Adding Affective Argumentation to the GenIE Assistant

Nancy L. Green*, Brian Stadler*, and Jennifer Kimbrough**

*Department of Computer Science and **Center for Youth, Family & Community Partnerships
University of North Carolina Greensboro
Greensboro, NC 27402 USA
nlgreen at uncg.edu

Abstract

This paper presents preliminary results of an empirical study to investigate effects of adding affective argumentation to the GenIE Assistant, an implemented proof-of-concept computational model of normative biomedical argument generation. The Assistant has been implemented in the domain of genetic counseling, a domain where human writers are advised to show empathy in addition to presenting clinical arguments for the diagnosis and source of a patient's condition.

Introduction

The GenIE Assistant is an implemented proof-of-concept computational model of normative biomedical argument generation informed by study of a corpus of letters written by genetic counselors to their clients (Green et al. 2011). The goal of the model is to generate *transparent* argumentation, i.e., natural language text in which the structure and functional components of the argumentation are accessible to a letter's audience. Transparency is necessary for a lay audience to comprehend, evaluate or challenge argumentation, or re-evaluate it in light of new findings about the patient or changes in medical science. However, that model did not address the use of statements in the corpus that seem designed to soften the impact of a message or to respond to anticipated objections to negatively valued messages. Such statements are referred to in this paper as *affective* argumentation. Genetic counselors are aware of the potential negative reaction of their client to information in a letter (e.g., low self-esteem, feeling of loss or sadness, blame, anger, fear, shame, worry, and guilt), which may adversely influence the recepient's comprehension and decision-making abilities (Kessler 1979). They are advised to use various writing strategies that address this problem (Baker et al. 2002).

Informed by strategies found in the corpus, the GenIE Assistant's computational model has been extended to produce several types of affective argumentation. This paper reports some preliminary results of an empirical study to investigate effects of adding affective argumentation. For background, extensions to the model and the affective strategies included in the study are described briefly in the next section.

Generating Affective Argumentation

Instances of four strategies were found in the corpus in the context of a claim that the patient inherited some adverse genetic change from one or both parents, the intended recipients of the letter. The strategies seem designed to mitigate guilt over the parents' role in their child's inheritance of a genetic condition. The names used to refer to the strategies in this paper and examples of each are listed below. All four apply to cases of autosomal recessive inheritance, while only the first two apply to cases of autosomal dominant inheritance. (The last two apply to autosomal recessive inheritance since a carrier of such a genetic change may not be affected himself.)

- *Not-uncommon: Many genetic conditions, such as …, are inherited this way*
- *Non-intentional*: *It is important to remember that we have no control over which genes our children inherit.*
- *Universal: It is estimated that everyone has a number of gene changes that can cause problems in our children if our partner is also a carrier*
- *Did-not-know*: *Most people do not know they are a carrier until they have a child with …*

To generate the above statements, it was necessary to add a *user model* (Zukerman 2001) to the GenIE Assistant, representing the system's model of the audience's attitudes (knowledge, beliefs, and values) towards certain facts in the GenIE Assistant's knowledge base. A generic user model has been created using default assumptions (Rich 1979). For example, it is assumed that a parent negatively values being a carrier of the genetic condition that resulted in his child's condition. If the

GenIE Assistant were to be deployed in a real-world healthcare application, the user model could be initialized by clinical staff with knowledge of the attitudes of each letter's intended recipients.

In the original GenIE Assistant (Green et al. 2011), a discourse grammar encoding the genre-specific format of patient letters extracts content from the knowledge base (KB) in non-linguistic form to create a hierarchical discourse structure described in terms of Rhetorical Structure Theory (RST) relations (Mann and Thompson 1988). Parts of the structure represent claims that are passed to an argument generator. The argument generator uses formal presumptive argumentation schemes to extract additional content from the KB to create arguments. The arguments are structured according to an extension of RST, ArgRST (Green 2010), and grafted onto the previous discourse structure. The resulting discourse structure is then transformed into English text using natural language generation techniques.

To produce the affective argumentation described in this paper, the Assistant's generation process has been modified. During generation, the RST structure is revised as follows. Given the user model and a certain node of the tree,

- if any of *Not-uncommon*, *Universal*, *Non-intentional* is/are licensed it/they are inserted into the tree.
- if *Did-not-know* is licensed, a discourse structure rooted by the RST relation of *Antithesis* is inserted into the tree. The inserted *Antithesis* tree represents the opposing viewpoints of the writer and of the audience.

