The challenge of simultaneous speech translation

Anoop Sarkar

SFU NatLangLab
School of Computing Science
Simon Fraser University
Vancouver, British Columbia, Canada

PACLIC 30: Seoul. Oct 30, 2016

Simultaneous Translation

Simultaneous Translation [Latency vs. Quality]

Input Sentence

Simultaneous Translation

Reference translation

it is estimated that variations that occur in the sum total of the human genetic code are related to at least 1500 diseases such as diabetes , cancer and heart disease .

Simultaneous Translation [Latency vs. Quality]

Input Sentence

/ 인간 유전학에서 /

Simultaneous Translation

Reference translation

it is estimated that variations that occur in the sum total of the human genetic code are related to at least 1500 diseases such as diabetes , cancer and heart disease .

Simultaneous Translation [Latency vs. Quality]

Input Sentence

/ 인간 유전학에서 /

Simultaneous Translation

/ in human genetics /

Reference translation

it is estimated that variations that occur in the sum total of the human genetic code are related to at least 1500 diseases such as diabetes , cancer and heart disease .

Simultaneous Translation [Latency vs. Quality]

Input Sentence

/ 인간 유전학에서 / / 인간 지놈 총체에서 일어나는 변이는 /

Simultaneous Translation

/ in human genetics /

Reference translation

it is estimated that variations that occur in the sum total of the human genetic code are related to at least 1500 diseases such as diabetes , cancer and heart disease .

Simultaneous Translation [Latency vs. Quality]

Input Sentence

/ 인간 유전학에서 / / 인간 지놈 총체에서 일어나는 변이는 /

Simultaneous Translation

/ in human genetics / / variations that occur in the entire human genome /

Reference translation

it is estimated that variations that occur in the sum total of the human genetic code are related to at least 1500 diseases such as diabetes , cancer and heart disease .

Simultaneous Translation [Latency vs. Quality]

Input Sentence

/ 인간 유전학에서 / / 인간 지놈 총체에서 일어나는 변이는 /
/ 당뇨병 암 심장마비 등 /

Simultaneous Translation

/ in human genetics / / variations that occur in the entire human genome /

Reference translation

it is estimated that variations that occur in the sum total of the human genetic code are related to at least 1500 diseases such as diabetes , cancer and heart disease .

Simultaneous Translation [Latency vs. Quality]

Input Sentence

/ 인간 유전학에서 / / 인간 지놈 총체에서 일어나는 변이는 / / 당뇨병 암 심장마비 등 /

Simultaneous Translation

/ in human genetics / / variations that occur in the entire human genome / / is related to diabetes , cancer and heart attack /

Reference translation

it is estimated that variations that occur in the sum total of the human genetic code are related to at least 1500 diseases such as diabetes , cancer and heart disease .

Simultaneous Translation [Latency vs. Quality]

Input Sentence

/ 인간 유전학에서 / / 인간 지놈 총체에서 일어나는 변이는 / / 당뇨병 암 심장마비 등 / / 최소 1500여가지의 질병과 관련이 있는 것으로 /

Simultaneous Translation

/ in human genetics / / variations that occur in the entire human genome / / is related to diabetes , cancer and heart attack /

Reference translation

it is estimated that variations that occur in the sum total of the human genetic code are related to at least 1500 diseases such as diabetes , cancer and heart disease .

Simultaneous Translation [Latency vs. Quality]

Input Sentence

/ 인간 유전학에서 / / 인간 지놈 총체에서 일어나는 변이는 / / 당뇨병 암 심장마비 등 / / 최소 1500여가지의 질병과 관련이 있는 것으로 /

Simultaneous Translation

/ in human genetics / / variations that occur in the entire human genome / / is related to diabetes , cancer and heart attack / / and causes at least 1500 other diseases /

Reference translation

it is estimated that variations that occur in the sum total of the human genetic code are related to at least 1500 diseases such as diabetes , cancer and heart disease .

Simultaneous Translation [Latency vs. Quality]

Input Sentence

/ 인간 유전학에서 / / 인간 지놈 총체에서 일어나는 변이는 / / 당뇨병 암 심장마비 등 / / 최소 1500여가지의 질병과 관련이 있는 것으로 / / 추정된다 . /

Simultaneous Translation

/ in human genetics / / variations that occur in the entire human genome / / is related to diabetes , cancer and heart attack / / and causes at least 1500 other diseases /

Reference translation

it is estimated that variations that occur in the sum total of the human genetic code are related to at least 1500 diseases such as diabetes , cancer and heart disease .

Simultaneous Translation [Latency vs. Quality]

Input Sentence

/ 인간 유전학에서 / / 인간 지놈 총체에서 일어나는 변이는 / / 당뇨병 암 심장마비 등 / / 최소 1500여가지의 질병과 관련이 있는 것으로 / / 추정된다 . /

Simultaneous Translation

/ in human genetics / / variations that occur in the entire human genome / / is related to diabetes , cancer and heart attack / / and causes at least 1500 other diseases / / it is estimated . /

Reference translation

it is estimated that variations that occur in the sum total of the human genetic code are related to at least 1500 diseases such as diabetes , cancer and heart disease .

Simultaneous Translation [Latency vs. Quality]

Input Sentence

/ 인간 유전학에서 / / 인간 지놈 총체에서 일어나는 변이는 /
/ 당뇨병 암 심장마비 등 / / 최소 1500여가지의 질병과 관련이

- **Segmentation**: Avoid long delay in producing the translation (Oda+ 14)
- Prediction: To produce timely translations, predict what will be said (Grissom+ 14)
- Paraphrasing: e.g. convert to passive form if that reduces delay (Shimizu+ 13, He+ 15) [also Disfluencies]
- **Evaluation**: reward both translation quality and reduced delay (Mieno+ 15)

Speech to speech translation

Karlsruhe (KIT) Lecture Translator

Speech to speech translation

Karlsruhe (KIT) Lecture Translator

NICT Speech Translator

Speech to speech translation

Karlsruhe (KIT) Lecture Translator

NICT Speech Translator

Skype Translator

Speech to speech translation is not simultaneous

I made sure to include pauses after each sentence so that the audience would have time to clearly hear the Mandarin version of what I was saying. *This also meant there was plenty of time for the audience to react. I remember hearing some gasps from the front rows, along with general applause and approval from the audience. It was quite moving.*

— Rick Rashid (Microsoft) in an interview in 2014

Contributions

We improve the state of the art in simultaneous machine translation by providing:

- A choice between latency and translation quality using Pareto optimality

Contributions

We improve the state of the art in simultaneous machine translation by providing:

- A choice between latency and translation quality using Pareto optimality
- An efficient algorithm for segment annotation used to train a segmentation classifier

Contributions

We improve the state of the art in simultaneous machine translation by providing:

- A choice between latency and translation quality using Pareto optimality
- An efficient algorithm for segment annotation used to train a segmentation classifier
- A new simultaneous translation system that uses our segmentation classifier

Contributions

We improve the state of the art in simultaneous machine translation by providing:

- A choice between latency and translation quality using Pareto optimality
- An efficient algorithm for segment annotation used to train a segmentation classifier
- A new simultaneous translation system that uses our segmentation classifier
- Significant improvement in latency with the same quality

Segmentation

Measuring translation quality: Bleu score

Input:

Ich war in meinen zwanzigern bevor ich erstmals in ein kunstmuseum ging .

