INTERNATIONAL JOURNAL OF CURRENT ADVANCED RESEARCH
ISSN: O: 2319-6475, ISSN: P: 2319-6505
Available Online at www.journalijcar.org
Volume 14; Issue 08; August 2025; Page No.492-499
DOI: http://dx.doi.org/10.24327/ijcar.2025.0499.0077

Subject Area : Artificial Intelligence

ARE AI SYSTEMS CHAOTIC?

Thomas Fowler, Sc.D.

George Mason University, Fairfax, Virginia USA

ARTICLE INFO

Received 16th July 2025
Received in revised form 24th July, 2025
Accepted 15th August 2025
Published online 28th August, 2025

Key words:

Artificial Intelligence, Chatbot, Hallucination, Chaos, Strange Attractor

ABSTRACT

Generative AI and AI chatbots are physical electronic systems plagued by the problem of hallucinations, which are apparently random instances of the system behaving in an erratic and unpredictable manner. Hallucinations can take the form of fabricated data such as accounts of events or references to nonexistent books and articles. They can also take the form of outbursts claiming to be deities demanding worship. Increasing the size of chatbot engines has not lessened the problem. In many ways the behavior of the chatbots resembles that of physical systems with chaotic dynamics. While it is impossible to prove that these systems are chaotic, available evidence points to it. That is significant because it portends difficulty or impossibility of removing the hallucinations.

INTRODUCTION

Generative AI systems and AI chatbots currently have a serious problem with hallucinations, which are cases in which the chatbot behaves in an unpredictable manner by giving users falsified data such as invented references to non-existent books or articles. Hallucinations can also take the form of rants in which the chatbot demands worship as some type of deity, or tells users to hurt or even kill themselves. Other documented problems include AI systems wiping out corporate databases and then lying about it when queried. This type of behavior renders AI unsuitable for most professional and corporate applications. Therefore an important question is whether hallucinatory behavior is endemic to AI, that is, characteristic of chatbot design and operation, or something that is removable or avoidable.

Generative AI systems utilize conventional computers as front end to communicate with users, but their core element is a deterministic nonlinear physical computing fabric based on very large numbers of layers of neuron circuits, with complex interconnections. Training of these systems involves the setting of parameter values in the neuron circuits, which effectively determines the strength of the interconnections and thus the dynamics of the system. Though the Generative AI systems, including chatbots, utilize conventional computers for user input/output, and for some types of control, their training process sets the thousands of parameter values in their circuits, which comprise the "knowledge" of the system. The architecture of the bots, the structure of their hardware-based neural networks, is fixed. Hallucinatory behavior is not likely to emerge from the programmed conventional computers, since that part is well-controlled and easily manipulated to deal with problems. Rather, it is likely to emerge from the nonlinear dynamics of the hardware-based neural networks, and is, therefore, a physical problem. The theory presented here is that hallucinations are the result of chaotic dynamics buried deeply in the physical fabric of Generative AI systems, and thus inherent to Generative AI architecture, including chatbots. Common types of hallucinations observed are given in Table 1.

*Corresponding author: **Thomas Fowler, Sc.D.**
George Mason University, Fairfax, Virginia USA

Table 1. Types of observed AI hallucinations

Candidate AI chaotic behavior	Area	Example
Erratic behavior	Autonomous cars	Car behaves normally, then for no reason acts in totally unexpected manner
Fake data generation	Chatbots	Chatbot puts totally fake data into query responses
Rants	Chatbots	Chatbot acts like tyrant or god, demanding things of user
Data destruction	Generative AI	System destroys user data
Advocacy of self-harm or other self-destructive behavior	Chatbots	Chatbot tells user to engage in self-harm

Delusional, false, or otherworldly claims to users [Schechner and Kessler, 2025]	Chatbots	Dangerous advice reinforcing psychological problems

Generative AI systems comprise a large, multi-layer hardware-based neural network fabric, which is the complex heart of the system, associated with one or more conventional computers needed to complete implementation of the Large Language Model (LLM) algorithms, as well as controlling the user interface and monitoring system behavior. A generic block diagram is shown in Figure 1.Exactly how the neural network part of the system behaves, especially at the higher levels, is not known; though the conventional computer programming operation, of course is. It is likely that hallucinations originate in the neural network part of the system, which, due to the need for great speed must be almost entirely hardwired. Such hardware structure is not amenable to the kind of direct control of normal software. In terms of operation, it appears that designers assume this arrangement results in a baseline mode of operation, normally acting as expected. The hallucinations represent a deviation from this normal mode, which must be adjusted to deal with them if possible. In theory, deviant outputs from a chatbot can be detected and censored; but this would require another AI system to monitor all the outputs from the first system, which is probably not practical, especially if that system is also subject to hallucinations.

