COGENT: Cognitive Agent for Cogent Analysis

Gheorghe Tecuci, Dorin Marcu, Mihai Boicu, David Schum
Learning Agents Center, George Mason University, Fairfax, VA 22030
{tecuci, dmarcu, mboicu}@gmu.edu, dschum398@earthlink.net

Abstract
Timely, relevant, and accurate intelligence analysis is critical to the national security, but it is astonishingly complex. This paper provides an intuitive overview of Cogent, a cognitive assistant that facilitates a synergistic integration of analyst’s imaginative reasoning with agent’s critical reasoning to draw defensible and persuasive conclusions from masses of evidence, in a world that is changing all the time. It presents Cogent’s design goals characterizing a new generation of structured analytical tools, introduces the evidence-based analysis concepts on which it is grounded, illustrates a sample session with its current version, and summarizes the cognitive assistance provided to its user.

Introduction
Intelligence analysts face the stunningly complex task of drawing defensible and persuasive conclusions from masses of evidence, in a world that is changing all the time (Schum, 1987; Tecuci et al., 2015). Because this requires both imaginative and critical reasoning, we believe that it can be best performed through the mixed-initiative integration of human imagination and computer knowledge-based reasoning (Tecuci et al., 2007a).

For several years we have worked on a computational theory of intelligence analysis (Tecuci et al., 2011) and, on this basis, we have developed a sequence of increasingly more practical cognitive assistants for the intelligence analysis education and practice. The first of these systems, Disciple-LTA (Tecuci et al., 2005; 2007b; 2008), is a unique and complex cognitive assistant that integrates powerful capabilities for analytic assistance, learning and tutoring, and is at the basis of the other developed systems.

TIACRITIS (Teaching intelligence analysts critical thinking skills) was develop primarily for teaching intelligence analysis and was experimentally used in many IC and DOD organizations (Tecuci et al., 2011). While praising its solid theoretical framework and deep evidentiary knowledge, the analysts desired a simplified interface and interaction.

Cogent Design Goals
The following are the main design goals of Cogent, aimed at developing a very powerful and easy to use cognitive assistant for a typical intelligence analyst:

- **Ease of use** (based on simplified interface, mixed-initiative interaction, simple yet rigorous logic and probability system).
- **Cogent analysis** (through an embedded computational theory enabling analysis checking, detection and mitigation of cognitive biases, and confidence evaluation).
- **Learning** (through the capturing of context-dependent hypotheses and argument patterns, from the analysts).
- **Rapid analysis** (through automatic reuse of learned analytical expertise, drill-down at various levels of detail, evidence-monitoring, and dynamic report generation).
- **Collaborative** (through multiple forms of collaborative analysis, information sharing, and analyses sharing).

Some of these goals have already been achieved in the current version of Cogent, as discussed in the following.
Evidence-based Analysis Concepts for Cogent

This section introduces the evidence-based analysis concepts that need to be understood in order to use Cogent. It presents the analysis framework, alternative assessment scales, and the evidence credentials (believability or credibility, relevance, and inferential force or weight) and their use in the direct assessment of hypotheses. It then discusses the need of the argumentation for assessing complex hypothesis, and presents the structure of the argumentation.

Analysis Framework

The analysis framework implemented in Cogent is illustrated in Figure 1: From the questions we ask about a situation of interest we generate alternative hypotheses as possible answers, we use these hypotheses to guide the collection of evidence, and we use the found evidence to test the hypotheses and determine the most probable answer.

Hypothesis Assessment Scales

Hypothesis assessment is necessarily probabilistic in nature because our evidence is always incomplete (we can look for more, if we have time), usually inconclusive (it is consistent with the truth of more than one hypothesis), frequently ambiguous (we cannot always determine exactly what the evidence is telling us), commonly dissonant (some of it favors one hypothesis but other evidence favors other hypotheses), and with various degrees of believability (Schum, 2001; Tecuci et al., 2015).

Cogent makes possible the use of different assessment scales, allowing the analysts to select the one that is the most natural to them. Figure 3 shows the three alternative assessment scales that have already been implemented in Cogent. All of them are based on the same system of Baconian probabilities (Cohen, 1977; 1989) with Fuzzy qualifiers (Zadeh, 1983), where the assessment values are on an ordered positive scale. This is illustrated below with the “Probability” scale:

lack of support<likely<very likely<almost certain<certain

In this case, there may be a lack of support from the available evidence to the considered hypothesis, or the evidence may indicate some level of support (e.g., likely).