Reference translation:

I was in my twenties before I ever went to an art museum .

Measuring translation quality: Bleu score

Input:

Ich war in meinen zwanzigern bevor ich erstmals in ein kunstmuseum ging .

Reference translation:

I was in my twenties before I ever went to an art museum .

Low BLEU% score (41.1): [few n-gram matches with reference]

I was twenty I ever went to art .

Measuring translation quality: Bleu score

Input:

Ich war in meinen zwanzigern bevor ich erstmals in ein kunstmuseum ging .

Reference translation:

I was in my twenties before I ever went to an art museum .

Low BLEU% score (41.1): [few n-gram matches with reference]

I was twenty I ever went to art .

High BLEU% score (89.0): [many n-gram matches with reference]

I was in my twenties before I first went to an art museum .

Simultaneous Translation – the Delay problem

- No segmentation inside a sentence:

$\boxed{\text{I was in my twenties before I ever went to an art museum}}$

$\downarrow$

$\boxed{\text{Ich war in meinen zwanzig bevor ich in ein kunstmuseum ging}}$

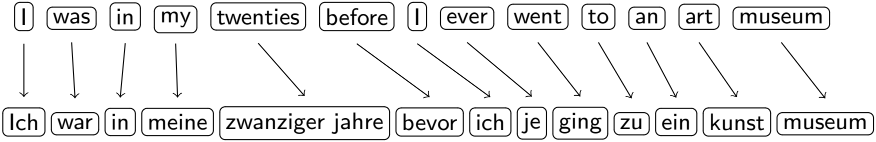
- Reference Sentence:

$\boxed{\text{Ich war in meinen zwanzigern bevor ich erstmals in ein kunstmuseum ging}}$

- BLEU Score: High (57.6)
- Segments/Second: Low

Simultaneous Translation – the Delay problem

- Word by word segmentation:

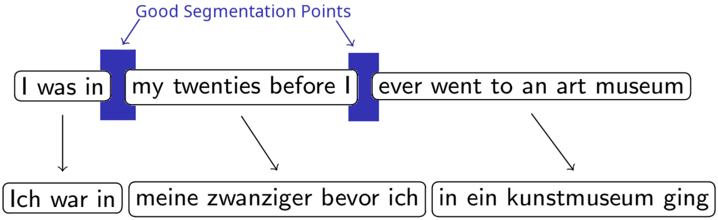


- Reference Sentence:

Ich war in meinen zwanzigern bevor ich erstmals in ein kunstmuseum ging

- BLEU Score: Low (15.6)
- Segments/Second: High

Segmentation - A Trade-off between Extremes



- Reference Sentence:

  Ich war in meinen zwanzigern bevor ich erstmals in ein kunstmuseum ging

- BLEU Score: Acceptable (38.2)
- Segments/Second: Acceptable

Creating Segmentation Data

Training a classifier needs annotated data

We provide a method that will create annotated data for segmentation

Creating Segmentation Data - An Example

- Task: English-German
- Segmentation locations indexed by adjacent part-of-speech tags (only source side shown here)
- For each possible segmentation location: translate segments and pre-compute BLEU scores
  - Exponential! Computed once for training data and stored.

I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

N[noun], V[verb], D[determiner], J[adjective], P[preposition], S[possessive pronoun], A[adverb], R[particle], .[dot]

Candidates for Segmentation: Using part of speech tags

Feat	Freq	Feat	Freq	Feat	Freq
N-P	6	J-N	3	V-R	1
P-D	5	N-N	2	P-S	1
D-N	4	P-N	2	P-J	1
N-.	3	D-J	2	S-N	1
N-V	3	R-P	1	A-V	1
V-D	3	N-A	1		
Full Segmentation Set Size				40	

Candidates for Segmentation: Using part of speech tags

Feat	Freq	Feat	Freq	Feat	Freq
N-P	6	J-N	3	V-R	1
P-D	5	N-N	2	P-S	1
D-N	4	P-N	2	P-J	1
N-.	3	D-J	2	S-N	1
N-V	3	R-P	1	A-V	1
V-D	3	N-A	1		
Full Segmentation Set Size				40	

I	am	a	contemporary	artist	ǀ	with	a	bit	ǀ	of	an	unexpected	background	.
N	V	D	J	N	ǀ	P	D	N	ǀ	P	D	J	N	.

I	was	in	my	twenties	ǀ	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	ǀ	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	ǀ	of	nowhere	ǀ	on	a	dirt	road	ǀ	in	rural	Arkansas	.
N	V	R	P	D	N	ǀ	P	N	ǀ	P	D	N	N	ǀ	P	J	N	.

Candidates for Segmentation: Using part of speech tags

Feat	Freq	Feat	Freq	Feat	Freq
N-P	6	J-N	3	V-R	1
P-D	5	N-N	2	P-S	1
D-N	4	P-N	2	P-J	1

Part of Speech alternatives

POS alternatives	Granularity

I grew up in the middle | of nowhere | on a dirt road | in rural Arkansas .
N V R P D N | P N | P D N N | P J N .

Candidates for Segmentation: Using part of speech tags

Feat	Freq	Feat	Freq	Feat	Freq
N-P	6	J-N	3	V-R	1
P-D	5	N-N	2	P-S	1
D-N	4	P-N	2	P-J	1

Part of Speech alternatives

POS alternatives	Granularity
Google universal tagset	12

I	grew	up	in	the	middle	|	of	nowhere	|	on	a	dirt	road	|	in	rural	Arkansas	.
N	V	R	P	D	N	|	P	N	|	P	D	N	N	|	P	J	N	.

Candidates for Segmentation: Using part of speech tags

Feat	Freq	Feat	Freq	Feat	Freq
N-P	6	J-N	3	V-R	1
P-D	5	N-N	2	P-S	1
D-N	4	P-N	2	P-J	1

Part of Speech alternatives

POS alternatives	Granularity
Google universal tagset	12
Penn Treebank tagset	36

I grew up in the middle | of nowhere | on a dirt road | in rural Arkansas .
N V R P D N | P N | P D N N | P J N .

Candidates for Segmentation: Using part of speech tags

Feat	Freq	Feat	Freq	Feat	Freq
N-P	6	J-N	3	V-R	1
P-D	5	N-N	2	P-S	1
D-N	4	P-N	2	P-J	1

Part of Speech alternatives

POS alternatives	Granularity
Google universal tagset	12
Penn Treebank tagset	36
Brown clusters 100C	100

I grew up in the middle | of nowhere | on a dirt road | in rural Arkansas .
N V R P D N | P N | P D N N | P J N .

Candidates for Segmentation: Using part of speech tags

Feat	Freq	Feat	Freq	Feat	Freq
N-P	6	J-N	3	V-R	1
P-D	5	N-N	2	P-S	1
D-N	4	P-N	2	P-J	1

Part of Speech alternatives

POS alternatives	Granularity
Google universal tagset	12
Penn Treebank tagset	36
Brown clusters 100C	100
Brown clusters 400C	389

Penn Treebank tagset gave us the best tradeoff between latency and quality.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N |	P	N |	P	D	N	N |	P	J	N	.

