An approach to visual interaction in mixed-initiative planning

David A. Pegram **Robert St. Amant** **Mark Riedl**
Department of Computer Science
North Carolina State University
Box 7534
Raleigh, NC 27695-7534
{dapegram | stamant | moriedl} @eos.ncsu.edu

Abstract

Researchers in mixed-initiative problem-solving have generally viewed interaction between the user and the system as a form of dialog, which provides an effective unifying framework for multimodal systems. For mixed-initiative interaction through a visual medium, however, an approach that exploits our visual perceptual abilities and the benefits of direct manipulation mechanisms is equally compelling. This paper explores the possibility of communication between human planners and intelligent planning systems via shared control of a three-dimensional graphical user interface. We are currently testing our early development efforts in the Visual Interaction Dialog (VID) system, which supports agent and user manipulation of camera placement for communicating plan structure and domain information.

Introduction

A view of human-computer interaction as dialog has come to dominate research on mixed-initiative systems. A dialog-based framework unifies the different types of communication and interaction supported by a multimodal system. For example, instructions entered via spoken input or typed commands or direct manipulation can all be interpreted in a common (symbolic) representation. The same applies to different modes of output, whether generated speech, natural language explanations, or any of a variety of graphical and tabular visualizations. The dialog perspective has generally been successful, resulting in a number of well-known systems in mixed-initiative planning, including TRAINS (Ferguson & Allen 1996), TRIPS (Ferguson & Allen 1998), COLLAGEN (Rich & Sidner 1998), and others.

In mixed-initiative planning, users collaborate with software agents to produce plans. Effective collaboration between human planners and automated software requires that participants work in areas where they perform best, use appropriate representations for communication, and acquire/transfer authority for planning tasks (Burstein & McDermott 1996). These system goals, along with previous studies (Allen 1994; Ferguson & Allen 1996; 1998), have motivated the use of dialog support in planning systems.

While the dialog-based approach has proven effective, it is not without its difficulties. User interface designers, for example, have argued for direct manipulation as an alternative to command line interaction for almost two decades. The central point of disagreement is not about the weakness of a command line interface in contrast to, say, the power and flexibility of unrestricted natural language, but rather about whether human-computer interaction is best viewed as a dialog or as action in an environment. In this paper we explore some of the issues raised by this alternative perspective, in which we concentrate on the ability of an interactive environment to constrain and guide the behavior of a human user as well as provide guidance to an automated planner.

Our work has some of the flavor of the ecological view of human-computer interaction (HCI) (Flach *et al.* 1995; J.Gibson 1979; St. Amant 1999). In human-computer interaction circles, interface designers are encouraged to provide cues in their environments that indicate how objects can be used, in order to improve ease of use, reduce the need for instructions, and enhance familiarity with the interface. These cues are sometimes referred to as *affordances*. Ideally, the affordances of an environment suggest appropriate responses at any point in time, such that one is led through the most effective sequences of actions toward one's goals. The ecological perspective suggests a few desirable properties for a mixed-initiative system:

- The system can adapt the environment such that some actions can be carried out more easily than others.
- The system can present the environment such that these actions appear (visually, aurally, etc.) to be easier or more direct than others.
- The system can convey goals, state information, and at least some task structure (e.g., focusing only on the objects related to a task) by changing the environment.
- Conversely, if the user makes changes to the environment, the system can interpret these appropriately.
- The system is accommodating, in that its suggestions rule out the user's choosing other possibilities.

Note that these can in principle be achieved by a dialog-based system, but the capabilities fit more naturally into a direct manipulation interpretation of interaction.

Our approach focuses on the importance of visual stimuli to human perception and understanding. We adopt the existing idea of using a direct manipulation interface for dialog among agents (Moo 1995). In our planning environment, currently under development, system agents manipulate the location and direction of cameras used for viewing three-dimensional (3-D) plans overlaid onto domain specific representations. Although a camera metaphor (Carroll, Mack, & Kellogg 1988) may not map directly to a planner's knowledge of the domain, it can exploit innate features of human attention and perception (Banks & Karjicek 1991; Kinchla 1991). Furthermore, not only does placing 3-D representations of plan components into a planning domain allow users to directly add, remove, and edit plan components and relationships, but some domain dependent characteristics become easily recognized, in what Woods (1991) calls "design for information extraction."

This paper is structured as follows. In the Related Work section we discuss three areas of research that have influenced our own: mixed-initiative planning, intelligent multimedia systems, and visual perception for interactive data analysis. Our work takes a step toward integration of disparate themes in these areas. The next section describes AFS, the Abstract Force Simulator in which our planning research takes place. In the section that follows we discuss visual dialog for inter-agent communications and coordination, concentrating on the potential of a visual system to support flexible, interactive visualizations, context registration, and dialog-based task management. In the final section we describe a prototype 3-D interface we have developed for AFS, called VID. The work in this last section is preliminary. We do not yet have a full implementation in which all components are integrated; we can automatically produce a variety of examples such as the one shown, on the fly, but they are currently canned in the sense that they are not produced automatically by the planner, but are instead controlled by an independent visualization module.

