Prof. Dr. Thomas Schultz
Mohammad Khatami, Ikram Jumakulyyev

University of Bonn
B-IT / Computer Science II
July 3, 2020

Trial exam for the lecture "Visual Data Analysis"

Summer Term 2020

Note:
You do not have to hand in your answers; we will discuss them during the exercise session on July 15.

Rules for the final exam – please read in advance:

- In case of attempted cheating, the exam will be immediately rated with 0 points without a warning.
- No auxiliary materials except writing utensils and a single A4 sheet of paper with hand-written notes are allowed. If you need paper for sketches or drafts, you can get empty paper from the supervising tutors, which has to be returned to us after the exam. Do not use your own paper.
- Please try to fit all your answers into the provided fields, or on the back of the paper. If you cannot avoid using additional paper, clearly state your name and student ID, the number of the task, and the text "Please grade" *at the top of the page.* Otherwise, we will assume that all additional sheets of paper have been used for sketches or drafts only, and we will throw them away ungraded.
- Do not detach the individual sheets of the exam from each other.
- If multiple solutions are provided for one task and it is not clear which of them you would like us to grade (e.g., by clearly striking out alternative solutions), the task will not be evaluated, and no points will be given.
- Use only permanent pens. Do NOT use red pens.
- Do not forget to clearly mark the exam with your name and student ID.
- Please write clearly. No points will be given for non-readable solutions.
- You will have 120 min time to write the exam.
- You will need 60 points for passing (i.e., 50% without counting the bonus task).

Task	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	$\sum$
max. points	17	14	13	16	15	15	15	15	15 (bonus)	**120**
score										

Task		**17 Points**
	Name:	(4+3+5+5)
1	Student ID:	Σ: ....

Visualization Design and Color Vision

a) Name two visualization channels that are expressive for ordered attributes, and rank them according to their effectiveness. Name another channel that is expressive for categorical attributes.

b) Is the rainbow (spectral) color map well-suited to encode ordinal data? Why?

c) Briefly explain what is meant by the term "lightness constancy". Name and briefly explain two mechanisms that contribute to it.

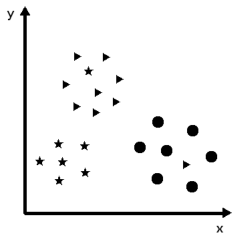
d) Sketch a CIE chromaticity diagram and mark the following features:

- Spectral locus
- Purple boundary
- D65 whitepoint (approximate location)

Task		14 Points
2	Name: .... Student ID:	(3+4+7) Σ: ....

Multidimensional Data Visualization

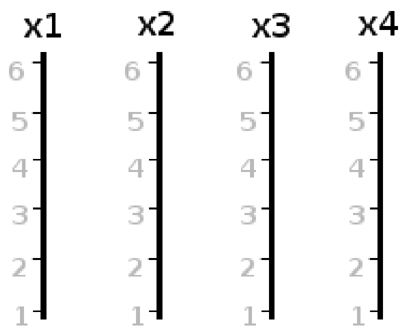
a) For the following scatter plot of 23 data items, belonging to 3 different classes, please compute the distance consistency.



b) Briefly name and explain an approach to facilitate navigation in scatterplot matrices.

c) For the following data set, please draw the parallel coordinates on the given axes and point out a pair of variables with a high **negative correlation.** Also, please comment on the importance of interaction when using parallel coordinates.

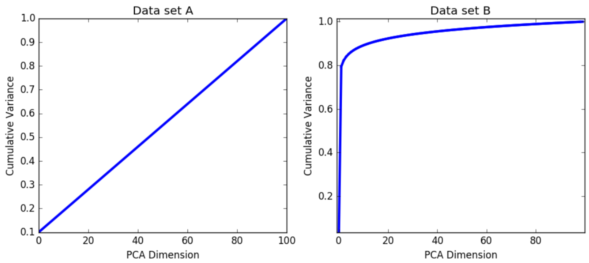
Data-ID	x_1	x_2	x_3	x_4	$class$
a	6	1	4	4	0
b	5	2	5	2	0
c	4	3	1	1	1
d	3	4	5	5	0
e	2	5	2	4	1
f	1	6	6	6	0



Task		13 Points
3	Name:	(3+4+6)
	Student ID:	Σ: ....

