The Sensemaking Process and Leverage Points for Analyst Technology as Identified Through Cognitive Task Analysis

Peter Pirolli and Stuart Card

PARC

3333 Coyote Hill Road

Palo Alto, California, 94304, USA

{card, pirolli}@parc.com

Keywords: Sensemaking, cognitive task analysis, schema, empirical studies

Abstract

There are a relatively few open literature reports that provide empirical descriptive studies of intelligence analysis and that link these into the context of expertise and work. This paper, based on first results from a cognitive task analysis and verbal protocols give a broad brush description of intelligence analysis as an example of sensemaking. It then suggests some possible leverage points where technology might be applied.

1. Introduction

A number of reports are emerging that provide normative or prescriptive views on intelligence analysis. There have been very few that provide empirical descriptive studies. Furthermore, there are only a limited number of studies that set the process of analysis in the context of the structure of the process involved. Despite the spottiness of the available literature, what does exist reveals that intelligence analysis is a widely variegated task domain. This means that it is important to be careful in making generalizations from any circumscribed types of intelligence tasks or types of analysts. It is equally important not to be daunted by the vastness of the domain, and to start the investigative venture somewhere. This paper reports on preliminary results from a study using cognitive task analysis and think aloud protocol analysis to help broadly characterize the processes used. The purpose is to not to work toward incremental improvements on the current process, but by empirically understanding that process in greater depth to find leverage points where improvement of that process could make a difference, including the radical replacement where warranted with new processes.

2. A Notional Model of Analyst Sensemaking

2.1 Expertise Schemas

It is useful to set intelligence analysis into a more general context. The fact that intelligence analysis is a form of expert behavior leads us to several expectations about the nature of that behavior based on what is known about experts in general. One of these expectations is that experts will have built up from extensive experience a set of patterns around the important elements of their tasks, which we here call schemas. The key to expert performance is more to be found in these domain specific schemas (Ericsson and Lehmann, 1996) than in personal capacities. In a classic study, for example, Simon and Chase (1973) showed how master-level skill in chess derived from patterns built up over 15,000 hours. With these patterns the masters could easily reproduce chess board positions from memory even though they had no better ordinary memory skills than novices.

Klein and associates (Klein 1989, 1998) have developed a model of recognition-primed decision making, as part of a program on naturalistic decision making that has been used as the basis of military command and control. This model emphasizes the role of these schematic knowledge structures built from expertise and experience in allowing a soldier or a firefighter to make sense of a situation and rapidly to formulate an action. In fact, they show that for an expert, it can actually be the lack of expected features of a situation that can trigger sensemaking and action. In a study of an expert business intelligence analyst, we found (Pirolli & Card, 1999) a set of schemas around market players, application opportunities, market development, report types, and the like. These were partly the basis of the analyst's skill that allowed him quickly to organize the flood of incoming information and present it in reports. As Ericsson & Lehmann (1966) summarize about the role of skill schemas,

Experts don't just automatically extract patterns and retrieve their response directly from memory. Instead, they select the relevant information and encode it in special representations . . . that allow planning, evaluation and reasoning about alternative courses of actions.

In our interviews and protocols with intelligence analysts, we found evidence of schemas used to organize the information. Information that came in was reorganized to these representations as an aid in planning, evaluation, and reasoning. It will not be surprising that some of these reorganizations centered around ways to understand people, organizations, tasks, and time:

- One analyst had organized a large collection of descriptions of people and organizations in order to predict whether a coup would occur.

- Another analyst had organized a map of people, institutions, projects, and biological mechanisms to understand future bio-warfare threats.

- Yet another mapped out relationships and methods among drug cartels.

- Still another made linkage maps among telephone numbers, persons, addresses, and other information to solve a terrorist exercise.

These efforts by analysis show attempts to amplify their ability to find patterns for the conceptual schemas they use in understanding their domains of analysis.

2.2 Sensemaking Process

The analyst's conceptual schema sometimes play a central role in the intelligence activities. Many forms of intelligence analysis are what we might call sensemaking tasks. Such tasks consist of information gathering, re-representation of the information in a schema that aids analysis, the development of insight through the manipulation of this representation, and the creation of some knowledge product or direct action based on the insight. In a formula

Information ➔ Schema ➔ Insight ➔ Product

The re-representation may be informally in the analyst's mind or aided by a paper and pencil or computer-based system.

