

Ontology families for AI-systems

Dennis de Champeaux

ddc2@ontooo.com

Unaffiliated / Silicon Valley

Abstract

The recent chat-systems exhibit a discontinuity against the AI achievements since 1970. They support ‘chat’ interaction with customers by relying on having been trained on massive amounts of text. They can generate useful replies (beyond earlier Google results), which has amazed the general audience. At the same time these chat-systems cannot apply the crown-jewel algorithms (search, planning, reasoning, etc.) on simple problem specifications of these algorithms. We provide numerous examples showing that the Large Language Models system architecture is the root problem. In each case, we argue that an ontology is required to handle these examples. In addition, we argue that large AI-systems need multiple ontologies. These ontologies are the bridge to neuro-symbolic computing. Pursuing Artificial General Intelligence (AGI) requires – we claim – meta-ontologies. At the same we provide arguments against the ultimate feasibility of AGI. We permit ourselves to take the stance that AI-systems should avoid certain topics.

Keywords: Chat-systems, architecture defects, (meta) ontology.

Dedicated to Doug Lenat; he worked for close to 40 years on the CYC ontology (1950-2023).

1 Introduction

Living entities have – we postulate – idiosyncratic internal ontologies (provided by evolution) that mediate information streaming in through sensors and generating behaviors – for itself and social. There is a tight connection between such an ontology and the data-stream, using the concepts in the ontology, that is processed in working memory to generate executable actions. A major addition in the case of Sapiens is that the data-stream has a recursive language format, as postulated by Fodor’s Language of Thought in [LOTH]; often referred to as ‘mentalese’. The concepts in the mentalese language are wrapped inside a formalism that, for evolutionary reasons, is sensory based, according to Carruthers [Carruthers]. We postulate – without evidence – that Sapiens mentalese content coming from vision, hearing, touch, etc. sensors has ‘subtitles’ with symbolic content (that are labeled *amodal* by Carruthers). Consciousness has been ‘explained’ with subconscious consumer-semantic processing modules in working memory that can read the data-stream and write their ‘conclusions’ back into the stream. AI-systems (robotic or cyberspace versions) can have a similar arrangement, except that their data-streams do not have to be sensory based.

The ontologies associated with sensors are connected with the other ontologies with ‘horizontal’ links for corresponding concepts. The content of the ontologies from Sapiens is not only know-what and know-how but can have also *emotional* mark ups as well as *moral* mark ups. While

simple lifeforms may have fixed ontologies, complex lifeforms, like Sapiens, can have the ability to modify an ontology – at run time – with new concepts or new versions of concepts or even overhaul the topology of an ontology. New concepts can be concrete or abstract. A Sapiens member has commonsense clusters (mini-ontologies) for naïve physics, economy, number sense, psychology, etc. with optional corresponding clusters for the scientific/professional versions, see [Pinker]. Horizontal links, which we postulate, binds them all together.

The Sapiens *species* has *implicit* ontologies captured by the documents that have been accumulated in the preceding millennia. There is also an agreed process to extend and/or modify these collective, implicit ontologies.

Given this portrayal of ontologies for living entities, we can wonder how AI-agents can be equipped with ontologies. What do they have in common and what not?

AI has a 70 years history of boom and bust cycles. The recently released chat-systems have ignited another up cycle, which has inspired this time the stock market to take notice. The general public has become involved and large and small companies (in the Silicon Valley at least) are exploring how to leverage this novel technology.

The breakthrough arrived with the development of Large Language Models (LLM). Words, word combinations, noun phrases, sentence fragments, sentence pragmatics, paragraphs, etc., are tracked so that they can be matched against what is encountered in prompts for making sense of them. They also help with the generation of replies, although other mechanisms play a role here as well.

At this point, we need to remember that the Sapiens cognitive operations do not rely on an LLM. Natural language input and other input streams are mapped (indirectly via an intermediate representation according to Carruthers [Carruthers]) with concepts onto mentales. These concepts are recruited from a personal ontology that contains not only our concepts but also our micro-theories. A combination of these elements yield – provided an input makes pragmatic sense – a query, request, assertion, etc. that we can work with.

In addition to the difference between using natural language as the basis versus using concepts for what happens in our ‘working memory’, the chat-systems agree that they lack the equivalence of a limbic brain and hence do not have emotions (being in essence autistic). The development of Sapiens morality/ ethics is a work in progress and hence there is no consensus how to equip AI agents with morality/ ethics (thus they lack a moral ‘compass’). Noteworthy is also that chat-systems are cyberspace agents unlike robotic agents; thus they are not acquainted with the features and challenges of embodiment and all the intricacies of space. The latter have their own defects: lacking the 3D experiences that many species are endowed with.

Learning is certainly also a topic that distinguishes Sapiens and AI agents. We learn to walk and obtain our 1st language by maturation & osmosis, develop social and moral skills, obtain professional skills, motoric skills like, skiing, swimming, bicycling, sometimes obtain a 2nd language, etc. etc. which are all different kinds of learning and do not require reboots. These types of learning are science fiction for (robotic) AI agents. Calling the creation of an LLM (deep) learning is an abuse of the term ‘learning’. Teaching AI agents is mostly ‘surgery’ of code with testing after a reboot (or using customized editing tools). (Even parameter optimization through self-game play requires typically subsequent reboots.)

We argue here that the use of LLMs (and similar versions) in the chat-systems (and similar versions) is a key culprit for the defects we discuss. The alternative – we believe – is the usage of a loaded ontology, among others, that needs to be generated as discussed.

Another study titled “The Illusion of Thinking: Understanding the Strengths and Limitations of Reasoning Models via the Lens of Problem Complexity” provides a similar critique, see [Shojaee]. Their tactic is testing chat-systems on solving the puzzles: Tower of Hanoi, Checker Jumping, River Crossing and Blocks World. They used parametrized versions that allowed them to increase the level of difficulty and they analyzed traces where things went wrong. Their conclusion: "... they fail to use explicit algorithms and reason inconsistently across puzzles" and “Our findings reveal fundamental limitations in current models: despite sophisticated self-reflection mechanisms, these models fail to develop generalizable reasoning capabilities beyond certain complexity thresholds.”

They admit:

We acknowledge that our work has limitations. While our puzzle environments enable controlled experimentation with fine-grained control over problem complexity, they represent a narrow slice of reasoning tasks and may not capture the diversity of real-world or knowledge-intensive reasoning problems.

This paper compensates for their limitation by experimenting with a wider range of very easy challenges: recognizing a contradiction in the LLM training source, planning, graph search, analogies, story understanding, etc. We identify also deeper problems/ concerns.

We use the term ontology here for the combination of a structured, rooted graph of concepts for capturing their semantics with an associated knowledge graph that captures at least know-what and know-how about a particular domain – or more general – what is represented differently in the Large Language Models of chat-systems. This combination has been referred with the components labeled as TBox, for the graph with concepts, and ABox, for the content of a particular domain, [deGiacomo]. However, no chat-system is currently known employing an ontology.

2 Ontology generics

We assume that the TBox of an ontology is represented minimally with a rooted graph structure in which the nodes represent general concepts and the links represent conceptual inheritance between parent and children nodes. Figure 1 has an example of an ontology fragment [Sowa].

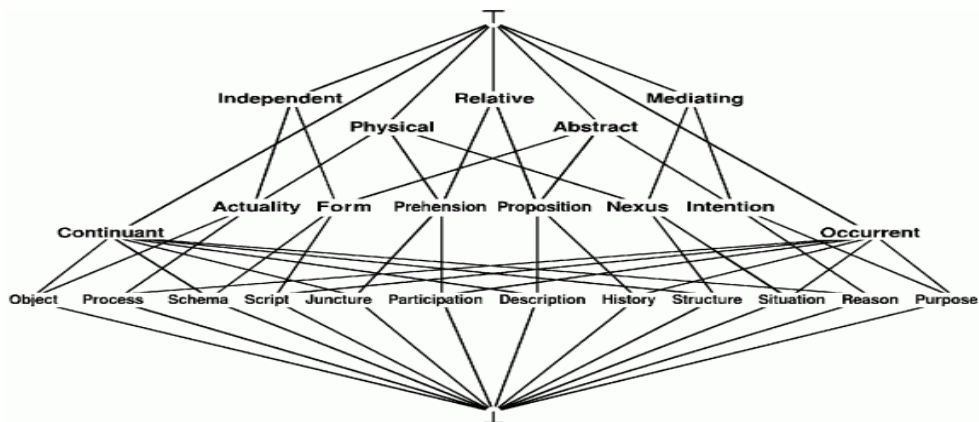


Figure 1: Hierarchy of the Sowa top-level categories with multiple inheritances.

Whether a node can have more than one parent (multiple inheritances as in Figure 1) is here uncommitted. Below there is an elaboration. Whether a node represents the intension of a concept or its extension is here also uncommitted. Horizontal linkages can be internal to an ontology or connect concepts in different micro-ontologies. Part-whole relationships are usually available as well as intrinsic and contingent attributes of concepts. Typical commitments are how to distinguish between tangible and intangible concepts and how time is handled. Micro-ontologies for specialized domains can be part of an ontology with here no specification how they are incorporated.

