Neural Lattice Search for Domain Adaptation in Machine Translation

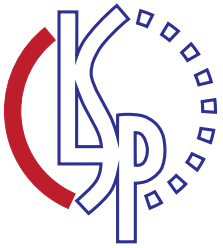
Huda Khayrallah, Gaurav Kumar
Kevin Duh, Matt Post, Philipp Koehn

This talk was presented at IJCNLP 2017
It is based on this paper:
http://aclweb.org/anthology/I17-2004
bib:
http://aclweb.org/anthology/I17-2004.bib

Neural Lattice Search for Domain Adaptation in Machine Translation

Huda Khayrallah, Gaurav Kumar
Kevin Duh, Matt Post, Philipp Koehn

JOHNS HOPKINS
UNIVERSITY



combine

adequacy of PBMT

with

fluency of NMT

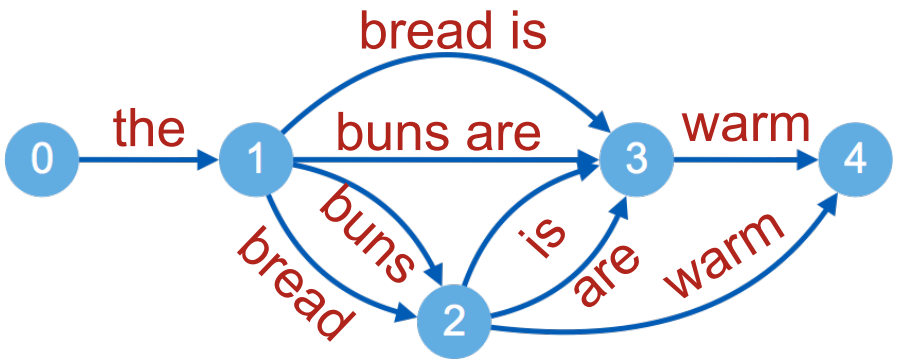
use PBMT
to constrain the search space
of NMT

Source

die
brötchen
sind warm

PBMT

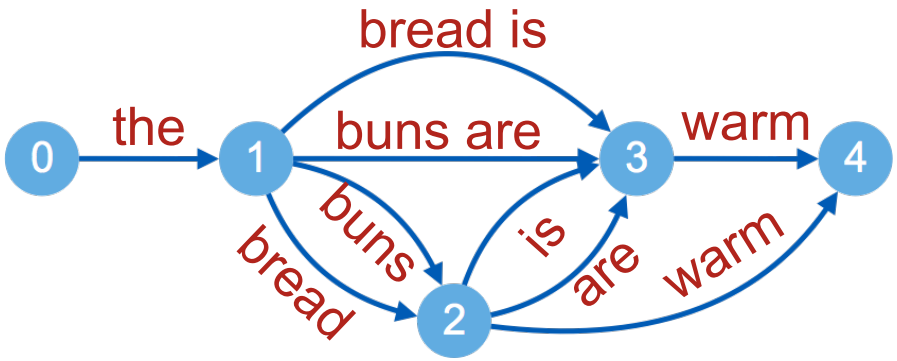
Lattice



Source

die brötchen sind warm

Lattice

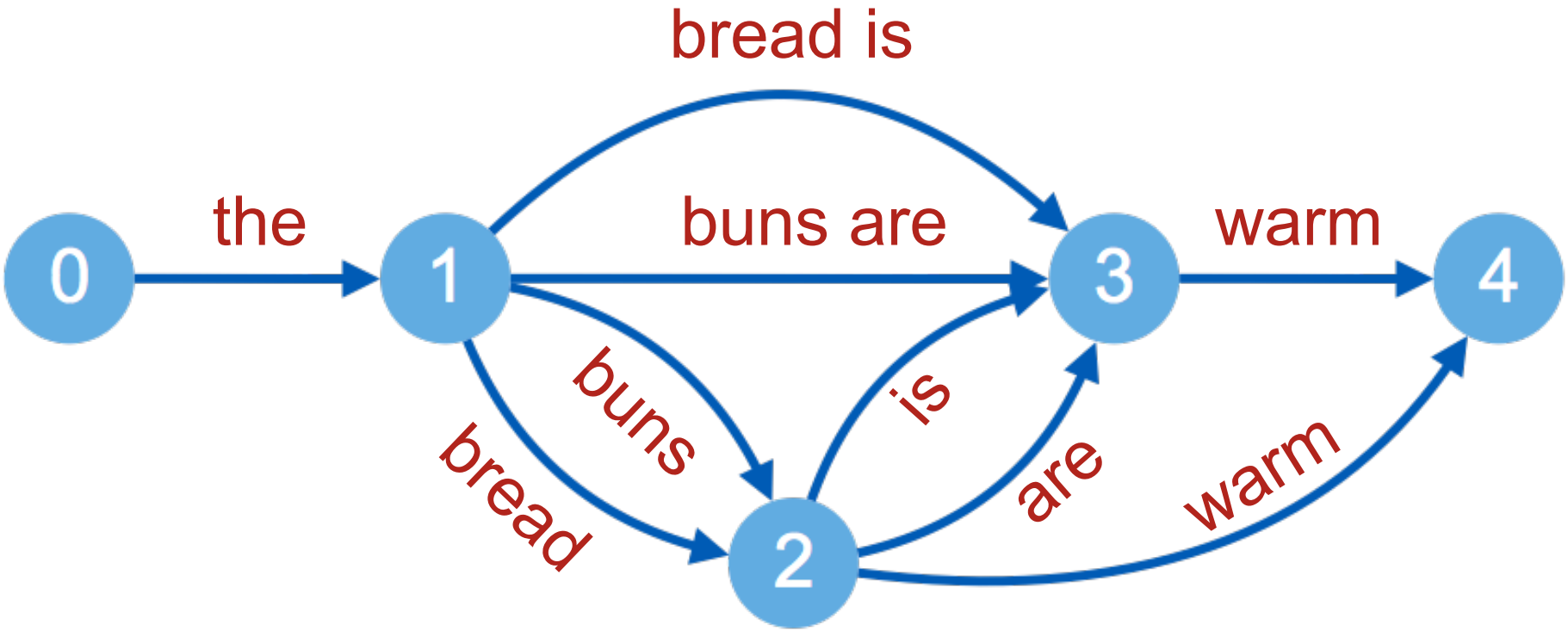


Neural Lattice Search

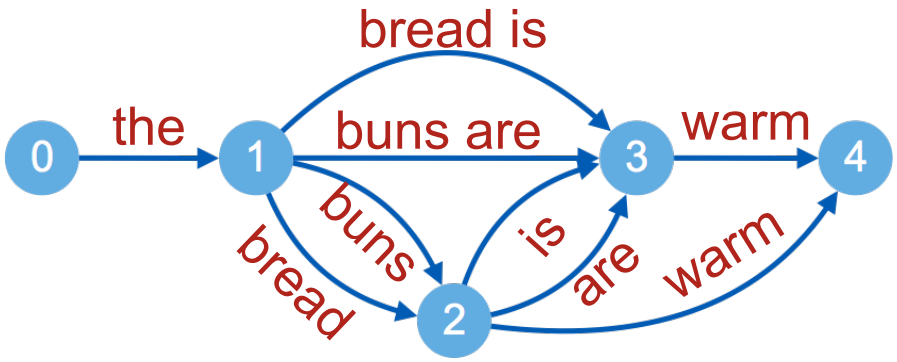
Target

the buns are warm

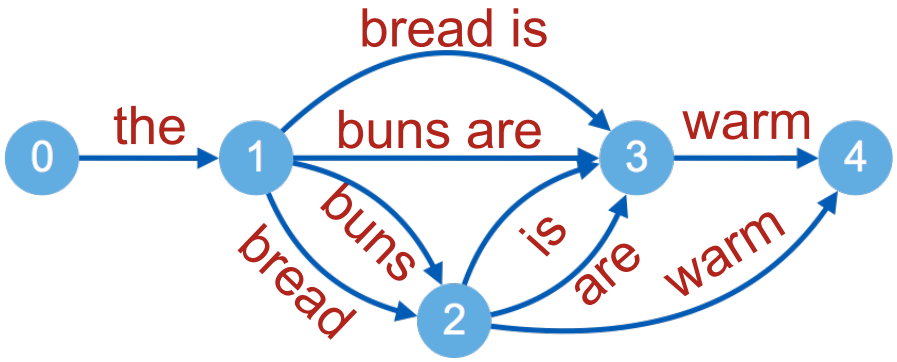
die brötchen sind warm *the buns are warm*