Russell et al. (1993) have explicated the inner re-representational sensemaking process for a case in which large amounts of information had to be digested to create a curriculum for printer repairmen. The core of the process is what they call a "learning loop complex" (Figure 1). First is a search for a good representation (the generation loop). Then there is an attempt to encode information in the representation (the data coverage loop). The attempt at encoding information in the representation identifies items that do not fit ("residue"). This gives rise to an attempt to adjust the representation so that it has

better coverage (the "representation shift loop"). The result is a more compact representation of the essence of the information relative to the intended task. In the case of experts and repeated tasks, the representation may not be problematic and most activity would concern the coverage loop.

Figure 2 represents our notional understanding of the analyst's process derived from our preliminary cognitive task analysis. This is a broad brush characterization of the whole process we have seen across several all-source analysts as a way of approximating the process and orienting our ongoing more detailed studies. The rectangular boxes represent an approximate data flow. The circles represent the process flow. The processes and data are arranged by degree of effort and degree of information structure. This is a process with lots of back loops and seems to have one set of activities that cycle around finding information and another that cycles around making sense of the information, with plenty of interaction between these. This process diagram summarizes how it is that an analyst comes up with novel information.
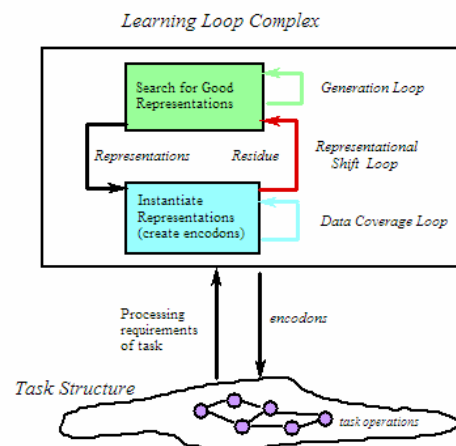

Figure 1. Learning Loop Complex theory of sensemaking (Russell, et al. 1993).

The data flow show the transformation of information as it flows from raw information to reportable results. *External data sources* are the raw evidence, largely text by the time it reaches the all sources analyst. The *"shoebox"* is the much smaller subset of that external data that is relevant for processing. The *evidence file* refers to snippets extracted from items in the shoebox. *Schemas* are the re-representation or organized marshalling of the information so that it can be used more easily to draw conclusions. *Hypotheses* are the tentative representation of those conclusions with supporting arguments. Ultimately there is a *presentation* or other work product. Basically the data flow represents the transducing of information from its raw state in to a form where expertise can apply and then out to another form suited for communication.
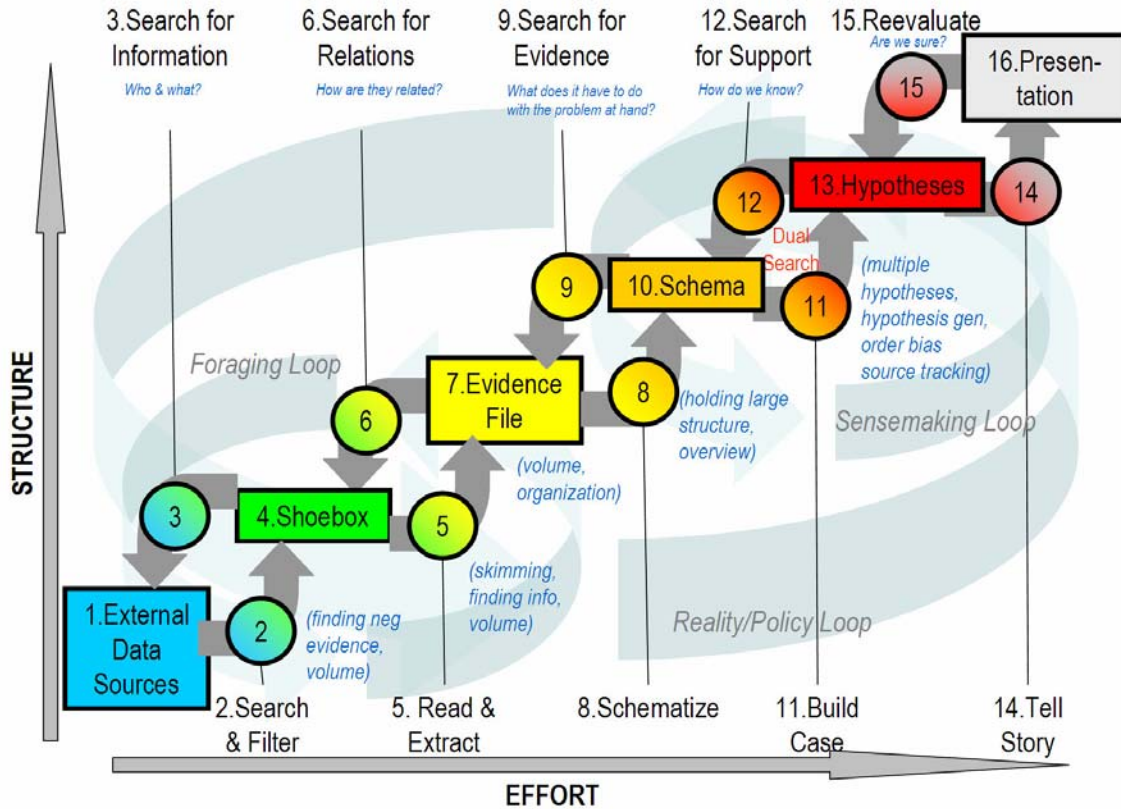
Figure 2. Notional model of sensemaking loop for intelligence analysis derived from CTA.

