

KitAi: Textual Entailment Recognition System for NTCIR-10 RITE2

Kazutaka Shimada, Yasuto Seto, Mai Omura and Kohei Kurihara

Kyushu Institute of Technology (Department of Artificial Intelligence)

NTCIR-10 RITE2 Poster ID - RITE05

Task and basic idea

BC subtask: Given a text pair $\langle t1, t2 \rangle$, identify whether $t1$ entails a hypothesis $t2$ or not

ExamBC subtask: Multiple-choice questions of real university entrance exams

The basic idea: Classification based on word correspondence and syntactic & semantic information

Features

Correspondence

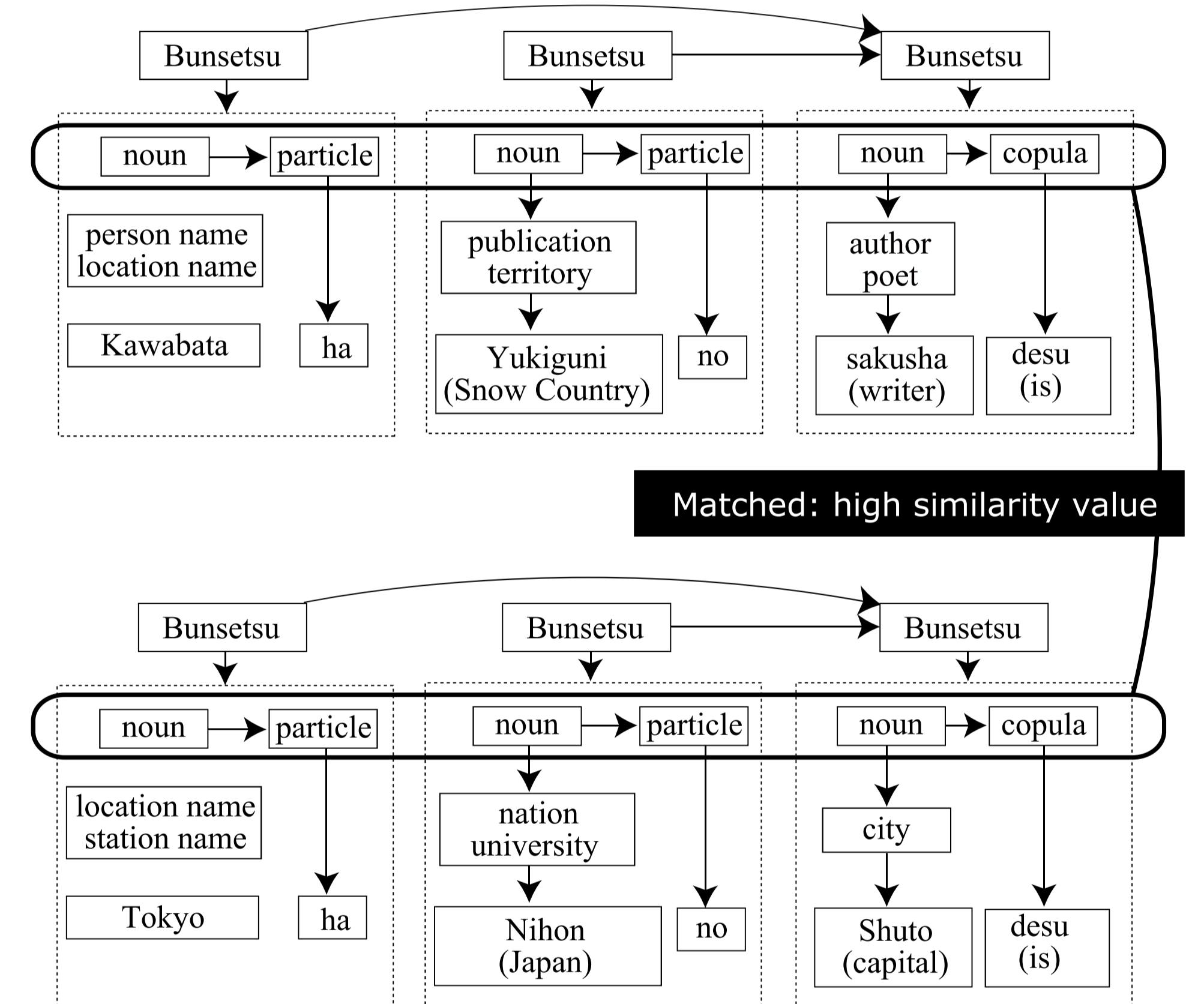
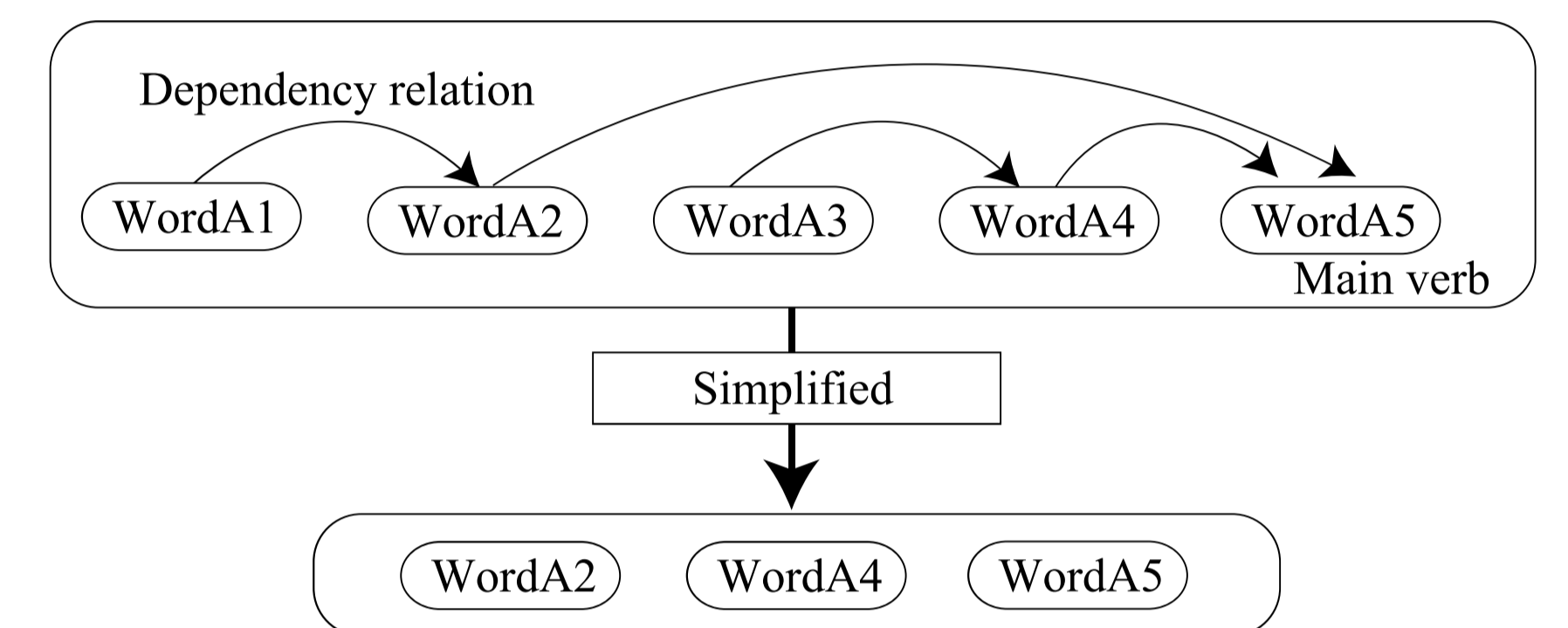
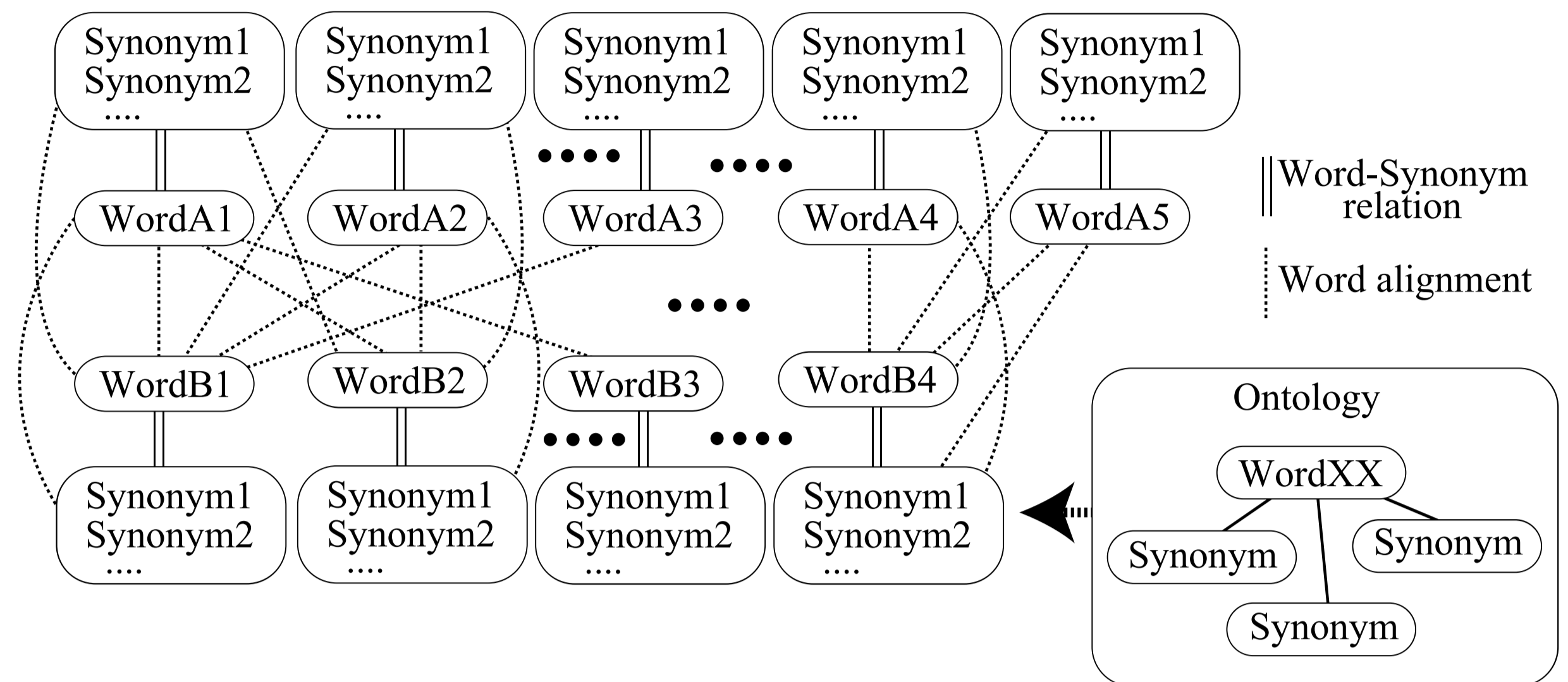
- Word overlap and edit distance between $t1$ and $t2$
- Expansion with ontologies; Japanese WordNet and Nihongo-Goi-Taikai
- All word, content word and simplified word lists