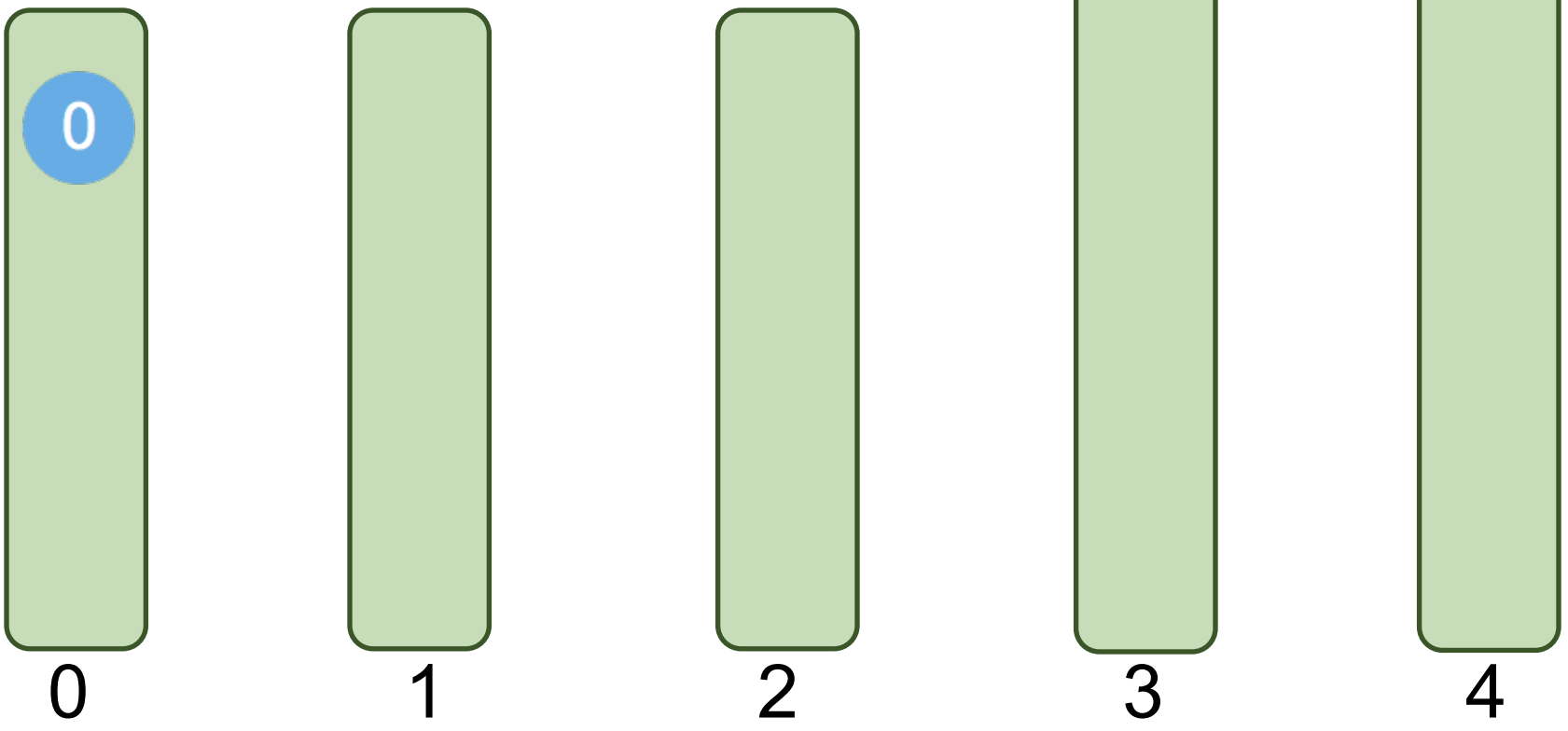


die brötchen sind warm

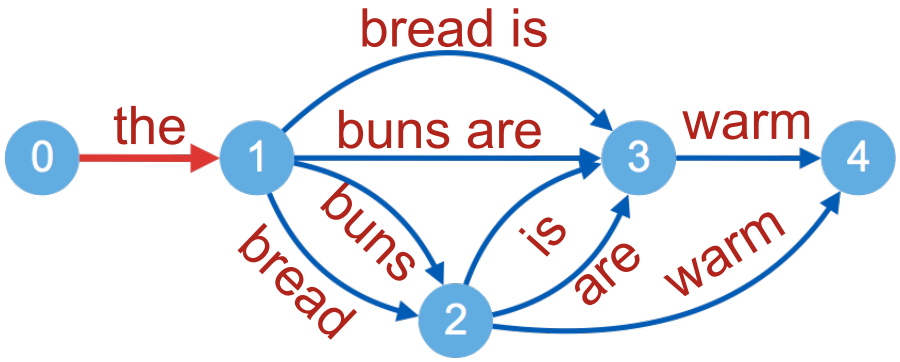


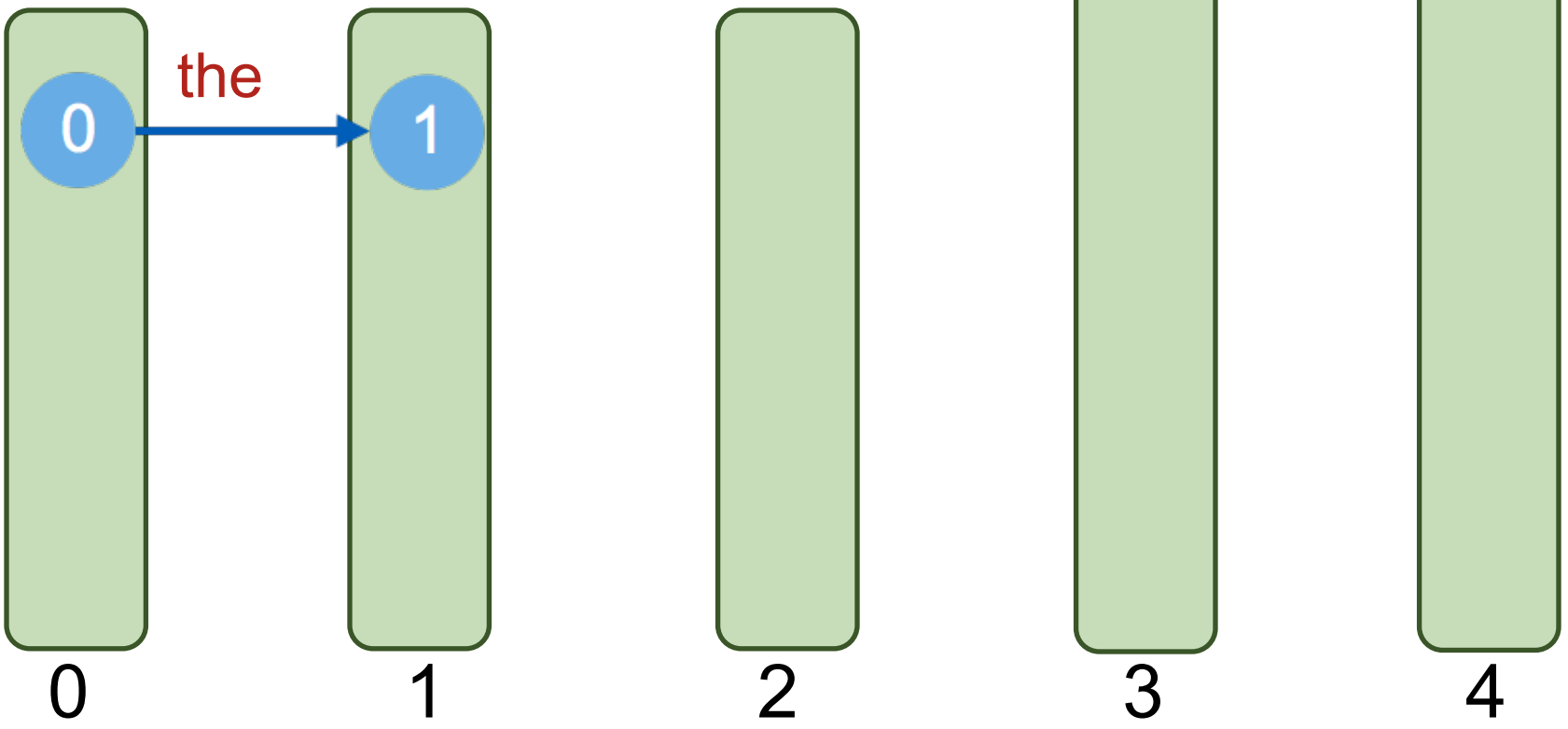
die brötchen sind warm



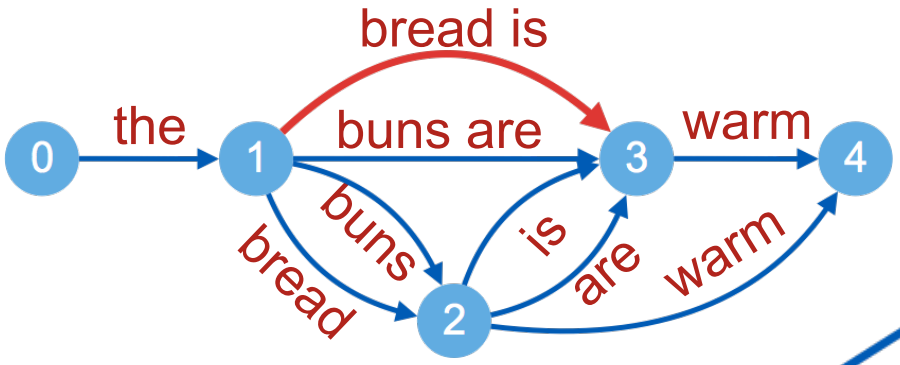


die brötchen sind warm

