Summary of Heuristic Evaluation of Prototypes (Group)

Due: End of class Tuesday, December 4, 2007

Goals

The goal of this assignment is to summarize the heuristic evaluations and assign severity ratings to violations so that next quarter the project teams can fix the most important problems found in their project prototype UI. A template file for your group evaluation can be found at:
http://www.cs.washington.edu/education/courses/cse440/CurrentQtr/assignments_files/PROJECTNAME_HE_prototype_group.doc

1. Problem Statement

One sentence description of the UI you are evaluating.

2. List of Heuristic Violations

You will be summarizing the HE reports along with the others who evaluated the same project.

I. List each distinct problem with a unique number.
> You may have to merge similar reports from different users. For example:

> report #1:
> **5. [H2-4 Consistency & Standards]**
> The interface used the string "Save" on the first screen for saving the user's file, but used the string "Write file" on the second screen. Users may be confused by this different terminology for the same function.

> report #2:
> **18. [H2-4 Consistency & Standards]**
> "Write file" and "Save" are used interchangeable in the interface. This is confusing.

> Those reports are listing the same violation. You should give it a unique number and only list it **once** in your summary report (with the **best** possible description). List the reviewer number of each reviewer who found the violation. This will help you with bookkeeping later.

II. Indicate the heuristic violated.

III. Assign it a severity rating based on the following ratings defined in lecture:
> **0 = not a problem, 1 = cosmetic, 2 = minor, 3 = major, 4 = UI catastrophe**

For the above example, you might write:
> **1. [H2-4 Consistency & Standards] [Severity 3]**
> The interface used the string "Save" on the first screen for saving the user's file, but used the string "Write file" on the second screen. Users may be confused by this different terminology for the same function.

Your report will list each of the problems found in the following format:
> **problem # [heuristic violated] [severity rating]**
> > description of problem and reasoning why it violates the heuristic

CSE440 Autumn 2007 Web site
http://www.cs.washington.edu/cse440

3 Summary of Violations

Your report will also summarize the number of violations found in each of the ten heuristic categories, the number of violations found in each of the severity categories, and give a total number of violations in the entire interface. Fill in the table provided for this in the template.

4 Evaluation Statistics

4.1 Your report will also give an analysis of how well each of the evaluators did (use reviewer #, not names) in finding problems in the interface. This will show you how well the technique works:

- Sum the total number of violations detected at each severity level
- Sum the number of violations each reviewer detected at each severity level
- For each evaluator and each severity rating, give the % of the total violations of that severity level found by that evaluator. *(e.g., A found one of the five "level 1" violations, so 20%)*
- For each evaluator, list the % of all violations found, and the % of level 3 & 4 violations.

For example:

severity\evaluator	reviewer A	reviewer B	reviewer C	reviewer D	reviewer E
level 0	10%	etc…			
level 1	20%				
level 2	5%				
level 3	10%				
level 4	10%				
total (levels 3 & 4)	10%	30%	20%	50%	80%
total (all levels)	25%	75%	80%	10%	40%

(note that the bottom rows are **not** calculated by adding the numbers above it.)

4.2 Graph

You should also draw a graph like that shown in the lecture on HE ("problems found" on the "Decreasing Returns" slide). Order the judges starting with the evaluator who found the most violations (from the final merged list). Then, if you remove all the violations that that person found, choose the next evaluator with the most violations. This is the "value added" of the second evaluator. Then the next evaluator with the most violations remaining that the first & second evaluators did not have... etc. So if you had the following reviewers with the listed number of violations found:

reviewer #	problems found	unique?
A	7	5 shared with #2, #4, #5 → 2
B	18	
C	5	all but one shared with everyone else → 1
D	9	3 shared with #2 → 6
E	6	1 shared with #2 & #4 → 5

your graph would plot the following points: (1, 18), (2, 18+6=24), (3, 24+5=29), (4, 29+2=31), (5, 31+1=32)

Deliverable

Your deliverable is the word document report. Your typed write-up should follow this outline with separate sections for the top-level items. It must be named appropriately and e-mailed to the teaching assistants at the end of class.

1. Problem (one sentence description of UI you are evaluating)
2. Violations found
3. Summary of violations
4. Evaluation statistics