

Developing a Decision-Theoretic User Model for Automatic Customization

Thesis Proposal

Bowen Hui

Need for Automatic Customization

- Problems from industry practices:
 - Varying user needs and preferences
 - One-size-fits-all solution
 - Lost in interface/functionality
- Most affected users
 - People with cognitive, sensory, motor impairments
 - Elderly people
 - Children
 - Novices



General Approaches

- Requirements engineering
 - Design solution
- Adaptable interfaces (HCI)
 - User has full control
 - Burden on user
- Adaptive interfaces (AI)
 - Automatic customization
 - Learns user-specific preferences
 - Mixed-initiative approach



Run-time solutions



Interface Customization

- Objectives
 - Minimize user effort
 - Maximize interaction experience
- Existing implementations
 - Auto-completions
 - Toolbars
 - Adding/hiding widgets
 - Hints (balloons)
 - Questions (dialog boxes)



Example: Adaptive Word Completion Typing Task

state of affairs are far b

- | |
|-------------------|
| 1. buses |
| 2. because |
| 3. bedroom |



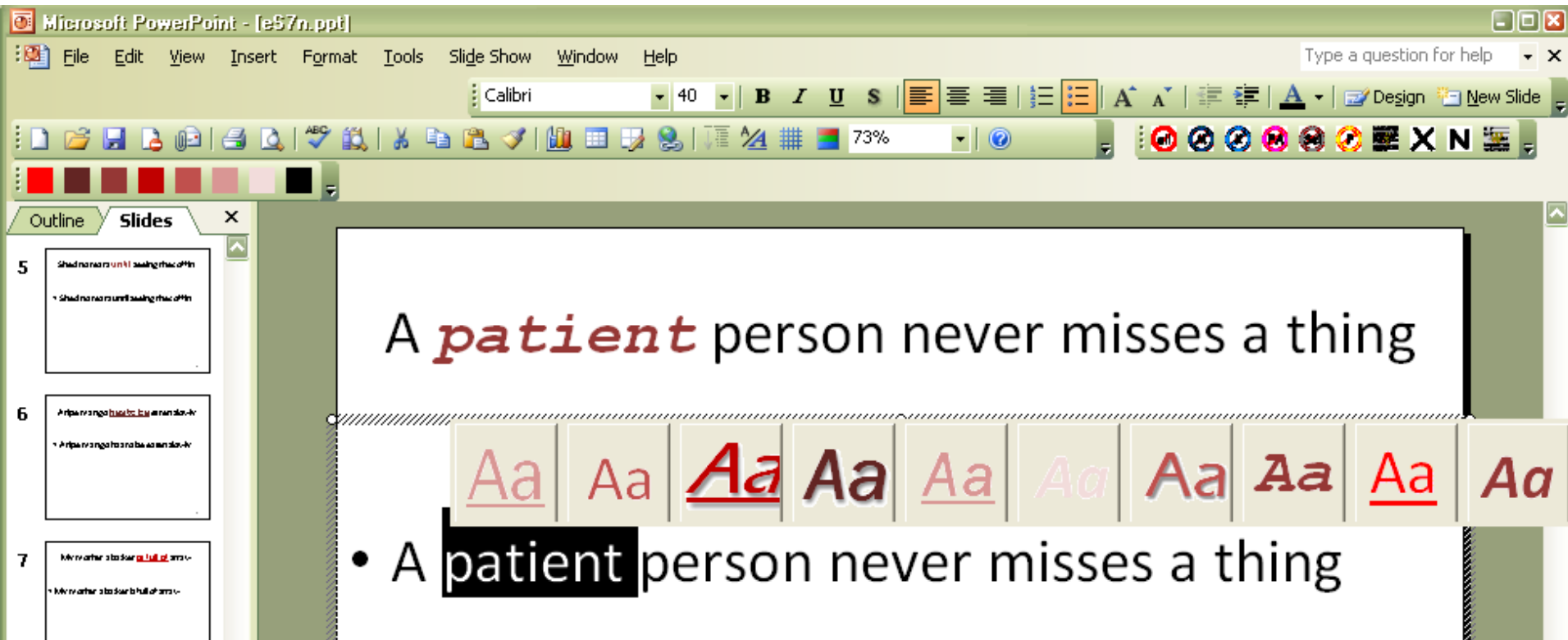
Example: Adaptive Layout Menu Selection Task

