

Discourse Relation Sense Classification with Two-Step Classifiers

Yusuke Kido[†] and Akiko Aizawa[‡]

Contact: y.k@is.s.u-tokyo.ac.jp

[†]Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

[‡]National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan

TASK & BACKGROUND

Discourse Relation Sense Classification =

- Classification task on *sense* of PDTB-style discourse relations (Prasad et al., 2007)
- Subtask of shallow discourse parsing

Previous Approaches

- NB, SVM, MaxEnt classifier with ~20 kinds of **carefully hand-crafted features**

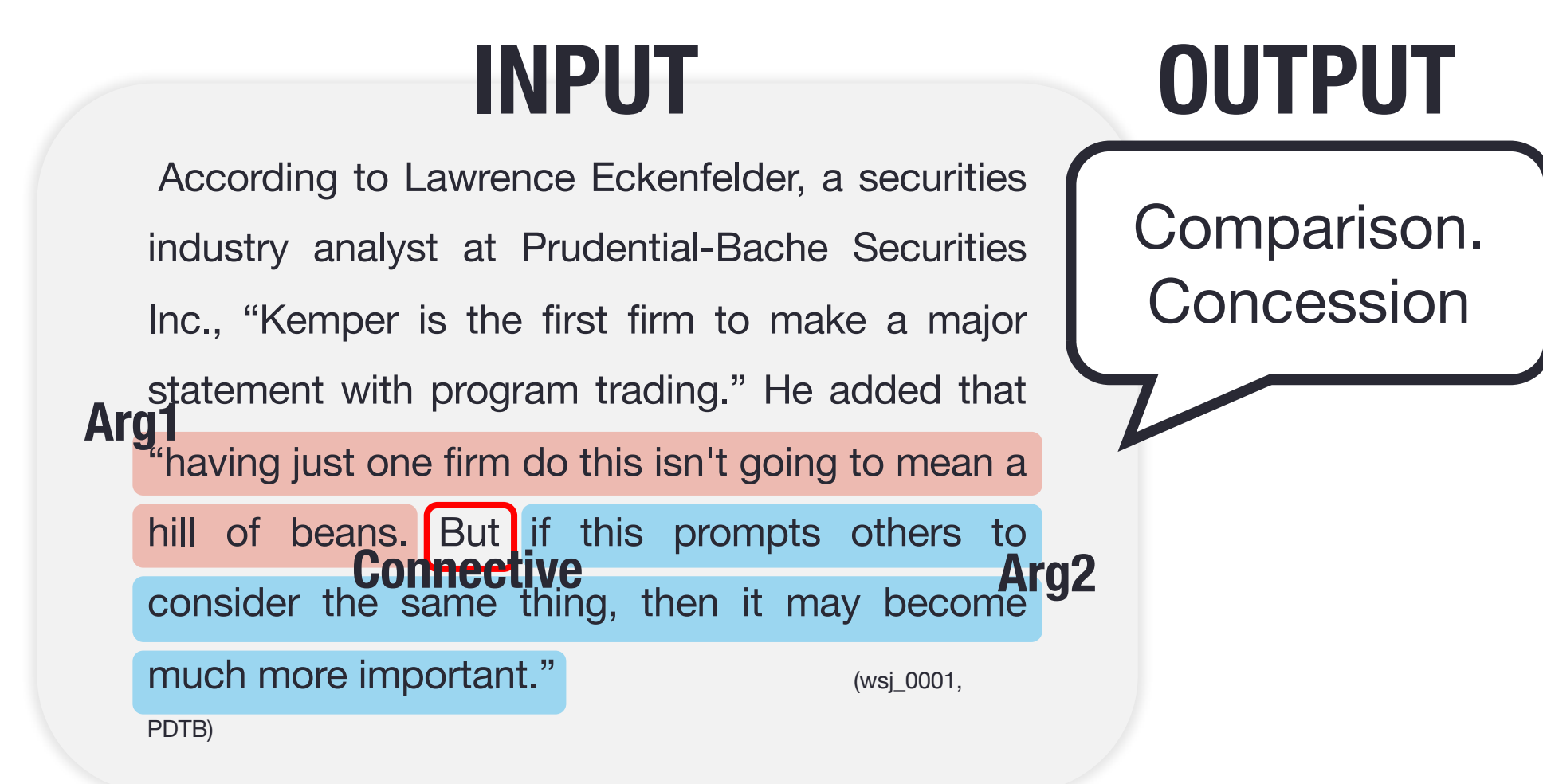


Figure A. Example input and corresponding output

DATA ANALYSIS

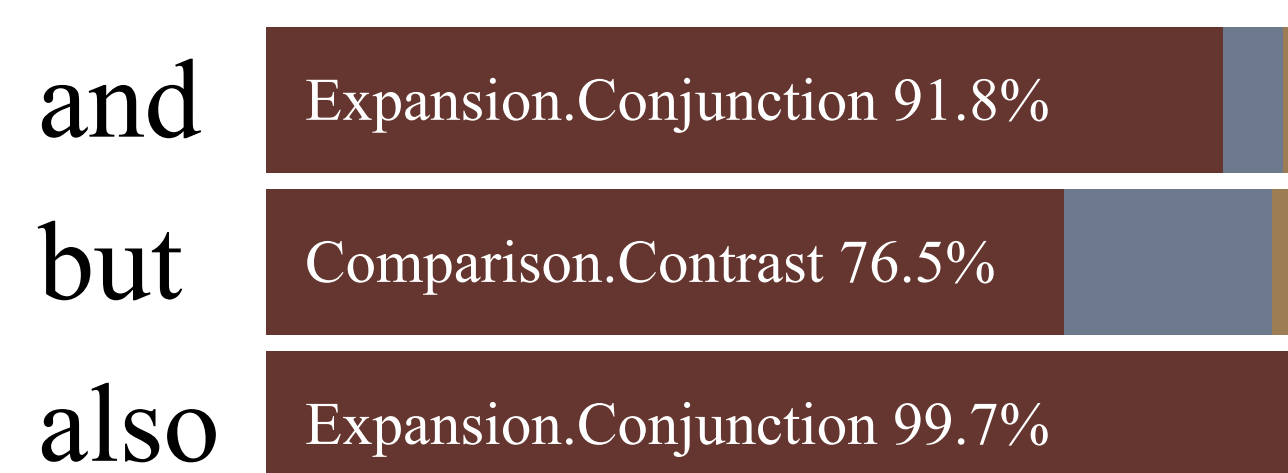


Figure B. Top 3 most frequent connectives and their dominant senses

