Tool for Teaching Visualization Techniques: Learning and Homework Assignments for Multivariate Data Visualization

Tatiana von Landesberger[1], Felix Brodkorb[1], Philipp Schneider[1] and Kathrin Ballweg[1]

Abstract— This paper presents a prototype tool for teaching information visualization techniques in a university class. It addresses a need for learning students how to create and compare various visualization techniques for a particular data type. Moreover, it deals with the challenge of creating interactive homework assignments.

The tool consists of two parts: a learning tool and a tool for assigning interactive homeworks to students. In the learning tool, the students get familiar with several visualization techniques for the same data type (e.g., parallel coordinates, scatterplot matrix and RadViz for multivariate data). They learn the mechanics and advantages and disadvantages of the techniques. The homework tool allows the teachers to create interactive homeworks for students (e.g., "change axis ordering in parallel coordinates"). It automatically grades the submissions and, thus, reduces the time cost of correcting assignments. The tool was used for teaching multivariate data visualization in a visual computing class with approximately 400 students.

Index Terms—information visualization, teaching, homework assignments, pedagogy of data visualization

---

1 INTRODUCTION

This paper presents a tool that we developed for teaching visualization techniques. Our information visualization lectures teach students techniques for various data types (multivariate data, hierarchies, etc.). For each data type, several techniques are explained. For example, the lectures describe parallel coordinates (PCP) [8], scatterplot matrix (SCM) [5, 12] and RadViz [6] for multivariate data. We wish to teach the students how to create and read the visualizations. We wish to illustrate the correspondences among these visualizations, when showing one dataset. It demonstrates them advantages and disadvantages of these techniques. We also teach students how to use interaction to overcome the limitations of a particular technique (e.g., how to use axis reordering to solve object ambiguity in parallel coordinates).

Our class is composed of lectures and practical exercises. The lectures explain various ways of data visualization for each data type (multivariate data, hierarchies etc). During practical exercises and weekly homework assignments, the students create visualizations of a dataset with various techniques and compare the results in order to learn about the pros and cons of these techniques.

In recent years, our class attendance has grown from approximately 20 students to approximately 100 students. The large size of the class and various student majors (computer science, economics, or psychology) pose challenges on teaching visualization and on correcting the homework assignments. In the past, the students were allowed to use any available tool (e.g., Tableau, or pen and paper) to do the homeworks. They submitted the results in Moodle as PDF documents. This resulted in a time consuming and cumbersome correcting of a large amount of assignments. It took up to 2 days per week. Moreover, students complained about spending a lot of time learning how to use an external tool instead of solving the task. We also observed that students had difficulties in assessing correspondences in visualizations and getting deeper knowledge about pros and contras solely from listening to lectures and solving the few tasks in the homeworks.

Therefore, we were looking for new ways how to allow students to get a deeper knowledge of the techniques outside the class and how to lower the burden of creating and correcting homeworks.

There are many sources of information visualization literature for teaching such as books, papers, and free online courses, e.g., [4,7,9–11], InfoVis wiki, lectures by Fraunhofer IAIS, a course in Oxford, UK, or course by TU Wien. They provide information about course content, not tools for teaching. The available tools for data visualization (e.g. Tableau, R, Excel) fitted to our needs only partly as they only allow to create visualizations, not to correct the submissions.

Therefore, we developed a tool that unifies data visualization learning with homework assignments and their correction. The tool has two parts: 1) a learning tool which shows a data set in interactive linked views with various visualization techniques. 2) a tool for homework assignments and correcting responses which allows for interactive homework solutions. It offers pseudo-randomized homework assignments and it automatically grades the responses.

The first prototype has been employed for teaching multivariate data in a beginners class of 400 students. The students could use the tool learn about three multivariate visualization techniques: parallel coordinates [8], scatterplot matrix [5, 12] and RadViz [6]. They had to solve several tasks within a homework assignment. The tool enhanced the knowledge of visualization techniques (approx. 90% correct answers with the tool compared to approx. 70% without). Moreover, it reduced the time cost of correcting assignments from two days to 1 hour inclusive filling in the grades to Moodle. The tool is still in a development. We plan to extend it with more techniques and homework types.

In the following, we describe the two parts of the tool and show exemplary homework assignments with the tool. We report on the usage and evaluation of the tool in a real world setting.