As indicated in Figure 3, each symbolic value (e.g., “very likely”) is abbreviated in the Cogent analysis whiteboard (i.e., “VL”) to reduce space usage and visualize larger argumentations.

Evidence-based Hypothesis Assessment

One can directly assess a hypothesis based on an item of evidence by assessing the credentials of evidence, as illustrated in Figure 4, where the probability of the hypothesis from the top (very likely) is assessed based on the evidence from the bottom.

One first assesses the believability or credibility of evidence by answering the question, “What is the probability
that what the evidence is telling us is true?” Let us assume this to be “almost certain” (AC), as in Figure 4.

![Figure 4: The credentials of evidence.](image)

Next one assesses the relevance of the item of evidence to the hypothesis by answering the question, “What would be the probability of the hypothesis if the evidence were true?” That is, assuming that “Chemicals of the purity required for sarin-based weapons are readily accessible at low visibility for plausibly legitimate business purposes” what is the probability that “It is safe [i.e., not suspicious] for a legitimate chemical business, such as that of Aum Shinrikyo, to acquire chemicals for sarin-based weapons?” Let’s say that it is very likely (VL).

Finally, Cogent determines the inferential force or weight of the evidence on the hypothesis. The inferential force answers the question, “What is the probability of the hypothesis, based only on this evidence?” Obviously, an irrelevant item of evidence will have no inferential force, and will not convince us that the hypothesis is true. An item of evidence that is not believable will have no inferential force either. Only an item of evidence that is both very relevant and very believable may convince us that the hypothesis is true. Consistent with both the Bayesian and the Fuzzy min/max probability combination rules, Cogent determines the inferential force of an item of evidence on a hypothesis as the minimum between its believability and its relevance which, in this illustration, is very likely (VL).

In general, there may be several items of evidence that are relevant to a given hypothesis, some favoring it and some disfavoring it, and each with a specific believability, relevance, and inferential force. Figure 5 illustrates a situation where there are two evidence items favoring Hypothesis1 (Evidence1 and Evidence2), and one evidence item disfavoring it (Evidence3).

The analyst needs to estimate the believability and relevance of each of these three items (e.g., believability almost certain and relevance very likely for Evidence1). Then Cogent automatically determines their individual and collective inferential forces. The individual inferential forces are determined as illustrated in Figure 4 and discussed above. For example, the inferential force of Evidence1 on Hypothesis1 is min(AC, VL) = VL.

The inferential force of all the favoring evidence (e.g., Evidence1 and Evidence2) on a hypothesis (i.e., Hypothesis1) is the maximum of the inferential forces of the evidence items (i.e., max(VL, AC) = AC). The inferential force of the disfavoring evidence (e.g., Evidence3) is computed in a similar way (i.e., min(L, VL) = L).

The Baconian probability system (Cohen, 1977; 1989) requires to consider either H or not H as probably true, but not both at the same time. Therefore, if the disfavoring evidence has higher inferential force than the favoring evidence, then this will be considered as lack of support for H. In this example, the inferential force of the favoring evidence (i.e., Evidence1 and Evidence2) is almost certain (AC), as specified inside the left (green) square under Hypothesis1. The inferential force of the disfavoring evidence (Evidence3) is likely (L), as specified inside the right (pink) square under it. Therefore Cogent automatically determines the inferential force of all these items of evidence on Hypothesis1 as very likely (VL), by balancing the inferential force of the favoring items (AC) with that of the disfavoring item (L), based on the function from Figure 6.

![Figure 5: Hypothesis assessment based on multiple items of evidence.](image)

![Figure 6: On balance function in Cogent.](image)
In general, as indicated in the right and upper side of Figure 6, if the inferential force of the disfavoring evidence is higher than or equal to that of the favoring evidence, then Cogent concludes that, based on all the available evidence, there is a lack of support for $H$. If, however, the inferential force the favoring evidence is strictly greater than that of the disfavoring evidence (and there is some force of the disfavoring evidence), then the probability of $H$ is lowered, based on the inferential force of the disfavoring evidence (see the left and lower side of Figure 6).

Assessing Complex Hypotheses

The previous section presented a simple way of directly assessing a hypothesis based on evidence. This works well when it is obvious that the evidence favors or disfavors the hypothesis, as was illustrated in Figure 4. But it does not work well for a complex hypothesis such as “Aum Shinrikyo has sarin-based weapons” from the bottom of Figure 2. Indeed, how confident would you be in assessing the relevance of the evidence from Figure 4 to this hypothesis?

And there is an additional question “How can you find evidence to assess such complex hypotheses?”

