The "Etiquette Quotient":
An Approach to Believable Social Interaction Behaviors

Dr. Christopher Miller
Mr. Marc Chapman
Ms. Peggy Wu
Smart Information Flow Technologies
311 First St. N. Ste. 300
Minneapolis, MN 55401
(612) 339-7438, (612) 339-7437
{cmiller, mchapman, pwu}@sift.info

Dr. Lewis Johnson
CARTE Labs USC/ISI
Address: 4676 Admiralty Way
Marina del Rey, CA 90292
(310) 448-8210
johnson@isi.edu

ABSTRACT: *Making avatars react appropriately in social interaction—-to take offense when reasonable, to give deference where appropriate, etc.-- is a more fundamental need for believability and cost-effectiveness than is accuracy in appearance, especially for military applications such as cross-cultural training. We are using a rich, universal theory of human-human "politeness" behaviors and the culture-specific interpretive frameworks for them (labeled "etiquette") from sociology, linguistics and anthropology to create a computational model of social behavior expectations. This model links observable and inferred aspects of power and familiarity relationships, the degree of imposition of an act (all of which have implications for roles and intents) and the actor's character to produce politeness behaviors expectations. By using observations of politeness behaviors (or its lack), the same model permits inferences and updates about those attributes. We are refining and implementing this model to provide a computational believability metric based on the delta between observed and expected politeness behaviors—-an "Etiquette Quotient" (EQ)—-of an actor in context. We see applicability of this model t interactive avatar behavior generation and adaptation through modular, cross-cultural etiquette libraries.*

1. Introduction

Cost effective training for the increasingly complex and unpredictable roles we ask our soldiers to play is rapidly becoming unattainable. This problem is made far worse as the diversity of roles and locales with which soldiers must cope increases. Beyond technical competence with a range of weapons and tactics, we are increasingly asking virtually all soldiers to be cultural ambassadors, conflict arbitrators, detectives and social workers. Worse still, these roles must be practiced across a largely unpredictable range of cultures, ethnicities, religions and social and political groupings, not to mention languages, throughout the soldier's career. This is, of course, an extraordinarily heavy burden to be laid upon the training resources of the military—not to mention the time available to the individual soldier.

A potential way out of this problem is provided by the use of increasingly sophisticated computerized gaming techniques in the training process—allowing humans to train with realistically-behaving computerized simulations not just of the contexts and equipment with which they will be required to work, but more importantly, with the types of people they will be required to work with. Users enjoy a well crafted game and willingly spend tens of hours working through it, acquiring new skills as needed to win. Massively multi-player games elicit even more devotion, and entire user communities have grown up around many games with individuals devoting months of their own leisure time to not only improving their own skills, but to assisting others to do the same through the creation of walk-throughs, FAQs, hints and cheats lists, and player support and conversation groups.

To achieve this goal, however, requires that the avatars used in training be "believable". That is, they must behave in important ways, exactly as the humans they are simulating would behave. But a simulation, by definition, does not behave in all ways like the thing it is simulating. Thus, it is critical to find those avatar behaviors that are significant for training purposes and ensure that they are performed believably.

Computerized avatars don't currently behave with the richness and fluency of behavior that we expect of them and are therefore, unbelievable in key ways. This is certainly true of their physical behavior and appearances. Skin textures, many body movements and facial expressions, etc. cannot be rendered as completely "real" even by all the might of Hollywood's computer graphics houses. But these are not the only factors that contribute the realism of a NPC, nor are they the most crucial behaviors for providing the type of training necessary for the DoD to maintain "Training Superiority". The recent Iraq and Afghanistan experiences demonstrate clearly that in order to bring peace to other nations, our battle elements need to be specifically trained for language skills and multicultural understandings. Future conflicts involve dealing with foreign local forces, civilians, and with their culture, politics, and institutions (Chatham & Braddock, 2003). Moore's law will allow the low-end video games of the future to render realistic skin and hair follicles, but it is the social interaction between a player and the avatar that adds a new and much more important dimension to the training experience. Achieving a cultural understanding involves mastery of the foreign language, but also a comprehension of different social interac-tions. The DARWARS project is currently striving to achieve such training-- for example, in one DARWARS scenario, an Army forward observer interacts with an avatar of a local ally while practicing his tactical Pashto language skills (Chatham, 2003).

