A SYSTEM OF LEARNING LINKS BETWEEN UTTERANCES WITH META-INFORMATION IN CHARACTER BASED MULTI-PARTY CONVERSATION BY DECISION TREE

Junpei Nakamura, Yasuhiro Tajima and Yoshiyuki Kotani
Department of Electronic and Information Engineering, Tokyo University of Agriculture and Technology
Nakacho 2-24-16, Koganei, Tokyo 184-8588, Japan

Keywords: Web mining, Web dialogue systems, Metadata and Metamodeling.

Abstract: It is important to find out interactive links between pairs of utterances in multi-party conversation like an online chat. Though the usage of linguistic information is necessary to do this, we showed the better performance to this criterion by using physical meta-information that consists of the number of conversation members, the distance between utterances, and the frequency of individual utterance. The result of the examination of Decision Tree learning showed the accuracy is 84.1%, the precision is 87.1% and the recall is 66.9% for link between same speaker's utterances, and the accuracy is 77.5%, the precision is 68.5% and the recall is 56.8% for link between others' utterances. The result of the examination without meta-information showed the accuracy is 62.7%, the precision is 86.7%, the recall is 64.5% for same speaker's utterances, and the accuracy is 77.9%, the precision is 73.8% and the recall is 49.9% for others'. These results showed we could find new links by using meta-information.

1 INTRODUCTION

There are many character data in the Internet, for example online chat log data. These are very important language resources. However, it is difficult to analyze these data, because its structure is complex. Therefore, the research analyzing its structure becomes more important. As a part of the research, finding a semantic links for the multi-party conversation is very necessary.

In this paper, we made a system of finding new links for character based multi-party conversation with linguistic information and some meta-data. We had the link learned by the Decision Tree. The result of the examination of Decision Tree learning showed the accuracy is 84.1%, the precision is 87.1% and the recall is 66.9% for link between same person's utterances, and the accuracy is 77.5%, the precision is 68.5% and the recall is 56.8% for link between others' utterances. The result of the examination without meta-information showed the accuracy is 62.7%, the precision is 86.7%, the recall is 64.5% for same person's utterances, and the accuracy is 77.9%, the precision is 73.8% and the recall is 49.9% for others'. These results showed we could find new links by using meta-information.

Our system will help users to understand complex character-based multi-party conversation.

2 DEFINE OF LINKS BETWEEN UTTERANCES

The dialogue is structured the list of conversation participants M and the list of utterances U. The list M is expressed the following formula (1). M_NUM is a number of participants, and 2 or more without fail.

$$M = \{m_1, m_2, \cdots, m_{M\_NUM}\} \quad (1)$$

The list U is expressed the following formula (2). In the formula (2), u_t is a utterance at a time t, and a Speaker of u_t is expressed $m(u_t)$.

$$U = \{u_1, u_2, \cdots, u_n\} \quad (2)$$

We define that there is a link between utterance u_i and utterance u_j, when there is semantic relation between u_i and u_j, or u_j is caused by u_i.

Nakamura J., Tajima Y. and Kotani Y. (2007).
A SYSTEM OF LEARNING LINKS BETWEEN UTTERANCES WITH META-INFORMATION IN CHARACTER BASED MULTI-PARTY CONVERSATION BY DECISION TREE.
In *Proceedings of the Third International Conference on Web Information Systems and Technologies - Web Interfaces and Applications*, pages 420-424
DOI: 10.5220/0001280304200424

We show an example of the links between utterances as follows (Figure1). In the Figure 1, the arrows indicate the links between utterances.

T	$m(u_t)$	u_t
1	*A*	You should ready your graduation thesis from now. > *C*
2	*B*	Form the second grade of junior high-school?
3	*C*	I'm still 13 years old. > *A*
4	*A*	Wow! You have 9 years! You can easy to write it. > *C*

Figure 1: Example of links between utterances.

In Japanese online chat, they specify the recipients of their utterance using the special sign (e.g. ">").

The links exist not only between others' utterances, but also between same person's utterances (Figure2).

t	$m(u_t)$	u_t
1	*A*	I must submit this paper by 15:00 o'clock.
2	*A*	Oh! The time limit is 14 o'clock, not 15.
3	*B*	Hurry!

