

IIT Bombay's English-Indonesian submission at WAT: Integrating neural language models with SMT

Sandhya Singh, Anoop Kunchukuttan, Pushpak Bhattacharyya
{sandhya, anoopk, pb}@cse.iitb.ac.in

Center for Indian Language Technology
IIT Bombay



Motivation

- At CFILT, English-Indonesian language pair is being experimented as a part of a Project.
- Relatively new language pair among Asian language Translations.

About English-Indonesian Language pair

- Script is Latin for both English and Indonesian
- Sentence structure followed is SVO (Subject Verb Object)
- Not much structural divergence between English and Indonesian
- Indonesian is highly agglutinative and morphologically rich as compared to English language.
- Indonesian is considered as resource poor language

Experimental Setup (1/2)

PBSMT system

(Moses baseline)

- *MGIZA++ for word alignment*
- *grow-diag-final-end heuristic*
- Lexicalized Reordering
- 5-gram LM with Kneser-Ney smoothing using SRILM

OSM system

- Integrates 5-gram based reordering and translation in a single generative process (Durrani et al., 2013)
- Deals with words along with context of source & target

Experimental Setup (2/2)

Integrating NPLM to PBSMT system

- Neural Language model with default NPLM settings (Vaswani et al., 2013)
- Word embedding size as 700, 750, 800 for 5 epochs
- One hidden layer
- Integrated as a feature in PBSMT system

Integrating NNJM to PBSMT system

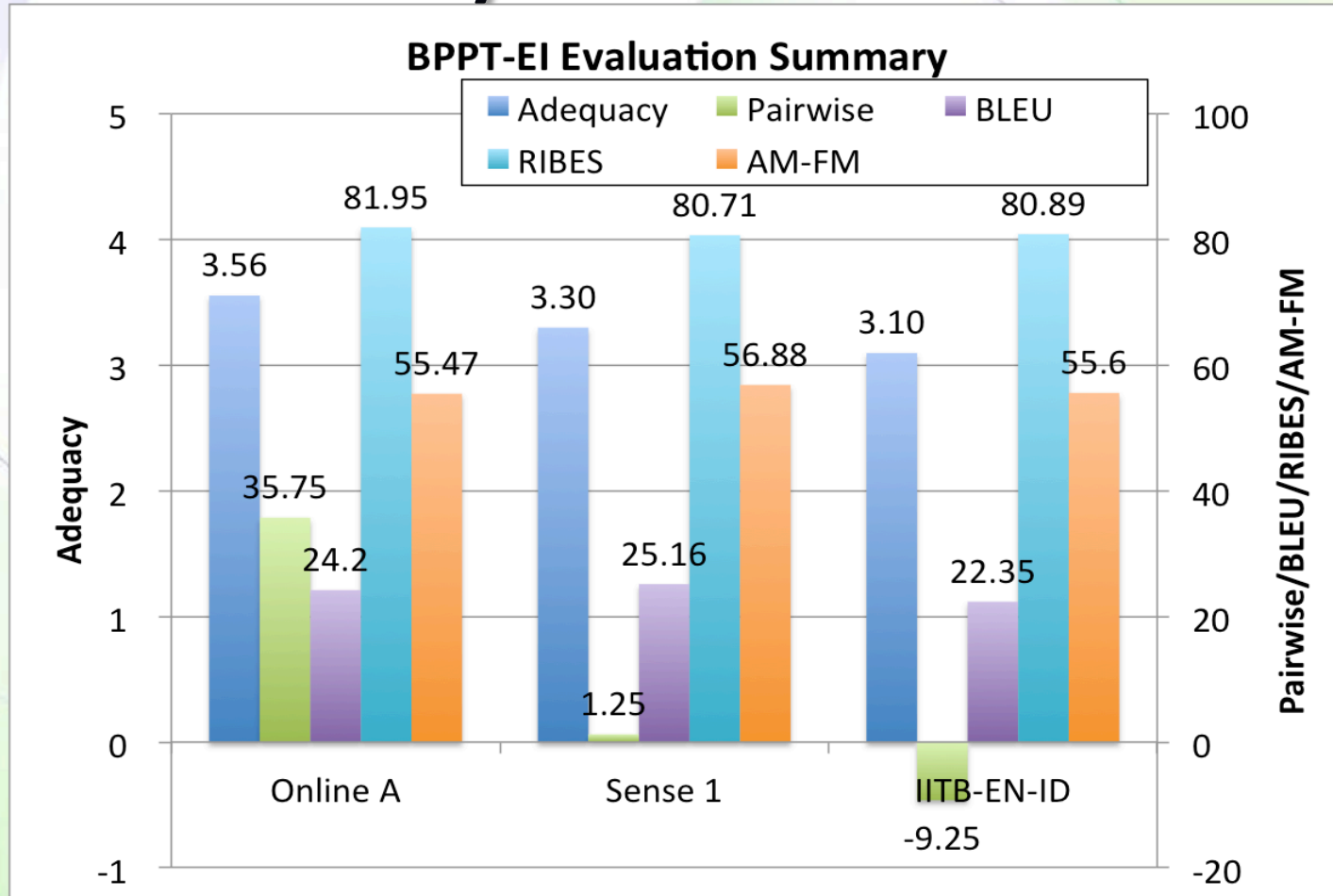
- Neural network joint LM using parallel data (*Devlin et al., 2014*)
- 5-gram LM with 9 source context word
- One hidden layer
- Integrated as a feature in PBSMT system

