AFRL-IF-RS-TR-2005-187
Final Technical Report
May 2005

SOCIO-CULTURALLY ORIENTED PLAN DISCOVERY ENVIRONMENT (SCOPE)

CHI Systems, Incorporated

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

STINFO FINAL REPORT

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-IF-RS-TR-2005-187 has been reviewed and is approved for publication

APPROVED: /s/

JAMES M. NAGY
Project Engineer

FOR THE DIRECTOR: /s/

JOSEPH CAMERA, Chief
Information & Intelligence Exploitation Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	MAY 2005	Final Aug 01 – Feb 05

4. TITLE AND SUBTITLE
SOCIO-CULTURALLY ORIENTED PLAN DISCOVERY ENVIRONMENT (SCOPE)

5. FUNDING NUMBERS
C - F30602-01-C-0200
PE - 62301E
PR - EELD
TA - 01
WU - 10

6. AUTHOR(S)
Jim Eilbert

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
CHI Systems, Incorporated
1035 Virginia Drive, Suite 300
Fort Washington Pennsylvania 19034

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory/IFED
525 Brooks Road
Rome New York 13441-4505

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
AFRL-IF-RS-TR-2005-187

11. SUPPLEMENTARY NOTES

AFRL Project Engineer: James M. Nagy/IFED/(315) 330-3173/ James.Nagy@rl.af.mil

12a. DISTRIBUTION / AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

Socio-Culturally Oriented Plan Discovery Environment (SCOPE) is a link discovery project in the Evidence Assessment, Grouping, Linking, and Evaluation (EAGLE) program. The primary objective was to model terrorist organization (TO) mission plans from a counter/anti-terrorist (C/AT) point of view by finding links between evidence from disparate sources that currently might be missed. The original SCOPE approach called for using the invariant properties of the TO planning process to create global patterns that could be matched against evidence in intelligence reports. The original SCOPE architecture called for a hybrid system that performed link discovery through a synthesis of cognitive modeling and case-based reasoning (CBR) techniques using CHI Systems' iGEN environment for cognitive modeling, and a variant of SHAI's Intelligent Correlation of Evidence (ICE) system for case-based reasoning (CBR). Chi and their subcontractors (SHAI and Sytex) worked on two very distinct sets of tasks. The first set involved working with intelligence analysts (IAs) and real reports in order to determine the internal patterns used by analysts and how they use them. The second set of tasks involved designing and constructing a cognitive model of the analytic process, and applying it to the plan discovery problem.

14. SUBJECT TERMS
Link Discovery, iGEN, Cognitive Modeling, Case-Based Reasoning, Intelligence Analysis, Analysts, Mission Plans, Evidence, Patterns

15. NUMBER OF PAGES
73

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

Socio-Culturally Oriented Plan discovery Environment (SCOPE) is a link discovery project in the Evidence Assessment, Grouping, Linking, and Evaluation (EAGLE) program administered by AFRL. Initially, CHI Systems was the prime contractor, and Sytex and Stottler Henke Associates, Inc. (SHAI) were the subcontractors. SHAI only participated in the first year of the three-year contract, while Sytex participated in all three years of the contract.

Our primary objective was to model terrorist organization (TO) mission plans from a counter/anti-terrorist (C/AT) point of view. SCOPE models are intended to support the work of intelligence analysts by finding links between evidence from disparate sources that currently might be missed. In order to build this type system, we realized it was necessary to characterize the combination of hard problems that analysts must solve simultaneously during the analytic process. It turns out there are problems associated with the evidence and with the types of internal models that analysts must match to the evidence in the course of the analytic process. In contrast to the other Link Discovery teams, we spent a considerable portion of our effort on characterizing the types of patterns used by analysts and how they use them. Sections 3, 4, and 5 discuss this work.

The original SCOPE approach called for using the invariant properties of the TO planning process to create global patterns that could be matched against evidence in intelligence reports. Underlying this approach was the belief that essentially all major attacks by sophisticated terrorist organizations utilize military-style mission plan. The requirement for secrecy and the cellular structure of most terrorist organizations make communication cumbersome and long-range planning necessary. Our initial conjecture was that we could find a hierarchy of global patterns in which essentially any instance of a planned terrorist operation could fit. We called these global patterns Mission Plan Templates (MPT). Over the course of the project, we came to realize that global patterns by themselves are not adequate for discovering the execution of terrorist plans in noisy, incomplete, and incrementally arriving evidence. We found that MPTs could only be used for pattern discovery if they were imbedded in a local context, and included patterns at several levels of detail. We developed a revised version of the MPT that we believe can be built by intelligence analysts, and used by automated analysis programs to do pattern discovery which is discussed in Section 6.

The original SCOPE architecture called for a hybrid system that performed link discovery through a synthesis of cognitive modeling and case-based reasoning (CBR) techniques. The novelty in the design was this way in which we combine the technologies, and shared the knowledge stored in the MPTs. Off-the-shelf technology was used as a starting point for each of the technologies that was integrated into the initial version of SCOPE: CHI Systems' iGEN^{TM} environment for cognitive modeling, and a variant of SHAI's Intelligent Correlation of Evidence (ICE) system for case-based reasoning (CBR). Our initial plan was cut back and most of the work on probability update (which we proposed to do through Bayesian networks) as well as the work on learning was largely eliminated. As a result there were a number of places where we relied on services provided by other EAGLE contractors in order to have a complete LD component: these include evidence correlation and deconfliction, and probability update of assertions. The Technical Information Exchange group that we worked with was called OddTIE, and included AlphaTech, SRI, NRL, NYU, and University of

Wisconsin. The exchange of information between OddTIE teammates was accomplished primarily by passing hypotheses or assertions about a TO mission plan over a WebDAV server provided by SRI.

1.1. SCOPE Tasks

During the EAGLE project CHI Systems and their subcontractors worked on two very distinct sets of tasks. The first set involved working with intelligence analysts (IAs) and real reports in order to determine the internal patternsused by analysts and how they use them. The three main tasks involving real data that the CHI team worked on were:

- Collecting open source reports and putting them in a machine-readable format.
- Cataloguing the types of a priori knowledge that IA add to evidence while creating analytic products, as well as the process control knowledge needed to create these products,
- Characterizing the problems that make creating analytic products difficult.

The second set of tasks involved designing and constructing a cognitive model of the analytic process, and applying it to the plan discovery problem. Our motivation for using a cognitive modeling approach was that people are the only working example of a system that can carry out the analytic process. We designed two cognitive models of the analytic process over the course of the project. The first was adequate for dealing with the class of problems presented by the EAGLE challenge problems in years 1 and 2 of the program. Like other cognitive modeling systems (e g., SOAR, ACT-R, etc.), the iGEN model we constructed focused on the decision making aspects of cognition. It was not well suited to modeling the complex interactions between bottom-up evidence linking and matching; and top-down reasoning about context (i.e., what subset of known patterns are currently relevant), selection of evidence, and behavior selection. The second SCOPE model is based on an extended version of iGEN designed to model the interaction of bottom-up and top-down processes. It works with patterns at several different levels of specificity and temporal time scales. It is intended to deal with the full range of interacting problems that make creating analytic products difficult. It has been tested against several EAGLE challenge problems.

We conclude this introductory section with a justification of the cognitive modeling approach as a basis for automated intelligence analysis. The key question is whether the IA's approach to analysis is an optimal or even a good solution, or is it simply the best solution given human limitations. We know that human analysts suffer from a range of biases and have limited capacities. However, we believe that the dependence of people on context knowledge is not due to people being poor information processors, but rather to the lack of information in intelligence reports. In current ATR systems that use template and graph matching techniques similar to those being employed by the EAGLE program only a small number of target types can be recognized, and they are not very effective outside of the backgrounds on which they have been trained. This is not that surprising when you consider the vast number of appearances that a single target (say a tank) can present over all poses, lighting conditions, self-occlusion, and backgrounds. By selecting the appropriate local context in which to do analysis, people are able to eliminate most of the objects that could give rise to the intensity distribution seen in a region of an image. By using this technique, people are capable of recognizing thousands of objects over a wide range of backgrounds and

different viewing conditions. The downside of this technique is that people often fail to recognize novel or unexpected objects. Thus, we believe that much of the bias that people show is due to selecting an inappropriate context in which to perform analysis. Our approach attempts to take advantage of the strengths and avoid some of the pitfall of human analysis.

In order to construct a SCOPE system that can be transitioned to analysts, it must be able to determine which pattern instances in evidence are important and bring these to the attention of analysts. Like an intelligence analyst, SCOPE incrementally receives ambiguous evidence, under pressure for early judgment about the details of the TO plans. The primary output of a SCOPE model is a set of assertions about a partially executed terrorist plan and its components, together with links to the original reports about terrorist activity that support each assertion. The assertions specify who and what is involved, as well as where and when TO activity took place or is predicted to take place. Based on a SCOPE model's current assertions about a terrorist plan, past activity can explained and future activity predicted.

2. Problem Space Characterization

Generally speaking terrorists are fanatically dedicated individuals who believe they are participants in a dynamic social or political process. These people cannot, or choose not to, achieve the changes they desire through the normal political process and resort to violence. Most acts of terrorism are committed to gain publicity for their organization and purpose, to achieve political goals, or to obtain arms or financing for future operations. By performing sensational acts that attract media attention and outrage from the public, terrorists seek a government reaction that will further their cause. The Department of Defense definition of terrorism is "the calculated use of violence or the threat of violence to inculcate fear; intended to coerce or to intimidate governments or societies in the pursuit of goals that are generally political, religious, or ideological."

This definition was carefully crafted to distinguish between terrorism and other kinds of violence. The act of terrorism is defined independent of the cause that motivates it. People employ terrorist violence in the name of many causes. The tendency to label as terrorism any violent act of which we do not approve is erroneous. Terrorism is a specific kind of violence. The official definition says that terrorism is calculated. Terrorists generally know what they are doing. Their selection of a target is planned and rational. They know the effect they seek. Terrorist violence is neither spontaneous nor random. Terrorism is intended to produce fear; by implication, that fear is engendered in someone other than the victim. In other words, terrorism is a psychological act conducted for its impact on an audience.

Finally, the definition addresses goals. Terrorism may be motivated by political, religious, or ideological objectives. In a sense, terrorist goals are always political, as extremists driven by religious or ideological beliefs usually seek political power to compel society to conform to their views. The objectives of terrorism distinguish it from other violent acts aimed at personal gain, such as criminal violence. However, the definition permits including violence by organized crime when it seeks to influence government policy. Some drug cartels and other international criminal organizations engage in political action when their activities influence governmental functioning. The essence of terrorism is the intent to induce fear in someone other than its victims and to make a government or another audience change its political behavior.

While the legal distinction is clear, it rarely inhibits terrorists who convince themselves that their actions are justified by a higher law. Their single-minded dedication to a goal, however poorly it may be articulated, renders legal sanctions relatively ineffective. In contrast, war is subject to rules of international law. Terrorists recognize no rules. No person, place, or object of value is immune from terrorist attack. There are no innocents.

The major objectives of TOs lead to several operating characteristics that are used to simplify the process of building SCOPE models. TOs share the following objectives:

- Exist as an Entity of Influence,
  - Organizational Structure (often using a cell model to enable continued operations if one cell is disrupted)
  - Secrecy (avoid CT detection or interdiction)
- Effect change or achieve goals through terrorist actions which requires,
  - Continuous fund raising
  - Ongoing recruitment
  - Communication and logistic plans and actions
- Attack High Value Targets
  - Mass casualties, destruction of government personnel or facilities, destruction of national symbols,
- Maximum media coverage

These objectives, in turn, produce several characteristics of TO operations, including: Lengthy planning through execution time cycles – TO activities often unfold over long periods of time, with few explicit or overt interactions between the asymmetrical forces involved (i.e., CTs and TOs). This, plus the need for secrecy makes detailed military-style mission planning a necessity, especially when the target of the attack is difficult.
Secrecy – TOs seek to remain as invisible as possible to CTs. With fewer resources than their opponents, TOs try to keep their plans and operations entirely hidden prior to the culminating event. Likewise, CTs seek to conceal their detection means and channels from the TOs. Thus, compared to other examples of asymmetric warfare, there are relatively few direct interactions between terrorist and CT/AT groups within a mission.

Ability to truncate plans/operations – TO operations unfold in discrete steps, culminating with some overt action (e.g. assassination, bombing, kidnapping, etc.). However, indicators of CT readiness or preparation can often lead the process to be truncated. Thus, operations may or may not produce a final behavioral outcome that can be predicted, but a truncated process is to some degree an outcome in favor of the CT, while a culminating outcome is largely an outcome in favor of the TO. The tendency to abort missions as soon as the TO see indications that their activities have been detected minimizes direct confrontations. Thus, at least initially, SCOPE does not worry about C/T interactions with TOs, or how TOs behaviors or actions are impacted or altered by the actions of their opponents. This in turn allows SCOPE to treat terrorist plans as relatively static objects.

All of the characteristics listed above are strongly associated with International TOs (State Department 2000). For trans-national terrorist planning and actions greater planning, secrecy, and skill are needed since operators do not fit into their surroundings, and when more difficult targets are attacked. The line between international and local terrorism is somewhat arbitrary (Anderson and Stone 1995). In some situations, the international TO may work with a local group.

In going from conception to execution, a terrorist attack can pass through a number of phases, and the degree to which different factors influence TO behavior varies in each of these phases. There are different types of observable indicators of activity associated with each phase. For example, there are indications that when Chechen rebels found other

military means cut off, they contacted established international terrorists (http://www.stratfor.com/CIS/commentary/0103162000) and began using tactics, such as plane hijacking to attempt to advance their cause.. In the process, this group moved through a number of different domains of activity. Early in the process, there were activity patterns associated with new policy formulations (e.g., the Chechen rebel decision to use terrorist tactics). We would expect cultural and organizational factors to have a strong influence in this domain. Later in the process, activity related to alliance formation took place as a top-level Chechen decision-maker made the decision to meet with leaders of an established terrorist organization. Individual psychological makeup and theories of negotiation were probably important factors in this domain of activity. Finally, the group planned and then executed the terrorist attack, i.e. hijacked an aircraft. In this domain, activity was largely dominated by the constraints of secretive military mission planning. There may have been other domains of activity that occurred, such as infrastructure building that did not lead to an immediate attack. We believe that cultural and organizational context could play a big role in detecting infrastructure building and recruitment. It is important to note that any piece of evidence collected by an agency monitoring the situation might plausibly have fit into several of these domains of activity.

The information CT analysts actually get to see is a small fraction of the observable activity. The relevant information is buried in vast amounts of noise, clutter, and deception. The analysts know that the TOs intend to do harm, however the "who, what, when where, and how" are the critical information items the TOs try to keep hidden. Another critical factor in the plan discovery type of intent inferencing the analysts do is that the earlier they discover the plan, the better the chance for disruption, interdiction, or apprehension prior to an incident.

Social, cultural, and situational factors play a major role in determining whether an organization will consider terrorism as a tool for furthering their goals, and in target selection (Hofstede, 1984). However, mission planning constraints dominate TO activity once a target is selected. Although the prevalent religious beliefs in the TO may impact the method of attack. For example, suicide bombings are a more viable option in some groups than in others.

2.1. Types of Evidence

Intelligence analysts (IAs) work on a variety of different classes of problem. They may work on investigative problems, such as finding all of the people involved in a recent terrorist attack or find Saddam Hussein. They may address more open ended questions, i.e. does Iraq have weapons of mass destruction or are members of FARC directly involved in drug trafficking. They may also work on predictive problems, i.e. is a known terrorist group planning an attack in the next few months or when and where will a terrorist attack occur and what weapons will be used.

The evidence seen by analysts working on predictive problems is the result of activity that is planned and secretive. While they are being executed, the planned actions interact with other plans and real world processes. As a result, portions of the plan fail and contingencies are invoked. Thus, the original plan is obscured and diverted by interactions with other plans and processes. Analysts will get to see some portion of the activity generated during the execution of the plan. Several observation processes (some of which can be directed by the analyst) examine the world and deliver their outputs to different databases. The primary objective of the SCOPE system is to examine the evidence available to the analyst and discover the underlying plan or at least important parts of it before the plan is fully executed. There are three main types of evidence available to the analyst:

- Primary evidence - standard intelligence collections
  - Reports continue to come in independent of the situation as part of a normal level of security
  - Focus is on potential targets that we are trying to protect
  - Should contain some starting point or standalone indicators (i.e. indicators of secretive activity are interesting in their own right although they may tell you very little about the type of plan being executed)
  - Quality of the observation process determines percentage of indicators actually seen
  - Multiple reports about the same event (may be highly correlated or uncorrelated)
  - Contradiction and deception are common
- Secondary evidence - Communications, finance, travel, medical records
  - Covers everyone and all types of transactions, so provides many unimportant cross links between cases or groups
  - Need additional info to utilize
- Actively collected evidence - Surveillance, wire taps
  - Highly directed and limited coverage in space and time
  - Collection only occurs after a particular hypothesis is formed

Observation processes also have an impact on plan execution. When an adversary understands the limits of your primary evidence collection, they will modify their behavior to avoid observation. A problem for the United States (during the recent invasion of Iraq) was that the Iraqis understood the strengths and weaknesses of U.S. intelligence. The United States is extremely strong in technical means of intelligence, including image and signal intelligence. So, the guerrillas avoided electromagnetic communications and hid from aerial reconnaissance. They were essentially invisible to the preferred U.S. intelligence methods (Dr. George Friedman (2003) Saddam Hussein and the Dollar War. THE STRATFOR WEEKLY 18 December)

2.2. Evidence Graphs and Link Criteria

An evidence graph is a set of reported entities or events that are linked by significant relationships. All of the links can be described by predicates. One of the most fundamental issues in the analytic process is determining when pieces of evidence should be linked. In some cases, links are explicitly described in the evidence. However, a large percentage of the links must be inferred. Some relationships obviously provide a valid reason for linking evidence in a particular domain, while other relationships do not. Two people who are brothers probably have a significant relationship with each other in the domain of organized crime, while people who each have five fingers on their right hand probably should not be linked.

One possible criteria for interesting graphs comes from the graphs that analysts currently build during (criminal) investigations (Andrews and Peterson 1990). These include:

- association charts (showing who has associated with who and how often,
- telephone record charts (showing the number of calls between numbers and numbers billed to),
- event flow charts or timelines,
- criminal incident charts (i.e. what criminal activity is the group under investigation responsible for)
- Investigative activity charts (e.g. which suspects have been wire-tapped or staked out already. This is a useful way of keeping track of negative as well as positive information.,
- commodity flow charts (showing where goods or money has gone),
- financial analysis charts (including checking account transactions, business financial and/or sales records, and net worth),
- Suspect lists.

