**MPML for Describing Multimodal Contents with Lifelike Agents**

**Mitsuru Ishizuka**
Dept. of Information & Communication Eng. & Dept. of Creative Informatics
School of Information Science and Technology
University of Tokyo

---

**Progress in Lifelike Embodied Agents**

- *Research Activities from approx. 1990 at*:
  - DFKI, USC/ISI, CMU, NCSU, Stanford, MIT, Univ. of Rome, Curtin Univ. of Tech., Microsoft, etc.,
  - and Univ. of Tokyo

  have been showing the feasibility and positive effect as new multimodal media and new educational media.

- *Necessary media components are becoming available.*
Some Cognitive Backgrounds

**Non-verbal Communication**
by Albert Mehrabian

- via Language (flat sentence) 7%
- via Speech with tone and intonation 38%
- via Facial expression and Gesture 55%

**The Media Equation**
by B. Reeves, C. Nass

*Media = Real Life*

**The Persona Effect**
The presence of a lifelike character even one that is not expressive - can have a strong positive effect on student’s perception of their learning experience.

**Dimensions:**
- motivation, entertainment,
- helpfulness, ...

Many Component Technologies are necessary to build a system.
Rich & Cool Multimodal Media not only for everyone, but also by everyone

Need for XML-based Description Language

- MPML (Multimodal Presentation Markup Language)
- VHML (Virtual Human Markup Language)
- CML/AML, APML, RRL-NECA, BEAT, ...
- HumanML (Human Markup Language)

Some Competitors
- VoiceXML
- Web3D (Shockwave3D, …….)
- MS Narrator

**MPML concept**
(Multimodal Presentation Markup Language)

- Multimodal Presentation Anytime, Anyplace through the network (even to mobile).
- Allows Anyone (ordinary people) to write effective/attractive Multimodal Presentation Contents easily.

- Serves as an extensible center integrating many advanced functional modules.
**MPML** as a Markup Language conformed to XML

A simple example of **MPML** script

```xml
<mpml>
  <head>
    <spot id="spot1" location="200,260"/>
    <agent id="simasan" system="MSAgent" character="simasan"
      voice="LH" agreeableness="50" activity="50" spot="spot1"/>
  </head>
  <body>
    <seq>
      <scene agents="simasan">
        <page ref="page0.html">
          <act agent="simasan" act="greet"/>
          <speak agent="SmArt1">
            <emotion assign="simasan:happy+">
              Hello! My name is Sima. Welcome to our Web.
            </emotion>
          </speak>
        </page>
      </scene>
    </seq>
  </body>
</mpml>
```
**History of MPML**

1998
- MPML Ver. 1.0
  - DWML
  - Dynamic Objects as well as dynamic characters, XSL

- MPML Ver. 2.0a
  - Multiple Agents
  - XSL implementation

- MPML Ver. 2.0e
  - Emotion based on OCC model

- MPML Ver. 3.0
  - SmArt Agent
  - Emotion Mechanism
  - Visual Editor

- MPML-Flash
  - 3D Agent in VRML

- MPML-VR
  - for Mobile phones

- MPML-mobile
  - Humanoid Robots

- MPML-HR

**MPML Play -- several ways**

- MPML Editor
- Converter 1
  - (ViewMpml)
- XSL-based Converter
  - (Plug-in to IE)
- Converter 2
- Converter 3
- Converter for Mobile Phones

- DWML
  - for Dynamic Web Contents

- MS-Agent
- 3D Agents in VRML
- SmAart Agent
- Charas for Mobile
MPML’s position in the taxonomy of description languages

- Medium level
dialogue management based on finite state machine and a memory mechanism.
- Easy authoring for ordinary people in language level (like HTML for Web contents)

MPML 2.0 featuring Full Presentation with Emotional Expressions

MPML Version 2.0c
The main extension in Version 2.0c:
- Emotion express function

How does lifelike agent express emotion?
- medium
- volume
- pitch
- emphasis
MPML3.0

