Visualization in an early Stage of the Problem Solving Process in GIS

Andreas D. Blaser[*], *Monika Sester*[+], *Max J. Egenhofer*[*]

[*]*Department of Spatial Information Science and Engineering and Analysis,*
National Center for Geographic Information and Analysis
University of Maine
Orono, ME 04469-5711, USA
abl@spatial.maine.edu, max@spatial.maine.edu
&
[+]*Institut für Photogrammetrie*
Universität Stuttgart
Geschwister-Scholl-Str. 24
D-70174 Stuttgart, Germany
monika.sester@ifp.uni-stuttgart.de

Abstract

Methods of user-computer interaction have not changed greatly since the introduction of graphical user interfaces and their popularization by the Apple Macintosh in the early 80's. Thus, most applications still rely on primitive modalities, such as typing and pointing for input generation. Although this approach works fine for a host of common applications, it can become difficult if more complex tasks are involved. To improve this situation we propose a concept that allows people to visualize their ideas, problems, or instructions during the initial phase of an interaction with a computer, by using traditional as well as alternative modalities, such as sketching, gesturing, or talking. This approach leads to a more natural user-computer interaction and it enhances a user's power of expression. We suggest further that computers become actively involved in the process of problem formulation and that they provide support and give advice where this is adequate. This leads to an incremental process of problem formulation, where both, user and computer, are able to better visualize the actual task and where misunderstandings occur less frequently. An important field of application that could largely benefit from improved user-interaction techniques is the domain of geographic information systems (GISs). These systems are inherently complex and an interaction is often tedious, because most such systems are based on sequential and non-spatial input methods, such as SQL or other query languages that are not capable to express spatial concepts appropriately. The present paper demonstrates the concept for a visualization in an early stage of the problem solving process in GIS and discusses advantages as well as challenges in this respect. The paper closes with some application examples and considerations about involved key-research topics.

Keywords: *Multimodal user interfaces, alternative user-interaction, sketching, geographic information systems, spatial information, computer-supported decision making, problem visualization.*

Introduction

If user and computer were a married couple in our present society, they had long gone consulting a psychiatrist, because their communication is full of misunderstandings, false interpretations, and frustration and because they seem frequently unable to express their problems, conclusions, or intentions to each other. Although many people have accepted this situation—the electronic pair could not be asked—they often seem to be discontented. This article proposes a conceptual therapy and suggests specific methods that can stimulate an interaction between user and computer. The proposed measures advocate for an increased, bi-directional interaction and they describe how such an interaction could actually look like and why some applications can benefit. In this context we suggest the introduction of alternative, more human-centered modalities that simplify the communication between human and machine and we propose an active integration of computers during the initial phase of user-computer interaction.

Traditionally the term *visualization* has been used to describe the process of graphically conveying or presenting end-results. Hence, the act of visualization is in general the last step in the processing sequence of a task, with the computer as information provider and the user as information consumer. We think that this interpretation of the term visualization is incomplete, first, because it is unidirectional and secondarily, because it focuses only on the final stage of a user-computer interaction. This interpretation is also conversely to the original definition of the term visualization, which is more general and rather in the sense of building a mental image of something than solely graphically representing results on a computer screen. Based on this generic definition, we consider visualization also as the process during which user *and* computer become aware of the actual task and in which they reach a stage, where both parties are in consent about the intention or objective of an interaction. This strong integration of computers into the process of problem formulation requires a system to apply human-like strategies to be able to perceive and interpret a user's input. Applications in the GIS domain will especially benefit from such an approach, because spatial concepts are inherently difficult to understand and manage if only traditional methods are applied.

Status Quo of User-Interaction in GIS

Today's methods for interacting with geographic information systems and geographic databases are often tedious (Egenhofer and Herring 1993). Although this is partially due to the fact that some important applications still heavily rely on command-line based interfaces, this is only half the truth. Another important reason why people frequently struggle with GIS applications is that many fundamental principles upon which today's GISs are built have non-spatial origins, such as relational databases, tables, or other flat, list-related data structures that have been developed for non-GIS purposes. This strong affinity to sequential and text-oriented procedures reflects also in many GIS user interfaces of today. To improve this situation, several commercial applications have substituted command-lines with pull-down menus. Unfortunately an interaction with such systems is still cumbersome, because only the method but not the principle of interaction has changed (Figure 1).

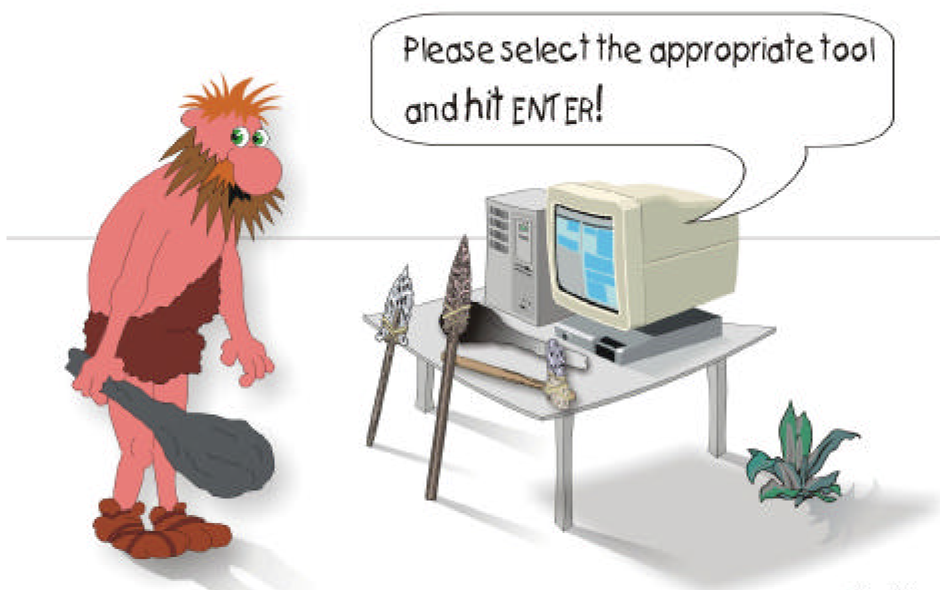

Figure 1 Today's user-computer interaction is often based on relative primitive user modalities.