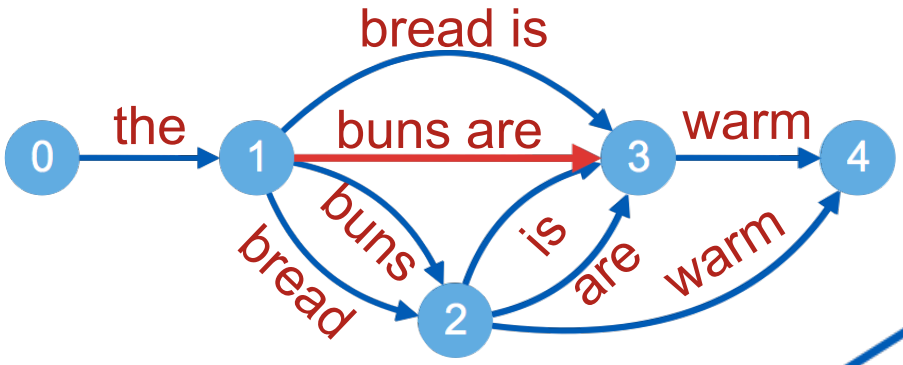




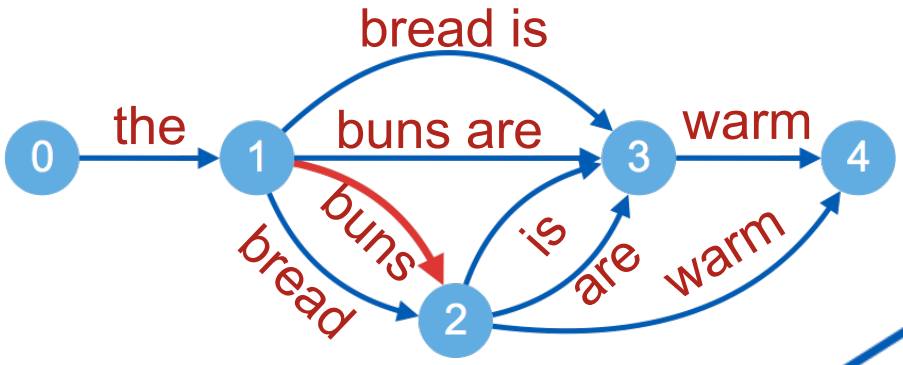
die brötchen sind warm



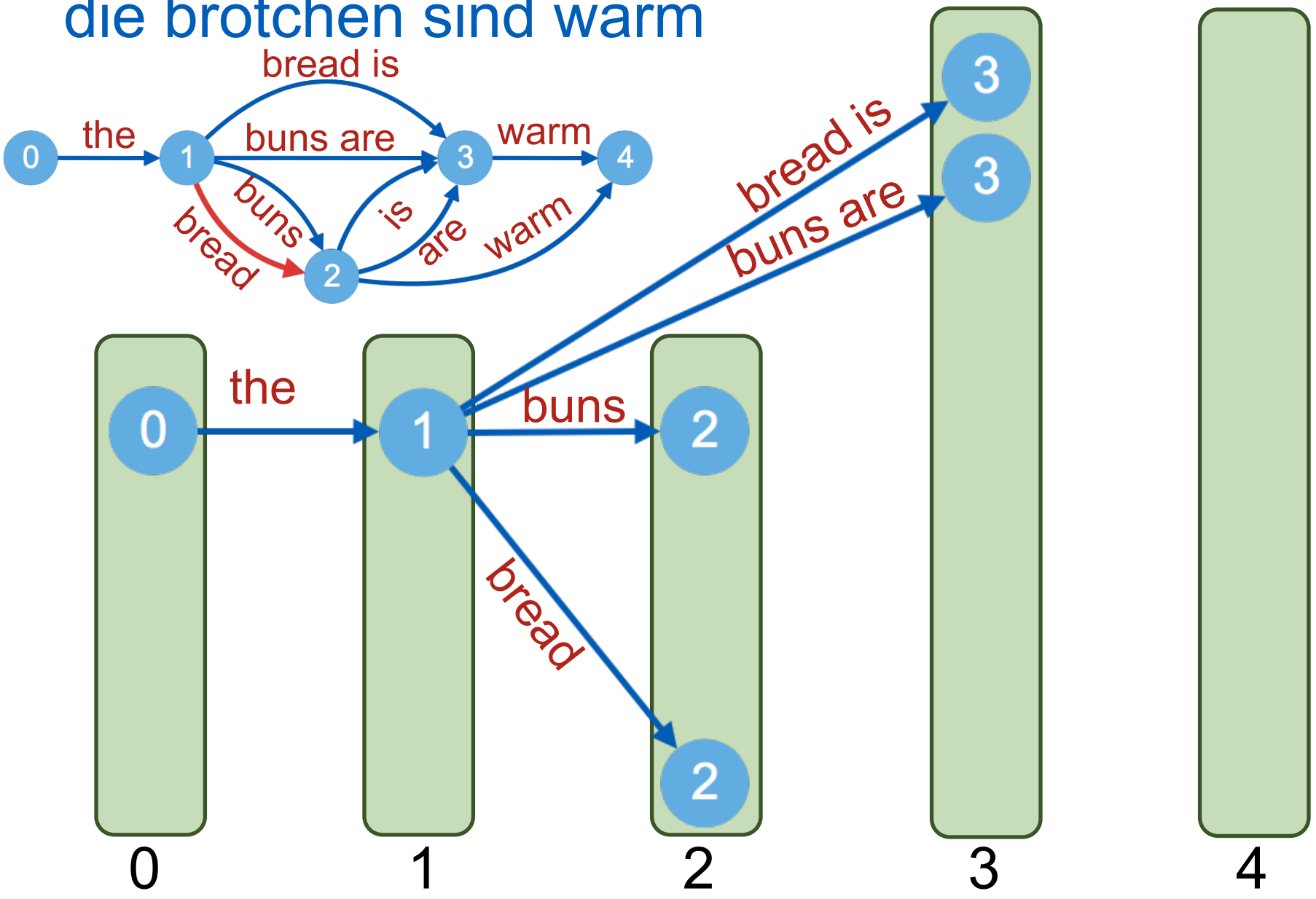
die brötchen sind warm



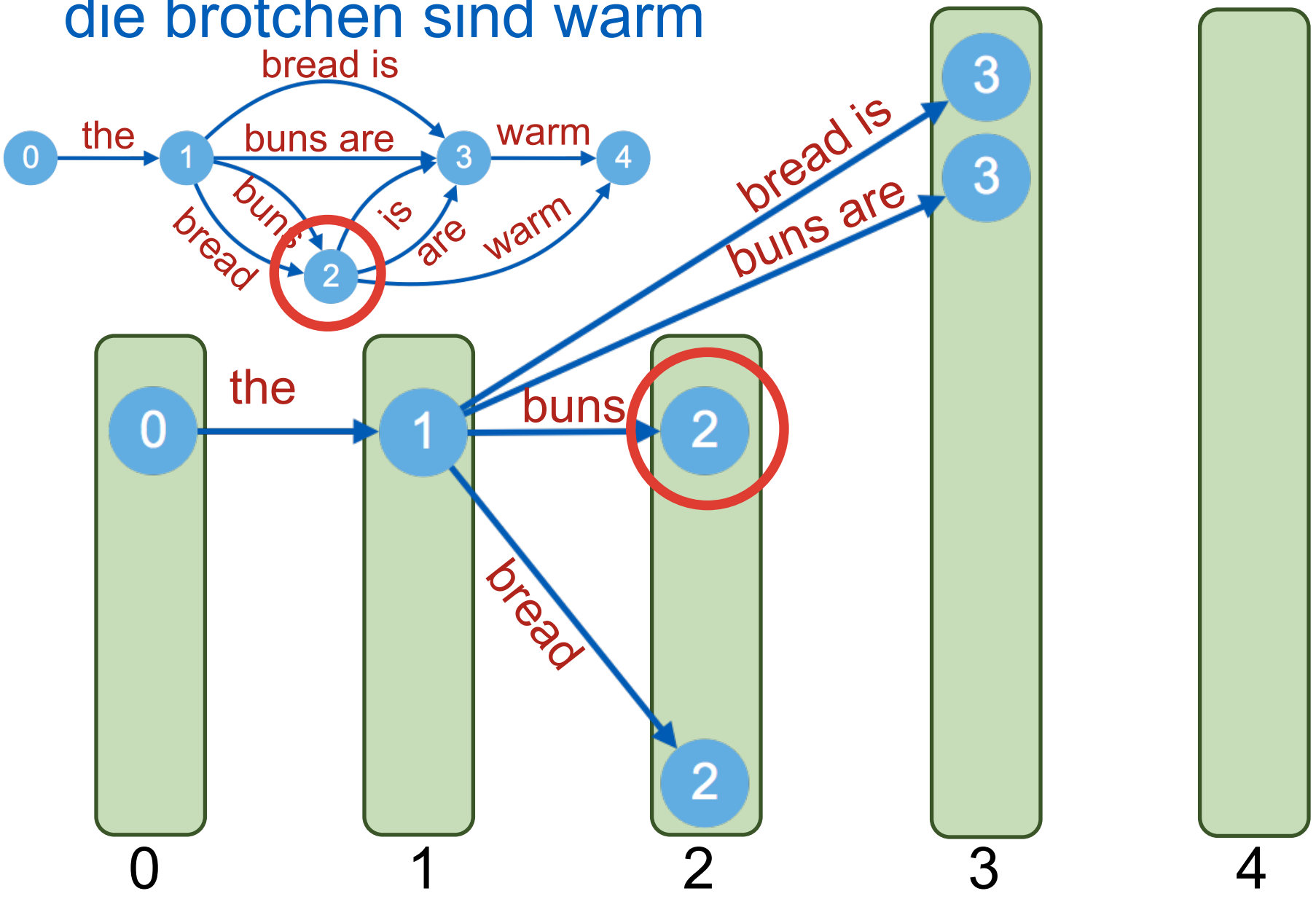
die brötchen sind warm