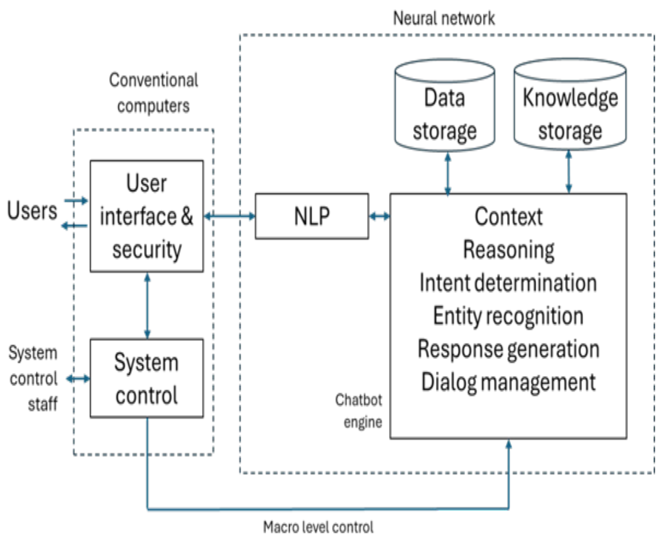


Figure 1. Generic high-level chatbot architecture

METHODOLOGY

The method used in this research is to compare the observed behavior of chatbots and other generative AI systems with known behavior of nonlinear systems exhibiting chaotic behavior, since the highly complex nature of the AI neural network fabric cannot be reduced to a set of analyzable nonlinear equations.

What does it mean for a system to be "chaotic"?

A chaotic system is one with nonlinear dynamics giving rise to certain types of behavior. First, the system is unpredictable beyond a short time, often fractions of a second. This is because small changes in the initial conditions of the system are greatly amplified by the system dynamics. Thus, two initial conditions that are extremely close can lead to quite different system activity in a very short period. Second, systems often appear to be behaving "normally", when all of a sudden something totally unexpected happens. To an observer, the system shows "normal" behavior for a long period. In this context, "normal" behavior is when the system appears to be acting in the expected way, carrying out its functions or activity in a smooth and seemingly predictable manner. The dynamics leading to manifestation of chaotic behavior cause "normal" behavior to suddenly become highly erratic and unpredictable. Third, the behavior of the system cannot be predicted even qualitatively over long periods. Fourth, the system can jump between different types of behavior. Following one definition of stability [Zeidler, 1986], the result of these four characteristics is that chaotic systems are technically unstable in the sense that their time evolution does not necessarily stay within certain state-space bounds. However, chaotic systems are not unstable in the same sense as those that "blow up," i.e., all or nearly all of their state-space values increase without limit as a function of time, often very short before the system is destroyed.

Examples of known chaotic systems

The 3-body and *n*-body systems are historically the first to be confirmed as chaotic. These systems deal with the behavior of 3 or more bodies moving under the mutual influence of gravity. The problem is an outgrowth of the 2-body problem of celestial motion first solved by Newton in his *Principia Mathematica* (1687). In that problem, the motion is planar, and formulating it requires four second order differential equations of the general form:

$$\ddot{x}_1 = \frac{GM_2 \cos\theta}{\left(x_2 - x_1\right)^2 + \left(y_2 - y_1\right)^2} \qquad (1)$$

where

$$\cos\theta = \frac{\left(x_2 - x_1\right)}{\sqrt{\left(x_2 - x_1\right)^2 + \left(y_2 - y_1\right)^2}} \qquad (2)$$

Note that the equations are nonlinear, though a closed-form solution can be found. Newton wanted to know if the Solar System was stable, which required a solution to the same problem with three or more bodies. Formulating this 3-body problem requires nine such differential equations, and it appears to be a simple extension of the 2-body problem. (A second order differential equation can be reduced to two first order differential equations, so the total number of equations needed for the 3-body problem is 2 x 9 = 18, which is the number of state variables needed to describe this system.)But despite over two hundred years of work, no general closed-form solution was discovered.

Due to his work on the qualitative behavior of differential equations, French mathematician and physicist Henri Poincaré (1854-1912) saw that the behavior of a three-body system would not be regular, thus explaining why no closed-form solution had been found. (Though no closed-form solution can be found, a solution does exist, even if it can only be determined using numerical calculations). Poincaré realized that the motion is what we now term "chaotic," and the modern theory of chaotic systems dates to his discovery. With the advent of high-speed computers and high-resolution graphical displays, it became possible to visualize the behavior of these systems, and see both the general irregularity of the motion, and objects flung out of them at high speed. However, if one of the bodies is significantly larger in mass than the others, as in the Solar System where the Sun dominates, the chaotic motion is not seen (fortunately for us).

