Estimation System of People's Freindship and Hierarchical Relationship from Sentence Types

NISHIHARA, Yoko
Graduate School of Engineering Science
Osaka University
1-3, Machikaneyama-cho
Toyonaka, JAPAN
nisihara@gmail.com

SUNAYAMA, Wataru
Faculty of Information Science
Hiroshima City University
3-4-1, Ozuka-Higashi
Asa-Minami-Ku, Hiroshima, JAPAN

YACHIDA, Masahiko
Graduate School of Engineering Science
Osaka University
1-3, Machikaneyama-cho
Toyonaka, JAPAN

Abstract

Many people cooperate for creating something new. It is important for the team's perfomance to build a trusting relationship. It is necessary for building the trusting relationship to know the present state of the team. People are able to understand other people's friendship and hierarchical relationship from their talk. We propose an estimation system of people's friendship and hierarchical relationship. People's uttered sentence types are used to estimate the two relationships. The types are identified from meanings of auxiliary verbs and particles that are included in a sentence. We experimented confirmed that this system is efficient to estimate the two relationships.

1. Introduction

For creating something new, many people must cooperate together. For working smoothly, people must communicate each other. The team's perfomance will decline when the team members don't trust each other[1]. It is important for exercising people's ability to build a good human relationship in the team[2]. People should not make a team from people who are not on friendly terms and who are in a hierarchical relationship. However, it is difficult to understand everyone's friendship and hierarchical relationship.

People understand their friend's friendship and hierarchical relationship from the friends' talks[3]. The utterance intentions are expressed by sentence types. We considered that the friendship and the hierarchical relationship are able to be estimated from the sentence types.

We propose an estimatin system of people's frinedship and hierarchical relationship from sentence types. The system identifies the uttered sentence types with the meaning of auxiliary verb and particle in a sentence. People's friendship and hierarchical relationship are estimated from the sentence types.

2. Related Works

In this section, we will introduce estimation methods of sentence types. We will also introduce human relationship visualization methods.

2.1. Sentence type Estimation

Utterance intentions are expressed between sentences[4]. It is difficult to estimate sentence types automatically only from written texts. Kumamoto et al. developed a voice answering system for a personal computer help system[5]. In this system, the sentence types are estimated from combinations of verb, particles and auxiliary verbs. Although the number of sentence types are not defined, Kumamoto et al. prepared 12 types for personal computer users. In most research, the sentence types are prepared according to each situation[6]. The proposed system estimates the sentece types from the meanings of auxiliary verb and particle. We will also prepared some sentence types for estimation.

2.2. Human Relationship Visualization

FOAF is a language for describing human relationships on the Web[7]. With FOAF, we give our name, affiliation, the URLs of our homepage, chat ID, our friends ' names and so on. The friendship is described " know " or " unknown ". RELATIONSHIP is also a language that is expanded from FOAF[8]. We are able to describe relationship strength[8]. As the relationships are written by each users , it is difficult to describe all relationships with FOAF or RELATIONSHIP.

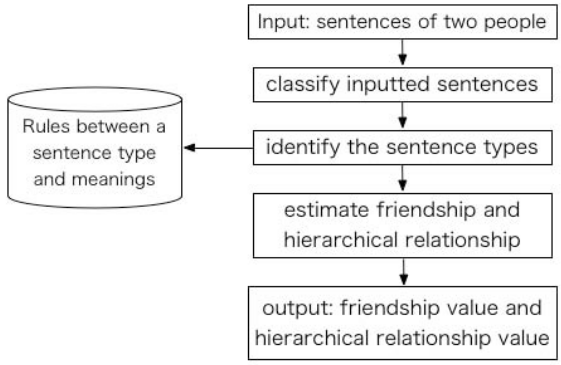

Figure 1. Estimation system of friendship and hierarchical relationship

Tyler developed a method that describes office member's relationships using e-mail exchange information[9]. In this method, relationships are displayed as a graph in which the nodes denote people and the lines denote e-mail exchanges. Important nodes are colored. Matsuo et al. developed a method of human relationship identification by calculating the hit number of a researcher 's name using a search engine and using the hit number to identify pairs who work together[10].

We don't identify important persons or a researcher's relationships. We identify people's friendships and hierarchical relationships.

3. Estimation System of Friendship and Hierarchical Relationship

We explain our estimation system of friendship and hierarchical relationship from uttered sentence types1. The sentences that were uttered by two people are inputted into the system. The inputted sentences are classified whether particles and auxiliary verbs are included or not. The sentence type is indentified from the meanings of auxiliary verb and particle. A sentence type database is used for indentification. The friendship and hierarchical relationship are estimated from the sentece types.

3.1. Input: Utterance Sentences of Two People

Utterance sentences are inputted to the system. The sentences are uttered by two people (in Table 1).