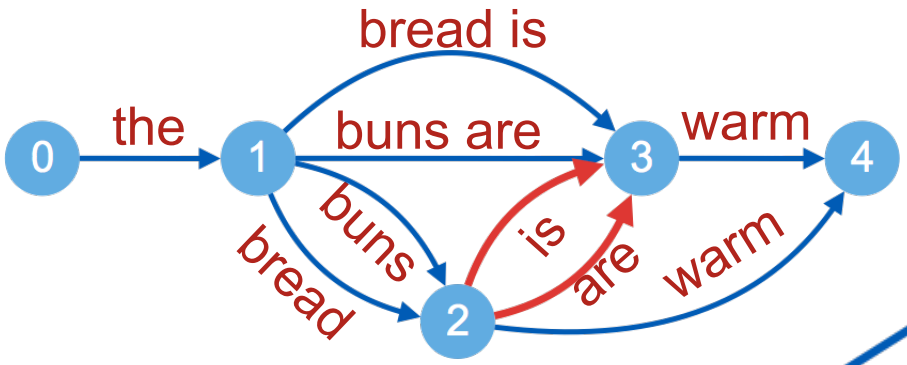
die brötchen sind warm



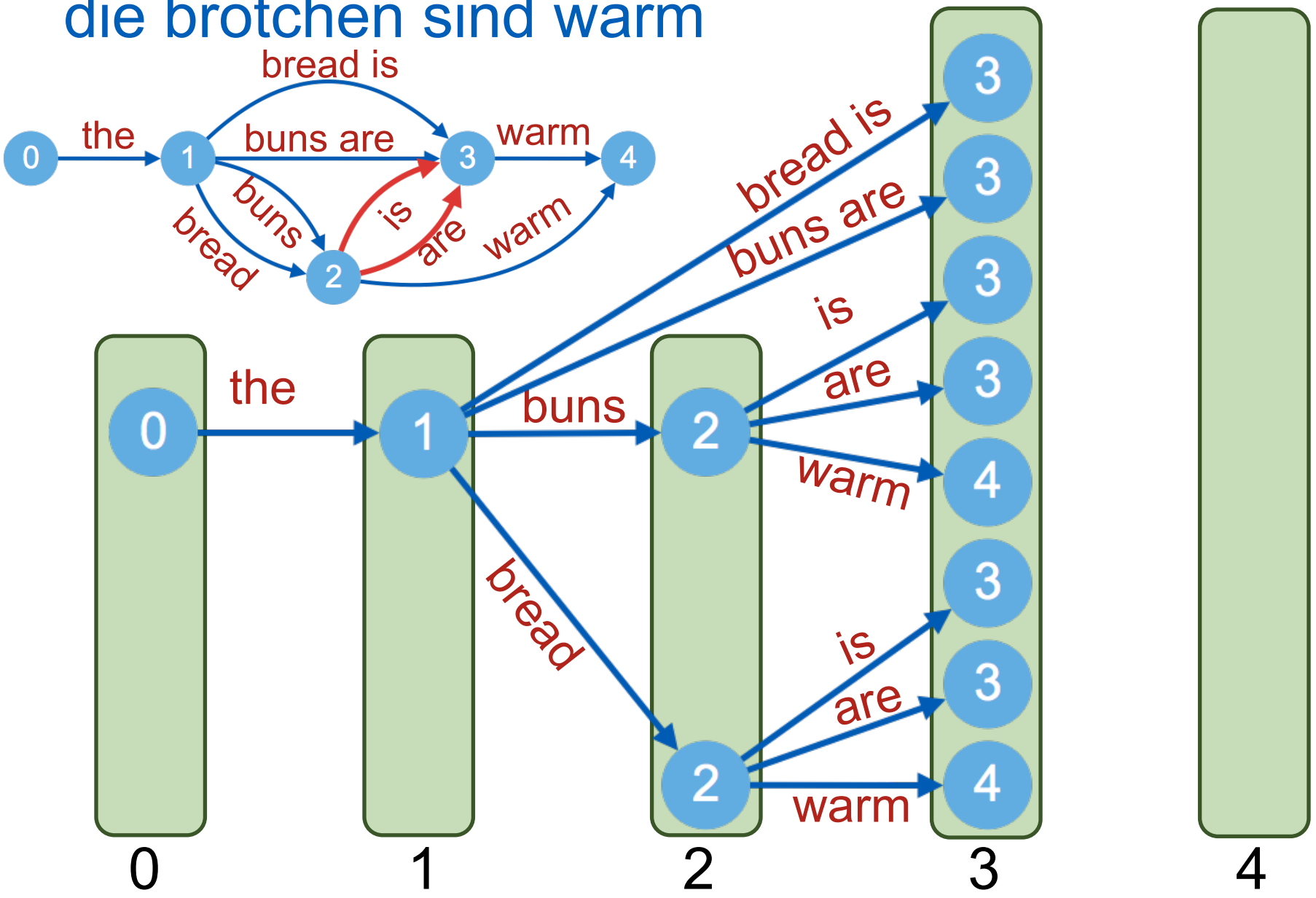
die brötchen sind warm



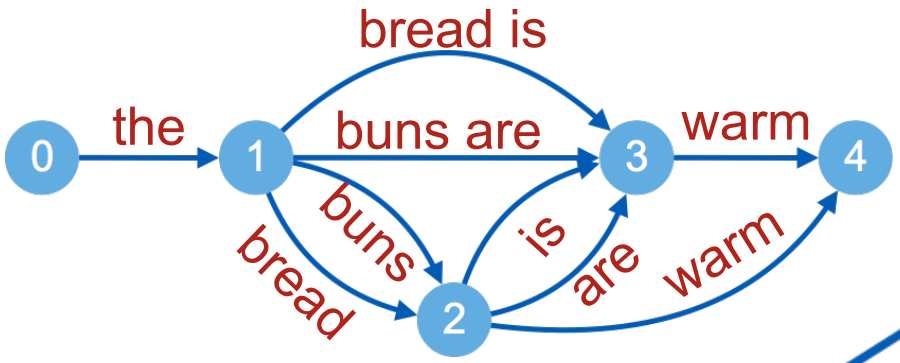
die brötchen sind warm



die brötchen sind warm



die brötchen sind warm



