

An Event-Based Conversational System for the Nao Robot



Ivana Kruijff-Korbayová, Georgios Athanasopoulos, Aryel Beck, Lola Cañamero, Piero Cosi, Heriberto Cuayáhuitl, Tomas Dekens, Valentin Enescu, Racquel Ros Espinoza, Antoine Hiolle, Bernd Kiefer, Giulio Paci, Stefania Racioppa, Hichem Sahli, Giacomo Sommavilla, Fabio Tesser, Werner Verhelst, Weiyi Wang

Aims

Build a long-term adaptive human-robot conversational system for interacting with hospitalized childen

- Robotic companions for child-robot interaction
- Robot has to handle multiple inputs and outputs simultaneously
- Parallelism as well as Synchronization necessary ⇒
 event-based component integration
- Distributed processing with flexible communication streams



Integrated System

Centralized Event Handler

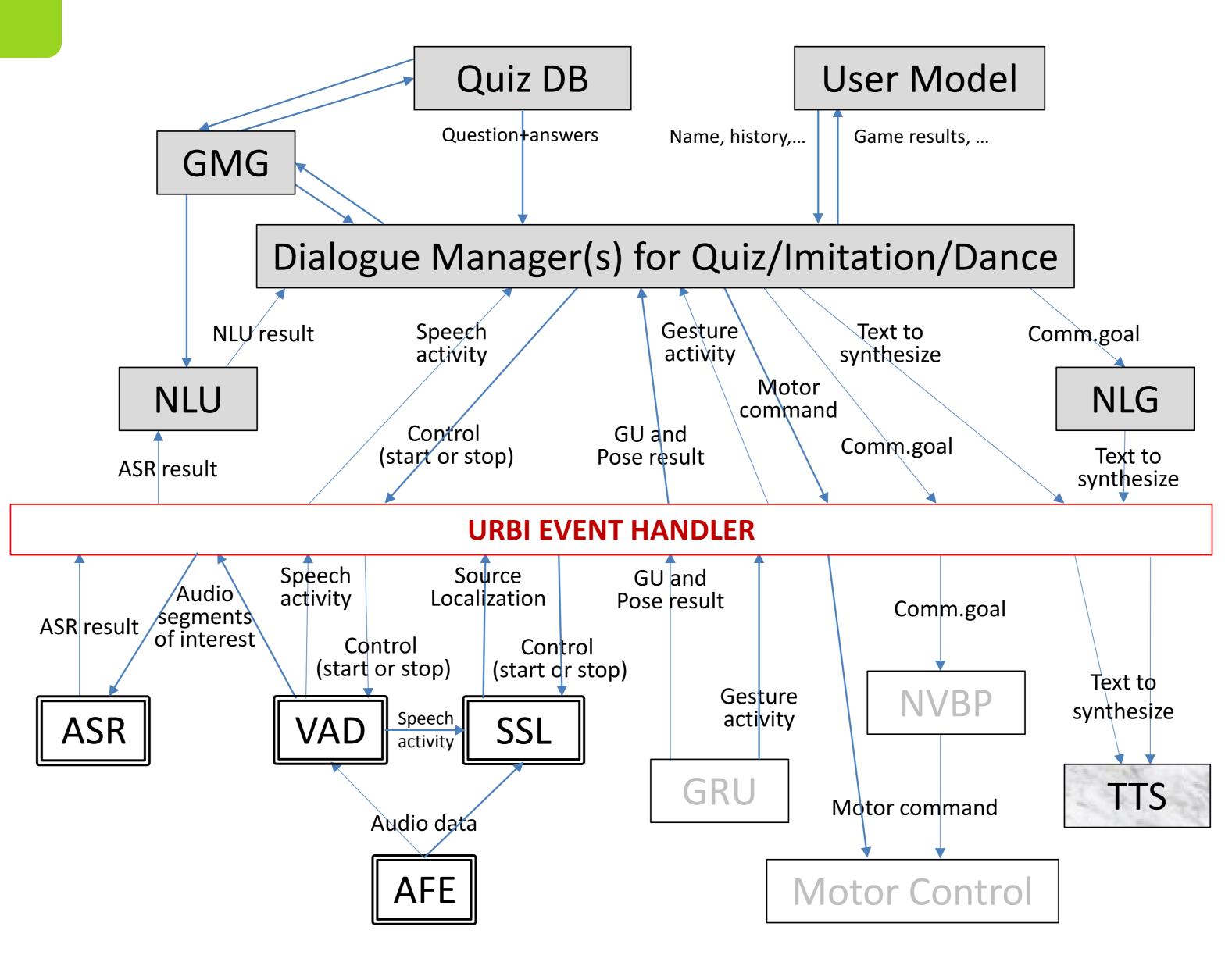
- Flexible event routing
- Provides central services
- -Implemented in UrbiScript