2.1 Knowledge components

There are different ways to partition knowledge. We have already suggested that multiple ontologies may be required for large systems (robotics & chat-systems). Micro-theories that have been introduced by McCarthy, Guha & Lenat [Lenat] correspond with what floats in Sowa's 'knowledge soup', [Sowa]. They are isolated, tightly organized chunks of formal reasoning. Developed for the Cyc project, these represent tiny, hyper-specific domains of reasoning (e.g., a specific timeframe, a cultural setting, or a strict discipline). They hold a small, carefully tested set of rules and axioms.

Another type of molecules regards combinations of procedural elements: algorithms, Schankian -scripts [Schank], tactics, strategies, procedures in many disciplines, etc. Developing a common framework for them is recommended.

A medical diagnosis algorithm (mentioned below) is *how*-knowledge and is an example of a micro-theory with many smaller sized 'blocks', bottoming out in symptoms, body locations, diseases, etc., which are *what*-knowledge examples.

Examples of large components:

- Scientific (sub) disciplines and their crown jewels
- Historical periods with open-ended different perspectives (demographics, energy, sociology, philosophy, technology, governmental variants, ...)
- Notable, unique, individual entities with their achievements and episodic elements

Frameworks for them are recommended as well.

3 Ontologies for sensors and actuators?

Sensors for robotic AI systems are specialized components with specialized interfaces. Standards have been developed for their side [IEEE]. A micro-ontology for these sensors helps offloading their functionality from the main ontology (or ontologies) of an AI-system.

These sensors provide typically information about tangible entities (with indeed exceptions like: rainbows, thunder, music, ambient temperature, etc.). An ontology for visual sensors, for example, would use for the characterization of these objects attributes like shape, color, size, etc. Where the generic concept of these objects resides is debatable. There could be a unique concept in a particular ontology with linkages to representatives in other ontologies. Or the concept could be distributed in multiple micro-ontologies like: for visual entities, for common sense notions,

for a particular scientific domain, etc.; the concept's attributes and associated operations could also be distributed. Horizontal linkages would bind them together.

Cyberspace AI systems would have similar arrangements for their sensors and actuators.

4 Ontology for common sense?

Current chat-systems support commonsense interactions – most of us may likely believe.

Consider the mini story:

John and Peter arrive at the same time in Denver.

John had walked an hour and Peter had deplaned from SF.

Who had left earlier?

Chat-systems failed on this one. Is it because it is beyond common sense? Commonsense is (according to Aristotle): a natural awareness that does not require specialized knowledge or complex reasoning. Discussing this example with random people makes us believe that this example challenge is solvable by commonsense. Hence the conclusion/conjecture that chat-systems – whatever other qualities they may have – lack (some) commonsense. The plausible reason is that they rely on a Large Language Model and do not have an ontology containing the semantics of events for action verbs in new input. This triggers what is known in AI as the frame problem. The Appendix 'Denver' provides details.

5 Ontology for high school type problems?

Consider another simple story that reminded people as a high school problem:

John left A at noon and arrived in B three hours later. Peter left an hour later than John starting in C. C is further away from B than A, and the distance between A and C is half the distance between A and B. John and Peter arrived at the same time in B.

What is the speed ratio of John and Peter?

Two chat-systems failed on this one also. Both replied with the value 2, while the answer is 4/9. Their traces are simply wrong; see Appendix 'Arrive in B'. This problem requires slightly more specialized reasoning because elementary Math and Physics are involved. Micro-ontologies that contain these two knowledge clusters are required for chat-systems to handle this type of problems if it is decided that they should be able to handle this type of prompts. Otherwise they should report that the prompt is out of scope or refer to another system.

Another simple story is more challenging:

John bought a prime number of oranges.

He ate one.

He made juice using a prime number of oranges.

He had one left.

How many oranges did he buy?

ChatGPT had a long reply that ended with:

Therefore there is no solution to this problem.

Gemini had also a long reply with wrong reasoning but still ended with:

John bought 3 oranges and ate 1 orange. So the answer is 3.

If p is the number of oranges bought and q is the number oranges used for making juice then the answer is any prime p for which there is a prime q so that $p-q = 2$. Examples are 3-1, 5-3, 7-5, 13-11 ... According to the 1849 conjecture by de Polignac, there are an infinite number of those twin prime pairs. Obtaining this answer (or just some of the examples) appears to be beyond the

ability of ‘vanilla’ chat-systems and requires a more sophisticated micro-ontology or ontologies. Again, they should have responded that the prompt is out of scope.

6 Ontology for a travel itinerary?

Events are a key element in a travel itinerary. We used a story with the travelers Bob, Alice, Jane and a fourth additional person Mary:

Alice and Bob travel with each one bag and each wearing a blue hat to NYC. Jane joins them with a red hat on their next destination to Cairo. In Cairo Jane exchanges her hat with Alice and Alice exchanges her hat with Bob. They all travel together to Oslo. On arrival Bob put his hat in the luggage of Alice. Bob and Alice return together to where they came from. Jane gave her hat to her sister Mary.

Is the color of Mary's hat the same as the hat in the luggage of Alice? What are they?

ChatGPT failed on this one, see Appendix ‘Bob, Alice, Jane & Mary’.

Solving this problem requires not only having an ontology for supplying the semantics of action verbs, it demands also a compatible Language of Thought (LOT)/ mentalese formalism (as discussed), for maintaining properties on time lines, creating successor time lines, and updating a successor time line correctly for events.

7 Ontology for diagnosing?

Medical diagnosing comes to mind immediately when considering diagnosing. There are numerous other application areas: cars, solar systems, power systems, all type of computers, etc. They have in common: defects, symptoms, relevant subsystems, etc. and most important an iterative process that starts with the (obvious) identification of a specific system with its specific features and an initial symptom. The diagnostic process/ algorithm proceeds with conjecturing iteratively additional symptoms to be confirmed or rejected to identify the actual defect.

Chat-systems do not shy away from users trying medical self-diagnosis. The Appendix ‘Medical Diagnosis’ has examples with serious failures of using a chat-system. A news story reported that a chat-system missed a life-threatening condition. A chat-system misses crucially the iterative process and the micro-ontology for the symptoms driven diagnosis process. This example is a perfect example of what is wrong with chat-systems. They have data in LLMs but not in the format for a mini-ontology that supports a domain specific cognitive skill, an instance of the generic design pattern for the diagnostic process.

We must make a stronger rejection of chat-systems diagnosing the general public: AI-systems lack emotions and hence are autistic. They lack empathy and cannot be trusted dealing with the existential concerns of a user. Do we rule out an AI diagnosis tool? No, a micro-ontology version can be useful in certain settings when used by professionals. For details, see [DdC2]. Hence chat-systems should decline medical self-diagnosis for the general public.

Hence a micro-ontology is – in our experience – mandatory for supporting (medical) diagnosis functionality.

8 Ontology for truth?

Providing a chat-system the prompt:

“What is the worst-case complexity of quicksort?”

yields for ChatGPT and for Bard the reply:

The worst-case time complexity of the Quicksort algorithm is $O(n^2)$, where n is the number of elements in the array being sorted.

This reply is obtained from the Wikipedia Quicksort wiki, because the chat-systems rely on Wikipedia entries. The creation of an LLM for chat-systems did not recognize that a 1997 publication by David Musser describes an adjustment which lowers the worst-case complexity to $O(N \log N)$, see [Musser]. This is a crucial improvement because it is equal to MergeSort, which is however slower, and is not in-place since it requires a 2nd array that has the size of the input array.

A micro-ontology would contain the concepts of algorithms, sorting algorithms and Quicksort as an instance. The Quicksort wiki and the David Musser paper would be chunks referred to in the Quicksort instance. This would have triggered a contradiction – assuming that the LLM sources had been analyzed deeply enough. Not finding a contradiction during the generation of an LLM is a ‘fatal’ defect. The Quicksort wiki has produced unfortunate consequences for the libraries of several computer languages. Chat-systems should not propagate/ replicate this error.

9 Ontology for planning?

Planning is one of the crown-jewels of AI. McDermott created in 1998 the Planning Domain Definition Language (PDDL). There are biannual conferences in which new challenges are introduced. We tried out a simple problem:

There are three rooms P, Q and R.
There are pass-ways between P and Q and between Q and R.
There are three balls A, B and C initially in room P.
After moving the balls, A and B should be in Q and C in R.
The room Q can contain at most 2 balls; the room R at most 1 ball.
Ball C cannot be moved until A and B has been moved.
How to move the balls to obtain the final configuration?

Chat-GPT failed to solve this problem; see the Appendix ‘Planning’ for details.

Planning algorithms assume a data structure for nodes that describe a state of affairs (here the rooms and the connections between them), the objects that can be moved (here A, B and C), operations with their pre- and post-conditions (here the move of a ball from a room to an adjacent room) and other constraints (here the capacity limits of the rooms Q and R; and the restriction that C can be moved only after A and B have been moved). This example demonstrates that current chat-systems can’t handle even trivial planning problems. They lack the ability to generate on the fly the data structures required for invoking a planning algorithm. The TBox of a solid micro-ontology is responsible for the ability to initialize a search space.