Figure 2: Example of links between same person's utterances.

Two or more sentences might be included in one utterance. And each sentence might have a link to different utterance. Therefore, we divided sentences so that an utterance consists of a sentence.

3 LEARNING OF LINKS BETWEEN UTTERANCES

In this paper, we tried to learn the links between utterances by C4.5 Decision Tree (J.R. Quinlan, 1993) from linguistic information and some meta-information. The meta-information should be easy to get from any dialogue data. We suggested 3 elements for meta-information, the number of conversation participants (NUM), the distance between utterances (DIS), and the frequency of individual utterance (FREQ).

We use 5 elements; consist of 2 linguistic elements and 3 elements of meta-information following Table1, as attribute for Decision Tree.

Table 1: 2 linguistic elements and 3 elements of meta-information.

Information	Elements
(A) Linguistic	(A-1) Relation of sentences
	(A-2) Information of recipients
(B) Meta	(B-1) NUM
	(B-2) DIS
	(B-3) FREQ

3.1 Linguistic Information

The linguistic information is obtained by natural language processing, e.g. morphological analysis, syntactic analysis, and pattern matching and so on. We use 2 elements from linguistic information; the relation of sentences and the information of Recipients.

3.1.1 Relation of Sentences

The relation of sentences expresses the possibility of relation between two utterances. In many other researches, this relation is main element to detect the links, and approximated by various elements, for example words, morpheme, heuristic rule, etc. In this paper, we simply use bi-gram of Dialogue Acts (DA) (Andress Stolcke, et al., 1998; Inui N., et al., 2001; Edward Ivanovic, 2005) as the relation of sentences. We use 18 kind of DA (Inui N., et al., 2001): greeting, farewell, opinion, intention, fact explanation, reason, question (wh), question (yes-no), check, request, suggest, affirmation, negation, deliberation, apology, surprise, gratitude, and NO_DA.

We consider that the relation of sentences is able to approximate by the bi-gram of DA (Figure4).

t	$m(u_t)$	u_t [DA]
1	*C*	Did you go to the park yesterday? [question (yes-no)]
		A has joined conversation.
2	*A*	Hello. > *all* [greeting]
		Low probability — High probability
3	*B*	No, I didn't. [negation]
4	*C*	Hello > *A* [greeting]

Figure 4: Relation of sentences and bi-gram of DA.

However, the values of bi-gram depend on the situation that the speakers of two utterances are the same or others (Figure5).

Hereafter, we express DA of the utterance at time t as "DAt".

t	$m(u_t)$	u_t [DA]
		A has joined conversation.
1	*A*	Hello. > *all* [greeting]
		High probability
2	*B*	Hi > *A* [greeting]
		C has joined conversation.
3	*C*	Hello [greeting] Low probability
4	*A*	Hello > *C* [greeting]

Figure 5: Difference of bi-gram probability by speakers.

3.1.2 Information of Recipients

The information of recipients is important to detect links from multi-party conversation. In Japanese online chat, they specify the recipients of their utterance using the special sign (e.g. ">"). We can easily gain information of recipients using this expression. We express recipients at utterance u_t as $r(u_t)$.

It is reasonable to suppose that if $r(u_i)$ includes $m(u_j)$ then the link is more likely to exist between utterance u_i and u_j (Figure6).

t	$m(u_t)$	u_t
1	*A*	What did you do yesterday?
2	*B*	I caught a cold. > *A*
3	***A***	**Excuse me. >*C***
4	*C*	Are you OK? > *B*
5	*C*	What? > *A*

u_3 is more likely to link to u_4 and u_5

Figure 6: Possibility of existing links from information of recipients at the utterance u_i.

However, it is possible that utterance u_i link to utterance u_j whose speaker is not recipient of u_i. Thus, we assume that if $r(u_j)$ includes $m(u_i)$ then the link is more likely to exist between utterance u_i and u_j (Figure7).

