Detecting the links between utterances in character based multi-party conversation by machine learning with meta-information

Junpei NAKAMURA, Yasuhiro TAJIMA, Yoshiyuki KOTANI
Department of Electronic and Information Engineering
Tokyo University of Agriculture and Technology
Nakacho 2-24-16, Koganei, Tokyo 184-8588
JAPAN

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 Support Vector Machine (SVM) learning showed the accuracy is 81.3%, the precision is 74.3% and the recall is 77.9% for link between same perosn's utterances, and the accuracy is 80.3%, the precision is 71.1% and the recall is 66.8% for link between others' utterances. The result of the examination without meta-information showed the accuracy is 63.9%, the precision is 50.9%, the recall is 53.3% for same person's utterances, and the accuracy is 79.5%, the precision is 74.7% and the recall is 57.0% for others'. These results showed we could find new links by using meta-information.

Key-Words: - Support Vector Machine, data mining, knowledge engneering, online chat, character base multi-party conversation, interactional structure analysis

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.

Several studies have been made on detecting the links between utterances [1][2][8][9]. In particular, it shows that the links could be learned by machine learning [1]. In these studies, the linguistic information; morpheme, words, and so on, is mainly used as detecting measure. However, it is quite likely that using meta-information is effective method to the character based dialogue data that lack the nonverbal information like an online chat.

In this paper, we propose a method of finding new links for character based multi-party conversation with linguistic information and some meta-data. We had the link learned by Support Vector Machine (SVM) [3]. The result of the examination of SVM showed the accuracy is 81.3%, the precision is 74.3% and the recall is 77.9% for link between same person's utterances, and the accuracy is 80.3%, the precision is 71.1% and the recall is 66.8% for link between others' utterances. The result of the examination without meta-information showed the accuracy is 63.9%, the precision is 50.9%, the recall is 53.3% for same person's utterances, and the accuracy is 79.5%, the precision is 74.7% and the recall is 57.0% for others'. These results showed we could find new links by using meta-information.

2 Define of links between utterances

First, we define the links between utterances, and next, we express the link between utterances.

2.1 Dialogue and utterance

The dialogue *D* is structured the list of conversation participants *M* and the list of utterances *U*, and expressed the following formula (1).

$$D = (M, U) \tag{1}$$

The list *M* is expressed the following formula (2). *M_NUM* is a number of participants, and 2 or more without fail.

$$M = \{m_1, m_2, \cdots, m_{M\_NUM}\} \tag{2}$$

The list U is expressed the following formula (3). In the formula (3), u_t is an utterance at a time t, and a Speaker of u_t is expressed $m(u_t)$.

$$U = \{u_1, u_2, \cdots, u_n\} \quad (3)$$

2.2 Links between utterances

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.

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

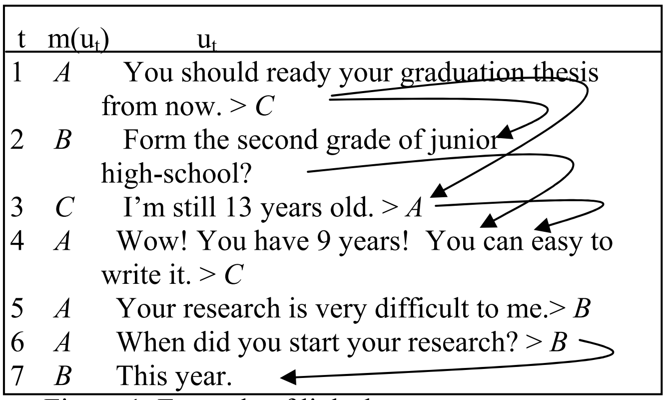

Figure 1. Example of links between utterances.

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 parson's utterances (Figure 2).

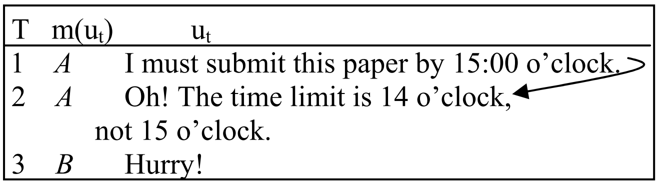

Figure 2. Example of links between same person's utterances.

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 (Figure 3).