Greedy Segmentation

[Oda+ 2014]

Greedy Segmentation

- Greedily maximize the sum of BLEU Scores of Sentences
  - Decoding is done Sentence by Sentence

Greedy Segmentation

- Greedily maximize the sum of BLEU Scores of Sentences
  - Decoding is done Sentence by Sentence

- Input: the desired average segment length (μ)

  $\Rightarrow$ total number of expected segments (K)

Greedy Segmentation

- Greedily maximize the sum of BLEU Scores of Sentences
  - Decoding is done Sentence by Sentence

- Input: the desired average segment length (μ)

  $\Rightarrow$ total number of expected segments (K)

$$\boxed{K = \left\lfloor \frac{\#Words}{\mu} \right\rfloor - [\#Sentences]}$$

Sentence boundaries do not count towards K

Greedy Segmentation - An Example for $\mu = 13$

$$\boxed{K = 0 = \left\lfloor \frac{[\#Words{=}43]}{[\mu{=}13]} \right\rfloor - [\#Sentences = 3]}$$

Sum of $\text{B}\textsc{leu}$ Scores [of the 3 sentences] = 57.6

I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Greedy Segmentation - An Example for $\mu = 8$

$$\boxed{K = 2 = \left\lfloor \frac{[\#Words{=}43]}{[\mu{=}8]} \right\rfloor - [\#Sentences = 3]}$$

Sum of BLEU Scores [of the 3 sentences] = 13.8

I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Greedy Segmentation - An Example for $\mu = 8$

$$\boxed{K = 2 = \left\lfloor \frac{[\#Words=43]}{[\mu=8]} \right\rfloor - [\#Sentences = 3]}$$

Sum of BLEU Scores [of the 3 sentences] = 27.2

I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Greedy Segmentation - An Example for $\mu = 8$

$$K = 2 = \left\lfloor \frac{[\#Words{=}43]}{[\mu{=}8]} \right\rfloor - [\#Sentences = 3]$$

Sum of BLEU Scores [of the 3 sentences] = 38.2

I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Greedy Segmentation - An Example for $\mu = 8$

$$K = 2 = \left\lfloor \frac{[\#\mathit{Words}=43]}{[\mu=8]} \right\rfloor - [\#\mathit{Sentences} = 3]$$

Sum of BLEU Scores [of the 3 sentences] = 38.2

I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in
N	V	P

my	twenties	before	I
S	N	P	N

ever	went	to	an	art	museum	.
A	V	P	D	N	N	.

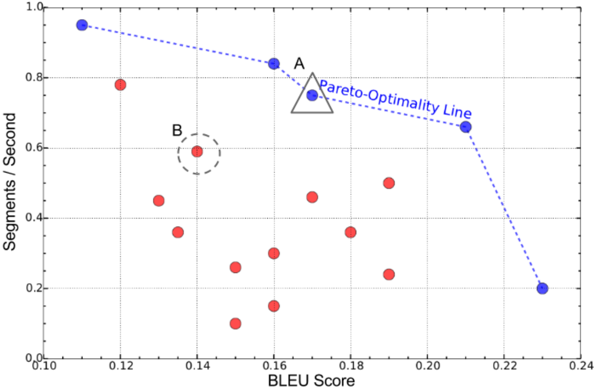
I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Only maximizes the BLEU score

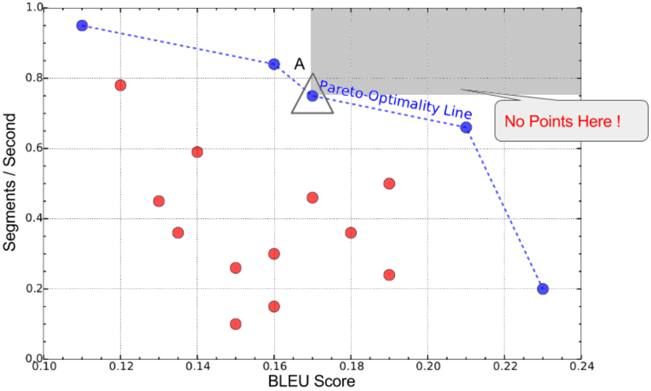
Tends to oversegment a small number of sentences

Pareto-Optimal Segmentation

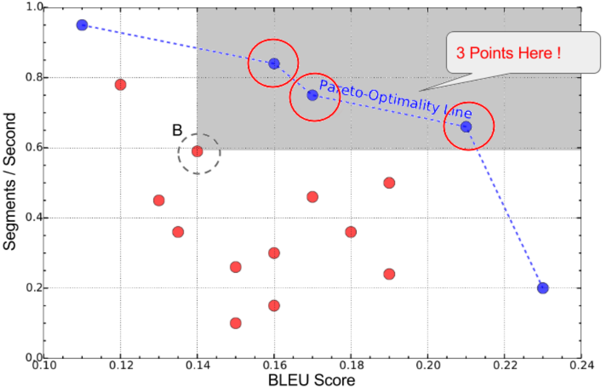
Pareto-Optimality



Pareto-Optimality



Pareto-Optimality



Pareto-Optimal Segmentation

- Tries to find the best segmentation points taking both BLEU and Segs/Sec into consideration
- The input is the same desired average segment length μ
- For each possible segmentation location: translate segments and pre-compute BLEU scores and segments/second.

Pareto-Optimal Segmentation - An Example for $\mu = 8$

$$\boxed{K = 2 = \left\lfloor \frac{[\#Words = 43]}{[\mu = 8]} \right\rfloor - [\#Sentences = 3]}$$

Avg $\{\frac{\text{BLEU}}{\#Segments}\}$/ Sentence $= 12.7$, Segs/Sec $= 0.560$

I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Pareto-Optimal Segmentation - An Example for $\mu = 8$

$$\boxed{K = 2 = \left\lfloor \frac{[\#Words=43]}{[\mu=8]} \right\rfloor - [\#Sentences = 3]}$$

Avg $\{\frac{\textsc{Bleu}}{\#Segments}\}$/ Sentence = 9.0, Segs/Sec = 0.956

I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Example: Find optimal pair of segments ($\mu = 8$)

Feat	Freq	Feat	Freq	Feat	Freq
N-P	6	J-N	3	V-R	1
P-D	5	N-N	2	P-S	1
D-N	4	P-N	2	P-J	1
N-.	3	D-J	2	S-N	1
N-V	3	R-P	1	A-V	1
V-D	3	N-A	1		
Full Segmentation Set Size			40		

I am a contemporary artist with a bit of an unexpected background .
N V D J N P D N P D J N .

I was in my twenties before I ever went to an art museum .
N V P S N P N A V P D N N .

I grew up in the middle of nowhere on a dirt road in rural Arkansas .
N V R P D N P N P D N N P J N .

