

THK's Natural Logic-based Compositional Textual Entailment Model at NTCIR-10 RITE-2

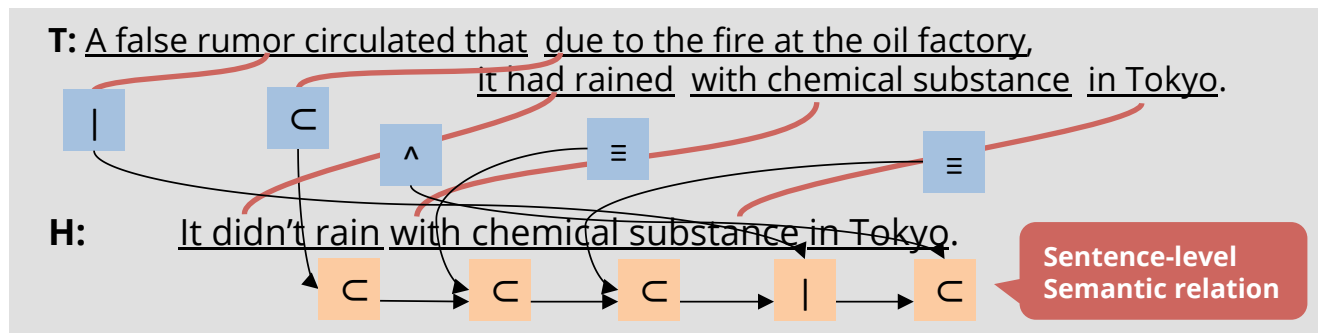
Yotaro Watanabe, Junta Mizuno and Kentaro Inui
 {yotaro-w,junta-m,inui}@ecei.tohoku.ac.jp

Overview

- ◆ Our system learns plausible transformations of pairs of *Text* t_1 and *Hypothesis* t_2 only from semantic labels of the pairs using a discriminative probabilistic model combined with the framework of **Natural Logic**
- ◆ We achieved the highest contradiction detection performance in MC subtask (28.57 of F1)

Natural Logic (NL) for RTE

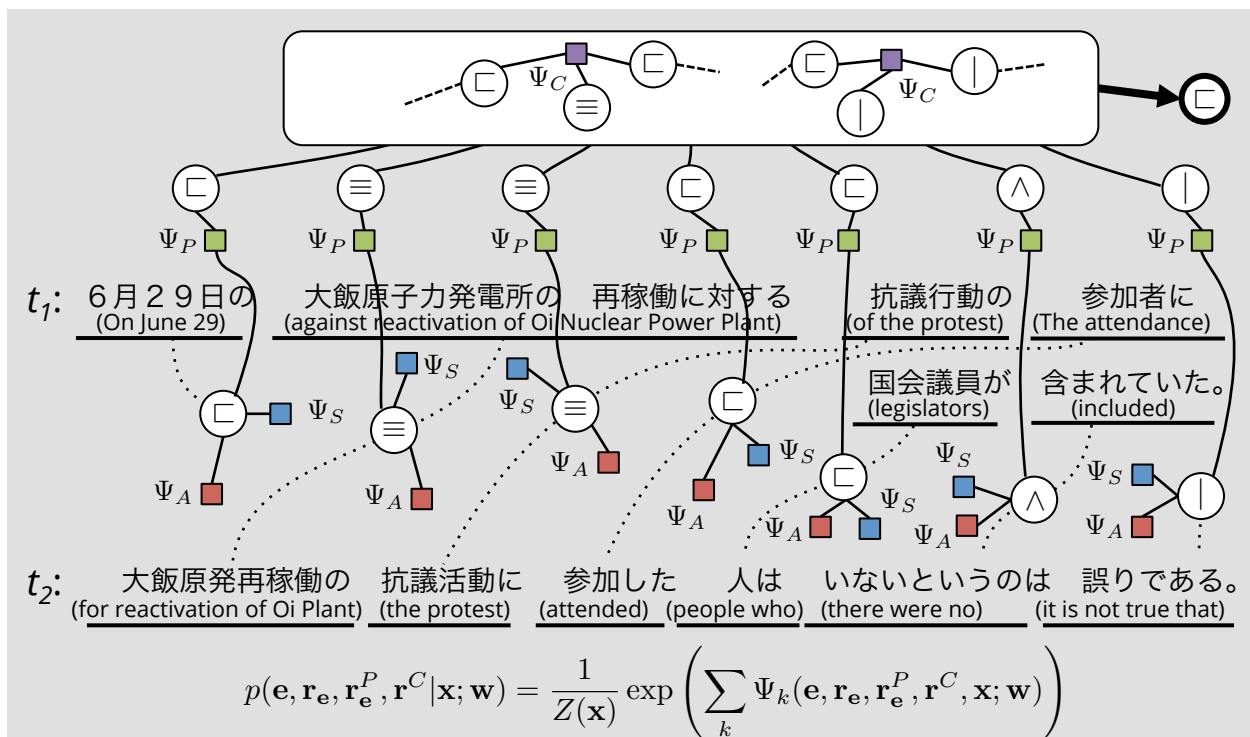
- ◆ Originally proposed by [MacCartney and Manning 2008]
- ◆ Assumption: the entailments of a compound expression are a function of the entailments of its parts (**compositionality**)
- ◆ A transformation from T to H represents a decomposition of its parts and is represented by a set of alignment edits (insertion, deletion and substitution)
- ◆ Sentence-level semantic relations are inferred from semantic relations of alignment edits



