

Yao-jie LU, Mu XU, Chang-xing WU, De-yi XIONG, Hong-ji WANG, Jin-song SU,
2018. Cross-lingual implicit discourse relation recognition with co-training.
Frontiers of Information Technology & Electronic Engineering, 19(5):651-661.
<https://doi.org/10.1631/FITEE.1601865>

Cross-lingual implicit discourse relation recognition with co-training

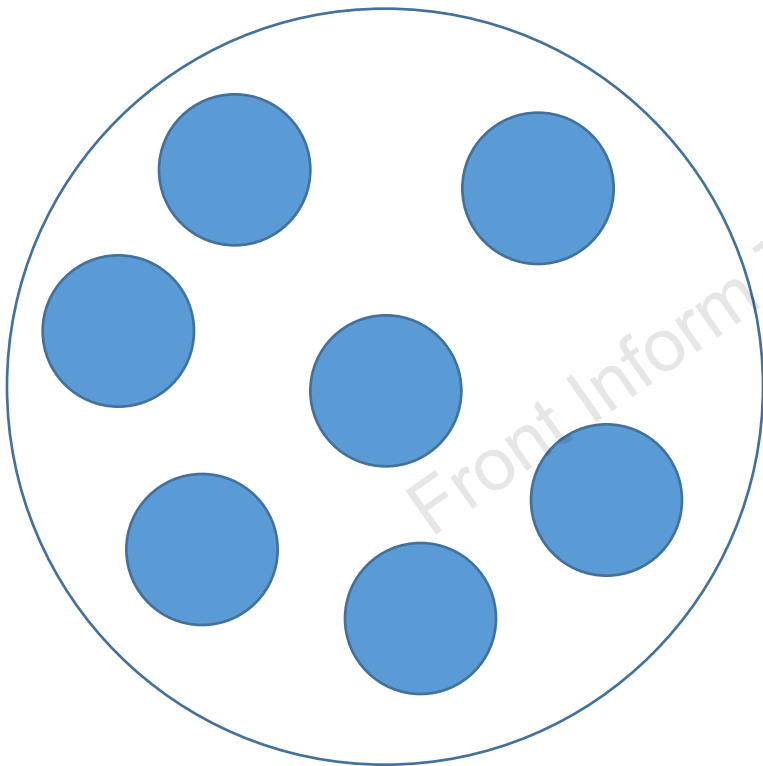
Key words: Cross-lingual; Implicit discourse relation recognition;
Co-training

Corresponding author: Jin-song SU
E-mail: jssu@xmu.edu.cn

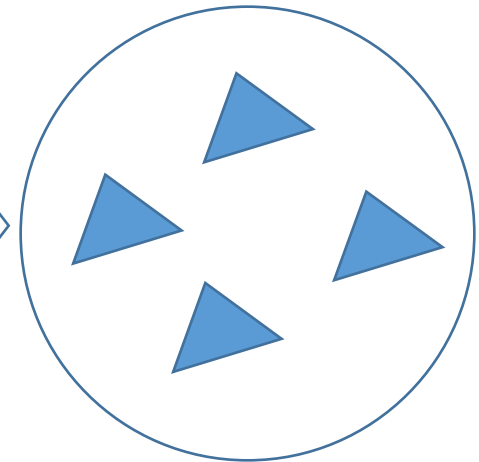
Motivation

Rich resource language
(English)

Low resource language
(Chinese)



Discourse
relation
knowledge



Methods

1. Translate the labeled English instances into the labeled Chinese ones and the unlabeled Chinese instances into the unlabeled English ones using machine translation services, respectively.