Syntactic and semantic information

- Hierarchical Directed Acyclic Graph (HDAG)

Others

- # of negation words and sentence length



Classifier

BC task

- SVM (MethodBC1), Logistic regression (MethodBC2), Rotation forest (MethodBC3)

ExamBC task

- AdaBoost with SVM (MethodEX1), Logistic regression (MethodEX2), Weighted Voting (MethodEX3)

$$Score = \frac{\alpha \times SVM + \beta \times Logistic + \gamma \times Rotation Forest}{3}$$

Experimental result

Results for development data

BC task

Features	Accuracy	MacroF1
WCall	65.6	58.8
WCont	77.1	76.7
WSimp	59.1	47.0
WCall+WCont	78.1	77.8
WCall+WCont+WSimp	77.9	77.5
W+HDAG	77.9	77.5
W+OTHER	79.5	79.3
ALL	79.4	79.2
ALL without Ontologies	78.4	78.2

ExamBC task

Features	Accuracy	MacroF1
WCall	58.6	43.8
WCont	65.9	64.6
WSimp	58.8	43.6
WCall+WCont	65.3	64.2
WCall+WCont+WSimp	65.4	64.6
W+HDAG	66.7	65.8
W+OTHER	65.5	64.4
ALL	66.5	65.5
ALL without Ontologies	64.1	63.3

Results for formal run

BC task

Method	HDAG	MacroF1
MethodBC1	With	77.1
	Out	76.9
MethodBC2	With	72.4
	Out	75.5
MethodBC3	With	76.2
	Out	76.0

