1.0 Situation Awareness in the 21st Century

Dr. Charles L. Morefield

Situation awareness depends upon trusted fusion[7] software, good sensors, adequate communication links, and an effective display capability. Early in the 21st century, the Air Force will be capable of developing *automated* fusion technology for its command and control centers, aircraft, and satellites. Automated fusion technology will also be directed at the military infosphere itself.

In the view of some, appending the word "automated" to the word "fusion" constructs an oxymoron. This is largely due to the long history of failure associated with such systems in 20th century military applications. Indeed, success in this area will be difficult, and will require coordinated research programs that span several engineering and scientific disciplines. The focus of this monograph is on fusion algorithms and their associated software. However, the future success of automated fusion systems depends in fundamental ways on improvements in communications (data links and compression), computer protection, and sensors.

Fusion technology has not yet addressed the issues of situation awareness within the infosphere: how can we both monitor the information flow of our military infosphere, and automatically determine its evolving configuration? The technical challenges are complex due to our need to monitor the military infosphere without crossing the privacy boundaries our nation guarantees its citizens. Many top level tools of fusion are applicable to the infosphere since they derive from general principles of information interpretation. However, most of the lower level tools and sensors still await development. It will be particularly important to develop adequate human computer interfaces and simulation tools. They will be critical in actual operations and training, as well as the research process.

This monograph describes several aspects of the data fusion problem, and suggests useful research threads for the Air Force. It provides a short survey of a broad topic, and leaves out many issues that a longer paper would address. However, this paper does address certain issues of particular importance to the Air Force:

- Sensor volumes have already reached the point where automation is needed, since human operators alone are not capable of evaluating all information available on the battlefield
- The increasing pace of modern warfare requires more automation
- Research should be focused on the task of monitoring our own military infosphere in order to protect it against both peacetime and wartime threats
- Slow but steady progress is being made in important areas, but research must be consistently funded over a very long period to reach the goal of automated fusion

7. We use the term *fusion* in a very general sense: the evaluation of data from one or more sources to extract knowledge about events or objects of interest. We will use the term *fusion* when even a single sensor is involved.

The Broad Themes of Fusion Technology

Fusion research has a long history: several decades as a well-defined technical specialty. It began as a set of mathematical techniques applied to aircraft and missile tracking. Its techniques are rooted in probability theory and optimization. Over the years, many of its technical problems have attracted the interest of workers in the field of artificial intelligence.

It is important for the Air Force to provide new capabilities for data fusion. Research can leverage commercial technology, capturing the momentum of computer power and bandwidth, as well as the continuing evolution of commercial tools and standards for software design. In addition, the fusion community's large body of existing theoretical work provides a reasonable starting point for future work.

Data fusion involves several key tasks:

- Identify on a non-cooperative basis sufficient militarily significant targets and threats to establish an accurate real-time picture of the battlespace
- Locate targets, threats, and friendly forces with sufficient precision to support attack with available weapon systems
- Obtain sufficient situation awareness to support attack vs. threat avoidance decisions
- If attacked, understand in time to respond
- Do all the above within the threats' decision loop.

This monograph addresses the development of tools to carry out these important military tasks. There are several general observations that can be made about fusion and its relationship to Air Force operations.

Observation 1: Current information architectures supporting battlefield situation awareness are confusing. On occasion, late or incorrect information is presented to operational users. Architecture is dictated by the preference of both the operator and producer communities for autonomous execution of their roles and missions.

Workstations supporting manual fusion represent the dominant thread of fusion research. Factors that have influenced current designs include:

- Absence of an integrated battlefield information architecture
- Conflicting goals among producer and user communities
- Lack of interoperable data links
- Delays associated with manual fusion
- Security policies

Current *architectures* involve the interactions of multiple, often overlapping fusion products (see Figure 4). Products propagate through the reporting systems with different speeds, leaving conflicting viewpoints in the minds of key participants involved in battle management and execution. Biased fusion products may result from the duplicative use of underpinning sensor data. Such architectural confusion may place too much emphasis on certain objects

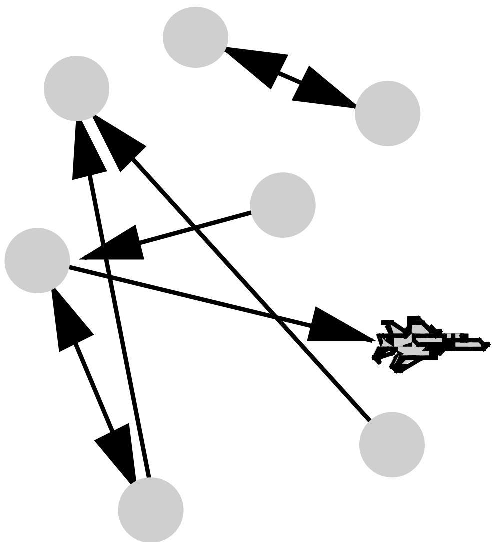

Figure 4. Current fusion architecture

causing duplicative reporting, or may lead to conflicting evaluations of the same object. Anecdotal stories abound, such as duplicative reports of missile launches. Future architectures need to be more coherently designed. In Figure 5, for example, each platform contains fusion software, and there are clear interactions among the fusion systems. The long term view shown is only one of many feasible coherent designs. The architectural paradigm shown (tree-like pathways to integrated fusion products) would work well in a very automated environment. Other sensible distributed fusion architectures are possible.

