



# HOKKAIDO UNIVERSITY

|                  |                                                                                                                                                |
|------------------|------------------------------------------------------------------------------------------------------------------------------------------------|
| Title            | When Your Users Are Not Serious : Using Web-based Associations, Affect and Humor for Generating Appropriate Utterances for Inappropriate Input |
| Author(s)        | Rzepka, Rafal; Higuchi, Shinsuke; Ptaszynski, Michal et al.                                                                                    |
| Citation         | 人工知能学会論文誌, 25(1), 114-121<br><a href="https://doi.org/10.1527/tjsai.25.114">https://doi.org/10.1527/tjsai.25.114</a>                           |
| Issue Date       | 2010-01-06                                                                                                                                     |
| Doc URL          | <a href="https://hdl.handle.net/2115/63795">https://hdl.handle.net/2115/63795</a>                                                              |
| Type             | journal article                                                                                                                                |
| File Information | Whenn Your Users Are Not Serious.pdf                                                                                                           |



## 特集論文

# When Your Users Are Not Serious

Using Web-based Associations, Affect and Humor  
for Generating Appropriate Utterances for Inappropriate Input

Rafal Rzepka

Hokkaido University  
kabura@media.eng.hokudai.ac.jp, <http://sig.media...jp/~kabura/>

Shinsuke Higuchi

(affiliation as previous author)  
shin\_h@media.eng.hokudai.ac.jp, [http://sig.media...jp/~shin\\_h/](http://sig.media...jp/~shin_h/)

Michal Ptaszynski

(affiliation as previous author)  
ptaszynski@media.eng.hokudai.ac.jp, <http://sig.media...jp/~ptaszynski/>

Pawel Dybala

(affiliation as previous author)  
paweldybala@media.eng.hokudai.ac.jp, <http://sig.media...jp/~paweldybala/>

Kenji Araki

(affiliation as previous author)  
araki@media.eng.hokudai.ac.jp, <http://sig.media...jp/~araki/>

**keywords:** non-task oriented dialogue, web-mining, affect analysis, pun generation

## Summary

In this paper we propose a method for generating simple but semantically correct replies to user inputs which are not related to a given task of a task-oriented information kiosk or any other natural language interface placed in a public place. We describe our method for retrieving meaningful associations from the Web and adding modality based on chatlog data. After showing the results of the evaluation experiments, we introduce an implementation of an affect analysis algorithm and pun generator to increase users' satisfaction level.

## 1. Introduction

Together with rapid technological development, more and more natural language processing systems are appearing in streets and buildings, where anyone passing by is able to ask an automatic guide or an information kiosk about route or place details. Considerable progress has been achieved in many subfields concerning task-oriented dialogue systems [Liu 03, Reitter 06]. However, there is no research on solutions to the common problem of replying to inputs which are not related to a task as described in the literature [Gustafson 00, Kopp 05]. We believe that the main reason for this is that an unrestricted domain is disproportionately difficult compared to the possible problems such input could cause. It is very hard to predict the contents and topics of user utterances, and therefore it is almost impossible to prepare conversational scenarios. Furthermore, scenarios usually need specific goals to be useful. However, in our opinion, by combining task-oriented dialogue systems with non-task-oriented ones, we are able to create more human-like architectures which should be more trustful, and give better impressions of a company or organization which sets the automatic informer. Utterances such as "what's the weather gonna

be?", "I'm in love" or "you are ugly" appearing in information kiosks logs should not simply be answered by "I don't understand", "please repeat" or "I have no info on this topic".

A system developer could add a chatbot which would react after discovering whether keywords belong or not to a task. However, such programs have obvious problems. Two well-known examples of non-task-oriented dialogue systems are ELIZA [Weizenbaum 66] and A.L.I.C.E.\*<sup>1</sup>. Both systems and their countless imitators use many rules coded by hand. ELIZA is able to generate a response to any input, but these responses are only information requests which do not provide any new information to the user. In the case of A.L.I.C.E., the knowledge resource is limited to the existing database. Creating such databases is costly and a programmer must learn the AIML mark-up language to build it. Although there have been attempts at updating AIML databases automatically, [De Pietro 05], the scale was rather limited.

As mentioned above, these examples and many other "chatbots" need hand-crafted rules, and are thus often ignored by computer scientists and rarely become a research

\*1 Wallace, R. The Anatomy of A.L.I.C.E.  
<http://www.alicebot.org/anatomy.html>.

topic. However, they have proved to be useful for e-learning [De Pietro 05] and language acquisition [Araki 06] support.

Building a system using automatic methods, like we do, seems to be the most realistic strategy for inputs of unrestricted domains. Considering the large cost of developing a program that can talk about any topic, it is appealing to turn to the huge, cheap textual resource that is the Internet.

