Knowledge Visualization: Redesigning the Human-Computer Interface

Kevin M. Dillon and Patrick J. Talbot
Northrop Grumman Mission Systems

W. Daniel Hillis
Applied Minds, Inc.

Knowledge-based computer algorithms help human operators make decisions. The complexity and sophistication of such algorithms provide sufficient computational support for successful outcomes. The real challenge is designing a visual delivery system that presents information clearly, thereby enhancing the decision-making process. The primary focus of our existing human-computer interface was on improving the processing performed by the initial algorithms, rather than presenting the information in a manner that facilitated human understanding. Our new goal was to redesign that human-computer interface so that the user could more easily access, understand, and use the information provided by the algorithms. Using a story metaphor, we identified and linked three basic visualization layers for delivery of the information: geospatial, temporal, and logical.

The underlying technology used visualization techniques that allowed humans to better understand the output of automated reasoning algorithms in support of military and intelligence community decision making. Metrics were defined to quantify the effectiveness of this technology for knowledge visualization in decision making. The design was demonstrated for a strategic command and control program. The knowledge visualization was found to be intuitive and easy to navigate. Our ongoing work will quantify the extent to which decisions are made faster, more accurately, and with smaller crews.

Introduction

Ever notice how a child quickly masters a sophisticated computer game? Yet an adult may require weeks or months to attain competency with an application no more complex than the computer game. What is the difference between the two scenarios? Adults view the game as a product designed for human interaction, while seeing the application as an example of a design geared to manage complex functionality.

The level and type of information supplied by the computer game is limited and typically low dimensional. A scenario is usually provided to the player and only relevant information is presented, so that decisions affect an anticipated outcome. The interface thus reflects the relatively simple nature of the system. Nevertheless, the gaming scenario—which tells a story—illustrates a concept that can be exploited very effectively to improve decision making in intelligence and military domains: computer-generated content can be delivered in a cohesive and easily understood manner.

Our Northrop Grumman–developed Advanced Reasoning and Information Extraction System (ARIES™) is an implementation geared to assist user understanding of computer-generated content. In sophistication, it compares with the most advanced games. Its complexity stems from the decision algorithms that transform information into actionable knowledge. With those algorithms, before ARIES, we had achieved the necessary computational support for successful military operations. However, through benign neglect, algorithm complexity had produced interface complexity. It was easier to show the output of the processing algorithms than to present information clearly. Delivery of the computer-generated content suffered as we developed a multifaceted suite of automated reasoning tools. We succeeded in merging various bits of information to tell a unified story but, at best, the style of the telling was visually confusing. With ARIES, we have overcome many of the limitations and easily applied the new capabilities to typical customer operational formats. Customer interest testifies to the potential effectiveness of the new concepts, as well as the ability to add them to operational systems.

Visualization Concept

Our goal was to find a cohesive way to articulate the knowledge produced by decision-support algorithms. Our solution was a primary display, or “dashboard” (Figure 1), that provides a mission-level box score, including indicators of mission success, status of defensive and offensive forces, and results of planning algorithms. The mission box score is driven by linked geospatial, temporal, and logical views, supported by dynamically changing evidence. The dashboard is unique in enabling the user to view the mission as an unfolding story. All mission personnel view the same dashboard, which provides access to underlying decision-support tools and enables drill-down to the supporting evidence.

The key feature that makes such a presentation effective is the display’s story-telling paradigm, covering a wide range of decision-support needs:

- It enables the mission analyst to determine (geospatially and temporally) what is going on, including highlighting the critical events and indicating how they are changing.
- It provides a logical view to assess how well we are doing, by considering the changing evidence associated with the situation and the effects of possible events.
- It enables the planners and those with approval authority to analyze the possible courses of action. The interface shows content as it supports those decisions.

To achieve the most effective decision support, we required the following:

- A situation display in the top panel with three essential features:
    - Efficient delivery of information through the use of icon color, shape, size, and movement to focus attention on key developments. The display background can be tailored to the situation, using maps, timelines, or data plots on which to present the information.
    - Visual consistency in the use of icon color, shape, size, and movement to capture critical events and their changes.
    - Ease of interaction, with metadata extracted automatically and posted.
- Real-time interaction.
- Access to detailed functionality (box score: drill down for explanation).