The overall process is organized into two major loops of activities: (1) a *foraging loop* that involves processes aimed at seeking information, searching and filtering it, and reading and extracting information (Pirolli & Card, 1999) possibly into some schema, and (2) a *sense making loop* (Russell, Stefik, Pirolli, & Card, 1993) that involves iterative development of a mental model (a conceptualization) from the schema that best fits the evidence. Two caveats are that (1) much day-to-day intelligence mainly consists of extracting information and repackaging it without much actual analysis and (2) schema-based expert skill can be used at all points in the process of Figure 2, for example in rapidly skimming and rejecting information in the early stages.

Information processing in Figure 2 can be driven by *bottom-up* processes (from data to theory) or *top-down* (from theory to data). Our analysis suggested that top-down and bottom-up processes are invoked in an opportunistic mix.

Bottom-up Processes

- *Search and filter*. External data sources, such as Pubmed, the Web, or classified databases, provide a repository that is searched (queried) by the analyst. Results of those searches are filtered (judged) for relevance. An analyst filters message traffic or does active search, collecting relevant documents into some store (the "shoebox" in the diagram) for further processing.

- *Read and extract*. Collections of shoebox evidence are read to extract nuggets of evidence that may be used to draw inferences, or support or disconfirm theory. Relevant snippets from this store and related low-level inferences are placed in evidence files. Evidence extracted at this stage may trigger new hypotheses and searches.

- *Schematize*. At this point the information may be re-represented in some schematic way. Given a lack of easily-used tools, this may be in the mind of the analyst, informal (as in Figure 3), a simple marshalling, or an elaborate computer-based method, for example, a time line visualization to coordinate many events. Evidence may be organized into small-scale stories about typical topics or in answer to typical questions (e.g., who, what, when, where, why, how) that are used to organize raw evidence.

- *Build case*. A theory or case is built by additional marshalling of evidence to support or disconfirm hypotheses (as in Figure 4).

- *Tell story*. A presentation or publication of a case is made to some audience (client).

Top-down processes

- *Re-evaluate*. Inquiries or feedback from clients of a presentation may generate re-evaluations of the current theory developed by an analyst requiring the marshalling of additional evidence to support or disconfirm the theory or the generation and testing of alternative theories.
- *Search for support*. Analysis or re-evaluation of theories may require re-examination of the lower-level schematic organization of basic facts.
- *Search for evidence*. Analysis or re-evaluation of theories may require re-examination of collected evidence, or the searches for new evidence.
- *Search for relations*. Nuggets of information in an evidence file may suggest new patterns (e.g., people linked to other people) that generate hypotheses about plausible relations among entities and events. These hypotheses may generate new searches and data extraction from the shoebox and raw data.
- *Search for information*. New hypotheses generated from processes at higher levels may cause the analyst to dig deeper in the raw data.