Conflicting goals have driven the evolution of battlefield information systems. Producers want wide consensus and integration before releasing information. Some users need a single vital datum immediately, while others need high confidence broader views. The producer community has (for good reasons) traditionally resisted the dissemination of information without very careful evaluation of its accuracy. Such considerations take time, frustrating end users in the heat of battle. The user and producer communities are currently discussing how to bridge these conflicting needs. They will enjoy greater success if both sides permit software mediation of the fusion process.

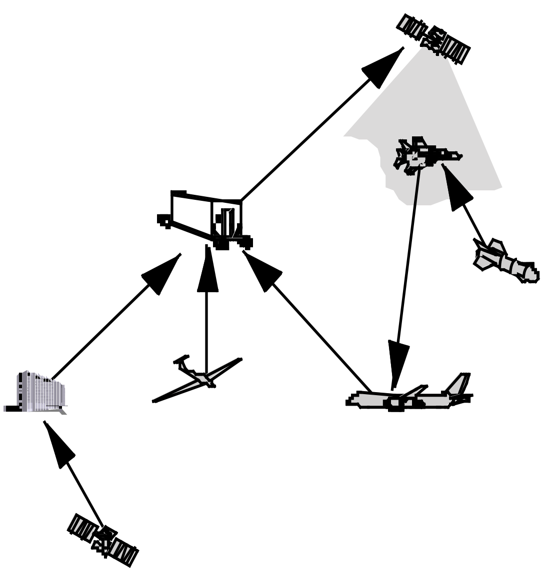

Figure 5. An example of a coherent future architecture

Data links continue to be a critical missing piece of the puzzle. We describe in other monographs a long term vision for Air Force data links. In the short and intermediate term, the Air Force already has plans to improve its connectivity. Adoption of community standard broadcast, tactical, and collection links is useful, although present versions of these links are expensive and less flexible than desirable. However, without thoughtful near term and intermediate term link planning, it will not be possible to improve situation awareness to the Air Force's satisfaction.

Queuing delays will continue as an issue in even well connected systems. The main thrust of the technology described in this monograph is automation of the fusion process, since this is the key technology for speeding up the dissemination of situation awareness information.

Security models will continue to place limits on the speed, accuracy, and wide dispersion of knowledge on the battlefield. Lacking key information, even the best analyst or fusion device reaches poor conclusions. Security (or lack thereof) has played an important role in 20th century wars, and will continue to do so in future wars. However, it is important to understand that the technical translation of security into software for situation awareness can drive architectures away from their optimal configuration by introducing quirky technical artifacts.[8] The easiest way to dispense with such issues is to give every reported fact the same security label (as is sometimes done in wartime). Absent such a possibility, we must deal with security so that that information products are not factually misleading to the end user. This will remain true even when multilevel secure systems are available.

Observation 2: Automatic fusion systems are needed to monitor the military infosphere. The advent of software munitions will require the development of new data fusion methodologies and sensors.

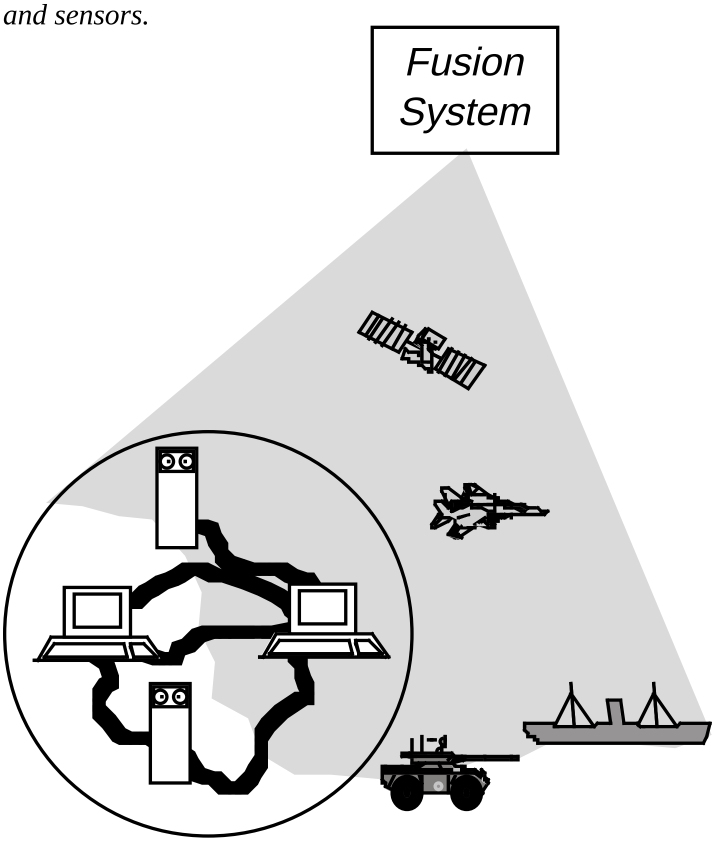

Figure 6. Fusion and sensor systems will monitor the military infosphere

Wide area computer networks have already become important for military purposes. As the military infosphere continues to evolve, a pressing need has emerged to see and understand software events as they unfold (Figure 6). Monitoring for context and intent is important for both peacetime and wartime use of the military infosphere. Current tools supporting fusion in the infosphere are very elementary. Many current tools require manual intervention and interpretation. Long term research will be required to develop automated means of advanced warning and attack assessment. As in classical applications of fusion, multiple software sensors will be required to track moving objects through networks, and to identify individual objects from their observed characteristics.