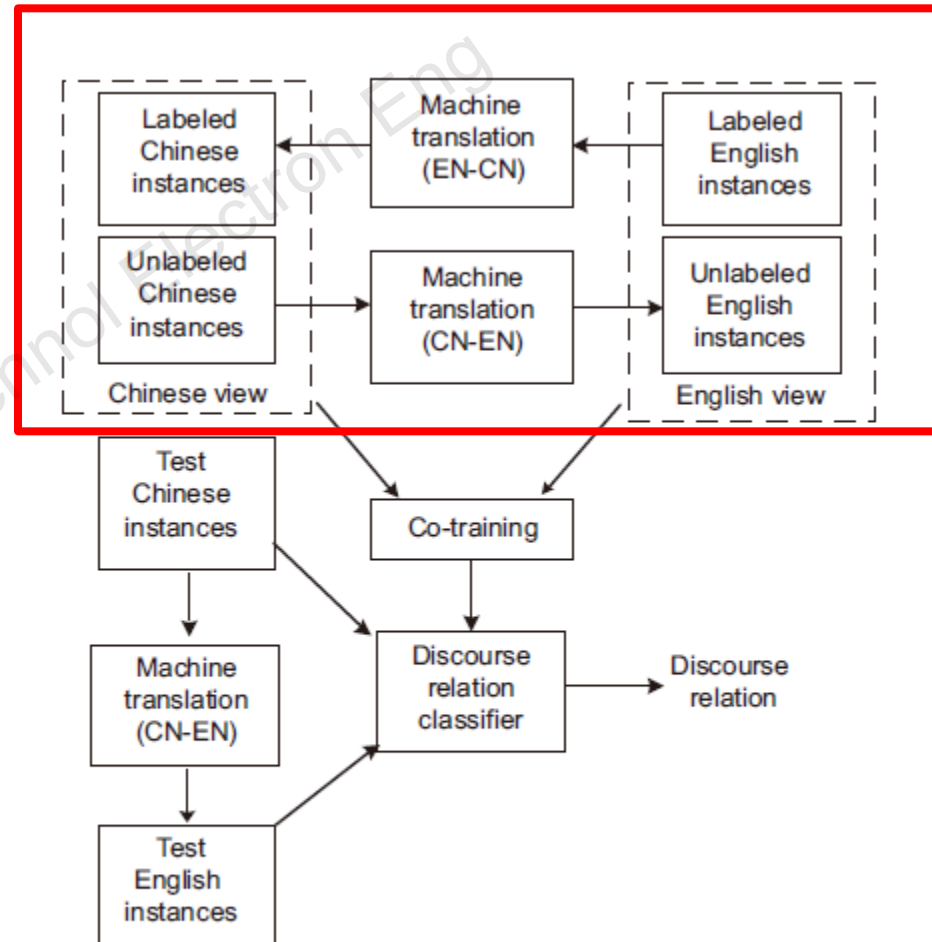


Fig. 1 Framework of co-training for cross-lingual implicit discourse relation recognition

