ to}\ X_{\boxed{5}} \rangle \quad \Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{6}} X_{\boxed{7}},\ \mathit{She\ wants\ to}\ X_{\boxed{7}} X_{\boxed{6}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee}\ X_{\boxed{7}},\ \mathit{She\ wants\ to}\ X_{\boxed{7}}\ \mathit{a\ cup\ of\ coffee} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee\ trinken},\ \mathit{She\ wants\ to\ drink\ a\ cup\ of\ coffee} \rangle$$

Using syntax in Machine Translation:

Synchronous Context Free Grammar (SCFG)

$S \rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle$

$X \rightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{1}}\ X_{\boxed{2}} \rangle$

$X \rightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{2}}\ X_{\boxed{1}} \rangle$

$X \rightarrow \langle \textit{Sie},\ \textit{She} \rangle$

$X \rightarrow \langle \textit{will},\ \textit{wants to} \rangle$

$X \rightarrow \langle \textit{eine Tasse Kaffee},\ \textit{a cup of coffee} \rangle$

$X \rightarrow \langle \textit{trinken},\ \textit{drink} \rangle$

Example Derivation

$$S \Rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle \quad \Rightarrow \langle X_{\boxed{2}}\ X_{\boxed{3}},\ X_{\boxed{2}}\ X_{\boxed{3}} \rangle$$

$$\Rightarrow \langle \textit{Sie}\ X_{\boxed{3}},\ \textit{She}\ X_{\boxed{3}} \rangle \quad \Rightarrow \langle \textit{Sie}\ X_{\boxed{4}}\ X_{\boxed{5}},\ \textit{She}\ X_{\boxed{4}}\ X_{\boxed{5}} \rangle$$

$$\Rightarrow \langle \textit{Sie will}\ X_{\boxed{5}},\ \textit{She wants to}\ X_{\boxed{5}} \rangle \quad \Rightarrow \langle \textit{Sie will}\ X_{\boxed{6}} X_{\boxed{7}},\ \textit{She wants to}\ X_{\boxed{7}} X_{\boxed{6}} \rangle$$

$$\Rightarrow \langle \textit{Sie will eine Tasse Kaffee}\ X_{\boxed{7}},\ \textit{She wants to}\ X_{\boxed{7}}\ \textit{a cup of coffee} \rangle$$

$$\Rightarrow \langle \textit{Sie will eine Tasse Kaffee trinken},\ \textit{She wants to drink a cup of coffee} \rangle$$

Using syntax in Machine Translation:

Synchronous Context Free Grammar (SCFG)

$S \rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle$

$X \rightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{2}}\ X_{\boxed{1}} \rangle$

$X \rightarrow \langle \mathit{Sie},\ \mathit{She} \rangle$

$X \rightarrow \langle \mathit{eine\ Tasse\ Kaffee},\ \mathit{a\ cup\ of\ coffee} \rangle$

$X \rightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{1}}\ X_{\boxed{2}} \rangle$

$X \rightarrow \langle \mathit{will},\ \mathit{wants\ to} \rangle$

$X \rightarrow \langle \mathit{trinken},\ \mathit{drink} \rangle$

Example Derivation

$$S \Rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle \quad \Rightarrow \langle X_{\boxed{2}}\ X_{\boxed{3}},\ X_{\boxed{2}}\ X_{\boxed{3}} \rangle$$

$$\Rightarrow \langle \mathit{Sie}\ X_{\boxed{3}},\ \mathit{She}\ X_{\boxed{3}} \rangle \quad \Rightarrow \langle \mathit{Sie}\ X_{\boxed{4}}\ X_{\boxed{5}},\ \mathit{She}\ X_{\boxed{4}}\ X_{\boxed{5}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{5}},\ \mathit{She\ wants\ to}\ X_{\boxed{5}} \rangle \quad \Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{6}} X_{\boxed{7}},\ \mathit{She\ wants\ to}\ X_{\boxed{7}} X_{\boxed{6}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee}\ X_{\boxed{7}},\ \mathit{She\ wants\ to}\ X_{\boxed{7}}\ \mathit{a\ cup\ of\ coffee} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee\ trinken},\ \mathit{She\ wants\ to\ drink\ a\ cup\ of\ coffee} \rangle$$