After insertion into the RST structure, the entire RST structure is processed by natural language generation modules for aggregation, sentence planning, lexical choice, and syntactic realization. The process by which affective arguments are generated will be covered in more detail in a future paper.

Empirical Study

Design

Participants were divided randomly into four groups. Each group read and answered questions on two letters generated by the GenIE Assistant. Two versions of two letters, one on a case of cystic fibrosis (CF) and the other on a case of achondroplasia (ACH) were generated. For each case, one version was generated by the original GenIE Assistant without affective argumentation (-AF) and another version by the extended Assistant with affective argumentation (+AF). The +AF version of CF used all four of the strategies listed in the previous section, but the +AF version of ACH used only the two applicable to autosomal dominant inheritance: *Not-uncommon* and *Non-intentional*. The design is summarized in Table 1.

Table 1. Experimental design.

Group (N)	First letter	Second letter
A (19)	CF, +AF	ACH, -AF
B (22)	CF, -AF	ACH, +AF
C (23)	ACH, -AF	CF, +AF
D (19)	ACH, +AF	CF, -AF

Participants

The participants were students age 18 or older recruited from university classes and were offered a chance to enter a drawing for a gift card as compensation. 93% of participants were age 18-34 and 65% were female. Only 2.4% had had no previous college courses, and 65% had at least a four-year college degree. 79% reported no or little familiarity with the genetic conditions covered in the study. At the end of the study, participants were asked to answer questions to measure their mathematical numeracy using the Subjective Numeracy Scale (SNS) instrument (Fagerlin et al. 2007). The mean (median) score was 4.6 (4.75). The purpose of using the SNS was to investigate interactions between numeracy (Ancker and Kaufman 2007) and inclusion of the affective strategies.

Procedure

After reading each letter the participant filled out a questionnaire on that letter without looking back at the letter. After finishing the questionnaire on the second letter, the participant filled out a third questionnaire to collect demographic and SNS data. The first and second questionnaires were designed to evaluate the effect of the affective strategies as follows.

The first question (Q1), *How accurate or true does the diagnosis that … sound to you*, was designed to measure perceived degree of truth of the diagnosis. The response choices were on a five-point Likert scale from *Very true* to *Very untrue.* This question was inspired by a study that found that perceived truth of data among female patients differed across certain graphic formats, e.g., pictorial display vs. bar graph format (Schapira et al. 2006). For our study, it was hypothesized that perceived degree of truth might be higher in the +AF condition since a reader might be more accepting of a message delivered with affective strategies.

The second question (Q2), *Imagine you are a parent receiving this letter. Please rate how it would make you feel*, listed terms from the PANAS inventory describing five positive and five negative emotions (Thompson 2007). Participants rated each term on a seven-point Likert scale from *Not at all* to *Very*. It was hypothesized that the +AF version would be more positively and less negatively rated than the –AF version.

The next several questions (Q3-Q6) were multiple choice questions asking for information stated directly in the letters: the reason the patient had been referred to the clinic (Q3), the patient's test result (Q4), the diagnosis (Q5), and the number of changed copies (none, one, or two) of the gene responsible for the patient's condition (Q6). The last two questions (Q7-Q8) were more difficult questions on arguments for the parents' possession of the genetic change affecting the child. It was hoped that the +AF versions would have no negative effect on comprehension in Q3-Q8.

Preliminary Results and Discussion

Because the analysis is in progress, this section covers preliminary findings only.

According to a paired-samples t-test for Q1 (perceived truthfulness of diagnosis), the mean response for +AF was 3.96 and for –AF was 3.92 (p=.335). It is interesting that the means in both conditions were closer to the response for 'true' (4) than for 'very true' (5).

Table 2 shows the results of the PANAS ratings (Q2). None of the differences were significant (p<.05). Since the participants were asked to imagine that they were parents receiving such a letter, rather than being the actual parents who might receive such a letter, it is perhaps not surprising that the scores were close to neutral (4) for the negative emotions 'upset' and 'afraid' and even lower for 'ashamed' (2.78) or 'hostile' (2.94). The results suggest that this type of data could be meaningful only if collected from actual recipients of genetic counseling letters, which poses practical challenges however. Interestingly, there was no significant difference in positive feelings either (the first five rows).

Table 2. Paired-samples *t*-test results for Q2.