Databases of biographical information and organizational structure are also used in conjunction with these graphs, charts, and tables.

These evidence graphs, currently built by analysts, give us an idea of the range of evidence that is available and considered important by analysts. The full range of evidence that could potentially be important is much larger. Note that these different charts can be considered different views into the same set of linked evidence.

Other possible criteria for link creation comes from Standing Information Requirements (SIR) which are general in nature and not dependant on particular situations, as well as Information Requirements (IR) and Priority Information Requirements (PIR) which are targeted against a specific analytical problem set.

There are two important points about the creation of links that we call attention to at this point. One is that the analyst's a priori knowledge is needed to make decisions about which links are meaningful and useful, and this can only be done on a particular local context. The other is that in creating new links and hypothesizing missing evidence analysts are adding information to the evidence that they received initially.

2.3. Characteristics of Evidence That Make Analysis Difficult

Analysis of intelligence reports requires the simultaneous solution of a number of hard interacting problems. The main interacting problems are shown in Table 1. Perception and monitoring automated systems share essentially the same set of problems with intelligence analysis, and some of the examples cited in motivating the framework for analysis developed for SCOPE will come from these domains.

Table 1: Problems Associated With the Analysis and Visual Perception

PROBLEM TYPE	LINK DISCOVERY (LD) PROBLEM	CAUSE OF LD PROBLEM	CAUSE OF PROBLEM - SCENE RECOGNITION
Matching Plans to evidence	Hard to determine if an event belongs to the evidence graph associated with a plan execution	Flexible plans are needed to adjust to interactions with processes or other plans	Variability of object appearance due to pose, lighting, and background influence
Clutter	Similar evidence graphs not associated with interesting plans	Plans with similar sub-missions (or deception)	Objects with similar sub-structures
Entity correlation	Hard to tell if two people named in different events are the same	Use of multiple names by one person, or incomplete info collected for person	Processes causing change in object (eg., dirt on a car)
Predicate correlation	Hard to tell if two reports are about the same event	Different ways to describe the same event	Object descriptions can use different characteristics
Incremental data	Must decide ID based on partial information	Reports come in over time. Decision called for at a specific time	Recognition of situation in a picture assembled over many eye movements
Segmentation	Which events belong to which cases, or which person belongs to which group	Many attributes can be used for linking. Sub-missions (ex. surveillance) occur in most secretive plans (noise makes it harder)	Many characteristic are computed at each point in a scene. Object & background may be very similar based on observed characteristics
Limited number of differentiable patterns in a domain	Large number of possible instances per patterns leads to pattern overlap	Contingent nature of plans	Variability of object appearance due to pose, lighting, and background influence

First, analysts or data fusion systems must deal with problems related to the data itself. They must extract useful signal from the noisy, cluttered, partial information from the sensors (or intelligence reports) that arrives incrementally over time. In incremental data, one must deal with the problem of correlating objects in pairs of images (or reports) arriving at different times.

Second, analysts must recognize objects and events in the stream of evidence, particularly objects or events that pose a threat. However, recognition of objects in incomplete, noisy data is known to be an inverse problem (Granlund & Moe, 1994). This implies that there is no general way to determine whether a pixel in an image (or a person in a report) is part of a threat object (or a threat organization) without recognizing the object (organization). To

solve this inverse problem, a hypothesis about a model or pattern must first be made. Actual evidence can then be compared to models likely to be relevant in the current situation. If there are too many potential models, making the necessary comparisons with the evidence becomes an intractable problem. So, the inverse nature of the problem requires a method of filtering all possible models down to a small set of potential models that must be checked at any point. In fact, the analyst must solve recognition problems on several scales, ranging from recognizing localized objects to recognizing the situation based on the current evidence. The number of plausible global models that could match the current situation is generally very large. Consider the execution of a military plan for a terrorist attack when you actually know the plan. Over time, events occur in the world that require replanning or the selection of contingent branches of the plan. After a short while, it becomes impossible to determine whether any particular event is related to the execution of that plan or not. Therefore, the set of plausible global models, when the analyst doesn't know the plan and there is limited evidence is generally very large. To make matters worse, many, if not most, of the models needed to recognize threats that are unknown.

One of the most basic functions of Link Discovery (LD) components is linking evidence to construct evidence graphs that correspond to planned activity. The LD systems must then decide which evidence graphs are associated with "bad" activity planned by opposition groups and which are not. This decision must involve knowledge about bad activity and a matching process. This section focuses on the hard problems that must be solved in order to make this decision. It is especially important to know the characteristics that make a problem hard when you are testing against simulated data that does not necessarily have all of these characteristics.

2.4. Examples of Analytic Problems

2.4.1. Problem: Breadth or Variability of Evidence Graphs Associated with a Plan

In both the ROC and the terrorist domains, the most important driver of dynamic behavior is the interaction between the contingent plans of hostile, asymmetric groups. While following a plan greatly constrains the decisions and actions that that the group supporting the plan will perform over an extended period of time, the interaction of plans (and the interaction of plans and natural processes) has just the opposite effect. Even if the analyst knows the exact plan initially formulated by a TO, it can still be difficult to determine whether a particular piece of evidence is actually the result of that plan being executed. This is because all plans have contingencies built into them, and as other plans and processes in the world interact with the original plan, parts of the plan cannot be executed and contingent plans must be used.

For example, when a group fails to deliver the nerve gas it promised, the terrorist group expecting the gas may change plans. They may decide to carry out a contingent plan for a bombing attack. Evidence that made it appear that a gas attack was being planned can change character, and start looking more like a bombing attack. Thus, the interactions of plans and processes and the contingencies that are included in most plans leads to a much broader range of activity that is consistent with the plan than one would expect if they were anticipating the simple execution of a plan. These issues make predicting key terminal events problematic and error prone.

2.4.2. Problem: Clutter and Density of Evidence Graphs

Another reason that it can be difficult to decide whether an event results from a particular secretive plan is that other secretive plans may be carried out in the same location by other groups. Some events resulting from a terrorist plan may look a lot like events resulting from smuggling, extortion, industrial espionage, or military spying. When such activities are intermingled in time and space the result is clutter for the analyst trying to determine whether an event is the result of a particular plan.

Clutter includes not only a notion of similarity between clutter and target objects, but also an implicit notion that clutter and target objects can be found close to each other. Unfortunately, terrorist activities are spread wide in both time and space. Clutter cannot be based on complete patterns (Nothing else can really be confused with something like the attack on the Cole.) Clutter in the terrorist domain must be based on indicators of sub-mission execution, i.e. sub-patterns. In the ROC domain, the most common clutter is criminal activity other than the types that you are looking for.

2.4.3. Problem: Entity Correlation Resolution-

The question is whether two people (or weapons, or vehicles, or meetings, etc.) are the same. Unknown persons generally comprise a large part of the entity correlation unknown in intelligence data. Many variations of this have been addresses in data fusion domains. One of the most thoroughly worked is the tract correlation problem, which asks “what is the probability that two radar tracks are produced by the same vehicle?”

There are known solutions to the tract correlation problem, which are probably applicable in our domain. This problem includes the name deconfliction problem, where there are multiple names for the same person.

2.4.4. Problem: Predicate Correlation Problems –

The basic issue is that different people may describe the same thing in different ways, and with different levels of detail. This problem was stated very clearly by Ben Rode in a document he put together for EE.

In general the IA will not have information about the specific plan being executed. They are likely to know about cases that resulted from the execution of similar plans, or patterns of evidence associated with a class of plans. However, in many, if not most cases, an analyst will not have a detailed top-level pattern associated with the plan that the terrorists of criminals are following. Therefore, it is important to be able to link evidence that is associated with the same cases without top-level patterns. Even when top-level patterns are available IA still need to be able to link evidence without patterns, because evidence comes in incrementally and only a small portion of the pattern is observable until most of a planned mission has been executed. Thus, the major problems associated with linking evidence relate to incremental data and to case segmentation.

2.4.5. Problem: Incremental Data (A Complete Picture of a Case Is Not Seen All At Once)

Information arrives incrementally, so the analyst must assemble the parts in order to see the complete structure of a case. People tend to misremember attributes of a scene when they are mentally reconstructing it. Thus, a blue sweater that someone may have worn during one event may be remembered as red, particularly if someone suggests that it was red. A computer program that assembles information from an inconsistent set of reports is likely to have a similar problem.

There is a difference between actively collected evidence that will never be seen unless it is collected in a particular time interval (ex. a wire tap), and evidence that is sitting in a database and can be retrieved with a query at any point after it is entered into the DB. The actively collected evidence requires a degree of planning and reasoning about asset allocation beyond what is needed to query databases of secondary evidence.

2.4.6. Problem: Case Segmentation

The general processing sequence used by analysts to find a set of linked evidence associated with the execution of one plan is to:

1. Build out from one or more SAI by finding evidence that can be linked to the SAI (These evidence graphs tend to be associated with sub-missions within the plan.).
2. Recognize the basic sub-mission corresponding to evidence graphs.
3. Group these basic components (with looser or few links between groups than within groups) into a larger evidence graph that may correspond to a single case.
4. Recognized or identified the evidence graphs as being an example of a particular type of plan or top-level pattern.

In Section 4, algorithms for these steps, the a priori knowledge employed by the analyst, and the interaction between the different steps are discussed in more detail.

There are several problems in finding evidence graphs tied together with relevant links. The first problem is finding SAI in the first place. However, the major problem is that there are large number reasons for linking any two pieces of primary evidence given the known attributes of the evidence. While on the other hand, attributes that may be a valid reason for linking may be unknown. This means that relevant links within the evidence graph corresponding to a mission will be missed, while at the same time evidence from different cases will be linked together. Thus, analysts generally face a case segmentation problem when looking at a stream of evidence. This problem and its consequences are more easily understood in relation to an analogous problem from computer vision, namely the difference in segmenting objects in the blocks world vs. the real world. In the simulated blocks world, objects clearly stand out from background based on color or intensity alone. Simple thresholding is sufficient to separate background pixels from object pixels. In addition, there is always a visible edge between any two blocks or a block and the background. Since there may also be interior edges, the hard part of the segmentation problem in a blocks world is determining whether an edge is interior or exterior. Edge and corner information is sufficient to determine shape in the blocks world. However, in the real world you do not necessarily have clear edges between objects and background. Edges between objects may only be indicated by a texture difference, or there may be no visible difference. In the general case, segmenting the objects in an image is an unsolved problem after 30 years of research.

An example of a segmentation problem for the evidence graphs is seen in the following evidence. “John, a suspected terrorist, is seen taking pictures with a camera. Some evidence has already been collected showing that John’s group may be involved in carrying out a terrorist activity.” Should this evidence be linked to an evidence graph that is hypothesized to represent the execution of a terrorist plan, or is it an irrelevant action. Is John doing surveillance in support of an upcoming terrorist attack, or is he just taking photos because it's his hobby? If this picture taking evidence is linked to the potential terrorist activity graph, what other pieces of evidence will end up linked to the evidence graph no matter how stringent your linking criteria? A myriad of different ways of linking evidence make this example the rule rather than the exception. This segmentation problem applies to both organizational or event graphs.

Our work with analysts suggests that what they use to decide whether two pieces of evidence belong to the same case are standalone indicators and a set of context-sensitive linking criteria. The standalone indicators (SAI) found in primary evidence generally do not tell the analyst what type of plan is being executed, just that the activity is much more likely to be associated with secretive, planned activity than with normal daily activity. Note that much of the evidence linked to an SAI will come from secondary evidence.

There are two parts to group identification problem: determining group membership and determining group structure. Group membership is a type of segmentation problem. It is relatively difficult because, in most organizations, the level of participation in group activities found among members varies widely. Organizations may have members who pay minimal dues, but do not participate in any group activities.

The group structure is a more complex problem, and there may be several types of structure within a single organization. The places where commands can originate and the paths they can follow determines the command structure of the group; the paths along which general information can be passed through the group form the communication structure; and the paths along which assets can be passed form the group's financial structure. Adding to the complication many group members participate in many groups. Social network analysis is generally applied to work environments where structure can be determined statistically, since people are working between certain hours.

A last point that needs to be made about the segmentation problem is that it is a completely different problem than the clutter problem. However, segmentation problems can interact with clutter making that problem more difficult to solve.

If you have identified a pattern, one way to use the pattern is find entities or events that fit into the pattern and check whether they are linked by the relationships specified in the pattern. In this case, the pattern is used to probe the evidence in order to find exemplars of the pattern. This implies that if you have a pattern, then you don't need to worry about pattern free segmentation (assuming enough incremental data has already been collected) you can go directly to probing the evidence for instance. This approach is made difficult by the following problem.

2.4.7. Problem: When Probing Data With a Pattern Only a Limited Number of Patterns Can Be Distinguished

This problem comes up in ATR, because while you have very specific models of targets, when you consider a 2D projection of the target into a scene the way that it looks depends on its pose and its illumination. This uncertainty in what a target will look like is analogous to the uncertainty in the evidence associated with a plan due to the contingencies that are invoked as the plan interacts with other plans and processes.

2.4.8. Problem: Non-stationary Nature of Plans

Over time the techniques that a TO organization will change. There is turn over in the membership. New techniques and weapons may be tried and adopted. Thus, the system carrying out the analytic process must be able to adapt to these changes.

2.4.9. Problem: Inferring Network Structure From Its Input Response

Group response to a stimulus depends on both the structure of the group and the input. In the study of terrorist groups, neither the group structure nor the input characteristics are known in detail. The communication structure within these group share many characteristics with models of activity within 2D interneuron sheet found in the mammalian cerebral cortex. These nets will support a variety of response modes to a point stimulus, including a response that dies away, a steady state response, traveling waves. Thus, it seems that small changes in the input to a social network can result in changing the response from waves of communication that continue to reverberate through the network for a long time to strong response throughout the net that dies away quickly. Thus, seeing a particular mode of response may say little about the network structure.

3. Situation Estimation from Perceptual Inputs

Over a number of years the US military has been very interested in situation awareness. This interest has recently intensified with the new emphasis on netcentric operations. The primary focus in essentially all of the work on situation awareness sponsored by the military has been on passing people the information that will make them aware of the current situation and improve their decision making, rather than on the process of making situation estimations. However, in analysis, the objective really is making good estimates of the current situation given the noisy, incomplete, incrementally arriving evidence available to the senses. While their objective may not be situation estimation, people who study situation awareness (SA) have realized that there is a cyclic process involving what you know and what you perceive. "People are very active participants in the situation assessment process, with SA guiding the process and the process resulting in the SA." (Endsley 2000).

Situation estimation, which lies at the heart of the perceptual and analytic processes, requires a framework that deals explicitly with key factors that are at the same time known yet ignored in cognitive architecture research:

- Plans and stories - encode large portions of human social and cultural knowledge, provide plausible models of future events, and highlight the events and objects that people believe are important in a local context.
- Cognitive maps or spatial memory - allow us to return to places that we've visited before, visualize places that we are going to, and encode the locations and relations among the important objects in a context.
- Context switching – cognitive and perceptual processing is done within a large number of local contexts that are organized into a hierarchical structure.
- Motivation and emotion - play huge roles in estimating the current situation, selecting strategies (fight/flight, lead/follow, etc.), and in determining what experience goes into memory.

The foundations for a computational representation of the phenomena described above converge from both psychological and neuromorphic directions, and are constrained by issues of computational efficiency.

Our estimate of the situation has two complementary components:

- Context knowledge that specifies the boundary of the region that is relevant, the important objects within that region, and the plans or stories that can be carried out in the region.
- An estimate of the current state of the world that results from matching recent sensations against models of objects, situations and stories that are relevant in the current context. (So, the estimate will contain an explanation of how the current state came about, and the probable trajectory of future states); and

Given the similarity of the problems involved in intelligence analysis and in recognizing the situation in a scene, we conjecture that people use similar techniques to recognize situations in either domain. The techniques and strategies that people employ in doing object recognition, as well as the neurological substrate supporting it, have been more thoroughly researched than those used in intelligence analysis. In the rest of the section, we discuss neural and psychological evidence about how people recognize a situation. We will also

look at some state-of-the-art image-processing techniques that have been developed to deal with various problems listed above. Automatically extracting feature and object information from individual images has proven to be one of the more difficult problems taken on by Artificial Intelligence (Fischler & Bolles, 1996). Researchers have been working on computer vision systems, with moderate success in certain limited domains, for the last 25 years. People are still far more successful than any automated system when operating on scenes with a broad range of natural objects.

Human visual processing leading to the recognition of the situation in an image (or a sequence of images) addresses essentially all of the problems associated with analysis cited above. Deciding whether a view of part of an object corresponds to an object for which you have a detailed model is made difficult by the many ways the object can look in different poses and types of lighting. Thus, an object's appearance varies in a way that is analogous to the variety of events that could result from the execution of a single plan. The clutter, noise, and segmentation problems in imagery that have been addressed by computer vision systems for more than 25 years. Entity correlation between models and objects in an image is made difficult by occlusion (i.e. partial information) as well as variability in the appearance on an individual object. Human vision involves incremental processing of information collected over many eye movements around the scene to build up an understanding of the situation. Many different scenes may correspond to one situation, just as many evidence graphs may correspond to a single plan.

Focusing on situation estimation requires using a different type of model of visual processing and image understanding than have been used in the past. Essentially, all previous models have focused on object recognition. We argue that for people object recognition is always an intermediate step on the way to situation recognition. Recognizing that an object is a tiger is not a biologically significant event, while recognizing that you are in an open space with a man-eating animal that you cannot fight or outrun is quite significant. Recognizing the tiger situation should cause surprise and an immediate change to the local context the person is operating in. The situation estimate is colored both by emotions and motivations. It includes expectations developed by comparing the sequence of situations leading to the present situation to known plans or stories.

This section introduces a framework for a cognitive architecture that focuses on situation estimation and the interaction of perceptual and cognitive processing. This framework can address the set of hard interacting problems described in Section 2. Two principles underlying the structure of the framework are:

- Situation estimate is almost always the result of an update or revision of a previous situation estimate.
- A combination of local "context" information is needed to drive the cycles of situation update and revision, and people learn a large number of these.

The framework also requires process control and attentional mechanisms to direct evidence collection and switching between local contexts during cycles of situation estimation.