The screenshot shows a window titled "SemFace" with a menu bar containing: Fruits, Emotions, Vehicles, Illnesses, Ordours, Teas, Singers, Whales, Housing. The "Vehicles" menu is expanded, showing a list of vehicle types. Pink arrows indicate a selection path starting from the "Vehicles" menu item, moving down to "Ambulance", then to "Trucks", and finally to "Kit Cars".

Menu Item	Sub-item
Fruits	Surprise
Emotions	Townhouse
Vehicles	Trucks
Vehicles	Trucks
Vehicles	Limousines
Vehicles	Car
Vehicles	Ambulance
Vehicles	Trucks
Vehicles	Trailers
Vehicles	Motorcycles
Vehicles	Motorhomes
Vehicles	Lowriders
Vehicles	Vans
Vehicles	Jeeps
Vehicles	Bus
Vehicles	Taxi
Vehicles	Kit Cars
Illnesses	
Ordours	
Teas	
Singers	
Whales	
Housing	



Example: Adaptive Toolbar Highlighting Task



The screenshot shows a Microsoft PowerPoint window with the title "Microsoft PowerPoint - [e57n.ppt]". The menu bar includes File, Edit, View, Insert, Format, Tools, Slide Show, Window, and Help. The ribbon shows the Font section with Calibri font, size 40, and options for Bold, Italic, Underline, and Strikethrough. The text "A patient person never misses a thing" is displayed on the slide, with the word "patient" highlighted in red. Below the text, an adaptive toolbar is shown with ten buttons, each containing the text "Aa". The third button from the left is highlighted in red, indicating it is the active tool for highlighting text.

Outline Slides

5

6

7

A *patient* person never misses a thing

Aa Aa **Aa** Aa Aa Aa Aa Aa Aa Aa

- A patient person never misses a thing



Customization Issues

- **Who:** user type,
varying preferences
- **What:** kinds of help
(e.g., auto-completion vs. toolbar)
- **When:** long term benefits and risks
- **Where:** near current focus,
off to the side
- **How:** display mechanism (e.g., animation)
- **Why:** rationale,
transparency



Specific Research Questions

- Individual differences?
- Influence on goal recognition?
- Influence on perceived value of help?



Specific Research Questions

- Individual differences?
 - General user characteristics?
 - Variation in “reactions” to adaptive help?
 - Explanatory model of differences?
 - Patterns of user behaviour?
 - Fast, online inference?
- Influence on goal recognition?
- Influence on perceived value of help?



Specific Research Questions

- Individual differences?
- Influence on goal recognition?
 - Modeling idiosyncrasies?
 - Classes of user-specific goals?
 - Quantifying quality of help?
 - Fast, online inference?
- Influence on perceived value of help?



Specific Research Questions

- Individual differences?
- Influence on goal recognition?
- Influence on perceived value of help?
 - Factors relevant to costs/benefits?
 - Predictive model of interaction factors?
 - Elicit preferences with real users?
 - Capture evolving preferences?