At this very moment, millions of people [Kumar 03] are updating their blogs and writing articles on every possible topic. These are available on the Web which we can access at any time, and in a faster and faster manner as search engines grow more and more efficient. Thus, the Web is well suited to extracting word associations from user utterances from conversations with a topic-free dialogue system. We describe a system making use of this, details of which were presented at [Higuchi 09] and demonstrated at [Rzepka 09]. It automatically extracts word association lists using all keywords in a given utterance without choosing a specific one (which most other systems that ignore the context do) then generates a reply using one association from the strongest associations found. Modality is then added to the reply, resulting in the system's output.

Our system is built upon the idea that human utterances consist of a proposition and a modality [Nitta 89]. In this paper we present an algorithm for extracting word associations from the Web and a method for adding modality to statements. We evaluate both the word associations and the use of modality. We also suggest some future possible extensions of the system and show the results of a small experiment with adding humor to the system.

The system described in this paper works for Japanese and uses text as input and output. Although the final goal of our research is to help develop a freely talking car navigation system which, by using chatting abilities, can help to prevent drowsiness while driving, at this stage of development we are concentrating on proposition generation and modality processing. Therefore, at present we work only with text. We plan to combine this project with research on in-car voice recognition and generation.

To the best of the authors' knowledge, their system is the only one that does not use preprocessed data, and generates new phrases without citing existing ones previously created by humans.

## 2. Extracting Word Associations

In this section, we present a method for automatic extraction of word associations based on keywords from user

utterances. We use Google<sup>\*2</sup> search engine snippets to extract word associations in real time without using earlier prepared resources, such as off-line databases.

### 2.1 Extracting Word Associations from the Web

In the first step, the system analyzes user utterances using the morphological analyzer MeCab<sup>\*3</sup> in order to spot query keywords for extracting word associations lists. We define nouns, verbs, adjectives, and unknown words as query keywords. The reason we chose these word classes is that these word classes can be treated as important and, to some extent, describe the context. We define a noun as the longest set of nouns in a compound noun. For example, the compound noun *shizen gengo shori*<sup>\*9</sup> (natural language processing) is treated by MeCab as three words: (*shizen* - natural), (*gengo* - language) and (*shori* - processing). Our system, however, treats it as one noun.

In the next step, the system uses these keywords as query words for the Google search engine. The system extracts the nouns from the search results and sorts them in frequency order. This process is based on the idea that words which co-occur frequently with the input words are of high relevance to them. The number of extracted snippets is 500. This value was set experimentally, taking the processing time and output quality into account. The top ten words of a list are treated as word associations, see Table 1 for an example of a noun group.

**Table 1** Example of word associations extracted for a user utterance

| <i>Sapporo wa samui.</i> (Sapporo (city) is cold.) |                           |    |
|----------------------------------------------------|---------------------------|----|
| Association frequency ranking:                     |                           |    |
| 1                                                  | <i>yuki</i> (snow)        | 52 |
| 2                                                  | <i>fuyu</i> (winter)      | 50 |
| 3                                                  | <i>kion</i> (temperature) | 16 |
| 4                                                  | <i>jiki</i> (season)      | 12 |
| 5                                                  | <i>Tokyo</i> (Tokyo)      | 12 |
| 6                                                  | <i>tenki</i> (weather)    | 11 |
| 7                                                  | <i>chiiki</i> (area)      | 10 |
| 8                                                  | <i>heya</i> (room)        | 10 |

### 2.2 Evaluation

We asked evaluators to use our system and to evaluate the correctness of word lists generated by the system. First, an evaluator freely inputs an utterance, for which the system retrieves ten association words. Next, he or she rated these words using a scale of one to three with

<sup>\*2</sup> Google, <http://www.google.co.jp/>

<sup>\*3</sup> MeCab: Yet Another Part-of-Speech and Morphological Analyzer, <http://mecab.sourceforge.jp/>

<sup>\*9</sup> All Japanese transcriptions will be written in italics.

3 meaning "perfectly correct", 2 - "partially correct" and 1 - "incorrect". In this research we consider words that receive 2 or 3 as usable. Three evaluators repeated the experiment ten times, so the final amount of evaluated words was 300. Table 2 shows the top 10 words, sorted by the frequency of appearance. Table 3 shows the top 5 words.

What constitutes a correct word association was left to each evaluator to decide subjectively, as in a casual conversation setting, associations are hard to define strictly.