To summarize the following three principles are incorporated into our framework:

- Humans address the visual perception problem is the critical role played by context. A variety of different types of information must be learned together in the process of building a local context. A local context must contain the various types of models needed to perform situation estimation.
- The situation estimation process can only work if there are cycles of situation update and revision. People are generally well aware of the sequence of events that lead to the current situation they are perceiving.
- Process control knowledge plays a large role in both analysis and perception.

3.1. Behavioral Evidence about the Mechanisms of Situation Estimation

People are capable of recognizing objects over a wide range of backgrounds and different viewing conditions. However, they are not particularly good at recognizing even familiar objects in unfamiliar contexts. For example, *Puzzle* magazine features a section on identifying common objects in real images taken at unusual scales. Similarly, finding familiar objects in a SAR image is quite difficult until one gets used to the differences in how objects look in an unfamiliar imaging modality. A researcher has told me that on his first flight on a fighter aircraft he was unable find the golf course that the pilot asked him to find. The pilot, of course, had no trouble spotting it. In this case, scale, viewing angle, and speed were all unfamiliar to the researcher, who became much better at spotting things after a few flights. These examples point out that people need to go through a learning process in order to extend their ability to recognize objects to unfamiliar settings.

While people must learn to operate effectively in unfamiliar environments, what they learn in order to become familiar with the new environments is not obvious. Looking at the routines that a person employs to find and utilize things during object recognition, shows that they have knowledge about what they are likely to see, about what these objects look like in the local context, and what characteristics of the imagery are important when segmenting objects. These types of knowledge must be learned in order to operate in a new context. People must also be able to recognize the appropriate local context in the first place. Fauconnier (1998) calls the knowledge that allows a person to function effectively in an environment a *mental space*. Our final version of an MPT has many similarities to Mental Spaces.

Knowledge about the visual routines that should be used to find and recognize objects is part of a mental space. In a familiar context, a person tends to switch seamlessly from step to step in a visual routine. He or she may find certain features before others to bootstrap finding the more difficult ones to recognize, or move back and forth between scales in a particular order. Individual operations in a processing sequence include activities such as finding landmarks, zooming in and out, following a road or a river, and comparing areas that may have the same classification.

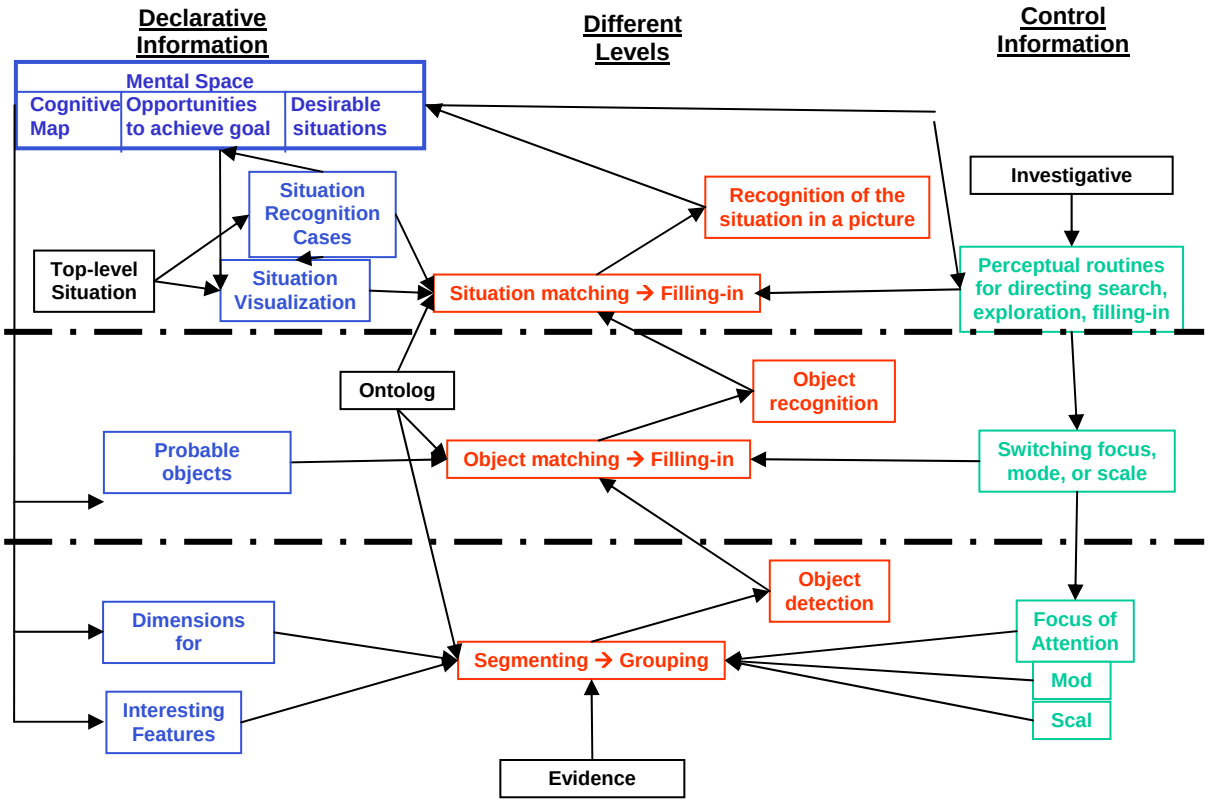

Figure 1: Model of Cognitive Process Leading to Recognition of the Situation in a Picture

In Figure 1, we present a model of situation recognition with three levels of processing and different time scales at each level. The levels are the rows in the Figure, and are separated with heavy dotted lines. The output, on thee lowest level of processing, is object components that are detected, but not recognized. Cycles on this level happen every time there is a saccadic eye movement. The major processes at this level are finding boundaries (or alternatively growing regions) and grouping regions. People operate in at least three different behavior modes while doing segmentation, and the components found can vary greatly from mode to mode. Important visual behavior modes include search for a particular object, exploration to locate interesting objects, and tracking. Very different eye movement patterns are associated with these behavior modes (Yarbus 1967). In addition, different modes are associated with differences in the number of objects that can be reported based on one glance at an image, as well as the amount of detail that can be reported for each object. In exploration, a person will report the largest number of peripheral objects and the most detail about these objects, while in tracking they can report the least. In tracking mode, the motion dimensions of an object dominate, while in visual search mode shape dimensions are emphasized. In exploration mode, overall commonalities in all dimensions are extracted

The output on the second level is recognized objects. Recognition implies a comparison between the detected object components and a known object or an object class. In search mode recognition is often achieved in a single saccade, while in exploration multiple looks may be required to determine how components are related in a novel object.

At the third level, people assess the situation in a picture. They employ visual routines or strategies that involve switching between modes and scales, in order to understand what is going on within the scene. Different visual routines are learned for different contexts. In a familiar context, a person tends to switch seamlessly from step to step in a visual routine. He or she may find certain features before others to bootstrap finding the more difficult ones to recognize, or move back and forth between scales in a particular order. Individual operations in a processing sequence include activities such as finding landmarks, zooming in and out, following a road or a river, and comparing areas that may have the same classification.

Switching from the rows to the columns in Figure 1, the middle column of Figure 1 labeled Different Levels shows the main operations at each level of processing. So, segmenting followed by grouping leads to component or object detection on the lowest level. The right hand column shows the context-sensitive control or procedural knowledge a person uses at the three levels of processing. The left hand column shows the context-sensitive patterns and ontological information used at the different levels.

Figure 1 illustrates the feedback relations between the different levels of processing. It suggests that search for an expected object can be limited to just the objects found by segmentation. It suggests that a relationship among recognized objects can imply a situation. The situation or context can be associated with routines that direct the focus of attention and determine which characteristics should be used to do the next segmentation. It also explains why people consider fewer image characteristics while in search mode for an expected object, than they use when exploring for anything unusual that might serve as the starting point for a new object..

People can overcome an initial segmentation by consciously following a routine that focuses on different image characteristics or a small portion of the image. This can be illustrated by the hidden object pictures found in many kid's magazines. These puzzles are difficult because one must overcome an initial segmentation of the scene that provides a consistent interpretation of the objects in the scene. To find the hidden objects, one must recognize that they are in a hidden picture context, and consciously focus on local portions of the image and re-segment that area, while ignoring the objects found initially.

3.2. Anatomic and Neurophysiological Evidence about the Mechanisms Used to Do Situation Estimation

There is neurological as well as behavioral evidence for context-sensitive object and situation recognition. Since Hubel and Weisel's (1977) original research on neural response in the visual cortex until recently, it was believed that, in the early areas of the visual pathway (see Figure 2), neurons simply respond to a small range of values in receptive field (RF) characteristics. In other words, the belief was that an individual neuron would only respond to a small range of values in spatial frequency, orientation, color, disparity, retinal position, line length, etc. In general, RF size increases for neurons in higher areas and is sensitive to more complex stimuli. However, it has recently been found that the RF properties of individual neurons change as the systemic response to a visual input evolves. In fact, an early indication of the involvement of recurrent connections is a change in the tuning of a neuron over the course of its response (Lamme & Roelfsema 2000). In the primary visual cortex (V1), tuning for orientation and color dynamically changes during the neuronal response. In other words, even in the early stages of visual processing, a mammal's neural response is affected by what the animal is doing. However, it is noteworthy that the anesthetized (i.e., unresponsive) animals examined by Hubel and Wiesel did not show these types of changes.

Figure 2 shows that there are many retinotopically-organized areas in the brain (i.e., organized so that neurons that are close together in the brain tend to respond to visual stimuli that are close together in the visual field). Each area specializes in different pairings of RF characteristics, such as orientation and spatial frequency, color and disparity, and so on. Segmentation can be based on almost any local RF characteristic (or combination of characteristics), including disparity. Normally, objects are recognized within about 200 msec., and one is not conscious of the separate processes leading to detection and recognition. However, in the cases of images in unfamiliar environments, such as random dot stereograms, the process is stretched out, and one becomes aware of the individual processes. For example, when looking at a random dot stereogram where all structural information is based on disparity, recognition can take more than 10 seconds. The brain is first aware of segmented components at different distances. Additional concentration may be needed to see all the components at once and to link them into coherent objects. Even after an object is seen, some additional seconds may pass before it is recognized.

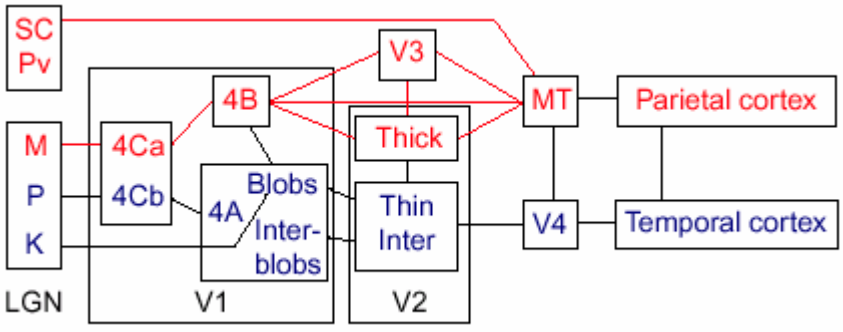

Figure 2: Visual Pathways (Lamme, Roelfsema, 2000) in active animals

In explaining situation estimation from imagery, our framework models a variety of processes with different substrates in the brain. It will have to model:

1. Processes starting with segmentation and going through object recognition to situation understanding. It will be based on operations in the visual pathways going from retina to inferotemporal cortex.
2. Formation and use of models in object recognition and situation estimation, and their organization into local contexts. In fact, we will argue that there is a combination of map knowledge, relevant object models, relevant stories or plot units, and motive potential of important situations that can occur in a local context that are learned together. This will involve hippocampus and the circuit of Papes, as well as frontal cortex.
3. Sensory collection. The process control knowledge needed to do this would include various types of attentional focus, as well as visual routines.
4. Emotional appraisal based on past, present and future situation estimations. These appraisals would impact future situation estimation and the learning of local context. This would involve hippocampus and various cortical regions
5. Generation of expectations about future situations based on stories or plot units relevant in the current context or rational processing. This would involve regions involved in verbal processing.
6. Behavior selection designed to bring the person to desirable future situations or to avoid undesirable situations. Again, various cortical regions would be involved.

The capability to do situation estimation is critically important in doing image analysis, automated analysis of intelligence reports, level 2 and 3 data fusion, and monitoring the status of autonomous systems over long periods of time. The number of request for proposals in these areas over the last few years by DOD, NGA, NASA, and others speaks to the importance of these areas, and the lack of general success to this point.

Brain damage to particular regions can also interfere with one or more of the individual processes leading to recognition. For example, temporal cortex damage can prevent segmented components from being assembled into a complete object (Luria, 1973). A person with this problem, when shown a picture of a bike, will recognize wheels, but not the complete bike.

While Figure 1 provides a framework for recognizing the situation in a scene, it does not specify algorithms for the various steps at different levels of processing, nor does it specify how to use the control or pattern knowledge that a person must bring to the task. The following sections look at some existing image-processing techniques that can be used to implement various operations in Figure 1.

3.3. Techniques from ATR and Image Understanding Relevant for Situation Estimation

Since automatic target recognition (ATR) has strong analogies to intelligence analysis, we quickly review some of the approaches used for ATR. In ATR, the segmentation process is short circuited and an algorithm probes a scene with a small set of templates or models, finding only objects similar in size and shape to the model. If an object sufficiently similar to the model is found it is recognized at the same time it is detected. Generally, some filtering is done so expensive matching isn't done where there is no possibility of a match. ATR in some ways is an extreme version of search mode where the system is totally uninterested in any object that does not match one of its templates.

A good review of the evolution of ATR algorithms is found in Ratches, Walters, Buser, and Guenther (1997). They characterize the algorithms of the early 1980s as heuristic. Typically, detection was based on some sort of threshold, determined by the contrast of an object compared to the local background in an arbitrary box drawn around the object (Walters, 1986). The second action in the process was a series of steps globally referred to as segmentation. The first step in the segmentation process was typically running one of the standard edge-finding operators on the region that had been detected (Poggio, 1986). The next step logically connected the edge segments and filled in the gaps to form a continuous line around the presumptive target. Finally, the region was converted to a binary image by assigning a high-bit value to all pixels inside of the line. Features were then calculated on the segmented area. The calculated values formed a vector in feature space that could be subsequently used for object sorting. Classification was usually the highest level of discrimination and was based upon some sort of statistical classifier, such as Bayesian (Fukunaga, 1990), k-nearest neighbor, or Parzen. Performance of these early ATR systems was found to be marginal in government testing. Detection in low clutter did not exceed 70 percent, and recognition was little better than that obtained by random guessing. False-alarm rates in all but the most benign clutter were unacceptable.

In the late 1980s, a new generation of algorithms was developed that used knowledge-based systems or template-matching approaches (Fu, 1980). The operation of this class of algorithms can be divided into two stages: a region-of-interest (ROI) generation stage and a target identification stage. The task of the ROI stage is to locate all target-sized objects above some minimum contrast in the image. This can be accomplished by convolving a target-sized filter (Burt, 1988) with the image. Typically, the ROIs produced by this stage are then subjected to a template matcher in which the contents of the inner window are compared to stored templates of the target set, after adjustment for pose and scale. The best match, usually in a mean-squared-error sense, is then identified as an object in the ROI. Each match between an ROI and a template results in a score that can be subjected to a thresholding procedure for false-alarm reduction.

Recent experiments indicate that multisensor integration may provide the level of ATR required in a variety of applications (Filippidis, Jain, & Martin, 2000). The utilization of independent parameters from the scene appears to give performance improvements that make ATR acceptable for some applications. Another recent improvement is the use of model-based algorithms. Verly et al. (1992) suggest a model-based approach that fuses range and

intensity information. Model-based algorithms contain libraries of models of the targets for scenarios of interest. Target models coupled with environmental effect models presumably can represent any state in which the target can occur. The images that the sensor produces are then compared to a library of known models until a match occurs with some level of confidence.

Most ATR algorithms have been developed to work on individual images. However, there is also a history of applying image-processing techniques to video image sequences in order to track moving objects (Farmer, Rein-Lien, Jain, 2002). Other techniques, such as using radar in "moving target indicator mode," serves to process motion directly. The primary use for images taken more than a few minutes apart has been to detect changes in land use (Benediktsson & Swain, 1989).

We note that, unlike people, ATR algorithms are not context sensitive nor do they make use of synergistic interactions between segmentation, knowledge-base search, and component linking. As a result, current ATR algorithms do not work well outside the range of backgrounds they were trained on, and can recognize a fairly small set of distinct targets (10-50)

3.4. Identifying the Context

We believe that ability to identify the proper context at any point in time is a core problem of common sense reasoning. In general, this is a very hard problem. Fortunately for people, the appropriate context changes relatively infrequently. Even when the appropriate local context changes, in most cases, people have intentionally moved into the new context, rather than being dropped into it. So people generally know what context they are in. However, as we noted above, when people are dropped into a new context they have all sorts of problems recognizing anything.

4. Applying the Framework to the Analytic Process

By applying the framework described in the previous section to the analytic process, we created an approach that deals with all of majors problems in that domain. We believe that a slightly revised version of the framework in Figure 1 can be applied to the analytic process.

4.1. Information That Intelligence Analysts (IAs) Add to Their Output Reports

People have a wealth of ontological information about what things are part of other things (an arm is part of a body), to what collection does a thing belong (Felix is a cat), and what collections of things are part of larger collections (all whales are mammals). When an IA reports illegal entry into a country after reading about sneaking across a border, he is adding ontological information to the new product.

Table 2: Types of Patterns Used by IAs

Pattern Properties	Recipes for discovering evidence generated by plan execution	Cases	Tells (a small part of a case that implies info about the full case)	Interesting evidence graph seeds and linking criteria
Mental process leading to pattern generation	Planning constraints generalization and chunking	Application of analysis, rules for adaptation	Simple inference, correlation	Generalization, analogies
Assumptions	Consistency in evidence is generated thru same planning process	Similar objectives leads to cases with similar time courses & links	Bad guy doesn't know that anyone can see his green nose	Rare small events are often part of a bigger threat
Representation	Rules, contingencies, constraints	Evidence graphs	Rules (Look for a man traveling with a gray cat)	Seed events plus linking rules for graph growing
Duration of relevance	Long (Months - years)	Medium to long (Weeks -years)	Short (Hours – weeks)	Very long (Years -)
Context sensitivity	MPTs include info about valid range of context	May include info about context	Limited to one person, group, place, or time	Little sensitivity

Patterns of various types also provide ways add information about the relations between events, people, weapons, vehicles, etc. to the basic information in the contained in a set of reports. Table 2 shows four important types of patterns used by IAs. "Recipes for discovering evidence generated by plan execution or causal sequences" is discussed in more detail in the next section. Cases consist of the full set of linked evidence associated with a planned activity. The most basic pattern consists of one or more seed events plus linking criteria for growing an evidence graph outward from the seed. The last type of pattern, which we call tells (For gamblers, a tell is a gesture or action that is unconsciously performed whenever someone is bluffing.) are a small piece of a case that by itself shows that a particular type of terrorist activity is underway.

