In *Evidence Dialogues*, 2023, https://evidencedialogues.wordpress.com/

How Intelligent Is Artificial Intelligence?

Gheorghe Tecuci

Learning Agents Center and Computer Science Department, School of Computing, George Mason University, Fairfax, VA 22030, USA; Former Chair of Artificial Intelligence, Center for Strategic Leadership, US Army War College, Carlisle, PA, USA; Member of the Romanian Academy, tecuci@gmu.edu, http://lac.gmu.edu

Introduction

Following the writing of my latest books on critical thinking in intelligence analysis and law, I was interviewed by Dr. Yvonne McDermott Rees (Professor of Law at Swansea University in the UK), from *Evidence Dialogues* (https://evidencedialogues.wordpress.com/). She asked me about ChatGPT and how this and other AI systems could be used in Law. These were, in essence, my answers.

Chat Generative Pre-training Transformer (ChatGPT)

People look at the very impressive accomplishments of AI, such as, Deep Blue (the IBM chess program that defeated world champion Gary Kasparov), AlphaGo that plays better Go than any human, IBM's Watson that defeated the best human players at Jeopardy, and attribute to AI systems super-intelligence abilities when, in fact, these are just computer programs that implement very sophisticated algorithms.

The latest of these "AI miracles" is ChatGPT. So impressive is its ability to answer complex questions in natural language, that many people forget its limits as a tool. Figure 1 shows its overall architecture.

ChatGPT ingested (i.e., represented and integrated internally) what was posted on the Internet, and combines the information related to the asked question to generate a well-documented answer. As a result, its answer is a kind of average of the information posted on the Internet.

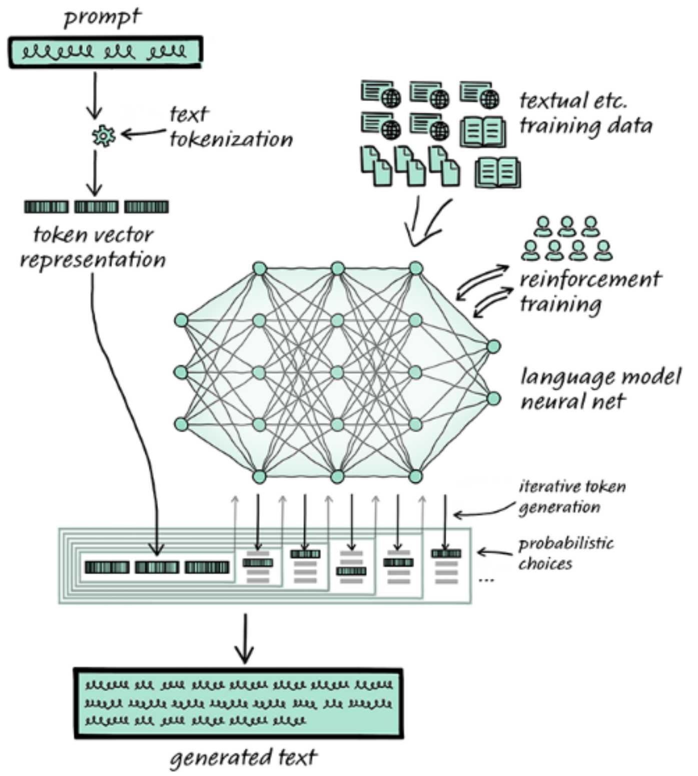

Figure 1. Overall architecture of ChatGPT (from https://writings.stephenwolfram.com/2023/01/wolframalpha-as-the-way-to-bring-computational-knowledge-superpowers-to-chatgpt/)

But if much of the information on a topic is wrong, its answer will also be wrong. The answer is also somewhat dated, because it takes time to represent the massive and continuously growing amounts of information available on the Internet. For example, ChatGPT 4 has ingested information up to 2021.

But remember that ChatGPT is just

a tool. It uses a very sophisticated algorithm and a deep neural network to learn and generate answers. To put it very simply, the answer is generated by using a highly intricate formula that operates on numerical values corresponding to the input, resulting in a set of numerical values that represent the output. ChatGPT has no "understanding" of *why* this is the answer, and therefore cannot explain it. This is the main drawback of neural networks, in general. Additionally, ChatGPT is not (yet) a sophisticated problem solver, and cannot correctly answer questions that require complex (multi-step) reasoning, such as design or planning.

If people use ChatGPT to generate answers, then they should exercise their critical reasoning to check the generated answers. But it is much simpler to check whether an answer is correct or not, than to find it in the first place, and that is really the power of a tool like ChatGPT.