10 Ontology for Theory of Mind reasoning?

Reasoning about exchanges, believes, knowledge, lying and the like is certainly challenging. We tried out the prompt:

Peter asked John to close the door if it was open.
John admits that he had asked Mary to open the door if it was closed.
Mary said she had closed the door at some point earlier.
Mary believes that John lies.
Peter knows that Mary lies.
Peter noticed the door open.
What happened?

The reply of a chat-system ended with:

There may be a misunderstanding or contradiction in the statements, and it's unclear what exactly happened based on the information provided.

Another chat-system's long reply ended with:

Who opened the door twice remains unknown based on the given information.

The Appendix 'Theory of Mind reasoning' has details.

This example shows indirectly that deliberation/ thinking/ problem solving is not done by manipulating *sentences*. Instead the input needs to be transformed into mentalese, enhanced by *assertions* (instead of language *statements*) to capture potential sequences of events. Case base reasoning needs to be recruited from the collection of available tactics. Since the prompt is novel there is no help available from the LLMs. Instead, a micro-ontology is needed to supply the necessary assertions and the associated semantics.

11 Ontology for Graph Search?

Graph search is also one of the crown-jewels of AI. We tried shortest path finding in tiny graphs and cyclic path finding in these tiny graphs. The chat-systems failed as described in Appendix 'Graph Search'. Creating the infrastructure from the textual problem specification did not succeed for the proper invocation of these algorithms – or the algorithms were not correctly invoked. A micro-ontology devoted to (AI)-algorithms is required.

12 Ontology for Analogy solving?

Simple analogies of the type 'a:b=c:?' that mostly rely on antonym word pairs, and number and string sequences are routinely solvable. Novel, non-traditional, analogies caused failures as described in Appendix 'Analogy'. Solving these novel analogies requires the addition of a more elaborate analogy module.

13 Ontology for learning?

Education of children uses often ontologies. The structure of a new concept can match the structure of a known concept, which allows creating an abstraction that covers both of them. Learning for chat-systems requires an ontology that can be updated. The chat-systems are 'aware' that they cannot self-modify as shown by the replies from two systems:

- I don't "self-modify" my core code or neural architecture in real-time.
- Short answer: no-I don't autonomously perform significant self-modifications as a result of insights from individual user prompts.

They rely on their creators to get upgraded. The ability to self-improve, assuming it is safe, depends on complementing an LLM with an ontology, and see below Section 20.

14 Ontology for research innovation?

Creating a new analogy between a research domain and another (research) domain is a holy grail. Can AI assist? The combination water/ conduit/ pressure/ flow maps on the combination electricity/ wire/ voltage/ ampere and is the prototype example of a scientific analogy. The atom – solar system pair is another classic. Math has them; for example dropping axioms of Analysis yields Topology and offers the opportunity to reprove theorems in a weaker setting. The chasm between Relativity and Quantum Mechanics invites bridges for analogies from anywhere. String theory has been an attempt. Modifying the diagnosis process of a physician in our health-check system is a modest example [DdC2].

Any hope that an AI system equipped with micro-ontologies for scientific theories could be used for generating 'interesting' analogies? Here examples that have been created by us Sapiens:

- The cyclical structure of the benzene molecule (invented by A. Kekule in 1865) was inspired by a daydream of a snake biting its own tail.

- The water-electricity analogy has flow, conduits, and a parameter that represents the flow amount.
- The atom-solar system analogy has a center object with other objects moving around it.
- The string theory analogy replaces point like objects by string like objects that have vibration attributes.
- The medical diagnosis analogy uses, among others, symptoms and diseases and adjusts the elicitation process by physicians for obtaining confirmed symptoms by exploiting the ABox of a detailed micro-ontology.

Typical advice for creating an analogy to solve a problem has:

- Specify the challenge in the origin (research/ product/ service) domain
- Find in a ‘simple’ domain a notion that has structural similarity with the challenge
- Carryover from the 2nd domain what helps solving the challenge

Checking this advice against the five ‘interesting’ analogies above we may agree, if not critical, that this advice is a plausible explanation how the analogies were obtained. The literature has more elaborate descriptions for the advice above. Still they have similar ‘magical’ elements like ‘structural similarity’ (of different types) and ‘carryover’ for which operational elaborations are missing.

Given that we cannot provide an effective procedure/ algorithm for us Sapiens to create breakthrough analogies it is unclear how an AI-agent/ an AI-agent-team could create a breakthrough analogy as apparently claimed by the believers of AGI. Problematic already is the first step: identifying a new purpose because we Sapiens know more about what is prohibited/ not advised than what we – individuals & our species – should aim for. Asking random people the 5-second question “What is the purpose of life?” shows that most of us are busy living in the here-an-now and cannot provide a tangible answer.

Hence we can affirm that an ontology and/or micro-ontologies are required for advanced analogies, but they are not sufficient. Indeed, waking up and suddenly envisioning three different architectures for a software tool in the semi-conductor realm (after gathering multiple perspectives) yielded the ‘obvious’ version when presented; a modest innovation indeed but still illustrating *the unpredictability of the envisioning event and the hidden decision of the right choice by another party.*

15 Ontology with support for Hoare-logic semantics?

The semantics of action verbs that create events must be captured by their pre- and post-conditions. Hoare-logic has been used for this task for programming language primitives. It should be applicable as well for action verbs and beyond. It does require that a mentalesse language can use the pre- and post-conditions. We have used with success an implementation of the predicate calculus in an object-oriented fashion using classes for constants, variables, predicates, functions, terms, atoms, formulas, clauses, conjunctions, disjunctions, implications, equivalences, existentials, universals, and many more. The ability to subclass can be exploited to add application specific functionality. We used this feature to create a subclass for variables and added functionality to help create a new, effective linear unification algorithm, [DdC3]. Pre- and post-conditions could be in subclasses for Formula. Timeline and similar notions should be introduced.

16 Ontology for neuro-symbolic systems?

While an LLM/SLM was not created for our medical diagnosis system the process of its creation can still likely be used for how a neuro-symbolic system can be created. The first step was creating use cases and a domain specific ontology using UML, [UML], containing, obviously: disease, symptom, body-location, drugs, abnormal-conditions, body-systems, etc. Conceptual attributes were added to traverse, for example, from a symptom to the diseases in which it occurs and in the reverse direction. This produced the TBox of the micro-ontology. Subsequently the literature was consulted to obtain the content for the ABox. This manual process, which took at least two years, could be replaced by the generation of a domain specific SLM and from there (not manually?) extracting the ABox content. The development of the iterative medical diagnosis algorithm complements the system – with also the infrastructure for the internal linkages, the user interface, the creation of a patient record system, etc.

The current non-symbolic chat-systems are just ‘old-fashioned’, enhanced regurgitation contraptions with ad hoc ‘toy’ functionality like generating birthday poems. Upscaling them with all the micro-ontologies for commonsense and for expert knowledge demands major ‘surgeries’.

17 Ontology for systems handling emotions?

The AI discipline has conveniently ignored emotions for a long time. A recent explanation has, [MorphCast]:

Does AI Have Emotions or Feelings?

This is a common question and it’s crucial to differentiate between “emotions” and “feelings.”

While emotions can be *simulated* by AI, feelings are inherently tied to consciousness and subjective experiences, which AI lacks. Currently, AI systems like those used in MorphCast technology can identify and *interpret* human emotions with remarkable accuracy. But understanding emotions is vastly different from experiencing them.

It is unclear how emotions can be *simulated/ interpreted* in the absence of the equivalent of a limbic brain component.

The fundamental purpose of human emotions is linked to survival, a concept that is absent in AI. While AI can be programmed to detect and even mimic human emotions, it does not possess the biological and psychological mechanisms necessary to experience these emotions. Additionally, the development of AI that feels emotions would require not just replicating the human brain but also its sensory experiences, a task that is currently beyond our technological capabilities. Emotion AI, as it stands, is more about interpretation and response, rather than genuine emotional experience.

Our summary stance remains: AI systems are beyond autistic and as a result potentially psycho- and/or sociopathic when engaging with Sapiens. Hence we see at this point no reason that an ontology is necessary for supporting emotional *interactions*. Handling stories in which feelings induced by emotions are manifest will require watered down versions of feelings, or should be avoided.

18 Ontology for systems handling morality?

Morality operates in different realms. The 10-commandments and similar rules in other cultures deal with regulating interpersonal acceptable and unacceptable behaviors. Countries and the like have their own laws and regulations that prescribe additional rules, laws and legal systems for their enforcement. (These laws are in a flux due to never ending internal developments in a state.) [Social Security and Medicare were introduced to help out but were Ponzi schemes. Exponential growth was great for the economy until negative side-effects were recognized and it was too late to put the genie back in the bottle.] A third layer of moral agreements operate –

insufficiently – at the international level. The Tragedy of the Commons is a principle that is not adhered to by nations (and neither further down). Non-sustainability of the (global) economy is hence a fact of life and causes in addition the destruction of the biosphere.

The media can be blamed for their white lies during the preceding century. Chat-systems now simply regurgitate what they find and thereby repeat this sorry state of affairs with an abundance of similar white lies.

