# From autonomy to interaction - A pladoyer for a paradigm shift

Britta Wrede

Applied Informatics

# Overview

- 3 Misconceptions of Robotics
- Some examples of how interaction influences learning
- Some Implications

# Misconception 1: Autonomous Learning

**„Robots should learn autonomously"**

But

- Human infants learn through interaction
- Some actions can not be learned through observation and imitation [Csibra & Gergely, 2003]
- The meaning of an action in terms of goal, means and restrictions needs to be communicated by a tutor

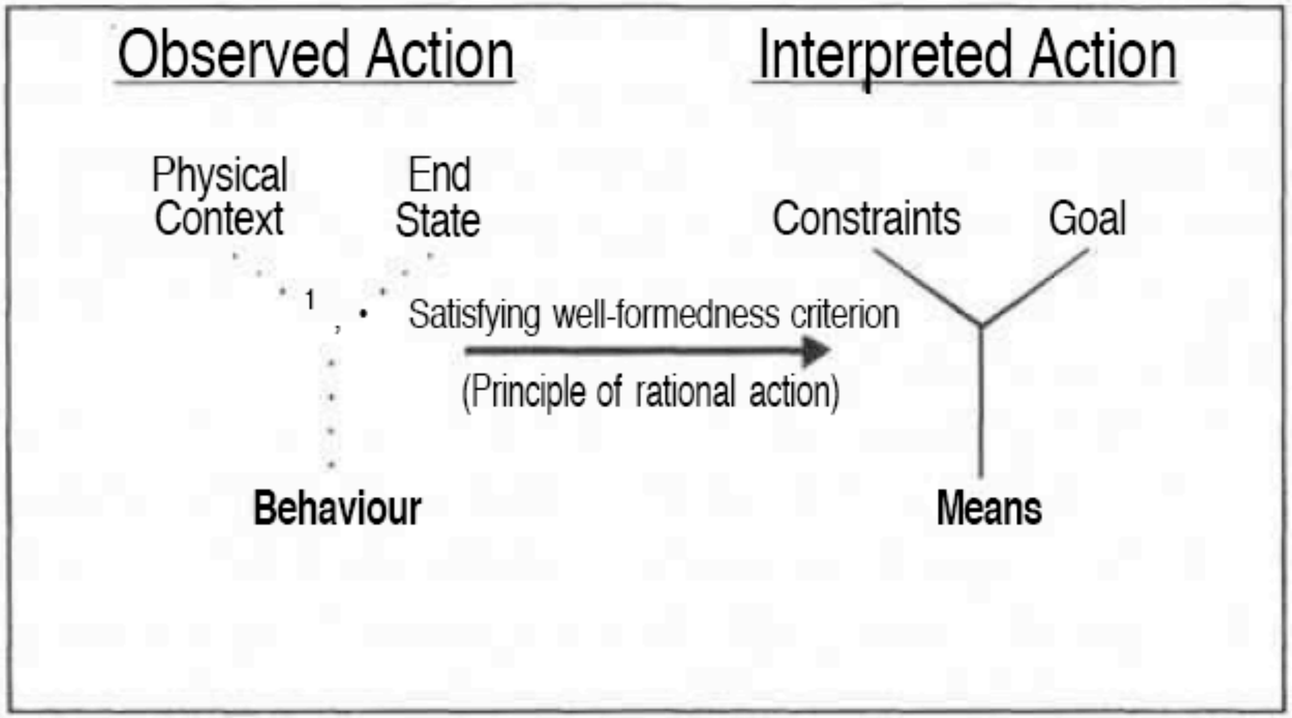
# How to draw inferences about other‘s goal directed actions? [Csibra & Gergeley 2009]

Children

- primarily **imitate causally efficacious means** to achieve goals,
- **ignore** apparently unnecessary actions
- unless the **demonstrator makes it manifest** for them that these cognitively opaque aspects are relevant

⇒ ostensive behavior of tutor important

# How to draw inferences about other's goal directed actions? [Csibra 2003]




Assumption (well-formedness criterion):

Observed behavior

– will bring about goal state

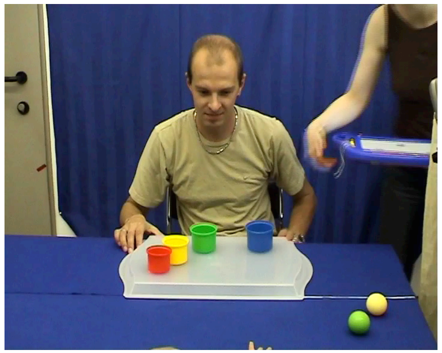
– is most efficient means to reach goal

# Support for Interpreting Actions in Infant Directed Actions

Constraints

Goal

Means



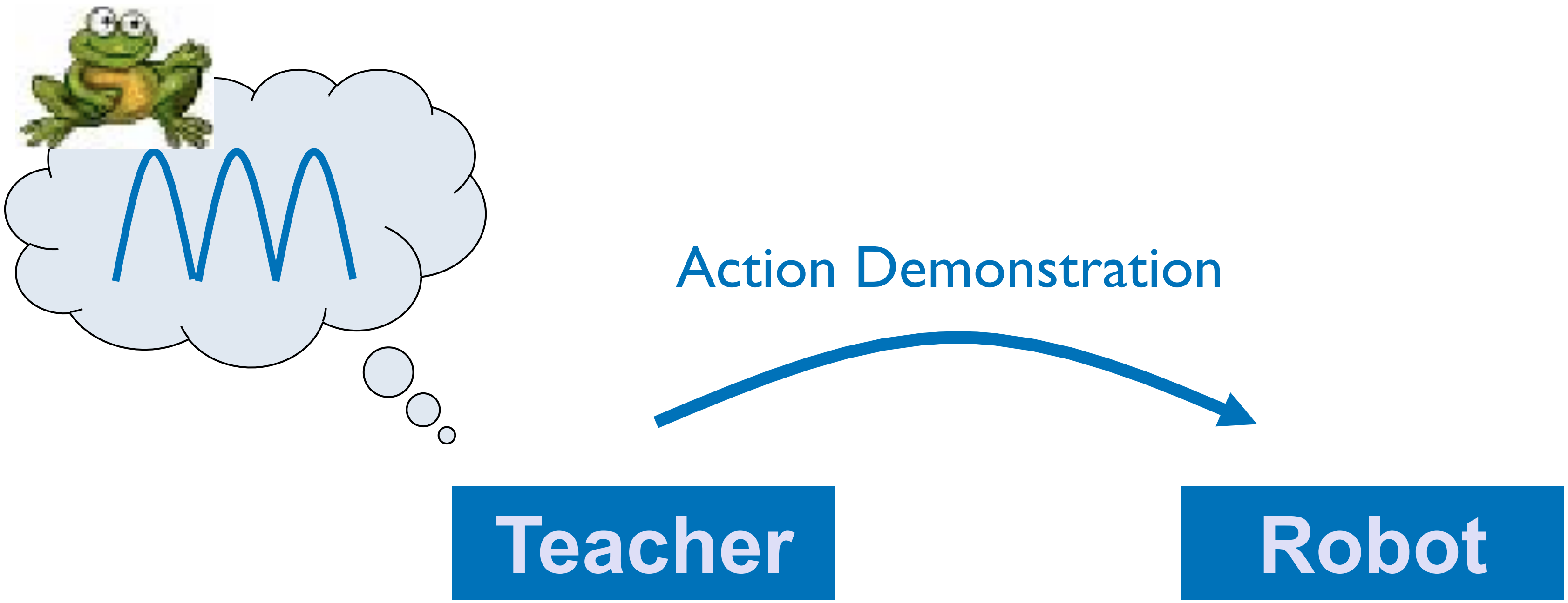
# Misconception 2: stable representation = static representation