4.2. Cycles of Situation Update and Revision

The framework for a cognitive architecture shown in Figure 3 is an attempt to summarize the steps taken by humans during each cycle of situation update. Cognitive task analysis (CTA) of intelligence analysts has shown that this hierarchical matching approach is the strategy followed by humans faced with analytic problems (Eilbert, et al., 2004). The central column in Figure 4 shows the processing that goes from segmentation to update or revision of situation understanding, i.e., the first three steps in Figure 3. The column on the left side of Figure 3 shows the feedback from the current estimate of the situation to the parameters to be used in the next cycle of segmentation and object recognition. The column on the right hand side of the figure shows feedback from the situation to the process control decisions relating to local context selection and evidence collection (i.e., what to search for and where to search).

The framework implies that object recognition can not be the ultimate purpose of biological vision. For example, recognizing a tiger has little survival value, while understanding that you are out in the open with an animal you cannot fight or outrun may critical knowledge. It is also important to know the difference between seeing a tiger when you are in the open vs. seeing a tiger that is safely behind bars.

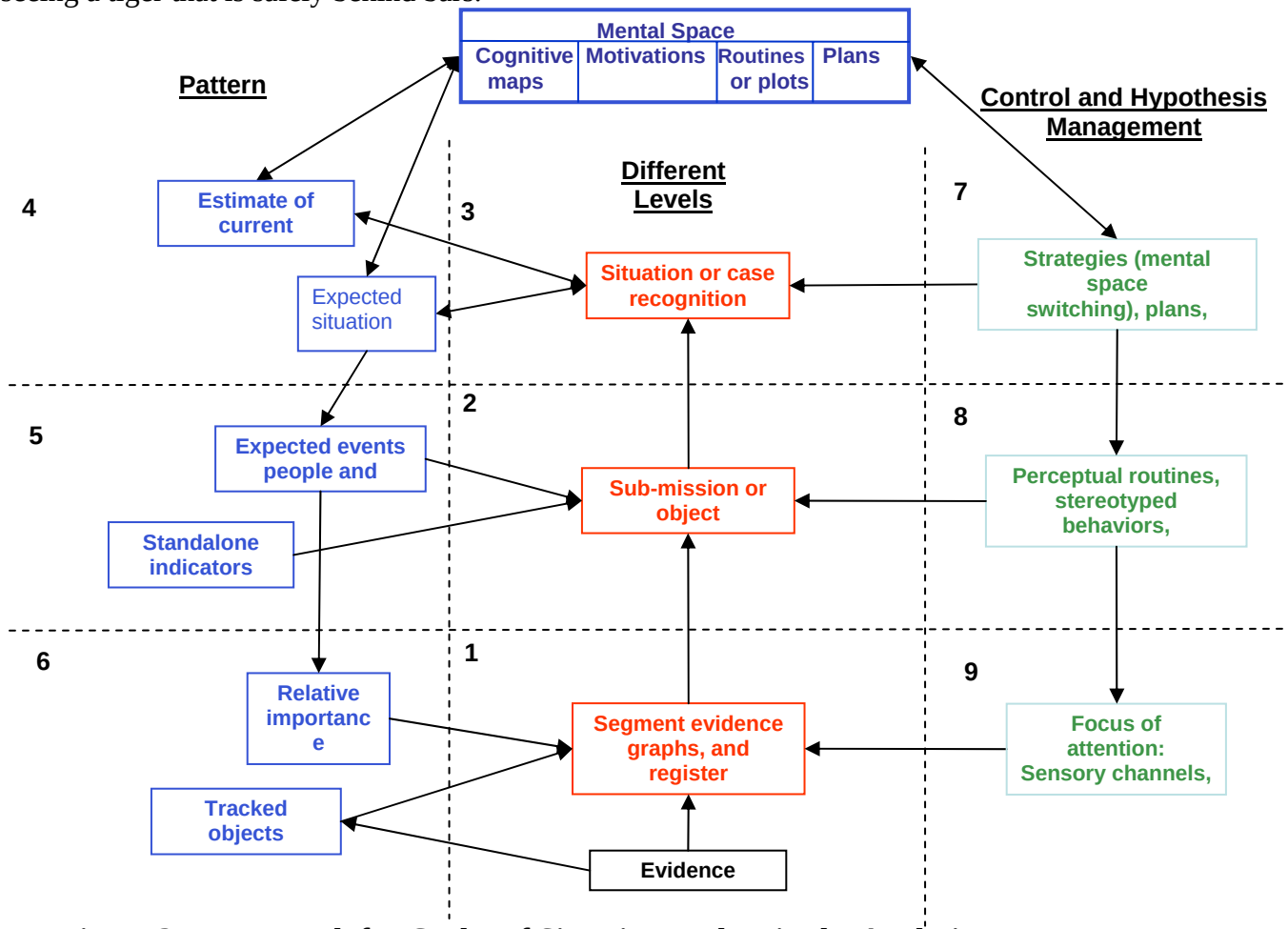

Figure 3: Framework for Cycles of Situation Update in the Analytic Process

There is a variety of neurophysiological evidence for cycles of the type shown in Figure 3. First, there is evidence that signals carrying context or control knowledge impact even the early stages of sensory processing. Lamme & Roelfse (2000) have observed the response of neurons in primary visual cortex change over the course of a reaction to a stimulus.

Normally, the cycles of situation update in response to visual stimulus occur without our being aware of them. Objects are generally recognized within about 150 ms (Thorpe, Fize, Marlot, 1996), and the estimate of the situation is updated in about 500 ms. In intelligence analysis, the cycles last much longer. However, the perceptual process can also be stretched out as in the case of random dot stereograms. For example, when looking at a random dot stereogram where all structural information is based on binocular disparity (i.e., depth information), recognition can take more than 10 seconds. The observer is first aware of segmented components at different distances. Additional concentration may be needed to see all the components at once and to link them into coherent objects. Even after an object is seen, some additional seconds may pass before it is recognized.

4.3. Steps in the Situation Update Cycle

In addition to psychological and neuroanatomical evidence about the steps involved in each situation update cycle, there are important computational reasons for this particular sequence of steps. Performing processing in a bottom-up fashion, starting with rare indicators rather than working top-down from situations, greatly reduces the amount of object matching that needs to be done. Thus, the cycle of situation update requires comparing the relatively small number of segmented regions (Step 1 in Figure 3) to a relatively small number of low-level models relevant in the current Mental Space (Step 2 in Figure 3). This greatly reduces the complexity of the partial matching problem. Simple components can then be boot strapped to find more complex objects. Semantic labels for objects and actions can then be compared to plots or plans, in order to build up to global descriptions of the situation. Thus, we claim that it is the feedback between our understanding of the situation and its support from sensory evidence that allows us to make use of our a priori knowledge and perceive the world as consistent.

The complete situation update cycle is shown in Figure 4. We next describe a single cycle beginning with arrival of new imagery (shown at the bottom of the middle column in Figure 3). The steps are:

1. Find rare indicators in the evidence that are important in their own right. In imagery, a hot spot in the intensity or a color boundary can serve as indicators. Pylyshyn (2000) has argued that human vision makes a direct, pre-conceptual connection between objects in the visual world and their representations in the visual system. In others words, people can track and locate a small set of landmarks or "deictic pointers" across eye movements. This would allow them to select a subset of seed points in the tracked objects, and use a context-sensitive segmentation technique (e.g., Udupa, Saha & Lofuto, 2001) to decide which pixels should be associated with each of the objects or with the background. The regions that emerge from segmentation depend on the seeds and the relative strengths associated with different kinds of links.
2. Recognition of known objects or sub-missions is done in two stages. Registration of landmarks or tracked objects from the previous situation to corresponding landmarks or objects in the newest evidence. Matching models of additional targets or objects of interest against regions found during segmentation.
3. Find groups of objects that should be linked together, and search for stories or plans in the current local context (or Mental Space) that are consistent with the evidence, and what was known from the previous situation estimate. The result of this match is an updated or revised situation estimate.
4. Based on the updated situation estimate generate an expected situation using the story from the current Mental Space with the best match with the evidence.
5. Based on the current Mental Space and the expected situation, generate expectations about the objects that should be observed after the next increment of evidence.
6. Based on the current Mental Space and the expected situation select parameter settings for segmentation on the next cycle
7. (and 8 and 9) Select visual routines that can be used to distinguish between competing explanations of the current evidence. Note that one visual routine may direct attention over many situation update cycles.

5. Algorithm Development

In the framework presented in Figure 3, there are there are three different kinds of patterns used on each of the three levels of processing. The lowest level of pattern utilized by analysts is indicators that can be matched to rare pieces of evidence generally found in one or two reports. Indicators serve as starting points for building a case. At the middle level, patterns correspond to sub-missions that tend to be carried out in routine ways. Using these sub-mission or routine patterns, an analyst can reason about what activities should have preceded an evidentiary event if that event is part of an instance of a particular global pattern, or what events are likely to follow. Thus, finding a partial match with a sub-mission pattern allows an analyst to develop a story that simultaneously explains the existing evidence and predicts future developments. Those portions of a the pattern with no supporting evidence can act as a guide for what evidence needs to be found, or targeted for collection, to increase the likelihood that the pattern is being executed. At the highest level, patterns correspond to complete missions based long-term military plans. All such plans must have considerable flexibility and/or built in contingencies, particularly at points between sub-missions connect to form a complete mission.

We have developed techniques for utilizing each of types of patterns. Section 5.1 describes the use of a context-sensitive segmentation technique that uses indicators to find people belonging to the same organization or events belonging to the same group. Section 5.2 presents a graph matching technique that uses sub-mission or object patterns to discover routine activity. Section 5.3 describes how graph matching can be applied to a hierarchy of sub-mission graphs to find more complex patterns.

Due to the fact that all of EAGLE challenge problems dealt with statistical variation of a few patterns that evolved without interactions with the world, we believe that all the cases test can be consider to be routine behavior. As a result, neither we nor any of the other link discovery contractors in the EAGLE program were forced we deal with patterns corresponding to activity directed by long-term military plans.

5.1. Context-Sensitive Segmentation of the Evidence in Intelligence Reports

A variety of context-sensitive segmentations techniques have recently by computer vision researchers. While work in computer vision performed segmentation earlier than object recognition, it was found that no single segmentation algorithm worked very well over a range of natural environments. While ATR research moved to template matching and model-based techniques, researchers in medical imaging, which deals with far more amorphous shapes than ATR, continued to focus on segmentation techniques. These researchers have recently developed a range of context-sensitive segmentation techniques that adjust their parameters based on the types of objects sought, the background, and the visual modality.

Researchers in a variety of areas of image processing have found that context information must be used to narrow the region where segmentation is being done and the image characteristics on which the segmentation is based. Algorithms that extract linear features from aerial imagery have often made assumptions about curvature and connectedness when looking for "expected" portions that are hidden in noise or clutter (Fischler and Bolles 1986). Eilbert et.al. (1994) made assumptions about the allowable position of small tabs on machine parts, which were marginally distinguishable from the noise, in order to detect important cues that could be used to classify the parts. In all of these cases, a relatively sophisticated understanding of the problem was needed in order to determine what contextual information could be used to make recognition of a particular class of structures easier.

One recent algorithm, grouping with bias (Yu and Shi 2001), uses a computationally efficient graph-partitioning algorithm that can take limited classes of prior knowledge into account. The three types of context considered by Yu and Shi are focus of attention (i.e. only segmenting over a limited portion of the image), a priori knowledge that some points are part of one group, and information about common characteristics shared by points in the same region.

Researchers doing content-based image access and retrieval (CBIR) for multimedia libraries (Pauwels and Frederix 1999) have begun using context in some of their techniques. In CBIR, the aim is to retrieve images that are similar to a query image. Extensive experimentation over the last few years has shown that matching natural images solely on the basis of global similarities is often too crude an approach to produce satisfactory results. What is required is some form of perceptually relevant segmentation that allows one to identify a (small) number of salient and semantically meaningful image regions, which can then serve as the basis for more discerning region-based matching. Thus, there are two distinct steps in this type of algorithm: locating a salient region, and the actual segmentation within that region. Well-known techniques for finding salient regions include figure–ground separation, blob-detection, and contour completion (Amir and Lindenbaum 1998; Howe 1998).

The most useful segmentation approach for incorporation into intelligence analysis is the approach of Upada, Saha & Lotufo (2002). They have developed a fuzzy membership technique that can be used to group an unclassified spatial image element (SPEL) with one of several objects. To begin the process, the system must have a set of seed SPELs that are assumed to belong to objects in the image. A 'local' fuzzy relation, called affinity, is defined

between the seed SPELs and the ungrouped SPELs in an image. The affinity value assigned to a pair of SPELs is based on how close the SPELs are spatially, and the similarity of their intensity and intensity-based properties. A second fuzzy relation called connectedness is defined on pairs SPELs based only on affinities. The connectedness value assigned to a pair of SPELs (say c and d) is the strength of the strongest of all paths from c to d. The 'strength' of a path is simply the smallest affinity along the path. It can be shown that fuzzy connectedness is a similitude relation and that the fuzzy components defined by this relation are an appropriate choice for characterizing objects in images. It can also be shown that, in spite of its enormous combinatorics, fuzzy component extraction can be done computationally elegantly via dynamic programming (Saha & Udupa, 2001; Zhuge, Upada & Saha, 2002; Upada & Saha, 2003).

In the EAGLE setting, we are using a modified version of the fuzzy segmentation algorithm developed by Udupa and his associates to perform context–sensitive segmentation of people into threat groups or organizations, and events into cases (Eilbert et al., 2004). To compute affinities between people, we had to define adjacency. In this application, people are considered adjacent if they participated in the same transaction. There is a notion of intensity that can be assigned to adjacent individuals, namely the number of shared transactions. In addition to the affinity related to transactions 'intensity,' there are two types of features that we want to use as input into our affinity calculation. One has to do with shared "social links", and the other with particular temporal patterns among transactions. We break the affinity for grouping people into threat groups into three components that are multiplied to get the overall affinity. Each component of the overall affinity is an affinity function on its own. The form of the affinity function is selected based on our knowledge of the problem context.

5.1.1. Intensity Component

The following function is used as a measure the intensity component of affinity.
I(c,d) = number of transactions involving c and d/number of transactions involving c or d.
Note that
I(c,c) = 1, 0 <= I(c,d) <= 1, and I(c,d) = I(d,c).
So, I(c,d) is a legal affinity function, and so is any power of I(c,d). Fractional powers increase the relative importance of intensity similarity, while integer powers decrease its importance.

This component of affinity is fairly discriminating in the simulated EAGLE data, since communications among members of the same group is at least twice as frequent as with outsiders, and up to sixteen times as frequent for teams within a group.

5.1.2. Feature Component – Social Networks

We define the social link component of affinity in the following way:
SN(c,d) = for each of the 4 social links that match in c and d, add .25 to SN(c,d)
We note again that this is a legal affinity
SN(c,c) = 1, 0 <= SN(c,d) <= 1, and I(c,d) = SN(d,c).
We expect that this will be a less important component of the overall affinity than intensity.

5.1.3. Feature Component – Exploitation Patterns

This component is based on subevents within a single attack on a target by a team from a threat group, or an exploitation event. The patterns in an exploitation are defined in the EAGLE Y3 PE Lab Documentation. An exploitation involves four sequential stages: communicate, observe target, acquire resources, and consummate exploitation. There are many variable length waits that appear in the top level of an exploitation as well as in each of the stages. This means that using timing to determine whether two events are part of the same exploitation is problematic unless they occur one or two steps apart in the long sequence of things that happen during an exploitation. Here are some ways to show that individuals are participating in the same exploitation, and thus belong to the same team within a group.

1. Given an exploitation in primary evidence where c and d are both agents, and there is at least one potential sub-events that also involve c and d.
2. Given an exploitation with only c or d, check for at least two potential sub-events that involve c and d.

The exploitation affinity is defined as follows:

VE(c,d) = 1, only when d = c.

If c and d participate in the same exploitation, and c and d participate in at least 1 other subevents or events within 150 ticks of the exploitation

VE(c,d) = .98,

VE(c,d) = .85 with just the exploitation having c and d.

If c or d participate in the same exploitation, and c and d participate at least 2 other subevents while the exploitation is going on

VE(c,d) = .9

VE(c,d) = .65 with just one event

5.1.4. Connectivity Computation

An affinity is computed for all pairs of people in a simulation by taking the product of the affinity components described in Sections 5.1, 5.2, and 5.3. Note that affinity is zero for any two people who are not adjacent. Affinity computations can be done incrementally as new evidence arrives. Only pairs of people who share a transaction in the current increment need to have their affinities updated. Otherwise, previously computed affinities are not affected. However, connectivity between any pair of people can change if any affinities are changed and so, connectivity is recalculated after each time step. The technique for calculating connectedness was described in Section 5 above.

5.1.5. Selecting Seeds

The final piece of information needed in order to run the threat group segmentation algorithm is seeds for each of the threat groups. In imagery, connectedness is computed between the seed for each object and every other pixel. A pixel can only be part of one object, and it is grouped with the object seed to which it has the strongest connectedness. However, people in the EAGLE data can belong to multiple groups, and there is a small chance that people listed as a group member in evidence may not really belong to that group. To get around multiple group membership, unknown groups, and corruption problems; group seeds were selected in three steps

1. Compute the pairwise connectivity of all members within each of the groups listed in evidence. Then, select three members who have the highest sum of connectivity as the seed of the group.

2. For all individuals, including members of the groups listed in evidence, compute the connectivity to each of the seeds for a group. Group connectivity for a person is the average to all of the seed members in a group. People are associated with all groups where the highest average connectivity to the seeds is above a threshold.

3. Make sure that all the members involved each threat exploitation case are contained in one of the groups obtained in Step 2. If Not then:

a. Create a new group containing all the individuals participating in the Vulnerability Exploitation

b. Compute the seed for the new group as in Step 1.

c. Compute the connectivity of all individuals to the new seeds.

d. If connectivity to a seed is greater than threshold, assign the individual to the group.