**Table 2** Top 10 word associations

| score        | evaluators (A, B, C) | total |
|--------------|----------------------|-------|
| 3            | 40, 52, 57           | 149   |
| 2            | 37, 17, 27           | 81    |
| 1            | 23, 31, 16           | 70    |
| usability[%] | 77, 69, 84           | 77    |

**Table 3** Top 5 word associations

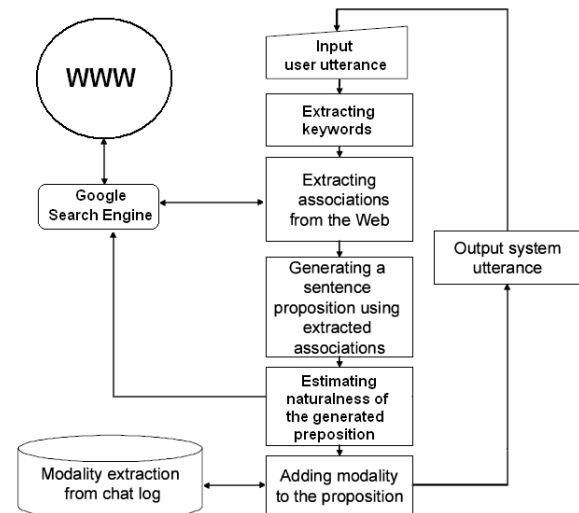
| score        | evaluators (A, B, C) | total |
|--------------|----------------------|-------|
| 3            | 20, 29, 36           | 85    |
| 2            | 17, 9, 10            | 36    |
| 1            | 13, 12, 4            | 29    |
| usability[%] | 74, 76, 92           | 81    |

As shown in Table 2, approximately 77% of the word associations were judged as usable, but there were individual differences between the evaluators. This shows that the definition of word associations is different for each participant. Table 3 shows that approximately 80% of the word associations were judged as usable. It is thus highly likely that the top words from the frequency lists are correct associations. The results show that automatic extracting of word associations using a Web search engine is feasible. The main reason for extracting word associations from the Web is that, due to this method, the system can handle new information, proper names, technical terms, etc. As the system uses snippets from the search engine, the word association extraction takes no more than few seconds.

### 3. General Description of the System

The system generates responses in the following pattern (Figure 1 shows the system flow):

- Extraction of keywords from user utterance
- Extraction of word associations from the Web
- Generation of sentence proposition using word associations
- Addition of modality to the sentence proposition



**Fig. 1** System flow

#### 3.1 Extraction of Keywords from User Utterances

The system applies morphological analysis to the user's utterances in the same way as described in Section 2.1 and extracts keywords based on part of speech. These keywords create association groups by using methods also introduced in the same section.

#### 3.2 Generation of Proposition Using Word Associations

Using the word associations, the system generates the proposition of a sentence to be used as a reply to the user input. A proposition is an expression representing an objective, declarative statement, which does not contain any form of affect the modality usually conveys. The proposition is generated by applying word associations to a proposition template like [(noun) (topic indicating particle *wa*) (adjective)]. We prepared 8 proposition templates manually (see Table 4). The templates were chosen subjectively after examining statistics from IRC<sup>\*4</sup> chat logs, checking their flexibility to be grammatically correct after combining them with different parts of speech. In order to ensure greater diversity of utterances, the proposition templates are applied in a predetermined order exactly the same as the one shown in Table 4. However, since the generated proposition is not always a natural statement, the system uses exact matching searches of the whole phrase in a search engine to check the naturalness of each proposition. If the frequency of occurrence of the proposition is low, it is defined as unnatural and deleted. This processing is based on the idea that the phrases existing on the Web in large numbers are most probably correct grammatically

\*4 Internet Relay Chat Protocol, <http://www.irchelp.org/irchelp/rfc/rfc.html>

and semantically. If an unnatural, low frequency proposition is generated, the system repeats the proposition generation using the same preposition template but with different, randomly chosen top associations. In this experiment the system used propositions for which the hit number exceeded 1,000 hits using Google.

**Table 4** Proposition templates

|                                       |
|---------------------------------------|
| (noun) ( <i>wa</i> ) (adjective)      |
| (noun) ( <i>ga</i> ) (adjective)      |
| (noun) ( <i>ga</i> ) (verb)           |
| (noun) ( <i>wa</i> ) (verb)           |
| ( <i>so-re</i> ) ( <i>wa</i> ) (verb) |
| (noun)                                |
| (adjective)                           |
| (verb)                                |

### 3.3 Adding Modality to the Propositions

Finally, the system adds modality to the generated proposition. By modality we mean a set of grammatical and pragmatic rules to express subjective judgments and attitudes. It is realized through adverbs at the end of a sentence [Nitta 89]. In our system, a pair of sentence-head and sentence-end auxiliary verbs are defined as "modality".