**„A robot needs a stable representation in order to act in its environment“**

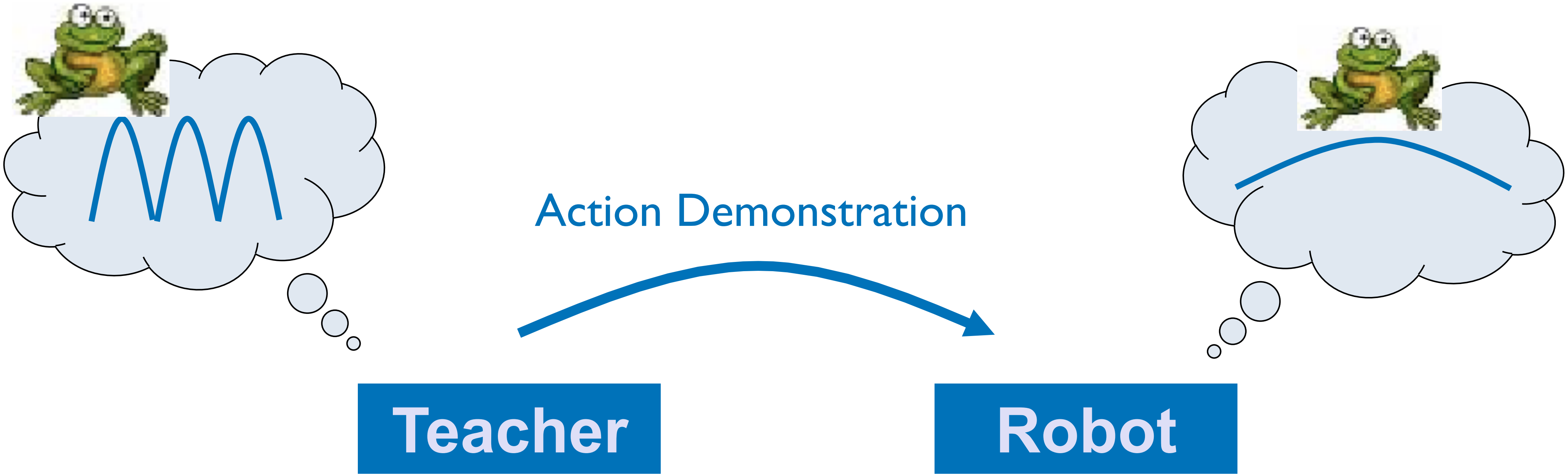
But

- Representations are emergent and shaped through interaction
- E.g. consider teaching a robot an action

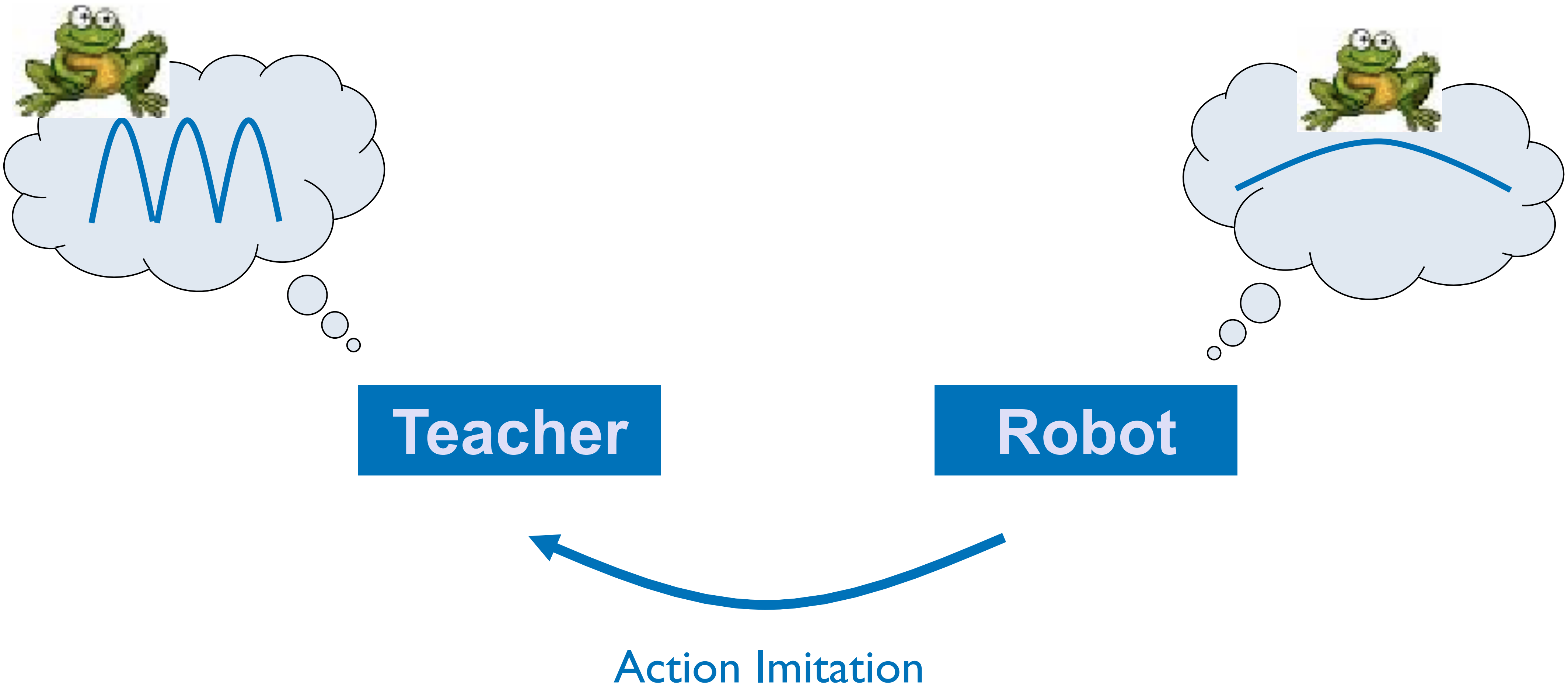
# Adapting Representations during Tutoring