8. Technical designs that mimic (in software) the historical US security approach include the Bell-LaPadula model. This well known model (whose adequacy is still debated within the software community) reaches an occasional impasse that impacts the fusion process. For example, it requires blind writing to databases when moving from a lower to a higher level of security.

Observation 3: Attempts to automate fusion for aircraft avionics and other difficult applications have not been widely successful, leading to great skepticism about its use on the battlefield.

Dependable knowledge of the battlefield comes from the fusion of disparate information carried by communication systems of varying capabilities. Limitations in the quality and flow rate of incoming data, as well as limitations imposed by computer power and security, make *automatic* fusion very difficult. However, we are approaching an era when development of automated fusion systems will be enabled by the evolution of computers, improved sensors, and more robust communication systems. It is not yet clear that the security models adopted for these systems will evolve sufficiently to support automated fusion.

Rigorous technical approaches to automated fusion are often couched in heavily mathematical terms. They use evaluation metrics, drawn mainly from probability theory, to guide the search for an accurate fusion product. These metrics are based in turn on detailed physical considerations (for example, radar reflectivity or Kalman filter motion models).

Complete automation of this process has failed so far because of its immense computational requirements, and because robust interpretations of sensor data are hard to develop from mathematical considerations alone. (These comments apply equally to one-, two-, and n-dimensional sensor data.)

To arrive at accurate fusion products, analysts involved in manual fusion intuitively select certain mental models and tools based on their evolving situation viewpoint. As time goes on, this activity (commonsense reasoning about the *process* of fusion) can be automated. Commonsense reasoning systems to manage fusion tools are an important goal for fusion research to seek. If feasible, they will provide an important improvement in the speed and volume of fusion processing.

Data fusion technology will become particularly important to the Air Force as new 21st century threats emerge. The set piece strategic warfare of the Cold War is giving way to a patchwork of many different threats and levels of engagement. The task of adjusting the parameters of our information systems to match enemy targets will be more difficult as the stable military equipment configurations of prior years give way to greater diversity.

Robustness (with respect to changes in target parameters) is key to effective military use of these powerful technologies. Therefore, variations in military equipment will be among the critical factors in evaluating the utility of new model-based fusion systems for signals and images. Current fusion systems, such as automatic target recognizers, are often fragile with respect to variations in the target's physical configuration. Using present-day technology, we are often embarrassed at late points in the development cycle as parameters of a key threat change, or new threats appear.

Observation 4: Automatic fusion is important since it will give a qualitative advantage to the US on future battlefields.

Our opponents on the battlefield have increasing access to such important military technologies as computers, data communications, accurate navigation systems, and cheap missiles. The integration of such threats into a meaningful military capability will be their primary goal. Staying ahead of such threats will require us to radically improve our situation assessment capabilities for both manned and unmanned aircraft.

Data fusion technology will also be critical to the employment of precision weapons. In an era of powerful, cheap computers, even our least capable weapons can carry sophisticated onboard fusion algorithms and data links. Because of precision weapons, our opponents will find it attractive to hide. Camouflage, cover, and deception will be one of the very few ways they can evade more powerful sensors. They will also focus on hidden facilities and dispersion within their civilian population, knowing that our democracy considers it important to avoid civilian casualties. With precision weapons as the trademark of US Air Force operations, data fusion will be a necessary enabling technology.

Substantial progress must be made in automation if the full gain of future information architectures is to be realized. Onboard and offboard sensor data rates will increase by substantial amounts in the next century. It will not be possible to use this deluge of sensor data without new automated fusion systems that can automatically integrate multiple sources of onboard and offboard data.

Finally, data fusion is one way to improve an existing fleet in a cost constrained environment. As the Air Force enters the 21st century, many 20th century airframes will remain in its fleet. Inserting new technology into existing airframes will become an important business. Data fusion for situation assessment and targeting is one of the most cost effective insertions the Air Force can make.

Typical Fusion Applications

Fusion is an extremely broad subject. Example topics include:

- Multiple target tracking of moving air, space, ground or software objects
- Automatic target recognition based upon images, bit patterns, or time-varying signals
- Multiple *sensor* data integration, important to the process of unambiguously identifying individual objects
- Finding the position, location, and intent of someone attacking the military infosphere, and tracking the movement of their software objects through the network
- Natural language processing of text (perhaps derived from a speech recognition algorithm)

Fusion is a complex layered process, so much so that in recent years the research community has developed a standard reference model for describing its core functions. Figure 7 indicates some of the defining characteristics of this reference model, and points out areas that merit particular attention. Each layer and each issue highlighted applies equally well within the domains of imagery, position data, and objects moving through the infosphere.