Failures in achieving believable behavior are arguably much more significant with regards to the simple "moves" in social interactions—what we refer to broadly as social interaction etiquette. It seems reasonable that an avatar that exhibits a perfectly natural appearance and movements, and yet which fails to behave socially as a "normal" human being, will provoke more disconnect and be less effective for training purposes than one whose appearance and movement appear somewhat abnormal and yet which exhibits social behaviors which are natu-ral and "human". We find it quite possible to interact with, trust and even be taught by humans whose appear-ance and movements are abnormal (perhaps through birth defects or injury). We even rapidly evolve methods for interacting with humans who have no "appearance" at all—for example, when interacting over radio or tele-phone communication, or in print. But we find it more difficult to interact with machines who fail to behave in accordance with our social rules for such matters as who should speak when, what sorts of information should be provided and which should be reserved until requested, who gets to dictate tasks to be performed, etc.

Moreover, it is becoming apparent that training for cultural awareness and appropriate social interaction is im-portant in assisting soldiers to work with local authorities and civilians. The Nobel Peace Laureate Forum has designated Religious and Cultural Conflicts as one of its five major issues. The DoD is well aware of problems that arise from cultural misunderstandings, as well as the value of educating troops in language and social inter-action prior to sending them abroad. One specific focus is cross-cultural training (CCT). An avatar who displays social characteristics consistent with its cultural background can provide CCT in an appropriate and cost-effective manner. Case in point: the Peace Operations Training Center hosted more than 200 soldiers from Fort Hood for a course on Arabic culture in early November of 2003 to prepare them for deployment to Jordan. The current state of cultural training involves foreign instructors covering everything from basic language to dealing with Arabian women during checkpoint inspections (Mares, 2003). While this is an excellent way to introduce the culture, it is resource-intensive, only available to a limited number of soldiers, and at the end of the day, there was little interaction between a trainee and a Jordanian civilian. Given the limitations in human resources required to provide such training, a computer-based avatar may be the only viable solution.

Accurately simulating cultural differences and social interactions relies on more than the accuracy of an avatar's movements, facial expressions, surface rendering, etc. While believability in these aspects will be important in some applications (e.g., close combat, medical diagnosis, etc.), cross-culture training can fit naturally with "socially-aware" avatars. Such avatars would take offense believably if not addressed in a culturally appropriate fashion, might appear recalcitrant or ignorant when they are merely trying to follow their culturally-derived notions of polite turn taking in discourse, etc. Additionally, a socially-aware avatar facilitates self-paced training, as well as auto-feedback of the training material. Both can lead to a reduction of instructional time, accelerated improvement in training materials, and substantial savings. The Defense Science Board estimates that a DoD-wide reduction in learning time in residential schools can easily amount to a savings of over a billion dollars per year (Chatham & Braddock, 2001). Individual human tutors either observe the trainee for signs of confusion and frustration, and offer appropriate assistance, or see signs of confidence and allow the

trainee to advance without interrupting. An auto-tutor with the same observant "nature" is more likely to offer help when the trainee needs it, and can keep track of where most trainees have problems with the training program.

Relevant social interaction behaviors, even those for different cultures and contexts, can frequently be emulated in hand-written scripts and simple, locally-relevant rules. But such approaches are time- and labor-intensive in their own right and brittle--only limited interaction complexity can be supported if every move has to be hand-scripted in advance. A general theory and model of social interactions would greatly enhance the usability and sophistication of avatars, while improving the speed and/or reducing the cost of their construction.

Therefore, our focus is on developing general models and methods of achieving and assessing believable social interactions between individuals and small groups. We are leveraging existing theoretical work by transferring sound socio-anthropological research on human-human social interactions as to what constitutes "appropriate" interaction etiquette to develop a computational model to adapt and/or score the interaction behavior of a computer-based avatar in a given role and with a given action intent, as described below.

