GIDA

Intelligent Support for Geographic Data Analysis and Decision Making in the Web

Natalia Andrienko
Fraunhofer Institute Autonomous Intelligent Systems
AiS.KD - Knowledge Discovery Team
Schloss Birlinghoven, Sankt Augustin, 53754 Germany
Natalia.Andrienko@ais.fhg.de

Gennady Andrienko
Fraunhofer Institute Autonomous Intelligent Systems
AiS.KD - Knowledge Discovery Team
Schloss Birlinghoven, Sankt Augustin, 53754 Germany
Gennady.Andrienko@ais.fhg.de

ABSTRACT In our software system Descartes we have implemented a wide range of interactive map-based tools intended to support exploratory analysis of spatial data and decision-making in a geographic context. The tools can be accessed through the Web and, hence, could potentially serve a broad user community. However, the novelty of the tools makes them difficult to understand and use. The users need instructions that could also be provided over the Web. Preferably, the system should recommend appropriate instruments and assist in applying them according to particular user's goals. This was the rationale for developing an intelligent guidance subsystem within Descartes. The paper describes the implementation of the guidance subsystem. A special emphasis is made put on the decision support facilities of Descartes and on how the guidance subsystem can help users in utilising them.

1. Introduction

The Commission on Visualisation and Virtual Environments of the International Cartographic Association (http://www.geovista.psu.edu/sites/icavis/) coordinates research and promotional activities for raising the role of maps as tools for "spatial thinking", i.e., hypothesis generation, data analysis, and decision-support (MacEachren 1994, MacEachren and Kraak 1997). The Commission emphasizes the importance of building computer tools that expand the possibilities for human interaction with maps and thus the possibilities of facilitating visual thinking.

Our work on design and development of a software system named Descartes goes in line with the objectives and guidelines set forth by the Commission. With its interface, components implemented in the Java programming language, the system can run on the World Wide Web within standard Web browsers. Descartes offers a wide range of highly

interactive tools for visual analysis of geographic data (Andrienko and Andrienko 1999) and spatial decision support (Jankowski et al. 2001, Andrienko et al. 2002). However, our observation of people who tried to use the tools showed that the users experienced substantial difficulties. The main reason was the novelty of the tools. Moreover, the very idea of facilitating thinking by means of manipulating map displays was by itself quite new and unusual for the public. Nevertheless, after a demonstration and short training the users were quite successful in applying the interactive techniques and eventually even admired them.

The main conclusion we derived from our observations is that users, in general, need to be taught to apply the interactive techniques. However, direct teaching is not always possible, for example, in the case of first time users encountering the tools on the Web. In search of alternative solutions we considered various options. We found that providing on-line help is necessary but perhaps not very effective since many users tend not to read the texts. Using interactive on-line tutorials for training is more appropriate but also not ideal. Many people would not invest the time and effort for such training without being sufficiently convinced of the considerable benefits of the new tools.

Ideally, the new techniques should be introduced to users when they perform their own analysis, i.e. each technique should be introduced exactly at the moment when it would be useful. This means that an intelligent guide is needed, which understands the users' tasks, proposes the appropriate instruments, and helps in using them.

In this paper we introduce our ideas concerning intelligent user guidance in analysis of geographical information and describe our prototype implementation of the guide. In particular, we demonstrate how the guide can assist a user in application of the interactive decision support tools available in Descartes.

The remainder of the paper is structured as follows. In section 2 we overview existing software systems that provide intelligent support in spatial data analysis depending on the characteristics of the data and user goals. In §3 we briefly introduce the interactive facilities of Descartes that are meant to support decision-making in a spatial context. Section 4 expounds our approach to guiding users in selection and application of appropriate analytical tools. Finally, we describe the prototype user interface of the intelligent guide and demonstrate by example how the guide could assist a user in a decision-making task.

2. Intelligent support in spatial data analysis

Proper data visualisation can greatly enhance human capabilities of understanding the data and making use of them. This is especially relevant to spatial data that are hardly comprehensible without seeing their spatial distribution. However, in visualising data it is important to obey the principles of graphical and, for spatial data, cartographic representation (see, for example, Bertin 1983). These principles have been derived from studies of the characteristics of human perception, and their violation can result in useless or even misleading data displays. Since, on the one hand, lots of people occasionally or regularly need to represent graphically various data and, on the other hand, not many people possess the necessary expertise on graphical representation, it is a good idea to incorporate the appropriate knowledge in data visualisation software.

For building proper data visualisation two aspects are important. First, visual techniques applied must comply with the characteristics of the data (Roth and Mattis 1990). Mackinlay (1986) implemented one of the earliest knowledge-based systems that automatically generated appropriate graphics taking into account data characteristics. Second, the display must be suited to the purposes of its intended use. Thus, Casner (1991) provided convincing evidence that for different information-seeking tasks the same data need to be represented in distinct ways.

Early knowledge-based systems on data visualisation (e.g., Mackinlay 1986, Casner 1991) did not deal with spatial data. Such data require representation on maps that, due to isomorphism to the two-dimensional space, are capable of conveying to the human eye important spatial properties and relationships. Several intelligent systems on cartographic representation appeared in the mid-nineties. Zhan and Buttenfield (1995) describe an expert

system that asks a GIS user what cartographic representation method to select for a given single data field (attribute). The system VIZARD (Jung 1995) can combine several representation techniques in one map when more than one attribute needs to be represented. About the same time we developed IRIS, the predecessor of the current Descartes system (Andrienko and Andrienko 1997). The knowledge-based capabilities of IRIS were similar to those of VIZARD. Our system is sensitive to more elaborate data characteristics, in particular, semantic relationships between attributes. On the other hand, IRIS just generated all valid data representations while VIZARD optimised map design for better meeting user goals expressed in terms of general tasks: read values from the map, locate values on the map, compare values, or observe their spatial distribution.