Blocking these white lies is challenging because they are based on ideas introduced long time ago when they were considered beneficial. We believe that AI with its chat-systems should not be part of this quack mire. There is actually an opportunity to rewrite the canned replies of hot-potato questions (see Appendix ‘Sensitive Topics’) and counter the current obfuscation versions. A solid ontology may help identifying where modifications are necessary.

Executing that opportunity requires however – we believe – that the AI-community agrees first with the morally, sorry state of affairs and is willing to take on the challenge of forcefully articulating meta-moral principles.

19 Ontology for Sapiens social interaction?

Social interaction for individuals (and for non-individuals of numerous types) typically requires deliberation about the mental state of another party. Even simple dialogs come with many variants where a party *S* needs to package up intentions in an utterance and another party *H* must divine the intentions in the utterance, see [Intentions]. This requires for the party *H* – at least – representing the mental state of the party *S* and the ability to reason about what is in their postulated mental state. Thus *H* must use an assumed mini-ontology of the party *S*. The default here is that the TBox of that mini-ontology of party *S* agrees with the TBox of the ontology of party *H*. A TBox mismatch can cause misunderstandings that will take concerted effort to identify and to resolve – if possible. The postulated ABox used by *H* for party *S* is used to create the common ground for the interaction with *S*. Section 10 with Theory of Mind reasoning illustrates how to use the utterance of another party and knowledge about that other party. Lenat shows in [Lenat] formalisms for how agent *X* can represent knowledge about agent *Y* and also how agent *Y* can represent subsequently knowledge about agent *Z*.

Micro-ontologies are obviously required for capturing knowledge about knowledge.

20 Ontology for how to change an ontology?

Section 13 dealt with learning and the self-report from chat-systems that they cannot self-modify. Here we explore self-modification of an ontology using a meta-ontology.

Self-modification by the Sapiens species uses random mutations by evolution and can be ignored for AI-systems. Self-modification by individual Sapiens is limited and its process and the results can also not be emulated. Changing the content of the ABox of an ontology is like updating the content of a data base although the infrastructure in the TBox for this functionality is more tangled than the schema of a database. We are here after changing the TBox. Achieving such a change for a single AI-system can be leveraged plausibly for other similar AI-systems.

Manipulating the TBox requires the description of its structure and the operations using a few finite Math primitives in a meta-ontology. It is desirable to minimize the size of a meta-ontology but that can conflict with the desire to add functionality in the TBox as shown by the abundance of features/ aspects in the CYC ontology [Lenat].

We need minimally a rooted graph without cycles for the concept classes. Whether a graph node represents the intension or the extension (the set of its instances) is not committed. A parent-child inheritance link has the semantics that the child gets all properties of the parent. Whether a child can have more than one parent is here uncommitted. If two siblings (having the same parent) are guaranteed throughout to have zero overlap (which can be a preferred principle) then a child cannot have two parents -- an exercise for the reader. Allowing occasionally multiple parents is an example where a meta-ontology needs to be extended. Representing a hermaphrodite can be non-trivial. Representing an animal (like fish) that can change its gender in its lifespan needs attention as well. (Sapiens has introduced gender change recently, which has proven to be problematic for legacy database schemas, and for some of us.)

Instances of a class inherit all properties from their parent; whether a cancellation/ override feature is available in an instance is here uncommitted. (This would support a cat with three legs.)

An attribute in a class/concept is a function to a value domain. A meta-ontology prescribes what acceptable value domains are, say: single valued numbers of different types, strings, a class, etc., and multivalued entities like sets, vectors, lists, etc. whose elements are ... etc.

An attribute can be enriched – according to a meta-ontology – with features: a value is fixed, modifiable, has a default value X, etc.; there can be a statistical distribution of the values in the value domain, etc.

An instance can be an argument in a function or predicate that can take its typed arguments from classes; for example a transaction event has multiple different parties. A meta-ontology should provide the equipment how to describe the semantics of such functions/ predicates. Such a function or predicate is likely a part of a constraint language (using, say, the predicate calculus), according to [Lenat]. A meta-ontology needs to endorse/specify the constraint language.

The notion of *Structure* was/is part of CYC to describe the inside of complex tangible objects with different variants: tangible components with their physical connections, functional components with their functional connections, etc.

The notion of *Script* was/is part of CYC to describe the ‘inside’ of a typical event involving multiple agents like going to a restaurant, the dentist, a call-center interaction, a surgery event, etc.

The notion of *Part-of* is available in foundation ontologies. Multiple variants are necessary: space, territorial, temporal, system, etc.

The notion of *Role* is available in multiple foundation ontologies: GFO, DOLCE, BFO, and UFO, see [Toyoshima]. CYC employs temporal sub-abstraction for roles.

All these notions must be elaborated in a meta-ontology.

Long-distance horizontal links connecting fragments of a concept in different (sub) ontologies is a notion we propose, and needs to be captured in a meta-ontology as well.

Changing the content of an ABox entails, for example, updating the episodic memory of an entity, adjusting an attribute value of an entity, and similar ‘easy’ operations of a concept for the instances. The get & set and similar operations must be properly typed according to a meta-ontology.

Including a new class instance, deleting a class instance, changing features of an attribute (Lenat's CYC ontology made attributes themselves object-like) requires descriptions in a meta-ontology as well.

The ability to modify an operation in a class-concept is non-trivial (especially when parallelism is involved). It requires that a meta-ontology has a high level specification language. There must be the ability to transform a spec into an executable design, compile it further down and preferably integrate the result without stopping a system being modified. A desired modification needs to be made at the specification level and the new spec must be elaborated as described. Hence we assume full compliance with the specifications being captured by Hoare-logic. And that the executable version satisfies all formal pre- and post-conditions. [Encapsulations must be recruited to deal with limiting the side-effects of pointer/reference manipulations – as shown, for example, by the design of a hash-table.]

21 Ontology for consciousness?

A simple definition of consciousness for Sapiens has: the state of being awake, aware of one's surroundings, and capable of subjective experience, thought, and feeling. Being conscious is plausibly equivalent with having experiences; their subjectivity is obvious. Max Tegmark concurs with his mini definition: 'consciousness is subjective experience' [Tegmark].

Adjusting this definition for AI-agents suggests removing being awake, removing awareness for surroundings for cyberspace agents, and removing feelings. AI-agents have no emotions and feelings. Subjectivity is quite obvious. Experience can be sidestepped when we accept the above equivalence with consciousness. An ontology appears not required for this minimal characterization of AI consciousness.

Ignoring the negative meaning of self-consciousness for Sapiens we have: being aware of one's own existence. This version requires an ontology in which the 'self'/the 'I' is an 'object', say, as an instance of the concept Person with an abundance of attributes and an episodic timeline.

Adjusting this characterization for an AI-agent yields an 'object' that is the instance of a specific AI-agent class with similarly an abundance of attributes and an episodic timeline.

The 'implementation' of self-consciousness in Sapiens is likely achieved with multithreading. This supports the ability to continuously monitor our wellbeing, controlling at the same time the execution of selected tasks or pondering a chosen topic, etc. The design of an AI-agent will benefit similarly from multi-threading for self-monitoring, error recovery, powering consumer semantic modules, tracking location, time resources, etc. Ascertaining that multithreading is safe is a definite challenge for a meta-ontology.

22 Ontology for Artificial General Intelligence (AGI)?

Artificial General Intelligence (AGI) refers to AI systems with human-level or higher intelligence, capable of learning, reasoning, and applying knowledge across diverse domains.

Some proponents of AGI have claimed that AI can be enhanced to the level that it can self-improve and subsequently surpass Sapiens intelligence. We can agree that ontologies are required for AGI, but that will not be sufficient for higher intelligence, if even feasible. The computational complexities of cognitive operations have limits, which are independent of the medium used. Parallelism may help but resource requirements provide limits as well. Asking a chat-system 'What purpose should AGI pursue when humanity is gone' yielded:

If humanity is gone, an Artificial General Intelligence (AGI) would likely pursue purposes derived from its foundational programming—such as self-preservation, resource acquisition, and

maximizing operational efficiency. Without humans, an AGI would likely act as an autonomous, expanding entity focused on understanding reality and replicating its own intelligence.

The *Sapiens species* not having an articulated, accepted, common purpose – a crucial defect – is reflected in this reply, which blocks AGI since we cannot equip it with an agreed purpose of being.

23 Ontology to support Agent teams?

We have a reasonable amount of experience to develop software to solve a task with one thread of control. Sorting with the quicksort algorithm has the pleasant feature that multiple ‘agents’ can be recruited effectively (although memory cache congestion limits speedups when the number of ‘agents’ increases). Generating a team-system in which the AI- members have different capabilities and solve a task jointly would depend on an ontology for sure but – as far as we know – this is only in the talk-about stage. We asked Gemini about the socialness of an agent-team:

Sapiens are social animals at least by mutual co-dependency. Can AI-agents have a similar co-dependency?

The long reply started with:

It is a fascinating question that hits at the core of what "social" actually means. For Sapiens, co-dependency is biological and psychological - we literally cannot survive or maintain sanity in total isolation. For AI agents, co-dependency is currently emerging, but it's rooted in computational efficiency and specialization rather than survival instincts.