Solutions to both these difficulties are provided by Cogent which guides the analyst in developing an argumentation structure that successively reduces a complex hypothesis to simpler and simpler hypotheses, down to the level of very simple hypotheses such as that from the top of Figure 4. Then the analyst has to search for evidence that is relevant to these very simple hypotheses, and to assess their believability and relevance, as was illustrated in the previous section. Once this is done, Cogent automatically composes these assessments, from bottom up, based on the logic embedded in the argumentation, finally assessing the probability of the top-level hypothesis.

The next section presents the structure of the argumentation in Cogent.

Argumentation Structure

Figure 7 provides an example of an argumentation for the hypothesis “Aum Shinrikyo has sarin-based weapons.” In general, a hypothesis is assessed by considering both favoring arguments (under the left, green square) and disfavoring arguments (under the right, pink square). There may be one argument of each type, several, or none. Each argument reduces the top-hypothesis to simpler hypotheses for which favoring and disfavoring arguments are again considered.

For example, there are two favoring arguments and no disfavoring argument for the top hypothesis “Aum Shinrikyo has sarin-based weapons.” The first favoring argument states that, if Aum Shinrikyo has expertise, production material, and funds, then it is very likely (VL) that it develops sarin-based weapons. Further down: If Aum Shinrikyo has created a legitimate chemical business (“legitimate business”), and if it is safe for such a legitimate chemical business to acquire chemicals for sarin-based weapons (“safe acquisition”), then it is almost certain (AC) that Aum Shinrikyo has production-material for sarin-based weapons.

Let us first consider the “develops” hypothesis. Its favoring argument is the following one: If Aum Shinrikyo has expertise, production material, and funds, then it is very likely (VL) that it develops sarin-based weapons. Further down: If Aum Shinrikyo has created a legitimate chemical business (“legitimate business”), and if it is safe for such a legitimate chemical business to acquire chemicals for sarin-based weapons (“safe acquisition”), then it is almost certain (AC) that Aum Shinrikyo has production-material for sarin-based weapons.

Now it should be quite easy to search for evidence which is relevant to these simpler hypotheses (i.e., “legitimate business” and “safe acquisition”). For example, it would be enough to check whether Aum has created any legitimate chemical business. One may easily discover that it has, in fact, created two dummy companies, Beck and Belle Epoch – both run by Niimi – under Hasegawa Chemical, an already existing Aum shell company. Similarly, one may easily find the other evidence from the bottom of Figure 7 (whose description is in Figure 4), just by checking the Japanese regulations on the buying of chemicals.
Thus, the reduction of complex hypotheses (such as “Aum Shinrikyo has sarin-based weapons”) to simpler ones (e.g., “legitimate business” and “safe acquisition”), guides the analyst to collect relevant evidence. Once such evidence is found, the probabilities of the simplest hypotheses can be determined, as discussed in the previous sections. Then the probabilities of the upper-level hypotheses are automatically determined by Cogent by using the logic embedded in the argumentation, in accordance with Baconian and Fuzzy min/max probability combination rules: AND (\&) structures in the argumentation require the use of the minimum (min) function, while OR (alternative) structures require the use of the maximum (max) function.

For example, based on E1 Chemicals business and E2 Chemicals acquisition, the probabilities of the “legitimate business” and “safe acquisition” are determined as almost certain (AC) and very likely (VL), respectively (see the bottom of Figure 7). Then the probability of “production material” is automatically determined as min(min(AC, VL), AC) = VL because “production material” requires both “legitimate business” and “safe acquisition,” and their combined relevance is almost certain (AC).

However, the analysts do not always find relevant evidence, or they may not even have time to look for it. In such cases they may make assumptions with respect to the probabilities of some sub-hypotheses. Let us consider, for example, the “buys sarin-based weapons” hypothesis. It has one favoring argument with relevance very likely: If Aum has funds then it is very likely that it buys sarin-based weapons. It also has one disfavoring argument with relevance certain: If there is no seller then it is certain that Aum does not buy sarin-based weapons. The analyst may now assume that it is “almost certain” that Aum has funds. This probability combines with the “very likely” relevance to produce a probability of “very likely” for the favoring argument (the minimum of “almost certain” and “very likely”). Similarly, if the analyst makes the assumption that “almost certain” there is no seller, then Cogent infers that the probability of the disfavoring argument is “almost certain.” Since this is greater than “very likely,” the probability of the favoring argument, Cogent concludes that there is a “lack of support” for the hypothesis that Aum Shinrikyo buys sarin-based weapons (see Figure 7).