2. Why Focus on "Etiquette" for Social Interactions?

The terms "etiquette" and "politeness" are likely to evoke notions of formal courtesies and which dinner fork to use—considerations of limited use in military applications. But politeness is a technical term and a well-studied phenomenon in anthropology, sociology and linguistics having to do with the processes by which we determine and manage the "threat" inherent in communication and interaction between two intentional agents in a social interaction—that is, agents that are presumed to have goals and the potential to take offense at having those goals thwarted in any interaction where those intentional attributes are relevant (cf. Dennet, 1989; Goffman, 1967). As we see below, politeness in this sense is the method by which we signal, interpret, maintain and alter power relationships, familiarity relationships and interpretations of the degree of imposition of an act.

We use the term etiquette to refer to the set of expectations about observable behaviors that allow interpretations to be made, in a cultural context, about those who do or do not exhibit them. Observable behaviors are inter-preted against a framework of etiquette expectations to allow conclusions about the

politeness of those we inter-act with, while simultaneously, we choose behaviors (consciously or unconsciously) on the basis of the same etiquette framework--which dictates how they will be interpreted by those who observe them. As such, the formal and prescriptive etiquettes formulated by Miss Manners and Emily Post are a particularly stilted viewpoint, but hardly the only one; more common are the unwritten (and descriptive) etiquettes we encounter, ma-nipulate and react to as we move through our lives—the etiquettes of the classroom, the locker room, the mar-ketplace, etc. Etiquette refers to the expected "moves" in context that allow participants to make inferences about group membership, power relationships, formality/informality, degree of friendship, importance of infor-mation conveyed, etc. Violation of etiquette can convey lack of regard, lack of acceptance of the proposed rela-tionships, or can convey overriding concerns such as a critical threat.

Etiquette enables the interpretation of observable behaviors—and thus it makes use of a wide range of verbal, physical, gestural and even more primitive modes of interaction. For example, deference can be expressed by posture, by quiet speech and/or by explicit markers such as titles and honorifics. The key is the set of cultural expectations which allow interactants to interpret the behavior, or lack of behavior, in a predetermined fashion. In this sense, there is a "cultural etiquette" associated with, say, infantry soldiers as opposed to clerical workers, just as there is a one for marketplace negotiations in the Middle East vs. an American shopping center.

As such, therefore, politeness and etiquette are very much at the forefront of determining the believability and effectiveness of avatars engaged interactions with other social actors in training applications in militarily rele-vant domains. Believable behavior is behavior that is understandable (i.e., the viewer can infer intent behind the behavior) and broadly consistent with the viewer's expectations. Understandability and expectations, in turn, depend upon the social and cultural context of the behavior. Etiquette provides a way of modeling interactions and moves within a social and cultural context, and of predicting their impact on observers' interpretations about the motives, understanding, knowledge and relationships of those who exhibit them. As we will develop below, believability in social interactions means behaving in accordance with expectations for an actor who knows the social conventions and has a personal stake (personal goals to be thwarted) in the outcomes. Therefore, we focus in this project on etiquette and its role in achieving believability. If avatars do not behave in accordance with etiquette-based expectations, one of two outcomes may result: either (1) they will not be perceived as believable, or (2)

they will be misinterpreted—the trainee will draw false inferences about their relationships, intentions, etc. In either event, they will be useless for training purposes—and worse yet, they may produce inaccurate expectations in students who interact with them.

3. A Model of Human-Human Etiquette for Politeness

A seminal body of work in the sociological and linguistic study of politeness is the cross-cultural studies and resulting model developed by Brown and Levinson (1978; 1987). Brown and Levinson were interested in cataloging and accounting regular deviations, across languages and cultures, from Grice's (1975) conversational maxims. Grice had formulated four "rules" or maxims that characterized efficient conversation. These were:

1. *Maxim of Quality:* Speak truthfully and sincerely
2. *Maxim of Quantity:* Be concise; say neither more nor less than required to convey your message
3. *Maxim of Relevance:* Don't introduce topics at random, follow the conversational "flow"
4. *Maxim of Manner:* Be clear in your statements, avoid ambiguity and obfustication.

By and large, these maxims do a reasonable job of accounting for many otherwise unusual interpretations that hearers make in social interactions. For example, if we've been talking about bass fishing and you suddenly tell me that your brother lives in Manitoba, the Maxim of Relevance explains why I am likely to think that that is somehow connected to the topic of bass fishing.