Adjective	+AF	-AF	p
Inspired	2.57	2.4	.168
Attentive	5.46	5.31	.157
Determined	4.3	4.31	.468
Active	4.18	4.1	.287
Alert	5.58	5.5	.278
Ashamed	2.78	2.78	1
Afraid	4.8	4.78	.446
Nervous	5.2	5.15	.371
Upset	4.59	4.68	.292
Hostile	2.94	2.94	1

Table 3 shows the proportion of participants who answered the comprehension questions, Q3 through Q8, correctly. The scores were high in both conditions for Q3-Q6, which may reflect the educational level of the participants. On the more difficult questions, Q7 and Q8, performance was similar in both conditions although lower than for the easier questions. Note that in practice, genetic counselors assume that recipients have no more than an eighth-grade education. Again, these results suggest that it would be informative to evaluate the impact of affect, if not on actual clients, at least on participants with less education. In the near future, the data will undergo further analysis to see if there are differences taking demographic factors such as gender, background, and numeracy (SNS score) into consideration, or differences related to the topic (CF or ACH) or order effects.

Table 3. Paired-samples *t*-test results for Q3-Q8.

Question	+AF	-AF	p
Q3	1	1	1
Q4	.904	.988	.0035
Q5	.988	1	.16
Q6	.916	.940	.265
Q7	.771	.771	1
Q8	.747	.747	1

.

Related Work

Previous work in generation of affective argumentation in dialogue has used affective strategies to enhance persuasiveness (e.g., Grasso et al. 2002; Mazotta and de Rosis 2006). However, it is not the goal of the GenIE Assistant to use affect in this way. The goal is to address negative side effects of the message, without having an adverse effect on transparency.

Conclusions

This paper presented the preliminary results of a study to determine the effect of adding affective strategies to the GenIE Assistant. The results of the analysis so far show that the affective strategies have not made much difference to the readers' affective state. However, since the study used university students rather than real-world recipients of this type of letter, i.e., people who have an emotional stake in the medical case, the results on affective state may not generalize to the target audience of such letters. The preliminary results also show no negative effect on comprehension, as we had hoped. However, the participants were well-educated compared to the public at large and this question merits further study.

Acknowledgments

Project GenIE was supported by the National Science Foundation under CAREER Award No. 0132821.

References

Ancker, J.S. and Kaufman, D. 2007. Rethinking Health Numeracy: A Multidisciplinary Literature Review. *Journal of the American Medical Informatics Association* 14(6):713-721.

Baker DL, Eash T, Schuette JL, Uhlmann WR. 2002. Guidelines for writing letters to patients. *J Genetic Counseling,* 11(5):399-418.

Fagerlin, A., Zikmund-Fisher,B.J., Ubel, P.A., Jankovic, A.,Derry, H.A., and Smith, D.M. 2007. Measuring numeracy without a math test: Development of the Subjective Numeracy Scale (SNS). *Medical Decision Making* 27: 672-680.

Grasso, F., Cawsey, A., and Jones, R. 2002. Dialectical argumentation to solve conflicts in advice giving: a case study in the promotion of healthy nutrition. *International Journal of Human-Computer Studies* 53: 1077-1115.

Green, N. 2010. Representation of argumentation in text with Rhetorical Structure Theory. *Argumentation* 24(2) (2010), 181-196.

Green, N., Dwight,R., Navoraphan, K., and Stadler, B. 2011. Natural language generation of transparent arguments for lay audiences. *Argument and Computation* 2(1), 23-50.

Kessler, S. (ed.) 1979. *Genetic Counseling: Psychological Dimensions*. Academic Press, NY.

Mann, W. C. and S. A. Thompson: 1988, 'Rhetorical Structure Theory: Toward a Functional Theory of Text Organization', *Text* 8, pp. 243-281

Mazotta, I. et al. Mazzotta, I. and de Rosis, F. 2006. Artifices for persuading to improve eating habits. In *Argumentation for Consumers of Healthcare: Papers from the AAAI Spring Symposium.* AAAI Press, Menlo Park, CA, 76-85.

Rich, E. 1979. User modeling via stereotypes. *Cognitive Science* 3(4), October-December 1979, Pages 329-354.

Schapira, M.M., Nattinger, A.B., and McAuliffe, T.L. 2006. The influence of graphic format on breast cancer risk communication. Journal *of Health Communication,* 11:569-582.

Thompson, E.R. 2007. Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology* March 2007, 38: 227-242.

Zukerman, I. 2001. Natural language processing and user modeling: Synergies and limitations. *User Modeling and User-Adapted Interaction.*