NL Processing

- Uses reversible CCG grammar
- Lattice parsing
- Canned Text as well as full generation with utterance planning

Speech Recognition

- Julius ASR, acoustic model from Italian corpus ChildIt
- Google ASR (eng), triggered by Voice Activity Detector



Non-verbal Planning/Motor Control

- Important for robot's acceptance
- On Nao: body language only

Dialogue Manager

- -MDPs with relational representations
- Information state approach
- Hierarchical reinforcement learn-ing

Text-To-Speech

-Acapela (on NAO)/MARY TTS

Gesture Recognition

- Hand gestures for Imitation
- OpenCV-based skin detection

Applications

Dance

Imitation



- R: Do you want to play a Quiz game?
- U: Sure
- R: One of us asks questions and offers possible answers. The other answers by selecting one of those answers. Do you want to ask me questions or shall I ask you?
- U: You ask me
- R: Okay, III ask you.

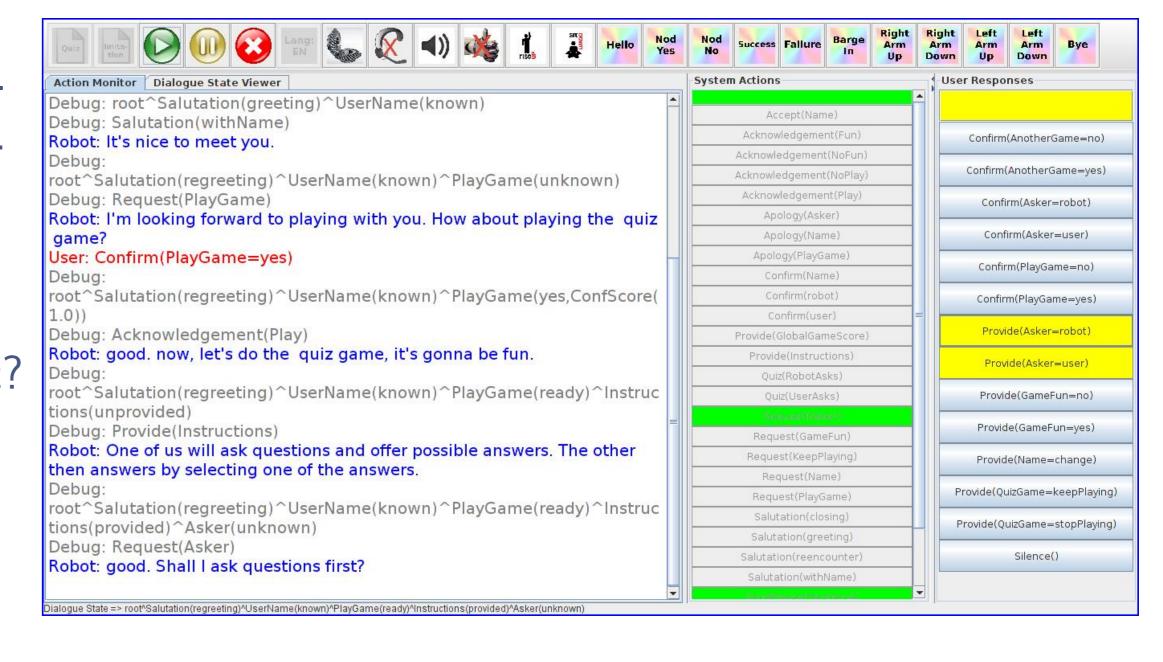
 Next question: What is the typical magic instrument?

 Your choices: 1, guitar. 2, drums. 3, violin. 4, flute.

 What is your choice?
- U: Number two
- R: That is not quite right.

 Do you have another choice?

WOz Interface



Conclusions

Integrated system implemented

- Various IO channels: speech, video, gestures
- Distributed, parallel and synchronized module execution

Quiz

- Event-based communication seems to be feasible
- Several experiments with children conducted, partially Wizard-of-Oz

Lessons learnt / Future Work

- Gesture recognition not reliable for untrained users
- On-board Nao sensors often too noisy
- Bandwidth problems using sensors w/ high data rate
- Striving towards ever more autonomy















