CSI 703/ IT 875
Mid-Term 1
Scope and Selected Answers

Parts in red will be excluded from this exam as they have not been covered or covered adequately. In a few cases answers may be provide here that were not covered in class but you are expected to know based their presentation here.

I mentioned the derivation of projection (available in the class notes) and stereo graphics as likely being on the exam in class. However I but did not get the topics in the lecture so have excluded the topics.

Lecture 2 and related materials on perception, cognition and graph design are particularly important.

1. Perception and Cognition

1.1 Know **Steven's law**.

Be able to give example exponents " b" in the proper intervals for
$p(x) = ax^b$
p(x) = perceived value
x = stimulus magnitude

x = length b in [.9 1.1]
x = area b in [.6 .9]
x = volume b in [.5 .8]
x = electrical shock b > 1

Consider comparing two areas, 4 and 1.
Do we judge the larger area to be less than 4
or more than 4 times bigger than the smaller area.

1.2 Understand **Weber's law** well enough to cite two examples of its application.
grid lines in a plot,
framed rectangles

1.3 Know **Cleveland's ranking of univariate encoding methods** based on perceptual accuracy of extraction as presented in class.

Best = position along a scale, aligned nor not
Good = Length and Angle
Poor = about everything else.

1.4 Be able to cite 4 **perceptual grouping principles**
Five of more are: Proximity, Similarity, Good continuation, Common fate, Enclosure

1.5 Know the four **general graphics design objectives** stated in class. Some overlap and some conflict.

- Encourage accurate comparisons
- Provide context for interpretation
- Strive for simple appearance
- Involve the reader

1.6 Does closer **proximity improve visual comparison accuracy?**
Yes. Sorting can be very helpful in bringing similar values color together where they are easier to compare accurately.

1.7a What are three basics approaches to comparative layouts?
Juxtaposition, superposition, computed differences (or ratios)
1.7b Are we good visually assessing the difference between two superimposed curves?
In what direction does our eye-brain system assess the difference?

1.8 Know about layering information in depth
Consider estimates, confidence intervals, and grid lines
What goes in front? What goes in in the middle? What goes in back?

1.9 List four depth cues.
What two cues are most relevant to the 2-D scatter plots used in this class?
What are three perceptual/cognitive issues related to the use depth cues in scatter plot construction?

1.9 Our working memory is quite limited.
What constitutes a chunk of information can vary.
We are adept at working with only ____ chunks of information and
sometimes only _____.

1.10 Visual working memory only handles about ___ chunks of information.

During eye saccades (eyes jump to focus on a new location) we are pretty much blind.
Our iconic visual memory in retinal coordinate fades in about 200 milliseconds.
By 40 milliseconds about all that is left of working memory.

A consequence of this is _____ blindness when we are comparing two side by side images or sequential images separated by say a gray flash that fires off a lot of our change detectors.

1.11 Some things are discriminated preattentive vision.
People don't have to work at this, it is automatic, working in parallel and very fast.

Historically Bela Julesz referred to these features as textons and cited items such as

color, angle, endpoints, crossings.

More recently Colin Ware provides conditions on preattentive color discrimination and there may be more current results in several areas.

List four features preattentively processed: See the lecture notes for a long list.

For the present just remember a few results cited in Kosslyn,
A dot plotted on a line should have _____ the diameter as the line width (twice)
Angles should differ by ___ degrees. (30)
And from other sources: good discriminable symbols are: o, +, and x.

2. Graphics

2.1 Know how to construct **quantiles plots**.

2.2 Know how to construct **QQplots** including details of interpolation and extrapolation.
Be prepared to construct a QQplot comparing two samples

2.3 Know how to interpret QQplots.
Departures from a straight line indicate discrepancies in what moments? _____
If a straight line fit is a reasonable fit to points in a normal QQ plot
Is the data necessarily from a normal distribution? _____
What does the intercept estimate? _____
What does the slope estimate? _____
Know how the obtain these estimates from a QQnormal plots without seeing the regression coefficients.
Be able to assess tail thickness in a normal QQplot
If the data values are positive, be prepared to propose a power transformation to reduce the thickness of one tail

2.3 Know how a **box plot is constructed**.
For example, know how adjacent values are determined.

2.4 Know how a construct a mean difference **(MD) plot**.
Know what motivates the transformation and how to interpret the plot.

