A Survey on Automatic Expressive Opinion Sentence Generation for Enjoyable Conversational Systems

Mahesh. B. Kate[1], Vijaysinh G. Chavan[2], Mahesh B. Kulkarni[3]
[1]ME- CSE, *SKNSCOE, Korti, Pandharpur, Solapur University, Solapur, India*
[2]Assistant Professor, *SKNSCOE, Korti, Pandharpur, Solapur University, Solapur, India*
[3]Assistant Professor, *SKNSCOE, Solapur, Solapur University, Solapur, MS, India*
[1]vijaysinh.chavan@sknscoe.ac.in, [2]mahesh.kulkarni@sknscoe.ac.in, [3]maheshkate2007@gmail.com

Abstract *— **In terms of functional conversations, Grice's Maxim of Quantity suggests that responses should contain no more information than was explicitly asked for. However, in our daily conversations, more informative response skills are usually employed in order to hold enjoyable conversations with interlocutors. These responses are usually produced as forms of one's additional opinions, which usually contain their original viewpoints as well as novel means of expression, rather than simple and common responses characteristic of the general public. In this paper, we propose automatic expressive opinion sentence generation mechanisms for enjoyable conversational systems. The generated opinions are extracted from a large number of reviews on the web, and ranked in terms of contextual relevance, length of sentences, and amount of information represented by the frequency of adjectives. The sentence generator also has an additional phrasing skill. Three controlled lab experiments were conducted, where subjects were requested to read generated sentences and watch videos filmed about conversations between the robot and a person. The results implied that mechanisms effectively promote users' enjoyment and interests.***

Keywords*— **Conversational robots, Natural sentence generation, Opinion generation, Question answering***

I. INTRODUCTION

In the age of technology and science there is very much importance for communication to each other. There are various ways to perform the task of communication. Conversation is one of them. But to achieve good communication enjoyable conversation system must be there. According to Grice [1] there are four maxims (Quality, Quantity, Relevance & Manner) for enjoyable conversation system. But if I consider Grice's Maxim of Quantity then responses should short but sweet. But in daily routine more response information is used to achieve enjoyable conversation with each other.

These responses are nothing but one's opinion which consists of one's viewpoint as well as novel means of expression. This paper is proposing automatic expressive opinion sentence generation for enjoyable conversation systems. The opinions generated are fetched from a vast number of reviews from the web. These extracted opinions are ranked in terms of contextual relevance, length of sentences & amount of information represented by frequency of adjectives. There is one extra quality for sentence generation called as additional phrasing skill. Actual three experiments were performed to see the conversation of sentence generation between robot & person. This result proves that mechanisms effectively fulfill enjoyment & interests. Let's take one example to illustrate for question answering system.

A: Who is your favorite actor or actress?

B: Yes my favorite actress is Audrey Hepburn. Audrey is charming & beautiful woman.

In the above example there is question answering between two persons namely A & B.A is asking about favorite actor to B.B answered the question not only saying the name of favorite actress but describing some qualities she has which is not expected by person A. Because According to Grice [1] conversation principle such long answer is not expected Because according Grice maxim of Quantity it suggests that conversation should be short but sweet. This principle is not followed here. But in daily routine more information is used for good conversation. Now I will analyze the elements for enjoyable conversation system at both discourse level & sentence level.

In the discourse level only two opinions are considered just like the above example. This is called as "Small talk" skills. This was initially studied by Malinowski[2] who first coined the term "Phatic Communication" to describe a type communication where only mere exchange of words take place which is only used for engagement of communication. Small talk is just used as conversation opener at the end of conversation to fill the space of silence. Schneider [3] did the first extensive study of small talk. He proposed that such conversation consists of number of moves such as topic initialization, agreeable phrasing, informative responding & acknowledgement.

According to Schneider's category person B's action of additionally responding with his/her own opinion in the above example is an informative responding move. In the

sentence level person B's simple opinion "Audrey is beautiful, isn't she?" have less information & fulfill the Grice's Maxim. It is not due to length of sentences but due to general opinion of public. So it cannot attract an interlocutor's interest in an effective manner.

Vice versa person B's actual second opinion "She has a lot of qualities such as she is charming & beautiful woman .having expresses novel opinion about Audrey with wealth of words which contains positive attitude. Further we will take another two examples "This is an erotic movie which also expresses elegance of ballet." "Dola's family & princess sheata with pure mind & are really cute, charming & innocent." These two opinions are about the movie "Black swan" & "Castle in the sky" respectively. Adjectives are shown in bold face. Sentences with less frequently adjectives & proper length are likely to provide a unique viewpoint which is different from that of long length sentences. From this one conclude that the amount of information in each sentence can be described as a level of generality of expressions which is mostly represented by the frequency of adjectives in documents on a certain topic. Some natural language processing techniques are sentiment analysis [4][5] & opinion mining[6] have direct relation to opinion generation. On the basis of the results of these analyses, automatic expressive opinion sentence generation mechanisms are proposed which has capacity of an additional phrasing skill & automatic expressive opinion sentence generation.