Explicit/AltLex Discourse Relations

- Have explicit connective words
- Each kind of connective is mostly used in very few different senses
→ **Even simple majority classifier is expected to work very well**
- Same trend for AltLex (389 of 405 connectives are used in single sense)
- Exceptionally low accuracy on **Comparison.Concession** (e.g. *while*, *however*), mostly confused with Comparison.Contrast

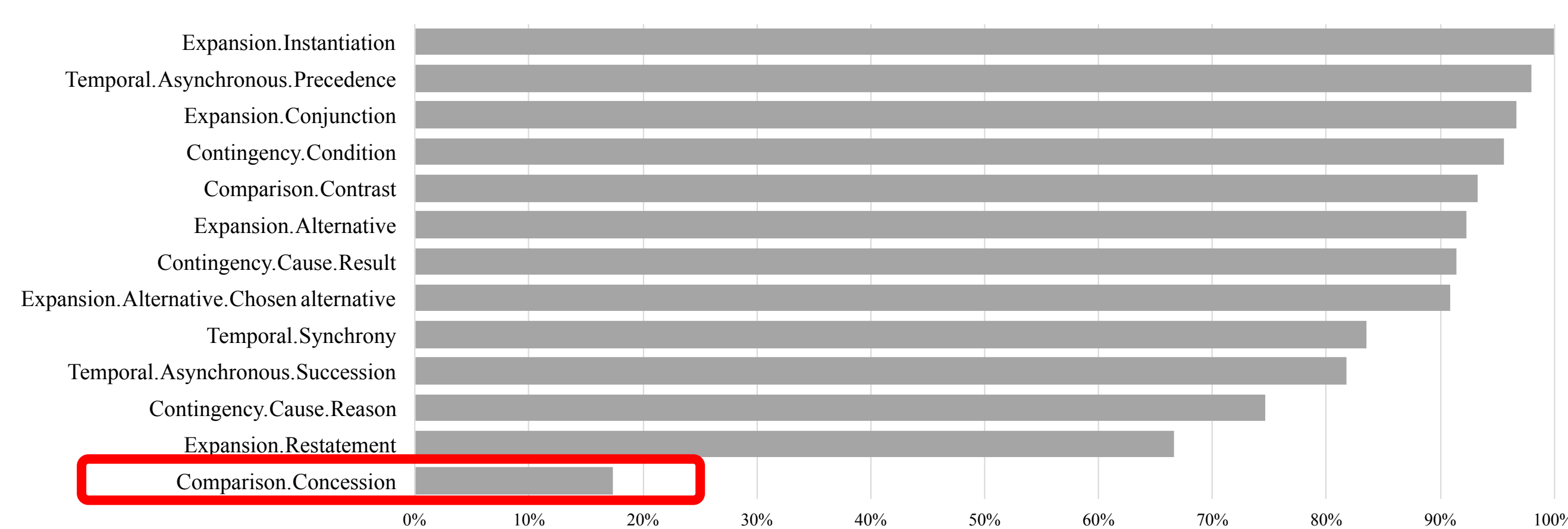


Figure C. Accuracy of baseline majority classifier

Implicit/EntRel Discourse Relations

- Most challenging part as they have no connective words in text
- More EntRel relations than Implicit relations with the most popular sense
→ Dealing with data imbalance is important task

