Unsupervised domain adaptation for neural machine translation with domain-aware feature embeddings

Zi-Yi Dou, Junjie Hu, Antonios Anastasopoulos, Graham Neubig
Language Technologies Institute, Carnegie Mellon University

Carnegie Mellon University
School of Computer Science

Language Technologies Institute

NEULAB

Unsupervised Domain Adaptation

- Neural machine translation (NMT) models perform poorly when domain mismatch occurs.
  - Unseen words
  - Different senses in different domains
- Often no access to in-domain parallel data.
- How can we adapt NMT models with *no in-domain* parallel training data?

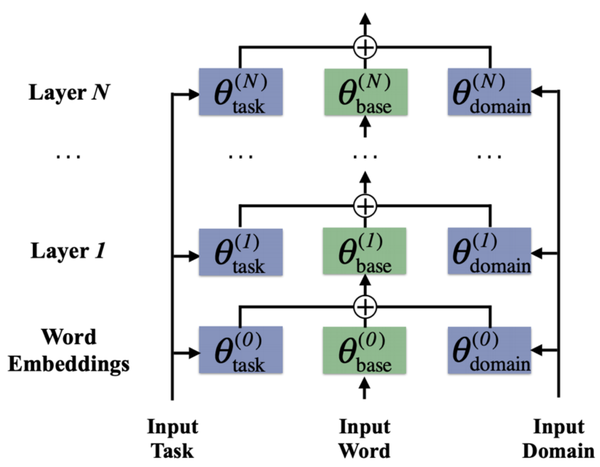
Domain-Aware Feature Embeddings (DAFE)

- DAFE performs unsupervised domain adaptation by separating the networks into three different parts.

Base Network
learns common features across different domains/tasks; its parameters are shared.

Domain/Task Embedding Learner
directly generates domain and task embeddings at each encoding layer with look-up operations.

Main Architecture

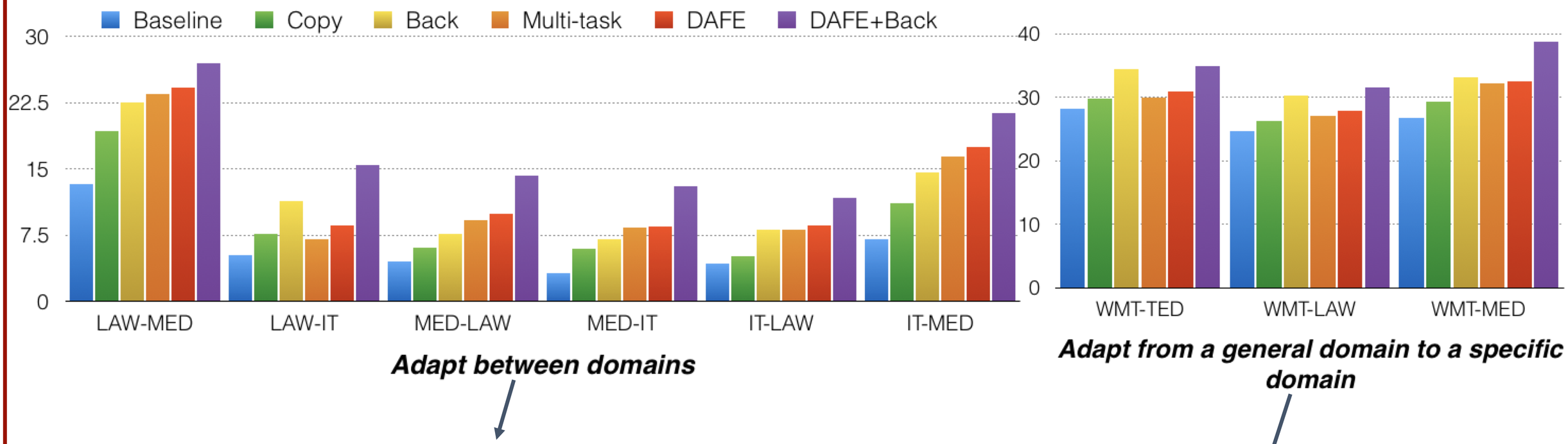


- The output of the *l*-th encoding layer is:
  BASE(input_word) +
  DOMAIN_EMBED(input_domain) +
  TASK_EMBED(input_task)

Training Strategy

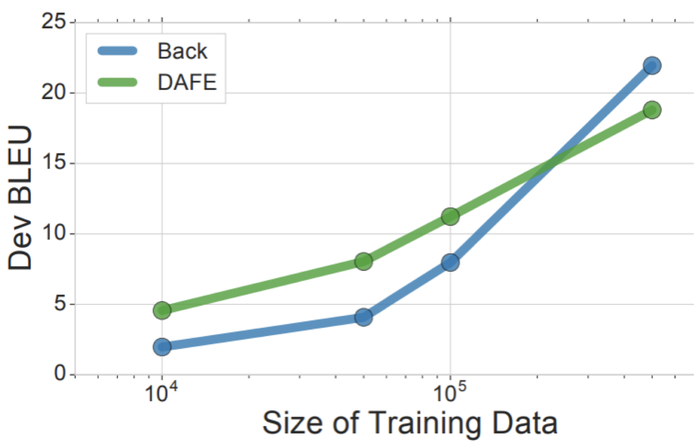
- Auxiliary objective: masked language modeling
- At each training step, train with:
  1. out-of-domain translation objective
  2. out-of-domain language modeling objective
  3. in-domain language modeling objective

How does DAFE compare with Unadapted, Back-translation, Copy baselines?



- DAFE outperforms all other baselines when adapting between domains (e.g. adapt from the law domain to the medical domain)
- DAFE cannot outperform back-translation when adapting from a general (WMT) to a specific (e.g. law) domain
- Combined DAFE and back-translation (DAFE+Back) further improves performance, outperforming all other methods

How do corpus size and lexicon coverage affect the performance of DAFE and Back-translation?



- DAFE outperform back-translation in *low-resource* settings
- Low quality back-translated data can harm performance

Can we control the output domain by feeding desired domain embeddings?

Reference	please report this bug to the developers .
MED-embed	please report this to the EMEA .
IT-embed	please report this bug to the developers .
Reference	for intramuscular use .
MED-embed	for intramuscular use .
IT-embed	for the use of the product .

- Input medical embeddings -> generate words like *"EMEA" (European Medicines Evaluation Agency)* and *"intramuscular"*.
- Input IT embeddings -> generate words like *"bug"* and *"developers"*.