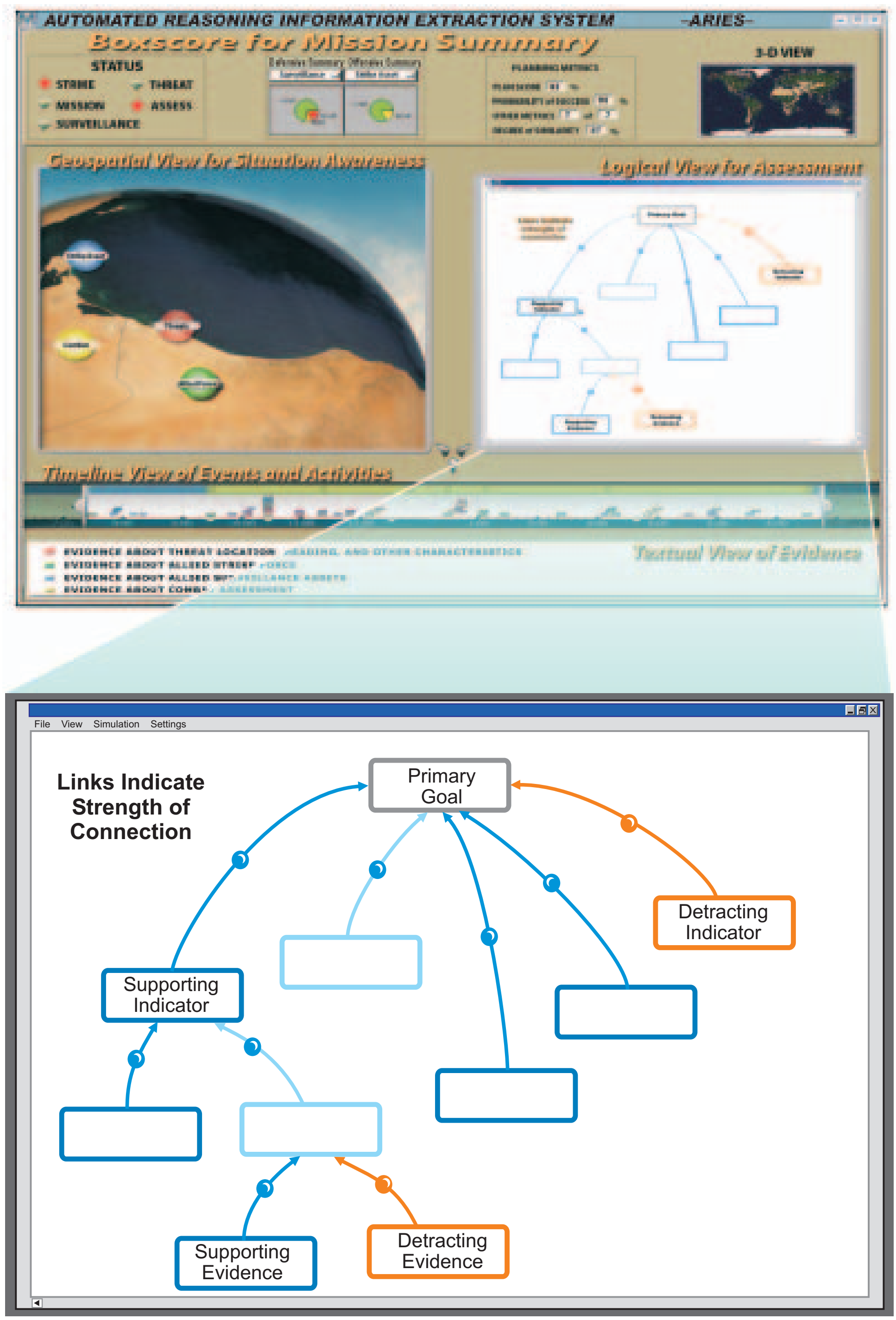


Figure 1. Visualization supports decision making

- Expression of uncertainty (logical view: shows belief, conflict, and ignorance).
- Identification of the smallest effective difference, making all visual distinctions as subtle as possible, but still clear and effective.

Uniting Algorithms with Knowledge Visualization through Storyboarding

We now describe military and intelligence community scenarios, prototype visualizations, and algorithm integration. The technology is focused on uniting the automated reasoning algorithms with knowledge visualization to produce user understanding based on the story metaphor.

We produced a *storyboard*, consisting of sketches and legacy displays, to describe a scenario as a story. Mission summaries, evidence, maps, graphs, and timelines were linked, based on a hypothetical setting, to make the story coherent. Like the dashboard shown in Figure 1 for a potential missile attack, the dashboard for this story (Figure 2) comprised a hypothetical mission summary, evidence log, and linked geospatial, logical, and temporal views. Depending on the mission parameters, those primary indicators can direct an operator's attention in a specific direction. We were able to enhance the operator's judgment, while minimizing keystrokes, as users accessed the interface. The system moved the operator through the knowledge set as he or she saw fit at whatever time he or she needed to do so, as explained next.

We developed a graphical representation of knowledge. The story (Figures 2, 3, and 4) began with situation monitoring. In this vignette, the hypothetical crew chief was briefed on the status of the potential missile attack**.** The crew monitored message traffic that was posted to a map display. The intelligence and combat assessment cells monitored the belief network simultaneously to determine how well goal values were being achieved. As tensions increased, the operations center received and accepted tasking directives from higher authority. In this vignette, the crew shifted focus to mission planning and plan optimization by interacting with optimizer metrics such as the Course of Action (COA) score (Figure 2, top panel). Authority to execute the plan was then received, the attack began, and combat assessment evidence was received at the operations center. In this vignette, the crew fused the evidence to perform mission assessment.

At this juncture, the operator had choices: depending on need or preference, he or she could view the information on each screen, or combine two or all three screens to consolidate the information. Each of the three primary screens, discussed below, offers unique benefits:

- For situational awareness—i.e., what is going on—the geospatial view provided by the *Attention Cuer* shows up to 16 characteristics of each event or activity, such as criticality, reflected in the size of the icon.
- For determining how well the crew is doing, the logical view provided by the *Belief Network Editor* (BNE) fuses uncertain evidence to present actionable knowledge.
- For determining what do, the mission optimizer metrics window driven by the *COA Analyzer* presents an image for interactive planning, which allows adjustments of the probabilities to analyze "what if" scenarios.

Attention Cuer for Situational Awareness. Many situations involve so many related, concurrent, real-time events requiring attention that available resources are stressed. To mitigate the stress, we developed a way to organize and prioritize numerous events, so

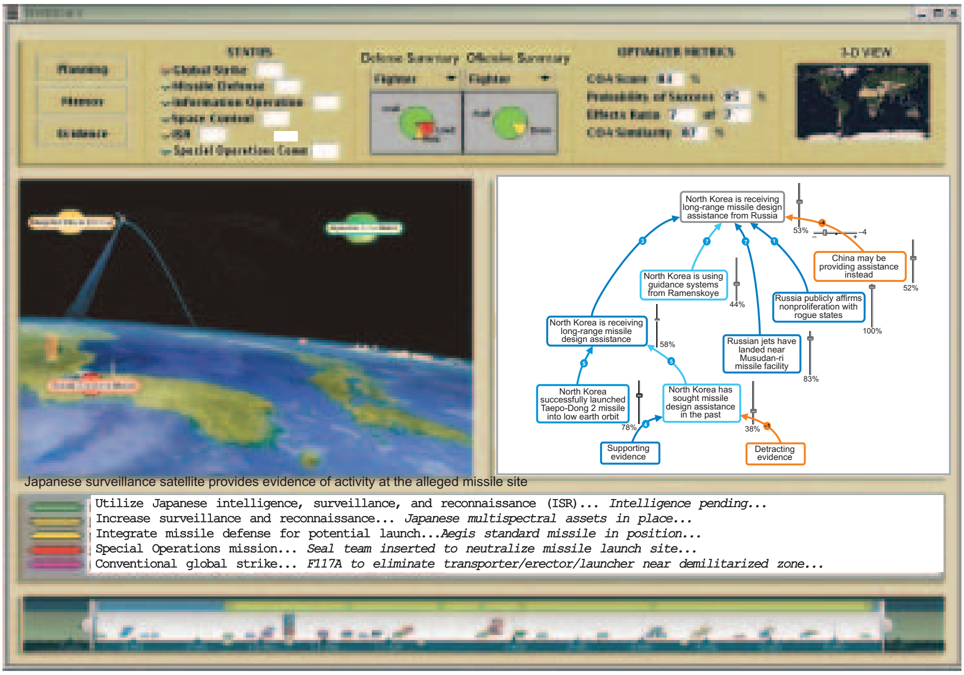


Figure 2. Primary display: The dashboard

that the operator and crew could more efficiently make timely decisions based on relevant information. The Attention Cuer applies a visual shorthand for organizing and prioritizing events. By allowing users to represent events as nodes with visual characteristics corresponding to metadata of interest to decision makers, it gives them more detailed information about what is happening. By combining information graphically, it also presents a clearer, simpler picture of the overall situation.