#### §1 Extracting Modality

There is no standard definition of what constitutes modality in Japanese. In this paper, modality of casual conversation is classified into questions and informative expressions. Questions are expressions that request information from the user. Informative expressions are expressions that transmit information to the user. Patterns for these modalities are extracted automatically from IRC chat logs (100,000 utterances) in advance. Modality patterns are extracted in the following ways:

- a pair of a grammatical particle and an auxiliary verb placed at the end of a sentence
- sentences with question marks are defined as questions
- adverbs, emotive words, and connectives at the beginning of a sentence are defined as informative expressions
- candidate patterns thus obtained are sorted by frequency

685 patterns were obtained for informative expressions. 550 of these informative expression patterns were considered by the authors as correct (80%). For questions, 396 patterns were obtained, and 292 patterns (73%) were evaluated as correct. We sorted these candidates in frequency

order. The words appearing at the top of the list were correct, but even the ones appearing only once were still deemed as usable. For example, the question expression "*janakatta deshita-kke?*" is a correct expression, but appeared only once in the 100,000 utterances. Hence, we confirmed that chat logs include various modality expressions, and only a few of them are incorrect. Therefore the system randomly chooses from the whole set of correct modalities and sets them with the flexible proposition patterns we picked up beforehand. See Table 5 and Table 6 for examples.

**Table 5** Examples of informative expression modality

| informative expression          | frequency |
|---------------------------------|-----------|
| <i>maa - kedo</i>               | 21        |
| (Well , it can be said - but -) |           |
| <i>maa - dana</i>               | 16        |
| (Well , it can be said -)       |           |
| <i>maa - desu-ga</i>            | 16        |
| (Well , it appears that -)      |           |
| <i>soko-de - desu-yo</i>        | 15        |
| (Here , it is said that -)      |           |
| <i>maa - da-ga</i>              | 14        |
| (Well , it can be said - but -) |           |
| <i>maa - desu-yo</i>            | 12        |
| (Well , it is that -)           |           |

**Table 6** Examples of question modality

| question               | frequency |
|------------------------|-----------|
| <i>desuka?</i>         | 232       |
| (Is it that - ?)       |           |
| <i>kana?</i>           | 90        |
| (Maybe - ?)            |           |
| <i>da-kke?</i>         | 87        |
| (Is it right that - ?) |           |
| <i>masu-ka?</i>        | 69        |
| (Is it that - ?)       |           |
| <i>nano?</i>           | 68        |
| (Is it that - ?)       |           |
| <i>toka?</i>           | 55        |
| ( - , isn't it ?)      |           |

#### §2 Adding Modality

The system adds the modality from **Section 3.3.1** to the proposition from **Section 3.2** to generate the system output. This process is based on the idea that human utterances consist of proposition and modality. A modality pattern is selected randomly. For example, if the system generates the proposition *fuyu wa samui* (Winter is

cold.) and selects the modality *iyaa ... desu-yo* (Ooh ... isn't it?), the generated output will be *iyaa, fuyu-wa samui desu-yo* (Winter is cold, you know). However, there is a possibility that the system generates unnatural output like *fuyu-wa samui dayo-ne*, depending on the pair of proposition and modality. Regarding this problem, the system uses the Google search engine to filter out unnatural output. The system performs a phrase search on the end of the sentence. If the number of search hits are higher than the threshold, the output is judged as correct. If the number of a search hits is lower than the threshold, the output is judged as incorrect and discarded, and a new reply is generated. Here, we set the threshold to 100 hits.

#### 4. Experimental Results

We used system  $\alpha$ , generating only the proposition, and system  $\beta$ , generating both proposition and modality. 5 participants used each system for conversations of 10 turns and evaluated the conversations on a 5-point scale. Evaluation criteria were "willingness to continue the conversation" (A), "grammatical naturalness of dialogues" (B), "semantical naturalness of dialogues" (C), "vocabulary richness" (D), "knowledge richness" (E), and "humanity of the system" (F)\*5. Table 7 shows average scores for the evaluations of each system. System  $\beta$  that uses modality scored much higher than system  $\alpha$ . Table 8 shows examples of actual dialogue. In the evaluation, the participants expressed the opinion that an utterance like (*xx ha yy*) is unnatural and using a modality like (*maa*), (*moo*) is very natural. Thus we can say that modality expressions make the utterances of the system more natural.

#### 5. System Expandability Examples

The simplicity of our system, real-time processing capabilities and promising results showing that users do not get bored quickly, encouraged us to perform trials with other ongoing projects, and to experiment with the system as a platform for adding various modules and algorithms in order to make the utterances more natural. By using our system, it is possible to perform tests determining whether a new idea will support or improve Human-Computer interaction or not. Here we will briefly describe two such trials - one on guessing emotive values of utterances, and one on improving the system's overall evaluation by adding a pun generator.

\*5 The evaluation sheets were created by setting following polarities: A - I do not want to continue (1) / I wish to continue (5); B,C - unnatural (1) / natural (5); D,E - not rich (1) / rich (5); F - not human-like (1) / human-like.

