Lecture 5

Usability Evaluation Methods (other than Usability Testing)

(chapter 4)

Usability Evaluation Summary

- Purpose: Evaluation for usability
- Methods
  - Without Users (analytic)
    - Guidelines (Chapter 2.2)
    - Interface Walkthrough
    - Expert Review (Chapter 4.2)
    - Model-Based analysis (Keystroke Model)
  - With Users (empirical)
    - Usability testing (Chapter 4.3)
    - Experiments (Chapter 4.7)
    - Field Studies (Chapter 4.6)
    - Surveys (Chapter 4.4)

Testing Goals vs. Method

	Guidelines	Walkthrough	Expert	Keystroke Level Model	Usability Testing
Overall Usability	√		√		√
Detailed Usability		√	√		√
Completeness		√	√		
Correctness		√	√		
Consistency		√	√		
Performance Time				√	√

Interface (Cognitive) Walkthrough

- Definition
  - Given an interface and a set of functions, for each function the developers walk through the user's action sequence noting any problems.
- Purpose
  - Checks usability by a detailed rehearsal of steps in user tasks
- Problems
  - Incomplete since the developers may have made oversights
  - Developers are not the users.

Expert (Heuristic) Evaluation

- Definition
  - Human factors expert evaluates interface
  - Uses Guidelines and experience
  - Suggests changes
- Purpose
  - Use human expertise to find usability problems
- Problems
  - Expert may be unavailable
  - Experts may disagree
  - An expert is not a user

Heuristic evaluation

(Niesen & Molich, CHI 1990, "Heuristic evaluation of user interfaces")

- Problem
  - Individual evaluators found between 20-51% of usability problems
- Solution
  - Use aggregated results from 3 to 5+ evaluators

Keystroke Level Model

- Definition
  - Predicts time to do a task for an expert user
- How to do it
  - Specify a task with low-level actions
  - Give times for each action and system response
    - key press = .2 sec; mouse point = 1.1 sec; reach = .4 sec
  - Sum to compute estimated time for the task
- Problems
  - Reliability of prediction is 80%
  - Cannot predict learning time

Guidelines

- Shared language
- Best practices
- Proponents
  - Encapsulates experience of real designers

Guidelines

- Definition
  - Guidelines are written "standards" and heuristics (rules of thumb) for interfaces. Following them should lead to more usable designs.
- Guidelines are used to evaluate the mockups, scenarios, storyboards and user interaction networks specifications

Guidelines (Apple Computer)

- Examples from Apple Computer's *Human Interface Guidelines*, 1985-1989
  - Friendly Dialog. Plain English, no jargon. Ask for clarification of risky operations.
  - Learnability. Use concrete metaphors.

Guidelines of Usability Experts (Arnold Lund, Ameritech, 1995)

- Experts in HCI design field suggested rules of thumb they found particularly useful during design
- 34 rules defined
- 31 HCI experts rated each of the rules of thumb by their estimate of magnitude of impact on usability of designs
- 5 is most impact; 1 is least
- Forced choice (20% of rules "5", 20% "4", etc.)
- 82% inter-rater correlation: high consensus

Guidelines of Usability Experts (5 is most impact on usability)

```
4.1     Know thy user, and YOU are not the user.
4       Things that look the same should act the same.
4       Everyone makes mistakes, so every mistake should be
fixable.
3.9     The information for the decision needs to be
        there when the decision is needed.
3.8     Error messages should actually mean something to
the user, and tell the user how to fix the
problem.
3.8     Every action should have a reaction.
3.7     Don't overload the user's buffers.
3.6     Consistency, consistency, consistency.
3.5     Minimize the need for a mighty memory.
3.5     Keep it simple.
```

Guidelines of Usability Experts - cont.

3.4 The more you do something, the easier it should be to do.

3.4 The user should always know what is happening.

3.4 The user should control the system. The system shouldn't control the user. The user is the boss, and the system should show it.

3.3 The idea is to empower the user, not speed up the system.

3.3 Eliminate unnecessary decisions, and illuminate the rest.

3.3 If I made an error, let me know about it before I get into REAL trouble.

3.3 The best journey is the one with the fewest steps. Shorten the distance between the user and their goal.

Guidelines

- Problems
  - Too specific, incomplete, hard to apply
  - Unsystematic folklore: sometimes wrong
  - Tradeoffs

Dilbert by Scott Adams

Testing & Evaluation: What is the best method? (Jeffries et al., CHI '91)

- Which method discovers the most severe usability problems?

Walkthrough	3.44
Expert Evaluation	3.59
Guidelines	3.61
Usability Testing	4.15

- Range: 1= trivial 9= critical
- mean shown for severity of problems discovered

Testing & Evaluation: What is the best method?

- Number of severe problems found

	Most Severe	Least Severe
Walkthrough	9	10
Guidelines	12	11
Usability Testing	18	2
Expert Evaluation	28	52

- most= top third least= bottom third

Usability Testing in Practice

(Nielsen, *Usability Engineering*, 1993)

Survey of 28 projects from broad variety of companies and industries

Findings:

	Median
Project size in person-years	24
Actual share of budget for usability	6%
Ideal share of budget for usability	10%
Actual usability effort in person-years	1.5
Ideal usability effort in person-years	2.1

Usability Testing in Practice
cont.

