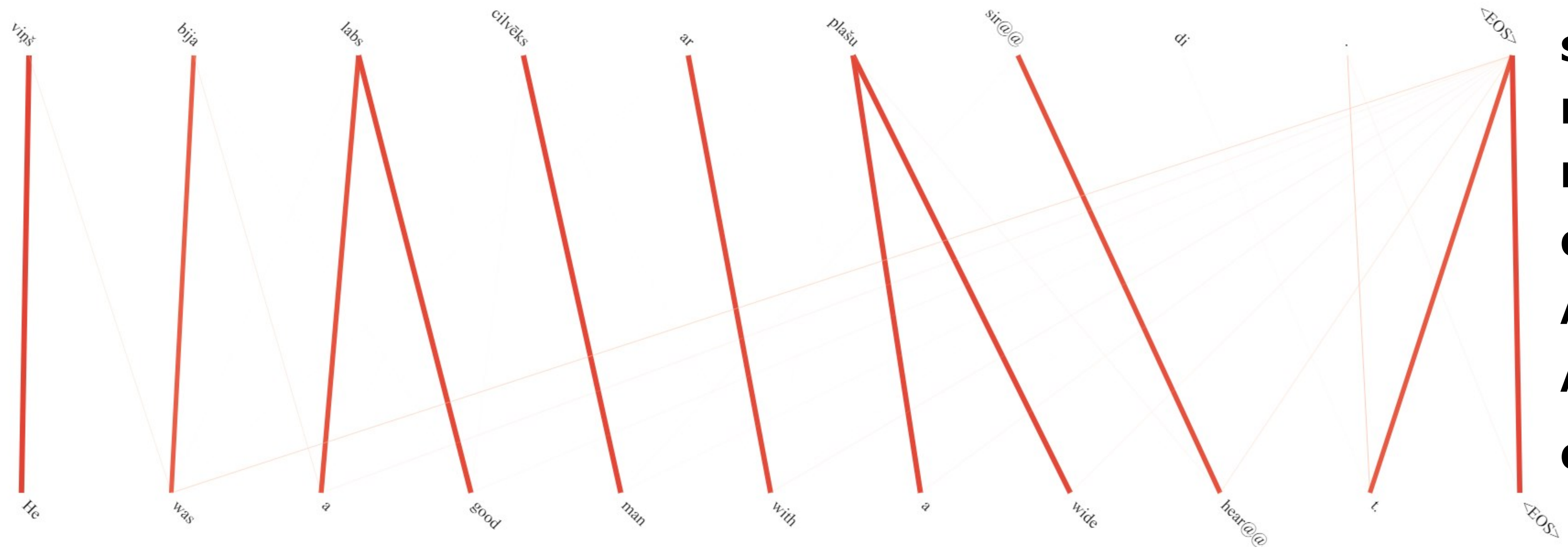
Confidence through Attention

Matīss Rikters
Faculty of Computing
University of Latvia
matiss@lielakeda.lv

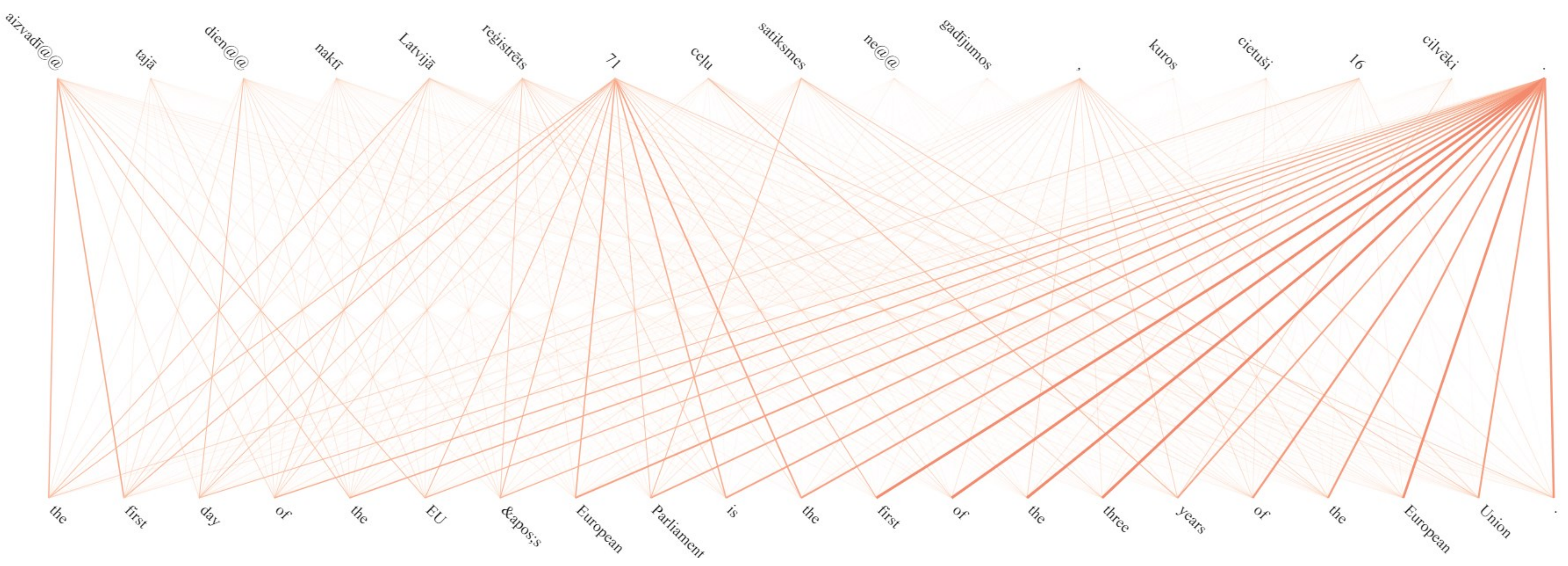
Mark Fishel
Institute of Computer Science
University of Tartu
fishel@ut.ee