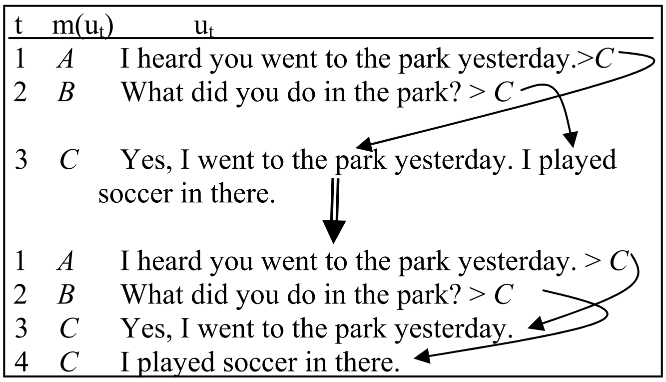

Figure 3. Example of dividing utterance including two or more sentences.

3 Learning of links between utterances

In this paper, we tried to learn the links between utterances by SVM [3] 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, the distance between utterances, and the frequency of individual utterance.

We use 5 elements; consist of 2 linguistic elements and 3 elements of meta-information following Table 1, as attribute for SVM.

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

Kind of information	Elements
(A)Linguistic	(A-1)Relation of sentences
	(A-2)Information of Recipients
(B)Meta	(B-1)Number of conversation participants
	(B-2)Distance between utterances
	(B-3)Frequency of individual utterance

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) [5][6][7] as the relation of sentences. We use 18 kind of DA [6]: 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. For example, when the bi-gram probability of greeting-greeting is high, it is quite likely that an utterance whose DA is "greeting" links to other utterance whose DA is "greeting", too. On the other hand, when the bi-gram probability of greeting-negation is low, it is not quite likely that an utterance whose DA is "greeting" links to other utterance whose DA is "negation" (Figure 4).

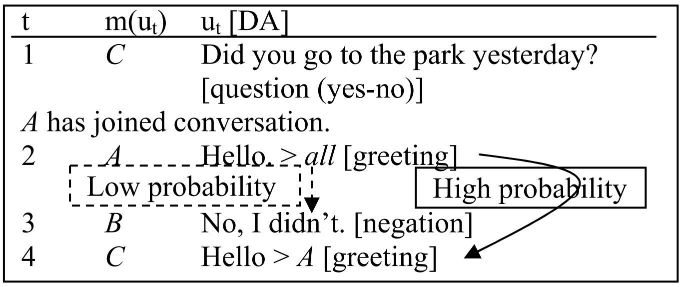

Figure 4. Relation of sentences and bi-gram of DA

However, the values of bi-gram depend on the situation that the speakers of two utterances are the same or others. For example, when the speakers are same person, the bi-gram probability of greeting-greeting is low, and when the speakers are not same person, the bi-gram probability of greeting-greeting is high (Figure 5).

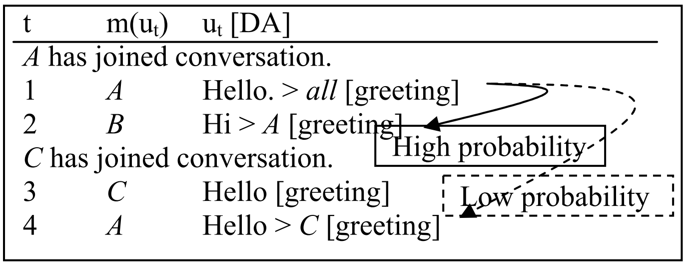

Figure 5. Difference of bi-gram probability by speakers

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

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 $r(u_t)$.

If an utterance includes this expression, we can consider that the utterance was spoken to others. For this reason, the information if $r(u_t)$ is NULL is important. Secondly, 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 (Figure 6).

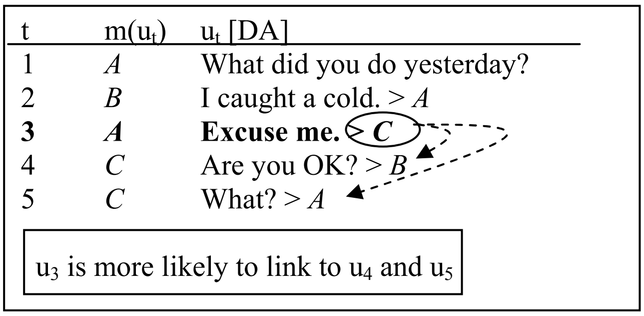

Figure 6. Possibility of existing links from information of recipients at the utterance u_i

In Figure 6, the utterance u_3 is more likely to link to the after utterances whose speaker is participant "*C*", e.g. utterance u_4 and u_5, because recipient of the utterance u_3 is "*C*".