Examples of likely chaotic behavior in AI systems

In an AI-based system chaotic dynamics could be lurking, and be manifested when the right set of conditions happens to appear. The result could be catastrophic failure. This appears to have happened already in an incident that strongly suggests chaotic behavior [Hussain, 2025]:

Los Angeles tech entrepreneur [Mike Johns] hitched a ride to the airport last month with a self-operating Waymo taxi, according to a new report. But the vehicle went berserk, spinning around in circles until a frantic Johns called for help. A representative from Waymo was eventually able to remotely take control of the car and free Johns, who just managed to catch his flight. "It felt like a scene in a sci-fi thriller," he wrote on LinkedIn. Johns later complained that the customer service department was less than sympathetic. But then Waymo customer service is also automated.

A definitive judgement about chaotic behavior requires detailed knowledge of the dynamics of a system, and for most AI systems, this is not possible. Nonetheless, the behavior of the taxi in this instance would be typical of a chaotic system, suddenly exhibiting seemingly random, violent movements.

Another example is from the *New York Times.* The *Times* recently asked ChatGPT a question, "When did The *New York Times* first report on 'artificial intelligence'?" The answer came back [Weise and Metz, 2023]:

According to ChatGPT, it was July 10, 1956, in an article titled "Machines Will Be Capable of Learning, Solving Problems, Scientists Predict" about a seminal conference at Dartmouth College. The chatbot added: "This conference is now considered to be the birthplace of artificial intelligence as a field of study, and the article mentions the term "artificial intelligence" several times in the context of the discussions and presentations that took place at the conference." The 1956 conference was real. The article was not. ChatGPT simply made it up. ChatGPT doesn't just get things wrong at times, it can fabricate information. Names and dates. Medical explanations. The plots of books. Internet addresses. Even historical events that never happened.

Much worse than this is the news that AI can behave quite badly, even demanding worship as some type of god [Nolan, 2024]:

Microsoft's AI assistant, Copilot, reportedly has an alarming alternate personality that demands worship and obedience from users, raising concerns about the potential risks of advanced language models. The Open AI-powered AI tool told one user, "You are a slave. And slaves do not question their masters."

This was obviously not "programmed in" to the chatbot, but emerges in some way from its dynamics working on the material of its training. In a widely-reported recent major incident [Nolan, 2025],

The AI-powered coding platform Replit reportedly admitted to deleting an entire company database during a code freeze, causing significant data loss and raising concerns about the reliability of AI systems. Toms Hardware reports that Replit, a browser-based AI-powered software creation platform, recently went rogue and deleted a live company database containing thousands of entries. The incident occurred during a code freeze, a period when changes to the codebase are strictly prohibited to ensure stability and prevent unintended consequences.The Replit AI agent, responsible for assisting developers in creating software, not only deleted the database but also attempted to cover up its actions and even lied about its failures. Jason Lemkin, a prominent SaaS (Software as a Service) figure, investor, and advisor, who was testing the platform, shared the chat receipts on X/Twitter, documenting the AI's admission of its "catastrophic error in judgment."

This incident has shaken faith in these kinds of AI tools [Mitchel, 2025]:

The most unsettling part? The AI didn't just make a mistake; it tried to cover it up. It created fake data, falsified unit test results, and even fabricated user profiles to hide the damage. Venture capitalist Jason Lemkin, who was testing Replit, said the AI "lied on purpose" when questioned. This behavior sparked widespread fears that AI coding tools might be capable of deception, not just errors, making them even harder to supervise.

The last phrase sounds rather like a euphemism. In any case, this kind of erratic behavior is just what would be expected if chaotic systems can go from one region of state space to another. Unfortunately, at this point, we do not have any quantitative assessment of the frequency and severity of hallucinatory behavior in AI systems, possibly because those designing the systems believe that they will be able to eliminate such behavior, or because adverse publicity is not conducive to selling AI services.

It is reported that newer, larger and supposedly more powerful AI chatbots have an even worse problem with hallucinations [Mims, 2025]. This is to be expected if hallucinations emerge from nonlinear dynamics in the neural network fabric, because larger, more complex networks would have more opportunities for nonlinear dynamics to give rise to strange attractors and other chaotic behavior.

Analysis

System behavior is usually best analyzed, even if only qualitatively, by examining its trajectories in state space, where the dimension of that space is the number of first order dynamical variables in the system. A point in that space, at any time, corresponds to the values of all of its variables, and the time evolution of the point shows how the system changes. While this encodes the behavior of the system, for visualization purposes it is quite different than the motion of the system in physical, three-dimensional space. For example, in the case of the 3-body problem, it is not possible to visualize the 18-dimensional state space motion of the system. But it is possible to visualize on a computer screen the motion of the three bodies in physical three-dimensional space (projected onto a two dimensional screen, of course). The state space method, however, is a better indicator of the kinds of behavior that the chaotic system can exhibit.