Attention Alignments



Source	Viņš bija labs cilvēks ar plašu sirdi.
Reference	He was a kind spirit with a big heart.
Hypothesis	He was a good man with a wide heart.
CDP	-0.099
AP$_{out}$	-1.077
AP$_{in}$	-0.847
Confidence	-2.024

Source	Aizvadītajā diennaktī Latvijā reģistrēts 71 ceļu satiksmes negadījumos, kuros cietuši 16 cilvēki.
Reference	71 traffic accidents in which 16 persons were injured have happened in Latvia during the last 24 hours.
Hypothesis	The first day of the EU' European Parliament is the first of the three years of the European Union .
CDP	-0.900
AP$_{out}$	-2.809
AP$_{in}$	-2.137
Confidence	-5.846



Confidence Scores

$$CDP = \frac{1}{J}\sum_{j}\log\left(1+\left(\sum_{i}\propto_{ji}\right)^{2}\right)$$

$$AP_{out} = -\frac{1}{I}\sum_{i}\sum_{j}\propto_{ji}\cdot \log \propto_{ji}$$

$$AP_{in} = -\frac{1}{I}\sum_{j}\sum_{i}\propto_{ij}\cdot \log \propto_{ij}$$

$$confidence = CDP + AP_{out} + AP_{in}$$

$\{J, I\}$ - source sentence length; i - output token index; j - input token index; α - attention weight

Experimental Settings

Filtered Synthetic Training Data

- Train baseline NMT systems
- Translate 4 million monolingual news sentences of each source language
- Obtain a confidence score for each of the translated sentences; drop the worst 50%
- Train the final NMT system with the remaining 50% added to parallel data

Hybrid System Combination

- Translate the same sentence with two different NMT systems
- Use the translation with the highest confidence score as te final output

Kendall's Tau Correlation

Language pair	CDP	AP_in	AP_out	Overall
En->Lv	0.099	0.074	0.123	0.086
Lv->En	-0.012	-0.153	-0.2	-0.153

Human Judgment Overlap

	En->Lv	Lv->En
LM-based overlap with human	58%	56%
Attention-based overlap with human	52%	60%
LM-based overlap with Attention-based	34%	22%

NMT with Differently Filtered Back-translated Data

	BLEU							
Dataset	Dev	Test	Dev	Test	Dev	Test	Dev	Test
System	En->Lv		Lv->En		En->De		De->En	
Baseline	8.36	11.90	8.64	12.40	25.84	20.11	30.18	26.26
+ Full Synthetic	9.42	13.50	9.01	13.81	28.97	22.68	34.82	29.35
+ LM-Filtered Synthetic	9.75	13.52	9.45	14.30	29.59	23.48	34.47	29.42
+ Attn.-Filtered Synth.	8.99	12.76	11.23	14.83	30.19	23.16	35.19	29.47

GitHub Poster

Acknowledgements

LATVIJAS UNIVERSITATE
ANNO 1919
UNIVERSITY OF LATVIA

TARTU ÜLIKOOL
UNIVERSITAS TARTUENSIS
1632

Hybrid Selections

	BLEU			
System	En->De	De->En	En->Lv	Lv->En
Neural Monkey	18.89	26.07	13.74	11.09
Nematus	22.35	30.53	13.80	12.64
Hybrid	20.19	27.06	14.79	12.65
Human	23.86	34.26	15.12	13.24