Experiments

Setting: Domain adaptation

Small **in-domain**

IT, Medical, Koran, Subtitles

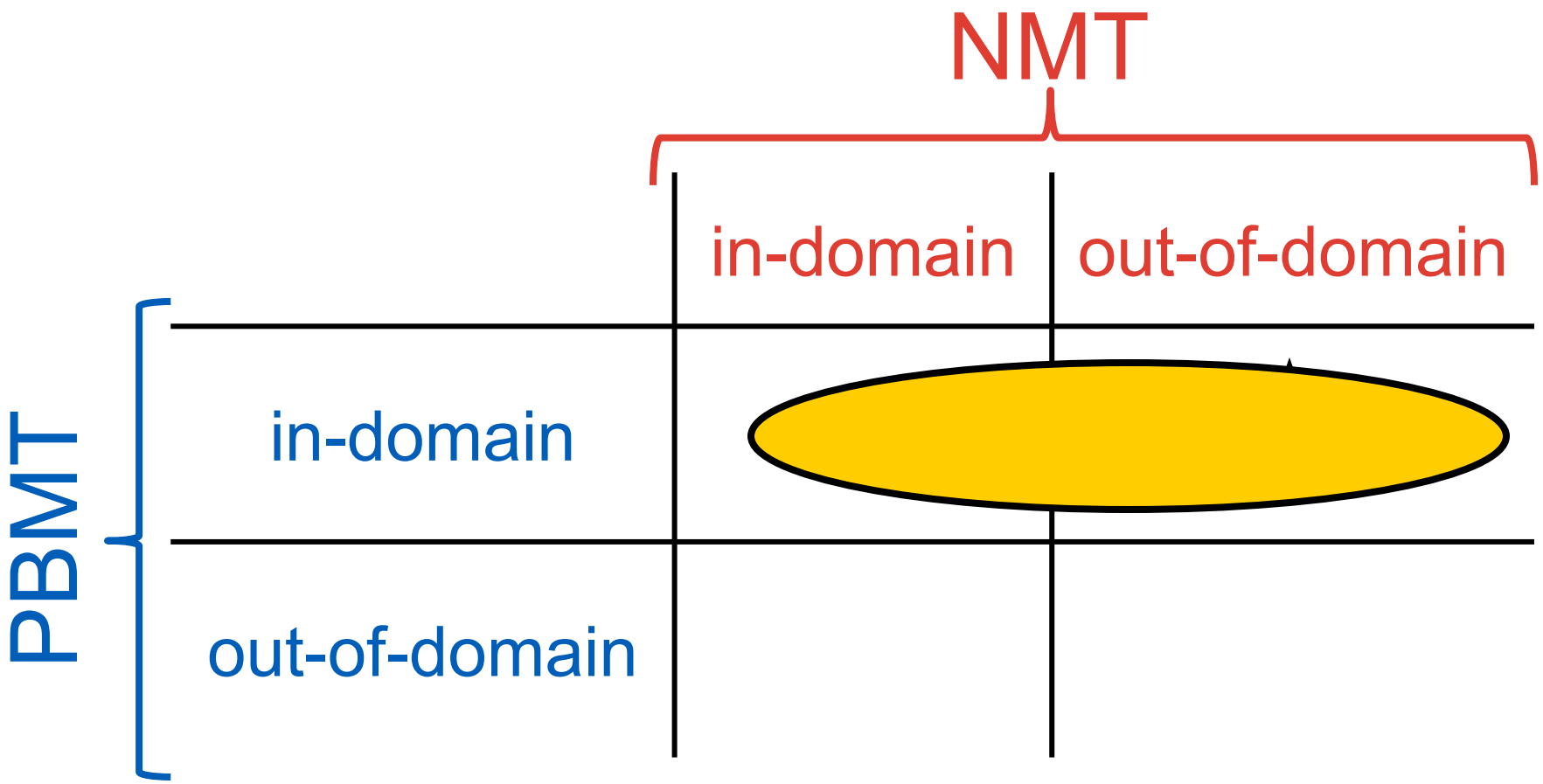
PBMT outperforms NMT

Large **out-of-domain**

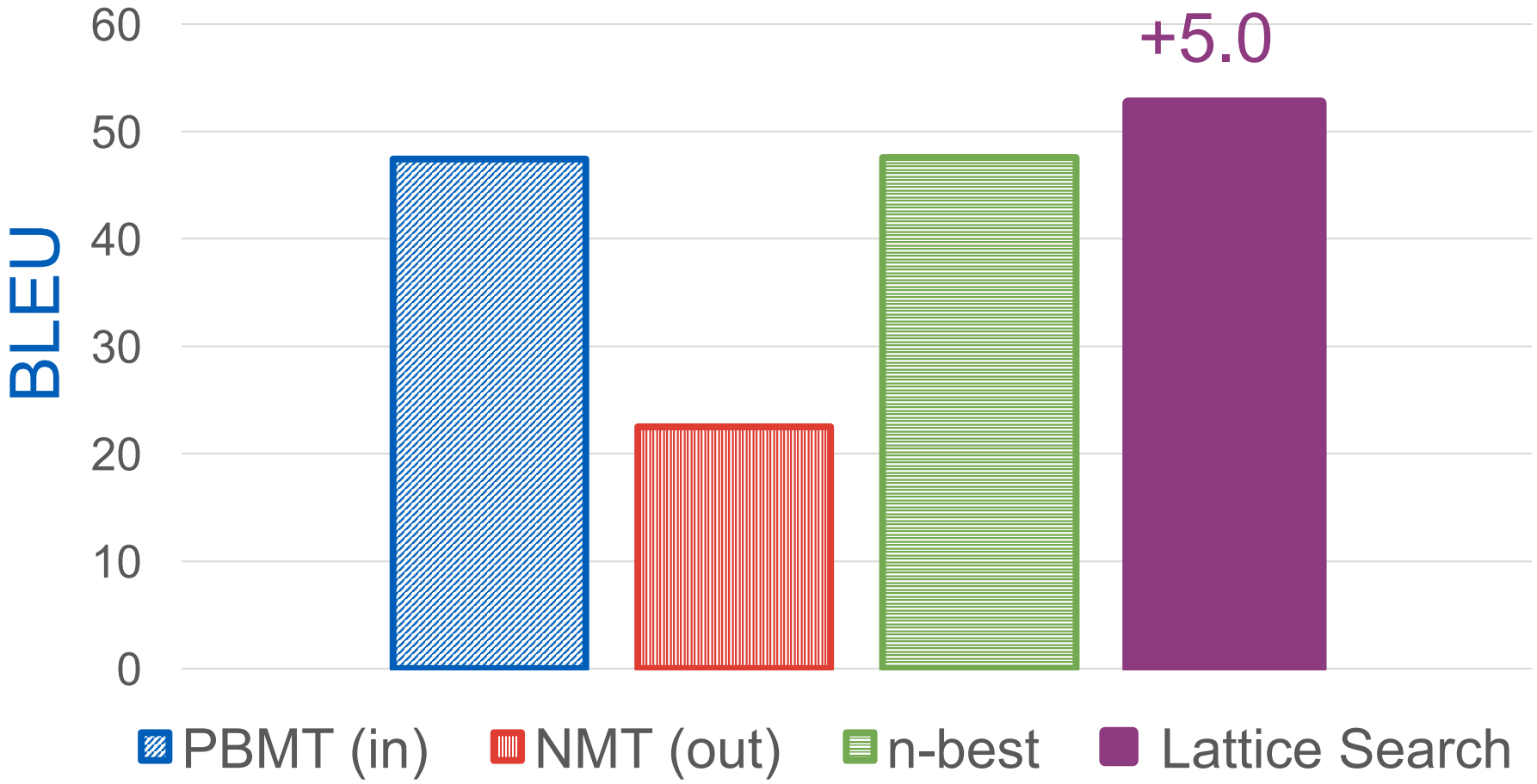
parliamentary proceedings (WMT)

