A Computational Approach to Etiquette and Politeness:
An “Etiquette Engine™” for Cultural Interaction Training

Christopher Miller
Peggy Wu
Harry Funk
Lewis Johnson
Hannes Vilhjalmsson
Military forces increasingly need training in the cultures they deploy to (esp. for urban ops)

Cross Cultural training helps, but is a huge drain on resources and time
- 200 soldiers trained for Iraq deployment (Mares, 2003)

Soldiers willingly spend off-duty time playing PC games

**Solution:** Make games that teach elements of culture
- Estimated DoD-wide savings of >$1B/year (Chatham, 2003)

**Problem:** Avatar behaviors are hand-scripted, hence brittle and expensive, or driven by first-principles models—hence expensive and unpredictable

**Solution:** A general, computational social interaction model with “pluggable” culture knowledge modules
To avoid expensive modeling solutions, we need a general, computational model.

We have access to a qualitative theory of how humans decide what politeness behaviors to exhibit in social interactions.

- Brown and Levinson

We have developed a computational model based on that theory.

We need to implement/integrate it in a game training environment.

We need to make predictions about scale up.
Brown and Levinson—cross-cultural, socio-linguistic, human-human, politeness model

- "Face Threat" is a function of Power, Familiarity and Imposition
- Redress must $\approx$ Face Threat for "Nominal Politeness"
  - More will be perceived as "over polite"; less as rude
- Redressive politeness actions fall into general categories
- Imbalance yields reinterpretation
  - of character (rudeness, intelligence, believability)
  - or of situation (urgency, power, familiarity, imposition)
B&L’s Politeness Strategies

1. w/o Redress, baldly
2. Positive Politeness
3. Negative Politeness
4. Off record
5. Don’t do the FTA

Do the FTA

On record

w/ Redress

Increased FTA Risk
Negative Politeness Strategies

1. Be conventionally indirect
2. Question, hedge
3. Be pessimistic
4. Minimize the imposition, Rx
5. Give deference
6. Apologize
7. Impersonalize: avoid pronouns I and you
8. State the FTA as a general rule
9. Nominalize
10. Go on record as incurring a debt, or as not indebting H

Be direct

Don’t presume/assume

Make minimal assumptions about H’s wants

Don’t coerce

Give H option not to act

Communicate S’s want not to impinge on H

Do FTAx (a) on record (b) plus redress to H’s want to be unimpinged upon

Minimize threat

Make explicit R,P,D values

Dissociate S, H from the particular infringement

Redress other wants of H, derivative from negative face

16th BRIMS Conference 26-29 March 2007
Adapting B&L 1

- B&L are attempting to explain/account for Action Production
- They note that P, D, and R affect redressive strategy usage— and that these factors are culturally informed
- They note that redressive strategies themselves are culturally-specific (though there are universal abstractions)
- We infer a “Character” variable— the predisposition of an individual to subvert his/her own face to others

\[
\begin{align*}
& D(S,H) \\
& P(H,S) \\
& R_x \\
\rightarrow & C(S) \\
\rightarrow & A_x
\end{align*}
\]

- Culture-specific Markers
- Culture-specific redressive strategies
Adapting B&L 2

For Normal/Believable/Unremarkable Interactions: \( W_x = A_x \)
Adapting B&L 3

*We need to account for interpretations of believability of observed redressive strategies*

**Goal of Phase I: an Etiquette Quotient**
- A Believability Metric based on the degree to which Face Threat = Redressive Actions
- **Assumption:** “Believable” means expected redress in context
- Unexpected Redress may mean “unbelievable” or reinterpretation of context
The Grand Vision

- Imbalance Metric can be used for scoring & evaluation
- More valuable when integrated for avatar behavior recognition and generation
- “Culture Modules” to swap in and out to give avatars culture-specific etiquette sensitivities and reactions

Familiarity Level (D)  Power Level (P)  Imposition Level (R)

Character  

Selected Politeness/Redress Level

Selected Actions for Redress

Culture-specific Knowledge Base
USC’s TLTS

The DARPA Tactical Language Training Project

Language and gesture exercises culminating in simulated missions

Learner speaks for, and chooses gestures for on-screen avatar. Other characters respond automatically.

System tracks learner proficiency and adapts character behavior accordingly.

Authoring and machine learning tools help with the creation of new modules.
Integration Architecture (simplified)

Who is saying what?
To Whom?
How (gesture, with hat off, etc.)?

What “etiquette imbalance” was used?

