

EXPRESSIVITY CONTROL

Enabling Emotional Displays and Individualized Actions in Embodied Conversational Agents (ECAs)

Bjoern Hartmann

(b.hartmann@iut.univ-paris8.fr),

Maurizio Mancini,

Catherine Pelachaud,

Vincent Maya, Myriam Lamolle

IUT Montreuil, Univ. Paris 8



Game plan

- 1) Problem definition and goal statement
- 2) State of the Art / Brief overview of our current ECA system
- 3) Theories/techniques considered, adopted, rejected
- 4) Solutions thus far
- 5) **Open Questions**

Problem Definition

- Towards believable, individualized ECAs:
 - **Believability**: multimodal, synchronized behaviors; adaptation to context
 - **Individuality**: set of complex characteristics such as culture, age, personality, profession, gender; but also expression of **emotion**, mood,...
 - We do not model internal characteristics per se, but **focus on their externalizations** – the observable behavior modulations resulting from the sum of the inner processes: automatic **generation** and **modulation** of behavior
 - Two aspects of behavior: **which** actions to select [w/ Maya, Lamolle] and **how** to execute them [this presentation]

Goal

- We want to define a useful **parametric model of behavior modulation** for synthesis and animation of multimodal nonverbal behavior.
 - Find a connection between high-level processes like emotion to low-level animation parameters
 - Current focus: gesture production;
future: multimodal application/integration

State of the Art

- ECA systems

- Animated Conversation
[Cassell et al., 94]
- 'Embodied Conversational Agents'
[Cassell et al., 00]
- 'Lifelike Characters'
[Prendinger, Ishizuka 04]

- Facial animation

- Much work based on FACS
[Ekman, Friesen 78]

- Gesture animation

- Hand and Mind [McNeill92] ,
BEAT [Cassell et al.02],
Max [Kopp02]

- Expressivity

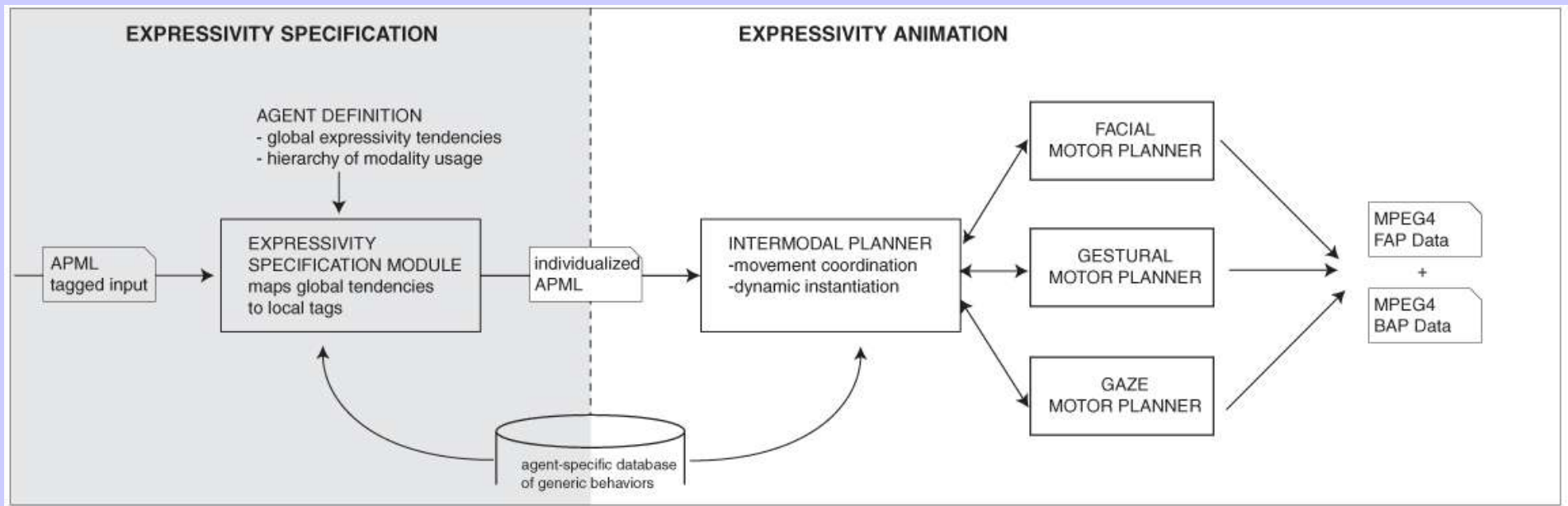
- HMS@UPenn:
EMOTE [Chi et al. 00],
FacEMOTE [Byun, Badler 02]
OCC/Ocean link? [Allbeck,Badler02]
- Pen-based avatar gesture
[Barrientos 02]
- GESTYLE [Noot, Ruttkay 04]
- Tension and relaxation: Physics
simulation [Fiume, Neff, 04]
- Egges...