3.2. Classification of Sentences

In this module, the sentences are classified whether particles and auxiliary verbs are included or not. The sentences with auxiliary verbs and particles are divided into two groups according to the speakers.

Table 1. Examples of inputted sentences that were used in the experiment of Section.5

speaker1	I have a meeting with M on July 27.
	If you come to join us, please let me know by e-mail.
speaker2	I will stay in Osaka at the time.
	I will join the meeting if I do not have another works.
	If I have another works, I will contact you.
speaker1	Would you come to the A station at 2 o'clock?
speaker2	I will go.
speaker1	Please come to the A station about 13:20.
	Let's go together.

3.3. Creation of Rules between the Sentence Type and the Meanings

We describe how to create the rules between a sentence type and the meanings of auxiliary verb and particle. The correspondences are shown in Table 2.

To create the rules, particles and auxiliary verbs are first divided into clusters. One cluster consists only of particles and auxiliary verbs whose meanings are similar. The clusters are grouped if the particles and auxiliary verbs in the clusters are used together in the same situation. The grouped cluster become a new cluster. When the grouping is finished, labels are given to each cluster. The sentence type is indicated by a label. In searching for lacked clusters, sentence tags[13] are assigned to proper clusters. New clusters are created from those with unassigned tags and given an appropriate cluster meaning and label.

3.4. Identification of Sentence Types

The types of sentence s in the sets $S1$ and $S2$ are estimated by using the rules. This module outputs a vector. The elements of the vector are the sentence types shown in Table 2. The value of the element is the number of correspondences between an sentence type and particles and auxiliary verbs in the database.

The sentence type vector I_s of the sentence s is described by Eq.(1),

$$I_s = (x_1, x_2, x_3, \ldots, x_{23}) \quad (1)$$

Table 2. Rules between a sentence type and meanings

sentence type	xi	Meanings
Sentimental	x1	will, spontaneous, possible, impossible, hope, emphasis
Fact	x2	negation, confirm, question, addition, side by side, choice, indefinite, recitation, irony, degree, limitation, slight
Supplement	x3	distinction, addition, side by side, indefinite, recitation, emphasis
Value judgement	x4	not naturally, restless mind, reason
Gain	x5	confirm, doubt, question, hope, desire, restless mind, ask oneself
knowledge providi	x6	emphasis, negation, assertion, value judgement
knowledge teach	x7	analogy, guess, estimate, euphemistic, negation, emphasis
Be taught	x8	confirm, doubt, question, desire, request, restless mind, ask oneself
Request	x9	doubt, question, hope, desire, request, restless mind
Confirtm	x10	restless mind, confirm
Request action	x11	hope, desire, request, appeal, prohibition, request agreement, invitation, restless mind, order, permission, cause
Request saying	x12	euphemistic,confirm, doubt, question, will, appeal, request agreement, invitation, order, ask oneself, permission
Propose	x13	distinction, irony, will, hope, emphasis
Thank	x14	respect,spontaneous
Apology	x15	respect, will, spontaneous
Agreement	x16	naturally, denial, affirmation
Opposite	x17	negation, will not to do, not emphasis, criticism, denial, affirmation, opposite
Accept action request	x18	negation, not naturally, confirm, question, will, spontaneous, possibile, impossible, will not to do, desire, request, past
Accept saying request	x19	proportion, hearsay, example, illustration, guess, estimate, euphemistic, negation, not naturally, confirm, question
Satisfaction	x20	euphemistic, confirm, assertion, past
Say reason	x21	reason, value judgement
Hold	x22	restless mind, confirm, recitation
Switch topic	x23	appeal, will, question

The set of particles and auxiliary verbs in the sentence s_i is described by P_{s_i} . The vector value of sentence type x_l is evaluated with Eq.(2),

$$I_{s_j}(x_l) = (\sum_{p_j \in P_{s_i}} freq(x_l, m_{p_j}^k))conf(x_l) \quad (2)$$

where, $freq(x, m)$ denotes the correspondence between x and m in the database. When m is at the end of a sentence, the value of $freq(x, m)$ is 1. When m is within a sentence, the value of $freq(x, m)$ is 0.5. $conf(x)$ denotes the degree of estimation confidence. The noise of the vector element is then removed. The three elements with the highest values are saved. The elements with the highest correlation values to these three elements are also saved. The correlation value between sentence type x_l and x_m is evaluated with Eq.(3),

$$correlation(x_l, x_m) = \frac{T(x_l)T(x_m)}{|T(x_l)||T(x_m)|} \quad (3)$$

In Eq.(3), $T(x)$ denotes the vector whose elements are the utterance tags of SWBD. $|T(x)|$ denotes the vector length.