Methods

2. Adapt the co-training algorithm to our scenario and train two classifiers in English and Chinese, respectively.

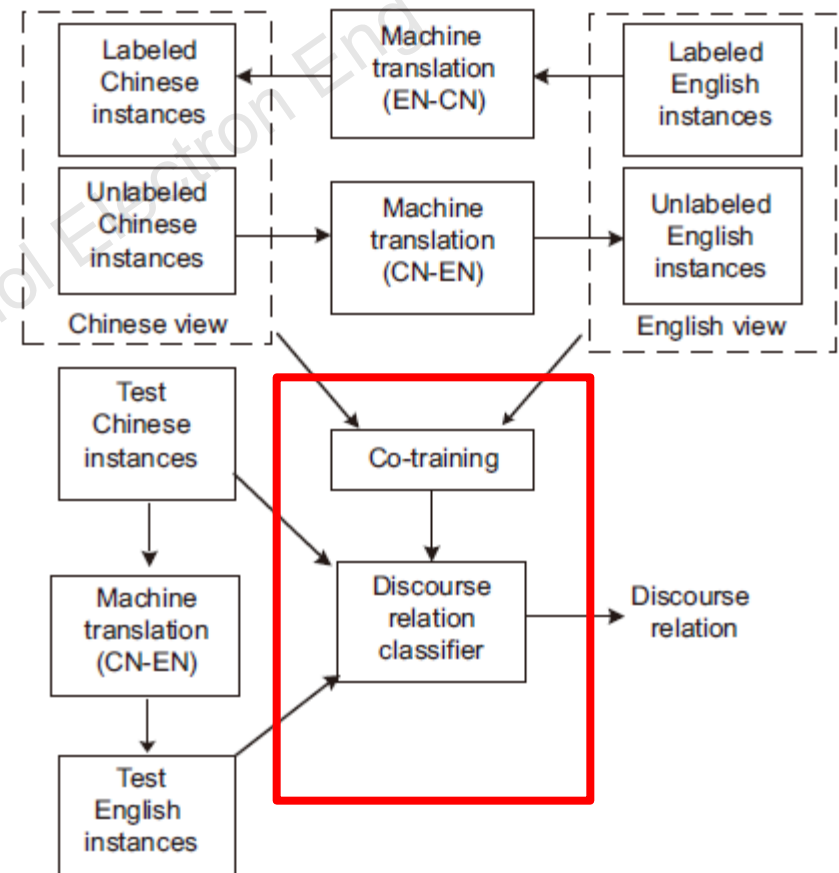


Fig. 1 Framework of co-training for cross-lingual implicit discourse relation recognition