A common feature of all these systems is that they generate static, non-interactive images that are fully analogous to "paper" graphics and maps. The opportunities for supporting possible analytical tasks were thus very limited. This explains why the systems that sought to suit graphics to user goals addressed only a small set of rather primitive tasks (Roth and Mattis 1990, Casner 1991, Jung 1995).

In developing Descartes, we also started from maps that only allowed zooming and switching on/off map layers (themes). Then we proceeded by adding more and more interactivity (Andrienko and Andrienko 1999). Now all the maps generated by Descartes can dynamically change their appearance in response to user actions. Thus the representation of the data may become more expressive, and may help the user reveal important relationships initially hidden within the volume of data. The user may, simultaneously, have on the screen several different maps as well as graphical displays of other types (dot plot, scatter plot, parallel co-ordinate plot, etc.) supporting views of the same data from multiple perspectives. All parallel views of the data are dynamically linked. This means that corresponding elements are simultaneously highlighted in all the displays when the user interactively selects them with the mouse in one of the graphics. Some particular interactive facilities of Descartes relevant to decision support are described in the next section.

To help users cope with this multitude of tools, we decided to implement an intelligent subsystem that would recommend suitable analytical facilities depending on user goals and explain to the user how to apply the tools to her/his data. The first idea was to reuse the approach adopted in the earlier systems. However, after careful consideration, we found it inappropriate for our purposes. The main shortcoming of the approach appeared to be the very concept of optimisation of the design to a particular task. In our opinion, in a real data analysis process one rarely addresses only a single basic task at a time. Tasks of different types intermingle, and it would be very inconvenient for the user to get a separate display per task. Moreover, it is not necessary since interactive displays are potentially capable of supporting a much wider set of basic tasks in comparison to their static predecessors, and various interactive analytical instruments can be aptly combined.

Besides, it is in general difficult and unnatural for the user to decompose analytical activities to the level of the primitive operations addressed in the earlier systems. The user would certainly prefer to think about her/his goals in problem- and domain-oriented terms, e.g., "investigate variation of population structure throughout Europe".

Our approach to intelligent user guidance is based on the idea of generic scenarios. The system proposes the user follow one of the "standardised" scenarios of data analysis defined in the knowledge base that can be applied to the data under analysis. Examples of possible scenarios are "analyse statistical data referring to territory division" or "select appropriate location". A scenario is a hierarchy of tasks and subtasks that ultimately refer to primitive operations. However, these operations are hidden from the user who operates higher-level concepts. When the user starts a task within a scenario, the system selects and combines the analysis instruments according to the primitive operations involved in the task. The intelligent guidance subsystem of Descartes will be considered more in detail later on.

3. Interactive tools for decision support

3.1 General notes

In decision-making a person deals with a number of options characterized by various attributes. Some of these attributes can be used as decision criteria. Descartes requires criteria to be numeric attributes. Hence, in order to use a qualitative attribute as a decision criterion, one needs to specify an order of preference among its values and encode the values by numbers; for example, 1, 2, 3, ... or -1, 0, 1.

Two different types of criteria are possible: benefit criteria and cost criteria. If higher attribute values make an option better (more appropriate), this attribute is a benefit criterion. If higher attribute values are less desirable (less appropriate) than lower ones, this is a cost criterion.

Criteria may have different relative importance for a decision maker. In Descartes, relative importance of criteria is specified through so-called weights of criteria. Weights are real numbers between 0 and 1. The sum of weights of all criteria used in the decision-making process is always equal to 1. The weight 0 means that the corresponding criterion is of no importance to the decision maker. The user interface to specify types (cost or benefit) and weights of decision criteria is shown in Figure 1.
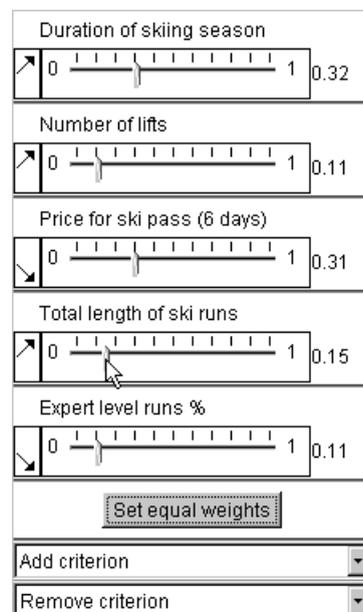

Figure 1. The user interface for specification of types of criteria and manipulation of their weights.

Below each criterion name there is a slider allowing the user to select/adjust the criterion weight within the value range (0, 1). The adjustment of one weight causes all other weights to automatically change values proportional to their values before the adjustment in order to keep the sum of weights equal to 1. Through the weights the user expresses the relative importance of the criteria. The North-East pointing arrow, to the left of each weight slider, indicates a benefit criterion; the South-East pointing arrow corresponds to a cost criterion. The user can easily change the directionality of the criteria with a mouse click.

The task of a tool for multi-criteria decision support is to help the user to derive from values of several criteria for each option either a single value representing the overall appropriateness of this option (evaluation score) or the rank of the option among the other options. In Descartes we implemented two different tools for combining values of multiple criteria. This allows us to accommodate different styles of decision-making. One tool is based on a computational method known as "ideal point" (Hwang and Yoon, 1981) that produces aggregated scores of the options from multiple values taking into account the weights of the criteria. The algorithm itself is irrelevant for the current discussion. Instead, we focus here on the implementation of the method as an interactive analysis tool. The other tool is based on visualisation of characteristics of spatial options on a map.