While it is possible to interpret input and output through a visual medium as simply another mode of dialog, HCI researchers have long argued that direct manipulation provides a qualitatively different interaction experience. We intend to incorporate findings from the literature on perception and direct manipulation into a dialog framework, with the goal of allowing shared control of a graphical user interface for inter-agent communication in a mixed-initiative planning system.

Related work

Our work on visual interaction dialog merges research from several areas including mixed initiative planning, intelligent multimedia systems, visual perception, and interactive data analysis.

Mixed-initiative planning. A mixed-initiative planning system can (perhaps inevitably) increase the amount of collaboration required between the system and user. At times the system may be employed as a tool for completing familiar and important tasks. However, at other times the system will need to function autonomously to complete unfamiliar or time consuming tasks. Although users delegate tasks to the system, they should not have to surrender the ability to guide and review the decision-making process. Work by James Allen (Allen 1994) characterizes mixed-initiative planning on the basis of three characteristics: the flexible and opportunistic exchange of initiative, shifting focus of attention to meet user needs, and providing mechanisms for maintaining shared implicit knowledge. These three characteristics are closely related to cognitive orientation, deep knowledge, intention sharing, and control plasticity, components of Silverman's model of collaboration processes (Silverman 1992). Burstein and McDermott, in their summary of mixed-initiative planning, additionally point out that research in inter-agent communication should provide flexible visualizations, context registration, and task management support (Burstein & McDermott 1996). These latter points are addressed in a later section.

Intelligent multimedia systems. Advances in graphics hardware and software technologies help reduce the cost of generating quality 3-D images, thus increasing the feasibility of immersing users into dynamic virtual worlds. Recent intelligent multimedia research has taken advantage of cinematography heuristics to produce systems for automatic explantation generation, intelligent tutoring, and other tasks (Feiner & McKeown 1991; Smith & Bates 1989; Karp & Feiner 1990; Seligmann & Feiner 1991; Gleicher & Witkin 1992; Phillips, Badler, & Granieri 1992; Drucker & Zelter 1994; 1995; Christianson *et al.* 1996; He, Cohen, & Salesin 1996; Bares & Lester 1997; 1999). Although some of these systems present a direct manipulation interface to the user, the camera is not considered a method for communication; instead, camera planning is simply used to orient the user's perspective in the virtual world. Our goal is slightly different, in that we want to support dynamic communication of task and domain information between the system and the user, with shared control of camera placement and direction. This may (we hope) have the additional benefit of identifying some of the lower-level foundations, based on principles of human attention, perception, and interaction, for current heuristic approaches to intelligent multimedia.

Visual perception and data analysis. If we consider that sign language and gestures have been used for communication for millenia, visual communication can hardly be regarded as a new concept. However, the principles behind effective communication through pictures, graphs, and computer-generated images have only much more recently been examined. Pioneering work from Tufte (Tufte 1983), Cleveland (Cleveland 1985), and Friedhoff (Friedhoff & Benzon 1989) provided fuel for a later generation of work by Keller (Keller & Keller 1993), Kosslyn (Kosslyn 1994), Brown (Brown *et al.* 1995), and Bertoline (Bertoline *et al.* 1997). We have a particular interest in effective graphic communication in statistical data analysis systems. The Visage environment (Roth *et al.*), for example, utilizes a direct manipulation interface allowing users to explore complex relationships among data.