As Figure 3 shows, the Attention Cuer organizes and presents up to 16 dimensions, or characteristics, of a large set of concurrent events. The characteristics included three dimensions of position, three dimensions of velocity (icons move on the screen), time, time duration, event identifier (icon label), event description (text field, shown in the figure as the line of type just below the map), importance (icon size), urgency (icon glow), event category (icon color), effort required (icon shape: round for little effort, square for significant effort), and event uncertainty and confidence interval (dashed red icon and link to the belief network—click on the icon).

The information is displayed both spatially (on a map) and temporally (on a timeline) and is updated in real time, presenting a dashboard summary of the overall situation. Potential applications include a wide range of situations, such as battlefield decision making, urban disaster emergency support, and National Park Service fire hazard planning.

Additional features include the ability to create icons with various characteristics at scenario-specified times. Icons can be created on a user-specified background and positioned anywhere on the dashboard. Changes to icons can be set at specific times on the timeline. Further, the user-created visualization can be animated—e.g., for mission rehearsal, to show the addition of new nodes and changes to existing icons.

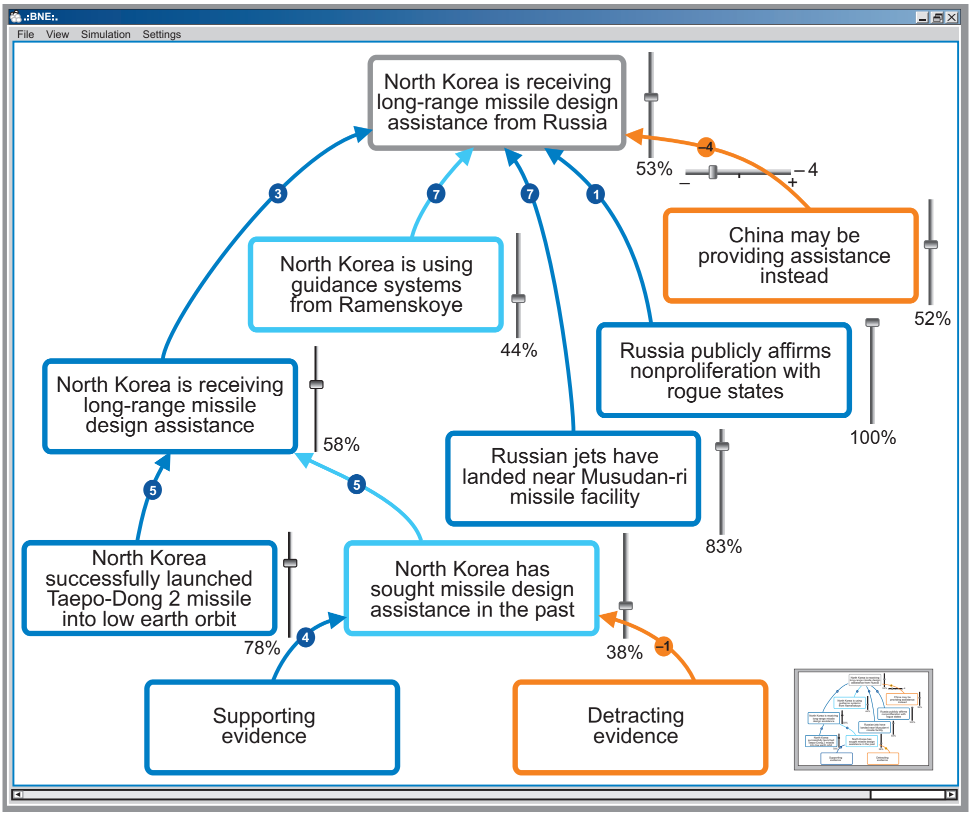
Story: The area of interest for the strategic command-and-control scenario was North Korea. A Japanese satellite (green icon) is collecting initial data in a wide-area search mode. The data reveal two convoys leaving a chemical facility in North Korea. The Integrated Missile Defense System (yellow icon) is alerted. The theater manager tasks Special Operations Forces to monitor a depot (red icon) where the two convoys may meet and load chemicals into a missile transporter, suggesting a launch may be imminent.

Figure 3. Attention Cuer for situational awareness

Belief Network Editor. Analysts, researchers, and investigators must often form hypotheses and structured arguments from imprecise, unstructured information. However, existing tools do not allow them to work intuitively with weak hypotheses, imprecise estimates, sparse evidence, and few links between hypotheses. The BNE, illustrated in Figure 4, allows users to structure arguments and evidence to produce quantitatively justified decisions in a more intuitive way.

The BNE is useful for visualizing and interacting with belief networks. It allows users to explore complex mathematical models via an easy-to-use interface in the absence of evidence. For example, "Ignorance" is the default value for any given hypothesis.

Belief networks consist of nodes and links, with each node and link having visual representations (size, color, value) corresponding to their numeric weights. The BNE is a tool for quantitative exploration of a space of lattice-shaped uncertainties (multiple influences on a hypothesis). Once the BNE is invoked, the user can perform sensitivity analyses—adjusting the sliders, which represent probabilities, to see how uncertainties in one hypothesis influence another hypothesis—as well as add or edit nodes and links,



Story: The area of interest for the fictitious scenario was North Korea. The outcome under assessment is whether North Korea is receiving missile design assistance from Russia. The assessment (53% probability) is based on four supporting indicators: design assistance, guidance systems from Ramenskoye, sightings of Russian jets near a missile facility, and Russian affirmation of a treaty. A contrary indicator is that China may be providing such assistance. The indicators are all based on lower-level postulates—e.g., North Korea has successfully launched a missile into orbit (a far-fetched idea). Notionally, indicators are shown as being driven by supporting and detracting evidence.

Note: Vertical bars show confidence in nodes (boxes). Numbers on curvilinear links show the strength of connection between nodes. In the upper right, a horizontal bar provides a slider to change the strength of the connection on a scale of –10 to +10, where –10 indicates strongest detracting evidence and +10 indicates strongest supporting evidence.

Figure 4. Belief Network Editor for combat assessment

propagate beliefs, arrange nodes in the network, collapse subtrees, and scroll through a large network.

We have used the BNE for demonstrations related to combat assessment and predictive battlespace awareness, and we have integrated it with Dempster-Shafer belief networks, which provide data fusion of uncertain evidence [1]. We have also demonstrated automated extraction of evidence from text to a knowledge base, followed by automated update of the network based on new evidence. Nearly all existing software packages with belief network visualization and editing capabilities suffer from highly burdensome interfaces and too many required inputs. We have learned that creating an easy-to-use, lightweight interface is more important than including every conceivable feature.