Cognitive Agent for Cogent Analysis (Cogent)

There are other approaches to AI that are able to "show their work" and explain their reasoning. For example, I have a career-long interest in developing instructable computer systems. These are systems that would be taught rather than programmed. Thus, a doctor, an engineer, or an intelligence analyst would be able to teach these agents to become their assistants, in a way that is similar to how they would teach a collaborator or student, through problem solving examples, and by verifying and correcting their problem-solving attempts. In time, these systems will learn to simulate the reasoning of their instructor, and therefore I called them Disciple agents. This work started in with my Ph.D. research, where I developed the first prototype Disciple agent that was taught how to manufacture loudspeakers (Tecuci, 1988). Since then, my students and I have demonstrated instructable agents in a wide variety of domains, including education (Tecuci 1998; Tecuci and Keeling, 1999; Tecuci 2021), engineering planning (Tecuci et al., 2000; Tecuci et al., 2016a), course of action critiquing (Tecuci et al., 2001; Tecuci et al., 2016a), center of gravity analysis (Tecuci et al., 2002; 2008a; 2016a), emergency response planning (Tecuci et al., 2008b; 2016a), intelligence analysis (Tecuci et al., 2016b; 2018a; Tecuci 2023a; 2023b; Tecuci and Schum, 2023a; 2023b), cybersecurity (Tecuci et al., 2018b; 2019), automatic sensemaking (Tecuci et al., 2020), knowledge discovery in agriculture (Tecuci et al, 2021), and physics (Tecuci, 2021).

Let me briefly explain this approach, as implemented in the latest Disciple-type system, the Cognitive agent for cogent analysis, or Cogent (Tecuci et al., 2018a). You demonstrate to Cogent how to solve a specific problem, and Cogent learns rules by generalizing your individual reasoning steps. Then it uses these rules to solve similar problems. You review and correct its mistakes, and Cogent refines the rules and learns additional ones. As a result, Cogent incrementally learns your problem-solving expertise. Figure 2 illustrates this process in the context of an investigative problem.

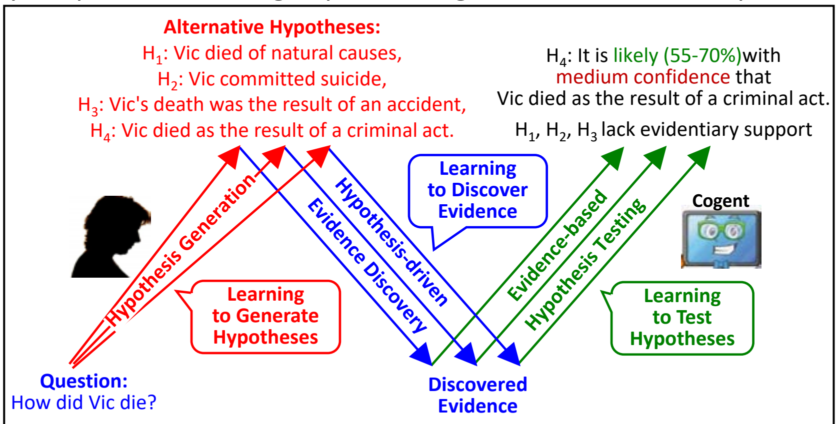

Figure 2. Cogent teaching and learning.

Hypothesis Generation

Imagine that you are a detective called to investigate the death of Vic, whose body was discovered on the floor of his garage. You ask yourself: How did Vic die?

Through *abductive (imaginative) reasoning* that shows that something is *possibly* true (Peirce, 1988; 1901), you imagine the following answers:

A_1: Vic died of natural causes,
A_2: Vic committed suicide,
A_3: Vic's death was the result of an accident,
A_4: Vic died as the result of a criminal act.

The imagined answers are the hypotheses to be tested. To determine which of these hypotheses (if any) is true, you need *evidence*.

Evidence Discovery

To discover evidence, you put each hypothesis to work by using *deductive reasoning* that shows that something is *necessarily* true. For each hypothesis, you ask:

What evidence would favor or disfavor the hypothesis?

To answer it, you decompose the hypothesis into simpler and simpler sub-hypotheses, until these sub-sub-hypotheses are simple enough to point directly to the evidence that favors it or that disfavors it.

Let us consider the hypothesis A_4: Vic died as the result of a criminal act. To prove first degree murder, the U.S. law requires proving the following sub-hypotheses (Schum and Tecuci, 2023):

- A_4^1 Vic was unlawfully killed.
- A_4^2 The defendant was the one who killed Vic.
- A_4^3 The defendant intended to kill Vic [it as not an accident or self-defense].
- A_4^4 The defendant intended to kill Vic beforehand [Premeditation].