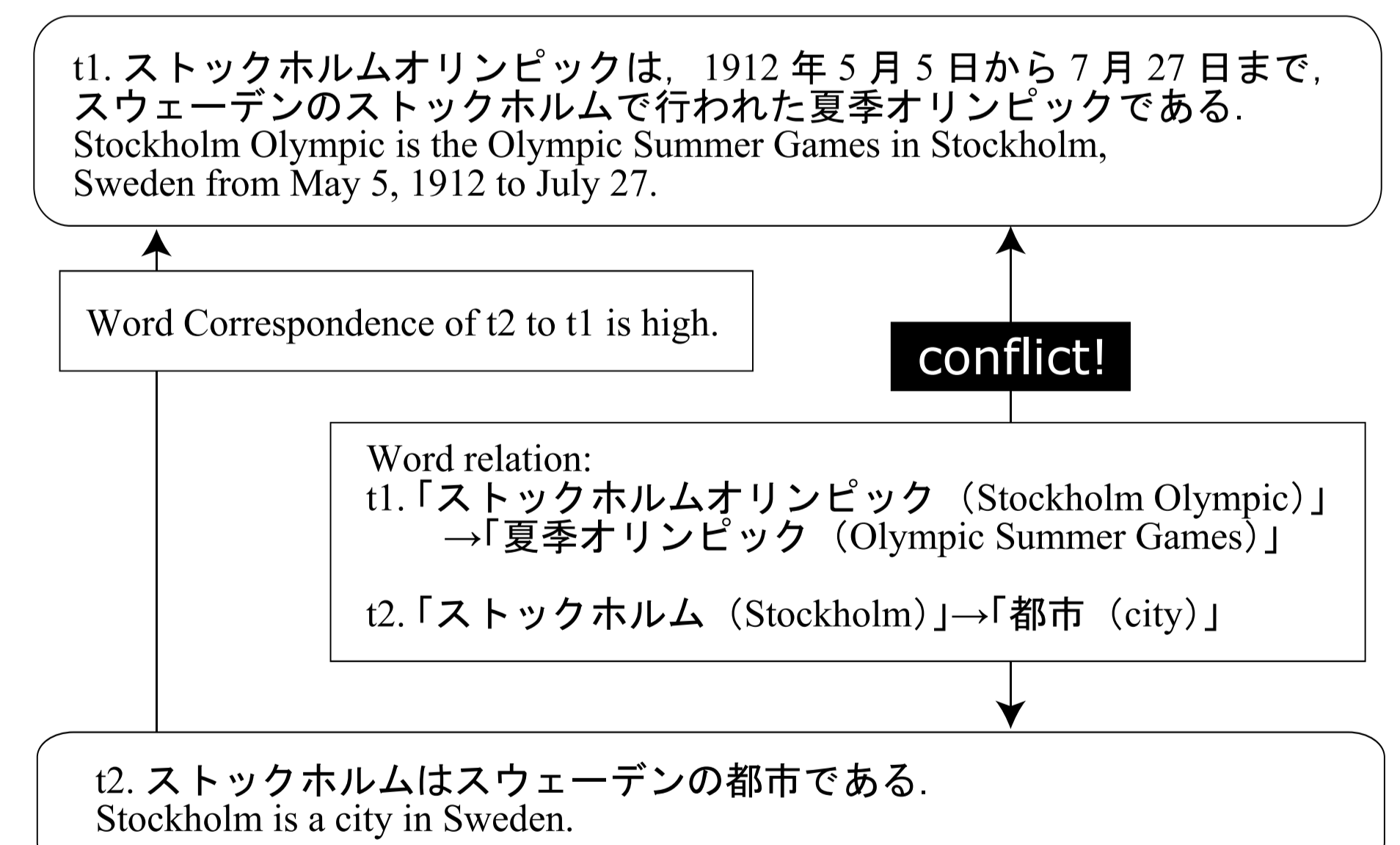
ExamBC task

Method	HDAG	MacroF1
MethodEX1	With	59.8
	Out	61.3
MethodEX2	With	57.2
	Out	57.4
MethodEX3	With	59.1
	Out	60.8

The correct answer ratio on the formal run

Method	MacroF1	CorrectAR
MethodEX1	59.8	36.1
MethodEX2	57.2	39.8
MethodEX3	59.1	45.4

An incorrect instance with HDAG



Discussion and Future Work

- Ontologies are important
- Confidence value is important for the ExamBC task
- Optimal parameters for the voting method
- Structure of HDAG