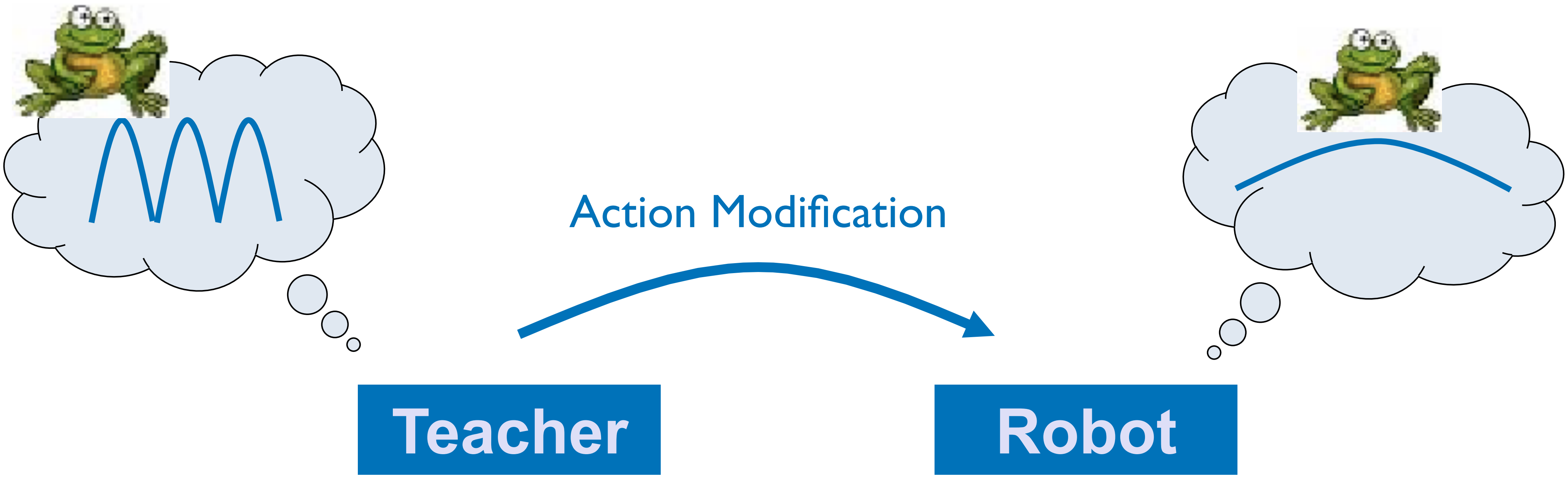
# Adapting Representations during Tutoring



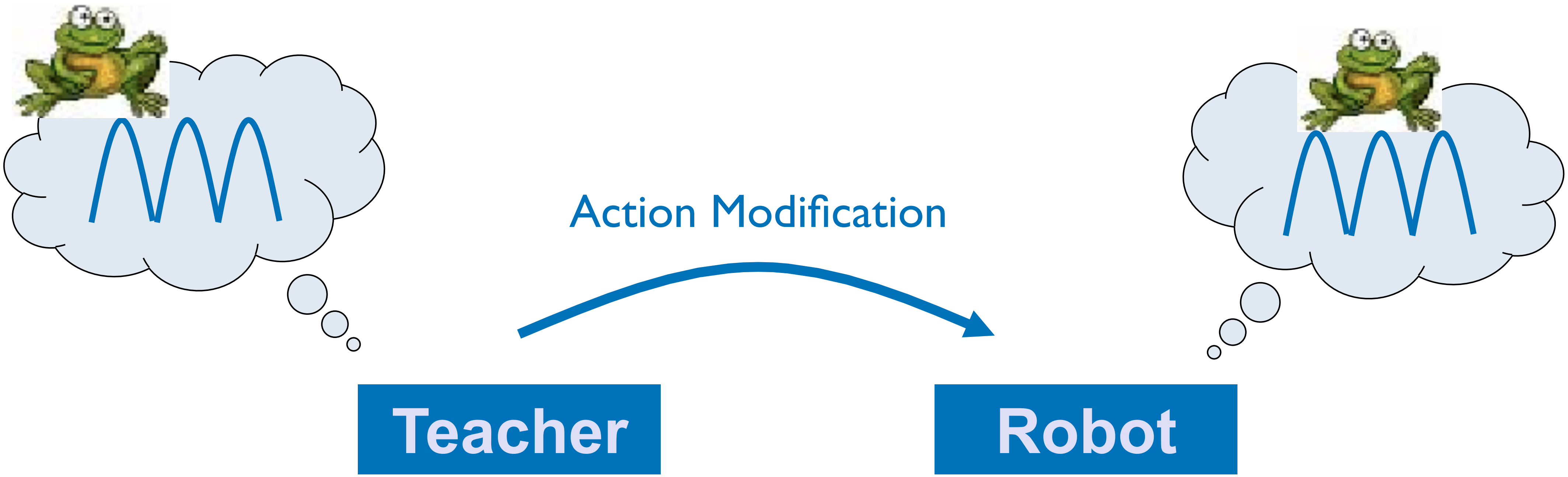
# Adapting Representations during Tutoring



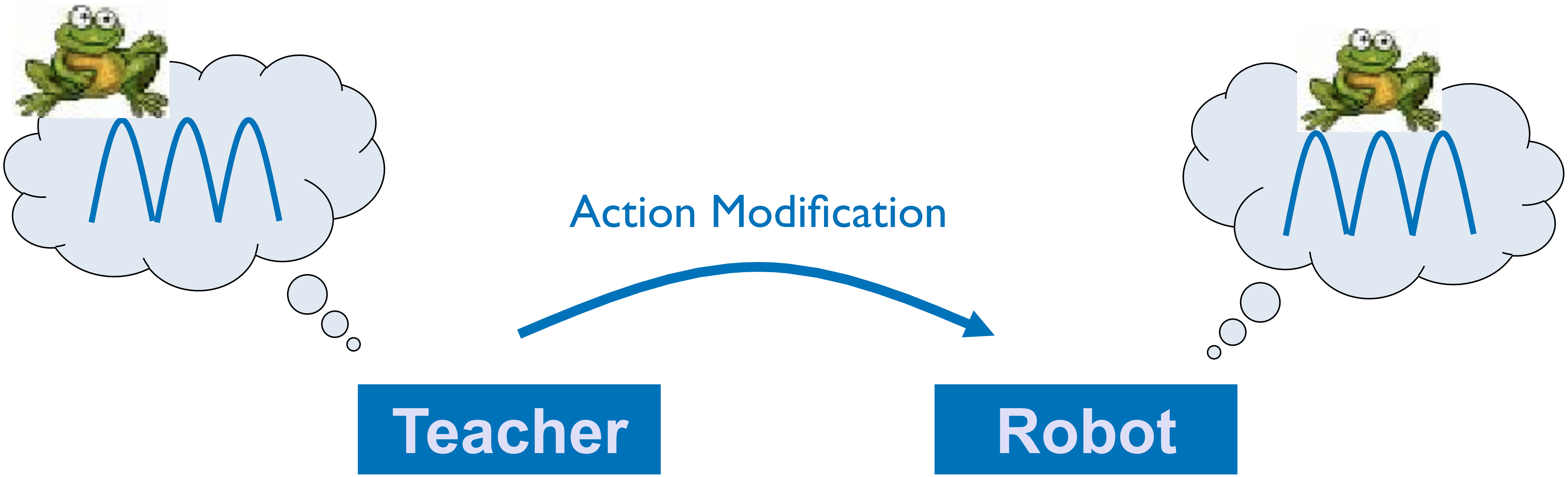
# Adapting Representations during Tutoring



# Adapting Representations during Tutoring



# Adapting Representations during Tutoring




- stable representation may produce variable action demonstrations
- learner‘s representation needs to be flexible