5.2. Matching Evidence to Patterns Associated with Routine Behaviors

In year 1 of the program, we developed an architecture that just utilized sub-mission or routine activity patterns to discover terrorist missions. In fact, we developed both a case-based (Fu, et al, 2003) and a cognitive model based (Eilbert, et al., 2002) algorithms, and a method for combining the results. (The method for combining results was not actually completed till Year 3 of the program.) The architecture for performing this graph matching is shown in Figure 4.

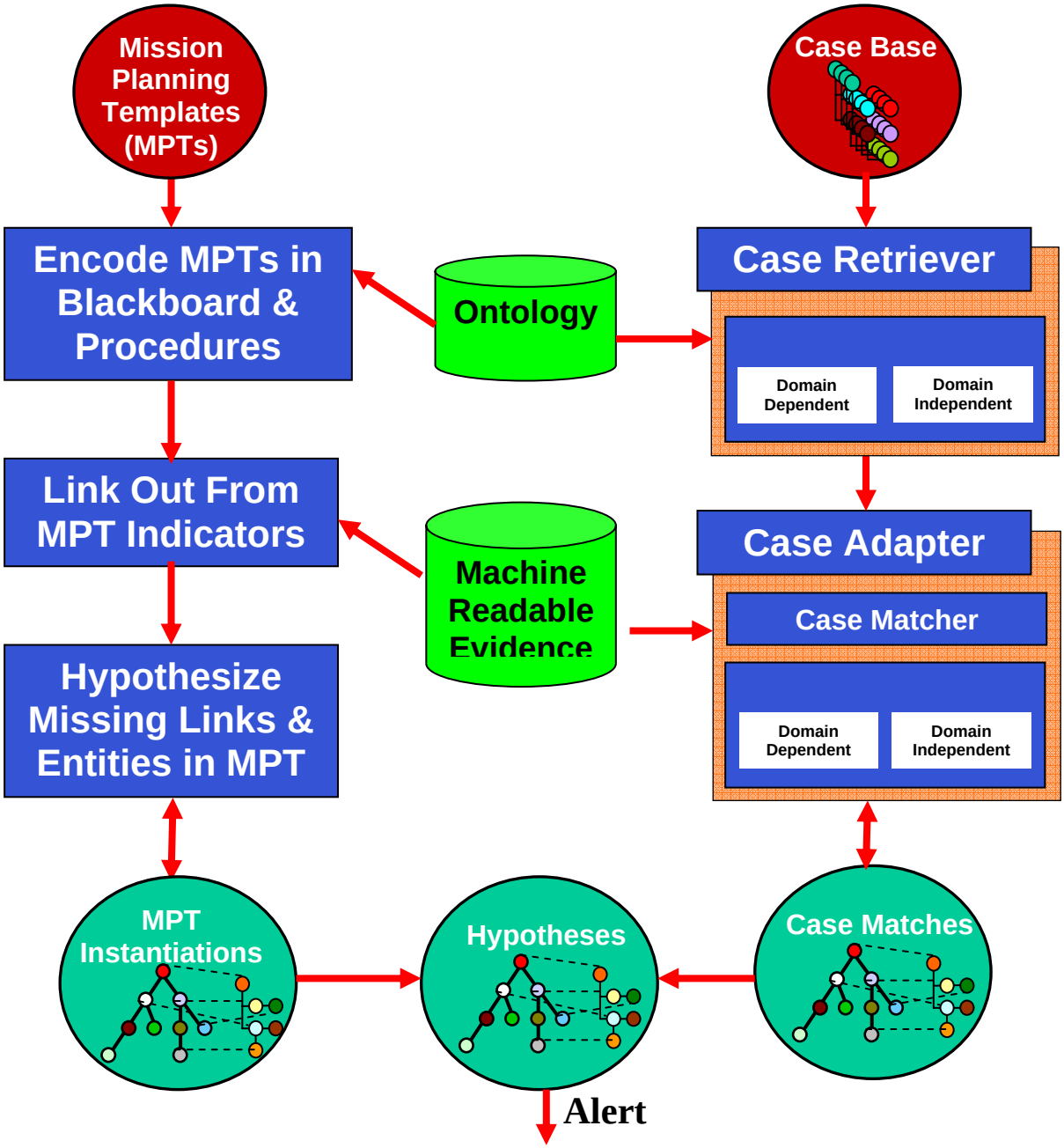

Figure 4: Probing the data with a pattern

With a relatively small number of cases or MPTs (i.e., recipes for finding instances of plan execution), the strategy we evolved in Year 1 was to probe a dataset with each MPT or case and determine if it contained one or more matching instances. If you start with relatively uncommon, highly observable events as a seed or focus with search time proportional to (number of cases times dataset size). There are much fewer MPTs than cases.

An ontology was used to extend the number of events that matched any individual node in the pattern. In the case of the cognitive model, the generalization hierarchy in the ontology was captured using levels within an object-oriented blackboard that was developed as part of this project.

Our system performed quite well on the Year 1 challenge problem. As shown in Figure 5a and 5b, CHI's cognitive modeling component got the best score of all of the teams working on the challenge problem, while SHAI's case-based reasoning module got the second highest scores. In figure 5a, iGEN Module of SCOPE is listed as column 5 and CBR Module of SCOPE is column 6. In figure 5b, The iGEN Module of SCOPE is identified as TIE 5 and CBR Module of SCOPE is CBR.

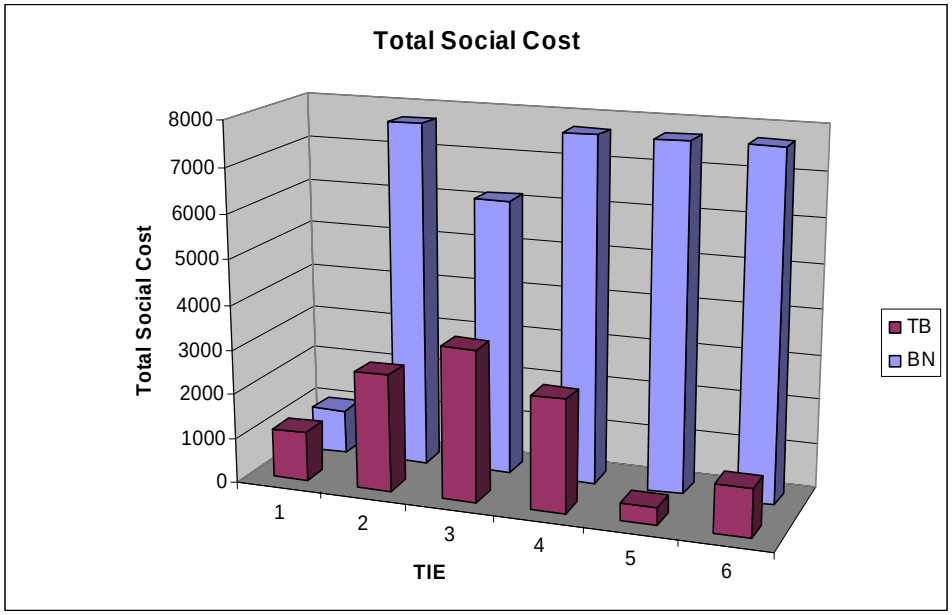

Figure 5a: Comparative total score

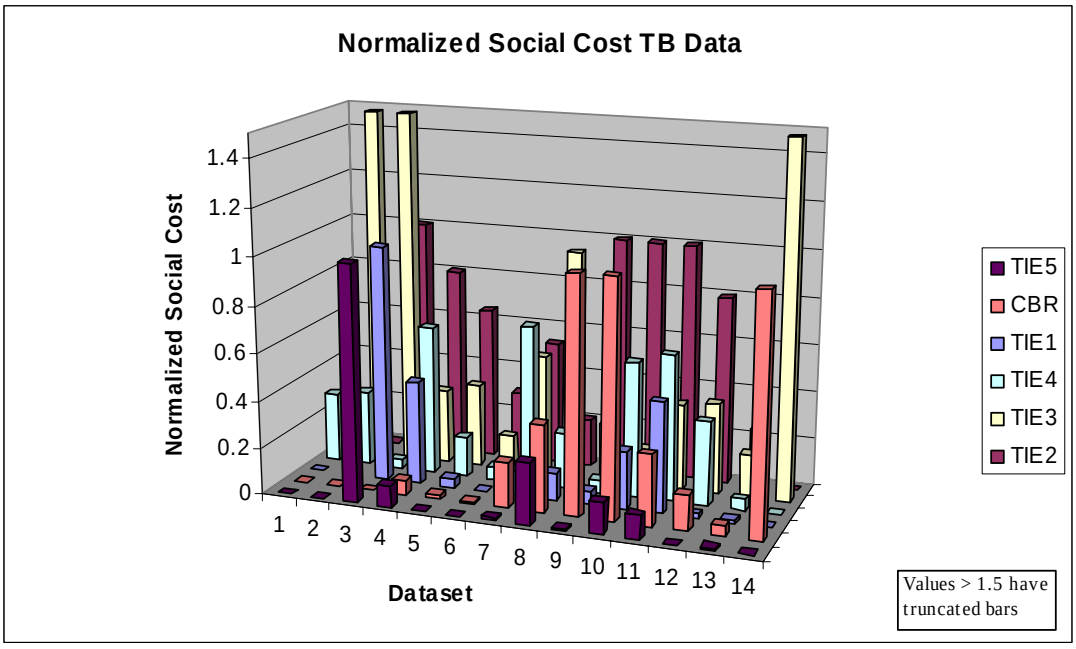

Figure 5b: Comparative scores on individual runs

At the end of Year 1 evaluation, there was a heated discussion about indicators that are in essence giveaways or "tells" for a whole mission, even though they are a small part of the overall mission. We feel that most of the sub-missions that the Sytex analysts have identified in constructing MPTs are routine behaviors. We argue that "tells" occur frequently in routine behaviors and that terrorist use secrecy to make the "tells" more difficult to discover. In most cases, "tells' are only "tells" for small pieces of routine behavior that appear in many contexts and missions. Thus, there are not really "tells" for complete terrorist plans. However, the EAGLE challenge problems had no contingencies in them and always could be characterized as routine behaviors.

5.3. Matching Evidence to Patterns Associated with Flexible Global Patterns

The flexibility of global patterns can make matching them to evidence difficult. The secretive plans used in major cases of terrorist activity extend over long periods of time and have many branches down which execution can proceed. As a result, almost any piece of evidence is a potential match with one or more of the contingent branches of the overall plan. As a result, brute force matches between large datasets of evidence and global patterns representing generalizations of extended global plans are computationally intractable. On the other hand, if an analyst can somehow obtain matches between an associated set of evidence and several of the sub-missions in a global plan, the probability that the evidence is an instance of the global pattern rises very quickly with the proportion of sub-mission patterns matched. This suggests the strategy of building up from rare starting points found using low-level pattern matches, to sub-mission pattern matches, and finally to global pattern matches in computationally tractable steps. In fact, that is reasonable description of how the situational logic technique described by Heuer works.

We note that direct matching of global patterns is not always computationally intractable. When plans are short, simple, and effective (e g., suicide bombers), they are often repeated without much variation. For activities based on these types of plans, a direct match with a complete pattern is potentially more effective than situational logic. One of a relatively small number of complete exemplar cases will be similar to most real cases of suicide bombing. With these short, simple plans each exemplar case can be used as a "probe" into the evidence that sequentially find matches with individual pieces of evidence. This analytic technique of probing with a small number of cases seems to be what Heuer calls case-based reasoning (CBR). Fewer pieces of evidence are needed to discriminate cases based on short, simple plans with a lot of possible branches from other types of activity, than are needed to build up logical support for a hypothesis about a unique situation.

Global patterns associated with flexible plans are not constrained enough to make good probes, and CBR is at a disadvantage compared to situational logic. A case library with a large number of cases would be needed to just cover all of the possible executions of a single plan. The number of cases needed for a large number of plans would be truly vast.

We believe that situational logic and case-based reasoning are quite complimentary, since even complex plans are made up of sub-missions that are short, simple, and effective. From a biological point of view, this is not surprising, since these types of behaviors are strongly reinforced and learned quickly. Lower animals exhibit a range of fixed-action patterns, while people have sets of stereotypes routines that are repeated frequently (although they may be restricted to a particular context).

5.4. Hypothesis Management

Threat exploitation reports consist of three levels of info: the claim that an exploitation has occurred, the attributes of the exploitation, and the underlying evidence (events) that supports the occurrence of the exploitation with those attributes. Just the threat and its attributes are reported to scoring.

When we have several sources for reports about threat exploitations we want to combine the reports into an approximation of ground truth that is as accurate as possible. The first decision that needs to be made is when do reports from different sources match. Then, individual parameters must be compared. For each parameter, sources can agree on the value, disagree on the value, or only one (or a subset) of the sources may have a value for that parameter. Similarly, there can be agreement, disagreement or missing information about the events supporting the threat exploitation and its parameter values.

For each hypothesis (or set of hypotheses about the same exploitation), we need to decide whether to report the exploitation to scoring or not. This decision depends on the quality of the evidence supporting it and the amount of contradiction among the sources. The amount of evidence needed for reporting should depend on both our estimate of ground truth, and the observability and amount of corruption (and/or aliasing) we believe is present in the data set. Observability and corruption estimations can use raw data plus historical information about observability and corruption in similar circumstances. Our ground truth estimate depends on the hypotheses coming from the various sources, and our estimates of the quality of those sources.

Decisions about what to report can be made given the estimates of ground truth, observability and corruption.

5.4.1. "Optimal" Estimates of Ground Truth

There are really two very different ground truth decisions that an analyst makes. The first decision is whether an exploitation is vulnerability or productivity. The second is given an exploitation that 2 or more experts identified (They just have to match, not necessarily have the same classification) which events are part of the case.

In considering whether to call an exploitation vulnerability or productivity, the input information is the a priori knowledge about the likelihood of these classifications, several expert decisions about the classification, estimates of the quality of expert decisions, and the supporting evidence supplied by each expert. The first question is whether level of support should count. We assume that the experts have already considered support in making their classification decision. On the other hand, we would expect classifications made with more supporting evidence are more likely to be correct. What this suggests is indexing expert quality on both level of support and whether their decision was retrospective or prospective.

For each exploitation, there is a decision about whether to call it a threat or non-threat. Suppose there are 2 experts.
p_1 = probability that expert 1 say an exploitation is a threat when it really is a threat
p_2 = probability that expert 2 say an exploitation is a threat when it really is a threat
q_1 = probability that expert 1 say an exploitation is a non-threat when it really is a non-threat

q_2 = probability that expert 2 say an exploitation is a non-threat when it really is a non-threat

Let D_{ijkl} be the decision made by expert j on the ith exploitation with the kth level of support and the lth mode (retrospective or prospective). By checking against ground truth we can determine if indexing over k and l make any difference.
Let

$\alpha_i = \prod_{j:D_{IJ}=1} p_j^0 \prod_{j:D_{IJ}=0} \left(1 - p_j^0\right)$ where the first product is over all experts who decided that exploitation i was a threat (D_{ij}=1),

$\beta_i = \prod_{j:D_{IJ}=0} q_j^0 \prod_{j:D_{IJ}=1} \left(1 - q_j^0\right)$

Using these we can compute the probability of the ground truth at exploitation i being equal to one.

1) $W_i = \dfrac{g(T_i = 1)\alpha_i}{g(T_i = 1)\alpha_i + (1 - g(T_i = 1))\beta_i}$ where $g(T_i)$ = the a priori probability of T_i

Since the values for p and q are just estimates, they can be improved if we can put better information into the calculation of the estimate. Using expectation maximization we find the following estimates of p and q:

2) $p_j = \dfrac{\sum_{i:D_{ij}=1} W_i}{\sum_{i:D_{ij}=1} W_i + \sum_{i:D_{ij}=0} W_i}$

3) $q_j = \dfrac{\sum_{i:D_{ij}=0} (1 - W_i)}{\sum_{i:D_{ij}=1} (1 - W_i) + \sum_{i:D_{ij}=0} (1 - W_i)}$

Using the results of plugging the estimate of ground truth W_i into equations 2 and 3 gives a better estimate of p and q. These estimates for p and q can then be plugged back into equation 1 to the estimate of ground truth. So, we can iterate back and forth for some fixed number of iteration or until the change in ground truth over a cycle is small enough.

Turning to the question of what values to plug into the slots in an exploitation, we can take a similar approach to the estimation of ground truth relative to threat or productivity. If multiple exports say to put different values in the same slot, and a majority agree on one value, we can take that as a candidate value. We will then use the expectation maximization technique to decide whether or not to fill the slot. (Experts who wanted to fill the slot with a different value will be dropped from the calculation of the slot value.)

5.4.2. Determining When Two Hypotheses Describe the Same Event

One of the lessons learned during the Y2.5 evaluation is even if a threat hypothesis is given away in Primary evidence; it may have a different name when it appears in the list of hypotheses coming from a particular source. The method we used in Y2.5 for determining

whether 2 hypotheses correspond is still reasonable, however all hypotheses found by a source need to be checked against all "give away" hypotheses.

The basic algorithm for checking whether 2 hypotheses are the same is as follows:

1- Check if the hypotheses have the same UID.
2- If not, do they have approximately the same starting date, plus a subset with 3 of the same directing agent, target, mode, min or max target application date, or two or more of the same actors
3- Otherwise, they are different

5.4.3. Using Estimates in Final Reports

Given the current best estimate of ground truth an additional decision has to be made about whether a hypothesis is good enough to report the hypothesis to scoring. This should depend in general on the completeness of the information in the slots of the hypothesis, and on the confidence we have in the individual values in the slots. In Y1, we tuned our reporting algorithm pretty finely in order to optimize our score. However, we only had one source of information in that case.

In issuing alerts we need essentially the same type of information needed in getting an optimal estimate of ground truth. In addition, we need to decide when to issue the alerts.

5.4.4. Estimating Observability and Corruption

Knowing observability and corruption is important in making a decision whether to report or alert. At 50% observability, one must report with much less complete patterns than if there were 100% observability. BAE is estimating these from the data. (See dataset-mix spreadsheet).

For example, how often use see the mode of a group given that you see the group varies in a consistent way between difficulty settings. This means that the fraction of the time that the mode shows up in PL is an estimate of which difficulty setting is being used.

