Modeling Affective Reactions for Training Adaptive Interviewing

Geoffrey A. Frank
Robert C. Hubal
RTI International
3040 Cornwallis Road, Research Triangle Park, NC 27709
919/541-{6629,6045}
{gaf,rhubal}@rti.org

Keywords:
Interview training, synthetic characters, affect, natural language

1. Interview Simulation Training

Modeling affective responses of an interviewee is key in providing realistic adaptive interview skills training. In real time, students must plan a sequence of questions or responses to elicit sufficient needed information, and exercise pragmatism in linguistic skills in formulating these questions or responses. The student must also pick up on visual and aural cues relating to the interviewee's affective state; an uncooperative (i.e., disaffected) interviewee will typically prevent the student from gaining the information required to assure an appropriate decision.

Traditionally, training interview skills has involved role-playing scenarios, a labor intensive approach typically providing limited practice for the student (Hubal, et al., 2000). Further, students may feel constrained by the presence of role-players and limit exploration of different solution strategies. An alternative uses simulations. For the training applications described here, an interview consists of verbal interactions between the student and a synthetic character that results in a decision on a course of action. The simulations leverage gaming technology for visualization and animation, but include psychological models developed from clinical research (Deterding, et al., 2003; Hubal, et al., 2004) and dialog derived from interviews with subject matter experts.

A language processor and a behavior engine employ these dialogs and models (Hubal, et al., 2003). Students engage in unscripted (but bounded) conversations with synthetic characters. The students' words are analyzed along four dimensions to determine: (1) content (which interview topic is being addressed); (2) form (command, query, request, informative statement, statement of appreciation, statement of understanding, threat, insult and/or use of profanity); (3) personalization (e.g., using *I* instead of *we*); and (4) politeness (e.g., finding *please* and *thank you*). The synthetic character responds with gesture, expression, and change in emotional state, driven by the verbal inputs of the student and by the nature of the initial conditions that define the character (e.g., personality or mental illness). The emotional state consists of graded levels across a range of emotions including anger, fear, depression, and confusion. If the student threatens, commands, or insults, then the character's fear and/or anger would increase, and the interaction would likely end on a failed note. In contrast, if the student speaks politely and personally, then the character's emotional state would attenuate and the character would become more cooperative.

2. Applications

Three applications highlight how an interviewee's affect affects the interview flow. The first is a field interview (survey) trainer. An application (Figure 1) was developed to simulate the first 30-60 seconds of a field interview. Interviewers began with a scripted introduction, then had to respond to a series of objections or questions raised by the character, a potential respondent. Ultimately, the character ended the conversation either by granting the interview or closing the door. A number of different initial emotional states for the character caused conversations to vary between scenarios. The second is a pediatric patient trainer (Figure 2), where a young girl presenting with otitis media might act whiny and initially uncooperative, and the student could provide the character with distractors (e.g., a teddy bear to calm the girl) or use medical devices to collect information, as well as verbally interview. Third is a law enforcement trainer for officers to learn to manage encounters with a schizophrenic consumer (Figure 3). Scenarios included a schizophrenic hearing voices, a paranoid afraid of a conspiracy of the police with federal agencies, and a normal individual angry because he was almost run over.

Fig 1. Survey Interviewer

Fig 2. Pediatric Patient

Fig 3. Schizophrenic

3. Modeling Challenges

A development challenge was the "curse of dimensionality", where the total range of discourse both for analyzing student inputs and for generating realistic character responses grows quickly. Simulations, fortunately, can deal with multidimensional response spaces much more efficiently than other forms of training, such as branching video, without the cost or storage space required for a huge number of minor variations in responses, and therefore scale well for complex interviews. So that the behavior engine could guide character behavior appropriately to any response in the four-dimensional input space, techniques were devised to analyze input at lexical, syntactic, and semantic levels. Thus a polite query on a particular topic might increase fear but not anger, while an impolite version of the same query might increase both. The situation might be stabilized by conversation to determine problems and solutions, or destabilized with inflammatory language or challenging statements. Tools were created for generating simulation scripts from a multidimensional database of possible inputs, desired emotional updates, and appropriate responses. Similarly, for adaptive response construction, to provide realistic audio cues on the emotional state of the characters, large (but manageable) corpora of character responses were recorded in different tones of voice for appropriate genders and ages. Thus, a little girl was recorded for the same sentence in happy, sad, and whiny voices.

To assess skills acquisition, three types of student actions are sought. First are correct decisions during the simulation: Did the student's actions cause the character to behave appropriately, given the context, and was the outcome a success? Second is acquisition of relevant information: Did the student determine all that was needed before making his/her decisions. Third is use of appropriate politeness, personalization, or other important lexical and syntactic forms. Incorrect student actions are analyzed for root cause: sins of commission (extraneous and erroneous actions), sins of omission (missing steps or actions), and serious violations (e.g., of safety or timing).

4. References

Hubal, R.C., Deterding, R.R., Frank, G.A., Schwetzke, H.F., & Kizakevich, P.N. (2003). Lessons learned in modeling pediatric patients. Studies in Health Technology and Informatics, 94, 127-130.

Hubal, R., Frank, G., Guinn, C., & Dupont, R. (2004). Integrating a crisis stages model into a simulation for training law enforcement officers to manage encounters with the mentally ill. Proceedings of the Workshop on Architectures for Modeling Emotion: Cross-Disciplinary Foundations, AAAI Spring Symposium Series (pp. 68-69). New York, NY: ACM Press.

Hubal, R.C., Frank, G.A., & Guinn, C.I. (2003). Lessons learned in modeling schizophrenic and depressed responsive virtual humans for training. Proceedings of the Intelligent User Interface Conference (pp. 85-92). New York, NY: ACM Press.

Hubal, R.C., Kizakevich, P.N., Guinn, C.I., Merino, K.D., & West, S.L. (2000). The virtual standardized patient: Simulated patient-practitioner dialogue for patient interview training. Studies in Health Technology and Informatics, 70, 133-138.

Author Biographies

GEOFF FRANK is Principal Scientist in RTI's Center for Distributed Learning. He is a member of the IEEE Learning Technologies Standards Committee and the ISO Subcommittee on Information Technology for Learning, Education, and Training.

ROB HUBAL is Senior Research Psychologist in RTI's Center for Distributed Learning. He has led synthetic character projects involving field and telephone interviewing, de-escalating during emotional conversations, and establishing trust to elicit information from medical patients.