NMT outperforms PBMT

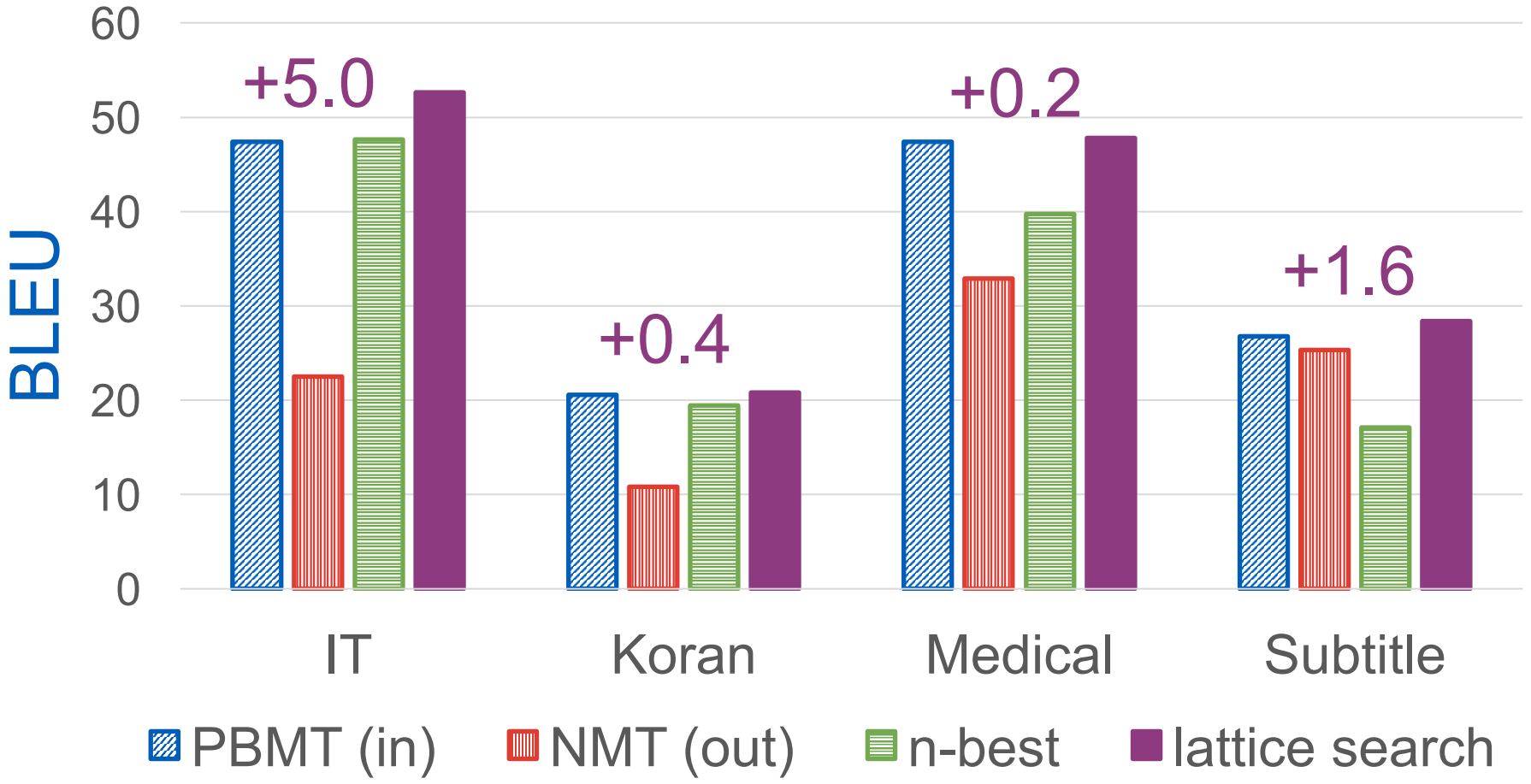
Setting: Domain adaptation



IT Results



Results


BLEU
60
50
40
30
20
10
0
+5.0
+0.4
+0.2
+1.6
IT
Koran
Medical
Subtitle
PBMT (in)
NMT (out)
n-best
lattice search

Conclusion

- Lattice search > *n*-best rescoring
- Use in-domain PBMT to constrain search space
- NMT can be in- or out-of-domain

Code:

github.com/khayrallah/nematus-lattice-search

Thanks!

This material is based upon work supported in part by the Defense Advanced Research Projects Agency (DARPA) under Contract No. HR0011-15-C-0113.

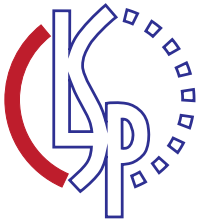
Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Defense Advanced Research Projects Agency (DARPA).

Neural Lattice Search for Domain Adaptation in Machine Translation

Huda Khayrallah, Gaurav Kumar
Kevin Duh, Matt Post, Philipp Koehn
{**huda**, gkumar, kevinduh, post, phi}@cs.jhu.edu