2 LEARNING TOOL

The learning tool (see Figure 1) shows a user-defined dataset (1) using various visualization techniques in linked interactive views. These three views teach students how the same data looks in various visualizations. They allow them to assess the pros and contras of the techniques.

Our tool prototype currently focuses on the visualization of multivariate data. The students can edit the dataset and see the results concurrently in Radviz, in a parallel coordinates plot and in a scatterplot matrix with one selected scatterplot in larger detail. The views are linked so that the students can see the correspondences among the visualizations. A user-selected object is highlighted in all views and students can change its value by dragging. While dragging in one view, they can watch how the point changes in other views. Moreover, the students can change settings of individual visualizations. For example, they can learn about the effect of relative and absolute scaling.

The web-based tool is available at the learning tool website. Visualizations are in d3js [2], for example, parallel coordinates view uses Parallel Coordinates library.

---

[1]*Interactive Graphics Systems, Technische Universität Darmstadt, Germany*

1: Data input

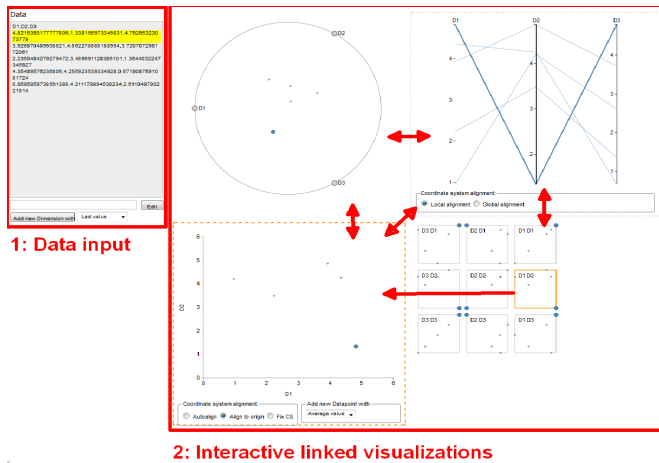**2: Interactive linked visualizations**

Fig. 1. Learning tool. 1: Data input 2: Interactive linked views.

3 TOOL FOR HOMEWORK ASSIGNMENTS

The tool for homework assignments allows the teachers to create homeworks for students and allows the students to interactively solve the assignments (see exemplary assignments in the next section). The submissions are then automatically graded.

The teacher can create a pool of assignments programmatically or using an exercise template (see next section). Then, she can decide which assignments should be solved in a particular homework. She can also assign the deadline for solving the task (e.g., 20 June 2016 5PM). These assignments are shown in homework overview viewed by the students (see Figure 2, each task is one row). After signing in with their university ID and password, the students see the task list.

Fig. 2. Overview of tasks in a homework assignment.

The task list shows which tasks are already solved and which still need to be solved ("noch nicht abgegeben"). It also includes a test task ("Probeaufgabe"), which students can solve to learn how to use the tool. The test task is not graded.

The students solve the tasks online. By clicking on a task, the task is shown and the students can solve it interactively (see Fig. 3). After solving the task, the students can submit it. Then the task is automatically graded. The result list is imported to Moodle, where the students can see their grades.

4 EXEMPLARY HOMEWORK TASKS

The tool currently supports two types of tasks:

- interactive tasks, where the students need to complete a task by interacting in the visualization (see Sec. 4.1)

- correspondence tasks, where the students need to match a visualizations of the same dataset in two techniques (see Sec. 4.2)

An important tool feature is that the students get randomized data within tasks. In this way, the students cannot pass the solution to their friends or student colleagues. This mitigates the problem of cheating, when having one assignment for all students. We needed to implement this feature, as cheating, unfortunately, occurs quite often otherwise. The randomization depends on the task type. It is explained below.

4.1 Interactive Tasks

In interactive tasks, the students have to solve a task by interacting with the visualization. For example, they should highlight specific points or they should change the visualization (e.g. by axis reordering). Interactive tasks should support gaining a deeper understanding of the particular type of visualization.

From a teacher's point of view, these tasks require an adjustment of the visualizations in the learning tool (e.g., restricting the interaction to highlighting only) and determining the dataset to be shown.

Two exemplary tasks are implemented in the tool currently. More tasks can be implemented in a similar vein.