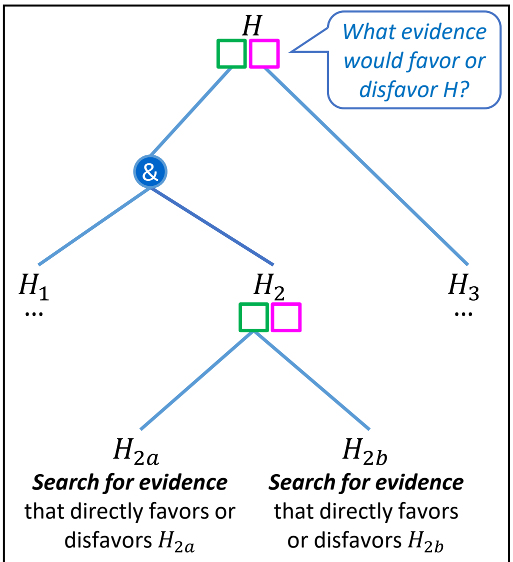

Figure 3. Hypothesis in search of evidence.

Each of these sub-hypotheses can be further decomposed, through both favoring and disfavoring arguments, as illustrated in Figure 3 with the hypothesis H:

- H would be true if both H_1 and H_2 were true.
- Then H_2 would be true if either H_{2a} or H_{2b} were true.
- Searching for evidence relevant to H_{2a} we discover the favoring evidence E_w^*.
- Searching for evidence relevant to H_{2b} we discover the favoring evidence E_x^* and the disfavoring evidence E_y^*.

We call this process *hypothesis in search of evidence*. It involves *deductive reasoning* that shows that something is *necessarily* true.

Hypothesis Testing

Through *inductive reasoning,* that shows that something is *probably* true, one tests the hypotheses. First one develops a Wigmorean argumentation (WigNet) as in Figure 4 (Wigmore, 1913; 1937; Schum and Tecuci 2023; Tecuci 2023a; Tecuci and Schum, 2023a; 2023b).

Hypothesis testing is necessarily probabilistic in nature because of five characteristics of evidence: Our evidence is always *incomplete*, no matter how much we have; It is commonly *inconclusive* in the sense that it is consistent with the truth of more than one hypothesis; It is frequently *ambiguous*, we cannot always say what the evidence is telling us; It is *dissonant* in most situations, some evidence favoring one hypothesis but other evidence favors other hypotheses; It comes from sources having any possible gradation of *credibility* short of perfection.

However, there is little consensus about how uncertainty should be expressed, combined, and reported. Most of us have learned in school about the conventional system of probability (Kolmogorov, 1933) in which uncertainty is expressed by percentages, or odds and odds ratios, as the only way in which uncertainty can be expressed about evidence-based conclusions. Many of us will have taken courses in statistics in which the probability of events is estimated based on the relative frequency of their observed occurrence in the past. This approach involves events that are the result of replicable or repeatable processes that can be counted. However, the problem is that there are many cases in intelligence analysis, law, medicine, science, history, and other domains where we have uncertainty about certain past or future events, but we have nothing to count because these events are unique, singular or one-of-a-kind (Twining, 2003). If they happened in the past, they did so on just one occasion. If they will occur in future, they will do so on just one occasion. Because there is nothing to count, the assessments must be judgmental or subjective in nature. This also means that different people may assess these probabilities differently and arrive at different answers. The point here is very simple: All statistical reasoning is probabilistic in nature, but not all probabilistic reasoning is statistical in nature.

There are some very interesting but difficult issues concerning the extent to which the concepts and methods so useful in statistical analyses continue to apply in situations in which we have uncertainty but no statistics. There are many issues regarding the assessment, combination, and reporting of uncertainty in these non-statistical situations that are very important but frequently go unrecognized.

Table 1 categorizes some alternative views of probability. We refer to the views in the first column as enumerative because they assume that the probabilities are the result of counting. We refer to the views in the second column as epistemic since they assume that probabilities are based on some kind of knowledge, whatever form it may take. In short, probabilities are the result of informed judgments.

Table 1. Alternative views of probability.

Enumerative	**Epistemic**
Aleatory (Chances) Relative Frequency and Statistics Bayesian Statistics	Subjective Bayesian Belief Functions Baconian Probability Fuzzy Probability

Complementariness of Probability Views

Questions about uncertainty are naturally linked to views about probability. Some analysts may have studied probability quite extensively; many others will have little or no formal tutoring on the subject. Still others may dislike the sometimes complicated formulas that the study of probability often requires. *However, words and pictures - rather than mathematics - can express very useful ideas about probability and uncertainty.* One does not need a background in mathematics or statistics to reason with uncertainty and draw conclusions from evidence. Probabilistic judgments can be expressed numerically in several ways, and also in terms of words. Speaking of numerical judgments of probability, a very wise and devoted scholar, Professor Glenn Shafer has correctly noted (Shafer, 1988, pp. 5 - 9):

> *Probability is more about structuring arguments than it is about numbers. All probabilities rest upon arguments. If the arguments are faulty, the probabilities however determined, will make no sense.*

We add the same concern about verbal assessments of probability, such as very probable, probable, unlikely, and so on. If these arguments are not defensible, no one will take seriously any numerical or verbal assessments we make concerning the force or weight of our evidence.