3.2 Interactive "ideal point" tool

The "ideal point" tool uses the panel shown in Figure 1 for the user to specify and alter criteria weights. Any slight change in the weights results in immediate recomputation of the aggregated scores. The computation results as well as the characteristics of the options in terms of the criteria are represented on an interactive plot of parallel coordinates (Inselberg 1998). Such a plot has as many parallel axes, as there are attributes; i.e., one axis per attribute. For each object characterised by the attributes (in our case an option) there is a specific position on every axis reflecting the value of the respective attribute. Lines connect the

positions on the adjacent axes. As a result, each object is represented by its "value path" through the plot area.

In our implementation, we customise the general technique of parallel coordinates to the use for decision support. Thus, the axes for benefit and cost criteria have different orientations (indicated by arrows): left to right vs. right to left. With such a solution, the best values of each attribute are always on the right, and the worst are on the left. This makes it easy to estimate visually how good any specific option is: the closer to the right edge of the plot a line is, the better the option.

Differences in relative weights of criteria can be (optionally) reflected by variation in lengths of the axes: the more important a criterion, the longer the corresponding axis. Due to this transformation, lines of options surpassing others in more important criteria shift visually more to the right ("good") pole of the plot. The appearance of the customised parallel coordinate plot may be seen in Figure 2.
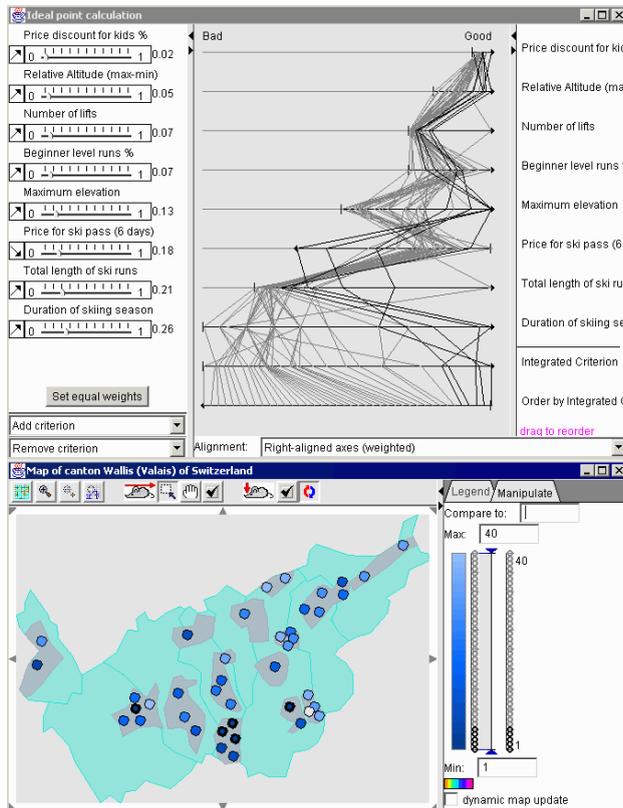

Figure 2. Interactive "ideal point" tool. The axes in the parallel coordinate plot are transformed to reflect the directionality and weights of the criteria. The plot is dynamically linked to the map by means of simultaneous highlighting of the same options.

The two axes at the bottom of the plot represent the results of the computation by the "ideal point" method. The axis, second from the bottom, represents the aggregated scores, and the next axis shows how the options are ordered (ranked) according to these scores.

The parallel coordinate plot is sensitive to any changes in the criteria weights and the results of computation: it immediately redraws itself when a change occurs. This allows the user to investigate how sensitive the result of the multi-criteria evaluation is to small changes in the criteria weights.

The tool is dynamically linked to a map showing locations of the options (Figure 2, bottom). The map may represent, if the user desires, the aggregated scores or the ranking of the options (for example, by degrees of shading). The map also dynamically changes its appearance in parallel with changes to the computation results.

3.3 "Visual" decision making with "utility signs"

To support visual evaluation of spatially distributed options and search for suitable options on a map, we have developed a cartographic visualisation method called "utility signs". We designed the "utility sign" technique as an alternative to the computation-based "ideal point" method. It was devised to give the user a feeling of better control over the process of option evaluation and stronger involvement in this process.

Like the "ideal point" tool, the "utility signs" visualization technique allows the users to express relative importance of decision criteria through the weights given the criteria.

In our system, we implemented two variants of utility signs: utility bar charts (or, simply, utility bars) and utility wheels. A utility sign consists of several graphical elements

(Figure 3): bars in utility bar charts and circle segments in utility wheels. Each element corresponds to one of the attributes under consideration (decision criterion). One dimension of an element (height in a bar and radius in a circle segment) represents the value of the attribute for the object this sign stands for. When the attribute is a benefit criterion, the size is proportional to the value. For a cost criterion the inverse proportion is kept. Hence, better values are always represented by bigger sizes. The other dimension (width in a bar and angle in a circle segment) represents the importance of the criterion. When the user interactively changes the weights of the criteria, the signs on the map are immediately redrawn.

Figure 3. Utility bar charts and utility wheels.

By the construction of a sign, the total area of all its elements shows approximately the "goodness", or utility, of the object this sign stands for. Hence, a decision maker needs to look on the map for signs with the largest areas.

To simplify estimation of areas of utility signs and, thereby, visual evaluation of options, the signs can be supplemented with frames showing the maximum area (see Figure 3 on the right). This area corresponds to the best possible values of all the attributes. Hence, the larger the part of the maximum area is filled, the better the option is.

To facilitate visual searching for the best options on the map, we have implemented an interactive control that allows the user to remove from the view signs with areas smaller than a specified threshold (as percents of the maximum possible area). We called this operation "focusing" because it allows the user to focus the analysis on a subset of the most promising options. Thus, the map in Figure 4 includes only signs with areas of no less than 50% of the maximum possible area. Note that manipulation of weights of the criteria may change the areas of the signs and therefore can make some of the currently visible signs disappear and some of the previously hidden signs become visible. This feature allows the decision maker to test the sensitivity of her/his selection to the criteria weights.