Other researchers have come to a similar conclusion about the nature of sensemaking for intelligence analysts and first responders. For example, Klein et al (in press) have a "data/frame"-based theory of sensemaking, which plays a similar role to "schema" in Figure 2. For Klein, a frame is a mental structure that organizes the data and sensemaking is the process of fitting information into that frame. Frames are a consequence of developed expertise. Bodnar (2003) describes a process similar to Figure 2 in his book on warning analysis for intelligence. Leedom (2001) has reviewed notions of sensemaking concepts in military decision making that range from situational awareness in air combat to naturalistic decision making to organizational sensemaking.
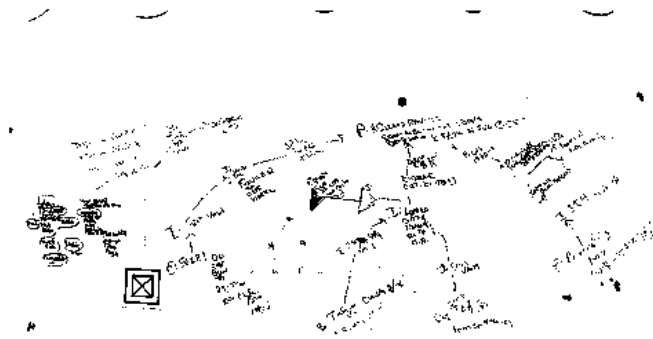

Figure 4. Hypothesized solution as drawn by analyst

LEVEREAGE POINTS

The cognitive task analysis suggests a set leverage points that we organize by the two major loops in Figure 2: (a) a foraging loop and (b) a sense making loop. We are particularly concerned with intelligence analysis involving massive amounts of data, so most of these leverage points are related to issues of data overload and attention management.

Foraging Loop

The foraging loop is essentially a tradeoff among three kinds of processes summarized schematically in Figure 5. Figure 5 is based on data visualizations of analysts behavior produced in Patterson, Roth, and Woods (2001). Patterson et al. observed that analysts tended to begin with a broad set of documents, for instance one that was retrieved by a high-recall/ low-precision query, and then proceeded to narrow that set down into successively smaller, higher-precision sets of data, before reading and analyzing the documents. The successively smaller rings in Figure 5 indicate successive narrowing activity by a hypothetical analyst with a broader universe of documents. The marks in Figure 5 indicate high profit, relevant items—some of which lie within the set of documents under consideration by the analyst and some lying outside. More generally, we may consider three processes that tradeoff against one another under deadline or data overload constraints:
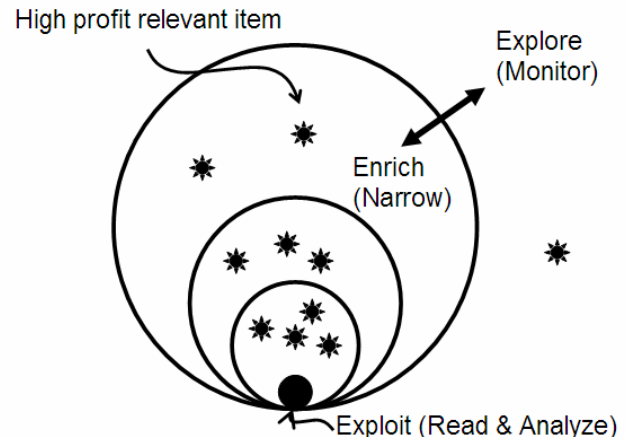

Figure 5. The exploration-enrichment-exploitation tradeoff in information foraging.

1. *Exploring* or monitoring more of the space, by which we mean increasing the span of new information items into the analysis process. In information retrieval terms this would correspond to increasing the recall of the information search.

2. *Enriching* (or narrowing) the set of items that has been collected for analysis (Pirolli & Card, 1999). This a process in which smaller, higher-precision sets of documents are created.

