

# Adversarial Training for Cross-Domain Universal Dependency Parsing

Motoki Sato Hitoshi Manabe Hiroshi Noji Yuji Matsumoto  
 {sato.motoki.sa7, manabe.hitoshi.me0, noji, matsu}@is.naist.jp

6<sup>th</sup> Rank NAIST-SATO



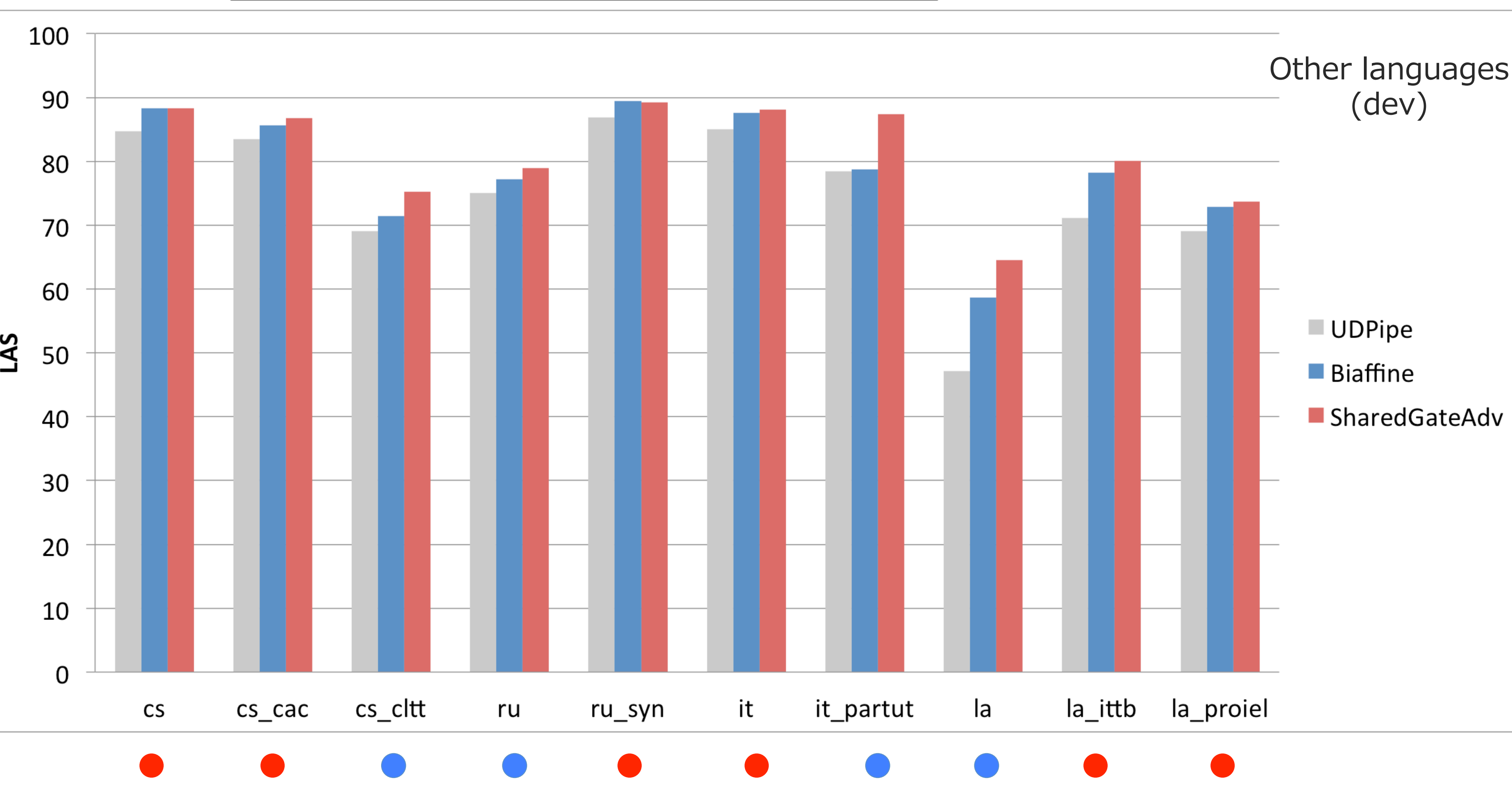
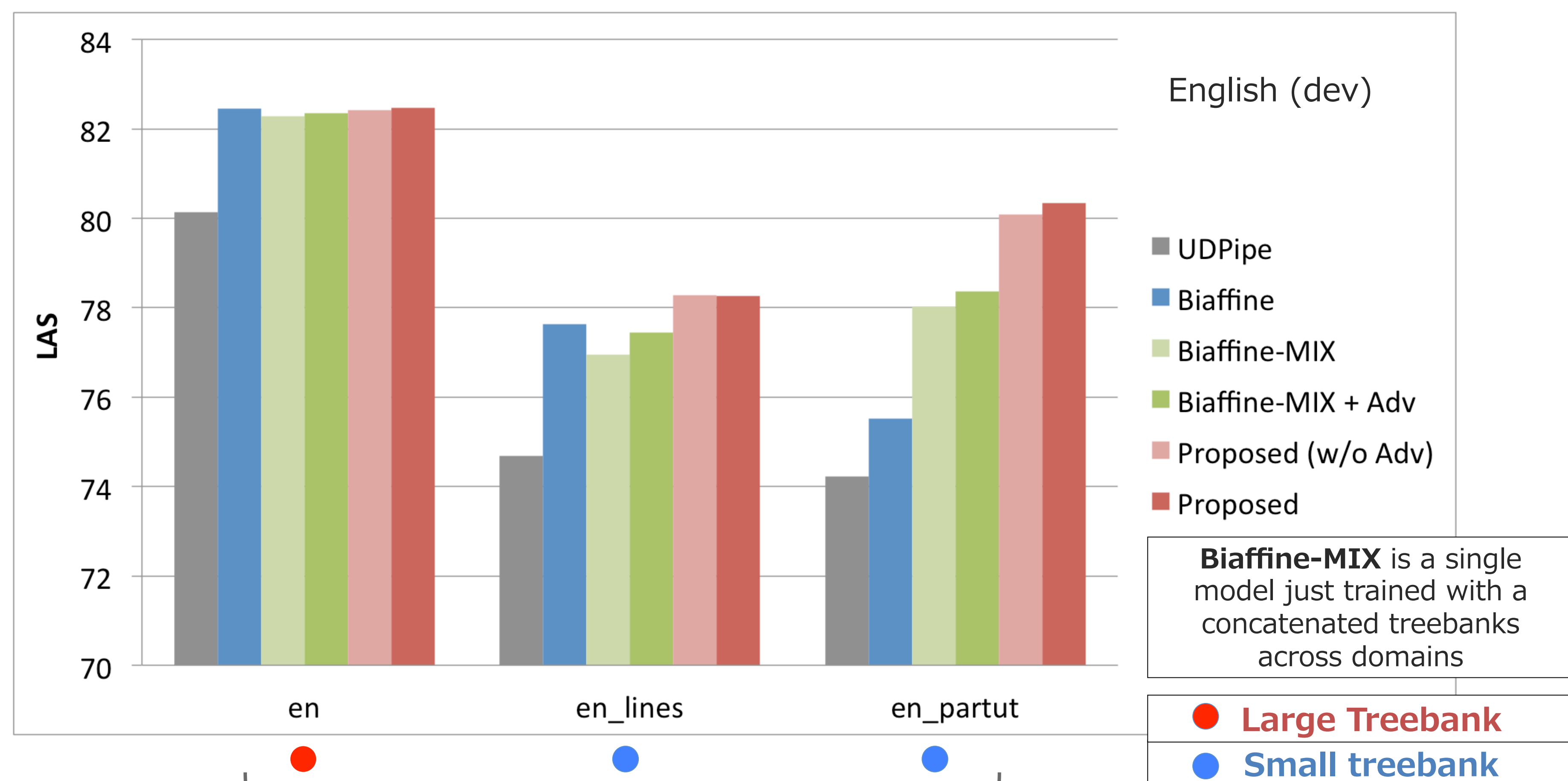
## 1. Cross-Domain Setting