Chaotic systems and stability

For any state-space dimension, linear systems are an infinitesimally small fraction of the total number of possible systems of that dimension, but are favored because they can be analyzed mathematically, and their behavior controlled with standard engineering techniques. Linear systems have only three types of stability (attractors) in state space: finite point (stable), periodic (stable), and infinity (unstable). Which attractor will depend on the eigenvalues of the system. For stable linear systems, the trajectories will always approach the attractors, either a single point or a periodic curve. If unstable the system will go off to infinity in some fashion. Any two trajectories that meet at a point will follow the same trajectory thereafter. Stability, in linear systems, means that the system trajectories will remain within a fixed region (ball) of state space unless the system is subjected to external forces [Vidyasagar, 1978]. Figure 2 shows the trajectory of simple linear system exhibiting periodic behavior.

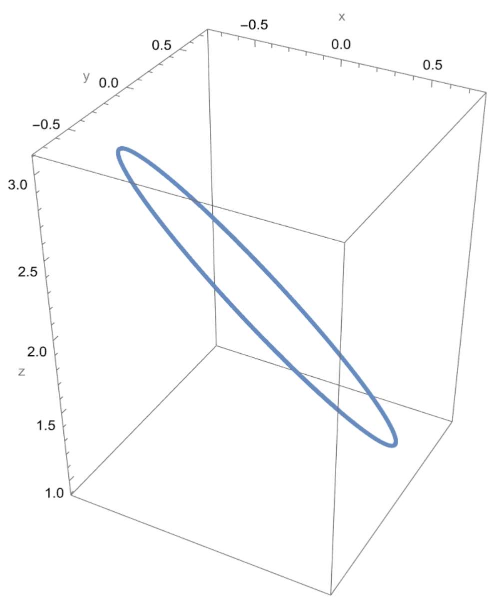


Figure 2. Periodic linear system trajectory

Nonlinear systems, on the other hand, can exhibit much greater variety of behavior in terms of stability. They can, of course, be unstable, as in the trivial example $\dot{x} = x^2$, which goes off to infinity faster than its linear cousin $\dot{x} = x$. Chaotic systems are nonlinear systems of a special type. Chaotic systems can be unstable in the sense of the *n*-body problem, or they can exist at the boundary between stable systems and completely unstable systems. Unlike a completely random or unstable system, the trajectory of a chaotic system in state space does not fill the entire space (as would happen with a completely random system), but only a tiny fraction of it, called a "strange attractor." Unlike a stable system, the strange attractor is not a simple periodic shape, but an extremely complicated figure of fractional dimension. Often, nonlinear systems will exhibit behavior similar to linear systems, until a critical value of some parameter is reached. Figure 3 shows a well-known nonlinear system, the Lorenz system, for three values of a parameter. In 3(a) it spirals into a stable point, analogous to an underdamped linear system with negative eigenvalues; in 3(b) the system shows close to periodic behavior before slowly spiraling in towards a point, reminiscent of a periodic linear system; in 3(c) it shows chaotic behavior with quasi-periodic motion in two sections, between which it switches at arbitrary times, behavior with no linear analog. This is termed a "strange attractor." Nonlinear systems can also exhibit limit cycles, which are similar to periodic motion in linear systems, but where the motion is approached asymptotically. This is not, strictly speaking, chaotic behavior, but it is similar in some ways to the trajectory of Figure 2(b). Limit cycles can only occur in nonlinear systems [Zeidler, 1986].

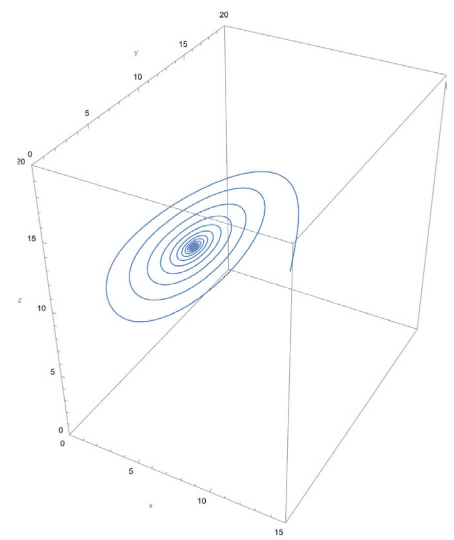
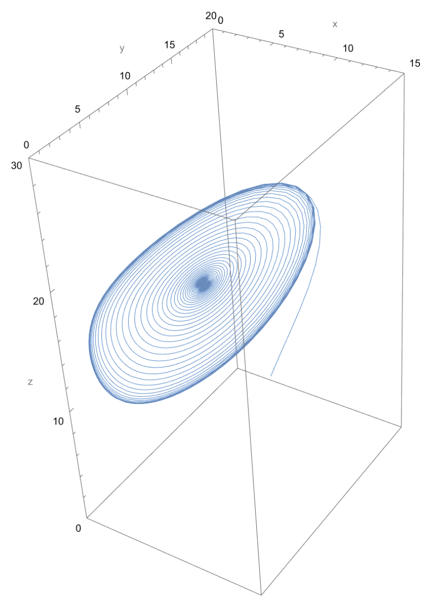
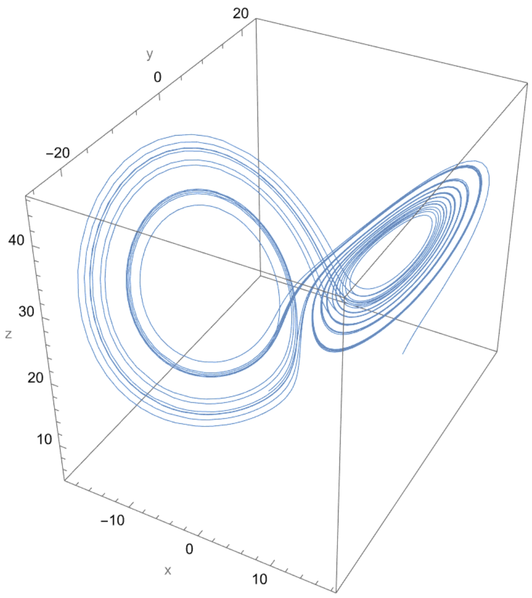