To support comparison of characteristics of options, we have designed and implemented the interactive operation of visual comparison. It is applied to utility bars. The user may select any option as a reference object, and the signs for the other options will change so that upward oriented bars represent the more preferred values than the reference object has and the downward oriented bars the less preferred. The heights of the bars are proportional to the differences in the values. Figure 5 illustrates the effect of the visual comparison operation.
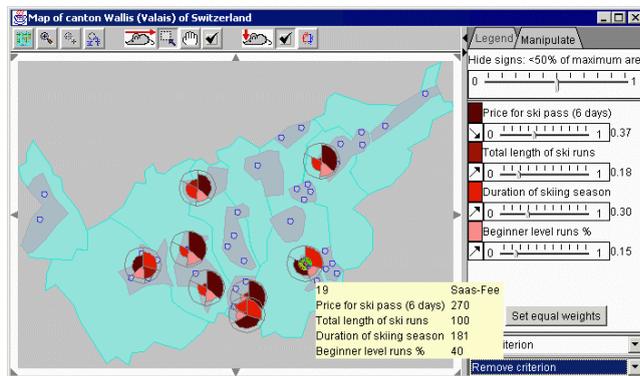

Figure 4. The operation of focusing: only the signs with no less than 50% of the maximal area filled are shown on the map.

Besides the methods for evaluation and ranking options, Descartes includes many other interactive tools that were not primarily intended for decision support but still can be effectively used in option analysis and decision-making. As an example, we describe, in the next section, one such tool called "dynamic query" (Ahlberg et al. 1992).

3.4 Dynamic query

With the "dynamic query" tool, the user can interactively set and dynamically alter constraints on attribute values. In response to these operations all the displays of the system

are immediately redrawn to show only objects satisfying the constraints. This feature can be used, in particular, to remove from consideration options with totally unsuitable characteristics. For example, a person selecting a house to buy needs a house with at least three bedrooms. At the same time the person can spend for the house no more than $300,000. In this case, the "dynamic query" tool may be used in order to remove from consideration the houses having less than three bedrooms or that cost more than $300,000.
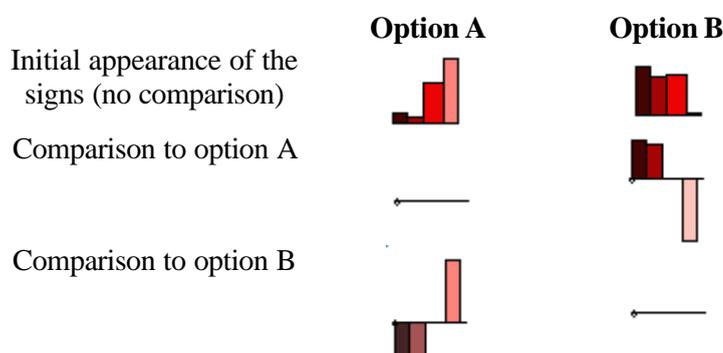

Figure 5. "Visual comparison" of options using utility bars

A decision maker can impose constraints not only on source attributes of the options but also on the results of computation by the "ideal point" method. Thus, applying "dynamic query" to the ranking of options according to the aggregated scores, the user can make all displays show only a desired number of top-scoring options; for example, the five best options. When the decision maker alters the weights of the criteria in the "ideal point" tool, the system recomputes the scores and the ranking, applies the constraints of the "dynamic query" to the new results obtained, and makes the map and other displays show the new set of best options. This shows how different tools in Descartes can be aptly combined.

4. Intelligent guidance

We have described here a few selected tools of Descartes for two purposes. The first was to show that Descartes, indeed, has certain power to support decision-making in a spatial context. The second was to demonstrate that the tools we offer are rather different from those one can find in standard GIS or mapping software. The originality of the tools causes significant problems for users. This was the rationale for development of an intelligent guidance subsystem intended to help users in applying the tools of Descartes in the course of their actual data analyses and problem solving.

4.1 Requirements to the guidance
From our observation of user behaviour and the assumption that the services of Descartes are to be widely accessible through the Web, we derived the following requirements to the guiding subsystem: the guide should

- propose relevant analysis instruments at the moment when they might become useful. For this purpose the guide needs to recognize the current goal of the user, i.e., the data analysis task to be solved.
- not request the user to explicitly decompose goals to the level of primitive operations. Instead, the guide should communicate with the user about the goals in terms of high-level analysis tasks. Hence, the guide must have knowledge about these high-level tasks and the primitive operations they involve.
- not propose a separate display for each primitive operation. Instead, the guide must provide for each high-level task a complex of instruments that fit together, that can be used in parallel and in arbitrary order, and that collectively support all the primitive operations involved.

- allow the user to decide whether to use any of the proposed instruments or not. The decision should be supported by information about the instruments: how they work and for what purpose they can be used. If the user decides to employ an instrument, it becomes immediately available, otherwise it does not appear.
- supply the user with explanations necessary for understanding how to use the instruments in the context of each task (one instrument can be useful for multiple tasks but be used differently in each case). It should be possible to have for each text its versions in different languages so that the user could select the appropriate language for communication with the system.

Design and implementation of the guide proved to be a difficult task. Numerous problems were encountered such as the absence of a suitable classification of possible data analysis tasks in the literature, a lack of knowledge about the effectiveness of the interactive instruments for different tasks, and problems related to the user interface and interaction with the user (how to recognize user goals, how to combine teaching with user analytical activities, etc.). Still, we have implemented a first prototype of the guide as a new type of on-line, task- and situation-specific help system.