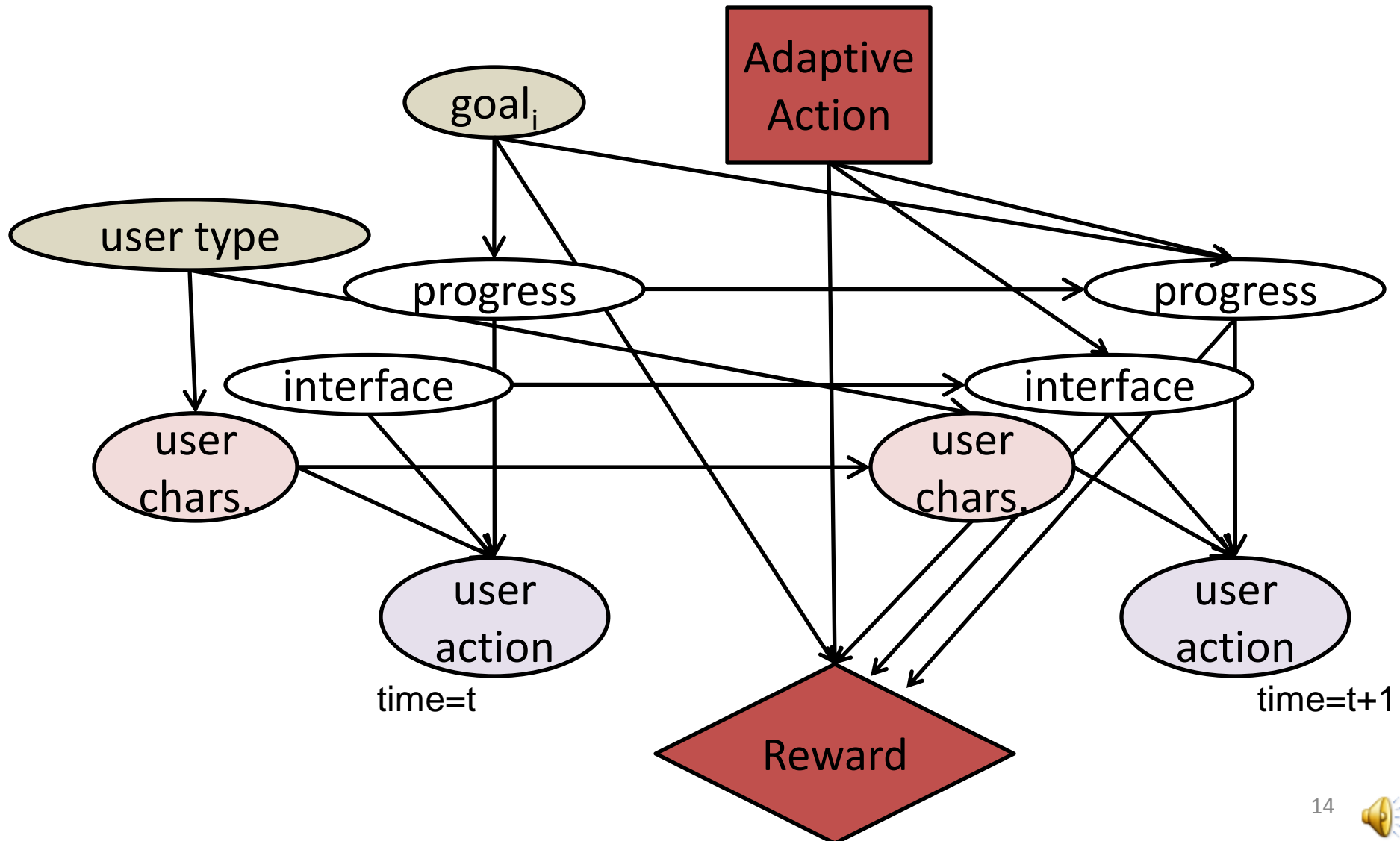


Our Approach/Thesis Structure

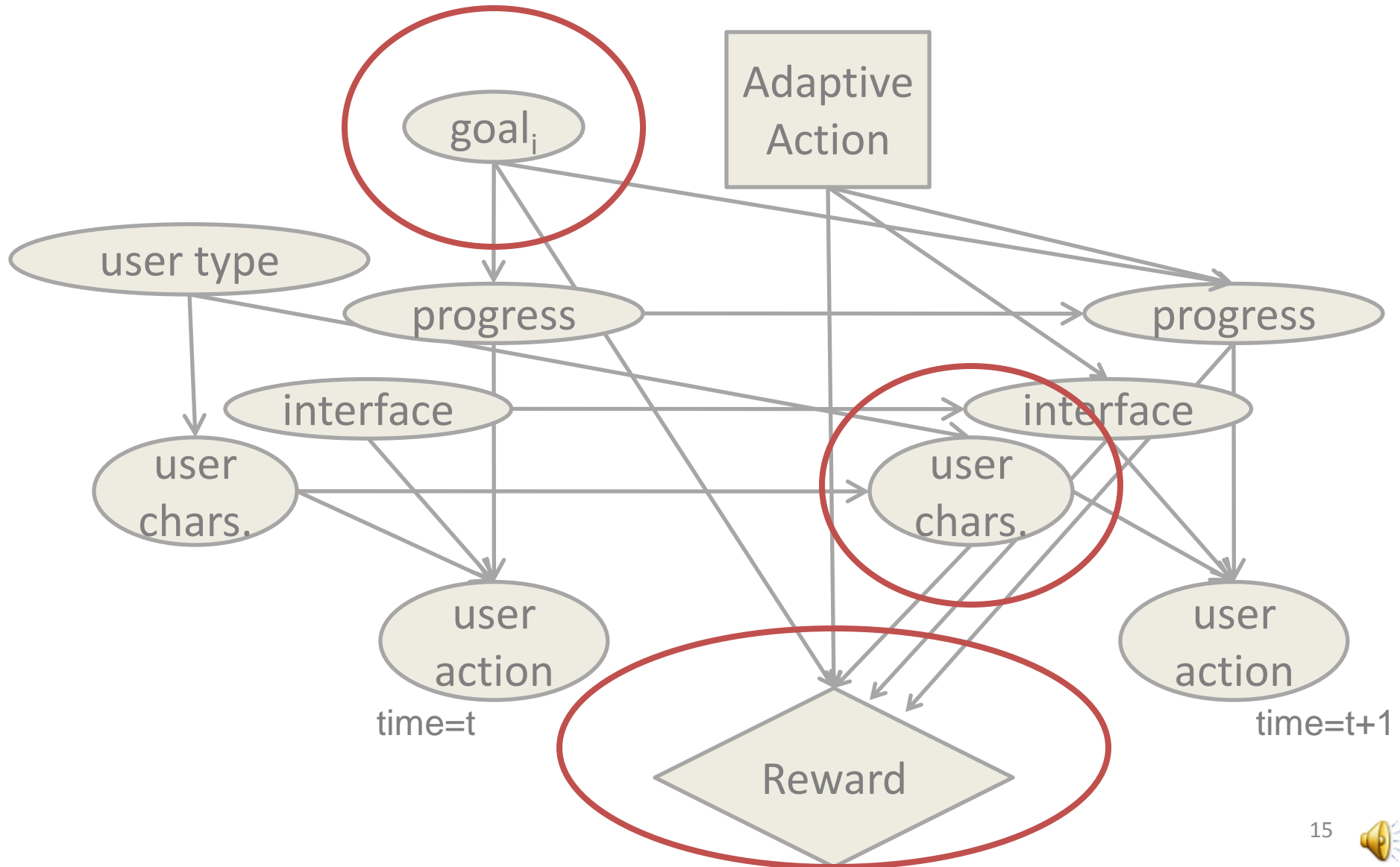
- Interface customization as a decision-theoretic planning problem
- General POMDP customization assistant **Ch 3**
- Canonical system architecture
 - User characteristics prediction **Ch 4**
 - User goal prediction **Ch 5**
 - Action selection **Ch 6**
- Evaluate integrated system **Ch 7**