Course of Action Analyzer. When planning military actions, many decisions must be made with incomplete, imprecise, and sometimes conflicting goals. Interdependencies among decisions cause further complications. Existing planning tools do not function adequately in such an uncertain environment: they are template-filling tools much like Turbo-Tax. The planner completes an endless set of rigidly structured templates, including a specified set of characteristics that may or may not be relevant. The emphasis is on the plan—that is, the tangible result of filling in the templates—rather than the planning process. We have proposed a revolution in planning that gives the planner automated help to access preplanned options and optimize effects-to-be-achieved against targets with probability sliders and drag-and-drop interactions.

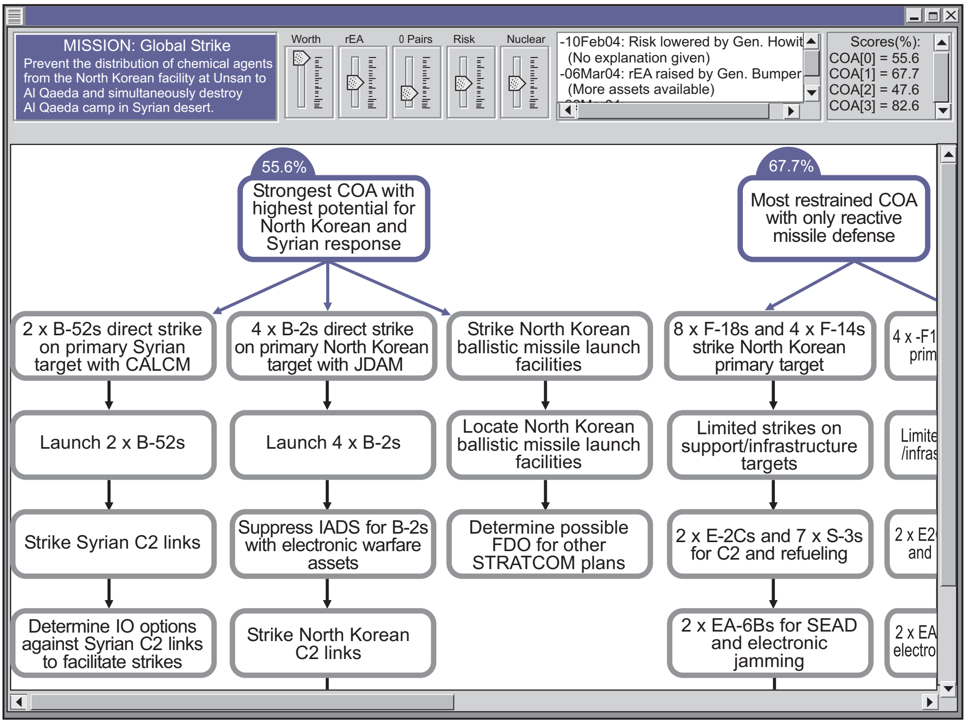
The COA Analyzer, illustrated in Figure 5, makes explicit the decision-making processes that result in strategic plans (COAs). It makes visible the mission objective, the interactions available to the planner, the history of interactions that have shaped the plan, and the library of on-the-shelf plan options used to construct the plan. It also shows quantitatively the components of the plan and its predicted effectiveness. As a result, the planner gains full control of the planning process and results—without filling templates.

Our goal was to give the user greater access to information when developing a strategic plan, or COA. We developed many innovations: visual evaluation, emphasis on planning rather than the plan, real-time "sensitivity analysis" of factors that influence decisions and COAs. The relative influence of each factor may be adjusted, and changes to the soundness of each COA decision become immediately and quantitatively visible. Simultaneous comparison of multiple COAs is possible, along with access to a library of preplanned options. We used the library to automatically extract plan fragments that were similar to planning directives.

Collaborative features allow multiple users to operate on COAs. Templates can be created and placed in a shared library to give others access to preplanned COAs. The shared templates are visible and copyable from a dedicated window. The history of COA operations is also visible, allowing the user to trace the decision-making process. Text annotations on operations can be made and shared.

Whereas the Attention Cuer and the BNE were integrated with information extraction software to demonstrate automated processing, the COA Analyzer provided a much richer opportunity. We received very favorable customer feedback at project reviews late in 2003, leading to a customer request for a working version of the nonfunctional mock-up. That development included a novel COA planning optimizer algorithm, illustrated in Figure 6, that combined text matching, case-based planning, and genetic optimization. We demonstrated the working application in spring 2004. Thus, a compelling need for knowledge visualization drove the development of an underlying algorithm to make it work. That approach was a departure from the common practice of first developing the processing modules and then determining how to show the data produced.

The COA planning optimizer algorithm included a text-matching algorithm to compare the text of an incoming planning order with the text of preplanned COAs. We then used a case-based planning algorithm [2] to compute similarity between the planning order and preplanned COAs to provide qualitative criteria, reinforced by a quantitative degree-of-similarity metric, identifying the best COAs. The best COAs provided feasible solutions to a genetic optimizer [3], which produced optimal plans for display in the COA Analyzer. Moving the sliders in the display dynamically changed both the qualitative criteria in the case-based planner and the quantitative criteria in the genetic optimizer. We also implemented a drag-and-drop feature that allowed a planner to move a planning component from one COA to another.