2.5 Know how to construct a spread location **(SL) plot** and how to interpret it.
Also be prepared to select a power transformation that would stabilize the spread.

2.6 Know how to construct a residual fit **(RF) plot** and how to interpret it.

3. Smoothes

3.1 If using loess, know from the scatterplot when to use a **local linear or a local quadratic fit**.

3.2 Be able to decide if a fit is way **too rough or too smooth**.

3.3 Know to look for **misfits at peaks and valleys**.

4. R Syntax

4.1 Be able to **write a function** that has both required arguments and arguments that default.

4.2 Know about **distribution functions**.
What do the prefixes q, p, r, and d stand for?
What are the required arguments for
qnorm(), pnorm(), rnorm() and dnorm()
Select from: cumulative probabilities, quantiles, sample size and distracting answers
What are the distribution parameters for the
uniform, normal and t distributions.
Know the names of 2 additional families of distributions
For example: weibull, gamma

4.3 Arguments of basic graphics functions

What are the argument key words to plot()
for the title, the x-axis label and the y-axis label

What are the argument key words to points() that control
character type, color, and size

What are the argument key words to lines() that control
the line width, dash style and color.

What are the arguments key words to polygon() that control
fill, outline and color. What three things happens when
one of the x or y values is NA.

Two argument key words to axis() are “at” and “label”. What do they control?

4.4 Subscripts
Know how to select subsets from a vector or a data.frame()
Know how to select elements from a list using the naming convention.
With two commands show how to sort the rows of a 7 column data.frame based
on the values in its third column, and breaking ties with values in the fifth column.
What additional function will change the order from ascending to descending.

4.5 Panel Layout Functions from your instructor
Know how the use the panel.layout function to define a matrix of panels
with given relative row and column sizes, separations in inches, and margins.

Know how to select a panel for plotting.
Know how to set the scale limits for a panel.

4.6 Know how to convert a continuous value into ordered categories using cut()
colvector <- cut(dat, 10,include.lowest=T) # one variation

4.7 Know how to quickly define a regular grid of rectangles using
expand.grid and rect().

```
dx= .5; dy=.5
nx=12; ny=9
mat = expand.grid(list(x=1:12,y=1:9) )
x = mat[,1]
y = mat[,2]
xlim = c(.5,nx+.5)
ylim = c(.5,ny+.5)
plot(xlim,ylim,type='n')
rect(x-dx, y-dx, x+dx, y+dx)
```

5. CCmaps

5.1 Be able to describe the grid search for a large R^2 values in a very few sentences.

5.2 What happens if there are missing data
in the data set?
in the variables to be used in the maps or plots?

5.3 Polygon id and case id mismatches.
What happens if there is no data for a polygon?
What happens if there is no polygon for a case?

5.4 What two features would you add to CCmaps to make it more useful?
What two features would you fix or change to make CCmaps better?
Reasonable answers are accepted.

6. Row-labeled Plots

6.1 List three methods for sorting cases with continuous multivariate observations.
Sorting by: the median, the first principal component, a minimal spanning tree traversal.

6.2 Should the variables generally be scaled to be unitless or in comparable units before ordering multivariate values using the median?

6.3 Is it possible to represent as many as 150 variables on a page? Yes

6.4 Be prepared to indicate graphics guidelines general and specific followed in row-labeled plots

construction. This includes linked micromap plots.

7. CrystalVision

7.1 In the 3-D plot does the 3rd variable in the selection list go on the z-axis?
Are the smallest values for the x-axis on the left?

7.2 Crystal Vision has alpha blending. With the blending on, what trick helps low count pixels to become readily evident?

8.0 Projection and Stereo Graphics

9.0 Grand and Random Tour