According to these two rules, the link is the most likely to exist between utterance u_3 and u_5 in Figure6 and 7.

t	$m(u_t)$	u_t
1	*A*	What did you do yesterday?
2	*B*	I caught a cold. > *A*
3	*A*	Excuse me. > *C*
4	*C*	Are you OK? > *B*
5	***C***	**What? > *A***

u_1 and u_3 are more likely to link to u_5

Figure 7: Possibility of existing links from information of recipients at the utterance u_j.

3.2 Meta-information

3.2.1 Number of Conversation Participants and Distance

We examined the relation between NUM and DIS where the links exist from the log data of the conversation whose NUM is 2 to 10. Table 2 shows the rate of utterance to each NUM and DIS. Figure8 is a graph of Table2. The rate of the utterance in which a link does not exist was about 40% by each NUM. In Table2, the shading cells show the maximum distance among cells having over 1%.

Table 2: Rate of utterance to each number of participants and distance between utterances.

Number of participants	Distance between utterances where link exists(j − i): 1	2	3	4	5	6	7	8	9	10	11	12	13	14
2	51	12	3	1	0.5	0.2	0.2	0	0	0	0	0	0	0
3	30	13	5	5	2.5	0	0	0	0	0	0	0	0	0
4	26	11	6	5	4.2	0.6	1.2	0	0.6	0	0	0	0	0
5	25	8	4	7	4.4	2.4	0	1.2	0.4	0	0.4	0.4	0	0
6	15	10	10	8	3.8	1.3	4.5	1.3	0	0	0	0.6	0	0.6
7	12	10	12	9	5.6	4.8	6	4	0.4	1.2	1.2	0.4	0	0
8	13	12	15	7	5.4	3.2	3.8	3.2	1.6	0.5	0	0.5	0	0
9	16	8	5	4	6.3	4.4	1.3	0.6	0.6	0.6	2.5	0.6	0	0
10	9	11	10	7	7.8	4.9	2.5	0	3.4	0.5	0.5	0	0.5	0.5

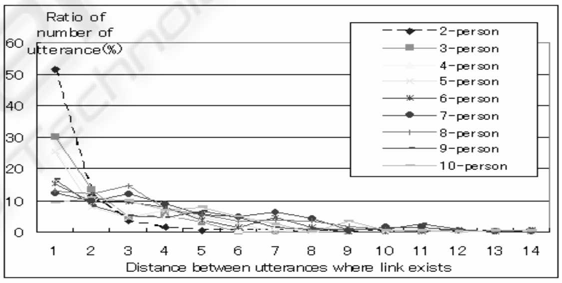

Figure 8: Graph of the rate of utterance to each number of participants and distance between utterances.

From these shading cells in Table2, we realize that the most of cells of maximum distance are located in about three plus each participant. Therefore, we define the maximum distance *N* expressed by the following formula (3).

$$N = M\_NUM + 3 \quad (3)$$

From Figure8, it seems quite probable that the exponential relation exist between NUM and DIS. We consider that NUM emphasizes the influence from DIS.

3.2.2 Frequency of Individual Utterances

We use the FREQ, because we consider that the main participant has lots of utterances. We define the FREQ to following;

"The FREQ is the rate of utterances of participant $m(u_j)$ from u_{i-Ni} to u_{i-1}. Here, Ni is the maximum distance at time *i*."

3.3 Case Data of Learning

First, we collected log data of online chat from the Internet. The log data included 533 utterances by 3-12 participants. Second, we added information of speaker, DA, recipients, and NUM. We attached all suitable tags from 18 kinds about information of DA. If a given utterance included an expression of recipients, we attached all of participants as information of recipients, and we attached "None" if the utterance did not include the expression. The utterances including the expression occupied 38.9% of all. Third, we made the learning data for Decision Tree by the following method (Figure9). As a result, we gained 600 positive samples and 4500 negative samples, and we used 600 positive samples and 1200 negative samples for learning.

```
repeat
   for i = 1 to End_Of_Data -1
   repeat
     for j = i + 1 to i + Ni
       if the link exist between ui and uj
         then positive samples
       else
         then negative samples
     end for
   until j > End_Of_Data
  end for
end
```

Figure 9: Algorithm of making learning data.