SYSTEM OVERVIEW

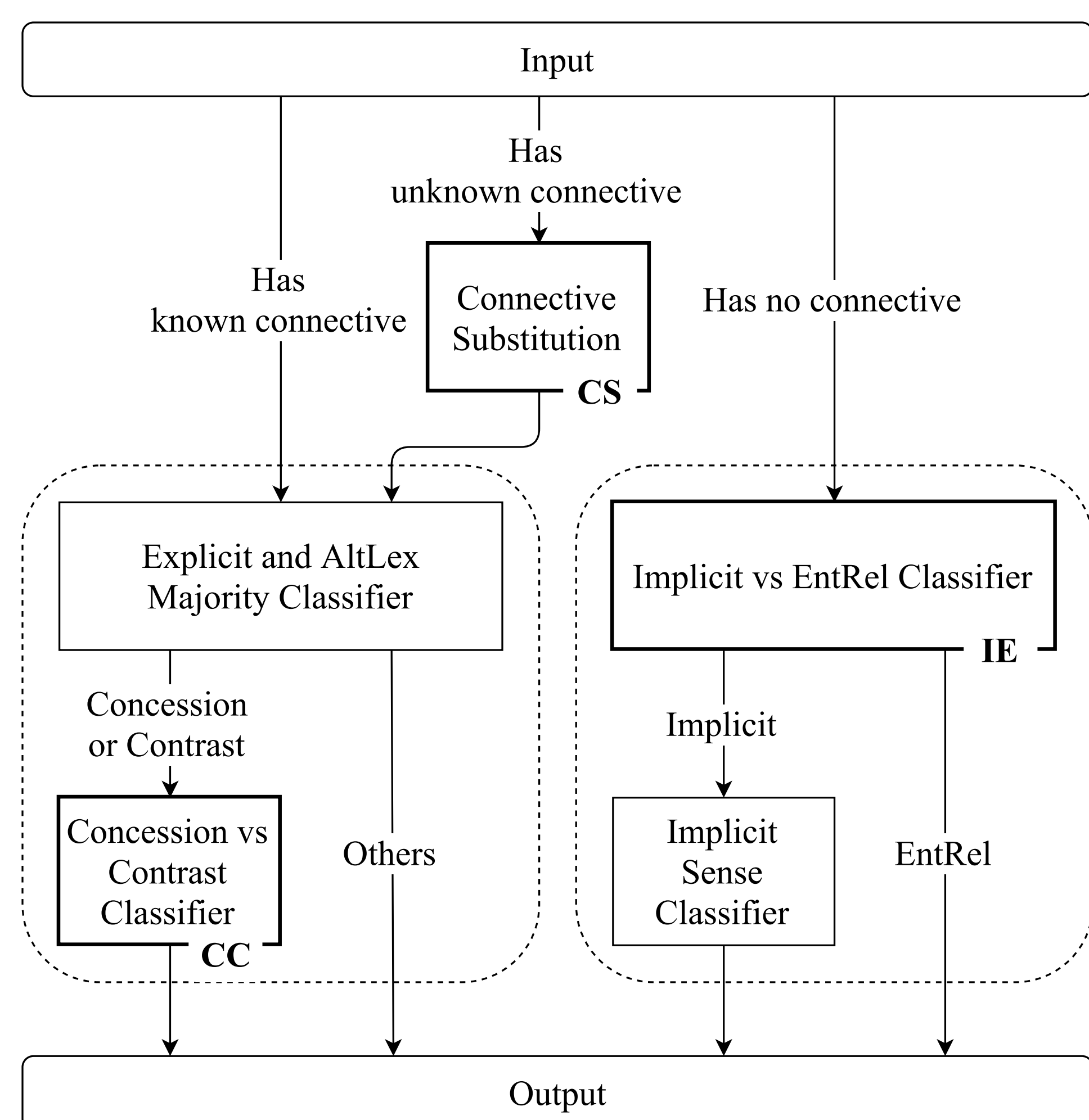


Figure C. Pipeline of our system

Connective Substitution

- If connective words were unseen in training data, majority voting can't classify them
- Unknown connective is substituted with the closest connective** that is known, using skip-gram neural word embeddings (Mikolov et al., 2013)
- Owing to this Explicit and AltLex can be dealt with in the same manner, which contributes to reusability of succeeding modules