Brown and Levinson noted, however, that there is at least one way in which people across cultures and languages very regularly depart from the efficient conversation characterized by Grice's Maxims. A simple example in English will illustrate the point: as we settle down to a meal together and I ask you "Please pass the salt." The use of "please" in that sentence is unnecessary for a truthful, relevant or clear expression of my wish and it in fact explicitly violates the Maxim of Quantity since it adds verbiage not required to express my propositional intent (to have the salt passed to me).

Over years of cross linguistic and cross cultural studies, Brown and Levinson collected and catalogued a huge database of such violations of efficient conversation. Their explanation for many of these violations is embodied in their model of politeness, which will be explained next.

3.1 Face threats in social interactions

The Brown and Levinson model assumes that social actors are motivated by a set of wants including two important social wants based on the concept of face (Goffman, 1967) or, loosely, the "positive social value a person effectively claims for himself" (cf. Cassell and Bickmore, 2002, p. 6). Face can be "saved" or lost, and it can be threatened or conserved in interactions. Brown and Levinson further refine the concept of face into two specific subgoals that all social actors can be presumed to have:

1. *Positive face*—an individual's desire to be held in high esteem, to have his/her actions and opinions valued, to be approved of by others, etc.
2. *Negative face*—an individual's desire for autonomy, to have his/her will hold sway, to direct his/her at-tention where and when desired, etc.

The problem is that virtually all interactions between social agents involve some degree of threat to the participants' face—what Brown and Levinson call Face Threatening Acts (FTAs). My simple act of speaking to you, regardless of the content of my words, places a demand on your attention that threatens your negative face, for example. This, then, is the reason for the "please" in my request for the salt above: If I simply state my desire that you give me the salt as bald propositional content (e.g., "Give me the salt") I may efficiently communicate that intent, but I have also been ambiguous about whether or not I have the power or right or can otherwise compel you to give me salt. You might well take offense at the implication that I could demand salt from you, and in this way, I would have threatened at least your negative face and perhaps your positive face as well.

The "please" in the example above is an example of a politeness strategy used to "redress" or mitigate the threat contained in the request for the salt. Furthermore, the expectation that such a strategy be used in certain contexts is an example of etiquette that enables interpretations. The etiquette which we believe to be in play entitles us to conclude that those who use "please" in an appropriate context are striving to play by the rules—striving to be seen as polite; those who do not are not striving to be polite for various reasons (perhaps they don't believe they need to be, perhaps their notions about politeness are different, perhaps they are just rude.)

3.2 Computing the severity of a face threat

The core of Brown and Levinson's model is the claim that the degree of face threat posed by an act is provided by the function:

$$Wx = D(S,H) + P(H,S) + Rx$$

❑ Wx is the 'weightiness' or severity of the FTA

❑ D(S,H) is the social distance between the speaker (S) and the hearer (H). Social distance is roughly equivalent to familiarity—it increases with contact and interaction, but may also with be based on a priori factors such as membership in the same family, clan or organization and perhaps on being in a "familiar" setting as opposed to a formal one—a sporting event rather than a church. Social distance is a symmetrical relationship—S and H share the same social distance. In training contexts, social distance might derive from familial or clan relationships among avatars, or it might be used to convey (or invite) a deeper degree of familiarity with an avatar tutor, sidekick or counselor.

❑ P(H,S) is the relative power that H has over S, the "degree to which H can impose his own plans and his own self-evaluation (face) at the expense of S's plans and self-evaluation" (Brown and Levinson, 1987, p. 77). Power comes from different sources in different cultures and organizations, but power relationships are likely to be of particular importance military tutoring environments. Clearly, a tutor needs to maintain some power over a student, but avatars representing commanders, subordinates, or high or low status citizens might all need to act, and to be handled according to different etiquettes if face threats are to be minimized. Power is an asymmetric relationship between S and H.

❑ Rx is the ranked imposition of the raw act itself. Some degree of imposition is culturally defined—it may be inherently more of an imposition to request food from a host in Western culture than in an Arabic one, for example. But imposition is also dependent upon the roles and duties of the parties involved. One reason a tutor can correct a pupil, even though s/he might have lower power in the society, is that the correction is expected from the tutor and is, therefore, less of an imposition.