Perfect		**Easy**		**Fair**		**Hard**	
Threat	*Non-threat*	*Threat*	*Non-threat*	*Threat*	*Non-threat*	*Threat*	*Non-threat*
0.75	1	0.75	1	0.5	0.875	0.3333	0.75

So, for threat groups, if the percentage of times a mode shows up for the group is 1/3, then it is probably a Hard setting. Other quantities differ in other ways. Analogous difference exist for corruption at different levels of difficulty.

Due to the way the simulator works, once we know that its setting is on Easy, then we know how observability is set for the whole range of events and entities.

6. Pattern Development and Editing

For many years the intelligence community as a whole has suffered from a glut of information and a dearth of analysis. With the restructuring of the intelligence community, and the redirection of information flow to widen dissemination, the amount of information that must be analyzed is increasing rapidly. Analysts supporting asymmetric warfare must mine not just the usual information sources (i.e., ELINT, SIGINT, MASINT, IMINT, HUMINT, FBI, CIA, NSA, State Department), but also the ever increasing and potentially critical important new sources of information from Immigration and Naturalization Service (INS); Customs; Border Patrol; Port Authorities; law enforcement agencies; random general or specific telephone and/or email intercepts; and airline, train, ship, and passenger travel logs. Large quantities of imagery are currently being collected by satellites and aircraft that could be exploited by the intelligence, military, or geographic information communities if the flood of information could be fully analyzed.

While there may be an overwhelming amount of data to look at, finding enough linked evidence to support an actionable conclusion is an equally difficult problem. Even if the future analyst has a greater depth of information management experience and is better trained to detect patterns and indicators across an ever expanding universe of information reports, the sheer volume of reports that must be interpreted is unmanageable without new and innovative user-friendly tools capable of autonomous or semi-autonomous processing of data.

Developing MPTs is an important step towards developing a semi-automated analysis capability using cycles of situation update. The MPTs capture three major types of a priori knowledge: 1) specific knowledge about the situation and context; 2) 'patterns' at three levels of complexity; and 3) process control knowledge about what sources should be looked at and in what order, as well as how and in what order to match patterns and evidence. We think identifying the current context can be done by the analyst using the SCOPE system. Since context does not change very frequently, this should certainly be feasible in the plan discovery domain.

The lowest level of pattern utilized by analysts can be matched to rare pieces of evidence generally found in one or two reports that can serve as indicators of problems worth pursuing, or the starting point for building a case. At the middle level, patterns correspond to sub-missions that are carried out in a fairly small number of routine ways. With these short, simple plans each exemplar case can be considered a pattern that can be used to "probe" the evidence for matches. This analytic technique of probing with a small number of cases seems to be what Heuer calls case-based reasoning (CBR). At the highest level, analysts have knowledge about global patterns that can be matched against whole cases. Using these global patterns, an analyst can reason about what activities must necessarily precede an evidentiary event, and what events are likely to follow a particular activity. Thus, finding a partial match with a global pattern allows an analyst to develop a complete story that simultaneously explains the existing evidence and predicts future developments.

The flexibility of global patterns can make matching them to evidence difficult. The secretive plans used in major cases of terrorist activity extend over long periods of time and have many branches down which execution can proceed. As a result, almost any piece of

evidence is a potential match with one or more of the contingent branches of the overall plan. As a result, brute force matches between large datasets of evidence and global patterns representing generalizations of extended global plans are computationally intractable. This makes it necessary to build up from rare starting points found using low-level pattern matches, to sub-mission pattern matches, and finally to global pattern matches in computationally tractable steps. In other words, it is necessary to sequentially perform steps 1, 2, and 3 in our framework.

6.1. Models of Plans: MPTs

The basic tenet of the SCOPE approach is that the process of creating TO mission plans can be characterized by a relatively small number of templates, and that these templates provide an invariant pattern that can be detected in evidence prior to terrorist attacks. The purpose of the SCOPE algorithm is thus to discover an existing instance of a TO mission planning template given the information available to an intelligence analyst in a C/AT organization. The SCOPE algorithm needs to model TO mission plans at a level of fidelity that will allow the security/military group within the C/AT organization to intervene in terrorist activity. We adopt one of the primary assumptions underlying the EAGLE project, i.e. the evidence available for collection is often adequate to distinguish between different instances of mission plans (as well as other patterns of activity).

Plans for major terrorist attacks are created using military style planning which means that the ordering of sub-missions used from plan to plan is very similar. We have developed a hierarchy of MPTs to capture much of the information that analysts use in discovering planned, secretive activity. The four levels we propose are as follows:

1. Generic mission planning template (MPT) –
    a. Only sub-missions of a plan would be specified.
    b. A generic template would apply to all terrorist activity, contract killing, security fraud, etc.
2. Specific MPTs –
    a. This would add indicators of various sub-mission, basic timeline information, and routines for searching for data and determining when to issue alerts
    b. A specific MPT would differentiate between bombings, assassinations, chemical, or biological attacks in a terrorist environment
3. Domain specific MPTs
    a. A much smaller set of indicators would be listed for each sub-mission than in the specific MPT level
    b. By adding a details about the type of target, the terrain where it is found, the method of attack etc., it is possible to come up with a much smaller, more focused set of indicators for each sub-missions. Also, information need to resolve competing hypotheses would be added
4. Individual Plan
    a. The analyst would actually create a single contingent plan for attacking a specific target. It should be possible to run one of these plans in a simulator and create appropriate evidence.

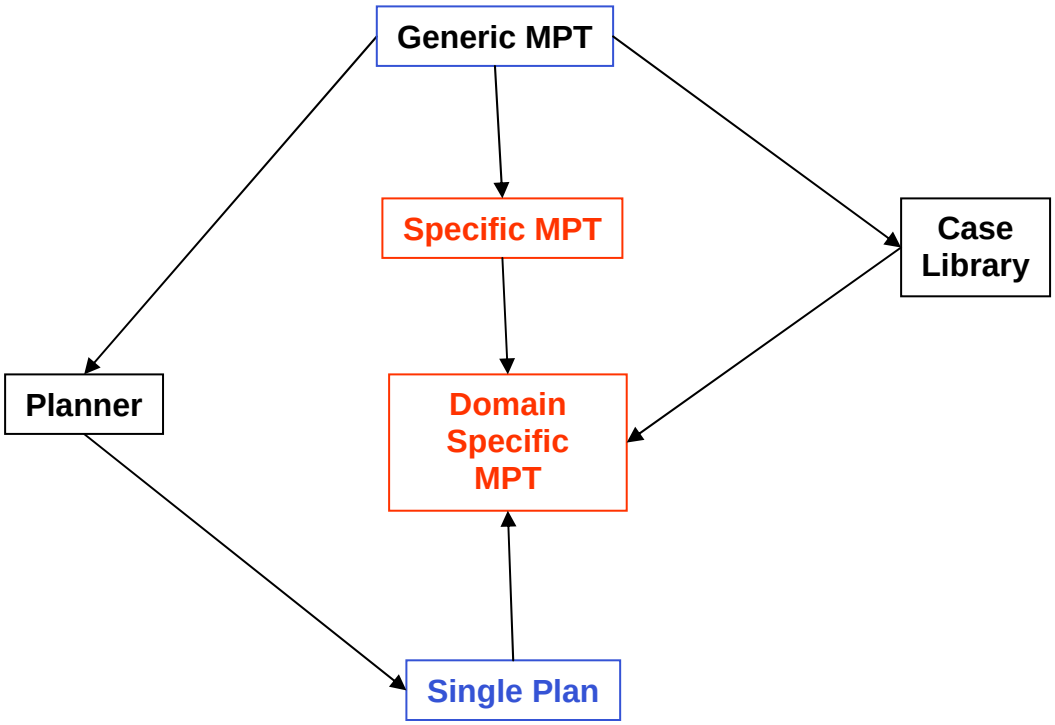

Figure 6: Levels of MPT Detail

In general, the information an analyst would be given to support the construction of MPTs includes:

- A set of generic MPTs
- A case library of executed MPTs
- A planning tool that can be used to create individual plans

A pictorial version of MPT creation is given in Figure 6. The natural level for a analyst to think about MPTs is the Domain Specific level. It is much easier to come up with realistic indicators if you have a type of target, the area around the target, and a delivery system for the attack in mind. In thinking about MPTs, its is important for the analyst to have actual case studies of missions and the reports seen during the mission. If cases are not available, the analyst can create his own plan and either simulate or think through its execution. The Specific MPT is then constructed using a union of the indicators found in the Domain Specific MPTs.

As part of our SCOPE effort, we have attempted to acquire knowledge about IAs recipes for discovering evidence generated by terrorist plan execution. An example of this type of pattern, which we call a Mission Planning Template (MPT) discovery process, is shown in Table 3. The basic assumption that makes an MPT a useful pattern, is that terrorist organizations (TOs) use a mission planning process that varies little between missions. We

believe that while TOs take pains to vary the weapons, delivery systems, and targets of their attacks; the mission planning process they use will be similar across TOs and missions because they have no other effective way to coordinate their long, distributed, secretive plans.

The common mission planning process produces many clearly definable, detectable, characteristics common to TO operations. We have captured in sets of MPTs. A rudimentary version of an MPT is illustrated in Table 3, which describes a terrorist bombing mission. The final versions will by necessity include a far more extensive list of indicators/activities in order to capture potentially pertinent information that may indicate a sub-mission. At the core of the MPT concept is the set of sub-missions shown in the left hand column of Table 1. Our discussions with analysts at our EAGLE subcontractor (Sytex) and INSCOM lead us to believe that some form of these basic sub-missions occur over the course of essentially any terrorist attack. In the right hand column, there are indicators/activity events that would lead an analyst to believe that the corresponding sub-mission was being, or has been executed.

Different TO missions require different MPTs. To date we have created general tables analogous to Table 3 for bio-terrorism, chemical terrorism, assassination, and kidnapping. We believe that MPTs can be organized in a hierarchy, with new ones added when new terrorist missions come to light, or an existing MPT needs to be split into two or more versions to improve performance. When moving to an unrelated domain, e.g. conventional military planning or criminal activity, a completely new hierarchy of MPTs would have to be developed.

When an IA makes use of his internal version of an MPT, various types of information are added to the set of reports being considered. Some events are labeled as indicators, event may be predicted assuming that evidence supporting a terrorist mission has been found, and evidence fitting into the MPT can be linked together into a graph.

In addition to the basic information shown in Table 3, a complete MPT should also specify temporal constraints on the relationships between the indicators and sub-missions in the MPT. The major terrorist attacks carried out by Al Qaeda are planned and executed in a 3-5 year time frame. There is also some evidence that the time intervals between sub-missions may be consistent enough to act as a discriminator between different types of mission, i.e. bombing vs. chemical vs. biological, etc. Other probabilistic information that should be linked to the indicators in an MPT include the quality or value of an indicator or its impact on belief in its sub-mission, and probability that an indicator of an event is available given that the event has actually occurred.

A final component of an MPT is rules for the combination of multiple pieces of evidence associated with a single sub-mission. There are generally several indicator of each sub-mission found when terrorist activity is uncovered. Within the set of indicators for any sub-mission some pairs of indicators are mutually supporting, others are conflicting, while still other are independent of each other. Therefore, a set of evidence combination rules that captures analyst expertise on these matters should be part of the MPT.

Table 3: Problems Associated With the Analysis and Visual Perception

Sub-mission	Indicators/Activities
Mission Initiation and Target Selection Criteria include: symbolic or tactical value, vulnerability, accessibility	1.Generalized overt threat by known or unknown group or personnel to attack American interests (Bin Laden's fatwa proclamation) 2. Increased "chatter" (SIGINT) intercepts discussing a mission 3. HUMINT sources indicate that there are plans to attack US interests 4. Increase in inflammatory rhetoric by known terrorist groups or countries or organizations sympathetic to them 5. Reports of possible surveillance conducted against US facility 6. Protests against the treatment of a group member (i.e., events following capture, trial, or death of a group member
Reconnaissance Intelligence, Surveillance or Target Analysis (RISTA)	1. Individual(s)s reported taking pictures, video, or sketching US embassy or area near the embassy 2. Embassy guards report vehicle driving near embassy on multiple occasions 3. Embassy personnel report they may have been followed (vehicle, foot) 4. Host nation US embassy employees approached by unknown individual asking for embassy building or compound information 5. Embassy employee asked by unknown individual for embassy phone book 6. Reports of people parking, standing, or loitering in the same area over a multiple day period
Establish Cell	1. Increased traffic (personnel, telephonic, radio, email) at known or suspected terrorist support or sympathetic organizations in or near target country/city (i.e. Islamic Relief Agency, Mercy International etc.) 2. Reports of known terrorist or "watch listed" personnel movement 3. Theft of passports, military/government identification cards or government license plates 4. Establish operations enabling commercial enterprise(s) (in the Kenya bombing and the USS Cole, a commercial fishing license was purchased to facilitate movement of Al Qaeda owned ships that were used for transporting personnel, material, and explosives)
Logistics	1. Rental (most often cash payment) of large, secure, isolated residence 2. Reported purchases of satellite telephones and airtime 3. Report of military, or mining company explosives stolen 4. Theft or purchase of trucks or vans (especially those with large weight or cargo capacity 5. Large theft/sales of combinations of ingredients for explosives (fuel oil, nitrates, other fertilizers, etc.) 6. Known or suspected terrorist receives money wired from an observed bank account or institution 7. Theft/purchases of a quantity of potential explosive containers (propane bottles, welding tanks etc)
Planning and Training	1. Imagery or HUMINT reports of refugee or training camps regarding explosions/craters 2. Imagery or HUMINT reports of mock-ups of US Embassies 3.Chemical fires, toxic odors, or brightly colored stains in apartments, hotel rooms, or self-storage units; 4. Radio, telephonic, or email intercepts indication general or detailed information of planned event 5. Source (credibility not determined) informs embassy or law enforcement official of a planned bombing 6. Terrorist organization publicly threatens an attack within a specific or general period of time 7. HUMINT reports of specific training (weapons, bombing, surveillance,

	communications etc) at terrorist camps 8. Terrorist training manuals describing various types of training (weapons, bombing, surveillance, communications, etc.)
Rehearse/Final RISTA	1. Small test explosions in rural or wooded areas 2. Embassy guards report unusual vehicle traffic in the vicinity of embassy (vehicle(s) driving to, then away rapidly, repeatedly drive bys, etc.) 3. One or a series of false alarms requiring law enforcement and or emergency services response 4. Reported surveillance (or suspected surveillance, all types) 5. Hospital reports of individual(s) with burn or explosion caused injuries
Pre-Execution/ Execution	1. Departure of official embassy personnel from terrorist sympathetic countries (i.e., days prior to the embassy bombing in Kenya, the Iranians evacuated most of their embassy personnel) 2. A series of "wrong number" or "hang up" phone calls to the embassy 3. Abrupt departure from rental property (often damaged, which can result in landlord making an official report) 4. Attack
Post Event	1. Escape – Reports of suspicious individuals detained at target country's exit points 2. Individuals with false/modified/stole documents attempt to board departing flights/ships/border control points 3. Maximum publicity, (claim responsibility) faxes, emails, phone calls to news organizations 4. Speeches by known terrorist of sympathizers proclaiming great victory (both for current member morale and to entice new recruits)

6.1.1. Different Planning Templates: Criminal and Terrorist Domains

In this section, we point out some of the significant differences underlying planning in terrorist vs. criminal domains. These difference leads to differences in the type of evidence available and the data collection techniques used by law enforcement vs. counterterrorism. Of course, there are huge differences within a domain, say between a complex Al Qaeda mission to using multiple explosive laden vehicles to destroy a chemical plant vs. a single Hezbollah suicide bomber killing patrons at a restaurant.

One major difference between criminal domains and terrorist domains is that in the criminal domain, there are usually many intermediate criminal incidents and the frequent absence of a discernable major terminal incident. On the other hand, in terrorist domains there are more subtle intermediate incidents (few that are actually criminal) and an intended major terminal event. (Again, we are generally speaking of a terrorist attack which requires more planning that a suicide bomber.) This difference has a major impact on how data is collected, as well as what data can be collected (based on resource and capabilities) in these environments. This also means that counterterrorism analysts are more focused towards prediction, interdiction, prevention, and subsequent eradication, while law enforcement analysts focus more on detection and evidence collection (in judicial terms). Law enforcement agencies accomplish prevention and eradication by removing individuals who have committed criminal acts from the picture, not generally by analyzing predictive indicators that allow them to interdict prior to an incident. There are obviously exceptions to this generalization. An example would be if law enforcement heard via a wiretap the X was going to kill Y on Friday the 13th, in the, library, with a candlestick. But that type of intelligence is the exception, not the rule.

In a real police investigation of Mafiya activity, much of the evidence would come from a data collection process that is planned and replanned. This type of active collection of evidence distinguishes police analysis from counterterrorism analysis (at least before an incident occurs). This is in large part is due to the much more stringent criteria for proof in prosecution than in detection and identification of a potential terrorist attack. In 2003, we will probably not be seeing information from active collection or historical information. However, in the long run, a TIE system that deals with law enforcement issues, such as contract killing, would need to get involved in planning evidence collection either directly or through analyst in-the-loop.

Another major difference between the terrorist domain and ROC is that terrorist attacks are the usually the result of executing a detailed military mission plan. In the criminal domain, there is often not a planned or stated terminal objective. Because the overriding goal of ROC is to make money, they can and do operate on a take what is available plan. If the end result is control of a specific industry, all the better. But if only part of the industry can be taken over/controlled, and ROC still makes money, so be it. While individual murders and programs of intimidation may be planned to some extent, the overall takeover may not be. It can be a result more closely resembling the spread of a fungal disease through a forest. Individual nodes of an industry are taken over as the opportunity presents itself. This lack of a terminal goal, (and the aforementioned differences in the types of information available to be collected and the types of collection itself), against which we search for indicators to build a predictive analytic assessments, fundamentally defines the process and goals of analysis in the two domains. The ROC domain is generally after the fact evidentiary analysis, while analysts in the terrorist domain prefer to be predictive.

6.1.2. Discriminating Patterns vs. Generative Patterns

The definition I would propose for a discriminative pattern is a set of rules about what entities can be, might be, and cannot be nodes; and which predicates can be, might be, and cannot be links in any evidence graph that is an instance of one type of pattern. In addition, the definition of a discriminative pattern can include rules and constraints on the values assumed by the nodes and links in any evidence graph that is an instance of that pattern. A discriminative category does not need to be homogeneous in its connections. In other words, a category may contain two groups of graphs in which each of the group is less similar to any other member of the other group than it is to some graph that is not a member of the category. The pattern should ideally separate every instance of a pattern from all other cases.

