KNDTE: A System for Deduction of Textual Entailment and Textual Fact at the NTCIR-11 RITE-VAL

Tao-Hsing Chang, Shih-Kuei Kao, Yen-Wen Lin, Li-Jia Wei, Chung-Yin Tsai
Department of Computer Science and Information Engineering
National Kaohsiung University of Applied Sciences, Taiwan

DETERMINING TEXTUAL ENTAILMENT RELATIONSHIPS

The structure of the decision tree that we used to predict textual entailment relationship is shown in Figure 1. Six features of textural pairs as proposed by Chang et al. are adopted and three of six features are improved by this paper. These features are used as the basis for the proposed decision tree. Improved features and new features are show in below.

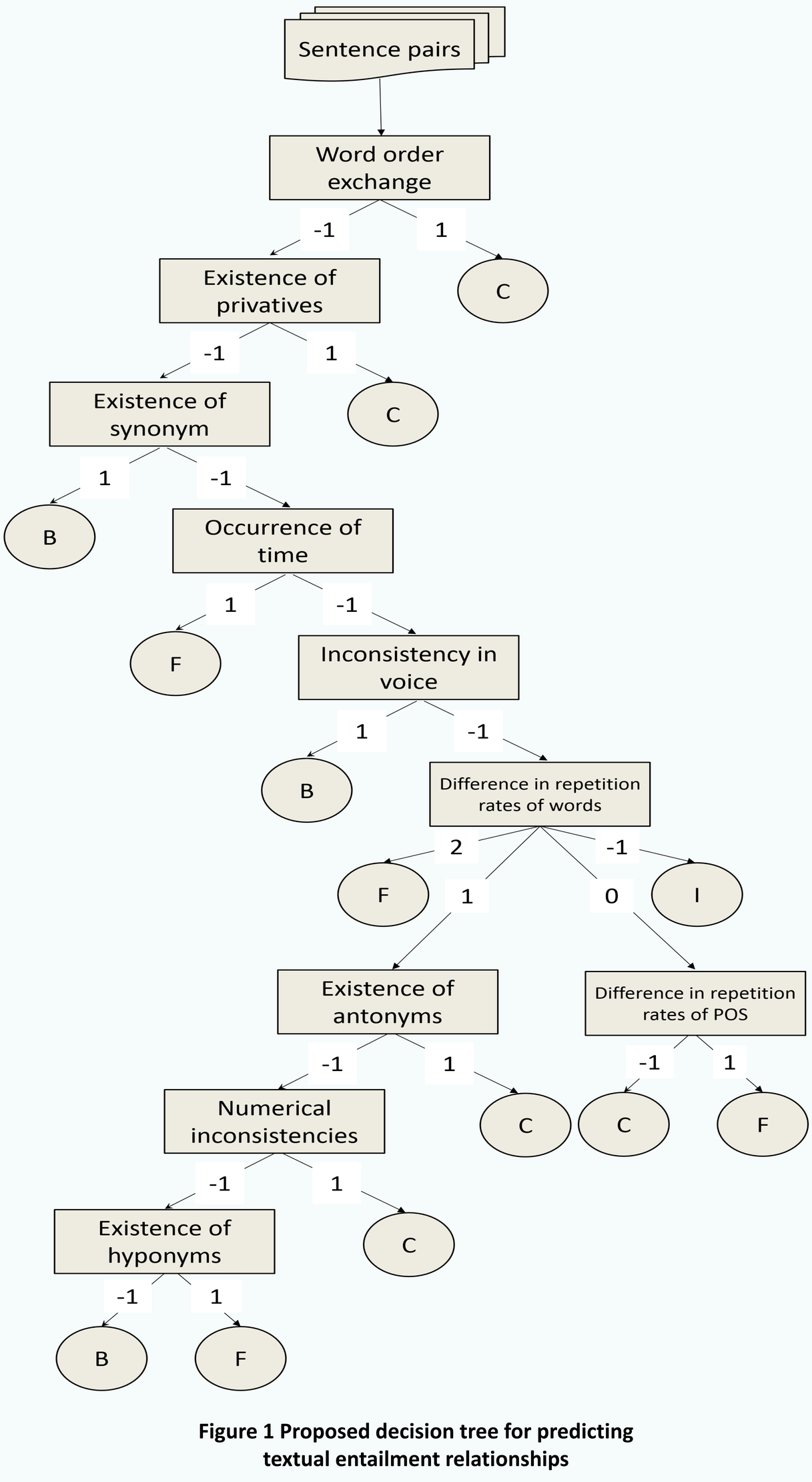

Figure 1 Proposed decision tree for predicting textual entailment relationships

1. Improved features

1.1 Existence of privatives (EPV)

This definition of this feature as follows: if the value for feature "consistency in nouns" of a textual pair is 1, and a privative appears is one sentence but not the other, the value for this feature is 1; otherwise, it is -1.

1.2 Difference in repetition rates of words (DRO)

The definition of this feature as follows:

$$DRO = \begin{cases} 2, if(RWF \geq TI \ or \ RWB \geq TI) \ and |RWB - RWF| \geq TQ \\ 1, if(RWF \geq TE \ or \ RWB \geq TE) and |RWF - RWB| \leq TD \\ 0, if(RWF \geq TI \ or \ RWB \geq TI) \ and (DRO \neq 2 \ and \ DRO \neq 1) \\ -1, otherwise \end{cases}$$

where the definition of RWF and RWB is the same as Chang et al., whereas TI, TE, and TD are threshold with the value of TE≥TI.

1.3 Difference in repetition rates of POS (DOP)

The definition of this feature as follows:

$$DOP = \begin{cases} 1, if(RPF \geq TP \ or \ RPB \geq TP) and |RPF - RPB| \geq TK \\ -1, otherwise \end{cases}$$

where the definition of RWF and RWB is the same as Chang et al., whereas TP and TK are threshold values.

2. New features

2.1 Inconsistency in voice (IVO)

The feature is defined as follows: if the value for feature " consistency in nouns" of a textual pair is 1, with one sentence containing the character "被" (by, a passive indicator) but not the other, the value for this feature is 1.

2.2 Existence of antonyms (EAN)