GRETA Overview

- Engine:
 - **Input:** APML (Affective Presentation Markup Language) annotated text
 - **Output:** Animation as MPEG4 FAP and BAP files + WAV audio
 - **Inbetween:**
 - Extraction of communicative functions (Poggi) from APML
 - Meaning-to-signal mapping for multiple modalities
 - Decision which signals to convey on which channels
 - Adaptation and synchronization of the signals with speech (prerecorded or synthesized)

[video demo I]

GRETA Overview II



+ Inverse kinematics, joint limits, TCB spline and quat slerp interpolation,...

Contributions Considered

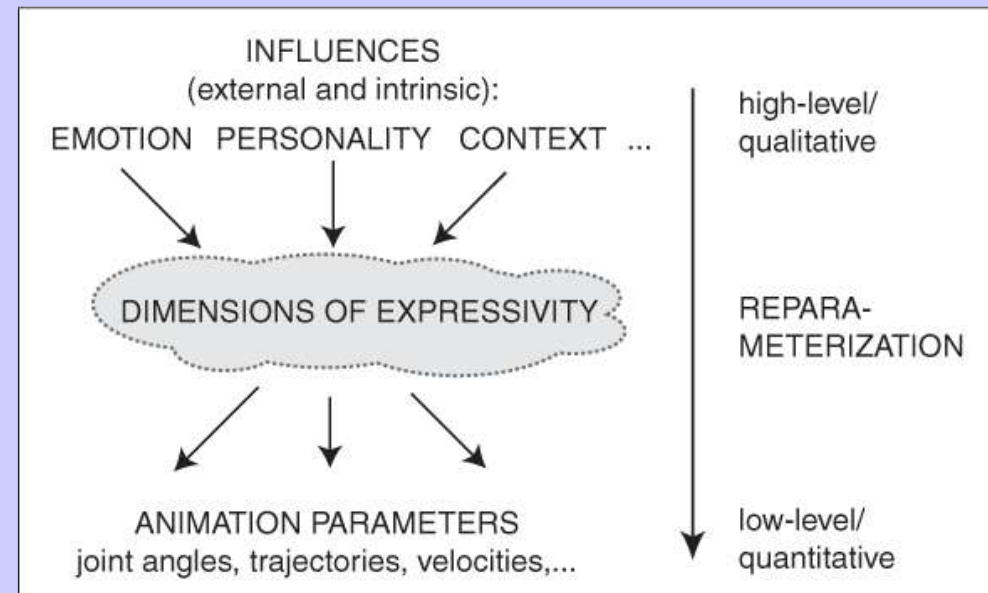
- Psychology
 - 'The expression of the emotions in man and animals' [Darwin 1872]
 - Bodily expression of emotion [Wallbot 98, Wallbott and Scherer 86]
 - Dimensions of personal style [Gallaher 92]
- Dance/Music
 - Laban Movement Analysis / Kinetography [Laban 74]
 - Benesh, Eshkol-Wachman notation systems
 - Eyesweb [Camurri et al.,04]
- Corpus analysis: Form2 [Martell et al., 03]
- Biomechanics?

Problems Encountered

- Granularity / level-of-detail mismatch between existing studies and requirements for animation systems (*we only truly understand a system when we can simulate it*)
- Lack of complete theories on bodily expression of emotion and personality
- Lack of data on **spontaneous**, coverbal gesture

Solutions Found Thus Far...

- Mid-level representation:
'Dimensions of expressivity'
 - **Overall Activation:** quantity of movement across modalities
 - **Spatial:** amplitude of movement
 - **Temporal:** duration of movement
 - **Power:** dynamic property of movement
 - **Fluidity:** smoothness and continuity of movement
 - **Repetitiveness:** tendency to rhythmic repeats



Solutions cont'd...

- Animation parameter mapping for gesture, face
 - *ad hoc / intuitive* at the moment – need more theoretical/experimental support
- Adjustments reach over multiple representation levels
- Tied to semantic features of underlying behaviors

[video demo 2]

Remaining Issues

- How can we link emotion models to our set of expressivity parameters?
- Multimodality:
 - Is it realistic/sound to have a single unified system for treating face, gaze and body (and posture and...)?
 - What are synchronization/conflict resolution functions between modalities? Redundant, complementary, additive, substitutive, masking?
- Alternative parameter/implementation choices (esp. power/tension)?
- Need evaluation studies to test results!