Evidence created by a simulator may contain emergent properties that are not specified in the simulator patterns or rules. In general, it is a hard job to determine whether a set of evidence is consistent with the patterns in a simulator.

A real world mission plan or process may also interact with unanticipated events in the world to create patterns not specified in the plan.

The rules in the simulator that describe the bad Mafiya activities that we will be evaluated on this year do not care if there are similar rules describing another activity done by the Mafiya or a different (non-Mafiya) group. They are also indifferent to whether the world-state they are reacting to resulted from the execution of other rules. Thus, we can expect interactions among the rules generating both the bad and not-bad activities that change some of the

characteristics of the simulated activity produced. Thus, knowing the simulator rules that generate bad activity does not mean that you know exactly what bad activity looks like. Hopefully, there will be some portion of an evidence graph associated with a bad activity that will discriminate it from all other activity before an attack takes place. Of course, after the attack takes place there will be a difference between threat and non-threat evidence graphs.

We believe is that the simulator will produce evidence whose analysis is a complex task, albeit a much more bounded one than finding bad activity in the real-world. In the real-world, an analyst has an idea of what is possible and what is normal in a domain that he/she is familiar with. This understanding is based on deep common-sense reasoning that no EAGLE component is trying to model. We do not expect EAGLE pattern learning to be able to reason outside of the evidence available to it as well analysts. Particularly in the terrorist domain, many of the patterns of activity have not yet been seen, and the ability of analyst to anticipate terrorist plans will be needed to even begin to cover the domain.

Although an analyst can capture much of their understanding about particular types of terrorist or criminal activity in a pattern or model, there are likely to be omissions and small mistakes about some characteristics of the pattern. We think that the simulated evidence can be analyzed in the same way as real-world evidence, that analysts can produce patterns that are close to discriminative patterns for the simulated ROC world. These patterns can then be improved by looking at the errors the patterns make in discriminating bad activity in historical/training runs of the simulator. The corrections could be made either by analysts or PL systems.

7. SCOPE Software Architecture

The software architecture that we are using to implement the framework shown in Figure 3 is shown in Figure 7. It shows the three types of databases that we are using. The Hypothesis Database in practice was a WebDAV server through which we exchange hypotheses in either XML or IET answer key format. All of the SCOPE algorithms run within an iGEN cognitive model and the object-oriented blackboard that the model uses. However, several algorithms including connectedness portion of the segmentation algorithm is a separate C++ algorithm that is called with the right parameters by the iGEN model. iGEN currently does the most of the hypothesis management and controls probing the evidence with global patterns.

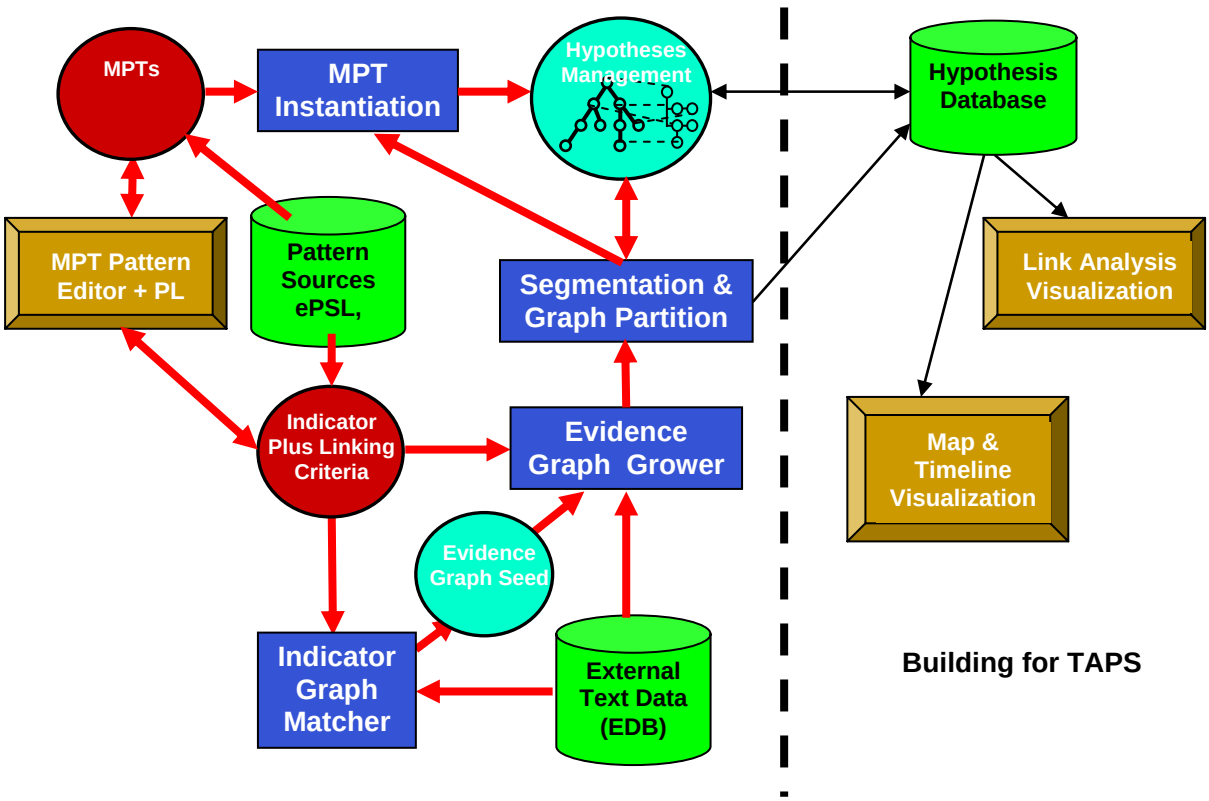

Figure 7: Year 2: SCOPE Architecture

8. Graph Similarity Metrics

The field where the similarity of graphs is currently receiving the most attention is content-based retrieval of images in large image databases. The reason for this is well stated by Lou and Hancock [1], "Graph matching is a task of pivotal importance in high-level vision since it provides a means by which abstract pictorial descriptions can be matched to one another." Thus, for image database access the task of eliciting graph structures from raw image is analogous to finding evidence in raw reports for the EAGLE project.

An important point in graph matching is that in most cases information can either be represented in the graph structure or in a set of attributes describing the node objects. Since the process of eliciting graph structures from raw image data is a task of some fragility due to noise and the limited effectiveness of the available segmentation algorithms, graph matching is invariably approached by inexact means. Because of this, many high-level matching techniques have weakened the role of structural information and have relied more heavily on the use of attribute relations. This is disappointing since structural graph representations provide abstractions that convey important visual invariance.

The next section describes the more structural approaches to matching and the following section talks about structural approaches.

8.1. Literature Review of Structural Similarity in Attributed Relational Graphs (ARG)

The following literature review is from Lou and Hancock [1]. Some of the pioneering work on graph matching was undertaken in the early 1970's by Barrow and Popplestone [2] and by Fischler and Enschlager [3]. These two studies provided proof of concept for the use of relational structures in high-level pictorial object recognition. Over the intervening three decades, there has been a sustained research activity. Broadly speaking, the work reported in the literature can be divided into three areas.

The first of these is concerned with defining a measure of relational similarity. Much of the early work here was undertaken in the structural pattern recognition literature. For instance, Shapiro and Haralick [4] showed how inexact structural representations could be compared by counting consistent subgraphs. This similarity measure was refined by Eshera and Fu [5] and by Sanfeliu and Fu [6] who showed how the concept of string edit distance could be extended to graphical structures. The formal basis of graph edit distance has recently been extended by Bunke and his coworkers [7], [8] who have shown, among other things, that the edit distance is related to the size of the maximum common subgraph. More recently, Tirthapura et al. have shown how the classical Levenshtein distance can be used to match shock graphs representing 2D skeletal shapes [9]. Much of this work adopts a heuristic or goal-directed approach to measuring graph similarity.

In the most practiced approach to content-based image retrieval, the visual appearance of each spatial entity is represented independently by a vector of features with predefined dimensionality. This permits entities to be regarded as independent points within a vectorial space of features, thus enabling efficient indexing based on consolidated point access methods (e.g., KD-trees and R-trees), which partition the space along an ordered hierarchical structure [23]. Mutual relationships between entities can be accounted for in this retrieval process through a cascade filter which evaluates the similarity in the arrangement of entities after these have been retrieved on the basis of their individual features [10], [28], [32], [33], [49]. However, this matching scheme is not able to select entities that attain a better fit in the mutual arrangement at the expense of a lower feature similarity.

To overcome the limit, the query must be matched against archived images so as to maximize a joint similarity balancing, the relevance of individual entity features, and mutual entity relationships. This requires entities and relationships to be represented and compared as parts of a global structure that captures mutual dependencies. In this case, the model of content takes the shape of an Attributed Relational Graph (ARG), with entity vectors and relationship descriptors attached to vertices and edges, respectively.

Unfortunately, representation of image contents as ARGs, radically increases the complexity of matching algorithms, and hinders the feasibility of indexing schemes. For this reason, though partially prospected in a few contributions [40], [50], no comprehensive solution has been yet proposed supporting the application of ARGs to content-based retrieval from image databases of practical interest. In fact, while the distance between two sets of independent vectors can be computed in polynomial time, the distance between two ARGs requires the identification of an optimal error correcting (sub)graph isomorphism [22], which is an NP-

complete problem with exponential time solution algorithms [26]. Besides, the lack of a structural correspondence between the entities of different images hinders the direct representation of graphs through vectors with predefined dimensionality and structure, thus preventing indexing based on point access methods.

The second issue addressed is how to develop more principled statistical measures of similarity that model the processes of structural error present in the graph-matching problem. Wong and You [10] made one of the first contributions here by defining an entropy measure for structural graph matching. Boyer and Kak [11] also adopted an information theoretic approach, but worked instead with attribute relations. Using a probabilistic relaxation framework Christmas et al. [12] have developed a statistical model for pairwise attribute relations. Working in the purely structural domain, Wilson and Hancock [13] derived probability distributions for the relational errors that occur when there is significant graph corruption. More recently, Cross and Hancock [14] have developed a variant of the EM algorithm in which the structural error model of Wilson and Hancock is used to improve the alignment of triangulated point-sets under perspective geometry.

The third issue is optimization. Here, there have been several attempts to use both continuous and discrete optimization methods to locate optimal graph matches. Turning our attention to discrete optimization methods, there have been several attempts to apply techniques such as simulated annealing, genetic search, and tabu search to the graph matching problem. However, continuous optimization methods provide attractive alternatives since their fixed points and convergence properties are usually better understood than their discrete counterparts. However, the main difficulty associated with mapping a discretely defined search problem onto a continuous optimization method is that of embedding. There are several ways in which this embedding can be effected for the problem of graph matching. The most straightforward of these is to pose the graph-matching problem as that of recovering a permutation matrix which preserves edge or adjacency structure. For instance, Kosowsky and Yuille have cast the problem into a statistical physics setting and have recovered a continuous representation of the permutation matrix using mean-field update equations [15]. Gold and Rangarajan [16] have exploited the stochastic properties of Sinkhorn matrices to recover the matches using a soft assign update algorithm. Umeyama [17] takes a more conventional least-squares approach and shows how an eigen decomposition method can be used to recover the permutation matrix. An alternative representation has recently been developed by Pelillo [18] which involves an embedding based on the association graph. Matches are located by using the replicator equations of evolutionary game-theory to locate the maximal clique of the association graph, i.e., the maximum common subgraph, of the two graphs being matched. Subsequently, this method has also been applied to shock-graph matching [19].

Closely related to this work on recovering permutation structure by continuous embedding is the literature on spectral graph theory. This is a term applied to a family of techniques that aim to characterize the global structural properties of graphs using the eigenvalues and eigenvectors of the adjacency matrix. In the computer vision literature, there have been a number of attempts to use spectral properties for graph matching, object recognition, and image segmentation. Umeyama has an eigen decomposition method that matches graphs of the same size [17]. Borrowing ideas from structural chemistry, Scott and Longuet-Higgins were among the first to use spectral methods for correspondence analysis [20]. They showed

how to recover correspondences via singular value decomposition on the point association matrix between different images. In keeping more closely with the spirit of spectral graph theory, yet seemingly unaware of the related literature, Shapiro and Brady [21] developed an extension of the Scott and Longuet-Higgins method, in which point sets are matched by comparing the eigenvectors of the point proximity matrix. Here, the proximity matrix is constructed by computing the Gaussian weighted distance between points. The eigenvectors of the proximity matrices can be viewed as the basis vectors of an orthogonal transformation on the original point identities. In other words, the components of the eigenvectors represent mixing angles for the transformed points. Matching between different point-sets is effected by comparing the pattern of eigenvectors in different images. Shapiro and Brady's method can be viewed as operating in the attribute domain rather than the structural domain. Horaud and Sossa [22] have adopted a purely structural approach to the recognition of line-drawings. Their representation is based on the immanental polynomials for the Laplacian matrix of the line-connectivity graph. By comparing the coefficients of the polynomials, they are able to index into a large database of line-drawings. In another application involving indexing into large databases, Sengupta and Boyer [23] have used property matrix spectra to characterize line-patterns. Various attribute representations are suggested and compared. Shokoufandeh et al. [24] have shown how graphs can be encoded using local topological spectra for shape recognition from large databases.

Although formally elegant, the main limitation of these matrix methods is their inability to cope with graphs of different sizes. This means that they cannot be used when significant levels of structural corruption are present. Also this scheme cannot be applied to graphs labeled with numerical (nonsymbolic) values.

8.2. Similarity of ARGs based on Attributes

Out of the specific context of content-based image retrieval, the problem of comparing an input graph against a large number of model graphs is addressed in [35], [36] using a decomposition approach. At archiving time, model graphs are repeatedly decomposed in subgraphs, which are organized by size in a global hierarchical index. At runtime, matching is accomplished by comparing the input graph against the subgraphs of the index in a bottom-up composition. In so doing, any subgraph appearing within multiple images is checked only once, thus reducing the dependency on the database size. However, this reduction does not provide significant advantages when graphs are labeled with numerical (nonsymbolic) values, a situation very common in content-based image retrieval. In this context, even in the comparison of models of apparently similar images, differences in numerical values affect all entities and relationships.

Following an opposite approach, graphs representing the structure of 3D objects are organized into a hierarchy of clusters, and each cluster is represented by a parametric supermodel that merges the structures of the models that it contains. In this case, the access is performed in a top-down manner, by pruning those clusters whose supermodels do not match the query and by descending the hierarchy to identify a specific model through repetitive refinement of its parametric representation. The approach can provide a significant speedup, but may cause false dismissals and blow up the size of supermodels.

In [40], ARGs that model medical images are reduced to a vectorial representation enabling R-tree indexing, under the assumption that all the graphs contain a set of anchor entities with predefined labels. Nonanchor entities are also allowed, but their number determines a linear degradation in the efficiency of the index. This prevents application when archived images do not share a dominant number of equal anchor entities. In particular, the approach is not applicable in the case in which entities are identified by numeric and densely changing features rather than by symbolic identifiers.

Metric indexing [53], has been proposed as a general solution to organize large databases of objects without reducing them into a vectorial representation. In fact, in this indexing scheme, objects are clustered and retrieved according to their mutual distances, rather than to their absolute position in the reference system of a vectorial space. This supports indexing of objects with high-dimensionality and enables searches based on complex metrics of distance. Apparently, metric indexing could also effectively fit the needs for the organization of an archive of ARGs. However, this solution, which was never practiced, faces a major difficulty due to the need for repeated computation of object distances during the access to the index. In the case of an archive of ARGs, each such distance involves the solution of an error correcting subgraph isomorphism problem, yielding a critical computational complexity.

8.3. Berritti, Bimbo, and Vicario Approach to Modeling Visual Content ARGs

Berritti, Bimbo, and Vicario [25] address the problem of efficient indexing and matching of ARGs as employed in the application context of content-based image retrieval. They describe image models that capture properties of entities and of their relationships as ARGs. They formulate the computation of distance between image models as a problem of optimal (sub)graph error-correcting isomorphism.

In the description of the visual content of an image, the identification of multiple spatial entities permits one to combine information on the visual appearance of salient parts and on their mutual relationships. In general, a spatial entity can be any set of pixels that is cohesive in the user perception of the image appearance. This can be the set of pixels constituting any object with a high-level semantics, such as a character, a face, or a geographic landmark. Or, it can be a set of pixels with low-level visual cohesion, induced by a common chrominance or texture, or by a common position within a predefined area of the image. As a limit case, the overall image itself can be regarded as a particular spatial entity.

Selecting the kind of entities in which content representation should be based entails a trade-off between the significance of the model and the complexity of its creation: Models containing high-level entities permit a closer fit to the users' expressive habits, but they also require manual assistance in the archiving stage for the identification and the classification of significant entities.

Information associated with each entity generally combines the salient features that identify the entity, along with additional indexes that can be measured once the entity has been extracted: A high-level object is usually associated with a symbolic type an image region derived through a color-based segmentation is associated with a chromatic descriptor, and both of them can be associated with a measure of size, or with any other shape index.

Relational information associated with multiple entities can capture high-level concepts, such as an action involving represented objects or spatial relationships between the pixel sets representing different entities. Relationships of the latter kind are the most commonly employed in content-based image retrieval due to the possibility to derive them automatically and to their capability to convey a significant semantics. Spatial relationships may address topological set-theoretical concepts (e.g., inclusion, adjacency, or distance) or directional constructs (e.g., above or below). In both cases, relationships can either be interpreted over a finite set of predefined (symbolic) classes, or they can be associated with numeric descriptors taking values in dense spaces. The latter approach enables the use of distance functions which change with continuity and avoid classification thresholds, thus better coping with the requirements of retrieval by visual similarity.

9. Collaboration and TIE Participation

In Year 2 and 3 of the SCOPE effort, CHI Systems worked closely with a subgroup of the EAGLE contractors called the OddTIE, where TIE stands for technical integration experiment. The group included Alphatech (now part of BAE), SRI, NRL, University of Wisconsin, and NYU. The OddTIE architecture is shown in Figure 8.

