Nonverbal Behavior Generator for Virtual Humans

Jina Lee, Stacy Marsella
USC Information Sciences Institute
{jinal, marsella}@isi.edu

Introduction

**Problem Description**

- Utterance
- Function
- Behaviors

- I’m glad to hear that. We may not trust each other well. I can’t believe you did that...
- Emphasize
- Contrast
- Emotional

- Smile
- Brow
- Lowered
- Head nod
- Headshake

How do we find the mapping between utterance + communicative function to behaviors?

**Nonverbal Behavior Generator**

- Read in surface text and markup of virtual character’s affective and intentional state and generate appropriate behaviors (e.g. facial expressions, body gestures, gaze)

**Goal**

- Robust NVB generation that can use markup of communicative function if provided, but can also extract/infer it if not
- Extraction that leverages syntactic and semantic analysis of text

Current Implementation

**System Architecture**

- NVB Generator
- Communicative Function Derivation
- Behavior Suggestion [NVB Rules]
- Behavior Realization
- Output: Nonverbal Behavior Execution Code

**Examples of NVB Rules [Lee & Marsella, 2006]**

<table>
<thead>
<tr>
<th>Derivation</th>
<th>Function</th>
<th>Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>No, not, nothing, cannot, none</td>
<td>Negation</td>
<td>Head shakes on phrase</td>
</tr>
<tr>
<td>Really, very, quite, great, absolutely, gorgeous…</td>
<td>Intensification</td>
<td>Head nod and brow brow on word</td>
</tr>
<tr>
<td>Um/uh/well + interjection from parser</td>
<td>Word Search</td>
<td>Head tilt, brow raise, gaze away</td>
</tr>
<tr>
<td>But, however</td>
<td>Contrast</td>
<td>Head moved to side and brow raise</td>
</tr>
<tr>
<td>Everything, all, whole, several, plenty, full…</td>
<td>Inclusivity</td>
<td>Lateral head sweep and brow flash on word</td>
</tr>
<tr>
<td>Have to, need to, ought to</td>
<td>Obligation</td>
<td>Head nod once on phrase</td>
</tr>
</tbody>
</table>

Ongoing Research: Machine Learning for Gesture Models

- Start with head gesture model (e.g. head nods, head shakes, etc.), then extend to other gestures
- Gesture Corpus: AMI Meeting Corpus [Carletta et al., 2006]
  - Multimodal data set of meeting recordings
  - Contains video, audio, transcripts, dialog acts, topic segmentations, focus of attention, head gestures, hand gestures, etc.

**Current Progress**

**Classifiers**

- HMM used for visual recognition of arm gestures [Brand, Oliver, Pentland, 1996], sign languages [Assan & Groebel, 1997]
- CRF used for human motion activities [Smichiescu et al., 2005]
- Variations of CRF (HCRF, LDCRF) used for head gesture recognition [Wang et al., 2006], [Morency et al., 2007]

**Feature Selection**

- Part of Speech
- Phrase boundary
- Dialog act
- Salient words (emotionally charged words, words used in current NVB rules)
- Speaker’s role (speaker vs. listener)
- Theme / rheme

Future Work

- Explore ML techniques for other gesture models (arm, posture, etc.)
- Evaluation of the system and behaviors generated
- Modify and customize the current behavior generation for different gender, cultures, or personalities