Methods

3. Apply these two classifiers to implement implicit DRR for the Chinese.

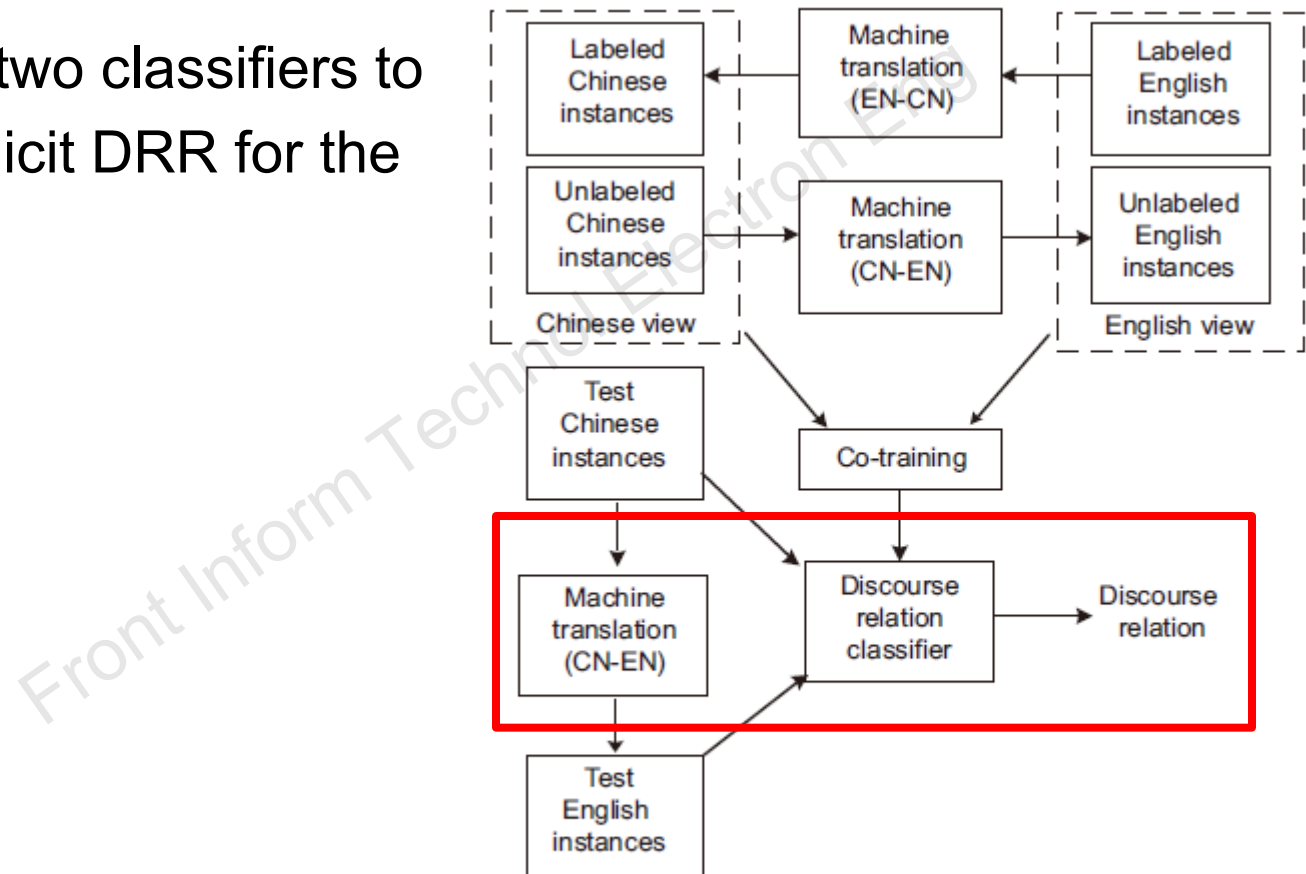


Fig. 1 Framework of co-training for cross-lingual implicit discourse relation recognition

Experiments

1. Instance numbers of data sets in our experiments

Table 2 Instance numbers of datasets in our experiments

Discourse relation	Number of labeled data			Unlabeled data (CN)
	Train positive/negative (EN)	Val positive/negative (CN)	Test positive/negative (CN)	
Comp	1942/1942	15/944	24/1172	10 691
Cont	3342/3342	146/813	175/1021	
Expa	7004/7004	772/187	772/187	
Temp	760/760	26/933	41/1155	

CN: Chinese; EN: English

2. Data used in baselines and the co-training method

Table 3 Data used in baselines and the co-training method

Method	Training set		Val/Test set		Unlabeled set	
	English	E-to-C	Chinese	C-to-E	Chinese	C-to-E
SVM (EN)	✓			✓		
SVM (CN)		✓	✓			
SVM (ENCN1)	✓	✓	✓	✓		
SVM (ENCN2)	✓	✓	✓	✓		
TSVM (EN)	✓			✓		✓
TSVM (CN)		✓	✓		✓	
TSVM (ENCN1)	✓	✓	✓	✓	✓	✓
TSVM (ENCN2)	✓	✓	✓	✓	✓	✓
Self-training SVM (EN)	✓			✓		✓
Self-training SVM (CN)		✓	✓		✓	
Co-training	✓	✓	✓	✓	✓	✓

E-to-C: Chinese translations of English instances; C-to-E: English translations of Chinese instances

Experiments

3. Performance (F_1 -score/accuracy) comparison on the test set

Table 5 Performance (F -score/accuracy) comparison on the test set

Method	F_1 (Acc)			
	Temp	Cont	Comp	Expa
SVM (EN)	7.46 (78.11)	19.51 (79.61)	10.45 (89.41)	68.26 (55.34)
SVM (CN)	7.12 (67.78)	19.38 (79.44)	6.14 (75.73)	66.84 (54.28)
SVM (ENCN1)	7.69 (70.34)	21.79 (82.26)	10.81 (88.35)	69.67 (57.28)
SVM (ENCN2)	9.18 (75.55)	20.48 (79.44)	10.53 (86.50)	63.41 (51.10)
TSVM (EN)	5.77 (91.35)	23.17 (42.63)	10.13 (93.73)	81.20 (69.55)
TSVM (CN)	5.77 (91.35)	20.99 (29.57)	4.17 (83.76)	82.40 (70.96)
TSVM (ENCN1)	8.47 (90.47)	21.29 (30.19)	9.64 (93.38)	82.37 (70.79)
TSVM (ENCN2)	8.47 (90.47)	21.29 (30.19)	9.64 (93.38)	82.37 (70.79)
Self-training SVM (EN)	8.42 (84.64)	21.79 (60.72)	10.45 (89.41)	83.87 (72.90)
Self-training SVM (CN)	8.21 (84.20)	22.00 (52.43)	9.62 (91.70)	87.24 (78.02)
Co-training	12.50 (93.82)	22.86 (83.32)	14.00 (92.41)	88.61 (79.96)