The opinions are extracted from a large number of reviews on the web, & ranked in terms of contextual relevance, length of sentences, & amount of information represented by frequency of adjectives. The additional phrasing skill is implemented as a mechanism of doing sentence combination of a simple preceding response & an additional opinion. Our typical scenario is as follows: when system is asked a factoid typed question, it first replies with a sufficient answer based on a structured database & then it adds another expressive opinion in line with the current context which might be informative to the user.

II. LITERATURE SURVEY

1. Paper presents a simple unsupervised learning algorithm for classifying reviews as recommended (thumbs up) or not recommended (thumbs down). The classification of a review is predicted by the average semantic orientation of the phrases in the review that contain adjectives or adverbs. A phrase has a positive semantic orientation when it has good associations (e.g., "subtle nuances") and a negative semantic orientation when it has bad associations (e.g., "very cavalier"). In this paper, the semantic orientation of a phrase is calculated as the mutual information between the given phrase and the word "excellent" minus the mutual information between the given phrase and the word "poor". A review is classified as recommended if the average semantic orientation of its phrases is positive. The algorithm achieves an average accuracy of 74% when evaluated on 410 reviews from opinions, sampled from four different domains (reviews of automobiles, banks, movies, and travel destinations). The accuracy ranges from 84% for automobile reviews to 66% for movie reviews

2. Recommender systems do not always generate good recommendations for users. In order to improve recommender quality, we argue that recommenders need a deeper understanding of users and the information seeking tasks. Human-Recommender Interaction (HRI) provides a framework and a methodology for understanding users, their tasks, and recommender algorithms using a common language. Further, by using an analytic process model, HRI becomes not only descriptive, but also constructive. It can help with the design and structure of a recommender system, and it can act as a bridge between user information seeking tasks and recommender algorithms.

3. In this paper, we present a dependency tree based method for sentiment classification of Japanese and English subjective sentences using conditional random fields with hidden variables. Subjective sentences often contain words which reverse the sentiment polarities of other words. Therefore, interactions between words need to be considered in sentiment classification, which is difficult to be handled with simple bag-of-words approaches, and the syntactic dependency structures of subjective sentences are exploited in our method. In the method, the sentiment polarity of each dependency subtree in a sentence, which is not observable in training data, is represented by a hidden variable. The polarity of the whole sentence is calculated inconsideration of interactions between the hidden variables. Sum- product belief propagation is used for inference. Experimental results of sentiment classification for Japanese and English subjective sentences showed that the method performs better than other methods based on bag-of-features.

4. IBM Research undertook a challenge to build a computer system that could compete at the human champion level in real time on the American TV quiz show, Jeopardy. The extent of the challenge includes fielding a real-time automatic contestant on the show, not merely a laboratory exercise. The Jeopardy Challenge helped us address requirements that led to the design of the DeepQA architecture and the implementation of Watson. After three years of intense research and development by a core team of about 20 researchers, Watson is performing at human expert levels in terms of precision, confidence, and speed at the Jeopardy quiz show. Our results strongly suggest that DeepQA is an effective and extensible architecture that can be used as a foundation for combining, deploying,

evaluating, and advancing a wide range of algorithmic techniques to rapidly advance the field of question answering (QA).

5. In this work we present topic diversification, a novel method designed to balance and diversify personalized recommendation lists in order to reflect the user's complete spectrum of interests. Though being detrimental to average accuracy, we show that our method improves user satisfaction with recommendation lists, in particular for lists generated using the common item-based collaborative filtering algorithm .Our work builds upon prior research on recommender systems, looking at properties of recommendation lists as entities in their own right rather than specifically focusing on the accuracy of individual recommendations.

We introduce the intra-list similarity metric to assess the topical diversity of recommendation lists and the topic diversification approach for decreasing the intra-list similarity. We evaluate our method using book recommendation data, including offline analysis on 361, 349 ratings and an online study involving more than 2, 100 subjects.