Dimensionality Reduction

a) Given the cumulative sum of the eigenvalues of the PCA analysis of data sets A and B, state for which data set it makes more sense to use linear dimensionality reduction and briefly explain why.



b) Name two advantages and two disadvantages of Kernel PCA compared to standard PCA.

c) State the definition of perplexity and explain how it relates to entropy. What is the relevance of perplexity in the context of t-SNE dimensionality reduction? What happens if it is set to 50 while applying t-SNE to a data set of size 40?

Task		16 Points
4	Name: .... Student ID:	(6+4+6) Σ: ....

Visualizing Graphs and Neural Networks

a) Name and sketch three main strategies for visualizing trees.

b) Briefly explain an advantage of occlusion sensitivity maps over class activation maps for visualizing the input-output relationship of neural networks for image classification. Also explain an advantage of class activation maps over occlusion sensitivity maps.

c) When visualizing class models with gradient ascent, regularization is important to achieve more natural and recognizable images. Name two possible regularizers, and briefly explain their effect.

Task		15 Points
5	Name:	(5+5+5)
	Student ID:	Σ: ....

Foundations of Scientific Visualization

a) Sketch the visualization pipeline and label its most important stages.

b) Name an advantage of structured grids, as well as an advantage of unstructured grids.

c) Briefly explain an advantage of using homogeneous coordinates.

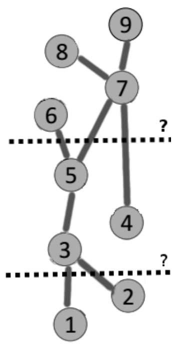
Task		15 Points
6	Name: .... Student ID:	(5+5+5) Σ: ....

Isosurface Extraction

a) Briefly explain the difference between the midpoint and the asymptotic decider in the Marching Cubes algorithm, and provide an example in which such a decider is needed.

b) Given a set of intervals on the real line, interval trees are a data structure that can be used to efficiently retrieve all intervals that contain a specific value. Suggest how such a data structure could be used to accelerate the Marching Cubes algorithm.

c) Briefly explain how a contour tree can be useful in isosurface visualization. In the contour tree shown below, what is the number of connected components at the iso-levels marked with the dashed lines?

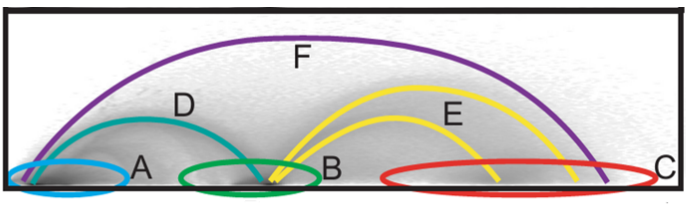


Task		**15 Points**
7	Name: .... Student ID:	(5+5+5) Σ: ....

Direct Volume Rendering

a) Name and briefly explain a method for numerically computing the volume rendering integral.

b) 2D transfer functions can be used to volume render material boundaries. What quantities are mapped onto the two axes of the corresponding 2D histogram? What do the arcs in the image indicate?



c) Briefly explain the concepts of pre- and post-classification in direct volume rendering. How does pre-classification differ from pre-segmentation?

Task		**15 Points**
8	Name:	(5+5+5)
	Student ID:	Σ: ....

Vector Field Visualization

a) Name two types of characteristic curves in an unsteady vector field and briefly explain their difference.

b) Briefly explain the advantage of an embedded Runge-Kutta integration scheme over regular Runge-Kutta integration.

c) In the context of stream line illumination, we have seen the equation $\sqrt{1-(L \cdot T)^2}$. Please state what role this equation plays in this context, and what L and T stand for.

Task		15 Points
9	Name:	(7 + 4 + 4)
	Student ID:	Σ: ....

BONUS TASK: Fast LIC

a) The following pseudocode for fast LIC is incomplete. Please fill in the missing parts:

```
For each pixel p in output image


   Initialize streamline computation with x0 = center of p
   Compute convolution I(x0)
   Add result to pixel p


   For m = 1 to Limit M
     Incremental convolution for I(x[m]) and I(x[-m])
     Add results to pixels containing x[m] and x[-m]


   End for

End for
```

b) Briefly explain what makes fast LIC faster than standard line integral convolution.

c) Briefly explain how LIC can be modified to convey the orientation of the vector field.