In recent years many areas in computer science have made significant progress. Examples are the emerging global computer network with the world-wide-web, the dawn of ubiquitous computing (Weiser 1998), or the exponential increase in performance of microchips, storage, or telecommunication devices that brings a live in real-time (Lewis 1995) within reach. Compared with these rapid advancements, it seems that the art of user-computer interaction stayed behind. While traditional user-interaction methods may still be adequate for many mainstream applications, such as word processing or other office applications, they are often inappropriate for typical applications in GIS, because they lack specific spatial expression mechanisms.

Increasing the Bandwidth of User-Computer Interaction

One of the mayor problems regarding the use of today's computer systems and their applications is the absence of what could be called "user-centeredness". Although computer user interfaces evolved considerably since the times of mainframes, when slow ASCII terminals were standard and an interface configuration was restricted to an alteration of the font-color, they are still lacking essential characteristics that would make them really user-friendly. Despite the apparent progressiveness of today's window oriented point and click interfaces, the computer still determines how a user has to interact with the system. Another impeding factor for an effective user-computer interaction is the relatively narrow interface between user and computer that leaves users in general only the choice between mouse and keyboard—if at all.

The term multi-modal user-interaction is not new and many scientists are searching for new, improved, and intuitive ways for people to interact with computers (Shneiderman 1987; Kuhn 1992; Tue Vo 1993; Waibel et al. 1995). Great attention has been paid to voice and handwriting recognition and there are also some promising projects that investigate real-time interpretation of natural language (Bolt 1980; McKevitt 1992; Allen 1995). While this is certainly good news, there are other expressive human modalities, such as sketching or gesturing that have not yet found their way into the mainstream of research.

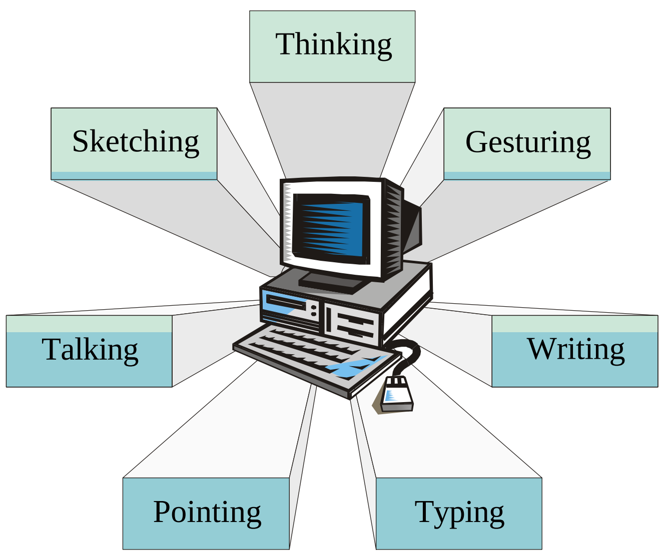

Figure 2 This graph shows traditional and alternative ways of user-computer interaction. Blue color indicates a good integration into today's systems, while green stands for prospective modalities.

Without claim of being complete, Figure 2 anticipates the diversity of a future user-computer interaction. Beside techniques that are already commonplace today, such as *pointing* and *typing* the graph includes modalities that people use frequently in their day to day interaction with other people, such as *talking*, *writing*, *gesturing*, and *sketching*. *Thinking*, though not directly a modality, could become a modality as well if it becomes possible to detect and most of all interpret human thought directly, e.g. by monitoring brain currents.

At a first glance, one might be tempted to reject such Orwellian ways to interact with machines, but we have to consider that the trend to miniaturize computing devices, render them mobile, and to integrate "intelligence" in a manifold of appliances is most likely to continue in the future, possibly at an even faster pace than today (Weiser 1998). Such, we might have as well once vivid conversations with devices, such as dishwashers, petrol pumps, or coffee-machines that have no connection to the term computer science, as we use it today. A primary incentive for research in the field of user-computer interaction is that it is obviously not possible to operate all devices and gadget with a keyboard and a mouse—a wrist-computer is an example. Verbal communication might be an appropriate solution, for some of those applications, because there is no need for any additional physical device to carry around and language is easy to use as well as relatively expressive. On the other hand, talking to devices might not be adequate if people in the vicinity get disturbed or if an interaction becomes too complex. In such cases, quieter modalities seem more suitable. Another reason that suggests an increase of bandwidth in human-computer interaction is that user interfaces should become simpler and more intuitive, so that they can be operated by people with different knowledge and skills. People should be able to freely choose and interchange their ways of interaction, depending on their ability and liking as well as depending on the situation. Hence, we argue that future systems have to adapt to the user and not vice versa and we anticipate that multi-modal user-computer interaction will be a key science for upcoming technologies.

While an improved user-interaction is beneficial for most every domain, it seems especially important for applications in GIS. GIS technology is increasingly used within a variety of originally non-GIS applications, such as biology or criminal prosecution. This diversification leads to diverging requirements and brings people from very different professional environments together. If user interfaces are narrow and too specialized, chances are high that few people will be able to operate such an application without learning the appropriate, specific dialect of the system. Therefore, GIS applications will especially benefit from simplified and more powerful human-computer interfaces that allow people to chose the type of interaction depending on their background and their preferences (Medyckyj-Scott and Hearnshaw 1993). A second reason why an introduction of alternative input modalities for GIS applications seems appropriate is the inherent spatial component, which is omnipresent in all geo-scientific applications. Using conventional modalities, such as typing and pointing alone we can not address space in its full extent.