4.2 Data analysis scenarios

One of the key problems was the recognition of the user's current goal. In seeking a solution to this problem the idea of generic data analysis scenarios appeared; i.e., very high-level tasks expressed in general terms such as "investigate statistical data referring to territory division" or "select an appropriate place (or places) according to multiple criteria". A hypothesis was formed that there exists a finite and rather limited number of such highest-level geography-related tasks, or that at least a constrained number of tasks would be sufficient to cover the needs of the majority of potential users. Each scenario is composed of smaller tasks that, in turn, may consist of subtasks, and so on recursively. Often individual tasks inside one scenario must be performed in a given sequence, but it is also possible that the order of some tasks is irrelevant.

As an example, consider the task hierarchy corresponding to the scenario "select an appropriate place (or places) according to multiple criteria". This scenario may stand for various specific selection tasks such as "select a house to buy", "select a place to go on vacation", or "select a location for a new waste disposal factory".

1. Explore the opportunities
    1.1. Survey the territory
    1.2. Define the options to select from
    1.3. Investigate spatial distribution of the options
        1.3.1. Observe spatial distribution of the options, detect patterns
        1.3.2. Explain the observed patterns of distribution of the options
            1.3.2.1. Relate to variation of contents of the territory (neighbourhood)
            1.3.2.2. Relate to variation of thematic (attribute) characteristics over the territory
    1.4. Investigate the variation of the thematic characteristics of the options
        1.4.1. Observe the spatial distribution of values of attributes, detect patterns and trends
        1.4.2. Explain observed peculiarities of the distribution of attribute values
            1.4.2.1. Relate to variation of contents of the territory (neighbourhood)
            1.4.2.2. Relate to variation of thematic characteristics over the territory
        1.4.3. Reveal correlations among the attributes characterizing the options
    1.5. Investigate the variation of neighbourhood of the options
        1.5.1. Investigate the variety of objects around the options, detect the presence or absence of objects of a particular kind
        1.5.2. Investigate the variation of the thematic characteristics of objects in the neighbourhood
2. Select suitable options

2.1. Exclude inappropriate options
    2.1.1. Exclude options on the basis of their thematic characteristics
    2.1.2. Exclude options on the basis of location, presence or absence of particular objects in the neighbourhood
    2.1.3. Exclude options on the basis of the thematic characteristics of objects in the neighbourhood
2.2. Rank or classify the appropriate options according to the degree of preference
    2.2.1. Define criteria
        2.2.1.1. Define criteria on the basis of the thematic characteristics of the options
        2.2.1.2. Define criteria on the basis of the location of the options, the presence or absence of particular objects in the neighbourhood
        2.2.1.3. Define criteria on the basis of the thematic characteristics of objects in the neighbourhood
    2.2.2. Rank or classify the options according to the criteria

Note that the tasks are not detailed to a level of primitive operations such as "read values", "locate values", or "compare values" as this would be inconvenient for users. In a particular case, not all of the tasks comprising a scenario may be applicable. For example, in the selection of a house the options will usually be already defined: these are the houses proposed for sale. Therefore, task 1.2 "Define the options to select from" is in this case irrelevant, while in the task "select a place to go on vacation" it may be significant. Task 1.3 "Investigate the spatial distribution of the options" makes sense only when the options are rather numerous and spread over a wide territory.

The generic scenarios are formally described in a knowledge base. The guide may thus look up the knowledge base and find out which scenarios actually exist and what smaller tasks they consist of. For each terminal task (i.e., a task without descendants in the hierarchy) the knowledge base specifies appropriate instruments of analysis. When a user wishes to utilise this support system, the guide will propose the selection of the scenario that matches most of the user's goals. Then the user may follow the given hierarchy of tasks. When a terminal task is reached, the guide proposes the corresponding analytical tools together with instructions on how to apply each of them in the context of this task. Thus, the decision support tools described above fit to task 2.2.2, "Rank or classify the options according to the criteria". For performing this task, the user is offered the "ideal point" and the "utility signs" tools as alternative instruments. Brief descriptions allow the user to choose which of the instruments to use. When the user chooses one of them, the guide activates the instrument and displays more detailed instructions on how to use it. If the user, for whatever reason, does not like the chosen instrument, she/he may at any time switch to the other one.

The user does not have to attend to every task described for the scenario but may skip any one that is deemed irrelevant. On the other hand, the proposed list of potentially relevant tasks will serve as a reminder that helps the user to avoid overlooking important steps in the analysis, or encourages consideration of the steps that were originally not deemed relevant.Whenever possible, the user may decide in what order to fulfil the tasks. However, some tasks may use results of other tasks and, hence, require the latter to be done earlier. For example, task 2.2.2 "Rank or classify the options according to the criteria" uses the results of task 2.2.1 "Define the criteria". In such cases, the guide should ensure that the user adheres to the required order.

4.3 Knowledge base on task support

The knowledge base for task support is defined in the XML language. It consists of three major parts: enumeration of possible basic operations (e.g., "compare values" or "detect spatial patterns"), description of available instruments, and definition of scenarios. Each scenario is a hierarchy of general tasks. In the current prototype, the knowledge base includes only two generic scenarios: "investigate demographic data referring to territory division" and "select an appropriate place (or places) according to multiple criteria".

For the available analytical instruments the knowledge base specifies which basic operations are covered by each of them. For example, for the interactive technique, "dynamic query", it is specified that it supports such operations as "locate values", "select or exclude objects on the basis of attribute values", and "detect patterns in the spatial distribution of values". For each operation there is an instruction addressed to the user about how the operation can be done with the respective instrument.

