

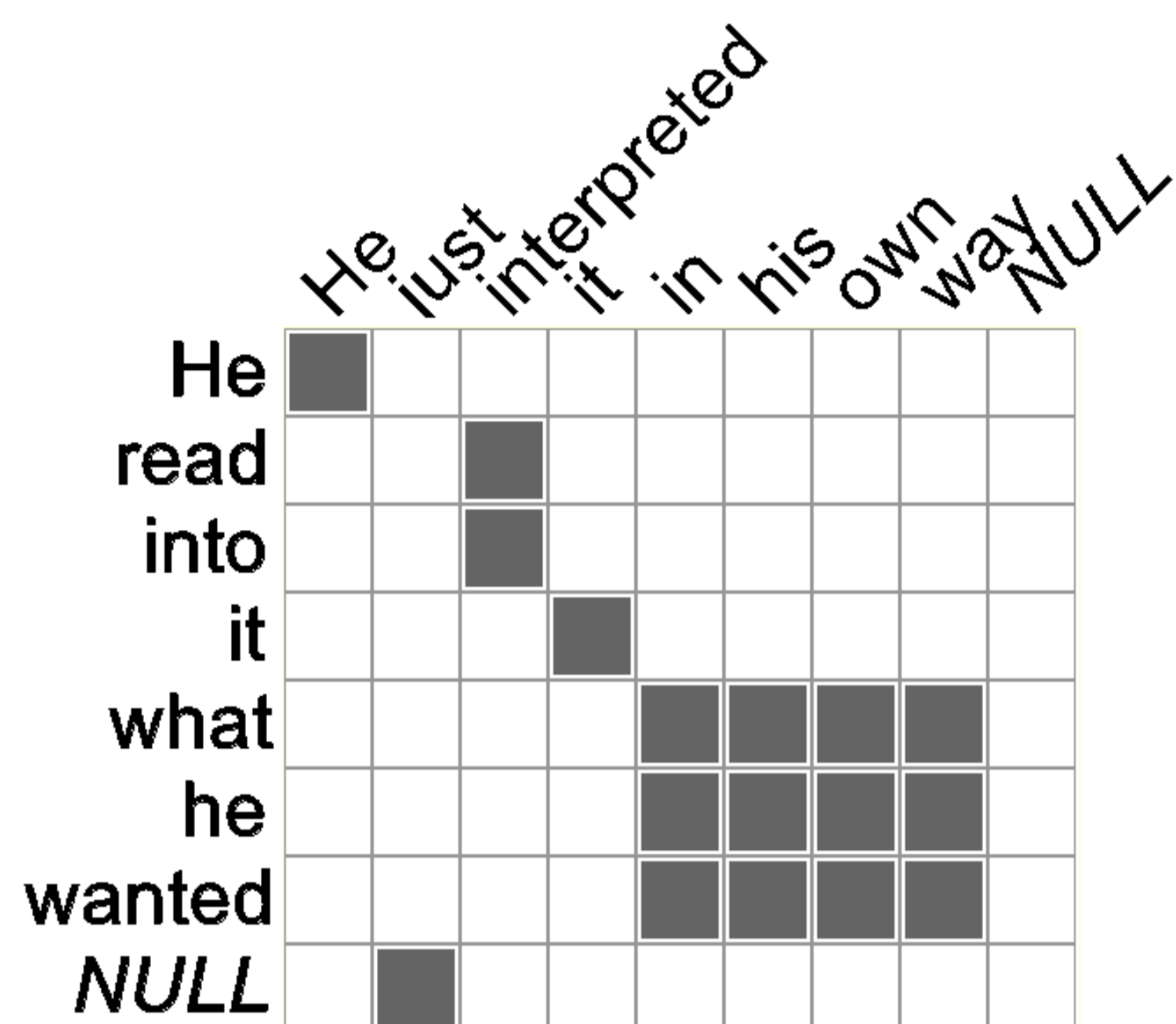
# BCMI-NLP Labeled-Alignment-Based Entailment System for NTCIR-10 RITE-2 Task