The bold numbers are the best results in each relation

Experiments

4. Illustration of co-training and self-training on the test set with different iteration numbers: (a) temporal; (b) contingency; (c) comparison; (d) expansion

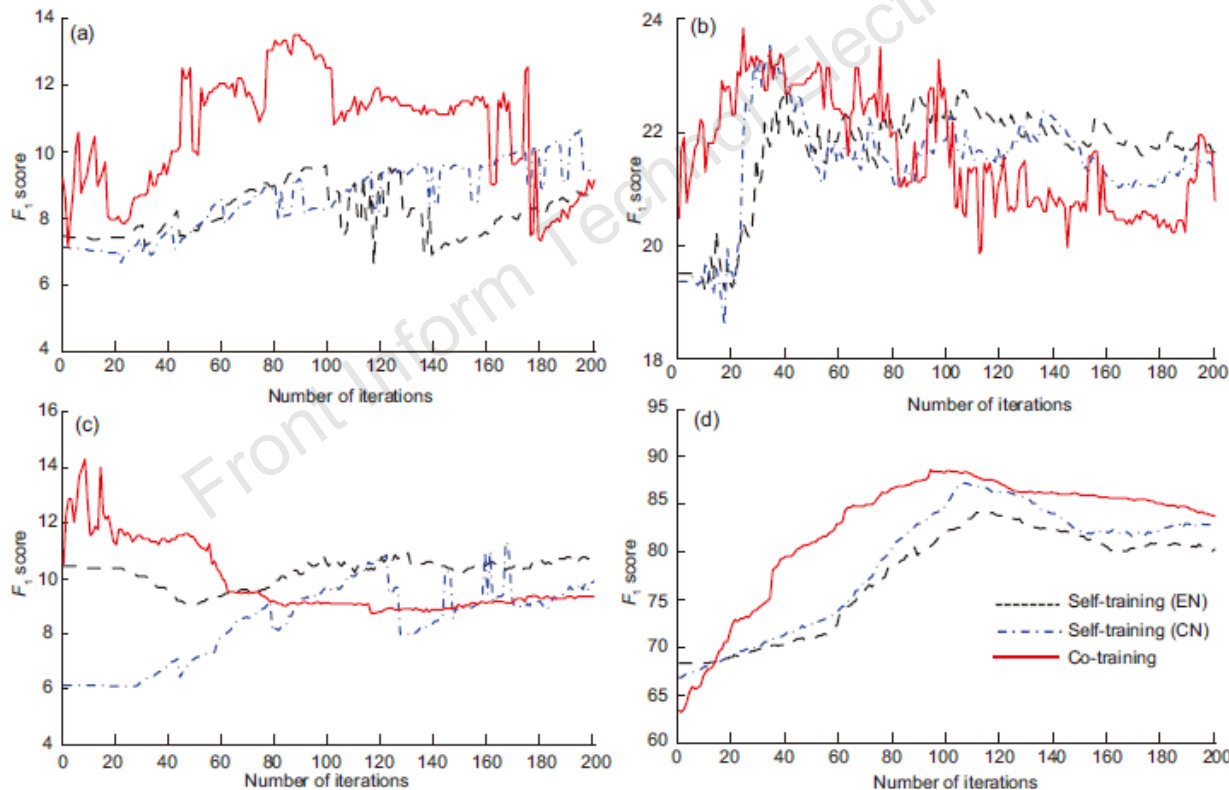


Fig. 2 Illustration of co-training and self-training on the test set with different iteration numbers: (a) temporal; (b) contingency; (c) comparison; (d) expansion

Conclusions

1. In this paper, we have proposed a co-training approach to improve the accuracy of Chinese implicit DRR by exploiting English labeled data and Chinese unlabeled data.
2. In our approach, each instance could be considered by two views of different languages and therefore various syntactic features in different languages could be exploited for Chinese implicit DRR.
3. In the future, we plan to investigate the effectiveness of our method for other languages and fine-grained subtype relations in PDTB.