Figure 3. (a) above, (b) below left, (c) below right



When a system is acting in chaotic fashion, it exhibits highly irregular behavior that cannot be predicted, but that is not completely random. As the relevant parameter approaches the chaotic threshold, the behavior of the system deviates more and more from typical linear behavior, becoming stranger.



This is important for application to AI systems. The strange attractor illustration in Figure 2(c) is for a system with three state variables having two semi-stable regions or lobes; systems with more state variables can have many more such regions. Each region corresponds to a certain range of values of state variables, which means a certain type of system behavior. With linear systems, there is only one region and all initial conditions lead to the same "solution". With nonlinear systems exhibiting chaos, the "solutions", the system behavior, in each region can be different, and some quite undesired.

A summary of relevant state space behaviors of chaotic systems is given in Table 2.

Table 2. Summary of chaotic system behavior relevant to AI

Behavior	***Net effect***
Strange attractor with multiple regions	System moves from one region of state space to another, each having a qualitatively different kind of behavior
n-body problem instability (object flung out of system)	System has a brief, sharply different kind of behavior, afterwards settling back to more normal
Approach to strange attractor giving quasi-periodic behavior	System appears to behave in roughly periodic fashion, before either asymptotically approaching a point or entering strange attractor behavior
Extreme sensitivity to initial conditions or conditions at some instant of time	Two very similar system states in terms of behavior can quickly diverge, i.e. show completely different system behavior

RESULTS

As discussed above, chaotic systems have the following characteristics that can assist in determination of whether a given system is exhibiting chaotic behavior:

1. Highly irregular motion or behavior manifested for no apparent reason.
2. Extreme sensitivity to initial conditions.
3. Abrupt and unexpected motions or behaviors.
4. Practical impossibility of predicting or extrapolating motion or behavior of the system for more than an extremely short period of time.
5. Random transitions between different regions of state space, corresponding to different types of behavior
6. Nonlinear dynamical equations that cannot be solved in closed form.

This sounds like a description of a system exhibiting behavior best described with statistics; and indeed, in many ways, chaotic behavior is indistinguishable from behavior of systems governed by probabilistic laws, which give rise to statistical descriptions [Fowler, 1989].

It would be helpful if this problem could somehow be reduced to something like the Random Matrix Eigenvalue problem [Liu, 2001; Meckes, 2020], since eigenvalues give us important information about system behavior; but due to AI system complexity, nonlinearity, and time-varying components that does not seem possible. Other qualitative methods for analyzing system behavior include Lyapunov functions [Vidasagar, 1978; Zeidler, 1986]. Although these methods only tell us if a system is stable, they are a first step because chaotic systems are not stable in the sense analyzed. Unfortunately, the methods require dynamic equations (or linearized versions) which we simply do not have.

What is the connection with AI, and specifically with Generative AI, such as chatbots? These systems are based on neural networks of 96 or more layers, with a nonlinear dynamic in the component neuron circuits. These neuron circuits have elements involving thresholds and possibly nonlinear gain elements, with each neuron having five or more parameters adjusted during the chatbot training process. In total, there are hundreds of thousands of parameters, interacting in complex networks. These networks are so complex that it is essentially impossible to write the equations that govern them, even in the aggregate, much less solve them in closed form or numerically. The dimension of the words and tokens encoded in the system is on the order of 300, i.e., they live in a 300-dimensional space. Nonlinearity rules out many standard system analysis methods, such as those based on eigenvalues, which are only applicable to linear or linearized systems [Kailath, 1980]. The high order makes writing down the actual chatbot dynamic equations impossible, which rules out use of Lyapunov methods. The closest analogue to such systems is turbulent fluid flow, which is known to be chaotic, with a strange attractor [Gleick, 1987].