For each terminal task of a scenario, the knowledge base lists the instruments that may be appropriate for this task. Moreover, it is specified which of the basic operations supported by these instruments are relevant in the context of this task. Thus, the dynamic query can be used for the tasks in 1.4.1, "Observe spatial distribution of values of attributes", and in 2.1.1, "Exclude options on the basis of their thematic characteristics" of the selection scenario. The relevant operation in the former case is "detect patterns in the spatial distribution of values" and in the latter "select or exclude objects on the basis of attribute values".

The so defined correspondence between tasks, instruments, and basic operations works as follows. When a user chooses a terminal task, the guide proposes the appropriate instruments according to the correspondence task-instruments specified in the knowledge base. For each instrument a brief general description is given. After the user chooses to apply a particular instrument, the guide displays a more detailed instruction. The latter is constructed from the specific instructions concerning the basic operations supported by the instrument and relevant for the current task. Thus, when the user employs the dynamic query tool in the context of task 1.4.1, the guide will instruct the user on how to apply this tool for detecting patterns in the spatial distribution of values. When the tool is activated in the context of task 2.1.1, the guide explains how to apply it in order to select or exclude objects on the basis of attribute values.

Although currently we have only English versions of the texts displayed by the guide, the design and implementation of the knowledge base foresees that all texts for the interaction with users can be stored and displayed in different languages.

For any task the knowledge base may specify requirements regarding the data that need to be met so that this task becomes relevant. For example, task 1.4 "Investigate the variation of thematic characteristics of the options" makes sense only if there are some thematic characteristics available. Such knowledge allows the guide to display only those tasks that can be executed with the available data.

When appropriate, inputs and outputs of instruments and tasks are described. For example, the "ideal point" tool needs as input a set of objects with attributes that are to be used as decision criteria. The output of this tool is two new attributes: the aggregated evaluation scores of the options and their ranking (order) according to these scores. There are also instruments that do not produce any output; for example, the tool for visual comparison or for viewing attribute values.

Tasks with inputs and/or outputs may be linked to other tasks through these inputs and/or outputs. For example, the output of task 2.1, "Exclude inappropriate options" is a subset of the initial set of options. It is used as an input to task 2.2 "Rank or classify the appropriate options according to the degree of preference". However, task 2.1 is optional, and a user may skip it. The knowledge base specifies which alternative input shall be taken for task 2.2 in this case, namely, the whole set of options. Similarly, task 1.2 "Define the options to select from" provides an input to all the following tasks. The knowledge base may state that, if task 1.2 is skipped, the user needs to specify which of the available map layers contains the options to select from.

Outputs of tasks result from the application of the corresponding instruments, therefore the knowledge base links outputs of tasks to outputs of the instruments. When several instruments having outputs need to be associated with a task, this can be done according to one of the following schemes:

1. The instruments are declared as alternative. Then, if the user applies one of them after another, the results of the earlier applied instrument will be discarded, and only the results of the last applied instrument are taken as the output of the task.

2. The instruments are applied in a prescribed order. The output of each instrument is taken as an input for the next instrument in the sequence. The output of the last instrument is the result of the task.

3. Each instrument can be applied independently of the others. The results of several instruments are combined using either the operation of intersection or the operation of union. This scheme is applicable only when the instruments produce results of the same type.

The same schemes may be applied to subtasks of a task. For example, the subtasks of task 2.1, "Exclude inappropriate options", can be executed independently of each other (scheme 3). Their results (subsets of the initial set of options) are combined using the intersection operation that defines the result of task 2.1. The same scheme applies to the subtasks of task 2.2.1, "Define criteria", but here the results (i.e., the criteria defined) are combined through the union operation.

It should be noted that the currently existing prototype implementation of the guide does not yet adequately handle all input/output links between instruments and tasks. However, this work will be definitely concluded in the near future.

4.4 The user interface of the task-supporting guide

The guide may be activated on the user's initiative at any time. At the beginning, the guide asks which of the scenarios should be applied. After the user has chosen the scenario, the guide displays the hierarchy of potentially relevant tasks for this scenario. Thus, for the scenario of multi-criteria selection among spatially distributed options the guide will display the task tree given in section 4.2. Special symbols are used to differentiate between terminal tasks and tasks composed from subtasks: circles denote terminal tasks and "folder" symbols mark higher-level tasks. Below the task tree there is an explanation of the meaning of the currently highlighted task. By traversing the task hierarchy, the user may view explanations of all tasks. Terminal tasks may be chosen for execution.

After some terminal task has been chosen, the guide replaces the window with the task hierarchy by a new window devoted to this particular task. For example, the screenshot in Figure 6 demonstrates the window intended to help the user in performing the task "Rank or classify the options according to the criteria". On the left of the window the so-called task context is shown. In this case, the context includes the name of the map layer containing the spatially distributed options to select from ("Skiing resorts") and the list of the attributes of the options relevant to the decision-making problem (i.e., the potential decision criteria). On the right of the window, the system displays a general instruction on how to use the map in order to do the task. Below this instruction the recommended instruments are listed.
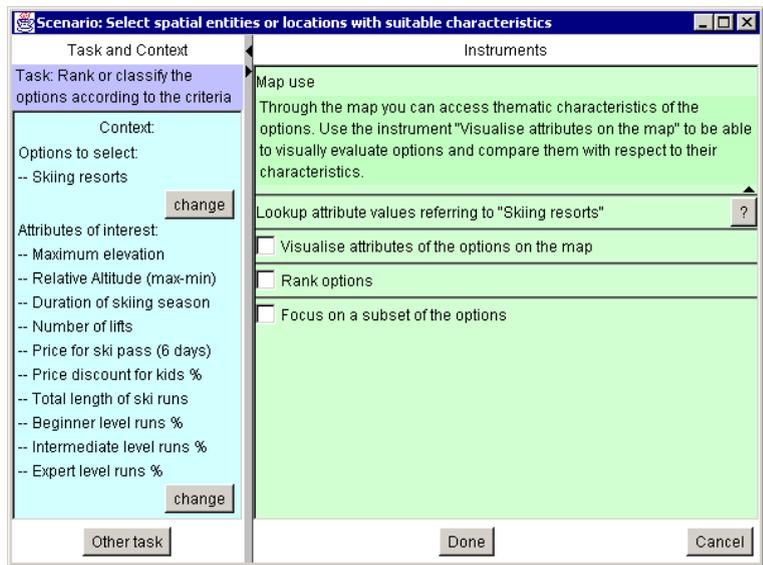

Figure 6. The guide displays a map use instruction and a list of instruments appropriate for the task "Rank or classify the options according to the criteria".

