GESTICON

Catherine Pelachaud,
LINC - University of Paris 8
Brigitte Krenn and Hannes Pirker
OFAI

Aims and Purposes

♦ Aims:
  - player and graphics model independent.
  - description of a behavior: not specific to any particular model geometry nor model animation parameterization.

♦ Purposes:
  - Creating gesture shape, facial expression, body posture etc. can be very time consuming.
  - Allow for the mutualisation of work
  - Share behavior definition would greatly help the agent community
Aims and Purposes

Several attempts

- 2002: AAMAS ws « Let’s represent and evaluate ECAs » (C. Pelachaud, T. Rist, Z. Ruttkay, K. Thorisson)
- ½ day ws at AAMAS 2002: B. Krenn, T. Rist
- 2003: 1st Gesticon meeting (B Krenn, C Pelachaud, T Rist, N Thalmann)
- 2005: 1st SAIBA meeting (N Badler, L Jonshon, B Krenn, S Kopp, S Marsella, C Pelachaud, K Thorisson, H Vilhamson)

Aims and Purposes

- Common ideas across these ws:
  - Study existing languages
  - Toward the design of a common Representation Language to control an agent and describe its behavior
Gesticon Representation Language

- Representation language for communicative behaviors only (no action, locomotion, etc)
- Refer to a variety of body behaviors such as facial expression, gaze behavior, head movements, hand-arm gesture, posture as well as combinations of these
- Existence of several dictionaries of emblematic gestures (Posner et al; Poggi; etc)

Gesticon Entry

- Units of nonverbal behaviors
- Hierarchical description of signals across modalities
- Different levels of description
  - very specified to under specified descriptions
  - what is the range from minimal information needed to maximal information depending on players technology
- Gesticon Entry:
  - descriptive part
  - link toward animation file
NOT a Gesticon Entry

- Mapping meaning/communicative functions to signals is done in a separate structure (e.g. table)
- No viseme specification
- Speech: use existing language representation (e.g. SSML)

Overall Structure of a Gesticon Entry

```xml
<GesticonEntry key ="UID" identifier="STRING" />
<verbatim/>
<form/>
</GesticonEntry>
```

- key: unique unique reference to form entry
- identifier: human readable description
Form Element

- `<form>` element: any communicative non-verbal behaviour
- sub-elements:
  - gesture: coordinated movement with arms and hands
  - hand_configuration: hand shape, fingers, orientation of the thumb
  - facial_expression: muscular contraction
  - gaze: eye and head direction (include neck)
  - head: movement of the head independent of eyes
  - upper_body: movement of the spine and shoulder
  - posture: movement of the body elements downward from the hip

Form Element

- Any combination of sub-element possible
- Motivation of decomposition in sub-elements
  - physiology: muscular contraction and joint articulation
  - existence of studies on communicative non-verbal behaviors
  - computational factors: same hand shape used in different arm movements
Gesture Entry

- One or a sequence of basic gesture elements, each of which describing a basic hand-arm movement trajectory
- Phase attribute: preparation, stroke, etc
- Account for different theories: phase attribute is optional (Martell 2002; Bielefeld GeWi)
- Hand description follows the ASL hand shape configuration description
- Movement can be specified in order to allow for the representation of complex hand and finger movements

Movement Element

- Describe the trajectory of any articulated element, i.e. finger, hand, upper body, legs, head
- Two sub-elements:
  - timing: specify both the duration of a movement as well as its temporal alignment in relation to other movements within the gesticonEntry. A minimum, maximum and default duration can be specified in msec.
  - point: one or several keypoint
  - trajectory_type specifies the linear realization of the movement.
- Adoption of SMIL to specify temporal synchrony between sub-elements (seq and par)
Examples of GestureEntry

- **Example of a continuous movement:**
  
  ```xml
  <base_gesture id="ID" phase="stroke">
    <movement trajectory_type="linear">
      <timing/>
      <point wrist_position=""/>
      <point wrist_position=""/>
      <point wrist_position=""/>
      <point wrist_position=""/>
    </movement>
  </base_gesture>
  ```

- **Example of a discontinuous movement:**
  
  ```xml
  <base_gesture id="ID" phase="stroke">
    <movement trajectory_type="linear">
      <timing/>
      <point wrist_position=""/>
    </movement>
  </base_gesture>
  ```

  ```xml
  <base_gesture id="ID" phase="stroke">
    <movement trajectory_type="linear">
      <timing/>
      <point wrist_position=""/>
    </movement>
  </base_gesture>
  ```

  ....