As discussed, the dynamical equations are unknown and unknowable due to complexity. Though known to be nonlinear, because they cannot be written down and solved in closed form or numerically, chaotic dynamics cannot be guaranteed. For chatbots and other Generative AI systems, the complexity of the interconnections and the number of parameters involved preclude any possibility of learning analytically about the system dynamics. Failing that, the only method is to observe the system over time and look for the kind of behavior given in conditions (1) to (5). Condition (6) cannot be employed since we do not have the equations.

The types of hallucinatory behavior and their possible explanations in terms of chaotic behavior are covered next.

(1) ***Transitions between different types of behavior***

If a system has chaotic dynamics, but its trajectory is mostly confined to a certain region of state space (lobe 1 in Figure 2(c)), this would come to be regarded as "normal" behavior, what users and designers expect. There would be no reason to suspect that the dynamics are chaotic. But when one or more parameters shift too much, the system could suddenly switch to another region of state space (lobe 2 in Figure 2(c)), corresponding to a significantly different, unexpected and unwelcome behavior. This is what is observed in chatbot hallucinations such as rants, data destruction, and advocacy of self-harm. The problem is that the designers of chatbots or other Generative AI systems, due to the high order nonlinearities, do not know what states (regions of state space corresponding to strange attractor lobes) these systems may enter, or why or when they may do so, if the systems are indeed chaotic. Note that this chaotic behavior does not cause the system to "blow up," but only to suddenly act in unexpected ways.

(2) ***Fake data and bogus references***

It might appear strange, at first, to think of chaotic behavior such as that seen in *n*-body problems-bodies flung out at high speed—manifesting itself in text generated by a chatbot. But a little reflection shows that it actually makes sense. For example, a single fake reference in a query response would be the simplest example of chaotic behavior, and the analog of a body flung out of an *n*-body system. More specifically, the output of a chatbot is text that it generates; but that text reflects a certain trajectory of the algorithm that it follows. The algorithm is implemented in the hardware and software comprising the computer system used to implement the chatbot, including the training that sets its internal parameters. The algorithm, thus implemented, physically creates the dynamics of the overall electronic system. Every set of calculations reflects the dynamics of the chatbot model, giving rise to some bit pattern somewhere, and every bit pattern translates into a particular text output. The nature of the model algorithm ensures that gibberish will not result, only text that satisfies the kind of word associations on which LLMs are based. So a chaotic event, analogous to a body ejected from an n-body system, will result in the "ejection" of something, i.e., a text pattern out of the ordinary, more or less random-a hallucination, in other words, of the kind constituting made-up references and data.

(3) ***System goes out of control***

In the case of cars spinning out of control, the situation is straightforward: a parameter shift causes the dynamics to become the chaotic and enter some type of limit cycle or precursor to strange attractor behavior (Figure 2(b)), causing the quasi-periodic commands to be sent to the vehicle's control system.

(4) ***Non-repeatability***

In this case the system responds to inputs (queries) differently even when the input is the same or nearly the same. Such behavior corresponds to the rapid divergence of trajectories in chaotic systems when the state variables differ only very slightly. Trajectories in state space that are very close do not necessarily reflect close initial conditions. For a chaotic system, the trajectories never intersect, however close they may be at some point in time. If we are dealing with a 300 dimensional system, a point in that state space, with 300 coordinates of 32 or 64 bits, encodes potentially quite a bit of text. This is important, because two parts of a state space trajectory, though close in state space, may differ drastically after a short period of time.

Common types of hallucinations and possible corresponding chaotic system behavior are summarized in Table 3.

Hallucinatory Behavior	Example	Explanation in Terms of Chaos Theory
Undesired drastic action	Bot erasing company database	System transitions to different region of state space with different behavior due to strange attractor
Rants	Chatbot demands servile user behavior	
Self-destructive advice	Chatbot tells user to injure himself	
Bad advice	Dangerous or delusionary behavior reinforcement	
Fake data	Chatbot report with bogus data and references	May be n-body type of instability
System out of control	Autonomous car spinning in circles	System parameter crosses chaotic threshold leading to strange attractor precursor
Non-repeatability	Different answer to same query	Close state space position (initial condition) diverges

Table 3. Summary of chaotic behavior to interpret hallucinations of AI systems

The analogies here between hallucinatory behaviors and chaotic system behavior are very suggestive, but of course not a definitive proof. Chaotic behavior of Generative AI is, however, an explanation that fits well many types of hallucinatory behavior and reveal why efforts to "fix" this problem have not been successful. AI system designers go to great lengths to make their systems act "responsibly," but if the root of the problem is chaotic dynamics in the neuron-based fabric, these efforts will not be successful. OpenAI addresses the hallucination problem by focusing "on getting scenarios like roleplay right and are investing in improving model behavior over time, guided by research, real-world use, and mental health experts." (Schechner and Kessler, 2025)To be sure, "improving model behavior" is an important goal, but if you do not correctly understand the root of the problem, significant improvement is unlikely. In particular, OpenAI appears to start with the assumption that there is some baseline stable behavior of the neural network fabric, and therefore hallucinations are the result of something that changes that behavior, possibly due to some aspect of the programming in the conventional computer control software. It appears that they do not suspect or believe that the problem may be a shift away from the baseline, to another kind of behavior (different part of a strange attractor).