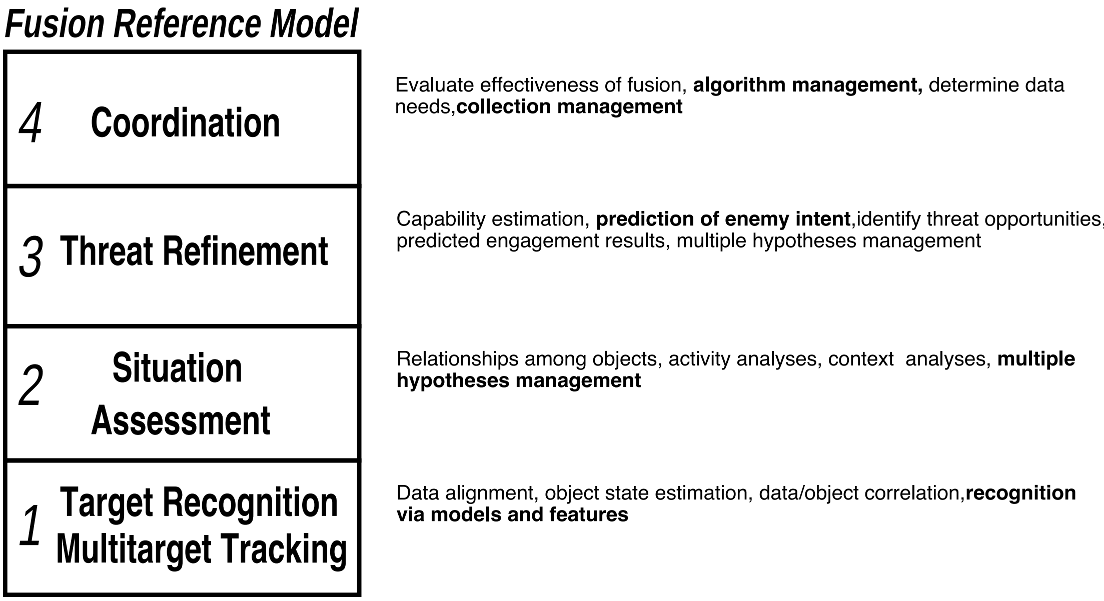

Figure 7. Multisensor fusion

Tracking Multiple Targets

Multitarget tracking is one of the primary requirements of the Air Force. Experience with large airborne radar aircraft has pointed to the utility of this function, particularly when the resulting information moves quickly to smaller attack aircraft. The problem of radar tracking of *complex closely spaced maneuvers by multiple targets* is representative of the issues involved. At the most fundamental level, the application involves one radar as it tracks multiple closely spaced targets. A basic issue is how to maintain track and identification through complex maneuvers. The algorithms for this application are decades old.[9] The fundamental problem addressed in multitarget tracking is the difficult combinatorial considerations related to initiating and maintaining tracks.

Automatic Target Recognition Based Upon Images

Automatic target recognition based upon images recognizable by a human operator has been a key problem in both military and commercial applications. In the university research community and the commercial world, its primary application has been to robot vision. Numerous government development groups attempted to bring early forms of this work to a useful product. To date, none has been successful. Research support should continue for an important range of two-dimensional sensors: synthetic aperture radar, multispectral sensors, and electro-optic sensors. Progress in automatic target recognition from two-dimensional geospatial data is important because of the great volume of such data that modern sensors can provide. Early

9. They involve calculations based upon the assignment algorithm, 0-1 integer programming, Lagrangian relaxation, and other methods.

success is possible for the more modest goal of simply assisting human operators. The goal of complete automation remains an important but longer term prospect.

Integrating Multiple Sensors

More general forms of fusion involve multiple sensors: images, signals, and others. Fusion of radar data from multiple platforms involved in multitarget tracking would be particularly useful in air combat. The viewpoints of several aircraft could be combined into a common air picture. The identification component can be supplied from analysis of other signals using statistical pattern recognition, neural nets, or other means.

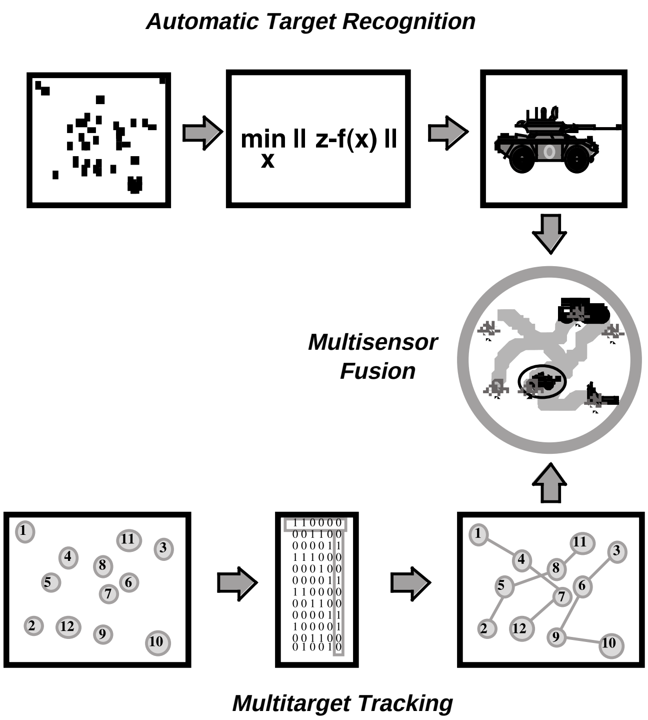

Figure 8. Core technologies for the fusion (more important denoted in bold)

The surveillance product is produced through an automated search among the likely ways data can fit together. Similar processes can be used for other sensor combinations. Figure 8 illustrates how this might come about when collecting two types of data: images and time varying position data. In this case, an automatic target recognizer (ATR) provides the identity of a key target, while a position tracker follows the time history of an entire collection of targets. In this illustration, two fusion sub-problems are solved first: one for ATR, and one for multitarget tracking. Final integration occurs at the object level at the conclusion of ATR processing.