All of this work, based on visual perception, emphasizes

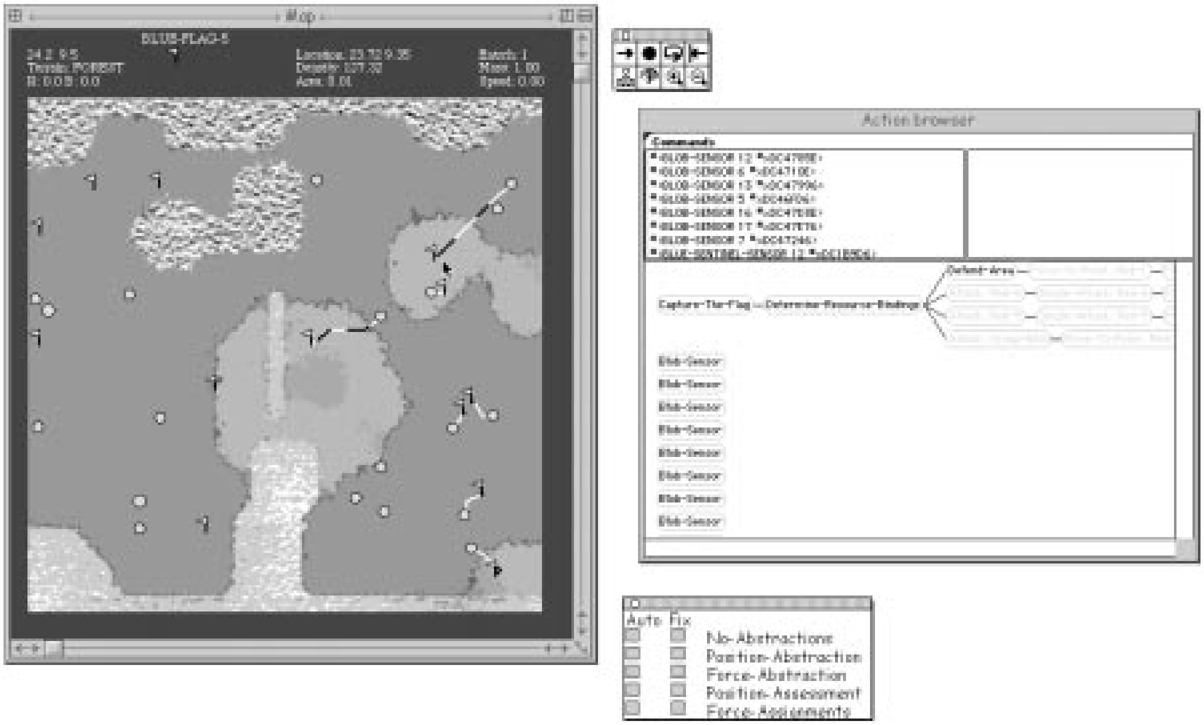

Figure 1: A Capture the Flag scenario in AFS

the correct usage of graphs and various forms of visual images for communication. Effective use of the graphical forms presented in a direct manipulation interface also relies on ecological concepts. Typical direct manipulation interfaces rely on the affordances provided by buttons, scrollbars, sliders, and other widgets for interaction. Due to the successful integration of these objects into common interfaces, some researchers have suggested that mapping the appearance directly to an object in the real world increases the likelihood that it will be perceived (Carroll, Mack, & Kellogg 1988; W.Gaver 1991; Anderson 1993). Psychological research on feature integration theory, grouping, continuity, and attention has also contributed to this area (Banks & Karjicek 1991; Kinchla 1991). From our perspective, however, existing approaches do not address the possibility of the user (or an intelligent automated assistant) adjusting the viewing perspective—distortions of objects or spatial relationships are ordinarily an effect that the developers of visualization systems would wish to avoid.

AFS

Our work takes place in the context of AFS, an abstract force simulator provided by Paul Cohen's lab at the University of Massachusetts (Atkin *et al.* 1998). AFS is a general-purpose simulation system that supports experimentation with interactive, distributed planning techniques and their relationship to physical processes. AFS provides a physical domain in which abstract agents (which for clarity we will call "force units" or "forces") can interact, based generally on Newtonian physics. Forces and inanimate objects have mass, size, and shape; they may be solid or permeable; they move with variable friction over a domain-dependent surface; they apply force to one another, causing damage/mass reduction. Figure 1 shows the existing 2-D interface to AFS, while Figure 2 shows the architecture of the system with its various components.

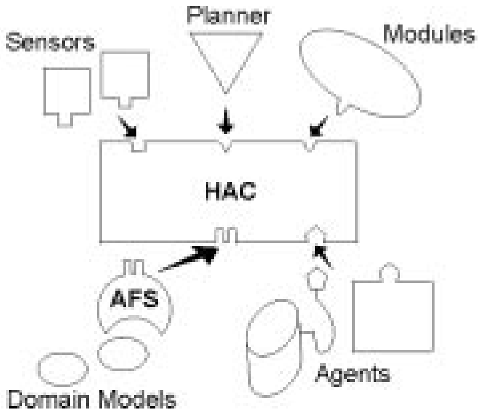

Figure 2: AFS architecture

In AFS's Capture the Flag (CTF) domain, two teams of forces move over a terrain, their travel constrained by mountains, water, and forests. Each team is responsible for defending a set of stationary flags, and successfully completes a scenario by destroying the members of the opposing team or capturing all of its flags. Figure 1 shows a sample scenario. In this domain, as in all AFS domains, force units rely on a small set of primitive physical actions: they may `move` from one location to another and `apply-force` to other forces and objects such as flags. These actions can be specialized and combined in various ways to form higher level strategies, such as blocking a pass, encircling a flag, attacking an opponent in a group, and so forth. Plan execution and monitoring is provided by HAC, the hierarchical planner at the center of the system.

We have taken steps toward mixed-initiative planning in AFS, mixing a navigational metaphor with mechanisms for direct manipulation (St. Amant 1997; St. Amant, Long, & Dulberg 1998). The user can direct the low-level actions of the teams of forces, and can view visualizations of decisions the planner makes, such as the tasks a team has taken on and how its members are assigned. Figure 1 shows a visualization of a partial plan, in which some force units are assigned to defend their flags. The figure also shows a plan browser that displays a more abstract view of the planning process. As we discuss in the next section, our current work extends this interface to a 3-D world, with the goal of providing the features described in the Introduction.