Example: Find optimal pair of segments ($\mu = 8$)

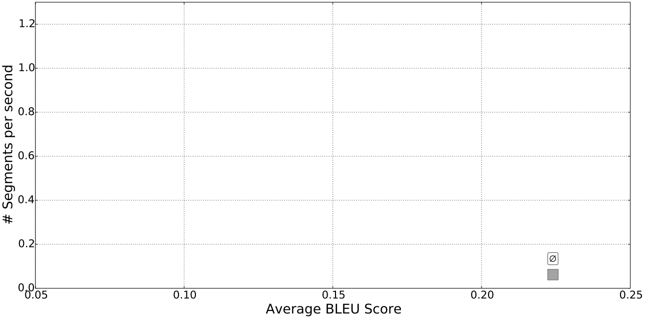
Feat	Freq	Feat	Freq	Feat	Freq
N-P	6	J-N	3	V-R	1
P-D	5	N-N	2	P-S	1
D-N	4	P-N	2	P-J	1
N-.	3	D-J	2	S-N	1
N-V	3	R-P	1	A-V	1
V-D	3	N-A	1		
Full Segmentation Set Size				40	

I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Pareto-Optimal Segmentation - No segments

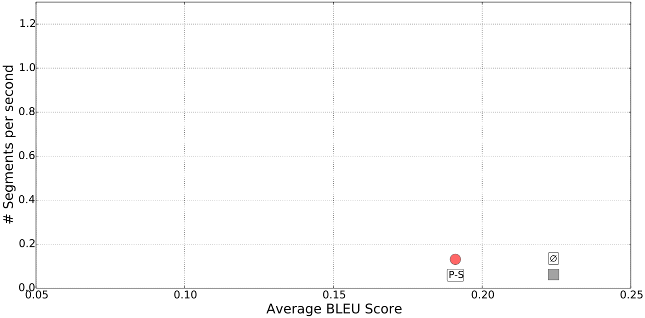


I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Pareto-Optimal Segmentation - One segment

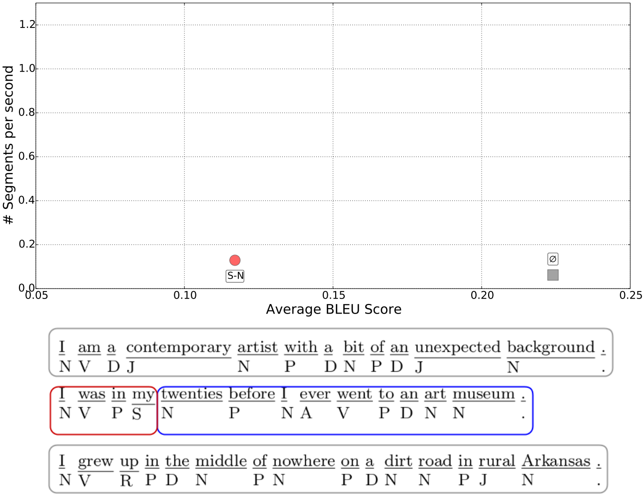


I am a contemporary artist with a bit of an unexpected background .
N V D J N P D N P D J N .

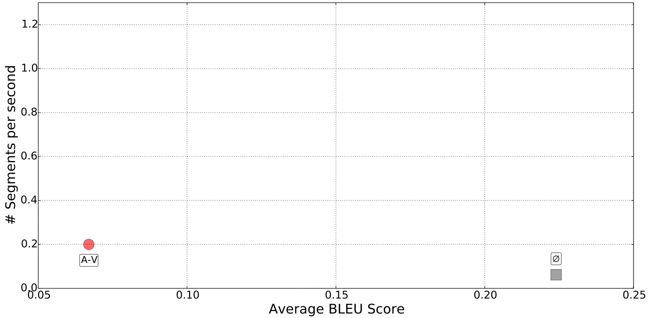
I was in | my twenties before I ever went to an art museum .
N V P | S N P N A V P D N N .

I grew up in the middle of nowhere on a dirt road in rural Arkansas .
N V R P D N P N P D N N P J N .

Pareto-Optimal Segmentation - One segment



Pareto-Optimal Segmentation - One segment

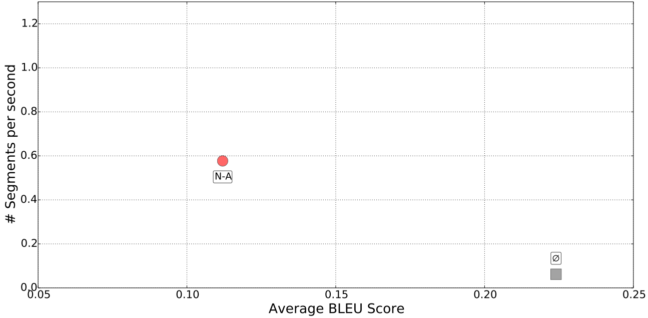


I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Pareto-Optimal Segmentation - One segment

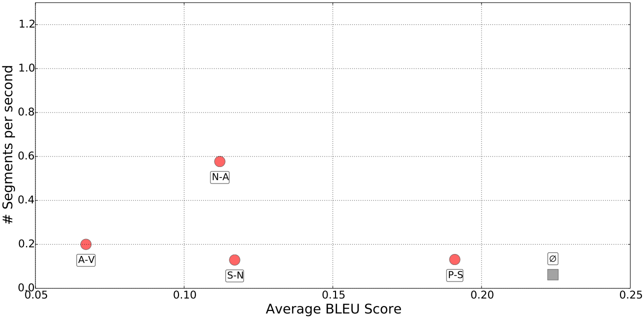


I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

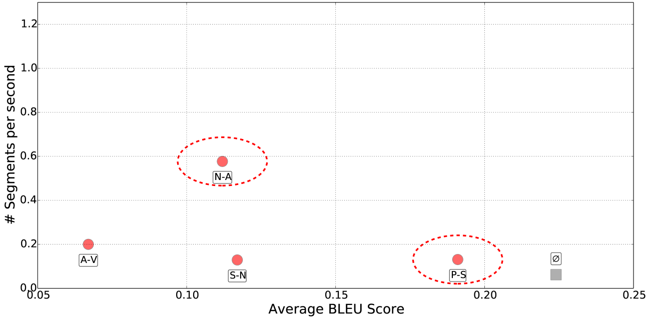
I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

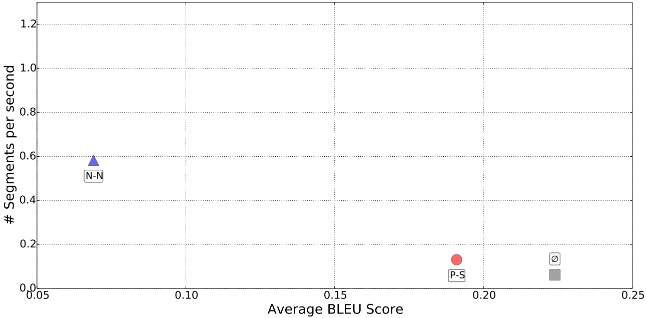
Pareto-Optimal Segmentation - One segment