# Misconception 3: Interaction is for giving commands

**„Interaction is not necessary but may be used for giving commands“**

But

- Interaction is bi-directional
- Feedback of the learner is important

# Analysing Infants' Feedback and its effect on the tutor's behavior


Anna-Lisa Vollmer
Karola Pitsch

Feedback of infants changes with infant's age (and capabilities)

Pre-lexical infants:
- Gazing behavior displays the infant's state of attention

⇒ Tutor attracts attention by e.g. waving
⇒ Tutor exaggerates movement

Early lexical infants:
- Anticipate next actions with the direction of gaze

⇒ Tutor's movements not exaggerated

Lexical infants:
- Give systematic feedback according to the structure of the action including instructions for the tutor's next actions

⇒ Tutor changes behavior

# Interaction is important for learning

- Learning needs to be both: autonomous and interactive
- Representations need to be both: stable and flexible
- Feedback of the learner influences how the next demonstration will be carried out
- Learner needs to be sensitive to tutor‘s ostensive signals

# Overview

- 3 Misconceptions of Robotics
- Some examples of how interaction influences learning
- Some Implications

# Overview

- 3 Misconceptions of Robotics
- Some examples of how interaction influences learning
  - Tutoring
  - Acoustic Packaging
- Some implications

# Tutoring

Tutoring plays an important role in learning [Csibra & Gergely, 2009]:

- Learners are sensitive to tutoring cues
- Learners prefer tutoring behavior (child-directed speech & action)
- Tutoring enhances learning

Quantitative Analyses:

- What are the characteristic modifications in infant-directed tutoring and how can they be used for learning?
- Are these modifications similar in robot-directed tutoring?
- How to benefit from these modifications?

# What is Tutoring?

Ostensive cues that are characteristic for teaching:

- Contingency
- Motherese
- Motionese

# Contingency

Reciprocity between physical events or social events

**What aspects does contingency contain?** [Watson, 1984]

**Temporal Contingency**

Sensitivity to reciprocal responsiveness and coordination of temporal parameters

**Spatial Relation**

Sensitivity to covariation of place of behavior and place of effect

**Sensorial Relation**

Covariation between magnitude of sensory effect of own behavior (proprioeption) and sensory consequences

# Innate Contingency Detection Module

[Gergely & Watson, 1999]

**How to perceive stimulus-response contingencies effectively?**

Sufficiency Index (SI)

P( **stimulus** (=perceived effect) | **response** (=own action))

P( moving_mobile | movement_of_right_leg)

Necessity Index (NI)

1 – P ( **stimulus** | **no_response** )

1 – P (moving_mobile | no_movement_of_right_leg)

Learning causality or relatedness:

If SI=1 and NI=1 then perfect contingency (= perceived causality)

If NI > SI, then reduce response class (e.g. right leg instead of both legs)

If NI < SI, then expand response class (e.g. both legs instead of right leg)

# Contingency

**Perfect Contingency**

Detection of Causality, e.g. learning of body schema

**Nearly Perfect Contingency**

Detection of Social Interaction

**Contingency during Infant Development**

2 months: preference of perfect contingency

3 months: preference of imperfect contingency (not in autistic children)

[Bahrick & Watson, 1985]

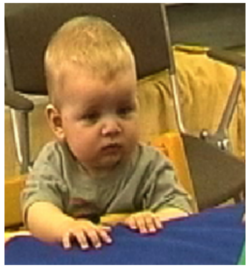
# Tutoring


Anna-Lisa Vollmer

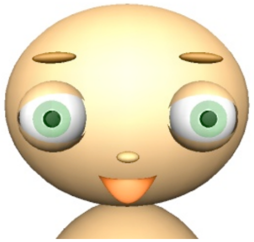

Katrin Lohan

## Motionese (velocity, pace, roundness and range of hand trajectories)

<



<

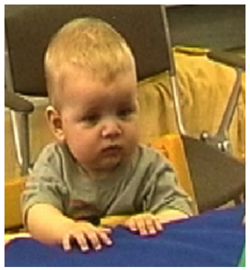


- Hand movements are **slower, less round and have more range** in infant-directed tutoring
- **Motionese strongest in ARI**

## Contingency (nmb and length of eye-gaze bouts)

<

<



- More contingency in infant-directed tutoring
- **Contingency weakest in ARI** (impaired!)

**Feedback-behavior of robot relevant for tutoring behavior**

# Overview

- 3 Misconceptions of Robotics
- Some examples of how interaction influences learning
    - Tutoring – Detect ostensive cues in order to determine **when** to imitate
    - Acoustic Packaging
- Some implications

# Acoustic Packaging [Brand et al., 2007]

- Children need to discover meaningful action units
  - Language helps to divide a sequence of events into units
  - Prerequisite: synchrony between language and events

- Described as **acoustic packaging** (AP) [Hirsh-Pasek and Golinkoff 1996]

- AP can provide a bottom-up action segmentation

# Analysis of Parent-Infant Interaction

**Characteristics of child-directed actions (Motionese)**

- Trajectories more straight, less smooth, e.g., higher arches
- Modulations at trajectory onset
- Lower velocity

**Characteristics of Child-Directed Speech (Motherese)**

- More structure
  - More and longer pauses
  - Different intonation patterns
- Repeated checking of learners attention

