

More Than Just a Pretty Face: Conversational Protocols and the Affordances of Embodiment

J. Cassell, T. Bickmore, L. Campell, H. Vilhjálmsón, H. Yan

Gesture and Narrative Language Group MIT Media Laboratory


Overview



- Embodied Human Communication Protocols
- Rea
- Architecture and Implementation
- Conclusions and Future Work



Rea the real estate agent

The background is a solid dark purple color. On the left side, there are several overlapping circles of varying shades of purple, creating a layered, abstract effect. The circles are positioned in the upper-left and lower-left quadrants of the slide.

Embodied Human Communication Protocols



Communication Protocols

- Complexity
 - Particular behaviors can be employed in various circumstances, meaning different things
 - Particular communication function can be expressed through various behaviors

- Conversation contributions can be
 - Propositional (e.g. meaningful speech and hand gesture)
 - Interactional (nonverbal e.g. head nods and “huh”’s)

FEMBOT Model



- **F**: propositional & interactional functions
- **M**: Multi-modal (speech, gesture ,eye gaze...)
- **B**: Separation of function and behavior
 - Simplifies implementation
 - Allows modularity with respect to personality, culture
- **T**: Real-time
 - Attention paid to overall responsiveness
 - Tight temporal synchronization in input and output

Conversational functions and Communicative Behavior



Conversational Functions	Communicative Behavior
<i>Initiation and termination:</i>	
Reacting	Short Glance
Inviting Contact	Sustained Glance, Smile
Distance Salutation	Looking, Head Toss/Nod, Raise Eyebrows, Wave, Smile
Close Salutation	Looking, Head Nod, Embrace or Handshake, Smile
Break Away	Glance Around
Farewell	Looking, Head Nod, Wave
<i>Turn-Taking</i>	
Give Turn	Looking, Raise Eyebrows (followed by silence)
Wanting Turn	Raise Hands into gesture space
Take Turn	Glance Away, Start talking
<i>Feedback</i>	
Request Feedback	Looking, Raise Eyebrows
Give Feedback	Looking, Head Nod

Based on North American nonverbal displays

Conversational functions and Communicative Behavior cont.



- Initiating conversation
 - Turning towards
 - Gazing at
 - Sign exchanging (smile, wave etc.)

- Termination
 - Orientation shifts
 - Glancing away

- Verbal: hi and bye

Conversational functions and Communicative Behavior cont.



- Turn-Taking and Interruption
 - Intonation and gaze
 - Gesture
- Elaboration and Emphasis
 - Gestures: beat, deictic, iconic, metaphoric
- Feedback
 - Head nods, “uh-huh”, “mmm” -
 - Confused face





An Embodied Conversational Agent

- Rea acts like a real estate salesperson
- She is equipped with a graphical body
- Implements conversational protocols based on the FEMBOT model

- **F**: propositional & interactional functions
- **M**: Multi-modal (speech, gesture, eye gaze...)
- **B**: Separation of function and behavior
 - Simplifies implementation
 - Allows modularity with respect to personality, culture
- **T**: Real-time
 - Attention paid to overall responsiveness
 - Tight temporal synchronization in input and output



Rea can...

- ...communicate naturally
- ...sense the user through camera and audio input
- ...speak with intonation
- ...show facial expressions
- ...make head and eye movement
- ...perform gestural output



Rea can...

- ...give and take turn
- ...initiate error correction when she misunderstands input
- Ergo: More than just a pretty face!