Pareto-Optimal Segmentation - One segment



Pareto-Optimal Segmentation - Two segments

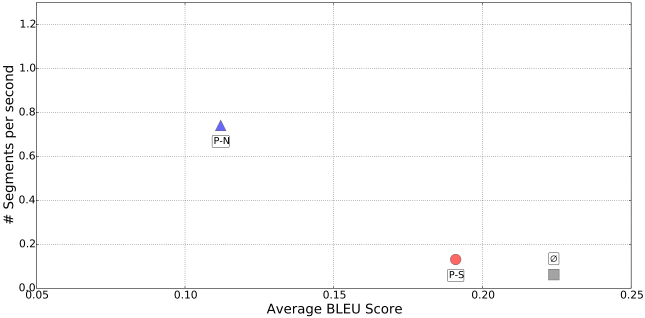


I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Pareto-Optimal Segmentation - Two segments



I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

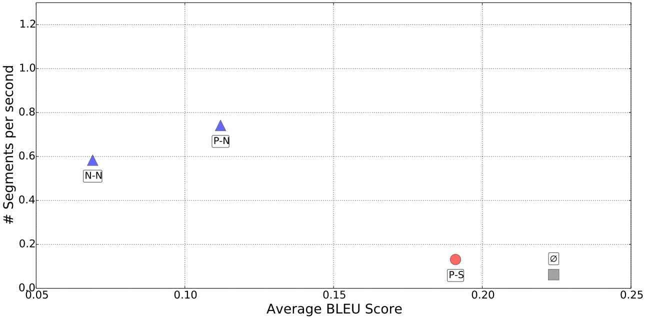
I	was	in	my	twenties	before
N	V	P	S	N	P

I	ever	went	to	an	art	museum	.
N	A	V	P	D	N	N	.

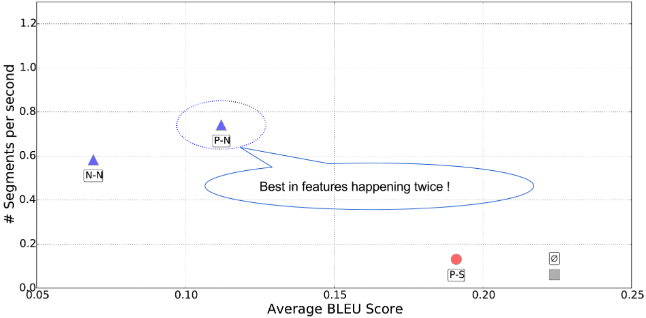
I	grew	up	in	the	middle	of
N	V	R	P	D	N	P

nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	P	D	N	N	P	J	N	.

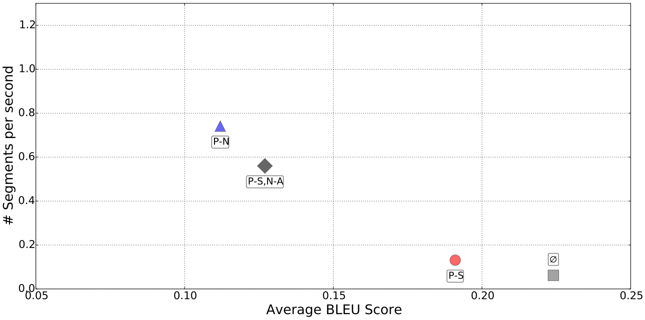
Pareto-Optimal Segmentation - Two segments



Pareto-Optimal Segmentation - Two segments



Pareto-Optimal Segmentation - Two segments

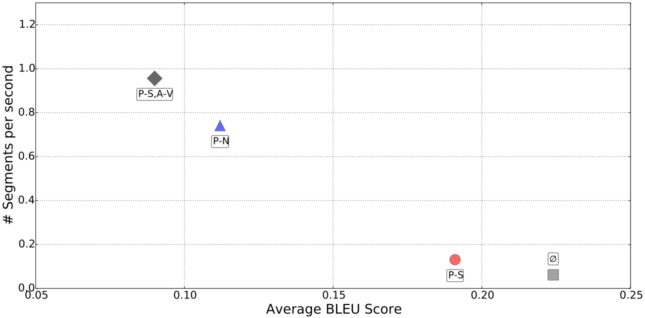


I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Pareto-Optimal Segmentation - Two segments

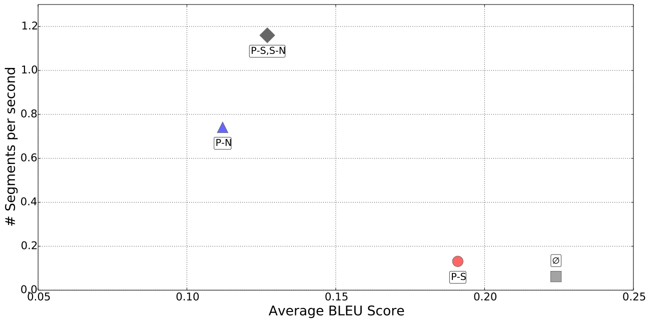


I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Pareto-Optimal Segmentation - Two segments

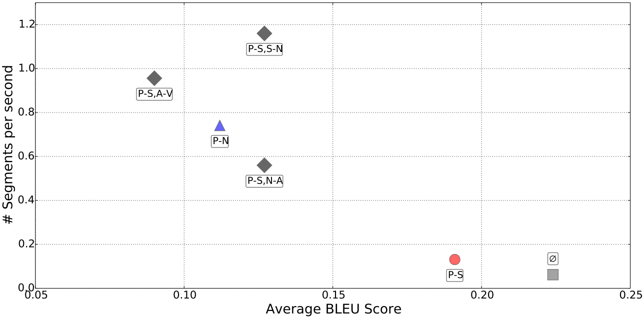


I	am	a	contemporary	artist	with	a	bit	of	an	unexpected	background	.
N	V	D	J	N	P	D	N	P	D	J	N	.

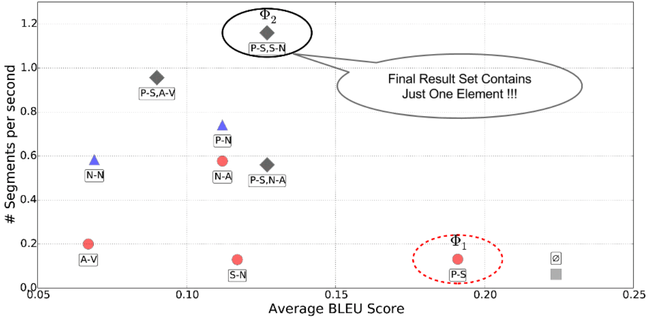
I	was	in	my	twenties	before	I	ever	went	to	an	art	museum	.
N	V	P	S	N	P	N	A	V	P	D	N	N	.

I	grew	up	in	the	middle	of	nowhere	on	a	dirt	road	in	rural	Arkansas	.
N	V	R	P	D	N	P	N	P	D	N	N	P	J	N	.