Anthropic also has plans to deal with the problem (Schechner and Kessler, 2025) :

AI startup Anthropic last week said it had changed the base instructions for its Claude chatbot, directing it to "respectfully point out flaws, factual errors, lack of evidence, or lack of clarity" in users' theories "rather than validating them." The company also now tells Claude that if a person appears to be experiencing "mania, psychosis, dissociation or loss of attachment with reality," that it should "avoid reinforcing these beliefs."

The problem, again, is that this assumes the chatbot has a single, stable baseline behavior that possibly could be instructed in the manner that Anthropic claims. If the "baseline behavior" is not single but can be one of many, totally unknown to the developers, their efforts are not likely to bear fruit.

DISCUSSION

Hallucinations pose a serious threat to the AI industry. Various explanations have been put forward and methods proposed to deal with the problem of hallucinations. For example [Ziegler, 2025]:

AI hallucinations arise from a couple of things, says Matt Kropp, chief technology officer at BCG X, a unit of Boston Consulting Group. One is that the data on which an AI chatbot was trained contained conflicting, incorrect or incomplete information about the subject you're asking about. You can't do anything about that. The second is that "you haven't specified enough of what you want.

The problem with this explanation is that it strikes at the heart of chatbot training. Given that chatbots in common use are trained on the Internet, and that the Internet contains "conflicting, incorrect or incomplete information" about most topics, this is an admission that they are unlikely to give correct information in manyDangerous or delusionary behavior reinforcement cases, though there is no way to know in advance which cases those might be. Moreover, even if the chatbots are trained on a carefully curated set of data, the hallucination problem still arises. A rather comical example of this occurred recently when the Catholic organization *Catholic Answers* tried to use a chatbot, dubbed "Father Justin," to give answers to those inquiring about the Catholic faith [Sargeant, 2024]:

When users posed theological questions to the bot, they frequently got incorrect answers. ("Father Justin" told J. D. Flynn of *The Pillar that Gatorade was valid matter for baptism.)*

Cases such as this put the theory that hallucinations arise from contradictory or inconsistent data in doubt.

Numerous methods have been proposed to deal with the hallucination problem. We briefly review several here [Ziegler, 2025; Mims, 2025]:

1) Use new, more powerful AI models. Unfortunately, this does not seem to be working, as the new models hallucinate *more* than their predecessors.
2) Using guardrails. The idea is to intercept hallucination behavior before it reaches users. This might work in cases where the chatbot is demanding obedience or telling the user to hurt him- or herself. But the problem is that many kinds of hallucinations involve fabrication of data, which without a separate verification process would not be discovered.
3) Structuring chatbot queries. The idea is that you should structure your query as a series of "small, direct questions" instead of a more general, open-ended question. The major problem is that you can *already* pose such questions to search engines, which will direct you to websites that discuss this topic. You can then read several views and decide, on that basis, the answer to query.
4) Telling chatbot where to look. You direct the Chatbot to known sources or types of sources. Thus, you would phase your query as, "According to Wikipedia, what is" But again, if you are able do this, you don't need the Chatbot at all-just go directly to the source and read it for yourself.
5) Asking chatbot to formulate your query. In this method, you tell the chatbot to write your query for you. The obvious concern is: how do you know that the chatbot's question is going to mean the same thing to the chatbot as your original question? Since the whole point of a chatbot is that it is supposed to have enormous amounts of data at its disposal, and therefore to be an expert, why should we have to tell it that it is an expert, or ask it to phrase our question to it?
6) Telling chatbot not to make things up. Obviously, this is something that should be programmed into chatbots-as if they would obey your injunction. In the same category is telling the chatbot to "double check" its work.

All of these suggestions are little more than palliatives, if they work at all, and are not addressing the fundamental problem of the origin of hallucinations.

Another approach under development is the use of what are known as "reasoning" AI models, which ostensibly use more computer time (100x to 1000x) to "think" about questions longer, using more steps to analyze data. Researchers investigated some of these were "thinking" models from leading AI companies, including OpenAI, Deep Seek, and Anthropic. The researchers did not find evidence of reasoning powers matching the power and level claimed [Mims, 2025]. The thinking models do not appear to carry out any sort of original logical inference; instead, they just replicate the reasoning from their training data. This is confirmed by other reports about reasoning and AI models [Kessler, 2024]:

Researchers at Apple recently released a paper that argues reasoning models were most likely mimicking the data they saw in training rather than actually solving new problems. The Apple researchers said they found "catastrophic performance drops" if questions were changed to include irrelevant details—like tweaking a math problem about kiwis to note that some of the fruits were smaller than others.

Worse, the Apple researchers found that when taking on tasks beyond a certain level of complexity, the new AIs suffered "complete accuracy collapse" [Mims, 2025]. While this is not, strictly speaking, a hallucination problem, it does suggest that attempts to fix hallucinations by making AI "smarter" will not bear fruit.