Graphical Editor

with SmArt Agents

MPML-VR

-- Presentation in 3D VRML Space
3D Agents in VRML space

Andy and Aya

MPML-HR (humanoid robot) version
MPML-HR for Honda’s ASIMO

MPML-mobile 0.5 for J-Phone (J-Sky) and 1.0 for DoCoMo’s i-mode

In Cooperation with Hottolink Inc.
**DWML**

*Dynamic Web Markup Language*

*Animation control not only for character agents, but also for all objects.*

---

**Features required for Affective Lifelike Agents**

<table>
<thead>
<tr>
<th>Embodiment</th>
<th>Artificial Emotional Mind</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synthetic Bodies</td>
<td>Affective-based Responses</td>
</tr>
<tr>
<td>Emotional Facial Display</td>
<td>Personality</td>
</tr>
<tr>
<td>Communicative Gestures</td>
<td>Response adjusted to Social Context</td>
</tr>
<tr>
<td>Posture</td>
<td>Social role awareness</td>
</tr>
<tr>
<td>Affective Voice</td>
<td>Adaptive Behavior</td>
</tr>
<tr>
<td></td>
<td>Social intelligence</td>
</tr>
</tbody>
</table>
More Affective and Social

Workable Model of Emotion/Personality

- Friendliness, Entertainment, Satisfaction, Motivation, Relax, ...  
- Lifelikeness (Illusion of Life) or Believability
- Emotion Elicitation
- Emotion Manifestation
- Many Other Factors

Affect expression is also an important feedback channel in communication. For example, when your interlocutor frowns, you know something is wrong in the conversation.

OCC (22 emotions) and Lang’s 2-dimensional Emotion Models

- Valence: positive or negative dimension of feeling
- Arousal: degree of intensity of emotional response

- Emotion Structure
- Consequences of Events
- Actions of Agent
- Aspects of Objects

- Consequences for Others
- Consequences for Self
- Self Agent
- Other Agent

- focusing on
- desirable for others
- undesirable for others
- happy-for / greeting / acceptance
- resentment / rejection
- pity / compassion

- focus of interest / involvement
- pride / confidence
- envy / disappointment
- well-being / attribution compounds

- gratification / gratitude
- remorse / anger
- self-hating / attribution compounds

- hope / fear
- satisfied / dissatisfied
- satisfaction / relief
- happy / confirmed / disappointment

- prospection / reflection
- confirmed / disconfirmed
McCrae and Costa’s 2-dimensional Personality Model (89)

- **Dominance**: individual’s disposition to control
- **Friendliness**: tendency to be warm and sympathetic

Scripting Emotion in *MPML2.0e*

```xml
<mpml>
<head>
<title>MPML Presentation</title>
</head>
<body>
<page id='first' ref='self_intro.html'>
<emotion type='happy-for'>
<speak>
I am Mitsu Ishizuka from the Univ. of Tokyo.
</speak>
</emotion>
</page>
</body>
</mpml>
```
Emotion and Voice Parameters

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Speech Rate</th>
<th>Average Pitch</th>
<th>Loudness</th>
<th>Pitch Range</th>
<th>Speech Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disgust</td>
<td>Slightly faster</td>
<td>Very much slower</td>
<td>Very much lower</td>
<td>Higher</td>
<td></td>
</tr>
<tr>
<td>Happiness</td>
<td>Faster or slower</td>
<td>Much higher</td>
<td>Much higher</td>
<td>Very much higher</td>
<td></td>
</tr>
<tr>
<td>Sadness</td>
<td>Slower</td>
<td>Very much lower</td>
<td>Much lower</td>
<td>Very much higher</td>
<td></td>
</tr>
<tr>
<td>Anger</td>
<td>Normal</td>
<td>Slightly higher</td>
<td>Slightly lower</td>
<td>Very much higher</td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>Normal</td>
<td>Normal</td>
<td>Normal</td>
<td>Normal</td>
<td></td>
</tr>
</tbody>
</table>

(The emotion of “grief” is omitted.)

Voices parameter changes for five emotions available for the Eloquent TTS system. Speech rate is words per minute (WPM). Average pitch (AP) in Hz. Loudness (G5) in dB.