Supporting People in Formulating their Intention

A simplified and more intuitive user interaction with computers is only one part in getting more out of the symbiosis between user and computer. As computers have become more and more powerful, allowing a host of sophisticated user applications to run smoothly and efficiently, the bottleneck has increasingly shifted from the computer to the user. The immense flow and amount of information available as well as the great number of options offered by modern applications results only too often in people's inability to deal with modern computer systems. Again this is especially true for GISs, because they are inherently complex, typically involve big volumes of inhomogeneous data, and often include multi-dimensional information. The initial phase of a user-computer interaction, where users have to outline a problem, formulate an idea, or sketch their thoughts, so that they can be processed by the system, seems to be a very frequent source of difficulties. Using predefined templates or routines to simplify this initial step is a valid approach, but it works only if the tasks are not too different from each other and if the problem is not too complex. However, templates are not the preferred choice for geo-scientific application. To improve this situation we propose a strategy that supports people already in the initial stage of an interaction with a GIS. Such, users receive help right from start and making the right choice of the problem solution strategy becomes much easier.

Visualizing People's Thoughts, Ideas, and Problems

Visualization in today's GISs starts only when the initial problem is processed and a result is generated. The technique of visualization is, therefore, solely used to explain and illustrate the outcome of a process. This approach is based on the common experience that the digestion of information within a complex context is simplified if the results are presented graphically. While this route is certainly appropriate and justified alone with the success of graphical user interfaces, the Internet, and all efforts concerning the trend to *go visual*, we think that the presentation of end-results is not the only phase where the concept of visualization can effectively be applied. Figure 3 shows the five typical steps that are necessary to assess and process a generic problem or a specific task.

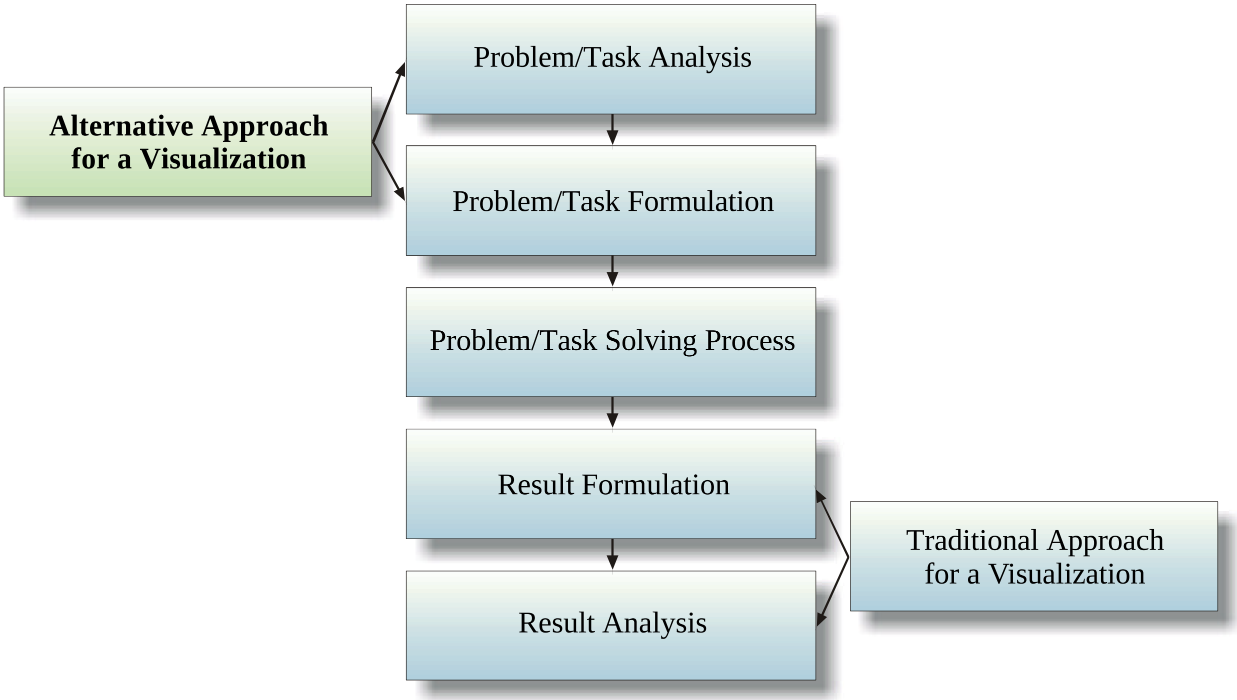

Figure 3 The generic sequence of steps that are involved in solving a problem or processing a task (blue/bold). Our proposed extension (green) offers visual help right from the start.

Figure 3 illustrates our idea of extending the generic scheme in so far that the technique of visualization is already introduced at the stage when problem or task are formulated (green/bold). This approach is based on a more general definition of the term visualization that can be formulated as

> *Visualization is the action of forming a mental image or becoming aware or conscious of something.*

Using this terminology it becomes clear that visualization techniques apply beyond presenting end-results and that intermediate results, task descriptions, or initial problem statements can be visualized as well. This last point is particularly important, because becoming aware of the general set-up of a problem is a critical part of the initial phase of the problem analysis. Hence if we can support people during this phase right at the beginning, we will enhance their ability to comprehend the initial settings, reduce the number of unsuccessful attempts to resolve a problem, and increase the overall quality of work done.