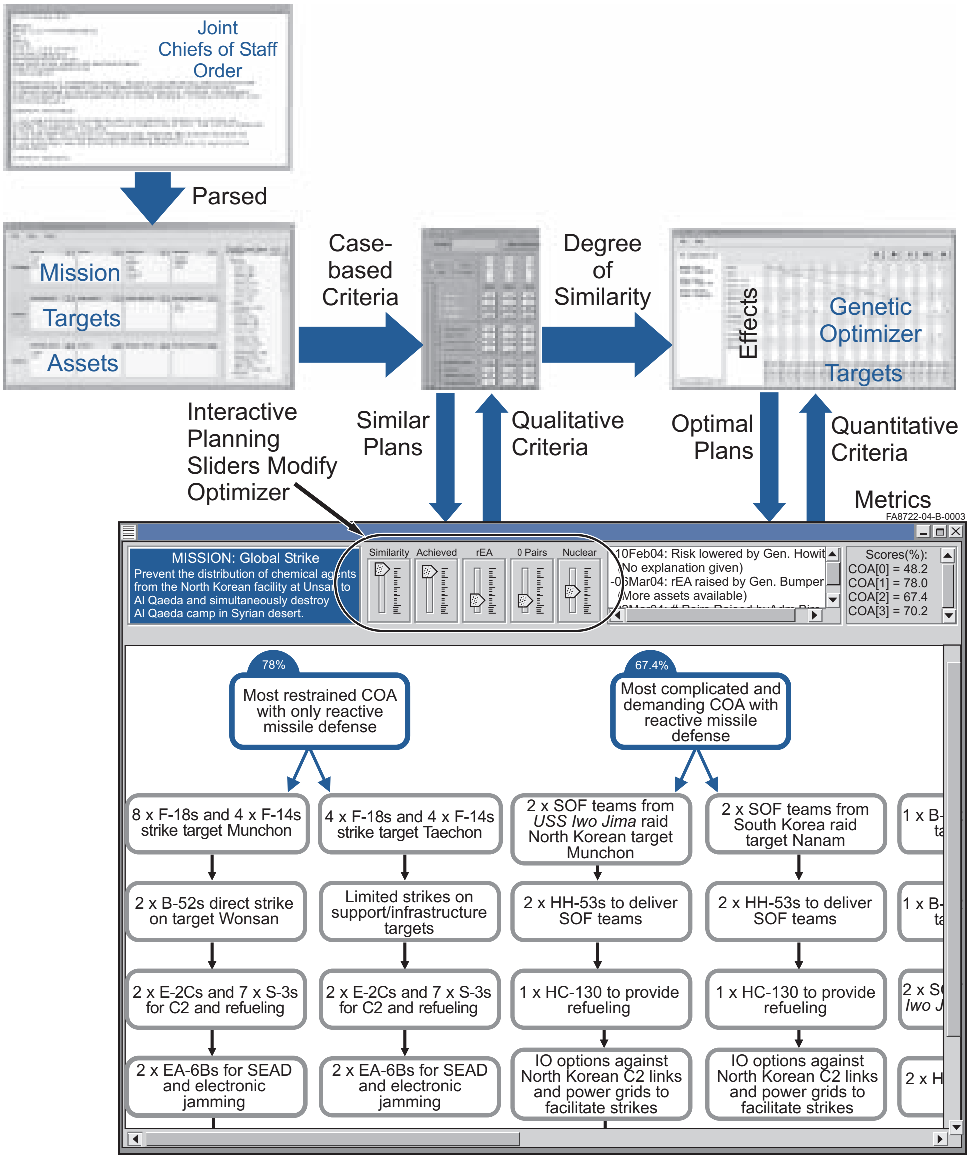
Story: The plan must address the Global Strike Mission (upper left). Relative influences (center) are adjustable by the planner and dynamically modify the optimization function. As shown, worth based on performance on similar plans is highly weighted. A low force exchange ratio (rEA) is desired. Three effect/target pairs are specified. An example of an *effect* is kinetic interceptor strategic defense; an example of a *target* is a long-range missile. Moderate risk is acceptable. A moderate nuclear weapon mix is specified. A log of previous planning activity is maintained (upper right), and scores of top-ranking courses of action (COAs) are shown (upper right). A COA (middle left) is predicted to be 55.6% effective in satisfying the criteria set by the sliders. The components of each COA are shown.

Note:
CALCM = Conventional air-launched cruise missile
JDAM = Joint Direct Attack Munition
C2 = Command and control
IADS = Integrated Air Defense System
FDO = Flexible deterrent option
STRATCOM = Strategic Command
IO = Information operations
SEAD = Suppression of enemy air defenses

Figure 5. Course of Action Analyzer for interactive planning

Figure 6 shows the processing thread, with emphasis on the underlying algorithms that produce the sample results shown in Figure 5. For example, when the sliders are moved, the genetic optimizer receives a new objective function, and results from both the case-based reasoner and the genetic optimizer are scored to reflect the modified criteria and the resulting choices of best COAs.

Common Display Format. In addition to the new visualization tools just described, we built a tool to produce commonality across already existing displays. Such a capability was required to produce consistency in legacy displays that were still necessary to provide algorithm-level detail and drill-down for explanations to more expert users and



Note: SOF = Special Operations Forces
C2 = Command and control
SEAD = Suppression of enemy air defenses
IO = Information operations

Figure 6. Course of Action planning optimizer algorithm: Optimization of effects on targets

analysts. The new tool, illustrated in Figure 7, provided a user-interactive display that allowed the user to set the foreground and background colors, font, style, borders, and many other screen attributes. The code standardized 12 attributes for about 90 screens built using the World Wide Web Consortium's interoperable XForms [4] graphical user interface (GUI) builder and about 10 displays built in Java™ code.

Development Process: A Storyboard Approach

We developed a six-step technical approach to improve knowledge visualization, as shown in Figure 8. We began with storyboards: visual representations of the flow of the story, tailored to our application as displays identifying the required sequence of decisions. Military and antiterrorism domains drove the scenarios. The legacy screens helped us to identify shortfalls and define a mitigation strategy. We derived requirements to produce objectives and define tasks. With goals in sight, Northrop Grumman and Applied Minds personnel teamed to produce prototype visualization mock-ups—typically in Adobe® Photoshop®, Macromedia® Flash®, or Java software. For visualizations of significant interest to our U.S. Strategic Command (USSTRATCOM) and intelligence community customers, we integrated existing algorithms with the displays to provide working demonstrations.

To quantify the effectiveness of our knowledge visualization design, we defined metrics and produced easily understood visualizations. The section "Visualization Metrics" below (page 51) discusses shortfalls, objectives, and effectiveness assessment. The emphasis is on measuring the benefits of knowledge visualization. The metrics work is well under way: we have quantified figures of merit, prototyped a data fusion experiment, and formulated an end-to-end set of metrics for a long processing thread. However, the high degree of automation needed to drive sophisticated logical views is illustrated by a completed workload study presented in Table 1, detailed in the "Visualization Metrics" section.

Table 1. Starlight workload study

Step	Time for Starlight Only	Time for Starlight + ARIES™
1. Data filtering	50 hr at 1 record/min	5 min at 10 records/s
2. Information extraction	5 hr at 1 record/min	0.5 min at 10 records/s
3. Reasoning	8 hr	0.5 min
4. Data conversion	25 hr at 1 record/5 min	0.5 min at 10 records/s
5. Data visualization	1 min	1 min
Total	*88 hr*	*7.5 min*

Note: Analysis assumes three intelligence collections, each containing 10^3 records with 10% of the content extracted, i.e., 300 records.