SCRipting Emotion-based Agent Minds

Emotions are derived from an agent’s beliefs, goals, attitudes, and expressed according to personality trait (agreeableness, extroversion).
SCREAM -- implementation
SCRipting Emotion-based Agent Minds

Emotion Regulation in SCREAM

Ekman & Friesen’s (69) “display rules”
- Expression of emotional state is governed by social and cultural norms, intensity of facial expression

Brown & Levinson (87) on linguistic style
- Assessment of seriousness of Face Threatening Acts (FTAs) considering agent’s desire for autonomy and approval
- Social variables: distance, power, imposition of speech acts
- … avoiding disharmony in conversation (Moulin 98)

Poggi & Pelachaud (01) on reflexive agents
- Hamlet module: to display or not to display (emotions)

J.Gross (98) on emotion regulation in psychology
- Emotion regulation refers to processes by which individuals influence which emotions they have, when they have them, and how they experience and express these emotions. (p.275)
Interface between **MPML** and **SCREAM**

<!--MPML script illustrating interface with SCREAM -->

```mpml
<mpml>
...
<consult target="[...].jamesApplet.askResponseComAct('james,'al','5')>
  <test value="response25">
    <act agent="james" act="pleased"/>
    <speak agent="james">I am so happy to hear that.</speak>
  </test>
  <test value="response26">
    <act agent="james" act="decline"/>
    <speak agent="james">We can talk about that another time.</speak>
  </test>
...
</consult>
...
</mpml>
```

**SCREAM technology** as a scripting language

![Graph showing the relationship between MPML, SCREAM, REA Architecture, JAM, BDI-Agent, Microsoft Agent, XML-based Markup Language, and Power (flexibility)]
SmArt Agent

SmArt Agent’s Faces with Emotion Expressions
Application to 3D Chat with Emotional Expressions

3D Chat
- facial exp. recognition
- agent + chat

Biophysical Emotion Sensors for Affective Interaction

- **Skin-conductivity** (associated with *Arousal*)
- **Heart-pulse rate** (associated with *Valence*)
- **Others**
  - Blood pressure, Temperature, Breath rate,
  - Electocardiogram(ECG), Brain waves(EEG), Electromyography(EMG)
Original Biophysical Emotion Sensing Device with Bluetooth Interface

- We have developed our original emotion sensing device with the Bluetooth wireless interface to detect:
  - SC (Skin Conduct)
  - HR (Heart rate)

- We tested it in the learning process, etc.

Effects appeared as Skin-conductivity

empathic interaction non- empathic interaction
Persona Effect Observation in Empathic and Non-empathic Agents

Emotion Mirror in Virtual Job Interview
Eye-tracker in addition to biophysical sensors for affective interactions

Biophysical sensors (skin conductance) → Main controller PC: Eye-tracking data capture card installed → Eye-mark recorder → (students wear the both)

Analysis from Eye-tracking Data

Our approach vs. approach without agent reaction to eye-tracking

With

Without
English Conversation Training using **MPML** and Character Agents

Towards  **MPML-mobile** version

- Small Display Area, Restricted Behaviors
- Menu Selection Inputs other than Voice or Text Inputs.
- Redesign of MPML Tags.
- Contents generation through Mobile Java Appli.
- MPML-mobile Converter to Mobile Java Appli.
- Small Memory (≤ 100KB) Still big problem
MPML-mobile0.5

Based on Hi-Corp’s Mascot Capsule Engine
(a light-weight 3D modeler for mobile phones)
**MPML-mobile** ➔ **J2ME** on mobile phone

---

**MPML-mobile script** ➔ **MPML-mobile Parser** ➔ **Tag Info.** ➔ **MPML-mobile Controller** ➔ **MIDlet (execution modules)** ➔ **MPML Canvas (chara display)**

---

**MPML mobile version**

```xml
<?xml version="1.0" encoding="shift_jis"?>
<?xml:stylesheet type="text/xsl" href="mpml.xsl"?>
<mpml>
<head>
<title>Hello World!</title>
<agent char="rokey" id="rokey" x="400" y="100"/>
</head>
<body>
<par>
<play id="rokey" act="Nod"/>
<speak id="rokey">Hello World!
You’re ready to proceed</speak>
</par>
</body>
</mpml>
```
MPML-mobile for KDDI-au’s EZ-web, DoCoMo’s i-mode, and Vodafone

Portrait Character Synthesis

- Eigen Vector Analysis
  - Large coefficient components
  - Small coefficient components

emphasis by dignity attachment
At present, the Cartoon-like Portrait SmArt Agents run only on D504i which has a graphic hardware.