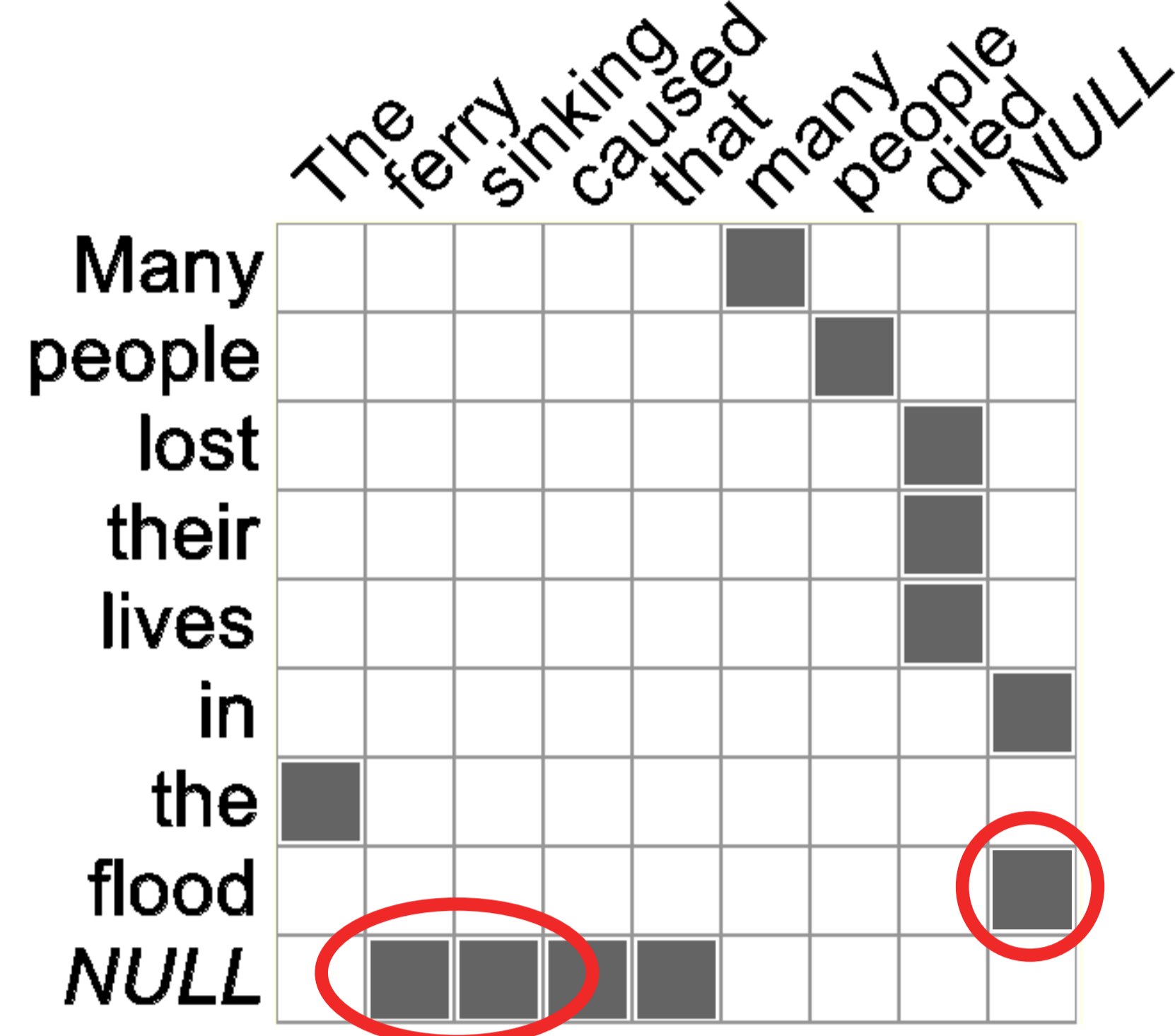
Xiao-lin Wang, Hai Zhao, and Bao-liang Lu  
 Center for Brain-like Computing and Machine Intelligence,  
 MOE-Microsoft Key Laboratory of Intelligent Computing and Intelligent System  
 Department of Computer Science and Engineering  
 Shanghai Jiao Tong University  
 arthur.xl.wang@gmail.com, {zhaohai,blu}@cs.sjtu.edu.cn