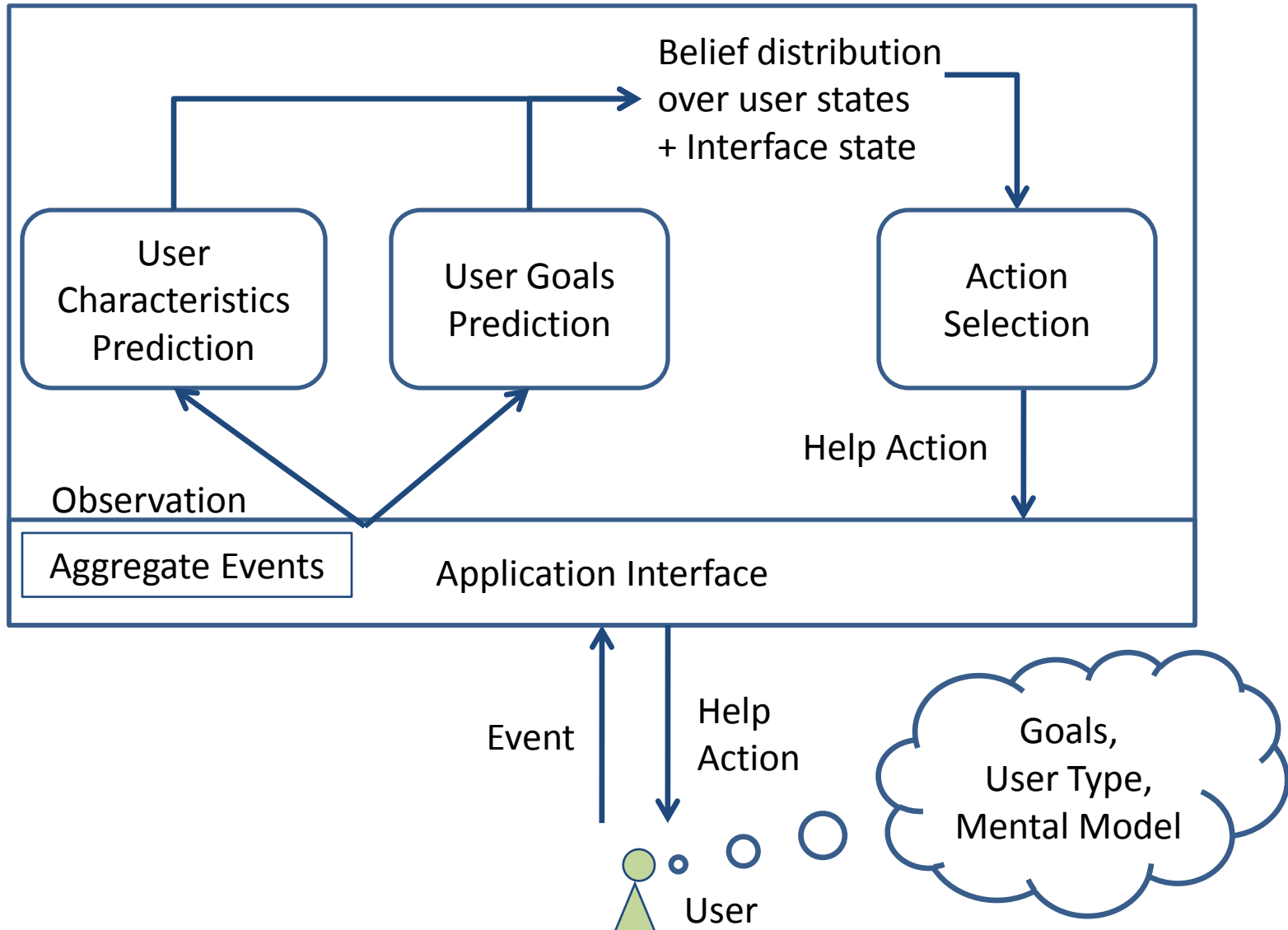
POMDP Assistance Model



POMDP Assistance Model



Canonical System Architecture



Chapter 4:

A Model of Individual Differences

- Evidence supporting individual differences
- Representing user characteristics
- Bayesian model of characteristics for savings
 - Domain: adaptive word prediction
 - Expm: data collection, simulation, usability
- Probabilistic model of mental model for disruption
 - Domain: adaptive menu
 - Expm: data collection, simulation, usability



Chapter 5:

Personalized Goal Recognition

- Common goal classes
- Finite state models
 - Indentation goals
 - Highlighting goals
- Inference procedure
- Evaluation
 - Domain: adaptive toolbar in PowerPoint
 - Expm: simulation **(incomplete)**



