

Tommi Kärkkäinen

Timo Aittokoski

Olli Hokkanen

A-KIT-VISUALIZATION SYSTEM DEVELOPMENT PROCESS

InBCT 2.4

Agora Learning Laboratory

17.4.2003

Tiivistelmä: Tekstiä...

Abstract: The A-KIT visualization system is used to process questionnaire data with specific mathematical algorithms and to analyze the results in a three-dimensional view. The purpose of the 3D-view is to find the most defining factors of the data set, e.g. separate data points or groups.

Contents

1	Introduction.....	1
2	Project summary	2
2.1	Initial briefing	2
3	Application instructions and help files.....	3
3.1	Instructions.....	3
3.1.1	Initial Instructions	3
3.2	Help files.....	5
4	Testing process	6
4.1	Requirement specifications.....	6
4.2	Use Cases	6
4.3	Usability tests.....	7
4.4	Testing plan.....	7
	References.....	8
	Appendixes	9
	Appendix 1. Title	9
	Appendix 2. Title	9

1 Introduction

The purpose of this report is to find concrete approaches to improve usability and functionality of the A-KIT visualization application developed in Agora Center's Learning Laboratory (ALL) project. This report concerns the visualization software, not the whole questionnaire creating process functionality. In this report 'A-KIT' means the A-KIT visualization application, unless stated otherwise.

At this point (April 2003) the prototype version of the software is working as planned. The aim is to create e.g. technical documentation, help files and instructions for the software. Requirement specifications and use cases are used to evaluate and improve the software usability and suitable software development methodologies are used to evaluate the application structure. After analyzing all aspects of the process some interfaces of the prototype are rebuilt, making the A-KIT system a powerful tool for analyzing questionnaire data. A-KIT system can be used later on to support research made in other InBCT-projects.

2 Project summary

The A-KIT is used to process questionnaire data with specific mathematical algorithms and to analyze the results in a three-dimensional view. The purpose of the 3D-view is to find the most defining factors of the data set, e.g. separate data points or groups.

2.1 Initial briefing

At a project meeting 3.4.2003 Tommi Kärkkäinen summarized the objectives of the development process and the concrete tasks for the future. The first ‘proof-of-concept’ initiation of the whole A-KIT system is working as planned. To improve the A-KIT application and to make the development process ‘visible’ e.g. following things were discussed.

A-KIT

- Technical documentation (object interfaces etc.)
- Usability documentation (testing, use cases, help files)
- Methods documentation

It was also decided to install a CVS (Concurrent Versions System) application on Agora Center’s server to keep the software code and all reports up-to-date and available to all participants and also to make it easy to retrieve older files if necessary.

3 Application instructions and help files

The application needs both instructions and 'on-line' help files and dialogues to achieve acceptable usability level. These are developed along with the rest of the tasks.

3.1 Instructions

At this point (February 2003) there exists only one quite informal version of instructions how to use the A-KIT software.

3.1.1 Initial Instructions

Compressed data is stored in files (file extension .dat). Three-letter identifier in the file name indicates the method that has been used for the 3D-dimension compression. The following methods are used:

- PCA – Principal component based analysis (CiscoPCA.dat)
- MED - Median based analysis (CiscoMED.dat)
- SMD - Spatial median based analysis (CiscoSMD.dat)

3.1.1.1 Main window

3d-näkymän tarkoituksena on pyrkiä löytämään datajoukon kuvaavimpia piirteitä, kuten pääjoukosta erillään olevat pisteet tahi ryhmittymät. Alkuperäinen datajoukko kuvataan koordinaatistoon siten, että alkuperäisen datan huomattavan suuri dimensioiden määrä on pudotettu kolmeen. Alkuperäisessä datassa dimensioiden määrä vastaa kymysten määrää, kun taas muunnoksen jälkeen dimensiot vastaavat datajoukon merkittävimpiä piirteitä. Kolme merkittävintä dimensiota kuvataan kolmiulotteiseen koordinaatistoon.

After opening the file a three dimensional view window of the data set is opened. The orbs represent individual respondents and the cubes represent specific values of the data set. The cubes on coordinate axes are three primary components of the data set, other cubes are spatial median, median and mean value.

The purpose of the 3D-view is to find the most defining factors of the data set, e.g. separate data points or groups. In the original data set the number of dimensions is equal to the number of questions. After data transformation the number of dimensions is reduced to three, which represent the most significant features of the data set. These dimensions are illustrated by coordinate axes.

Pallot ovat yksittäisiä vastaajia ja kuutiot ovat datajoukon erityisarvoja. Akseleille sijoittuneet kuutiot ovat kolme tärkeintä pääkomponenttia. Muut kolme kuutiota ovat spatiaalinen mediaani, mediaani ja keskiarvo.

The data set view can be rotated with mouse. Buttons ‘XY’, ‘XZ’ and ‘YZ’ can be used to rotate the data set to main coordinate levels. Zoom buttons enable zooming in and out of the view.

3.1.1.2 Respondent data view

A specific respondent’s answers can be examined by clicking one of the orbs representing respondents in the main window. This opens a window where that respondent’s answers are shown as a bar chart. Numbers beneath the bars indicate question numbers.