Visual dialog

In VID, the Visual Interaction Dialog system, system agents manipulate the location and direction of cameras used for evaluation and editing in a three-dimensional (3-D) planning environment. After positioning the camera, they may add, remove, and edit hierarchical plan components overlaid onto domain specific representations. The effectiveness of the interaction between the user and the HAC planner depends on the ability of the system to provide flexible visualizations, context registration, and task management support (Burstein & McDermott 1996).

We emphasize that the current system is under development; our discussion in this section and the next is of the design and early prototypes of components. AFS and HAC are robust and support all the planning activity and object manipulation we discuss. However, the figures illustrating our example in the next section were generated programmatically, rather than entirely by hand, but are not yet completely automated. Two major tasks for future development are the extension and refinement of prototypes for the various visual components in VID and their integration into AFS.

Flexible, interactive visualizations. AFS is designed as a general simulator of physical processes. Since it can simulate many different domains, the visual representation of plan components can exclude domain specific knowledge. To achieve this, VID will display plan goals, sub-goals, primitive actions, and their relationships as semi-transparent diamonds and cylinders. The diamonds represent goals at various levels in the plan hierarchy. Cylinders, connecting hierarchical goals and primitive actions, represent the relationships between them. These simple geometric shapes ease the requirements of rendering hardware allowing real-time manipulation and feedback in the virtual world. Semi-transparent plan components avoid total occlusion of other information from any single viewing perspective. Additional information can be encoded into the representations using other visual dimensions like color, pattern, etc. For example, in the CTF domain VID will color goals and their relationships according to the team and current user selection. To avoid confusion, VID does not interchange the meaning of domain specific assignments among the visual dimensions. Encodings persist until another domain is loaded or, the human planner explicitly changes them. In complex domains there are many important domain and plan elements needing representation. If only a single dimension is used to represent each of these elements we would quickly surpass the number of commonly used visual dimensions or, we might create a multi-dimensional image to complex for comprehension. Conveniently, VID can combine several dimensions to form a glyph, an abstract visual feature, which represents one or more domain characteristics.

Placement of these glyphs in VID's 3-D planning space allows the user to specify the spatial and temporal characteristics of the plans. Figure 3 demonstrates how spatial and temporal information is organized in VID's 3-D planning world. The spatial information is obtained by extrapolating from the glyph to a point on one of the three surfaces used to represent up to six dimensions in the physical domain. These domain surfaces may vary both in scale and size. For domains requiring fewer than six dimensions, the additional surface axes do not correspond to a dimension. The point extrapolated to the surface locates the plan component using two coordinates for each plane, the longitude and latitude. Again, the longitude and latitude of each plane may or may not map to a dimension in the physical domain. In some instances it may be convenient to have similar axes on multiple domain surfaces represent the same physical dimension. For example, in the CTF domain it is convenient to use the longitude of the secondary and tertiary domain surfaces to represent time. The distance from the domain surface to the plan component represents the amount of time before the simulator achieves the plan fragment. Thus, the more abstract components of hierarchical plans appear further away from each surface in the 3-D plan space.

Although the VID is flexible in representing multidimensional planning spaces, this is of little use to human planners if they cannot recognize critical features that dictate the success or failure of the plan. To assist, VID's direct manipulation interface permits system agents to dynamically alter the location and direction of cameras used for viewing the planning space. VID typically constrains the interface by providing one camera for the system agents. However, multiple cameras could be used to compare specific features in distant locations of the plan space. Each shared view permits agents to communicate with one another. When one of the system agents orients a particular camera they are allowing the other to perceive the plan space as they would (i.e. as a first person perspective). During the collaborative planning process human users and software agents take turns serving as an audience or, as a director focusing on areas of interest and eliminating visual clutter. When VID's software agents attempt to illustrate key features, they position the camera according to heuristics taken from visual perception. For convenience, a single example from a CTF scenario is lo-

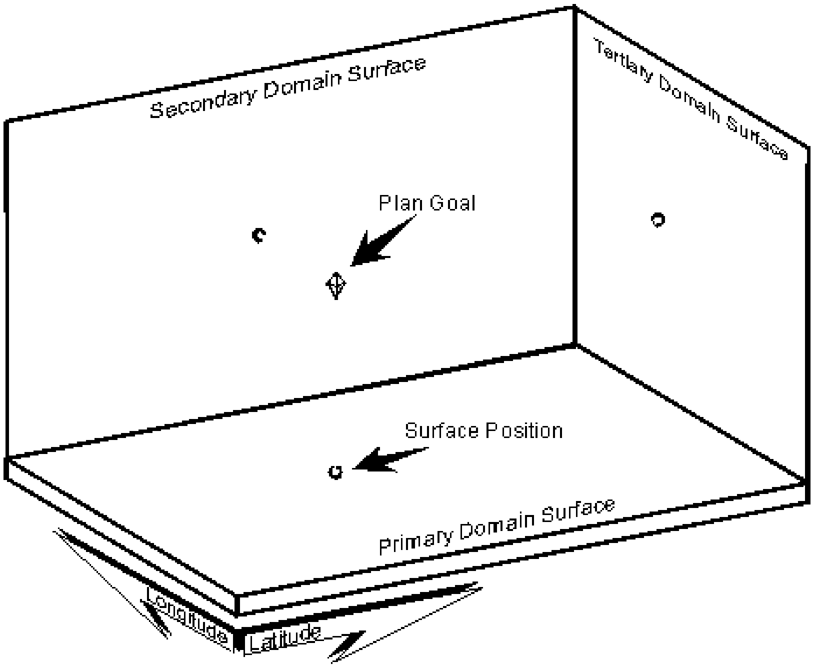

Figure 3: Organization of spatial and temporal information in VID

cated at the end of this section.