### Example:

en	12,543 sentences	} <b>Large treebank</b>
en_lines	2,738 sentences	
en_partut	1,090 sentences	} <b>Small treebank</b>

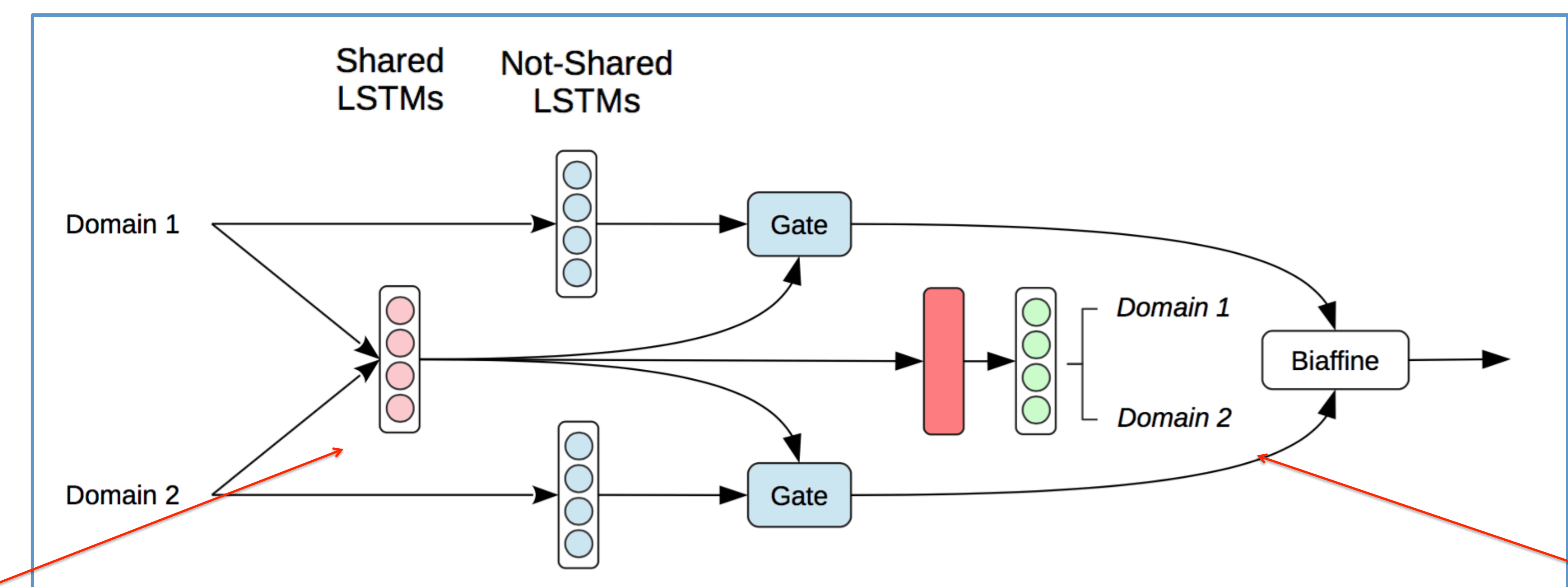
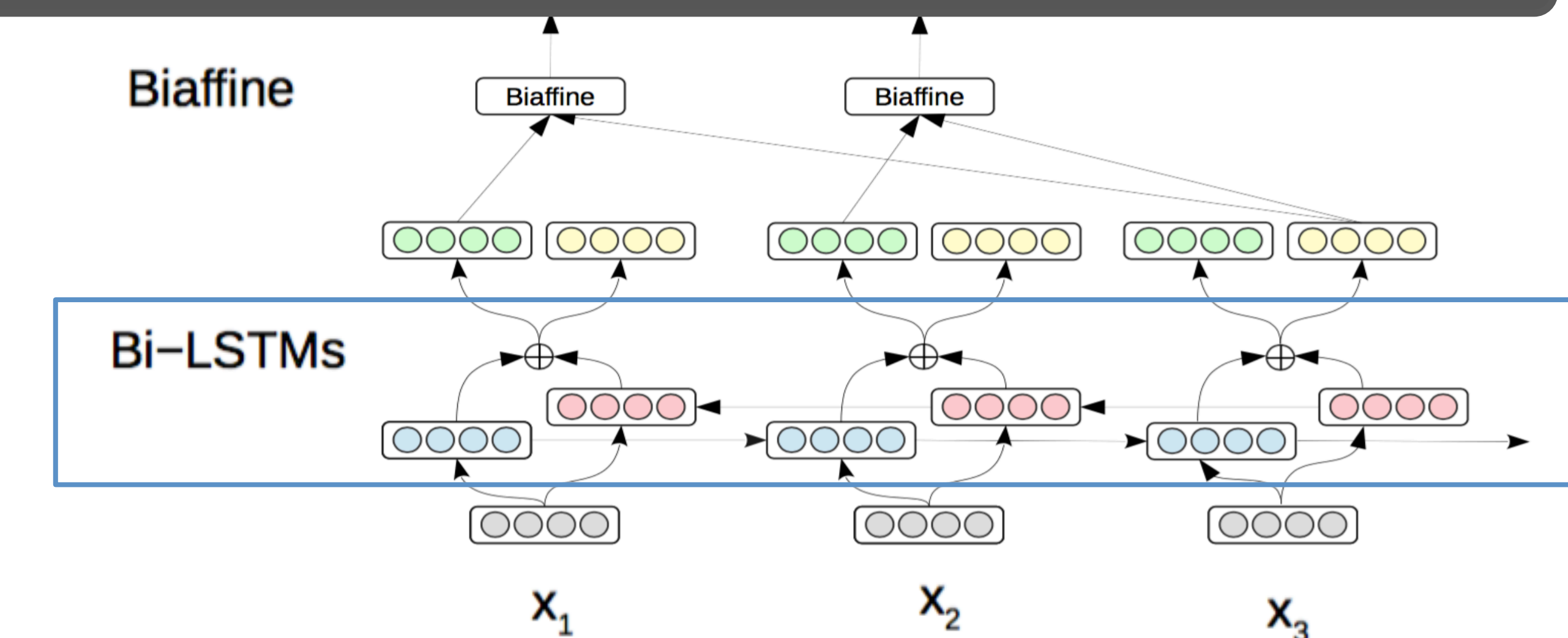
**Our goal :** Improve performance for **small treebanks**