code:
github.com/khayrallah/nematus-lattice-search

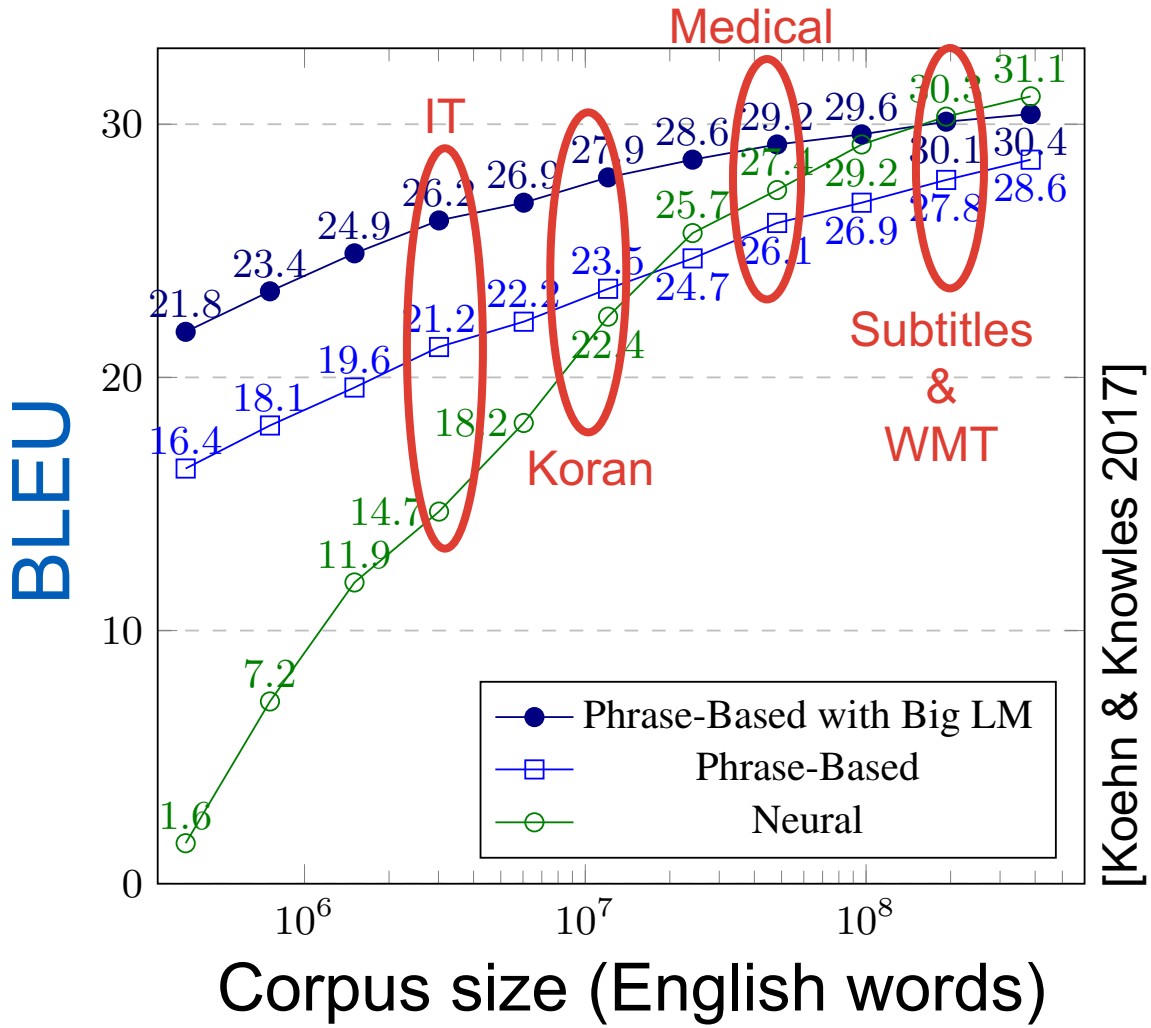
JOHNS HOPKINS
UNIVERSITY



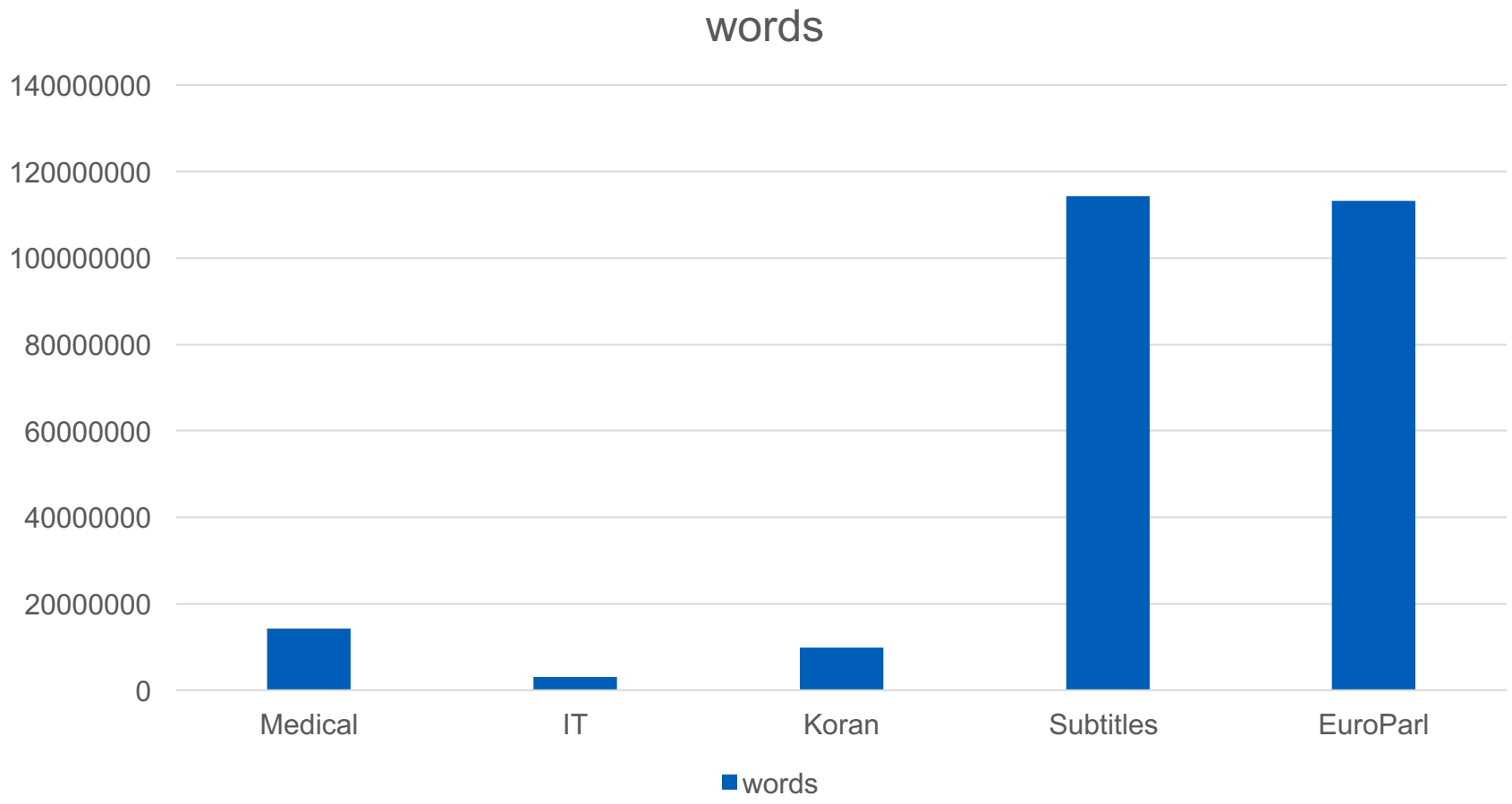
Corpus Sizes

Corpus	Words	Sentences	W/S
Medical	14,301,472	1,104,752	13
IT	3,041,677	337,817	9
Koran	9,848,539	480,421	21
Subtitles	114,371,754	13,873,398	8
EuroParl	113,165,079	4,562,102	25

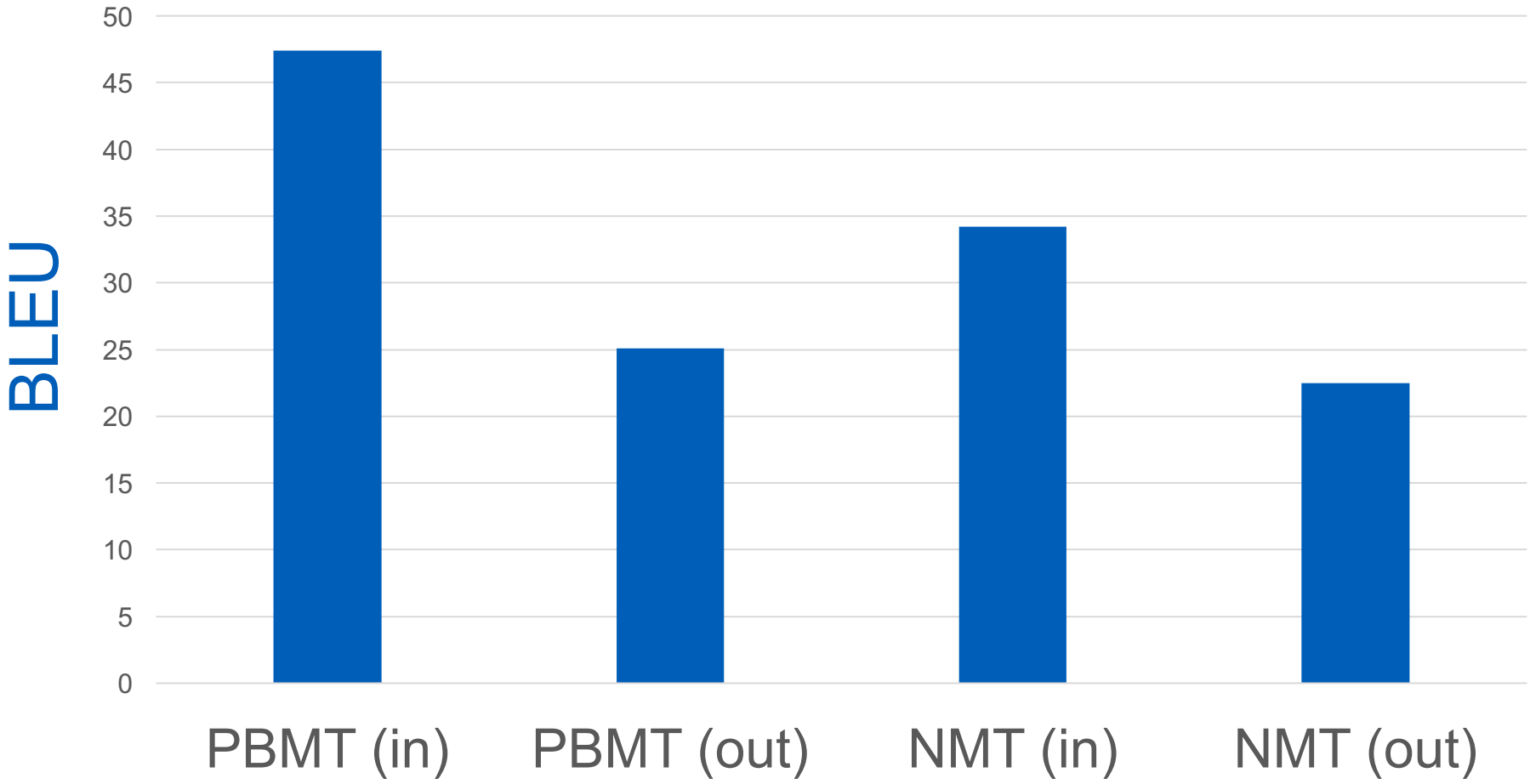
How much text do we have?



