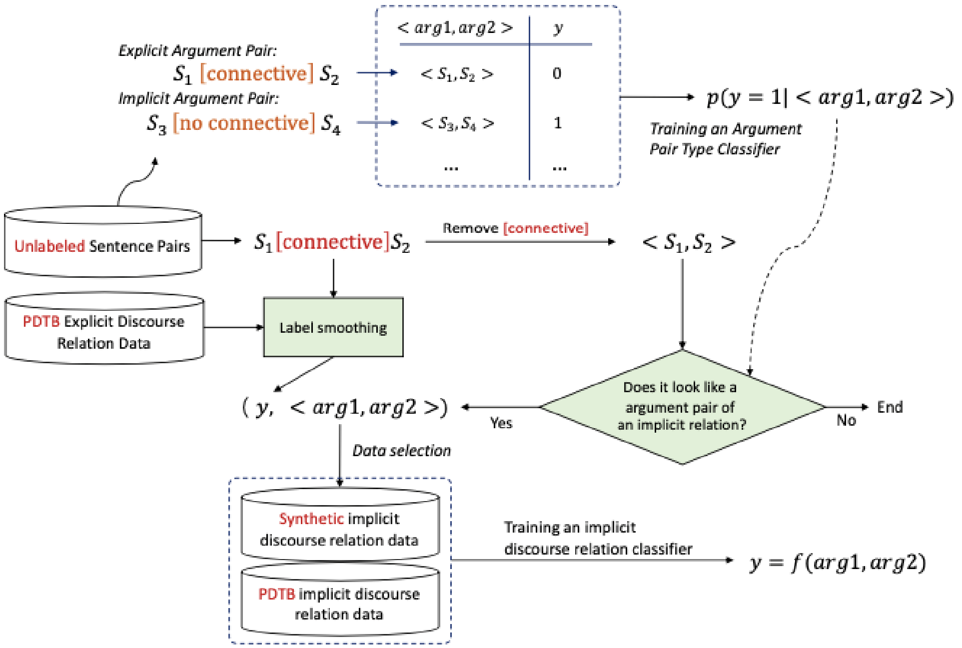
Discriminative Explicit Instance Selection for Implicit Discourse Relation Classification

Wei Song, Hongfei Han, Xu Han, Miaomiao Cheng, Jiefu Gong, Shijin Wang, Ting Liu

Frontiers of Computer Science, DOI: 10.1007/s11707-023-3058-2

Problems & Ideas

- Implicit Discourse Relation Classification is challenging due to the limited training data. A potential solution is to expand implicit discourse relations from explicit discourse relations. However,
  - There is linguistic dissimilarity between explicit and implicit argument pairs.
  - Dropping the discourse connective between explicit argument pairs could lead to ambiguous and wrong labels, resulting in problematic synthetic implicit data.
- Ideas:
  - Can we select explicit argument pairs that are similar to natural implicit argument pairs for data expansion?
  - Can we adopt the label-smoothing strategy to reduce noise during sense assignment?



An overview of the proposed method. We train an argument pair type classifier, which gives the probability of being naturally implicit for a given argument pair. A label-smoothing strategy is also proposed for automatic sense assignment, reducing the noise labels in synthetic data expansion.

Main Contributions

- Contributions: we propose
  - A discriminative explicit instance selection method based on argument pair type classification, reducing inappropriate argument pairs.
  - A simple label-smoothing method for reducing noise during automatic sense assignment.

Method	Top-level Macro-F1	Top-level ACC	Second-level ACC
PDTB 2.0			
Domain-Adaptation [10]	39.46	-	-
Freely omissible [9]	40.50	51.70	-
Multi-task Learning [11]	44.98	57.27	-
Active Learning [35]	44.48	60.63	-
Knowledge [36]	52.89	59.66	48.23
Knowledge [26]	51.24	59.94	-
Multi-task Learning [12]	58.48	65.26	54.32
XLNet-large [37]	59.10	68.70	61.29
BMGF-RoBERTa [27]	63.39	69.06	58.13
Our Method	**64.29**	**69.75**	**61.69**
PDTB 3.0			
LSTM [38]	-	-	43.41
XLNet-large [37]	68.30	73.80	**64.83**
MTL [39]	-	-	63.30
BMGF-RoBERTa	66.57	71.00	61.02
Our Method	**69.94**	**74.98**	64.20

Our method obtains competitive and even better performance compared with previous approaches.

Model	Top-level Macro-F1	Top-level ACC	Second-level ACC
PDTB 2.0			
Baseline	62.83	69.46	58.96
Low ambiguity	62.23	68.79	58.83
Freely omissible	61.55	67.90	-
Our method	**64.29**	**69.75**	**61.69**
PDTB 3.0			
Baseline	68.86	74.41	62.93
Low ambiguity	68.93	73.83	63.16
Freely omissible	66.73	71.95	-
Our method	**69.94**	**74.98**	**64.20**

Our method outperforms previous data expansion approaches on PDTB 2.0 and PDTB 3.0.

Model	Top-level Com	Con	Exp	Tem
PDTB 2.0				
Baseline	61.06	61.91	**77.00**	51.32
Low ambiguity	58.79	62.85	76.20	51.09
Freely omissible	60.15	60.49	75.64	49.93
Our method	**62.23**	**63.49**	76.73	**54.71**
PDTB 3.0				
Baseline	62.83	77.30	**78.55**	56.76
Low ambiguity	63.19	75.42	77.98	**59.13**
Freely omissible	59.10	74.27	76.43	57.14
Our method	**64.80**	**77.77**	78.48	58.69

Our method gains obvious improvements on the minority relations, while Expansion (Exp) is most frequent.