Brown and Levinson themselves do not operationalize these parameters; instead, they are offered as qualitative constructs. Recent work by Cassell and Bickmore (2002) and by Johnson (2003) has created numerical representations for them. In Cassell and Bickmore's work, the resulting computational model was used as a component in a conversational avatar (a real estate agent) whose goal is to use small talk to increase familiarity to the point where a more face threatening conversational topic (such as personal income level) can be introduced. Johnson has used a similar model to create a pedagogical agent that is designed to maintain and enhance learner confidence and motivation, by offering advice and criticism in ways that protect the learner's face.

3.3   Redressing face threats

Since FTAs are potentially disruptive to human-human relationships, and since we generally wish to avoid disruption, we generally make use of redressive strategies to mitigate the degree of face threat imposed by our actions. Brown and Levinson offer an extensive catalogue of universal strategies for redressing, organized according to 5 broad strategies. These are illustrated in Figure 1 ranked from least to most threatening.

❑ The least threatening approach is simply not to do the FTA. At some threshold, in some contexts and cultures, it will simply be too threatening for some FTAs to be performed, regardless of the amount of redress offered. At this point, the only viable strategy is to avoid doing the act.

❑ If one is to do the FTA at all, then the least threatening way to do it is "off record". Off record FTA strategies are means of doing the act with a sort of "plausible deniability" by means of innuendo and hints. If done successfully, S can accomplish the goals of the act without running the risk of face threat because, after all, the FTA was never overtly done. An "off record" method of asking for the salt from the example above might be "I find this food a bit bland". By using this approach, I have not implied that I have any right to demand salt from you, or anyone—in fact, I haven't even asked for salt.

❑ If one does FTA overtly, then one can still undercut its degree of threat by offering redress aimed at either positive or negative face. Brown and Levinson
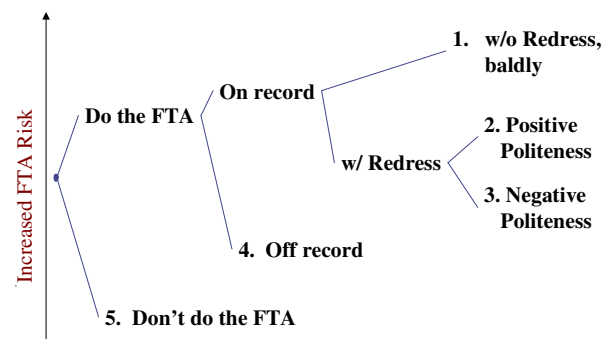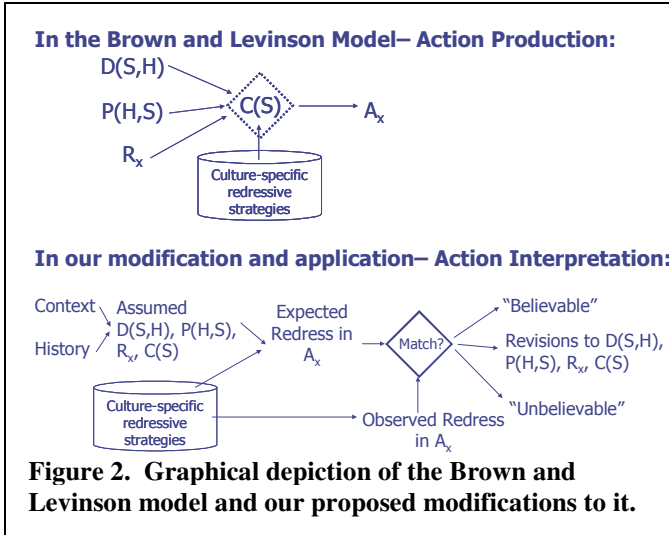

Figure 1. Universal redress strategies as ranked by Brown and Levinson (1987).

Figure 2. Graphical depiction of the Brown and Levinson model and our proposed modifications to it.

suggest that negative redress will be more effective (less threatening) than positive. Negative redressive strategies focus on H's negative face needs— independence of action and attention. They minimize the impact on H by being direct and simple in making the request, offering apologies and deference, minimizing the magnitude of the imposition and/or explicitly incurring a debt. "I'm sorry, but I'd be very grateful if you could just pass me the salt" includes many negative redress strategies (apology, incurred debt, minimization of the imposition).

