Basic components of a face-to-face interaction with a conversational agent Mutual attention and deixis

Stephan Raidt, Gérard Bailly & Frédéric Elisei
ICP Grenoble

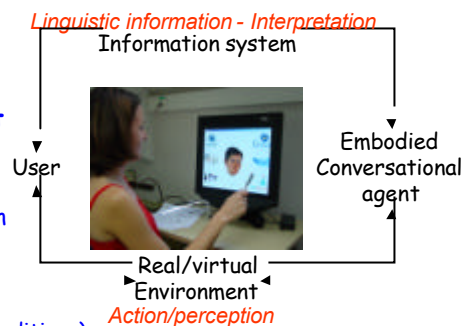Talking Machines - ICP

Framework

- **Presence in AmI environments**
- **Need for enhanced HCI**
    - Input and output
- **Embodied Conversational Agent**
    - Natural face-to-face interaction
    - Delivers high-level information but also give implicit and explicit signs of
        - awareness of
            where (environmental conditions)
            who (human partners) and
            what (activities) is intended in the scene
        - mutual interest and attention when contact is established
    - Keys for grounded interaction, information (precision, trustiness, credibility) and empathy
- **Applications to companion robots, intelligent homes, meeting assistants…**

Linguistic information - Interpretation
Information system

User

Embodied Conversational agent

Real/virtual Environment
Action/perception

Focus

- **Eye-gaze control and scene analysis**
  - – Provides indirect cues to context-awareness
  - – Illustration : scrutinize audiovisual scenes
  - – Implementation: coupling a salience map (saturated colours, movement), a pertinence map (all) & an attention map (WTA scrutation)
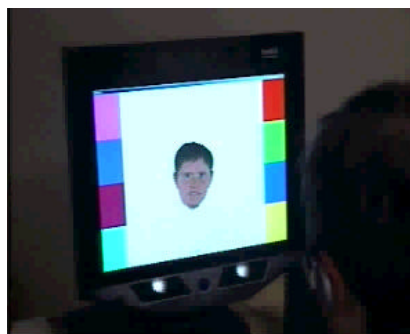
… let you know that I am aware of the environment

Focus

- **Eye-gaze control, its relation to other communication-motivated body movements (head, face, hands) is crucial**
  - – Provides indirect cues to user(s)-awareness and empathy
  - – Illustration: mutual attention & deixis
  - – Implementation: coupling ECA and user's focus of interest via eyetracking
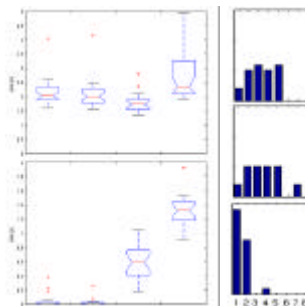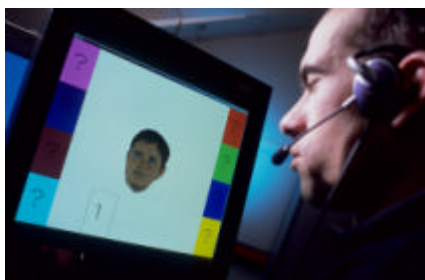
… let you know that I know that you are aware of the environment

Experiments

- **Do our ECA have it? Cues to the direction of social attention**
    - Use case: simple game play
    - Cues : reaction times for target selection, number of scrutated targets
    - Variables: good/bad hints provided by multimodal deixis

- **Answer: Yes… for most users**