Why Automatic Fusion Hasn’t Worked Yet

What is the current state of the art in fusion systems?[10] After two decades of research by groups in several different technical communities, automatic fusion remains an elusive goal. It has proven difficult even in the benign environment of a research

10. Level 1 functions are the most mature, with numerous algorithms tested on workstations with simulated and real data. Automatic target recognition systems today remain fragile with respect to variations in target parameters. Most Level 2 and 3 software is immature and fragile, due in part to heavy dependence on *a priori* expectations concerning target characteristics (for example the use of scripts). Expert reasoning engines are typically not integrated with the underlying Level 1 functions. Level 4 consists of mission planning, collection management, algorithm management, and response management. The mission planning and algorithm management layer is not mature. The typical system is manual, or uses a fixed algorithm execution path. Collection management has practical difficulties at the sensor level since there are typically few application program interfaces made available to system designers.

laboratory.[11] Robust solutions will require coordinated, well-funded, long term research programs. Attempts at a short term harvest of existing ideas from one or another of the various communities of interest will not produce broad solutions. Automation will require more than powerful algorithms. Attention must be paid to the clarification of surveillance architectures, sensor design, network communication systems, and simulation frameworks. Even if unique solutions are found for specific environments, the constant cat and mouse game of surveillance versus deception will require continuous evolution of our systems.

The difficulty of automatic fusion in conventional battlefield situations is replicated in the infosphere. In these applications, an added problem is the uncharted nature of the waters. We are in a fundamentally new environment, and tools that allow us to analyze information are only now emerging. Since the military infosphere intersects the private world in many ways, we will need to exercise great care that our tools are appropriate to their specific tasks.

We have found it hard to build manually guided data fusion software for workstations. We will find it even harder to design hands-off fusion software for maneuvering attack aircraft or satellites. Success in fully automated aircraft and satellite avionics fusion will depend upon a carefully coordinated effort managed by several Air Force laboratories over a long period.

Today’s fusion systems, whether embedded or built upon an open standards workstation architecture, are primarily custom-built stand-alone software designs built to the unique needs of a particular project or military task. Underneath these seemingly different designs is a core of common functions. This commonality should be exploited in future designs that fit a broader range of applications.

As this monograph is written, most operational data fusion systems are designed to provide computer support for what is still a manual process. Although the theory of data fusion is well developed, most attempts at heavy automation have been too fragile (in an algorithmic sense) for actual battlefield applications. Even in the benign environment of workstation based intelligence analysis, fundamental improvements must be made before machine based fusion becomes a useful tool.

It now appears that serious improvement in the accuracy of automatic machine-based fusion will require careful integration of functional capabilities drawn from several different technical communities. For example, sensors are needed with broader spectral diversity, better resolution, more diverse viewing geometries, longer persistence, and higher observation rates. (These issues are described in detail in monographs by the Sensor Panel.)

Research Threads for Automatic Fusion

The following paragraphs describe some of the more important of research issues.

System Architecture

New, carefully integrated architectures are needed for Air Force information systems, both at a macro-level (Joint theater) and at a micro-level (onboard aircraft or other nodes).

11. Many of the successes of commercial vision systems have come as the result of carefully controlling lighting, target orientation, or other parts of the environment. It is precisely in these areas that military fusion systems are at a disadvantage.

They should provide rational processing and dataflows for knowledge integration and dissemination to all users. Many architectural issues involving fusion and command centers, sensors, and users have been traditionally decided by invocation of a security model, or by the dedicated requirements of specific users. Less importance has been attached to the possibility of incorrect or late results being transmitted to the tactical user. Other historical impediments include lack of adequate data links, and a military acquisition process that induces stovepipe system designs. As new technology and new policy provide solutions to these problems, reengineering of military fusion systems will be accomplished more effectively.

A careful architectural comparison reveals that many fusion or command and control systems have common components. As a result, systems built for aircraft, spacecraft, or ground nodes could have an important number of common algorithms and software objects. Proper development of architecture will permit greater utilization of shared components, leading to greater affordably.

Situation Awareness in the Military Infosphere[12]

The US military will increasingly operate its computer nets across a mixed military and civilian infrastructure.[13] Security flaws exist in such systems, derived as they are from experimental research that has been quickly taken into everyday use. More potential flaws are introduced by the extension of networks to aircraft in flight. Threats to one part of the military infosphere might therefore affect the overall system as it becomes more densely interconnected through both wired and wireless links. Current military concerns are often focused on illegal entry into government ground-based computing systems. More disturbing is the issue of tampering with software objects in transit between secure systems, particularly in the wireless linkages required on the battlefield. The intent of such tampering might be to copy information contained within, or to insert some type of software munition.