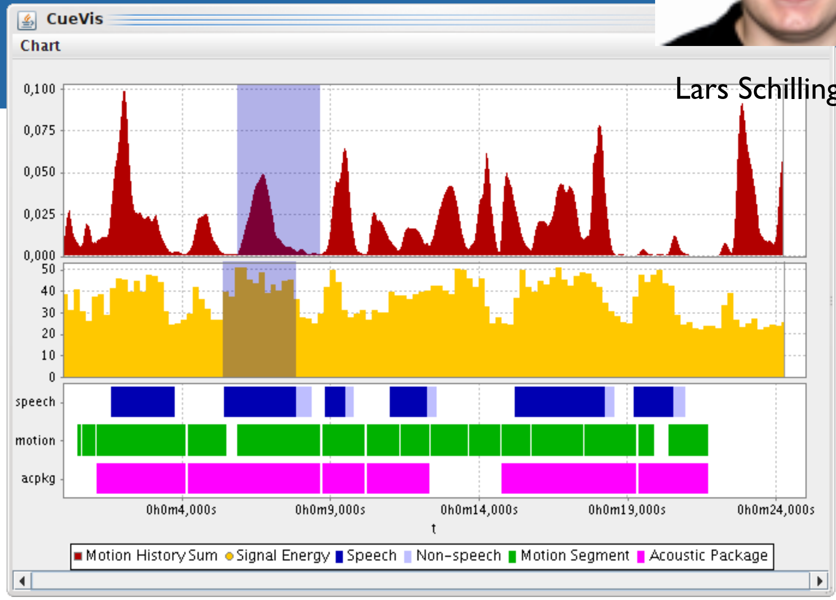
Universität Bielefeld

CoR-Lab – Research Institute for Cognition and Robotics

CoR-Lab

# A Computational Model of Acoustic Packaging

Lars Schillingmann

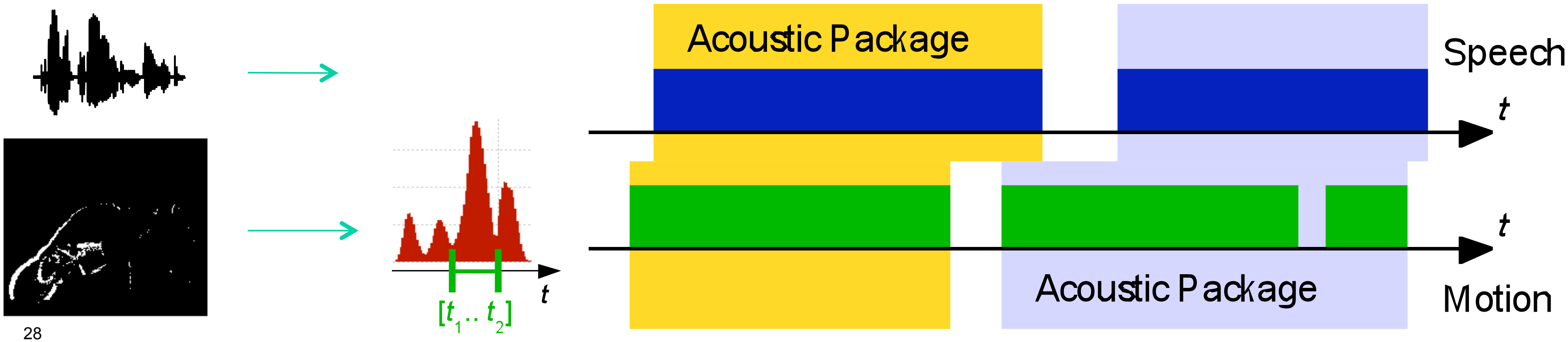



## Segmentation of input cues

- Acoustic temporal segmentation
- Visual temporal segmentation

## Cue fusion

- Temporal association of multi-modal input streams
- Results of the association process are Acoustic Packages



# Evaluation

- Videos from Motionese corpus (11 AAI, 11 ACI) and from babyface study (11 ARI)

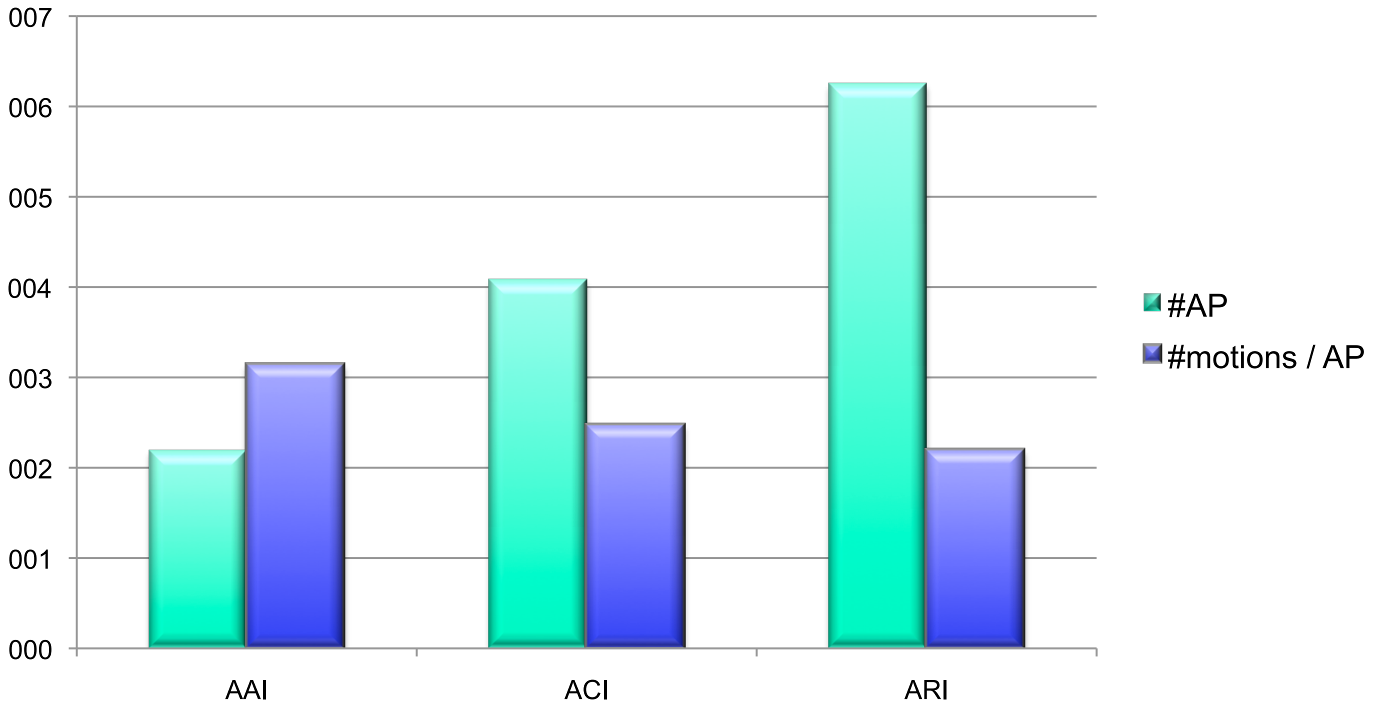
**Analysis**

- Automatic detection of Acoustic Packages
- Measurements:
    - **number of Acoustic Packages (#AP)**
    - mean number of **motions per Acoustic Package (#motions / AP)**

**Hypothesis**

- ACI more structured than AAI
- **More #AP and less #motions / AP in ACI**

# Results




- Sig. more Acoustic Packages in ACI and ARI
- Sig. less Motions per Acoustic Packages in ACI and ARI

# Goals and Challenges

Evaluation showed AP is able to reflect differences between adult-adult and adult-child interaction