Seven types of Relations in NL

- Equivalence** (couch \equiv sofa)
- Forward-ent.** (crow \subset bird)
- Backward-ent.** (bird \supset crow)
- Negation** (human \wedge non-human)
- Alternation** (dog | cat)
- Cover** (animal \cup non-human)
- Independence** (hungry # hippo)

Model



- ◆ The model provides a conditional joint distribution of alignment edits, their semantic relations, their projected relations and the final semantic relation between T and H
- ◆ Two sentences are aligned using an extended MANLI algorithm [Watanabe+ 12]
- ◆ Training the model: maximization of marginal likelihood $\mathcal{L}_\lambda = \sum_n \log p(r_T^C = l^n | \mathbf{x}^n; \lambda)$

Factors

- Alignment factor Ψ_A** TRAIN provides plausibility of each (unlabeled) alignment edits
- Alignment Semantic Relation Factor Ψ_S** TRAIN provides plausibility of semantic relation r_e of each alignment edit e
- Projection Factor Ψ_P** INITIAL provides an appropriate projection from r_e to r_e^P by considering the context of e
- Composition Factor Ψ_C** INITIAL encodes the set of composition rules of semantic relations defined in [MacCartney 2009]

TRAIN: parameters in the factors are learned from training data
INITIAL: parameters in the factors are left to initial values

Results

- ◆ THK-01: the model was trained with the corresponding development data of each subtask.
- ◆ THK-02: the model was trained with the MC-dev data

* : unofficial results

BC	Macro F1	Acc.	Y-F1	Y-Prec	Y-Rec	N-F1	N-Prec	N-Rec
THK-02 (*)	58.34	58.69	62.16	50.49	80.86	54.51	75.50	42.66
THK-01	52.40	53.28	45.92	44.65	47.27	58.87	60.18	57.63

ExamBC	Macro F1	Acc.	Y-F1	Y-Prec	Y-Rec	N-F1	N-Prec	N-Rec
THK-02 (*)	46.59	46.65	48.38	38.62	64.74	44.80	61.39	35.27
THK-01	43.77	62.28	11.52	61.11	6.36	76.03	62.33	97.45

UnitTest	Macro F1	Acc.	Y-F1	Y-Prec	Y-Rec	N-F1	N-Prec	N-Rec
THK-02 (*)	56.59	73.86	83.93	91.16	77.83	29.21	21.67	44.83
THK-01	53.26	71.37	82.35	89.94	75.94	24.18	17.74	37.93

- ◆ Fine-grained semantic labels provided effective information for estimating better parameters in alignment
- ◆ The model failed to assign the relation to delete or insert meaning-less expressions (e.g. 主に (mainly))

MC	Macro F1	Acc.	B-F1	B-Prec.	B-Rec.	F-F1	F-Prec.	F-Rec.	C-F1	C-Prec.	C-Rec.	I-F1	I-Prec	I-Rec.
THK-01	30.98	49.09	21.95	75.00	12.86	60.75	47.77	83.41	28.57	52.17	19.67	43.63	54.61	36.32