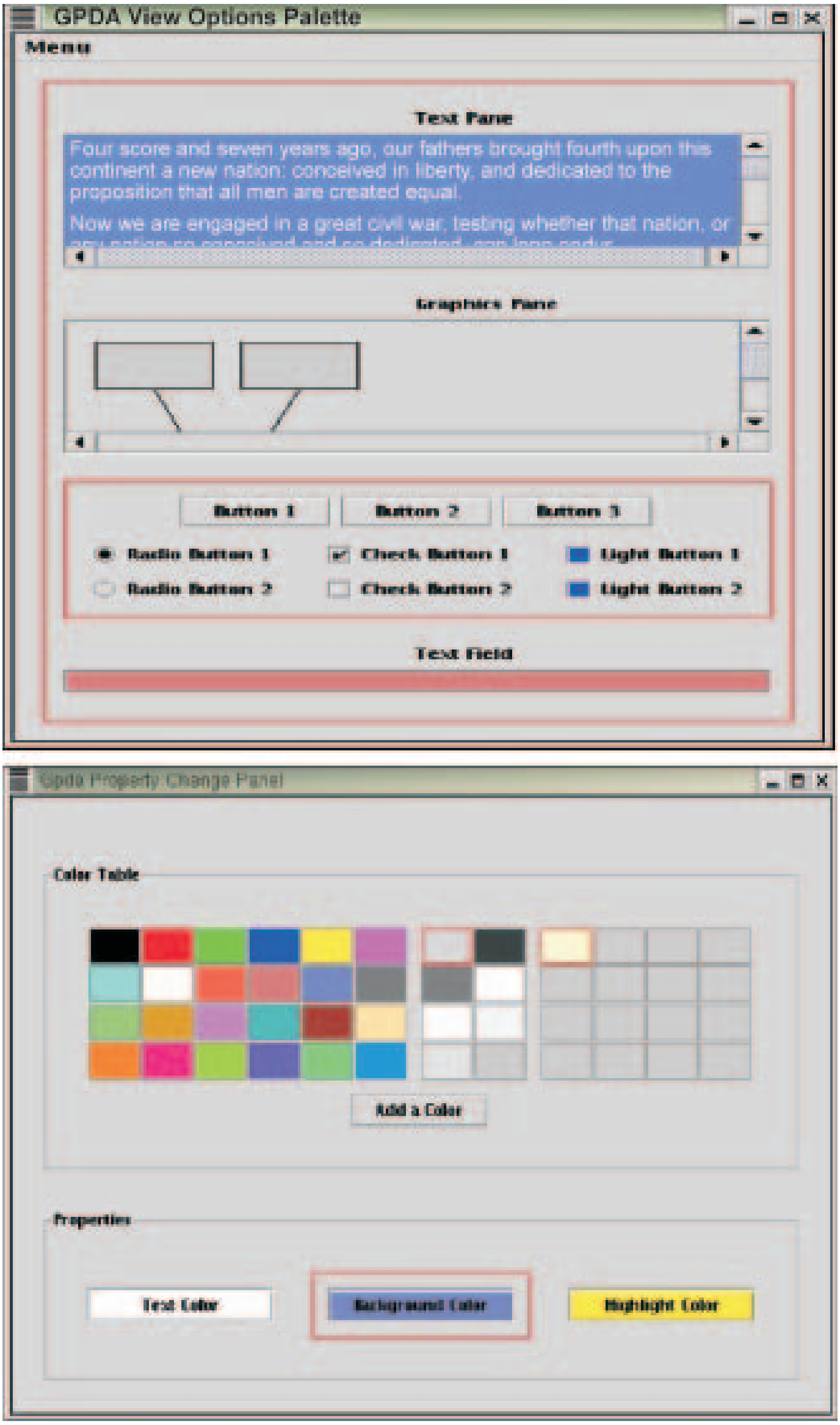


Figure 7. Common display format makes legacy displays consistent

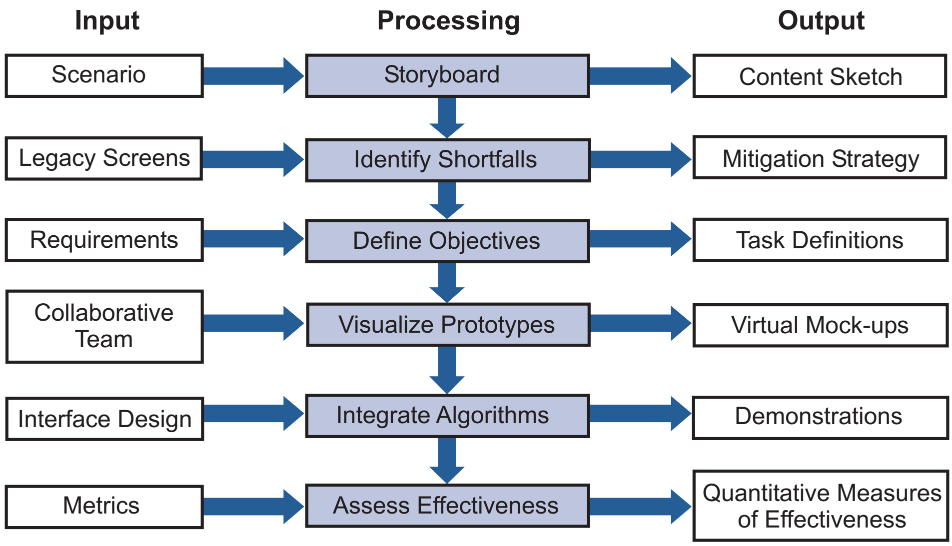


Figure 8. Technical approach: Storyboarding to improve knowledge visualization

Step 1. Constructing the Storyboard. In combination, the legacy displays showed computer-generated content. We saw the display of computer-generated content as telling a dynamically changing story, delivered in structured and textual formats. In the film industry, the use of storyboards is universal for designing the flow of the story. The paradigm was used here to aid in our interface design. It gave us a visual flow of the decisions (what's going on, what to do, how well are we doing) to be made. Storyboarding was used from the early stages of the project and helped in ongoing refinement of the look and usability of the interface, as well as the presentation of knowledge in the context of telling a story.

Step 2. Identifying Visualization Shortfalls. We then defined typical shortfalls of the storyboarded views and identified mitigation strategies, shown in Figure 9. Each display was developed in proprietary applications that generated disparate interfaces for accessing, viewing, and editing content. The only way to extract understanding from the system was to manually search out each display, extract the pertinent information, exit that application, enter another, extract another data set, exit that application, etc.

Compounding problems were the confusing nature of the inconsistent displays and the lack of flexibility inherent in a single input/output layer. As a final issue, we looked at the text-based visualization of the knowledge: the system relied heavily on delivering knowledge textually; even system navigation depended on textual input. We abandoned that tedious, error-prone interface for a metaphor-based system that lent itself to a more intuitive operator interaction. By mitigating the shortfalls, we were able to develop ideas for a robust system for visualizing a given story.