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 (Figure 7).

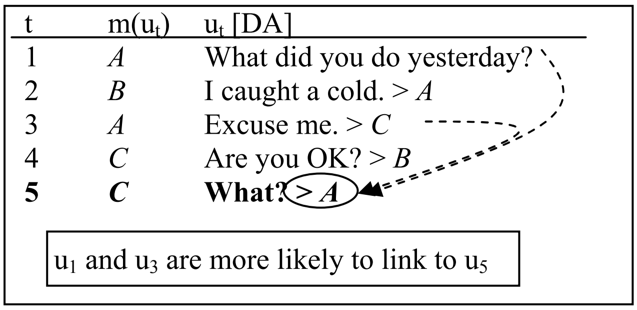

Figure 7. Possibility of existing links from information of recipients at the utterance u_j

In Figure 7, the utterance u_5 is more likely to link from the before utterances whose speaker is participant "*A*", e.g. utterance u_1 and u_3, because recipient of the utterance u_5 is "*A*".

According to these two rules, the link is the most likely to exist between utterance u_3 and u_5 in Figure 6 and 7.

3.2 Meta-information

We suggested 3 elements as meta-information, the number of conversation participants, the distance between utterances, and the frequency of individual utterance. The elements of meta-information should be gained from any dialogue systems. These three elements can be gained from any dialogue system, although some elements like the time of utterance depend on the system.

3.2.1 Number of conversation participants and the distance

We examined the relation between number of participants and distance of utterances where the links exist from the log data of the conversation whose number of participants is 2 to 10. Table 2 shows the rate of utterance to each number of participants and distance between utterances. Figure 8 is a graph of Table 2. The rate of the utterance in which a link does not exist was about 40% by each number of participants. In Table 2, 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

	Distance between utterances where link exists(j − i)													
Number of participants	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

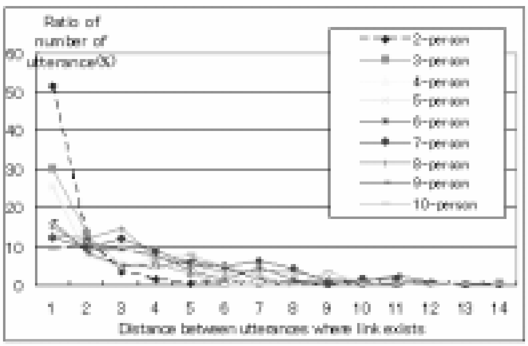

Figure 8. Graph of the rate of utterance to each number of participants and distance between utterances

From these shading cells in Table 2, we realize that the most of cells of maximum distance are located in about three plus each participant. Therefore, we define the maximum distance at time t "*Nt"* expressed by the following formula (4).

$$Nt = M\_NUMt + 3 \quad (4)$$

*M_NUMt is number of participants at time t

From Figure 8, it seems quite probable that the exponential relation exist between number of participants and distance. We consider that the number of participants emphasizes the influence from distance.

3.2.2 Frequency of individual utterance

We assume that the main participant exists in the conversation, and the links are more likely to exist between the main participant's utterances and others'. We use the frequency of individual utterance, because we consider that the main participant has lots of utterances. We define the frequency of individual utterance to following;