The next question that arises is the question of methodology, or *how can we support the process of visualization already in an early stage*? We have already seen that an advanced communication between people and computer devices will be a key-feature for future systems and that this communication is likely to include a variety of modalities. While this process of two way communication and incremental problem formulation is obviously beneficial for users as they can develop more easily a mental picture of the initial problem statement, it is as well very advantageous for the interacting system. This is, because the system is actually integrated in the problem formulation process and has, therefore, access and knowledge of the genesis and evolution of the problem. The proposed method of interaction comprehends, therefore, not only a visualization for the user but also for the processing system.

Advantages and Challenges for Applications in GIS

Although it is possible to describe objects and relations in space mathematically, this is not the usual way for people to outline a spatial situation in their everyday life. They prefer figurative expressions, relative referencing, landmarks, symbols, or the use of metaphors (Kuhn 1993). The traditional and abstract mathematical world in contrast favors a fixed referencing system, coordinates, entities with discrete attributes and properties or in short, things that are not directly compatible with people's preferences

(Egenhofer 1992). Working within a spatial context requires, therefore, that people are allowed to use adapted technologies that are capable to address spatial issues as necessary (Beard and Mackaness 1998). The proposed solution of an early visualization encourages people to formulate their thoughts by applying natural and familiar modalities that they use for their everyday communication with others people. A visual interaction supports more directly human spatial thinking and is, therefore, more appropriate for the formulation and comprehension of problems in a spatial environment than the usual, rather abstract, text- or menu-based methods. The computer assists and actively supports the user during the phase of problem formulation so that the initial phase becomes an incremental and gradual process, where much of the original problem formulation task is delegated to the system. This is important, because people are inherently vague and subjective in specifying spatial relations and because it helps to detect errors and inconsistencies during the problem formulation already in an early stage, which leads to a more focused and successful user-interaction.

An early and active involvement of a GIS during the problem formulation process allows the system to obtain more information about the intention of the user and supports, therefore, the overall understanding of the initial settings of this specific situation. An intelligent system may even be able to deduce a user's intention prior to the completion of the problem description and inform the user that the focused object is not part of the database or initiate anticipated processes, such as preparing potential datasets for processing, or starting to download remotely stored data records. The involvement of the system throughout the entire user-interaction offers yet another advantage. Conversely to traditional GISs, where chronological information about the genesis of a problem statement gets irrevocable lost, we suggest that the system keeps a record of all user statements to get further insight about a user's actual intention. So for instance if a user sketches an object, then erases it and draws the same object at a different location, then the system can infer that the original situation is unwanted and, therefore, must be excluded from possible solutions. The proposed technique of problem visualization offers, therefore, a route away from snapshot-style problem statements to descriptions with a rich informational content that address not only issues of space but also temporal aspects of a problem.

Increasing the bandwidth of user-interaction will obviously increase the workload for a GIS as well. This extended involvement is technically challenging, because the system has to trace, synchronize, and evaluate multiple input sources in real time, it has to retrieve and evaluate information and it has to give advice on the other hand. The enormous amount of involved data, a global distribution of spatial databases, and the very likely inter-connectivity of future GISs, render the task of accessing, retrieving, and processing spatial information even more challenging. Therefore and in order to render such complex structured digital libraries accessible, a GIS will have to employ powerful inference mechanisms and rely on sophisticated metadata models (Beard et al. 1998).

Besides such rational challenges, we have to consider the human factor as well. People's formulations of spatial relations are often vague, imprecise, and strongly context dependent, as well as influenced by the environment and background of a GIS operator. This and the increase of bandwidth between user and computer, involving a variety of modalities may further challenge such implementations. Attention must also be paid to the fact that we know little about how people respond to an increased and active involvement of computers and about how much interaction and system intervention is necessary to improve, but not impede, the communication between user and computer. Similar concerns have been raised by Norman (Norman 1994). Finding the right balance between passive and active user support is, therefore, certainly a challenging task as well.

Sketching the Future of User-Interaction in GIS

This chapter presents some practical examples of applications that we believe can benefit from the proposed visualization technique.

Sketching for Input Generation in GIS

Spatial-Query-by-Sketch (SQbS) (Egenhofer 1996; Blaser 1997; Egenhofer 1997) is an ongoing project with the main goal to develop the foundations for a sketch-based user interface that allows a user to sketch a spatial query directly on an electronic sketching device, from where the sketch is translated into a query statement and processed against a spatial database. Figure 4 shows the basic principle for a sketch-based query with SQbS.