Task Type 1: Axes reordering
Task: The student has to reorder axes in parallel coordinates (PC) so that the data objects are uniquely identifiable.
Input & randomization: The tool shows the student a randomized dataset where the values of all objects on one random axis are the same. The values on other axes are random.
Explanation and Solution: It is well known that object crossing on an axis in PC disables the identification of individual objects (see Figure 3a). Sorting the crossing axis to the very left or right solves this problem (see Figure 3b).
Task alternative: Alternatively, the task of axes reordering can be posed in RadViz visualization . Then, the students should reorder axes to resolve object overplotting. This will be implemented in future.

Task Type 2: Highlighting of the corresponding objects between visualizations
Task: Select objects in parallel coordinates view, that are overlapping in the middle of RadViz.
Input & Randomization: The dataset has several points in the middle of RadViz. They have random equal values in all dimensions. Other points in the dataset are fully randomized.
Explanation and solution: The student needs to know that objects with equal values in each dimension are mapped to the middle of RadViz. Thus, they need to highlight them in parallel coordinates.
Task alternative: This task can be modified by varying the data value structure. For example, objects with values in all but one dimension equal zero are shown on the periphery (at the axis anchor) of RadViz.

4.2 Multiple Choice Visualization Correspondence Tasks

The tool offers a template for multiple choice tasks, where the students have to identify the correspondences between two visualizations of the same dataset. This teaches the students to read various visualizations, to identify correspondences among them and to see the advantages and disadvantages of the techniques.

Task Type 3: Identify the corresponding visualization
Task: The students see a data visualization using one technique (e.g., PCP) and three visualizations using another technique (e.g. SC, SCM or RadViz). The students have to identify, which of the three visualizations correspond to the first visualization (see Figure 5). This task usually contains several subtasks of the same type (e.g., 3 subtasks with different datasets).
Input: Each task shows a user-defined dataset and a set of visualizations. This is usually created by the teacher in the learning tool. The learning tool allows to easily create datasets, the correct solutions, and alternative visualizations in an interactive way.
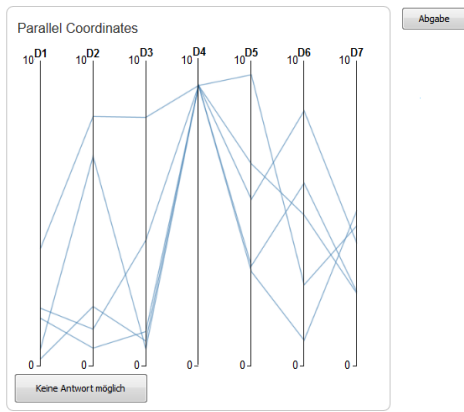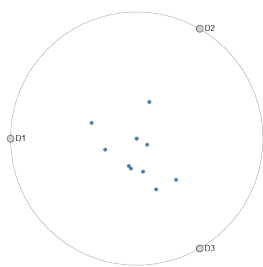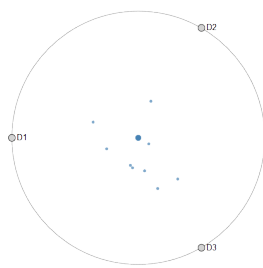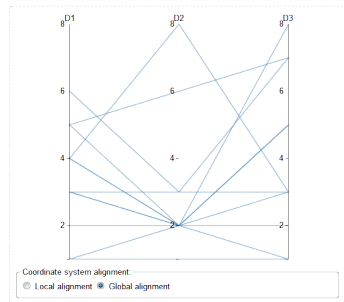
(a) Task: order axes
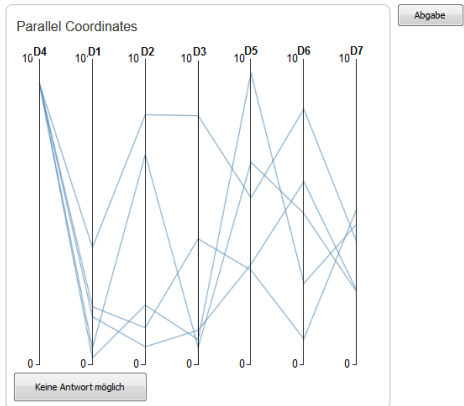
(b) Solution by interaction with an axis

Fig. 3. Interactive solving of the task to order axis in parallel coordinates. "Abgabe" button submits the solution.