Automatic Evaluation Results for English-Indonesian MT Systems

Approach Used	BLEU score	RIBES score	AMFM score
Phrase based SMT	21.74	0.804986	0.55095
Operation Sequence Model	21.70	0.806182	0.552480
Neural LM with OE = 700	22.12	0.804933	0.5528
Neural LM with OE =750	21.64	0.806033	0.555
Neural LM with OE = 800	22.08	0.806697	0.55188
Joint neural LM*	22.35	0.808943	0.55597

* WAT submission , OE : Output Embedding

Summary of evaluations for English-Indonesian system

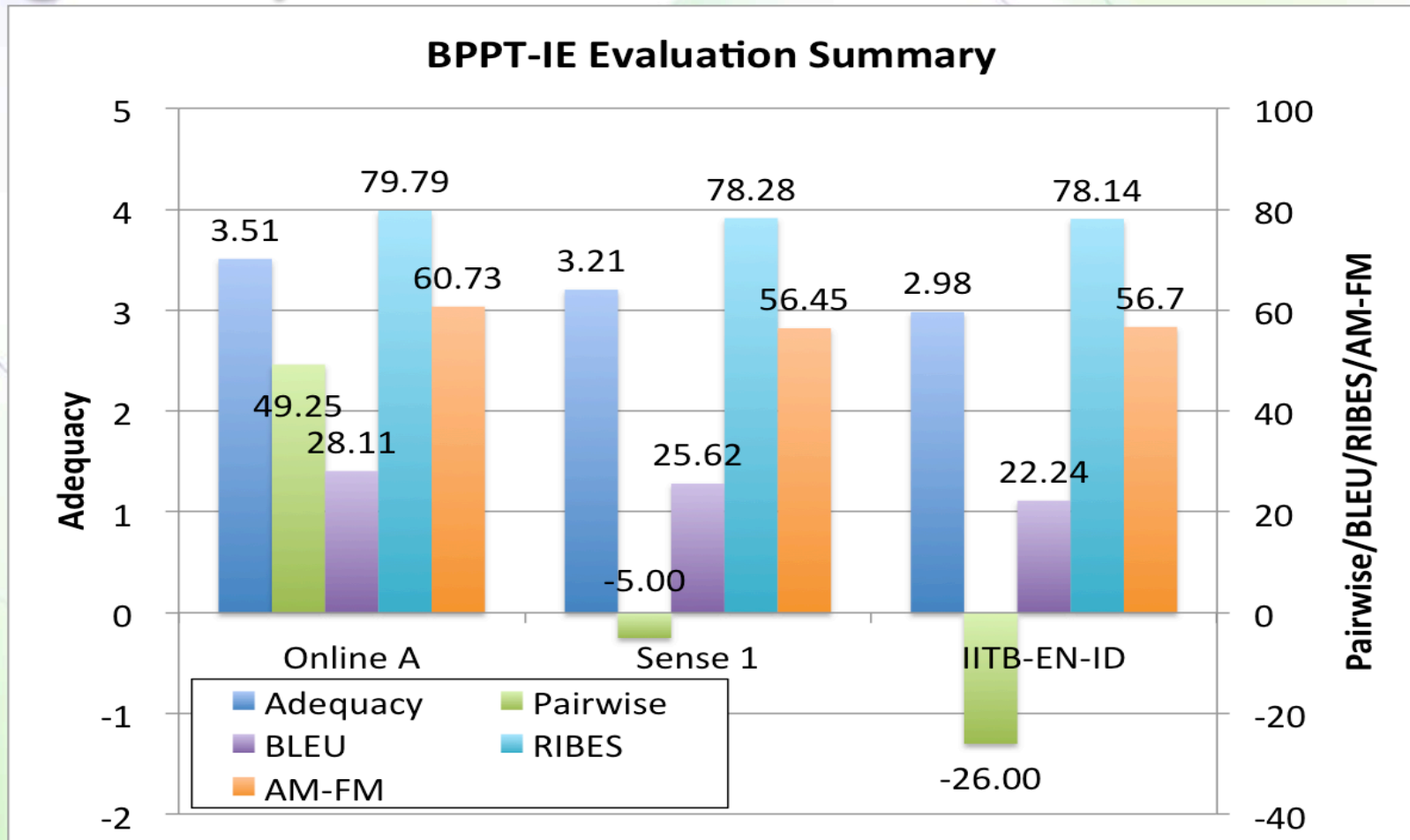


Automatic Evaluation Results for Indonesian-English MT Systems

Approach Used	BLEU score	RIBES score	AMFM score
Phrase based SMT	22.03	0.78032	0.564580
Operation Sequence Model*	22.24	0.781430	0.566950
Neural LM with OE= 700	22.58	0.781983	0.569330
Neural LM with OE = 750	21.99	0.780901	0.56340
Neural LM with OE = 800	22.15	0.782302	0.566470
Joint Neural LM	22.05	0.781268	0.565860

*WAT submission, OE: Output Embedding

Summary of evaluations for Indonesian-English system



Translations from Indonesian-English system

Reference Sentence	Translated Sentence	Error Analysis
Moreover, syariah banking has yet to become a national agenda, Riawan said.	In addition, the banking industry had not so national agenda, said Riawan who also director of the main BMI.	Phrase insertion