❑ Positive redressive strategies target the hearer's positive face needs—the desire that his/her needs and wants be seen as desirable. These strategies emphasize common ground between S and H by noticing and attending to H, by invoking in-group identity, by joking and assuming agreement and/or by explicitly offering rewards/promises. "Hey buddy, you want to pass me that salt, don't you?" is a positive redressive strategy that incorporates both an in-group identity marker and assumed compliance.

❑ Finally, the most threatening way of performing an FTA is "baldly, on record" and without any form of redress. In some cases where power of S over H is high, familiarity is high and/or imposition is extremely low, doing an FTA with no form of redress may be the expected thing to do. The "Give me the salt" example used above is a bald, unredressed form of performing that FTA.

Brown and Levinson's model doesn't stop at that level, however. For positive and negative redressive and off record strategies, they offer a host of well-researched examples from at least three different language/culture groups (English, Tamil and Tzeltal) organized into a structure of mutually supporting and incompatible

approaches. We do not have space to present their findings in depth, but we note as an example that their categorization of negative redress strategies contains 10 alternate approaches, some of which are mutually supporting or conflicting, including:

❑ *Be Pessimistic*—"You're not going to pass me the salt, are you?"

❑ *Minimize the Imposition*—"Could you just nudge that salt shaker over here?"

❑ *Give Deference*—"Excuse me, sir, would you pass the salt?"

❑ *Apologize*—"I'm sorry to interrupt, but would you pass the salt?"

4. An "Etiquette Quotient"— Believable levels of Politeness

According to the Brown and Levinson model described above, people generally want to accomplish their goals expeditiously-- and this argues for minimizing redressive strategies. But they also experience a range of social and personal pressures to not threaten the face of those they interact with (especially those with greater power or shared familiarity)-- and this argues for extensive redressive strategies. The balance between these pressures yields the selection of specific strategies in context. Brown and Levinson allude to, but don't explicitly include a factor representing the relative weighting that an individual puts on his/her own goals vs. the face goals of others-- his/her general willingness (independent of the other factors) to place others needs first. For want of a better term, we'll call that "character" and introduce a term for it, abbreviated as C, with the character of speaker (S) being C(S). In other words, the degree of redress that a speaker S chooses to use will be a function of the degree of face threat inherent in the act (itself a function of P,D and R) and the speaker's character C(S).

But the above description, and indeed Brown and Levinson's primary focus, is from the perspective of the speaker/actor (S) interested in achieving interaction goals and, presumably, in avoiding face threat to hearers (Hs). We can characterize Brown and Levinson's model graphically as in the top portion of Figure 2. A Speaker with a given character C(S), uses his/her knowledge of the D, P, and R of a given context and desired FTA in order to select a strategy from among a knowledge base of culture-specific strategies resulting in a specific action A_x which is designed to both further S's goals and to avoid undue face threat to his/her interlocutors.

In order to implement and make use of this model in believable human-computer interactions (i.e., with

avatars), however, we need to take the perspective of an observer/hearer (who may or may not be the one the speaker is actually interacting with). This perspective is represented graphically at the bottom of Figure 2. Here, an observer (O) perceives an utterance that has bald content as a speech act and may or may not contain culturally-recognized redressive strategies. O also has access to additional cues from his/her perception of the context and perhaps memory for past history. Given these cues, O's goal is to construct a picture of the "politeness" character of S and, through that, to the P, D and R of the interaction between S & H. Fundamental to our approach is the claim that this construction process is based largely on the degree of match or mismatch between the redressive strategies actually used by S (as perceived by O) and those expected by O.