(a) Task: Highlight overplotted objects in middle of Radviz in PC plot

(b) Solution

Fig. 4. Interactive task of highlighting overplotted objects in middle of RadViz in the corresponding PC plot.

Randomization: The order of subtasks, as well as the order of the alternative visualizations, is randomized for each student. The datasets are stable.

Solution: The student has to select the correct corresponding visualization in a multiple choice field below the visualizations. The student can also determine "no corresponding visualization" ("keine der angegebenen Darstellungen") as the correct solution.
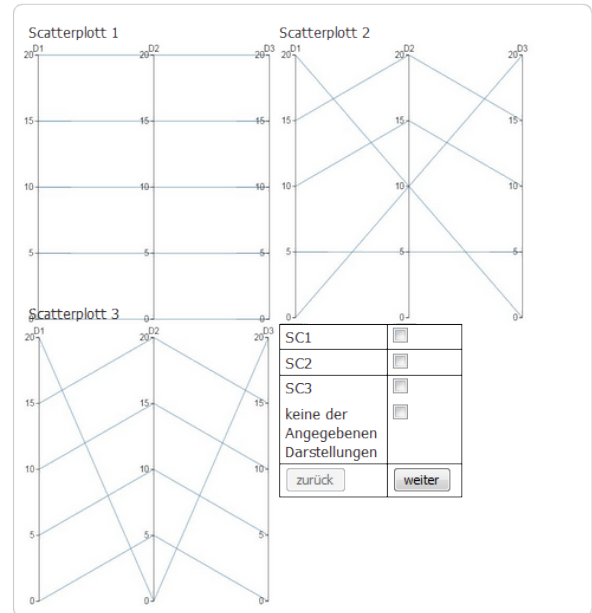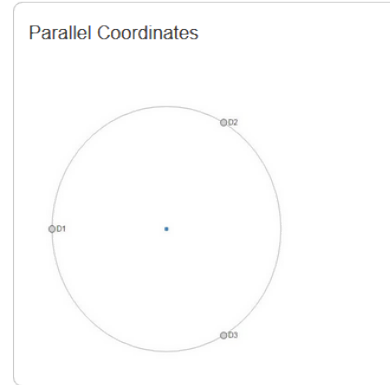

Fig. 5. Example task of identifying the corresponding visualization of the dataset in RadViz and PCP.

Task Type 4: Identify the partly corresponding visualization

Task: The students see a full data visualization using one technique (e.g. PCP) and a three 'partial' visualizations using another technique (e.g. SC), where only a subset of data dimensions is displayed. The students have to identify which of the three visualizations correspond to the first visualization. This is an advanced version of Task 3, as several alternative visualizations can be correct. The students have to identify all of them.

Input, randomization, and solutions are the same as in Task 3.

5 PRACTICAL APPLICATION AND STUDENT FEEDBACK

The current tool prototype has been applied to teaching and homework assignment withing the bachelor course "Visual Computing", the lecture on information visualization. The course is attended by approx. 400 students, who are assigned to groups of 4 students. Each group has to solve weekly assignments.
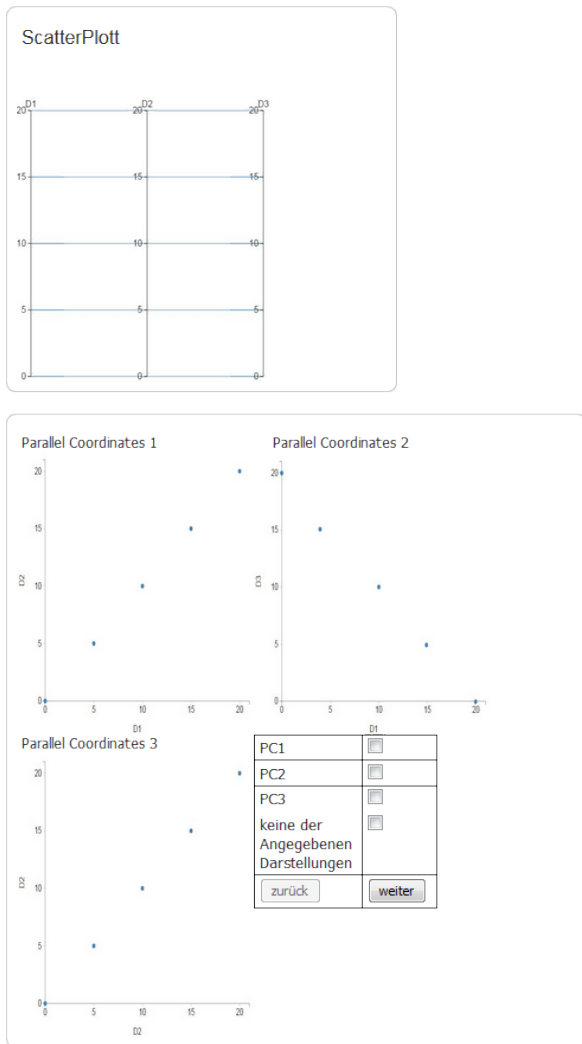
Fig. 6. Example task of identifying corresponding visualization of the dataset in PCP and SC, which shows only a subset of dimensions.