Table 8 Examples of dialogues with system  $\beta$

|         |                                                                                                                              |
|---------|------------------------------------------------------------------------------------------------------------------------------|
| user:   | <i>Nanika sukina tabemono aru?</i><br>(What food do you like?)                                                               |
| system: | <i>Maa, tabemono-wa oishii desu.</i><br>(Well, food tastes good.)                                                            |
| user:   | <i>Saikin-wa osake-mo sukini nattekitanda.</i><br>(Recently, I begin to like alcohol too.)                                   |
| system: | <i>Demo, sake-wa yowai-no-yo-ne.</i><br>(But, I can't drink much.)                                                           |
| user:   | <i>Yowai-n-da... Sugu yochau-no?</i><br>(Really...?<br>Do you get drunk quickly?)                                            |
| system: | <i>Maa, sore-wa sugoi-wa-ne.</i><br>(Whow, that's amazing.)                                                                  |
| user:   | <i>Boku-wa wari-to osake-wa tsuyoi hou-nan-da.</i><br>(I have a rather high tolerance for alcohol.)                          |
| system: | <i>Iya, sore-wa tsuyoi-n-de.</i><br>(Well, it is strong, I can say.)                                                         |
| user:   | <i>Kyou-wa sukoshi hadazamui-yo-ne.</i><br><i>Natsu-nanoni.</i><br>(It is a little cold today.<br>Even though it is summer.) |
| system: | <i>Asa-wa hadazamui-toka?</i><br>(Isn't it chilly in morning?)                                                               |
| user:   | <i>Asa-kara hadazamukatta-ne.</i><br>(It was chilly since morning.)                                                          |
| system: | <i>Aa kyou-wa hayai-na.</i><br>(Ah, it's early. )                                                                            |

##### 5.1 Testing Affect Analysis

Ptaszynski et al. [Ptaszynski 08] have developed a method for affect analysis of Japanese text called ML-Ask. Their method is based on cross-referencing lexical emotive elements with emotive expressions appearing in text. In the process of analysis, first a general emotive context is determined and then the specific types of emotional states conveyed in an utterance are extracted. Their method achieved human level performance in determining the emotiveness of utterances, and 85% of human level performance in extracting the specific types of emotion was achieved by adding a Web mining technique. They used our baseline and humor-equipped systems to prove that their affect analysis could replace human evaluators. Ability to recognize user emotions is a very important indication of

**Table 7** Evaluation Results

|                     | System $\alpha$ (proposition) |     |     |     |     |     | system $\beta$ (proposition + modality) |     |     |     |     |     |
|---------------------|-------------------------------|-----|-----|-----|-----|-----|-----------------------------------------|-----|-----|-----|-----|-----|
| Evaluation criteria | A                             | B   | C   | D   | E   | F   | A                                       | B   | C   | D   | E   | F   |
| Evaluator a         | 1                             | 3   | 2   | 2   | 4   | 2   | 4                                       | 4   | 3   | 4   | 3   | 5   |
| Evaluator b         | 1                             | 3   | 1   | 2   | 1   | 1   | 4                                       | 4   | 4   | 5   | 4   | 3   |
| Evaluator c         | 1                             | 2   | 1   | 2   | 1   | 1   | 1                                       | 2   | 1   | 2   | 1   | 1   |
| Evaluator d         | 1                             | 3   | 1   | 3   | 1   | 2   | 4                                       | 3   | 1   | 3   | 3   | 4   |
| Evaluator e         | 1                             | 4   | 1   | 1   | 2   | 1   | 3                                       | 2   | 2   | 4   | 5   | 4   |
| Average             | 1.0                           | 3.0 | 1.2 | 2.0 | 1.8 | 1.4 | 3.2                                     | 3.0 | 2.2 | 3.6 | 3.2 | 3.4 |

intelligence, and is becoming a crucial part of our system, needed for modules such as the one introduced in the next subsection.

### 5.2 Improving the System Using Humor

In this trial, an experiment showing that humor can improve a non-task oriented conversational system's overall performance was conducted.

#### §1 Implementing PUNDA system

By using a simplified version of Dybala's PUNDA system [Dybala 08], a pun-generator was added to our baseline system. The PUNDA algorithm consists of two parts: a Candidate Selection Algorithm and a Sentence Integration Engine. The former generates a candidate for a pun by analyzing an input utterance and selecting words or phrases that could be transformed into a pun by one of four generation patterns: homophony, initial mora addition, internal mora addition or final mora addition. The latter part generates a sentence including the candidate extracted in the previous step. To make the system's response more relevant to the user's input, each sentence which includes a joke starts with the pattern [base phrase] to *ieba* ("Speaking of [base phrase]"). The remaining part of the sentence is extracted from the Web, where the candidate is used as a query word and the list of sentences including this word is retrieved. Then the shortest sentence with an exclamation mark is selected, as most jokes convey some emotions. When the candidate list is empty, the system selects one random pun from a pun database.