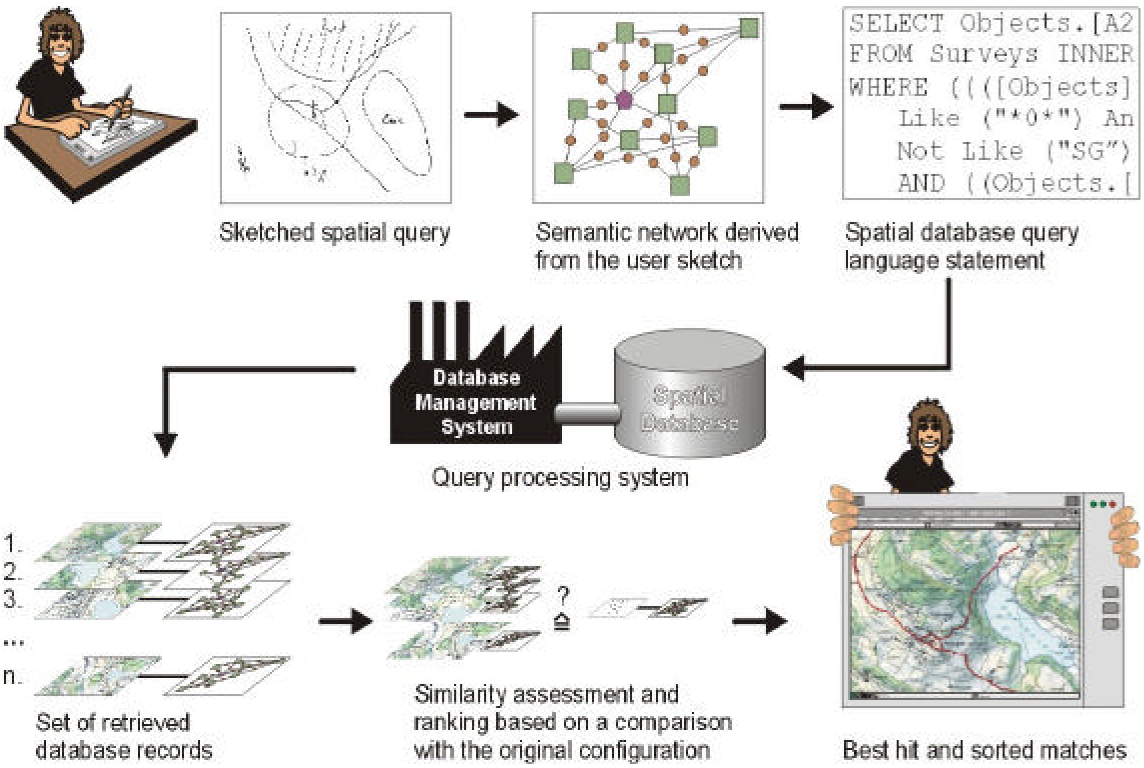

Figure 4 General overview of a sketched, spatial query, generated with Spatial-Query-by-Sketch.

Objects in a sketch are typically drawn, but they can also be specified by verbal input. For instance if the user would draw an object and say “This forest is on the east-coast of the US”, the system would label the drawn object as forest and create two other objects (east-coast and US) with a link to the initially dawn object. During the interactive process of sketching the system builds up a prioritized, semantic network that contains all objects and their relations and it checks for inconsistencies and obvious user errors, such as if a building is drawn in a lake. When the sketch is completed, the semantic network is translated into a spatial query statement that can be processed against a spatial database. The next step involves the resulting database records being transformed into a representation that can be compared with the initial semantic network of the query, whereupon the retrieved records are ranked and presented to the user.

The methodology of SQbS is particularly useful for applications that use a spatial referencing system or other relational systems with non-spatial objects, such as organizational charts. This makes the approach of this novel form of visual interaction appropriate for querying spatial, hierarchical, and conceptual structures and thus interesting for a host of other applications in GIS, but also in other domains, where the relationship of objects is important (Blaser 1997).

Practical Examples in GIS

A civil engineer has a fixed initial set-up, defined by the prevailing topography, geology, and other given facts, which may span on non-geographical issues as well, such as the available budget. In our example the engineer has to project a fresh-water pipe that has to lead over a recently build landfill site to its maintenance building. The problem is that those new landfills have a strong tendency to subside, so that the freshwater pipe is very likely to break. To get an idea of what others have done in similar situations, the engineer makes a rough sketch with some written annotations. Of course such annotations could as well be made verbally. Figure 5 shows the sketch for this simple scenario.

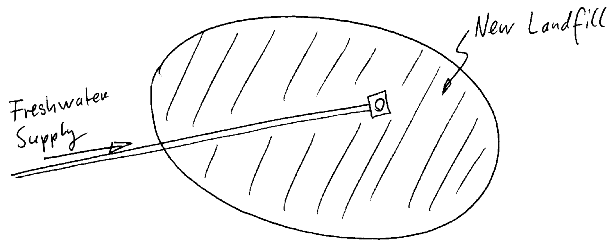

Figure 5 Sketch of fresh-water pipe crossing a recently built landfill.

By processing this sketch against a spatial database the engineer gains access to similar scenarios and thus also access to associated information, such as name and address of the engineering company in charge or even construction plans if these are available.

Our next example originates from the field of earth sciences. Scientists of a big oil company have found a strong correlation between the concentration of Jewfish, some specific geological structures of the seafloor and places, where in the past large oil fields have been discovered. To get a quick overview of locations with potential oil deposits in coastal regions, the scientists make a rough sketch, including necessary information, which is in this case a cross section through the upper seafloor. After processing their request the scientists are presented with a map, indicating similar spatial configurations within the selected area. Specifying the same query with one of today's spatial query languages would be much more complicated, because it is difficult to circumscribe spatial characteristics of a scene, such as the variable thickness of the clay layer in Figure 6 with traditional, text-based methods alone.

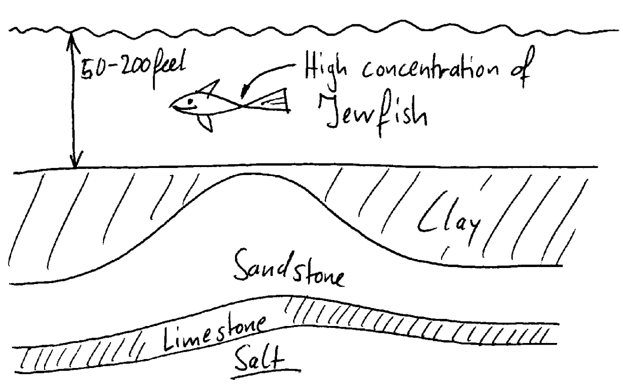

Figure 6 This sketch describes the spatial configuration of a typical oil field.

This example shows that other than map-like representations are feasible to express and convey spatial scenarios as well. However, a simple sketch may not be expressive enough for more complex three dimensional configurations, in which case other visualization techniques, such as on virtual reality based object composition toolkits, might be more appropriate.