Using syntax in Machine Translation:

Synchronous Context Free Grammar (SCFG)

$S \rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle$

$X \rightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{2}}\ X_{\boxed{1}} \rangle$

$X \rightarrow \langle \mathit{Sie},\ \mathit{She} \rangle$

$X \rightarrow \langle \mathit{eine\ Tasse\ Kaffee},\ \mathit{a\ cup\ of\ coffee} \rangle$

$X \rightarrow \langle X_{\boxed{1}}\ X_{\boxed{2}},\ X_{\boxed{1}}\ X_{\boxed{2}} \rangle$

$X \rightarrow \langle \mathit{will},\ \mathit{wants\ to} \rangle$

$X \rightarrow \langle \mathit{trinken},\ \mathit{drink} \rangle$

Example Derivation

$$S \Rightarrow \langle X_{\boxed{1}},\ X_{\boxed{1}} \rangle \quad \Rightarrow \langle X_{\boxed{2}}\ X_{\boxed{3}},\ X_{\boxed{2}}\ X_{\boxed{3}} \rangle$$

$$\Rightarrow \langle \mathit{Sie}\ X_{\boxed{3}},\ \mathit{She}\ X_{\boxed{3}} \rangle \quad \Rightarrow \langle \mathit{Sie}\ X_{\boxed{4}}\ X_{\boxed{5}},\ \mathit{She}\ X_{\boxed{4}}\ X_{\boxed{5}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{5}},\ \mathit{She\ wants\ to}\ X_{\boxed{5}} \rangle \quad \Rightarrow \langle \mathit{Sie\ will}\ X_{\boxed{6}} X_{\boxed{7}},\ \mathit{She\ wants\ to}\ X_{\boxed{7}} X_{\boxed{6}} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee}\ X_{\boxed{7}},\ \mathit{She\ wants\ to}\ X_{\boxed{7}}\ \mathit{a\ cup\ of\ coffee} \rangle$$

$$\Rightarrow \langle \mathit{Sie\ will\ eine\ Tasse\ Kaffee\ trinken},\ \mathit{She\ wants\ to\ drink\ a\ cup\ of\ coffee} \rangle$$

Models of translation

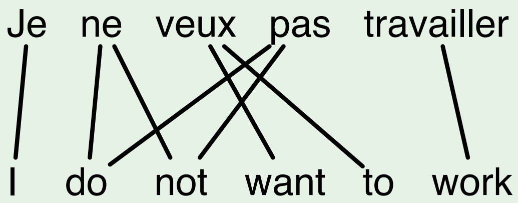
Phrase extraction:

Je ne veux pas travailler

I do not want to work

Models of translation

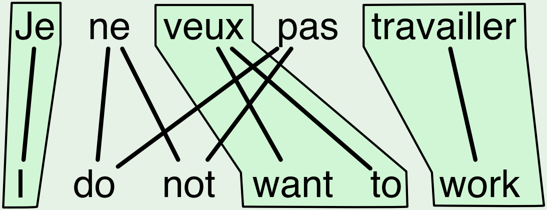
Phrase extraction:



- Use a word-based translation model to annotate the parallel corpus with word-alignments

Models of translation

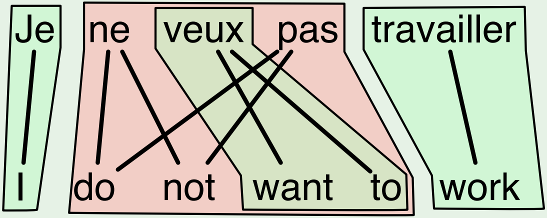
Phrase extraction:



- ⟨ Je, I ⟩, ⟨ veux, want to ⟩, ⟨ travailler, work ⟩

Models of translation

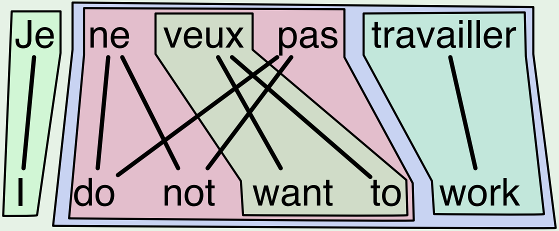
Phrase extraction:



- ⟨ Je, I ⟩, ⟨ veux, want to ⟩, ⟨ travailler, work ⟩, ⟨ ne veux pas, do not want to ⟩

Models of translation

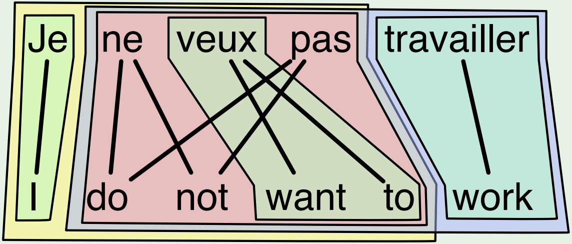
Phrase extraction:



- ⟨ Je, I ⟩, ⟨ veux, want to ⟩, ⟨ travailler, work ⟩, ⟨ ne veux pas, do not want to ⟩, ⟨ ne veux pas travailler, do not want to work ⟩

Models of translation

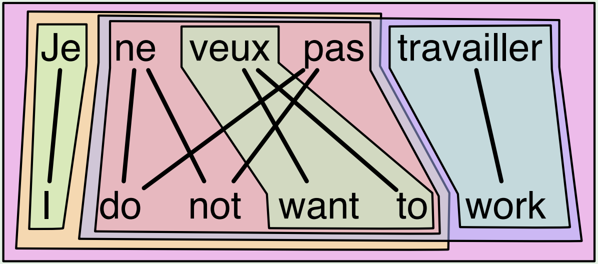
Phrase extraction:



- ⟨ Je, I ⟩, ⟨ veux, want to ⟩, ⟨ travailler, work ⟩, ⟨ ne veux pas, do not want to ⟩, ⟨ ne veux pas travailler, do not want to work ⟩, ⟨ Je ne veux pas, I do not want to ⟩

Models of translation

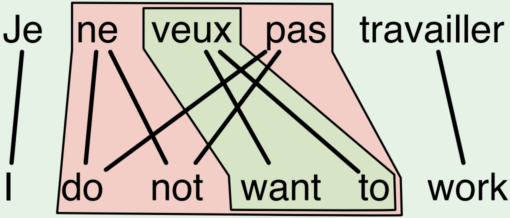
Phrase extraction:



- ⟨ Je, I ⟩, ⟨ veux, want to ⟩, ⟨ travailler, work ⟩, ⟨ ne veux pas, do not want to ⟩, ⟨ ne veux pas travailler, do not want to work ⟩, ⟨ Je ne veux pas, I do not want to ⟩, ⟨ Je ne veux pas travailler, I do not want to work ⟩

Models of translation

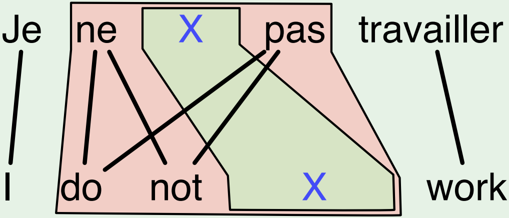
SCFG Rule extraction:



- X -> ⟨ ne veux pas, do not want to ⟩

Models of translation

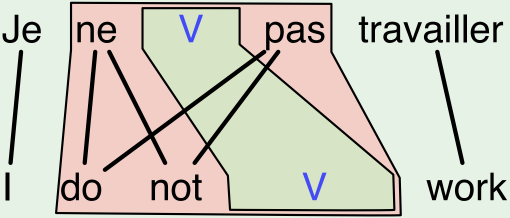
SCFG Rule extraction:



- X -> ⟨ ne veux pas, do not want to ⟩,
- X -> ⟨ ne $X_{\boxed{1}}$ pas, do not $X_{\boxed{1}}$ ⟩

