

University of Maryland Department of Computer Science Spring 2005
CMSC 838S Information Visualization: Prof. Ben Shneiderman

Final exam: 2 hours, open book (papers), open notes, NO discussion, NO laptops

1) A linguistic summary of a social network could be a useful alternative to adjacency matrices and node-link diagrams. Give a brief argument in favor of linguistic summaries for social networks. Describe three features you might like to read about in a linguistic summary with examples of typical sentences. (20 points)

2) Imagine a database with height, weight, age, and gender (M or F) for 100 people. A scattergram uses height and weight as x and y axes, with marker size indicating age and marker color indicating gender (blue for M, red for F).

a) Describe an interactive information sonification program that would enable a blind user to query this database. (10 points)

b) Give two examples of tasks and how a user could accomplish them. (10 points)

3) Assuming a relational database with a

- person table having unique name, plus age, gender, and job title
- car table having unique Vehicle#, owner name, model, year

Explain your answer in each case

a) If the relationship between person and car is 1 to 1, describe and draw a visualization that will enable users to explore whether there is a relationship between the models owned, gender, and age. Assume 50 people and 50 cars. (10 points)

b) If the relationship between person and car is 1 to many, describe and draw a visualization that will enable users to explore whether there is a relationship between the models owned, gender, and age. Assume 50 people and 87 cars. Users should be able to select gender or age ranges and see which models are preferred; or select models and see which genders or age ranges are most common among owners. (10 points)

4) A new tree visualization program shows values at nodes by size and uncertainty of the value by gray-scale coding the nodes, with less certain values being lighter gray and more certain values being darker gray. But a critic suggests that gray-scale coding for value and size coding for uncertainty would be better, making uncertain items smaller. You have been asked to conduct a **controlled experiment** to provide evidence in this matter.

a) Draw examples of the two versions with 3 levels and 6-8 leaf nodes that you might use in this study. (5 points)

b) Describe how you would go about designing such an experiment: describe independent variables and hypotheses, and discuss choosing subjects, training, tasks, dependent variables, and statistical tests. (15 points)

5) (This is the challenging creative question) A computer science professor is working on an information visualization program to help assign final grades to students. The data for each student include five grades in the range 0-100: two exams (weights are .15 each), two projects (weights are .20 each), and class participation (weight is .30). Final grades are usually assigned as A (90-100), B (80-89), C (70-79), and F (below 70), but maybe there needs to be some room for flexibility.

a) Give two reasons why a straight algorithmic approach based on five grades might be inadequate or inappropriate. (10 points)

b) Draw an information visualization to help this professor assign fairer final grades A, B, C, and F. (10 points)