The next example is different from the previous ones in that the primary user, a real estate agent, has no geo-scientific education. Customers interested in buying real estate are invited for an interview, during which the agent sketches the desired spatial configuration of the property, according to the description of the customer. Because a customer's preferences concerning a piece of land may involve various aspects, this is a very convenient way for a description that helps both parties to understand the desired configuration and minimizes the likelihood of misunderstandings between agent and customer. The system can detect and learn custom symbols with a specific meaning, such as the little rotating arrows in Figure 7 that indicate an irrelevance concerning the pointing arrow's direction. This allows a real estate agent to develop a personal interaction language, which extends the basic sketch-based language of the system (Blaser 1998).

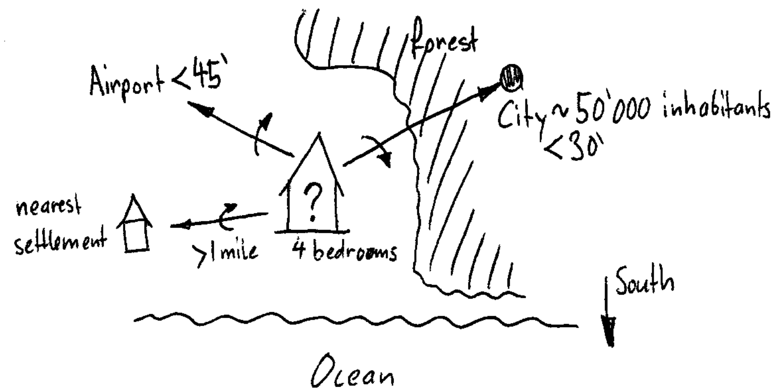

Figure 7 Sketch showing a customer's description of the desired spatial configuration of a property.

A similar and non less attractive field of application is marketing. Imagine, for instance a company that produces soundproof glazing. In order to save resources the company wants to mail their brochures only to a specific group of households along high-traffic roads. While a sketched query that fulfills the companies requirement includes only three simple sketched objects—a big road, a house along this road, and the specification of the region, e.g. Massachusetts—, it would be quite difficult to formulate the same query with traditional query languages.

Applications beyond GIS

With the ongoing trend of integrating GIS technologies in systems that we do not yet mentally connect with GIS applications, its technology will evolve from an independent science to an integrated, service rendering science, which makes it for may applications less obvious that GISs are involved at all. A prominent example is the equipment of high-end cars with GPS receivers and small on-the-road GISs (Holkstein 1997). The tendency to hide technology from users is nothing new and it usually flags the first indication that a technology has reached a certain stage of maturity. The subsequent step after a successful introduction of a technology into real world products, can be characterized by a phase of specialization and customization, where the technology is prepared to be used within a larger number of applications. The same will happen to GISs as well.

Architecture is a field of application in which space and its visualization are of great importance. An architect has to be able to convey his of her ideas convincingly and to describe the concepts of a projected building in a consistent way. Today, the preferred method is to sketch, using paper and pencil. Although paper has many advantages, such as being thin, light, and inexpensive there are also some major disadvantages, such as the fact that sketches on paper are not easy to reuse and editing capabilities are rather limited. Gross (Gross 1996) has introduced these issues with the *Electronic Cocktail Napkin*—a sketching device that provides users with pen-based gestural commands and with limited sketch interpretation capabilities. The device can already be used to retrieve plans and an extension, such as proposed in this paper, could support the detection of structural or logical flaws of a design already in the conceptual phase. Such, the system could point out that the windows in a room are too small or that with the current layout a support beam in the middle of the kitchen becomes necessary.

Our last example is hosted in a space very different from GIS, but it shares a distinct spatial component as well. The sketch of Figure 8 depicts an elementary substructure of chitin with two question marks indicating open connections to other elements. The goal of this visual query is to find other molecules that contain the specified acetamid compound ($NHCOCH_3$). Although this query could have been easily formulated with a single verbal statement, such as "Find me all molecules that contain chitin as a substructure", there are other, more complex situations, where language alone is predestined to fail. This is, for instance, the case for larger molecular structures that can fold or tangle in a specific and relevant way. If it is once possible to understand and interpret such higher order relations among molecules, then it will be extremely important to have access to appropriate methods to describe and query such spatial constellations of chemical substructures. A domain that seems to be very promising in this respect is the field of genome research, where up to now investigations are restricted to a sequential analysis of the stretched DNA.

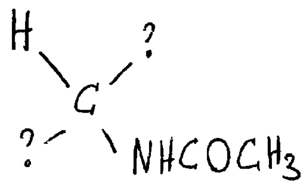

Figure 8 A researcher is looking for molecular structures similar to the one this sketch.

Beside these examples there are many other application areas, where it seems advantageous to make computers understand and actively support users, such as proposed in this paper. Since there is no space for a comprehensive listing in the scope of this publication, here just a selection of other promising fields of applications: medicine (retrieving X-rays with pathological organs (Hasegawa et al. 1989; Petrakis and Faloutsos 1994)), criminal investigations (reconstruction of criminal incidents or simulation of accidents, introducing time as an important factor), military-, and justice applications.

Key-Research Issues

The proposed approach for a visualization during the early stage of user-computer interaction involves a host of challenging research topics. Unfortunately there is no space for an in depth analysis in this paper. We limit ourselves, therefore, to providing a brief overview and discussing some important key-research issues. The focus is on spatial database queries, because this is the most typical form of user-interaction in many geo-scientific applications.

Parsing and Segmentation

During the initial phase of user-computer interaction the system has to parse and segment a user's input. Such a distinction of graphically generated objects can be based on geometric primitives, such as points, lines, areas, or more complex object properties, such as attributes, or geometrical and topological relations. Furthermore, so-called Gestalt-information like collinearity, rectangularity, or parallelity can be extracted using grouping procedures. This bottom-up interpretation process can proceed without domain specific knowledge up to a certain degree (Marr 1982). At a certain stage, however, top-down processes have to start, involving domain specific knowledge. In image understanding the differentiation between objects is a considerable problem, since segmentation results are never unique, complete, or certain. The same process is significantly easier for sequentially generated user-input with a temporal component, such as sketching, because a separation of objects can be achieved by a temporal sequencing of the input-stream. To annotate objects becomes easier as well, because the semantic of objects can be assigned either graphically or verbally.