Notice that, as summarized in Table 2, neither the Subjective Bayesian (Schum, 2001), nor the Belief Functions (Schafer, 1974), or the Fuzzy probability view (Negoita and Ralescu 1975; Zaheh, 1983) can account for the incompleteness of the coverage of evidence. The only view that can account for this incompleteness is the Baconian view (Schum, 2001; Tecuci 2023a; Tecuci and Schum, 2023a). Consider, for example, the scenario where three competing hypotheses on an adversary, H_1, H_2, and H_3, were analysed for a battlefield commander. A body of evidence was examined and the Bayesian, Belief Functions, and Fuzzy methods have been employed. Each method showed that, based on the existing evidence, H_3 has the highest probability (and very close to certainty):

Table *2*. Complementariness of probability views.

Evidence Characteristic	Subjective Bayes	Belief Functions	Baconian	Fuzzy
Incompleteness			☑	
Inconclusiveness	☑	☑	☑	☑
Ambiguity		☑		☑
Dissonance	☑	☑	☑	☑
Credibility	☑		☑	

$P(H_3)$ = 0.998   Belief(H_3) = 0.989   Fuzzy(H_3) = almost certain

As time passes and action is taken based on H_3, it becomes painfully clear to the battlefield commander that H_1 was the true hypothesis, not H_3. What went wrong - after all Subjective Bayes, Belief Functions, and Fuzzy is each a highly respected probability view? Then an analyst having knowledge of the Baconian probability view enters the debate and makes the following comments:

> *"It is true that your analysis rested on quite a bit of evidence. But how many relevant questions you can think of were not answered by the evidence you had? If you believed that these unanswered questions would supply evidence that also favored H_3, you were misleading yourself since you did not obtain any answers to them. It is now clear that the answers to these questions did not favor H_3, but favored H_1. So, the posterior probability you determined, by itself, is not a good indicator of the weight of evidence. What makes better sense is to say that the weight of evidence depends on the amount of favorable evidence you have and on how completely it covers matters you said were relevant. In your analysis you completely overlooked the inferential importance of the questions your existing evidence did not answer."*

The Bayesian, Belief Functions, and Fuzzy probability views all answer the question: *How strong is the evidence we do have about this hypothesis?*

But the Baconian view answers a different question: *How much evidence do we have about this hypothesis, and how many questions about it remain unanswered?*

Answering the Baconian question would have revealed that there are unanswered questions whose answers may make H_3 less likely. Thus, it would have helped to acknowledge that the answers to these unanswered questions may not favor H_3.

Cogent uses an integrated logic and probability view that uses the min/max probability combination rules common to the Fuzzy probability view and the Baconian probability view:

$$P(H_1 \text{ and } H_2) = \min\{P(H_1), P(H_2)\};$$

$$P(H_1 \text{ or } H_2) \;=\; \max\{P(H_1), P(H_2)\}.$$

Notice the probability scale table from the top-left of Figure 4. This is a slight refinement of the symbolic probability scale used in the U.S. Intelligence Community, where the wide interval "55-80% likely" was split into two intervals, "55-70% likely (L)" and "70-80% more than likely (ML)", in order to enable more precise assessments (ODNI, 2007).

Wigmorean Probabilistic Inference Networks (WigNets)

To assess the probability of the hypothesis H in Figure 4, one decomposes it into simpler hypotheses by considering both favoring arguments (supporting the truthfulness of H), under the left (green) square, and disfavoring arguments (supporting the falsehood of H), under the right (pink) square. Each argument is an independent strategy of showing that H is true or false, and is characterized by a specific relevance or strength. The argument consists either of a single hypothesis (e.g.,H_3) or a conjunction of hypotheses (e.g., H_1 & H_2). These sub-hypotheses are further decomposed through other arguments, leading to simpler and simpler hypotheses that can be more accurately assessed based on evidence.