3. *Exploiting* the items in the set, by which we mean more thorough reading of documents, extraction of in-

formation, generation of inferences, noticing of patterns, etc.

Some of the leverage points concerning the foraging loop that emerge from the cognitive task analysis include:

- **The cost structure of the exploration-enrichment-exploitation tradeoff, which is related to the precision-recall trade-off in information retrieval.**

By cost structure, we mean the absolute and relative time costs of information operations. These costs structure user information behavior (Pirolli & Card, 1999) and can often been altered (positively or negatively) by compute or methodological innovations. It will generally be desirable to explore as much of the information space as possible (because there may be a cost to missing something novel in the data) but this comes at the cost of having to actually work through the material and eventually exploit it. One class of solutions (that comes by analogy to human sensory and perceptual systems such as vision) would be to produce human-information interaction systems that foster broad-band, low-fidelity assessments of incoming data coupled with narrow-band, high-fidelity processing. Focus+context techniques (Furnas, 1986; Furnas, 1983) are examples of these kinds of solutions.

- **The cost structure of scanning, recognizing (assessing), and selecting items for further attention.**

Our analysts spent considerable time scanning data seeking relevant entities (names, numbers, locations, etc.). The assessment of whether or not an item is relevant also takes time. Techniques for highlighting important information with pre-attentive codings, or re-representing documents (e.g., by summaries) appropriate to the task can improve these costs.

- **The cost of shifting attentional control.**

The need to shift attention to a novel domain of information may arise from top-down goals (e.g., new taskings) or bottom-up from noticing something of interest in the data (e.g., an anomaly or novel connection). Starting up on a new task (whether goal-initiated or data-driven) is usually quite costly.

- **The cost of follow-up searches.**

As information is extracted and analyzed it often generates new questions and hypotheses requiring additional search.

One simple way of looking at these leverage points, that at least provides some guidance and a generalized metric for measuring progress, is presented in Figure 6. Expert analysts often set their filters for information *lower* (thereby accepting more irrelevant information) because they want to make sure that they don't miss something that is relevant. You might say that they look for fainter signals by accepting more noise. Since the expert analyst has a lot of prior and tacit knowledge, they are *faster* in rejecting the unimportant messages than a less expert analyst. Figure 6 shows a metric for describing this phenomenon as a graph. The figure illustrates

that the expert analyst can get more of the relevant information in a shorter amount of time (the higher curve). This can be exploited as the obtaining the same information in less time, or more information in the same time, or a combination (the green arrows).
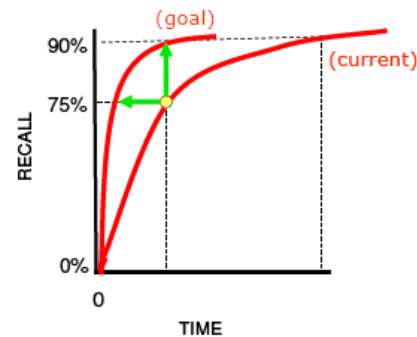

Figure 6. A cost structure metric for the foraging loop.

Sense making loop

Many of the leverage points (pain points) associated with the sense making loop concern problem structuring (the generation, exploration, and management of hypotheses), evidentiary reasoning (marshalling evidence to support or disconfirm hypotheses), and decision making (choosing a prediction or course of action from the set of alternatives). These processes are affected by many well-known cognitive biases.

- **Span of attention for evidence and hypotheses.**

Human working memory has inherent capacity limits and transient storage properties that limit the number of hypotheses, the amount of evidence, and the number of evidentiary relations that can be simultaneously heeded. One leverage point is to improve the capacity of analysts to attend to more of the structure of organized evidence and hypotheses. Reasoning about evidence and hypotheses has an exponential cost structure. The verbal protocols of analysts revealed how rapidly the number of relations (patterns) among data can grow, such as the social networks of Russian scientists, or the calling networks among suspected terrorists. Techniques aimed at expanding the working memory capacity of analysts by offloading information patterns onto external memory (e.g., visual displays) may ameliorate these problems. Information visualization techniques (Card, Mackinlay, & Schneiderman, 1999) and broad band displays are examples of such techniques.