## 1. Labeled Alignment Scheme

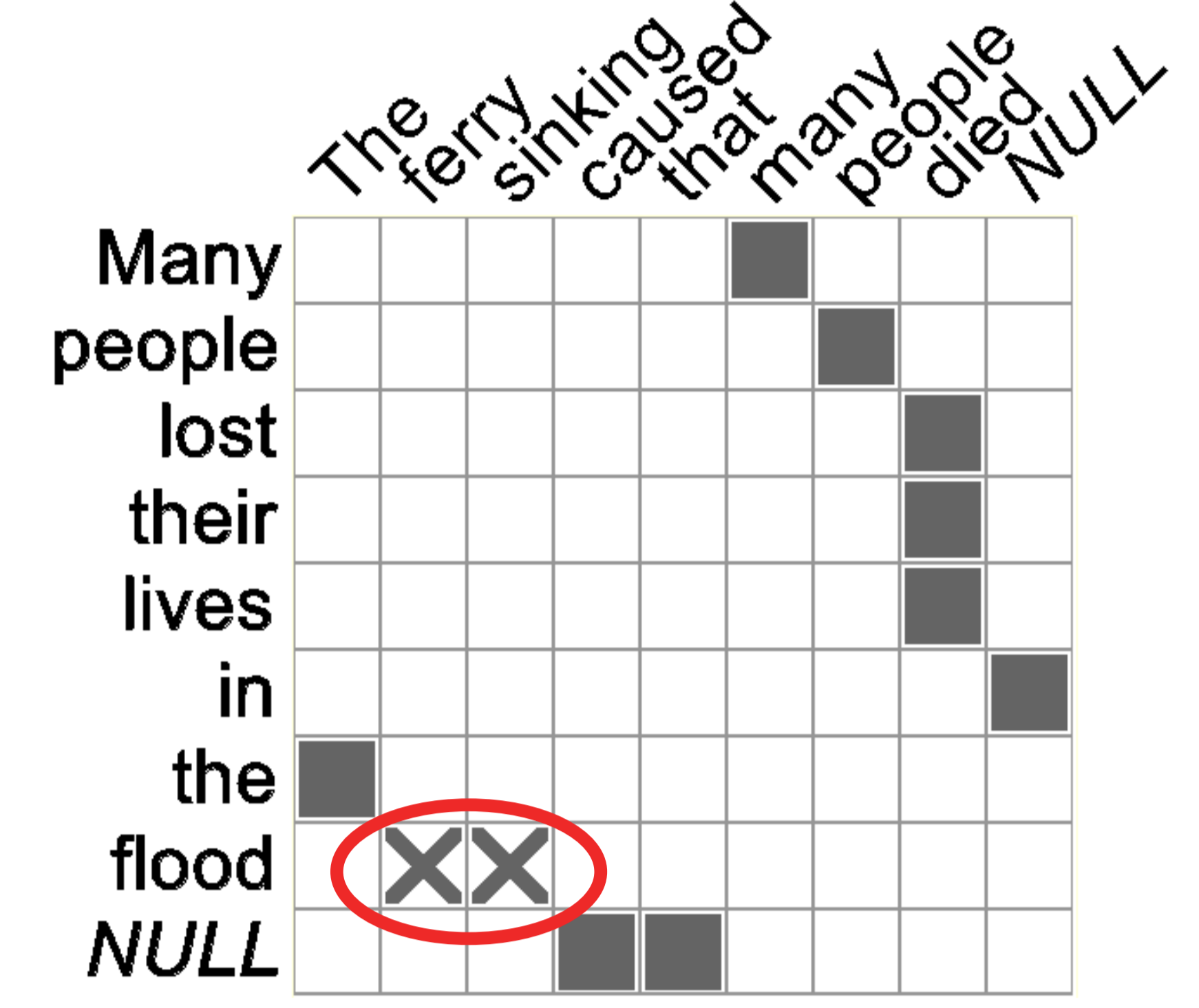
Suppose each subfigure presents an RTE sample. The vertical text is the  $t_1$ , and the horizontal text is the  $t_2$ . The solid squares represent normal links, and the crosses represent negative links which is introduced by labeled alignment scheme.