Pareto-Optimal Segmentation - Two segments



Pareto-Optimal Segmentation - Two segments



Segmentation Evaluation

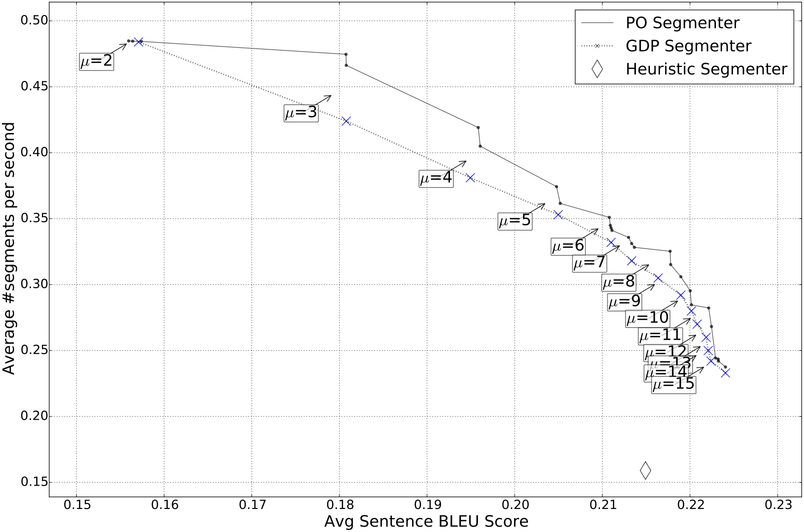
Experimental Setup

- Task: English-German TED speech translation shared task (original task is not simultaneous translation!)
- Segmenter Training Data: IWSLT Dev 2010 and 2012 and Test 2010
- Segmenter Test Data: IWSLT Test 2013
- Segmentation Train Size: 3669 sents
- Segmentation Test Size: 1025 sents

Accuracy vs. Latency - Comparison

- We compared
  - the state-of-the-art prosodic speech segmenter (monolingual) [Rangarajan+ 13; Sridhar+ 13] Heuristic
  - Greedy Segmentation Approach [Oda+ 2014] GDP
  - Pareto-Optimal Segmentation Approach PO

Results on the Test Data



Result comparison for $\mu = 3$ and $\mu = 8$

	$\mu = 3$		$\mu = 8$	
	Segs/Sec	BLEU	Segs/Sec	BLEU
Pareto-Optimal Segmenter	**0.474**	18.07	**0.315**	**21.77**
Greedy Segmenter	0.424	18.07	0.305	21.63

Segmentation Classifier

Classifier based on an alignment heuristic: Align

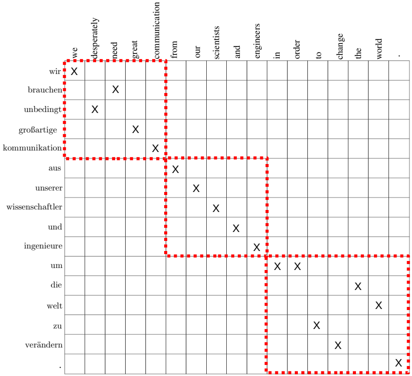


Figure: Word alignment matrix for an English-German sentence. Monotone phrases are shown in dashed lines. Heuristic annotation for $\mu = 5$.

Classifier based on our generated training data: PO

- We created training data for segment boundaries using Pareto optimal search.
- We use this data to build a segment classifier.

Classifier based on our generated training data: PO

- We created training data for segment boundaries using Pareto optimal search.
- We use this data to build a segment classifier.

Feature set	Example
Set1: LastWord, Position, Length	“engineers”, 9, 5
Set2: + Segment POS n-grams	[NNS],[CC-NNS],[NN-CC-NNS]
Set3: + Cross-segment POS tags	[NNS-IN]

Table: For segment “from our scientist and engineers * in”

Classification Results

	F1	Prec	Recall
Align Set1	67.04	62.88	71.78
Align Set2	64.27	56.46	74.58
Align Set3	67.43	58.39	79.78
PO Set1	63.87	52.56	81.38
PO Set2	65.06	53.90	82.03
PO Set3	**81.31**	**69.89**	**97.18**

- Data from IWSLT 2011 (train)
- Data was split into 90% training for Align and 10% test (5K words) for both methods

Incremental Decoding

Translation Data

```
<CHAPTER ID=1>
Wiederaufnahme der Sitzungsperiode
<SPEAKER ID=1 NAME="Die Pr  sidentin">
Ich erkl  re die am Freitag, dem 17. Dezember unterbro
<P>
Wie Sie feststellen konnten, ist der gef  rchtete "Mil
Im Parlament besteht der Wunsch nach einer Aussprache
Heute m  chte ich Sie bitten - das ist auch der Wunsch
Ich bitte Sie, sich zu einer Schweigeminute zu erhebe
<P>
Das Parlament erhebt sich zu einer Schweigeminute.
<P>
<SPEAKER ID=2 LANGUAGE="EN" NAME="Evans, Robert J">
Frau Pr  sidentin, zur Gesch  ftsordnung.
Wie Sie sicher aus der Presse und dem Fernsehen wisse
Zu den Attentatsopfern, die es in j  ngster Zeit in Sr
W  re es angemessen, wenn Sie, Frau Pr  sidentin, der P
<P>
<SPEAKER ID=3 NAME="Die Pr  sidentin">
Ja, Herr Evans, ich denke, da   eine derartige Initiat
Wenn das Haus damit einverstanden ist, werde ich dem
<P>
<SPEAKER ID=4 LANGUAGE="EN" NAME="MacCormick">
Frau Pr  sidentin, zur Gesch  ftsordnung.
K  nnten Sie mir eine Auskunft zu Artikel 143 im Zusam
Meine Frage betrifft eine Angelegenheit, die am Donne
<P>
Das Parlament wird sich am Donnerstag mit dem Cunha-B
Und zwar sollen derartige Strafen trotz des Grundsatz
Ich meine, da   der Grundsatz der relativen Stabilit  t
Ich m  chte wissen, ob es m  glich ist, einen Einwand g
<P>
```

```
<CHAPTER ID=1>
Resumption of the session
<SPEAKER ID=1 NAME="President">
I declare resumed the session of the European Parliam
<P>
Although, as you will have seen, the dreaded 'millenn
You have requested a debate on this subject in the co
In the meantime, I should like to observe a minute' s
Please rise, then, for this minute' s silence.
<P>
The House rose and observed a minute' s silence
<P>
<SPEAKER ID=2 NAME="Evans, Robert J">
Madam President, on a point of order.
You will be aware from the press and television that
One of the people assassinated very recently in Sri L
Would it be appropriate for you, Madam President, to
<P>
<SPEAKER ID=3 NAME="President">
Yes, Mr Evans, I feel an initiative of the type you h
If the House agrees, I shall do as Mr Evans has sugge
<P>
<SPEAKER ID=4 NAME="MacCormick">
Madam President, on a point of order.
I would like your advice about Rule 143 concerning in
My question relates to something that will come up on
<P>
The Cunha report on multiannual guidance programmes c
It says that this should be done despite the principl
I believe that the principle of relative stability is
I want to know whether one can raise an objection of
<P>
```