Models of translation

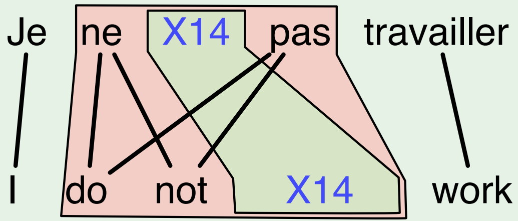
SCFG Rule extraction:



- VP/NN -> $\langle$ ne veux pas, do not want to $\rangle$,
- VP/NN -> $\langle$ ne $V_{\boxed{1}}$ pas, do not $V_{\boxed{1}}$ $\rangle$

Models of translation

SCFG Rule extraction:



- X10 -> ⟨ ne veux pas, do not want to ⟩,
- X10 -> ⟨ ne $X14_{\boxed{1}}$ pas, do not $X14_{\boxed{1}}$ ⟩

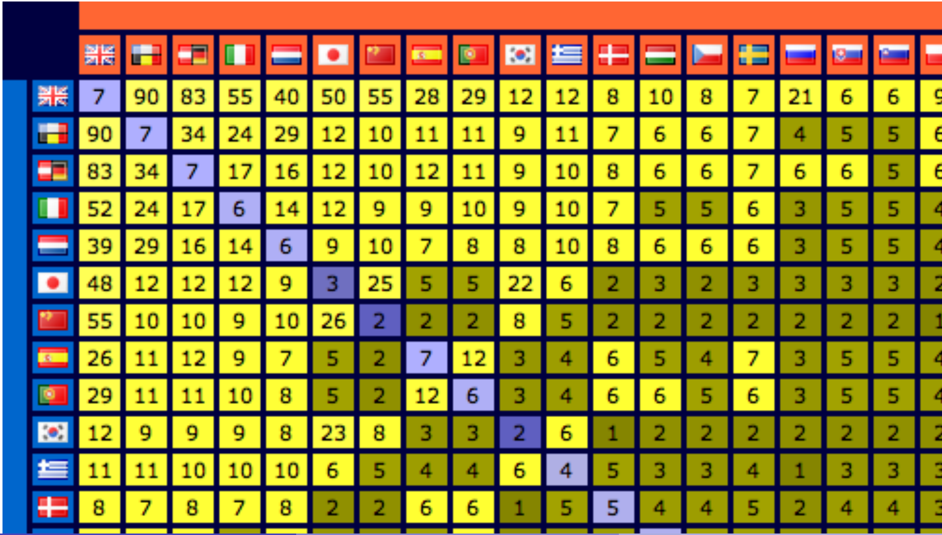
Impact

Language	Words	Domain
English	4.5M	Financial news
Chinese	0.5M	Broadcasting news
Arabic	300K (1M planned)	News
Korean	54K	Military

Table: Major treebanks: data size and domain

Impact

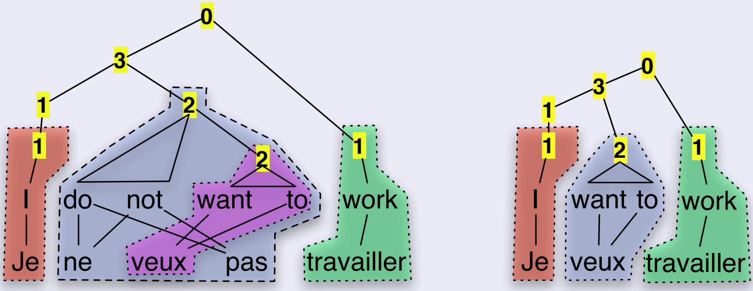
Parallel corpora far exceed treebanks (millions of words):