Context registration To avoid confusion, agents should communicate their changes as they construct a plan. Context registration involves conveying areas of interest and notification of new changes to collaborators. Clearly, human planners can only add, remove, and edit plan components through VID's direct manipulation interface. Like most interfaces, camera placement and component manipulation are the result of keyboard or mouse actions. Events that are easily recognized by most windowing systems. The processing of these events then relays the current state of the world to other interested agents (software or remotely located human planners). Although user modifications are always mediated through the system, it might be possible for software agents to perform tasks outside the viewing region of the shared camera. To avoid any disparities of knowledge, VID's fundamental principle requires software agents to reorient the camera to view any changes they make. During some automated tasks this additional camera movement may distract planners. To correct this problem, users may optionally specify the amount and type of information the system agents should communicate.

Dialog-based task management Although multiple agents collaborate through VID, software agents yield control of the dialog to human planners. Any time VID interrupts their planning activities by attempting to reposition the camera they easily reposition it by activating one of the camera controls. This memory feature provides a degree of control plasticity so VID can achieve a simple, quick style of interaction (Silverman 1992). Additionally, VID does not attempt to change the location or orientation of the shared camera while human agents manipulate its controls or plan components. The flexibility incorporated into VID allows the user to control the search through the possible plan space by focusing the shared camera on smaller regions. Within these focused regions users may specify plan objectives at a variety of abstract levels and delegate the lower level details to other system agents.

An example

As a simple demonstration of the previously discussed interaction it will help to walk through an example from the CTF domain. Figure 4 shows the initial settings for a scenario in this domain. Forces are represented using colored spheres. The color of the spheres and flags denotes team membership. Forces, flags, and terrain are represented on the primary domain surface.

Each of the planning decisions we discuss in our example can be made by the planner or the user, the interaction managed with conventional direct manipulation mechanisms, including icon, menu, and button selections. For example, the system can accept instructions from the user concerning which flags should be targeted; alternatively, the user might ask for a suggestion by pressing a button. In either case the planner generates plans on its own to determine its goals, possibly deferring its execution of actions to the user's decisions. For simplicity, we will assume in our discussion that the system makes all the planning decisions, each approved by the user, and the goal of the interaction is to convey its planning intentions to the user.

The system begins by constructing a plan to attack one of the opponent's forces defending a flag. VID, along with the information provided by the planner, conveys the goal/task