Of course, we will adhere to the rules, Bimo said.	We will certainly patuhi regulations, Bimo said.	All words not translated
The Indonesian government last year canceled 11 foreign-funded projects across the country for various reasons, the Finance Ministry said.	The government has cancel foreign loans from various creditors to 11 projects in 2006 because various reasons.	Phrase dropped

Conclusion

- Due to structural similarity, translation outputs are adequate to understand.
- Integrating Neural Probabilistic LM (NPLM) with additional data as a feature in PBSMT system improves the translation quality.
- Integrating Neural Network Joint Model (Bilingual LM) trained on parallel data as a feature in PBSMT system improves translation quality.

Future Work

- Experiment with pure neural machine translation system

References

1. Devlin, Jacob, Rabih Zbib, Zhongqiang Huang, Thomas Lamar, Richard M. Schwartz, and John Makhoul. 2014. "*Fast and Robust Neural Network Joint Models for Statistical Machine Translation.*" In conference of the Association of Computational Linguistics.
2. Durrani, Nadir, Alexander M. Fraser, and Helmut Schmid. 2013. "*Model With Minimal Translation Units, But Decode With Phrases.*" HLT-NAACL
3. Vaswani, Ashish, Yinggong Zhao, Victoria Fossum, and David Chiang. 2013. "*Decoding with Large-Scale Neural Language Models Improves Translation.*" In EMNLP.