(a) an entailment pair can be well justified by normal alignment scheme

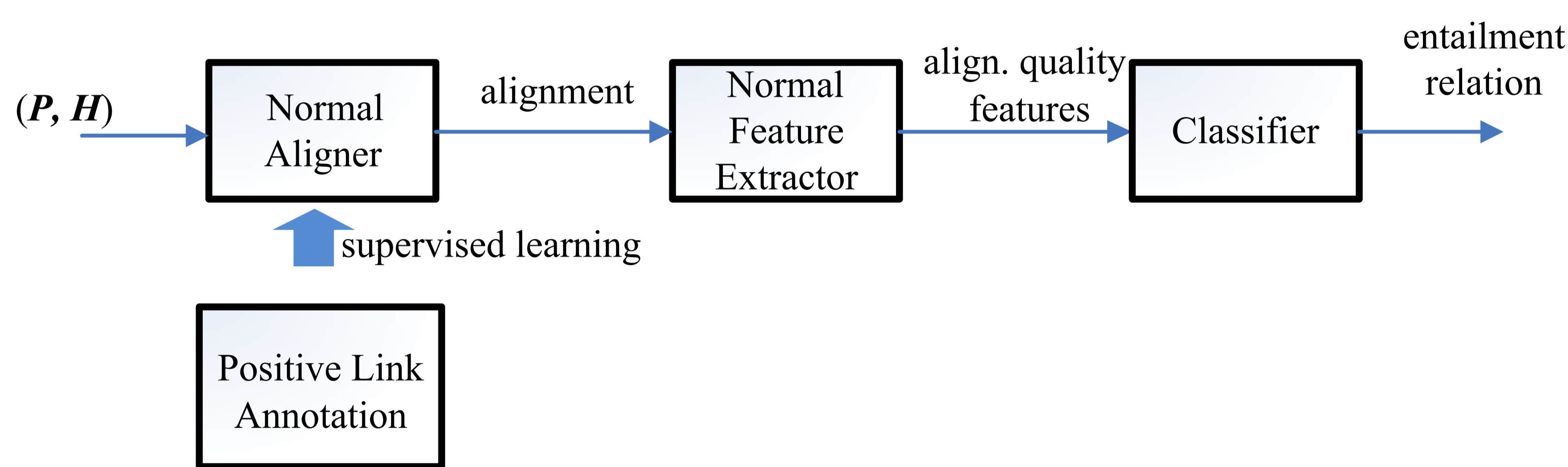


(b) a non-entailment pair is not justified so well by normal alignment scheme.