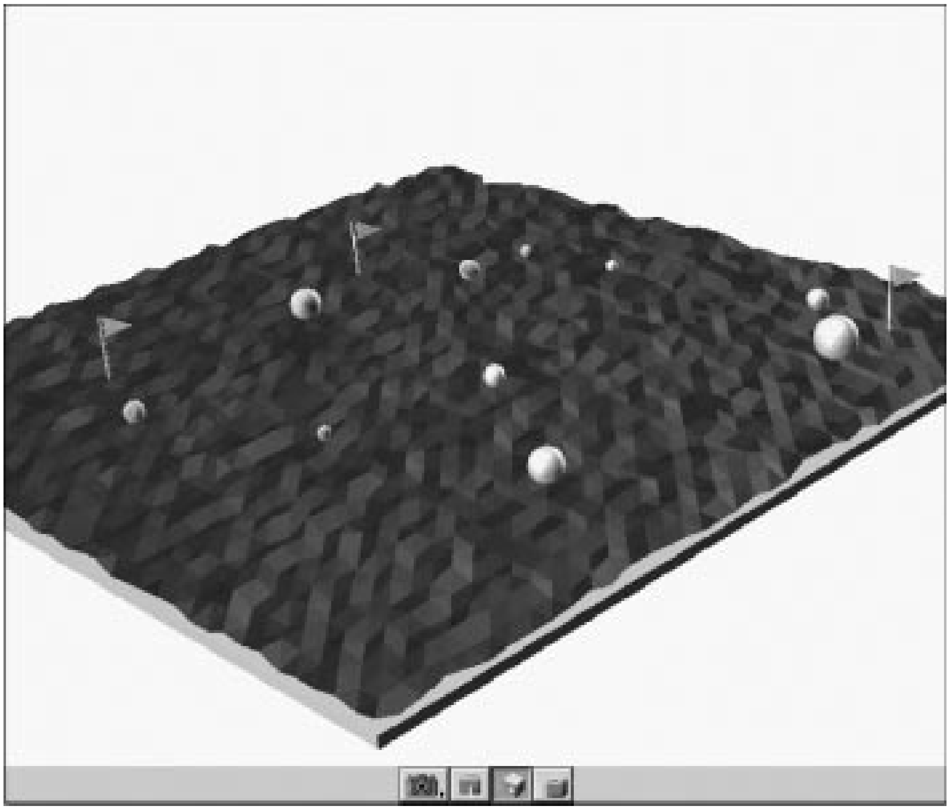

Figure 4: Initial state in a CTF scenario

combination by placing the camera at a position indicating a high likelihood of mounting a successful attack. This is shown in Figure 5. In addition to conveying the selection of a specific target, the figure shows the planner's selection of a specific offensive force for the task. Even limiting VID's interaction to camera placement, we find some useful benefits in the richness of visual cues. For example, the system shows the scope of the action: off-screen forces are implicitly considered irrelevant. Even the speed at which VID moves the camera position into place can influence the user's assessment of the required tempo of the action. Thus with a single camera positioning, the system can convey an acting force, a target object, intervening forces, and a good deal of further implicit information. This conciseness is possible partly due to the simplicity of the environment, but it is also because the system relies on the power of our visual interpretation.

In the completed system, VID will additionally add a goal icon in the plan space above the final location. The height of the goal above the primary domain surface will indicate the amount of time until the goal is reached. Initially, the system estimates the amount of time required for goal completion. After establishing this goal the user can adjust the plan tempo by dragging the height of the goal up or down. The higher it is the slower the tempo; a lower goal icon implies a faster tempo.

Let's complicate the situation. Assume that the planner has formed a plan in which the main attack is provided by a single force unit, but that its offensive power is insufficient to overcome its opponent's defense. To convey this, VID first identifies the need for further refinement of the plan (a flaw, in an informal sense), and shows it to the user as in Figure 6. VID positions the camera so the user can make a comparison between the two forces. Based on perceptual heuristics, VID zooms the camera in to eliminate as much distracting information as possible. During the zoom, the camera also pans so the two forces are shown along a common baseline, to allow comparison of aligned distances (Cleveland 1985). The planner's solution to this mismatch is to select another force unit for assistance. The planner conveys this decision by repositioning the camera as it did for the first force unit, showing the new unit's contribution to the main action.

Part of the power of this approach is that we are able to exploit the physical nature of planning with AFS. It might seem that for some purposes the interaction is at too low a level, however—how can one manage abstractions, such as some number of forces occupying some area? Fortunately, AFS supports aggregation of objects and spatial regions. The planner or the user can group and characterize agents, for example, to allow visualizations at different levels of abstraction.

