

Applications of BICA to Intelligence Analysis

Paul Kogut, Norris Heintzelman, and Greg Stachnick

Lockheed Martin Information Systems & Global Services

P.O. Box 8048 Philadelphia, PA 19101, USA

paul.a.kogut@lmco.com, norris.h.heintzelman@lmco.com, gregory.l.stachnick@lmco.com

Abstract

This paper describes potential applications of Biologically Inspired Cognitive Architectures to Intelligence Analysis. The focus of our efforts is on higher level reasoning rather than low level perception. We will never have enough human analysts to read, filter and make sense of all the text data out there. Can some form of BICA help? In this paper we discuss issues related to knowledge acquisition, natural language processing and cognitive architectures that we have encountered in an ongoing project to apply the Sandia Cognitive Framework to analysis problems. We believe that studying intelligence analysis will lead to new insights into BICA.

- reasoning about situations that they do not recognize to attempt to find an explanation?
- learning knowledge that will help them recognize future situations?

A situation is the state of the world usually in an unfolding story (e.g., a terrorist plot, a political movement or an evolving startup company). Recent results from cognitive science and neuroscience are beginning to reveal answers to these questions. This paper discusses some initial experience with building BICA systems and applying them to intelligence analysis.

Why BICA?

Human analysts leverage large amounts of context, previous cases, domain specific expertise and common sense knowledge to assemble explanations of incoming real world data. It is difficult and time consuming to capture this knowledge in the form of rules, ontologies and Bayesian networks. These conventional Artificial Intelligence technologies are not well suited for representing and reasoning with the complex narrative knowledge that analysts use. Some of the knowledge that the analyst uses evolves rapidly and is acquired as part of the task. So traditional manual knowledge engineering approaches cannot keep up with a dynamically changing world. In intelligence analysis and law enforcement, new individuals, groups and social/cultural/political trends emerge constantly. The same is true for other forms of analysis such as competitive intelligence and venture capital investment. There is clearly a need for some form of system that can apply rich knowledge and can be quickly trained by reading natural language text. One potential path to overcoming the limitations of conventional AI approaches is to look into how the brain does these higher level functions and build a system based on these insights. What parts of an analyst's brain are involved in:

- recognizing a situation based on previous situations?

Intelligence Analysis and BICA

The volume of text, images and video that is available for analysis is rapidly increasing. There are not enough analysts to keep up with this volume. A recent model of the intelligence analysis process includes two heavily interacting loops of sensemaking (generating a hypothesis that explains the evidence) and foraging (searching and filtering data) (Bodnar 2005). The evidence found during the foraging process is fed into sensemaking. For example, startup company X has a strong technical person and a strong business person along with investment from a prestigious venture capital firm which leads to the hypothesis that company X has high potential. Additional foraging is triggered and guided by sensemaking to fill in gaps in the evidence and support or refute the hypothesis. For example, search for more information about company X's technology and market forecast.

Some form of BICA could be applied to both foraging and sensemaking to help analysts analyze more data. The degree of human involvement could vary from a highly interactive personal assistant to a highly autonomous system that runs overnight and presents the analyst with a prioritized list of potential evidence and hypotheses. Even a system that has a significant rate of false positive results would be valuable because much of this data would be otherwise ignored due to manpower limitations.

From the cognitive perspective, sensemaking involves recognizing situations and reasoning about situations that are not immediately recognized. Foraging involves recognizing data that is relevant to the situation and reasoning about how to find more relevant data. Both

sensemaking and foraging require large amounts of knowledge which implies the need for the general cognitive processes of learning a large repertoire of situations. So we believe that sensemaking, foraging and the forms of learning associated with these processes involve similar areas and pathways in the brain. The focus of this investigation is on the cognitive processes that happen in the hippocampus and cortex after low level perception occurs. Sandia National Labs has been focusing on these areas of the brain in the development of their cognitive framework.

Intelligence Analysis BICA prototype

The goal of our initial prototype effort was to determine the feasibility of applying BICA to analysis. The focus of our initial prototype was sensemaking on evidence drawn from text sources. The primary component for situation recognition was the Sandia Cognitive Runtime Engine with Active Memory (SCREAM) (Bernard 2006). SCREAM has a semantic memory that supports associative priming between concepts (e.g., *doctor* activates *nurse*). SCREAM also has a hierarchical context module that supports bottom up activation of contexts (e.g., *doctor* and *nurse* and *bed* activate *hospital*) and top down priming of other expected concepts (e.g., *hospital* activates *emergency room*).

In the initial prototype, the cognitive models used by SCREAM were developed by traditional manual knowledge engineering techniques. Perceptions in the form of incrementally arriving text reports were processed with hybrid natural language processing (NLP) techniques into a form that would trigger activation of SCREAM concepts and contexts. The NLP subsystem includes AeroText a rule-based information extraction system and STANLEY a statistical clustering tool (Bauer 2005). AeroText recognizes multi-word named entities which STANLEY uses as a basis for determining concept associations which are then passed to SCREAM as perceptual inputs.

The output of SCREAM was an indication that one or more higher level contexts (hypotheses) were activated. The analyst could examine the hypothesis along with the chain of activations that provides justification. SCREAM tolerates inconsistencies unlike logic based rules and ontologies. The highest level of hypothesis in the experiment was supported by 30 pieces of evidence with multiple inference chains that had up to 6 steps.

We also experimented with a novel approach for reasoning about situations that are not immediately recognized by SCREAM (e.g., a previously unseen situation). We investigated the use of an abductive reasoning system called the Peircean Decision Aid (PDA) (Senglaub 2008) to activate additional contexts based on past experience derived by induction from data in previous situations rather than actual evidence in the current situation. For example, it is plausible that an object can be used for a certain purpose but there is no evidence yet that there is intention to use it that way.

Future Research

There are many open issues to explore. First, the interaction between recognizing known situations (e.g., with SCREAM) and reasoning about novel situations (e.g., with abduction) is not well understood. This is a fertile area for additional cognitive science and neuroscience experiments.

Second, BICA systems for any form of analysis application require large amounts of knowledge including previous cases/situations, linguistic rules, domain specific expertise and common sense. We believe the form of knowledge representation in SCREAM (i.e., cognitive models) is well suited for automated knowledge capture from raw text such as previously written analyst reports, textbooks, case studies, and Wikipedia. SCREAM does not require relations/associations to have explicit names and it tolerates noise and inconsistency. We have done some initial experiments in processing of analyst reports with AeroText and STANLEY to understand the linguistic challenges in automatically generating SCREAM cognitive models. This will be a major focus for future research.

Finally, we plan to explore the application of BICA to interacting loops of foraging and sensemaking for open source intelligence. Analysts still spend large amounts of time searching for information. The control of searching and filtering is a complex cognitive task that will drive advances in BICA research.

References

- Bauer, T., "Text Analysis and Dimensionality Reduction for Automatically Populating a Cognitive Model Framework" in Forsythe, C., Bernard, M., Goldsmith, T.eds. 2005, *Cognitive Systems: Human Cognitive Models In Systems Design*, Lawrence Erlbaum Associates
- Bodnar, J., "Making Sense of Massive Data by Hypothesis Testing", International Conference on Intelligence Analysis, McLean, Virginia, 2005
- Sandia Report SAND2006-7812 "Simulating Behavior for National Security Human Interactions" Bernard, M., Dec. 2006
- Senglaub, M., "Automated Decision Support in a Complex Information Space", Int. C2 Research and Technology Symposium, June 17-19, 2008