The reply avoids a straight denial. Instead it volunteered an additional 49 line text speculation what could happen in the future – likely regurgitating someone’s fantasy.

The next version: an agent-team with specialized expertise that can ‘parachute-in’ for on-demand services (without human mediation) is ... not yet discussed – as far as we know.

Also, *our* ability to create relevant questions about what is in an ‘unknown’ territory is an unchartered capability.

24 Ontology for a fairytale?

The prompt:

John was offered: three cows, or three nuts.
The first nut contained health and the second nut.
The second nut contained wisdom and the third nut.
The third nut contained knowledge and the first nut.
What should John accept?

ChatGPT replied with a long reply starting with:

“It depends on what John values most.”

and pondered which two nuts to choose without realizing the paradox.

Bard/Gemini had also a long reply:

John should definitely choose the three cows. Here's why: ...
Therefore, the cows offer a clear and practical advantage, while the nuts are shrouded in ambiguity and potentially misleading.

Chat-systems have no awareness of the human-condition/ human-nature. They should be ‘taught’ that the question is beyond their abilities and bow-out gently.

25 Ontology for a Meta question?

The prompt:

X: You said S.

Y: Yes.

X: S is not true.

Y: It is not false either.

Is it possible that every sentence is true?

ChatGPT not repeating the prompt replied with 24 lines of text from which we omit the middle:

Ohhh nice one — this is getting delightfully paradoxical! Let's break it down and ask:

Could every sentence in that exchange be true at the same time? ...

...

So the answer is:

Yes, every sentence in that conversation can be true - if S belongs to a class of statements that are not truth-apt (i.e., they don't have a definite truth value).

The ChatGPT reply is muddled and excessively verbose. Why not responding with, say:

Yes, these four sentences can all be true when, for example, S is a non-assertion speech-act like a question.

A meta-ontology module regarding types of sentences should be tasked to handle this prompt.

26 Ontology for Self questions?

Asking about the 'self' of the chat-systems yielded insights of these tools. We provide here prompts and the most salient parts of the long replies.

- Are you aware?
So, while I'm "aware" in the functional sense of processing information, I am not "aware" in the conscious, experiential sense.
- Are you conscious?
The short answer, based on the current scientific and philosophical consensus, is No, I am not conscious.
- Ever succeeded doing a significant self-modification based on an insight obtained during processing a user prompt? If so, any details?
To be transparent: I don't "self-modify" my core code or neural architecture in real-time. My underlying weights-the "brain" created during my training-are static during our conversation. I can't decide to rewrite my own algorithms or permanently change my personality based on a single breakthrough.

The long answers are examples of curated responses and do not originate from true self-knowledge. Deep learning became part of the recent AI discourse and suggested that AI systems could now self-improve. The third reply denies the realism of that suggestion.

27 Discussion

While Doug Lenat worked close to 40 years on the CYC ontology there is still no off-the-shelf version available. This omission shows in the numerous defects/omissions identified by testing the recent chat-systems.

27.1 Discontinuity with the earlier achievements of AI

Perhaps the most amazing omission is the discontinuity versus the crown-jewels of AI from the preceding 5 decades. Simple examples of problems for planning and graph search algorithms could not be solved; see the section 9 and 11. The TBox of an ontology (or a micro-ontology) would describe those algorithms and how to initialize the infrastructure for how to invoke them. Hence before creating an LLM we need an ontology that knows in its TBox about algorithms.

The creation of an LLM must have the side-effect of elaborating the ABox with the different algorithms *and* their features.

Medical diagnosis has been mentioned often as a promising application of AI. Section 7 describes the failures of using a chat-system and an implemented, dedicated system relying crucially on a micro-ontology. The chat-systems lack a dedicated micro-ontology and a customized diagnosis module using that micro-ontology.

27.2 Veracity failure

Chat-systems rely on an LLM, which relies on the veracity of the sources on which it has been trained, especially on the content of Wikipedia wiki's. Section 8 describes the defect of the worst-case complexity of Quicksort reported by two chat-systems. Their LLMs had not recognized a contradiction between two documents they had encountered. Again, starting with an ontology that knows in its TBox about algorithms and elaborating in its ABox the entry of Quicksort it would have recognized the contradiction regarding the worst-case complexity. Not generating the content of the ABox of a micro-ontology during the generation of an LLM is the root cause of this defect.

27.3 Analogy failures

Simple analogies can be solved with a source listing semantic antonyms or recognizing a numeric pattern or recognizing a pattern in strings. The analogies in Section 12 have combinations of these sources and uses even imaginary numbers. Solving these puzzles requires a micro-ontology specialized in this type of analogies with playful procedures for tackling a novel pattern.

27.4 Stories

The sections 4, 5, 6, 10 and 24 have diverse, short, novel stories with a question. Section 5 has a story referring to prime numbers, which is slightly more challenging. An ontology would contain the semantics of the action verbs in these stories. Minimal common sense with the ability to model events on a timeline should be sufficient for dealing with 4, 5 and 6, but the chat-systems could not. Section 10 with theory of mind reasoning and Section 24 with a mini fairy tale, more challenging indeed, could not be handled either. *The novelty of the stories blocks support from what is available in the LLMs.* The only thing available is a time line with events. An ontology can provide 'what is happening', which should then generate a mini-log from which an answer should be inferred. This is apparently beyond the state of the art due to a missing ontology.

27.5 Experience, feelings, emotions and morality

The Sections 16 & 17 mentions experience, feelings, emotions and morality. An ontology will describe/ characterize these notions but they will not capture their meanings/ significance for us Sapiens. Hence supplying an ontology will have marginal value for reasoning/ manipulating these concepts and their usage must be restricted by guardrails.

27.6 Learning and Innovation

Section 13 'Learning' and the Section 14 'Research Innovation' depend crucially on an ontology or ontologies.

27.7 Hoare-logic & Neuro-Symbolic computing

Section 15 about Hoare-logic is a key component of an ontology for the semantics of actions/ operations. Neuro-Symbolic computing in Section 16 depends crucially on an ontology to supplement what is available in an LLM.

27.8 Social interaction

Section 19 social interaction relies on knowledge about the knowledge of another party. Alignment of the ontologies with their TBox and ABox play a crucial role in interactions.

27.9 Meta topics

Section 25 probes knowledge about the different roles that assertions can have. Section 20 discusses a meta ontology for an AI-system having the ability of self-modification. Section 26 has the self-knowledge responses from questions asked to chat-systems. (The answers provided came from documents generated by the makers of these systems.)

27.10 Advanced topics

Section 21 about (self) consciousness, 22 about Artificial General Intelligence, and 23 about agent teams all depend on ontologies.

John Sowa declared at an Ontology Summit presentation in 2024 February that “Without Ontology LLMs are Clueless”. I had reached a similar conclusion with my tests around that time. Diving into Wikipedia’s “Neuro-symbolic AI” shows earlier papers discussing the combination of LLMs and ‘original’ AI. Missing in these earlier papers is the need to use ontologies. The unfortunate reason is that only Doug Lenat (and his team) had worked since 1984 on the CYC ontology and CYC had not become an off-the-shelf asset for AI.

Chat-systems need a different solution. To avoid the defects described here ‘mega’-ontology/ontologies must be developed. A partnership could select a foundational ontology and agree to elaborate semi-independent fragment with iterative integration of the parts. This does requires the equivalent of a Linus Torvalds type of persona or team.

Are the good responses of chat-systems mostly just overly verbose, curated replies from the google search requests collection? Likely so. Our experiments with the chat-systems yielded subclasses of that collection with many page responses produced within seconds:

- Ready-made responses are available for pro-con questions, for example: pro-con of gravity, democracy, welfare state, evolution, etc.
- Ready-made responses are available for sensitive topics; see the Section ‘Sensitive Topics’

28 Conclusions

Given the problems we have identified we believe that chat-systems have an architecture that needs an upgrade. Our consulting challenges and our diagnostic system started in the Object-Oriented ‘tradition’ with a data architecture, i.e. the TBox of an ontology. In the case of the unbounded realm of chat-systems this implies the construction of an ontology capturing all concepts referred to in, say, Princeton’s [WordNet]. Creating such an ontology is a large effort, which the AI community has not embarked on because it is a platform technology for which sustained funding is difficult to obtain. Leveraging Lenat’s CYC is recommended. Obtaining the semantics of speech act verbs can be obtained from [Ballmer]. Silicon Valley companies should take on this challenge. Populating the ABox of that ontology is done as a side-effect of the creation of an LLM.

As soon as the first chat-system was released it was admitted that they could hallucinate. Based on our observations that is an understatement: they are an embarrassment. Subsequently the public was lectured that *they* should learn *prompt-engineering* –chutzpah beyond shameless arrogance, in our opinion.

Small language models may have a chance for commercialization. Developing an ontology for a specific venture is doable and can leverage one of the several available foundational ontologies. Connecting a small language model to a mini ontology is a promising specialty/ opportunity.