What would the character want to do in response (content plus etiquette imbalance)?

What CACT will achieve that intention (for these participants)?

Realize it (in animation and speech)
Resulting Representation

- Each agent has beliefs about all pairwise power difference and social distance
  - Examples:
    - # Trainee's beliefs (Trainee, everyone)
      - belief Trainee Trainee Malek P -20
      - belief Trainee Trainee Malek D 15
    - # Trainee's beliefs (Malek, everyone)
      - belief Trainee Malek Trainee P 20
  - Mismatches can certainly exist
    - Could be inherited (“cultural modules”). Could be dynamic.

- Each agent has beliefs about the imposition and redressive value of various “Communicative Acts” (speech + gesture)
  - Examples:
    - # utterances
      - ## imposition, pos redress, neg redress
        - utterance GREET pashto 5 30 30 | salaam aleekum. | Peace be upon you
        - utterance GREET pashto 5 30 70 | salaam aleekum. [HOH] | Peace be upon you
        - utterance GREET_RESP pashto -3 30 30 | wa aleekum salaam.| And upon you be peace.
        - utterance GREET_RESP pashto -3 30 70 | wa aleekum salaam.[HOH] | And upon you be peace
    - When a CACT type is required (by conversational flow) all available CACTS of that type are considered
      - Known CACTS and assumed values may mismatch
        - Could be inherited (“Cultural Modules”)
## Redressive Action Types Used in Scoring

<table>
<thead>
<tr>
<th>CACT</th>
<th>R</th>
<th>V(Ax)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>wa salaam aleekum + HOH</code></td>
<td>-3</td>
<td>30 + 30 + 40 = 100</td>
<td>30 Pos Redress, cite H’s interests</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30 Neg Redress, use formality</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>40 neg redress, HOH gesture</td>
</tr>
<tr>
<td><code>wa salaam aleekum</code></td>
<td>-3</td>
<td>30 + 30 = 60</td>
<td>30 Pos Redress, cite H’s interests</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30 Neg Redress, use formality</td>
</tr>
<tr>
<td><code>wa aleekum + HOH</code></td>
<td>-3</td>
<td>30 + 40 = 70</td>
<td><strong>Less formality than above, but the HOH gesture makes it more redressive than wa salaam aleekum w/o the gesture</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30 Pos Redress, cite H’s interests</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>40 Neg redress, HoH gesture</td>
</tr>
<tr>
<td><code>wa aleekum</code></td>
<td>-3</td>
<td>30</td>
<td>30 Pos Redress, cite H’s interests</td>
</tr>
<tr>
<td><code>steRi meshey + HOH</code></td>
<td>-3</td>
<td>20 + 40 = 60</td>
<td>20 Pos Redress, cite H’s interests</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(less powerful)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>40 neg redress, HOH gesture</td>
</tr>
<tr>
<td><code>steRi meshey</code></td>
<td>-3</td>
<td>20</td>
<td>20 Pos Redress, cite H’s interests</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(less powerful)</td>
</tr>
</tbody>
</table>
Etiquette Inspection and Interaction Tool

wxPython GUI

File

Last Utterance:

CACT type: RECOGNITION

Observers:

Trainee

Speaker:

Trainee

Addressee:

Malek

Culture:

North American

C(S):

0.0

C(A):

0.0

O8:P(S,A):

-20.0

O8:D(S,A):

15.0

Utterance:

xoshala shwem. | (informal) Pleased.

Vx-Wx:

Vx:

Wx:

-10.0

-50.0

40.0
"Imbalance" Metric: \( I_x = \text{OP}:V(A_x) - \text{OP}:W_x \) where:
- Imbalance \( (I_x) \) of an interaction \( x \) is the
- Observer’s Perception (OP) of the
- Summed Value (V) of all Redressive Actions in the interaction \( (A_x) \)
- Minus the Observer’s Perception of the Weightiness (or degree of Face Threat) of the interaction \( (W_x) \)

Behavior:
- Politeness is balanced/nominal (about what was expected) when \( I_x = 0 \)
- When more redress is used than needed/expected ("over polite" interaction) \( I_x >> 0 \)
- When less redress is used than needed/expected (rude) \( I_x << 0 \)
- "Threshold of Believability"

Keep in mind that:
- "Imbalance" means no need to rethink initial, nominal assumptions
- All scores are as perceived/expected by the Observer
Scoring Threat and Redress

Relative Values for Face Threat are a function of:
- Power of Hearer over Speaker
- Social Distance between Speaker and Hearer
- Ranked (Raw) Imposition of the Interaction/Utterance
- Knowledge of the general Character of the Speaker

Redress values are scored according to the categories and relative values laid out by Brown and Levinson.