The teaching team (the lecturer and two student assistants) assigned the students a homework in the tool It contained one interactive task of axis reordering in PCP (Task 1) and three correspondence visualization tasks (Task 3). The students were also asked to give us voluntary feedback about their experience with the tool. We posed them questions about the tool usability according to the System Usability Scale [1, 3].

81 student groups used our tool to solve their homework assignment. We were positively surprised by the high accuracy of the answers. The tasks were correctly solved by more than 97% of students. This is a significant increase with respect to previous years, where the students had to solve the tasks without the tool (reached on average only 7.3 out of 9 points). This increase can be attributed to various reasons, e.g., learning effect with the tool, or the question types.

27 student groups answered questions about the tool's usability and one group gave a free text feedback. The usability scores showed that the system was easy to use (mode = 4 of 5 points) without the need to learn how to use it (mode= 5 of 5 points). Due to several bugs in the prototype, the students felt rather uncertain using the tool and thus also the feedback on the future regular use was divided into positive and negative (about half positive and half negative, mode=3). The students wished to have a possibility not only to submit the solutions but also to save the current state of solving the task for future. This was especially asked for tasks with several subtasks. Some students had difficulties solving the tasks in Debian, while it worked well on Windows or MacOS using Firefox or Safari. We hope that solving the prototype bugs and improving the tool with students' feedback will improve the incentive for tools usage.

6 CONCLUSIONS AND FUTURE EXTENSIONS

This paper has presented a tool for teaching students various visualization techniques and the correspondences among them.

We plan to extend the tool in three directions. First, we wish to simplify task creation for the teacher with new task templates and improve the functionality of the tool based on students' feedback. Second, we wish to integrate more multivariate visualizations such as star coordinates. Third, we wish to extend the tool with techniques for other data types such as hierarchic data and their visualization techniques (treemap, icicle plot, sunburst, a node-link diagram with various layouts and an adjacency matrix).

REFERENCES

[1] A. Bangor, P. T. Kortum, and J. T. Miller. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, 24(6):574–594, 2008.

[2] M. Bostock, V. Ogievetsky, and J. Heer. D^3 data-driven documents. *IEEE transactions on visualization and computer graphics*, 17(12):2301–2309, 2011.

[3] J. Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[4] S. K. Card, J. D. Mackinlay, and B. Shneiderman. *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.

[5] D. B. Carr, R. J. Littlefield, W. Nicholson, and J. Littlefield. Scatter-plot matrix techniques for large n. *Journal of the American Statistical Association*, 82(398):424–436, 1987.

[6] L. Di Caro, V. Frias-Martinez, and E. Frias-Martinez. Analyzing the role of dimension arrangement for data visualization in radviz. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 125–132. Springer, 2010.

[7] N. Elmqvist and D. S. Ebert. Leveraging multidisciplinarity in a visual analytics graduate course. *IEEE computer graphics and applications*, 32(3):84–87, 2012.

[8] A. Inselberg and B. Dimsdale. Parallel coordinates. In *Human-Machine Interactive Systems*, pp. 199–233. Springer, 1991.

[9] A. Kerren, J. T. Stasko, and J. Dykes. Teaching information visualization. In *Information visualization*, pp. 65–91. Springer, 2008.

[10] T. Munzner. *Visualization Analysis and Design*. CRC Press, 2014.

[11] C. Ware. *Information visualization: perception for design*. Elsevier, 2012.

[12] P. C. Wong and R. D. Bergeron. 30 years of multidimensional multivariate visualization. In *Scientific Visualization*, pp. 3–33, 1994.