Prediction of the User’s Reply
Using Emotional Information Retrieved from Internet Resources

Rafal RZEPEKA† Kenji ARAKI† Koji TOCHINAI†
† Graduate School of Engineering, Hokkaido University
† Graduated School of Business Administration, Hokkaido-Gakuen University

1. Introduction
Many researchers tend to think of WWW as an enormous database, which unfortunately is full of useless information that makes the web mining or knowledge-bases creation difficult. However, when we started to rummage through those seemingly useless for AI purposes personal homepages with very similar contents, we imagined the human brain cells might look exactly the same. Not only the stored pieces of semantic information are important but also the number of how many times such similar data was stored. We assumed that Internet is an interesting material for retrieving a common sense, beliefs, opinions and emotional information for various types of agents(1). Without any sophisticated method, our system is able to easily find out that in most cases being cold is not pleasant and cold bear almost always “sounds nice” or that one singer is being hated and the other is being loved. We introduce ideas for our project (GENTA — GENERAL belief reTrieving Agent) and the results of initial experiments with implementing a primitive method for retrieving basic feelings towards human utterances and applying this emotional information in Inductive Learning of the speech acts (2).

2. Main features of GENTA system
Assuming that conversational information could be divided into two kinds — information about the world and discourse information, we concentrated on following two parts of our system — “General Belief” and “Conversation Keeper” which is based on Inductive Learning from chat conversations.

> General Belief
By this expression, we mean a mixture of common sense based on retrieved opinions and simplified environmental knowledge.

> Environment
Because its informal character and “world simplicity” we chose IRC (Internet Relay Chat) for experiments.

> Main topic detection
While detecting speech act GENTA tries to guess the leading keyword(s) from the first user’s utterance since the domain of conversation is still unknown. First, GENTA searches Internet for the whole utterance and its grammatically connected parts previously parsed by link-parser, trying to establish what he can associate with given verbs, nouns, noun phrases or conditional expressions concentrating on feelings-based opinions. For example when input was “Do you like playing soccer when it rains?” GENTA counts how many sentences “I like playing soccer when it rains”, “I love playing soccer”, “I hate playing soccer”, “I love when it rains”, “I hate when it rains” and so on, appear on the Web. This lets our system achieve “own” opinion about playing soccer, when it rains. Then, paraphrasing Shannon’s information theory, we assume that the keyword with less frequency is more interesting for interlocutors and GENTA chooses playing soccer for a leading topic (37,263 vs. 33,368 hits). System believes that it should continue discourse in this “semantic direction”. But before that “Conversation Keeper” must establish which linguistic behavior will be proper for a reply.

> Conversation Keeper
Taking research results of the last decade into consideration, we decided to find such a combination of web-mining and Inductive Learning methods that would help us to create a dialog system which does not require big initial data prepared by hand and does need sophisticated modules. As the first step, before creating dialog manager, NLG module and so on, we decided to confirm that our system is able to learn from above-word level information. We assumed that peculiarity and emotional load of given expressions could support intention recognition which is one of the most important tasks of human discourse management, therefore we divided General Beliefs here into two above word-level values that we call positiveness and usualness which are also measured by counting above mentioned string frequencies upon WWW. We prepared chat logs with about 100 turns and tagged it manually with previously prepared dialog act tags, ([information || opinion || reason || handling || demanding || advising] (warning) (greeting) (nodding). Then GENTA declared usualness and positiveness for those utterances to make them learning data useful for our experiments, as in example utterance: Do you like playing soccer when it rains? becomes a DAPU string
(Dialog+Act+Positiveness+Usualness)
which means that it was a Demand of Opinion consisting of two positive and usual expressions connected by subordinate clause conjunctor (SCC) “cond” (conditional clause). What is characteristic for our method, even if positiveness of expression seems to be doubtful (most Web pages creators like when it rains what does not have to mean that most human beings would state so) it does
not disturb the process since the opinions stays logic. Values of *usualness* and *positiveness* are calculated by comparing frequencies of "I don't like ..." / "I hate ..." and "I like ..." / "I love ...") searching frames. The frequency thresholds are different depending how long the searched string is.

> Inductive Learning

GENTA system has an ability to learn from human spontaneous conversations. We use Inductive Learning method to predict which utterance should be used and to make new rules. The system represents dialog discourse as connected *DAPU strings*.

\[ \text{OH P6 U2 : OD P5 U3 : ND : } (A1 B1 A2 B2 A3 B3) \]

which are divided into double rules

\[(A1 B1) (B1 A2) (A2 B2) (B2 A3) (A3 B3) \]

stored in a *Dictionary*. When new input is done, GENTA parses the utterance to *DAPU string* by recognizing a dialog act *determinator* which are words attributed to every dialog act tag (for example *should* determines *advising* tag). If subordinate clause conjunctor (SCC) is detected, both clauses are parsed into *DAPU strings* and they become an individual element for learning. When there are more than one sentence during one turn, GENTA confirm if they are of the same dialog act. If not, the input is divided - last rule is changed and a new one is created, for example:

*H1: Do you care?*

*H2: Well, I don't care. What about ya?*

*G: Me neither, man!*

creates *A1:B1a (B1b: A2) instead of (A1:B1) (B1: A2).* Learning concentrates on dialog acts tagging and conjunctors, and their coexistence with *positiveness* and *usualness*. For example if not known dialog act determinator appears, our system decides the most probable tag and if a user does not cancel computer's output new rule is created in the dictionary.

3. Experiment and Evaluation

> Method

There are two interlocutors A and B. They converse through IRC channel which is monitored by our system (G). G listens to A's utterances and proposes its own answers (as *DAPU strings*). B's utterances are changed into *DAPU strings* which are compared with G's ones.

A person evaluates naturalness of strings when the system chose different dialog act as there are more possibilities than one.

> Results

Two non-native English speakers took part in our experiment. There was no given topic of conversation and emotions were not prohibited. They made 128 turns and they talked mostly of sports.

By comparing user B and GENTA's *DAPU strings* we understood that:

The systems started to use learned data early as the chat was mostly question - answer style, but finally less than half (37.5%) of dialog acts were chosen the same way by human user. Although 81.25% of those different ones were evaluated as natural by a person.

*Positiveness* (5 grade scale: 1-negative, 2-slightly negative, 3-neutral, 4-slightly positive, 5-positive) of system's output that had the same dialog act tag as human's was in 78.3% the same as human user's.

*Usualness* (6 grade scale: 5-very usual, 4-usual, 3-slightly peculiar, 2-peculiar, 1-very peculiar) of system's output was in only in 20.8% the same as human user's, because all parser errors due to misspells were detected as the most peculiar expression.

4. Conclusion and Future Work

We have described a new approach to WWW statistical information usage in the dialog system while it is able to achieve information that is not obvious to the machine without using logic programming. For the reason that it is the preliminary stage of our project, we could evaluate our idea only indirectly and grammatically built sentences were not outputted but the results are convincing enough to continue walking upon chosen path - the system was guessing interlocutor's intentions properly even if utterances were not complete or its contents were very sophisticated. What seems promising for future tasks, our program based only on automatically retrieved knowledge of common opinion and peculiarity of users utterance.

Thus there is a need of creating mechanisms which allow GENTA learn other things from the biggest and growing rapidly database in the world which is Internet and try to apply those methods to commonly explored areas as for instance qualitative spatial reasoning. First we plan to concentrate on implementing substitution of *imagination* which must be an elastic plan retrieval mechanism supported by commonsense libraries created through search frames as *I always (verb) when it rains*.

References