Step 3. Designing to Objectives. The primary goal was to provide an interface to deliver knowledge, i.e., tell a story, so that decisions could be made quickly and with the highest confidence. To satisfy that requirement:

- *The interface had to be consistent throughout.* From screen to screen, dialogue box to dialogue box, and algorithm to algorithm, the "look and feel" had to be consistent.

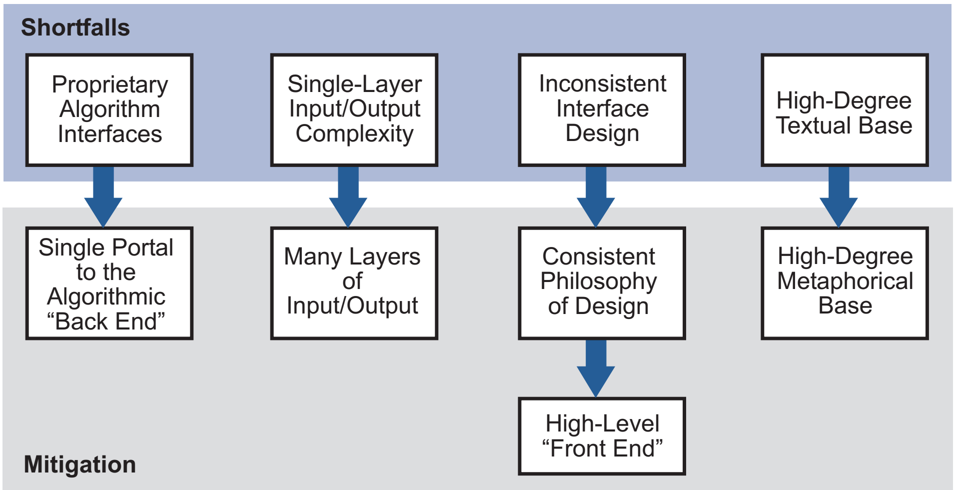


Figure 9. Visualization shortfalls and mitigation

- *The interface had to be intuitive, easy to use.* In Figure 3, for example, the Japanese surveillance icon on the map is elaborated with a text phrase (located just below the map) to highlight the most significant problem. However, the interface could not be so simple that it failed to deliver understandable information.
- *The interface had to allow the user to delve deeper into the details.* The user had to be able to drill down to an explanation of results computed by automated reasoning tools. In Figure 1, for example, the belief network allows drill-down to evidence supporting the degree of belief in the hypotheses and the input from which the evidence was derived.
- *The interface had to preserve the integrity of the evidential information as it was processed*. In Figure 1, for example, the story metaphor used maps, graphs, and timelines linked together to allow more robust user understanding of the underlying questions that a story answers: the who, what, where, when, why, and how of the computer-generated data. In our interface design we included domain-specific geospatial (what, where), logical (what, who, how, why) and temporal (what, when) visualization.

The concept of the *smallest effective difference* [5] was also part of our design philosophy In Figure 3, for example, the icons gradually change size, shape, and position. That allows us to get the message across, without requiring the observer to reconstruct the entire analysis as changes occur. Within the network of algorithms lies the information needed to make accurate decisions. The key is to be able to discern the information and not be distracted by the graphics delivering that information.

Step 4. Prototyping Visualizations. The prototyping effort was collaborative, based on a relationship forged with Applied Minds, Inc., through Northrop Grumman's Futures Laboratory initiative. The ongoing collaboration produced a design synthesis that included visual consistency (color, shape, glow) in a system that incorporated all of our knowledge-based algorithms into a clear, intuitive interface from which information can be accessed and edited by the operator. It also produced three compelling visualizations and associated algorithms: the Attention Cuer, BNE, and COA Analyzer.

Step 5. Integrating Algorithms. We identified applications that form a processing thread. Processing was automated by integrating information-extraction and knowledge-based technology to automatically populate the displays, integrated into three sequenced applications: the Attention Cuer for situational awareness, the BNE for mission assessment, and the COA Analyzer for strategic planning. The resulting demonstration was vetted both in-house and with external customers (e.g., USSTRATCOM and intelligence agencies).

Step 6. Assessing Effectiveness. To date, our primary means for assessing display effectiveness has been the "cool factor." However, though attractive and easy to work with, a "cool" display may not ensure full communication. Therefore, we sought engineering discipline by defining quantitative metrics to assess effectiveness. Of particular interest was determining how much faster, cheaper, and more accurate were the decisions based on knowledge visualizations, compared with legacy displays and ground truth (e.g., statistical results or empirical wisdom):

- *Speed:* We designed stopwatch experiments that time human decision makers in performing relevant tasks using both legacy and knowledge-based visualization.
- *Expense:* We identified workflow analysis to compute differences in man-hours to perform pertinent tasks.
- *Accuracy:* We postulated trials to validate statistical or empirical correctness.

Visualization Metrics

We have made progress in quantifying how much faster, cheaper (smaller crews), and more accurate our knowledge visualizations are for decision making, compared with legacy displays. We produced an engineering trade matrix (Table 2) that distills the requirements discussed earlier (pages 38–40) into seven categories, each of which has effectiveness criteria weighted by importance.

To date, values in the trade matrix (Table 2) are goals. We are designing human-in-the-loop experiments, based on cognitive tasks derived from a scenario. A series of tasks will be presented to human subjects with legacy displays that have knowledge visualization plus either limited or substantial automation. Subjects will then be trained on the system and asked to make decisions based on interactions with the displays. Results in the form of decision quality and time to arrive at decisions will be recorded. The evaluation process will be as automated as possible using such systems as the Automatic Mental Model Evaluator (AMME) [6,7]. Extensions of the same experiments will be tested in a crew setting to decide whether man-hours can be saved via automated knowledge visualizations using workflow tools such as IBM WebSphere®.

As an example, we performed an analytical task to compute the benefit of automated visualization for the Starlight [8] data mining and clustering tool. Results are presented in Table 1 above (page 47). Orders of magnitude in time savings are shown to be possible via automated information extraction. However, that level of visualization cannot be achieved without significant automation to relieve the knowledge acquisition bottleneck. ARIES provides that capability. The Starlight results were obtained for functions similar to ARIES. Based on our prototype work, we expect those results to be achieved or exceeded in the full ARIES implementation.