"The frequency of individual utterance is the rate of utterances of participant $m(u_j)$ from $u_{i\text{-}Ni}$ to $u_{i\text{-}1}$. 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 number of participants. 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 "NULL" 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 SVM by following method (Figure 9). In Figure 9, the line which has begun from the expression '//' means a comment out. 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
  // End_Of_Data = 533
    repeat
      for j = i + 1 to i + Ni
        if the link exist between u_i and u_j
          then positive sample
        else
          then negative sample
      end for
    until j > End_Of_Data
  end for
end
```

Figure 9. Algorithm of making learning data

3.4 Learning of links by SVM

We used the systems of SVM [3] (TinySVM [4]) in the 6 fold cross-validation with the attribute showed Table 3. The system of SVM learned with the following attribute pattern in order to examine influence to the links by each attribute.

(1) Linguistic + Meta: (A) + (B)
(2) Linguistic: (A)
(3) Linguistic + Number of participants: (A) + (B-1)
(4) Linguistic + Distance: (A) + (B-2)
(5) Linguistic + Frequency of utterances: (A) + (B-3)
(6) Linguistic + Number of participants + Distance: (A) + (B-1) + (B-2)
(7) Linguistic + Number of participants + Frequency of utterances: (A) + (B-1) + (B-3)
(8) Linguistic + Distance + Frequency of utterances: (A) + (B-2) + (B-3)

We illustrated the result of learning in Table 4 and Table 5. Table 4 shows the result of learning links between same person's utterances, and Table 5 shows the result of learning between others' utterances.

Table 3. Attribute for SVM

Kind of information	Elements	Attribute for SVM
(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) ≠ Null ?
		(A-2-b) r(ui) ∈ m(uj) ?
		(A-2-d) r(uj) ≠ Null ?
		(A-2-d) 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			
Attribute	accuracy(%)	precision(%)	recall(%)	F value
linguistic + meta	81.3	74.3	77.9	0.76
only linguistic	63.9	50.9	53.3	0.52
linguistic + number of participants	60.5	47.2	50.0	0.48
linguistic + distance	80.0	72.1	74.6	0.73
linguistic + frequency of utterance	63.6	50.6	54.1	0.52
linguistic + number of participants + distance	80.2	72.9	74.6	0.74
linguistic + number of participants + frequency of utterance	61.7	48.0	49.9	0.49
linguistic + distance + frequency of utterance	83.0	77.3	76.9	0.77

Table 5. Result of learning links between others' utterances

	others			
Attribute	accuracy(%)	precision(%)	recall(%)	F value
linguistic + meta	80.3	71.1	66.8	0.69
only linguistic	79.5	74.7	57.0	0.65
linguistic + number of participants	79.5	74.1	58.2	0.65
linguistic + distance	79.8	71.3	63.8	0.67
linguistic + frequency of utterance	79.2	73.6	57.4	0.64
linguistic + number of participants + distance	78.9	68.7	65.1	0.67
linguistic + number of participants + frequency of utterance	79.4	73.4	58.8	0.65
linguistic + distance + frequency of utterance	80.9	73.1	65.7	0.69

In Table 4 and 5, the accuracy, the precision and the recall are defined the following formulas.

$$accuracy = \frac{pp + nn}{pp + np + pn + nn} \quad (5)$$

$$precision = \frac{pp}{pp + pn} \quad (6)$$

$$recall = \frac{pp}{pp + np} \quad (7)$$

$$F\_value = \frac{2 \times precision \times recall}{precision + recall} \quad (8)$$

*Where pp: true positive, pn: false positive, np: false negative and nn: true negative.

4 Consideration

In the matter of detecting the links between same person's utterances, the result of SVM clearly shows that the information of distance is an important element to detect the links. We understood that the links between same person's utterances did not depend on the number of participants from the result of an attribute "Linguistic + Number of participants" in Table 4. We can read that the information of the frequency of utterances backs up the influence from the information of the distance between utterances.

In the matter of detecting the links between others' utterances, the result of SVM clearly shows that the information of distance greatly influences the links with the information of number of participants or the frequency of utterances.

We can see the recall is rising by using meta-information compared with only linguistic information in Table 4 and Table 5. 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.

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 Conclusion

We proposed the method of learning the links between two utterances with SVM 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 SVM learning showed the accuracy is 81.3%, the precision is 74.3% and the recall is 77.9% for link between same person's utterances, and the accuracy is 80.3%, the precision is 71.1% and the recall is 66.8% for link between others' utterances. The result of the examination without meta-information showed the accuracy is 63.9%, the precision is 50.9%, the recall is 53.3% for same person's utterances, and the accuracy is 79.5%, the precision is 74.7% and the recall is 57.0% for others'.

References:

[1] YASUHIRO TOKUNAGA, KENTARO INUI, and YUJI MATSUMOTO, 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, 2006, pp. 79-105.

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

[3] V.N.Vapnik., The Nature of Statistical Learning Theory, *Springer*, 1995.

[4] Taku Kudoh, TinySVM: Support Vector Machines, http://www.chasen.org/~taku/software/TinySVM/

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

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

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

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

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