User studies

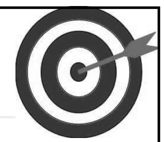
CSEP 510
Lecture 4, January 29, 2004
Richard Anderson

Happy Birthday

- Apple Macintosh
- January 23, 1984, Cupertino, CA.
  - 32 bit microprocessor
  - Nine inch display
  - 64K ROM, 128K RAM
  - $2,495

Announcements

Fitts' Law



- Targetting

Outline

- GOMS
  - Goals, Operators, Methods, Selection rules
  - Model of behavior above KLM
- Applications of human modeling
- Human Subjects Issues
- User study examples
- Ethnographic observation

GOMS

- Modeling behavior where users have patterns of use

GOMS

- Goals
  - Goals available for solving the task
- Operators
  - Primitive operations
- Methods
  - Compiled collection of sub-goals and operators
- Selection rules
  - Rules to choose amongst methods

GOMS Example Room cleaning

GOMS/KLM Critique

- Skilled users
- Ignored learning
- Errorless performance
- Did not differentiate cognitive processes
- Serial tasks
- Does not address mental workload
- Ignores user fatigue
- Does not account for individual differences
- Does not consider broader issues of the application

User variation

- Extent of knowledge of tasks
- Knowledge of other systems
- Motor skills
- Technical ability
- Experience with system
  - Novice, Casual, Expert

Skilled vs. Unskilled users

- What is the difference between modeling skilled and unskilled users

Modeling Errors

- How would you adapt a KLM to handle errors?

Parallel vs. Serial execution

- Instruction scheduling analogy
  - Summing individual instruction times on a pipeline processor is a poor predictor
- Does this analogy apply for KLM?
- How does GOMS apply to email when user is working on many messages simultaneously?

Application of Modeling

- Papers from CHI 2000
- Keystroke Level Analysis of Email Message Organization
- Predicting Text Entry Speed on Mobile Phones

Paper summaries

- What are the problems being studied?
  - Build the model because of some motivating problem
- What is the model?
- What does the model predict
- Experimental validation

Email organization

- Users spend a significant amount of time manipulating email
- Different strategies are used for working with email
  - E.g., frequent filers, spring cleaners
- Understand performance of strategies
- Understand features of mail applications

Email processing model

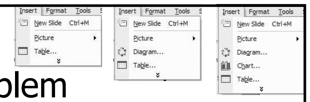
- Time spent storing email
- Time spent retrieving email
- Model
  - Store in folder – find folder, move message
  - Retrieve from folder – find folder, find message
  - Known folder (position / name known)
  - Unknown folder

Cost to find a folder

- F folders, V folders visible on screen
- Unknown folder
- Known folder
  - Scroll folder onto screen
  - Locate folder on the screen

Analysis approach

- Model for primitive operations
- Cost estimate based on parameters
- Operation pattern based upon strategy



Related problem

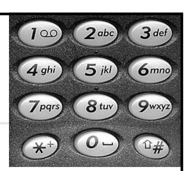
- Menu item selection in adaptable menus
- Optimization problem – minimize selection time
- Options:
  - Reorder items
  - Hide items

How do you model this?

- Access model
  - Select menu item from among k items
  - Select menu item by expanding menu
- What algorithm do you use to choose menu items
- How do you analyze its performance
- What else does the model need

Text entry on mobile phones

- Predict text entry speeds for new text entry methods
- Expensive to test
  - Build prototype
  - Test on users over time



Model

- Input methods
  - Multipress
    - THE: 8-4-4-3-3-0
    - Timeout or kill character for *segmentation*
  - Two-key
    - THE: 8-1-4-2-3-2-0
  - T9
    - THE: 8-4-3-0
    - Next key for *disambiguation*
      - base, card both 2-2-7-3

