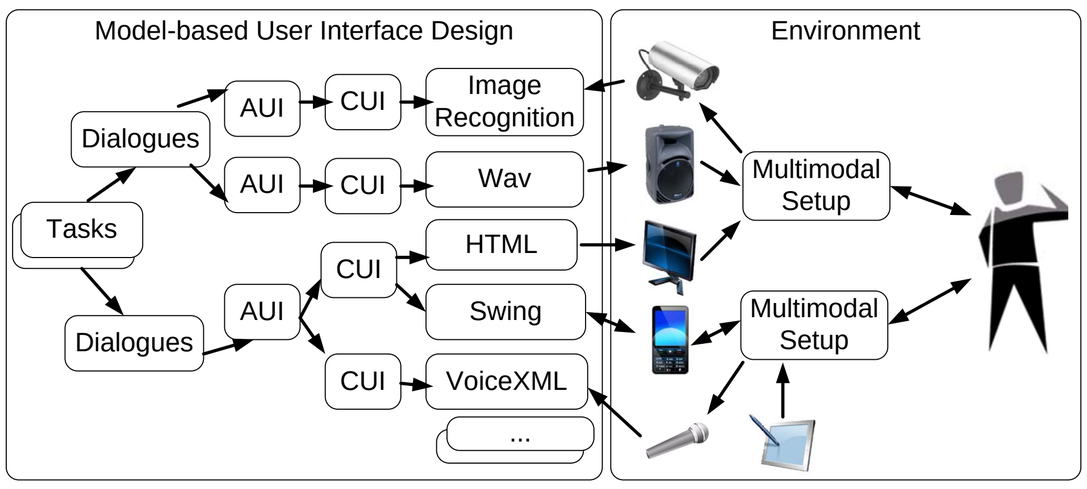
Interactive Dialogue Modeling for Designing Multimodal Applications

Sebastian Feuerstack
Universidade Federal de São Carlos
sfeu@cs.tu-berlin.de

Motivation

Multimodal systems that support the user by a combination of speech, gesture and graphical-driven interaction are already part of our everyday life. They rely on a static, predefined multimodal interaction setup, where the interaction devices, paradigms and the possibilities of controlling their applications are predefined.



Modeling multimodal systems that support various multimodal setups is an open research issue.

Challenges

1. How to model real multimodal interaction that includes fusion and fission?
2. Have different modalities more in common than they differ?
   - What is a suitable dialogue modeling abstraction?
   - Can design models be executed?
3. What is an appropriate design process?
   - Graceful Degradation vs. Abstract-to-Concrete

Related Work

The CARE Properties

***C**omplementary* denotes several modalities that convey complementary chunks of information.

***A**ssignment* implies that a modality is fixed in the way that the user has no choice in for performing a task with another modality.

***R**edundancy* indicates that the same piece of information is conveyed by several modalities.

***E**quivalence* of modalities implies that the user can perform a task using a modality chosen amongst a set of equivalent modalities.

Cameleon Reference Framework and UsiXML

Our work considers the user interface development abstraction levels of the Cameleon Framework. Different to UsiXML, which offers a standardized markup language for several modalities, we add a behavior specification of interactors and focus on designing multimodal interaction.

Evaluation in E-Learning

Multimodal approaches to learning have been proven to be extremely effective since information introduced aurally, visually and kinesthetically can significantly increase the possibility of understand and remembering information. By a case study we want to prove that our approach can be applied to model a multimodal learning application that can be run in different multimodal setups.