The feature is defined as follows: if antonyms exist in the two sentences of a textual pair, the value for this feature is 1; otherwise, it is -1. The vocabulary of antonyms used in our study comprises a total of 445 groups of words collected from multiple sources.

2.3 Numerical inconsistencies (NIC)

The feature is defined as follows: if the numerals that appear in the same position for both sentences are inconsistent, the value is 1; otherwise, it is -1.

2.4 Existence of hyponyms (EHY)

This feature is defined as follows: for different words in two sentences of a textual pair, if one word is a hyponym of the other, the value for this feature is 1. If that is not the case, the value is -1. For two different words located in the same position in the sentences , their knowledge categories are separately checked using the Wikipedia query function to determine whether one is a hyponym of the other. If their categories both appear under the same upper knowledge category within a limited number of layers, a hyponym exists.

AN ALGORITHM FOR CONFIRMATION OF TEXTUAL FACT

Step 1: Use WECAn to segment the text into words and tag the POS of the words, While retaining all nouns.

Step 2: Collate from the knowledge base all pages that contain those nouns.

Step 3: Tag the text according to the following rules:

Stet 3.1: E if any of the pages contains more than two-thirds of its nouns and half of its verbs.

Step 3.2: C if any of the pages contains more than two-thirds of its nouns, but less than half of its verbs.

Nouns refer to the POS Na, Nb, and Nc as defined in WECAn, whereas verbs refer to VA, VB, and VC.

EXPERIMENTAL RESULTS

The test data used for this study are from the formal run dataset of the RITE-VAL task recorded during the NTCIR-11 Conference. When the proposed method is applied to the formal run stage data set of the SV and FV subtask, the results of the predictions are as listed in Table1.

Table 1 Data set from formal-run stage of SV and FV subtasks: prediction results using proposed method

Indicator \ Task	SV-CT-BC		SV-CT-MC				FV-CT		
	Y	N	B	C	F	I	E	C	U
F1	62.59	45.60	50.50	43.00	49.09	13.74	48.99	25.59	27.33
Precision	54.14	58.99	43.95	42.04	42.21	39.06	41.43	26.92	37.27
Recall	74.17	37.17	59.33	44.00	58.67	8.33	59.91	24.38	21.58