3.4 Learning of Links by Decision Tree

We made the systems of Decision Tree learning in the 10 fold cross-validation with the attribute showed Table3. The system of Decision Tree learned with the following attribute pattern in order to examine influence to the links by each attribute.

(1) Linguistic + Meta: (A) + (B)
(2) Only Linguistic: (A)
(3) Linguistic + NUM: (A) + (B-1)
(4) Linguistic + DIS: (A) + (B-2)
(5) Linguistic + FREQ: (A) + (B-3)
(6) Linguistic + NUM + DIS: (A) + (B-1) + (B-2)
(7) Linguistic + NUM + FREQ: (A) + (B-1) + (B-3)
(8) Linguistic + DIS + FREQ: (A) + (B-2) + (B-3)

We illustrated the result of learning in Table4 and Table5. Table 4 shows the result of learning links between same person's utterances, and Table5 shows the result of learning between others' utterances.

Table 3: Attribute for Decision Tree.

Kind of information	Elements	Attribute for Decision Tree
(A)Linguistic	(A-1)Relation of sentences	(A-1-a) DAi
		(A-1-b) DAj
	(A-2)Information of Recipients	(A-2-a) r(ui) = m(uj) ?
		(A-2-b) r(uj) = m(ui) ?
(B)Meta	(B-1)Number of conversation participants	(B-1)M_NUM at time i
	(B-2)Distance between utterances	(B-2) j − i
	(B-3)Frequency of individual utterance	(B-3)Frequency of individual utterance of m(uj) at time i

Table 4: Result of learning links between same person's utterances.

		same person			
	Attributes	accuracy(%)	precision(%)	recall(%)	F value
(1)	ling + meta	83.2	86.7	64.5	0.74
(2)	ling	62.7	49.1	21.5	0.30
(3)	ling + NUM	62.7	49.2	26.4	0.34
(4)	ling + DIS	83.8	87.8	65.3	0.75
(5)	ling + FREQ	63.6	51.4	31.4	0.39
(6)	ling + NUM + DIS	83.2	86.7	64.5	0.74
(7)	ling + NUM + FREQ	63.3	50.6	35.5	0.42
(8)	ling + DIS +FREQ	84.1	87.1	66.9	0.76

Table 5: Result of learning links between others' utterances.

		others			
	Attribute	accuracy(%)	precision(%)	recall(%)	F value
(1)	ling + meta	77.5	68.5	56.8	0.62
(2)	ling	77.9	73.8	49.9	0.60
(3)	ling + NUM	76.6	70.1	49.1	0.58
(4)	ling + DIS	78.2	78.2	45.7	0.58
(5)	ling + FREQ	77.0	71.7	48.2	0.58
(6)	ling + NUM + DIS	77.6	69.6	55.3	0.62
(7)	ling + NUM + FREQ	76.5	69.5	49.5	0.58
(8)	ling + DIS +FREQ	77.8	74.1	48.9	0.59

In Table4 and 5, the accuracy, the precision, the recall and the F_value are defined the following formulas. In these formulas, pp: true positive, pn: false positive, np: false negative and nn: true negative.

$$accuracy = \frac{pp + nn}{pp + np + pn + nn} \quad (4)$$

$$precision = \frac{pp}{pp + pn} \quad (5)$$

$$recall = \frac{pp}{pp + np} \quad (6)$$

$$F\_value = \frac{2 \times precision \times recall}{precision + recall} \quad (7)$$

The sub-tree of Decision Tree is pictured Figure 10 and 11. Figure 10 shows a sub-set of Decision

Tree when the attribute pattern is (8) in Table4, and Figure 11 shows a sub-set when the attribute pattern is (1) in Table 5.

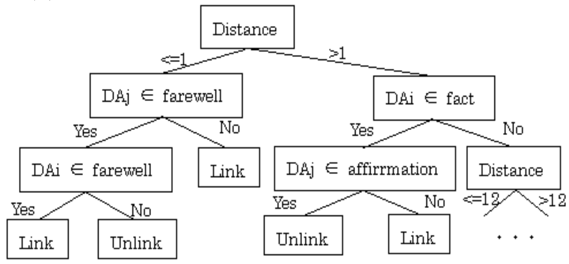

Figure 10: Sub-set of Decision Tree when attribute pattern is (8) in Table 4.