Context-dependent Analysis

Notice in the argumentation structure from Figure 7 that some words, such as Aum Shinrikyo, appear in blue and are underlined. This is because they are part of Cogent’s knowledge base, and are recognized by it.

Notice also that only the top-level hypothesis is completely specified, while the sub-hypotheses are abstracted, and are understood in the context of their upper level hypotheses. For example, “develops sarin-based weapons” is understood as “Aum Shinrikyo develops sarin-based weapons.” Similarly, “production material” is understood as “Aum Shinrikyo has production material to develop sarin-based weapons.” Also, “legitimate business” is understood as “Aum Shinrikyo has a legitimate business for obtaining production material to develop sarin-based weapons.”

Context-dependent analysis enables a very succinct representation of an argumentation, helping the analyst to visualize a larger portion of it in the Cogent whiteboard. At the same time, Cogent is aware of the complete representation of each hypothesis which is necessary for the learning and reuse of analytic expertise, as will be discussed in a follow-on section.

A Sample Session with Cogent

Having presented the evidence-based analysis concepts, this section illustrates a sample session with the current version of Cogent, to provide an intuitive understanding of its use by a typical analyst.

Starting the Analysis

The analyst starts by defining, in natural language, the situation of interest, the intelligence question(s), and the alternative hypotheses, as was illustrated in Figure 2. While editing a statement, the analyst may select a phrase (e.g., sarin-based weapons) and ask Cogent to introduce it into its knowledge base as a new entity. In general, entities in a hypothesis are those phrases that can be replaced with other phrases to obtain similar hypotheses. Consider, for example, the hypothesis “Aum Shinrikyo has sarin-based weapons.” One may replace Aum Shinrikyo and sarin-based weapons with Al Qaeda and botulinum-based weapons, respectively, to obtain a similar hypothesis: “Al Qaeda has botulinum-based weapons.” When the analyst starts typing a word, Cogent proposes its completion with known entities. Recognizing entities enables Cogent to learn and reuse not only hypotheses patterns from specific hypotheses, such as “<agent> has <weapon>,” but also argument patterns, as will be discussed in a follow-on section.

Developing the Argumentation

Once a top-level hypothesis is defined, the analyst interacts with Cogent to develop an argumentation, such as the one from Figure 7, through easy and intuitive operations, for example by dragging and dropping building blocks (e.g., a new hypothesis) under, above or next to existing hypotheses, and by updating them.

Then the analyst looks for evidence relevant to the simplest hypotheses and attaches it to them. This process is illustrated in Figure 8. The analyst selects a paragraph from a document which represents favoring evidence for the “legitimate business” hypothesis. Then it drags and drops it
on the left (green) square under the hypothesis. Similarly, disfavoring evidence is dropped on the right (pink) square.

As a result, Cogent automatically defines the item of evidence in the Evidence Assistant (see the upper right side of Figure 8), and attaches it to the hypothesis (see the bottom left side of Figure 8). The automatically generated evidence name in the Evidence Assistant is selected in case the analyst desires to replace it with a more suggestive one.

Figure 8: Attaching evidence to hypothesis.

Notice that the believability and relevance of the newly defined item of evidence are NS (Not Set). Once the analyst assesses them, Cogent automatically assesses the probability of the elementary hypothesis and of the upper-level ones, based on the defined structure of the argumentation, as was discussed in the previous section.

The evidence may be attached not only to an elementary hypothesis, but also to a higher level one, in which case it represents an additional (favoring or disfavoring) argument for that hypothesis.

The analyst has to also make assumptions with respect to the probabilities of the elementary hypotheses that have no evidence. Alternative assumptions correspond to alternative what if scenarios.

While argument development may seem a laborious process, it is greatly facilitated by the reuse of learned patterns, as will be discussed in the next section.

**Pattern Learning and Reuse**

Once the analysis is completed, the analyst may wish to request Cogent to learn hypotheses and argument patterns to be reused in future analyses. This is done by simply right-clicking on a hypothesis, such as “Aum Shinrikyo has sarin-based weapons” from the top of Figure 7, and selecting “Learn.” As a result, Cogent learns both a hypothesis pattern (“[actor] has [weapon]”), and two argument patterns (one for each of the two favoring arguments), as illustrated in Figure 9.

The patterns are obtained by using Cogent’s ontology as a generalization hierarchy, where individual entities from the analysis in Figure 7 (e.g., sarin-based weapons) are replaced with more general concepts from the ontology (Tecuci et al., 2007b). The analyst may change the pattern by clicking on a concept (e.g., [weapon]) and selecting another concept (e.g., [WMD]) from a list presented by Cogent.