Input Interpretation

The process of input interpretation runs parallel to the parsing process. Whenever new information is available—that is when objects have successfully been detected—this information has to be verified and integrated into the bigger picture in order to perceive a user's intention. By applying domain specific models it is then possible to assign a specific meaning to that information. Interpretation is a well known problem in other domains of spatial reasoning (Haralick and Shapiro 1993). There is a growing awareness, however, that also in GIS such interpretation capabilities have to be provided. This leads to an automation of analysis procedures, which are up to now solely controlled and triggered by human interaction (Anders and Sester 1997). The need for automation evolves from the ever increasing amount of digital data available that has to be structured and exploited in an efficient way. Database interpretation can also be seen in the context of Spatial Data Mining (Ester et al. 1995)—the process of finding "interesting structures" in a spatial data set, such as spatial association rules (Koperski et al. 1996). Data Mining techniques either rely on existing hypotheses that are checked against the data set or on unsupervised approaches, where the system tries to detect structures that the user did not think of.

Modeling

Domain specific knowledge includes object models, as well as higher order relations between objects. The first problem in modeling is the choice of an adequate representation that captures the relevant object aspects and properties. In principle, all representation schemes of knowledge representation can be applied, varying from simple attribute-value-lists, to rule sets or semantic networks. Models can be described in different ways: one possibility is to use specific models that unambiguously characterize and discriminate individual objects. The use of parametric models gives a certain degree of freedom, describing objects in terms of a fixed set of attributes, with some unassigned parameters. More flexibility is offered by so-called generic models that specify object classes instead of individual objects. An important aspect for the interpretation of the input-stream is the level of specification at which the information is processed. Pictorial representation can have many possible interpretations and vice versa also a spatial situation can be depicted in many ways. For instance, the sketch in Figure 4 can be interpreted that a general location is searched, containing village, forest, and lake (matching objects), but it can also mean that the forest should be North of the village and East of the lake (matching objects, geometry, and topology). Therefore, methods to asses and quantify the similarity between spatial scenes are an essential part of every evaluation.

The acquisition of models involves typically some hard-coding by experts. This is appropriate, as long as objects are exactly known and their characteristic and discriminating features can clearly be named. However, for geo-scientific applications this is not always an easy task. One way to overcome this situation is the use of Machine Learning techniques. The conceptual model of these techniques is based on examples instead of comprehensive and exact a priori descriptions of a domain. Examples can be given in terms of manuals, rules of thumb, expert queries, or existing data sets. A phase of supervised learning is used to extract characteristic elements in a set of given examples. Beside this, Machine Learning can also reveal the internal structure of examples (Michalski et al. 1984). Another important characteristic of Machine Learning is its ability to adapt to new situations or environments. For sketch-interpretation this is an important prerequisite, since drawings can vary considerably from user to user so that no general model can be given prior to the process.

Matching and Evaluation

The matching process, finally, establishes the link between the given models and the user-input and thus it is the final step of any interpretation of spatial structures. For user-input with a rich informational content, such as described in this paper, it is often appropriate to establish a correspondence between model and user-input on the semantic level, using object features, such as shape, object name, or location and by analyzing object relations with a focus on topology or metric. It is important, however that this analysis is able to reflect the inherited hierarchy among different objects and object groups so that different levels of specification in the data are preserved (e.g. many trees = forest). The matching itself is a typical search problem, where the best correspondence between user-input and model is searched and a similarity coefficient is determined. Subsequent steps involve an evaluation of the results upon their plausibility, quality, and reliability. This post-processing phase is essential, because of the significant portion of uncertain data that is involved. Suitable methods for this purpose include probabilistic theories

(Stassopoulou 1998), fuzzy sets theories (Pardalos et al. 1994) or theories of evidence. For the presentation of the results, finally, it seems appropriate to include additional explanations and descriptions so that the user is able to comprehend what was processed and so that the essence of the result can be captured.

Conclusions and Future Work

In this paper we have presented a concept for improving communication between people and computers. The main focus of our investigations concentrated on the initial stage of a user-computer interaction, when ideas, thought, or problems are formalized and communicated to the system. We have shown that for a more efficient interaction it is important to increase the communication bandwidth between user and computer. By introducing new and alternative modalities, such as sketching, gesturing, or talking, people have less difficulties to interact with a computer and their ability to formulate their intentions can be improved. The computer is actively involved during the initial phase of problem formulation in that the system is tracking the genesis of the problem, checking consistency issues, giving advice, and preparing for a user's most likely future steps. Through this mutual and early involvement of user *and* computer, it is possible that both pairs can better visualize a specific problem. This in turn, allows user and computer to become aware of initial set-up of a situation in a more consistence way and thus leaving less room for misunderstandings and misinterpretations.

We have shown further that the concept of visualization in an early stage of the problem solving process is particularly advantageous for applications with non-transparent procedures and complex data structures. Typical applications in GIS and other geo-scientific fields have such characteristics. The usability for such systems can be considered even higher, because the proposed technique is especially suitable to describe and formulate spatial issues. Beside examples that involved geo-scientific application we have shown some applications beyond GIS, where the proposed forms of interaction seem to be promising as well. Finally, we have discussed some important key-research issues with a focus on spatial database queries and outlined four research fields, where increased research activities appear desirable and necessary in order to transform the proposed concept of an early visualization of a problem or task description into reality.