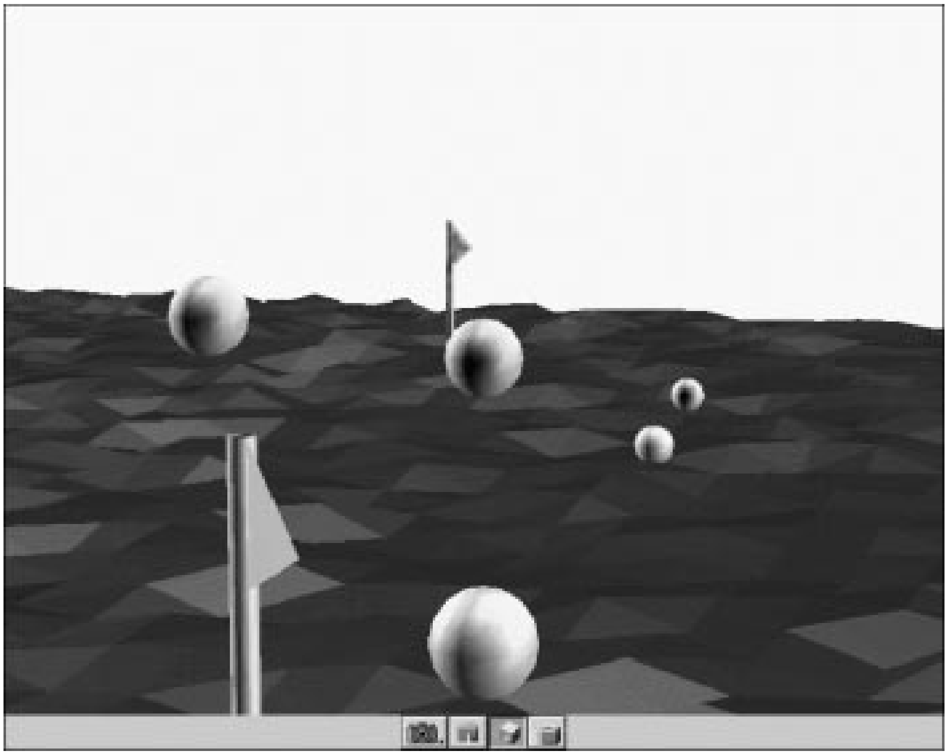

Figure 5: VID conveys the goal of capturing a specific flag

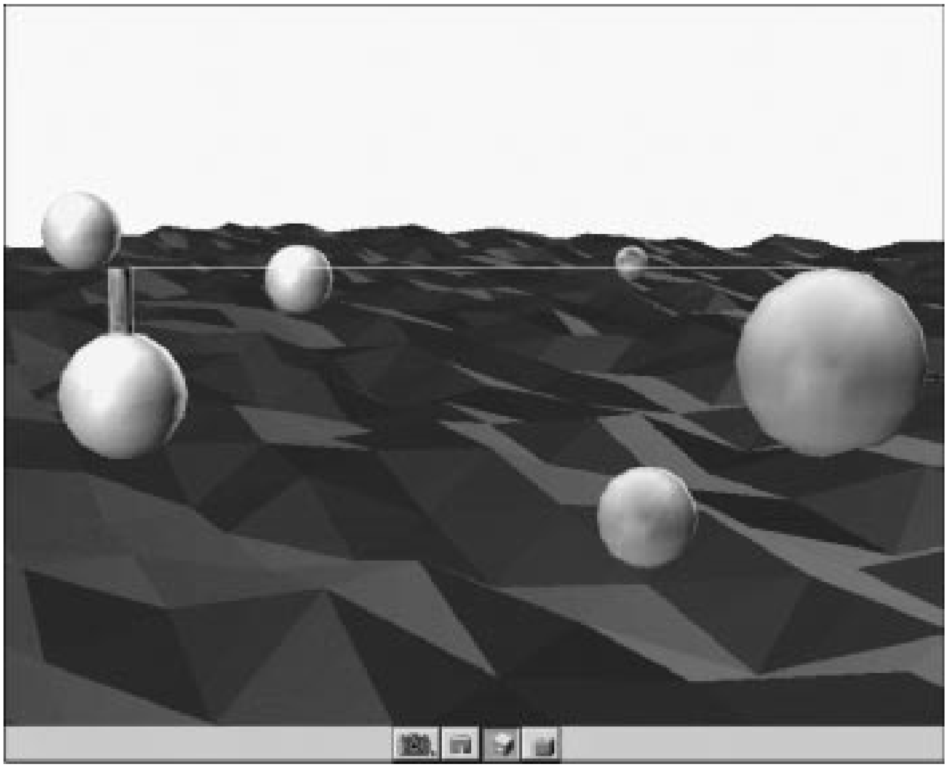

Figure 6: VID conveys a problem with the basic plan

Conclusion

In the previous example we saw how VID's shared manipulation of camera is used for communication among various system agents. VID makes this style of interaction possible by providing flexible, interactive visualization based on visual perception, context registration, and dialog based task management.

The visual approach has limitations, and we do not propose it as an exclusive modality for interaction. For example, VID does not provide a direct means for users to generate system queries, and it is not entirely clear how visualizations can address temporal reasoning. However, we believe the approach has significant promise. We hope that the completed system, like the Magic Lens filters (Stone, Fishkin, & Bier 1995; Fishkin & Stone 1995), will offer users a quick, easy-to-use interface with which they can find answers to many of their questions.

Acknowledgments

The graphical interface could not have been implemented without the assistance and donated code of John Wiseman. Our work with AFS has been helped greatly by interaction with Paul Cohen, David Westbrook, and Marc Atkin. This research was supported by ARPA/Rome Laboratory under contract F30602-97-1-0289. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes not withstanding any copright notation hereon.

References

Allen, J. F. 1994. Mixed initiative planning: Position paper. [WWW document]. Presented at the ARPA/Rome Labs Planning Initiative Workshop. URL http://www.cs.rochester.edu/research/trains/mip/.

Anderson, B. 1993. Graphical interfaces considered as representations of the real world: Implications of an affordances-based model. In Valenti, S. S., and Pittenger, J. B., eds., *Studies in Perception and Action II*, 89–93. Lawrence Erlbaum.

Atkin, M.; Westbrook, D. L.; Cohen, P. R.; and Jorstad, G. D. 1998. Afs and hac: Domain-general agent simulation and control. In *AAAI-98 Workshop on Software Tools for Developing Agents*, 89–95.

Banks, W. P., and Karjicek, D. 1991. Perception. In *Annual Review of Psychology*, 305–331. Annual Reviews.

Bares, W. H., and Lester, J. C. 1997. Realtime generation of customized 3d animated explanations for knowledge-based learning environments. In *Proceedings of the Fourteenth National Conference on Artificial Intelligence*, 347–354.

Bares, W. H., and Lester, J. C. 1999. Intelligent multi-shot visualization interfaces for dynamic 3d worlds. In *Proceedings of the 1999 International Conference on Intelligent User Interfaces*, 119–126.

Bertoline, G. R.; Wiebe, E. N.; Miller, C.; and Mohler, J. L. 1997. *Technical graphics communications (2nd Ed.)*. McGraw-Hill.

Brown, J. R.; Earnshaw, R.; Jern, M.; and Vince, J. 1995. *Visualization: Using computer graphics to explore data and present information*. Wiley.