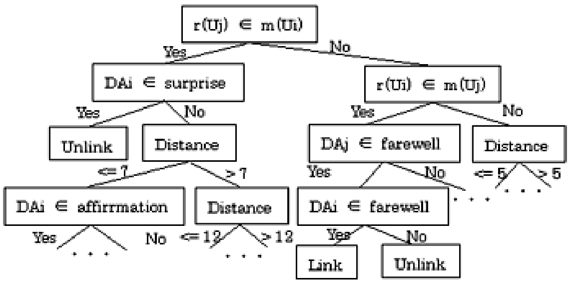

Figure 11: Sub-set of Decision Tree when attribute pattern is (1) in Table 5.

4 CONSIDERATION

In the matter of detecting the links between same person's utterances, the result of Decision Tree clearly shows that the information of DIS is an important element to detect the links. The information of recipients is not useful element because the speakers do not specify themselves as recipients. Therefore, the linguistic elements are especially important, it is necessary improve the estimation accuracy from the linguistic elements. In the matter of detecting the links between others' utterances, the result of Decision Tree clearly shows that the information of DIS greatly influences the links with the information of NUM.From these viewpoints, one may say that our system could find the new links that it was difficult to find from only linguistic information by using meta-information.

We can recognize from Figure10 that an attribute (A-2-b) is top of Decision Tree. Hence, we may say that our assumption that shown in Figure 7 is significant.

A further direction of this study will be to improve the estimation accuracy, the precision, and the recall from the linguistic elements. For example, it is considered that an effective method is to use the information of Rhetorical Relation. In addition to this, it may also be effective method to add other meta-information.

5 CONCLUSIONS

We proposed the method of learning the links between two utterances with Decision Tree using meta-information. Our system could find the new links that it was difficult to find from only linguistic information by using meta-information.

The result of the examination of Decision Tree learning showed the accuracy is 84.1%, the precision is 87.1% and the recall is 66.9% for link between same person's utterances, and the accuracy is 77.5%, the precision is 68.5% and the recall is 56.8% for link between others' utterances. The result of the examination without meta-information showed the accuracy is 62.7%, the precision is 86.7%, the recall is 64.5% for same person's utterances, and the accuracy is 77.9%, the precision is 73.8% and the recall is 49.9% for others'.

REFERENCES

Tokunaga Y., Inui K., and Matsumoto Y., 2006, *Identifying Continuation and Response Relations between Utterances in Computer-Mediated Chat Dialogues*, The Association for Natural Language Processing, Journal of Natural Language Processing, Vol.12, No.1, pp. 79-105.

Ogura K., Ishizaki M., 2003, *The Analysis of Surface Information to Identifying Relevant Utterances in Chat Conversations*, The Japanese Society for Artificial Intelligence, SIG-SLUD-A203-P05, pp.75-80.

J.R. Quinlan, 1993, *C4.5 Programs for Machine Learning*, Morgan Publishers San Mated, California, pp.302.

Andress Stolcke, Elizabeth Shriberg, et al., 1998, *Dialogue Act Modeling for Conversational Speech*, AAAI Spring Symposium, AAAI Press, pp.98-105.

Inui N., Ebe T., Indurkhya B., and Kotani Y., 2001, *A Case-Based Natural Language Dialogue System Using Dialogue Acts*, IEEE International Conference on Systems, Man and Cybernetics, pp.193-198.

Edward Ivanovic, 2005, *Dialogue Act Tagging for Instant Messaging Chat Sessions*, Proceeding of the ACL Student Research Workshop, pp.79-84.

Nakamura J., Inui N., and Kotani Y., 2005, *A method to extract dialogue relations in the polylogue*, IPSJ Programing symposium, Vol. 46, pp.155-163.

Nakamura J., Inui N., and Kotani Y., 2006, *A system of Neural Network detecting link of utterances in the polylogue*, IPSJ SIG Technical Reports, 2006-NL-173, pp.53-58.