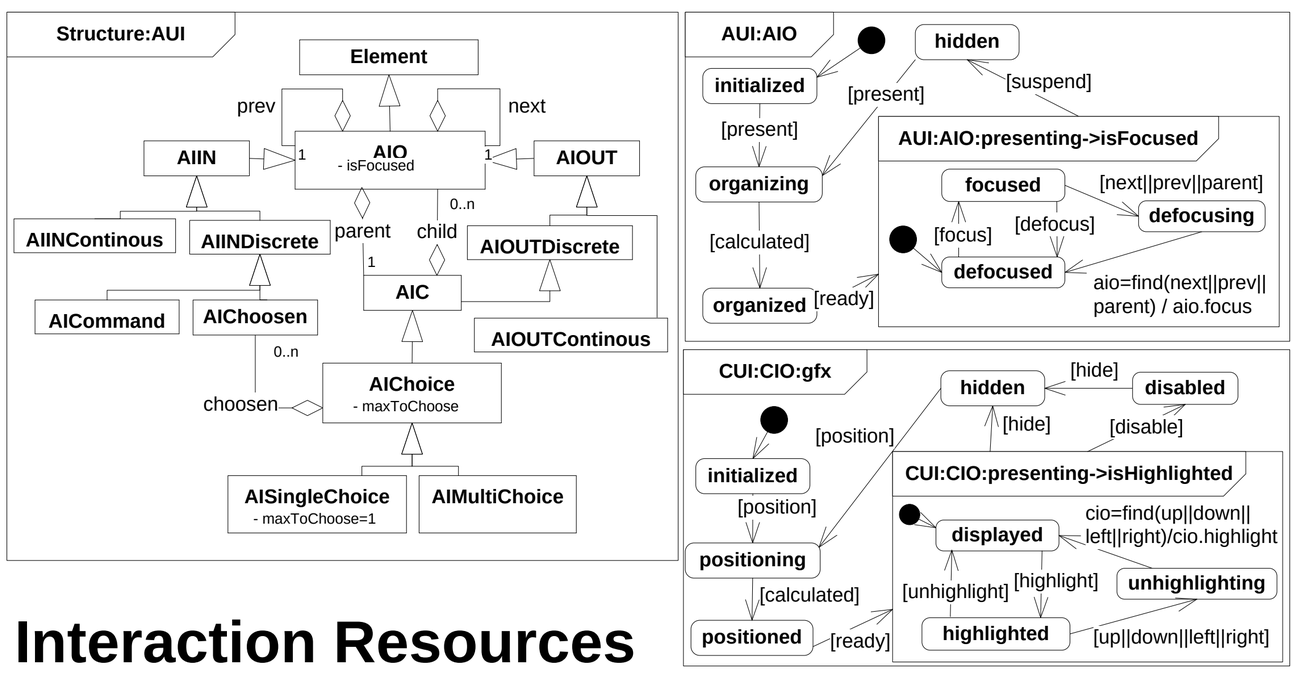
Approach

Bottom-up approach by an extensive description of abstract and concrete user interface interactors for graphical, voice, and gesture-driven interfaces.

=> Specification of structure and behavior.

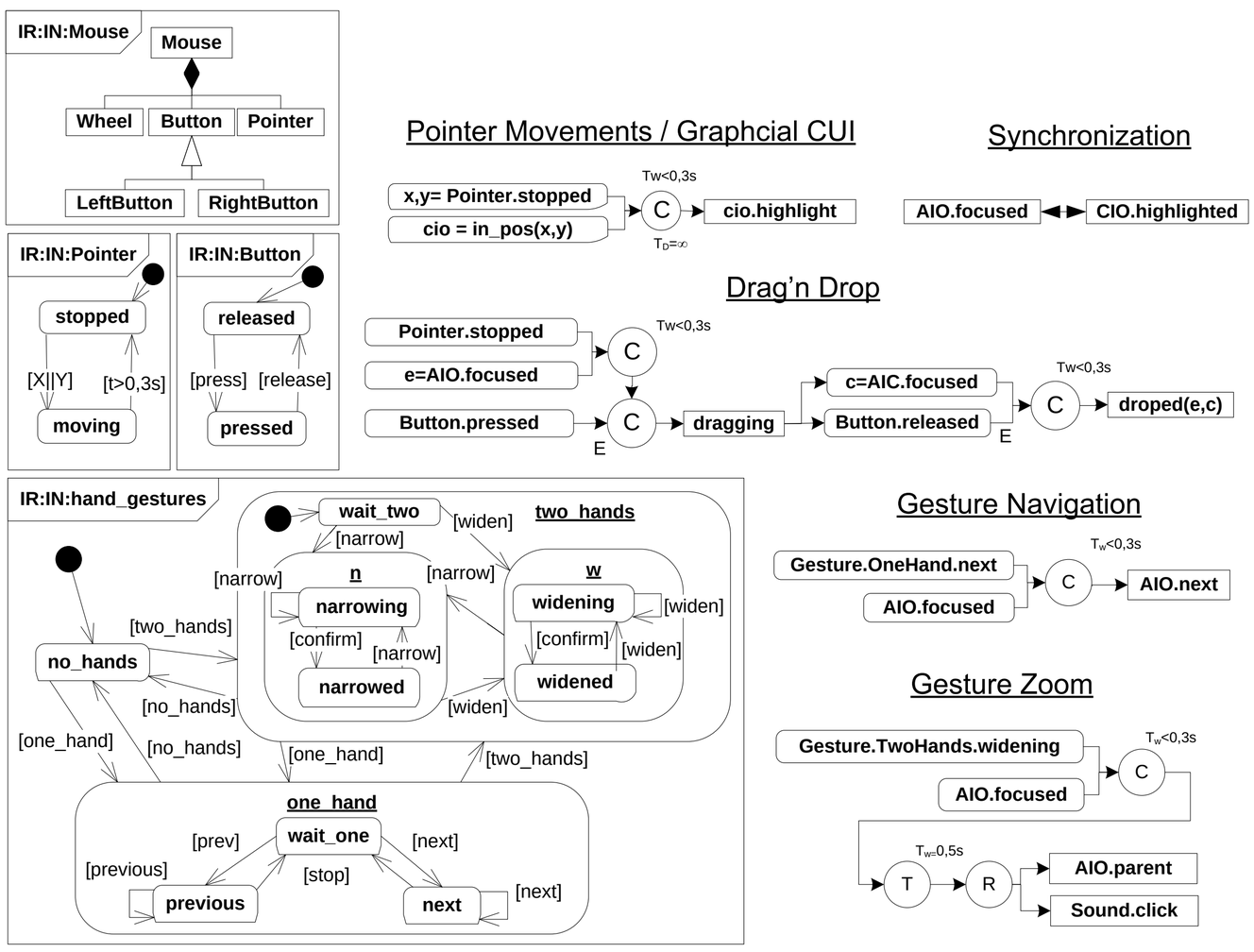
User Interface Interactors

Specification of interactors on modal-independent (abstract), and concrete (modal-dependent) level with state machines and class diagrams.



Interaction Resources

Flexible aggregation of components to form interaction resources that can be aggregated to multimodal setups. Mappings describe multimodal relations between interactors and interaction resources.



Editor and Interpreter

An Eclipse-based editor enables task-based design and is currently extended to support assembling of abstract and concrete interactors to dialogues. An initial interpreter has been implemented that loads the state machines and mapping definitions into software agents that communicate through a tuplespace.