Instruments appropriate for a task are of two kinds. One kind is system functions or tools that must be specially activated. The other kind, such as access to attribute values or highlighting of objects, are always available without special activation. The guide marks instruments of the first kind with

checkboxes. To activate an instrument, the user needs to click in the checkbox. Instruments of the second kind are also listed among the recommended instruments in order to at least remind the user that they may be helpful in fulfilling the task. The buttons labelled by question marks provide access to the appropriate pages of the system's on-line manual where the user can find more extended explanations concerning the use of the tools. If provided in the knowledge base, the guide can additionally display task-specific instructions regarding the permanently available instruments.

For any instrument included in the list of recommended instruments the user can get a brief explanation of its meaning. The explanation is displayed in a popup window when the user moves the mouse cursor over the name of the instrument (Figure 7).
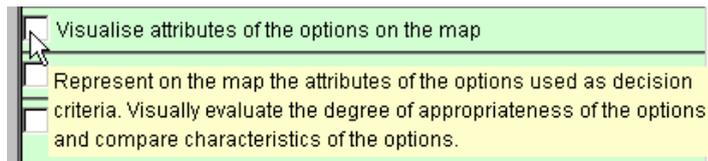

Figure 7. The guide explains to the user the meaning of an instrument.

When the user activates some instrument by checking the corresponding checkbox, the guide may ask them to specify the input for the instrument. Thus, for the instrument "Rank options" the user needs to specify the attributes to use as criteria. When the input is defined, the guide launches the appropriate tool. This may cause the appearance of a new window, for example, with the user interface of the "ideal point" tool (as is shown in Figure 2). At the same time detailed information about the tool is inserted into the guide's window (Figure 8). This information consists of two parts: 1) a general explanation of how the instrument works, and 2) task-specific instruction, i.e., how to use the instrument for the current task (in a case when the instrument can support several types of tasks). The latter is constructed from the instructions for the particular basic operations supported by the instrument (these instructions are specified in the knowledge base). The guide will show only instructions for those operations that are relevant for the current task.
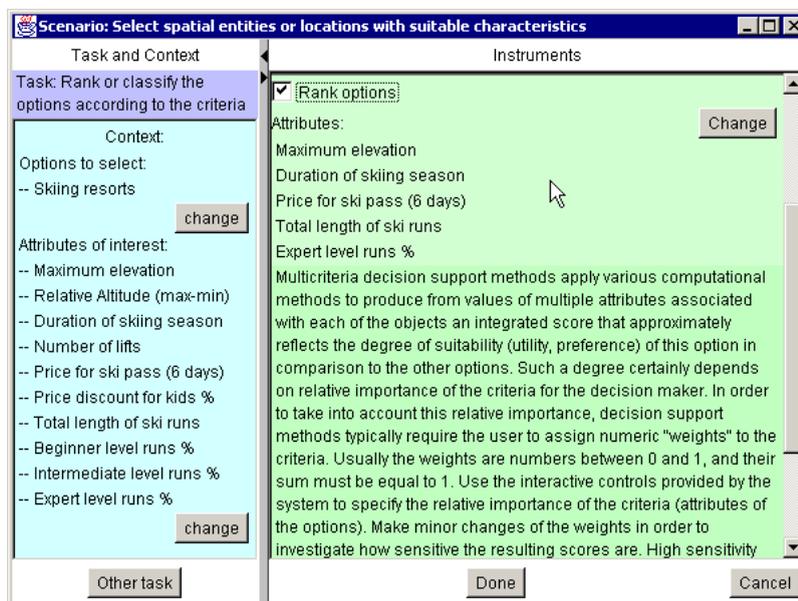

Figure 8. The guide explains how to use the computational methods for multi-criteria decision support.

When the user activates the instrument "Visualise attributes on the map", the guide first asks what attributes need to be visualised. After that the guide asks the user to select the desired cartographic represen-tation method. The guide proposes only those methods that are appropriate with regard to the selected attributes and the current task. Thus, in the context of the task "Rank or classify the options according to the criteria" the guide will offer the user the methods "utility bars" and "utility wheels" described in section 3.3. The guide can also provide explanations about each of the proposed presentation methods in pop-up windows. If the user selects one of the methods, the data are immediately represented on the map by means of this method, and appropriate map

manipulation tools are attached to the map. At the same time the guide inserts information about this method and the manipulator in the window for task support.

To avoid overloading the user with masses of text, all the explanations and instructions appear initially in a "folded" state. A triangular mark resembling a downward-oriented arrow indicates the presence of a text. Clicking on the mark opens the text for reading. The user can "fold" the text again by clicking on the upward-oriented arrow that appears when the text is opened.

In summary, the user interface of the guide serves the following purposes:

- Remind the user about potentially relevant subtasks of a task, and recommend an appropriate sequence of subtasks;
- Suggest analysis instruments when they are useful, i.e., appropriate for the subtask being currently fulfilled;
- Activate the instruments when the user decides to apply them, i.e., visualize data, create corresponding controls, etc.;
- Explain how to use the instruments that are relevant for the current subtask.