#### §2 Experiment results

After using one of the systems (baseline or humor-equipped), if an utterance is positive, negative or neutral. During conversation with the humor-equipped talking system, each user's utterance would be analyzed with ML-Ask, then based on the analysis results, the system would decide whether it is appropriate to tell a pun. The decision about

that the system was human-like? G) Do you think the system tried to make the dialogue more funny and interesting? and H) Did you find the system's utterances interesting and funny?\*

<sup>6</sup> Answers were given on a 5-point scale and the results are shown in Table 9.

A third-person evaluation experiment was also performed. Again, the humor-equipped system scored higher than the non-humor one. The question asked in this evaluation was: "Which dialogue do you find most interesting and funny?". Evaluators could choose between 3 options: Dialogue 1 (Baseline system's first 3 turns), Dialogue 2 (Humor-equipped system's first 3 turns, with system's third response replaced by pun generator's output) and Dialogue 3 (the first 3 turns of the baseline system with joking ability). Among 25 evaluators, only 5 (20%) responded that Dialogue 1 was most interesting and funny. 10 chose Dialogue 2 and the other 10 chose Dialogue 3 (40% respectively). This means that both of the humor equipped dialogues received evaluations double that of non-humor dialogue.

### 5.3 Timing Problem - Combining Affect and Humor

In the experiment described above, the system tells jokes (puns) at every third turn of dialogue. In future, timing problems could be solved by replacing this rule with a timing algorithm based on emotiveness analysis of users' utterances. To perform the analysis, Ptaszynski's idea [Ptaszynski 08] mentioned in the "Testing Affect Analysis" subsection could be useful as it detects user's emotional states from the textual layer of speech, by which it can discover

\*6 Polarity answers were the same as in section 4 for A-F; G - did not try at all (1) / tried hard (5); H - not funny at all (1) / very funny (5).

**Table 9** Results of humor experiments

| Evaluation Criteria | A   | B   | C   | D   | E   | F   | G   | H   |
|---------------------|-----|-----|-----|-----|-----|-----|-----|-----|
| Baseline System     | 3.0 | 2.2 | 2.4 | 2.4 | 2.0 | 2.8 | 2.2 | 2.8 |
| With pun generator  | 3.2 | 3.0 | 2.8 | 2.8 | 2.2 | 3.0 | 3.4 | 3.6 |

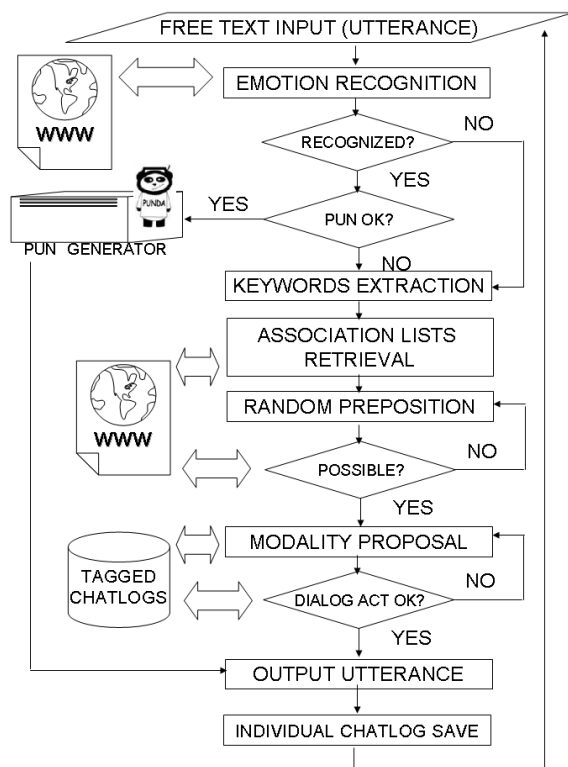
appropriateness of pun-telling would be made based on conclusions drawn from broad humor literature [Martin 96, Newman 96, Cann 99, Danzer 90, Labott 87, Dienstbier 96, Vilaythong 03, Takayanagi 07, Robinson 91, Frankl 60, Hennman 00, Bonanno 97]:

a) If the user's emotive state is negative (stress, depression, anxiety etc.), a pun can be told to help him/her deal with it. For example, if the user says: "You know, I'm feeling kind of down today...", the system, after detecting negative emotion (sadness), could tell a joke to make the user's mood better.

b) If the user's state is neutral, a pun can be told to induce a good mood.

These rules, however, should be limited to situations in which there would be no risk of inducing negative instead of positive reaction.

The flow of such a combined algorithm is shown in Figure 2.