Probability scale	
100%	certain (C)
95-99%	almost certain (AC)
80-95%	very likely (VL)
70-80%	more than likely (ML)
55-70%	likely (L)
50-55%	barely likely (BL)
0-50%	lacking support (LS)

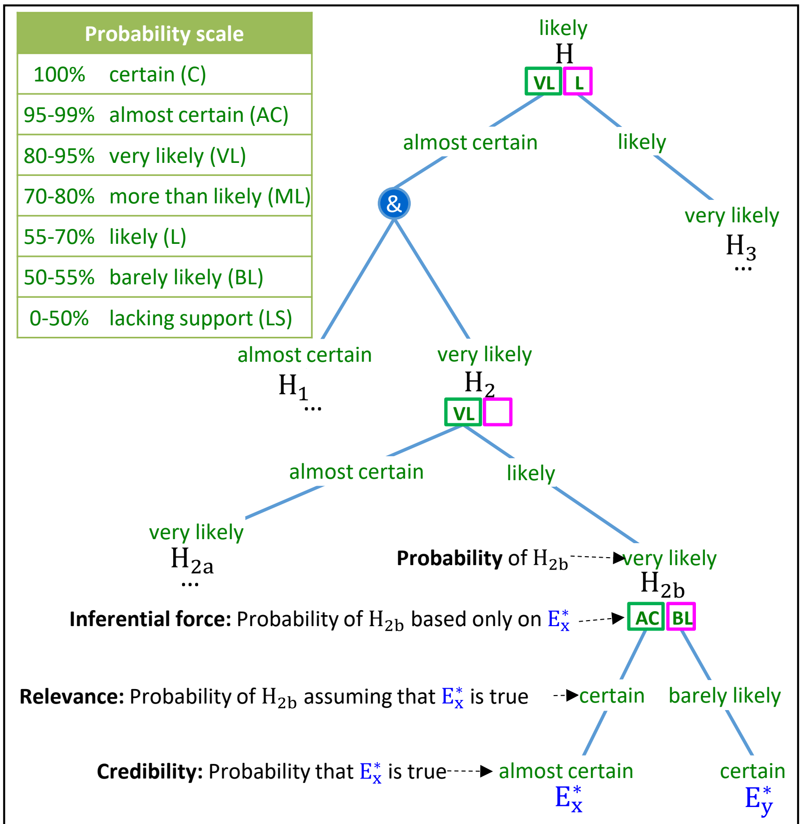

Figure 4. Wigmorean network for hypothesis testing.

Consider, for example, sub-sub-hypothesis H_{2b}. There are two items of evidence relevant to this hypothesis, the favoring item E^*_x, and the disfavoring item E^*_y. Each item of evidence has three credentials that need to be assessed: credibility, relevance, and inferential force.

The *credibility* of evidence answers the question:

What is the probability that the evidence is true?

The *relevance* of evidence to a hypothesis answers the question:

What would the probability of the hypothesis be if the evidence were true?

The *inferential force* of the evidence on the hypothesis that answers the question:

What is the probability of the hypothesis, based only on this evidence?

The inferential force of an evidence item is determined as the smaller between its credibility and its relevance. Indeed, an evidence item that is not credible would not convince us that the hypothesis is true, no matter how relevant the provided information is. Therefore, the inferential force in this circumstance would be low. Similarly, it is not enough for the evidence item to be credible, if the information provided is not relevant to the hypothesis. The inferential force will be high only if the evidence item is both highly relevant and credible.

The probabilities of the upper-level hypotheses are assessed from the probabilities of the leaf hypotheses based on the argumentation logic, using the simple min/max probability composition rules common to the Baconian and Fuzzy probability views.

Table 3. On-balance function.

Inferential force of disfavoring evidence/arguments (columns); Inferential force of favoring evidence/arguments (rows)

H	lacking support	barely likely	likely	more than likely	very likely	almost certain	certain
lacking support	lacking support	lacking support	lacking support	lacking support	lacking support	lacking support	lacking support
barely likely	barely likely	lacking support	lacking support	lacking support	lacking support	lacking support	lacking support
likely	likely	barely likely	lacking support	lacking support	lacking support	lacking support	lacking support
more than likely	more than likely	likely	barely likely	lacking support	lacking support	lacking support	lacking support
very likely	very likely	more than likely	likely	barely likely	lacking support	lacking support	lacking support
almost certain	almost certain	very likely	more than likely	likely	barely likely	lacking support	lacking support
certain	certain	almost certain	very likely	more than likely	likely	barely likely	lacking support

The probability of the hypothesis H is determined by balancing the combined inferential force of the favoring evidence (almost certain), and the inferential force of the disfavoring evidence (likely). The Baconian probability view (Cohen, 1977; 1989) requires considering either **H** or **not H** as probably true, but not both at the same time. To assess a hypothesis that has both favoring and disfavoring evidence, such as hypothesis H in Figure 4, we have introduced an <u>on-balance</u> function shown in Table 3 that balances the inferential force of the favoring evidence with that of the disfavoring evidence. As indicated in the right and upper side of Table 3, if the inferential force of the disfavoring evidence is higher than or equal to that of the favoring evidence, then the hypothesis H is lacking support. If, however, the inferential force of 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 Table 3).