**Goals**

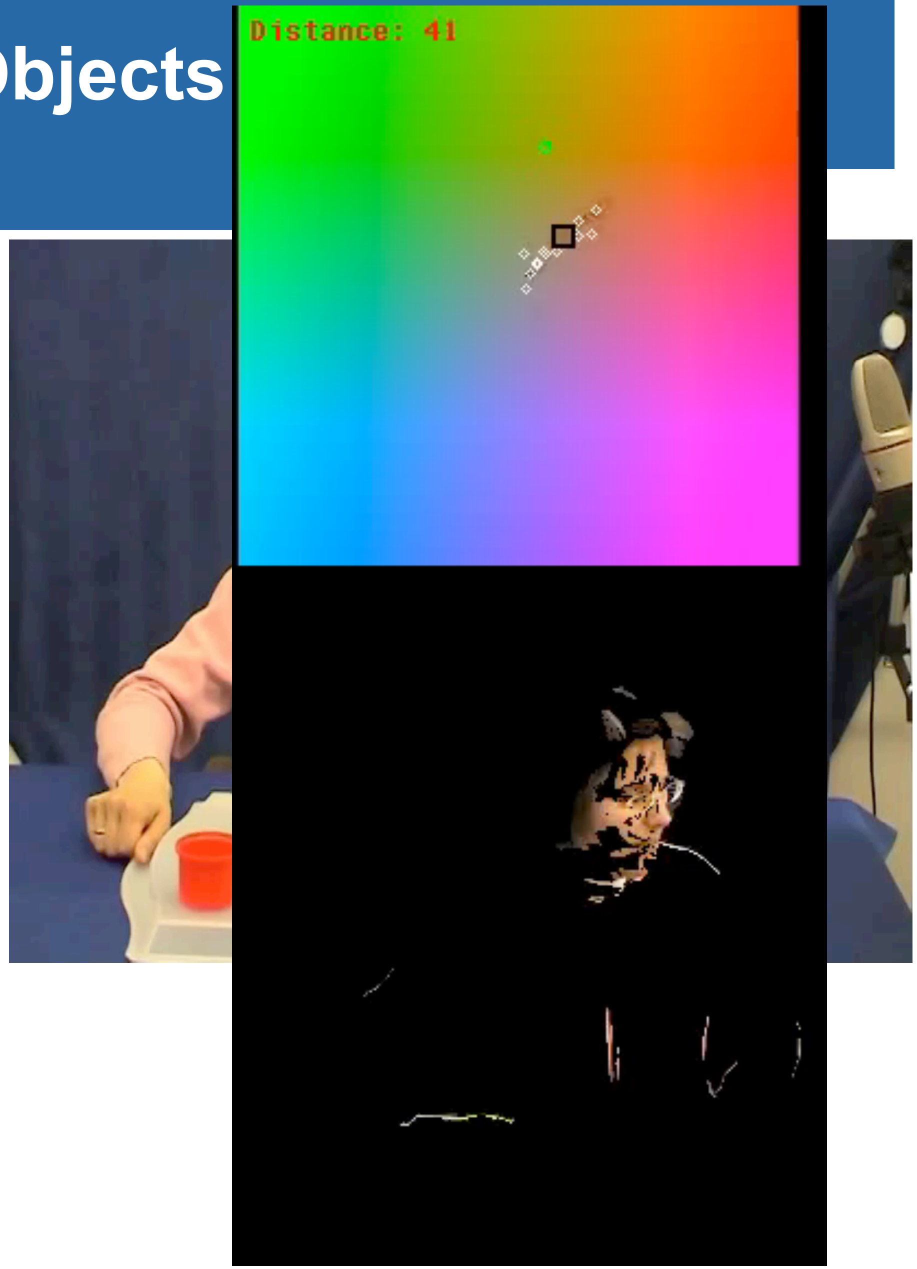
- Generating feedback events
- Providing learning units for further processes

**Challenges**

- How to generate Feedback?
- What is visually interesting?
- What is acoustically interesting?
- How to discriminate human motion against object motion?

# Detecting Moving Colored Objects

- Detecting changing regions
  - Masking delayed image with Motion History Image
  - Labeling
- Clustering in YUV color space
- Ranking according to color distance  (U,V) to centroid of all clusters
- Heuristical filtering
  - Detecting background by region growing on current frame
  - Skin color filtering
  - Deviation of Pixels
  - Density: Pixels / Variance ratio
- Trajectory accumulation
  - Multiple hypothesis
  - Ranked by average color distance to centroid



