Nonverbal Behavior Generator

Jina Lee
Stacy Marsella
Information Sciences Institute
NVBG Overview

- **Goal**
  - Reads FML + BML (speech) message and generates NVB specified in BML
  - Research goal: Robust NVB generation that can use markup of communicative function if provided, but can also extract/infer it if not
  - Clear distinction of function and behavior
Architecture of NVBG

Input
Surface Text + Agent's Emotional State (FML)

Communicative Function Derivation
Behavior Suggestion
Behavior Realization

NVBGenerator

Behavior Filtering
Output
Nonverbal Behavior Execution Code (BML)

Function Rules
NVB Rules
Behavior Description

Cache

Surface Text
Parse Tree

Natural Language Parser (Charniak Parser)
Example of Nonverbal Behavior Generation

Surface Text:
I was mad at him.

Function Rules:
- first_NP Rule
  Priority 5
- Me Rule
  Priority 5
- first_VP Rule
  Priority 5
- noun_phrase Rule
  Priority 5

Behavior Rules:
- Head nod
- Head nod
- Beat gesture
- Head nod
# Input Message Types

<table>
<thead>
<tr>
<th>Message Type</th>
<th>FML Includes</th>
<th>BML Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech</td>
<td>&lt;turn&gt; in FML &lt;speech&gt; in BML</td>
<td>Various NVB according to syntactic/semantic structure of surface text</td>
</tr>
<tr>
<td>Gaze</td>
<td>&lt;visualAttention&gt; (rationale behind what’s affecting gaze behavior)</td>
<td>Parameters for gaze behaviors (eg. angle, offset, etc.)</td>
</tr>
<tr>
<td>Emotion</td>
<td>&lt;affect&gt;</td>
<td>Emotion doesn’t directly generate BML msg. Instead, it implicitly modifies NVBs</td>
</tr>
<tr>
<td>Listener Feedback</td>
<td>&lt;listenerFeedback&gt;</td>
<td>Different facial expression and head movements according to the listener’s agreement with the speaker</td>
</tr>
<tr>
<td>Negotiation Stance</td>
<td>&lt;negotiationStance&gt;</td>
<td>Changes the agent’s posture</td>
</tr>
</tbody>
</table>
Message Formats

<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<act><participant id="elder-al-hassan" role="actor"/>
<bml>
<speech id="sp1" ref="" type="application/ssml+xml">
  <mark name="T0"/>I
  <mark name="T1"/>was
  <mark name="T2"/>mad
  <mark name="T3"/>at
  <mark name="T4"/>him.
</speech>
</bml>
</act>

<!--first_VP Animation-->
<animation name="HandsAtSide_RArm_MidBeat" priority="5" ready="sp1:T0" stroke="sp1:T3"/>

<!--Noun clause nod-->
<head amount="0.10" priority="5" ready="sp1:T8" relax="sp1:T9" repeats="0.5" type="NOD"/>
</bml>

FML

BML (speech)

BML for behaviors generated by NVBG

Output
Behavior Generation Process

Agent → FML (+ BML) → Communicative Function Derivation → Behavior Suggestion → Behavior Realization → BML → SmartBody

rule_input_general.xsl → nvb_rules.xsl → behavior_description.xsl → transform.xsl
# NVB Rules – from surface text

<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 frown on word</td>
</tr>
<tr>
<td>Yes, yeah, I do, We have, It’s true, OK</td>
<td>Affirmation</td>
<td>Head nods and brow raise on phrase</td>
</tr>
<tr>
<td>I guess, I suppose, I think, maybe, probably, perhaps, could</td>
<td>Assumption / Possibility</td>
<td>Head nods on phrase</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>You, me</td>
<td>Believability</td>
<td>Gestures pointing towards self or other</td>
</tr>
</tbody>
</table>
## NVB Rules – from parse tree

<table>
<thead>
<tr>
<th>Derivation</th>
<th>Function</th>
<th>Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>INT (interjection)</td>
<td>Interjection</td>
<td>Head nod on word</td>
</tr>
<tr>
<td>First &lt;NP&gt; (noun phrase)</td>
<td>Believability</td>
<td>Big head nod on start of the noun phrase</td>
</tr>
<tr>
<td>&lt;NP&gt; (noun phrase)</td>
<td>Believability</td>
<td>Head nod on start of noun phrase</td>
</tr>
<tr>
<td>First &lt;VP&gt; (verb phrase)</td>
<td>Believability</td>
<td>Beat gesture on start of the first verb phrase</td>
</tr>
</tbody>
</table>
rule_input_general.xml

- Specifies NVB rule, keywords, priority, gesture clips to be called

- If simply changing the animation clips called (without modifying the NVB rules), modify this file and it’s done!

- Adding a rule (ex. ‘emo_negative’)

```xml
<!-- negative -->
<rule keyword="emo_negative" priority="1">
  <pattern>mad</pattern>
  <pattern>argument</pattern>
  <pattern>harm</pattern>
  <pattern>trouble</pattern>
</rule>
```
transform.xsl

- Creates `<rule>` element
- Specifies the parameters needed for each rule (ex. `speech_id`, `ready_time`, `relax_time`, `priority`, etc.)

```xml
<xsl:when test="$rule_name='emo_negative'">
  <xsl:element name="rule">
    <xsl:attribute name="ready">…</xsl:attribute>
    <xsl:attribute name="relax">…</xsl:attribute>
    <xsl:attribute name="type">…</xsl:attribute>
    <xsl:attribute name="priority">…</xsl:attribute>
    <xsl:text>&#10;</xsl:text>
    <xsl:call-template name="emo_negative" >
      <xsl:with-param name="speech_id" select="$sp_id" />
      <xsl:with-param name="ready_time" select="$ready" />
      <xsl:with-param name="relax_time" select="$relax" />
      <xsl:with-param name="priority" select="./@priority" />
    </xsl:call-template>
  </xsl:element>
</xsl:when>
```

Call template with these parameters
Specifies which (generic) behavior is associated with each nvb rule

<!-- emo_negative: head nod and brow frown -->
<xsl:template name="emo_negative">
    <xsl:param name="speech_id" />
    <xsl:param name="ready_time" />
    <xsl:param name="relax_time" />
    <xsl:param name="priority"/>
    <xsl:comment>Emotion Negative</xsl:comment>
    <xsl:call-template name="nod">
        <xsl:with-param name="speech_id" select="$speech_id" />
        <xsl:with-param name="ready_time" select="$ready_time" />
        <xsl:with-param name="relax_time" select="$relax_time" />
        <xsl:with-param name="priority" select="$priority" />
    </xsl:call-template>

    <xsl:call-template name="brow_frown">
        <xsl:with-param name="speech_id" select="$speech_id" />
        ...
    </xsl:call-template>
</xsl:template>
behavior_description.xsl

- Creates the actual BML code for each generic behavior

- Specified behaviors
  - Nods:
    - big_nod, nod, small_nod
  - Shakes:
    - big_shake, shake, small_shake, shake_twice
  - Head orientations:
    - head_to_left, head_to_right, head_up, head_down, head_tilt_right, head_tilt_left
  - Eye brow movements:
    - brow_frown, brow_raise
  - Mouth movements:
    - smile
  - Animations:
    - ani_beat, ani_greeting, ani_you, ani_me, ani_negation, ani_contrast, ani_assumption, ani_theoretical, ani_inclusivity, ani_question, ani_obligation