Consider, for example, the hypothesis H at the top of Figure 4:

Inferential force of favoring argument of H = very likely
Inferential force of disfavoring argument of H = likely
Probability of H is obtained from Table 3 as likely.

The Future of AI: Cognitive Augmentation

What I see in the future are human-machine systems that synergistically integrate their complementary capabilities. Computer systems are fast, rigorous, precise, explicit, and objective, but they lack common sense and the ability to deal with new situations. Humans are indeed slow, sloppy, forgetful, implicit, and subjective, but have common sense and intuition, and may find creative solutions in new situations (Turoff, 2007; Tecuci et al., 2007).

In the Cogent example just provided, the human investigator will imagine the hypotheses, Cogent will develop the routine (i.e., learned) parts of the argumentation and the human investigator will complete it with the novel parts. The human investigator will assess the credibility and relevance of the evidence items and Cogent will compute the probabilities of all the hypotheses.

Notice also the following features:

- The analysis makes very clear the logic, what evidence was used and how, what is not known, and what assumptions have been made. It can be shared with other users, subjected to critical review, and correspondingly improved. As a result, this systematic process leads to the development of defensible and persuasive conclusions.
- Rapid analysis, not only through the reuse of learned patterns, but also through a drill-down process where a hypothesis may be decomposed to different levels of detail, depending on the available time.
- Analysis of what-if scenarios, where the user may make various assumptions and the assistant automatically determines their influence on the analytic conclusion.
- Rapid updating of the analysis based on new (or revised) evidence and assumptions.

To give a very recent example, Dr. Steven Rieber of the Intelligence Advanced Research Projects Activity (IARPA) has just announced a very interesting and challenging research program, called REASON, an acronym for Rapid Explanation, Analysis, and Sourcing Online:

> "Decision makers rely on the Intelligence Community to help them understand a wide variety of complex issues. Intelligence analysts face numerous challenges in their efforts to produce high-quality analytic reports. One major challenge is finding all relevant evidence from an ever-growing collection of often uncertain and conflicting information drawn from classified and unclassified sources. A second challenge is making well-reasoned judgments in the face of uncertainty. REASON will develop technology that analysts can use to discover additional relevant evidence (including contrary evidence) and to identify strengths and weaknesses in reasoning.
>
> REASON is not designed to replace analysts, write complete reports, or to increase their workload. The technology will work within the analyst's current workflow. It will function in the same manner as an automated grammar checker but with a focus on evidence and reasoning.
>
> Performer teams will conduct research and development to build systems that will be evaluated by an independent testing and evaluation (T&E) team. Independent T&E will ensure that REASON is effective in helping analysts discover valuable evidence, identify strengths and weaknesses in reasoning, and produce higher quality reports." (https://www.iarpa.gov/research-programs/reason).

Conclusion

Elon Musk and Steve Wozniak recently called for a six-month pause in developing new AI tools more powerful than GPT-4 (Durden, 2023). They wrote:

> "Contemporary AI systems are now becoming human-competitive at general tasks, and we must ask ourselves: *Should* we let machines flood our information channels with propaganda and untruth? *Should* we automate away all the jobs, including the fulfilling ones? *Should* we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? *Should* we risk loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders. Powerful

> AI systems should be developed only once we are confident that their effects will be positive and their risks will be manageable." (futureoflife.org)

I could not disagree more strongly with this statement. Of course, any powerful tool, such as ChatGPT, may produce a lot of damage, if not used properly. So could nuclear power.

None of the artificial intelligence systems mentioned above (including ChatGPT) has the attributes unique to human intelligence, such as, creativity, intuition, wisdom, distinguishing between good and evil, consciousness, the capacity to empathize, love, and many others. But until now, human intelligence has no equal in the biological and artificial world from the point of view of versatility, evolutionary goals and intentions, reasoning skills, understanding and generating language, perception and response to sensory inputs, creating art and music, or writing literature and poems.

The scary, futuristic presentations of AI by the media (and now even by Elon Musk and Steve Wozniak) have no basis in reality. There is no competition between humans and AI robots on the horizon, and probably it is not even possible (AI 100, 2016; Lauchbury, 2017;).

References

AI 100 (2016). *One Hundred Year Study on Artificial Intelligence*, Report of the 2015 Study Panel, Stanford University, September 2016. https://ai100.stanford.edu/2016-report

Cohen, L. J. (1977). *The probable and the provable.* Oxford, UK: Clarendon Press.