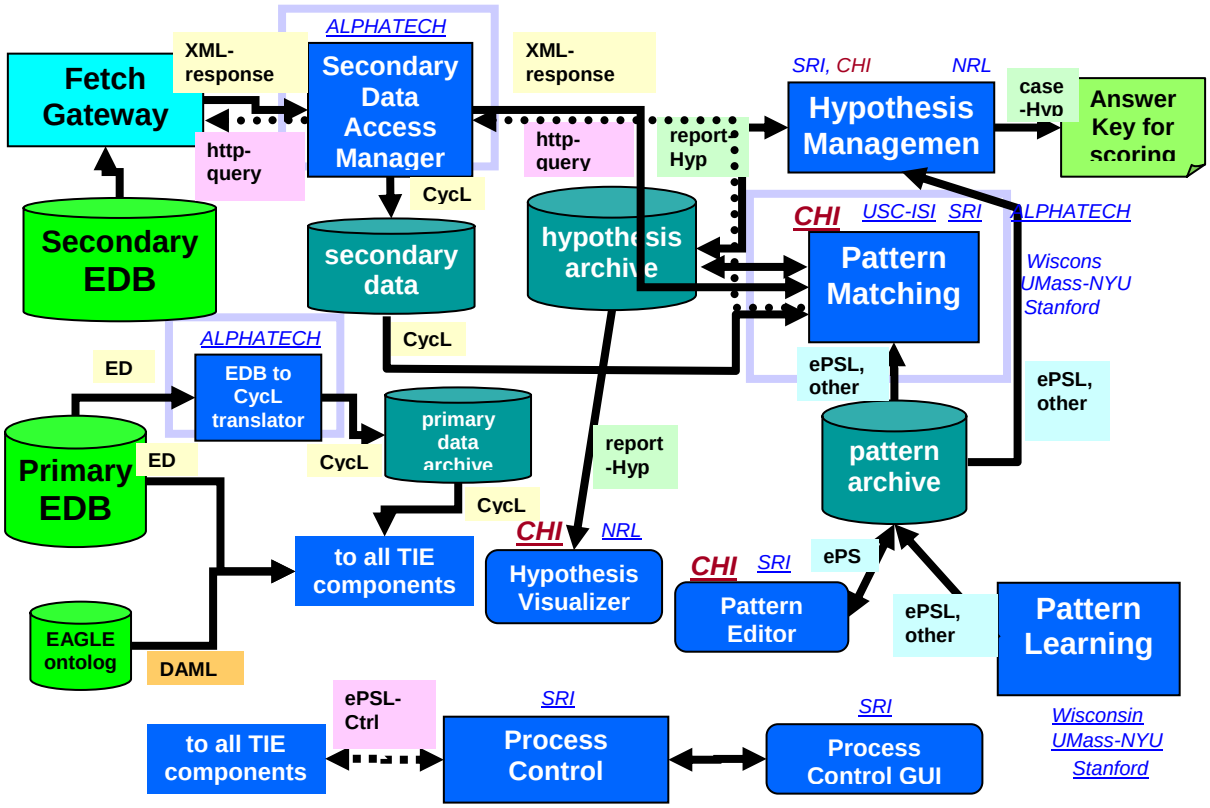

Figure 8: The OddTIE Architecture Diagram

The hypothesis management diagram is shown in figure 9. In Year 2, much of this was done by NRL, but in Year 3, CHI Systems took over and expanded on that role.

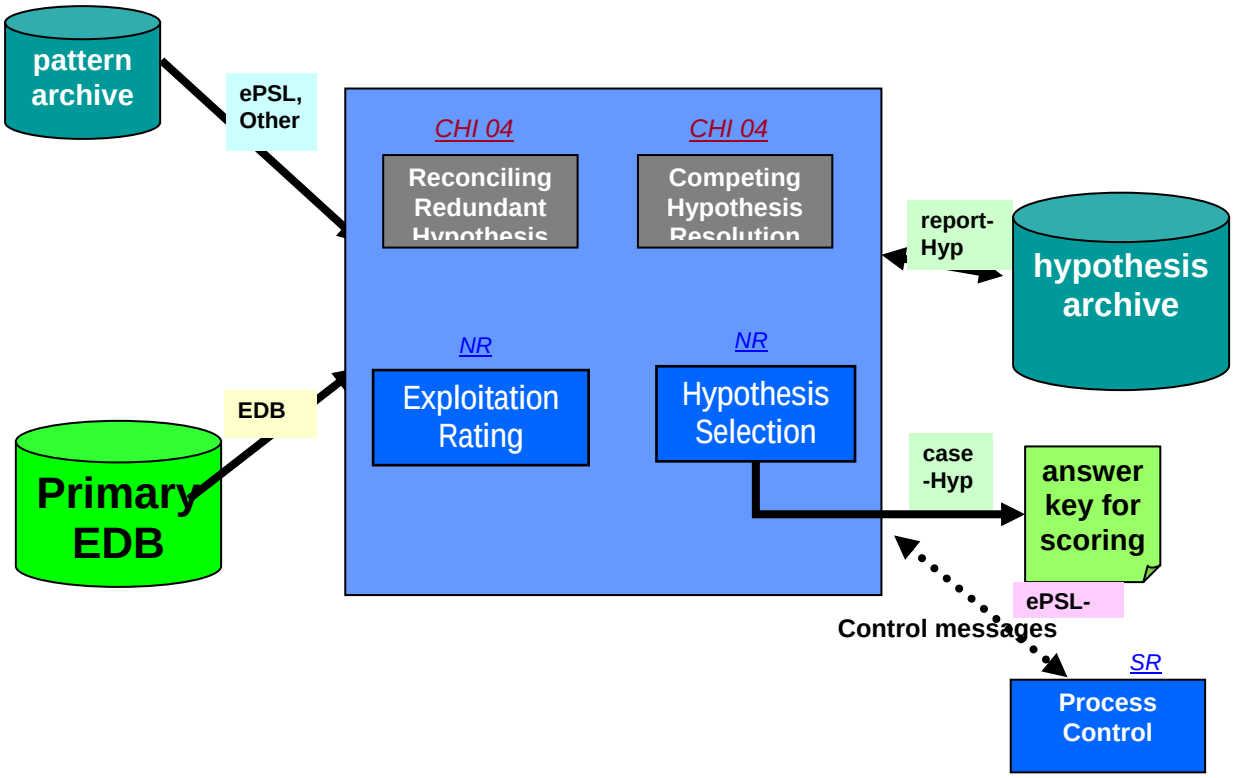

Figure 9: The OddTIE Hypothesis Management Diagram

10. Conclusions

Initially, we believed that the primary benefit of using SCOPE would be in bringing patterns of evidence that might otherwise be missed to the attention of human intelligence analysts. As a result of the CTA we've conducted, we have come to realize that an equally important benefit is the ability to prioritize reports based on their impact on current assertions, and to show how the incoming reports can be linked to those assertions. This realization has important consequences both in term of how SCOPE output should be introduced into an analysis environment, and how SCOPE should be tested. SCOPE can support better analysis by:

- Using multiple strategies (e.g., application of theory, situational logic, and comparison with cases) to elaborate and track alternative assertions;
- Delineating assumptions and chains of inference and specifying the degree and source of uncertainty involved in conclusions about assertions;
- Re-examining key assertions from the ground up periodically in order to avoid the pitfalls of the incremental approach.

11. References

1. Anderson, S., Stone, S. 1995. Historical Dictionary of Terrorism. Metuchen, NJ: The Scarecrow Press.
2. Barrow, H.G. and Popplestone, R.J. (1971) ªRelational Descriptions in Picture Processing,º Machine Intelligence, vol. 5, pp. 377-396.
3. Berretti, S. Bimbo, A.D., and Vicario, E. (2001) "Efficient Matching and Indexing of Graph Models in Content-Based Retrieval" IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 23, NO. 10:1089-1106.
4. Bunke, H. (1999) Error Correcting Graph Matching: On the Influence of the Underlying Cost Function, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21, no. 9, pp. 917-922..
5. Bunke, H. and Shearer, K. (1998) A Graph Distance Metric Based on the Maximal Common Subgraph, Pattern Recognition Letters, vol. 19, pp. 255-259.
6. Eilbert, J.L., Carmody, D.M, Fu, D., Santarelli, T., Wischusen, D., Donmoyer, J. (2002) Reasoning About Adversarial Intent in Asymmetric Situations. Proceeding Intent Inference for Users, Teams, and Adversaries. AAAI 2002 Fall Symposium Series
7. Eilbert, J.L., Carmody, D.M, Fu, D., Santarelli, T., Wischusen, D., Donmoyer, J. (2002) Reasoning About Adversarial Intent in Asymmetric Situations. Proceeding Intent Inference for Users, Teams, and Adversaries. AAAI 2002 Fall Symposium Series
8. Eilbert, J.L., Hicinbothom, J., Karnavat, A. (Submitted) Finding Organization Membership With a New Model of the Analytic Process. 2005 International Conference on Intelligence Analysis Methods and Tools (IA 2005). McLean, VA.
9. Eilbert, J.L., Reed, F., Bracken, K. Campbell, G.E. (2001) Proactive Delivery of Human Factors Information to a Design Team by an Automated Advisor, Proc. Human Systems Integration Symposium, Arlington, VA.
10. Eilbert, J.L., Santarelli, T., Carmody, D.M, (2004) A new framework for cognitive/perceptual problems. ICCM.
11. Eilbert, J.L., Santarelli, T., Carmody, D.M, (2004) A new framework for cognitive/perceptual problems. ICCM.
12. Endsley, M.R (2000) Theoretical underpinnings of situation awareness:A critical review, In: Situation Awareness Analysis and Measurement. Endsley, M.R and Garland, D.J. (eds.), Mahwah,NJ, Lawrence Erlbaum Assoc.
13. Eshera, M.A. and Fu, K.S. (1986) An Image Understanding System Using Attributed Symbolic Representation and Inexact Graph-Matching. J. Assoc. Computing Machinery, vol. 8, no. 5, pp. 604-618.
14. Fauconnier. G. (1998). Mental spaces, language modalities, and conceptual integration. The New Psychology of Language: Cognitive and Functional Approaches to Language Structure. Ed. M. Tomasello. Lawrence Erlbaum
15. Fischler, M. and Elschlager, R., (1973) The Representation and Matching of Pictorical Structures, IEEE Trans. Computers, vol. 22, no. 1, pp. 67-92, Jan..

16. Gettys, C. 1980. Hypothesis Generation: A Final Report on Three Years of Research, Technical Report 15-10-80. University of Oklahoma, Decision Processes Laboratory.
17. Gold,S. and Rangarajan, A. (1996) A Graduated Assignment Algorithm for Graph Matching, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18, no. 4, pp. 377-388.
18. Heuer, R. (1999) The Psychology of Intelligence Analysis. Center for the Study of Intelligence, Central Intelligence Agency. http://www.cia.gov/csi/books/19104/index.html.
19. Kaplan, S., Kaplan, R. (1982) Cognition and Environment. New York:Praeger
20. Lamme, V.A.F., Roelfsema, P.R. (2000) The distinct modes of vision offered by feedforward and recurrent processing. Trends in Neuroscience 23(11): 571-579.
21. Le Mentec, J.-C., Zachary, W., Iordanov, V. 1999. "Knowledge Tracing of Cognitive Tasks for Model-based Diagnosis." Proceeding of 9th International Conference on Artificial Intelligence in Education, Le Mans, France.
22. Lehnert, W.G., Loiselle, C.L. (1989) An introduction to plot units. In Waltz, D.L. (ed.) Semantic Structures: Advances in Natural Language Processing. Lawrence Erlbaum Associates, Inc. Hillsdale, NJ.
23. Lou, B., Hancock, E.R. (2001) Structural graph matching using the EM algorithm and singular value decomposition. IEEE Trans. On Pattern Analysis and Machine Intelligence. VOL. 23, NO. 10:1120-1136
24. Pelillo, M. Siddiqi, K. and Zucker, S. W. (1999) Matching Hierarchical Structures Using Association Graphs,º IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21, no. 11, pp. 1105-1120.
25. Pylyshyn, Z.W. (2000) Situating vision in the world. Trends in Cognitive Sciences. Volume 4(5): 197-207.
26. Ratches, J.A., Walters, C.P., Buser, R.G. and Guenther, B. D (1997) Aided and Automatic Target Recognition Based Upon Sensory Inputs From Image Forming Systems. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 19, NO. 9
27. Ratches, J.A., Walters, C.P., Buser, R.G.and Guenther, B. D (1997) Aided and Automatic Target Recognition Based Upon Sensory Inputs From Image Forming Systems. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 19, NO. 9
28. Saha, P.K. and Udupa, J.K. (2001) "Relative Fuzzy Connectedness among Multiple Objects: Theory, Algorithms, and Applications in Image Segmentation," Computer Vision and Image Understanding, vol. 82, pp. 42-56.
29. Sanfeliu, A. and Fu, K. S. (1983) A Distance Measure between Attributed Relational Graphs for Pattern Recognition, IEEE Trans. Systems, Man, and Cybernetics, vol. 13, no. 3, pp. 353-362.
30. Scott, G.L. and Longuet-Higgins, H.C. (1991) An Algorithm for Associating the Features of 2 Images. Proc. Royal Soc. London Series B, vol. 244, no. 1309, pp. 21-26.
31. Shapiro, L.G. and Haralick, R.M., (1985) A Metric for Comparing Relational Descriptions, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 7, no. 1, pp. 90-94..

32. State Department Report 2000. Patterns of Global Terrorism 1999 – Middle East Overview. http://www.usinfo.state.gov/topical/pol/terror/00050107.htm.
33. Stottler, R.H., Henke, A.L., King, J.A., 1989. “Rapid Retrieval Algorithms, for Case-Based Reasoning," Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), Detroit, MI.
34. Szczepkowski, M.A., Wischusen, D., & Hicinbothom, James (2003). Work-Centered Support Development Methodology in Tactical Aviation. In *Proceedings of the 22st Digital Avionics Systems Conference (DASC)*, Indianapolis, IN, October, 2003.
35. Thorpe, S., Fize, D. & Marlot, C. (1996) Speed of processing in the human visual system. Nature 381, 520-522.
36. Udupa, J.K., Saha, P.K., Lofuto (2002) Relative Fuzzy Connectedness and Object Definition: Theory, Algorithms, and Applications in Image Segmentation. IEEE Transactions On Pattern Analysis and Machine Intelligence, Vol. 24, No. 11
37. Umeyama S., (1988) An Eigen Decomposition Approach to Weighted Graph Matching Problems,º IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 10, pp. 695-703.
38. Upada, J.K., Saha, P.K., Lofuto (2002) Relative Fuzzy Connectedness and Object Definition: Theory, Algorithms, and Applications in Image Segmentation. IEEE Transactions On Pattern Analysis and Machine Intelligence, Vol. 24, No. 11
39. Vargas-Vera, M., Domingue, J., Kalfoglou, Y., Motta, E., Shum, S.B. (2001) Template-driven information extraction for populating ontologies. Proc. IJCAI 2001 Workshop on Ontology Learning.
40. Vicente, K. J. and J. Rasmussen (1992). Ecological interface design: theoretical foundations. IEEE Transactions on Systems Man Cybernetics 22(4): pp. 589-606.
41. Wilson, R.C. and Hancock, E.R. (1997) Structural Matching by Discrete Relaxation, IEEE Pattern Analysis and Machine Intelligence, vol. 19, no. 6, pp. 634-648, June 1997.
42. Young, M., Eggleston, R., & Whitaker, R. (2000). Direct Manipulation Interface Techniques for Users Interacting with Software Agents. Proceedings NATO/TRO Symposium on Usability of Information in Battle Management Operations. Oslo, Norway.
43. Yu, S.X., and Shi, J. (2003) Object-Specific Figure-Ground Segregation. IEEE Conference on Computer Vision and Pattern Recognition, Madison, Wisconsin.
44. Zachary W., and Bell, B. (2003) Work-centered Infomediary Layer (WIL): An architecture for adaptive interfaces. Human Interaction With Autonomous Systems in Complex Environments. Papers from the 2003 AAAI spring symposium. Technical Report SS-03-04. Menlo park, CA: American Association for Artificial Intelligence. Pp. 208-212.
45. Zachary, W, Eggleston, R. (2002). A Development Environment and Methodology for the Design of Work-Centered User Interface Systems, *Proceedings of Human Factors And Ergonomics Society 46th Annual Meeting*, Santa Monica, CA, Human Factors and Ergonomics Society, pp. 656-660.
46. Zachary, W. (1985). Beyond user-friendly: Building decision-aid interfaces for expert end users. Cybernetics and Society Conference Proceedings, Tucson, Arizona. IEEE: New York. pp. 641-647.

47. Zachary, W. W., Ryder, J. M., Ross, L., & Weiland, M. Z. (1992). Intelligent computer-human interaction in real-time multi-tasking process control and monitoring systems. In M. Helander & M. Nagamachi (Eds.), Human factors.
48. Zachary, W., Eggleston, R., Donmoyer, J., & Schremmer, S. (2003). A work-centered cognitively-based architecture for decision support: the work-centered infomediary layer (WIL) model. In *Proceedings of the SPIE's AeroSense 2003 Conference - Enabling Technologies for Simulation Science VII*, Orlando, Florida.
49. Zachary, W., Schremmer, S., & Donmoyer, J. (2002). *Design Analysis for A Work-centered Infomediary Layer Application Toolkit (WAT)*, Springhouse, PA, CHI Systems, Inc. Technical Report 011214.01011.

12. Publications

1. Eilbert, J.L., Hicinbothom, J., Karnavat, A. (submitted) A Cognitive Architecture Focusing on Situation Estimation: Merging Perceptual and Cognitive Processes, Behavior Representation in Modeling and Simulation (BRIMS 2005).
2. Eilbert, J.L., Hicinbothom, J., Karnavat, A., Santarelli, T. (submitted) Finding Organization Membership With a New Model of the Analytic Process. International Conference on Intelligence Analysis (IA'05).
3. Eilbert, J., Santarelli, T., Carmody, D. (2004) A New Framework for Cognitive/Perceptual Problems. ICCM Conference.
4. Fu, D., Remolina, E., Eilbert, J. (In Press) A CBR Approach to Asymmetric Plan Detection. Proc. Link2003 workshop at KDD'2003.
5. Eilbert, J.L., Carmody, D.M, Fu, D., Santarelli, T., Wischusen, D., Donmoyer, J. (2002) Reasoning About Adversarial Intent in Asymmetric Situations. Proceeding Intent Inference for Users, Teams, and Adversaries. AAAI 2002 Fall Symposium Series.
6. Eilbert, J.L., Glenn, F., Hicinbothom, J., Reed, F. (2001) Using Cognitive Models of Teams to Determine the Need for Intent Inferencing in Various Situations. AAAI Fall Symposium on Intent Inferencing.