III. SYSTEM DESIGN

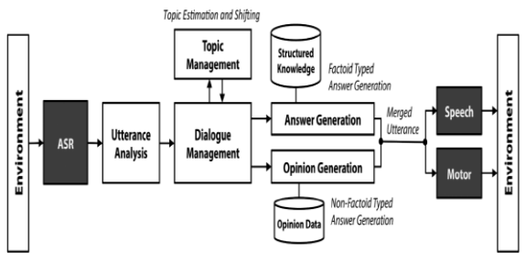

Fig.:- System Architecture

We describe the architecture of our system for informative question answering based on the consideration, and depicted in Figure the main processes in the framework are the natural language understanding (NLU) process, the dialogue management process, and the sentence generation process. The NLU process includes topic estimation and utterance (question) analysis. The sentence generation process is divided into factoid and non-factoid typed answer generation modules. The factoid typed answer generation module refers to structured knowledge databases organized using Semantic Web techniques. The non-factoid typed answer generation module generates the system's own opinions automatically extracted from a large indefinite number of reviews on the Web. The framework also has an utterance combination mechanism that combines factoid and non-factoid typed responses to realize the additional phrasing function.

IV. CONCLUSION

Our proposed automatic expressive opinion sentence generation model is nothing but Question Answering model. It has large capacity to QA recommender system. Our system has extra feature as expressive opinion generation to offer various ideas which is user's requirement for enjoyable conversation system. Ziegler et al. has achieved high user satisfaction by making topic diversified recommendendations & reducing the similarity of recommendation lists. Further Murakami et al assumed that user satisfaction depends on whether a recommender system suggests unexpected items that are relevant to the user's preferences. For application of our framework for more serendipitous recommender systems we will take into consideration learning mechanisms of opinion ranking algorithms & sentence combination which is based on conversational contexts & user models to achieve user's satisfaction.

REFERENCES

[1] P.D.Turnkey, "Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews," in Proc. 40th Annual .Meeting Assoc. Computational. Linguist. 2002, pp. 417–424, Association for Computational Linguistics.

[2] S.M.McNee, J.Riedl, and J AKonstan, "Making recommendations better: An analytic model for human-recommender interaction," in CHI'06 Extended Abstracts on Human Factors in Computing Systems. New York, NY, USA: ACM, 2006, pp. 1103–1108.

[3] T. Nakagawa, K. Inui, and S. Kurohashi, "Dependency tree-based sentiment classification using crfs with hidden variables," in Proc. Human Lang. Technol.: Annu. Conf. North Amer. Chap. Assoc. Comput.Linguist., 2010, pp. 786–794, Association for Computational Linguistics.

[4] D. Ferrucci, E.Brown, J.Chu-Carroll, J.Fan, D.Gondek, A.A.Kalyanpur, A.Lally, J.W. Murdock, E.Nyberg, and J.Prageretal., "Building Watson: An overview of the DeepQA project," AI Mag., vol. 31, no. 3, pp. 59–79, 2010.

[5] Y. Matsuyama, K.Hosoya, H.Taniyama, H.Tsuboi, S.Fujie, and T.Kobayashi, "Schema: Multi-party interaction-oriented humanoid robot," in ACM SIGGRAPH ASIA Art Gallery Emerging Technol.: Adapt. 2009, pp. 82–82, ACM.

[6] T. Nakagawa, T. Kawada, K. Inui, and S. Kurohashi, "Extracting subjective and objective evaluative expressions from the web," in Proc. IEEE 2nd Int.

Symp. Universal Commun. (ISUC'08), 2008, pp.251–258

[7] T.Kobayashi and S.Fujie, "Conversational robots: An approach to conversation protocol issues that utilizes the paralinguistic information available in a robot-human setting," Acoust. Sci. Technol., vol. 34, no. 2, pp. 64–72, 2013

[8] N.Kobayashi, K.Inui, Y.Matsumoto, K.Tateishi, and T.Fukushima, "Collecting evaluative expressions for opinion extraction," in Proc. Nat. Lang. Process. (IJCNLP '04), 2005, pp. 596–605.

[9] M.Higashiyama, K.Inui, and Y.Matsumoto, "Acquiring noun polarity knowledge using selectional preferences," in Proc. 14th Annu.Meeting Assoc. Nat. Lang. Process., 2008, pp. 584–587.

[10] C.-N.Ziegler, S.M.McNee, J.A.Konstan, and G.Lausen, "Improving recommendation lists through topic diversification," in Proc. 14th Int. Conf. World Wide Web, 2005, pp. 22–32, ACM.

[11]T.Murakami, K.Mori, and R.Orihara, "Metrics for evaluating the serendipity of recommendation lists," in New Frontiers in Artificial Intelligence. New York, NY, USA: Springer, 2008, pp. 40–46.