**Fig. 2** Three systems combined into one chatbot.

## 6. Conclusion and Future Work

In this research, we investigated if word associations extracted automatically from the Web are reasonable (i.e., semantically on topic) and if they can be successfully used in non-task-oriented dialogue systems. We also implemented an extraction module which is able to automatically generate in real-time responses to user utterances, by generating a proposition and adding modality retrieved from IRC chat logs. We conducted evaluation experiments on the overall influence of the modality usage and it improved the system. Therefore, we showed that it is possible to construct a dialogue system that automatically generates understandable on-topic utterances without the need to create vast amounts of rules and data beforehand. We also confirmed that our system's performance can be improved by joke generation and affect analysis and we introduced an idea regarding how these two topics could be combined to achieve an even more natural human-computer interface.

There is still a lot of work to be done. It is necessary for a non-task-oriented dialogue system to obtain not only word associations, but also different kinds of knowledge - of the user's preferences or of dialogue itself - for example, conversational strategies. At this moment, the system generates utterances by applying word associations to the proposition templates and adding modality. We also need to consider semantics, speech acts and context more deeply to create a more advanced system. Finally, the system needs to recognize not only keywords, but also the user's modality. We assume that the affect recognition mentioned above will help us to achieve this goal in near future and this will be our next step.

## Acknowledgments

This work was partially supported by the Research Grant from the Nissan Science Foundation. We are especially grateful to Mr. Takashi Sunda from Mobility Laboratory at Nissan Research Center for his valuable comments.

## ◆ References ◆

- [Araki 06] K. Araki, and M. Kuroda (2006) Generality of a Spoken Dialogue System Using SeGA-IL for Different Languages, Proceedings of the IASTED International Conference COMPUTER INTELLIGENCE, pp. 70-75, San Francisco, CA, U.S.A. , 2006.

- [Bonanno 97] Bonanno, G. A. and Keltner, D.: Facial expressions of emotion and the course of conjugal bereavement, *Journal of Abnormal Psychology*, 106(1), pp.126-137, 1997.
- [Cann 99] Cann, A., Holt, K. and Calhoun, L. G.: The roles of humor and sense of humor in response to stressors, *Humor: International Journal of Humor Research*, 12(2), pp. 177-193, 1999.
- [Danzer 90] Danzer, A., Dale, J. A. and Klions, H. L.: Effect of exposure to humorous stimuli on induced depression, *Psychological Reports*, 66(3, Pt 1), pp. 1027-1036, 1990.
- [De Pietro 05] M. De Pietro, M. De Rose, and G. Frontera(2005) Automatic Update of AIML Knowledge Base in E-Learning Environment, In *Proceedings of Computers and Advanced Technology in Education.*, Oranjestad, Aruba, August 29-31, 2005.
- [Dienstbier 96] Dienstbier, R. A., The impact of humor on energy, tension, task choices, and attributions: Exploring hypotheses from toughness theory, *Motivation and Emotion*, 19(4), pp. 255-267, 1995.
- [Dybala 08] P. Dybala, M. Ptaszynski, R. Rzepka and K. Araki (2008) *Proceedings of LIBM'08 First International Workshop on Laughter in Interaction and Body Movement*, pp. 46-51, 2008
- [Frankl 60] Frankl, V. E.: *Man's search for meaning*, New York Washington Press, 1960.
- [Gustafson 00] J. Gustafson and L. Bell (2000). Speech technology on trial: Experiences from the August system. In *Natural Language Engineering*, 1 (1), pp. 1-15.
- [Hennman 00] Hennman, L. D.: Humor as a coping mechanism: Lessons from POWs, *Humor: International Journal of Humor Research*, 14(1), pp. 83-04, 2000.
- [Higuchi 09] Higuchi, S., Rzepka, R. and Araki, K. 2009 A Casual Conversation System Using Modality and Word Associations Retrieved from the Web, *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, pp. 382-390, Honolulu, USA, October 2008
- [Kopp 05] Kopp, S., Gesellensetter, L., Kramer, N., & Wachsmuth, I. (2005). A Conversational Agent as Museum Guide - Design and Evaluation of a Real-World Application. In Panayiotopoulos et al., (Eds.), *Intelligent Virtual Agents*, LNAI 3661, pp. 329-343.
- [Kumar 03] Ravi Kumar, Jasmine Novak, Prabhakar Raghavan, Andrew Tomkins :On the Bursty Evolution of Blogspace, *The Twelfth International World Wide Web Conference* (2003).
- [Labott 87] Labott, S. M. and Martin, R. B.: The stressmoderating effects of weeping and humor, *Journal of Human Stress*, 13(4), pp. 159-164, 1987.
- [Liu 03] B. Liu, L. Du, S. Yu. The method of building expectation model in task-oriented dialogue systems and its realization algorithms. *Natural Language Processing and Knowledge Engineering*, 2003. *Proceedings. 2003 International Conference on*, pp. 174- 179
- [Martin 96] Martin, R. A. and Lefcourt, H. M.: *Humor and life stress. Antidote to adversity*. New York, Springer Verlag, 1996
- [Newman 96] Newman, M. G. and Stone, A. A.: Does humor moderate the effects of experimentally induced stress? *Annals of Behavioral Medicine*, 18(2), pp. 101-109, 1996.
- [Nitta 89] Nitta, Y. and T. Masuoka, "Japanese modality (Nihongo no modariti)" Kuroshio, 1989.
- [Ptaszynski 08] Ptaszynski, M. DYBALA, P. Rzepka, R. and Araki, K. Double Standpoint Evaluation Method for Affect Analysis Systems The 22nd Annual Conference of Japanese Society for Artificial Intelligence (JSAI 2008), CD-ROM Proceedings 2P2-07, 2008
- [Reitter 06] D. Reitter, J. D. Moore and F. Keller. Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. In *Proc. 28th Annual Conference of the Cognitive Science Society (CogSci)*, Vancouver, Canada, 2006.
- [Robinson 91] Robinson, V. M.: *Humor and the Health Professions*. Thorofare, NJ, Slack, Inc., 1991
- [Rzepka 09] Rzepka, R., Dybala, P., Shi, W., Higuchi, S., Ptaszynski, M. and Araki, K. 2009 Serious Processing for Frivolous Purpose - A Chatbot Using Web-mining Supported Affect Analysis and Pun Generation, *Proceedings of IUI'09 - International Conference on Intelligent User Interfaces*, pp 487-488, Sanibel Island, FL, USA. (ACM Press).
- [Takayanagi 07] Takayanagi, K.: *The Laughter Therapy*, Japanese Journal of Complementary and Alternative Medicine, Vol. 4, 2007,