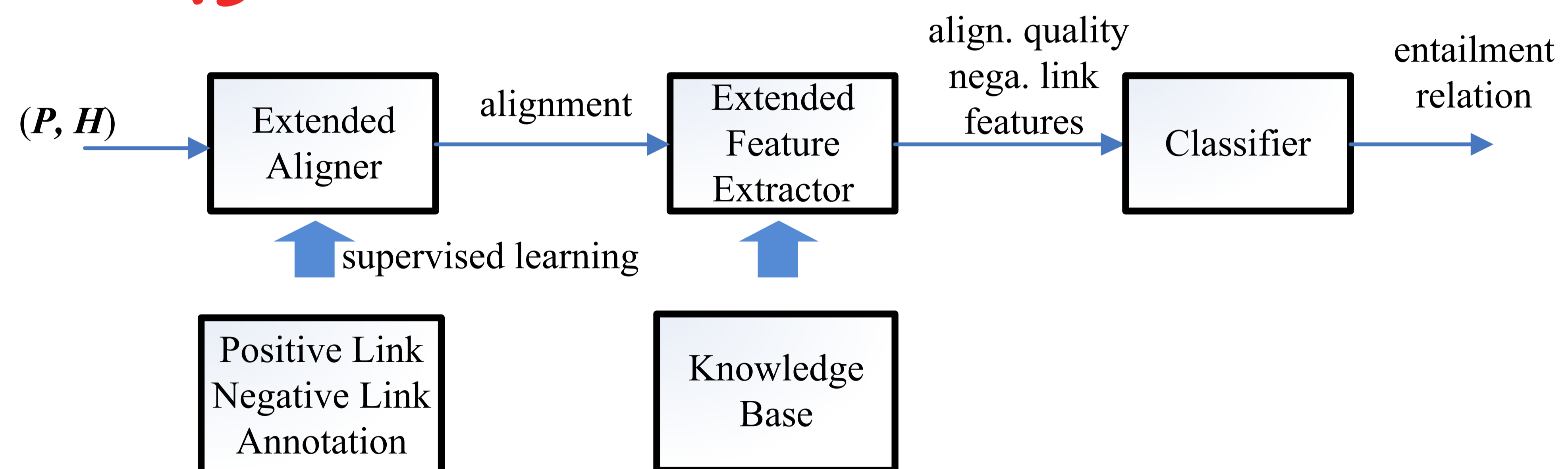
(c) a non-entailment pair can be justified by labeled alignment scheme

## 2. Alignment-based Entailment Systems



(a) Baseline RITE system based on normal alignment scheme

VS



(b) Proposed RITE system based on labeled alignment scheme that better solves non-entailment pairs

## 3. Classification Component

### Link Type Features:

- Whether  $e_1$  and  $e_2$  are in an antonym list
- Whether  $e_1$  and  $e_2$  are in an synonym list
- Whether  $e_1$  and  $e_2$  are unequal numbers
- Whether  $e_1$  and  $e_2$  are different named entities
- Relation of  $e_1$  and  $e_2$  in an ontology (hyponym, sibling, etc.)
- Ontology-based similarities of  $e_1$  and  $e_2$  (CiLin, Hownet)
- Count of common characters
- Length of the common prefix
- Length of the common suffix
- Tuple of the syntactic tags
- Tuple of the ancestors in an ontology
- Tuple of whether  $e_1$  or  $e_2$  is in a list of negative expressions
- Tuple of whether  $e_1$  or  $e_2$  is the head of a noun phrase

Open question :

Why a cascaded system that  
 - first classifies the link type  
 - then classifies the RITE relation  
 performs poorly ?

(Its results are worse than the baseline according to our experiment)

### Sample Representaion:

Single flat vector that combines the features extracted from all the links of the automated alignment

### Classifier:

RBF-kernel SVM (LibSVM) with the default parameters  
 1-vs-rest framework for the MC task

## 4. Evaluation Results

Run	Macro-F1 on BC	Macro-F1 on MC	Worse Rank. on RITE4QA
Run01(char-overlap)	67.04	39.95	6.67 *
Run02(normal align.)	66.89	44.88	0.00 *
Run03(labeled align.)	73.84	56.82	3.67 *

(\* not checked during the formal run due to the limit of time, might be improved in the future)