Cohen, L. J. (1989). *An introduction to the philosophy of induction and probability*. Oxford, UK: Clarendon Press.

Durden T. (2023). Musk, Wozniak Call for Pause in Developing 'More Powerful' AI than GPT-4, *Zero Hedge*, March 29.

ODNI (2007). *Intelligence Community Directive 203*, Office of the Director of National Intelligence.

Kolmogorov, A. N. (1956). *Foundations of a Theory of Probability* (1933), 2nd English edition, Chelsea Publishing, New York, NY, pp. 3-4.

Lauchbury J., *A DARPA Perspective on Artificial Intelligence*, YouTube 2017, https://www.youtube.com/watch?v=-O01G3tSYpU

Negoita, C. V., and Ralescu, D. A. (1975). *Applications of Fuzzy Sets to Systems Analysis*. Wiley: New York.

Peirce, C. S. (1898). *Reasoning and the Logic of Things*, 1898, in Ketner, K. (ed.), Harvard University Press, Cambridge, MA, 1992.

Peirce, C. S. (1901). Abduction and Induction, in Buchler, J. (ed.) (1995). *Philosophical Writings of Peirce*, 150-156, Dover: New York, New York.

Schum, D. A. (2001). *The Evidential Foundations of Probabilistic Reasoning*. Evanston, IL: Northwestern University Press.

Schum D. A., and Tecuci G. (2023), *The Art and Science of the Judicial Proof* (to appear).

Shafer, G. (1976). *A Mathematical Theory of Evidence,* Princeton University Press, Princeton, NJ.

Shafer, G. (1988). *Combining AI and OR*. University of Kansas School of Business Working

Paper No. 195, April.

Tecuci, G. (1988). Disciple: A Theory, Methodology and System for Learning Expert Knowledge, *Thèse de Docteur en Science,* University of Paris-Sud (now Paris-Saclay University) http://lac.gmu.edu/publications/1988/TecuciG_PhD_Thesis.pdf

Tecuci, G. (1998). *Building Intelligent Agents: An Apprenticeship Multistrategy Learning Theory, Methodology, Tool and Case Studies*. London, England: Academic Press. http://lac.gmu.edu/publications/1998/TecuciG_Building_Intelligent_Agents/default.htm

Tecuci, G. and Keeling, H. (1999). Developing an Intelligent Educational Agent with Disciple, *International Journal of Artificial Intelligence in Education*, vol. 10, no. 3-4. http://lac.gmu.edu/publications/1999/TecuciG_Intelliget_Educational_Agent.pdf

Tecuci, G., Boicu, M., Wright, K., Lee, S. W., Marcu, D. and Bowman, M. (2000). A Tutoring Based Approach to the Development of Intelligent Agents, in Teodorescu, H.N., Mlynek, D., Kandel, A. and Zimmermann, H.J. (editors). *Intelligent Systems and Interfaces,* Kluwer Academic Press. http://lac.gmu.edu/publications/data/2000/2000_Disciple-Planning.pdf

Tecuci, G., Boicu, M., Bowman, M., Marcu, D., with commentary by Burke, M. (2001). An Innovative Application from the DARPA Knowledge Bases Programs: Rapid Development of a Course of Action Critiquer, *AI Magazine,* 22(2), 43-61. http://lac.gmu.edu/publications/2001/TecuciG_Disciple_COA_IAAI.pdf

Tecuci, G., Boicu, M., Marcu, D., Stanescu, B., Boicu, C., Comello, J. (2002). Training and Using Disciple Agents: A Case Study in the Military Center of Gravity Analysis Domain, in *AI Magazine,* 24(4), 51-68, AAAI Press, CA.

http://lac.gmu.edu/publications/2002/TecuciG_Disciple_COG_IAAI.pdf

Tecuci, G., Boicu, M., Cox M.T. (2007). Seven Aspects of Mixed-Initiative Reasoning: An Introduction to the Special Issue on Mixed-Initiative Assistants, *AI Magazine*, 28(2), pp.11-18, Summer. http://lac.gmu.edu/publications/2007/BoicuM_AIMagazine_Intro.pdf

Tecuci, G., Boicu, M., and Comello, J. (2008a). *Agent-Assisted Center of Gravity Analysis*, CD with Disciple-COG and Lecture Notes used in courses at the U.S. Army War College and Air War College, GMU Press, ISBN 978-0-615-23812-8. http://lac.gmu.edu/cog-book/