Fusion research is important in the infosphere, since attacks against a network must often be *inferred* from uncertain data. Situations of this nature require analytical methods that can operate effectively even when multiple uncertain hypotheses must be maintained about the underlying situation. Many tools of fusion have been derived from general mathematical or psychological principles to deal with such problems. Some will be directly applicable to the infosphere. For most levels of the fusion model (see Figure 8), and in regard to sensors, new methods will be required.

Typical questions of concern are: Who is in our system? What are they doing now? What do they intend to do? Has a software object been tampered with in transit? Our sensors for detection are meager in comparison to the deluge of information pouring through the military infosphere. Software objects come and go in military networks in unimaginable numbers, without the explicit knowledge of their operators. Sensor design is difficult since attacks can be mounted from points of intersection with civilian open networks.

---

12. See also the monograph “Defensive Information Warfare in the 21st Century.”
13. Other monographs describe the type of robust distributed network communications that will couple various system elements regardless of their physical location, and permit knowledge delivery to the end user.

If and when the pace of computer evolution slows, it is possible that many of the vulnerabilities in open commercial systems can be closed. Today's rapid rate of development introduces vulnerabilities with each modification of software or hardware substrates.

Over time, it may also be possible to develop computing systems that are trusted by virtue of the design procedures used in the manufacture of software. Even then it will be important to understand the health of the military infosphere and the threats arrayed against it.

Fusion in Distributed Networks[14]

Fusion will take place in future years across distributed networks of sensors, computing servers, and platforms. Figure 9 indicates the environment that will be common in the next century.

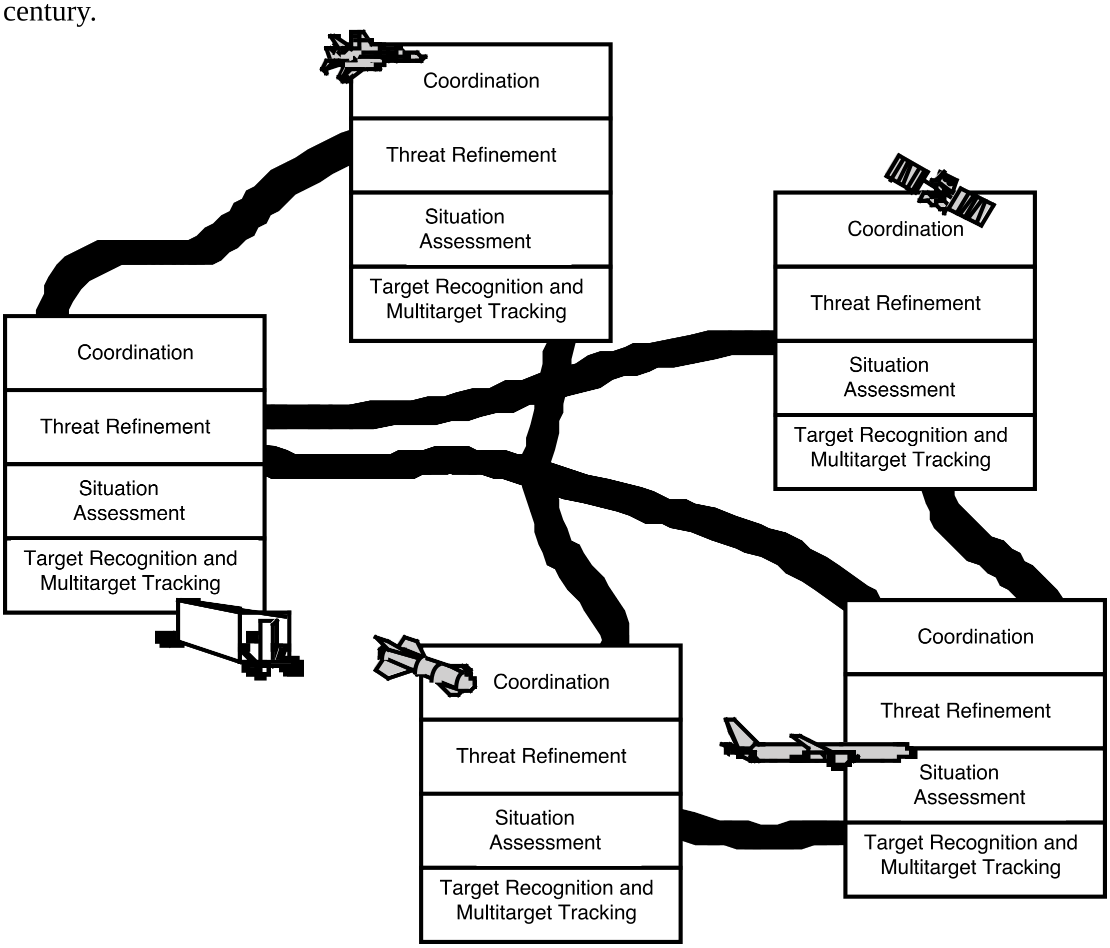

Figure 9. Fusion systems constructed to a common reference model, operating in concert across a computer communications network

14. See the Information Technology Panel's monographs for a discussion of scaleable and network-oriented computation.

There will be a continued evolution toward smarter sensors that draw more of the fusion process into the sensor system itself. Such an evolution may lead to lower bandwidth requirements throughout the data networks supporting the fusion process. This will make it possible to employ direct communication between sensor and shooter.

Current government research is developing the tools required to tightly integrate fusion systems across multiple nodes using parallel coordination languages. This approach will be particularly suitable in ground or other applications where communication links between processing nodes are relatively stable.