Given his/her own observations or knowledge of the context, O can construct an understanding of the parameters P, D, and R. For example, if S is noticeably older, richer, or is wearing insignia that make it clear that s/he outranks H, then O might reasonably conclude that the power distance (P) between them is large and favors S. If S and H are behaving familiarly (standing close together, interacting jovially, using nicknames, etc.), are known to be related as family members or friends, etc., then O might conclude that the social distance (D) is comparatively small between them. Finally, O will have his/her own culturally-based beliefs about the degree of imposition (R) of a given act (e.g., asking for money is a greater imposition than asking for help finding a location, which is a greater imposition than asking for the time), but observed or known characteristics of the interaction may also serve to reduce the perceived R. For example, if S is known to have a duty (perhaps based on his/her role) or a standing request to provide certain information or advice to H, or if H is not apparently en-gaged in any ongoing activity.

Then, given his/her beliefs about these parameters, O can construct an estimate of the degree of face threat associated with the bald content of the act. Furthermore, given whatever information s/he possess about C(S) , O can adjust his/her predictions about the degree, and therefore the types, of redressive actions that s/he might expect to see used. Let us call this product the expected act (A_x).

But at the same time, O can actually perceive an observed act (Ax). S performs an act that O will perceive as having a degree of imposition and, perhaps, various associated redressive actions. If the observed act and the expected act are the same (perhaps within certain degrees of tolerance), then the actor will be seen as believable—at least with regards to his/her/its politeness-producing etiquette behaviors.

Therefore, conceptually at least, one metric for believability is the delta between the expected act and the perceived act. And yet, other humans fail to behave as we expect them to behave all the time without our labeling them "unbelievable". This seems to be because humans are generally aware that predicting politeness behaviors is far from an exact science. We are generally more willing to revise our beliefs about aspects of the context or character that produced our initial predictions and then reassess that prediction than we are to conclude that S is acting artificially. This metric may be computed over time as well. If successive actions, with their associated degrees of redress employed, continue to violate O's notions of the avatars' context and P,D,R and C values, O may choose to revise the assumed characteristics seeking a set of P,D,R and C values that minimizes the delta between expected and observed degrees of redress. If no such model is found, or if violations are extreme, s/he may give the game up and simply declare those behaviors to be "unbelievable".

We are ultimately interested in two aspects of O's behavior. First, we want to be able to create avatars that behave in such a way as to avoid that reassessment process entirely by being believable in the context in which they are observed. As discussed above, believability in this sense means conforming to O's beliefs about how S should (in the sense of expectations, not necessarily prescriptive norms) behave. Second, and logically subsequent to the above, we also want to create avatars which can control or affect O's reassessment process through aspects of S's behavior (i.e., to cause O to believe that S has a different power status than might have been expected, or that the context is such that the imposition of this intrusion-- say, a warning-- is less than might otherwise be the case). Before we can get to that, however, we have to be able to minimize the chance that O will simply conclude that the actor/avatar is unbelievable or artificial.

Thus, the Brown and Levinson model provides the seed for a computational model for predicting, assessing and ultimately creating believable politeness behaviors in avatars in a wide range of social and cultural contexts. Their model does not, by itself, provide these capabilities, but the general modifications outlined above will take us in the direction of adapting their model to provide predictions about believability.

A completed and validated model would provide the following benefits:

1. An objective and verified predictive metric for scoring the degree of believability associated with various potential actions in an interaction between two or more social agents.

2. The opportunity to use that metric predictively to determine how a proposed behavior will be perceived—whether as believable, unbelievable, or as implying a revision in assumptions about of D, P, R and C(S)).

3. The opportunity to use that metric adaptively to shape the etiquette behavior of an avatar to adjust how it is perceived—both initially in context and in an ongoing fashion, making it adapt to polite or impolite behaviors from other avatars or human actors in believable ways.

4. Ultimately, the opportunity to adapt this basic, underlying model of believability in etiquette to the specific cultural methods by which these universal politeness dimensions are realized in different cultures. The fact that Brown and Levinson's model was developed from a diverse set of cultures and has abstracted away from them to universals makes it possible for us to move back from those universals to cultural specifics "simply" by plugging in a library of culture-specific observed etiquette behaviors—for example, the specific facial expressions, body postures, gestures, utterances and tone of voice used to issue an apology in Iraq vs. one in Somalia.