# Acoustic Prominence

Relative ranking of syllables within an utterance

Syllable Segmentation

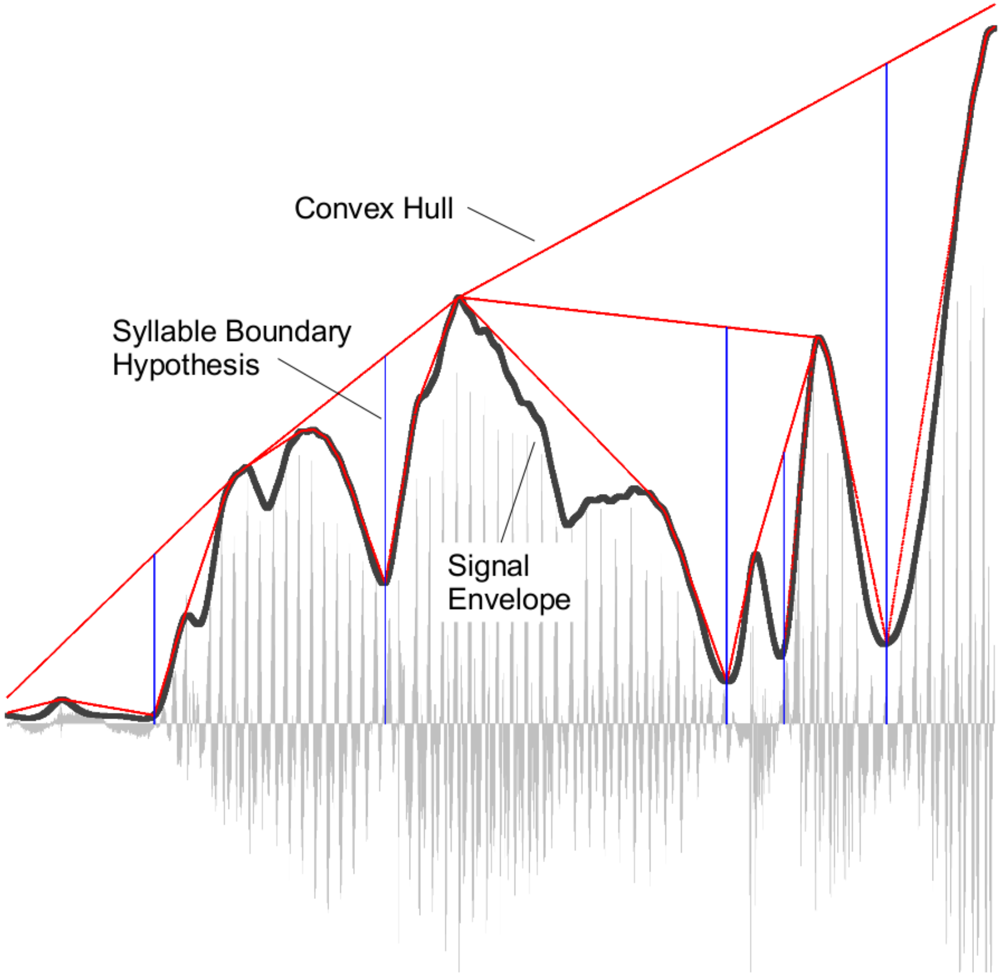
- Mermelstein algorithm

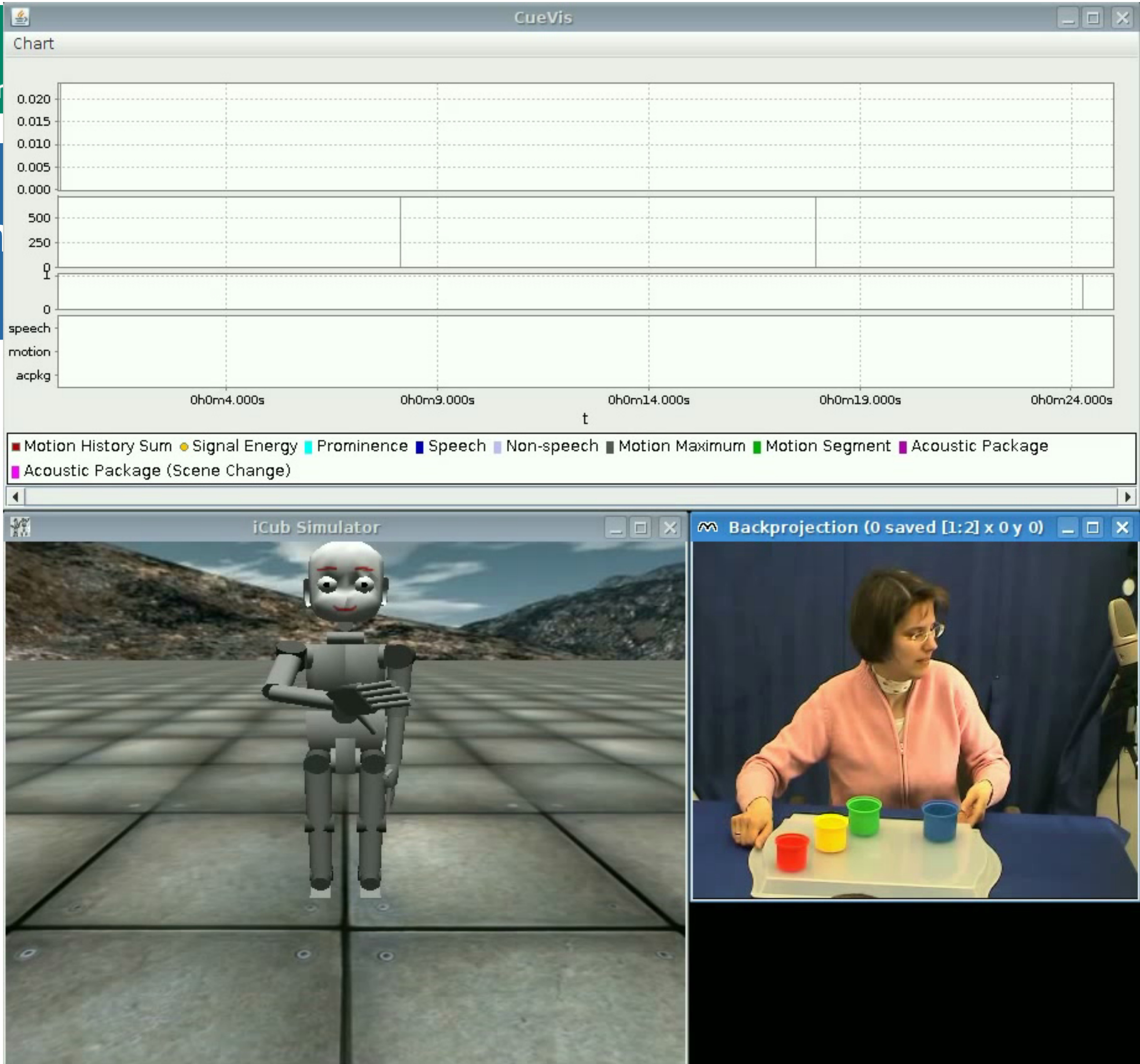
Features [Tamburini, Wagner 2007]

- Nucleus duration
- Spectral emphasis
- Pitch movements
- Overall intensity

Currently used: Spectral emphasis

Examples with 3 syllables context




CueVis
Chart
0.020
0.015
0.010
0.005
0.000
500
250
0
1
0
speech
motion
acpkg
0h0m4.000s
0h0m9.000s
0h0m14.000s
0h0m19.000s
0h0m24.000s
t
Motion History Sum
Signal Energy
Prominence
Speech
Non-speech
Motion Maximum
Motion Segment
Acoustic Package
Acoustic Package (Scene Change)
iCub Simulator
Backprojection (0 saved [1:2] x 0 y 0)

# Overview

- 3 Misconceptions of Robotics
- Some examples of how interaction influences learning
  - Tutoring – Detect ostensive cues in order to determine **when** to imitate
  - Acoustic Packaging – Detect stucture in demonstration to determine **what** to imitate
- Some implications

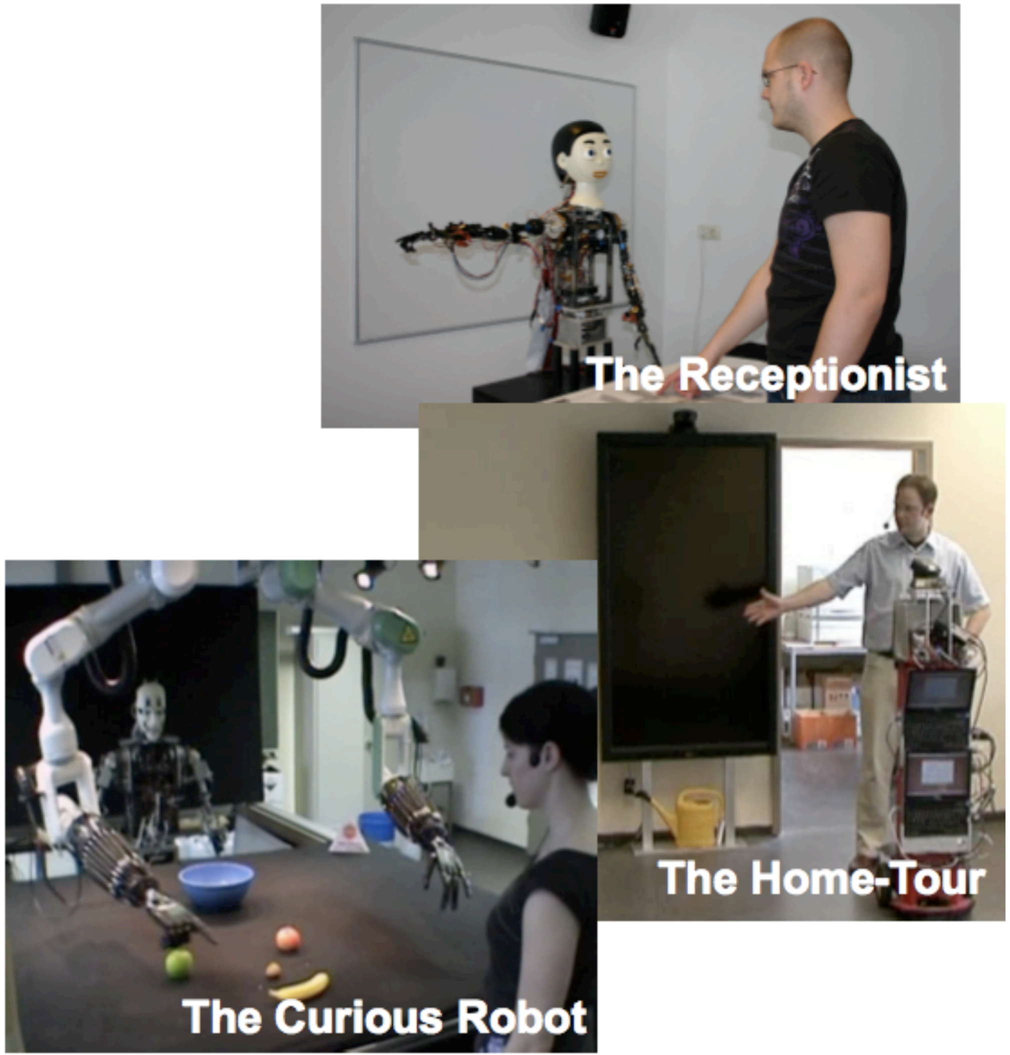
# Overview

- 3 Misconceptions of Robotics
- Some examples of how interaction influences learning
  - Tutoring – Detect ostensive cues in order to determine **when** to imitate
  - Acoustic Packaging – Detect stucture in demonstration to determine **what** to imitate
- Some implications
  - Modeling Dialog on a Robot in order to provide feedback