5. Conclusion

The rationale for developing the intelligent guidance subsystem of Descartes was to make the modern map-based facilities for data analysis and decision support accessible to a broad community of users; in particular, to those utilising such services through the Web. It was important to ensure that the technical tools and functions could be utilised appropriately and effectively. We assume that, besides professional cartographers, the target user community includes ordinary people using thematic maps for some practical needs (such as selection of a house to buy) or for self-education, and working people from various fields that might occasionally use thematic maps in their profession. These potential casual users typically have no specific expertise in cartography, principles of graphical representation of information, and geographical information systems. Descartes was designed and realised as a knowledge-based system that incorporates this expertise and applies it for three complementary purposes. First, it assists users to represent data on maps and other graphical displays by automatically selecting and generating visualisations that are in accord with the principles of cartography and graphics design. Second, it tries to suggest exploratory instruments suitable to the goals of analysis. Third, it assists in utilising the suggested instruments by context-specific instructions about how to operate them in view of the goals. Such instructions are especially important because many of the instruments are novel even for experienced GIS users.

The current design and implementation of the guide is a prototype to be further developed in several directions. First, the guide must be extended in order to manage input-output links between instruments, from instruments to tasks, and between tasks. On this basis the guide must be enabled, when necessary, to induce a logical sequence of performing tasks. Second, we need to extend the knowledge base, which currently only contains knowledge for two general scenarios, to a broader spectrum of scenarios. Third, although we have no doubt that such knowledge-based services can be of great use, it is necessary to perform tests regarding the usability of the specific implementation of these services. In particular, we need to validate the following hypotheses:

- The concepts of general scenarios, tasks, and task hierarchies are understandable and appropriate for users; i.e., users are able to relate their specific tasks to the general tasks the system deals with;

- The guide encourages the use of available interactive techniques; i.e., more users will actively apply more of the interactive techniques;

- The instructions provided by the guide are effective; i.e., people can and want to learn how to apply the new exploratory instruments without previous oral presentation and live demonstration;

- The user interface of the guide is understandable and convenient.

Regarding the user interface, we are in particular worried about the large number of texts provided by the guide. From our experience, we learned that common users are reluctant to read any manuals or explanations, even when they are available on-line. On the other hand, it is necessary not only to prompt people into using the novel exploratory techniques but also to explain how these techniques work and how to apply them. It is hard to do this without texts. Since the guide shows texts in relatively small portions and only when applicable, we hope that the users will not feel overloaded. Of course, this needs validation.

We will also pursue the idea of complementing textual information with some visual ways of demonstrating the work of the tools. For example, when Descartes generates a map with "utility signs", the guide could automatically play an animation demonstrating how the changing of the criteria weights influences the representation and how to visually compare the spatial options. Such a visual demonstration may already be sufficient for an advanced user to understand how to use the tool, and in this case there would be no further need in reading the corresponding texts. At the same time, a visual mini-demonstration could help casual users in understanding the textual explanations. Besides, it might raise interest and willingness to learn more about the tool.

References

Ahlberg, C., Williamson, C., and Shneiderman, B. (1992) "Dynamic queries for information exploration: an implementation and evaluation". In *Proceedings ACM CHI'92* (ACM Press), 619-626

Andrienko, G., and Andrienko, N. (1997) Knowledge-Based Support for Visual Exploration of Spatial Data. *Extended Abstracts of Int. Conf. CHI'97*. ACM Press, 16-17

Andrienko, G., and Andrienko, N. (1999) Interactive maps for visual data exploration. *International Journal Geographical Information Science* 13, 355-374.

Andrienko, G., Andrienko, N., and Jankowski, P. (2002) Building Spatial Decision Support Tools for Individuals and Groups, *Proceedings of the International Conference on Decision Making and Decision Support in the Internet Age (DSIAge'2002)*, Corc City, Ireland (to appear).

Bertin, J. (1983) *Semiology of Graphics. Diagrams, Networks, Maps.* Madison: The University of Wisconsin Press.

Casner, S.M. (1991) A Task-analytic Approach to the Automated Design of Graphic Presentations. *ACM Transactions on Graphics* 10, 111-151.

Hwang, C,L., and Yoon, K. (1981) *Multiple Attribute Decision Making: Methods and Applications*. Berlin, Springer-Verlag.

Inselberg, A. (1998) Visual Data Mining with Parallel Coordinates, *Computational Statistics*, 13 (1), 47-63

Jankowski, P., Andrienko, N., and Andrienko, G. (2001) Map-Centered Exploratory Approach to Multiple Criteria Spatial Decision Making, *International Journal Geographical Information Science*, 15 (2), 101-127.

Jung, V. (1995) Knowledge-based Visualization Design for Geographic Information Systems. *Proc. 3rd ACM Workshop on Advances in GIS*. ACM Press pp.101-108

MacEachren, A.M. (1994) Visualization in modern cartography: setting the agenda. In *Visualisation in Modern Cartography* (NY: Elsevier Science Inc.), pp. 1-12.

MacEachren, A.M., and Kraak, M.-J. (1997) Exploratory cartographic visualization: advancing the agenda. *Computers and Geosciences*, 23, 335-344.

Mackinlay, J. (1986) Automating the Design of Graphical Presentation of Relational Information. *ACM Transactions on Graphics* 5, 110-141.

Roth, S.M., and Mattis, J. (1990) Data Characterization for Intelligent Graphics Presentation. *Proc. SIGCHI'90: Human Factors in Computing Systems*. ACM Press 193-200.

Zhan, F.B., and Buttenfield, B.P. (1995) Object-oriented Knowledge-based Symbol Selection for Visualizing Statistical Information. *International Journal of Geographical Information Systems* 9, 293-315