1. w/o Redress, boldly
2. Positive Politeness
3. Negative Politeness
4. Off record
5. Don’t do the FTA

Increased FTA Risk

Do the FTA
On record
w/ Redress
1. w/o Redress, boldly
2. Positive Politeness
3. Negative Politeness
4. Off record
5. Don’t do the FTA

Increased FTA Risk

Do the FTA
On record
w/ Redress
1. w/o Redress, boldly
2. Positive Politeness
3. Negative Politeness
4. Off record
5. Don’t do the FTA

Increased FTA Risk

Do the FTA
On record
w/ Redress
1. w/o Redress, boldly
2. Positive Politeness
3. Negative Politeness
4. Off record
5. Don’t do the FTA

Increased FTA Risk

Do the FTA
On record
w/ Redress
1. w/o Redress, boldly
2. Positive Politeness
3. Negative Politeness
4. Off record
5. Don’t do the FTA

Increased FTA Risk

Do the FTA
On record
w/ Redress
1. w/o Redress, boldly
2. Positive Politeness
3. Negative Politeness
4. Off record
5. Don’t do the FTA

Increased FTA Risk

Do the FTA
On record
w/ Redress
1. w/o Redress, boldly
2. Positive Politeness
3. Negative Politeness
4. Off record
5. Don’t do the FTA

Increased FTA Risk

Do the FTA
On record
w/ Redress
1. w/o Redress, boldly
2. Positive Politeness
3. Negative Politeness
4. Off record
5. Don’t do the FTA

Increased FTA Risk

Do the FTA
On record
w/ Redress
1. w/o Redress, boldly
2. Positive Politeness
3. Negative Politeness
4. Off record
5. Don’t do the FTA

Increased FTA Risk

Do the FTA
On record
w/ Redress
1. w/o Redress, boldly
2. Positive Politeness
3. Negative Politeness
4. Off record
5. Don’t do the FTA

Increased FTA Risk

Do the FTA
On record
w/ Redress
1. w/o Redress, boldly
2. Positive Politeness
3. Negative Politeness
4. Off record
5. Don’t do the FTA
Formalize & Implement Model 2

**Threat Weight** = \( \Sigma (Social\ Distance,\ Power\ Difference,\ Imposition,\ Character) \)

\[
OP: W_x = \{[\text{OP:SP:w1} \cdot \text{OP:SP:D}(S,H)] + [\text{OP:SP:w2} \cdot \text{OP:SP:P}(H,S)] + [\text{OP:SP:w3} \cdot \text{OP:SP:R}_x]\} + \text{OP:CS}
\]

**Behavior:**
- Threat is greater:
  - The more unfamiliar interactants are
  - The more power the Hearer has over the Speaker
  - The greater the imposition
- Perceived threat can be mitigated by the perceived nature of the Speaker’s character

**Keep in mind that:**
- Each factor is the Observer’s perception of the Speaker’s perception of the attribute (except for Character)
  - This is because the Observer is making a judgment about believability of Speaker’s behavior
- All values, therefore, rely on O’s understanding of S’s cultural norms about D,P and R.
  - Potential for differential weights on factors
Parameters

\[ OP: W_x = \{[OP:SP:w_1 \cdot OP:SP:D(S,H)] + [OP:SP:w_2 \cdot OP:SP:P(H,S)] + [OP:SP:w_3 \cdot OP:SP:R_x]\} + OP:CS \]

- All weights \((w_1, w_2, w_3)\) set to 1 for now (i.e., ignored)
- Character term \((OP:CS)\) simply added or subtracted
  - Nominal character = 0; Good is pos. (increases perceived \(W_x\) – the “Mother Theresa” effect)
- \(D, P\) and \(R\) use scales range from \(-\infty\) to \(+\infty\); 0 is nominal/equal
The Power Scale

-1000 -- the power a CEO has (as S) over/relative to a janitor or a parent over a small child.
-100 -- the power a professor has over a freshman or a parent over an early teenage child.
-10 -- the power that a project manager in an informal research team over project members or a parent has over an older teenager.
0 -- equal power between S and H; no or negligible difference (i.e., two co-workers at same level/seniority)
10 -- power that an older teenager or work team project member (as S) has relative to a parent or project manager (as H), respectively.
100 -- the inverse of the power described for -100 above.
1000 -- the inverse of the power described for -1000 above.
Formalize & Implement Model 4