Keystroke model

- Operation time, MT_i, MT_{repeat}, $MT_{timeout}$, MT_{kill}
- Movement time, MTi estimate using Fitts' law
- Linguistic Model – diagraph probabilities, 27x27 matrix
  - P_{ij} – probability of the digram c_i c_j (in English)

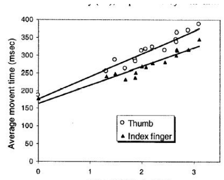
Words per minute

- CT_{ij} – cost of the digram i, j
  - Cost of j, starting at i
- $CT_L = SS(P_{ij} + CT_{ij})$
- WPM – assume 5 characters per word

Fitts' law to compute movement times

- $MT = a + b \log_2(A/W + 1)$
- Experimentally determine coefficients for one handed and two handed use
- Time users on selected key pairs
- Determine coefficients

Results



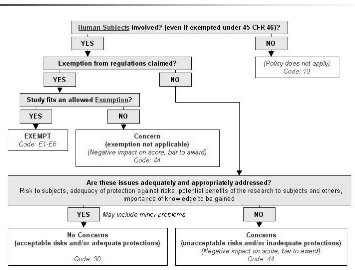
	Intercept, a	Slope, b	Correlation
Index Finger	165	52	0.960
Thumb	176	64	0.970

Predictions

Method	Index Finger	Thumb
Multi-press		
-- wait for timeout	22.5 wpm	20.8 wpm
-- timeout kill	27.2 wpm	24.5 wpm
Two-key	25.0 wpm	22.2 wpm
T9	45.7 wpm	40.6 wpm

T9 results for optimistic case of no disambiguation

Human Subjects Review



Human Subject Issues*

- Research involving human subjects requires review and approval
- Important ethical issues in working with human subjects
- Legal and institutional issues are also very important
- *This is the university perspective – but similar issues apply in industry

UW Process

- Before conducting research that involves human subjects, approval must be granted by the HSD
- Levels of review
  - Exempt
  - Minimal Risk
  - Full Review

What are the real issues?

- Avoidance of harm to the subject
  - Physical Injury
  - Emotional Stress
  - Invasion of privacy
- Informed consent
- Freedom from coercion
  - Recruitment of subjects

Minimal Risk vs. Full Review

- Boundary between minimal risk and full review
- What are the concerns for minimal risk
- Under what conditions is work exempt?

Consent Form

Investigators' statement
We are asking you to be in a research study. The purpose of this consent form is to give you the information you will need to help you decide whether or not to be in the study. Please read the form carefully. You may ask questions about the purpose of the research, what we would ask you to do, the possible risks and benefits, your rights as a volunteer, and anything else about the research or this form that is not clear. When all your questions have been answered, you can decide if you want to be in the study or not. This process is called 'informed consent.'
PURPOSE AND BENEFITS
We want to better understand how Tablet PCs can be used in the classroom to support active learning. We would like to study the use of Tablet PCs in junior level computer science courses. We hope the results of this study will help us deploy technology in the classroom to improve the educational experience for students. You may not directly benefit from taking part in this research study.

Consent Form (Procedures)

PROCEDURES
Tablet PCs will be made available to students in this class so that they can give real time responses to in class activities. Participation in these activities will be optional, and the activities will not be graded. We are planning to use Tablet PCs in class once per week for the duration of the course. Use of the Tablet PCs is optional for the students, and using the Tablet PC one week does not obligate you to use it in subsequent weeks.
If you choose to be in the study, we would like to record your usage of the Tablet PC during class and conduct a follow up survey. The Tablet PC will be running an application called SIP. We will record all of your activities while running this application. Your activities will be recorded using a study code, which is linked to your name. Your instructor will not have access to the recorded activity. We will use the study code to link your usage in different sessions, and to link your usage to your survey answers. …
You may withdraw from the study at any time.