Chapter 6:

Utility of Intelligent Assistance

- Interaction factors as influence diagrams
- Bloat
 - Domain: adaptive menu
 - Expm: simulation
- Savings (summary)
- Disruption (summary)
- Eliciting subjective value of help
 - Domain: adaptive toolbar in PowerPoint
 - Expm: experiential elicitation



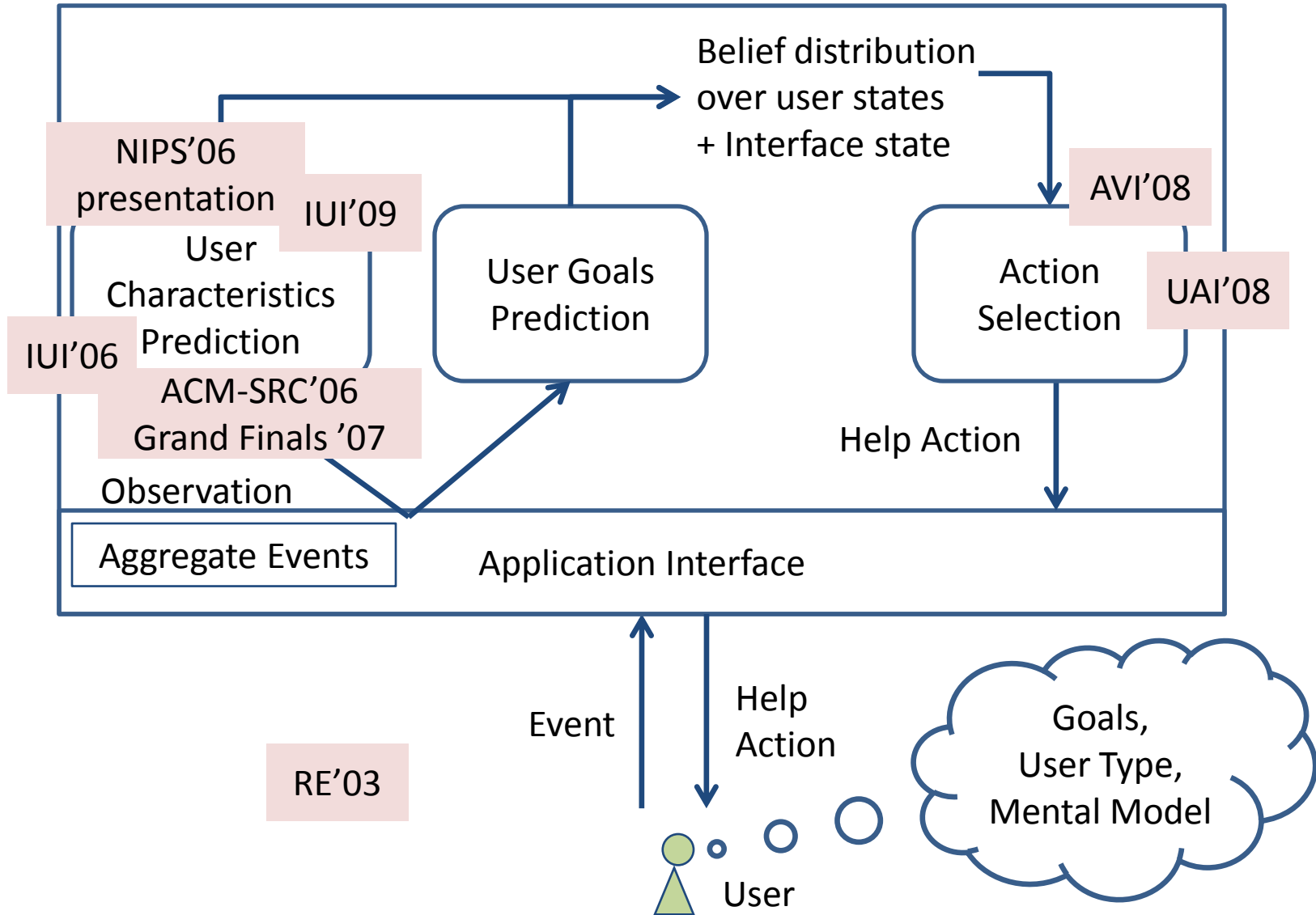
Chapter 7:

Integrated System and Evaluation

- Main questions:
 - What is the impact of the user characteristics component on overall utility?
 - What is the impact of a decision-theoretic policy on overall utility?
- Measure utility as a function of changing accuracy
- Integrated system as test application
 - Domain: adaptive toolbar in PowerPoint
 - Expm: usability **(incomplete)**



Summary of Contributions



Major Steps Remaining

- Chapter 5 simulation
 - Set-up automatic simulation environment
- Chapter 7 evaluation
 - Modify design?
 - Experiments with people?
- Chapter 8 conclusions
 - Writing
- Editing throughout



Details on Integrated Evaluation

- Understand contribution to overall help utility
- Questions of interest:
 - impact of user characteristics component?
 - impact of decision-theoretic policy?
- Test Policies
 - Heur-G, Heur-GC, DT-G, DT-GC, Always, Never
- Prediction Accuracy
 - acc_G = distribution (or ranks) of goals
 - acc_C = distribution (or ranks) of user types



Design Options

acc_G	acc_C	Always	DT-G	DT-GC	Never	Heur-G	Heur-GC
100	100	C1	C2	C3	C4	C5	C6
50	100	C7	C8	C9	(C4)	C10	C11
100	50	(C1)	(C2)	C12	(C4)	(C5)	C13

- Within: ~100 people per 13 conditions
 - More controlled goal recognition
 - Allows comparison of multiple policies
- Across: ~10 people per condition (x13)
 - More realistic goal recognition
 - Experiences one policy only



Simplifying Assumptions

acc_G	acc_C	Always	DT-G	DT-GC	Never	Heur-G	Heur-GC
100	100	C1	C2	C3	C4	C5	C6
50	100	C7	C8	C9	(C4)	C10	C11
100	50	(C1)	(C2)	C12	(C4)	(C5)	C13

- Explore char comp impact in best case
 - Ignore $acc_C=50\%$ → 11 conditions
- Assume char comp works best with DT policy
 - Ignore Heur-GC → 9 conditions
- Drop one of the two questions?
 - Ignore impact of DT/Char → 7 conditions
- Drop one accuracy setting?
 - Ignore $acc_G=50\%, acc_C=100\%$ → 5 conditions



Pilot Study

- **Within-subjects** design per accuracy setting
- 5 systems (1 accuracy setting) per experiment
- 10 highlighting tasks in augmented PowerPoint
- 4 events per goal (minimal)
- ~60 minutes per experiment
- Hypotheses to compare mean “utilities” across system policies



Sample Slide

Facts about *Frogs*



- Some frogs have suction disks on the tip of their toes to help them climb



“H0”: Decreasing Goal Accuracy

- mean utility of sys A at $\text{acc}_G=100\%, \text{acc}_C=100\%$ > mean utility of sys A at $\text{acc}_G=50\%, \text{acc}_C=100\%$
- Cases:
 - System policy = Always
 - System policy = Heur-G
 - System policy = DT-G
 - System policy = DT-GC
- Objective: replicability of literature results



H1: Impact of Char Comp

- mean utility of DT-G < mean utility of DT-GC
- Cases:
 - Setting at $\text{acc}_G=100\%, \text{acc}_C=100\%$
 - Setting at $\text{acc}_G=50\%, \text{acc}_C=100\%$
- Objective: benefit of characteristics component



H2: Impact of DT Policy

- mean utility of Heur-G < mean utility of DT-G
- Cases:
 - Setting at $\text{acc}_G=100\%, \text{acc}_C=100\%$
 - Setting at $\text{acc}_G=50\%, \text{acc}_C=100\%$
- Objective: benefit of decision-theoretic policy



Remaining Steps in Evaluation

- Testing/Debugging?
- Modify design?
- More experiments?
- Results analysis?