Facial Expression

- Difficulty to describe expression as a small set of parameters
- FACS: not so much used in ECA community, difficulty to become a FACS coder
- Our aim: encodes variability of larger clusters (smile, raise eyebrow, open mouth)
- (optional) decomposition of face into areas: eg based on FACS, Ekman 78
  - forehead_eyebrows, eye_lids
  - nose_cheeks, mouth,
  - chin_jaw, tongue
Facial Expression

- Temporal information: onset/apex/offset; attack/decay/sustain/release
- Temporal specification should cover:
  - static approach: full-blown expression of emotion (Ekman)
  - dynamic approach: gradual display of expression of emotion (Kaiser, Scherer)

Gaze Entry

- Complex element:
  1. only eye direction,
  2. neck, head and eyes showing one direction
  3. neck, head and eyes showing individual directions
- Gaze direction specified through a mechanism for relative reference (a target in real/virtual world) or through spatial area
Head Entry

- Neck and head direction: specified via Gaze Entry,
- Head and neck movement: given by means of a movement trajectory and timing

Upper Body Entry

- Upper_body element:
  - specification for the neck, shoulders, spine and global movements of the torso.
  - The manner attribute of spine: posture of the upper body
    - upright and collapsed
Posture Entry

- The posture element involves the body elements downward from the hip
- Pelvis, hip, legs, (including knee, toes and ankle) are characterized by movement trajectory and duration.
- The specification of legs is analogous to the arms.

SAIBA Framework

- Towards a Common Framework for Multimodal Generation
- SAIBA: Situation, Agent, Intention, Behavior, Animation
Agenda

Tuesday 7.11.2006 (chair Brigitte)
09:30 - 10:00: Gesticon Talk, Catherine
10:00 - 10:30: Basic ideas behind SAIBA/BML (recap IVA presentation), Stefan
10:30 - 11:00: Recap of previous discussions, Hannes Högni
11:00 - 11:30: Coffee break
11:30 - 12:00: Gesticon, Hannes
12:00 - 12:45: General discussion (model, architecture, etc)
12:45 - 14:00: Lunch
14:00 - 16:00: BML discussion (chair Hannes Högni)
   -- synchronization
   -- levels and ways of behavior specifications
   -- interplay with animation engines, how much is due to decisions of a specific animation engine? Experience from SmartBody, Greta, ACE, Nebula, Twente, ...
16:00 - 16:30: coffee break
16:30 - 18:00: Reusable behavior repositories like Gesticon (chair Stefan)
   -- pros and cons
   -- Gesticon: fundamental ideas behind it, what are the problems
18:00 - 18:30: wrap-up & recap, suggestions for day 2
20:30: Social dinner Heuriger

Agenda

Wednesday 8.11.2006
09:00 - 09:30: W3C Emotion Incubator group (Marc Schröder)
09:30 - 11:00: BML - pick up open issues, next steps (chair Hannes)
11:00 - 11:30: Coffee break
11:30 - 13:00: Towards FML (chair Catherine)
   -- what are the major parts/aspects of FML
   -- how flexible must FML be
   -- can we learn something from designing EARL (the emotion representation language) for designing FML
13:00 - 14:00: Cooperation and organizational issues (chair Brigitte)
14:00 - 16:00: lunch and transfer to OFAI
16:00 - ...: meeting at OFAI (if needed)
Conclusion

- To which extent can Gesticon RL be
  - player independent?
  - body independent?
  - animation independent?
- Which constraints need to be specified for such aims?
- Can it use for other tasks (eg annotation of behaviors, behaviors recognition)?