Acknowledgement

This work was partially supported by Rome Laboratory under grant number F30602-95-1-0042 and the National Science Foundation under grant IRI-9613646. Max Egenhofer's work is further supported by NSF grants IRI-9309230, SBR-9600465, and BDI-9723873; by grants from the National Imagery and Mapping Agency under grant number NMA202-97-1-1023 and the National Aeronautics and Space Administration under grant number COE/97-0015; and by a Massive Digital Data Systems contract sponsored by the Advanced Research and Development Committee of the Community Management Staff and administered by the Office of Research and Development.

References

Allen, J. (1995). *Natural Language Understanding*. The Benjamin/Cummings Publishing Company, Inc., Redwood City, Ca.

Anders, K.-H. and M. Sester (1997) Methods of data base interpretation - applied to model generalization from large to medium scale. *SMATI '97*, Birkhäuser.

Beard, K., T. Smith and L. Hill (1998) Meta-Information Models for Georeferenced Digital Library Collections. *2nd International IEEE Conference on Metadata*, Silver Springs, MD,

Beard, M. K. and W. A. Mackaness (1998) Visual Access to Data Quality in Geographic Information Systems in *Cartographica* **(in press)**,

Blaser, A. (1997) *SQbS - A Sketch-Based User Interface for GIS*. Orono, University of Maine, National Center of Geographic Information and Analysis, *NCGIA*.

Blaser, A. (1997) User Interaction in a Sketch-Based GIS User Interface. *COSIT '97 (Poster Session)*, Laurel Highlands, Springer.

Blaser, A. (1998) *How People sketch geographic Scenes*. Orono, University of Maine, National Center of Geographic Information and Analysis, *NCGIA*.

Bolt, R. (1980) Put-That-There: Voice and Gesture at the Graphics Interface in *Computer Graphics* **14**(3), 262-270.

Egenhofer, M. (1992) Why not SQL! in *International Journal of Geographical Information Systems* **6**(2), 71-85.

Egenhofer, M. (1996) Spatial-Query-by-Sketch. *VL '96: IEEE Symposium on Visual Languages*, Boulder, CO, IEEE Computer Society.

Egenhofer, M. (1997) Query Processing in Spatial-Query-by-Sketch in *Journal of Visual Languages and Computing* **8**(4), 403-424.

Egenhofer, M. and J. Herring (1993) Querying a Geographical Information System. In *Human Factors in Geographical Information Systems*, ed. D. M.-S. a. H. Hearnshaw, pp. 124-136. Belhaven Press, London.

Ester, M., H.-P. Kriegel and X. Xu (1995) Knowledge Discovery in Large Spatial Databases: Focusing Techniques for Efficient Class Identification. *Advances in Spatial Databases (Proc. 4th Symp. SSD-95),* Portland, Maine,

Gross, M. D. (1996) The Electronic Cocktail Napkin - computer support for working with diagrams in *Design Studies* **17**(1), 53-69.

Haralick, R. and L. Shapiro (1993). *Computer and Robot Vision*. Addison-Wesley,

Hasegawa, J., N. Okada and J. Toriwaki (1989) Intelligent Retrieval of Chest X-Ray Image Database Using Sketches in *Systems and Computers in Japan* **20**(7), 29-42.

Holkstein, W. J. (1997) The 'smart car' reves up in *U.S. News & World* **November 24**,pp 64-65.

Koperski, K., J. Adhikary and J. Han (1996) Knowledge Discovery in Spatial Databases: Progress and Challenges. *Proceedings of Workshop on Research Issues on Data Mining and Knowledge Discovery*, Montreal, Quebec,

Kuhn, W. (1992) Paradigms of GIS Use. *Fifth International Symposium on Spatial Data Handling*, Charleston, SC,

Kuhn, W. (1993) Metaphors Create Theories for Users. In *Spatial Information Theory, European Conference COSIT '93, Marciana Marina, Elba Island, Italy*, ed. A. F. a. I. Campari, pp. 366-376. Springer-Verlag, New York, NY.

Lewis, T. (1995) Living in real time, side A in *IEEE, Computer* 8-10.

Marr, D. (1982). *Vision*. W.H. Freeman, San Francisco.

McKevitt, P. (1992) Natural Language Processing in *Artificial Intelligence Review* **6**(4), 327-331.

Medyckyj-Scott, D. and H. Hearnshaw (1993). *Human Factors in Geographical Information Systems*. Belhaven Press, London.

Michalski, R., J. Carbonell and T. Mitchell (1984). *Machine Learning - An Artificial Intelligence Approach*. Springer, Berlin.

Norman, D. (1994) How Might People Interact with Agents in *Communications of the ACM* **37**(7), 68-71.

Pardalos, P. M., E. Triantaphyllou, F. Lootsma and S. H. Mann (1994) On the Evaluation and Application of Different Scales for Quantifying Pairwise Comparisons in Fuzzy Sets in *Journal of Multi-Criteria Decision Analysis* **3**,133-155.

Petrakis, E. G. M. and C. Faloutsos (1994) *Similarity Searching in Large Image Databases*. College Park, University of Maryland Institute for Advanced Computer Studies.

Shneiderman, B. (1987). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Addison-Wesley, Reading, MA.

Stassopoulou, A., M. Petrou and J. Kittler (1998) Application of a Bayesian network in a GIS based decision making system in *International Journal of Geographical Information Science* **12**(1), 23-45.

Tue Vo, M. (1993) A Multi-modal Human-Computer Interface: Combination of gesture and speech recognition. *InterCHI 1993*, Amsterdam, The Netherlands,

Waibel, A., M. Tue Vo, P. Duchnowski and S. Manke, Eds. (1995) *Multimodal Interfaces*. Artificial Intelligence Review, Special Volume on Integration of Natural Language and Vision Processing,

Weiser, M. (1998) The Future of Ubiquitous Computing on Campus in *Communications of the ACM* **41**(1), 41-42.