- **Generation of alternative hypotheses.**

Human perception is biased towards interpretation of information into existing schemas and existing expectations. Human reasoning is subject to a variety of well-documented heuristics and biases (Tversky & Kahneman, 1974) that deviate from normative rationality. In problem structuring and decision analysis, people typically fail to generate hypotheses. Time pressures and data overload

work against the individual analyst's ability to rigorously follow effective methods for generating, managing, and evaluating hypotheses. Improving the space of possibilities covered by a set of generated hypotheses is another leverage point.

- **Confirmation bias.**

People typically fail to consider the diagnosticity of evidence, and fail to focus on the disconfirmation of hypotheses. Evidence is fit to existing schemas of thought. A leverage point for new tools is to distribute more attention of the analyst to highly diagnostic evidence and to the search for disconfirming relations.

Conclusion

The notional model presented in Figure 2 provides an organization for identifying new technologies for improving the production of novel intelligence from massive data. Intelligence analysis is seen as a form of sensemaking and expert skill. The leverage points identified above provide a framework for new technology design principles and the development of evaluation studies and metrics. Proposals for tools or new methods for intelligence analysis can often profitably be discussed in terms of where in this process they propose to intervene and which leverage point they seek to improve. More radical proposals can be discussed in terms of the effects of, for example, changing data and threat characteristics and how they lead to modifications of this process. We intend to pursue these issues more in later analyses.

Acknowledgements

References

Card, S. K., Mackinlay, J. D., & Schneiderman, B. (1999). *Information visualization: Using vision to think*. San Francisco: Morgan-Kaufmann.

Chipman, S. F., Schraagen, J. M., & Shalin, V. L. (2000). Introduction to cognitive task analysis. In J. M. Schraagen & S. F. Chipman & V. L. Shalin (Eds.), *Cognitive task analysis* (pp. 3-23). Mahwah, NJ: Lawrence Erlbaum Associates.

Endsley, M. R. 1995. Situation awareness and the cognitive management of complex systems. Human Factors Special Issue 37(1): 85-104.

Ericsson, D. Anders 1999. Expertise. In Wilson, R. A. and Keil, Frank, 1999, The MIT Encyclopedia of the Cognitive Sciences. Cambridge, MA: MIT Press.

Ericsson, K. A. and Lehmann, A. C., 1996. Expert and exceptional performance: evidence on maximal adaptations on task constraints. Annual Review of Psychology 47: 273-305.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol Analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Furnas, G. (1986). Generalized fisheye views, *CHI '86, Proceedings of the ACM conference on human factors in software* (pp. 16–23). New YORK: ACM.

Furnas, G. W. (Artist). (1983). *Small views of large structure.*

Klein, G. A. 1989. Recognition-primed decisions. In W. B. Rouse, ed. Advances in Man-machine Systems Research Greenwich, CT: JAI Press, Inc.

Klein, G. A. (1998). Sources of Power: How People Make Decisions. Cambridge, MA: MIT Press.

Klein, G. et al, in press. A data/frame theory of sensemaking. In R. Hoffman., ed. Expertise Out of Context. Mahwah, NJ: Lawrence Erlbaum Associates.

Leedom, D. K. 2001. Final report: Sensemaking Symposium (Technical Report prepared under contract for Office of Assistant secretary of Defense for Command and Control.)

Militello, L. G., Hutton, R. J. B., Pliske, R. M., Knight, B. J., & G. Klein. (1997). *Applied cognitive task analysis (ACTA) methodology* (Tech. Rep. No. NPRDC-TN-98-4). Fairborn, OH: Klein Associates.

Patterson, E. S., Roth, E. M., & Woods, D. D. (2001). Predicting vulnerabilities in computer-supported inferential analysis under data overload. *Cognition Technology and Work, 3*, 224-237.

Pirolli, P., & Card, S. K. (1999). Information foraging. *Psychological Review, 106*, 643-675.

Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993). *The cost structure of sensemaking.* Paper presented at the INTERCHI '93 Conference on Human Factors in Computing Systems, Amsterdam.

Simon, A. H., and Chase, W. g., 1973. Skill in chess. American Scientist, 61: 3940-403.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131.