In more difficult environments (between aircraft and ground nodes) it may be advantageous to recast fusion algorithms on an agent-based substrate. This would represent a looser integration than parallel coordination languages, one that is more appropriate when communications can be delayed or interrupted for various reasons.

Symbol-Based Algorithms

The fundamental questions addressed in fusion overlap the artificial intelligence community in many ways. Among the more important overlaps is the following: how can common sense be integrated with numerical search algorithms to produce trusted systems for rapidly changing environments? Automation of the fusion process will depend in part upon our ability to integrate many separate tools (target models, search, and filtering algorithms) with very large amounts of domain-specific commonsense knowledge. Algorithm control (a Level 4 process in the fusion reference model) is cumbersome and fragile with traditional programming languages. As artificial intelligence technology continues its progress, reasoning systems with integrated truth maintenance should appear that are robust and practical for fusion applications.

In current systems, symbolic algorithms are used primarily on a stand-alone basis (often at Levels 2 or 3 of the reference model). Future systems will address the problem of integrating all levels of the fusion process with control algorithms drawn from artificial intelligence research.

Metrics and Models in the Fusion Process

This thread is closely aligned with the nature of specific sensor and target characteristics. It involves mathematical models and associated measurement interpretation metrics (signal processing, probability theory, model based vision and allied methods).

Other research is directed at the details of decisions made under uncertainty, which in strictly Bayesian terms becomes a lengthy calculation. Various theoretical approaches are used to alleviate this condition, and these detailed mathematical approaches will need continued support if automated fusion systems are to be built.

Search Methods

Intrinsically, data fusion is a search across uncertain data. Numerical searches and filtering techniques are required for its solution. Methods include statistical pattern recognition, neural networks, and combinatorial optimization. Relaxation methods are emerging for the decomposition of very large search processes. Tools that accelerate the search problems of fusion are still far from complete.

Planning and Resource Management

We need automated planning and scheduling systems for collection management. These systems couple sensors to the needs of end users and fusion algorithms. Early research in sensor management is currently underway within Air Force laboratories, and should be continued, particularly with a view towards its use on airborne platforms.

It will be important for sensor vendors to provide adequate software interfaces for external sensor commands at the outset. Application programming interfaces that permit external algorithms to easily couple to the sensor system are a necessary part of automatic fusion.

Security Models for Fusion

One of the major issues in the use of netted fusion systems concerns protection. Problems include those associated with the use of agents and parallel coordination languages across a distributed network where information attacks can occur.

Further development of the security and protection models applied to machine based fusion is needed. Current systems usually employ the Bell-LaPadula information disclosure model. This inhibits (downward) write and (upward) read of objects with different security labels. Other government agencies are pursuing multilevel network security research. The Air Force should also participate in this work.

Database Design

Trends toward higher volume sensor systems seem well established, making it necessary to develop the means to efficiently integrate data from very large distributed databases. Detailed and careful analysis of algorithms concerning synchronization, truth maintenance, garbage collection, and queuing delays will be required in such systems.

Commercial vendors will be able to supply much of the required database technology for workstation-based applications. However, specialized systems will be necessary for applications such as aircraft avionics.

Human Interfaces

Careful analysis is needed of the human-computer interface at the level of the end user. The key end users are commanders, analysts at workstations, and aircrews under stress in a battlefield environment. Each of these presents different issues to the interface designer.

The never ending games of camouflage, cover, and deception make it likely that fusion will many times not be able to produce a single unique opinion with high confidence. Therefore, systems will be needed that permit the presentation of multiple hypotheses to users under stress. For the airborne user, it will be particularly important to develop an approach that avoids information overload in the cockpit.

Simulation

Fusion systems are complex and require a significant amount of testing before use in operational environments. Space and airborne avionics applications are particularly stressing for these systems since they must operate in a hands-off mode when embedded in fighter, bomber, or spacecraft. Fusion applications need a carefully thought out test approach, involving several levels of activity:

- Static workstation simulations with standardized data sets
- Dynamic workstation simulations
- Distributed simulations with realistic sensor data
- Hardware in the loop simulations

Since fusion devices will be employed in highly netted circumstances, the complexity of their operational environment should be represented. For example, in fighter applications new tactics might be employed that reflect a capability for rapid automatic fusion. The pilot community is justifiably skeptical of interposing complex software modules between pilots and their sensors. Laboratory demonstrations, with pilot in the loop, will be required to effectively evaluate and justify these new approaches to sensor data processing.

New Paradigms for Computation[15]

Computational complexity is often the fundamental issue facing Air Force information applications. For example, combinatorial optimization for multitarget tracking (a key problem in information fusion) is hard in the sense that cryptography is hard. Most Air Force algorithms implicitly respond to computational complexity with approximation, since optimal solutions are often impractical in a computational sense. Special designs based on conventional microelectronics (parallel processors, digital signal processors) have been used to accelerate computation in certain military applications.

However, a fundamentally new paradigm for hard problems is emerging: computation based upon deoxyribonucleic acid (DNA) molecules. It deserves to be nurtured by the Air Force. The DNA paradigm departs radically from current computer designs. It promises extraordinary benefits to the Air Force should it prove viable: it may be possible to build DNA based machines that operate at billions of tera-operations per second.