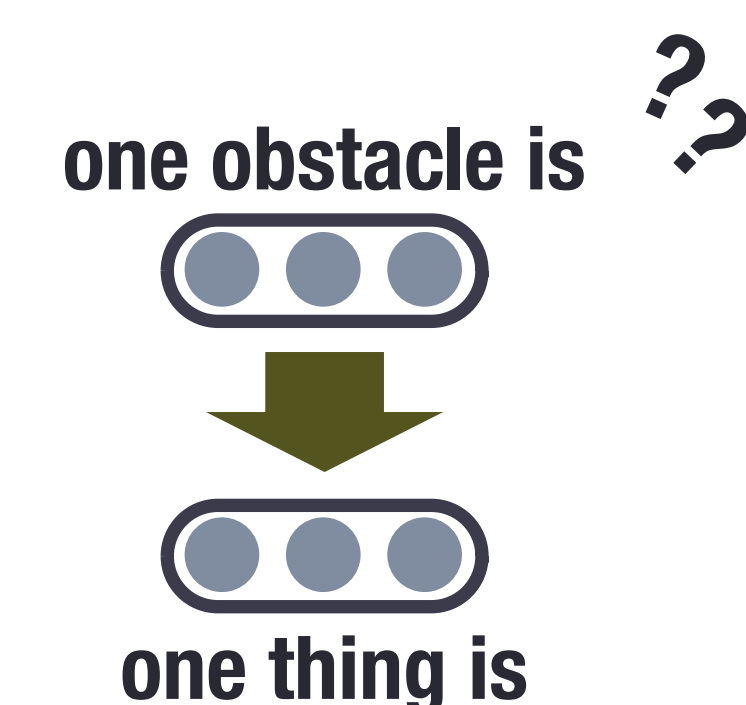


Figure D. Finding closest connective

Concession vs Contrast Classifier

- Redo classification** after majority classifier when its output was Comparison.Concession or Comparison.Contrast in order to reduce confusion
- SVM/MaxEnt Classifier using following as features:

- connective words
- Arg1/Arg2 text bag-of-words
- Arg1/Arg2 parse tree nodes
- MPQA subjectivity lexicon

Implicit vs EntRel Classifier

- Separate EntRel from Implicit before classifying sense in order to deal with unbalanced data better
- Features and classification algorithm shared with CC classifier

RESULTS

F_1 -score	test			blind-test		
	All	Explicit	NonExp	All	Explicit	NonExp
Maj+SVM (Baseline)	0.5116	0.8991	0.1589	0.4404	0.7495	0.1776
Maj+SVM (TIRA Official)	0.5473	0.9022	0.2261	0.5188	0.7543	0.3231
Maj+MaxEnt	0.6093	0.9002	0.3445	0.5215	0.7532	0.3247
Maj+MaxEnt+CS	0.6145	0.9046	0.3504	0.5257	0.7622	0.3241
Maj+MaxEnt+CS+IE	0.5540	0.9046	0.2340	0.5290	0.7622	0.3308
Maj+MaxEnt+CS +CC	0.5866	0.8460	0.3504	0.5357	0.7838	0.3241
Maj+MaxEnt+CS+IE+CC	0.5261	0.8460	0.2340	0.5389	0.7838	0.3308

Experimental Results (Official Evaluator)

CS: Unknown connective substitution, IE: Implicit vs EntRel classifier, CC: Concession vs Contrast classifier

Unknown Connective Substitution

(Top 3) These connective words were unseen in training data, but correct senses were guessed successfully using closest known connective words (Bottom 2) Module failed to find the right connective for substitution, thus wrong sense was output

Unknown Connective	Closest Connective	Output	Golden Sense
the delay resulted from that change will obviously impact that rise came on top of one reason is that one reason is	the rise resulted from that will cinch on top of that that is why one is	Contingency.Cause.Reason Contingency.Cause.Result Expansion.Conjunction Contingency.Cause.Result Expansion.Instantiation	Contingency.Cause.Reason Contingency.Cause.Result Expansion.Conjunction Contingency.Cause.Reason Contingency.Cause.Reason

CONTRIBUTIONS

- Simpler but effective classification algorithm for Explicit and AltLex discourse relations based on their data characteristics
- Effectiveness of majority classifier
- Method for dealing with unknown AltLex connective words using word embeddings
- 8th in official ranking, 1st in (unofficial) test data ranking on Explicit sense classification