Tecuci, G., Boicu, M., Marcu, D., Barbulescu, M., Boicu, C., Le, V., Hajduk, T. (2008b). Teaching Virtual Experts for Multi-Domain Collaborative Planning, *Journal of Software*, Vol. 3, No. 3, pp. 38-59, March. http://lac.gmu.edu/publications/2008/TecuciG_Disciple_VE_JS.pdf

Tecuci, G., Marcu, D., Boicu, M., Schum, D. A. (2016a). *Knowledge Engineering: Building Cognitive Assistants for Evidence-based Reasoning*. New York, NY: Cambridge University Press. http://lac.gmu.edu/KEBook/

Tecuci, G., Schum, D. A., Marcu, D., Boicu, M. (2016b). *Intelligence Analysis as Discovery of Evidence, Hypotheses, and Arguments: Connecting the Dots*. New York, NY: Cambridge University Press. http://lac.gmu.edu/IABook/

Tecuci, G., Kaiser, L., Marcu, D., Uttamsingh, C., Boicu, M. (2018a). Evidence-based Reasoning in Intelligence Analysis: Structured Methodology and System, Special Issue on Evidence-based Reasoning and Applications, *Computing in Science and Engineering*, 20(6), pp.9-21, November/December. http://lac.gmu.edu/publications/2018/Cogent-in-CiSE-2018.pdf

Tecuci, G., Meckl, S., Marcu, D., Boicu, M. (2018b). Evidence-based Detection of Advanced Persistent Threats, Special Issue on Evidence-based Reasoning and Applications, *Computing in Science and*

Engineering, 20(6), pp.54-65, November/December. http://lac.gmu.edu/publications/2018/CAAPT-in-CiSE-2018.pdf

Tecuci, G., Meckl, S., Marcu, D., Boicu, M. (2019). Instructable Cognitive Agents for Autonomous Evidence-Based Reasoning, *Advances in Cognitive Systems*, Vol. 8, 2019. Also in *Proceedings of the Seventh Annual Conference on Advances in Cognitive Systems*, Technical Report Number COLAB^2-TR-4, pp.183-204, August 2-5, 2019, Massachusetts Institute of Technology, Cambridge, MA. http://lac.gmu.edu/publications/2019/Automated_EBR.pdf

Tecuci, G., Marcu, D., Boicu, M., Kaiser, L. (2020). Instructable Cognitive Agent to Perform Sensemaking in Intelligence, Surveillance and Reconnaissance. T*he Eight Annual Conference on Advances in Cognitive Systems*, August 8-10, 2020. http://lac.gmu.edu/publications/2020/ACS_2020_Paper.pdf

Tecuci, G., Marcu, D., Mirsky, S., Robertson, A. (2021). Prediction-Driven Knowledge Discovery from Data and Prior Knowledge, *The 2021 AAAI Fall Symposium "Science-Guided AI,"* Arlington, VA, November 4-6. http://lac.gmu.edu/publications/2021/Tecuci-KDDK-2021.pdf

Tecuci, G. (2021). *Inquiry-based Teaching of Critical Thinking Skills in Science with sInvestigator*, Learning Agents Center Report, George Mason University, January. http://lac.gmu.edu/publications/2021/sInvestigator-Report.pdf

Tecuci G. (2023a), *Critical Thinking for Intelligence Analysis: A Gentle Introduction* (to appear).

Tecuci G. (2023b), *Instructable Cognitive Agents For Critical Thinking* (to appear).

Tecuci and Schum (2023a), *Critical Thinking for Intelligence Analysis: Connecting the Dots*, (to appear).

Tecuci and Schum (2023b), *A Practicum in Evidence Marshaling and Argument Construction: Connecting the Dots*, (to appear).

Turoff, M. (2007), Design of Interactive Systems, in *Emergency Management Information Systems Tutorial*, The Hawaii International Conference on System Sciences, HICSS-40, Hawaii, 3 January 2007.

Twining W. (2003). Evidence as a Multi-Disciplinary Subject, *Law, Probability and Risk*, 2, 91-107.

Wigmore, J. H. (1913). The Problem of Proof. *Illinois Law Review*, Vol. 8, No. 2, 1913, 77-103.

Wigmore, J. H. (1937). *The Science of Judicial Proof: As Given by Logic, Psychology, and General Experience, and Illustrated in Judicial Trials.* 3rd Edition, Little, Brown & Co., Boston, MA, 1937.

Zadeh, L. (1983). The Role of Fuzzy Logic in the Management of Uncertainty in Expert Systems, *Fuzzy Sets and Systems*, Vol.11, pp. 199 - 227.

Acknowledgements

David Schum (1932 – 2018) has significantly influenced our work on critical thinking and evidence-based reasoning.

The views and conclusions contained herein are those of the author and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the George Mason University or the U.S. Government.