Attached with each pattern, Cogent also maintains the example from which it was learned, which will be used to refine the pattern when new examples are encountered (e.g., when the pattern is reused).

Figure 9: Learned patterns.

Figure 10 illustrates the reuse of the learned hypothesis pattern “[actor] has [WMD] (...).” The analyst simply drags it from the Argument assistant, and drops it on “Question(Q): Is Aum Shinrikyo a threat?” Then the analyst clicks on each concept in the hypothesis pattern (e.g., [WMD]) and selects an entity to instantiate it from a list provided by Cogent, or defines a new entity (e.g., “botul-
num-based weapons”). As a result, Cogent generates the analysis from the right hand side of Figure 10. It also automatically introduces the user-defined entity (botulinum-based weapons) in its ontology, as an instance of WMD.

Notice that here Cogent has applied both argument patterns from Figure 9 because the analyst has reused the hypothesis pattern. But the analyst may also reuse only one argument, or even only a part of it, by dragging that argument (or part of it) and dropping it on one of the squares under a hypothesis. To complete the analysis of the new hypothesis, the analyst needs to look for evidence for the leaf hypotheses, and to analyze it.

Notice in this illustration how much faster the analysis is completed through the reuse of learned patterns.

Cogent is distributed with two manuals. Getting Started with Cogent is a very short manual that introduces the basic elements of the theory and the operation of Cogent through a simple analysis example, enabling the analysts to immediately start using the system. Cogent: Operations presents the argumentation structure in more details, and describes all the available operations.

**Cognitive Assistance**

The computational theory which is embedded into Cogent guides the analyst through a systematic analysis process which synergistically integrates analyst’s imaginative reasoning with agent’s critical reasoning. The analyst imagines the questions to be asked and hypothesizes their possible answers. Cogent helps developing the arguments by reusing previously learned patterns, and guides the evidence collection by the analyst. The analyst assesses the believability and relevance of the found evidence, and Cogent determines its inferential force, as well as the probabilities of the upper-level hypotheses in the argumentation. This jointly-developed analysis makes very clear the argumentation logic, what evidence was used and how, what is not known, and what assumptions have been made. It can be shared with other analysts, subjected to critical analysis, and correspondingly improved. As a result, this systematic and theoretically justified process leads to the development of defensible and persuasive conclusions.

Cogent also enables rapid analysis, not only through the reuse of patterns, but also through a drill-down process where a hypothesis may be decomposed to different levels of detail, depending on the available time. It facilitates the analysis of what-if scenarios, where the analyst may make various assumptions, Cogent automatically determining their influence on the analytic conclusion. Cogent also makes possible the rapid updating of the analysis based on new evidence and assumptions.

![Figure 10: Argument development through pattern reuse.](image-url)
Beyond Intelligence Analysis

This paper has focused on the use of Cogent for intelligence analysis. However, the presented analysis framework is applicable to evidence-based reasoning in a wide variety of domains, including cyber security, forensics, medicine, law, natural science (physics, chemistry, etc.), education, history, archaeology, and many others.

In cyber security, a Cogent-like agent may be integrated into a Cybersecurity Operations Center (CSOC) to automate the investigation of intrusion alerts from a variety of intrusion detection devices, integrating multiple detection techniques with automated network forensics, to significantly increase the probability of accurately detecting intrusion activity while drastically reducing the workload of the CSOC operators.

In science education, a Cogent-like agent may facilitate inquiry-based teaching and learning, engaging students in understanding, extending, creating, critiquing, and debating evidence-based scientific arguments in real-life scientific investigations, giving the students opportunities to exercise imagination and creativity, and develop critical scientific practices, particularly: asking questions; constructing explanations; engaging in argument from evidence; and obtaining, evaluating, and communicating explanations. (National Research Council, 2012, p.3).

Conclusions and Future Work

This paper has presented a brief overview of Cogent, emphasizing the user experience of a typical analyst.

Future work will continue the development of Cogent in accordance with the presented design goals. In particular, we plan to focus on the development of pattern learning for automatic reuse, on collaborative analysis, and on the integration of advanced analytics (e.g., confidence assessment, automatic detection of key evidence and assumption).

We also plan to further develop the computational theory of intelligence analysis which is at the basis of Cogent, and to generalize it into a computational theory of evidence-based reasoning. On this basis, we plan to develop cognitive assistants for other evidence-based reasoning domains, particularly cyber security and education.

Acknowledgements

This research was partially supported by George Mason University, the Intelligence Community, and the Department of Defense. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of any agency of the U.S. Government.

References