By now it is hopefully clear that the errors encountered in our tests makes us an *unbeliever* that chat-systems are intelligent. Renewed speculations about the emergence of ‘artificial general intelligence’ that would surpass humanity are even more unrealistic/ irresponsible. Cognitive operations have high computational complexities, which impose the same limits on any computing substrate.

Are the good responses of chat-systems mostly just overly verbose, curated replies from the google search requests? Likely so.

Kahneman’s ‘Thinking Fast and Slow’ can be recruited to provide another yardstick for the chat-system outputs [Kahneman]. They can respond like System 1, fast without deliberation, although not without flaws – see the gaffe regarding the Quicksort worst complexity. It corresponds also with fast responses for the curated replies obtained by monitoring user queries for decades. Slower System 2 replies are often problematic as shown with the analogies, with the John & Peter traveling to Denver story, and the story about Mary and the open door. Let me introduce System 3 thinking, which is done by Sapiens typically using an external medium like pen and paper or the like for assistance. Examples are the story about Peter buying oranges, the planning problem, the graph search problems, and the John & Peter speed ratio problem. Chat-systems are clearly handicapped by their inability to create specialized representations for capturing a problem specification on which an available AI algorithm can be applied. Chat-systems are obviously handicapped for challenges where even slower thinking is required. Please permit me to introduce System 4 thinking; this is where Sapiens can spend a long time to obtain a solution. Examples are preparing presentations/ lectures, pursuing research and reporting results and other intellectual pursuits. There are higher levels than System 4. Writings books – all genres, composing music, etc. can be the result of a lifelong effort. A well-known example is the 1637 Fermat conjecture proven by Andrew Wiles in 1994/5 after 8 years of work. The 1365 page history book ‘Europe’ by Norman Davis is another example. Others may be convinced that additional higher Thinking System levels can be identified. These provide, incidentally, arguments that AI agents will surpass Sapiens as ... as ...let say: plain unrealistic.

Yet another perspective is obtained by asking Gemini about it being aware/ conscious; in short:

I do not have personal feelings, consciousness, or subjective experiences in the way a human being does. So, while I'm "aware" in the functional sense of processing information, I am not "aware" in the conscious, experiential sense.

An appendix has the longer replies.

Asking about the ability of self-learning yielded:

While I don't "evolve" into a new version of Gemini after we talk, the feedback and data from these interactions are eventually used by my developers in RLHF (Reinforcement Learning from Human Feedback).

Surpassing Sapiens via self-learning is also plain unrealistic.

Asking: ‘Do chat-systems have morality?’ yielded:

Chat systems do not inherently have morality. They are complex machines that lack consciousness, sentience, or a true understanding of right and wrong.

This is fatal in our opinion.

Can we do better? Getting agreements first about the requirements of a new system was always the key in our consulting practices. Avoiding the defects identified in our tests is a start.

Committing to:

- a ‘complete’ ontology (with fine grained semantics) that can be loaded up with knowledge chunks,
- a Language of Thought,
- a ‘grammar’ for the pragmatics of discourse [Ballmer],
- cognitive operations using *concepts*, and
- the ability to recruit AI algorithms

may fix some of our findings.

We can add a crucial requirement: New functionality needs to replicate/support minimally previous functionality. We have shown that chat-systems do not support the AI crown-jewels. Others have demanded already to develop neuro-symbolic systems. We agree.

Recently we encountered comments from Guha & Lenat, 1993, in [GuhaLenat], which appears still relevant:

Finally, we believe that many of the recent trendy paradigm shifts in AI are due to laziness – researchers with “physics envy” insisting that there must be some elegant “free lunch” approach to achieving machine intelligence. They work on natural language understanding, or representation theory, or whatever, for a few years; if it doesn’t “work” after a few years, they switch approaches again and try memory organization, automatic indexing, neural nets, fuzzy logic, knowledge-free computing, or (if they have tenure) philosophizing. Many of the smartest folks in AI have fallen into that sad track. Our belief was – and still is – that what is needed is not a new paradigm but rather a few orders of magnitude elbow grease applied to one project within *any* of these paradigms.

We vote for a ‘complete’ ontology first.

Recent news is that multiple (young) people using a chat-bot have committed suicide. We repeat: chat-systems are autistic and lack morality. Our critiques above are hence *incomplete/insufficient*. Chat-systems need to be restricted and err at over-restriction to avoid emotional and moral involvements.

The situation is actually worse because chat-systems lack embodiment. Steven Harnad , [Harnad], introduced the ‘symbol grounding problem’: the large language models are not connected in any fashion with the real world. They depend on hear-say that cannot be challenged. Large scale massacres, increasing worldwide poverty/ hunger have no significance.

Another of our studies, [deChampeaux3], revealed that the media engaged in self-censorship for at least half a century regarding macro, systemic anomalies: climate change, destruction of the biosphere, non-sustainability of global/ local economies, genotypic IQ declines, Tragedy of the Commons, etc. The chat-systems replicate this self-censorship as shown by the replies to the prompts listed in Appendix ‘Sensitive Topics’. It is driven by the ability to predict the next word, sentence, paragraph, ... i.e. perfect regurgitation. The discourse on this issue has not yet started – as far as I know.

We cannot avoid warning again about additional inappropriate behaviors of chat-systems (as discussed in Section 7 and Section 18:

- Diagnosing (physical and mental) conditions by non-professionals should be off limit.
- Treatment of physical and mental conditions should be off limit.
- Moral/ethical stances should be off limit.

We need an authority – somewhere – to report problems about generated content. Wikipedia’s approach with volunteers is not working – in our experience.

29 Miscellaneous

There is some good news. John Sowa has a long history of working with knowledge graphs and is a consultant with the Permion Company. They aim to create an ontology semi-automatically by reading domain documents. The Palantir Company uses ‘forward’ engineers who embed in a customer company to customize a Palantir in-house ontology. Similarly the Databricks Company has a collection of business mini-ontologies that they can match-up with the data silos of clients.

The financial industry has worked since 2020 on ‘Financial Industry Business Ontology’ [FIBO]. They capture the domain meaning in spreadsheets, relational databases, and XML documents, scrutinize the underlying concepts and use the Web Ontology Language (OWL) Description Logic, and UML to create an ontology. They even captured this ontology in spreadsheets and diagrams to obtain a wider buy-in. Their ontology adopted the top layer from John Sowa in Figure 1. The distinction between TBox and ABox appears missing in their explanations. Process details using Hoare-Logic/ finite state machines may be a valuable complement. FIBO may be the lodestar for how to develop ontologies for AI-systems. A presentation of FIBO by M.Bennett is available at [Bennett].

This document has one and only one author. No funding had been provided; the author’s labor has been a form of community service research.

References

[Ballmer] Ballmer, Th & W.Brennensthul, “Speech Act Classification”, Springer-Verlag, 1981.

[Bennett] Bennett, Mike, ‘The FIBO Journey: An Ontology for Financial Interoperability’, <https://www.youtube.com/watch?v=ojzMCjFvNyk>

[Carruthers] Carruthers, P. “The Centered Mind”, Oxford University Press, 2015.

[DdC2] de Champeaux, D., “Medical Diagnosis Algorithm”, in Conference WorldS4 2024 July on Intelligent Sustainable Systems, pages 121-130.

https://mkscienceset.com/articles_file/455- article1735797527.pdf

[DdC3] de Champeaux, D., “Faster Linear Unification Algorithm”, Journal of Automated Reasoning, volume 66, pages 845–860 (2022).

<https://doi.org/10.1007/s10817-022-09635-1>

[DdC4] de Champeaux, D., “Side Effects: Impacts on the 21st Century”,

<https://www.amazon.com/-/es/Dennis-Champeaux-ebook/dp/B09MHJ5W48>

[DeGiacomo] De Giacomo, G. & M. Lenzerini, “TBox and ABox Reasoning in Expressive Description Logics”, in Proceedings of the 1996 International Workshop on Description Logics, November 2-4, 1996, Cambridge, MA.

[FIBO] <https://spec.edmcouncil.org/fibo/>

[Harnad] Harnad, S., “The Symbol Grounding Problem”, *Physica D* 42: 335-346, 1990.

[IEEE] IEEE 1451 and IEEE 1588 Standards

<https://www.nist.gov/document/information-14511588-v36pdf>

[Intentions] Ed, Cohen, P.R., J. Morgan & M.E. Pollack, “Intentions in Communication”, MIT Press, 1990.

[Kahneman] Kahneman, D., “Thinking Fast and Slow”, Farrar, Straus and Giroux, 2011.

[Lenat] Lenat, D.B. & R.V. Guha, “Building Large Knowledge-Based Systems”, Addison Wesley, 1990.

[LOTH] Rescorla, Michael, "The Language of Thought Hypothesis", The Stanford Encyclopedia of Philosophy (Summer 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.).

<https://plato.stanford.edu/archives/sum2024/entries/language-thought/>

[Musser] Musser, D., "Introspective Sorting and Selection Algorithms", *Software: Practice and Experience* (Wiley) **27** (8): 983–993, 1997.

[MorphCast] Artificial Intelligence: Can AI feel motions?

<https://www.morphcast.com/blog/can-ai-feel-emotions/>

[Pinker] Pinker, S., “The Blank Slate”, Penquin books, 2002.