- Findings:
  - Usability effort is independent of project size!
  - Why? Many usability activities take about the same time to perform, no matter how difficult the program is to develop

Testing and Evaluation of Users
Summary

- Methods
  - Without Users
    - Guidelines
    - Interface (Cognitive) walkthrough
    - Heuristic evaluation
    - Keystroke Level Model
  - With Users
    - Usability testing
    - Questionnaire & Interview
- Use *all* methods for a more usable interface

Beyond Guidelines: Principles

- More fundamental, widely applicable, and enduring than guidelines
- Need more clarification
- Fundamental principles
  - Principle 1: Determine user's skill levels
  - Principle 2: Identify the tasks
  - Principle 3: Five primary interaction styles
  - Principle 4: Eight golden rules of interface design
  - Principle 5: Prevent errors
  - Principle 6: Automation and human control

Principle 1: Determine user's skill levels

- "Know thy user" Hansen (1971)
- Age, gender, physical and cognitive abilities, education, cultural or ethnic background, training, motivation, goals and personality
- Design goals based on skill level
  - Novice or first-time users
  - Knowledgeable intermittent users
  - Expert frequent users
- Multi-layer designs

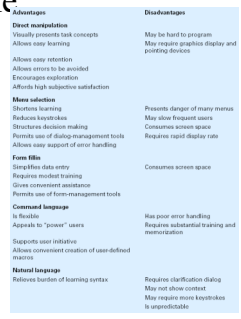
Principle 2: Identify the tasks

- Task Analysis usually involves long hours observing and interviewing users
- Decomposition of high level tasks
- Relative task frequencies

	TASK				
Job title	Query by Patient	Update Data	Query across Patients	Add Relations	Evaluate System
Nurse	0.14	0.11			
Physician	0.06	0.04			
Supervisor	0.01	0.01	0.04		
Appointment personnel	0.26				
Medical-record maintainer	0.07	0.04	0.04	0.01	
Clinical researcher			0.08		
Database programmer			0.02	0.02	0.05

Principle 3: Choose an interaction style

- **Direct Manipulation**
- **Menu selection**
- **Form fillin**
- **Command language**
- **Natural language**



Advantages	Disadvantages
Direct manipulation	
Visually presents task concepts	May be hard to program
Allows easy learning	May require graphics display and pointing devices
Allows easy retention	
Allows errors to be avoided	
Encourages exploration	
Affords high subjective satisfaction	
Menu selection	
Shortens learning	Presents danger of many menus
Reduces keystrokes	May slow frequent users
Structures decision making	Consumes screen space
Permits use of dialog-management tools	Requires rapid display rate
Allows easy support of error handling	
Form fills	
Simplifies data entry	Consumes screen space
Requires modest training	
Gives convenient assistance	
Permits use of form-management tools	
Command language	
Is flexible	Has poor error handling
Appeals to "power" users	Requires substantial training and memorization
Supports user initiative	
Allows convenient creation of user-defined macros	
Natural language	
Relieves burden of learning syntax	Requires clarification dialog
	May not show context
	May require more keystrokes
	Is unpredictable

Principle 4: The 8 golden rules of interface design

1. Strive for consistency
2. Cater to universal usability
3. Offer informative feedback
4. Design dialogs to yield closure
5. Prevent errors
6. Permit easy reversal of actions
7. Support internal locus of control
8. Reduce short term memory

Principle 5: Prevent errors

- Make error messages specific, positive in tone, and constructive
- Mistakes and slips (Norman, 1983)
- Correct actions
  - Gray out inappropriate actions
  - Selection rather than freestyle typing
  - Automatic completion
- Complete sequences
  - Single abstract commands
  - Macros and subroutines

Principle 6: Automation and human control

Humans Generally Better	Machines Generally Better
Sense low-level stimuli	Sense stimuli outside human's range
Detect stimuli in noisy background	Count or measure physical quantities
Recognize constant patterns in varying situations	Store quantities of coded information accurately
Sense unusual and unexpected events	Monitor prespecified events, especially infrequent ones
Remember principles and strategies	Make rapid and consistent responses to input signals
Retrieve pertinent details without a priori connection	Recall quantities of detailed information accurately
Draw on experience and adapt decisions to situation	Process quantitative data in prespecified ways
Select alternatives if original approach fails	Reason deductively: infer from a general principle
Reason inductively: generalize from observations	Perform repetitive preprogrammed actions reliably
Act in unanticipated emergencies and novel situations	Exert great, highly controlled physical force
Apply principles to solve varied problems	Perform several activities simultaneously
Make subjective evaluations	Maintain operations under heavy information load
Develop new solutions	Maintain performance over extended periods of time
Concentrate on important tasks when overload occurs	
Adapt physical response to changes in situation	

Automation and human control (cont.)

- Successful integration:
  - Users can avoid:
    - Routine, tedious, and error prone tasks
  - Users can concentrate on:
    - Making critical decisions, coping with unexpected situations, and planning future actions

Beyond Principles: Theories

- Beyond the specifics of guidelines
- Principles are used to develop theories
- Descriptions: explanatory or predictive
  - Example: Fitts Law for predicting pointing time
- Motor task, perceptual, or cognitive