Redress = \( \sum \) Value of Redressive Actions

\( \text{OP:SP } V(A_x) = \text{OP:SP } V(A_1) + \text{OP:SP } V(A_2) + ... + \text{OP:SP } V(A_n) \)

Behavior:
- Threat is greater:
  - The more unfamiliar interactants are
  - The more power the Hearer has over the Speaker
  - The greater the imposition
- Perceived threat can be mitigated by the perceived nature of the Speaker’s character

Keep in mind that:
- Each factor is the Observer’s perception of the Speaker’s perception of the attribute (except for Character)
  - This is because the Observer is making a judgment about believability of Speaker’s behavior
- All values, therefore, rely on O’s understanding of S’s cultural norms about D,P and R.

Potential for differential weights on factors

16th BRIMS Conference 26-29 March 2007
Redressive Action Value Scales

1. **w/o Redress, boldly**
   - Do FTA
   - On record
   - w/ Redress

2. **Positive Politeness**
   - Do FTA
   - (a) on record
   - (b) plus redress to H's want to be unimpinged upon

3. **Negative Politeness**
   - Don't do FTA
   - Off record

4. **Off record**

5. **Don’t do the FTA**

---

**Individual positive redress actions** → 1-40 redress “units”

**Individual negative redress actions** → 20-60 units

**Individual off record redress actions** → 40-80 units

**Multiple strategies simply add**

- But some encounters are so face threatening that even profuse apology just “digs S in deeper”
- Some need to reflect time and other goals here
The Etiquette Algorithm:

Enhancing Social Interaction in the Tactical Language Training System (TLTS)
TLTS EQ Demo

Available at:

http://www.sift.info/demos/eqdemo.html

Login: DemoAccess
Password: miufia
Scale up and Coverage

Novel PSs

Traditional Scripting Methods

Effort to Represent

- Modularity in Representation permits “syntactic” recombination of interactions
  - “Semantic” categories guide Knowledge Acquisition
- Computational algorithm allows ANY (scored) utterance used in ANY (scored) context
  - Explosive growth in coverage
  - Ease of character and utterance modification/insertion/alteration
- Payoffs:
  - Breadth of coverage
  - Inspectability
  - Reusability

EQM Representation & Scoring

Graph: Hours vs. Utterance Perceptions Scored
Conclusions and Future Work

- We’ve demonstrated the ability to integrate with a visual simulation architecture
  - Providing NPC interpretation and generation of etiquette
- Evidence indicates substantial payoff in scalability
- Validation of algorithm—experiments show it tracks expectations

Future Work:
- Language Training and Rehearsal
- Knowledge Acquisition Framework
- Directive Compliance models (both human and machine)
Cultural Modules for Rapid Creation of Training Simulations

SIFT, LLC—Dr. Chris Miller
USC/ISI—Dr. Lewis Johnson

Main Objectives

- **Phase I**—Develop algorithm for dynamic, culture-specific social interactions based on an abstract model of “face threats”
- **Out Phases**—
  - Make model interactive with users
  - Investigate methods for embedding culture modules
  - Demo portability in militarily-relevant game/simulation environment

Key Innovations

- Abstract, modular approach to social interaction “etiquette” knowledge
  - Supported by theory and 20 years of empirical observation
- Embedding in gaming/sim technology
  - Rapid generation of diverse Non-Player Characters (NPCs) that behave like culture-specific individuals
    - Take offense realistically for their culture
    - Offer redress realistically

Expected Impact

- 10x improvement in ability to generate NPC behaviors
- Improvements in game generation speed and cultural interaction
  - Accurate depiction of cultural norms
- More playable games → 100 to 10,000x improvement in distribution of Cross-Cultural Training for soldiers
  - Saves resources, time, dollars and lives
Does “politeness” make a difference in directive performance?
Does it vary cross-culturally?
“Cultural Factors” (Nisbett, Hofstede, etc.) too abstract for specific interaction models

**APPROACH:**
Quantitative etiquette model
Testbed for single human operator
  - Resource management task (UAVs/Fire Patrol)
  - “Chat channel” for simulated interactions
  - Vary politeness (and “culture”) of directives

**GOAL:** Predictive models of cultural differences in directive compliance

**STATUS:**
Phase I designed testbed and experimental procedures
Beginning Phase II: Conduct multi-cultural experiments