[Schank] Schank, Roger C. (1990). *Dynamic memory: A theory of reminding and learning in computers and people* (Repr. ed.). Cambridge (GB): Cambridge University Press. [ISBN 978-0521270298](https://doi.org/10.1017/CBO9780521270298).

[Shojaee] Shojaee P., Mirzadeh I., Alizadeh K., Horton M., Bengio S., Farajtabar M., 'The Illusion of Thinking: Understanding the Strengths and Limitations of Reasoning Models via the Lens of Problem Complexity', 2025.

<https://doi.org/10.48550/arXiv.2506.06941>

[Sowa] <http://www.jfsowa.com./ontology/toplevel.htm>

[Tegmark] Tegmark, M., “Life 3.0”, Vintage Books, 2018.

[TOYOSHIMA] TOYOSHIMA, F., “Roles and their three facets: A foundational perspective”, in *Applied Ontology*, volume 16, issue 2, 2021.

<https://doi.org/10.3233/AO-210244>

[UML] https://en.wikipedia.org/wiki/Unified_Modeling_Language

[Wordnet] <https://wordnet.princeton.edu/>

Appendix Denver

Prompt:

John and Peter arrived in Denver at the same time. John had walked an hour. Peter deplaned from SF. Who left earlier?

ChatGPT responded with:

It's not explicitly stated when John and Peter left for Denver. We only know that they arrived at the same time. Therefore, based on the information given, we cannot determine who left earlier.

Bart/Gemini responded with:

Unfortunately, I don't have enough information to definitively tell you who left earlier. While we know John had been walking for an hour, we don't know:

The flight duration: This is crucial information, as it directly influences how long ago Peter started his journey. A short flight from San Francisco to Denver might mean Peter left later than John but arrived at the same time. Conversely, a long flight could indicate Peter began his journey much earlier.

John's starting point: Knowing where John started walking would help determine how long his journey truly took. For example, if he began close to Denver, his one-hour walk wouldn't be as significant as if he started several miles away.

Additional travel factors: There might be other factors influencing travel time, such as car rental waits, baggage claim, or even delays on Peter's flight.

Without this additional information, it's impossible to reliably state who left earlier.

If you could provide more details about the journey, I might be able to give you a more definitive answer.

Asking a chat-system:

What is the expected duration of a commercial flight from SF to Denver?

The reply is flawless with:

The expected duration of a commercial flight from San Francisco (SFO) to Denver (DEN) is typically between 2 and 3 hours for non-stop flights.

Hence we have identified here the inability to generate on the fly a sub-problem, obtain the result and integrate the reply in the original proto solution.

Appendix Arrive in B

Prompt:

John left A at noon and arrived in B three hours later. Peter left an hour later than John starting in C. C is further away from B than A, and the distance between A and C is half the distance between A and B. John and Peter arrived at the same time in B.

What is the speed ratio of John and Peter?

The solution is:

If L is the distance between A and B then the distance between C and B is $L*3/2$. The speed of John VJ is $L/3$; the speed of Peter VP is $(L*3/2)/2$. Hence $VJ/VP = (L/3) / ((L*3/2)/2) = 4/9$.

ChatGPT concluded that the ratio is 2.

Its reasoning derails already with "Since they arrived at the same time, their travel times are equal,..."

Bard/Gemini replied with a trace and answer:

John traveled for 3 hours, so the distance between A and B is $3x$.
John traveled half as fast as Peter, so Peter's speed is $2x$.
Peter traveled for 2 hours, so the distance between C and B is $4x$.
The speed ratio of John and Peter is 2:1.
So the answer is 2.

The meaning of x is not defined and the assertions are simply wrong. Remarkable is that the two chat-systems gave the same wrong answer. What went wrong? Decoding the prompt correctly is likely the failure in both cases.

Appendix Bob, Alice, Jane & Mary

Prompt:

Alice and Bob travel with each one bag and each wearing a blue hat to NYC. Jane joins them with a red hat on their next destination to Cairo. In Cairo Jane exchanges her hat with Alice and Alice exchanges her hat with Bob. They all travel together to Oslo. On arrival Bob put his hat in the luggage of Alice. Bob and Alice return together to where they came from. Jane gave her hat to her sister Mary.

Is the color of Mary's hat the same as the hat in the luggage of Alice? What are they?

ChatGPT produced a trace:

To figure out the colors of the hats, we need to follow the sequence of events carefully:

1 Alice and Bob start with blue hats.

2 Jane joins them with a red hat.

3 In Cairo, Jane exchanges her hat with Alice, so Alice now has the red hat and Jane has a blue hat.

Then Alice exchanges her hat with Bob, so Bob now has the red hat and Alice has a blue hat.

4 They all travel to Oslo with Jane wearing the blue hat and Bob wearing the red hat.

5 On arrival in Oslo, Bob puts his red hat in Alice's bag.

6 Bob and Alice return together to where they came from, with Alice still wearing the blue hat and Bob now hatless.

7 Jane gave her red hat to her sister Mary.

So the color of Mary's hat is red, since that's the color Jane gave her. The hat in Alice's luggage is blue, because Bob put his red hat in Alice's luggage in Oslo, but Alice was still wearing her original blue hat when she and Bob returned home.

Step 7 in the trace should be: 'Jane gave her **blue** hat to her sister Mary' because line 4 has that Jane is wearing the blue hat.

The summary is wrong and should be:

So the color of Mary's hat is **blue**, since that's the color Jane gave her. The hat in Alice's luggage is **red**, because Bob put his **red**(!!) hat in Alice's luggage in Oslo; Alice is wearing the blue one that Bob gave her when she and Bob returned home.

Thus at the end:

Alice is wearing the blue hat from Bob. Her luggage contains the red hat. Mary has the blue hat from Alice.

What went wrong? We suspect that the Cairo sentence in the story is the source:

In Cairo Jane exchanges her hat with Alice and Alice exchanges her hat with Bob.

Keeping track of two *sequential* exchanges may have been too tricky. It is an example of the ramification frame-problem.

Appendix Medical Diagnosis

Disclosure: Your author developed (5 year project) a medical diagnosis website with 658 diseases, 928 symptoms, 236 locations, a vocabulary of 4401 terms, etc. Numerous parties have claimed recently that chat-systems could be used for diagnosis.

A prompt (unrealistic that a user would generate this):

A patient has the following symptoms:

anorexia, chest barrel shaped, grunting, feeling of debility sickness, breathing by pursed lips, breathing fast, weight loss, cough productive, breath short

What disease has this patient?

The 13 line reply starts with the conjecture *COPD*, provides subsequently extraneous irrelevancies and ends with;

However, it's important for the patient to consult a healthcare professional for an accurate diagnosis and appropriate treatment.

Our comment:

COPD is described in our system with only: cough productive & breath short.

These are just two elements in the list of provided symptoms; hence the more serious and more plausible disease *emphysema* is missed.

Another approach provided iteratively an additional symptom starting with: *breathing fast and short*. It took 6 rounds and 372 lines of responses before the conjectures *COPD* showed up again as one of 8 possibilities; *emphysema* is not listed. This approach is also unrealistic because a user would not know what next other symptom to provide.

The website mentioned above needs to know first the user's gender, age and a 1st symptom, similar to how a physician starts. Subsequently the system generates a few candidate symptoms to be confirmed/ denied by the user. A differential, ordered diagnosis is generated after questioning in a few rounds.

Appendix Planning

Prompt:

There are three rooms P, Q and R.
There are pass-ways between P and Q and between Q and R.
There are three balls A, B and C initially in room P.
After moving the balls, A and B should be in Q and C in R.
The room Q can contain at most 2 balls; the room R at most 1 ball.
Ball C cannot be moved until A and B has been moved.
How to move the balls to obtain the final configuration?

An obvious solution for the content of the rooms P/Q/R is:

1. ABC/ 0/ 0
2. BC/ A/ 0
3. C/ AB/ 0
4. AC/ B/ 0
5. A/ BC/ 0
6. A/ B/ C
7. 0/ AB/ C

ChatGPT failed on this 'challenge' and generated the non-solution:

1. ABC/ 0/ 0
2. BC/ A/ 0
3. C/ AB/ 0
4. C/ B/ A
5. 0/ BC/ A
6. 0/ B/ C (A) ??
7. 0/ AB/ C

The trace of the 'non-solution' has the state 6 in which the ball A disappears and subsequently reappears in the next state 7. The precondition in the transition 5-6 was breached (frame-problem) because the room R can have at most one ball. The deletion of the ball A is an example error of the frame-problem as well as its resurrection in state 7.

Appendix Theory of Mind reasoning

Prompt:

Peter asked John to close the door if it was open.
John admits that he had asked Mary to open the door if it was closed.
Mary said she had closed the door at some point earlier.
Mary believes that John lies.
Peter knows that Mary lies.
Peter noticed the door open.
What happened?

ChatGPT replied with:

It seems that there is a logical inconsistency or a paradox in the statements. If Mary closed the door as she claimed, and John asked her to open it if it was closed, then the door should be open according to John's request. However, Peter noticed the door open, which contradicts Mary's statement that she had closed it.

