Deciding When To React To Incremental User Input In Human-Robot Interaction

Kathrin Janowski, Elisabeth André

Human Centered Multimedia
Institute of Computer Science
Augsburg University
Universitätsstr. 6a
86159 Augsburg, Germany
Motivation

input becomes more natural ↔ interaction flow stays robotic

„Move the green piece to field three.“

„The triangle.“

„No, to field three.“

„Well done.“

[looks at green objects]

„The triangle or the square?“

[moves triangle to field four]

[moves triangle to field three]
Motivation
Motivation

goal: react as early as possible
   → avoiding delays
   → intercepting mistakes

„Move the green piece to field three.“
„The triangle.“
„No, to field three.“
„Well done.“

[looks at green objects]
[moves triangle towards field four]
[changes path to field three]
Problem: Uncertainty

- many different interdependent factors and goals
- situation only partially observable
- different positive or negative effects

„Virtual butler: What can we learn from adaptive user interfaces?“

*Conati (2013)*
Problem: Uncertainty

„Put the green object on field three.“

„The triangle.“

„Oh, okay.“

„Put the green object on field three.“ [looking at objects]

„The triangle.“

„Do you m...“

„Huh? Pardon?“
Problem: Uncertainty

„Put the green object on field three.“ [looking at objects]

„The triangle.“

„Yeah, I know!“ [looking at robot]

„Hello?“
Approach: Influence Diagram

- Bayesian Network
  - modeling dependencies between factors
  - inferring probabilities from partial observations

- Utility Nodes
  - modeling effects for different situations
  - trade-off between costs and benefits
Example – Observation
Example – Utilities

user comfort =
  + helpfulness(action)
  + is_mistake(target)
  + help_wanted(certainty)
- obtrusiveness(action)

time saved =
  + rollback(manipulation, target)
  + avoided_hesitation(certainty)
Example – Inference
Example – Decision
Flexibility

- draw information from any available channel
- use interpretations on any abstraction layer
  - low level raw data at the beginning
  - high level semantic data near the end

[gaze at triangle] [at square] [back at triangle]

„Do you mean the green triangle?“
Changes over time

- utility scales with confidence in interpretation
  - more reliable data $\rightarrow$ utility increases
  - conflicting data $\rightarrow$ utility decreases

- implicit priorities
  - uncertain interpretation
    $\rightarrow$ prefer actions with low benefit but also low risk
  - certain interpretation
    $\rightarrow$ prefer actions with high benefit
Ongoing Work

• general framework for sensor processing
  • „The Social Signal Interpretation (SSI) Framework: Multimodal Signal Processing and Recognition in Real-time“
    *Johannes Wagner, Florian Lingenfelser, Tobias Baur, Ionut Damian, Felix Kistler, and Elisabeth André (2013)*
  • „Modeling multimodal integration with event logic charts“
    *Gregor Ulrich Mehlmann and Elisabeth André (2012)*

• interactive system for data collection
  • adding different modalities
    (speech, eye tracking, gestures, tangible interaction...)
  • creating different situations
    (eg. grounding problems)