## 3. Experiments

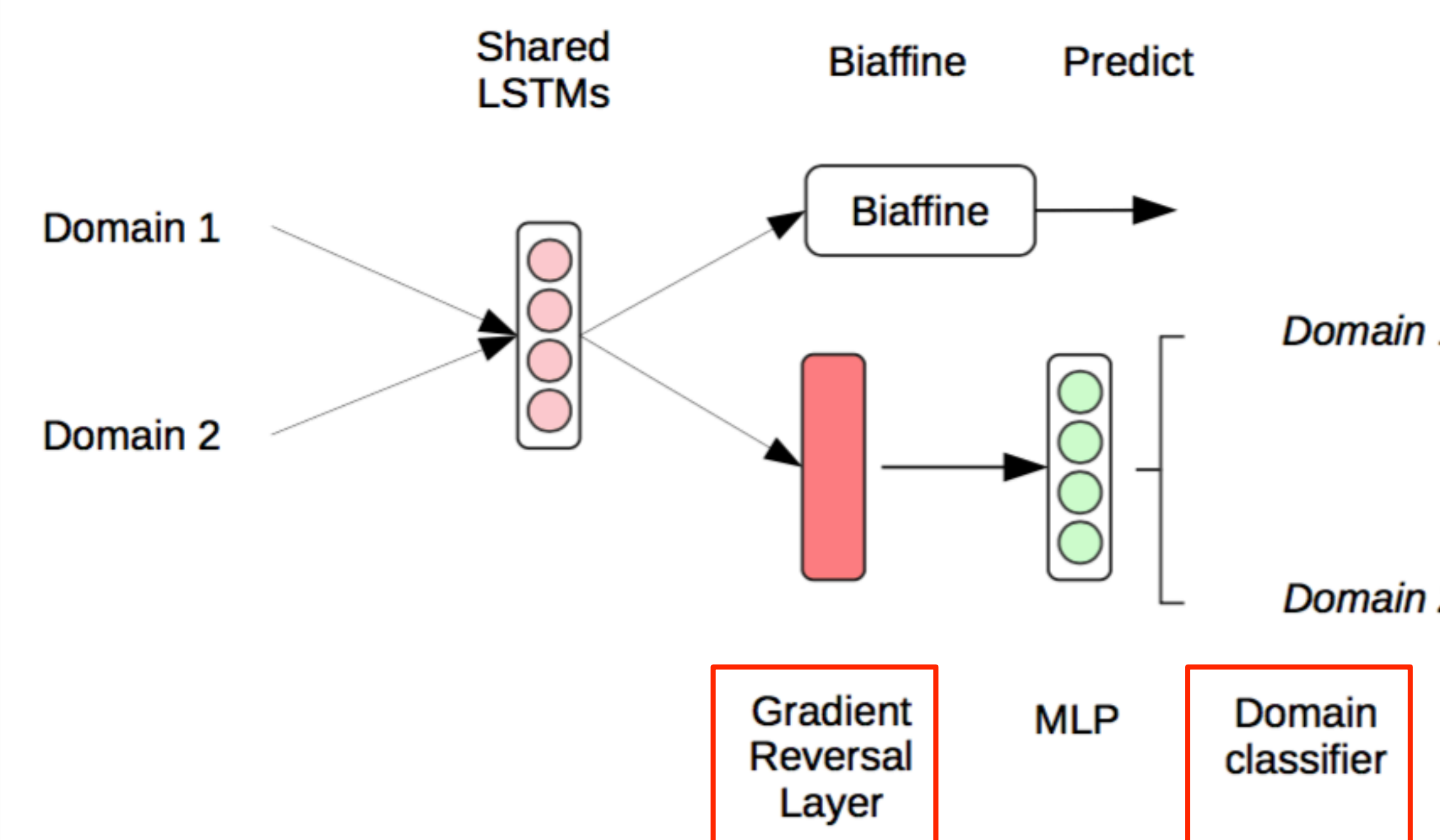


## 2. Our proposed method (SharedGateAdvNet)

- Our proposed method is based on **Biaffine model** [Dozat and Manning, 2017].
- We jointly train **domain-specific models** that are comprised of two components:
  - Shared LSTMs : **Domain-invariant features**
  - Not-Shared LSTMs : **Domain-specific features**



### Adversarial Training



### Gradient Reversal Layer:

$$R_{\lambda}(\mathbf{x}) = \mathbf{x}; \quad \frac{dR_{\lambda}}{d\mathbf{x}} = -\lambda \mathbf{I}.$$

hyper-parameter:  $\lambda$  (=0.5)

The model is encouraged to learn **shared parameters** that are not specific to a particular domain (**domain-invariant**)

## 4. Conclusion

- We have proposed a domain adaptation technique with adversarial training for parsing.
- By applying it on the recent state-of-the-art graph-based dependency parsing model with Bi-LSTMs, we obtained a consistent score improvement, especially for the treebanks having **less training data**.