The research thread described below is oriented toward DNA computing, rather than other important paradigms (such as quantum computing). There are several reasons for this. First, such computers may be able to solve military problems of great interest to the Air Force. Second, the Air Force needs to develop a stronger capability in the biological sciences. Third, a simple DNA computing model has already been demonstrated in Len Adleman's USC laboratory. Fourth, Dick Lipton's theoretical work at Princeton suggests that DNA models support broader computational applications than the graph algorithm demonstrated by Adleman. (These include fusion, planning, and other problems of Air Force interest.) Last, studying DNA from the viewpoint of computer science may result in new tools for biologists. Whether or not computational applications of these models prove viable, the research will add to the base of biological understanding available to the Air Force and our nation.

15. The Air Force is grateful to Dick Lipton for his help in developing the thoughts in this section.

This work is very early research, and as such may best be supported under AFOSR sponsorship. Success in understanding DNA computations will require teams composed of both computer scientists and biochemists. A simple way to make this point is that DNA is not just a sequence of letters from a four letter biochemical alphabet. It has fine structure that plays an important role in computations. Computer scientists must become aware of this finer structure if they are to make important contributions.

DNA calculations are realized by performing feasible biochemical operations. Such operations may be simple (pouring two test tubes together), or complex (such as a polymerase chain reaction). A series of biochemical operations forms the basis for a single instruction multiple data (SIMD) computer. The state of this computer consists of a set of binary sequences encoded as DNA. Biochemical operations are carried out concurrently on each strand of DNA. Each strand executes the same computation, leading to a highly parallel SIMD computer. The minimum time required to complete a problem is likely to be measured in minutes or hours. The number of operations completed within these minutes or hours is so large that the number of DNA operations per unit of time dwarfs that of classical electronic computers.

There are several important research issues to resolve before DNA computation becomes a reality. These involve physical details of the calculations. DNA computations are based on processes that are neither perfect nor infinitely scaleable. Given these physical realities, what is the correct model for DNA computations? There are many operations that are possible to perform on DNA. What are the tradeoffs among them? There are subtle differences among the many feasible algorithms for computation that impact issues such as error resistance. What is the power of certain subsets of operations? Which operations can simulate others? What paths are available to speed up DNA computations? It may be possible to uncover new methods to make improvements to individual steps in DNA computations. Such improvements may dictate the ultimate viability of this paradigm.

Another question concerns the written notation for describing DNA computations. Computer science may be able to provide molecular biology a notation for complex operations on DNA. Currently, biochemists use graphics to describe operations on molecules. Such graphics are not a useful language for describing the operation of a computer. Building computer programs for simulation will require a new formal language for the various operations. Having such a notation raises many interesting questions. Is the notation powerful enough to describe most DNA operations? Does the notation have adequate formal properties? (For example, are certain questions about it decidable?)

Can DNA computations be made error tolerant, since they rely on physical processes that are not perfect? DNA computations are based on different physical processes than those in conventional electronic machines. Conventional machines must also be concerned with error. However, the inherent reliability of electronic switches is now so good that this is not the limiting factor in modern machines. What is the error tolerance of DNA computers? What is the cost of designing for better error resistance? What is the best method?

What can we do with DNA computers? There is an exciting difference between DNA computations and classical electronic ones. DNA machines are highly parallel, but relatively slow. Lipton refers to this unexplored part of computational space as that of “ultra parallel computation" (UPC).[16] Which problems belong to UPC? Long term surveillance and planning problems should be very well matched to DNA computation. Such problems are large and complex, but do not require response times that are measured in seconds.[17] Other potential applications include cryptography and computer aided design

How can we simulate DNA-based computations? One of the lessons learned in the design and construction of electronic machines is the major role played by simulation. The same will be true in the area of DNA-based machines. Simulations can be performed on many levels.

Low level simulation of molecular dynamics will be computationally intensive and require powerful conventional machines. Such simulations may help answer a number of questions. For example, in some algorithms it is necessary to assume that particular DNA products are formed. The probability that they will form might be answered by such simulations.

Simulations at a higher level will be needed for algorithm design. For example, a simulator based upon some type of formal notation could be quite useful. Such a simulator would take as inputs simple rules that describe the computation. This would make biochemical programs much easier to describe, and increase the likelihood that computations will be specified correctly.

Conclusion

We have described a number of research threads that will, if followed over the long term, support the development of automated fusion supporting a variety of applications. These are:

- System architecture
- Situation awareness in the military infosphere
- Distributed networks and symbol based algorithms
- Metrics and models in the fusion process, and search methods
- Planning and resource management
- Security models for fusion
- Database design, human interfaces, and simulation
- New paradigms for computation

These research threads will require a long gestation period to produce the ultimate goal: fusion systems that automatically produce a very clear picture of large surveillance areas. Consistent programs with stable long term objectives are as important as funding levels.

An integrated objective architecture that spans the universe of surveillance systems employed by the Air Force should be one of the first steps. Other areas deserving immediate support include: prototype fusion systems developed for onboard applications, automatic target recognition, and techniques for distributed fusion. The area of DNA-based computation deserves early funding since years of research are needed to evaluate its potential.

16. R. Lipton, personal correspondence.
17. *Real-time* fusion is not in UPC because DNA computers need more than seconds to respond.