Translation Data

```
<CHAPTER ID=1>
Wiederaufnahme der Sitzungsperiode
<SPEAKER ID=1 NAME="Die Pr  sidentin">
Ich erkl  re die am Freitag, dem 17. Dezember unterbro
<P>
Wie Sie feststellen konnten, ist der gef  rchtete "Mil
Im Parlament besteht der Wunsch nach einer Aussprache
Heute m  chte ich Sie bitten - das ist auch der Wunsch
Ich bitte Sie, sich zu einer Schweigeminute zu erhebe
<P>
Das Parlament erhebt sich zu einer Schweigeminute.
<P>
<SPEAKER I
Frau Pr  si
Wie Sie si
Zu den Att
W  re es ar
<P>
<SPEAKER I
Ja, Herr Evans, ich denke, da   eine derartige Initiat
Wenn das Haus damit einverstanden ist, werde ich dem
<P>
<SPEAKER ID=4 LANGUAGE="EN" NAME="MacCormick">
Frau Pr  sidentin, zur Gesch  ftsordnung.
K  nnten Sie mir eine Auskunft zu Artikel 143 im Zusam
Meine Frage betrifft eine Angelegenheit, die am Donne
<P>
Das Parlament wird sich am Donnerstag mit dem Cunha-B
Und zwar sollen derartige Strafen trotz des Grundsatz
Ich meine, da   der Grundsatz der relativen Stabilit  t
Ich m  chte wissen, ob es m  glich ist, einen Einwand g
<P>
```

```
<CHAPTER ID=1>
Resumption of the session
<SPEAKER ID=1 NAME="President">
I declare resumed the session of the European Parliam
<P>
Although, as you will have seen, the dreaded 'millenn
You have requested a debate on this subject in the co
In the meantime, I should like to observe a minute' s
Please rise, then, for this minute' s silence.
<P>
The House rose and observed a minute' s silence
<P>
elevision that
cently in Sri L
 President, to
Yes, Mr Evans, I feel an initiative of the type you h
If the House agrees, I shall do as Mr Evans has sugge
<P>
<SPEAKER ID=4 NAME="MacCormick">
Madam President, on a point of order.
I would like your advice about Rule 143 concerning in
My question relates to something that will come up on
<P>
The Cunha report on multiannual guidance programmes c
It says that this should be done despite the principl
I believe that the principle of relative stability is
I want to know whether one can raise an objection of
<P>
```

Parallel Text:

(Web, United Nations, European/Canadian Parliament, Wikipedia, etc.)

Statistical Machine Translation

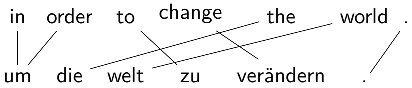


Figure: Learn alignments from parallel text

Statistical Machine Translation

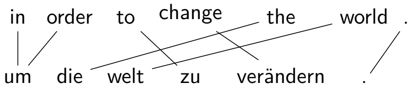


Figure: Learn alignments from parallel text

Id	Source	Target	Weight
r_1	in order	um	-5.3
r_2	X_1 the world X_2	die welt X_1 X_2	-2.8
r_3	to change	verändern	-3.1

Figure: Learn weighted translation rules from parallel text

Statistical Machine Translation

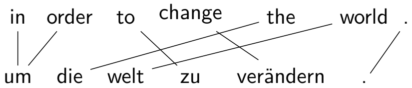


Figure: Learn alignments from parallel text

Id	Source	Target	Weight
r_1	in order	um	-5.3
r_2	X_1 the world X_2	die welt X_1 X_2	-2.8
r_3	to change	verändern	-3.1

Figure: Learn weighted translation rules from parallel text

$$\mathbf{t}^* = \operatorname*{argmax}_{\mathbf{t} \in \mathcal{Y}(\mathbf{d})} \sum_{r \in d} \mathbf{w} \cdot \mathbf{f}(\mathbf{r})$$

Statistical Machine Translation

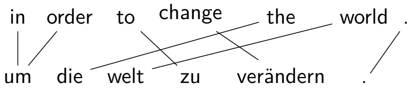


Figure: Learn alignments from parallel text

Id	Source	Target	Weight
r_1	in order	um	-5.3
r_2	X_1 the world X_2	die welt X_1 X_2	-2.8
r_3	to change	verändern	-3.1

Figure: Learn weighted translation rules from parallel text

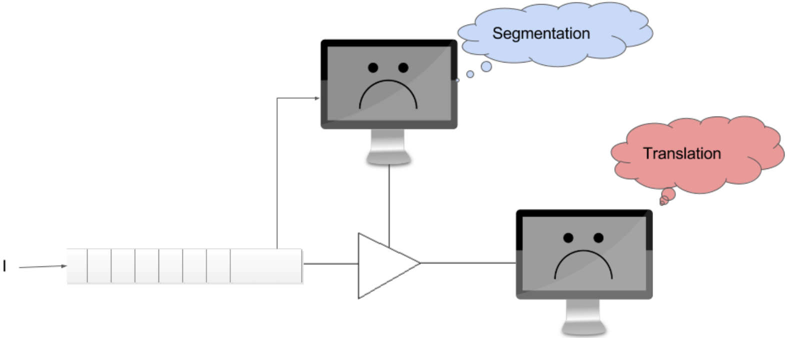
$$\mathbf{t}^* = \operatorname*{argmax}_{\mathbf{t} \in \mathcal{Y}(\mathbf{d})} \sum_{r \in d} \mathbf{w} \cdot \mathbf{f}(\mathbf{r})$$

- Exponential time $\Rightarrow$ CKY dynamic programming $\mathcal{O}(n^3)$
- Our algorithm: **Earley** style decoding $\mathcal{O}(n^2 b)$
- `https://github.com/sfu-natlang/lrhiero`

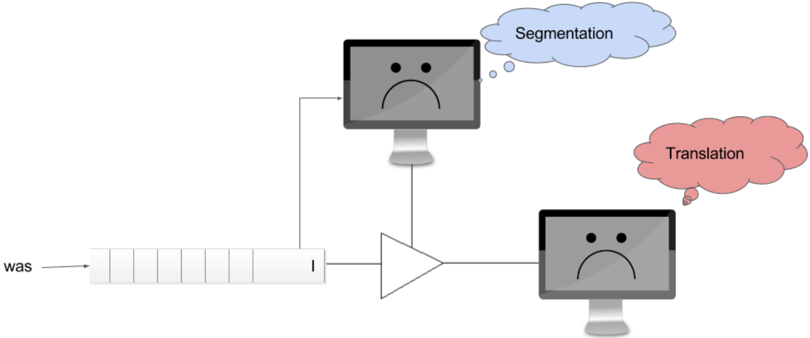
Figure: Decoder: produces most likely translation

Segmentation Classifier Integrated with Decoder

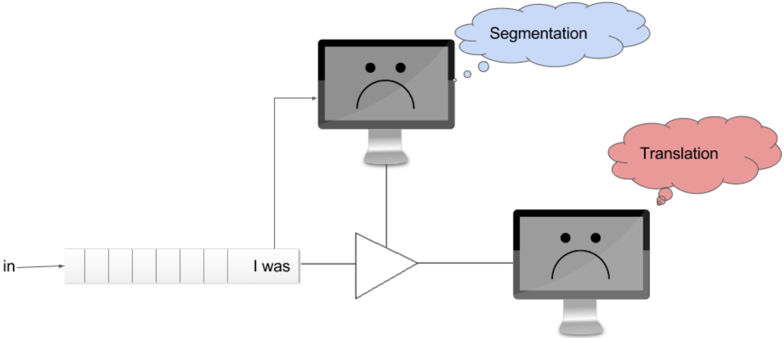
Segmentation Classifier Integrated with Decoder