Ikkunan alosassa näkyvät vastaajan tiedot. Tiettyä kysymystä voidaan tarkastella painamalla sitä vastaavaa pylvästä. Tällöin avautuu kysymyksen vastausjakaumaa kuvaava histogrammi sekä kysymysteksti. Useita kysymyksiä voidaan tarkastella samanaikaisesti. Kysymyksen voi poistaa histogrammiosiosta painamalla sitä uudelleen.

The lower part of the window shows information about the selected respondent. A certain question’s data can be examined by clicking that question’s bar chart. This opens a histogram representing question’s response distribution.

Multiple data points (i.e. respondents) can be examined at the same time; the colour of the bar indicates which data point is represented in the chart. Selecting two questions from the question history diagram opens a scatter plot –view, which shows the correspondence of the selected questions.

3.1.1.3 Correlation matrix window

Clicking the ‘Correlations’ button opens the correlation matrix associated with the data set. In this matrix the number of rows and columns indicate the question number. The upper triangular matrix indicates the correlation of all questions numerically, meaning that the rows and columns in the matrix represent question data correlation. The lower triangular matrix represents the significance level, which is indicated by a number of stars in the matrix cells.

The histogram in the lower part of the window represents the distribution of correlation coefficients at different magnitude levels. The number of bars in the graph can be adjusted by right-clicking mouse button on the histogram and inputting the wanted number of bars.

Clicking on a coefficient factor’s numeric value in the correlation matrix turns the corresponding bar, coefficient factor value and the significance level, if applicable, green, showing the correspondence of that correlation coefficient in the histogram.

Ikkunan alaosassa olevassa histogrammissa on esitetty korrelaatiokertoimien jakautuminen niiden vaihteluvälillä. Painamalla kahden kysymyksen välistä korrelaatiokerrointa

esitetään sen sijainti histogrammissa vihreällä värillä. Jos ko. korrelaatiokertoimen merkittävyys on ilmaistu matriisissa myös sen väri vaihtuu vihreäksi. Myös merkittävyytensä ilmaisevan tähtisarakkeen painaminen muuttaa sitä vastaavan korrelaatiokertoimen värin vihreäksi.

Clicking on a correlation bar in the histogram turns all coefficient values distributed within that bar blue. Multiple bars can be selected simultaneously. Clicking mouse button on a selected factor reverses selection.

3.1.1.4 Distance matrix window

Clicking the 'Distances' button opens the distance matrix associated with the data set. In this matrix the number of rows and columns indicate the respondent number and the cell values represent respondents' distance from each other. Histogram functionality is similar to that of the correlation matrix –window.

3.1.1.5 Cluster matrix window

Datajoukon klusteroinnilla pyritään jakamaan data selvästi erottuviin osajoukkoihin eli klustereihin. Klusterointi on suoritettu yleisesti käytetyllä K-Means–algoritmilla.

The purpose of clustering is to divide the data set in clearly distinguished (??) subsets, clusters. In this application clustering is made using the well known K-Means algorithm.

Matriisissa kukin rivi kuvaa yhtä klusteria. Rivillä ko. klusteriin kuuluvat alkion on esitetty punaisella värillä. Alkion numeerinen arvo kertoo alkion etäisyyden ko. klusterin keskipisteestä.

In the cluster matrix view each row represents a different cluster and the values distributed within that cluster are coloured red. The numeric value in each cell indicates that cell's distance from corresponding cluster's centre point. Histogram functionality is similar to that of the correlation matrix –window.

3.2 Help files

Help files....

4 Testing process

This chapter is about the application testing process.

4.1 Requirement specifications

These are the tasks the program is required to be used for.

To run the A-KIT application Java3D and JRE (Java Runtime Environment) have to be installed on the client machine.

Hercules g400-näytönohjain ei ole yhteensopiva Java3D-ohjelmiston kanssa, näytönohjaimen ajureiden tai DirectX:n päivitys eivät vaikuta asiaan.

Agoranet-ympäristö: Jos java-ohjelmien käynnistyksessä ongelmia, syynä voi olla esimerkiksi: PATH-ympäristömuuttuja asetettu oletuksena hakemaan ensimmäisenä java.exeä hakemistosta C:\WINNT\System32 -> myöhemmin polussa määritellyt javat eivät vaikuta tilanteeseen. Korjaus: admin-oikeuksilla pathin alkuun haluttu polku, esim. C:\Program Files\j2sdk1.4.1_01\bin

Pathin näkee komentokehoteessa käskyllä SET.

4.2 Use Cases

Tehdään DTD, joka kuvaa kunkin käyttötapauksen rakenteen. Käyttötapaukset XML-formaatissa.

Määrittely

- Käyttötapauksen nimi
- Numero
- Metatiedot (tekijä, pvm,..)
- Toimijat (ei aktiivinen, koska vain yksi käyttäjä)
- Esiehdot
- Kuvaus
- Poikkeukset
- Jälkiehdot

4.3 Usability tests

...

4.4 Testing plan

sdjks

References

Google: <http://www.google.com>, 17.4.2003.

Appendixes

Appendix 1. Title

Appendix 2. Title