Flexibility is not enough at present in Interactive Dialogue

<table>
<thead>
<tr>
<th>Technique used</th>
<th>Example Task</th>
<th>Task Complexity</th>
<th>Dialogue Phenomena Handled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finite-state script</td>
<td>Long-distance calling</td>
<td>Least complex</td>
<td>User asks questions, simple clarifications by system</td>
</tr>
<tr>
<td>Frame based</td>
<td>Getting train arrival/dept. info</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set of contexts</td>
<td>Travel booking agent</td>
<td></td>
<td>Shifts between predetermined topics</td>
</tr>
<tr>
<td>Plan-based models</td>
<td>Kitchen design consultant</td>
<td></td>
<td>Dynamically generated topics, collaborative negotiation sub-dialogues</td>
</tr>
<tr>
<td>Agent-based models</td>
<td>Disaster relief management</td>
<td>Most complex</td>
<td>Different modalities, planed/actual world</td>
</tr>
</tbody>
</table>
Dialogue Control -- switching between In-domain and Out-of-domain (Chatterbot)

Enhancement of Conversational Flexibility through Chatbot technology

**ALICE Chatbot**
(by Richard Wallace, Winner of the 2000 & 2001 Loebner prizes)

**AIMA** (Artificial Intelligence Markup Language)
Auto Presentation with Web Intelligence Functions

- Understand the presentation topic from input query.
- Search the topic in Wikipedia, or Search by Google, Yahoo and AltaVista.
- Text segment summarization (extraction), and associate with relevant outline.
- Generation of a scene-based MPML script with affective support.

The topic is “Big Bang” here.

MPML basic tools are available at

http://www.miv.t.u-tokyo.ac.jp/MPML/
**MPML’s International Publicity**

by Robin Cover at [http://www.oasis-open.org/cover/mpml.html](http://www.oasis-open.org/cover/mpml.html)

Our edited Book published from Springer in 2004
Summary of the Talk

- Background and Related Work
- Overview of MPML
- Various Versions of MPML
- Emotion Expressions (SCREAM)
- An Original Character Agent with Rich Expressions (SmArt)
- Conversational Flexibility
- MPML-VR (virtual reality)
- MPML-mobile
- MPML-HR (humanoid robot)
- Applications (web presentations, entertainments, language learning, etc.)

Current Issues

- More Autonomy
  - Extraction of Emotion from Texts ➔ Emotional Behavior Generation.
  - A Combination of a Chatbot for flexible conversation.
  - Behavior Plan Generation based on the Intention, Goal .... of an Agent.
  - Storytelling.

- Affective Communication
  - Emotion Sensors (face, voice, skin conductivity, ...).
  - Modification of Output Sentences.

- Multimodal Content Business
  - Mobile Contents.
  - Multimodal Educational Contents.
# Acknowledgments to the MPML members

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Helmut Prendinger</td>
<td>Masanori Mitsuzumi</td>
</tr>
<tr>
<td>Hiroshi Dohi</td>
<td>Jun’ichiro Mori</td>
</tr>
<tr>
<td>Santi Saeyor</td>
<td>Du Peng</td>
</tr>
<tr>
<td>Istvan Barakonyi</td>
<td>Naoaki Okazaki</td>
</tr>
<tr>
<td>Sylvain Decamps</td>
<td>Yuan Zhong</td>
</tr>
<tr>
<td>Kyoshi Mori</td>
<td>Zhenglu Yang</td>
</tr>
<tr>
<td>Arturo Nakasone</td>
<td>Takayuki Tsutsui</td>
</tr>
<tr>
<td>Mostafa Al Aasum</td>
<td>Ma Chunling</td>
</tr>
</tbody>
</table>