5. Conclusions and Future Work

In the coming months we will be developing an implementation of the Brown and Levinson model and testing its predictions about the believability of a variety of social interactions. Preliminary work will be performed within nominal American cultural settings and we will begin with non-interactive, scripted avatar-avatar interactions. Future work, however, will incorporate the believability metric we derive in an interactive avatar architecture and enable such avatars to ascertain the relative threat and redress of behaviors directed at them, as well as generating redressive behaviors in keeping with their goals and relationships with human player- or trainee-characters in gamelike interactions or simulations. Beyond that, we hope to begin work on developing "culture modules"—the representation and culture-specific knowledge to be integrated into our basic Brown and Levinson-based computational algorithm to give the resulting avatars the specific library of observable cultural redressive behaviors and sensitivities to behaviors directed at them that they might possess if they were, say, Iraqi, Kosovar, German or French.

The outcome of our proposed developments will be a dramatic increase in the ability to rapidly create computer training simulations or games with realistic, culture-specific social interaction models for their NPCs. We anticipate at least a 10x reduction in the time required to generate equivalently rich social interactions included in a 30 minute game episode as compared to the use of traditional scripting approaches. Our approach will also provide for much greater flexibility in the interactions which NPCs can exhibit—and thus, richer interaction capabilities which will, by allowing greater and more user-driven exploration capabilities, have payoffs in terms of the engagement and training effects in the applications in which they are used.

The ability to provide every individual soldier with a training simulation or, better a game which s/he would want to play in spare time, prior or during the early phases of deployment to a new culture would vastly increase the dissemination of training about the peoples with whom the soldiers will need to interact. Instead of training only 200 soldiers prior to their deployment, as was done at the beginning of Operation Iraqi Freedom, at the cost of weeks in training, transportation to Jordan and the full time involvement of those few cultural experts available (Mares, 2003), *all* soldiers to be deployed to Iraq could easily have been provided with game-based materials for the cost of initial game generation and digital media duplication. If only 10% of the roughly 200,000 U.S. troops who were initially deployed were to have played even portions of the game (an extraordinarily low estimate, given anecdotal evidence of game usage during relaxation hours for deployed troops), the result would be a hundredfold increase in the number of soldiers exposed to Iraqi politeness assumptions and, therefore, better equipped to interpret, navigate and manipulate them as needed—with no increase in training time or resources.

Such training would be expected to improve the efficiency of intelligence, stability and even combat operations. Including the cultural dimension enables a more complete intelligence picture. As Dr. Paul Bellutowski, of the U.S. Marine Corps Command and Staff College, wrote: "Understanding culture may help to answer important military and civil questions such as the will of the enemy to fight, the determination of resistance groups to persevere, or the willingness of the populace to support insurgents or warlords." (Bellutowski, 1996).

6. Acknowledgements

7. REFERENCES

Brown, P. & Levinson, S. *Politeness: Some Universals in Language Usage*. Cambridge Univ. Press; Cambridge,UK. 1987.

Cassell, J. and T. Bickmore, Negotiated Collusion: Modeling Social Language and its Relationship Effects in Intelligent Agents. *User Modeling and User-Adapted Interaction. 13*(1): 89-132

Chatham, R. *A Possible Future for Military Training*, http://www.t2net.org/briefs/dakc/DARPA_TrainingVision_5D8BF.pdf, July 2003

Chatham, R. and J. Braddock, *Training Superiority and Training Surprise*, Report of the Defense Science Board, 2001, http://www.acq.osd.mil/dsb/trainingsuperiority.pdf

Chatham, R., and J. Braddock, *Training for Future Conflicts*, Report of the Defense Science Board, 2003.

Dennet, D., *The Intentional Stance*, MIT Press; Cambridge, MA. 1989.

Goffman, E. *Interaction Ritual: Essays on Face to Face Behavior.* Garden City; New York. 1967

Grice, H.P., Logic and Conversation, In P. Cole and J. Morgan (Eds.), *Syntax and Semantics, vol. 3., Speech Acts*, Academic Press; New York, 1975, pp. 41-58.

Johnson, L. (2003). Attentiveness and Politeness in Learner-Computer Communications.

Mares, Pvt. K. L. Jordanian course preps soldiers on Arabic culture, *Army News Service*, Nov. 10, 2003. http://www4.army.mil/ocpa/read.php?story_id_key=5395