No. 2, pp.51-57

- [Vilaythong 03] Vilaythong, A. P., R. C. Arnau, D. H. Rosen, and Mascard, N. Humor and hope: Can humor increase hope? In: *Humor: International Journal of Humor Research* 16(1), pp. 79.89, 2003.

- [Weizenbaum 66] J. Weizenbaum, ELIZA - computer program for the study of natural language communication between man and machine, *Communication (ACM)*, vol.9, no.1 ,pp. 36-45,1966.

〔担当委員：山本 健一〕

Received May 1, 2009.

## Author's Profile



**Rzepka, Rafal** (Member)

Born in Szczecin, Poland in 1974. Received M.A. from Adam Mickiewicz University in Poznan, Poland in 1999 and Ph.D. from Hokkaido University, Japan in 2004. Currently he is an assistant professor at the Graduate School of Information Science and Technology, Hokkaido University. Research interests: NLP, Web-mining, Commonsense Retrieval, Affect Analysis and Machine Ethics. Member of: AAAI, ACL, IPSJ, IEICE, JCSS and NLP.



**Higuchi, Shinsuke**

Born in Sapporo, Japan in 1983. He is a second year Master Course student at the Graduate School of Information Science and Technology, Hokkaido University, Japan. His research interests include Natural Language Processing, Dialogue Processing, and Web-mining.



**Ptaszynski, Michal** (Member)

Born in Wroclaw, Poland in 1981. He received his M.A. from Adam Mickiewicz University in Poznan, Poland in 2006. He was a research student at Otaru University of Commerce, and since 2007 he has been studying towards his Ph.D. degree at the Graduate School of Information Science and Technology, Hokkaido University, Japan. His research interests include Natural Language Processing, Dialogue Processing, Affect Analysis, Sentiment Analysis, HCI and Information Retrieval. He is a member of SOFT and NLP.



**Dybala, Pawel**

Born in Ostrow Wielkopolski, Poland in 1981. He received his M.A. from Jagiellonian University in Krakow, Poland in 2006. He was a research student at Hokkaido University, and since 2007 he has been studying towards his Ph.D. degree at the Graduate School of Information Science and Technology, Hokkaido University, Japan. His research interests include Natural Language Processing, Dialogue Processing, Humor Processing, HCI and Information Retrieval.



**Araki, Kenji** (Member)

Born in 1959 in Otaru, Japan. Received B.E., M.E. and Ph.D. degrees in electronics engineering from Hokkaido University, Sapporo, Japan in 1982, 1985 and 1988, respectively. In April 1988, he joined Hokkai Gakuen University, Sapporo, Japan. He was a professor of Hokkai Gakuen University. He joined Hokkaido University in 1998 as an associate professor of the Division of Electronics and Information Engineering. He was a professor of the Division of Electronics and Information Engineering of Hokkaido University from 2002. Currently, he is a professor of the Division of Media and Network Technologies of Hokkaido University. Research interests: NLP, Spoken Dialogue Processing, Machine Translation and Language Acquisition. Member of: AAAI, IEEE, IPSJ, IEICE and JCSS.