Rea – Technical info

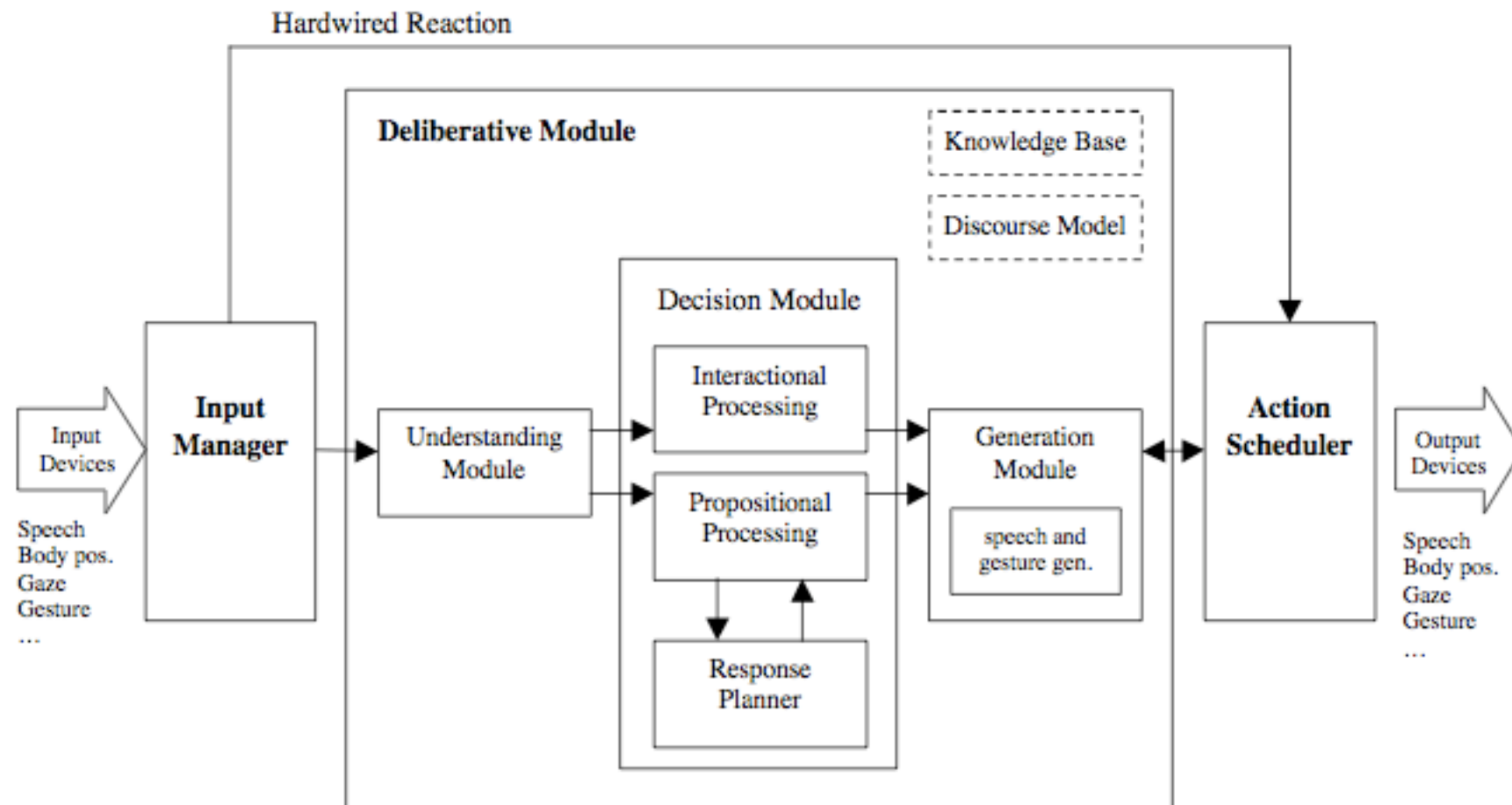
- A single SGI Octane computer for graphics and conversation engine
- Speech recognition and generation along with image processing on several other computers
- NL generation engine for responses – both verbal and gestures
- A simple discourse model
- C++ and CLIPS

The background of the slide is a solid dark purple color. On the left side, there are several overlapping circles in lighter shades of purple, creating a decorative graphic element. The circles vary in size and are arranged in a way that they appear to be layered on top of each other.

Architecture and Implementation



Overall architecture



Communication Protocols Implementation



- Initiating conversation
 - 1 Vision system detects a user
 - 2 Rea turns face-to-face to the user
 - 3 Exchanges verbal and nonverbal greetings
- Termination
 - 1 Vision system detects the user turning away
 - 2 Rea suspends speech input processing
 - 3 Exchanges verbal and nonverbal farewells

Communication Protocols Implementation cont.



- Turn-Taking and Interruption
 - Verbal interruption detected => Rea yields the turn
 - Gestures from user => Rea halts at the nearest sentence boundary
 - She looks away when planning a response



Elaboration and Emphasis

- Rea uses iconic, metaphoric and deictic gestures as well as beats
- She uses the following to choose gesture:
 - Semantics
 - Structure (rheme vs. theme)
 - Focus
 - Visibility

Elaboration and Emphasis cont.



- Gestures support rhematic material: new visible entity => deictic
- Semantic content mapped into iconic or metaphoric gesture
- Annotated utterance is the output of the text generator
- Scheduling module then processes the data and prepares it for the animation

Communication Protocols Implementation cont.



- Feedback and Error Correction
 - Rea nods her head at the end of users utterances
 - She ask clarifying questions if she doesn't understand
- Rea in action

The background of the slide is a solid dark purple color. On the left side, there are several overlapping circles of varying sizes and shades of purple, creating a decorative graphic element. The text "Conclusion and future work" is centered on the right side of the slide in a white, sans-serif font.

Conclusion and future work



Conclusions

- If embodied interfaces leverage knowledge of human communication behavior they provide a qualitative advantage over non-embodied interfaces



Future work

- Increasing symmetry of input and output
- A sensor measuring head movements and eye gaze
- Implementing a larger set of conversational functions

The image features a dark purple rectangular background with rounded corners. On the left side, there are several overlapping circles in various shades of purple, ranging from light to dark. The text "Thank you" is positioned on the right side of the rectangle in a white, sans-serif font.

Thank you