3.5. Estimation of the Friendship Degree

The friendship degree is estimated with the identified sentence types. We define the friendship degree as **the degree of friendly terms**. If people are on friendly terms, they talk about many things. It is considered that the number of sentence types become big. If people are on friendly terms, they speak alteranately. The estimation function of friendship degree is created by using the two ideas.

Table 3. the number of sentences of mail texts and the number of sentence type estimated

	pair	total	estimated
data1	A B	114	29
data2	A C	109	38
data3	A D	131	41

The sentence types $type(h)$ from speaker h's all uttered senteces is calculated by Eq.(4),

$$type(h) = \sum_{i=1...23} delta(x_i) \quad (4)$$

$delta(x)$ returns 1 or 0. When the sum of $I_{s_j}(x_l)$ is larger than a threshold[1], 1 is returned. When the sum of $I_{s_j}(x_l)$ is smaller than a threshold, 1 is returned.

The friendship value of two people $h1$ and $h2$ is calculated with Eq.(5),

$$friendship(h1, h2) = \frac{type(h1)type(h2)}{line(h1)line(h2)} \quad (5)$$

The $line()$ denotes the number of inputted sentences.

3.6. Estimation of the Hierarchical Relationship Degree

People's hierarchical relationship degree is estimated from identified sentence types. We define the hierarchical relationship degree as **rank difference among the talking people**. If people are in hierarchical relationship, an inferior is not able to disobey his superior. The inferior don't use sentence types " action request" and "speak request" that are identified with the meaning "order" and "prohibition" (in Fig.2). It is expected that the amount of superior's utterance is bigger than that of inferior's. We made the estimation function for hierarchical relationship. People's hierarchical relationship degreee $hierarchy(h1, h2)$ are estimated with Eq.(6),

$$hierarchy(h1, h2) = \frac{(order(h1)+1)}{(order(h2)+1)} \frac{(word(h1)+1)}{(word(h2)+1)} - 1 \quad (6)$$

$h1$ and $h2$ mean speaker. $order()$ means the number of sentences whose sentence types are "action request" and

[1] In this paper, the threshold was 1.0.

"speak request". $word()$ means the number of uttered words. When the value of Eq.(6) is minus, $h1$ is an inferior. When the value of Eq.(6) is plus, $h1$ is an superior.

3.7. Output: the value of Friendship Degree and the Value of Hierarchical Relationship Degree

The friendship value estimated in Section3.5 and the hierarchical relationship value estimated in Section3.6are outputted. Examples are shown in Table 1's 3rd and 4th row.

4. Friendship and Hierarchical Relationship Estimation Experiment

4.1. Experimental Preparation

We used seven data in Table 4 for this experiment. We prepared text-based conversation data and voice-based conversation data. Mail, chat1, chat2, BBS, and voice1 data are the two people's conversation. One fixed speaker and the ohter speaker talked in each conversation pairs. Mixed and voice2 data are also two people's conversation data. Various pairs talked in these two data. Examples of uttered sentences are shown in Table 1.

We prepared three systems for comparing the results. Each system's friendship value is (1) the number of two peoples utter words, (2)the types of auxiliary verbs and particles that are included in two people's utter sentences and (3) the number of nouns that two people utter. Each system's hierarchical relationship value is (4) the rate of Japanese respectable words "desu" and "masu", (5)the rate of words two people utter and (6) the rate of sentences whose sentece type is requested one. The rate in (4), (5), (6) is calculated with Eq.(7).

$$rate(h1, h2) = \frac{num2 + 1}{num1 + 1} \tag{7}$$

$num1$ denotes the $h1$'s number. $num2$ denotes the $h2$'s number.

The correct values of friendship and hierarchical relationship are the results from questionaire survey. The participatns were the university students whose majors are information science. The number of participants were ten. We asked the participants to read the sentence in each convesations and set scores to each pairs. The score is between 1 to 5. If the pairs are on friendly terms, the score is high. If the pairs are in hierarchical relationship, the score is high. We used the average of scores as the correct values.

We calculated the correlation values between the system outputs and the correct values for evaluation.

Table 4. Used data for the proposed system evaluation experiment

data	pairs	average sentences a person
Mail	3	59
BBS	5	53
Chat1	4	42.5
Chat2	9	75.8
Voice1	47	95
Mixed	21	55.4
Voice2	8	97

Table 5. Correlations between the system output value and the correct value in estimation of friendship degree

Data	proposed	words	types	numbers
mail	**0.785**	-0.300	-0.774	-0.677
BBS	**0.587**	-0.459	-0.442	-0.372
chat1	**0.789**	-0.686	-0.623	-0.650
chat2	**0.778**	-0.926	-0.748	-0.733
voice1	**0.448**	0.414	0.213	0.030
mixed	**0.413**	-0.350	-0.514	-0.509
voice2	0.522	**0.530**	-0.332	-0.257
average	**0.646**	-0.253	-0.460	-0.452

4.2. Experimental Results

Table 5 shows the correlation values from estimating friendship degrees. Table 6 shows the correlation values from estimating hierarchical relationship degrees.