	🇬🇧	🇫🇷	🇩🇪	🇮🇹	🇳🇱	🇯🇵	🇨🇳	🇪🇸	🇵🇹	🇰🇷	🇬🇷	🇩🇰	🇭🇺	🇨🇿	🇸🇪	🇷🇺	🇸🇰	🇸🇮	
🇬🇧	7	90	83	55	40	50	55	28	29	12	12	8	10	8	7	21	6	6	9
🇫🇷	90	7	34	24	29	12	10	11	11	9	11	7	6	6	7	4	5	5	6
🇩🇪	83	34	7	17	16	12	10	12	11	9	10	8	6	6	7	6	6	5	6
🇮🇹	52	24	17	6	14	12	9	9	10	9	10	7	5	5	6	3	5	5	4
🇳🇱	39	29	16	14	6	9	10	7	8	8	10	8	6	6	6	3	5	5	4
🇯🇵	48	12	12	12	9	3	25	5	5	22	6	2	3	2	3	3	3	3	2
🇨🇳	55	10	10	9	10	26	2	2	2	8	5	2	2	2	2	2	2	2	1
🇪🇸	26	11	12	9	7	5	2	7	12	3	4	6	5	4	7	3	5	5	4
🇵🇹	29	11	11	10	8	5	2	12	6	3	4	6	6	5	6	3	5	5	4
🇰🇷	12	9	9	9	8	23	8	3	3	2	6	1	2	2	2	2	2	2	2
🇬🇷	11	11	10	10	10	6	5	4	4	6	4	5	3	3	4	1	3	3	3
🇩🇰	8	7	8	7	8	2	2	6	6	1	5	5	4	4	5	2	4	4	3

Models of translation

Hierarchical



- AIM: Implement a large scale open-source synchronous constituent learning system.
- AIM: Investigate and understand the relationship between the choice of synchronous grammar and SMT performance,
- AIM: and fix our decoders accordingly.

Evaluation goals

We will predominately evaluate using BLEU, but also use automatic structured metrics and perform small scale human evaluation:

- Evaluate phrasal, syntactic, unsupervised syntactic,
- Aim 1: Do no harm (not true of existing syntactic approach)
- Aim 2: Exceed the performance of current non-syntactic systems.
- Aim 3: Meet or exceed performance of existing syntactic systems.

Workshop Streams

- Implement scalable SCFG grammar extraction algorithms.
- Improve SCFG decoders to effieciently handle the grammars produce.
- Investigate discriminative training regimes the leverage features extracted from these grammars.

Language pairs (small)

- BTEC Chinese-English:
  - 44k sentence pairs, short sentences
  - Widely reported ‘prototyping’ corpus
  - Hiero baseline score: 52.4 (16 references)
  - Prospects: BTEC always gives you good results
- NIST Urdu-English:
  - 50k sentence pairs
  - Hiero baseline score: MT05 - 23.7 (4 references)
  - Major challenges: major long-range reordering, SOV word order
  - Prospects: small data, previous gains with supervised syntax

Language pairs (large)

- NIST Chinese-English:
  - 1.7M sentence pairs, Standard NIST test sets
  - Hiero baseline score: MT05 - 33.9 (4 references)
  - Major challenges: large data, mid-range reordering, lexical ambiguity
  - Prospects: supervised syntax gains reported
- NIST Arabic-English:
  - 900k sentence pairs
  - Hiero baseline score: MT05 - 48.9 (4 references)
  - Major challenges: strong baseline, local reordering, VSO word order
  - Prospects: difficult
- Europarl Dutch-French:
  - 1.5M sentence pairs, standard Europarl test sets
  - Hiero baseline score: Europarl 2008 - 26.3 (1 reference)
  - Major challenges: V2 / V-final word order, many non-literal translations
  - Prospects: ???

Summary

- Scientific Merit:
  - A systematic comparison of existing syntactive approaches to SMT.
  - An empirical study of how constituency is useful in SMT.
  - An evaluation of existing theories of grammar induction in a practical application (end-to-end evaluation).
- Potential Impact:
  - Better MT systems, for more languages, across a range of domains.
  - More accessible high performance translation models for researchers.
- Feasibility:
  - A great team with a wide range of both theoretical and practical experience.
  - Solid preparation.
- Novelty:
  - First attempt at large scale unsupervised synchronous grammar induction.