Corpus Sizes



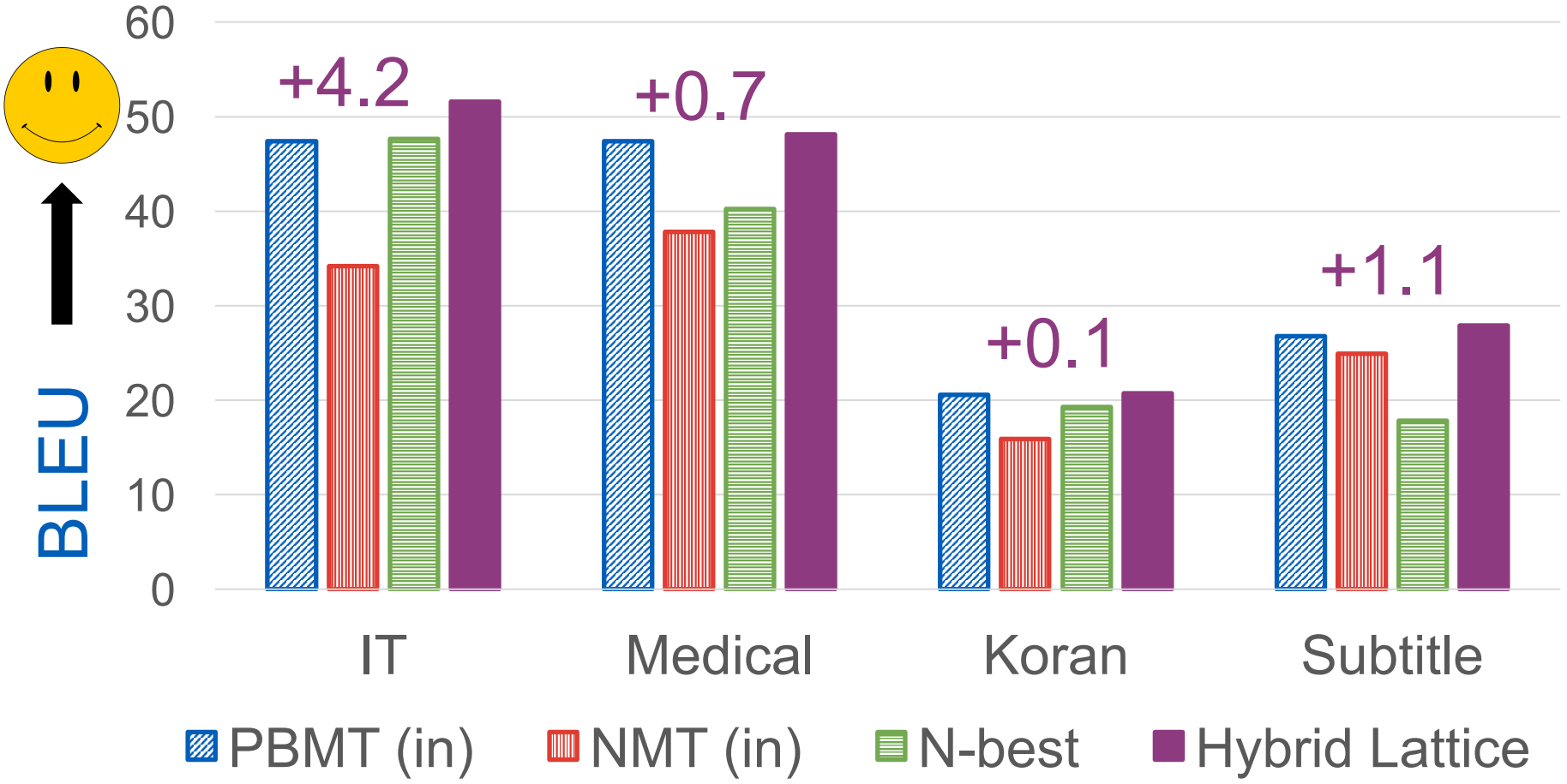
IT Baselines



Test Domain	Training Configuration	PBMT 1-best	NMT Standard Search	N-best Rescoring	NMT Lattice Search
IT	$\text{PBMT}_{out} \times \text{NMT}_{out}$	25.1 (-0.3)	22.5 (-2.9)	22.2 (-3.2)	25.4
	$\text{PBMT}_{in} \times \text{NMT}_{in}$	47.4 (-4.2)	34.2 (-17.4)	47.6 (-4.0)	51.6
	$\text{PBMT}_{in} \times \text{NMT}_{out}$	47.4 (-5.2)	22.5 (-30.1)	47.6 (-5.0)	**52.6***
	$\text{PBMT}_{out} \times \text{NMT}_{in}$	25.1 (-2.2)	34.2 (6.9)	22.4 (-4.9)	27.3
Medical	$\text{PBMT}_{out} \times \text{NMT}_{out}$	33.3 (-0.9)	32.9 (-1.3)	30.8 (-3.4)	34.2
	$\text{PBMT}_{in} \times \text{NMT}_{in}$	47.4 (-0.7)	37.8 (-10.3)	40.2 (-7.9)	**48.1***
	$\text{PBMT}_{in} \times \text{NMT}_{out}$	47.4 (-0.4)	32.9 (-14.9)	39.7 (-8.1)	**47.8**
	$\text{PBMT}_{out} \times \text{NMT}_{in}$	33.3 (-2.7)	37.8 (1.8)	31.2 (-4.8)	36.0
Koran	$\text{PBMT}_{out} \times \text{NMT}_{out}$	14.7 (-0.2)	10.8 (-4.1)	13.9 (-1.0)	14.9
	$\text{PBMT}_{in} \times \text{NMT}_{in}$	**20.6 (-0.1)**	15.9 (-4.8)	19.3 (-1.4)	**20.7**
	$\text{PBMT}_{in} \times \text{NMT}_{out}$	**20.6 (-0.2)**	10.8 (-10.0)	19.4 (-1.4)	**20.8***
	$\text{PBMT}_{out} \times \text{NMT}_{in}$	14.7 (-1.4)	15.9 (-0.2)	13.9 (-2.2)	16.1
Subtitle	$\text{PBMT}_{out} \times \text{NMT}_{out}$	26.6 (-0.9)	25.3 (-2.2)	19.7 (-7.8)	27.5
	$\text{PBMT}_{in} \times \text{NMT}_{in}$	26.8 (-1.1)	24.9 (-3.0)	17.8 (-10.1)	**27.9**
	$\text{PBMT}_{in} \times \text{NMT}_{out}$	26.8 (-1.6)	25.3 (-3.1)	17.1 (-11.3)	**28.4***
	$\text{PBMT}_{out} \times \text{NMT}_{in}$	26.6 (-1.0)	24.9 (-2.7)	19.8 (-7.8)	27.6

Source	Versionsinformationen ausgeben und beenden
Reference	output version information and exit
PBMT	Spend version information and end
NMT	Spend and end versionary information
lattice	Print version information and exit

Results


60
50
40
30
20
10
0
BLEU
+4.2
+0.7
+0.1
+1.1
IT
Medical
Koran
Subtitle
PBMT (in)
NMT (in)
N-best
Hybrid Lattice

Stack Based Decoding

- Stacks based on number of target words translated
- Keep track of:
  - Score
  - Current lattice node
  - Current neural state
  - incoming arc
  - length