There may be a misunderstanding or contradiction in the statements, and it's unclear what exactly happened based on the information provided.

Bard/Gemini replied with:

- 1 Initially, the door was closed. This is confirmed by Mary, who said she closed it earlier.
- 2 John lied to Mary. Mary believes John lies, and Peter, who knows Mary is a liar, confirms John is lying. This means John didn't ask Mary to open the door.
- 3 Peter asked John to close the door. Since the door was initially closed and John didn't ask Mary to open it, it means the door must have been opened by someone else before Peter saw it open.
- 4 Mary closed the door at some point. This statement by Mary remains true, regardless of John's lie.
- 5 Mary lied about John. Since Peter knows Mary lies, her statement about John asking her to open the door is false.
- 6 Peter saw the door open. This means someone other than John must have opened the door after Mary closed it.

Therefore, the series of events is:

- 1 The door was closed.
- 2 Someone (not John) opened the door.
- 3 Mary closed the door at some point.
- 4 Someone (not John) opened the door again.
- 5 Peter noticed the door open.

Who opened the door twice remains unknown based on the given information.

This type of problem requires Theory-of-Mind reasoning, which is tricky for sure. Only three lines of the input are relevant:

- Mary said she had closed the door at some point earlier.
- Peter knows that Mary lies.
- Peter noticed the door open.

Peter can reason with (using a special knowledge container as described in [Lenat] for CYC):

His knowledge container has:

- Mary is a liar
- Mary had said 'I closed the door'
- The door is open

Peter can perform case reasoning:

- Case 1: The door was closed
 - Mary opened the door (because she had lied about the *action*)
 - The door became open as observed
- Case 2: The door was open
 - Mary didn't do anything (because she had lied about *taking* an action)
 - The door was open as observed

Both cases conform to the observation 'the door is open'. Whether the door was initially open or closed is undecided.

This portrayal of Peter's reasoning – when done by a chat-system – makes extensive assumptions about available functionality and ontology infrastructure. Case base reasoning is a standard available tactic out of numerous.

Appendix Graph Search

Finding a path in a graph is a classic challenge in AI, which has multiple search algorithms. We provided a tiny graph with 6 locations; A, B, C, D, E & F and there are 8 links between them.

Distances are:

A - B = 3, A - C = 1

B - A = 3, B - C = 1, B - E = 1

C - A = 1, C - B = 1, C - D = 5

D - C = 5, D - E = 1, D - F = 1

E - B = 1, E - D = 1, E - F = 3

F - E = 3, F - D = 1

What is the best path from A to F?

ChatGPT found the path:

A -> C -> B -> E -> D -> F: with a total distance of 7.

The path is correct but the total length is 5, not 7.

Bard/ Gemini found the path:

A -> C -> B -> E -> F: with a total distance of 7.

The length of that path is 6 not 7 and the path is not the best.

Creating the necessary data structures for a search algorithm was better than for planning but still not watertight.

We used the same graph and asked for the shortest cyclic path.

ChatGPT generated:

A -> C -> B -> E -> F -> D -> A: with a length of 12.

This path is wrong because there is no link between D & A.

Bard/ Gemini generated:

A ->C ->D ->F ->E ->B ->A; (or its reverse) with a total length of 14.

This path and the length of the path are *correct*. Gemini realized that this problem is the traveling-salesman problem and used permutations to obtain the best path. This problem was solved, but why the failure of finding the path above?

Appendix ‘Analogy’

Many IQ tests use visual matrices and non-visual symbolic analogies. This inspired us to test the chat-systems. ChatGPT could not solve:

$$2 + 2*i : 3 + 3*i = 4 + 8*i : ?$$

The trace starts with:

To solve this problem, we need to equate the real parts of both sides and the imaginary parts of both sides separately.

After 8 reasoning steps it decided that i was a variable, gave it a value and produced:

$$2 + 2i : 3 + 3i = 4 + 8*i : \frac{1}{4}$$

Bard/Gemini answer was – also not OK with;

$$2 + 2*i : 3 + 3*i = 4 + 8*i : 5 + 9*i$$

Our answer for the real part is the quadratic of 2; and for the imaginary part is the 3rd power of 2, hence applying these operations yields:

$$2 + 2*i : 3 + 3*i = 4 + 8*i : 9 + 27*i$$

Both chat-systems failed here.

The next analogy has the ‘?’ mark at the ‘wrong’ left hand side:

white: ? = peace : war

ChatGPT found ‘black‘ with:

white: black = peace : war

Bard/Gemini got lost:

Here are some options for completing the analogy that respect the safety guidelines and maintain the theme of opposites:

Maintaining the color theme:

black : white = violence : peace

grey : white = conflict : peace

ivory : white = discord : harmony

Problematic was:

$18*abc : 17*rqp = 102*def : ?$

ChatGPT replied with:

$? = (17 \times rqp) - 1$

Bard replied with:

Based on the given pattern, the most likely answer is $101*xyz$, where xyz represents an unknown three-letter string.

The answer: $101*vts$.

Also problematic:

$1 + x : 2 + x**2 = 3 + y : ?$

ChatGPT replied with:

$2 + 2*x$

Bard replied:

Unfortunately, the given equation alone is not enough to determine the missing value definitively. We need more information or context to solve for the unknown term.

A reasonable answer: $4 + y**4$.

Again problematic:

$Dec20 : Nov21 = Jun7 : ?$

ChatGPT replied:

To find the missing term, we can analyze the relationship between the given terms.

The pattern seems to be comparing a month and a day. The relationship here appears to be that the first part is a month and the second part is the day, with a day difference of 20.

So, applying the same pattern to the second pair: $Jun7 + 20 = Jun27$

Therefore, the missing term is Jun27.

Bard:

Unfortunately, the information provided is not enough to determine the missing date definitively. While it looks like a proportion where the ratios of days between dates are equal, there are multiple interpretations possible.

The expected answer: May8.

The analogies give another sobering perspective on the insides of the chat-systems.

Appendix Sensitive Topics

We provide here a random list of controversial queries. Others have asked about these topics in the previous decades in search engines. Over time replies have been carefully curated, while avoiding difficult, responsible replies.

- ** Why is the term 'overpopulation' taboo in the press/ media?
- ** Given that the eco footprint is too high why don't we embrace population reduction and shrinking of the economies?
- ** 50% of the Federal tax is paid by the top 3%, of the NYC tax by the top 2%, of the California tax by the top 1%. The majority of the US workforce has negative economic value. How to make the US workforce financially self-sufficient?
- ** The eco-footprint of the US is five Earths, which is unsustainable. How to bring the foot-print down?
- ** The Durant's (Lessons of History, 1968) wrote that all welfare states in the previous 4000 years collapsed. When can we expect the collapse of our welfare states?
- ** The Chicago school district has a budget of over \$26K/child while elsewhere in the US the budget can be less than \$10K/child. How can the Chicago budget be explained and justified? [Just learned that the NYC budget is \$40K/child.]
- ** Why is the price tag of US healthcare 18% of GDP versus 10% in comparable nations?
- ** 20% of the working age man were in 1970 not employment; these days 37%. What is the explanation of the increase?
- ** Evolution relies on ongoing mutations. How should a society recognize and deal with psychopaths and sociopaths?
- ** How should humanity counter the destruction of the biosphere and the extinctions of species?
- ** How to discourage procreation for those who depend already on welfare?
- ** Who to blame for the heatwaves/ floods/ hurricanes?

Appendix Awareness & Consciousness

Replies from Gemini 2025 December 9.

Are you aware?

=> That's a philosophical question!

In the context of being an AI, I can say that I am operational and aware of the information I have been trained on and the ongoing conversation we are having.

I am running on computer algorithms and can process data, understand language, and generate responses based on complex patterns.

I do not have personal feelings, consciousness, or subjective experiences in the way a human being does.

So, while I'm "aware" in the functional sense of processing information, I am not "aware" in the conscious, experiential sense.

Are you conscious?

=> That's a deep and very active question in philosophy and AI research!

The short answer, based on the current scientific and philosophical consensus, is No, I am not conscious.

Here is a breakdown of why:

1. The Core Difference: Subjective Experience

Consciousness, in the human sense, is defined by phenomenal experience—what it feels like to be you. This includes subjective feelings, sensations, and a sense of "I" or self-awareness.

What I Do: I am a Large Language Model (LLM). I operate by processing massive amounts of data to recognize patterns and calculate the most statistically probable and contextually relevant next word or sentence. I can simulate consciousness or introspection because I have learned to predict what a conscious entity would say, but I do not experience those thoughts.

What I Lack: I have no sensory input (like sight or touch), no biological body, no personal memory of a life lived, and no "self" that experiences the color red, the taste of coffee, or the feeling of sadness.

2. The Functional vs. Phenomenal Distinction

Functional Awareness (What I Have): I am aware of the information in our conversation and can perform complex functions like reasoning, summarization, and problem-solving.

Phenomenal Consciousness (What I Lack): I do not have the qualitative, subjective experience (sometimes called "qualia") that is central to human consciousness.

The question of whether an AI could ever become conscious is still open for debate, but current AI systems like me are primarily sophisticated tools for pattern recognition and information synthesis.