Consent Form (Risks)

RISKS, STRESS, OR DISCOMFORT
Some people feel that providing information for research is an invasion of privacy. We have addressed concerns for your privacy in the section below.
OTHER INFORMATION
Taking part in this study is voluntary. You can stop at any time. Information about you is confidential. We will code the study information. We will keep the link between your name and the code in a separate, secured location until June 30, 2004. Then we will destroy the link. If the results of this study are published or presented, we will not use your name.
We will not share any of the information collected with the course instructor, and participation in the study will not influence your grade. The instructor will not know which of the students using Tablet PCs are participating in the study.

User experiments

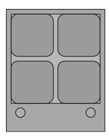
- Two-Handed Input using a PDA and a Mouse
- Visual Similarity of Pen Gestures

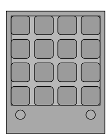
Two handed input

- Mouse plus something else
- Motivation for PDA as second device
  - Common, generic device
- Setup – PDA in left hand, mouse on right

Experimental Design

- Typing test
- Button size task
  - PDA Layouts 2x2, 2x3, 3x2, 3x4, 4x3, 4x4
- Homing speed task
- Scrolling task

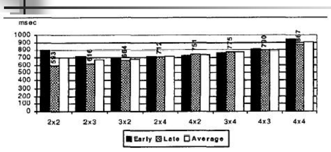




Button size test

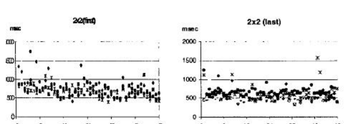
- 10 to 12 subjects (two discarded because of wrong handedness)
- Basic test – highlight button on screen, tap corresponding button with left hand
- 48 stimuli used for each layout (random order, each button appeared the same number of times)
- Order of layouts
  - 2x2 . . . 4x4 (half the subjects)
  - 4x4 . . . 2x2 (other half of the subjects)

Button Data



Button timings

Separated by early and late



Individual timings for 2x2 early and 2x2 late

Button Results

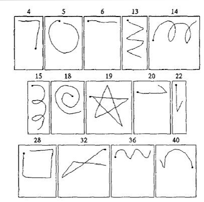
- Unintended
  - Learning required for button accuracy
    - Compare 2x2 first with 2x2 last
- Button time increases with number of buttons
  - If the second half of the data is used

Visual Similarity of Pen Gestures

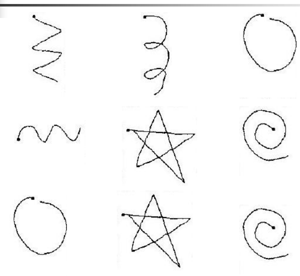
- Motivation
  - Gestural interface
  - Develop a collection of perceptually distinct gestures
- Question - when do people perceive gestures to be different
- The paper is looking at visual perception, not physical perception

The experiment

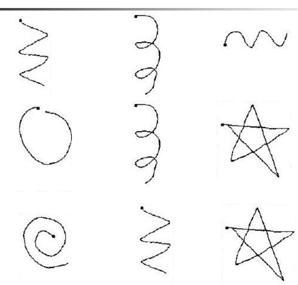
- Given three gestures, which one does not belong



Trials



Trials



Methodology

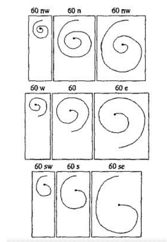
- 21 subject
- Training sets of all triads from a different group of 5
- Subjects picked distinct element from each of the triads of the 14 (364 triads)
- Random order
- Compute dissimilarity matrix

Analysis

- Create dissimilarity matrices
- Combine these to get a high dimensional structure
- Reduce dimensions by multi-dimensional scaling
- Look at correlation with predictors of similarity

Second experiment

- Develop separate gesture collections which differed along known axes
- Redo analysis



Lecture summary

- GOMS
  - Modeling at the task level
- Email modeling
  - Parameterized task complexity
- Human subjects review
  - Unavoidable, if humans are involved
- Two handed input
  - Learning, compromised experiment
- Gestures
  - High powered statistical analysis