Table 2. Metrics tracking matrix

	Criterion	Weight	Legacy: Baseline	ARIES™: Goal
Quantitative	Decision performance			
	Speed	1	2 min/report	10 s/report
	Efficiency	1	200 keystrokes	20 keystrokes
	Cost effectiveness	0.8	Crew of 6	Crew of 1
	Accuracy	1	2 errors/template	0 errors/template
	Standards compliance			
	Specifications	1	None	100% defined
	Portability	0.75	Poor: C, some Java	Excellent: Web portal
	Layout	1	Ad hoc	Dynamically linked views
	Map icons	0.6	Ad hoc	15 characteristics (Attention Cuer)
	Best practices	0.6	None	100% defined
	Usability testing			
	Laboratory setup	0.75	Not available	Cyber Warfare Information Network, El Segundo
	Measuring standards	0.65	Not available	Dynamic, repeatable
Qualitative	Integration			
	Minimum requirements	0.9	Poor: Stovepipe	Excellent: Web portal
	Processing thread design	0.85	Short threads	Long threads
	Decisions made	1	Not the focus	Focal point of design
	Displays	0.85	Ad hoc	Actionable
	Algorithms	1	Ad hoc representation	Semantic net structure
	Data	1	Proprietary	Dynamically linked
	Psychological/biological			
	Bias	0.5	Unaccounted for	Preventive
	"Cool factor"	0.5	Antiquated	State of the art
	Compelling	0.85	Degrades focus	Enhances understanding
	"Uncertainty" variables	1	Belief/conflict	Adds estimate/margin of error
	Transparency	1	Displays mimic applications	Applications are transparent
	Training			
	Simulation	1	Ad hoc	Framework for simulation integration/interoperation infrastructure
	Process	1	Ad hoc	Computer aided
	Procedures	1	Ad hoc	Repeatable/evoluntionary
	Documentation			
	System comments	0.85	Seldom provided	Documented/updated

Note: Experimentation, testing, and metrics evaluation is in progress and is based on decisions made and simulated actions taken.

Conclusions and Recommendations

Decision algorithms were combined with knowledge visualization to construct decision-support systems. Those systems exploited the unique information-crunching capabilities of the algorithms, while providing understanding of computer-generated content to the user. We rated the knowledge visualization presented to the military customer against our objectives and found our requirements to be satisfied. Those requirements and other metrics were our criteria for success.

Lessons learned are summarized as follows:

- The classic Windows® environment is poorly suited to risky, real-time decision making. In mission-critical settings, no time is available to drill down through multiple menus. Uncertainty must be explicitly presented.
- Simplicity versus features is often a good trade-off when the application domain is complex or hard to understand. We relegated many “bells and whistles” to expert menus.
- *What-if* analyses are important to planners. Real-time sensitivity analysis in the BNE and COA Analyzer was very useful for detecting important influences and eliminating unimportant factors.
- Scalable GUIs are critical to making complex decision processes more tractable and accessible. We focused on scalable belief networks and attention cueing.
- Metrics for quantifying the effectiveness of user interfaces are difficult to find in the literature. A promising resource, however, is the work done to quantify the effectiveness, e.g., navigability, of Web pages.

We have described an interface that harnesses the storytelling metaphor to present computer-generated data so that it conveys understanding to an operator. We are experimenting with metrics to quantify the effectiveness of knowledge visualization.

References

1. P.J. Talbot and D.R. Ellis, “Computational Antiterrorism Solutions Using a General-Purpose Evidence Fusion Engine,” *Technology Review Journal*, Vol. 10, No. 1, Spring/Summer 2002, pp. 23–36.
2. P.J. Talbot, “Semantic Networks: A Unifying Framework for Multistrategy Reasoning,” *Technology Review Journal*, Vol. 11, No. 1, Spring/Summer 2003, pp. 59–76.
3. D.R. Ellis, “Offensive/Defensive Mission Planning: Detailed Plan Optimization Using Genetic Algorithms,” *Technology Review Journal*, Vol. 9, No. 2, Fall/Winter 2001, pp. 29–40.
4. World Wide Web Consortium (W3C®), *XForms: The Next Generation of Web Forms, www.w3.org/MarkUp/Forms/.* Accessed May 5, 2005.
5. E.R. Tufte, *Visual Explanations*, Graphics Press, Cheshire, Conn., 1997.
6. M.Y. Ivory and M.A. Hearst, “The State of the Art in Automating Usability Evaluation of User Interfaces,” *ACM Comput. Surveys*, Vol. 33, No. 4, December 2001, pp. 470–516.
7. G.W.M. Rauterberg, Automatic Mental Model Evaluator (AMME) home page, *http://www.idemployee.id.tue.nl/g.w.m.rauterberg/amme.html*. Accessed May 5, 2005.
8. Pacific Northwest National Laboratory (operated by Battelle for U.S. Department of Energy), Starlight Information Visualization Technologies, *http://starlight.pnl.gov/*. Accessed April 2, 2005.

Author Profiles

Kevin M. Dillon is a human factors engineer for Northrop Grumman Mission Systems, currently with the ARIES group in Colorado Springs. The work reported here represents his effort to bridge the gap between the graphics and engineering communities. He is now working on a wide variety of visualization projects, including the Command and Control of the Future Cross Sector Independent Research and Development project, where he is designing nongeospatial displays for the 2.5 D table. He is also working on an uncertainty management project, where he is developing ways to visualize "holes": areas in which we do not know what a potential adversary may be doing. He holds a BA in media arts and animation from the Art Institute of Colorado in Denver, Colorado.

kevin.dillon@ngc.com

Patrick J. Talbot has extensive experience with Northrop Grumman, primarily in Independent Research and Development efforts. He is currently the Chief Technologist in Northrop Grumman Mission Systems' Colorado Springs Engineering Organization. He has performed on Minuteman development, spacecraft projects, battle management tasks, antisatellite systems, and the development of decision algorithms. In addition, he has managed the Technology Insertion Studies and Analysis delivery order at the Joint National Test Facility in Colorado Springs, Colorado, where his research included parallel computing, simulated human behavior, and strategic planning. He holds BS degrees in applied mathematics and physics and an MS in physics, all from Pennsylvania State University.

patrick.talbot@ngc.com

W. Daniel (Danny) Hillis is Chairman and Chief Technology Officer of Applied Minds, Inc., a research and development company providing advanced technology, creative design, and consulting services to a variety of clients. Previously, he was Vice President, Research and Development, at Walt Disney Imagineering and a Disney Fellow. He developed new technologies and business strategies for Disney's theme parks, television, motion pictures, and Internet and consumer products businesses. Dr. Hillis is an inventor, scientist, author, and engineer. He pioneered the concept of parallel computers that is now the basis for most supercomputers, as well as the RAID disk array technology used to store large databases. He holds over 40 U.S. patents, covering parallel computers, disk arrays, forgery prevention methods, and various electronic and mechanical devices. He holds a BS in mathematics and a PhD in computer science, both from the Massachusetts Institute of Technology (MIT), where he also served as adjunct professor at the MIT Media Laboratory.

danny@appliedminds.net