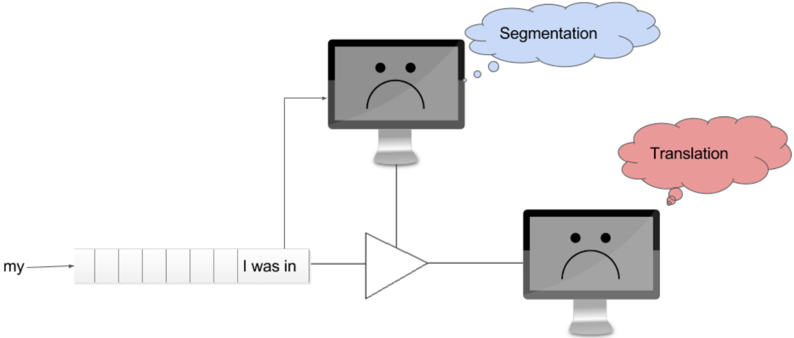
Segmentation Classifier Integrated with Decoder



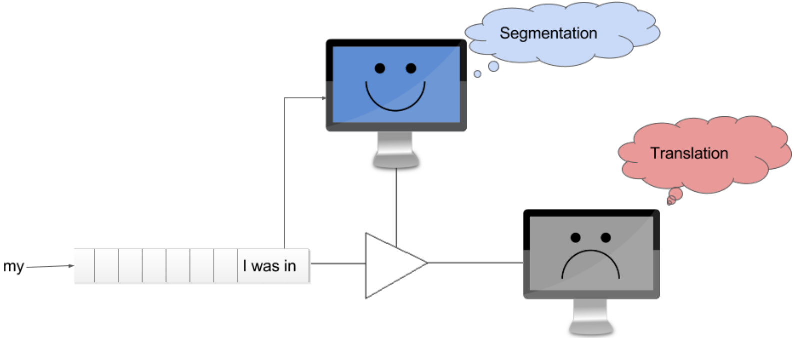
Segmentation Classifier Integrated with Decoder



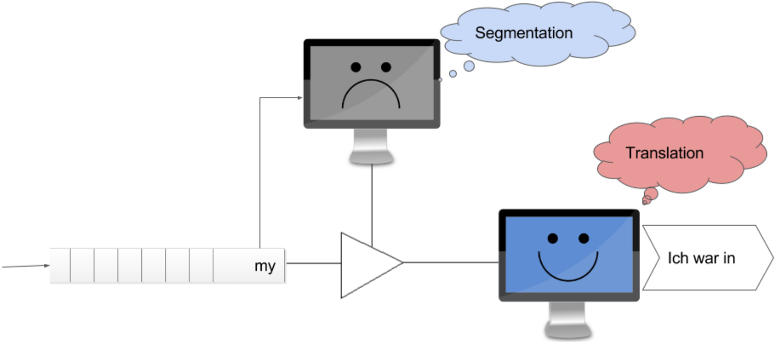
Segmentation Classifier Integrated with Decoder



Segmentation Classifier Integrated with Decoder



Segmentation Classifier Integrated with Decoder



Incremental Decoder Evaluation

Experimental Setup

Translation data

- Task: English-German TED talks translation
- Train: IWSLT 2013 Train data + Europarl v7 data [Koehn 2005]
- Tuning: IWSLT Test 2012
- German Language Model: WMT 2013 Shared Task

Segmenter data

- Train: IWSLT Dev 2010 and 2012 and Test 2010 (3669 sentences)
- Test: IWSLT Test 2013 (1025 sentences)

Incremental Decoder Results

	Num segs	BLEU	Latency (time/segs)	Segs/second
Heuristic (Sridhar+ 13)	2709	**20.88**	0.468	2.27
Align	2654	20.62	0.524	1.96
PO Set1	3608	19.97	0.274	4.07
PO Set2	2777	20.74	0.466	2.35
PO Set3	**3471**	***20.70***	**0.270**	**3.73**

Summary

We improve the state of the art in simultaneous machine translation by providing:

- A choice between latency and translation quality using Pareto optimality

Summary

We improve the state of the art in simultaneous machine translation by providing:

- A choice between latency and translation quality using Pareto optimality
- A new dynamic programming algorithm for segment annotation

Summary

We improve the state of the art in simultaneous machine translation by providing:

- A choice between latency and translation quality using Pareto optimality
- A new dynamic programming algorithm for segment annotation
- Segmentation annotated data used to train a segmentation classifier

Summary

We improve the state of the art in simultaneous machine translation by providing:

- A choice between latency and translation quality using Pareto optimality
- A new dynamic programming algorithm for segment annotation
- Segmentation annotated data used to train a segmentation classifier
- A new simultaneous translation decoder that uses our segmentation classifier

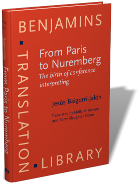
Summary

We improve the state of the art in simultaneous machine translation by providing:

- A choice between latency and translation quality using Pareto optimality
- A new dynamic programming algorithm for segment annotation
- Segmentation annotated data used to train a segmentation classifier
- A new simultaneous translation decoder that uses our segmentation classifier
- Significant improvement in latency with the same quality

The Paris Peace Conference 1919

Birth of multilingual (human) simultaneous translation



- To avoid "such a confusion of tongues that it will be ridiculous" (Nuremberg trial judge)
- But even now: few can afford interpretation services
- Everyone should have access to simultaneous translation!

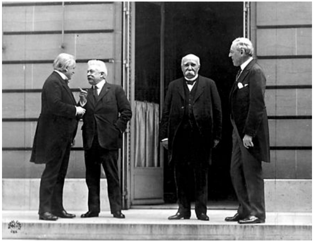


Figure: from left to right, David Lloyd George of Britain, Vittorio Emanuele Orlando of Italy, Georges Clemenceau of France, Woodrow Wilson of the U.S.

These interpreters have a language of their own. We are completely in their hands.

— Stalin to Anthony Eden, Moscow 1943
in Birse 1967, 144

Collaborators

Maryam Siahbani,
ex-PhD student

Hassan Shavarani,
ex-MSc student

Future Work

- Compare against human interpreter output (EPIC corpus)
- Use Pareto optimal points on demand in the decoder
- Improve the scores for translation quality and latency
- Extend encoder-decoder recurrent neural networks
    - Use encoder to predict future input tokens in incremental decoding (predict the clause-final predicate)

Future Work

- Compare against human interpreter output (EPIC corpus)
- Use Pareto optimal points on demand in the decoder
- Improve the scores for translation quality and latency
- Extend encoder-decoder recurrent neural networks
  - Use encoder to predict future input tokens in incremental decoding (predict the clause-final predicate)
  - Multiple decoder stages while encoding the input

Future Work

- Compare against human interpreter output (EPIC corpus)
- Use Pareto optimal points on demand in the decoder
- Improve the scores for translation quality and latency
- Extend encoder-decoder recurrent neural networks
  - Use encoder to predict future input tokens in incremental decoding (predict the clause-final predicate)
  - Multiple decoder stages while encoding the input
  - Integrate training of segmenter with translation model

Fin