Burstein, M. H., and McDermott, D. V. 1996. Issues in the development of human-computer mixed-initiative planning. In *Cognitive Technology: In Search of a Humane Interface*, 285–303. Elsevier.

Carroll, J. M.; Mack, R. L.; and Kellogg, W. A. 1988. *Interface metaphors and user interface design*. North Holland. 45–65.

Christianson, D. B.; Anderson, S. E.; He, L. W.; Salesin, D. H.; Weld, D.; and Cohen, M. F. 1996. Declarative camera control for automatic cinematography. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, 148–155.

Cleveland, W. S. 1985. *The elements of graphing data*. Wadsworth.

Drucker, S., and Zelter, D. 1994. Intelligent camera control in a virtual environment. In *Proceedings of Graphics Interface*, 190–199.

Drucker, S., and Zelter, D. 1995. Camdroid: A system for implementing intelligent camera control. In *Proceedings of the 1995 Symposium on Interactive 3D Graphics*, 139–144.

Feiner, S. K., and McKeown, K. 1991. Automating the generation of coordinated multimedia explanations. *IEEE Computer* 24(10):33–41.

Ferguson, G., and Allen, J. 1996. Trains-95: Towards a mixed-initiative planning assistant. In *Proceedings of the Third International Conference on Artificial Intelligence Planning Systems*, 70–77.

Ferguson, G., and Allen, J. 1998. Trips: An intelligent integrated problem-solving assistant. In *Proceedings of the Fifth National Conference on Artificial Intelligence*.

Fishkin, K., and Stone, M. C. 1995. Enhanced dynamic queries via movable filters. In *Proceedings of the CHI '95 Conference*, 415–420. ACM Press.

Flach, J.; Hancock, P.; Caird, J.; and Vicente, K. 1995. *Global perspectives on the ecology of human-machine systems*. Lawrence Erlbaum.

Friedhoff, R. M., and Benzon, W. 1989. *Visualization: The second computer revolution*. Abrams.

Gleicher, M., and Witkin, A. 1992. Through-the-lens camera control. *Computer Graphics* 26(2):331–340.

He, L. W.; Cohen, M. F.; and Salesin, D. H. 1996. The virtual cinematographer: a paradigm for automatic real-time camera control and directing. In *Proceedings of the ACM SIGGRAPH '96*, 217–224. ACM Press.

J.Gibson, J. 1979. *The Ecological Approach to Visual Perception*. Houghton Mifflin.

Karp, P., and Feiner, S. 1990. Issues in the automated generation of animated presentations. In *Proceedings of Graphics Interface '90*, 39–48.

Keller, P. R., and Keller, M. M. 1993. *Visual Cues*. IEEE Press.

Kinchla, R. A. 1991. Attention. In *Annual Review of Psychology*, 711–742. Annual Reviews.

Kosslyn, S. M. 1994. *Elements of graph design*. Freeman.

1995. *Participating in Explanatory Dialogs*. MIT Press. chapter A direct manipulation interface for user feedback.

Phillips, C. B.; Badler, N.; and Granieri, J. 1992. Automatic viewing control for 3d direct manipulation. In *Proceedings of the 1992 Symposium on Interactive 3D Graphics*, 71–74.

Rich, C., and Sidner, C. L. 1998. COLLAGEN: a collaboration manager for software interface agents. *User Modeling and User-Adapted Interaction* 8(3/4).

Roth, S. F.; Chuah, M. C.; Kerpedjiev, S.; and Kolojejchick, J. A. Toward an information visualization workspace: combining multiple means of expression. *Human-Computer Interaction*.

Seligmann, D., and Feiner, S. 1991. Automated generation of intent-based 3d illustrations. *Computer Graphics* 25(4):123–132.

Silverman, B. G. 1992. Human-computer collaboration. *Human-Computer Interaction* 7:165–196.

Smith, S., and Bates, J. 1989. Towards a theory of narrative for interactive fiction. School of Computer Science CMU-CS-89-121, Carnegie Mellon University.

St. Amant, R.; Long, T.; and Dulberg, M. S. 1998. Evaluation in a navigation testbed. *Knowledge-Based Systems* 11(1):61–70.

St. Amant, R. 1997. Navigation and planning in a mixed-initiative user interface. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence*. AAAI Press. 64–69.

St. Amant, R. 1999. User interface affordances in a planning representation. *Human Computer Interaction*. In press.

Stone, M. C.; Fishkin, K.; and Bier, E. A. 1995. The movable filter as a user interface tool. In *Proceedings of the CHI '94 Conference*, 306–312. ACM Press.

Tufte, E. R. 1983. *The Visual Display of Quantitative Information*. Graphics Press.

W.Gaver, W. 1991. Technology affordances. In *Proceedings of the CHI'91 Conference*, 79–84. ACM Press.

Woods, D. D. 1991. *Human-Computer Interaction and Complex Systems*. Academic Press. chapter The cognitive engineering of problem representations, 169–188.