# Dialog Modeling for Robots: The PaMini Framework

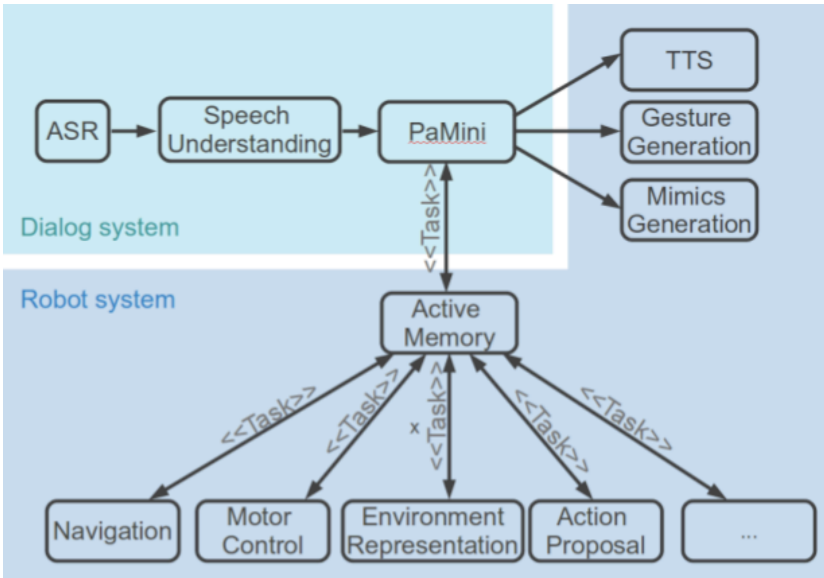
Julia Peltason

- Concepts
  - Modeling Robot Tasks: Task State Protocol
  - Modeling Dialog States: Interaction Patterns
- Usability Test



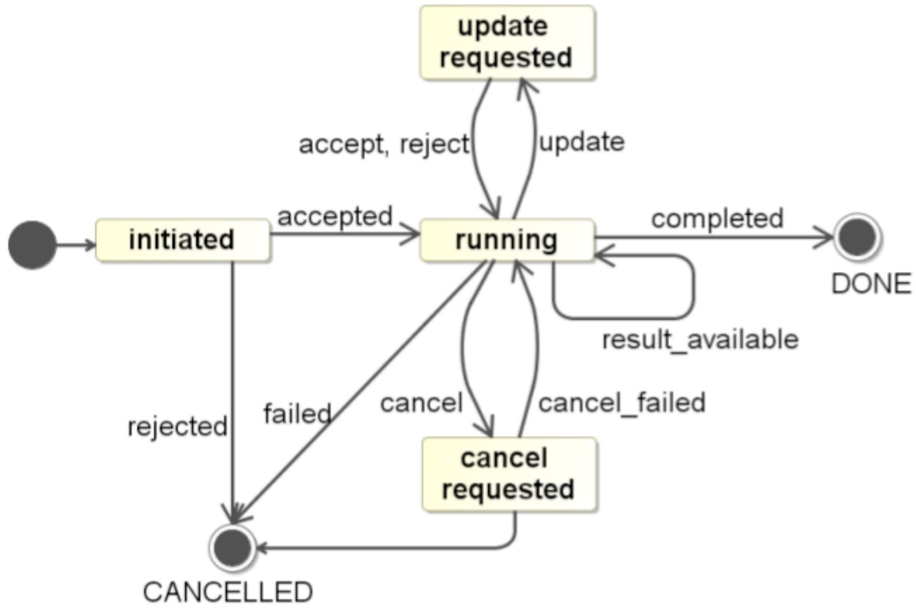
# The Task-State Protocol

- Complex application back-end on robots
  - Multiple compoments
  - Temporally extended actions
  - ...that may fail
- A uniform interface for coordination needed



# The Task-State Protocol

- Fine-grained interface to robotic subsystem
  - Tight integration of action and perception
  - Basis for verbalizing the robot's actions and internal state
- Supports task update during execution
- Gives the robot the ability to react to comments and corrections on-the-fly.



# Interaction Patterns

Dialog modeling on robots often relies on simple command-control techniques.

- Roboticists want to build HRI scenarios,
- but they do not want to bother about subtleties of dialog modeling

**Interaction Patterns**

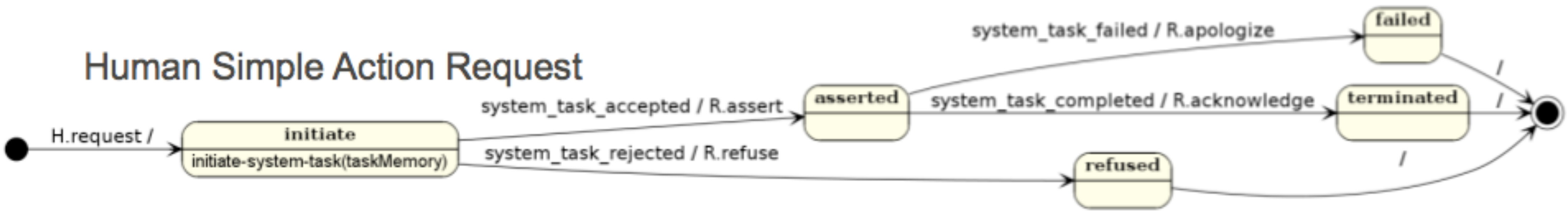
- describe recurring conversational structures
- provide configurable building blocks of interaction
- support rapid prototyping

# Interaction Patterns

Transducer-like notation

- Input: human dialog acts or task events
- Outputs: robot dialog acts
- Actions: task and variable updates

Defined at an abstract level and configured for specific situations



# Usability Evaluation: Do interaction patterns ease dialog modeling?

- Usability test: performance measure + think-aloud
- 2 groups: roboticists, non-roboticists
- 5 tasks with increasing complexity; 1 hour time-limit

| Task | Description | # DA | Challenge | Interaction Pattern |
|---|---|---|---|---|
| 1 | Greeting | 2 | | Interaction Opening |
| 2 | Parting | 2 | | Interaction Closing |
| 3 | Following | 11 | Task communication | Cancellable Action Request |
| 4 | Low Battery warning | 1 | Task communication, Variable definition, Parameterized output | Notification |
| 5 | Acquire name | 6 | Task communication, Variable definition, Task update | Correctable Information Request |

# Usability Evaluation: Do interaction patterns ease dialog modeling?

## Results & Observations

- All participants were able to solve task 1-3
- Half of participants were able to solve task 5 partially
- Roboticists slightly faster
- Steep learning curve
- Roboticists rely on task events,
  Non-Roboticists rely on dialog acts

# Conclusion

Paradigm shift towards interaction necessary:

- Autonomy neccessary but not sufficient for learning
- Learning implies interacting

Implications

- Dialog modeling on robots requires abstraction from action and dialog steps
- System integration necessary

# Merci beaucoup!