The correlations value from the proposed system were highest in six data in Table 5The average of correlation was 0.669 in one-fixed convesation data. As the average of correlation was 0.403 in various pair conversation data, the average was higher in using one-fixed conversation data. It was because nobody speaks alike. The proposed system are able to estimate robastly in using one-fixed conversation data.

The correlations value from the proposed system were highest in five data in Table 6. The average of correlation from the proposed sytem was 0.564 in one-fixed convesation data . As the average of correlation was 0.467 in various pair conversation data, the average was higher in using one-fixed conversation data. It was because nobody speaks alike. The proposed system are able to estimate robastly in using one-fixed conversation data.

Table 6. Correlations between the system output value and the correct value in estimation of hierarchical relationship degree

data	proposed	respect	words	intention
mail	**0.873**	0.754	-0.795	0.567
BBS	0.614	-0.674	**0.714**	0.509
chat1	**0.795**	-0.470	0.653	0.640
chat2	**0.694**	-0.729	0.486	0.650
voice1	**0.370**	-0.494	0.025	0.010
mixed	**0.585**	-0.674	0.514	0.509
voice2	**0.327**	0.221	0.228	0.030
average	**0.710**	-0.229	0.218	0.401

5. Conclusion

We proposed an estimation system for people's friendship and hierarchical relationship from sentence types. The system identifies the sentence types from the meanings of auxiliary verbs and particles. The friendship degree and hierarchical relationship degree are estimated from the sentence types. We confirmed that the proposed system can estimates two relationships robastly in one-person fixed conversation.

References

[1] Hidemasa Asahi and Yukio Ohsawa: Discovery of the Evaluation Measure from Mixed Data 360-degree Feedback in Personnel Evaluation, Journal of Japan Society of Artificial Intelligence, Vol.20, No.3, pp.167–176 (2005)

[2] Akira Ishibashi: Human factors and error prevention, Japan National Institute Public Health, Vol.51, No.4, pp.232–244 (2002)

[3] Kazuyuki Matsumoto, Junko Minato, Fuji Ren and Shingo Kuroiwa: Estimating Human Emotions Using Wording and Sentence Patterns, IEEE International Conference on Information Acquisition, Vol.2005, No.1, pp.421–426 (2005)

[沼 04] Kosuke Numa, Ikki Ohmukai, Masahiro Hamasaki and Hideaki Takeda: Egocentric Search based on RSS, In Poster Proceedings of 3rd International Semantic Web Conference (ISWC2004)

[Neustaedter05] Neustaedter, C., Brush, A., Smith, M., Beyond "From" and "Received": Exploring the Dynamics of Email Triage. In Proceedings of CHI 2005, pp. 1977- 1980.

[4] Takasi MASUOKA: "MODARITY NO BUNPO", Kuroshio Syuppan, 1991.

[5] Tadahiko KUMAMOTO , Akira ITO and Tsuyoshi EBINA: "Recognizing User Communicative Intention in a Dialogue-Based Consultant System — A Statistical Approach Based on the Analysis of Spoken Japanese Sentences —", Systems and Computers in Japan, Vol.25, No.14, pp.91-105, 1994.

[6] Hideki Mima, Kazuaki Ando and Jun-Ichi Aoe: "A Robust Method for Understanding NL Interface Commands in the ICI Intelligent Command Interpreter", Inf. Sci., Vol.85, No.1-3, pp.185-198, 1995.

[7] the foaf project,
(URL) http://www.foaf-project.org/.

[8] RELATIONSHIP: A vocabulary for describing relationships between people,
(URL) http://vocab.org/relationship/.

[9] Josh Tyler, Dennis Wilkinson, and Bernardo A. Huberman: "Email as Spectroscopy: Automated Discovery of Community Structure within Organizations", HP Laboratories,Palo Alto, CA 94304.

[10] Yutaka Matsuo and Yuki Yasuda: "An Analysis of Researcher Network Evolution on the Web" , ANDI'05, 2005.

[11] (URL) http://jiten.www.infoseek.co.jp/

[12] Yuji Matsumoto, Akira Kitauchi, Tatsuo Yamashita, Yoshitaka Hirano, Hiroshi Matsuda, Kazuma Takaoka, Masayuki Asahara: "Japanese Morphological Analysis System ChaSen version 2.2.", Dec, 2000.

[13] Dan Jurafsky, Liz Shriberg, Debra Biasca: "Switchboard SWBD-DAMSL, Shallow-Discourse-Function Annotation; Coders Manual, Draft 13".