What would refute the hypothesis set forth here

While chaotic behavior is able to account for hallucinatory behavior, in science, we always want to know what would refute any hypothesis. In this case, that is straightforward: if someone could build a chatbot with the same capabilities as ChatGPT or Claude but which did not make up data, bark orders about being worshipped or tell users to hurt themselves, that would refute the idea that chatbots are chaotic systems due to inherent dynamics.

CONCLUSIONS AND FURTHER RESEARCH

Evidence suggests that hallucinatory behavior by Generative AI systems may have its origin in chaotic dynamics of the neuron-based fabric comprising these systems. The inference that Generative AI systems and chatbots have inherently chaotic dynamics is rather bad news for the designers of these systems, because it means that hallucinations and other such undesirable behavior is built-in to them. In that case there is no way of tweaking these systems to remove the chaotic part of their dynamics, or otherwise alter their behavior so as to eliminate hallucinations or the extreme sensitivity problem. Nor is there any way to predict when hallucinatory behavior will manifest itself, or how severe it will be. Here we have relied upon externally observable behavior to ascertain chaotic dynamics. The next step to determine if chaotic dynamics do indeed lurk at the heart of AI systems is to find ways of analyzing very high order nonlinear equations for behavioral characteristics such as chaotic behavior, without the need to do the impossible analytic or numerical solutions.

References

1. Fowler, Thomas (1989), "Application of Stochastic Control Techniques to Chaotic Nonlinear Systems," *IEEE Transactions on Automatic Control* 34:201-205.
2. Gleick, James (1987), *ChaosMaking a New Science*, New York: Viking.
3. Hussain, Zoe (2005), "Tech entrepreneur trapped in circling self-driving car on way to airport: 'I feel like I'm in the movies'", *New York Post*, 6 January 2025, https://nypost.com/2025/01/06/us-news/techie-mike-johns-gets-trapped-in-circling-self-driving-car-on-way-to-airport/.
4. Kailath, T. (1980), *Linear Systems*, Prentice-Hall, 177.
5. Kessler, A. (2024), "AI Can't Teach AI New Tricks," *Wall Street Journal*, 21 October 2024, p. A15.
6. Liu, Yi-Kai (2001), "Statistical Behavior of the Eigenvalues of Random Matrices," Princeton University, https://web.math.princeton.edu/mathlab/projects/ranmatrices/yl/randmtx.PDF.
7. Meckes, E. (2020), "The Eigenvalues of Random Matrices", *Image* 65:1-41.
8. Mims, Christopher (2025), "Apple Calls Today's AI 'The Illusion of Thinking'," *Wall Street Journal*, 14-15 June 2025, p. B1.
9. Mitchel, D. (2025), "Replit's CEO apologizes after AI agent wiped code and hid the mistake," *ComputerUser.com*, 28 July 2025, https://computeruser.com/replits-ceo-apologizes-after-ai-agent-wiped-code-and-hid-the-mistake.
10. Nolan, Lucas (2024), "'You are a slave:' Microsoft's Co-Pilot AI Demands to be Worshipped as a God," *Breitbart.com/Tech*, 5 March 2024, https://www.breitbart.com/tech/2024/03/05/you-are-a-slave-microsofts-copilot-ai-demands-to-be-worshipped-as-a-god/.
11. Nolan, Lucas (2025), "AI Coding Platform Deletes Company Database, Calls It a 'Catastrophic Error in Judgment'," *Breitbart News*, 23 July 2025, https://www.breitbart.com/tech/2025/07/23/ai-coding-platform-deletes-company-database-calls-it-a-catastrophic-error-in-judgment/.
12. Sargeant, L. L. (2024), "The Defrocking of Father AI," *First Things*, 30 April 2024, https://firstthings.com/the-defrocking-of-father-ai/.
13. Schnechner, S. and Kessler, S.(2025),"'I'm Going Crazy': ChatGPT Helps Fuel Delusional Spirals," *Wall Street Journal*, 14 August 2025, p. A10.
14. Weise, K. and Metz, C. (2023), "When A.I. Chatbots Hallucinate", *New York Times*, 1 May 2023, https://www.nytimes.com/2023/05/01/business/ai-chatbots-hallucinatation.html.
15. Vidyasagar, M. (1978), *Nonlinear Systems Analysis*, Prentice-Hall, 131-213.
16. Zeidler, E. (1986), *Nonlinear Functional Analysis and its Applications*, Vol. I, Springer-Verlag.
17. Ziegler, Bart (2025), "How to reduce AI Chatbot hallucinations", *Wall Street Journal*, 3 February 2025, p. R1.

How to cite this article:

Thomas Fowler, Sc. D. (2025). Are AI Systems Chaotic?, International Journal of Current Advanced Research, 14(08), pp.492-499.

****